WIRELESS COMMUNICATIONS AND NETWORKS

SECOND EDITION

William Stallings

PEARSON
Prentice Hall

Upper Saddle River, NJ 07458

Library of Congress Cataloging-in-Publication Data on file

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Publisher: *Alan Apt*
Associate Editor: *Toni D. Holm*
Editorial Assistant: *Patrick Lindner*
Vice President and Director of Production and Manufacturing, ESM: *David W. Riccardi*
Executive Managing Editor: *Vince O'Brien*
Managing Editor: *Camille Trentacoste*
Production Editor: *Rose Keman*
Cover Photo: *Brand X Pictures*
Director of Creative Services: *Paul Belfanti*
Art Director: *Heather Scott*
Cover Designer: *Tamara Newnam*
Managing Editor, AV Management and Production: *Patricia Burns*
Art Editor: *Gregory Dulles*
Manufacturing Manager: *Trudy Pisciotti*
Manufacturing Buyer: *Lynda Castillo*
Marketing Manager: *Pamela Hersperger*
Marketing Assistant: *Barrie Reinhold*

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: 0-13-191835-4

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*

As always,
for my loving wife
and her constant companions
Geoffroi and Helma

CONTENTS

PREFACE

OBJECTIVES

Wireless technology has become the most exciting area in telecommunications and networking. The rapid growth of mobile telephone use, various satellite services, and now the wireless Internet and wireless LANs are generating tremendous changes in telecommunications and networking. This book explores the key topics in the field in the following general categories:

- **Technology and architecture:** There is a small collection of ingredients that serves to characterize and differentiate wireless communication and networking, including frequency band, signal encoding technique, error correction technique, and network architecture.
- **Network type:** This book covers the important types of wireless networks, including satellite, cellular, fixed wireless access, and wireless LANs.
- **Design approaches:** The book examines alternative design choices and assesses their relative merits.
- **Applications:** A number of key technologies and applications have been developed on top of wireless infrastructures, especially mobile IP and wireless Web access.

Throughout, there is an emphasis on both technology and on standards. The book provides a comprehensive guide to understanding specific wireless standards, such as those promulgated by ITU and IEEE 802, as well as standards developed by other organizations. This emphasis reflects the importance of such standards in defining the available products and future research directions in this field.

INTENDED AUDIENCE

This book is intended for a broad range of readers who will benefit from an understanding of wireless communications and networks, and the associated technologies. This includes students and professionals in the fields of data processing and data communications, designers and implementers, and data communication and networking customers and managers. For the professional interested in this field, the book serves as a basic reference volume and is suitable for self-study.

As a textbook, it is suitable for an advanced undergraduate or graduate course. It covers the material in the CS332 Wireless and Mobile Computing advanced course of the joint ACM/IEEE Computing Curricula 2001. The chapters and parts of the book are sufficiently modular to provide a great deal of flexibility in the design of courses.

PLAN OF THE BOOK

The book treats a number of advanced topics and provides a brief survey of the required elementary topics. For the reader with little or no background in data communications, Part One and the appendices cover a number of basic topics. The book is divided into four parts:

- Technical Background
- Wireless Communication Technology

- Wireless Networking
- Wireless LANs

In addition, the book includes an extensive glossary, a list of frequently used acronyms, and a bibliography. Each chapter includes problems, suggestions for further reading, and a list of relevant Web sites. Each chapter also includes, for review, a list of key words and a number of review questions.

INTERNET SERVICES FOR INSTRUCTORS AND STUDENTS

There is a Web site for this book that provides support for students and instructors. The site includes links to other relevant sites, transparency masters of figures and tables from the book in PDF (Adobe Acrobat) format, PowerPoint slides, and sign-up information for the book's Internet mailing list. The Web page is at WilliamStallings.com/Wireless/Wireless2e.html; see Section 1.8 for more information. An Internet mailing list has been set up so that instructors using this book can exchange information, suggestions, and questions with each other and with the author. As soon as typos or other errors are discovered, an errata list for this book will be available at WilliamStallings.com. I also maintain the Computer Science Student Resource Site at WilliamStallings.com/StudentSupport.html.

WHAT'S NEW IN THE SECOND EDITION

In the three years since the first edition of this book was published, the field has seen continued innovations and improvements. In this new edition, I try to capture these changes while maintaining a broad and comprehensive coverage of the entire field. To begin the process of revision, the first edition of this book was extensively reviewed by a number of professors who teach the subject. The result is that, in many places, the narrative has been clarified and tightened, and illustrations have been improved. Also, a number of new "field-tested" problems have been added.

Beyond these refinements to improve pedagogy and user friendliness, the technical content of the book has been updated throughout, to reflect the ongoing changes in this exciting field. Every chapter has been revised. Highlights include the following:

- **Minimum shift keying:** MSK is a form of modulation that is found in some mobile communications systems. This material is now covered.
- **CDMA2000:** The first 3G (third generation) wireless system to be deployed commercially is known as CDMA2000 1xEV-DO. A discussion of this important standard is included.
- **WiMAX and IEEE 802.16a:** Work on wireless local loop has evolved, including the introduction of the WiMAX specification to provide interoperability specifications for 802.16. Chapter 11 includes new material on 802.16, including the recent 802.16a standard.
- **Orthogonal frequency division multiplexing:** The popularity of OFDM is increasing and is used in a variety of local and wide area wireless standards. The material on OFDM has been updated and expanded.
- **Wi-Fi and IEEE 802.11:** The coverage of 802.11a and 802.11b has been expanded significantly, and treatment of 802.11g had been added.
- **Data scrambling:** Scrambling is a technique often used to improve signal quality. An overview of data scrambling is provided in Chapter 14.

- **Wi-Fi protected access:** WPA has replaced Wireless Equivalent Privacy (WEP) as the specification for providing security in wireless LANs. Chapter 14 provides coverage of WPA.
- **IEEE 802.15 and personal area networks:** The initial 802.15.1 standard provides an official specification for Bluetooth, which was covered in the first edition as well as this edition. This edition also covers two new standards: the 802.15.3 high-speed wireless PAN standard and the 802.15.4 low-speed wireless PAN standard.
- **Trellis-coded modulation:** TCM is a technique that provides for efficient use of bandlimited channels; it is described in Chapter 15.

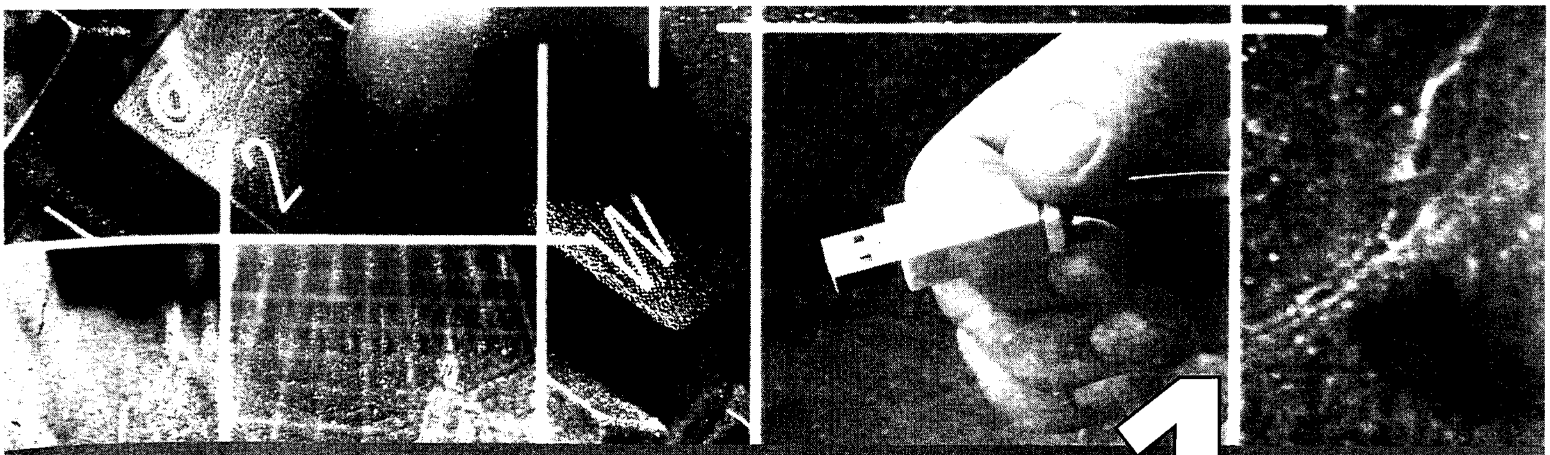
In addition, throughout the book, virtually every topic has been updated to reflect the developments in standards and technology that have occurred since the publication of the first edition.

ACKNOWLEDGMENTS

This new edition has benefited from review by a number of people, who gave generously of their time and expertise. The following people reviewed all or a large part of the manuscript: Dr. Albert Cheng (University of Houston-University Park), Dale W. Callahan (University of Alabama, Birmingham), Ravi Sankar (University of South Florida, Tampa), Pei Zheng (Arcadia University, Pennsylvania), and Anne Cox (Austin Community College, Texas).

Thanks also to the many people who provided detailed technical reviews of a single chapter: Lars Poulsen, Howard Eisenhauer, D. E. Jennings, Paul Robichaux, John Adams, Jerry Huang, Andreas Kasenides, Munira Ahmed, Hossein Izadpanah, Aaron King, Benoit d'Udekem, Marco Casole, Kevin Peterson, Dinesh Lal Pradhan, and Cathal Mc Daid.

Finally, I would like to thank the many people responsible for the publication of the book, all of whom did their usual excellent job. This includes the staff at Prentice Hall, particularly my editors Alan Apt and Toni Holm; their assistant Patrick Lindner: production manager Rose Kernan; and supplements manager Sarah Parker. Also, Jake Warde of Warde Publishers managed the reviews; and Patricia M. Daly did the copy editing.



