Modern Wireless Communications

SIMON HAYKIN • MICHAEL MOHER


THE LIBRARY OF CONGRESS
1800

Modern Wireless Communications

Simon Haykin
McMaster University
Hamilton, Ontario, Canada

and

Michael Moher
Space-Time DSP Inc.
Ottawa, Ontario, Canada

Upper Saddle River, NJ 07458

Library of Congress Cataloging-in-Publication

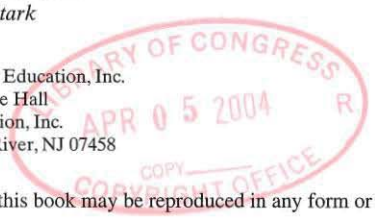
Haykin, Simon S.,
Modern wireless communications / Simon Haykin and Michael Moher.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-022472-3
1. Wireless communication systems. 2. Spread spectrum communications. I. Moher, Michael. II. Title

TK5103.2.H39 2003
621.382--dc22

2003061139

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Vice President and Director of Production and Manufacturing, ESM: *David W. Riccardi*
Executive Managing Editor: *Vince O'Brien*
Managing Editor: *David A. George*
Production Editor: *Craig Little*
Director of Creative Services: *Paul Belfanti*
Art Director: *Jayne Conte*
Cover Designer: *Bruce Kenselaar*
Art Editor: *Greg Dulles*
Manufacturing Manager: *Trudy Pisciotti*
Manufacturing Buyer: *Lisa McDowell*
Marketing Manager: *Holly Stark*

Pearson Prentice Hall
Pearson Education, Inc.
Upper Saddle River, NJ 07458



Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-022472-3

Pearson Education Ltd., *London*
Pearson Education Australia Pty. Ltd., *Sydney*
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd., *Hong Kong*
Pearson Education Canada, Inc., *Toronto*
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education—Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd.
Pearson Education, Inc., *Upper Saddle River, New Jersey*

In Memory of William S. Hart

and

To Dianne

Contents

Preface

The rapid growth of wireless communications and its pervasive use in all walks of life are changing the way we communicate in some fundamental ways. Most important, reliance on radio propagation as the physical mechanism responsible for the transport of information-bearing signals from the transmitter to the receiver has endowed communications with a distinctive feature, namely, mobility.

Modern Wireless Communications is a new book aimed at the teaching of a course that could follow a traditional course on communication systems, as an integral part of an undergraduate program in electrical engineering or as the first graduate course on wireless communications. The primary focus of the book is on the physical layer, emphasizing the fundamentals of radio propagation and communication-theoretic aspects of multiple-access techniques. Many aspects of wireless communications are covered in an introductory level and book form for the first time.

1. ORGANIZATION OF THE BOOK

The book is organized in seven chapters, nine appendices, and a bibliography.

Chapter 1 motivates the study of wireless communications. It begins with a brief historical account of wireless communications, and then goes on to describe the OSI model of communication networks. The discussion, however, focuses on the issues that arise in the study of the physical layer, which is the mainstay of the book.

Chapter 2 on radio propagation starts with an explanation of the physical mechanisms of the propagation process, including free-space propagation, reflection, and diffraction. These physical mechanisms provide insight into the statistical models that are employed for terrestrial and indoor propagation effects that follow. The small-scale effects of fading and uncorrelated scattering are discussed, leading up to a careful classification of the different wireless channel types. The second half of the chapter describes noise and interference, and how combined with propagation, we may determine wireless communication system performance through a link-budget analysis.

Chapter 3 reviews the modulation process with emphasis on digital transmission techniques. This introductory treatment of modulation paves the way for discussions of the following issues:

- Complex baseband representation of linear modulated signals, and the corresponding input/output descriptions of linear wireless communication channels and linear band-pass filters.
- Practical problems concerning adjacent channel interference and nonlinearities in transmit power amplifiers.

The stage is then set for comparative evaluation of various modulation strategies for wireless communications, discussion of receiver performance in the presence of channel noise and Rayleigh fading, and discussion of frequency-division multiple-access (FDMA).

Chapter 4 focuses on coding techniques and time-division multiple-access (TDMA). After a brief review of Shannon's classical information theory, the source coding of speech signals is discussed, which is then followed by fundamental aspects of convolutional codes, interleavers, and turbo codes. The relative merits of convolutional codes and turbo codes are discussed in the context of wireless communications. The various aspects of channel-estimation, tracking, and channel equalization are treated in detail. The discussion then moves onto TDMA and the advantages it offers over FDMA.

Chapter 5 discusses spread spectrum, code-division multiple-access (CDMA), and cellular systems. It first presents the basics of spread-spectrum systems, namely, direct-sequence, and frequency-hopped systems, and their tolerance to interference. A fundamental component of spread-spectrum systems is the spreading code: a section of the chapter is devoted to explaining Walsh-Hadamard, maximal-length sequences, Gold codes, and random sequences. This discussion is then followed with a description of RAKE receivers, channel estimation, code synchronization, and the multipath performance of direct-sequence systems. This leads naturally to a discussion of how direct-sequence systems perform in a cellular environment.

Chapter 6 is devoted to the notion of space diversity and related topics. It starts with diversity on receive, which represents the traditional technique for mitigating the fading problem that plagues wireless communications. Then the chapter introduces the powerful notion of multiple-input, multiple-output (MIMO) wireless communications, which includes space diversity on receive and space diversity on transmit as special cases. Most important, the use of MIMO communications represents the "spatial frontier" of wireless communications in that, for prescribed communication resources in the form of fixed transmit power and channel bandwidth, it provides the practical means for significant increases in the spectral efficiency of wireless communications at the expense of increased computational complexity. The discussion of MIMO wireless communications also includes orthogonal space-time block codes (STBC), best exemplified by the Alamouti code and its differential form. The discussion then moves onto space-division multiple access (SDMA), and smart antennas.

Chapter 7 links the physical layer and multiple-access topics of the previous chapters with the higher layers of the communications network. This final chapter of the book begins with a comparison of the different multiple-access strategies. The discussion then leads to a consideration of various link-management functions associated with wireless systems, namely, signaling, power control, and handover. The differences between systems used for telephony and those used for data transmission are clearly delineated. This is then followed by a discussion of wireless network architectures, both for telephony and data applications.

1.1 Theme Examples

An enriching feature of the book is the inclusion of Theme Examples within each of the chapters in the book, except for Chapter 1. In a loose sense, they may be viewed as

"Chapters within Chapters" that show the practical applications of the topics discussed in the pertinent chapters. Specifically, the following Theme Examples are discussed:

Chapter 2: Empirical propagation model, wireless local area networks (LANs), and impulse radio and ultra-wideband

Chapter 4: Global system for mobile (GSM) communications, joint equalization and decoding, and random-access techniques

Chapter 5: Code-division multiple access (CDMA) Standard IS-95, GPSS, bluetooth, wideband CDMA and WiFi

Chapter 6: BLAST architectures, diversity, space-time block codes, and V-BLAST, and keyhole channels

Chapter 7: Wireless telephone network standards, wireless data network standards, and IEEE 801.11 MAC

1.2 Appendices

To provide supplementary material for the book, nine appendices are included:

- Fourier theory
- Bessel functions
- Random variables and random processes
- Matched filters
- Error function
- Maximum a posteriori probability (MAP) decoding
- Capacity of MIMO links
- Eigendecomposition
- Adaptive antenna array

The inclusion of these appendices is intended to make the book essentially self-sufficient.

1.3 Other Features of the Book

Each chapter includes "within-text" problems that are intended to help the reader develop an improved understanding of the issues being discussed in the text. "End-of-chapter" problems provide an abundance of additional problems, whose solutions will further help the reader develop a deeper understanding of the material covered in the pertinent chapter.

Moreover, each chapter includes examples with detailed solutions covering different aspects of the subject matter.

"Notes and References" included at the end of the chapter provide explanatory notes, and they guide the reader to related references for further reading. All the references so made are assembled in the Bibliography placed at the end of the book.

2. SUPPLEMENTARY MATERIAL ON THE BOOK

The "within-text" problems are provided with answers alongside the problems. The solutions to the additional "end-of-chapter" problems are assembled in the Instructor's Manual, copies of which are available to qualified instructors. Contact your Pearson Prentice Hall representative.

The MATLAB codes used to plot results of the experiments and the graphs in the book are also included in the Instructor's Manual. M files are available on the Companion Website at

`http://www.prenhall.com/haykin`

Electronic copies in JPG format of selected figures from the book can also be accessed on the website. This material can be used for PowerPoint™ presentations of lectures based on the book.

3. ACKNOWLEDGMENTS

We are indebted to many colleagues for background material and insightful inputs, which, in one way or another, have helped us write certain parts of the book. In this context, the inputs of the following contributors (in alphabetical order) are gratefully acknowledged:

Dr. Claude Berrou, *ENST Bretagne, Brest, France*
Dr. Stewart Crozier, *Communications Research Centre, Ottawa, Ontario*
Dr. Suhas Diggavi, *Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland*
Dr. David Gesbert, *University of Oslo, Sweden*
Dr. Lajos Hanzo, *University of Southampton, United Kingdom*
Dr. John Lodge, *Communications Research Centre, Ottawa, Ontario*
Dr. Mathini Sellathurai, *Communications Research Centre, Ottawa, Ontario*
Dr. Abbas Yongaçoglu, *University of Ottawa, Ottawa, Canada*
Dr. Mansoor Shafi, *Telecom, Wellington, New Zealand*

The critical inputs and recommendations made by several anonymous reviewers of early versions of the manuscript are deeply appreciated. They provided us with a great deal of food for thought.

Moreover, we are grateful to Kris Huber, McMaster University, and Blair Simpson, Space-Time DSP Inc., for their assistance in generating many of the figures in the book.

We thank the Institute of Electrical and Electronics Engineers, Inc. and the European Telecommunications Standards Institute, for giving us the permissions to produce certain figures in the book.

We are particularly grateful to Tom Robbins, Craig Little, and the production staff at Prentice Hall, Pearson Education for their contributions, which, individually and collectively, have indeed made the production of the book possible.

The continued support and encouragement of our respective wives, Nancy and Dianne, and our families were an essential ingredient in the completion of the book. We are grateful to them all.

Last but by no means least, we are deeply grateful to Lola Brooks, McMaster University, for her tireless effort in all matters relating to the preparation and production of many different versions of the manuscript for the book.

SIMON HAYKIN
Ancaster, Ontario

MICHAEL MOHER
Ottawa, Ontario

CHAPTER 1

Introduction

You see, wire telegraph is a kind of a very, very long cat. You pull his tail in New York and his head is meowing in Los Angeles. Do you understand this? And radio operates exactly the same way: you send signals here, they receive them there. The only difference is that there is no cat. *Albert Einstein (1879–1955)*

1.1 BACKGROUND

The understanding of radio waves is fundamental to wireless communications, but simply knowing that electromagnetic waves exist is a relatively recent historical event. In the short period since that time, there have been numerous milestones in the development of radio communications. Some of these milestones are the following:

- In 1864, James Clerk Maxwell formulated the electromagnetic theory of light and *predicted the existence of radio waves.* In his honor, the set of equations basic to the propagation of electromagnetic waves is known as *Maxwell's equations.*
- The physical existence of radio waves was first demonstrated by Heinrich Hertz in 1887.
- In 1894, building on the pioneering works of Maxwell and Hertz, Oliver Lodge demonstrated wireless communications, albeit over the relatively short distance of 150 meters.
- During the period from 1895 to 1901, Guglielmo Marconi developed an apparatus for transmitting radio waves over longer distances, culminating in a transmission across the Atlantic Ocean on December 12, 1901, from Cornwall, England, to Signal Hill in Newfoundland, Canada. Similar work was being done by A.S. Popoff of Russia during this time period.
- In 1902, the first point-to-point radio link in the United States was established from California to Catalina Island. These first radio systems were often referred to as wireless telegraphy.
- In 1906, Reginald Fessenden made history by conducting the first radio broadcast, transmitting music and voice using a technique that came to be known as *amplitude modulation* (AM) radio.
- In those early days, the military and merchant marine were quick to adapt wireless techniques. Wireless communications is often given credit for saving over 700 lives during the sinking of the *Titanic* in 1912.

- The early history of *land-mobile wireless* communication is a history of police pioneering. In 1921, the Detroit Police Department made the earliest significant use of wireless communications in a vehicle, operating a radio system at a carrier frequency close to 2 MHz. By 1934, 194 municipal and 58 state police forces were using AM radio for mobile voice communications.
- Parallel work performed on both sides of the Atlantic resulted in the first television broadcasts in 1927. Bell Labs demonstrated television broadcasts in the New York area, and John Baird made similar demonstrations in the United Kingdom.
- *Spread spectrum* techniques made their first appearance just before and during World War II. There were two applications in particular: encryption and ranging. Spread spectrum encryption techniques were often analog in nature with a noiselike signal multiplying a voice signal. The noise signal or information characterizing it was often sent on a separate channel to allow decryption at the receiver.
- In 1946, the first *public mobile telephone systems* were introduced in five American cities.
- In 1947, the first microwave relay system consisting of seven towers connecting New York and Boston became operational. This relay system was capable of carrying 2400 simultaneous conversations between the two cities.
- In 1958, a new era in wireless communications was initiated with the launch of the SCORE (Signal Communication by Orbital Relay Equipment) satellite. This satellite had only the capacity of one voice channel, but its success initiated the start of a new area of radio communications.
- In 1981, the first analog cellular system, known as the Nordic Mobile Telephone (NMT), was introduced in Scandinavia. This was soon followed by the Advanced Mobile Phone Service (AMPS) in North America in 1983.
- In 1988, the first digital cellular system was introduced into Europe. It was known as the *Global System for Mobile (GSM) Communications.* Originally intended to provide a pan-European standard to replace the myriad of incompatible analog systems in operation at the time, GSM was soon followed by the North American IS-54 digital standard.

These are just a few of the accomplishments in wireless communications over the past 150 years. Today, wireless devices are everywhere. Cellular telephones are commonplace. Satellites broadcast television direct to the home. Offices are replacing Ethernet cables with wireless networks. The introduction of these wireless services has increased the mobility and service area for many existing applications and numerous unthought-of applications. Wireless is a growing area of public networks and it plays an equally important role in private and dedicated communications systems. It is an exciting time in radio communications.

1.2 COMMUNICATION SYSTEMS

Today's public communications networks are complicated systems. Networks such as the public switched telephone network (PSTN) and the Internet provide seamless connections between cities, across oceans, and between different countries, languages, and cultures. Wireless is only one component of these complex systems, but there are three areas or layers where it can affect the design of such systems:

1. *Physical layer.* This layer provides the physical mechanism for transmitting *bits* (i.e., binary digits) between any pair of nodes. In wireless systems, it performs the modulation and demodulation of the electromagnetic waves used for transmission, and it includes the transmission medium as well. The module for performing *mo*dulation and *demo*dulation is often called a *modem.*
2. *Data-link layer.* Wireless links can often be unreliable. One purpose of the data-link layer is to perform error correction or detection, although this function is shared with the physical layer. Often, the data-link layer will retransmit packets that are received in error but, for some applications, it discards them. This layer is also responsible for the way in which different users share the transmission medium. In wireless systems, the transmission medium is the radio spectrum. A portion of the data-link layer called the *Medium Access Control (MAC)* sublayer is responsible for allowing frames to be sent over the shared media without undue interference with other nodes. This aspect is referred to as *multiple-access* communications.
3. *Network layer.* This layer has several functions, one of which is to determine the *routing* of the information, to get it from the source to its ultimate destination. A second function is to determine the *quality of service.* A third function is *flow control*, to ensure that the network does not become congested. Wireless systems, in which some of the nodes can be mobile, place greater demands on the network layer, since the associations among nodes are continually changing.

These are three layers of a seven-layer model for the functions that occur in the communications process. The model is called the *open system interconnection* (OSI) reference model[1] and we will discuss it in greater detail later in the book. The physical layer is the lowest layer of this model and is the one directly connected to the transmission medium. Higher layers are proportionately less affected by the transmission medium.

The book will concentrate mainly on the physical and data-link layers. We will finish by illustrating how these are integrated with wireless networks.

1.3 THE PHYSICAL LAYER

The physical layer provides the equivalent of a communications pipe between the source and destination of the information. In wireless and other communications systems, the physical layer has three basic components; the transmitter, the channel, and the receiver, as depicted in Fig. 1.1. *How are these different in a wireless system?*

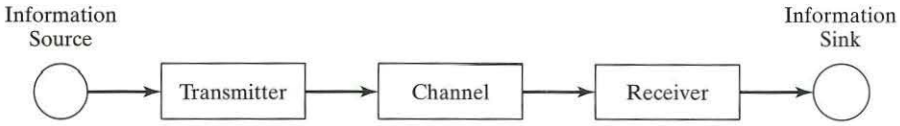


FIGURE 1.1 Block diagram of a communication system linking a source of information to a user of the information.

1. *Transmitter.* The basic function of the transmitter is to take the information-bearing signal produced by the source of information and modify it into a form suitable for transmission over the channel. In wireless systems,

- **(a)** the transmitter shapes the signal so that it can pass reliably through the channel, while efficiently using the limited *transmission medium resources* (i.e., *the radio spectrum*)
- **(b)** the terminals are often mobile and limited by battery power, so the transmitter must use modulation techniques that are both robust and power-efficient; and
- **(c)** because the medium is shared with other users, the design should minimize interference with other users.

2. *Channel.* The channel provides the physical means for transporting the signal produced by the transmitter and delivering it to the receiver. In wireless systems, the channel impairments

- **(a)** include *channel distortion* that may take the form of multipath—the constructive and destructive interference between many copies of the same signal received;
- **(b)** often have a *time-varying nature,* due to either the terminal mobility or a change in conditions along the propagation path;
- **(c)** include *interference*, which is produced (accidentally or intentionally) by other sources whose output signals occupy the same frequency band as the transmitted signal; and
- **(d)** include *receiver noise*, which is produced by electronic devices at the front end of the receiver. Although produced at the receiver, this noise is often considered a channel effect and therefore also referred to as *channel noise.* The effect of receiver noise will depend on the received signal strength, which in turn will depend significantly on the propagation path between the transmitter and receiver.

3. *Receiver.* The receiver operates on the received signal to produce an *estimate* of the original information-bearing signal. We say "estimate" rather than an exact reproduction of the original information-bearing signal because of the unavoidable presence of *impairments* in the channel. In wireless systems, the receiver

- **(a)** frequently must *estimate the time-varying nature of the channel* in order to implement compensation techniques;

(b) implements *error-correction techniques* to improve the often poor reliability of wireless channels relative to other media; and

(c) must *maintain synchronization* under rapidly varying channel conditions.

All of these facets make the physical layer very challenging in wireless systems.

1.4 THE DATA-LINK LAYER

Conceptually, at the highest level of the data link is the *multiple access strategy.* This describes the general approach to sharing the physical resources among the different users. For wireless systems, the key physical resource is the radio spectrum. In general, for any wireless service, only a fixed amount of spectrum is allotted. We will consider four multiple access strategies for sharing this spectrum:

1. FDMA – *Frequency-division multiple access* refers to sharing the spectrum by assigning specific frequency channels to specific users on either a permanent or temporary basis.
2. TDMA – *Time-division multiple access* refers to allowing all users access to all of the available spectrum, but users are assigned specific time intervals during which they can access it, on either a temporary or permanent basis.
3. CDMA – *Code-division multiple access* is a form of spread spectrum modulation in which users are allowed to use the available spectrum, but their signal must be spread ("encrypted") with a specific code to distinguish it from other signals.
4. SDMA – *Space-division multiple access* refers to sharing the spectrum between the users by exploiting the spatial distribution of user terminals through the use of "smart" directional antennas that minimize the interference between the terminals.

The objective of all these strategies is to maximize the spectrum utilization—that is, to provide service to more users with the same amount of spectrum. In practice, most wireless systems are a combination of one or more of these multiple access strategies. Historically, due to the existing technology and network topologies, these multiple access strategies developed approximately in the order just presented.

1.4.1 FDMA

Early wireless systems tended to be point-to-point systems with one transmitter communicating with only one receiver. The natural approach was to assign each pair of *user terminals* (UTs) a frequency channel, a form of frequency-division multiple access (FDMA). In the frequency domain, FDMA can be visualized as shown in Fig. 1.2. The radio spectrum is divided into a number of channels, and each pair of users is assigned a different channel. In some cases, a different channel is assigned for each direction of transmission. From an interference viewpoint, it is important that each user signal be kept confined to the assigned band.

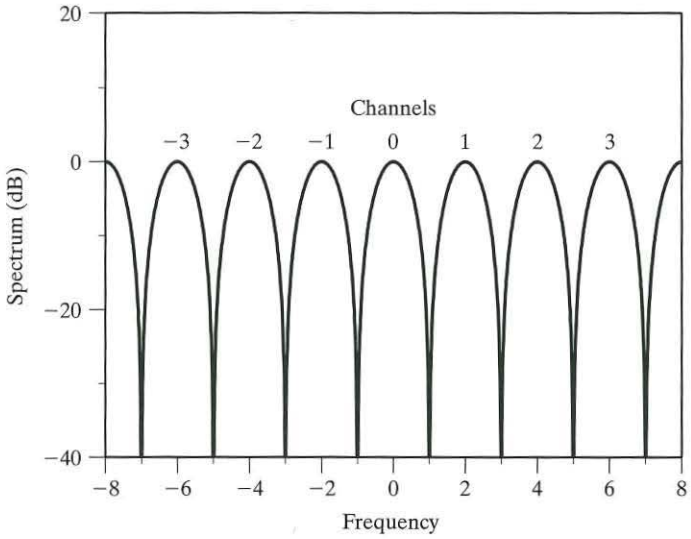


FIGURE 1.2 A frequency-domain representation of FDMA.

1.4.2 TDMA

Later systems were designed to operate in a point-to-multipoint manner, with many users sharing a central base station. The *point-to-multipoint* architecture shown in Fig. 1.3 consists of multiple UTs communicating with a single *base station* (BS). The multiple UTs time-share one or a small number of radio channels. These early time-division multiple access (TDMA) systems used analog modulation with a simple *push-to-talk* protocol. A user waited until the channel seemed unoccupied and then pushed a button on his or her microphone to talk. An example of this type of system would be a taxi-dispatch system.

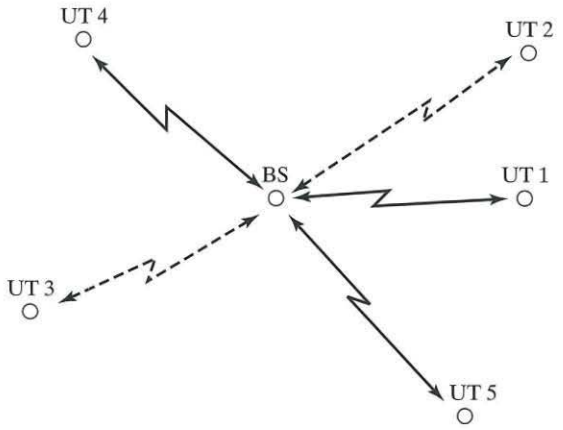


FIGURE 1.3 Point-to-multipoint network architecture.

FIGURE 1.4 TDMA frame structure with N slots per frame.

With the advent of digital technology, TDMA systems appeared with more complex and efficient sharing strategies. An example of this is shown in Fig. 1.4, in which time is divided up into frames and each frame is divided into slots. Each active UT is assigned one or more slots from each frame.

From the structure of Fig. 1.4, it would appear that TDMA is suited only to data applications. In fact, many initial applications were voice applications, and the key to such an application is recognizing that the transmission rate for a digital TDMA channel is typically N times faster than that required for a single channel.

1.4.3 CDMA

The object of a *cellular system* is to reuse the radio spectrum over a large area as many times as possible. In a cellular system, the cells are modeled with a hexagon pattern as shown in Fig. 1.5. User terminals (UTs) in each cell communicate with a base station located at the center of the cell as in a point-to-multipoint arrangement, but the UTs can freely roam between cells, changing their base station as they move.

Within a cell, either FDMA or TDMA techniques can be used for sharing the transmission medium. Distinct sets of channel frequencies are assigned to each cell, with the channels being reused in sufficiently separated cells; an example of these *cochannel cells* are shaded in Fig. 1.5.

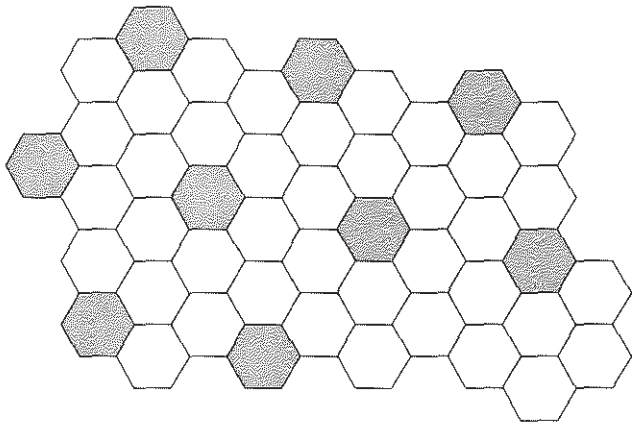


FIGURE 1.5 Hexagonal pattern of cells used in a cellular telephone system.

A drawback of using FDMA and TDMA in cellular systems is that the *reuse distance* is typically limited by worst-case interference. Cochannel cells must be assigned so that the worst-case interference is at an acceptable level. Intuitively, we would expect that reuse could be improved if cochannel cells were assigned on the *basis of average interference. Spread spectrum* is a modulation strategy that is quite tolerant of interference, and it forms the basis for the access technique known as code-division multiple access (CDMA). Spread spectrum spreads the information-bearing data over a large bandwidth, achieving the desired averaging effect. As a result, it allows this same spectrum to be used simultaneously by many UTs in adjacent cells, under appropriate conditions.

1.4.4 SDMA

The three multiple access techniques of FDMA, TDMA, and CDMA have increased spectral efficiency by increasing reuse in frequency, time, and space. So far, we implicitly have assumed antennas that are *omnidirectional* (i.e., antennas that operate essentially in a uniform manner along all directions). If the transmit and receive antenna could be focused directly at the other end of the link, then this would provide a number of improvements:

- It would reduce the total power needed to be transmitted, as all power would be flowing in the right direction.
- It would reduce the amount of interference generated by each transmitter, because total transmit power is reduced and localized.
- The receiver would receive a stronger signal due to both antenna gain and less interference.

The development of such "smart" antennas could allow even greater reuse of the radio spectrum. Space-division multiple access (SDMA) improvements in spectral efficiency are thus achieved by exploiting the angular separation of the individual UTs. In particular, *multibeam antennas* are used to separate radio signals by pointing them along different directions. Thus, different users are able to reuse the same radio spectrum, as long as they are separated in angle.

1.5 OVERVIEW OF THE BOOK

The fundamental physical resource in wireless systems is the radio spectrum. *The development of wireless communications has often been a search for more efficient ways of using the radio spectrum.* This evolution is a consequence of both increasing demand for a limited spectral resource and new technologies that have made spectrally efficient techniques a practical reality. This book follows this theme of evolving spectral efficiency.

The fundamental constraints on this evolutionary process are due to the transmission media—the *physical properties of the wireless channel.* These properties shape and influence the transmission techniques that are both practical and reliable. For that reason, understanding the physical media is a key to understanding wireless communications, and this is the topic of Chapter 2.

Early wireless systems were point-to-point systems using FDMA. As the radio spectrum became more crowded, there was a great interest in improving the efficiency of these individual channels. The role that *modulation* plays in the spectral efficiency of individual channels is investigated in Chapter 3. For mobile UTs, the modulation options are constrained by power limitations of the terminal and the need for the signal to be robust and easily detectable in fading.

The introduction of digital technology resulted in large improvements in efficiency for strategies such as TDMA. Similar technology can be applied to reduce the amount of information that needs to be transmitted; in particular, *digital speech-coding* techniques, which encode speech signals intelligently to minimize the number of bits required, are another way of increasing spectral efficiency. Once wireless networks started being used for digitized voice, it was a small step to also use them for data traffic. This created a problem, since existing networks and mobile terminals were designed for voice applications. Voice has a relatively high tolerance for bit errors, whereas data can be highly sensitive to bit errors—just imagine if someone moved the decimal point in your bank account! Data transmission led to the use of *forward error-correction coding* in wireless systems. TDMA and these pertinent technologies are investigated in Chapter 4.

After the time and frequency dimensions, the next area for further improvements in spectral efficiency is the spatial dimension—that is, the reuse of the same spectrum, but in different localities. This leads to the introduction of cellular systems in Chapter 5. The key factor limiting spectrum reuse is interference between user signals, and one approach to managing the effects of interference to optimize spectral efficiency is the use of a *spread spectrum*. This is why Chapter 5 begins with a discussion of spread spectrum modulation and CDMA before introducing the idea behind cellular systems.

While cellular systems produced significant increases in spectral efficiency, further increases were still possible through the *reuse of the spectrum in different angular directions*. This angular reuse requires "smart antenna" technology and is the motivation for SDMA—the topic of Chapter 6. One of the other challenges of wireless communications is the unpredictable and potentially unreliable nature of the propagation path. One well-known method of increasing the reliability of communications is *diversity*. Diversity is the transmission of the same information-bearing data by different methods in order to improve reliability. One of the many forms of diversity is spatial diversity, which usually implies the transmission or reception of the signal through the use of multiple antennas, forcing the signal to travel different paths in space. For this reason, spatial diversity is closely related to SDMA techniques—hence the discussion of both diversity and SDMA in Chapter 6.

The focus of this book is the physical-layer and multiple-access techniques associated with wireless communications. It is important that the reader understand how these technologies fit into the wireless networks that are currently evolving. Toward that end, Chapter 7 includes a comparison of the different multiple-access strategies and link-management functions and discusses how they fit into the architecture of wireless networks.

Notes and References

[1] The OSI reference model was developed as a paradigm for computer-to-computer communication. It has since been applied in a qualitative manner to many other networks. The first chapter of the book by Bertsekas and Gallagher (1992) provides an overview of the OSI model, and later chapters describe its application to data networks.

CHAPTER 2

Propagation and Noise

2.1 INTRODUCTION

The study of propagation is important to wireless communications because it provides prediction models for estimating the power required *to close a communications link* and provide reliable communications. The study of propagation also provides clues to receiver techniques for compensating the impairments introduced through wireless transmission. In this chapter, we will look at physical models for various propagation phenomena and use these models to predict their effects on transmission. For the most part, the physical models are limited to simple scenarios, but they do provide a basic understanding of how signals propagate. Although the major emphasis of this chapter will be on terrestrial wireless communications, we will consider satellite communications as a baseline because it often provides close to ideal propagation conditions.

The propagation effects and other signal impairments are often collected and categorically referred to as the *channel*. The additive white Gaussian noise (AWGN) channel is the channel model most often used in communications theory development. With this model, zero-mean noise having a Gaussian distribution is added to the signal. The noise is usually assumed to be *white* over the bandwidth of interest; that is, samples of the noise process are uncorrelated with each other. This implies that the noise has an autocorrelation function given by $(N_0/2)\,\delta(t)$, where $\delta(t)$, is the Dirac delta function, or equivalently, it has a flat two-sided power spectral density $N_0/2$ (watts/Hz) across the frequency range $-\infty < f < \infty$. The AWGN channel model is shown in Fig. 2.1.

For fixed satellite communications in which there is a direct line-of-sight path between the transmitter and satellite, and the satellite and receiver, the additive white Gaussian noise (AWGN) channel often provides a reasonably good model[1]. In general, channel models for wireless communications take one of two forms:

1. *Physical models* take into account the exact physics of the propagation environment. Consequently, for any particular situation, site geometry has to be taken into consideration. This method provides the most reliable estimates of propagation behavior but it is computationally intensive.
2. *Statistical models* take an empirical approach, measuring propagation characteristics in a variety of environments and then developing a model based on the

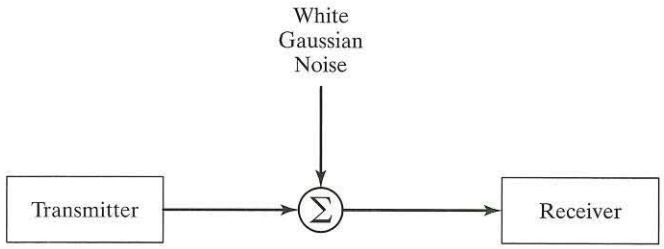


FIGURE 2.1 The Gaussian noise channel.

measured statistics for a particular class of environments. These models are easier to describe and use than the physical models, but do not provide the same accuracy.

