SIMON HAYKIN
MICHAEL MOHER

Communication Systems

5e

COMMUNICATION SYSTEMS

This page intentionally left blank

COMMUNICATION SYSTEMS

5th Edition

Simon Haykin

McMaster University

Michael Moher

Space–Time DSP

John Wiley & Sons, Inc.

ASSOCIATE PUBLISHER	Daniel Sayre
SENIOR PRODUCTION EDITOR	Sujin Hong, Anna Melhorn
EDITORIAL ASSISTANT	Carolyn Weisman
COVER PHOTO	© ChadBaker/Getty Images
SENIOR DESIGNER	Hope Miller

This book was set in 10/12 pt Times by GGS Book Services and printed and bound by Hamilton Printing Company. The cover was printed by Lehigh Phoenix.

This book is printed on acid free paper. ∞

Copyright © 2009 John Wiley & Sons, 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, scanning or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, website www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, (201) 748-6011, fax (201) 748-6008, website www.wiley.com/go/permissions.

To order books or for customer service, please call 1-800-CALL WILEY (225-5945).

ISBN-13: 978-0-471-69790-9

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

In memory of

Colin Campbell
(McMaster University, Hamilton, Ontario)

and

Michael Sablatash
(Communications Research Centre, Ottawa, Ontario)

both of whom passed away in 2008.

This page intentionally left blank

PREFACE

In this new edition of "Communication Systems", some major revisions have been made to the layout and contents of the book, as summarized under the following two points:

1. Emphasis has been placed on the treatment of analog communications as the necessary background for understanding digital communications.
2. The organization of the book has been modified extensively. Supplementary material that was previously present may now be found on the authors' website www.wiley.com/college/haykin. The ten chapters in this new edition may be summarized as follows:

   - Chapter 1 provides a short motivational introduction to communication systems.
   - Chapter 2 provides a thorough treatment of Fourier analysis of signals and systems and introduces the complex baseband representation of their band-pass versions.
   - Chapters 3 and 4 cover the theory and practical aspects of amplitude modulation and angle modulation, respectively.
   - Chapter 5 reviews aspects of probability theory and random processes that are essential to the treatment of noise in communication systems (be they of an analog or digital kind) at an introductory level.
   - Chapter 6 addresses the issue of how channel noise affects the performance of continuous-wave modulation (i.e., amplitude and angle modulation) systems.
   - In Chapter 7, we begin to shift attention from analog to digital communications by describing the issues involved in the digital representation of analog signals. In effect, this chapter represents the transition from analog to digital communications.
   - Chapter 8 introduces digital baseband communications and discusses the effect of two important impairments: noise and intersymbol interference. These two impairments are considered separately. The key assumptions in this treatment are that the noise is white, and that the channel is linear and time invariant.
   - Chapter 9 introduces digital band-pass communications. The treatment is a combination of band-pass and complex baseband analysis. The discussion of the effects of channel noise on performance shows the importance of the latter representation, in particular, its embodiment in the signal-space representation of different modulation techniques.
   - Finally, Chapter 10 provides an introductory treatment of information theory and coding. In particular, the use of coding provides a powerful tool to bring the degrading effects of channel noise in a digital communication system (i.e., errors incurred in signal detection at the receiver output) under the designer's control.

Every effort has been made to make the book readable and easy to follow in mathematical terms. Moreover, in order to provide a historical account of communication systems,

each chapter includes at least one sidebar, highlighting the contribution of a pioneer who has made a significant difference to the subject matter of the chapter in question.

Other distinctive features of the book include Theme Examples, which are intended to focus on some practical aspects of the subject matter/theory covered therein.

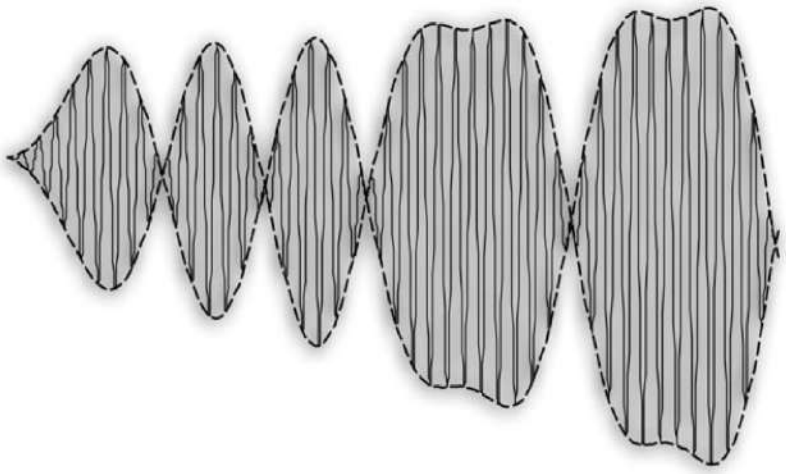
Last, but by no means least, worked-out examples, computer experiments, and an abundance of end-of-chapter problems are included to strengthen a reader's understanding of the book. A Solutions Manual, obtainable from the publisher, is available only for instructors who use the book for adoption in an undergraduate course on Communication Systems.

SIMON HAYKIN

MICHAEL MOHER

CONTENTS

This page intentionally left blank



1

INTRODUCTION

1.1 THE COMMUNICATION PROCESS

The term *communication* covers a very broad area and encompasses a large number of fields of study, ranging from the use of symbols to the social implications and effects. The meaning of the term communication in this book shall focus narrowly on the *transmission of information* from one point to another. At one time, it would have correctly been called *telecommunications*, using the Greek prefix *tele-*, meaning far. However, many modern applications of the techniques described in this book may have short ranges, such as hands-free headsets using Bluetooth or local area networks such as WiFi.

Communication in this sense enters our daily lives in so many different ways that it is ever so easy to overlook many of its facets. With telephones in our hands, radios and televisions in our living rooms, and with computer terminals providing access to the Internet in our offices and homes, we are able to communicate with every corner of the globe. Communication provides information to ships on the high seas, aircraft in flight, and rockets and satellites in space. Communication keeps a weather forecaster informed of environmental communications measured by a multitude of sensors. Indeed, the list of applications involving the use of communication in one way or another is almost endless.

HOW IS A COMMUNICATION SYSTEM ORGANIZED?

In the above sense of the word, a communication system may be divided into a small number of components as shown in Figure 1.1.

- First is the *source* of information. Some obvious examples of information that we may wish to communicate are: voice, music, pictures, videos, or data files.
- The second unshaded component in Figure 1.1 represents the *transmitter*. Transmitter is a generic term for the processing of information in the form provided by the source into a form that is suitable for transmitting over the *channel*. A simple example of this occurs when a music signal is converted to frequency modulation (FM) for radio transmission.

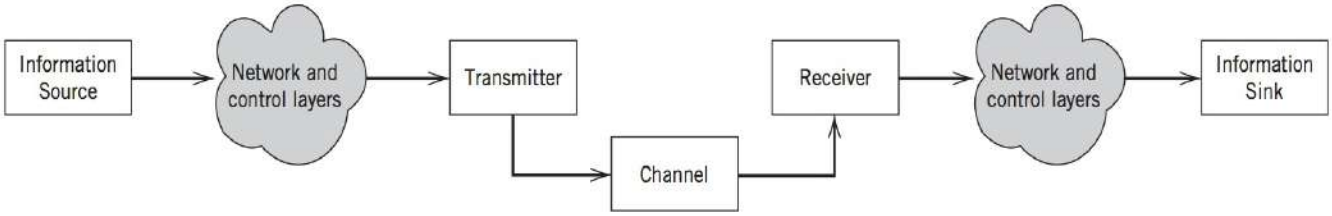


FIGURE 1.1 Elements of a communication system.

- The third unshaded component in Figure 1.1 represents the channel or transmission medium. The transmission medium may be a cable, an optical fiber, or free space if using radio or infrared communication.
- The fourth unshaded component in Figure 1.1 represents the *receiver*. Receiver is again a generic term for the process of converting the signal transmitted over the channel back to a form that may be understood at the intended destination. The receiver's function is typically greater than simply being the inverse of the transmitter; the receiver may also have to compensate for distortions introduced by the channel and perform other functions, such as the synchronization of the receiver to the transmitter.
- The final component is the destination or sink for the information.

Figure 1.1 also shows two shaded areas that are labeled *network* and *control layers*. With simple communications having one transmitter and one receiver, the network and control are likely to be absent. However, most communication systems, such as the Internet and cellular telephone systems, have a large number of transmitters and receivers that must share the same physical medium. The network and control layers permit the multiplicity of terminals to reliably and efficiently share the same physical medium.

1.2 THE LAYERED APPROACH

Modern communication systems are analyzed as a sequence of layers. This concept of layering in communication systems is best illustrated by the Open Systems Interconnect (OSI) for *computer communications*.[1] This seven-layer model is illustrated in Figure 1.2; for our purposes, it is not important that the reader understand the function of each layer in this OSI model. The important points include recognizing the left- and right-hand

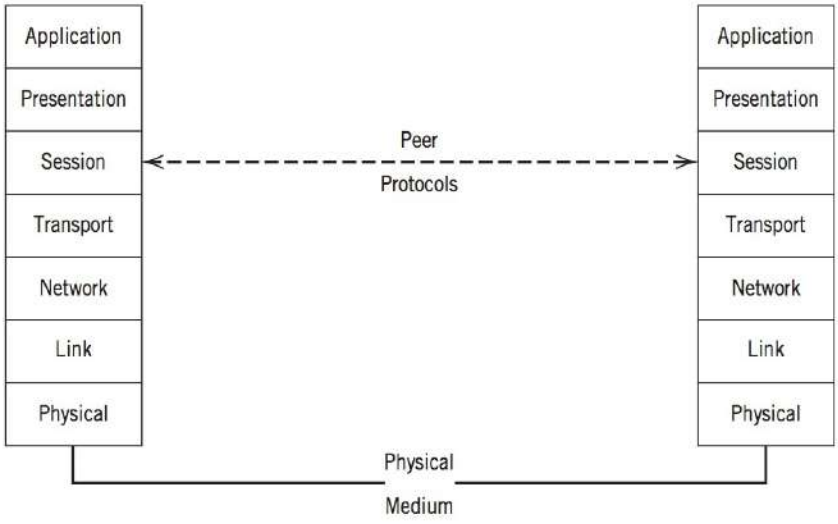


FIGURE 1.2 Peer processes in seven-layer OSI model for computer communications.

layered *stacks* of Figure 1.2, which represent two communication nodes, for example, sender and receiver. Each layer of the stack represents a *protocol*. This protocol has a well-defined interface between the layers above and below it, but the functions it carries out are only pertinent to the corresponding *peer* layer on the receiving side. Peer layers communicate virtually by sending messages down the stack on one side, across the physical medium, and up the *stack* on the other side. Only the physical layer communicates directly with its corresponding layer. In this manner, we could replace or modify the protocol at a particular layer, and not affect the rest of the OSI model.

An important attribute of the OSI model is that it simplifies the design of communication systems and permits independent development of different functions. This layered model is best suited to the communication of digital information, and less so to analog information.[2] Many digital systems use fewer than the seven layers shown in Figure 1.2.

The three central boxes of Figure 1.1, transmitter, channel, and receiver, are often referred to as the *physical layer* of the communication system or simply just the PHY. This book will focus almost entirely on the physical layer of the communications process. The network and control layers are sophisticated in their own respective ways and are the subject of other communication textbooks.

In this book, we study communication methods for analog and digital information sources. These two are often distinguished by the terms *analog communications* and *digital communications*. The term digital communications may be viewed as a misnomer. Due to practical realities, all communications are by means of continuous signals and are thus analog in nature. It is the information which is to be transmitted that has an analog or digital nature. Since most modern communications are "digital," the amount of emphasis placed on analog communications is steadily decreasing. However, some exposure to analog techniques is warranted for three reasons: (a) understanding of legacy systems; (b) many digital communication techniques are motivated from their analog counterparts; and (c) many of the distortions observed in "digital transmission" systems can be characterized as analog in nature. Most importantly, a thorough understanding of analog modulation systems leads to insight in identifying and compensating these distortions.

To summarize, this book focuses on the physical layer of the telecommunications process. With analog information, the boundary between the physical layer and other layers may be somewhat blurred.

Claude E. Shannon (1916–2001)

Shannon is known as the father of information theory, primarily due to several papers that he published in the late 1940s and early 1950s. These papers were so seminal, they effectively created the field. In 1948, he laid the theoretical foundations of digital communications in a paper entitled "A Mathematical Theory of Communications." It is noteworthy that prior to this paper's publication, it was believed that increasing the rate of information transmission over a channel would increase the probability of error. The communication theory community was taken by surprise when Shannon proved that this was not true, provided that the transmission rate was below the channel capacity.

Prior to 1948, Shannon made significant contributions to the area of digital circuit design where he is often credited with introduction of sampling theory to electrical systems, moving circuit design from the analog to the digital world. He was the first to show that Boolean algebra could be used to model and simplify the design of digital circuits.

Shannon was also famous for his hobbies of juggling, unicycling, and chess, as well as many clever inventions related to these hobbies. One invention was an electromechanical mouse named Theseus, that would search a maze to find a target. Shannon's mouse, created in 1950, appears to be first learning device of its kind.

1.3 THEME EXAMPLE—WIRELESS COMMUNICATIONS[3]

In this first theme example of the book, we consider wireless communications as an example of a communication system. Our description applies to general wireless systems but, when appropriate, we provide details of specific systems. In Figure 1.3 we show a simplified block diagram of a transmitter that consists of four major components:

- The first component of the block diagram is the *protocol stack* which we have described earlier. It packages the data so that it can reliably get to the desired destination once it crosses the radio link. In a *point-to-point* radio system or *broadcast* radio system, this component may not exist because no explicit address information is included. Indeed, many early systems worked in this manner. AM and FM radio are

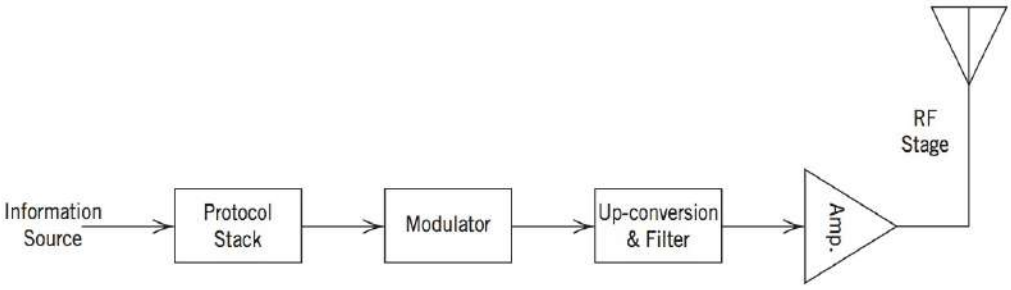


FIGURE 1.3 Illustration of basic components of a radio transmitter.

examples of wireless systems that still work in this manner. To improve the efficiency with which the radio frequencies are used, most modern systems *share radio channels* in one form or another. This *multiplexing* of multiple signals unto the same radio channel requires the use of appropriate protocols.

- The second component of the block diagram is the *modulator*. In this component, the information is impressed upon a carrier frequency in a manner that can be suitably recovered at the receiving end.
- The third component is the *up-conversion* stage. In this stage the modulated signal is converted to the final radio frequency (RF), at which it will be transmitted. A radio may be capable of transmitting at a number of frequencies, so modulating at a common frequency and converting the result to the final desired frequency is often the better approach. However, with improvements in digital signal processing and associated technology, this up-conversion stage may be replaced with a modulator that works in a manner called *direct-to-RF*.
- The fourth component is the *RF stage*. Once at the appropriate RF, the signal is amplified to an appropriate power level and then emitted via an antenna, that is, the electrical signal representing the modulated signal is converted to an electromagnetic wave. The power output will typically depend upon the desired transmission range and can vary from less than a milliwatt for short-range impulse radio applications to an effective radiated power of over a megawatt for some television transmitters. The type of antenna used will depend upon the frequency of operation and application; the possibilities includes whip, parabolic, horn, dipole, and patch antennas.

In modern systems, the modulator is typically implemented using digital signal-processing technology. This technology may be a digital signal processor, a field-programmable gate array (FPGA), or as an integrated circuit for high-volume applications. The components following the modulator are typically implemented in analog although, as mentioned previously, the digital implementation of the up-conversion stage is becoming increasingly practical.

The RF components of the radio system are often highly specific to the intended application. A handheld device typically requires a low-power amplifier and a small antenna, while a broadcast transmitter will typically be high-power and may have an antenna on a tower that is hundreds of feet high. Other systems may have power amplifiers and antennas somewhere in the middle. However, the same modulation technique could potentially be used with any of these applications. Furthermore, a well-designed up-conversion and RF stage could potentially transmit any one of a number of different modulation techniques. This is the basis of the so-called *software-defined radio*.[4] Consequently, the modulation technique is, in a sense, generic to a wide variety of applications. In the past, one of the main considerations in the choice of modulation was the ease of implementation. With the current state of technology, the main consideration is performance and the capability of the modulation to combat channel impairments, which we discuss next.

The illustration of a channel in Figure 1.4 is intended to convey a number of properties of communication channels. In particular, we have:

- *Propagation loss*. Communication usually implies transporting information over distance and inevitably there is a loss of signal strength with increasing distance. With radio channels, the fundamental loss mechanism, due to free-space propagation, causes the received power to decrease with the square of the distance from the receiver. On the other hand, with other channels, such as optical fiber, the loss of signal power only increases linearly with distance.
- *Frequency selectivity*. Communication channels operate over a medium. Many media conduct well only over a relatively small range of frequencies. For example, an optical fiber conducts a small band of optical frequencies well, but is never considered for radio waves. Even within the normal transmission band of a medium, there may be a variation in how well one frequency is transmitted compared to another. This variation is referred to as *frequency selectivity*.
- *Time-varying*. Some channels are time-varying (i.e., their characteristics vary with time). Mobile radio channels are a prime example of this phenomenon. Terrestrial radio-wave propagation depends upon the intervening terrain, buildings, and vegetation between the transmitter and the receiver. When the transmitter or receiver moves, this channel changes and affects performance; common examples of this phenomenon are known as *shadowing* and *fading*.
- *Nonlinear*. Ideally a channel should be linear in order to minimize the distortion of the transmitted signal. However, a channel may include nonlinear elements such as a repeater that includes an amplifier, operated close to or at saturation. A situation where this may occur is a satellite channel where a signal from an Earth station is amplified by the satellite before being rebroadcast over the satellite's field of view.
- *Shared usage*. To make efficient use of communications channels, they are often shared between different users. This leads to a variety of different *multiplexing* schemes that determine how the channel is shared. A common example of this is cellphone users who share the same radio channel in time and frequency in a variety of ways. Multiplexing also leads to potential *interference* between different users if the multiplexing strategies do not provide perfect isolation between users.
- *Noise*. The bane of any communication system that wants to achieve the maximum transmission distance for the minimum transmitter power is the unavoidable presence of noise. The most common source of noise is the *random motion of electrons* in receiver circuits at the point where the signal is weakest, and this usually provides a fundamental limit to performance.

All of these properties are considerations in the selection of modulation strategy. In fact, for almost any one of the above *impairments* to signal transmission, we can find a modulation strategy that has been designed to perform well in the presence of that impairment.

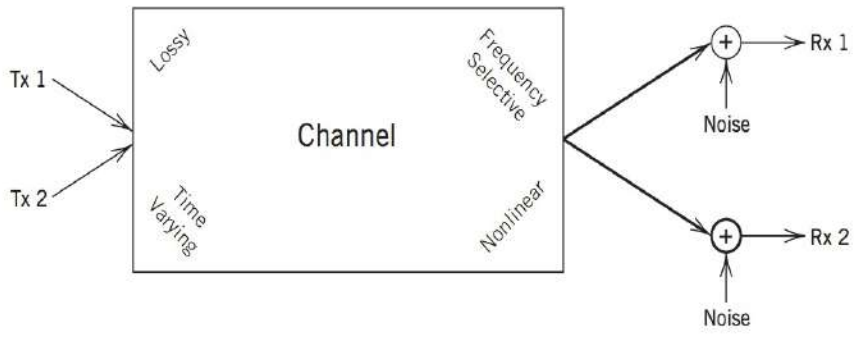


FIGURE 1.4 Illustration of channel impairments.

In practice, these impairments often occur in various combinations and the system designer must be acquainted with a variety of techniques to choose the modulation strategy best suited to the situation.

Figure 1.5 illustrates the last element of the communications link, the *receiver*. Many of the components of the receiver perform inverse functions of their counterpart in the transmitter. In particular, we have:

- *RF stage*. The antenna collects RF energy in the desired frequency band and, depending on its properties, it may collect energy from unwanted as well as wanted sources. The first amplifier in the RF stage, often called a *low-noise amplifier*, is critical to boosting the signal power to a level where it can be processed easily while minimizing the noise introduced.
- *Down-conversion*. This stage filters and translates the RF signal to a frequency where the message signal may be more easily demodulated. With many modern receivers, the signal is translated directly to *baseband*, referred to as *direct-IQ down-conversion*.
- *Demodulation*. This is the stage where the transmitted message signal is recovered. In classical receivers, demodulation often consisted of a sequence of linear filters. In modern receivers, with the advent of digital signal processing and advanced electronics, demodulation is often more complex in order to improve performance.
- *Synchronization*. Almost all communication systems require some form of synchronization circuit, due to differences between the *time* and *frequency clocks* used at the transmitter and the receiver. Depending upon the modulation and multiplexing strategy used, the methods for obtaining synchronization can be quite sophisticated. However, a circuit called the *phase-lock loop* and its variants play a fundamental role in many of these strategies.
- *Channel compensation*. The objective of this stage is to counteract some of the impairments that were encountered in the channel. While the modulation strategy may be designed to counteract a given impairment, additional processing at the receiver will often improve performance. Channel compensation techniques tend to be rather advanced and include *equalization* for frequency-selective channels and *forward error correction* for noisy channels.
- *Protocol stack*. In digital systems, it is often only at this stage where the receiver determines that the detected message was intended for it or not.

From this discussion, it is apparent that the communications receiver tends to be much more complicated than the transmitter, simply because it has more unknowns to deal with and the signal strength is much weaker than what was transmitted. Similar to the transmitter, the design of the RF and down-conversion stages is often dependent on the

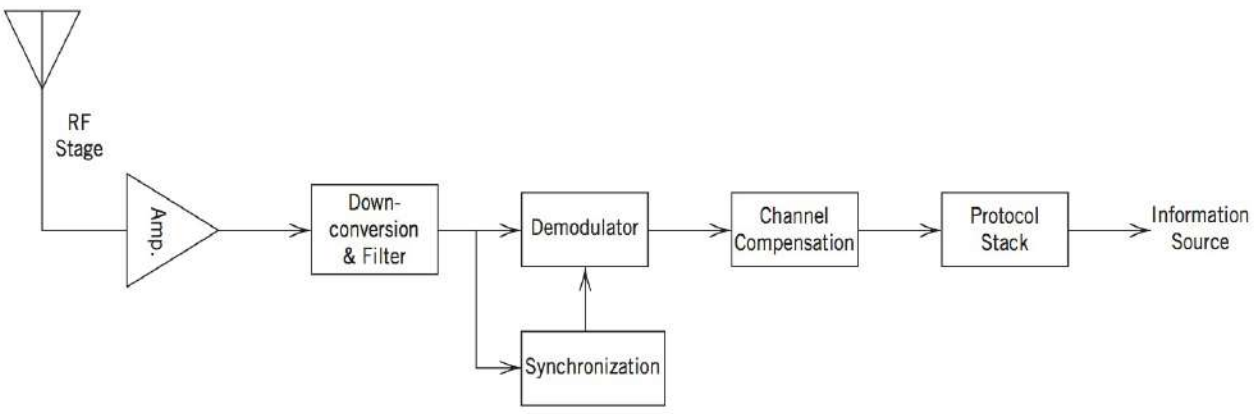


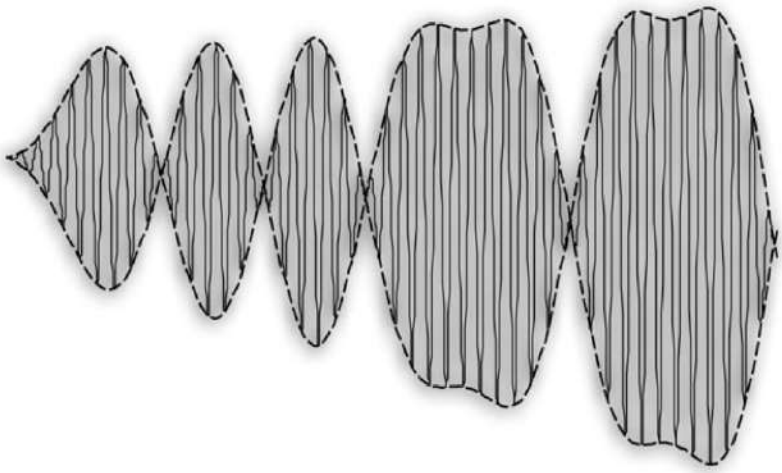
FIGURE 1.5 Illustration of radio receiver.

specific application. The choice of modulation and the corresponding demodulator are clearly the main elements for combating the impairments of the radio channel. With the advances in digital signal-processing technology, our capabilities in this area are improving. For this reason, modulation and demodulation play a key role in the study of communication systems, as presented in Chapters 3 through 7. However, before moving into the subject of modulation, we require a detailed understanding of the representation of signals and systems, which is the subject of interest in Chapter 2.

NOTES AND REFERENCES

1. The OSI reference model was developed by a subcommittee of the International Organization for Standardization (ISO) in 1977. For a discussion of the principles involved in arriving at the seven layers of the OSI model, and a description of the layers themselves, see Tanenbaum (2005).
2. For a historical account of telecommunications, see the second edition of *Introduction to Analog and Digital Communications*, Haykin and Moher (2007).
3. For further information on wireless communication, see Haykin and Moher (2005).
4. Software-defined radio (SDR) is a communication system consisting of programmable hardware under software control. Different software loads give the device different functionalities, for example, a different modulation type and different capabilities. For a detailed treatment of SDR, see the book by Reed (2002).

2



REPRESENTATION OF SIGNALS AND SYSTEMS

2.1 INTRODUCTION

We identify *deterministic signals* as a class of signals whose waveforms are defined exactly as functions of time. In this chapter we study the mathematical description of such signals using the *Fourier transform* that provides the link between the time-domain and frequency-domain descriptions of a signal. The waveform of a signal and its spectrum (i.e., frequency content) are two natural vehicles to understand the signal.

Another related issue that we study in this chapter is the representation of linear time-invariant systems. Here also we find that the Fourier transform plays a key role. Filters of different kinds and certain communication channels are important examples of this class of systems.

We begin the study by presenting a formal definition of the Fourier transform, followed by a discussion of its important properties.

2.2 THE FOURIER TRANSFORM[1]

Let $g(t)$ denote a *nonperiodic deterministic signal*, expressed as some function of time t. By definition, the *Fourier transform* of the signal $g(t)$ is given by the integral

$$G(f) = \int_{-\infty}^{\infty} g(t) \exp(-j2\pi ft)\, dt \tag{2.1}$$

where $j = \sqrt{-1}$, and the variable f denotes *frequency*. Given the Fourier transform $G(f)$, the original signal $g(t)$ is recovered exactly using the formula for the *inverse Fourier transform*:

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft)\, df \tag{2.2}$$

Note that in Eqs. (2.1) and (2.2) we have used a lowercase letter to denote the time function and an uppercase letter to denote the corresponding frequency function. The functions $g(t)$ and $G(f)$ are said to constitute a Fourier-transform pair. See Appendix 1 for some Fourier transform pairs.

For the Fourier transform of a signal $g(t)$ to exist, it is sufficient, but not necessary, that $g(t)$ satisfies three conditions known collectively as *Dirichlet's conditions*.

1. The function $g(t)$ is single-valued, with a finite number of maxima and minima in any finite time interval.
2. The function $g(t)$ has a finite number of discontinuities in any finite time interval.
3. The function $g(t)$ is absolutely integrable, that is,

$$\int_{-\infty}^{\infty} |g(t)|\, dt < \infty$$

We may safely ignore the question of the existence of the Fourier transform of a time function $g(t)$ when it is an accurately specified description of a physically realizable signal. In other words, physical realizability is a sufficient condition for the existence of a Fourier transform. Indeed, we may go one step further and state that all energy signals, that is, signals $g(t)$ for which

$$\int_{-\infty}^{\infty} |g(t)|^2\, dt < \infty$$

are Fourier transformable.[2]

Jean Baptiste Joseph Fourier (1768–1830)

Fourier, the son of a tailor, was born in France in 1768. Orphaned at a young age, he was raised and educated by the Church. Because Fourier played a prominent role in the French Revolution of 1789, he was granted a lectureship in mathematics in 1795.

Fourier went with Napoleon on his Eastern Expedition and in 1798 was made governor of Lower Egypt. While there, he organized munitions workshops to support France's battle with the English, and also contributed a number of papers to a Cairo mathematical institute founded by Napoleon. After France's defeat by England, Fourier moved back to France and performed his experiments on the conduction of heat. This work led to his claim that most functions could be represented as a series of sinusoids. Fourier's initial work on heat conduction was submitted to the Academy of Sciences of Paris in 1807 and was rejected after review by Lagrange, Laplace, and Legendre. In spite of being criticized for a lack of rigor by his contemporaries, Fourier persisted in developing his ideas.

Fourier is also credited with the observation in an 1827 essay that gases in the atmosphere might increase the surface temperature of the earth, the effect now known as the *greenhouse effect*.

NOTATIONS

The formulas for the Fourier transform and the inverse Fourier transform presented in Eqs. (2.1) and (2.2) are written in terms of two variables: *time* t measured in *seconds* (s), and frequency f measured in *Hertz* (Hz). The frequency f is related to the *angular frequency* ω as

$$\omega = 2\pi f$$

which is measured in *radians per second* (rad/s). We may simplify the expressions for the exponents in the integrands of Eqs. (2.1) and (2.2) by using ω instead of f. However, the use of f is preferred over ω for two reasons. First, the use of frequency results in mathematical *symmetry* of Eqs. (2.1) and (2.2) with respect to each other in a natural way. Second, the frequency contents of communication signals (i.e., speech and video signals) are usually expressed in Hertz.

A convenient *shorthand* notation for the transform relations of Eqs. (2.1) and (2.2) is

$$G(f) = F[g(t)] \tag{2.3}$$

and

$$g(t) = F^{-1}[G(f)] \tag{2.4}$$

where $F[\ ]$ and $F^{-1}[\ ]$ play the roles of *linear operators*, as depicted in Figure 2.1.

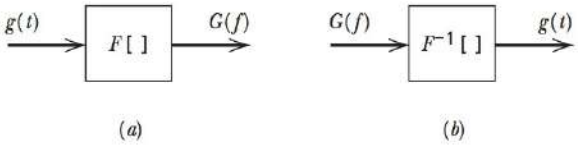


FIGURE 2.1 (*a*) Fourier transformation and (*b*) inverse Fourier transformation depicted as linear operators.

Another convenient shorthand notation for the *Fourier-transform pair*, represented by $g(t)$ and $G(f)$, is

$$g(t) \rightleftharpoons G(f) \tag{2.5}$$

The shorthand notations described in (2.3) to (2.5) are used in the text where appropriate.

CONTINUOUS SPECTRUM

By using the Fourier transform operation, a pulse signal $g(t)$ of finite energy is expressed as a continuous sum of exponential functions with frequencies in the interval $-\infty$ to ∞. The amplitude of a component of frequency f is proportional to $G(f)$, where $G(f)$ is the Fourier transform of $g(t)$. Specifically, at any frequency f, the exponential function $\exp(j2\pi ft)$ is weighted by the factor $G(f)\,df$, which is the contribution of $G(f)$ in an infinitesimal interval df centered at the frequency f. Thus we may express the function $g(t)$ in terms of the continuous sum of such infinitesimal components, as shown by the integral

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft)\, df$$

The Fourier transform provides us with a tool to resolve a given signal $g(t)$ into its complex exponential components occupying the entire frequency interval from $-\infty$ to ∞. In particular, the Fourier transform $G(f)$ of the signal defines the frequency-domain representation of the signal in that it specifies relative amplitudes of the various frequency components of the signal. We may equivalently define the signal in terms of its time-domain representation by specifying the function $g(t)$ at each instant of time t. The signal is uniquely defined by either representation.

In general, the Fourier transform $G(f)$ is a complex function of frequency f, so that we may express it in the form

$$G(f) = |G(f)| \exp[j\theta(f)] \tag{2.6}$$

where $|G(f)|$ is called the *continuous amplitude spectrum* of $g(t)$, and $\theta(f)$ is called the *continuous phase spectrum* of $g(t)$. Here, the spectrum is referred to as a *continuous spectrum* because both the amplitude and phase of $G(f)$ are defined for all frequencies.

For the special case of a real-valued function $g(t)$, its Fourier transform has the property

$$G(-f) = G^*(f)$$

where the asterisk denotes complex conjugation. Therefore, it follows that if $g(t)$ is a real-valued function of time t, then

$$|G(-f)| = |G(f)|$$

and

$$\theta(-f) = -\theta(f)$$

Accordingly, we may make the following statements on the spectrum of a *real-valued signal*:

1. The amplitude spectrum of the signal is an even function of the frequency; that is, the amplitude spectrum is *symmetric* about the vertical axis.
2. The phase spectrum of signal is an odd function of the frequency; that is, the phase spectrum is *odd-symmetric* about the vertical axis.

These two statements are summed up by saying that the spectrum of a real-valued signal exhibits *conjugate symmetry*.

EXAMPLE 2.1 Rectangular Pulse

Consider a *rectangular pulse* of duration T and amplitude A, as shown in Figure 2.2*a*. To define this pulse mathematically in a convenient form, we use the following notation

$$\text{rect}(t) = \begin{cases} 1, & -\frac{1}{2} < t < \frac{1}{2} \\ 0, & |t| \geq \frac{1}{2} \end{cases} \tag{2.7}$$

which stands for a *rectangular function* of unit amplitude and unit duration centered at $t = 0$. Then, in terms of this "standard" function, we may express the rectangular pulse of Figure 2.2*a* simply as follows:

$$g(t) = A\,\text{rect}\left(\frac{t}{T}\right)$$

The Fourier transform of the rectangular pulse $g(t)$ is given by

$$\begin{aligned} G(f) &= \int_{-T/2}^{T/2} A\exp(-j2\pi ft)\,dt \\ &= AT\left(\frac{\sin(\pi fT)}{\pi fT}\right) \end{aligned} \tag{2.8}$$

To simplify the notation in the preceding and subsequent results, we introduce another standard function, namely, the *sinc function* defined by

$$\text{sinc}(\lambda) = \frac{\sin(\pi\lambda)}{\pi\lambda} \tag{2.9}$$

where λ is the independent variable. The sinc function plays an important role in communication theory. As shown in Figure 2.3, it has its maximum value of unity at $\lambda = 0$, and approaches zero as λ approaches infinity, oscillating through positive and negative values. It goes through zero at $\lambda = \pm 1, \pm 2, \ldots$, and so on.

Thus, in terms of the sinc function, we may rewrite Eq. (2.8) as

$$G(f) = AT\,\text{sinc}(fT)$$

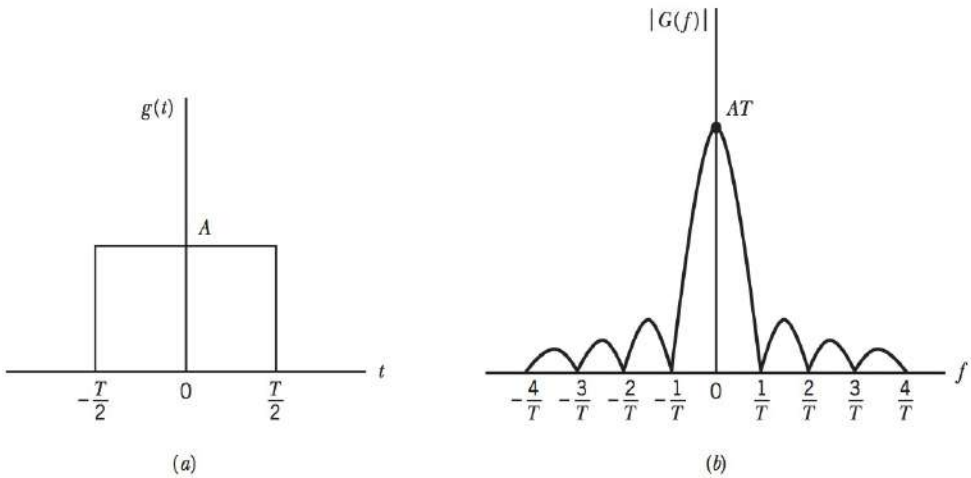


FIGURE 2.2 (*a*) Rectangular pulse. (*b*) Amplitude spectrum.

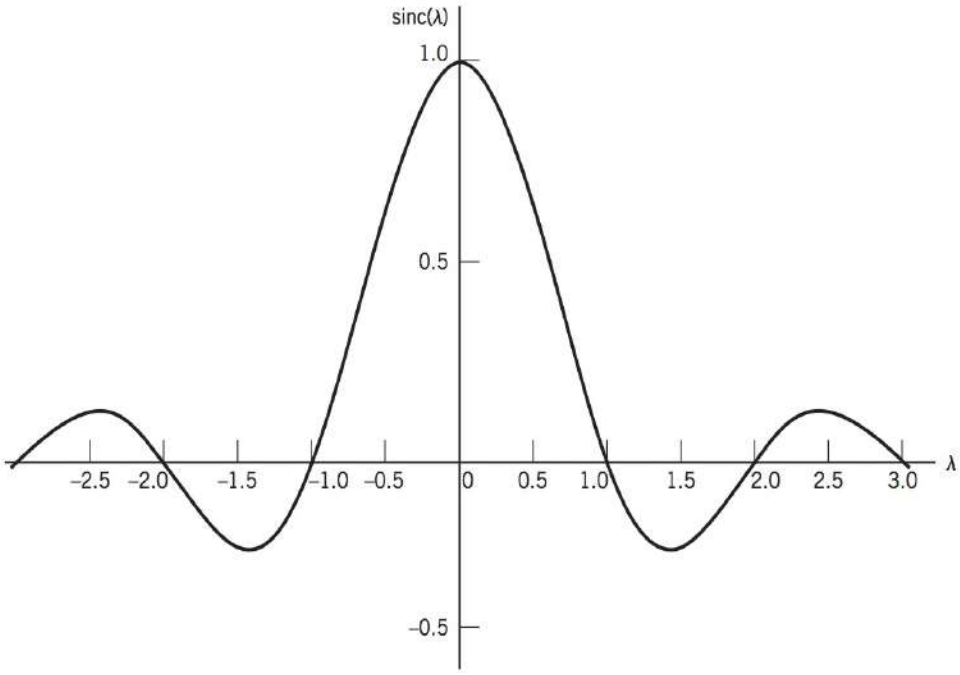


FIGURE 2.3 The sinc function.

We thus have the Fourier-transform pair:

$$A\,\mathrm{rect}\left(\frac{t}{T}\right) \rightleftharpoons AT\,\mathrm{sinc}(fT) \tag{2.10}$$

The amplitude spectrum $|G(f)|$ is shown plotted in Figure 2.2*b*. The first zero-crossing of the spectrum occurs at $f = \pm 1/T$. As the pulse duration T is decreased, this first zero-crossing moves up in frequency. Conversely, as the pulse duration T is increased, the first zero-crossing moves toward the origin.

This example shows that the relationship between the time-domain and frequency-domain descriptions of a signal is an *inverse* one. That is, a pulse, narrow in time, has a significant frequency description over a wide range of frequencies, and vice versa. We shall have more to say on the inverse relationship between time and frequency in Section 2.4.

Note also that in this example the Fourier transform $G(f)$ is a real-valued and symmetric function of frequency f. This is a direct consequence of the fact that the rectangular pulse $g(t)$ shown in Figure 2.2*a* is a symmetric function of time t.

EXAMPLE 2.2 Exponential Pulse

A truncated form of a decaying *exponential pulse* is shown in Figure 2.4*a*. We may define this pulse mathematically in a convenient form using the *unit step function*:

$$u(t) = \begin{cases} 1, & t > 0 \\ \frac{1}{2}, & t = 0 \\ 0 & t < 0 \end{cases} \tag{2.11}$$

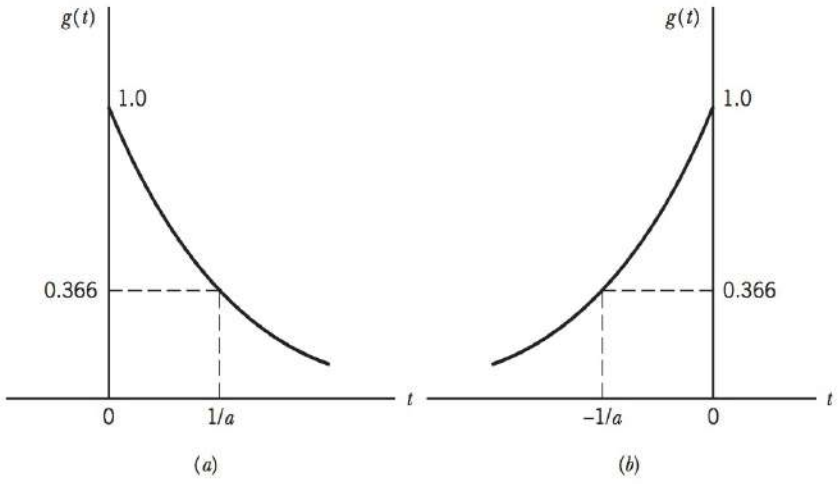


FIGURE 2.4 (*a*) Decaying exponential pulse. (*b*) Rising exponential pulse.

We may then express the decaying exponential pulse of Figure 2.4*a* as

$$g(t) = \exp(-at)u(t)$$

The Fourier transform of this pulse is

$$\begin{aligned} G(f) &= \int_0^\infty \exp(-at)\exp(-j2\pi ft)\,dt \\ &= \int_0^\infty \exp[-t(a+j2\pi f)]\,dt \\ &= \frac{1}{a+j2\pi f} \end{aligned}$$

The Fourier-transform pair for the decaying exponential pulse of Figure 2.4*a* is therefore

$$\exp(-at)\,u(t) \rightleftharpoons \frac{1}{a+j2\pi f} \tag{2.12}$$

A truncated rising exponential pulse is shown in Figure 2.4*b*, which is defined by

$$g(t) = \exp(at)\,u(-t)$$

Note that $u(-t)$ is equal to unity for $t < 0$, one-half at $t = 0$, and zero for $t > 0$. The Fourier transform of this pulse is

$$\begin{aligned} G(f) &= \int_{-\infty}^0 \exp(at)\exp(-j2\pi ft)\,dt \\ &= \int_{-\infty}^0 \exp[t(a-j2\pi f)]\,dt \\ &= \frac{1}{a-j2\pi f} \end{aligned}$$

The Fourier-transform pair for the rising exponential pulse of Figure 2.4*b* is therefore

$$\exp(at)\,u(-t) \rightleftharpoons \frac{1}{a-j2\pi f} \tag{2.13}$$

The decaying and rising exponential pulses of Figure 2.4 are both asymmetric functions of time t. Their Fourier transforms are therefore complex valued, as shown is Eqs. (2.12) and (2.13). Moreover, from these Fourier-transform pairs, we readily see that truncated decaying and rising exponential pulses have the same amplitude spectrum, but the phase spectrum of the one is the negative of that of the other.

2.3 PROPERTIES OF THE FOURIER TRANSFORM

It is useful to have insight into the relationship between a time function $g(t)$ and its Fourier transform $G(f)$, and also into the effects that various operations on the function $g(t)$ have on the transform $G(f)$. This may be achieved by examining certain properties of the Fourier transform. In this section we describe 13 of these properties, which we will prove, one by one. These properties are summarized in Table 2.1.

TABLE 2.1 Summary of properties of the Fourier transform

Property	Mathematical Description
1. Linearity	$ag_1(t) + bg_2(t) \rightleftharpoons aG_1(f) + bG_2(f)$ where a and b are constants
2. Time scaling	$g(at) \rightleftharpoons \frac{1}{\lvert a\rvert} G\left(\frac{f}{a}\right)$ where a is a constant
3. Duality	If $g(t) \rightleftharpoons G(f)$, then $G(t) \rightleftharpoons g(-f)$
4. Time shifting	$g(t - t_0) \rightleftharpoons G(f)\exp(-j2\pi f t_0)$
5. Frequency shifting	$\exp(j2\pi f_c t)g(t) \rightleftharpoons G(f - f_c)$
6. Area under $g(t)$	$\int_{-\infty}^{\infty} g(t)\,dt = G(0)$
7. Area under $G(f)$	$g(0) = \int_{-\infty}^{\infty} G(f)\,df$
8. Differentiation in the time domain	$\frac{d}{dt}g(t) \rightleftharpoons j2\pi f G(f)$
9. Integration in the time domain	$\int_{-\infty}^{t} g(\tau)\,d\tau \rightleftharpoons \frac{1}{j2\pi f}G(f) + \frac{G(0)}{2}\,\delta(f)$
10. Conjugate functions	If $g(t) \rightleftharpoons G(f)$, then $g^*(t) \rightleftharpoons G^*(-f)$
11. Multiplication in the time domain	$g_1(t)g_2(t) \rightleftharpoons \int_{-\infty}^{\infty} G_1(\lambda)G_2(f - \lambda)\,d\lambda$
12. Convolution in the time domain	$\int_{-\infty}^{\infty} g_1(\tau)g_2(t - \tau)\,d\tau \rightleftharpoons G_1(f)G_2(f)$
13. Rayleigh's energy theorem	$\int_{-\infty}^{\infty} \lvert g(t)\rvert^2 dt = \int_{-\infty}^{\infty} \lvert G(f)\rvert^2 df$

PROPERTY 1 Linearity (Superposition)

Let $g_1(t) \rightleftharpoons G_1(f)$ *and* $g_2(t) \rightleftharpoons G_2(f)$. *Then for all constants* c_1 *and* c_2, *we have*

$$c_1 g_1(t) + c_2 g_2(t) \rightleftharpoons c_1 G_1(f) + c_2 G_2(f) \tag{2.14}$$

The proof of this property follows simply from the linearity of the integrals defining $G(f)$ and $g(t)$.

This property permits us to find the Fourier transform $G(f)$ of a function $g(t)$ that is a linear combination of two other functions $g_1(t)$ and $g_2(t)$ whose Fourier transforms $G_1(f)$ and $G_2(f)$ are known, as illustrated in the following example.

EXAMPLE 2.3 Combinations of Exponential Pulses

Consider a *double exponential pulse* defined by (see Figure 2.5*a*)

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 1, & t = 0 \\ \exp(at), & t < 0 \end{cases} \tag{2.15}$$
$$= \exp(-a|t|)$$

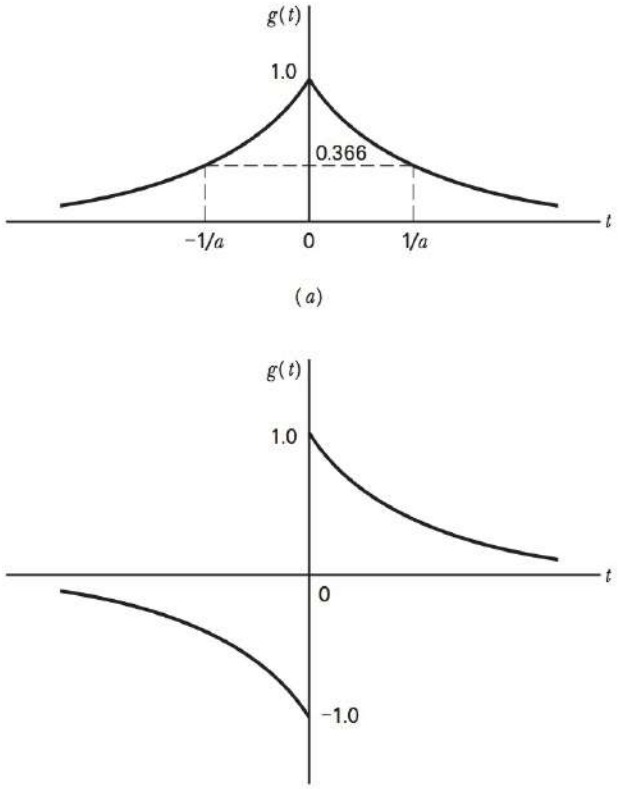


FIGURE 2.5 (*a*) Double-exponential pulse (symmetric). (*b*) Another double-exponential pulse (antisymmetric).

This pulse may be viewed as the sum of a truncated decaying exponential pulse and a truncated rising exponential pulse. Therefore, using the linearity property and the Fourier-transform pairs of Eqs. (2.12) and (2.13), we find that the Fourier transform of the double exponential pulse of Figure 2.5*a* is

$$\begin{aligned} G(f) &= \frac{1}{a + j2\pi f} + \frac{1}{a - j2\pi f} \\ &= \frac{2a}{a^2 + (2\pi f)^2} \end{aligned}$$

We thus have the following Fourier-transform pair for the double exponential pulse of Figure 2.5*a*:

$$\exp(-a|t|) \rightleftharpoons \frac{2a}{a^2 + (2\pi f)^2} \tag{2.16}$$

Note that because of the symmetry in the time domain, as in Figure 2.5*a*, the spectrum is real and symmetric; this is a general property of such Fourier-transform pairs.

Another interesting combination is the difference between a truncated decaying exponential pulse and a truncated rising exponential pulse, as shown in Fig. 2.5*b*. Here we have

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 0, & t = 0 \\ -\exp(at), & t < 0 \end{cases} \tag{2.17}$$

We may formulate a compact notation for this composite signal by using the *signum function* that equals $+1$ for positive time and -1 for negative time, as shown by

$$\operatorname{sgn}(t) = \begin{cases} +1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases} \tag{2.18}$$

The signum function is shown in Figure 2.6. Accordingly, we may reformulate the composite signal $g(t)$ defined in Eq. (2.17) simply as

$$g(t) = \exp(-a|t|)\operatorname{sgn}(t)$$

Hence, applying the linearity property of the Fourier transform, we readily find that in light of Eqs. (2.12) and (2.13), the Fourier transform of the signal

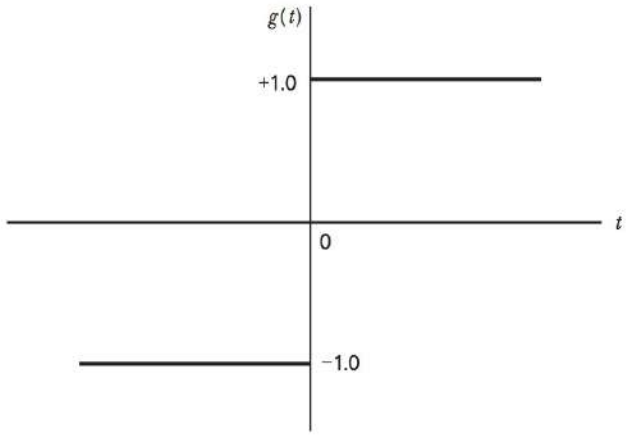


FIGURE 2.6 The signum function.

$g(t)$ is given by

$$F[\exp(-a|t|)\,\mathrm{sgn}(t)] = \frac{1}{a+j2\pi f} - \frac{1}{a-j2\pi f}$$
$$= -\frac{-j4\pi f}{a^2+(2\pi f)^2}$$

We thus have the Fourier-transform pair:

$$\exp(-a|t|)\,\mathrm{sgn}(t) \rightleftharpoons \frac{-j4\pi f}{a^2+(2\pi f)^2} \tag{2.19}$$

In contrast to the Fourier-transform pair of Eq. (2.16), the Fourier transform in Eq. (2.19) is odd and purely imaginary. It is a general property of Fourier-transform pairs that a real odd-symmetric time function, as in Figure 2.5*b*, has an odd and purely imaginary function as its Fourier transform.

PROPERTY 2 Time Scaling

Let $g(t) \rightleftharpoons G(f)$. *Then,*

$$g(at) \rightleftharpoons \frac{1}{|a|}G\left(\frac{f}{a}\right) \tag{2.20}$$

To prove this property, we note that

$$F[g(at)] = \int_{-\infty}^{\infty} g(at)\exp(-j2\pi ft)\,dt$$

Set $\tau = at$. There are two cases that can arise, depending on whether the scaling factor a is positive or negative. If $a > 0$, we get

$$F[g(at)] = \frac{1}{a}\int_{-\infty}^{\infty} g(\tau)\exp\left[-j2\pi\left(\frac{f}{a}\right)\tau\right]\,d\tau$$
$$= \frac{1}{a}G\left(\frac{f}{a}\right)$$

On the other hand, if $a < 0$, the limits of integration are interchanged so that we have the multiplying factor $-(1/a)$ or, equivalently, $1/|a|$. This completes the proof of Eq. (2.20).

Note that the function $g(at)$ represents $g(t)$ *compressed* in time by a factor a, whereas the function $G(f/a)$ represents $G(f)$ *expanded* in frequency by the same factor a. Thus, the scaling property states that the compression of a function $g(t)$ in the time domain is equivalent to the expansion of its Fourier transform $G(f)$ in the frequency domain by the same factor, or vice versa.

For the special case when $a = -1$, we readily find from Eq. (2.20) that

$$g(-t) \rightleftharpoons G(-f) \tag{2.21}$$

In words, if a function $g(t)$ has a Fourier transform given by $G(f)$, then the Fourier transform of $g(-t)$ is $G(-f)$.

PROPERTY 3 Duality

If $g(t) \rightleftharpoons G(f)$, *then*

$$G(t) \rightleftharpoons g(-f) \tag{2.22}$$

This property follows from the relation defining the inverse Fourier transform by writing it in the form:

$$g(-t) = \int_{-\infty}^{\infty} G(f)\exp(-j2\pi ft)\, df$$

and then interchanging t and f.

EXAMPLE 2.4 Sinc Pulse

Consider a signal $g(t)$ in the form of a sinc function, as shown by

$$g(t) = A\,\mathrm{sinc}(2Wt)$$

To evaluate the Fourier transform of this function, we apply the duality and time-scaling properties to the Fourier-transform pair of Eq. (2.10). Then, recognizing that the rectangular function is an even function of time, we obtain the following result:

$$A\,\mathrm{sinc}(2Wt) \rightleftharpoons \frac{A}{2W}\,\mathrm{rect}\left(\frac{f}{2W}\right) \tag{2.23}$$

which is illustrated in Figure 2.7. We thus see that the Fourier transform of a sinc pulse is zero for $|f| > W$. Note also that the sinc pulse itself is only asymptotically limited in time in the sense that it approaches zero as time t approaches infinity.

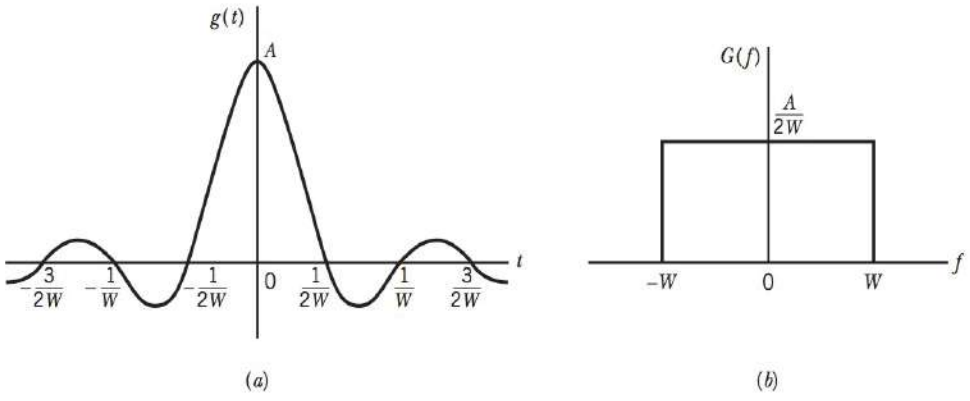


FIGURE 2.7 (*a*) Sinc pulse $g(t)$. (*b*) Fourier transform $G(f)$.

PROPERTY 4 Time Shifting

If $g(t) \rightleftharpoons G(f)$, *then*

$$g(t - t_0) \rightleftharpoons G(f)\exp(-j2\pi f t_0) \tag{2.24}$$

To prove this property, we take the Fourier transform of $g(t - t_0)$ and then set $\tau = (t - t_0)$ to obtain

$$\begin{aligned} F[g(t - t_0)] &= \exp(-j2\pi f t_0) \int_{-\infty}^{\infty} g(\tau)\exp(-j2\pi f \tau)\, d\tau \\ &= \exp(-j2\pi f t_0)\, G(f) \end{aligned}$$

The time-shifting property states that if a function $g(t)$ is shifted in the positive direction by an amount t_0, the effect is equivalent to multiplying its Fourier transform $G(f)$ by the factor $\exp(-j2\pi f t_0)$. This means that the amplitude of $G(f)$ is unaffected by the time shift, but its phase is changed by the linear factor $-2\pi f t_0$.

PROPERTY 5 Frequency Shifting

If $g(t) \rightleftharpoons G(f)$, *then*

$$\exp(j2\pi f_c t)\, g(t) \rightleftharpoons G(f - f_c) \tag{2.25}$$

where f_c *is a real constant.*

This property follows from the fact that

$$\begin{aligned} F[\exp(j2\pi f_c t)g(t)] &= \int_{-\infty}^{\infty} g(t)\exp[-j2\pi t(f - f_c)]\, dt \\ &= G(f - f_c) \end{aligned}$$

That is, multiplication of a function $g(t)$ by the factor $\exp(j2\pi f_c t)$ is equivalent to shifting its Fourier transform $G(f)$ in the positive direction by the amount f_c. This property is called the *modulation theorem*, because a shift of the range of frequencies in a signal is accomplished by using modulation. Note the duality between the time-shifting and frequency-shifting operations described in Eqs. (2.24) and (2.25).

EXAMPLE 2.5 Radio Frequency (RF) Pulse

Consider the pulse signal $g(t)$ shown in Figure 2.8*a* which consists of a sinusoidal wave of amplitude A and frequency f_c, extending in duration from $t = -T/2$ to $t = T/2$. This signal is sometimes referred to as an *RF pulse* when the frequency f_c falls in the radio-frequency band. The signal $g(t)$ of Figure 2.8*a* may be expressed mathematically as follows:

$$g(t) = A\,\text{rect}\left(\frac{t}{T}\right)\cos(2\pi f_c t) \tag{2.26}$$

To find the Fourier transform of this signal, we note that

$$\cos(2\pi f_c t) = \frac{1}{2}\left[\exp(j2\pi f_c t) + \exp(-j2\pi f_c t)\right]$$

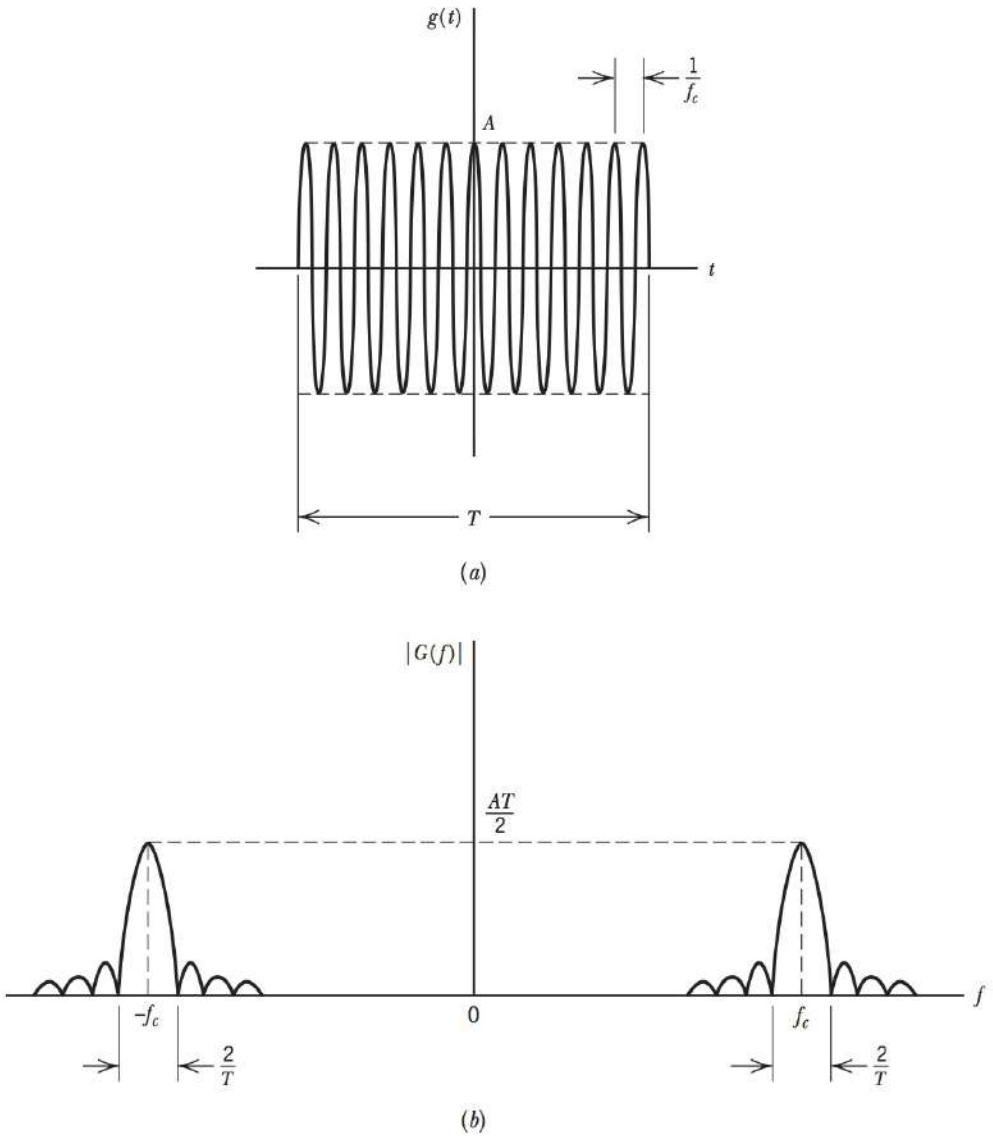


FIGURE 2.8 (*a*) RF pulse. (*b*) Amplitude spectrum.

Therefore, applying the frequency-shifting property to the Fourier-transform pair of Eq. (2.10), we get the desired result

$$G(f) = \frac{AT}{2}\{\text{sinc}[T(f - f_c)] + \text{sinc}[T(f + f_c)]\} \tag{2.27}$$

In the special case of $f_c T \gg 1$, we may use the approximate result

$$G(f) \simeq \begin{cases} \dfrac{AT}{2}\,\text{sinc}[T(f - f_c)], & f > 0 \\ 0, & f = 0 \\ \dfrac{AT}{2}\,\text{sinc}[T(f + f_c)], & f < 0 \end{cases} \tag{2.28}$$

The amplitude spectrum of the RF pulse is shown in Figure 2.8*b*. This diagram, in relation to Figure 2.2*b*, clearly illustrates the frequency-shifting property of the Fourier transform.

PROPERTY 6 Area Under $g(t)$

If $g(t) \rightleftharpoons G(f)$, *then*

$$\int_{-\infty}^{\infty} g(t)\,dt = G(0) \tag{2.29}$$

That is, the area under a function $g(t)$ *is equal to the value of its Fourier transform* $G(f)$ *at* $f = 0$.

This result is obtained simply by putting $f = 0$ in the formula defining the Fourier transform of the function $g(t)$.

PROPERTY 7 Area Under $G(f)$

If $g(t) \rightleftharpoons G(f)$, *then*

$$g(0) = \int_{-\infty}^{\infty} G(f)\,df \tag{2.30}$$

That is, the value of a function $g(t)$ *at* $t = 0$ *is equal to the area under its Fourier transform* $G(f)$.

The result is obtained simply by putting $t = 0$ in the formula defining the inverse Fourier transform of $G(f)$.

PROPERTY 8 Differentiation in the Time Domain

Let $g(t) \rightleftharpoons G(f)$, *and assume that the first derivative of* $g(t)$ *is Fourier transformable. Then*

$$\frac{d}{dt} g(t) \rightleftharpoons j2\pi f\, G(f) \tag{2.31}$$

That is, differentiation of a time function $g(t)$ *has the effect of multiplying its Fourier transform* $G(f)$ *by the factor* $j2\pi f$.

This result is obtained simply by taking the first derivative of both sides of the integral defining the inverse Fourier transform of $G(f)$, and then interchanging the operations of integration and differentiation.

We may generalize Eq. (2.31) as follows:

$$\frac{d^n}{dt^n} g(t) \rightleftharpoons (j2\pi f)^n\, G(f) \tag{2.32}$$

Equation (2.32) assumes that the Fourier transform of the higher-order derivative exists.

EXAMPLE 2.6 Gaussian Pulse

In this example we use the differentiation property of the Fourier transform to derive the particular form of a pulse signal that has the same mathematical form as its own Fourier transform.

Let $g(t)$ denote the pulse expressed as a function of time, and $G(f)$ its Fourier transform. We note that by differentiating the formula for the Fourier transform $G(f)$ with respect to f, we have

$$-j2\pi t g(t) \rightleftharpoons \frac{d}{df} G(f) \tag{2.33}$$

which expresses the effect of differentiation in the frequency domain.

If we add Eq. (2.31) plus j times Eq. (2.33) we obtain the relationship

$$\frac{dg(t)}{dt} + 2\pi t g(t) \rightleftharpoons j\left[\frac{dG(f)}{df} + 2\pi f G(f)\right] \tag{2.34}$$

In particular, the relationship (2.34) holds if both the left-hand side and right-hand side are zero. That is, if

$$\frac{dg(t)}{dt} = -2\pi t g(t) \tag{2.35}$$

then an identical differential equation holds for $G(f)$ with the indeterminate f. Under this condition, the pulse signal $g(t)$ and its Fourier transform $G(f)$ satisfy the same differential equation; hence they are the same function. In other words, provided that the pulse signal $g(t)$ satisfies the differential equation (2.35), then $G(f) = g(f)$, where $g(f)$ is obtained from $g(t)$ by substituting f for t. Solving Eq. (2.35) for $g(t)$, we obtain

$$g(t) = \exp(-\pi t^2) \tag{2.36}$$

The pulse defined by Eq. (2.36) is called a *Gaussian pulse*, the name being derived from the similarity of the function to the Gaussian density function of probability theory (see Chapter 5). It is shown plotted in Figure 2.9. By applying Eq. (2.29), we find that the area under this Gaussian pulse is unity, as shown by

$$\int_{-\infty}^{\infty} \exp(-\pi t^2)\, dt = 1 \tag{2.37}$$

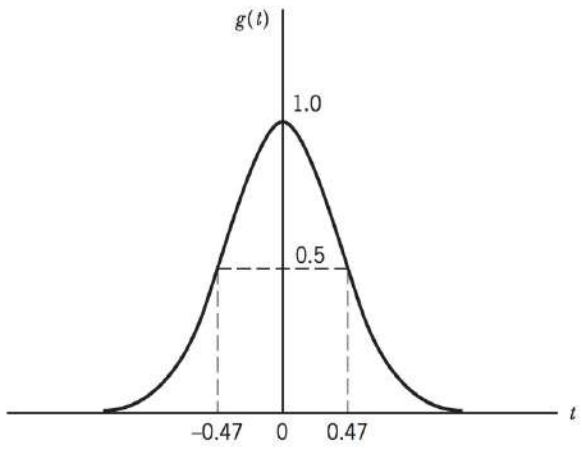


FIGURE 2.9 Gaussian pulse.

When the area under the curve of a pulse is unity, as in the case of the Gaussian pulse of Eq. (2.36), we say that the pulse is *normalized*. We conclude therefore that the normalized Gaussian pulse is its own Fourier transform, as shown by

$$\exp(-\pi t^2) \rightleftharpoons \exp(-\pi f^2) \tag{2.38}$$

PROPERTY 9 Integration in the Time Domain

Let $g(t) \rightleftharpoons G(f)$. *Then provided that* $G(0) = 0$, *we have*

$$\int_{-\infty}^{t} g(\tau)\, d\tau \rightleftharpoons \frac{1}{j2\pi f} G(f) \tag{2.39}$$

That is, integration of a time function $g(t)$ *has the effect of dividing its Fourier transform* $G(f)$ *by the factor* $j2\pi f$, *assuming that* $G(0)$ *is zero.*

The result is obtained by expressing $g(t)$ as

$$g(t) = \frac{d}{dt}\left[\int_{-\infty}^{t} g(\tau)\, d\tau\right]$$

and then applying the time-differentiation property of the Fourier transform to obtain

$$G(f) = j2\pi f\left\{F\left[\int_{-\infty}^{t} g(\tau)\, d\tau\right]\right\}$$

from which Eq. (2.39) follows readily.

It is a straightforward matter to generalize Eq. (2.39) to multiple integration; however, the notation becomes rather cumbersome.

Equation (2.39) assumes that $G(0)$, that is, the area under $g(t)$, is zero. The more general case pertaining to $G(0) \neq 0$ is considered in Section 2.5.

EXAMPLE 2.7 Triangular Pulse

Consider the *doublet pulse* $g_1(t)$ shown in Figure 2.10*a*. By integrating this pulse with respect to time, we obtain the *triangular pulse* $g_2(t)$ shown in Figure 2.10*b*. We note that the doublet pulse $g_1(t)$ consists of two rectangular pulses: one of amplitude A, defined for the interval $-T \leq t \leq 0$; and the other of amplitude $-A$, defined for the interval $0 \leq t \leq T$. Applying the time-shifting property of the Fourier transform to Eq. (2.10), we find that the Fourier transforms of these two rectangular pulses are equal to $AT\,\mathrm{sinc}(fT)\exp(j\pi fT)$ and $-AT\,\mathrm{sinc}(fT)\exp(-j\pi fT)$, respectively. Hence, invoking the linearity property of the Fourier transform, we find that the Fourier transform $G_1(f)$ of the doublet pulse $g_1(t)$ of Figure 2.10*a* is given by

$$\begin{aligned} G_1(f) &= AT\,\mathrm{sinc}(fT)[\exp(j\pi fT) - \exp(-j\pi fT)] \\ &= 2jAT\,\mathrm{sinc}(fT)\sin(\pi fT) \end{aligned} \tag{2.40}$$

We further note that $G_1(0)$ is zero. Hence, using Eqs. (2.39) of Property 9 and (2.40), we find that the Fourier transform $G_2(f)$ of the triangular pulse $g_2(t)$ of Figure 2.10*b* is given by

$$\begin{aligned} G_2(f) &= \frac{1}{j2\pi f} G_1(f) = AT \frac{\sin(\pi f T)}{\pi f} \operatorname{sinc}(fT) \\ &= AT^2 \operatorname{sinc}^2(fT) \end{aligned} \tag{2.41}$$

Note that the doublet pulse of Figure 2.10*a* is real and odd-symmetric and its Fourier transform is therefore odd and purely imaginary, whereas the triangular pulse of Figure 2.10*b* is real and symmetric and its Fourier transform is therefore symmetric and purely real.

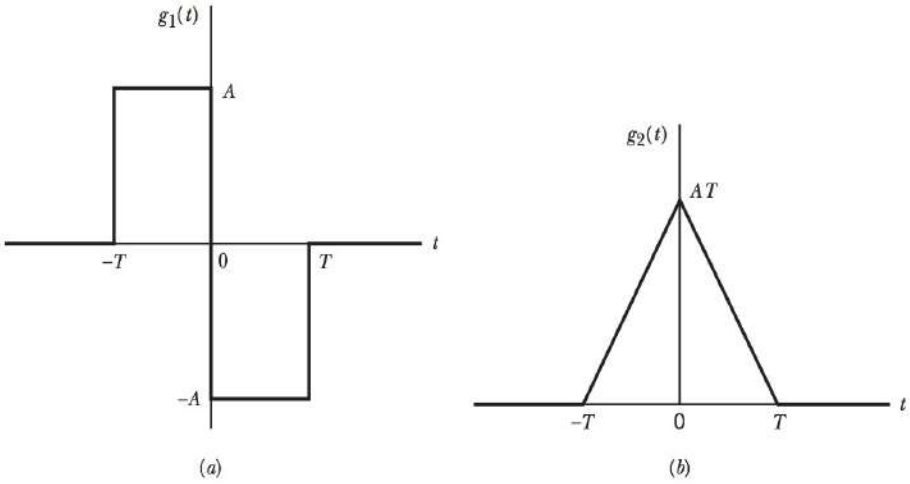


FIGURE 2.10 (*a*) Doublet pulse $g_1(t)$. (*b*) Triangular pulse $g_2(t)$ obtained by integrating $g_1(t)$.

PROPERTY 10 Conjugate Functions

If $g(t) \rightleftharpoons G(f)$, *then for a complex-valued time function* $g(t)$ *we have*

$$g^*(t) \rightleftharpoons G^*(-f), \tag{2.42}$$

where the asterisk denotes the complex conjugate operation.

To prove this property, we know from the inverse Fourier transform that

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft)\, df$$

Taking the complex conjugates of both sides yields

$$g^*(t) = \int_{-\infty}^{\infty} G^*(f) \exp(-j2\pi ft)\, df$$

Next, replacing f with $-f$ gives

$$\begin{aligned} g^*(t) &= -\int_{\infty}^{-\infty} G^*(-f) \exp(j2\pi ft)\, df \\ &= \int_{-\infty}^{\infty} G^*(-f) \exp(j2\pi ft)\, df \end{aligned}$$

That is, $g^*(t)$ is the inverse Fourier transform of $G^*(-f)$, which is the desired result.

As a corollary to Property 10, we may state that

$$g^*(-t) \rightleftharpoons G^*(f) \tag{2.43}$$

This result follows directly from Eq. (2.42) by applying the special form of the scaling property described in Eq. (2.21).

EXAMPLE 2.8 Real and Imaginary Parts of a Time Function

Expressing a complex-valued function $g(t)$ in terms of its real and imaginary parts, we may write

$$g(t) = \mathrm{Re}[g(t)] + j\,\mathrm{Im}[g(t)] \tag{2.44}$$

where Re denotes "the real part of" and Im denotes the "imaginary part of." The complex conjugate of $g(t)$ is

$$g^*(t) = \mathrm{Re}[g(t)] - j\,\mathrm{Im}[g(t)] \tag{2.45}$$

Adding Eqs. (2.44) and (2.45) gives

$$\mathrm{Re}[g(t)] = \frac{1}{2}[g(t) + g^*(t)] \tag{2.46}$$

and subtracting them yields

$$\mathrm{Im}[g(t)] = \frac{1}{2j}[g(t) - g^*(t)] \tag{2.47}$$

Therefore, applying Property 10, we obtain the following two Fourier-transform pairs:

$$\begin{aligned} \mathrm{Re}[g(t)] &\rightleftharpoons \frac{1}{2}[G(f) + G^*(-f)] \\ \mathrm{Im}[g(t)] &\rightleftharpoons \frac{1}{2j}[G(f) - G^*(-f)] \end{aligned} \tag{2.48}$$

From Eq. (2.48), it is apparent that in the case of a real-valued time function $g(t)$, we have $G(f) = G^*(-f)$, that is, $G(f)$ exhibits *conjugate symmetry*, confirming a result that we stated previously in Section 2.2.

PROPERTY 11 Multiplication in the Time Domain

Let $g_1(t) \rightleftharpoons G_1(f)$ *and* $g_2(t) \rightleftharpoons G_2(f)$. *Then*

$$g_1(t)\,g_2(t) \rightleftharpoons \int_{-\infty}^{\infty} G_1(\lambda)\,G_2(f-\lambda)\,d\lambda \tag{2.49}$$

To prove this property, we first denote the Fourier transform of the product $g_1(t)\,g_2(t)$ by $G_{12}(f)$, so that we may write

$$g_1(t)\,g_2(t) \rightleftharpoons G_{12}(f)$$

where

$$G_{12}(f) = \int_{-\infty}^{\infty} g_1(t)g_2(t)\exp(-j2\pi ft)\,dt$$

For $g_2(t)$, we next substitute the inverse Fourier transform

$$g_2(t) = \int_{-\infty}^{\infty} G_2(f')\exp(j2\pi f't)\,df'$$

in the integral defining $G_{12}(f)$ to obtain

$$G_{12}(f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(t)\,G_2(f')\exp[-j2\pi\,(f-f')t]\,df'\,dt$$

Define $\lambda = f - f'$. Then, interchanging the order of integration, we obtain

$$G_{12}(f) = \int_{-\infty}^{\infty} d\lambda G_2(f-\lambda)\int_{-\infty}^{\infty} g_1(t)\exp(-j2\pi\lambda t)\,dt$$

The inner integral is recognized simply as $G_1(\lambda)$, so we may write

$$G_{12}(f) = \int_{-\infty}^{\infty} G_1(\lambda)\,G_2(f-\lambda)\,d\lambda$$

which is the desired result. This integral is known as the *convolution integral* expressed in the frequency domain, and the function $G_{12}(f)$ is referred to as the *convolution* of $G_1(f)$ and $G_2(f)$. We conclude that *the multiplication of two signals in the time domain is transformed into the convolution of their individual Fourier transforms in the frequency domain*. This property is known as the *multiplication theorem*.

In a discussion of convolution, the following shorthand notation is frequently used:

$$G_{12}(f) = G_1(f) \star G_2(f)$$

Accordingly, we may reformulate Eq. (2.49) in the following symbolic form:

$$g_1(t)\,g_2(t) \rightleftharpoons G_1(f) \star G_2(f) \tag{2.50}$$

Note that convolution is *commutative*, that is,

$$G_1(f) \star G_2(f) = G_2(f) \star G_1(f)$$

which follows directly from Eq. (2.50).

PROPERTY 12 Convolution in the Time Domain

Let $g_1(t) \rightleftharpoons G_1(f)$ *and* $g_2(t) \rightleftharpoons G_2(f)$. *Then*

$$\int_{-\infty}^{\infty} g_1(\tau)g_2(t-\tau)\,d\tau \rightleftharpoons G_1(f)G_2(f) \tag{2.51}$$

This result follows directly by combining Property 3 (duality) and Property 11 (time-domain multiplication). We may thus state that the *convolution of two signals in the time domain is transformed into the multiplication of their individual Fourier*

transforms in the frequency domain. This property is known as the *convolution theorem.* Its use permits us to exchange a convolution operation for a transform multiplication, an operation that is ordinarily easier to manipulate.

Using the shorthand notation for convolution, we may rewrite Eq. (2.51) in the form

$$g_1(t) \star g_2(t) \rightleftharpoons G_1(f)G_2(f) \tag{2.52}$$

where the symbol $\star$ denotes convolution.

Note that Properties 11 and 12, described by Eqs. (2.49) and (2.51), respectively, are the dual of each other.

PROPERTY 13 Rayleigh's Energy Theorem

Let $g(t)$ be defined over the entire interval $-\infty < t < \infty$ and assume its Fourier transform $G(f)$ exists. If the energy of the signal satisfies

$$E = \int_{-\infty}^{\infty} |g(t)|^2 \, dt < \infty \tag{2.53}$$

then

$$\int_{-\infty}^{\infty} |g(t)|^2 \, dt = \int_{-\infty}^{\infty} |G(f)|^2 \, df \tag{2.54}$$

This result follows from noting that the energy intensity $|g(t)|^2$ may be expressed as the product of two time functions, namely, $g(t)$ and its complex conjugate $g^*(t)$. The Fourier transform of $g^*(t)$ is equal to $G^*(-f)$, by virtue of Property 10 (complex conjugation). Then, applying Property 11 (the multiplication theorem) or, more specifically, applying Eq. (2.49) to the product $g(t)\, g^*(t)$ and evaluating the result for $f = 0$, we obtain the relation

$$\int_{-\infty}^{\infty} g(t)g^*(t) \, dt = \int_{-\infty}^{\infty} G(\lambda)G^*(\lambda) d\lambda \tag{2.55}$$

which is equivalent to Eq. (2.54).

Let $\mathcal{E}_g(f)$ denote the squared amplitude spectrum of the signal $g(t)$, as shown by

$$\mathcal{E}_g(f) = |G(f)|^2 \tag{2.56}$$

The quantity $\mathcal{E}_g(f)$ is referred to as the *energy spectral density*[3] of the signal $g(t)$. To explain this meaning of the definition, suppose $g(t)$ denotes the voltage of a source connected to a 1-ohm load resistor. Then the quantity

$$\int_{-\infty}^{\infty} |g(t)|^2 \, dt$$

equals the total energy delivered by the source. According to Rayleigh's theorem, this energy equals the total area under the $\mathcal{E}_g(f)$ curve. If follows, therefore, that the function

$\varepsilon_g(f)$ is a measure of the density of the energy contained in $g(t)$ in joules per Hertz. Note that since in the special case of a real-valued signal the amplitude spectrum is an even function of f, the energy spectral density of such a signal is symmetrical about the vertical axis passing through the origin.

EXAMPLE 2.9 Sinc Pulse (continued)

Consider again the sinc pulse $A\,\text{sinc}(2Wt)$. The energy of this pulse equals

$$E = A^2 \int_{-\infty}^{\infty} \text{sinc}^2(2Wt)\,dt$$

The integral in the right-hand side of this equation is rather difficult to evaluate. However, we note from Example 2.4 that the Fourier transform of the sinc pulse $A\text{sinc}(2Wt)$ is equal to $(A/2W)\text{rect}(f/2W)$; hence, applying Rayleigh's energy theorem to the problem at hand, we readily obtain the desired result:

$$\begin{aligned} E &= \left(\frac{A}{2W}\right)^2 \int_{-\infty}^{\infty} \text{rect}^2\left(\frac{f}{2W}\right) df \\ &= \left(\frac{A}{2W}\right)^2 \int_{-W}^{W} df \\ &= \frac{A^2}{2W} \end{aligned} \tag{2.57}$$

This example clearly illustrates the usefulness of Rayleigh's energy theorem.

2.4 THE INVERSE RELATIONSHIP BETWEEN TIME AND FREQUENCY

The properties of the Fourier transform discussed in Section 2.3 clearly show that the time-domain and frequency-domain description of a signal are *inversely* related. In particular, we may make the following important statements:

1. If the time-domain description of a signal is changed, then the frequency-domain description of the signal is changed in an *inverse* manner, and vice versa. This inverse relationship prevents arbitrary specifications of a signal in both domains. In other words, *we may specify an arbitrary function of time or an arbitrary spectrum, but we cannot specify both of them together.*
2. If a signal is strictly limited in frequency, then the time-domain description of the signal will trail on indefinitely, even though its amplitude may assume a progressively smaller value. We say a signal is *strictly limited in frequency* or *strictly band limited* if its Fourier transform is exactly zero outside a finite band of frequencies. The sinc pulse is an example of a strictly band-limited signal, as illustrated in Figure 2.7. This figure also shows that the sinc pulse is only *asymptotically limited in time*, which confirms the opening statement we made for a strictly band-limited signal. In an inverse

manner, if a signal is *strictly limited in time* (i.e., the signal is exactly zero outside a finite time interval), then the spectrum of the signal is infinite in extent, even though the amplitude spectrum may assume a progressively smaller value. This behavior is exemplified by both the rectangular pulse (described in Figure 2.2) and the triangular pulse (described in Figure 2.10*b*). Accordingly, we may state that *a signal cannot be strictly limited in both time and frequency*.

BANDWIDTH

The *bandwidth* of a signal provides a measure of the *extent of significant spectral content of the signal for positive frequencies*. When the signal is strictly band limited, the bandwidth is well defined. For example, the sinc pulse described in Figure 2.7 has a bandwidth equal to W. When, however, the signal is not strictly band limited, which is generally the case, we encounter difficulty in defining the bandwidth of the signal. The difficulty arises because the meaning of "significant" attached to the spectral content of the signal is mathematically imprecise. Consequently, there is no universally accepted definition of bandwidth.

Nevertheless, there are some commonly used definitions for bandwidth. In this section, we consider three such definitions; the formulation of each definition depends on whether the signal is low-pass or band-pass. A signal is said to be *low-pass* if its significant spectral content is centered around the origin. A signal is said to be *band-pass* if its significant spectral content is centered around $\pm f_c$, where f_c is a nonzero frequency.

When the spectrum of a signal is symmetric with a *main lobe* bounded by well-defined *nulls* (i.e., frequencies at which the spectrum is zero), we may use the main lobe as the basis for defining the bandwidth of the signal. Specifically, if the signal is low-pass, the bandwidth is defined as one half the total width of the main spectral lobe, since only one half of this lobe lies inside the positive frequency region. For example, a rectangular pulse of duration T seconds has a main spectral lobe of total width $2/T$ Hertz centered at the origin, as depicted in Figure 2.2. Accordingly, we may define the bandwidth of this rectangular pulse as $1/T$ Hertz. If, on the other hand, the signal is band-pass with main spectral lobes centered around $\pm f_c$, where f_c is large, the bandwidth is defined as the width of the main lobe for positive frequencies. This definition of bandwidth is called the *null-to-null bandwidth*. For example, an RF pulse of duration T seconds and frequency f_c has main spectral lobes of width $2/T$ Hertz centered around $\pm f_c$, as depicted in Figure 2.8 and Figure 2.11*a*. Hence, we may define the null-to-null bandwidth of this RF pulse as $(2/T)$ Hertz. On the basis of the definitions presented here, we may state that shifting the spectral content of a low-pass signal by a sufficiently large frequency has the effect of doubling the bandwidth of the signal; such a frequency translation is attained by using modulation.

Another popular definition of bandwidth is the *3-dB bandwidth*. Specifically, if the signal is low-pass, the 3-dB bandwidth is defined as the separation between zero frequency, where the amplitude spectrum attains its peak value, and the *positive frequency* at which the amplitude spectrum drops to $1/\sqrt{2}$ of its peak value. For example, the decaying exponential and rising exponential pulses defined in Figure 2.4 have a 3-dB bandwidth of $a/2\pi$ Hertz. If, on the other hand, the signal is band-pass, centered at $\pm f_c$, the 3-dB bandwidth is defined as the separation (along the positive frequency axis) between the two frequencies at which the amplitude spectrum of the signal

Power Ratios and Decibels

In communications systems, the power levels may vary from megawatts for the case of a television transmitter to 10^{-12} watts at a satellite receiver. To manage this wide range of powers, it is customary practice to use a unit called the *decibel*. The decibel, commonly abbreviated as dB, is one tenth of a larger unit, the *bel*. The unit bel is named in honor of Alexander Graham Bell. In addition to inventing the telephone, Bell was the first to use logarithmic power measurements in sound and hearing research. In practice, however, we find that for most applications the bel is too large a unit, hence, the wide use of decibel as the unit for expressing power ratios.

Let P denote the power at some point of interest in a system. Let P_0 denote the reference power level with respect to which power P is compared. The number of decibels in the power ratio P/P_0 is defined as

$$\left(\frac{P}{P_0}\right)_{dB} = 10 \log_{10}\left(\frac{P}{P_0}\right)$$

For example, a power ratio of 2 approximately corresponds to 3 dB, and a power ratio of 10 exactly corresponds to 10 dB.

We may also express the signal power relative to one watt or one milliwatt. In the first case, we express the signal power P in dBW as $10 \log_{10}(P/1\text{W})$ where W is the abbreviation for watt. In the second case, we express the signal power P in dBm as $10 \log_{10}(P/1\text{ mW})$, where mW is the abbreviation for milliwatt.

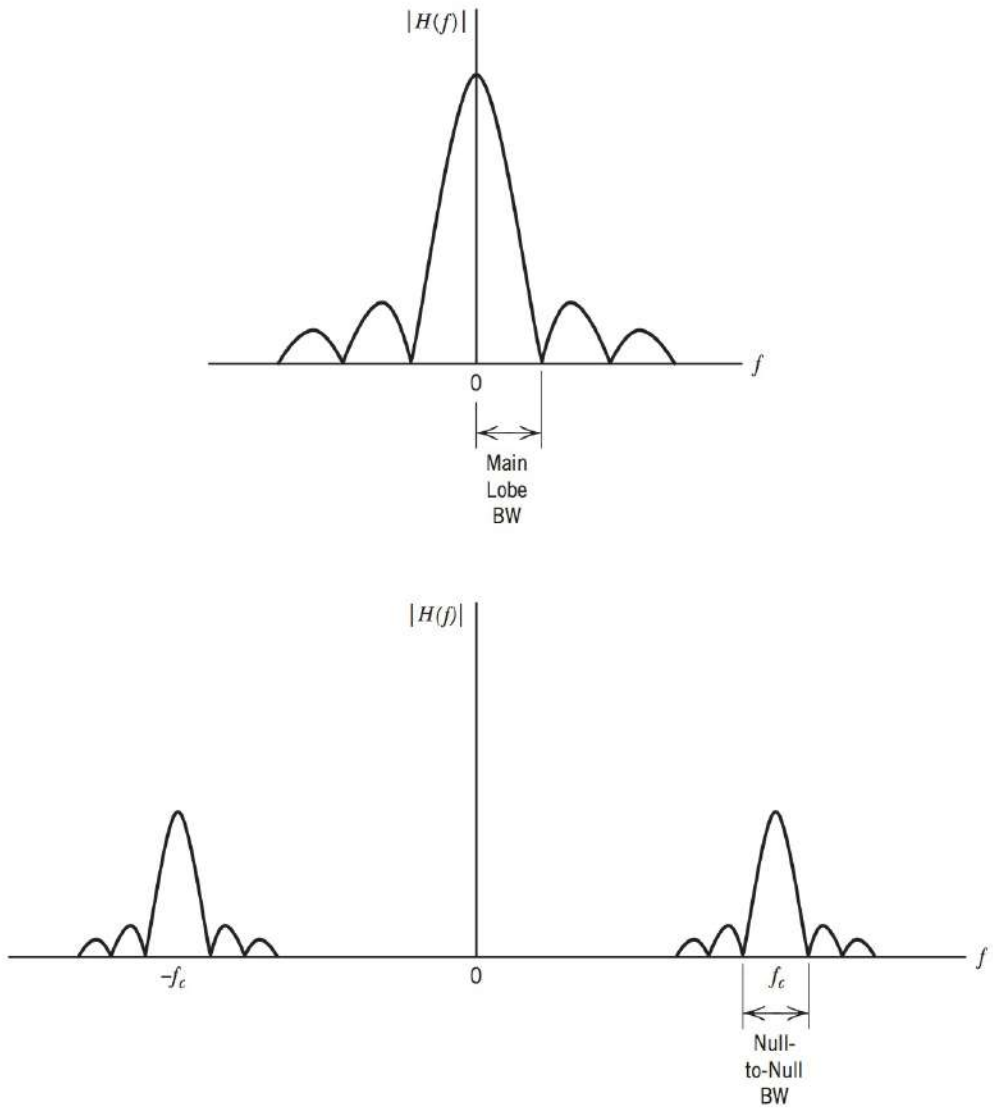


FIGURE 2.11 (*a*) Illustration of null-to-null bandwidth in the baseband and pass-band cases.

drops to $\frac{1}{\sqrt{2}}$ of the peak value at f_c, as depicted in Figure 2.11*b*. The 3-dB bandwidth has the advantage, similar to the null-to-null bandwidth, in that it can be read directly from a plot of the amplitude spectrum. However, it has the disadvantage in that it may be misleading if the amplitude spectrum has slowly decreasing tails.

Yet another measure for the bandwidth of a signal is the *root mean square (rms) bandwidth*, defined as the square root of the second moment of a properly normalized form of the squared amplitude spectrum of the signal about a suitably chosen point. We assume that the signal is low-pass, so that the second moment may be taken about the origin. As for the normalized form of the squared amplitude spectrum, we use the nonnegative function $|G(f)|^2 / \int_{-\infty}^{\infty} |G(f)|^2 \, df$, in which the denominator applies the correct normalization in the sense that the integrated value of this ratio over the entire frequency axis is unity. We may thus formally define the rms bandwidth of a low-pass signal $g(t)$ with Fourier transform $G(f)$ as follows:

$$W_{\text{rms}} = \left(\frac{\displaystyle\int_{-\infty}^{\infty} f^2 |G(f)|^2 \, df}{\displaystyle\int_{-\infty}^{\infty} |G(f)|^2 \, df} \right)^{1/2} \tag{2.58}$$

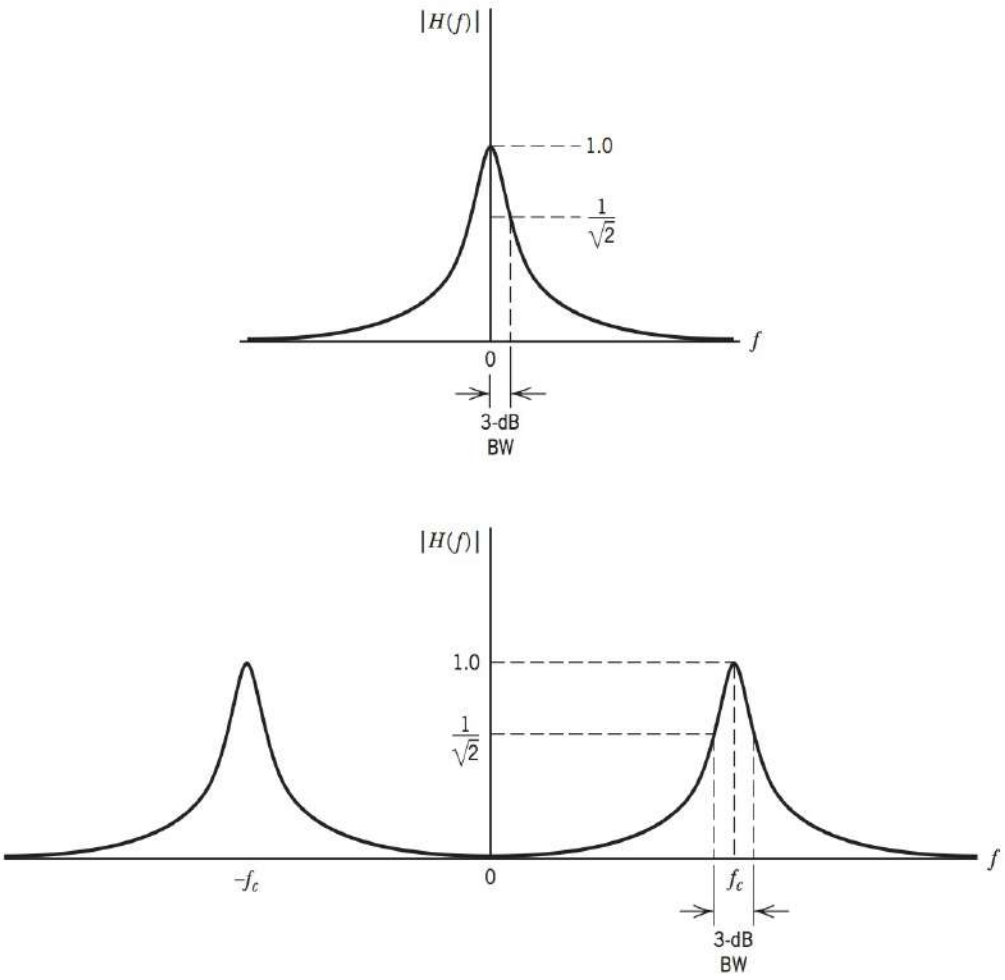


FIGURE 2.11 (*b*) Illustration of 3-dB bandwidth in the baseband and pass-band cases. *Note*: The term "baseband" refers to the frequency band occupied by an original source of information.

An attractive feature of the rms bandwidth W_{rms} is that it lends itself more readily to mathematical evaluation than the other two definitions of bandwidth, but it is not as easily measurable in the laboratory.

TIME–BANDWIDTH PRODUCT

For any family of pulse signals (e.g., the exponential pulse of Example 2.2) that differ in time scale, the product of the signal's duration and its bandwidth is always a constant, as shown by

$$(\text{duration}) \cdot (\text{bandwidth}) = \text{constant}$$

The product is called the *time–bandwidth product* or *bandwidth–duration product*. The constancy of the time–bandwidth product is another manifestation of the inverse relationship that exists between the time-domain and frequency-domain descriptions of a signal. In particular, if the duration of a pulse signal is decreased by reducing the time scale by a factor a, the frequency scale of the signal's spectrum, and therefore the bandwidth of the signal, is increased by the same factor a, by virtue of Property 2 (time scaling), and the time–bandwidth product of the signal is thereby maintained constant. For example, a rectangular pulse of duration T seconds has a bandwidth (defined on the basis of the positive-frequency part of the main lobe) equal to $1/T$ Hertz, making the time–bandwidth product of the pulse equal unity. Whatever definition we use for the bandwidth of a signal, the time–bandwidth product remains constant over certain classes of pulse signals. The choice of a particular definition for bandwidth merely changes the value of the constant.

To be more specific, consider the rms bandwidth defined in Eq. (2.58). The corresponding definition for the *rms duration* of the signal $g(t)$ is

$$T_{\text{rms}} = \left(\frac{\int_{-\infty}^{\infty} t^2 |g(t)|^2 \, dt}{\int_{-\infty}^{\infty} |g(t)|^2 \, dt} \right)^{1/2} \tag{2.59}$$

where it is assumed that the signal $g(t)$ is centered around the origin. It may be shown that using the rms definitions of Eqs. (2.58) and (2.59), the time–bandwidth product has the following form:

$$T_{\text{rms}} W_{\text{rms}} \geq \frac{1}{4\pi} \tag{2.60}$$

where the constant is $1/4\pi$. It may also be shown that the Gaussian pulse satisfies this condition with the equality sign. For the details of these calculations, the reader is referred to Problem 2.14.

2.5 DIRAC DELTA FUNCTION

Strictly speaking, the theory of the Fourier transform, as described in Sections 2.2 to 2.4, is applicable only to time functions that satisfy the Dirichlet conditions. Such functions include energy signals. However, it would be highly desirable to extend this theory in two ways:

1. To combine the Fourier series and Fourier transform into a unified theory, so that the Fourier series may be treated as a special case of the Fourier transform.
2. To include power signals in the list of signals to which we may apply the Fourier transform.

It turns out that both of these objectives can be met through the "proper use" of the *Dirac delta function* or *unit impulse*.

The Dirac delta function,[4] denoted by $\delta(t)$, is defined as having zero amplitude everywhere except at $t = 0$, where it is infinitely large in such a way that it contains unit area under its curve; this is,

$$\delta(t) = 0, \qquad t \neq 0 \tag{2.61}$$

and

$$\int_{-\infty}^{\infty} \delta(t) \, dt = 1 \tag{2.62}$$

An implication of this pair of relations is that the delta function $\delta(t)$ must be an even function of time t.

For the delta function to have meaning, however, it has to appear as a factor in the integrand of an integral with respect to time and then, strictly speaking, only when the other factor in the integrand is a continuous function of time. Let $g(t)$ be such a function, and consider the product of $g(t)$ and the time-shifted delta function $\delta(t - t_0)$. In light of the two defining equations (2.61) and (2.62), we may express the integral of this product as follows:

$$\int_{-\infty}^{\infty} g(t) \, \delta(t - t_0) \, dt = g(t_0) \tag{2.63}$$

The operation indicated on the left-hand side of this equation sifts out the value $g(t_0)$ of the function $g(t)$ at time $t = t_0$, where $-\infty < t < \infty$. Accordingly, Eq. (2.63) is referred to as the *sifting property* of the delta function. This property is sometimes used as the defining equation of a delta function; in effect, it incorporates Eqs. (2.61) and (2.62) into a single relation.

Noting that the delta function $\delta(t)$ is an even function of t, we may rewrite Eq. (2.63) in a way emphasizing its resemblance to the convolution integral, as shown by

$$\int_{-\infty}^{\infty} g(\tau)\,\delta(t-\tau)\,d\tau = g(t) \tag{2.64}$$

or, using the notation for convolution:

$$g(t) \star \delta(t) = g(t)$$

In words, the convolution of any function with the delta function leaves that function unchanged. We refer to this statement as the *replication property* of the delta function.

By definition, the Fourier transform of the delta function is given by

$$F[\delta(t)] = \int_{-\infty}^{\infty} \delta(t)\exp(-j2\pi ft)\,dt$$

Hence, using the sifting property of the delta function and noting that $\exp(-j2\pi ft)$ is equal to unity at $t = 0$, we obtain

$$F[\delta(t)] = 1$$

We thus have the Fourier-transform pair for the Dirac delta function:

$$\delta(t) \rightleftharpoons 1 \tag{2.65}$$

This relation states that the spectrum of the delta function $\delta(t)$ extends uniformly over the entire frequency interval, as shown in Figure 2.12.

It is important to realize that the Fourier-transform pair of Eq. (2.65) exists only in a limiting sense. The point is that no function in the ordinary sense has the two properties of Eqs. (2.61) and (2.62) or the equivalent sifting property of Eq. (2.63). However, we can imagine a sequence of functions that have progressively taller and thinner peaks at $t = 0$, with the area under the curve remaining equal to unity, whereas the value of the function tends to zero at every point except $t = 0$, where it tends to infinity. That is, we may view the delta function as *the limiting form of a pulse of unit area as the duration of the pulse approaches zero*. It is immaterial what sort of pulse shape is used.

In a rigorous sense, the Dirac delta function belongs to a special class of functions known as *generalized functions* or *distributions*. Indeed, in some situations its use requires that we exercise considerable care. Nevertheless, one beautiful aspect of the Dirac delta function lies precisely in the fact that a rather intuitive treatment of the function along the lines described herein often gives the correct answer.

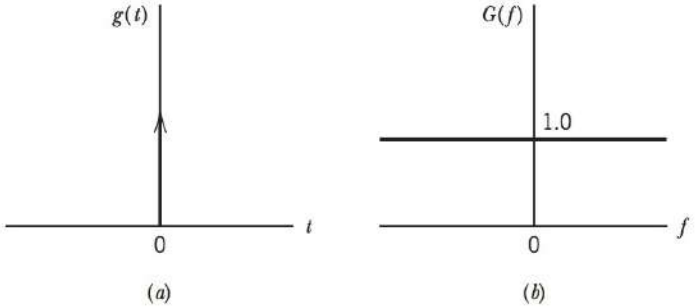


FIGURE 2.12 (*a*) The Dirac delta function $\delta(t)$. (*b*) Spectrum of $\delta(t)$.

EXAMPLE 2.10 The Delta Function as a Limiting Form of the Gaussian Pulse

Consider a Gaussian pulse of unit area, defined by

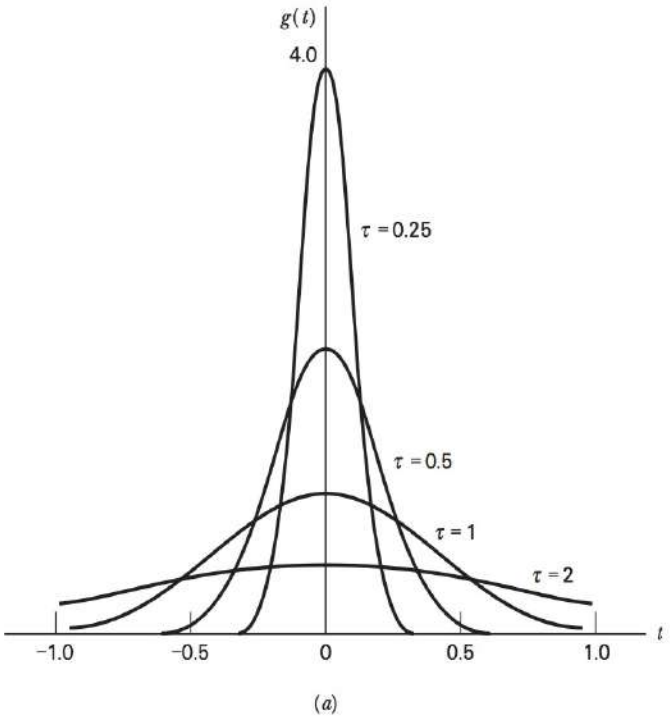
$$g(t) = \frac{1}{\tau}\exp\left(-\frac{\pi t^2}{\tau^2}\right) \tag{2.66}$$

where τ is a variable parameter. The Gaussian function $g(t)$ has two useful properties: (1) its derivatives are all continuous, and (2) it dies away more rapidly than any power of t. The delta function $\delta(t)$ is obtained by taking the limit $\tau \to 0$. The Gaussian pulse then becomes infinitely narrow in duration and infinitely large in amplitude, and yet its area remains finite and fixed at unity. Figure 2.13*a* illustrates the sequence of such pulses as the parameter τ varies.

The Gaussian pulse $g(t)$, defined here, is the same as the normalized Gaussian pulse $\exp(-\pi t^2)$ derived in Example 2.6, except for the fact that it is now expanded in time by the factor τ and compressed in amplitude by the same factor. Therefore, applying the linearity and time-scaling properties of the Fourier transform to the transform pair of Eq. (2.38), we find that the Fourier transform of the Gaussian pulse $g(t)$ defined in Eq. (2.66) is also Gaussian, as shown by

$$G(f) = \exp(-\pi\tau^2 f^2)$$

Figure 2.13*b* illustrates the effect of varying the parameter τ on the spectrum of the Gaussian pulse $g(t)$. Thus, putting $\tau = 0$, we find, as expected, that the Fourier transform of the delta function is unity.



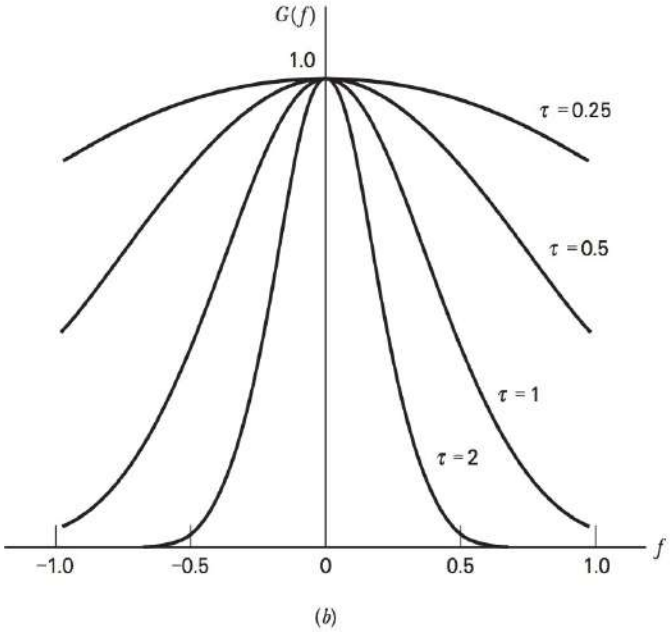
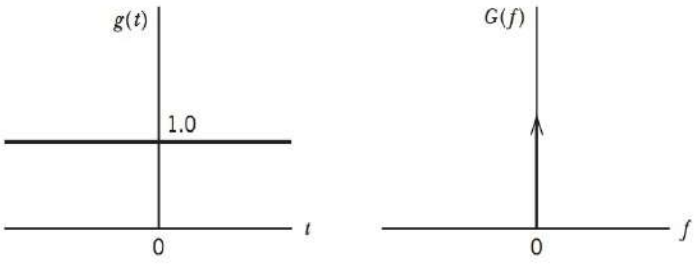


FIGURE 2.13 (*a*) Gaussian pulse of varying duration. (*b*) Corresponding spectrum.

APPLICATIONS OF THE DELTA FUNCTION

1. dc Signal. By applying the duality property to the Fourier-transform pair of Eq. (2.65) and noting that the delta function is an even function, we obtain

$$1 \rightleftharpoons \delta(f) \tag{2.67}$$



(a) (b)

FIGURE 2.14 (*a*) dc signal. (*b*) Spectrum.

Equation (2.67) states that a *dc signal* is transformed in the frequency domain into a delta function $\delta(f)$ occurring at zero frequency, as shown in Figure 2.14. Of course, this result is intuitively satisfying.

Invoking the definition of Fourier transform, we readily deduce from Eq. (2.67) the useful relation

$$\int_{-\infty}^{\infty} \exp(-j2\pi ft)\, dt = \delta(f)$$

Recognizing that the delta function $\delta(f)$ is real valued, we may simplify this relation as follows:

$$\int_{-\infty}^{\infty} \cos(2\pi ft)\, dt = \delta(f) \tag{2.68}$$

which provides yet another definition for the delta function, albeit in the frequency domain.

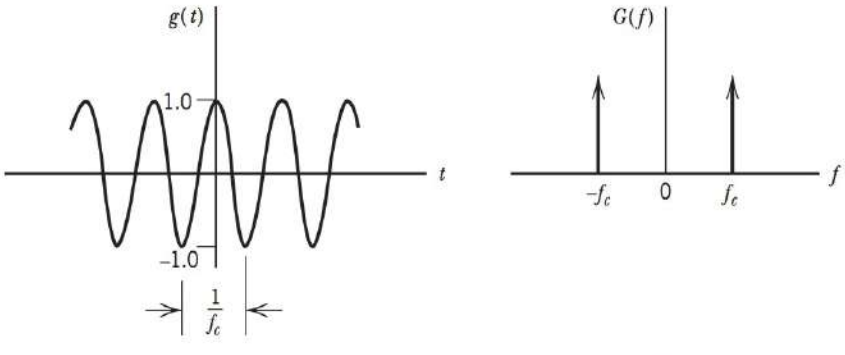
2. Complex Exponential Function. Next, by applying the frequency-shifting property to Eq. (2.67), we obtain the Fourier-transform pair

$$\exp(j2\pi f_c t) \rightleftharpoons \delta(f - f_c) \tag{2.69}$$

for a complex exponential function of frequency f_c. Equation (2.69) states that the complex exponential function $\exp(j2\pi f_c t)$ is transformed in the frequency domain into a delta function $\delta(f - f_c)$ occurring at $f = f_c$.

3. Sinusoidal Functions. Consider next the problem of evaluating the Fourier transform of the cosine function $\cos(2\pi f_c t)$. We first use Euler's formula to write

$$\cos(2\pi f_c t) = \frac{1}{2}\left[\exp(j2\pi f_c t) + \exp(-j2\pi f_c t)\right] \tag{2.70}$$



(a) (b)

FIGURE 2.15 (*a*) Cosine function. (*b*) Spectrum.

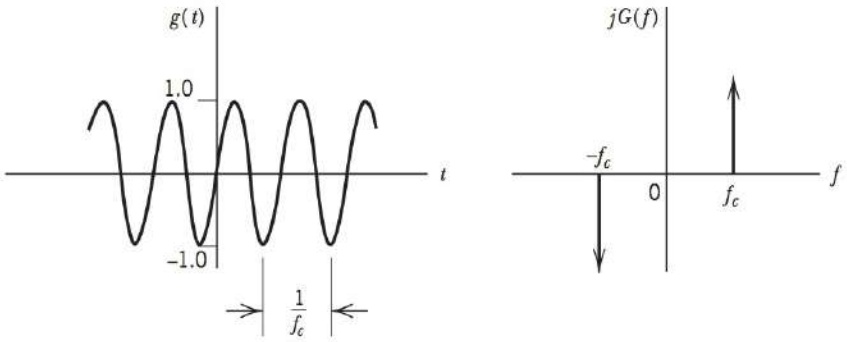


FIGURE 2.16 (*a*) Sine function. (*b*) Spectrum.

Therefore, using Eq. (2.69), we find that the cosine function $\cos(2\pi f_c t)$ is represented by the Fourier-transform pair

$$\cos(2\pi f_c t) \rightleftharpoons \frac{1}{2}[\delta(f - f_c) + \delta(f + f_c)] \tag{2.71}$$

In other words, the spectrum of the cosine function $\cos(2\pi f_c t)$ consists of a pair of delta functions occurring at $f = \pm f_c$, each of which is weighted by the factor 1/2, as shown in Figure 2.15.

Similarly, we may show that the sine function $\sin(2\pi f_c t)$ is represented by the Fourier-transform pair

$$\sin(2\pi f_c t) \rightleftharpoons \frac{1}{2j}[\delta(f - f_c) - \delta(f + f_c)] \tag{2.72}$$

which is illustrated in Figure 2.16.

4. Signum Function. The *signum function* $\text{sgn}(t)$ equals $+1$ for positive time and -1 for negative time, as shown by the solid curve in Figure 2.17*a*. The signum function was defined previously in Eq. (2.18); this definition is reproduced here for convenience of presentation:

$$\text{sgn}(t) = \begin{cases} +1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases}$$

The signum function does not satisfy the Dirichlet conditions and therefore, strictly speaking, it does not have a Fourier transform. However, we may define a Fourier transform for the signum function by viewing it as the limiting form of the antisymmetric double-exponential pulse

$$g(t) = \begin{cases} \exp(-at), & t > 0 \\ 0, & t = 0 \\ -\exp(at), & t < 0 \end{cases} \tag{2.73}$$

as the parameter a approaches zero. The signal $g(t)$, shown as the dashed curve in Figure 2.17*a*, does satisfy the Dirichlet conditions. Its Fourier transform was derived in Example 3; the result is given by [see Eq. (2.19)]:

$$G(f) = \frac{-j4\pi f}{a^2 + (2\pi f)^2}$$

The amplitude spectrum $|G(f)|$ is shown as the dashed curve in Figure 2.17*b*. In the limit as a approaches zero we have

$$\begin{aligned} F[\text{sgn}(t)] &= \lim_{a \to 0} \frac{-4j\pi f}{a^2 + (2\pi f)^2} \\ &= \frac{1}{j\pi f} \end{aligned}$$

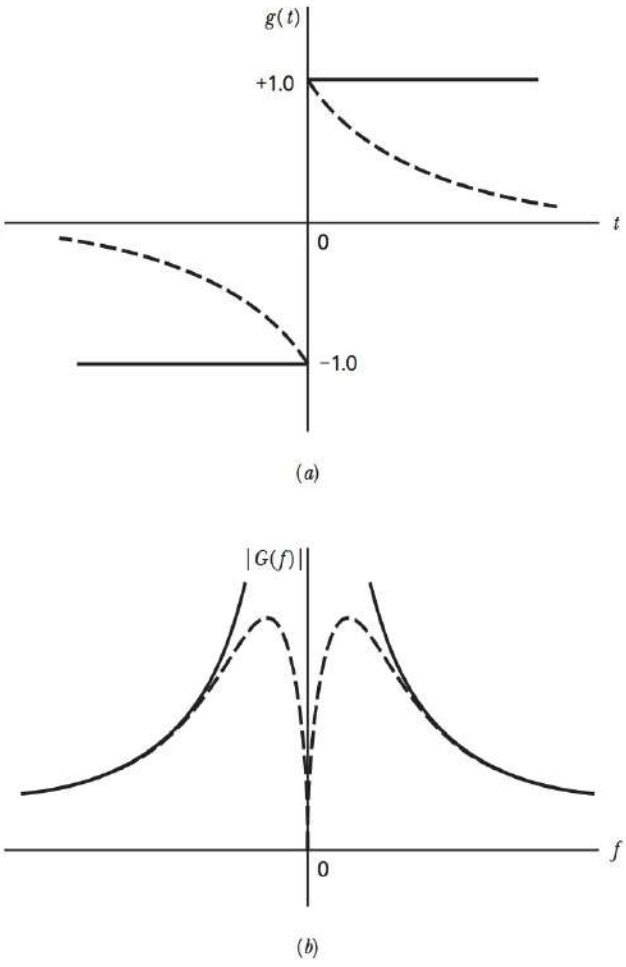


FIGURE 2.17 (*a*) Signum function (continuous curve), and double-exponential pulse (dashed curve). (*b*) Amplitude spectrum of signum function (continuous curve), and that of double-exponential pulse (dashed curve).

That is,

$$\operatorname{sgn}(t) \rightleftharpoons \frac{1}{j\pi f} \tag{2.74}$$

The amplitude spectrum of the signum function is shown as the continuous curve in Figure 2.17*b*. Here we see that for small a, the approximation is very good except near the origin on the frequency axis. At the origin, the spectrum of the approximating function $g(t)$ is zero for $a > 0$, whereas the spectrum of the signum function goes to infinity. It should also be noted that, although the Fourier-transform pair of Eq. (2.74) does not involve a delta function, the Fourier transform $1/j\pi f$ can be obtained from the signum function $\operatorname{sgn}(t)$ only if it is given a special meaning that implies the use of the delta function.

5. Unit Step Function. The *unit step function* $u(t)$ equals +1 for positive time and zero for negative time. Previously defined in Eq. (2.11), it is reproduced here for convenience:

$$u(t) = \begin{cases} 1, & t > 0 \\ \dfrac{1}{2}, & t = 0 \\ 0, & t < 0 \end{cases}$$

The waveform of the unit step function is shown in Figure 2.18*a*. From this defining equation and that of the signum function, we see that the unit step function and signum function are related by

$$u(t) = \frac{1}{2}\left[\operatorname{sgn}(t) + 1\right] \tag{2.75}$$

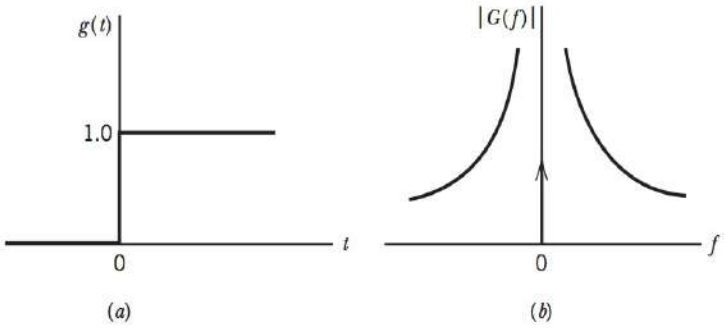


FIGURE 2.18 (*a*) The unit step function. (*b*) Amplitude spectrum.

Hence, using the linearity property of the Fourier transform and the Fourier-transform pairs of Eqs. (2.67) and (2.75), we find that the unit step function is represented by the Fourier-transform pair

$$u(t) \rightleftharpoons \frac{1}{j2\pi f} + \frac{1}{2}\delta(f) \tag{2.76}$$

This means that the spectrum of the unit step function contains a delta function weighted by a factor of 1/2 and occurring at zero frequency, as shown in Figure 2.18*b*.

6. Integration in the Time Domain (Revisited). The relation of Eq. (2.39) describes the effect of integration on the Fourier transform of a signal $g(t)$, assuming that $G(0)$ is zero. We now consider the more general case, with no such assumption made.

Let

$$y(t) = \int_{-\infty}^{t} g(\tau)\, d\tau \tag{2.77}$$

The integrated signal $y(t)$ can be viewed as the convolution of the original signal $g(t)$ and the unit step function $u(t)$, as shown by

$$y(t) = \int_{-\infty}^{\infty} g(\tau)u(t-\tau)\, d\tau$$

where the time-shifted unit step function $u(t-\tau)$ is defined by

$$u(t-\tau) = \begin{cases} 1, & \tau < t \\ \frac{1}{2}, & \tau = t \\ 0, & \tau > t \end{cases}$$

Recognizing that convolution in the time domain is transformed into multiplication in the frequency domain in accordance with Property 12, and using the Fourier-transform pair of Eq. (2.76) for the unit step function $u(t)$, we find that the Fourier transform of $y(t)$ is

$$Y(f) = G(f)\left[\frac{1}{j2\pi f} + \frac{1}{2}\delta(f)\right] \tag{2.78}$$

where $G(f)$ is the Fourier transform of $g(t)$. Since

$$G(f)\delta(f) = G(0)\delta(f)$$

we may rewrite Eq. (2.77) in the equivalent form:

$$Y(f) = \frac{1}{j2\pi f}G(f) + \frac{1}{2}G(0)\delta(f)$$

In general, the effect of integrating the signal $g(t)$ is therefore described by the Fourier-transform pair

$$\int_{-\infty}^{t} g(\tau)\,d\tau \rightleftharpoons \frac{1}{j2\pi f}\,G(f) + \frac{1}{2}\,G(0)\delta(f) \tag{2.79}$$

This is the desired result, which includes Eq. (2.39) as a special case (i.e., $G(0) = 0$).

2.6 FOURIER TRANSFORMS OF PERIODIC SIGNALS

It is well known that by using the Fourier series, a periodic signal can be represented as a sum of complex exponentials. Also, in a limiting sense, Fourier transforms can be defined for complex exponentials. Therefore, it seems reasonable that a periodic signal can be represented in terms of a Fourier transform, provided that this transform is permitted to include delta functions.

Consider then a periodic signal $g_{T_0}(t)$ of *period* T_0. We can represent $g_{T_0}(t)$ in terms of the *complex exponential Fourier series*:

$$g_{T_0}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_0 t) \tag{2.80}$$

where c_n is the *complex Fourier coefficient* defined by

$$c_n = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} g_{T_0}(t)\exp(-j2\pi n f_0 t)\,dt \tag{2.81}$$

and f_0 is the *fundamental frequency* defined as the reciprocal of the period T_0; that is,

$$f_0 = \frac{1}{T_0} \tag{2.82}$$

Let $g(t)$ be a pulselike function, which equals $g_{T_0}(t)$ over one period and is zero elsewhere; that is,

$$g(t) = \begin{cases} g_{T_0}(t), & -\dfrac{T_0}{2} \le t \le \dfrac{T_0}{2} \\ 0, & \text{elsewhere} \end{cases} \tag{2.83}$$

The periodic signal $g_{T_0}(t)$ may now be expressed in terms of the function $g(t)$ as an infinite summation, as shown by

$$g_{T_0}(t) = \sum_{m=-\infty}^{\infty} g(t - mT_0) \tag{2.84}$$

Based on this representation, we may view $g(t)$ as a *generating function*, which generates the periodic signal $g_{T_0}(t)$.

The function $g(t)$ is Fourier transformable. Accordingly, we may rewrite the formula for the complex Fourier coefficient as follows:

$$\begin{aligned} c_n &= f_0 \int_{-\infty}^{\infty} g(t)\exp(-j2\pi n f_0 t)\,dt \\ &= f_0 G(n f_0) \end{aligned} \tag{2.85}$$

where $G(nf_0)$ is the Fourier transform of $g(t)$, evaluated at the frequency nf_0. We may thus rewrite the formula for the reconstruction of the periodic signal $g_{T_0}(t)$ as

$$g_{T_0}(t) = f_0 \sum_{n=-\infty}^{\infty} G(n f_0)\exp(j2\pi n f_0 t) \tag{2.86}$$

or, equivalently, in light of Eq. (2.84)

$$\sum_{m=-\infty}^{\infty} g(t - mT_0) = f_0 \sum_{n=-\infty}^{\infty} G(nf_0)\exp(j2\pi nf_0 t) \tag{2.87}$$

Equation (2.87) is one form of *Poisson's sum formula*.

Finally, using Eq. (2.69), which defines the Fourier transform of a complex exponential function, and Eq. (2.87), we deduce the following Fourier-transform pair for a periodic signal $g_{T_0}(t)$ with a generating function $g(t)$ and period T_0:

$$\sum_{m=-\infty}^{\infty} g(t - mT_0) \rightleftharpoons f_0 \sum_{n=-\infty}^{\infty} G(nf_0)\delta(f - nf_0) \tag{2.88}$$

where f_0 is the fundamental frequency. This relation simply states that the Fourier transform of a periodic signal consists of delta functions occurring at integer multiples of the fundamental frequency $f_0 = 1/T_0$, including the origin, and that each delta function is weighted by a factor equal to the corresponding value of $G(nf_0)$. Indeed, this relation merely provides a method to display the frequency content of a periodic signal $g_{T_0}(t)$.

It is of interest to observe that the function $g(t)$, constituting one period of the periodic signal $g_{T_0}(t)$, has a continuous spectrum defined by $G(f)$. On the other hand, the periodic signal $g_{T_0}(t)$ itself has a discrete spectrum. We conclude, therefore, that *periodicity in the time domain has the effect of changing the frequency-domain description or spectrum of the signal into a discrete form defined at integer multiples of the fundamental frequency*.

EXAMPLE 2.11 Ideal Sampling Function

An *ideal sampling function*, or *Dirac comb*, consists of an infinite sequence of uniformly spaced delta functions, as shown in Figure 2.19*a*. We denote this waveform by

$$\delta_{T_0}(t) = \sum_{m=-\infty}^{\infty} \delta(t - mT_0) \tag{2.89}$$

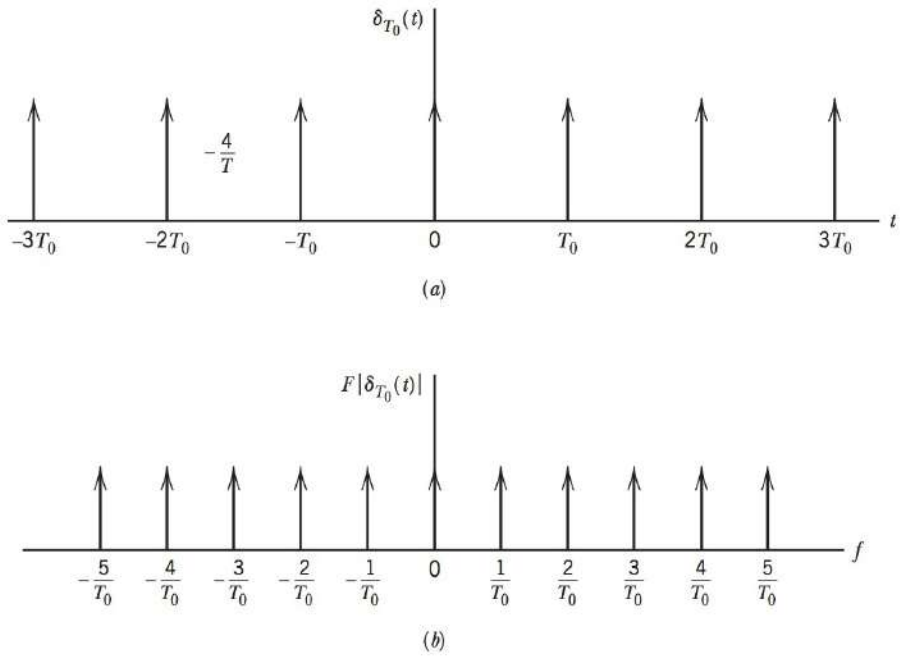


FIGURE 2.19 (*a*) Dirac comb. (*b*) Spectrum.

We observe that the generating function $g(t)$ for the ideal sampling function $\delta_{T_0}(t)$ consists simply of the delta function $\delta(t)$. We therefore have $G(f) = 1$, and

$$G(nf_0) = 1 \qquad \text{for all } n$$

Thus the use of Eq. (2.88) yields the result

$$\sum_{m=-\infty}^{\infty} \delta(t - mT_0) \rightleftharpoons f_0 \sum_{n=-\infty}^{\infty} \delta(f - nf_0) \tag{2.90}$$

Equation (2.90) states that the Fourier transform of a periodic train of delta functions, spaced T_0 seconds apart, consists of another set of delta functions weighted by the factor $f_0 = 1/T_0$ and regularly spaced f_0 Hz apart along the frequency axis as in Figure 2.19*b*. In the special case of $T_0 = 1$, a periodic train of delta functions is, like a Gaussian pulse, its own transform.

We also deduce from Poisson's sum formula, Eq. (2.87), the following useful relation:

$$\sum_{m=-\infty}^{\infty} \delta(t - mT_0) = f_0 \sum_{n=-\infty}^{\infty} \exp(j2\pi n f_0 t) \tag{2.91}$$

The dual of this relation is as follows:

$$\sum_{m=-\infty}^{\infty} \exp(j2\pi m f T_0) = f_0 \sum_{n=-\infty}^{\infty} \delta(f - nf_0) \tag{2.92}$$

2.7 TRANSMISSION OF SIGNALS THROUGH LINEAR SYSTEMS

With the Fourier transform theory presented in the previous sections at our disposal, we are ready to turn our attention to the study of a special class of systems known to be linear. A *system* refers to any physical device that produces an output signal in response to an input signal. It is customary to refer to the input signal as the *excitation* and to the output signal as the *response*. In a *linear* system, the *principle of superposition* holds; that is, the response of a linear system to a number of excitations applied simultaneously is equal to the sum of the responses of the system when each excitation is applied individually. Important examples of linear systems include *filters* and *communication channels* operating in their linear region. A filter refers to a frequency-selective device that is used to limit the spectrum of a signal to some band of frequencies. A channel refers to a transmission medium that connects the transmitter and receiver of a communication system. We wish to evaluate the effects of transmitting signals through linear filters and communication channels. This evaluation may be carried out in two ways, depending on the description adopted for the filter or channel. That is, we may use time-domain or frequency-domain ideas, as described below.

TIME RESPONSE

In the time domain, a linear system is described in terms of its *impulse response, which is defined as the response of the system (with zero initial conditions) to a unit impulse or delta function $\delta(t)$ applied to the input of the system*. If the system is *time invariant*, then

the shape of the impulse response is the same no matter when the unit impulse is applied to the system. Thus, assuming that the unit impulse or delta function is applied at time $t = 0$, we may denote the impulse response of a linear time-invariant system by $h(t)$. Let this system be subjected to an arbitrary excitation $x(t)$, as in Figure 2.20*a*. To determine the response $y(t)$ of the system, we begin by first approximating $x(t)$ by a staircase function composed of narrow rectangular pulses, each of duration $\Delta\tau$, as shown in Figure 2.20*b*. Clearly the approximation becomes better for smaller $\Delta\tau$. As $\Delta\tau$ approaches zero, each pulse approaches, in the limit, a delta function weighted by a factor equal to the height of the pulse times $\Delta\tau$. Consider a typical pulse, shown shaded in Figure 2.20*b*, which occurs at $t = \tau$. This pulse has an area equal to $x(\tau)\Delta\tau$. By definition, the response of the system to a unit impulse or delta function $\delta(t)$, occurring at $t = 0$, is $h(t)$. It follows, therefore, that the response of the system to a delta function, weighted by the factor $x(\tau)\Delta\tau$ and occurring at $t = \tau$, must be $x(\tau)h(t - \tau)\Delta\tau$. To find the total response $y(t)$ at some time t, we apply the principle of superposition. Thus, summing the various infinitesimal responses due to the various input pulses, we obtain in the limit, as $\Delta\tau$ approaches zero,

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\,d\tau \tag{2.93}$$

This relation is called the *convolution integral.*

In Eq. (2.93), three different time scales are involved: *excitation time* τ, *response time* t, and *system-memory time* $t - \tau$. This relation is the basis of time-domain analysis of linear time-invariant systems. It states that the present value of the response of a linear time-invariant system is a weighted integral over the past history of the input signal, weighted according to the impulse response of the system. Thus, the impulse response acts as a *memory function* for the system.

In Eq. (2.93), the excitation $x(t)$ is convolved with the impulse response $h(t)$ to produce the response $y(t)$. Since convolution is commutative, it follows that we may also write

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\,d\tau \tag{2.94}$$

where $h(t)$ is convolved with $x(t)$.

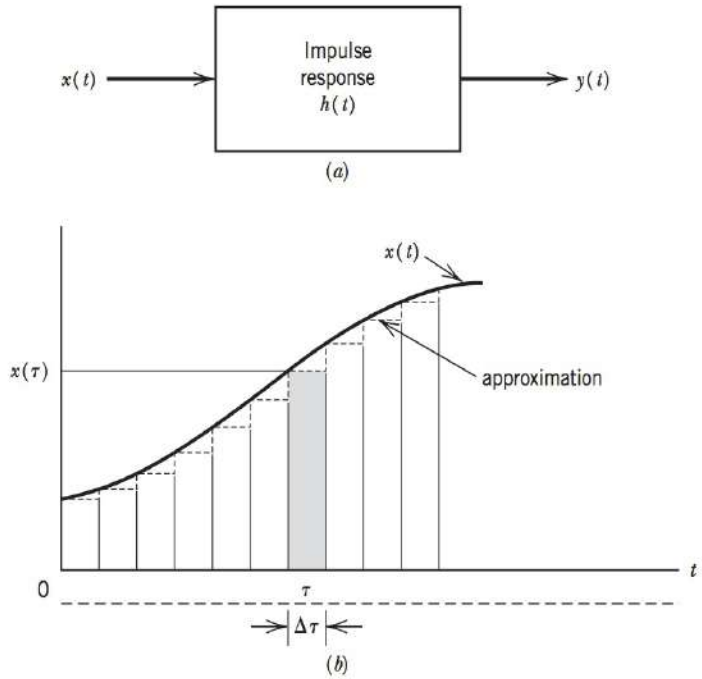


FIGURE 2.20 (*a*) Linear system. (*b*) Approximation of input $x(t)$.

EXAMPLE 2.12 Tapped-Delay-Line Filter

Consider a linear time-invariant filter with impulse response $h(t)$. We make the following two assumptions:

1. The impulse response $h(t) = 0$ for $t < 0$.
2. The impulse response of the filter is of some finite duration T_f, so that we may write $h(t) = 0$ for $t \geq T_f$.

Then we may express the filter output $y(t)$ produced in response to the input $x(t)$ as follows:

$$y(t) = \int_0^{T_f} h(\tau)x(t-\tau)\,d\tau \tag{2.95}$$

Let the input $x(t)$, impulse response $h(t)$, and output $y(t)$ be *uniformly sampled* at the rate $1/\Delta\tau$ samples per second, so that we may put

$$t = n\Delta\tau$$

and

$$\tau = k\Delta\tau$$

where k and n are integers, and $\Delta\tau$ is the *sampling period*. Assuming that $\Delta\tau$ is small enough for the product $h(\tau)x(t-\tau)$ to remain essentially constant for $k\Delta\tau \leq \tau \leq (k+1)\,\Delta\tau$ for all values of k and t of interest, we may approximate Eq. (2.95) by the *convolution sum*:

$$y(n\Delta\tau) = \sum_{k=0}^{N-1} h(k\Delta\tau)x(n\Delta\tau - k\Delta\tau)\,\Delta\tau \tag{2.96}$$

where $N\Delta\tau = T_f$. Define

$$w_k = h(k\Delta\tau)\,\Delta\tau$$

We may then rewrite Eq. (2.96) as

$$y(n\Delta\tau) = \sum_{k=0}^{N-1} w_k x(n\Delta\tau - k\Delta\tau) \tag{2.97}$$

Equation (2.97) may be realized using the circuit shown in Figure 2.21, which consists of a set of *delay elements* (each producing a delay of

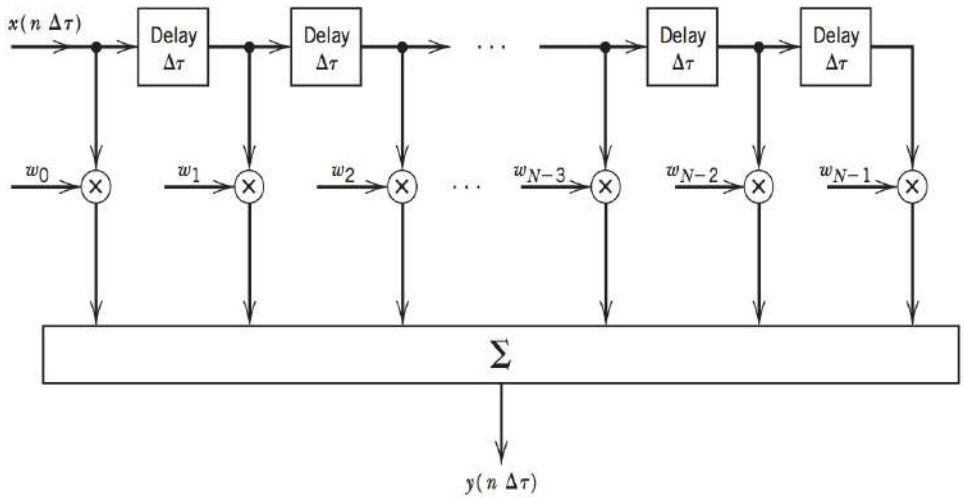


FIGURE 2.21 Tapped-delay-line filter.

$\Delta\tau$ seconds), a set of *multipliers* connected to the *delay-line taps*, a corresponding set of *weights* applied to the multipliers, and a *summer* for adding the multiplier outputs. This circuit is known as a *tapped-delay-line filter* or *transversal filter*. Note that in Figure 2.21 the tap-spacing or basic increment of delay is equal to the sampling period of the input sequence $\{x(n\Delta\tau)\}$.

CAUSALITY AND STABILITY

A system is said to be *causal* if it does not respond before the excitation is applied. For a linear time-invariant system to be causal, it is clear that the impulse response $h(t)$ must vanish for negative time. That is, the necessary and sufficient condition for causality is

$$h(t) = 0, \qquad t < 0 \tag{2.98}$$

Clearly, for a system operating in *real time* to be physically realizable, it must be causal. However, there are many applications in which the signal to be processed is available in stored form; in these situations the system can be noncausal and yet physically realizable.

The system is said to be *stable* if the output signal is bounded for all bounded input signals; we refer to this as the *bounded input-bounded output* (BIBO) *stability criterion*, which is well suited for the analysis of linear time-invariant systems. Let the input signal $x(t)$ be bounded, as shown by

$$|x(t)| \leq M \tag{2.99}$$

where M is a positive real finite number. Substituting Eq. (2.99) in (2.94), we get

$$|y(t)| \leq M \int_{-\infty}^{\infty} |h(\tau)|\, d\tau$$

It follows, therefore, that for a linear time-invariant system to be stable, the impulse response $h(t)$ must be absolutely integrable. That is, the necessary and sufficient condition for BIBO stability is

$$\int_{-\infty}^{\infty} |h(t)|\, dt < \infty \tag{2.100}$$

FREQUENCY RESPONSE

Consider a linear time-invariant system of impulse response $h(t)$ driven by a complex exponential input of unit amplitude and frequency f, that is,

$$x(t) = \exp(j2\pi ft) \tag{2.101}$$

Using Eq. (2.101) in (2.94), the response of the system is obtained as

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)\exp[j2\pi f(t-\tau)]\, d\tau \\ &= \exp(j2\pi ft) \int_{-\infty}^{\infty} h(\tau)\exp(-j2\pi f\tau)\, d\tau \end{aligned} \tag{2.102}$$

Define the *transfer function* of the system as the Fourier transform of its impulse response, as shown by

$$H(f) = \int_{-\infty}^{\infty} h(t)\exp(-j2\pi ft)\, dt \tag{2.103}$$

The integral in the last line of Eq. (2.102) is the same as that of Eq. (2.103), except that τ is used in place of t. Hence, we may rewrite Eq. (2.102) in the form

$$y(t) = H(f)\exp(j2\pi ft) \tag{2.104}$$

The response of a linear time-invariant system to a complex exponential function of frequency f is, therefore, the same complex exponential function multiplied by a constant coefficient $H(f)$.

An alternative definition of the transfer function may be deduced by dividing Eq. (2.104) by (2.101) to obtain

$$H(f) = \left.\frac{y(t)}{x(t)}\right|_{x(t)=\exp(j2\pi ft)} \tag{2.105}$$

Consider next an arbitrary signal $x(t)$ applied to the system. The signal $x(t)$ may be expressed in terms of its inverse Fourier transform as

$$x(t) = \int_{-\infty}^{\infty} X(f)\exp(j2\pi ft)\, df \tag{2.106}$$

or, equivalently, in the limiting form

$$x(t) = \lim_{\substack{\Delta f \to 0 \\ f = k\Delta f}} \sum_{k=-\infty}^{\infty} X(f)\exp(j2\pi ft)\Delta f \tag{2.107}$$

That is, the input signal $x(t)$ may be viewed as a superposition of complex exponentials of incremental amplitude. Because the system is linear, the response to this superposition of complex exponential inputs is

$$\begin{aligned} y(t) &= \lim_{\substack{\Delta f \to 0 \\ f = k\Delta f}} \sum_{k=-\infty}^{\infty} H(f)X(f)\exp(j2\pi ft)\Delta f \\ &= \int_{-\infty}^{\infty} H(f)X(f)\exp(j2\pi ft)\, df \end{aligned} \tag{2.108}$$

The Fourier transform of the output signal $y(t)$ is therefore

$$Y(f) = H(f)X(f) \tag{2.109}$$

A linear time-invariant system may thus be described quite simply in the frequency domain by noting that *the Fourier transform of the output is equal to the product of the transfer function of the system and the Fourier transform of the input.*

The result of Eq. (2.109) may, of course, be deduced directly by recognizing that the response $y(t)$ of a linear time-invariant system of impulse response $h(t)$ to an arbitrary input $x(t)$ is obtained by convolution $x(t)$ with $h(t)$, or vice versa, and by the fact that the convolution of a pair of time functions is transformed into the multiplication of their Fourier transforms. The derivation above is presented primarily to develop an understanding of why the Fourier representation of a time function as a superposition of complex exponentials is so useful in analyzing the behavior of linear time-invariant systems.

The transfer function $H(f)$ is a characteristic property of a linear time-invariant system. It is, in general, a complex quantity, so that we may express it in the form

$$H(f) = |H(f)|\exp[j\beta(f)] \tag{2.110}$$

where $|H(f)|$ is called the *amplitude response*, and $\beta(f)$ the *phase* or *phase response*. In the special case of a linear system with a real-valued impulse response $h(t)$, the transfer

function $H(f)$ exhibits conjugate symmetry, which means that

$$|H(f)| = |H(-f)|$$

and

$$\beta(f) = -\beta(-f)$$

That is, the amplitude response $|H(f)|$ of a linear system with real-valued impulse response is an even function of frequency, whereas the phase $\beta(f)$ is an odd function of frequency.

In some applications it is preferable to work with the logarithm of $H(f)$, expressed in polar form, rather than with $H(f)$ itself. Define

$$\ln H(f) = \alpha(f) + j\beta(f) \tag{2.111}$$

where

$$\alpha(f) = \ln|H(f)| \tag{2.112}$$

The function $\alpha(f)$ is called the *gain* of the system. It is measured in *nepers*, whereas $\beta(f)$ is measured in *radians*. Equation (2.111) indicates that the gain $\alpha(f)$ and phase $\beta(f)$ are the real and imaginary parts of the (natural) logarithm of the transfer function $H(f)$, respectively. The gain may also be expressed in *decibels* (dB) by using the definition

$$\alpha'(f) = 20\log_{10}|H(f)| \tag{2.113}$$

The two gain functions $\alpha(f)$ and $\alpha'(f)$ are related by

$$\alpha'(f) = 8.69\alpha(f) \tag{2.114}$$

That is, 1 neper is equal to 8.69 dB.

As a means of specifying the constancy of the amplitude response $|H(f)|$ or gain $\alpha(f)$ of a system, we use a parameter called the *bandwidth* of the system. The definitions of bandwidth described for signals in Section 2.5 also apply to systems. In the case of a *low-pass system*, the 3-dB bandwidth is defined as the frequency at which the amplitude response $|H(f)|$ is $1/\sqrt{2}$ times its value at zero frequency or, equivalently, the frequency at which the gain $\alpha'(f)$ drops by 3 dB below its value at zero frequency, as illustrated in Figure 2.22a. In the case of a band-pass system, the bandwidth is defined as the range of frequencies over which the amplitude response $|H(f)|$ remains within $1/\sqrt{2}$ times its value at the mid-band frequency, as illustrated in Figure 2.22*b*.

PALEY–WIENER CRITERION

A necessary and sufficient condition for a function $\alpha(f)$ to be the gain of a causal filter is the convergence of the integral

$$\int_{-\infty}^{\infty} \frac{|\alpha(f)|}{1+f^2}\,df < \infty \tag{2.115}$$

This condition is known as the *Paley–Wiener criterion*.[5] It states that provided the gain $\alpha(f)$ satisfies the condition of Eq. (2.115), then we may associate with this gain a suitable phase $\beta(f)$, such that the resulting filter has a causal impulse response that is zero for negative time. In other words, the Paley–Wiener criterion is the frequency-domain equivalent of the causality requirement. A system with a realizable gain characteristic may have infinite attenuation for a discrete set of frequencies, but it cannot have infinite attenuation over a band of frequencies; otherwise, the Paley–Wiener criterion is violated.

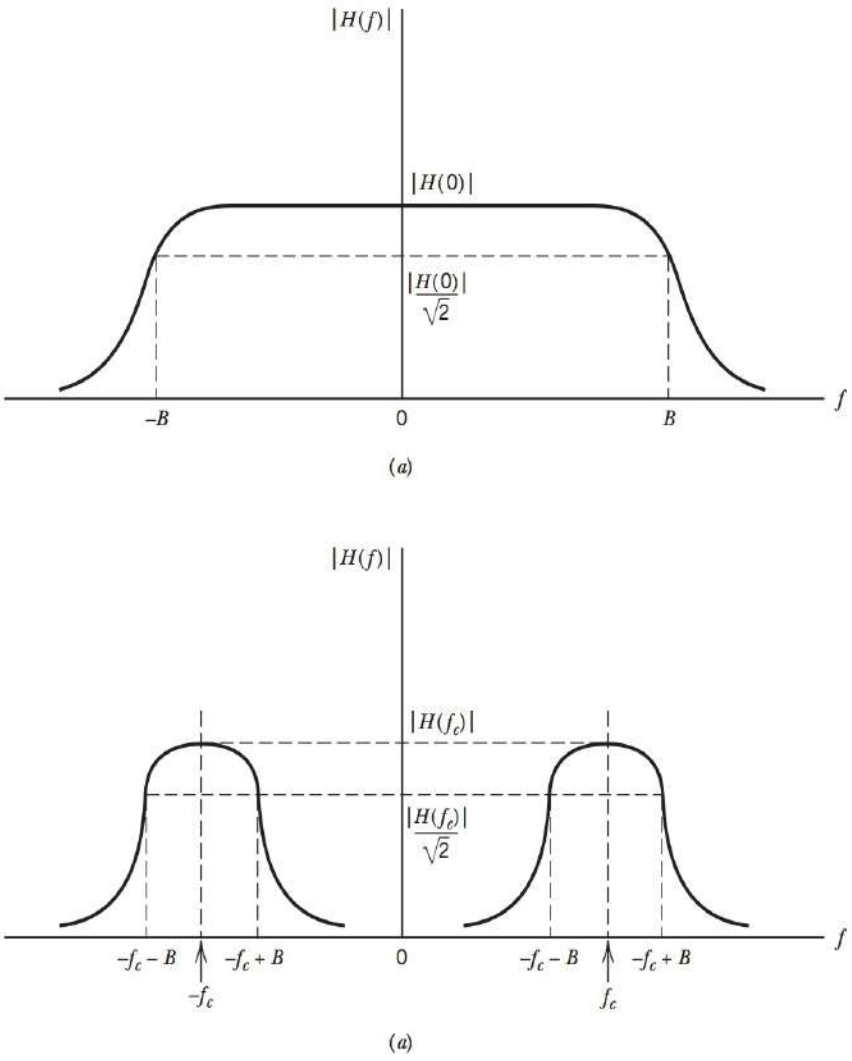


FIGURE 2.22 Illustrating the definition of system bandwidth: (*a*) low-pass system of bandwidth B; (*b*) band-pass system of bandwidth $2B$.

2.8 FILTERS

As previously mentioned, a *filter* is a frequency-selective device that is used to limit the spectrum of a signal to some specified band of frequencies. Its frequency response is characterized by a *passband* and a *stopband*. The frequencies inside the passband are transmitted with little or no distortion, whereas those in the stopband are rejected. The filter may be of the *low-pass*, *high-pass*, *band-pass*, or *band-stop* type, depending on whether it transmits low, high, intermediate, or all but intermediate frequencies, respectively. We have already encountered examples of low-pass and band-pass systems in Figure 2.22.

Filters, in one form or another, represent an important functional block in building communication systems. In this book, we will be largely concerned with the use of low-pass and band-pass filters.

In this section we study the time response of the *ideal low-pass filter*, which transmits, without any distortion, all frequencies inside the passband and completely rejects all frequencies inside the stopband, as illustrated in Figure 2.23. The transfer function of an ideal low-pass filter is therefore defined by

$$H(f) = \begin{cases} \exp(-j2\pi f t_0), & -B \le f \le B \\ 0, & |f| > B \end{cases} \tag{2.116}$$

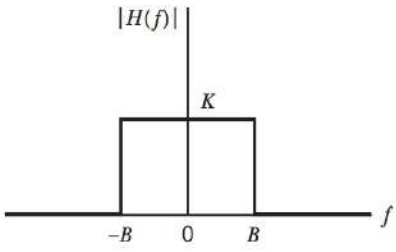


FIGURE 2.23 Frequency response of ideal low-pass filter.

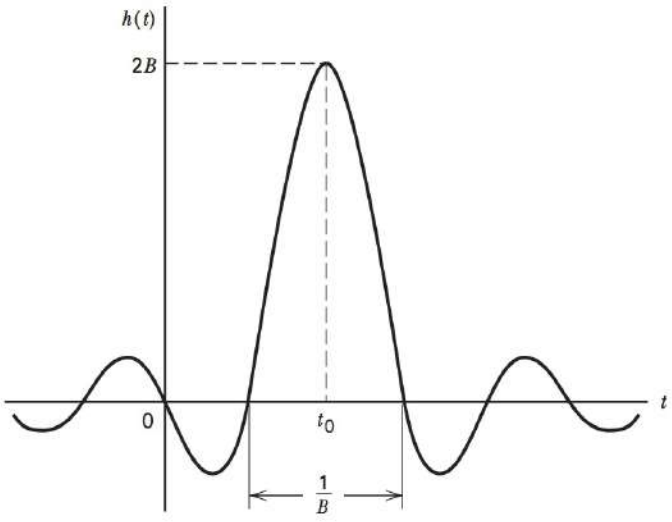


FIGURE 2.24 Impulse response of ideal low-pass filter.

The parameter B defines the bandwidth of the filter. The ideal low-pass filter is, of course, noncausal because it violates the Paley-Wiener criterion. This observation may also be confirmed by examining the impulse response $h(t)$. Thus, by evaluating the inverse Fourier transform of the transfer function of Eq. (2.116), we get

$$h(t) = \int_{-B}^{B} \exp[j2\pi f(t - t_0)]\, df \tag{2.117}$$

where the limits of integration have been reduced to the frequency band inside which $H(f)$ does not vanish. Equation (2.117) is readily integrated, yielding

$$\begin{aligned} h(t) &= \frac{\sin[2\pi B(t - t_0)]}{\pi(t - t_0)} \\ &= 2B\,\mathrm{sinc}[2B(t - t_0)] \end{aligned} \tag{2.118}$$

The impulse response has a peak amplitude of $2B$ centered on time t_0, as shown in Figure 2.24 for $t_0 = 1/B$. The duration of the main lobe of the impulse response is $1/B$, and the build-up time from the zero at the beginning of the main lobe to the peak value is $1/2B$. We see from Figure 2.24 that, for any finite value of t_0, there is some response from the filter before the time $t = 0$ at which the unit impulse is applied to the input, confirming that the ideal low-pass filter is noncausal. Note, however, that we can always make the delay t_0 large enough for the condition

$$|\mathrm{sinc}[2B(t - t_0)]| \ll 1 \qquad \text{for } t < 0$$

to be satisfied. By so doing, we are able to build a causal filter that closely approximates an ideal low-pass filter.

EXAMPLE 2.13 Pulse Response of Ideal Low-Pass Filter

Consider a rectangular pulse $x(t)$ of unit amplitude and duration T, which is applied to an ideal low-pass filter of bandwidth B. The problem is to determine the response $y(t)$ of the filter.

The impulse response $h(t)$ of the filter is defined by Eq. (2.118). The delay t_0 has no effect on the *shape* of the filter response $y(t)$. Without loss of

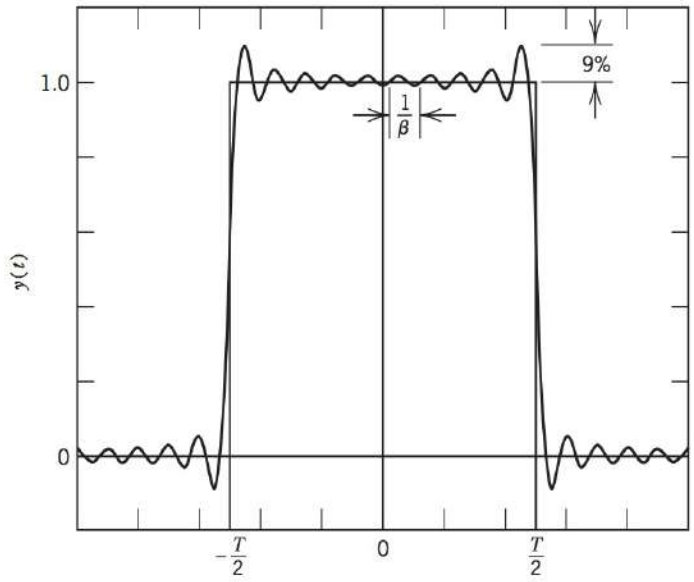


FIGURE 2.25 Filter response for a square pulse.

generality, we may therefore simplify the exposition by setting $t_0 = 0$, in which case the impulse of Eq. (2.118) reduces to

$$h(t) = 2B\,\mathrm{sinc}(2Bt) \tag{2.119}$$

The resulting filter response is thus given by the convolution integral

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \\ &= 2B\int_{-T/2}^{T/2} \frac{\sin[2\pi B(t-\tau)]}{2\pi B(t-\tau)}\,d\tau \end{aligned} \tag{2.120}$$

This integral does not have a closed form solution in terms of elementary functions. For illustration we plot the simulated filter response to a square pulse in Figure 2.25. The response shows an overshoot of approximately 9% (known as the *Gibbs phenomenon*) and an oscillatory pattern that is related to the bandwidth of the filter. This behavior is investigated further in Problem 2.24.

DESIGN OF FILTERS

A filter may be characterized by specifying its impulse response $h(t)$ or, equivalently, its transfer function $H(f)$. However, the application of a filter usually involves the separation of signals on the basis of their spectra (i.e., frequency content). This, in turn, means that the design of filters is usually carried out in the frequency domain. There are two basic steps involved in the design of a filter:

1. The *approximation* of a prescribed frequency response (i.e., amplitude response, phase response, or both) by a realizable transfer function.
2. The *realization* of the approximating transfer function by a physical device.

For an approximating transfer function $H(f)$ to be physically realizable, it must represent a *stable* system. Stability is defined here on the basis of the bounded input–bounded output criterion described in Eq. (2.100) that involves the impulse response $h(t)$. To specify the corresponding condition for stability in terms of the transfer function, the traditional approach is to replace $j2\pi f$ with s and recast the transfer function in terms of s.

The new variable s is permitted to have a real part as well as an imaginary part. Accordingly, we refer to s as the *complex frequency*. Let $H'(s)$ denote the transfer function of the system, defined in the manner described herein. Ordinarily, the approximating transfer function $H'(s)$ is a rational function, which may therefore be expressed in a *factored* form as

$$H'(s) = H(f)|_{j2\pi f = s}$$
$$= K\frac{(s - z_1)(s - z_2)\cdots(s - z_m)}{(s - p_1)(s - p_2)\cdots(s - p_n)}$$

where K is a scaling factor; $z_1, z_2, \ldots, z_m$ are the *zeros* of the transfer function; $p_1, p_2, \ldots, p_n$ are its *poles*. For low-pass and band-pass filters, the number of zeros, m, is less than the number of poles, n. If the system is causal, then the bounded input–bounded output condition for stability of the system is satisfied by restricting all the poles of the transfer function $H'(s)$ to be inside the left half of the s-plane; that is to say,

$$\text{Re}[p_i] < 0 \qquad \text{for all } i$$

DIFFERENT TYPES OF FILTERS

In the case of low-pass filters where the principal requirement is to approximate the ideal amplitude response shown in Figure 2.23, we may mention two popular families of filters: *Butterworth filters* and *Chebyshev filters*, both of which have all their zeros at $s = \infty$. In a Butterworth filter, the poles of the transfer function $H'(s)$ lie on a circle with origin as the center and $2\pi B$ as the radius, where B is the 3-dB bandwidth of the filter. In a Chebyshev filter, on the other hand, the poles lie on an ellipse. In both cases, of course, the poles are confined to the left half of the s-plane.

Examples of Butterworth and Chebyshev filters are shown in Figure 2.26. The Butterworth filter is said to have a *maximally flat* passband response, which is excellent for passing the amplitude spectrum of a signal with little distortion. Chebyshev filters, on the

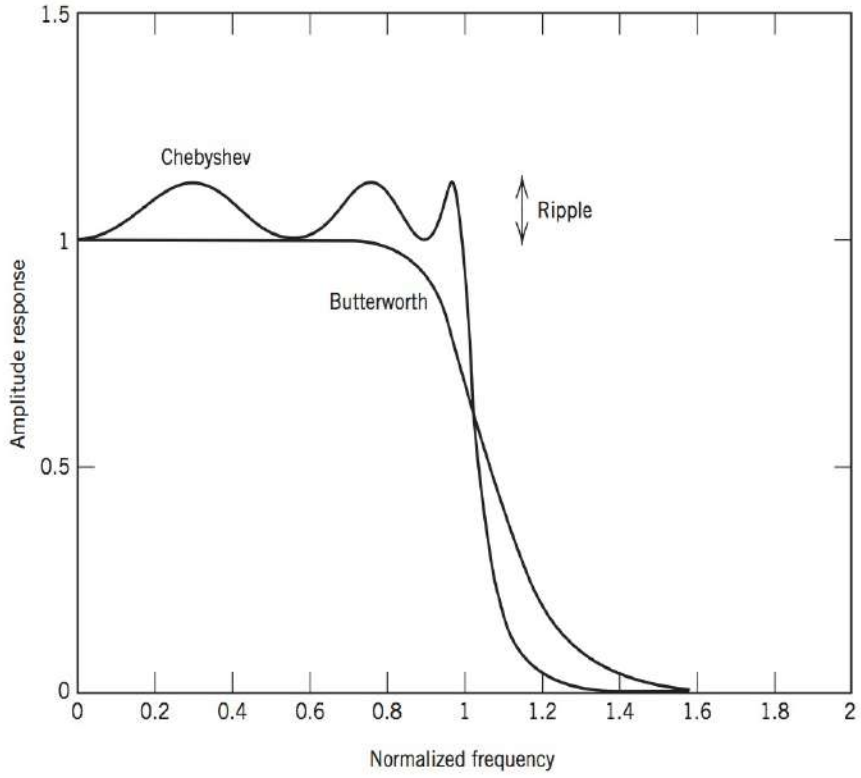


FIGURE 2.26 Comparison of the amplitude response of 6th order Butterworth low-pass filter with that of 6th order Chebyshev filter.

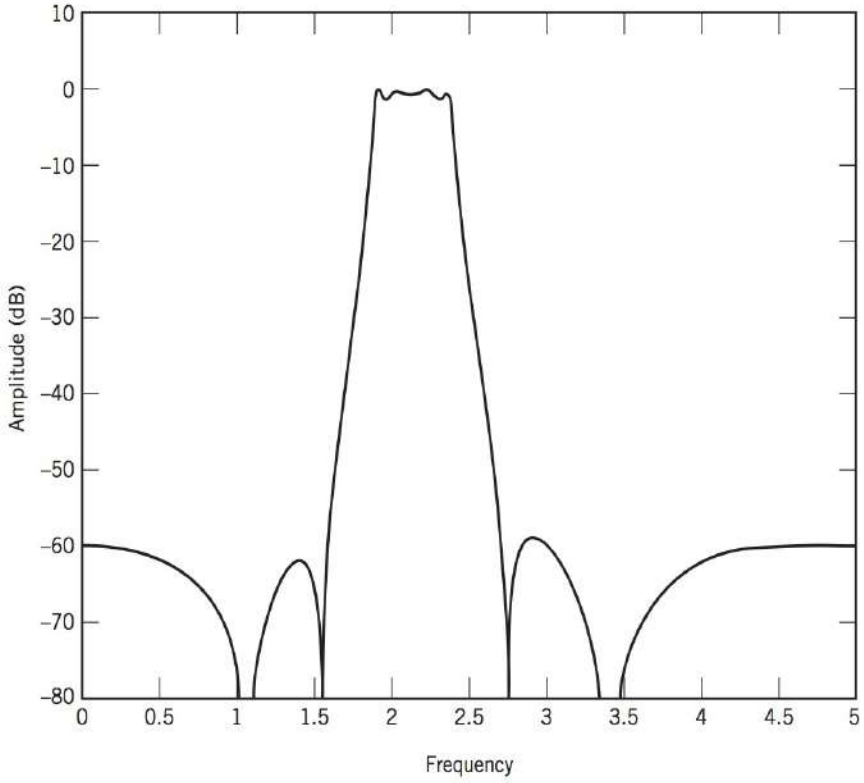


FIGURE 2.27 Amplitude response of 8th order elliptic bandpass filter.

other hand, provide faster *roll-off* than Butterworth filters by allowing *ripple* in the frequency response. There are two types of Chebyshev filters. Type 1 filters have ripple only in the passband; Type 2 filters have ripple only in the stopband and are seldom used. More specifically, Type 1 Chebyshev filters are equi-ripple in the passband and monotonic in the stopband. As the ripple increases in Chebyshev filters, the roll-off becomes sharper (better). The Chebyshev response is an optimal trade-off between these two parameters. When the ripple is set to zero, the Chebyshev filter reduces to a Butterworth filter.

A common alternative to both the Butterworth and Chebyshev filters is the *elliptic filter*, which has ripple in both the passband and the stopband. Elliptic filters provide even faster roll-off for a given number of poles but at the expense of ripple in both the passband and the stopband. The amplitude spectrum of a bandpass elliptic filter is shown in Figure 2.27. To achieve comparable roll-off characteristics, a bandpass filter must have twice as many poles as a low-pass filter.

In considering the use of Butterworth, Chebyshev, and elliptic filters, it should be noted that Butterworth filters are the simplest and elliptic filters are the more complicated to design in mathematical terms.

An alternative filtering strategy, the *finite-duration impulse response* (FIR) filter, is often used in digital signal processing. The FIR filter is the equivalent of the tapped delay-line filter described in the previous section. This filter has the advantage that it has only zeros; it is thus inherently stable. An amplitude spectrum of an example FIR filter is shown in Figure 2.28. The illustrated filter was designed using an equi-ripple approach, which produces equal amounts of ripple in the passband and the stopband. The disadvantage of FIR filters is the large number of coefficients (taps) required to achieve performance similar to the other approaches.

This is not a book on filter design but the previous examples illustrate the distortions that a communication signal may encounter on its path from the sender to the receiver. This section has illustrated the potential amplitude distortion caused by filters. There is a second kind of distortion, referred to as *group-delay* distortion; the notion of group delay is discussed in Section 2.11.

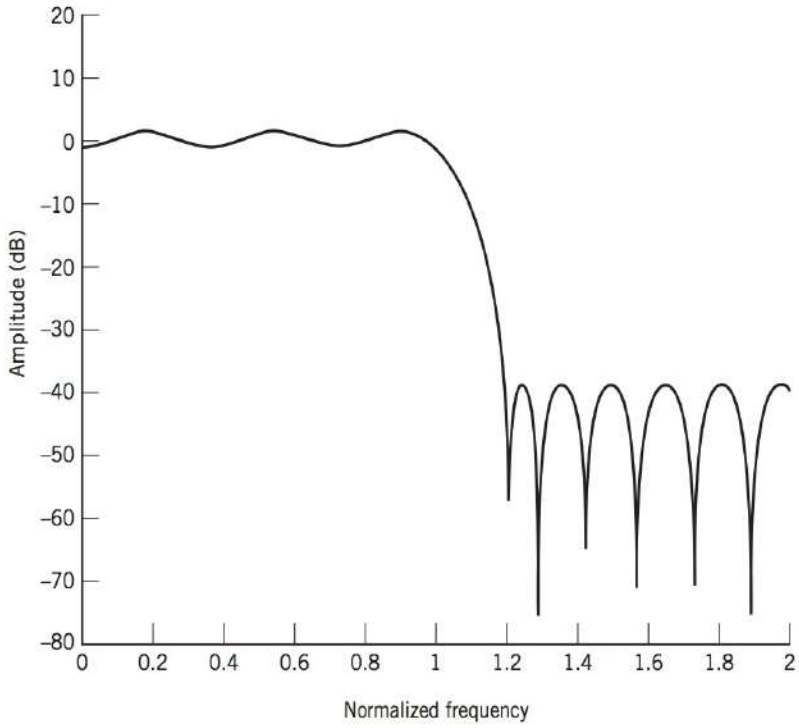


FIGURE 2.28 Amplitude response of 29-tap FIR low-pass filter.

COMMUNICATION LINK VIEWED AS A FILTER

The filters described above are usually part of the communication link that is under the control of the system designer. In addition to these filters, the channel ordinarily acts as a filter. A simple example of this occurs in radio transmission where the receiver may receive the transmitted signal over two paths: one path being the direct path between the transmitter and the receiver, and the second path being a reflection from an intervening object. Since the second path undergoes a reflection, the path length (duration) is longer by some duration τ, and the signal often undergoes an attenuation of α, and a phase rotation of ϕ, due to the reflection process. This is an example of a *multipath channel*. The impulse response of this channel may be modeled as

$$h(t) = \delta(t) + \alpha e^{j\phi}\delta(t - \tau) \tag{2.121}$$

where the first term on the right-hand side represents the unaffected direct path, and the second term represents the reflected path. The corresponding amplitude spectrum of this channel is shown in Figure 2.29 for $\alpha = 0.2$, $\phi = 180°$, and $\tau = 0.2$ microseconds. The amplitude response of the channel shows a slow but significant variation from *dc* to 5 MHz. If the bandwidth of the signal transmitted over this channel is narrow relative to the rate of variation of this channel, e.g., a signal with a 100 kHz bandwidth, then there will be very little distortion but simply an attenuation or gain. However, if the signal is relatively wideband (e.g., greater than 3 MHz in bandwidth) then it will encounter significant distortion. This is an example of a *frequency-selective* channel that was referred to in the Theme Example of Chapter 1.

2.9 LOW-PASS AND BAND-PASS SIGNALS

In the previous sections, most of the development was focused on signals whose frequency content was centered at the origin. These signals are said to be *low-pass* since their non-zero frequency content is limited by $|f| < W$. Communication using low-pass signals is referred to as *baseband communication*. Baseband communication has its

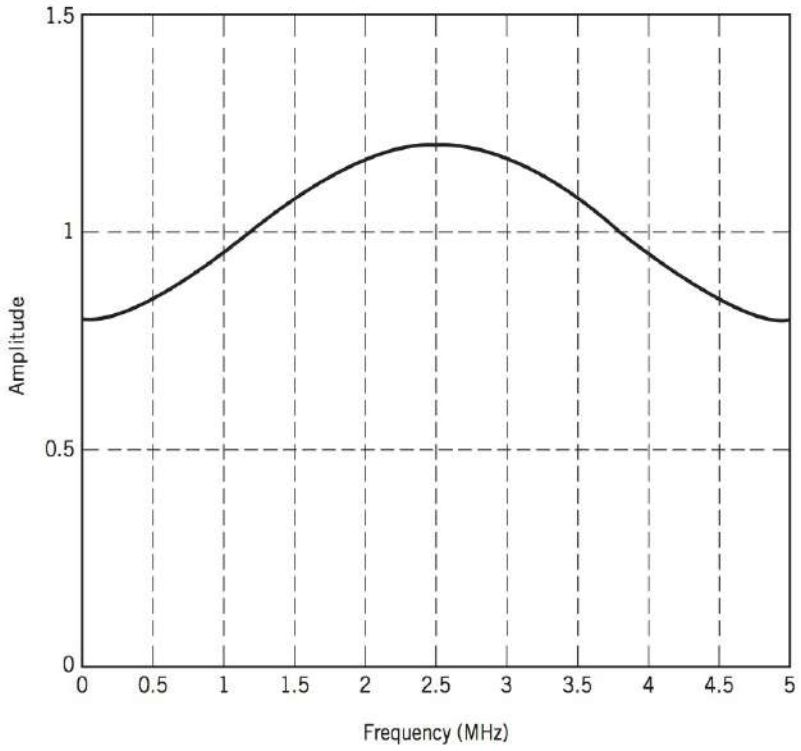


FIGURE 2.29 Amplitude response of multipath channel $h(t) = \delta(t) + \alpha e^{j\phi}\delta(t - \tau)$ with $\alpha = 0.2$, $\phi = 180°$, and $\tau = 0.2$ microseconds.

application but often is limited to wired or cabled installations. In some transmission media (e.g., radio waves) there is insufficient spectrum at baseband to share among potential applications. For other transmission media, the properties of media are not conductive to conducting signal at baseband but at other frequencies (e.g., optical fibers). In these cases, we employ *band-pass communications*. Fortunately, there are many parallels between the design of baseband and band-pass systems that simplify design techniques.

What is a band-pass signal? We say that a signal $g(t)$ is a band-pass signal if its Fourier transform $G(f)$ is non-negligible only in a band of frequencies of total extent $2W$, say, centered about the positive and negative frequency $\pm f_c$. This amplitude spectrum is illustrated in Figure 2.30*a*. We refer to f_c as the *carrier frequency*. In the majority of communication signals, we find that the bandwidth $2W$ is small compared to f_c, so we refer to such a signal as a *narrow-band signal*. However, a precise statement about how small the bandwidth must be in order for the signal to be considered narrow-band is not necessary for our present discussion.

In Figure 2.30*b*, we show a band-pass signal in the time domain. As shown in the figure, the signal is sinusoidal with approximate frequency f_c and an amplitude that varies with time. Based on this observation, a real-valued band-pass signal $g(t)$ with non-zero

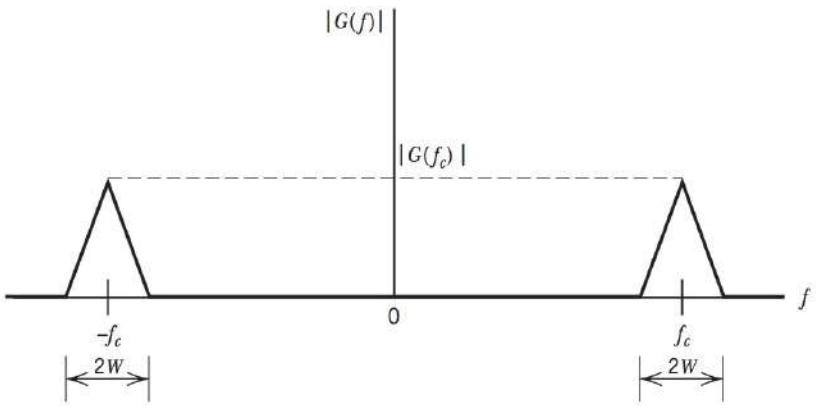


FIGURE 2.30 (*a*) Illustration of spectrum of band-pass signal.

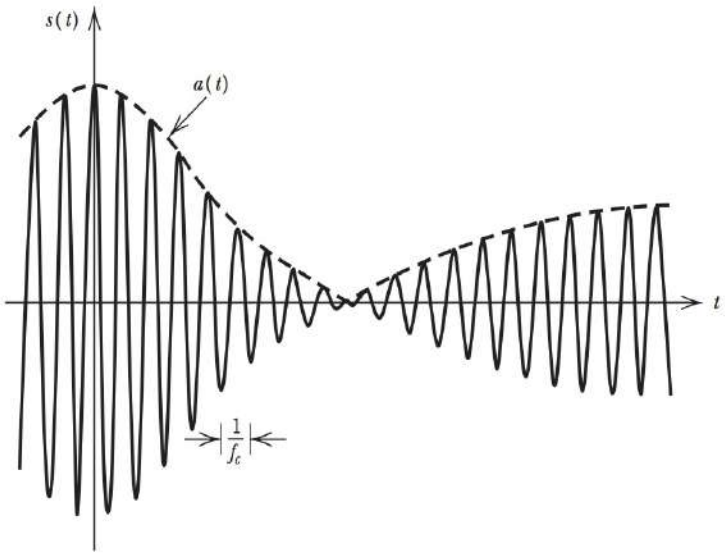


FIGURE 2.30 (*b*) A bandpass signal.

spectrum $G(f)$ in the vicinity of f_c may be expressed in the form

$$g(t) = a(t)\cos[2\pi f_c t + \phi(t)] \tag{2.122}$$

We refer to $a(t)$ as the *envelope* of the band-pass signal $g(t)$ and to $\phi(t)$ as the *phase* of the signal. The envelope is defined to be non-negative. Any point where the envelope crosses zero, the phase is adjusted by 180° to keep the envelope positive. Equation (2.122) represents the hybrid form of *amplitude modulation* and *angle modulation*, on both of which we shall have more to say in Chapters 3 and 4.

The phasor representation of the bandpass signal is a vector in the complex plane with length equal to $a(t)$ and phase $2\pi f_c t + \phi(t)$ as shown in Figure 2.31*a*. On the other hand, using the relationship $\cos(A + B) + \cos(A)\cos(B) - \sin(A)\sin(B)$, we may expand Eq. (2.122) to obtain

$$g(t) = g_I(t)\cos(2\pi f_c t) - g_Q(t)\sin(2\pi f_c t) \tag{2.123}$$

where

$$g_I(t) = a(t)\cos\phi(t) \quad \text{and} \quad g_Q(t) = a(t)\sin\phi(t) \tag{2.124}$$

are called the *in-phase* and *quadrature* components of $g(t)$, respectively. Equation (2.121) is referred to as the *canonical representation of a bandpass signal*. Returning to the phasor diagram of Figure 2.31*a*, since the term $2\pi f_c t$ represents a constant angular rotation, it may be suppressed and we are left with the phasor diagram of Figure 2.31*b*, where the phasor has been decomposed into the two orthogonal components. These two orthogonal components are the in-phase and quadrature signals defined above.

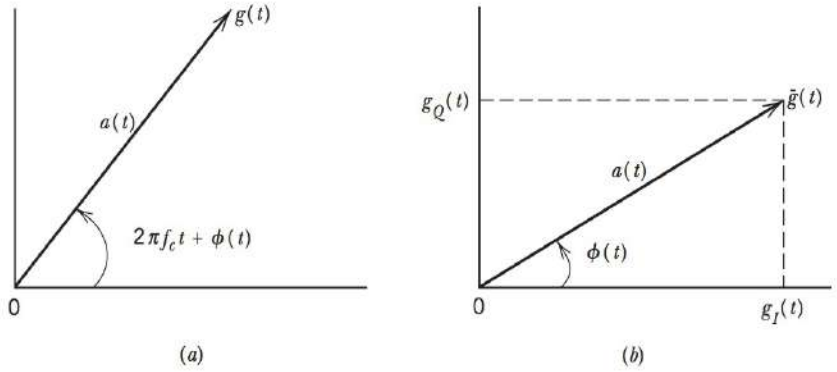


FIGURE 2.31 (a) Phasor representation of a band-pass signal $g(t)$. (b) Phasor representation of the corresponding complex envelope $\tilde{g}(t)$.

The two representations of band-pass signals, namely, the *envelope and phase description* of Eq. (2.122) and the *in-phase* and *quadrature description* of Eq. (2.123), have different advantages in different applications. The transformation from the latter to the former relies on the relations

$$a(t) = \sqrt{g_I^2(t) + g_Q^2(t)} \quad \text{and} \quad \phi(t) = \tan^{-1}\left(\frac{g_Q(t)}{g_I(t)}\right) \tag{2.125}$$

Thus, each of the quadrature components of a band-pass signal contains both amplitude and phase information. Both components are required for a unique definition of the phase $\phi(t)$, modulo 2π.

COMPLEX BASEBAND REPRESENTATION

Inspection of Eq. (2.123) indicates that it may also be written as

$$g(t) = \text{Re}[\tilde{g}(t)\exp(j2\pi f_c t)] \tag{2.126}$$

if we define the quantity $\tilde{g}(t)$ to be

$$\tilde{g}(t) = g_I(t) + jg_Q(t) \tag{2.127}$$

where both $g_I(t)$ and $g_Q(t)$ are real-valued signals. In Eq. (2.126), Re[ ] refers to the *real part* of the enclosed quantity. We refer to $\tilde{g}(t)$ as the *complex envelope* of the band-pass signal, since, in general, $\tilde{g}(t)$ is a complex-valued quantity. The complex envelope corresponds to the phasor of Figure 2.31*b*, that has the constant phase rotation $2\pi f_c t$ suppressed. Note that Eq. (2.126) may be expanded as

$$g(t) = \frac{1}{2}[\tilde{g}(t)\exp(j2\pi f_c t) + \tilde{g}^*(t)\exp(-j2\pi f_c t)] \tag{2.128}$$

If we represent the Fourier transform of $\tilde{g}(t)$ by $\tilde{G}(f)$, then by Property 5 (frequency shifting), the transform of the band-pass signal may be written as the sum of shifted versions of $\tilde{G}(f)$, as shown by

$$G(f) = \frac{1}{2}[\tilde{G}(f - f_c) + \tilde{G}^*(-f - f_c)] \tag{2.129}$$

Since the non-zero portion of the spectrum $G(f)$ is concentrated in the vicinity of f_c, Eq. (2.129) implies that the non-zero portion of $\tilde{G}(f)$ must be concentrated near the origin; thus $\tilde{G}(f)$ is the spectrum of a low-pass signal. This relationship between $G(f)$ and $\tilde{G}(f)$ is illustrated in Figures 2.32*b* and *c*.

The real-valued, low-pass nature of the signals $g_I(t)$ and $g_Q(t)$ implies that their Fourier transforms are symmetric about the origin and non-zero only for $|f| < W$ as suggested in Figure 2.32*a*. Furthermore, the Fourier transform of the band-pass signal is also guaranteed to be symmetric about the origin, since it is real. But it is not guaranteed to be symmetric about f_c as depicted in Figure 2.32*b*. Based on Eq. (2.129), the Fourier transform of the complex envelope $\tilde{g}(t)$ of a band-pass signal corresponds to a *low-pass signal* whose Fourier transform is depicted in Figure 2.32*c*.

The low-pass properties of $g_I(t)$ and $g_Q(t)$ imply that they may be derived from the band-pass signal $g(t)$ using the scheme shown in Figure 2.33*a* (with appropriate scaling), where both low-pass filters are identical, each having a bandwidth equal to W. To reconstruct $g(t)$ from its in-phase and quadrature components, we may use the scheme of Figure 2.33*b*.

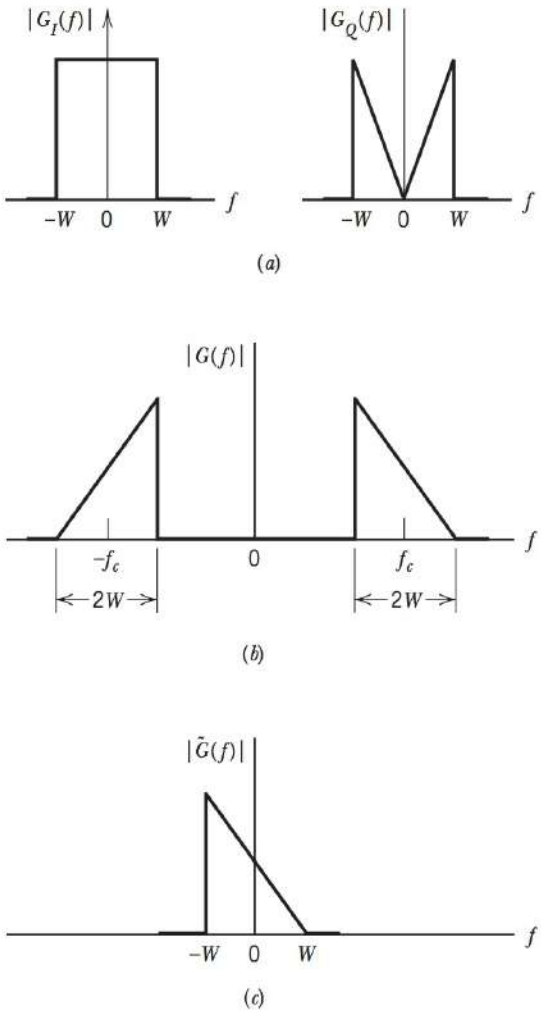


FIGURE 2.32 (*a*) In-phase $G_I(f)$ and quadrature $G_Q(f)$ component spectra; (*b*) corresponding band-pass spectrum $G(f)$; and (*c*) spectrum of complex envelope $\tilde{G}(f)$.