CHAPTER 1

INTRODUCTION

This book is a survey of the technology of wireless communications and networks. Many factors, including increased competition and the introduction of digital technology, have led to unprecedented growth in the wireless market. In this chapter, we discuss some of the key factors driving this new telecommunications revolution.

This book, and the accompanying Web site, covers a lot of material. Following the general discussion, this chapter gives the reader an overview of the book.

1.1 WIRELESS COMES OF AGE

Guglielmo Marconi invented the wireless telegraph in 1896.[1] In 1901, he sent telegraphic signals across the Atlantic Ocean from Cornwall to St. John's Newfoundland; a distance of about 3200 km. His invention allowed two parties to communicate by sending each other alphanumeric characters encoded in an analog signal. Over the last century, advances in wireless technologies have led to the radio, the television, the mobile telephone, and communications satellites. All types of information can now be sent to almost every corner of the world. Recently, a great deal of attention has been focused on satellite communications, wireless networking, and cellular technology.

Communications satellites were first launched in the 1960s. Those first satellites could only handle 240 voice circuits. Today, satellites carry about one-third of the voice traffic and all of the television signals between countries [EVAN98]. Modern satellites typically introduce a quarter-second propagation delay to the signals they handle. Newer satellites in lower orbits, with less inherent signal delay, have been deployed to provide data services such as Internet access.

Wireless networking is allowing businesses to develop WANs, MANs, and LANs without a cable plant. The IEEE has developed 802.11 as a standard for wireless LANs. The Bluetooth industry consortium is also working to provide a seamless wireless networking technology.

The cellular or mobile telephone is the modern equivalent of Marconi's wireless telegraph, offering two-party, two-way communication. The first-generation wireless phones used analog technology. These devices were heavy and coverage was patchy, but they successfully demonstrated the inherent convenience of mobile communications. The current generation of wireless devices is built using digital technology. Digital networks carry much more traffic and provide better reception and security than analog networks. In addition, digital technology has made possible value-added services such as caller identification. Newer wireless devices connect to the Internet using frequency ranges that support higher information rates.

The impact of wireless communications has been and will continue to be profound. Very few inventions have been able to "shrink" the world in such a manner. The standards that define how wireless communication devices interact are quickly

[1] The actual invention of radio communications more properly should be attributed to Nikola Tesla, who gave a public demonstration in 1893. Marconi's patents were overturned in favor of Tesla in 1943 [ENGE00].

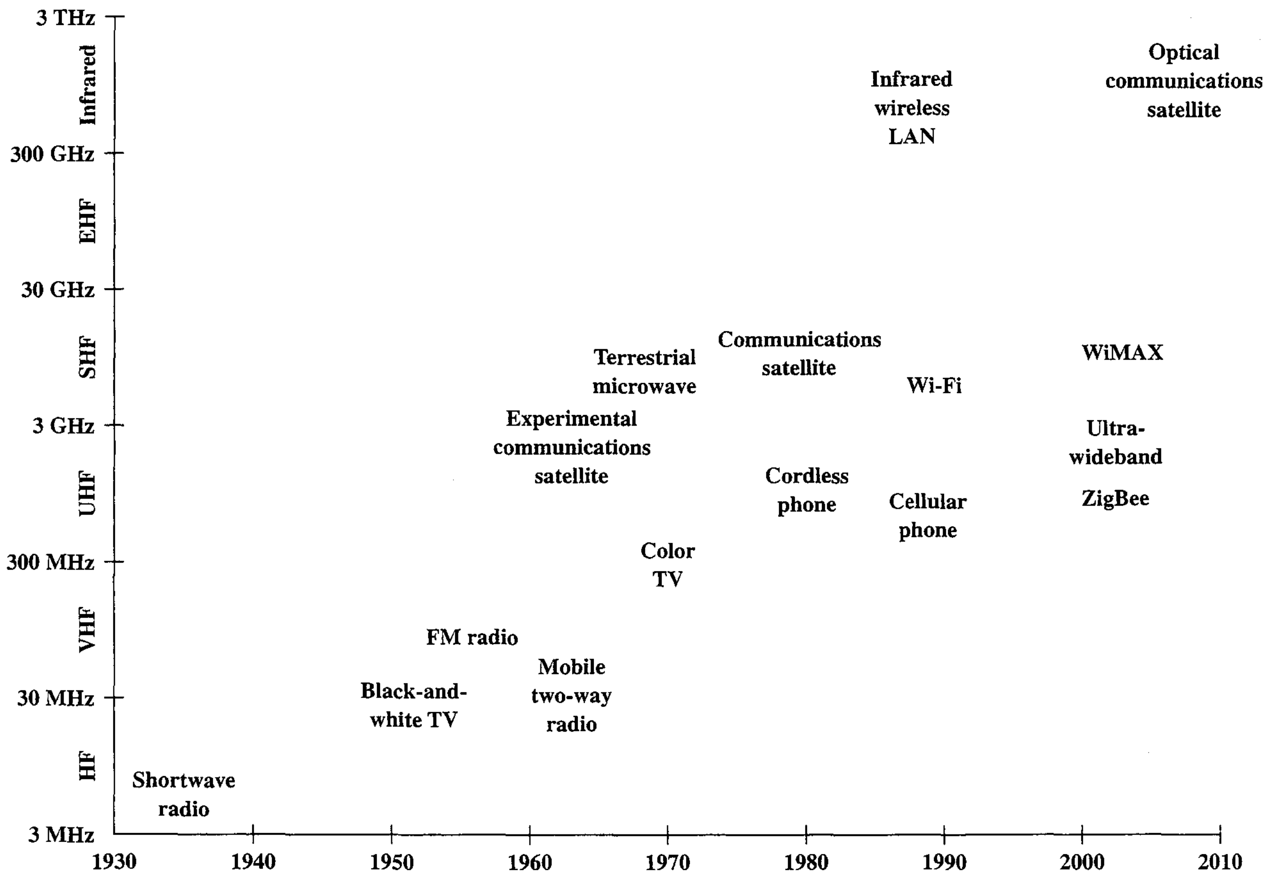


Figure 1.1 Some Milestones in Wireless Communications

converging and soon will allow the creation of a global wireless network that will deliver a wide variety of services.

Figure 1.1 highlights some of the key milestones in the development of wireless communications.[2] Wireless technologies have gradually migrated to higher frequencies. As will be seen in later chapters, higher frequencies enable the support of greater data rates and throughput.

1.2 THE CELLULAR REVOLUTION

The cellular revolution is apparent in the growth of the mobile phone market alone. In 1990, the number of users was approximately 11 million [ECON99]. Today, that number is in the billions. According to the ITU (International Telecommunications Union),[3] the number of mobile phones worldwide outnumbered fixed-line phones for the first time in 2002. The newer generation devices, with access to the Internet and built-in digital cameras, add to this momentum. There are a number of reasons

[2]Note the use of a log scale for the *y*-axis. A basic review of log scales is in the math refresher document at the Computer Science Student Resource Site at **WilliamStallings.com/StudentSupport.html.**

[3]A description of ITU and other standards-making bodies is contained in a supporting document at this book's Web site.

for the increasing dominance of mobile phones. Mobile phones are convenient; they move with people. In addition, by their nature, they are location aware. A mobile phone communicates with regional base stations that are at fixed locations.

Technical innovations have contributed to the success of mobile phones. The handsets have become smaller and lighter, battery life has increased, and digital technology has improved reception and allowed better use of a finite spectrum. As with many types of digital equipment, the costs associated with mobile telephones have been decreasing. In areas where competition flourishes, prices have dropped dramatically since 1996.

In many geographic areas, mobile telephones are the only economical way to provide phone service to the population. Operators can erect base stations quickly and inexpensively when compared with digging up ground to lay copper in harsh terrain.

Mobile telephones are only the tip of the cellular revolution. Increasingly, new types of wireless devices are being introduced. These new devices have access to the Internet. They include personal organizers and telephones, but now they have Web access, instant messaging, e-mail, and other services available on the Internet. Wireless devices in automobiles allow users to download maps and directions on demand. Soon, the devices may be able to call for help when an accident has occurred or perhaps notify the user of the lowest-priced fuel in the immediate area. Other conveniences will be available as well. For example, refrigerators may one day be able to order groceries over the Internet to replace consumed items.

The first rush to wireless was for voice. Now, the attention is on data. A big part of this market is the "wireless" Internet. Wireless users use the Internet differently than fixed users. Wireless devices have limited displays and input capabilities compared with typical fixed devices such as the PC. Transactions and messaging will be the rule instead of lengthy browsing sessions. Because wireless devices are location aware, information can be tailored to the geographic location of the user. Information will be able to find users, instead of users searching for information.

1.3 THE GLOBAL CELLULAR NETWORK

Today there is no single cellular network. Devices support one or two of a myriad of technologies and generally work only within the confines of a single operator's network. To move beyond this model, more work must be done to define and implement standards.

The ITU is working to develop a family of standards for the next-generation wireless devices. The new standards will use higher frequencies to increase capacity. The new standards will also help overcome the incompatibilities introduced as the different first- and second-generation networks were developed and deployed over the last decade.

The dominant first-generation digital wireless network in North America was the Advanced Mobile Phone System (AMPS). This network offers a data service using the Cellular Digital Packet Data (CDPD) overlay network, which provides a 19.2-kbps data rate. The CPDP uses idle periods on regular voice channels to provide the data service.

The key second-generation wireless systems are the Global System for Mobile Communications (GSM), Personal Communications Service (PCS) IS-136, and PCS

IS-95. The PCS standard IS-136 uses time division multiple access (TDMA) while IS-95 uses code division multiple access (CDMA). The GSM and PCS IS-136 use dedicated channels at 9.6 kbps to deliver the data service.