We will provide examples of both types of models.

When considering physical models, we will consider the following three basic modes of propagation:

1. Free-space or line-of-sight transmission, as the name implies, corresponds to a clear transmission path between the transmitter and the receiver. Satellite communications generally rely on line-of-sight paths between the transmitter and satellite, and between the satellite and receiver.
2. *Reflection* refers to the bouncing of electromagnetic waves from surrounding objects such as buildings, mountains, and passing vehicles. In terrestrial wireless communications, there is often no direct line-of-sight path between the transmitter and receiver, and communications rely on reflection and diffraction, described next.
3. *Diffraction* refers to the bending of electromagnetic waves around objects such as buildings, or terrain such as hills, and through objects such as trees and other forms of vegetation.

As a result of another propagation phenomenon called *refraction*, electromagnetic waves are bent as they move from one medium to another. An example is the bending of light when it enters water from the air. The occurrence of refraction with wireless communications is limited to special circumstances, such as ionospheric communications.

Often, the received signal is the combination of many of these modes of propagation. That is, the transmitted signal may arrive at the receiver over many paths. The signals on these different paths can constructively or destructively interfere with each other. This is referred to as *multipath*. If either the transmitter or receiver is moving, then these propagation phenomena will be time varying, and *fading* occurs. One advantage of statistical models is their succinct way of describing these complicated situations.

In addition to propagation impairments, the other phenomena that limit communication are *noise* and *interference*. There are many sources for these impairments, including the receiver, human-made sources, and other transmitters.

This chapter has two objectives. The first is to provide a basic understanding of the physics behind various propagation phenomena and noise and how they affect communications. To that end, Section 2.2 explains free-space propagation, Sections 2.3 and 2.4 respectively present physical and statistical models for terrestrial wireless propagation, and Section 2.5 discusses indoor propagation. These explanations are followed by a discussion of local propagation effects such as multipath and Rayleigh fading in Section 2.6 and statistical scattering models in Section 2.7.

The second objective is to provide the reader with a system-level appreciation of the importance of propagation and noise effects in wireless communications. To achieve this, noise and interference effects are described in Section 2.8, and link budgets are presented in Section 2.9. The chapter ends with three theme examples—one describing an empirical model for land mobile propagation, the second one describing how all the facets examined in this chapter fit together in the design of a wireless local area network, and the third describing propagation advantages of ultra-wideband radio.

2.2 FREE-SPACE PROPAGATION

Wireless transmission is characterized by the generation, in the transmitter, of an electric signal representing the desired information, the 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. The transmission effects are most completely described by Maxwell's equations; however, in this text we will model the propagation method solely on the basis of the electrical field, and we shall *assume a linear medium in which all distortions can be characterized by attenuation or superposition of different signals.*

2.2.1 Isotropic Radiation

A fundamental concept in the understanding of radio transmission systems is that of an *isotropic antenna.* This is an antenna that transmits equally in all directions. In reality, an isotropic antenna does not exist, and all antennas have some *directivity* associated with them. Still, the isotropic antenna is a reference to which other antennas are compared, and it is useful for explaining fundamentals.

Consider an isotropic source that radiates power P_T watts equally in all directions, as illustrated in Fig. 2.2. The power per unit area, or *power flux density*, on the surface of a sphere of radius R centered on the source is given by

$$\Phi_R = \frac{P_T}{4\pi R^2} \tag{2.1}$$

where $4\pi R^2$ is the surface area of the sphere. Typical units for power flux density are watts per square meter. Equation (2.1) follows directly from the conservation of energy; that is, the energy radiated per unit time by the source must appear at the surface of sphere (or any closed surface) enclosing that source.

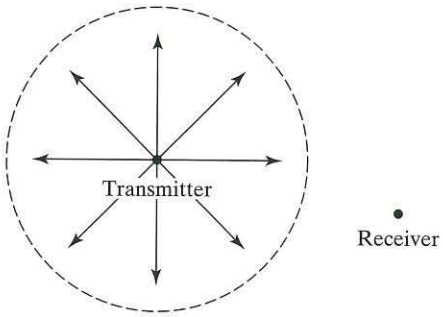


FIGURE 2.2 Illustration of isotropic radiation.

The power P_R that is captured by the receiving antenna depends on the size and orientation of the antenna with respect to the transmitter. The power received by an antenna of *effective area* or *absorption cross section* A_e is given by

$$P_R = \Phi_R A_e = \frac{P_T}{4\pi R^2} A_e \tag{2.2}$$

An antenna's physical area A and its effective area A_e are related by the *antenna efficiency*

$$\eta = \frac{A_e}{A} \tag{2.3}$$

This parameter indicates how well the antenna converts incident electromagnetic energy into usable electrical energy. For parabolic antennas, the efficiency typically ranges from 45% to 75%; for horn antennas, the efficiency can range from 50% to 80%.

It is intuitively clear that an isotropic transmit antenna can be considered a point source. What is the effective area of an isotropic receive antenna? If it is considered a *point sink*, then the answer is not obvious. However, from electromagnetic theory, we note that the *effective area of an isotropic antenna in any direction* is given by[2]

$$A_{\text{iso}} = \frac{\lambda^2}{4\pi} \tag{2.4}$$

where λ is the wavelength of radiation. Substituting Eq. (2.4) into (2.2), we obtain the following relationship between the transmitted and received power for isotropic antennas:

$$P_R = \frac{P_T}{(4\pi R/\lambda)^2} = \frac{P_T}{L_p} \tag{2.5}$$

In this equation, the path loss is defined as

$$L_p = \left(\frac{4\pi R}{\lambda}\right)^2 \tag{2.6}$$

The quantity L_p is the *free-space path loss* between two isotropic antennas. This definition of path loss depends, somewhat surprisingly, on the wavelength of transmission—a consequence of the dependence of the effective area of an isotropic antenna on the wavelength.

EXAMPLE 2.1 Receiver Sensitivity

Sensitivity is a receiver parameter that indicates the minimum signal level required at the antenna terminals in order to provide reliable communications. The factors it depends on include the receiver design, modulation format, and transmission rate. Receiver sensitivity is often expressed in dBm—that is, the power in milliwatts, expressed in decibels. For example, a commercial mobile receiver for data transmission may be specified with a sensitivity of –90 dBm. Assuming a 100-milliwatt transmitter and free-space path loss between the transmitting and receiving isotropic antennas, what is the radius of the service area of this receiver at a transmission frequency of 800 MHz?

To answer this question, we first note that –90dBm is equivalent to 10^{-9} milliwatts of power. From Eq. (2.5), the maximum tolerable path loss is

$$L_p = \frac{P_T(\text{mW})}{P_R(\text{mW})} = \frac{100}{10^{-9}} = 10^{11}$$

Consequently, observing that $\lambda = c/f$, where c is the speed of light, and using Eq. (2.6), we find that the maximum range is given by

$$R = \frac{\lambda}{4\pi}\sqrt{L_p} = \frac{c}{4\pi f}\sqrt{L_p} = \frac{3\times 10^8 \text{m/s}}{4\pi \times 800 \times 10^6 \text{s}^{-1}}\sqrt{10^{11}} = 9.2\text{km}$$

The conclusion drawn from this example is that it takes very little transmit power to provide a large service area under free-space propagation conditions. ■

2.2.2 Directional Radiation

While the isotropic antenna is a useful illustrative device, most antennas are not isotropic. Instead, they have *gain* or *directivity* $G(\theta,\phi)$ that is a function of the *azimuth* angle θ and the *elevation* angle ϕ:

1. The azimuth angle θ is the look angle in the horizontal plane of the antenna relative to a reference horizontal direction. The reference direction could, for example, be due north.
2. The elevation angle ϕ is the look angle of the antenna above the horizontal plane.

With these definitions, the transmit gain of an antenna is defined as

$$G_T(\theta, \phi) = \frac{\text{Power flux density in direction } (\theta, \phi)}{\text{Power flux density of an isotropic antenna}} \tag{2.7}$$

for the same transmit power. Clearly, the transmit gain of an isotropic antenna is unity.

The corresponding definition for the receive antenna gain is

$$G_R(\theta, \phi) = \frac{\text{Effective area in direction } (\theta, \phi)}{\text{Effective area of an isotropic antenna}} \tag{2.8}$$

Equations (2.7) and (2.8) define the gain for the transmit and receive antennas, respectively. With bidirectional communications, the same antenna is often used to both transmit and receive. *How are the two gains related*? To answer this question, we introduce what is known as the *principle of reciprocity*, which may be stated as follows:[3]

> *Signal transmission over a radio path is reciprocal in the sense that the locations of the transmitter and receiver can be interchanged without changing the transmission characteristics.*

The principle of reciprocity is based on Maxwell's equations of electromagnetic theory. What it implies is that, if the direction of propagation is reversed, the energy of the signal would follow exactly the same paths and suffer the same effects, but in the reverse direction. This implies that the *transmit and receive antenna gains are equal.* The intuitive explanation of why this is the case is the observation that the transmit gain is a measure of how well an antenna emits the radiated energy in a certain area (defined by the azimuth and elevation), while the receive gain is a measure of how well the antenna collects the radiated energy in that area.

Consequently, from the principle of reciprocity and Eq. (2.8), the *maximum transmit or receive gain*, of an antenna, in any direction is given by

$$G = \frac{4\pi}{\lambda^2} A_e \tag{2.9}$$

The antenna pattern for a parabolic (dish-shaped) antenna is shown in Fig. 2.3. The maximum gain of the antenna is along the axis of the parabola and is given by Eq. (2.9). The gain decreases in the off-axis direction. The *beamwidth* of the antenna is defined as the angular width of the antenna beam at the 3-dB points, as shown in Fig. 2.3. The illustration shows the main lobe of the antenna pattern plus two *sidelobes.* An antenna may also have a *backlobe.* The sidelobes and backlobe are typically not considered for use in the communications link, but they often have to be considered when analyzing interference.

EXAMPLE 2.2 Parabolic Antenna Gain

For a parabolic antenna, looking directly at the antenna boresight, we find that the area is simply $\pi D^2/4$, where D is the diameter of the dish. Consequently, the effective area of the antenna is given by

$$A_e = \eta \frac{\pi D^2}{4}$$

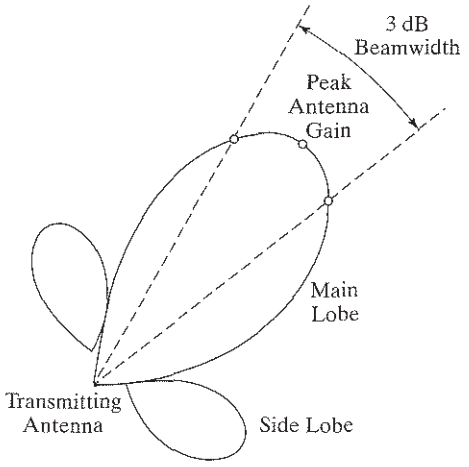


FIGURE 2.3 Illustration of gain pattern of a parabolic antenna.

and the corresponding antenna gain (both transmit and receive) is

$$G = \frac{A_e}{A_{\text{isotropic}}} = \frac{\eta \pi D^2/4}{\lambda^2/4\pi} = \eta\left(\frac{\pi D}{\lambda}\right)^2$$

This equation illustrates that the antenna gain depends on the wavelength of transmission. For example, a 0.6-m parabolic dish used for receiving a direct broadcast satellite television signal at 12 GHz has a gain of

$$G = \eta\left(\frac{\pi D}{\lambda}\right)^2 = 0.5\left(\frac{\pi \times 0.6 \times f}{c}\right)^2 = 2842.4$$

If we express this gain in decibels, it corresponds to a gain of 34.5 dB. We have assumed an antenna efficiency of 50%. ■

The previous example described how to calculate the maximum gain for a parabolic antenna when one is looking directly at the center of the antenna. For such an antenna, the theoretical gain G varies with the off-axis angle ψ, as illustrated by Fig. 2.3. Normalized to unity on-axis gain, the theoretical gain of a parabolic antenna is

$$G(\psi) = \left(\frac{2J_1((\pi D/\lambda)\sin\psi)}{\sin\psi}\right)^2\left(\frac{\pi D}{\lambda}\right)^{-2} \tag{2.10}$$

where $J_1(x)$ is the Bessel function of the first kind and D is the diameter of the antenna. (The Bessel function is discussed in Appendix D.) This antenna gain characteristic is shown in Fig. 2.4, plotted as a function of the off-axis angle ψ.

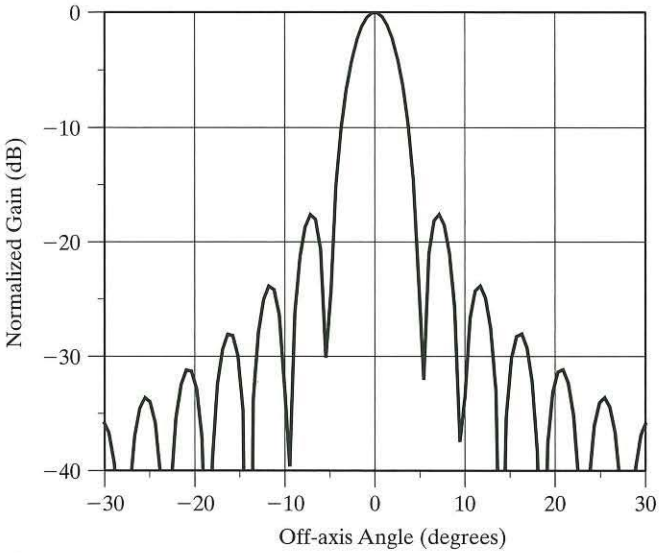


FIGURE 2.4 Off-axis gain characteristic of a 1-meter parabolic antenna at 4 GHz.

The *3-dB beamwidth* is a measure of the size of the antenna beam. For the parabolic antenna, it is the off-axis angle for which the gain falls to 3 dB below its peak value. For a parabolic antenna with an ideal gain characteristic as described by Eq. (2.10), the 3-dB beamwidth[4] is $58.4° \times (\lambda/D)$. From Eq. (2.9), it should be clear that the larger the antenna, the greater the gain it provides. One important further result inferred from Eq. (2.10) is that the larger the antenna, the narrower is the beamwidth of the antenna. That is, large antennas have high gain, but they must be pointed accurately to make use of the gain. This result illustrates a general rule of thumb that *the 3-dB beamwidth is linearly proportional to the wavelength and inversely proportional to the antenna diameter.*

Problem 2.1 Early satellite communications systems often used large 20-m-diameter parabolic dishes with an efficiency of approximately 60% to receive a signal at 4 GHz. What is the gain of one of these dishes in dB?
Ans. 56.2 dB. ■

2.2.3 The Friis Equation

When nonisotropic antennas are used, the free-space loss relating the received and transmitted power for general antennas is determined by direct substitution of the gain definitions Eqs. (2.7) and (2.8) into Eq. (2.5) to obtain

$$P_R = \frac{P_T G_T G_R}{L_p} \tag{2.11}$$

Equation (2.11) is referred to as the *Friis equation*. To simplify its evaluation, we write it as the decibel relation

$$P_R(\text{dB}) = P_T(\text{dB}) + G_R(\text{dB}) + G_T(\text{dB}) - L_P(\text{dB}) \tag{2.12}$$

where $X(\text{dB}) = 10 \log_{10}(X)$. The Friis equation is the fundamental *link budget* equation. It relates power received to power transmitted, taking into account transmission characteristics of the radio link. *Closing the link* refers to the requirement that the right-hand side provide enough power at the receiver to detect the transmitted information reliably. That is, the right-hand side must be greater than the receiver sensitivity. In a later section, a more formal link budget will be presented that takes into account a greater number of transmission effects.

Problem 2.2 In terrestrial microwave links, line-of-sight transmission limits the separation of transmitters and receivers to about 40 km. If a 100-milliwatt transmitter at 4 GHz is used with transmitting and receiving antennas of 0.5-m^2 effective area, what is the received power level in dBm? If the receiving antenna terminals are matched to a 50-ohm impedance, what voltage would be induced across these terminals by the transmitted signal?
Ans. $P_R = -55.5$ dBm and $V_{\text{rms}} = 0.37$ millivolt. ■

2.2.4 Polarization

Electromagnetic waves are transmitted in two orthogonal dimensions, referred to as *polarizations*. There are two commonly used orthogonal sets of polarizations:

1. *Horizontal and vertical polarization.* Vertical polarization is generally used for terrestrial mobile radio communications. At frequencies in the VHF band, vertical polarization is better than horizontal polarization because it produces a higher field strength near the ground. Furthermore, mobile antennas for vertical polarization are more robust and convenient to implement.
2. *Left-hand and right-hand circular polarizations.* These are often used in satellite communications. For well-designed fixed communication links, we can use the two orthogonal polarizations to double the transmission capacity in a given frequency band.

In mobile communications, there is often little frequency reuse advantage to employing both polarizations. *Scattering effects*, described later in the chapter, tend to create a cross-polar component that causes interference between the original orthogonal components and that limits the benefit of polarization.

2.3 TERRESTRIAL PROPAGATION: PHYSICAL MODELS

The free-space propagation equation assumes that the effect of the Earth's surface can be neglected and that there is a clear line of sight between the transmitting and receiving antennas. This is usually the situation with satellite communications. With terrestrial communications, however, buildings, terrain, or vegetation may obstruct the line-of-sight path between the transmit and receive antennas. In this case,

communication relies on either the *reflection* of electromagnetic waves from various surfaces or *diffraction* around various objects. With these additional modes of propagation, a multitude of possible paths arise between the transmitter and receiver, and the receiver often receives a significant signal from more than one path. This phenomenon is referred to as *multipath* propagation. With multiple waves arriving at the same location, there is the possibility of either *destructive* or *constructive interference*, and consequently, the propagation properties can be quite different from those obtained with free-space propagation. In the next two subsections, we will look at the physical mechanisms of reflection and diffraction.

Other propagation modes, such as *ducting*, occur when the physical characteristics of the environment create a waveguidelike effect. One example of ducting occurs in high-frequency (HF) radio, for which the carrier frequencies range from 3 to 30 MHz. At these carrier frequencies, differences between the layers of the atmosphere are most pronounced, and the layers act like different media to the transmission. The HF signal can be trapped within a layer of the ionosphere, reflecting off the layers above and below in a manner similar to a stone skipping across a lake. This allows the signal to travel long distances with very little attenuation and is sometimes called *skywave*. However, the skywave form of wireless communications tends to be unreliable, because the boundaries between the atmospheric layers do not have long-term stability.

2.3.1 Reflection and the Plane-Earth Model[5]

Except for very short distances, the presence of the Earth plays an important role in terrestrial propagation. For distances less than a few tens of kilometers, we may often neglect the curvature of the earth and use a flat-Earth model, as illustrated in Fig. 2.5.

The flat-Earth model shows a fixed transmitter with an antenna of height h_T transmitting to a fixed receiving antenna of height h_R. It also shows a two-ray propagation model between the transmitter and the receiver, consisting of a direct path and a path reflected by the Earth's surface. The objective is to determine the power level at the receiving antenna terminals. First, we note the following relationship between the

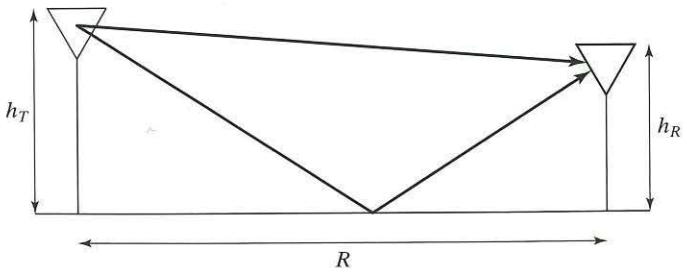


FIGURE 2.5 Plane-Earth reflection model.

power flux density and the *field strength* E:

$$\Phi = \frac{|E|^2}{\eta_0} \tag{2.13}$$

The field strength is measured in volts per meter (or an equivalent), and η_0 is the *characteristic wave impedance of free space*.[6] This characteristic impedance, 120π ohms, is the ratio of the complex amplitude of the electric field to that of the magnetic field in free space. The antenna acts as an impedance transformer to convert the free-space impedance to the impedance seen at the antenna terminals. We shall assume that the electric field is generated by a continuous wave (CW) signal with a transmission frequency f; that is, at a given point in space,

$$\begin{aligned} E(t) &= \sqrt{2}E_0 \cos(2\pi ft + \theta) \\ &= \sqrt{2}\mathrm{Re}\left\{E_0 e^{j(2\pi ft + \theta)}\right\} \end{aligned} \tag{2.14}$$

where the field strength E_0 and the phase θ depend on the location in space, and Re { } denotes the real part of the quantity inside the parentheses. To simplify the notation, we define the complex phasor

$$\tilde{E} = E_0 e^{j\theta} \tag{2.15}$$

and work extensively with this complex phasor notation in the discussion that follows.

In the case of the reflection of a single ray, if $\tilde{E}_d$ is the incident electric field and $\tilde{E}_r$ is the reflected field, then the relationship between the two fields is given by

$$\tilde{E}_r = \tilde{E}_d \rho e^{j\psi} \tag{2.16}$$

where ρ is the attenuation of the electric field and ψ is the phase change caused by the reflection. These parameters depend on both the angle of incidence of the field and the nature of the reflecting surface, including its smoothness and absorption properties.

Differences between the direct and reflected paths will depend on path length differences. From Fig. 2.5, by the Pythagorean theorem, the length of the direct path is

$$R_d = \sqrt{R^2 + (h_T - h_R)^2} \tag{2.17}$$

and, similarly, the length of the reflected path is

$$R_r = \sqrt{R^2 + (h_T + h_R)^2} \tag{2.18}$$

(This is most easily observed by reflecting the receiving antenna into the Earth.) To calculate the field strength at the receiving antenna, we assume that the difference in attenuation caused by different path lengths between the direct and ground-reflected

waves is negligible; that is,

$$|\tilde{E}_r| \approx |\tilde{E}_d| \tag{2.19}$$

However, the phase difference between the two paths is much more sensitive to the path length, and it cannot be neglected. The path difference between the reflected and incident rays is

$$\Delta R = R_r - R_d \tag{2.20}$$

We note that if R is large compared with both h_T and h_R, then the length of the direct path can be approximated by

$$\begin{aligned} R_d &= R\sqrt{1+\frac{(h_T-h_R)^2}{R^2}} \\ &\approx R\left(1+\frac{(h_T-h_R)^2}{2R^2}\right) \end{aligned} \tag{2.21}$$

In Eq. (2.21), we have used the approximation $\sqrt{1+x} \approx (1+x/2)$ for $x<<1$. If the same approximation technique is applied to R_r, then it follows by substituting Eq. (2.21) and its equivalent for R_r into Eq.(2.20) that

$$\Delta R \approx \frac{(h_T+h_R)^2-(h_T-h_R)^2}{2R} = 2\frac{h_T h_R}{R} \tag{2.22}$$

The corresponding phase difference between the two paths is proportional to the transmission wavelength and is given by

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta R = \frac{4\pi h_T h_R}{\lambda R} \tag{2.23}$$

Since $\tilde{E}_d$ represents the field strength at the receiving antenna due to the direct wave, it follows from using Eq. (2.16) that the total received field is

$$\begin{aligned} \tilde{E} &= \tilde{E}_d + \tilde{E}_r \\ &= \tilde{E}_d(1+\rho\exp(j\psi)\exp(-j\Delta\phi)) \end{aligned} \tag{2.24}$$

If the Earth's surface is assumed to be smooth and flat, and the reflected ray is at grazing incidence so that the reflection coefficient is $\rho\exp(j\psi) = -1$, then the total received electric field is

$$\tilde{E} = \tilde{E}_d(1-\exp(-j\Delta\phi)) \tag{2.25}$$

The magnitude of the combined electric field is given by

$$\begin{aligned} |\tilde{E}| &= |\tilde{E}_d||1 - \exp(j\Delta\phi)| \\ &= |\tilde{E}_d|[1 + \cos^2\Delta\phi - 2\cos\Delta\phi + \sin^2\Delta\phi]^{1/2} \\ &= |\tilde{E}_d|[2 - 2\cos\Delta\phi]^{1/2} \\ &= 2|\tilde{E}_d|\sin\left(\frac{\Delta\phi}{2}\right) \end{aligned} \tag{2.26}$$

Substituting Eq. (2.23) into (2.26) yields

$$|\tilde{E}| = 2|\tilde{E}_d|\sin\left(\frac{2\pi h_T h_R}{\lambda R}\right) \tag{2.27}$$

From Eqs. (2.2) and (2.13), the received power is given by

$$\begin{aligned} P_R &= \Phi A_e \\ &= \frac{|\tilde{E}|^2}{\eta_0} A_e \\ &= 4\frac{|\tilde{E}_d|^2}{\eta_0} A_e \sin^2\left(\frac{2\pi h_T h_R}{\lambda R}\right) \end{aligned} \tag{2.28}$$

The factor $|\tilde{E}_d|^2 A_e/\eta_0$ in Eq. (2.28) is the power received via the direct path. This direct-path power is governed by free-space propagation conditions and thus is equivalent to the right-hand side of Eq. (2.11). Substituting Eq. (2.6) into (2.11), and using the result to replace the factor $|\tilde{E}_d|^2 A_e/\eta_0$ in Eq. (2.28), we have the modified Friis' equation

$$P_R = 4P_T\left(\frac{\lambda}{4\pi R}\right)^2 G_T G_R \sin^2\left(\frac{2\pi h_T h_R}{\lambda R}\right) \tag{2.29}$$

If the product λR is much greater than the product $h_T h_R$, we can approximate $\sin\theta$ by θ and Eq. (2.29) becomes

$$P_R \approx P_T G_T G_R\left(\frac{h_T h_R}{R^2}\right)^2 \tag{2.30}$$

Equation (2.30) is the *plane-Earth propagation equation*, which differs from the free-space equation in three ways:

1. As a consequence of the assumption that $R >> h_T, h_R$, the angle $\Delta\phi$ is small, and λ cancels out of the equation, leaving it to be essentially *frequency independent.*
2. It shows an *inverse fourth-power law,* rather than the inverse-square law of free-space propagation. This points to a far more rapid attenuation of the power received.
3. It shows the effect of the transmit and receive antenna heights on propagation losses. The *dependence on antenna height* makes intuitive sense.

The conclusion is that an apparently minor difference in the model can make a significant change in the behavior of radio propagation.

Problem 2.3 Plot and compare the path loss (dB) for the free-space and plane-Earth models at 800 MHz versus distance on a logarithmic scale for distances from 1m to 40 km. Assume that the antennas are isotropic and have a height of 10 m. ■

2.3.2 Diffraction[7]

In the previous section, we saw how reflections can significantly affect the propagation characteristics of electromagnetic waves. In this section, we consider a second significant propagation effect for terrestrial propagation, namely, *diffraction*.

In physics, *diffraction* refers to the phenomenon whereby, when electromagnetic waves are forced to travel through a small slit, they tend to spread out on the far end of the slit, as illustrated in Fig. 2.6. The diffraction mechanism is often explained by *Huygens's principle*, which may be stated as follows:

> *Each point on a wave front acts as a point source for further propagation. However, the point source does not radiate equally in all directions, but favors the forward direction, of the wave front.*

Huygens's principle can be derived from Maxwell's equations. Most important, the principle can be used to explain why electromagnetic waves bend over hills and around buildings, as illustrated in Fig. 2.7, to provide communications where we may have thought that there should have been an *electromagnetic shadow*. One portion of the wave front directly hitting the obstacle is blocked (absorbed or perfectly reflected), while the remaining portion tends to bend and illuminate the shadow region.

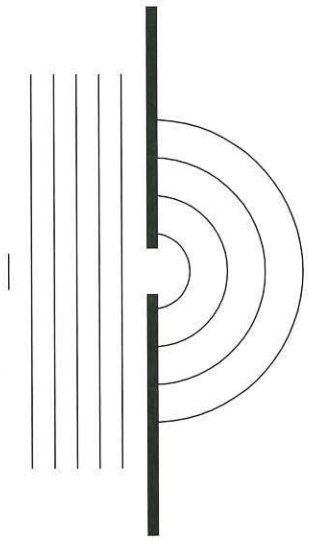


FIGURE 2.6 Behavior of plane wave passing through a slit from left to right.

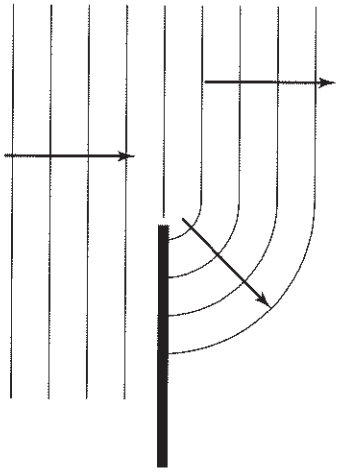


FIGURE 2.7 Illustration of knife-edge diffraction.

To introduce the concepts associated with diffraction, consider a transmitter **T** and receiver **R** in free space, as shown in Fig. 2.8. We also consider a plane normal to the line of sight path at a point **O** between **T** and **R**. On the plane, we construct circles of arbitrary radii. Any wave that propagates from **T** to **R** via a point on any of these circles traverses a longer path than the path **TOR**.

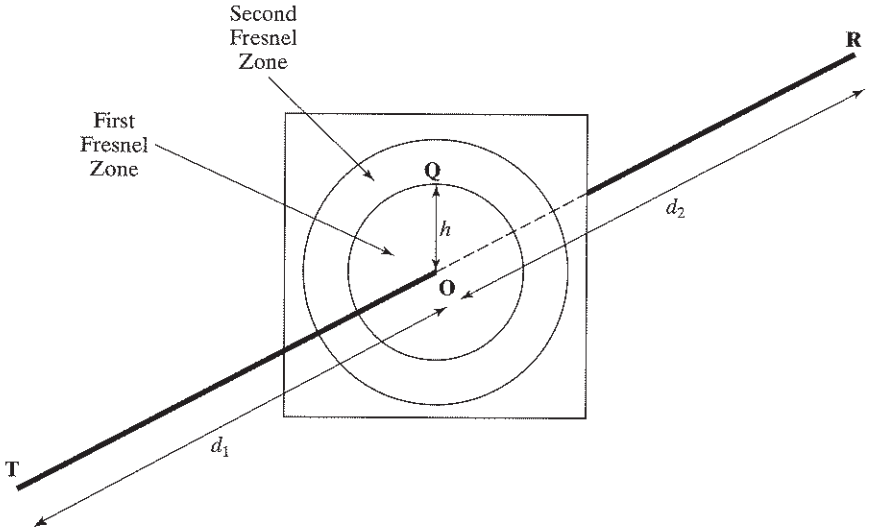


FIGURE 2.8 Fresnel zones of transmitted signal.

In terms of the geometry of Fig. 2.8, the excess path length between the direct path and the nondirect path through the point **Q** on the circle of radius h is given by

$$\begin{aligned}\Delta R &= |\mathbf{TQR}| - |\mathbf{TOR}| \\ &= \sqrt{d_1^2 + h^2} + \sqrt{d_2^2 + h^2} - (d_1 + d_2) \\ &= d_1\sqrt{1 + \left(\frac{h}{d_1}\right)^2} + d_2\sqrt{1 + \left(\frac{h}{d_2}\right)^2} - (d_1 + d_2)\end{aligned} \tag{2.31}$$

Assuming that $h << d_1, d_2$ and using the approximation $\sqrt{1+x} \approx 1 + x/2$ for $x<<1$, we find that Eq. (2.31) reduces to

$$\Delta R \approx \left(\frac{h^2}{2}\frac{d_1 + d_2}{d_1 d_2}\right) \tag{2.32}$$

As with the plane-Earth model, the phase difference between the two paths is the critical parameter. This phase difference, corresponding to a transmission wavelength λ, is given by

$$\begin{aligned}\Delta\phi &= \frac{2\pi}{\lambda}\Delta R \\ &\approx \frac{2\pi h^2}{\lambda}\frac{1}{2}\left(\frac{d_1 + d_2}{d_1 d_2}\right) \\ &= \frac{\pi}{2}h^2\left(\frac{2(d_1 + d_2)}{\lambda d_1 d_2}\right)\end{aligned} \tag{2.33}$$

We next define the *Fresnel–Kirchhoff diffraction parameter*

$$v = h\sqrt{\frac{2(d_1 + d_2)}{\lambda d_1 d_2}} \tag{2.34}$$

The phase difference of Eq. (2.33) can then be written as

$$\Delta\phi \approx \frac{\pi}{2}v^2 \tag{2.35}$$

The Fresnel–Kirchhoff diffraction parameter is a dimensionless quantity that characterizes the phase difference between two propagation paths. It is used in the discussion that follows to characterize diffraction losses in a general situation.