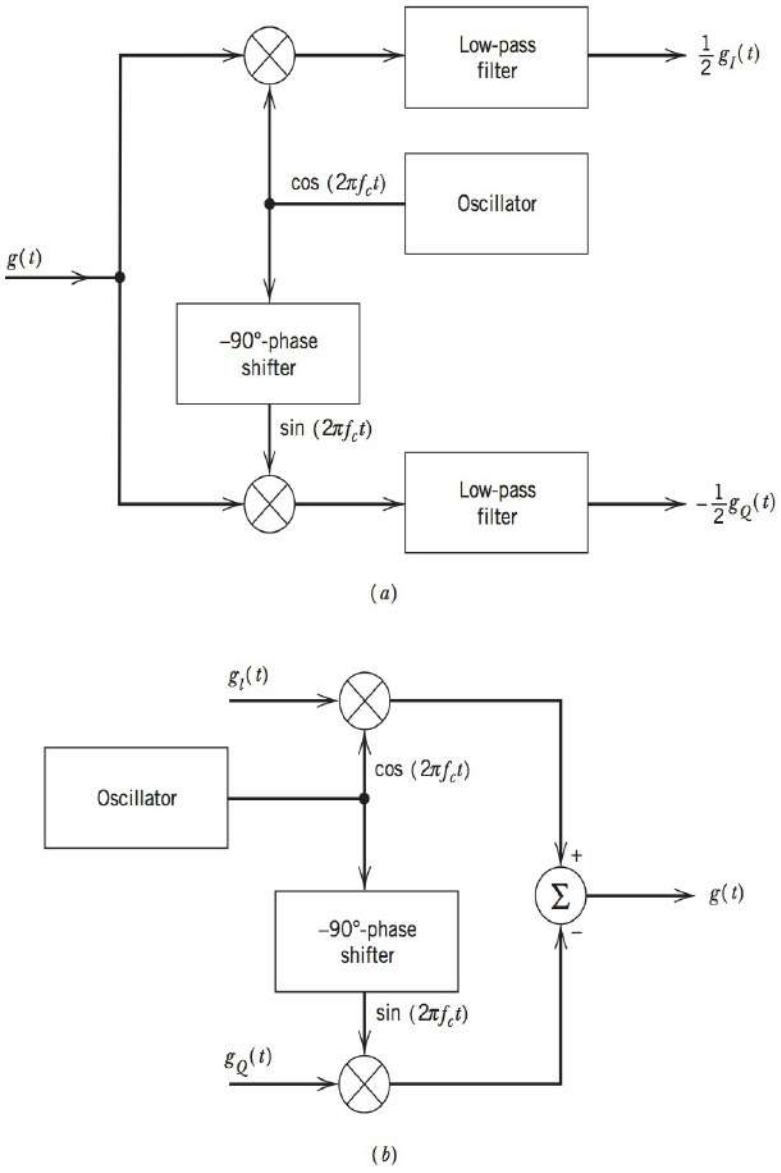


FIGURE 2.33 (*a*) Scheme for deriving the in-phase and quadrature components of a band-pass signal. (*b*) Scheme for reconstructing the band-pass signal from its in-phase and quadrature components.

From this discussion it is apparent that whether we represent a band-pass (modulated) signal $g(t)$ in terms of its in-phase and quadrature components as in Eq. (2.123) or in terms of its envelope and phase as in Eq. (2.122), the information content of the signal $g(t)$ is completely represented by the complex envelope $\tilde{g}(t)$. The particular virtue of using the complex envelope $\tilde{g}(t)$ to represent the band-pass signal is an analytical one, which centers on the suppression of the term $2\pi f_c t$; this will become evident in the following chapters.

EXAMPLE 2.14 RF Pulse (continued)

Suppose we wish to determine the complex envelope of the RF pulse defined by

$$g(t) = A\,\text{rect}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$$

We assume that $f_c T \gg 1$, so that the RF pulse $g(t)$ may be considered narrow-band. Using complex notation, we may write

$$g(t) = \text{Re}\left[A\,\text{rect}\left(\frac{t}{T}\right)\exp(j2\pi f_c t)\right]$$

From this representation, it is clear that the complex envelope is

$$\tilde{g}(t) = A\,\mathrm{rect}\left(\frac{t}{T}\right)$$

and the envelope equals

$$a(t) = |\tilde{g}(t)| = A\,\mathrm{rect}\left(\frac{t}{T}\right)$$

The latter result is intuitively satisfying. Note that in this example the complex envelope is real valued and has the same value as the envelope.

2.10 BAND-PASS SYSTEMS

With low-pass systems, we have the fundamental result that the output $y(t)$ of a linear system with impulse response $h(t)$ and input $x(t)$ is given by the convolution integral

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\,d\tau \tag{2.130}$$

In the frequency-domain, we may use the convolution property of the Fourier transform to make the equivalent statement:

$$Y(f) = H(f)\,X(f) \tag{2.131}$$

where $X(f) = F[x(t)]$, $H(f) = F[x(t)]$, and $Y(f) = F[y(t)]$. In baseband communications, $x(t)$ typically represents the message signal, while $h(t)$ is the impulse response of an element of the transmission path. Such an element may be a filter or it may be a characteristic of the channel, for example, the transmission cable.

If $x(t)$ corresponds to a band-pass signal and $h(t)$ corresponds to a filter in the transmission path, then clearly, for a linear system the relationships of Eqs. (2.130) and (2.131) still apply. When $h(t)$ is the impulse response of a bandpass filter, by analogy with $g(t)$ of Eq. (2.126), it may be represented as

$$h(t) = \mathrm{Re}[\tilde{h}(t)\exp(j2\pi f_c t)] \tag{2.132}$$

where $\tilde{h}(t)$ is the *complex impulse response* of the bandpass filter. This response and its Fourier transform may be expressed in the following way:

$$\begin{aligned} h(t) &= \frac{1}{2}[\tilde{h}(t)\exp(j2\pi f_c t) + \tilde{h}^*(t)\exp(-j2\pi f_c t)] \\ H(f) &= \frac{1}{2}[\tilde{H}(f-f_c) + \tilde{H}^*(f+f_c)] \end{aligned} \tag{2.133}$$

where we have used the relationship, $\mathrm{Re}[A] = \frac{1}{2}[A + A^*]$. Since $\tilde{H}(f)$ is low-pass limited to $|f| < B$, we deduce that the first and second terms in each of the right-hand expressions represent the positive and negative frequency portions of $H(f)$, respectively. Consequently, we have

$$\tilde{H}(f) = \begin{cases} 2H(f-f_c), & f - f_c > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.134}$$

This low-pass filter response is the frequency-domain equivalent of the complex impulse response of the filter. It is most easily understood by replacing $G(f)$ by $H(f)$ in Figure 2.32.

By linearity, the output $y(t)$ is also a band-pass signal, thus, it may be represented as

$$y(t) = \mathrm{Re}[\tilde{y}(t)\exp(j2\pi f_c t)] \tag{2.135}$$

where $\tilde{y}(t)$ is the complex envelope of $y(t)$. For the band-pass case, we may rewrite Eq. (2.131) as

$$\begin{aligned} Y(f) &= H(f)X(f) \\ &= \frac{1}{2}[\tilde{H}(f-f_c)+\tilde{H}^*(-f-f_c)] \times \frac{1}{2}[\tilde{X}(f-f_c)+\tilde{X}^*(-f-f_c)] \end{aligned} \tag{2.136}$$

where the substitutions for $X(f)$ and $H(f)$ come from Eqs. (2.129) and (2.133), respectively. Since $\tilde{H}(f)$ and $\tilde{X}(f)$ are both low-pass, it follows that

$$\tilde{H}^*(f-f_c)\tilde{X}(-f-f_c) = \tilde{H}(f-f_c)\tilde{X}^*(-f-f_c) = 0 \tag{2.137}$$

Substituting this result into Eq. (2.136) and simplifying, we obtain

$$Y(f) = \frac{1}{2}[\tilde{Y}(f-f_c)+\tilde{Y}^*(-f-f_c)] \tag{2.138}$$

where

$$\tilde{Y}(f) = \frac{1}{2}\tilde{H}(f)\tilde{X}(f) \tag{2.139}$$

Taking the inverse Fourier transform of Eq. (2.139) gives

$$\tilde{y}(t) = \frac{1}{2}\tilde{h}(t) * \tilde{x}(t) \tag{2.140}$$

That is, the complex envelope of the band-pass output is the convolution of the complex envelopes of the filter and the input, scaled by the factor 1/2. Except for the factor of 1/2, this result is the same as the baseband result.

The significance of the complex representation of Eq. (2.140) for the output signal is that, when dealing with band-pass signals and systems, we need only concern ourselves with the low-pass functions $\tilde{x}(t)$, $\tilde{y}(t)$, and $\tilde{h}(t)$, representing the excitation, the response, and the system, respectively. That is, the analysis of a band-pass system, which is complicated by the presence of the multiplying factors, $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$, is replaced by an equivalent but much simpler low-pass analysis that completely retains the essence of the filtering process. The procedure is illustrated schematically in Figure 2.34.

To sum up, the procedure for evaluating the response of a band-pass system (with mid-band frequency f_c) to an input band-pass signal (of carrier frequency f_c) is as follows:

1. The input band-pass signal $x(t)$ is replaced by its complex envelope $\tilde{x}(t)$, which is related to $x(t)$ by

$$x(t) = \text{Re}[\tilde{x}(t)\exp(j2\pi f_c t)]$$

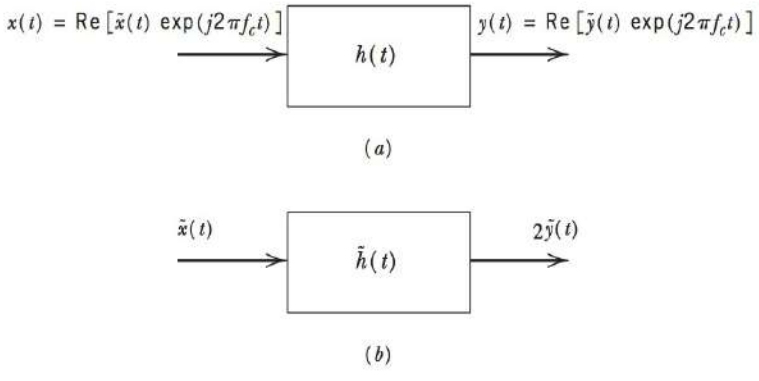


FIGURE 2.34 (*a*) Narrow-band filter of impulse response $h(t)$ with narrow-band input signal $x(t)$. (*b*) Equivalent low-pass filter of complex impulse response $\tilde{h}(t)$ with complex low-pass input $\tilde{x}(t)$.

2. The band-pass system, with impulse response $h(t)$, is replaced by a low-pass equivalent, which is characterized by a complex impulse response $\tilde{h}(t)$ related to $h(t)$ by

$$h(t) = \mathrm{Re}[\tilde{h}(t)\exp(j2\pi f_c t)]$$

3. The complex envelope $\tilde{y}(t)$ of the output band-pass signal $y(t)$ is obtained by convolving $\tilde{h}(t)$ with $\tilde{x}(t)$, as shown by

$$\tilde{y}(t) = \frac{1}{2}\tilde{h}(t) * \tilde{x}(t)$$

4. The desired output $y(t)$ is finally derived from the complex envelope $\tilde{y}(t)$ by using the relation

$$y(t) = \mathrm{Re}[\tilde{y}(t)\exp(j2\pi f_c t)]$$

EXAMPLE 2.15 Response of an Ideal Band-Pass Filter to a Pulsed RF Wave

Consider an ideal band-pass filter of mid-band frequency f_c, the amplitude response of which is band limited to $f_c - B \le |f| \le f_c + B$, as in Figure 2.35*a*, with $f_c > B$. To simplify the exposition, the effect of delay in the filter is ignored, as it has no effect on the shape of the filter response. We wish to compute the response of this filter to an RF pulse of duration T and carrier frequency f_c; it is defined by (see Figure 2.36*a*)

$$x(t) = A\ \mathrm{rect}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$$

where $f_c T \gg 1$.

Retaining the positive frequency part of the transform function $H(f)$, defined in Figure 2.35*a*, and then shifting it to the origin, we find that the transfer function $\tilde{H}(f)$ of the low-pass equivalent filter is given by (see Figure 2.35*b*)

$$\tilde{H}(f) = \begin{cases} 2, & -B < f < B \\ 0, & |f| > B \end{cases}$$

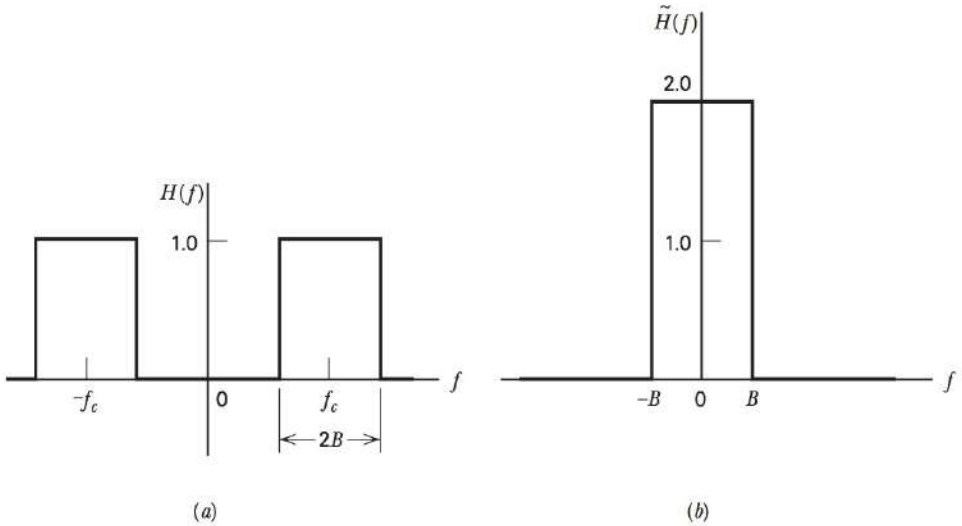


FIGURE 2.35 (*a*) Amplitude response $H(f)$ of an ideal band-pass filter. (*b*) Corresponding complex transfer function $\tilde{H}(f)$.

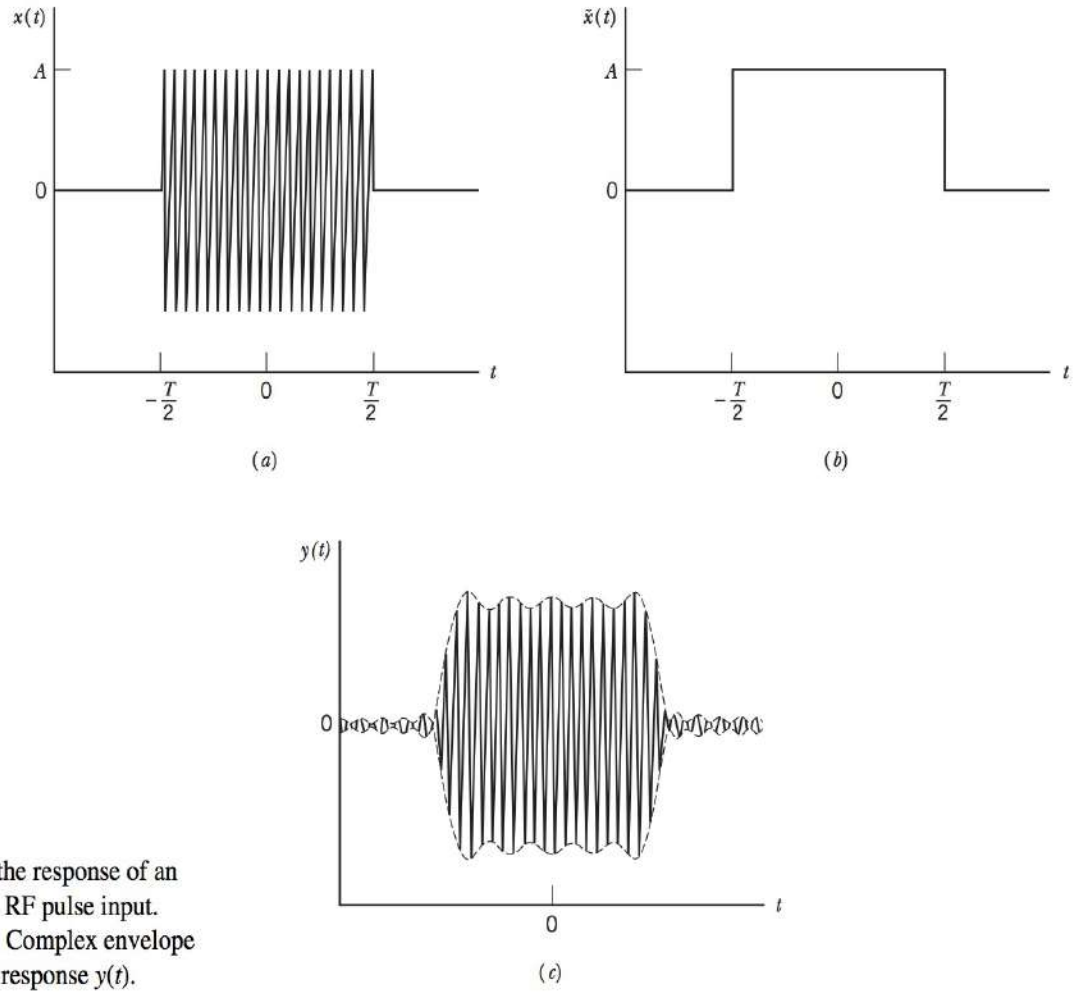


FIGURE 2.36 Illustrating the response of an ideal band-pass filter to an RF pulse input. (*a*) RF pulse input $x(t)$. (*b*) Complex envelope $\tilde{x}(t)$ of RF pulse. (*c*) Filter response $y(t)$.

The complex impulse response in this example has only a real component, as shown by

$$\tilde{h}(t) = 4B \text{ sinc}(2Bt)$$

From Example 2.14, we recall that the complex envelope $\tilde{x}(t)$ of the input RF pulse also has only a real component, as shown by (see Figure 2.36*b*)

$$\tilde{x}(t) = A \text{ rect}\left(\frac{t}{T}\right)$$

The complex envelope $\tilde{y}(t)$ of the filter output is obtained by convolving $\tilde{h}(t)$ with $\tilde{x}(t)$ and then scaling it by the factor 1/2. This convolution is exactly the same as the low-pass filtering operation that was studied in Example 2.13.

As described in Example 2.13, there is no simple closed-form analytical solution for this convolution. For illustration, the simulated response is provided in Figure 2.36 for the case of a time-bandwidth product $BT = 5$.

2.11 PHASE AND GROUP DELAY

Whenever a signal is transmitted through a dispersive (frequency-selective) device such as a filter or communication channel, some *delay* is introduced into the output signal in relation to the input signal. In an ideal low-pass filter or ideal band-pass filter, the phase

response varies *linearly* with frequency inside the passband of the filter, in which case the filter introduces a constant delay equal to t_0, say; in effect, the parameter t_0 controls the slope of the linear phase response of the filter. Now, what if the phase response of the filter is nonlinear, which is frequently the case in practice? The purpose of this section is to address this important question.

To begin the discussion, suppose that a steady sinusoidal signal at frequency f_c is transmitted through a dispersive channel or filter that has a total phase-shift of $\beta(f_c)$ radians at that frequency. By using two phasors to represent the input signal and the received signal, we see that the received signal phasor lags the input signal phasor by $\beta(f_c)$ radians. The time taken by the received signal phasor to sweep out this phase lag is simply equal to $\beta(f_c)/2\pi f_c$ seconds. This time is called the *phase delay* of the channel.

It is important to realize, however, that the phase delay is not necessarily the true signal delay. This follows from the fact that a steady sinusoidal signal does not carry information, and so it would be incorrect to deduce from the above reasoning that the phase delay is the true signal delay. In actual fact, as we will see in subsequent chapters, information can be transmitted only be applying some appropriate form of modulation to the sinusoidal wave. Suppose then that a slowly varying signal is multiplied by a sinusoidal carrier wave, so that the resulting modulated signal consists of a narrow group of frequencies centered around the carrier frequency; the waveform of Figure 2.36*c* illustrates such a modulated signal. When this modulated signal is transmitted through a communication channel, we find that there is a delay between the envelope of the input signal and that of the received signal. This delay is called the *envelope* or *group delay* of the channel; it represents the true signal delay.

Assume that the dispersive channel is described by the transfer function

$$H(f) = K \exp[j\beta(f)] \tag{2.141}$$

where the amplitude K is a constant and the phase $\beta(f)$ is a nonlinear function of frequency. The input signal $x(t)$ consists of a narrow-band signal defined by

$$x(t) = m(t)\cos(2\pi f_c t)$$

where $m(t)$ is a low-pass (information-bearing) signal with its spectrum limited to the frequency interval $|f| \leq W$. We assume that $f_c \gg W$. By expanding the phase $\beta(f)$ in a Taylor series about the point $f = f_c$, and retaining only the first two terms, we may approximate $\beta(f)$ as

$$\beta(f) \simeq \beta(f_c) + (f - f_c)\left.\frac{\partial\beta(f)}{\partial f}\right|_{f=f_c} \tag{2.142}$$

Define

$$\tau_p = -\frac{\beta(f_c)}{2\pi f_c} \tag{2.143}$$

and

$$\tau_g = -\frac{1}{2\pi}\left.\frac{\partial\beta(f)}{\partial f}\right|_{f=f_c} \tag{2.144}$$

Then we may rewrite Eq. (2.142) in the simple form

$$\beta(f) \simeq -2\pi f_c\tau_p - 2\pi(f - f_c)\tau_g \tag{2.145}$$

Correspondingly, the transfer function of the channel takes the form

$$H(f) \simeq K \exp[-j2\pi f_c\tau_p - j2\pi(f - f_c)\tau_g]$$

Following the procedure described in Section 2.10, in particular, using Eq. (2.134), we may replace the channel described by $H(f)$ by an equivalent low-pass filter whose transfer function is approximately given by

$$\tilde{H}(f) \simeq 2K \exp(-j2\pi f_c \tau_p - j2\pi f \tau_g)$$

Similarly, we may replace the input narrow-band signal $x(t)$ by its low-pass complex envelope $\tilde{x}(t)$ (for the problem at hand), which equals

$$\tilde{x}(t) = m(t)$$

The Fourier transform of $\tilde{x}(t)$ is simply

$$\tilde{X}(f) = M(f)$$

where $M(f)$ is the Fourier transform of $m(t)$. Therefore, the Fourier transform of the complex envelope of the received signal is given by

$$\begin{aligned}\tilde{Y}(f) &= \frac{1}{2}\tilde{H}(f)\tilde{X}(f) \\ &\simeq K \exp(-j2\pi f_c \tau_p)\exp(-j2\pi f \tau_g)M(f)\end{aligned}$$

We note that the multiplying factor $K \exp(-j2\pi f_c \tau_p)$ is a constant for fixed values of f_c and τ_p. We also note, from the time-shifting property of the Fourier transform, that the term $\exp(-j2\pi f \tau_g)M(f)$ represents the Fourier transform of the delayed signal $m(t - \tau_g)$. Accordingly, the complex envelope of the received signal is

$$\tilde{y}(t) \simeq K \exp(-j2\pi f_c \tau_p) m(t - \tau_g)$$

Finally, we find that the received signal is itself given by

$$\begin{aligned}y(t) &= \mathrm{Re}[\tilde{y}(t)\exp(j2\pi f_c t)] \\ &= Km(t - \tau_g)\cos[2\pi f_c(t - \tau_p)]\end{aligned} \tag{2.146}$$

Equation (2.146) shows that, as a result of transmission through the channel, two delay effects occur:

1. The sinusoidal carrier wave $\cos(2\pi f_c t)$ is delayed by τ_p seconds; hence τ_p represents the *phase delay*. Sometimes, τ_p is also referred to as the *carrier delay*.
2. The envelope $m(t)$ is delayed by τ_g seconds; hence, τ_g represents the *envelope* or *group delay*.

Note that τ_g is related to the slope of the phase $\beta(f)$, measured at $f = f_c$, as in Eq. (2.144). Note also that when the phase response $\beta(f)$ varies linearly with frequency, the signal is delayed but undistorted. When this linear condition is violated, we get group-delay distortion.

2.12 SOURCES OF INFORMATION

Communication systems support a wide variety of information sources including speech, music, television, video, facsimile, personal computers, and so on. Some of these sources are analog in the sense that the information varies continuously as a function of time, such as that shown in Figure 2.37. Speech is the prime example of an analog source. In the following section, we shall represent a generic analog waveform by the function $m(t)$.

Some sources are digital in the sense that the information can be naturally represented as a sequence of zeros and ones. Personal computers are prime examples of digital sources. A common conception of a digital signal is the random binary wave shown in Figure 2.38*a*. This digital signal consists of a sequence of rectangular pulses with

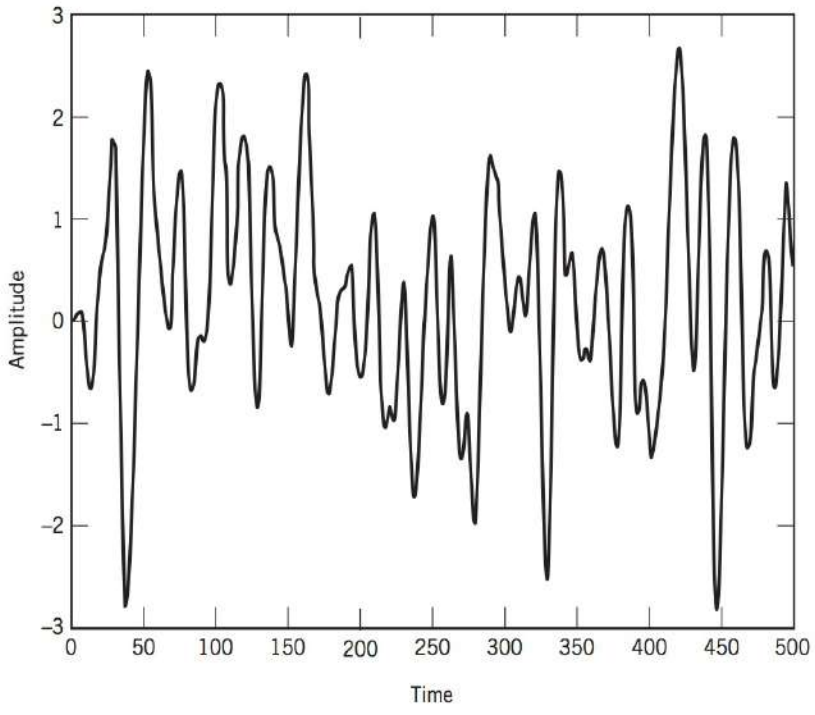


FIGURE 2.37 An example of waveform that represents an analog source of information.

amplitude either 0 or 1 corresponding to whether the corresponding bit is a logical "0" or a logical "1." Mathematically, we may represent this digital waveform as

$$g(t) = \sum_{k=t}^{K} b_k p(t - kT) \tag{2.147}$$

In this expression, $p(t)$ represents the pulse shape which could be rectangular or otherwise. The argument $t - kT$ shifts the center of pulse to kT where T is the period or duration of an individual pulse. This time-shifting is illustrated in Figure 2.38*b*. The coefficient b_k represents the kth data bit. The coefficient b_k may take the values 0 and 1, ± 1, or $\pm A$, depending upon the application. The expression for $g(t)$ mathematically describes the waveform corresponding to a sequence of K bits. The function $g(t)$ is an analog waveform although it represents a sequence of bits $\{b_k\}$.

As a consequence, any modulation technique that can be used with an analog signal $m(t)$, which may come from a voice or video source, can also be applied to the digital waveform $g(t)$. In general, the analog waveform $m(t)$ has a somewhat random nature (i.e., unpredictable). The digital waveform $g(t)$ has a predictable component $p(t)$ and a random

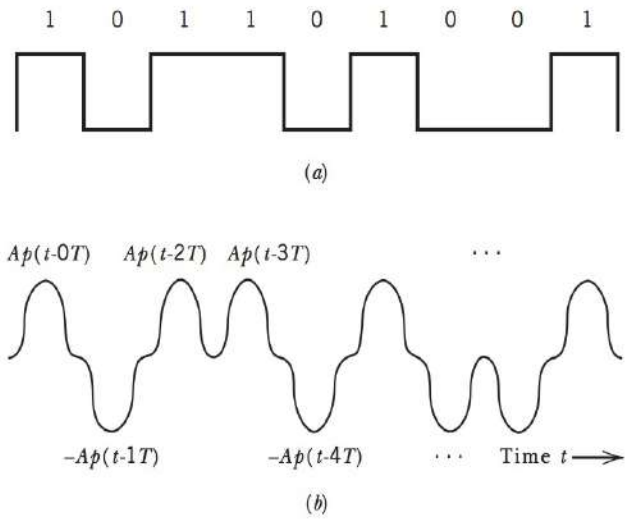


FIGURE 2.38 Illustration of random binary wave: (*a*) rectangular pulse shape and (*b*) non-rectangular pulse shape.

component b_k and, consequently, also has an overall random nature. This is what we meant in the introduction when we said that all modulation was analog. It is the sources which are either analog or digital in nature.

In general, almost all analog sources of information are being represented digitally these days. The reader should not be misled by this observation. Modulation is inherently an analog process regardless of whether the information source is analog or digital.

2.13 NUMERICAL COMPUTATION OF THE FOURIER TRANSFORM

The material presented in this chapter clearly testifies to the importance of the Fourier transform as a theoretical tool for the representation of deterministic signals and linear time-invariant systems. The importance of the Fourier transform is further enhanced by the fact that there exists a class of algorithms called fast Fourier transform algirithms[12] for the numerical computation of the Fourier transform in a highly efficient manner.

The fast Fourier transform algorithm is derived from the discrete Fourier transform in which, as the name implies, both time and frequency are represented in discrete form. The discrete Fourier transform provides an *approximation* to the Fourier transform. In order to properly represent the information content of the original signal, we have to take special care in performing the sampling operations involved in defining the discrete Fourier transform. A detailed treatment of the sampling process is presented in Chapter 5. For the present, it suffices to say that given a band-limited signal, the sampling rate should be greater than twice the highest frequency component of the input signal. Moreover, if the samples are uniformly spaced by T_s seconds, the spectrum of the signal becomes periodic, repeating every $f_s = (1/T_s)$ Hz. Let N denote the number of frequency samples contained in an interval f_s. Hence, the *frequency resolution* involved in the numerical computation of the Fourier transform is defined by

$$\Delta f = \frac{f_s}{N} - \frac{1}{NT_s} = \frac{1}{T} \tag{2.148}$$

where T is the total duration of the signal.

Consider then a *finite data sequence* $\{g_0, g_1, \ldots, g_{N-1}\}$. For brevity, we refer to this sequence as g_n, in which the subscript is the *time index* $n = 0, 1, \ldots, N-1$. Such a sequence may represent the result of sampling an *analog signal* $g(t)$ at times $t = 0, T_s, \ldots, (N-1)T_s$, where T_s is the sampling interval. The ordering of the data sequence defines the sample time in that $g_0, g_1, \ldots, g_{N-1}$ denote samples of $g(t)$ taken at times $0, T_s, \ldots, (N-1)T_s$, respectively. Thus we have

$$g_n = g(nT_s) \tag{2.149}$$

We formally define the *discrete Fourier transform* (DFT) of g_n as

$$G_k = \sum_{n=0}^{N-1} g_n \exp\left(-\frac{j2\pi}{N}kn\right) \qquad k = 0, 1, \ldots, N-1 \tag{2.150}$$

The sequence $\{G_0, G_1, \ldots, G_{N-1}\}$ is called the *transform sequence*. For brevity, we refer to this sequence as G_k, in which the subscript is the *frequency index* $k = 0, 1, \ldots, N-1$. Correspondingly, we define the *inverse discrete Fourier transform* (IDFT) of G_k as

$$g_n = \frac{1}{N}\sum_{k=0}^{N-1} G_k \exp\left(\frac{j2\pi}{N}kn\right) \qquad n = 0, 1, \ldots, N-1 \tag{2.151}$$

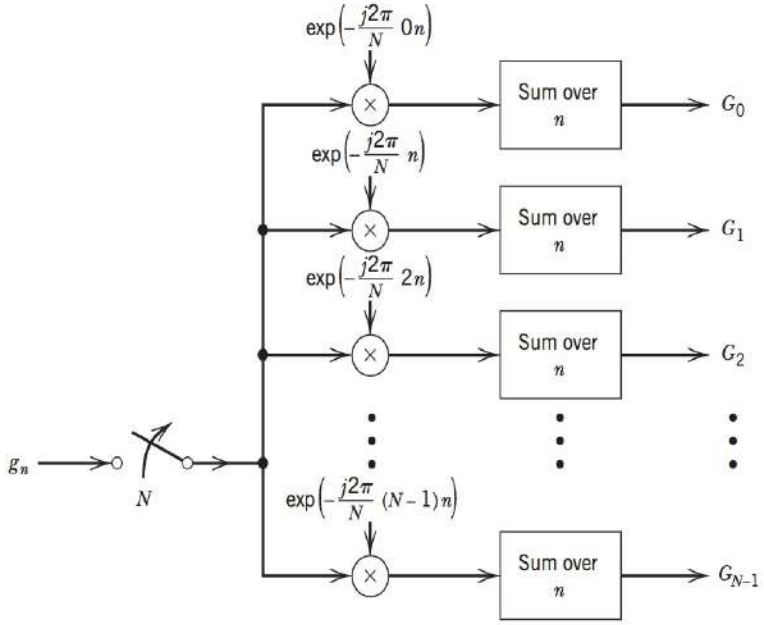
The DFT and the IDFT form a transform pair. Specifically, given a data sequence g_n, we may use the DFT to compute the transform sequence G_k; and given the transform sequence G_k, we may use the IDFT to recover the original data sequence g_n.

A distinctive feature of the DFT is that for the finite summations defined in Eqs. (2.150) and (2.151), there is no question of convergence.

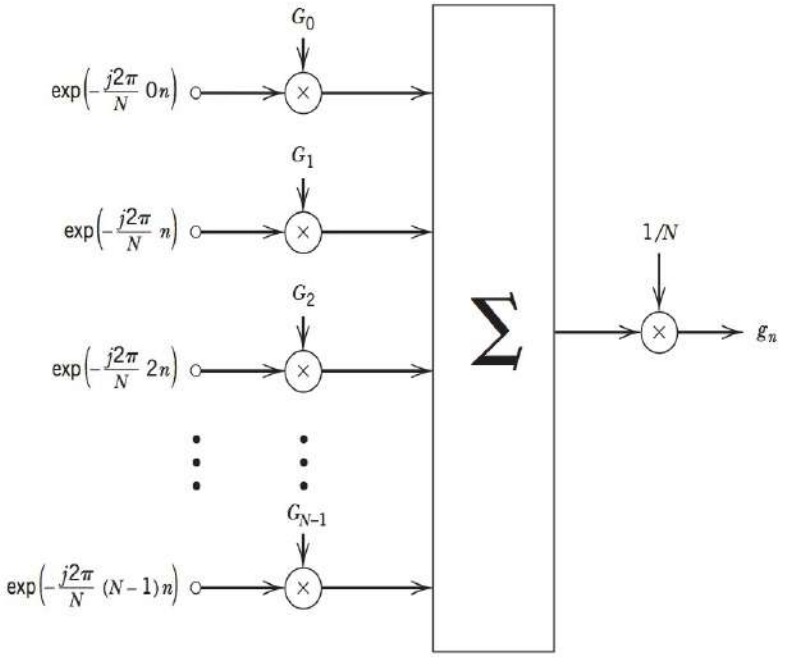
When discussing the DFT (and algorithms for its computation), the words "sample" and "point" are used interchangeably to refer to a sequence value. Also, it is common practice to refer to a sequence of length N as an *N-point sequence*, and to refer to the DFT of a data sequence of length N as an *N-point DFT*.

INTERPRETATION OF THE DFT AND THE IDFT

We may visualize the DFT process, described in Eq. (2.150), as a collection of N *complex frequency-shift* and *averaging* operations, as shown in Figure 2.39*a*. We say that the *frequency-shift* is complex in that samples of the data sequence are multiplied by *complex*



(a)



(b)

FIGURE 2.39 Interpretation of (*a*) DFT and (*b*) the IDFT.

exponential sequences. There are a total of N complex exponential sequences to be considered, corresponding to the frequency index $k = 0, 1, \ldots, N-1$. Their periods have been selected in such a way that each complex exponential sequence has precisely an integer number of cycles in the total interval 0 to $N-1$. The zero-frequency response, corresponding to $k = 0$, is the only exception.

For the interpretation of the IDFT process, described in Eq. (2.151), we may use the scheme shown in Figure 2.39*b*. Here we have a collection of N *complex signal generators*, each of which produces a complex exponential sequences:

$$\begin{aligned}\exp\left(\frac{j2\pi}{N}kn\right) &= \cos\left(\frac{2\pi}{N}kn\right) + j\sin\left(\frac{2\pi}{N}kn\right) \\ &= \left\{\cos\left(\frac{2\pi}{N}kn\right), \sin\left(\frac{2\pi}{N}kn\right)\right\}\end{aligned} \tag{2.152}$$

where $k = 0, 1, \ldots, N-1$.

Thus, each complex signal generator, in reality, consists of a pair of generators that outputs a cosinusoidal and a sinusoidal sequence of k cycles per observation interval. The output of each complex signal generator is weighted by the complex Fourier coefficient G_k. At each time index n, an output is formed by summing the weighted complex generator outputs.

Also, the addition of harmonically related periodic signals, as in Figures 2.39*a* and 2.39*b*, suggests that their outputs G_k and g_n must both be periodic. Moreover, the processors shown in Figures 2.39*a* and 2.39*b* must be linear, suggesting that the DFT and IDFT are both linear operations. This important property is also obvious from the defining equations (2.150) and (2.151).

FAST FOURIER TRANSFORM ALGORITHMS

In the discrete Fourier transform (DFT) both the input and the output consist of sequences of numbers defined at uniformly spaced points in time and frequency, respectively. This feature makes the DFT ideally suited for direct numerical evaluation on a digital computer. Moreover, the computation can be implemented most efficiently using a class of algorithms called *fast Fourier transform (FFT) algorithms*. An algorithm refers to a "recipe" that can be written in the form of a computer program.

FFT algorithms are efficient because they use a greatly reduced number of arithmetic operations as compared to the brute force computation of the DFT. Basically, an FFT algorithm attains its computational efficiency by following a "divide and conquer" strategy, whereby the original DFT computation is decomposed successively into smaller DFT computations.

A detailed treatment of the FFT algorithm may be found in any one of several books on digital signal processing.[6]

2.14 THEME EXAMPLE—CHANNEL ESTIMATION OF A WIRELESS LAN CHANNEL

One of the modulation standards used to transmit data over wireless local area networks (WLANs) is referred to as IEEE 802.11a. This standard also forms a component of the popular IEEE 802.11g standard. A practical issue associated with wireless networks is

the multipath that can occur due to receiving multiple reflections of the same signal as described in Section 2.8. One of the keys to efficient receiver design is estimating the effect that the channel has had upon the signal. The channel estimation technique that is used in the 802.11a standard is based on several of the ideas we have studied in this chapter.

In many systems, the channel is estimated by sending a known sequence or *training sequence* as part of the transmission. At the receiver, the effect of the channel on the training sequence is determined. Then, knowing the channel, the receiver compensates for the channel's effect on the remainder of the signal. For this WLAN system, it is easiest to conceptualize the training sequence in the frequency domain. In particular, the training sequence consists of a sequence of unit impulses in the frequency domain as shown in Figure 2.40*a*. If this signal is passed through the channel which has a frequency domain response shown in Figure 2.40*b*, then the impulse sequence detected at the receiver is that shown in Figure 2.40*c* with its amplitude (and phase) modified by the channel.

If the impulses are spaced sufficiently closely in the frequency domain, then they should provide a good estimate of the channel characteristics. Now we consider how such a channel estimation scheme would be implemented in practice.

The unit impulses of Figure 2.40 represent a sampling of the signal in the frequency domain; hence, it is a form of the digital signal processing described in Section 2.13. In particular, for the WLAN example, the sequence of delta functions in the frequency domain does not extend to infinity but only over the bandwidth of the signal, which is approximately ± 10 MHz. In particular, we may use the sampled representation

$$G_k = G(kf_s)$$

This equation represents the frequency-domain analog of Eq. (2.149) where f_s is the frequency spacing between impulses. Consequently, we may determine the time-domain equivalent of the sequence of frequency impulses by using the *inverse* Discrete Fourier Transform (DFT) of Eq. (2.177). The corresponding sampled time-domain signal is given by

$$g_n = \frac{1}{N}\sum_{k=0}^{N=1} G_k \exp\left[\frac{j2\pi}{N}kn\right] \qquad n = 0, \ldots, N-1$$

Thus, for the input sequence of Figure 2.40*a*, $G_k = 1$ for all k and the sampled time-domain representation of the corresponding training sequence is

$$g_n = \frac{1}{N}\sum_{k=0}^{N-1} \exp\left[\frac{j2\pi}{N}kn\right] \qquad n = 0, \ldots, N-1$$

For the WLAN case, $N = 64$ delta functions are used to span the ± 10 MHz of interest, so the spacing of the delta functions in the frequency domain is 312.5 kHz. For this application, the spacing provides sufficient resolution for estimating most of the expected channel variation. It can be shown that the inverse DFT of a sequence of impulses in the frequency domain is a unit impulse in the time domain (see Problem 2.23). When the unit impulse is limited to a bandwidth of 20 MHz, the resulting time-domain training sequence is shown in Figure 2.41.

Consequently, if the channel frequency response at the sample points is represented by the coefficients $H_k = H(kf_s)$ in Figure 2.40*b*,

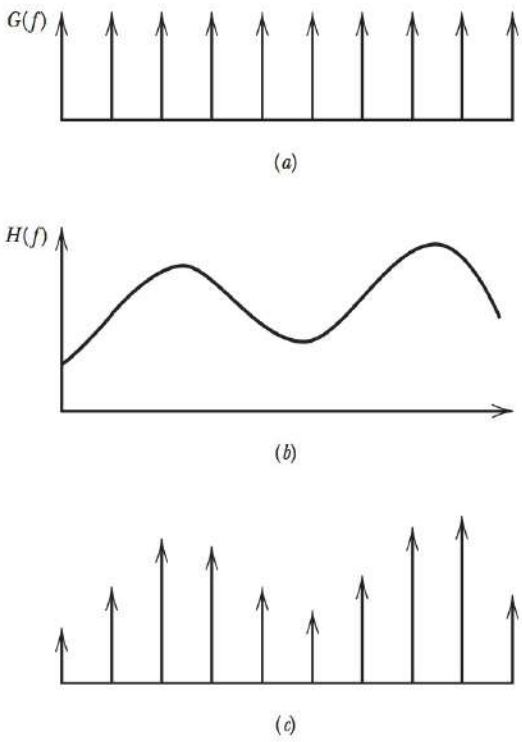


FIGURE 2.40 Illustration of channel estimation with sequence of delta functions in frequency domain: (*a*) training sequence in frequency domain; (*b*) channel frequency response; and (*c*) estimate of channel.

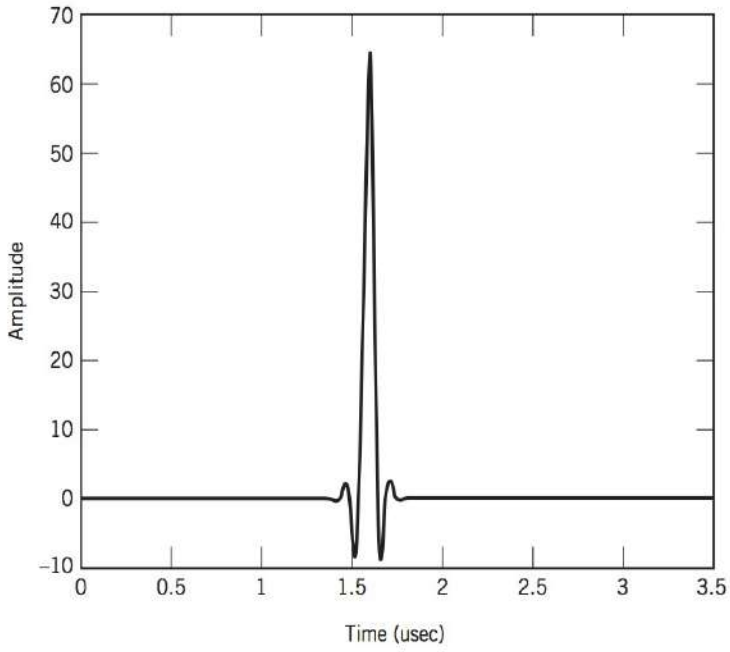


FIGURE 2.41 Illustration of a WLAN training sequence in the time domain.

then letting $\{R_k\}$ represent the DFT of the received training sequence, we have

$$\begin{aligned} R_k &= H_k G_k \\ &= H_k, \qquad k = 0, \ldots, N-1 \end{aligned}$$

which is a direct estimate of the channel frequency response. From these observations, we may derive the channel estimation and compensation process shown in Figure 2.42. In this diagram, the received signal is processed in blocks of $N = 64$ samples. If the 64 samples correspond to the training sequence, a 64-point DFT is applied to the block, and the channel estimate $\{R_k\}$ is obtained. If the 64 samples correspond to data, the DFT is applied to the block, and then the block is multiplied by the inverse of the channel estimate obtained with the previous training sequence, effectively undoing the effects of the channel. The inverse DFT is then applied to the compensated signal to convert it back to the time domain.

This type of channel compensation is a form of *equalization*, which we address in greater detail in later chapters. The compensation technique of simply inverting the channel is often *not* the most effective equalization method. However, for the choice of modulation strategy used for IEEE 802.11a, it is particularly effective.

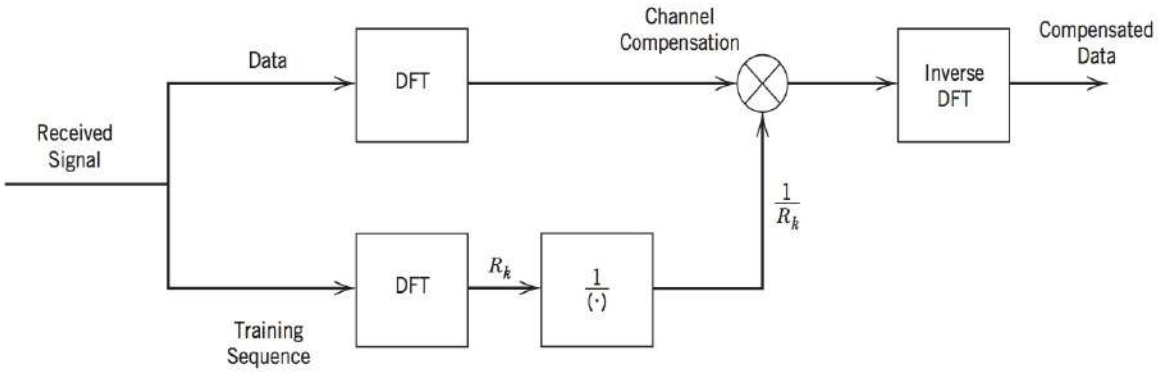


FIGURE 2.42 Block diagram of channel compensation technique based on training sequence that is a set of impulses in the frequency domain.

2.15 SUMMARY AND DISCUSSION

In this chapter we have described the Fourier transform as a fundamental mathematical tool for relating the time-domain and frequency-domain descriptions of a deterministic signal. The signal of interest may be an energy signal or a power signal. By an *energy signal* we mean a signal whose energy is finite; a *power* signal is similarly defined. The Fourier transform includes the exponential Fourier series as a special case, provided that we permit the use of the Dirac delta function.

An inverse relationship exists between the time-domain and frequency-domain descriptions of a signal. Whenever an operation is performed on the waveform of a signal in the time domain, a corresponding modification is applied to the spectrum of the signal in the frequency domain. An important consequence of this inverse relationship is the fact that the time-bandwidth product of an energy signal is a constant; the definitions of signal duration and bandwidth merely affect the value of the constant.

An important signal-processing operation frequently encountered in communication systems is that of linear filtering. This operation involves the convolution of the input signal with the impulse response of the filter or, equivalently, the multiplication of the Fourier transform of the input signal by the transfer function (i.e., Fourier transform of the impulse response) of the filter. Low-pass and band-pass filters represent two commonly used types of filters. Band-pass filtering is usually more complicated than low-pass filtering.

A key result developed in this chapter was the representation of band-pass signals in a complex baseband equivalent form. In many instances, the complex baseband form simplifies both the analysis and simulation of band-pass communication systems. It is also a key tool for the application of digital signal-processing techniques to communication systems.

The final part of the chapter was concerned with the discrete Fourier transform and its numerical computation. Basically, the discrete Fourier transform is obtained from the standard Fourier transform by uniformly sampling both the input and the output. The fast Fourier transform algorithm provides a practical tool for the efficient implementation of the discrete Fourier transform on a digital computer. This makes the fast Fourier transform algorithm a powerful computational tool for spectral analysis and linear filtering.

NOTES AND REFERENCES

1. The books by Bracewell (1986) and Champeney (1973) provide detailed treatments of the Fourier transform with emphasis on the physical aspects of the subject.
2. If a time function $g(t)$ is such that the value of the energy $\int_{-\infty}^{\infty} |g(t)|^2 dt$ is defined and finite, then the Fourier transform $G(f)$ of the function $g(t)$ exists and

$$\lim_{A\to\infty}\left[\int_{-\infty}^{\infty}\left|g(t)-\int_{-A}^{A}G(f)\exp(j2\pi ft)\,df\right|^2 dt\right]=0$$

   This result is known as *Plancherel's theorem.*
3. The energy intensity $|g(t)|^2$ and the energy spectral density $\mathcal{E}_g(f) = |G(f)|^2$ do *not* always tell the whole story about the energy content of a signal $g(t)$. This is particularly so when the spectral characteristics of the signal (e.g., speech signal) vary with time. Such signals are often referred to as *time-varying* or *nonstationary signals*. For an accurate spectral analysis of this important class of signals, we cannot use the standard Fourier transform. Rather, we need to use *time-frequency analysis*; see, for example, the book by Cohen (1995).
4. The notation $\delta(t)$ for a delta function was first introduced into quantum mechanics by Dirac. This notation is now in general use in the signal processing literature. For detailed discussions of the delta function, see Bracewell (1986, Chapter 5) and Papoulis (1984).

5. For a discussion of the Paley–Wiener criterion, see the book by Papoulis (1984).
6. Fast Fourier transform (FFT) algorithms were brought into prominence by the publication of the paper by Cooley and Tukey (1965). For discussions of FFT algorithms, see the books by Oppenheim, Schafer, and Buck (1999, Ch. 9) and Haykin and Van Veen (2005). For a discussion of how the FFT algorithm may be used to perform linear filtering, see the book *Numerical Recipes in C++* by Press et al. (2002).

PROBLEMS

2.1

(a) Find the Fourier transform of the half-cosine pulse shown in Figure P2.1*a*.

(b) Apply the time-shifting property to the result obtained in part (*a*) to evaluate the spectrum of the half-sine pulse shown in Figure P2.1*b*.

(c) What is the spectrum of a half-sine pulse having a duration equal to aT?

(d) What is the spectrum of the negative half-sine pulse shown in Figure P2.1*c*?

(e) Find the spectrum of the single sine pulse shown in Figure P2.1*d*.

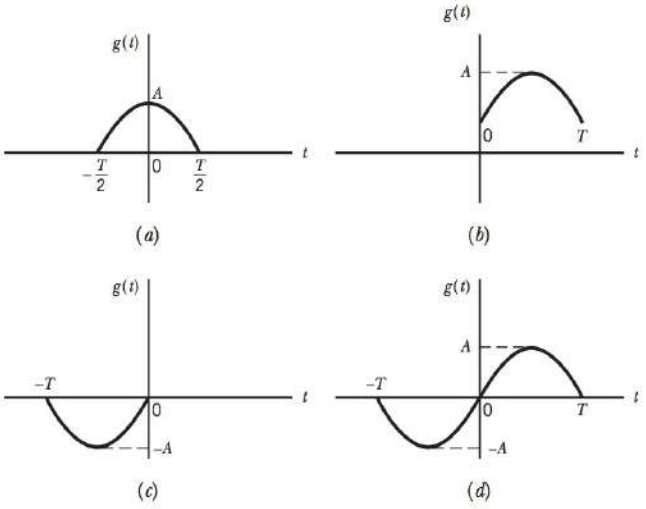


Figure P2.1

2.2 Evaluate the Fourier transform of the damped sinusoidal wave

$$g(t) = \exp(-t)\sin(2\pi f_c t)u(t)$$

where $u(t)$ is the unit step function.

2.3 Any function $g(t)$ can be split unambiguously into an *even part* and an *odd part*, as shown by

$$g(t) = g_e(t) + g_o(t)$$

The even part is defined by

$$g_e(t) = \frac{1}{2}[g(t) + g(-t)]$$

and the odd part is defined by

$$g_o(t) = \frac{1}{2}[g(t) - g(-t)]$$

(a) Evaluate the even and odd parts of a rectangular pulse defined by

$$g(t) = A\,\text{rect}\left(\frac{t}{T} - \frac{1}{2}\right)$$

(b) What are the Fourier transforms of these two parts of the pulse?

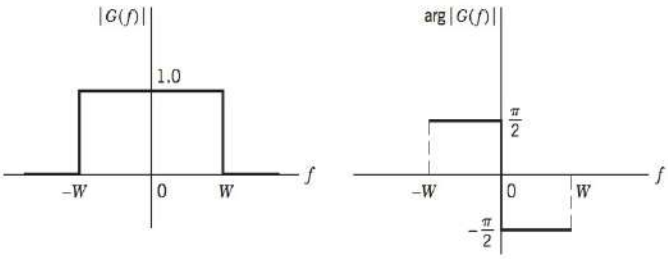


Figure P2.4

2.4 Determine the inverse Fourier transform of the frequency function $G(f)$ defined by the amplitude and phase spectra shown in Figure P2.4.

2.5 The following expression may be viewed as an approximate representation of a pulse with finite rise time:

$$g(t) = \frac{1}{\tau}\int_{t-T}^{t+T} \exp\left(-\frac{\pi u^2}{\tau^2}\right) du$$

where it is assumed that $T >> \tau$. Determine the Fourier transform of $g(t)$. What happens to this transform when we allow τ to become zero? *Hint*: Express $g(t)$ as the superposition of two signals, one corresponding to integration from $t - T$ to 0, and the other from 0 to $t + T$.

2.6 The Fourier transform of a signal $g(t)$ is denoted by $G(f)$. Prove the following properties of the Fourier transform:

(a) If a real signal $g(t)$ is an even function of time t, the Fourier transform $G(f)$ is purely real. If a real signal $g(t)$ is an odd function of time t, the Fourier transform $G(f)$ is purely imaginary.

(b)

$$t^n g(t) \rightleftharpoons \left(\frac{j}{2\pi}\right)^n G^{(n)}(f)$$

where $G^{(n)}(f)$ is the nth derivative of $G(f)$ with respect to f.

(c) $$\int_{-\infty}^{\infty} t^n g(t)\,dt = \left(\frac{j}{2\pi}\right)^n G^{(n)}(0)$$

(d) $$g_1(t)g_2^*(t) \rightleftharpoons \int_{-\infty}^{\infty} G_1(\lambda)G_2^*(\lambda - f)\,d\lambda$$

(e) $$\int_{-\infty}^{\infty} g_1(t)g_2^*(t)\,dt = \int_{-\infty}^{\infty} G_1(f)G_2^*(f)\,df$$

2.7 The Fourier transform $G(f)$ of a signal $g(t)$ is bounded by the following three inequalities:

$$|G(f)| \leq \int_{-\infty}^{\infty} |g(t)|dt$$

$$|j2\pi fG(f)| \leq \int_{-\infty}^{\infty} \left|\frac{dg(t)}{dt}\right| dt$$

and

$$|(j2\pi f)^2 G(f)| \leq \int_{-\infty}^{\infty} \left|\frac{d^2g(t)}{dt^2}\right| dt$$

where it is assumed that the first and second derivatives of $g(t)$ exist.

Construct these three bounds for the triangular pulse shown in Figure P2.7 and compare your results with the actual amplitude spectrum of the pulse.

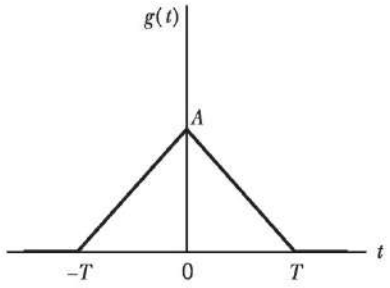


Figure P2.7

2.8 Prove the following properties of the convolution process:

(a) The commutative property:

$$g_1(t) \star g_2(t) = g_2(t) \star g_1(t)$$

(b) The associative property:

$$g_1(t) \star [g_2(t) \star g_3(t)] = [g_1(t) \star g_2(t)] \star g_3(t)$$

(c) The distributive property:

$$g_1(t) \star [g_2(t) + g_3(t)] = g_1(t) \star g_2(t) + g_1(t) \star g_3(t)$$

2.9 Consider the convolution of two signals $g_1(t)$ and $g_2(t)$. Show that

(a) $$\frac{d}{dt}[g_1(t) \star g_2(t)] = \left[\frac{d}{dt}g_1(t)\right] \star g_2(t)$$

(b) $$\int_{-\infty}^{t} [g_1(\tau) \star g_2(\tau)]d\tau = \left[\int_{-\infty}^{t} g_1(\tau)d\tau\right] \star g_2(t)$$

2.10 A signal $x(t)$ of finite energy is applied to a square-law device whose output $y(t)$ is defined by

$$y(t) = x^2(t)$$

The spectrum of $x(t)$ is limited to the frequency interval $-W \leq f \leq W$. Hence, show that the spectrum of $y(t)$ is limited to $-2W \leq f \leq 2W$. *Hint*: Express $y(t)$ as $x(t)$ multiplied by itself.

2.11 Evaluate the Fourier transform of the delta function by considering it as the limiting form of (1) a rectangular pulse of unit area, and (2) a sinc pulse of unit area.

2.12 The Fourier transform $G(f)$ of a signal $g(t)$ is defined by

$$G(f) = \begin{cases} 1, & f > 0 \\ \frac{1}{2}, & f = 0 \\ 0, & f < 0 \end{cases}$$

Determine the signal $g(t)$.

2.13 Show that the two different pulses defined in parts (*a*) and (*b*) of Figure P2.1 have the same energy spectral density:

$$\varepsilon_g(f) = \frac{4A^2T^2\cos^2(\pi Tf)}{\pi^2(4T^2f^2 - 1)^2}$$

2.14

(a) The *root mean-square* (*rms*) *bandwidth* of a low-pass signal $g(t)$ of finite energy is defined by

$$W_{\text{rms}} = \left[\frac{\int_{-\infty}^{\infty} f^2|G(f)|^2\,df}{\int_{-\infty}^{\infty} |G(f)|^2\,df}\right]^{1/2}$$

where $|G(f)|^2$ is the energy spectral density of the signal. Correspondingly, the *root mean-square (rms) duration* of the signal is defined by

$$T_{\text{rms}} = \left[\frac{\int_{-\infty}^{\infty} t^2|g(t)|^2\,dt}{\int_{-\infty}^{\infty} |g(t)|^2\,dt}\right]^{1/2}$$

Using these definitions, show that

$$T_{\text{rms}}W_{\text{rms}} \geq \frac{1}{4\pi}$$

Assume that $|g(t)| \to 0$ faster than $1\sqrt{|t|}$ as $|t| \to \infty$.

(b) Consider a Gaussian pulse defined by

$$g(t) = \exp(-\pi t^2)$$

Show that, for this signal, the equality

$$T_{\text{rms}}W_{\text{rms}} \equiv \frac{1}{4\pi}$$

can be reached.

Hint: Use Schwarz's inequality:

$$\left\{\int_{-\infty}^{\infty}[g_1^*(t)g_2(t)+g_1(t)g_2^*(t)]dt\right\}^2 \le 4\int_{-\infty}^{\infty}|g_1(t)|^2\,dt \times \int_{-\infty}^{\infty}|g_2(t)|^2\,dt$$

in which we set

$$g_1(t) = tg(t)$$

and

$$g_2(t) = \frac{dg(t)}{dt}$$

2.15 Let $x(t)$ and $y(t)$ be the input and output signals of a linear time-invariant filter. Using Rayleigh's energy theorem, show that if the filter is stable and the input signal $x(t)$ has finite energy, then the output signal $y(t)$ also has finite energy. That is, given that

$$\int_{-\infty}^{\infty}|x(t)|^2\,dt < \infty$$

then show that

$$\int_{-\infty}^{\infty}|y(t)|^2\,dt < \infty$$

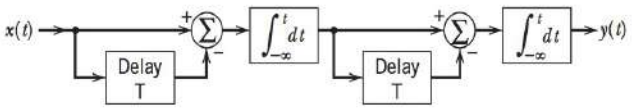


Figure P2.16

2.16 Evaluate the transfer function of a linear system represented by the block diagram shown in Figure P2.16.

2.17

(a) Determine the overall amplitude response of the cascade connection shown in Figure P2.17 consisting of N identical stages, each with a time constant RC equal to τ_0.

(b) Show that as N approaches infinity, the amplitude response of the cascade connection approaches the Gaussian function $\exp\left(-\frac{1}{2}f^2T^2\right)$, where for each value of N, the time constant τ_0 is selected so that

$$\tau_0^2 = \frac{T^2}{4\pi^2 N}$$

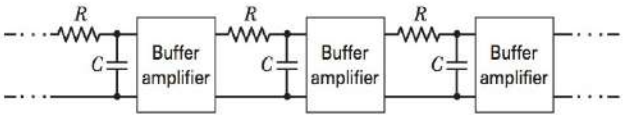


Figure P2.17

2.18 Suppose that, for a given signal $x(t)$, the integrated value of the signal over an interval T is required, as shown by

$$y(t) = \int_{t-T}^{T} x(\tau)\,d\tau$$

(a) Show that $y(t)$ can be obtained by processing the signal $x(t)$ with a filter having the transfer function

$$H(f) = T\,\mathrm{sinc}(fT)\exp(-j\pi fT)$$

(b) An adequate approximation to this transfer function is obtained by using a low-pass filter with a bandwidth equal to $1/T$, passband amplitude response T, and delay $T/2$. Assuming this low-pass filter is created by the combination of an RC circuit followed by a gain of T, determine the filter output at time $t = T$ due to a unit step function applied to the filter at $t = 0$, and compare the result with the corresponding output of the ideal integrator.

2.19 A tapped-delay-line filter consists of N weights, where N is odd. It is symmetric with respective to the center tap, that is, the weights satisfy the condition

$$w_n = w_{N-1-n} \qquad 0 \le n \le N-1$$

(a) Find the amplitude response of the filter.

(b) Show that this filter has a linear phase response.

2.20 Consider an ideal band-pass filter with frequency midband f_c and bandwidth $2B$, as defined in Figure P2.20. The carrier wave $A\cos(2\pi f_0 t)$ is suddenly applied to this filter at time $t = 0$. Assuming that $|f_c - f_0|$ is large compared to the bandwidth $2B$, determine the response of the filter.

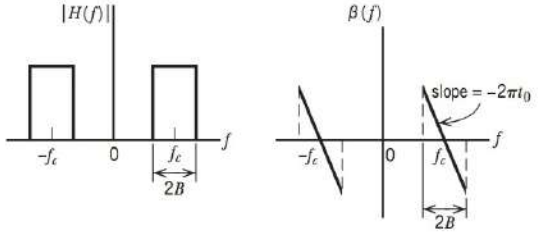


Figure P2.20

2.21 The rectangular RF pulse

$$x(t) = \begin{cases} A\cos(2\pi f_c t), & 0 \le t \le T \\ 0, & \text{elsewhere} \end{cases}$$

is applied to a linear filter with impulse response

$$h(t) = x(T-t)$$

Assume that the frequency f_c equals a large integer multiple of $1/T$. Determine the response of the filter and sketch it.

2.22 Show that the inverse DFT of a sequence of constant impulses in the frequency domain is a corresponding sequence of impulses in the time domain when the DFT length is even.

Computer Problems

2.23 A rectangular pulse $x(t)$ of unit amplitude and duration T is applied to an ideal low-pass filter of bandwidth B.

(a) What is the impulse response of the ideal low-pass filter?

(b) Determine and plot the response $y(t)$ of the filter for $BT = 5$, 10, 20. Using the following Matlab script,

```
% - - - Simulation parameters - - - -
BT      = 5; %BT product
T       = 1;
B       = BT/T;
Delta_t = T/100;
t       = [-6*T: Delta_t: 6*T];
```

```
% - - - Pulse of unit amplitude and duration T - -
x       = zeros(size(t));
index   = find (abs(t)<T/2);
x(index) = 1;

% - - - (truncated) impulse response of ideal filter - - -
t1 = [-3*T: Delta_t: 3*T];
h = 2*B*sinc(2*B*t1);

% - - - filter response - - -
y = filter(h, 1, x) * Delta_t; % has delay of 3T

% - - - plot results (removing delay) - - - -
subplot(2,1,1), plot(t,x), axis([-T T -0.25 1.25]), grid on;
subplot(2,1,2), plot(t - 3*T, y), axis([-T T -0.25 1.25]),
grid on;
```

(c) Prepare a table with columns of BT, the oscillation frequency of the response, and the percentage overshoot.

(d) Repeat the experiment for the case of $BT = 100$ but vary the sample period Δt. What is observed? Add the value $BT = 100$ to the table of part (c). Draw conclusions from the table.

2.24 Repeat the experiment of Problem 2.23 using an input consisting of the raised cosine pulse:

$$x(t) = \begin{cases} 1 + \cos\left(\dfrac{\pi t}{T}\right), & -T \le t \le T \\ 0, & \text{otherwise} \end{cases}$$

In particular, evaluate the filter response for the time-bandwidth product $BT = 5$, 10, 100. Compare the results of this experiment with that of Problem 2.24.

2.25 Repeat the experiment of Problem 2.23 using an input consisting of the periodic square-wave input.

(a) Set $B = 1$ and $T = 0.5/f_0$; determine and plot the response $y(t)$ of the filter for a square-wave of frequencies: $f_0 = 0.1, 0.25, 0.5$, 1.0, and 1.1 Hz. Use the following to generate a square wave and modify the Matlab script of Problem 2.23 accordingly.

```
% - - - Square wave of frequency f0 - -
x          = zeros(size(t));
index_1    = find(mod(t, 2*T) < T);
x(index_1) = 1;
index_ml   = find(mod(t, 2*T) > = T);
x(index_ml) = -1;
```

(b) Prepare a table with columns of BT and maximum amplitude of the filter response. Draw conclusions.

(c) Increase the length of truncated impulse response from $6T$ to $12T$. (Increase the simulation and observation times accordingly.) What happens to the results of part (b)? Explain.

2.26 Use the sound card on a PC with the following Matlab script to capture the sound output of an MP3 player, radio, microphone, or similar device.

```
Fs = 8000;                  % sample rate: eg. 2250, 8000,
                            11025, or 44100 Hz
N = Fs*10;                  % number of samples in 10s of data
FFTsize = 1024;
y = wavrecord(N, Fs);       % collect data
Y = spectrum(y, FFTsize);   % compute average amplitude spec-
                            trum
Freq = [0:Fs/FFTsize:Fs/2]; % frequency scale
Time = [1:N]/Fs;            % time scale
subplot(2,1,1), plot(Time, y),
ylabel('Amplitude'), xlabel('Time(s)');
subplot(2,1,2), plot(Freq,10*log10(Y/max(Y))),
ylabel('Spectrum(dB)'), xlabel('Frequency(Hz)');
```

(a) Collect several samples from the same type of source, for example, music. Compare and comment on your results.

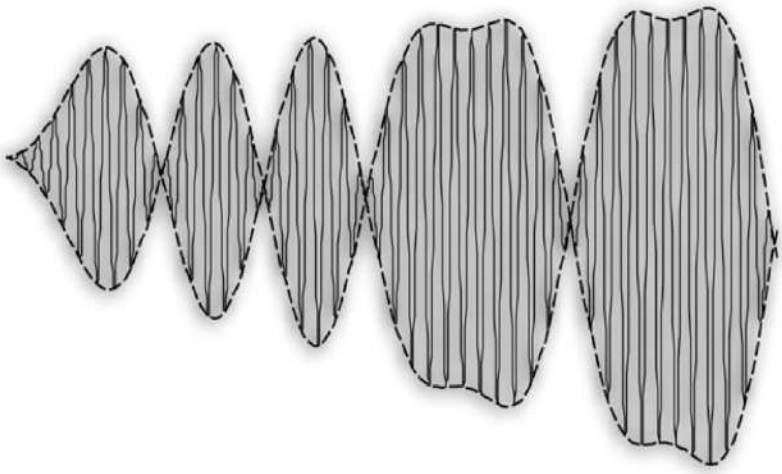
(b) Collect samples from different sources, for example, music and speech. Compare and comment. What happens if the sample rate is changed? Explain.

2.27 In this problem, we perform the WLAN channel estimation of the Section 2.14 Theme Example for a specific case using the following Matlab code.

```
% - - - WLAN problem
alpha = 0.5;
t = [0:63]; % microseconds
G = ones(1,64);
g = ifft(G);
h = exp(- alpha*t);
r = filter(h, l, g);
R = fftshift(fft(r));
stem(abs(R))
```

What is the time-domain response of the training sequence? Plot the estimated amplitude spectrum of the channel and comment. What is the estimated phase spectrum of channel? Explain.

3



AMPLITUDE MODULATION

3.1 INTRODUCTION

The purpose of a communication system is to convey *information* through a medium or communication channel separating the transmitter from the receiver. The information is often represented as a baseband signal, that is, a signal whose spectrum extends from 0 to some maximum frequency. Proper utilization of the communication channel often requires a shift of the range of baseband frequencies into other frequency ranges suitable for transmission, and a corresponding shift back to the original frequency range after reception. For example, a radio system must operate with frequencies of 30 kHz and upward, whereas the baseband signal may contain frequencies in the audio frequency range, so some form of frequency-band shifting must be used for the system to operate satisfactorily. A shift of the range of frequencies in a signal is accomplished by using *modulation*, which is defined as *the process by which some characteristic of a carrier is varied in accordance with a modulating wave (signal)*. A common form of the carrier is a *sinusoidal wave*, in which case we speak of *continuous-wave modulation*. The baseband signal is referred to as the *modulating wave*, and the result of the modulation process is referred to as the *modulated wave*. Modulation is performed at the transmitting end of the communication system. At the receiving end of the system, we usually require the original baseband signal to be restored. This is accomplished by using a process known as *demodulation*, which is the reverse of the modulation process.

In this chapter we study the first of two widely used families of continuous-wave (CW) modulation systems, namely, *amplitude modulation*. In amplitude modulation, the amplitude of a sinusoidal carrier wave is varied in accordance with the baseband signal. Sections 3.2 through 3.6 are devoted to the standard form of amplitude modulation and its variants. In Sections 3.7 and 3.8 we discuss the ideas of *frequency translation* and *frequency-division multiplexing* for sharing a common channel among a multitude of different users.

3.2 AMPLITUDE MODULATION

Consider a *sinusoidal carrier wave* $c(t)$ defined by

$$c(t) = A_c \cos(2\pi f_c t) \tag{3.1}$$

where A_c is the *carrier amplitude* and f_c is the *carrier frequency*. To simplify the exposition without affecting the results obtained and conclusions reached, we have assumed that the phase of the carrier wave is zero in Eq. (3.1). Let $m(t)$ denote the baseband signal that carries the specification of the message. The source of carrier wave $c(t)$ is physically independent of the source responsible for generating $m(t)$. *Amplitude modulation (AM) is defined as a process in which the amplitude of the carrier wave $c(t)$ is varied about a mean value, linearly with the baseband signal $m(t)$.* An amplitude-modulated (AM) wave may thus be described, in its most general form, as a function of time as follows:

$$s(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t) \tag{3.2}$$

where k_a is a constant called the *amplitude sensitivity* of the modulator responsible for the generation of the modulated signal $s(t)$. Typically, the carrier amplitude A_c and the message signal $m(t)$ are measured in volts, in which case the amplitude sensitivity k_a is measured in volt^{-1}.

Figure 3.1*a* shows a baseband signal $m(t)$, and Figures 3.1*b* and 3.1*c* show the corresponding AM wave $s(t)$ for two values of amplitude sensitivity k_a and a carrier amplitude $A_c = 1$ volt. We observe that the *envelope* of $s(t)$ has essentially the same shape as the baseband signal $m(t)$ provided that two requirements are satisfied:

1. The amplitude of $k_a m(t)$ is always less than unity, that is,

$$|k_a m(t)| < 1 \qquad \text{for all } t \tag{3.3}$$

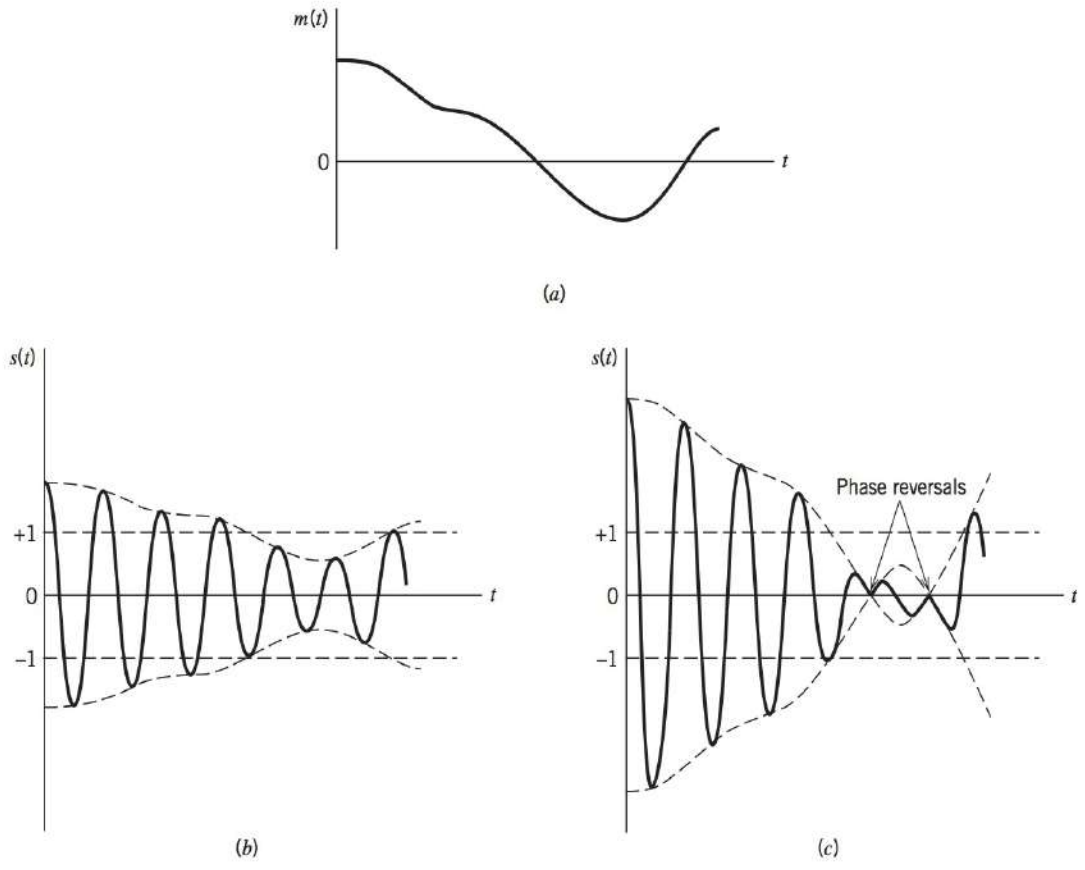


FIGURE 3.1 Illustrating the amplitude modulation process. (*a*) Baseband signal $m(t)$. (*b*) AM wave for $|k_a m(t)| < 1$ for all t. (*c*) AM wave for $|k_a m(t)| > 1$ for some t.

This condition is illustrated in Figure 3.1*b*; it ensures that the function $1 + k_a m(t)$ is always positive, and since an envelope is a positive function, we may express the envelope of the AM wave $s(t)$ of Eq. (3.2) as $A_c[1 + k_a m(t)]$. When the amplitude sensitivity k_a of the modulator is large enough to make $|k_a m(t)| > 1$ for any t, the carrier wave becomes *overmodulated*, resulting in carrier phase reversals whenever the factor $1 + k_a m(t)$ crosses zero. The modulated wave then exhibits *envelope distortion*, as in Figure 3.1*c*. It is therefore apparent that by avoiding overmodulation, a one-to-one relationship is maintained between the envelope of the AM wave and the modulating wave for all values of time—a useful feature, as we shall see later on. The absolute maximum value of $k_a m(t)$ multiplied by 100 is referred to as the *percentage modulation*.

Guglielmo Marconi (1874–1937)

Marconi was an Italian inventor famous for his development of a radio-telegraph system. In 1864, James Clerk Maxwell had formulated the electromagnetic theory of light and predicted the existence of radio waves. The existence of radio waves was established experimentally in 1887 by Heinrich Hertz. From that time forward many experimenters were investigating the practical applications of radio waves with a reliable radio-telegraphy system being a common goal. Beginning in 1896, Marconi began demonstrating radiotelegraphy over longer and longer distances, culminating in the first transatlantic transmission from Cornwall, England to Signal Hill, Newfoundland in 1901.

Marconi's radio-telegraph used a device known as a spark-gap transmitter to generate radio waves. A simple spark gap consists of the conducting electrodes separated by a gap that is usually filled with air. When sufficient voltage is applied a spark is formed, ionizing the air between the electrodes and exciting an inductance-capacitance (LC) circuit that is connected to an antenna. Although the LC-circuit does provide some tuning, the spark tends to produce a broad spectrum of electromagnetic radiation, quite unlike the narrowband bandpass signals discussed in this chapter. Spark gap transmitters were eventually legislated out of existence due to the interference they caused. It is interesting that recent developments in communications, specifically, impulse radio techniques, generate a broad spectrum signal, not unlike a spark-gap transmitter.

Karl Braun and Marconi were awarded the 1909 Nobel Prize in Physics for their development of a radio-telegraph system. In later years, Marconi became an active Italian Fascist, and his speech-making in that regard was notorious to the extent that he was banned from the BBC that he had helped create.

2. The carrier frequency f_c is much greater than the highest frequency component W of the message signal $m(t)$, that is

$$f_c \gg W \tag{3.4}$$

We call W the *message bandwidth*. If the condition of Eq. (3.4) is not satisfied, an envelope cannot be visualized (and therefore detected) satisfactorily.

From Eq. (3.2), we find that the Fourier transform of the AM wave $s(t)$ is given by

$$\begin{aligned} S(f) &= \frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] \\ &\quad + \frac{k_a A_c}{2}[M(f - f_c) + M(f + f_c)] \end{aligned} \tag{3.5}$$

Suppose that the baseband signal $m(t)$ is band-limited to the interval $-W \le f \le W$, as in Figure 3.2*a*. The shape of the spectrum shown in this figure is intended for the purpose of illustration only. We find from Eq. (3.5) that the spectrum $S(f)$ of the AM wave is as shown in Figure 3.2*b* for the case when $f_c > W$. This spectrum consists of two delta functions weighted by the factor $A_c/2$ and occurring at $\pm f_c$, and two versions of the baseband spectrum translated in frequency by $\pm f_c$ and scaled in amplitude by $k_a A_c/2$. From the spectrum of Figure 3.2*b*, we note the following:

1. As a result of the modulation process, the spectrum of the message signal $m(t)$ for negative frequencies extending from $-W$ to 0 becomes completely visible for positive (i.e., measurable) frequencies, provided that the carrier frequency satisfies the condition $f_c > W$; herein lies the importance of the idea of "negative" frequencies.
2. For positive frequencies, the portion of the spectrum of an AM wave lying above the carrier frequency f_c is referred to as the *upper sideband*, whereas the symmetric portion below f_c is referred to as the *lower sideband*. For negative frequencies, the upper sideband is represented by the portion of the spectrum below $-f_c$ and the lower sideband by the portion above $-f_c$. The condition $f_c > W$ ensures that the sidebands do not overlap.
3. For positive frequencies, the highest frequency component of the AM wave equals $f_c + W$, and the lowest frequency component equals $f_c - W$. The difference between these two frequencies defines the *transmission bandwidth* B_T for an AM wave, which is exactly twice the message bandwidth W, that is,

$$B_T = 2W \tag{3.6}$$

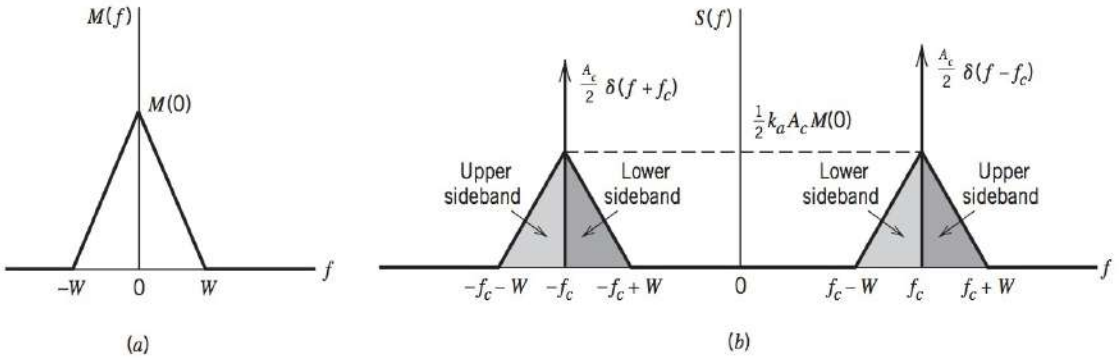


FIGURE 3.2 (*a*) Spectrum of baseband signal. (*b*) Spectrum of AM wave.

EXAMPLE 3.1 Single-Tone Modulation

Consider a modulating wave $m(t)$ that consists of a single tone or frequency component, that is,

$$m(t) = A_m \cos(2\pi f_m t)$$

where A_m is the amplitude of the sinusoidal modulating wave and f_m is its frequency (see Figure 3.3*a*). The sinusoidal carrier wave has amplitude A_c and frequency f_c (see Figure 3.3*b*). The corresponding AM wave is therefore given by

$$s(t) = A_c[1 + \mu \cos(2\pi f_m t)]\cos(2\pi f_c t) \tag{3.7}$$

where

$$\mu = k_a A_m$$

The dimensionless constant μ is the *modulation factor*, or the percentage modulation when it is expressed numerically as a percentage. To avoid envelope distortion due to overmodulation, the modulation factor μ must be kept below unity.

Figure 3.3*c* shows a sketch of $s(t)$ for μ less than unity. Let $A_{\max}$ and $A_{\min}$ denote the maximum and minimum values of the envelope of the modulated wave. Then, from Eq. (3.7) we get

$$\frac{A_{\max}}{A_{\min}} = \frac{A_c(1+\mu)}{A_c(1-\mu)}$$

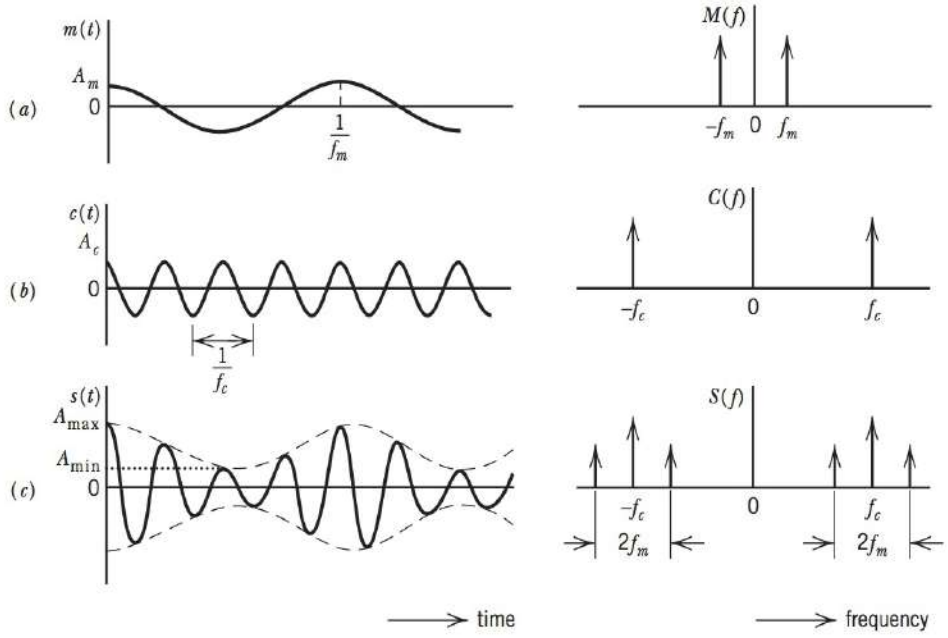


FIGURE 3.3 Illustrating the time-domain (on the left) and frequency-domain (on the right) characteristics of standard amplitude modulation produced by a single tone. (*a*) Modulating wave. (*b*) Carrier wave. (*c*) AM wave.

That is,

$$\mu = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}}$$

Expressing the product of the two cosines in Eq. (3.7) as the sum of two sinusoidal waves, one having frequency $f_c + f_m$ and the other having frequency $f_c - f_m$, we get

$$s(t) = A_c \cos(2\pi f_c t) + \tfrac{1}{2}\mu A_c \cos[2\pi(f_c + f_m)t] + \tfrac{1}{2}\mu A_c \cos[2\pi(f_c - f_m)t]$$

The Fourier transform of $s(t)$ is therefore

$$\begin{aligned} S(f) = {} & \tfrac{1}{2}A_c[\delta(f - f_c) + \delta(f + f_c)] \\ & + \tfrac{1}{4}\mu A_c[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)] \\ & + \tfrac{1}{4}\mu A_c[\delta(f - f_c + f_m) + \delta(f + f_c - f_m)] \end{aligned}$$

Thus the spectrum of an AM wave, for the special case of sinusoidal modulation, consists of delta functions at $\pm f_c$, $f_c \pm f_m$, and $-f_c \mp f_m$, as in Figure 3.3*c*.

In practice, the AM wave $s(t)$ is a voltage or current wave. In either case, the average power delivered to a 1-ohm resistor by $s(t)$ is comprised of three components:

$$\text{Carrier power} = \tfrac{1}{2}A_c^2$$

$$\text{Upper side-frequency power} = \tfrac{1}{8}\mu^2 A_c^2$$

$$\text{Lower side-frequency power} = \tfrac{1}{8}\mu^2 A_c^2$$

For a load resistor R different from 1 ohm, which is usually the case in practice, the expressions for carrier power, upper side-frequency power, and lower side-frequency power are merely scaled by the factor $1/R$ or R, depending on whether the modulated wave $s(t)$ is a voltage or a current, respectively. In any case, the ratio of the total sideband power to the total power in the modulated wave is equal to $\mu^2/(2 + \mu^2)$, which depends only on the modulation factor μ. If $\mu = 1$, that is, 100 percent modulation is used, the total power in the two side frequencies of the resulting AM wave is only one-third of the total power in the modulated wave.

Figure 3.4 shows the percentage of total power in both side frequencies and in the carrier plotted versus the percentage modulation. Note that when the percentage modulation is less than 20 percent, the power in one side frequency is less than 1 percent of the total power in the AM wave.

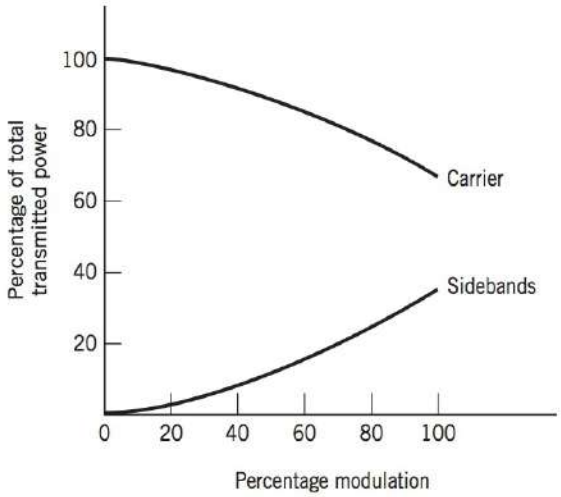


FIGURE 3.4 Variations of carrier power and total sideband power with percentage modulation.

SWITCHING MODULATOR

The generation of an AM wave may be accomplished using various devices; here we describe one such device called a *switching modulator*. Details of this modulator are shown in Figure 3.5*a*, where it is assumed that the carrier wave $c(t)$ applied to the diode is large in amplitude, so that it swings right across the characteristic curve of the diode. We assume that the diode acts as an *ideal switch*, that is, it presents zero impedance when it is forward-biased [corresponding to $c(t) > 0$]. We may thus approximate the transfer characteristic of the diode–load resistor combination by a *piecewise-linear characteristic*,

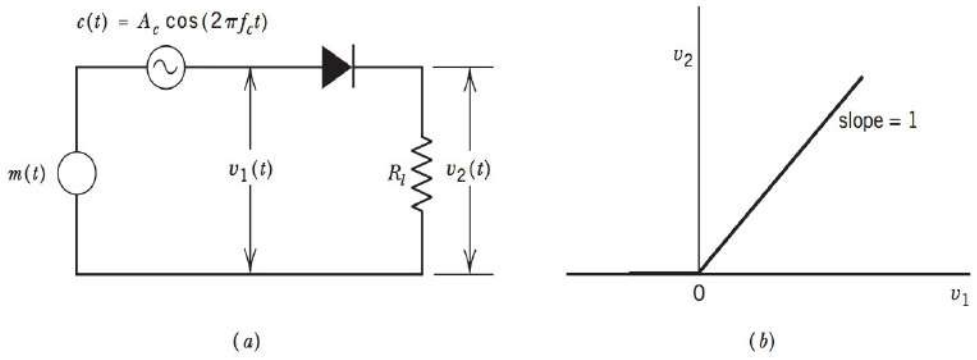


FIGURE 3.5 Switching modulator: (*a*) circuit diagram, (*b*) idealized input–output characteristic curve.

as shown in Figure 3.5*b*. Accordingly, for an input voltage $v_1(t)$ consisting of the sum of the carrier and the message signal:

$$v_1(t) = A_c \cos(2\pi f_c t) + m(t) \tag{3.8}$$

where $|m(t)| \ll A_c$, the resulting load voltage $v_2(t)$ is

$$v_2(t) \approx \begin{cases} v_1(t), & c(t) > 0 \\ 0, & c(t) < 0 \end{cases} \tag{3.9}$$

That is, the load voltage $v_2(t)$ varies periodically between the values $v_1(t)$ and zero at a rate equal to the carrier frequency f_c. In this way, by assuming a modulating wave that is weak compared with the carrier wave, we have effectively replaced the nonlinear behavior of the diode by an approximately equivalent piecewise-linear time-varying operation.

We may express Eq. (3.9) mathematically as

$$v_2(t) \simeq [A_c \cos(2\pi f_c t) + m(t)] g_{T_0}(t) \tag{3.10}$$

where $g_{T_0}(t)$ is a periodic pulse train of duty cycle equal to one-half, and period $T_0 = 1/f_c$, as in Figure 3.6. Representing this $g_{T_0}(t)$ by its Fourier series, we have

$$g_{T_0}(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos[2\pi f_c t(2n-1)] \tag{3.11}$$

Therefore, substituting Eq. (3.11) in (3.10), we find that the load voltage $v_2(t)$ consists of the sum of two components:

1. The component

$$\frac{A_c}{2}\left[1 + \frac{4}{\pi A_c} m(t)\right] \cos(2\pi f_c t)$$

which is the desired AM wave with amplitude sensitivity $k_a = 4/\pi A_c$. The switching modulator is therefore made more sensitive by reducing the carrier amplitude A_c; however, it must be maintained large enough to make the diode act like an ideal switch.

2. Unwanted components, the spectrum of which contains delta functions at 0, $\pm 2f_c$, $\pm 4f_c$, and so on, and which occupy frequency intervals of width $2W$ centered at 0, $\pm 3f_c$, $\pm 5f_c$, and so on, where W is the message bandwidth.

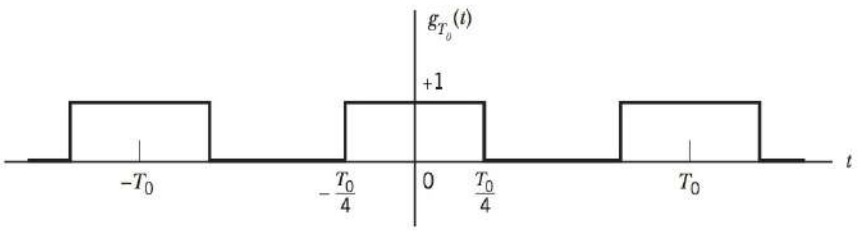


FIGURE 3.6 Periodic pulse train.

The unwanted terms are removed from the load voltage $v_2(t)$ by means of a band-pass filter with mid-band frequency f_c and bandwidth $2W$, provided that $f_c > 2W$. This latter condition ensures that the frequency separations between the desired AM wave and the unwanted components are large enough for the band-pass filter to suppress the unwanted components.

ENVELOPE DETECTOR

As mentioned in the introductory section, the process of *demodulation* is used to recover the original modulating wave from the incoming modulated wave; in effect, demodulation is the reverse of the modulation process. As with modulation, the demodulation of an AM wave can be accomplished using various devices; here, we describe a simple, yet highly effective device known as the *envelope detector*. Some version of this demodulator is used in almost all commercial AM radio receivers. For it to function properly, however, the AM wave has to be narrow-band, which requires that the carrier frequency be large compared to the message bandwidth. Moreover, the percentage modulation must be less than 100 percent.

An envelope detector of the series type is shown in Figure 3.7*a*, which consists of a diode and a resistor–capacitor (RC) filter. The operation of this envelope detector is as follows. On a positive half-cycle of the input signal, the diode is forward-biased and the capacitor C charges up rapidly to the peak value of the input signal. When the input signal falls below this value, the diode becomes reverse-biased and the capacitor C discharges slowly through the load resistor R_l. The discharging process continues until the next positive half-cycle. When the input signal becomes greater than the voltage across the capacitor, the diode conducts again and the process is repeated. We assume that the diode is ideal, presenting resistance r_f to current flow in the forward-biased region and infinite resistance in the reverse-biased region. We further assume that the AM wave applied to the envelope detector is supplied by a voltage source of internal impedance R_s. The charging time constant $(r_f + R_s)C$ must be short compared with the carrier period $1/f_c$, that is,

$$(r_f + R_s)C \ll \frac{1}{f_c} \tag{3.12}$$

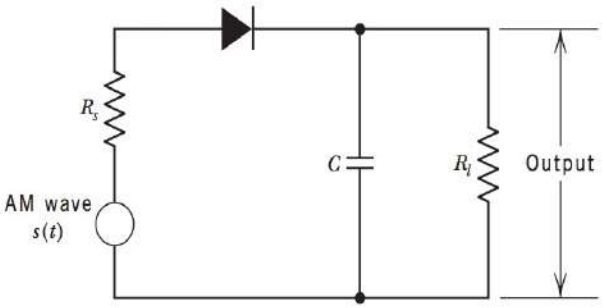
so that the capacitor C charges rapidly and thereby follows the applied voltage up to the positive peak when the diode is conducting. On the other hand, the discharging time constant R_lC must be long enough to ensure that the capacitor discharges slowly through the load resistor R_l between positive peaks of the carrier wave, but not so long that the capacitor voltage will not discharge at the maximum rate of change of the modulating wave, that is

$$\frac{1}{f_c} \ll R_lC \ll \frac{1}{W} \tag{3.13}$$

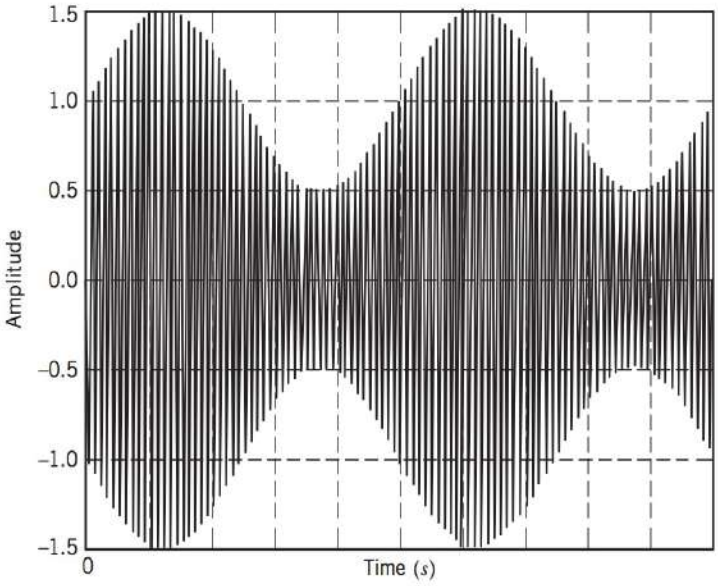
where W is the message bandwidth. The result is that the capacitor voltage detector output is nearly the same as the envelope of the AM wave, as demonstrated next.

EXAMPLE 3.2 Sinusoidal Amplitude Modulation

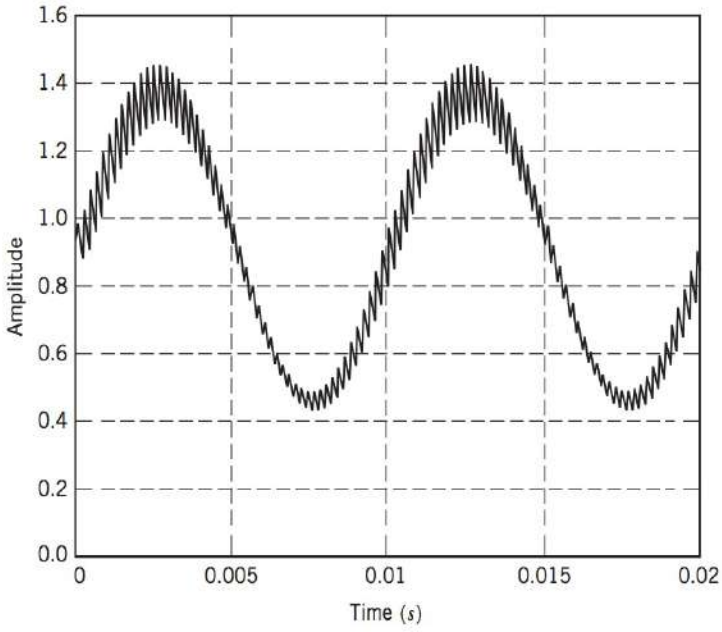
In Figure 3.7*b*, a sinusoidal AM wave is shown corresponding to 50 percent modulation. If this signal is applied to an envelope detector, the output shown in Figure 3.7*c* is obtained. This latter waveform is computed assuming the diode is ideal, having a constant resistance r_f when forward biased and infinite resistance when reverse-biased. Figure 3.7*c* shows that the envelope detector output contains a small amount of ripple at the carrier frequency; this ripple is easily removed by low-pass filtering. (See Problem 3.25.)



(*a*)



(*b*)



(*c*)

FIGURE 3.7 Envelope detector. (*a*) Circuit diagram. (*b*) AM wave input. (*c*) Envelope detector output.

VIRTUES, LIMITATIONS, AND MODIFICATIONS OF AMPLITUDE MODULATION

Reginald Fessenden (1866–1932)

Fessenden, a Canadian inventor, was trained in the classics, but took a position with the Edison Machine Works of New York to increase his knowledge of electricity. He became interested in radio after hearing of Marconi's success. Fessenden's initial radio work focused on improving detectors. In 1900, under a contract with the U.S. Weather Bureau, he became the first to transmit voice by radio. Because he used a spark-gap transmitter, the quality was too poor to be commercially useful.

Fessenden's subsequent research was sponsored by two wealthy Pittsburgh businessmen. This led to development of the rotary-spark transmitter and the first two-way transatlantic radio (telegraph) transmission in 1906. Fessenden was convinced that the spark-gap transmitter was not the best approach. In particular, he believed that a *continuous wave* transmitter would be better. In this vein, he let several contracts to General Electric to develop a high-speed alternator that would develop electrical signals at frequencies of 10 kHz (which turned out to be of limited use) and then 50 kHz. By using a carbon microphone in line with the source, Fessenden was able to *amplitude modulate* the signal, providing far better audio quality than with a spark-gap transmitter. In December 1906, Fessenden broadcast a music program using an alternator-transmitter, becoming the first to broadcast entertainment and music by radio. However, initial alternator transmitters were much less powerful than spark-gap transmitters, and these early demonstrations of radio were soon forgotten. It was not until the 1920s that radio broadcasts using amplitude modulation became commonplace.

In the early 1900s, Fessenden also developed the *heterodyne* principle whereby two signals are mixed (multiplied) to produce a third frequency. However, due to hardware limitations, it was over a decade before heterodyning was practical. Fessenden was a prolific inventor, accumulating over 500 patents in a wide variety of fields including microfilm, seismology, and sonar.

Amplitude modulation is the oldest method of performing modulation. Its biggest virtue is the ease with which it is generated and reversed. Modulation is accomplished rather simply in the transmitter using a switching modulator (described earlier) or a square-law modulator (described in Problem 3.4). Demodulation is accomplished just as easily in the receiver using an envelope detector (described earlier) or a square-law detector (described in Problem 3.6). The net result is that an amplitude modulating system is relatively cheap to build, which is the reason that AM radio broadcasting has been popular for so long and is quite likely to remain so well into the future.

However, from Chapter 1 we recall that transmitted power and channel bandwidth are our two primary communication resources and they should be used efficiently. In this context, we find that the standard form of amplitude modulation defined in Eq. (3.2) suffers from two major limitations:

1. *Amplitude modulation is wasteful of power.* The carrier wave $c(t)$ is completely independent of the information-bearing signal of baseband signal $m(t)$. The transmission of the carrier wave therefore represents a waste of power, which means that in amplitude modulation only a fraction of the total transmitted power is actually affected by $m(t)$.
2. *Amplitude modulation is wasteful of bandwidth.* The upper and lower sidebands of an AM wave are uniquely related to each other by virtue of their symmetry about the carrier frequency; hence, given the amplitude and phase spectra of either sideband, we can uniquely determine the other. This means that insofar as the transmission of information is concerned, only one sideband is necessary, and the communication channel therefore needs to provide only the same bandwidth as the baseband signal. In light of this observation, amplitude modulation is wasteful of bandwidth as it requires a transmission bandwidth equal to twice the message bandwidth.

To overcome these limitations, we must make certain changes, which result in increased system complexity of the amplitude modulation process. In effect, we trade off system complexity for improved utilization of communication resources. Starting with amplitude modulation as the standard, we can distinguish three modified forms of amplitude modulation:

1. *Double sideband–suppressed carrier (DSB-SC) modulation*, in which the transmitted wave consists of only the upper and lower sidebands. Transmitted power is saved here through the suppression of the carrier wave, but the channel bandwidth requirement is the same as before (i.e., twice the message bandwidth).
2. *Vestigial sideband (VSB) modulation*, in which one sideband is passed almost completely and just a trace, or *vestige*, of the other sideband is retained. The required channel bandwidth is therefore in excess of the message bandwidth by an amount equal to the width of the vestigial sideband. This form of modulation is well suited for the transmission of wideband signals such as television signals that contain significant components at extremely low frequencies. In commercial television broadcasting, a sizable carrier is transmitted together with the modulated wave, which makes it possible to demodulate the incoming modulated signal by an envelope detector in the receiver and thereby simplify the receiver design.

3. *Single sideband (SSB) modulation*, in which the modulated wave consists only of the upper sideband or the lower sideband. The essential function of SSB modulation is therefore to translate the spectrum of the modulating signal (with or without inversion) to a new location in the frequency domain. Single sideband modulation is particularly suited for the transmission of voice signals by virtue of the *energy gap* that exists in the spectrum of voice signals between zero and a few hundred hertz. It is an optimum form of modulation in that it requires the minimum transmitted power and minimum channel bandwidth: its principal disadvantage is increased cost and complexity.

In Section 3.3 we discuss DSB-SC modulation, followed by discussions of VSB and SSB forms of modulation in later sections.

3.3 DOUBLE SIDEBAND–SUPPRESSED CARRIER MODULATION

Basically, *double sideband–suppressed carrier (DSB-SC) modulation* consists of the product of the message signal $m(t)$ and the carrier wave $c(t)$, as follows:

$$\begin{aligned} s(t) &= c(t)\, m(t) \\ &= A_c \cos(2\pi f_c t) m(t) \end{aligned} \tag{3.14}$$

Consequently, the modulated signal $s(t)$ undergoes a *phase reversal* whenever the message signal $m(t)$ crosses zero, as indicated in Figure 3.8*b* for the message signal of Figure 3.8*a*. The envelope of a DSB-SC modulated signal is therefore different from the message signal.

From Eq. (3.14), the Fourier transform of $s(t)$ is obtained as

$$S(f) = \frac{1}{2} A_c [M(f - f_c) + M(f + f_c)] \tag{3.15}$$

For the case when the baseband signal $m(t)$ is limited to the interval $-W \le f \le W$, as in Figure 3.9*a*, we thus find that the spectrum $S(f)$ of the DSB-SC wave $s(t)$ is as illustrated in

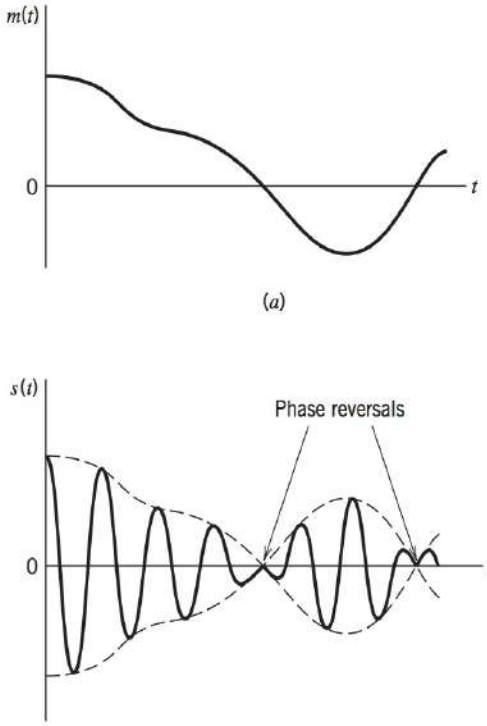


FIGURE 3.8 (*a*) Baseband signal. (*b*) DSB-SC modulated wave.

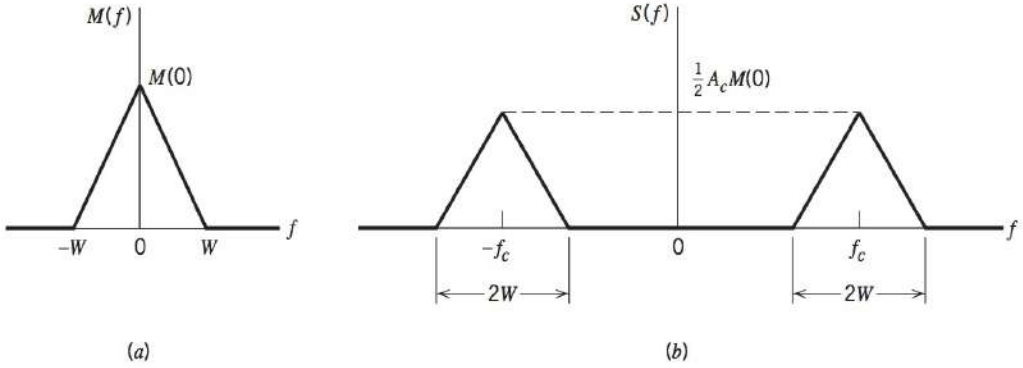


FIGURE 3.9 (*a*) Spectrum of baseband signal. (*b*) Spectrum of DSB-SC modulated wave.

Figure 3.9*b*. Except for a change in scale factor, the modulation process simply *translates* the spectrum of the baseband signal by $\pm f_c$. Of course, the transmission bandwidth required by DSB-SC modulation is the same as that for amplitude modulation, namely, $2W$.

RING MODULATOR

One of the most useful product modulators, well suited for generating a DSB-SC wave, is the ring modulator shown in Figure 3.10*a*. The four diodes form a ring in which they all point in the same way—hence the name. The diodes are controlled by a square-wave carrier $c(t)$ of frequency f_c, which is applied longitudinally by means of two center-tapped transformers. If the transformers are perfectly balanced and the diodes are identical, there is *no* leakage of the modulation (switching) frequency into the modulator output. To understand the operation of the circuit, we assume that the diodes have a constant forward resistance r_f when switched on and a constant backward resistance r_b when switched off and that they switch as the carrier wave $c(t)$ goes through zero. On one half-cycle of the carrier wave, the outer diodes are switched to their forward resistances r_f and the inner diodes are switched to their backward resistances r_b, as indicated in Figure 3.10*b*. On the other half-cycle of the carrier wave, the diodes operate in the opposite condition, as shown in Figure 3.10*c*. Typically, the terminating resistances at the input and output ends of the modulator are the same (assuming ideal 1 : 1 transformers). Under the conditions

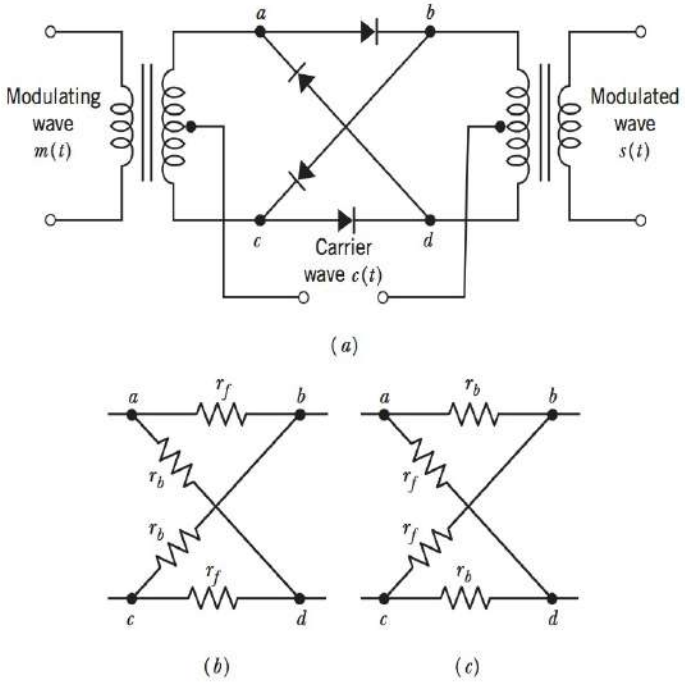


FIGURE 3.10 Ring modulator. (*a*) Circuit diagram. (*b*) Illustrating the condition when the outer diodes are switched on and the inner diodes are switched off. (*c*) Illustrating the condition when the outer diodes are switched off and the inner diodes are switched on.

described herein, it is a straightforward matter to show that the output voltage in Figure 3.10*b* has the same magnitude as the output voltage in Figure 3.10*c*, but they have opposite polarity. In effect, the ring modulator acts as a *commutator*.

Figure 3.11*c* shows the idealized waveform of the modulated signal $s(t)$ produced by the ring modulator, assuming a sinusoidal modulating wave $m(t)$ as in Figure 3.11*a* and a square carrier wave $c(t)$ as in Figure 3.11*b*. Now, the square-wave carrier $c(t)$ can be represented by a Fourier series as follows:

$$c(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos[2\pi f_c t(2n-1)] \tag{3.16}$$

The ring modulator output is therefore

$$\begin{aligned} s(t) &= c(t)m(t) \\ &= \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos[2\pi f_c t(2n-1)]m(t) \end{aligned} \tag{3.17}$$

We see that there is no output from the modulator at the carrier frequency; that is, the modulator output consists entirely of modulation products. The ring modulator is sometimes referred to as a *double-balanced modulator*, because it is balanced with respect to both the baseband signal and the square-wave carrier.

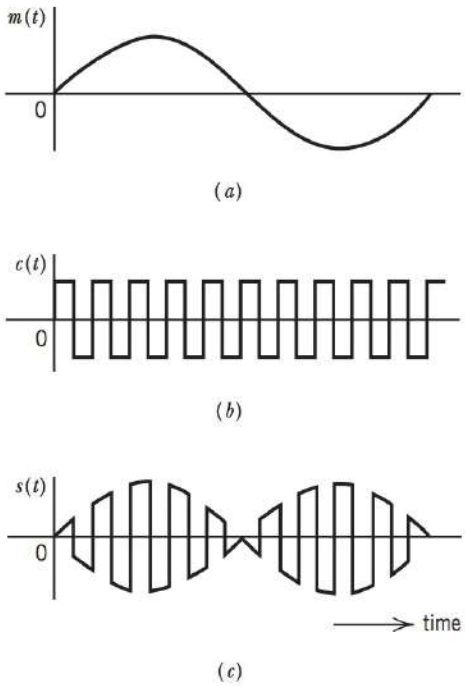


FIGURE 3.11 Waveforms illustrating the operation of the ring modulator for a sinusoidal modulating wave. (*a*) Modulating wave. (*b*) Square-wave carrier. (*c*) Modulated wave.

Assuming that $m(t)$ is limited to the frequency band $-W \le f \le W$, the spectrum of the modulator output consists of sidebands around each of the odd harmonics of the square-wave carrier $m(t)$, as illustrated in Figure 3.12. Here it is assumed that $f_c > W$ so as to prevent *sideband overlap*, which arises when sidebands belonging to the adjacent harmonic frequency f_c and $3f_c$ overlap each other. Thus, provided that we have $f_c > W$, we may use a band-pass filter of mid-band frequency f_c and bandwidth $2W$ to select the desired pair of sidebands around the carrier frequency f_c. Accordingly, the circuitry needed for the generation of a DSB-SC modulated wave consists of a ring modulator followed by a band-pass filter.

COHERENT DETECTION

The baseband signal $m(t)$ can be uniquely recovered from a DSB-SC wave $s(t)$ by first multiplying $s(t)$ with a locally generated sinusoidal wave and then low-pass filtering the product, as in Figure 3.13. It is assumed that the local oscillator signal is exactly coherent or synchronized, in both frequency and phase, with the carrier wave $c(t)$ used in the product modulator to generate $s(t)$. This method of demodulation is known as *coherent detection* or *synchronous demodulation*.

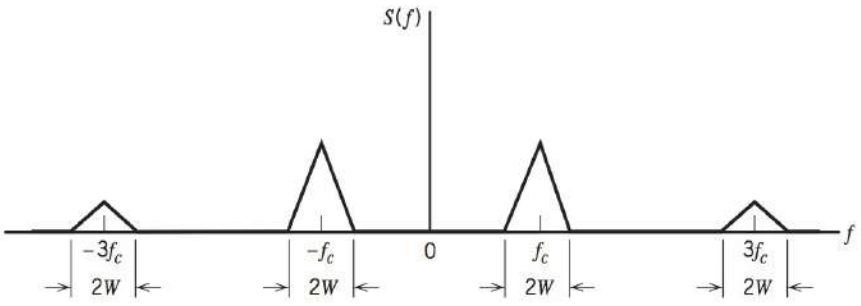


FIGURE 3.12 Illustrating the spectrum of ring modulator output.

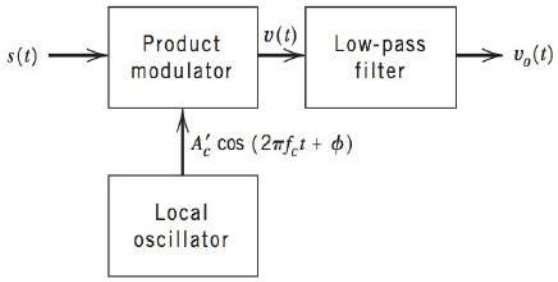


FIGURE 3.13 Coherent detection of DSB-SC modulated wave.

It is instructive to derive coherent detection as a special case of the more general demodulation process using a local oscillator signal of the same frequency but arbitrary phase difference ϕ, measured with respect to the carrier wave $c(t)$. Thus, denoting the local oscillator signal by $A_c' \cos(2\pi f_c t + \phi)$, and using Eq. (3.14) for the DSB-SC wave $s(t)$, we find that the product modulator output in Figure 3.13 is

$$\begin{aligned} v(t) &= A_c' \cos(2\pi f_c t + \phi) s(t) \\ &= A_c A_c' \cos(2\pi f_c t) \cos(2\pi f_c t + \phi) m(t) \\ &= \frac{1}{2} A_c A_c' \cos(4\pi f_c t + \phi) m(t) + \frac{1}{2} A_c A_c' \cos\phi \, m(t) \end{aligned} \tag{3.18}$$

The first term in Eq. (3.18) represents a DSB-SC modulated signal with a carrier frequency $2f_c$, whereas the second term is proportional to the baseband signal $m(t)$.This is further illustrated by the spectrum $V(f)$ shown in Figure 3.14, where it is assumed that the baseband signal $m(t)$ is limited to the interval $-W \leq f \leq W$. It is therefore apparent that the first term in Eq. (3.18) is removed by the low-pass filter in Figure 3.13, provided that the cut-off frequency of this filter is greater than W but less than $2f_c - W$. This is satisfied by choosing $f_c > W$. At the filter output we then obtain a signal given by

$$v_o(t) = \frac{1}{2} A_c A_c' \cos\phi \, m(t) \tag{3.19}$$

The demodulated signal $v_o(t)$ is therefore proportional to $m(t)$ when the phase error ϕ is a constant. The amplitude of this demodulated signal is maximum when $\phi = 0$, and it is minimum (zero) when $\phi = \pm\pi/2$. The zero demodulated signal, which occurs for $\phi = \pm\pi/2$, represents the *quadrature null effect* of the coherent detector. Thus the phase error ϕ in the local oscillator causes the detector output to be attenuated by a factor equal to $\cos\phi$. As long as the phase error ϕ is constant, the detector output provides an undistorted version of the original baseband signal $m(t)$. In practice, however, we usually find that the phase error ϕ varies randomly with time, due to random variations in the communication channel. The result is that at the detector output, the multiplying factor $\cos\phi$ also varies randomly with time, which is obviously undesirable. Therefore, provision must be made in the system to maintain the local oscillator in the receiver in perfect synchronism, in both frequency and phase, with the carrier wave used to generate the DSB-SC modulated signal in the transmitter. The resulting system complexity is the price that must be paid for suppressing the carrier wave to save transmitter power.

COSTAS RECEIVER

One method of obtaining a practical synchronous receiver system, suitable for demodulating DSB-SC waves, is to use the *Costas receiver* shown in Figure 3.15. This receiver

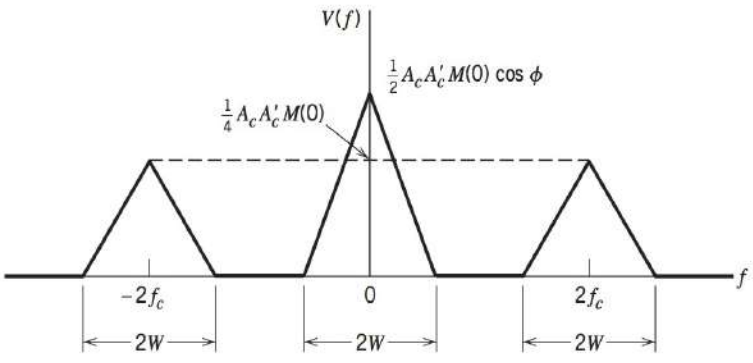


FIGURE 3.14 Illustrating the spectrum of a product modulator output with a DSB-SC modulated wave as input.

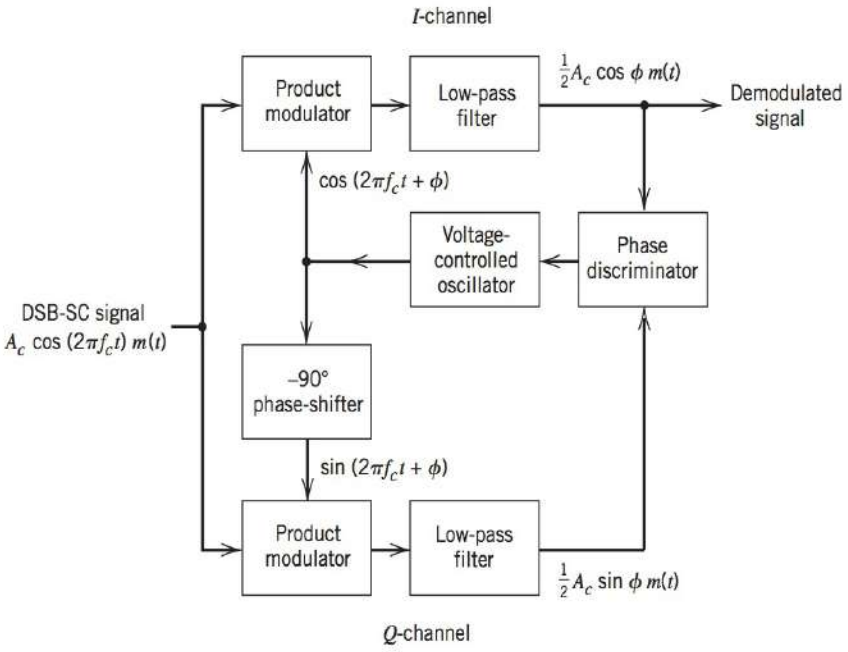


FIGURE 3.15 Costas receiver.

consists of two coherent detectors supplied with the same input signal, namely, the incoming DSB-SC wave $A_c \cos(2\pi f_c t)m(t)$, but with individual local oscillator signals that are in phase quadrature with respect to each other. The frequency of the local oscillator is adjusted to be the same as the carrier frequency f_c, which is assumed known *a priori*. The detector in the upper path is referred to as the *in-phase coherent detector* or *I-channel*, and that in the lower path is referred to as the *quadrature-phase coherent detector* or *Q-channel*. These two detectors are coupled together to form a *negative feedback* system designed in such a way as to maintain the local oscillator synchronous with the carrier wave.

To understand the operation of this receiver, suppose that the local oscillator signal is of the same phase as the carrier wave $A_c \cos(2\pi f_c t)$ used to generate the incoming DSB-SC wave. Under these conditions, we find that the *I*-channel output contains the desired demodulated signal $m(t)$, whereas the *Q*-channel output is zero due to the quadrature null effect of the *Q*-channel. Suppose next that the local oscillator phase drifts from its proper value by a small angle ϕ radians. The *I*-channel output will remain essentially unchanged, but there will now be some signal appearing at the *Q*-channel output, which is proportional to $\sin\phi \simeq \phi$ for small ϕ. This *Q*-channel output will have the same polarity as the *I*-channel output for one direction of local oscillator phase drift and opposite polarity for the opposite direction of local oscillator phase drift. Thus, by combining the *I*- and *Q*-channel outputs in a *phase discriminator* (which consists of a multiplier followed by a low-pass filter), as shown in Figure 3.15, a dc control signal is obtained that automatically corrects for local phase errors in the *voltage-controlled oscillator*.

It is apparent that phase control in the Costas receiver ceases with the modulation and that phase-lock has to be reestablished with the reappearance of modulation. This is not a serious problem when receiving voice transmission, because the lock-up process normally occurs so rapidly that no distortion is perceptible.

3.4 QUADRATURE-CARRIER MULTIPLEXING

The quadrature null effect of the coherent detector may also be put to good use in the construction of the so-called *quadrature-carrier multiplexing* or *quadrature-amplitude modulation* (QAM). This scheme enables two DSB-SC modulated waves (resulting from the application of two physically *independent* message signals) to occupy the same

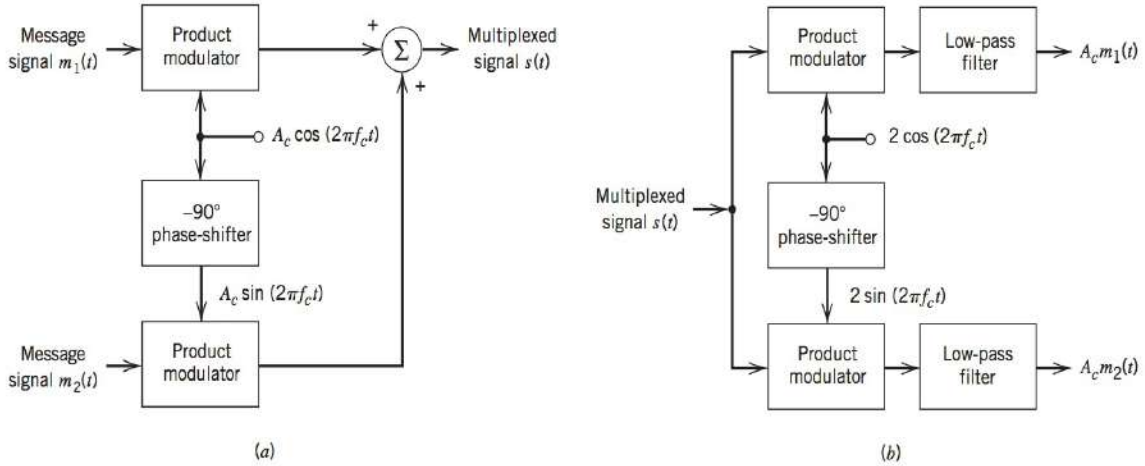


FIGURE 3.16 Quadrature-carrier multiplexing system. (*a*) Transmitter. (*b*) Receiver.

channel bandwidth, and yet it allows for the separation of the two message signals at the receiver output. It is therefore a *bandwidth-conservation scheme*.

A block diagram of the quadrature-carrier multiplexing system is shown in Figure 3.16. The transmitter part of the system, shown in Figure 3.16*a*, involves the use of two separate product modulators that are supplied with two carrier waves of the same frequency but differing in phase by −90 degrees. The transmitted signal $s(t)$ consists of the sum of these two product modulator outputs, as shown by

$$s(t) = A_c m_1(t)\cos(2\pi f_c t) + A_c m_2(t)\sin(2\pi f_c t) \tag{3.20}$$

where $m_1(t)$ and $m_2(t)$ denote the two different message signals applied to the product modulators. Thus $s(t)$ occupies a channel bandwidth of $2W$ centered at the carrier frequency f_c, where W is the message bandwidth of $m_1(t)$ or $m_2(t)$. According to Eq. (3.20), we may view $A_c m_1(t)$ as the in-phase component of the multiplexed band-pass signal $s(t)$ and $-A_c m_2(t)$ as its quadrature component.

The receiver part of the system is shown in Figure 3.16*b*. The multiplexed signal $s(t)$ is applied simultaneously to two separate coherent detectors that are supplied with two local carriers of the same frequency, but differing in phase by −90 degrees. The output of the top detector is $\frac{1}{2}A_c' m_1(t)$, whereas the output of the bottom detector is $\frac{1}{2}A_c' m_2(t)$. For the system to operate satisfactorily, it is important to maintain the correct phase and frequency relationships between the local oscillators used in the transmitter and receiver parts of the system.

To maintain this synchronization, we may use the Costas receiver described above. Another commonly used method is to send a *pilot signal* outside the passband of the modulated signal. In the latter method, the pilot signal typically consists of a low-power sinusoidal tone whose frequency and phase are related to the carrier wave $c(t)$; at the receiver, the pilot signal is extracted by means of a suitably tuned circuit and then translated to the correct frequency for use in the coherent detector.

3.5 SINGLE-SIDEBAND AND VESTIGIAL-SIDEBAND METHODS OF MODULATION

Our investigation of band-pass signals and quadrature amplitude modulation indicated that we could transmit two independent signals, each of baseband bandwidth W, in a band-pass channel of bandwidth $2W$. With double-sideband modulation, we are transmitting only one such signal and the question that comes to mind is whether the band-pass bandwidth of $2W$ is actually required. In actual fact, it can be shown that due to the symmetry of the DSB signal about the carrier frequency, the same information is transmitted in the *upper* and *lower sidebands*, and only one of the sidebands needs to be transmitted. In this

section, we discuss bandwidth conservation measures known as *single-sideband* (SSB) *modulation* and *vestigial sideband* (VSB) *modulation.*

SINGLE-SIDEBAND MODULATION

Conceptually, the generation of a SSB signal is straightforward. We first generate a double-sideband signal and then apply an ideal band-pass filter to the result with cutoff frequencies of f_c and $f_c + W$ for the upper sideband, for instance. Practically, the approximate construction of an ideal filter is very difficult.[1]

Where SSB finds its greatest application is in the transmission of an analog voice signal. Analog voice has very little energy at low frequencies (<300 Hz), that is, there is an energy gap in the spectrum near the origin as depicted in Figure 3.17*a*. In this case, the ideal SSB filter is shown in Figure 3.17*b*, resulting in the band-pass spectra shown in Figure 3.17*c*. In this situation, the design of the SSB filter does not need to be as stringent as depicted in Figure 3.17*b*. In particular, the filter must only satisfy the following requirements:

- The desired sideband lies inside the passband of the filter
- The unwanted sideband lies inside the stopband of the filter

This indicates that the filter's transition band, separating the passband from the stopband, is twice the lowest frequency component ($2f_a$) of the message signal. This nonzero transition bandwidth greatly simplifies the design of the SSB filter. The analysis of a SSB signal uses a technique known as the Hilbert transform[2]. Similar to a DSB signal, coherent demodulation is required to detect an SSB signal. The synchronization information required to perform coherent demodulation is often obtained by one of two methods:

- Transmitting a low power pilot carrier in addition to the selected sideband, or
- Using highly stable oscillators in both the transmitter and receiver for generating the carrier frequency.

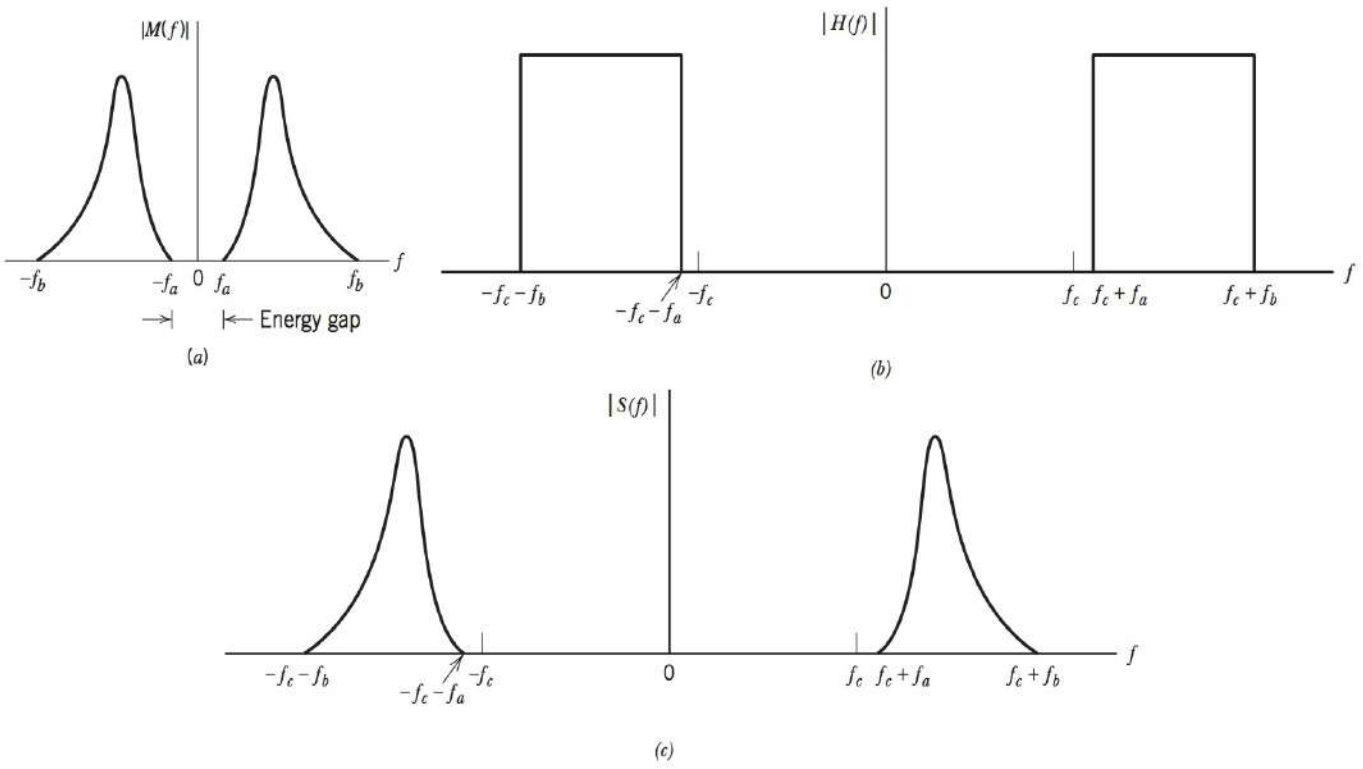


FIGURE 3.17 (a) Spectrum of a message signal $m(t)$ with an energy gap centered around the origin. (b) Idealized frequency response of band-pass filter. (c) Spectrum of SSB signal containing the upper sideband.

With the second approach, there is inevitably some phase error between the transmit and receive oscillators. This oscillator phase error leads to phase distortion of the demodulated signal, and gives rise to the so-called Donald Duck voice effect.

VSB MODULATION

The practical application of the idea of single-sideband transmission to signals that do not have an energy gap at the origin leads to *vestigial-sideband* (VSB) transmission. With VSB, all of the one sideband is transmitted and a small amount (vestige) of the other sideband is transmitted as well, as illustrated in Figure 3.18. With VSB, the filter is allowed to have a nonzero transition band but the question remains, what restrictions are placed on the filter, if any, to permit accurate recovery of the message signal? To answer this question, we consider the models shown in Figure 3.19 for generating the VSB signal and for coherently detecting it.

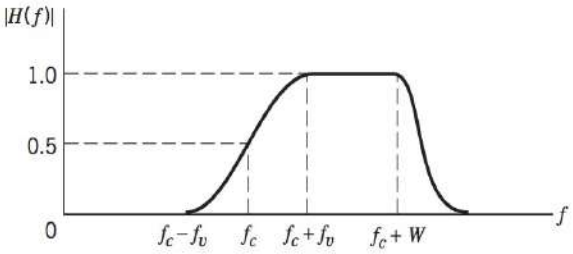


FIGURE 3.18 Amplitude response of VSB filter; only positive-frequency portion is shown.

Let $H(f)$ denote the transfer function of the filter following the product modulator of Figure 3.19*a*. The spectrum of the modulated signal $s(t)$ produced by passing the frequency-shifted signal $u(t)$ through the filter $H(f)$ is given by

$$\begin{aligned} S(f) &= U(f)H(f) \\ &= \frac{A_c}{2}[M(f-f_c) + M(f+f_c)]H(f) \end{aligned} \tag{3.21}$$

where $M(f)$ is the Fourier transform of the baseband signal $m(t)$ and $U(f)$ is the Fourier transform of $u(t)$. The problem we wish to address is to determine the particular $H(f)$ required to produce a modulated signal $s(t)$ with desired spectral characteristics, such that the original baseband signal $m(t)$ may be recovered from $s(t)$ by coherent detection.

The first step in the coherent detection process involves multiplying the modulated signal $s(t)$ by a locally generated sinusoidal wave $A_c' \cos(2\pi f_c t)$, which is synchronous with the carrier wave $A_c \cos(2\pi f_c t)$, in both frequency and phase as in Figure 3.19*b*. We may thus write

$$v(t) = A_c' \cos(2\pi f_c t) s(t)$$

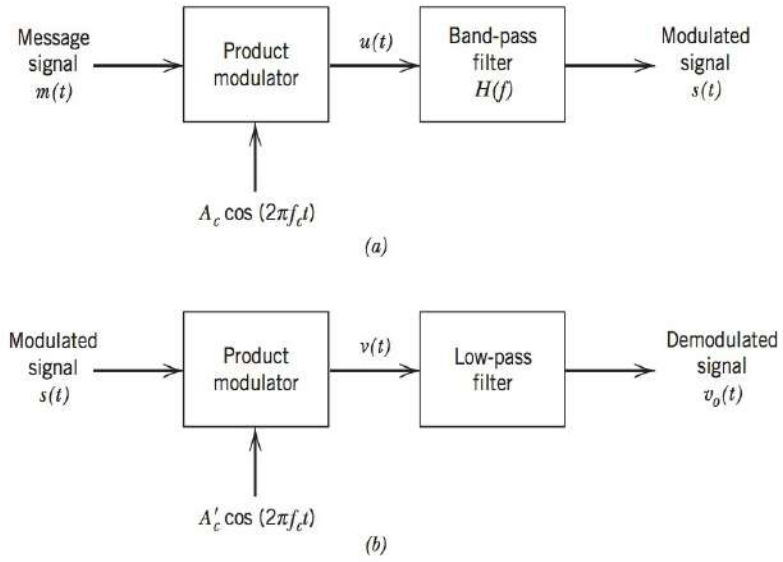


FIGURE 3.19 (*a*) Filtering scheme for processing sidebands. (*b*) Coherent detector for recovering the message signal.

Transforming this relation into the frequency domain gives the Fourier transform of $v(t)$ as

$$V(f) = \frac{A_c'}{2}[S(f - f_c) + S(f + f_c)] \tag{3.22}$$

Therefore, substitution of Eq. (3.21) in (3.22) yields

$$\begin{aligned} V(f) = {} & \frac{A_c A_c'}{4} M(f)[H(f - f_c) + H(f + f_c)] \\ & + \frac{A_c A_c'}{4}[M(f - 2f_c)H(f - f_c) + M(f + 2f_c)H(f + f_c)] \end{aligned} \tag{3.23}$$

The high-frequency components of $v(t)$ represented by the second term in Eq. (3.23) are removed by the low-pass filter in Figure 3.19*b* to produce an output $v_o(t)$, the spectrum of which is given by the remaining components:

$$V_o(f) = \frac{A_c A_c'}{4} M(f)[H(f - f_c) + H(f + f_c)] \tag{3.24}$$

For a distortionless reproduction of the original baseband signal $m(t)$ at the coherent detector output, we require $V_o(f)$ to be a scaled version of $M(f)$. This means, therefore, that the transfer function $H(f)$ must satisfy the condition

$$H(f - f_c) + H(f + f_c) = 2H(f_c) \tag{3.25}$$

where $H(f_c)$, the value of $H(f)$ at $f = f_c$, is a constant. When the baseband spectrum $M(f)$ is zero outside the frequency range $-W \le f \le W$, we need only satisfy Eq. (3.25) for values of f in this interval. Also, to simplify the exposition, we set $H(f_c) = 1/2$. We thus require that $H(f)$ satisfies the condition:

$$H(f - f_c) + H(f + f_c) = 1, \qquad -W \le f \le W \tag{3.26}$$

There is a great deal of flexibility in the selection of $H(f)$ to satisfy this condition, as discussed later in Section 3.6. In any event, under the condition described in Eq. (3.26), we find from Eq. (3.24) that the coherent detector output in Figure 3.19*b* is given by

$$v_o(t) = \frac{A_c A_c'}{2} m(t)$$

Thus, if the transfer function of filter satisfies Eq. (3.26), we can recover the original baseband signals without distortion.

Assuming that the requirement is to generate a *vestigial sideband* (VSB) modulated signal containing a vestige of the lower sideband, we find that Eq. (3.26) is satisfied by using a band-pass filter whose transfer function $H(f)$ is as shown in Figure 3.18; to simplify matters, only the response for positive frequencies is shown here. This frequency response is normalized, so that $|H(f)|$ is one-half at the carrier frequency f_c. The important feature to note, however, is that the cutoff portion of the frequency response around the carrier frequency f_c exhibits *odd symmetry*. That is to say, inside the transition interval $f_c - f_v \le |f| \le f_c + f_v$ the sum of the values of $|H(f)|$ at any two frequencies equally displaced above and below f_c is unity; f_v is the width of the vestigial sideband. Note also that outside the frequency band of interest (i.e., $|f| > f_c + W$), the transfer function $H(f)$ may have arbitrary specification.

Figure 3.18 applies to a VSB modulated signal containing a vestige of the lower sideband. For a VSB modulated signal containing a vestige of the upper sideband, we have similar results except for the following differences: The upper cutoff portion of $H(f)$ is controlled to exhibit odd symmetry around the carrier frequency f_c, whereas the lower cutoff portion is arbitrary.[3]

3.6 THEME EXAMPLE—VSB TRANSMISSION OF ANALOG AND DIGITAL TELEVISION

A discussion of vestigial sideband modulation would be incomplete without a mention of its role in commercial television broadcasting, both analog and digital.[4] The channel bandwidth used for TV broadcasting in North America is 6 MHz, as indicated in Figure 3.20. In the case of analog transmission, this channel bandwidth not only accommodates the bandwidth requirement of the VSB-modulated video signal but also provides for the accompanying sound signal that modulates a carrier of its own. The values presented on the frequency axis of Figure 3.20 pertain to a specific TV channel. According to this figure, for analog VSB transmission, the picture carrier frequency is at 55.25 MHz and the sound carrier is at 59.75 MHz. Note, however, that the information content of the TV signal lies in a baseband spectrum extending from 1.25 MHz below the picture carrier to 4.5 MHz above it.

The choice of the VSB modulation format for transmission of analog television was influenced by two factors:

1. The video signal exhibits a large bandwidth and significant low-frequency content, which suggests the use of vestigial sideband modulation.
2. The circuitry used for demodulation in the receiver should be simple and therefore cheap. This suggests the use of envelope detection, which requires the addition of a carrier to the VSB modulated wave.

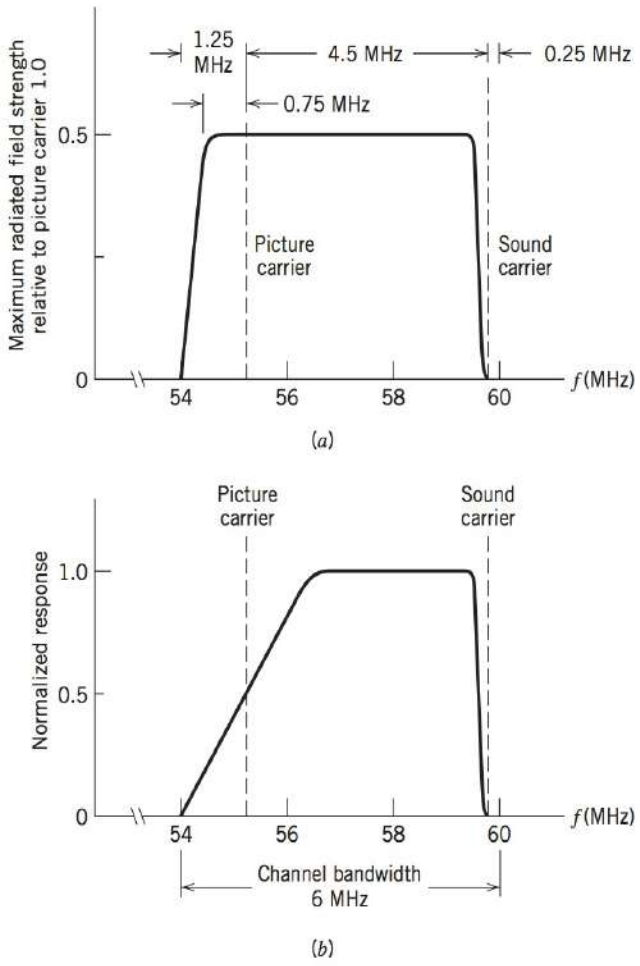


FIGURE 3.20 (*a*) Idealized amplitude spectrum of a transmitted TV signal, (*b*) Amplitude response of VSB shaping filter in the receiver.

With regard to point 1, it should be stressed that although there is indeed a basic desire to conserve bandwidth, in commercial TV broadcasting the transmitted signal is not quite VSB modulated. The reason is that at the transmitter the power levels are high, with the result that it would be expensive to rigidly control the filtering of sidebands. Instead, a VSB filter is inserted in each receiver where the power levels are low. The overall performance is the same as conventional vestigial-sideband modulation, except for some wasted power and bandwidth. These remarks are illustrated by Figure 3.20. In particular, Figure 3.20*a* shows the idealized spectrum of a transmitted TV signal. The upper sideband, 25 percent of the lower sideband, and the picture carrier are transmitted. The frequency response of the VSB filter used to do the required spectrum shaping in the receiver is shown in Figure 3.20*b*. With regard to point 2, the use of envelope detection (applied to a VSB modulated wave plus carrier) produces waveform distortion in the video signal recovered at the detector output.

Since we have shown that with VSB, the baseband signal can be faithfully recovered with the proper filtering, the VSB technique can be applied to digital signals as well as analog. With the evolution from analog to digital transmission of television signals in North America, a common factor is the continued use of VSB modulation. The choice of VSB modulation for transmission of digital *high-definition television signals* (HDTV) was influenced by two closely related factors:

1. The transmitted signal should be compatible in terms of bandwidth with the existing analog format. With modern digital techniques, data rates of 20 megabits per second (Mbps) and higher may be transmitted in a bandwidth of 6 MHz. These data rates are consistent with the requirements for the digital encoding of video signals.
2. The circuitry for demodulation in the receiver should be simple and relatively cheap. With the evolution of electronics, the complexity implied by the term "simple" has increased several orders of magnitude compared to the original analog television receivers. However, it is still an important design factor.

The spectrum for the digitally-modulated VSB signal is shown in Figure 3.21. Due to advances in technology, in particular, digital signal processing, the transmitted signal is the true VSB in this case. The spectrum has a shape known as the *raised cosine*, about which we shall have more to say in Chapter 8. With the digital approach, the sound, video, and color are all integrated as one data stream. The spectrum is shaped to extend 0.31 MHz below the carrier and 5.69 MHz above the carrier, which is shown dashed in Figure 3.21. Similar to analog television, a carrier component is added to the digital VSB signal at the frequency of 54.155 MHz but only at a low level. This carrier component simplifies the data detection and reduces the receiver cost.

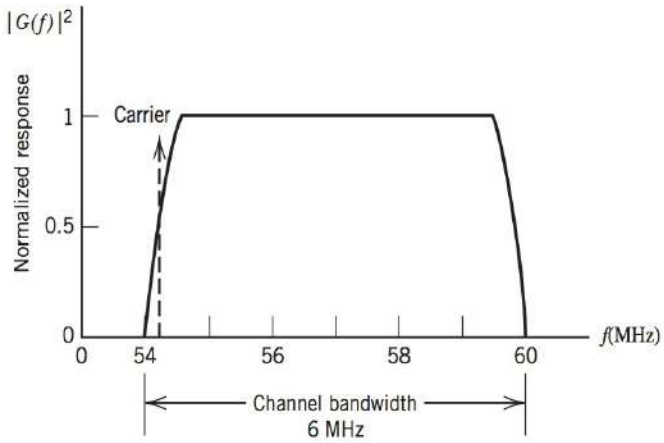


FIGURE 3.21 Idealized amplitude spectrum of VSB modulated digital television signal.

3.7 FREQUENCY TRANSLATION

The basic operation involved in single-sideband modulation is in fact a form of *frequency translation*, which is why single-sideband modulation is sometimes referred to as *frequency changing, mixing*, or *heterodyning*. This operation is clearly illustrated in the spectrum of the signal shown in Figure 3.17*c* compared to that of the original message signal in Figure 3.17*a*. Specifically, we see that a message spectrum occupying the band from f_a to f_b for positive frequencies in Figure 3.17*a* is shifted upward by an amount equal to the carrier frequency f_c in Figure 3.17*c*. The message spectrum for negative frequencies is translated downward in a symmetric fashion.

The idea of frequency translation described herein may be generalized as follows. Suppose that we have a modulated wave $s_1(t)$ whose spectrum is centered on a carrier frequency f_1, and the requirement is to translate it upward in frequency, such that its carrier frequency is changed from f_1 to a new value of f_2. This requirement may be accomplished using the *mixer* shown in Figure 3.22, which is similar to the scheme of Figure 3.19*a*. Specifically, the *mixer* is a device that consists of a product modulator followed by a band-pass filter. The band-pass filter is designed to have a bandwidth equal to that of the modulated signal $s_1(t)$ used as input. The key issue to be resolved is the frequency of the

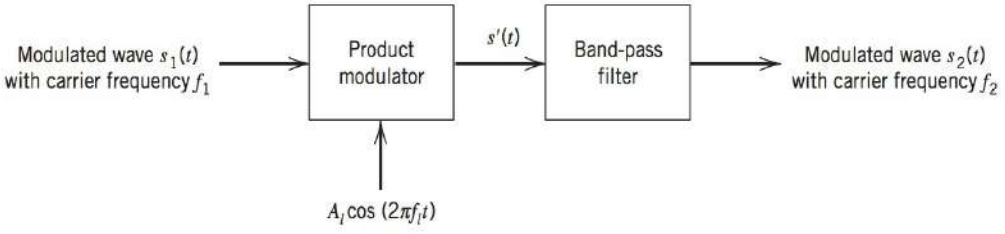


FIGURE 3.22 Block diagram of mixer.

local oscillator connected to the product modulator. Let f_l denote this frequency. Due to the frequency translation performed by the mixer, the carrier frequency f_1 of the incoming modulated signal is changed by an amount equal to f_l; hence, we may set

$$f_2 = f_1 + f_l$$

Solving for f_l, we thus have

$$f_l = f_2 - f_1$$

This relation assumes that $f_2 > f_1$, in which case the carrier frequency is translated *upward*. If, on the other hand, we have $f_1 > f_2$, the carrier frequency is translated *downward*, for which the corresponding frequency of the local oscillator is

$$f_l = f_1 - f_2$$

The reason for the band-pass filter is that the product modulator of Figure 3.23 produces two terms:

$$\begin{aligned} s_1(t) \times A_l \cos(2\pi f_l t) &= m(t)\cos(2\pi f_1 t) \times A_l \cos(2\pi f_l t) \\ &= \frac{1}{2} A_l m(t)[\cos(2\pi(f_1 + f_l)t) + \cos(2\pi(f_1 - f_l)t)] \end{aligned}$$

The band-pass filter rejects the unwanted frequency and keeps the desired one.

It is important to note that mixing is a linear operation. Accordingly, the relation of the sidebands of the incoming modulated wave to the carrier is completely preserved at the mixer output.

3.8 FREQUENCY-DIVISION MULTIPLEXING[5]

Another important signal processing operation is *multiplexing*, whereby a number of independent signals can be combined into a composite signal suitable for transmission over a common channel. Voice frequencies transmitted over telephone systems, for example, range from 300 to 3100 Hz. To transmit a number of these signals over the same channel, the signals must be kept apart so that they do not interfere with each other, and thus they

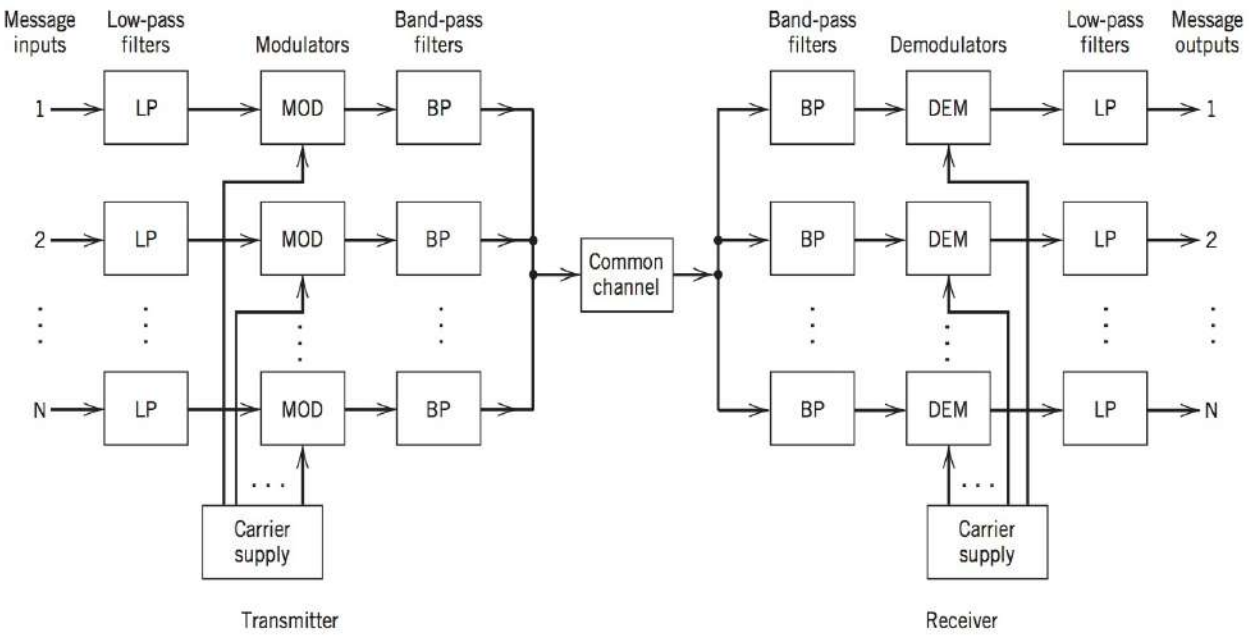


FIGURE 3.23 Block diagram of FDM system.

can be separated at the receiving end. This is accomplished by separating the signals either in frequency or in time. The technique of separating the signals in frequency is referred to as *frequency-division multiplexing* (FDM), whereas the technique of separating the signals in time is called *time-division multiplexing* (TDM). In this section, we discuss FDM systems, and TDM systems are discussed in Chapter 7.

A block diagram of an FDM system is shown in Figure 3.23. The incoming message signals are assumed to be of the low-pass type, but their spectra do not necessarily have nonzero values all the way down to zero frequency. Following each signal input, we have shown a low-pass filter, which is designed to remove high-frequency components that do not contribute significantly to signal representation but are capable of disturbing other message signals that share the common channel. These low-pass filters may be omitted only if the input signals are sufficiently band-limited initially. The filtered signals are applied to modulators that shift the frequency ranges of the signals so as to occupy mutually exclusive frequency intervals. The necessary carrier frequencies needed to perform these frequency translations are obained from a carrier supply. For the modulation, we may use any one of the methods described in this book. If the information source is analog voice, then a common method of modulation used with frequency-division multiplexing is single-sideband modulation. In this case, each voice input is usually assigned a bandwidth of 4 kHz. The band-pass filters following the modulators are used to restrict the band of each modulated wave to its prescribed range. The resulting band-pass filter outputs are next combined in parallel to form the input to the common channel. At the receiving terminal, a bank of band-pass filters, with their inputs connected in parallel, is used to separate the message signals on a frequency-occupancy basis. Finally, the original message signals are recovered by individual demodulators. Note that the FDM system shown in Figure 3.23 operates in only one direction. To provide for two-way transmission, as in telephony for example, we have to completely duplicate the multiplexing facilities, with the components connected in reverse order and with the signal waves proceeding from right to left.

3.9 SUMMARY AND DISCUSSION

In this chapter we studied the principles of amplitude modulation, which is one form of continuous-wave (CW) modulation, and important methods of generating and demodulating amplitude-modulated signals. Amplitude modulation uses a sinusoidal carrier whose amplitude is varied in accordance with the message signal.

Amplitude modulation may itself be classified into four types, depending on the spectral content of the modulated signal. The four types of amplitude modulation and their practical merits are as follows:

1. Standard amplitude modulation (AM) is the type in which the upper and lower sidebands are transmitted in full, accompanied by the carrier wave. Accordingly, demodulation of an AM signal is accomplished rather simply in the receiver by using an envelope detector, for example. It is for this reason we find that full AM is commonly used in commercial AM radio *broadcasting*, which involves a single powerful transmitter and numerous receivers that are relatively cheap to build.
2. Double sideband-suppressed carrier (DSB-SC) modulation is the type in which only the upper and lower sidebands are transmitted. The suppression of the carrier wave means that DSB-SC modulation requires much less power than standard AM to transmit the same message signal; this advantage of DSB-SC modulation over full AM is, however, attained at the expense of increased receiver complexity. DSB-SC modulation is therefore well suited for *point-to-point communication* involving one transmitter and

one receiver; in this form of communication, transmitted power is at a premium and the use of a complex receiver is therefore justifiable.

3. Single-sideband (SSB) modulation is the type in which only the upper sideband or lower sideband is transmitted. It is optimum in the sense that it requires the minimum transmitted power and the minimum channel bandwidth for conveying a message signal from one point to another.
4. Vestigial sideband modulation is the type in which "almost" the whole of one sideband and a "vestige" of the other sideband are transmitted in a prescribed complementary fashion. VSB modulation requires a channel bandwidth that is intermediate between that required for SSB and DSB-SC systems, and the saving in bandwidth can be significant if modulating signals with large bandwidths are being handled, as in the case of television signals and high-speed data.

DSB-SC, SSB, and VSB are examples of linear modulation in the sense that if $s_1(t)$ and $s_2(t)$ are two modulated signals corresponding to messages $m_1(t)$ and $m_2(t)$, then the sum $s_1(t) + s_2(t)$ corresponds to the message $m_1(t) + m_2(t)$. Ordinary AM does not have this property due to the presence of the carrier component.

We also introduced two important principles for detecting amplitude modulated signals. Envelope detection is an example of the first principle of *noncoherent detection* since it does not require that the receiver recover the phase of the sinusoidal carrier. Noncoherent detection has the advantage that it is simple to implement but has the disadvantage that it is wasteful of power. The second detection principle was *coherent demodulation*, whereby the receiver includes circuitry to recover the phase of transmitted carrier. This method may be applied in cases where the carrier is suppressed and when it is not. The advantage of coherent demodulation is that it can be applied to suppressed carrier modulations and thus can be more power efficient than noncoherent detection. Coherent demodulation has the disadvantage that it requires a more complex receiver design.

In the next chapter, we will consider a second family of continuous-wave modulation techniques, namely, angle modulation. We will find that angle modulation has properties significantly different from those of amplitude modulation.

NOTES AND REFERENCES

1. For a discussion of the filtering requirements in generating SSB modulated signals, see the paper by Kurth (1976).
2. See Haykin (2001) for a discussion of the Hilbert transform.
3. In Section 3.5 we described one method for the generation of a VSB signal. In an insightful article, Hill (1974) describes another time-domain method for the representation of VSB signals. Specifically, the VSB signal is expressed as the product of a narrow-band "envelope" function and an SSB signal.
4. For a collection of papers on television technology, see the book edited by Rzeszewski (1984). For a description of the VSB transmission of digital television system, see the paper by Challapali et al. (1995).
5. For a discussion of the performance of multiplex transmission, see Bennett (1970, pp. 213–218). For additional information on FDM systems, see "Transmission Systems for Communications," Bell Telephone Laboratories, pp. 128–137 (Western Electric, 1971).

PROBLEMS

3.1 A carrier wave of frequency 1 MHz is modulated 50 percent by a sinusoidal frequency 5 kHz. The resulting AM signal is transmitted through the resonant circuit of Figure P3.1, which is tuned to the carrier frequency and has a Q-factor of 175. Determine the modulated signal after transmission through this circuit. What is the percentage modulation of this modulated signal?

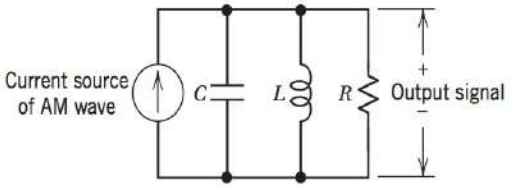


Figure P3.1

3.2 For a *p-n* junction diode, the current i through the diode and the voltage v across it are related by

$$i = I_0\left[\exp\left(-\frac{v}{V_T}\right) - 1\right]$$

where I_0 is the reverse saturation current, and V_T is the volt-equivalent of temperature defined by

$$V_T = \frac{kT}{e}$$

where k is Boltzmann's constant in joules per degree Kelvin, T is the absolute temperature in degrees Kelvin, and e is the charge of an electron. At room temperature, $V_T = 0.026$ volts.

(a) Expand i as a power series in v, retaining terms up to v^3.

(b) Let

$$v = 0.01\cos(2\pi f_m t) + 0.01\cos(2\pi f_c t) \text{ volts}$$

where $f_m = 1$ kHz and $f_c = 100$ kHz. Determine the spectrum of the resulting diode current i.

(c) Specify the band-pass filter required to extract from the diode current an AM signal with carrier frequency f_c.

(d) What is the percentage modulation of this AM signal?

3.3 Suppose that nonlinear devices are available for which the output current i_o and input voltage v_i are related by

$$i_o = a_1 v_1 + a_3 v_i^3$$

where a_1 and a_3 are constants. Explain how these devices may be used to provide: (a) a product modulator and (b) an amplitude modulator.

3.4 Figure P3.4 shows the circuit diagram of a *square-law modulator*. The signal applied to the nonlinear device is relatively weak, such that it can be represented by a square law:

$$v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$$

where a_1 and a_2 are constants, $v_1(t)$ is the input voltage, and $v_2(t)$ is the output voltage. The input voltage is defined by

$$v_1(t) = A_c\cos(2\pi f_c t) + m(t)$$

where $m(t)$ is a message signal and $A_c\cos(2\pi f_c t)$ is the carrier wave.

(a) Evaluate the output voltage $v_2(t)$.

(b) Specify the frequency response that the tuned circuit in Figure P3.4 must satisfy in order to generate an AM signal with f_c as the carrier frequency.

(c) What is the amplitude sensitivity of this AM signal?

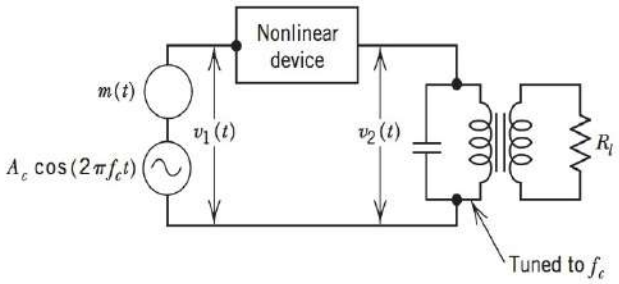


Figure P3.4

3.5 Consider the AM signal

$$s(t) = A_c[1 + \mu\cos(2\pi f_m t)]\cos(2\pi f_c t)$$

produced by a sinusoidal modulating signal of frequency f_m. Assume that the modulation factor is $\mu = 2$, and the carrier frequency f_c is much greater than f_m. The AM signal $s(t)$ is applied to an ideal envelope detector, producing the output $v(t)$.

(a) Determine the Fourier series representation of $v(t)$.

(b) What is the ratio of second-harmonic amplitude to fundamental amplitude in $v(t)$?

3.6 Consider a *square-law detector*, using a nonlinear device whose transfer characteristic is defined by

$$v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$$

where a_1 and a_2 are constants, $v_1(t)$ is the input, and $v_2(t)$ is the output. The input consists of the AM wave

$$v_1(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t)$$

(a) Evaluate the output $v_2(t)$.

(b) Find the conditions for which the message signal $m(t)$ may be recovered from $v_2(t)$.

3.7 The AM signal

$$s(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t)$$

is applied to the system shown in Figure P3.7. Assuming that $|k_a m(t)| < 1$ for all t and the message signal $m(t)$ is limited to the interval $-W \le f \le W$, and that the carrier frequency $f_c > 2W$, show that $m(t)$ can be obtained from the square-rooter output $v_3(t)$.

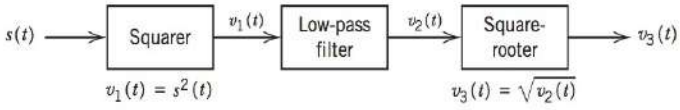


Figure P3.7

3.8 Consider a message signal $m(t)$ with the spectrum shown in Figure P3.8. The message bandwidth $W = 1$ kHz. This signal is applied to a product modulator, together with a carrier wave $A_c\cos(2\pi f_c t)$, producing the DSB-SC modulated signal $s(t)$. The modulated signal is next applied to a coherent detector. Assuming perfect synchronism between the carrier waves in the modulator and detector, determine the spectrum of the detector output when: (a) the carrier frequency $f_c = 1.25$ kHz and (b) the carrier frequency $f_c = 0.75$ kHz. What is the lowest carrier frequency for

which each component of the modulated signal $s(t)$ is uniquely determined by $m(t)$?

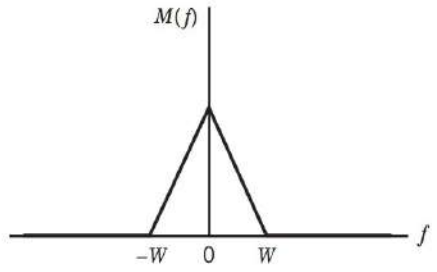


Figure P3.8

3.9 Figure P3.9 shows the circuit diagram of a *balanced modulator*. The input applied to the top AM modulator is $m(t)$, whereas that applied to the lower AM modulator is $-m(t)$; these two modulators have the same amplitude sensitivity. Show that the output $s(t)$ of the balanced modulator consists of a DSB-SC modulated signal.

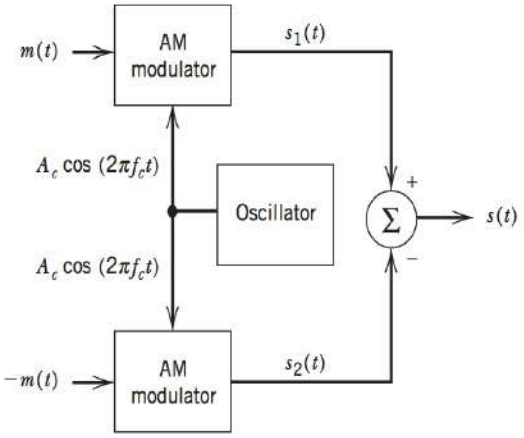


Figure P3.9

3.10 Figure 3.10 shows the circuit details of the ring modulator. Assume that the diodes are identical and the transformers are perfectly balanced. Let R denote the terminating resistance at the input end and output end of the modulator (assuming ideal 1 : 1 transformers). Determine the output voltage of the modulator for each of the two conditions described in Figures 3.10*b* and 3.10*c*. Hence, show that these two output voltages are equal in magnitude and opposite in polarity.

3.11 A DSB-SC modulated signal is demodulated by applying it to a coherent detector.

(a) Evaluate the effect of a frequency error Δf in the local carrier frequency of the detector, measured with respect to the carrier frequency of the incoming DSB-SC signal.

(b) For the case of a sinusoidal modulating wave, show that because of this frequency error, the demodulated signal exhibits *beats* at the error frequency. Illustrate your answer with a sketch of this demodulated signal.

3.12 Consider the DSB-SC signal

$$s(t) = A_c \cos(2\pi f_c t) m(t)$$

where $A_c \cos(2\pi f_c t)$ is the carrier wave and $m(t)$ is the message signal. This modulated signal is applied to a square-law device characterized by

$$y(t) = s^2(t)$$

The output $y(t)$ is next applied to a narrow-band filter with a passband amplitude response of one, mid-band frequency $2f_c$, and bandwidth Δf. Assume that Δf is small enough to treat the spectrum of $y(t)$ as essentially constant inside the passband of the filter.

(a) Determine the spectrum of the square-law device output $y(t)$.

(b) Show that the filter output $v(t)$ is approximately sinusoidal, given by

$$v(t) \simeq \frac{A_c^2}{2} E \Delta f \cos(4\pi f_c t)$$

where E is the energy of the message signal $m(t)$.

3.13 Consider the quadrature–carrier multiplex system of Figure 3.16. The multiplexed signal $s(t)$ produced at the transmitter output in Figure 3.16*a* is applied to a communication channel of transfer function $H(f)$. The output of this channel is in turn applied to the receiver input in Figure 3.16*b*. Prove that the condition

$$H(f_c + f) = H^*(f_c - f), \qquad 0 \le f \le W$$

is necessary for recovery of the message signals $m_1(t)$ and $m_2(t)$ at the receiver outputs; f_c is the carrier frequency, and W is the message bandwidth. *Hint:* Evaluate the spectra of the two receiver outputs.

3.14 Suppose that in the receiver of the quadrature–carrier multiplex system of Figure 3.16 the local carrier available for demodulation has a phase error ϕ with respect to the carrier source used in the transmitter. Assuming a distortionless communication channel between transmitter and receiver, show that this phase error will cause *cross-talk* to arise between the two demodulated signals at the receiver outputs. By cross-talk we mean that a portion of one message signal appears at the receiver output belonging to the other message signal, and vice versa.

3.15 A particular version of *AM stereo* uses quadrature multiplexing. Specifically, the carrier $A_c \cos(2\pi f_c t)$ is used to modulate the sum signal

$$m_1(t) = V_0 + m_l(t) + m_r(t)$$

where V_0 is a dc offset included for the purpose of transmitting the carrier component, m_ℓ is the left-hand audio signal, and $m_r(t)$ is the right-hand audio signal. The quadrature carrier $A_c \sin(2\pi f_c t)$ is used to modulate the difference signal

$$m_2(t) = m_l(t) - m_r(t)$$

(a) Show that an envelope detector may be used to recover the sum $m_r(t) + m_l(t)$ from the quadrature-multiplexed signal. How would you minimize the signal distortion produced by the envelope detector?

(b) Show that a coherent detector can recover the difference $m_l(t) - m_r(t)$.

(c) How are the desired $m_l(t)$ and $m_r(t)$ finally obtained?

3.16 The single tone modulating signal $m(t) = A_m \cos(2\pi f_m t)$ is used to generate the VSB signal

$$s(t) = \frac{1}{2}aA_mA_c \cos[2\pi(f_c + f_m)t] + \frac{1}{2}A_mA_c(1-a)\cos[2\pi(f_c - f_m)t]$$

where a is a constant, less than unity, representing the attenuation of the upper side frequency.

(a) If we represent this VSB signal as a quadrature carrier multiplex

$$s(t) = A_c m_1(t)\cos(2\pi f_c t) + A_c m_2(t)\sin(2\pi f_c t)$$

What is $m_2(t)$?

(b) The VSB signal, plus the carrier $A_c \cos(2\pi f_c t)$, is passed through an envelope detector. Determine the distortion produced by the quadrature component, $m_2(t)$.

(c) What is the value of constant a for which this distortion reaches its worst possible condition?

3.17 Using the message signal

$$m(t) = \frac{1}{1+t^2}$$

determine and sketch the modulated waves for the following methods of modulation:

(a) Amplitude modulation with 50 percent modulation.

(b) Double sideband-suppressed carrier modulation.

3.18 The local oscillator used for the demodulation of an SSB signal $s(t)$ has a frequency error Δf measured with respect to the carrier frequency f_c used to generate $s(t)$. Otherwise, there is perfect synchronism between this oscillator in the receiver and the oscillator supplying the carrier wave in the transmitter. Evaluate the demodulated signal for the following two situations:

(a) The SSB signal $s(t)$ consists of the upper sideband only.

(b) The SSB signal $s(t)$ consists of the lower sideband only.

3.19 Figure P3.19 shows the block diagram of *Weaver's method* for generating SSB modulated waves. The message (modulating) signal $m(t)$ is limited to the frequency band $f_a \le |f| \le f_b$. The auxiliary carrier applied to the first pair of product modulators has a frequency f_0, which lies at the center of this band, as shown by

$$f_0 = \frac{f_a + f_b}{2}$$

The low-pass filters in the upper and lower branches are identical, each with a cutoff frequency equal to $(f_b - f_a)/2$. The carrier applied to the second pair of product modulators has a frequency f_c that is greater than $(f_b - f_a)/2$. Sketch the spectra at the various points in the modulator of Figure P3.19, and hence show that:

(a) For the lower sideband, the contributions of the upper and lower branches are of opposite polarity, and by adding them at the modulator output, the lower sideband is suppressed.

(b) For the upper sideband, the contributions of the upper and lower branches are of the same polarity, and by adding them, the upper sideband is transmitted.

(c) How would you modify the modulator of Figure P3.19, so that only the lower sideband is transmitted?

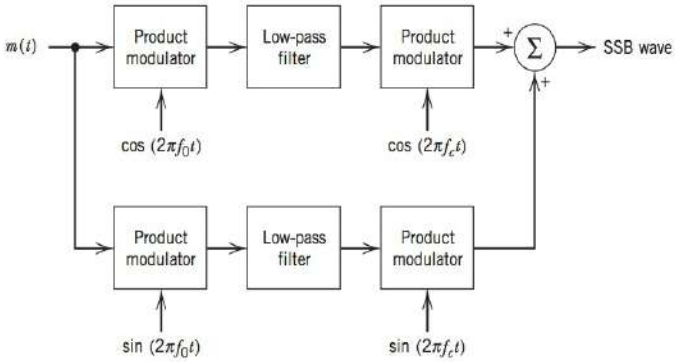


Figure P3.19

3.20

(a) Consider a message signal $m(t)$ containing frequency components at 100, 200, and 400 Hz. This signal is applied to an SSB modulator together with a carrier at 100 kHz, with only the upper sideband retained. In the coherent detector used to recover $m(t)$, the local oscillator supplies a sine wave of frequency 100.02 kHz. Determine the frequency components of the detector output.

(b) Repeat your analysis, assuming that only the lower sideband is transmitted.

3.21 The spectrum of a voice signal $m(t)$ is zero outside the interval $f_a \le |f| \le f_b$. In order to ensure communication privacy, this signal is applied to a *scrambler* that consists of the following cascade of components: a product modulator, a high-pass filter, a second product modulator, and a low-pass filter. The carrier wave applied to the first product modulator has a frequency equal to f_c, whereas that applied to the second product modulator has a frequency equal to $f_b + f_c$; both of them have unit amplitude. The high-pass and low-pass filters have the same cutoff frequency at f_c. Assume that $f_c > f_b$.

(a) Derive an expression for the scrambler output $s(t)$, and sketch its spectrum.

(b) Show that the original voice signal $m(t)$ may be recovered from $s(t)$ by using an *unscrambler* that is identical to the unit described above.

3.22 A method that is used for carrier recovery in SSB modulation systems involves transmitting two pilot frequencies that are appropriately positioned with respect to the transmitted sideband. This is illustrated in Figure P3.22*a* for the case when only the lower sideband is transmitted. In this case, the pilot frequencies f_1 and f_2 are defined by

$$f_1 = f_c - W - \Delta f$$

and

$$f_2 = f_c + \Delta f$$

where f_c is the carrier frequency and W is the message bandwidth. The Δf is chosen so as to satisfy the relation

$$n = \frac{W}{\Delta f}$$

where n is an integer. Carrier recovery is accomplished by using the scheme shown in Figure P3.22*b*. The outputs of the two narrow-band filters centered at f_1 and f_2 are defined by, respectively,

$$v_1(t) = A_1 \cos(2\pi f_1 t + \phi_1)$$

and

$$v_2(t) = A_2 \cos(2\pi f_2 t + \phi_2)$$

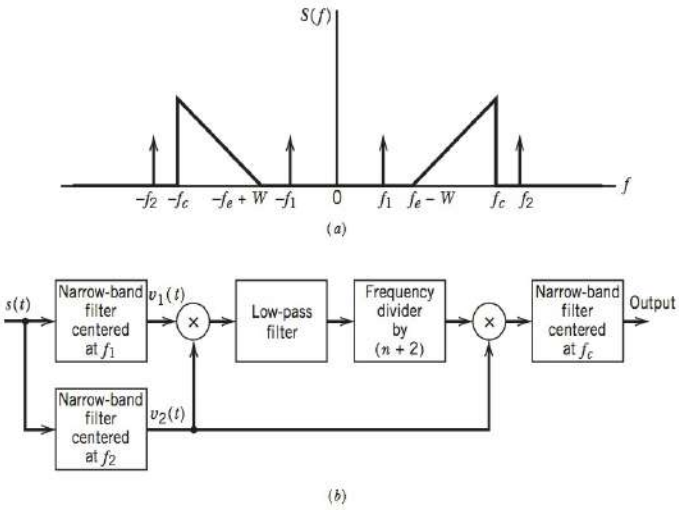


Figure P3.22

The low-pass filter is designed to select the difference frequency component of the first multiplier output due to $v_1(t)$ and $v_2(t)$.

(a) Show that the output signal of the circuit in Figure P3.22*b* is proportional to the carrier wave $A_c \cos(2\pi f_c t)$ if the phase angles ϕ_1 and ϕ_2 satisfy the relation

$$\phi_2 = -\frac{\phi_1}{1+n}$$

(b) For the case when only the upper sideband is transmitted, the two pilot frequencies are defined by

$$f_1 = f_c - \Delta f$$

and

$$f_2 = f_c + W + \Delta f$$

How would you modify the carrier recovery circuit of Figure P3.22*b* in order to deal with this case? What is the corresponding relation between ϕ_1 and ϕ_2 for the circuit output to be proportional to the carrier wave?

3.23 Figure P3.23 shows the block diagram of a *frequency synthesizer*, which enables the generation of many frequencies, each with the same high accuracy as the *master oscillator*. The master oscillator of frequency 1 MHz feeds two *spectrum generators*, one directly and the other through a *frequency divider*. Spectrum generator 1 produces a signal rich in the following harmonics: 1, 2, 3, 4, 5, 6, 7, 8, and 9 MHz. The frequency divider provides a 100-kHz output, in response to which spectrum generator 2 produces a second signal rich in the following harmonics: 100, 200, 300, 400, 500, 600, 700, 800, and 900 kHz. The harmonic selectors are designed to feed two signals into the mixer, one from spectrum generator 1 and the other from spectrum generator 2. Find the range of possible frequency outputs of this synthesizer and its resolution.

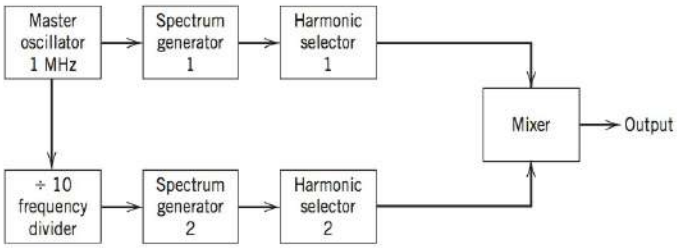


Figure P3.23

3.24 Consider a multiplex system in which four input signals $m_1(t)$, $m_2(t)$, $m_3(t)$, and $m_4(t)$, are respectively multiplied by the carrier waves

$$[\cos(2\pi f_a t) + \cos(2\pi f_b t)]$$
$$[\cos(2\pi f_a t + \alpha_1) + \cos(2\pi f_b t + \beta_1)]$$
$$[\cos(2\pi f_a t + \alpha_2) + \cos(2\pi f_b t + \beta_2)]$$
$$[\cos(2\pi f_a t + \beta_3) + \cos(2\pi f_b t + \beta_3)]$$

and the resulting DSB-SC signals are summed and then transmitted over a common channel. In the receiver, demodulation is achieved by multiplying the sum of the DSB-SC signals by the four carrier waves separately and then using filtering to remove the unwanted components.

(a) Determine the conditions that the phase angles α_1, α_2, α_3 and β_1, β_2, β_3 must satisfy in order that the output of the kth demodulator is $m_k(t)$, where $k = 1, 2, 3, 4$.

(b) Determine the minimum separation of carrier frequencies f_a and f_b in relation to the bandwidth of the input signals so as to ensure a satisfactory operation of the system.

Computer Problems

3.25 In this computer experiment, we simulate the modulation and demodulation of an AM wave.

(a) Develop a Matlab script to simulate the modulation of a 20 kHz carrier with a 0.4 kHz modulating wave. Use a modulation index of 50 percent and a sampling rate of 160 kHz.

(b) An envelope detector is assumed to have a forward resistance r_f of 25 Ω and a capacitance of 0.01 μF. The source resistance is 75 Ω and the load resistance is 10 kΩ.

(i) What is the charging time constant? Compare this time constant to one period of the modulating wave, and comment on how well it should track the envelope.

(ii) What is the capacitor discharge time constant? Compare this to one period of the carrier wave. Using a linear approximation, what fraction of the capacitor voltage decays during one sample period?

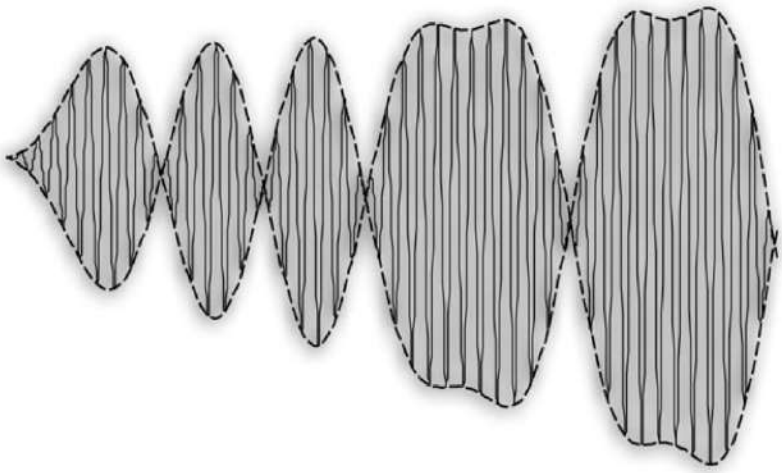
(iii) Based on these results, justify the following sampled model of an envelope detector.

```
Vc(1) = 0;                % initial capacitor voltage
for i = 2: length(s)
  if s(i) > Vc(i − 1) % diode on
    Vc(i) = s(i);
  else                    % diode off
    Vc(i) = Vc(i − 1) − 0.023*Vc(i − 1);
  end
end
```

(iv) Apply the envelope detector to the modulated signal.