The ITU is developing International Mobile Telecommunications-2000 (IMT-2000). This family of standards is intended to provide a seamless global network. The standards are being developed around the 2-GHz frequency band. The new standards and frequency band will provide data rates up to 2 Mbps.

In addition to defining frequency usage, encoding techniques, and transmission, standards also need to define how mobile devices will interact with the Internet. Several standards bodies and industry consortiums are working to that end. The Wireless Application Protocol (WAP) Forum is developing a common protocol that allows devices with limited display and input capabilities to access the Internet. The Internet Engineering Task Force (IETF) is developing a mobile IP standard that adapts the ubiquitous IP protocol to work within a mobile environment.

1.4 BROADBAND

The Internet is increasingly a multimedia experience. Graphics, video, and audio abound on the pages of the World Wide Web. Business communications are following the same trend. For example, e-mail frequently includes large multimedia attachments. In order to participate fully, wireless networks require the same high data rates as their fixed counterparts. The higher data rates are obtainable with broadband wireless technology.

Broadband wireless service shares the same advantages of all wireless services: convenience and reduced cost. Operators can deploy the service faster than a fixed service and without the cost of a cable plant. The service is also mobile and can be deployed almost anywhere.

There are many initiatives developing broadband wireless standards around many different applications. The standards cover everything from the wireless LAN to the small wireless home network. Data rates vary from 2 Mbps to well over 100 Mbps. Many of these technologies are available now and many more will become available in the next several years.

Wireless LANs (WLANs) provide network services where it is difficult or too expensive to deploy a fixed infrastructure. The primary WLAN standard is IEEE 802.11, which provides for data rates as high as 54 Mbps.

A potential problem with 802.11 is compatibility with Bluetooth. Bluetooth is a wireless networking specification that defines wireless communications between devices such as laptops, PDAs, and mobile phones. Bluetooth and some versions of 802.11 use the same frequency band. The technologies would most likely interfere with each other if deployed in the same device.

1.5 FUTURE TRENDS

Much of the development effort in new wireless technology makes use of portions of the frequency spectrum that do not, in many countries, require licensing. In the United States, two such frequency bands are Industrial, Scientific, and Medical

(ISM) band near 2.4 GHz and the newly allocated unlicensed radio band, the Unlicensed National Information Infrastructure (UNII) band. UNII was created by the FCC (Federal Communications Commission) to allow manufacturers to develop high-speed wireless networks. In order to find enough bandwidth to satisfy needs, the band was established at 5 GHz, making it incompatible with 2.4-GHz equipment. The free, unlicensed portions of the radio spectrum enable manufacturers to avoid billions of dollars in licensing fees.

For years, these radio frequencies were neglected, the lonely domain of cordless phones and microwave ovens. In recent years however, spurred by consumer demand and active standards bodies, considerable research and development is underway. The first significant fruit of this activity is **Wi-Fi** (Wireless Fidelity), the very popular wireless LAN technology based on the IEEE 802.11 standards. In essence, Wi-Fi refers to 802.11-compatible products that have been certified as interoperable by the Wi-Fi Alliance, a body specifically set up for this certification process. Wi-Fi covers not only office-based LANs, but also home-based LANs and publicly available *hot spots*, which are areas around a central antenna in which people can wirelessly share information or connect to the Internet with a properly equipped laptop. Wi-Fi is examined in some detail in Chapter 14.

Wi-Fi is just the first major step in utilizing these bands. Four other innovative technologies are working their way through the research, development, and standardization efforts: WiMAX, Mobile-Fi, ZigBee, and Ultrawideband. We survey these technologies briefly in this section.

WiMAX is similar to Wi-Fi. Both create hot spots, but while Wi-Fi can cover several hundred meters, WiMAX has a range of 40 to 50 km. Thus, WiMAX provides a wireless alternative to cable, DSL, and T1/E1 for last-mile broadband access. It will also be used as complimentary technology to connect 802.11 hot spots to the Internet. Initial deployments of WiMAX are in fixed locations, but a mobile version is under development. WiMAX is an interoperability specification based on IEEE 802.16 and is discussed in more detail in Chapter 11.

Mobile-Fi is similar to the mobile version of WiMAX in terms of technology. The objective with Mobile-Fi is to provide Internet access to mobile users at data rates even higher than those available in today's home broadband links. In this context, mobile truly means mobile, not just movable. Thus, a Mobile-Fi user could enjoy broadband Internet access while traveling in a moving car or train. Mobile-Fi is based on the IEEE 802.20 specifications.

ZigBee functions at a relatively low data rate over relatively short distances, compared to Wi-Fi. The objective is to develop products that are very low cost, with low power consumption and low data rate. ZigBee technology enables the coordination of communication among thousands of tiny sensors, which can be scattered throughout offices, farms, or factories, picking up bits of information about temperature, chemicals, water, or motion. They're designed to use little energy because they'll be left in place for 5 or 10 years and their batteries need to last. ZigBee devices communicate efficiently, passing data over radio waves from one to the other like a bucket brigade. At the end of the line the data can be dropped into a computer for analysis or picked up by another wireless technology like Wi-Fi or WiMAX.

Ultrawideband serves a very different purpose than the other technologies mentioned in this section. Ultrawideband enables the movement of massive files at high data rates over short distances. For example, in the home, Ultrawideband would allow the user to transfer hours of video from a PC to a TV without any messy cords. On the road, a passenger who has a laptop in the trunk receiving data over Mobile-Fi could use Ultrawideband to pull that information up to a handheld computer in the front seat.

1.6 THE TROUBLE WITH WIRELESS

Wireless is convenient and often less expensive to deploy than fixed services, but wireless is not perfect. There are limitations, political and technical difficulties that may ultimately prevent wireless technologies from reaching their full potential. Two issues are incompatible standards and device limitations.

As mentioned previously, in North America there are two standards for digital cellular service. Internationally, there is at least one more. A device using PCS IS-136 will not work in an area where the deployed technology is PCS IS-95. Also mentioned previously is the inability to use Bluetooth and 802.11 in the same device. These are just two examples of problems that arise when industrywide standards do not exist. The lack of an industrywide standard holds the technologies back from delivering one of the true ideals of wireless: ubiquitous access to data.

Device limitations also restrict the free flow of data. The small display on a mobile telephone is inadequate for displaying more than a few lines of text. In addition, most mobile wireless devices cannot access the vast majority of WWW sites on the Internet. The browsers use a special language, wireless markup language (WML), instead of the de facto standard HTML.

Most likely, no one wireless device will be able to meet every need. The potential of wireless can be met but not with a single product. Wireless will succeed because it will be integrated into a variety of devices that can meet a variety of needs.

1.7 OUTLINE OF THE BOOK

The objective of this book is to provide a comprehensive technical survey of wireless communications fundamentals, wireless networks, and wireless applications. The book is organized into four parts (Figure 1.2). The reader who is already familiar with data communications and networking technology can safely skip or just skim Part One. Part Two discusses underlying principles common to all of the material covered in the remainder of the book and should be read next. Parts Three and Four are independent and may be covered in either order. Within Part Three, all of the chapters are more or less independent and can be read in any order depending on your level of interest. The same is true of Chapters 14 and 15 in Part Five.

Part One: Background

Part One provides a preview and context for the remainder of the book, covering basic topics in data communications as well as TCP/IP. Part One, together with the appendices at the end of the book, is intended to make the book as self-contained as possible.

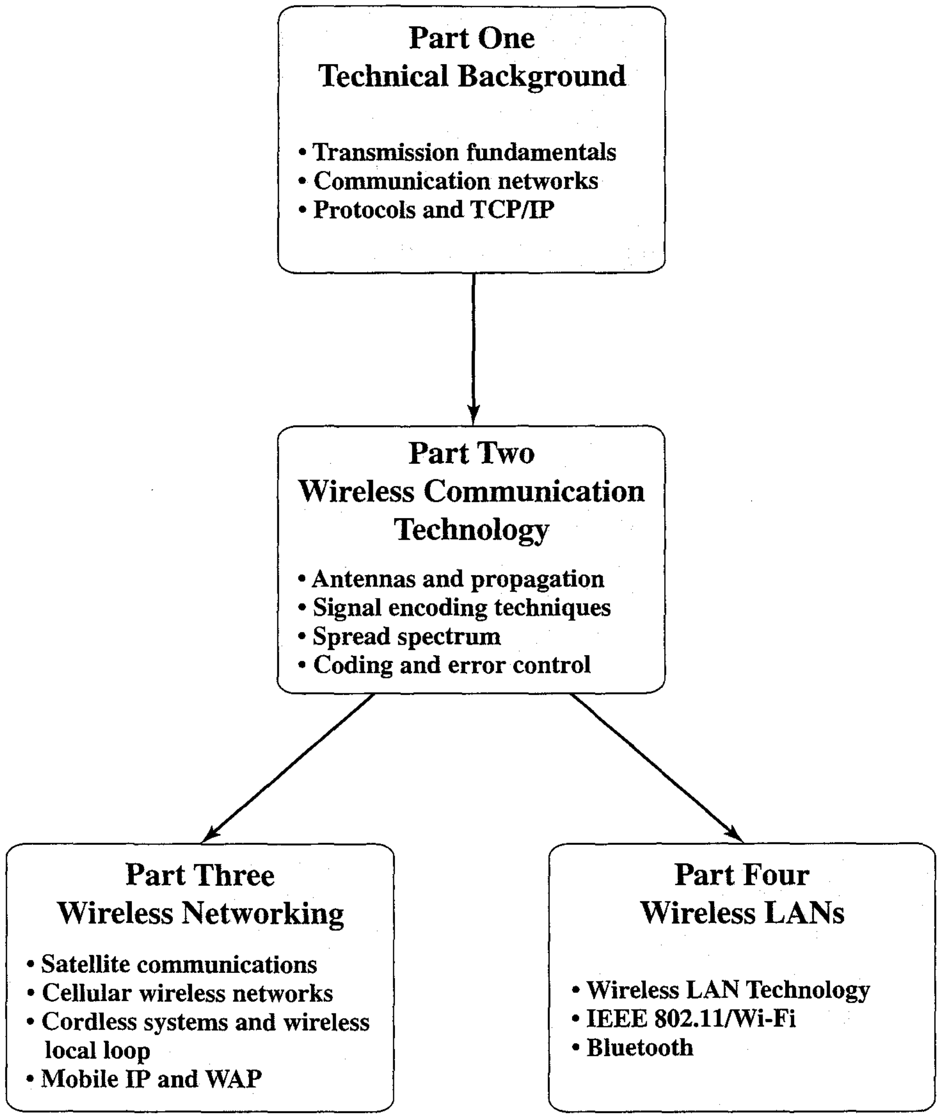


Figure 1.2 Wireless Topics

Chapter 2: Transmission Fundamentals Chapter 2 provides a basic overview of transmission topics. The chapter begins with a look at some data communications concepts, including signaling techniques and analog and digital data transmission. The chapter then covers channel capacity, transmission media, and the concept of multiplexing.