Returning to Fig. 2.8, we construct a family of hypothetical circles having the property that the total path length from **T** to **R** via each circle is $\Delta R = q\lambda/2$ longer than the path **TOR**, where q is an integer. The radii of these circles depend on the location of this imaginary plane along the **TR**-axis. The set of points for which the excess path length is an integer number of half wavelengths defines a family of ellipsoids, with **TOR** being the axis of revolution. If we slice a specific ellipsoid perpendicular to **TOR**, then, with $\Delta R = q\lambda/2$, it follows from Eq. (2.32) that the radius of the resulting circle is

$$h = r_q = \sqrt{\frac{q\lambda d_1 d_2}{d_1 + d_2}} \tag{2.36}$$

and the corresponding diffraction parameter is $v_q = \sqrt{2q}$. From Eq. (2.32), this approximation is valid for $d_1, d_2 >> r_q$. In words, *a particular Fresnel–Kirchhoff parameter v_q defines a corresponding ellipsoid of constant excess path.*

The volume enclosed by the first ellipsoid, $q = 1$, is known as the *first Fresnel zone.* The volume between the first ellipsoid and the second one, $q = 2$, is known as the *second Fresnel zone*, and so on.

Since the qth Fresnel zone is defined as a differential path length

$$\frac{(q-1)\lambda}{2} \le \Delta R \le \frac{q\lambda}{2} \tag{2.37}$$

relative to the line-of-sight path, the corresponding phase difference Eq. (2.33) pertaining to the qth Fresnel zone is $(q-1)\pi \le \Delta\phi \le q\pi$. Consequently, the contributions to the electric field at the receiver from successive Fresnel zones tend to be in phase opposition and therefore interfere destructively rather than constructively. Accordingly, we may state the following:

> *As a general rule of thumb, we must keep the "first Fresnel zone" free of obstructions in order to obtain transmission under free-space conditions.*

EXAMPLE 2.3 Fresnel Zones

Consider the case shown in Fig. 2.9, where the transmitter and receiver are at a height of 10 m and separated from each other by 500 m. From Eq.(2.34), at a point midway between the transmitter and receiver the radius of the first Fresnel zone is given by

$$r_1 = \sqrt{\frac{\lambda d_1 d_2}{d_1 + d_2}} = \sqrt{\frac{\lambda(250)(250)}{500}} = 11.18\sqrt{\lambda}\ \text{m} \tag{2.38}$$

At 800 MHz, $\lambda = 0.375$ m, and from Eq. (2.38), the first Fresnel zone has a radius of 6.84 m at the midway point. At a distance of 100 m from either the transmitter or receiver, the same calculation indicates that the first Fresnel zone has a radius of 5.48 m. Consequently, all objects between the transmitter and receiver must be shorter than 3 to 4 meters in order to have the approximate equivalent of free-space propagation over the 500-m path at 800 MHz. ■

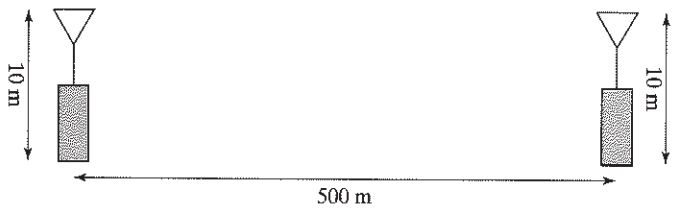


FIGURE 2.9 Scenario for Example 2.3.

Problem 2.4 A company owns two office towers in a city and wants to set up a 4-GHz microwave link between the towers. The two towers have heights of 100 m and 50 m, respectively, and are separated by 3 km. In the line of sight (LOS) and midway between the two towers is a third tower of height 70 meters. Will line-of-sight transmission be possible between the two towers? Justify your answer. Describe an engineering solution to obtain line-of-sight transmission.
Ans. No. ■

2.3.3 Diffraction Losses

If an ideal, straight, perfectly absorbing screen is placed between **T** and **R** in Fig. 2.8, it will have little effect if the top of the screen is well below the line-of-sight path. The field at **R** will then be the "free-space" value of $\tilde{E}_d$. As the height of the screen is increased, the field strength will vary up and down as the screen blocks more of the Fresnel zones below the path. The amplitude of the oscillation increases until the screen is just in line with **T** and **R** and the field strength is exactly one-half of the unobstructed value. When an obstruction is present, the Fresnel–Kirchhoff parameter is still given by

$$v = h\sqrt{\frac{2(d_1 + d_2)}{\lambda d_1 d_2}} \tag{2.39}$$

but in this case h refers to the height of the obstruction. *The height h, and thus also v, is considered positive if the obstruction extends above the line of sight and negative if it does not.*

From advanced diffraction theory,[8] the field strength at the point **R** is determined by the sum of all secondary Huygens's sources in the plane above the obstruction and can be expressed as

$$E = E_d \frac{1+j}{2} \int_v^\infty \exp\left(-j\frac{\pi}{2}t^2\right) dt \tag{2.40}$$

where v is the Fresnel–Kirchhoff parameter and $j = \sqrt{-1}$. Equation (2.40) is known as the *complex Fresnel integral*. There is no known closed-form solution to this equation, but it can be evaluated numerically. In Fig. 2.10, we plot the diffraction loss over a knife edge as a function of the Fresnel–Kirchhoff parameter v. Note how the loss oscillates around 0 dB when the knife edge is below the first Fresnel zone. Once the screen obstructs that zone and beyond, there is a steady decrease in field strength. When the knife edge is even with the line-of-sight path ($h = 0$), the electric field is reduced by one-half and there is a 6-dB loss in signal power.

For a fixed obstruction at a distance d_1 from the transmitter and a transmission wavelength λ, the diffraction parameter

$$v = h\sqrt{\frac{2(d_1 + d_2)}{\lambda d_1 d_2}} \tag{2.41}$$

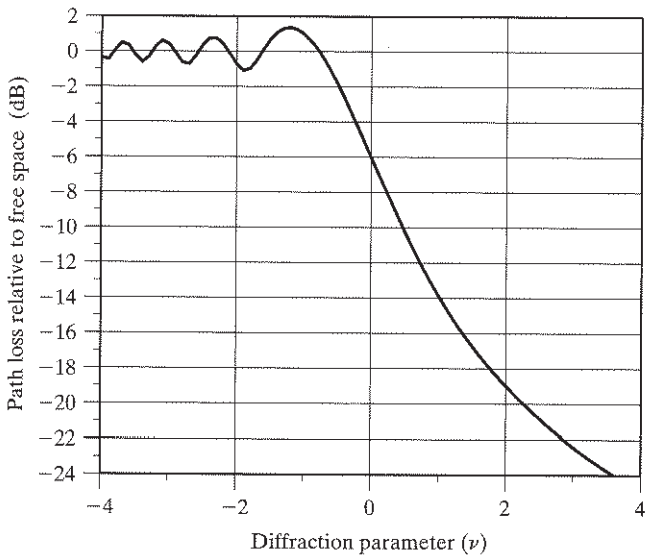


FIGURE 2.10 Diffraction loss over a single knife edge as a function of the Fresnel–Kirchhoff parameter (ν).

increases toward infinity as the distance d_2 from the obstruction to the receiver shrinks. This relationship implies a loss becoming infinitely large as the receiver moves more and more into the shadow of the obstruction. The free-space loss referred to in Fig. 2.10 is that corresponding to the straight-line distance between the transmitter and receiver in the absence of the obstruction. Note that the diffraction loss depends on frequency through the wavelength λ that appears in the definition of the diffraction parameter. In addition to the loss introduced relative to the direct path, there will be a phase rotation, as suggested by the presence of the imaginary unit j in Eq.(2.40).

This subsection has described the diffraction loss due to an abrupt "knife-edge" obstruction. The loss provides a second physical model as to why losses may be greater than those predicted by the free-space model. In practice, there will often be obstructions of other shapes and multiple obstructions along the transmission path. The principles for calculating the diffraction losses are the same as described herein, although the calculations involved are more difficult and are frequently done by computer simulations.

Problem 2.5 In Problem 2.4, suppose the middle tower was 80 m and the shorter tower was only 30 m. The separation between the two communicating towers is 2 km. What would the increase in path loss be in this case relative to the free-space loss? How would the diffraction loss be affected if the transmission frequency is decreased from 4 GHz to 400 MHz?
Ans. The losses are approximately 22 dB for 4 GHz and 13 dB for 400 MHz. ■

2.4 TERRESTRIAL PROPAGATION: STATISTICAL MODELS[9]

The previous sections have identified three modes of propagation:

1. *Free-space propagation,* in which the received power decreases as the square of the distance from the transmitter.
2. *Reflection,* wherein, for the plane-Earth model, the received power decreases as the fourth power of distance. For other reflection models, we may expect power decreases at other powers, typically greater than two.
3. *Diffraction*, which, for a knife-edge model, introduces a constant attenuation that depends on the proportion of the direct path that is blocked. For a terrestrial radio link, the signal may be diffracted a number of times along its path; consequently, we would expect diffraction losses to also exhibit some dependence on distance.

We have seen that, even with simple physical models for these propagation modes, the analysis can be quite complicated. With the *physical approach*, we build up a model of the environment, including the terrain, buildings, and other features that may affect propagation. With this physical model, the possible propagation paths are determined; the process is often referred to as *ray tracing*. For the different paths, the techniques described in the previous sections are used to estimate the path losses between the transmitter and the receiver. In some cases, signals may undergo multiple reflections, multiple diffractions, or a combination of both. This complicated calculation is often programmed and determined by computer. The results, however, apply only to the particular physical model being analyzed.

Alternatively, we may use a *statistical approach* in which the propagation characteristics are empirically approximated on the basis of measurements in certain general types of environments, such as urban, suburban, and rural. The statistical approach is broken down into two components: an estimate of the *median path loss* and a component representing *local variations.* These issues are discussed in the subsections that follow.

2.4.1 Median Path Loss

In the general case, the received signal is the sum of several versions of the transmitted signal received over different transmissions paths. That is, if $\tilde{E}$ is the total received electrical field and $\tilde{E}_d$ is the electrical field of an equivalent direct path, then, assuming that there are N different paths between the transmitter and receiver due to different reflections, etc., the total electric field is

$$\tilde{E} = \tilde{E}_d \sum_{k=1}^{N} L_k e^{j\phi_k} \tag{2.42}$$

The $\{L_k\}$ represent the relative losses for the different paths, and the $\{\phi_k\}$ represent the relative phase rotations. If a direct path exists, then it would nominally be characterized by $L_0 = 1$ and $\phi_0 = 0$.

Empiricists have performed measurements of the field strength in various environments as a function of the distance from the transmitter to the receiver. Together, these investigations motivate a general propagation model for median path loss having the form

$$\frac{P_R}{P_T} = \frac{\beta}{r^n} \tag{2.43}$$

where the *path-loss exponent n* typically ranges from 2 to 5, depending on the propagation environment. The parameter β represents a loss that is related to frequency and that may also be related to antenna heights and other factors. The right-hand side of Eq. (2.43) is sometimes written in the equivalent logarithmic form

$$L_p = \beta_0(\text{dB}) - 10n \log_{10}(r/r_0) \tag{2.44}$$

In this representation, β_0 *represents the measured path loss at the reference distance* r_0, typically 1 meter. *In the absence of other information,* β_0 *is often taken to be the free-space path loss at a distance of 1 meter.*

Numerous propagation studies have been carried out in an attempt to identify the different environmental effects. The conclusion drawn from these studies is that most of these effects are locally dependent and difficult, if not impossible, to characterize in general. Some sample values for n in the model of Eq. (2.43) are shown in Table 2.1. An example of a more complex model, called the *Okumura–Hata model*, is provided in Section 2.10.

While this table indicates general trends, there are exceptions. For example, if the propagation path is along a straight street with skyscrapers on either side, there may be a ducting (waveguide) effect, and propagation losses may be similar to those occurring in free space or be even less. In general, performance in open spaces is not as poor as is predicted by the plane-Earth model.

This model for median path loss is quite flexible, and it is intended for the analytical study of problems, as it allows us to parameterize the performance of various system-related factors. For commercial applications in a terrestrial environment, either a field measurement campaign or detailed modeling of the environment is preferred.

TABLE 2.1 Sample path-loss exponents.

Environment	n
Free space	2
Flat rural	3
Rolling rural	3.5
Suburban, low rise	4
Dense urban, skyscrapers	4.5

2.4.2 Local Propagation Loss

In the previous section, a model for predicting the median path loss was presented. For any particular site, there will be a variation from this median value that will depend on the local characteristics. Several investigators have measured the variation about the median and have suggested that it can be modeled as a *lognormal distribution*. In particular, if μ_{50} is the median value of the path loss (in dB) at a specified distance r from the transmitter, then the distribution x_{dB} of observed path losses at this distance has the probability density function

$$f(x_{\text{dB}}) = \frac{1}{\sqrt{2\pi}\sigma_{\text{dB}}} e^{-(x_{\text{dB}} - \mu_{50}/2\sigma_{\text{dB}}^2)} \tag{2.45}$$

Equation (2.45) is known as the *lognormal model* for local *shadowing*. It is the Gaussian distribution in which all quantities are measured in decibels. (See Appendix C for a description of the Gaussian distribution.) Typical values for the standard deviation σ_{dB} range from 5 to 12 dB. In Fig. 2.11, the lognormal distribution

$$\text{Prob}(x_{\text{dB}} > x) = \int_x^{\infty} f(x_{\text{dB}}) dx_{\text{dB}} \tag{2.46}$$

is plotted. For a σ_{dB} of 6 dB, the figure indicates that the loss relative to the median path is greater than 10 dB only 1% of the time. Since, by definition of the median, 50% of the samples are expected to be less than the median and 50% greater, regardless of the shadowing, we find that all curves intersect at the median point.

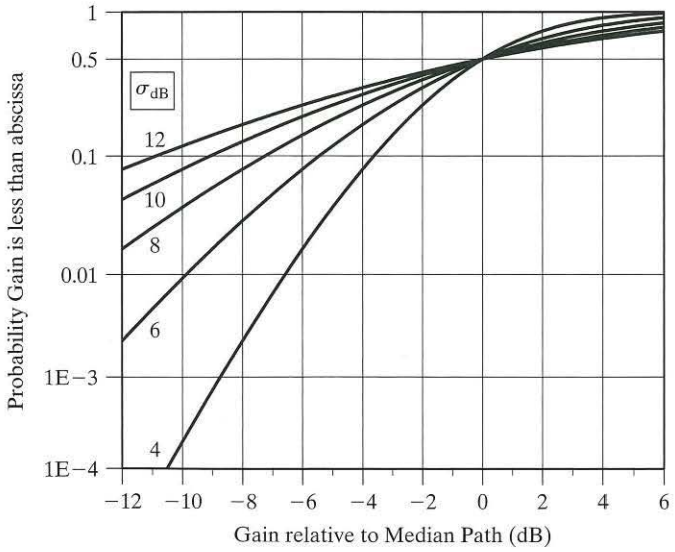


FIGURE 2.11 The lognormal distribution.

We say that a service has 90% *availability* if the received signal level is above receiver sensitivity over 90% of the region of coverage. The availability requirement will depend on the kind of service being offered. For example, safety services require a higher availability than non-safety services.

EXAMPLE 2.4 Availability

A measurement campaign in a large city indicates that the propagation can be reasonably well modeled with a loss exponent of $n = 2.9$. The shadowing deviation about this loss is 6 dB. What is the range of coverage if 99% availability is required for a public-safety radio application? Assume that the receiver sensitivity is –100 dBm and the measured power at 10 meters is 2 milliwatts.

If 99% availability is required with a 6-dB shadowing deviation, then Fig. 2.11 indicates that an extra 10 dB of margin on top of the median path loss is required. Consequently, the median signal power at the edge of coverage must be –90 dBm. Assuming the use of isotropic antennas, the statistical model for median path loss is

$$\begin{aligned} P_R(\text{dBm}) &= P_T(\text{dBm}) + L_p(\text{dB}) \\ &= P_T + \beta_0 - 10n\log_{10}(r/r_0) \\ &= 3 - 10n\log 10(r/10) \end{aligned}$$

where we have used $P_T + \beta_0 = 3$ dBm at a distance $r_0 = 10$ meters.

Solving for r, we obtain

$$\begin{aligned} 10n\log_{10}(r/10) &= 3 - P_R(\text{dBm}) \\ &= 3 - (-90) \\ &= 93 \text{ dB} \end{aligned}$$

or

$$r = 10\times 10^{93/10n} = 16.1 \text{ km.}$$ ■

Problem 2.6 A brief measurement campaign indicates that the median propagation loss at 420 MHz in a midsize North American city can be modeled with $n = 2.8$ and a fixed loss (β) of 25 dB; that is,

$$L_p = 25\text{dB} + 10\log_{10}(r^{2.8})$$

Assuming a cell phone receiver sensitivity of –95 dBm, what transmitter power is required to service a circular area of radius 10 km? Suppose the measurements were optimistic and $n = 3.1$ is more appropriate. What is the corresponding increase in transmit power that would be required?

Ans. A transmit power of 12 dBW is required for $n = 2.8$ and 24 dBW for $n = 3.1$. ■

2.5 INDOOR PROPAGATION[10]

With the growth of cellular telephone usage, these appliances are being used more and more in indoor locations such as shopping malls, office buildings, train stations, and

airports. It is important that wireless design take into account the propagation characteristics in these high-density locations.

In other applications, wireless local area networks (LANs) are being implemented in buildings and on campuses so as to eliminate the cost of wiring or rewiring buildings to upgrade or provide new services. For these applications, understanding the physics of indoor applications is a key to determining how the locations and numbers of transmitters and receivers will affect the desired quality of service to the user.

In parallel with previous sections, there are two common methods of analyzing indoor propagation: ray tracing, described in Section 2.4, and modifying the free-space model to fit empirical/statistical data and then applying the model for prediction purposes.

The statistical approach provides a more easily understood model. A simple model for the indoor path loss is given by

$$L_p(\text{dB}) = \beta(\text{dB}) + 10\log_{10}\left(\frac{r}{r_0}\right)^n + \sum_{p=1}^{P} \text{WAF}(p) + \sum_{q=1}^{Q} \text{FAF}(q) \qquad (2.47)$$

where r is the distance separating the transmitter from the receiver, r_0 is the nominal reference distance (typically, 1m), n is the path-loss exponent, WAF(p) is the wall attenuation factor, FAF(q) is the floor attenuation factor, and P and Q are the number of walls and floors, respectively, separating the transmitter and the receiver.

With indoor propagation, the path-loss exponent n is often chosen to be 2 for small separations, and higher values are selected for greater separations. A typical wall attenuation factor for hollow plaster walls is about 5 dB at 900 MHz; for concrete walls, the attenuation factor is typically 10 dB at the same frequency. The attenuation of concrete floors is usually slightly higher than that of concrete walls, most likely due to the presence of steel beams and re-enforcing rods. These attenuation factors increase as the frequency increases. Some measurements at 1700 MHz indicate a 6-dB increase in the attenuation due to floors and walls over the 900-MHz case. This path-loss model is a relatively simple extension of the models described in earlier sections. It tends to be most accurate when the number of walls and floors is small (less than five).

The shortcomings of the statistical model are that it does not take into account such matters as the dependence of the path-loss exponent on distance and the dependence of the WAF on the angle of incidence. If the transmitter is outside the building, then there is an additional building penetration loss, the size of which depends on the frequency of operation and what floor of the building the receiver is located on.

EXAMPLE 2.5 Indoor Propagation

Suppose, in an office building, a 2.4-GHz transmitter located at a workstation is separated from the network access node (receiver) by a distance of 35 m. The transmission must pass through 5 m of an office, through a plasterboard wall, and then through a large open area. The propagation is modeled as free space for the first 5 m and with a loss exponent of 3.1 for the remainder of the distance. The plasterboard wall causes 6-dB attenuation of the signal. The isotropic transmitter radiates 20 dBm. Can the link be closed if the receiver has a sensitivity of –75 dBm?

TABLE 2.2 Link budget for Example 2.5.

Parameter	Value	Comments
Transmit Power	20 dBm	
Free-space loss	54 dB	$L_p = \left(\frac{4\pi R}{\lambda}\right)^2$
Wall attenuation	6 dB	
Open-area loss	26.2 dB	$L_p = (r/r_0)^{3.1}$ with $r/r_0 = 35/5$
Received power	-66.2 dBm	assuming 0-dB receive antenna gain

The propagation path is summarized in Table 2.2. The line items in the table are as follows:

1. The first line of the table is the transmitted power, in dBm.
2. The second line is the free-space loss in the 5 m surrounding the transmitter, using the formula for free-space path loss as shown.
3. The third line is the loss through the plasterboard wall.
4. The fourth line is the relative loss in the open area. To calculate this loss, we assume that the reference distance for the path-loss calculation in the open area is $r_0 = 5$ m. The fourth line is the relative loss between a distance of 5 m and a distance of 35 m.
5. The fifth line is the received power level, obtained by subtracting the previous three lines from the first.

The answer is yes, the link can be closed. ■

Problem 2.7 Using the same model as in Example 2.5, predict the path loss for the site geometry shown in Fig. 2.12. Assume that the walls cause an attenuation of 5dB and the floors 10 dB.
Ans. –67.2 dBm. ■

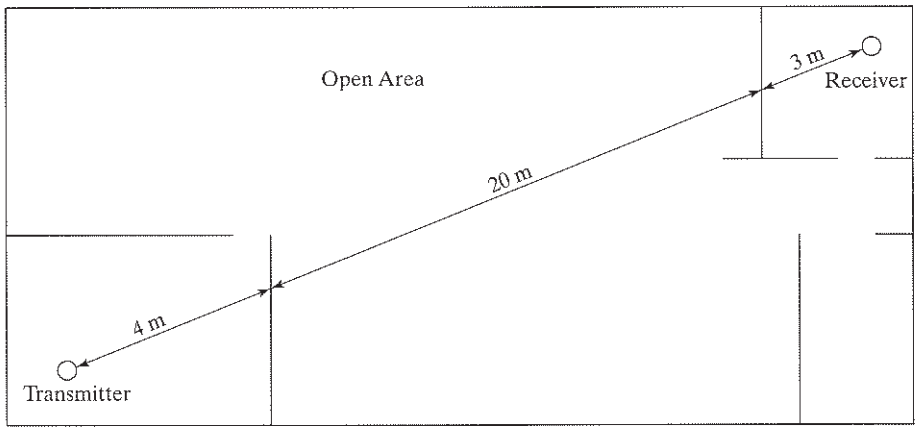


FIGURE 2.12 Site geometry for Problem 2.7.

2.6 LOCAL PROPAGATION EFFECTS WITH MOBILE RADIO[11]

Most mobile communication systems are used in and around centers of population. The major difficulties are caused by the fact that the mobile antenna is well below the surrounding buildings; thus, most communication is via scattering of electromagnetic waves from surfaces or diffraction over and around buildings. These multiple propagation paths, or *multipaths*, have both slow and fast aspects:

1. *Slow fading* arises from the fact that most of the large reflectors and diffracting objects along the transmission path are distant from the terminal. The motion of the terminal relative to these distant objects is small; consequently, the corresponding propagation changes are slow. These factors contribute to the median path losses between a fixed transmitter and a fixed receiver that were described in the previous sections. The statistical variation of these mean losses due to variation of the intervening terrain, vegetation, etc., was modeled as a lognormal distribution for terrestrial applications. The slow-fading process is also referred to as *shadowing* or *lognormal fading*.
2. *Fast fading* is the rapid variation of signal levels when the user terminal moves short distances. Fast fading is due to reflections of local objects and the motion of the terminal relative to those objects. That is, the received signal is the sum of a number of signals reflected from local surfaces, and these signals sum in a constructive or destructive manner, depending on their relative phase relationships, as illustrated in Fig. 2.13. The resulting phase relationships are dependent on relative path lengths to the local objects, and they can change significantly over short distances. In particular, the phase relationships depend on the speed of motion and the frequency of transmission.

In this section, we begin with a statistical analysis of the effects of local reflectors and then consider the case in which the terminal is moving in the environment.

2.6.1 Rayleigh Fading

In wireless communications, a distinction is made between portable terminals and mobile terminals. *Portable terminals* can be moved easily from one place to another, but communications occur when the terminal is stationary. *Mobile terminals* can also be easily moved, but, with them, communications may occur while the terminal is moving. In this section, we consider portable terminals and how local propagation effects may differ at different positions. Subsequently, we will consider mobile terminals.

Consider a stationary receiver that might be at position I_1 or I_2, as shown in Fig. 2.13. We wish to characterize the amplitude distribution of the received signal over a variety of such positions. We model the case in which the transmitted signal reaches a stationary receiver via multiple paths, but the path differences are due only to local reflections. The complex phasor of the N signal rays (reflections) is given by

$$\tilde{E} = \sum_{n=1}^{N} E_n e^{j\theta_n} \tag{2.48}$$

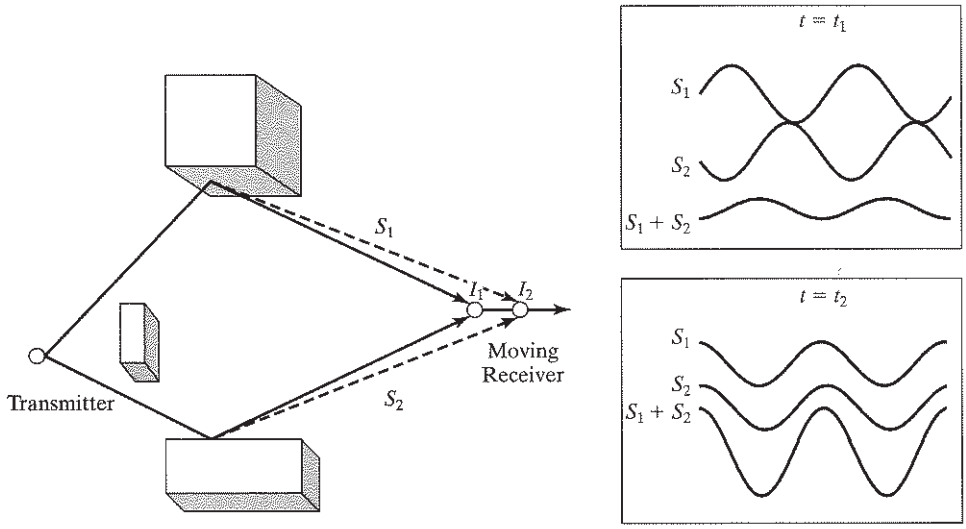


FIGURE 2.13 Illustration of constructive and destructive interference.

where E_n is the electric field strength of the nth path and θ_n is its relative phase. The phasor $\tilde{E}$ is a random variable representing the multiplicative effects of the multipath channel. As seen previously, small differences in path length can make large differences in phase. Since the reflections can arrive from any direction, we assume that the relative phases are independent and uniformly distributed over $[0,2\pi]$.

With Eq. (2.48) as the model for the complex channel envelope, we see that it is a sum of independent and identically distributed (i.i.d.) complex random variables. From the *central limit theorem* of statistics, the sum of a number of i.i.d. random variables approaches the Gaussian distribution as the number of random variables gets large. (See Appendix C for a review of probability theory and random processes.) That is,

$$\sum_{n=1}^{N} E_n e^{j\theta_n} \rightarrow Z_r + jZ_i \text{ as } N \rightarrow \infty \tag{2.49}$$

where Z_r and Z_i are real Gaussian random variables. If we consider one of the components of the sum, then we find that the expectation of each component is

$$\begin{aligned} \mathbf{E}[E_n e^{j\phi_n}] &= \mathbf{E}[E_n]\mathbf{E}[e^{j\theta_n}] \\ &= \mathbf{E}[E_n]\frac{1}{2\pi}\int_0^{2\pi} e^{j\theta} d\theta \\ &= 0 \end{aligned} \tag{2.50}$$

since the mean of a complex phasor with a uniformly distributed phase is zero; **E** denotes the *statistical expectation operator*. Consequently, the mean of the complex envelope is given by

$$\begin{aligned}\mathbf{E}[\tilde{E}] &= \mathbf{E}\left[\sum_{n=1}^{N} E_n e^{j\theta_n}\right] \\ &= \sum_{n=1}^{N} E_n \mathbf{E}[e^{j\theta_n}] \\ &= 0\end{aligned} \tag{2.51}$$

The variance (power) in the complex envelope is given by the mean-square value

$$\begin{aligned}\mathbf{E}[|\tilde{E}|^2] &= \mathbf{E}\left[\sum_{n=1}^{N} E_n e^{j\theta_n} \sum_{m=1}^{N} E_m e^{-j\theta_m}\right] \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N} E_n E_m \mathbf{E}[e^{j(\theta_n - \theta_m)}] \\ &= \sum_{n=1}^{N} E_n^2 \\ &= P_0\end{aligned} \tag{2.52}$$

where we have used the fact that the difference of two random phases is also a random phase, unless the two phases are equal. By symmetry, this power is equally distributed between the real and imaginary parts of the complex envelope.

Since the complex envelope has zero mean, the probability density function of Z_r in Eq. (2.49) is given by the Gaussian density function

$$f_{z_r}(z_r) = \frac{1}{\sqrt{2\pi}\sigma} e^{-z_r^2/2\sigma^2} \tag{2.53}$$

where $\sigma^2 = P_0/2$, and similarly for Z_i. We define the amplitude of the complex envelope as

$$R = \sqrt{Z_i^2 + Z_r^2} \tag{2.54}$$

The probability density function of the amplitude is determined from the density function of Eq. (2.53) by performing the appropriate change of variables to obtain

$$f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \tag{2.55}$$

(See Appendix C for details.) Equation (2.55) is known as the *Rayleigh probability density function*. Integrating this density function yields the corresponding cumulative probability distribution function:

$$\begin{aligned} \text{Prob}(r < R) &= \int_0^R f_R(r)dr \\ &= 1 - e^{-R^2/2\sigma^2} \end{aligned} \tag{2.56}$$

The mean value of the Rayleigh distribution (the mean value of the absolute envelope) is given by

$$\begin{aligned} \mathbf{E}[R] &= \int_0^\infty r f_R(r)dr \\ &= \sigma\sqrt{\frac{\pi}{2}} \end{aligned} \tag{2.57}$$

and the mean-square value is given by

$$\begin{aligned} \mathbf{E}[R^2] &= \int_0^\infty r^2 f_R(r)dr \\ &= 2\sigma^2 \\ &= R_{\text{rms}}^2 \end{aligned} \tag{2.58}$$

That is, the root-mean-square (rms) amplitude is $R_{\text{rms}} = \sqrt{2}\sigma$.

The Rayleigh distribution is plotted in Fig. 2.14. The median of the distribution is $R = R_{\text{rms}}$, which implies that there is constructive interference ($R > R_{\text{rms}}$) for 50% of locations and destructive interference ($R < R_{\text{rms}}$) for 50% of locations. Conceptually, each location corresponds to a different set of $\{\phi_n\}$. Deep fades of 20 dB or more ($R < 0.1R_{\text{rms}}$) occur only rarely (with a probability of 1%). However, Fig. 2.14 indicates that there can be a wide variation in received signal strength due to local reflections.

These results are for portable receivers. In any one position, the received power will be constant. If we take measurements at a number of random locations, then the received power measurements would show a Rayleigh distribution. For a truly stationary receiver, we would choose a location that minimizes the local reflections and provides the maximum received signal strength.