(c) The output of the envelope detector is applied to a unit-gain RC filter with time constant 1 ms. Develop a discrete-time truncated impulse response model of this filter. Apply this filter model to the envelope detector output. Comment on the results.

4



ANGLE MODULATION

4.1 INTRODUCTION

In the previous chapter, we investigated the effect of slowly varying the amplitude of a sinusoidal carrier wave in accordance with the baseband (information-carrying) signal. In this chapter, we study a second family of continuous-wave (CW) modulation systems, namely, *angle modulation*, in which the angle of the carrier wave is varied according to the baseband signal. In this method of modulation, the amplitude of the carrier wave is maintained constant. There are two common forms of angle modulation, namely, *phase modulation and frequency modulation*. An important feature of angle modulation is that it can provide better discrimination against noise and interference than amplitude modulation. As will be shown in Chapter 6, however, this improvement in performance is achieved at the expense of increased transmission bandwidth. That is, angle modulation provides us with a practical means of exchanging channel bandwidth for improved noise performance. Such a trade-off is not possible with amplitude modulation. Moreover, the improvement in noise performance with angle modulation is achieved at the expense of increased system complexity in both the transmitter and receiver.

4.2 BASIC DEFINITIONS

Let $\theta_i(t)$ denote the *angle* of a modulated sinusoidal carrier at time t; it is assumed to be a function of the information-bearing signal or message signal. We express the resulting *angle-modulated wave* as

$$s(t) = A_c \cos[\theta_i(t)] \tag{4.1}$$

where A_c is the carrier amplitude. A complete oscillation occurs whenever $\theta_i(t)$ changes by 2π radians. If $\theta_i(t)$ increases monotonically with time, the average frequency in Hertz over an interval from t to $t + \Delta t$ is given by

$$f_{\Delta t}(t) = \frac{\theta_i(t + \Delta t) - \theta_i(t)}{2\pi \Delta t} \tag{4.2}$$

We may thus define the *instantaneous frequency* of the angle-modulated signal $s(t)$ as follows:

$$\begin{aligned} f_i(t) &= \lim_{\Delta t \to 0} f_{\Delta t}(t) \\ &= \lim_{\Delta t \to 0} \left[\frac{\theta_i(t + \Delta t) - \theta_i(t)}{2\pi \Delta t} \right] \\ &= \frac{1}{2\pi} \frac{d\theta_i(t)}{dt} \end{aligned} \tag{4.3}$$

where, in the last line, we have invoked the definition for the derivative of the angle with respect to time t.

Thus, according to Eq. (4.1), we may interpret the angle-modulated signal $s(t)$ as a rotating phasor of length A_c and angle $\theta_i(t)$. The angular velocity of such a phasor is $d\theta_i(t)/dt$, measured in radians per second in accordance with Eq. (4.3). In the simple case of an unmodulated carrier, the angle $\theta_i(t)$ is

$$\theta_i(t) = 2\pi f_c t + \phi_c$$

and the corresponding phasor rotates with a constant angular velocity equal to $2\pi f_c$. The constant ϕ_c is the value of $\theta_i(t)$ at $t = 0$.

There are an infinite number of ways in which the angle $\theta_i(t)$ may be varied in some manner with the message (baseband) signal. However, we shall consider only two commonly used methods, phase modulation and frequency modulation, as defined below:

1. *Phase modulation* (PM) *is that form of angle modulation in which the instantaneous angle* $\theta_i(t)$ is varied linearly with the message signal as shown by

$$\theta_i(t) = 2\pi f_c t + k_p m(t) \tag{4.4}$$

The term $2\pi f_c t$ represents the *angle of the unmodulated* carrier; the constant k_p represents the *phase sensitivity* of the modulator, expressed in radians per volt on the assumption that $m(t)$ is a voltage waveform. For convenience, we have assumed in Eq. (4.4) that the angle of the unmodulated carrier is zero at $t = 0$. The phase-modulated signal $s(t)$ is thus described in the time domain by

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)] \tag{4.5}$$

2. *Frequency modulation* (FM) *is that form of angle modulation in which the instantaneous frequency* $f_i(t)$ *is varied linearly with the message signal* $m(t)$*, as shown by*

$$f_i(t) = f_c + k_f m(t) \tag{4.6}$$

The term f_c represents the frequency of the unmodulated carrier, and the constant k_f represents the *frequency sensitivity* of the modulator, expressed in Hertz per volt on the assumption that $m(t)$ is a voltage waveform. Integrating Eq. (4.6) with respect to time and multiplying the result by 2π, we get

$$\theta_i(t) = 2\pi f_c t + 2\pi k_f \int_0^t m(\tau)\, d\tau \tag{4.7}$$

where, for convenience, we have assumed that the angle of the unmodulated carrier wave is zero at $t = 0$. The frequency-modulated signal is therefore described in the time domain by

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau)\, d\tau\right] \tag{4.8}$$

PROPERTIES OF ANGLE-MODULATED WAVES

Angle-modulated waves are characterized by some important properties that put angle-modulated waves in a family of their own, and distinguish them from the family of amplitude modulated waves, as illustrated by the waveforms shown in Figure 4.1 for the case of sinusoidal modulation. Figures 4.1*a* and 4.1*b* are the sinusoidal carrier and modulating waves, respectively. Figures 4.1*c*, 4.1*d*, and 4.1*e* display the corresponding amplitude-modulated (AM), phase-modulated (PM), and frequency-modulated (FM) waves, respectively.

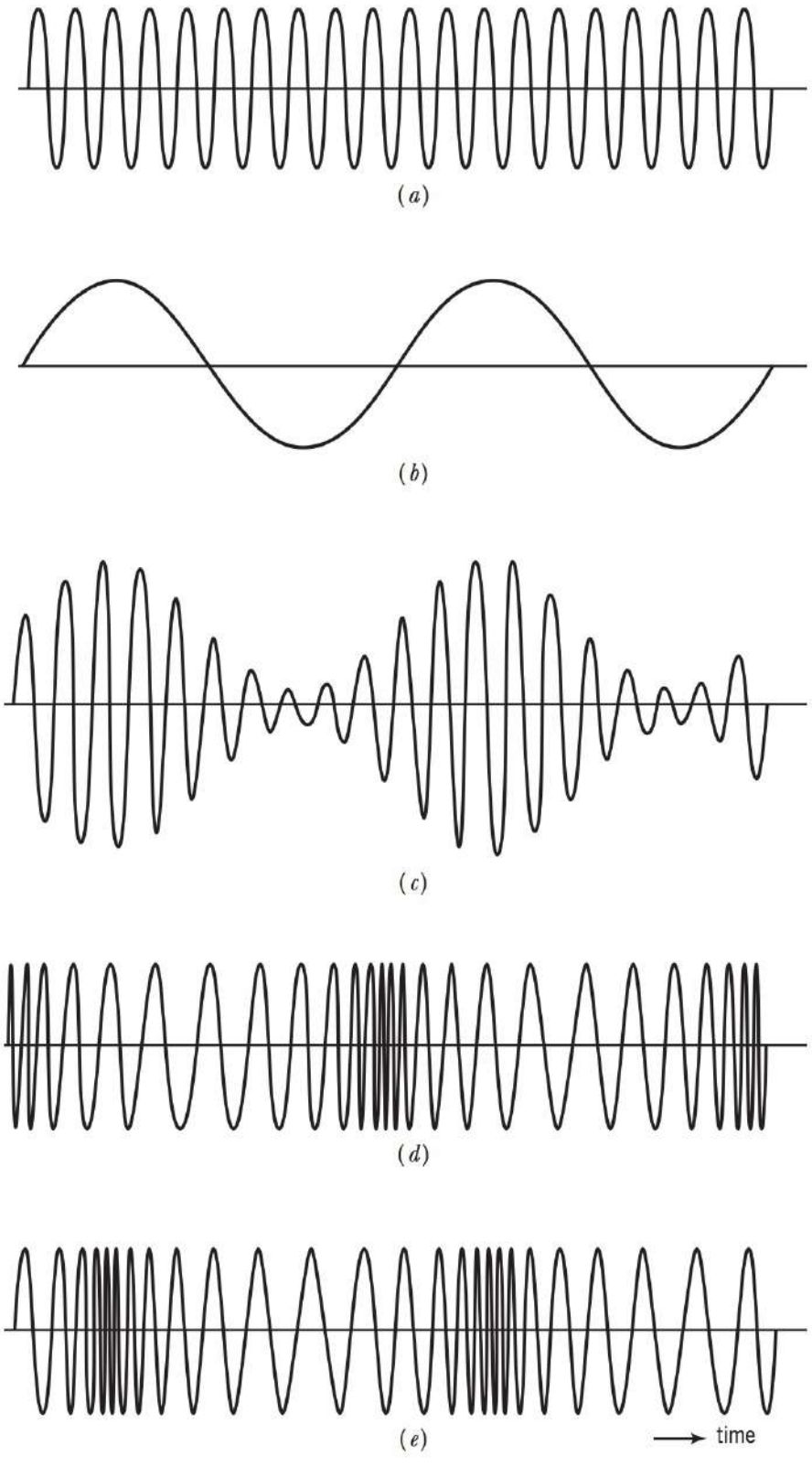


FIGURE 4.1 Illustrating AM, PM, and FM signals produced by a single tone. (*a*) Carrier wave, (*b*) sinusoidal modulating signal, (*c*) amplitude-modulated signal, (*d*) phase-modulated signal, (*e*) frequency-modulated signal.

PROPERTY 1 Constancy of Transmitted Power

From both Eqs. (4.4) and (4.7), we readily see that the amplitude of PM and FM waves is maintained at a constant value equal to the carrier amplitude A_c for all time t, irrespective of the sensitivity factors k_p and k_f. This property is well demonstrated by the PM wave of Figure 4.1*d* and FM wave of Figure 4.1*e*. Consequently, the average transmitted power of angle-modulated waves is a constant, as shown by

$$P_{av} = \frac{1}{2}A_c^2 \tag{4.9}$$

where it is assumed that the load resistance is 1 ohm.

PROPERTY 2 Nonlinearity of the Modulation Process

Another distinctive property of angle modulation is its nonlinear character. We say so because both PM and FM waves violate the principle of superposition. Suppose, for example, that the message signal $m(t)$ is made up of two different components, $m_1(t)$ and $m_2(t)$, as shown by

$$m(t) = m_1(t) + m_2(t)$$

Let $s(t)$, $s_1(t)$, and $s_2(t)$ denote the PM waves produced by $m(t)$, $m_1(t)$, and $m_2(t)$ in accordance with Eq. (4.4), respectively. In light of this equation, we may express these PM waves as follows:

$$\begin{aligned} s(t) &= A_c \cos[2\pi f_c t + k_p(m_1(t) + m_2(t))] \\ s_1(t) &= A_c \cos[2\pi f_c t + k_p m_1(t)] \end{aligned}$$

and

$$s_2(t) = A_c \cos[2\pi f_c t + k_p m_2(t)]$$

From these expressions, despite the fact that $m(t) = m_1(t) + m_2(t)$, we readily see that the principle of superposition is violated because

$$s(t) \neq s_1(t) + s_2(t)$$

A similar result holds for FM waves. The fact that the angle-modulation process is nonlinear complicates the spectral analysis and noise analysis of PM and FM waves, compared to amplitude modulation. By the same token, the angle-modulation process has practical benefits of its own. For example, frequency modulation offers superior noise performance compared to amplitude modulation, which is attributed to the nonlinear character of frequency modulation.

PROPERTY 3 Irregularity of Zero-Crossings

A consequence of allowing the instantaneous angle $\theta_i(t)$ to become dependent on the message signal as in Eq. (4.4) or its integral $\int_0^t m(\tau)\,d\tau$ as in Eq. (4.7) is that, in general, the zero-crossings of a PM or FM wave no longer have a perfect regularity in their spacing across the time-scale. *Zero-crossings* are defined as the instants of time at which a waveform changes its amplitude from a positive to negative value or the other way around. In a way, the irregularity of zero-crossings in angle-modulated waves is also attributed to the nonlinear character of the modulation process. To illustrate this property, we may contrast the PM wave of Figure 4.1*d* and the FM wave of Figure 4.1*e* to Figure 4.1*c* for the corresponding AM wave.

It is important to note that in angle modulation, the information content of the message signal $m(t)$ resides in the zero-crossings of the modulated wave. This statement holds provided the carrier frequency f_c is large compared to the highest frequency component of the message signal $m(t)$.

PROPERTY 4 Visualization Difficulty of Message Waveform

In AM, we see the message waveform as the envelope of the modulated wave, provided, of course, the percentage modulation is less than 100 percent, as illustrated in Figure 4.1*c* for sinusoidal modulation. This is not so in angle modulation, as illustrated by the corresponding waveforms of Figures 4.1*d* and 4.1*e* for PM and FM, respectively. In general, the difficulty in visualizing the message waveform in angle-modulated waves is also attributed to the nonlinear character of angle-modulated waves.

PROPERTY 5 Trade-Off of Increased Transmission Bandwidth for Improved Noise Performance

An important advantage of angle modulation over amplitude modulation is the realization of improved noise performance. This advantage is attributed to the fact that the transmission of a message signal by modulating the angle of a sinusoidal carrier wave is less sensitive to the presence of additive noise than transmission by modulating the amplitude of the carrier. The improvement in noise performance is, however, attained at the expense of a corresponding increase in the transmission bandwidth requirement of angle modulation. In other words, the use of angle modulation offers the possibility of exchanging an increase in transmission bandwidth for an improvement in noise performance. Such a trade-off is not possible with amplitude modulation since the transmission bandwidth of an amplitude-modulated wave is fixed somewhere between the message bandwidth W and $2W$, depending on the type of modulation employed. The effect of noise on angle modulation is discussed in Chapter 6.

EXAMPLE 4.1 Zero-Crossings

Consider a modulating wave $m(t)$ that increases linearly with time t, starting at $t = 0$, as shown by

$$m(t) = \begin{cases} at, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

where a is the slope parameter (see Figure 4.2*a*). In what follows, we study the zero-crossings of the PM and FM waves produced by $m(t)$ for the following set of parameters:

$$f_c = \frac{1}{4} \text{ Hz}$$
$$a = 1 \text{ volt/s}$$

1. *Phase modulation*: phase-sensitivity factor $k_p = \frac{\pi}{2}$ radians/volt. Applying Eq. (4.5) to the given $m(t)$ yields the PM wave

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t + k_p a t), & t \geq 0 \\ A_c \cos(2\pi f_c t), & t < 0 \end{cases}$$

which is plotted in Figure 4.2*b* for $A_c = 1$ volt.

Let t_n denote the instant of time at which the PM wave experiences a zero crossing; this occurs whenever the angle of the PM wave is an odd multiple of $\pi/2$. Then, we may set up

$$2\pi f_c t_n + k_p a t_n = \frac{\pi}{2} + n\pi, \qquad n = 0, 1, 2, \ldots$$

as the *linear* equation for t_n. Solving this equation for t_n, we get the linear formula

$$t_n = \frac{\frac{1}{2} + n}{2f_c + \frac{k_p}{\pi} a}$$

Substituting the given values for f_c, a, and k_p into this linear formula, we get

$$t_n = \frac{1}{2} + n, \qquad n = 0, 1, 2, \ldots$$

where t_n is measured in seconds.

2. *Frequency modulation:* frequency-sensitivity factor, $k_f = 1$ Hz/volt. Applying Eq. (4.8) yields the FM wave

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t + \pi k_f a t^2), & t \geq 0 \\ A_c \cos(2\pi f_c t), & t < 0 \end{cases}$$

which is plotted in Figure 4.2*c*.

Invoking the definition of a zero-crossing, we may set up

$$2\pi f_c t_n + \pi k_f a t_n^2 = \frac{\pi}{2} + n\pi, \qquad n = 0, 1, 2, \ldots$$

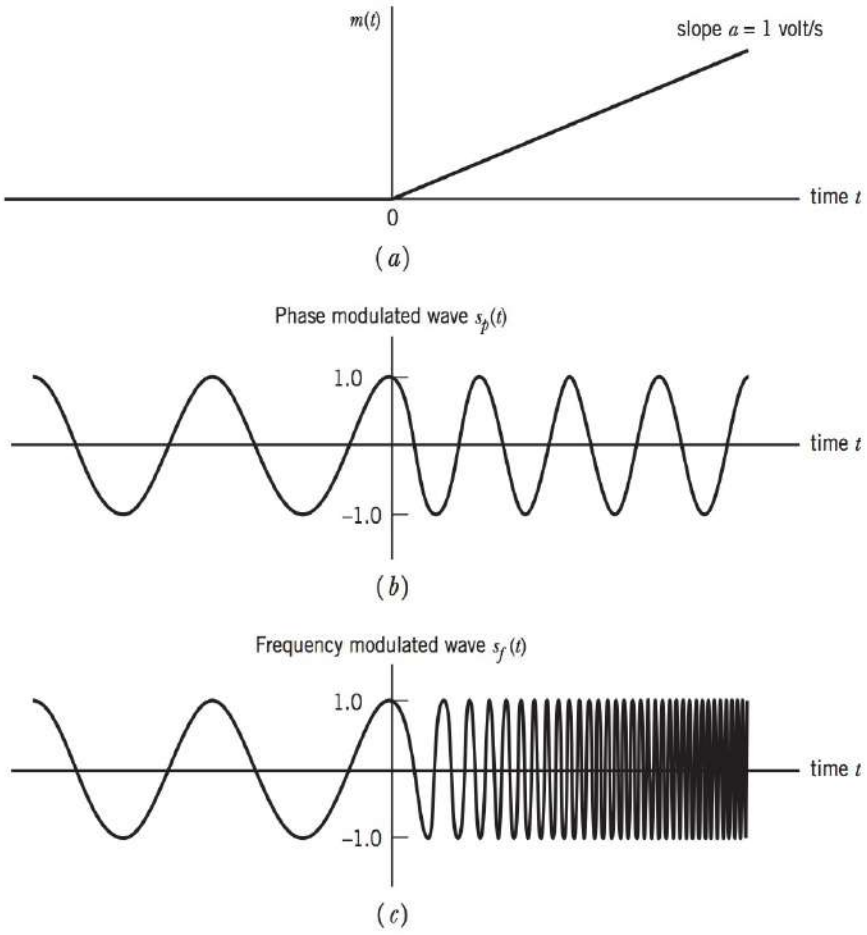


FIGURE 4.2 Starting at time $t = 0$, the figure displays (*a*) linearly increasing message signal $m(t)$, (*b*) phase-modulated wave, and (*c*) frequency-modulated wave.

as the *quadratic* equation for t_n. The positive root of this equation—namely,

$$t_n = \frac{1}{ak_f}\left(-f_c + \sqrt{f_c^2 + ak_f\left(\frac{1}{2} + n\right)}\right), \qquad n = 0, 1, 2, \ldots$$

defines the formula for t_n. Substituting the given values of f_c, a, and k_f into this quadratic formula, we get

$$t_n = \frac{1}{4}(-1 + \sqrt{9 + 16n}), \qquad n = 0, 1, 2, \ldots$$

where t_n is again measured in seconds.

Comparing the zero-crossing results derived for PM and FM waves, we may make the following observations once the linear modulating wave begins to act on the sinusoidal carrier wave:

1. For PM, regularity of the zero-crossings is maintained; the instantaneous frequency changes from the unmodulated value of $f_c = \frac{1}{4}$ Hz to the new constant value of

$$f_c + k_p(a/2\pi) = \frac{1}{2}\text{Hz}$$

2. For FM, the zero-crossings assume an irregular form; as expected, the instantaneous frequency increases linearly with time t.

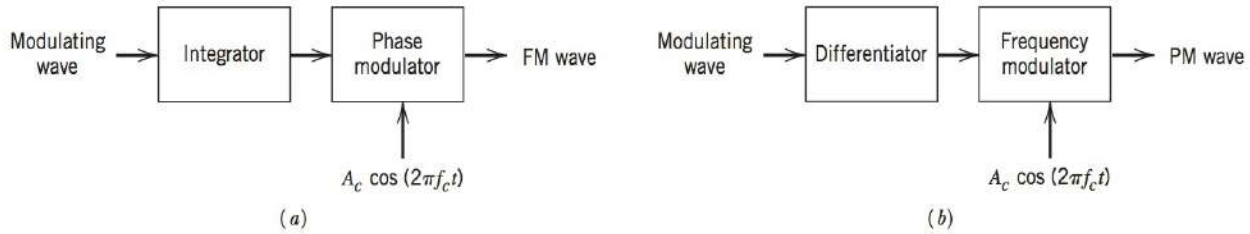


FIGURE 4.3 Illustrating the relationship between frequency modulation and phase modulation. (*a*) Scheme for generating an FM wave by using a phase modulator, (*b*) scheme for generating a PM wave by using a frequency modulator.

The angle-modulated waveforms of Figure 4.2 should be contrasted with the corresponding ones of Figure 4.1. Whereas in the case of sinusoidal modulation depicted in Figure 4.1 it is difficult to discern the difference between PM and FM, this is not so in the case of Figure 4.2 In other words, depending on the modulating wave, it is possible for PM and FM to exhibit entirely different waveforms.

Comparing Eq. (4.5) with (4.8) reveals that an FM signal may be regarded as a PM signal in which the modulating wave is $\int_0^t m(\tau)\, dz$ in place of $m(t)$. This means that an FM signal can be generated by first integrating $m(t)$ and then using the result as the input to a phase modulator, as in Figure 4.3*a*. Conversely, a PM signal can be generated by first differentiating $m(t)$ and then using the result as the input to a frequency modulator, as in Figure 4.3*b*. We may thus deduce all the properties of PM signals from those of FM signals and vice versa. Henceforth, we concentrate our attention on FM signals.

4.3 FREQUENCY MODULATION

The FM signal $s(t)$ defined by Eq. (4.8) is a nonlinear function of the modulating signal $m(t)$, which makes frequency modulation a *nonlinear modulation process*. Consequently, unlike amplitude modulation, the spectrum of an FM signal is not related in a simple manner to that of the modulating signal; rather, its analysis is much more difficult than that of an AM signal.

How then can we tackle the spectral analysis of an FM signal? We propose to provide an empirical answer to this important question by proceeding in the same manner as with AM modulation, that is, we consider the simplest case possible, namely, single-tone modulation.

We could, of course, go on and consider the more elaborate case of a multitone FM signal. However, we propose not to do so, because our immediate objective is to establish an empirical relationship between the transmission bandwidth of an FM signal and the message bandwidth. As we shall subsequently see, the spectral analysis described above provides us with enough insight to propose a solution to the problem.

Consider then a sinusoidal modulating signal defined by

$$m(t) = A_m \cos(2\pi f_m t) \tag{4.10}$$

The instantaneous frequency of the resulting FM signal is

$$\begin{aligned} f_i(t) &= f_c + k_f A_m \cos(2\pi f_m t) \\ &= f_c + \Delta f \cos(2\pi f_m t) \end{aligned} \tag{4.11}$$

where

$$\Delta f = k_f A_m \tag{4.12}$$

The quantity Δf is called the *frequency deviation*, representing the maximum departure of the instantaneous frequency of the FM signal from the carrier frequency f_c. A

fundamental characteristic of an FM signal is that the frequency deviation Δf is proportional to the amplitude of the modulating signal and is independent of the modulating frequency.

Using Eq. (4.11), the angle $\theta_i(t)$ of the FM signal is obtained as

$$\begin{aligned}\theta_i(t) &= 2\pi \int_0^t f_i(t)\, dt \\ &= 2\pi f_c t + \frac{\Delta f}{f_m}\sin(2\pi f_m t)\end{aligned} \tag{4.13}$$

The ratio of the frequency deviation Δf to the modulation frequency f_m is commonly called the *modulation index* of the FM signal. We denote it by β, and so write

$$\beta = \frac{\Delta f}{f_m} \tag{4.14}$$

and

$$\theta_i(t) = 2\pi f_c t + \beta \sin(2\pi f_m t) \tag{4.15}$$

From Eq. (4.15) we see that, in a physical sense, the parameter β represents the phase deviation of the FM signal, that is, the maximum departure of the angle $\theta_i(t)$ from the angle $2\pi f_c t$ of the unmodulated carrier, hence; β is measured in radians.

The FM signal itself is given by

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_m t)] \tag{4.16}$$

Depending on the value of the modulation index β, we may distinguish two cases of frequency modulation:

- *Narrow-band* FM, for which β is small compared to one radian.
- *Wide-band* FM, for which β is large compared to one radian.

These two cases are considered next, in that order.

NARROW-BAND FREQUENCY MODULATION

Consider Eq. (4.16), which defines an FM signal resulting from the use of a sinusoidal modulating signal. Expanding this relation, we get

$$\begin{aligned}s(t) = {} & A_c \cos(2\pi f_c t)\cos[\beta \sin(2\pi f_m t)] \\ & - A_c \sin(2\pi f_c t)\sin[\beta \sin(2\pi f_m t)]\end{aligned} \tag{4.17}$$

Assuming that the modulation index β is small compared to one radian, we may use the following two approximations:

$$\cos[\beta \sin(2\pi f_m t)] \simeq 1$$

and

$$\sin[\beta \sin(2\pi f_m t)] \simeq \beta \sin(2\pi f_m t)$$

Hence, Eq. (4.17) simplifies to

$$s(t) \simeq A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t)\sin(2\pi f_m t) \tag{4.18}$$

Equation (4.18) defines the approximate form of a narrow-band FM signal produced by the sinusoidal modulating signal $A_m \cos(2\pi f_m t)$.

We now expand Eq. (4.18) as follows:

$$s(t) \simeq A_c \cos(2\pi f_c t) + \frac{1}{2}\beta A_c\{\cos[2\pi(f_c + f_m)t] - \cos[2\pi(f_c - f_m)t]\} \tag{4.19}$$

This expression is somewhat similar to the corresponding one defining an AM signal, which is reproduced from Example 3.1 as follows:

$$s_{AM}(t) = A_c \cos(2\pi f_c t) + \frac{1}{2}\mu A_c\{\cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t]\} \qquad (4.20)$$

where μ is the modulation factor of the AM signal. Comparing Eqs. (4.19) and (4.20), we see that in the case of sinusoidal modulation, the basic difference between an AM signal and a narrow-band FM signal is that the algebraic sign of the lower side frequency in the narrow-band FM is reversed. Thus, a narrow-band FM signal requires essentially the same transmission bandwidth (i.e., $2f_m$) as the AM signal.

We may represent the narrow-band FM signal with a phasor diagram as shown in Figure 4.4*a*, where we have used the carrier phasor as reference. We see that the resultant of the two side-frequency phasors is always at right angles to the carrier phasor. The effect of this is to produce a resultant phasor representing the narrow-band FM signal that is approximately of the same amplitude as the carrier phasor, but out of phase with respect to it. This phasor diagram should be contrasted with that of Figure 4.4*b*, representing an AM signal. In this latter case we see that the resultant phasor representing the AM signal has an amplitude different from that of the carrier phasor, but always in phase with it.

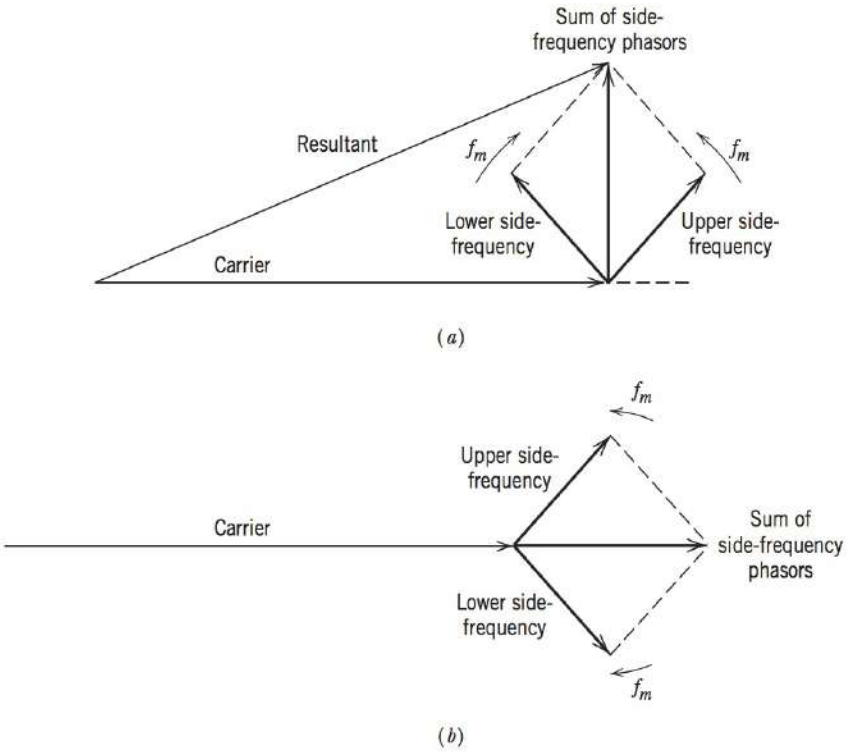


FIGURE 4.4 A Phasor comparison of narrow-band FM and AM waves for sinusoidal modulation, (*a*) narrow-band FM wave. (*b*) AM wave.

EXAMPLE 4.2 Phase Noise

While the intentional use of narrow-band FM for analog information sources is not common, unintentional narrow-band phase modulation is quite common. This unintentional phase modulation is commonly referred to as *phase noise*. Phase noise is often introduced by oscillators in band-pass communications and has a number of causes. Some causes are deterministic, such as those created by changes in oscillator temperature, supply voltage, physical vibration, magnetic field, humidity, or output load impedance. The phase noise due to these sources may be minimized by good design. Other sources

are categorized as random, which can be controlled but not eliminated by appropriate circuitry, such as phase-lock loops. Phase-locked loops are considered later in Section 4.4.

Oscillators play a fundamental role in band-pass communications and most systems include several of them. The phase noise introduced by oscillators has a multiplicative effect on an angle-modulated signal. For example, if $s(t)$ is an angle-modulated signal, and $c(t)$ is the receiver oscillator, having phase noise $\phi_n(t)$, then when translating the signal from f_c to f_b (see Section 3.7), the output is

$$\begin{aligned} s(t)c(t) &= A_c \cos[2\pi f_c t + \phi(t)] \times \cos[2\pi(f_c - f_b)t + \phi_n(t)] \\ &= \frac{A_c}{2}[\cos(2\pi f_b + \phi(t) - \phi_n(t)) + \cos(2\pi(2f_c - f_b) + \phi(t) + \phi_n(t))] \\ &\approx \frac{A_c}{2}\cos[2\pi f_b + \phi(t) - \phi_n(t)] \end{aligned}$$

where we have assumed that the high-frequency term in the second line has been removed by a band-pass filter centered around f_b, following the mixer. Thus the phase noise of the oscillator directly affects the information component of the angle-modulated signal.

Phase noise due to oscillators and other random sources tends to be slowly varying with most of its energy concentrated at low frequencies. Thus we may use the results of our narrowband FM analysis to characterize it. An example oscillator spectrum that includes phase noise is shown in Figure 4.5, where the oscillator spectrum has been shifted to *dc* for convenient representation.

A common practical concern is the root-mean-square (*rms*) phase error introduced by the phase noise on this carrier. To determine this *rms* phase error, we first make the following three observations:

1. For small modulation indices, the spectrum of the PM signal is similar to that of the modulating signal plus a carrier component (see Problem 4.7).
2. A phase detection system often includes a phase-locked loop (see Section 4.4) that tracks the carrier and those frequency components of the phase (noise) variations (noise) below a certain maximum frequency (the loop

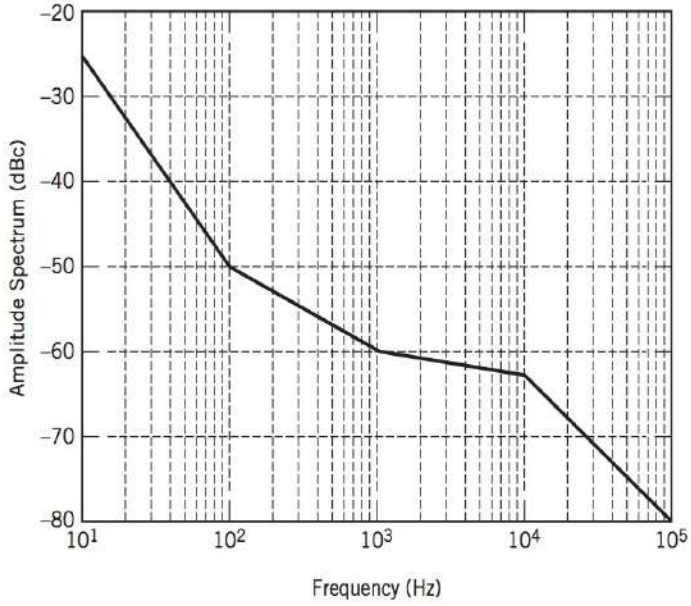


FIGURE 4.5 Phase noise amplitude spectrum. (*Note:* dBc implies dB relative to carrier level.)

bandwidth), say f_1. This phase-tracking effectively nulls those frequency components of the phase noise less than f_1.

3. Phase noise, which is outside the bandwidth of the message bandwidth W, is eliminated when the message is low-pass filtered.

From observation (1), if Figure 4.5 represents the spectrum of the noisy oscillator (excluding the carrier), then the amplitude spectrum of the modulating signal $\phi_n(t)$ is also approximately given by Figure 4.5. From the Rayleigh energy theorem of Fourier transforms discussed in Chapter 2, time-domain energies and frequency-domain energies are equivalent, so we write

$$\int_{-\infty}^{\infty} |\phi_n(t)|^2 \, dt = \int_{-\infty}^{\infty} |\Phi_n(f)|^2 \, df$$

where $\Theta(f)$ is the Fourier transform of $\phi_n(t)$. After the phase-locked loop, we have

$$\int_{-\infty}^{\infty} |\overline{\phi}_n(t)|^2 \, dt = \int_{-\infty}^{-f_1} |\Phi_n(f)|^2 \, df + \int_{+f_1}^{\infty} |\Phi_n(f)|^2 \, df$$

where $\overline{\phi}_n(t)$ excludes the carrier and frequency components of $\phi_n(t)$ below f_1. The left-hand side represents the energy of the phase variations, so combining these results with the third observation, we find that the *root-mean-square phase error* is

$$\overline{\phi}_{rms} = \sqrt{2 \int_{f_1}^{W} |\Phi_n(f)|^2 \, df} \text{ radians}$$

Numerically integrating the phase noise spectrum of Figure 4.5 (see Problem 4.27) for $f_1 = 10$ Hz and $W = 10$ kHz, shows that the rms phase error is 6.5°. Since this value is less than 0.3 radians, the use of the narrowband FM analysis in the solution is justified.

WIDE-BAND FREQUENCY MODULATION

We next wish to determine the spectrum of the single-tone FM signal of Eq. (4.16) for an arbitrary value of the modulation index β. In general, an FM signal produced by a sinusoidal modulating signal, as in Eq. (4.16), is in itself nonperiodic unless the carrier frequency f_c is an integral multiple of the modulation frequency f_m. However, we may simplify matters by using the complex representation of band-pass signals described in Chapter 2. Specifically, we assume that the carrier frequency f_c is large enough (compared to the bandwidth of the FM signal) to justify rewriting Eq. (4.16) in the form

$$\begin{aligned} s(t) &= \text{Re}[A_c \exp(j2\pi f_c t + j\beta \sin(2\pi f_m t))] \\ &= \text{Re}[\tilde{s}(t) \exp(j2\pi f_c t)] \end{aligned} \tag{4.21}$$

where $\tilde{s}(t)$ is the complex envelope of the FM signal $s(t)$, defined by

$$\tilde{s}(t) = A_c \exp[j\beta \sin(2\pi f_m t)] \tag{4.22}$$

Thus, unlike the original FM signal $s(t)$, the complex envelope $\tilde{s}(t)$ is a periodic function of time with a fundamental frequency equal to the modulation frequency f_m. We may therefore expand $\tilde{s}(t)$ in the form of a complex Fourier series as follows:

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_m t) \tag{4.23}$$

where the complex Fourier coefficient c_n is given by

$$\begin{aligned} c_n &= f_m \int_{-1/2f_m}^{1/2f_m} \tilde{s}(t)\exp(-j2\pi n f_m t)\, dt \\ &= f_m A_c \int_{-1/2f_m}^{1/2f_m} \exp[j\beta \sin(2\pi f_m t) - j2\pi n f_m t]\, dt \end{aligned} \tag{4.24}$$

Define a new variable:

$$x = 2\pi f_m t \tag{4.25}$$

Hence, we may rewrite Eq. (4.24) in the new form

$$c_n = \frac{A_c}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin x - nx)]\, dx \tag{4.26}$$

The integral on the right-hand side of Eq. (4.26), except for a scaling factor, is recognized as the *n*th *order Bessel function of the first kind*[1] and argument β. This function is commonly denoted by the symbol $J_n(\beta)$, as shown by

$$J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin x - nx)]\, dx \tag{4.27}$$

Accordingly, we may reduce Eq. (4.26) to

$$c_n = A_c J_n(\beta) \tag{4.28}$$

Substituting Eq. (4.28) in (4.23), we get, in terms of the Bessel function $J_n(\beta)$, the following expansion for the complex envelope of the FM signal:

$$\tilde{s}(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta)\, \exp(j2\pi n f_m t) \tag{4.29}$$

Next, substituting Eq. (4.29) in (4.21), we get

$$s(t) = A_c \cdot \mathrm{Re}\left[\sum_{n=-\infty}^{\infty} J_n(\beta)\, \exp[j2\pi(f_c + nf_m)t] \right] \tag{4.30}$$

Interchanging the order of summation and evaluation of the real part in the right-hand side of Eq. (4.30), we get

$$s(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + nf_m)t] \tag{4.31}$$

This is the desired form for the Fourier series representation of the single-tone FM signal $s(t)$ for an arbitrary value of β. The discrete spectrum of $s(t)$ is obtained by taking the Fourier transforms of both sides of Eq. (4.31). We thus have

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta)[\delta(f - f_c - nf_m) + \delta(f + f_c + nf_m)] \tag{4.32}$$

In Figure 4.6 we have plotted the Bessel function $J_n(\beta)$ versus the modulation index β for different positive integer values of n. We can develop further insight into the behavior of the Bessel function $J_n(\beta)$ by making use of the following properties:

1. For n even, we have $J_n(\beta) = J_{-n}(\beta)$; on the other hand, for n odd, we have $J_n(\beta) = -J_{-n}(\beta)$. That is,

$$J_n(\beta) = (-1)^n J_{-n}(\beta) \qquad \text{for all } n \tag{4.33}$$

2. For small values of the modulation index β, we have

$$\left.\begin{aligned} J_0(\beta) &\simeq 1 \\ J_1(\beta) &\simeq \frac{\beta}{2} \\ J_n(\beta) &\simeq 0, \qquad n > 2 \end{aligned}\right\} \tag{4.34}$$

3.

$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \tag{4.35}$$

Thus, using Eqs. (4.32) through (4.35) and the curves of Figure 4.6, we may make the following observations:

1. The spectrum of an FM signal contains a carrier component and an infinite set of side frequencies located symmetrically on either side of the carrier at frequency separations of f_m, $2f_m$, $3f_m$, ... In this respect, the result is unlike that which prevails in an AM system, since in an AM system a sinusoidal modulating signal gives rise to only one pair of side frequencies.

2. For the special case of β small compared with unity, only the Bessel coefficients $J_0(\beta)$ and $J_1(\beta)$ have significant values, so that the FM signal is effectively composed of a carrier and a single pair of side frequencies at $f_c \pm f_m$. This situation corresponds to the special case of narrowband FM that was considered previously.

3. The amplitude of the carrier component varies with β according to $J_0(\beta)$. That is, unlike an AM signal, the amplitude of the carrier component of an FM signal is dependent on the modulation index β. The physical explanation for this property is that the envelope of an FM signal is constant, so that the average power of such a signal developed across a 1-ohm resistor is also constant, as shown by

$$P = \frac{1}{2}A_c^2 \tag{4.36}$$

When the carrier is modulated to generate the FM signal, the power in the side frequencies may appear only at the expense of the power originally in the carrier, thereby

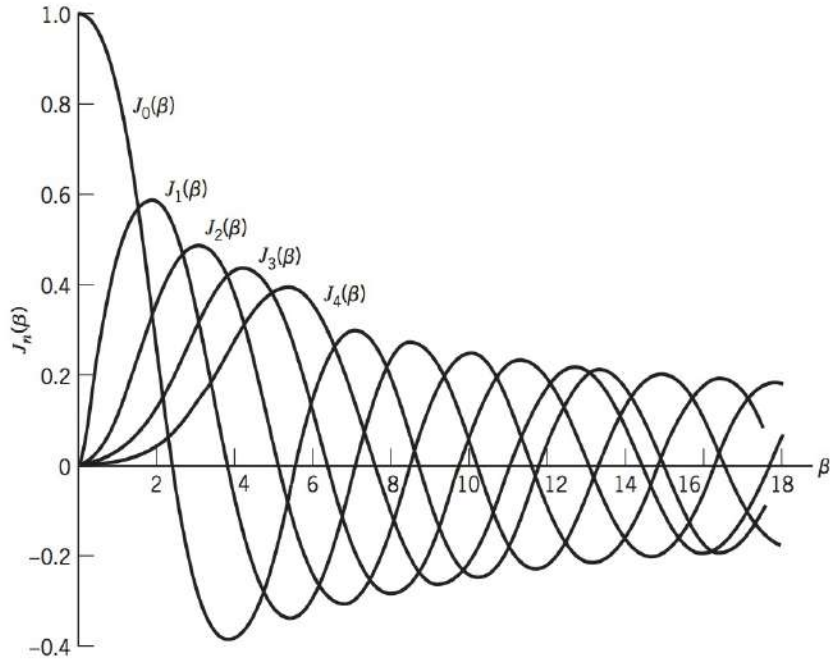


FIGURE 4.6 Plots of Bessel functions of the first kind.

making the amplitude of the carrier component dependent on β. Note that the average power of an FM signal may also be determined from Eq. (4.31), obtaining

$$P = \frac{1}{2}A_c^2 \sum_{n=-\infty}^{\infty} J_n^2(\beta) \tag{4.37}$$

Substituting Eq. (4.35) in (4.37), the expression for the average power P reduces to Eq. (4.36), and so it should.

EXAMPLE 4.3 Spectra of FM Signals

In this example, we wish to investigate the ways in which variations in the amplitude and frequency of a sinusoidal modulating signal affect the spectrum of the FM signal. Consider first the case when the frequency of the modulating signal is fixed, but its amplitude is varied, producing a corresponding variation in the frequency deviation Δf. Thus, keeping the modulation frequency

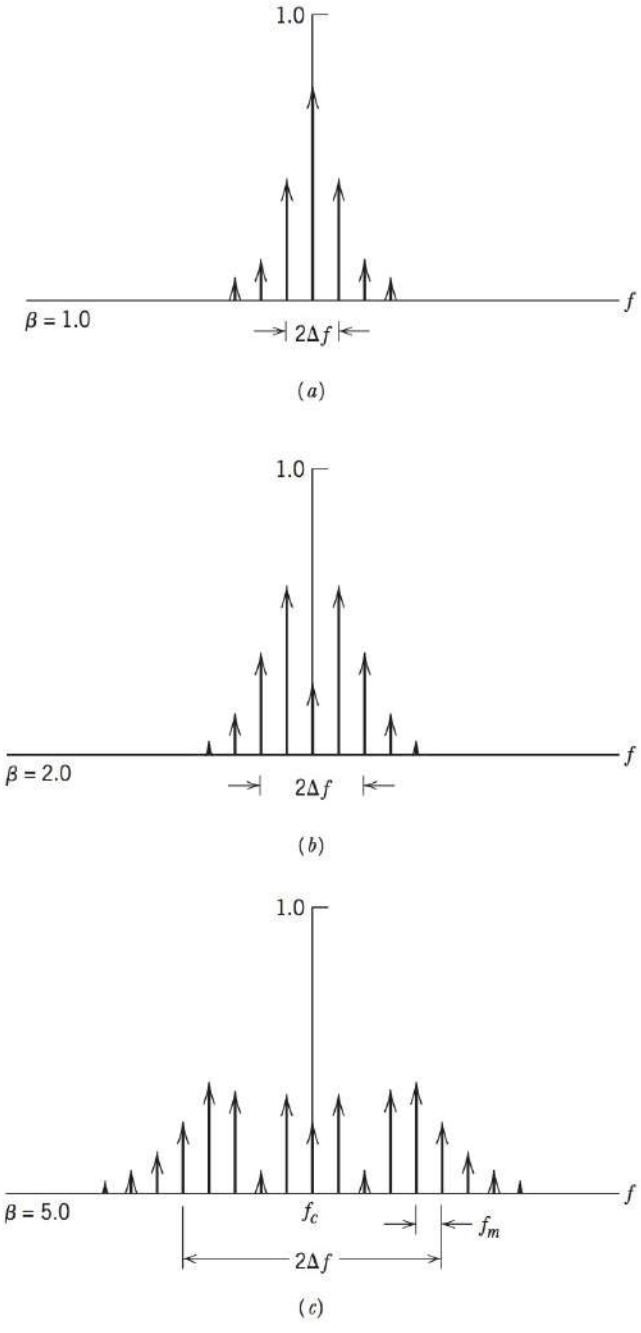


FIGURE 4.7 Discrete amplitude spectra of an FM signal, normalized with respect to the carrier amplitude, for the case of sinusoidal modulation of fixed frequency and varying amplitude. Only the spectra for positive frequencies are shown.

f_m fixed, we find that the amplitude spectrum of the resulting FM signal is as shown plotted in Figure 4.7 for $\beta = 1$, 2, and 5. In this diagram we have normalized the spectrum with respect to the unmodulated carrier amplitude.

Consider next the case when the amplitude of the modulating signal is fixed; that is, the frequency deviation Δf is maintained constant, and the modulation frequency f_m is varied. In this case we find that the amplitude spectrum of the resulting FM signal is as shown plotted in Figure 4.8 for $\beta = 1$, 2, and 5. This time we see that when Δf is fixed and β is increased, we have an increasing number of spectral lines crowding into the fixed frequency interval $f_c - \Delta f < |f| < f_c + \Delta f$. That is, when β approaches infinity, the bandwidth of the FM wave approaches the limiting value of $2\Delta f$, which is an important point to keep in mind.

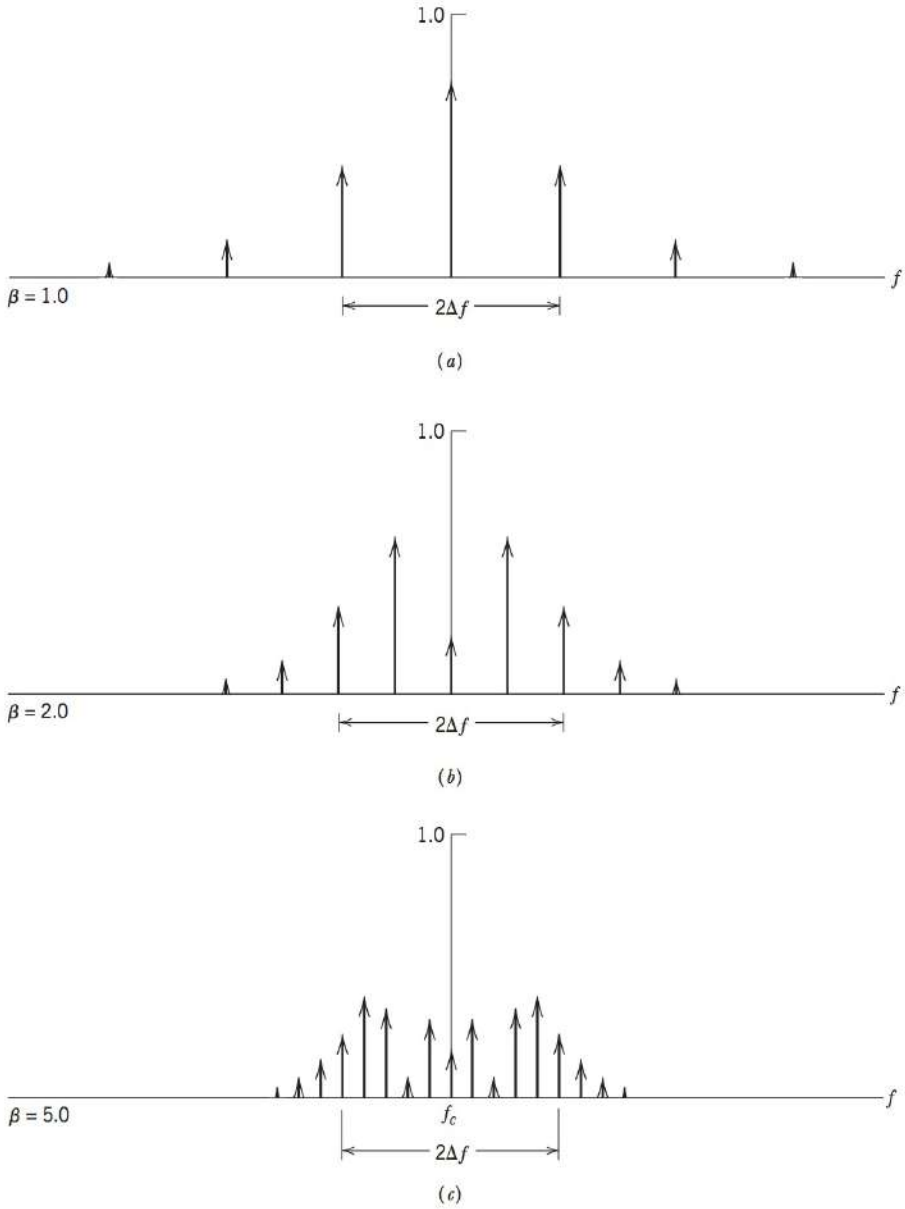


FIGURE 4.8 Discrete amplitude spectra of an FM signal, normalized with respect to the carrier amplitude, for the case of sinusoidal modulation of varying frequency and fixed amplitude. Only the spectra for positive frequencies are shown.

TRANSMISSION BANDWIDTH OF FM SIGNALS

In theory, an FM signal contains an infinite number of side frequencies so that the bandwidth required to transmit such a signal is similarly infinite in extent. In practice, however, we find that the FM signal is effectively limited to a finite number of significant

side frequencies compatible with a specified amount of distortion. We may therefore specify an effective bandwidth required for the transmission of an FM signal. Consider first the case of an FM signal generated by a single-tone modulating wave of frequency f_m. In such an FM signal, the side frequencies that are separated from the carrier frequency f_c by an amount greater than the frequency deviation Δf decrease rapidly toward zero, so that the bandwidth always exceeds the total frequency excursion, but nevertheless is limited. Specifically, for large values of the modulation index β, the bandwidth approaches, and is only slightly greater than, the total frequency excursion $2\Delta f$. On the other hand, for small values of the modulation index β, the spectrum of the FM signal is effectively limited to the carrier frequency f_c and one pair of side frequencies at $f_c \pm f_m$, so that the bandwidth approaches $2f_m$. We may thus define an approximate rule for the transmission bandwidth of an FM signal generated by a single-tone modulating signal of frequency f_m as follows:

$$B_T \simeq 2\Delta f + 2f_m = 2\Delta f\left(1 + \frac{1}{\beta}\right) \tag{4.38}$$

This empirical relation is known as *Carson's rule*.

For a more accurate assessment of the bandwidth requirement of an FM signal, we may use a definition based on retaining the maximum number of significant side frequencies whose amplitudes are all greater than some selected value. A convenient choice for this value is 1 percent of the unmodulated carrier amplitude. We may thus define *the transmission bandwidth of an FM wave as the separation between the two frequencies beyond which none of the side frequencies is greater than 1 percent of the carrier amplitude obtained when the modulation is removed*. That is, we define the transmission bandwidth as $2n_{\max} f_m$, where f_m is the modulation frequency and $n_{\max}$ is the largest value of the integer n that satisfies the requirement $|J_n(\beta)| > 0.01$. The value of $n_{\max}$ varies with the modulation index β and can be determined readily from tabulated values of the Bessel function $J_n(\beta)$. Table 4.1 shows the total number of significant side frequencies (including both the upper and lower side frequencies) for different values of β, calculated on the 1 percent basis explained herein. The transmission bandwidth B_T calculated using this procedure can be presented in the form of a *universal curve* by normalizing it with respect to the frequency deviation Δf and then plotting it versus β. This curve is shown in Figure 4.9, which is drawn as a best fit through the set of points obtained by using Table 4.1. In Figure 4.9 we note that as the modulation index β is increased, the bandwidth occupied by the

Edwin H. Armstrong (1890–1954)

Among Armstrong's many inventions related to radio were the superheterodyne receiver (1918) and FM radio (1933). Previously, in 1922, J. H. Carson (of Carson's rule) published a paper claiming there was no advantage to frequency modulation. To Armstrong's credit, in the face of such criticism, he was able to show that wideband FM offered much clearer transmission than the amplitude modulation then in current use.

In 1945, RCA won a petition to the regulatory agencies to have the FM radio band moved from 40–52 MHz to 88–108 MHz. The objective was to protect its AM radio business and promote the fledgling television business. This change rendered all of Armstong's FM radios useless overnight. Furthermore, RCA argued Armstrong's FM patent, and prevented his collection of royalties on new FM stations. Penniless and distraught, Armstrong committed suicide by jumping from his fourteenth story balcony. The actions by RCA are considered to have set back FM radio for decades. Even so, FM became one of the dominant transmission methods by the end of the 20th century.

TABLE 4.1 Number of significant side frequencies of a wide-band FM signal for varying modulation index

Modulation Index β	Number of Significant Side Frequencies $2n_{\max}$
0.1	2
0.3	4
0.5	4
1.0	6
2.0	8
5.0	16
10.0	28
20.0	50
30.0	70

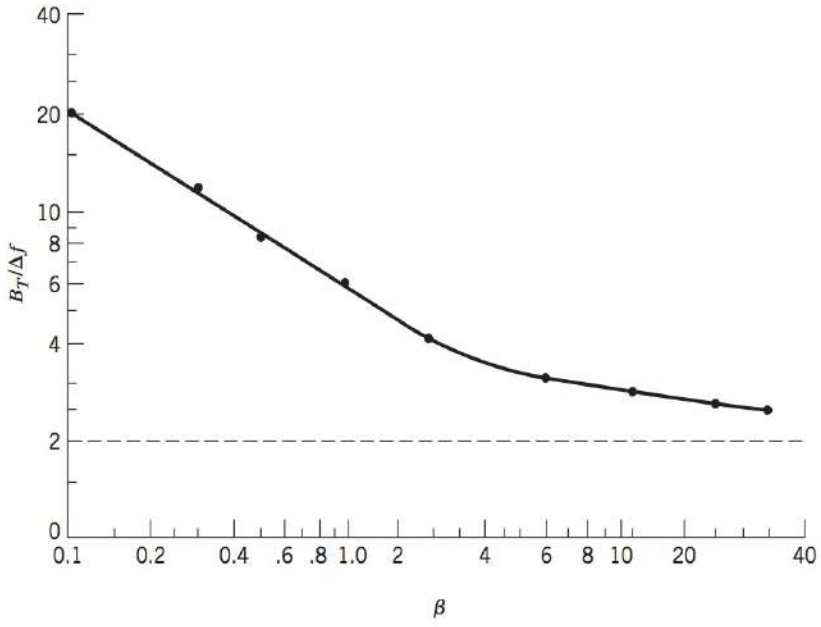


FIGURE 4.9 Universal curve for evaluating the 1 percent bandwidth of an FM wave.

significant side frequencies drops toward that over which the carrier frequency actually deviates. This means that small values of the modulation index β are relatively more extravagant in transmission bandwidth than are the larger values of β.

Consider next the more general case of an arbitrary modulating signal $m(t)$ with its highest frequency component denoted by W. The bandwidth required to transmit an FM signal generated by this modulating signal is estimated by using a worst-case tone-modulation analysis. Specifically, we first determine the so-called *deviation ratio D*, defined as the ratio of the frequency deviation Δf, which corresponds to the maximum possible amplitude of the modulation signal $m(t)$, to the highest modulation frequency W; these conditions represent the extreme cases possible. *The deviation ratio D plays the same role for nonsinusoidal modulation that the modulation index β plays for the case of sinusoidal modulation.* Then, replacing β by D and replacing f_m with W, we may use Carson's rule given by Eq. (4.38) or the universal curve of Figure 4.9 to obtain a value for the transmission bandwidth of the FM signal. From a practical viewpoint, Carson's rule somewhat underestimates the bandwidth requirement of an FM system, whereas using the universal curve of Figure 4.9 yields a somewhat conservative result. Thus, the choice of a transmission bandwidth that lies between the bounds provided by these two rules of thumb is acceptable for most practical purposes.

EXAMPLE 4.4

In North America, the maximum value of frequency deviation Δf is fixed at 75 kHz for commercial FM broadcasting by radio. If we take the modulation frequency $W = 15$ kHz, which is typically the "maximum" audio frequency of interest in FM transmission, we find that the corresponding value of the deviation ratio is

$$D = \frac{75}{15} = 5$$

Using Carson's rule of Eq. (4.38), replacing β by D, and replacing f_m by W, the approximate value of the transmission bandwidth of the FM signal is obtained as

$$B_T = 2(75 + 15) = 180 \text{ kHz}$$

On the other hand, use of the curve of Figure 4.9 gives the transmission bandwidth of the FM signal to be

$$B_T = 3.2\,\Delta f = 3.2 \times 75 = 240 \text{ kHz}$$

Thus Carson's rule underestimates the transmission bandwidth by 25 percent compared with the result of using the universal curve of Figure 4.9.

GENERATION OF FM SIGNALS

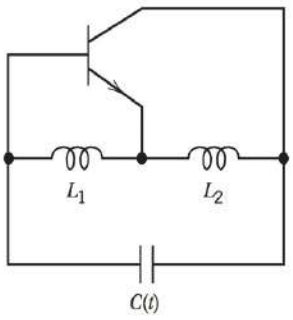


FIGURE 4.10 Hartley oscillator.

In a direct FM system, the instantaneous frequency of the carrier wave is varied directly in accordance with the message signal by means of a device known as a *voltage-controlled oscillator*. One way of implementing such a device is to use a sinusoidal oscillator having a highly selective frequency-determining resonant network and to control the oscillator by symmetrical incremental variation of the reactive components of this network. An example of such a scheme is shown in Figure 4.10, depicting a *Hartley oscillator*. We assume that the capacitive component of the frequency-determining network in the oscillator consists of a fixed capacitor shunted by a voltage-variable capacitor. The resultant capacitance is represented by $C(t)$ in Figure 4.10. A voltage-variable capacitor, commonly called a *varactor* or *varicap*, is one whose capacitance depends on the voltage applied across its electrodes. The variable-voltage capacitance may be obtained, for example, by using a *p-n* junction diode that is biased in the reverse direction; the larger the reverse voltage applied to such a diode, the smaller the transition capacitance of the diode. The frequency of oscillation of the Hartley oscillator of Figure 4.10 is given by

$$f_i(t) = \frac{1}{2\pi\sqrt{(L_1 + L_2)C(t)}} \tag{4.39}$$

where $C(t)$ is the total capacitance of the fixed capacitor and the variable-voltage capacitor, and L_1, and L_2 are the two inductances in the frequency-determining network of the oscillator. Assume that for a sinusoidal modulating wave of frequency f_m, the capacitance $C(t)$ is expressed as

$$C(t) = C_0 + \Delta C \cos(2\pi f_m t) \tag{4.40}$$

where C_0 is the total capacitance in the absence of modulation and ΔC is the maximum change in capacitance. Substituting Eq. (4.40) in (4.39), we get

$$f_i(t) = f_0\left[1 + \frac{\Delta C}{C_0}\cos(2\pi f_m t)\right]^{-1/2} \tag{4.41}$$

where f_0 unmodulated frequency of oscillation, that is,

$$f_0 = \frac{1}{2\pi\sqrt{C_0(L_1 + L_2)}} \tag{4.42}$$

Provided that the maximum change in capacitance ΔC is small compared with the unmodulated capacitance C_0, we may approximate Eq. (4.41) as

$$f_i(t) \simeq f_0\left[1 - \frac{\Delta C}{2C_0}\cos(2\pi f_m t)\right] \tag{4.43}$$

Let

$$\frac{\Delta C}{2C_0} = -\frac{\Delta f}{f_0} \tag{4.44}$$

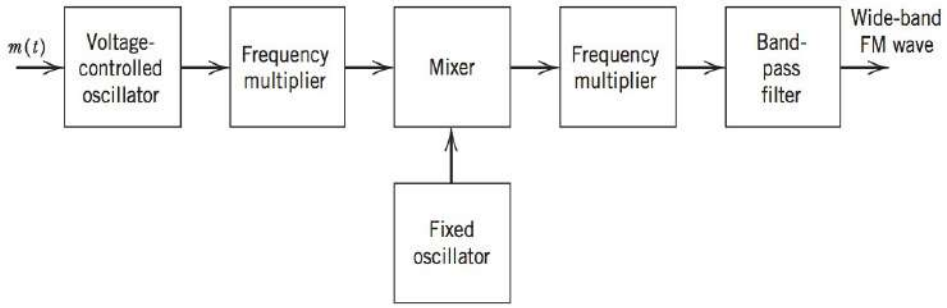


FIGURE 4.11 Block diagram of wide-band frequency modulator using a voltage-controlled oscillator.

Hence, the instantaneous frequency of the oscillator, which is being frequency-modulated by varying the capacitance of the frequency-determining network, is approximately given by

$$f_i(t) \simeq f_0 + \Delta f \cos(2\pi f_m t) \tag{4.45}$$

Equation (4.45) is the desired relation for the instantaneous frequency of an FM wave, assuming sinusoidal modulation.

In order to generate a wide-band FM wave with the required frequency deviation, we may use the configuration shown in Figure 4.11 consisting of a voltage-controlled oscillator, followed by a series of frequency multipliers and mixers. This configuration permits the attainment of good oscillator stability, constant proportionality between output frequency change and input voltage change, and the necessary frequency deviation to achieve wide-band FM.

An FM transmitter using the direct method as described, however, has the disadvantage that the carrier frequency is not obtained from a highly stable oscillator. It is therefore necessary, in practice, to provide some auxiliary means by which a very stable frequency generated by a crystal will be able to control the carrier frequency. One method of effecting this control is illustrated in Figure 4.12. The output of the FM generator is applied to a mixer together with the output of a crystal-controlled oscillator, and the difference frequency term is extracted. The mixer output is next applied to a frequency discriminator and then low-pass filtered. A frequency discriminator is a device whose output voltage has an instantaneous amplitude that is proportional to the instantaneous frequency of the FM signal applied to its input; this device is described in the next subsection. When the FM transmitter has exactly the correct carrier frequency, the low-pass filter output is zero. However, deviations of the transmitter carrier frequency from its assigned value will cause the frequency discriminator–filter combination to develop a dc output voltage with a polarity determined by the sense of the transmitter frequency drift. This dc voltage, after suitable amplification, is applied to the voltage-controlled oscillator of the FM transmitter in such a way as to modify the frequency of the oscillator in a direction that tends to restore the carrier frequency to its correct value.

The feedback scheme of Figure 4.12 is an example of a *frequency-locked loop*, which is closely related to a phase-locked loop. We will explain the phase-locked loop in Section 4.4.

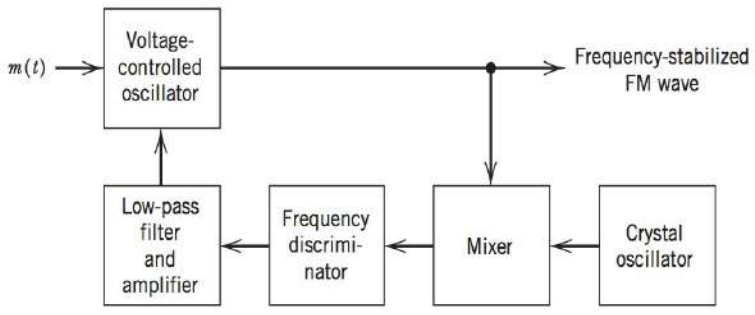


FIGURE 4.12 A feedback scheme for the frequency stabilization of a frequency modulator.

DEMODULATION OF FM SIGNALS

Frequency demodulation is the process that enables us to recover the original modulating signal from a frequency-modulated signal. The objective is to produce a transfer characteristic that is the inverse of that of the frequency modulator, which can be realized directly or indirectly. Here we describe a direct method of frequency demodulation involving the use of a popular device known as a frequency discriminator, whose instantaneous amplitude is directly proportional to the instantaneous frequency of the input FM signal. In the next section, we describe an indirect method of frequency demodulation that uses another popular device known as a phase-locked loop.

Basically, the *frequency discriminator* consists of a *slope circuit* followed by an *envelope detector*. An ideal slope circuit is characterized by a transfer function that is purely imaginary, varying linearly with frequency inside a prescribed frequency interval. Consider the transfer function plotted in Figure 4.13*a*, which is defined by

$$H_1(f) = \begin{cases} j2\pi a\left(f - f_c + \dfrac{B_T}{2}\right), & f_c - \dfrac{B_T}{2} \le f \le f_c + \dfrac{B_T}{2} \\ j2\pi a\left(f + f_c - \dfrac{B_T}{2}\right), & -f_c - \dfrac{B_T}{2} \le f \le -f_c + \dfrac{B_T}{2} \\ 0, & \text{elsewhere} \end{cases} \tag{4.46}$$

where a is a constant parameter. We wish to evaluate the response of this slope circuit, denoted by $s_1(t)$, which is produced by an FM signal $s(t)$ of carrier frequency f_c and transmission bandwidth B_T. It is assumed that the spectrum of $s(t)$ is essentially zero outside the frequency interval $f_c - B_T/2 \le |f| \le f_c + B_T/2$. For evaluation of the response $s_1(t)$, it is convenient to use the procedure described in Section 2.10, which involves replacing the slope circuit with an equivalent low-pass filter and driving this filter with the complex envelope of the input FM signal $s(t)$.

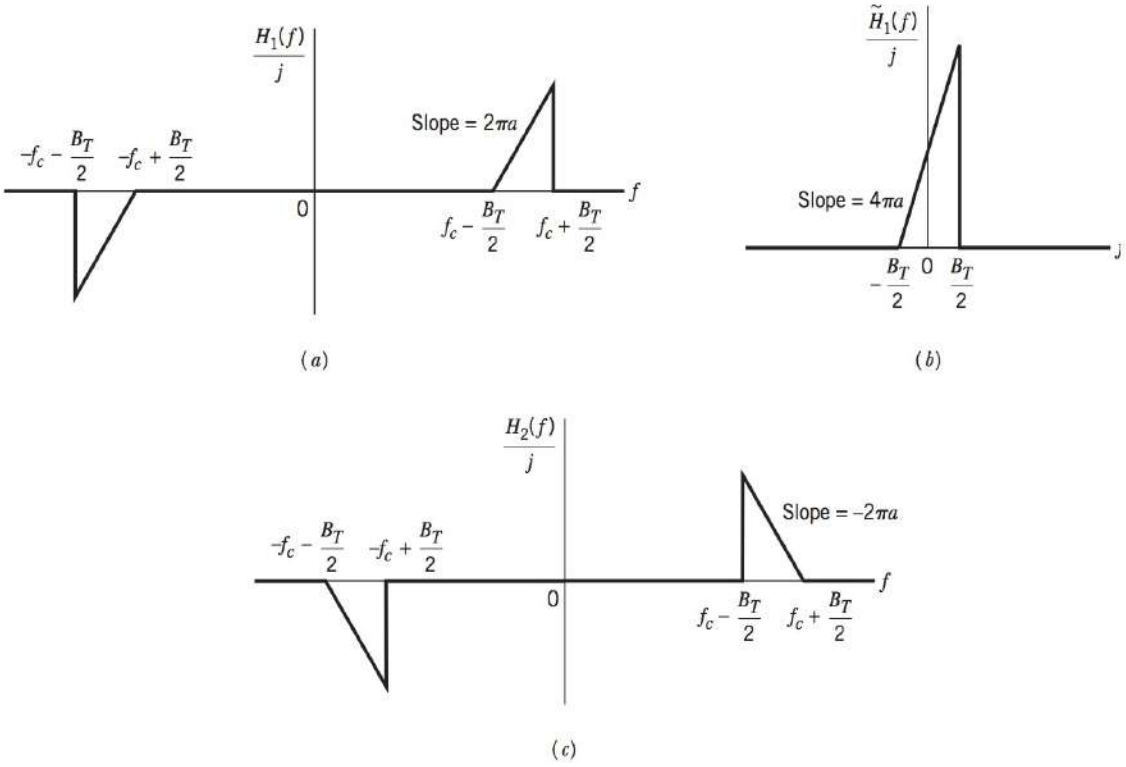


FIGURE 4.13 (*a*) Frequency response of ideal slope circuit. (*b*) The slope circuit's response. (*c*) Frequency response of the complex low-pass filter equivalent to ideal slope circuit complementary to that of part (*a*).

Let $\tilde{H}_1(f)$ denote the complex transfer function of the slope circuit defined by Figure 4.13*a*. This complex transfer function is related to $H_1(f)$ by

$$\tilde{H}_1(f - f_c) = 2H_1(f), \qquad f > 0 \tag{4.47}$$

Hence, using Eqs. (4.46) and (4.47), we get

$$\tilde{H}_1(f) = \begin{cases} j4\pi a\left(f + \dfrac{B_T}{2}\right), & -\dfrac{B_T}{2} \le f \le \dfrac{B_T}{2} \\ 0, & \text{elsewhere} \end{cases} \tag{4.48}$$

which is plotted in Figure 4.13*b*.

The incoming FM signal $s(t)$ is defined by Eq. (4.8), which is reproduced here for convenience:

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau)\, d\tau\right]$$

Given that the carrier frequency f_c is high compared to the transmission bandwidth of the FM signal $s(t)$, the complex envelope of $s(t)$ is

$$\tilde{s}(t) = A_c \exp\left[j2\pi k_f \int_0^t m(\tau)\, d\tau\right] \tag{4.49}$$

Let $\tilde{s}_1(t)$ denote the complex envelope of the response of the slope circuit defined by Figure 4.13*b* due to $\tilde{s}(t)$. Then, according to the theory described in Section 2.10, we may express the Fourier transform of $\tilde{s}_1(t)$ as follows:

$$\begin{aligned} \tilde{S}_1(f) &= \frac{1}{2}\tilde{H}_1(f)\tilde{S}(f) \\ &= \begin{cases} j2\pi a\left(f + \dfrac{B_T}{2}\right)\tilde{S}(f), & -\dfrac{B_T}{2} \le f \le \dfrac{B_T}{2} \\ 0, & \text{elsewhere} \end{cases} \end{aligned} \tag{4.50}$$

where $\tilde{S}(f)$ is the Fourier transform of $\tilde{s}(t)$. Since multiplication of the Fourier transform of a signal by the factor $j2\pi f$ is equivalent to differentiating the signal in the time domain (see Section 2.3), we deduce from Eq. (4.50) that

$$\tilde{s}_1(t) = a\left[\frac{d\tilde{s}(t)}{dt} + j\pi B_T \tilde{s}(t)\right] \tag{4.51}$$

Substituting Eq. (4.49) in (4.51) we get

$$\tilde{s}_1(t) = j\pi B_T a A_c \left[1 + \frac{2k_f}{B_T} m(t)\right] \exp\left[j2\pi k_f \int_0^t m(\tau)\, d\tau\right] \tag{4.52}$$

The desired response of the slope circuit is therefore

$$\begin{aligned} s_1(t) &= \mathrm{Re}[\tilde{s}_1(t)\exp(j2\pi f_c t)] \\ &= \pi B_T a A_c \left[1 + \frac{2k_f}{B_T} m(t)\right] \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau)\, d\tau + \frac{\pi}{2}\right] \end{aligned} \tag{4.53}$$

The signal $s_1(t)$ is a hybrid-modulated signal, in which both amplitude and frequency of the carrier wave vary with the message signal $m(t)$. However, provided that we choose

$$\left|\frac{2k_f}{B_T} m(t)\right| < 1 \qquad \text{for all } t$$

then we may use an envelope detector to recover the amplitude variations and thus, except for a bias term, obtain the original message signal. The resulting envelope-detector

output is therefore

$$|\tilde{s}_1(t)| = \pi B_T a A_c \left[1 + \frac{2k_f}{B_T} m(t)\right] \tag{4.54}$$

The bias term $\pi B_T a A_c$ in the right-hand side of Eq. (4.54) is proportional to the slope a of the transfer function of the slope circuit. This suggests that the bias may be removed by subtracting from the envelope-detector output $|\tilde{s}_1(t)|$ the output of a second envelope detector preceded by the *complementary slope circuit* with the transfer function $H_2(f)$ plotted in Figure 4.13*c*. That is, the respective complex transfer functions of the two slope circuits are related by

$$\tilde{H}_2(f) = \tilde{H}_1(-f) \tag{4.55}$$

Let $s_2(t)$ denote the response of the complementary slope circuit produced by the incoming FM signal $s(t)$. Then, following a procedure similar to that just described, we find that the envelope of $s_2(t)$ is

$$|\tilde{s}_2(t)| = \pi B_T a A_c \left[1 - \frac{2k_f}{B_T} m(t)\right] \tag{4.56}$$

where $\tilde{s}_2(t)$ is the complex envelope of the signal $s_2(t)$. The difference between the two envelopes in Eqs. (4.54) and (4.56) is

$$\begin{aligned} s_o(t) &= |\tilde{s}_1(t)| - |\tilde{s}_2(t)| \\ &= 4\pi k_f a A_c m(t) \end{aligned} \tag{4.57}$$

which is free from bias, as desired.

We may thus model the *ideal frequency discriminator* as a pair of slope circuits with their complex transfer functions related by Eq. (4.55), followed by envelope detectors and finally a summer, as in Figure 4.14*a*. This scheme is called a *balanced frequency discriminator*.

The idealized scheme of Figure 4.14*a* can be closely realized using the circuit shown in Figure 4.14*b*. The upper and lower resonant filter sections of this circuit are tuned to frequencies above and below the unmodulated carrier frequency f_c, respectively. In Figure 4.14*c* we have plotted the amplitude responses of these two tuned filters, together with their total response, assuming that both filters have a high Q-factor. The *quality factor* or *Q-factor* of a resonant circuit is a measure of goodness of the whole circuit. It is formally defined as 2π times the ratio of maximum energy stored in the circuit during one cycle to the energy dissipated per cycle. In the case of an RLC parallel (or series) resonant circuit, the Q-factor is equal to the resonant frequency divided by the 3-dB bandwidth of the circuit. In the RLC parallel resonant circuits shown in Figure 4.14*b*, the resistance R is contributed largely by imperfections in the inductive elements of the circuits.

The linearity of the useful portion of the total response in Figure 4.14*c*, centered at f_c, is determined by the separation of the two resonant frequencies. As illustrated in Figure 4.14*c*, a frequency separation of $3B$ gives satisfactory results, where $2B$ is the 3-dB bandwidth of either filter. However, there will be distortion in the output of this frequency discriminator due to the following factors:

1. The spectrum of the input FM signal $s(t)$ is not exactly zero for frequencies outside the range $f_c - B_T/2 \le f \le f_c + B_T/2$.
2. The tuned filter outputs are not strictly band limited, and so some distortion is introduced by the low-pass RC filters following the diodes in the envelope detectors.
3. The tuned filter characteristics are not linear over the whole frequency band of the input FM signal $s(t)$.

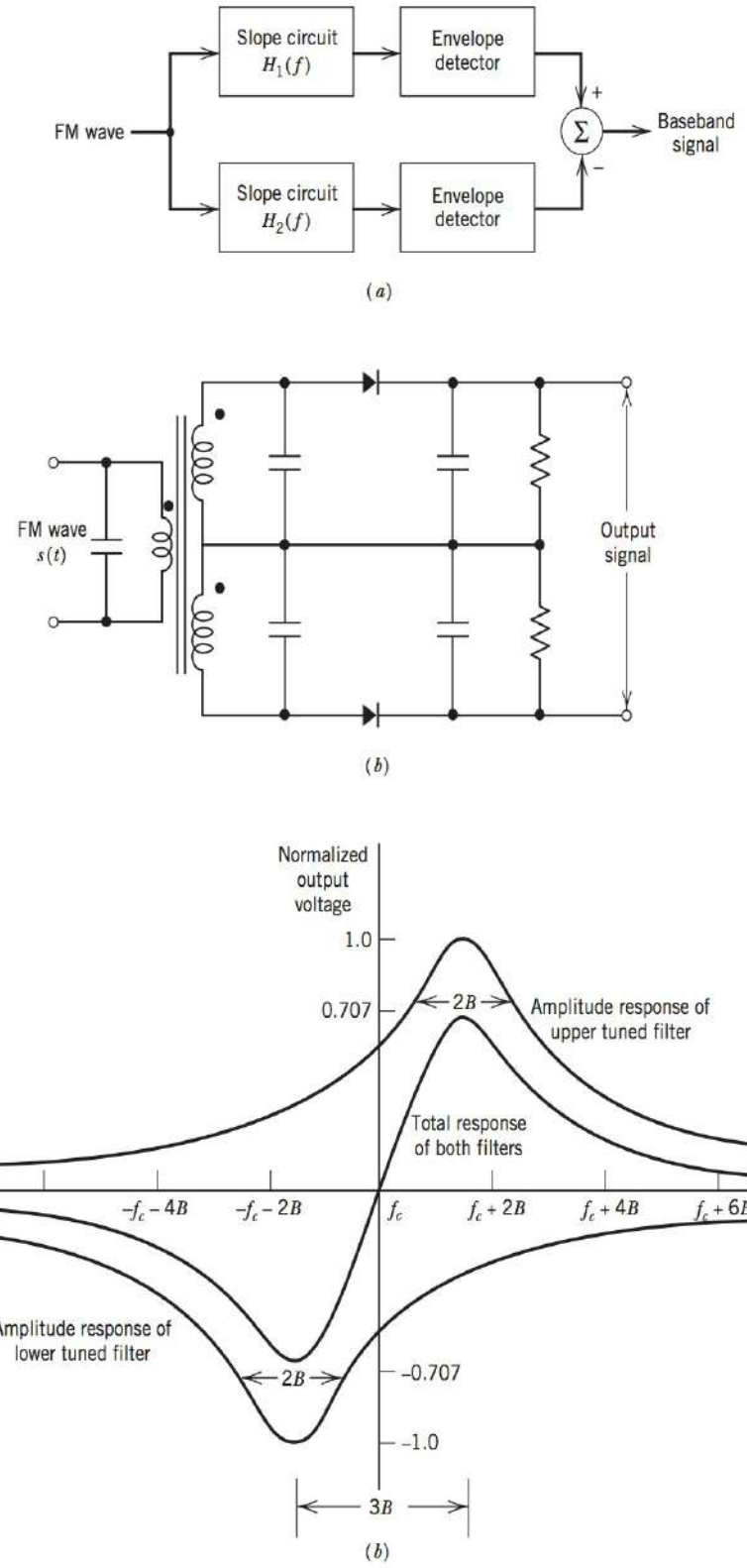


FIGURE 4.14 Balanced frequency discriminator. (*a*) Block diagram. (*b*) Circuit diagram. (*c*) Frequency response.

Nevertheless, by proper design, it is possible to maintain the FM distortion produced by these factors within tolerable limits.

FM STEREO MULTIPLEXING

Stereo multiplexing is a form of frequency-division multiplexing (FDM) designed to transmit two separate signals via the same carrier. It is widely used in FM radio broadcasting to send two different elements of a program (e.g., two different sections of an orchestra, a vocalist and an accompanist) so as to give a spatial dimension to its perception by a listener at the receiving end.

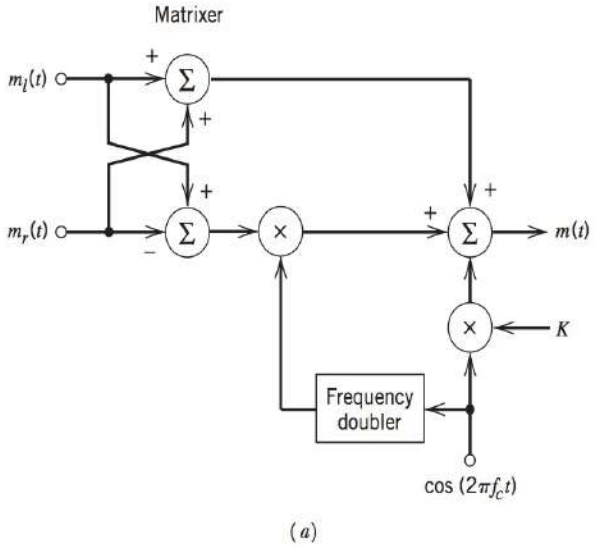
The specification of standards for FM stereo transmission is influenced by two factors:

1. The transmission has to operate within the allocated FM broadcast channels.
2. It has to be compatible with monophonic radio receivers.

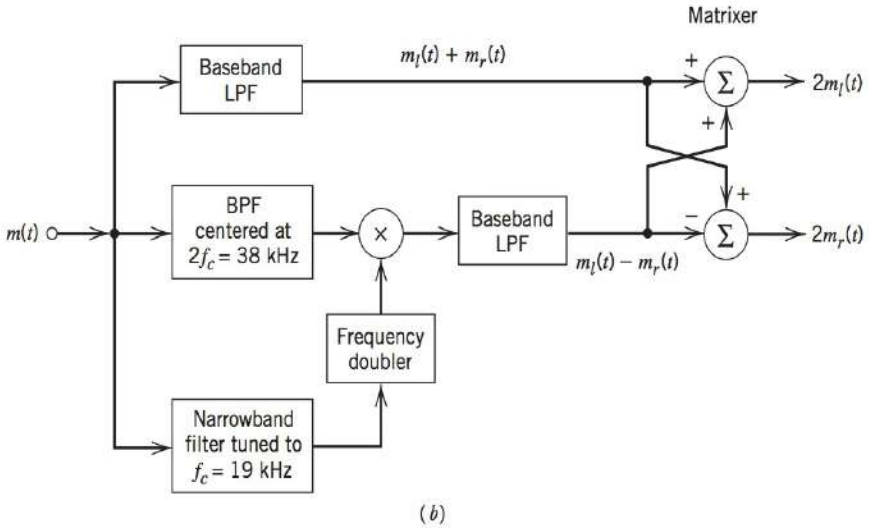
The first requirement sets the permissible frequency parameters, including frequency deviation. The second requirement constrains the way in which the transmitted signal is configured.

Figure 4.15*a* shows the block diagram of the multiplexing system used in an FM stereo transmitter. Let $m_l(t)$ and $m_r(t)$ denote the signals picked up by left-hand and right-hand microphones at the transmitting end of the system. They are applied to a simple *matrixer* that generates the *sum signal*, $m_l(t) + m_r(t)$, and the *difference signal*, $m_l(t) - m_r(t)$. The sum signal is left unprocessed in its baseband form; it is available for monophonic reception. The difference signal and a 38-kHz subcarrier (derived from a 19-kHz crystal oscillator by frequency doubling) are applied to a product modulator, thereby producing a DSB-SC modulated wave. In addition to the sum signal and this DSB-SC modulated wave, the multiplexed signal $m(t)$ also includes a 19-kHz pilot to provide a reference for the coherent detection of the difference signal at the stereo receiver. Thus the multiplexed signal is described by

$$m(t) = [m_l(t) + m_r(t)] + [m_l - m_r(t)]\cos(4\pi f_c t) + K\cos(2\pi f_c t) \tag{4.58}$$



(*a*)



(*b*)

FIGURE 4.15 (*a*) Multiplexer in transmitter of FM stereo. (*b*) Demultiplexer in receiver of FM stereo.

where f_c = 19 kHz, and K is the amplitude of the pilot tone. The multiplexed signal $m(t)$ then frequency-modulates the main carrier to produce the transmitted signal. The pilot is allotted between 8 and 10 percent of the peak frequency deviation; the amplitude K in Eq. (4.58) is chosen to satisfy this requirement.

At a stereo receiver, the multiplexed signal $m(t)$ is recovered by frequency demodulating the incoming FM wave. Then $m(t)$ is applied to the *demultiplexing system* shown in Figure 4.15*b*. The individual components of the multiplexed signal $m(t)$ are separated by the use of three appropriate filters. The recovered pilot (using a narrowband filter tuned to 19 kHz) is frequency doubled to produce the desired 38-kHz subcarrier. The availability of this subcarrier enables the coherent detection of the DSB-SC modulated wave, thereby recovering the difference signal, $m_l(t) - m_r(t)$. The baseband low-pass filter in the top path of Figure 4.15*b* is designed to pass the sum signal, $m_l(t) + m_r(t)$. Finally, the simple matrixer reconstructs the left-hand signal $m_l(t)$ and right-hand signal $m_r(t)$ and applies them to their respective speakers.

4.4 PHASE-LOCKED LOOP

The *phase-locked loop* (PLL) is a negative feedback system, the operation of which is closely linked to frequency modulation. It can be used for synchronization, frequency division/multiplication, frequency modulation, and indirect frequency demodulation. The latter application is the subject of interest here.

Basically, the phase-locked loop consists of three major components: a *multiplier,* a *loop filter*, and a *voltage-controlled oscillator* (VCO) connected together in the form of a feedback loop, as in Figure 4.16. The VCO is a sinusoidal generator whose frequency is determined by a voltage applied to it from an external source. In effect, any frequency modulator may serve as a VCO.

We assume that initially we have adjusted the VCO so that when the control voltage is zero, two conditions are satisfied:

1. The frequency of the VCO in precisely set at the unmodulated carrier frequency f_c.
2. The VCO output has a 90-degree phase-shift with respect to the unmodulated carrier wave.

Suppose then that the input signal applied to the phase-locked loop is an FM signal defined by

$$s(t) = A_c \sin[2\pi f_c t + \phi_1(t)] \tag{4.59}$$

where A_c is the carrier amplitude. With a modulating signal $m(t)$, the angle $\phi_1(t)$ is related to $m(t)$ by the integral

$$\phi_1(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau \tag{4.60}$$

where k_f is the frequency sensitivity of the frequency modulator. Let the VCO output in the phase-locked loop be defined by

$$r(t) = A_v \cos[2\pi f_c t + \phi_2(t)] \tag{4.61}$$

where A_v is the amplitude. With a control voltage $v(t)$ applied to the VCO input, the angle $\phi_2(t)$ is related to $v(t)$ by the integral

$$\phi_2(t) = 2\pi k_v \int_0^t v(t)\, dt \tag{4.62}$$

where k_v is the frequency sensitivity of the VCO, measured in Hertz per volt.

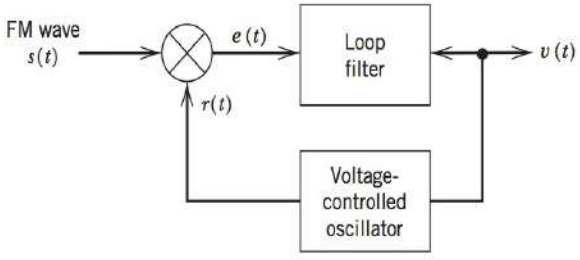


FIGURE 4.16 Phase-locked loop.

The object of the phase-locked loop is to generate a VCO output $r(t)$ that has the same phase angle (except for the fixed difference of 90 degrees) as the input FM signal $s(t)$. The time-varying phase angle $\phi_1(t)$ characterizing $s(t)$ may be due to modulation by a message signal $m(t)$ as in Eq. (4.60), in which case we wish to recover $\phi_1(t)$ in order to estimate $m(t)$. In other applications of the phase-locked loop, the time-varying phase angle $\phi_1(t)$ of the incoming signal $s(t)$ may be an unwanted phase shift caused by fluctuations in the communication channel; in this latter case, we wish to *track* $\phi_1(t)$ so as to produce a signal with the same phase angle for the purpose of coherent detection (synchronous demodulation).

To develop an understanding of the phase-locked loop, it is desirable to have a *model* of the loop. In what follows, we first develop a nonlinear model, which is subsequently linearized to simplify the analysis.

NONLINEAR MODEL OF THE PHASE-LOCKED LOOP[2]

According to Figure 4.16, the incoming FM signal $s(t)$ and the VCO output $r(t)$ are applied to the multiplier, producing two components:

1. A high-frequency component, represented by the *double-frequency* term

$$k_m A_c A_v \sin[4\pi f_c t + \phi_1(t) + \phi_2(t)]$$

2. A low-frequency component represented by the *difference-frequency* term

$$k_m A_c A_v \sin[\phi_1(t) - \phi_2(t)]$$

where k_m is the *multiplier gain*, measured in volt^{-1}.

The loop filter in the phase-lock loop is a low-pass filter, and its response to the high-frequency component will be negligible. The VCO also contributes to the attenuation of this component. Therefore, discarding the high-frequency component (i.e., the double-frequency term), the input to the loop filter is reduced to

$$e(t) = k_m A_c A_v \sin[\phi_e(t)] \tag{4.63}$$

where $\phi_e(t)$ is the *phase error* defined by

$$\begin{aligned} \phi_e(t) &= \phi_1(t) - \phi_2(t) \\ &= \phi_1(t) - 2\pi k_v \int_0^t v(\tau)\, d\tau \end{aligned} \tag{4.64}$$

The loop filter operates on the input $e(t)$ to produce an output $v(t)$ defined by the convolution integral

$$v(t) = \int_{-\infty}^{\infty} e(\tau) h(t - \tau)\, d\tau \tag{4.65}$$

where $h(t)$ is the impulse response of the loop filter.

Using Eqs. (4.62) to (4.64) to relate $\phi_e(t)$ and $\phi_1(t)$, we obtain the following nonlinear integro-differential equation as the descriptor of the dynamic behavior of the phase-locked loop:

$$\frac{d\phi_e(t)}{dt} = \frac{d\phi_1(t)}{dt} - 2\pi K_0 \int_{-\infty}^{\infty} \sin[\phi_e(\tau)] h(t - \tau)\, d\tau \tag{4.66}$$

where K_0 is a *loop-gain parameter* defined by

$$K_0 = k_m k_v A_c A_v \tag{4.67}$$

The amplitudes A_c and A_v are both measured in volts, the multiplier gain k_m in volt^{-1} and the frequency sensitivity k_v in Hertz per volt. Hence, it follows from Eq. (4.67) that K_0

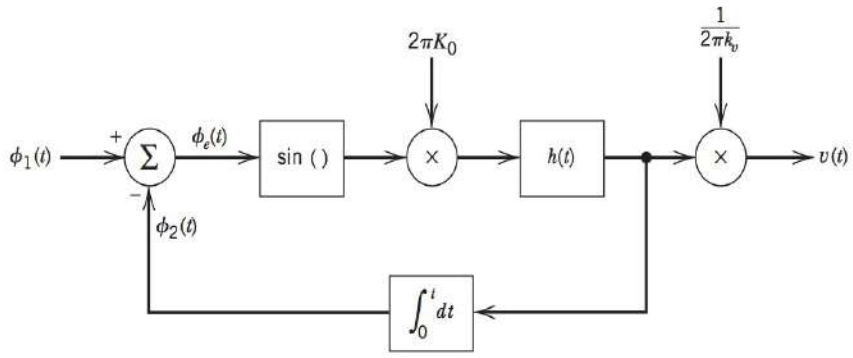


FIGURE 4.17 Nonlinear model of the phase-locked loop.

has the dimensions of frequency. Equation (4.66) suggests the model shown in Figure 4.17 for a phase-locked loop. In this model we have also included the relationship between $v(t)$ and $e(t)$ as represented by Eqs. (4.63) and (4.65). We see that the model resembles the block diagram of Figure 4.16. The multiplier at the input of the phase-locked loop is replaced by a subtracter and a sinusoidal nonlinearity, and the VCO by an integrator.

The sinusoidal nonlinearity in the model of Figure 4.17 greatly increases the difficulty of analyzing the behavior of the phase-locked loop. It would be helpful to *linearize* this model to simplify the analysis, yet give a good approximate description of the loop's behavior in certain modes of operation. This we do next.

LINEAR MODEL OF THE PHASE-LOCKED LOOP

When the phase error $\phi_e(t)$ is zero, the phase-locked loop is said to be in *phase-lock*. When $\phi_e(t)$ is at all times small compared with one radian, we may use the approximation

$$\sin[\phi_e(t)] \simeq \phi_e(t) \tag{4.68}$$

which is accurate to within 4 percent for $\phi_e(t)$ less than 0.5 radians. In this case, the loop is said to be *near phase-lock*, and the sinusoidal nonlinearity of Figure 4.17 may be disregarded. Thus, we may represent the phase-locked loop by the linearized model shown in Figure 4.18*a*. According to this model, the phase error $\phi_e(t)$ is

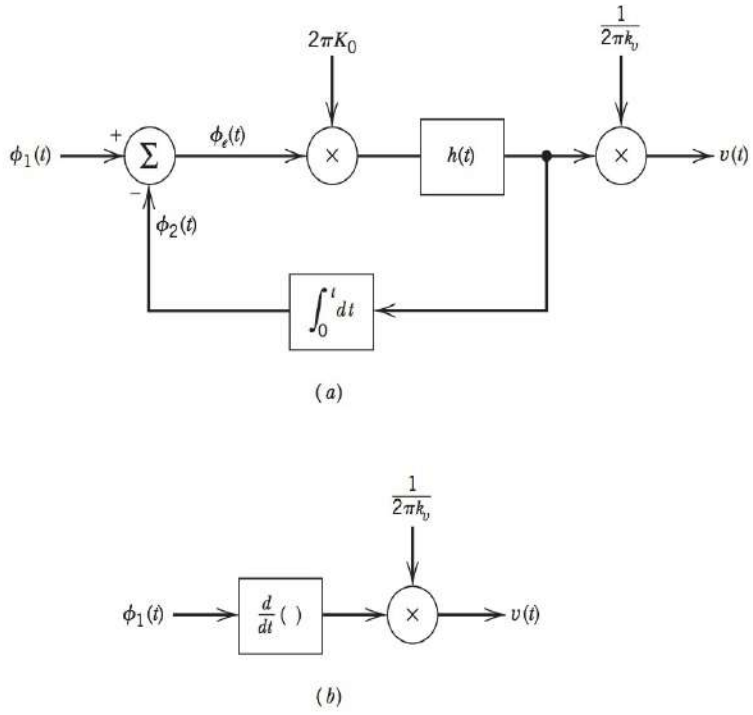


FIGURE 4.18 Models of the phase-locked loop. (*a*) Linearized model. (*b*) Simplified model when the loop gain is very large compared to unity.

related to the input phase $\phi_1(t)$ by the *linear integro-differential equation*

$$\frac{d\phi_e(t)}{dt} + 2\pi K_0 \int_{-\infty}^{\infty} \phi_e(\tau) h(t-\tau)\, d\tau = \frac{d\phi_1(t)}{dt} \tag{4.69}$$

Transforming Eq. (4.69) into the frequency domain and solving for $\Phi_e(f)$, the Fourier transform of $\phi_e(t)$, in terms of $\Phi_1(f)$, the Fourier transform of $\phi_1(t)$, we get

$$\Phi_e(f) = \frac{1}{1 + L(f)} \Phi_1(f) \tag{4.70}$$

The function $L(f)$ in Eq. (4.70) is defined by

$$L(f) = K_0 \frac{H(f)}{jf} \tag{4.71}$$

where $H(f)$ is the transfer function of the loop filter. The quantity $L(f)$ is called the *open-loop transfer function* of the phase-locked loop. Suppose that for all values of f inside the baseband we make the magnitude of $L(f)$ very large compared with unity. Then from Eq. (4.70) we find that $\Phi_e(f)$ approaches zero. That is, the phase of the VCO becomes asymptotically equal to the phase of the incoming signal. Under this condition, phase-lock is established, and the objective of the phase-locked loop is thereby satisfied.

From Figure 4.18*a* we see that $V(f)$, the Fourier transform of the phase-locked loop output $v(t)$, is related to $\Phi_e(f)$ by

$$V(f) = \frac{K_0}{k_v} H(f) \Phi_e(f) \tag{4.72}$$

Equivalently, in light of Eq. (4.71), we may write

$$V(f) = \frac{jf}{k_v} L(f) \Phi_e(f) \tag{4.73}$$

Therefore, substituting Eq. (4.70) in (4.73), we get

$$V(f) = \frac{(jf/k_v) L(f)}{1 + L(f)} \Phi_1(f) \tag{4.74}$$

Again, when we make $|L(f)| \gg 1$ for the frequency band of interest, we may approximate Eq. (4.74) as follows:

$$V(f) \simeq \frac{jf}{k_v} \Phi_1(f) \tag{4.75}$$

The corresponding time-domain relation is

$$v(t) \simeq \frac{1}{2\pi k_v} \frac{d\phi_1(t)}{dt} \tag{4.76}$$

Thus, provided that the magnitude of the open-loop transfer function $L(f)$ is very large for all frequencies of interest, the phase-locked loop may be modeled as a *differentiator* with its output scaled by the factor $1/2\pi k_v$, as in Figure 4.18*b*.

The simplified model of Figure 4.18*b* provides an indirect method of using the phase-locked loop as a frequency demodulator. When the input is an FM signal as in Eq. (4.59), the angle $\phi_1(t)$ is related to the message signal $m(t)$ as in Eq. (4.60). Therefore, substituting Eq. (4.60) in (4.76), we find that the resulting output signal of the phase-locked loop is approximately

$$v(t) \simeq \frac{k_f}{k_v} m(t) \tag{4.77}$$

Equation (4.77) states that when the loop operates in its phase-locked mode, the output $v(t)$ of the phase-locked loop is approximately the same, except for the scale factor k_f/k_v, as the original message signal $m(t)$; frequency demodulation of the incoming FM signal $s(t)$ is thereby accomplished.

A significant feature of the phase-locked loop acting as a demodulator is that the bandwidth of the incoming FM signal can be much wider than that of the loop filter characterized by $H(f)$. The transfer function $H(f)$ can and should be restricted to the baseband. Then the control signal of the VCO has the bandwidth of the baseband (message) signal $m(t)$, whereas the VCO output is a wideband frequency-modulated signal whose instantaneous frequency tracks that of the incoming FM signal. Here we are merely restating the fact that the bandwidth of a wide-band FM signal is much larger than the bandwidth of the message signal responsible for its generation.

The complexity of the phase-locked loop is determined by the transfer function $H(f)$ of the loop filter. The simplest form of a phase-locked loop is obtained when $H(f) = 1$; that is, there is no loop filter, and the resulting phase-locked loop is referred to as a *first-order phase-locked loop*. For higher-order loops, the transfer function $H(f)$ assumes a more complex form. The order of the phase-locked loop is determined by the order of the denominator polynomial of the *closed-loop transfer function*, which defines the output transform $V(f)$ in terms of the input transform $\Phi_1(f)$, as shown in Eq. (4.74).

A major limitation of a first-order phase-locked loop is that the loop gain parameter K_0 controls both the loop bandwidth as well as the hold-in frequency range of the loop; the *hold-in frequency range* refers to the range of frequencies for which the loop remains phase-locked to the input signal. It is for this reason that a first-order phase-locked loop is seldom used in practice. Accordingly, in the remainder of this section we deal only with a second-order phase-locked loop.

SECOND-ORDER PHASE-LOCKED LOOP

To be specific, consider a *second-order phase-locked loop* using a loop filter with the transfer function

$$H(f) = 1 + \frac{a}{jf} \tag{4.78}$$

where a is a constant. The filter consists of an integrator (using an operational amplifier) and a direct connection, as shown in Figure 4.19. For this phase-locked loop, the use of Eqs. (4.70) and (4.78) yields

$$\Phi_e(f) = \frac{(jf)^2/aK_0}{1 + [(jf)/a] + [(jf)^2/aK_0]}\Phi_1(f) \tag{4.79}$$

Define the *natural frequency* of the loop:

$$f_n = \sqrt{aK_0} \tag{4.80}$$

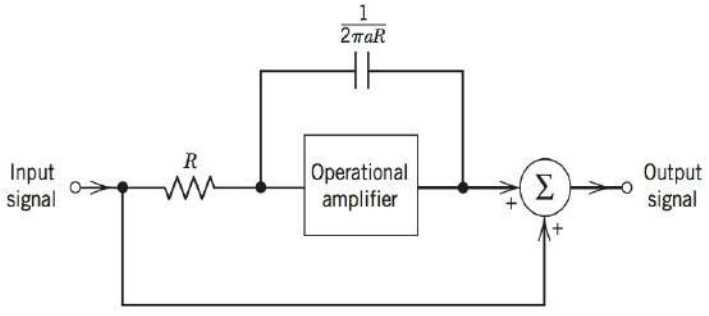


FIGURE 4.19 Loop filter for second-order phase-locked loop.

and the *damping factor*:

$$\zeta = \sqrt{\frac{K_0}{4a}} \tag{4.81}$$

Then we may recast Eq. (4.79) in terms of the parameters f_n and ζ as follows:

$$\Phi_e(f) = \left(\frac{(jf/f_n)^2}{1 + 2\zeta(jf/f_n) + (jf/f_n)^2} \right) \Phi_1(f) \tag{4.82}$$

Assume that the incoming FM signal is produced by a single-tone modulating wave, for which the phase input is

$$\phi_1(t) = \beta \sin(2\pi f_m t) \tag{4.83}$$

Hence, from Eq. (4.82) we find that the corresponding phase error is

$$\phi_e(t) = \phi_{e0} \cos(2\pi f_m t + \psi) \tag{4.84}$$

where the amplitude ϕ_{e0} and phase ψ are, respectively, defined by

$$\phi_{e0} = \frac{(\Delta f/f_n)(f_m/f_n)}{\{[1 - (f_m/f_n)^2]^2 + 4\zeta^2(f_m/f_n)^2\}^{1/2}} \tag{4.85}$$

and

$$\psi = \frac{\pi}{2} - \tan^{-1}\left[\frac{2\zeta f_m/f_n}{1 - (f_m/f_n)^2} \right] \tag{4.86}$$

In Figure 4.20 we have plotted the phase error amplitude ϕ_{e0}, normalized with respect to $\Delta f/f_n$, versus f_m/f_n for different values of ζ. It is apparent that for all values of the damping factor ζ, and assuming a fixed frequency deviation Δf, the phase error is small at low modulation frequencies, rises to a maximum at $f_m = f_n$, and then falls off at higher modulation frequencies. Note also that the maximum value of phase error amplitude decreases with increasing ζ.

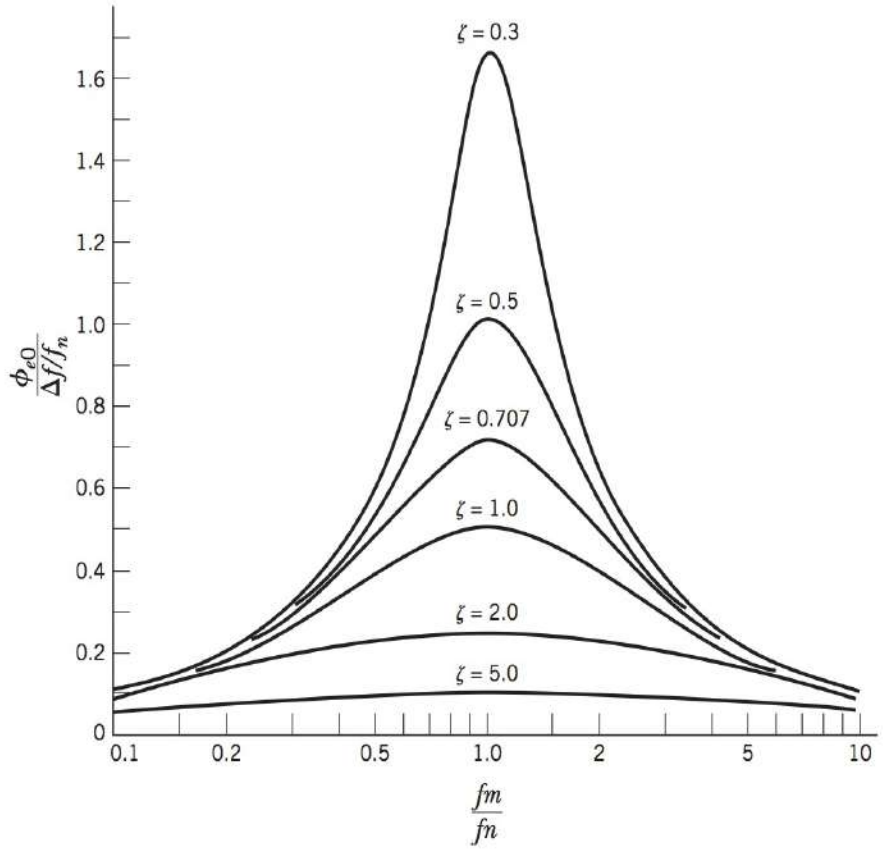


FIGURE 4.20 Phase-error amplitude characteristic of second-order phase-locked loop.

The Fourier transform of the loop output is related to $\Phi_e(f)$ by Eq. (4.72); hence, with $H(f)$ as defined in Eq. (4.78), we get

$$V(f) = \frac{K_0}{k_v}\left(1 + \frac{a}{jf}\right)\Phi_e(f) \tag{4.87}$$

In light of the definitions given in Eqs. (4.80) and (4.81), we have

$$V(f) = \left(\frac{f_n^2}{jfk_v}\right)\left[1 + 2\zeta\left(\frac{jf}{f_n}\right)\right]\Phi_e(f) \tag{4.88}$$

Substituting Eq. (4.82) in (4.88), we get

$$V(f) = \left(\frac{(jf/k_v)[1 + 2\zeta(jf/f_n)]}{1 + 2\zeta(jf/f_n) + (jf/f_n)^2}\right)\Phi_1(f) \tag{4.89}$$

Therefore, for the phase input $\phi_1(t)$ of Eq. (4.83), we find that the corresponding loop output is

$$v(t) = A_0 \cos(2\pi f_m t + \alpha) \tag{4.90}$$

where the amplitude A_0 and phase α are, respectively, defined by

$$A_0 = \frac{(\Delta f/k_v)[1 + 4\zeta^2(f_m/f_n)^2]^{1/2}}{\{[1 - (f_m/f_n)^2]^2 + 4\zeta^2(f_m/f_n)^2\}^{1/2}} \tag{4.91}$$

and

$$\alpha = \tan^{-1}\left[2\zeta\left(\frac{f_m}{f_n}\right)\right] - \tan^{-1}\left[\frac{2\zeta(f_m/f_n)}{1 - (f_m/f_n)^2}\right] \tag{4.92}$$

From Eq. (4.91), we see that the amplitude A_0 attains its maximum value of $\Delta f/k_v$ at $(f_m/f_n) = 0$; it decreases with increasing f_m/f_n, dropping to zero at $(f_m/f_n) = \infty$.

The important feature of the second-order phase-locked loop is that with an incoming FM signal produced by a modulating sinusoidal wave of fixed amplitude (corresponding to a fixed frequency deviation) and varying frequency, the frequency response that defines the phase error $\phi_e(t)$ is representative of a band-pass filter [see Eq. (4.85)], but the frequency response that defines the loop output $v(t)$ is representative of a low-pass filter [see Eq. (4.91)]. Therefore, by appropriately choosing the parameters ζ and f_n, which determine the frequency response of the loop, it is possible to restrain the phase error to always remain small and thereby lie within the linear range of the loop, whereas at the same time the modulating (message) signal is reproduced at the loop output with minimum distortion. This restraint is, however, conservative with respect to the hold-in capabilities of the loop. As a reasonable rule of thumb, the loop should remain locked if the maximum value of the phase error ϕ_{e0} (which occurs when the modulation frequency f_m is equal to the loop's natural frequency f_n) is always less than 90 degrees.

Phase-locked loop performance is explored experimentally in Problem 4.29.

4.5 NONLINEAR EFFECTS IN FM SYSTEMS

In the preceding three sections, we studied frequency modulation theory and methods for its generation and demodulation. We complete the discussion of frequency modulation by considering nonlinear effects in FM systems.

Nonlinearities, in one form or another, are present in all electrical networks. There are two basic forms of nonlinearity to consider:

1. The nonlinearity is said to be *strong* when it is introduced intentionally and in a controlled manner for some specific application. Examples of strong nonlinearity include square-law modulators, limiters, and frequency multipliers.
2. The nonlinearity is said to be *weak* when a linear performance is desired, but nonlinearities of a parasitic nature arise due to imperfections. The effect of such weak nonlinearities is to limit the useful signal levels in a system and thereby become an important design consideration.

In this section we examine the effects of weak nonlinearities on frequency modulation.[3]

Consider a communications channel, the transfer characteristic of which is defined by the nonlinear input–output relation

$$v_o(t) = a_1 v_i(t) + a_2 v_i^2(t) + a_3 v_i^3(t) \tag{4.93}$$

where $v_i(t)$ and $v_o(t)$ are the input and output signals, respectively, and a_1, a_2 and a_3 are constants. The channel described in Eq. (4.93) is said to be *memoryless* in that the output signal $v_o(t)$ is an instantaneous function of the input signal $v_i(t)$ (i.e., there is no energy storage involved in the description). We wish to determine the effect of transmitting a frequency-modulated wave through such a channel. The FM signal is defined by

$$v_i(t) = A_c \cos[2\pi f_c t + \phi(t)]$$

where

$$\phi(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau$$

For this input signal, the use of Eq. (4.93) yields

$$\begin{aligned} v_o(t) = {} & a_1 A_c \cos[2\pi f_c t + \phi(t)] + a_2 A_c^2 \cos^2[2\pi f_c t + \phi(t)] \\ & + a_3 A_c^3 \cos^3[2\pi f_c t + \phi(t)] \end{aligned} \tag{4.94}$$

Expanding the squared and cubed cosine terms in Eq. (4.94) and then collecting common terms, we get

$$\begin{aligned} v_o(t) = {} & \frac{1}{2} a_2 A_c^2 + \left(a_1 A_c + \frac{3}{4} a_3 A_c^3\right) \cos[2\pi f_c t + \phi(t)] \\ & + \frac{1}{2} a_2 A_c^2 \cos[4\pi f_c t + 2\phi(t)] \\ & + \frac{1}{4} a_3 A_c^3 \cos[6\pi f_c t + 3\phi(t)] \end{aligned} \tag{4.95}$$

Thus the channel output consists of a dc component and three frequency-modulated signals with carrier frequencies f_c, $2f_c$, and $3f_c$; the sinusoidal components are contributed by the linear, second-order, and third-order terms of Eq. (4.93), respectively.

To extract the desired FM signal from the channel output $v_o(t)$, that is, the particular component with carrier frequency f_c, it is necessary to separate the FM signal with this carrier frequency from the one with the closest carrier frequency: $2f_c$. Let Δf denote the frequency deviation of the incoming FM signal $v_i(t)$, and W denote the highest frequency component of the message signal $m(t)$. Then, applying Carson's rule and noting that the frequency deviation about the second harmonic of the carrier frequency is doubled, we find that the necessary condition for separating the desired FM signal with the carrier frequency f_c from that with the carrier frequency $2f_c$ is

$$2f_c - (2\Delta f + W) > f_c + \Delta f + W$$

or

$$f_c > 3\Delta f + 2W \tag{4.96}$$

Thus, by using a band-pass filter of mid-band frequency f_c and bandwidth $2\Delta f + 2W$, the channel output is reduced to

$$v_o'(t) = \left(a_1 A_c + \frac{3}{4} a_3 A_c^3\right) \cos[2\pi f_c t + \phi(t)] \tag{4.97}$$

We see therefore that the only effect of passing an FM signal through a channel with amplitude nonlinearities, followed by appropriate filtering, is simply to modify its amplitude. That is, unlike amplitude modulation, frequency modulation is not affected by distortion produced by transmission through a channel with amplitude nonlinearities. It is for this reason that we find frequency modulation widely used in microwave radio and satellite communication systems: It permits the use of highly nonlinear amplifiers and power transmitters, which are particularly important to producing a maximum power output at radio frequencies.

An FM system is extremely sensitive to *phase nonlinearities*, however, as we would intuitively expect. A common type of phase nonlinearity that is encountered in microwave radio systems is known as *AM-to-PM conversion.* This is the result of the phase characteristic of repeaters or amplifiers used in the system being dependent on the instantaneous amplitude of the input signal. In practice, AM-to-PM conversion is characterized by a constant K, which is measured in degrees per dB and may be interpreted as the peak phase change at the output for a 1-dB change in envelope at the input. When an FM wave is transmitted through a microwave radio link, it picks up spurious amplitude variations due to noise and interference during the course of transmission, and when such an FM wave is passed through a repeater with AM-to-PM conversion, the output will contain unwanted phase modulation and resultant distortion. It is therefore important to keep the AM-to-PM conversion at a low level. For example, for a good microwave repeater, the AM-to-PM conversion constant K is less than 2 degrees per dB.

4.6 THE SUPERHETERODYNE RECEIVER

In a communication system, irrespective of whether it is based on amplitude modulation or frequency modulation, the receiver not only has the task of demodulating the incoming modulated signal, but it is also required to perform some other system functions:

- *Carrier-frequency tuning*, the purpose of which is to select the desired signal (i.e., desired radio or TV station).
- *Filtering*, which is required to separate the desired signal from other modulated signals that may be picked up along the way.
- *Amplification*, which is intended to compensate for the loss of signal power incurred in the course of transmission.

The *superheterodyne receiver*, or *superhet* as it is often referred to, is a special type of receiver that fulfils all three functions, particularly the first two, in an elegant and practical fashion. Specifically, it overcomes the difficulty of having to build a tunable high- (and variable-) Q filter. Indeed, practically all analog radio and TV receivers are of the superheterodyne type.

Basically, the receiver consists of a radio-frequency (RF) section, a mixer and local oscillator, an intermediate frequency (IF) section, demodulator, and power amplifier. Typical frequency parameters of commercial AM and FM radio receivers are listed in Table 4.2. Figure 4.21 shows the block diagram of a superheterodyne receiver for amplitude modulation using an envelope detector for demodulation.

TABLE 4.2 Typical frequency parameters of AM and FM radio receivers

	AM Radio	FM Radio
RF carrier range	0.535–1.605 MHz	88–108 MHz
Mid-band frequency of IF section	0.455 MHz	10.7 MHz
IF bandwidth	10 kHz	200 kHz

The incoming amplitude-modulated wave is picked up by the receiving antenna and amplified in the RF section that is tuned to the carrier frequency of the incoming wave. The combination of mixer and local oscillator (of adjustable frequency) provides a *heterodyning* function, whereby the incoming signal is converted to a predetermined fixed *intermediate frequency,* usually lower than the incoming carrier frequency. This frequency translation is achieved without disturbing the relation of the sidebands to the carrier. The result of the heterodyning is to produce an intermediate-frequency carrier defined by

$$f_{IF} = f_{RF} - f_{LO} \tag{4.98}$$

where f_{LO} is the frequency of the local oscillator and f_{RF} is the carrier frequency of the incoming RF signal. We refer to f_{IF} as the intermediate frequency (IF), because the signal is neither at the original input frequency nor at the final baseband frequency. The mixer-local oscillator combination is sometimes referred to as the *first detector*, in which case the demodulator is called the *second detector.*

The IF section consists of one or more stages of tuned amplification, with a bandwidth corresponding to that required for the particular type of signal that the receiver is intended to handle. This section provides most of the amplification and selectivity in the receiver. The output of the IF section is applied to a demodulator, the purpose of which is to recover the baseband signal. If coherent detection is used, then a coherent signal source must be provided in the receiver. The final operation in the receiver is the power amplification of the recovered message signal.

In a superheterodyne receiver the mixer will develop an intermediate frequency output when the input signal frequency is greater or less than the local oscillator frequency by an amount equal to the intermediate frequency. That is, there are two input frequencies, namely, $|f_{LO} \pm f_{IF}|$, which will result in f_{IF} at the mixer output. This introduces the possibility of simultaneous reception of two signals differing in frequency by twice the intermediate frequency. For example, a receiver tuned to 1 MHz and having an IF of 0.455 MHz is subject to an *image interference* at 1.910 MHz; indeed, any receiver with this value of IF, when tuned to any station, is subject to image interference at a frequency of 0.910 MHz higher than the desired station. Since the function of the mixer is to produce the difference between two applied frequencies, it is incapable of distinguishing between the desired signal and its image in that it produces an IF output from either one of

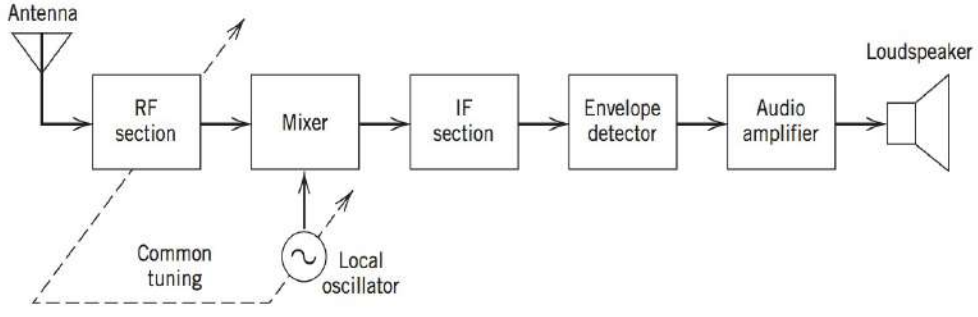


FIGURE 4.21 Basic elements of an AM receiver of the superheterodyne type.

them. The only practical cure for image interference is to employ highly selective stages in the RF section (i.e., between the antenna and the mixer) in order to favor the desired signal and discriminate against the undesired or *image signal*. The effectiveness of suppressing unwanted image signals increases as the number of selective stages in the radio-frequency section increases, and as the ratio of intermediate to signal frequency increases.

The basic difference between AM and FM superheterodyne receivers lies in the use of an FM demodulator such as limiter-frequency discriminator. In an FM system, the message information is transmitted by variations of the instantaneous frequency of a sinusoidal carrier wave, and its amplitude is maintained constant. Therefore, any variations of the carrier amplitude at the receiver input must result from noise or interference. An *amplitude limiter,* following the IF section, is used to remove amplitude variations by clipping the modulated wave at the IF section output almost to the zero axis. The resulting rectangular wave is rounded off by a band-pass filter that suppresses harmonics of the carrier frequency. Thus the filter output is again sinusoidal, with an amplitude that is practically independent of the carrier amplitude at the receiver input (see Problem 4.20).

4.7 THEME EXAMPLE—ANALOG AND DIGITAL FM CELLULAR TELEPHONES

In this example, we consider two applications of an FM modulator that are both related to cellular telephone service. The initial cellular telephone system in North America was known as the *Advanced Mobile Phone Service* (AMPS) and went into operation in 1983. The AMPS system uses 30 kHz channel spacing, that is, two 30 kHz channels, one in each direction is assigned to each user for the duration of call. This method of sharing the radio spectrum is known as *frequency-division multiple access* (FDMA). The two channels (uplink and downlink) are separated by 45 MHz in a frequency band from 824 to 894 MHz.

In AMPS, analog frequency modulation is used for voice transmission, and frequency-shift keying (see Chapter 9) is used for data transmission. As with wired telephone service, the voice bandwidth (W) is limited to approximately 3 kHz before transmission. The FM modulator is designed such that the peak deviation due to the voice is limited to 12 kHz. Using Carson's rule of Eq. (4.38), with $\Delta f = 12$ kHz, and replacing f_m by W, the approximate value of the transmission bandwidth of the AMPS signal is obtained as

$$\begin{aligned} B_T &= 2(\Delta f + W) \\ &= 2(12 + 3) = 30\,\text{kHz} \end{aligned}$$

This transmission bandwidth estimate agrees with the assigned channel spacing of 30 kHz. Since FM is a constant-envelope modulation technique, AMPS mobile units can use high-efficiency power amplifiers. In particular, power amplifiers could be operated in saturation (facilitating high efficiency) without distorting the output envelope, since it is constant. The constant-envelope property also has advantages for combating the fading that occurs on mobile radio links. One serious drawback of the analog FM system was that it provided no protection from eavesdroppers.

The AMPS was the first system to introduce the *cellular* concept for frequency reuse. However, the success of AMPS was the progenitor of its own demise since greater demand for the limited radio spectrum meant that more bandwidth-efficient transmission techniques had to be found.

One of the successors to AMPS is the digital cell-phone standard known as GSM (*Global System for Mobile Communications*). GSM built on some of the FM-related advantages of AMPS but uses a more complex multiplexing strategy and a digital representation of the data to reduce the bandwidth requirements. To understand the

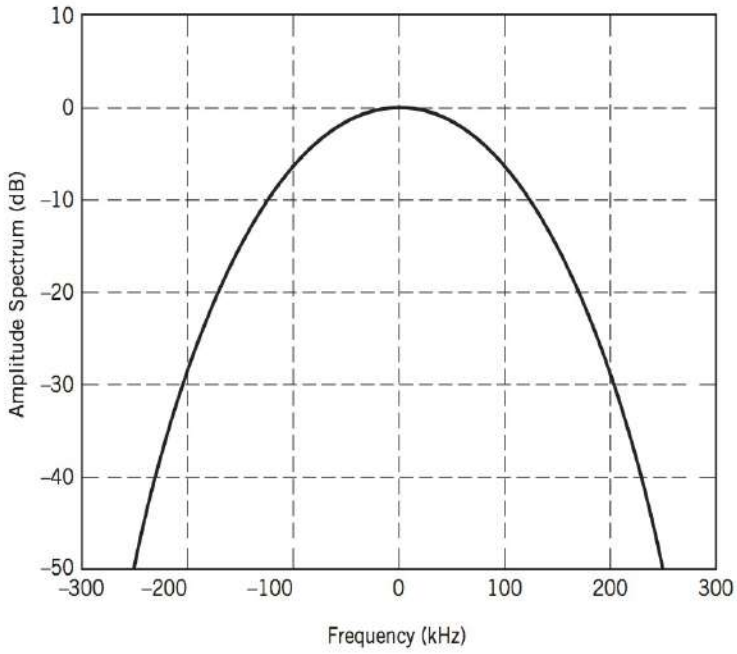


FIGURE 4.22 Spectrum of baseband pulse used in GSM.

FM nature of GSM, recall that the general FM equation is

$$s(t) = A_c \cos\left[2\pi f_c t + k_f \int_0^t m(\tau)\, d\tau\right]$$

where f_c is the carrier frequency and $m(t)$ is the modulating signal. With GSM, the modulating signal is given by the digital signal [see Eq. (2.120)]

$$m(t) = \sum_{k=0}^{K} b_k p(t - kT)$$

where the bits $\{b_k\}$ are the digital representation of an audio (voice) source. The data bits are modulated by a pulse shape that is described by the convolution of two functions

$$p(t) = c \exp[-\pi c^2 t^2] * \text{rect}[t/T]$$

where $*$ denotes convolution, $c = B\sqrt{2\pi/\log(2)}$, and the logarithm is the natural logarithm. For GSM, the product BT is set to 0.3 where the symbol period T is 3.77 microseconds. The amplitude spectrum of the baseband pulse $p(t)$ is shown in Figure 4.22. When the sensitivity coefficient k_f is set to $\pi/2$, this digital modulation is referred to as *Gaussian minimum shift-keying* (GMSK), which is discussed in Chapter 9.

In Figure 4.23, we plot the simulated spectrum of the modulated GSM signal. The 3-dB (one-sided) bandwidth of this signal is approximately 60 kHz. The similarity of the modulated spectrum to that of the baseband pulse shape indicates that GMSK is a form of narrowband frequency modulation.

The GMSK signal of GSM is allocated a bandwidth of 200 kHz, which is substantially larger than the 30 kHz channel allocation of AMPS. However, due to the digital representation of voice, the 200 kHz channel can be simultaneously shared among 32 voice calls in one direction. This multiplexing strategy provides a (30/200) $\times$ 32 = 4.8 times improvement over AMPS in the number of telephone calls that can be serviced per unit bandwidth, which is a greatly improved bandwidth efficiency.

In addition to the frequency band used by AMPS, GSM uses a number of other frequency bands as indicated in Table 4.3. These bands are shared using a FDMA similar to

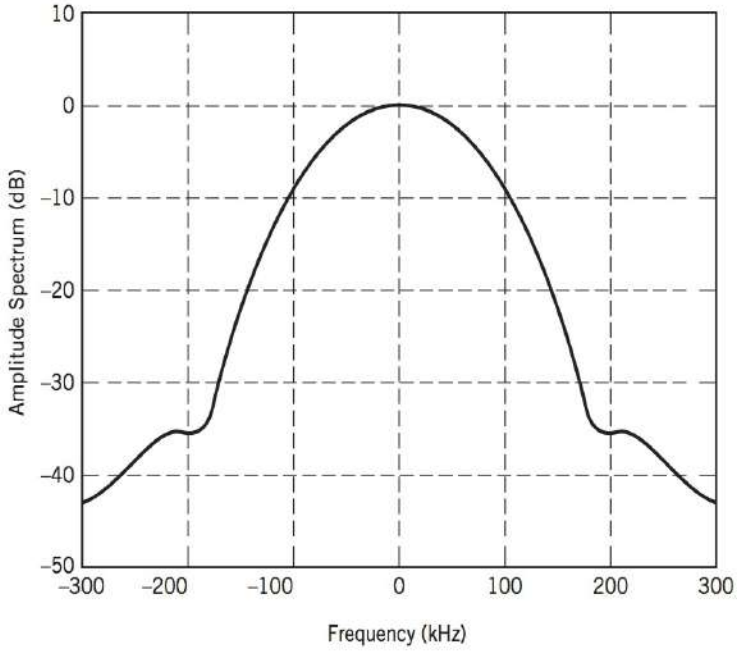


FIGURE 4.23 Spectrum of a GSM signal.

TABLE 4.3 Frequency bands for GSM

Band	Uplink Frequency (MHz)	Downlink Frequency (MHz)	Usage
GSM-850	824–849	869–894	United States, Canada, and most of Americas (also used for AMPS)
GSM-900	890–915	935–960	Europe, Africa, and most of Asia
GSM-1800	1710–1785	1805–1880	Europe, Africa, and most of Asia
GSM-1900	1850–1910	1930–1990	United States, Canada, and most of Americas

AMPS. The individual GSM channels are also shared in time using a strategy known as *time-division multiple access* (TDMA), which will be explained in Chapter 7.

4.8 SUMMARY AND DISCUSSION

In this chapter, we studied the principles of angle modulation, which is a second form of continuous-wave (CW) modulation. Angle modulation uses a sinusoidal carrier whose angle is varied in accordance with a message signal.

Angle modulation may be classified into frequency modulation (FM) and phase modulation (PM). In FM, the instantaneous frequency of a sinusoidal carrier is varied in proportion to the message signal. In PM, on the other hand, it is the phase of the carrier that is varied in proportion to the message signal. The instantaneous frequency is defined as the derivative of phase with respect to time, except for the scaling factor $1/(2\pi)$. Accordingly, FM and PM are closely related to each other. If we know the properties of the one, we can determine those of the other. For this reason, and because FM is commonly used in broadcasting, much of the material on angle modulation in the chapter was devoted to FM.

Unlike amplitude modulation, FM is a nonlinear modulation process. Accordingly, spectral analysis of FM is more difficult than for AM. Nevertheless, by studying single-tone FM, we were able to develop a great deal of insight into the spectral properties of

FM. In particular, we derived an empirical rule known as Carson's rule for an approximate evaluation of the transmission bandwidth B_T of FM. According to this rule, B_T is controlled by a single parameter: the modulation index β for sinusoidal FM, or the deviation radio D for nonsinusoidal FM.

In FM, the carrier amplitude, and therefore the transmitted average power, is maintained constant. Herein lies the important advantage of FM over AM in combating the effects of noise or interference at reception, an issue that we study in Chapter 6, after familiarizing ourselves with probability theory and random processes in Chapter 5. This advantage becomes increasingly more pronounced as the modulation index (deviation ratio) is increased, which has the effect of increasing the transmission bandwidth in a corresponding way. Thus, frequency modulation provides a practical method for the tradeoff of channel bandwidth for improved noise performance, which is not feasible with amplitude modulation.

NOTES AND REFERENCES

1. Bessel functions play an important role in the solution of a certain differential equation and also in the mathematical formulation of many physical problems. For a detailed treatment of the subject, see Wylie and Barrett (1982, pp. 572–625).
2. When a phase-locked loop is used to demodulate an FM wave, the loop must first lock onto the incoming FM wave and then follow the variations in its phase. During the lock-up operation, the phase error $\phi_e(t)$ between the incoming FM wave and the VCO output will be large, thus requiring the use of the nonlinear model of Figure 4.17. For a full treatment of the nonlinear analysis of a phase-locked loop, see Gardner (1979), Egan (1998), and Best (2003).
3. For a detailed discussion of the characterization and system effects of weak nonlinearities, see "Transmission Systems for Communication," Bell Telephone Laboratories, pp. 237–278 (Western Electric, 1971).

PROBLEMS

4.1 Sketch the PM and FM waves produced by the sawtooth wave shown in Figure P4.1.

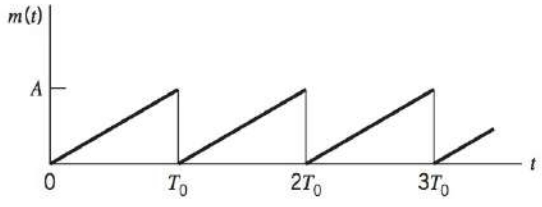


Figure P4.1

4.2 In a *frequency-modulated radar,* the instantaneous frequency of the transmitted carrier is varied as in Figure P4.2, which is obtained by using a triangular modulating signal. The instantaneous frequency of the received echo signal is shown dashed in Figure P4.2, where τ is the round-trip delay time. The transmitted and received echo signals are applied to a mixer, and the difference frequency component is retained. Assuming that $f_0\tau \ll 1$, determine the number of beat cycles at the mixer output, averaged over one second, in terms of the peak deviation Δf of the carrier frequency, the delay τ, and the repetition frequency f_0 of the transmitted signal.

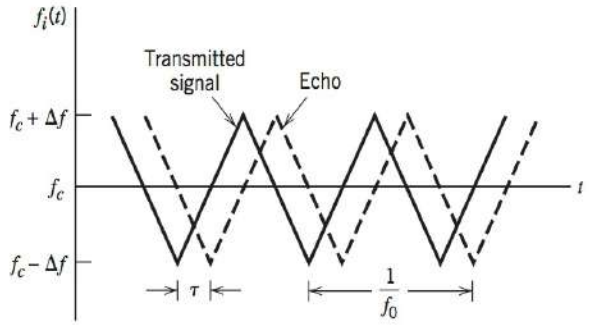


Figure P4.2

4.3 The instantaneous frequency of a sine wave is equal to $f_c - \Delta f$ for $|t| \le T/2$, and f_c for $|t| > T/2$. Determine the spectrum of this frequency-modulated wave. *Hint:* Divide up the time interval of interest into three regions: $-\infty < t < -T/2$, $-T/2 \le t \le T/2$, and $T/2 < t < \infty$.

4.4 Consider a narrow-band FM signal approximately defined by

$$s(t) \simeq A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t)\sin(2\pi f_m t)$$

(a) Determine the envelope of this modulated signal. What is the ratio of the maximum to the minimum value of this envelope? Plot this ratio versus β, assuming that β is restricted to the interval $0 \leq \beta \leq 0.3$.

(b) Determine the average power of the narrow-band FM signal, expressed as a percentage of the average power of the unmodulated carrier wave. Plot this result versus β, assuming that β is restricted to the interval $0 \leq \beta \leq 0.3$.

(c) By expanding the angle $\theta_i(t)$ of the narrow-band FM signal $s(t)$ in the form of a power series, and restricting the modulation index β to a maximum value of 0.3 radians, show that

$$\theta_i(t) \simeq 2\pi f_c t + \beta \sin(2\pi f_m t) - \frac{\beta^3}{3}\sin^3(2\pi f_m t)$$

What is the value of the harmonic distortion for $\beta = 0.3$?

4.5 The sinusoidal modulating wave

$$m(t) = A_m \cos(2\pi f_m t)$$

is applied to a phase modulator with phase sensitivity k_p. The unmodulated carrier wave has frequency f_c and amplitude A_c.

(a) Determine the spectrum of the resulting phase-modulated signal, assuming that the maximum phase deviation $\beta_p = k_p A_m$ does not exceed 0.3 radians.

(b) Construct a phasor diagram for this modulated signal, and compare it with that of the corresponding narrow-band FM signal.

4.6 Suppose that the phase-modulated signal of Problem 4.5 has an arbitrary value for the maximum phase deviation β_p. This modulated signal is applied to an ideal band-pass filter with mid-band frequency f_c and a passband extending from $f_c - 1.5f_m$ to $f_c + 1.5f_m$. Determine the envelope, phase, and instantaneous frequency of the modulated signal at the filter output as functions of time.

4.7 In Section 4.3, we showed how the output of a narrowband modulator may be approximated when the input is a tone of frequency f_m. Using a similar approach, derive an approximation to the output of a phase modulator with input $m(t)$, under the condition that $\max\{|m(t)|\} < 0.3$ radians and $k_p = 1$. What is the approximate spectrum of this phase-modulated signal?

4.8 A carrier wave is frequency-modulated using a sinusoidal signal of frequency f_m and amplitude A_m.

(a) Determine the values of the modulation index β for which the carrier component of the FM signal is reduced to zero. See the Appendix for calculating $J_0(\beta)$.

(b) In a certain experiment conducted with $f_m = 1$ kHz and increasing A_m (starting from 0 volts), it is found that the carrier component of the FM signal is reduced to zero for the first time when $A_m = 2$ volts. What is the frequency sensitivity of the modulator? What is the value of A_m for which the carrier component is reduced to zero for the second time?

4.9 An FM signal with modulation index $\beta = 1$ is transmitted through an ideal band-pass filter with mid-band frequency f_c and bandwidth $5f_m$, where f_c is the carrier frequency and f_m is the frequency of the sinusoidal modulating wave. Determine the amplitude spectrum of the filter output.

4.10 A carrier wave of frequency 100 MHz is frequency-modulated by a sinusoidal wave of amplitude 20 volts and frequency 100 kHz. The frequency sensitivity of the modulator is 25 kHz per volt.

(a) Determine the approximate bandwidth of the FM signal, using Carson's rule.

(b) Determine the bandwidth by transmitting only those side frequencies whose amplitudes exceed 1 percent of the unmodulated carrier amplitude. Use the universal curve of Figure 4.9 for this calculation.

(c) Repeat your calculations, assuming that the amplitude of the modulating signal is doubled.

(d) Repeat your calculations, assuming that the modulation frequency is doubled.

4.11 Consider a wide-band PM signal produced by a sinusoidal modulating wave $A_m \cos(2\pi f_m t)$, using a modulator with a phase sensitivity equal to k_p radians per volt.

(a) Show that if the maximum phase deviation of the PM signal is large compared with one radian, the bandwidth of the PM signal varies linearly with the modulation frequency f_m.

(b) Compare this characteristic of a wide-band PM signal with that of a wideband FM signal.

4.12 Figure P4.12 shows the block diagram of a real-time *spectrum analyzer* working on the principle of frequency modulation. The given signal $g(t)$ and a frequency-modulated signal $s(t)$ are applied to a multiplier and the output $g(t)s(t)$ is fed into a filter of impulse response $h(t)$. The $s(t)$ and $h(t)$ are *linear FM signals* whose instantaneous frequencies vary at opposite rates, as shown by

$$s(t) = \cos(2\pi f_c t - \pi k t^2),$$
$$h(t) = \cos(2\pi f_c t + \pi k t^2)$$

where k is a constant. Show that the envelope of the filter output is proportional to the amplitude spectrum of the input signal $g(t)$ with kt playing the role of frequency f. *Hint:* Use the complex notations described in Chapter 2 for the analysis of band-pass signals and band-pass filters.

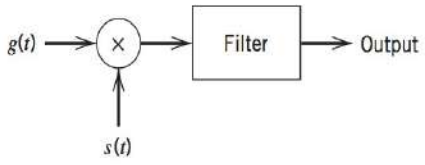


Figure P4.12

4.13 An FM signal with a frequency deviation of 10 kHz at a modulation frequency of 5 kHz is applied to two frequency multipliers connected in cascade. The first multiplier doubles the frequency and the second multiplier triples the frequency. Determine the frequency deviation and the modulation index of the FM signal obtained at the second multiplier output. What is the frequency separation of the adjacent side frequencies of this FM signal?

4.14 An FM signal is applied to a square-law device with output voltage v_2 related to input voltage v_1 by

$$v_2 = av_1^2$$

where a is a constant. Explain how such a device can be used to obtain an FM signal with a greater frequency deviation than that available at the input.

4.15 Figure P4.15 shows the frequency-determining network of a voltage-controlled oscillator. Frequency modulation is produced by applying the modulating signal $A_m \sin(2\pi f_m t)$ plus a bias V_b to a pair of varactor diodes connected across the parallel combination of a 200-μH inductor and 100-pF capacitor. The capacitor of each varactor diode is related to the voltage V (in volts) applied across its electrodes by

$$C = 100V^{-1/2}\text{pF}$$

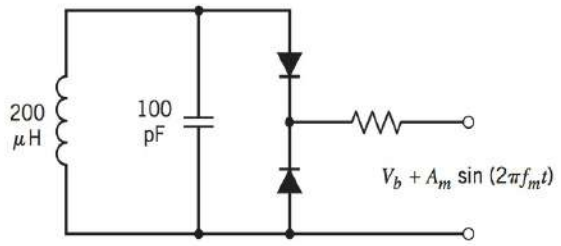


Figure P4.15

The unmodulated frequency of oscillation is 1 MHz. The VCO output is applied to a frequency multiplier to produce an FM signal with a carrier frequency of 64 MHz and a modulation index of 5. Determine (a) the magnitude of the bias voltage V_b and (b) the amplitude A_m of the modulating wave, given that $f_m = 10$ kHz.

4.16 The FM signal

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)\, dt\right]$$

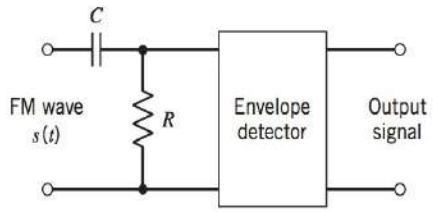


Figure P4.16

is applied to the system shown in Figure P4.16 consisting of a high-pass RC filter and an envelope detector. Assume that (a) the resistance R is small compared with the reactance of the capacitor C for all significant frequency components of $s(t)$ and (b) the envelope detector does not load the filter. Determine the resulting signal at the envelope detector output, assuming that $k_f|m(t)| < f_c$ for all t.

4.17 In the frequency discriminator of Figure 4.14, let the frequency separation between the resonant frequencies of the two parallel-tuned LC filters be denoted by $2kB$, where $2B$ is the 3-dB bandwidth of either filter and k is a scaling factor. Assume that both filters have a high Q-factor.

(a) Show that the total response of both filters has a slope equal to $2k/B(1 + k^2)^{3/2}$ at the center frequency f_c.

(b) Let D denote the deviation of the total response measured with respect to a straight line passing through $f = f_c$ with this slope. Plot D versus δ for $k = 1.5$ and $-kB \le \delta \le kB$, where $\delta = f - f_c$.

4.18 Consider the frequency demodulation scheme shown in Figure P4.18 in which the incoming FM signal $s(t)$ is passed through a delay line that produces a phase-shift of $\pi/2$ radians at the carrier frequency f_c. The delay-line output is subtracted from the incoming FM signal, and the resulting composite signal is then envelope-detected. This demodulator finds application in demodulating microwave FM signals. Assuming that

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_m t)]$$

analyze the operation of this demodulator when the modulation index β is less than unity and the delay T produced by the delay line is sufficiently small to justify making the approximations

$$\cos(2\pi f_m T) \simeq 1$$

and

$$\sin(2\pi f_m T) \simeq 2\pi f_m T$$

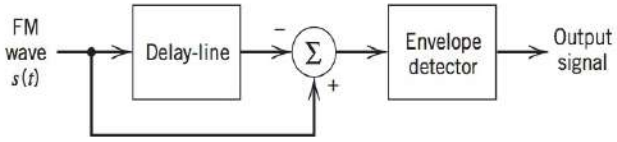


Figure P4.18

4.19 Figure P4.19 shows the block diagram of a *zero-crossing detector* for demodulating an FM signal. It consists of a limiter, a pulse generator for producing a short pulse at each zero-crossing of the input, and a low-pass filter for extracting the modulating wave.

(a) Show that the instantaneous frequency of the input FM signal is proportional to the number of zero crossings in the time interval $t - (T_1/2)$ to $t + (T_1/2)$, divided by T_1. Assume that the modulating signal is essentially constant during this time interval.

(b) Illustrate the operation of this demodulator, using the sawtooth wave of Figure P4.1 as the modulating wave.

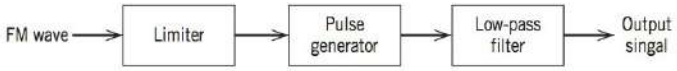


Figure P4.19

4.20 Suppose that the received signal in an FM system contains some residual amplitude modulation of positive amplitude $a(t)$, as shown by

$$s(t) = a(t)\cos[2\pi f_c t + \phi(t)]$$

where f_c is the carrier frequency. The phase $\phi(t)$ is related to the modulating signal $m(t)$ by

$$\phi(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau$$

where k_f is a constant. Assume that the signal $s(t)$ is restricted to a frequency band of width B_T, centered at f_c, where B_T is the transmission bandwidth of the FM signal in the absence of amplitude modulation, and that the amplitude modulation is slowly varying compared with $\phi(t)$. Show that the output of an ideal frequency discriminator produced by $s(t)$ is proportional to $a(t)m(t)$. *Hint:* Use the complex notation described in Chapter 2 to represent the modulated wave $s(t)$.

4.21

(a) Let the modulated wave $s(t)$ in Problem 4.20 be applied to an ideal amplitude *limiter*, whose output $z(t)$ is defined by

$$z(t) = \text{sgn}[s(t)] = \begin{cases} +1, & s(t) > 0 \\ -1, & s(t) < 0 \end{cases}$$

Show that the limiter output may be expressed in the form of a Fourier series as follows:

$$z(t) = \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\cos[2\pi f_c t(2n+1) + (2n+1)\phi(t)]$$

(b) Suppose that the limiter output is applied to a band-pass filter with a passband amplitude response of one and bandwidth B_T centered about the carrier frequency f_c, where B_T is the transmission bandwidth of the FM signal in the absence of amplitude modulation. Assuming that f_c is much greater than B_T, show that the resulting filter output equals

$$y(t) = \frac{4}{\pi}\cos[2\pi f_c t + \phi(t)]$$

By comparing this output with the original modulated signal $s(t)$ defined in Problem 4.20, comment on the practical usefulness of the result.

4.22 In this problem we study the idea of mixing in a superheterodyne receiver. To be specific, consider the block diagram of the *mixer* shown in Figure P4.22 that consists of a product modulator with a local oscillator of *variable frequency* f_l, followed by a band-pass filter. The input signal is an AM wave of bandwidth 10 kHz and carrier frequency that may lie anywhere in the range 0.535–1.605 MHz; these parameters are typical of AM radio broadcasting. It is required to translate this signal to a frequency band centered at a fixed *intermediate frequency* (IF) of 0.455 MHz. Find the range of tuning that must be provided in the local oscillator in order to achieve this requirement.

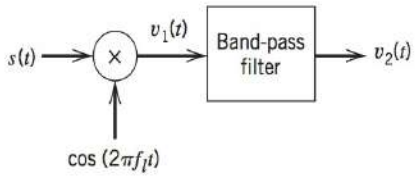


Figure P4.22

4.23 Figure P4.23 shows the block diagram of a *heterodyne spectrum analyzer*. It consists of a variable-frequency oscillator, multiplier, band-pass filter, and root mean-square (rms) meter. The oscillator has an amplitude A and operates over the range f_0 to $f_0 + W$, where f_0 is the mid-band frequency of the filter and W is the signal bandwidth. Assume that $f_0 = 2W$, the filter bandwidth Δf is small compared with f_0, and that the passband amplitude response of the filter is one. Determine the value of the rms meter output for a low-pass input signal $g(t)$.

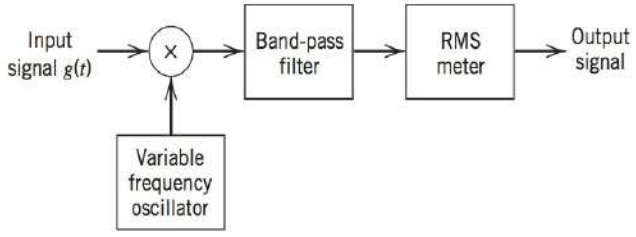


Figure P4.23

4.24 What is the analytical description of the amplitude spectrum corresponding to the Gaussian pulse of the Theme Example in Section 4.7? Justify your answer.

4.25 What are the Carson rule and 1 percent FM bandwidth for GSM? Justify your answer.

Computer Problems

4.26 In this problem, we simulate the spectrum produced by an FM modulator with input $A_m \sin(2\pi f_m t)$. It is suggested that the following Matlab script be used to simulate the FM modulator.

```
fc    = 100;                    % Carrier frequency (kHz)
Fs    = 1024;                   % Sampling rate (kHz)
fm    = 1;                      % Modulating frequency (kHz)
Ts    = 1/Fs;                   % Sample period (ms)
t     = [0:Ts:120];             % Observation period (ms)
m     = cos(2*pi*fm*t);         % modulating signal
beta  = 1.0;                    % modulation index
theta = 2*pi*fc*t+ 2*pi*beta*   % integrate signal
        cumsum(m)*Ts;
s     = cos(theta);
FFTsize = 4096;
S     = spectrum(s,FFTsize);
```

```
Freq = [0:Fs/FFTsize:Fs/2];
subplot(2,1,1), plot(t,s), xlabel('Time (ms)'), ylabel('Amplitude');
axis([0 0.5 −1.5 1.5]), grid on
subplot(2,1,2), stem(Freq,sqrt(S/682))
xlabel('Frequency (kHz)'), ylabel('Amplitude Spectrum');
axis([95 105 0 1), grid on
```

(a) For modulation indices of 1, 2, 5, and 10, determine the power in the harmonics of the modulating frequency about the carrier (ignore sidelobes). How many side-frequencies are required for 90% of the power in each case?

(b) For what minimum modulation index is the power at the carrier frequency reduced to zero?

4.27 Use the following Matlab script to numerically estimate the rms phase error for a signal having the spectrum of Figure 4.5.

```
%—One-sided phase noise spectrum
f    = [1 10 100 1000 10000];      %Hz
SdB = [−30 −40 −50 −65 −70];       % Spectrum (dBc)

%—interpolate spectrum on linear scale—
del_f = 1;                         % integration step size (Hz)
f1 = [10: del_f: 10000];           %Hz
S1 = interp1(f, 10.^(SdB/10), f1); % absolute power

%—numerically (Riemann) integrate from 10 Hz to 10 kHz—
Int_S = sum(S1)* del_f;
Theta_rms = sqrt(Int_S);           % in radians
Theta_rms = Theta_rms*180/pi       % in degrees
```

Repeat for the phase noise spectrum of Figure P4.27 where the range of frequencies of interest is from 100 Hz to 1 MHz.

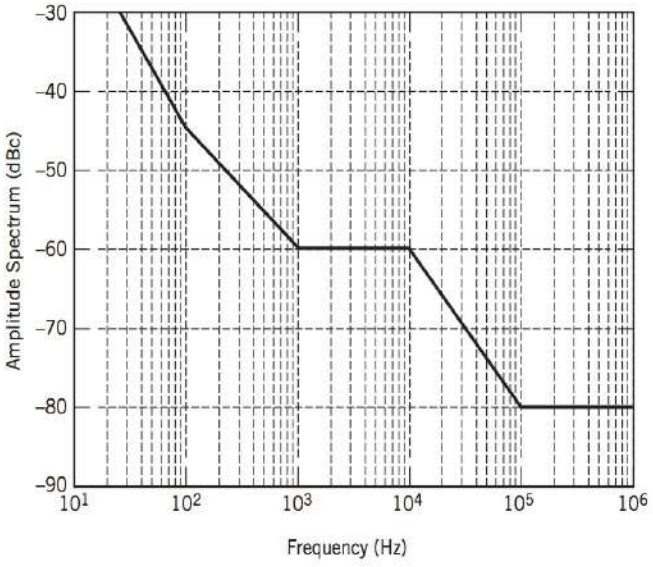


Figure P4.27

4.28 In this problem, we will simulate the behavior of an FM discriminator. Use the FM modulator of Problem 4.26 to generate a test signal and use the following Matlab script as a guide for performing discriminator detection. This will also require use of the envelope detector of Problem 3.25.

```
fc = 100;                          %kHz
Fs = 1024;                         % kHz
Ts = 1/Fs;
 t = [0:Ts:10*Ts];

%— FIR differentiator (Fs = 1024 kHz. BT/2 = 10 kKhz) —
FIRdiff = [1.60385 0.0 0.0 0.0 −0.0 0.0 0.0 −0.0 −0.0 −0.0
  −1.60385];
BP_diff = FIRdiff * exp(j*2*pi*fc*t);

%—FIR Lowpass Butterworth filter - Fs = 1024 kHz, f3dB =
  5 kHz—
LPF_B = 1E-4*[0.0706  0.2117  0.2117  0.0706];
LPF_A = [1.0000  −2.9223  2.8476  −0.9252];

D1 = filter(BP_diff, 1, S);        % Bandpass discriminator
D2 = EnvDetect(D1);                % Envelope detection
D3 = filter(LPF_B,LPF_A, D2);      % Low-pass filtering
```

(a) What differences are observed between the detected signal and the original? Why?

(b) Modify the modulating signal to the following

$$s(t) = \sin(2\pi f_m t) + 0.5\cos(2\pi f_m t/3)$$

Are any differences observed between the detected signal and the original?

(c) As the modulating frequency f_m increases, when does distortion in the output appear? Why?

4.29 The following Matlab script is a digital model for simulating the behavior of a phase-locked loop (PLL). For this model, do the following:

(a) Compare the simulation models for the phase detector, VCO, and loop filter to the analytical models presented in the text. Justify the digital approximation.

(b) How does the behavior change if we modify the gain k_v of the VCO? How is the lock-in time affected?

(c) If the loop filter is replaced with simple proportional control, how is the behavior affected?

(d) What is the maximum modulating frequency for which the PLL tracks the signal?

(e) If the phase detector is replaced by the multiply operator $s(t)r(t)$, how does the PLL behave? Why?

```
fc = 100;                  % carrier frequency (kHz)
fm = 1;                    % frequency of modulation signal (kHz)
Fs = 32*fc;                % Sampling rate (kHz)
Ts = 1/Fs;                 % sampling period (ms)
 t = [0:Ts:5];             % observation period (ms)
Ac = 1; Av = 1;            % Output amplitudes of FM modulator
                             and VCO
kf = 10; kv = 20;          % frequency sensitivities of FM modulator
                             and VCO
FilterState = 0;           % initial state of loop filter
VCOstate = 0;              % initial phase state of VCO

%———FM modulator———————
m  = 0.2* sin(2*pi*fm*t) + 0.3*cos(2*pi*fm/3*t);
```

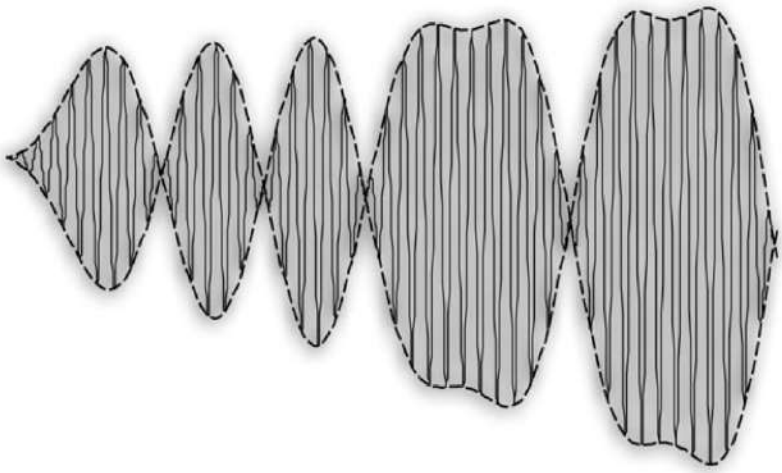
```
phi = cumsum(m) * Ts;
s   = Ac * sin(2*pi*fc*t + kf*phi);

%—— Simulate nonlinear PLL ——
v = zeros(size(t)); r(l) = 0; e(l) = 0;
for i = 2: length(t)
%—— VCO ———
VCOstate = VCOstate + 2*pi*kv*v(i – l)*Ts;
r(i)        = Av*cos(2*pi*fc*t(i) + VCOstate);

%——— Phase Detector ———
e(i) = sin(phi(i) – VCOstate);
%— Loop Filter ——
FilterState = FilterState + e(i);      % integrator
v(i)        = FilterState + e(i);      % integral plus proportional
                                          control

end
subplot(4,1,1), plot(t,m)             % modulating signal
subplot(4,1,2), plot(t,phi)           % phase of transmitted signal
subplot(4,1,3), plot(t,e)             % phase detector output
subplot(4,1,4), plot(t,v)             % recovered signal
```

5



PROBABILITY THEORY AND RANDOM PROCESSES

5.1 INTRODUCTION

In Chapter 2, we dealt with the Fourier transform as a mathematical tool for the representation of *deterministic signals* and the transmission of such signals through linear time-invariant filters; by deterministic signals we mean the class of signals that may be modeled as completely specified functions of time. In this chapter, we resume the development of background material necessary for a more detailed understanding of communication systems. Specifically, we deal with the statistical characterization of *random signals,* which may be viewed as the second pillar of communication theory.

Examples of *random signals* are encountered in every practical communication system. We say a signal is "random" if it is not possible to predict its precise value in advance. Consider, for example, a radio communication system. The received signal in such a system usually consists of an *information-bearing signal* component, a random *interference* component, and *receiver noise.* The information-bearing signal component may represent, for example, a voice signal that, typically, consists of randomly spaced bursts of energy of random duration. The interference component may represent spurious electromagnetic waves produced by other communication systems operating in the vicinity of the radio receiver. A major source of receiver noise is *thermal noise,* which is caused by the random motion of electrons in conductors and devices at the front end of the receiver. We thus find that the received signal is completely random in nature.

Although it is not possible to predict the precise value of a random signal in advance, it may be described in terms of its *statistical* properties such as the average power in the random signal, or the average spectral distribution of this power. The mathematical discipline that deals with the statistical characterization of random signals is *probability theory.*

We begin this chapter on random processes by reviewing some basic definitions in probability theory, followed by a review of the notions of a random variable and random

process. A random process consists of an ensemble (family) of sample functions, each of which varies randomly with time. A random variable is obtained by observing a random process at a fixed instant of time.

5.2 PROBABILITY

Probability theory[1] is rooted in phenomena that, explicitly or implicitly, can be modeled by an experiment with an outcome that is subject to *chance*. Moreover, if the experiment is repeated, the outcome can differ because of the influence of an underlying random phenomenon or chance mechanism. Such an experiment is referred to as a *random experiment*. For example, the experiment may be the observation of the result of tossing a fair coin. In this experiment, the possible outcomes of a trial are "heads" or "tails."

There are two approaches to the definition of probability. The first approach is based on the *relative frequency of occurrence*: in n trials of a random experiment, if we expect an event A to occur m times, then we assign the probability $\frac{m}{n}$ to the event A. This definition of probability is straightforward to apply in games of chance and many engineering situations.

However, there are situations where experiments are not repeatable and yet the concept of probability has intuitive applicability. In this second case, we use a more general definition of probability based on *set theory* and a set of related mathematical *axioms*. In situations where the experiment is repeatable, the set theory approach agrees completely with the relative frequency of occurrence approach.

In general, when we perform a random experiment, it is natural for us to be aware of the various outcomes that are likely to arise. In this context, it is convenient to think of an experiment and its possible outcomes as defining a space and its points. If an experiment has K possible outcomes, then for the kth possible outcome we have a point called the *sample point*, which we denote by s_k. With this basic framework, we make the following definitions:

- The set of all possible outcomes of the experiment is called the *sample space*, which we denote by S.
- An *event* corresponds to either a single sample point or a set of sample points in the space S.
- A single sample point is called an *elementary event*.
- The entire sample space S is called the *sure event*; and the null set ϕ is called the *null or impossible event*.
- Two events are *mutually exclusive* if the occurrence of one event precludes the occurrence of the other event.

The sample space S may be *discrete* with a countable number of outcomes, such as the outcomes when tossing a die. Alternatively, the sample space may be *continuous*, such as the voltage measured at the output of a noise source.

A *probability measure* $\mathbf{P}$ is a function that assigns a non-negative number to an event A in the sample space S and satisfies the following three properties (axioms):

1. $0 \leq \mathbf{P}[A] \leq 1$ (5.1)

2. $\mathbf{P}[S] = 1$ (5.2)

3. If A and B are two mutually exclusive events, then

$$\mathbf{P}[A \cup B] = \mathbf{P}[A] + \mathbf{P}[B] \tag{5.3}$$

This abstract definition of a probability system is illustrated in Figure 5.1. The sample space S is mapped to events via the random experiment. The events may be elementary outcomes of the sample space or larger subsets of the sample space. The probability function assigns a value between 0 and 1 to each of these events. The probability value is not unique to the event; mutually exclusive events may be assigned the same probability. However, the probability of the union of all events—that is, the sure event—is always unity.

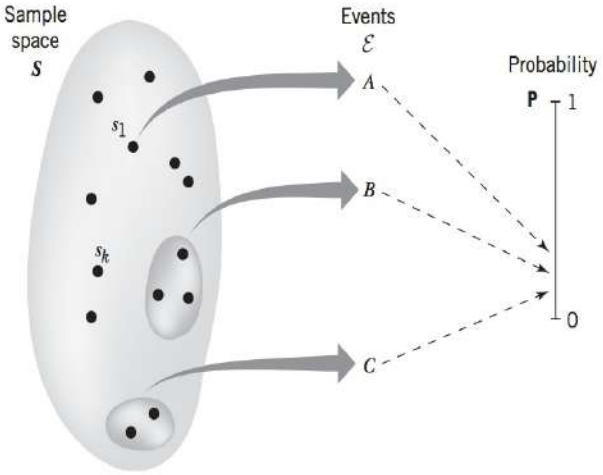


FIGURE 5.1 Illustration of the relationship between sample space, events, and probability.

The three axioms and their relationship to the relative frequency approach may be illustrated by the Venn diagram of Figure 5.2. If we equate **P** to a measure of the area in the Venn diagram with the total area of S equal to one, then the axioms are simple statements of familiar geometric results regarding area.

The following properties of probability measure **P** may be derived from the above axioms:

1. $\mathbf{P}[\overline{A}] = 1 - \mathbf{P}[A]$ (5.4)
where $\overline{A}$ is the *complement* of the event A.

2. When events A and B are not mutually exclusive, then the probability of the union event "A or B" satisfies

$$\mathbf{P}[A \cup B] = \mathbf{P}[A] + \mathbf{P}[B] - \mathbf{P}[A \cap B] \tag{5.5}$$

where $\mathbf{P}[A \cap B]$ is the probability of the *joint event* "A and B."

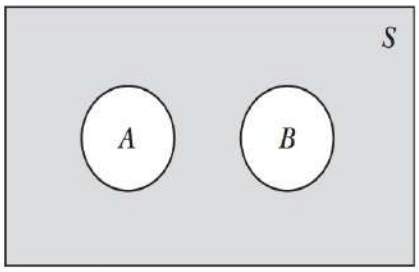


FIGURE 5.2 Venn diagram presenting a geometric interpretation of the three axioms of probability.

3. If $A_1, A_2, \ldots, A_m$ are mutually exclusive events that include all possible outcomes of the random experiment, then

$$\mathbf{P}[A_1] + \mathbf{P}[A_2] + \ldots + \mathbf{P}[A_m] = 1 \tag{5.6}$$

CONDITIONAL PROBABILITY

Suppose we perform an experiment that involves a pair of events A and B. Let $\mathbf{P}[B|A]$ denote the probability of event B, given that event A has occurred. The probability $\mathbf{P}[B|A]$ is called the *conditional probability* of B given A. Assuming that A has nonzero probability, the conditional probability $\mathbf{P}[B|A]$ is defined by

$$\mathbf{P}[B|A] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[A]} \tag{5.7}$$

where $\mathbf{P}[A \cap B]$ is the joint probability of A and B. We leave it to the reader to justify this definition based on relative frequency of occurrence. We may write Eq. (5.7) as

$$\mathbf{P}[A \cap B] = \mathbf{P}[B|A]\mathbf{P}[A] \tag{5.8}$$

It is apparent that we may also write

$$\mathbf{P}[A \cap B] = \mathbf{P}[A|B]\mathbf{P}[B] \tag{5.9}$$

Accordingly, we may state that *the joint probability of two events may be expressed as the product of the conditional probability of one event given the other, and the*

elementary probability of the other. Note that the conditional probabilities $\mathbf{P}[B|A]$ and $\mathbf{P}[A|B]$ have essentially the same properties as the various probabilities previously defined.

Situations may exist where the conditional probability $\mathbf{P}[A|B]$ and the probabilities $\mathbf{P}[A]$ and $\mathbf{P}[B]$ are easily determined directly, but the conditional probability $\mathbf{P}[B|A]$ is desired. From Eqs. (5.8) and (5.9), it follows that, provided $\mathbf{P}[A] \neq 0$, we may determine $\mathbf{P}[B|A]$ by using the relation

$$\mathbf{P}[B|A] = \frac{\mathbf{P}[A|B]\mathbf{P}[B]}{\mathbf{P}[A]} \tag{5.10}$$

This relation is a special form of *Bayes' rule.*

Suppose that the conditional probability $\mathbf{P}[B|A]$ is simply equal to the elementary probability of occurrence of event B, that is,

$$\mathbf{P}[B|A] = \mathbf{P}[B] \tag{5.11}$$

Under this condition, the probability of occurrence of the joint event $A \cap B$ is equal to the product of the elementary probabilities of the events A and B:

$$\mathbf{P}[A \cap B] = \mathbf{P}[A]\mathbf{P}[B] \tag{5.12}$$

so that

$$\mathbf{P}[A|B] = \mathbf{P}[A] \tag{5.13}$$

That is, the conditional probability of event A, assuming the occurrence of event B, is simply equal to the elementary probability of event A. We thus see that in this case a knowledge of the occurrence of one event tells us no more about the probability of occurrence of the other event than we knew without that knowledge. Events A and B that satisfy this condition are said to be *statistically independent.*

EXAMPLE 5.1 Binary Symmetric Channel

Consider a *discrete memoryless channel* used to transmit binary data. The channel is said to be *discrete* in that it is designed to handle discrete messages. It is *memoryless* in the sense that the channel output at any time depends only on the channel input at that time. Owing to the unavoidable presence of *noise* in the channel, *errors* are made in the received binary data stream. Specifically, when symbol 1 is sent, *occasionally* an error is made and symbol 0 is received and vice versa. The channel is assumed to be symmetric, which means that the probability of receiving symbol 1 when symbol 0 is sent is the same as the probability of receiving symbol 0 when symbol 1 is sent.

To describe the probabilistic nature of this channel fully, we need two sets of probabilities.

1. The *a priori probabilities* of sending binary symbols 0 and 1: they are

$$\mathbf{P}[A_0] = p_0$$

and

$$\mathbf{P}[A_1] = p_1$$

where A_0 and A_1 denote the events of transmitting symbols 0 and 1, respectively. Note that $p_0 + p_1 = 1$.

2. The *conditional probabilities of error*: they are

$$\mathbf{P}[B_1|A_0] = \mathbf{P}[B_0|A_1] = p$$

where B_0 and B_1 denote the events of receiving symbols 0 and 1, respectively. The conditional probability $\mathbf{P}[B_1|A_0]$ is the probability of receiving symbol 1, given that symbol 0 is sent. The second conditional probability $\mathbf{P}[B_0|A_1]$ is the probability of receiving symbol 0, given that symbol 1 is sent.

The requirement is to determine the *a posteriori probabilities* $\mathbf{P}[A_0|B_0]$ and $\mathbf{P}[A_1|B_1]$. The conditional probability $P[A_0|B_0]$ is the probability that symbol 0 was sent, given that symbol 0 is received. The second conditional probability $P[A_1|B_1]$ is the probability that symbol 1 was sent, given that symbol 1 is received. Both these conditional probabilities refer to events that are observed "after the fact"; hence, the name "a posteriori" probabilities.

Since the events B_0 and B_1 are mutually exclusive, we have from axiom (3)

$$\mathbf{P}[B_0|A_0] + \mathbf{P}[B_1|A_0] = 1$$

That is to say,

$$\mathbf{P}[B_0|A_0] = 1 - p$$

Similarly, we may write

$$\mathbf{P}[B_1|A_1] = 1 - p$$

Accordingly, we may use the *transition probability diagram* shown in Figure 5.3 to represent the binary communication channel specified in this example; the term "transition probability" refers to the conditional probability of error. Figure 5.3 clearly depicts the (assumed) symmetric nature of the channel; hence, the name "binary symmetric channel."

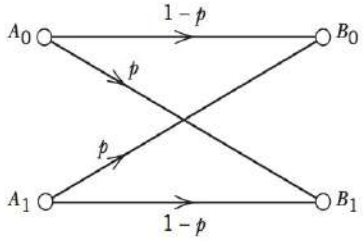


FIGURE 5.3 Transition probability diagram of binary symmetric channel.

From Figure 5.3, we deduce the following results:

1. The probability of receiving symbol 0 is given by

$$\begin{aligned}\mathbf{P}[B_0] &= \mathbf{P}[B_0|A_0]\mathbf{P}[A_0] + \mathbf{P}[B_0|A_1]\mathbf{P}[A_1] \\ &= (1-p)p_0 + pp_1\end{aligned}$$

2. The probability of receiving symbol 1 is given by

$$\begin{aligned}\mathbf{P}[B_1] &= \mathbf{P}[B_1|A_0]\mathbf{P}[A_0] + \mathbf{P}[B_1|A_1]\mathbf{P}[A_1] \\ &= pp_0 + (1-p)p_1\end{aligned} \tag{5.14}$$

Therefore, applying Bayes' rule, we obtain

$$\begin{aligned}\mathbf{P}[A_0|B_0] &= \frac{\mathbf{P}[B_0|A_0]\mathbf{P}[A_0]}{\mathbf{P}(B_0)} \\ &= \frac{(1-p)p_0}{(1-p)p_0 + pp_1}\end{aligned}$$

$$\begin{aligned}\mathbf{P}[A_1|B_1] &= \frac{\mathbf{P}[B_1|A_1]\mathbf{P}[A_1]}{\mathbf{P}[B_1]} \\ &= \frac{(1-p)p_1}{pp_0 + (1-p)p_1}\end{aligned}$$

These two *a posteriori* probabilities are the desired results.

5.3 RANDOM VARIABLES

While the meaning of the outcome of a random experiment is clear, such outcomes are often not the most convenient representation for mathematical analysis. For example, heads or tails is not a convenient mathematical representation. As another example, consider a random experiment where we draw colored balls from an urn; the color is not directly amenable to mathematical analysis.

In these cases, it is often more convenient if we assign a number or a range of values to the outcomes of a random experiment. For example, a head could correspond to 1 and a tail to 0. We use the expression *random variable* to describe this process of assigning a number to the outcome of a random experiment.

In general, a function whose domain is a sample space and whose range is a set of real numbers is called a random variable of the experiment. That is, for events in $\mathscr{E}$, a random variable assigns a subset of the real line. Thus, if the outcome of the experiment is s, we denote the random variable as $X(s)$ or just X. Note that X is a function, even if it is, for historical reasons, called a random variable. We denote a particular outcome of a random experiment by x; that is, $X(s_k) = x$. *There may be more than one random variable associated with the same random experiment.*

The concept of a random variable is illustrated in Figure 5.4, where we have suppressed the events but show subsets of the sample space being mapped directly to a subset or the real line. The probability function applies to this random variable in exactly the same manner that it applies to the underlying events.

The benefit of using random variables is that probability analysis can now be developed in terms of real-valued quantities regardless of the form or shape of the underlying events of the random experiment. Random variables may be *discrete* and take only a finite number of values, such as in the coin-tossing experiment. Alternatively, random variables may be *continuous* and take a range of real values. For example, the random variable that represents the amplitude of a noise voltage at a particular instant in time is a continuous random variable because, in theory, it may take on any value between plus and minus infinity. Random variables may also be complex valued, but a complex-valued random variable may always be treated as a vector of two real-valued random variables.

To proceed further, we need a probabilistic description of random variables that works equally well for discrete as well as continuous random variables. To this end, let us consider the random variable X and the probability of the event $X \leq x$. We denote this probability by $\mathbf{P}[X \leq x]$. It is apparent that this probability is a function of the *dummy variable* x. To simplify our notation, we write

$$F_X(x) = \mathbf{P}[X \leq x] \tag{5.15}$$

The function $F_X(x)$ is called the *cumulative distribution function* (cdf) or simply the *distribution function* of the random variable X. Note that $F_X(x)$ is a function of x, not of the random variable X. However, it depends on the assignment of the random variable X, which accounts for the use of X as subscript. For any point x, the distribution function $F_X(x)$ expresses a probability.

The distribution function $F_X(x)$ has the following properties, which follow directly from Eq. (5.15):

1. The distribution function $F_X(x)$ is bounded between zero and one.

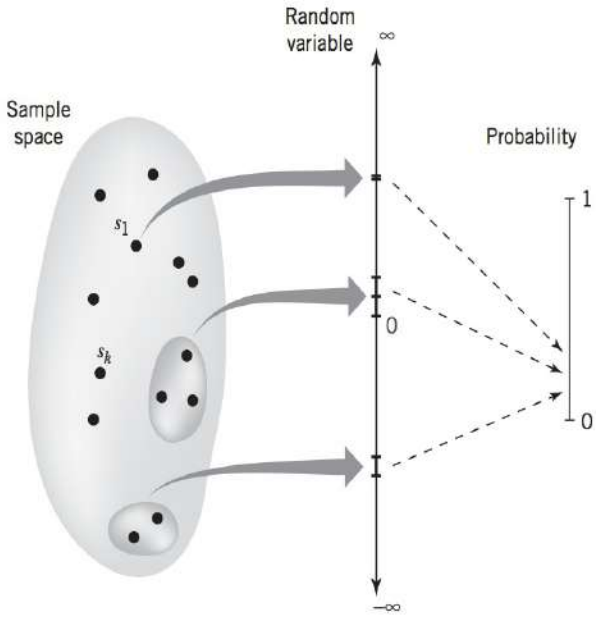


FIGURE 5.4 Illustration of the relationship between sample space, random variable, and probability.

2. The distribution function $F_X(x)$ is a monotone-nondecreasing function of x; that is,

$$F_X(x_1) \leq F_X(x_2) \qquad \text{if } x_1 < x_2 \tag{5.16}$$

The distribution function of a random variable always exists. If the distribution function is continuously differentiable then an alternative description of the probability of the random variable X is often useful. The derivative of the distribution function, as shown by

$$f_X(x) = \frac{d}{dx} F_X(x) \tag{5.17}$$

is called the *probability density function* (pdf) of the random variable X. Note that the differentiation in Eq. (5.17) is with respect to the dummy variable x. The name density function arises from the fact that the probability of the event $x_1 < X \leq x_2$ equals

$$\begin{aligned} \mathbf{P}[x_1 < X \leq x_2] &= \mathbf{P}[X \leq x_2] - \mathbf{P}[X \leq x_1] \\ &= F_X(x_2) - F_X(x_1) \\ &= \int_{x_1}^{x_2} f_X(x)\, dx \end{aligned} \tag{5.18}$$

The probability of an interval is therefore the area under the probability density function in that interval. Putting $x_1 = -\infty$ in Eq. (5.18), and changing the notation somewhat, we

EXAMPLE 5.2 Uniform Distribution

A random variable X is said to be *uniformly distributed* over the interval (a,b) if its probability density function is

$$f_X(x) = \begin{cases} 0, & x \leq a \\ \dfrac{1}{b-a}, & a < x \leq b \\ 0, & x > b \end{cases} \tag{5.21}$$

The cumulative distribution function of X is therefore

$$F_X(x) = \begin{cases} 0, & x \leq a \\ \dfrac{x-a}{b-a}, & a < x \leq b \\ 0, & x > b \end{cases} \tag{5.22}$$

Figure 5.5 shows plots of the probability density function and the cumulative distribution function of the uniformly distributed random variable X.

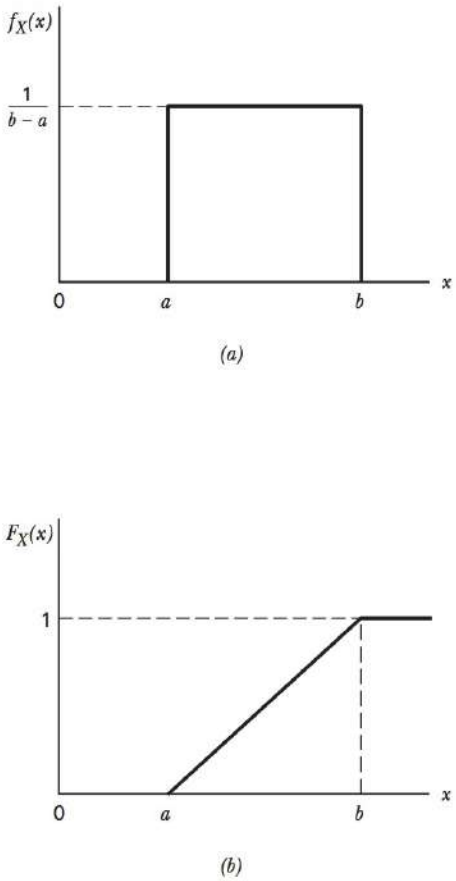


FIGURE 5.5 Uniform distribution.
(*a*) Probability density function.
(*b*) Cumulative distribution function.

readily see that the distribution function is defined in terms of the probability density function as follows:

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\, d\xi \tag{5.19}$$

Since $F_X(\infty) = 1$, corresponding to the probability of a certain event, and $F_X(-\infty) = 0$, corresponding to the probability of an impossible event, we also find from Eq. (5.18) that

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1 \tag{5.20}$$

Earlier we mentioned that a distribution function must always be monotone nondecreasing. This means that its derivative or the probability density function must always be nonnegative. Accordingly, we may state that a *probability density function must always be a nonnegative function, and with a total area of one*.

SEVERAL RANDOM VARIABLES

Thus far we have focused attention on situations involving a single random variable. However, we find frequently that the outcome of an experiment requires several random variables for its description. We now consider situations involving two random variables. The probabilistic description developed in this way may be readily extended to any number of random variables.

Consider two random variables X and Y. We define *the joint distribution function* $F_{X,Y}(x,y)$ *as the probability that the random variable X is less than or equal to a specified value x and that the random variable Y is less than or equal to a specified value y*. The variables X and Y may be two separate one-dimensional random variables or the components of a single two-dimensional random variable. In either case, the joint sample space is the xy-plane. The joint distribution function $F_{X,Y}(x,y)$ is the probability that the outcome of an experiment will result in a sample point lying inside the quadrant ($-\infty < X \leq x$, $-\infty < Y \leq y$) of the joint sample space. That is,

$$F_{X,Y}(x,y) = \mathbf{P}[X \leq x, Y \leq y] \tag{5.23}$$

Suppose that the joint distribution function $F_{X,Y}(x,y)$ is continuous everywhere, and that the partial derivative

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} \tag{5.24}$$

exists and is continuous everywhere. We call the function $f_{X,Y}(x,y)$ the *joint probability density function* of the random variables X and Y. The joint distribution function $F_{X,Y}(x,y)$ is a monotone-nondecreasing function of both x and y. Therefore, from Eq. (5.24) it follows that the joint probability density function $f_{X,Y}(x,y)$ is always nonnegative. Also the total volume under the graph of a joint probability density function must be unity, as shown by

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(\xi,\eta)\, d\xi\, d\eta = 1 \tag{5.25}$$

The probability density function for a single random variable (X, say) can be obtained from its joint probability density function with a second random variable (Y, say) in the following way. We first note that

$$F_X(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{X,Y}(\xi, \eta)\, d\xi\, d\eta \tag{5.26}$$

Therefore, differentiating both sides of Eq. (5.26) with respect to x, we get the desired relation:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \eta)\, d\eta \tag{5.27}$$

Thus the probability density function $f_X(x)$ is obtained from the joint probability density function $f_{X,Y}(x,y)$ by simply integrating it over all possible values of the undesired random variable, Y. The use of similar arguments in the other dimension yields $f_Y(y)$. The probability density functions $f_X(x)$ and $f_Y(y)$ are called *marginal densities*. Hence, the joint probability density function $f_{X,Y}(x,y)$ contains all the possible information about the joint random variables X and Y.

Suppose that X and Y are two continuous random variables with joint probability density function $f_{X,Y}(x,y)$. The *conditional probability density function* of Y given that $X = x$ is defined by

$$f_Y(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \tag{5.28}$$

provided that $f_X(x) > 0$, where $f_X(x)$ is the marginal density of X. The function $f_Y(y|x)$ may be thought of as a function of the variable y, with the variable x arbitrary, but *fixed*. Accordingly, it satisfies all the requirements of an ordinary probability density function, as shown by

$$f_Y(y|x) \geq 0 \tag{5.29}$$

and

$$\int_{-\infty}^{\infty} f_Y(y|x)\, dy = 1 \tag{5.30}$$

If the random variables X and Y are *statistically independent*, then knowledge of the outcome of X can in no way affect the distribution of Y. The result is that the conditional probability density function $f_Y(y|x)$ reduces to the *marginal density* $f_Y(y)$, as shown by

$$f_Y(y|x) = f_Y(y) \tag{5.31}$$

In such a case, we may express the joint probability density function of the random variables X and Y as the product of their respective marginal densities, as shown by

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \tag{5.32}$$

Equivalently, we may state that if the joint probability density function of the random variables X and Y equals the product of their marginal densities, then X and Y are statistically independent. This last equation is one way of expressing the general statement that

$$\mathbf{P}[X \in A, Y \in B] = \mathbf{P}[X \in A]\mathbf{P}[Y \in B] \tag{5.33}$$

or $\mathbf{P}[X,Y] = \mathbf{P}[X]\mathbf{P}[Y]$ for statistically independent random variables X and Y.

EXAMPLE 5.3 Binomial Random Variable

Consider a sequence of coin-tossing experiments where the probability of a head is p and let X_n be the Bernoulli random variable representing the outcome of the nth toss. Since the outcome of one coin toss is not expected to influence the outcome of subsequent coin tosses, this may be referred to as a set of *independent Bernoulli trials.*

Let Y be the number of heads that occur on N tosses of the coins. This new random variable may be expressed as

$$Y = \sum_{n=1}^{N} X_n \tag{5.34}$$

What is the probability mass function of Y?