Chapter 3: Communication Networks This chapter provides an overview and comparison of basic communication network technologies, including circuit switching, packet switching, and ATM.

Chapter 4: Protocols and the TCP/IP Protocol Suite Data network communication and distributed applications rely on underlying communications software that is independent of application and relieves the application of much of the burden of reliably exchanging data. This communications software is organized into a protocol architecture, the most important incarnation of which is the TCP/IP protocol suite. Chapter 4 introduces the concept of a protocol architecture and provides an overview of TCP/IP. Another architecture, the Open Systems Interconnection (OSI) reference model, is briefly described. Finally, the concept of internetworking and the use of TCP/IP to achieve internetworking are discussed.

Part Two: Wireless Communication Technology

This part is concerned with the underlying technology of wireless transmission and the encoding of analog and digital data for wireless transmission.

Chapter 5: Antennas and Propagation Chapter 5 examines the fundamental principles of radio and microwave. The chapter discusses relevant aspects of antenna performance, then looks at wireless transmission modes, and finally examines the key issue of fading.

Chapter 6: Signal Encoding Techniques Data come in both analog (continuous) and digital (discrete) form. For transmission, input data must be encoded as an electrical signal that is tailored to the characteristics of the transmission medium. Both analog and digital data can be represented by either analog or digital signals; the relevant cases for wireless transmission are discussed in Chapter 6.

Chapter 7: Spread Spectrum An increasingly popular form of wireless communications is known as spread spectrum. Two general approaches are used: frequency hopping and direct sequence spread spectrum. Chapter 7 provides an overview of both techniques. The chapter also looks at the concept of code division multiple access (CDMA), which is an application of spread spectrum to provide multiple access.

Chapter 8: Coding and Error Control Wireless communications systems are highly prone to error, and virtually all wireless transmission schemes include techniques for forward error correction (FEC) by adding redundancy to the transmitted data so that bit errors can be corrected at the receiver. Chapter 8 examines FEC in detail. In addition, Chapter 8 looks at the use of redundancy for error detection, which is also found in many wireless schemes. Finally, error detection is often combined with automatic repeat request (ARQ) techniques that enable a transmitter to retransmit blocks of data in which the receiver has detected an error.

Part Three: Wireless Networking

This part examines the major types of wireless networks. These include satellite-based networks, cellular networks, cordless systems, fixed wireless access schemes, and the use of mobile IP and the Wireless Application Protocol (WAP) to provide Internet and Web access.

Chapter 9: Satellite Communications This chapter covers the basic principles of satellite communications. It looks at geostationary satellites (GEOS), low-earth orbiting satellites (LEOS), and medium-earth orbiting satellites (MEOS). The key design issue of capacity allocation is examined in detail.

Chapter 10: Cellular Wireless Networks Chapter 10 begins with a discussion of the important design issues related to cellular wireless networks. Next, the chapter covers the traditional mobile telephony service, now known as first-generation analog. Chapter 10 then examines second-generation digital cellular networks, looking at the two principal approaches: time division multiple access (TDMA) and code division multiple access (CDMA). Finally, an overview of third-generation networks is provided.

Chapter 11: Cordless Systems and Wireless Local Loop Chapter 11 looks at two technologies that bring wireless access into the residence and office: cordless systems and wireless local loop (WLL). Cordless systems have evolved from the simple single-user cordless telephones used within the home to accommodate multiple users over much larger ranges. Sometimes called radio in the loop (RITL) or fixed wireless access (FWA), WLL is a system that connects subscribers to the public switched telephone network (PSTN) using radio signals as a substitute for copper for all or part of the connection between the subscriber and the switch. Chapter 11 looks at the design issues related to WLL and then examines the IEEE 802.16 standard.

Chapter 12: Mobile IP and Wireless Access Protocol Chapter 12 examines the modifications to IP to accommodate wireless access to the Internet. The chapter then examines the Wireless Application Protocol (WAP). WAP provides mobile users of wireless phones and other wireless terminals, such as pagers and personal digital assistants (PDAs), access to telephony and information services, including the Internet and the Web.

Part Four: Wireless Local Area Networks

In recent years, a whole new class of local area networks have arrived to provide an alternative to LANs based on twisted pair, coaxial cable, and optical fiber—wireless LANs. The key advantages of the wireless LAN are that it eliminates the wiring cost, which is often the most costly component of a LAN, and that it accommodates mobile workstations. This part examines underlying wireless LAN technology and then examines two standardized approaches to local wireless networking.

Chapter 13: Wireless LAN Technology Wireless LANs use one of three transmission techniques: spread spectrum, narrowband microwave, and infrared. Chapter 13 provides an overview of LANs and wireless LAN technology and applications.

Chapter 14: IEEE 802.11 Wireless LAN Standard The most significant set of standards defining wireless LANs are those defined by the IEEE 802.11 committee. Chapter 14 examines this set of standards in depth.

Chapter 15: Bluetooth Bluetooth is an open specification for wireless communication and networking among PCs, mobile phones, and other wireless devices. Bluetooth is one of the fastest growing technology standards ever. It is intended for use within a local area. Chapter 15 examines this specification in depth.

1.8 INTERNET AND WEB RESOURCES

There are a number of resources available on the Internet and the Web to support this book and to help one keep up with developments in this field.

Web Sites for This Book

A special Web page has been set up for this book at **WilliamStallings.com/Wireless/Wireless2e.html.** The site includes the following:

- **Useful Web sites:** There are links to other relevant Web sites, including the sites listed in this section and throughout this book.
- **Errata sheet:** An errata list for this book will be maintained and updated as needed. Please e-mail any errors that you spot to me. Errata sheets for my other books are at **WilliamStallings.com.**
- **Documents:** Includes a number of documents that expand on the treatment in the book. Topics include standards organizations and the TCP/IP checksum.
- **Figures:** All of the figures in this book in PDF (Adobe Acrobat) format.
- **Tables:** All of the tables in this book in PDF format.
- **Slides:** A set of PowerPoint slides, organized by chapter.
- **Internet mailing list:** The site includes sign-up information for the book's Internet mailing list.
- **Wireless courses:** There are links to home pages for courses based on this book; these pages may be useful to other instructors in providing ideas about how to structure their course.

I also maintain the Computer Science Student Resource Site, at **WilliamStallings.com/StudentSupport.html;** the purpose of this site is to provide documents, information, and useful links for computer science students and professionals. Links are organized into four categories:

- **Math:** Includes a basic math refresher, a queuing analysis primer, a number system primer, and links to numerous math sites
- **How-to:** Advice and guidance for solving homework problems, writing technical reports, and preparing technical presentations
- **Research resources:** Links to important collections of papers, technical reports, and bibliographies
- **Miscellaneous:** A variety of useful documents and links

Other Web Sites

There are numerous Web sites that provide information related to the topics of this book. In subsequent chapters, pointers to specific Web sites can be found in the "Recommended Reading and Web Sites" section. Because the addresses for Web sites tend to change frequently, I have not included these in the book. For all of the Web sites listed in the book, the appropriate link can be found at this book's Web site.

The following Web sites are of general interest related to wireless communications:

- **Vendors:** Links to thousands of hardware and software vendors who currently have WWW sites, as well as a list of thousands of computer and networking companies in a Phone Directory
- **Wireless Developer Network:** News, tutorials, and discussions on wireless topics
- **Wireless.com:** An amazing list of links to all aspects or wireless communications, networking, and standards