EXAMPLE 2.6 Margin for Rayleigh Fading

Recall Example 2.4, in which we showed that the system would need 10 dB extra power (*margin*) to provide 99% availability when 6-dB lognormal fading (slow fading) was taken into

account. Suppose this same system also undergoes fast fading. What additional margin would be required to maintain the same availability?

From Fig. 2.14, we see that, if the receiver is to operate above threshold for 99% of the locations, then it must be able to tolerate the amplitude dropping to as low as 10% of the rms value, or a fade depth of 20 dB. That is, a 20-dB additional margin would be required for Rayleigh fading in order to maintain 99% availability.

In practice, a well-designed receiver that uses forward error-correction coding (discussed in Chapter 4) can usually significantly reduce the margin required to compensate Rayleigh fading. ■

Problem 2.8 What are the required margins for lognormal and Rayleigh fading in Example 2.6 if the availability requirement is only 90%?

Ans. The required margins are 5.5 dB for shadowing and 10 dB for Rayleigh fading. ■

2.6.2 Rician Fading

The Rayleigh-fading model assumes that all paths are relatively equal—that is, that there is no dominant path. Despite the fact that Rayleigh fading is the most popular model, occasionally there is a direct line-of-sight path in mobile radio channels and in indoor wireless as well. The presence of a direct path is usually required to close the link in satellite communications. In this case, the reflected paths tend to be weaker than the direct path, and we model the complex envelope as

$$\tilde{E} = E_0 + \sum_{n=1}^{N} E_n e^{j\theta_n} \tag{2.59}$$

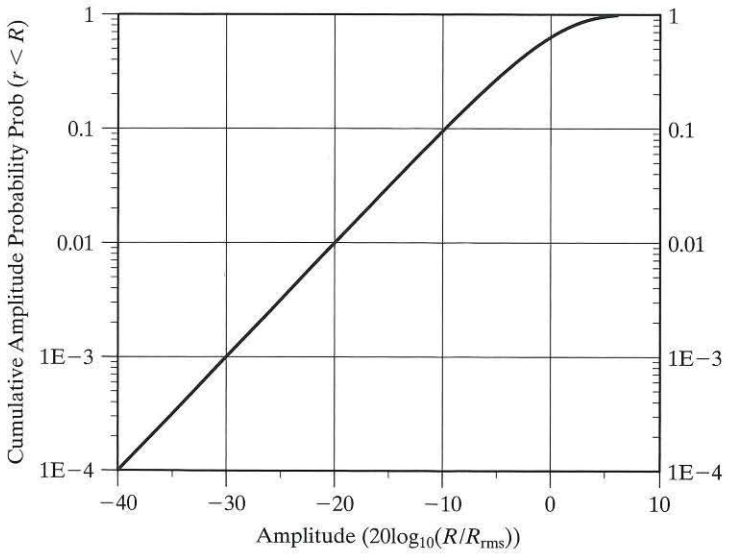


FIGURE 2.14 Rayleigh amplitude distribution (R/R_{rms}).

where the constant term represents the direct path and the summation represents the collection of reflected paths. This model is referred to as a *Rician fading* model. The analysis proceeds in a manner similar to that of the Rayleigh fading case, but with the addition of a constant term. A key factor in the analysis is the ratio of the power in the direct path to the power in the reflected paths. This ratio is referred to as the *Rician K-factor*, defined as

$$K = \frac{s^2}{\sum_{n=1}^{N} |E_n|^2} \tag{2.60}$$

where $s^2 = |E_0|^2$. The Rican K-factor is often expressed in dB.

The calculation of the amplitude density function in the Rician fading case is more involved than with Rayleigh fading, so we merely give the result here. We have

$$f_R(r) = \frac{r}{\sigma^2} e^{-(r^2+s^2)/2\sigma^2} I_0\left(\frac{rs}{\sigma^2}\right) \quad r \geq 0 \tag{2.61}$$

where $I_0(.)$ is the *modified Bessel function* of zeroth order.[8] (See Appendix C for more details on the Rician distribution.)

The amplitude distributions of the Rician fading channel for different K factors are shown in Fig. 2.15. Deep fades are clearly less probable than with the Rayleigh channel, and the probability of their occurrence decreases as the K factor increases.

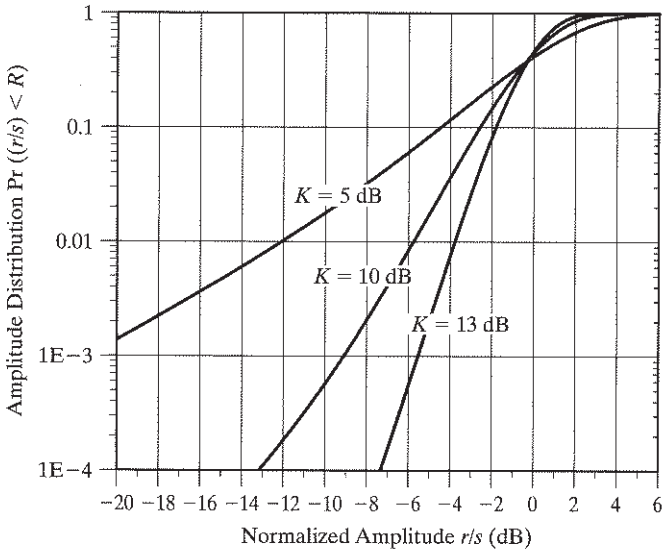


FIGURE 2.15 Amplitude distribution for the Rician channel.

2.6.3 Doppler

Just as a train whistle or car horn appears to have a different pitch, depending on whether it is moving toward or away from one's location, radio waves demonstrate the same phenomenon. If a receiver is moving toward the source, then the zero crossings of the signal appear faster, and consequently, the received frequency is higher. The opposite effect occurs if the receiver is moving away from the source. The resulting change in frequency is known as the *Doppler shift*.

To model the Doppler shift phenomenon, consider Fig. 2.16, which shows a fixed transmitter and a receiver moving at a constant velocity away from the transmitter. If the complex envelope of the signal emitted by the transmitter is $Ae^{j2\pi f_0 t}$, then the signal at a point along the x-axis is given by

$$\tilde{r}(t, x) = A(x)e^{j2\pi f_0(t - x/c)} \tag{2.62}$$

where $A(x)$ represents the amplitude as a function of distance and c is the speed of light. Equation (2.62) shows the signal has a phase rotation that depends on the distance of the signal from the source. If x represents the position of the constant-velocity receiver, then we may write

$$x = x_0 + vt \tag{2.63}$$

where x_0 is the initial position of the receiver and v is its velocity away from the source. Substituting Eq. (2.63) into (2.62), we find that the signal at the receiver is

$$\begin{aligned} \tilde{r}(t) &= A(x_0 + vt)e^{j2\pi f_0\left(t - \frac{x_0 + vt}{c}\right)} \\ &= A(x_0 + vt)e^{-j2\pi f_0 x_0/c} e^{j2\pi f_0(1 - v/c)t} \end{aligned} \tag{2.64}$$

If we focus on the frequency term in the last exponent of Eq. (2.64), we discover that the received frequency is given by

$$f_r = f_0\left(1 - \frac{v}{c}\right) \tag{2.65}$$

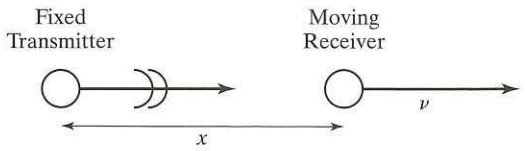


FIGURE 2.16 Illustration of Doppler effect.

where f_0 is the carrier transmission frequency. The *Doppler shift* is given by

$$f_D = f_r - f_0 = -f_0 \frac{v}{c} \tag{2.66}$$

Consequently, the relationship between the observed Doppler frequency shift and the velocity away from the source is

$$\frac{v}{c} = -\frac{f_D}{f_0} \tag{2.67}$$

Equation (2.67) describes the case when the direction of terminal motion and the radiation are collinear. This is actually the maximum Doppler shift that can be observed. If the terminal motion and the direction of radiation are at an angle ψ, the more general expression for the shift is

$$f_D = -\frac{f_0}{c} v \cos\psi \tag{2.68}$$

For operating frequencies between 100 MHz and 2 GHz, and for speeds up to 100 km/hr, the Doppler shift can be as large as 185 Hz. This Doppler shift appears as a carrier offset that varies with the speed and direction of the vehicle. Consequently, any mobile receiver must have the capability of both receiving a signal with frequency offsets and tracking fast changes in the frequency offset as the receiver terminal changes velocity.

EXAMPLE 2.7 Aircraft Doppler

An aircraft is headed toward an airport control tower with a speed of 500 km/hr at an elevation of 20°. Safety communications between the aircraft tower and the plane occur at a frequency of approximately 128 MHz. What is the expected Doppler shift of the received signal?

The velocity of the aircraft in the direction away from the tower is

$$v = -500 \text{ km/hr} \times \cos 20° = -130 \text{m/s}$$

The corresponding Doppler shift is

$$\begin{aligned} f_D &= -\frac{f_0}{c} v \cos\alpha \\ &= -\frac{128\times10^6}{3\times10^8} \times (-130) \\ &= 55 \text{ Hz} \end{aligned}$$

If the plane banks suddenly and heads in the other direction, the Doppler shift will change from 55 Hz to –55 Hz. This rapid change in frequency, df_D/dt, is sometimes referred to as *frequency slewing*. A good mobile receiver must be capable of tracking the frequency slew rates that can be generated by the receiver's motion. ■

Problem 2.9 Suppose that the aircraft in Example 2.7 has a satellite receiver operating in the aeronautical mobile–satellite band at 1.5 GHz. What is the Doppler shift observed at this receiver? Assume that the geostationary satellite has a 45° elevation with respect to the airport.
Ans. The Doppler shift for this receiver is 491 Hz. ■

2.6.4 Fast Fading

If we measure the received signal strength as a function of time for a mobile terminal, we may observe a rapid variation in the signal strength about a median value. This phenomenon is known as *fast fading*. From the previous section on stationary receivers, we would expect the signal power at different locations to have a Rayleigh distribution. In Fig. 2.17, we show a sample trace of the received signal power for a terminal moving at 100 km/hr in a Rayleigh fading environment. The frequency of transmission is 900 MHz.

Over short periods, the received power has a relatively smooth behavior, due to the correlation of the fading at adjacent positions. However, there are the occasional deep fades in the level of the received signal, as would be expected with a Rayleigh distribution. It is the correlation of the fading process that we will investigate in this section.

To model fading with a moving receiver, we expand the model used for stationary fading, given by Eq.(2.48), to include the effects of the Doppler shift. Let

$$\tilde{E}(t) = \sum_{n=1}^{N} E_n e^{j(2\pi f_n t + \phi_n)} \tag{2.69}$$

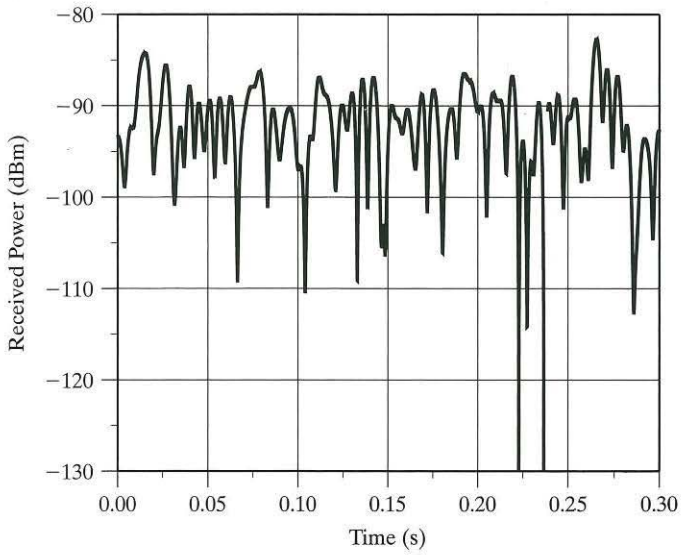


FIGURE 2.17 Illustration of variations in received signal power due to Rayleigh fading.

represent the *complex envelope* of N signal rays. The complex envelope of the channel $\tilde{E}(t)$ is time varying. Since the receiver is moving, each ray has a different Doppler shift $\{f_n\}$ because the relative directions of motion of the rays will be different. The rays are assumed to come from arbitrary angles ψ_n surrounding the receiver, as shown in Fig. 2.18. We assume that all rays are arriving from a horizontal direction; this is known as the *Clarke model.*[12] As before, we assume that the relative phase of each ray is i.i.d.

To study the time dependency of the fading phenomenon, we look at the autocorrelation of the complex envelope and so write

$$R_{\tilde{E}}(\tau) = \mathbf{E}[\tilde{E}(t)\tilde{E}^*(t+\tau)] \tag{2.70}$$

where the asterisk denotes complex conjugation and τ is a time offset. We begin by computing the cross-correlation between a pair of signal rays. This correlation is given by

$$\begin{aligned}\mathbf{E}\left[E_n e^{j(2\pi f_c t+\phi_n)} E_m e^{-j(2\pi f_m(t+\tau)+\phi_m)}\right] &= \mathbf{E}[E_n e^{j2\pi f_n t} E_m e^{-j2\pi f_m(t+\tau)}]\mathbf{E}[e^{j\theta_n} e^{j\theta_m}] \\ &= \begin{cases} 0 & \text{if } m \neq n \\ \mathbf{E}[E_n^2 e^{-j2\pi f_n \tau}] & \text{if } m = n \end{cases}\end{aligned} \tag{2.71}$$

due to the i.i.d. assumption. Using the result of Eq.(2.71) with Eq. (2.69) in Eq. (2.70),

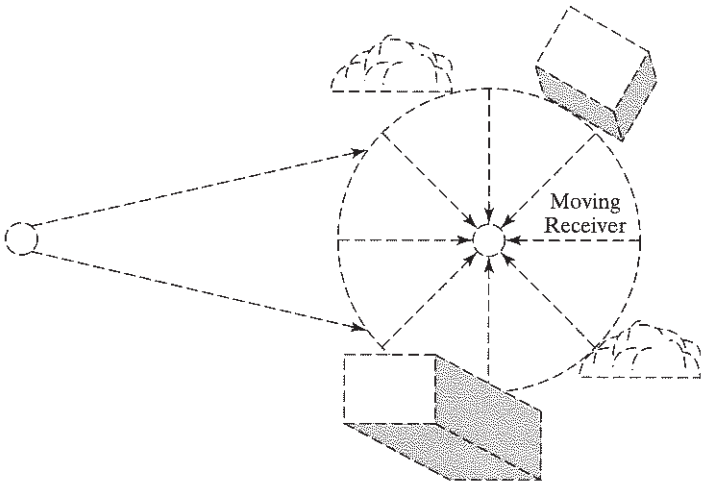


FIGURE 2.18 Illustration of local multipath arriving at a receiver.

we obtain the *autocorrelation of the complex envelope*, given by

$$\begin{aligned} R_E(\tau) &= \mathbf{E}\left[\sum_{n=1}^{N} E_n e^{j(2\pi f_n t + \theta_n)} \sum_{m=1}^{N} E_m e^{-j(2\pi f_m (t+\tau) + \theta_m)}\right] \\ &= \sum_{n=1}^{N} \mathbf{E}[E_n^2 e^{-j2\pi f_n \tau}] \\ &= \sum_{n=1}^{N} \mathbf{E}[E_n^2]\mathbf{E}[e^{-j2\pi f_n \tau}] \end{aligned} \tag{2.72}$$

where we assume that the effects of small changes in distance on amplitude $\{E_n\}$ are negligible. Recall from Section 2.6.3 that the Doppler shift is proportional to the angle of the radiation relative to the direction of motion; that is,

$$f_n = f_D \cos \psi_n \tag{2.73}$$

where f_D is the *maximum Doppler shift*. The maximum shift occurs for signal rays that are in the same direction as the motion of the terminal. Under the assumption that the multipath is uniformly distributed over the range $[-\pi,\pi]$, the expectation of Eq. (2.72) is independent of n, in which case we have

$$R_E(\tau) = P_0 \mathbf{E}[e^{-j2\pi f \tau}] \tag{2.74}$$

where $P_0 = \sum_{n=1}^{N} \mathbf{E}[E_n^2]$ is the average received power. Substituting Eq. (2.73) into (2.74), we finally obtain

$$\begin{aligned} R_E(\tau) &= P_0 \mathbf{E}[e^{-j2\pi f_D \tau \cos \psi}] \\ &= P_0 \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j2\pi f_D \tau \cos \psi} d\psi\right) \\ &= P_0 J_0(2\pi f_D \tau) \end{aligned} \tag{2.75}$$

where $J_0(x)$ is the *zeroth-order Bessel function* of the first kind, defined by

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} e^{-jx\cos\theta} d\theta \tag{2.76}$$

In summary, the correlation of the fading with time (or position) is described by Eq. (2.75). Figure 2.19 plots this autocorrelation function of the complex envelope as a function of the normalized parameter $2\pi f_D \tau$. As expected, the autocorrelation peaks at $\tau = 0$, but there is strong correlation for normalized delays as large as $2\pi f_D \tau = 1$. The autocorrelation is symmetric in τ.

The power spectrum of the fading process is determined by the Fourier transform of the autocorrelation function of the envelope and is given by

$$\begin{aligned} S_E(f) &= \mathbf{F}[R_E(\tau)] \\ &= \mathbf{F}[P_0 J_0(2\pi f_D \tau)] \\ &= \begin{cases} \dfrac{P_0}{\sqrt{1-(f/f_D)^2}} & f<f_D \\ 0 & f>f_D \end{cases} \end{aligned} \tag{2.77}$$

where $\mathbf{F}$ is the Fourier transform operator. Figure 2.20 plots Eq. (2.77) for $P_0 = 1$. The power spectrum is zero at frequencies greater than the Doppler shift f_D and has discontinuities at the edge of the nonzero band. This spectrum is quite similar to what is actually observed, although, in practice, the discontinuities are not present.

Implicit in the Clarke model for fast fading in mobile terrestrial communications is the assumption that all rays are arriving from a horizontal direction. In practice, however, there may be a vertical component to the waves. More sophisticated models for mobile terrestrial communications have taken this component into account, but their conclusions are similar to those presented here. Empirical measurements tend to support the Clarke model, with a Doppler bandwidth related to the frequency of transmission and the mobile velocity. The measured spectra usually show the peaks near the Doppler frequencies, as depicted in Fig. 2.20.

Empirical investigations have indicated that the Clarke model is not always a valid description of the fading process between a mobile terminal and satellite. For example, with aeronautical terminals and satellites, it has been found that a Gaussian

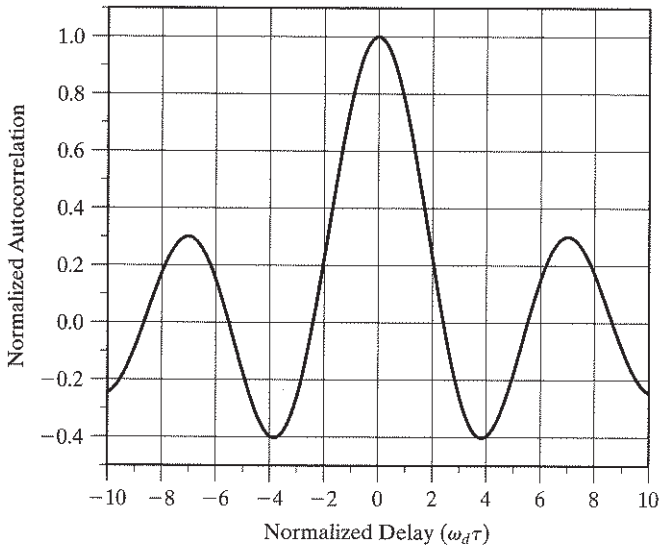


FIGURE 2.19 Autocorrelation of the complex envelope of the received signal according to Clarke's model.

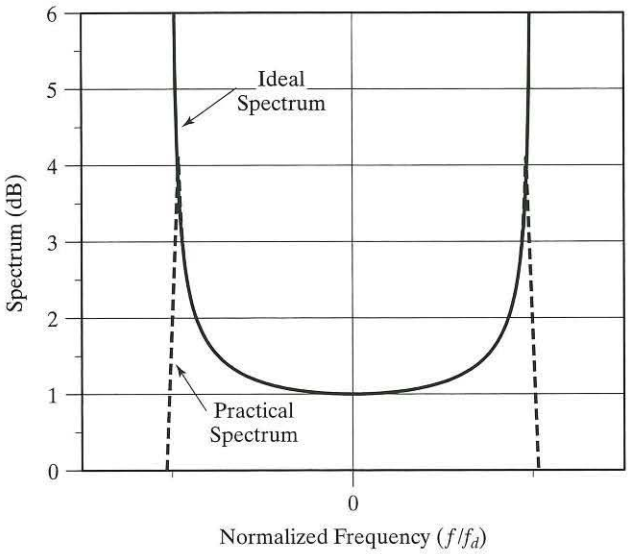


FIGURE 2.20 Power spectrum of the fading process for Clarke's model.

spectrum is a better model of the spectrum of the fading process and that the maximum Doppler frequency of the fading process is *not* proportional to the aircraft's speed, but is typically between 20 Hz and 100 Hz. There are a couple of factors causing this slower rate: (i) Nearby reflections come from the aircraft's fuselage and wings, which are moving (vibrating) at a much slower rate relative to the motion of the antenna, and (ii) more distant reflections from the sea surface tend to be directional rather than omnidirectional. In the case of maritime mobile terminal–satellite communications, a Gaussian fading spectrum with Doppler bandwidths of 1 Hz and less is a more accurate model of the empirical results. This is due not only to the slower motion of the ship through the water, but also to the effects of ocean waves and their movement as reflective surfaces relative to the motion of the ship.

Problem 2.10 A data signal with a bandwidth of 100 Hz is transmitted at a carrier frequency of 800 MHz. The signal is to be reliably received in vehicles travelling at speeds of up to 100 km/hour. What can we say about the minimum bandwidth of the filter at the receiver input?
Ans. This filter should be at least 175 Hz wide. ■

2.7 CHANNEL CLASSIFICATION

In previous sections, we investigated the effect that various propagation phenomena have on the received electric field. Some of the observations we made were as follows:

- The destructive interference caused by a single reflection can cause a significant reduction in the received power. This interference causes the power loss as a function of distance to change from a square-law loss, characteristic of free-space propagation, to a fourth-power attenuation.

- The presence of large objects such as hills, buildings, and so on does not prevent a signal from being propagated. Rather, diffraction effects allow the signal to propagate around these large objects, albeit at a reduced level.
- A large number of reflections from objects near the receiver can cause large variations in the received electric field that are highly sensitive to position. Again, this is due to the superposition of multiple electric fields associated with the different reflections.
- Movement of the receiver causes a variation in the received electric field, due to changes in constructive and destructive interference of the various reflections. The rate of variation of this electric field depends on the speed of motion of the receiver.

The first two of these observations can be categorized as large-scale propagation effects, the last two as small-scale effects:

1. *Large-scale effects* are due to the general terrain and the density and height of buildings and vegetation. These effects are characterized statistically by the median path loss and lognormal shadowing. Both of these phenomena have a behavior that varies relatively slowly with time. Large-scale effects are important for predicting the coverage and availability of a particular service.
2. *Small-scale effects* are due to the local environment, nearby trees, buildings, etc., and the movement of the radio terminal through that environment. These effects have a much shorter timescale. They have been characterized statistically as fast Rayleigh fading. A consideration of small-scale effects is important for the design of the modulation format and for general transmitter and receiver design.

When we investigated the behavior of these different physical phenomena, the analysis was based on the transmitted electric field, which was modeled at a point in space as

$$E(t) = E_0 \cos(2\pi ft + \theta) \tag{2.78}$$

and by its equivalent complex phasor notation, $\tilde{E}$. Consequently, the analysis describes the response of the channel at a single frequency f. In practice, the transmitted signal has a nonzero bandwidth and can be represented as

$$\begin{aligned} s(t) &= m(t)\cos(2\pi ft + \phi(t)) \\ &= \mathrm{Re}\{\tilde{s}(t)e^{j2\pi ft}\} \end{aligned} \tag{2.79}$$

where $m(t)$ is the amplitude and $\phi(t)$ is the phase. As Eq. (2.79) indicates, the transmitted signal also has a complex phasor representation, $\tilde{s}(t) = m(t)e^{j\phi(t)}$. We shall refer to the complex phasor as the modulating signal. In this section, we consider both the combination of large- and small-scale propagation effects and the effect of nonzero signal bandwidth.

In Section 2.2, we made the assumption that the channel was linear. According to this assumption, all distortions can be characterized by the attenuation or superposition of different signals. In addition, we allow the possibility that the propagation channel

may be time varying. As a consequence of these assumptions, the channel can be presented by a dual-time time-varying impulse response $h(t,\tau)$, defined as the response of the system measured at time t due to a unit impulse applied at time $t - \tau$. Hence, unlike the impulse response of a time-invariant system, the impulse response $h(t,\tau)$ changes shape with τ. For an input $s(t)$, the channel output is given by the convolution integral

$$x(t) = \int_{-\infty}^{\infty} h(t, \tau)s(t-\tau)d\tau \tag{2.80}$$

It can be shown (see Chapter 3) that this expression can also be written in complex phasor notation as

$$\tilde{x}(t) = \int_{-\infty}^{\infty} \tilde{h}(t, \tau)\tilde{s}(t-\tau)d\tau \tag{2.81}$$

Channels are classified on the basis of the properties of the time-varying impulse response $\tilde{h}(t, \tau)$. The effects of noise are not considered in classifying channels.

2.7.1 Time-Selective Channels

We begin by extending our model of a fast-fading channel to the case in which the electric field is modulated by the complex phasor $s(t)$. When the signal is modulated, the received signal is given by

$$\begin{aligned} x(t) &= \sum_{n=1}^{N} \alpha_n(t)m(t)\cos(2\pi ft + \phi(t) + \theta_n(t)) \\ &= \sum_{n=1}^{N} \mathrm{Re}\left\{ \alpha_n(t)e^{j\theta_n(t)}\tilde{s}(t)e^{j2\pi ft} \right\} \end{aligned} \tag{2.82}$$

where $\alpha_n(t)$ is the attenuation and $\theta_n(t)$ is the phase rotation of the nth reflected path. The channel model for fast fading assumes that the different paths of Eq. (2.82) are all of the same length and thus also have the same delay. Equation (2.82) can be written in complex phasor notation as

$$\begin{aligned} \tilde{x}(t) &= \sum_{n=1}^{N} \tilde{\alpha}_n(t)s(t) \\ &= \tilde{\alpha}(t)\tilde{s}(t) \end{aligned} \tag{2.83}$$

where we have made an equivalence between the received signal $\tilde{x}(t)$ and the electric field and where $\tilde{\alpha}_n(t) = \alpha_n(t)e^{j\theta_n(t)}$. That is, the complex gain $\tilde{\alpha}(t)$ (often assumed to have a Rayleigh-distributed magnitude) is the result of the constructive and destructive interference of multiple *local reflections*. It is straightforward to show that $\tilde{\alpha}(t)$ has the same statistical properties as the electric field phasor developed in Section 2.6.4.

Comparing Eq. (2.83) with Eq. (2.81), we can identify the time-varying impulse response. For the model given by Eq. (2.83), it can be verified that the channel impulse response is

$$\tilde{h}(t, \tau) = \tilde{\alpha}(t)\delta(\tau) \tag{2.84}$$

where δ(t) is the *Dirac delta function* or *unit-impulse function*. Since $\tilde{\alpha}(t)$ in Eq. (2.83) is time varying, the received signal strength is also time varying. For this reason, the channel described by Eq. (2.84) is referred to as a *time-selective* channel, in the sense that the channel is better at selected times than at other times.

From Fourier-transform theory (see Appendix A), we know that multiplication in the time domain is equivalent to convolution in the frequency domain. Accordingly, the power spectrum of the received signal of Eq. (2.83) is the convolution of the power spectra of the transmitted signal and the fading process; that is,

$$S_r(f) = S_\alpha(f) \otimes S_s(f) \tag{2.85}$$

where $\otimes$ is the *convolution operator*. Figure 2.21 depicts the result of convolving the power spectrum of the fading process given by Eq. (2.82) with a nominal transmit spectrum, also shown in the figure. The 3-dB bandwidth of the nominal spectrum is approximately five times the maximum Doppler shift f_D. The conclusion to be drawn from the figure is that, when the Doppler shift is a small fraction of the total signal bandwidth, there is very little effect on the received spectrum. In particular, each frequency component is received at approximately the same relative level as that at which it was transmitted. This type of channel is referred to as a *frequency-flat* channel; that is, the frequency response of the channel is approximately constant and does not change the spectrum of the transmitted signal much.

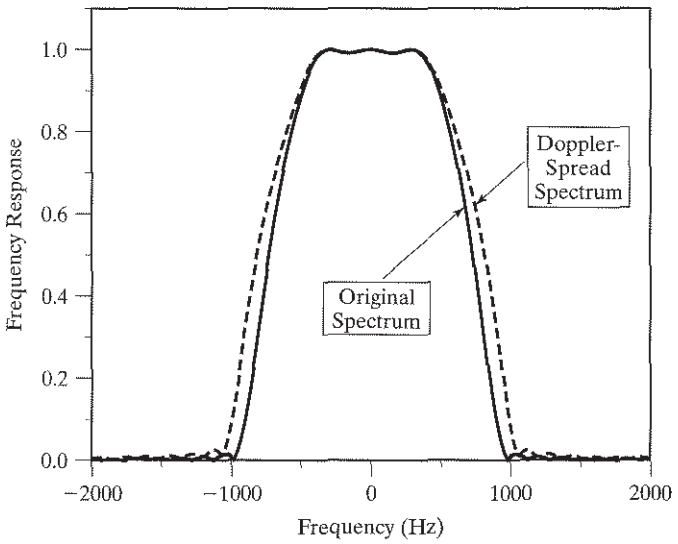


FIGURE 2.21 Comparison of original spectrum with Doppler-spread spectrum. Doppler bandwidth is 20% of signal bandwidth.

As we shall see, a frequency-flat channel is not necessarily time selective; and similarly, a time-selective channel is not necessarily frequency flat. Nevertheless, the combination of a frequency-flat and time-selective channel is common in practice, and we shall refer to it as a *flat-fading* channel.

2.7.2 Frequency-Selective Channels

With large-scale effects, the signal may arrive at the receiver via paths with different path lengths. Consequently, in complex phasor notation, a model for these large-scale effects is

$$\tilde{x}(t) = \sum_{l=1}^{L} \tilde{\alpha}_l \tilde{s}(t - \tau_l) \tag{2.86}$$

where τ_l represents the relative delay associated with the lth path and, for the moment, the complex gains $\{\tilde{\alpha}_l\}$ are assumed constant. (In practice, the complex gains could be constant due to the transmitter and receiver being physically stationary.) With this transmission model, the channel impulse response can be represented as

$$\tilde{h}(t, \tau) = \sum_{l=1}^{L} \tilde{\alpha}_l \delta(\tau - \tau_l) \tag{2.87}$$

This channel is time invariant, but it does show a frequency-dependent response, as illustrated by Example 2.8.

EXAMPLE 2.8 Frequency-Selective Channel

Consider the two-ray channel with the impulse response

$$\tilde{h}(t, \tau) = \delta(\tau) + \tilde{\alpha}_2 \delta(\tau - \tau_2) \tag{2.88}$$

where τ_2 is the relative delay of the second path. If we simulate this channel digitally, with τ_2 equal to the sample period, then the frequency response of the channel for various values of $\tilde{\alpha}_2$ is shown in Fig. 2.22.

For $\tilde{\alpha}_2 = 0.5$, there is attenuation at the edges of the channel. When $\tilde{\alpha}_2 = j/2$, the attenuation is on the left-hand side, and when $\tilde{\alpha}_2 = -j$, there is a null at $+0.25f_s$, where f_s is the sampling rate. (Recall from Section 2.3 that complex values for α imply a phase rotation.) ■

In general, channels with an impulse response characterized by Eq. (2.87) show a frequency dependence. For this reason, these channels are called *frequency selective.* By contrast, the channel of Eq. (2.88) is time invariant, and is thus also referred to as a *time-flat* channel.