First consider the probability of obtaining y heads in a row followed by $N - y$ tails. Using the independence of the trials, repeated application of Eq. (5.33) implies that this probability is given by

$$\begin{aligned}\mathbf{P}[y \text{ heads followed by } N-y \text{ tails}] &= ppp\ldots pp(1-p)(1-p)\ldots(1-p)\\ &= p^y(1-p)^{N-y}\end{aligned}$$

By symmetry, this is the probability of any sequence of N trials that has y heads. To determine the probability of obtaining y heads anywhere in the N trials, the relative frequency definition of probability implies we simply have to count the number of possible arrangements of N tosses that have y heads and $N - y$ tails. That is, the probability that $Y = y$ is given by

$$\mathbf{P}[Y = y] = \binom{N}{y} p^y(1-p)^{N-y} \tag{5.35}$$

where

$$\binom{N}{y} = \frac{N!}{y!(N-y)!}$$

is the combinatorial function. Equation (5.35) defines the probability mass function of Y and the random variable Y is said to have a *binomial distribution.* The binomial distribution derives its name from the fact that the values of $\mathbf{P}[Y = y]$ are the successive terms in the expansion of the binomial expression

$$[p + (1-p)]^n$$

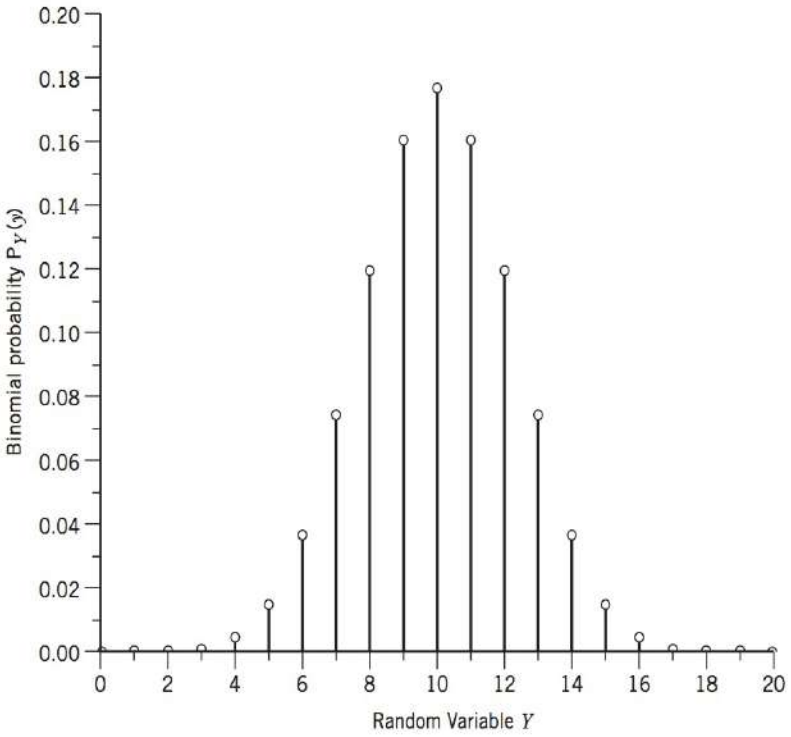


FIGURE 5.6 The binomial probability mass function for $N = 20$ and $p = \frac{1}{2}$.

where the $(y + 1)$st term of the expansion corresponds to $\mathbf{P}[Y = y]$. The binomial probability mass function for $N = 20$ and $p = \frac{1}{2}$ is illustrated in Figure 5.6.

5.4 STATISTICAL AVERAGES

Having discussed probability and some of its ramifications, we now seek ways for determining the *average* behavior of the outcomes arising in random experiments.

The *expected value* or *mean* of a random variable X is defined by

$$\mu_X = \mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx \tag{5.36}$$

where $\mathbf{E}$ denotes the *statistical expectation operator*. That is, the mean μ_X locates the center of gravity of the area under the probability density curve of the random variable X.

FUNCTION OF A RANDOM VARIABLE

Let X denote a random variable, and let $g(X)$ denote a real-valued function defined on the real line. The quantity obtained by letting the argument of the function $g(X)$ be a random variable is also a random variable which we denote as

$$Y = g(X) \tag{5.37}$$

To find the expected value of the random variable Y, we could of course find the probability density function $f_Y(y)$ and then apply the standard formula

$$\mathbf{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy$$

A simpler procedure, however, is to write

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_x(x)\, dx \tag{5.38}$$

Indeed, Eq. (5.38) may be viewed as generalizing the concept of expected value to an arbitrary function $g(X)$ of a random variable X.

EXAMPLE 5.4 Cosinusoidal Random Variable

Let

$$Y = g(X) = \cos(X)$$

where X is a random variable uniformly distributed in the interval $(-\pi,\pi)$; that is,

$$f_X(x) = \begin{cases} \dfrac{1}{2\pi}, & -\pi < x < \pi \\ 0, & \text{otherwise} \end{cases}$$

According to Eq. (5.38), the expected value of Y is

$$\begin{aligned} \mathbf{E}[Y] &= \int_{-\pi}^{\pi} (\cos x)\left(\frac{1}{2\pi}\right) dx \\ &= -\frac{1}{2\pi} \sin x\Big|_{x=-\pi}^{\pi} \\ &= 0 \end{aligned}$$

This result is intuitively satisfying.

MOMENTS

For the special case of $g(X) = X^n$, using Eq. (5.38) we obtain the nth *moment* of the probability distribution of the random variable X; that is,

$$\mathbf{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\, dx \tag{5.39}$$

By far the most important moments of X are the first two moments. Thus putting $n = 1$ in Eq. (5.39) gives the mean of the random variable as discussed above, whereas putting $n = 2$ gives the *mean-square value* of X:

$$\mathbf{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx \tag{5.40}$$

We may also define *central moments*, which are simply the moments of the difference between a random variable X and its mean μ_X. Thus, the nth central moment is

$$\mathbf{E}[(X - \mu_X)^n] = \int_{-\infty}^{\infty} (x - \mu_X)^n f_X(x)\, dx \tag{5.41}$$

For $n = 1$, the central moment is, of course, zero, whereas for $n = 2$ the second central moment is referred to as the *variance* of the random variable X, written as

$$\text{var}[X] = \mathbf{E}[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx \tag{5.42}$$

The variance of a random variable X is commonly denoted as σ_X^2. The square root of the variance, namely, σ_X, is called the *standard deviation* of the random variable X.

The variance σ_X^2 of a random variable X in some sense is a measure of the variable's "randomness." By specifying the variance σ_X^2, we essentially constrain the effective width of the probability density function $f_X(x)$ of the random variable X about the mean

μ_X. A precise statement of this constraint is due to Chebyshev. The *Chebyshev inequality* states that for any positive number ε, we have

$$P(|X - \mu_X| \geq \varepsilon) \leq \frac{\sigma_X^2}{\varepsilon^2} \tag{5.43}$$

From this inequality we see that the mean and variance of a random variable give a *partial description* of its probability distribution.

We note from Eqs. (5.40) and (5.42) that the variance σ_X^2 and mean-square value $\mathbf{E}[X^2]$ are related by

$$\begin{aligned}\sigma_X^2 &= \mathbf{E}[X^2 - 2\mu_X X + \mu_X^2] \\ &= \mathbf{E}[X^2] - 2\mu_X \mathbf{E}[X] + \mu_X^2 \\ &= \mathbf{E}[X^2] - \mu_X^2\end{aligned} \tag{5.44}$$

where, in the second line, we have used the *linearity* property of the statistical expectation operator $\mathbf{E}$. Equation (5.44) shows that if the mean μ_x is zero, then the variance σ_X^2 and the mean-square value $\mathbf{E}[X^2]$ of the random variable X are equal.

CHARACTERISTIC FUNCTION

Another important statistical average is the *characteristic function* $\phi_X(v)$ of the probability distribution of the random variable X, which is defined as the expectation of the complex exponential function $\exp(jvX)$, as shown by

$$\begin{aligned}\phi_X(v) &= \mathbf{E}[\exp(jvX)] \\ &= \int_{-\infty}^{\infty} f_X(x) \exp(jvx)\, dx\end{aligned} \tag{5.45}$$

where v is real and $j = \sqrt{-1}$. In other words, the characteristic function $\phi_X(v)$ is (except for a sign change in the exponent) the Fourier transform of the probability density function $f_X(x)$. In this relation we have used $\exp(jvx)$ rather than $\exp(-jvx)$, so as to conform with the convention adopted in probability theory. Recognizing that v and x play analogous roles to the variables $2\pi f$ and t of Fourier transforms, respectively, we deduce the following inverse relation, analogous with the inverse Fourier transform:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_X(v) \exp(-jvx)\, dv \tag{5.46}$$

This relation may be used to evaluate the probability density function $f_X(x)$ of the random variable X from its characteristic function $\phi_X(v)$.

EXAMPLE 5.5 Gaussian Random Variable

The *Gaussian random variable* is commonly encountered in the statistical analysis of a large variety of physical systems, including communication systems. Let X denote a Gaussian-distributed random variable of mean μ_X and variance σ_X^2. The probability density function of such a random variable is defined by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} \exp\left(-\frac{(x - \mu_X)^2}{2\sigma_X^2}\right), \qquad -\infty < x < \infty \tag{5.47}$$

Given this probability density function, we may readily show that the mean of the random variable X so defined is indeed μ_X and its variance is σ_X^2; these

evaluations are left as an exercise for the reader. In this example, we wish to use the characteristic function to evaluate the higher-order moments of the Gaussian random variable X.

Differentiating both sides of Eq. (5.45) with respect to v a total of n times, and then setting $v = 0$, we get the result

$$\frac{d^n}{dv^n}\,\phi_X(v)\Big|_{v=0} = (j)^n \int_{-\infty}^{\infty} x^n f_X(x)\,dx$$

The integral on the right-hand side of this relation is recognized as the nth moment of the random variable X. Accordingly, we may write

$$\mathbf{E}[X^n] = (-j)^n \frac{d^n}{dv^n}\,\phi_X(v)\Big|_{v=0} \tag{5.48}$$

Now, the characteristic function of a Gaussian random variable X of mean μ_X and variance σ_X^2 is given by (see Problem 5.1)

$$\phi_X(v) = \exp\left(jv\mu_X - \tfrac{1}{2}v^2\sigma_X^2\right) \tag{5.49}$$

Equations (5.48) and (5.49) show clearly that the higher-order moments of the Gaussian random variable are uniquely determined by the mean μ_X and variance σ_X^2. Indeed, a straightforward manipulation of this pair of equations shows that the central moments of X are as follows:

$$\mathbf{E}[(X-\mu_X)^n] = \begin{cases} 1\times 3\times 5\ldots(n-1)\sigma_X^n & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd} \end{cases} \tag{5.50}$$

JOINT MOMENTS

Consider next a pair of random variables X and Y. A set of statistical averages of importance in this case are the *joint moments,* namely, the expected value of X^iY^k, where i and k may assume any positive integer values. We may thus write

$$\mathbf{E}[X^iY^k] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i y^k f_{X,Y}(x,y)\,dx\,dy \tag{5.51}$$

A joint moment of particular importance is the *correlation* defined by $\mathbf{E}[XY]$, which corresponds to $i = k = 1$ in Eq. (5.51).

The correlation of the centered random variables $X - \mathbf{E}[X]$ and $Y - \mathbf{E}[Y]$, that is, the joint moment

$$\text{cov}[XY] = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])] \tag{5.52}$$

is called the *covariance* of X and Y. Letting $\mu_X = \mathbf{E}[X]$ and $\mu_Y = \mathbf{E}[Y]$, we may expand Eq. (5.52) to obtain the result

$$\text{cov}[XY] = \mathbf{E}[XY] - \mu_X\mu_Y \tag{5.53}$$

Let σ_X^2 and σ_Y^2 denote the variances of X and Y, respectively. Then the covariance of X and Y, normalized with respect to $\sigma_X\sigma_Y$, is called the *correlation coefficient* of X and Y:

$$\rho = \frac{\text{cov}[XY]}{\sigma_X\sigma_Y} \tag{5.54}$$

Jacob Bernoulli (1754–1801)

The Bernoullis were a family of Swiss scholars and traders that produced many renowned artists and scientists in the eighteenth century. Jacob (also known as James) Bernoulli is given credit for the notion of a *Bernoulli trial* and it is also after him that *Bernoulli numbers*, a sequence of rational numbers that have deep significance for number theory, are named. Jacob Bernoulli was the uncle of Daniel Bernoulli who was responsible for many significant developments in the theory of fluid dynamics and after whom *Bernoulli's principle* is named.

Jacob had two other nephews, both named Nicolas, who worked as mathematicians making contributions in the areas of probability, geometry, and differential equations. Much of Jacob's most well known work on probability and number theory was published after his death, including his introduction of the *law of large numbers*.

We say that the two random variables X and Y are *uncorrelated* if and only if their covariance is zero, that is, if and only if

$$\text{cov}[XY] = 0$$

We say that they are *orthogonal* if and only if their correlation is zero, that is, if and only if

$$\mathbf{E}[XY] = 0$$

From Eq. (5.53) we observe that if one of the random variables X and Y or both have zero means, and if they are orthogonal, then they are uncorrelated, and vice versa. Note also that if X and Y are statistically independent, then they are uncorrelated; however, the converse of this statement is not necessarily true.

EXAMPLE 5.6 Moments of a Bernouli Random Variable

Consider the coin-tossing experiment where the probability of a head is p. Let X be a random variable that takes the value 0 if the result is a tail and 1 if it is a head. We say that X is a *Bernoulli random variable*. The probability mass function of a Bernoulli random variable is

$$\mathbf{P}(X = x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

The expected value of X is

$$\begin{aligned} \mathbf{E}[X] &= \sum_{k=0}^{1} k\mathbf{P}(X = k) \\ &= 0 \cdot (1-p) + 1 \cdot p \\ &= p \end{aligned}$$

With $\mu_X = \mathbf{E}[X]$, the variance of X is given by

$$\begin{aligned} \sigma_X^2 &= \sum_{k=0}^{1} (k - \mu_X)^2 \, \mathbf{P}[X = k] \\ &= (0-p)^2(1-p) + (1-p)^2 p \\ &= (p^2 - p^3) + (p - 2p^2 + p^3) \\ &= p(1-p) \end{aligned}$$

Let $\{X_1, \ldots, X_N\}$ be a set of independent Bernouli random variables each with parameter p. Then the joint second moments are

$$\begin{aligned} \mathbf{E}[X_j X_k] &= \begin{cases} \mathbf{E}[X_j]\mathbf{E}[X_k] & j \neq k \\ \mathbf{E}[X_j^2] & j = k \end{cases} \\ &= \begin{cases} p^2 & j \neq k \\ p & j = k \end{cases} \end{aligned}$$

where the $\mathbf{E}[X_j^2] = \sum_{k=0}^{1} k^2 \mathbf{P}[X = k]$.

5.5 RANDOM PROCESSES

A basic concern in the statistical analysis of communication systems is the characterization of random signals such as voice signals, television signals, computer data, and electrical noise. These random signals have two properties. First, the signals are functions of time, defined on some observation interval. Second, the signals are random in the sense that before conducting an experiment, it is not possible to describe exactly the waveforms that will be observed. Accordingly, in describing random signals, we find that each sample point in our sample space is a function of time. The sample space or ensemble comprised of functions of time is called a *random* or *stochastic process*. As an integral part of this notion, we assume the existence of a probability distribution defined over an appropriate class of sets in the sample space, so that we may speak with confidence of the probability of various events.

Consider then a random experiment specified by the outcomes s from some *sample space* S, by the events defined on the sample space S, and by the probabilities of these events. Suppose that we assign to each sample point s a function of time in accordance with the rule:

$$X(t,s), \qquad -T \le t \le T \tag{5.55}$$

where $2T$ is the *total observation interval*. For a fixed sample point s_j, the graph of the function $X(t,s_j)$ versus time t is called a *realization* or *sample function* of the random process. To simplify the notation, we denote this sample function as

$$x_j(t) = X(t,s_j) \tag{5.56}$$

Figure 5.7 illustrates a set of sample functions $\{x_j(t) \mid j = 1, 2, \ldots, n\}$. From this figure, we note that for a fixed time t_k inside the observation interval, the set of numbers

$$\{x_1(t_k), x_2(t_k), \ldots, x_n(t_k)\} = \{X(t_k,s_1), X(t_k,s_2), \ldots, X(t_k,s_n)\}$$

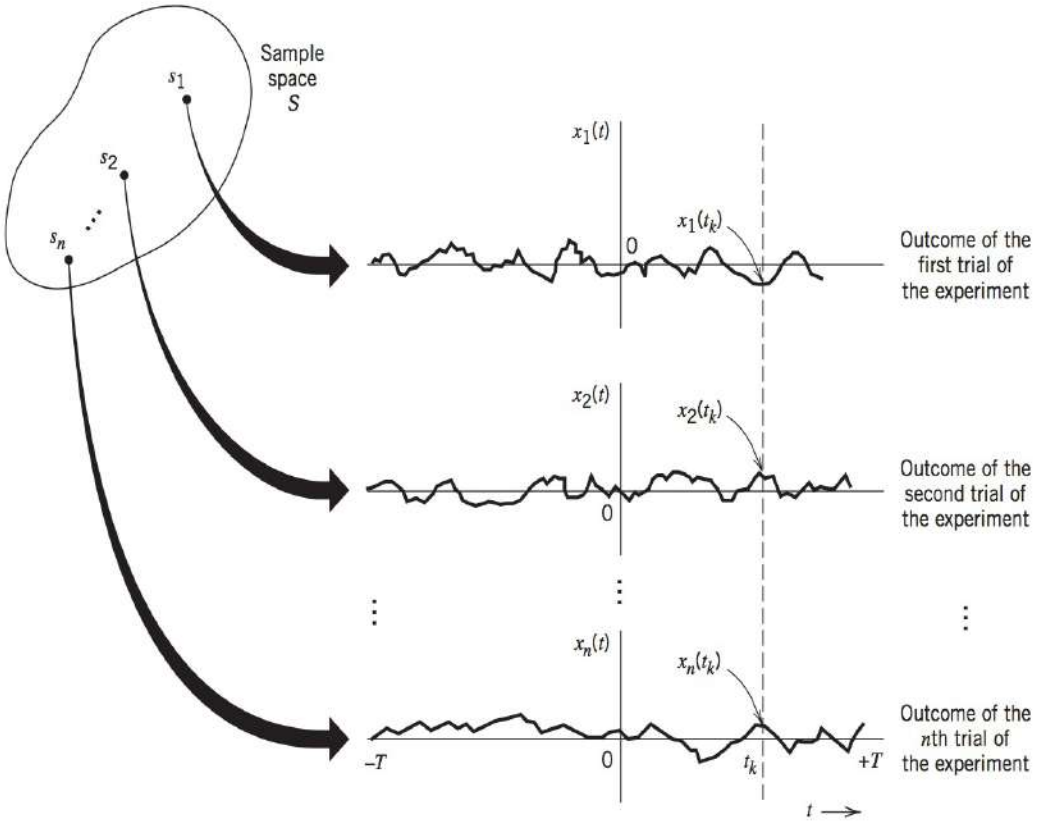


FIGURE 5.7 An ensemble of sample functions.

constitutes a *random variable.* Thus we have an indexed ensemble (family) of random variables $\{X(t,s)\}$, which is called a *random process.* To simplify the notation, the customary practice is to suppress the s and simply use $X(t)$ to denote a random process. We may then formally define a random process $X(t)$ as *an ensemble of time functions together with a probability rule that assigns a probability to any meaningful event associated with an observation of one of the sample functions of the random process.* Moreover, we may distinguish between a random variable and a random process as follows:

- For a random variable, the outcome of a random experiment is mapped into a number.
- For a random process, the outcome of a random experiment is mapped into a waveform that is a function of time.

5.6 MEAN, CORRELATION, AND COVARIANCE FUNCTIONS

Consider a random process $X(t)$. We define the mean of the process $X(t)$ as the expectation of the random variable obtained by observing the process at some time t, as shown by

$$\begin{aligned}\mu_X(t) &= \mathbf{E}[X(t)] \\ &= \int_{-\infty}^{\infty} x f_{X(t)}(x)\,dx\end{aligned} \tag{5.57}$$

where $f_{X(t)}(x)$ is the probability density function of the process at time t. A random process is said to be *stationary to first order* if the distribution function (and therefore density function) of $X(t)$ does not vary with time. That is, the density functions for the random variables $X(t_1)$ and $X(t_2)$ satisfy

$$f_{X(t_1)}(x) = f_{X(t_2)}(x) \tag{5.58}$$

for all t_1 and t_2. Consequently, we deduce that, for a process that is stationary to first order, the *mean of the random process is a constant*, as shown by

$$\mu_X(t) = \mu_X \qquad \text{for all } t \tag{5.59}$$

By a similar argument, we may also deduce that the variance of such a process is also constant.

We define the *autocorrelation function* of the process $X(t)$ as the expectation of the product of two random variables $X(t_1)$ and $X(t_2)$, obtained by observing $X(t)$ at times t_1 and t_2, respectively. Specifically, we write

$$\begin{aligned}R_X(t_1,t_2) &= \mathbf{E}[X(t_1)X(t_2)] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1),X(t_2)}(x_1,x_2)\,dx_1\,dx_2\end{aligned} \tag{5.60}$$

where $f_{X(t_1),X(t_2)}(x_1,x_2)$ is the joint probability density function of the random variables $X(t_1)$ and $X(t_2)$. We say a random process $X(t)$ is *stationary to second order* if the joint distribution function $f_{X(t_1),X(t_2)}(x_1,x_2)$ depends only on the difference between the observation times t_1 and t_2. This, in turn, implies that the autocorrelation function of a stationary (to second order) random process depends only on the time difference $t_2 - t_1$, as shown by

$$R_X(t_1,t_2) = R_X(t_2 - t_1) \qquad \text{for all } t_1 \text{ and } t_2 \tag{5.61}$$

Similarly, the *autocovariance function* of a stationary random process $X(t)$ is written as

$$\begin{aligned} C_X(t_1, t_2) &= \mathbf{E}[(X(t_1) - \mu_X)(X(t_2) - \mu_X)] \\ &= R_X(t_2 - t_1) - \mu_X^2 \end{aligned} \tag{5.62}$$

Equation (5.62) shows that, like the autocorrelation function, the autocovariance function of a stationary random process $X(t)$ depends only on the time difference $t_2 - t_1$. This equation also shows that if we know the mean and autocorrelation function of the process, we can readily determine the autocovariance function. The mean and autocorrelation function are therefore sufficient to describe the first two moments of the process.

However, there are two important points that should be carefully noted:

1. The mean and autocorrelation function only provide a *partial description* of the distribution of a random process $X(t)$.
2. The conditions of Eqs. (5.59) and (5.61) involving the mean and autocorrelation function, respectively, are *not* sufficient to guarantee that the random process $X(t)$ is stationary.

Nevertheless, practical considerations often dictate that we simply content ourselves with a partial description of the process given by the mean and autocorrelation function. A random process for which the conditions of Eqs. (5.59) and (5.61) hold is said to be *wide-sense stationary*.[2] Clearly, all strictly stationary processes are wide-sense stationary, but not all wide-sense stationary processes are strictly stationary.

PROPERTIES OF THE AUTOCORRELATION FUNCTION

For convenience of notation, we redefine the autocorrelation function of a stationary process $X(t)$ as

$$R_X(\tau) = \mathbf{E}[X(t + \tau)X(t)] \qquad \text{for all } t \tag{5.63}$$

This autocorrelation function has several important properties:

1. The mean-square value of the process may be obtained from $R_X(\tau)$ simply by putting $\tau = 0$ in Eq. (5.63), as shown by

$$R_X(0) = \mathbf{E}[X^2(t)] \tag{5.64}$$

2. The autocorrelation function $R_X(\tau)$ is an even function of τ, that is,

$$R_X(\tau) = R_X(-\tau) \tag{5.65}$$

This property follows directly from the defining equation (5.63). Accordingly, we may also define the autocorrelation function $R_X(\tau)$ as

$$R_X(\tau) = \mathbf{E}[X(t)X(t - \tau)] \tag{5.66}$$

3. The autocorrelation function $R_X(\tau)$ has its maximum magnitude at $\tau = 0$, that is,

$$|R_X(\tau)| \leq R_X(0) \tag{5.67}$$

To prove this property, consider the nonnegative quantity

$$\mathbf{E}[(X(t + \tau) \pm X(t))^2] \geq 0$$

Expanding terms and taking their individual expectations, we readily find that

$$\mathbf{E}[X^2(t + \tau)] \pm 2\mathbf{E}[X(t + \tau)X(t)] + \mathbf{E}[X^2(t)] \geq 0 \tag{5.68}$$

which, in light of Eqs. (5.63) and (5.64), reduces to

$$2R_X(0) \pm 2R_X(\tau) \geq (0) \tag{5.69}$$

Equivalently, we may write

$$-R_X(0) \leq R_X(\tau) \leq R_X(0) \tag{5.70}$$

from which Eq. (5.67) follows directly.

The physical significance of the autocorrelation function $R_X(\tau)$ is that it provides a means of describing the "interdependence" of two random variables obtained by observing a random process $X(t)$ at times τ seconds apart. It is therefore apparent that the more rapidly the random process $X(t)$ changes with time, the more rapidly will the autocorrelation function $R_x(\tau)$ decrease from its maximum $R_X(0)$ as τ increases, as illustrated in Figure 5.8. This decrease may be characterized by a *decorrelation time* τ_0, such that for $\tau > \tau_0$, the magnitude of the autocorrelation function $R_X(\tau)$ remains below some prescribed value. We may thus define the decorrelation time τ_0 of a wide-sense stationary random process $X(t)$ of zero mean as the time taken for the magnitude of the autocorrelation function $R_X(\tau)$ to decrease to 1 percent, say, of its maximum value $R_X(0)$.

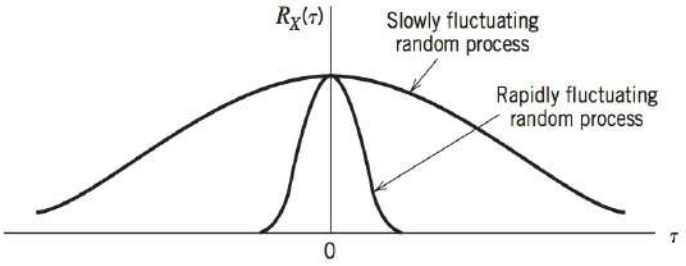


FIGURE 5.8 Illustrating the autocorrelation functions of slowly and rapidly fluctuating random processes.

EXAMPLE 5.7 Sinusoidal Signal with Random Phase

Consider a sinusoidal signal with random phase, defined by

$$X(t) = A\cos(2\pi f_c t + \Theta) \tag{5.71}$$

where A and f_c are constants and Θ is a random variable that is *uniformly distributed* over the interval $(-\pi,\pi)$, that is,

$$f_\Theta(\theta) = \begin{cases} \dfrac{1}{2\pi}, & -\pi \leq \theta \leq \pi \\ 0, & \text{elsewhere} \end{cases} \tag{5.72}$$

This means that the random variable Θ is equally likely to have any value in the interval $(-\pi,\pi)$. The autocorrelation function of $X(t)$ is

$$\begin{aligned} R_X(\tau) &= \mathbf{E}[X(t+\tau)X(t)] \\ &= \mathbf{E}[A^2\cos(2\pi f_c t + 2\pi f_c \tau + \Theta)\cos(2\pi f_c t + \Theta)] \\ &= \frac{A^2}{2}\mathbf{E}[\cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)] + \frac{A^2}{2}\mathbf{E}[\cos(2\pi f_c \tau)] \\ &= \frac{A^2}{2}\int_{-\pi}^{\pi}\frac{1}{2\pi}\cos(4\pi f_c t + 2\pi f_c \tau + 2\theta)\,d\theta + \frac{A^2}{2}\cos(2\pi f_c \tau) \end{aligned} \tag{5.73}$$

The first term integrates to zero, and so we get

$$R_X(\tau) = \frac{A^2}{2}\cos(2\pi f_c \tau) \tag{5.74}$$

which is shown plotted in Figure 5.9. We see therefore that the autocorrelation function of a sinusoidal signal with random phase is another sinusoid at the same frequency in the "τ domain" rather than the original time domain.

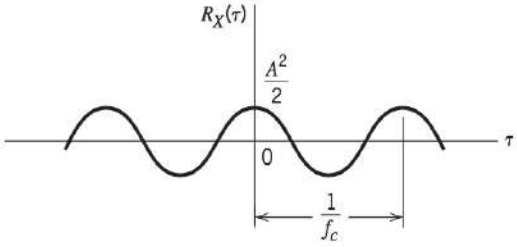


FIGURE 5.9 Autocorrelation function of a sinusoidal signal with random phase.

EXAMPLE 5.8 Random Binary Signal

Figure 5.10 shows the sample function $x(t)$ of a process $X(t)$ consisting of a random sequence of *binary symbols* 1 and 0. The following assumptions are made:

1. The symbols 1 and 0 are represented by pulses of amplitude $+A$ and $-A$ volts, respectively, and duration T seconds.
2. The pulses are not synchronized, so that the starting time t_d of the first complete pulse for positive time is equally likely to lie anywhere between zero and T seconds. That is, t_d is the sample value of a uniformly distributed random variable T_d, with its probability density function defined by

$$f_{T_d}(t_d) = \begin{cases} \dfrac{1}{T}, & 0 \le t_d \le T \\ 0, & \text{elsewhere} \end{cases}$$

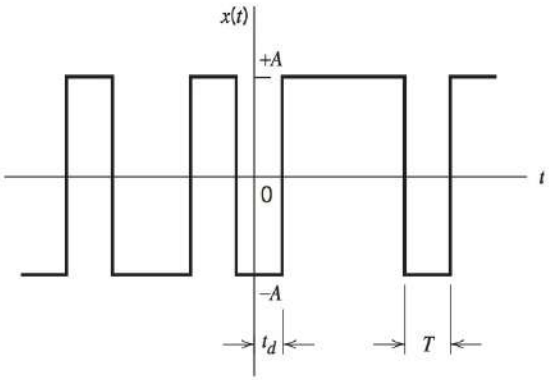


FIGURE 5.10 Sample function of random binary signal.

3. During any time interval $(n-1)\,T < t - t_d < nT$, where n is an integer, the presence of a 1 or a 0 is determined by tossing a fair coin; specifically, if the outcome is "heads," we have a 1 and if the outcome is "tails," we have a 0. These two symbols are thus equally likely, and the presence of a 1 or 0 in any one interval is independent of all other intervals.

Since the amplitude levels $-A$ and $+A$ occur with equal probability, it follows immediately that $\mathbf{E}[X(t)] = 0$ for all t, and the mean of the process is therefore zero.

To find the autocorrelation function $R_X(t_k,t_i)$, we have to evaluate $\mathbf{E}[X(t_k)X(t_i)]$, where $X(t_k)$ and $X(t_i)$ are random variables obtained by observing the random process $X(t)$ at times t_k and t_i, respectively.

Consider first the case when $|t_k - t_i| > T$. Then the random variables $X(t_k)$ and $X(t_i)$ occur in different pulse intervals and are therefore independent. We thus have

$$\mathbf{E}[X(t_k)X(t_i)] = \mathbf{E}[X(t_k)]\mathbf{E}[X(t_i)] = 0, \qquad |t_k - t_i| > T$$

Consider next the case when $|t_k - t_i| < T$, with $t_k = 0$ and $t_i < t_k$. In such a situation we observe from Figure 5.10 that the random variables $X(t_k)$ and $X(t_i)$ occur in the same pulse interval if and only if the delay t_d satisfies the condition $t_d < T - |t_k - t_i|$. We thus obtain the *conditional expectation*:

$$\mathbf{E}[X(t_k)X(t_i)|t_d] = \begin{cases} A^2, & t_d < T - |t_k - t_i| \\ 0, & \text{elsewhere} \end{cases} \tag{5.75}$$

Averaging this result over all possible values of t_d, we get

$$\begin{aligned} \mathbf{E}[X(t_k)X(t_i)] &= \int_0^{T-|t_k-t_i|} A^2 f_{T_d}(t_d)\, dt_d \\ &= \int_0^{T-|t_k-t_i|} \frac{A^2}{T}\, dt_d \\ &= A^2\left(1 - \frac{|t_k - t_i|}{T}\right), \qquad |t_k - t_i| < T \end{aligned} \tag{5.76}$$

By similar reasoning for any other value of t_k, we conclude that the autocorrelation function of a random binary signal, represented by the sample function depicted in Figure 5.10 is only a function of the time difference $\tau = t_k - t_i$, as shown by

$$R_X(\tau) = \begin{cases} A^2\left(1 - \dfrac{|\tau|}{T}\right), & |\tau| < T \\ 0, & |\tau| \geq T \end{cases} \tag{5.77}$$

This result is shown plotted in Figure 5.11.

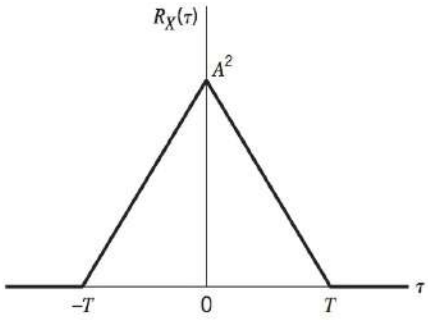


FIGURE 5.11 Autocorrelation function of random binary signal.

CROSS-CORRELATION FUNCTIONS

Consider next the more general case of two random processes $X(t)$ and $Y(t)$ with autocorrelation functions $R_X(t,u)$ and $R_Y(t,u)$, respectively. The *cross-correlation function* of $X(t)$ and $Y(t)$ is defined by

$$R_{XY}(t,u) = \mathbf{E}[X(t)Y(u)] \tag{5.78}$$

If the random processes $X(t)$ and $Y(t)$ are each wide-sense stationary and, in addition, are jointly wide-sense stationary, the cross-correlation may be written as

$$R_{XY}(t,u) = R_{XY}(\tau) \tag{5.79}$$

where $\tau = t - u$.

The cross-correlation function is not generally an even function of τ, as is true for the autocorrelation function, nor does it have a maximum at the origin. However, it does obey a certain symmetry relationship as follows (see Problem 5.12):

$$R_{XY}(\tau) = R_{YX}(-\tau) \tag{5.80}$$

EXAMPLE 5.9 Quadrature-Modulated Processes

Consider a pair of *quadrature-modulated processes* $X_1(t)$ and $X_2(t)$ that are related to a wide-sense stationary process $X(t)$ as follows:

$$\begin{aligned} X_1(t) &= X(t)\cos(2\pi f_c t + \Theta) \\ X_2(t) &= X(t)\sin(2\pi f_c t + \Theta) \end{aligned} \tag{5.81}$$

where f_c is a carrier frequency, and the random variable Θ is uniformly distributed over the interval $(0,2\pi)$. Moreover, Θ is independent of $X(t)$. One cross-correlation function of $X_1(t)$ and $X_2(t)$ is given by

$$\begin{aligned} R_{12}(\tau) &= \mathbf{E}[X_1(t)X_2(t-\tau)] \\ &= \mathbf{E}[X(t)X(t-\tau)\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t - 2\pi f_c \tau + \Theta)] \\ &= \mathbf{E}[X(t)X(t-\tau)]\mathbf{E}[\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t - 2\pi f_c \tau + \Theta)] \\ &= \frac{1}{2}R_X(\tau)\mathbf{E}[\sin(4\pi f_c t - 2\pi f_c t + 2\Theta) - \sin(2\pi f_c \tau)] \\ &= -\frac{1}{2}R_X(\tau)\sin(2\pi f_c \tau) \end{aligned} \tag{5.82}$$

where, in the last line, we have made use of the uniform distribution of the random variable Θ representing phase. Note that at $\tau = 0$, the factor $\sin(2\pi f_c \tau)$ is zero and therefore

$$\begin{aligned} R_{12}(0) &= \mathbf{E}[X_1(t)X_2(t)] \\ &= 0 \end{aligned} \tag{5.83}$$

Equation (5.83) shows that the random variables obtained by simultaneously observing the quadrature-modulated processes $X_1(t)$ and $X_2(t)$ at some fixed value of time t are orthogonal to each other.

ERGODIC PROCESSES[3]

The expectations of a stochastic process $X(t)$ are averages "across the process." For example, the mean of a stochastic process at some fixed time t_k is the expectation of the random variable $X(t_k)$ that describes all possible values of the sample functions of the process observed at time $t = t_k$. For this reason, the expectations of a stochastic process are often referred to as *ensemble averages*.

In many instances, it is difficult or impossible to observe all sample functions of a random process at a given time. It is often more convenient to observe a single sample function for a long period of time. For a single sample function, we may compute the *time average* of a particular function. For example, for a sample function $x(t)$, the time average of the mean value over an observation period $2T$ is

$$\mu_{x,T} = \frac{1}{2T}\int_{-T}^{T} x(t)\,dt \tag{5.84}$$

Fortunately, for many stochastic processes of interest in communications, the time averages and ensemble averages are equal, a property known as *ergodicity*. This property

implies that whenever an ensemble average is required, we may estimate it by using a time average. In what follows, we shall consider all random processes to be *ergodic*.

5.7 TRANSMISSION OF A RANDOM PROCESS THROUGH A LINEAR FILTER

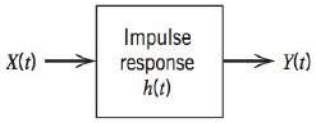


FIGURE 5.12 Transmission of a random process through a linear filter.

Suppose that a random process $X(t)$ is applied as input to a linear time-invariant filter of impulse response $h(t)$, producing a new random process $Y(t)$ at the filter output, as in Figure 5.12. In general, it is difficult to describe the probability distribution of the output random process $Y(t)$, even when the probability distribution of the input random process $X(t)$ is completely specified for $-\infty < t < \infty$.

In this section, we wish to determine the time-domain form of the input–output relations of the filter for defining the mean and autocorrelation functions of the output random process $Y(t)$ in terms of those of the input $X(t)$, assuming that $X(t)$ is a wide-sense stationary random process.

Consider first the mean of the output random process $Y(t)$. By definition, we have

$$\mu_Y(t) = \mathbf{E}[Y(t)] = \mathbf{E}\left[\int_{-\infty}^{\infty} h(\tau_1)X(t-\tau_1)\,d\tau_1\right] \tag{5.85}$$

where τ_1 is a dummy variable. Provided that the expectation $\mathbf{E}[X(t)]$ is finite for all t and the system is stable, we may interchange the order of the expectation and the integration with respect to τ_1 in Eq. (5.85), and so we write

$$\begin{aligned}\mu_Y(t) &= \int_{-\infty}^{\infty} h(\tau_1)\,\mathbf{E}[X(t-\tau_1)]\,d\tau_1 \\ &= \int_{-\infty}^{\infty} h(\tau_1)\mu_X(t-\tau_1)\,d\tau_1\end{aligned} \tag{5.86}$$

When the input random process $X(t)$ is wide-sense stationary, the mean $\mu_X(t)$ is a constant μ_X, so that we may simplify Eq. (5.86) as follows:

$$\begin{aligned}\mu_Y &= \mu_X \int_{-\infty}^{\infty} h(\tau_1)\,d\tau_1 \\ &= \mu_X H(0)\end{aligned} \tag{5.87}$$

where $H(0)$ is the zero-frequency (dc) response of the system. Equation (5.87) states that the mean of the random process $Y(t)$ produced at the output of a linear time-invariant system in response to $X(t)$ acting as the input process is equal to the mean of $X(t)$ multiplied by the dc response of the system, which is intuitively satisfying.

Consider next the autocorrelation function of the output random process $Y(t)$. By definition, we have

$$R_Y(t,u) = \mathbf{E}[Y(t)Y(u)]$$

where t and u denote two values of the time at which the output process is observed. We may therefore use the convolution integral to write

$$R_Y(t,u) = \mathbf{E}\left[\int_{-\infty}^{\infty} h(\tau_1)X(t-\tau_1)\,d\tau_1 \int_{-\infty}^{\infty} h(\tau_2)X(u-\tau_2)\,d\tau_2\right] \tag{5.88}$$

Here again, provided that the mean-square value $\mathbf{E}[X^2(t)]$ is finite for all t and the system is stable, we may interchange the order of the expectation and the integrations with respect to τ_1 and τ_2 in Eq. (5.88), obtaining

$$\begin{aligned} R_Y(t,u) &= \int_{-\infty}^{\infty} d\tau_1 h(\tau_1) \int_{-\infty}^{\infty} d\tau_2 h(\tau_2) \mathbf{E}[X(t-\tau_1)X(u-\tau_2)] \\ &= \int_{-\infty}^{\infty} d\tau_1 h(\tau_1) \int_{-\infty}^{\infty} d\tau_2 h(\tau_2) R_X(t-\tau_1, u-\tau_2) \end{aligned} \tag{5.89}$$

When the input $X(t)$ is a wide-sense stationary random process, the autocorrelation function of $X(t)$ is only a function of the difference between the observation times $t - \tau_1$ and $u - \tau_2$. Thus putting $\tau = t - u$ in Eq. (5.89), we may write

$$R_Y(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\tau_1) h(\tau_2) R_X(\tau - \tau_1 + \tau_2) d\tau_1 \, d\tau_2 \tag{5.90}$$

On combining this result with that involving the mean μ_Y, we see that *if the input to a stable linear time-invariant filter is a wide-sense stationary random process, then the output of the filter is also a wide-sense stationary random process.*

5.8 POWER SPECTRAL DENSITY

We found in Chapter 2, when analyzing deterministic time-domain signals, that the frequency-domain representation, the amplitude spectrum, is often very useful. The time- and frequency-domain representations of a signal are related by the Fourier transform. Since a sample function of a random process $X(t)$ is also a time-domain signal, we may define its Fourier transform. However, an individual sample function $x(t)$ may not be representative of the whole ensemble of sample functions that comprise a random process. A statistical average of the sample functions, such as the autocorrelation function $R_X(\tau)$, is often a more useful representation. The Fourier transform of the autocorrelation function is called the *power spectral density* $S_X(f)$ of the random process $X(t)$.

The power spectral density $S_X(f)$ and the autocorrelation function $R_X(\tau)$ of a wide-sense stationary random process $X(t)$ form a Fourier-transform pair with f and τ as the variables of interest, as shown by the pair of relations:

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) \, d\tau \tag{5.91}$$

and

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) \exp(j2\pi f\tau) \, df \tag{5.92}$$

Equations (5.91) and (5.92) are basic relations in the theory of spectral analysis of random processes, and together they constitute what are usually called the Einstein–Wiener–Khintchine relations.[4]

The Einstein–Wiener–Khintchine relations show that if either the autocorrelation function or power spectral density of a random process is known, the other can be found exactly. These functions display different aspects of the correlation properties of the process.

PROPERTIES OF THE POWER SPECTRAL DENSITY

We now use this pair of relations to derive some general properties of the power spectral density of a wide-sense stationary process.

PROPERTY 1

The zero-frequency value of the power spectral density of a wide-sense stationary random process equals the total area under the graph of the autocorrelation function; that is,

$$S_X(0) = \int_{-\infty}^{\infty} R_X(\tau)\, d\tau \tag{5.93}$$

This property follows directly from Eq. (5.91) by putting $f = 0$.

PROPERTY 2

The mean-square value of a wide-sense stationary random process equals the total area under the graph of the power spectral density; that is,

$$\mathbf{E}[X^2(t)] = \int_{-\infty}^{\infty} S_X(f)\, df \tag{5.94}$$

This property follows directly from Eq. (5.92) by putting $\tau = 0$ and noting that $R_X(0) = \mathbf{E}[X^2(t)]$.

PROPERTY 3

The power spectral density of a wide-sense stationary random process is always nonnegative; that is,

$$S_X(f) \geq 0 \qquad \text{for all } f \tag{5.95}$$

This property arises from the fact that the power spectral density $S_X(f)$ is closely related to the expected value of the magnitude squared of the amplitude spectrum of the random process $X(t)$, as shown by

$$S_X(f) \approx \mathbf{E}[|P(f)|^2]$$

The random process $P(f)$ with parameter f is the Fourier transform of $X(t)$. That is, every sample function of $P(f)$ is the Fourier transform of a sample function of $X(t)$.

PROPERTY 4

The power spectral density of a real-valued random process is an even function of frequency; that is,

$$S_X(-f) = S_X(f) \tag{5.96}$$

This property is readily obtained by substituting $-f$ for f in Eq. (5.91):

$$S_X(-f) = \int_{-\infty}^{\infty} R_X(\tau)\exp(j2\pi f\tau)\, d\tau$$

Next, substituting $-\tau$ for τ, and recognizing that $R_X(-\tau) = R_X(\tau)$, we get

$$S_X(-f) = \int_{-\infty}^{\infty} R_X(\tau)\exp(-j2\pi f\tau)\, d\tau = S_X(f)$$

which is the desired result.

EXAMPLE 5.10 Sinusoidal Signal with Random Phase (continued)

Consider the random process $X(t) = A\cos(2\pi f_c t + \Theta)$, where Θ is a uniformly distributed random variable over the interval $(-\pi,\pi)$. The autocorrelation function of this random process is given by Eq. (5.74), which is reproduced here for convenience:

$$R_X(\tau) = \frac{A^2}{2}\cos(2\pi f_c\tau)$$

Taking the Fourier transform of both sides of this relation, we find that the power spectral density of the sinusoidal process $X(t)$ is

$$S_X(f) = \frac{A^2}{4}\left[\delta(f - f_c) + \delta(f + f_c)\right] \tag{5.97}$$

which consists of a pair of delta functions weighted by the factor $A^2/4$ and located at $\pm f_c$, as in Figure 5.13. We note that the total area under a delta function is one. Hence, the total area under $S_X(f)$ is equal to $A^2/2$, as expected.

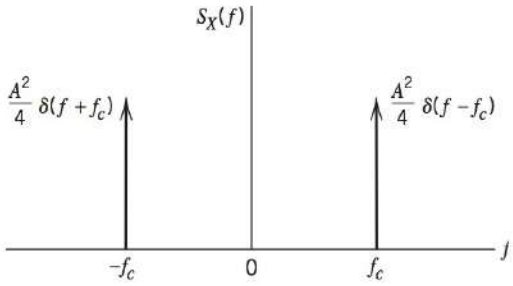


FIGURE 5.13 Power spectral density of sinusoidal signal with random phase.

EXAMPLE 5.11 Random Binary Signal (continued)

Consider again a random binary signal consisting of a sequence of 1s and 0s represented by the values $+A$ and $-A$, respectively. In Example 5.8 we showed that the autocorrelation function of this random process has a triangular wave-form, as shown by

$$R_X(\tau) = \begin{cases} A^2\left(1 - \dfrac{|\tau|}{T}\right), & |\tau| < T \\ 0, & |\tau| \geq T \end{cases}$$

The power spectral density of the process is therefore

$$S_X(f) = \int_{-T}^{T} A^2 \left(1 - \frac{|\tau|}{T}\right) \exp(-j2\pi f\tau)\, d\tau$$

Using the Fourier transform of a triangular function evaluated in Example 2.7 of Chapter 2, we obtain

$$S_X(f) = A^2 T \operatorname{sinc}^2(fT) \tag{5.98}$$

which is shown plotted in Figure 5.14. Here again we see that the power spectral density is nonnegative for all f and that it is an even function of f. Noting that $R_X(0) = A^2$ and using Property 3, we find that the total area under $S_X(f)$ or the average power of the random binary wave described here, is A^2.

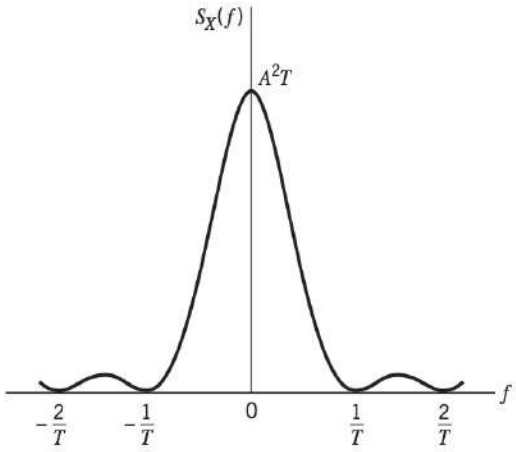


FIGURE 5.14 Power spectral density of random binary signal.

The result of Eq. (5.98) may be generalized as follows. We note that the energy spectral density of a rectangular pulse $g(t)$ of amplitude A and duration T is given by

$$\mathscr{E}_g(f) = A^2T^2 \operatorname{sinc}^2(fT) \tag{5.99}$$

We may therefore rewrite Eq. (5.98) in terms of $\mathscr{E}_g(f)$ as

$$S_X(f) = \frac{\mathscr{E}_g(f)}{T} \tag{5.100}$$

Equation (5.100) states that for a random binary wave in which binary symbols 1 and 0 are represented by pulses $g(t)$ and $-g(t)$, respectively, the power spectral density $S_X(f)$ is equal to the energy spectral density $\mathscr{E}_g(f)$ of the *symbol shaping pulse* $g(t)$, divided by the *symbol duration* T.

EXAMPLE 5.12 Mixing of a Random Process with a Sinusoidal Process

A situation that often arises in practice is that of *mixing* (i.e., multiplication) of a wide-sense stationary random process $X(t)$ with a sinusoidal signal $\cos(2\pi f_c t + \Theta)$, where the phase Θ is a random variable that is uniformly distributed over the interval $(0,2\pi)$. The addition of the random phase Θ in this manner merely recognizes the fact that the time origin is arbitrarily chosen when $X(t)$ and $\cos(2\pi f_c t + \Theta)$ come from physically independent sources, as is usually the case. We are interested in determining the power spectral density of the random process $Y(t)$ defined by

$$Y(t) = X(t)\cos(2\pi f_c t + \Theta) \tag{5.101}$$

Using the definition of autocorrelation function of a wide-sense stationary process, and noting that the random variable Θ is independent of $X(t)$, we

find that the autocorrelation function of $Y(t)$ is given by

$$\begin{aligned} R_Y(\tau) &= \mathbf{E}[Y(t+\tau)Y(t)] \\ &= \mathbf{E}[X(t+\tau)\cos(2\pi f_c t + 2\pi f_c \tau + \Theta)X(t)\cos(2\pi f_c t + \Theta)] \\ &= \mathbf{E}[X(t+\tau)X(t)]\mathbf{E}[\cos(2\pi f_c t + 2\pi f_c \tau + \Theta)\cos(2\pi f_c t + \Theta)] \\ &= \frac{1}{2}R_X(\tau)\mathbf{E}[\cos(2\pi f_c \tau) + \cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)] \\ &= \frac{1}{2}R_X(\tau)\cos(2\pi f_c \tau) \end{aligned} \tag{5.102}$$

Because the power spectral density is the Fourier transform of the autocorrelation function, we find that the power spectral densities of the random process $X(t)$ and $Y(t)$ are related as follows:

$$S_Y(f) = \frac{1}{4}[S_X(f - f_c) + S_X(f + f_c)] \tag{5.103}$$

According to Eq. (5.103), the power spectral density of the random process $Y(t)$ defined in Eq. (5.101) is obtained as follows: we shift the given power spectral density $S_X(f)$ of random process $X(t)$ to the right by f_c, shift it to the left by f_c, add the two shifted power spectra, and divide the result by 4.

RELATION AMONG THE POWER SPECTRAL DENSITIES OF THE INPUT AND OUTPUT RANDOM PROCESSES

Let $S_Y(f)$ denote the power spectral density of the output random process $Y(t)$ obtained by passing the random process $X(t)$ through a linear filter of transfer function $H(f)$. Then, recognizing by definition that the power spectral density of a random process is equal to the Fourier transform of its autocorrelation function and using Eq. (5.90), we obtain

$$\begin{aligned} S_Y(f) &= \int_{-\infty}^{\infty} R_Y(\tau)\exp(-j2\pi f\tau)\, d\tau \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau - \tau_1 + \tau_2)\exp(-j2\pi f\tau)\, d\tau_1\, d\tau_2\, d\tau \end{aligned} \tag{5.104}$$

Let $\tau - \tau_1 + \tau_2 = \tau_0$, or, equivalently, $\tau = \tau_0 + \tau_1 - \tau_2$. Then, by making this substitution in Eq. (5.104), we find that $S_Y(f)$ may be expressed as the product of three terms: the transfer function $H(f)$ of the filter, the complex conjugate of $H(f)$, and the power spectral density $S_X(f)$ of the input random process $X(t)$. We may thus simplify Eq. (5.104) as

$$S_Y(f) = H(f)H^*(f)S_X(f) \tag{5.105}$$

Finally, since $|H(f)|^2 = H(f)H^*(f)$, we find that the relationship among the power spectral densities of the input and output random processes is expressed in the frequency domain by writing

$$S_Y(f) = |H(f)|^2 S_X(f) \tag{5.106}$$

Equation (5.106) states that *the power spectral density of the output process $Y(t)$ equals the power spectral density of the input process $X(t)$ multiplied by the squared magnitude of the transfer function $H(f)$ of the filter*. By using this relation, we can therefore determine the effect of passing a random process through a stable, linear, time-invariant, filter. In computational terms, Eq. (5.106) is usually easier to handle than its time-domain counterpart of Eq. (5.90), involving the autocorrelation function.

EXAMPLE 5.13 Comb Filter

Consider the filter of Figure 5.15*a* consisting of a delay line and a summing device. We wish to evaluate the power spectral density of the filter output $Y(t)$, given that the power spectral density of the filter input $X(t)$ is $S_X(f)$.

The transfer function of this filter is

$$\begin{aligned} H(f) &= 1 - \exp(-j2\pi fT) \\ &= 1 - \cos(2\pi fT) + j\sin(2\pi fT) \end{aligned}$$

The squared magnitude of $H(f)$ is

$$\begin{aligned} |H(f)|^2 &= [1 - \cos(2\pi fT)]^2 + \sin^2(2\pi fT) \\ &= 2[1 - \cos(2\pi fT)] \\ &= 4\sin^2(\pi fT) \end{aligned}$$

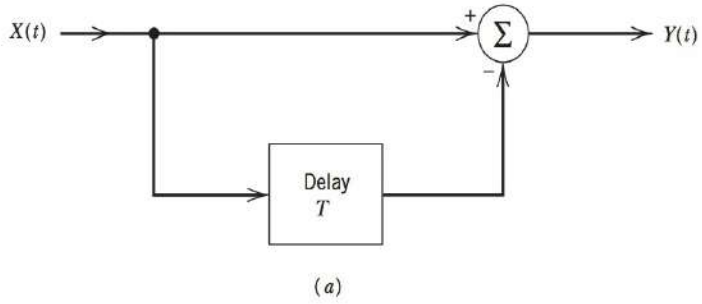
which is shown plotted in Figure 5.15*b*. Because of the periodic form of this frequency response, the filter of Figure 5.15*a* is sometimes referred to as a *comb filter*.

The power spectral density of the filler output is therefore

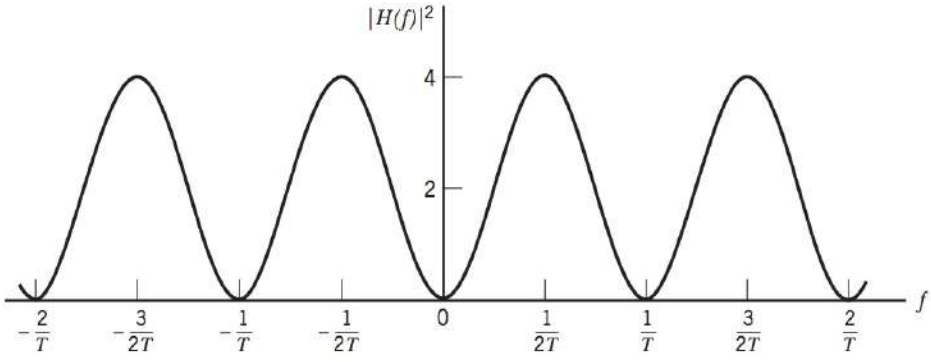
$$S_Y(f) = 4\sin^2(\pi fT)S_X(f)$$

For values of frequency f that are small compared to $1/T$, we have

$$\sin(\pi fT) \simeq \pi fT$$



(*a*)



(*b*)

FIGURE 5.15 Comb filter. (*a*) Block diagram. (*b*) Frequency response.

Under this condition, we may approximate the expression for $S_Y(f)$ as follows:

$$S_Y(f) \simeq 4\pi^2 f^2 T^2 S_X(f) \tag{5.107}$$

Because differentiation in the time domain corresponds to multiplication by $j2\pi f$ in the frequency domain, we see from Eq. (5.107) that the comb filter of Figure 5.15*a* acts as a differentiator for low-frequency inputs.

5.9 GAUSSIAN PROCESS

The material we have presented on random processes up to this point in the discussion has been of a fairly general nature. In this section, we consider an important family of random processes known as Gaussian processes.

Let us suppose that we observe a random process $X(t)$ for an interval that starts at time $t = 0$ and lasts until $t = T$. Suppose also that we weight the random process $X(t)$ by some function $g(t)$ and then integrate the product $g(t)X(t)$ over this observation interval, thereby obtaining a random variable Y defined by

$$Y = \int_0^T g(t)X(t)\, dt \tag{5.108}$$

We refer to Y as a *linear functional of* $X(t)$. The distinction between a function and a functional should be carefully noted. For example, the sum $Y = \Sigma_{i=1}^{N} a_i X_i$, where the a_i are constants and the X_i are random variables, is a linear *function* of the X_i; for each observed set of values for the random variables X_i, we have a corresponding value for the random variable Y. On the other hand, in Eq. (5.108) the value of the random variable Y depends on the course of the *argument function* $g(t)X(t)$ over the entire observation interval from 0 to T. Thus a functional is a quantity that depends on the entire course of one or more functions rather than on a number of discrete variables. In other words, the domain of a functional is a set or space of admissible functions rather than a region of a coordinate space.

If in Eq. (5.108) the weighting function $g(t)$ is such that the mean-square value of the random variable Y is finite, and if the random variable Y is a *Gaussian-distributed* random variable for every $g(t)$ in this class of functions, then the process $X(t)$ is said to be a *Gaussian process*. In other words, the process $X(t)$ is a Gaussian process if every linear functional of $X(t)$ is a Gaussian random variable.

In Example 5.5 we described the characterization of a Gaussian random variable. We say that the random variable Y has a Gaussian distribution if its probability density function has the form

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left[-\frac{(y-\mu_Y)^2}{2\sigma_Y^2} \right] \tag{5.109}$$

where μ_Y is the mean and σ_Y^2 is the variance of the random variable Y. A plot of this probability density function is given in Figure 5.16 for the special case when the Gaussian random variable Y is *normalized* to have a mean μ_Y of zero and a variance σ_Y^2 of one, as shown by

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2} \right)$$

Such a normalized Gaussian distribution is commonly written as $\mathcal{N}(0,1)$.

Carl F. Gauss (1777–1855)

Gauss was a child prodigy who later made numerous contributions in many areas of mathematics and science. Legend has it that, while in school, his teacher assigned the class the problem of adding up the numbers from 1 to 100. The teacher was quite shocked when Gauss provided the answer in seconds, having discovered the formula for arithmetic sums.

At age 18, Gauss invented the *method of least squares* for finding the best value of a sequence of measurements of some quantity. Gauss later used the method of least squares in fitting orbits of planets to data measurements, a procedure that was published in 1809 in his book entitled *Theory of Motion of the Heavenly Bodies*. In Gauss's early years many of his contributions were in the area of mathematics and then in astronomy. While involved in the survey of an area of Germany near Hanover, Gauss formulated the Gaussian distribution to describe measurement errors.

In 1833 Gauss, in collaboration with Wilhem Weber, made a contribution to communications when they constructed the first magnetic telegraph. Gauss is so renowned in Germany that his likeness and the Gaussian distribution appear on the former German ten Mark bill.

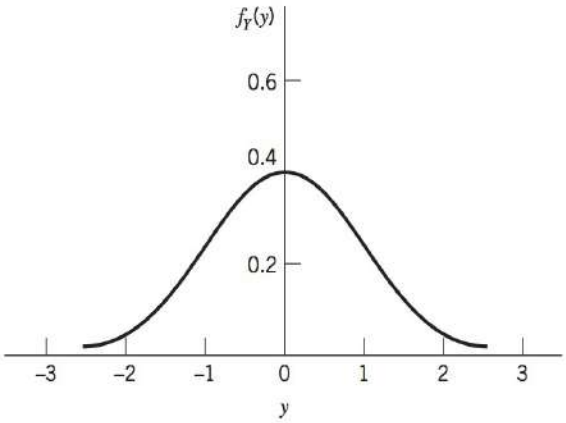


FIGURE 5.16 Normalized Gaussian distribution.

A Gaussian process has two main virtues. First, the Gaussian process has many properties that make analytic results possible; we will discuss these properties later in the section. Second, the random processes produced by physical phenomena are often such that a Gaussian model is appropriate. Furthermore, the use of a Gaussian model to describe such physical phenomena is usually confirmed by experiments. Thus the widespread occurrence of physical phenomena for which a Gaussian model is appropriate, together with the case with which a Gaussian process is handled mathematically, make the Gaussian process very important in the study of communication systems.

CENTRAL LIMIT THEOREM

The *central limit theorem* provides the mathematical justification for using a Gaussian process as a model for a large number of different physical phenomena in which the observed random variable, at a particular instant of time, is the result of a large number of individual random events. To formulate this important theorem, let X_i, $i = 1, 2, \ldots, N$, be a set of random variables that satisfies the following requirements:

1. The X_i are statistically independent.
2. The X_i have the same probability distribution with mean μ_X and variance σ_X^2.

The X_i so described are said to constitute a set of *independently and identically distributed* (i.i.d.) random variables. Let these random variables be *normalized* as follows;

$$Y_i = \frac{1}{\sigma_X}(X_i - \mu_X), \qquad i = 1, 2, \ldots, N$$

so that we have

$$\mathbf{E}[Y_i] = 0$$

and

$$\text{var}[Y_i] = 1$$

Define the random variable

$$V_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y_i$$

The central limit theorem states that the probability distribution of V_N approaches a normalized Gaussian distribution $\mathcal{N}(0,1)$ in the limit as N approaches infinity. That is, regardless of the distribution of the Y_i, the sum V_N approaches a Gaussian distribution.

It is important to realize, however, that the central limit theorem gives only the "limiting" form of the probability distribution of the normalized random variable V_N as N approaches infinity. When N is finite, the Gaussian limit is most accurate in the central portion of the density function (hence, the central limit) and less accurate in the "tails" of the density function (see Problem 5.36).

PROPERTIES OF A GAUSSIAN PROCESS

A Gaussian process has some useful properties that are described in the sequel.

PROPERTY 1

If a Gaussian process $X(t)$ is applied to a stable linear filter, then the random process $Y(t)$ developed at the output of the filter is also Gaussian.

This property is readily derived by using the definition of a Gaussian process based on Eq. (5.108). Consider the situation depicted in Figure 5.12, where we have a linear time-invariant filter of impulse response $h(t)$, with the random process $X(t)$ as input and the random process $Y(t)$ as output. We assume that $X(t)$ is a Gaussian process. The random processes $Y(t)$ and $X(t)$ are related by the convolution integral

$$Y(t) = \int_0^T h(t-\tau)X(\tau)\,d\tau, \qquad 0 \le t < \infty \tag{5.110}$$

We assume that the impulse response $h(t)$ is such that the mean-square value of the output random process $Y(t)$ is finite for all t in the range $0 \le t < \infty$ for which $Y(t)$ is defined. To demonstrate that the output process $Y(t)$ is Gaussian, we must show that any linear functional of it is a Gaussian random variable. That is, if we define the random variable

$$Z = \int_0^\infty g_Y(t) \int_0^T h(t-\tau)X(\tau)\,d\tau\,dt \tag{5.111}$$

then Z must be a Gaussian random variable for every function $g_Y(t)$, such that the mean-square value of Z is finite. Interchanging the order of integration in Eq. (5.111), we get

$$Z = \int_0^T g(\tau)X(\tau)\,d\tau \tag{5.112}$$

where

$$g(\tau) = \int_0^\infty g_Y(t)h(t-\tau)\,dt \tag{5.113}$$

Since $X(t)$ is a Gaussian process by hypothesis, it follows from Eq. (5.112) that Z must be a Gaussian random variable. We have thus shown that if the input $X(t)$ to a linear filter is a Gaussian process, then the output $Y(t)$ is also a Gaussian process. Note, however, that although our proof was carried out assuming a time-invariant linear filter, this property is true for any arbitrary stable linear system.

PROPERTY 2

Consider the set of random variables or samples $X(t_1), X(t_2), \ldots, X(t_n)$, obtained by observing a random process $X(t)$ at times $t_1, t_2, \ldots, t_n$. If the process $X(t)$ is Gaussian, then this set of random variables is jointly Gaussian

for any n, with their n-fold joint probability density function being completely determined by specifying the set of means

$$\mu_{X(t_i)} = \mathbf{E}[X(t_i)], \qquad i = 1, 2, \ldots, n$$

and the set of autocovariance functions

$$C_X(t_k,t_i) = \mathbf{E}[(X(t_k) - \mu_{X(t_k)})(X(t_i) - \mu_{X(t_i)})], \qquad k, i = 1, 2, \ldots, n$$

Property 2 is frequently used as the definition of a Gaussian process. However, this definition is more difficult to use than that based on Eq. (5.108) for evaluating the effects of filtering on a Gaussian process.

We may extend Property 2 to two (or more) random processes as follows. Consider the composite set of random variables $X(t_1), X(t_2), \ldots, X(t_n), Y(u_1), Y(u_2), \ldots, Y(u_m)$ obtained by observing a random process $X(t)$ at times $\{t_i, i = 1, 2, \ldots, n\}$, and a second random process $Y(t)$ at times $\{u_k, k = 1, 2, \ldots, m\}$. We say that the processes $X(t)$ and $Y(t)$ are *jointly Gaussian* if this composite set of random variables are jointly Gaussian for any n and m. Note that in addition to the mean and correlation functions of the random processes $X(t)$ and $Y(t)$ individually, we must also know the cross-covariance function

$$\mathbf{E}[(X(t_i) - \mu_{X(t_i)})(X(u_k) - \mu_{X(u_k)})] = R_{XY}(t_i,u_k) - \mu_{X(t_i)}\mu_{Y(u_k)}$$

for any pair of observation instants (t_i,u_k). This additional knowledge is embodied in the cross-correlation function, $R_{XY}(t_i,u_k)$, of the two processes $X(t)$ and $Y(t)$.

PROPERTY 3

If a Gaussian process is wide-sense stationary, then the process is also stationary in the strict sense.

This follows directly from Property 2.

PROPERTY 4

If the random variables $X(t_1), X(t_2), \ldots, X(t_n)$, obtained by sampling a Gaussian process $X(t)$ at times $t_1, t_2, \ldots, t_m$, are uncorrelated, that is,

$$\mathbf{E}[(X(t_k) - \mu_{X(t_k)})(X(t_i) - \mu_{X(t_i)})] = 0, \qquad i \neq k$$

then these random variables are statistically independent.

The implication of this property is that the joint probability density function of the set of random variables $X(t_1), X(t_2), \ldots, X(t_n)$ can be expressed as the product of the probability density functions of the individual random variables in the set.

5.10 NOISE

The term *noise* is used customarily to designate unwanted waves that tend to disturb the transmission and processing of signals in communication systems and over which we have incomplete control. In practice, we find that there are many potential sources of noise in a communication system. The sources of noise may be external to the system (e.g., atmospheric noise, galactic noise, man-made noise), or internal to the system. The second category includes an important type of noise that arises from *spontaneous fluctuations* of current or voltage in electrical circuits.[6] This type of noise represents a basic limitation on the transmission or detection of signals in communication systems involving the use of electronic devices. The two most common examples of spontaneous fluctuations in electrical circuits are *shot noise* and *thermal noise*.

SHOT NOISE

Shot noise arises in electronic devices such as diodes and transistors because of the discrete nature of current flow in these devices. For example, in a *photodetector* circuit a current pulse is generated every time an electron is emitted by the cathode due to incident light from a source of constant intensity. The electrons are naturally emitted at random times denoted by τ_k, where $-\infty < k < \infty$. It is assumed that the random emissions of electrons have been going on for a long time. Thus, the total current flowing through the photodetector may be modeled as an infinite sum of current pulses, as shown by

$$X(t) = \sum_{k=-\infty}^{\infty} h(t - \tau_k) \tag{5.114}$$

where $h(t - \tau_k)$ is the current pulse generated at time τ_k. The process $X(t)$ defined by Eq. (5.114) is a stationary process, called *shot noise*.[7]

The number of electrons, $N(t)$, emitted in the time interval $(0, t)$ constitutes a discrete stochastic process, the value of which increases by one each time an electron is emitted. Figure 5.17 shows a sample function of such a process. Let the mean value of the number of electrons, v, emitted between times t and $t + t_0$ be defined by

$$\mathbf{E}[v] = \lambda t_0 \tag{5.115}$$

The parameter λ is a constant called the *rate* of the process. The total number of electrons emitted in the interval $(t, t + t_0)$, that is,

$$v = N(t + t_0) - N(t)$$

follows a *Poisson distribution* with a mean value equal to λt_0. In particular, the probability that k electrons are emitted in the interval $(t, t + t_0)$ is defined by

$$\mathbf{P}[v = k] = \frac{(\lambda t_0)^k}{k!} e^{-\lambda k} \qquad k = 0, 1, \ldots \tag{5.116}$$

Unfortunately, a detailed statistical characterization of the shot-noise process $X(t)$ defined in Eq. (5.114) is a difficult mathematical task. Here we simply quote the results pertaining to the first two moments of the process:

- The mean of $X(t)$ is

$$\mu_X = \lambda \int_{-\infty}^{\infty} h(t)\, dt \tag{5.117}$$

where λ is the rate of the process and $h(t)$ is the waveform of a current pulse.

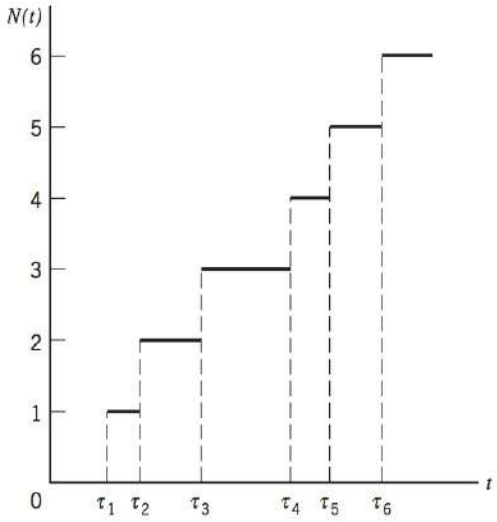


FIGURE 5.17 Sample function of a Poisson counting process.

- The autocovariance function of $X(t)$ is

$$C_X(\tau) = \lambda \int_{-\infty}^{\infty} h(t)h(t+\tau)\,dt \tag{5.118}$$

This latter result is known as *Campbell's theorem.*

For the special case of a waveform $h(t)$ consisting of a rectangular pulse of amplitude A and duration T, the mean of the shot-noise process $X(t)$ is λAT, and its autocovariance function is

$$C_X(\tau) = \begin{cases} \lambda A^2(T - |\tau|), & |\tau| < T \\ 0, & |\tau| \geq T \end{cases}$$

which has a triangular form similar to that shown in Figure 5.11.

THERMAL NOISE

Thermal noise[8] is the name given to the electrical noise arising from the random motion of electrons in a conductor. The mean-square value of the thermal noise voltage V_{TN} appearing across the terminals of a resistor, measured in a bandwidth of Δf Hertz, is, for all practical purposes, given by

$$\mathbf{E}[V_{TN}^2] = 4kTR\,\Delta f \text{ volts}^2 \tag{5.119}$$

where k is *Boltzmann's constant* equal to 1.38×10^{-23} joules per degree Kelvin, T is the *absolute temperature* in degrees Kelvin, and R is the resistance in ohms. We may thus model a noisy resistor by the *Thévenin equivalent circuit* consisting of a noise voltage generator of mean-square value $\mathbf{E}[V_{TN}^2]$ in series with a noiseless resistor, as in Figure 5.18*a*. Alternatively, we may use the *Norton equivalent circuit* consisting of a noise current generator in parallel with a noiseless conductance, as in Figure 5.18*b*. The mean-square value of the noise current generator is

$$\begin{aligned} \mathbf{E}[I_{TN}^2] &= \frac{1}{R^2}\mathbf{E}[V_{TN}^2] \\ &= 4kTG\,\Delta f \text{ amps}^2 \end{aligned} \tag{5.120}$$

where G = 1/R is the conductance. It is also of interest to note that because the number of electrons in a resistor is very large and their random motions inside the resistor are statistically independent of each other, the central limit theorem indicates that thermal noise is Gaussian distributed with zero mean.

Noise calculations involve the transfer of power, and so we find that the use of the *maximum-power transfer theorem* is applicable to such calculations. This theorem states that the maximum possible power is transferred from a source of internal resistance R to a load of resistance R_l when $R_l = R$. Under this *matched condition*, the power produced by the source is divided equally between the internal resistance of the source and the load resistance, and the power delivered to the load is referred to as the *available power.* Applying the maximum-power transfer theorem to the Thévenin equivalent circuit of Figure 5.18*a* or the Norton equivalent circuit of Figure 5.18*b*, we find that a noisy resistor produces an *available noise power* equal to $kT\,\Delta f$ watts.

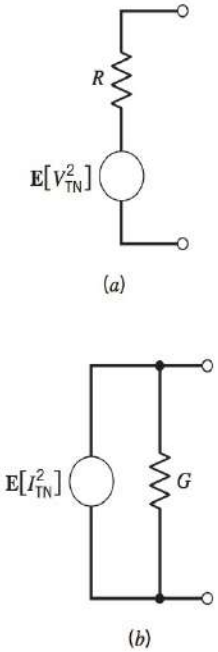


FIGURE 5.18 Models of a noisy resistor. (*a*) Thévenin equivalent circuit. (*b*) Norton equivalent circuit.

The noise analysis of communication systems is customarily based on an idealized form of noise called *white noise*, the power spectral density of which is independent of the operating frequency. The adjective *white* is used in the sense that white light contains equal amounts of all frequencies within the visible band of electromagnetic radiation. We express the power spectral density of white noise, with a sample function denoted by $w(t)$, as

$$S_W(f) = \frac{N_0}{2} \tag{5.121}$$

which is illustrated in Figure 5.19*a*. The dimensions of N_0 are in watts per Hertz. The parameter N_0 is usually referenced to the input stage of the receiver of a communication system. It may be expressed as

$$N_0 = kT_e \tag{5.122}$$

where k is Boltzmann's constant and T_e is the *equivalent noise temperature* of the receiver.[9] *The equivalent noise temperature of a system is defined as the temperature at which a noisy resistor has to be maintained such that, by connecting the resistor to the input of a noiseless version of the system, it produces the same available noise power at the output of the system as that produced by all the sources of noise in the actual system.* The important feature of the equivalent noise temperature is that it depends only on the parameters of the system.

Since the autocorrelation function is the inverse Fourier transform of the power spectral density, it follows that for white noise

$$R_W(\tau) = \frac{N_0}{2}\delta(\tau) \tag{5.123}$$

That is, the autocorrelation function of white noise consists of a delta function weighted by the factor $N_0/2$ and occurring at $\tau = 0$, as in Figure 5.19*b*. We note that $R_W(\tau)$ is zero for $\tau \neq 0$. Accordingly, any two different samples of white noise, no matter how closely together in time they are taken, are uncorrelated. If the white noise $w(t)$ is also Gaussian, then the two samples are statistically independent. In a sense, white Gaussian noise represents the ultimate in "randomness."

Strictly speaking, white noise has infinite average power and, as such, it is not physically realizable. Nevertheless, white noise has simple mathematical properties exemplified by Eqs. (5.121) and (5.123), which make it useful in statistical system analysis.

The utility of a white noise process is parallel to that of an impulse function or delta function in the analysis of linear systems. Just as we may observe the effect of an impulse only after it has been passed through a system with a finite bandwidth, so it is with white noise whose effect is observed only after passing through a similar system. We may state, therefore, that as long as the bandwidth of a noise process at the input of a system is appreciably larger than that of the system itself, then we may model the noise process as white noise.

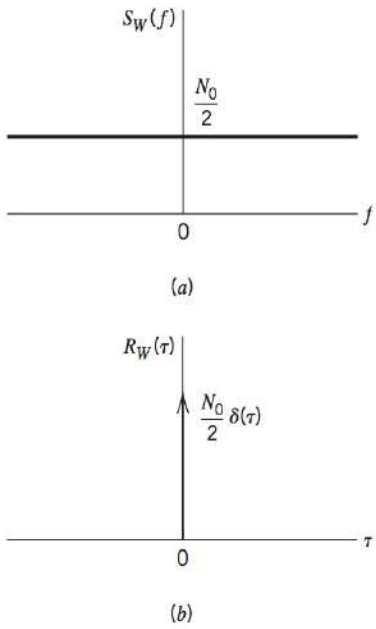


FIGURE 5.19 Characteristics of white noise. (*a*) Power spectral density. (*b*) Autocorrelation function.

EXAMPLE 5.14 Ideal Low-Pass Filtered White Noise

Suppose that a white Gaussian noise $w(t)$ of zero mean and power spectral density $N_0/2$ is applied to an ideal low-pass filter of bandwidth B and passband amplitude response of one. The power spectral density of the noise $n(t)$

appearing at the filter output is therefore (see Figure 5.20*a*)

$$S_N(f) = \begin{cases} \dfrac{N_0}{2}, & -B < f < B \\ 0, & |f| > B \end{cases} \tag{5.124}$$

The autocorrelation function of $n(t)$ is the inverse Fourier transform of the power spectral density shown in Figure 5.20*a*:

$$\begin{aligned} R_N(\tau) &= \int_{-B}^{B} \frac{N_0}{2} \exp(j2\pi f\tau)\, df \\ &= N_0 B \operatorname{sinc}(2B\tau) \end{aligned} \tag{5.125}$$

This autocorrelation function is shown plotted in Figure 5.20*b*. We see that $R_N(\tau)$ has its maximum value of N_0B at the origin, and it passes through zero at $\tau = \pm k/2B$, where $k = 1, 2, 3, \ldots$.

Since the input noise $w(t)$ is Gaussian (by hypothesis), it follows that the band-limited noise $n(t)$ at the filter output is also Gaussian. Suppose now that $n(t)$ is sampled at the rate of $2B$ times per second. From Figure 5.20*b*, we see that the resulting noise samples are uncorrelated and, being Gaussian, they are statistically independent. Accordingly, the joint probability density function of a set of noise samples obtained in this way is equal to the product of the individual probability density functions. Note that each such noise sample has a mean of zero and variance of N_0B.

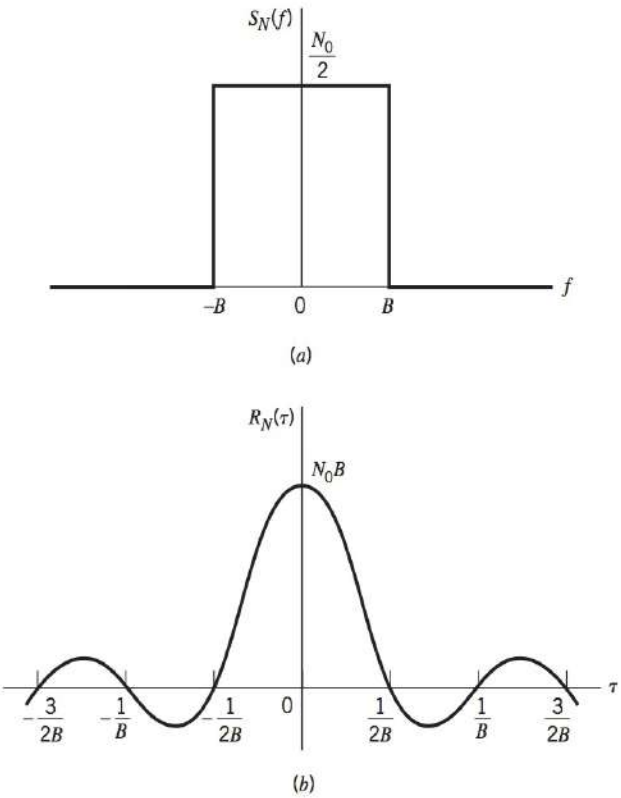


FIGURE 5.20 Characteristics of low-pass filtered white noise. (*a*) Power spectral density. (*b*) Autocorrelation function.

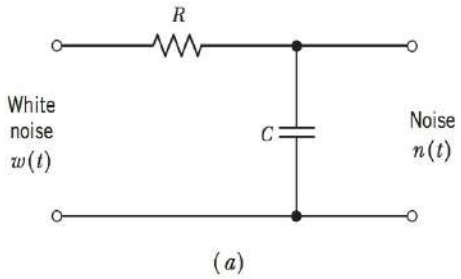
EXAMPLE 5.15 RC Low-Pass Filtered White Noise

Consider next a white Gaussian noise $w(t)$ of zero mean and power spectral density $N_0/2$ applied to a low-pass *RC* filter, as in Figure 5.21*a*. The transfer function of the filter is

$$H(f) = \frac{1}{1 + j2\pi f\, RC}$$

The power spectral density of the noise $n(t)$ appearing at the low-pass *RC* filter output is therefore (see Figure 5.21*b*)

$$S_N(f) = \frac{N_0/2}{1 + (2\pi f\, RC)^2}$$



(a)

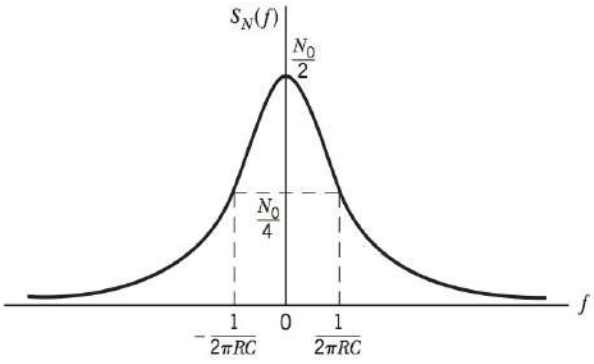
From Example 2.3 of Chapter 2, we recall the following Fourier-transform pair (using τ in place of t as the time variable to suit the problem at hand):

$$\exp(-a|\tau|) \rightleftharpoons \frac{2a}{a^2 + (2\pi f)^2} \tag{5.126}$$

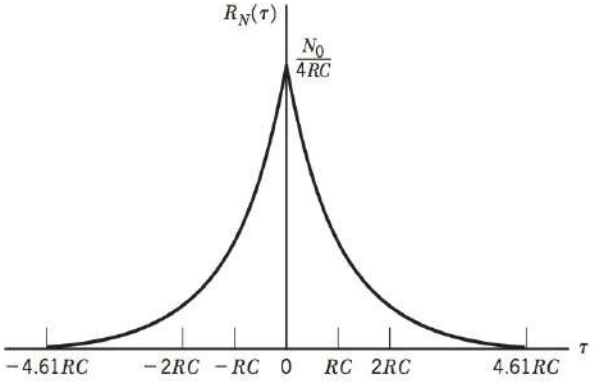
where a is a constant. Therefore, setting $a = 1/RC$, we find that the autocorrelation function of the filtered noise $n(t)$ is

$$R_N(\tau) = \frac{N_0}{4RC}\exp\left(-\frac{|\tau|}{RC}\right) \tag{5.127}$$

which is shown plotted in Figure 5.21. The decorrelation time τ_0 for which $R_N(\tau)$ drops to 1 percent, say, of its maximum value of $N_0/4RC$ is equal to $4.61RC$. Thus, if the noise appearing at the filter output is sampled at a rate equal to or less than $0.217/RC$ samples per second, the resulting samples are essentially uncorrelated and, being Gaussian, they are statistically independent.



(b)



(c)

FIGURE 5.21 Characteristics of RC-filtered white noise. (*a*) Low-pass RC filter. (*b*) Power spectral density of filter output $n(t)$. (*c*) Autocorrelation function of $n(t)$.

EXAMPLE 5.16 Autocorrelation of a Sinusoidal Signal Plus White Gaussian Noise

In this computer experiment, we study the statistical characterization of a random process $X(t)$ consisting of a sinusoidal signal $A\cos(2\pi f_c t + \Theta)$ and a white Gaussian noise process $W(t)$ of zero mean and power spectral density $N_0/2$. That is, we have

$$X(t) = A\cos(2\pi f_c t + \Theta) + W(t) \tag{5.128}$$

where Θ is a uniformly distributed random variable over the interval $(-\pi, \pi)$. Clearly, the two components of the process $X(t)$ are independent. The autocorrelation function of $X(t)$ is therefore the sum of the individual autocorrelation functions of the sinusoidal signal component and the noise component, as shown by

$$R_X(\tau) = \frac{A^2}{2}\cos(2\pi f_c \tau) + \frac{N_0}{2}\delta(\tau) \tag{5.129}$$

This equation shows that for $|\tau| > 0$, the autocorrelation function $R_X(\tau)$ has the same sinusoidal waveform as the signal component. We may generalize this result by stating that the presence of a periodic signal component corrupted by additive white noise can be detected by computing the autocorrelation function of the composite process $X(t)$.

The purpose of the example described here is to perform this computation using two different methods: (1) ensemble averaging and (2) time averaging. The trace of Figure 5.22*a* shows a sinusoidal signal of frequency $f_c = 0.002$ Hz and phase $\theta = -\pi/2$, truncated to a finite duration $T = 1000$ seconds; the amplitude A of the sinusoidal signal is set to $\sqrt{2}$ to give unity average power. The trace of Figure 5.22*b* shows a particular realization $x(t)$ of the random process $X(t)$ consisting of this sinusoidal signal and additive white Gaussian noise; the power spectral density of the noise for this realization is $(N_0/2) = 1$. The original sinusoid is barely recognizable in $x(t)$. The trace of Figure 5.22*c* shows the theoretical autocorrelation function of Eq. (5.129).

For ensemble-averaged computation of the autocorrelation function, we may proceed as follows:

- Compute the product $x(t + \tau)\, x(t)$ for some fixed time t and specified time shift τ, where $x(t)$ is a particular realization of the random process $X(t)$.
- Repeat the computation of the product $x(t + \tau)\, x(t)$ for M independent realizations (i.e., sample functions) of the random process $X(t)$.
- Compute the average of these computations over M.
- Repeat this sequence of computations for different values of τ.

The results of this computation are plotted in Figure 5.22*d* for $M = 500$ realizations. The picture portrayed here is in perfect agreement with the theory defined by Eq. (5.129). The important point to note here is that the ensemble-averaging process yields an estimate of the autocorrelation function $R_X(\tau)$ of the random process $X(t)$. Moreover, the presence of the sinusoidal signal is clearly visible in the plot of $R_X(\tau)$ versus τ.

For the time-averaged estimation of the autocorrelation function of the process $X(t)$, we invoke ergodicity and use the formula

$$R_X(\tau) = \lim_{T \to \infty} R_x(t, T) \tag{5.130}$$

where $R_X(\tau,T)$ is the time-average autocorrelation function:

$$R_x(\tau, T) = \frac{1}{2T} \int_{-T}^{T} x(t + \tau) x(t)\, dt \tag{5.131}$$

applied to a single sample function $x(t)$. Figure 5.22*e* presents the estimation of $R_X(\tau)$ using the time-averaging approach; it is also in close agreement of Figure 5.22*c*.

The fact that the ensemble-averaging and time-averaging approaches yield similar results for the autocorrelation function $R_X(\tau)$ signifies the fact that the random process $X(t)$ described in this example is indeed ergodic.

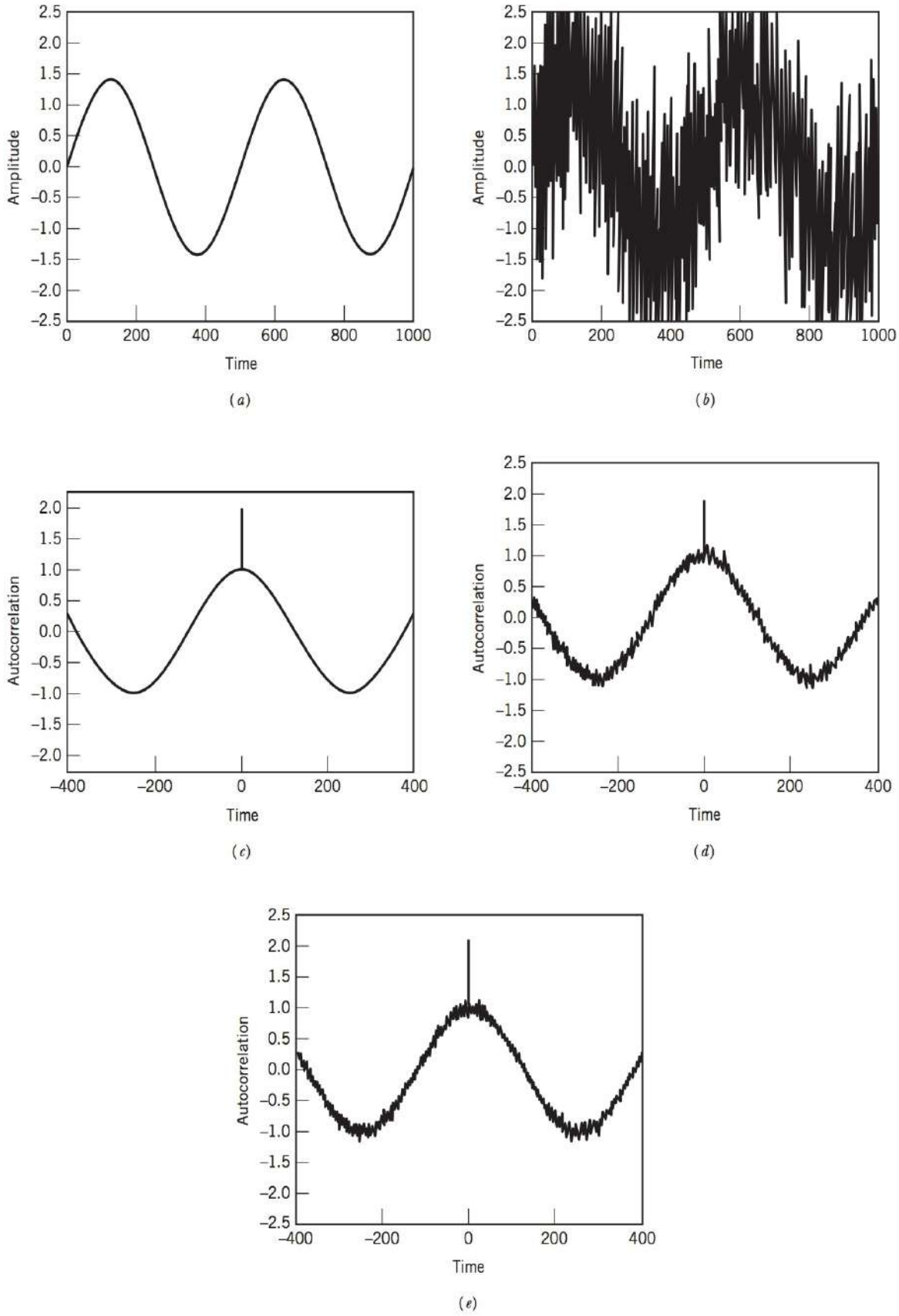


FIGURE 5.22 (a) The original truncated sinusoidal signal $A\cos(2\pi f_c t+\theta)$. (b) The noisy version $x(t)$ of the sinusoidal signal. (c) The theoretical autocorrelation of $X(t)$. (d) The estimated autocorrelation $R_X(\tau)$ using ensemble averaging. (e) The estimated autocorrelation $R_X(\tau)$ using time averaging.

NOISE EQUIVALENT BANDWIDTH

In Example 5.14 we observe that when a source of white noise of zero mean and power spectral density $N_0/2$ is connected across the input of an ideal low-pass filter of bandwidth B and passband amplitude response of one, the average output noise power [or equivalently $R_N(0)$] is equal to N_0B. In Example 5.15 we observe that when such a noise source is connected to the input of the simple RC low pass filter of Figure 5.21*a*, the corresponding value of the average output noise power is equal to $N_0/(4RC)$. For this filter, the half-power or 3-dB bandwidth is equal to $1/(2\pi RC)$. Here again we find that the average output noise power of the filter is proportional to the bandwidth.

We may generalize this statement to include all kinds of low-pass filters by defining a noise equivalent bandwidth as follows. Suppose that we have a source of white noise of zero mean and power spectral density $N_0/2$ connected to the input of an arbitrary low-pass filter of transfer function $H(f)$. The resulting average output noise power is therefore

$$\begin{aligned} N_{\text{out}} &= \frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2\, df \\ &= N_0 \int_0^{\infty} |H(f)|^2\, df \end{aligned} \tag{5.132}$$

where, in the last line, we have made use of the fact that the amplitude response $|H(f)|$ is an even function of frequency.

Consider next the same source of white noise connected to the input of an *ideal* low-pass filter of zero-frequency response $H(0)$ and bandwidth B. In this case, the average output noise power is

$$N_{\text{out}} = N_0 B H^2(0) \tag{5.133}$$

Therefore, equating this average output noise power to that in Eq. (5.132), we may define the *noise equivalent bandwidth* as

$$B = \frac{\displaystyle\int_0^{\infty} |H(f)|^2\, df}{H^2(0)} \tag{5.134}$$

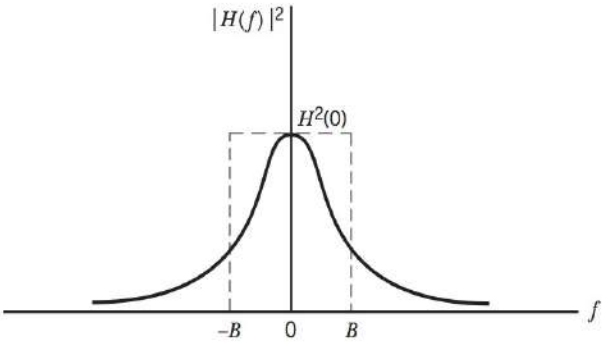


FIGURE 5.23 Illustrating the definition of noise-equivalent bandwidth.

Thus, the procedure for calculating the noise equivalent bandwidth consists of replacing the arbitrary low-pass filter of transfer function $H(f)$ by an equivalent ideal low-pass filter of zero-frequency response $H(0)$ and bandwidth B, as illustrated in Figure 5.23. In a similar way, we may define a noise equivalent bandwidth for band-pass filters.

5.11 NARROWBAND NOISE

The receiver of a communication system usually includes some provision for *preprocessing* the received signal. The preprocessing may take the form of a narrowband filter whose bandwidth is just large enough to pass the modulated component of the received signal essentially undistorted but not so large as to admit excessive noise through the

receiver. The noise process appearing at the output of such a filter is called *narrowband noise*. With the spectral components of narrowband noise concentrated about some midband frequency $\pm f_c$ as in Figure 5.24*a*, we find that a sample function $n(t)$ of such a process appears somewhat similar to a sine wave of frequency f_c, which undulates slowly in both amplitude and phase, as illustrated in Figure 5.24*b*.

To analyze the effects of narrowband noise on the performance of a communication system, we need a mathematical representation of it. Depending on the application of interest, there are two specific representations of narrowband noise:

1. The narrowband noise is defined in terms of a pair of components called the *in-phase* and *quadrature* components.
2. The narrowband noise is defined in terms of two other components called the *envelope* and *phase*.

These two representations are described in what follows. For now it suffices to say that given the in-phase and quadrature components, we may determine the envelope and phase components, and vice versa. Moreover, in their own individual ways, the two representations are not only basic to the noise analysis of communication systems but also to the characterization of narrowband noise itself.

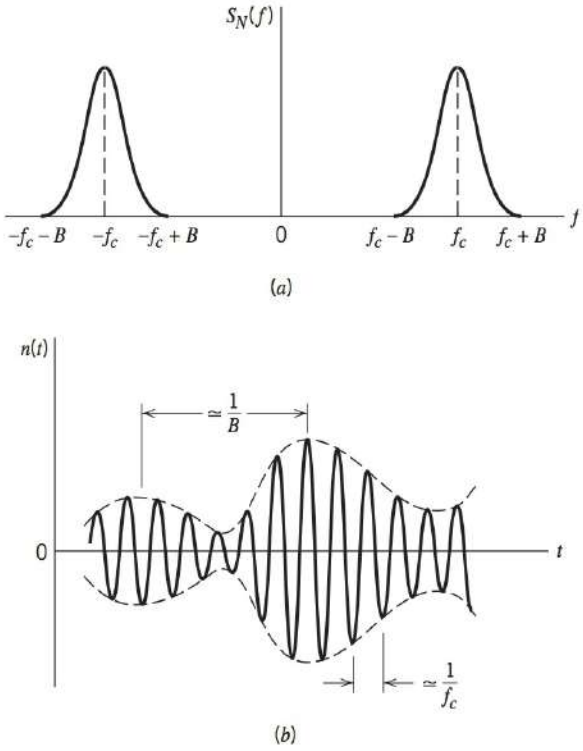


FIGURE 5.24 (*a*) Power spectral density of narrow-band noise. (*b*) Sample function of narrow-band noise.

REPRESENTATION OF NARROWBAND NOISE IN TERMS OF IN-PHASE AND QUADRATURE COMPONENTS

Consider a narrowband noise $n(t)$ of bandwidth $2B$ centered on frequency f_c, as illustrated in Figure 5.24. In light of the theory of band-pass signals and systems, we may represent $n(t)$ in the canonical (standard) form:

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \tag{5.135}$$

where $n_I(t)$ is called the *in-phase component* of $n(t)$, and $n_Q(t)$ is called the *quadrature component* of $n(t)$. Both $n_I(t)$ and $n_Q(t)$ are low-pass signals. Except for the midband frequency f_c, these two components are fully representative of the narrowband noise $n(t)$.

Given the narrowband noise $n(t)$, we may extract its in-phase and quadrature components using the scheme shown in Figure 5.25*a*. It is assumed that the two low-pass filters used in this scheme are ideal, each having a bandwidth equal to B (i.e., one-half the bandwidth of the narrowband noise $n(t)$). The scheme of Figure 5.25*a* follows from the representation of Equation (5.135). We may, of course, use this equation directly to generate the narrowband noise $n(t)$, given its in-phase and quadrature components, as shown in Figure 5.25*b*. The schemes of Figures 5.25*a* and 5.25*b* may thus be viewed as narrowband noise *analyzer* and *synthesizer*, respectively.

The in-phase and quadrature components of a narrowband noise have important properties that are summarized here:

1. The in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ of narrowband noise $n(t)$ have zero mean.

2. If the narrowband noise $n(t)$ is Gaussian, then its in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ are jointly Gaussian.
3. If the narrowband noise $n(t)$ is stationary, then its in-phase component $n_I(t)$ and quadrature component n_Q(t) are jointly stationary.
4. Both the in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ have the same power spectral density, which is related to the power spectral density $S_N(f)$ of the narrowband noise $n(t)$ as

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c), & -B \le f \le B \\ 0, & \text{otherwise} \end{cases} \tag{5.136}$$

where it is assumed that $S_N(f)$ occupies the frequency interval $f_c - B \le |f| \le f_c + B$, and $f_c > B$.

5. The in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ have the same variance as the narrowband noise $n(t)$.
6. The cross-spectral density of the in-phase and quadrature components of narrowband noise $n(t)$ is purely imaginary, as shown by

$$\begin{aligned} S_{N_I N_Q}(f) &= -S_{N_Q N_I}(f) \\ &= \begin{cases} j[S_N(f + f_c) - S_N(f - f_c)], & -B \le f \le B \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{5.137}$$

7. If the narrowband noise $n(t)$ is Gaussian and its power spectral density $S_N(t)$ is symmetric about the mid-band frequency f_c, then the in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ are statistically independent.

For further discussions of these properties, the reader is referred to Problems 5.31 and 5.32.

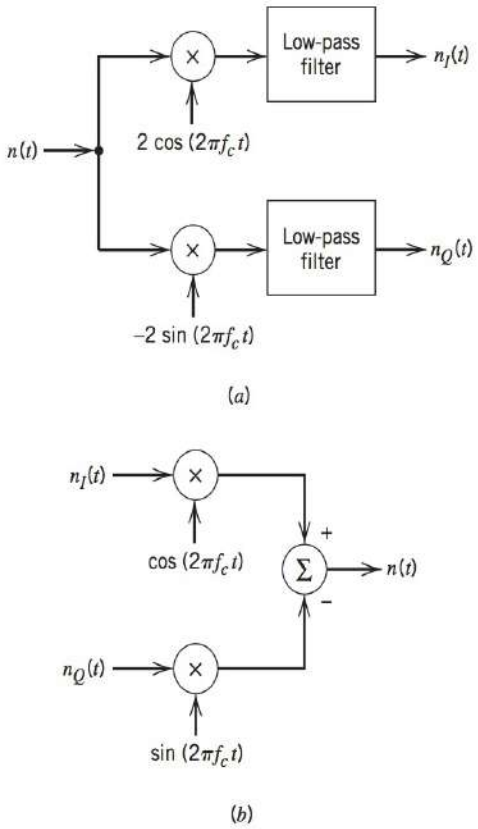
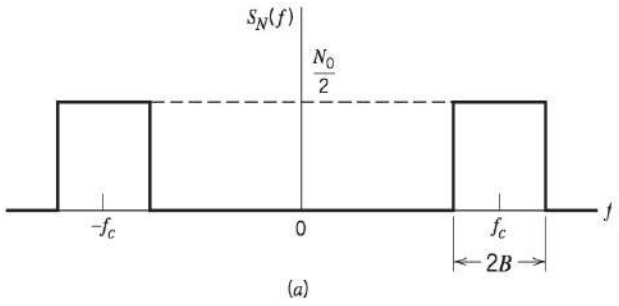


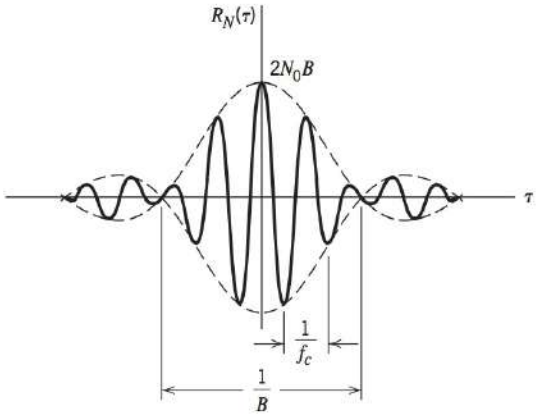
FIGURE 5.25 (*a*) Extraction of in-phase and quadrature components of a narrowband process. (*b*) Generation of a narrowband process from its in-phase and quadrature components.

EXAMPLE 5.17 Ideal Band-Pass Filtered White Noise

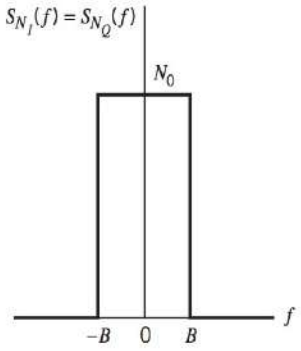
Consider a white Gaussian noise of zero mean and power spectral density $N_0/2$, which is passed through an ideal band-pass filter of passband magnitude response equal to one, mid-band frequency f_c, and bandwidth $2B$. The power spectral density characteristic of the filtered noise $n(t)$ will therefore be as shown in Figure 5.26*a*. The problem is to determine the autocorrelation functions of $n(t)$ and its in-phase and quadrature components.



(a)



(b)



(c)

FIGURE 5.26 Characteristics of ideal band-pass filtered white noise. (*a*) Power spectral density. (*b*) Autocorrelation function. (*c*) Power spectral density of in-phase and quadrature components.

The autocorrelation function of $n(t)$ is the inverse Fourier transform of the power spectral density characteristic shown in Figure 5.26*a*:

$$\begin{aligned} R_N(\tau) &= \int_{-f_c-B}^{-f_c+B} \frac{N_0}{2}\exp(j2\pi f\tau)df + \int_{f_c-B}^{f_c+B} \frac{N_0}{2}\exp(j2\pi f\tau)\,df \\ &= N_0B\,\text{sinc}(2B\tau)[\exp(-j2\pi f_c\tau) + \exp(j2\pi f_c\tau)] \\ &= 2N_0B\,\text{sinc}(2B\tau)\cos(2\pi f_c\tau) \end{aligned} \tag{5.138}$$

which is plotted in Figure 5.26*b*.

The spectral density characteristic of Figure 5.26*a* is symmetric about $\pm f_c$. Therefore, we find that the corresponding spectral density characteristic of the in-phase noise component $n_I(t)$ or the quadrature noise component $n_Q(t)$ is as shown in Figure 5.26*c*. The autocorrelation function of $n_I(t)$ or $n_Q(t)$ is therefore (see Example 5.14):

$$R_{N_I}(\tau) = R_{N_Q}(\tau) = 2N_0B\,\text{sinc}(2B\tau) \tag{5.139}$$

REPRESENTATION OF NARROWBAND NOISE IN TERMS OF ENVELOPE AND PHASE COMPONENTS

Previously, we considered the representation of a narrowband noise $n(t)$ in terms of its in-phase and quadrature components. We may also represent the noise $n(t)$ in terms of its envelope and phase components as follows:

$$n(t) = r(t)\cos[2\pi f_c t + \psi(t)] \tag{5.140}$$

where

$$r(t) = [n_I^2(t) + n_Q^2(t)]^{1/2} \tag{5.141}$$

and

$$\psi(t) = \tan^{-1}\left[\frac{n_Q(t)}{n_I(t)}\right] \tag{5.142}$$

The function $r(t)$ is called the *envelope* of $n(t)$, and the function $\psi(t)$ is called the *phase* of $n(t)$.

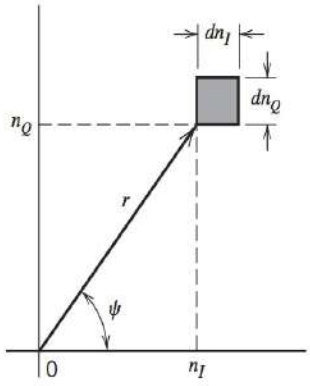
The envelope $r(t)$ and phase $\psi(t)$ are both sample functions of low-pass random processes. As illustrated in Figure 5.24*b*, the time interval between two successive peaks of the envelope $r(t)$ is approximately $1/B$, where $2B$ is the bandwidth of the narrowband noise $n(t)$.

The probability distributions of $r(t)$ and $\psi(t)$ may be obtained from those of $n_I(t)$ and $n_Q(t)$ as follows. Let N_I and N_Q denote the random variables obtained by observing (at some fixed time) the random processes represented by the sample functions $n(t)$ and $n_Q(t)$, respectively. We note that N_I and N_Q are independent Gaussian random variables of zero mean and variance σ^2, and so we may express their joint probability density function by

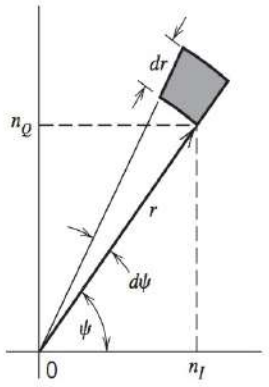
$$f_{N_I,N_Q}(n_I, n_Q) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{n_I^2 + n_Q^2}{2\sigma^2}\right) \tag{5.143}$$

Accordingly, the probability of the joint event that N_I lies between n_I and $n_I + dn_I$ and that N_Q lies between n_Q and $n_Q + dn_Q$ (i.e., the pair of random variables N_I and N_Q lies jointly inside the shaded area of Figure 5.27*a*) is given by

$$f_{N_I,N_Q}(n_I, n_Q)\, dn_I\, dn_Q = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{n_I^2 + n_Q^2}{2\sigma^2}\right) dn_I\, dn_Q \tag{5.144}$$



(*a*)



(*b*)

FIGURE 5.27 Illustrating the coordinate system for representation of narrowband noise: (*a*) in terms of in-phase and quadrature components, and (*b*) in terms of envelope and phase.

Define the transformation (see Figure 5.27*a*)

$$n_I = r\cos\psi \tag{5.145}$$

$$n_Q = r\sin\psi \tag{5.146}$$

In a limiting sense, we may equate the two incremental areas shown shaded in Figures 5.27*a* and 5.27*b* and thus write

$$dn_I\, dn_Q = r\, dr\, d\psi \tag{5.147}$$

Now, let R and Ψ denote the random variables obtained by observing (at some time t) the random processes represented by the envelope $r(t)$ and phase $\psi(t)$, respectively. Then, substituting Equations (5.145)–(5.147) into (5.144), we find that the probability of the random variables R and Ψ lying jointly inside the shaded area of Figure 5.27*b* is equal to

$$\frac{r}{2\pi\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right) dr\, d\psi$$

That is, the joint probability density function of R and Ψ is

$$f_{R,\Psi}(r, \psi) = \frac{r}{2\pi\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{5.148}$$

This probability density function is independent of the angle ψ, which means that the random variables R and Ψ are statistically independent. We may thus express $f_{R,\Psi}(r, \psi)$ as the product of $f_R(r)$ and $f_\Psi(\psi)$. In particular, the random variable Ψ representing phase is *uniformly distributed* inside the range 0 to 2π, as shown by

$$f_\Psi(\psi) = \begin{cases} \dfrac{1}{2\pi}, & 0 \le \psi \le 2\pi \\ 0, & \text{elsewhere} \end{cases} \tag{5.149}$$

This leaves the probability density function of the random variable R as

$$f_R(r) = \begin{cases} \dfrac{r}{\sigma^2}\exp\left(-\dfrac{r^2}{2\sigma^2}\right), & r \geq 0 \\ 0, & \text{elsewhere} \end{cases} \tag{5.150}$$

where σ^2 is the variance of the original narrowband noise $n(t)$. A random variable having the probability density function of Eq. (5.150) is said to be *Rayleigh distributed.*

For convenience of graphical presentation, let

$$v = \frac{r}{\sigma} \tag{5.151}$$

$$f_V(v) = \sigma f_R(r) \tag{5.152}$$

Then we may rewrite the Rayleigh distribution of Eq. (5.150) in the *normalized* form

$$f_V(v) = \begin{cases} v\exp\left(-\dfrac{v^2}{2}\right), & v \geq 0 \\ 0, & \text{elsewhere} \end{cases} \tag{5.153}$$

Eq. (5.153) is plotted in Figure 5.28. The peak value of the distribution $f_V(v)$ occurs at $v = 1$ and is equal to 0.607. Note also that, unlike the Gaussian distribution, the Rayleigh distribution is zero for negative values of v. This is because the envelope $r(t)$ can assume only nonnegative values.

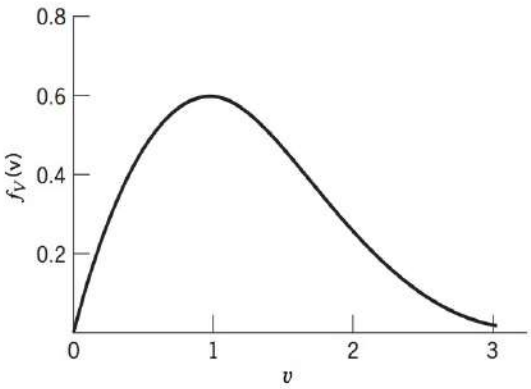


FIGURE 5.28 Normalized Rayleigh distribution.

EXAMPLE 5.18 Sinusoidal Signal Plus Narrowband Noise

Suppose that we add the sinusoidal signal $A\cos(2\pi f_c t)$ to narrowband noise $n(t)$, where A and f_c are both constants. We assume that the frequency of the sinusoidal wave is the same as the nominal carrier frequency of the noise. A sample function of the sinusoidal wave plus noise is then expressed by

$$x(t) = A\cos(2\pi f_c t) + n(t) \tag{5.154}$$

Representing the narrowband noise $n(t)$ in terms of its in-phase and quadrature components around the carrier frequency f_c, we may write

$$x(t) = n_I'(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \tag{5.155}$$

where

$$n_I'(t) = A + n_I(t) \tag{5.156}$$

We assume that $n(t)$ is Gaussian with zero mean and variance σ^2. Accordingly, we may state the following:

1. Both $n_I'(t)$ and $n_Q(t)$ are Gaussian and statistically independent.
2. The mean of $n_I'(t)$ is A and that of $n_Q(t)$ is zero.
3. The variance of both $n_I'(t)$ and $n_Q(t)$ is σ^2.

We may therefore express the joint probability density function of the random variables N_I' and N_Q, corresponding to $n_I'(t)$ and $n_Q(t)$, as follows:

$$f_{N_I',N_Q}(n_I',n_Q)=\frac{1}{2\pi\sigma^2}\exp\left[-\frac{(n_I'-A)^2+n_Q^2}{2\sigma^2}\right] \tag{5.157}$$

Let $r(t)$ denote the envelope of $x(t)$ and $\psi(t)$ denote its phase. From Equation (5.155), we thus find that

$$r(t)=\{[n_I'(t)]^2+n_Q^2(t)\}^{1/2} \tag{5.158}$$

and

$$\psi(t)=\tan^{-1}\left[\frac{n_Q(t)}{n_I'(t)}\right] \tag{5.159}$$

Following a procedure similar to that described for the derivation of the Rayleigh distribution, we find that the joint probability density function of the random variables R and Ψ, corresponding to $r(t)$ and $\psi(t)$ for some fixed time t, is given by

$$f_{R,\Psi}(r,\psi)=\frac{r}{2\pi\sigma^2}\exp\left(-\frac{r^2+A^2-2Ar\cos\psi}{2\sigma^2}\right) \tag{5.160}$$

We see that in this case, however, we cannot express the joint probability density function $f_{R,\Psi}(r,\psi)$ as a product $f_R(r)f_\Psi(\psi)$. This is because we now have a term involving the values of both random variables multiplied together as $r\cos\psi$. Hence, R and Ψ are dependent random variables for nonzero values of the amplitude A of the sinusoidal wave component.

Unfortunately, the presence of the sinusoidal component in the input signal $x(t)$ complicates the mathematical step that takes us from the joint probability density function $f_{R,\Psi}(r,\psi)$ to the associated marginal distributions: $f_R(r)$ and $f_\Psi(\psi)$. To simplify matters, we use intuitive reasoning to get a feel for the limiting forms of the marginal $f_R(r)$, depending on the value assigned to the amplitude A of the sinusoidal component:

1. When A is small compared to the noise envelope $r(t)$ for all t, that is, the "signal-to-noise ratio" of $x(t)$ is low, then Eq. (5.160) leads us to make the approximation

$$f_R(r)\approx\frac{r}{2\pi\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right),\qquad r(t)\ll A$$

2. When A is large compared to the noise envelope $r(t)$ for all t, that is, the "signal-to-noise ratio" is high, then we may neglect the composite term $2Ar\cos\psi$ compared to the sum (r^2+A^2) in the exponent of Eq. (5.160). This equation leads us to make the approximation.

$$f_R(r)=\frac{r}{2\pi\sigma^2}\exp\left(-\frac{r^2+A^2}{2\sigma^2}\right),\qquad r(t)\gg A$$

which may be viewed as "approximately Gaussian" in the vicinity of $r=A$.

Figure 5.29 plots the actual marginal distribution $f_R(r)$, versus r for varying sinusoidal amplitude A, where we have introduced the following definitions:

$$v = \frac{r}{\sigma} \tag{5.161}$$

$$a = \frac{A}{\sigma} \tag{5.162}$$

$$f_V(v) = \sigma f_R(r) \tag{5.163}$$

The normalized distribution of Figure 5.29 is called the *Rician distribution*. This figure clearly shows the evolution of the Rician distribution from a Rayleigh distribution (for small A) to an approximate Gaussian distribution (for large A).

Derivation of the Rician distribution, $f_R(r)$, requires knowledge of the modified Bessel function, which is discussed in the Appendix. For details of the derivation of $f_R(r)$, the reader is referred to Problem 5.34.

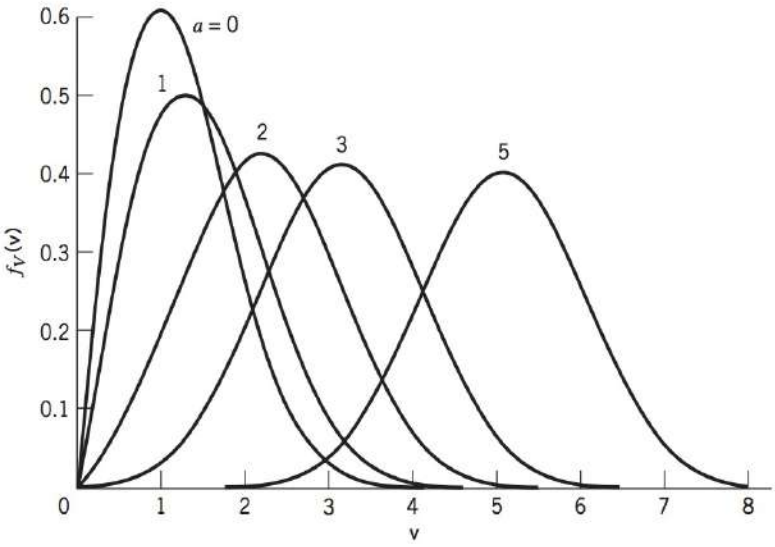


FIGURE 5.29 Normalized Rician distribution.

PHYSICAL RELEVANCE OF THE RAYLEIGH AND RICIAN DISTRIBUTIONS

In physical terms the Rayleigh distribution, plotted in Figure 5.28, is closely realized in an environment that consists of a large number of *scatterers* randomly distributed in the environment. The environment could be excited by a high-frequency (i.e., short-wave) signal generated by a transmitter, and the composite sum of the reflection from the scatterers constitutes the input of a receiver located at some distance away from the transmitter.

By the same token, the Rician distribution, plotted in Figure 5.29, is closely realized in the environment just described plus a path that directly connects the receiver to the transmitter. What, in effect, we are saying is that if we have an environment consisting of a large number of scatterers that are randomly located in space, then the underlying probabilistic distribution of the environment is closely realized by Rayleigh distribution. If, on the other hand, the environment of randomly located scatterers also includes a *direct path* from the transmitter to the receiver, then the underlying distribution of this new environment is closely realized by the Rician distribution.

5.12 THEME EXAMPLE—STOCHASTIC MODEL OF A MOBILE RADIO CHANNEL

One situation where probability theory and random processes play an important role in communications is in the analysis of mobile radio performance. We use the term "mobile radio" to encompass indoor and outdoor forms of wireless communications where a radio transmitter or receiver is capable of being moved, regardless of whether it actually moves or not. Due to the complex and variable nature of the mobile radio channel, it is not feasible to use a deterministic approach for its characterization. Rather, it is necessary to resort to the use of measurements and statistical analysis.[10]

The major propagation problems encountered in the use of cellular radio in built-up areas are due to the fact that the antenna of a mobile unit may lie well below the surrounding

buildings. Simply put, there is no "line-of-sight" path to the base station. Instead, radio propagation takes place mainly by way of scattering from the surfaces of the surrounding buildings and by diffraction over and/or around them, as illustrated in Figure 5.30. The important point to note from Figure 5.30 is that energy reaches the receiving antenna via more than one path. Accordingly, we speak of a *multipath phenomenon* in that the various incoming radio waves reach their destination from different directions and with different time delays.

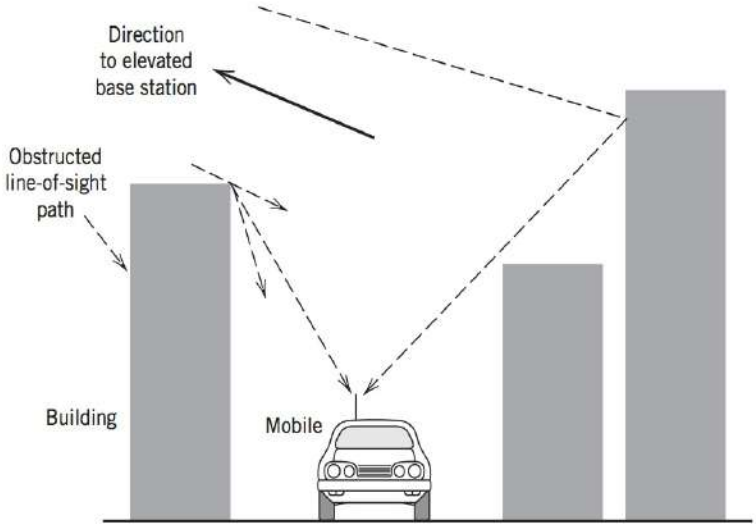


FIGURE 5.30 Illustrating the mechanism of radio propagation in urban areas. (From Parsons, 1992.)

To understand the nature of the multipath phenomenon, consider first a "static" multipath environment involving a stationary receiver and a transmitted signal that consists of a narrow-band signal (e.g., unmodulated sinusoidal carrier). Let it be assumed that two attenuated versions of the transmitted signal arrive sequentially at the receiver. The effect of the differential time delay is to introduce a relative phase shift between the two components of the received signal. We may then identify one of two extreme cases that can arise:

- The relative phase shift is zero, in which case the two components add constructively, as illustrated in Figure 5.31*a*.
- The relative phase shift is 180 degrees, in which case the two components add destructively, as illustrated in Figure 5.31*b*.

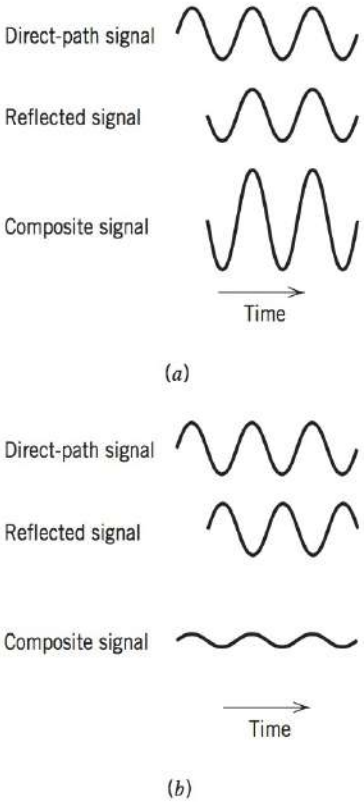


FIGURE 5.31 (*a*) Constructive and (*b*) destructive forms of the multipart phenomenon for sinusoidal signals.

Between these two extreme cases are a variety of situations where we may obtain partial constructive or destructive interference. Note that the relative phase shift of the two signals will vary with position since the relative time delay also varies with position.

The net result is that the envelope of the received signal varies with position in a complicated fashion, as shown by the experimental record of received signal envelope in an urban area that is presented in Figure 5.32. This figure clearly displays the fading nature of the received signal. The received signal envelope in Figure 5.32 is measured in dBm. The unit dBm is defined as $10 \log_{10}(P/P_0)$, with P denoting the power being measured and $P_0 = 1$ milliwatt. In the case of Figure 5.32, P is the instantaneous power in the received signal.

ENVELOPE DISTRIBUTION FOR FADED NARROWBAND SIGNALS

In the general case typical of Figure 5.32, there are N *versions* of the transmitted signal $s(t)$ arriving at the receiver, where at most one of them can be a direct path. The attenuation and phase of the received signals is usually randomly distributed. In this case, the received signal may be modeled, in the absence of noise, as

$$r(t) = \sum_{n=1}^{N} \mathrm{Re}[A_n \tilde{s}(t) \exp(j2\pi f_c t + \theta_n)] \tag{5.164}$$

where $\tilde{s}(t)$ is the complex envelope of the transmitted signal; A_n and θ_n are the attenuation and phase rotation of the nth signal path. The N different versions of the signal will travel along different paths from the transmitter to the receiver and thus, in general, will have different path lengths. The path length difference translates into

a relative time delay Δt_n, for $n > 1$, relative to the shortest path (assumed to correspond to $n = 1$). In practice, this relative delay is usually on the order of microseconds or less. The effect of the time delay on the phase is $\Delta\phi_n = f_c \Delta t_n$ and this phase difference is assumed to be included in the random component θ_n. As for its effect on the signal, we assume the signal is sufficiently narrowband such that $s(t) \approx s(t + \Delta t_n)$ for the relative time delays expected.

With these assumptions, the complex envelope of the received signal $r(t)$ may then be represented as

$$\tilde{r}(t) = \sum_{n=1}^{N} A_n \exp(j\theta_n)\tilde{s}(t) \tag{5.165}$$

The right-hand side of Eq. (5.165) may be expressed as

$$\tilde{r}(t) = \tilde{s}(t) \sum_{n=1}^{N} [a_n + jb_n] \tag{5.166}$$

where $a_n = A_n \cos(\theta_n)$ and $b_n = A_n \sin(\theta_n)$. It is reasonable to assume that a_n and b_n have approximately the same distribution, so that for large N, their sums approach Gaussian random variables due to the *central limit theorem*. Thus, we make the approximation

$$\tilde{r}(t) \approx [X + jY]\tilde{s}(t) \tag{5.167}$$

where X and Y are independent, identically-distributed, Gaussian random variables. Of practical interest is distribution of the signal amplitude. From Section 5.12, the amplitude $Z = \sqrt{X^2 + Y^2}$ where X and Y are independent, zero-mean Gaussian random variables has a *Rayleigh distribution*. That is, the amplitude due to fading at any instant in time is a random variable with a Rayleigh distribution.

In Figure 5.33, we plot the Rayleigh distribution on a log-scale for comparison with the experimental measurements of signal power shown in Figure 5.32. From inspection of Figure 5.32, the median power level appears to be about −73 dBm. As a rough approximation, the signal drops 10 dB below this level, that is, below −83 dBm, about 10% of the time. From the theoretical curve we find that the probability that a Rayleigh faded signal is 10 dB or more below the rms value is 10 percent, thus agreeing with the measurements. The signal envelope of Figure 5.32 drops 20 dB below the median level much less often and, from the theoretical curve, this situation should only occur 1 percent of the time. This theoretical result agrees qualitatively with the observations.

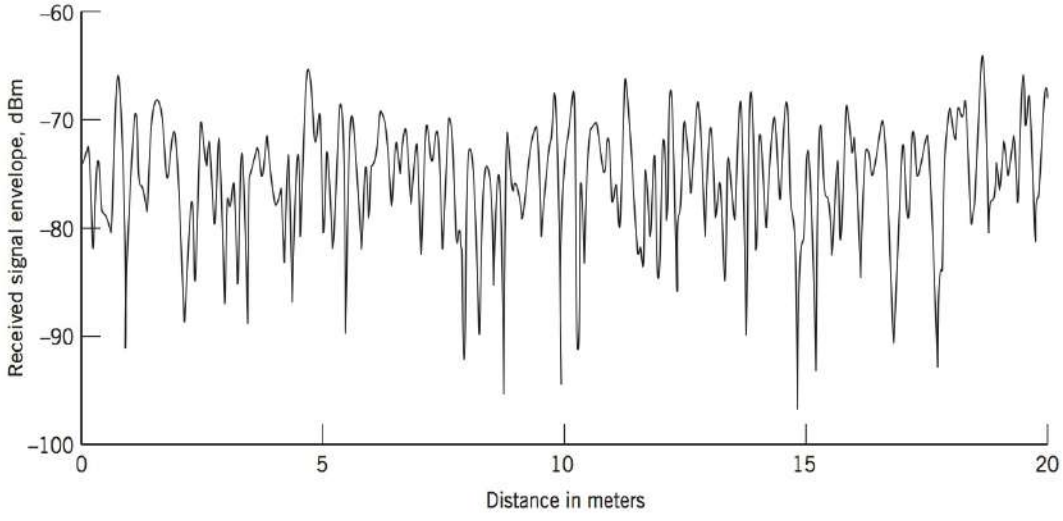


FIGURE 5.32 Experimental record of received signal envelope in an urban area. (From Parsons, 1992.)

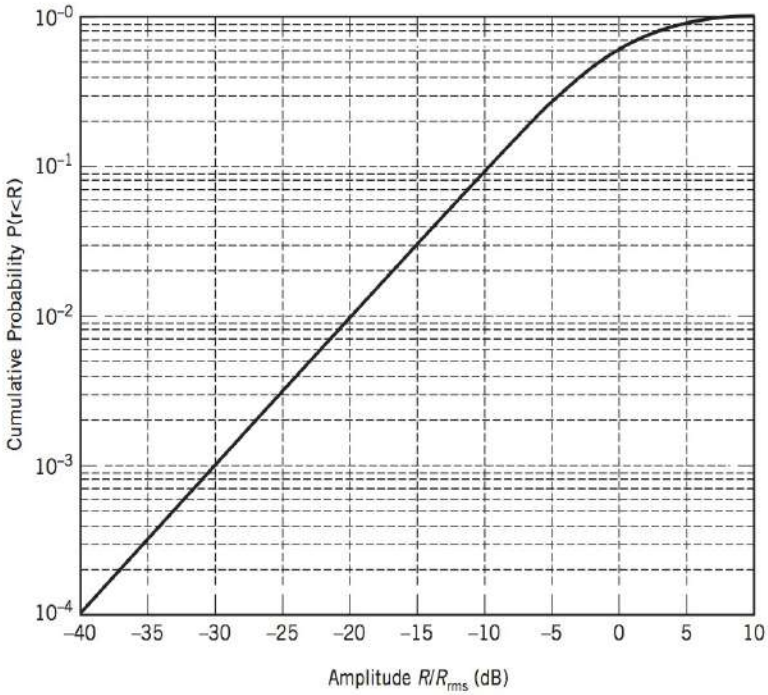


FIGURE 5.33 Rayleigh fading distribution function.

AUTOCORRELATION OF ENVELOPE OF FADED NARROWBAND SIGNALS

If the radio receiver is moving, then the fading varies with time as well, and the fading can be considered a random process. To characterize the fading process, we must adjust the model to account for the movement of the receiver. When a receiver moves relative to a signal source, there is a change in the frequency of the received signal proportional to the receiver's velocity in the direction of the source. To be specific, consider the situation illustrated in Figure 5.34, where the receiver is assumed to be moving along the line AA' with a constant velocity, v. It is also assumed that the received signal is due to a radio wave from a scatterer labeled S. Let Δt denote the time taken for the receiver to move from point A to A'. Using the notation described in Figure 5.34, the incremental change in the path length of the radio wave is deduced to be

$$\begin{aligned} \Delta l &= d \cos \alpha \\ &= v \Delta t \cos \alpha \end{aligned} \tag{5.168}$$

where α is the spatial angle between the incoming radio wave and the direction of motion of the receiver. Correspondingly, the change in the phase angle of the received signal at point A' with respect to that at point A is given by

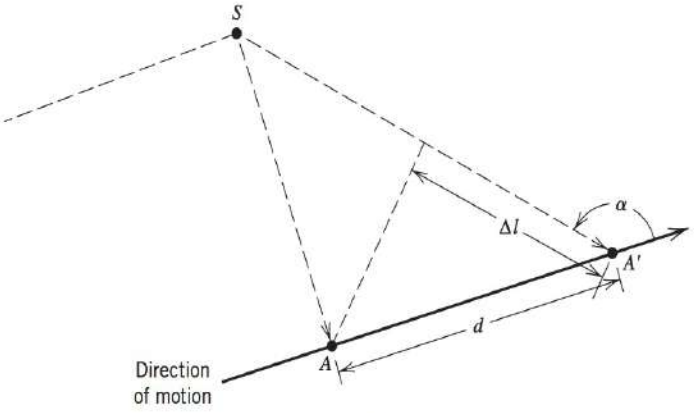


FIGURE 5.34 Illustrating Doppler shift. (From Parsons, 1992.)

$$\begin{aligned} \Delta \phi &= -\frac{2\pi}{\lambda} \Delta l \\ &= -\frac{2\pi v \Delta t}{\lambda} \cos \alpha \end{aligned} \tag{5.169}$$

where λ is the radio wavelength. The apparent change in frequency, or the *Doppler-shift*, is therefore

$$\begin{aligned} \Delta f &= -\frac{1}{2\pi} \frac{\Delta \phi}{\Delta t} \\ &= \frac{v}{\lambda} \cos \alpha \end{aligned} \tag{5.170}$$

The Doppler shift Δf is positive (resulting in an increase in frequency) when the radio waves arrive from ahead of the mobile unit, and it is negative when the radio waves arrive from behind the mobile unit.

With multiple reflected paths, each will have a slightly different frequency based on its angle of arrival at the receiver compared to the direction of the receiver motion. Consequently, the model for the *complex envelope* of the received signal for a moving receiver is

$$\tilde{r}(t) = \sum_{n=1}^{N} A_n \exp[j(2\pi f_n t + \theta_n)]\tilde{s}(t) \tag{5.171}$$

where f_n is the Doppler frequency of the nth scattered ray. We may characterize the behavior of the random process by computing the autocorrelation of the complex envelope.

$$\begin{aligned} R_r(\tau) &= \mathbf{E}[\tilde{r}(t)\tilde{r}^*(t+\tau)] \\ &= \mathbf{E}\left[\left\{\sum_{n=1}^{N} A_n \exp(j(2\pi f_n t + \theta_n))\right\}\left\{\sum_{n=1}^{N} A_n \exp(-j(2\pi f_n(t+\tau) + \theta_n))\right\}\right] \\ &\quad \times \mathbf{E}[\tilde{s}(t)\tilde{s}^*(t+\tau)] \end{aligned} \tag{5.172}$$

In the second line of Eq. (5.176) we are able to separate the autocorrelation of the fading process from that of the signal due to their assumed independence. Focusing on the fading process $F = \Sigma_{n=1}^{N} A_n \exp[j(2\pi f_n t + \theta_n)]$, we find that

$$R_F(\tau) = \mathbf{E}\left[\sum_{n=1}^{N} A_n^2 \exp[-j(2\pi f_n \tau)]\right] \tag{5.173}$$

which results from simplifying the second last line of Eq. (5.172). Under suitable assumptions, we may evaluate the expectation in Eq. (5.173) (see Problem 5.33) to obtain

$$R_F(\tau) = P_0 J_0(2\pi f_D \tau) \tag{5.174}$$

where P_0 is the average received power, $J_0(\,)$ is the *zeroth-order Bessel function*, and f_D is the *maximum Doppler frequency* for the given receiver velocity. The maximum Doppler is obtained by setting $\alpha = 0$ in Eq. (5.170). In Figure 5.35 we plot the normalized version of the autocorrelation function of Eq. (5.174) versus the parameter $f_D\tau$. The function is symmetric in τ and shows that over short distances (small $f_D\tau$) the fading signal is strongly correlated.

Again, it is insightful to compare the measured result of Figure 5.32 with the theoretical result of Figure 5.35. From Eq. (5.172) the distance, d, traveled at velocity v is

$$d = v\tau = \lambda(f_D\tau)$$

If, for example, the carrier frequency for the measurement of Figure 5.32 is f_c = 900 MHz, then $\lambda = 0.33$ meters. Since the theoretical autocorrelation indicates a strong correlation for $|f_D\tau| < 0.25$, on the distance axis of Figure 5.32, we expect a strong correlation over the range $0.25\lambda \approx 0.08$ meters; and this is indeed the case. Over greater distances, there is very little correlation of the fading process.

In engineering applications, these statistical models are useful design tools in a number of areas. If, for instance, narrowband communication over a mobile channel is to be reliable 99 percent of the time, then these results indicate that the design must include 20 dB power margin, unless there are other methods of compensating the signal loss due to fading. One method of compensating the fading loss is through *forward error correction (FEC) coding* (discussed in Chapter 10) in combination with a device known as an *interleaver*, which pseudo-randomly distributes the bits before transmission (and an inverse interleaver is applied to the bits at the receiver). Many forms of FEC work best if adjacent bits are independently faded. Thus, the autocorrelation results for the fading envelope provide an important parameter for the design of the interleaver.

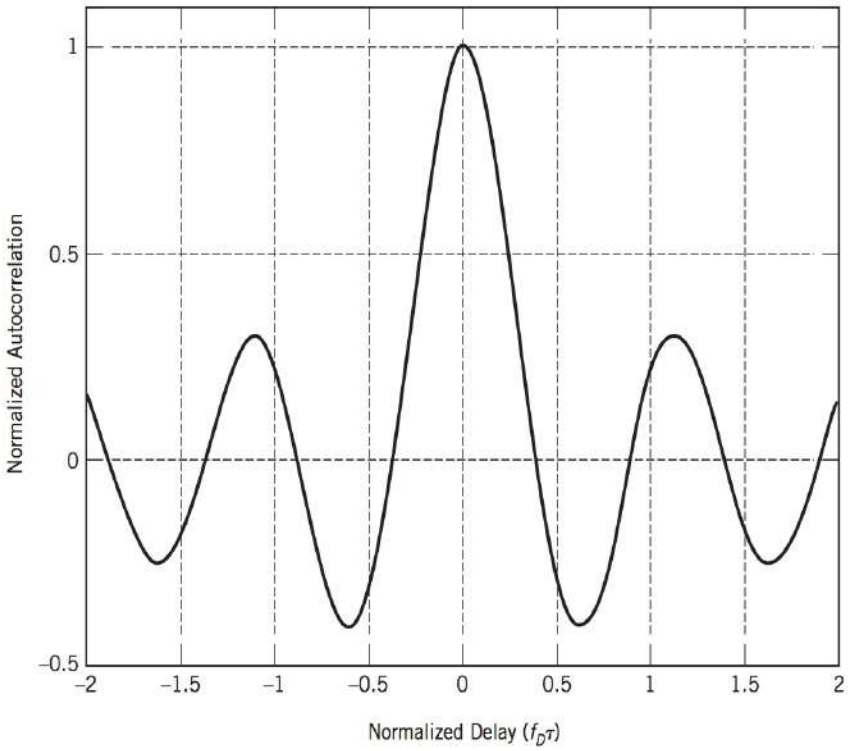


FIGURE 5.35 Autocorrelation of fading process.

In this example, we have only considered the fading characteristics of a narrowband signal where the relative path-length differences of the different multipath rays had negligible effect. This situation is often referred to as *frequency-flat* or simply *flat fading*, since the effects are uniform across all of the signal's frequencies. In the more general case, with wider bandwidth signals, the multipath channel must be modeled with an impulse response $h(t)$ in the static case and a time-varying impulse response $h(t,\tau)$ in the dynamic case, thus leading to more complicated fading characteristics.

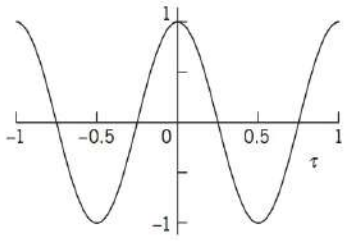
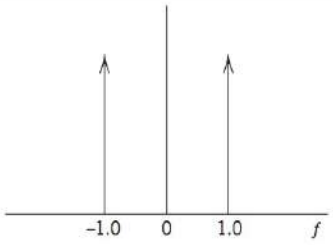
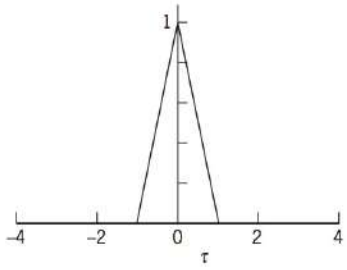
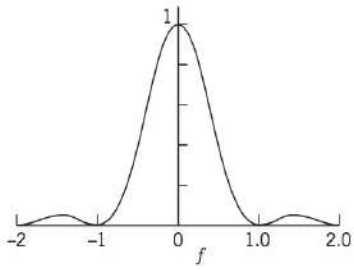
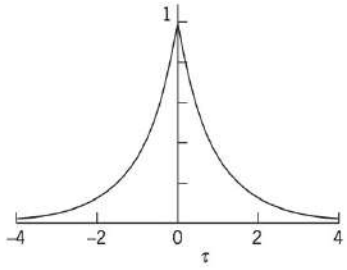
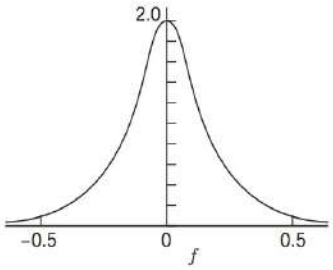
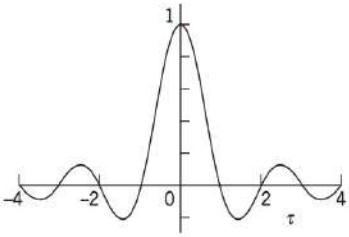
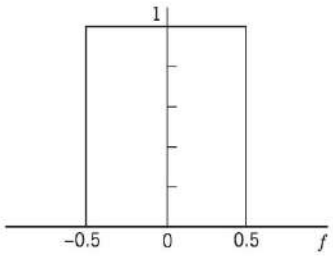
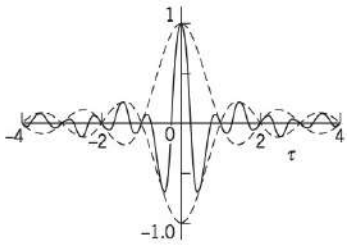
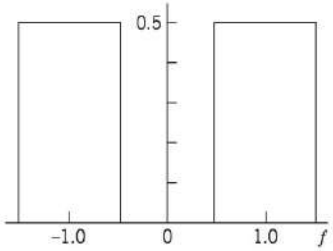
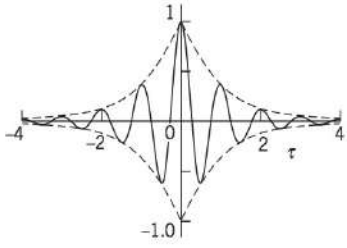
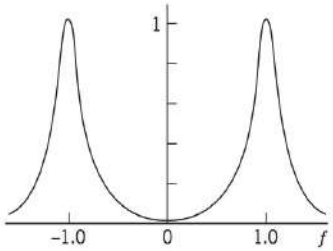
5.13 SUMMARY AND DISCUSSION

Much of the material presented in this chapter has dealt with the characterization of a particular class of random processes known to be wide-sense stationary and ergodic. The implication of wide-sense stationarity is that we may develop a partial description of a random process in terms of two ensemble-averaged parameters: (1) a mean that is independent of time, and (2) an autocorrelation function that depends only on the difference between the times at which two observations of the process are made.[11] Ergodicity enables us to use time averages as "estimates" of these parameters. The time averages are computed using a sample function (i.e., single realization) of the random process.

Another important parameter of a random process is the power spectral density. The autocorrelation function and the power spectral density constitute a Fourier-transform pair. The formulas that define the power spectral density in terms of the autocorrelation function and vice versa are known as the Einstein–Wiener–Khintchine relations.

In Table 5.1 we present a graphical summary of the autocorrelation functions and power spectral densities of some important random processes studied in the chapter. All the processes described in this table are assumed to have zero mean and unit variance. This table should give the reader a feeling for (1) the interplay between the autocorrelation function and power spectral density of a random process, and (2) the role of linear filtering in shaping the autocorrelation function or, equivalently, the power spectral density of a white noise process.

TABLE 5.1 Graphical summary of autocorrelation functions and power spectral densities of random processes of zero mean and unit variance

Type of Process, $X(t)$	Autocorrelation Function, $R_X(\tau)$	Power Spectral Density, $S_X(f)$
Sinusoidal process of unit frequency and random phase		
Random binary wave of unit symbol-duration		
RC low-pass filtered white noise		
Ideal low-pass filtered white noise		
Ideal band pass filtered white noise		
RLC-filtered white noise		

The latter part of the chapter dealt with noise processes that are Gaussian and narrow band, which is the kind of filtered noise encountered at the front end of an idealized form of communication receiver. Gaussianity means that the random variable obtained by observing the output of the filter at some fixed time has a Gaussian distribution. The narrowband nature of the noise means that it may be represented in terms of an in-phase and a quadrature component. These two components are both low-pass, Gaussian processes, each with zero mean and a variance equal to that of the original narrowband noise. Alternatively, a Gaussian narrow-band noise may be represented in terms of a Rayleigh-distributed envelope and a uniformly distributed phase. Each of these representations has its own specific area of application, as shown in subsequent chapters of the book.

The material presented in this chapter has been confined entirely to *real* random processes. It may be generalized for *complex* random processes. A commonly encountered complex random process is a complex Gaussian low-pass process, which arises in the equivalent representation of a Gaussian narrowband noise $n(t)$. From Section 5.11 we note that $n(t)$ is uniquely defined in terms of the in-phase component $n_I(t)$ and the quadrature component $n_Q(t)$. Equivalently, we may represent the narrowband noise $n(t)$ in terms of the complex envelope $\tilde{n}(t)$ defined as $n_I(t) + jn_Q(t)$.

NOTES AND REFERENCES

1. For introductory treatment of probability theory by itself, see Hamming (1991). For introductory treatment of probability and random processes with an engineering emphasis, see Leon-Garcia (1994), Helstrom (1990), and Papoulis (1984).
2. There is another important class of random processes commonly encountered in practice, the mean and autocorrelation function of which exhibit *periodicity*, as in

$$\mu_X(t_1 + T) = \mu_X(t_1)$$
$$R_X(t_1 + T, t_2 + T) = R_X(t_1, t_2)$$

   for all t_1 and t_2. A random process $X(t)$ satisfying this pair of conditions is said to be *cyclostationary* (in the wide sense). Modeling the process $X(t)$ as cyclostationary adds a new dimension, namely, period T to the partial description of the process. Examples of cyclostationary processes include a television signal obtained by raster-scanning a random video field, and a modulated process obtained by varying the amplitude, phase, or frequency of a sinusoidal carrier. For detailed discussion of cyclostationary processes, see Franks (1969), pp. 204–214, and the paper by Gardner and Franks (1975).
3. For a more detailed treatment of ergodicity, see Gray and Davisson (1986).
4. Traditionally, Eqs. (5.91) and (5.92) have been referred to in the literature as the Wiener–Khintchine relations in recognition of pioneering work done by Norbert Wiener and A. I. Khintchine. A discovery of a forgotten paper by Albert Einstein on time-series analysis (delivered at the Swiss Physical Society's February 1914 meeting in Basel) reveals that Einstein had discussed the autocorrelation function and its relationship to the spectral content of a time series many years before Wiener and Khintchine. An English translation of Einstein's paper is reproduced in the *IEEE ASSP Magazine*, vol. 4, October 1987. This particular issue also contains articles by W. A. Gardner and A. M. Yaglom, which elaborate on Einstein's original work.
5. For further details of power spectrum estimation, see Box and Jenkins (1976), Marple (1987), and Kay (1988).
6. For a detailed treatment of electrical noise, see Van der Ziel (1970) and the collection of papers edited by Gupta (1977).
7. An introductory treatment of shot noise is presented in Helstrom (1990).
8. Thermal noise was first studied experimentally by J. B. Johnson in 1928, and for this reason it is sometimes referred to as the "Johnson noise." Johnson's experiments were confirmed theoretically by Nyquist (1928).

9. The noisiness of a receiver may also be measured in terms of the so-called *noise figure*. The relationship between the noise figure and the equivalent noise temperature may be found in Haykin and Moher (2005).
10. Discussion of both analytical and statistical techniques of characterizing propagation may be found in Parsons (1992).
11. The statistical characterization of communication systems presented in this book is confined to the first two moments, mean and autocorrelation function (equivalently, autocovariance function) of the pertinent random process. However, when a random process is transmitted through a nonlinear system, valuable information is contained in higher-order moments of the resulting output process. The parameters used to characterize higher-order moments in the time domain are called *cumulants*; their multidimensional Fourier transforms are called *polyspectra*. For a discussion of higher-order cumulants and polyspectra and their estimation, see the papers by Brillinger (1965) and Nikias and Raghuveer (1987).

PROBLEMS

5.1

(a) Show that the characteristic function of a Gaussian random variable X of mean μ_X and variance σ_X^2 is

$$\phi_X(v) = \exp\left(jv\mu_X - \tfrac{1}{2}v^2\sigma_X^2\right)$$

(b) Using the result of part (a), show that the nth central moment of this Gaussian random variable is

$$E[(X-\mu_X)^n] = \begin{cases} 1\times 3\times 5\ldots(n-1)\sigma_X^n & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd} \end{cases}$$

5.2 A Gaussian-distributed random variable X of zero mean and variance σ_X^2 is transformed by a piecewise-linear rectifier characterized by the input–output relation (see Figure P5.1):

$$Y = \begin{cases} X, & X \geq 0 \\ 0, & X < 0. \end{cases}$$

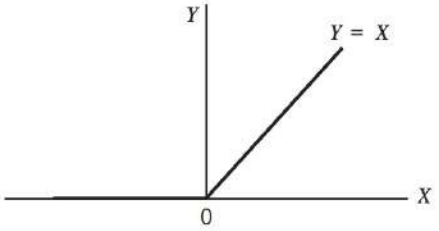


Figure P5.2

The probability density function of the new random variable Y is described by

$$f_Y(y) = \begin{cases} 0, & y < 0 \\ k\delta(y), & y = 0 \\ \dfrac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\dfrac{y^2}{2\sigma_X^2}\right), & y > 0 \end{cases}$$

(a) Explain the physical reasons for the functional form of this result.

(b) Determine the value of the constant k by which the delta function $\delta(y)$ is weighted.

5.3 A binary signal having the value of ± 1 is detected in the presence of additive white Gaussian noise of zero mean and variance σ^2. What is the probability density function of the signal observed at the input to the detector? Derive an expression for the probability that the observed signal is larger than a specified threshold α.

5.4 Consider a random process $X(t)$ defined by

$$X(t) = \sin(2\pi ft)$$

in which the frequency f is a random variable uniformly distributed over the interval $(0,W)$. Show that $X(t)$ is nonstationary. *Hint*: Examine specific sample functions of the random process $X(t)$ for the frequency $f = W/4$, $W/2$, and W, say.

5.5 For a complex random process $Z(t)$, define the autocorrelation function as

$$R_Z(\tau) = \mathbf{E}[Z^*(t)Z(t+\tau)]$$

where $*$ represents complex conjugation. Derive the properties of this complex autocorrelation corresponding to Eqs. (5.64), (5.65), and (5.67).

5.6 For the complex random process $Z(t) = Z_I(t) + jZ_Q(t)$ where $Z_I(t)$ and $Z_Q(t)$ are real-valued random processes given by

$$Z_I(t) = A\cos(2\pi f_1 t + \theta_1)$$

and

$$Z_Q(t) = A\cos(2\pi f_2 t + \theta_2)$$

where θ_1 and θ_2 are uniformly distributed over $[-\pi,\pi]$. What is the autocorrelation of $Z(t)$? Suppose $f_1 = f_2$? Suppose $\theta_1 = \theta_2 = \theta$?

5.7 Let X and Y be statistically independent Gaussian-distributed random variables, each with zero mean and unit variance. Define the Gaussian process

$$Z(t) = X\cos(2\pi t) + Y\sin(2\pi t)$$

(a) Determine the joint probability density function of the random variables $Z(t_1)$ and $Z(t_2)$ obtained by observing $Z(t)$ at times t_1 and t_2, respectively.

(b) Is the process $Z(t)$ stationary? Why?

5.8 Prove the following two properties of the autocorrelation function $R_X(\tau)$ of a random process $X(t)$:

(a) If $X(t)$ contains a dc component equal to A, then $R_X(\tau)$ will contain a constant component equal to A^2.

(b) If $X(t)$ contains a sinusoidal component, then $R_X(\tau)$ will also contain a sinusoidal component of the same frequency.

5.9 The square wave $x(t)$ of Figure P5.9 of constant amplitude A, period T_0, and delay t_d, represents the sample function of a random process $X(t)$. The delay is random, described by the probability density function

$$f_{T_d}(t_d) = \begin{cases} \dfrac{1}{T_0}, & -\dfrac{1}{2}T_0 \le t_d \le \dfrac{1}{2}T_0 \\ 0, & \text{otherwise} \end{cases}$$

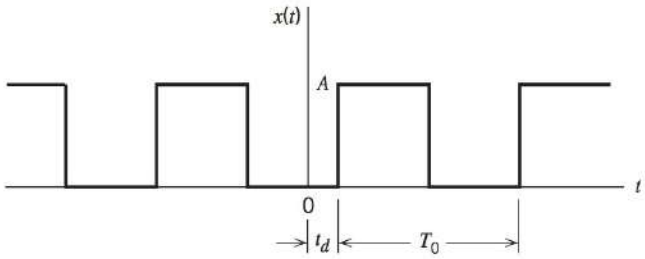


Figure P5.9

(a) Determine the probability density function of the random variable $X(t_k)$ obtained by observing the random process $X(t)$ at time t_k.

(b) Determine the mean and autocorrelation function of $X(t)$ using ensemble averaging.

(c) Determine the mean and autocorrelation function of $X(t)$ using time-averaging.

(d) Establish whether or not $X(t)$ is wide-sense stationary. In what sense is it ergodic?

5.10 A binary wave consists of a random sequence of symbols 1 and 0, similar to that described in Example 5.8, with one basic difference: symbol 1 is now represented by a pulse of amplitude A volts and symbol 0 is represented by zero volts. All other parameters are the same as before. Show that for this new random binary wave $X(t)$:

(a) The autocorrelation function is

$$R_X(\tau) = \begin{cases} \dfrac{A^2}{4} + \dfrac{A^2}{4}\left(1 - \dfrac{|\tau|}{T}\right), & |\tau| < T \\ \dfrac{A^2}{4}, & |\tau| \ge T \end{cases}$$

(b) The power spectral density is

$$S_X(f) = \frac{A^2}{4}\delta(f) + \frac{A^2T}{4}\,\text{sinc}^2(fT)$$

What is the percentage power contained in the dc component of the binary wave?

5.11 A random process $Y(t)$ consists of a dc component of $\sqrt{3/2}$ volts, a periodic component $g(t)$, and a random component $X(t)$. The autocorrelation function of $Y(t)$ is shown in Figure P5.11.

(a) What is the average power of the periodic component $g(t)$?

(b) What is the average power of the random component $X(t)$?

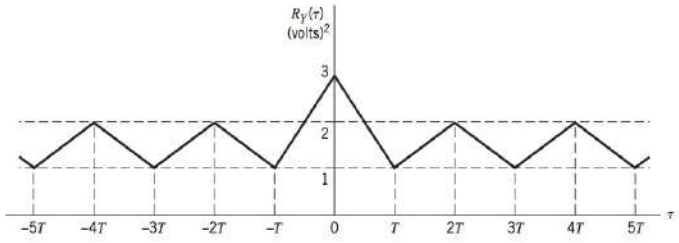


Figure P5.11

5.12 Consider a pair of wide-sense stationary random processes $X(t)$ and $Y(t)$. Show that the cross-correlations $R_{XY}(\tau)$ and $R_{YX}(\tau)$ of these processes have the following properties:

(a) $R_{XY}(\tau) = R_{YX}(-\tau)$

(b) $|R_{XY}(\tau)| \le \frac{1}{2}[R_X(0) + R_Y(0)]$

5.13 Consider two linear filters connected in cascade as in Figure P5.13. Let $X(t)$ be a wide-sense stationary process with autocorrelation function $R_X(\tau)$. The random process appearing at the first filter output is $V(t)$ and that at the second filter output is $Y(t)$.

(a) Find the autocorrelation function of $Y(t)$.

(b) Find the cross-correlation function $R_{VY}(\tau)$ of $V(t)$ and $Y(t)$.

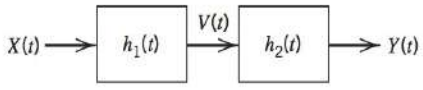


Figure P5.13

5.14 A wide-sense stationary random process $X(t)$ is applied to a linear time invariant filter of impulse response $h(t)$, producing an output $Y(t)$.

(a) Show that the cross-correlation function $R_{YX}(\tau)$ of the output $Y(t)$ and the input $X(t)$ is equal to the impulse response $h(\tau)$

convolved with the autocorrelation function $R_X(\tau)$ of the input, as shown by

$$R_{YX}(\tau) = \int_{-\infty}^{\infty} h(u)R_X(\tau - u)\,du$$

Show that the second cross-correlation function $R_{XY}(\tau)$ equals

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} h(-u)R_X(\tau - u)\,du$$

(b) Find the cross-spectral densities $S_{YX}(f)$ and $S_{XY}(f)$.

(c) Assuming that $X(t)$ is a white noise process with zero mean and power spectral density $N_0/2$, show that

$$R_{YX}(\tau) = \frac{N_0}{2}h(\tau)$$

Comment on the practical significance of this result.

5.15 The power spectral density of a random process $X(t)$ is shown in Figure P5.15.

(a) Determine and sketch the autocorrelation function $R_X(\tau)$ of $X(t)$.

(b) What is the dc power contained in $X(t)$?

(c) What is the ac power contained in $X(t)$?

(d) What sampling rates will give uncorrelated samples of $X(t)$? Are the samples statistically independent?

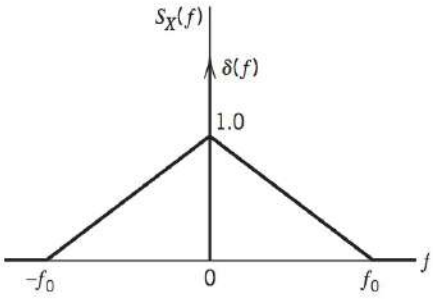


Figure P5.15

5.16 A pair of noise processes $n_1(t)$ and $n_2(t)$ are related by

$$n_2(t) = n_1(t)\cos(2\pi f_c t + \theta) - n_1(t)\sin(2\pi f_c t + \theta)$$

where f_c is a constant, and θ is the value of a random variable Θ whose probability density function is defined by

$$f_\Theta(\theta) = \begin{cases} \dfrac{1}{2\pi}, & 0 \le \theta \le 2\pi \\ 0, & \text{otherwise} \end{cases}$$

The noise process $n_1(t)$ is stationary and its power spectral density is as shown in Figure P5.16. Find and plot the corresponding power spectral density of $n_2(t)$.

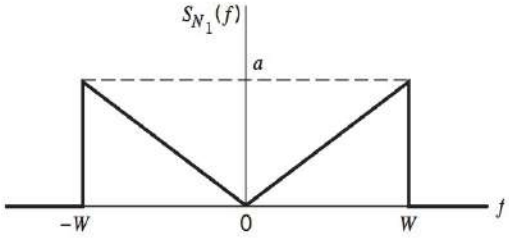


Figure P5.16

5.17 A *random telegraph signal* $X(t)$, characterized by the autocorrelation function

$$R_X(\tau) = \exp(-2v|\tau|)$$

where v is a constant, is applied to the low-pass RC filter of Figure P5.17. Determine the power spectral density and autocorrelation function of the random process at the filter output.

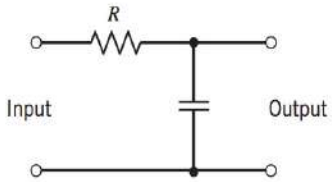


Figure P5.17

5.18 The output of an oscillator is described by

$$X(t) = A\cos(2\pi ft - \Theta),$$

where A is a constant, and f and Θ are independent random variables, The probability density function of Θ is defined by

$$f_\Theta(\theta) = \begin{cases} \dfrac{1}{2\pi}, & 0 \le \theta \le 2\pi \\ 0, & \text{otherwise} \end{cases}$$

Find the power spectral density of $X(t)$ in terms of the probability density function of the frequency f. What happens to this power spectral density when the frequency f assumes a constant value?

5.19 A stationary, Gaussian process $X(t)$ has zero mean and power spectral density $S_X(f)$. Determine the probability density function of a random variable obtained by observing the process $X(t)$ at some time t_k.

5.20 A Gaussian process $X(t)$ of zero mean and variance σ_X^2 is passed through a full-wave rectifier, which is described by the input–output relation of Figure P5.20. Show that the probability density function of the random variable $Y(t_k)$, obtained by

observing the random process $Y(t)$ at the rectifier output at time t_k, is as follows.

$$f_{Y(t_k)}(y) = \begin{cases} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_X} \exp\left(-\frac{y^2}{2\sigma_X^2}\right), & y \geq 0 \\ 0, & y < 0 \end{cases}$$

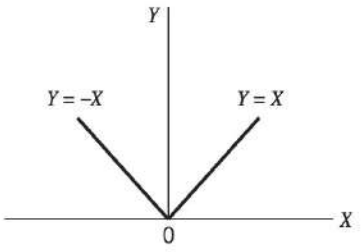


Figure P5.20

5.21 Let $X(t)$ be a zero mean, stationary, Gaussian process with autocorrelation function $R_X(\tau)$, This process is applied to a square-law device defined by the input-output relation

$$Y(t) = X^2(t)$$

where $Y(t)$ is the output.

(a) Show that the mean of $Y(t)$ is $R_X(0)$.
(b) Show that the autocovariance function of $Y(t)$ is $2R_X^2(\tau)$.

5.22 A stationary, Gaussian process $X(t)$ with mean μ_X and variance σ_X^2 is passed through two linear filters with impulse responses $h_1(t)$ and $h_2(t)$, yielding processes $Y(t)$ and $Z(t)$, as shown in Figure P5.22.

(a) Determine the joint probability density function of the random variables $Y(t_1)$ and $Z(t_2)$.
(b) What conditions are necessary and sufficient to ensure that $Y(t_1)$ and $Z(t_2)$ are statistically independent?

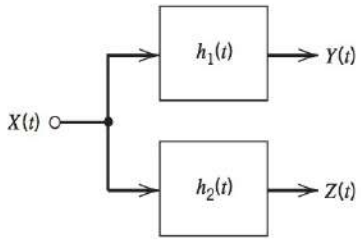


Figure P5.22

5.23 A stationary, Gaussian process $X(t)$ with zero mean and power spectral density $S_X(f)$ is applied to a linear filter whose impulse response $h(t)$ is shown in Figure P5.23. A sample Y is taken of the random process at the filter output at time T.

(a) Determine the mean and variance of Y.
(b) What is the probability density function of Y?

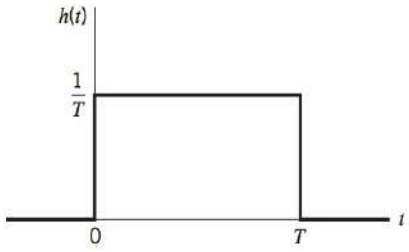


Figure P5.23

5.24 Consider a white Gaussian noise process of zero mean and power spectral density $N_0/2$ that is applied to the input of the high-pass RL filter shown in Figure P5.24.

(a) Find the autocorrelation function and power spectral density of the random process at the output of the filter.
(b) What are the mean and variance of this output?

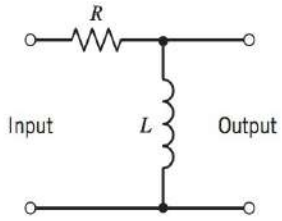


Figure P5.24

5.25 A white noise $w(t)$ of power spectral density $N_0/2$ is applied to a *Butterworth* low-pass filter of order n, whose amplitude response is defined by

$$|H(f)| = \frac{1}{[1 + (f/f_0)^{2n}]^{1/2}}$$

(a) Determine the noise equivalent bandwidth for this low-pass filter.
(b) What is the limiting value of the noise equivalent bandwidth as n approaches infinity?

5.26 The shot-noise process $X(t)$ defined by Eq. (5.114) is stationary. Why?

5.27 White Gaussian noise of zero mean and power spectral density $N_0/2$ is applied to the filtering scheme shown in Figure P5.27. The noise at the low-pass filter output is denoted by $n(t)$.

(a) Find the power spectral density and the autocorrelation function of $n(t)$.
(b) Find the mean and variance of $n(t)$.
(c) What is the rate at which $n(t)$ can be sampled so that the resulting samples are essentially uncorrelated?

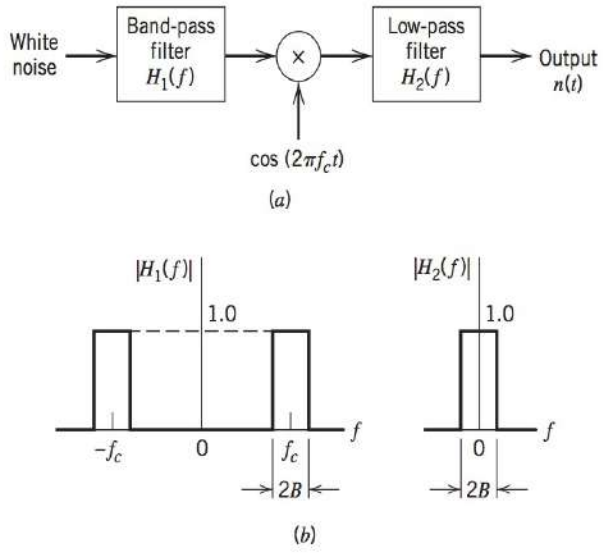


Figure P5.27

5.28 Let $X(t)$ be a stationary process with zero mean, autocorrelation function $R_X(\tau)$, and power spectral density $S_X(f)$. We are required to find a linear filter with impulse response $h(t)$, such that the filter output is $X(t)$ when the input is white noise of power spectral density $N_0/2$.

(a) Determine the condition that the impulse response $h(t)$ must satisfy in order to achieve this requirement.

(b) What is the corresponding condition on the transfer function $H(f)$ of the filter?

(c) Using the Paley–Wiener criterion (see Section 2.7), find the requirement on $S_X(f)$ for the filter to be causal.

5.29 The power spectral density of a narrowband noise $n(t)$ is as shown in Figure P5.29. The carrier frequency is 5 Hz.

(a) Find the power spectral densities of the in-phase and quadrature components of $n(t)$.

(b) Find their cross-spectral densities.

5.30 Consider a Gaussian noise $n(t)$ with zero mean and the power spectral density $S_N(f)$ shown in Figure P5.30.

(a) Find the probability density function of the envelope of $n(t)$.

(b) What are the mean and variance of this envelope?

5.31 In the noise analyzer of Figure 5.25*a*, the low-pass filters are ideal with a bandwidth equal to one-half that of the narrowband noise $n(t)$ applied to the input. Using this scheme, derive the following results:

(a) Equation (5.136), defining the power spectral densities of the in-phase noise component $n_I(t)$ and quadrature noise component $n_Q(t)$ in terms of the power spectral density of $n(t)$.

(b) Equation (5.137), defining the cross-spectral densities of $n_I(t)$ and $n_Q(t)$.

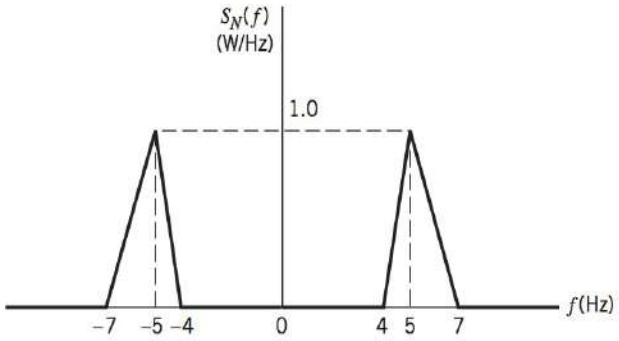


Figure P5.29

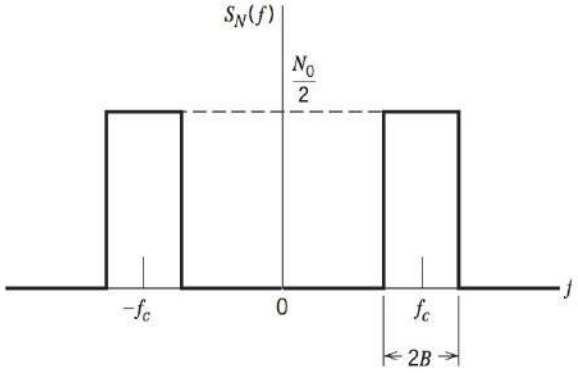


Figure P5.30

5.32 Assume that the narrowband noise $n(t)$ is Gaussian and its power spectral density $S_N(f)$ is symmetric about the mid-band frequency f_c. Show that the in-phase and quadrature components of $n(t)$ are statistically independent.

5.33

(a) A transmitter at position $x = 0$ emits the signal $A\cos(2\pi f_c t)$. The signal travels at the velocity of light such that a signal at a point on the x-axis is given by

$$r(t,x) = A(x)\cos\left[2\pi f_c\left(t - \frac{x}{c}\right)\right]$$

If the receiver starts at position x_0 and moves at velocity v along the x-axis, what Doppler shift in frequency f_D is observed?

(b) The frequencies of the Doppler shift of the reflected paths in Eq. (5.173) are proportional to the angle of radiation relative to the direction of motion, that is

$$f_n = f_D \cos\psi_n$$

where f_D is maximum Doppler shift. If the multipath angle ψ_n is uniformly distributed over $[-\pi,\pi]$. Compute $\mathbf{E}\,[\exp(j2\pi f_n\tau)]$. Use this result to prove Eq. (5.174).

5.34 The *modified Bessel function of the first kind zero order* is defined by

$$I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \exp(x\cos\psi)\,d\psi$$

Using this formula, show that the marginal distribution $f_R(r)$ derived from the joint distribution of Eq. (5.160) is defined by

$$f_R(r) = \frac{r}{\sigma^2}\exp\left(-\frac{r^2 + A^2}{2\sigma^2}\right) I_0\left(\frac{Ar}{\sigma^2}\right)$$

Hence, derive the normalized Rician distribution

$$f_V(v) = v\exp\left(\frac{v^2 + a^2}{2}\right) I_0(av)$$

which is used to plot the curves of Figure 5.29.

Computer Problems

5.35 To demonstrate the central limit theorem, use Matlab to compute 20,000 samples of $Z = \Sigma_{n=1}^{5} X_n$, where X_n is a random variable uniformly distributed over $[-1, +1]$. Estimate the corresponding probability density function by forming a histogram of the results. Compare this histogram (scaled for unit area) to the Gaussian density function having the same mean and variance. What is the relative error between the two density functions at 0σ, 1σ, 2σ, 3σ, and 4σ?

5.36 A narrowband Gaussian noise process is sampled at the Nyquist rate. The samples of the complex envelope of this process are given by

$$\tilde{n}^k = n_I^k + jn_Q^k$$

where $\{n_I^k\}$ and $\{n_Q^k\}$ are independent, white, Gaussian random variables with variance $\sigma^2 = 1$. These samples are processed by the discrete-time filter

$$\hat{y}^{k+1} = a\hat{y}^k + \hat{n}^k$$

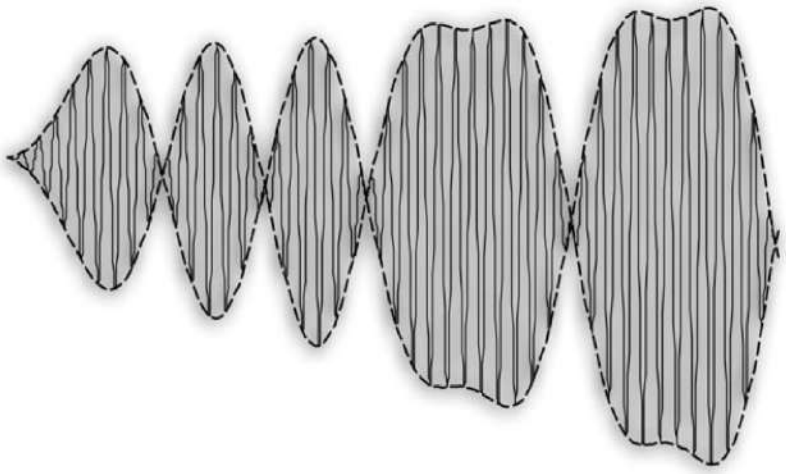
using the following Matlab script

```
a = 0.8;
sigma = 1;
K = 1000;
n = sigma * randn(K,1) + j * sigma * randn(K,1);
y = filter(1, [1 -a], n);
```

(a) What is mean and variance of output? What are the theoretical values?

(b) Is the output Gaussian? Justify your answer from simulation results and theoretically. (*Hint:* Use the *hist* function in Matlab.)

(c) What is the discrete-time autocorrelation function of the filter output? Compute theoretically and using a time-average of the filter output. Plot the latter result.



6

NOISE IN CW MODULATION SYSTEMS

6.1 INTRODUCTION

In Chapters 3 and 4, we characterized continuous-wave (CW) modulation [i.e., amplitude modulation (AM) and frequency modulation (FM)] techniques,] entirely from a deterministic perspective. Then in Chapter 5, we equipped ourselves with the mathematical tools needed for the statistical characterization of random signals and noise. We are now ready to resume the study of CW modulation systems by evaluating the effects of noise on their performance, and thereby develop a deeper understanding of analog communications.

To undertake an analysis of noise in CW modulation systems, we need to do a number of things. First and foremost, however, we must have a *receiver model.* In formulating such a model, the customary practice is to model the receiver noise (channel noise) as *additive, white,* and *Gaussian.* These simplifying assumptions enable us to obtain a basic understanding of the way in which noise affects the performance of the receiver. Moreover, it provides a framework for the comparison of the noise performances of different CW modulation–demodulation schemes.

The material of the chapter is organized as follows. In Section 6.2 we describe a receiver model and define some related quantitative measures of noise performance. This is followed by two sections on noise in AM receivers, namely, double sideband–suppressed carrier, and standard amplitude modulation receivers. Next, we discuss noise in FM receivers, the analysis of which is a more difficult task. The chapter concludes with a comparative evaluation of the noise performance of AM and FM systems.

6.2 RECEIVER MODEL

The idea of *modeling* is fundamental to the study of all physical systems, including communication systems. Through modeling, we improve our understanding of the capabilities

and limitations of a system. In formulating a receiver model for the study of noise in CW modulation systems, we need to keep the following points in mind:

- The model provides an adequate description of the form of receiver noise that is of common concern.
- The model accounts for the inherent filtering and modulation characteristics of the system.
- The model is simple enough for a statistical analysis of the system to be possible.

For the situation at hand, we propose to use the *receiver model* of Figure 6.1, shown in its most basic form. In this figure, $s(t)$ denotes the incoming *modulated signal* and $w(t)$ denotes *front-end receiver noise.* The *received signal* is therefore made up of the sum of $s(t)$ and $w(t)$; this is the signal that the receiver has to work on. The *band-pass filter* in the model of Figure 6.1 represents the combined filtering action of the tuned amplifiers used in the actual receiver for the purpose of signal amplification prior to demodulation. The bandwidth of this band-pass filter is just wide enough to pass the modulated signal $s(t)$ without distortion. As for the *demodulator* in the model of Figure 6.1, its details naturally depend on the type of modulation used.

In performing the noise analysis of a communication system, the customary practice is to assume that the noise $w(t)$ is *additive, white,* and *Gaussian.* For many communication receivers, this assumption is very accurate for the band of frequencies of interest. It also conveniently simplifies some of the mathematical computations. We thus let the power spectral density of the noise $w(t)$ be denoted by $N_0/2$, defined for both positive and negative frequencies. That is, N_0 *is the average noise power per unit bandwidth measured at the front end of the receiver.* We also assume that the band-pass filter in the receiver model of Figure 6.1 is ideal, having a bandwidth equal to the transmission bandwidth B_T of the modulated signal $s(t)$ and a mid-band frequency equal to the carrier frequency f_c. The latter assumption is justified for double sideband–suppressed carrier (DSB–SC) modulation, standard amplitude modulation (AM), and frequency modulation (FM). The cases of single sideband (SSB) modulation and vestigial sideband (VSB) modulation require special considerations. Taking the mid-band frequency of the band-pass filter to be the same as the carrier frequency f_c, we may model the power spectral density $S_N(f)$ of the noise $n(t)$, resulting from the passage of the white noise $w(t)$ through the filter, as shown in Figure 6.2. Typically, the carrier frequency f_c is large compared to the transmission bandwidth B_T. We may therefore treat the *filtered noise* $n(t)$ as a narrowband noise represented in the canonical form

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \tag{6.1}$$

where $n_I(t)$ is the *in-phase noise component* and $n_Q(t)$ is the *quadrature noise component,* both measured with respect to the carrier wave $A_c\cos(2\pi f_c t)$. The filtered signal $x(t)$ available for demodulation is defined by

$$x(t) = s(t) + n(t) \tag{6.2}$$

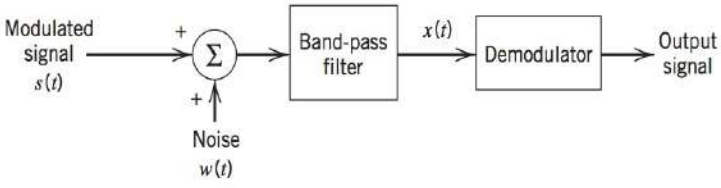


FIGURE 6.1 Noisy receiver model.

The details of $s(t)$ depend on the type of modulation used, in any event, the average noise power at the demodulator input is equal to the total area under the curve of the power spectral density $S_N(f)$. From Figure 6.2 we readily see that this average noise power is equal

to $N_0 B_T$. Given the format of $s(t)$, we may also determine the average signal power at the demodulator input. With the modulated signal $s(t)$ and the filtered noise $n(t)$ appearing additively at the demodulator input in accordance with Eq. (6.2), we may go on to define an *input signal-to-noise ratio*, $(\text{SNR})_I$, *as the ratio of the average power of the modulated signal* $s(t)$ *to the average power of the filtered noise* $n(t)$.

A more useful measure of noise performance, however, is the *output signal-to-noise ratio,* $(\text{SNR})_O$, *defined as the ratio of the average power of the demodulated message signal to the average power of the noise, both measured at the receiver output.* The output signal-to-noise ratio provides an intuitive measure for describing the fidelity with which the demodulation process in the receiver recovers the message signal from the modulated signal in the presence of additive noise. For such a criterion to be well defined, the recovered message signal and the corruptive noise component must appear *additively* at the demodulator output. This condition is perfectly valid in the case of a receiver using coherent detection. On the other hand, when the receiver uses envelope detection as in full AM or frequency discrimination as in FM, we have to assume that the average power of the filtered noise $n(t)$ is relatively low to justify the use of output signal-to-noise ratio as a measure of receiver performance.

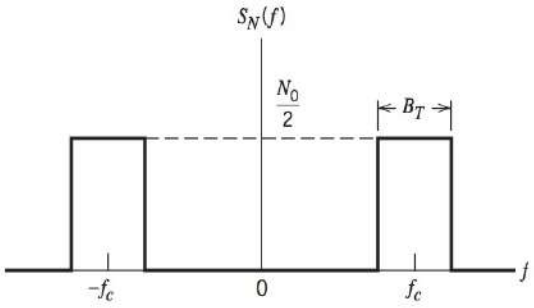


FIGURE 6.2 Idealized characteristic of band-pass filtered noise.

The output signal-to-noise ratio depends, among other factors, on the type of modulation used in the transmitter and the type of demodulation used in the receiver. Thus it is informative to compare the output signal-to-noise ratios for different modulation-demodulation systems. However, for this comparison to be of meaningful value, it must be made on an equal basis as described here:

- The modulated signal $s(t)$ transmitted by each system has the same average power.
- The front-end receiver noise $w(t)$ has the same average power measured in the message bandwidth W.

Accordingly, as a frame of reference we define the *channel signal-to-noise ratio,* $(\text{SNR})_C$, as *the ratio of the average power of the modulated signal to the average power of noise in the message bandwidth, both measured at the receiver input.* This ratio may be viewed as the signal-to-noise ratio that results from *baseband (direct) transmission* of the message signal $m(t)$ without modulation, as modeled in Figure 6.3. Here it is assumed that (1) the message power at the low-pass filter input is adjusted to be the same as the average power of the modulated signal, and (2) the low-pass filter passes the message signal and rejects out-of-band noise.

For the purpose of comparing different continuous-wave (CW) modulation systems, we *normalize* the receiver performance by dividing the output signal-to-noise ratio by the channel signal-to-noise ratio. We thus define a *figure of merit* for the receiver as follows:

$$\text{Figure of merit} = \frac{(\text{SNR})_O}{(\text{SNR})_C} \tag{6.3}$$

Clearly, the higher the value of the figure of merit, the better will the noise performance of the receiver be. The figure of merit may equal one, be less than one, or be greater than one, depending on the type of modulation used.

In the next three sections, we use the ideas described herein to perform a noise analysis of (1) DSB–SC receivers using coherent detection, (2) AM receivers using envelope detection, and (3) FM receivers using frequency discrimination. We also

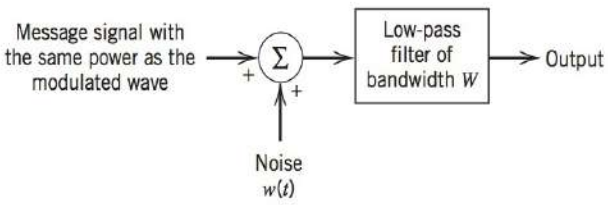


FIGURE 6.3 The baseband transmission model, assuming a message signal of bandwidth W, used for calculating channel signal-to-noise ratio.

consider related issues that arise under high noise levels. These receivers pertain to typical examples of CW modulation systems, which exhibit different noise behavior.

6.3 NOISE IN DSB–SC RECEIVERS

The noise analysis of a DSB–SC receiver using coherent detection is the simplest of the above-mentioned cases. Figure 6.4 shows the model of a DSB–SC receiver using a coherent detector. The use of coherent detection requires multiplication of the filtered signal $x(t)$ by a locally generated sinusoidal wave $\cos(2\pi f_c t)$ and then low-pass filtering the product. To simplify the analysis, we assume that the amplitude of the locally generated sinusoidal wave is unity. For this demodulation scheme to operate satisfactorily, however, it is necessary that the local oscillator be synchronized both in phase and in frequency with the oscillator generating the carrier wave in the transmitter. We assume that this synchronization has been achieved.

The DSB–SC component of the filtered signal $x(t)$ is expressed as

$$s(t) = CA_c \cos(2\pi f_c t) m(t) \tag{6.4}$$

where $A_c \cos(2\pi f_c t)$ is the sinusoidal carrier wave and $m(t)$ is the message signal. In the expression for $s(t)$ in Eq. (6.4) we have included a *system-dependent scaling factor C*, the purpose of which is to ensure that the signal component $s(t)$ is measured in the same units as the additive noise component $n(t)$. We assume that $m(t)$ is the sample function of a stationary process of zero mean, whose power spectral density $S_M(f)$ is limited to a maximum frequency W; that is, W is the *message bandwidth.* The average power P of the message signal is the total area under the curve of power spectral density, as shown by

$$P = \int_{-W}^{W} S_M(f)\, df \tag{6.5}$$

The carrier wave is statistically independent of the message signal. To emphasize this independence, the carrier should include a random phase that is uniformly distributed over 2π radians. In the defining equation for $s(t)$ this random phase angle has been omitted for convenience of presentation. Using the result of Example 12 of Chapter 5 on a modulated random process, we may express the average power of the DSB–SC modulated signal component $s(t)$ as $C^2A_c^2P/2$. With a noise spectral density of $N_0/2$, the average noise power in the message bandwidth W is equal to WN_0. The channel signal-to-noise ratio of the DSB–SC modulation system is therefore

$$(\text{SNR})_{C,\text{DSB}} = \frac{C^2A_c^2P}{2WN_0} \tag{6.6}$$

where the constant C^2 in the numerator ensures that this ratio is dimensionless.