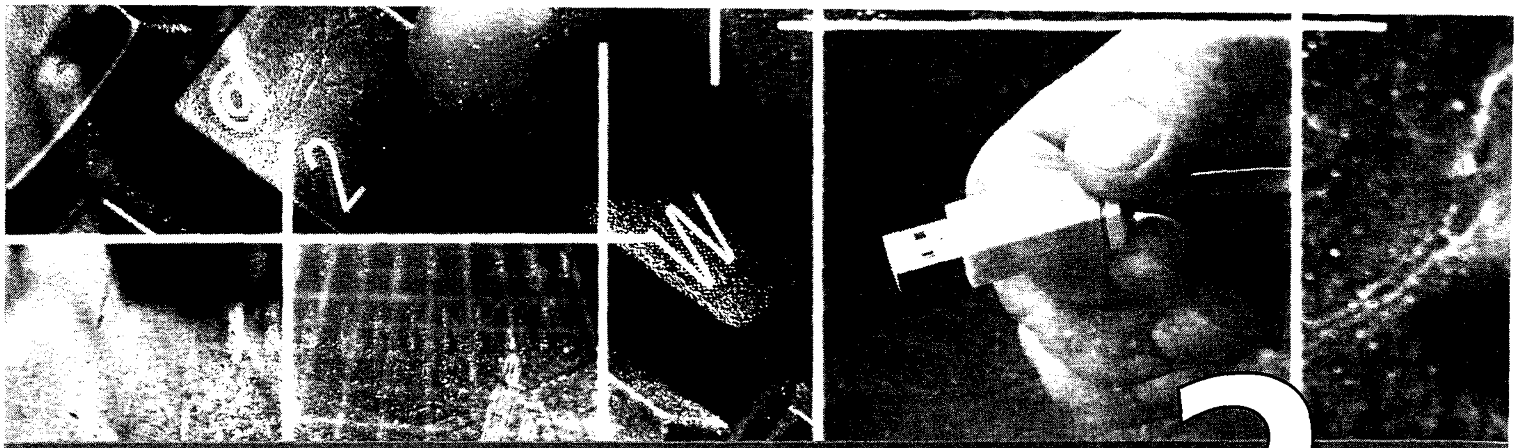
USENET Newsgroups

A number of USENET newsgroups are devoted to some aspect of data communications or networking. As with virtually all USENET groups, there is a high noise-to-signal ratio, but it is worth experimenting to see if any meet your needs. The most relevant are

- **comp.std.wireless:** General discussion of wireless standards for wide area and local area networks. This is a moderated group, which keeps the discussion focused.
- **comp.dcom.*:** There are a number of data communications related newsgroups that begin with "comp.dcom."

PART ONE

Technical Background



CHAPTER 2

TRANSMISSION FUNDAMENTALS

The purpose of this chapter is to make this book self-contained for the reader with little or no background in data communications. For the reader with greater interest, references for further study are supplied at the end of the chapter.

2.1 SIGNALS FOR CONVEYING INFORMATION

In this book, we are concerned with electromagnetic signals used as a means to transmit information. An electromagnetic signal is a function of time, but it can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies. It turns out that the frequency domain view of a signal is far more important to an understanding of data transmission than a time domain view. Both views are introduced here.

Time Domain Concepts

Viewed as a function of time, an electromagnetic signal can be either analog or digital. An **analog signal** is one in which the signal intensity varies in a smooth fashion over time. In other words, there are no breaks or discontinuities in the signal. A **digital signal** is one in which the signal intensity maintains a constant level for some period of time and then changes to another constant level.[1] Figure 2.1 shows

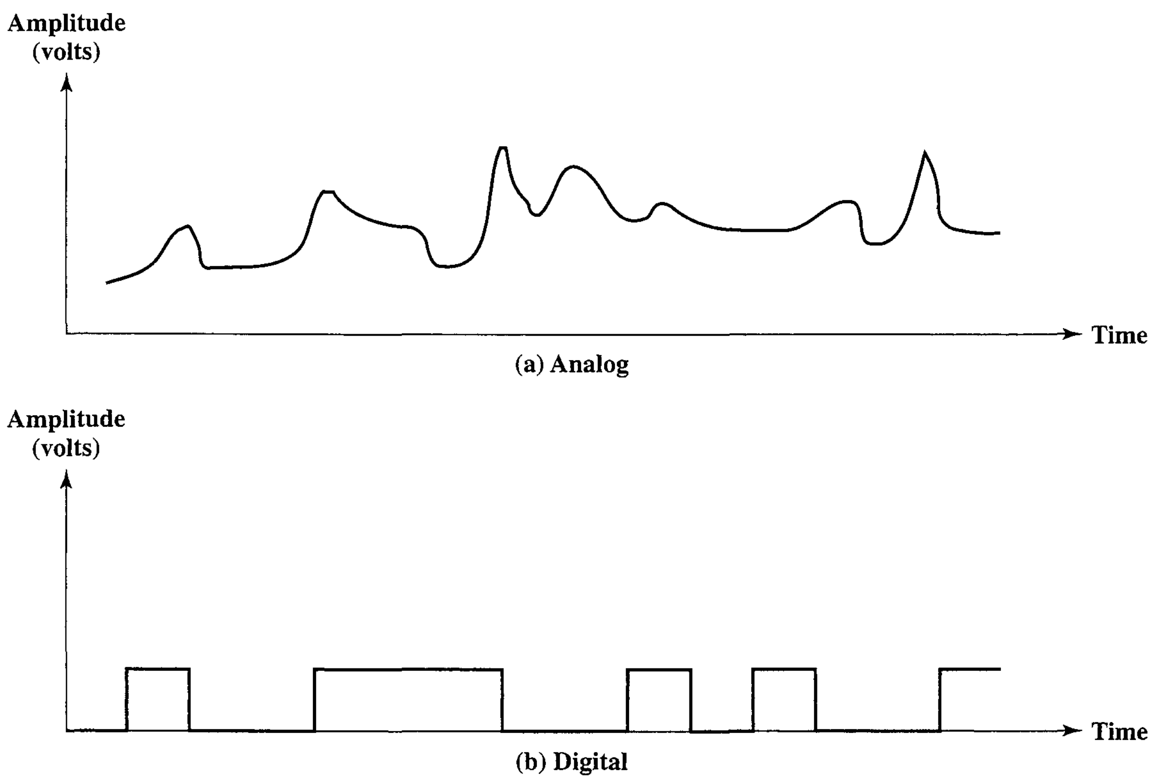


Figure 2.1 Analog and Digital Waveforms

[1]This is an idealized definition. In fact, the transition from one voltage level to another will not be instantaneous, but there will be a small transition period. Nevertheless, an actual digital signal approximates closely the ideal model of constant voltage levels with instantaneous transitions.

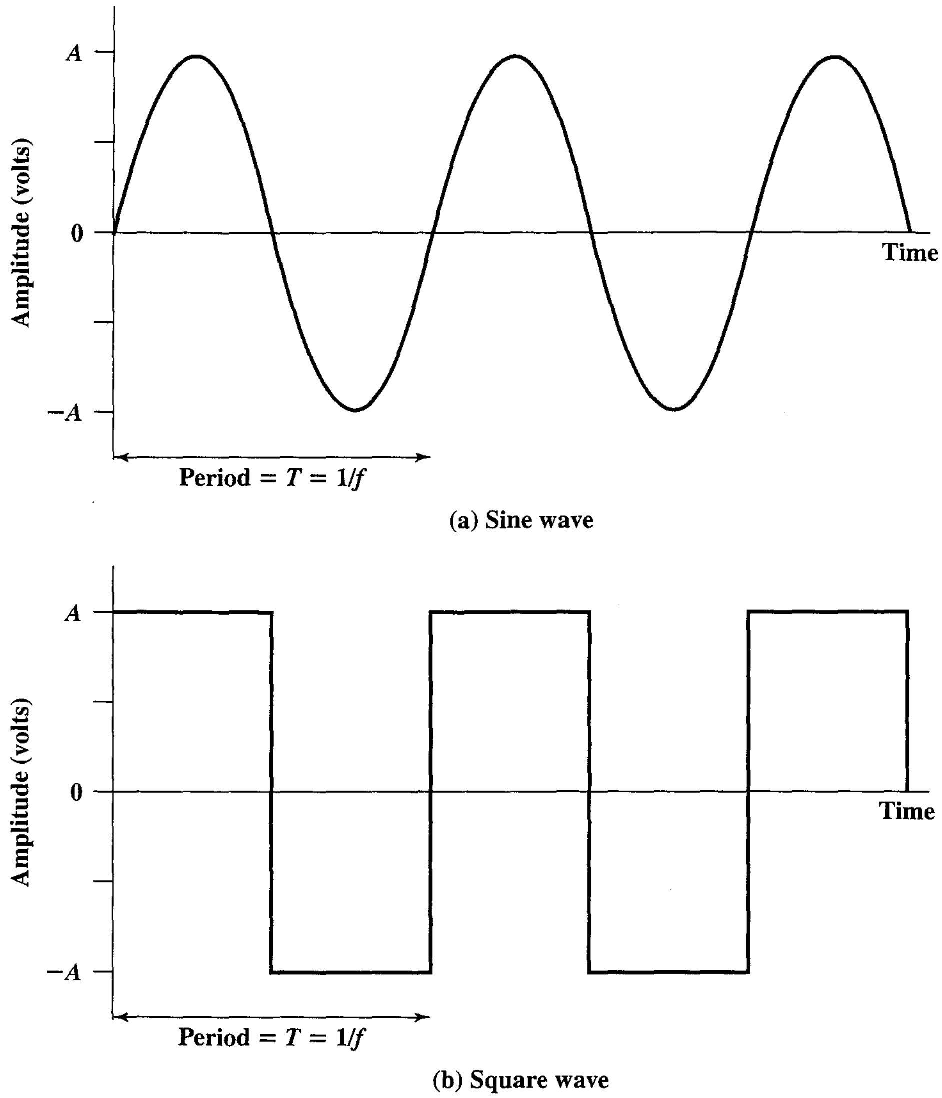


Figure 2.2 Examples of Periodic Signals

examples of both kinds of signals. The analog signal might represent speech, and the digital signal might represent binary 1s and 0s.

The simplest sort of signal is a **periodic signal**, in which the same signal pattern repeats over time. Figure 2.2 shows an example of a periodic analog signal (sine wave) and a periodic digital signal (square wave). Mathematically, a signal $s(t)$ is defined to be periodic if and only if

$$s(t + T) = s(t) \qquad -\infty < t < +\infty$$

where the constant T is the period of the signal (T is the smallest value that satisfies the equation). Otherwise, a signal is **aperiodic**.