2.7.3 General Channels

In the preceding two sections, we have described time-selective and frequency-selective channels. A simpler case occurs when the channel is neither time varying nor frequency varying; this channel is referred to as *flat–flat*, since the response is flat in both the time and frequency domains.

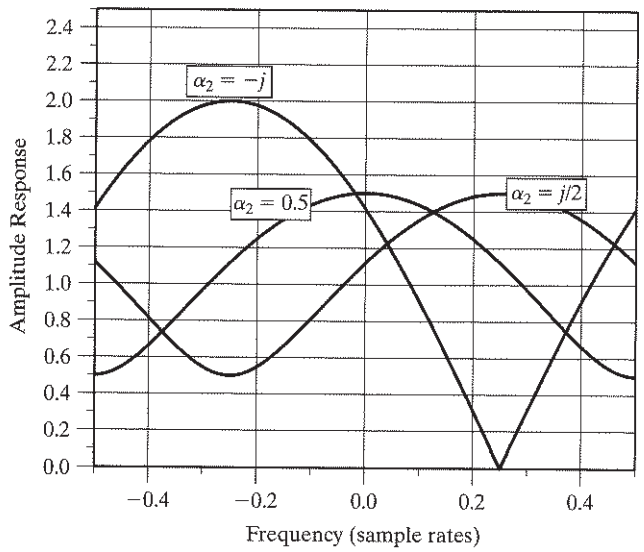


FIGURE 2.22 Frequency response of static two-ray channel $\tilde{h}(t, \tau) = \delta(\tau) + \alpha_2\delta(\tau - T_s)$ where T_s is the sample period.

A more interesting case is characterized by the situation in which the received signal can be represented as

$$\tilde{x}(t) = \sum_{l=1}^{L} \tilde{\alpha}_l(t)\tilde{s}(t - \tau_l(t)) \tag{2.89}$$

That is, the signal arrives at the receiver via paths having different lengths, and each of these paths undergoes time-varying fading. The path lengths can also be slowly time varying. Under these conditions, the channel is characterized by the time-varying impulse response

$$\tilde{h}(t, \tau) = \sum_{l=1}^{L} \tilde{\alpha}_l(t)\delta(\tau - \tau_l(t)) \tag{2.90}$$

The channel model of Eq. (2.89) includes both large- and small-scale effects. The expression for $h(t,\tau)$ given in Eq.(2.90) is for a finite number of signal paths; in the general case, *we may consider $h(t,\tau)$ as representing a continuum of signal paths with arbitrarily small differences.* A channel characterized by Eq. (2.90) is both *time selective* and *frequency selective.*

In the case of either a continuum or a discrete number of signal paths, the received signal is equal to the convolution of the channel impulse response $\tilde{h}(t, \tau)$ with the transmitted signal; that is,

$$\tilde{x}(t) = \int_{-\infty}^{\infty} \tilde{h}(t, \tau)\tilde{s}(t - \tau)d\tau \tag{2.91}$$

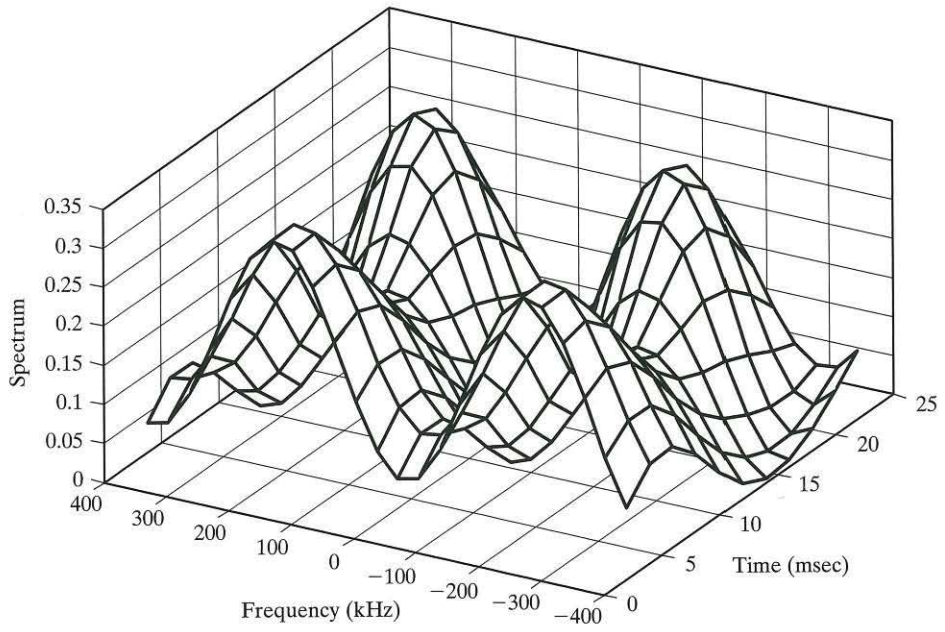


FIGURE 2.23 Time-varying transfer function of a frequency-selective channel.

and related to the time-varying impulse response $h(t,\tau)$, we may define the time-varying frequency response or transfer function

$$H(t,f) = \mathbf{F}[h(t,\tau)] \tag{2.92}$$

where $\mathbf{F}$ is the Fourier transform operator. The quantity $H(t,f)$ represents the time-varying frequency response of the channel, as illustrated in Example 2.9.

EXAMPLE 2.9 Time-Varying Impulse Responses

Consider a channel with the time-varying impulse response

$$\tilde{h}(t,\tau) = \tilde{\alpha}_1(t)\delta(\tau) + \tilde{\alpha}_2(t)\delta(\tau-\tau_2) \tag{2.93}$$

where the $\{\tilde{\alpha}_i(t)\}$ are independent Rayleigh processes. At each time $t=t_0$, we can Fourier transform the impulse response and determine the power spectrum of the result. In Fig. 2.23, we show the progression in time of the transfer function of this fading channel. The channel shows some frequencies with strong responses and other frequencies that correspond to nulls, both of which change with time. ■

2.7.4 WSSUS Channels[13]

Two conclusions may be drawn from the channel models presented in the previous sections:

- The mobile channel introduces delay spread into the received signal. That is, the received signal has a longer duration than that of the transmitted signal, due to the different delays of the signal paths. This phenomenon is referred to as *time dispersion*.
- The mobile channel introduces Doppler spread into the received signal. That is, the received signal has a larger bandwidth than that of the transmitted signal, due to the different Doppler shifts introduced by the components of the multipath. This second phenomenon is referred to as *frequency dispersion*.

Both time dispersion and frequency dispersion introduce distortion into the received signal. The amount of degradation caused by this distortion depends on how the signal is designed.

Doppler spreading or frequency dispersion causes variations of the received signal in the time domain. Of interest for the design of data transmission systems is the maximum duration for which the channel can be assumed to be approximately constant. A transmitted data symbol that has a duration less than this time should suffer little distortion from the effects of frequency dispersion. However, it will still suffer the effects of reduced signal levels.

Just as fading does in the time domain, time dispersion causes slow variations in the received signal in the frequency domain. (See Example 2.8.) Of interest in the design of digital transmissions systems is the maximum transmission bandwidth over which there is little variation. A signal contained within that bandwidth should suffer little distortion from the effects of time dispersion.

For design purposes, two parameters characterize the effects of frequency and time dispersion:

- *Coherence time* refers to the time separation at which the amplitudes of two time-domain samples of the channel become decorrelated.
- *Coherence bandwidth* refers to the frequency separation at which the attenuation of two frequency-domain samples of the channel becomes decorrelated.

In other words, the coherence time is a measure of the length of time for which the channel can be assumed to be approximately constant in the time domain. The coherence bandwidth is a measure of the approximate bandwidth within which the channel can be assumed to be nearly constant. The parameters of coherence time and coherence bandwidth are useful for assessing both the performance of various modulation techniques and the distortion that will be caused by the channel. The threshold for declaring decorrelation in the definitions of coherence time and coherence bandwidth is a matter for debate. Although not definitive, a threshold of 0.5 is used by Jakes (1974) and Lee (1982), and we shall use this definition in the discussion that follows.

The task ahead of us is to develop expressions for estimating the coherence time and coherence bandwidth for a general channel impulse response $\tilde{h}(t, \tau)$. We begin the task by assuming that $\tilde{h}(t, \tau)$ is *wide-sense stationary (WSS)*. As explained in Appendix C, a random process is wide-sense stationary if it has a mean that is time independent and a correlation function $R(t_1, t_2) = R(t_1 - t_2)$. What does this mean for

the time-varying impulse response? It means that the autocorrelation function of the channel response takes the form

$$\mathbf{E}[\tilde{h}(t_1, \tau_1)\tilde{h}^*(t_2, \tau_2)] = R_h(t_1 - t_2; \tau_1, \tau_2) \tag{2.94}$$

That is, the channel autocorrelation function depends only on the time difference, $t_1 - t_2$.

It is also commonly assumed that, in multipath channels, the gain and phase shift at one delay are uncorrelated with the gain and phase shift at another delay. This type of behavior is referred to as *uncorrelated scattering (US)*, a phenomenon in which the $\{\tilde{\alpha}_1(t)\}$ are uncorrelated random processes. Under the uncorrelated-scattering assumption, we find that

$$R_h(t_1 - t_2; \tau_1, \tau_2) = R_h^w(t_1 - t_2; \tau_1)\delta(\tau_1 - \tau_2) \tag{2.95}$$

That is, the autocorrelation is nonzero only when $\tau_1 = \tau_2$. The combination of a wide-sense stationary signal and uncorrelated scattering is referred to as a *wide-sense stationary uncorrelated scattering (WSSUS) channel.*

EXAMPLE 2.10 Uncorrelated Scattering

Suppose that the channel impulse response is given by

$$\tilde{h}(t, \tau) = \sum_{l=1}^{L} \tilde{\alpha}_l(t)\delta(\tau - \xi_1) \tag{2.96}$$

and the scattering processes are uncorrelated. Determine the autocorrelation function of the channel.

The autocorrelation function of the channel is given by

$$\begin{aligned} R_h(t_1, t_2; \tau_1, \tau_2) &= \mathbf{E}[h(t_1, \tau_1)h^*(t_2, \tau_2)] \\ &= \mathbf{E}\left[\sum_{l=1}^{L} \tilde{\alpha}_l(t_1)\delta(\tau_1 - \xi_l) \sum_{k=1}^{L} \tilde{\alpha}_k^*(t_2)\delta(\tau_2 - \xi_k)\right] \\ &= \sum_{l=1}^{L}\sum_{k=1}^{L} \mathbf{E}[\tilde{\alpha}_l(t_1)\tilde{\alpha}_k^*(t_2)]\delta(\tau_1 - \xi_l)\delta(\tau_2 - \xi_k) \\ &= \sum_{l=1}^{L} \mathbf{E}[\tilde{\alpha}_l(t_1)\tilde{\alpha}_l^*(t_2)]\delta(\tau_1 - \xi_l)\delta(\tau_2 - \xi_l) \end{aligned} \tag{2.97}$$

where the last equality follows from the third by the uncorrelated-scattering assumption. Then, with a Clarke model for the scattering process described in Eq. (2.75), the autocorrelation of the fading for the lth path is given by

$$\mathbf{E}[\tilde{\alpha}_l(t_1)\tilde{\alpha}_l^*(t_2)] = P_l J_0(2\pi f_D(t_1 - t_2)) \tag{2.98}$$

where P_l is the average power of the lth path. Consequently, the autocorrelation function of the channel reduces to

$$R_h(t_1, t_2;\tau_1, \tau_2) = \left[\sum_{l=1}^{L} P_l J_0(2\pi f_D(t_1 - t_2))\delta(\tau_1 - \xi_l)\right]\delta(\tau_1 - \tau_2) \tag{2.99}$$

Equation (2.99) has the form indicated by Eq. (2.95). ■

2.7.5 Coherence Time

In Section 2.7.4, the coherence time is defined as the period over which there is a strong correlation of the channel time response. For a WSSUS process, this is the range of Δt over which $R_h^w(\Delta t;\tau_1) \equiv R_h(t, t + \Delta t;\tau_1, \tau_2)$ of Eq. (2.95) is significant. If we set $\tau_1 = 0$ in $R_h^w(\Delta t;\tau_1)$, then we find that the Fourier transform of the resulting autocorrelation function is given by

$$S_D(f) = \mathbf{F}[R_h^w(\Delta t; 0)] \tag{2.100}$$

The function $S_D(f)$ is referred to as the *Doppler power spectrum* of the channel. A general property of the Fourier transform is that the width of a signal in the time domain is inversely proportional to the width of the signal in the frequency domain; see Appendix A. Consequently, if the Doppler spread $2f_D$, where f_D is the maximum Doppler frequency, is the approximate width of the Doppler power spectrum, then the coherence time is given approximately by

$$T_{\text{coherence}} \approx \frac{1}{2f_D} \tag{2.101}$$

That is, the coherence time of the channel is approximately the inverse of the Doppler spread of the channel.

EXAMPLE 2.11 Coherence Time for a Flat-Fading Channel

For the Clarke model for fast fading described in Section 2.6.4, what is the coherence time of the channel?

From Eq. (2.75), the normalized autocorrelation of the fading process of a flat-fading channel described by the Clarke model is given by

$$R_h^w(\Delta t, 0) = R_h(\Delta t) = J_0(2\pi f_D \Delta t) \tag{2.102}$$

If we define the coherence time as that range of values, Δt, over which the correlation is greater than 0.5, then substituting $R_h^w(T_{\text{coherence}};0) = 0.5$ into the left-hand side of Eq. (2.102) implies that the coherence time is given by

$$\begin{aligned} T_{\text{coherence}} &= \frac{1}{2\pi f_D} J_0^{-1}(0.5) \\ &= \frac{0.30}{2f_D} \end{aligned}$$

where $J_0^{-1}(.)$ is the inverse Bessel function of order zero. As predicted by the preceding analysis, the coherence time is inversely related to the Doppler spread. ■

The two inversely related parameters of coherence time and Doppler spread provide one measure of the channel characteristics. If the coherence time is finite, then the channel is *time selective*, that is, *time varying*. If the coherence time is infinite, then the channel is *time flat*, that is, *time invariant*.

2.7.6 Power-Delay Profile

With a WSSUS channel, we found that the autocorrelation of the channel impulse response is given by Eq. (2.95); that is, the wide-sense stationary assumption implies that there is a dependence only on Δt and not on the pair (t_1,t_2), while the uncorrelated scattering assumption implies that τ_1 and τ_2 must align for nonzero correlation. We define the function

$$P_h(\tau) = R_h^w(0;\tau) \tag{2.103}$$

based on $R_h^w(\Delta t;\tau)$ with $\Delta t = 0$. This can be shown to be equivalent to $\mathbf{E}[|h(t, \tau)|^2]$ for the WSSUS channel. The function $P_h(\tau)$, which is known as the *multipath intensity profile* or *power-delay profile* of the channel, provides an estimate of the average multipath power as a function of the relative delay τ.

EXAMPLE 2.12 Discrete Power-Delay Profile

Consider the situation in which the channel impulse response is given by

$$\tilde{h}(t, \tau) = \sum_{l=1}^{L} \tilde{\alpha}_l(t)\delta(\tau - \xi_l) \tag{2.104}$$

Then, from Example 2.10, the corresponding autocorrelation function of $\tilde{h}(t, \tau)$ is given by Eq. (2.99); that is,

$$R_h(t_1, t_2;\xi_1, \xi_2) = \left[\sum_{l=1}^{L} P_l J_0(2\pi f_D \Delta t)\delta(\tau_1 - \xi_l)\right]\delta(\tau_1 - \tau_2) \tag{2.105}$$

and the power-delay profile is given by

$$P_h(\tau) = \sum_{l=1}^{L} P_l\delta(\tau - \xi_l) \tag{2.106}$$

That is, at the delay ξ_l, the power-delay profile is the average power of the lth path, as would be expected from the definition. ■

The different rays that compose a multipath signal will travel different paths between the transmitter and the receiver. These paths will often have different lengths, and there will be correspondingly different transmission delays. Intuitively, we would

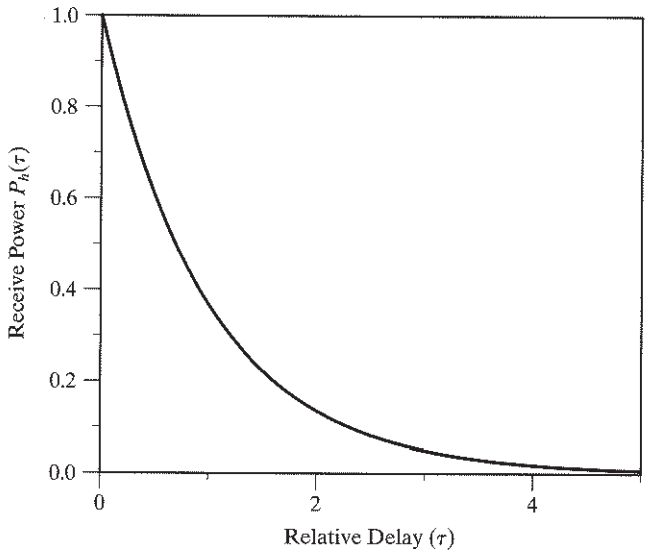


FIGURE 2.24 Sample average power-delay profile.

expect signals at greater delays to be weaker simply because they have traveled greater distances. This is often the case, but not always: the actual propagation losses depend on the reflection coefficients and diffraction effects. The multipath intensity is sometimes modeled with an exponential power-delay profile, as shown in Fig. 2.24. That is, the average power decreases exponentially with relative delay. Sample channels may vary significantly, but the power-delay profile refers to the power as a function of delay, averaged over a large number of channels.

The power-delay profile may be characterized by two statistical parameters. With $P_h(\tau)$ representing the profile, the *average delay* is given by

$$T_D = \frac{1}{P_m}\int_0^\infty \tau P_h(\tau)d\tau \tag{2.107}$$

where the *average power* is

$$P_m = \int_0^\infty P_h(\tau)d\tau \tag{2.108}$$

The second central moment of the average power-delay profile is given by

$$\begin{aligned}\mu_2 &= \frac{1}{P_m}\int_0^\infty (\tau - T_D)^2 P_h(\tau)d\tau \\ &= \frac{1}{P_m}\int_0^\infty \tau^2 P_h(\tau)d\tau - T_D^2\end{aligned} \tag{2.109}$$

There are two parameters in common use to characterize this time dispersion or excess delay. The first is *delay spread,* defined by

$$S = \sqrt{\mu_2} \tag{2.110}$$

which is simply the root-mean-square (rms) delay. The second parameter is the *multipath spread,* given by

$$T_M = 2\sqrt{\mu_2} = 2S \tag{2.111}$$

The multipath spread is the two-sided rms delay, or the approximate width of the excess delay. The multipath spread is effectively the time interval over which the power-delay profile is nonzero.

2.7.7 Coherence Bandwidth

Just as the coherence time of a channel is related to the autocorrelation function of the channel impulse response, the coherence bandwidth of a channel is related to the autocorrelation function of the time-varying frequency response. Recall from Section 2.7.3 that the time-varying frequency response of the channel is given by the Fourier transform

$$H(t,f) = \mathbf{F}[\tilde{h}(t,\tau)] \tag{2.112}$$

With this definition of the time-varying frequency response, we may define the autocorrelation function

$$R_H(f_1,f_2;\Delta t) = \mathbf{E}[H(t,f_1)H^*(t+\Delta t,f_2)] \tag{2.113}$$

This is the *frequency-spaced, time-spaced correlation function*; it depends only on the time difference Δt due to the assumed WSS properties of the channel. It can be shown that the frequency-spaced, time-spaced correlation also depends only on Δf for a WSSUS channel, not on the individual f_1 and f_2, so we can define

$$P_H(\Delta f,\Delta t) = R_H(f,f+\Delta f;\Delta t) \tag{2.114}$$

It can also be shown that there is a Fourier transform relationship between the two quantities $[P_h(\tau)]$ and $(P_H(\Delta f,0))$, as shown by

$$P_H(\Delta f,0) = \mathbf{F}[P_h(\tau)] \tag{2.115}$$

The function on the left is the *coherence spectrum* of the channel. The function on the right is the Fourier transform of the power-delay profile. The coherence spectrum represents the average correlation between frequencies separated by Δf at any particular instant in time ($\Delta t = 0$). Consequently, the coherence spectrum is closely related to

the coherence bandwidth of the signal. In particular, because of the inverse relationship between the time and frequency domains, we have the relationship

$$\mathrm{BW}_{\mathrm{coh}} \approx \frac{1}{T_M} \tag{2.116}$$

That is, *the coherence bandwidth is inversely proportional to the multipath spread* of the channel.

The coherence bandwidth of the channel is the bandwidth over which the frequency response is strongly correlated—that is, relatively flat. If the coherence bandwidth is small with respect to the bandwidth of the transmitted signal, then the channel is said to be *frequency selective.* That is, insofar as the transmitted signal is concerned, the channel has a nonflat frequency response. When the coherence bandwidth is large with respect to the bandwidth of the transmitted signal, the channel is said to be *frequency nonselective* or *frequency-flat.*

2.7.8 Stationary and Nonstationary Channels

The key feature of wide-sense stationary uncorrelated-scattering (WSSUS) channels is that correlation of the channel response depends only on the time difference and not on the absolute time. These stationary models for channel characteristics are convenient for analysis, but often, except for short time intervals, they are not an accurate description of reality. For example, terrestrial mobile channels are usually highly nonstationary, for the following reasons, among others:

1. The propagation path often consists of several discontinuities, such as buildings, that can cause significant changes in the propagation characteristics (e.g., when a terminal moves in and out of the electromagnetic shadow of a building).
2. The environment itself is physically nonstationary. There may be moving trucks, moving people, or other elements of the environment that can significantly affect propagation. An example is the effect of a person's head on the propagation path for a cell phone user as the user nods or changes position.
3. The interference caused by other users sharing the same frequency channel will vary dynamically as these other users come onto and leave the system.

All of these factors contribute to the nonstationarity of the link. Different types of radio links have different degrees of nonstationarity; nevertheless, some wireless links, such as fixed satellite links, can often be considered essentially stationary.

Stationary models still have their application. Even though many communications links are highly nonstationary, the WSSUS model provides a reasonably accurate account of the propagation characteristics for short periods of time. For many applications, this is sufficient. For example, with digitized voice transmissions, the size of a voice frame is often on the order of 20 milliseconds. Since communications link can frequently be considered stationary over many voice frames, the WSSUS model is an important tool for analyzing system performance.

2.7.9 Summary of Channel Classification

This section introduced several classifications of wireless channels:

- *Time-flat channels* are time-invariant channels. An example is a transmitter and receiver that are both physically stationary, with the propagation environment unchanging.
- *Frequency-flat channels* have a frequency response that is approximately flat over a bandwidth greater than or equal to the bandwidth of the transmitted signal.
- *Time-selective channels* are time-varying channels. An example is a wireless terminal moving through the environment and undergoing Rayleigh fading.
- *Frequency-selective channels* have a frequency response that cannot be assumed to be flat over the bandwidth of the signal. Frequency selectivity is due to a multipath that has significant delay spread relative to the symbol period of the transmission.

The classification of a channel depends on the signal that is to be transmitted. Narrowband mobile channels are often flat-fading (flat frequency, time selective), because the coherence bandwidths of most radio channels at transmission frequencies less than 1 gigahertz are typically tens of kilohertz. The exception to this is HF radio, wherein due to propagation modes such as ducting through the ionosphere, even very narrow channels can be frequency selective. Wideband channels, in which either the transmitter or the receiver is moving, are often both time selective and frequency selective.

For data transmission systems, there are two parameters of interest:

- *Coherence time* is the maximum duration for which the channel can be assumed to be approximately constant. If a transmitted data symbol has a duration that is less than the coherence time, it should suffer little distortion from the effects of frequency dispersion.
- *Coherence bandwidth* is the maximum transmission bandwidth over which there is little variation. A signal contained within the coherence bandwidth should suffer little distortion from the effects of time dispersion.

If we assume that the channel is wide-sense stationary with uncorrelated scattering (WSSUS), then the coherence time is inversely proportional to the Doppler spread of the channel and the coherence bandwidth is inversely proportional to the delay spread of the channel.

Problem 2.11 Measurements of a radio channel in the 800 MHz frequency band indicate that the coherence bandwidth is approximately 100 kHz. What is the maximum symbol rate that can be transmitted over this channel that will suffer minimal intersymbol interference?
Ans. The multipath spread of the channel is approximately 10 microseconds. If we assume that spreading of the symbol by less than 10% causes negligible interference, then maximum symbol period is 100 microseconds and the corresponding symbol rate is 10 kHz. ■

Problem 2.12 Calculate the *rms* delay spread for an HF radio channel for which

$$P(\tau) = 0.6\delta(\tau) + 0.3\delta(\tau - 0.2) + 0.1\delta(\tau - 0.4)$$

where τ is measured in milliseconds. Assume that signaling with a 5-kHz bandwidth is to use the channel. Will delay spread be a problem? That is, is it likely that some form of compensation (an equalizer) will be necessary?

Ans. The rms delay spread is 0.29 millisecond and the coherence bandwidth is 1720 Hz, so an equalizer is likely necessary. ■

Problem 2.13 Show that the time-varying impulse response of Eq.(2.84) and a time-invariant impulse response are related by

$$\tilde{h}_{\text{time-invariant}}(t) = \tilde{h}_{\text{time-varying}}(t, t)$$

Explain, in words, what the preceding equation means. ■

2.8 NOISE AND INTERFERENCE

The primary objective of the previous sections is to estimate the received power level and characterize its behavior as a function of the environment. *What sets the minimum acceptable received power level is the receiver noise and interference levels.* In general, noise can be defined as unwanted (and usually uncontrollable) electrical signals interfering with the desired signal. Unwanted signals arise from a variety of sources, both natural and artificial. Artificial sources include noise from automobile ignition circuits, commutator sparking in electric motors, 60-cycle hum, and signals from other communications systems. Numerous artificial sources result from harmonics of the natural frequency. For example, spark plug firings in car engines have a frequency on the order of thousands of rpm; consequently, although the fundamental frequency is less than 1 kHz, the energy emitted by this excitation is often so strong that high-order harmonics can cause significant interference in radio systems. Natural sources of noise include circuit noise, atmospheric disturbances, and extraterrestrial radiation. In this section, we will characterize various sources of noise. Later sections will show how they influence communication system design.

2.8.1 Thermal Noise[14]

Thermal noise can be considered a fundamental property of matter above the absolute temperature of 0°K. In an insulator, electrons are bound to the atom. In a conductor, such as a wire or a resistor, the attraction between the electrons and the nucleus is not as strong, and electrons are free to move from one atom to another. In any conductor at a temperature above 0°K, these free electrons are never really stationary, but move about with random velocities in all directions. These random motions have an average velocity that is zero in any direction over a long period of time. Over short intervals, there are statistical fluctuations from this zero average—equivalent to intermittent currents. We call these spontaneous fluctuations *thermal noise.*

Consider the voltage that is observed between the two terminals of a metallic resistor. The average voltage is zero. The mean-square voltage is nonzero, due to thermal noise. Quantum physics provides an expression for this mean-square voltage namely,

$$\bar{v}^2 = \frac{2\pi^2 k^2 T^2}{3h} R \tag{2.117}$$

where

R is the resistance in ohms
$k = 1.37 \times 10^{-23}$ W-s/°K is Boltzman's constant
$h = 6.62 \times 10^{-34}$ W-s^2 is Planck's constant
T is the absolute temperature of the resistor in degrees Kelvin

Consequently, the resistor can be modeled as equivalent to the Thévenin circuit shown in Fig. 2.25, in which the resistor is ideal (noiseless) and the voltage source has zero mean and a mean-square voltage given by Eq. (2.117). Since the voltage is the sum of the motions of a large number of electrons, it is accurately modeled as a *Gaussian distribution* with zero mean. The mean-square (m.s.) voltage across the resistor terminals is finite and depends on the resistance R. The maximum power delivered by the circuit of Fig. 2.25 occurs when the load resistance is matched to R. By dividing both sides of Eq. (2.117) by R and scaling by a factor of 1/4, we obtain an expression for the maximum available noise power.

Any communications link occupies only a very narrow portion of the whole electromagnetic spectrum. What is of interest to communications engineers is the distribution of thermal noise power across frequency, or $S_n(f)$—that is, the *noise spectral density*. Quantum physics provides an expression for this as well. The important result is that, *for frequencies up to 10^{12} Hz, the available thermal noise spectral density (watts/Hz) is approximately constant* and is given by

$$\begin{aligned} S_n(f) &= kT/2 \\ &\equiv N_0/2, \qquad -\infty < f < \infty \end{aligned} \tag{2.118}$$

Equation (2.118) describes the noise power delivered to a matched load; it is independent of the resistance. This upper limit of 10^{12} Hz falls in the near-infrared portion of the electromagnetic spectrum, far above where conventional electrical components have ceased to respond. Above these frequencies, the noise power decreases exponentially as a function of frequency.

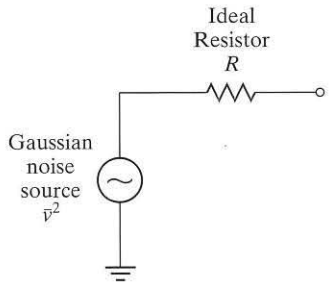


FIGURE 2.25 Thévenin model of a resistor.

On the basis of Eq. (2.118), we use the *two-sided spectral model* for thermal noise depicted in Fig. 2.26. This model shows that thermal noise has a flat spectrum across all frequencies (positive as well as negative), with density $N_0/2$. This type of noise is referred to as *white noise*, since it contains all frequencies at an equal level. The amplitude distribution at any frequency is Gaussian, and there is no correlation between samples of the equivalent time-domain noise process $n(t)$. That is, for a *real* white-noise process, the autocorrelation function is given by

$$\begin{aligned} R_n(\tau) &= E[n(t)n(t+\tau)] \\ &= \frac{N_0}{2}\delta(\tau) \end{aligned} \tag{2.119}$$

The noise spectral density is conventionally defined as a two-sided quantity $N_0/2$ applicable over the range $-\infty < f < \infty$. In practice, the single-sided value N_0 is often used, for two reasons:

- Conventional instrumentation measures only positive frequencies.
- In Chapter 3, it will be shown that noise has a complex envelope representation similar to that of any other signal. In the case of noise, both the in-phase and quadrature components are white, Gaussian, and independent of each other. The corresponding complex baseband process $\tilde{n}(t)$ has a two-sided noise spectral density of N_0, and autocorrelation is $R_{\tilde{n}}(\tau) = N_0\delta(\tau)$.

The noise power in a bandwidth B is

$$\begin{aligned} P &= (N_0/2)(2B) \\ &= N_0 B \end{aligned} \tag{2.120}$$

where the first equality corresponds to the two-sided representation and the second to the single-sided representation. If a filter has a frequency response $H(f)$, then we define the *noise-equivalent bandwidth*, or *noise bandwidth*, of the filter as

$$B_{\text{eq}} = \frac{\int_{-\infty}^{\infty} |H(f)|^2 (N_0/2)\,df}{N_0/2} \tag{2.121}$$

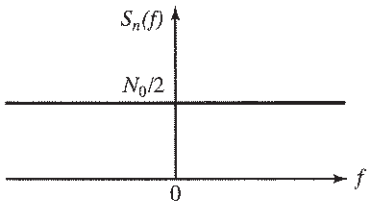


FIGURE 2.26 Spectral model for white noise.

The noise bandwidth is equivalent to the bandwidth of an ideal rectangular filter that would pass the same amount of noise power. For many filters, the noise bandwidth is approximately equal to their *3-dB bandwidth*. We may define the noise-equivalent bandwidth of a signal in a similar manner.