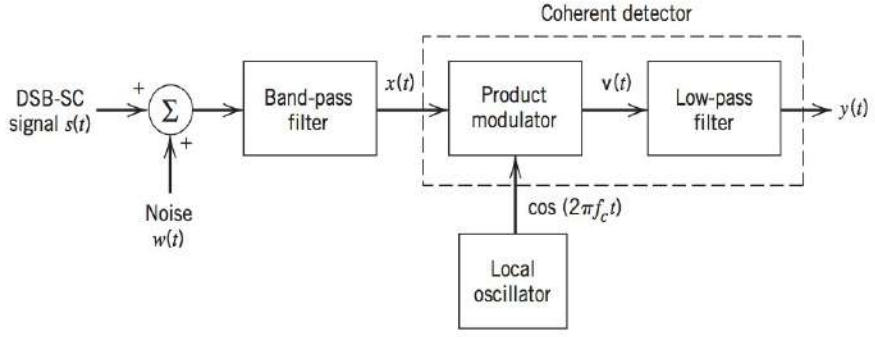


FIGURE 6.4 Model of DSB–SC receiver using coherent detection.

Next, we wish to determine the output signal-to-noise ratio of the system. Using the narrowband representation of the filtered noise $n(t)$, the total signal at the coherent detector input may be expressed as

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= CA_c \cos(2\pi f_c t)\, m(t) + n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \end{aligned} \tag{6.7}$$

where $n_I(t)$ and $n_Q(t)$ are the in-phase and quadrature components of $n(t)$ with respect to the carrier. The output of the product–modulator component of the coherent detector is therefore

$$\begin{aligned} v(t) &= x(t)\cos(2\pi f_c t) \\ &= \frac{1}{2}CA_c m(t) + \frac{1}{2}n_I(t) \\ &\quad + \frac{1}{2}[CA_c m(t) + n_I(t)]\cos(4\pi f_c t) - \frac{1}{2}A_c n_Q(t)\sin(4\pi f_c t) \end{aligned}$$

The low-pass filter in the coherent detector removes the high frequency components of $v(t)$, yielding a receiver output

$$y(t) = \frac{1}{2}CA_c m(t) + \frac{1}{2}n_I(t) \tag{6.8}$$

Equation (6.8) indicates the following:

1. The message signal $m(t)$ and in-phase noise component $n_I(t)$ of the filtered noise $n(t)$ appear additively at the receiver output.
2. The quadrature component $n_Q(t)$ of the noise $n(t)$ is completely rejected by the coherent detector.

These two results are independent of the input signal-to-noise ratio. Thus, coherent detection distinguishes itself from other demodulation techniques in the important property: the output message component is unmutilated and the noise component always appears additively with the message, irrespective of the input signal-to-noise ratio.

The message signal component at the receiver output is $CA_c m(t)/2$. Therefore, the average power of this component may be expressed as $C^2A_c^2P/4$, where P is the average power of the original message signal $m(t)$ and C is the system-dependent scaling factor referred to earlier.

In the case of DSB–SC modulation, the band-pass filter in Figure 6.4 has a bandwidth B_T equal to $2W$ in order to accommodate the upper and lower sidebands of the modulated signal $s(t)$. It follows therefore that the average power of the filtered noise $n(t)$ is $2WN_0$. From Property 5 of narrowband noise described in Section 5.11, we know that the average power of the (low-pass) in-phase noise component $n_I(t)$ is the same as that of the (band-pass) filtered noise $n(t)$. Since from Eq. (6.8) the noise component at the receiver output is $n_I(t)/2$, it follows that the average power of the noise at the receiver output is

$$\left(\frac{1}{2}\right)^2 2WN_0 = \frac{1}{2}WN_0$$

The output signal-to-noise for a DSB–SC receiver using coherent detection is therefore

$$\begin{aligned} (\text{SNR})_O &= \frac{C^2A_c^2P/4}{WN_0/2} \\ &= \frac{C^2A_c^2P}{2WN_0} \end{aligned} \tag{6.9}$$

Using Eqs. (6.6) and (6.9), we obtain the figure of merit

$$\left.\frac{(\text{SNR})_O}{(\text{SNR})_C}\right|_{\text{DSB–SC}} = 1 \tag{6.10}$$

Note that the factor C^2 is common to both the output and channel signal-to-noise ratios, and therefore cancels out in evaluating the figure of merit.

Note also that at the coherent detector output in the receiver of Figure 6.4 using DSB–SC modulation, the translated signal sidebands sum coherently, whereas the translated noise sidebands sum incoherently. This means that the output signal-to-noise ratio in this receiver is twice the signal-to-noise ratio at the coherent detector input.

6.4 NOISE IN AM RECEIVERS

The next noise analysis we perform is for an amplitude modulation (AM) system using an *envelope detector* in the receiver, as shown in the model of Figure 6.5. In a full AM signal, both sidebands and the carrier wave are transmitted as shown by

$$s(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t) \tag{6.11}$$

where $A_c \cos(2\pi f_c t)$ is the carrier wave, $m(t)$ is the message signal, and k_a is a constant that determines the percentage modulation. In the expression for the amplitude-modulated signal in Eq. (6.11), it is reasonable to assume that the carrier amplitude A_c has the same units as the additive noise. The factor k_a is then assumed to have the units necessary to make the remainder of the expression dimensionless.

As with the DSB–SC receiver, we perform the noise analysis of the AM receiver by first determining the channel signal-to-noise ratio, and then the output signal-to-noise ratio.

The average power of the carrier component in the AM signal $s(t)$ is $A_c^2/2$. The average power of the information-bearing component $A_c k_a m(t)\cos(2\pi f_c t)$ is $A_c^2 k_a^2 P/2$, where P is the average power of the message signal $m(t)$.The average power of the full AM signal $s(t)$ is therefore equal to $A_c^2(1 + k_a^2 P)/2$. As for the DSB–SC system, the average power of noise in the message bandwidth is WN_0. The channel signal-to-noise ratio for AM is therefore

$$(\text{SNR})_{C,\text{AM}} = \frac{A_c^2(1 + k_a^2 P)}{2WN_0} \tag{6.12}$$

To evaluate the output signal-to-noise ratio, we first represent the filtered noise $n(t)$ in terms of its in-phase and quadrature components. We may therefore define the filtered signal $x(t)$ applied to the envelope detector in the receiver model of Figure 6.5 as follows:

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= [A_c + A_c k_a m(t) + n_I(t)]\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \end{aligned} \tag{6.13}$$

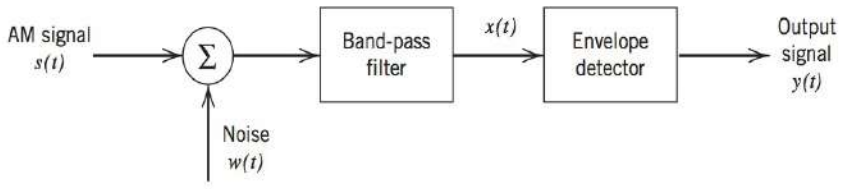


FIGURE 6.5 Noisy model of AM receiver.

It is informative to represent the components that comprise the signal $x(t)$ by means of phasors, as in Figure 6.6*a*. From this phasor diagram, the receiver output is readily obtained as

$$\begin{aligned} y(t) &= \text{envelope of } x(t) \\ &= \{[A_c + A_c k_a m(t) + n_I(t)]^2 + n_Q^2(t)\}^{1/2} \end{aligned} \tag{6.14}$$

The signal $y(t)$ defines the output of an ideal envelope detector. The phase of $x(t)$ is of no interest to us, because an ideal envelope detector is totally insensitive to variations in the phase of $x(t)$.

The expression defining $y(t)$ is somewhat complex and needs to be simplified in some manner in order to permit the derivation of insightful results. Specifically, we would like to approximate the output $y(t)$ as the sum of a message term plus a term due to noise. In general, this is quite difficult to achieve. However, when the average carrier power is large compared with the average noise power, so that the receiver is operating satisfactorily, then the signal term $A_c[1 + k_a m(t)]$ will be large compared with the noise terms $n_I(t)$ and $n_Q(t)$, at least most of the time. Then we may approximate the output $y(t)$ as (see Problem 6.7):

$$y(t) \simeq A_c + A_c k_a m(t) + n_I(t) \tag{6.15}$$

The presence of the dc or constant term A_c in the envelope detector output $y(t)$ of Eq. (6.15) is due to demodulation of the transmitted carrier wave. We may ignore this term, however, because it bears no relation whatsoever to the message signal $m(t)$. In any case, it may be removed simply by means of a blocking capacitor. Thus if we neglect the dc term A_c in Eq. (6.15), we find that the remainder has, except for scaling factors, a form similar to the output of a DSB–SC receiver using coherent detection. Accordingly, the output signal-to-noise ratio of an AM receiver using an envelope detector is approximately

$$(\text{SNR})_{O,\text{AM}} \simeq \frac{A_c^2 k_a^2 P}{2WN_0} \tag{6.16}$$

Equation (6.16) is, however, valid only if the following two conditions are satisfied:

1. The average noise power is small compared to the average carrier power at the envelope detector input.
2. The amplitude sensitivity k_a is adjusted for a percentage modulation less than or equal to 100 percent.

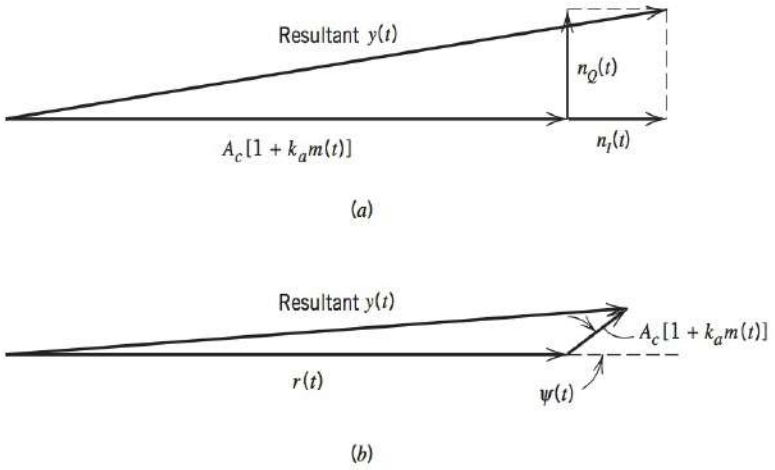


FIGURE 6.6 (*a*) Phasor diagram for AM wave plus narrowband noise for the case of high carrier-to-noise ratio. (*b*) Phasor diagram for AM wave plus narrowband noise for the case of low carrier-to-noise ratio.

Using Eqs. (6.12) and (6.16), we obtain the following figure of merit for amplitude modulation:

$$\left.\frac{(\text{SNR})_O}{(\text{SNR})_C}\right|_{\text{AM}} \simeq \frac{k_a^2 P}{1 + k_a^2 P} \tag{6.17}$$

Thus, whereas the figure of merit of a DSB–SC receiver or that of an SSB receiver using coherent detection is always unity, the corresponding figure of merit of an AM receiver using envelope detection is always less than unity. In other words, *the noise performance of an AM receiver is always inferior to that of a DSB–SC receiver.* This is due to the wastage of transmitter power, which results from transmitting the carrier as a component of the AM wave.

EXAMPLE 6.1 Single-Tone Modulation

Consider the special case of a sinusoidal wave of frequency f_m and amplitude A_m as the modulating wave, as shown by

$$m(t) = A_m \cos(2\pi f_m t)$$

The corresponding AM wave is

$$s(t) = A_c[1 + \mu \cos(2\pi f_m t)]\cos(2\pi f_c t)$$

where $\mu = k_a A_m$ is the modulation factor. The average power of the modulating wave $m(t)$ is (assuming a load resistor of 1 ohm)

$$P = \frac{1}{2}A_m^2$$

Therefore, using Eq. (6.17), we get

$$\begin{aligned}\left.\frac{(\text{SNR})_O}{(\text{SNR})_C}\right|_{\text{AM}} &= \frac{\frac{1}{2}k_a^2 A_m^2}{1 + \frac{1}{2}k_a^2 A_m^2} \\ &= \frac{\mu^2}{2 + \mu^2}\end{aligned} \tag{6.18}$$

When $\mu = 1$, which corresponds to 100 percent modulation, we get a figure of merit equal to 1/3. This means that, other factors being equal, an AM system (using envelope detection) must transmit three times as much average power as a suppressed-carrier system (using coherent detection) in order to achieve the same quality of noise performance.

THRESHOLD EFFECT

When the carrier-to-noise ratio is small compared with unity, the noise term dominates and the performance of the envelope detector changes completely from that described above. In this case it is more convenient to represent the narrowband noise $n(t)$ in terms of its envelope $r(t)$ and phase $\psi(t)$, as shown by

$$n(t) = r(t)\cos[2\pi f_c t + \psi(t)] \tag{6.19}$$

The corresponding phasor diagram for the detector input $x(t) = s(t) + n(t)$ is shown in Figure 6.6*b*, where we have used the noise as reference, because it is now the dominant term. To the noise phasor $r(t)$ we have added a phasor representing the signal term $A_c[1 + k_a m(t)]$, with the angle between them being equal to the relative phase $\psi(t)$ between the noise $n(t)$ carrier $\cos(2\pi f_c t)$. In Figure 6.6*b* it is assumed that the carrier-to-noise ratio is so low that the carrier amplitude A_c is small compared with the noise envelope $r(t)$, at least most of the time. Then we may neglect the quadrature component of the signal with respect to the noise, and thus find directly from Figure 6.6*b* that the envelope detector output is approximately

$$y(t) \simeq r(t) + A_c \cos[\psi(t)] + A_c k_a m(t) \cos[\psi(t)] \tag{6.20}$$

This relation reveals that when the carrier-to-noise ratio is low, the detector output has no component strictly proportional to the message signal $m(t)$. The last term of the expression defining $y(t)$ contains the message signal $m(t)$ multiplied by noise in the form of $\cos[\psi(t)]$. From Section 5.12 we recall that the phase $\psi(t)$ of the narrowband noise $n(t)$ is uniformly distributed over 2π radians. It follows therefore that we have a complete loss of information in that the detector output does not contain the message signal $m(t)$ at all. The loss of a message in an envelope detector that operates at a low carrier-to-noise ratio is referred to as the *threshold effect.* By threshold we mean *a value of the carrier-to-noise ratio below which the noise performance of a detector deteriorates much more rapidly than proportionately to the carrier-to-noise ratio.* It is important to recognize that every nonlinear detector (e.g., envelope detector) exhibits a threshold effect. On the other hand, such an effect does *not* arise in a coherent detector.

6.5 NOISE IN FM RECEIVERS

We finally turn our attention to the noise analysis of a frequency modulation (FM) system, for which we use the receiver model shown in Figure 6.7. As before, the noise $w(t)$ is modeled as white Gaussian noise of zero mean and power spectral density $N_0/2$. The received FM signal $s(t)$ has a carrier frequency f_c and transmission bandwidth B_T, such that only a negligible amount of power lies outside the frequency band $f_c \pm B_T/2$ for positive frequencies.

As in the AM case, the band-pass filter has a mid-band frequency f_c and bandwidth B_T and therefore passes the FM signal essentially without distortion. Ordinarily, B_T is small compared with the mid-band frequency f_c, so that we may use the narrowband representation for $n(t)$, the filtered version of receiver noise $w(t)$, in terms of its in-phase and quadrature components.

In an FM system, the message information is transmitted by variations of the instantaneous frequency of a sinusoidal carrier wave, and its amplitude is maintained constant. Therefore, any variations of the carrier amplitude at the receiver input must result from

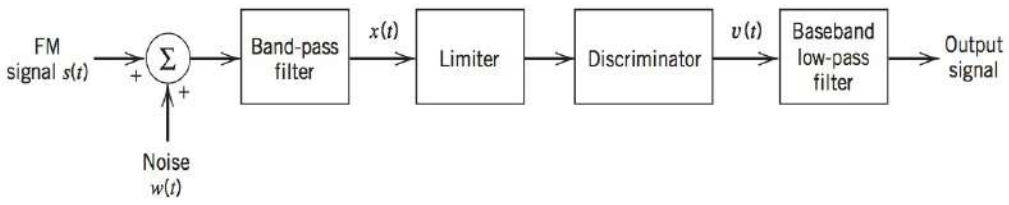


FIGURE 6.7 Noisy model of an FM receiver.

noise or interference. The amplitude *limiter,* following the bandpass filter in the receiver model of Figure 6.7, is used to remove amplitude variations by clipping the modulated wave at the filter output almost to the zero axis. The resulting rectangular wave is rounded off by another bandpass filter that is an integral part of the limiter, thereby suppressing harmonics of the carrier frequency. Thus, the filter output is again sinusoidal, with an amplitude that is practically independent of the carrier amplitude at the receiver input.

The discriminator in the model of Figure 6.7 consists of two components:

1. A *slope network* or *differentiator* with a purely imaginary transfer function that varies linearly with frequency. It produces a hybrid-modulated wave in which both amplitude and frequency vary in accordance with the message signal.
2. An envelope detector that recovers the amplitude variation and thus reproduces the message signal.

The slope network and envelope detector are usually implemented as integral parts of a single physical unit.

The *postdetection filter,* labeled "baseband low-pass filter" in Figure 6.7, has a bandwidth that is just large enough to accommodate the highest frequency component of the message signal. This filter removes the out-of-band components of the noise at the discriminator output and thereby keeps the effect of the output noise to a minimum.

The filtered noise $n(t)$ at the band-pass filter output in Figure 6.7 is defined in terms of its in-phase and quadrature components by

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)$$

Equivalently, we may express $n(t)$ in terms of its envelope and phase as

$$n(t) = r(t)\cos[(2\pi f_c t) + \psi(t)] \tag{6.21}$$

where the envelope is

$$r(t) = [n_I^2(t) + n_Q^2(t)]^{1/2} \tag{6.22}$$

and the phase is

$$\psi(t) = \tan^{-1}\left[\frac{n_Q(t)}{n_I(t)}\right] \tag{6.23}$$

The envelope $r(t)$ is Rayleigh distributed, and the phase $\psi(t)$ is uniformly distributed over 2π radians (see Section 5.11).

The incoming FM signal $s(t)$ is defined by

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau)\, d\tau\right] \tag{6.24}$$

where A_c is the carrier amplitude, f_c is the carrier frequency, k_f is the frequency sensitivity, and $m(t)$ is the message signal. Note that, as with the standard AM, in FM there is no need to introduce a scaling factor in the definition of the modulated signal $s(t)$, since it is reasonable to assume that its amplitude A_c has the same units as the additive noise component $n(t)$. To proceed, we define

$$\phi(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau \tag{6.25}$$

We may thus express $s(t)$ in the simple form

$$s(t) = A_c \cos[2\pi f_c t + \phi(t)] \tag{6.26}$$

The noisy signal at the band-pass filter output is therefore

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= A_c \cos[2\pi f_c t + \phi(t)] + r(t)\cos[2\pi f_c t + \psi(t)] \end{aligned} \tag{6.27}$$

It is informative to represent $x(t)$ by means of a phasor diagram, as in Figure 6.8. In this diagram we have used the signal term as reference. The phase $\theta(t)$ of the resultant phasor representing $x(t)$ is obtained directly from Figure 6.8 as

$$\theta(t) = \phi(t) + \tan^{-1}\left\{\frac{r(t)\sin[\psi(t) - \phi(t)]}{A_c + r(t)\cos[\psi(t) - \phi(t)]}\right\} \tag{6.28}$$

The envelope of $x(t)$ is of no interest to us, because any envelope variations at the band-pass output are removed by the limiter.

Our motivation is to determine the error in the instantaneous frequency of the carrier wave caused by the presence of the filtered noise $n(t)$. With the discriminator assumed ideal, its output is proportional to $\theta'(t)/2\pi$ where $\theta'(t)$ is the derivative of $\theta(t)$ with respect to time. In view of the complexity of the expression defining $\theta(t)$, however, we need to make certain simplifying approximations, so that our analysis may yield useful results.

We assume that the carrier-to-noise ratio measured at the discriminator input is large compared with unity. Let R denote the random variable obtained by observing (at some fixed time) the envelope process with sample function $r(t)$ [due to the noise $n(t)$]. Then, at least most of the time, the random variable R is small compared with the carrier amplitude A_c, and so the expression for the phase $\theta(t)$ simplifies considerably as follows:

$$\theta(t) \simeq \phi(t) + \frac{r(t)}{A_c}\sin[\psi(t) - \phi(t)] \tag{6.29}$$

or, using the expression for $\phi(t)$ given in Eq. (6.25),

$$\theta(t) \simeq 2\pi k_f \int_0^t m(t)\,dt + \frac{r(t)}{A_c}\sin[\psi(t) - \phi(t)] \tag{6.30}$$

The discriminator output is therefore

$$\begin{aligned} v(t) &= \frac{1}{2\pi}\frac{d\theta(t)}{dt} \\ &\simeq k_f m(t) + n_d(t) \end{aligned} \tag{6.31}$$

where the noise term $n_d(t)$ is defined by

$$n_d(t) = \frac{1}{2\pi A_c}\frac{d}{dt}\{r(t)\sin[\psi(t) - \phi(t)]\} \tag{6.32}$$

We thus see that provided the carrier-to-noise ratio is high, the discriminator output $v(t)$ consists of the original message or modulating wave $m(t)$ multiplied by the constant factor k_f, plus an additive noise component $n_d(t)$. Accordingly, we may use the output signal-to-noise ratio as previously defined to assess the quality of performance of the FM receiver. Before doing this, however, it is instructive to see if we can simplify the expression defining the noise $n_d(t)$.

From the phasor diagram of Figure 6.8, we note that the effect of variations in the phase $\psi(t)$ of the narrowband noise appear in

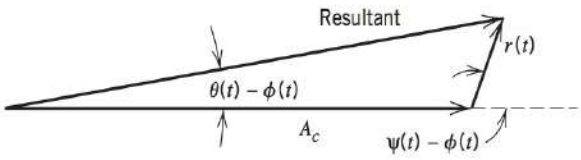


FIGURE 6.8 Phasor diagram for FM wave plus narrowband noise for the case of high carrier-to-noise ratio.

combination with the signal term $\phi(t)$. We know that the phase $\psi(t)$ is uniformly distributed over 2π radians. It would therefore be tempting to assume that the phase difference $\psi(t) - \phi(t)$ is also uniformly distributed over 2π radians. If such an assumption were true, then the noise $n_d(t)$ at the discriminator output would be independent of the modulating signal and would depend only on the characteristics of the carrier and narrowband noise. Theoretical considerations show that this assumption is justified provided that the carrier-to-noise ratio is high. Then we may simplify Eq. (6.32) as:

$$n_d(t) \simeq \frac{1}{2\pi A_c} \frac{d}{dt}\{r(t)\sin[\psi(t)]\} \tag{6.33}$$

However, from the defining equations for $r(t)$ and $\psi(t)$, we note that the quadrature component $n_Q(t)$ of the filtered noise $n(t)$ is

$$n_Q(t) = r(t)\sin[\psi(t)] \tag{6.34}$$

Therefore, we may rewrite Eq. (6.33) as

$$n_d(t) \simeq \frac{1}{2\pi A_c} \frac{dn_Q(t)}{dt} \tag{6.35}$$

This means that *the additive noise $n_d(t)$ appearing at the discriminator output is determined effectively by the carrier amplitude A_c and the quadrature component $n_Q(t)$ of the narrowband noise $n(t)$.*

The output signal-to-noise ratio is defined as the ratio of the average output signal power to the average output noise power. From Eq. (6.31), we see that the message component in the discriminator output, and therefore the low-pass filter output, is $k_f m(t)$. Hence, the average output signal power is equal to $k_f^2 P$, where P is the average power of the message signal $m(t)$.

To determine the average output noise power, we note that the noise $n_d(t)$ at the discriminator output is proportional to the time derivative of the quadrature noise component $n_Q(t)$. Since the differentiation of a function with respect to time corresponds to multiplication of its Fourier transform by $j2\pi f$, it follows that we may obtain the noise process $n_d(t)$ by passing $n_Q(t)$ through a linear filter with a transfer function equal to

$$\frac{j2\pi f}{2\pi A_c} = \frac{jf}{A_c}$$

This means that the power spectral density $S_{N_d}(f)$ of the noise $n_d(t)$ is related to the power spectral density $S_{N_Q}(f)$ of the quadrature noise component $n_Q(t)$ as follows:

$$S_{N_d}(f) = \frac{f^2}{A_c^2} S_{N_Q}(f) \tag{6.36}$$

With the band-pass filter in the receiver model of Figure 6.7 having an ideal frequency response characterized by bandwidth B_T and mid-band frequency f_c, it follows that the narrowband noise $n(t)$ will have a power spectral density characteristic that is similarly shaped. This means that the quadrature component $n_Q(t)$ of the narrowband noise $n(t)$ will have the ideal low-pass characteristic shown in Figure 6.9*a*. The corresponding power spectral density of the noise $n_d(t)$ is shown in Figure 6.9*b*; that is,

$$S_{N_d}(f) = \begin{cases} \dfrac{N_0 f^2}{A_c^2}, & |f| \le \dfrac{B_T}{2} \\ 0, & \text{otherwise} \end{cases} \tag{6.37}$$

In the receiver model of Figure 6.7, the discriminator output is followed by a low-pass filter with a bandwidth equal to the message bandwidth W. For wide-band FM, we usually find that W is smaller than $B_T/2$, where B_T is the transmission bandwidth of the FM signal. This means that the out-of-band components of noise $n_d(t)$ will be rejected. Therefore, the power spectral density $S_{N_o}(f)$ of the noise $n_o(t)$ appearing at the receiver output is defined by

$$S_{N_o}(f) = \begin{cases} \dfrac{N_0 f^2}{A_c^2}, & |f| \le W \\ 0, & \text{otherwise} \end{cases} \tag{6.38}$$

as shown in Figure 6.9*c*. The average output noise power is determined by integrating the power spectral density $S_{N_o}(f)$ from $-W$ to W. We thus get the result:

$$\begin{aligned} \text{Average power of output noise} &= \frac{N_0}{A_c^2}\int_{-W}^{W} f^2\, df \\ &= \frac{2N_0 W^3}{3A_c^2} \end{aligned} \tag{6.39}$$

Note that the average output noise power is inversely proportional to the average carrier power $A_c^2/2$. Accordingly, in an FM system, increasing the carrier power has a *noise-quieting effect.*

Earlier we determined the average output signal power as $k_f^2 P$. Therefore, provided the carrier-to-noise ratio is high, we may divide this average output signal power by the average output noise power of Eq. (6.39) to obtain the output signal-to-noise ratio

$$(\text{SNR})_{O,\text{FM}} = \frac{3A_c^2 k_f^2 P}{2N_0 W^3} \tag{6.40}$$

The average power in the modulated signal $s(t)$ is $A_c^2/2$, and the average noise power in the message bandwidth is WN_0. Thus the channel signal-to-noise ratio is

$$(\text{SNR})_{C,\text{FM}} = \frac{A_c^2}{2WN_0} \tag{6.41}$$

Dividing the output signal-to-noise ratio by the channel signal-to-noise ratio, we get the following figure of merit for frequency modulation:

$$\left.\frac{(\text{SNR})_O}{(\text{SNR})_C}\right|_{\text{FM}} = \frac{3k_f^2 P}{W^2} \tag{6.42}$$

From Section 4.3 we recall that the frequency deviation Δf is proportional to the frequency sensitivity k_f of the modulator. Also, by definition, the deviation ratio D is equal to the frequency deviation Δf divided by the message bandwidth W. In other words, the deviation ratio D is proportional to the ratio $k_f P^{1/2}/W$. It follows therefore from Eq. (6.42) that the figure of merit of a wide-hand FM system is a quadratic function of the deviation ratio. Now, in wide-band FM, the transmission bandwidth B_T is approximately proportional to the deviation ratio D. Accordingly, we may state that *when the*

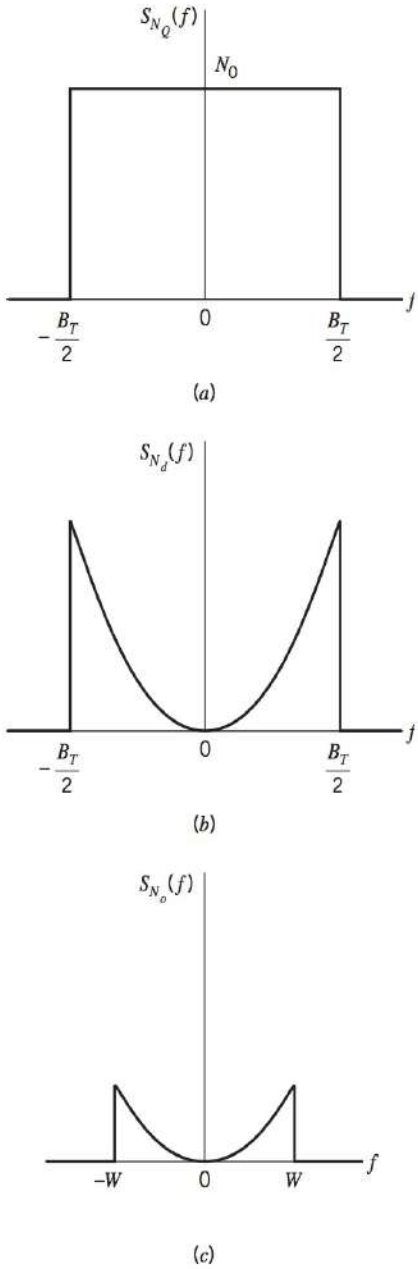


FIGURE 6.9 Noise analysis of FM receiver. (*a*) Power spectral density of quadrature component $n_Q(t)$ of narrowband noise $n(t)$. (*b*) Power spectral density of noise $n_d(t)$ at discriminator output. (*c*) Power spectral density of noise $n_o(t)$ at receiver output.

carrier-to-noise ratio is high, an increase in the transmission bandwidth B_T provides a corresponding quadratic increase in the output signal-to-noise ratio or figure of merit of the FM *system.* The important point to note from this statement is that, unlike amplitude modulation, the use of frequency modulation does provide a useful mechanism for the exchange of increased transmission bandwidth for improved noise performance.

EXAMPLE 6.2 Single-Tone Modulation

Consider the case of a sinusoidal wave of frequency f_m as the modulating signal, and assume a peak frequency deviation Δf. The modulated FM signal is thus defined by

$$s(t) = A_c \cos\left[2\pi f_c t + \frac{\Delta f}{f_m}\sin(2\pi f_m t)\right]$$

Therefore, we may write

$$2\pi k_f \int_0^t m(\tau)\, d\tau = \frac{\Delta f}{f_m}\sin(2\pi f_m t)$$

Differentiating both sides with respect to time and solving for $m(t)$, we get

$$m(t) = \frac{\Delta f}{k_f}\cos(2\pi f_m t)$$

Hence, the average power of the message signal $m(t)$, developed across a 1-ohm load, is

$$P = \frac{(\Delta f)^2}{2k_f^2}$$

Substituting this result into the formula for the output signal-to-noise ratio given in Eq. (6.40), we get

$$\begin{aligned}(\text{SNR})_{O,\text{FM}} &= \frac{3A_c^2(\Delta f)^2}{4N_0W^3}\\ &= \frac{3A_c^2\beta^2}{4N_0W}\end{aligned}$$

where $\beta = \Delta f/W$ is the modulation index. Using Eq. (6.42) to evaluate the corresponding figure of merit, we get

$$\begin{aligned}\left.\frac{(\text{SNR})_O}{(\text{SNR})_C}\right|_{\text{FM}} &= \frac{3}{2}\left(\frac{\Delta f}{W}\right)^2\\ &= \frac{3}{2}\beta^2\end{aligned} \tag{6.43}$$

It is important to note that the modulation index $\beta = \Delta f/W$ is determined by the bandwidth W of the postdetection low-pass filter and is not related to the sinusoidal message frequency f_m, except insofar as this filter is usually chosen so as to pass the spectrum of the desired message; this is merely a matter of

consistent design. For a specified system bandwidth W, the sinusoidal message frequency f_m may lie anywhere between 0 and W and would yield the same output signal-to-noise ratio.

It is of particular interest to compare the noise performances of AM and FM systems. An insightful way of making this comparison is to consider the figures of merit of the two systems based on a sinusoidal modulating signal. For an AM system operating with a sinusoidal modulating signal and 100 percent modulation, we have (from Example 6.1):

$$\left.\frac{(\text{SNR})_O}{(\text{SNR})_C}\right|_{\text{AM}} = \frac{1}{3}$$

Comparing this figure of merit with the corresponding result described in Eq. (6.43) for an FM system, we see that the use of frequency modulation offers the possibility of improved noise performance over amplitude modulation when

$$\frac{3}{2}\beta^2 > \frac{1}{3}$$

that is,

$$\beta > \frac{\sqrt{2}}{3} = 0.471$$

We may therefore consider $\beta = 0.5$ as defining roughly *the transition between narrowband FM and wide-band FM*. This statement, based on noise considerations, further confirms a similar observation that was made in Section 4.3 when considering the bandwidth of FM waves.

CAPTURE EFFECT

The inherent ability of an FM system to minimize the effects of unwanted signals (e.g., noise, as discussed above) also applies to *interference* produced by another frequency-modulated signal whose frequency content is close to the carrier frequency of the desired FM wave. However, interference suppression in an FM receiver works well only when the interference is weaker than the desired FM input. When the interference is the stronger one of the two, the receiver locks onto the stronger signal and thereby suppresses the desired FM input. When they are of nearly equal strength, the receiver fluctuates back and forth between them. This phenomenon is known as the *capture effect*, which describes another distinctive characteristic of frequency modulation.

FM THRESHOLD EFFECT

The formula of Eq. (6.40) defining the output signal-to-noise ratio of an FM receiver is valid only if the carrier-to-noise ratio, measured at the discriminator input, is high compared with unity. It is found experimentally that as the input noise power is increased so that the carrier-to-noise ratio is decreased, the FM receiver *breaks*. At first, individual clicks are heard in the receiver output, and as the carrier-to-noise ratio decreases still further, the clicks rapidly merge into a *crackling* or *sputtering sound*. Near the breaking point, Eq. (6.40) begins to fail by predicting values of output signal-to-noise ratio larger than the actual ones. This phenomenon is known as the *threshold effect*.[1] The threshold is defined as the minimum carrier-to-noise ratio yielding an FM improvement that is not significantly deteriorated from the value predicted by the usual signal-to-noise formula assuming a small noise power.

For a qualitative discussion of the FM threshold effect, consider first the case when there is no signal present, so that the carrier wave is unmodulated. Then the composite signal at the frequency discriminator input is

$$x(t) = [A_c + n_I(t)]\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \tag{6.44}$$

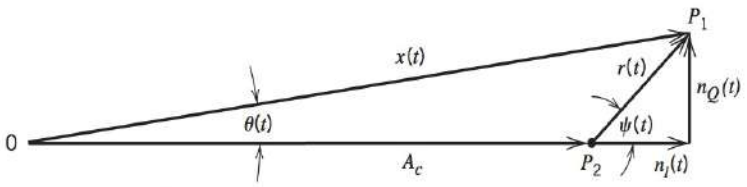


FIGURE 6.10 A Phasor diagram interpretation of Eq. (6.44).

where $n_I(t)$ and $n_Q(t)$ are the in-phase and quadrature components of the narrowband noise $n(t)$ with respect to the carrier wave. The phasor diagram of Figure 6.10 displays the phase relations between the various components of $x(t)$ in Eq. (6.44). As the amplitudes and phases of $n_I(t)$ and $n_Q(t)$ change with time in a random manner, the point P_1 [the tip of the phasor representing $x(t)$] wanders around the point P_2 (the tip of the phasor representing the carrier). When the carrier-to-noise ratio is large, $n_I(t)$ and $n_Q(t)$ are usually much smaller than the carrier amplitude A_c, and so the wandering point P_1 in Figure 6.10 spends most of its time near point P_2. Thus the angle $\theta(t)$ is approximately $n_Q(t)/A_c$ to within a multiple of 2π. When the carrier-to-noise ratio is low, on the other hand, the wandering point P_1 occasionally sweeps around the origin and $\theta(t)$ increases or decreases by 2π radians. Figure 6.11 illustrates how in a rough way the excursions in $\theta(t)$, depicted in Figure 6.11*a*, produce impulselike components in $\theta'(t) = d\theta/dt$. The discriminator output $v(t)$ is equal to $\theta'(t)/2\pi$. These impulselike components have different heights depending on how close the wandering point P_1 comes to the origin O, but all have areas nearly equal to $\pm 2\pi$ radians, as illustrated in Figure 6.11*b*. When the signal shown in Figure 6.11*b* is passed through the postdetection low-pass filter, corresponding but wider impulselike components are excited in the receiver output and are heard as clicks. The clicks are produced only when $\theta(t)$ changes by $\pm 2\pi$ radians.

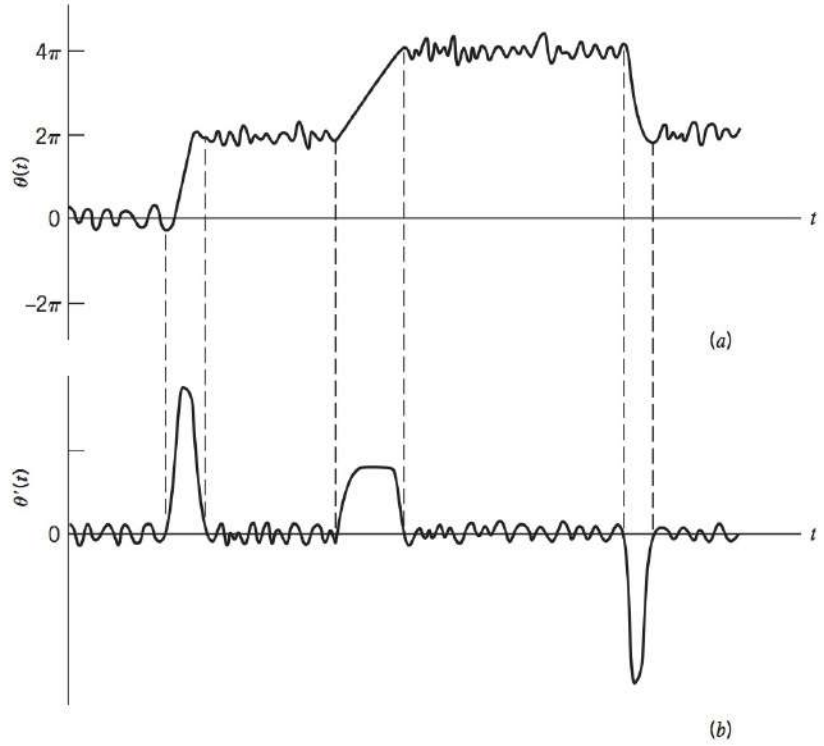


FIGURE 6.11 Illustrating impulselike components in $\theta'(t) = d\theta(t)/dt$ produced by changes of 2π in $\theta(t)$; (*a*) and (*b*) are graphs of $\theta(t)$ and $\theta'(t)$, respectively.

From the phasor diagram of Figure 6.10, we may deduce the conditions required for clicks to occur. A positive-going click occurs when the envelope $r(t)$ and phase $\psi(t)$ of the narrowband noise $n(t)$ satisfy the following conditions:

$$\begin{aligned} r(t) &> A_c \\ \psi(t) < \pi &< \psi(t) + d\psi(t) \\ \frac{d\psi(t)}{dt} &> 0 \end{aligned}$$

These conditions ensure that the phase $\theta(t)$ of the resultant phasor $x(t)$ changes by 2π radians in the time increment dt, during which the phase of the narrowband noise increases by the incremental amount $d\psi(t)$. Similarly, the conditions for a negative-going click to occur are as follows:

$$\begin{aligned} r(t) &> A_c \\ \psi(t) > -\pi &> \psi(t) + d\psi(t) \\ \frac{d\psi(t)}{dt} &< 0 \end{aligned}$$

These conditions ensure that $\theta(t)$ changes by -2π radians during the time increment dt.

To characterize threshold performance, let the *carrier-to-noise ratio* be defined by

$$\rho = \frac{A_c^2}{2B_T N_0} \tag{6.45}$$

As ρ is decreased, the average number of clicks per unit time increases. When this number becomes appreciably large, the threshold is said to occur.

The output signal-to-noise ratio is calculated as follows:

1. The output signal is taken as the receiver output measured in the absence of noise. The average output signal power is calculated assuming a sinusoidal modulation that produces a frequency deviation Δf equal to $B_T/2$, so that the carrier swings back and forth across the entire input frequency band.
2. The average output noise power is calculated when there is no signal present; that is, the carrier is unmodulated, with no restriction imposed on the value of the carrier-to-noise ratio ρ.

On this basis, theory[2] yields Curve I of Figure 6.12 presenting a plot of the output signal-to-noise ratio versus the carrier-to-noise ratio when the ratio $B_T/2W$ is equal to 5. This curve shows that the output signal-to-noise ratio deviates appreciably from a linear function of the carrier-to-noise ratio ρ when ρ is less than about 10 dB. Curve II of Figure 6.12 shows the effect of modulation on the output signal-to-noise ratio when the modulating signal (assumed sinusoidal) and the noise are present at the same time. The average output signal power pertaining to curve II may be taken to be effectively the same as for

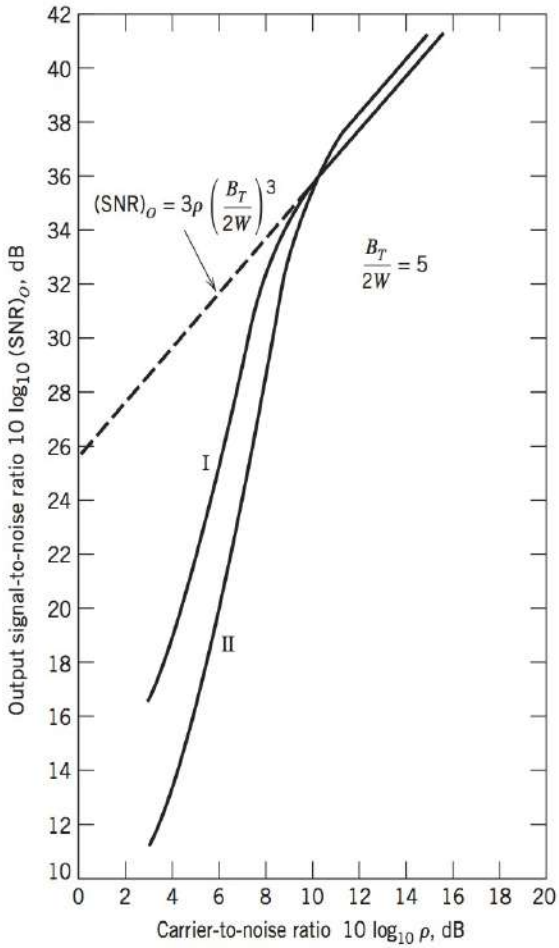


FIGURE 6.12 Dependence of output signal-to-noise ratio on input carrier-to-noise ratio. In curve I, the average output noise power is calculated assuming an unmodulated carrier. In curve II, the average output noise power is calculated assuming a sinusoidally modulated carrier. Both curves I and II are calculated from theory.

curve I. The average output noise power, however, is strongly dependent on the presence of the modulating signal, which accounts for the noticeable deviation of curve II from curve I. In particular, we find that as ρ decreases from infinity, the output signal-to-noise ratio deviates appreciably from a linear function of ρ when ρ is about 11 dB. Also when the signal is present, the resulting modulation of the carrier tends to increase the average number of clicks per second. Experimentally, it is found that occasional clicks are heard in the receiver output at a carrier-to-noise ratio of about 13 dB, which appears to be only slightly higher than what theory indicates. Also it is of interest to note that the enhanced increase in the average number of clicks per second tends to cause the output signal-to-noise ratio to fall off somewhat more sharply just below the threshold level in the presence of modulation.

From the foregoing discussion we may conclude that threshold effects in FM receivers may be avoided in most practical cases of interest if the carrier-to-noise ratio ρ is equal to or greater than 20 or, equivalently, 13 dB. Thus, using Eq. (6.45) we find that the loss of message at the discriminator output is negligible if

$$\frac{A_c^2}{2B_T N_0} \geq 20$$

or, equivalently, if the average transmitted power $A_c^2/2$ satisfies the condition

$$\frac{A_c^2}{2} \geq 20B_T N_0 \tag{6.46}$$

To use this formula, we may proceed as follows:

1. For a specified modulation index β and message bandwidth W, we determine the transmission bandwidth of the FM wave, B_T, using the universal curve of Figure 4.9 or Carson's rule.
2. For a specified average noise power per unit bandwidth, N_0, we use Eq. (6.46) to determine the minimum value of the average transmitted power $A_c^2/2$ that is necessary to operate above threshold.

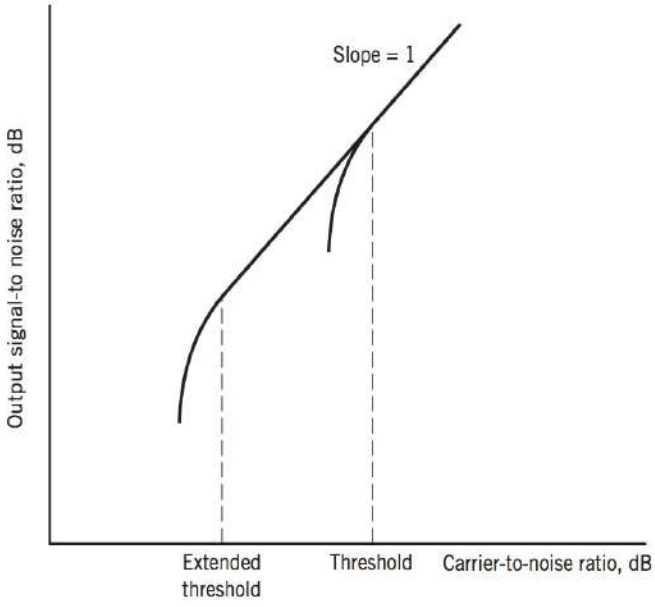


FIGURE 6.13 FM threshold extension.

FM THRESHOLD REDUCTION

In certain applications such as space communications using frequency modulation, there is particular interest in reducing the noise threshold in an FM receiver so as to satisfactorily operate the receiver with the minimum signal power possible. Threshold reduction in FM receivers may be achieved by using an FM demodulator with negative feedback (commonly referred to as an *FMFB demodulator*), or by using a *phase-locked loop demodulator*. Such devices are referred to as *extended-threshold demodulators*, the idea of which is illustrated in Figure 6.13. The threshold extension shown in this figure is measured with respect to the standard frequency discriminator (i.e., one without feedback).

The block diagram of an FMFB demodulator[3] is shown in Figure 6.14. We see that the local oscillator of the conventional FM receiver has been replaced by a voltage-controlled oscillator (VCO) whose instantaneous output frequency is

controlled by the demodulated signal. In order to understand the operation of this receiver, suppose for the moment that the VCO is removed from the circuit and the feedback loop is left open. Assume that a wide-band FM signal is applied to the receiver input, and a second FM signal, from the same source but whose modulation index is a fraction smaller, is applied to the VCO terminal of the mixer. The output of the mixer would consist of the difference frequency component, because the sum frequency component is removed by the band-pass filter. The frequency deviation of the mixer output would be small, although the frequency deviation of both input FM waves is large, since the difference between their instantaneous deviations is small. Hence, the modulation indices would subtract and the resulting FM wave at the mixer output would have a smaller modulation index. The FM wave with reduced modulation index may be passed through a band-pass filter, whose bandwidth need only be a fraction of that required for either wide-band FM, and then frequency demodulated. It is now apparent that the second wide-band FM signal applied to the mixer may be obtained by feeding the output of the frequency discriminator back to the VCO.

It will now be shown that the signal-to-noise ratio of an FMFB receiver is the same as that of a conventional FM receiver with the same input signal and noise power if the carrier-to-noise ratio is sufficiently large. Assume for the moment that there is no feedback around the demodulator. In the combined presence of an unmodulated carrier $A_c \cos(2\pi f_c t)$ and narrowband noise

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)$$

the phase of the composite signal $x(t)$ at the limiter-discriminator input is approximately equal to $n_Q(t)/A_c$, assuming that the carrier-to-noise ratio is high. The envelope of $x(t)$ is of no interest to us, because the limiter removes all variations in the envelope. Thus the composite signal at the frequency discriminator input consists of a small index phase-modulated wave with the modulation derived from the component $n_Q(t)$ of noise that is in phase quadrature with the carrier. When feedback is applied, the VCO generates a frequency-modulated signal that reduces the phase-modulation index of the wave in the band-pass filter output, that is, the quadrature component $n_Q(t)$ of noise. Thus we see that as long as the carrier-to-noise ratio is sufficiently large, the FMFB receiver does not respond to the in-phase noise component $n_I(t)$, but that it would demodulate the quadrature noise component $n_Q(t)$ in exactly the same fashion as it would demodulate signal modulation. Signal and quadrature noise are reduced in the same proportion by the applied feedback, with the result that the baseband signal-to-noise ratio is independent of feedback. For large carrier-to-noise ratios, the baseband signal-to-noise ratio of an FMFB receiver is then the same as that of a conventional FM receiver.

The reason that an FMFB receiver is able to extend the threshold is that, unlike a conventional FM receiver, it uses a very important piece of *a priori* information, namely, that even though the carrier frequency of the incoming FM wave will usually have large

Stephen O. Rice (1907–1986)

Stephen Rice was a member of the technical staff of New Jersey's Bell Laboratories in 1945 when he published the first paper on the effects of noise on analog-communication signals. This classic paper, consisting of three parts, is remarkable for its detailed mathematical exposition of random noise and its properties. It was in Part III of the paper where the distribution of a sinusoidal signal plus Gaussian noise was derived for the first time. The distribution, now known as the Rice distribution, bears his name.

In 1963, Rice made another significant contribution in formulating a heuristic approach to analyze the threshold phenomenon experienced at the output of an FM receiver due to noise. This phenomenon, difficult to analyze in mathematical terms, manifests itself in the increased number of "clicks" that are produced at the output of the FM receiver as the signal-to-noise at its input is reduced. This second contribution is remarkable for the close agreement between theory and practice.

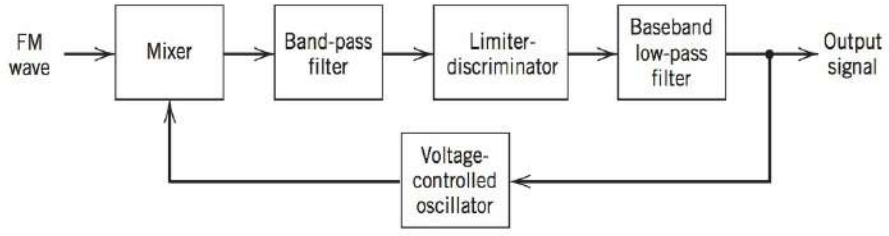


FIGURE 6.14 FMFB demodulator.

frequency deviations, its rate of change will be at the baseband rate. An FMFB demodulator is essentially a *tracking filter* that can track only the slowly varying frequency of a wide-band FM signal, and consequently it responds only to a narrow band of noise centered about the instantaneous carrier frequency. The bandwidth of noise to which the FMFB receiver responds is precisely the band of noise that the VCO tracks. The end result is that an FMFB receiver is capable of realizing a threshold extension on the order of 5–7 dB, which represents a significant improvement in the design of minimum power FM systems.

Like the FMFB demodulator, the phase-locked loop (discussed in Chapter 4) is also a tracking filter and, as such, the noise bandwidth to which it responds is precisely the band of noise tracked by the VCO. Indeed, the phase-locked loop demodulator[4] offers a threshold extension capability with a relatively simple circuit. Unfortunately, the amount of threshold extension is not predictable by any existing theory, and it depends on signal parameters. Roughly speaking, improvement by a few (on the order of 2 to 3) decibels is achieved in typical applications, which is not as good as an FMFB demodulator.

6.6 PRE-EMPHASIS AND DE-EMPHASIS IN FM

In Section 6.5 we showed that the power spectral density of the noise at the output of an FM receiver has a square-law dependence on the operating frequency; this is illustrated in Figure 6.15*a*. In Figure 6.15*b*, we have included the power spectral density of a typical message source; audio and video signals typically have spectra of this form. In particular, we see that the power spectral density of the message usually falls off appreciably at higher frequencies. On the other hand, the power spectral density of the output noise increases rapidly with frequency. Thus, around $f = \pm W$, the relative spectral density of the message is quite low, whereas that of the output noise is quite high in comparison. Clearly, the message is not utilizing the frequency band allotted to it in an efficient manner. It may appear that one way of improving the noise performance of the system is to slightly reduce the bandwidth of the post-detection low-pass filter so as to reject a large amount of noise power while losing only a small amount of message power. Such an approach, however, is usually not satisfactory because the distortion of the message caused by the reduced filter bandwidth, even though slight, may not be tolerable. For example, in the case of music, we find that although the high-frequency notes contribute only a very small fraction of the total power, nonetheless, they contribute a great deal from an esthetic viewpoint.

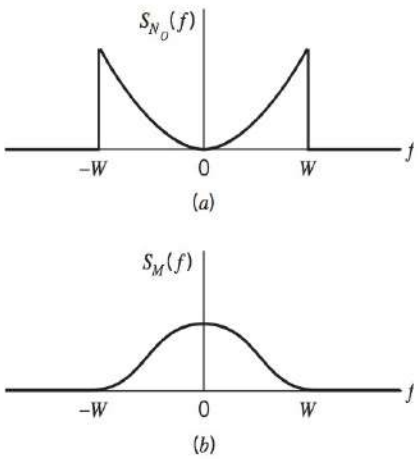


FIGURE 6.15 (*a*) Power spectral density of noise at FM receiver output. (*b*) Power spectral density of a typical message signal.

A more satisfactory approach to the efficient utilization of the allowed frequency band is based on the use of *pre-emphasis* in the transmitter and *de-emphasis* in the receiver, as illustrated in Figure 6.16. In this method, we artificially emphasize the high-frequency components of the message signal prior to modulation in the transmitter, and therefore before the noise is introduced in the receiver. In effect, the low-frequency and high-frequency portions of the power spectral density of the message are equalized in such a way that the message fully occupies the frequency band allotted to it. Then, at the discriminator output in the receiver, we perform the inverse operation by de-emphasizing the high-frequency components, so as to restore the original signal-power distribution of the message. In such a process, the high-frequency components of the noise at the discriminator output are also reduced, thereby effectively increasing the output signal-to-noise ratio of the system. Such a pre-emphasis and de-emphasis process is widely used in commercial FM radio transmission and reception.

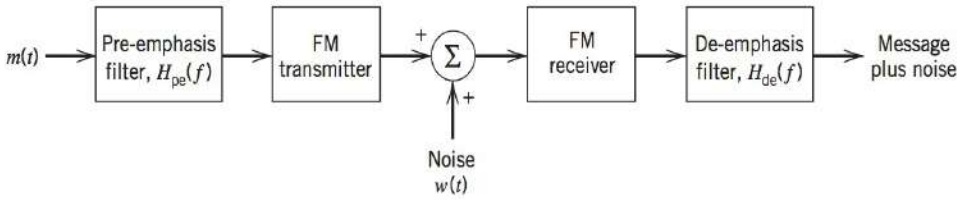


FIGURE 6.16 Use of pre-emphasis and de-emphasis in an FM system.

In order to produce an undistorted version of the original message at the receiver output, the pre-emphasis filter in the transmitter and the de-emphasis filter in the receiver must ideally have transfer functions that are the inverse of each other. That is, if $H_{pe}(f)$ designates the transfer function of the pre-emphasis filter, then the transfer function $H_{de}(f)$ of the de-emphasis filter must ideally be (ignoring transmission delay)

$$H_{de}(f) = \frac{1}{H_{pe}(f)}, \qquad -W \le f \le W \tag{6.47}$$

This choice of transfer functions makes the average message power at the receiver output independent of the pre-emphasis and de-emphasis procedure.

From our previous noise analysis in FM systems, assuming a high carrier-to-noise ratio, the power spectral density of the noise $n_d(t)$ at the discriminator output is

$$S_{N_d}(f) = \begin{cases} \dfrac{N_0 f^2}{A_c^2}, & |f| \le \dfrac{B_T}{2} \\ 0, & \text{otherwise} \end{cases} \tag{6.48}$$

Therefore, the modified power spectral density of the noise at the de-emphasis filter output is equal to $|H_{de}(f)|^2 S_{N_d}(f)$. Recognizing, as before, that the post-detection low-pass filter has a bandwidth W that is, in general, less than $B_T/2$, we find that the average power of the modified noise at the receiver output is as follows:

$$\begin{pmatrix} \text{Average output noise} \\ \text{power with de-emphasis} \end{pmatrix} = \frac{N_0}{A_c^2} \int_{-W}^{W} f^2 |H_{de}(f)|^2 \, df \tag{6.49}$$

Because the average message power at the receiver output is ideally unaffected by the pre-emphasis and de-emphasis procedure, it follows that the improvement in output signal-to-noise ratio produced by the use of pre-emphasis in the transmitter and de-emphasis in the receiver is defined by

$$I = \frac{\text{average output noise power without pre-emphasis and de-emphasis}}{\text{average output noise power with pre-emphasis and de-emphasis}} \tag{6.50}$$

Earlier we showed that the average output noise power without pre-emphasis and de-emphasis is equal to $(2N_0W^3/3A_c^2)$. Therefore, after cancellation of common terms, we may express the improvement factor I as

$$I = \frac{2W^3}{3\displaystyle\int_{-W}^{W} f^2 |H_{de}(f)|^2 \, df} \tag{6.51}$$

It must be emphasized that this improvement factor assumes the use of a high carrier-to-noise ratio at the discriminator input in the receiver.

EXAMPLE 6.3

A simple pre-emphasis filter that emphasizes high frequencies and is commonly used in practice is defined by the transfer function

$$H_{pe}(f) = 1 + \frac{jf}{f_0}$$

This transfer function is closely realized by the RC-amplifier network shown in Figure 6.17*a*, provided that $R \ll r$ and $2\pi fCr \ll 1$ inside the frequency band of interest. The amplifier in Figure 6.17*a* is intended to make up for the attenuation introduced by the RC network at low frequencies. The frequency parameter f_0 is $1/(2\pi Cr)$.

The corresponding de-emphasis filter in the receiver is defined by the transfer function

$$H_{de}(f) = \frac{1}{1 + jf/f_0}$$

which can be realized using the simple RC network of Figure 6.17*b*.

The improvement in output signal-to-noise ratio of the FM receiver, resulting from use of the pre-emphasis and de-emphasis filters of Figure 6.17, is therefore

$$I = \frac{2W^3}{3\displaystyle\int_{-W}^{W} \frac{f^2\,df}{1 + (f/f_0)^2}} = \frac{(W/f_0)^3}{3[(W/f_0) - \tan^{-1}(W/f_0)]} \tag{6.52}$$

In commercial FM broadcasting, we typically have $f_0 = 2.1$ kHz, and we may reasonably assume $W = 15$ kHz. This set of values yields $I = 22$, which corresponds to an improvement of 13 dB in the output signal-to-noise ratio of the receiver. The output signal-to-noise ratio of an FM receiver without pre-emphasis and de-emphasis is typically 40–50 dB. We see, therefore, that by using the simple pre-emphasis and de-emphasis filters shown in Figure 6.17, we can realize a significant improvement in the noise performance of the receiver.

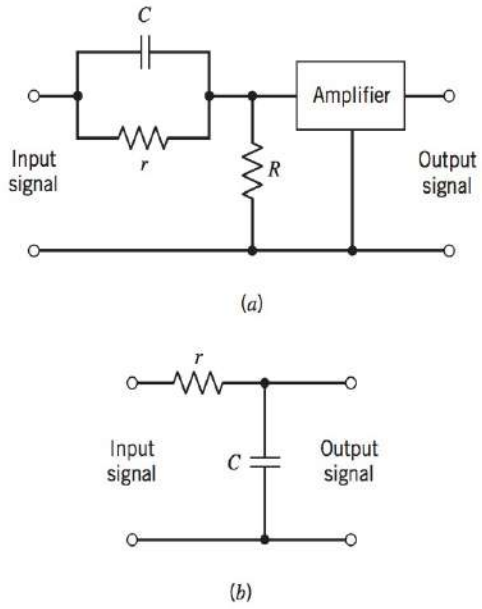


FIGURE 6.17 (*a*) Pre-emphasis filter. (*b*) De-emphasis filter.

The use of the simple *linear* pre-emphasis and de-emphasis filters described above is an example of how the performance of an FM system may be improved by utilizing the differences between characteristics of signals and noise in the system. These simple filters also find application in audio tape-recording. Specifically, *nonlinear* pre-emphasis and de-emphasis techniques have been applied successfully to tape-recording. These techniques[5] (known as *Dolby-A*, *Dolby-B*, and *DBX* systems) use a combination of filtering and dynamic range compression to reduce the effects of noise, particularly when the signal level is low.

6.7 THEME EXAMPLE—LINK BUDGET OF FM SATELLITE LINK

In the previous sections of this chapter we have shown how to determine the output SNR of a demodulator, knowing the SNR at the input to the receiver for a number of modulation strategies. A critical portion of the physical layer design of a communication system is determining the SNR at the input to the receiver. This portion of the communication link is addressed in this section by using a geostationary satellite system as the example.

In a geostationary satellite communication system, a signal is transmitted from an Earth station via an uplink to a satellite, amplified in a transponder (i.e., electronic circuitry) on board the satellite, and then retransmitted from the satellite via a downlink to another earth station, as illustrated in Figure 6.18. The first generation of communication satellites often used the 6 GHz frequency band for the uplink and 4 GHz for the downlink. The use of these frequency bands offers the following advantages:

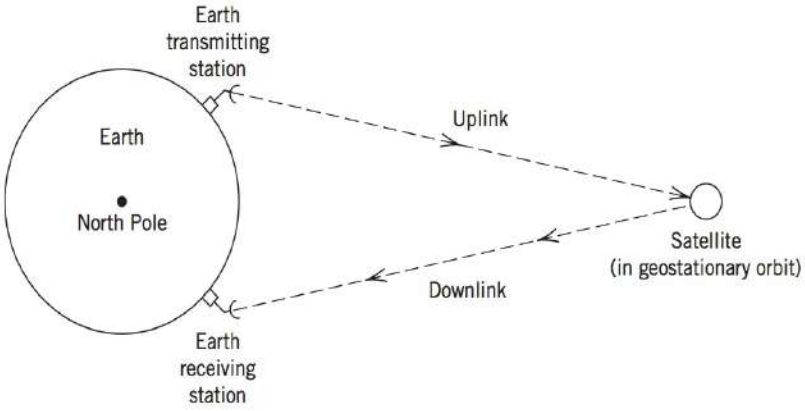


FIGURE 6.18 Satellite communications system.

- Relatively inexpensive microwave equipment.
- Low attenuation due to rainfall—rainfall is the primary atmospheric cause of signal degradation.
- Insignificant sky background noise—the sky background noise (due to random noise emissions from galactic, solar, and terrestrial sources) reaches its lowest level between 1 and 10 GHz.

However, radio interference limits the applications of communication satellites operating in the 6/4 GHz band, because the transmission frequencies of this band coincide with those used for terrestrial microwave systems. This problem is eliminated in the more powerful "second-generation" communication satellites that operate in the 14/12 GHz band. Moreover, the use of these higher frequencies makes it possible to build smaller and therefore less expensive antennas.

The block diagram of Figure 6.19 shows the basic components of a single transponder channel of a simple communication satellite. This structure is sometimes referred to as a "bent-pipe" satellite since it simply amplifies and redirects the incoming signal. More advanced satellites include on-board digital signal processing. Specifically, the

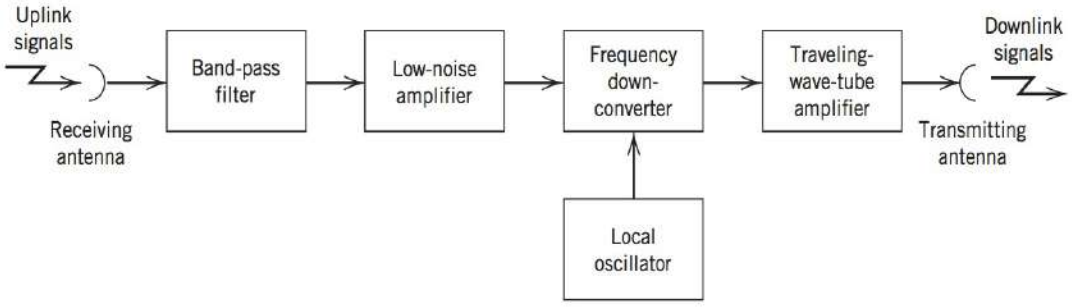


FIGURE 6.19 Block diagram of transponder.

receiving antenna output of the uplink is applied to the cascade connection of the following components:

- Band-pass filter, designed to separate the received signal from among the different radio channels.
- Low-noise amplifier.
- Frequency down-converter, the purpose of which is to convert the received radio frequency (RF) signal to the desired downlink frequency.
- An amplifier, which provides high gain over a wide band of frequencies. Early satellites used traveling-wave tube (TWT) amplifiers. Modern satellites often use solid-state amplifiers.

The channel configuration shown in Figure 6.19 uses a single frequency translation. Other transponder configurations are possible.

RADIO LINK ANALYSIS

An important issue that arises in the design of satellite communication systems is that of link budget analysis. As its name implies, a link budget, or more precisely "link power budget," is the totaling of all the gains and losses incurred in operating a communication link. In particular, the balance sheet constituting the link budget provides a detailed accounting of three broadly defined items:

1. Apportionment of the resources available to the transmitter and the receiver.
2. Sources responsible for the loss of signal power.
3. Sources of noise.

Putting all these items together into the link budget, we end up with an estimation procedure for evaluating the performance of a radio link, which could be the uplink or downlink of a satellite communication system.

THE FRIIS EQUATION

The first step in formulating the link budget is to calculate the received signal power. Radio transmission is characterized by the generation of an electric signal in the transmitter representing the desired information, propagation of corresponding radio waves through space, and a receiver that estimates the transmitted information from the recovered electrical signal. The transmission system is characterized by the antennas that convert between electrical signals and radio waves, and the propagation of the radio waves through space.

For *free-space* transmission, implying there is an unobstructed direct line-of-sight path between the transmitter and the receiver, the received signal power is given by the *Friis equation*

$$P_R = \frac{P_T G_T G_R}{L_p} \tag{6.53}$$

where P_T is the transmitted power and P_R is the received power. The parameters G_T and G_R represent the gains of the transmitting and receiving antennas, respectively. For a short whip antenna, these gains are often close to unity; however, for the parabolic ("dish") antennas often used with satellite communications, the gains can be quite large. The denominator represents the signal attenuation, referred to as the *path loss*, between and transmitter and the receiver. For free-space transmission, the path loss L_p between

two antennas is given by

$$L_p = \left(\frac{4\pi R}{\lambda}\right)^2 \tag{6.54}$$

where R is the distance between the transmitter and the receiver, and λ is the wavelength of the transmission. To simplify the evaluation of the Friis equation, we often write it as the following decibel relation

$$P_R(dB) = P_T(dB) + G_T(dB) + G_R(dB) - L_p(dB) \tag{6.55}$$

where $X(\text{dB}) = 10\log_{10}(X)$. The Friis' equation is the fundamental *link budget* equation. This equation relates received power to transmitted power taking into account transmission characteristics of the radio link. The term $P_T G_T$ is sometimes expressed as

$$\text{EIRP} = P_T G_T \tag{6.56}$$

where EIRP is the *effective radiated power referenced to an isotropic source*. An isotropic source transmits uniformly in all directions, whereas a directional antenna achieves its gain by focusing the power in one direction. So EIRP describes the equivalent power an antenna would radiate if it transmitted uniformly in all directions. There is an implicit assumption, in Eqs. (6.55) and (6.56), that the transmitting and receiving antennas are pointing at each other.

RECEIVER NOISE

To complete the link budget analysis, we need to determine the average noise power in the received signal. As described in Section 5.10, often the major source noise in a transmission is the thermal noise, due to random movement electrons in the receiver front end, where the signal is weakest. This is certainly the case for satellite links, and receiver front ends are designed to include very low noise amplifiers (LNAs). The noise level in satellite receivers is minimized to the extent that the noise contribution of these amplifiers is simply expressed as an equivalent temperate, T_e. Saying a receiver has an equivalent noise temperature of T_e implies that the noise due to the receiver is white over the bandwidth of interest and has a noise density given by

$$N(f) = kT_e. \tag{6.57}$$

Since the noise is white, we may combine it with the received signal to define a quantity known as the carrier-to-noise ratio, C/N_0. The received C/N_0 may be expressed as

$$\frac{C}{N_0} = \frac{P_R}{kT_e} \tag{6.58}$$

The carrier-to-noise ratio is often used in link budgets because it is a generic measure of performance that is independent of the modulation used and the bandwidth. It can equally well be used in both analog and digital systems.

DEMODULATOR SNR

One of the goals of a communication system is to deliver the message with a specified quality. The quality measure, which we have used in this chapter, is the signal-to-noise ratio at the output of the demodulator. Once the C/N_0 is

Noise Figure

An alternative method of expressing the noise contributed by an amplifier or receiver is the *noise figure*. The noise figure F is defined as the *ratio of the total available output noise power per unit bandwidth to the portion thereof solely due to source*. The noise figure is, in general, frequency dependent, being defined mathematically as

$$F = \frac{S_{NO}(f)}{G(f)S_{NS}(f)}$$

where $S_{NO}(f)$ is the available noise power at the device output, $S_{NS}(f)$ is the available noise power from the source at the device input and $G(f)$ is the available device gain. *Available power* implies the maximum power that may be delivered to an external load when the source and load impedances are matched. This definition recognizes that the output noise of a two-port device is made of two contributions, one due to the source and other due to the device itself.

In many instances, the source and device noise are white and the device gain is approximately constant over the frequency range of interest, and consequently F is a constant, and the noise power in a fixed bandwidth B may be expressed as

$$N_0 B = FkT_0 B$$

where kT_0 is the thermal noise power density at nominal temperature T_0. Thus the noise density added by the device is $(F-1)kT_0$ which is equal to kT_e where T_e is the *equivalent noise temperature*.

known, we may use the results of this chapter for AM and FM modulation (and in later chapters for digital modulation) to determine the output SNR. Knowing the required quality, we can determine the signal-to-noise ratio required at the input to the demodulator, and knowing the channel bandwidth, the corresponding carrier-to-noise density ratio, C/N_0.

EXAMPLE SATELLITE LINK

To illustrate the concepts introduced under radio-link analysis, consider the following problem. A satellite transponder with 30 MHz bandwidth relays a television signal using FM modulation. The downlink EIRP of the satellite is 32 dBW. Is this sufficient power to provide video SNR of 36 dB at the receiver output? Assume the receiving earth station has a system noise temperature of 100° K and uses a parabolic-dish antenna having a gain of 52 dB.

We begin solving this problem by using the Friis equation to compute the received signal strength; the only unknown in Eq. (6.55) is the path loss, which is given by Eq. (6.54). For a geostationary satellite, the range from the satellite to the Earth station may range from 36,000 km when the satellite is directly overhead to 41,000 km when the satellite is at 10° elevation. We assume the worst case range of 41,000 km. We also assume that transmission is at 4 GHz, as would be typical of a first generation satellite using FM transmission. Substituting these assumed values into Eq. (6.54) with $\lambda = c/f$, we find that the path loss is

$$\begin{aligned} L_p &= \left(\frac{4\pi R}{\lambda}\right)^2 \\ &= \left(\frac{4\pi(4.1\times 10^7)}{(3\times 10^8/4\times 10^9)}\right)^2 \\ &= 4.7\times 10^{19} \\ &\approx 196.7\ \text{dB} \end{aligned} \tag{6.59}$$

Substituting this result into the Friis equation [Eq. (6.55)], we find that the received power, expressed in decibels, is

$$\begin{aligned} P_R &= 32 + 52 - 196.7 \\ &= -112.7\ \text{dBW} \end{aligned} \tag{6.60}$$

To determine the received carrier-to-noise density ratio, we must combine P_R with the receiver noise according to Eq. (6.58) to obtain, in decibels,

$$\begin{aligned} \frac{C}{N_0}(\text{dBHz}^{-1}) &= P_R(\text{dBW}) - 10\log_{10}k(\text{dBWHz}^{-1}\text{K}^{-1}) - 10\log_{10}T_e(\text{dBK}) \\ &= -112.7 + 228.6 - 20 \\ &= 95.9\,\text{dBHz}^{-1} \end{aligned} \tag{6.61}$$

Knowing the received C/N_0, we use the FM equation [Eq. (6.40)] to compute the video SNR at the output of the FM demodulator, obtaining

$$\begin{aligned} (\text{SNR})_{O,\text{FM}} &= \frac{3k_f^2 P}{W^3}\left(\frac{A_c^2}{2N_0}\right) \\ &\simeq \frac{3D^2}{W}\left(\frac{C}{N_0}\right) \end{aligned} \tag{6.62}$$

where the deviation ratio D is given by [see the discussion following Eq. (6.42)]

$$D = \frac{\Delta f}{W} \approx \frac{k_f P^{1/2}}{W} \tag{6.63}$$

TABLE 6.1 Link budget for FM satellite downlink

Parameter	Units	Value
EIRP	dBW	32
Free-space loss	dB	196.7
Receiver antenna gain	dB	52
System noise temperature	dBK	20
Boltzmann constant	$\text{dBWHz}^{-1}\text{K}^{-1}$	−228.6
Received C/N_0	dBHz^{-1}	95.9
Demodulated video SNR	dB	41.5
Required video SNR	dB	36

To estimate the deviation ratio we recall from Section 3.6 that the baseband bandwidth of an analog television signal is $W = 4.5$ MHz (including the audio component). Furthermore, the satellite transponder bandwidth of 30 MHz limits the transmission bandwidth to $B_T = 30$ MHz. Using Carson's rule of Eq. (4.38) with f_m replaced by W (see the discussion preceding Example 4.4) we determine the peak deviation Δf from the formula

$$B_T = 2(W + \Delta f) \tag{6.64}$$

to be 10.5 MHz. Thus $D = \Delta f/W = 2.33$. Substituting D and W into Eq. (6.62), we obtain

$$\begin{aligned}(\text{SNR})_{O,\text{FM}} &= 10 \log_{10}\left(\frac{3D^2}{W}\right) + 95.9 \\ &= 41.5 \text{ dB}\end{aligned} \tag{6.65}$$

as the video SNR. We may summarize these calculations in the link budget of Table 6.1.

The link budget indicates that, for the given satellite EIRP, the demodulated video SNR would be 5.5 dB larger than required. This excess power or SNR is often referred to as *margin*. In practice, there are several factors that will decrease the margin available. Some of these factors include: (a) losses due to non-ideal implementation of the demodulator; (b) contributions to the noise due to the uplink portion of the transmission; and (c) losses due to atmospheric attenuation of the signal. In a more detailed link budget, these factors would be included.

Needless to say, the essence of the communication link analysis presented in this section also applies to other radio links. It is for this reason that the treatment of radio link analysis presented in this section is of a generic nature. In practice, the design is usually iterative with the output SNR being given as a requirement, and tradeoffs are then made between satellite EIRP, antenna size, and system noise temperature to achieve this requirement.

6.8 SUMMARY AND DISCUSSION

We conclude the noise analysis of CW modulation systems by presenting a comparison of the relative merits of the different modulation techniques. For this comparison, we assume that the modulation is produced by a sinusoidal wave. For the comparison to be meaningful, we also assume that all the different modulation systems operate with exactly the same channel signal-to-noise ratio. In making the comparison, it is informative to keep in mind the transmission bandwidth requirement of the modulation system in question. In this regard, we use a *normalized transmission bandwidth* defined by

$$B_n = \frac{B_T}{W} \tag{6.66}$$

where B_T is the transmission bandwidth of the modulated signal and W is the message bandwidth. We may thus make the following observations:

1. In a full AM system using envelope detection, the output signal-to-noise ratio, assuming sinusoidal modulation, is given by [see Eq. (6.20)]

$$(\text{SNR})_O = \frac{\mu^2}{2+\mu^2}(\text{SNR})_C$$

This relation is shown plotted as curve I in Figure 6.20, assuming $\mu = 1$. In this curve we have also included the AM threshold effect. Since in a full AM system both sidebands are transmitted, the normalized transmission bandwidth B_n equals 2.

2. In the case of a DSB–SC modulation system using coherent detection, the output signal-to-noise ratio is given by [see Eq. (6.10)]

$$(\text{SNR})_O = (\text{SNR})_C$$

This relation is shown plotted as curve II in Figure 6.20. We see, therefore, that the noise performance of a DSB–SC system, using coherent detection, is superior by 4.8 dB to that of a full AM system using envelope detection. It should also be noted that the DSB–SC system does not exhibit a threshold effect.

3. In an FM system using a conventional discriminator, the output signal-to-noise ratio, assuming sinusoidal modulation, is given by [see Eq. (6.43)]

$$(\text{SNR})_O = \frac{3}{2}\beta^2(\text{SNR})_C$$

where β is the modulation index. This relation is shown plotted as curves III and IV in Figure 6.20, corresponding to $\beta = 2$ and $\beta = 5$, respectively. In each case, we have included a 13-dB improvement, which is typically obtained by using pre-emphasis in the transmitter and de-emphasis in the receiver as described in Section 6.6. To determine the transmission bandwidth requirement, we use the universal curve of Figure 4.9 and so find that

$$B_n = 8 \quad \text{for } \beta = 2$$
$$B_n = 16 \quad \text{for } \beta = 5$$

Thus we see that, compared with the DSB–SC system, by using wide-band FM, we obtain an improvement in output signal-to-noise ratio equal to 20.8 dB for a normalized bandwidth $B_n = 8$, and an improvement of 28.8 dB for $B_n = 16$. This clearly illustrates the improvement in noise performance that is achievable by using wide-band FM. However, the price that we have to pay for this improvement is excessive transmission bandwidth. It is, of course, assumed that the FM system operates above threshold for the noise improvement to be realizable.

An important point to conclude the discussion is that, unlike amplitude modulation, frequency modulation offers the capability to trade off transmission bandwidth for improved noise performance. The trade-off follows a square law, which is the best that we can do with CW modulation (i.e., analog communications). In the next chapter we describe pulse-code modulation, which is

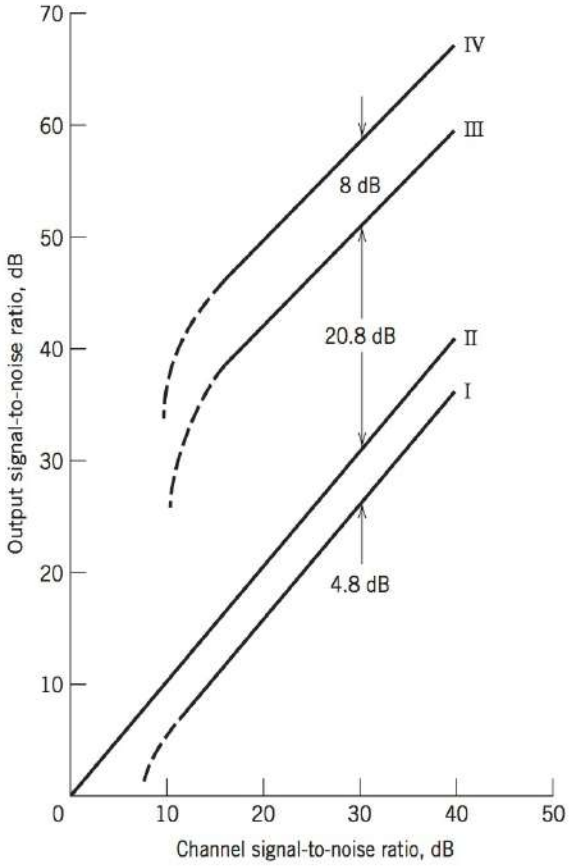


FIGURE 6.20 Comparison of the noise performance of various CW modulation systems. Curve I: Standard AM, $\mu = 1$. Curve II: DSB–SC. Curve III: FM, $\beta = 2$. Curve IV: FM, $\beta = 5$. (Curves III and IV include 13-dB pre-emphasis, de-emphasis improvement.)

basic to the transmission of analog information-bearing signals by a digital communication system, and which can indeed do better.

NOTES AND REFERENCES

1. For a detailed discussion of the threshold effect in FM receivers, see Rice (1963) and Schwartz, Bennett, and Stein (1966, pp. 129–163).
2. Figure 6.12 is adapted from Rice (1963). The validity of the theoretical curve II in this figure has been confirmed experimentally; see Schwartz, Bennett, and Stein (1966, p. 153).
3. The treatment of the FMFB demodulator presented in Section 6.5 is based on the paper by Enloe (1962); see also Roberts (1977, pp. 166–181).
4. For a full discussion of threshold effects in phase-locked loops, see Gardner (1979, pp. 178–196) and Roberts (1977, pp. 200–202).
5. For a detailed description of Dolby systems mentioned in the latter part of Section 6.7, see Stremler (1990, pp. 732–734).

PROBLEMS

6.1 The sample function

$$x(t) = A_c \cos(2\pi f_c t) + w(t)$$

is applied to the low-pass RC filter of Figure P6.1. The amplitude A_c and frequency f_c of the sinusoidal component are constants, and $w(t)$ is a white Gaussian noise of zero mean and power spectral density $N_0/2$. Find an expression for the output signal-to-noise ratio with the sinusoidal component of $x(t)$ regarded as the signal of interest.

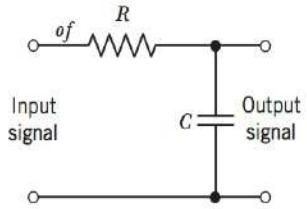


Figure P6.1

6.2 Suppose next the sample function $w(t)$ of Problem 6.1 is applied to the band-pass LCR filter of Figure P6.2, which is tuned to the frequency f_c of the sinusoidal component. Assume that the Q-factor of the filter is high compared with unity. Find an expression for the output signal-to-noise ratio by treating the sinusoidal component of $x(t)$ as the signal of interest.

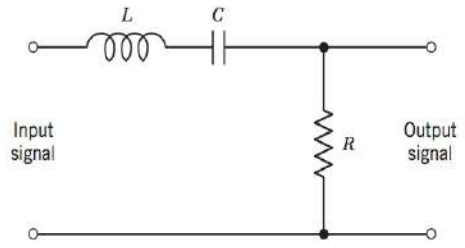


Figure P6.2

6.3 A DSB–SC modulated signal is transmitted over a noisy channel, with the power spectral density of the noise being as shown in Figure P6.3. The message bandwidth is 4 kHz and the carrier frequency is 200 kHz. Assuming that the average power of the modulated wave is 10 watts, determine the output signal-to-noise ratio of the receiver.

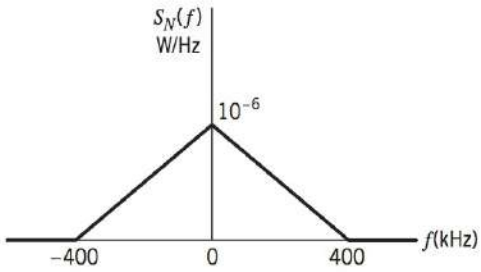


Figure P6.3

6.4 Evaluate the autocorrelation functions and cross-correlation functions of the in-phase and quadrature components of the narrowband noise at the coherent detector input for the DSB–SC system.

6.5 In a receiver using coherent detection, the sinusoidal wave generated by the local oscillator suffers from a phase error $\theta(t)$ with respect to the carrier wave $\cos(2\pi f_c t)$. Assuming that $\theta(t)$ is a sample function of a zero-mean Gaussian process of variance σ_Θ^2, and that most of the time the maximum value *of* $\theta(t)$ is small compared with unity, find the mean-square error of the receiver output for DSB–SC modulation. The mean-square error is defined as the expected value of the squared difference between the receiver output and the message signal component of the receiver output.

6.6 The average noise power per unit bandwidth measured at the front end of an AM receiver is 10^{-3} watt per Hertz. The modulating wave is sinusoidal, with a carrier power of 80 kilowatts, and a sideband power of 10 kilowatts per sideband. The message bandwidth is 4 kHz. Assuming the use of an envelope detector in the receiver, determine the output signal-to-noise ratio of the system. By how many decibels is this system inferior to a DSB–SC modulation system?

6.7 Consider the output of an envelope detector defined by Eq. (6.14), which is reproduced here for convenience:

$$y(t) = \{[A_c + A_c k_a m(t) + n_I(t)]^2 + n_Q^2(t)\}^{1/2}$$

(a) Assume that the probability of the event

$$|n_Q(t)| > \varepsilon A_c |1 + k_a m(t)|$$

is equal to or less than δ_1, where $\varepsilon \ll 1$. What is the probability that the effect of the quadrature component $n_Q(t)$ is negligible?

(b) Suppose that k_a is adjusted relative to the message signal $m(t)$ such that the probability of the event

$$A_c[1 + k_a m(t)] + n_I(t) < 0$$

is equal to δ_2. What is the probability that the approximation

$$y(t) \simeq A_c[1 + k_a m(t)] + n_I(t)$$

is valid?

(c) Comment on the significance of the result in part (b) for the case when δ_1 and δ_2 are both small compared with unity.

6.8 An unmodulated carrier of amplitude A_c and frequency f_c and band-limited white noise are summed and then passed through an ideal envelope detector. Assume the noise spectral density to be of height $N_0/2$ and bandwidth $2W$, centered about the carrier frequency f_c. Determine the output signal-to-noise ratio for the case when the carrier-to-noise ratio is high.

6.9 An AM receiver, operating with a sinusoidal modulating signal and 80 percent modulation, has an output signal-to-noise ratio of 30 dB.

(a) What is the corresponding carrier-to-noise ratio?

(b) By how many decibels can we decrease the carrier-to-noise ratio so that the system is operating just above threshold?

6.10 Consider a phase modulation (PM) system, with the modulated wave defined by

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)]$$

where k_p is a constant and $m(t)$ is the message signal. The additive noise $n(t)$ at the phase detector input is

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)$$

Assuming that the carrier-to-noise ratio at the detector input is high compared with unity, determine (a) the output signal-to-noise ratio and (b) the figure of merit of the system. Compare your results with the FM system for the case of sinusoidal modulation.

6.11 An FDM system uses single-sideband modulation to combine 12 independent voice signals and then uses frequency modulation to transmit the composite baseband signal. Each voice signal has an average power P and occupies the frequency band 0.3–3.4 kHz; the system allocates it a bandwidth of 4 kHz. For each voice signal, only the lower sideband is transmitted. The sub-carrier waves used for the first stage of modulation are defined by

$$c_k(t) = A_k \cos(2\pi k f_0 t), \qquad 0 \le k \le 11$$

The received signal consists of the transmitted FM signal plus white Gaussian noise of zero mean and power spectral density $N_0/2$.

(a) Sketch the power spectral density of the signal produced at the frequency discriminator output, showing both the signal and noise components.

(b) Find the relationship between the subcarrier amplitudes A_k so that the modulated voice signals have equal signal-to-noise ratios.

6.12 In the discussion on FM threshold effect presented in Section 6.5, we described the conditions for positive-going and negative-going clicks in terms of the envelope $r(t)$ and phase $\psi(t)$ of the narrowband noise $n(t)$. Reformulate these conditions in terms of the in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ of $n(t)$.

6.13 By using the pre-emphasis filter shown in Figure 6.17*a* and with a voice signal as the modulating wave, an FM transmitter produces a signal that is essentially frequency-modulated by the lower audio frequencies and phase-modulated by the higher audio frequencies. Explain the reasons for this phenomenon.

6.14 Suppose that the transfer functions of the pre-emphasis and de-emphasis filters of an FM system are scaled as follows:

$$H_{pe}(f) = k\left(1 + \frac{jf}{f_0}\right)$$

and

$$H_{de}(f) = \frac{1}{k}\left(\frac{1}{1 + jf/f_0}\right)$$

The scaling factor k is to be chosen so that the average power of the emphasized message signal is the same as that of the original message signal $m(t)$.

(a) Find the value of k that satisfies this requirement for the case when the power spectral density of the message signal $m(t)$ is

$$S_M(f) = \begin{cases} \dfrac{S_0}{1 + (f/f_0)^2}, & -W \le f \le W \\ 0, & \text{elsewhere} \end{cases}$$

(b) What is the corresponding value of the improvement factor I produced by using this pair of pre-emphasis and de-emphasis filters? Compare this ratio with that obtained in Example 6.3. The improvement factor I is defined by Eq. (6.50).

6.15 A phase modulation (PM) system uses a pair of pre-emphasis and de-emphasis filters defined by the transfer functions

$$H_{pe}(f) = 1 + \frac{jf}{f_0}$$

and

$$H_{de}(f) = \frac{1}{1 + (jf/f_0)}$$

Show that the improvement in output signal-to-noise ratio produced by using this pair of filters is

$$I = \frac{W/f_0}{\tan^{-1}(W/f_0)}$$

where W is the message bandwidth. Evaluate this improvement for the case when $W = 15$ kHz and $f_0 = 2.1$ kHz, and compare your result with the corresponding value for an FM system.

Computer Problems

6.16 Using the Matlab model of an AM modulator and envelope detector developed in Problem 3.25, do the following:

(a) Simulate the modulation and detection of a 400 Hz modulating wave having a modulation index of 50 percent. Keep a copy of this noise-free demodulator output as a reference.

(b) Add narrowband noise to the signal prior to demodulation such that the channel SNR is 30 dB.

(c) Compare the output of the demodulator with the noisy input to the noise free output. Compute the mean-square error between the two and estimate the output SNR using this.

Repeat the problem for channel SNRs of 20 dB and 10 dB. At what channel SNR does envelope detection break down?

6.17 Using the Matlab model of an FM modulator and discriminator developed in Problems 4. 26 and 4.28, do the following:

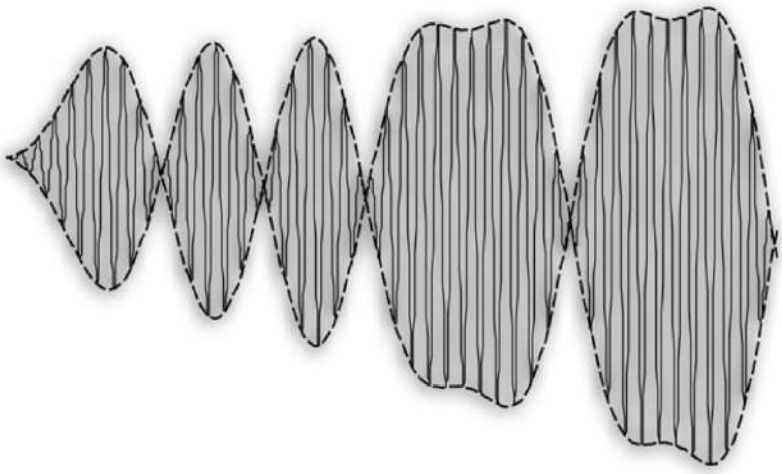
(a) Simulate the modulation and detection of a 1 kHz modulating wave having a modulation index of 2. Keep a copy of this noise-free demodulator output as a reference.

(b) Add narrowband noise to the signal prior to demodulation such that the channel SNR is 30 dB.

(c) Compare the output of the demodulator with the noisy input to the noise free output. Compute the mean-square error between the two and estimate the output SNR using this.

Repeat for channel SNRs of 20, 15 and 10 dB. At what channel SNR does frequency discrimination break down?

7



THE TRANSITION FROM ANALOG TO DIGITAL

7.1 INTRODUCTION

In continuous-wave (CW) modulation, which we studied in Chapters 3 and 4, some parameter of a sinusoidal carrier wave is varied continuously in accordance with the message signal. The amplitude and angle forms of CW modulation were originally developed in the early 1900s with analog sources, usually voice, in mind. As the understanding of digital transmission improved through the period from the 1930s to the 1960s, there came the recognition that digital transmission has many advantages over the transmission of analog information. However, it was not until the developments in solid state electronics, micro-electronics and large-scale integration in the 1970s that the right capabilities existed to take advantage of digital transmission in an efficient and economical fashion.

The first step in this evolution from analog to digital transmission is the conversion of common information sources, such as voice and music, which are inherently analog to digital representation. This analog to digital conversion and the representation of the analog information as a sequence of pulses is the focus of this chapter.

In the first step from analog to digital, an analog source is sampled at discrete times. The resulting analog samples are then transmitted by means of analog pulse modulation. Consequently, the chapter begins with a description of the sampling process and this is followed by a discussion of pulse-amplitude modulation, which is the simplest form of analog pulse modulation. This is followed by a description of pulse-position modulation, which is another important form of pulse modulation.

In the second step from analog to digital, an analog source is not only sampled at discrete times but the samples themselves are also quantized to discrete levels. Hence, we begin this second step with a discussion of quantization. We then describe two methods of digitally representing an analog source: pulse-code modulation and delta modulation.

Historically, the conversion from an analog information source, such as voice or video, to a digital representation and subsequent transmission, was often implemented as a single step. This is in contrast to the modern layered approach to communications where the different aspects of the communication process are clearly separated. As a result of this history, the nomenclature may be confusing at times. For example, pulse-code modulation describes both a method of digitally representing an analog source and a method for transmitting that information over a baseband channel. For the purposes of this chapter, we consider the described techniques as methods of digitally representing analog sources. These digital representations can then be used with a variety of modulation techniques, not just baseband transmission.

7.2 WHY DIGITIZE ANALOG SOURCES?

The introduction to this chapter alluded to various advantages that the transmission of digital information has over analog. Many of these advantages are described in technical detail in the following chapters but we will explain them briefly here:

- Digital systems are less sensitive to noise than analog. For long transmission lengths, the signal may be regenerated effectively error-free at different points along the path, and the original signal transmitted over the remaining length.
- With digital systems, it is easier to integrate different services, for example, video and the accompanying soundtrack, into the same transmission scheme.
- The transmission scheme can be relatively independent of the source. For example, a digital transmission scheme that transmits voice at 10 kbps could also be used to transmit computer data at 10 kbps.
- Circuitry for handling digital signals is easier to repeat and digital circuits are less sensitive to physical effects such as vibration and temperature.
- Digital signals are simpler to characterize and typically do not have the same amplitude range and variability as analog signals. This makes the associated hardware easier to design.

While almost all media (e.g., cable, radio waves, optical fiber) may be used for either analog or digital transmission, digital techniques offer strategies for more efficient use of those media:

- Various media sharing strategies, known as multiplexing techniques, are more easily implemented with digital transmission strategies.
- There are techniques for removing redundancy from a digital transmission, so as to minimize the amount of information that has to be transmitted. These techniques fall under the broad classification of source coding and we discuss some of these techniques in Chapter 10.
- There are techniques for adding controlled redundancy to a digital transmission, such that errors that occur during transmission may be corrected at the receiver without any additional information. These techniques fall under the general category of channel coding, which is described in Chapter 10. As an example, a forward error-correcting technique that is relatively simple by today's standards can reduce an error rate of 7 percent at the channel output to as little as 0.001 percent at the decoder output.
- Digital techniques make it easier to specify complex standards that may be shared on a worldwide basis. This allows the development of communication components with

many different features (e.g., a cellular handset) and their interoperation with a different component (e.g., a base station) produced by a different manufacturer.

- Other channel compensations techniques, such as equalization, especially adaptive versions, are easier to implement with digital transmission techniques.

It should be emphasized that the majority of these advantages for digital transmission rely on the availability of low-cost microelectronics. This counterbalances the original advantage for analog transmission of transporting a large amount of information in a very simple fashion.

7.3 THE SAMPLING PROCESS

Much of the material on the representation of signals and systems covered up to this stage in the book has been devoted to signals and systems that are continuous in both time and frequency. At various points in Chapter 2, however, we did consider the representation of periodic signals. In particular, referring to Eq. (2.88), we see that the Fourier transform of a periodic signal with period T_0 consists of an infinite sequence of delta functions occurring at integer multiples of the fundamental frequency $f_0 = 1/T_0$. On the basis of this observation, we may state that making a signal periodic in the time domain has the effect of sampling the spectrum of the signal in the frequency domain. We may go one step further by invoking the duality property of the Fourier transform, and thus make the observation that sampling a signal in the time domain has the effect of making the spectrum of the signal periodic in the frequency domain. This latter issue is the subject of this section.

The *sampling process* is usually described in the time domain. As such, it is an operation that is basic to digital signal processing and digital communications. Through use of the sampling process, an analog signal is converted into a corresponding sequence of samples that are usually spaced uniformly in time. Clearly, for such a procedure to have practical utility, it is necessary that we choose the sampling rate properly, so that the sequence of samples uniquely defines the original analog signal. This is the essence of the sampling theorem, which is derived in what follows.

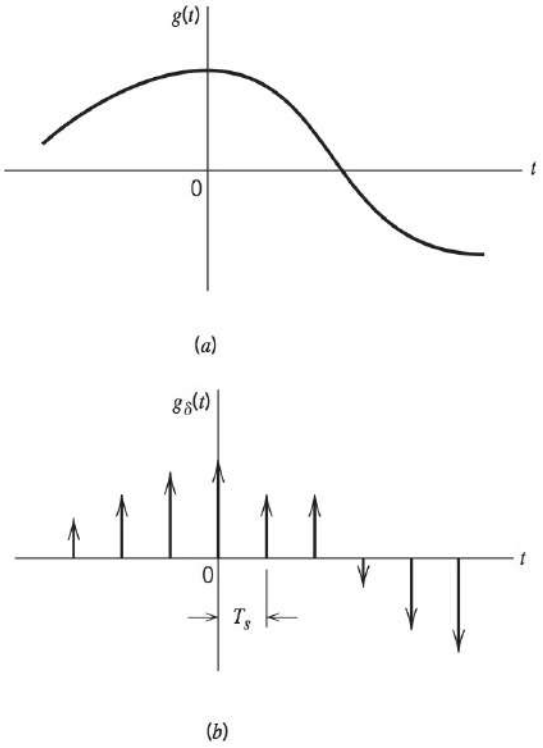


FIGURE 7.1 The sampling process. (*a*) Analog signal, (*b*) instantaneously sampled version of the signal.

Consider an arbitrary signal $g(t)$ of finite energy, which is specified for all time. A segment of the signal $g(t)$ is shown in Figure 7.1*a*. Suppose that we sample the signal $g(t)$ instantaneously and at a uniform rate, once every T_s seconds. Consequently, we obtain an infinite sequence of samples spaced T_s seconds apart and denoted by $\{g(nT_s)\}$, where n takes on all possible integer values. We refer to T_s as the *sampling period*, and to its reciprocal $f_s = 1/T_s$ as the *sampling rate*. This ideal form of sampling is called *instantaneous sampling*.

Let $g_\delta(t)$ denote the signal obtained by individually weighting the elements of a periodic sequence of Dirac delta functions spaced T_s seconds apart by the sequence of numbers $\{g(nT_s)\}$, as shown by (see Figure 7.1*b*)

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\delta(t - nT_s) \tag{7.1}$$

We refer to $g_\delta(t)$ as the *ideal sampled signal*. The term $\delta(t - nT_s)$ represents a delta function positioned at time $t = nT_s$. From the definition of the delta function presented in Chapter 2, we recall that such an idealized function has unit area. We may therefore view the multiplying factor $g(nT_s)$ in Eq. (7.1) as a "mass" assigned to the delta function $\delta(t - nT_s)$. A delta function weighted in this manner is closely approximated by a

rectangular pulse of duration Δt and amplitude $g(nT_s)/\Delta t$; the smaller we make Δt, the better will be the approximation.

The ideal sampled signal $g_\delta(t)$ has a mathematical form similar to that of the Fourier transform of a periodic signal. This is readily established by comparing Eq. (7.1) for $g_\delta(t)$ with the Fourier transform of a periodic signal given in Eq. (2.88). This correspondence suggests that we may determine the Fourier transform of the ideal sampled signal $g_\delta(t)$ by applying the duality property of the Fourier transform to the transform pair of Eq. (2.88). By so doing, and using the fact that a delta function is an even function of time, we get the result:

$$g_\delta(t) \rightleftharpoons f_s \sum_{m=-\infty}^{\infty} G(f - mf_s) \tag{7.2}$$

where $G(f)$ is the Fourier transform of the original signal $g(t)$, and f_s is the sampling rate. Equation (7.2) states that *the process of uniformly sampling a continuous-time signal of finite energy results in a periodic spectrum with a period equal to the sampling rate.*

Another useful expression for the Fourier transform of the ideal sampled signal $g_\delta(t)$ may be obtained by taking the Fourier transform of both sides of Eq. (7.1) and noting that the Fourier transform of the delta function $\delta(t - nT_s)$ is equal to $\exp(-j2\pi fT_s)$. Let $G_\delta(f)$ denote the Fourier transform of $g_\delta(t)$. We may therefore write

$$G_\delta(f) = \sum_{n=-\infty}^{\infty} g(nT_s)\exp(-j2\pi nfT_s) \tag{7.3}$$

This relation is called the *discrete-time Fourier transform* and was briefly discussed in Chapter 2. It may be viewed as a complex Fourier series representation of the periodic frequency function $G_\delta(f)$, with the sequence of samples $\{g(nT_s)\}$ defining the coefficients of the expansion.

The relations, as derived here, apply to any continuous-time signal $g(t)$ of finite energy and infinite duration. Suppose, however, that the signal $g(t)$ is *strictly band-limited*, with no frequency components higher than W Hertz. That is, the Fourier transform $G(f)$ of the signal $g(t)$ has the property that $G(f)$ is zero for $|f| \geq W$, as illustrated in Figure 7.2*a*; the shape of the spectrum shown in this figure is intended for the purpose of illustration only. Suppose also that we choose the sampling period $T_s = 1/2W$. Then the corresponding spectrum $G_\delta(f)$ of the sampled signal $g_\delta(t)$ is as shown in Figure 7.2*b*. Putting $T_s = 1/2W$ in Eq. (7.3) yields

$$G_\delta(f) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\exp\left(-\frac{j\pi nf}{W}\right) \tag{7.4}$$

From Eq. (7.2), we readily see that the Fourier transform of $g_\delta(t)$ may also be expressed as

$$G_\delta(f) = f_s G(f) + f_s \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} G(f - mf_s) \tag{7.5}$$

Hence, under the following two conditions:

1. $G(f) = 0$ for $|f| \geq W$
2. $f_s = 2W$

we find from Eq. (7.5) that

$$G(f) = \frac{1}{2W} G_\delta(f), \qquad -W < f < W \tag{7.6}$$

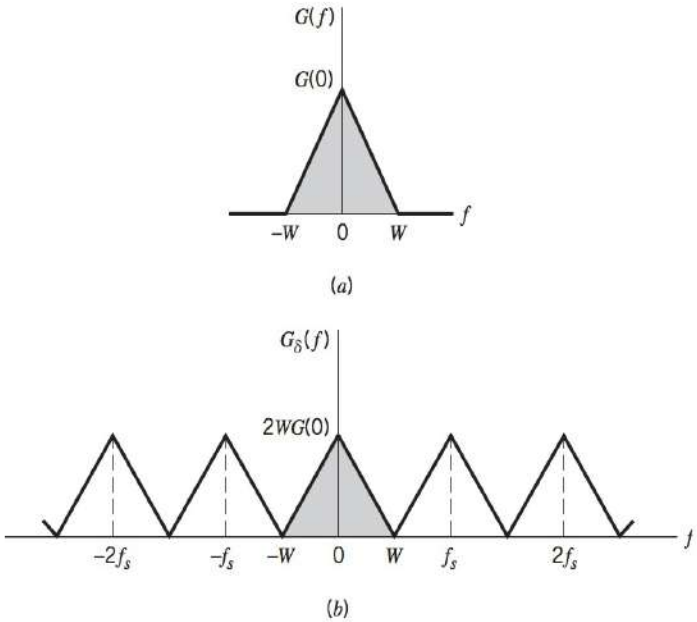


FIGURE 7.2 (a) Spectrum of a strictly band-limited signal $g(t)$. (*b*) Spectrum of sampled version of $g(t)$ for a sampling period $T_s = 1/2W$.

Substituting Eq. (7.4) in Eq. (7.6), we may also write

$$G(f) = \frac{1}{2W}\sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\exp\left(-\frac{j\pi nf}{W}\right), \qquad -W < f < W \tag{7.7}$$

Therefore, if the sample values $g(n/2W)$ of a signal $g(t)$ are specified for all time, then the Fourier transform $G(f)$ of the signal is uniquely determined by using the discrete-time Fourier transform of Eq. (7.7). Because $g(t)$ is related to $G(f)$ by the inverse Fourier transform, it follows that the signal $g(t)$ is itself uniquely determined by the sample values $g(n/2W)$ for $-\infty < n < \infty$. In other words, the sequence $\{g(n/2W)\}$ has all the information contained in $g(t)$.

Consider next the problem of reconstructing the signal $g(t)$ from the sequence of sample values $[g(n/2W)]$. Substituting Eq. (7.7) in the formula for the inverse Fourier transform defining $g(t)$ in terms of $G(f)$, we get

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} G(f)\exp(j2\pi ft)\,df \\ &= \int_{-W}^{W} \frac{1}{2W}\sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\exp\left(-\frac{j\pi nf}{W}\right)\exp(j2\pi ft)\,df \end{aligned}$$

Interchanging the order of summation and integration:

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\frac{1}{2W}\int_{-W}^{W}\exp\left[j2\pi f\left(t-\frac{n}{2W}\right)\right]df \tag{7.8}$$

The integral term in Eq. (7.8) is readily evaluated, yielding the final result

$$\begin{aligned} g(t) &= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\frac{\sin(2\pi Wt - n\pi)}{(2\pi Wt - n\pi)} \\ &= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\mathrm{sinc}(2Wt - n), \qquad -\infty < t < \infty \end{aligned} \tag{7.9}$$

Equation (7.9) provides an *interpolation formula* for reconstructing the original signal $g(t)$ from the sequence of sample values $\{g(n/2W)\}$, with the sinc function $\mathrm{sinc}(2Wt)$ playing the role of an *interpolation function*. Each sample is multiplied by a delayed version of the interpolation function, and all the resulting waveforms are added to obtain $g(t)$. Looking at Eq. (7.9) in another way, it represents the convolution (or filtering) of the impulse train $g_\delta(t)$ given by Eq. (7.1) with the impulse response $\mathrm{sinc}(2Wt)$. Consequently, any impulse response that plays the same role as $\mathrm{sinc}(2Wt)$ is also referred to as a *reconstruction filter*.

We may now state the *sampling theorem* for strictly band-limited signals of finite energy in two equivalent parts:

1. *A band-limited signal of finite energy, which only has frequency components less than W Hertz, is completely described by specifying the values of the signal at instants of time separated by 1/2W seconds.*
2. *A band-limited signal of finite energy, which only has frequency components less than W Hertz, may be completely recovered from a knowledge of its samples taken at the rate of 2W samples per second.*

The sampling rate of $2W$ samples per second, for a signal bandwidth of W Hertz, is called the *Nyquist rate*; its reciprocal $1/2W$ (measured in seconds) is called the *Nyquist interval*.

The Whittakers, Father and Son

The exact origin of the sampling theorem has an intriguing history of its own. The earliest and most highly cited paper is that of E. T. Whittaker, published in 1915. In that paper, Whittaker described an idea that he termed the *cardinal function*, which was subsequently, in 1929 renamed the *cardinal series* by his son, J. M. Whittaker. In his 1915 paper, the senior Whittaker showed (among other findings) that if a function of time is band-limited, then the cardinal series is applicable to that function.

The *sampling theorem*, under that very name, is mentioned (perhaps for the first time) in Shannon's 1949 paper on information theory. For the derivation of the theorem, the reader is referred to another Shannon paper written in 1949 on "Communication in the presence of noise." In this latter paper, Shannon does make reference to a book by J. M. Whittaker on *Interpolation Function Theory*, published in 1935.

For a more detailed account of the history of the sampling theorem, see Chapter 1 of the book by Marks (1991), which, interestingly enough, is entitled *Introduction to Shannon Sampling and Interpolation Theory*.

The derivation of the sampling theorem, as described herein, is based on the assumption that the signal $g(t)$ is strictly band limited. In practice, however, an information-bearing signal is *not* strictly band limited, with the result that some degree of undersampling is encountered. Consequently, some *aliasing* is produced by the sampling process. Aliasing refers to the phenomenon of a high frequency component in the spectrum of the signal seemingly taking on the identity of a lower frequency in the spectrum of its sampled version, as illustrated in Figure 7.3. The aliased spectrum shown by the solid curve in Figure 7.3*b* pertains to an "undersampled" version of the message signal represented by the spectrum of Figure 7.3*a*. To combat the effects of aliasing in practice, we may use two corrective measures, as described here:

1. Prior to sampling, a low-pass *pre-alias filter* is used to attenuate those high-frequency components of the signal that are not essential to the information being conveyed by the signal.
2. The filtered signal is sampled at a rate slightly higher than the Nyquist rate.

The use of a sampling rate higher than the Nyquist rate also has the beneficial effect of easing the design of the *reconstruction filter* used to recover the original signal from its sampled version. Consider the example of a message signal that has been pre-alias (low-pass) filtered, resulting in the spectrum shown in Figure 7.4*a*. The corresponding spectrum of the instantaneously sampled version of the signal is shown in Figure 7.4*b*, assuming a sampling rate higher than the Nyquist rate. According to Figure 7.4*b*, we readily see that the design of the reconstruction filter may be specified as follows (see Figure 7.4*c*):

- The reconstruction filter is low-pass with a passband extending from $-W$ to W, which is itself determined by the pre-alias filter.
- The filter has a transition band extending (for positive frequencies) from W to $f_s - W$, where f_s is the sampling rate.

The fact that the reconstruction filter has a well-defined transition band means that it is physically realizable. This is to be compared to the implementation of the ideal reconstruction filter corresponding to $\text{sinc}(2Wt)$ that would be necessary if the signal was not oversampled.

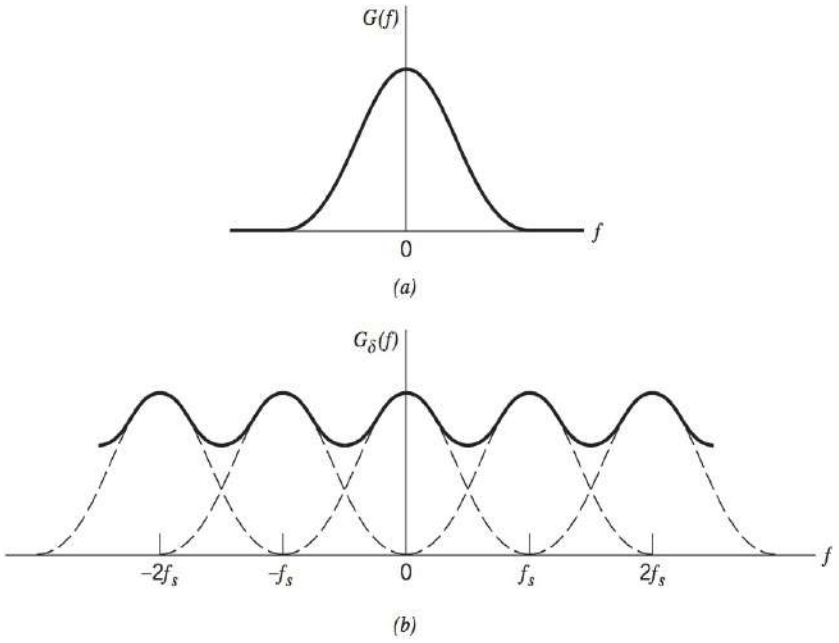


FIGURE 7.3 (*a*) Spectrum of a signal, (*b*) spectrum of an undersampled version of the signal exhibiting the aliasing phenomenon.

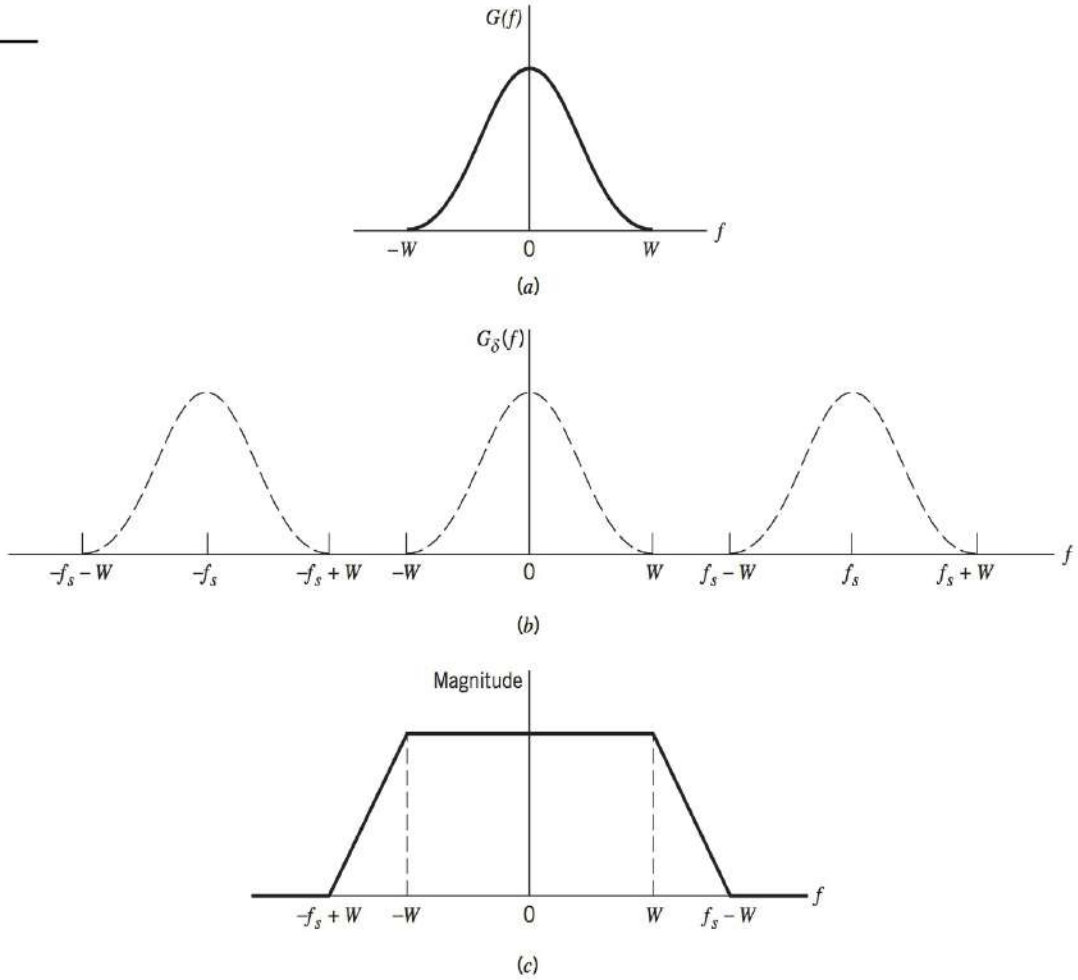


FIGURE 7.4 (*a*) Pre-alias filtered spectrum of an information-bearing signal. (*b*) Spectrum of instantaneously sampled version of the signal, assuming the use of a sampling rate greater than the Nyquist rate. (*c*) Amplitude response of reconstruction filter.

7.4 PULSE-AMPLITUDE MODULATION

Now that we understand the essence of the sampling process, we are ready to formally define pulse-amplitude modulation, which is the simplest and most basic form of analog pulse modulation. In *pulse-amplitude modulation* (PAM), *the amplitudes of regularly spaced pulses are varied in proportion to the corresponding sample values of a continuous message signal*; the pulses can be of a rectangular form or some other appropriate shape. Pulse-amplitude modulation as defined here is somewhat similar to natural sampling, where the message signal is multiplied by a periodic train of rectangular pulses. However, in natural sampling the top of each modulated rectangular pulse varies with the message signal, whereas in PAM it is maintained flat; natural sampling is explored further in Problem 7.1.

The waveform of a PAM signal is illustrated in Figure 7.5. The dashed curve in this figure depicts the waveform of a message signal $m(t)$, and the sequence of amplitude-modulated rectangular pulses shown as solid lines represents the corresponding PAM signal $s(t)$. There are two operations involved in the generation of the PAM signal:

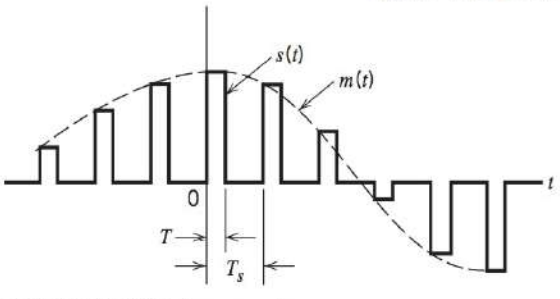


FIGURE 7.5 Flat-top samples.

1. *Instantaneous sampling* of the message signal $m(t)$ every T_s seconds, where the sampling rate $f_s = 1/T_s$ is chosen in accordance with the sampling theorem.
2. *Lengthening* the duration of each sample so obtained to some constant value T.

In digital circuit technology, these two operations are jointly referred to as "sample and hold." One important reason for intentionally lengthening the duration of each sample is to avoid the use of an excessive channel bandwidth, since bandwidth is inversely proportional to pulse duration. However, care has to be exercised in how long we make the sample duration T, as the following analysis reveals.

Let $s(t)$ denote the sequence of flat-top pulses generated in the manner described in Figure 7.5. Hence, we may express the PAM signal as

$$s(t) = \sum_{n=-\infty}^{\infty} m(nT_s)h(t - nT_s) \tag{7.10}$$

where T_s is the *sampling period* and $m(nT_s)$ is the sample value of $m(t)$ obtained at time $t = nT_s$. The $h(t)$ is a standard rectangular pulse of unit amplitude and duration T, defined as follows (see Figure 7.6*a*):

$$h(t) = \begin{cases} 1, & 0 < t < T \\ \dfrac{1}{2}, & t = 0,\ t = T \\ 0, & \text{otherwise} \end{cases} \tag{7.11}$$

By definition, the instantaneously sampled version of $m(t)$ is given by

$$m_\delta(t) = \sum_{n=-\infty}^{\infty} m(nT_s)\delta(t - nT_s) \tag{7.12}$$

where $\delta(t - nT_s)$ is a time-shifted delta function. Therefore, convolving $m_\delta(t)$ with the pulse $h(t)$, we get

$$\begin{aligned} m_\delta(t) \star h(t) &= \int_{-\infty}^{\infty} m_\delta(\tau)h(t - \tau)\, d\tau \\ &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} m(nT_s)\delta(\tau - nT_s)h(t - \tau)\, d\tau \\ &= \sum_{n=-\infty}^{\infty} m(nT_s) \int_{-\infty}^{\infty} \delta(\tau - nT_s)h(t - \tau)\, d\tau \end{aligned} \tag{7.13}$$

Using the sifting property of the delta function, we thus obtain

$$m_\delta(t) \star h(t) = \sum_{n=-\infty}^{\infty} m(nT_s)h(t - nT_s) \tag{7.14}$$

From Eqs. (7.10) and (7.14) it thus follows that the PAM signal $s(t)$ is mathematically equivalent to the convolution of $m_\delta(t)$, the instantaneously sampled version of $m(t)$, and the pulse $h(t)$, as shown by

$$s(t) = m_\delta(t) \star h(t) \tag{7.15}$$

Taking the Fourier transform of both sides of Eq. (7.15) and recognizing that the convolution of two time functions is transformed into the multiplication of their respective Fourier transforms, we get

$$S(f) = M_\delta(f)H(f) \tag{7.16}$$

where $S(f) = F[s(t)], M_\delta(f) = F[m_\delta(t)]$, and $H(f) = F[h(t)]$ as illustrated in Figures 7.7*a*, *b*, and *c*. From Eq. (7.2) we note that the Fourier transform $M_\delta(f)$ is related to the Fourier transform $M(f)$ of the original message signal $m(t)$ as follows:

$$M_\delta(f) = f_s \sum_{k=-\infty}^{\infty} M(f - kf_s) \tag{7.17}$$

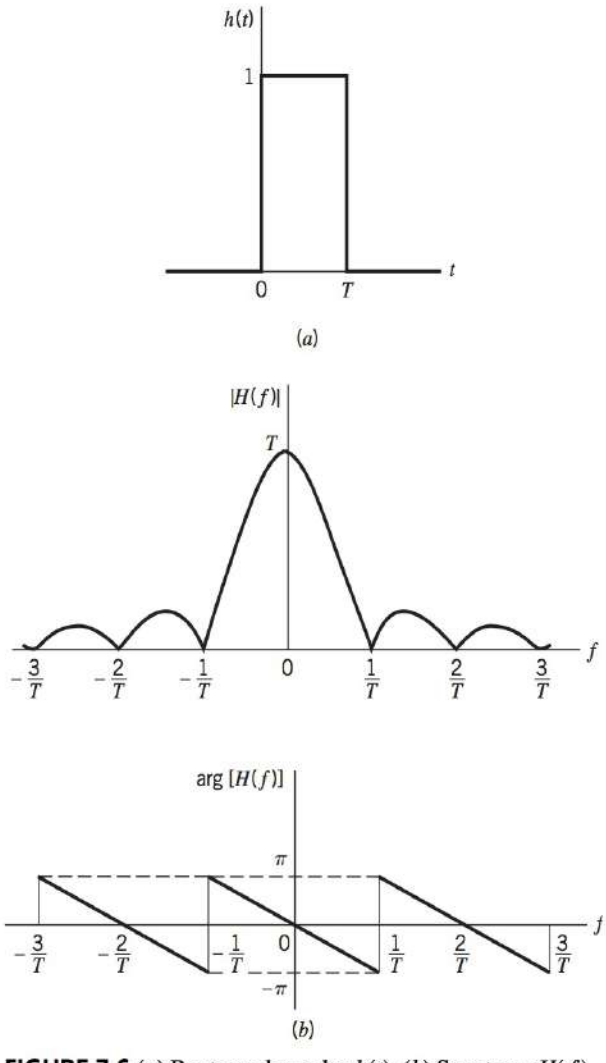


FIGURE 7.6 (*a*) Rectangular pulse $h(t)$. (*b*) Spectrum $H(f)$.

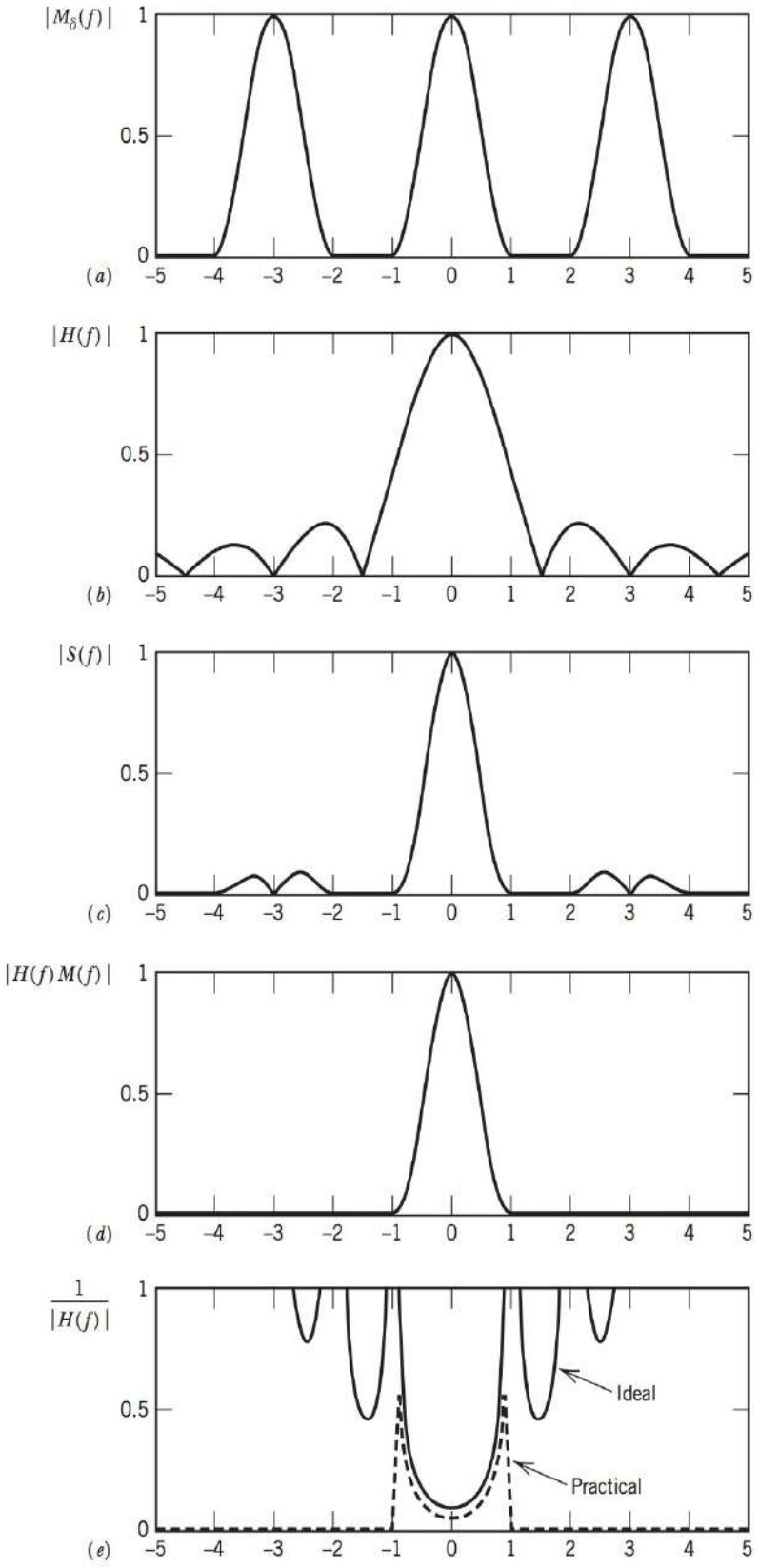


FIGURE 7.7 (*a*) Spectrum of sampled signal. (*b*) Spectrum of low-pass filter. (*c*) Transmitted spectrum. (*d*) Spectrum after receiver filtering. (*e*) Equalizer spectrum.

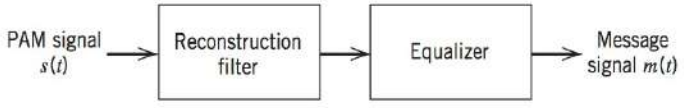


FIGURE 7.8 Recovering $m(t)$ from PAM signal $s(t)$.

where $f_s = 1/T_s$ is the sampling rate. Therefore, the substitution of Eq. (6.17) in (7.16) yields

$$S(f) = f_s \sum_{k=-\infty}^{\infty} M(f - kf_s)H(f) \tag{7.18}$$

Given a PAM signal $s(t)$ whose Fourier transform $S(f)$ is as defined in Eq. (7.18), how do we recover the original message signal $m(t)$? As a first step in this reconstruction, we may pass $s(t)$ through a low-pass filter whose frequency response is defined in Figure 7.4*c*; here it is assumed that the message signal is limited to bandwidth W and the sampling rate f_s is larger than the Nyquist rate $2W$. Then, from Eq. (7.18) we find that the spectrum of the resulting filter output is equal to $M(f)H(f)$. This output is equivalent to passing the original message signal $m(t)$ through another low-pass filter of transfer function $H(f)$ as illustrated in Figure 7.7*d*.

From Eq. (7.11) we note that the Fourier transform of the rectangular pulse $h(t)$ is given by

$$H(f) = T\text{sinc}(fT)\exp(-j\pi fT) \tag{7.19}$$

which is shown plotted in Figure 7.6*b*. We see therefore that by using flat-top samples to generate a PAM signal, we have introduced *amplitude distortion* as well as a *delay* of $T/2$. This effect is rather similar to the variation in transmission with frequency that is caused by the finite size of the scanning aperture in television. Accordingly, the distortion caused by the use of pulse-amplitude modulation to transmit an analog information-bearing signal is referred to as the *aperture effect*.

This distortion may be corrected by connecting an *equalizer* in cascade with the low-pass reconstruction filter, as shown in Figure 7.8. The equalizer has the effect of decreasing the in-band loss of the reconstruction filter as the frequency increases in such a manner as to compensate for the aperture effect. Ideally, the amplitude response of the equalizer is given by

$$\frac{1}{|H(f)|} = \frac{1}{T\text{sinc}(fT)} = \frac{\pi f}{\sin(\pi fT)}$$

as illustrated in Figure 7.7*e*. This figure displays the ideal form of equalization needed for the example being considered herein. The dashed curve included in this figure illustrates the framework for a practical equalizer.

Some final remarks are in order: The transmission of a PAM signal imposes rather stringent requirements on the amplitude and phase responses of the channel, because of the relatively short duration of the transmitted pulses. Furthermore, the noise performance of a PAM system can never be better than baseband-signal transmission. Accordingly, we find that for transmission over long distances, PAM would be used only as a means of message processing for time-division multiplexing, from which conversion to some other form of pulse modulation is subsequently made. The concept of time-division multiplexing is discussed in the next section.

7.5 TIME-DIVISION MULTIPLEXING

The sampling theorem provides the basis for transmitting the information contained in a band-limited message signal $m(t)$ as a sequence of samples of $m(t)$ taken uniformly at a rate that is usually slightly higher than the Nyquist rate. An important feature of the sampling process is a *conservation of time*. That is, the transmission of the message samples engages the communication channel for only a fraction of the sampling interval on a periodic basis, and in this way some of the time interval between adjacent samples is

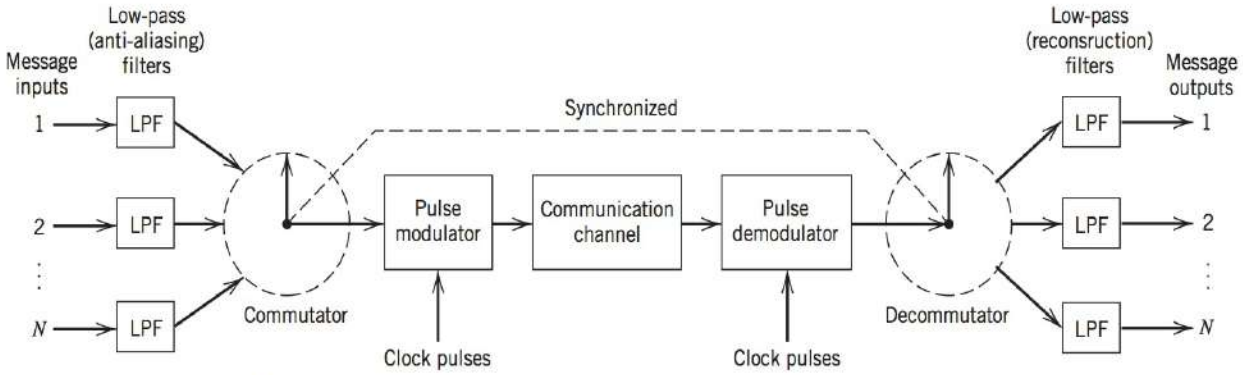


FIGURE 7.9 Block diagram of TDM system.

cleared for use by other independent message sources on a time-shared basis. We thereby obtain a *time-division multiplex* (TDM) *system*, which enables the joint utilization of a common communication channel by a plurality of independent message sources without mutual interference among them.

The concept of TDM is illustrated by the block diagram shown in Figure 7.9. Each input message signal is first restricted in bandwidth by a low-pass pre-alias filter to remove the frequencies that are nonessential to an adequate signal representation. The low-pass filter outputs are then applied to a *commutator*, which is usually implemented using electronic switching circuitry. The function of the commutator is twofold: (1) to take a narrow sample of each of the N input messages at a rate f_s that is slightly higher than $2W$, where W is the cutoff frequency of the pre-alias filter, and (2) to sequentially interleave these N samples inside the sampling interval T_s. Indeed, this latter function is the essence of the time-division multiplexing operation. Following the commutation process, the multiplexed signal is applied to a *pulse modulator*, the purpose of which is to transform the multiplexed signal into a form suitable for transmission over the common channel. It is clear that the use of time-division multiplexing introduces a bandwidth expansion factor N, because the scheme must squeeze N samples derived from N independent message sources into a time slot equal to one sampling interval. At the receiving end of the system, the received signal is applied to a *pulse demodulator*, which performs the reverse operation of the pulse modulator. The narrow samples produced at the pulse demodulator output are distributed to the appropriate low-pass reconstruction filters by means of a *decommutator*, which operates in *synchronism* with the commutator in the transmitter. This synchronization is essential for a satisfactory operation of the system. The way this synchronization is implemented depends naturally on the method of pulse modulation used to transmit the multiplexed sequence of samples.

7.6 PULSE-POSITION MODULATION

In a pulse modulation system, we may use the increased bandwidth consumed by pulses to obtain an improvement in noise performance by representing the sample values of the message signal by some property of the pulse other than amplitude. *In pulse-duration modulation (PDM), the samples of the message signal are used to vary the duration of the individual pulses.* This form of modulation is also referred to as *pulse-width modulation* or *pulse-length modulation*. The modulating signal may vary the time of occurrence of the leading edge, the trailing edge, or both edges of the pulse. In Figure 7.10*c* the trailing edge of each pulse is varied in accordance with the message signal, assumed to be sinusoidal as shown in Figure 7.10*a*. The periodic pulse carrier is shown in Figure 7.10*b*.

In PDM, long pulses expend considerable power during the pulse while bearing no additional information. If this unused power is subtracted from PDM, so that only time transitions are preserved, we obtain a more efficient type of pulse modulation known as *pulse-position modulation* (PPM). In *PPM, the position of a pulse relative to its unmodulated time of occurrence is varied in accordance with the message signal*, as illustrated in Figure 7.10*d* for the case of sinusoidal modulation.

Let T_s denote the sample duration. Using the sample $m(nT_s)$ of a message signal $m(t)$ to modulate the position of the nth pulse, we obtain the PPM signal

$$s(t) = \sum_{n=-\infty}^{\infty} g(t - nT_s - k_p m(nT_s)) \tag{7.20}$$

where k_p is the *sensitivity* of the pulse-position modulator and $g(t)$ denotes a standard pulse of interest. Clearly, the different pulses constituting the PPM signal $s(t)$ must be *strictly non-overlapping*; a sufficient condition for this requirement to be satisfied is to have

$$g(t) = 0, \qquad |t| > \frac{T_s}{2} - k_p|m(t)|_{\max} \tag{7.21}$$

which, in turn, requires that

$$k_p|m(t)|_{\max} < \frac{T_s}{2} \tag{7.22}$$

The closer $k_p|m(t)|_{\max}$ is to one half the sampling duration T_s, the narrower must the standard pulse $g(t)$ be in order to ensure that the individual pulses of the PPM signal $s(t)$ do not interfere with each other, and the wider will the bandwidth occupied by the PPM signal be. Assuming that Eq. (7.21) is satisfied, and that there is no interference between adjacent pulses of the PPM signal $s(t)$, then the signal samples $m(nT_s)$ can be recovered perfectly. Furthermore, if the message signal $m(t)$ is strictly band limited, it follows from the sampling theorem that the original message signal $m(t)$ can be recovered from the PPM signal $s(t)$ without distortion.

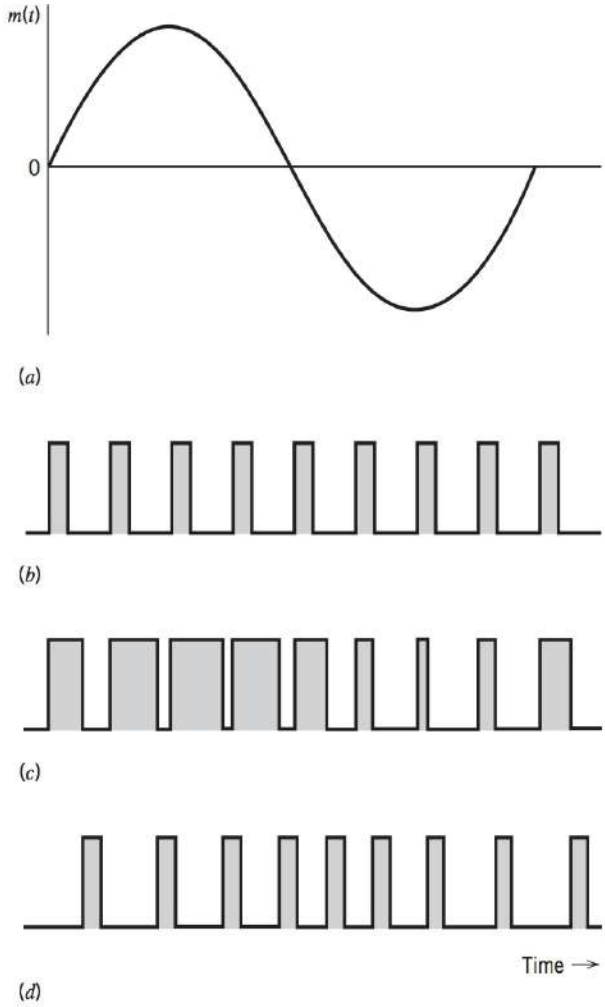


FIGURE 7.10 Illustrating two different forms of pulse-time modulation for the case of a sinusoidal modulating wave. (*a*) Modulating wave. (*b*) Pulse carrier. (*c*) PDM wave. (*d*) PPM wave.

GENERATION OF PPM WAVES

The PPM signal described by Eq. (7.20) may be generated using the system described in Figure 7.11. The message signal $m(t)$ is first converted into a PAM signal by means of a

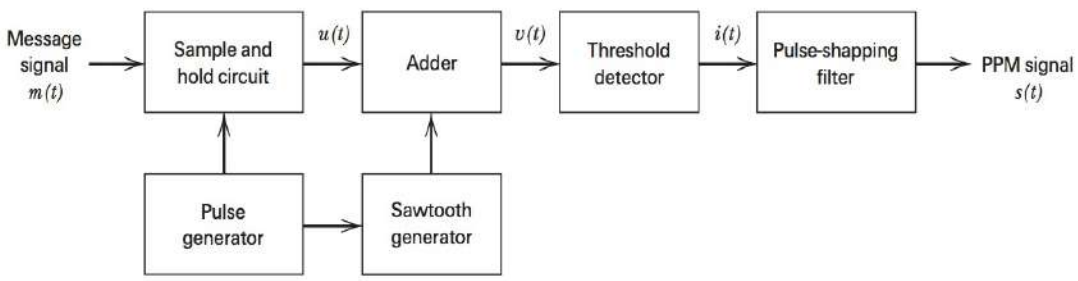


FIGURE 7.11 Block diagram of PPM generator.

sample-and-hold circuit, generating a staircase waveform $u(t)$. Note that the pulse duration T of the sample-and-hold circuit is the same as the sampling duration T_s. This operation is illustrated in Figure 7.12*b* for the message signal $m(t)$ shown in Figure 7.12*a*. Next, the signal $u(t)$ is added to a sawtooth wave (shown in Figure 7.12*c*), yielding the combined signal $v(t)$ shown in Figure 7.12*d*. The combined signal $v(t)$ is applied to a *threshold detector* that produces a very narrow pulse (approximating an impulse) each time $v(t)$ crosses zero in the negative-going direction. The resulting sequence of "impulses" $i(t)$ is shown in Figure 7.12*e*. Finally, the PPM signal $s(t)$ is generated by using this sequence of impulses to excite a filter whose impulse response is defined by the standard pulse $g(t)$.

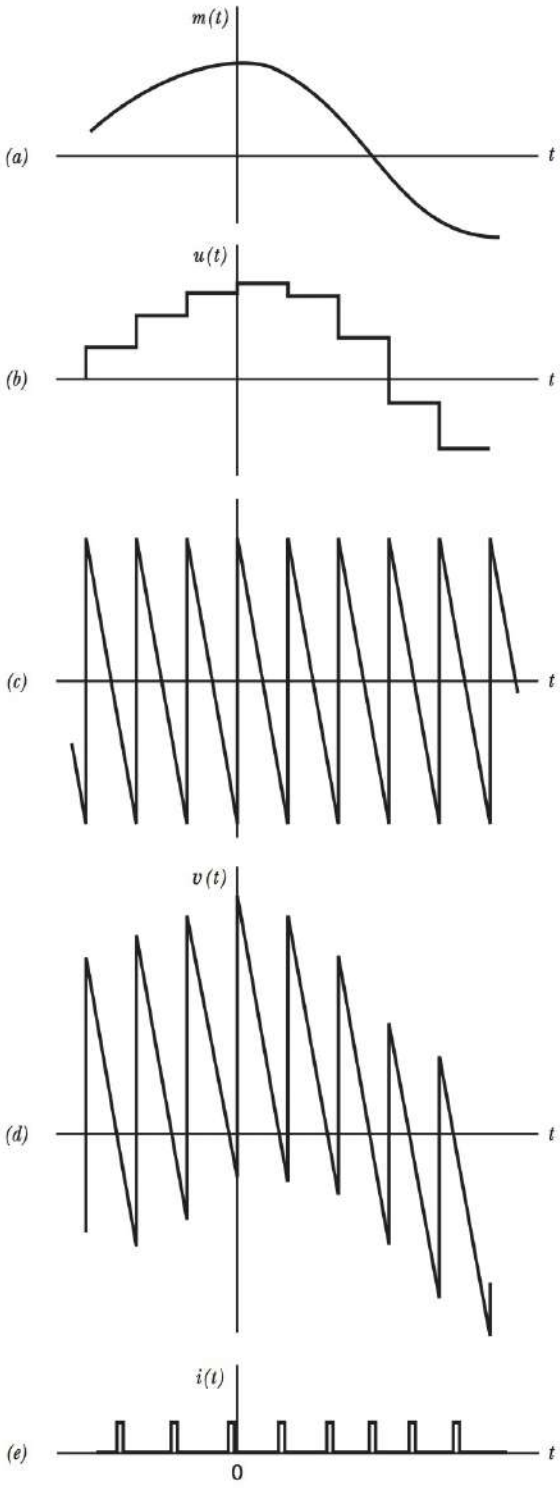


FIGURE 7.12 Generation of PPM signal. (*a*) Message signal. (*b*) Staircase approximation of the message signal. (*c*) Sawtooth wave. (*d*) Composite wave obtained by adding (*b*) and (*c*). (*e*) Sequence of "Impulses" used to generate the PPM signal.

DETECTION OF PPM WAVES

Consider a PPM wave $s(t)$ with uniform sampling, as defined by Eqs. (7.20) and (7.21), and assume that the message (modulating) signal $m(t)$ is strictly band-limited. The operation of one type of PPM receiver may proceed as follows:

- Convert the received PPM wave into a PDM wave with the same modulation.
- Integrate this PDM wave using a device with a finite integration time, thereby computing the area under each pulse of the PDM wave.
- Sample the output of the integrator at a uniform rate to produce a PAM wave, whose pulse amplitudes are proportional to the signal samples $m(nT_s)$ of the original PPM wave $s(t)$.
- Finally, demodulate the PAM wave to recover the message signal $m(t)$.

All the operations described here are linear. In addition, a practical PPM receiver includes a nonlinear device called a *slicer* at its input end. The input–output characteristic of an ideal slicer is shown in Figure 7.13, where the slicing level is normally set at approximately half the peak pulse amplitude of the received PPM wave. The function of the slicer is to preserve the positions of the edges of the received pulses (as modified by noise) and remove everything else. It does so by producing almost "rectangular" pulses with fairly sharp leading and trailing edges at the same instants as the corresponding edges of the received pulses. Thus, in a loose sense, the slicer acts as a "noise cleaning device" in that the final noise level at the output of the receiver is greatly reduced by eliminating all the noise in the received PPM wave except in the neighborhood of the leading and trailing edges.

The output of the slicer is differentiated and then half-wave rectified, yielding a very short pulse (approximating an impulse) each time the amplitude of a pulse in the received PPM wave passes through the slicing level. Figure 7.14*a* shows the nth pulse of a PPM wave, and Figure 7.14*b* shows the short pulse produced (by the operations described herein) as the pulse passes through

the slicing level. In Figure 7.14*c* an appropriate delay is applied to the short pulse, and the corresponding PDM pulse is shown in Figure 7.14*d*.

Having converted the received (noisy) PPM wave into a PDM wave with the same modulation, the receiver then proceeds to reconstruct the original baseband signal $m(t)$ in the manner described above.

NOISE IN PULSE-POSITION MODULATION

In a PPM system, the transmitted information is contained in the relative positions of the modulated pulses. The presence of additive noise affects the performance of such a system by falsifying the time at which the modulated pulses are judged to occur. Immunity to noise can be established by making the pulse build up so rapidly that the time interval during which noise can exert any perturbation is very short. Indeed, additive noise would have no effect on the pulse positions if the received pulses were perfectly rectangular, because the presence of noise introduces only vertical perturbations. However, the reception of perfectly rectangular pulses implies an infinite channel bandwidth, which is of course impractical. Thus, with a finite channel

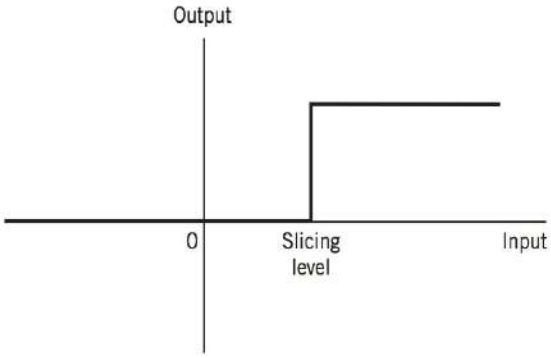


FIGURE 7.13 Input–output relation of slicer.

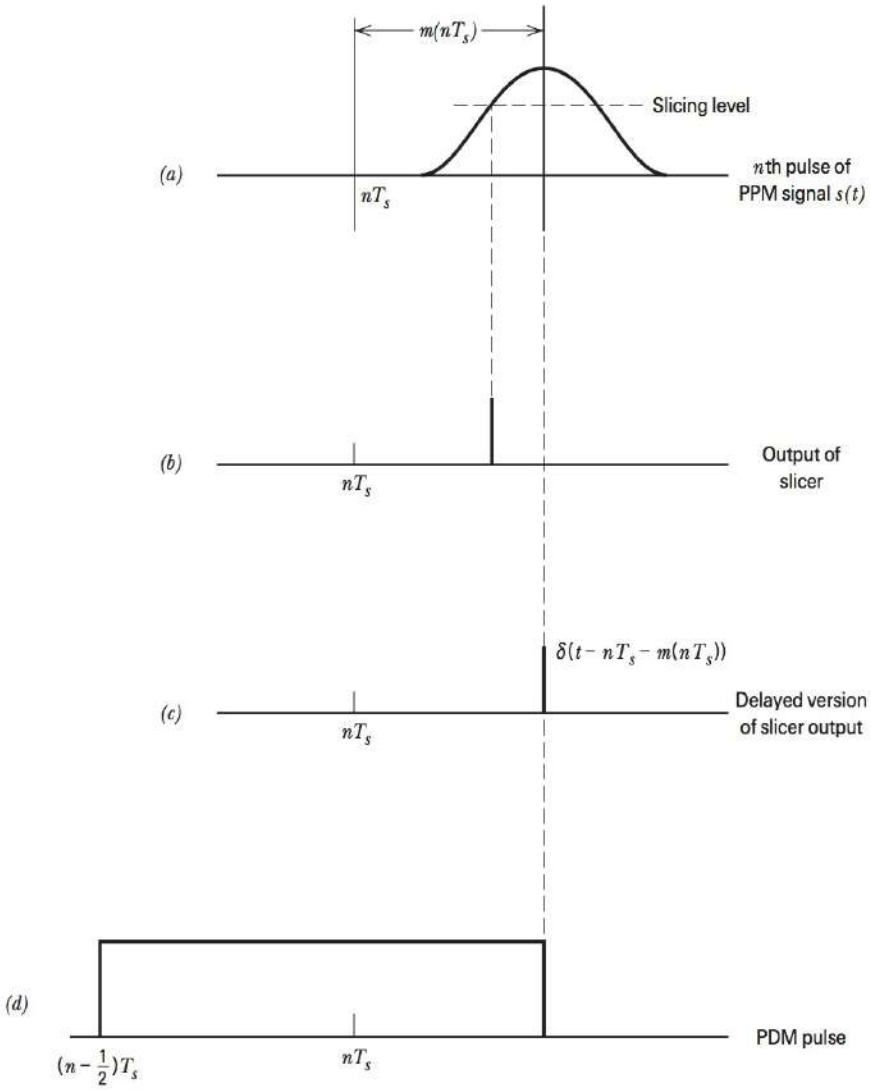


FIGURE 7.14 Detection of a noiseless PPM signal.

bandwidth in practice, we find that the received pulses have a finite rise time, and so the performance of the PPM receiver is affected by noise.

As with a CW modulation system, the noise performance of a PPM system may be described in terms of the output signal-to-noise ratio. Also, to find the noise improvement produced by PPM over baseband transmission of a message signal, we may use the figure of merit defined as the output signal-to-noise ratio of the PPM system divided by the channel signal-to-noise ratio. We illustrate this evaluation by considering the example of a PPM system using a raised cosine pulse and sinusoidal modulation.

EXAMPLE 7.1 Signal-to-Noise Ratios of a PPM System Using Sinusoidal Modulation

Consider a PPM system whose pulse train, in the absence of modulation, is as shown in Figure 7.15*a*. The standard pulse of the carrier is assumed to be a *raised cosine pulse*, which is a convenient type of pulse for analysis. This pulse, centered at $t = 0$ and denoted as $g(t)$, is defined by

$$g(t) = \frac{A}{2}[1 + \cos(\pi B_T t)], \qquad -T \le t \le T \tag{7.23}$$

where $B_T = 1/T$. The time of occurrence of such a pulse may be determined by applying the pulse to an ideal slicer with the input–output amplitude characteristic shown in Figure 7.13 and then observing the slicer output. We assume that the slicing level is set at half the peak pulse amplitude, namely, $A/2$, as in Figure 7.15*a*. For inputs below the slicing level the output is zero, and for inputs above the slicing level the output is constant.

The Fourier transform of the pulse $g(t)$ is given by

$$G(f) = \frac{A \sin(2\pi f/B_T)}{2\pi f(1 - 4f^2/B_T^2)}$$

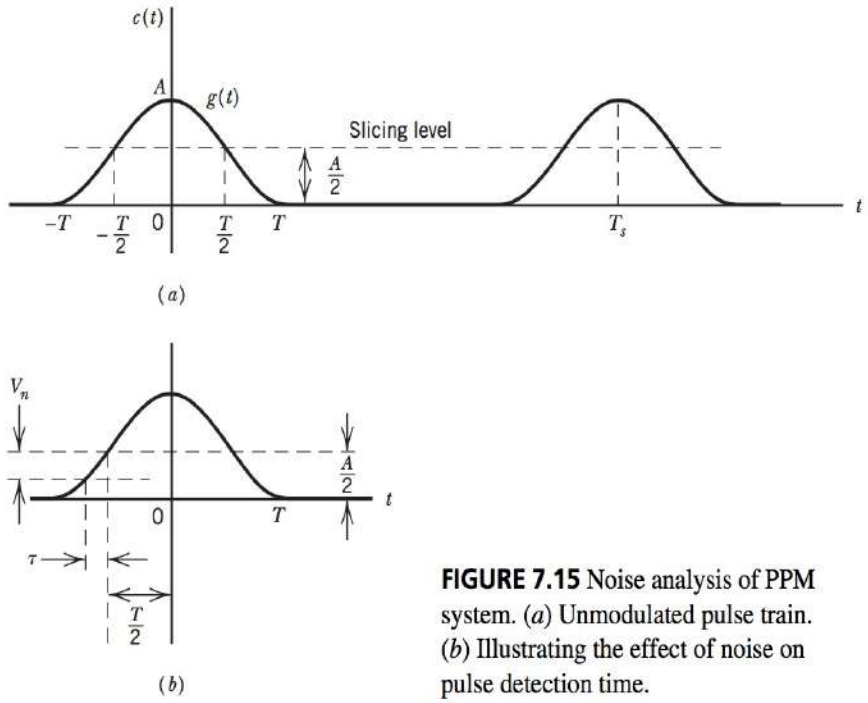


FIGURE 7.15 Noise analysis of PPM system. (*a*) Unmodulated pulse train. (*b*) Illustrating the effect of noise on pulse detection time.

As indicated in Figure 7.16, this transform has its first nulls at $f = \pm B_T$ and is small outside this interval, so that the transmission bandwidth required to pass such a pulse may be taken as essentially equal to B_T.

Let the peak-to-peak swing in the position of a pulse be denoted by T_s. Then, in response to a full-load sinusoidal modulating wave, the peak-to-peak amplitude of the receiver output will be KT_s, where K is a constant determined by the receiver circuitry. The root-mean-square (rms) value of the receiver output is $KT_s/2\sqrt{2}$, and the corresponding average signal power at the receiver output (assuming a 1-ohm load) is given by

$$\left(\frac{KT_s}{2\sqrt{2}}\right)^2 = \frac{K^2T_s^2}{8}$$

In the presence of additive noise, both amplitude and position of the pulse will be perturbed. Random variations in the pulse amplitude are removed by the slicer. Random variations in the pulse position due to noise will remain, however, thereby contributing to noise at the receiver output. We assume that, at the receiver input, the noise power is small compared with the peak pulse power. Then, if at a particular instant of time the noise amplitude is V_n, the time of pulse detection will be replaced by a small amount τ as depicted in Figure 7.15*b*. To a first order of approximation, V_n/τ is equal to the slope of the pulse $g(t)$ at time $t = -T/2$. Thus, using Eq. (7.23), we get

$$\frac{V_n}{\tau} = \left.\frac{dg(t)}{dt}\right|_{t=-T/2} = \frac{\pi B_T A}{2}$$

Solving this equation for τ, we have

$$\tau = \frac{2V_n}{\pi B_T A} \tag{7.24}$$

The error τ in the position of the pulse $g(t)$ will produce an average noise power at the receiver output equal to $K^2\mathbf{E}[\tau^2]$, where $\mathbf{E}$ is the statistical expectation operator. Assuming that the noise at the front end of the receiver has a power spectral density $N_0/2$, we find that the mean-square value of V_n in a bandwidth B_T is given by

$$\mathbf{E}[V_n^2] = B_T N_0 \tag{7.25}$$

Using Eqs. (7.24) and (7.25), we obtain the following result:

$$\begin{aligned}\text{Average power of output noise} &= K^2\mathbf{E}[\tau^2] \\ &= \frac{4K^2N_0}{\pi^2 B_T A^2}\end{aligned} \tag{7.26}$$

The output signal-to-noise ratio, assuming a full-load sinusoidal modulation, is therefore

$$\begin{aligned}(\text{SNR})_O &= \frac{K^2T_s^2/8}{4K^2N_0/\pi^2 B_T A^2} \\ &= \frac{\pi^2 B_T T_s^2 A^2}{32N_0}\end{aligned} \tag{7.27}$$

The average transmitted power P in a PPM system is independent of the applied modulation. Accordingly, we may determine P by averaging the power in a single pulse of the PPM wave over the sampling period T_s, as shown by

$$\begin{aligned}P &= \frac{1}{T_s}\int_{-T_s/2}^{T_s/2} g^2(t)\,dt \\ &= \frac{3A^2}{4T_s B_T}\end{aligned} \tag{7.28}$$

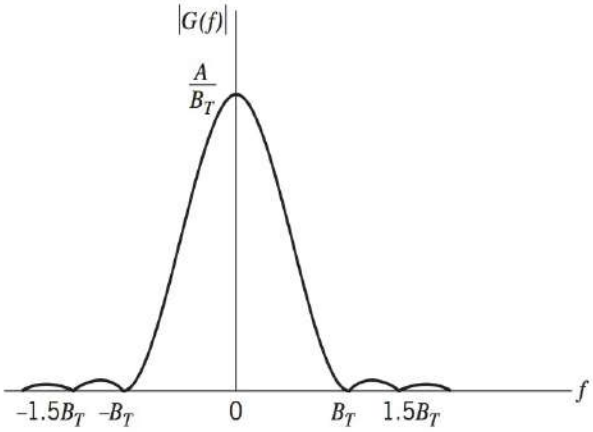


FIGURE 7.16 Amplitude spectrum of a raised cosine pulse.

The average noise power in a message bandwidth W is equal to WN_0. The channel signal-to-noise ratio is therefore

$$(\text{SNR})_C = \frac{3A^2/4T_sB_T}{WN_0} = \frac{3A^2}{4T_sB_TWN_0} \tag{7.29}$$

Thus the figure of merit of a PPM system using a raised cosine pulse is as follows:

$$\text{Figure of merit} = \frac{(\text{SNR})_O}{(\text{SNR})_C} = \frac{\pi^2}{24}B_T^2T_s^3W \tag{7.30}$$

Assuming that the message signal is sampled at its Nyquist rate, we have $T_s = 1/2W$. Then, we find from Eq. (7.30) that the corresponding value of the figure of merit is $(\pi^2/192)\,(B_T/W)^2$, which is greater than unity if $B_T > 4.41\,W$. We also see that the figure of merit of a PPM system is proportional to the square of the normalized transmission bandwidth B_T/W.

In the noise analysis presented here for PPM, we assumed that the average power of the additive noise at the front end of the receiver is small compared with the peak pulse power. In particular, it is assumed that there are two crossings of the slicing level for each pulse, one for the leading edge and one for the trailing edge. A Gaussian noise will have occasional peaks that produce additional crossings of the slicing level, however, and so the occasional noise peaks are mistaken for message pulses. The analysis neglects the *false pulses* produced by high noise peaks. It is apparent that these false pulses have a finite though small probability of occurrence when the noise is Gaussian, no matter how small its standard deviation is compared with the peak amplitude of the pulses. As the transmission bandwidth is increased indefinitely, the accompanying increase in average noise power eventually causes the false pulses to occur often enough, thereby causing loss of the wanted message signal at the receiver output. We thus find, in practice, that a PPM system suffers from a *threshold effect.*

BANDWIDTH-NOISE TRADE-OFF

In the context of noise performance, a PPM system represents the optimum form of analog pulse modulation. The noise analysis of a PPM system presented in the preceding example reveals that pulse-position modulation (PPM) and frequency modulation (FM) systems exhibit a similar noise performance, as summarized here:

1. Both systems have a figure of merit proportional to the square of the transmission bandwidth normalized with respect to the message bandwidth.
2. Both systems exhibit a threshold effect as the signal-to-noise ratio is reduced.

The practical implication of point 1 is that, in terms of a trade-off of increased transmission bandwidth for improved noise performance, the best that we can do with continuous-wave (CW) modulation and analog pulse modulation systems is to follow a *square law.* A question that arises at this point in the discussion is: Can we produce a trade-off better than a square law? The answer is an emphatic yes, and *digital pulse modulation* is the way to do it. The use of such a method is a radical departure from CW modulation.

There are two fundamental processes involved in the generation of a digital pulse representation of an analog signal: *sampling* and *quantization*. The sampling process takes care of the discrete-time representation of the message signal; for its proper application, we have to follow the sampling theorem described in Section 7.3. The quantization process takes care of the discrete-amplitude representation of the message signal; quantization is a new process, the details of which are described in Section 7.8. For now it suffices to say that the combined use of sampling and quantization permits the transmission of a message signal in coded form. This, in turn, makes it possible to realize an *exponential law* for the bandwidth-noise trade off, which is also demonstrated in Section 7.8.

7.7 THEME EXAMPLE—PPM IN IMPULSE RADIO[1]

Traditional digital transmission systems attempt to minimize the bandwidth of the transmitted signal. Hence, filtering is often applied to rectangular pulses to reduce the occupied bandwidth. However, a method that does not follow this philosophy and has captured attention recently is known as *impulse radio*. With this technique, information is sent by means of very narrow pulses that are widely separated in time. Since the pulses' widths are very narrow, the spectrum of the resulting signal is very broad and, consequently, this technique is a form of *ultra-wideband* (*UWB*) *radio transmission* or *impulse radio*.

One type of pulse used for impulse radio is the *Gaussian monocycle*, which is based on the first derivative of the Gaussian pulse. The waveform of the Gaussian monocycle is given by

$$v(t) = A\frac{t}{\tau}\exp\left\{-6\pi\left(\frac{t}{\tau}\right)^2\right\} \tag{7.31}$$

where A is an amplitude scaling factor and τ is the time constant of the pulse. This signal, normalized to unit amplitude, is depicted in Figure 7.17. It consists of a negative lobe followed by a positive lobe, with a total pulse width of approximately τ. For impulse radio applications, the pulse width τ is typically between 0.20 and 1.50 nanoseconds.

The spectrum of a sequence of these pulses can be obtained from the Fourier transform of an individual pulse and this spectrum is shown in Figure 7.18. The frequency axis in Figure 7.18 has been normalized in terms of the time constant τ. For $\tau = 1.0$ nanosecond, this frequency axis ranges from 0 to 5 GHz.

There are several methods for digitally modulating such an impulse wave. One method is *pulse-position modulation*, as depicted in Figure 7.19. With this method, there is a nominal time separation T_p between successive pulses. To transmit a binary symbol "0," the pulse is transmitted slightly early ($-T_c$). To transmit a binary symbol "1," the pulse is transmitted slightly late ($+T_c$). The receiver detects this *early/late timing* and demodulates the data accordingly. Typical separations between pulses (T_p) range from 1000 nanoseconds to as little as 25 nanoseconds, resulting in data rates ranging from 1 to 40 megabits per second.

The ultra-wideband nature of the modulated signal has good and bad aspects. Since the signal power is spread over a large bandwidth, the amount of power that falls in any particular narrowband channel is small.

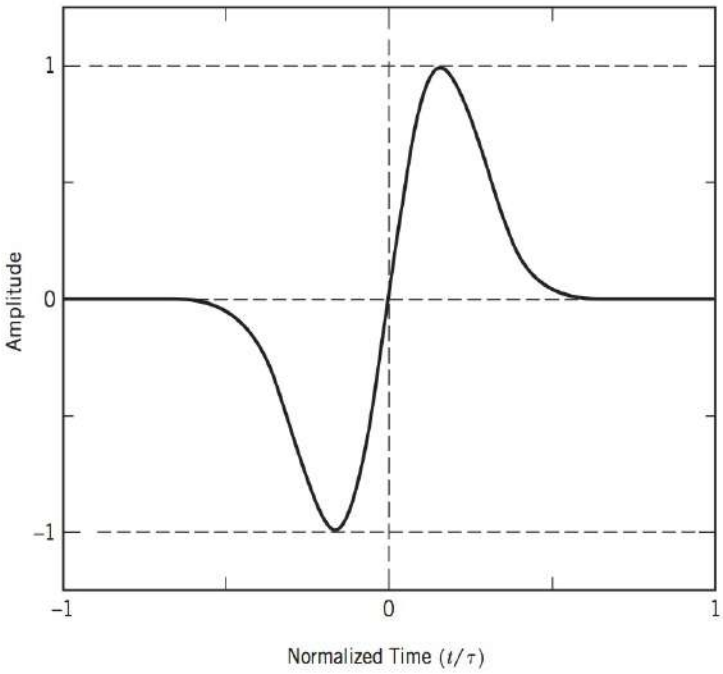


FIGURE 7.17 Illustration of Gaussian monocycle used for impulse radio.

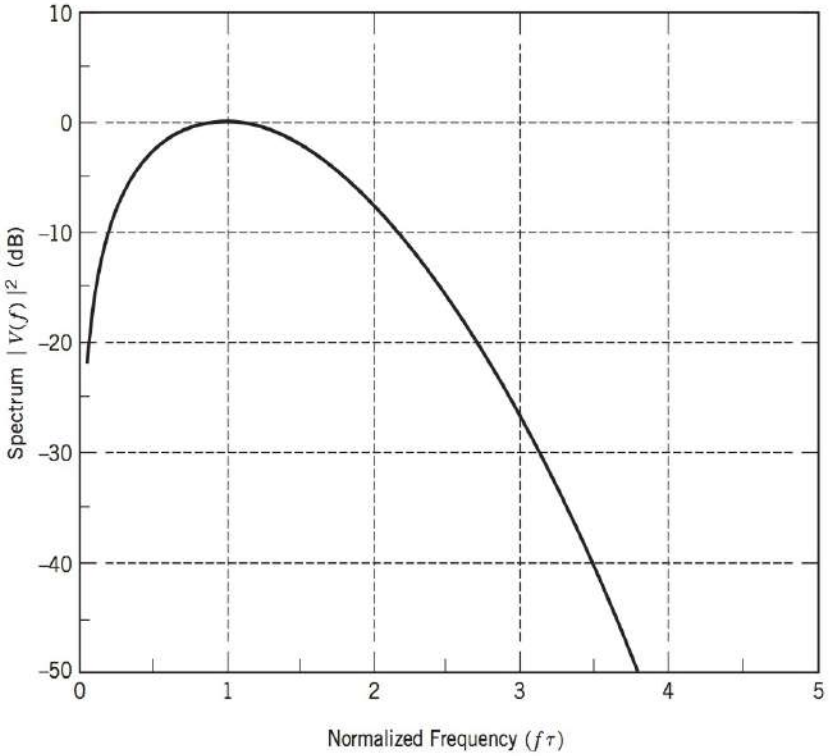


FIGURE 7.18 Amplitude spectrum of Gaussian monocycle.

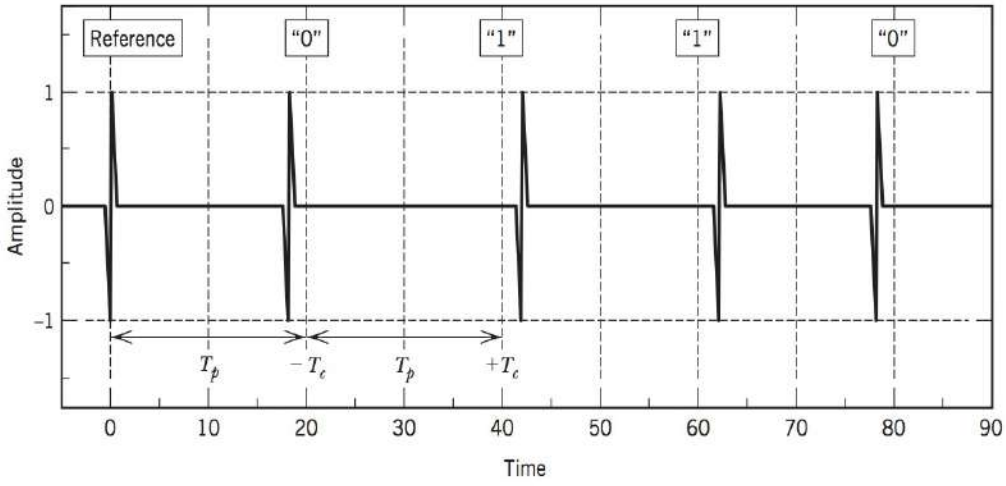


FIGURE 7.19 Pulse-position modulation of impulse radio.

However, such power falls in all such narrowband channels. Consequently, there is a concern that ultra-wideband radios will cause harmful interference into existing narrowband radio services occupying the same radio spectrum. As a consequence, although ultra-wideband radio has been allowed in various jurisdictions, there are strict limits on the power that may be transmitted. Due to these limitations on the transmitted power, ultra-wideband radio is limited to short-range applications, typically less than a few hundred meters.

7.8 THE QUANTIZATION PROCESS

A continuous signal, such as voice, has a continuous range of amplitudes and therefore its samples have a continuous amplitude range. In other words, within the finite amplitude range of the signal, we find an infinite number of amplitude levels. It is not

necessary in fact to transmit the exact amplitudes of the samples. Any human sense (the ear or the eye), as ultimate receiver, can detect only finite intensity differences. This means that the original continuous signal may be *approximated* by a signal constructed of discrete amplitudes selected on a minimum error basis from an available set. Clearly, if we assign the discrete amplitude levels with sufficiently close spacing, we may make the approximated signal practically indistinguishable from the original continuous signal.

Amplitude *quantization* is defined as the process of transforming the sample amplitude $m(nT_s)$ of a message signal $m(t)$ at time $t = nT_s$ into a discrete amplitude $v(nT_s)$ taken from a *finite* set of possible amplitudes. In this book, we assume that the quantization process is *memoryless* and *instantaneous*, which means that the transformation at time $t = nT_s$ is not affected by earlier or later samples of the message signal. This simple form of quantization, though not optimum, is commonly used in practice.

When dealing with a memoryless quantizer, we may simplify the notation by dropping the time index. We may thus use the symbol m in place of $m(nT_s)$, as indicated in the block diagram of a quantizer shown in Figure 7.20*a*. Then, as shown in Figure 7.20*b*, the signal amplitude m is specified by the index k if it lies inside the interval

$$\mathcal{I}_k: \{m_k < m \leq m_{k+1}\}, \qquad k = 1, 2, \ldots, L \tag{7.32}$$

where L is the total number of amplitude levels used in the quantizer. The amplitudes m_k, $k = 1, 2, \ldots, L$, are called *decision levels* or *decision thresholds*. At the quantizer output, the index k is transformed into an amplitude v_k that represents all amplitudes of the interval $\mathcal{I}_k$; the amplitudes v_k, $k = 1, 2, \ldots, L$, are called *representation levels* or *reconstruction levels*, and the spacing between two adjacent representation levels is called a *quantum* or *step-size*. Thus, the quantizer output v equals v_k if the input signal sample m belongs to the interval $\mathcal{I}_k$. The mapping (see Figure 7.20*a*)

$$v = g(m) \tag{7.33}$$

is the *quantizer characteristic*, which is a staircase function by definition.

Quantizers can be of a *uniform* or *nonuniform* type. In a uniform quantizer, the representation levels are uniformly spaced; otherwise, the quantizer is non-uniform. In this section, we consider only uniform quantizers; nonuniform quantizers are considered in the next section. The quantizer characteristic can also be of *midtread* or *midrise type*. Figure 7.21*a* shows the input–output characteristic of a uniform quantizer of the midtread type, which is so called because the origin lies in the middle of a tread of the staircaselike graph. Figure 7.21*b* shows the corresponding input–output characteristic of a uniform quantizer of the midrise type, in which the origin lies in the middle of a rising part of the staircaselike graph. Note that both the midtread and midrise types of uniform quantizers illustrated in Figure 7.21 are *symmetric* about the origin.

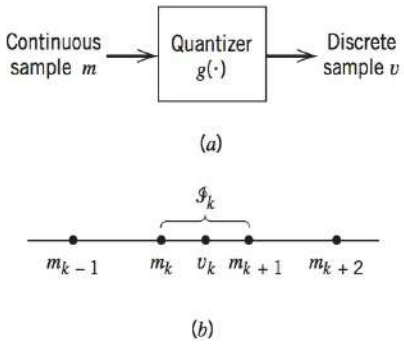


FIGURE 7.20 Description of a memoryless quantizer.

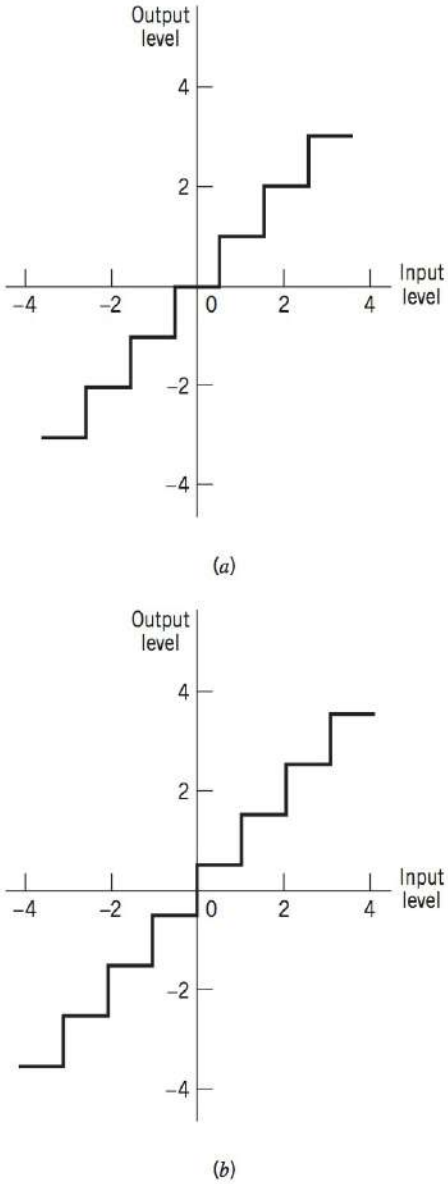


FIGURE 7.21 Two types of quantization: (*a*) midtread and (*b*) midrise.

QUANTIZATION NOISE

The use of quantization introduces an error defined as the difference between the input signal m and the output signal v. This error is called *quantization noise*. Figure 7.22 illustrates a typical variation of the quantization noise as a function of time, assuming the use of a uniform quantizer of the midtread type.

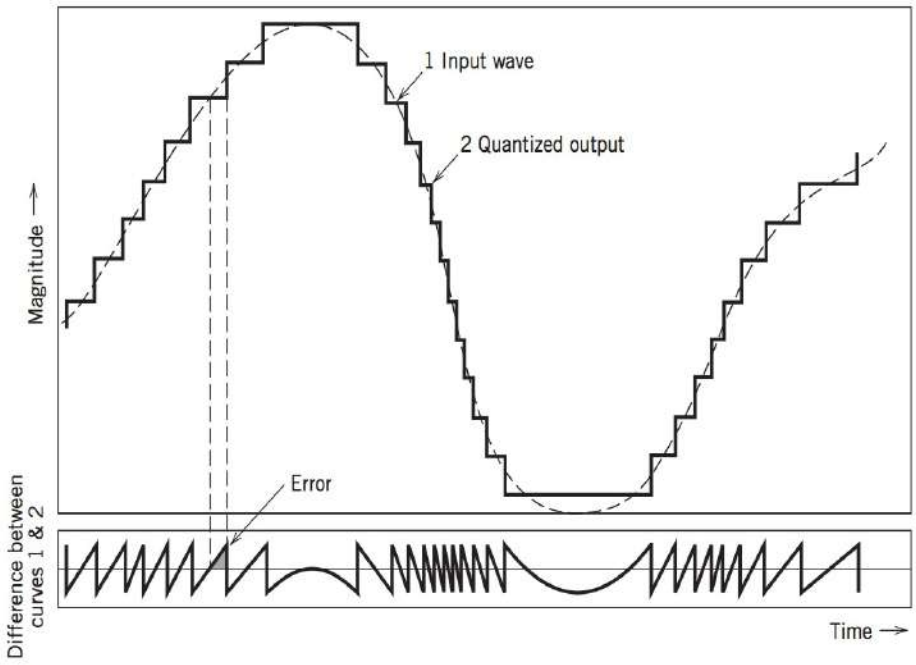


FIGURE 7.22 Illustration of the quantization process. (Adapted from Bennett, 1948, with permission of AT&T.)

Let the quantizer input m be the sample value of a zero-mean random variable M. (If the input has a nonzero mean, we can always remove it by subtracting the mean from the input and then adding it back after quantization.) A quantizer $g(\cdot)$ maps the input random variable M of continuous amplitude into a discrete random variable V; their respective sample values m and v are related by Eq. (7.33). Let the quantization error be denoted by the random variable Q of sample value q. We may thus write

$$q = m - v \tag{7.34}$$

or, correspondingly,

$$Q = M - V \tag{7.35}$$

With the input M having zero mean, and the quantizer assumed to be symmetric as in Figure 7.21, it follows that the quantizer output V and therefore the quantization error Q, will also have zero mean. Thus, for a partial statistical characterization of the quantizer in terms of output signal-to-noise ratio, we need to find only the mean-square value of the quantization error Q.

Consider then an input m of continuous amplitude in the range $(-m_{\max}, m_{\max})$. Assuming a uniform quantizer of the midrise type illustrated in Figure 7.21*b*, we find that the step-size of the quantizer is given by

$$\Delta = \frac{2m_{\max}}{L} \tag{7.36}$$

where L is the total number of representation levels. For a uniform quantizer, the quantization error Q will have its sample values bounded by $-\Delta/2 \leq q \leq \Delta/2$. If the step-size Δ is sufficiently small (i.e., the number of representation levels L is sufficiently large), it is reasonable to assume that the quantization error Q is a *uniformly distributed* random variable, and the interfering effect of the quantization noise on the quantizer input is similar to that of thermal noise. We may thus express the probability density function of the

quantization error Q as follows:

$$f_Q(q) = \begin{cases} \dfrac{1}{\Delta}, & -\dfrac{\Delta}{2} < q \le \dfrac{\Delta}{2} \\ 0, & \text{otherwise} \end{cases} \tag{7.37}$$

For this to be true, however, we must ensure that the incoming signal does *not* overload the quantizer. Then, with the mean of the quantization error being zero, its variance σ_Q^2 is the same as the mean-square value

$$\sigma_Q^2 = \int_{-\Delta/2}^{\Delta/2} q^2 f_Q(q)\, dq = \mathbf{E}[Q^2] \tag{7.38}$$

Substituting Eq. (7.37) in (7.38), we get

$$\begin{aligned} \sigma_Q^2 &= \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} q^2\, dq \\ &= \frac{\Delta^2}{12} \end{aligned} \tag{7.39}$$

Typically, the L-ary number k, denoting the kth representation level of the quantizer, is transmitted to the receiver in binary form. Let R denote the number of *bits per sample* used in the construction of the binary code. We may then write

$$L = 2^R \tag{7.40}$$

or, equivalently,

$$R = \log_2 L \tag{7.41}$$

Hence, substituting Eq. (7.40) in (7.36), we get the step size

$$\Delta = \frac{2m_{\max}}{2^R} \tag{7.42}$$

Thus, the use of Eq. (7.42) in (7.39) yields

$$\sigma_Q^2 = \frac{1}{3} m_{\max}^2 2^{-2R} \tag{7.43}$$

Let P denote the average power of the message signal $m(t)$. We may then express the *output signal-to-noise ratio* of a uniform quantizer as

$$\begin{aligned} (\text{SNR})_O &= \frac{P}{\sigma_Q^2} \\ &= \left(\frac{3P}{m_{\max}^2}\right) 2^{2R} \end{aligned} \tag{7.44}$$

Equation (7.44) shows that the output signal-to-noise ratio of the quantizer increases *exponentially* with increasing number of bits per sample, R. Recognizing that an increase in R requires a proportionate increase in the channel (transmission) bandwidth B_T, we thus see that the use of a binary code for the representation of a message signal (as in pulse-code modulation) provides a more efficient method than either frequency modulation (FM) or pulse-position modulation (PPM) for the trade-off of increased channel bandwidth for improved noise performance. In making this statement, we presume that the FM and PPM systems are limited by receiver noise, whereas the binary-coded modulation system is limited by quantization noise. We have more to say on the latter issue in Section 7.9.

EXAMPLE 7.2 Sinusoidal Modulating Signal

Consider the special case of a full-load sinusoidal modulating signal of amplitude A_m, which utilizes all the representation levels provided. The average signal power is (assuming a load of 1 ohm)

$$P = \frac{A_m^2}{2}$$

The total range of the quantizer input is $2A_m$, because the modulating signal swings between $-A_m$ and A_m. We may therefore set $m_{max} = A_m$, in which case the use of Eq. (7.43) yields the average power (variance) of the quantization noise as

$$\sigma_Q^2 = \tfrac{1}{3} A_m^2 2^{-2R}$$

Thus the output signal-to-noise ratio of a uniform quantizer, for a full-load test tone, is

$$(\text{SNR})_O = \frac{A_m^2/2}{A_m^2 2^{-2R}/3} = \frac{3}{2}(2^{2R}) \qquad (7.45)$$

Expressing the signal-to-noise ratio in decibels, we get

$$10 \log_{10}(\text{SNR})_O = 1.8 + 6R \qquad (7.46)$$

For various values of L and R, the corresponding values of signal-to-noise ratio are as given in Table 7.1. From Table 7.1 we can make a quick estimate of the number of bits per sample required for a desired output signal-to-noise ratio.