The sine wave is the fundamental analog signal. A general sine wave can be represented by three parameters: peak amplitude (A), frequency (f), and phase (ϕ). The **peak amplitude** is the maximum value or strength of the signal over time; typically, this value is measured in volts. The **frequency** is the rate [in cycles per second, or Hertz (Hz)] at which the signal repeats. An equivalent parameter is the

period (T) of a signal, which is the amount of time it takes for one repetition; therefore, $T = 1/f$. **Phase** is a measure of the relative position in time within a single period of a signal, as illustrated later.

The general sine wave can be written

$$s(t) = A \sin(2\pi ft + \phi) \tag{2.1}$$

A function with the form of Equation (2.1) is known as a **sinusoid.** Figure 2.3 shows the effect of varying each of the three parameters. In part (a) of the figure, the frequency is 1 Hz; thus the period is $T = 1$ second. Part (b) has the same frequency and phase but a peak amplitude of 0.5. In part (c) we have $f = 2$, which is equivalent to $T = 1/2$. Finally, part (d) shows the effect of a phase shift of $\pi/4$ radians, which is 45 degrees (2π radians $= 360° = 1$ period).

In Figure 2.3 the horizontal axis is time; the graphs display the value of a signal at a given point in space as a function of time. These same graphs, with a change of scale, can apply with horizontal axes in space. In that case, the graphs display the value of a signal at a given point in time as a function of distance. For example, for a sinusoidal transmission (say, an electromagnetic radio wave some distance from a radio antenna or sound some distance from loudspeaker) at a particular instant of time, the intensity of the signal varies in a sinusoidal way as a function of distance from the source.

There is a simple relationship between the two sine waves, one in time and one in space. The **wavelength** (λ) of a signal is the distance occupied by a single cycle, or, put another way, the distance between two points of corresponding phase of two consecutive cycles. Assume that the signal is traveling with a velocity v. Then the wavelength is related to the period as follows: $\lambda = vT$. Equivalently, $\lambda f = v$. Of particular relevance to this discussion is the case where $v = c$, the speed of light in free space, which is approximately 3×10^8 m/s.

Frequency Domain Concepts

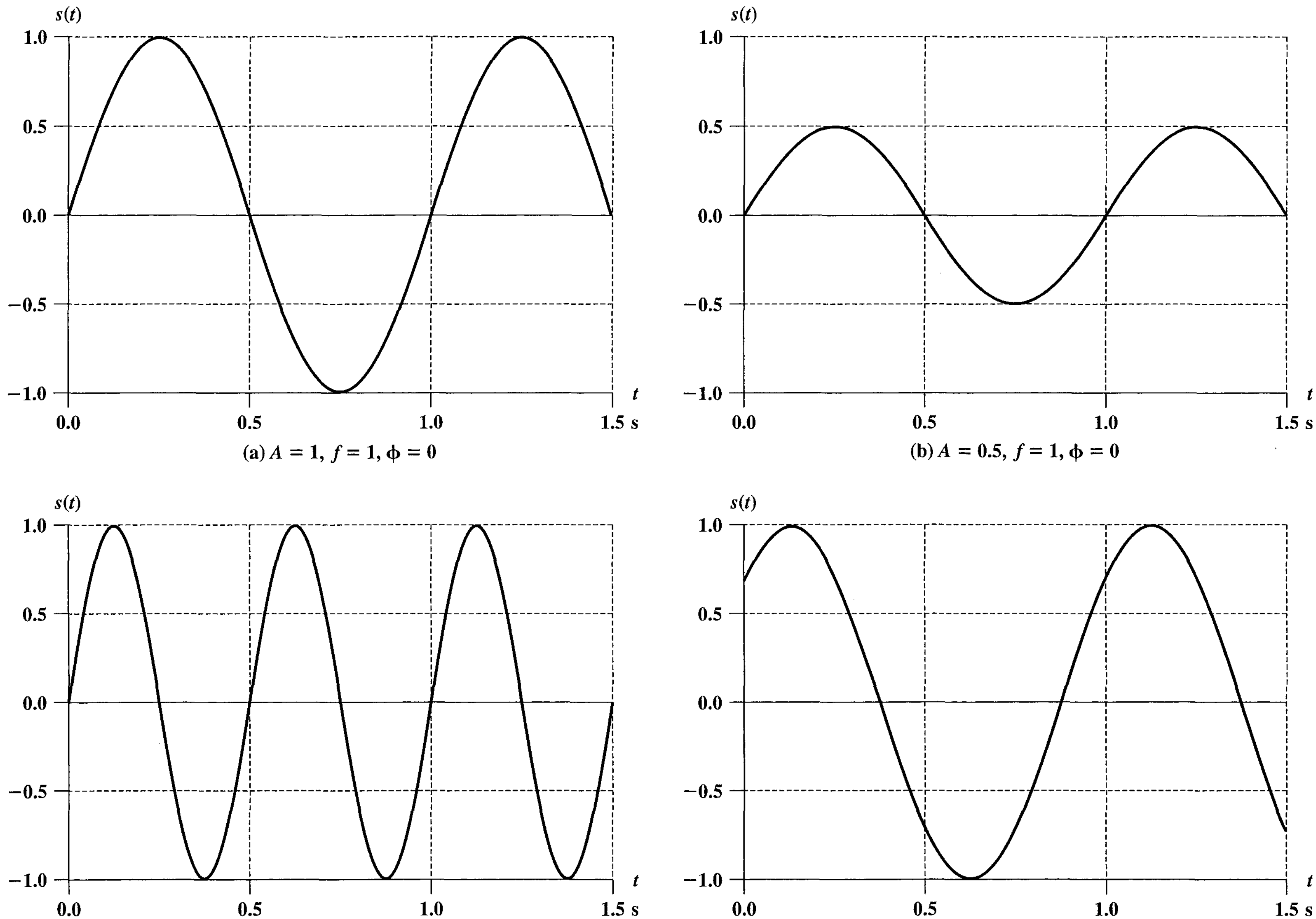
In practice, an electromagnetic signal will be made up of many frequencies. For example, the signal

$$s(t) = (4/\pi) \times (\sin(2\pi ft) + (1/3)\sin(2\pi(3f)t))$$

is shown in Figure 2.4c. The components of this signal are just sine waves of frequencies f and $3f$; parts (a) and (b) of the figure show these individual components. There are two interesting points that can be made about this figure:

- The second frequency is an integer multiple of the first frequency. When all of the frequency components of a signal are integer multiples of one frequency, the latter frequency is referred to as the **fundamental frequency**.
- The period of the total signal is equal to the period of the fundamental frequency. The period of the component $\sin(2\pi ft)$ is $T = 1/f$, and the period of $s(t)$ is also T, as can be seen from Figure 2.4c.

It can be shown, using a discipline known as Fourier analysis, that any signal is made up of components at various frequencies, in which each component is a sinusoid. By adding together enough sinusoidal signals, each with the appropriate



(a) $A = 1,\ f = 1,\ \phi = 0$

(b) $A = 0.5,\ f = 1,\ \phi = 0$

(c) $A = 1,\ f = 2,\ \phi = 0$

(d) $A = 1,\ f = 1,\ \phi = \pi/4$

Figure 2.3 $s(t) = A \sin(2\pi ft + \phi)$

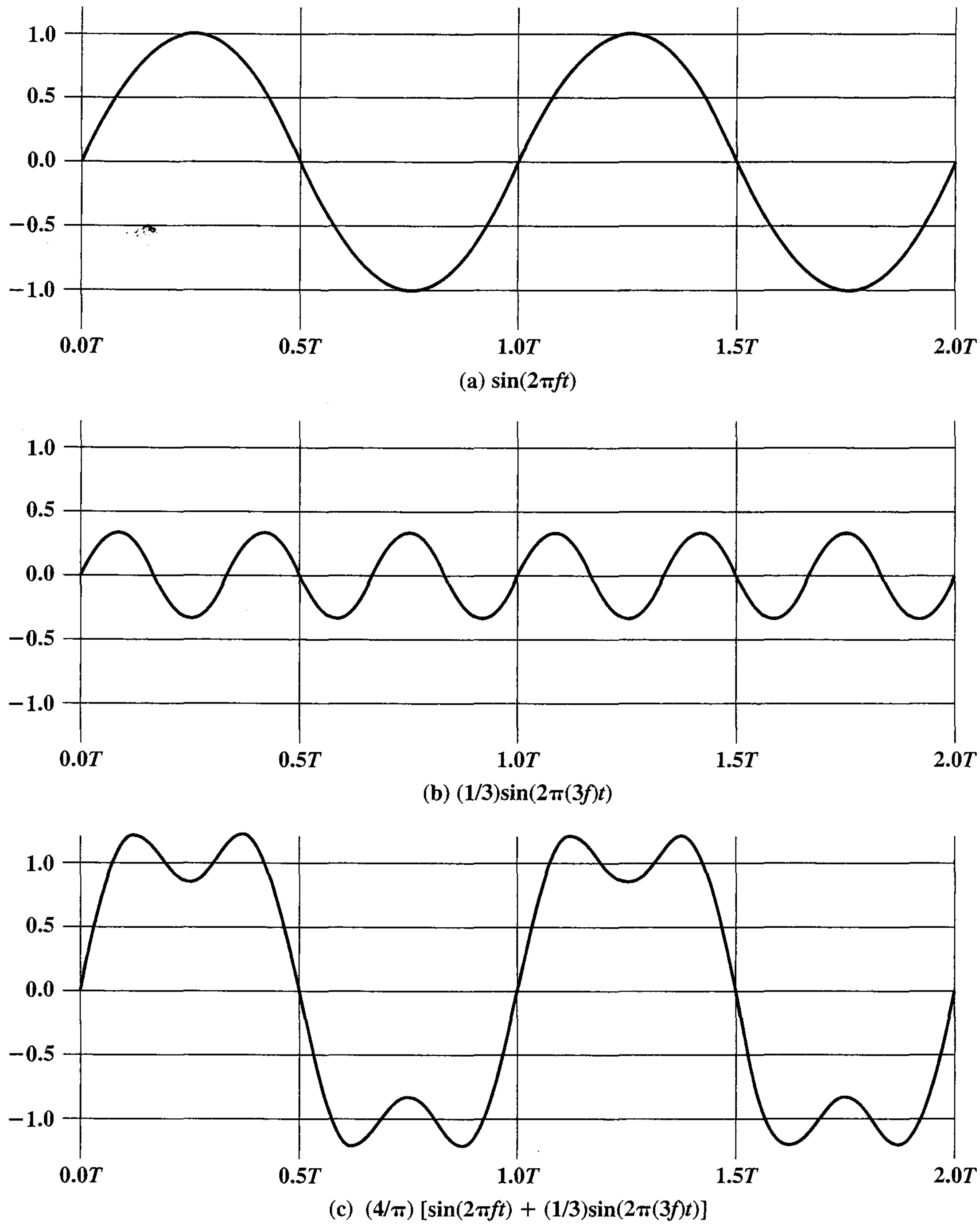


Figure 2.4 Addition of Frequency Components ($T = 1/f$)