Problem 2.14 What would be the rms voltage observed across a 10-MΩ metallic resistor at room temperature? Suppose the measuring apparatus has a bandwidth of 1GHz, with an input impedance of 10 MΩ. What voltage would be measured then? Compare your answer with the voltage generated across the antenna terminals by the signal defined in Problem 2.5. Why is the avoidance of large resistors recommended for circuit design? What is the maximum power density (W/Hz) that a thermal resistor therefore delivers to a load?
Ans. The maximum rms voltage is 1.6 volts. Over 1 GHz into a 10-MΩ load, the voltage is approximately 6 millivolts. The maximum power density delivered to a matched load is $kT = 4 \times 10^{-21}$ *watt/Hz at room temperature.* ■

2.8.2 Equivalent Noise Temperature and Noise Figure

There are other sources of noise that exhibit behavior similar to that of thermal noise. That is, the noise produced by these sources is Gaussian in distribution and white over the range of frequencies of interest. For this reason, the concept of an *equivalent noise temperature* is often assigned to these sources. The equivalent noise temperature of a device is not necessarily its physical temperature, but rather a measure of the noise that the device produces.

EXAMPLE 2.13 Earth's Radiation

The antenna of a satellite looks directly at the Earth in order to receive signals emitted by ground-based terminals. The Earth has an average temperature of approximately $T_0 = 290°K$ and, like any warm body, radiates energy that is proportional to this temperature. The radiated energy is effectively white across the frequency band. Any of this energy that falls within the antenna aperture is collected and added to the received signal, in which case we say the antenna has an *equivalent noise temperature* of T_0. ■

Amplifiers add noise to the signal in much the same way that passive devices such as resistors do. The noise added depends on the amplifier design and must be calibrated for each device. Although amplifiers could be specified in terms of their equivalent noise temperature, they are conventionally specified in terms of their *noise figure*. One reason for this alternative representation is that amplifiers usually have a finite operational bandwidth; consequently, the noise effects may be frequency dependent. For applications considered in this book, we will assume that noise effects are white.

The *noise figure* represents the increase in noise at the output of the amplifier, referenced to the input. Let $G(f)$ be the available power gain of the device as a function of frequency. G is defined as the ratio of the available signal power at the output of the device to the available signal power of the source when the source is a sinusoid

of frequency f. The noise figure is then defined as

$$F = \frac{S_{no}(f)}{G(f)S_{ni}(f)} \tag{2.122}$$

where $S_{no}(f)$ is the spectrum of the output noise power and $S_{ni}(f)$ is the spectrum of the input noise power. For an ideal noiseless amplifier, the noise figure is unity; that is, the amplifier simply amplifies the input noise, but adds no noise itself. Equation (2.122) permits the calculation of the noise figure dependence on the frequency f if necessary.

The noise figure is usually expressed in decibels, but for very low-noise amplifiers such as those used as receiver front ends in satellite communications, the noise figure is often very close to 0 dB. The decibel scale makes it difficult to distinguish between the performance of the different amplifiers in this case. In these situations, it is often preferable to use the equivalent noise temperature of the amplifier. That is, equivalent noise temperature and noise figure are equivalent methods of specifying the noise contribution of a device.

To show the relationship between noise figure and equivalent noise temperature, assume that the amplifier has a power gain G that is constant; this is referred to as a *flat spectral response*. A metallic resistor is connected across the input terminals. The available noise power at the output due to the input noise is GkT_0. The total output noise power is defined as $G(kT_e + kT_0)$, where T_e is the equivalent noise temperature of the device. This noise model is shown in Fig. 2.27. Consequently, the noise figure can be written as

$$F = \frac{G(kT_e + kT_0)}{GkT_0} = \frac{T_e + T_0}{T_0} \tag{2.123}$$

where T_0 is the standard temperature of 290°K. The relationship between the equivalent noise temperature T_e and the noise figure is given by

$$T_e = (F-1)T_0 \tag{2.124}$$

EXAMPLE 2.14 Noise Figure and Receiver Sensitivity

Suppose a wireless receiver has a noise figure of 8 dB and includes a modem that requires an SNR of 12 dB for proper operation in a 5-kHz bandwidth. What is the receiver sensitivity?

The noise power due to the receiver in a bandwidth B is given by

$$N = F(kT_0)B$$

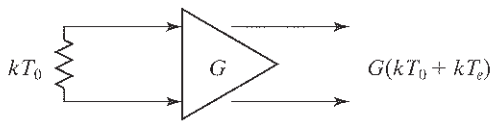


FIGURE 2.27 Noise model of amplifier.

where F is the noise figure of the receiver. In the decibel representation, this equation becomes

$$\begin{aligned} N(\text{dBm}) &= k(\text{dBm} - \text{sK}^{-1}) + T_0(\text{dBK}) + F(\text{dB}) + B(\text{dBHz}) \\ &= -198.6 + 10\log_{10}290 + 8 + 10\log_{10}5000 \\ &= -129.0 \text{ dBm} \end{aligned}$$

where we have used the following equivalences for Boltzmann's constant:

$$\begin{aligned} k &= 1.37\times10^{-23}\text{Watt} - \text{sec}^\circ\text{K} \\ &= -228.6\text{dBW} - \text{s/K} \\ &= -198.6\text{dBm} - \text{s/K} \end{aligned}$$

The corresponding receiver sensitivity is given by

$$S = \text{SNR} \times N$$

So, in a decibel representation, the sensitivity (i.e., the minimum signal level that the receiver can reliably detect) is

$$\begin{aligned} S(\text{dB}) &= \text{SNR(dB)} + N(\text{dBm}) \\ &= 12 + (-129) \\ &= -117 \text{ dBm} \end{aligned}$$

■

Problem 2.15 The noise figure of a cell phone receiver is specified as 16 dB. What is the equivalent noise temperature? Assume that reliable detection of a 30-kHz FM signal by this receiver requires an SNR of 13 dB. What is the receiver sensitivity in dBm?
Ans. The equivalent noise temperature is 11,255°K and the receiver sensitivity is –100 dBm. ■

2.8.3 Noise in Cascaded Systems

A communications receiver is composed of a number of components, including an antenna, amplifiers, filters, and mixers, and these components may be connected by transmission lines. The overall *system temperature* involves noise contributions from these components in a weighted manner. In this section, we show how to combine the noise contributions of these different portions of the system.

We model each of the aforementioned devices with the two-port model shown in Fig. 2.28, where each device has a noise factor F and a gain (or loss) G. With this model, the input to the device is summed with an internal noise source. The spectral density of the internal noise source is $(F-1)kT_0$. The combined signal is then processed

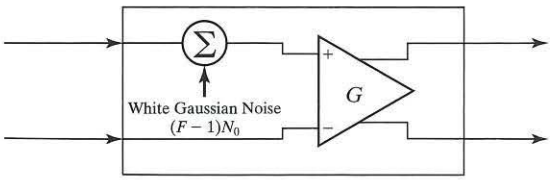


FIGURE 2.28 Two-port model for a communications system component with $N_0 = kT_0$.

by an ideal (noiseless) amplifier with gain G. If $S(f)$ is the device input, then the output of the circuit is given by

$$Y(f) = GS(f) + G(F-1)kT_0 \tag{2.125}$$

That is, the output has both a signal component and a noise component.

Now consider a system with two such devices, having noise factors F_1 and F_2 and gains G_1 and G_2, respectively, as shown in Fig. 2.29. To determine the noise figure of the cascaded system, we assume that the input is connected to a metallic resistor with noise density N_0.

Under these assumptions, the output of the first stage and input to the second stage is $F_1G_1N_0$. The output of the second stage is $G_2\{(F_2\text{-}1)N_0 + F_1G_1N_0\}$. In accordance with the definition of noise figure given in Eq. (2.122), if we compare the input and output of the cascaded devices, the overall noise figure is given by

$$\begin{aligned} F &= \frac{G_2\{(F_2-1)N_0 + F_1G_1N_0\}}{G_1G_2N_0} \\ &= F_1 + \frac{F_2-1}{G_1} \end{aligned} \tag{2.126}$$

That is, *the combined noise figure is the noise figure of the first device, plus the noise figure of the second device attenuated by the gain of the first device.* In the general case, the noise figure for a multistage system is given by

$$F = F_1 + \frac{F_2-1}{G_1} + \frac{F_3-1}{G_1G_2} + \ldots, \tag{2.127}$$

Correspondingly, the equivalent noise temperature of the system is

$$T_{\text{sys}} = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2} + \ldots. \tag{2.128}$$

where T_k and G_k are the noise temperature and gain of the kth stage of the receiver amplification chain, with $k = 1,2,\ldots$. If we include the equivalent noise temperature of the antenna in Eq. (2.128), the equation becomes

$$T_{\text{sys}} = T_A + T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2} + \ldots. \tag{2.129}$$

where T_A is the equivalent noise temperature of the antenna.

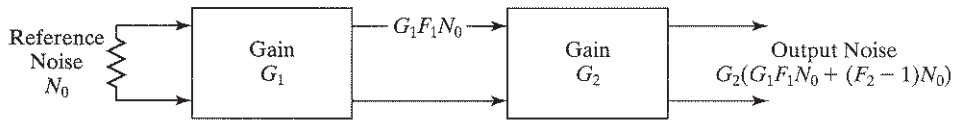


FIGURE 2.29 Illustration of cascaded noise calculation for two-device system.

EXAMPLE 2.15 System Noise Figure Calculation

Suppose the antenna of Example 2.13 is connected to an amplifier stage with a gain of 40 dB and an equivalent noise temperature of 50°K. A subsequent amplifier stage has a gain of 30 dB and noise figure of 5. What is the system noise figure?

From Example 2.13, the antenna has an equivalent noise temperature of T_0 or, equivalently, a noise figure of 2 and a gain of 1. For the first amplifier, the noise figure is

$$F_1 = 1 + T_e/T_0 = 1.17$$

The system noise figure is then

$$F_{\text{sys}} = F_{\text{ant}} + (F_1 - 1) + \frac{(F_2 - 1)}{G_1}$$

$$= 2 + (1.17 - 1) + \frac{5}{10^4} = 2.17$$

From this example, it should be clear that it is the noise performance of the first amplifier stage, often referred to as the *low-noise amplifier* (LNA), that is critical to communications performance. ■

Problem 2.16 Show that the system temperature, in general, is given by Eq. (2.129). ■

2.8.4 Man-Made Noise[15]

Natural background and internal noise sources may be minimized with good system and antenna design. Beyond this, artificial noise levels may set the sensitivity limit of a receiver. Some artificial sources of noise may be other communications systems operating in the same frequency band; other sources may be systems that operate in other frequency bands, but unintentionally generate RF signals in the band of interest. In this section, we will consider three such sources.

Devices that produce large electrical discharges often cause *impulse noise*. The electromagnetic pulses may have durations of 5 to 10 nanoseconds and may occur at rates up to several hundred pulses per second. Due to the short pulse duration, impulse noise is often modeled as a broadband type of noise. The impulses vary in their strength and frequency. Some impulses have power levels well above the receiver noise floor, but may occur infrequently. Impulses of weaker strength occur more frequently. The spectral density of the impulse power decreases with frequency, being greater in frequency bands below 500 MHz and decreasing rapidly at high frequencies.

The sources of impulse noise include the following:

- Noise from electrical machinery (particularly commutating motors)
- Noise from spark ignition systems in automobile or other internal combustion engines
- Switching transients
- Discharge lighting.

A general characterization of impulse noise is difficult to come by, but impulse noise is a nonstationary phenomenon that depends on the frequency of operation of the

source, the time of day, the surrounding environment, and various other factors. Accurate modeling requires measurements in the frequency band of interest, in differing environments, and at different times of day, week, and year.

The impact of impulse noise on digital communications systems depends on the impulse characteristics, such as pulse power, duration, and repetition rate. The impulse noise can affect both the front-end and intermediate stages of a receiver and can cause variations in signal strength, synchronization loss, and message errors.

A second type of man-made noise or interference is *out-of-band transmissions* from other communicationslike services. These out-of-band transmissions include the harmonics of high-power radars and television signals, services that may be using a highly separate frequency band, but that have a second or third harmonic which may fall directly into the communications band of interest. Of particular concern are examples where one service transmits at much higher levels than another. For example, the Global Positioning Satellite Service (GPSS) operates at very low power levels in the frequency band centered on 1.476 GHz. The band at 738 MHz is assigned for wideband public safety messages, with base station transmit powers of up to 30W permitted. Clearly, the second harmonic of the latter signal must be sufficiently attenuated to prevent interference with the former. (See Problem 2.44.)

A third type of interference related to system design, *multiple access interference*, is discussed in the next section.

Problem 2.17 Microwave ovens operate at a natural frequency of the water molecule at approximately 2.45 GHz. This frequency falls in the middle of a band from 2.41 to 2.48 GHz that has been allocated for low-power unlicensed radio use, including wireless local area networks (sometimes referred to as Wi-Fi—see Section 5.16). The oscillators used in some microwave ovens have poor stability and have been observed to vary ±10 MHz around their nominal frequency. Discuss how such variability would affect the signal design for the use of this band. ■

2.8.5 Multiple-Access Interference

Due to a limited frequency spectrum, communication frequencies are reused the world over. Often, the rules for frequency reuse are subject to international agreements to minimize and control the effects of cross-border interference. An example is the assignment of television broadcast frequencies along international borders. In other cases, international agreements are in place for the use of common frequencies. For example, the VHF radio band, 118 MHz to 130 MHz, is used for aeronautical safety communication the world over. The channel frequencies and permitted modulation types within that band are part of the agreement.

For wireless communications, frequencies are often assigned by federal authorities on a regional basis. These same frequencies can then be assigned to another user or service provider in a different region, sufficiently distant that there is no interference. To maximize frequency reuse, reuse distances are often made as small as possible. As a consequence, a service is not interference free, but rather must have a certain interference tolerance. A typical example is an FM receiver that must tolerate a signal-to-interference ratio (SIR) of 20 dB at its edge of coverage, such as that illustrated in Fig. 2.30.

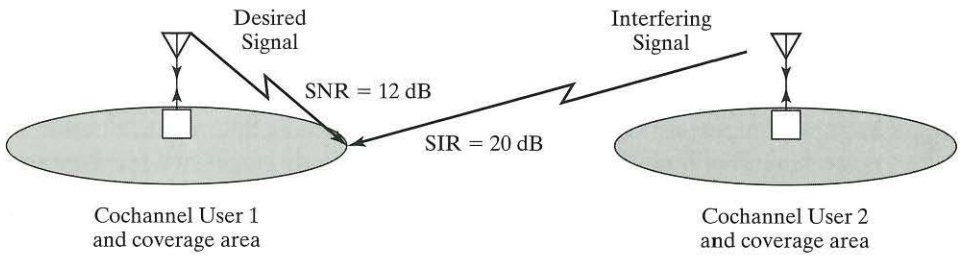


FIGURE 2.30 Example of traditional frequency reuse.

Frequency reuse policies have made significant advances in recent years with the introduction of *cellular systems*, wherein a single service provider is given a band of frequencies for a given service area. The service provider tries to reuse the frequencies as many times as possible within the given service area in order to meet customer demands. In a cellular system, the cells are often modeled with a hexagon pattern, as shown in Fig. 2.31, with users in each cell communicating with a base station located at the center of the cell. Cellular telephones connect the mobile user wirelessly to the base station, which acts as an interface to the wired telephone network; or it could potentially connect the user to another mobile user. The term *cellular* is usually applied to terrestrial systems, but similar considerations apply to multibeam satellites. For nongeostationary satellite systems, there is the added issue that the cells move with respect to the earth as the satellite moves.

The hexagonal cell distribution is an approximation of practical cellular patterns that is useful for illustration because of its ease of analysis. With a hexagon geometry,

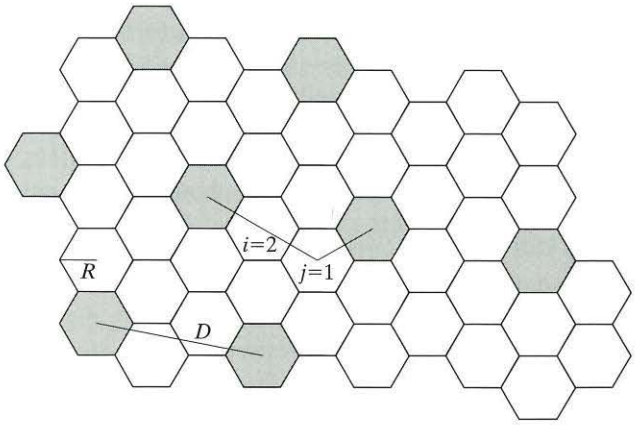


FIGURE 2.31 Hexagonal pattern of cells used in cellular telephone system.

the reuse pattern can be defined relative to a given reference cell as follows: *Move i cells along any chain of hexagons, and turn clockwise 60 degrees; then move j cells along the chain that lies on this new heading,* as shown in Fig. 2.31. The *j*th cell and the reference cell are *cochannel cells.* That is, the same frequency channels are used in each of these cells. With this hexagonal geometry, the cells form natural *clusters* around the reference cell and each of its cochannel cells. Each cell in a cluster uses different frequencies, but all available frequencies are used in each cluster. The number of cells per cluster is given by

$$N = i^2 + ij + j^2 \tag{2.130}$$

For the geometry shown in Fig. 2.31, with $i = 2$ and $j = 1$, there are seven cells per reuse cluster. The ratio of D, the distance between the centers of nearest-neighboring cochannel cells, to R, the cell radius, is the *normalized reuse distance* and is given by

$$\frac{D}{R} = \sqrt{3N} \tag{2.131}$$

With different choices of i and j, we obtain *reuse factors* of 1-in-3, 1-in-4, 1-in-7, 1-in-9, 1-in-12, etc. In terrestrial systems, cochannel isolation is determined by propagation losses; consequently, the reuse distance is a function of the propagation loss. (For satellite systems, the reuse distance is a function of the isolation between the spot beams that form the cells.) In the preceding expression, the reuse distance D is a function of both the cell radius and N, the number of cells per cluster.

Recall our model for the median propagation loss between a transmitter and a receiver from Section 2.4.1, which is given by

$$P_R = \frac{\beta P_T}{(r/r_0)^n} \tag{2.132}$$

where β is the loss at distance r_0 and n is a path-loss exponent that typically ranges from two to five, depending on the propagation environment (see Table 2.1). For users with similar modulation and transmit power, and with similar propagation environments, the mean carrier-to-interference ratio can be approximated by

$$\begin{aligned} \frac{C}{I} &= \frac{\text{Received power of desired user}}{\sum \text{Received powers of interfering users}} \\ &= \frac{\beta P_T/(r_d/r_0)^{n_d}}{\sum_{k \neq d} \beta P_T/(r_d/r_0)^{n_k}} \\ &= \frac{r_d^{-n_d}}{\sum_{k \neq d} r_k^{-n_k}} \end{aligned} \tag{2.133}$$

where the subscript d corresponds to the desired user and $k \neq d$ corresponds to interfering cochannel users. Equation (2.133) assumes that βP_T is the same for all users. The range r_k is the range of the kth transmitter to the receiver of interest.

With frequency-division multiple access (FDMA) strategies, frequencies are not reused in each cell, since that would cause excessive *cochannel interference.* (FDMA was discussed briefly in Chapter 1, and will be discussed in greater detail in Chapter 3.)

EXAMPLE 2.16 Minimum Reuse Pattern

Consider a cellular system with a one-in-N reuse pattern, where the base station receivers are at the center of each cell. The desired user is at the edge of the cell. Assume that the six closest interferers in a cellular system cause most of the interference and that these interferers are at the centers of their cells. What is the best reuse pattern that achieves a carrier-to-interference ratio of $(C/I)_{\min}$?

From Eq. (2.133), for the desired user at the edge of the cell, $r_d = R$, and for all cochannel users, $r_k \approx D$. So

$$\frac{R^{-n}}{6D^{-n}} \geq \left(\frac{C}{I}\right)_{\min} \tag{2.134}$$

where R is the cell radius, D is the reuse distance, and n is the path-loss exponent. It follows from Eqs. (2.131) and (2.134) that the number of cells per cluster (the inverse of the *frequency reuse factor*) has a lower bound of

$$N \geq \frac{1}{3}\left[6\left(\frac{C}{I}\right)_{\min}\right]^{2/n} \tag{2.135}$$

If the minimum tolerable carrier-to-interference ratio for an analog communication system is 18 dB and the path-loss exponent is 3.1, then

$$N \geq \frac{1}{3}[6\times 10^{18/10}]^{2/3.1} = 15,$$

The smallest value of N satisfying both this equation and Eq. (2.130) is $N = 19$, with $i = 2$ and $j = 3$. That is, to achieve the desired C/I ratio, the smallest cluster size and reuse factor is 19. ■

Problem 2.18 The minimum tolerable C/I ratio depends on the modulation and coding strategy and the quality of service required. Suppose a narrowband digital system requires a $C/I = 12$ dB. What would be the maximum frequency reuse factor? If the addition of forward error-correction coding would reduce this to 9 dB without increasing the signal bandwidth, what would be the relative improvement in reuse factor? Assume that the propagation loss exponent is 2.6.

Ans. For a C/I ratio of 12 dB, the minimum reuse is 12. For a C/I ratio of 9 dB, the minimum reuse is 7. ■

2.9 LINK CALCULATIONS

2.9.1 Free-Space Link Budget

In designing a system for reliable communications, we must perform a link budget calculation to ensure that sufficient power is available at the receiver to close the link and meet the SNR requirement.

The basis for the link budget is the Friis equation (2.11) described in Section 2.2.3. This equation describes the relationship between the received power and the transmitted power under given assumptions about propagation. The equation does not include the effects of noise. As we have seen, there are many sources of noise in wireless communications: receiver noise, antenna noise, artificial noise, and multiple access interference. In what follows, we assume that the dominant effect is receiver noise and that this noise is characterized by the single-sided noise spectral density N_0.

To include the effects of noise, we divide each side of the Friis equation by the noise density N_0. For free-space propagation, the basic link budget equation is given by

$$\frac{P_R}{N_0} = \frac{P_T G_R G_T}{L_p k T_e} \tag{2.136}$$

where $N_0 = kT_e$ and T_e is the equivalent noise temperature of the system. For satellite applications, this equation is often written as

$$\frac{C}{N_0} = \text{EIRP} - L_p + (G/T) - k \tag{2.137}$$

where all quantities are expressed in decibels and

$C/N_0 = P_R/N_0$	is the received carrier-to-noise density ratio (dB-Hz);
$\text{EIRP} = G_T P_T$	is the *equivalent isotropic radiated power* of the transmitter (dBW);
L_p	is the path loss (dB);
$G/T = G_R/T_e$	is the ratio of receiver antenna gain to noise temperature (dB-K^{-1});
k	is Boltzmann's constant (-228.6 dBW-sK^{-1}).

The C/N_0 ratio is one of a number of equivalent ways of expressing the SNR. Indeed, in many cases it is the preferred choice, because it makes no assumption about the underlying modulation strategy.

EXAMPLE 2.17 *G/T* of a Satellite

Consider a geosynchronous satellite whose global beam covers all of the visible Earth's surface. The radius of the Earth is about 6400 km, and the altitude of the satellite is 36,000 km.

Consequently, the satellite antenna gain, relative to the isotropic situation, is equivalent to the inverse ratio of the cross-sectional area of the earth to the surface area of a sphere at 36,000 km, assuming 100 percent efficiency; that is,

$$G_{\text{sat}} \approx \frac{4\pi R_{\text{alt}}^2}{\pi R_{\text{earth}}^2} = \frac{4\pi \times 36{,}000^2}{\pi \times 6400^2} = 126.6 \tag{2.138}$$

Expressed in decibels, the satellite antenna (power) gain is 21.0 dB. The satellite antenna is looking directly at the Earth and thus has an equivalent noise temperature of approximately 290°K. Usually, the satellite receiver noise can be considered small relative to the antenna noise. Consequently, the G/T ratio of the satellite receiving system is

$$\begin{aligned}(G/T) &= G_{\text{sat}}(dB) - 10\log_{10}T_{\text{ant}} \\ &= 21.0 - 24.6 \\ &= -3.6 \text{ dBK}^{-1}\end{aligned}$$

The system has a G/T ratio of -3.6 dBK^{-1}. ■

Problem 2.19 A fixed satellite terminal has a 10-m parabolic dish with 60% efficiency and a system noise temperature of 70°K. Find the G/T ratio of this terminal at 4 GHz. Suppose it was a terrestrial mobile radio using an omnidirectional antenna. What would you expect the equivalent noise temperature of the mobile antenna to be?
Ans. The satellite terminal G/T is 31.7 dBK^{-1}. The mobile antenna has a noise temperature of approximately 290°K. ■

EXAMPLE 2.18 Link Budget from Earth Station to Satellite.

A simple link budget for transmitting from a ground Earth station to a satellite is provided in Table 2.3. The table is broken down into four subsections that can be loosely identified with the transmitter, the transmission path, the receiver, and the combined system. We break down the line items of the table according to these divisions:

1. *Earth station transmitter*, which identifies:
   - the transmit frequency from the ground Earth station to the satellite.
   - the transmitted power transmitted by the Earth station (as measured at the transmit antenna terminals).
   - the antenna gain. If the antenna is a parabolic dish, then the power gain can be determined from the diameter and the transmit frequency.
   - the transmit EIRP is the dB sum of the transmit power and the antenna gain.
2. *Losses*, which identify:
   - the satellite elevation. For a geostationary satellite, this is the only geometric information required to compute the range of the satellite.
   - the satellite range, as computed from geometrical considerations. (See Problem 2.20.)
   - the free-space loss, which is determined from the satellite range and transmit frequency.

TABLE 2.3 Sample link budget from Earth station to satellite.

Parameter	Units		Comments
Earth station transmitter			
Transmit frequency	GHz	6.0	
Transmit power	dBW	−10	100-mW transmitter
Transmit antenna gain	dBi	30	60% efficiency for a 10-m parabolic dish
Transmit EIRP	dBW	20	
Losses			
Satellite elevation	°	45	
Satellite range	km	37,630	
Free-space loss	dB	193.5	$=20 \times \log_{10}(2\pi r/\lambda)$
Satellite receiver			
Antenna gain	dB	23.0	
Received power	dBm	−150.5	
Receive G/T	dB-K	−2.5	Global beam looking at Earth (290°K)
System			
Boltzmann's constant	dBW-K	−228.6	
Uplink C/N_0	dB-Hz	52.6	From link equation (2.137)

3. *The satellite receiver*, which identifies:
   - the satellite antenna gain. This also depends on the transmit frequency.
   - the received power level. This is the dB sum of the transmit EIRP, the free-space loss, and the receive antenna gain.
   - the satellite G/T. This empirically provided parameter characterizes the satellite design. In satellite systems, a terminal is usually specified in terms of its G/T rather than in terms of separate gain and system noise components.
4. *System*, which identifies
   - Boltzmann's constant, which is necessary to convert a system temperature into an equivalent noise density.
   - the uplink C/N_0, as determined by Eq. (2.137).

The conclusion from this table is that the link from Earth station to satellite will support any signaling system that requires a C/N_0 ratio of 52.6 dB-Hz or less. ■

Problem 2.20 For a geostationary satellite at altitude h (36,000 km), determine a formula relating the range r from the satellite to an earth station to the satellite elevation ϕ relative to the earth station. (Let $R_e = 6400$ km be the radius of the Earth.)
Ans. The range is

$$r = R_e \sin\phi + \sqrt{(R_e + h)^2 + R_e \cos^2\phi}$$ ■

EXAMPLE 2.19 Satellite-to-Mobile Terminal Link Budget

In Table 2.4, we show a more detailed link budget for the link from the satellite to a mobile Earth station and combine it with the link results of the previous example. In this case, the satellite transmit power is proportional to the received power obtained from Table 2.3. Table 2.4 is broken down into subsections similar to those in Table 2.3, but we have included a number of additional effects corresponding to a more realistic link budget:

1. *Satellite transmitter*, which includes
   - the transmit frequency.
   - the receive power. Since this is a two-hop system, the transmit power is determined by the power received at the satellite from the earth station (Example 2.18).
   - the satellite gain. This depends on the satellite design and is an empirically provided parameter.
   - the transmit antenna gain. This again depends on the antenna size and the transmit frequency.
   - transmit EIRP. For the satellite, this is the dB sum of the satellite-received power, the satellite amplifier gain, and the satellite transmit antenna gain.
2. *Losses*, which include
   - the elevation of the satellite with respect to the mobile terminal. This is typically different from the elevation of the satellite with respect to the (base) Earth station.
   - the range of the downlink. For mobile terminals, the link budget must consider the worst-case path loss. This is equivalent to the largest range or smallest elevation angle.
   - the free-space loss.
   - the absorption loss due to the atmosphere. There are losses in RF signal strength due to atmospheric absorption, which will be largest at the lowest elevation angles, because, at those angles, the signal passes through more of the atmosphere.
3. *Earth station receiver*, which includes
   - the receive G/T.
   - implementation losses. These losses, which are due to nonideal implementation of the receiver/modem, can include filtering losses, synchronization losses, and distortion losses due to slight nonlinearities in receiver processing, finite precision effects, etc. They could also include losses due to interference, or the latter could be included as a separate line item.
4. *System*, which includes
   - Boltzmann's constant.
   - the downlink C/N_0 calculated with the use of Eq. (2.137).
   - the uplink C/N_0, which is included for reference purposes.
   - the combined C/N_0 for the two links. This calculation is described next.

The overall link has two components in this example: Earth station–to-satellite (uplink) and then satellite-to–mobile terminal (downlink). *The use of the terms* uplink *and* downlink *in satellite systems differs from those used in terrestrial systems.* We assume the satellite acts as a "bent pipe"; that is, it simply retransmits whatever it receives, and there is no onboard processing for regenerating the signal. If it is a bent-pipe satellite, the overall link carrier-to-noise ratio is given by

$$\frac{C}{N_0} = \left(\left(\frac{C}{N_0}\right)^{-1}_{\text{uplink}} + \left(\frac{C}{N_0}\right)^{-1}_{\text{downlink}}\right)^{-1} \tag{2.139}$$

TABLE 2.4 Downlink budget from satellite to mobile Earth station.

Parameter	Units		Comments
Satellite transmitter			
Transmit frequency	GHz	1.5	
Receive power	dBW	−150.5	From uplink budget
Satellite gain	dB	150	Internal satellite amplifiers
Transmit Antenna gain	dBi	26	Corresponds to 10-m parabolic dish on the satellite
Transmit EIRP	dBW	25.5	
Losses			
Elevation	°	20	Worst-case elevation angle
Range	km	39,809	
Free-space loss	dB	181.9	
Absorption loss	dB	0.2	Due to atmosphere
Earth station receiver			
Receive G/T	dBi	−22.6	Approximately omnidirectional in upper hemisphere
Receive implementation loss	dB	1.0	Loss due to nonideal receiver
System			
Boltzmann's constant	dBW-sK^{-1}	−228.6	
Downlink C/N_0	dB-Hz	48.6	From link equation
Uplink C/N_0	dB-Hz	52.6	From uplink budget
Overall C/N_0	dB-Hz	47.1	

This calculation sums the noise of both the uplink and the downlink, weighted according to the received power level. The calculation is done in absolute terms, not in decibels. The result is shown in the system section of Table 2.4. In this case, the C/N_0 ratio of the downlink is dominant, and we say that the *channel is downlink limited*.

The final step is to compare the overall C/N_0 ratio with that required by the modem delivering the service. If the C/N_0 ratio is less than the required one, then reliable communication is not possible. If the C/N_0 ratio is greater than the required one, we say that the link has a *margin*. Margin is desirable for a variety of reasons. Transmit power that varies with temperature and losses that vary with weather conditions are examples of other considerations that have not been included in the link budget. In the absence of further information, the margin is a tolerance available for uncharacterized transmission impairments. If further information were available about any of these effects, it could be included as part of the link budget. ■

Problem 2.21 Numerous other quantities may be included in the satellite link budget. For example, if the satellite amplifier (which is typically nonlinear) is shared with a number of carriers, then *intermodulation distortion* will be generated. When there is a large number of equal-power carriers present, intermodulation distortion can be modeled as white noise with spectral density I_0. This will produce a term C/I_0 at the satellite that must be combined with the uplink and downlink C/N_0 to produce the overall $C/(N_0 + I_0)$. Given the uplink and downlink C/N_0 of Examples 2.18 and 2.19, what is the overall $C/(N_0 + I_0)$ if $C/I_0 = 50$ dB-Hz?

Ans. The $C/(N_0 + I_0)$ is 45.3 dB-Hz. ■

2.9.2 Terrestrial Link Budget

For a terrestrial link budget, many of the line items are similar to those in the free-space link budget examples, but their calculations are different.