TABLE 7.1 Signal-to-quantization noise ratio for varying number of representation levels

Number of Representation Levels, L	Number of Bits/Sample, R	Signal-to-Noise Ratio (dB)
32	5	31.8
64	6	37.8
128	7	43.8
256	8	49.8

7.9 PULSE-CODE MODULATION[2]

With the sampling and quantization processes at our disposal, we are now ready to describe pulse-code modulation which, as mentioned previously, is the most basic form of digital pulse modulation. In *pulse-code modulation (PCM) a message signal is represented by a sequence of coded pulses, which is accomplished by representing the signal in discrete form in both time and amplitude*. The basic operations performed in the transmitter of a PCM system are *sampling*, *quantizing*, and *encoding*, as shown in Figure 7.23*a*; the low-pass filter prior to sampling is included to prevent aliasing of the message signal. The quantizing and encoding operations are usually performed in the same circuit, which is called an *analog-to-digital converter*. The basic operations in the receiver are *regeneration* of impaired signals, *decoding*, and *reconstruction* of the train of quantized samples, as shown in Figure 7.23*c*. Regeneration also occurs at intermediate points along the transmission path as necessary, as indicated in Figure 7.23*b*. When time-division multiplexing is used, it becomes necessary to synchronize the receiver to the transmitter for the overall system to operate satisfactorily. In what follows we describe the various operations that constitute a PCM system.

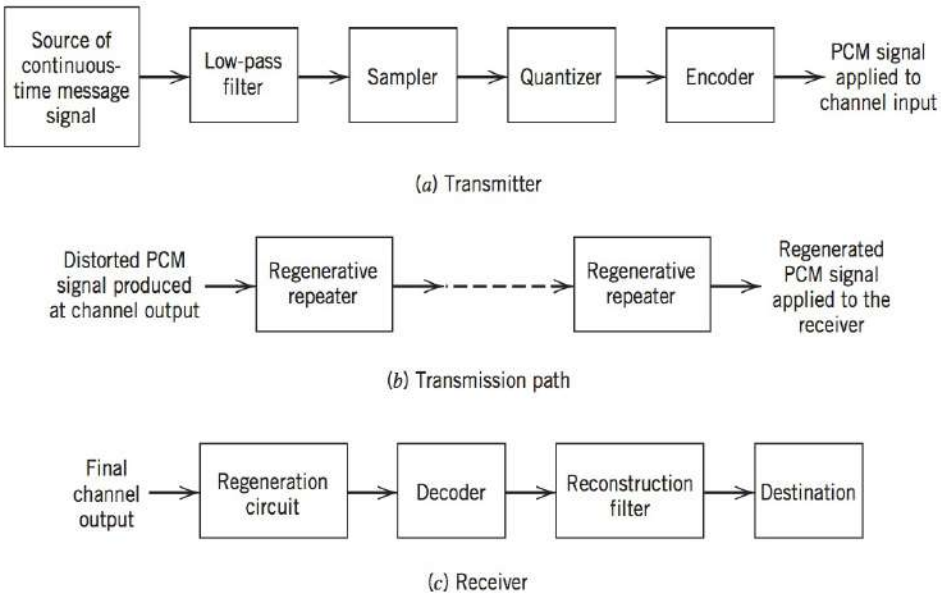


FIGURE 7.23 The basic elements of a PCM system.

SAMPLING

The incoming message signal is sampled with a train of narrow rectangular pulses so as to closely approximate the instantaneous sampling process. In order to ensure perfect reconstruction of the message signal at the receiver, the sampling rate must be greater than twice the highest frequency component W of the message signal in accordance with the sampling theorem. In practice, a pre-alias (low-pass) filter is used at the front end of the sampler in order to exclude frequencies greater than W before sampling. Thus the application of sampling permits the reduction of the continuously varying message signal (of some finite duration) to a limited number of discrete values per second.

QUANTIZATION

The sampled version of the message signal is then quantized, thereby providing a new representation of the signal that is discrete in both time and amplitude. The quantization process may follow a uniform law as described in Section 7.8. In certain applications, however, it is preferable to use a variable separation between the representation levels. For example, the range of voltages covered by voice signals, from the peaks of loud talk to the weak passages of weak talk, is on the order of 1000 to 1. By using a *nonuniform quantizer* with the feature that the step-size increases as the separation from the origin of the input–output amplitude characteristic is increased, the large end step of the quantizer can take care of possible excursions of the voice signal into the large amplitude ranges that occur relatively infrequently. In other words, the weak passages, which need more protection, are favored at the expense of the loud passages. In this way, a nearly uniform percentage precision is achieved throughout the greater part of the amplitude range of the input signal, with the result that fewer steps are needed than would be the case if a uniform quantizer were used.

The use of a nonuniform quantizer is equivalent to passing the baseband signal through a *compressor* and then applying the compressed signal to a uniform quantizer. A particular form of compression law that is used in practice is the so called μ-*law*[3] defined by

$$|v| = \frac{\log(1 + \mu|m|)}{\log(1 + \mu)} \tag{7.47}$$

where m and v are the normalized input and output voltages, and μ is a positive constant. In Figure 7.24*a*, we have plotted the μ-law for varying μ. The case of uniform quantization corresponds to $\mu = 0$. For a given value of μ, the reciprocal slope of the compression curve, which defines the quantum steps, is given by the derivative of $|m|$ with respect to $|v|$; that is,

$$\frac{d|m|}{d|v|} = \frac{\log(1+\mu)}{\mu}(1+\mu|m|) \tag{7.48}$$

We see therefore that the μ-law is neither strictly linear nor strictly logarithmic, but it is approximately linear at low input levels corresponding to $\mu|m| \ll 1$, and approximately logarithmic at high input levels corresponding to $\mu|m| \gg 1$.

Another compression law that is used in practice is the so-called *A-law* defined by

$$|v| = \begin{cases} \dfrac{A|m|}{1+\log A}, & 0 \le |m| \le \dfrac{1}{A} \\ \dfrac{1+\log(A|m|)}{1+\log A}, & \dfrac{1}{A} \le |m| \le 1 \end{cases} \tag{7.49}$$

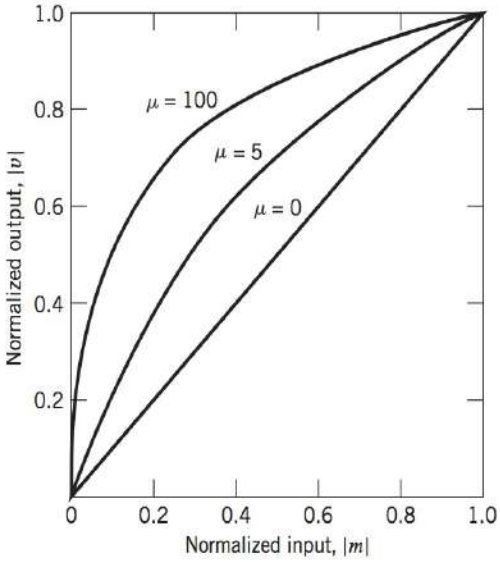
which is shown plotted in Figure 7.24*b*. Practical values of A (as of μ in the μ-law) tend to be in the vicinity of 100. The case of uniform quantization corresponds to $A = 1$. The reciprocal slope of this compression curve is given by the derivative of $|m|$ with respect to $|v|$, as shown by

$$\frac{d|m|}{d|v|} = \begin{cases} \dfrac{1+\log A}{A}, & 0 \le |m| \le \dfrac{1}{A} \\ (1+\log A)|m|, & \dfrac{1}{A} \le |m| \le 1 \end{cases} \tag{7.50}$$

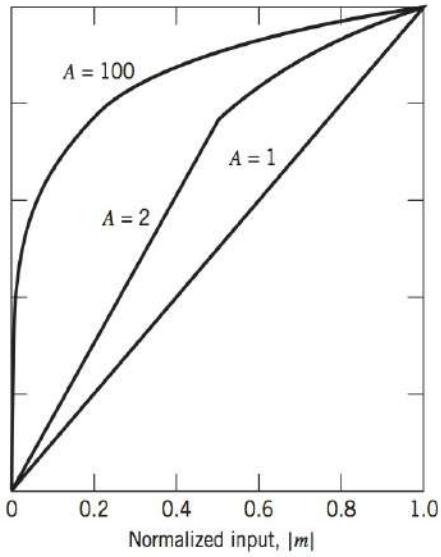
Thus the quantum steps over the central linear segment, which have the dominant effect on small signals, are diminished by the factor $A/(1 + \log A)$. This is typically about 25 dB in practice, as compared with uniform quantization.

In order to restore the signal samples to their correct relative level, we must, of course, use a device in the receiver with a characteristic complementary to the compressor. Such a device is called an *expander*. Ideally, the compression and expansion laws are exactly inverse so that, except for the effect of quantization, the expander output is equal to the compressor input. The combination of a *compressor* and an exp*ander* is called a *compander*.

In actual PCM systems, the companding circuitry does not produce an exact replica of the nonlinear compression curves shown in Figure 7.24. Rather, it provides a *piecewise linear* approximation to the desired curve. By using a large enough number of linear segments, the approximation can approach the true compression curve very closely. This form of approximation is illustrated in Example 7.3 at the end of the section.



(*a*)



(*b*)

FIGURE 7.24 Compression laws, (*a*) μ-law. (*b*) A-law.

ENCODING

In combining the processes of sampling and quantizing, the specification of a continuous message (baseband) signal becomes limited to a discrete set of values, but not in the form best suited to transmission over a line or radio path. To exploit the advantages of sampling and quantizing for the

purpose of making the transmitted signal more robust to noise, interference, and other channel degradations, we require the use of an *encoding process* to translate the discrete set of sample values to a more appropriate form of signal. Any plan for representing each of this discrete set of values as a particular arrangement of discrete events is called a *code*. One of the discrete events in a code is called a *code element* or *symbol*. For example, the presence or absence of a pulse is a symbol. A particular arrangement of symbols used in a code to represent a single value of the discrete set is called a *code word* or *character*.

In a *binary code*, each symbol may be either of two distinct values or kinds, such as the presence or absence of a pulse. The two symbols of a binary code are customarily denoted as 0 and 1. In a *ternary code*, each symbol may be one of three distinct values or kinds, and so on for other codes. However, *the maximum advantage over the effects of noise in a transmission medium is obtained by using a binary code, because a binary symbol withstands a relatively high level of noise and is easy to regenerate.* Suppose that, in a binary code, each code word consists of R *bits:* the bit is an acronym for *binary digit*; thus R denotes the number of *bits per sample*. Then, using such a code, we may represent a total of 2^R distinct numbers. For example, a sample quantized into one of 256 levels may be represented by an 8-bit code word.

There are several ways of establishing a one-to-one correspondence between representation levels and code words. A convenient method is to express the ordinal number of the representation level as a binary number. In the binary number system, each digit has a place-value that is a power of 2, as illustrated in Table 7.2 for the case of four bits per sample (i.e., $R = 4$).

TABLE 7.2 Binary number system for $R = 4$.

Ordinal number of Representation Level	Level Number Expressed as Sum of Powers of 2	Binary Number
0		0000
1	2^0	0001
2	2^1	0010
3	$2^1 + 2^0$	0011
4	2^2	0100
5	$2^2 + 2^0$	0101
6	$2^2 + 2^1$	0110
7	$2^2 + 2^1 + 2^0$	0111
8	2^3	1000
9	$2^3 + 2^0$	1001
10	$2^3 + 2^1$	1010
11	$2^3 + 2^1 + 2^0$	1011
12	$2^5 + 2^2$	1100
13	$2^3 + 2^2 + 2^0$	1101
14	$2^3 + 2^2 + 2^1$	1110
15	$2^3 + 2^2 + 2^1 + 2^0$	1111

LINE CODES

It is in a *line code* that a binary stream of data takes on an electrical representation. Any one of several line codes can be used for the electrical representation of a binary data stream. Figure 7.25 displays the waveforms of five important line codes for the example data stream 01101001. The lines codes often used the terminology *nonreturn-to-zero* (NRZ) or *return-to-zero* (RZ). Return-to-zero implies that the pulse shape used to represent the bit always returns to the 0 volts or the neutral level before the end of the bit. Nonreturn-to-zero indicates that the pulse does not necessarily return to the neutral level before the end of the bit. The five line codes illustrated in Figure 7.25 are described here:

1. Unipolar Nonreturn-to-Zero (NRZ) Signaling. In this line code, symbol 1 is represented by transmitting a pulse of amplitude A for the duration of the symbol, and symbol 0 is represented by switching off the pulse, as in Figure 7.25*a*. This line code is also referred to as *on-off signaling*. A disadvantage of on-off signaling is the waste of power due to the transmitted DC level.

2. Polar Nonreturn-to-Zero (NRZ) Signaling. In this second line code, symbols 1 and 0 are represented by transmitting pulses of amplitudes $+A$ and $-A$, respectively, as illustrated in Figure 7.25*b*. This line code is relatively easy to generate and is more power-efficient than its unipolar counterpart.

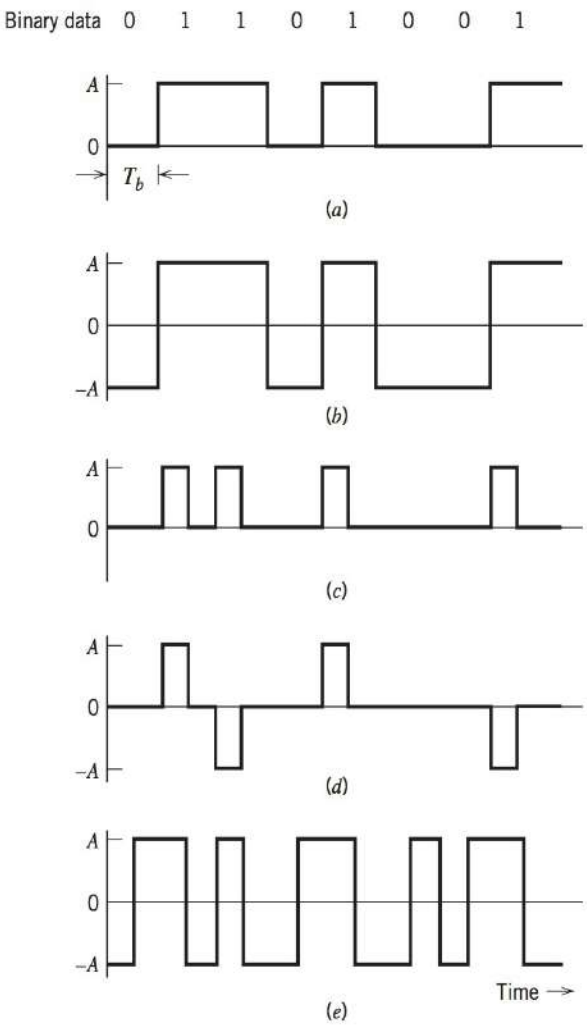


FIGURE 7.25 Line codes for the electrical representations of binary data. (*a*) unipolar NRZ signaling, (*b*) polar NRZ signaling, (*c*) unipolar RZ signaling, (*d*) bipolar RZ signaling, (*e*) split-phase or Manchester code.

3. Unipolar Return-to-Zero (RZ) Signaling. In this other line code, symbol 1 is represented by a rectangular pulse of amplitude A and half-symbol width, and symbol 0 is represented by transmitting no pulse, as illustrated in Figure 7.25*c*. An attractive feature of this line code is the presence of delta functions at $f = 0, \pm 1/T_b$ in the power spectrum of the transmitted signal, which can be used for bit-timing recovery at the receiver. However, its disadvantage is that it requires 3 dB more power than polar return-to-zero signaling for the same probability of symbol error. This issue is addressed in Chapter 8.

4. Bipolar Return-to-Zero (BRZ) Signaling. This line code uses three amplitude levels as indicated in Figure 7.25*d*. Specifically, positive and negative pulses of equal amplitude (i.e., $+A$ and $-A$) are used alternately for symbol 1, with each pulse having a half-symbol width. No pulse is always used for symbol 0. A useful property of the BRZ signaling is that the power spectrum of the transmitted signal has no DC component and relatively insignificant low-frequency components for the case when symbols 1 and 0 occur with equal probability. This line code is also called *alternate mark inversion* (AMI) signaling.

5. Split-Phase (Manchester Code). In this method of signaling, illustrated in Figure 7.25*e*, symbol 1 is represented by a positive pulse of amplitude A followed by a negative pulse of amplitude $-A$, with both pulses being a half-symbol wide. For symbol 0, the polarities of these two pulses are reversed. The Manchester code suppresses the DC component and has relatively insignificant low-frequency components, regardless of the signal statistics. This property is essential in some applications.

DIFFERENTIAL ENCODING

This method is used to encode information in terms of signal transitions. In particular, a transition is used to designate symbol 0 in the incoming binary data stream, while no transition is used to designate symbol 1, as illustrated in Figure 7.26. In Figure 7.26*b* we show the differentially encoded data stream for the example data specified in Figure 7.26*a*. The original binary data stream used here is the same as that used in Figure 7.25. The waveform of the differentially encoded data is shown in Figure 7.26*c*, assuming the use of unipolar nonreturn-to-zero signaling. From Figure 7.26 it is apparent that a differentially encoded signal may be inverted without affecting its interpretation. The original binary information is recovered simply by

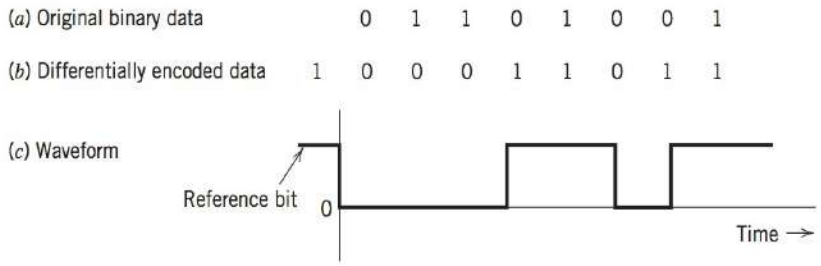


FIGURE 7.26 (*a*) Original binary data. (*b*) Differentially encoded data, assuming reference bit 1. (*c*) Waveform of differentially encoded data using unipolar NRZ signaling.

comparing the polarity of adjacent binary symbols to establish whether or not a transition has occurred. Note that differential encoding requires the use of a reference bit before initiating the encoding process. In Figure 7.26, symbol 1 is used as the reference bit.

REGENERATION

The most important feature of any digital system lies in the ability to control the effects of distortion and noise produced by transmitting a digital signal through a channel. This capability is accomplished by reconstructing the signal by means of a chain of *regenerative repeaters* located at sufficiently close spacing along the transmission route. As illustrated in Figure 7.27, three basic functions are performed by a regenerative repeater: *equalization, timing*, and *decision making*. The equalizer shapes the received pulses so as to compensate for the effects of amplitude and phase distortions produced by the transmission characteristics of the channel. The timing circuitry provides a periodic pulse train, derived from the received pulses, for sampling the equalized pulses at the instants of time where the signal-to-noise ratio is a maximum. The sample so extracted is compared to a predetermined *threshold* in the decision-making device. In each bit interval a decision is then made whether the received symbol is a 1 or a 0 on the basis of whether the threshold is exceeded or not. If the threshold is exceeded, a clean new pulse representing symbol 1 is transmitted to the next repeater. Otherwise, another clean new pulse representing symbol 0 is transmitted. In this way, the accumulation of distortion and noise in a repeater span is completely removed, provided that the disturbance is not too large to cause an error in the decision-making process. Ideally, except for delay, the regenerated signal is exactly the same as the signal originally transmitted. In practice, however, the regenerated signal departs from the original signal for two main reasons:

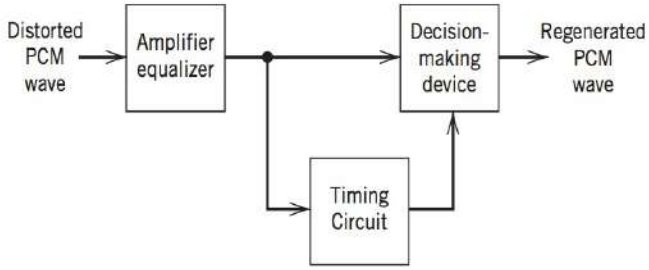


FIGURE 7.27 Block diagram of a regenerative repeater.

1. The unavoidable presence of channel noise and interference causes the repeater to make wrong decisions occasionally, thereby introducing *bit errors* into the regenerated signal.
2. If the spacing between received pulses deviates from its assigned value, a *jitter* is introduced into the regenerated pulse position, thereby causing distortion.

DECODING

The first operation in the receiver is to regenerate (i.e., reshape and clean up) the received pulses one last time. These clean pulses are then regrouped into code words and decoded (i.e., mapped back) into a quantized PAM signal. The *decoding* process involves generating a pulse the amplitude of which is the linear sum of all the pulses in the code word, with each pulse being weighted by its place value ($2^0, 2^1, 2^2, 2^3, \ldots, 2^{R-1}$) in the code, where R is the number of bits per sample.

FILTERING

The final operation in the receiver is to recover the message signal wave by passing the decoder output through a low-pass reconstruction filter whose cutoff frequency is equal to the message bandwidth W. Assuming that the transmission path is error free, the recovered signal includes no noise with the exception of the initial distortion introduced by the quantization process.

MULTIPLEXING

In applications using PCM, it is natural to multiplex different message sources by time division, whereby each source keeps its individuality throughout the journey from the transmitter to the receiver. This individuality accounts for the comparative ease with which message sources may be dropped or reinserted in a time-division multiplex system. As the number of independent message sources is increased, the time interval that may be allotted to each source has to be reduced, since all of them must be accommodated into a time interval equal to the reciprocal of the sampling rate. This in turn means that the allowable duration of a code word representing a single sample is reduced. However, pulses tend to become more difficult to generate and to transmit as their duration is reduced. Furthermore, if the pulses become too short, impairments in the transmission medium begin to interfere with the proper operation of the system. Accordingly, in practice, it is necessary to restrict the number of independent message sources that can be included within a time-division group.

EXAMPLE 7.3 Time-division Multiplexing, PCM, and the T1 System

In this example, we describe the important characteristics of a system known as the T1 carrier system, which is designed to accommodate 24 voice channels, primarily for short-distance, heavy usage in metropolitan areas. The T1 system was pioneered by the Bell System in the United States in the early 1960s, and with its introduction the shift to digital communications facilities started. The T1 system was adopted for use throughout the United States, Canada, and Japan.

A voice signal is essentially limited to a band from 300 to 3100 Hz in that frequencies outside this band do not contribute much to articulation efficiency. Indeed, telephone circuits that respond only to this range of frequencies give quite satisfactory service. Accordingly, it is customary to pass the voice signal through a low-pass filter with a cutoff frequency of about 3.1 kHz prior to sampling. Hence, with $W = 3.1$ KHz, the nominal value of the Nyquist rate is 6.2 kHz. The filtered voice signal is usually sampled at a slightly higher rate, namely 8 kHz, which is the standard sampling rate in telephone systems.

For companding, the T1 system uses a piece-linear characteristic (consisting of 15 linear segments) to approximate the logarithmic μ-law of Eq. (7.47) with the constant $\mu = 255$. There are a total of 255 amplitude levels associated with this companding characteristic. To accommodate this number of representation levels, each of the 24 voice channels uses a binary code with an 8-bit word of the PCM signal.

The different voice channels are combined using a time-division multiplex (TDM) strategy as described in Section 7.5. Each frame of the multiplexed signal consists of twenty-four 8-bit words, one for each voice source, plus a single bit that is added to the end of the frame for synchronization. Hence, each frame consists of a total of $(24 \times 8) + 1 = 193$ bits. With a sampling rate of 8 kHz for each voice channel, this implies that each frame has a period of 125 μs. Correspondingly, the duration of each bit is 0.647 μs and the resultant transmission rate is 1.544 megabits per second.

The T1 carrier system is slow by today's wireline standards, but the basic principles still apply. Modern metropolitan networks are more likely to use fiber optic transmission lines, and the T1 network is replaced by *synchronous optical network* (SONET)[4]. However, SONET uses a digital time-division multiplex scheme with a basic frame size of 125 μs. The basic frame contains 6480 bits, and the minimum SONET rate is 51.84 megabits per second.

7.10 DELTA MODULATION[5]

In some applications, the increased bandwidth requirement of PCM is a reason for concern. In this section, we discuss an alternative method of digitally representing analog information sources, called *delta modulation.*

In *delta modulation* (DM), an incoming message signal is oversampled (i.e., at a rate much higher than the Nyquist rate) to purposely increase the correlation between adjacent samples of the signal. This is done to permit the use of a simple quantizing strategy for constructing the encoded signal.

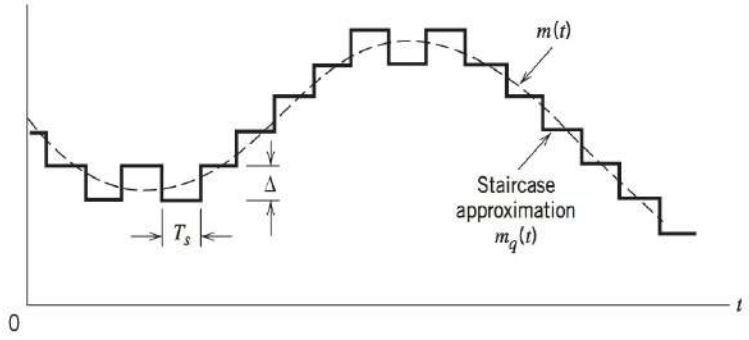
In its basic form, DM provides a *staircase approximation* to the oversampled version of the message signal, as illustrated in Figure 7.28*a*. The difference between the input and the approximation is quantized into only two levels, namely, $\pm\Delta$, corresponding to positive and negative differences, respectively. Thus, if the approximation falls below the signal at any sampling epoch, it is increased by Δ. If, on the other hand, the approximation lies above the signal, it is diminished by Δ. Provided that the signal does not change too rapidly from sample to sample, we find that the staircase approximation remains within $\pm\Delta$ of the input signal.

Denoting the input signal as $m(t)$, and its staircase approximation as $m_q(t)$, the basic principle of delta modulation may be formalized in the following set of discrete-time relations:

$$e(nT_s) = m(nT_s) - m_q(nT_s - T_s) \tag{7.51}$$

$$e_q(nT_s) = \Delta\,\mathrm{sgn}[e(nT_s)] \tag{7.52}$$

$$m_q(nT_s) = m_q(nT_s - T_s) + e_q(nT_s) \tag{7.53}$$



(*a*)

(*b*)

FIGURE 7.28 Illustration of delta modulation.

where T_s is the sampling period; $e(nT_s)$ is an error signal representing the difference between the present sample value $m(nT_s)$ of the input signal and the latest approximation to it, that is, $m_q(nT_s - T_s)$; and $e_q(nT_s)$ is the quantized version of $e(nT_s)$. The quantizer output $e_q(nT_s)$ is finally coded to produce the desired DM signal.

Figure 7.28*a* illustrates the way in which the staircase approximation $m_q(t)$ follows variations in the input signal $m(t)$ in accordance with Eqs. (7.51) to (7.53), and Figure 7.28*b* displays the corresponding binary sequence at the delta modulator output. It is apparent that in a delta modulation system the rate of information transmission is simply equal to the sampling rate $f_s = 1/T_s$.

The principal virtue of delta modulation is its simplicity. It may be generated by applying the sampled version of the incoming message signal to a digital encoder that involves a *comparator*, *quantizer*, and *accumulator* interconnected as shown in Figure 7.29*a*. Details of the modulator follow directly from Eqs. (7.51) to (7.53). The comparator computes the difference between its two inputs. The quantizer consists of a *hard limiter* with an input–output relation that is a scaled version of the signum function. The quantizer output is then applied to an accumulator, producing the result

$$\begin{aligned} m_q(nT_s) &= \Delta \sum_{i=1}^{n} \operatorname{sgn}[e(iT_s)] \\ &= \sum_{i=1}^{n} e_q(iT_s) \end{aligned} \tag{7.54}$$

which is obtained by solving Eqs. (7.52) and (7.53) for $m_q(nT_s)$. Thus, at the sampling instant nT_s, the accumulator increments the approximation by a step Δ in a positive or negative

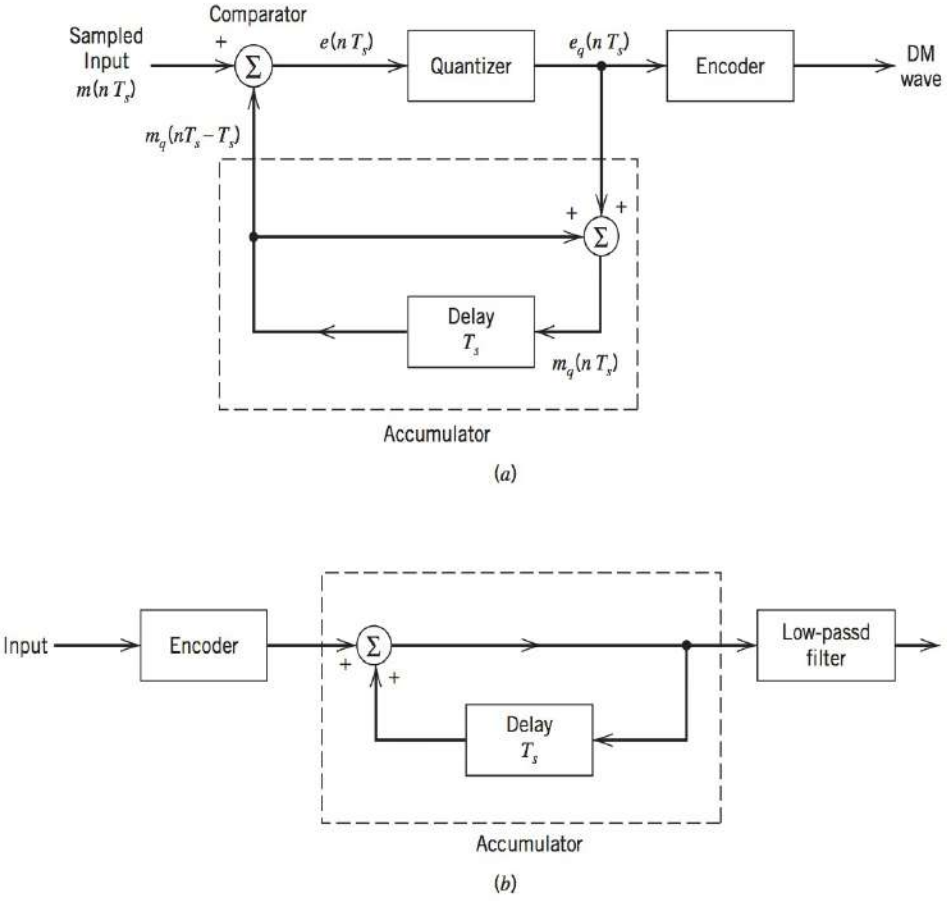


FIGURE 7.29 Delta modulation system. (*a*) Analog to digital encoder. (*b*) Digital to analog decoder.

direction, depending on the algebraic sign of the error signal $e(nT_s)$. If the input signal $m(nT_s)$ is greater than the most recent approximation $m_q(nT_s)$, a positive increment $+\Delta$ is applied to the approximation. If, on the other hand, the input signal is smaller, a negative increment $-\Delta$ is applied to the approximation. In this way, the accumulator does the best it can to track the input samples by one step (of amplitude $+\Delta$ or $-\Delta$) at a time. In the receiver shown in Figure 7.29*b*, the staircase approximation $m_q(t)$ is reconstructed by passing the sequence of positive and negative pulses, produced at the decoder output, through an accumulator in a manner similar to that used in the transmitter. The out-of-band quantization noise in the high-frequency staircase waveform $m_q(t)$ is rejected by passing it through a low-pass filter, as in Figure 7.29*b*, with a bandwidth equal to the original message bandwidth.

Delta modulation is subject to two types of quantization error: (1) slope overload distortion and (2) granular noise. We first discuss the cause of slope overload distortion, and then granular noise.

We observe that Eq. (7.53) is the digital equivalent of integration in the sense that it represents the accumulation of positive and negative increments of magnitude Δ. Also, denoting the quantization error by $q(nT_s)$, as shown by,

$$m_q(nT_s) = m(nT_s) + q(nT_s) \tag{7.55}$$

we observe from Eq. (7.51) that the input to the quantizer is

$$e(nT_s) = m(nT_s) - m(nT_s - T_s) - q(nT_s - T_s) \tag{7.56}$$

Thus, except for the quantization error $q(nT_s - T_s)$, the quantizer input is a *first backward difference* of the input signal, which may be viewed as a digital approximation to the derivative of the input signal or, equivalently, as the inverse of the digital integration process. If we consider the maximum slope of the original input waveform $m(t)$, it is clear that in order for the sequence of samples $\{m_q(nT_s)\}$ to increase as fast as the input sequence of samples $\{m(nT_s)\}$ in a region of maximum slope of $m(t)$, we require that the condition

$$\frac{\Delta}{T_s} \geq \max\left|\frac{dm(t)}{dt}\right| \tag{7.57}$$

be satisfied. Otherwise, we find that the step-size Δ is too small for the staircase approximation $m_q(t)$ to follow a steep segment of the input waveform $m(t)$, with the result that $m_q(t)$ falls behind $m(t)$, as illustrated in Figure 7.30. This condition is called *slope overload*, and the resulting quantization error is called *slope-overload distortion* (noise). Note that since the maximum slope of the staircase approximation $m_q(t)$ is fixed by the step-size Δ, increases and decreases in $m_q(t)$ tend to occur along straight lines. For this reason, a delta modulator using a fixed step-size is often referred to as a *linear delta modulator*.

In contrast to slope-overload distortion, *granular noise* occurs when the step-size Δ is too large relative to the local slope characteristics of the input waveform $m(t)$, thereby causing the staircase approximation $m_q(t)$ to hunt around a relatively flat segment of the input waveform; this phenomenon is also illustrated in Figure 7.30. Granular noise is analogous to quantization noise in a PCM system.

We thus see that there is a need to have a large step-size to accommodate a wide dynamic range, whereas a small step-size is required for the accurate representation of relatively low-level signals. It is therefore clear that the choice of the optimum step-size that minimizes the mean-square value of the quantization error in a linear delta modulator will be the result of a compromise between slope overload distortion and granular noise. To improve performance, we need to make the delta modulator "adaptive," in the sense that the step-size is made to vary in accordance with the input signal.

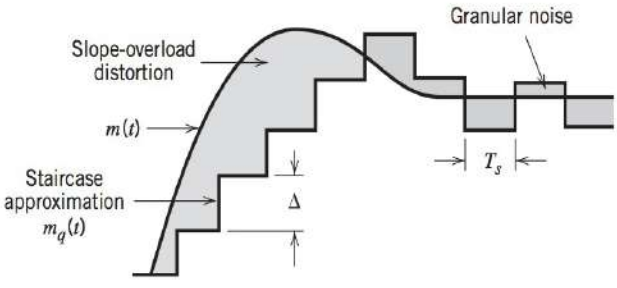


FIGURE 7.30 Illustration of quantization error in delta modulation.

DELTA-SIGMA MODULATION

As mentioned previously, the quantizer input in the conventional form of delta modulation may be viewed as an approximation to the *derivative* of the incoming message signal. This behavior leads to a drawback of delta modulation in that transmission disturbances such as noise result in an accumulative error in the demodulated signal. This drawback can be overcome by *integrating* the message signal prior to delta modulation. The use of integration in the manner described here has also the following beneficial effects:

- The low-frequency content of the input signal is pre-emphasized.
- Correlation between adjacent samples of the delta modulator input is increased, which tends to improve overall system performance by reducing the variance of the error signal at the quantizer input.
- Design of the receiver is simplified.

A delta modulation scheme that incorporates integration at its input is called *delta-sigma modulation* (D-ΣM). To be more precise, however, it should be called *sigma-delta modulation*, because the integration is in fact performed before the delta modulation. Nevertheless, the former terminology is the one commonly used in the literature.

Figure 7.31*a* shows the block diagram of a delta-sigma modulation system. In this diagram, the message signal $m(t)$ is defined in its continuous-time form, which means that the pulse modulator now consists of a hard-limiter followed by a multiplier; the latter component is also fed from an external pulse generator (clock) to produce a 1-bit encoded

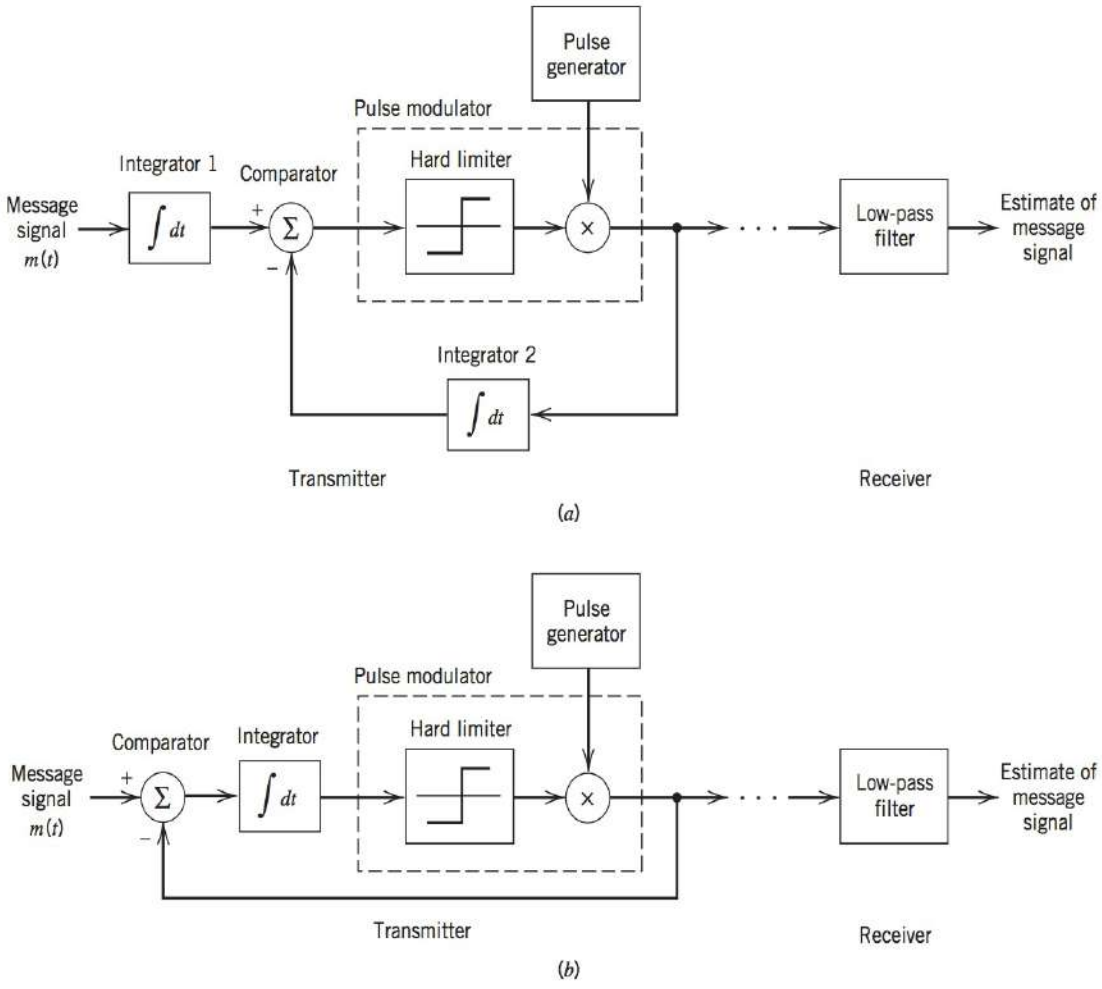


FIGURE 7.31 Two equivalent versions of delta-sigma modulation system.

signal. The use of integration at the transmitter input clearly requires an inverse signal emphasis, namely, differentiation, at the receiver. The need for this differentiation is, however, eliminated because of its cancellation by integration in the conventional DM receiver. Thus, the receiver of a delta-sigma modulation system consists simply of a low-pass filter, as indicated in Figure 7.31*a*.

Moreover, we note that integration is basically a linear operation. Accordingly, we may simplify the design of the transmitter by combining the two integrators 1 and 2 of Figure 7.31*a* into a single integrator placed after the comparator, as shown in Figure 7.31*b*. This latter form of the delta-sigma modulation system is not only simpler than that of Figure 7.31*a*, but it also provides an interesting interpretation of delta-sigma modulation as a "smoothed" version of 1-bit pulse-code modulation. The term *smoothness* refers to the fact that the comparator output is integrated prior to quantization, and the term *1-bit* merely restates that the quantizer consists of a hard-limiter with only two representation levels.

The reason for investigating delta modulation is its reduced bandwidth requirements compared to PCM. For telephone applications, a typical PCM system may use an 8-kHz sampling rate with an 8-bit representation for an overall binary symbol rate of 64 kHz. As we have seen, delta modulation uses a 1-bit representation but oversamples the signal. Typical oversampling rates for delta modulation range from 16 to 32 kHz depending on the desired voice quality. Thus, delta modulation may provide a net bandwidth saving of 50 to 75 percent over PCM at the expense of a more complicated implementation.

There are other methods for converting analog sources, particularly voice, to a digital representation. For example, *differential pulse-code modulation* is a technique that is a combination of delta modulation and PCM as described above. All of the techniques considered here, except delta modulation, consider the samples of the analog source to be independent of one other. More advanced voice coding, or *vocoder* techniques, make use of the correlated nature of voice samples to reduce the bandwidth requirements even further. These techniques often process a 10 to 30 millisecond *voice spurt*, and construct a model of the voice spurt that may be communicated to the other end with just a few bits. In this way, the standard 64 kbps requirement of PCM can be reduced to as little as 2.4 kbps depending upon the voice quality that is acceptable.

7.11 THEME EXAMPLE—DIGITIZATION OF VIDEO AND MPEG[6]

As the preceding discussion indicated, when a source produces correlated samples, a digitization scheme which takes advantage of this correlation can be more efficient than one that does not. In this theme example, we consider how an analog video source may be efficiently converted to a digital representation, in preparation for digital transmission.

Modern video-compression techniques offer a way to represent video in an efficient and robust manner. On a simplistic level, video can be represented by three dimensions. Two dimensions are spatial and represent a still image, while the third dimension is temporal and represents how the image evolves with time. In practice, the still image is actually represented further by three dimensions referred to as the luminance (brightness) and two chrominance (color) components [similar to the three red-green-blue (RGB) components of an analog video signal]. The MPEG standard takes advantage of the high degree of spatial and temporal correlation that is expected in a video signal in order to reduce the number of bits required to represent the signal.

What we describe in the following is a simplistic interpretation of the complex processing that occurs in the MPEG-1 video compression standard. This interpretation is displayed in the block diagram of Figure 7.32. The processing that is shown in Figure 7.32 is applied to each of the three luminance and chrominance components separately.

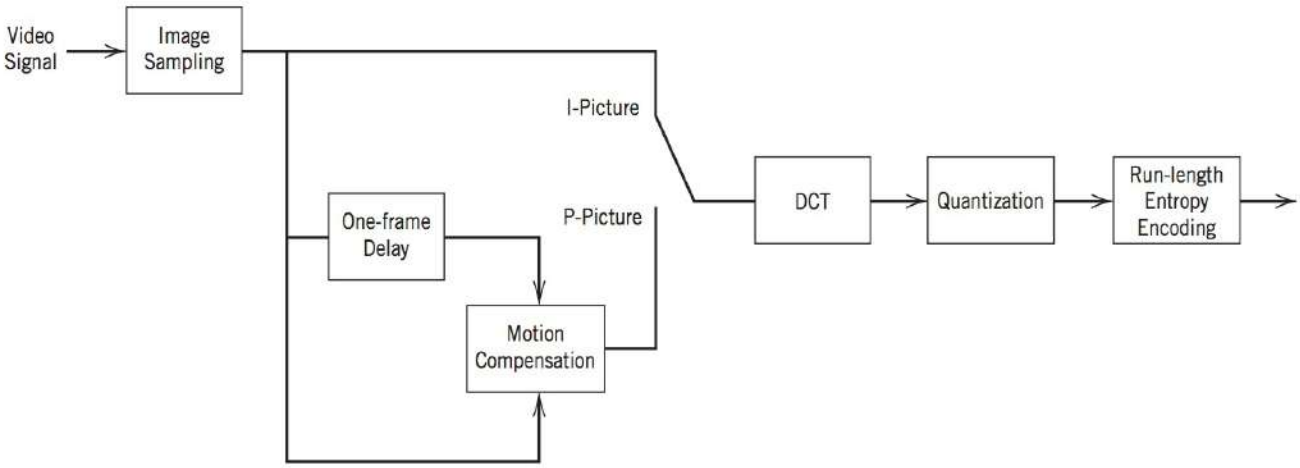


FIGURE 7.32 Simplified block diagram of video signal processing.

The first step is sampling the video signal. In contrast to sampling a voice signal, which consists of a single sample per unit time, a sample of a video signal is an $N \times M$ matrix of picture elements or *pels*[7] per unit time corresponding to a complete still image. This matrix sample is referred to as a *video frame*. In fact, three such matrices must be obtained, one for each of the luminance and two chrominance components.

As with voice signals, the quality of the reconstructed signal depends on the sample rate or frame rate. However, this signal quality must also be traded against the bandwidth required to transmit or store the signal. The MPEG standard takes advantage of the fact that the human eye is less sensitive to changes in chrominance than in luminance and uses a lower frame rate for chrominance signals. Typical frame rates for luminance signals may range 15 to 60 per second and those for chrominance signals may be one-quarter of this value. At the receiver, a decoder uses interpolation[8] to construct the missing chrominance samples and recreate the video signal.

The MPEG-encoding algorithm encodes the first frame in a video sequence in the *intraframe coding mode* (I-picture). Each subsequent frame is coded using *interframe prediction* (P-picture)—where data from the previously coded I- or P-frame are used for prediction.

For the first frame (I-picture), the *discrete cosine transform* (DCT) is applied to each 8×8 block of pels. Mathematically, this two-dimensional transform is defined by

$$X(k_1, k_2) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} x(m_1, m_2) \cos\left(\frac{k_1 \pi}{M_1}(m_1 + 0.5)\right) \cos\left(\frac{k_2 \pi}{M_2}(m_2 + 0.5)\right) \qquad (7.58)$$

where (m_1,m_2) are the coordinates in the spatial domain, and (k_1,k_2) are the coordinates in the transform domain. For an 8×8 transform, $M_1 = M_2 = 8$ and, k_1, and k_2 range from 0 to 7. The DCT is closely related to the *discrete Fourier transform* and the reason for applying the DCT is that it accurately identifies the spatial correlation of the pels within the 8×8 block, similar to how a Fourier transform identifies the frequency components of a signal. The output of the two-dimensional DCT may look something like that shown in Figure 7.33. Due to the high spatial correlation, only a small number of the DCT coefficients are significant, typically those near the (0,0) element.

The coefficients for each DCT block are then quantized, and then only the resulting non-zero coefficients (and their position) are transmitted for each 8×8 block of the first frame. The method for encoding the quantized coefficients and their position uses an advanced data-compression technique known as *run-length entropy encoding* (see Chapter 10).

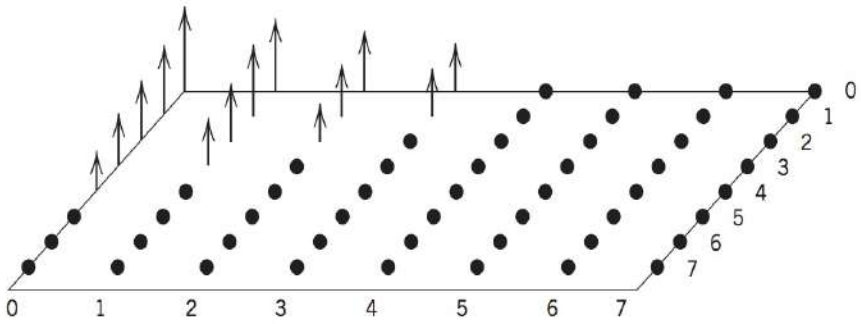


FIGURE 7.33 Example 8 × 8 discrete cosine transform.

For coding subsequent P-pictures, the previous I- or P-picture frame is stored. The first step in this process is to identify correlation in the time domain. Correlation in the time domain corresponds to motion of a pel or a block of pels in the same direction. To identify this motion, the picture is broken down into 16 × 16 pel *macroblocks*. The stored image and the new image are correlated and a motion vector is identified for each 16 × 16 macroblock.

A motion-compensated prediction error is calculated by subtracting each pel in a macroblock from its motion-shifted counterpart in the previous frame. The processing of the prediction errors then follows the same steps as the first picture. A DCT of each 8 × 8 block is performed, and the results are quantized. Since the temporal correlation is expected to be high, the prediction errors are small and most of the quantized DCT coefficients are zero. For the P-pictures, only the motion vectors for each macroblock and the small number of non-zero DCT coefficients have to be transmitted. The transmitted data for P-pictures is typically much less than for I-pictures.

The MPEG standards −1, −2, −3, and so on, provide features to further reduce the bit rate and robustness of the transmission. For example:

- The transmitter may indicate that there is no change to a particular macroblock and therefore there is no need to transmit the corresponding information.
- Over a poor transmission channel, the transmitter may transmit I-pictures at regular intervals to compensate for any accumulation of transmission errors at the decoder.

From this simple description, it should be clear that there are many ways to adjust the video quality and the corresponding bandwidth required for transmission. These include the choice of $N \times M$ picture resolution, the frame rate, and the quantization used for the DCT coefficients. In this way, the family of MPEG standards offers varying video qualities with bit rates ranging from 64 kbps to 10 Mbps.

This theme example illustrates many of the techniques that we have described in this and previous chapters including signal analysis and processing using Fourier transforms, sampling and quantization theory, as well the efficient representation of data by using differential encoding (prediction) techniques.

7.12 SUMMARY AND DISCUSSION

In this chapter, we introduced two fundamental and complementary processes:

- *Sampling*, which operates in the time domain. The sampling process is the link between an analog waveform and its discrete-time representation.
- *Quantization*, which operates in the amplitude domain. The quantization process is the link between an analog waveform and its discrete amplitude representation.

The sampling process builds on the *sampling theorem*, which states that a strictly band-limited signal with no frequency components higher than W Hz is represented uniquely by a sequence of samples taken at a uniform rate greater than the Nyquist rate of $2W$ samples per second. The quantization process exploits the fact that any human sense, as ultimate receiver, can only detect finite intensity differences.

The sampling process is the first step in converting an analog information source into a digital representation. *Analog pulse modulation* is the natural way of representing the sampled signal. The distinguishing feature between analog and digital pulse modulation is that analog pulse modulation maintains a continuous amplitude representation of the information, whereas *digital pulse modulation* also employs quantization to provide a representation of the message signal that is discrete in both time and amplitude.

Analog pulse modulation results from varying some parameter of a pulse, such as amplitude, duration, or position, in which case we speak of pulse-amplitude modulation (PAM), pulse-duration modulation (PDM), or pulse-position modulation (PPM), respectively. Despite the fact that PPM is more efficient than both PPM and PAM when used for transmitting the information, it still falls short of the ideal system for exchanging transmission bandwidth for improved noise performance.

Digital pulse modulation represents analog information sources as a sequence of quantized pulses which is made possible through the combined use of sampling and quantization. Pulse-code modulation, where the quantized signal is often represented as a binary code word, is a common method of representing analog signals such as voice and video. Delta modulation is a second useful method of digitally representing an analog source that has the advantage of reduced bandwidth requirements.

In a strict sense, the term "pulse modulation" is a misnomer in that in all of its different forms, be they analog or digital, are in fact *source coding* techniques, that is, methods of digitally representing analog information. The resulting digital representation is a baseband signal and, hence, can be transmitted over a baseband channel of adequate bandwidth. Indeed, this is the reason that these techniques are often considered transmission schemes as well as source encoding schemes.

It is also important to recognize that pulse modulation techniques are *lossy* in the sense that some information is lost as a result of the signal representation that they perform. For example, in pulse-amplitude modulation, the customary practice is to use prealias (low-pass) filtering prior to sampling; in so doing, information is lost by virtue of the fact that high-frequency components considered to be unessential are removed by the filter. The lossy nature of pulse modulation is most vividly seen in pulse-code modulation that is characterized by the generation of quantization noise (i.e., distortion). The sequence of encoded pulses does not have the infinite precision needed to represent continuous samples exactly. Nevertheless, the loss of information incurred by the use of a pulse modulation process is *under the designer's control* in that it can be made small enough for it not to be discernible by the end user.

This chapter has been an introduction to methods of digitally representing analog sources. In Chapter 8, we consider methods for transmitting this digital information over baseband channels.

NOTES AND REFERENCES

1. For a detailed treatment of mathematically modeling impulse radio, see the paper by Win and Scholtz (1998).
2. The book by Jayant and Noll (1984) presents the most complete treatment of pulse-code modulation, differential pulse-code modulation, delta modulation, and their variants.

3. The μ-law used for signal compression is described in Smith (1957). The μ-law is used in the United States, Canada, and Japan. In Europe, the A-law is used for signal compression. This compression law is described in Cattermole (1969, pp. 133–140). For a discussion of the μ-law and A-law, see also the paper by Kaneko (1970).
4. For more information on optical fiber transmission and SONET, see Keiser (2000).
5. See the book by Jayant and Noll (1984) for further discussion of the various forms of delta modulation.
6. A more detailed description of the MPEG coding standards may be found in Sikora (1997).
7. Initially, the term pixel was the abbreviation for a picture element. Now the term pel is more commonly used.
8. Linear interpolation is a well-known method of estimating interim samples. More complex interpolation strategies can have beneficial effects such as reducing the harm caused by noise. See Crochiere and Rabiner (1983).

PROBLEMS

7.1 A narrowband signal has a bandwidth of 10 kHz centered on a carrier frequency of 100 kHz. It is proposed to represent this signal in discrete-time form by sampling its in-phase and quadrature components individually. What is the minimum sampling rate that can be used for this representation? Justify your answer. How would you reconstruct the original narrowband signal from the sampled versions of its in-phase and quadrature components?

7.2 In *natural sampling*, an analog signal $g(t)$ is multiplied by a periodic train of rectangular pulses $c(t)$. Given that the pulse repetition frequency of this periodic train is f_s and the duration of each rectangular pulse is T (with $f_sT \gg 1$), do the following:

(a) Find the spectrum of the signal $s(t)$ that results from the use of natural sampling. You may assume that time $t = 0$ corresponds to the midpoint of a rectangular pulse in $c(t)$.
(b) Show that the original signal $m(t)$ may be recovered exactly from its naturally sampled version, provided that the conditions embodied in the sampling theorem are satisfied.

7.3 Specify the Nyquist rate and the Nyquist interval for each of the following signals:

(a) $g(t) = \text{sinc}(200t)$
(b) $g(t) = \text{sinc}^2(200t)$
(c) $g(t) = \text{sinc}(200t) + \text{sinc}^2(200t)$

7.4

(a) Plot the spectrum of a PAM wave produced by the modulating signal

$$m(t) = A_m \cos(2\pi f_m t)$$

assuming a modulation frequency $f_m = 0.25$ Hz, sampling period $T_s = 1$s, and pulse duration $T = 0.45$ s.
(b) Using an ideal reconstruction filter, plot the spectrum of the filter output. Compare this result with the output that would be obtained if there were no aperture effect.

7.5 In this problem we evaluate the equalization needed for the aperture effect in a PAM system. The operating frequency is $f = f_s/2$, which corresponds to the highest frequency component of the message signal for a sampling rate equal to the Nyquist rate. Plot $1/\text{sinc}(0.5T/T_s)$, versus T/T_s, and hence find the equalization needed when $T/T_s = 0.1$.

7.6 Consider a PAM wave transmitted through a channel with white Gaussian noise and bandwidth $B_T = 1/2T_s$, where T_s is the sampling period. The noise is of zero mean and power spectral density $N_0/2$. The PAM signal uses a standard pulse $g(t)$ with its Fourier transform defined by

$$G(f) = \begin{cases} \dfrac{1}{2B_T}, & |f| < B_T \\ 0, & |f| > B_T \end{cases}$$

By considering a full-load sinusoidal modulating wave, show that PAM and baseband-signal transmission have equal signal-to-noise ratios for the same average transmitted power.

7.7 Twenty-four voice signals are sampled uniformly and then time-division multiplexed. The sampling operation uses flat-top samples with 1 μs duration. The multiplexing operation includes provision for synchronization by adding an extra pulse of sufficient amplitude and also 1 μs duration. The highest frequency component of each voice signal is 3.4 kHz.

(a) Assuming a sampling rate of 8 kHz, calculate the spacing between successive pulses of the multiplexed signal.
(b) Repeat your calculation assuming the use of Nyquist rate sampling.

7.8 Twelve different message signals, each with a bandwidth of 10 kHz, are to be multiplexed and transmitted. Determine the minimum bandwidth required for each method if the multiplexing/modulation method used is

(a) FDM, SSB.
(b) TDM, PAM.

7.9 A PAM *telemetry* system involves the multiplexing of four input signals: $s_i(t)$, $i = 1, 2, 3, 4$. Two of the signals $s_1(t)$ and $s_2(t)$

have bandwidths of 80 Hz each, whereas the remaining two signals $s_3(t)$ and $s_4(t)$ have bandwidths of 1 kHz each. The signals $s_3(t)$ and $s_4(t)$ are each sampled at the rate of 2400 samples per second. This sampling rate is divided by 2^R (i.e., an integer power of 2) in order to derive the sampling rate for $s_1(t)$ and $s_2(t)$.

(a) Find the maximum value of R.
(b) Using the value of R found in part (a), design a multiplexing system that first multiplexes $s_1(t)$ and $s_2(t)$ into a new sequence, $s_5(t)$, and then multiplexes $s_3(t)$, $s_4(t)$, and $s_5(t)$.

7.10 The unmodulated pulse train in a PPM system is as shown in Figure P7.10. The slicing level in the receiver is set at $A/2$.

(a) Assuming a full-load sinusoidal modulating wave and front-end receiver noise of zero mean and power spectral density $N_0/2$, determine the output signal-to-noise ratio and figure of merit of the system. Assume a high peak pulse-to-noise ratio.
(b) For the case when the message signal is sampled at its Nyquist rate, find the value of the transmission bandwidth for which the figure of merit of the system is greater than unity.

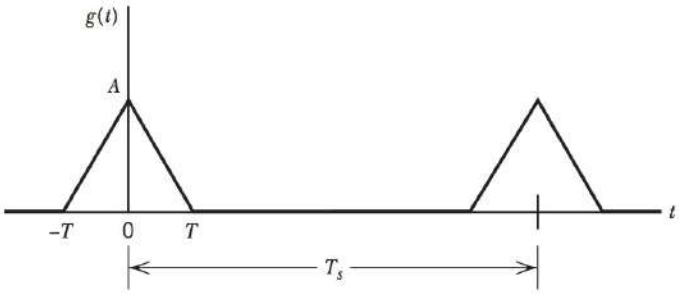


Figure P7.10

7.11 Consider Figure P7.11, which shows the unmodulated pulse train for PDM. The PDM pulse consists of a rectangular pulse of duration D, which is preceded and followed by leading and trailing segments that are identical to the corresponding halves of the PPM pulse shown in Figure 7.15. The slicer in the receiver is set to half the peak pulse amplitude, removing all noise effects except for the displacement of edge detection time by a small amount τ similar to that evaluated for the PPM system in Example 7.1. Assume that one edge of the duration-modulation pulse is fixed by means of a noise-free reference.

(a) Find the output signal-to-noise ratio of the PDM system.
(b) Find its channel signal-to-noise ratio.
(c) Compare the figure of merit for the PDM system to that of the corresponding PPM system.

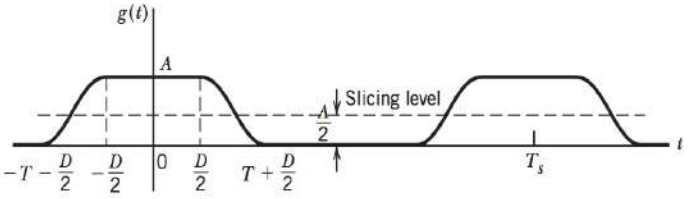


Figure P7.11

7.12
(a) A sinusoidal signal, with an amplitude of 3.25 volts, is applied to a uniform quantizer of the midtread type whose output takes on the values 0, ±1, ±2, ±3 volts. Sketch the waveform of the resulting quantizer output for one complete cycle of the input.
(b) Repeat this evaluation for the case when the quantizer is of the midrise type whose output takes on the values ±0.5, ±1.5, ±2.5, ±3.5 volts.

7.13 Consider the following sequences of 1s and 0s:

(a) An alternating sequence of 1s and 0s.
(b) A long sequence of 1s followed by a long sequence of 0s.
(c) A long sequence of 1s followed by a single 0 and then a long sequence of 1s.

Sketch the waveform for each of these sequences using the following methods of representing symbols 1 and 0:

(1) On-off signaling.
(2) Bipolar return-to-zero signaling.

7.14 The signal

$$m(t) = 6\sin(2\pi t) \text{ volts}$$

is transmitted using a 4-bit binary PCM system. The quantizer is of the midrise type, with a step-size of 1 volt. Sketch the resulting PCM wave for one complete cycle of the input. Assume a sampling rate of four samples per second, with samples taken at $t = \pm 1/8, \pm 3/8, \pm 5/8, \ldots$, seconds.

7.15 Figure P7.15 shows a PCM signal in which the amplitude levels of +1 volt and −1 volt are used to represent binary symbols 1 and 0, respectively. The code word used consists of three bits. Find the sampled version of an analog signal from which this binary signal is derived.

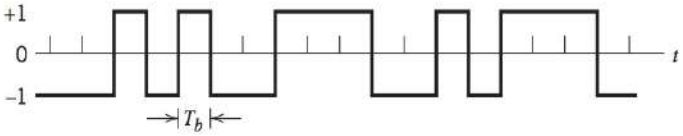


Figure P7.15

7.16 Consider a uniform quantizer characterized by the input–output relation illustrated in Figure 7.21*a*. Assume that a Gaussian-distributed random variable with zero mean and unit variance is applied to this quantizer input.

(a) What is the probability that the amplitude of the input lies outside the range −4 to +4?
(b) Using the result of part (a), show that the output signal-to-noise ratio of the quantizer is given by

$$(\text{SNR})_O = 6R - 7.2\text{dB}$$

where R is the number of bits per sample. Specifically, you may assume that the quantizer input extends from −4 to +4. Compare the result of part (b) with that obtained in Example 7.2.

7.17 A PCM system that uses a uniform quantizer is followed by a 7-bit binary encoder. The bit rate of the system is equal to 50×10^6 b/s.

(a) What is the maximum message bandwidth for which the system operates satisfactorily?

(b) Determine the output signal-to-quantization noise ratio when a full-load sinusoidal modulating wave of frequency 1 MHz is applied to the input.

7.18 Show that, with a nonuniform quantizer, the mean-square value of the quantization error is approximately equal to $(1/12)\Sigma_i \Delta_i^2 p_i$ where Δ_i is the ith step size and p_i is the probability that the input signal amplitude lies within the ith interval. Assume that the step-size Δ_i is small compared with the excursion of the input signal.

7.19 Consider a chain of $(n-1)$ regenerative repeaters, with a total of n sequential decisions made on a binary wave, including the final decision made at the receiver. Assume that any binary symbol transmitted through the system has an independent probability p_1 of being inverted by any repeater. Let p_n represent the probability that a binary symbol is in error after transmission through the complete system.

(a) Show that

$$p_n = \tfrac{1}{2}[1 - (1 - 2p_1)^n]$$

(b) If p_1 is very small and n is not too large, what is the corresponding value of p_n?

7.20 Consider a test signal $m(t)$ defined by a hyperbolic tangent function:

$$m(t) = A \tanh(\beta t)$$

where A and β are constants. Determine the minimum step-size Δ for delta modulation of this signal which is required to avoid slope overload.

7.21 Consider a sine wave of frequency f_m and amplitude A_m, applied to a delta modulator of step-size Δ. Show that slope-overload distortion will occur if

$$A_m > \frac{\Delta}{2\pi f_m T_s}$$

where T_s is the sampling period. What is the maximum power that may be transmitted without slope-overload distortion?

7.22 A linear delta modulator is designed to operate on speech signals limited to 3.4 kHz. The specifications of the modulator are as follows:

- Sampling rate $= 10 f_{Nyquist}$, where $f_{Nyquist}$ is the Nyquist rate of the speech signal.
- Step-size $\Delta = 100$ mV.

The modulator is tested with a 1-kHz sinusoidal signal. Determine the maximum amplitude of this test signal permissible to avoid slope overload.

Computer Problems

7.23 The signal

$$s(t) = \sin(400\,\pi t) + 0.5\cos(12000\,\pi t)$$

is sampled at a 10 kHz rate.

(a) What is the spectrum of the sampled signal?

(b) Use the following Matlab script to simulate the sampled spectrum. Explain the results

```
Fs = 10;                %Sample rate (kHz)
Ts = 1/Fs;              %Sample period (ms)

t = [0: Ts: 100];       %Observation period (ms)
s = sin(2*pi*2*t)+ 0.5*cos(2*pi*6*t);

FFTsize = 1024;
spec = fftshift(abs(fft(s,FFTsize)).^2);
freq = [-Fs/2 : Fs/FFTsize: Fs/2];
freq = freq(1:end-1);
plot(freq,spec)
xtabel('Frequency(kHz)')
ylabel('Amplitude Spectrum')
```

(c) How does the sampled spectrum change if the sample rate is changed to 11 kHz? Why?

7.24

(a) Determine the mathematical expression for the expanding portion of a μ-law compander.

(b) The signal

$$s(t) = 10\exp(-t) + \sin(2\pi t)$$

is sampled at an 20 Hz rate over the interval from 0 to 20 seconds. The signal is then quantized.

(i) If 8-bit quantization is performed without companding, determine the root-mean-square (rms) error between the unquantized and quantized signals.

(ii) If 8-bit quantization is perform with a μ-law compander with $\mu = 255$, determine the rms error between the unquantized signal and quantized signals after expansion.

Use the following Matlab script as a guide for computing the mean-square error.

```
Fs = 20;        % Sample rate (Hz)
Ts = 1/Fs;      % Sample period (s)

t = [0: Ts: 20]; % Observation period (s)
s = 10*exp(-t) + sin(2*pi*t);

% - - - Compression - - - -
mu = 100;
s = s/max(abs(s)+eps);      % normalize signal level
s_mu = log(1+mu*abs(s))/log(1+mu).* sign(s);
                            % mu-law compression

% - - - Quantization - - - -
Q = 8;                              % number of bits of quantization
s_mu_q = floor(2^(Q-1)*s_mu);   % non-uniform
                    % quantization to 256 levels (-128...127)
```

```
% --- Expansion ------
s_mu_r = (exp(log(1+mu)*abs(s_mu_q)/2^(Q-1))-1)/...
   mu.* sign(s_mu_q);

% --- Compare --------
plots(s-s_mu_r)
rms_mu = sqrt(mean((s-s_mu_r).^2))
```

(c) How do the results of part (b) change if we shorten or lengthen the observation period? Why? What does this imply about companding?

(d) How do the results of part (b) change if we decrease or increase the number of quantization levels?

7.25 Use the Matlab scripts of the previous two problems to compute the spectrum of quantized and unquantized versions of the following signal

$$s(t) = \sin(2\pi t) + 0.5\cos(\pi t/2)$$

Assume a 20 Hz sampling rate with 8-bit uniform quantization and 20 second observation window. Describe the differences between the spectra of the quantized and unquantized signals. How do the results change if 6-bit quantization is used instead of 8-bit? 4-bit? Suggest what may be done to the quantized samples to more closely approximate the original signal spectrum.

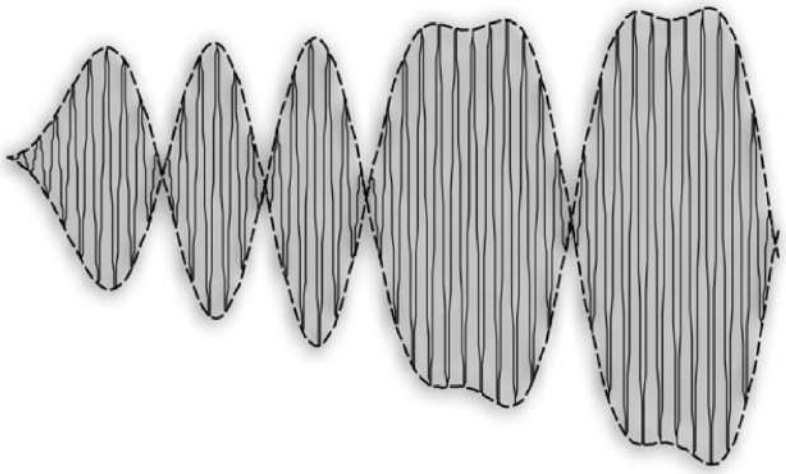
7.26 Delta modulation is applied to the signal of Problem 7.25. The sample rate is increased by a factor of four to $F_s = 80$ Hz. Using the Matlab scripts of previous problems as a guide, construct a Matlab script to determine the delta-modulation representation of the signal, and to reconstruct the signal.

(a) What step-size minimizes the rms error between the reconstructed signal and the unquantized signal?

(b) What is the rms error with delta modulation over an observation window of 20 seconds? Compare this to the rms error with uniform quantization and $F_s = 20$ Hz.

(c) For what sample rate does the rms error of delta modulation approximately equal that of uniform sampling? Comment on the type of signals where delta modulation would be most suitable.

(d) What difficulties would a delta modulator have with the signal of Problem 7.24?



8

BASEBAND DIGITAL TRANSMISSION

8.1 INTRODUCTION

In the previous chapter, we described techniques for converting analog information, such as voice and video, into digital form. Other information, such as text, computer files and programs, have a natural digital (usually binary) representation. In this chapter we study the transmission of digital data (of whatever origin) over a *baseband channel.*[1] Data transmission over a band-pass channel using modulation is covered in the next chapter.

Digital data have a broad spectrum with a significant low-frequency content. Baseband transmission of digital data therefore requires the use of a low-pass channel with a bandwidth large enough to accommodate the essential frequency content of the data stream. Typically, however, the channel is *dispersive* in that its frequency response deviates from that of an ideal low-pass filter. The result of data transmission over such a channel is that each received pulse is affected somewhat by adjacent pulses, thereby giving rise to a common form of interference called *intersymbol interference* (ISI). Intersymbol interference is a major source of bit errors in the reconstructed data stream at the receiver. To correct for it, control has to be exercised over the pulse shape in the overall system. Thus much of the material covered in this chapter is devoted to *pulse shaping* in one form or another.

Another source of bit errors in a baseband data transmission system is the ubiquitous *receiver noise* (channel noise). Naturally, noise and ISI arise in the system simultaneously. However, to understand how they affect the performance of the system, we propose to consider them separately. We thus begin the chapter by describing a fundamental result in communication theory, which deals with the *detection* of a known waveform that is immersed in additive white noise. The device for the optimum detection of such a pulse involves the use of a linear-time-invariant filter known as a *matched filter,*[2] which is so called because its impulse response is matched to the pulse signal.

8.2 BASEBAND PULSES AND MATCHED FILTER DETECTION

In Chapter 7, we presented a number of line codes for transmitting binary data. Examples included on-off signaling and bipolar return-to-zero signaling. These line codes were introduced in the context of PCM but may be used with any stream of binary data. Each line code has its advantages and disadvantages, but they can be characterized in general as different forms of baseband pulses. In Figure 8.1 we show the power spectra of the several line codes introduced earlier. Note that:

- These power spectra are for long sequences of random bits with a given line code where 0 and 1 are equally probable. The resulting signal forms a random process and the computation of the power spectrum relies on the techniques introduced in Chapter 5. Example 5.1 is one instance. For the analytical computation of the power spectra of Figure 8.1, the reader is referred to Problem 8.3.
- The frequency axes of the power spectra have been normalized with respect to T_b denoting, the bit period, and the average power is normalized to unity. The nominal bandwidth of the signal is the same order of magnitude as $1/T_b$ and is centered around the origin.

For the moment, we shall assume that the frequency response of the channel is relatively ideal and has little effect on the transmitted pulse shape. That is, the transmitted signal spectra shown in Figure 8.1 are unchanged at the receiver. Transmission at a low data rate over a short cable is an example of such an ideal baseband channel but there are other situations where this model is applicable. In this ideal case, the transmitted pulse $g(t)$ for each bit is unaffected by the transmission except for the addition of white noise at the receiver front end, as illustrated in Figure 8.2. This illustrates the basic problem of detection—a pulse transmitted over a channel that is corrupted by additive noise at the front end of the receiver.

MATCHED FILTER

Consider the receiver model shown in Figure 8.1, involving a linear time-invariant filter of impulse response $h(t)$. The filter input $x(t)$ consists of a pulse signal $g(t)$ corrupted by additive noise $w(t)$, as shown by

$$x(t) = g(t) + w(t) \qquad 0 \le t \le T \tag{8.1}$$

where T is an arbitrary observation interval. The pulse signal $g(t)$ may represent a binary symbol 1 or 0 in a digital communication system. The $w(t)$ is the sample function of a white noise process of zero mean and power spectral density $N_0/2$. It is assumed that the receiver has knowledge of the waveform of the pulse signal $g(t)$. The source of uncertainty lies in the noise $w(t)$. The function of the receiver is to detect the pulse signal $g(t)$ in an optimum manner, given the received signal $x(t)$. To satisfy this requirement, we have to optimize the design of the filter so as to minimize the effects of noise at the filter output, and thereby enhance the detection of the pulse signal $g(t)$.

Since the filter is linear, the resulting output $y(t)$ may be expressed as

$$y(t) = g_o(t) + n(t) \tag{8.2}$$

where $g_o(t)$ and $n(t)$ are produced by the signal and noise components of the input $x(t)$, respectively. A simple way of describing the requirement that the output signal component $g_o(t)$ be considerably greater than the output noise component $n(t)$ is to have the filter make the instantaneous power in the output signal $g_o(t)$, measured at time $t = T$, as large as possible compared with the average power of the output noise $n(t)$. This is

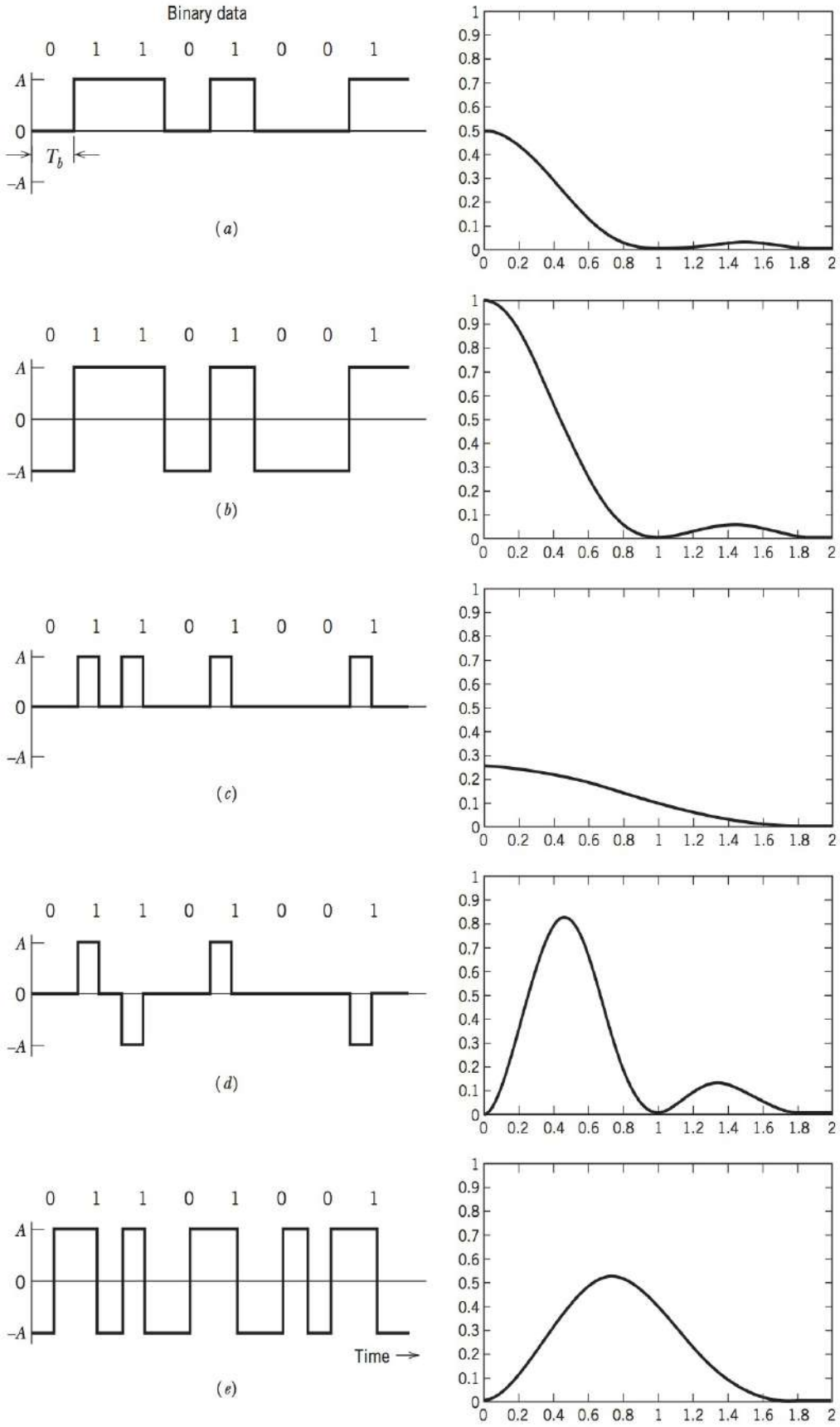


FIGURE 8.1 (a) Unipolar NRZ line code and its amplitude spectrum. (b) Polar NRZ line code and its amplitude spectrum. (c) Unipolar RZ line code and its amplitude spectrum. (d) Bipolar RZ line code and its amplitude spectrum. (E) Manchester line code and its amplitude spectrum.

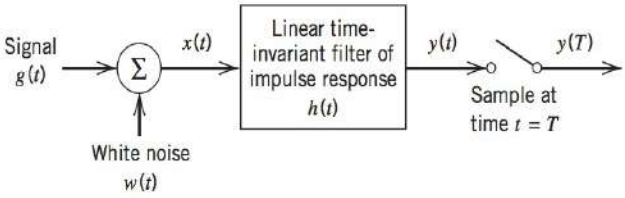


FIGURE 8.2 Linear receiver.

equivalent to maximizing the *peak pulse signal-to-noise ratio*, defined as

$$\eta = \frac{|g_o(T)|^2}{\mathbf{E}[n^2(t)]} \tag{8.3}$$

where $|g_o(T)|^2$ is the instantaneous power in the output signal, $\mathbf{E}$ is the statistical expectation operator, and $\mathbf{E}[n^2(t)]$ is a measure of the average output noise power. The requirement is to specify the impulse response $h(t)$ of the filter such that the output signal-to-noise ratio in Eq. (8.3) is maximized.

Let $G(f)$ denote the Fourier transform of the known signal $g(t)$, and $H(f)$ denote the transfer function of the filter. Then the Fourier transform of the output signal $g_o(t)$ is equal to $H(f)\,G(f)$, and $g_o(t)$ is itself given by the inverse Fourier transform

$$g_o(t) = \int_{-\infty}^{\infty} H(f)G(f)\exp(j2\pi ft)\,df \tag{8.4}$$

Hence, when the filter output is sampled at time $t = T$, we have (in the absence of receiver noise)

$$|g_o(T)|^2 = \left|\int_{-\infty}^{\infty} H(f)G(f)\exp(j2\pi fT)\,df\right|^2 \tag{8.5}$$

Consider next the effect on the filter output due to the noise $w(t)$ acting alone. The power spectral density $S_N(f)$ of the output noise $n(t)$ is equal to the power spectral density of the input noise $w(t)$ times the squared magnitude of the transfer function $H(f)$ (see Section 4.10). Since $w(t)$ is white with constant power spectral density $N_0/2$, it follows that

$$S_N(f) = \frac{N_0}{2}|H(f)|^2 \tag{8.6}$$

The average power of the output noise $n(t)$ is therefore

$$\begin{aligned}\mathbf{E}[n^2(t)] &= \int_{-\infty}^{\infty} S_N(f)\,df \\ &= \frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2\,df\end{aligned} \tag{8.7}$$

Thus, substituting Eqs. (8.5) and (8.7) into (8.3), we may rewrite the expression for the peak pulse signal-to-noise ratio as

$$\eta = \frac{\left|\int_{-\infty}^{\infty} H(f)G(f)\exp(j2\pi T)\,df\right|^2}{\frac{N_0}{2}\int_{-\infty}^{\infty}|H(f)|^2 df} \tag{8.8}$$

Our problem is to find, for a given $G(f)$, the particular form of the transfer function $H(f)$ of the filter that makes η a maximum. To find the solution to this optimization problem, we apply a mathematical result known as Schwarz's inequality to the numerator of Eq. (8.8).

Schwarz's inequality states that if we have two complex functions $\phi_1(x)$ and $\phi_2(x)$ in the real variable x, satisfying the conditions

$$\int_{-\infty}^{\infty} |\phi_1(x)|^2\,dx < \infty$$

and

$$\int_{-\infty}^{\infty} |\phi_2(x)|^2 \, dx < \infty$$

then we may write

$$\left| \int_{-\infty}^{\infty} \phi_1(x)\phi_2(x) \, dx \right|^2 \leq \int_{-\infty}^{\infty} |\phi_1(x)|^2 \, dx \int_{-\infty}^{\infty} |\phi_2(x)|^2 \, dx \tag{8.9}$$

The equality in (8.9) holds if, and only if, we have

$$\phi_1(x) = k\phi_2^*(x) \tag{8.10}$$

where k is an arbitrary constant, and the asterisk denotes complex conjugation.

Returning to the problem at hand, we readily see that by invoking Schwarz's inequality (8.9), and setting $\phi_1(x) = H(f)$ and $\phi_2(x) = G(f)\exp(j\pi fT)$, the numerator in Eq. (8.8) may be rewritten as

$$\left| \int_{-\infty}^{\infty} H(f)G(f)\exp(j2\pi fT) \, df \right|^2 \leq \int_{-\infty}^{\infty} |H(f)|^2 \, df \int_{-\infty}^{\infty} |G(f)|^2 \, df \tag{8.11}$$

Using this relation in Eq. (8.8), we may redefine the peak pulse signal-to-noise ratio as

$$\eta \leq \frac{2}{N_0} \int_{-\infty}^{\infty} |G(f)|^2 \, df \tag{8.12}$$

The right-hand side of this relation does not depend on the transfer function $H(f)$ of the filter but only on the signal energy and the noise power spectral density. Consequently, the peak pulse signal-to-noise ratio η will be a maximum when $H(f)$ is chosen so that the equality holds; that is,

$$\eta_{\max} = \frac{2}{N_0} \int_{-\infty}^{\infty} |G(f)|^2 \, df \tag{8.13}$$

Correspondingly, $H(f)$ assumes its optimum value denoted by $H_{\text{opt}}(f)$. To find this optimum value we use Eq. (8.10), which, for the situation at hand, yields

$$H_{\text{opt}}(f) = kG^*(f)\exp(-j2\pi fT) \tag{8.14}$$

where $G^*(f)$ is the complex conjugate of the Fourier transform of the input signal $g(t)$, and k is a scaling factor of appropriate dimensions. This relation states that, except for the factor $k \exp(-j2\pi fT)$, the transfer function of the optimum filter is the same as the complex conjugate of the spectrum of the input signal.

Equation (8.14) specifies the optimum filter in the frequency domain. To characterize it in the time domain, we take the inverse Fourier transform of $H_{\text{opt}}(f)$ in Eq. (8.14) to obtain the impulse response of the optimum filter as

$$h_{\text{opt}}(t) = k \int_{-\infty}^{\infty} G^*(f)\exp[-j2\pi f(T-t)] \, df \tag{8.15}$$

Since for a real signal $g(t)$ we have $G^*(f) = G(-f)$, we may rewrite Eq. (8.15) as

$$\begin{aligned} h_{\text{opt}}(t) &= k \int_{-\infty}^{\infty} G(-f)\exp[-j2\pi f(T-t)] \, df \\ &= kg(T-t) \end{aligned} \tag{8.16}$$

Equation (8.16) shows that the impulse response of the optimum filter, except for the scaling factor k, is a time-reversed and delayed version of the input signal $g(t)$; that is, it is "matched" to the input signal. A linear time-invariant filter defined in this way is called a *matched filter*. Note that in deriving the matched filter the only assumption we have

made about the input noise $w(t)$ is that it is stationary and white with zero mean and power spectral density $N_0/2$.

PROPERTIES OF MATCHED FILTERS

We note that a filter, which is matched to a pulse signal $g(t)$ of duration T, is characterized by an impulse response that is a time-reversed and delayed version of the input $g(t)$, as shown by

$$h_{\text{opt}}(t) = kg(T - t)$$

In other words, the impulse response $h_{\text{opt}}(t)$ is uniquely defined, except for the delay T and the scaling factor k, by the waveform of the pulse signal $g(t)$ to which the filter is matched. In the frequency domain, the matched filter is characterized by a transfer function that is, except for a delay factor, the complex conjugate of the Fourier transform of the input $g(t)$, as shown by

$$H_{\text{opt}}(f) = kG^*(f)\exp(-j2\pi fT)$$

The most important result in the calculation of the performance of signal processing systems using matched filters is perhaps the following:

- *The peak pulse signal-to-noise ratio of a matched filter depends only on the ratio of the signal energy to the power spectral density of the white noise at the filter input.*

To demonstrate this property, consider a filter matched to a known signal $g(t)$. The Fourier transform of the resulting matched filter output $g_o(t)$ is

$$\begin{aligned} G_o(f) &= H_{\text{opt}}(f)G(f) \\ &= kG^*(f)G(f)\exp(-j2\pi fT) \\ &= k|G(f)|^2\exp(-j2\pi fT) \end{aligned} \tag{8.17}$$

Using Eq. (8.17) in the formula for the inverse Fourier transform, we find that the matched filter output at time $t = T$ is

$$\begin{aligned} g_o(T) &= k\int_{-\infty}^{\infty} G_o(f)\exp(j2\pi fT)\,df \\ &= k\int_{-\infty}^{\infty} |G(f)|^2\,df \end{aligned}$$

Using Rayleigh's energy theorem, this result reduces to

$$g_o(T) = kE \tag{8.18}$$

where E is the energy of the pulse signal $g(t)$. Next, substituting Eq. (8.14) in (8.7), we find that the average output noise power is

$$\begin{aligned} \mathbf{E}[n^2(t)] &= \frac{k^2N_0}{2}\int_{-\infty}^{\infty} |G(f)|^2 df \\ &= k^2N_0E/2 \end{aligned} \tag{8.19}$$

where again we have made use of Rayleigh's energy theorem. Therefore, the peak pulse signal-to-noise ratio has the maximum value

$$\eta_{\max} = \frac{(kE)^2}{(k^2N_0E/2)} = \frac{2E}{N_0} \tag{8.20}$$

From Eq. (8.20) we see that dependence on the waveform of the input $g(t)$ has been completely removed by the matched filter. Accordingly, in evaluating the ability of a

matched-filter receiver to combat additive white Gaussian noise, we find that all signals that have the same energy are equally effective. Note that the signal energy E is in joules and the noise spectral density $N_0/2$ is in watts per hertz, so that the ratio $2E/N_0$ is dimensionless; however, the two quantities have different physical meaning. We refer to E/N_0 as the *signal energy-to-noise spectral density ratio*.

EXAMPLE 8.1 Matched Filter for Rectangular Pulse

Consider a signal $g(t)$ in the form of a rectangular pulse of amplitude A and duration T, as shown in Figure 8.3*a*. In this example, the impulse response $h(t)$ of the matched filter has exactly the same waveform as the signal itself. The output signal $g_o(t)$ of the matched filter produced in response to the input signal $g(t)$ has a triangular waveform, as shown in Figure 8.3*b*.

The maximum value of the output signal $g_o(t)$ is equal to kA^2T, which is the energy of the input signal $g(t)$ scaled by the factor k; this maximum value occurs at $t = T$, as indicated in Figure 8.3*b*.

For the special case of a rectangular pulse, the matched filter may be implemented using a circuit known as the *integrate-and-dump circuit*, a block diagram of which is shown in Figure 8.4. The integrator computes the area under the rectangular pulse, and the resulting output is then sampled at time $t = T$, where T is the duration of the pulse. Immediately after $t = T$, the integrator is restored to its initial condition; hence the name of the circuit. Figure 8.3*c* shows the output waveform of the integrate-and-dump circuit for the rectangular pulse of Figure 8.3*a*. We see that for $0 \leq t \leq T$, the output of this circuit has the *same waveform* as that appearing at the output of the matched filter.

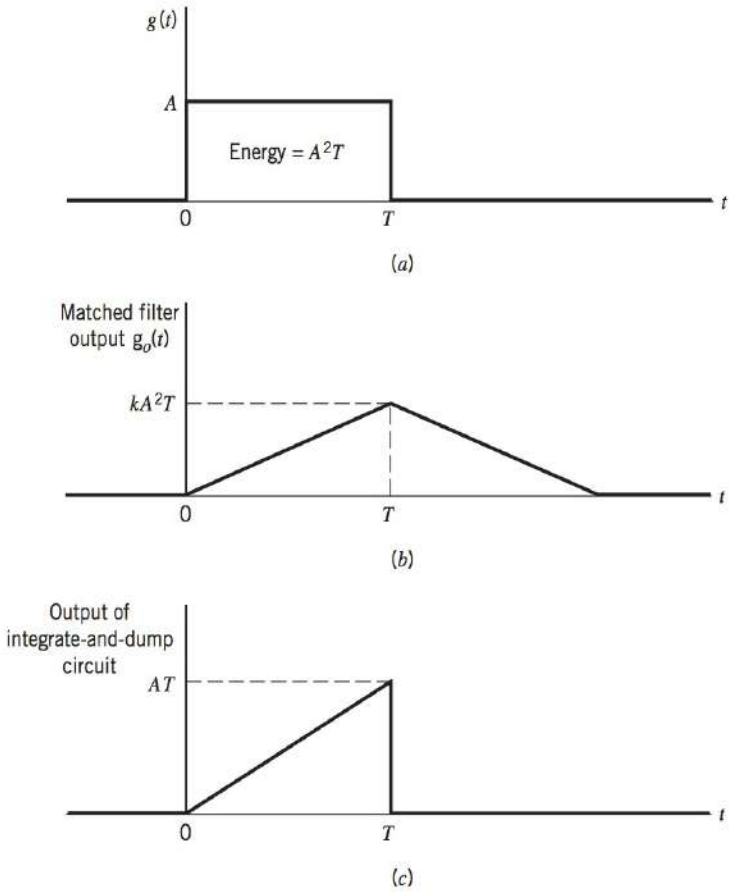


FIGURE 8.3 (*a*) Rectangular pulse. (*b*) Matched filter output. (*c*) Integrator output.

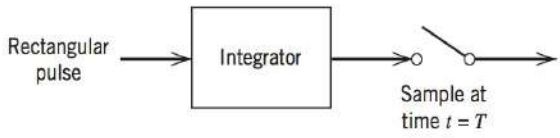


FIGURE 8.4 Integrate-and-dump circuit.

8.3 PROBABILITY OF ERROR DUE TO NOISE

Now that we are equipped with the matched filter as the optimum detector of a known pulse in additive white noise, we are ready to derive a formula for the error rate in such a system due to noise.

To proceed with the analysis, consider a binary transmission system based on polar *nonreturn-to-zero (NRZ) signaling*. In this form of signaling, symbols 1 and 0 are represented by positive and negative rectangular pulses of equal amplitude and equal duration. The noise is modeled as *additive white Gaussian noise* $w(t)$ of zero mean and power spectral density $N_0/2$; the Gaussian assumption is needed for later calculations. In the signaling interval $0 \le t \le T_b$, the received signal is thus written as follows:

$$x(t) = \begin{cases} +A + w(t), & \text{symbol 1 was sent} \\ -A + w(t), & \text{symbol 0 was sent} \end{cases} \tag{8.21}$$

where T_b is the *bit duration*, and A is the *transmitted pulse amplitude*. It is assumed that the receiver has acquired knowledge of the starting and ending times of each transmitted pulse; in other words, the receiver has prior knowledge of the pulse shape, but not its polarity. Given the noisy signal $x(t)$, the receiver is required to make a decision in each signaling interval as to whether the transmitted symbol is a 1 or a 0.

The structure of the receiver used to perform this decision-making process is shown in Figure 8.5. It consists of a matched filter followed by a sampler, and then finally a decision device. The filter is matched to a rectangular pulse of amplitude A and duration T_b, exploiting the bit-timing information available to the receiver. The resulting matched filter output is sampled at the end of each signaling interval. The presence of receiver noise $w(t)$ adds randomness to the matched filter output.

Let y denote the sample value obtained at the end of a signaling interval. The sample value y is compared to a preset *threshold* λ in the decision device. If the threshold is exceeded, the receiver makes a decision in favor of symbol 1; if not, a decision is made in favor of symbol 0. We adopt the convention that when the sample value y is exactly equal to the threshold λ, the receiver just makes a guess as to which symbol was transmitted; such a decision is the same as obtained by flipping a fair coin, the outcome of which will not alter the average probability of error.

There are two possible kinds of error to be considered:

1. Symbol 1 is chosen when a 0 was actually transmitted; we refer to this error as an *error of the first kind.*
2. Symbol 0 is chosen when a 1 was actually transmitted; we refer to this error as an *error of the second kind.*

To determine the average probability of error, we consider these two situations separately.

Suppose that symbol 0 was sent. Then, according to Eq. (8.21) the received signal is

$$x(t) = -A + w(t), \qquad 0 \le t \le T_b \tag{8.22}$$

Correspondingly, the matched filter output, sampled at time $t = T_b$, is given by (in light of Example 8.1 with kAT_b set equal to unity for convenience of presentation)

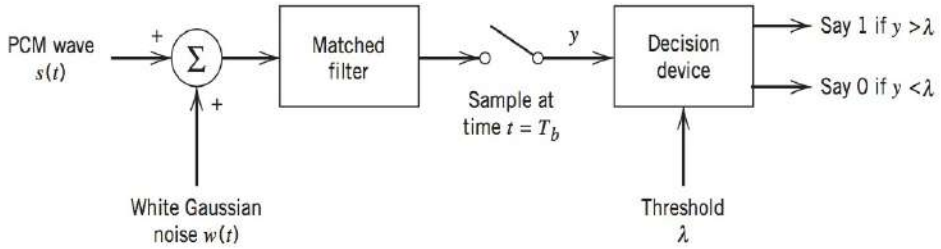


FIGURE 8.5 Receiver for baseband transmission of binary wave using NRZ signaling.

$$\begin{aligned} y &= \int_0^{T_b} x(t)\,dt \\ &= -A + \frac{1}{T_b}\int_0^{T_b} w(t)\,dt \end{aligned} \tag{8.23}$$

which represents the sample value of a random variable Y. By virtue of the fact that the noise $w(t)$ is white and Gaussian, we may characterize the random variable Y as follows:

- The random variable Y is Gaussian distributed with a mean of $-A$.
- The variance of the random variable Y is

$$\begin{aligned} \sigma_Y^2 &= \mathbf{E}[(Y+A)^2] \\ &= \frac{1}{T_b^2}\mathbf{E}\left[\int_0^{T_b}\int_0^{T_b} w(t)w(u)\,dt\,du\right] \\ &= \frac{1}{T_b^2}\int_0^{T_b}\int_0^{T_b} \mathbf{E}[w(t)w(u)]\,dt\,du \\ &= \frac{1}{T_b^2}\int_0^{T_b}\int_0^{T_b} R_W(t,u)\,dt\,du \end{aligned} \tag{8.24}$$

where $R_W(t,u)$ is the autocorrelation function of the white noise $w(t)$. Since $w(t)$ is white with a power spectral density $N_0/2$, we have

$$R_W(t,u) = \frac{N_0}{2}\delta(t-u) \tag{8.25}$$

where $\delta(t-u)$ is a time-shifted Dirac delta function. Hence, substituting Eq. (8.25) in (8.24) yields

$$\begin{aligned} \sigma_Y^2 &= \frac{1}{T_b^2}\int_0^{T_b}\int_0^{T_b} \frac{N_0}{2}\delta(t-u)\,dt\,du \\ &= \frac{N_0}{2T_b} \end{aligned} \tag{8.26}$$

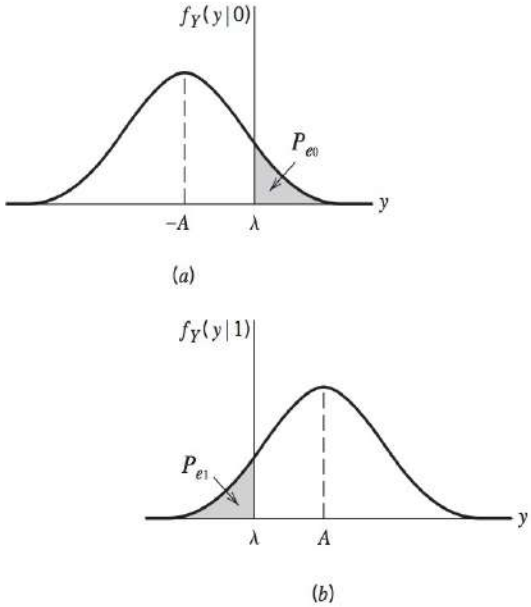


FIGURE 8.6 Analysis of the effect of channel noise on a binary system. (*a*) Probability density function of random variable Y at matched filter output when a 0 is transmitted. (*b*) Probability density function of Y when a 1 is transmitted.

The probability density function of the random variable Y, given that symbol 0 was sent, is therefore

$$f_Y(y|0) = \frac{1}{\sqrt{\pi N_0/T_b}}\exp\left(-\frac{(y+A)^2}{N_0/T_b}\right) \tag{8.27}$$

This function is shown plotted in Figure 8.6*a*. Let P_{e0} denote the *conditional probability of error, given that symbol 0 was sent*. This probability is defined by the shaded area under the curve of $f_Y(y|0)$ from the threshold λ to infinity, which corresponds to the range of values assumed by y for a decision in favor of symbol 1. In the absence of noise, the matched filter output y sampled at time $t = T_b$ is equal to $-A$. When noise is present, y occasionally assumes a value greater than λ, in which case an error is made. The probability of this error, conditional on sending symbol 0, is defined by

$$\begin{aligned} P_{e0} &= P(y > \lambda|\text{symbol 0 was sent}) \\ &= \int_\lambda^\infty f_Y(y|0)\,dy \\ &= \frac{1}{\sqrt{\pi N_0/T_b}}\int_\lambda^\infty \exp\left(-\frac{(y+A)^2}{N_0/T_b}\right)dy \end{aligned} \tag{8.28}$$

To proceed further we need to assign an appropriate value to the threshold λ. Such an assignment requires knowledge of the *a priori probabilities* of binary symbols 0 and 1, denoted by p_0 and p_1, respectively. It is clear that we must always have

$$p_0 + p_1 = 1 \tag{8.29}$$

In what follows we assume that symbols 0 and 1 occur with equal probability, in which case we have

$$p_0 = p_1 = \frac{1}{2} \tag{8.30}$$

Moreover, in the absence of noise, we note that the sampled value of the matched filter output is $-A$ when symbol 0 is sent, and $+A$ when symbol 1 is sent. Thus, in light of Eq. (8.30) it is reasonable to set the threshold halfway between these two values, that is,

$$\lambda = 0 \tag{8.31}$$

Accordingly, Eq. (8.28) for the conditional probability of error of the first kind takes the form

$$P_{e0} = \frac{1}{\sqrt{\pi N_0/T_b}} \int_0^\infty \exp\left(-\frac{(y+A)^2}{N_0/T_b}\right) dy \tag{8.32}$$

Define a new variable

$$z = \frac{y+A}{\sqrt{N_0/2T_b}} \tag{8.33}$$

Then we may reformulate Eq. (8.32) as

$$P_{e0} = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2E_b/N_0}}^\infty \exp(-z^2/2)\, dz \tag{8.34}$$

where E_b is the *transmitted signal energy per bit*, defined by

$$E_b = A^2 T_b \tag{8.35}$$

At this point we find it convenient to introduce the definition of the so-called Q-function.

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp(-z^2/2)\, dz \tag{8.36}$$

We may finally reformulate the conditional probability of error P_{e0} in terms of the Q-function as follows:

$$P_{e0} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{8.37}$$

Assume next that symbol 1 was transmitted. This time the Gaussian random variable Y represented by the sample value y of the matched filter output has a mean $+A$ and variance $N_0/2T_b$. Note that, compared to the situation when symbol 0 was sent, the mean of the random variable Y has changed but its variance is exactly the same as before. The conditional probability density function of Y, given that symbol 1 was sent, is therefore

$$f_Y(y|1) = \frac{1}{\sqrt{\pi N_0/T_b}} \exp\left(-\frac{(y-A)^2}{N_0/T_b}\right) \tag{8.38}$$

Complementary Error Function

The Q-function for determining the area under the tails of the Gaussian distribution is most commonly used by communication engineers. The Gaussian distribution plays a role in many fields and the *complementary error function* defined by

$$\text{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^\infty \exp(-z^2)\, dz$$

is also often used. The complementary error function is closely related to the Q-function, as shown by

$$Q(u) = \frac{1}{2}\text{erfc}\left(\frac{u}{\sqrt{2}}\right)$$

With this transformation, tables and approximations for erfc(u) may also be used to calculate $Q(u)$.

which is plotted in Figure 8.6*b*. Let P_{e1} denote the *conditional probability of error, given that symbol 1 was sent*. This probability is defined by the shaded area under the curve of $f_Y(y|1)$ extending from $-\infty$ to the threshold λ, which corresponds to the range of values assumed by y for a decision in favor of symbol 0. In the absence of noise, the matched filter output y sampled at time $t = T_b$ is equal to $+A$. When noise is present, y occasionally assumes a value less than λ, and an error is then made. The probability of this error, conditional on sending symbol 1, is defined by

$$\begin{aligned} P_{e1} &= P(y < \lambda | \text{symbol 1 was sent}) \\ &= \int_{-\infty}^{\lambda} f_Y(y|1)\, dy \\ &= \frac{2}{\sqrt{\pi N_0/T_b}} \int_{-\infty}^{\lambda} \exp\left(-\frac{(y-A)^2}{N_0/T_b}\right) dy \end{aligned} \tag{8.39}$$

Setting the threshold $\lambda = 0$, as before, and putting

$$\frac{y - A}{\sqrt{N_0/2T_b}} = -z$$

we readily find that $P_{e1} = P_{e0}$. This result is indeed a consequence of setting the threshold at the midpoint between $-A$ and $+A$, which was justified earlier on the assumption that symbol 0 and 1 are equiprobable. A channel for which the conditional error probabilities P_{e1} and P_{e0} are equal is said to be *binary symmetric*.

To determine the average probability of symbol error in the receiver, we note that the two possible kinds of error considered above are mutually exclusive events in that if the receiver, at a particular sampling instant, chooses symbol 1, then symbol 0 is excluded from appearing, and vice versa. Furthermore, P_{e0} and P_{e1} are conditional probabilities with P_{e0} assuming that symbol 0 was sent and P_{e1} assuming that symbol 1 was sent. Thus, the *average probability of symbol error* P_e in the receiver is given by

$$P_e = p_0 P_{e0} + p_1 P_{e1} \tag{8.40}$$

where p_0 and p_1 are the *a priori* probabilities of binary symbols 0 and 1, respectively. Since $P_{e1} = P_{e0}$, and $p_0 = p_1 = \frac{1}{2}$ in accordance with Eq. (8.30), we obtain

$$P_e = P_{e1} = P_{e0}$$

or

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{8.41}$$

We have thus shown that *the average probability of symbol error with binary signaling depends solely on E_b/N_0, the ratio of the transmitted signal energy per bit to the noise spectral density*.

In Figure 8.7 we have used the formula of Eq. (8.41) to plot the average probability of symbol error P_e versus the dimensionless ratio E_b/N_0. This figure shows that the error probability P_e decreases very rapidly as the ratio E_b/N_0 is increased, so that eventually a very small increase in transmitted signal energy will make the reception of binary pulses almost error free. Note, however, that in practical terms the increase in signal energy has to be viewed in the context of what the bias level is; for example, a 3-dB increase in E_b/N_0 is much easier to implement when E_b has a small value than when its value is orders of magnitude larger.

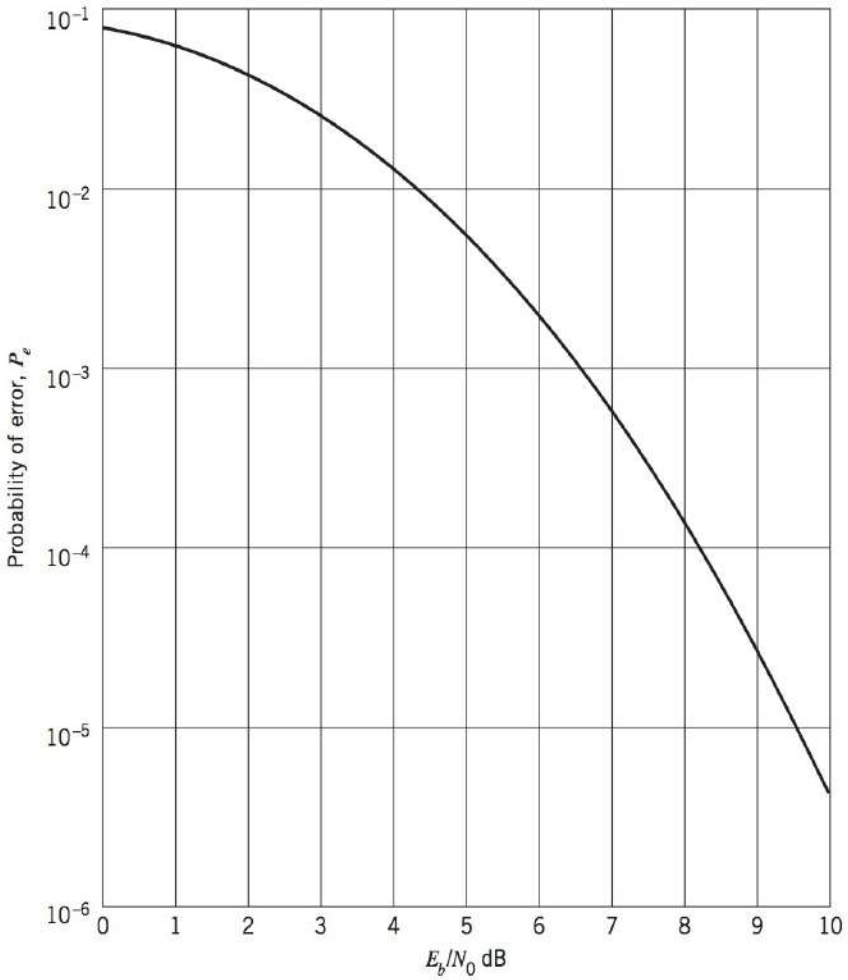


FIGURE 8.7 Probability of error in additive white Gaussian noise with binary signaling.

8.4 INTERSYMBOL INTERFERENCE

The next source of bit errors in a baseband-pulse transmission system that we wish to study is intersymbol interference (ISI), which arises when the communication channel is *dispersive*. When we say a channel is dispersive, we mean the channel has a frequency-dependent amplitude spectrum. In the previous section, we assumed the channel was ideal which means the amplitude spectrum of the channel is a constant in the frequency domain. The simplest example of a dispersive channel is the *band-limited channel*. For example, a brickwall band-limited channel passes all frequencies $|f| < \mathrm{W}$ without distortion, while it blocks all frequencies $|f| > W$. While communication media do not often have this type of abrupt characteristic, the band-limited channel is a good model for many practical situations where many signals must share the communication medium using an FDM strategy, and thus each signal must be limited in bandwidth to avoid interfering with signals adjacent in frequency.

First of all, however, we need to address a key question: Given a pulse shape of interest, how do we use it to transmit data in M-ary form? The answer lies in the use of *discrete pulse modulation*, in which the amplitude, duration, or position of the transmitted pulses is varied in a discrete manner in accordance with the given data stream. However, for the baseband transmission of digital data, the use of *discrete pulse-amplitude modulation* (PAM) is the most efficient one in terms of power and bandwidth utilization. Accordingly, we confine our attention to discrete PAM systems. We begin the study by first considering the case of binary data; later in the chapter, we consider the more general case of M-ary data.

Consider then a *baseband binary PAM system*, a generic form of which is shown in Figure 8.8. The incoming binary sequence $\{b_k\}$ consists of symbols 1 and 0, each of duration T_b. The *pulse-amplitude modulator* transforms this binary sequence into a new sequence of short pulses (approximating a unit impulse), whose amplitude a_k is represented in the polar form

$$a_k = \begin{cases} +1 & \text{if symbol } b_k \text{ is 1} \\ -1 & \text{if symbol } b_k \text{ is 0} \end{cases} \tag{8.42}$$

The sequence of short pulses so produced is applied to a *transmit filter* of impulse response $g(t)$, producing the transmitted signal (see Section 2.11).

$$s(t) = \sum_k a_k g(t - kT_b) \tag{8.43}$$

The signal $s(t)$ is modified as a result of transmission through the *channel* of impulse response $h(t)$. In addition, the channel adds random noise to the signal at the receiver input. The noisy signal $x(t)$ is then passed through a *receive filter* of impulse response $c(t)$. The resulting filter output $y(t)$ is sampled *synchronously* with the transmitter, with the sampling instants being determined by a *clock* or *timing signal* that is usually extracted from the receive-filter output. Finally, the sequence of samples thus obtained is used to reconstruct the original data sequence by means of a *decision device*. Specifically, the amplitude of each sample is compared to a *threshold* λ. If the threshold λ is exceeded, a decision is made in favor of symbol 1. If the threshold λ is not exceeded, a decision is made in favor of symbol 0. If the sample amplitude equals the threshold exactly, the flip of a fair coin will determine which symbol was transmitted (i.e., the receiver simply makes a guess).

The receive filter output is written as

$$y(t) = \mu \sum_k a_k p(t - kT_b) + n(t) \tag{8.44}$$

where μ is a scaling factor, and the pulse $p(t)$ is to be defined. To be precise, an arbitrary time delay t_0 should be included in the argument of the pulse $p(t - kT_b)$ in Eq. (8.44) to represent the effect of transmission delay through the system. To simplify the exposition, we have put this delay equal to zero in Eq. (8.44) without loss of generality.

The scaled pulse $\mu p(t)$ is obtained by a double convolution involving the impulse response $g(t)$ of the transmit filter, the impulse response $h(t)$ of the channel, and the impulse response $c(t)$ of the receive filter, as shown by

$$\mu p(t) = g(t) \star h(t) \star c(t) \tag{8.45}$$

where the star denotes convolution. We assume that the pulse $p(t)$ is *normalized* by setting

$$p(0) = 1 \tag{8.46}$$

which justifies the use of μ as a scaling factor to account for amplitude changes incurred in the course of signal transmission through the system.

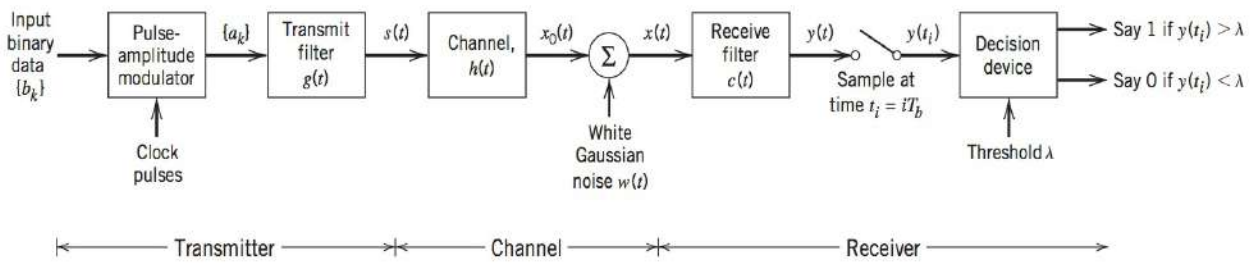


FIGURE 8.8 Baseband binary data transmission system.

Since the convolution in the time domain is transformed into multiplication in the frequency domain, we may use the Fourier transform to change Eq. (8.45) into the equivalent form

$$\mu P(f) = G(f)H(f)C(f) \tag{8.47}$$

where $P(f)$, $G(f)$, $H(f)$, and $C(f)$ are the Fourier transforms of $p(t)$, $g(t)$, $h(t)$, and $c(t)$, respectively.

Finally, the term $n(t)$ in Eq. (8.44) is the noise produced at the output of the receive filter due to the additive noise $w(t)$ at the receiver input. It is customary to model $w(t)$ as a white Gaussian noise of zero mean.

The receive filter output $y(t)$ is sampled at time $t_i = iT_b$ (with i taking on integer values), yielding [in light of Eq. (8.46)]

$$\begin{aligned} y(t_i) &= \mu \sum_{k=-\infty}^{\infty} a_k p[(i-k)T_b)] + n(t_i) \\ &= \mu a_i + \mu \sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} a_k p[(i-k)T_b)] + n(t_i) \end{aligned} \tag{8.48}$$

In Eq. (8.48), the first term μa_i represents the contribution of the ith transmitted bit. The second term represents the residual effect of all other transmitted bits on the decoding of the ith bit; this residual effect due to the occurrence of pulses before and after the sampling instant t_i is called intersymbol interference (ISI). The last term $n(t_i)$ represents the noise sample at time t_i.

In the absence of both ISI and noise, we observe from Eq. (8.48) that

$$y(t_i) = \mu a_i$$

which shows that, under these ideal conditions, the ith transmitted bit is decoded correctly. The unavoidable presence of ISI and noise in the system, however, introduces errors in the decision device at the receiver output. Therefore, in the design of the transmit and receive filters, the objective is to minimize the effects of noise and ISI and thereby deliver the digital data to its destination with the smallest error rate possible.

When the signal-to-noise ratio is high, as is the case in a telephone system, for example, the operation of the system is largely limited by ISI rather than noise; in other words, we may ignore $n(t_i)$. In the next couple of sections, we assume that this condition holds so that we may focus our attention on ISI and the techniques for its control. In particular, we wish to determine the pulse waveform $p(t)$ for which the ISI is completely eliminated. Before doing this, we consider an ISI example and a method for characterising ISI.

EXAMPLE 8.2 The Dispersive Nature of a Telephone Channel

A baseband communication channel that readily comes to mind for data transmission is a telephone channel. The telephone channel is generally characterized by a high signal-to-noise ratio. However, the channel is *band-limited*, as illustrated in Figure 8.9 for a typical toll connection. Figure 8.9*a* shows the insertion loss of the channel plotted against frequency; *insertion loss* (in dB) is defined as $10 \log_{10} (P_0/P_2)$ where P_2 is the power delivered to a load by the channel, and P_0 is the power delivered to the same load when it is connected directly to the source (i.e., the channel is removed). Figure 8.9*b* shows the corresponding plots of the phase response and envelope (group) delay versus frequency; for the definition of envelope delay, see Section 2.11. Figure 8.9 clearly illustrates the *dispersive* nature of the telephone channel.

This dispersive nature often leads to intersymbol interference. For illustration, consider choosing one of the aforementioned line codes for

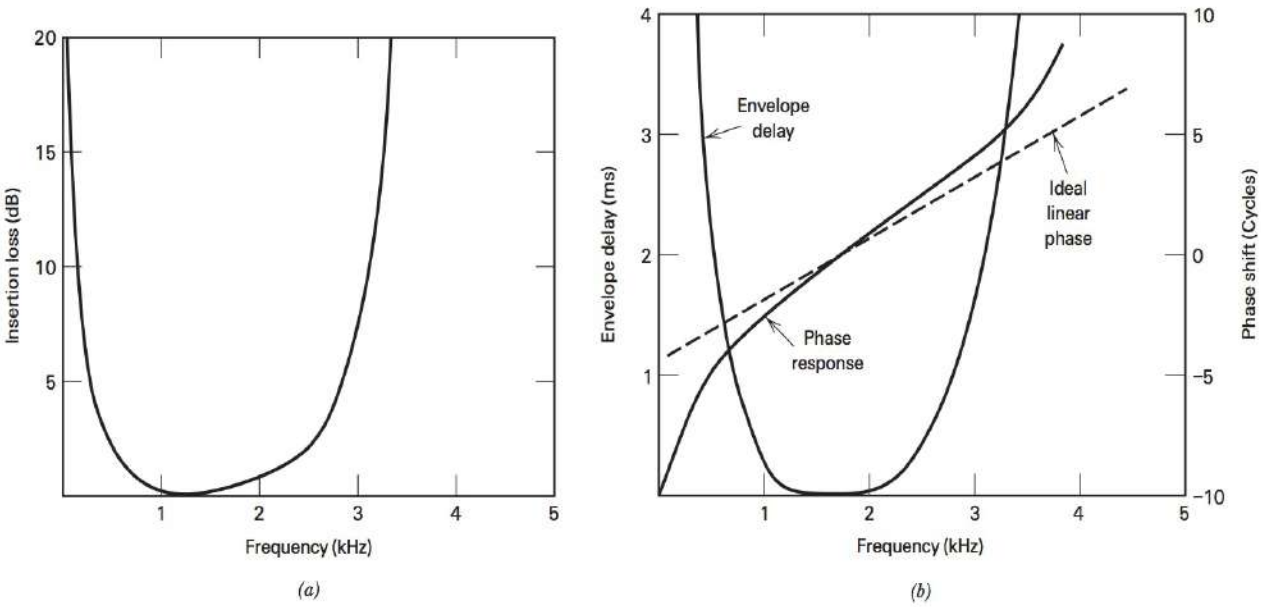


FIGURE 8.9 (*a*) Amplitude response of typical toll connection. (*b*) Envelope delay and phase response of typical toll connection. (Bellamy, 1982).

transmitting data over this channel. Before choosing the line code, we note two properties of the telephone channel of Figure 8.9:

- The pass-band of the channel cuts off rapidly above 3.5 kHz. This fact suggests the use of a line code with a narrow spectrum, so that we can maximize the data rate.
- The channel does not pass *dc*. From this fact, it may be preferable to use a line code that has no *dc* component such as bipolar (RZ) signaling or the Manchester code.

These two properties suggest two contradictory choices for the line code: (a) the polar NRZ, which has a narrow spectrum; and (b) the Manchester code, which has no *DC* component. In Figure 8.10*a* we show the original data signal using the polar NRZ line code and the resulting signal after it has passed through the channel in the absence of noise when the bit rate is 1600 bits per second (bps). There is clearly distortion of the signal. In particular,

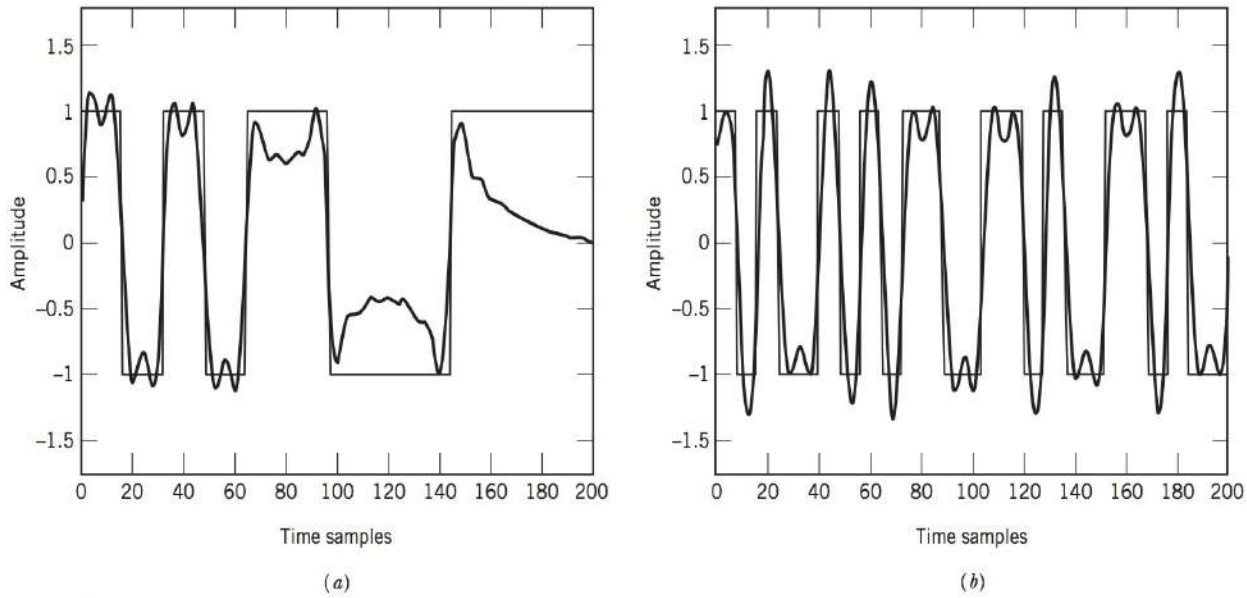


FIGURE 8.10 Data transmission over a telephone channel at 1600 bps: (*a*) polar NRZ line code and (*b*) Manchester line code.

long strings of the symbols of the same polarity cause the signal to drift toward zero volts. This drift is due to the fact that the channel does not pass *dc*; however, the transmitted signal is still recognizable. When using the Manchester line code at 1600 bps, this drift problem is not evident but there is a small amount of signal distortion.

In Figure 8.11, we provide the results for the same two line codes but at a rate of 3200 bps. With the faster rate, the *dc* drift with the polar NRZ line code is less evident but there is still noticeable signal distortion. With the Manchester line code, there is considerable distortion of the signal compared to the original. This is attributed to the fact that the spectrum of the 3200 bps Manchester line code has significant spectral content outside the band limits of this telephone channel.

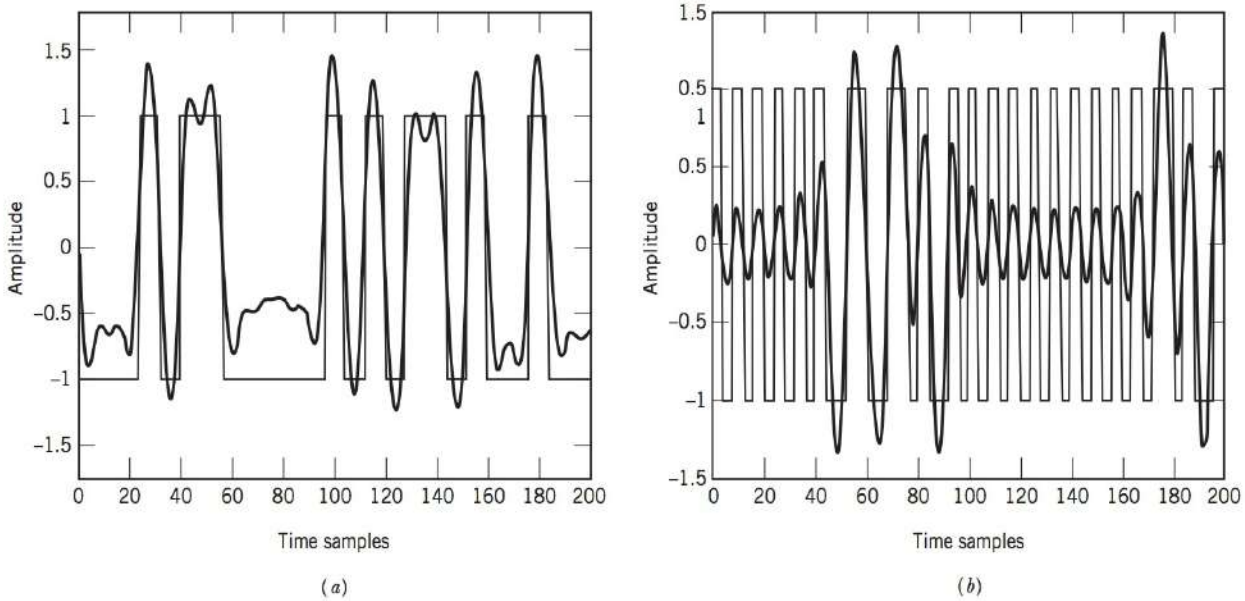


FIGURE 8.11 Data transmission over a telephone channel at 3200 bps: (*a*) polar NRZ line code and (*b*) Manchester line code.

8.5 EYE PATTERN

The discussion of Section 8.4 and Example 8.1 have illustrated qualitatively the effects of intersymbol interference (ISI) on the performance of a baseband pulse-transmission system. An operational tool for evaluating the effects of ISI in an insightful manner is the so-called *eye pattern*. The eye pattern is defined as the synchronized superposition of all possible realizations of the signal of interest (e.g., received signal, receiver output) viewed within a particular signaling interval. The eye pattern derives its name from the fact that it resembles the human eye for binary waves. The interior region of the eye pattern is called the *eye opening*.

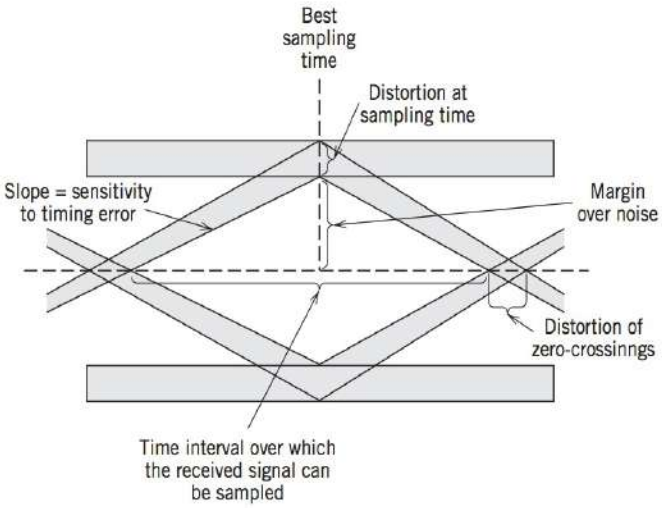


FIGURE 8.12 Interpretation of the eye pattern.

An eye pattern provides a great deal of useful information about the performance of a data transmission system, as described in Figure 8.12. Specifically, we may make the following statements:

- The width of the eye opening defines the *time interval over which the received signal can be sampled without error from intersymbol interference.* It is apparent that

the preferred time for sampling is the instant of time at which the eye is open the widest.

- The *sensitivity of the system to timing errors* is determined by the rate of closure of the eye as the sampling time is varied.
- The height of the eye opening, at a specified sampling time, defines the *noise margin* of the system.

When the effect of intersymbol interference is severe, traces from the upper portion of the eye pattern cross traces from the lower portion, with the result that the eye is completely closed. In such a situation, it is impossible to avoid errors due to the presence of intersymbol interference in the system.

In the case of an M-ary system, the eye pattern contains $(M - 1)$ eye openings stacked up vertically one on the other, where M is the number of discrete amplitude levels used to construct the transmitted signal. In a strictly linear system with truly random data, all these eye openings would be identical. In practice, however, it is often possible to discern asymmetries in the eye pattern, which are caused by nonlinearities in the communication channel.

EXAMPLE 8.3 Eye Diagrams for Binary and Quaternary Systems

Figures 8.13*a* and 8.13*b* show the eye diagrams for a simulated baseband PAM transmission system using $M = 2$ and $M = 4$, respectively. The channel has no bandwidth limitation, and the source symbols used are randomly generated. A pulse having a raised cosine spectrum is used in both cases. We will say more about this type of pulse in the following section. In both cases, we see that the eyes are open, indicating reliable operation of the system. In fact, at the ideal sampling point, there is no intersymbol interference, which is one of the properties of this pulse shape.

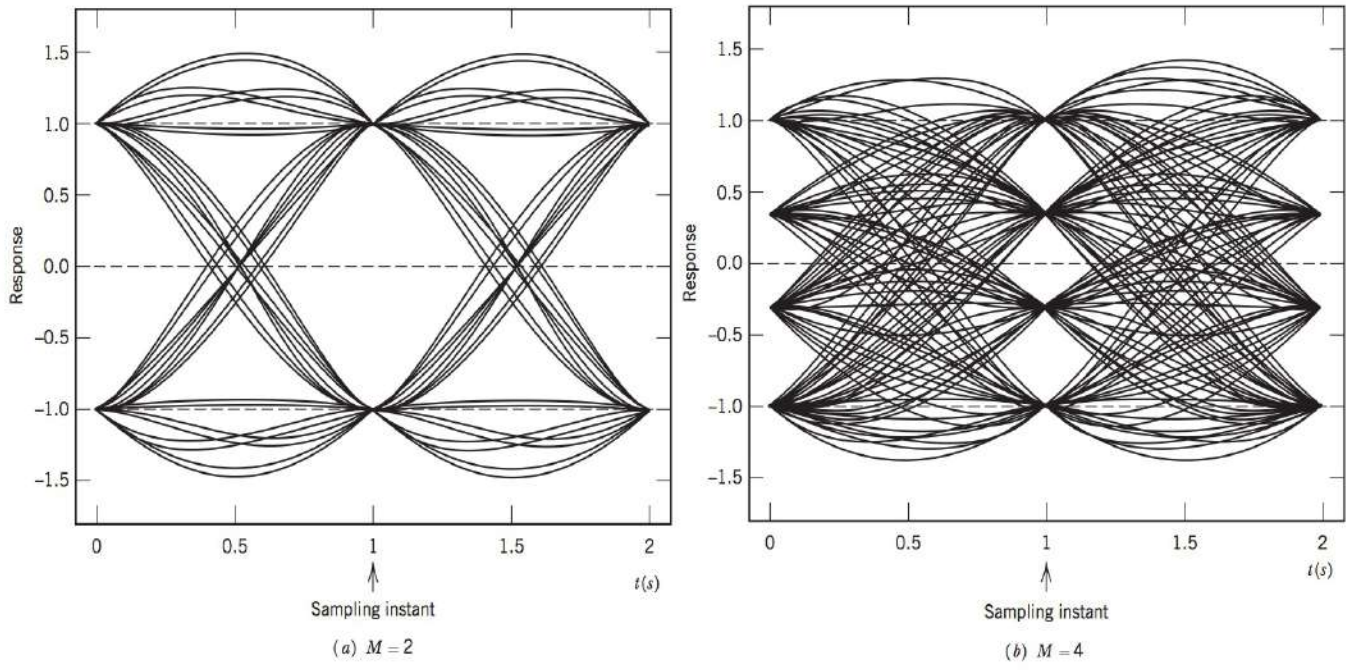


FIGURE 8.13 Eye diagram of received signal with no bandwidth limitation.

Figures 8.14*a* and 8.14*b* show the eye diagrams for these baseband-pulse transmission systems using the same system parameters as before, but this time under a bandwidth-limited condition. Specifically, the channel is now modeled by a low-pass *Butterworth filter*, whose frequency response is defined by

$$|H(f)| = \frac{1}{1 + (f/f_0)^{2N}}$$

where N is the order of the filter, and f_0 is its 3-dB cutoff frequency. For the simulation results shown in Figure 8.14, a filter of order 5 and f_0 equal to 55 percent of the symbol rate was used. Note that the 3-dB bandwidth of transmitted pulse is 50 percent of the symbol rate, so although the channel bandwidth cutoff frequency is greater than 3-dB bandwidth of the signal, its effect on the passband is observed in a decrease in the size of the eye opening. Instead of the distinct values at the ideal sampling time (as shown in Figure 8.13), now there is a blurred region. If the channel bandwidth were reduced further, the eye would close even more until finally no distinct eye opening would be recognizable.

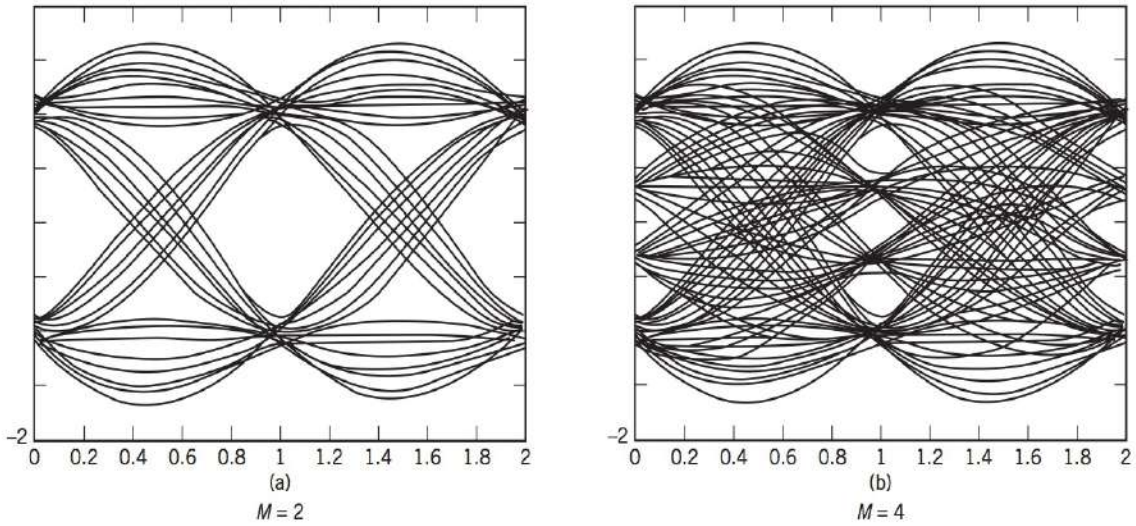


FIGURE 8.14 Eye diagram of received signal using a bandwidth-limited channel response.

8.6 NYQUIST'S CRITERION FOR DISTORTIONLESS TRANSMISSION

In practice, we typically find that the transfer function of a channel and the transmitted pulse shape are specified, and the problem is to determine the transfer functions of the transmit and receive filters so as to reconstruct the original binary data sequence $\{b_k\}$. The receiver does this by *extracting* and then *decoding* the corresponding sequence of coefficients. $\{a_k\}$, from the output $y(t)$. The *extraction* involves sampling the output $y(t)$ at time $t = iT_b$. The *decoding* requires that the weighted pulse contribution $a_k p(iT_b - kT_b)$ for $k = i$ be *free* from ISI due to the overlapping tails of all other weighted pulse contributions represented by $k \neq i$. This, in turn, requires that we *control* the overall pulse $p(t)$, as shown by

$$p(iT_b - kT_b) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \tag{8.49}$$

where $p(0) = 1$, by normalization. If $p(t)$ satisfies the condition of Eq (8.49), the receiver output $y(t_i)$ given in Eq. (8.48) simplifies to (ignoring the noise term)

$$y(t_i) = \mu a_i \qquad \text{for all } i$$

which implies zero intersymbol interference. Hence, the condition of Eq. (8.49) ensures *perfect reception in the absence of noise.*

From a design point of view, it is informative to transform the condition of Eq. (8.49) into the frequency domain. Consider then the sequence of samples $\{p(nT_b)\}$, where $n = 0, \pm 1, \pm 2, \ldots$. From the discussion presented in Chapter 7 on the sampling process, we recall that sampling in the time domain produces periodicity in the frequency domain. In particular, we may write

$$P_\delta(f) = R_b \sum_{n=-\infty}^{\infty} P(f - nR_b) \tag{8.50}$$

where $R_b = 1/T_b$ is the *bit rate* in bits per second (b/s); $P_\delta(f)$ is the Fourier transform of an infinite periodic sequence of delta functions of period T_b, whose areas are weighted by the respective sample values of $p(t)$. That is, $P_\delta(f)$ is given by

$$P_\delta(f) = \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} [p(mT_b)\,\delta(t - mT_b)]\exp(-j2\pi ft)\,dt \tag{8.51}$$

Let the integer $m = i - k$. Then, $i = k$ corresponds to $m = 0$, and likewise $i \neq k$ corresponds to $m \neq 0$. Accordingly, imposing the condition of Eq. (8.49) on the sample values of $p(t)$ in the integral of Eq. (8.51), we get

$$\begin{aligned} P_\delta(f) &= \int_{-\infty}^{\infty} p(0)\,\delta(t)\exp(-j2\pi ft)\,dt \\ &= p(0) \end{aligned} \tag{8.52}$$

where we have made use of the sifting property of the delta function. Since from Eq. (8.46) we have $p(0) = 1$, it follows from Eqs. (8.50) and (8.52) that the condition for zero intersymbol interference is satisfied if

$$\sum_{n=-\infty}^{\infty} P(f - nR_b) = T_b \tag{8.53}$$

We may now state the *Nyquist criterion for distortionless baseband transmission* in the absence of noise: *The frequency function $P(f)$ eliminates intersymbol interference for samples taken at intervals T_b provided that it satisfies Eq. (8.53).* Note that $P(f)$ refers to the overall system, incorporating the transmit filter, the channel, and the receive filter in accordance with Eq. (8.47).

IDEAL NYQUIST CHANNEL

The simplest way of satisfying Eq. (8.53) is to specify the frequency function $P(f)$ to be in the form of a *rectangular function*, as shown by

$$\begin{aligned} P(f) &= \begin{cases} \dfrac{1}{2W}, & -W < f < W \\ 0, & |f| > W \end{cases} \\ &= \frac{1}{2W}\,\mathrm{rect}\left(\frac{f}{2W}\right) \end{aligned} \tag{8.54}$$

where the overall system bandwidth W is defined by

$$W = \frac{R_b}{2} = \frac{1}{2T_b} \tag{8.55}$$

Harry Nyquist (1889–1976)

Nyquist was born in Sweden and immigrated to the United States in 1907. As an engineer at Bell Labs he made important contributions in many areas, including Johnson–Nyquist noise and the feedback stability of amplifiers (Nyquist stability theorem). The two areas where he made significant contributions to communications theory resulted from his investigations of transmission over a telegraph channel.

In 1927, Nyquist determined that the number of pulses that can be transmitted in a telegraph channel is limited to twice the bandwidth of the channel—now known as the Nyquist–Shannon sampling theorem. In the same investigation, Nyquist also determined the criteria that pulse shapes must have to achieve this limit—known as the Nyquist first, second, and third criterion. Nyquist also proposed raised cosine filters as examples of a pulse shape that satisfies his first criterion.

Nyquist was a prolific inventor and was credited with over 150 patents in his career. Some have credited Nyquist and Claude Shannon with the majority of the theoretical advances of modern communications.

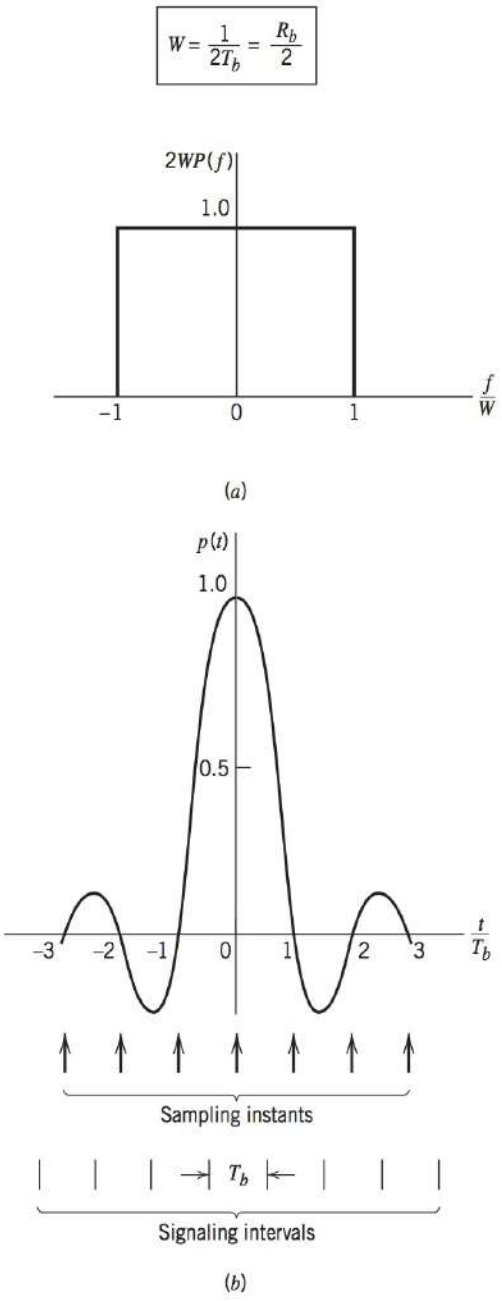


FIGURE 8.15 (*a*) Ideal amplitude response. (*b*) Ideal basic pulse shape.

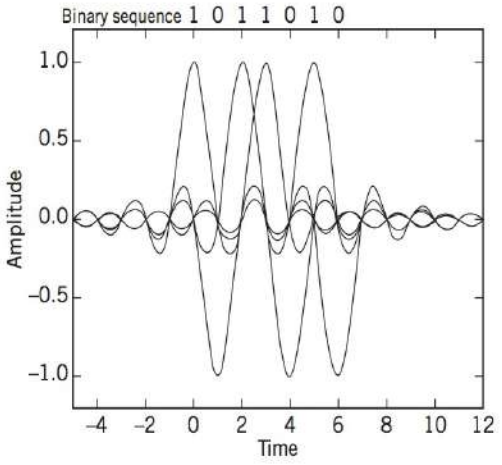


FIGURE 8.16 A series of sine pulses corresponding to the sequence 1011010.

According to the solution described by Eqs. (8.54) and (8.55), no frequencies of absolute value exceeding half the bit rate are needed. Hence, one signal waveform that produces zero intersymbol interference is defined by the *sinc function*:

$$p(t) = \frac{\sin(2\pi Wt)}{2\pi Wt} = \text{sinc}(2Wt) \tag{8.56}$$

The special value of the bit rate $R_b = 2W$ is called the *Nyquist rate*, and W is itself called the *Nyquist bandwidth*. Correspondingly, the ideal baseband pulse transmission system described by Eq. (8.54) in the frequency domain or, equivalently, Eq. (8.56) in the time domain is called the *ideal Nyquist channel*.

Figures 8.15*a* and 8.15*b* show plots of $P(f)$ and $p(t)$, respectively. In Figure 8.15*a*, the normalized form of the frequency function $P(f)$ is shown plotted for positive and negative frequencies. In Figure 8.15*b*, we have also included the signaling intervals and the corresponding centered sampling instants. The function $p(t)$ can be regarded as the impulse response of an ideal low-pass filter with passband amplitude response $1/2W$ and bandwidth W. The function $p(t)$ has its peak value at the origin and goes through zero at integer multiples of the bit duration T_b. It is apparent that if the received waveform $y(t)$ is sampled at the instants of time $t = 0, \pm T_b, \pm 2T_b, \ldots$, then the pulses defined by $\mu p(t - iT_b)$ with arbitrary amplitude μ and index $i = 0, \pm 1, \pm 2, \ldots$, will not interfere with each other. This condition is illustrated in Figure 8.16 for the binary sequence 1011010.

Although the use of the ideal Nyquist channel does indeed achieve economy in bandwidth in that it solves the problem of zero intersymbol interference with the minimum bandwidth possible, there are two practical difficulties that make it an undesirable objective for system design:

1. It requires that the amplitude characteristic of $P(f)$ be flat from $-W$ to W, and zero elsewhere. This is physically unrealizable because of the abrupt transitions at the band edges $\pm W$.
2. The function $p(t)$ decreases as $1/|t|$ for large $|t|$, resulting in a slow rate of decay. This is also caused by the discontinuity of $P(f)$ at $\pm W$. Accordingly, there is practically no margin of error in sampling times in the receiver.

To evaluate the effect of this *timing error*, consider the sample of $y(t)$ at $t = \Delta t$, where Δt is the timing error. To simplify the exposition, we may put the correct sampling time t_i equal to zero. In the absence of noise, we thus have

$$\begin{aligned} y(\Delta t) &= \mu \sum_k a_k p(\Delta t - kT_b) \\ &= \mu \sum_k a_k \frac{\sin[2\pi W(\Delta t - kT_b)]}{2\pi W(\Delta t - kT_b)} \end{aligned} \tag{8.57}$$

Since $2WT_b = 1$, by definition, we may rewrite Eq. (8.57) as

$$y(\Delta t) = \mu a_0 \,\text{sinc}(2\pi W \Delta t) + \frac{\mu \sin(2\pi W \Delta t)}{\pi} \sum_{\substack{k \\ k \neq 0}} \frac{(-1)^k a_k}{(2W\Delta t - k)} \tag{8.58}$$

The first term on the right-hand side of Eq. (8.58) defines the desired symbol, whereas the remaining series represents the intersymbol interference caused

by the timing error Δt in sampling the output $y(t)$. Unfortunately, it is possible for this series to diverge, thereby causing erroneous decisions in the receiver.

RAISED COSINE SPECTRUM

We may overcome the practical difficulties encountered with the ideal Nyquist channel by extending the bandwidth from the minimum value $W = R_b/2$ to an adjustable value between W and $2W$. We now specify the frequency function $P(f)$ to satisfy a condition more elaborate than that for the ideal Nyquist channel. Specifically, we retain three terms of Eq. (8.53) and restrict the frequency band of interest to $[-W,W]$, as shown by

$$P(f) + P(f - 2W) + P(f + 2W) = \frac{1}{2W}, \quad -W \le f \le W \tag{8.59}$$

We may devise several band-limited functions that satisfy Eq. (8.59). A particular form of $P(f)$ that embodies many desirable features is provided by a *raised cosine spectrum*. This frequency characteristic consists of a flat portion and a *rolloff* portion that has a sinusoidal form, as follows:

$$P(f) = \begin{cases} \dfrac{1}{2W}, & 0 \le |f| < f_1 \\ \dfrac{1}{4W}\left\{1 - \sin\left[\dfrac{\pi(|f| - W)}{2W - 2f_1}\right]\right\}, & f_1 \le |f| < 2W - f_1 \\ 0, & |f| \ge 2W - f_1 \end{cases} \tag{8.60}$$

The frequency parameter f_1 and bandwidth W are related by

$$\alpha = 1 - \frac{f_1}{W} \tag{8.61}$$

The parameter α is called the *rolloff factor*; it indicates the *excess bandwidth* over the ideal solution, W. Specifically, the transmission bandwidth B_T is defined by $2W - f_1 = W(1 + \alpha)$.

The frequency response $P(f)$, normalized by multiplying it by $2W$, is shown plotted in Figure 8.17*a* for three values of α, namely, 0, 0.5, and 1. We see that for $\alpha = 0.5$ or 1, the function $P(f)$ cuts off gradually as compared with the ideal Nyquist channel (i.e., $\alpha = 0$) and is therefore easier to implement in practice. Also the function $P(f)$ exhibits odd symmetry with respect to the Nyquist bandwidth W, making it possible to satisfy the condition of Eq. (8.59).

The time response $p(t)$ is the inverse Fourier transform of the function $P(f)$. Hence, using the $P(f)$ defined in Eq. (8.60), we obtain the result (see Problem 8.10)

$$p(t) = \left[\text{sinc}(2Wt)\right]\left(\frac{\cos(2\pi\alpha Wt)}{1 - 16\alpha^2 W^2 t^2}\right) \tag{8.62}$$

which is shown plotted in Figure 8.17*b* for $\alpha = 0$, 0.5, and $|t| < 3T_b$.

The function $p(t)$ consists of the product of two factors: the factor $\text{sinc}(2Wt)$ characterizing the ideal Nyquist channel and a second factor that decreases as $1/|t|^2$ for large $|t|$. The first factor ensures zero crossings of $p(t)$ at the desired sampling instants of time $t = iT$ with i an integer (positive and negative). The second factor reduces the tails of the pulse considerably below that obtained from the ideal Nyquist channel, so that the transmission of binary waves using such pulses is relatively insensitive to sampling time errors. In fact, for $\alpha = 1$ we have the most gradual rolloff in that the amplitudes of the oscillatory tails of $p(t)$ are smallest. Thus, the amount of intersymbol interference resulting from timing error decreases as the rolloff factor α is increased from zero to unity.

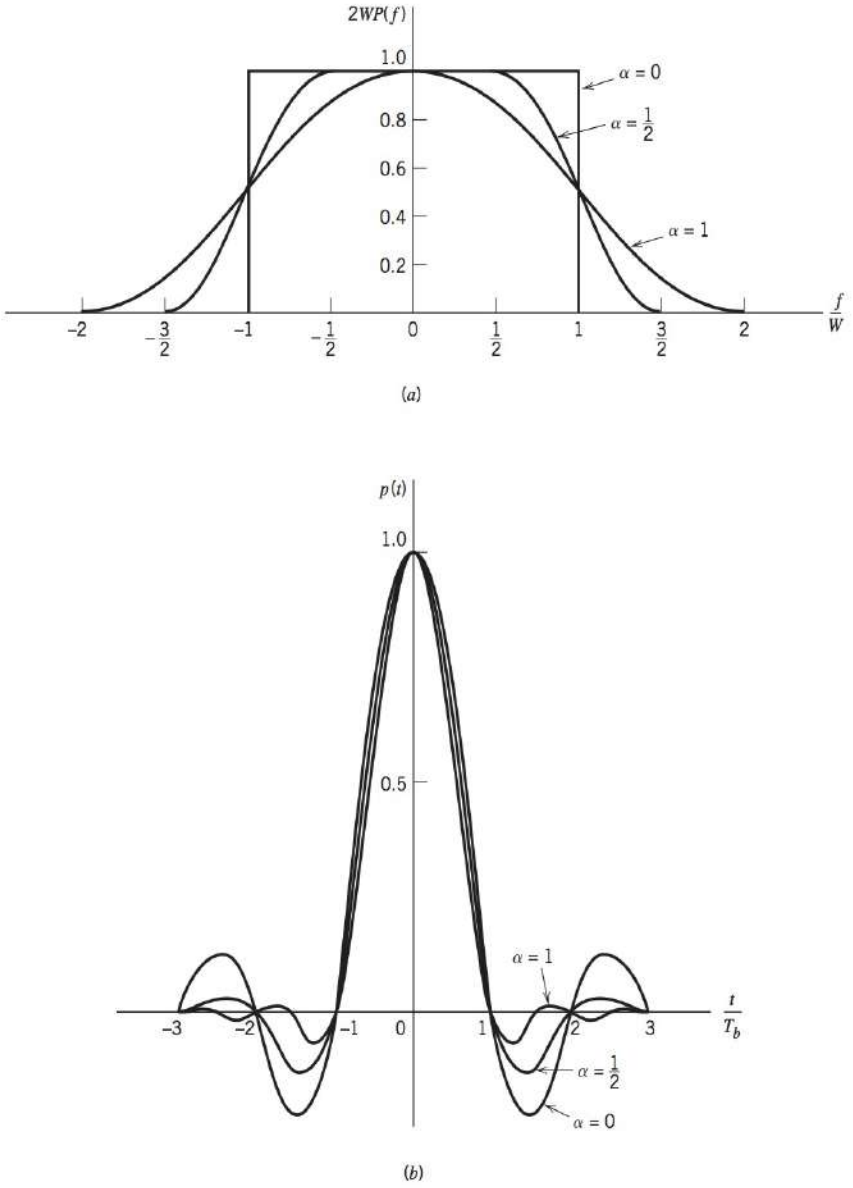


FIGURE 8.17 Responses for different rolloff factors. (*a*) Frequency response. (*b*) Time response.

The special case with $\alpha = 1$ (i.e., $f_1 = 0$) is known as the *full-cosine rolloff* characteristic, for which the frequency response of Eq. (8.60) simplifies to

$$P(f) = \begin{cases} \dfrac{1}{4W}\left[1 + \cos\left(\dfrac{\pi f}{2W}\right)\right], & 0 < |f| < 2W \\ 0, & |f| \geq 2W \end{cases} \tag{8.63}$$

Correspondingly, the time response $p(t)$ simplifies to

$$p(t) = \frac{\text{sinc}(4Wt)}{1 - 16W^2t^2} \tag{8.64}$$

This time response exhibits two interesting properties:

1. At $t = \pm T_b/2 = \pm 1/4W$, we have $p(t) = 0.5$; that is, the pulse width measured at half amplitude is exactly equal to the bit duration T_b.
2. There are zero crossings at $t = \pm 3T_b/2, \pm 5T_b/2, \ldots$ in addition to the usual zero crossings at the sampling times $t = \pm T_b, \pm 2T_b, \ldots$.

These two properties are extremely useful in extracting a timing signal from the received signal for the purpose of synchronization. However, the price paid for this desirable

property is the use of a channel bandwidth double that required for the ideal Nyquist channel corresponding to $\alpha = 0$.

EXAMPLE 8.4 Bandwidth Requirement of the T1 System

In Example 7.3 of Chapter 7, we described the signal format for the T1 carrier system that is used to multiplex 24 independent voice inputs, based on an 8-bit PCM word. It was shown that the bit duration of the resulting time-division multiplexed signal (including a framing bit) is

$$T_b = 0.647\ \mu s$$

Assuming the use of an ideal Nyquist channel, it follows that the minimum transmission bandwidth B_T of the T1 system is (for $\alpha = 0$)

$$B_T = W = \frac{1}{2T_b} = 772\ \text{kHz}$$

However, a more realistic value for the necessary transmission bandwidth is obtained by using a full-cosine rolloff characteristic with $\alpha = 1$. In this case, we find that

$$B_T = W(1 + \alpha) = 2W = \frac{1}{T_b} = 1.544\ \text{MHz}$$

It is interesting to compare the transmission bandwidth requirement of the T1 system with the minimum bandwidth requirement of a corresponding frequency-division multiplexing (FDM) system. We recall from Chapter 3 that of all the CW modulation techniques, the use of single sideband (SSB) modulation requires the minimum bandwidth possible. Thus, to accommodate an FDM system using SSB modulation to transmit 24 independent voice inputs, and assuming a bandwidth of 4 kHz for each voice input, the channel must provide the transmission bandwidth

$$B_T = 24 \times 4 = 96\ \text{kHz}$$

This is more than an order of magnitude smaller than the bandwidth requirement of the T1 system.

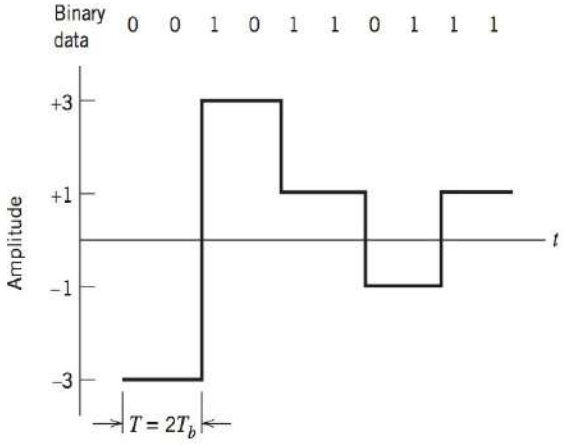
Controlled ISI

We have treated intersymbol interference as an undesirable phenomenon that produces degraded system performance. Nevertheless, we may design systems with controlled intersymbol interference such that it is possible to achieve a signaling rate equal to the Nyquist rate of $2W$ symbols per second in a channel of bandwidth W Hertz. Such schemes are called *correlative-level coding* or *partial response signaling*. The design of these schemes is based on the following premise: Since intersymbol interference introduced into the transmitted signal is known, its effect can be interpreted at the receiver in a deterministic way.

Correlative-level coding may be regarded as a practical method of achieving the theoretical maximum signaling rate using realizable filters.

8.7 BASEBAND *M*-ARY PAM TRANSMISSION

In the baseband binary PAM system of Figure 8.8, the pulse-amplitude modulator produces binary pulses, that is, pulses with one of two possible amplitude levels. On the other hand, in a *baseband M-ary PAM system*, the pulse-amplitude modulator produces one of M possible amplitude levels, with $M > 2$, as illustrated in Figure 8.18*a* for the case of a *quaternary* ($M = 4$) system and the binary data sequence 0010110111. The corresponding electrical representation for each of the four possible pairs of bits is shown in Figure 8.18*b*. In an M-ary system, the information source emits a sequence of symbols from an alphabet that consists of M symbols. Each amplitude level at the pulse-amplitude modulator output corresponds to a distinct symbol, so that there are M distinct amplitude levels to be transmitted. Consider then an M-ary PAM system with a signal alphabet that contains M equally likely and statistically independent symbols, with the symbol duration denoted by T seconds. We refer to $1/T$ as the *signaling rate* of the system, which is expressed in *symbols per second* or *bauds*. It is informative to relate the signaling rate of this system to that of an equivalent binary PAM system for which the value of M is 2 and the successive binary symbols 1 and 0 are equally likely and statistically independent,



(a)

Dibit	Amplitude
00	−3
01	−1
11	+1
10	+3

(b)

FIGURE 8.18 Output of a quaternary system. (*a*) Waveform. (*b*) Representation of the four possible pairs of bits.

with the duration of either symbol denoted by T_b seconds. Under the conditions described here, the binary PAM system produces information at the rate of $1/T_b$ bits per second. We also observe that in the case of a quaternary PAM system, for example, the four possible symbols may be identified with the bit pairs 00, 01, 10, and 11. We thus see that each symbol represents 2 bits of information, and 1 baud is equal to 2 bits per second. We may generalize this result by stating that in an M-ary PAM system, one baud is equal to $\log_2 M$ bits per second, and the symbol duration T of the M-ary PAM system is related to the bit duration T_b of the equivalent binary PAM system as

$$T = T_b \log_2 M \tag{8.65}$$

Therefore, in a given channel bandwidth, we find that by using an M-ary PAM system, we are able to transmit information at a rate that is $\log_2 M$ faster than the corresponding binary PAM system. However, to realize the same average probability of symbol error, an M-ary PAM system requires more transmitted power. Specifically, we find that for M much larger than 2 and an average probability of symbol error small compared to 1, the transmitted power must be increased by a factor of $M^2/\log_2 M$, compared to a binary PAM system.

In a baseband M-ary system, first of all, the sequence of symbols emitted by the information source is converted into an M-level PAM pulse train by a pulse-amplitude modulator at the transmitter input. Next, as with the binary PAM system, this pulse train is shaped by a transmit filter and then transmitted over the communication channel, which corrupts the signal waveform with both noise and distortion. The received signal is passed through a receive filter and then sampled at an appropriate rate in synchronism with the transmitter. Each sample is compared with preset *threshold* values (also called *slicing* levels), and a decision is made as to which symbol was transmitted. We therefore find that the designs of the pulse-amplitude modulator and the decision-making device in an M-ary PAM are more complex than those in a binary PAM system. Intersymbol interference, noise, and imperfect synchronization cause errors to appear at the receive output. The transmit and receive filters are designed to minimize these errors. Procedures used for the design of these filters are similar to those discussed in Sections 8.4 and 8.6 for baseband binary PAM systems. In particular, the raised cosine pulse shape, which is ISI-free for binary signaling, is also ISI-free for M-ary signaling.

8.8 TAPPED-DELAY-LINE EQUALIZATION

In Example 8.2, we saw how a band-limited channel, such as the telephone channel, can affect the high-speed transmission of digital data. An efficient approach to high-speed transmission of digital data over such a channel uses a combination of two basic signal-processing operations:

- Discrete PAM, which involves encoding the amplitudes of successive pulses in a periodic pulse train with a discrete set of possible amplitude levels.
- A linear modulation scheme, which offers bandwidth conservation to transmit the encoded pulse train over the telephone channel.

At the receiving end of the system, the received signal is demodulated and synchronously sampled, and then decisions are made as to which particular symbols were transmitted.

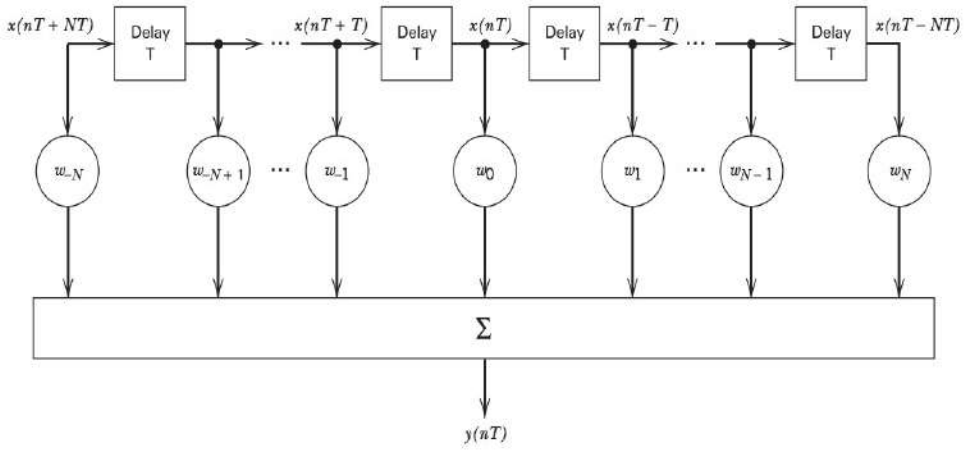


FIGURE 8.19 Tapped-delay-line filter.

As a result of dispersion of the pulse shape by the band-limited channel, we find that the number of detectable amplitude levels is often limited by intersymbol interference rather than by additive noise. In principle, if the channel is known precisely, it is virtually always possible to make the intersymbol interference at the sampling instants arbitrarily small by using a suitable pair of transmit and receive filters, so as to control the overall pulse shape in the manner described previously. The transmit filter is placed directly before the modulator, whereas the receive filter is placed directly after the demodulator.

In practice, however, we seldom have prior knowledge of the exact channel characteristics. Also, there is the unavoidable problem of imprecision that arises in the physical implementation of the transmit and receive filters. The net result of all these effects is that there will be some residual distortion for ISI to be a limiting factor on the data rate of the system. To compensate for the intrinsic residual distortion, we may use a process known as equalization. The filter used to perform such a process is called an equalizer.

A device well-suited for the design of a linear equalizer is the tapped-delay-line filter, as depicted in Figure 8.19. For symmetry, the total number of taps is chosen to be $(2N + 1)$, with the weights denoted by $w_{-N}, \ldots, w_{-1}, w_0, w_1, \ldots, w_N$. The impulse response of the tapped-delay-line equalizer is therefore

$$h(t) = \sum_{k=-N}^{N} w_k \delta(t - kT) \tag{8.66}$$

where $\delta(t)$ is the Dirac delta function, and the delay T is chosen equal to the symbol duration.

Suppose that the tapped-delay-line equalizer is connected in cascade with a linear system whose impulse response is $c(t)$, as depicted in Figure 8.20. Let $p(t)$ denote the impulse response of the equalized system. Then $p(t)$ is equal to the convolution of $c(t)$ and $h(t)$, as shown by

$$\begin{aligned} p(t) &= c(t) \star h(t) \\ &= c(t) \star \sum_{k=-N}^{N} w_k \delta(t - kT) \end{aligned}$$

Interchanging the order of summation and convolution:

$$\begin{aligned} p(t) &= \sum_{k=-N}^{N} w_k c(t) \star \delta(t - kT) \\ &= \sum_{k=-N}^{N} w_k c(t - kT) \end{aligned} \tag{8.67}$$

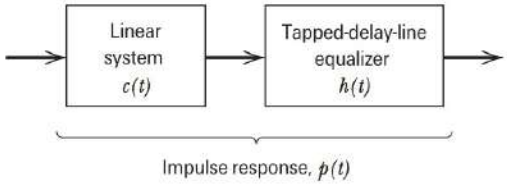


FIGURE 8.20 Cascade connection of linear system and tapped-delay-line equalizer.

where we have made use of the sifting property of the delta function. Evaluating Eq. (8.67) at the sampling times $t = nT$, we get the *discrete convolution sum*

$$p(nT) = \sum_{k=-N}^{N} w_k\, c((n-k)T) \tag{8.68}$$

Note that the sequence $\{p(nT)\}$ is longer than $\{c(nT)\}$.

To eliminate intersymbol interference completely, we must satisfy the Nyquist criterion for distortionless transmission described in Eq. (8.49), with T used in place of T_b. It is assumed that the $p(t)$ is defined in such a way that the normalized condition $p(0) = 1$ is satisfied in accordance with Eq. (8.46). Thus, for no intersymbol interference we require that

$$p(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

But from Eq. (8.68) we note that there are only $(2N + 1)$ adjustable coefficients at our disposal. Hence, this ideal condition can only be satisfied approximately as follows:

$$p(nT) = \begin{cases} 1, & n = 0 \\ 0, & n = \pm 1, \pm 2, \ldots, \pm N \end{cases} \tag{8.69}$$

To simplify the notation, we let the nth sample of the impulse response $c(t)$ be written as

$$c_n = c(nT) \tag{8.70}$$

Then, imposing the condition of Eq. (8.69) on the discrete convolution sum of Eq. (8.68), we obtain a set of $(2N + 1)$ simultaneous equations:

$$\sum_{k=-N}^{N} w_k\, c_{n-k} = \begin{cases} 1, & n = 0 \\ 0, & n = \pm 1, \pm 2, \ldots, \pm N \end{cases} \tag{8.71}$$

Equivalently, in matrix form we may write

$$\begin{bmatrix} c_0 & \cdots & c_{-N+1} & c_{-N} & c_{-N-1} & \cdots & c_{-2N} \\ \\ c_{N-1} & \cdots & c_0 & c_{-1} & c_{-2} & \cdots & c_{-N-1} \\ c_N & \cdots & c_1 & c_0 & c_{-1} & \cdots & c_{-N} \\ c_{N+1} & \cdots & c_2 & c_1 & c_0 & \cdots & c_{-N+1} \\ \\ c_{2N} & \cdots & c_{N+1} & c_N & c_{N-1} & \cdots & c_0 \end{bmatrix} \begin{bmatrix} w_{-N} \\ \vdots \\ w_{-1} \\ w_0 \\ w_1 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{8.72}$$

A tapped-delay-line equalizer described by Eq. (8.71) or, equivalently Eq. (8.72), is referred to as a *zero-forcing equalizer*. Such an equalizer is optimum in the sense that it minimizes the peak distortion (intersymbol interference). It also has the nice feature of being relatively simple to implement. In theory, the longer we make the equalizer (i.e., permit N to approach infinity), the more closely will the equalized system approach the ideal condition specified by the Nyquist criterion for distortionless transmission.

The zero-forcing strategy described above works well in the laboratory, where we have access to the system to be equalized and we know the system coefficients $c_{-N}, \ldots, c_{-1}, c_0\, c_1, \ldots, c_N$ that are needed for the solution of Eq. (8.72). In a telecommunications network, the channel is often time varying. To realize the full transmission capability of a time varying channel there is a need for *adaptive equalization*. The process of equalization is said to be adaptive when the equalizer adjusts itself continuously and automatically, using information extracted from the input signal. Adaptive equalization is beyond the scope of this introduction to equalization, but suffice it to say that the majority of equalizers used in practice are adaptive.

8.9 THEME EXAMPLE—100BASE-TX—TRANSMISSION OF 100 Mbps OVER TWISTED PAIR[3]

One of the dominant forms of fast Ethernet, referred to as 100BASE-TX, transmits up to 100 Mbps over two *pairs of twisted copper wires*, commonly referred to as a Category 5 cable. The communications link is a maximum of 100 meters in length and typically one pair of twisted wires is used in each direction.

With 100BASE-TX, the data stream is processed through several stages of encoding before it is transmitted. In the first stage, four bits are binary encoded to five bits to produce a series of zeros and ones in NRZ format. This encoding is referred to as 4B5B encoding and the output bits are clocked at a 125 MHz symbol rate. This mapping from four bits to five bits creates extra signal transitions that provide clocking information for the signal. For example, a run of four bits such as 0000 contains no transitions and that causes clocking problems for the receiver. The 4B5B encoding solves this problem by assigning each block of four consecutive bits an equivalent word of five bits. These five-bit words are predetermined in a dictionary and they are chosen to ensure that there will be at least one transition per block of bits. A drawback of the 4B5B code is that more bits are needed to send the same information. The NRZ output of a 4B5B encoding is shown in Figure 8.21*a*.

The second encoding stage converts the NRZ to a format known as NRZ invert on one or NRZI. With the NRZI format, the information is contained in the signal transitions and not the voltage levels, hence it is a form of differential encoding. With NRZI format, a 1 produces a half-width rectangular pulse at the current voltage level followed by a half-width pulse at the other voltage level; and a 0 is represented by no change in the current voltage level. An example of an NRZI encoding of a NRZ signal is shown in Figure 8.21*b*. The differential encoding means that absolute level is not important in the detection process, but rather changes in level determine the transmitted bits.

In the third encoding stage, the NRZI bits are converted to a three-level format known as MLT-3. With this three-level encoding, the zero portion of the signal is unchanged but the positive portion of the NRZI signal is alternated between positive and negative voltage levels after every zero. An example of the conversion of the NRZI bits to a MLT-3 representation is shown in Figure 8.21*c*. The multilevel format reduces the fundamental frequency of the data to 31.25 MHz from 62.5 MHz. This reduction in the signal spectrum makes the transmitted signal less sensitive to the bandwidth limitations of the channel.

Transmitting over *twisted pairs* of copper wires at rates of 100 Mbps and higher has many challenges. Among these challenges are:

- The signal attenuation over 100 meters of twisted pairs is considerable and increases significantly with frequency and temperature and is also affected by magnetics.
- There is a potential *echo* from the opposite cable end that can degrade performance.

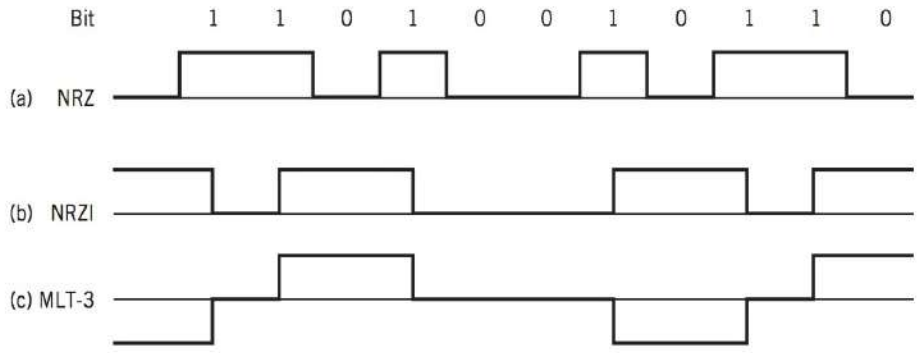


FIGURE 8.21 The translation between different line codes used with 100BASE-TX.

- Since the system transmits and receives simultaneously (full duplex), through adjacent twisted pairs, there is the possibility of crosstalk between the transmitting pair and the receiving twisted pair. If the crosstalk coupling occurs near the receiving end, it is referred to as the *near-end crosstalk*; otherwise it is called *far-end crosstalk*.

By performing appropriate signal processing such as equalization, echo and crosstalk cancellation, many of these problems can be mitigated. To illustrate, we address the dispersive nature of the channel through equalization in the following.

The pulse shape used for 100BASE-TX is not specified analytically but rather in terms of its duration, rise and fall times. In particular, the pulse is 16 nanoseconds wide and must have rise and fall times between 3 and 5 nanoseconds. Rise (and fall) time is defined as the time between 10 percent and 90 percent of full amplitude. An example of such a pulse is shown in Figure 8.22.

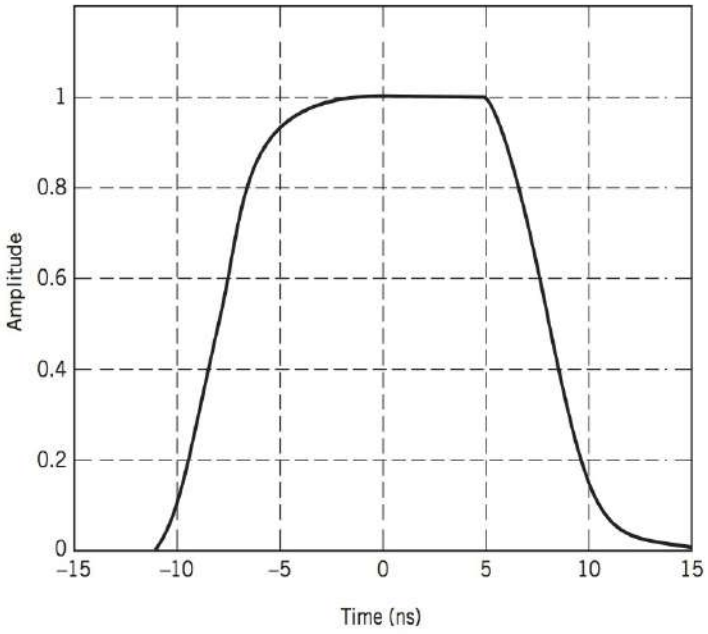


FIGURE 8.22 Example pulse shape for 100BASE-TX.

With 100BASE-TX, the receiver must tolerate a channel (that is, a twisted-pair cable) which has the attenuation characteristic shown in Figure 8.23 for a cable of 100 meters. This characteristic indicates a small amount of attenuation at low frequencies (where twisted pairs are used for analog telephone communications over much longer distances) but the attenuation increases rapidly when the frequency exceeds 1 MHz.

To illustrate the effect that such a channel would have on performance, we simulated the effect of transmitting data having the pulse shape of Figure 8.22 through a channel having an impulse response approximately corresponding to the amplitude spectrum of Figure 8.23. The eye diagram of the transmitted signal (prior to the channel) is shown in Figure 8.24*a*, while the eye diagram after the channel is shown in Figure 8.24*b*. The channel has caused significant closure of the eye, reducing its tolerance to noise and other forms of distortion.

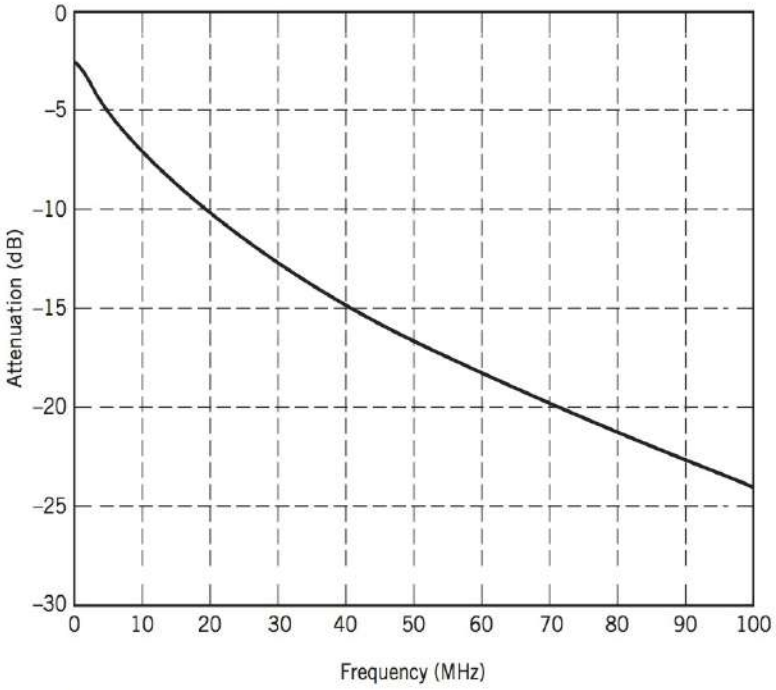


FIGURE 8.23 Worst case cable-attenuation characteristic for 100BASE-TX.

The 100BASE-TX standard assumes that the receiver will use equalization to compensate for the effects of the channel. For example, we can apply the techniques of the zero-forcing equalizer described in Section 8.8 to this case. In particular, the system response is simply the convolution of the transmitted pulse shape and the impulse response of the channel. The result of the convolution is then sampled at T-spaced intervals about the peak where T is the symbol period of 16 nanoseconds. This sampling provides the samples $\{c_{-2N}, \ldots, c_0, \ldots, c_{2N}\}$ to be used in the computation of the equalizer in Eq. (8.72). For this example, we selected $N = 1$, corresponding to a 3-tap equalizer. The computed equalizer taps are $\{w_{-1}, w_0, w_1\} = \{-0.41, 1.87, -0.39\}$, subject to scaling. If this 3-tap T-spaced equalizer is applied to the received signal, then the eye-diagram of Figure 8.24*c* is obtained. The

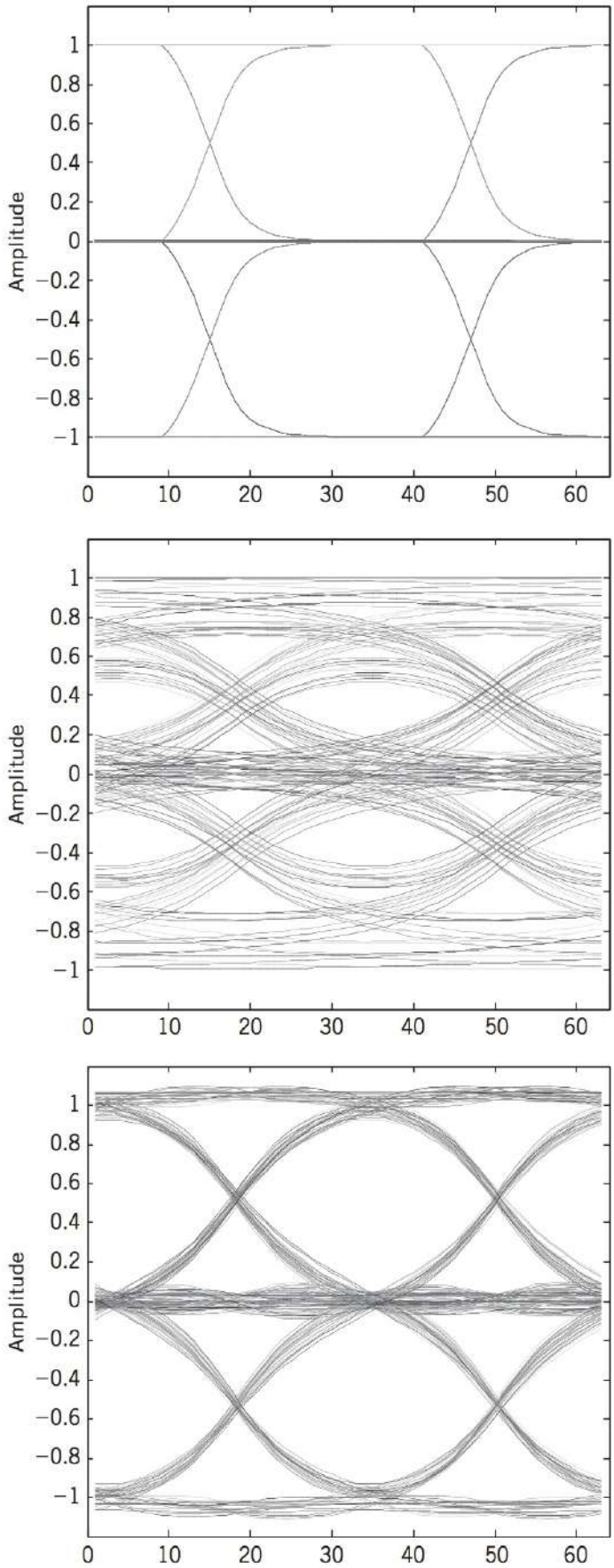


FIGURE 8.24 Eye diagrams for (*a*) transmitted signal; (*b*) received signal; and (*c*) equalized received signal.

equalization clearly increases the eye opening, providing more robust data detection by increasing the tolerance of the signal to noise and other distortions.

This theme example has demonstrated that with proper design and appropriate digital signal processing, high data rates can be transmitted over very nonideal channels.

8.10 SUMMARY AND DISCUSSION

In this chapter, we studied the effects of noise and intersymbol interference on the performance of baseband-pulse transmission systems. Intersymbol interference (ISI) is different from noise in that it is a *signal-dependent* form of interference that arises because of deviations in the frequency response of a channel from the ideal low-pass filter (Nyquist channel). It disappears when the transmitted signal is switched off. The result of these deviations is that the received pulse corresponding to a particular data symbol is affected by (1) the tail ends of the pulses representing the previous symbols and (2) the front ends of the pulses representing the subsequent symbols.

Depending on the received signal-to-noise ratio, we may distinguish three different situations that can arise in baseband-pulse transmission systems for channels with fixed characteristics:

1. *The effect of ISI is negligible in comparison to that of the noise.* The proper procedure in this case is to use a matched filter, which is the optimum linear time-invariant filter for maximizing the peak pulse signal-to-noise ratio.
2. *The received signal-to-noise ratio is high enough to ignore the effect of noise.* In this case, we need to guard against the effects of ISI on the reconstruction of the transmitted data at the receiver. In particular, control must be exercised over the shape of the received pulse. This design objective can be achieved in one of two different ways:
   - Using a raised cosine spectrum for the overall frequency response of the baseband-pulse transmission system.
   - Using correlative-level coding or partial-response signaling that adds ISI to the transmitted signal in a controlled manner.
3. *The ISI and noise are both significant.* The solution for this third situation requires the joint optimization of the transmit and receive filters. In brief, a suitable pulse shape is first used to reduce the ISI to zero at the sampling instants, and then Schwarz's inequality is invoked to maximize the output signal-to-noise ratio at the sampling instants.