amplitude, frequency, and phase, any electromagnetic signal can be constructed. Put another way, any electromagnetic signal can be shown to consist of a collection of periodic analog signals (sine waves) at different amplitudes, frequencies, and phases. The importance of being able to look at a signal from the frequency perspective (frequency domain) rather than a time perspective (time domain) should become clear as the discussion proceeds. For the interested reader, the subject of Fourier analysis is introduced in Appendix B.

The **spectrum** of a signal is the range of frequencies that it contains. For the signal of Figure 2.4c, the spectrum extends from f to $3f$. The **absolute bandwidth** of a signal is the width of the spectrum. In the case of Figure 2.4c, the bandwidth is $3f - f = 2f$. Many signals have an infinite bandwidth, but with most of the energy contained in a relatively narrow band of frequencies. This band is referred to as the **effective bandwidth**, or just **bandwidth**.

Relationship between Data Rate and Bandwidth

There is a direct relationship between the information-carrying capacity of a signal and its bandwidth: The greater the bandwidth, the higher the information-carrying capacity. As a very simple example, consider the square wave of Figure 2.2b. Suppose that we let a positive pulse represent binary 0 and a negative pulse represent binary 1. Then the waveform represents the binary stream 0101.... The duration of each pulse is $1/(2f)$; thus the data rate is $2f$ bits per second (bps). What are the frequency components of this signal? To answer this question, consider again Figure 2.4. By adding together sine waves at frequencies f and $3f$, we get a waveform that begins to resemble the square wave. Let us continue this process by adding a sine wave of frequency $5f$, as shown in Figure 2.5a, and then adding a sine wave of frequency $7f$, as shown in Figure 2.5b. As we add additional odd multiples of f, suitably scaled, the resulting waveform approaches that of a square wave more and more closely.

Indeed, it can be shown that the frequency components of the square wave with amplitudes A and $-A$ can be expressed as follows:

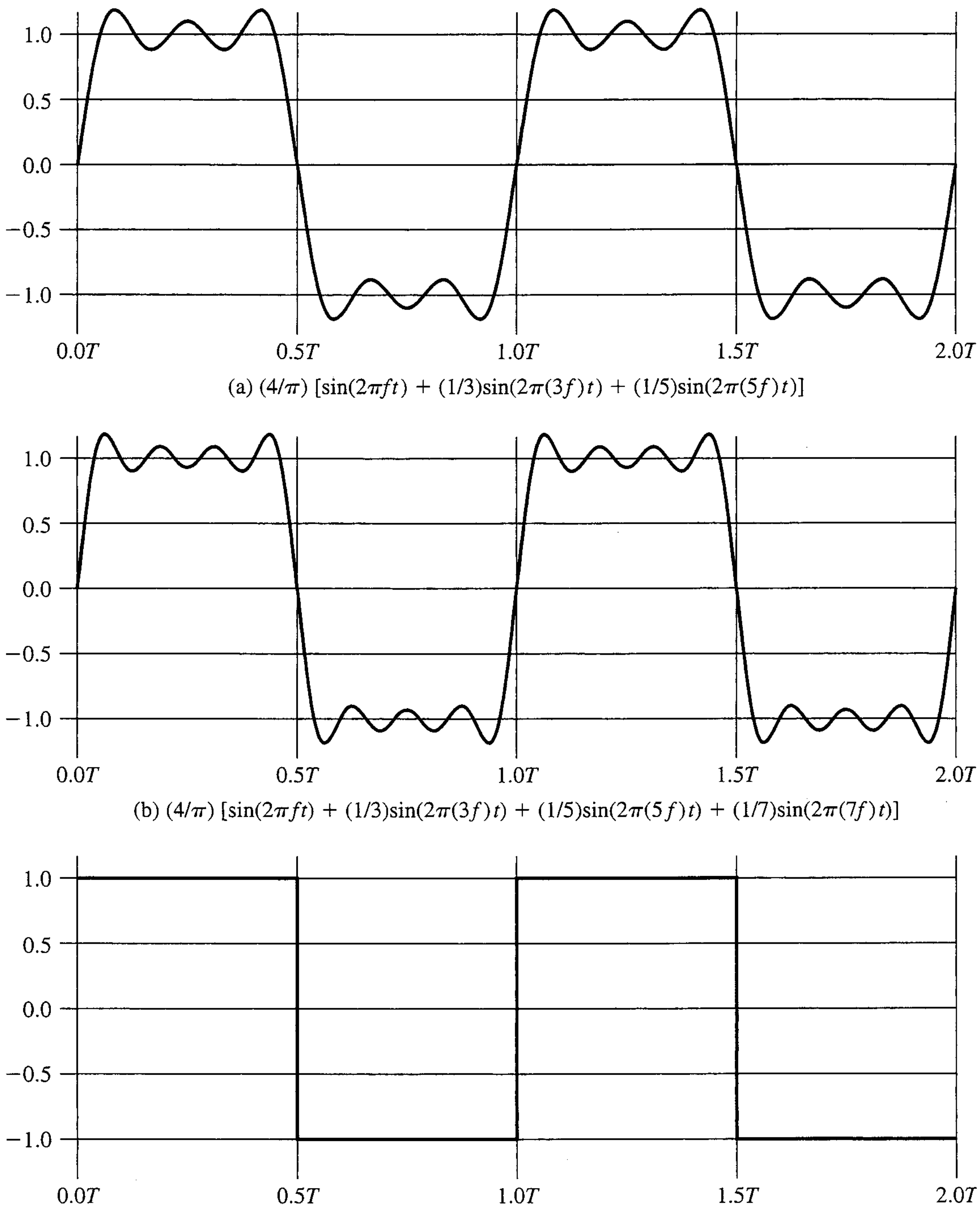
$$s(t) = A \times \frac{4}{\pi} \sum_{k \text{ odd}, k=1}^{\infty} \frac{\sin(2\pi k f t)}{k}$$

This waveform has an infinite number of frequency components and hence an infinite bandwidth. However, the peak amplitude of the kth frequency component, kf, is only $1/k$, so most of the energy in this waveform is in the first few frequency components. What happens if we limit the bandwidth to just the first three frequency components? We have already seen the answer, in Figure 2.5a. As we can see, the shape of the resulting waveform is reasonably close to that of the original square wave.

We can use Figures 2.4 and 2.5 to illustrate the relationship between data rate and bandwidth. Suppose that we are using a digital transmission system that is capable of transmitting signals with a bandwidth of 4 MHz. Let us attempt to transmit a sequence of alternating 0s and 1s as the square wave of Figure 2.5c. What data rate can be achieved? We look at three cases.

Case I. Let us approximate our square wave with the waveform of Figure 2.5a. Although this waveform is a "distorted" square wave, it is sufficiently close to the square wave that a receiver should be able to discriminate between a binary 0 and a binary 1. If we let $f = 10^6$ cycles/second $= 1$ MHz, then the bandwidth of the signal

$$s(t) = \frac{4}{\pi} \times \left[\sin((2\pi \times 10^6)t) + \frac{1}{3}\sin((2\pi \times 3 \times 10^6)t) + \frac{1}{5}\sin((2\pi \times 5 \times 10^6)t)\right]$$



(a) $(4/\pi)\,[\sin(2\pi ft) + (1/3)\sin(2\pi(3f)t) + (1/5)\sin(2\pi(5f)t)]$

(b) $(4/\pi)\,[\sin(2\pi ft) + (1/3)\sin(2\pi(3f)t) + (1/5)\sin(2\pi(5f)t) + (1/7)\sin(2\pi(7f)t)]$

(c) $(4/\pi)\,\Sigma\,(1/k)\sin(2\pi(kf)t)$, for k odd

Figure 2.5 Frequency Components of Square Wave ($T = 1/f$)

is $(5 \times 10^6) - 10^6 = 4$ MHz. Note that for $f = 1$ MHz, the period of the fundamental frequency is $T = 1/10^6 = 10^{-6} = 1\ \mu$s. If we treat this waveform as a bit string of 1s and 0s, one bit occurs every $0.5\ \mu$s, for a data rate of $2 \times 10^6 = 2$ Mbps. Thus, for a bandwidth of 4 MHz, a data rate of 2 Mbps is achieved.

Case II. Now suppose that we have a bandwidth of 8 MHz. Let us look again at Figure 2.5a, but now with $f = 2$ MHz. Using the same line of reasoning as before, the bandwidth of the signal is $(5 \times 2 \times 10^6) - (2 \times 10^6) = 8$ MHz. But in this case $T = 1/f = 0.5\ \mu s$. As a result, one bit occurs every 0.25 μs for a data rate of 4 Mbps. Thus, other things being equal, by doubling the bandwidth, we double the potential data rate.