EXAMPLE 2.20 Terrestrial Link Budget

The link budget in Table 2.5 is modeled after a police radio service that is intended to service the city core and the surrounding suburban and rural areas. The base station is assumed to be located close to the city core, with an antenna mounted on a high tower that provides a relatively unobstructed view of the surrounding rural area. In this case, although we would expect large path-loss exponents in the built-up areas, the median path losses in the outlying rural areas are assumed to be close to free-space losses. Shadowing losses are still expected, due to small undulations of the terrain, trees, bushes, and other forms of vegetation. Table 2.5 is broken down into the sections described on the next page.

TABLE 2.5 Terrestrial link budget from base to mobile.

Parameter	Units		Comments
Base station transmitter			
Transmit frequency	MHz	705	Mobile public safety band
Transmit power	dBW	15	
Transmit Antenna gain	dBi	2	Uniform radiation in azimuth
Transmit EIRP	dBW	17	Maximum EIRP of 30 dBW
Power at 1 m	dBm	17.6	$P_0 = P_T/(4\pi/\lambda)^2$
Losses			
Path-loss exponent		2.4	Applicable at edge of coverage
Range	km	10	Range at edge of coverage
Median path loss	dB	96	$2.4 \times 10\log_{10}(r/r_0)$
Lognormal shadowing	dB	8	standard deviation of log-normal shadowing
Shadowing margin	dB	13.2	for 95% availability ($1.65\sigma_{dB}$)
Received Signal			
Receive antenna gain	dBi	1.5	Vertically polarized whip antenna
Received signal strength	dBm	–90.1	$P_R = P_0 + G_R - 21\log_{10}(r/r_0) - M_{\text{shadow}}$
Receiver characteristics			
Required C/N_0	dB-Hz	69.8	from modem characteristics
Boltzmann's constant	dBm-K	–198.6	
Receiver noise figure	dB	6.0	provided
Receiver sensitivity	dBm	–98.2	$S = C/N_0 + NF + kT_0$
Margin	dB	8.1	Margin $= P_R - S$

1. Base station transmitter, which includes
   - the transmit frequency.
   - the transmit power at the antenna terminals.
   - the transmit antenna gain. We assume a donut-shaped radiation pattern that is approximately uniform in azimuth.
   - the transmit EIRP is the dB sum of the previous two quantities.
2. Losses, which include:
   - the path-loss exponent. Slightly worse than free-space propagation is assumed for the rural areas at the edge of coverage.
   - the range. This is the expected radius of the service area.
   - the median-path loss. This is obtained from the general statistical model, assuming a path-loss exponent of 2.4.
   - lognormal shadowing. This describes the standard deviation of the shadowing expected at the edge of coverage.
   - shadowing margin. This is the margin required to maintain 95% availability. The shadowing margin is obtained from the lognormal distribution with the preceding deviation. Since police radio is a safety service, it typically has higher availability requirements than nonsafety services.
3. *Received Signal*, which includes
   - the receive antenna gain. Here, we assume a whip antenna; a *whip antenna* is a flexible rod or wire antenna, such as commonly used for cellular telephones or car radios.
   - the received signal strength. This is the reference power at 1 meter, minus the path losses, minus the shadowing margin, plus the receive antenna gain.
4. *Receiver characteristics*
   - computes the receiver sensitivity from modem characteristics and the receiver noise figure.
   - compares the receiver sensitivity with the received signal level.

The receiver noise has a spectral density that is characterized by the product of the receiver noise factor F and thermal noise at the nominal temperature, namely, kT. Consequently, noting that $N_0 = F(kT)$, we observe that the received signal strength S and the carrier-to-noise ratio are related by

$$S(\text{dBm}) = \frac{C}{N_0}(\text{dBHz}) + kT(\text{dBHz}) + F(\text{dB}) \tag{2.140}$$

where T is the nominal temperature (290°). Equation (2.140) is used for the calculation in the receiver characteristics section of Table 2.5. Equation (2.140) also defines the relationship between the receiver sensitivity and the threshold C/N_0 of the receiver. ■

Problem 2.22 Repeat the link budget of Table 2.5, analyzing the performance in the city core. Assume that the maximum range within the city core is 2 km, but that the path-loss exponent is 3.5 and the lognormal shadowing deviation is 10 dB. Is this service limited by the receiver sensitivity? What is the expected service availability for the city core?

Ans. Yes, it is limited by the receiver sensitivity; only 11.5 dB of margin is available to compensate shadowing, and this implies an availability of approximately 75%. ■

2.10 THEME EXAMPLE 1: OKUMURA–HATA EMPIRICAL MODEL[16]

We now discuss an example of a land mobile propagation model that is based on empirical measurements. This model applies to propagation in the UHF/VHF frequency band from 150 MHz to 1 GHz. The original data were collected by Okumura and others in several areas of Japan. Numerous charts were provided illustrating the many factors that affect land mobile propagation, including building characteristics and antenna height. Hata later provided analytical approximations to these data that captured most of the major effects. This analytical representation is described here.

This *Okumura–Hata model* predicts the median-path loss in three types of environment: urban, suburban, and open. The median-path losses, in dB, for the three environments are given by the equations

$$\begin{aligned} L_p &= A + B\log_{10} r && \text{Urban} \\ L_p &= A + B\log_{10} r - C && \text{Suburban} \\ L_p &= A + B\log_{10} r - D && \text{Open} \end{aligned} \tag{2.141}$$

where r is the range in kilometers. The parameters in these equations depend on the frequency of operation, f_c, the height of the transmitting station, h_b, and the height of the receiving station, h_m. These parameters are given by the empirical formulas

$$\begin{aligned} A &= 69.55 + 26.16\log_{10} f_c - 13.82\log_{10} h_b - a(h_m) \\ B &= 44.9 - 6.55\log_{10} h_b \\ C &= 5.4 + 2[\log_{10}(f_c/28)]^2 \\ D &= 40.94 + 4.78(\log_{10} f_c)^2 - 18.33\log_{10} f_c \end{aligned} \tag{2.142}$$

where f_c is measured in MHz, h_b and h_m are in meters, and $a(h_m)$ is a correction factor that is defined in what follows. This model is valid for the following range of parameter values:

$$\begin{aligned} 150\ \text{MHz} < f_c < 1000\ \text{MHz} \\ 30\ \text{m} < h_b < 200\ \text{m} \\ 1\ \text{m} < h_m < 10\ \text{m} \\ 1\ \text{km} < r < 20\ \text{km} \end{aligned} \tag{2.143}$$

The parameter A represents a fixed loss that depends on the frequency of the signal and the heights of the base station and mobile station antennas. Its dependence on frequency is approximately a 2.6 power law, slightly higher than the square-law dependency we would expect for free-space propagation and certainly not independent of frequency as predicted by the plane-Earth model. The dependence on antenna heights has a term proportional to $h_b^{1.382}$ that is similar to the plane-Earth model, which has a quadratic dependence on h_b. The term $a(h_m)$ is a correction factor based on the mobile antenna height and is a function of the environment. For a large city, it is given by

$$\begin{aligned} a(h_m) &= 8.29(\log 1.54h_m)^2 - 1.1\text{dB} && \text{for } f_c \le 300\ \text{MHz} \\ a(h_m) &= 3.2(\log 11.75h_m)^2 - 4.97\text{dB} && \text{for } f_c > 300\ \text{MHz} \end{aligned} \tag{2.144}$$

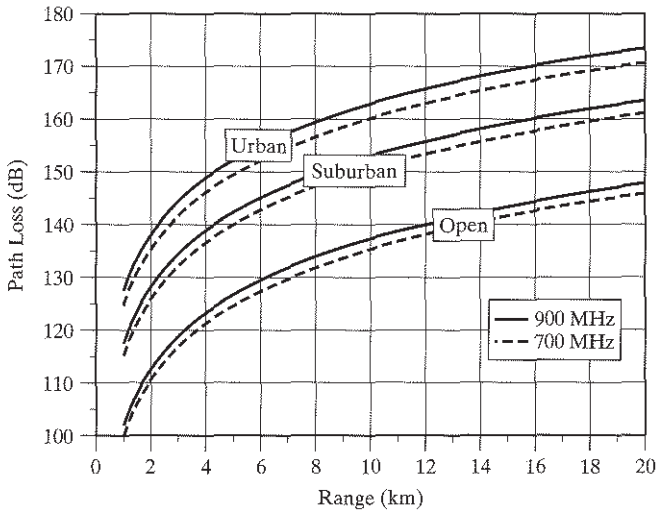


FIGURE 2.32 Performance loss predictions with Okumura–Hata propagation model with 30-m base antenna and 1-m mobile antenna for a midsized city.

For small and medium-sized cities, the correction factor is given by

$$a(h_m) = (1.1\log f_c - 0.7)h_m - (1.56\log f_c - 0.8)dB \tag{2.145}$$

for all frequencies. The dependence on the mobile antenna height as described by $a(h_m)$ differs significantly from the quadratic dependence of the plane-Earth model.

From Eq. (2.141), the parameter B represents the path-loss exponent. The worst-case path-loss exponent is approximately 4.5, improving with higher base station antennas, as we would intuitively expect.

The parameters C and D represent reductions in the fixed losses for the less demanding suburban and open propagation environments. These improvements are frequency dependent.

Although this model is mathematically complicated, it is straightforward to evaluate on a computer and estimate path losses. In Fig. 2.32, the computed path loss for frequencies of 700 and 900 MHz is plotted for the three different environments of the Okumura–Hata model. We see two expected trends:

1. Propagation losses increase with frequency.
2. Propagation losses increase in built-up areas.

There have been enhancements and modifications to the Okumura–Hata model to extend the frequencies, the distances, and the environments over which it is valid.

EXAMPLE 2.21 Base Station Antenna Height

For public safety applications in the 700-MHz band, the maximum transmitter power is 30 W. What base station height would be needed to communicate with a patrol car at a distance of 10 km in a suburban environment? Assume that the mobile station antenna height is 1 m.

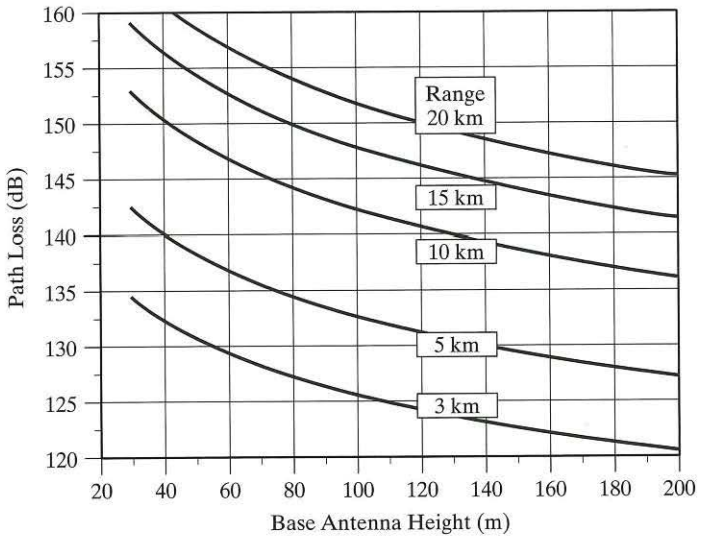


FIGURE 2.33 Path loss for Okamura–Hata model in a midsized suburban environment at 700 MHz as a function of base antenna height.

In Fig. 2.33, we plot the path loss from Eq. (2.141) of the Okumura–Hata model as a function of the base station antenna height for the suburban portion of a midsized city. Assuming a receiver sensitivity of –105 dBm and a fading margin of 10 dB, the allowable path loss would be

$$\begin{aligned} L_p(\text{dB}) &= P_T(\text{dBm}) - P_R(\text{dBm}) - M_f(\text{dB}) \\ &= 45 - (-105) - 10 \\ &= 140\text{dB} \end{aligned}$$

From Fig. 2.33, we find that the base station height would need to be at least 130 meters. ■

EXAMPLE 2.22 Receiver Range

A mobile phone operating at 900 MHz has a receiver sensitivity of –95 dBm and an antenna with an effective antenna height of 1 m and gain of 0 dBi. If the base antenna transmits 2 watts with an antenna height of 60 m, what is the range of the cell phone in an urban environment?

At the edge of coverage, the path loss (assuming no shadowing margin) is

$$\begin{aligned} L_p(\text{dB}) &= P_T(\text{dBm}) - P_R(\text{dBm}) \\ &= 33 - (-95) \\ &= 128\text{dB} \end{aligned}$$

Using the Okumura–Hata model and, in particular, the results of Fig. 2.33, we find that the expected range is slightly less than 3 km. ■

Problem 2.23 Evaluate the path loss at 900 MHz, using the Okumura–Hata model for the suburban environment. Assume that the base and mobile station antenna heights are 30 m and 1 m, respectively. ■

2.11 THEME EXAMPLE 2: WIRELESS LOCAL AREA NETWORKS[17]

Wireless local area networks (LANs) are being implemented in offices and plants to reduce the cost and inconvenience of wiring network connections. There are also applications, such as airport lounges, libraries, etc., in which the users are transient and a wireless connection to a local node may be convenient. Of the various approaches to wireless LANs, many fall under the umbrella of the IEEE 802.11 standard. The one in most common use is IEEE 802.11b: however, we will discuss the more future-oriented IEEE 802.11a. This standard operates in indoor environments in the 5-GHz band and offers information rates ranging from 6 Mbps up to 54 Mbps. In this section, we will consider propagation and noise issues related to this service.

2.11.1 Propagation Model[18]

The propagation model that the International Telecommunications Union (ITU) has recommended for evaluating indoor systems in the 5-GHz band is given by

$$P_R = P_T - 41\text{dB} - 31\log_{10} r - \sum_q \text{WAF}(q) - \sum_q \text{FAF}(q) \tag{2.146}$$

where the wall attenuation factor (WAF) and floor attenuation factor (FAF) were described in Section 2.5. Let us consider each of the terms on the right-hand side of Eq. (2.146) in turn. The second term is a fixed loss of 41 dB. The third term indicates a path-loss exponent of 3.1, where r is the distance in meters. The fourth and fifth terms represent the sum of losses the signal encounters, depending on how many walls and floors are in the path, as described in Section 2.5.

2.11.2 Receiver Sensitivity

The IEEE 802.11a wireless LAN standard recommends that the receiver have a noise figure of 10 dB or better. On the basis of this figure, we can calculate a nominal value for the noise floor of the receiver of

$$N_0 = FkT_0 \tag{2.147}$$

or, expressed in decibels,

$$\begin{aligned} N_0 &= 10\log_{10} F + k(\text{dBm} - sK^{-1}) + T(\text{dBK}) \\ &= 10 + (-198.6) + 24.6 \\ &= -164 \text{ dBm/Hz} \end{aligned} \tag{2.148}$$

The 6-Mbps service in this standard uses binary phase-shift keying (BPSK) modulation. We shall discuss BPSK modulation in Chapter 3—but for now, we note that, for digital services, the SNR is often expressed as E_b/N_0, where E_b is the energy transmitted per bit. Theoretically, BPSK modulation requires an E_b/N_0 of 11 dB to achieve a bit error rate of 10^{-6}. In practice, however, with a low-cost solution, there will be implementation losses, which we will assume are no larger than 3 dB. Consequently, the practical requirement for E_b/N_0 is 14 dB.

With the foregoing information, we can estimate the receiver sensitivity. Equation (2.140) defines the sensitivity, and we note that the carrier power is related to the energy per bit by $C = E_b R_b$, where R_b is the bit rate. Substituting the expression for C into Eq. (2.140), we can relate the sensitivity to the required E_b/N_0 by

$$S(\text{dBm}) = E_b/N_0(\text{dB}) + R_b(\text{dBHz}) + N_0(\text{dBm/Hz}) \tag{2.149}$$

That is, the sensitivity is equal to the required E_b/N_0 times the data rate R_b and multiplied by the noise density N_0. For this example, the required receiver sensitivity is

$$\begin{aligned} S &= 14 + 10\log_{10} 6\times10^6 + (-164) \\ &= -82 \text{ dBm} \end{aligned} \tag{2.150}$$

Although we have made some simplifying assumptions, the result given in Eq. (2.150) agrees with the required sensitivity specified in the standard for reception under AWGN conditions.

Problem 2.24 The 54-Mbps service of IEEE 802.11a uses 64-quadrature-amplitude-modulation (QAM). (This method of modulation is considered in Chapter 3) Suppose the practical E_b/N_0 required to achieve a bit error rate of 10^{-6} with 64 QAM is 30 dB. For this data rate, what is the sensitivity of the receiver just discussed?
Ans. The required sensitivity is –53 dBm, including a 3-dB implementation loss. ■

2.11.3 Range

Radios for 5-GHz wireless LAN applications must transmit 200 milliwatts (23 dBm) or less. What is the expected range of this service? If we assume an open office environment with no walls, we suppose that and the transmitter and receiver are on the same floor, then the propagation model of Eq. (2.144) simplifies to

$$P_R = P_T - 41(\text{dB}) - 31\log_{10} r \tag{2.151}$$

To close the link with the 6-Mbps service, the received power must be at least –82 dBm from Section 2.11.2, so

$$\begin{aligned} 31\log_{10} r(m) &= P_T(\text{dBm}) - P_R(\text{dBm}) - 41\,\text{dB} \\ &= 23 - (-82) - 41 \\ &= 64 \text{ dB} \end{aligned} \tag{2.152}$$

Solving this equation for r gives a maximum range of 116 meters.

Problem 2.25 From the results of Problem 2.24, what is the maximum range expected with the 54-Mbps service?
Ans. The expected range is 13.5 m. ■

2.11.4 Power-Delay Profile

Empirical measurements of the multipath intensity profile indicate that the rms delay spread for indoor applications in the 5-GHz band is typically between 40 and

60 nanoseconds. A discrete-time model that has been suggested for the multipath power-delay profile is depicted in Fig. 2.34. The model shows an exponential decrease in the average power of a multipath ray as a function of the delay.

With this model, the impulse response $\{\tilde{h}_k\}$ of the channel is composed of complex samples with a random uniformly distributed phase and Rayleigh-distributed magnitude, with average power calculable from the power-delay profile shown in the figure. Mathematically, the samples of the impulse response are given by

$$\tilde{h}_k = N(0, \sigma_k^2/2) + jN(0, \sigma_k^2/2) \tag{2.153}$$

where $N(\mu,\sigma^2)$ is a Gaussian random variable with mean μ and variance σ^2. The variances in Eq. (2.153) are given by

$$\sigma_k^2 = \sigma_0^2 e^{-kT_s/T_{\text{rms}}} \tag{2.154}$$

where T_s is the sampling period of the impulse response and T_{rms} is the rms delay spread. The constant $\sigma_0^2 = 1 - e^{-T_s/T_{\text{rms}}}$ is chosen to normalize the average power; that is,

$$\sum_{k=0}^{\infty} \sigma_k^2 = 1 \tag{2.155}$$

For a sampling period of 12.5 nanoseconds and T_{rms} of 50 nanoseconds, an example of the multipath spectrum for this application is shown in Fig. 2.35. The spectrum is

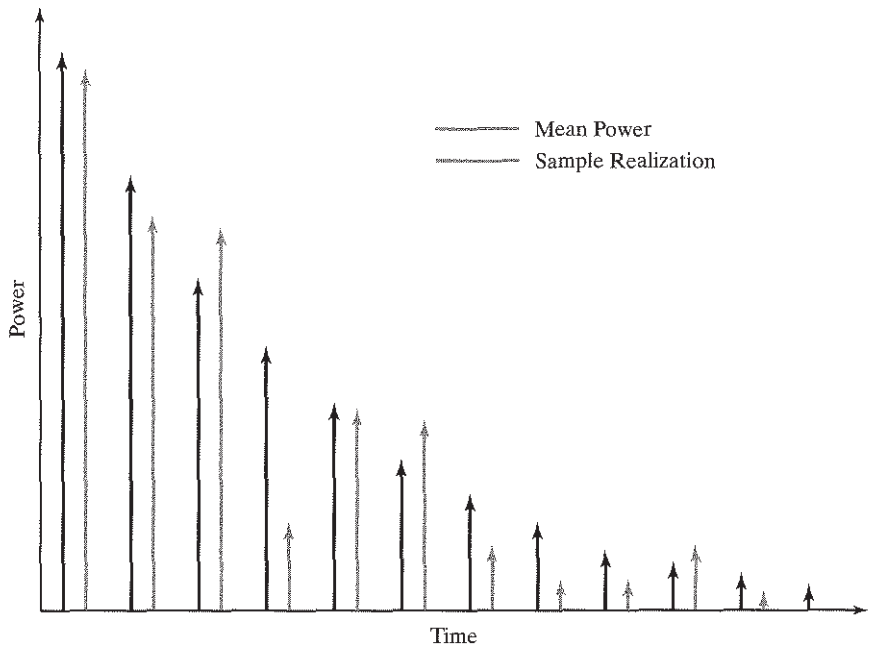


FIGURE 2.34 Impulse response model for multipath channel.

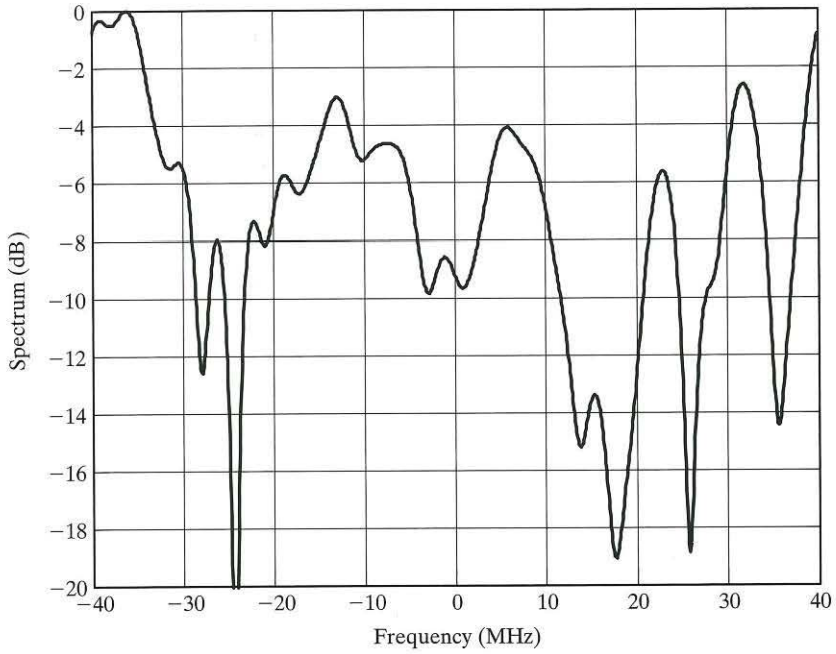


FIGURE 2.35 Spectral characteristics of 50-nanosecond rms delay spread channel.

depicted at complex baseband and ranges from –40 MHz to +40 MHz. The signal bandwidth for Fig. 2.35, IEEE 802.11a is approximately 20 MHz for all data rates. We see that there is significant variation over any 20-MHz band in this spectrum. The channel is frequency selective and, from Fig. 2.35, the coherence bandwidth is on the order of 1 Megahertz.

2.11.5 Modulation

As the results in the previous section indicate, the wireless LAN application has significant *frequency-selective characteristics.* One method of compensating for these characteristics is to include in the receiver a device known as an *equalizer,* which effectively estimates the channel and tries to invert its effects at the receiver. Unfortunately, these devices are difficult to implement for this application. The alternative approach is to design the modulation so that it is easier to compensate for the channel characteristics. This is what has been done in the IEEE 802.11a standard.

The standard uses a form of *multicarrier modulation* known as *orthogonal frequency division multiplex* (*OFDM*). This technique, which will be described in Chapter 3, consists of a number of narrowband carriers transmitted in a synchronous fashion. When all the details in the IEEE 802.11a standard are accounted for, it is seen that there are 52 carriers transmitted, each of which has a nominal bandwidth of 312.5 kHz.

By inspection of the spectrum plot of Fig. 2.34, we see that 312.5 kHz is significantly less than the coherence bandwidth of the channel. That is, over most 312.5-kHz pieces of the spectrum, the channel is approximately constant. This means that the

channel appears like a flat-fading channel to an individual carrier. As we shall see in Chapter 3, it is easier to track and compensate for a flat-fading channel than for a frequency-selective channel. Theme Example 2 is a precursor of the discussions in the chapters that follow, where we will see a number of cases in which the modulation strategy has been selected to take advantage of, or to compensate for, the wireless channel characteristics.

2.12 THEME EXAMPLE 3: IMPULSE RADIO AND ULTRA-WIDEBAND[19]

Traditional radio transmission strategies impress information on an RF carrier, a process known as modulation, for propagation from the transmitter to the receiver. The bandwidth of the resulting signal is typically much less than the carrier frequency. The result is known as a bandpass signal. As we shall see in the following chapters, the analysis of bandpass signals and systems is fundamental to much of communications theory. However, there are other methods of transmitting information via radio waves.

One method that 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. This is reminiscent of the "spark-gap" transmitter used by Marconi when he made the first radio transmission across the Atlantic Ocean. Since the pulse 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*.

There are several possible definitions for ultra-wideband transmission, but two popular ones are:

- any transmission where the RF bandwidth exceeds 1 GHz, or
- any transmission where the RF bandwidth at the –10 dB points of the spectrum exceeds 25% of the center frequency.

Ultra-wideband transmission is considered best for low-power indoor applications where there is high clutter, that is, the surrounding environment causes significant amounts of multipath.

Impulse radio is a form of ultra-wideband radio that is based on the use of very short pulses. One type of pulse used for this application is the *Gaussian monocycle*, which is based on the first derivative of the Gaussian function. The waveform of the Gaussian monocycle is given by

$$v(t) = A\frac{t}{\tau}\exp\left\{-6\pi\left(\frac{t}{\tau}\right)^2\right\} \tag{2.156}$$

where A is an amplitude scale factor and τ is the time constant of the pulse. This signal is depicted in Fig. 2.36. It consists of a positive lobe followed by a negative 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 Fig. 2.37. (Fourier theory is reviewed in Appendix A.) The frequency axis in Fig. 2.37 has been normalized in

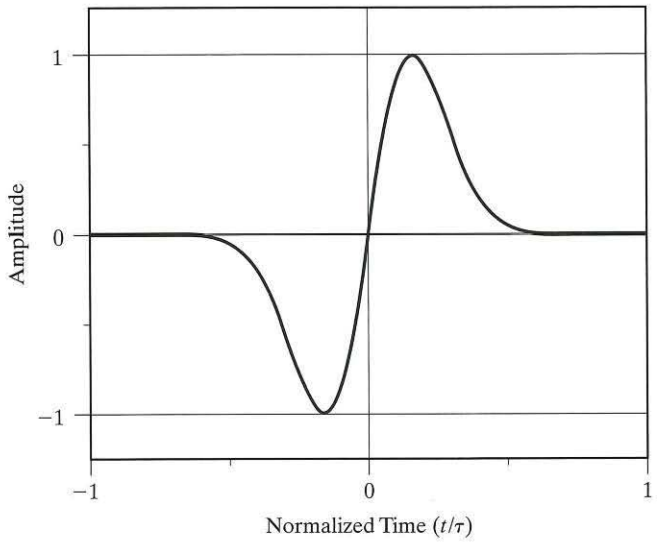


FIGURE 2.36 Illustration of a Gaussian monopulse used for impulse radio.

terms of the time constant τ; for $\tau = 1.0$ nanosecond, this frequency axis ranges from 0 to 5 GHz. In Fig. 2.37, the center frequency of the signal is approximately $1/\tau$, and at the –10 dB points the bandwidth of the signal is approximately $2/\tau$. Thus we see that the bandwidth of the UWB signal is greater than its center frequency, which certainly violates a fundamental requirement for a bandpass signal.

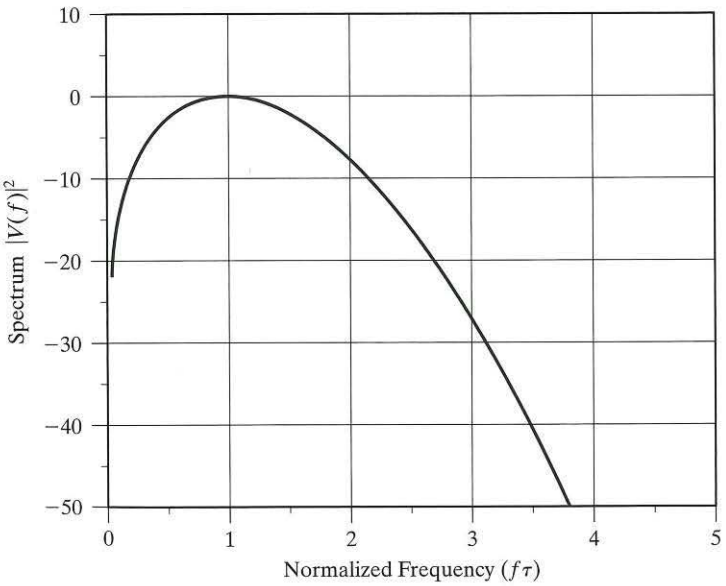


FIGURE 2.37 Spectrum of a Gaussian monopulse.

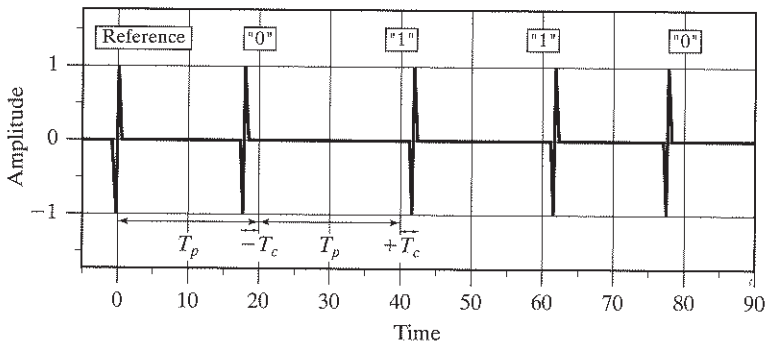


FIGURE 2.38 Pulse-position modulation of impulse radio.

There are several methods for modulating such an impulse wave. One method, known as *pulse-position modulation*, is illustrated in Fig. 2.38. With this method, there is a nominal time separation, T_p, between successive pulses. To transmit a binary signal "0," the pulse is transmitted slightly early ($-T_c$). To transmit a binary signal "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 25 nanoseconds to 1000 nanoseconds, resulting in a range of data rates from 40 Mbits/s to 1 Mbits/s.

The ultra-wideband nature of the modulated signal has both good aspects and bad. Since the signal power is spread over such a large bandwidth, the amount of power that falls in any particular narrowband channel is small. 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 spectra that may be transmitted. Due to this limitation on transmit power, ultra-wideband radio is limited to short-range applications, typically less than a few hundred meters.

EXAMPLE 2.23 Spectral Density of UWB Compared to Noise Floor

Compare the interference caused by a one-milliwatt impulse radio ($\tau = 1$ ns) into an 800 MHz radio to the noise floor of the narrowband radio. Assume the 800 MHz radio has a noise figure of 10 dB and the separation between the impulse and narrowband radios is 10 meters.

In the problem statement, one milliwatt refers to the average power transmitted by the radio. Since the impulse radio has a low duty cycle, the peak power is much higher than this average. However, the switching rate of the impulse radio is much greater than the bandwidth of most narrowband signals. This has a smoothing effect, implying that we can use the average power for analysis.