When, however, the channel is random in the sense of being one of an ensemble of possible physical realizations, which is frequently the case in a telecommunications environment, the use of fixed filter designs based on average channel characteristics may not be adequate. In situations of this kind, the preferred approach is to use an equalizer, positioned after the receive filter in the receiver. The purpose of the equalizer is to compensate for variations in the frequency response of the channel. We have introduced the tapped-delay-line filter as a highly effective method of implementing equalization. The design of such an equalizer requires knowledge of the channel and the transmit and receive filter characteristics. In practice, one method of estimating the combined impulse response of the system is to transmit a known sequence at the start of transmission. For channels that may be time-varying, algorithms have been developed to automatically adjust the taps during the course of transmission; such a process is known as *adaptive equalization*. An adaptive equalizer is capable of dealing with the combined effects of ISI and receiver noise in a nonstationary environment. Its practical value lies in the fact that almost every modem (modulator-demodulator) in commercial use today for the transmission of data over voice-grade telephone channel uses an adaptive equalizer as an integral part.

NOTES AND REFERENCES

1. The classic books on baseband-pulse transmission are Lucky, Salz, and Weldon (1968) and Sunde (1969). For detailed treatment of different aspects of the subject, see Gitlin, Hayes, and Weinstein (1992), Proakis (2001), and Benedetto, Biglieri, and Castellani (1987).
2. For review material on the matched filter and its properties, see the papers by Turin (1960, 1976).
3. The requirements for 100 Mbps transmission for internet applications are described in the standards: IEEE Std. 802.3 – 2005, Part 3; and ANSI X3.263 (1995).

PROBLEMS

8.1 Consider the signal $s(t)$ shown in Figure P8.1.

(a) Determine the impulse response of a filter matched to this signal and sketch it as a function of time.

(b) Plot the matched filter output as a function of time.

(c) What is the peak value of the output?

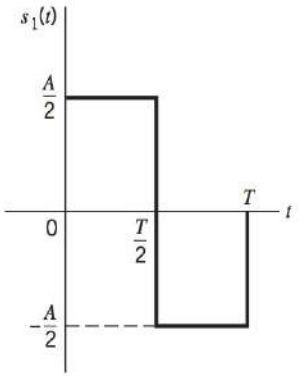


Figure P8.1

8.2 It is proposed to implement a matched filter in the form of a tapped-delay-line filter with a set of tap-weights $\{w_k, k = 0, 1, \ldots, K\}$. Given a signal $s(t)$ of duration T seconds to which the filter is matched, find the value of w_k. Assume that the signal is uniformly sampled.

8.3 In this problem, we derive the formulas used to compute the power spectra of Figure 8.1 for the five line codes described in Section 8.2, In the case of each line code, the bit duration is T_b and the pulse amplitude A is conditioned to normalize the average power of the line code to unity as indicated in Figure 8.1. Assume that the data stream is randomly generated, and symbols 0 and 1 are equally likely.

Derive the power spectral densities of these line codes as summarized here:

(a) Unipolar nonreturn-to-zero signals:

$$S(f) = \frac{A^2T_b}{4}\operatorname{sinc}^2(fT_b)\left(1 + \frac{1}{T_b}\delta(f)\right)$$

(b) Polar nonreturn-to-zero signals:

$$S(f) = A^2T_b \operatorname{sinc}^2(fT_b)$$

(c) Unipolar return-to-zero signals:

$$S(f) = \frac{A^2T_b}{16}\operatorname{sinc}^2\left(\frac{fT_b}{2}\right)\left[1 + \frac{1}{T_b}\sum_{n=-\infty}^{\infty}\delta\left(f - \frac{n}{T_b}\right)\right]$$

(d) Bipolar return-to-zero signals:

$$S(f) = \frac{A^2T_b}{4}\operatorname{sinc}^2\left(\frac{fT_b}{2}\right)\sin^2(\pi fT_b)$$

(e) Manchester-encoded signals:

$$S(f) = A^2T_b \operatorname{sinc}^2\left(\frac{fT_b}{2}\right)\sin^2\left(\frac{\pi fT_b}{2}\right)$$

Hence, confirm the spectral plots displayed in Figure 8.1.

8.4 Consider a rectangular pulse defined by

$$g(t) = \begin{cases} A, & 0 \le t \le T \\ 0, & \text{otherwise} \end{cases}$$

It is proposed to approximate the matched filter for $g(t)$ by an ideal low-pass filter of bandwidth B; maximization of the peak pulse signal-to-noise ratio is the primary objective.

(a) Determine the optimum value of B for which the ideal low-pass filter provides the best approximation to the matched filter.

(b) By how many decibels is the ideal low-pass filter worse off than the matched filter?

8.5 A binary PCM wave uses on–off signaling to transmit symbols 1 and 0; symbol 1 is represented by a rectangular pulse of amplitude A and duration T_b. The additive noise at the receiver input is white and Gaussian, with zero mean and power spectral density $N_0/2$. Assuming that symbols 1 and 0 occur with equal probability, find an expression for the average probability of error at the receiver output, using a matched filter as described in Section 8.3.

8.6 A binary PCM system, using NRZ signaling, operates with an average probability of error equal to 10^{-6}. Suppose that the signaling rate is doubled. Find the new value of the average probability of error. See the Appendix for methods of evaluating the Q-function or Figure 8.7.

8.7 A continuous-time signal is noise sampled and then transmitted as a PCM signal. The random noise at the input of the decision device in the receiver has a variance of 0.01 volts^2.

(a) Assuming the use of NRZ signaling, determine the pulse amplitude that must be received for the average error rate not to exceed 1 bit in 10^8 bits.

(b) If the added presence of interference causes the error rate to increase to 1 bit in 10^6 bits, what is the variance of the interference?

8.8 In a binary system, symbols 0 and 1 have *a priori probabilities* p_0 and p_1, respectively. The conditional probability density function of the random variable Y (with sample value y) obtained by sampling the matched filter output in the receiver of Figure 8.5 at the end of a signaling interval, given that symbol 0 was transmitted, is denoted by $f_Y(y|0)$. Similarly, $f_Y(y|1)$ denotes the conditional probability density function of Y, given that symbol 1 was transmitted. Let λ denote the threshold used in the receiver, so that if the sample value y exceeds λ, the receiver decides in favor of symbol 1; otherwise, it decides in favor of symbol 0. Show that the optimum threshold λ_{opt}, for which the average probability of error is a minimum, is given by the solution of

$$\frac{f_Y(\lambda_{opt}|1)}{f_Y(\lambda_{opt}|0)} = \frac{p_0}{p_1}$$

8.9 The overall pulse shape $p(t)$ of a baseband binary PAM system is defined by

$$p(t) = \text{sinc}\left(\frac{t}{T_b}\right)$$

where T_b is the bit duration of the input binary data. The amplitude levels at the pulse modulator output are $+1$ or -1, depending on whether the binary symbol at the input is 1 or 0, respectively. Sketch the waveform at the output of the receive filter in response to the input data 001101001.

8.10 Determine the inverse Fourier transform of the frequency function $P(f)$ defined in Eq. (8.60).

8.11 An analog signal is sampled, quantized, and encoded into a binary PCM wave. The specifications of the PCM system include the following:

Sampling rate = 8 kHz
Number of representation levels = 64

The PCM wave is transmitted over a baseband channel using discrete pulse-amplitude modulation. Determine the minimum bandwidth required for transmitting the PCM wave if each pulse is allowed to take on the following number of amplitude levels: 2, 4, or 8.

8.12 Consider a baseband binary PAM system that is designed to have a raised-cosine spectrum $P(f)$. The resulting pulse $p(t)$ is defined in Eq. (8.62). How would this pulse be modified if the system was designed to have a linear phase response?

8.13 A computer puts out binary data at the rate of 56 kb/s. The computer output is transmitted using a baseband binary PAM system that is designed to have a raised-cosine spectrum. Determine the transmission bandwidth required for each of the following roll-off factors: $\alpha = 0.25, 0.5, 0{,}75, 1.0$.

8.14 Repeat Problem 8.13, given that each set of three successive binary digits in the computer output are coded into one of eight possible amplitude levels, and the resulting signal is transmitted using an eight-level PAM system designed to have a raised-cosine spectrum.

8.15 An analog signal is sampled, quantized, and encoded into a binary PCM wave. The number of representation levels used is 128. A synchronizing pulse is added at the end of each code word representing a sample of the analog signal. The resulting PCM wave is transmitted over a channel of bandwidth 12 kHz using a quaternary PAM system with raised-cosine spectrum. The roll-off factor is unity.

(a) Find the rate (b/s) at which information is transmitted through the channel.

(b) Find the rate at which the analog signal is sampled. What is the maximum possible value for the highest frequency component of the analog signal?

8.16 A binary PAM wave is to be transmitted over a baseband channel with an absolute maximum bandwidth of 75 kHz. The bit duration is 10 μs. Find a raised-cosine spectrum that satisfies these requirements.

8.17 A digital communication system uses the one-sided exponential for the pulse shape to transmit data

$$p(t) = \begin{cases} 0 & t < 0 \\ \exp(-t/\tau) & t \geq 0 \end{cases}$$

(a) What is the worst case intersymbol interference with this pulse shape if $\tau = T$ where T is the symbol period?

(b) If a 20 percent reduction of the eye opening due to intersymbol interference is tolerable what is the optimum value of τ? What is the difference in 3-dB bandwidth of the signal with this τ compared to $\tau = T$?

8.18 The combination of the transmitter and channel produces a pulse with its amplitude spectrum defined by

$$H(f) = \exp(-|f|T)$$

Determine the spectrum of the ideal receive filter that would eliminate intersymbol interference.

8.19 The raised cosine pulse spectrum is not the only one that satisfies Nyquist criteria. The trapezoid pulse spectrum shown in Figure P8.19 also satisfies this criteria.

(a) Compute the time-domain pulse corresponding to the spectrum shown in Figure P8.19. Compare the zero crossings of this pulse with those of the raised cosine pulses and the sinc pulse.

(b) Suggest another pulse spectrum that satisfies Nyquist's criterion.

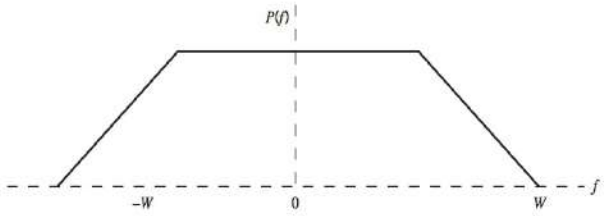


Figure P8.19

8.20 Consider a baseband M-ary system using M discrete amplitude levels. The receiver model is as shown in Figure P8.20; the operation of which is governed by the following assumptions:

(a) The signal component in the received wave is

$$m(t) = \sum_n a_n \operatorname{sinc}\left(\frac{t}{T} - n\right)$$

where $1/T$ is the signaling rate in bauds.

(b) The amplitude levels are $a_n = \pm A/2, \pm 3A/2, \ldots, \pm(M-1)A/2$ if M is even, and $a_n = 0, \pm A, \ldots, \pm(M-1)A/2$ if M is odd.

(c) The M levels are equiprobable, and the symbols transmitted in adjacent time slots are statistically independent.

(d) The noise $w(t)$ at the receiver input is white and Gaussian with zero mean and power spectral density $N_0/2$.

(e) The low-pass filter is ideal with bandwidth $B = 1/2T$.

(f) The threshold levels used in the decision device are $0, \pm A, \ldots, \pm(M-3)A/2$ if M is even, and $\pm A/2, \pm 3A/2, \ldots, \pm(M-3)A/2$ if M is odd.

The average probability of symbol error in this system is defined by

$$P_e = 2\left(1 - \frac{1}{M}\right)Q\left(\frac{A}{2\sigma}\right)$$

where σ is the standard deviation of the noise at the input of the decision device. Demonstrate the validity of this general formula by determining P_e for the following three cases: $M = 2, 3, 4$.

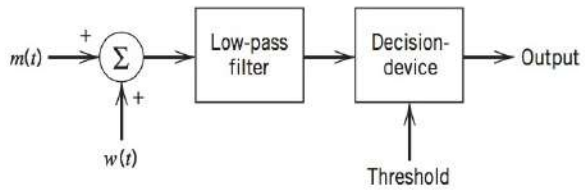


Figure P8.20

8.21 Suppose that in a baseband M-ary PAM system with M equally likely amplitude levels, as described in Problem 8.20, the average probability of symbol error P_e is less than 10^{-6} so as to make the occurrence of decoding errors negligible. Show that the minimum value of received signal-to-noise ratio in such a system is approximately given by

$$(\text{SNR})_{R,\min} \simeq 7.8(M^2 - 1)$$

8.22 Some radio systems suffer from *multipath distortion*, which is caused by the existence of more than one propagation path between the transmitter and the receiver. Consider a channel the output of which, in response to a signal $s(t)$, is defined by (in the absence of noise)

$$x(t) = K_1 s(t - t_{01}) + K_2 s(t - t_{02})$$

where K_1 and K_2 are constants, and t_{01} and t_{02} represent transmission delays. It is proposed to use the three-tap delay-line-filter of Figure P8.22 to equalize the multi-path distortion produced by this channel.

(a) Evaluate the transfer function of the channel.

(b) Evaluate the parameters of the tapped-delay-line filter in terms of K_1, K_2, t_{01} and t_{02}, assuming that $K_2 \ll K_1$ and $t_{02} > t_{01}$.

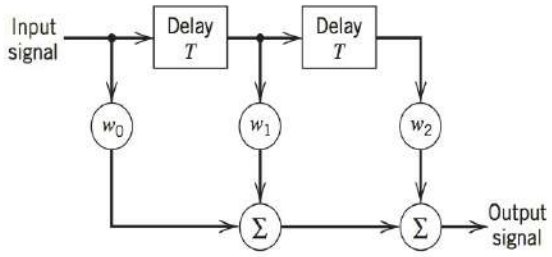


Figure P8.22

8.23 Let the sequence $\{x(nT)\}$ denote the input applied to a tapped-delay-line qualizer. Show that intersymbol interference is eliminated completely by the equalizer provided that its transfer function satisfies the condition

$$H(f) = \frac{T}{\sum_k X(f - k/T)}$$

where T is the symbol duration.

As the number of taps in the equalizer approaches infinity, the transfer function of tile equalizer becomes a Fourier series with real coefficients and can therefore approximate any function in the interval $(-1/2T, 1/2T)$. Demonstrate this property of the equalizer.

Computer Problems

8.24 The following Matlab script simulates the transmission of data through the telephone channel of Example 8.2 using a bipolar NRZ line code with a symbol rate of 1.6 kHz.

```
Fs = 32;          % sample rate (kHz)
Rs = 1.6;         % symbol rate (kHz)
Ns = Fs/Rs;       % number of samples per symbol
Nb = 30;          % number of bits to simulate

% - - - Discrete B(z)/A(z) model of telephone channel - - -
A = [1.00,-2.838,3.143,-1.709,0.458,-0.049];
B = 0.1*[1,-1];

% - - - Pulse shape the data - - - - -
pulse = [ones(1,Ns)];   % bipolar NRZ pulse
data = sign(randn(1,Nb));
```

```
Sig = pulse' * data;
Sig = Sig(:);

% - - - Pass signal through the channel - - - -
RxSig = filter(B,A,Sig);

% - - - Plot results - - - - - - - - - - - - - - - - - - - - - - -
plot(real(RxSig))
hold on, plot(Sig,'r'). hold off
xtabel ('Time samples'), ylabel('Amplitude')
```

(a) Modify the script to simulate transmission of a Manchester line code at 1.6 kHz. Compare your results with those of Example 8.2. Increase the data rate to 3.2 kHz and repeat.

(b) Process the channel output with a filter matched to the pulse shape. Use the following Matlab function to plot the eye diagram of the signal. Increase the number of bits simulated when plotting eye diagrams. Comment on the eye opening for the 1.6 and 3.2 kHz cases.

```
function b = ploteye(s,Ns);
% - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
% Inputs
%   s - real signal
%   Ns - oversample rate
% - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
f = mod(length(s), Ns);
s = real(s(l:end-f));   % make length multiple of Ns

%- - - extract individual symbol periods from signal - - -
EyeSigRef   = reshape(s, Ns, length(s)/Ns);   % one
                                              symbol per column
EyeSigm1    = EyeSigRef(Ns/2;Ns,1:end-2);   % last half
                                              of preceding symbol
EyeSig0     = EyeSigRef(:  , 2;end-1);   % current
                                              symbol
EyeSigp1    = EyeSigRef(1: Ns/2, 3:end),   % first half of
                                              following symbol
EyeSig      = [EyeSigm1; EyeSig0; EyeSigp1];   % tack
                                              together for curve

L = size(EyeSig,1);
plot([0:L-1]/Ns,EyeSig)                       % plot multiple curves
```

(c) Modify the Manchester code script to produce an M-ary line code with $M = 4$. Compare the eye opening with $M = 4$ for the 1.6 and 3.2 kHz symbol rates. Comment on the drawbacks of transmitting data at the higher rate.

8.25 The following Matlab script simulates the transmission of a pulse having a raised cosine spectrum through a channel with an impulse response given by

$$h(t) = \begin{cases} 0 & t < 0 \\ \exp(-t/\tau) & t \geq 0 \end{cases}$$

```
T     = 1;          % symbol period
Rs    = l/T;        % symbol rate
Ns    = 16;         % number of samples per symbol
Fs    = Rs*Ns;      % sample rate (kHz)
Nb    = 3000;       % number of bits to simulate
alpha = 1.0;        % rolloff of raised cosine

% - - - Discrete model of telephone channel - - -
t = [0: 1/Fs: 5*T];
h = exp(-t / (T/2)) /Fs;   % impulse response scaled for
                           % sample rate

% - - - Pulse shape the data - - - - -
pulse   = firrcos(5*Ns, Rs/2, Rs*alpha, Fs);   % 100
                                               % raised cosine filter
data    = sign(randn(1,Nb));                   % random binary data
Udata   = [1; zeros(Fs-1, 1)] * data;          % upsample data
Udata   = Udata(:);                            % "
Sig     = filter (pulse, 1, Udata);            % pulse shape data
Sig     = Sig( (length(pulse)-1)/2: end);      % remove filter
                                               % delay

% - - - Pass signal through the channel - - - -
RxSig = filter(h, 1, Sig);

% - - - Plot results - - - - - - - - - - - - - - - - - - -
plot(real(RxSig))
hold on, plot(Sig, 'r'), hold off
xlabel('Time samples'), ylabel('Amplitude')
```

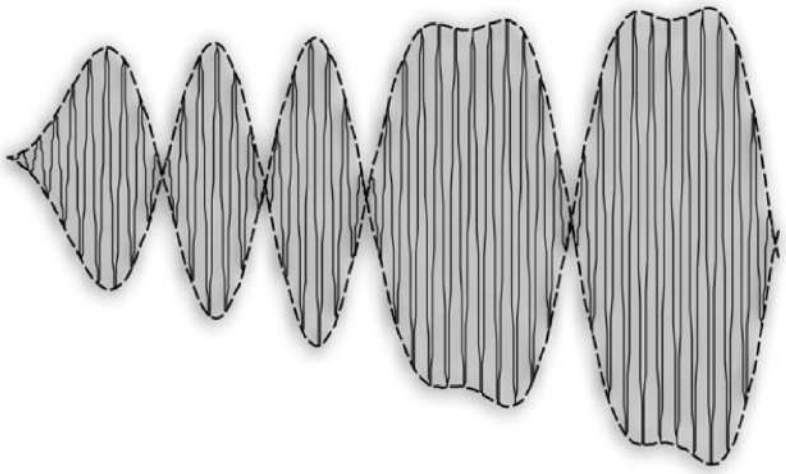
(a) For $\tau = T/2$ where T is the symbol period, simulate the transmission for rolloff factors $\alpha = 0.5$ and 1.0. Determine the amount by which the eye opening closes in each case and comment.

(b) Repeat (a) for the case $\tau = T$.

8.26 For part (a) of the Problem with $\alpha = 0.5$, compute a fixed tapped equalizer using the technique described in Section 8.8. (The set of linear equations $\mathbf{C}w = b$ may be solved in Matlab using the statement $w = \text{inv}(\mathbf{C})\, b$.)

(a) Compute the coefficients of a 3-tap T-spaced equalizer and apply it to the simulation output and plot the eye diagram. Quantify the improvement in the eye opening.

(b) Repeat part (a) for a 5-tap equalizer and comment on the tradeoffs of using a 3- versus 5-tap equalizer.



9

DIGITAL BAND-PASS TRANSMISSION TECHNIQUES

9.1 INTRODUCTION

In baseband pulse transmission, which we studied in the previous chapter, a data stream in the form of a discrete pulse-amplitude modulated (PAM) signal is transmitted directly over a low-pass channel. An issue of particular concern in baseband pulse transmission is the design of pulse-shaping to bring the intersymbol interference (ISI) problem under control. In digital band-pass transmission, on the other hand, the incoming data stream is modulated onto a carrier (usually sinusoidal) with fixed frequency limits imposed by a band-pass channel of interest. The major issue of concern here is the optimum design of the receiver so as to minimize the average probability of symbol error in the presence of channel noise. This does not mean, of course, that noise is of no concern in baseband pulse transmission, nor does it mean that ISI is of no concern in digital carrier modulation. It merely points out the issues that are of high priority in these two different domains of data transmission.

The communication channel used for band-pass data transmission may be a wireless link in a local area network, a satellite channel, or the like. In any event, the modulation process making the transmission possible involves switching (keying) the amplitude, frequency, or phase of a sinusoidal carrier in some fashion in accordance with the incoming data. Thus, there are three basic signaling schemes known as amplitude-shift keying (ASK), frequency-shift keying (FSK), and phase-shift keying (PSK), which may be viewed as special cases of amplitude modulation, frequency modulation, and phase modulation, respectively. A distinguishing feature of FSK and PSK signals is that ideally they both have a constant envelope. This feature makes them impervious to amplitude nonlinearities commonly encountered in radio links and satellite channels. It is for this reason that we find that, in practice, FSK and PSK signals are preferred to ASK signals for band-pass transmission over nonlinear channels.

In this chapter we study digital carrier modulation techniques with an emphasis on the following issues: (1) optimum design of the receiver in the sense that it will make fewer

errors in the long run than any other receiver, and (2) calculation of the average probability of symbol error at the receiver output. Two different cases are considered in the study: coherent receivers and noncoherent receivers. In a coherent receiver the receiver is phase-locked to the transmitter, whereas in a noncoherent receiver there is no phase synchronization between the local oscillator used in the receiver for demodulation and the oscillator supplying the sinusoidal carrier in the transmitter for modulation.

9.2 BAND-PASS TRANSMISSION MODEL

Before we discuss specific band-pass transmission strategies, let us review band-pass signals and systems. In Chapter 2, we introduced band-pass signals and showed how they could be studied by means of their equivalent *complex envelope* representations. In particular, we illustrated how a band-pass signal could be constructed from its low-pass *in-phase* and *quadrature* components. This construction is repeated in Figure 9.1 where the in-phase and quadrature components, $g_I(t)$ and $g_Q(t)$, are used to modulate the orthogonal carriers $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$ to produce the band-pass signal $g(t)$. The *signal encoder* shown on the left-hand side of Figure 9.1 maps the data from the message source onto the in-phase and quadrature components. It is the differences in how this mapping is performed that determine the specific band-pass transmission strategy.

The band-pass signal $g(t)$ is transmitted over a communications channel before it arrives at its destination. In most practical situations, the channel attenuates the signal before it reaches the receiver as shown in Figure 9.2. We represent this attenuation by the factor A_c. In addition to the attenuation, the bandpass communication channel coupling the transmitter to the receiver is assumed to have two characteristics:

1. The channel is linear. Often we shall assume the channel bandwidth is wide enough to accommodate the transmission of the modulated signal $g(t)$ with negligible or no distortion. In other situations, the effects of the band-pass channel may be modeled directly by the effect of its complex baseband impulse response on the complex envelope of the signal, as described in Section 2.10.
2. The received signal $s(t)$ is perturbed by an additive stationary white Gaussian noise process of zero mean and power spectral density $N_0/2$. A sample function of this noise process is denoted by $w(t)$.

The situation where the channel only attenuates the signal and adds noise, that is,

$$\begin{aligned} x(t) &= s(t) + w(t) \\ &= A_c g(t) + w(t) \end{aligned} \tag{9.1}$$

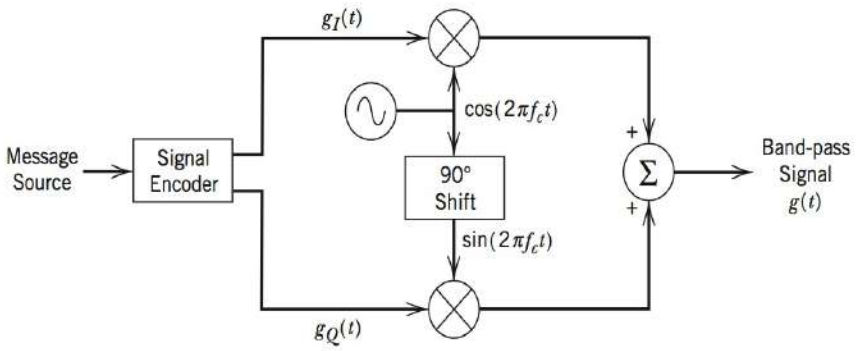


FIGURE 9.1 Block diagram showing construction of band-pass signal from its in-phase and quadrature components.

is a reasonable model for many practical channels. We refer to such an idealized channel as an *additive white Gaussian noise* (AWGN) *channel.*

In Figure 9.3, we show the receiver portion of the band-pass transmission model. All receivers include a bandpass filter at the front end that passes the signal undistorted and converts the white noise to *narrowband noise* $n(t)$ as discussed in Section 5.11. The in-phase and quadrature components of this band-pass signal plus narrowband noise are derived using the I-Q down-converter that was presented in Section 2.9; this down-converter consists of an oscillator that produces two local sinusoids in phase with the orthogonal carriers $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$, which are mixed with the incoming signal. The mixer outputs are then low-pass filtered to remove high-frequency components, leaving the received in-phase and quadrature signals, $\frac{1}{2}[A_c g_I(t) + n_I(t)]$ and $\frac{1}{2}[A_c g_Q(t) + n_Q(t)]$, respectively. Practical receivers will also include several amplification stages, but since these components amplify both the signal and the noise, we can exclude them without affecting the results.

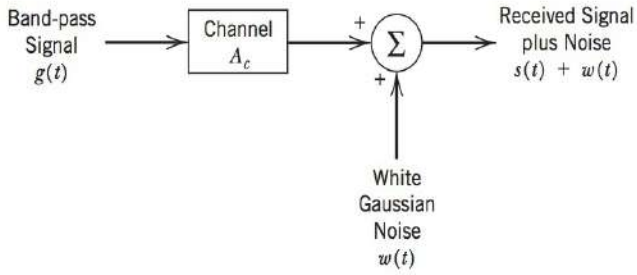


FIGURE 9.2 Block diagram showing channel model for band-pass transmission.

As described in Section 5.11, the noise processes $n_I(t)$ and $n_Q(t)$ have low-pass nature. In general, if the filters in the system pass the signal undistorted, then these noise processes will be *white over the bandwidth of the signal.* Since the detector is only interested in the signal, we may assume the noise processes $n_I(t)$ and $n_Q(t)$ are white without bandwidth limitation, and analyze the complex envelope in a manner similar to the analysis of a baseband signal received in white noise: the difference is that a complex envelope and the associated noise are complex-valued processes, while a baseband signal is real valued.

The distinguishing feature of the band-pass receiver is the *signal detector*, which depends upon how the signal encoding is performed, that is, on the specific transmission strategy. The receiver has the task of observing the complex representation of the received signal, $[g_I(t) + n_I(t)] + j[g_Q(t) + n_Q(t)]$, for a duration of T seconds and making its best estimate of the corresponding transmitted signal $g_I(t) + jg_Q(t)$ or, equivalently, the data symbol 0 or 1, for binary data.

To simplify the treatment, we shall assume that the receiver is *time synchronized* with the transmitter, which means that the receiver knows the instants of time when the modulation changes state. In practice, a receiver must include a timing recovery circuit. Sometimes, it is also assumed that the receiver is phase-locked to the transmitter. In such a case, we speak of *coherent detection*, and we refer to the receiver as a *coherent receiver.* On the other hand, there may be no phase synchronism between transmitter and receiver. In this second case, we speak *of noncoherent detection*, and we refer to the receiver as a *noncoherent receiver.* In this chapter, we assume the existence of time synchronism. However, we shall distinguish between *coherent* and *noncoherent* detection.

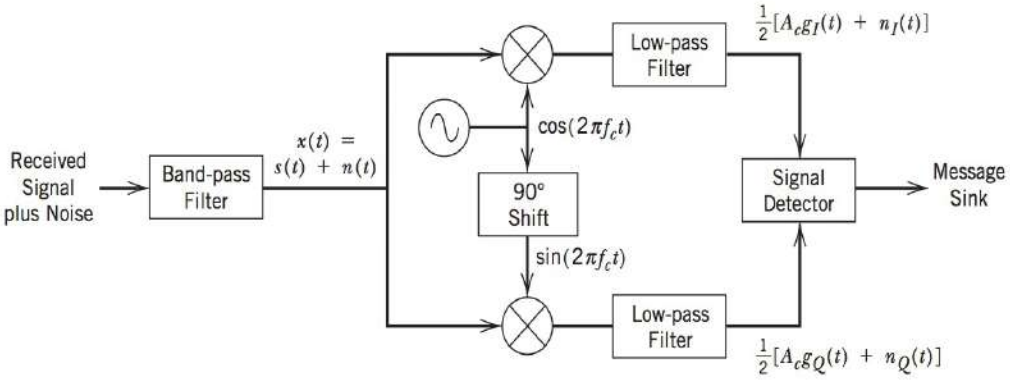


FIGURE 9.3 Block diagram showing analysis of a band-pass signal into its in-phase and quadrature components.

In summary, it should be pointed out that not all band-pass transmitters and receivers are implemented exactly as shown in Figures 9.1 and 9.3. Some simplifications or modifications may be made depending upon the available technology and transmission strategy. For example:

- Some transmission strategies use only in-phase signaling. This usually implies that the hardware associated with processing the quadrature component is removed.
- A receiver may perform noncoherent detection as opposed to coherent detection. Noncoherent detection implies that the oscillators used to derive the in-phase and quadrature components do not need to be phase-locked to the incoming carrier. In fact, with noncoherent detection of some transmission strategies, the message may often be recovered directly from the band-pass signal without deriving the in-phase and quadrature components. However, this altered implementation does not improve the performance obtained.
- In many modern receivers, the derivation of the in-phase and quadrature components is performed with oscillators that are not phase-locked to the incoming signal. This implies that there may be a phase rotation or even a small frequency error on the recovered in-phase and quadrature components. These potential phase and frequency errors are corrected by digital signal processing algorithms that are part of the signal detection strategy.

As the following will show, bandpass communications are richer than baseband communications in the sense that the quadrature modulation offers more possibilities for signaling techniques.

9.3 TRANSMISSION OF BINARY PSK AND FSK[1]

When it is required to transmit binary data (e.g., obtained by digitizing voice or video signals) over band-pass communication channels such as radio links or satellite channels, it is necessary to modulate the signal onto a carrier wave (usually sinusoidal) with fixed frequency limits set by the particular channel. The modulation process corresponds to switching or keying the amplitude, frequency, or phase of the carrier between either of two possible values corresponding to binary symbols 0 and 1. This results in three basic signaling techniques, namely, *amplitude-shift keying* (ASK), *frequency-shift keying* (FSK), and *phase-shift keying* (PSK), as described below:

1. In an ASK system, binary symbol 1 is represented by transmitting a sinusoidal carrier wave of fixed amplitude and fixed frequency for the bit duration T_b seconds, whereas binary symbol 0 is represented by switching off the carrier for T_b seconds, as illustrated in Figure 9.4*a*. For example, an ASK signal may be generated by using the on–off form of representation for the input binary data and then applying it, together with the carrier, to a product modulator.
2. In an FSK system, two sinusoidal waves of the same amplitude but different frequencies are used to represent binary symbols 1 and 0. An FSK signal may be generated by using the bipolar form of representation for the input binary data and then applying it to a voltage controlled oscillator, or by switching between two oscillators. The mathematical distinction between these two methods may be considered as shifting the frequency with *continuous phase* or *discontinuous phase*, respectively. The case of an FSK signal with continuous phase is illustrated in Figure 9.4*b*.
3. In a PSK system, a sinusoidal carrier wave of fixed amplitude and fixed frequency is used to represent both symbols 1 and 0, except that whenever symbol 0 is transmitted

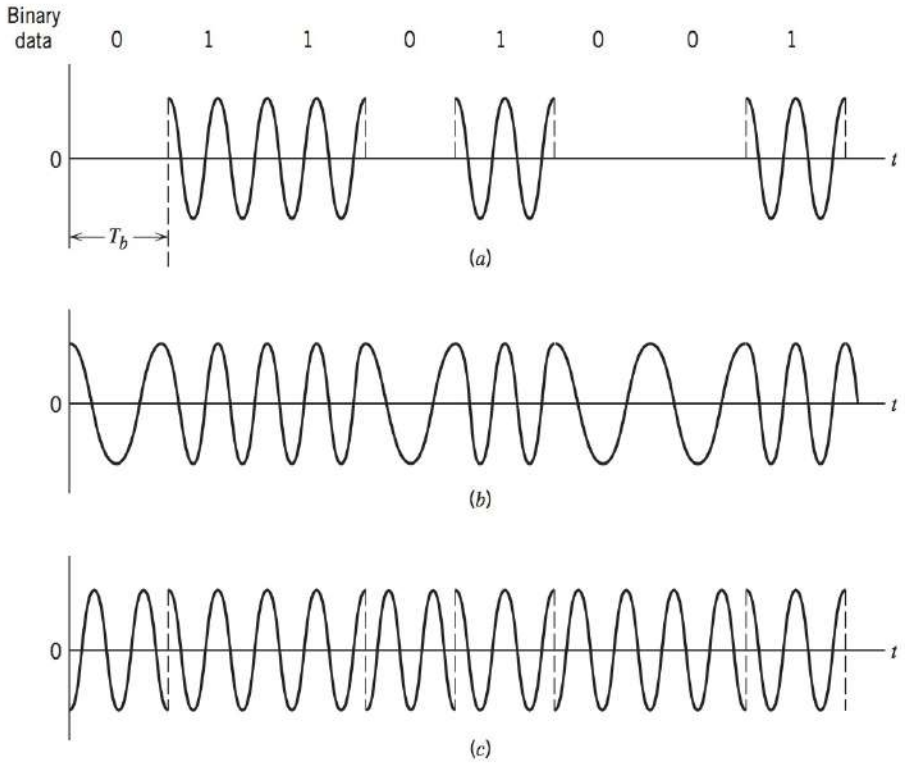


FIGURE 9.4 Illustrating the three basic forms of signaling binary information: (*a*) amplitude-shift keying; (*b*) frequency-shift keying with continuous phase; and (*c*) phase-shift keying.

the carrier phase is shifted by 180 degrees, as illustrated in Figure 9.4(c). For example, a PSK signal may be generated by representing the input binary data in bipolar form and applying it, together with the carrier, to a product modulator.

It is apparent, therefore, that ASK, FSK, and PSK signals are special cases of amplitude-modulated, frequency-modulated, and phase-modulated waves, respectively. It is of interest to note that although, in general, it is not easy to distinguish between frequency-modulated and phase-modulated waves (on an oscilloscope, say), this is not so in the case of FSK and PSK signals, as it may be readily ascertained by comparing the waveforms in parts *b* and *c* of Figure 9.4.

For binary PSK (BPSK) modulation, the steps of the modulation process may be broken down as shown in Figure 9.5. The binary data are used to create a bipolar line code of Figure 9.5*b*, which represents the in-phase component of the complex baseband signal $g_I(t)$. For binary PSK, $g_Q(t) = 0$. The in-phase component modulates a carrier as shown in Figure 9.1, to produce the band-pass signal $g(t) = g_I(t)\cos(2\pi f_c t)$ of Figure 9.5*c*. In particular, the binary PSK (BPSK) signal is described by

$$\begin{aligned} s_1(t) &= A_c \cos(2\pi f_c t) && \text{for symbol 1} \\ s_0(t) &= -A_c \cos(2\pi f_c t) && \text{for symbol 0} \\ &= A_c \cos(2\pi f_c t + \pi) \end{aligned} \tag{9.2}$$

For binary ASK modulation, the steps are similar except that a unipolar line code is used to represent the data at baseband.

For binary FSK modulation, the complex baseband equivalent cannot be represented as a simple line code. Binary FSK may be represented as a pair of related line codes, $g_I(t)$ and $g_Q(t)$, which is a level of complexity beyond that observed with BPSK and FSK. For this reason, early FSK transmitters used FM modulators rather than the linear modulation scheme shown in Figure 9.1.

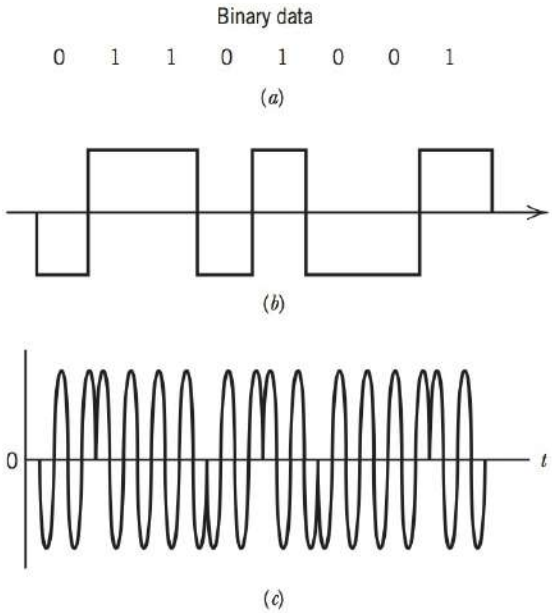


FIGURE 9.5 The steps of binary PSK modulation: (*a*) binary data, (*b*) bipolar line code; and (*c*) binary PSK waveform.

The band-pass binary FSK signal may be described directly as

$$\begin{aligned} s_1(t) &= A_c \cos(2\pi f_1 t) \qquad \text{for symbol 1} \\ s_0(t) &= A_c \cos(2\pi f_0 t) \qquad \text{for symbol 0} \end{aligned} \tag{9.3}$$

If we define the carrier frequency f_c as the midpoint between f_1 and f_0, that is, $f_c = (f_1 + f_0)/2$, and assuming $f_1 > f_0$, also define $\Delta f = (f_1 - f_0)/2$, then we may represent the FSK signal of Eq. (9.3) as

$$\begin{aligned} s_1(t) &= A_c \cos[2\pi(f_c + \Delta f)t] \\ &= \text{Re}\{A_c \exp[j2\pi(f_c + \Delta f)t]\} \end{aligned} \qquad \text{for symbol 1} \tag{9.4}$$

$$\begin{aligned} s_0(t) &= A_c \cos[2\pi(f_c - \Delta f)t] \\ &= \text{Re}\{A_c \exp[j2\pi(f_c - \Delta f)t]\} \end{aligned} \qquad \text{for symbol 0} \tag{9.5}$$

By inspection of Eqs. (9.4) and (9.5), we identify the complex baseband equivalent signals as

$$\begin{aligned} g_I(t) + jg_Q(t) &= A_c \exp[-j2\pi\Delta f t] \qquad \text{for symbol 1} \\ g_I(t) + jg_Q(t) &= A_c \exp[+j2\pi\Delta f t] \qquad \text{for symbol 0} \end{aligned} \tag{9.6}$$

While this complex baseband equivalent has a compact description, it certainly differs from the baseband line codes that were discussed in Chapter 8.

In the following, we will evaluate the performance of different forms of FSK and PSK receivers in the presence of additive white Gaussian noise. For BPSK we may apply the baseband analysis of Section 8.3 directly. However, for the reason mentioned above, this analysis cannot be extrapolated to FSK. Consequently, in this section, we shall use a more generic approach to analyzing receiver performance and apply it to both PSK and FSK. For analysis of an ASK system, the reader is referred to Problem 9.4.

COHERENT DETECTION OF FSK AND PSK SIGNALS

Let $s_0(t)$ and $s_1(t)$ denote the signals used to represent binary symbols 0 and 1, respectively. We may then distinguish between FSK and PSK signals, as follows:

(a) *FSK Signals*

$$\begin{aligned} s_1(t) &= A_c \cos(2\pi f_1 t), \qquad \text{for symbol 1} \\ s_0(t) &= A_c \cos(2\pi f_0 t), \qquad \text{for symbol 0} \end{aligned} \tag{9.7}$$

(b) *BPSK Signals*

$$\begin{aligned} s_1(t) &= A_c \cos(2\pi f_c t), \qquad \text{for symbol 1} \\ s_0(t) &= A_c \cos(2\pi f_c t + \pi), \qquad \text{for symbol 0} \end{aligned} \tag{9.8}$$

In both Eqs. (9.7) and (9.8), we have $0 \leq t \leq T_b$. We usually find that in the case of FSK signals, the frequencies f_1 and f_0 are both large compared with the bit rate $1/T_b$, whereas in the case of PSK signals, f_c is large compared with $1/T_b$. Then, it is apparent that in both FSK and PSK signals, the same signal energy E_b is transmitted in a bit interval T_b, as shown by

$$\begin{aligned} E_b &= \int_0^{T_b} s_0^2(t)dt = \int_0^{T_b} s_1^2(t)dt \\ &= \frac{A_c^2 T_b}{2} \end{aligned} \tag{9.9}$$

Assuming that the transmitted carrier phase and frequency are known exactly, we may use the two-path correlation receiver shown in Figure 9.6. This receiver uses two correlators or matched filters, one for the transmitted signal $s_0(t)$ and the other for $s_1(t)$.

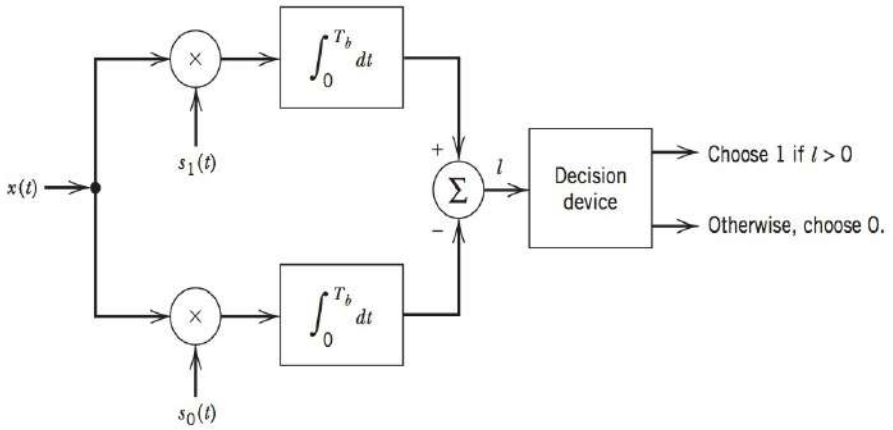


FIGURE 9.6 Two-path correlation receiver for the general case.

For FSK and PSK signals, the corresponding receiver structures are as shown in parts (*a*) and (*b*) of Figure 9.7, respectively. In the case of PSK signals, the receiver reduces to a single path, because $s_0(t)$ is the negative of $s_1(t)$. In Figure 9.7, it is also assumed that the integrator knows when a bit interval begins and ends.

To evaluate the performance of the receiver of Figure 9.7*a*, we will assume that the additive front-end receiver noise $w(t)$ is white Gaussian noise of zero-mean and spectral density $N_0/2$. Thus, the received signal is defined by

$$\begin{aligned} H_0&: \quad x(t) = s_0(t) + w(t) \\ H_1&: \quad x(t) = s_1(t) + w(t) \end{aligned} \tag{9.10}$$

where hypotheses H_0 and H_1 correspond to the transmission of symbols 0 and 1, respectively.

The receiver output l is given by

$$l = \int_0^{T_b} x(t)[s_1(t) - s_0(t)]\, dt \tag{9.11}$$

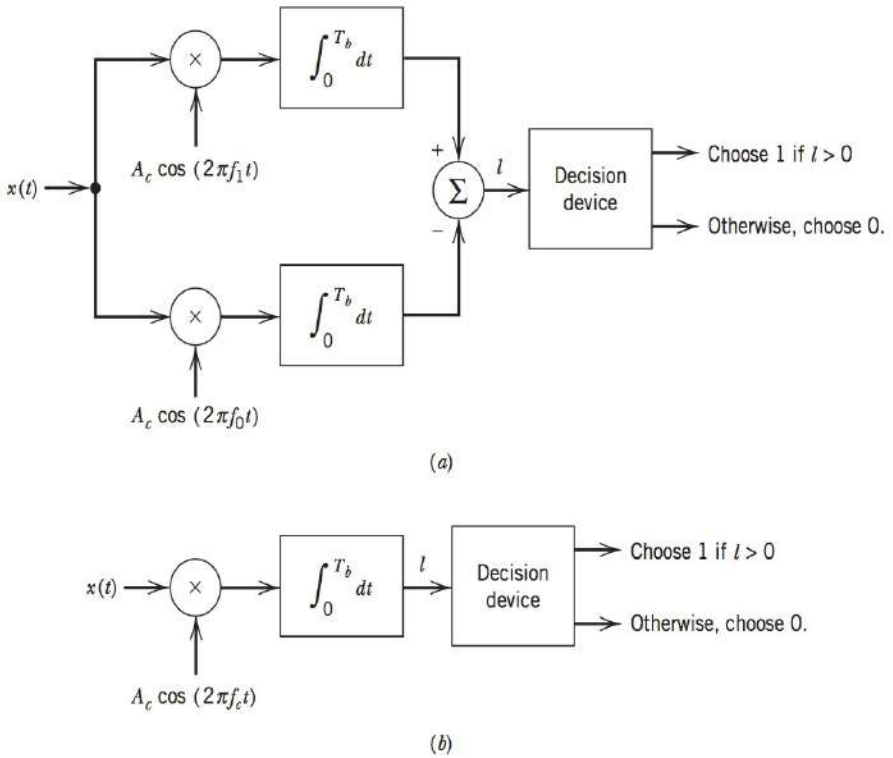


FIGURE 9.7 (*a*) Coherent receiver for FSK signals. (*b*) Coherent receiver for PSK signals.

The output l is compared with a decision level of zero volts. If l is greater than zero, the receiver chooses symbol 1; otherwise, it chooses symbol 0. Since the noise $w(t)$ is a sample function of a Gaussian process, it follows from the definition of a Gaussian process that the receiver output l is a Gaussian random variable. The mean value of l is dependent on whether symbol 1 or 0 is transmitted. Suppose we know that symbol 1 was transmitted. Then, we may write

$$H_1: \qquad l = \int_0^{T_b} s_1(t)[s_1(t) - s_0(t)]\,dt + \int_0^{T_b} w(t)[s_1(t) - s_0(t)]\,dt \tag{9.12}$$

Since the noise $w(t)$ is of zero-mean, it follows from Eq (9.12) that the random variable L, whose value is l, has the conditional mean

$$\begin{aligned}\mathbf{E}[L|H_1] &= \int_0^{T_b} s_1(t)[s_1(t) - s_0(t)]\,dt \\ &= E_b(1-\rho)\end{aligned} \tag{9.13}$$

The parameter ρ is the *correlation coefficient* of the signals $s_0(t)$ and $s_1(t)$, defined by

$$\begin{aligned}\rho &= \frac{\displaystyle\int_0^{T_b} s_0(t)s_1(t)\,dt}{\left[\displaystyle\int_0^{T_b} s_0^2(t)\,dt \int_0^{T_b} s_1^2(t)\,dt\right]^{1/2}} \\ &= \frac{1}{E_b}\int_0^{T_b} s_0(t)s_1(t)\,dt\end{aligned} \tag{9.14}$$

which has an absolute value that is less than or equal to unity. Similarly, we can show that the conditional mean of L, given that symbol 0 was transmitted, is defined by

$$\mathbf{E}[L|H_0] = -E_b(1-\rho) \tag{9.15}$$

The random variable L has the same variance, regardless of whether symbol 1 or 0 was transmitted, as shown by

$$\begin{aligned}\mathrm{Var}[L] &= \mathbf{E}[\{L - E[L]\}^2] \\ &= \mathbf{E}\left[\int_0^{T_b}\int_0^{T_b} w(t)w(u)[s_1(t) - s_0(t)][s_1(u) - s_0(u)]\,dt\,du\right] \\ &= \int_0^{T_b}\int_0^{T_b} [s_1(t) - s_0(t)][s_1(u) - s_0(u)]R_W(t,u)\,dt\,du\end{aligned} \tag{9.16}$$

where $R_W(t, u) = \mathbf{E}[W_t W_u]$ is the autocorrelation function of $w(t)$. Since $w(t)$ is a white noise process of spectral density $N_0/2$, we have

$$R_W(t,u) = \frac{N_0}{2}\delta(t-u) \tag{9.17}$$

Therefore, substituting Eq. (9.17) in (9.16), and using the sifting property of the delta function, we get

$$\begin{aligned}\mathrm{Var}[l] &= \frac{N_0}{2}\int_0^{T_b}\int_0^{T_b} [s_1(t) - s_0(t)][s_1(u) - s_0(u)]\delta(t-u)\,dt\,du \\ &= \frac{N_0}{2}\int_0^{T_b} [s_1(t) - s_0(t)]^2 dt \\ &= N_0E_b(1-\rho)\end{aligned} \tag{9.18}$$

We assume that symbols 1 and 0 occur with equal probability. An *error of the first kind* occurs whenever we transmit symbol 0, but the receiver output l is greater than zero volts, and the receiver therefore chooses symbol 1. An *error of the second kind* occurs whenever we transmit symbol 1, but the receiver output l is less that zero volts, and the receiver therefore chooses symbol 0. From the symmetry of the receiver of Figure 9.6*a*, it is apparent that probabilities of both kinds of error are equal. Thus, because l is a Gaussian random variable of mean $\pm E_b(1 - \rho)$ and variance $N_0E_b(1 - \rho)$, we find that the average probability of error in the receiver of Figure 9.6 is given in terms of the Q-function by the formula (by analogy with Section 8.3)

$$\begin{aligned} P_e &= P(l>0|H_0) = P(l<0|H_1) \\ &= Q\left(\sqrt{\frac{E_b(1-\rho)}{N_0}}\right) \end{aligned} \tag{9.19}$$

In the case of a coherent PSK receiver, we see from Eq. (9.8) that $s_0(t) = -s_1(t)$. Substituting this relation in Eq. (9.14), we obtain $\rho = -1$. A pair of signals $s_0(t)$ and $s_1(t)$ for which the correlation coefficient is $\rho = -1$ are called *antipodal signals*. Thus, putting $\rho = -1$ in Eq (9.19) gives the probability of error in a PSK system using coherent detection as follows:

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{9.20}$$

On the other hand, in a coherent FSK receiver for which the carrier frequencies f_0 and f_1 are spaced far enough apart to justify treating $s_0(t)$ and $s_1(t)$ as *orthogonal signals*, we have $\rho = 0$ (see Problem 9.8). Therefore, putting $\rho = 0$ in Eq. (9.19), we find that the probability of error in an FSK system using coherent detection is given by

$$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{9.21}$$

COHERENT DETECTION OF CONTINUOUS PHASE FREQUENCY-SHIFT KEYING (CPFSK) SIGNALS BASED ON PHASE DECODING

In the coherent detection of FSK signals described above, the phase information contained in the received signal was not fully exploited, other than to provide synchronization of the receiver to the transmitter. We will now show that by using a *continuous phase frequency-shift keying* (CPFSK) signal and by proper utilization of the phase information contained in the signal, it is possible to improve the noise performance of the receiver significantly, but at the expense of increased receiver complexity.

Let the CPFSK signal be defined by

$$s(t) = A_c \cos[2\pi f_c t + \phi(t)] \tag{9.22}$$

where the phase $\phi(t)$ is a continuous function of time t. The nominal carrier frequency f_c is equal to the arithmetic mean of the two frequencies f_1 and f_0 that are used to represent symbols 1 and 0, respectively; that is,

$$f_c = \frac{1}{2}(f_1 + f_0) \tag{9.23}$$

The CPFSK signal distinguishes between binary symbols 1 and 0 as follows:

$$s(t) = \begin{cases} A_c \cos[2\pi f_1 t + \phi(0)], & \text{for symbol 1} \\ A_c \cos[2\pi f_0 t + \phi(0)], & \text{for symbol 0} \end{cases} \tag{9.24}$$

where $0 \le t \le T_b$. The phase $\phi(0)$, denoting the value of $\phi(t)$ at time $t = 0$, depends on the past history of the modulation process. Comparing Eqs. (9.22) and (9.24), and using

(9.23), we find that in the interval $0 \leq t \leq T_b$ the phase $\phi(t)$ is a linear function of time, as shown by

$$\phi(t) = \phi(0) \pm \frac{\pi h}{T_b} t \tag{9.25}$$

where the plus sign corresponds to symbol 1 and the minus sign corresponds to symbol 0. The parameter h is defined by

$$h = T_b(f_1 - f_0) \tag{9.26}$$

We refer to h as the *deviation ratio* of the frequency-shift keying signal, measured with respect to the bit rate $1/T_b$.

When the phase $\phi(t)$ is a continuous function of time, the CPFSK signal $s(t)$ itself is also continuous at all times, including the inter-bit switching instants. The spectral density of a CPFSK signal, produced by a random binary sequence, falls off at least as the inverse fourth power of frequency at frequencies remote from the center of the signal band. On the other hand, in an FSK signal with discontinuous phase, the spectral density ultimately falls off as the inverse square of frequency. *Accordingly, a CPFSK signal does not produce as much interference outside the signal band as an FSK signal with discontinuous phase.* This is a desirable characteristic when operating with a bandwidth limitation.

From Eq. (9.25), we find that at time $t = T_b$,

$$\phi(T_b) - \phi(0) = \begin{cases} \pi h, & \text{for symbol 1} \\ -\pi h, & \text{for symbol 0} \end{cases} \tag{9.27}$$

That is, the transmission of symbol 1 increases the phase of the CPFSK signal $s(t)$ by πh radians, whereas the transmission of symbol 0 reduces it by an equal amount. The possible values of $\phi(t)$ are shown in Figure 9.8. It is therefore evident that the phase shift of the CPFSK signal is an odd or even multiple of πh radians at odd or even multiples of the bit duration T_b, respectively. Since all phase shifts are modulo 2π, the case of $h = 1/2$ is of particular interest, because then the phase can take on only the two values $\pm\pi/2$ at odd multiples of T_b and only the two values 0 and π at even multiples of T_b. This is illustrated in Figure 9.9 for t equal to $-T_b$, 0, T_b, and $2T_b$. We call this special case with $h = 1/2$ *minimum shift keying* (MSK).[2] Each path from left to right through the trellis corresponds to a specific binary sequence input. For example, the bold-face line shown in Figure 9.9 corresponds to the binary sequence 011, with $\phi(-T_b) - \phi(0) = \pi/2$.

Using Eq. (9.22), we may express the MSK signal $s(t)$ in terms of its in-phase and quadrature

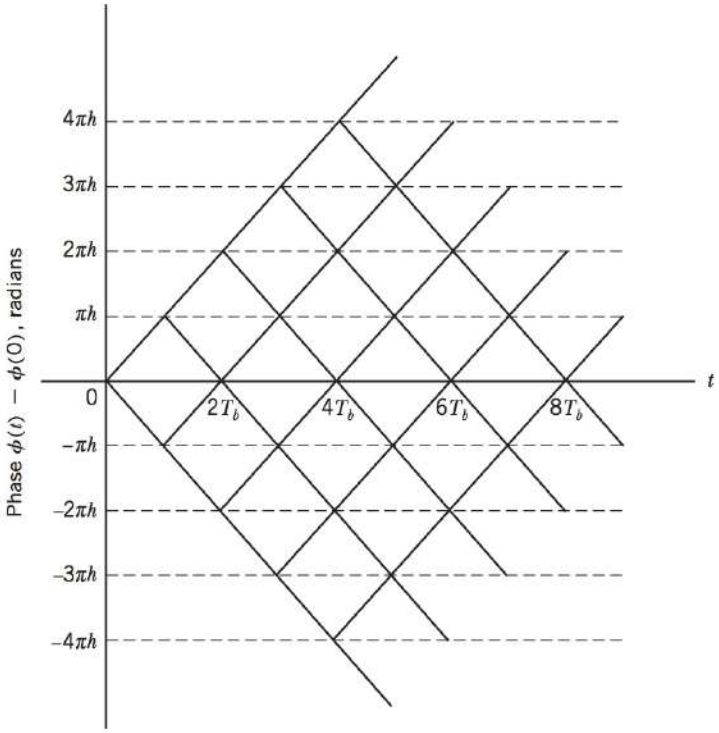


FIGURE 9.8 Possible values of the phase shift $\phi(t) - \phi(0)$.

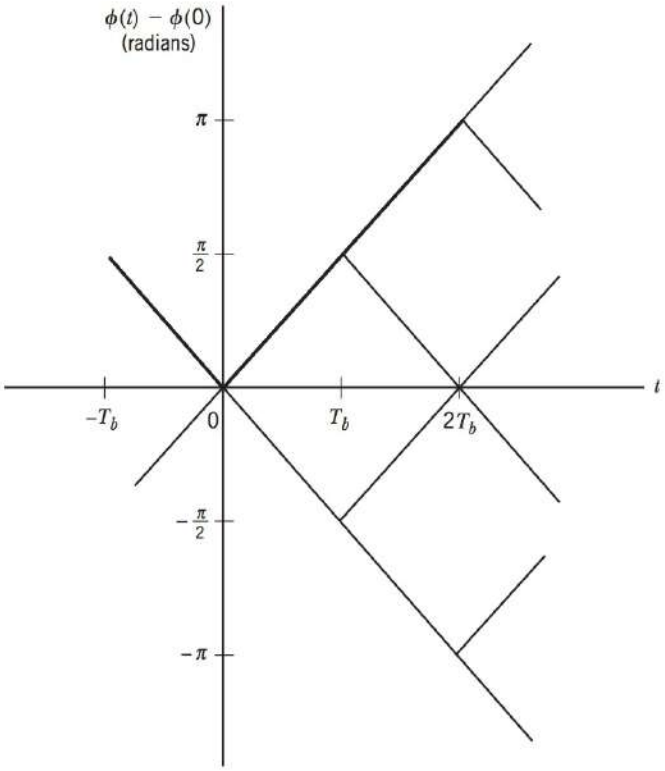


FIGURE 9.9 Possible values of the phase shift $\phi(t) - \phi(0)$ for the special case $h = \frac{1}{2}$.

components as follows:

$$s(t) = A_c \cos[\phi(t)]\cos(2\pi f_c t) - A_c \sin[\phi(t)]\sin(2\pi f_c t) \tag{9.28}$$

Consider first the in-phase component $A_c \cos[\phi(t)]$. With the deviation ratio $h = 1/2$, we have from Eq. (9.25)

$$\phi(t) = \phi(0) \pm \frac{\pi}{2T_b} t, \qquad 0 \le t \le T_b \tag{9.29}$$

where the plus sign corresponds to symbol 1 and the minus sign corresponds to symbol 0. A similar result holds for $\phi(t)$ in the interval $-T_b \le t \le 0$, except that the algebraic sign is not necessarily the same in both intervals. Since the phase $\phi(0)$ is 0 or π, depending on the past history of the modulation process, we find that in the given interval $-T_b \le t \le T_b$, the polarity of $\cos[\phi(t)]$ depends only upon $\phi(0)$, regardless of the sequence of 1's and 0's transmitted before or after $t = 0$. Thus, the in-phase component consists of a half-cosine pulse defined as follows:

$$A_c \cos[\phi(t)] = \pm A_c \cos\left(\frac{\pi}{2T_b} t\right), \qquad -T_b \le t \le T_b, \tag{9.30}$$

where the plus sign corresponds to $\phi(0) = 0$ and the minus sign corresponds to $\phi(0) = \pi$. In a similar way, we may show that, in the interval $0 \le t \le 2T_b$, the quadrature component $A_c \sin[\phi(t)]$ consists of a half-sine pulse, whose polarity depends only upon $\phi(T_b)$, as follows:

$$A_c \sin[\phi(t)] = \pm A_c \sin\left(\frac{\pi}{2T_b} t\right), \qquad 0 \le t \le 2T_b, \tag{9.31}$$

where the plus sign corresponds to $\phi(T_b) = \pi/2$ and the minus sign corresponds to $\phi(T_b) = -\pi/2$. Figure 9.10 illustrates the waveforms of the in-phase and quadrature components of $s(t)$ for the input binary sequence 011010, assuming that $\phi(-T_b) = \pi/2$. Note that both components have a bit rate equal to one-half that of the original binary sequence.

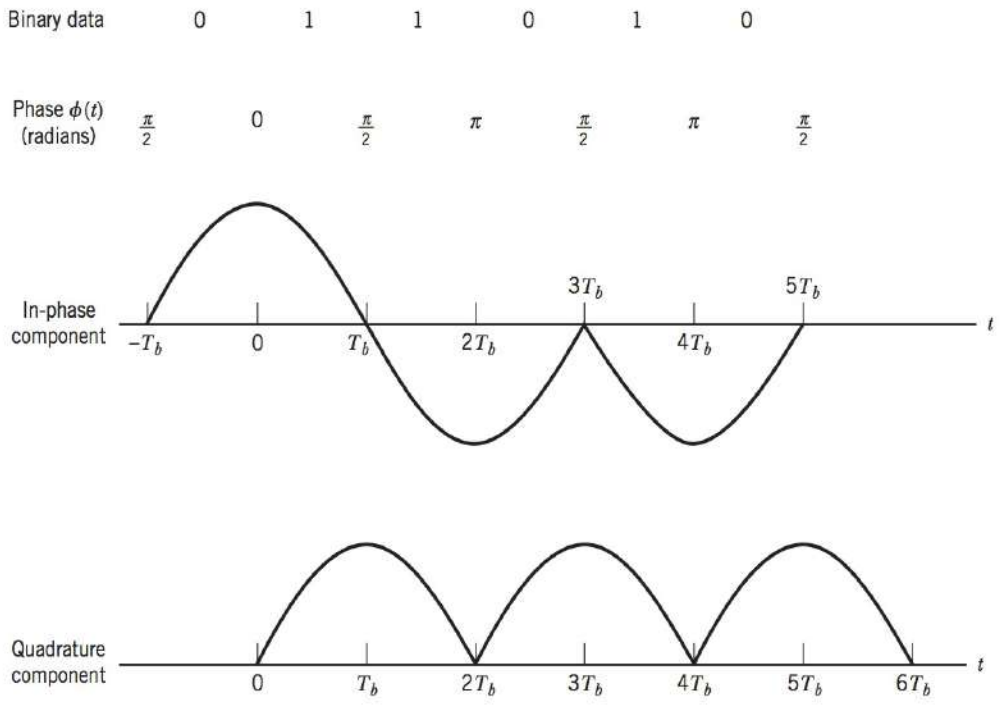


FIGURE 9.10 Waveforms illustrating the in-phase and quarantine components of the CPFSK signal with $h = \frac{1}{2}$.

We assume that the additive noise $w(t)$ at the receiver input is white and Gaussian, with zero-mean and spectral density $N_0/2$. Then, the received MSK signal $x(t)$ may be expressed in the form:

$$x(t) = \pm A_c \cos\left(\frac{\pi}{2T_b}t\right)\cos(2\pi f_c t) \pm A_c \sin\left(\frac{\pi}{2T_b}t\right)\sin(2\pi f_c t) + w(t) \tag{9.32}$$

In the right-hand side of Eq. (9.32), the first algebraic sign is positive if $\phi(0) = 0$ and negative if $\phi(0) = \pi$, whereas the second algebraic sign is positive if $\phi(T_b) = -\pi/2$ and negative if $\phi(T_b) = \pi/2$. For the optimum detection of the phase states $\phi(0)$ and $\phi(T_b)$, we use a pair of matched filters or correlators, as in Figure 9.11. The correlator in the in-phase channel compares $x(t)$ with the coherent reference signal $\cos(\pi t/2T_b)\cos(2\pi f_c t)$ in the interval $-T_b \leq 1 \leq T_b$, resulting in the output

$$l_1 = \int_{-T_b}^{T_b} x(t)\cos\left(\frac{\pi}{2T_b}t\right)\cos(2\pi f_c t)\, dt \tag{9.33}$$

If $l_t > 0$, the receiver chooses: $\phi(0) = 0$; otherwise, it chooses $\phi(0) = \pi$. The correlator in the quadrature channel compares $x(t)$ with the coherent reference signal $\sin(\pi t/2T_b)$ $\sin(2\pi f_c t)$ in the interval $0 \leq t \leq 2T_b$, resulting in the output

$$l_2 = \int_{0}^{2T_b} x(t)\sin\left(\frac{\pi}{2T_b}t\right)\sin(2\pi f_c t)dt \tag{9.34}$$

If $l_2 > 0$, the receiver chooses $\phi(T_b) = -\pi/2$; otherwise, it chooses $\phi(T_b) = \pi/2$. The original binary sequence is reconstructed by appropriately interleaving the phase decisions at the in-phase and quadrature correlator outputs.

From Eq. (9.32), we note that the two values $\phi(0) = 0$ and $\phi(0) = \pi$ are represented by a pair of antipodal signals with an energy equal to $A_c^2 T_b/2$ in a duration $2T_b$, which is the same as the energy E_b of the incoming MSK signal per bit (of duration T_b). We assume that these two values of $\phi(0)$ occur with equal probability. Then, applying Eq. (9.19) with $\rho = -1$; we find that the overall probability of error at the in-phase correlator

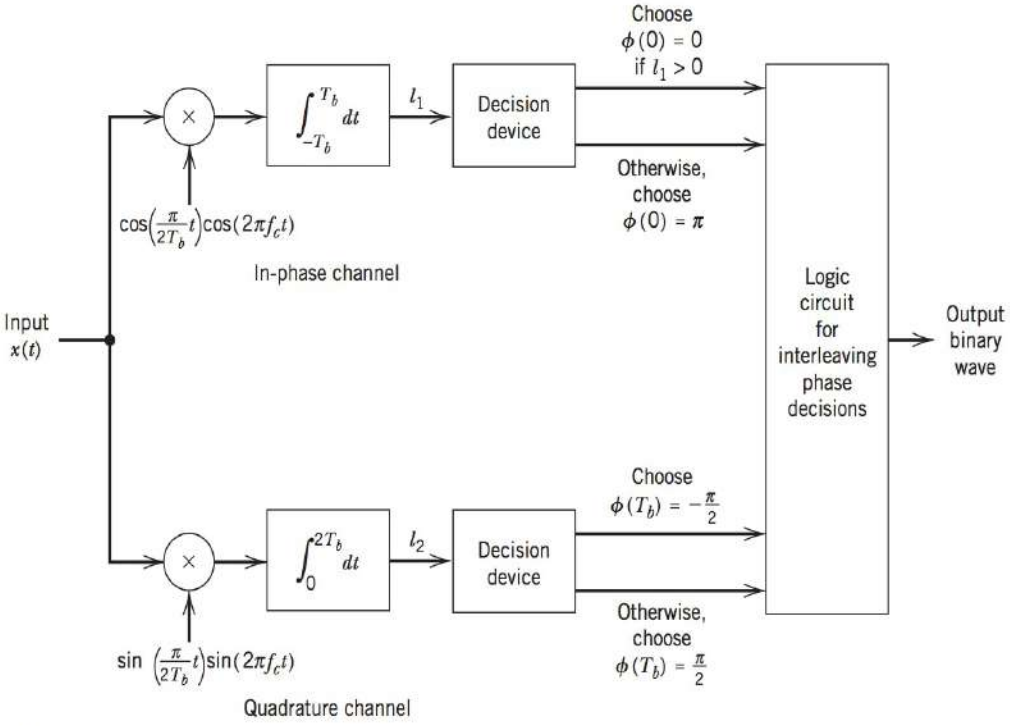


FIGURE 9.11 Coherent receiver for the detection of MSK.

output is given by

$$P_{e1} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{9.35}$$

Similarly, assuming that the two values $\phi(T_b) = -\pi/2$ and $\phi(T_b) = \pi/2$ occur with equal probability, the average probability of error P_{e2} at the quadrature correlator output is the same as P_{e1}.

It may be shown that the random variables L_1 and L_2, whose values are denoted by l_1 and l_2, are uncorrelated. They are also Gaussian because they are derived from the Gaussian noise process $w(t)$ by a linear filtering operation. Accordingly, they are statistically independent. This means that the errors at the two correlator outputs are also independent.

Thus, comparing Eqs. (9.20) and (9.35), we see that the average probability of error in a MSK system is the same as that of a coherent PSK system. It is also of interest to note that this MSK signal occupies much less bandwidth than a conventional FSK signal. In particular, assuming that all the baseband signaling pulses are equally likely, and that symbols transmitted during different time slots are statistically independent and identically distributed, we find that for rectangular signaling, 99 percent of the mean power of a MSK signal is contained in a bandwidth of $1.17/T_b$. The result is that it can transmit data at twice the rate of other FSK signals.

The receiver of Figure 9.11 assumes the availability of the coherent reference signals $\cos(\pi t/2T_b)\cos(2\pi f_c t)$ and $\sin(\pi t/2T_b)\sin(2\pi f_c t)$. This pair of reference signals may be recovered from the received signal in several ways; however, if the modulation scheme is to be successfully applied, we must provide an efficient and accurate method for establishing the reference signals in the receiver that are essentially independent of the modulation. Except for a 180-degree phase ambiguity, this requirement may be achieved by using a *carrier recovery circuit* that consists of a squaring device, a pair of phase-locked loops, a pair of frequency dividers, a summer and subtractor.

NONCOHERENT DETECTION OF FSK SIGNALS

When simplicity of receiver implementation is of prime concern, we completely disregard the phase information in the received signal and use noncoherent detection. This simplicity is achieved, however, at the expense of some degradation in the noise performance of the system. For the noncoherent detection of FSK signals, the receiver consists of a pair of matched filters followed by envelope detectors, as in Figure 9.12. The envelopes

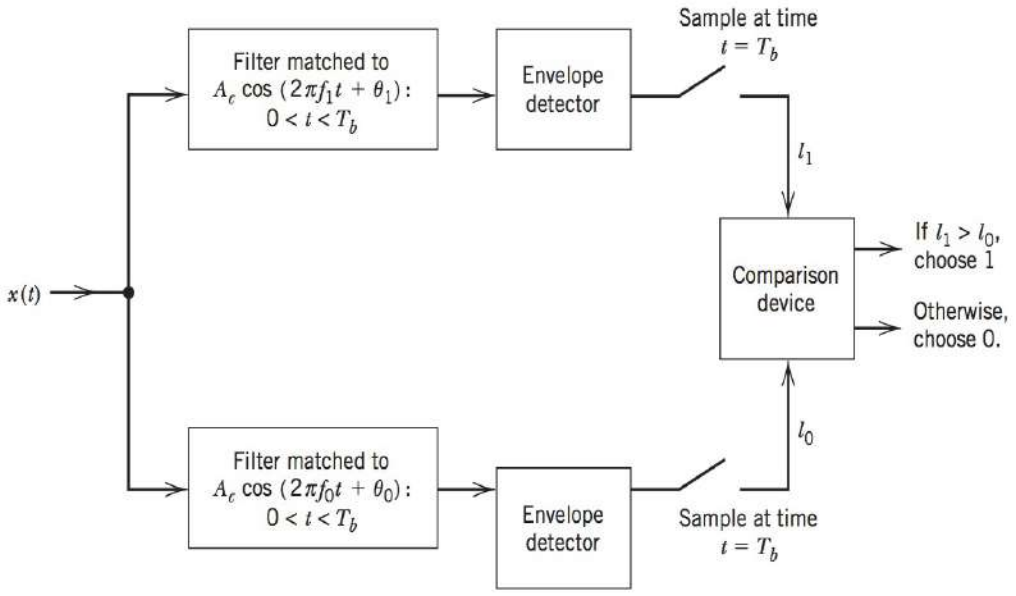


FIGURE 9.12 Noncoherent receiver for the detection of FSK signals.

thus obtained are sampled once every T_b seconds. Let l_0 and l_1 denote the envelope samples of the lower and upper paths of the receiver, respectively. Then, if $l_1 > l_0$, the receiver chooses symbol 1. Otherwise, it chooses symbol 0.

The calculation of the error rate for noncoherent detection of FSK involves application of the Rayleigh and Rician distribution functions;[3] these distributions pertain respectively to the random variables associated with l_0 and l_1. This calculation is addressed in Problem 9.23 but we quote the result that the average probability of error for noncoherent binary FSK is

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right) \tag{9.36}$$

This formula of Eq. (9.36) and noncoherent FSK correspond to a special case of noncoherent orthogonal modulation.

DIFFERENTIAL PHASE-SHIFT KEYING (DPSK)[4]

In the coherent PSK receiver of Figure 9.4*c*, it was assumed that the receiver was perfectly synchronized in frequency and had exact knowledge of the transmitted carrier phase. In practice, however, we frequently find that the receiver does not have exact knowledge of this carrier phase, although it may be able to establish a phase reference to some value of θ radians of the exact phase. Provided that θ remains essentially constant over a period of two-bit intervals, we may resolve this phase ambiguity by using *differential encoding*. As described in Section 7.9, in differential encoding we encode the digital information content of a binary wave in terms of signal transitions. For example, we may use symbol 0 to represent transition in a given binary sequence (with respect to the previous bit) and symbol 1 to represent no transition. A signaling technique that combines differential encoding with phase-shift keying is known as *differential phase-shift keying* (DPSK). Thus, by using DPSK, digital information is encoded, not by the absolute identification of zero carrier phase with symbol 1 and 180 degrees phase with symbol 0 (say), but rather in terms of the phase change between successive pulses in the given binary data stream. For example, symbol 1 is represented by zero phase change from the previous pulse of the binary sequence, whereas symbol 0 is represented by a phase change of 180 degrees, as illustrated in Figure 9.13, where we have arbitrarily chosen zero phase to represent the reference bit.

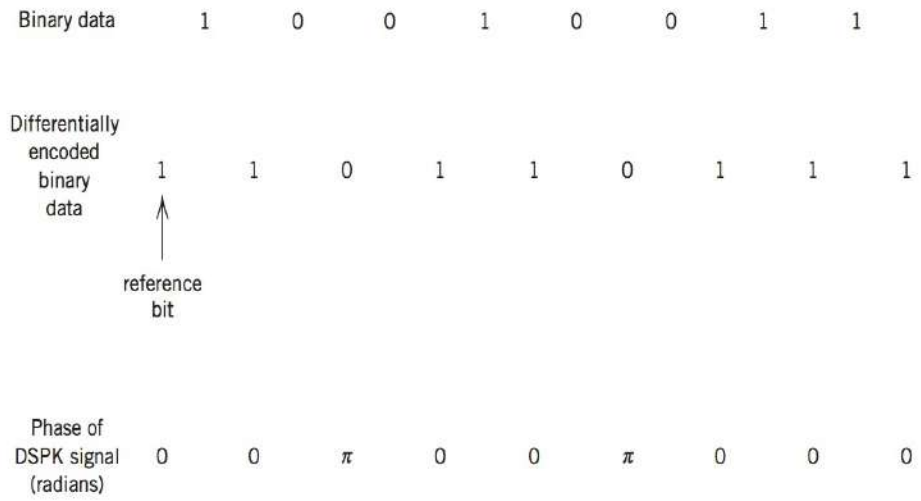


FIGURE 9.13 Illustrating the relationship between a binary sequence, and its differentially encoded and DPSK versions.

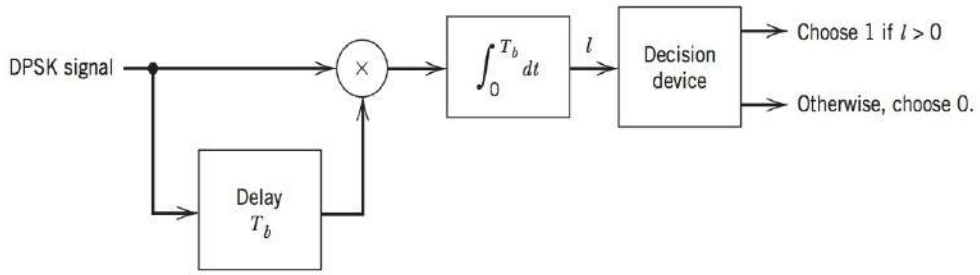


FIGURE 9.14 Receiver for the detection of DPSK signals.

For the differentially coherent detection of a DPSK signal, we may use the receiver shown in Figure 9.14. At any particular instant of time, we have the received DPSK signal as one input into the multiplier and a delayed version of this signal, delayed by the bit duration T_b, as the other input. The integrator output is proportional to cos ϕ, where ϕ is the difference between the carrier phase angles in the received DPSK signal and its delayed version, measured in the same bit interval. Therefore, when $\phi = 0$ (corresponding to symbol 1), the integrator output is positive; on the other hand, when $\phi = \pi$ (corresponding to symbol 0), the integrator output is negative. Thus, by comparing the integrator output with a decision level of zero volts, the receiver of Figure 9.14 can reconstruct the binary sequence, which, in the absence of noise, is exactly the same as the original binary data input.

The major difference between a DPSK system as described above and a coherent PSK system is not in the differential encoding, which can be used in any case, but rather it lies in the way in which the reference signal is derived for the phase detection of the received signal. Specifically, in a DPSK receiver the reference is contaminated by additive noise to the same extent as the information pulse; that is, both have the same signal-to-noise ratio. This makes the determination of the overall probability of error using differentially coherent detection of differentially encoded PSK signals somewhat complicated. Therefore, it will not be given here. The result is, however,

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right) \tag{9.37}$$

It is of interest to note that, since in a DPSK receiver decisions are made on the basis of the signal received in two successive bit intervals, there is a tendency for bit errors to occur in pairs.

9.4 *M*-ARY DATA TRANSMISSION SYSTEMS

In the binary data transmission systems considered in the previous section, we may send only one of two possible signals, $s_0(t)$ or $s_1(t)$, during each bit interval T_b. On the other hand, in an *M-ary data transmission* system we may send any one of M possible signals, $s_0(t)$, $s_1(t)$, $s_2(t), \ldots, s_{M-1}(t)$, during each signaling interval T. For almost all applications, the number of possible signals $M = 2^n$, where n is an integer, and the signaling interval is $T = nT_b$. It is apparent that a binary data transmission system is a special case of an M-ary data transmission system. Each of the M signals, $s_0(t)$, $s_1(t)$, $s_2(t), \ldots, s_M(t)$, is called a *symbol* of the system. The rate of at which these symbols are transmitted through the communication channel is expressed in units of *bauds*.

As we have seen, binary PSK and ASK may be represented at complex baseband as simple two-level line codes. Geometrically, we may illustrate the complex baseband representation as the two signal points shown in Figure 9.15*a* for binary ASK and Figure 9.15*b* for binary PSK. By extension M-ary ASK corresponds to M-ary PAM at complex baseband and could be represented as in Figure 9.15*c*.

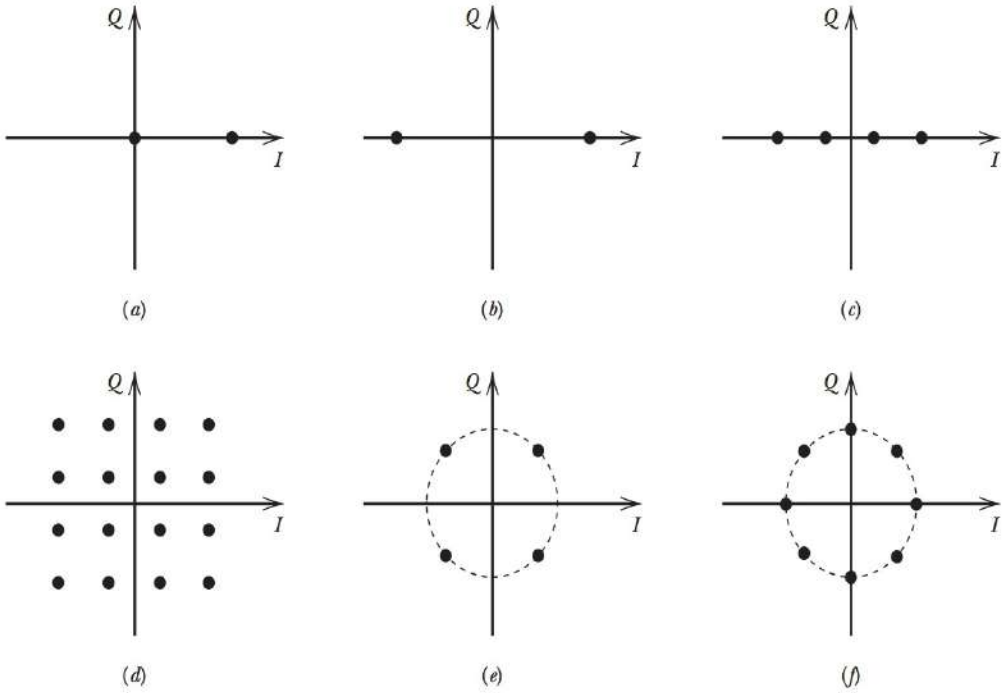


FIGURE 9.15 Representation of different band-pass modulations in signal space: (*a*) binary ASK; (*b*) binary PSK; (*c*) *M*-ASK ($M = 4$); (*d*) *M*-QAM ($M = 16$); (*e*) 4-QAM and QPSK; and (*f*) *M*-PSK ($M = 8$).

Now consider a modulation scheme that uses independent *M*-ary PAM sequences for the in-phase and quadrature components, $g_I(t)$ and $g_Q(t)$. This could be represented by the two-dimensional signal space diagram of Figure 9.15*d*. This type of band-pass modulation is referred to as quadrature amplitude modulation (QAM). Quadrature modulation for analog signals was discussed in Section 3.3. Figure 9.15*d* illustrates the example of 16-ary QAM (16-QAM). Thus, for this modulation scheme there are $M = 16$ different signals $s_i(t)$, one for each point in the figure, and each signal represents four bits.

If we consider the case of 4-QAM illustrated in Figure 9.15*e*, we observe that all of the points have the same amplitude relative to the origin. A constant amplitude implies that only the phase of the carrier changes so this diagram could also be considered quadrature phase shift keying or QPSK. From Figure 9.15*e*, it is a small step to Figure 9.12*f* for *M*-ary PSK. Clearly, *M*-ary PSK, similar to FSK, does not have a simple interpretation in terms of line codes.

This geometric representation of a digital signal as a point in space is referred to as the *signal space* representation. It is analogous to using a vector, i.e., a phasor, to represent a sinusoid in traditional circuit analysis. The set of points that characterize a given modulation technique is known as the *constellation* for that modulation.

There are *M*-ary forms of FSK as well, which are not as simple to represent geometrically. With *M*-FSK, there are *M* possible signals $s_i(t) = A_c \cos(2\pi f_i t)$, where each signal corresponds to a different frequency f_i. Often the frequencies are chosen so that the spacing is a multiple of the symbol rate,

$$\begin{aligned} \Delta f &= f_{i+1} - f_i \\ &= \frac{n}{T_b} \end{aligned} \tag{9.38}$$

This choice has detection advantages that we have already described.

Vladimir Kotelnikov (1908–2005)

Kotelnikov was a Russian pioneer in the fields of information and detection theory. He is also known for having discovered, independently of others (e.g., Whittaker, Nyquist, and Shannon), the sampling theorem in 1933. He was the first to write down a precise statement of this theorem in relation to signal transmission. He was a pioneer in the use of signal theory in modulation and communications, and his work was fundamental to the geometric interpretation of signals and the use of the *signal-space* representation to develop detector structures that are widely used in both radar and communications. This work was illuminated by his treatise on the theory of optimum noise immunity. Kotelnikov also played a leading role in radio astronomy. In 1961, he oversaw one of the first efforts to probe the planet Venus with radar.

QUADRIPHASE-SHIFT KEYING (QPSK)

In a quadriphase-shift keying system, one of four possible signals is transmitted during each signaling interval T, with each signal uniquely related to pairs of bits. For example, we may represent the four possible pairs of bits 10, 00, 01, and 11 as follows:

$$\begin{aligned}
s_0(t) &= \sqrt{2}A_c \cos\left(2\pi f_c t + \frac{\pi}{4}\right), && \text{for bit pair 11} \\
s_1(t) &= \sqrt{2}A_c \cos\left(2\pi f_c t + \frac{3\pi}{4}\right), && \text{for bit pair 01} \\
s_2(t) &= \sqrt{2}A_c \cos\left(2\pi f_c t + \frac{5\pi}{4}\right), && \text{for bit pair 00} \\
s_3(t) &= \sqrt{2}A_c \cos\left(2\pi f_c t + \frac{7\pi}{4}\right), && \text{for bit pair 10}
\end{aligned} \tag{9.39}$$

where $0 \le t \le T$. That is, the carrier is transmitted with one of four possible phase values, $\pm\pi/4$, $\pm 3\pi/4$, with each phase corresponding to a unique pair of bits, as illustrated in Figure 9.16.

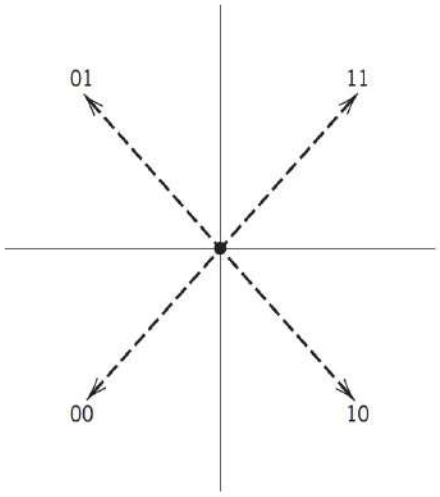


FIGURE 9.16 Illustrating four possible phase values, with each one corresponding to a unique pair of bits.

Figure 9.17 shows the block diagram of a QPSK transmitter, which consists of a *serial to parallel converter*, a pair of *product modulators*, an oscillator and phase shifter to generate the two carrier waves in phase quadrature, and *a summer*. The function of the serial to parallel converter is to represent each successive pair of bits of the incoming binary data stream $m(t)$ in parallel form. Typical signal waveforms are shown in Figure 9.18. It is apparent that the signaling interval T in a QPSK system is twice as long as the bit duration T_b of the input binary data stream $m(t)$. That is, for a given bit rate $1/T_b$, a QPSK system requires half the transmission bandwidth of the corresponding binary PSK system. Equivalently, for a given transmission bandwidth, a QPSK system carries twice as many bits of information as the corresponding binary PSK system. The QPSK receiver consists of two binary detectors or correlators connected in parallel as in Figure 9.19. One correlator computes the cosine of the carrier phase, whereas the other correlator computes the sine of the carrier phase. By computing the signs of the two correlator outputs, a unique resolution of one of the transmitted phase angles is made. We may thus view a QPSK system as two binary PSK systems operating in parallel, with the two carrier waves in phase quadrature. This is another example of *quadrature multiplexing*, considered in Chapter 3.

We assume that the additive noise at the receiver input is white and Gaussian, with zero-mean and spectral density $N_0/2$. We may express the received signal as

$$x(t) = \pm A_c \cos(2\pi f_c t) \pm A_c \sin(2\pi f_c t) + w(t), \quad 0 \le t \le T \tag{9.40}$$

depending on which particular pair of bits is transmitted. Therefore, at the end of a signaling interval T, we find that the output of the correlator in the in-phase channel is

$$l_1 = \pm\frac{1}{2}A_c T + \int_0^T w(t)\cos(2\pi f_c t)\,dt, \tag{9.41}$$

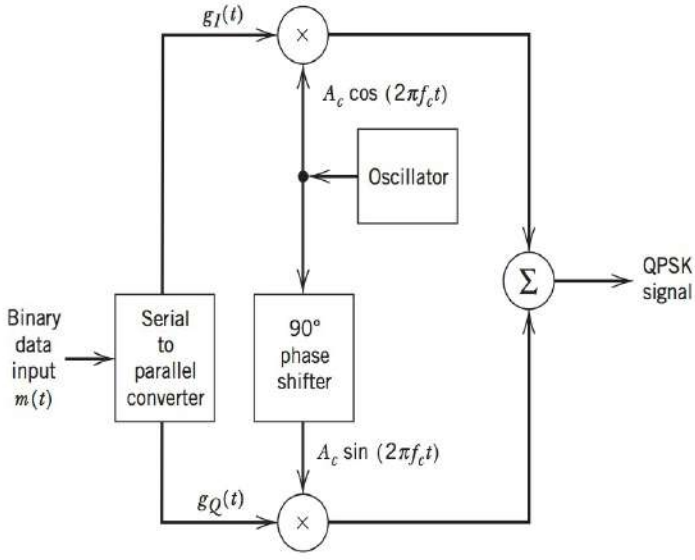


FIGURE 9.17 QPSK transmitter.

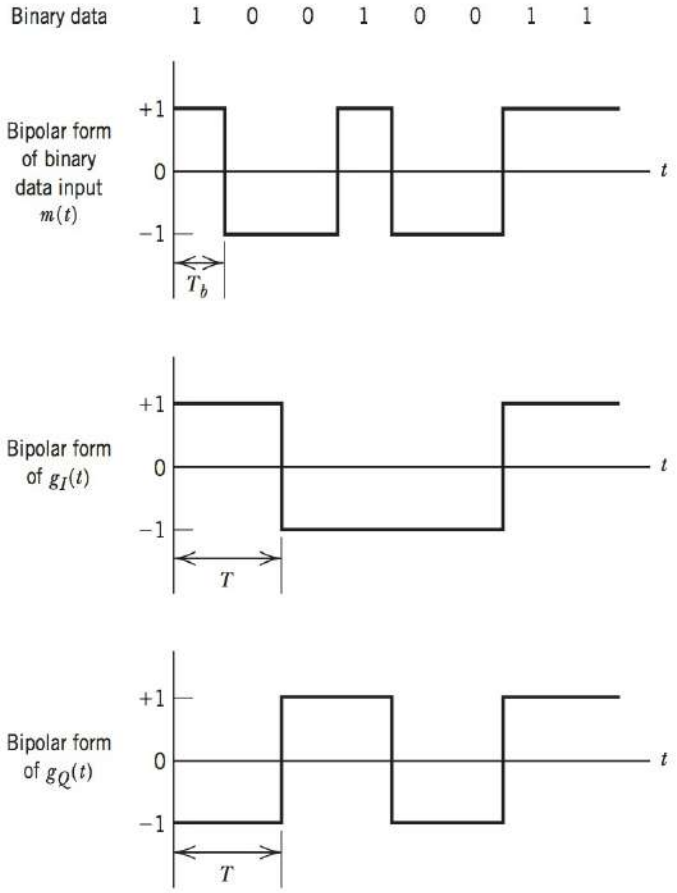


FIGURE 9.18 Illustrating the serial-to-parallel conversion process.

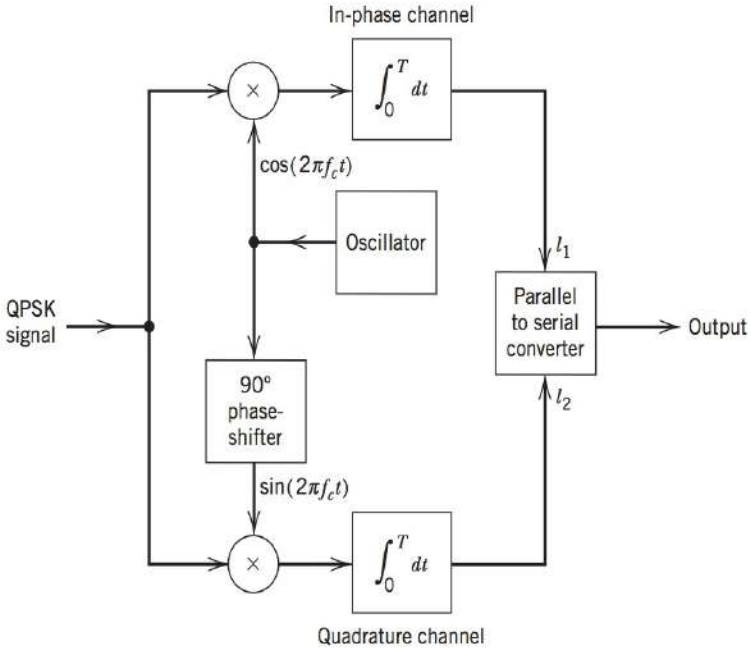


FIGURE 9.19 QPSK receiver.

whereas the output of the correlator in the quadrature channel is

$$l_2 = \pm\frac{1}{2}A_c T + \int_0^T w(t)\sin(2\pi f_c t)\,dt \tag{9.42}$$

The random variables L_1 and L_2, whose values are denoted by l_1 and l_2, are uncorrelated. They are also Gaussian because they are derived from the Gaussian process $w(t)$ by a linear filtering operation. Accordingly, they are statistically independent.

The mean of L_1 is given by the expected value

$$\mathbf{E}[L_1] = \pm\frac{A_c T}{2}, \tag{9.43}$$

depending on whether at the input of the upper product modulator in the transmitter of Figure 9.14 we have a binary symbol 1 or 0. The variance of L_1 is given by

$$\begin{aligned}\mathrm{Var}[L_1] &= \mathbf{E}\left[\left(\int_0^T w(t)\cos(2\pi f_c t)\,dt\right)^2\right] \\ &= \mathbf{E}\left[\int_0^T\int_0^T w(t)w(u)\cos(2\pi f_c t)\cos(2\pi f_c u)\,dt\,du\right] \\ &= \int_0^T\int_0^T \frac{N_0}{2}\delta(t-u)\cos(2\pi f_c t)\cos(2\pi f_c u)\,dt\,du \\ &= \frac{N_0}{2}\int_0^T \cos^2(2\pi f_c t)\,dt \\ &= \frac{N_0 T}{4}\end{aligned} \tag{9.44}$$

Similarly, for L_2 we have

$$\mathbf{E}[L_2] = \pm\frac{A_c T}{2} \tag{9.45}$$

$$\mathrm{Var}[L_2] = \frac{N_0 T}{4} \tag{9.46}$$

Let P_{ei} denote the probability of error at the output of the ith correlator in the receiver of Figure 9.19 where $i = 1$ corresponds to the upper correlator and $i = 2$ corresponds to the lower one. We then find that

$$P_{e1} = P_{e2} = Q\left(\sqrt{\frac{A_c^2 T}{N_0}}\right) \tag{9.47}$$

using an analysis similar to Section 8.3.

We note from Eq. (9.39) that the signal energy per symbol is

$$E = A_c^2 T \tag{9.48}$$

We may therefore rewrite Eq. (7.124) as

$$P_{e1} = P_{e2} = Q\left(\sqrt{\frac{E}{N_0}}\right) \tag{9.49}$$

In a QPSK system, there are two bits per symbol, so that the signal energy per symbol is twice the signal energy per bit, that is,

$$E = 2E_b \tag{9.50}$$

Thus, expressing the average probability of bit error in terms of the ratio E_b/N_0, we may write

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{9.51}$$

For a satisfactory operation of the receiver in Figure 9.19, we need an efficient carrier recovery circuit capable of tracking the carrier phase without concern for which of the data signals phase modulates the carrier wave. A carrier recovery circuit that satisfies this requirement, except for a phase ambiguity, is the *four-phase Costas loop*, which is an extension of the conventional Costas loop that was considered in Chapter 2. Alternatively, we may use the *fourth-power loop*, which involves raising the received signal to the fourth power, and then using a phase-locked loop to track the fourth harmonic of the carrier so produced.

9.5 COMPARISON OF NOISE PERFORMANCES OF VARIOUS PSK AND FSK SYSTEMS

Throughout this chapter, we have used the overall probability of committing a bit error as the figure of merit for evaluating the noise performance of a digital communication system. It should be realized, however, that even if two systems yield the same symbol error probability, their performances, from the users' viewpoint, may be quite different. In particular, the greater the number of bits per symbol, the more the bit errors will cluster together. For example, if the symbol error probability is 10^{-3}, the expected number of symbols occurring between any two erroneous symbols is 1000. If each symbol represents one bit of information (as in a binary PSK or binary FSK system), the expected number of bits separating two erroneous bits is 1000. If, on the other hand, there are 2 bits per symbol (as in a QPSK system), the expected separation is 2000 bits. Of course, a symbol error generally creates more bit errors in the second case. Nevertheless, this clustering effect may make one system more attractive than another, even at the same symbol error rate. In the final analysis, which system is preferable will depend upon the particular situation.

Two systems having an unequal number of symbols may be compared in a meaningful way only if they use the same amount of energy to transmit each bit of information. It is the total amount of energy needed to transmit the complete message that represents the cost of the transmission, not the amount of energy needed to transmit satisfactorily a particular symbol. Accordingly, in comparing the different data transmission systems considered above, we will use, as the basis of our comparison, the probability of bit error expressed as a function of the signal energy per bit-to-mean noise power per unit bandwidth ratio: that is, E_b/N_0.

In Table 9.1, we have summarized the expressions for the bit error probability P_e for the coherent PSK, conventional coherent FSK with 1-bit decoding, coherent MSK, noncoherent FSK, DPSK, and QPSK. In Figure 9.20 we have used these expressions to

TABLE 9.1 Summary of formulas for the bit error probability P_e for different data transmission systems

	P_e
Coherent PSK	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$
Coherent FSK (with 1-bit decoding)	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$
MSK	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$
QPSK	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$
Noncoherent FSK	$\frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$
DPSK	$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$

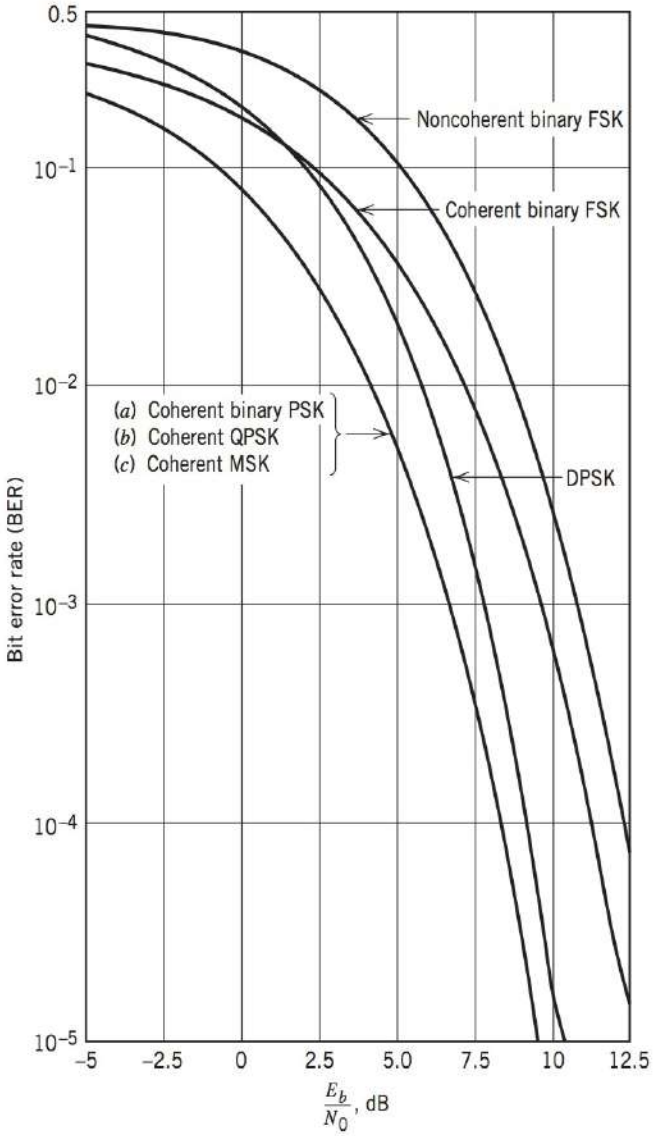


FIGURE 9.20 Comparison of the noise performances of different PSK and FSK systems.

plot P_e as a function of E_b/N_0. In practice, the error probability is typically of the order of 10^{-5}. On the basis of the curves in Figure 9.20, we may state the following:

1. The error rates for all the systems decrease montonically with increasing values of E_b/N_0.
2. For any value of E_b/N_0, coherent BPSK, QPSK, and MSK produce a smaller error rate than the other systems.
3. The coherent PSK and the DPSK require an E_b/N_0 that is 3 dB less than the corresponding values for the conventional coherent FSK and the noncoherent FSK, respectively, to realize the same error rate.
4. At high values of E_b/N_0, the DPSK and the noncoherent FSK perform almost as well (to within about 1 dB) as the coherent PSK and the conventional coherent FSK, respectively, for the same bit rate and signal energy per bit.
5. The QPSK system transmits, in a given bandwidth, twice as many bits of information as a conventional coherent BPSK system with the same error rate performance. Here again we find that a QPSK system requires a more sophisticated carrier recovery circuit than a BPSK system.

From Figure 9.20, we also see that at high values of E_b/N_0, we have approximately a 4-dB difference between the best signaling methods and the worst signaling method. It may appear that this represents a small improvement in signal-to-noise ratio in return for increased receiver complexity in going from the noncoherent FSK to the coherent PSK. However, in some applications where power is at a premium (e.g., as in digital satellite communications) even a 1-dB saving in signal-in-noise ratio is well worth the effort.

9.6 THEME EXAMPLE—ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING (OFDM)

One of the assumptions made at the beginning of this chapter was that the band-pass channel is linear and passes the signal undistorted. Indeed, the analysis made throughout the chapter is based on this assumption. In practice, the validity of the assumption depends upon the application of interest and becomes less true as the signal bandwidth increases. An example where this assumption breaks down is the wireless local area networks (WLAN) that provide so-called WiFi service.

The assumption that the channel passes the signal undistorted implies that the amplitude response of the channel is flat in the frequency domain. In Figure 9.21 we provide an example of the amplitude response of a typical WLAN channel. The WiFi networks are designed to carry high data rates, up to 54 megabits per second and more. As a consequence, the corresponding signals have been designed to occupy approximately 20 MHz of bandwidth. Over a bandwidth of 20 MHz, the amplitude response shown in Figure 9.21 is clearly not constant. However, we can also make the observation that over a small bandwidth, say 300 kHz, the amplitude response is approximately constant.

This last observation suggests an approach known as *multicarrier modulation*. With this modulation technique, a number of carriers are transmitted in a synchronous fashion. This is illustrated in Figure 9.22*a* for the case of unmodulated carriers, and in Figure 9.22*b* for the case of modulated carriers. For the sake of clarity, we will refer to these individual carriers as *subcarriers* with frequencies f_i, and reserve the term carrier for the center frequency of the aggregate signal, f_c, when it is modulated to a bandpass signal. If the bandwidth of one of the modulated subcarriers is on the order of 300 kHz or less, then its behavior will be similar to that analyzed in this chapter.

For illustration purposes, let us assume that there are 48 subcarriers. The modulated bandwidth of each of these subcarriers is set at 312.5 kHz. Let us denote the subcarrier frequencies as $f_0, f_1, \ldots, f_{47}$ and consider all of the subcarriers in their complex form

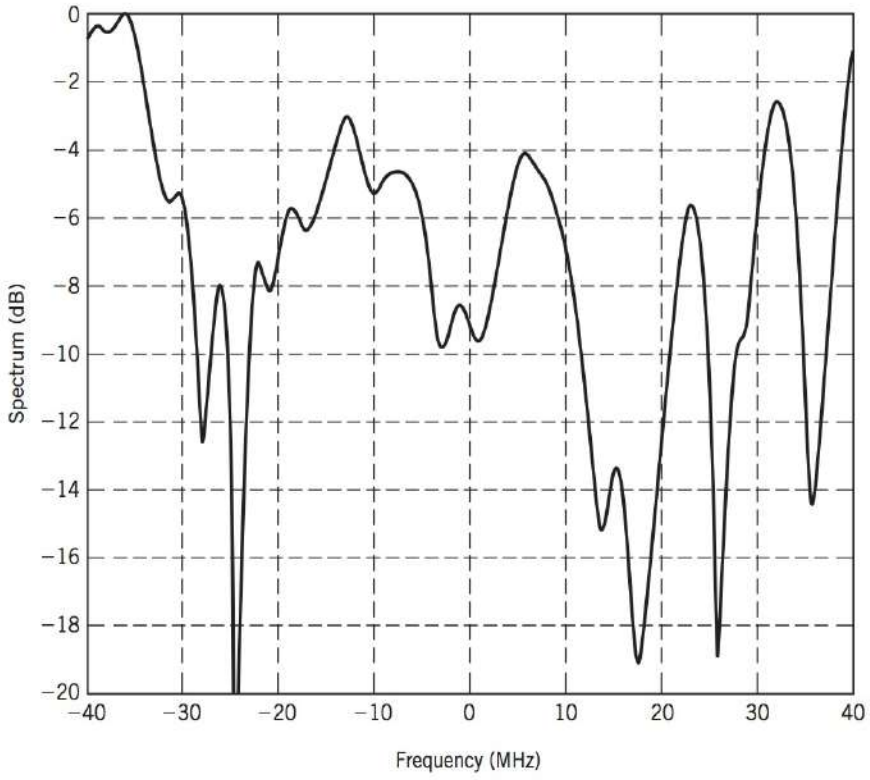


FIGURE 9.21 Example amplitude spectrum of wireless LAN channel.

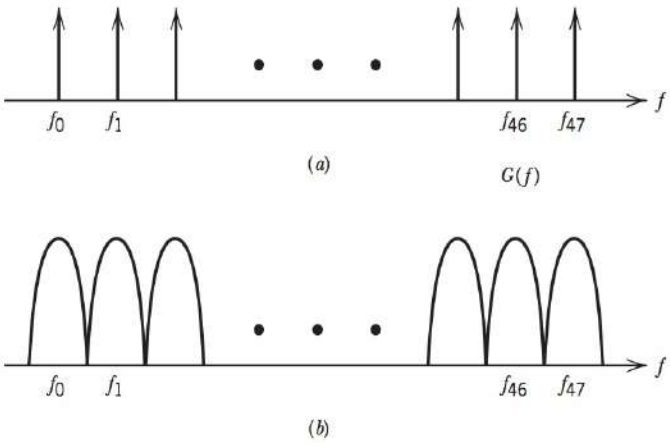


FIGURE 9.22 Conceptual illustration of complex baseband subcarriers: (*a*) unmodulated, and (*b*) modulated.

$$\tilde{c}_n(t) = \exp(j2\pi f_n t) \qquad n = 0, 2, \ldots, 47 \tag{9.52}$$

In general, the complex envelope of the nth subcarrier may be represented as by

$$\tilde{g}_n(t) = b_{k,n}p(t - kT), \qquad (k-1)T \le t < kT \tag{9.53}$$

where $p(t)$ is a rectangular pulse and T is the symbol period. The complex coefficient $b_{k,n}$ is selected in accordance with the selected constellation. For example, for each subcarrier, the modulation could be BPSK, QPSK, or M-ary QAM. If the selected modulation is 16-QAM, then the $b_{k,n}$ are selected from the 16 elements of the set

$$\{\pm 1 \pm j,\ \pm 3 \pm j,\ \pm 1 \pm 3j,\ \pm 3 \pm 3j\};$$

At each symbol time kT, a different symbol is selected to be transmitted.

In practice, the incoming data stream $\{d_l\}$ is demultiplexed into 48 parallel streams represented by $\{b_{k,n}\}_{n=1}^{48}$, running at 1/48 of the incoming data rate. Then the combination of the data modulation and the subcarrier modulation may be represented in the complex envelope form as follows:

$$\left.\begin{aligned} &\vdots \\ &\tilde{s}_n(t) = b_{k,n}p(t - kT)\exp(j2\pi f_n t) \\ &\tilde{s}_{n+1}(t) = b_{k,n+1}p(t - kT)\exp(j2\pi f_{n+1} t) \\ &\vdots \end{aligned}\right\} \quad \text{for } (k-1)T \le t < kT,\ n = 0, 1, \ldots, 47 \tag{9.54}$$

Each term of Eq. (9.54) contributes to the band-pass signal, and the aggregate is the overall signal to be transmitted over the wireless channel. The complex envelope representation of the combination of all 48 subcarriers for one symbol period T is given by

$$\tilde{s}(t) = \sum_{n=0}^{47} \tilde{s}_n(t) \tag{9.55}$$

This process is illustrated in Figure 9.23. The incoming data stream is 16-QAM modulated and demultiplexed into 48 separate data streams. After modulating the individual subcarriers, the individual data streams are once again combined.

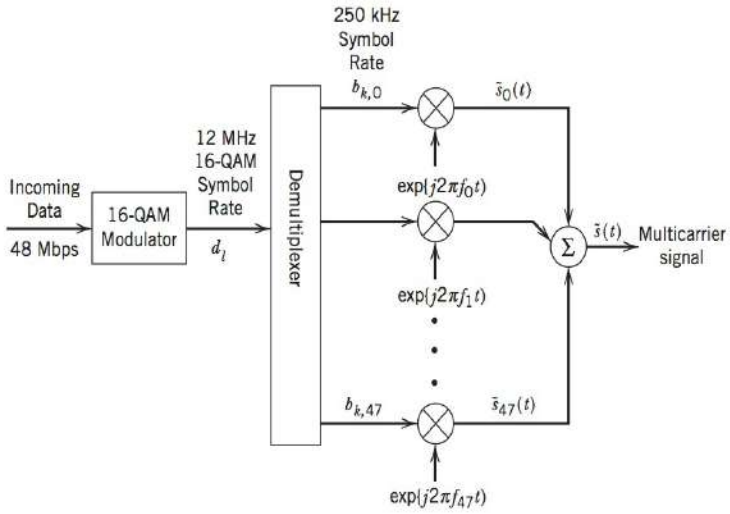


FIGURE 9.23 Illustration of conceptual OFDM modulation process.

The aggregate complex baseband signal is then converted to a band-pass signal using the conventional approach (see Figure 9.1*a*), and mathematically represented as

$$s(t) = \mathrm{Re}[\tilde{s}(t)\exp(j2\pi f_c t)] \tag{9.56}$$

where $\tilde{s}(t)$ is the complex envelope. Note that with this multicarrier modulation scheme, the number of bits transmitted per symbol period, T, is the number of subcarriers times the number of bits per modulated symbol. With 16-QAM used in this example, the number of bits per modulated symbol is $48 \times 4 = 192$ bits.

At first sight, the combination of Eqs. (9.54) and (9.55) appears to be a complicated modulation strategy to implement. However, recall the discrete Fourier transform (DFT), which was discussed in Chapter 2. The DFT transforms a set of samples in the time domain into an equivalent set of samples in the frequency domain. The inverse discrete Fourier transform (IDFT) performs the reverse operation. This transform-pair is described mathematically by

$$\left.\begin{aligned} &\text{DFT:} && b_n = \sum_{m=0}^{M-1} B_m \exp(-j2\pi mn/M) && n = 0, 1, \ldots, M-1 \\ &\text{IDFT:} && B_m = \frac{1}{M}\sum_{n=0}^{M-1} b_n \exp(j2\pi mn/M) && m = 0, 1, \ldots, M-1 \end{aligned}\right\} \tag{9.57}$$

where the sequences $\{b_n\}$ and $\{B_m\}$ are the frequency-domain and time-domain samples, respectively. Considering Eq. (9.54) in the context of the IDFT, let us make the following assumptions:

1. The pulse shape $p(t)$ is rectangular

$$p(t) = \begin{cases} 1, & 0 \le t < T \\ 0, & \text{elsewhere} \end{cases}$$

2. The subcarrier frequencies are selected such that

$$f_n = \frac{n}{T}, \qquad \text{for } n = 0, 1, 2, \ldots, 47$$

3. Each subcarrier and the output is sampled at M times per symbol interval, that is, for $0 \le t < T$ the samples are given by

$$t = \frac{m}{M}T \qquad \text{for } m = 0, 1, \ldots, M-1$$

Combining these three assumptions with Eqs. (9.54) and (9.55), we obtain for the kth symbol period, the following M samples of modulated waveform

$$\tilde{s}\left(\frac{mT}{M}\right) = \sum_{n=0}^{47} b_{k,n} \exp(j2\pi mn/M), \qquad m = 0, 1, \ldots, M-1 \tag{9.58}$$

To provide the symmetry of Eq. (9.57), we should choose $M = 48$. However, the DFT may be efficiently implemented as a fast Fourier transform (FFT) if M is a power of two. In a practical system, the number of subcarriers may be increased to 64 to provide: (a) the 48 data subcarriers mentioned above; (b) a number of additional subcarriers used for synchronization purposes at the receiver; and (c) a number of zero carriers to provide a guard band to protect against adjacent channel interference.

To summarize, the samples of the complex envelope of the multicarrier signal are given by the IDFT of the subcarriers. The typical implementation of the transmitter is shown in Figure 9.24*a*. First, the incoming binary data stream is forward error-correction encoded (as will be discussed in Chapter 10), followed by 16-QAM modulation. Then the data stream undergoes a serial-to-parallel conversion to create 48 independent data

streams. Next, these independent data streams are combined using the inverse fast Fourier transform (IFFT) algorithm. The output of the IFFT algorithm consists of the time-domain samples to be transmitted over the channel. Besides displaying the 48 data-bearing subcarriers, Figure 9.24*a* shows additional subcarriers, used by the receiver for synchronization, tracking, and guard band purposes.

The output of the IFFT algorithm consists of 64 time samples of a complex envelope for each period T, which are parallel-to-serial converted and then finally digital-to-analog converted to facilitate transmission of the multicarrier signal over the wireless channel. Figure 9.24*b* presents the corresponding implementation of the receiver, which follows a sequence of operations in the reverse order to those performed in the transmitter of Figure 9.24*a*. Specifically, to recover the original input binary data stream, the received signal is passed through the following processors:

- Analog-to-digital converter
- Serial-to-parallel converter
- 64-point FFT algorithm
- Parallel-to-serial converter
- 16-QAM demodulator
- Forward error-correction decoder

The modulation strategy described in this theme example certainly has a frequency division multiplex aspect, as can be seen from Figure 9.20. The fact that the individual subcarriers are orthogonal, we leave as an exercise for the reader. Putting these two statements together justifies referring to the communication system of Figure 9.24 as an *orthogonal frequency-division multiplexing (OFDM) system*.

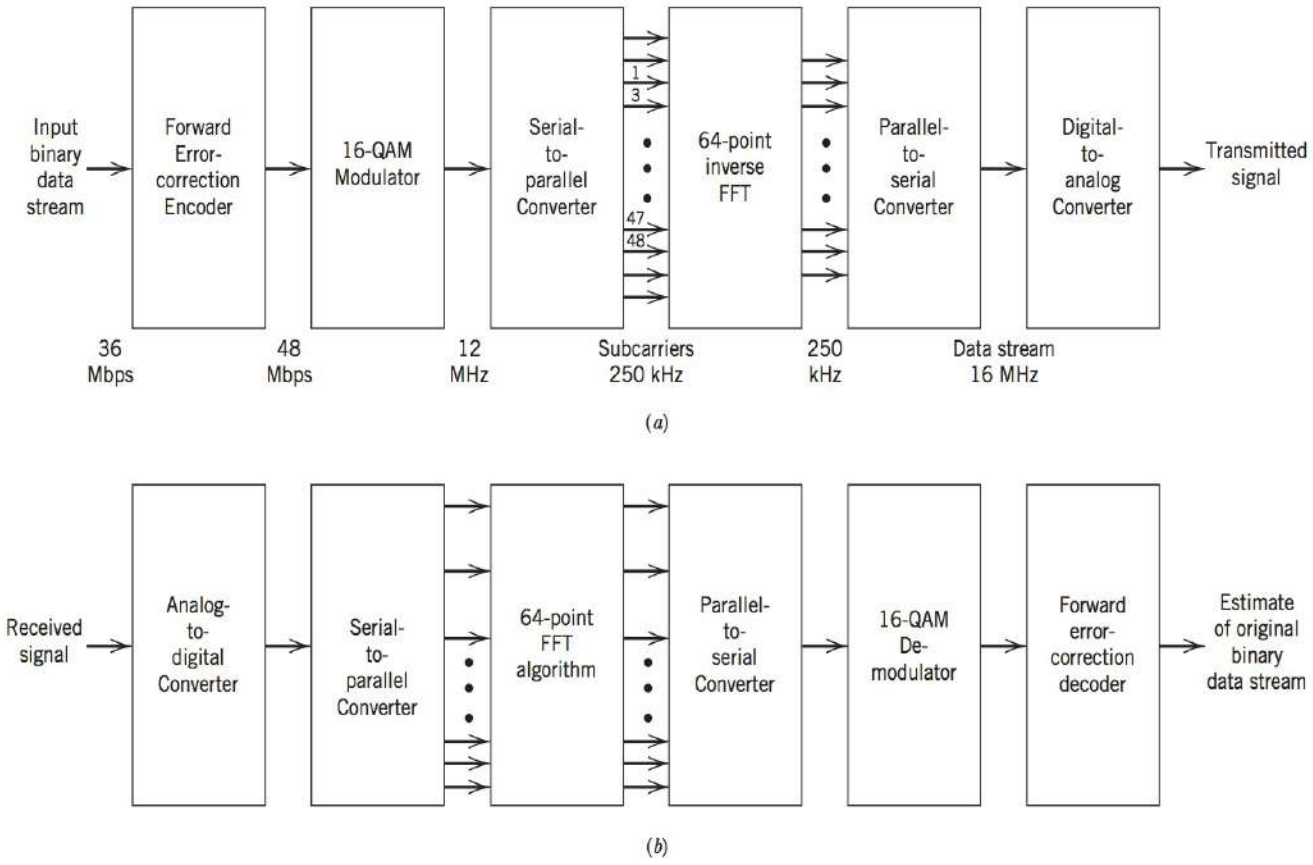


FIGURE 9.24 Block diagram of (*a*) OFDM transmitter, and (*b*) OFDM receiver.

OFDM is an example of a multilayer modulation strategy. On one level, the subcarriers form an orthonormal basis for a signal space. On a second level, each subcarrier has its own signal space that is 16-QAM modulated. Multilayer modulation systems of this sort are quite common. One layer is designed to achieve the throughput requirements. In this case, 16-QAM is chosen to provide the desired throughput. A second layer is designed to take advantage of, or compensate for, properties of the transmission medium. In this case, to compensate for the wireless channel characteristics. OFDM modulation has been selected for use in many wireless LAN standards including IEEE 802.11a, g, and n; as well as digital audio broadcast standards. Some of the values taken for this example come from IEEE 802.11a and g standards.

To summarize, the theme example on OFDM has demonstrated the following desirable features:

1. Spectrally efficient digital modulation schemes, such as 16-QAM, can be represented simply by their complex low-pass equivalents.
2. Modulations of even greater complexity, such as OFDM, can be presented and understood easily in terms of their complex low-pass equivalents.
3. The clear understanding of these modulation schemes allows us to take advantage of digital signal-processing techniques such as the fast Fourier transform algorithm so as to simplify the practical implementation of some complicated modulation schemes.

9.7 SUMMARY AND DISCUSSION

In this chapter, we presented a systematic analysis of the effects of noise on the performance of band-pass data transmission systems. The analysis proceeded by first reviewing the complex envelope representation of band-pass signals. An important result emerging from Chapter 8 on baseband pulse transmission was the idea of a correlation receiver or, equivalently, a matched filter receiver for the optimum detection of a known signal in an AWGN channel. This result extends directly to band-pass systems where there is a greater variety of modulation schemes. In particular, we showed the application of this basic principle to the analysis of the bit error rate performance for some important digital modulation techniques in an AWGN channel:

1. Coherent modulation techniques:
   - Coherent binary phase-shift keying (BPSK)
   - Coherent binary frequency-shift keying (BFSK)
   - Coherent minimum shift keying (MSK)
   - Coherent quadriphase-shift keying (QPSK)
2. Noncoherent binary modulation techniques:
   - Noncoherent binary frequency-shift keying
   - Differential phase-shift keying (DPSK)

This presentation was followed by a brief discussion of coherent M-ary modulation techniques: M-ary phase-shift keying, M-ary quadrature amplitude modulation, and M-ary frequency-shift keying.

From the discussion presented in this chapter, we conclude that the performance analysis of band-pass data transmission systems in the presence of additive white Gaussian noise (AWGN) is well understood for both coherent and noncoherent receivers. In general, we find that the bit error rate decreases exponentially as the signal-to-noise ratio,

E_b/N_0, increases in an AWGN channel. Coherent techniques offer performance advantages ranging from one to three decibels over their noncoherent counterparts, but at the expense of increased complexity required at the receiver to recover synchronization information contained in the received signal.

A final comment is in order: When explicit performance analysis of a band-pass transmission system defies a satisfactory solution, for example, when nonideal effects such as intersymbol interference or adjacent channel interference are present, the use of computer simulation provides the only alternative approach to actual hardware evaluation. The simulation procedure involves the formulation of a complex baseband equivalent model for the system, along the lines described Chapter 2.

NOTES AND REFERENCES

1. For a detailed tutorial review of different digital modulation techniques (ASK, FSK, and PSK) using a geometric viewpoint, see Arthurs and Dym, 1962. See also the following list of books:

   Proakis (2001, Chapter 5),
   Sklar (2001, Chapter 4),
   Gibson (1989, Chapter 11), and
   Viterbi and Omura (1979, pp. 47–127).

2. The MSK signal was first described in Doelz and Heald (1961). For a tutorial review of MSK and comparison with QPSK, see Pasupathy (1979). Since the frequency spacing is only half as much as the conventional spacing of $1/T_b$ that is used in the coherent detection of binary FSK signals, this signaling scheme is also referred to as fast FSK. See deBuda (1972).
3. The standard method of deriving the bit error rate for noncoherent binary FSK, presented in Whalen (1971), and that for differential phase-shift keying presented in Arthurs and Dym (1962), involves the use of the Rician distribution. This distribution was discussed in Chapter 5.
4. The optimum receiver for differential phase-shift keying is discussed in Simon and Divsalar (1992).

PROBLEMS

9.1 The sequence 101101011 is used to modulate a band-pass carrier. Sketch the waveforms for the three cases of binary ASK, FSK, and PSK modulation.

9.2 A band-pass carrier $\cos(2\pi f_c t)$ is modulated using a linear mixer (multiplier) by a real-valued digital signal $g(t)$. If the signal $g(t)$ has baseband spectrum $G(f)$, what is the spectrum of the modulated spectrum? Sketch this spectrum for the case where $g(t)$ is a bipolar NRZ line code.

9.3 A pair of orthogonal carriers $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$ are linearly modulated by the digital baseband signals $g_I(t)$ and $g_Q(t)$ and then combined.

(a) Develop an expression for the spectrum of the band-pass signal if $G_I(f)$ and $G_Q(f)$ are the corresponding baseband spectra.

(b) Suppose $g_I(t)$ and $g_Q(t)$ are independent bipolar NRZ line codes. Sketch the corresponding band-pass spectrum.

(c) Suppose $g_I(t) = -g_Q(t)$. How does this affect the band-pass spectrum?

(d) Suppose $g_I(t)$ and $g_Q(t)$ correspond to independent sequences of pulses, each pulse having the raised cosine pulse shape with rolloff factor 1.0. Sketch the corresponding band-pass spectrum.

9.4 In the on-off version of an ASK system, symbol 1 is represented by transmitting a sinusoidal carrier of amplitude $\sqrt{2E_b/T_b}$ where E_b is the signal energy per bit and T_b is the bit duration. Symbol 0 is represented by switching off the carrier. Assume that symbols 1 and 0 occur with equal probability. For an AWGN channel:

(a) Provide a block diagram for a coherent receiver for this ASK signal.

(b) Determine the average probability of error for this ASK system with coherent reception.

(c) Suppose symbol 1 occurs with probability $\frac{2}{3}$ and symbol 0 occurs with probability $\frac{1}{3}$. How would the receiver design and probability of error change if the objective is to minimize the overall probability of error?

9.5 A PSK signal is applied to a correlator supplied with a phase reference that lies within ϕ radians of the exact carrier phase. Determine the effect of the phase error ϕ on the average probability of error of the receiver.

9.6 The signal component of a coherent PSK system is defined by

$$s(t) = A_c k \sin(2\pi f_c t) \pm A_c\sqrt{1-k^2}\cos(2\pi f_c t)$$

where $0 \le t \le T_b$, and the plus sign corresponds to symbol 1 and the minus sign corresponds to symbol 0. The first term represents a carrier component included for the purpose of synchronizing the receiver to the transmitter.

(a) Draw the spectrum for the scheme described here. What observations can you make about this spectrum?

(b) Show that, in the presence of additive white Gaussian noise of zero mean and power spectral density $N_0/2$, the average probability of error is

$$P_e = \frac{1}{2}Q\left(\sqrt{\frac{2E_b}{N_0}(1-k^2)}\right)$$

where

$$E_b = \frac{1}{2}A_c^2 T_b$$

(c) Suppose that 10 percent of the transmitted signal power is allocated to the carrier component. Determine the E_b/N_0 required to realize a probability of error equal to 10^{-4}.

(d) Compare this value of E_b/N_0 with that required for a conventional PSK system with the same probability of error.

9.7 An FSK system transmits binary data at the rate of 2.5×10^6 bits per second. During the course of transmission, white Gaussian noise of zero mean and power spectral density 10^{-20} watts per hertz is added to the signal. In the absence of noise, the amplitude of the received sinusoidal wave for digit 1 or 0 is 1 microvolt. Determine the average probability of symbol error for the following system configurations:

(a) Coherent binary FSK.

(b) Coherent MSK.

(c) Noncoherent binary FSK.

9.8

(a) In a coherent FSK receiver, the signals $s_1(t)$ and $s_0(t)$ representing symbols 1 and 0, respectively, are defined by

$$s_1(t),\ s_0(t) = A_c \cos\left[2\pi\left(f_c \pm \frac{\Delta f}{2}\right)t\right], \qquad 0 \le t \le T_b$$

Assuming that $f_c > \Delta f$, show that the correlation coefficient of the signals $s_1(t)$ and $s_0(t)$ is approximately given by

$$\rho = \frac{\int_0^{T_b} s_1(t)s_0(t)\,dt}{\int_0^{T_b} s_1^2(t)\,dt} \simeq \operatorname{sinc}(2\,\Delta f T_b)$$

(b) What is the minimum value of frequency shift Δf for which the signals $s_1(t)$ and $s_0(t)$ are orthogonal?

(c) What is the value of Δf that minimizes the average probability of symbol error?

(d) For the value of Δf obtained in part (c), determine the increase in E_b/N_0 required so that this coherent FSK receiver has the same noise performance as a coherent binary PSK receiver.

9.9 A binary FSK signal with *discontinuous phase* is defined by

$$s(t) = \begin{cases} \sqrt{\dfrac{2E_b}{T_b}}\cos\left[2\pi\left(f_c + \dfrac{\Delta f}{2}\right)t + \theta_1\right] & \text{for symbol 1} \\ \sqrt{\dfrac{2E_b}{T_b}}\cos\left[2\pi\left(f_c - \dfrac{\Delta f}{2}\right)t + \theta_2\right] & \text{for symbol 0} \end{cases}$$

where E_b is the signal energy per bit, T_b is the bit duration, and θ_1 and θ_2 are sample values of uniformly distributed random variables over the interval 0 to 2π. In effect, the two oscillators supplying the transmitted frequencies $f_c \pm \Delta f/2$ operate independently of each other. Assume that $f_c >> \Delta f$.

(a) Evaluate the power spectral density of the FSK signal.

(b) Show that for frequencies far removed from the carrier frequency f_c, the power spectral density falls off as the inverse square of frequency.

9.10 Set up a block diagram for the generation of CPFSK signal $s(t)$ by using the representation given here:

$$s(t) = \sqrt{\frac{2E_b}{T_b}}\cos\left(\frac{\pi t}{T_b}\right)\cos(2\pi f_c t) \mp \sqrt{\frac{2E_b}{T_b}}\sin\left(\frac{\pi t}{T_b}\right)\sin(2\pi f_c t)$$

9.11 Binary data are transmitted over a microwave link at the rate of 10^6 bits per second and the power spectral density of the noise at the receiver input is 10^{-10} watts per hertz. Find the average carrier power required to maintain an average probability of error $P_e \le 10^{-4}$ for (a) coherent binary PSK, and (b) DPSK.

9.12 The values of E_b/N_0 required to realize an average probability of bit error $P_e = 10^{-4}$ using coherent binary PSK and coherent FSK (conventional) systems are equal to 7.2 and 13.5, respectively. Using the approximation

$$Q(u) \approx \frac{1}{\sqrt{2\pi}u}\exp\left(-\frac{u^2}{2}\right)$$

determine the separation in the values of E_b/N_0 for $P_e = 10^{-4}$, using

(a) Coherent binary PSK and DPSK.

(b) Coherent binary PSK and QPSK.

(c) Coherent binary PSK (conventional) and noncoherent binary FSK.

(d) Coherent binary FSK (conventional) and coherent MSK.

9.13 A binary ASK system uses on-off signaling. Sketch a receiver block diagram for detecting this signal.

9.14 Construct the block diagram of a DPSK transmitter that corresponds to the DPSK receiver of Figure 9.14. Then, do the following:

(a) Apply the binary sequence 1100100010 to this transmitter and sketch the resulting waveform at the transmitter output.

(b) Applying this waveform to the DPSK receiver of Figure 9.14, show that, in the absence of noise, the original binary sequence is reconstructed at the receiver output.

9.15

(a) Given the input binary sequence 1100100010, sketch the waveforms of the in-phase and quadrature components of a modulated wave obtained by using the QPSK based on the signal set of Figure 9.16.

(b) Sketch the QPSK waveform itself for the input binary sequence specified in part (a).

9.16 Let P_{eI} and P_{eQ} denote the probabilities of symbol error for the in-phase and quadrature channels of a narrowband system. Show that the average probability of symbol error for the overall system is given by

$$P_e = P_{eI} + P_{eQ} - P_{eI}P_{eQ}$$

9.17 There are two ways of detecting an MSK signal. One way is to use a coherent receiver to take full account of the phase information content of the MSK signal. Another way is to use a noncoherent receiver and disregard the phase information. The second method offers the advantage of simplicity of implementation, at the expense of a degraded noise performance. By how many decibels do we have to increase the bit energy-to-noise density ratio E_b/N_0 in the second case so as to realize an average probability of symbol error equal to 10^{-5} in both cases?

9.18

(a) Sketch the waveforms of the in-phase and quadrature components of the MSK signal in response to the input binary sequence 1100100010.

(b) Sketch the MSK waveform itself for the binary sequence specified in part (a).

9.19 In Section 9.5 we compared the noise performances of coherent binary PSK, coherent binary FSK, QPSK, MSK, DPSK, and noncoherent FSK by using the bit error rate as the basis of comparison. In this problem we take a different viewpoint and use the average probability of symbol error, P_e, to do the comparison. Plot P_e versus E_b/N_0 for each of these schemes, and comment on your results.

9.20 Figure P9.20*a* shows a noncoherent receiver using a matched filter for the detection of a sinusoidal signal of known frequency but random phase, in the presence of additive white Gaussian noise. An alternative implementation of this receiver is its mechanization in the frequency domain as a *spectrum analyzer receiver*, as in Figure P9.20*b*, where the correlator computes the finite time autocorrelation function $R_x(\tau)$ defined by

$$R_x(\tau) = \int_0^{T-\tau} x(t)x(t+\tau)\,dt, \qquad 0 \le \tau \le T$$

Show that the square-law envelope detector output sampled at time $t = T$ in Figure P9.20*a* is twice the spectral output of the Fourier transformer sampled at frequency $f = f_c$ in Figure P9.20*b*.

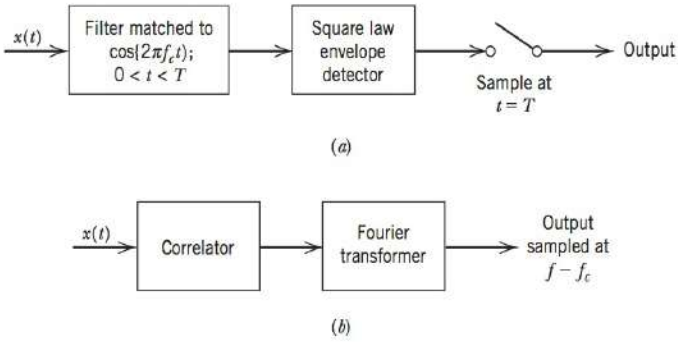


Figure P9.20

9.21

(a) Determine an expression for the baseband equivalent spectrum of a binary PSK signal. (Assume a bipolar NRZ line code is used for modulation.)

(b) What is the analytical expression for the MSK pulse shape? Assuming that the in-phase and quadrature components of an MSK signal are independent, develop an analytical expression for the spectrum of an MSK signal.

9.22 The *noise equivalent bandwidth* of a band-pass signal is defined as the value of bandwidth that satisfies the relation

$$2BS(f_c) = P/2$$

where $2B$ is the noise equivalent bandwidth centered around the midband frequency f_c, $S(f_c)$ is the maximum value of the power spectral density of the signal at $f = f_c$, and P is the average power of the signal. Show that the noise equivalent bandwidths of binary PSK, QPSK, and MSK are as follows:

Type of Modulation	Noise Bandwidth/Bit Rate
Binary PSK	1.0
QPSK	0.5
MSK	0.62

9.23 For the noncoherent FSK detector of Figure 9.12, assume that the signaling frequencies f_0 and f_1 are orthogonal over the symbol period T_b. Show that when the received signal is $x(t) = A_c \cos(2\pi f_1 t) + n(t)$ where $n(t)$ is white Gaussian noise with density $N_0/2$:

(a) The output of the lower envelope detector has a Rayleigh density function given by

$$P_{L_0}(l_0) = \begin{cases} \dfrac{2l_0}{N_0}\exp\left(-\dfrac{l_0^2}{N_0}\right) & l_0 \geq 0 \\ 0 & \textit{elsewhere} \end{cases}$$

(b) The output of the upper envelope detector has a Rician distribution given by

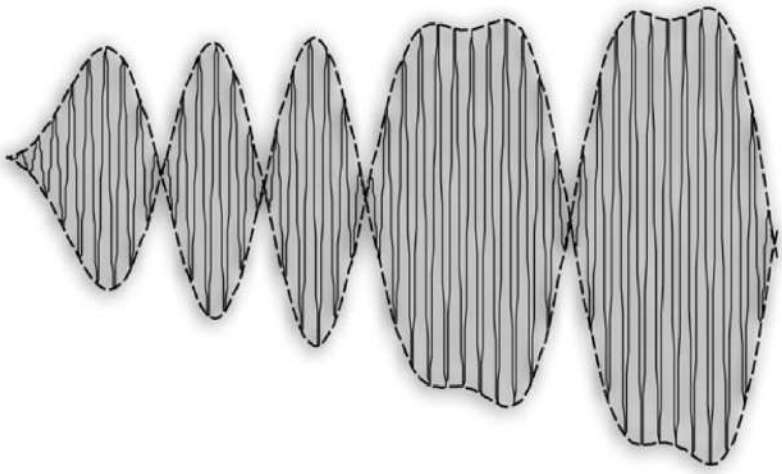
$$P_{L_1}(l_1) = \begin{cases} \dfrac{2l_1}{N_0}\exp\left(-\dfrac{l_1^2 + A_c^2}{N_0}\right) I_0\left(\dfrac{2A_c l_1}{N_0}\right) & l_1 \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

(c) Show that

$$\begin{aligned} P(L_0 > L_1) &= \int_0^\infty P(L_0 > l_1 | l_1) p_{L_1}(l_1)\, dl_1 \\ &= \frac{1}{2}\exp\left(-\frac{A_c^2}{4N_0}\right) \end{aligned}$$

Hint: Use the complex baseband equivalents of $l_0^2 = x_I^2 + x_Q^2$ and $l_I^2 = (x_I + A_c)^2 + x_Q^2$.

10



INFORMATION THEORY AND CODING

10.1 INTRODUCTION

As mentioned in Chapter 1 and reiterated along the way, the purpose of a communication system is to carry information-bearing baseband signals from one place to another over a communication channel. In preceding chapters of the book, we have described a variety of modulation schemes for accomplishing this objective. But what do we mean by the term information? To address this issue, we need to invoke information theory.[1] This broadly based mathematical discipline has made fundamental contributions, not only to communications, but also computer science, statistical physics, statistical inference, and probability.

In the context of communications, information theory deals with mathematical modeling and analysis of a communication system rather than with physical sources and physical channels. In particular, it provides answers to two fundamental questions (among others):

- What is the irreducible complexity below which a signal cannot be compressed?
- What is the ultimate transmission rate for reliable communication over a noisy channel?

The answers to these questions lie in the *entropy* of a source and the *capacity* of a channel, respectively. Entropy is defined in terms of the probabilistic behavior of a source of information. It is so named in deference to the parallel use of this concept in thermodynamics. Capacity is defined as the intrinsic ability of a channel to convey information. It is naturally related to the noise characteristics of the channel. A remarkable result that emerges from information theory is that if the entropy of the source is less than the capacity of the channel, then error-free communication over the channel can be achieved.

This chapter is devoted to two topics: information theory and error-control coding. The study of information theory provides the fundamental limits on the performance of a communication system by specifying the minimum number of bits per symbol required

discrete memoryless source, memoryless in the sense that the symbol emitted at any time is independent of previous choices.

Can we find a measure of how much information is produced by such a source? To answer this question, we note that the idea of information is closely related to that of uncertainty or surprise, as described next.

Consider the event $S = s_k$, describing the emission of symbol s_k by the source with probability p_k, as defined in Eq. (10.2). Clearly, if the probability $p_k = 1$ and $p_i = 0$ for all $i \neq k$, then there is no surprise and therefore no information when symbol s_k is emitted, since we know what the message from the source must be. If, on the other hand, the source symbols occur with different probabilities, and the probability p_k is low, then there is more surprise and therefore information when symbol s_k is emitted by the source than when symbol s_i, $i \neq k$, with higher probability is emitted. Thus, the words uncertainty, surprise, and information are all related. Before the event $S = s_k$ occurs, there is an amount of uncertainty. When the event $S = s_k$ occurs there is an amount of surprise. After the occurrence of the event $S = s_k$, there is gain in the amount of information, the essence of which may be viewed as the *resolution of uncertainty*. Moreover, the amount of information is related to the *inverse* of the probability of occurrence.

We define the amount of information gained after observing the event $S = s_k$, which occurs with probability p_k, as the *logarithmic* function

$$I(s_k) = \log\left(\frac{1}{p_k}\right) \tag{10.4}$$

as illustrated in Figure 10.1. The definition of Eq. (10.4) exhibits the following important properties that are intuitively satisfying:

1. $$I(s_k) = 0 \quad \text{for} \quad p_k = 1 \tag{10.5}$$

Obviously, if we are absolutely *certain* of the outcome of an event, even before it occurs, there is *no* information gained.

2. $$I(s_k) \geq 0 \quad \text{for} \quad 0 \leq p_k \leq 1 \tag{10.6}$$

That is to say, the occurrence of an event $S = s_k$ either provides some or no information, but never brings about a *loss* of information.

3. $$I(s_k) > I(s_i) \quad \text{for} \quad p_k < p_i \tag{10.7}$$

That is, the less probable an event is, the more information we gain when it occurs.

4. $I(s_k s_l) = I(s_k) + I(s_l)$ if s_k and s_l are statistically independent.

The base of the logarithm in Eq. (10.4) is quite arbitrary. Nevertheless, it is the standard practice today to use a logarithm to base 2. The resulting unit of information is called the *bit* (a contraction of *bi*nary digi*t*). We thus write

$$\begin{aligned} I(s_k) &= \log_2\left(\frac{1}{p_k}\right) \\ &= -\log_2 p_k \quad \text{for } k = 0, 1, \ldots, K-1 \end{aligned} \tag{10.8}$$

When $p_k = 1/2$, we have $I(s_k) = 1$ bit. Hence, *one bit is the amount of information that we gain when one of two*

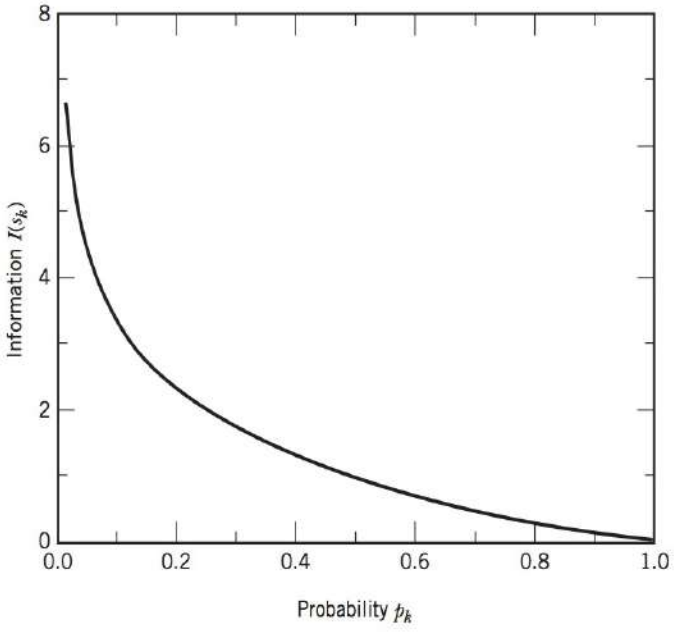


FIGURE 10.1 Information associated with the event $S = s_k$ of probability p_k.

possible and equally likely (*i.e., equiprobable*) *events occurs*. Note that the information $I(s_k)$ is positive, since the logarithm of a number less than one, such as a probability, is negative.

The amount of information $I(s_k)$ produced by the source during an arbitrary signaling interval depends on the symbol s_k emitted by the source at that time. Indeed, $I(s_k)$ is a discrete random variable that takes on the values $I(s_0), I(s_1), \ldots, I(s_{K-1})$ with probabilities $p_0, p_1, \ldots, p_{K-1}$ respectively. The mean of $I(s_k)$ over the source alphabet $\mathcal{S}$ is given by

$$\begin{aligned} H(\mathcal{S}) &= \mathbf{E}[I(s_k)] \\ &= \sum_{k=0}^{K-1} p_k I(s_k) \\ &= \sum_{k=0}^{K-1} p_k \log_2\left(\frac{1}{p_k}\right) \end{aligned} \tag{10.9}$$

The important quantity $H(\mathcal{S})$ is called the *entropy*[2] of a discrete memoryless source with source alphabet $\mathcal{S}$. It is a measure of the *average information content per source symbol*. Note that the entropy $H(\mathcal{S})$ depends only on the probabilities of the symbols in the alphabet $\mathcal{S}$ of the source. Thus, the symbol $\mathcal{S}$ in $H(\mathcal{S})$ is not an argument of a function but rather a label for a source.

SOME PROPERTIES OF ENTROPY

Consider a discrete memoryless source whose mathematical model is defined by Eqs. (10.1) and (10.2). The entropy $H(\mathcal{S})$ of such a source is bounded as follows:

$$0 \leq H(\mathcal{S}) \leq \log_2 K \tag{10.10}$$

where K is the number of symbols of the source alphabet $\mathcal{S}$. Furthermore, we may make two statements:

1. $H(\mathcal{S}) = 0$, if and only if the probability $p_k = 1$ for some k, and the remaining probabilities in the set are all zero; this lower bound on entropy corresponds to *no uncertainty*.
2. $H(\mathcal{S}) = \log_2 K$, if and only if $p_k = 1/K$ for all k (i.e., all the symbols in the alphabet $\mathcal{S}$ are *equiprobable*); this upper bound on entropy corresponds to *maximum uncertainty*.

EXAMPLE 10.1 Entropy of Binary Memoryless Source

To illustrate the properties of $H(\mathcal{S})$, we consider a binary source for which symbol 0 occurs with probability p_0 and symbol 1 with probability $p_1 = 1 - p_0$. We assume that the source is memoryless so that successive symbols emitted by the source are statistically independent.

The entropy of such a source equals

$$\begin{aligned} H(\mathcal{S}) &= -p_0 \log_2 p_0 - p_1 \log_2 p_1 \\ &= -p_0 \log_2 p_0 - (1 - p_0)\log_2(1 - p_0) \text{ bits} \end{aligned} \tag{10.11}$$

We note that

1. When $p_0 = 0$, the entropy $H(\mathcal{S}) = 0$; this follows from the fact that $x \log x \to 0$ as $x \to 0$.
2. When $p_0 = 1$, the entropy $H(\mathcal{S}) = 0$.
3. The entropy $H(\mathcal{S})$ attains its maximum value, $H_{\max} = 1$ bit, when $p_1 = p_0 = 1/2$, that is, symbols 1 and 0 are equally probable.

The function of p_0 given on the right-hand side of Eq. (10.11) is frequently encountered in information-theoretic problems. It is therefore customary to assign a special symbol to this function. Specifically, we define

$$\mathcal{H}(p_0) = -p_0 \log_2 p_0 - (1 - p_0)\log_2(1 - p_0) \tag{10.12}$$

We refer to $\mathcal{H}(p_0)$ as the *entropy function*. The distinction between Eq. (10.11) and Eq. (10.12) should be carefully noted. The $H(\mathcal{S})$ of Eq. (10.11) gives the entropy of a discrete memoryless source with source alphabet $\mathcal{S}$. The $\mathcal{H}(p_0)$ of Eq. (10.12), on the other hand, is a function of the prior probability p_0 defined on the interval [0, 1]. Accordingly, we may plot the entropy function $\mathcal{H}(p_0)$ versus p_0, defined on the interval [0, 1], as in Figure 10.2. The curve in Figure 10.2 highlights the observations made under points 1, 2, and 3.

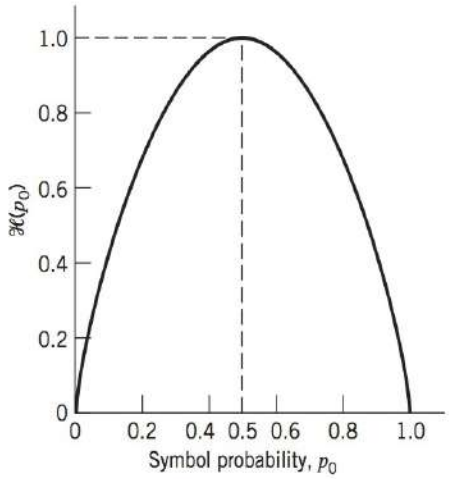


FIGURE 10.2 Entropy function $\mathcal{H}(p_0)$.

EXTENSION OF A DISCRETE MEMORYLESS SOURCE

In discussing information-theoretic concepts, we often find it useful to consider *blocks* rather than individual symbols, with each block consisting of n successive source symbols. We may view each such block as being produced by an *extended source* with a source alphabet $\mathcal{S}^n$ that has K^n *distinct* blocks, where K is the number of distinct symbols in the source alphabet $\mathcal{S}$ of the original source. In the case of a discrete memoryless source, the source symbols are statistically independent. Hence, the probability of a source symbol in $\mathcal{S}^n$ is equal to the product of the probabilities of the n source symbols in $\mathcal{S}$ constituting the particular source symbol in $\mathcal{S}^n$. We may thus intuitively expect that $H(\mathcal{S}^n)$, the entropy of the extended source, is equal to n times $H(\mathcal{S})$, the entropy of the original source. That is, we may write

$$H(\mathcal{S}^n) = nH(\mathcal{S}) \tag{10.13}$$

EXAMPLE 10.2 Second-order Extension of Discrete Memoryless Source

Consider a discrete memoryless source with source alphabet $\mathcal{S} = \{s_0, s_1, s_2\}$ with respective probabilities

$$p_0 = \frac{1}{4}$$

$$p_1 = \frac{1}{4}$$

$$p_2 = \frac{1}{2}$$

Hence, the use of Eq. (10.9) yields the entropy of the source as

$$\begin{aligned} H(\mathcal{S}) &= p_0 \log_2\left(\frac{1}{p_0}\right) + p_1 \log_2\left(\frac{1}{p_1}\right) + p_2 \log_2\left(\frac{1}{p_2}\right) \\ &= \frac{1}{4}\log_2(4) + \frac{1}{4}\log_2(4) + \frac{1}{2}\log_2(2) \\ &= \frac{3}{2} \text{ bits} \end{aligned}$$

Consider next the second-order extension of the source. With the source alphabet $\mathcal{S}$ consisting of three symbols, it follows that the source alphabet $\mathcal{S}^2$ of the extended source has nine symbols. The first row of Table 10.1 presents the nine symbols of $\mathcal{S}^2$, denoted as $\sigma_0, \sigma_1, \ldots, \sigma_8$. The second row of the table presents the compositions of these nine symbols in terms of the corresponding sequences of source symbols s_0, s_1, and s_2, taken two at a time. The probabilities of the nine source symbols of the extended source are presented in the last row of the table. Accordingly, the use of Eq. (10.9) yields the entropy of the extended source as

$$\begin{aligned} H(\mathcal{S}^2) &= \sum_{i=0}^{8} p(\sigma_i)\log_2\frac{1}{p(\sigma_i)} \\ &= \frac{1}{16}\log_2(16) + \frac{1}{16}\log_2(16) + \frac{1}{8}\log_2(8) + \frac{1}{16}\log_2(16) \\ &\quad + \frac{1}{16}\log_2(16) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8) + \frac{1}{4}\log_2(4) \\ &= 3 \text{ bits} \end{aligned}$$

We thus see that $H(\mathcal{S}^2) = 2H(\mathcal{S})$ in accordance with Eq. (10.13).

10.3 SOURCE-CODING THEOREM

An important problem in communications is the *efficient* representation of data generated by a discrete source. The process by which this representation is accomplished is called *source encoding*. The device that performs the representation is called a *source encoder*. For the source encoder to be *efficient*, we require knowledge of the statistics of the source. In particular, if some source symbols are known to be more probable than others, then we may exploit this feature in the generation of a *source code* by assigning *short* codewords to *frequent* source symbols, and *long* codewords to *rare* source symbols. We refer to such a source code as a *variable-length code*. The *Morse code* is an example of a variable-length code. In the Morse code, the letters of the alphabet and numerals are

TABLE 10.1 Alphabet particulars of second-order extension of a discrete memoryless source

Symbols of $\mathcal{S}^2$	σ_0	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
Corresponding sequences of symbols of $\mathcal{S}$	s_0s_0	s_0s_1	s_0s_2	s_1s_0	s_1s_1	s_1s_2	s_2s_0	s_2s_1	s_2s_2
Probability $p(\sigma_i)$, $i = 0, 1, \ldots, 8$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

encoded into streams of *marks* and *spaces,* denoted as dots "." and dashes "-", respectively. Since in the English language, the letter E occurs more frequently than the letter Q for example, the Morse code encodes E into a single dot ".", the shortest codeword in the code, and it encodes Q into "--.-", the longest codeword in the code.

Our primary interest is in the development of an efficient source encoder that satisfies two functional requirements:

1. The codewords produced by the encoder are in *binary* form.
2. The source code is *uniquely decodable*, so that the original source sequence can be reconstructed perfectly from the encoded binary sequence.

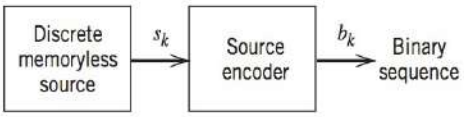


FIGURE 10.3 Source encoding.

Consider then the scheme shown in Figure 10.3, which depicts a discrete memoryless source whose output s_k is converted by the source encoder into a block of 0s and 1s, denoted by b_k. We assume that the source has an alphabet with K different symbols, and that the kth symbol s_k occurs with probability p_k, $k = 0, 1, \ldots, K-1$. Let the binary codeword assigned to symbol s_k by the encoder have length l_k, measured in bits. We define the average codeword length, $\bar{L}$, of the source encoder as

$$\bar{L} = \sum_{k=0}^{K-1} p_k l_k \tag{10.14}$$

In physical terms, the parameter $\bar{L}$ represents the *average number of bits per source symbol* used in the source encoding process. Let $L_{\min}$ denote the *minimum* possible value of $\bar{L}$. We then define the *coding efficiency* of the source encoder as

$$\eta = \frac{L_{\min}}{\bar{L}} \tag{10.15}$$

With $\bar{L} \geq L_{\min}$, we clearly have $\eta \leq 1$. The source encoder is said to be *efficient* when η approaches unity.

But how is the minimum value $L_{\min}$ determined? The answer to this fundamental question is embodied in Shannon's first theorem: the *source-coding theorem*, which may be stated as follows:

Given a discrete memoryless source of entropy $H(\mathcal{S})$, the average codeword length $\bar{L}$ for any distortionless source encoding is bounded as

$$\bar{L} \geq H(\mathcal{S}) \tag{10.16}$$

Accordingly, the entropy $H(\mathcal{S})$ represents a *fundamental limit* on the average number of bits per source symbol necessary to represent a discrete memoryless source in that it can be made as small as, but no smaller than, the entropy $H(\mathcal{S})$. Thus with $L_{\min} = H(\mathcal{S})$, we may rewrite the efficiency of a source encoder in terms of the entropy $H(\mathcal{S})$ as

$$\eta = \frac{H(\mathcal{S})}{\bar{L}} \tag{10.17}$$

10.4 LOSSLESS DATA COMPRESSION

A common characteristic of signals generated by physical sources is that, in their natural form, they contain a significant amount of information that is *redundant*, the transmission of which is therefore wasteful of primary communication resources. For *efficient* signal transmission, the *redundant information should be removed from the signal prior to transmission.* This operation is ordinarily performed on a signal in digital form, in which

case we refer to it as *lossless data compression*. The code resulting from such an operation provides a representation of the source output that is not only efficient in terms of the average number of bits per symbol but also exact in the sense that the original data can be reconstructed with no loss of information. The entropy of the source establishes the fundamental limit on the removal of redundancy from the data. Basically, data compression is achieved by assigning short descriptions to the most frequent outcomes of the source output and longer descriptions to the less frequent ones.

In this section, we discuss some source-coding schemes for data compression. We begin the discussion by describing a type of source code known as a prefix code, which is not only decodable but also offers the possibility of realizing an average codeword length that can be made arbitrarily close to the source entropy.

PREFIX CODING

Consider a discrete memoryless source of source alphabet $\{s_0, s_1, \ldots, s_{K-1}\}$ and source statistics $\{p_0, p_1, \ldots, p_{k-1}\}$. For a source code representing the output of this source to be of practical use, the code has to be uniquely decodable. This restriction ensures that for each finite sequence of symbols emitted by the source, the corresponding sequence of codewords is different from the sequence of codewords corresponding to any other source sequence. We are specifically interested in a special class of codes satisfying a restriction known as the *prefix condition*. To define the prefix condition, let the codeword assigned to source symbol s_k be denoted by $(m_{k_1}, m_{k_2}, \ldots, m_{k_n})$, where the individual elements $m_{k_1}, \ldots, m_{k_n}$ are 0s and 1s, and n is the code-word length. The initial part of the codeword is represented by the elements $m_{k_1}, \ldots, m_{k_i}$ for some $i \leq n$. Any sequence made up of the initial part of the codeword is called a *prefix* of the codeword. A *prefix code* is defined as a code in which no codeword is the prefix of any other codeword.

To illustrate the meaning of a prefix code, consider the three source codes described in Table 10.2. Code 1 is not a prefix code since the bit 0, the codeword for s_0, is a prefix of 00, the codeword for s_2. Likewise, the bit 1, the codeword for s_1, is a prefix of 11, the codeword for s_3. Similarly, we may show that code III is not a prefix code, but code II is.

To decode a sequence of codewords generated from a prefix source code, the *source decoder* simply starts at the beginning of the sequence and decodes one codeword at a time. Specifically, it sets up what is equivalent to a *decision tree*, which is a graphical portrayal of the codewords in the particular source code. For example, Figure 10.4 depicts the decision tree corresponding to code II in Table 10.2. The tree has an *initial state* and four *terminal states* corresponding to source symbols s_0, s_1, s_2, and s_3. The decoder always starts at the initial state. The first received bit moves the decoder to the terminal state s_0 if it is 0, or else to a second decision point if it is 1. In the latter case, the second bit moves the decoder one step further down the tree, either to terminal state s_1 if it is 0, or else to a third decision point if it is 1, and so on. Once each terminal state emits its symbol, the decoder is reset to its initial state. Note also that

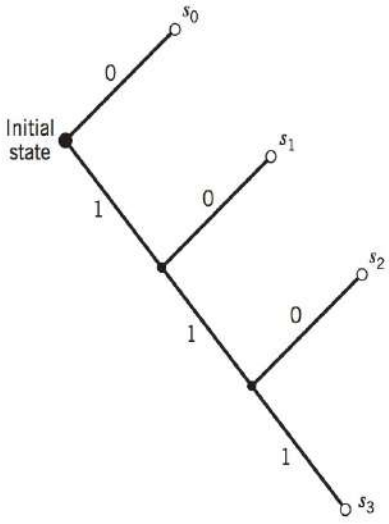


FIGURE 10.4 Decision tree for code II of Table 10.2.

TABLE 10.2 Illustrating the definition of a prefix code

Source Symbol	Probability of Occurrence	Code I	Code II	Code III
s_0	0.5	0	0	0
s_1	0.25	1	10	01
s_2	0.125	00	110	011
s_3	0.125	11	111	0111

each bit in the received encoded sequence is examined only once. For example, the encoded sequence 1011111000 ... is readily decoded as the source sequence $s_1 s_3 s_2 s_0 s_0 \ldots$. The reader is invited to carry out this decoding.

A prefix code has the important property that it is *always* uniquely decodable. Indeed, if a prefix code has been constructed for a discrete memoryless source with source alphabet $\{s_0, s_1, \ldots, s_{K-1}\}$ and source statistics $\{p_0, p_1, \ldots, p_{K-1}\}$ and the codeword for symbol s_k has length l_k, $k = 0, 1, \ldots, K-1$, then the codeword lengths of the code satisfy a certain inequality known as the *Kraft–McMillan inequality*. In mathematical terms, we may state that

$$\sum_{k=0}^{K-1} 2^{-l_k} \leq 1 \tag{10.18}$$

where the factor 2 refers to the radix (number of symbols) in the binary alphabet. Conversely, we may state that if the codeword lengths of a code for a discrete memoryless source satisfy the Kraft–McMillan inequality, then a prefix code with these codeword lengths can be constructed.

Although all prefix codes are uniquely decodable, the converse is not true. For example, code III in Table 10.2 does not satisfy the prefix condition, and yet it is uniquely decodable since the bit 0 indicates the beginning of each codeword in the code.

Prefix codes are distinguished from other uniquely decodable codes by the fact that the end of a codeword is always recognizable. Hence, the decoding of a prefix can be accomplished as soon as the binary sequence representing a source symbol is fully received. For this reason, prefix codes are also referred to as *instantaneous codes*.

Given a discrete memoryless source of entropy $H(\mathcal{L})$, the average codeword length $\bar{L}$ of a prefix code is bounded as follows:

$$H(\mathcal{L}) \leq \bar{L} < H(\mathcal{L}) + 1 \tag{10.19}$$

The left-hand bound of (10.19) is satisfied with equality under the condition that symbol s_k is emitted by the source with probability

$$p_k = 2^{-l_k} \tag{10.20}$$

where l_k is the length of the codeword assigned to source symbol s_k. We then have

$$\sum_{k=0}^{K-1} 2^{-l_k} \leq \sum_{k=0}^{K-1} p_k = 1$$

Under this condition, the Kraft–McMillan inequality of Eq. (10.18) implies that we can construct a prefix code, such that the length of the codeword assigned to source symbol s_k is l_k. For such a code, the average codeword length is

$$\bar{L} = \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}} \tag{10.21}$$

and the corresponding entropy of the source is

$$\begin{aligned} H(\mathcal{L}) &= \sum_{k=0}^{K-1} \left(\frac{1}{2^{l_k}} \right) \log_2(2^{l_k}) \\ &= \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}} \end{aligned} \tag{10.22}$$

Hence, in this special (rather meretricious) case, we find from Eqs. (10.21) and (10.22) that the prefix code is *matched* to the source in that $\bar{L} = H(\mathcal{L})$.

HUFFMAN CODING

We next describe an important class of prefix codes known as Huffman codes. The basic idea behind *Huffman coding* is to assign to each symbol of an alphabet a sequence of bits roughly equal in length to the amount of information conveyed by the symbol in question. The end result is a source code whose average codeword length approaches the fundamental limit set by the entropy of a discrete memoryless source, namely, $H(\mathcal{L})$. The essence of the *algorithm* used to synthesize the Huffman code is to replace the prescribed set of source statistics of a discrete memoryless source with a simpler one. This *reduction* process is continued in a step-by-step manner until we are left with a final set of only two source statistics (symbols), for which (0, 1) is an optimal code. Starting from this trivial code, we then work backward and thereby construct the Huffman code for the given source.

Specifically, the Huffman *encoding algorithm* proceeds as follows:

1. The source symbols are listed in order of decreasing probability. The two source symbols of lowest probability are assigned a 0 and a 1. This part of the step is referred to as a *splitting* stage.
2. These two source symbols are regarded as being *combined* into a new source symbol with probability equal to the sum of the two original probabilities. (The list of source symbols, and therefore source statistics, is thereby *reduced* in size by one.) The probability of the new symbol is placed in the list in accordance with its value.
3. The procedure is repeated until we are left with a final list of source statistics (symbols) of only two, for which a 0 and a 1 are assigned.

The code for each (original) source symbol is found by working backward and tracing the sequence of 0s and 1s assigned to that symbol as well as its successors.

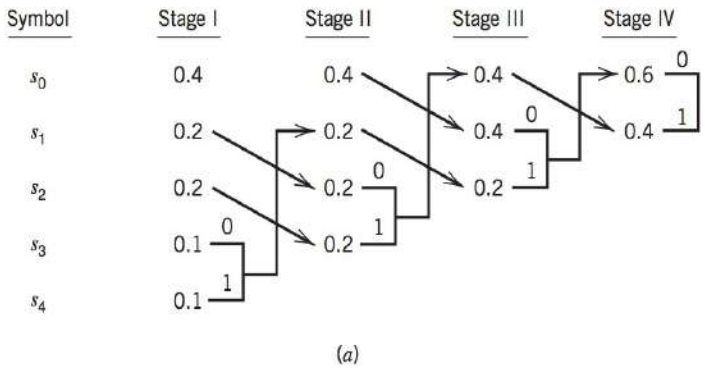
EXAMPLE 10.3 Huffman Algorithm

The five symbols of the alphabet of a discrete memoryless source and their probabilities are shown in the left-most two columns of Figure 10.5*a*. Following through the Huffman algorithm, we reach the end of the computation in four steps, resulting in the *Huffman tree* shown in Figure 10.5*a*. The codewords of the Huffman code for the source are tabulated in Figure 10.5*b*. The average codeword length is therefore

$$\begin{aligned}\bar{L} &= 0.4(2) + 0.2(2) + 0.2(2) + 0.1(3) + 0.1(3) \\ &= 2.2\end{aligned}$$

The entropy of the specified discrete memoryless source is calculated as follows [see Eq. (10.9)]:

$$\begin{aligned}H(\mathcal{L}) &= 0.4\log_2\left(\frac{1}{0.4}\right) + 0.2\log_2\left(\frac{1}{0.2}\right) + 0.2\log_2\left(\frac{1}{0.2}\right) \\ &\quad + 0.1\log_2\left(\frac{1}{0.1}\right) + 0.1\log_2\left(\frac{1}{0.1}\right) \\ &= 0.52877 + 0.46439 + 0.46439 + 0.33219 + 0.33219 \\ &= 2.12193\end{aligned}$$



Symbol	Probability	Codeword
s_0	0.4	00
s_1	0.2	10
s_2	0.2	11
s_3	0.1	010
s_4	0.1	011

(b)

FIGURE 10.5 (*a*) Example of the Huffman encoding algorithm, (*b*) source code.

For the example at hand, we may make two observations:

1. The average codeword length $\bar{L}$ exceeds the entropy $H(\mathcal{S})$ by only 3.67 percent.
2. The average codeword length $\bar{L}$ does indeed satisfy Eq. (10.19).

It is noteworthy that the Huffman encoding process (i.e., the Huffman tree) is not unique. In particular, we may site two variations in the process that are responsible for the nonuniqueness of the Huffman code. First, at each splitting stage in the construction of a Huffman code, there is arbitrariness in the way a 0 and a 1 are assigned to the last two source symbols. Whichever way the assignments are made, however, the resulting differences are trivial. Second, ambiguity arises when the probability of a *combined* symbol (obtained by adding the last two probabilities pertinent to a particular step) is found to equal another probability in the list. We may proceed by placing the probability of the new symbol as *high* as possible, as in Example 10.3. Alternatively, we may place it as *low* as possible. (It is presumed that whichever way the placement is made, high or low, it is consistently adhered to throughout the encoding process.) But this time, noticeable differences arise in that the codewords in the resulting source code can have different lengths. Nevertheless, the average codeword length remains the same.

As a measure of the variability in codeword lengths of a source code, we define the *variance* of the average codeword length $\bar{L}$ over the ensemble of source symbols as

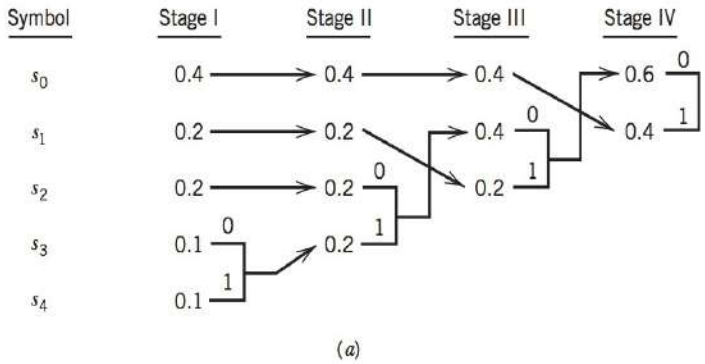
$$\sigma^2 = \sum_{k=0}^{K-1} p_k (l_k - \bar{L})^2 \tag{10.23}$$

where $p_0, p_1, \ldots, p_{K-1}$ are the source statistics, and l_k is the length of the codeword assigned to source symbol s_k. It is usually found that when a combined symbol is moved as high as possible, the resulting Huffman code has a significantly smaller variance σ^2 than when it is moved as low as possible. On this basis, it is reasonable to choose the former Huffman code over the latter.

In Example 10.3, a combined symbol was moved as high as possible. In Example 10.4, presented next, a combined symbol is moved as low as possible. Thus, by comparing the results of these two examples, we are able to appreciate the subtle differences and similarities between the two Huffman codes.

EXAMPLE 10.4 Nonuniqueness of the Huffman Algorithm

Consider again the same discrete memoryless source described in Example 10.3. This time, however, we move the probability of a combined symbol as low as possible. The resulting Huffman tree is shown in Figure 10.6*a*. Working backward through this tree and tracing through the various steps, we find



Symbol	Probability	Codeword
s_0	0.4	1
s_1	0.2	01
s_2	0.2	000
s_3	0.1	0010
s_4	0.1	0011

(b)

FIGURE 10.6 (*a*) Example illustrating nonuniqueness of the Huffman encoding algorithm. (*b*) Another source code.

that the codewords of this second Huffman code for the source are as tabulated in Figure 10.6*b*. The average codeword length for the second Huffman code is therefore

$$\begin{aligned}\bar{L} &= 0.4(1) + 0.2(2) + 0.2(3) + 0.1(4) + 0.1(4) \\ &= 2.2\end{aligned}$$

which is exactly the same as that for the first Huffman code of Example 10.3. However, as remarked earlier, the individual codewords of the second Huffman code have different lengths, compared to the corresponding ones of the first Huffman code.

The use of Eq. (10.30) yields the variance of the first Huffman code obtained in Example 10.3 as

$$\begin{aligned}\sigma_1^2 &= 0.4(2-2.2)^2 + 0.2(2-2.2)^2 + 0.2(2-2.2)^2 \\ &\quad + 0.1(3-2.2)^2 + 0.1(3-2.2)^2 = 0.16\end{aligned}$$

On the other hand, for the second Huffman code obtained in this example, we have from Eq. (10.30):

$$\begin{aligned}\sigma_2^2 &= 0.4(1-2.2)^2 + 0.2(2-2.2)^2 + 0.2(3-2.2)^2 \\ &\quad + 0.1(4-2.2)^2 + 0.1(4-2.2)^2 = 1.36\end{aligned}$$

These results confirm that the minimum variance Huffman code is obtained by moving the probability of a combined symbol as high as possible.

10.5 THEME EXAMPLE—THE LEMPEL–ZIV ALGORITHM AND FILE COMPRESSION[3]

A drawback of the Huffman algorithm is that it requires knowledge of a probabilistic model of the source. Unfortunately, in practice, source statistics are not always known *a priori*. Moreover, in modeling text we find that storage limitations prevent the Huffman algorithm from capturing the higher-order relationships between words and phrases, thereby compromising the efficiency of the code. To overcome these practical limitations, we may use the *Lempel–Ziv algorithm*, which is intrinsically adaptive.

The Lempel–Ziv algorithm is a dictionary-based scheme, although it differs from similar schemes in that it uses the previously seen input stream as the dictionary. The

encoder maintains a *sliding window* that shifts from left to right over the input stream as the data sequence is being encoded. This sliding window is illustrated in Figure 10.7. The window is divided in two parts. The part on the left is called the *search buffer*. The search buffer includes the symbols that have been recently encoded. It represents the dictionary. The part on the right is called the *look-ahead buffer*. The vertical bar | in Figure 10.7*a* is the dividing line (pointer) between the two buffers. In practice, the search buffer is typically kilobytes long, while the look-ahead buffer is only tens of bytes in length. The algorithm is best illustrated by way of an example.

- The encoder reads the first symbol of the look-ahead buffer, in this case a **y**. Then it scans the search buffer from right to left looking for a match.
- Once the encoder finds a match for the **y**, it determines how long the match is. In this case, the first match is found at offset from the pointer of 43, and the match is of length 6: **y pick**.
- The encoder continues to scan the search buffer for matches and records the longest string that matches. In this example, the search buffer is only of length 47 bytes and the longest match occurs at an offset of 43 from the pointer.

Once the search buffer has been completely scanned, the encoder produces a codeword that consists of three parts:

- The *offset* from the pointer to the longest string that matches in the search buffer
- The *length* of the longest matching string
- The *next symbol* in the input string after the match.

Then, the sliding window is advanced so that the pointer is positioned one character beyond the longest matching string as shown in the Figure 10.7*b*. In practice, the search buffer may be 4 kilobytes in length; thus a maximum of 12 bits is required to represent the *offset*. The maximum match that can occur is the full length of the look-ahead buffer. If the look-ahead buffer is limited to 64 bytes, this length can thus be represented by 6 bits. For text applications, the next symbol is usually represented as a byte or 8 bits. Thus, a codeword is of total fixed length of 26 bits but may represent up to 64 bytes of input text.

There are several practical situations which we have ignored in the above description of the algorithm:

- If the first character in the look-ahead buffer has not been seen before, i.e., it is not present in the search buffer, then the codeword is simply represented by an offset 0, a length 0, and the next symbol is the new character.
- The maximum-length matching sequence may occur at several positions in the search buffer. Any one of these positions may be used in the codeword. It does not

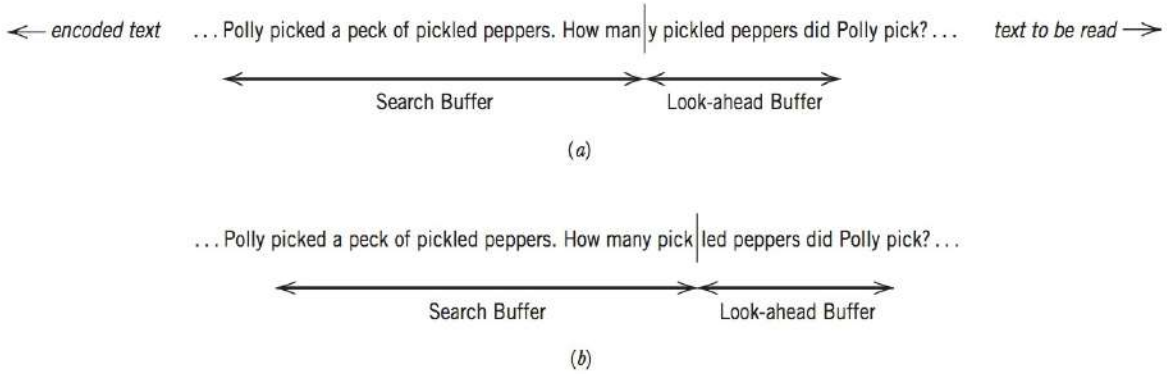


FIGURE 10.7 Illustration of search and look-ahead buffers.

matter to the decoder. Any one of them will produce the same output sequence. (Some more advanced algorithms use the first occurrence as it may be represented by fewer bits in general.)

The decoding algorithm is much simpler than the encoding algorithm, as the decoder knows exactly where to look in the decoded stream (search buffer) to find the matching string. The decoder starts with an empty (all zeros) search buffer, then:

- For each codeword received, the decoder reads the string from the search buffer of the indicated position and length and appends it to the right-hand end of the search buffer.
- The next character is then appended to the search buffer,
- The search buffer is then slid to the right so the pointer occurs immediately after the last known symbol and the process is repeated.

From the example described here, we note that, in contrast to Huffman coding, the Lempel–Ziv algorithm uses fixed length codes to represent a variable number of source symbols. If errors occur in the transmission of a data sequence that has been encoded with the Lempel–Ziv algorithm, the decoding is susceptible to error propagation. For short sequences of characters, the matching strings found in the search buffer are unlikely to be very long. In this case, the output of the Lempel–Ziv algorithm may be a "compressed" sequence, which is longer than the input sequence. The Lempel–Ziv algorithm only achieves its true advantage when processing long strings of data, for example, large files.

For a long time, Huffman coding was unchallenged as the algorithm of choice for lossless data compression. Then, the Lempel–Ziv algorithm took over almost completely from the Huffman algorithm and became the standard algorithm for file compression. In recent years, more advanced data compression algorithms have been developed building upon the ideas of Huffman, Lempel, and Ziv. Some of these techniques build adaptive, statistical models of the input text as it is processed and use this to construct codewords on an entropy-minimizing basis somewhat similar to the Huffman approach. These new approaches can more than double the compression provided by the original Lempel–Ziv algorithm in some cases, but at the expense of much increased memory and processing requirements for the encoder and decoder.

10.6 DISCRETE MEMORYLESS CHANNELS

Up to this point in the chapter, we have been preoccupied with discrete memoryless sources responsible for information generation. We next consider the issue of information transmission, with particular emphasis on reliability. We start the discussion by considering a discrete memoryless channel, the counterpart of a discrete memoryless source.

A *discrete memoryless channel* is a statistical model with an input X and an output Y that is a *noisy* version of X; both X and Y are random variables. Every unit of time, the channel accepts an input symbol X selected from an alphabet $\mathcal{X}$ and, in response, it emits an output symbol Y from an alphabet $\mathcal{Y}$. The channel is said to be discrete when both of the alphabets $\mathcal{X}$ and $\mathcal{Y}$ have *finite* sizes. It is said to be memoryless when the current output symbol depends *only* on the current input symbol and *not* any of the previous ones.

Figure 10.8 depicts a view of a discrete memoryless channel. The channel is described in terms of an *input alphabet*

$$\mathcal{X} = \{x_0, x_1, \ldots, x_{J-1}\}, \tag{10.24}$$

an *output alphabet*,

$$\mathcal{Y} = \{y_0, y_1, \ldots, y_{K-1}\}, \tag{10.25}$$

and a set of *transition probabilities*

$$p(y_k|x_j) = P(Y = y_k|X = x_j) \qquad \text{for all } j \text{ and } k \tag{10.26}$$

Naturally, we have

$$0 \le p(y_k|x_j) \le 1 \qquad \text{for all } j \text{ and } k \tag{10.27}$$

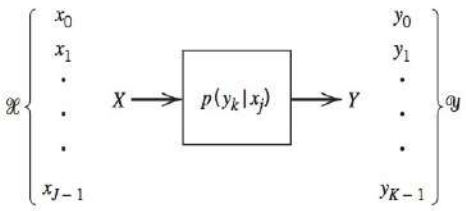


FIGURE 10.8 Discrete memoryless channel.

Also, the input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ need not have the same size. For example, in channel coding, the size K of the output alphabet $\mathcal{Y}$ may be larger than the size J of the input alphabet $\mathcal{X}$; thus, $K \geq J$. On the other hand, we may have a situation in which the channel emits the same symbol when either one of two input symbols is sent, in which case we have $K \leq J$.

A convenient way of describing a discrete memoryless channel is to arrange the various transition probabilities of the channel in the form of a matrix as follows:

$$\mathbf{P} = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) & \cdots & p(y_{K-1}|x_0) \\ p(y_0|x_1) & p(y_1|x_1) & \cdots & p(y_{K-1}|x_1) \\ \vdots & & & \vdots \\ p(y_0|x_{J-1}) & p(y_1|x_{J-1}) & & p(y_{K-1}|x_{J-1}) \end{bmatrix} \tag{10.28}$$

The J-by-K matrix $\mathbf{P}$ is called the *channel matrix*. Note that each *row* of the channel matrix $\mathbf{P}$ corresponds to a *fixed channel input*, whereas each column of the matrix corresponds to a *fixed channel output*. Note also that a fundamental property of the channel matrix $\mathbf{P}$, as defined here, is that the sum of the elements along any row of the matrix is always equal to one; that is,

$$\sum_{k=0}^{K-1} p(y_k|x_j) = 1 \qquad \text{for all } j \tag{10.29}$$

Suppose now that the inputs to a discrete memoryless channel are selected according to the *probability distribution* $\{p(x_j), j = 0, 1, \ldots, J-1\}$. In other words, the event that the channel input $X = x_j$ occurs with probability

$$p(x_j) = P(X = x_j) \qquad \text{for } j = 0, 1, \ldots, J-1 \tag{10.30}$$

Having specified the random variable X denoting the channel input, we may now specify the second random variable Y denoting the channel output. The *joint probability distribution* of the random variables X and Y is given by

$$\begin{aligned} p(x_j, y_k) &= P(X = x_j, Y = y_k) \\ &= P(Y = y_k|X = x_j)P(X = x_j) \\ &= p(y_k|x_j)p(x_j) \end{aligned} \tag{10.31}$$

The *marginal probability distribution* of the output random variable Y is obtained by averaging out the dependence of $p(x_j,y_k)$ on x_j, as shown by

$$\begin{aligned} p(y_k) &= P(Y = y_k) \\ &= \sum_{j=0}^{J-1} P(Y = y_k|X = x_j)P(X = x_j) \\ &= \sum_{j=0}^{J-1} p(y_k|x_j)p(x_j) \qquad \text{for } k = 0, 1, \ldots, K-1 \end{aligned} \tag{10.32}$$

The probabilities $p(x_j)$ for $j = 0, 1, \ldots, J-1$, are known as the *a priori probabilities* of the various input symbols. Eq. (10.32) states that if we are given the input *a priori* probabilities $p(x_j)$ and the channel matrix [i.e., the matrix of transition probabilities $p(y_k|x_j)$], then we may calculate the probabilities of the various output symbols, the $p(y_k)$.

EXAMPLE 10.5 Binary Symmetric Channel

The *binary symmetric channel* is of great theoretical interest and practical importance. It is a special case of the discrete memoryless channel with $J = K = 2$. The channel has two input symbols ($x_0 = 0, x_1 = 1$) and two output symbols ($y_0 = 0, y_1 = 1$). The channel is symmetric because the probability of receiving a 1 if a 0 is sent is the same as the probability of receiving a 0 if a 1 is sent. This conditional probability of error is denoted by p. The *transition probability diagram* of a binary symmetric channel is as shown in Figure 10.9.

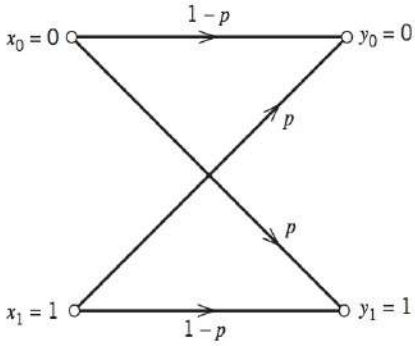


FIGURE 10.9 Transition probability diagram of binary symmetric channel.