From Fig. 2.37, the average power of the impulse radio is spread over a bandwidth of between $1/\tau$ to $2/\tau$. If we use the lower of these two bandwidths, the effective power spectrum is one milliwatt spread over 1 GHz. Assuming an approximate uniform distribution for this power, this is equivalent to one picowatt per hertz, or –90 dBm/Hz. Assuming free-space propagation from an isotropic antenna, the path loss between the impulse radio and the narrowband receiver at 800 MHz is, from Eq. (2.6),

$$\begin{aligned} L_p &= \left(\frac{4\pi R}{\lambda}\right)^2 \\ &= \left(\frac{4 \times 10}{c/800 \times 10^6}\right)^2 \\ &= 112,212 \\ &= 50.5 \text{ dB} \end{aligned} \tag{2.157}$$

The resulting interference density at the narrowband receiver input is then

$$\begin{aligned} J_0 &= -90 \text{ dBm/Hz} - 50.5 \text{ dB} \\ &= -(140.5 \text{ dBm/Hz}) \end{aligned} \tag{2.158}$$

We compare this interference density to the noise floor of the 800 MHz radio, which is given by

$$\begin{aligned} N_0 &= kTF \\ &\sim -174 \text{ dBm/Hz} + 10 \text{ dB} \\ &= -164 \text{ dBm/Hz} \end{aligned} \tag{2.159}$$

where F is the noise figure of the receiver. Since the interference density is approximately 30 dB higher than the noise density of the receiver, the 800 MHz radio will see a significant degradation of performance in the presence of the impulse radio, unless it has at least 24 dB margin. ■

This example illustrates a number of interesting points. First, even though impulse radio spreads power over a large bandwidth, the transmit powers must still be very low to be comparable to the noise floor, typically in the microwatt range. Second, it reminds us that free-space path loss depends on frequency, which means it varies considerably over a wideband signal. However, the receiver antenna gain usually has an inverse dependency on frequency that compensates for this loss.

The advantage of impulse radio is that it has a very simple transmitter and receiver design. There is no up-conversion to, and down-conversion from, carrier frequencies; transmission is performed at baseband. In addition, impulse radio has the capability to provide very high data rates, albeit over short distances.

As mentioned previously, the major application of ultra-wideband radio is in indoor environments where there is significant multipath. In this application, impulse radio has a significant advantage in its ability to mitigate the effects of multipath. Multipath is due to echoes of the transmitted signal arriving slightly later than the transmitted signal. In indoor environments, the multipath differential delay, or delay spread, tends to be small, often on the order of nanoseconds.[20] Since most of the time

an impulse radio is not transmitting, by proper design, these echoes will arrive at the receiver in between the signaling epochs, as illustrated in Fig. 2.39. Consequently, for a receiver that is attuned to the timing of the main signal, the echoes will have no effect.[21] Note that the speed of light is approximately 0.3 meters per nanosecond. Hence, reflections, even with a path difference as little as one meter, can be resolved by the impulse radio receiver. The upper limit on allowable path differences is determined by the pulse interval T_p.

In summary, impulse radio is a reincarnation of very early radio transmission techniques that did not use an RF carrier. For this reason, it results in a simple implementation for the transmitter and receiver. It also permits the transmission of very high data rates over short distances in dense multipath environments. There are a number of issues related to ultra-wideband radio. A practical issue is the interference that impulse radio may cause into existing narrowband radio services. Impulse radio also poses a modeling challenge. The time resolution of impulse radio is such that it can distinguish individual multipath components, and thus the Rayleigh model for amplitude distribution of the multipath is no longer valid.

Impulse radio is not the only ultra-wideband transmission technique. Other methods, based on an extension of the direct-sequence spread spectrum (to be discussed in Chapter 5), have also been proposed. Other modifications of the basic technique move the signal spectrum to frequencies above 2 GHz to reduce some of the concerns with interference.

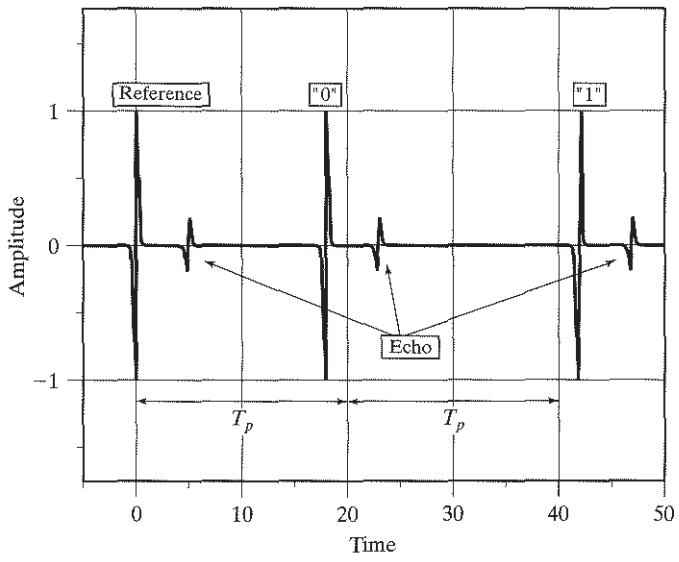


FIGURE 2.39 Illustration of multipath effects on ultra-wideband radio.

2.13 SUMMARY AND DISCUSSION[22]

In this chapter, we have looked at propagation and noise phenomena and their effect on a wireless communications link.

(a) Three principal *propagation modes* were identified:

- Free-space propagation, which corresponds to the ideal situation and a path loss proportional to the square of the distance between transmitter and receiver.
- Reflection, which is common in terrestrial propagation and shows significantly greater path loss than that of free-space propagation.
- Diffraction, which is also common in terrestrial propagation and explains why links may be closed in apparent electromagnetic shadows, albeit with some significant path losses.

(b) *Statistical propagation models* were then presented, always with the recognition that the computation of the exact physical behavior of a signal is too complex in most scenarios. Statistical models were discussed for

- Median-path loss that depends upon the range, with a path-loss exponent that is related to the environment in which propagation is occuring.
- Shadowing losses that are variations about the median-path loss due to the particular propagation path. These shadowing losses are often modeled with a log-normal distribution.
- Fast fading due to local reflections that quickly change their relative phase as the user terminal moves. Fast fading is often modeled with a Rayleigh distribution.

(c) *Noise* and *interference* were characterized as constraints on the performance of a communications system. Various forms of noise and their sources were examined, including

- Receiver noise that is due to the electronics of the receiver and the fundamental properties of circuits. This is ordinarily modeled as white noise.
- Antenna noise due to the natural radiation from the local environment. This noise contribution depends upon the antenna pattern, its height, and the environment.
- Artificial noise due to electrical machinery and discharges that produce harmonics or impulse noise in the radio frequencies used for communications. The effects of artificial noise decrease as the frequency of a signal increases.
- Multiple-access noise is due to other terminals accessing the same radio frequencies as those used by a signal, either in an adjacent channel or even in the same channel. Multiple-access noise is a system issue that relates the individual performance of a particular terminal to the number of user terminals that are allowed to access the system.

(d) The final topic of the chapter dealt with combining all of the previous elements in a *link budget* describing the communications link and its performance. The link budget depends on the following parameters:

- Propagation characteristics of the link, which can vary widely, depending upon the environment. The propagation determines the average signal power that the receive terminal can expect at its antenna port for a given transmit power.
- Noise characteristics determined by the receiver design and also by the environment in which the receiver operates.
- The SNR required by the modem to provide reliable communications. The combination of the received signal strength and the noise determines whether the communications link can be closed.

NOTES AND REFERENCES

[1] The additive white Gaussian model is often an excellent model for fixed satellite channels. More details on satellite channels and the factors involved in satellite link budgets can be found in Chapter 3 and 10 of Roddy (1996) and Chapter 6 of Spilker (1977).

[2] The derivation of the effective area of an isotropic antenna is provided in Chapter 2 of Stutzman (1998).

[3] There are various reciprocity theorems in electromagnetics, and a description of the theorems applicable to antennas may be found in Chapter 12 of Ramo (1968) and Chapter 9 of Stutzman(1998).

[4] The characteristics of parabolic antennas are discussed in Chapter 6 of Spilker (1977).

[5] More detailed analysis of the effects of ground-reflected waves and diffraction may be found in Chapters 2 and 3 of Parsons (1992).

[6] The characteristic impedance of free space, or wave impedance, is defined in analogy to the characteristic impedance of an infinite transmission line. The characteristic impedance of free space is the ratio of the total electric field to the total magnetic field in a plane perpendicular to the direction of propagation of a wave. For any such plane, the impedance is $\eta_0 = 120\pi$ ohms, as described in Chapter 6 of Ramo (1966) and Chapter 1 of Stutzman (1998).

[7] Further explanation of diffraction effects may be found in Chapter 3 of Rapport (1999) and in Chapter 3 of Parsons (1992). Further information on Fresnel integrals may be found in Chapter 7 of Abramowitz and Stegun (1964).

[8] A more advanced discussion of diffraction and Fresnel integrals may be found in Chapter 12 of Stutzman (1998).

[9] Mobile terrestrial channel propagation characteristics are discussed at a similar level to that of ours in the text by Rappaport (1999).

[10] The International Telecommunications Union (ITU) is an international body that makes recommendations regarding communications, both wired and wireless. The purpose of the recommendations is to promote common standards and interoperability among the communications systems of different nations. The ITU recommendations are published and are often a good source for propagation and interference models. ITU Recommendation P.1238-2 (2001) describes models for indoor propagation at various transmission frequencies. Detailed

information on indoor propagation models may be found in the papers by Saleh and Valenzuela (1987) and Cheung et al. (1998). The latter paper discusses extensions to the statistical model proposed in the text to more accurately model the effects of in-building diffraction.

[11] Further discussion of local propagation effects can be found in Chapter 1 of the classic book by Jakes (1974) that has been reprinted by the IEEE Press. Further information on Bessel functions may be found in Appendix B and in Chapter 9 of Abramowitz and Stegun (1964).

[12] Chapter 1 of Jakes (1974) provides a description of the Clarke model. The original presentation may be found in Clarke (1968).

[13] A more detailed description of WSSUS channels may be found in Chapter 7 of Proakis (1995). This book describes Bello's system functions (including the spaced-time spaced-frequency correlation and coherence spectrum functions) and the Fourier transform relationships between them. Similar material may be found in Chapter 6 of Parsons (1992).

[14] A more detailed description of the physics behind thermal noise may be found in Carlson (1975).

[15] More details on the analysis, characteristics, and measurement of artificial noise—particularly, impulse noise—can be found in Chapter 9 of Parsons (1992).

[16] The Okumura–Hata model for land-mobile propagation is often considered the "Bible" in this area. The original empirical investigations behind the model are described in the paper by Okumura et al. (1968), while the analysis and model fitting are described in the paper by Hata (1980).

[17] The Institute of Electrical and Electronic Engineers (IEEE) publishes a number of standards related to communications. IEEE Standard 802 was initially developed to specify the interchange of data on Ethernet networks. It has since been expanded manyfold to describe the interchange of data on many types of networks. IEEE Std. 802.11 (1999), in particular, describes the specifications for a number of wireless local area networks operating at different frequencies.

[18] This propagation model is part of ITU Recommendation P.1238-2 (2001) for indoor applications.

[19] The details of mathematically modeling impulse radio are described in the paper by Win and Scholtz (1998).

[20] Channel models for the ultra-wideband transmission are described in the paper by Cassioli et al (2002).

[21] Alternatively, one could use a device known as a Rake receiver to collect the energy in the echoes to improve performance. Rake receivers will be discussed in Section 5.6.

[22] Much of the material presented in this chapter can be found at a more advanced level in the book by Stüber (2001), which also considers in detail the effects on propagation of different modulation strategies for mobile radio.

ADDITIONAL PROBLEMS

Propagation

Problem 2.26 Consider a communications link with a geostationary satellite such that the transmitter–receiver separation is 40,000 km. Assume the same transmitter and receiver characteristics as described in Problem 2.2. What is the received power level in dBm? What implications does this power level have on the receiver design? With a land-mobile satellite

terminal, the typical antenna gain is 10 dB or less. What does this imply about the data rates that may be supported by such a link?

Problem 2.27 Consider a 10-watt transmitter communicating with a mobile receiver having a sensitivity of –100 dBm. Assume that the receiver antenna height is 2 m, and the transmitter and receiver antenna gains are 1 dB. What height of base station antenna would be necessary to provide a service area of radius 10 km? If the receiver is mobile, and the maximum radiated power is restricted by regulation to be 10 watts or less, what realistic options are there for increasing the service area?

Problem 2.28 In Problem 2.26, the satellite–receiver separation was 40,000 km. Assume the altitude of a geostationary satellite was said to be 36,000 km. What is the elevation angle from the receiver to the satellite in Problem 2.26? What was the increased path loss, in decibels, relative to a receiver when the satellite is in the *zenith* position (directly overhead)? If the transmitter–receiver separation in Problem 2.2 had been 20 km, what would the path loss have been? What can be said about comparing the dB path losses in satellite and terrestrial scenarios as a function of absolute distance?

Problem 2.29 Suppose that, by law, a service operator is not allowed to radiate more than 30 watts of power. From the plane-Earth model, what antenna height is required for a service radius of 1 km? 10 km? Assume that the receiver sensitivity is –100 dBm.

Problem 2.30 A satellite is in a geosynchronous orbit (i.e., an orbit in which the satellite appears to be fixed relative to the Earth). The satellite must be at an altitude of 36,000 km above the equator to achieve this synchrony. Such a satellite may have a global beam that illuminates all of the Earth in its view. What is the approximate 3-dB beamwidth of this global beam if the radius of the earth is 6400 km. What is the antenna gain?

Problem 2.31 A 100-m base station tower is located on a plateau as depicted in Fig. 2.40. In one directions there is a valley bounded by two (nonreflective) mountain ranges. If the transmit power is 10 watts, then, assuming ideal diffraction losses, what is the expected received signal strength. Assume free-space loss over the plains and a transmission frequency of 400 MHz.

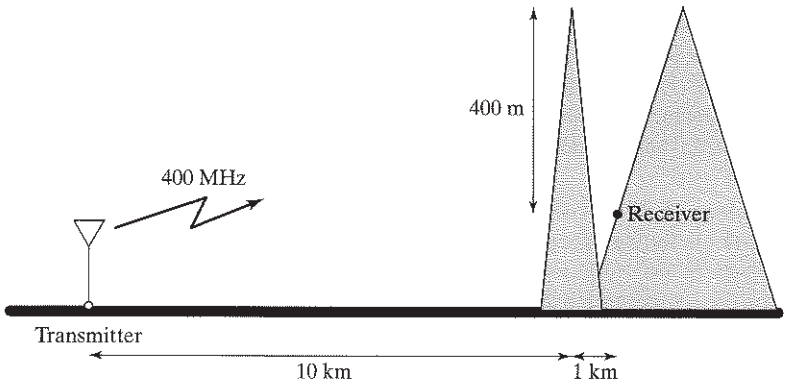


FIGURE 2.40 Site geometry for Problem 2.31.

Problem 2.32 In Problem 2.6, the estimates were based on median path loss. In the same city, the deviation about the median-path loss was estimated to be $\sigma_{dB} = 8$ dB. Assuming a log-normal model, how much additional power must be transmitted to cover the same service area with 90% *availability* at the edge of coverage when local shadowing is taken into account?

Problem 2.33 In Problem 2.7, we assume that the path-loss exponent depends on distance. For the first 10 m the path-loss exponent is 2, while beyond that distance this it is 3.5. What is the expected path loss in this case?

Problem 2.34 Suppose satellites are placed in geosynchronous orbit around the equator at 3° intervals. An Earth station with a 1-m parabolic antenna transmits 100 watts of power at 4 GHz toward the intended satellite. How much power is radiated toward the adjacent satellites on either side of the intended satellite? What implications does this amount of power have about low-power and high-power signals on adjacent satellites? What does it imply about the diameter of Earth station antennas?

Multipath

Problem 2.35 Prove that the Rayleigh density function is given by Eq. (2.55). (*Hint*: Let $Z_i = R\cos\theta$ and $Z_j = R\sin\theta$.)

Problem 2.36 Suppose the aircraft in Example 2.7 were accelerating at a rate of 0.25 g (i.e., approximately 2.5 m/s^2). What would the frequency slew rate be?

Problem 2.37 Suppose an aircraft is at an altitude of 4000 m in a circular holding pattern above an airport. The radius of the holding pattern is 4 km. If the aircraft speed is 400 km/hr, what is the frequency slew rate, df_r/dt (Hz/s), that an aircraft receiver must be capable of tracking. Assume that the transmitter is located at ground level at the center of the holding pattern. Aircraft safety communications use the VHF frequency band from 118 to 130 MHz.

Problem 2.38 A common device used in digital communications over fading channels is an *interleaver*, which is discussed in detail in Chapter 4. It boosts the performance of many forward error-correction codes. An interleaver takes a block of data from the encoder and permutes the order of the bits before transmitting them. At the receiver, the inverse operation, de-interleaving of the bits, is applied before the bits are decoded. The objective of the de-interleaver is to make the fading on adjacent bits (as seen by the decoder) appear uncorrelated. What is the ideal (minimum) spacing of bits in an interleaver to achieve this objective for a terminal moving at 100 km/hr and transmitting at 800 MHz?

Problem 2.39 From the development of Section 2.6, we could construct a model to simulate a Rayleigh fading process and the effects of fast fading. This model is illustrated in Fig. 2.41. The outputs of two independent white Gaussian noise generators are added in quadrature and then processed with a filter that approximates the desired fading spectrum. In practice, these operations are often done digitally. Since the fading process is frequently much slower than desired signalling rates, the spectrum-shaping filter may be very narrowband, requiring many interpolation stages to model accurately in the digital domain.

Develop a MATLAB program to generate a Rayleigh fast-fading signal.

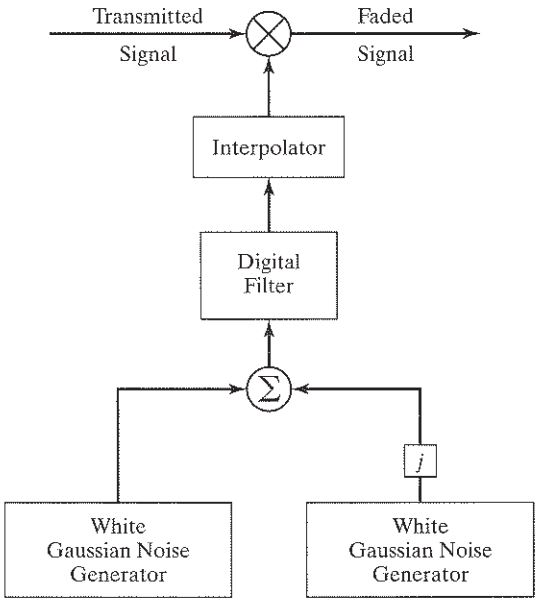


FIGURE 2.41 Illustration of digital model for simulating Rayleigh fading in Problem 2.39.

Problem 2.40 Generate a Rayleigh distribution with a 100-Hz fading bandwidth. A suggested approximation is to filter white Gaussian noise samples at 1 kHz with the one-pole filter $y_n = 0.7y_{n-1} + 0.3u_n$. For the output samples,

(a) plot the distribution of the output samples. Does this change with fading rate?

(b) plot the spectrum of the output samples. What is the 3-dB fading bandwidth?

(c) how much of the time is the signal –4dB or lower? How does this time compare with theoretical values?

(d) determine the average 3-dB fade duration. That is, each time the signal drops 3 dB below average, how long does it stay 3 dB or more below average?

Problem 2.41 Plot the frequency response of the channel $\tilde{h}(t) = \delta(t) - \alpha_2\delta(t - \tau)$ for $\alpha_2 = 0.3j$ and $\tau = 10$ microseconds. From the results obtained, what is the approximate maximum bandwidth over which the channel could be considered frequency flat?

Noise

Problem 2.42 Assume that a terrestrial radio receiver has a sensitivity of –100 dBm. What other information is necessary to compare this receiver to a satellite receiver that has a G/T of –22 dB/K?

Problem 2.43 Suppose the receiver in Problem 2.42 has a noise figure of 6 dB and an IF bandwidth of 30 kHz. What is its equivalent G/T ratio?

Problem 2.44 The sensitivity of a commercial GPS receiver is –130 dBm. If the receiver should function within 100 m of a 738–MHz base station transmitting 30 watts of power, how much attenuation of the second harmonic of the transmitted signal is required? Assume free-space loss between the transmitter and the GPS receiver, and that the GPS receiver requires a carrier-to-interference ratio, C/I, of 10 dB.

Problem 2.45 Suppose a laboratory spectrum analyzer has a noise figure of 25 dB. What N_0 would you expect to see when measuring a noiseless signal?

Problem 2.46 A satellite antenna is installed on the tail of an aircraft and connected to the receiver in an equipment bay located behind the cockpit. The antenna has a noise temperature of 50°K, and the transmission cable connecting the antenna to the low-noise amplifier (LNA) in the bay has a loss of 7 dB. The LNA has a gain of 60 dB and an equivalent noise temperature of 70°K. A secondary gain stage with 40 dB gain and a noise temperature of 1500°K follows. What is the noise temperature of the overall system? Where would be a better location for the LNA? What is the noise temperature of the new system?

Problem 2.47 A satellite is in a geosynchronous orbit at an altitude of 36,000 km. The satellite has three spot beams that illuminate approximately one-third of the visible Earth each. What is the approximate 3-dB beamwidth of the satellite spotbeam if the radius of the Earth is 6400 km. What is the antenna gain? Since the average temperature of the Earth is 290°K, what is the satellite G/T ratio?

Problem 2.48 A handheld radio has a small omnidirectional antenna. The first amplifier stage of the radio provides a 20-dB gain and has a noise temperature of 1200°K. The second amplifier stage provides another 20-dB gain and has a noise temperature of 3500°K. What is the combined noise figure of the antenna and first two stages of the radio? If the baseband processing of the radio requires an SNR of 9 dB in 5 kHz, what is the receiver sensitivity?

Problem 2.49 When a directional antenna is pointing toward the empty sky, the noise temperature falls to about 3°K at frequencies between 1 GHz and 10 GHz. The 3°K temperature represents the residual background radiation of the universe. Cosmological theory suggests that it is a remnant of the initial Big Bang. If the antenna is connected directly to a low-noise amplifier with a gain of 60 dB and a noise temperature of 100°K, what is the system noise temperature? What is the equivalent noise figure for the system?

Problem 2.50 Prove that, with a 1-in-N reuse pattern, the number of cochannel cells at the minimum distance is six.

Problem 2.51 Consider a dense urban environment in which the path-loss exponent is 4.5 and the required C/I is 13 dB. What reuse factor is permissible? If the cell size is reduced to increase capacity, what other changes would also be required?

Problem 2.52 Consider the previous problem, 2.51, but in a rolling rural environment with a path-loss exponent of 3. What reuse factor is permissible? What are the disadvantages and advantages related to the path-loss exponent?

Link Budgets

Problem 2.53 Suppose that a particular 100-kbps service requires a bit error rate (BER) of 10^{-5} in AWGN. Could this service be supported by the link budget described in Table 2.3? Justify your answer.

Problem 2.54 What is the power flux density on the surface of the Earth for Example 2.17, assuming that the satellite delivers 50 watts to the antenna terminals. How much power is collected by a 1.8-meter antenna with 50% efficiency? What voltage level would be generated across a 50-ohm resistor by that amount of power? What is the equivalent rms thermal noise voltage of this resistor in a 10-kHz bandwidth?

Problem 2.55 Construct a link budget for the return link from mobile to base of a commercial cell phone that transmits at most 600 milliwatts. Find the range of the service for the propagation models shown in Table 2.6, assuming that a service availability of 95% is required. Assume that the base station receiver noise figure is 3 dB and the required SNR is 8 dB for a 9.6-kbps service.

How do these results change if the required availability is 90%? What additional things could be done at the base station to help close the return link at greater distances?

Problem 2.56 Consider the design of a radio-controlled model airplane with a maximum range of 300 m. The receiver requires a C/N_0 ratio of 47 dB-Hz. Due to poor isolation from the aircraft engine, the receiver has a noise figure of 22 dB. What EIRP would have to be transmitted to achieve the maximum range? Assume line-of-sight transmission at 45 MHz, and assume that transmit and receive antennas have gains of –3 dB relative to an isotropic antenna.

Problem 2.57 A typical 1.5-volt "AA" cell has 1.5 ampere-hours (Ah) of energy. Assuming that a 3-V transmitter in the model airplane of Problem 2.56 has a 50% efficiency in converting input energy into EIRP, how long would a pair of batteries last? How long would they last if the transmitter had a 10% duty cycle?

Problem 2.58 You are a delegate to an ITU-R meeting representing your country. Your country's objective is to have the frequency band from 1910 to 1930 MHz designated worldwide for personal communication services (PCS). You have prepared submissions detailing the suitability of this band for this application, bearing in mind propagation characteristics, technology, and the need for the designation of this band, based on commercial projections of growth in the PCS market. The proposal has support from a number of participants at the meeting, but a number of countries that have existing fixed microwave services operating in the band from 1910 to 1920 MHz object to the proposal. Suggest a compromise proposal that may achieve your long-term objectives.

TABLE 2.6 Propagation models for Problem 2.55.

n	β(dB)	σ_{dB}
2	0	8
3	20	8
3.5	10	10
4	6	12

Problem 2.59 When G. Marconi made the first radio transmission in 1899 across the Atlantic Ocean, he used all of the spectrum available worldwide to transmit a few bits per second. It has been suggested that, in the period since then, spectrum usage (bits/s/Hz worldwide) has increased by a factor of a million. List the factors that have resulted in this substantial increase. Which factor will likely result in the largest increase in the future?

Problem 2.60 In Example 2.23, to what would the transmit power need to be reduced in order to produce an interference level equivalent to the noise floor of the 800 MHz receiver? If the transmit power remains at one milliwatt, what is the minimum separation of the impulse radio from the 800 MHz radio to produce equivalent interference and noise floor densities? Assume free-space propagation conditions.

CHAPTER 3

Modulation and Frequency-Division Multiple Access

3.1 INTRODUCTION

As mentioned in Chapter 1, *multiple access* is a technique that permits the communication resources of a wireless channel to be shared by a large number of users in different locations. There are four basic forms of multiple access applied to wireless communications, depending on which particular resource is exploited:

1. Frequency-division multiple access (FDMA)
2. Time-division multiple access (TDMA)
3. Code-division multiple access (CDMA)
4. Space-division multiple access (SDMA)

In a way, this list also orders the evolution of the spectral efficiency of wireless communication systems through the years, with the motivation for change being driven by improved utilization of the available spectrum through the use of increasingly sophisticated modulation and coding techniques. In this chapter, the focus is on FDMA.

As the name implies, FDMA operates by dividing the bandwidth of a wireless channel equally among a number of users wanting to access the channel. Specifically, FDMA can be visualized as shown in Fig. 1.2, reproduced here as Fig. 3.1 for convenience of presentation. To facilitate the kind of frequency division portrayed in this chapter, we resort to the use of *modulation*, which is a process of transforming the frequency content of a particular user's information-bearing signal so as to lie inside the frequency band allotted to that user. The block diagram of Fig. 3.2 presents the basic functional blocks that constitute the transmitter and receiver of the wireless communication system serving a single user; the figure also includes the idealized power spectra at different points in the system. The baseband processor performs operations such as filtering on the speech signal and thereby prepares it for modulation. These operations are followed by transmission over the wireless channel via a pair of antennas, one at the transmitter and the other at the receiver. The receiver performs inverse operations on the received signal, with the objective of delivering an estimate of the speech signal to a user.

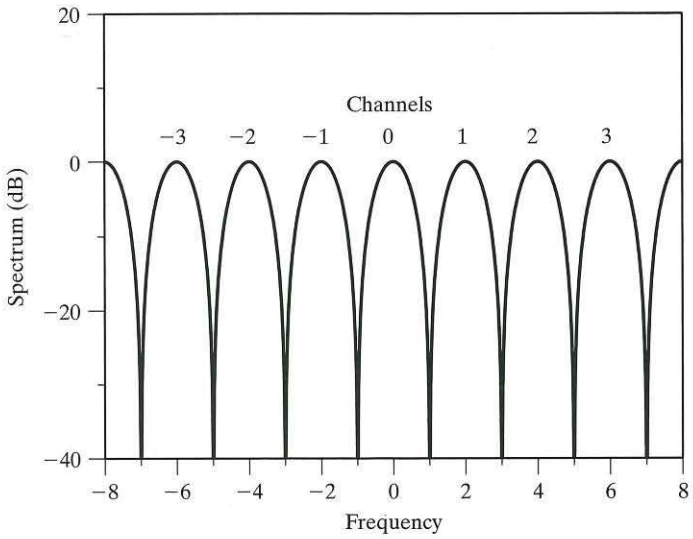


FIGURE 3.1 A frequency-domain representation of FDMA.

Section 3.2 reviews the essentials of the modulation process and the different approaches to impressing an information-bearing signal onto a carrier. A discussion of linear modulation techniques follows in Section 3.3, with emphasis on binary data transmission. Section 3.4 describes the raised cosine (RC) spectrum for pulse shaping so as to mitigate the so-called intersymbol interference problem. The material presented sets the stage for the complex representation of linear modulated signals and linear bandpass systems in Section 3.5. Section 3.6 discusses issues relating to the geometric representation of digitally modulated signals. Nonlinear modulation techniques are examined in Section 3.7, again with emphasis on binary data transmission. Section 3.8 presents two practical issues, adjacent channel interference and nonlinearity,

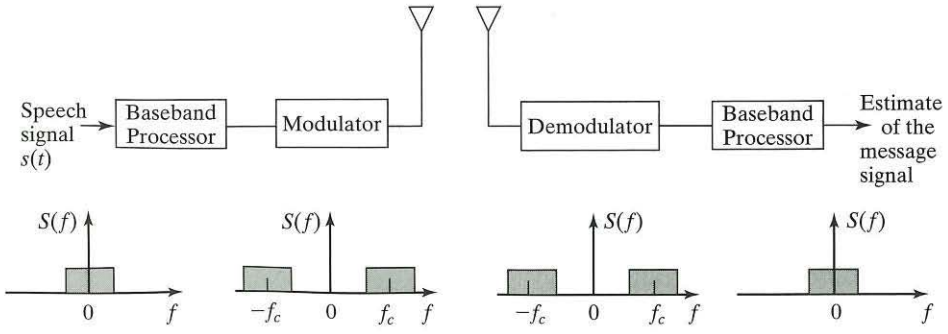


FIGURE 3.2 Illustration of the bandpass characteristic of a wireless system.

that are of practical concern in wireless communications. With this material at hand, the stage is set for a comparison of modulation techniques in the context of wireless communications, which is presented in Section 3.9. Section 3.10 takes up the issue of channel estimation and tracking at the receiver. The discussion of modulation techniques is completed in Section 3.11 with a look at the noise performance of digital modulation schemes.

Section 3.12 discusses the frequency-division multiple access technique. As with other chapters in the book, Sections 3.13 and 3.14 discuss two theme examples. The first, on orthogonal frequency-division multiplexing (OFDM), builds on a computationally efficient algorithm, namely, the fast Fourier transform, that is widely used in digital signal processing. The second theme example discusses the ubiquitous cordless telephone.

3.2 MODULATION

Modulation is formally defined as *the process by which some characteristic of a carrier wave is varied in accordance with an information-bearing signal.* In this context, the information-bearing signal is referred to as the *modulating signal*, and the output of the modulation process is referred to as the *modulated signal.* The device that performs the modulation process in the transmitter is referred to as a *modulator*, and the device used to recover the information-bearing signal in the receiver is referred to as a *demodulator.*

We can identify three practical benefits that result from the use of modulation in a wireless communication system:

1. *Modulation is used to shift the spectral content of a message signal so that it lies inside the operating frequency band of the wireless communication channel.*
   Consider, for example, telephonic communication over a cellular radio channel. In such an application, the frequency components of a speech signal from about 300 to 3100 Hz are considered adequate for the purpose of communication. In North America, one band of frequencies assigned to cellular radio systems is 800–900 MHz. The subband of 824–849 MHz is used to receive signals from mobile users, and the subband of 869–894 MHz is used for transmitting signals to mobile users. For this form of telephonic communication to be feasible, we clearly need to do two things: shift the essential spectral content of a speech signal so that it lies inside the prescribed subband for transmission, and shift it back to its original frequency band on reception. The first of these two operations is one of modulation, and the second is one of demodulation.
2. *Modulation provides a mechanism for putting the information content of a message signal into a form that may be less vulnerable to noise or interference.*
   In a wireless communication system, the received signal is ordinarily corrupted by noise generated at the front end of the receiver or by interference picked up in the course of transmission. Some specific forms of modulation (e.g., frequency modulation, considered in Section 3.4) have the inherent ability to trade