10.7 CHANNEL CAPACITY

Of practical interest in many communication applications is the number of bits that may be reliably transmitted per second through a given communications channel. In this section we shall provide a theoretical definition of this *channel capacity*, but before doing so we must define the two concepts of *relative entropy* and *mutual information.*

Given that we think of the channel output Y (selected from alphabet $\mathcal{Y}$) as a noisy version of the channel input X (selected from alphabet $\mathcal{X}$), and that the entropy $H(\mathcal{X})$ is a measure of the prior uncertainty about X, how can we measure the uncertainty about X after observing Y? To answer this question, we extend the ideas developed in Section 10.2 by defining the *conditional entropy* of X selected from alphabet $\mathcal{X}$, given that $Y = y_k$. Specifically, we write

$$H(\mathcal{X}|Y = y_k) = \sum_{j=0}^{J-1} p(x_j|y_k) \log_2 \left[\frac{1}{p(x_j|y_k)} \right] \tag{10.33}$$

This quantity is itself a random variable that takes on the values $H(\mathcal{X}|Y = y_0), \ldots, H(\mathcal{X}|Y = y_{K-1})$ with probabilities $p(y_0), \ldots, p(y_{K-1})$, respectively. The mean of entropy $H(\mathcal{X}|Y = y_k)$ over the output alphabet $\mathcal{Y}$ is therefore given by

$$\begin{aligned} H(\mathcal{X}|\mathcal{Y}) &= \sum_{k=0}^{K-1} H(\mathcal{X}|Y = y_k) p(y_k) \\ &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j|y_k) p(y_k) \log_2 \left[\frac{1}{p(x_j|y_k)} \right] \\ &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j|y_k)} \right] \end{aligned} \tag{10.34}$$

where, in the last line, we have made use of the relation

$$p(x_i, y_k) = p(x_i|y_k) p(y_k) \tag{10.35}$$

The quantity $H(\mathcal{X}|\mathcal{Y})$ is called a *conditional entropy*. It represents *the amount of uncertainty remaining about the channel input after the channel output has been observed.*

Since the entropy $H(\mathcal{X})$ represents our uncertainty about the channel input *before* observing the channel output, and the conditional entropy $H(\mathcal{X}|\mathcal{Y})$ represents our uncertainty about the channel input *after* observing the channel output, it follows that the

difference $H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y})$ must represent our uncertainty about the channel input that is *resolved* by observing the channel output. This important quantity is called the *mutual information* of the channel. Denoting the mutual information by $I(\mathcal{X};\mathcal{Y})$, we may thus write

$$I(\mathcal{X};\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) \tag{10.36}$$

Similarly, we may write

$$I(\mathcal{Y};\mathcal{X}) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X}) \tag{10.37}$$

where $H(\mathcal{Y})$ is the entropy of the channel output and $H(\mathcal{Y}|\mathcal{X})$ is the conditional entropy of the channel output given the channel input.

Mutual information has a number of properties:

- It is nonnegative

$$I(\mathcal{X};\mathcal{Y}) \geq 0 \tag{10.38}$$

- It is symmetric

$$I(\mathcal{X};\mathcal{Y}) = I(\mathcal{Y};\mathcal{X}) \tag{10.39}$$

- By combining the expressions for $H(\mathcal{X})$ and $H(\mathcal{X}|\mathcal{Y})$ it may be shown that

$$\begin{aligned} I(\mathcal{X};\mathcal{Y}) &= \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} p(x_j,y_k)\log_2\left[\frac{p(x_j|y_k)}{p(x_j)}\right] \\ &= \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} p(x_j,y_k)\log_2\left[\frac{p(y_k|x_j)}{p(y_k)}\right] \end{aligned} \tag{10.40}$$

The relationship between the source entropy $H(\mathcal{X})$, the conditional entropy $H(\mathcal{X}|\mathcal{Y})$, and the mutual information $I(\mathcal{X};\mathcal{Y})$ is illustrated conceptually in Figure 10.10.

Consider a discrete memoryless channel with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and transition probabilities $p(y_k|x_j)$. The mutual information of the channel is given by Eq. (10.40). Here we note that [see Eq. (10.31)]

$$p(x_j,y_k) = p(y_k|x_j)p(x_j) \tag{10.41}$$

Also, from Eq. (10.32), we have

$$p(y_k) = \sum_{j=0}^{J-1} p(y_k|x_j)p(x_j) \tag{10.42}$$

From Eqs. (10.40), (10.41), and (10.42) we see that it is necessary for us to know the input probability distribution $\{p(x_j)|\, j = 0, 1, \ldots, J-1\}$ so that we may calculate the mutual information $I(\mathcal{X};\mathcal{Y})$. The mutual information of a channel therefore depends not only on the channel but also on the way in which the channel is used.

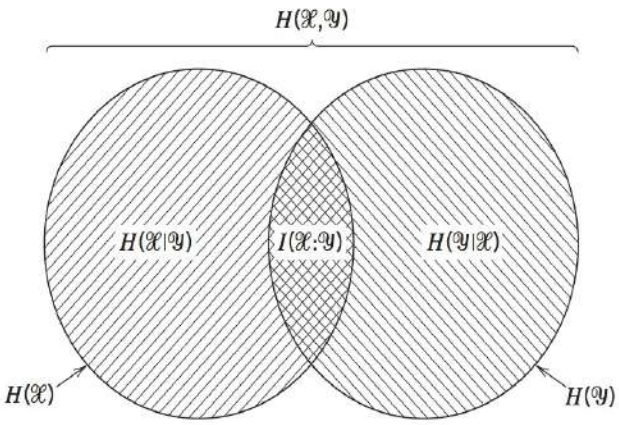


FIGURE 10.10 Illustrating the relations among various channel parameters.

The input probability distribution $\{p(x_j)\}$ is obviously independent of the channel. We can then maximize the average mutual information $I(\mathcal{X};\mathcal{Y})$ of the channel with respect to $\{p(x_j)\}$. Hence, *we define the channel capacity of a discrete memoryless channel as the maximum average mutual information $I(\mathcal{X};\mathcal{Y})$ in any single use of the channel (i.e., signaling interval), where the maximization is over all possible input probability distributions $\{p(x_j)\}$ on $\mathcal{X}$.* The channel capacity is commonly denoted by C. We thus write

$$C = \max_{\{p(x_j)\}} I(\mathcal{X};\mathcal{Y}) \tag{10.43}$$

The channel capacity C is measured in *bits per channel use.*

Note that the channel capacity C is a function only of the transition probabilities $p(y_k|x_j)$, which define the channel. The calculation of C involves maximization of the average mutual information $I(\mathscr{X};\mathscr{Y})$ over J variables [i.e., the input probabilities $p(x_0), \ldots, p(x_{J-1})$] subject to two constraints:

$$p(x_j) \geq 0 \text{ for all } j$$

and

$$\sum_{j=0}^{J-1} p(x_j) = 1$$

In general, the variational problem of finding the channel capacity C is a challenging task.

EXAMPLE 10.6 Binary Symmetric Channel (Revisited)

Consider again the *binary symmetric channel*, which is described by the *transition probability diagram* of Figure 10.9. This diagram is uniquely defined by the conditional probability of error p.

The entropy $H(X)$ is maximized when the channel input probability $p(x_0) = p(x_1) = 1/2$, where x_0 and x_1 are each 0 or 1. The mutual information $I(\mathscr{X};\mathscr{Y})$ is similarly maximized, so that we may write

$$C = I(\mathscr{X};\mathscr{Y})_{|p(x_0)=p(x_1)=\frac{1}{2}}$$

From Figure 10.9, we have

$$p(y_0|x_1) = p(y_1|x_0) = p$$

and

$$p(y_0|x_0) = p(y_1|x_1) = 1 - p$$

Therefore, substituting these channel transition probabilities into Eq. (10.40) with $J = K = 2$, and then setting the input probability $p(x_0) = p(x_1)$ in accordance with Eq. (10.43), we find that the capacity of the binary symmetric channel is

$$C = 1 + p\log_2 p + (1-p)\log_2(1-p) \tag{10.44}$$

Using the definition of the entropy function given in Eq. (10.12), we may reduce Eq. (10.44) to

$$C = 1 - H(p)$$

The channel capacity C varies with the probability of error (transition probability) p as shown in Figure 10.11, which is symmetric about $p = 1/2$. Comparing the curve in this figure with that in Figure 10.2, we may make the following observations:

1. When the channel is *noise free,* permitting us to set $p = 0$, the channel capacity C attains its maximum value of one bit per channel use, which is exactly the information in each channel input. At this value of p, the entropy function $H(p)$ attains its minimum value of zero.
2. When the conditional probability of error $p = 1/2$ due to noise, the channel capacity C attains its minimum value of zero, whereas the entropy function $H(p)$ attains its maximum value of unity; in such a case the channel is said to be *useless*.

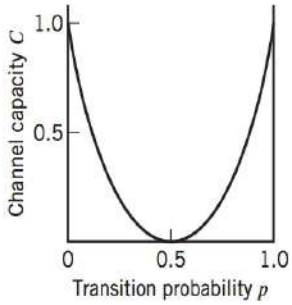


FIGURE 10.11 Variation of channel capacity of a binary symmetric channel with transition probability p.

10.8 CHANNEL CODING THEOREM

The inevitable presence of *noise* in a channel causes discrepancies (errors) between the output and input data sequences of a digital communication system. For a relatively noisy channel, the probability of error may have a value higher than 10^{-2}, which means that less than 99 out of 100 transmitted bits are received correctly. For many applications, this *level of reliability* is found to be far from adequate. Indeed, a probability of error equal to 10^{-6} or even lower is often a necessary requirement. To achieve such a high level of performance, we may have to resort to the use of channel coding.

The design goal of channel coding is to increase the resistance of a digital communication system to channel noise. Specifically, *channel coding* consists of *mapping* the incoming data sequence into a channel input sequence, and *inverse mapping* the channel output sequence into an output data sequence in such a way that the overall effect of channel noise on the system is minimized. The first mapping operation is performed in the transmitter by a *channel encoder*, whereas the inverse mapping operation is performed in the receiver by a *channel decoder*, as shown in the block diagram of Figure 10.12; to simplify the exposition, we have not included source encoding (before channel encoding) and source decoding (after channel decoding) in Figure 10.12.

The channel encoder and channel decoder in Figure 10.12 are both under the designer's control and should be designed to optimize the overall effectiveness of the communication system. The approach taken is to introduce *redundancy* in the channel encoder so as to reconstruct the original source sequence as accurately as possible. Thus, in a rather loose sense, we may view channel coding as the *dual* of source coding in that the former introduces controlled redundancy to improve reliability, whereas the latter reduces redundancy to improve efficiency.

For the purpose of our present discussion of channel coding, it suffices to confine our attention to *block codes*. In this class of codes, the message sequence is subdivided into sequential blocks each k bits long, and each k-bit block is *mapped* into an n-bit block, where $n > k$. The number of redundant bits added by the encoder to each transmitted block is $n - k$ bits. The ratio k/n is called the *code rate*. Using r to denote the code rate, we may thus write

$$r = \frac{k}{n}$$

where, of course, r is less than unity.

The accurate reconstruction of the original source sequence at the destination requires that the *average probability of symbol error* be arbitrarily low. This raises the following important question: Does there exist a sophisticated channel coding scheme such that the probability that a message bit will be in error is less than any positive number ε (i.e., as small as we want it), and yet the channel coding scheme is efficient in that the code rate need not be too small? The answer to this fundamental question is an emphatic "yes." Indeed, the answer to the question is provided by Shannon's second theorem in

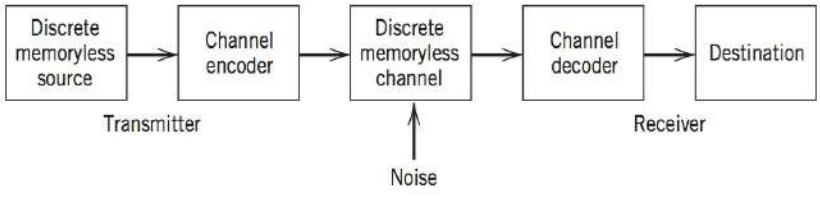


FIGURE 10.12 Block diagram of digital communication system.

terms of the channel capacity C, as described in what follows. Up until this point, *time* has not played an important role in our discussion of channel capacity. Suppose then the discrete memoryless source in Figure 10.12 has the source alphabet $\mathcal{S}$ and entropy $H(\mathcal{S})$ bits per source symbol. We assume that the source emits symbols once every T_s seconds. Hence, the *average information rate* of the source is $H(\mathcal{S})/T_s$ bits per second. The decoder delivers decoded symbols to the destination from the source alphabet $\mathcal{S}$ and at the same source rate of one symbol every T_s seconds. The discrete memoryless channel has a channel capacity equal to C bits per use of the channel. We assume that the channel is capable of being used once every T_c seconds. Hence, the *channel capacity per unit time* is C/T_c bits per second, which represents the maximum rate of information transfer over the channel. We are now ready to state Shannon's second theorem, known as the channel coding theorem.

Specifically, the *channel coding theorem* for a discrete memoryless channel is stated in two parts as follows.

(a) *Let a discrete memoryless source with an alphabet $\mathcal{S}$ have entropy $H(\mathcal{S})$ and produce symbols once every T_s seconds. Let a discrete memoryless channel have capacity C and be used once every T_c seconds. Then, if*

$$\frac{H(\mathcal{S})}{T_s} \leq \frac{C}{T_c} \tag{10.45}$$

there exists a coding scheme for which the source output can be transmitted over the channel and be reconstructed with an arbitrarily small probability of error. The parameter C/T_c is called the critical rate. *When Eq. (10.45) is satisfied with the equality sign, the system is said to be signaling at the critical rate.*

(b) *Conversely, if*

$$\frac{H(\mathcal{S})}{T_s} > \frac{C}{T_c}$$

it is not possible to transmit information over the channel and reconstruct it with an arbitrarily small probability of error.

The channel coding theorem is the single most important result of information theory. The theorem specifies the channel capacity C as a *fundamental limit* on the rate at which the transmission of reliable error-free messages can take place over a discrete memoryless channel.

It is important to note that the channel coding theorem does not show us how to construct a good code. Rather, the theorem can be characterized as an *existence proof* in the sense that it tells us that if the condition of Eq. (10.45) is satisfied, then good codes do exist.

APPLICATION OF THE CHANNEL CODING THEOREM TO BINARY SYMMETRIC CHANNELS

Consider a discrete memoryless source that emits equally likely binary symbols (0s and 1s) once every T_s seconds. With the source entropy equal to one bit per source symbol (see Example 10.1), the information rate of the source is $(1/T_s)$ bits per second. The source sequence is applied to a channel encoder with *code rate r*. The channel encoder produces a symbol once every T_c seconds. Hence, the *encoded symbol transmission rate* is $(1/T_c)$ symbols per second. The channel encoder engages a binary symmetric channel once every T_c seconds. Hence, the channel capacity per unit time is (C/T_c) bits per second, where C is determined by the prescribed channel transition probability p in accordance with Eq. (10.44). Accordingly, the channel coding theorem [part (i)] implies that if

$$\frac{1}{T_s} \leq \frac{C}{T_c} \tag{10.46}$$

the probability of error can be made arbitrarily low by the use of a suitable channel encoding scheme. But the ratio T_c/T_s equals the code rate of the channel encoder:

$$r = \frac{T_c}{T_s} \tag{10.47}$$

Hence, we may restate the condition of Eq. (10.46) as

$$r \leq C \tag{10.48}$$

That is, for $r \leq C$, there exists a code (with code rate less than or equal to C) capable of achieving an arbitrarily low probability of error.

EXAMPLE 10.7 Repetition Code

In this example, we present a graphical interpretation of the channel coding theorem. We also bring out a surprising aspect of the theorem by taking a look at a simple coding scheme.

Consider first a binary symmetric channel with transition probability $p = 10^{-2}$. For this value of p, we find from Eq. (10.44) that the channel capacity $C = 0.9192$. Hence, from the channel coding theorem, we may state that for any $\varepsilon > 0$ and $r \leq 0.9192$, there exists a code of large enough length n and code rate r, and an appropriate decoding algorithm, such that when the coded bit stream is sent over the given channel, the average probability of channel decoding error is less than ε. This result is illustrated in Figure 10.13, where we have plotted the average probability of error versus the code rate r. In this figure, we have arbitrarily set the limiting value $\varepsilon = 10^{-8}$.

To put the significance of this result in perspective, consider next a simple coding scheme that involves the use of a *repetition code*, in which each bit of the message is repeated several times. Let each bit (0 or 1) be repeated n times, where $n = 2m + 1$ is an odd integer. For example, for $n = 3$, we transmit 0 and 1 as 000 and 111, respectively. Intuitively, it would seem logical to use a *majority rule* for decoding, which operates as follows: *If in a block of n received bits (representing one bit of the message), the number of 0s exceeds the number of 1s, the decoder decides in favor of a 0. Otherwise, it decides in favor of a 1.* Hence, an error occurs when $m + 1$ or more bits out of $n = 2m + 1$ bits are received incorrectly. Because of the assumed symmetric nature of the channel, the *average probability of error* P_e is independent of the *a priori* probabilities of 0 and 1. Accordingly, we find that P_e is given by (see Problem 10.24)

$$P_e = \sum_{i=m+1}^{n} \binom{n}{i} p^i (1-p)^{n-i} \tag{10.49}$$

where p is the transition probability of the channel.

Table 10.3 gives the average probability of error P_e for a repetition code, which is calculated by using Eq. (10.49) for different values of the code rate r. The values given here assume the use of a binary symmetric channel with transition probability

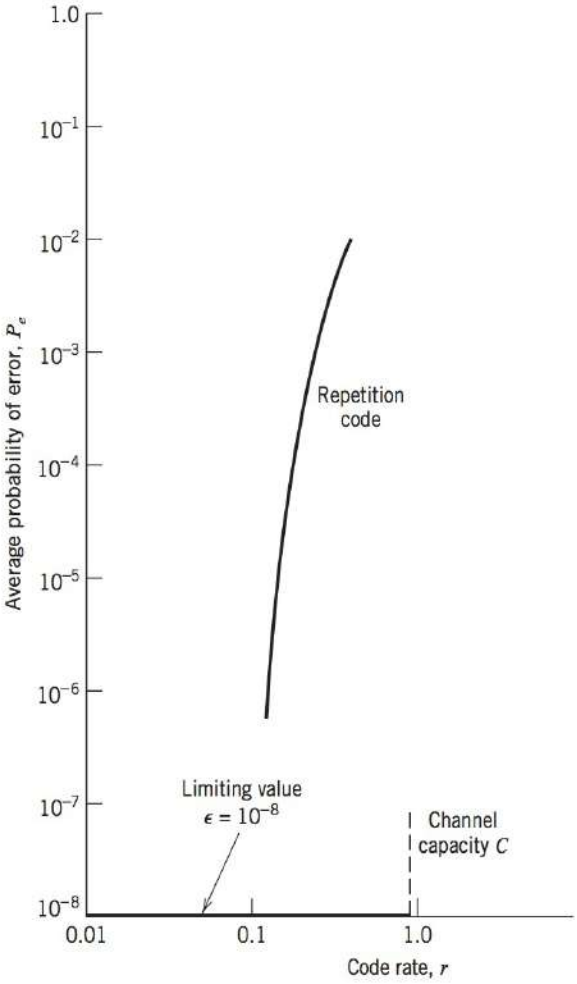


FIGURE 10.13 Illustrating significance of the channel coding theorem.

$p = 10^{-2}$. The improvement in reliability displayed in Table 10.3 is achieved at the cost of decreasing code rate. The results of this table are also shown plotted as the curve labeled repetition code in Figure 10.13. This curve illustrates the *exchange of code rate for message reliability*, which is a characteristic of repetition codes.

This example highlights the unexpected result presented to us by the channel coding theorem. The result is that it is not necessary to have the code rate r approach zero (as in the case of repetition codes) so as to achieve more and more reliable operation of the communication link. The theorem merely requires that the code rate be less than the channel capacity C.

TABLE 10.3 Average probability of error for repetition code

Code Rate $r = 1/n$	Average Probability of Error, P_e
1	10^{-2}
$\frac{1}{3}$	3×10^{-4}
$\frac{1}{5}$	10^{-6}
$\frac{1}{7}$	4×10^{-7}
$\frac{1}{9}$	10^{-8}
$\frac{1}{11}$	5×10^{-10}

10.9 CAPACITY OF A GAUSSIAN CHANNEL

In this section, we use the idea of average mutual information to formulate the information capacity theorem for *band-limited, power-limited Gaussian channels*. To be specific, consider a zero-mean stationary process $X(t)$ that is band-limited to B Hertz. Let X_k, $k = 1, 2, \ldots, K$, denote the continuous random variables obtained by uniform sampling of the process $X(t)$ at the Nyquist rate of $2B$ samples per second. These samples are transmitted in T seconds over a noisy channel, also band-limited to B Hertz. Hence, the number of samples, K, is given by

$$K = 2BT \tag{10.50}$$

We refer to X_k as a sample of the *transmitted signal*. The channel output is perturbed by *additive white Gaussian noise* of zero mean and power spectral density $N_0/2$. The noise is band-limited to B Hertz. Let the continuous random variables Y_k, $k = 1, 2, \ldots, K$ denote samples of the received signal, as shown by

$$Y_k = X_k + N_k, \qquad k = 1, 2, \ldots, K \tag{10.51}$$

The noise sample N_k is Gaussian with zero mean and variance given by

$$\sigma^2 = N_0 B \tag{10.52}$$

We assume that the samples Y_k, $k = 1, 2, \ldots, K$ are statistically independent.

A channel for which the noise and the received signal are as described in Eqs. (10.51) and (10.52) is called a *discrete-time, memoryless Gaussian channel*. It is modeled as in Figure 10.14. To make meaningful statements about the channel, however, we have to constrain the channel input. Typically, the transmitter is *power limited*; it is therefore reasonable to define the constraint as

$$\mathbf{E}[X_k^2] = P, \qquad k = 1, 2, \ldots, K \tag{10.53}$$

where P is the *average transmitted power*. The *power-limited Gaussian channel* described herein is of not only theoretical but also practical importance in that it models many communication channels, including radio and satellite links.

The *information capacity* of the channel is defined as the maximum of the mutual information between the channel input X_k and the channel output Y_k over all distributions on the input X_k that satisfy the power constraint of Eq. (10.53). Let $I(X_k;Y_k)$ denote the average mutual information between X_k and Y_k. We may then define the information capacity of the channel as

$$C = \max_{f_{X_k}(x)} \{I(X_k;Y_k): \ \mathbf{E}[X_k^2] = P\} \tag{10.54}$$

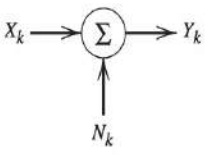


FIGURE 10.14 Model of discrete-time, memoryless Gaussian channel.

where the maximization is performed with respect to $f_{X_k}(x)$, the probability density function of X_k.

Performing this optimization is beyond the scope of this text but the result is

$$C = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}\right) \text{bits per use} \tag{10.55}$$

With the channel used K times for the transmission of K samples of the process $X(t)$ in T seconds, we find that the information *capacity per unit time* is (K/T) times the result given in Eq. (10.55). The number K equals $2BT$, as in Eq. (10.50). Accordingly, we may express the information capacity per transmission as

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right) \text{bits per second} \tag{10.56}$$

where we have used Eq. (10.52) for the noise variance σ^2.

Based on the formula of Eq. (10.56), we may now state Shannon' s third (and most famous) theorem, the *information capacity theorem*, as follows:

The information capacity of a continuous channel of bandwidth B Hertz, perturbed by additive white Gaussian noise of power spectral density $N_0/2$ and limited in bandwidth to B, is given by

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right) \text{bits per second}$$

where P is the average transmitted power.

The information capacity theorem is one of the most remarkable results of information theory for, in a single formula, it highlights most vividly the interplay among three key system parameters: channel bandwidth, average transmitted power (or, equivalently, average received signal power), and noise power spectral density at the channel output.

The theorem implies that, for given average transmitted power P and channel bandwidth B, we can transmit information at the rate C bits per second, as defined in Eq. (10.56), with arbitrarily small probability of error by employing sufficiently complex encoding systems. It is not possible to transmit at a rate higher than C bits per second by any encoding system without a definite probability of error. Hence, the channel capacity theorem defines the *fundamental limit* on the rate of error-free transmission for a power-limited, band-limited Gaussian channel. To approach this limit, however, the transmitted signal must have statistical properties approximating those of white Gaussian noise.

Now that we have an intuitive feel for the information capacity theorem, we may go on to discuss its implications in the context of a Gaussian channel that is limited in both power and bandwidth. For the discussion to be useful, however, we need an ideal framework against which the performance of a practical communication system can be assessed. To this end, we introduce the notion of an *ideal system* defined as one that transmits data at a bit rate R_b equal to the information capacity C. We may then express the average transmitted power as

$$P = E_b C \tag{10.57}$$

where E_b is the transmitted energy per bit. Accordingly, the ideal system is defined by the equation

$$\frac{C}{B} = \log_2\left(1 + \frac{E_b}{N_0}\frac{C}{B}\right) \tag{10.58}$$

Equivalently, we may define the *signal energy-per-bit to noise power spectral density ratio* E_b/N_0 in terms of the *bandwidth efficiency* C/B for the ideal system as

$$\frac{E_b}{N_0} = \frac{2^{C/B} - 1}{C/B} \tag{10.59}$$

A plot of bandwidth efficiency R_b/B versus E_b/N_0 is called the *bandwidth-efficiency diagram*. A generic form of this diagram is displayed in Figure 10.15, where the curve labeled capacity boundary corresponds to the ideal system for which $R_b = C$. Based on Figure 10.15, we can make the following observations:

1. For *infinite bandwidth*, the ratio E_b/N_0 approaches the limiting value

$$\left(\frac{E_b}{N_0}\right)_\infty = \lim_{B\to\infty}\left(\frac{E_b}{N_0}\right) = \ln 2 = 0.693 \tag{10.60}$$

This value is called the *Shannon limit*. Expressed in decibels, it equals −1.6 dB. The corresponding limiting value of the channel capacity is obtained by letting the channel bandwidth B in Eq. (10.56) approach infinity; we thus find that

$$C_\infty = \lim_{B\to\infty} C = \frac{P}{N_0}\log_2 e \tag{10.61}$$

2. The *capacity boundary*, defined by the curve for the critical bit rate $R_b = C$, separates combinations of system parameters that have the potential for supporting error-free transmission ($R_b < C$) from those for which error-free transmission is not possible ($R_b > C$). The latter region is shown shaded in Figure 10.15.

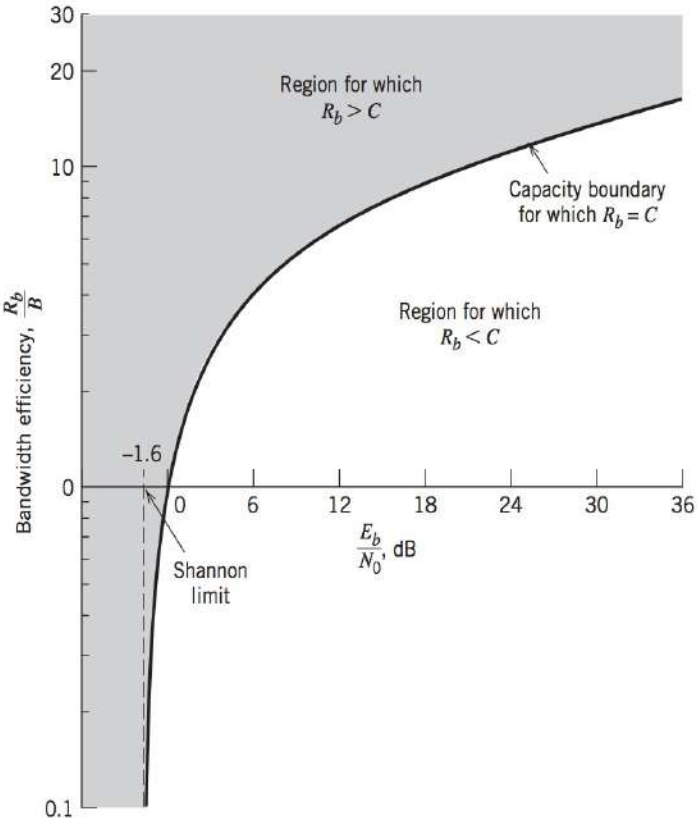


FIGURE 10.15 Bandwidth-efficiency diagram.

3. The diagram highlights potential *trade-offs* among E_b/N_0, R_b/B, and probability of symbol error P_e. In particular, we may view movement of the operating point along a horizontal line as trading P_e versus E_b/N_0 for a fixed R_b/B. On the other hand, we may view movement of the operating point along a vertical line as trading P_e versus R_b/B for a fixed E_b/N_0.

EXAMPLE 10.8 *M*-ary PSK and *M*-ary FSK

In this example, we compare the bandwidth-power exchange capabilities of M-ary PSK and M-ary FSK signals in light of Shannon's information capacity theorem. Consider first a coherent M-ary PSK system that employs a *nonorthogonal* set of M phase-shifted signals for the transmission of binary data. Each signal in the set represents a symbol with $\log_2 M$ bits. Using the definition of null-to-null bandwidth, we may express the bandwidth efficiency of M-ary PSK as follows

$$\frac{R_b}{B} = \frac{\log_2 M}{2}$$

In Figure 10.16*a*, we show the operating points for different numbers of phase levels $M = 2, 4, 8, 16, 32, 64$. Each point corresponds to an average probability of symbol error $P_e = 10^{-5}$. In the figure we have also included the capacity boundary for the ideal system. We observe from Figure 10.16 that as M is

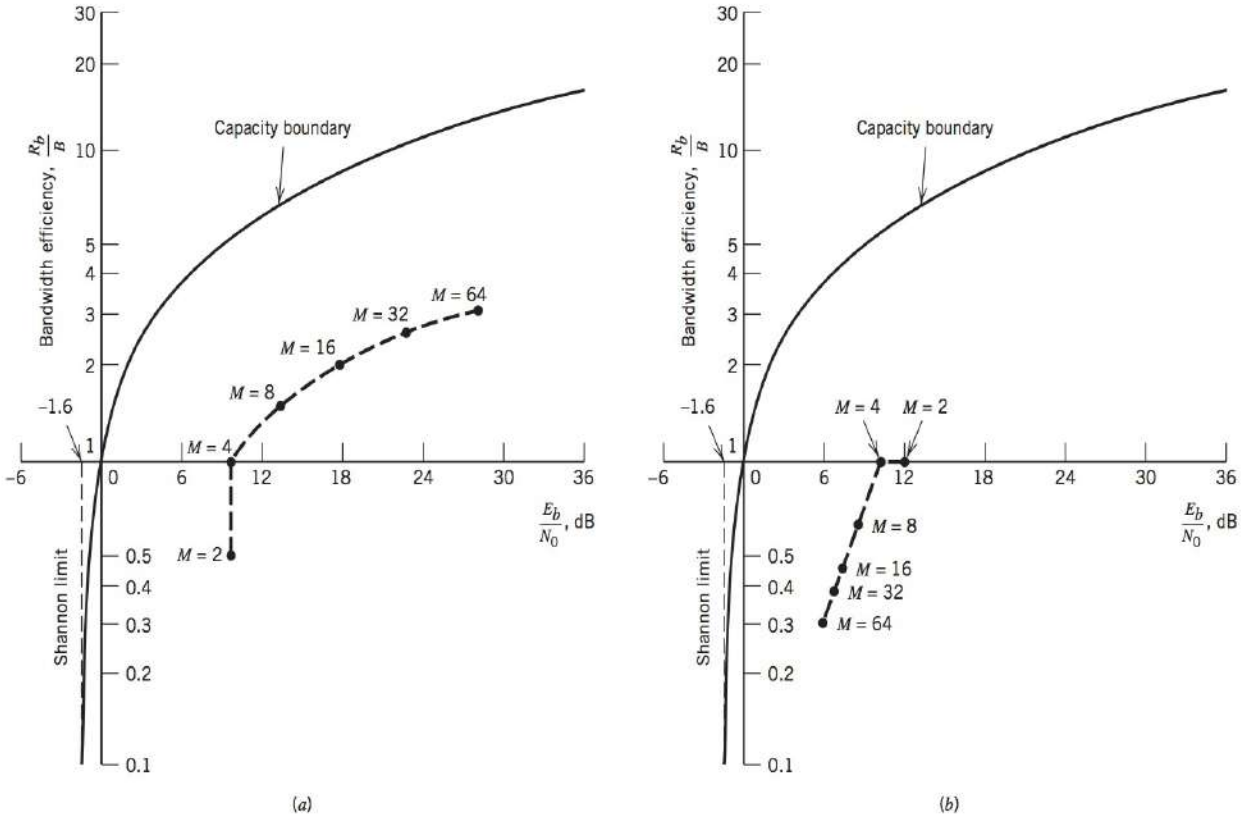


FIGURE 10.16 (*a*) Comparison of *M*-ary PSK with the ideal system for $P_e = 10^{-5}$. (*b*) Comparison of *M*-ary FSK with the ideal system for $P_e = 10^{-5}$.

increased, the bandwidth efficiency is improved, but the value of E_b/N_0 required for error-free transmission moves away from the Shannon limit.

Consider next a coherent *M*-ary FSK system that uses an *orthogonal* set of *M* frequency-shifted signals for the transmission of binary data, with the separation between adjacent signal frequencies set at $B = 1/2T$, where T is the symbol period. The bandwidth of an *M*-FSK system is proportional to *M*, the number of frequencies. As with the *M*-ary PSK, each signal in the set represents a symbol with $\log_2 M$ bits. The bandwidth efficiency of *M*-ary FSK is as follows

$$\frac{R_b}{B} = \frac{2 \log_2 M}{M}$$

In Figure 10.16*b*, we show the operating points for different numbers of frequency levels $M = 2, 4, 8, 16, 32, 64$ for an average probability of symbol error $P_e = 10^{-5}$. In the figure, we have also included the capacity boundary for the ideal system. We see that increasing *M* in (orthogonal) *M*-ary FSK has the opposite effect to that in (nonorthogonal) *M*-ary PSK. In particular, as *M* is increased, which is equivalent to increased bandwidth requirement, the operating point moves closer to the Shannon limit. Now that we understand the channel coding theorem and its implications, the stage is set for the study of error-control coding techniques, which occupy our attention for the rest of the chapter.

10.10 ERROR CONTROL CODING[4]

The *channel coding theorem* states that if a discrete memoryless channel has capacity *C* and a source generates information at a rate less than *C*, then there exists a coding

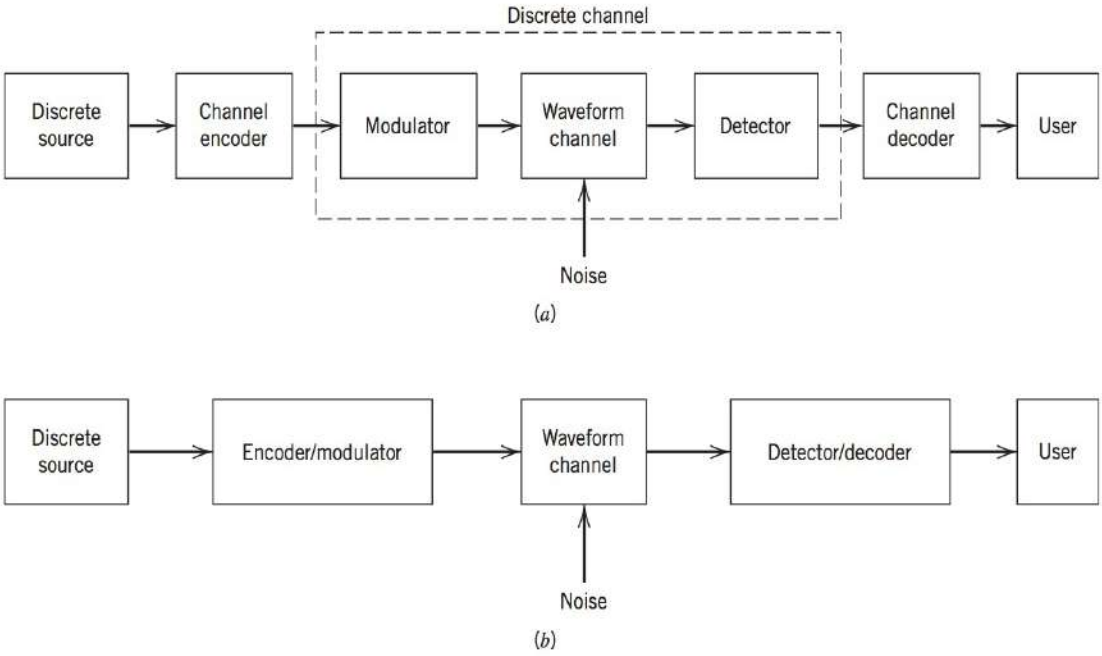


FIGURE 10.17 Simplified models of digital communication system. (*a*) Coding and modulation performed separately. (*b*) Coding and modulation combined.

technique such that the output of the source may be transmitted over the channel with an arbitrarily low probability of error.

The channel coding theorem thus specifies the channel capacity C as a *fundamental limit* on the rate at which the transmission of reliable (error-free) messages can take place over a discrete memoryless channel. The issue that matters is not the signal-to-noise ratio, so long as it is large enough, but how the channel input is encoded.

Figure 10.17*a* shows one model of how this encoding (and corresponding decoding) could be included in a digital communication system, an approach known as *forward error correction* (FEC).

The discrete source generates information in the form of binary symbols. The *channel encoder* in the transmitter accepts message bits and adds *redundancy* according to a prescribed rule, thereby producing encoded data at a higher bit rate. The *channel decoder* in the receiver exploits the redundancy to decide which message bits were actually transmitted. The combined goal of the channel encoder and decoder is to minimize the effect of channel noise. That is, the number of errors between the channel encoder input (derived from the source) and the channel decoder output (delivered to the user) is minimized.

The addition of redundancy in the coded messages implies the need for increased transmission bandwidth. Moreover, the use of error-control coding adds *complexity* to the system, especially for the implementation of decoding operations in the receiver. Thus, the design trade-offs in the use of error-control coding to achieve acceptable error performance include considerations of bandwidth and system complexity.

In the model depicted in Figure 10.17*a*, the operations of channel coding and modulation are performed separately. When, however, bandwidth efficiency is of major concern, the most effective method of implementing forward error-control correction coding is to combine it with modulation as a single function, as shown in Figure 10.17*b*. In such an approach, coding is redefined as a process of imposing certain patterns on the transmitted signal.

The most unsatisfactory feature of the channel coding theorem, however, is its nonconstructive nature. The theorem asserts the *existence of good codes* but does not tell us how to find them. We are still faced with the task of finding a good code that ensures reliable transmission of information over the channel. The error-control coding

techniques presented in this chapter provide different methods of achieving this important system requirement.

Returning to the model of Figure 10.17*a*, the waveform channel is said to be memoryless if the detector output in a given interval depends only on the signal transmitted in that interval, and not on any previous transmission. Under this condition, we may model the combination of the modulator, the waveform channel, and the detector as a *discrete memoryless channel*. The discrete memoryless channel is completely described by the set of transition probabilities $p(j|i)$, where i denotes a modulator input symbol, j denotes a demodulator output symbol, and $p(j|i)$ denotes the probability of receiving symbol j, given that symbol i was sent. (Discrete memoryless channels were described previously at some length in Section 10.6.)

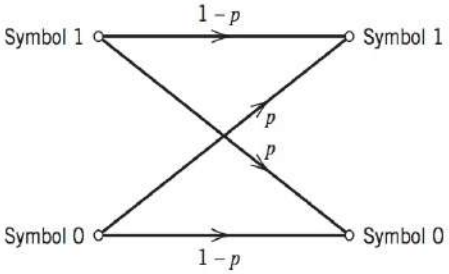


FIGURE 10.18 Transition probability diagram of binary symmetric channel.

The simplest discrete memoryless channel results from the use of binary input and binary output symbols. When binary coding is used, the modulator has only the binary symbols 0 and 1 as inputs. Likewise, the decoder has only binary inputs if binary quantization of the demodulator output is used, that is, a *hard decision* is made on the demodulator output as to which symbol was actually transmitted. In this situation, we have a *binary symmetric channel* (BSC) with a *transition probability diagram* as shown in Figure 10.18. The binary symmetric channel, assuming a channel noise modeled as additive white Gaussian noise (AWGN) channel, is completely described by the *transition probability p*.

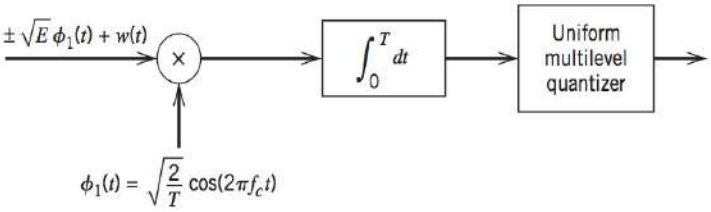


FIGURE 10.19 Binary input Q-ary output discrete memoryless channel.

The use of hard decisions prior to decoding causes an irreversible loss of information in the receiver. To reduce this loss, *soft-decision* coding is used. This is achieved by including a multilevel quantizer at the demodulator output, as illustrated in Figure 10.19 for the case of binary PSK signals. The input–output characteristic of the quantizer is shown in Figure 10.20*a*. The modulator has only the binary symbols 0 and 1 as inputs, but the demodulator output now has an alphabet with Q symbols. Assuming the use of the quantizer as described in Figure 10.20*a*, we have $Q = 8$. Such a channel is called a *binary input Q-ary output discrete memoryless channel*. The corresponding channel transition probability diagram is shown in Figure 10.20*b*. The form of this distribution, and consequently the decoder performance, depends on the location of the representation levels of the quantizer, which in turn depends on the signal level and noise variance. Accordingly, the demodulator must incorporate automatic gain control if an effective multilevel quantizer is to be realized. Moreover, the use of soft decisions complicates the implementation of the decoder. Nevertheless, soft-decision decoding offers significant improvement in performance over hard-decision decoding.

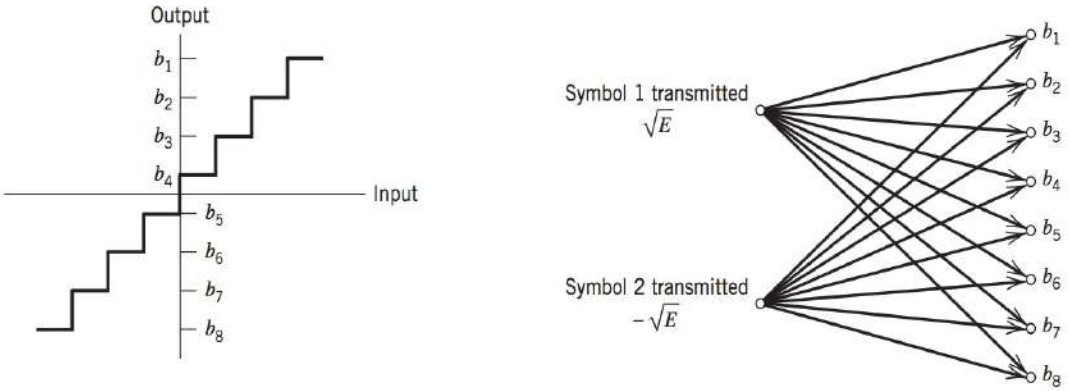


FIGURE 10.20 (*a*) Transfer characteristic of multilevel quantizer. (*b*) Channel transition probability diagram.

10.11 LINEAR BLOCK CODES

A code is said to be *linear* if any two codewords in the code can be added in modulo-2 arithmetic to produce a third codeword in the code. Consider then an (n, k) linear block code in which k bits of the n code bits are always identical to the message sequence to be transmitted. The $n - k$ bits in the remaining portion are computed from the message bits in accordance with a prescribed encoding rule that determines the mathematical structure of the code. Accordingly, these $n - k$ bits are referred to as *generalized parity check bits* or simply *parity bits*. Block codes in which the message bits are transmitted in unaltered form are called *systematic codes*. For applications requiring *both* error detection and error correction, the use of systematic block codes simplifies implementation of the decoder.

Let $m_0, m_1, \ldots, m_{k-1}$ constitute a block of k arbitrary message bits. Thus we have 2^k distinct message blocks. Let this sequence of message bits be applied to a linear block encoder, producing an n-bit codeword whose elements are denoted by $c_0, c_1, \ldots, c_{n-1}$. Let $b_0, b_1, \ldots, b_{n-k-1}$ denote the $(n - k)$ parity bits in the codeword. For the code to possess a systematic structure, a codeword is divided into two parts, one of which is occupied by the message bits and the other by the parity bits. Clearly, we have the option of sending the message bits of a codeword before the parity bits, or vice versa. The former option is illustrated in Figure 10.21, and its use is assumed in the sequel.

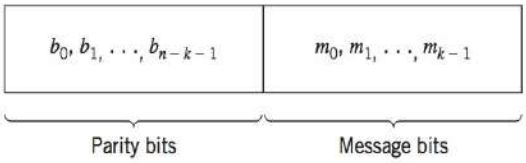


FIGURE 10.21 Structure of codeword.

According to the representation of Figure 10.21, the $(n–k)$ left-most bits of a codeword are identical to the corresponding parity bits, and the k right-most bits of the codeword are identical to the corresponding message bits. We may therefore write

$$c_i = \begin{cases} b_i, & i = 0, 1, \ldots, n-k-1 \\ m_{i+k-n}, & i = n-k, n-k+1, \ldots, n-1 \end{cases} \tag{10.62}$$

The $(n - k)$ parity bits are *linear sums* of the k message bits, as shown by the generalized relation where + refers to modulo-2 addition

$$b_i = p_{0i}\, m_0 + p_{1i}\, m_1 + \cdots + p_{k-1,i}\, m_{k-1} \tag{10.63}$$

where the coefficients are defined as follows:

$$p_{ij} = \begin{cases} 1 & \text{if } b_i \text{ depends on } m_j \\ 0 & \text{otherwise} \end{cases} \tag{10.64}$$

The coefficients p_{ij} are chosen in such a way that the rows of the generator matrix are linearly independent and the parity equations are *unique*.

The system of Eqs. (10.62) and (10.63) defines the mathematical structure of the (n, k) linear block code. This system of equations may be rewritten in a compact form using matrix notation. To proceed with this reformulation, we define the 1-by-k *message vector* $\mathbf{m}$, the 1-by-$(n - k)$ parity vector $\mathbf{b}$, and the 1-by-n code vector $\mathbf{c}$ as follows:

$$\mathbf{m} = [m_0, m_1, \ldots, m_{k-1}] \tag{10.65}$$

$$\mathbf{b} = [b_0, b_1, \ldots, b_{n-k-1}] \tag{10.66}$$

$$\mathbf{c} = [c_0, c_1, \ldots, c_{n-1}] \tag{10.67}$$

Note that all three vectors are *row vectors*. The use of row vectors is adopted in this chapter for the sake of being consistent with the notation commonly used in the coding literature. We may thus rewrite the set of simultaneous equations defining the parity bits in the compact matrix form:

$$\mathbf{b} = \mathbf{mP} \tag{10.68}$$

where $\mathbf{P}$ is the k-by-$(n-k)$ *coefficient matrix* defined by

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0,n-k-1} \\ p_{10} & p_{11} & \cdots & p_{1,n-k-1} \\ \vdots & \vdots & & \vdots \\ p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} \end{bmatrix} \tag{10.69}$$

where p_{ij} is 0 or 1.

From the definitions given in Eqs. (10.65) – (10.66), we see that $\mathbf{c}$ may be expressed as a partitioned row vector in terms of the vectors $\mathbf{m}$ and $\mathbf{b}$ as follows:

$$\mathbf{c} = [\mathbf{b} \mid \mathbf{m}] \tag{10.70}$$

Hence, substituting Eq. (10.68) in Eq. (10.70) and factoring out the common message vector $\mathbf{m}$, we get

$$\mathbf{c} = \mathbf{m}[\mathbf{P} \mid \mathbf{I}_k] \tag{10.71}$$

where $\mathbf{I}_k$ is the k-by-k *identity matrix*:

$$\mathbf{I}_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{10.72}$$

Define the k-by-n *generator matrix*

$$\mathbf{G} = [\mathbf{P} \mid \mathbf{I}_k] \tag{10.73}$$

The generator matrix $\mathbf{G}$ of Eq. (10.73) is said to be in the *echelon canonical form* in that its k rows are linearly independent; that is, it is not possible to express any row of the matrix $\mathbf{G}$ as a linear combination of the remaining rows. Using the definition of the generator matrix $\mathbf{G}$, we may simplify Eq. (10.71) as

$$\mathbf{c} = \mathbf{mG} \tag{10.74}$$

The full set of codewords, referred to simply as *the code*, is generated in accordance with Eq. (10.74) by letting the message vector $\mathbf{m}$ range through the set of all 2^k binary k-tuples (1-by-k vectors). Moreover, the sum of any two codewords is another codeword. This basic property of linear block codes is called *closure*. To prove its validity, consider a pair of code vectors $\mathbf{c}_i$ and $\mathbf{c}_j$ corresponding to a pair of message vectors $\mathbf{m}_i$ and $\mathbf{m}_j$, respectively. Using Eq. (10.74) we may express the sum of $\mathbf{c}_i$ and $\mathbf{c}_j$ as

$$\begin{aligned} \mathbf{c}_i + \mathbf{c}_j &= \mathbf{m}_i\mathbf{G} + \mathbf{m}_j\mathbf{G} \\ &= (\mathbf{m}_i + \mathbf{m}_j)\mathbf{G} \end{aligned}$$

The modulo-2 sum of $\mathbf{m}_i$ and $\mathbf{m}_j$ represents a new message vector. Correspondingly, the modulo-2 sum of $\mathbf{c}_i$ and $\mathbf{c}_j$ represents a new code vector.

There is another way of expressing the relationship between the message bits and parity-check bits of a linear block code. Let $\mathbf{H}$ denote an $(n-k)$-by-n matrix, defined as

$$\mathbf{H} = \left[\mathbf{I}_{n-k} | \mathbf{P}^T\right] \tag{10.75}$$

where $\mathbf{P}^T$ is an $(n-k)$-by-k matrix, representing the transpose of the coefficient matrix $\mathbf{P}$, and $\mathbf{I}_{n-k}$ is the $(n-k)$-by-$(n-k)$ identity matrix. Accordingly, we may perform the following multiplication of partitioned matrices:

$$\begin{aligned} \mathbf{HG}^T &= \left[\mathbf{I}_{n-k} | \mathbf{P}^T\right] \left[\frac{\mathbf{P}^T}{\mathbf{I}_k}\right] \\ &= \mathbf{P}^T + \mathbf{P}^T \end{aligned}$$

where we have used the fact that multiplication of a rectangular matrix by an identity matrix of compatible dimensions leaves the matrix unchanged. In modulo-2 arithmetic, we have $\mathbf{P}^T + \mathbf{P}^T = \mathbf{0}$, where $\mathbf{0}$ denotes an $(n - k)$-by-k null matrix (i.e., a matrix that has zeros for all of its elements). Hence,

$$\mathbf{H}\mathbf{G}^T = \mathbf{0} \tag{10.76}$$

Equivalently, we have $\mathbf{G}\mathbf{H}^T = \mathbf{0}$. Postmultiplying both sides of Eq. (10.74) by $\mathbf{H}^T$, the transpose of $\mathbf{H}$, and then using Eq. (10.76), we get

$$\begin{aligned}\mathbf{c}\mathbf{H}^T &= \mathbf{m}\mathbf{G}\mathbf{H}^T \\ &= \mathbf{0}\end{aligned} \tag{10.77}$$

The matrix $\mathbf{H}$ is called the *parity-check matrix* of the code, and the set of equations specified by Eq. (10.77) are called *parity-check equations*.

The generator equation (10.74) and the parity-check detector equation (10.77) are basic to the description and operation of a linear block code. These two equations are depicted in the form of block diagrams in Figure 10.22*a* and *b*, respectively.

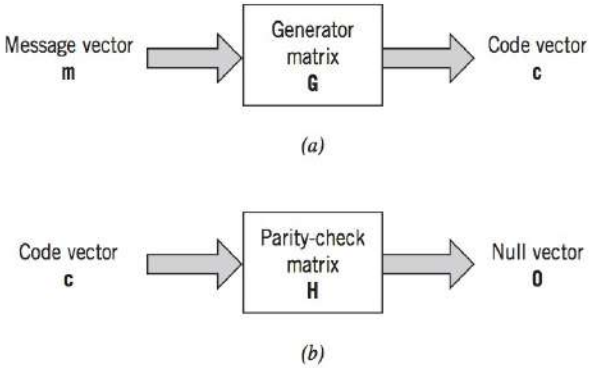


FIGURE 10.22 Block diagram representations of the generator equation (10.74) and the parity-check equation (10.77).

EXAMPLE 10.9 Repetition Codes

Repetition codes represent the simplest type of linear block codes. In particular, a single message bit is encoded into a block of n identical bits, producing an $(n, 1)$ block code. Such a code allows provision for a variable amount of redundancy. There are only two codewords in the code: an all-zero codeword and an all-one codeword.

Consider, for example, the case of a repetition code with $k = 1$ and $n = 5$. In this case, we have four parity bits that are the same as the message bit. Hence, the identity matrix $\mathbf{I}_k = 1$, and the coefficient matrix $\mathbf{P}$ consists of a 1-by-4 vector that has 1 for all of its elements. Correspondingly, the generator matrix equals a row vector of all 1s, as shown by

$$\mathbf{G} = [1 \quad 1 \quad 1 \quad 1 \quad | \quad 1]$$

The transpose of the coefficient matrix $\mathbf{P}$, namely, matrix $\mathbf{P}^T$, consists of a 4-by-1 vector that has 1 for all of its elements. The identity matrix $\mathbf{I}_{n-k}$ consists of a 4-by-4 matrix. Hence, the parity-check matrix equals

$$\mathbf{H} = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Since the message vector consists of a single binary symbol, 0 or 1, it follows from Eq. (10.74) that there are only two codewords: 00000 and 11111 in the (5, 1) repetition code, as expected. Note also that $\mathbf{H}\mathbf{G}^T = \mathbf{0}$, modulo-2, in accordance with Eq. (10.76).

SYNDROME DECODING — I

The generator matrix **G** is used in the encoding operation at the transmitter. On the other hand, the parity-check matrix **H** is used in the decoding operation at the receiver. In the context of the latter operation, let **r** denote the 1-by-*n* *received vector* that results from sending the code vector **c** over a noisy channel. We express the vector **r** as the sum of the original code vector **c** and a vector **e**, as shown by

$$\mathbf{r} = \mathbf{c} + \mathbf{e} \tag{10.78}$$

The vector **e** is called the *error vector* or *error pattern*. The ith element of **e** equals 0 if the corresponding element of **r** is the same as that of **c**. On the other hand, the ith element of **e** equals 1 if the corresponding element of **r** is different from that of **c**, in which case an error is said to have occurred in the ith location. That is, for $i = 1, 2, \ldots, n$, we have

$$e_i = \begin{cases} 1 & \text{if an error has occurred in the } i\text{th location} \\ 0 & \text{otherwise} \end{cases} \tag{10.79}$$

The receiver has the task of decoding the code vector **c** from the received vector **r**. The algorithm commonly used to perform this decoding operation starts with the computation of a 1-by-$(n - k)$ vector called the *error-syndrome vector* or simply the *syndrome*.[5] The importance of the syndrome lies in the fact that it depends only upon the error pattern.

Given a 1-by-n received vector **r**, the corresponding syndrome is formally defined as

$$\mathbf{s} = \mathbf{r}\mathbf{H}^T \tag{10.80}$$

Accordingly, the syndrome has the following important properties.

PROPERTY 1

The syndrome depends only on the error pattern, and not on the transmitted codeword.

To prove this property, we first use Eqs. (10.78) and (10.80) and then Eq. (10.77) to obtain

$$\begin{aligned} \mathbf{s} &= (\mathbf{c} + \mathbf{e})\mathbf{H}^T \\ &= \mathbf{c}\mathbf{H}^T + \mathbf{e}\mathbf{H}^T \\ &= \mathbf{e}\mathbf{H}^T \end{aligned} \tag{10.81}$$

Hence, the parity-check matrix **H** of a code permits us to compute the syndrome **s**, which depends only upon the error pattern **e**.

PROPERTY 2

All error patterns that differ by a codeword have the same syndrome.

For k message bits, there are 2^k distinct code vectors denoted as $\mathbf{c}_i$, $i = 1, \ldots, 2^k$. Correspondingly, for any error pattern **e**, we define the 2^k distinct vectors $\mathbf{e}_i$ as

$$\mathbf{e}_i = \mathbf{e} + \mathbf{c}_i, \qquad i = 1, \ldots, 2^k \tag{10.82}$$

The set of vectors $\{\mathbf{e}_i, i = 1, \ldots, 2^k\}$ so defined is called a *coset* of the code. In other words, a coset has exactly 2^k elements that differ at most by a code vector. Thus, an (n, k)

linear block code has 2^{n-k} possible cosets. In any event, multiplying both sides of Eq. (10.82) by the matrix $\mathbf{H}^T$, we get

$$\begin{aligned}\mathbf{e}_i\mathbf{H}^T &= \mathbf{e}\mathbf{H}^T + \mathbf{c}_i\mathbf{H}^T \\ &= \mathbf{e}\mathbf{H}^T\end{aligned} \tag{10.83}$$

which is independent of the index i. Accordingly, we may state that each coset of the code is characterized by a unique syndrome.

We may put Properties 1 and 2 in perspective by expanding Eq. (10.81). Specifically, with the matrix $\mathbf{H}$ having the systematic form given in Eq. (10.75), where the matrix $\mathbf{P}$ is itself defined by Eq. (10.69), we find from Eq. (10.81) that the $(n - k)$ elements of the syndrome $\mathbf{s}$ are linear combinations of the n elements of the error pattern $\mathbf{e}$ as shown by

$$\begin{aligned}s_1 &= e_0 + e_{n-k}p_{00} + e_{n-k+1}p_{10} + \cdots + e_{n-1}p_{k-1,1} \\ s_2 &= e_1 + e_{n-k}p_{01} + e_{n-k+1}p_{11} + \cdots + e_{n-1}p_{k-1,2} \\ &\vdots \\ s_{n-k} &= e_{n-k} + e_{n-k}p_{0,n-k+1} + \cdots + e_{n-1}p_{k-1,n-k}\end{aligned} \tag{10.84}$$

This set of $(n - k)$ linear equations clearly shows that the syndrome contains information about the error pattern and may therefore be used for error detection. However, it should be noted that the set of equations is *underdetermined* in that we have more unknowns than equations. Accordingly, there is *no* unique solution for the error pattern. Rather there are 2^k error patterns that satisfy Eq. (10.84) and therefore result in the same syndrome, in accordance with Property 2 and Eq. (10.83); the true error pattern is just one of the 2^k possible solutions. In other words, the information contained in the syndrome $\mathbf{s}$ about the error pattern $\mathbf{s}$ is *not* enough for the decoder to compute the exact value of the transmitted code vector. Nevertheless, knowledge of the syndrome $\mathbf{s}$ reduces the search for the true error pattern $\mathbf{e}$ from 2^n to 2^{n-k} possibilities. In particular, the decoder has the task of making the best selection from the coset corresponding to $\mathbf{s}$.

MINIMUM DISTANCE CONSIDERATIONS

Consider a pair of code vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ that have the same number of elements. The *Hamming distance* $d(\mathbf{c}_1,\mathbf{c}_2)$ between such a pair of code vectors is defined as the number of locations in which their respective elements differ.

The *Hamming weight* $w(\mathbf{c})$ of a code vector $\mathbf{c}$ is defined as the number of nonzero elements in the code vector. Equivalently, we may state that the Hamming weight of a code vector is the distance between the code vector and the all-zero code vector.

The *minimum distance* $d_{\min}$ of a linear block code is defined as the smallest Hamming distance between any pair of code vectors in the code. That is, the minimum distance is the same as the smallest Hamming weight of the difference between any pair of code vectors. From the closure property of linear block codes, the sum (or difference) of two code vectors is another code vector. Accordingly, we may state that *the minimum distance of a linear block code is the smallest Hamming weight of the nonzero code vectors in the code.*

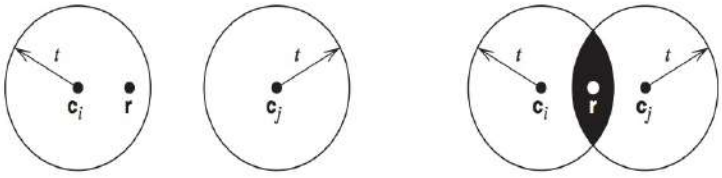
The minimum distance $d_{\min}$ is related to the structure of the parity-check matrix $\mathbf{H}$ of the code in a fundamental way. From Eq. (10.77) we know that a linear block code is defined by the set of all code vectors for which $\mathbf{c}\mathbf{H}^T = \mathbf{0}$, where $\mathbf{H}^T$ is the transpose of the parity-check matrix $\mathbf{H}$. Let the matrix $\mathbf{H}$ be expressed in terms of its columns as follows:

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n] \tag{10.85}$$

Then, for a code vector $\mathbf{c}$ to satisfy the condition $\mathbf{cH}^T = \mathbf{0}$, the vector $\mathbf{c}$ must have 1s in such positions that the corresponding rows of $\mathbf{H}^T$ sum to the zero vector $\mathbf{0}$. However, by definition, the number of 1s in a code vector is the Hamming weight of the code vector. Moreover, the smallest Hamming weight of the nonzero code vectors in a linear block code equals the minimum distance of the code. Hence, *the minimum distance of a linear block code is defined by the minimum number of rows of the matrix* $\mathbf{H}^T$ whose sum is equal to the zero vector.

The minimum distance of a linear block code, $d_{\min}$, is an important parameter of the code. Specifically, it determines the error-correcting capability of the code. Suppose an (n, k) linear block code is required to detect and correct all error patterns (over a binary symmetric channel), whose Hamming weight is less than or equal to t. That is, if a code vector $\mathbf{c}_i$ in the code is transmitted and the received vector is $\mathbf{r} = \mathbf{c}_i + \mathbf{e}$, we require that the decoder output $\hat{\mathbf{c}} = \mathbf{c}_i$, whenever the error pattern $\mathbf{e}$ has a Hamming weight $w(\mathbf{e}) \leq t$. We assume that the 2^k code vectors in the code are transmitted with equal probability. The best strategy for the decoder then is to pick the code vector closest to the received vector $\mathbf{r}$, that is, the one for which the Hamming distance $d(\mathbf{c}_i, \mathbf{r})$ is the smallest. With such a strategy, the decoder will be able to detect and correct all error patterns of Hamming weight $w(\mathbf{e}) \leq t$, provided that the minimum distance of the code is equal to or greater than $2t + 1$. We may demonstrate the validity of this requirement by adopting a geometric interpretation of the problem. In particular, the 1-by-n code vectors and the 1-by-n received vector are represented as points in an n-dimensional space. Suppose that we construct two spheres, each of radius t, around the points that represent code vectors $\mathbf{c}_i$ and $\mathbf{c}_j$. Let these two spheres be disjoint, as depicted in Figure 10.23*a*. For this condition to be satisfied we require that $d(\mathbf{c}_i,\mathbf{c}_j) \geq 2t + 1$. If then the code vector $\mathbf{c}_i$ is transmitted and the Hamming distance $d(\mathbf{c}_i,\mathbf{r}) \leq t$, it is clear that the decoder will pick $\mathbf{c}_i$ as it is the code vector closest to the received vector $\mathbf{r}$. If, on the other hand, the Hamming distance $d(\mathbf{c}_i,\mathbf{c}_j) \leq 2t$, the two spheres around $\mathbf{c}_i$ and $\mathbf{c}_j$ intersect, as depicted in Figure 10.23*b*. Here we see that if $\mathbf{c}_i$ is transmitted, there exists a received vector $\mathbf{r}$ such that the Hamming distance $d(\mathbf{c}_i,\mathbf{r}) \leq t$, and yet $\mathbf{r}$ is as close to $\mathbf{c}_j$ as it is to $\mathbf{c}_i$. Clearly, there is now the possibility of the decoder picking the vector $\mathbf{c}_j$, which is wrong. We thus conclude that *an* (n,k) *linear block code has the power to correct all error patterns of weight t or less if, and only if,*

$$d(\mathbf{c}_i, \mathbf{c}_j) \geq 2t + 1 \qquad \text{for all } \mathbf{c}_i \text{ and } \mathbf{c}_j$$



(*a*) (*b*)

FIGURE 10.23 (a) Hamming distance $d(c_ic_j)$. (b) Hamming distance $d(c_ic_j)$.

By definition, however, the smallest distance between any pair of code vectors in a code is the minimum distance of the code, $d_{\min}$. We may therefore state that *an* (n,k) *linear block code of minimum distance* $d_{\min}$ *can correct up to t errors if, and only if,*

$$t \leq \left\lfloor \frac{1}{2}(d_{\min} - 1) \right\rfloor \tag{10.86}$$

where $\lfloor\ \rfloor$ *denotes the largest integer* less than or equal to the enclosed quantity. Equation (10.86) gives the error-correcting capability of a linear block code a quantitative meaning.

SYNDROME DECODING — II

We are now ready to describe a syndrome-based decoding scheme for linear block codes. Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{2^k}$ denote the 2^k code vectors of an (n,k) linear block code. Let $\mathbf{r}$ denote the received vector, which may have one of 2^n possible values. The receiver has

the task of partitioning the 2^n possible received vectors into 2^k disjoint subsets $D_1, D_2, \ldots, D_{2^k}$ in such a way that the ith subset D_i corresponds to code vector $\mathbf{c}_i$ for $1 \leq i \leq 2^k$. The received vector $\mathbf{r}$ is decoded into $\mathbf{c}_i$ if it is in the ith subset. For the decoding to be correct, $\mathbf{r}$ must be in the subset that belongs to the code vector $\mathbf{c}_i$ that was actually sent.

The 2^k subsets described herein constitute a *standard array* of the linear block code. To construct it, we may exploit the linear structure of the code by proceeding as follows:

1. The 2^k code vectors are placed in a row with the all-zero code vector $\mathbf{c}_1$ as the leftmost element.
2. An error pattern $\mathbf{e}_2$ is picked and placed under $\mathbf{c}_1$, and a second row is formed by adding $\mathbf{e}_2$ to each of the remaining code vectors in the first row; it is important that the error pattern chosen as the first element in a row not have previously appeared in the standard array. (Note that $\mathbf{e}_1 = 0$.)
3. Step 2 is repeated until all the possible error patterns have been accounted for.

Figure 10.24 illustrates the structure of the standard array so constructed. The 2^k columns of this array represent the disjoint subsets $D_1, D_2, \ldots, D_{2^k}$. The 2^{n-k} rows of the array represent the cosets of the code, and their first elements $\mathbf{e}_2, \ldots, \mathbf{e}_{2^{n-k}}$ are called *coset leaders*.

For a given channel, the probability of decoding error is minimized when the most likely error patterns (i.e., those with the largest probability of occurrence) are chosen as the coset leaders. In the case of a binary symmetric channel, the smaller the Hamming weight of an error pattern the more likely it is to occur. Accordingly, the standard array should be constructed with each coset leader having the minimum Hamming weight in its coset.

We may now describe a decoding procedure for a linear block code:

1. For the received vector $\mathbf{r}$, compute the syndrome $\mathbf{s} = \mathbf{r}\mathbf{H}^T$.
2. Within the coset characterized by the syndrome $\mathbf{s}$, identify the coset leader (i.e., the error pattern with the largest probability of occurrence); call it $\mathbf{e}_0$.
3. Compute the code vector

$$\mathbf{c} = \mathbf{r} + \mathbf{e}_0 \tag{10.87}$$

as the decoded version of the received vector $\mathbf{r}$.

This procedure is called *syndrome decoding*.

$\mathbf{c}_1 = \mathbf{0}$	$\mathbf{c}_2$	$\mathbf{c}_3$	$\cdots$	$\mathbf{c}_i$	$\cdots$	$\mathbf{c}_{2^k}$
$\mathbf{e}_2$	$\mathbf{c}_2 + \mathbf{e}_2$	$\mathbf{c}_3 + \mathbf{e}_2$	$\cdots$	$\mathbf{c}_i + \mathbf{e}_2$	$\cdots$	$\mathbf{c}_{2^k} + \mathbf{e}_2$
$\mathbf{e}_3$	$\mathbf{c}_2 + \mathbf{e}_3$	$\mathbf{c}_3 + \mathbf{e}_3$	$\cdots$	$\mathbf{c}_i + \mathbf{e}_3$	$\cdots$	$\mathbf{c}_{2^k} + \mathbf{e}_3$
$\mathbf{e}_j$	$\mathbf{c}_2 + \mathbf{e}_j$	$\mathbf{c}_3 + \mathbf{e}_j$	$\cdots$	$\mathbf{c}_i + \mathbf{e}_j$	$\cdots$	$\mathbf{c}_{2^k} + \mathbf{e}_j$
$\mathbf{e}_{2^{n-k}}$	$\mathbf{c}_2 + \mathbf{e}_{2^{n-k}}$	$\mathbf{c}_3 + \mathbf{e}_{2^{n-k}}$		$\mathbf{c}_i + \mathbf{e}_{2^{n-k}}$	$\cdots$	$\mathbf{c}_{2^k} + \mathbf{e}_{2^{n-k}}$

FIGURE 10.24 Standard array for an (n,k) block code.

EXAMPLE 10.10 Hamming Codes

Consider a family of (n,k) linear block codes that have the following parameters:

$$\begin{aligned} &\text{Block length:} && n = 2^m - 1 \\ &\text{Number of message bits:} && k = 2^m - m - 1 \\ &\text{Number of parity bits:} && n - k = m \end{aligned}$$

where $m \geq 3$. These are the so-called *Hamming codes*.

Consider, for example, the (7, 4) Hamming code with $n = 7$ and $k = 4$, corresponding to $m = 3$. The generator matrix of the code must have a structure that conforms to Eq. (10.73). The following matrix represents an appropriate generator matrix for the (7, 4) Hamming code.

$$\mathbf{G} = \left[\begin{array}{ccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] = [\underbrace{\mathbf{P}}_{} \; \underbrace{\mathbf{I}_k}_{}]$$

The corresponding parity-check matrix is given by

$$\mathbf{H} = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right] = [\underbrace{\mathbf{I}_{n-k}}_{} \; \underbrace{\mathbf{P}^T}_{}]$$

With $k = 4$, there are $2^k = 16$ distinct message words, which are listed in Table 10.4. For a given message word, the corresponding codeword is obtained by using Eq. (10.74). Thus, the application of this equation results in the 16 codewords listed in Table 10.4.

In Table 10.4 we have also listed the Hamming weights of the individual codewords in the (7, 4) Hamming code. Since the smallest of the Hamming weights for the nonzero codewords is 3, it follows that the minimum distance of the code is 3. Indeed, Hamming codes have the property that the minimum distance $d_{\min} = 3$, independent of the value assigned to the number of parity bits m.

To illustrate the relation between the minimum distance $d_{\min}$ and the structure of the parity-check matrix **H**, consider the codeword 0110100. In the matrix multiplication defined by Eq. (10.77), the nonzero elements of this

TABLE 10.4 Codewords of a (7, 4) Hamming code

Message Word	Codeword	Weight of Codeword	Message Word	Codeword	Weight of Codeword
0000	0000000	0	1000	1101000	3
0001	1010001	3	1001	0111001	4
0010	1110010	4	1010	0011010	3
0011	0100011	3	1011	1001011	4
0100	0110100	3	1100	1011100	4
0101	1100101	4	1101	0001101	3
0110	1000110	3	1110	0101110	4
0111	0010111	4	1111	1111111	7

codeword sift out the second, third, and fifth columns of the matrix **H** yielding

$$\begin{bmatrix}0\\1\\0\end{bmatrix}+\begin{bmatrix}0\\0\\1\end{bmatrix}+\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

We may perform similar calculations for the remaining 14 nonzero codewords. We thus find that the smallest number of columns in **H** that sums to zero is 3, confirming the earlier statement that $d_{\min} = 3$.

An important property of Hamming codes is that they satisfy the condition Eq. (10.86) with the equality sign, assuming that $t = 1$. This means that Hamming codes are *single-error correcting binary perfect codes*.

Assuming single-error patterns, we may formulate the seven coset leaders listed in the right-hand column of Table 10.5. The corresponding syndromes, listed in the left-hand column, are calculated in accordance with Eq. (10.81). The zero syndrome signifies no transmission errors.

Suppose, for example, the code vector [1110010] is sent, and the received vector is [1100010] with an error in the third bit. Using Eq. (10.80), the syndrome is calculated to be

$$\mathbf{s} = [1100010]\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\1&1&0\\0&1&1\\1&1&1\\1&0&1\end{bmatrix} = [0\quad 0\quad 1]$$

From Table 10.5 the corresponding coset leader (i.e., error pattern with the highest probability of occurrence) is found to be [0010000], indicating correctly that the third bit of the received vector is erroneous. Thus, adding this error pattern to the received vector, in accordance with Eq. (10.87), yields the correct code vector actually sent.

Richard W. Hamming (1915–1998)

When Richard W. Hamming joined Bell Laboratories, he shared an office with Claude Shannon. While Shannon worked on information theory, Hamming worked on coding theory at the same time and in the same place.

In an interview taped in 1977, just about three decades after the discovery of the first binary codes, Hamming recalled his frustration working on a mechanical relay computer, to which he had access only on weekends: "Two weekends in a row I came in and found that all my stuff had been dumped and nothing was done ... And so I said: "Damn it, if the machine can detect an error, why can't it locate the position of the error and correct it?" It was that question that led to the discovery of the first binary error-correcting codes by Hamming.

The history of the origin of coding theory has a controversy of its own. The publication of Hamming's paper in the Bell System Technical Journal in 1949 was held up for some time due to patent reasons. In that same year, Golay published a paper in the *Proceedings of the IRE* (later renamed the IEEE), in which his (23,12) and (11,6) codes were described. For an interesting exposé of how this controversy played out, see the last section of Chapter 1 of Thompson's book (1983).

Dual Code. Given a linear block code, we may define its *dual* as follows. Taking the transpose of both sides of Eq. (10.76), we have

$$\mathbf{G}\mathbf{H}^T = 0$$

where $\mathbf{H}^T$ is the Transpose of the parity-check matrix of the code, and **0** is a new zero matrix. This equation suggests that every (n,k) linear block code with generator matrix **G** and parity-check matrix **H** has a *dual code* with parameters $(n, n - k)$, generator matrix **H** and parity-check matrix **G**.

TABLE 10.5 Decoding table for the (7, 4) Hamming code defined in Table 10.4

Syndrome	Error Pattern
000	0000000
100	1000000
010	0100000
001	0010000
110	0001000
011	0000100
111	0000010
101	0000001

CYCLIC CODES

The set of linear block codes is large. One important subclass of linear block codes is known as *cyclic codes*, as they are characterized by the fact that any cyclic shift a codeword is also a codeword. Important examples of cyclic codes are:

- Hamming codes, of which we have already given an example.

- Maximal length codes, which have very good autocorrelation properties and find many applications outside of forward error correction.
- Cyclic redundancy check (CRC) codes, which add parity bits to a transmission with the primary purpose of allowing the receiver to reliably, determine if any errors occurred in the transmission. Thus, these are *error-detection* codes.
- Bose–Chaudhuri–Hocquenghem (BCH) codes is a large family of cyclic codes. The BCH codes offer flexibility in the choice of code parameters, namely, block length and code rate.
- Reed–Solomon (RS) codes are an important subclass of nonbinary BCH codes. The encoder for an RS code differs from a binary encoder in that it operates on multiple bits rather than individual bits. An RS (n,k) always satisfies the condition $n - k = 2t$; this property makes the class of RS codes very powerful in an error-correcting sense.

A detailed treatment of these different cyclic coding techniques is beyond the scope of our present discussion.[6]

EXAMPLE 10.11 BER Performance of Cyclic Codes

In Figure 10.25, we compare the bit error rate (BER) performance of three cyclic coding techniques when applied to binary PSK transmitted over a Gaussian channel. The three codes are the Hamming (7,4) code, the BCH (31,16) code, and the RS(31,15) code. We have discussed the Hamming (7,4) code previously. The BCH (31,15) code adds 15 parity bits to 16 information bits to obtain a code capable of correcting $t = 3$ errors. The RS (31,15) code adds 16 parity symbols, not bits, to 15 information symbols, where each symbol represents 5 bits, to produce a code with $n - k = 16$; thus this code can correct $t = 8$ erroneous symbols (not bits) in a codeword.

The first observation from Figure 10.25 is that forward error-correction does not always improve performance. At low signal-to-noise ratios, the addition of forward error correction actually degrades bit error rate performance; with stronger codes, such as the RS (31,15) causing more degradation than the weaker codes. With sufficient SNR, the FEC does provide advantage. The SNR threshold where this crossover occurs depends upon the code and the decoding technique, as we shall see in the following. FEC techniques that are decoded using soft decisions will usually have a lower crossover point than block codes, which typically use hard decisions.

Above the SNR threshold, we see there is a progressive improvement in bit error rate performance relative to uncoded PSK, with the three codes. The code rates for the three codes are 4/7, 16/31, and 15/31; that is, all the code rates are approximately 1/2. Thus, for each code, approximately the same proportion of redundant bits is added to the transmission. Even

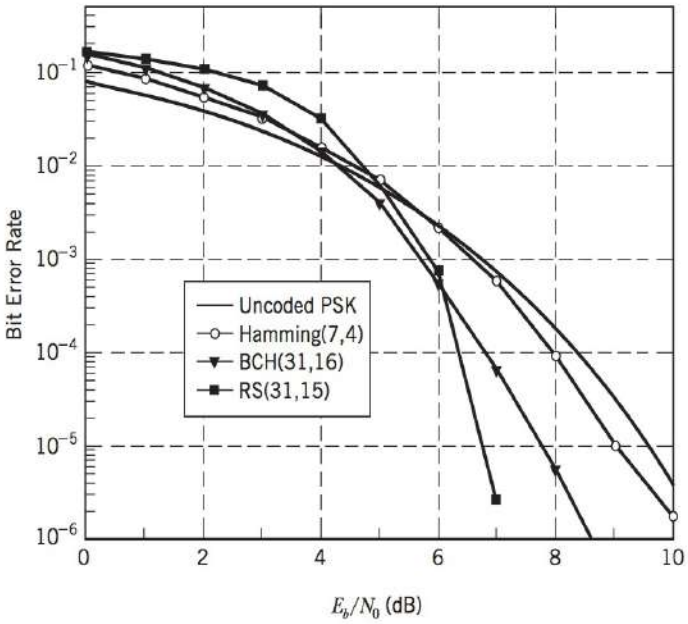


FIGURE 10.25 Simulated BER performance of uncoded and FEC-encoded PSK transmitted over a Gaussian channel.

though the three codes have approximately the same code rate, this improvement in performance is due to the codeword length. The numbers of bits in each codeword are 7, 31, and 155, respectively, for each of the three codes. With proper design, a longer code can have proportionately greater error-correcting capabilities due to the statistical properties of the channel. This improved performance is obtained at the expense of increased decoder complexity.

We must carefully interpret the horizontal axis of Figure 10.25. When dealing with uncoded modulation, the quantity E_b has always represented the energy per bit. When using coded modulation, there are two types of bits: *information bits* that are the input to FEC encoder; and *channel bits* that are the output of the FEC encoder. Since we are fundamentally interested in the information, we let E_b represent the energy per information bit. The energy per channel bit is then given by $E_c = r\,E_b$, where r is the code rate. For uncoded modulation, information bits and channel bits are the same. With this definition, the use of E_b/N_0 as the horizontal axis allows a fair comparison of the bit error rate performance of uncoded modulation and FEC codes having different code rates.

When interpreted in terms of the capacity curve of Figure 10.15, the three codes, of approximately the same rate correspond to a horizontal constant line $r/B = 0.5$, below the uncoded line $r/B = 1$. The improved performance with the different codes means moving to the left on this line towards the theoretical capacity limit.

In the above example, we did not make full use of the capabilities of the Reed–Solomon encoding scheme, since it is a nonbinary encoding scheme applied to a binary modulation scheme. If, for example, we had applied the RS(31,16) code to an M-ary modulation scheme such as 32-FSK, which has 5 bits per FSK tone, then an even greater improvement in bit error rate would have been realized.

10.12 CONVOLUTIONAL CODES[7]

In block coding, the encoder accepts a k-bit message block and generates an n-bit codeword. Thus, codewords are produced on a block-by-block basis. Clearly, provision must be made in the encoder to buffer an entire message block before generating the associated codeword. There are applications, however, where the message bits come in *serially* rather than in large blocks, in which case the use of a buffer may be undesirable. In such situations, the use of *convolutional coding* may be the preferred method. A convolutional encoder operates on the incoming message sequence continuously in a serial manner.

The encoder of a binary convolutional code with rate $1/n$, measured in bits per symbol, may be viewed as a *finite-state machine* that consists of an M-stage shift register with prescribed connections to n modulo-2 adders, and a multiplexer that serializes the outputs of the adders. An L-bit message sequence produces a coded output sequence of length $n(L + M)$ bits. The *code rate* is therefore given by

$$r = \frac{L}{n(L+M)} \quad \text{bits/symbol} \tag{10.88}$$

Typically, we have $L >> M$. Hence, the code rate simplifies to

$$r \simeq \frac{1}{n} \quad \text{bits/symbol} \tag{10.89}$$

The *constraint length* of a convolutional code, expressed in terms of message bits, is defined as the number of shifts over which a single message bit can influence the encoder output. In an encoder with an M-stage shift register, the *memory* of the encoder equals

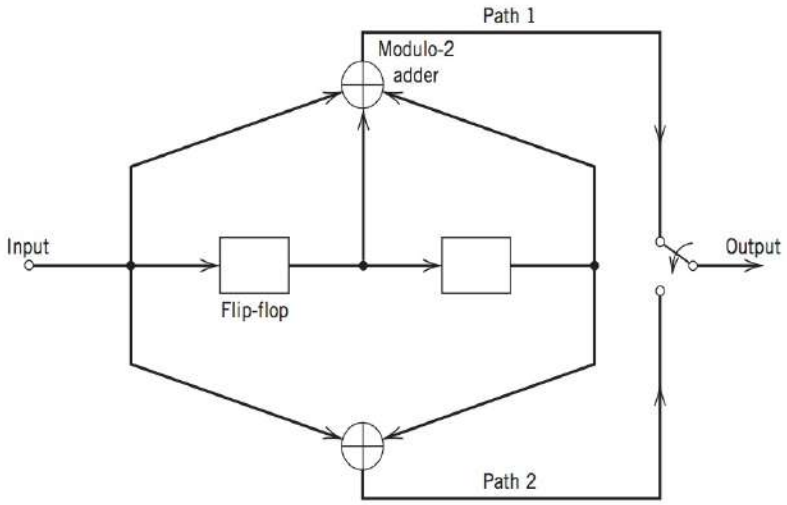
M message bits, and $K = M + 1$ shifts are required for a message bit to enter the shift register and finally come out. Hence, the constraint length of the encoder is K.

Figure 10.26*a* shows a convolutional encoder with $n = 2$ and $K = 3$. Hence, the code rate of this encoder is 1/2. The encoder of Figure 10.26*a* operates on the incoming message sequence, one bit at a time.

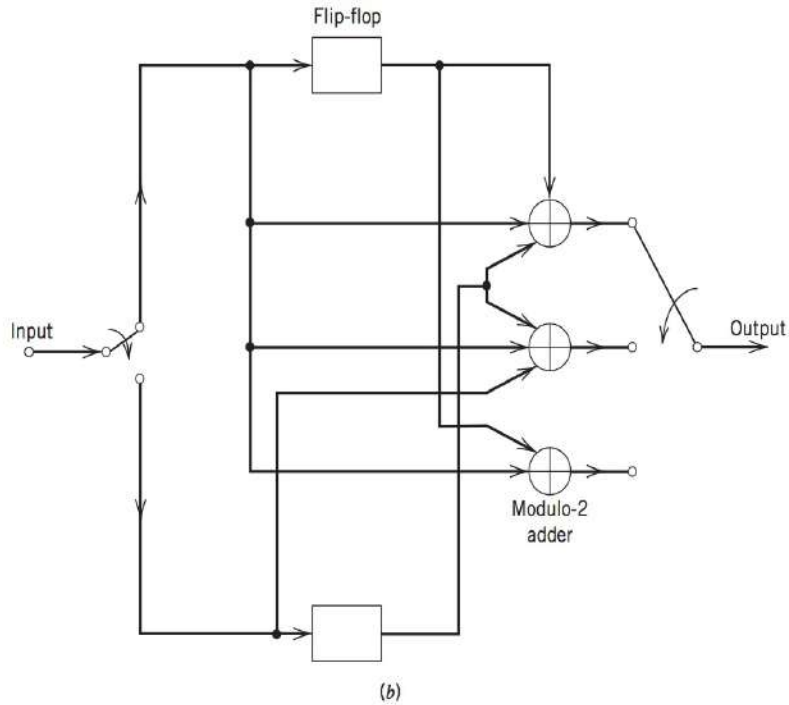
We may generate a binary convolutional code with rate k/n by using k separate shift registers with prescribed connections to n modulo-2 adders, an input multiplexer and an output multiplexer. An example of such an encoder is shown in Figure 10.26*b*, where $k = 2$, $n = 3$, and the two shift registers have $K = 2$ each. The code rate is 2/3. In this second example, the encoder processes the incoming message sequence two bits at a time.

The convolutional codes generated by the encoders of Figure 10.26 are *nonsystematic* codes. Unlike block coding, the use of nonsystematic codes is ordinarily preferred over systematic codes in convolutional coding.

Each path connecting the output to the input of a convolutional encoder may be characterized in terms of its *impulse response*, defined as the response of that path to a



(*a*)



(*b*)

FIGURE 10.26 (*a*) Constraint length-3, rate-$\frac{1}{2}$ convolutional encoder. (*b*) Constraint length-2, rate-$\frac{2}{3}$ convolutional encoder.

symbol 1 applied to its input, with each flip-flop in the encoder set initially in the zero state. Equivalently, we may characterize each path in terms of a *generator polynomial*, defined as the *unit-delay transform* of the impulse response. To be specific, let the *generator sequence* $(g_0^{(i)}, g_1^{(i)}, g_2^{(i)}, \ldots, g_M^{(i)})$ denote the impulse response of the ith path, where the coefficients $g_0^{(i)}, g_1^{(i)}, g_2^{(i)}, \ldots, g_M^{(i)}$ equal 0 or 1. Correspondingly, the *generator polynomial* of the ith path is defined by

$$g^{(i)}(D) = g_0^{(i)} + g_1^{(i)}D + g_2^{(i)}D^2 + \cdots + g_M^{(i)}D^M \tag{10.90}$$

where D denotes the unit-delay variable and + corresponds to modulo 2 addition. The complete convolutional encoder is described by the set of generator polynomials $\{g^{(1)}(D), g^{(2)}(D), \ldots, g^{(n)}(D)\}$. Traditionally, different variables are used for the description of convolutional and cyclic codes, with D being commonly used for convolutional codes and X for cyclic codes.

EXAMPLE 10.12

Consider the convolutional encoder of Figure 10.26*a*, which has two paths numbered 1 and 2 for convenience of reference. The impulse response of path 1 is (1, 1, 1). Hence, the corresponding generator polynomial is given by

$$g^{(1)}(D) = 1 + D + D^2$$

The impulse response of path 2 is (101). Hence, the corresponding generator polynomial is given by

$$g^{(2)}(D) = 1 + D^2$$

For the message sequence (10011), say, we have the polynomial representation

$$m(D) = 1 + D^3 + D^4$$

As with Fourier transformation, convolution in the time domain is transformed into multiplication in the D domain. Hence, the output polynomial of path 1 is given by

$$\begin{aligned} c^{(1)}(D) &= g^{(1)}(D)m(D) \\ &= (1 + D + D^2)(1 + D^3 + D^4) \\ &= 1 + D + D^2 + D^3 + D^6 \end{aligned}$$

From this we immediately deduce that the output sequence of path 1 is (1111001). Similarly, the output polynomial of path 2 in Figure 10.26*a* is

$$\begin{aligned} c^{(2)}(D) &= g^{(2)}(D)m(D) \\ &= (1 + D^2)(1 + D^3 + D^4) \\ &= 1 + D^2 + D^3 + D^4 + D^5 + D^6 \end{aligned}$$

The output sequence of path 2 is therefore (1011111). Finally, multiplexing the two output sequences of paths 1 and 2, we get the encoded sequence

$$\mathbf{c} = (11,\ 10,\ 11,\ 11,\ 01,\ 01,\ 11)$$

Note that the message sequence of length $L = 5$ bits produces an encoded sequence of length $n(L + K - 1) = 14$ bits. Note also that for the shift register to be restored to its zero initial state, a terminating sequence of $K - 1 = 2$ zeros is appended to the last input bit of the message sequence. The terminating sequence of $K - 1$ zeros is called the *tail of the message* or *flush bits*.

TABLE 10.6 State table for the convolutional encoder of Figure 10.26*a*

State	Binary Description
a	00
b	10
c	01
d	11

The structural properties of a convolutional encoder may be portrayed in graphical form as a trellis diagram. A trellis brings out explicitly the fact that the associated convolutional encoder is a finite-state machine. We define the *state* of a convolutional encoder of rate $1/n$ as the $(K-1)$ message bits stored in the encoder's shift register. At time j, the portion of the message sequence containing the most recent K bits is written as $(m_{j-K+1}, \ldots, m_{j-1}, m_j)$, where m_j is the *current* bit. The $(K-1)$-bit state of the encoder at time j is therefore written simply as $(m_{j-1}, \ldots, m_{j-K+2}, m_{j-K+1})$. In the case of the simple convolutional encoder of Figure 10.26*a* we have $(K-1)=2$. Hence, the state of this encoder can assume any one of four possible values, as described in Table 10.6. The trellis contains $(L+K)$ *levels*, where L is the length of the incoming message sequence, and K is the constraint length of the code. The levels of the trellis are labeled as $j=0, 1, \ldots, L+K-1$ in Figure 10.27 for $K=3$. Level j is also referred to as *depth* j; both terms are used interchangeably. The first $(K-1)$ levels correspond to the encoder's departure from the initial state a, and the last $(K-1)$ levels correspond to the encoder's return to the state a. Clearly, not all the states can be reached in these two portions of the trellis. However, in the central portion of the trellis, for which the level j lies in the range $K-1 \leq j \leq L$, all the states of the encoder are reachable. Note also that the central portion of the trellis exhibits a fixed periodic structure.

DECODING OF CONVOLUTIONAL CODES

Now that we understand the operation of a convolutional encoder, the next issue to be considered is the decoding of a convolutional code. Let $\mathbf{m}$ denote a message vector, and $\mathbf{c}$ denote the corresponding code vector applied by the encoder to the input of a discrete memoryless channel. Let $\mathbf{r}$ denote the received vector, which may differ from the transmitted code vector due to channel noise. Given the received vector $\mathbf{r}$, the decoder is required to make an estimate $\hat{\mathbf{m}}$ of the message vector. Since there is a one-to-one correspondence between the message vector $\mathbf{m}$ and the code vector $\mathbf{c}$, the decoder may equivalently produce an estimate $\hat{\mathbf{c}}$ of the code vector. We may then put $\hat{\mathbf{m}} = \mathbf{m}$ if and only if $\hat{\mathbf{c}} = \mathbf{c}$.

The object of the decoder is to minimize the probability of decoding error. For a binary symmetric channel the optimum decoding rule is: *Choose the estimate* $\hat{\mathbf{c}}$ *that has the minimum Hamming distance from the received vector* $\mathbf{r}$. This is often referred to as a *minimum distance decoder* and it is intuitively appealing. This decoding strategy is also known to be optimum in a likelihood sense; thus, it is also referred to as a *maximum likelihood decoder*.

VITERBI ALGORITHM

Recall the trellis description of a convolutional code that was provided in the previous section. A codeword represents one path through the trellis that outputs a symbol on each

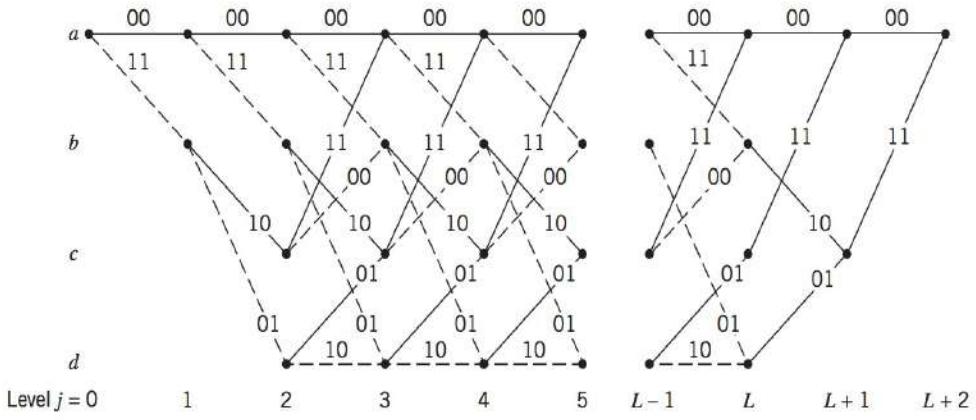


FIGURE 10.27 Trellis for the convolutional encoder of Figure 10.26*a*.

transition between nodes. The equivalence between maximum likelihood decoding and minimum distance decoding for a binary symmetric channel implies that we may decode a convolutional code by choosing a path in the code trellis whose coded sequence differs from the received sequence in the fewest number of places. That is, we limit our choice to the possible paths in the trellis representation of the code.

Consider, for example, the trellis diagram of Figure 10.27 for a convolutional code with rate $r = 1/2$ and constraint length $K = 3$. We observe that at level $j = 3$, there are two paths entering any of the four nodes in the trellis. Moreover, these two paths will be identical onward from that point. Clearly, a minimum distance decoder may make a decision at that point as to which of those two paths to retain, without any loss of performance. A similar decision may be made at level $j = 4$, and so on. This sequence of decisions is exactly what the *Viterbi algorithm* does as it walks through the trellis. The algorithm operates by computing a metric or discrepancy for every possible path in the trellis. The metric for a particular path is defined as the Hamming distance between the coded sequence represented by that path and the received sequence. Thus, for each node (state) in the trellis of Figure 10.27 the algorithm compares the two paths entering the node. The path with the lower metric is retained, and the other path is discarded. This computation is repeated for every level j of the trellis in the range $M \leq j \leq L$, where $M = K - 1$ is the encoder's memory and L is the length of the incoming message sequence. The paths that are retained by the algorithm are called *survivor* or *active paths*. For a convolutional code of constraint length $K = 3$, for example, no more than $2^{K-1} = 4$ survivor paths and their metrics will ever be stored. This list of 2^{K-1} paths is always guaranteed to contain the maximum likelihood choice.

The Viterbi algorithm may also be applied to decoding of convolutional algorithms over other channels, such as the Gaussian channel. For the Gaussian channel, distance is measured in terms of the geometric distance between the transmitted symbol and the received estimate of that symbol, rather than the Hamming distance. The Viterbi algorithm is commonly used in digital communication systems. In fact, many digital signal processors include special instructions to assist Viterbi decoding.

FREE DISTANCE AND ASYMPTOTIC CODING GAIN OF A CONVOLUTIONAL CODE

The bit error rate performance of a convolutional code depends not only on the decoding algorithm used but also on the distance properties of the code. In this context, the most important single measure of a convolutional code's ability to combat channel noise is the *free distance*, denoted by d_{free}. The free distance of a convolutional code is defined as the minimum Hamming distance between any two codewords in the code. Similar to a block code, a convolutional code with free distance d_{free} can correct t errors if and only if d_{free} is greater than $2t$. Investigations have shown that the free distance of systematic convolutional codes are usually smaller than for the case of nonsystematic convolutional codes as indicated in Table 10.7.

TABLE 10.7 Maximum free distances attainable with systematic and nonsystematic convolutional codes of rate $\frac{1}{2}$

Constraint Length K	Systematic	Nonsystematic
2	3	3
3	4	5
4	4	6
5	5	7
6	6	8
7	6	10
8	7	10

A bound on the bit error rate for convolutional codes may be obtained analytically, but the details of this evaluation are, however, beyond the scope of our present discussion. Here we simply summarize an asymptotic result for the binary-input additive white Gaussian noise (AWGN) channel, assuming the use of binary phase-shift keying (PSK) with coherent detection. For the case of a memoryless binary-input AWGN channel with no output quantization, theory shows that for large values of E_b/N_0 the bit error rate for binary PSK with convolutional coding is dominated by the exponential factor $\exp(-d_{free} r E_b/N_0)$, where the parameters are as previously defined. Accordingly, in this case, we find that the asymptotic coding gain,

that is, the advantage over uncoded transmission at high SNR, is defined by

$$G_a = 10 \log_{10}(d_{free} r) \text{dB} \tag{10.91}$$

As mentioned above, this result assumes an unquantized demodulator output. If hard decisions are made on the channel outputs before decoding then both theory and practice show an approximate 2 dB loss in performance. The improvement without quantization, however, is attained at the cost of increased decoder complexity due to the requirement for accepting analog inputs. The asymptotic coding gain for a binary-input AWGN channel is approximated to within about 0.25 dB by a binary input Q-ary output discrete memoryless channel with the number of representation levels $Q = 8$. This means that we may avoid the need for an analog decoder by using a soft-decision decoder that performs finite output quantization (typically, $Q = 8$), and yet realize a performance close to the optimum.

EXAMPLE 10.13 BER Performance of Convolutional Codes

To illustrate these asymptotic results, we have plotted the simulated performance of convolutional codes of constraint length: 3, 5, 7, and 9 in Figure 10.28 for comparison with uncoded performance. All the codes have rate $\frac{1}{2}$. If we compare the advantage these codes have over uncoded performance at low bit error rates, say 10^{-5}, we find that the gain approaches that predicted by Eq. (10.91) when we use the corresponding values from Table 10.7 and $r = \frac{1}{2}$. Note that E_b represents the energy per information bit as described in Example 10.11. At low signal-to-noise ratios, the performance curves of convolutional codes cross, with weaker codes performing better than more complex codes; this type of behavior is often seen at low SNR. Note that at very low SNR, forward error coding performs worse than uncoded PSK.

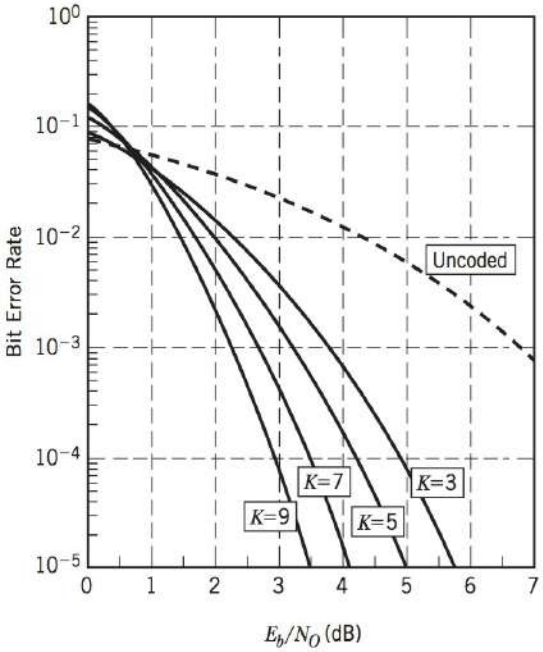


FIGURE 10.28 Simulated BER performance of rate $\frac{1}{2}$ convolutional codes with constraint lengths 3, 5, 7, and 9. Also included is BER of uncoded PSK. (The uncoded BER is 10^{-5} at E_b/N_0 of 9.6 dB.)

10.13 TRELLIS-CODED MODULATION[8]

In the traditional approach to channel coding described in the preceding sections of the chapter, encoding is performed separately from modulation in the transmitter; likewise for decoding and detection in the receiver. Moreover, error control is provided by transmitting additional redundant bits in the code, which has the effect of lowering the information bit rate per channel bandwidth. That is, bandwidth efficiency is traded for increased power efficiency.

To attain a more effective utilization of the available bandwidth and power, coding and modulation have to be treated as a single entity. We may deal with this new situation by redefining coding as *the process of imposing certain patterns on the transmitted signal*. Indeed, this definition includes the traditional idea of parity coding.

Trellis codes for band-limited channels result from the treatment of modulation and coding as a *combined* entity rather than as two separate operations. The combination is itself referred to as *trellis-coded modulation* (TCM). This form of signaling has three basic features:

1. The number of signal points in the constellation used is larger than what is required for the modulation format of interest with the same data rate; the additional points allow redundancy for forward error-control coding without sacrificing bandwidth.
2. Convolutional coding is used to introduce a certain dependency between successive signal points, such that, only certain *patterns* or *sequences of signal points* are permitted.
3. Soft-decision decoding is performed in the receiver, in which the permissible sequence of signals is modeled as a trellis structure; hence, the name trellis codes.

This latter requirement is the result of using a larger signal constellation. By increasing the size of the constellation the probability of symbol error increases for a fixed signal-to-noise ratio. Hence, with hard-decision demodulation we would face a loss before we begin. Performing soft-decision decoding on the combined code and modulation trellis ameliorates this problem.

In the presence of AWGN, maximum likelihood decoding of trellis codes consists of finding that particular path through the trellis with *minimum squared Euclidean distance* to the received sequence. Thus, in the design of trellis codes, the emphasis is on maximizing the Euclidean distance between code vectors (or, equivalently, codewords) rather than maximizing the Hamming distance of an error-correcting code. The reason for this approach is that, except for conventional coding with binary PSK and QPSK, maximizing the Hamming distance is not the same as maximizing the squared Euclidean distance. Accordingly, in what follows, the Euclidean distance is adopted as the distance measure of interest. Moreover, while a more general treatment is possible, the discussion is (by choice) confined to the case of *two-dimensional constellations of signal points*. The implication of such a choice is to restrict the development of trellis codes to multilevel amplitude and/or phase modulation schemes such as M-ary PSK and M-ary QAM.

The approach used to design this type of trellis codes involves partitioning an M-ary constellation of interest successively into 2, 4, 8, ... subsets with size $M/2$, $M/4$, $M/8$, ..., and having progressively larger increasing minimum Euclidean distance between their respective signal points. Such a design approach by *set partitioning* represents the key idea in the construction of efficient coded modulation techniques for band-limited channels.

In Figure 10.29, we illustrate the partitioning procedure by considering a circular constellation that corresponds to 8-PSK. The figure depicts the constellation itself and the 2 and 4 subsets resulting from two levels of partitioning. These subsets share the common property that the minimum Euclidean distances between their individual points follow an increasing pattern: $d_0 < d_1 < d_2$.

Figure 10.30 illustrates the partitioning of a rectangular constellation corresponding to 16-QAM. Here again we see that the subsets have increasing within-subset Euclidean distances: $d_0 < d_1 < d_2 < d_3$.

Based on the subsets resulting from successive partitioning of a two-dimensional constellation, we may devise relatively simple and yet highly effective coding schemes. Specifically, to send n bits/symbol with *quadrature modulation* (i.e., one that has in-phase and quadrature components), we start with a two-dimensional constellation of 2^{n+1} signal points appropriate for the modulation format of interest; a circular grid is used for M-ary PSK, and a rectangular one for M-ary QAM. In any event, the constellation is partitioned into 4 or 8 subsets. One or two incoming bits per symbol enter a rate-1/2 or rate-2/3 binary convolutional encoder, respectively; the resulting two or three coded bits per symbol determine the selection of a particular subset. The remaining uncoded data bits determine which particular point from the selected subset is to be signaled. This class of trellis codes is known as *Ungerboeck codes*.

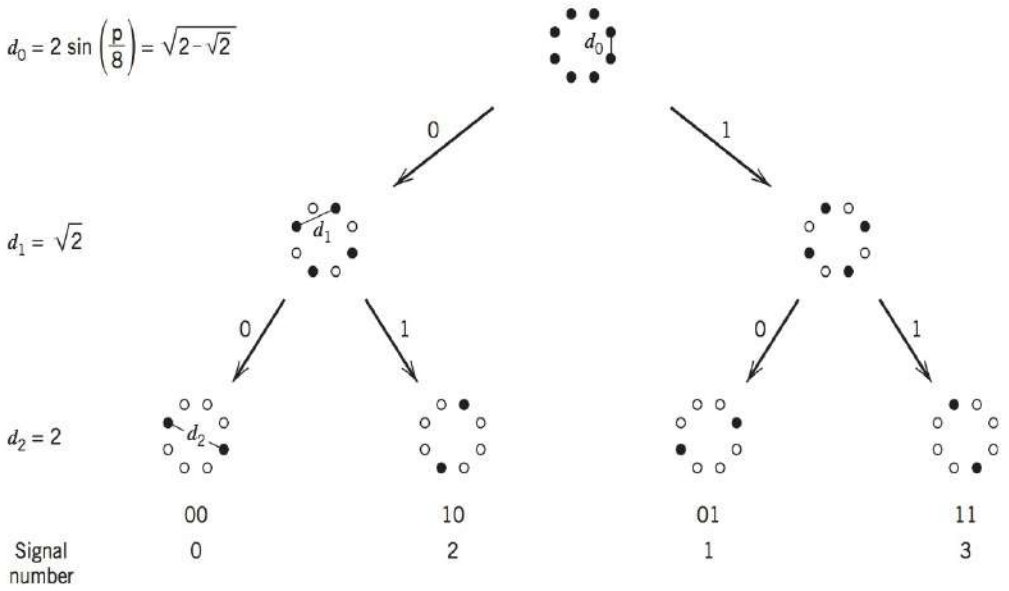


FIGURE 10.29 Partitioning of 8-PSK constellation.

Since the modulator has memory, we may use the Viterbi algorithm to perform maximum likelihood sequence detection at the receiver. Each branch in the trellis of the Ungerboeck code corresponds to a subset rather than an individual signal point. The first step in the detection is to determine the signal point within each subset that is closest to the received signal point in the Euclidean sense. The signal point so determined and its metric (i.e., the squared Euclidean distance between it and the received point) may be used thereafter for the branch in question, and the Viterbi algorithm may then proceed in the usual manner.

UNGERBOECK CODES FOR 8-PSK

The scheme of Figure 10.31*a* depicts the simplest Ungerboeck 8-PSK code for the transmission of 2 bits/symbol. The scheme uses a rate-1/2 convolutional encoder. The corresponding trellis of the code is shown in Figure 10.31*b*, which has four states. Note that

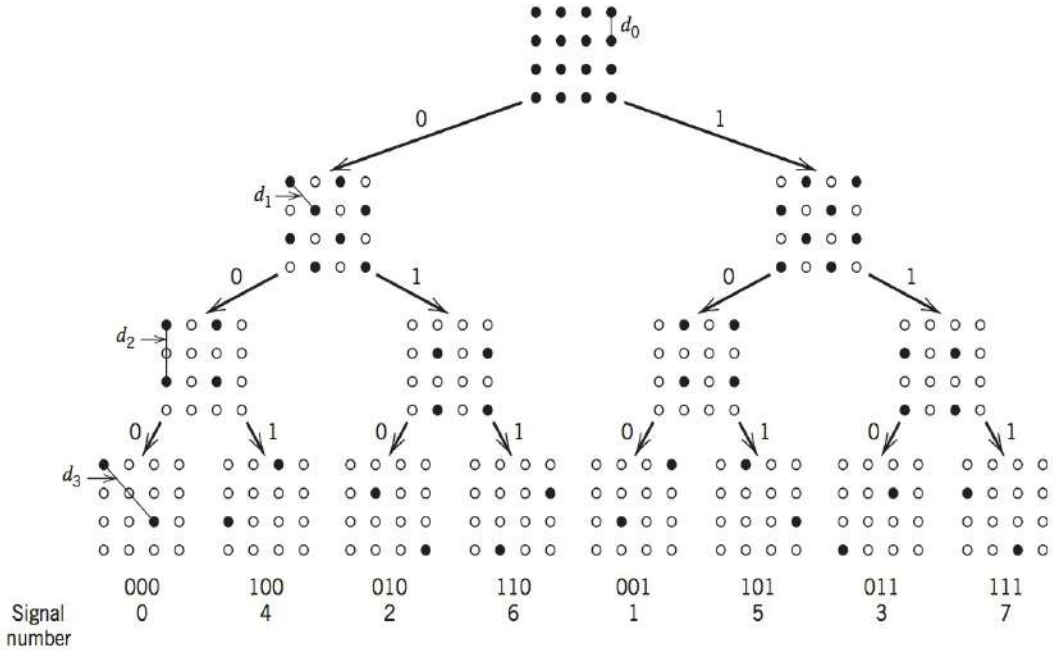


FIGURE 10.30 Partitioning of 16-QAM constellation.

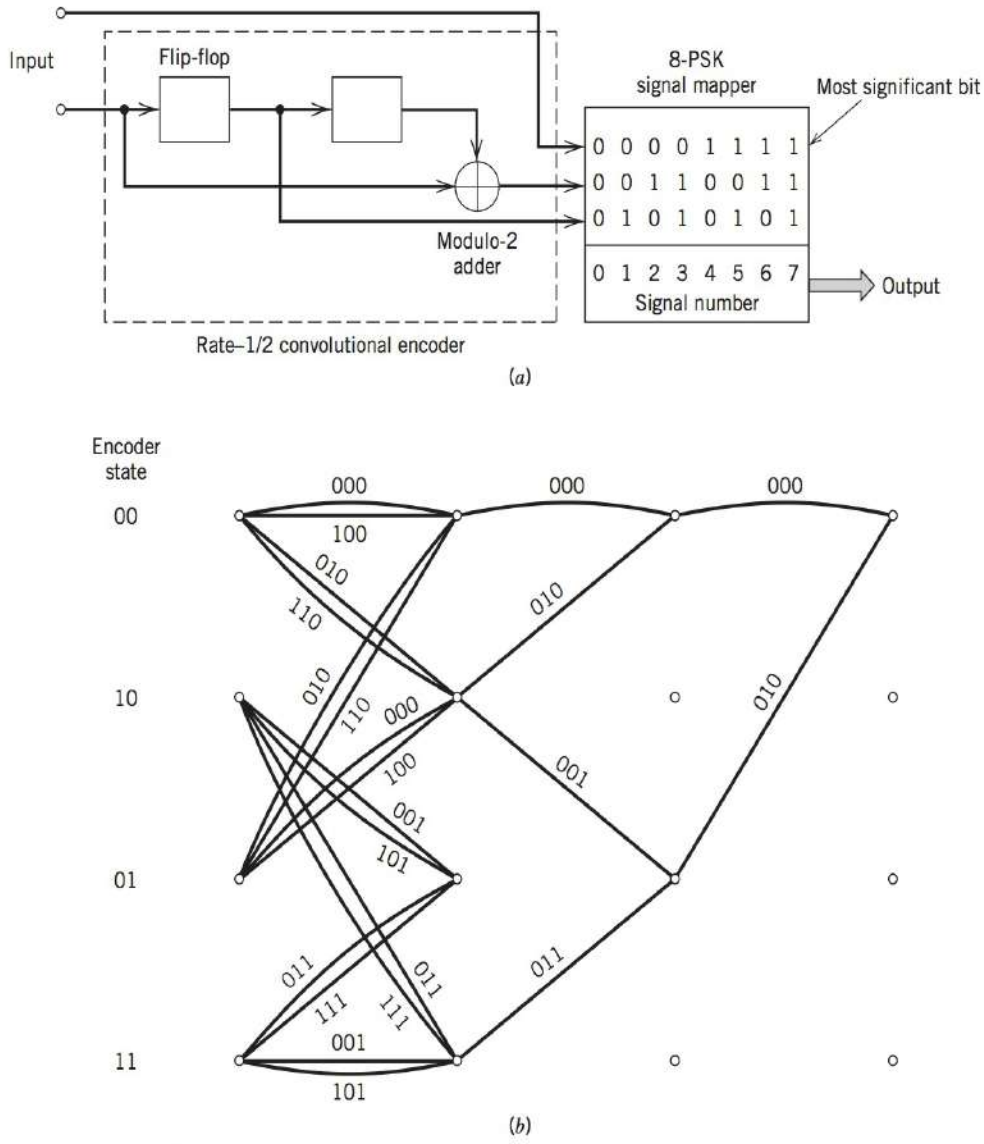


FIGURE 10.31 (*a*) Four-state Ungerboeck code for 8-PSK. (*b*) Trellis.

the most significant bit of the incoming binary word is left uncoded. Therefore, each branch of the trellis may correspond to two different output values of the 8-PSK modulator or, equivalently, to one of the four 2-point subsets shown in Figure 10.29. The trellis of Figure 10.31*b* also includes the minimum distance path.

Figure 10.31*b* and 10.32*b* also includes the encoder states. In Figure 10.31, the state of the encoder is defined by the contents of the two-stage shift register.

ASYMPTOTIC CODING GAIN

Following the discussion in Section 10.12, we define the *asymptotic coding gain* of Ungerboeck codes as

$$G_a = 10 \log_{10}\left(\frac{d_{\text{free}}^2}{d_{\text{ref}}^2}\right) \tag{10.92}$$

where d_{free} is the *free Euclidean distance* of the code and d_{ref} is the minimum Euclidean distance of an uncoded modulation scheme operating with the same signal energy per bit. For example, by using the Ungerboeck 8-PSK code of Figure 10.31*a*, the signal constellation has 8 message points, and we send 2 message bits per point. Hence, uncoded transmission requires a signal constellation with 4 message points. We may therefore regard uncoded 4-PSK as the reference for the Ungerboeck 8-PSK code of Figure 10.31*a*.

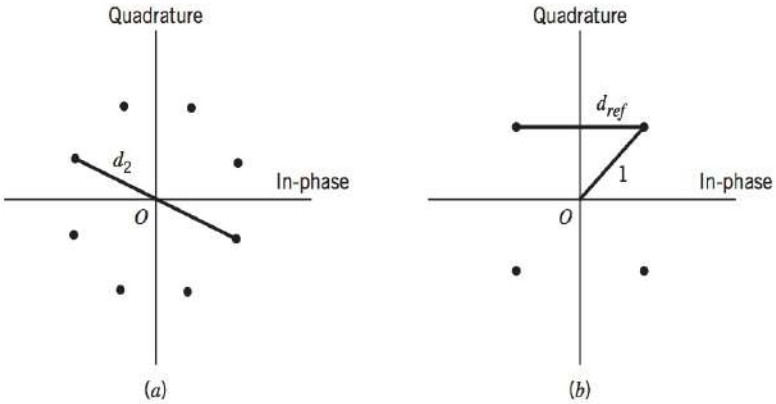


FIGURE 10.32 Signal-space diagrams for calculation of asymptotic coding gain.

The Ungerboeck 8-PSK code of Figure 10.31*a* achieves an asymptotic coding gain of 3 dB, calculated as follows:

1. Each branch of the trellis in Figure 10.31*b* corresponds to a subset of two antipodal signal points. Hence, the free Euclidean distance d_{free} of the code can be no larger than the Euclidean distance d_2 between the antipodal signal points of such a subset. We may therefore write

$$d_{\text{free}} = d_2 = 2$$

   where the distance d_2 is shown defined in Figure 10.32*a*; see also Figure 10.29.

2. The minimum Euclidean distance of an uncoded QPSK, viewed as a reference operating with the same signal energy per bit, equals (see Figure 10.32*b*)

$$d_{\text{ref}} = \sqrt{2}$$

Hence, as previously stated, the use of Eq. (10.92) yields an asymptotic coding gain of $10 \log_{10} 2 = 3$ dB.

10.14 TURBO CODES[9]

Traditionally, the design of good codes has been tackled by constructing codes with a great deal of algebraic structure, for which there are feasible decoding schemes. Such an approach is exemplified by the linear block codes and convolutional codes discussed in preceding sections. The difficulty with these traditional codes is that, in an effort to approach the theoretical limit for Shannon's channel capacity, we need to increase the codeword length of a linear block code or the constraint length of a convolutional code, which, in turn, causes the computational complexity of a maximum likelihood decoder to increase exponentially. Ultimately, we reach a point where complexity of the decoder is so high that it becomes physically unrealizable.

Various approaches have been proposed for the construction of powerful codes with large equivalent block lengths structured in such a way that the decoding can be split into a number of manageable steps. Building on these previous approaches, the development of turbo codes has been by far most successful. Indeed, this development has opened a brand new and exciting way of constructing good codes and decoding them with feasible complexity.

TURBO CODING

In its most basic form, the encoder of a turbo code consists of two constituent systematic encoders joined together by means of an interleaver, as illustrated in Figure 10.33.

An *interleaver* is an input-output mapping device that permutes the ordering of a sequence of symbols from a fixed alphabet in a completely deterministic manner. That is, it takes the symbols at the input and produces identical symbols at the output but in a different temporal order. The interleaver can be of many types, of which the periodic and pseudo-random are two. Turbo codes use a pseudo-random interleaver, which operates only on the systematic bits. There are two reasons for the use of an interleaver in a turbo code:

- To tie together errors that are easily made in one half of the turbo code to errors that are exceptionally unlikely to occur in the other half. This is indeed the main reason why the turbo code performs better than a traditional code.
- To provide robust performance with respect to mismatched decoding, which is a problem that arises when the channel statistics are not known or have been incorrectly specified.

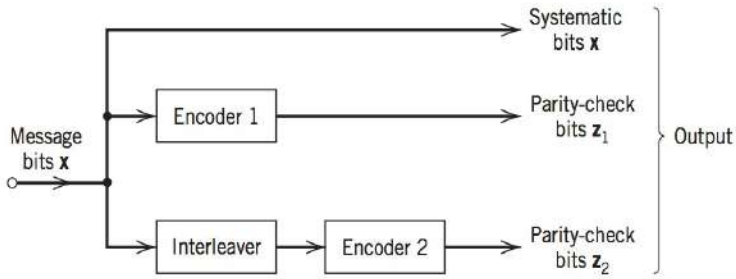


FIGURE 10.33 Block diagram of turbo encoder.

Typically, but not necessarily, the same code is used for both constituent encoders in Figure 10.33. The constituent codes recommended for turbo codes are short *constraint-length recursive systematic convolutional (RSC) codes*. The reason for making the convolutional codes recursive (i.e., feeding one or more of the tap outputs in the shift register back to the input) is to make the internal state of the shift register depend on past outputs. This affects the behavior of the error patterns (a single error in the systematic bits produces an infinite number of parity errors), with the result that a better performance of the overall coding strategy is attained.

EXAMPLE 10.14 Eight-state RSC Encoder

Figure 10.34 show an example eight-state RSC encoder. The generator matrix for this recursive convolutional code is

$$g(D) = \left[1, \frac{1 + D + D^2 + D^3}{1 + D + D^3}\right] \qquad (10.93)$$

where D is the delay variable. The second entry of the matrix $g(D)$ is the transfer function of the feedback shift register, defined as the transform of the output divided by the transform of the input. Let $X(D)$ denote the transform of the message sequence $\{x_i\}$ and $Z(D)$ denote the transform of the parity sequence $\{z_i\}$. By definition, we have

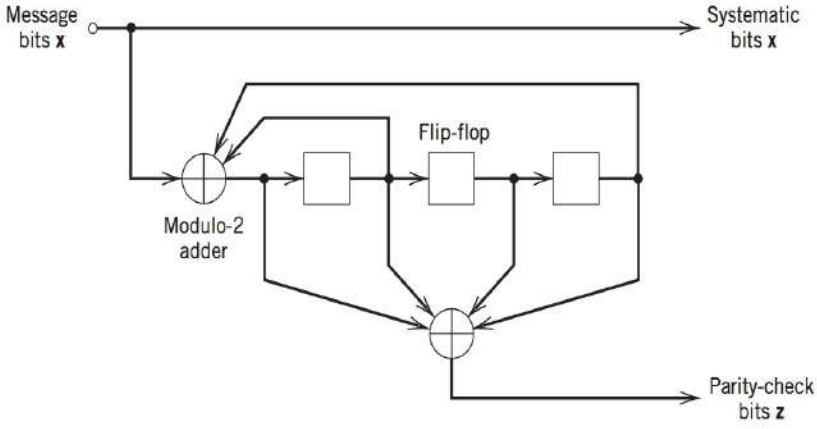


FIGURE 10.34 Example eight-state recursive systematic convolutional (RSC) encoder.

$$\frac{Z(D)}{X(D)} = \frac{1 + D + D^2 + D^3}{1 + D + D^3}$$

Cross-multiplying, we get

$$(1 + D + D^2 + D^3)X(D) = (1 + D + D^3)Z(D)$$

which, on inversion into the time domain, yields

$$x_i + x_{i-1} + x_{i-2} + x_{i-3} + z_i + z_{i-1} + z_{i-3} = 0 \qquad (10.94)$$

where the addition is modulo 2. Equation (10.94) is the parity-check equation, which the convolutional encoder of Figure 10.34 satisfies at each time step t.

In Figure 10.33, the input data stream is applied directly to encoder 1, and the pseudo-randomly reordered version of the same data stream is applied to encoder 2. The systematic bits (i.e., original message bits) and the two sets of parity-check bits generated by the two encoders constitute the output of the turbo encoder. Although the constituent codes are convolutional, in reality turbo codes are block codes with the block size being determined by the size of the interleaver. Moreover, since both RSC encoders in Figure 10.33 are linear, we may describe turbo codes as *linear block codes*.

In the original version of the turbo encoder, the parity-check bits generated by the two encoders in Figure 10.33 were punctured prior to data transmission over the channel to provide a rate $\frac{1}{2}$ code. A *punctured code* is constructed by deleting certain parity-check bits, thereby increasing the data rate. It should, however, be emphasized that the use of a puncture map is not a necessary requirement for the generation of turbo codes.

The novelty of the parallel encoding scheme of Figure 10.33 is in the use of recursive systematic convolutional (RSC) codes and the introduction of a pseudo-random interleaver between the two encoders. Thus, a turbo code appears essentially random to the channel by virtue of the pseudo-random interleaver, yet it possesses sufficient structure for the decoding to be physically realizable. Coding theory asserts that a code chosen at random is capable of approaching Shannon's channel capacity, provided that the block size is sufficiently large.

TURBO DECODING

Turbo codes derive their name from analogy of the decoding algorithm to the turbo engine principle. Figure 10.35 shows the basic structure of an iterative turbo decoder, corresponding to the parallel turbo encoder shown in Figure 10.33. Basically, the turbo decoder consists of two constituent decoders, linked together in a closed-loop structure by means of an interleaver and de-interleaver.

Examining this structure reveals the following distinctive features of the turbo decoder:

1. Each constituent decoder operates on three inputs:

- The noisy *systematic (message) bits.*
- The noisy *parity-check bits* produced by the corresponding constituent encoder.
- The *a priori information* produced by the other constituent decoder.

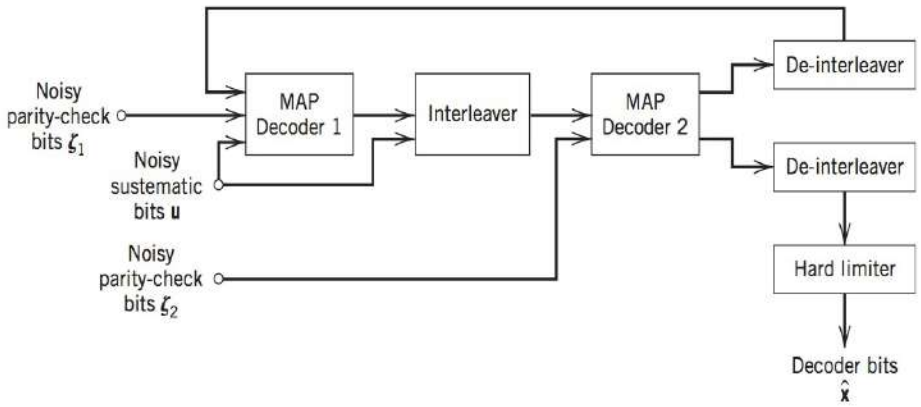


FIGURE 10.35 Block diagram of turbo decoder.

2. For their operations, the two constituent decoders use an algorithm called the *maximum a posteriori* (*MAP*) *decoding algorithm*, hence their labeling as MAP decoders in Figure 10.35. This decoding algorithm is designed to minimize the bit error rate, an objective that is achieved in accordance with the following criterion:

   Let $P(\hat{m}_l = m_l|\mathbf{r})$ denote the conditional probability that the decoded bit $\hat{m}_l$ is equal to the original message bit m_l, given the sequence of noisy bits received at the channel output, $\mathbf{r}$. The requirement for the MAP decoding algorithm is to maximize the probability $P(\hat{m}_l = m_l|\mathbf{r})$.

   The term a posteriori in the MAP decoding algorithm refers to the fact that the decoding is performed after receiving the noisy sequence $\mathbf{r}$.

3. The two constituent decoders, the interleaver and the de-interleaver, constitute a *closed-loop feedback system*, which operates in an iterative manner across time. In other words, the decoding process continues on an iteration-by-iteration basis, until there are no significant adjustments to be made to the decoded message bits, at which point in time the decoding process is terminated and a *hard decision* is made on the decoded bit $\hat{m}_l$, the decision being whether $\hat{m}_l$ represents bit 1 or 0.

The MAP decoders are also commonly referred to as *soft-input soft-output* (*SISO*) *decoders*, in the sense that their input and output signals retain their *analog* (i.e., *unquantized*) characterization throughout the iterative decoding process. The use of *hard decisions*, intended to recover estimates of the original message bits, is carried out at the very end of the sequence of iterations involved in the decoding process. In practice, we typically find that the decoding process is terminated after about five to ten iterations.

LOG-LIKELIHOOD RATIO

Basic to the operation of turbo decoding is the notion of log-likelihood ratio (LLR). In words, the LLR is the natural logarithm of the ratio of two conditional probabilities that an information bit assumes one of two possible values, $+1$ or -1. Let u_l denote an information bit to be decoded into $\hat{m}_l$, given the complete received sequence of bits $\mathbf{r}$. The *log-likelihood ratio*, or *L-value* for short, of the information bit u_l is formally defined by

$$L(u_l|\mathbf{r}) = \ln\left[\frac{P(u_l = +1|\mathbf{r})}{P(u_l = -1|\mathbf{r})}\right] \tag{10.95}$$

where, of course, we know from probability theory that the conditional probability in the numerator has its value constrained as follows

$$0 \le P(u_l = +1|\mathbf{r}) \le 1 \qquad \text{for all } \mathbf{r} \tag{10.96}$$

and the sum

$$P(u_l = +1|\mathbf{r}) + P(u_l = -1|\mathbf{r}) = 1 \qquad \text{for all } \mathbf{r} \tag{10.97}$$

Note that the information bit u_l is assumed to take the possible values $+1$ or -1, rather than 1 or 0.

Figure 10.36 plots the L-value $L(u_l = +1|\mathbf{r})$ versus the conditional probability $P(u_l = +1|\mathbf{r})$ using Eq. (10.95). From this figure we immediately make two practical observations:

1. The sign of the L-value $L(u_l|\mathbf{r})$ indicates whether the information bit u_l takes the possible value $+1$ or -1.
2. The magnitude of the L-value indicates the likelihood of the information bit u_l taking the value $+1$ or -1.

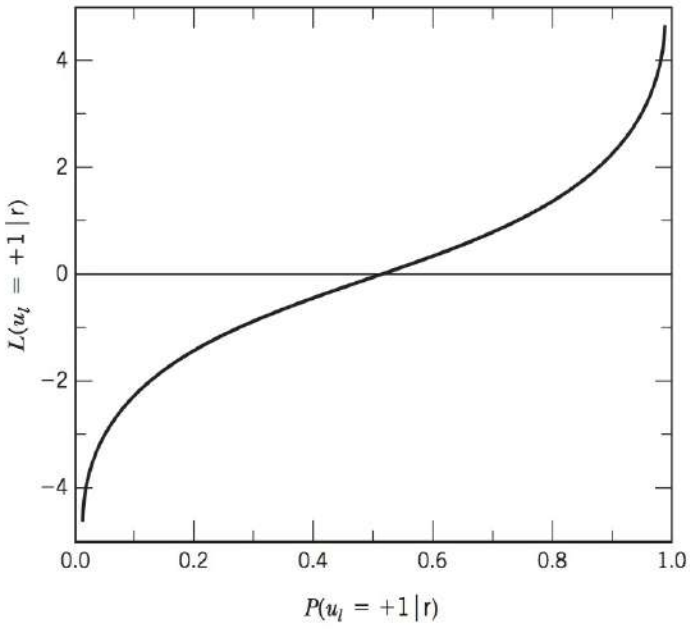


FIGURE 10.36 Log-likelihood ratio versus bit probability.

We may equivalently reformulate Eq. (10.95) as

$$\frac{P(u_l = +1|\mathbf{r})}{P(u_l = -1|\mathbf{r})} = \exp[L(u_l|\mathbf{r})]$$

Hence, solving Eqs. (10.95) and (10.97) for $P(u_l = +1|\mathbf{r})$, we get

$$\begin{aligned} P(u_l = +1|\mathbf{r}) &= \frac{\exp[L(u_l|\mathbf{r})]}{1 + \exp[L(u_l|\mathbf{r})]} \\ &= \frac{1}{1 + \exp[-L(u_l|\mathbf{r})]} \end{aligned} \tag{10.98}$$

Similarly, we have

$$P(u_l = -1|\mathbf{r}) = \frac{1}{1 + \exp[L(u_l|\mathbf{r})]} \tag{10.99}$$

The two conditional probabilities defined in Eqs. (10.98) and (10.99) are called *a posteriori probabilities* of the information bit u_l.

EXTRINSIC INFORMATION

Consider the situation described in Figure 10.35, involving the use of a MAP (i.e., soft-input soft-output) decoder. The *extrinsic information*, generated by a MAP decoder for a set of systematic (message) bits, is defined as the difference between the log-likelihood ratio computed at the output of the MAP decoder and the *intrinsic information* computed at its input. In effect, the extrinsic information generated by the MAP decoder is the incremental information gained by exploiting the dependencies that exist between a message bit of interest and incoming raw data to be processed by the decoder.

There are two basic steps involved in the operation of the MAP decoders:

Step 1. Consider the operation of MAP decoder 1, for example. For the first iteration, there is no extrinsic information from the MAP decoder 2. For second and subsequent iterations, the extrinsic information generated by MAP decoder 2 is reordered to compensate for the pseudo-random interleaving introduced in the turbo encoder. In addition, the received (noisy) parity-check bits, due to encoder 1, are used as inputs. With this combination of inputs, MAP decoder 1 is enabled to produce a refined soft estimate of the message bits.

Step 2. Estimates of the message bits are refined further when MAP decoder 2 takes over its part of the decoding process. This time around, the extrinsic information produced by MAP decoder 1 (and the received systematic bits) are interleaved before its application to MAP decoder 2; the interleaving is done, so that the resulting information sequence corresponds to the original message applied to encoder 2. Then, with the received (noisy) parity-check bits (due to encoder 2) as additional inputs, a further refinement of the message bits is realized.

Thus, the two MAP decoders work together in the turbo engine-like manner, each enhancing the approach of the other.

SUMMARIZING COMMENTS ON TURBO CODES

Turbo codes are capable of delivering impressive performance, which is attributed to two novel features, one in the transmitter and the other in the receiver:

1. The use of a parallel encoder, involving a pair of encoders separated by an interleaver; a pseudo-random encoded version of the message bits is thereby produced for transmission across the channel.

2. The ingenious use of feedback around a corresponding pair of decoders, producing a maximum a posteriori (MAP) probability estimate of the original message bits.

Most important, turbo codes are capable of approaching the Shannon limit in a computationally feasible manner.

10.15 SUMMARY AND DISCUSSION

In this chapter, we began by establishing two fundamental limits on different aspects of a communication system. The limits are embodied in the source coding theorem and the channel coding theorem.

The source coding theorem, Shannon's first theorem, provides the mathematical tool for assessing lossless compression of data generated by a discrete memoryless source. The theorem tells us that we can make the average number of bits per source symbol as small as, but no smaller than, the entropy of the source measured in bits. The entropy of a source is a function of the probabilities of the source symbols that constitute the alphabet of the source. Since entropy is a measure of uncertainty, the entropy is maximum when the associated probability distribution generates maximum uncertainty.

The channel coding theorem, Shannon's second theorem, is the most surprising as well as the single most important result of information theory. The channel coding theorem tells us that for any code rate r less than or equal to the channel capacity C, codes do exist such that the average probability of error is as small as we want it. For the all important Gaussian channel, this theorem implies that the capacity is proportional to the bandwidth of the channel and approximately proportional to the logarithm of the signal-to-noise ratio. When the system operates at a rate greater than the channel capacity, it is condemned to a high probability of error, regardless of the choice of signal set used for transmission or the receiver used for processing the received signal.

The channel coding theorem leads naturally to the study of error-control coding. These techniques represent current approaches to approaching the limits imposed by the channel capacity theorem for reliable digital communication over noisy channels. With error-control coding, the effect of errors occurring during the transmission is reduced by adding redundancy to the data prior to transmission in a controlled manner. The redundancy is used to enable the decoder in the receiver to detect and correct errors.

Error-control coding techniques may be divided into two broadly defined families:

1. *Algebraic codes*, which rely on abstract algebraic structure built into the design of the codes for decoding at the receiver. Algebraic codes include Hamming codes, maximal length codes, BCH codes, and Reed–Solomon codes.
2. *Probabilistic codes*, which rely on probabilistic methods for their decoding at the receiver. Probabilistic codes include convolutional codes and turbo codes. In particular, the decoding is based on one or the other of two basic methods, as summarized here:

   - *Soft input–hard output*, which is exemplified by the *Viterbi algorithm* that performs maximum likelihood sequence estimation in the decoding of trellis-based codes.
   - *Soft input–soft output*, which is exemplified by the *maximum a posteriori* (*MAP*) *algorithm* that performs maximum a posteriori estimation on a bit-by-bit basis in the decoding of turbo codes. The soft output is a necessity for the iterative way in which the MAP algorithm is used.

Trellis-coded modulation combines linear convolutional encoding and modulation to permit significant coding gains over conventional uncoded multilevel modulation without

sacrificing bandwidth efficiency. Turbo codes feature random encoding of a linear block kind, and error performance within a hair's breadth of Shannon's theoretical limit on channel capacity in a physically realizable fashion.

In practical terms, turbo codes have made it possible to achieve coding gains on the order of 10 dB, which is unmatched previously. These coding gains may be exploited to dramatically extend the range of digital communication receivers, substantially increase the bit rates of digital communication systems, or significantly decrease the transmitted signal energy per symbol. These benefits have significant implications for the design of wireless communications and deep-space communications, just to mention two important applications of digital communications. Indeed, turbo codes have already been standardized for use on deep-space communication links and wireless communication systems.

NOTES AND REFERENCES

1. According to Lucky (1989), the first mention of the phrase information theory by Shannon occurs in a 1945 memorandum entitled "A Mathematical Theory of Cryptography." It is rather curious that the phrase information theory was never used in the classic 1948 paper by Shannon, which laid down the foundations of information theory. For an introductory treatment of information theory, see Chapter 2 of Lucky (1989) and the paper by Wyner (1981). See also the books of Adámek (1991), Hamming (1980), and Abramson (1963). For more advanced treatments of the subject, see the books of Cover and Thomas (2006), Blahut (1987), and McEliece (1977).
2. In statistical physics, the entropy of a physical system is defined by (Reif, 1967, p. 147)

$$\mathscr{L} = k \ln \Omega$$

where k is Boltzmann's constant, Ω is the number of states accessible to the system, and ln denotes the natural logarithm. This entropy has the dimensions of energy because its definition involves the constant k. In particular, it provides a *quantitative measure of the degree of randomness of the system*. Comparing the entropy of statistical physics with that of information theory, we see that they have a similar form. For a detailed discussion of the relation between them, see Pierce (1961, pp. 184–207) and Brillouin (1962).
3. For further description of the Lempel–Ziv algorithm and other data compression schemes, see Salomon (2006).
4. For an introductory discussion of error correction by coding, see Chapter 2 of Lucky (1989); see also the book by Adámek (1991), and the paper by Bhargava (1983). The classic book on error-control coding is Peterson and Weldon (1972). The books of Lin and Costello (2004), Michelson and Levesque (1985), MacWilliams and Sloane (1977), and Wilson (1996) are also devoted to error-control coding.
5. In medicine, the term *syndrome* is used to describe a pattern of symptoms that aids in the diagnosis of a disease. In coding, the error pattern plays the role of the disease and parity-check failure that of a symptom. This use of *syndrome* was coined by Hagelbarger.
6. For further information on cyclic codes the interested reader may consult Lin and Costello (2004), and Blahut (1987).
7. Convolutional codes were first introduced, as an alternative to block codes, by P. Elias. Further information on convolutional codes may be found in Proakis (2001).
8. Trellis-coded modulation was invented by G. Ungerboeck. Further discussion of this technique may be found in Lee and Messerschmitt (1994), Biglieri, Divsalar, McLane, and Simon (1991), and Schlegel (1997).
9. Turbo codes were originated by C. Berrou and A. Glavieux. Work on these codes was motivated by two papers on error-correcting codes: Battail (1987), and Hagenauer and Hoeher (1989). The first description of turbo codes using heuristic arguments was presented at a conference paper by Berrou, Glavieux, and Thitamajshima (1993); see also Berrou and Glavieux (1996).

PROBLEMS

10.1 Let p denote the probability of some event. Plot the amount of information gained by the occurrence of this event for $0 \leq p \leq 1$.

10.2 A source emits one of four possible symbols during each signaling interval. The symbols occur with the probabilities:

$$p_0 = 0.4$$
$$p_1 = 0.3$$
$$p_2 = 0.2$$
$$p_3 = 0.1$$

Find the amount of information gained by observing the source emitting each of these symbols.

10.3 A source emits one of four symbols s_0, s_1, s_2, and s_3 with probabilities 1/3, 1/6, 1/4, and 1/4, respectively. The successive symbols emitted by the source are statistically independent. Calculate the entropy of the source.

10.4 Let X represent the outcome of a single roll of a fair die. What is the entropy of X?

10.5 The sample function of a Gaussian process of zero mean and unit variance is uniformly sampled and then applied to a uniform quantizer having the input–output amplitude characteristic shown in Figure P10.5. Calculate the entropy of the quantizer output.

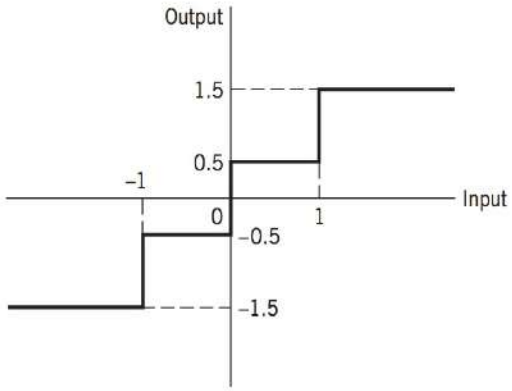


Figure P10.5

10.6 Consider a discrete memoryless source with source alphabet $\mathcal{L} = \{s_0, s_1, s_2\}$ and source statistics $\{0.7, 0.15, 0.15\}$.

(a) Calculate the entropy of the source.

(b) Calculate the entropy of the second-order extension of the source.

10.7 Consider the four codes listed below:

Symbol	Code I	Code II	Code III	Code IV
s_0	0	0	0	00
s_1	10	01	01	01
s_2	110	001	011	10
s_3	1110	0010	110	110
s_4	1111	0011	111	111

Two of these four codes are prefix codes. Identify them, and construct their individual decision trees.

10.8 Consider a sequence of letters of the English alphabet with their probabilities of occurrence as given here:

Letter	a	i	l	m	n	o	p	y
Probability	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.1

Compute two different Huffman codes for this alphabet. Hence, for each of the two codes, find the average codeword length and the variance of the average codeword length over the ensemble of letters.

10.9 A discrete memoryless source has an alphabet of seven symbols whose probabilities of occurrence are as described here:

Symbol	s_0	s_1	s_2	s_3	s_4	s_5	s_6
Probability	0.25	0.25	0.125	0.125	0.125	0.0625	0.0625

Compute the Huffman code for this source, moving a combined symbol as high as possible. Explain why the computed source code has an efficiency of 100 percent.

10.10 Consider a discrete memoryless source with alphabet $\{s_0, s_1, s_2\}$ and statistics $\{0.7, 0.15, 0.15\}$ for its output.

(a) Apply the Huffman algorithm to this source. Hence, show that the average codeword length of the Huffman code equals 1.3 bits/symbol.

(b) Let the source be extended to order two. Apply the Huffman algorithm to the resulting extended source, and show that the average codeword length of the new code equals 1.1975 bits/symbol.

(c) Compare the average codeword length calculated in part (b) with the entropy of the original source.

10.11 A computer executes four instructions that are designated by the codewords (00, 01, 10, 11). Assuming that the instructions are used independently with probabilities (1/2, 1/8, 1/8, 1/4), calculate the percentage by which the number of bits used for the instructions may be reduced by the use of an optimum source code. Construct a Huffman code to realize the reduction.

10.12 Consider the following binary sequence

$$11101001100010110100\ldots$$

Use the Lempel–Ziv algorithm to encode this sequence. Assume that the binary symbols 0 and 1 are already in the codebook.

10.13 Consider the transition probability diagram of a binary symmetric channel shown in Figure 10.9. The input binary symbols 0 and 1 occur with equal probability. Find the probabilities of the binary symbols 0 and 1 appearing at the channel output.

10.14 Repeat the calculation in Problem 10.13, assuming that the input binary symbols 0 and 1 occur with probabilities 1/4 and 3/4, respectively.

10.15 Consider a digital communication system that uses a *repetition code* for the channel encoding/decoding. In particular, each transmission is repeated n times, where $n = 2m + 1$ is an odd integer. The decoder operates as follows. If in a block of n received bits, the number of 0s exceeds the number of 1s, the decoder decides in favor of a 0. Otherwise, it decides in favor of a 1. An error occurs when $m + 1$ or more transmissions out of $n = 2m + 1$ are incorrect. Assume a binary symmetric channel.

(a) For $n = 3$, show that the average probability of error is given by

$$P_e = 3p^2(1-p) + p^3$$

where p is the transition probability of the channel.

(b) For $n = 5$, show that the average probability of error is given by

$$P_e = 10p^3(1-p)^2 + 5p^4(1-p) + p^5$$

(c) For the general case, show that the average probability of error is given by

$$P_e = \sum_{i=m+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

10.16 A voice-grade channel of the telephone network has a bandwidth of 3.4 kHz.

(a) Calculate the information capacity of the telephone channel for a signal-to-noise ratio of 30 dB.

(b) Calculate the minimum signal-to-noise ratio required to support information transmission through the telephone channel at the rate of 9600 b/s.

10.17 Alphanumeric data are entered into a computer from a remote terminal through a voice-grade telephone channel. The channel has a bandwidth of 3.4 kHz and output signal-to-noise ratio of 20 dB. The terminal has a total of 128 symbols. Assume that the symbols are equiprobable and the successive transmissions are statistically independent.

(a) Calculate the information capacity of the channel.

(b) Calculate the maximum symbol rate for which error-free transmission over the channel is possible.

10.18 In a *single-parity-check code*, a single parity bit is appended to a block of k message bits $(m_1, m_2, \ldots, m_k)$. The single parity bit, b_1 is chosen so that the codeword satisfies the *even parity rule*:

$$m_1 + m_2 + \cdots + m_k + b_1 = 0, \qquad \text{mod } 2$$

For $k = 3$, set up the 2^k possible codewords in the code defined by this rule.

10.19 Compare the parity-check matrix of the (7, 4) Hamming code considered in Example 10.10 with that of a (4, 1) repetition code.

10.20 Consider the (7, 4) Hamming code of Example 10.10. The generator matrix G and the parity-check matrix H of the code are described in that example. Show that these two matrices satisfy the condition

$$HG^T = 0$$

10.21

(a) For the (7, 4) Hamming code described in Example 10.10, construct the eight codewords in the dual code.

(b) Find the minimum distance of the dual code determined in part (a).

10.22 Consider the (5, 1) repetition code of Example 10.9, Evaluate the syndrome s for the following error patterns:

(a) All five possible single-error patterns

(b) All 10 possible double-error patterns

10.23 A convolutional encoder has a single-shift register with two stages, (i.e., constraint length $K = 3$), three modulo-2 adders, and an output multiplexer. The generator sequences of the encoder are as follows:

$$g^{(1)} = (1, 0, 1)$$
$$g^{(2)} = (1, 1, 0)$$
$$g^{(3)} = (1, 1, 1)$$

Draw the block diagram of the encoder H.

10.24 Consider the rate $r = 1/2$, constraint length $K = 2$ convolutional encoder of Figure P10.24. The code is systematic. Find the encoder output produced by the message sequence 10111.

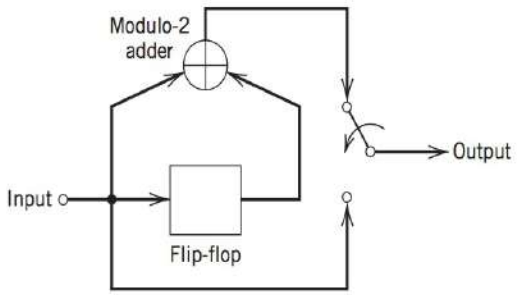


Figure P10.24

10.25 Figure P10.25 shows the encoder for a rate $r = 1/2$, constraint length $K = 4$ convolutional code. Determine the encoder output produced by the message sequence 10111....

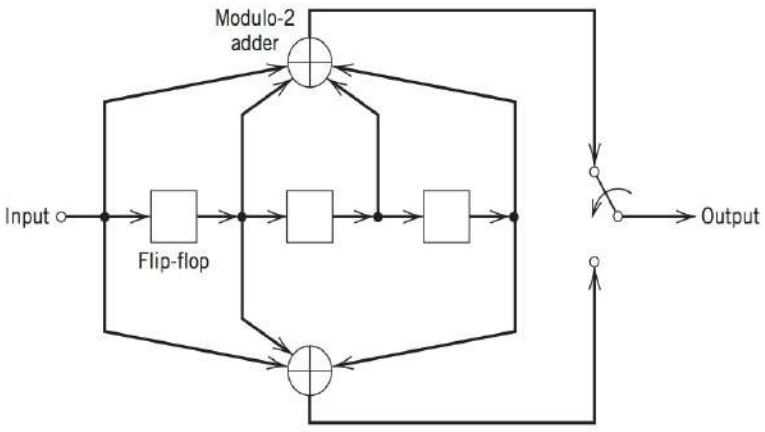


Figure P10.25

10.26 Consider the encoder of Fig 10.26*b* for a rate $r = 2/3$, constraint length $K = 2$ convolutional code. Determine the code sequence produced by the message sequence 10111....

10.27 Consider a rate-1/2, constraint length-7 convolutional code with free distance $d_{\text{free}} = 10$. Calculate the asymptotic coding gain for the following two channels:

(a) Binary symmetric channel
(b) Binary input AWGN channel

10.28 Let $r_c^{(1)} = p/q_1$ and $r_c^{(2)} = p/q_2$ be the code rates of RSC encoders 1 and 2 in the turbo encoder of Figure 10.33. Find the code rate of the turbo code.

10.29 The feedback nature of the constituent codes in the turbo encoder of Figure 10.33 has the following implication: A single bit error corresponds to an infinite sequence of channel errors. Illustrate this phenomenon by using a message sequence consisting of symbol 1 followed by an infinite number of symbols 0.

10.30 Consider the following generator matrices for rate 1/2 turbo codes:

$$\text{4-state encoder:} \quad g(D) = \left[1, \frac{1 + D + D^2}{1 + D^2}\right]$$

$$\text{8-state encoder:} \quad g(D) = \left[1, \frac{1 + D^2 + D^3}{1 + D + D^2 + D^3}\right]$$

$$\text{16-state encoder:} \quad g(D) = \left[1, \frac{1 + D^4}{1 + D + D^2 + D^3 + D^4}\right]$$

(a) Construct the block diagram for each one of these RSC encoders.
(b) Setup the parity-check equation associated with each encoder.

APPENDIX MATHEMATICAL TABLES

TABLE A.1 Summary of properties of the Fourier transform

Property	Mathematical Description
1. Linearity	$ag_1(t) + bg_2(t) \rightleftharpoons aG_1(f) + bG_2(f)$ where a and b are constants
2. Time scaling	$g(at) \rightleftharpoons \frac{1}{\lvert a\rvert} G\left(\frac{f}{a}\right)$ where a is a constant
3. Duality	If $g(t) \rightleftharpoons G(f)$, then $G(t) \rightleftharpoons g(-f)$
4. Time shifting	$g(t - t_0) \rightleftharpoons G(f)\exp(-j2\pi f t_0)$
5. Frequency shifting	$\exp(j2\pi f_c t)g(t) \rightleftharpoons G(f - f_c)$
6. Area under $g(t)$	$\int_{-\infty}^{\infty} g(t)\,dt = G(0)$
7. Area under $G(f)$	$g(0) = \int_{-\infty}^{\infty} G(f)\,df$
8. Differentiation in the time domain	$\frac{d}{dt}g(t) \rightleftharpoons j2\pi f G(f)$
9. Integration in the time domain	$\int_{-\infty}^{t} g(\tau)d\tau \rightleftharpoons \frac{1}{j2\pi f} G(f) + \frac{G(0)}{2}\delta(f)$
10. Conjugate functions	If $g(t) \rightleftharpoons G(f)$, then $g^*(t) \rightleftharpoons G^*(-f)$
11. Multiplication in the time domain	$g_1(t)g_2(t) \rightleftharpoons \int_{-\infty}^{\infty} G_1(\lambda)G_2(f - \lambda)\,d\lambda$
12. Convolution in the time domain	$\int_{-\infty}^{\infty} g_1(\tau)g_2(t - \tau)d\tau \rightleftharpoons G_1(f)G_2(f)$
13. Rayleigh's energy theorem	$\int_{-\infty}^{\infty} \lvert g(t)\rvert^2 dt = \int_{-\infty}^{\infty} \lvert G(f)\rvert^2 df$

TABLE A.2 Fourier-transform pairs

Time Function	Fourier Transform
$\text{rect}\left(\frac{t}{T}\right)$	$T\,\text{sinc}(fT)$
$\text{sinc}(2Wt)$	$\frac{1}{2W}\text{rect}\left(\frac{f}{2W}\right)$
$\exp(-at)u(t), \quad a>0$	$\frac{1}{a+j2\pi f}$
$\exp(-a\lvert t\rvert), \quad a>0$	$\frac{2a}{a^2+(2\pi f)^2}$
$\exp(-\pi t^2)$	$\exp(-\pi f^2)$
$\begin{cases} 1-\frac{\lvert t\rvert}{T}, & \lvert t\rvert < T \\ 0, & \lvert t\rvert \geq T \end{cases}$	$T\,\text{sinc}^2(fT)$
$\delta(t)$	1
1	$\delta(f)$
$\delta(t-t_0)$	$\exp(-j2\pi f t_0)$
$\exp(j2\pi f_c t)$	$\delta(f-f_c)$
$\cos(2\pi f_c t)$	$\frac{1}{2}[\delta(f-f_c)+\delta(f+f_c)]$
$\sin(2\pi f_c t)$	$\frac{1}{2j}[\delta(f-f_c)-\delta(f+f_c)]$
$\text{sgn}(t)$	$\frac{1}{j\pi f}$
$\frac{1}{\pi t}$	$-j\,\text{sgn}(f)$
$u(t)$	$\frac{1}{2}\delta(f)+\frac{1}{j2\pi f}$
$\sum_{i=-\infty}^{\infty}\delta(t-iT_0)$	$\frac{1}{T_0}\sum_{n=-\infty}^{\infty}\delta\left(f-\frac{n}{T_0}\right)$

Notes: $u(t)$ = unit step function
$\delta(t)$ = Dirac delta function
$\text{rect}(t)$ = rectangular function
$\text{sgn}(t)$ = signum function
$\text{sinc}(t)$ = sinc function

TABLE A.3 Summary of Bessel functions

Bessel Functions of the First Kind

1. Equivalent representations

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(jx\sin\theta - jn\theta)d\theta$$
$$= \frac{1}{\pi}\int_{0}^{\pi} \cos(x\sin\theta - n\theta)d\theta$$
$$= \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{2}x\right)^{n+2m}}{m!(n+m)!}$$

2. Properties
   a. $J_n(x) = (-1)^n J_{-n}(x)$
   b. $J_n(x) = (-1)^n J_n(-x)$
   c. $J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x} J_n(x)$
   d. For small x,
   $$J_n(x) \approx \frac{x^n}{2^n n!}$$
   e. For large x,
   $$J_n(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\pi}{4} - \frac{n\pi}{2}\right)$$
   f. For real x,
   $$\lim_{n\to\infty} J_n(x) = 0$$
   g. $$\sum_{n=-\infty}^{\infty} J_n(x)\exp(jn\phi) = \exp(jx\sin\phi)$$
   h. $$\sum_{n=-\infty}^{\infty} J_n^2(x) = 1$$
   i. Matlab function call
   $$J_n(x) = \text{besselj(n, x)}$$

Modified Bessel Functions of the First Kind

1. Equivalent representations

$$I_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(x\cos\theta)\cos(n\theta)d\theta$$
$$= \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}x\right)^{n+2m}}{m!(n+m)!}$$
$$= j^{-n} J_n(jx)$$

2. Properties
   a. For small x,
   $$I_0(x) \approx 1$$
   b. For large x,
   $$I_0(x) \approx \frac{\exp(x)}{\sqrt{2\pi x}}$$
   c. Matlab function call
   $$I_n(x) = \text{besseli(n, x)}$$

TABLE A.4 Summary of Q-function

1. Equivalent representations

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-z^2/2)dz$$
$$= \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)$$

where

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-z^2)dz$$

2. Properties
   a. $Q(-x) = 1 - Q(x)$
   b. For small x,
   $$\lim_{x \to 0} Q(x) = 0.5$$
   c. For large x,
   $$Q(x) \approx \frac{1}{\sqrt{2\pi}x} \exp(-x^2/2)$$
   d. Matlab function call
   $$Q(x) = 0.5*\text{erfc}(x/\text{sqrt}(2))$$

TABLE A.5 Trigonometric identities

$\exp(\pm j\theta) = \cos\theta \pm j \sin\theta$

$\cos\theta = \frac{1}{2}[\exp(j\theta) + \exp(-j\theta)]$

$\sin\theta = \frac{1}{2j}[\exp(j\theta) - \exp(-j\theta)]$

$\sin^2\theta + \cos^2\theta = 1$

$\cos^2\theta - \sin^2\theta = \cos(2\theta)$

$\cos^2\theta = \frac{1}{2}[1 + \cos(2\theta)]$

$\sin^2\theta = \frac{1}{2}[1 - \cos(2\theta)]$

$2\sin\theta\cos\theta = \sin(2\theta)$

$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$

$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$

$\tan(\alpha \pm \beta) = \frac{\tan\alpha \pm \tan\beta}{1 \mp \tan\alpha\tan\beta}$

$\sin\alpha\sin\beta = \frac{1}{2}[\cos(\alpha - \beta) - \cos(\alpha + \beta)]$

$\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha - \beta) + \cos(\alpha + \beta)]$

$\sin\alpha\cos\beta = \frac{1}{2}[\sin(\alpha - \beta) + \sin(\alpha + \beta)]$

TABLE A.6 Series expansions

Taylor series

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

where

$$f^{(n)}(a) = \left.\frac{d^n f(x)}{dx^n}\right|_{x=a}$$

MacLaurin series

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots$$

where

$$f^{(n)}(0) = \left.\frac{d^n f(x)}{dx^n}\right|_{x=0}$$

Binomial series

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots, \qquad |nx| < 1$$

Exponential series

$$\exp(x) = 1 + x + \frac{1}{2!}x^2 + \cdots$$

Logarithmic series

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots$$

Trigonometric series

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots$$

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots$$

$$\sin^{-1} x = x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \cdots$$

$$\tan^{-1} x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots, \qquad |x| < 1$$

$$\operatorname{sinc} x = 1 - \frac{1}{3!}(\pi x)^2 + \frac{1}{5!}(\pi x)^4 - \cdots$$

TABLE A.7 Summations

$$\sum_{k=1}^{K} k = \frac{K(K+1)}{2}$$

$$\sum_{k=1}^{K} k^2 = \frac{K(K+1)(2K+1)}{6}$$

$$\sum_{k=1}^{K} k^3 = \frac{K^2(K+1)^2}{4}$$

$$\sum_{k=0}^{K-1} x^k = \frac{(x^K - 1)}{x - 1}, \quad |x| \neq 1$$

TABLE A.8 Integrals

Indefinite integrals

$$\int x \sin(ax)\, dx = \frac{1}{a^2}[\sin(ax) - ax\cos(ax)]$$

$$\int x \cos(ax)\, dx = \frac{1}{a^2}[\cos(ax) + ax\sin(ax)]$$

$$\int x \exp(ax)\, dx = \frac{1}{a^2}\exp(ax)(ax - 1)$$

$$\int x \exp(ax^2)\, dx = \frac{1}{2a}\exp(ax^2)$$

$$\int \exp(ax)\sin(bx)\, dx = \frac{1}{a^2 + b^2}\exp(ax)[a\sin(bx) - b\cos(bx)]$$

$$\int \exp(ax)\cos(bx)\, dx = \frac{1}{a^2 + b^2}\exp(ax)[a\cos(bx) + b\sin(bx)]$$

$$\int \frac{dx}{a^2 + b^2x^2} = \frac{1}{ab}\tan^{-1}\left(\frac{bx}{a}\right)$$

$$\int \frac{x^2\, dx}{a^2 + b^2x^2} = \frac{x}{b^2} - \frac{a}{b^3}\tan^{-1}\left(\frac{bx}{a}\right)$$

Definite integrals

$$\int_0^\infty \frac{x\sin(ax)}{b^2 + x^2}\, dx = \frac{\pi}{2}\exp(-ab), \qquad a > 0, b > 0$$

$$\int_0^\infty \frac{\cos(ax)}{b^2 + x^2}\, dx = \frac{\pi}{2b}\exp(-ab), \qquad a > 0, b > 0$$

$$\int_0^\infty \frac{\cos(ax)}{(b^2 - x^2)^2}\, dx = \frac{\pi}{4b^3}[\sin(ab) - ab\cos(ab)], \qquad a > 0, b > 0$$

$$\int_0^\infty \operatorname{sinc} x\, dx = \int_0^\infty \operatorname{sinc}^2 x\, dx = \frac{1}{2}$$

$$\int_0^\infty \exp(-ax^2)dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}, \qquad a > 0$$

$$\int_0^\infty x^2 \exp(-ax^2)dx = \frac{1}{4a}\sqrt{\frac{\pi}{a}}, \qquad a > 0$$

TABLE A.9 Useful constants

Physical Constants	
Boltzmann's constant	$k = 1.38 \times 10^{-28}$ joule/degree Kelvin $= -228.6$ dBW K^{-1}
Planck's constant	$h = 6.626 \times 10^{-34}$ joule-second
Electron (fundamental) charge	$q = 1.602 \times 10^{-19}$ coulomb
Speed of light in vacuum	$c = 2.998 \times 10^{8}$ meters/second
Standard (absolute) temperature	$T_0 = 273$ degree Kelvin
Thermal voltage	$V_T = 0.026$ volt at room temperature
Thermal energy kT at standard temperature	$kT_0 = 3.77 \times 10^{-21}$ joule $= -204.2$ dBW Hz^{-1}
One Hertz (Hz) = 1 cycle/second; 1 cycle = 2π radians	
One watt (W) = 1 joule/second	
Mathematical Constants	
Base of natural logarithm	$e = 2.7182818$
Logarithm of e to base 2	$\log_2 e = 1.442695$
Logarithm of 2 to base e	In $2 = 0.693147$
Logarithm of 2 to base 10	$\log_{10} 2 = 0.30103$
Pi	$\pi = 3.1415927$

TABLE A.10 Recommended unit prefixes

Multiples and Submultiples	Prefixes	Symbols
10^{12}	tera	T
10^{9}	giga	G
10^{6}	mega	M
10^{3}	kilo	K *(k)*
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p

GLOSSARY

CONVENTIONS AND NOTATIONS

1. The symbol $|\ |$ means the magnitude of the complex quantity contained within.
2. The symbol arg( ) means the phase angle of the complex quantity contained within.
3. The symbol Re[ ] means the "real part of" and Im[ ] means the "imaginary part of."
4. The symbol ln( ) denotes the natural logarithm of the quantity contained within, whereas the logarithm to base a is denoted by $\log_a(\ )$.
5. The use of an asterisk as superscript denotes complex conjugate, e.g., x^* is the complex conjugate of x.
6. The symbol $\rightleftharpoons$ indicates a Fourier-transform pair, e.g., $g(t) \rightleftharpoons G(f)$, where a lowercase letter denotes the time function and a corresponding uppercase letter denotes the frequency function.
7. The symbol $F[\ ]$ indicates the Fourier-transform operation, e.g., $F[g(t)] = G(f)$, and the symbol $F^{-1}[\ ]$ indicates the inverse Fourier-transform operation, e.g., $F^{-1}[G(f)] = g(t)$.
8. The symbol $\star$ denotes convolution, e.g.,
$$x(t) \star h(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau$$
9. The symbol $\oplus$ denotes modulo-2 addition; except in Chapter 10 where modulo-2 addition is denoted by an ordinary plus sign.
10. The use of subscript T_0 indicates that the pertinent function $g_{T_0}(t)$, say, is a periodic function of time t with period T_0.
11. The use of a hat over a function indicates the estimate of an unknown parameter, e.g., the quantity $\hat{\alpha}(\mathbf{x})$ is an estimate of the unknown parameter α, based on the observation vector $\mathbf{x}$.
12. The use of a tilde over a function indicates the complex envelope of a narrowband signal, e.g., the function $\tilde{g}(t)$ is the complex envelope of the narrowband signal $g(t)$.
13. The use of subscripts I and Q indicates the in-phase and quadrature components of a narrowband signal, a narrow-band random process, or the impulse response of a narrowband filter, with respect to the carrier $\cos(2\pi f_c t)$.
14. For a low-pass message signal, the highest frequency component or message bandwidth is denoted by W. The spectrum of this signal occupies the frequency interval $-W \leq f \leq W$ and is zero elsewhere. For a band-pass signal with carrier frequency f_c, the spectrum occupies the frequency intervals, $f_c - W \leq f \leq f_c + W$ and $-f_c - W \leq f \leq -f_c + W$, and so $2W$ denotes the bandwidth of the signal. The (low-pass) complex envelope of this band-pass signal has a spectrum that occupies the frequency interval $-W \leq f \leq W$.

For a low-pass filter, the bandwidth is denoted by B. A common definition of filter bandwidth is the frequency at which the amplitude response of the filter drops by 3 dB below the zero-frequency value. For a band-pass filter of mid-band frequency f_c the bandwidth is denoted by $2B$ centered on f_c. The complex low-pass equivalent of this band-pass filter has a bandwidth equal to B.

The transmission bandwidth of a communication channel, required to transmit a modulated wave, is denoted by B_T.

15. Random variables or random vectors are uppercase (e.g., X or $\mathbf{X}$), and their sample values are lowercase (e.g., x or $\mathbf{x}$).
16. A vertical bar in an expression means "given that," e.g., $f_x(x|H_0)$ is the probability density function of the random variable X, given that hypothesis H_0 is true.
17. The symbol $\mathbf{E}[\,]$ means the expected value of the random variable enclosed within.
18. The symbol var[ ] means the variance of the random variable enclosed within.
19. The symbol cov[ ] means the covariance of the two random variables enclosed within.
20. The average probability of symbol error is denoted by P_e.

    In the case of binary signaling techniques, P_{e0} denotes the conditional probability of error given that symbol 0 was transmitted, and P_{e1} denotes the conditional probability of error given that symbol 1 was transmitted. The *a priori* probabilities of symbols 0 and 1 are denoted by p_0 and p_1, respectively.
21. Boldface letter denotes a vector or matrix. The inverse of a square matrix $\mathbf{R}$ is denoted by $\mathbf{R}^{-1}$. The transpose of a vector $\mathbf{w}$ is denoted by $\mathbf{w}^T$.
22. The length of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\|$. The Euclidean distance between the vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ is denoted by $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$.
23. The inner product of two vectors $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\mathbf{x}^T\mathbf{y}$; their outer product is denoted by $\mathbf{x}\mathbf{y}^T$.

FUNCTIONS

1. Rectangular function:
$$\text{rect}(t) = \begin{cases} 1, & -\frac{1}{2} < t < \frac{1}{2} \\ 0, & |t| \geq \frac{1}{2} \end{cases}$$

2. Unit step function:
$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

3. Signum function:
$$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \end{cases}$$

4. Dirac delta function:
$$\delta(t) = 0, \quad t \neq 0$$
$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1$$
or equivalently
$$\int_{-\infty}^{\infty} g(t)\, \delta(t - t_0)\, dt = g(t_0)$$

5. Sinc function:
$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

6. Q function:
$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^{\infty} \exp(-z^2/2)\, dz$$
Complementary error function:
$$\text{erfc}(u) = 1 - \text{erf}(u).$$

7. Bessel function of the first kind of order n:
$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(jx \sin\theta - jn\theta)\, d\theta$$

8. Modified Bessel function of the first kind of zero order:
$$I_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(x\cos\theta)\, d\theta$$

9. Binomial coefficient
$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

ABBREVIATIONS

ac:	alternating current
ANSI:	American National Standards Institute
AM:	amplitude modulation
ARQ:	automatic-repeat-request
ASCII:	American National Standard Code for Information Interchange
ASK:	amplitude-shift keying
ATM:	asynchronous transfer mode
BER:	bit error rate
BPF:	band-pass filter
BPSK:	binary phase shift keying
BSC:	binary symmetric channel
CCD:	charge-coupled device
CCITT:	Consultative Committee for International Telephone and Telegraph
CPFSK:	continuous-phase frequency-shift keying
CW:	continuous wave
dB:	decibel
dc:	direct current
DFT:	discrete Fourier transform
DM:	delta modulation
DPCM:	differential pulse-code modulation
DPSK:	differential phase-shift keying
DSB-SC:	double sideband-suppressed carrier
exp:	exponential
FDM:	frequency-division multiplexing
FDMA:	frequency-division multiple access
FFT:	fast Fourier transform
FMFB:	frequency modulator with feedback
FSK:	frequency-shift keying
HDTV:	high definition television
Hz:	Hertz
IDFT:	inverse discrete Fourier transform
IF:	intermediate frequency
I/O:	input/output
ISI:	intersymbol interference
ISO:	International Organization for Standardization
LAN:	local-area network
LED:	light emitting diode
LMS:	least-mean-square
ln:	natural logarithm
log:	logarithm
LPF:	low-pass filter
MAP:	maximum *a posteriori* probability
ms:	millisecond
μs:	microsecond
ML:	maximum likelihood
modem:	modulator–demodulator
MSK:	minimum shift keying

nm:	nanometer
NRZ:	nonreturn-to-zero
NTSC:	National Television Systems Committee
OOK:	on–off keying
OSI:	open systems interconnection
PAM:	pulse-amplitude modulation
PCM:	pulse-code modulation
PCN:	personal communication network
PLL:	phase-locked loop
PN:	pseudo-noise
PSK:	phase-shift keying
QAM:	quadrature amplitude modulation
QOS:	quality of service
QPSK:	quadriphase-shift keying
RF:	radio frequency
rms:	root-mean-square
RS:	Reed–Solomon
RS-232	Recommended standard-232 (port)
RZ:	return-to-zero
s:	second
SDH:	synchronous digital hierarchy
SDR:	signal-to-distortion ratio
SONET:	synchronous optical network
SNR:	signal-to-noise ratio
TCM:	trellis-coded modulation
TDM:	time-division multiplexing
TDMA:	time-division multiple access
TV:	television
UHF:	ultra high frequency
VCO:	voltage-controlled oscillator
VHF:	very high frequency
VLSI:	very-large-scale integration

BIBLIOGRAPHY

BOOKS

N. Abramson, *Information Theory and Coding* (New York: McGraw-Hill, 1963).

J. Adamek, *Foundations of Coding* (New York: Wiley, 1991).

Bell Telephone Laboratories, *Transmission Systems for Communications* (1971).

S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory* (Englewood Cliffs, N.J.: Prentice-Hall, 1987).

W. R. Bennett, *Introduction to Signal Transmission* (New York: McGraw-Hill, 1970).

R. E. Best, *Phase-locked Loops: Design, simulation and applications*, 5th ed. (New York: McGraw-Hill, 2003).

E. Biglieri, D. Divsalar, P. J. Mclane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications* (New York: Macmillan, 1991).

H. S. Black, *Modulation Theory* (Princeton, N.J.: Van Nostrand, 1953).

R. E. Blahut, *Principles and Practice of Information Theory* (Reading, Mass: Addison-Wesley, 1987).

G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control* (San Francisco: Holden-Day, 1976).

R. N. Bracewell, *The Fourier Transform and Its Applications*, 2nd ed., rev. (New York: McGraw-Hill, 1986).

L. Brillouin, *Science and Information Theory*, 2nd ed. (New York: Academic Press, 1962).

K. W. Cattermole, *Principles of Pulse-code Modulation* (New York: American Elsevier, 1969).

D. C. Champeney, *Fourier Transforms and Their Physical Applications* (London: Academic Press, 1973).

L. Cohen, *Time-frequency analysis* (New Jersey: Prentice Hall, 1994).

T. M. Cover and J. B. Thomas, *Elements of Information Theory* (New York: Wiley, 1991).

R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing* (Englewood Cliffs, N.J.: Prentice-Hall, 1983).

W. F. Egan, *Phase-Lock Basics* (New York: Wiley, 1998).

D. F. Elliott and K. R. Rao, *Fast Transforms: Algorithms, Analyses, Applications* (New York: Academic Press, 1982).

L. E. Franks, *Signal Theory* (Englewood Cliffs, N.J.: Prentice-Hall, 1969).

R. G. Gallagher, *Information Theory and Reliable Communication* (New York: Wiley, 1968).

F. M. Gardner, *Phaselock Techniques*, 2nd ed. (New York: Wiley, 1979).

J. D. Gibson, *Principles of Digital and Analog Communications* (New York: Macmillan, 1989).

R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Data Communications Principles* (New York: Plenum, 1992).

R. M. Gray and L. D. Davisson, *Random Processes: A Mathematical Approach for Engineers* (Englewood Cliffs, N.J.: Prentice-Hall, 1986).

M. S. Gupta (editor), *Electrical Noise: Fundamentals and Sources* (New York: IEEE Press, 1977).

R. W. Hamming, *The Art of Probability for Scientists and Engineers* (Reading, Mass.: Addison-Wesley, 1991).

R. W. Hamming, *Coding and Information Theory* (Englewood Cliffs, N.J.: Prentice-Hall, 1980).

S. Haykin. *Adaptive Filter Theory*, 2nd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1991).

S. Haykin, *Communication Systems* 4th ed. (New York: Wiley, 2001).

S. Haykin and M. Moher, *Introduction to Analog and Digital Communications*, 2nd ed. (New Jersey: Wiley, 2007).

S. Haykin and M. Moher, *Modern Wireless Communications* (New Jersey: Prentice Hall, 2005).

S. Haykin and B. Van Veen, *Signals and Systems* 2nd ed., (New York: Wiley, 2003).

C. Heegard and S. B. Wicker, *Turbo Coding* (Boston: Kluwer, 1999).

C. W. Helstrom, *Probability and Stochastic Processes for Engineers*, 2nd ed. (New York: Macmillan, 1990).

N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video* (Englewood Cliffs, N.J.: Prentice-Hall, 1984).

M. C. Jeruchim, B. Balaban, and J. S. Shanmugan, *Simulation of Communication Systems* (New York: Plenum, 1992).

S. M. Kay, *Modern Spectral Estimation: Theory and Applications* (Englewood Cliffs, N.J.: Prentice-Hall, 1988).

G. Keiser, *Optical Fiber Communications*, 3rd ed. (New York: McGraw-Hill, 2000).

E. A. Lee and D. G. Messerschmitt, *Digital Communications*, 2nd ed. (Boston: Kluwer Academic, 1994).

A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering* (Reading, Mass.: Addison-Wesley, 1989).

S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, 2nd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 2004).

W. C. Lindsey, *Synchronization Systems in Communication and Control* (Englewood Cliffs, N.J.: Prentice-Hall, 1972).

R. W. Lucky, *Silicon Dreams: Information, Man, and Machine* (New York: St. Martin's Press, 1989).

R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication* (New York: McGraw-Hill, 1968).

F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-correcting Codes* (Amsterdam: North-Holland, 1977).

R. J. Marks, *Introduction to Shannon Sampling and Interpolation Theory* (New York/Berlin: Springer-Verlag, 1991).

S. L. Marple, *Digital Spectral Analysis with Applications* (Englewood Cliffs, N.J.: Prentice-Hall, 1987).

R. J. McEliece, *The Theory of Information and Coding*, 2nd ed. (Cambridge: Cambridge University Press, 2002).

A. M. Michelson and A. H. Levesque, *Error-control Techniques for Digital Communication* (New York: Wiley, 1985).

A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. (New Jersey: Prentice Hall, 1999).

A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Englewood Cliffs, N.J.: Prentice-Hall, 1975).

A. Papoulis, *Probablity, Random Variables, and Stochastic Processes*, 2nd ed. (New York: McGraw-Hill, 1984).

J. D. Parsons, The Mobile Radio Propagation Channel (New York: Wiley, 1992).

W. W. Peterson and E. J. Weldon, Jr., *Error Correcting Codes*, 2nd ed. (Boston: MIT Press, 1972).

J. R. Pierce, *Symbols, Signals and Noise: The Nature and Process of Communication* (New York: Harper 1961).

W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. VeHerling, (editors), *Numerical Recipes in C: The Art of Scientific Computing* (New York: Cambridge University Press, 1988).

J. G. Proakis, *Digital Communications*, 2nd ed. (New York: McGraw-Hill, 1989).

L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing* (Englewood Cliffs, N.J.: Prentice-Hall, 1975).

J. H. Reed, *Software Radio: A Modern Approach to Radio Engineering* (New Jersey: Prentice Hall, 2002).

J. H. Roberts, *Angle Modulation: The Theory of System Assessment*, IEE Communication Series 5 (London: Institution of Electrical Engineers, 1977).

H. E. Rowe, *Signals and Noise in Communication Systems* (Princeton, N.J.: Van Nostrand, 1965).

T. S. Rzeszewski (editor), *Television Technology Today* (New York: IEEE Press, 1985).

D. Salomon, G. Motta, and D. Bryant, *Data Compression: The Complete Reference*, 4th ed., (London: Springer, 2006).

C. Schlegel, *Trellis Coding*, (New Jersey: IEEE Press, 1997).

M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques* (New York: McGraw-Hill, 1966).

J. M. Senior, *Optical Fiber Communications: Principles and Practice*, 2nd ed. (Englewood Cliffs, N.J.: Prentice Hall, 1992).

B. Sklar, *Digital Communications: Fundamentals and Applications* 2nd ed. (New Jersey: Prentice Hall, 2001).

F. G. Stremler, *Introduction to Communication Systems*, 3rd ed. (Reading, M. A.: Addison-Wesley, 1990).

E. D. Sunde, *Communication Systems Engineering Theory* (New York: Wiley, 1969).

A. S. Tannenbaum, *Computer Networks*, 3rd ed. (Englewood Cliffs, N.J.: Prentice Hall, 2005).

T. M. Thompson, *From Error-correcting Codes through Sphere Packing to Simple Groups* (The Mathematical Association of America, 1983).

A. Van der Ziel, *Noise: Source, Characterization, Measurement* (Englewood Cliffs, N.J.: Prentice-Hall, 1970).

H. F. Vanlandingham, *Introduction to Digital Control Systems* (New York: Macmillan, 1985).

A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding* (New York: McGraw-Hill, 1979).

A. D. Whalen, *Detection of Signals in Noise* (New York: Academic Press, 1971).

B. Widrow and S. D. Stearns, *Adaptive Signal Processing* (Englewood Cliffs, N.J.: Prentice-Hall, 1985).

S. G. Wilson, *Data Modulation and Coding* (Englewood Cliffs, N.J.: Prentice Hall, 1996).

C. R. Wylie and L. C. Barrett, *Advanced Engineering Mathematics*, 5th ed. (New York: McGraw-Hill, 1982).

R. E. Ziemer and W. H. Tranter, *Principles of Communications*, 3rd ed. (Boston: Houghton Mifflin, 1990).

PAPERS/REPORTS/PATENTS

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 Trans. on Communication Systems*, vol. CS-10, pp. 336–372, 1962.

G. Battail, "Pondération des symbols décodés par l'algorithme de Viterbi," *Ann. Télécommunication*, vol. 42, pp. 31–38, 1987.

C. Berrou, A. Glavieux, and P. Thitmajshima, "Near Shannon limit error-correction coding and decoding: turbo codes," *Int. Conf. Communications*, pp. 1064–1090, Geneva, Switzerland, May 1993.

C. Berrou and A. Glavieux, "Near optimum error correction coding and decoding: turbo codes," *IEEE Trans. Communications*, vol. 44, pp. 1261–1271, 1996.

V. K. Bhargava, "Forward error correction schemes for digital communications," *IEEE Communications Magazine*, vol. 21, no. 1, pp. 11–19, 1983.

D. R. Brillinger, "An introduction to polyspectra," *Annals of Mathematical Statistics*, pp. 1351–1374, 1965.

D. Cassioli, M. Z. Win, and A. F. Molisch, "The Ultra-Wide Bandwidth Indoor Channel: From Statistical Model to Simulations," *IEEE J. Selected Areas in Commun.*, vol. 20, pp. 1247–1257, 2002.

K. Challapali, X. Lebegue, J. S. Lim, W. H. Paik, and P. A. Snopko, "The Grand Alliance system for US HDTV," Proc. IEEE, Vol. 83, No. 2, February 1995, Pages: 158–174.

J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, pp. 297–801, 1965.

R. deBuda, "Coherent demodulation of frequency-shift keying with low deviation ratio," *IEEE Trans. on Communications*, vol. COM-20, pp. 429–535, 1972.

M. I. Doelz and E. H. Heald, "Minimum shift data communication system," U.S. Patent No. 2977417, March 1961.

L. H. Enloe, "Decreasing the threshold in FM by frequency feedback," *Proceedings of the IRE*, vol. 50, pp. 18–30, 1962.

W. A. Gardner and L. E. Franks,"Characterization of cyclostationary random signal processes," *IEEE Transactions on Information Theory*, vol. IT-21, pp. 4–14, 1975.

J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," *IEEE Globecom 89*, pp. 47.11–47.17, November 1989, Dallas, Texas.

F. S. Hill, Jr., "On time-domain representations for vestigial sideband signals," *Proceedings of the IEEE*, vol. 62, pp. 1032–1033, 1974.

H. Kaneko, "A unified formulation of segment companding laws and synthesis of codes and digital companders," *Bell System Tech. J.*, vol. 49, pp. 1555–1588, 1970.

C. F. Kurth, "Generation of single-sideband signals in multiplex communication systems," *IEEE Transactions on Circuits and Systems*, vol. CAS-23, pp. 1–17, Jan. 1976.

C. L. Nikas and M. R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proceedings of the IEEE*, vol. 75, pp. 869–891, 1987.

S. Pasupathy, "Minimum shift keying—A spectrally efficient modulation," *IEEE Communications Magazine*, vol. 17, no. 4, pp. 14–22, 1979.

S. O. Rice, "Noise in FM receivers," in M. Rosenblatt, (editor), *Proceedings of the Symposium on Time Series Analysis* (New York: Wiley, 1963), pp. 395–411.

T. Sikora, "MPEG Digital Video-coding standards," *IEEE Signal Processing Magazine*, pp. 82–99, September 1997.

M. K. Simon and D. Divsalar, "On the implementation and performance of single and double differential detection schemes," *IEEE Trans. on Communications*, vol. 40, pp. 278–291, 1992.

B. Smith, "Instantaneous compounding of quantized signals," *Bell System Tech, J.*, vol. 36, pp. 653–709, 1957.

G. L. Turin, "An introduction to matched filters," *IRE Transactions on Information Theory*, vol. IT-6, pp. 311–329, 1960.

G. L. Turin, "An introduction to digital matched filters," *Proceedings of the IEEE*, vol. 64, pp. 1092–1112, 1976.

M. Z. Win and R. A. Scholtz, "Impulse radio: how it works," *IEEE Comm. Letters*, vol. 2, pp. 36–38, 1998.

A. D. Wyner, "Fundamental limits in information theory," *Proceedings of the IEEE*, vol. 69, pp. 239–251, 1981.

INDEX

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V