Case III. Now suppose that the waveform of Figure 2.4c is considered adequate for approximating a square wave. That is, the difference between a positive and negative pulse in Figure 2.4c is sufficiently distinct that the waveform can be used successfully to represent a sequence of 1s and 0s. Assume as in Case II that $f = 2$ MHz and $T = 1/f = 0.5\ \mu s$, so that one bit occurs every 0.25 μs for a data rate of 4 Mbps. Using the waveform of Figure 2.4c, the bandwidth of the signal is $(3 \times 2 \times 10^6) - (2 \times 10^6) = 4$ MHz. Thus, a given bandwidth can support various data rates depending on the ability of the receiver to discern the difference between 0 and 1 in the presence of noise and other impairments.

To summarize,

- **Case I:** Bandwidth = 4 MHz; data rate = 2 Mbps
- **Case II:** Bandwidth = 8 MHz; data rate = 4 Mbps
- **Case III:** Bandwidth = 4 MHz; data rate = 4 Mbps

We can draw the following conclusions from the preceding discussion. In general, any digital waveform will have infinite bandwidth. If we attempt to transmit this waveform as a signal over any medium, the transmission system will limit the bandwidth that can be transmitted. Furthermore, for any given medium, the greater the bandwidth transmitted, the greater the cost. Thus, on the one hand, economic and practical reasons dictate that digital information be approximated by a signal of limited bandwidth. On the other hand, limiting the bandwidth creates distortions, which makes the task of interpreting the received signal more difficult. The more limited the bandwidth, the greater the distortion and the greater the potential for error by the receiver.

2.2 ANALOG AND DIGITAL DATA TRANSMISSION

The terms *analog* and *digital* correspond, roughly, to *continuous* and *discrete*, respectively. These two terms are used frequently in data communications in at least three contexts: data, signals, and transmission.

Briefly, we define **data** as entities that convey meaning, or information. **Signals** are electric or electromagnetic representations of data. **Transmission** is the communication of data by the propagation and processing of signals. In what follows, we try to make these abstract concepts clear by discussing the terms *analog* and *digital* as applied to data, signals, and transmission.

Analog and Digital Data

The concepts of analog and digital data are simple enough. Analog data take on continuous values in some interval. For example, voice and video are continuously

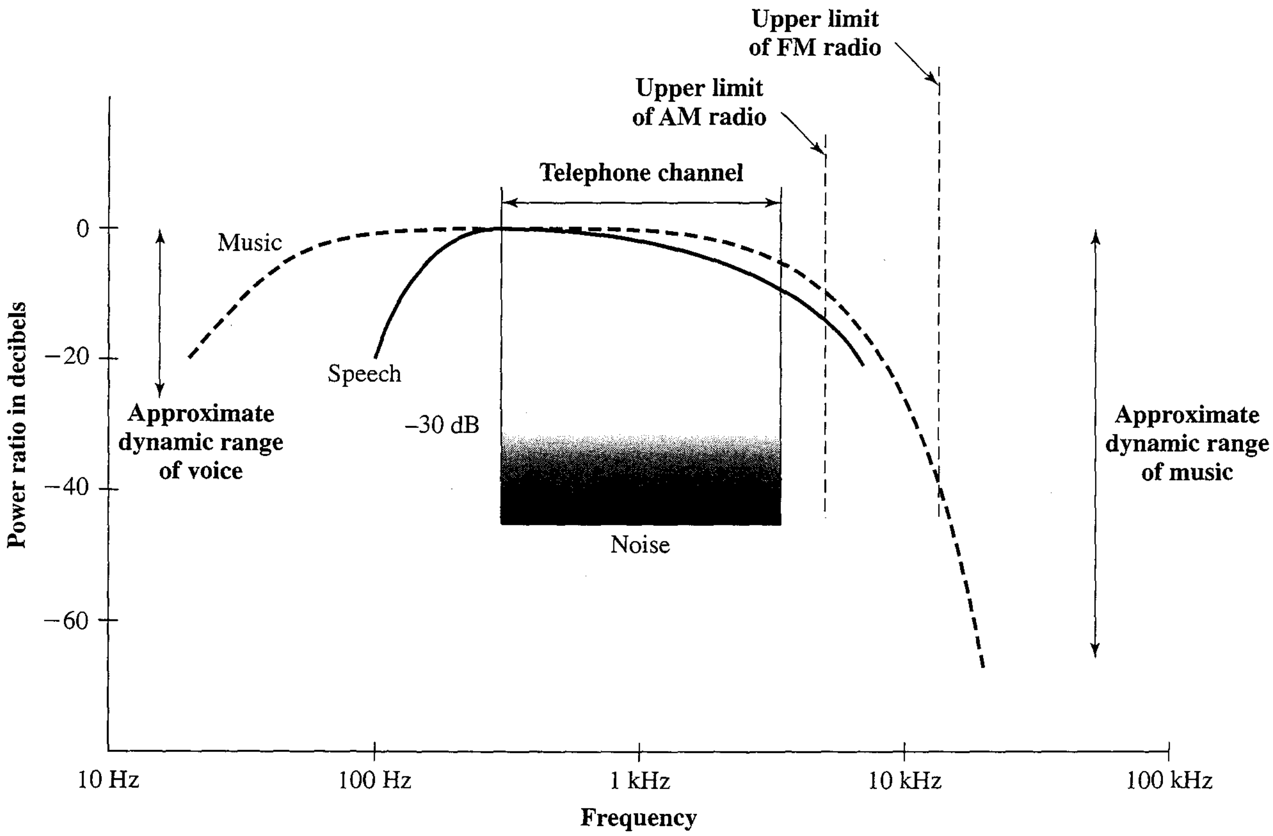


Figure 2.6 Acoustic Spectrum of Speech and Music [CARN99]

varying patterns of intensity. Most data collected by sensors, such as temperature and pressure, are continuous valued. Digital data take on discrete values; examples are text and integers.

The most familiar example of analog data is **audio**, which, in the form of acoustic sound waves, can be perceived directly by human beings. Figure 2.6 shows the acoustic spectrum for human speech and for music. Frequency components of typical speech may be found between approximately 100 Hz and 7 kHz. Although much of the energy in speech is concentrated at the lower frequencies, tests have shown that frequencies below 600 or 700 Hz add very little to the intelligibility of speech to the human ear. Typical speech has a dynamic range of about 25 dB;[2] that is, the power produced by the loudest shout may be as much as 300 times greater than that of the least whisper.

Analog and Digital Signaling

In a communications system, data are propagated from one point to another by means of electromagnetic signals. An **analog signal** is a continuously varying electromagnetic wave that may be propagated over a variety of media, depending on frequency; examples are copper wire media, such as twisted pair and coaxial cable;

[2]The concept of decibels is explained in Appendix 2A.

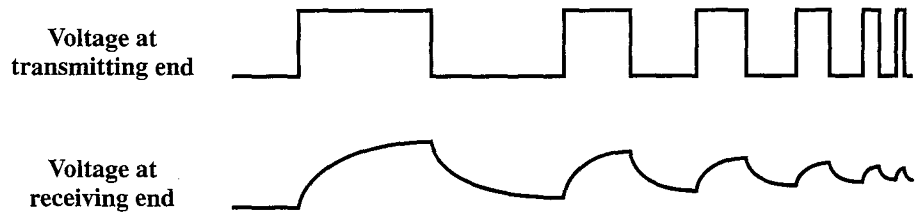


Figure 2.7 Attenuation of Digital Signals

fiber optic cable; and atmosphere or space propagation (wireless). A **digital signal** is a sequence of voltage pulses that may be transmitted over a copper wire medium; for example, a constant positive voltage level may represent binary 0 and a constant negative voltage level may represent binary 1.

The principal advantages of digital signaling are that it is generally cheaper than analog signaling and is less susceptible to noise interference. The principal disadvantage is that digital signals suffer more from attenuation than do analog signals. Figure 2.7 shows a sequence of voltage pulses, generated by a source using two voltage levels, and the received voltage some distance down a conducting medium. Because of the attenuation, or reduction, of signal strength at higher frequencies, the pulses become rounded and smaller. It should be clear that this attenuation can lead rather quickly to the loss of the information contained in the propagated signal.

Both analog and digital data can be represented, and hence propagated, by either analog or digital signals. This is illustrated in Figure 2.8. Generally, analog data are a function of time and occupy a limited frequency spectrum. Such data can be directly represented by an electromagnetic signal occupying the same spectrum. The best example of this is voice data. As sound waves, voice data have frequency components in the range 20 Hz to 20 kHz. As was mentioned, most of the speech energy is in a much narrower range, with the typical speech range of between 100 Hz and 7 kHz. The standard spectrum of voice signals is even narrower, at 300 to 3400 Hz, and this is quite adequate to propagate speech intelligibly and clearly. The telephone instrument does just that. For all sound input in the range of 300 to 3400 Hz, an electromagnetic signal with the same frequency–amplitude pattern is produced. The process is performed in reverse to convert the electromagnetic energy back into sound.

Digital data can also be represented by analog signals by use of a modem (modulator-demodulator). The modem converts a series of binary (two-valued) voltage pulses into an analog signal by modulating a carrier frequency. The resulting signal occupies a certain spectrum of frequency centered about the carrier and may be propagated across a medium suitable for that carrier. The most common modems represent digital data in the voice spectrum and hence allow those data to be propagated over ordinary voice-grade telephone lines. At the other end of the line, a modem demodulates the signal to recover the original data.

In an operation very similar to that performed by a modem, analog data can be represented by digital signals. The device that performs this function for voice data is a codec (coder-decoder). In essence, the codec takes an analog signal that directly represents the voice data and approximates that signal by a bit stream. At the other end of the line, a codec uses the bit stream to reconstruct the analog data. This topic is explored subsequently.

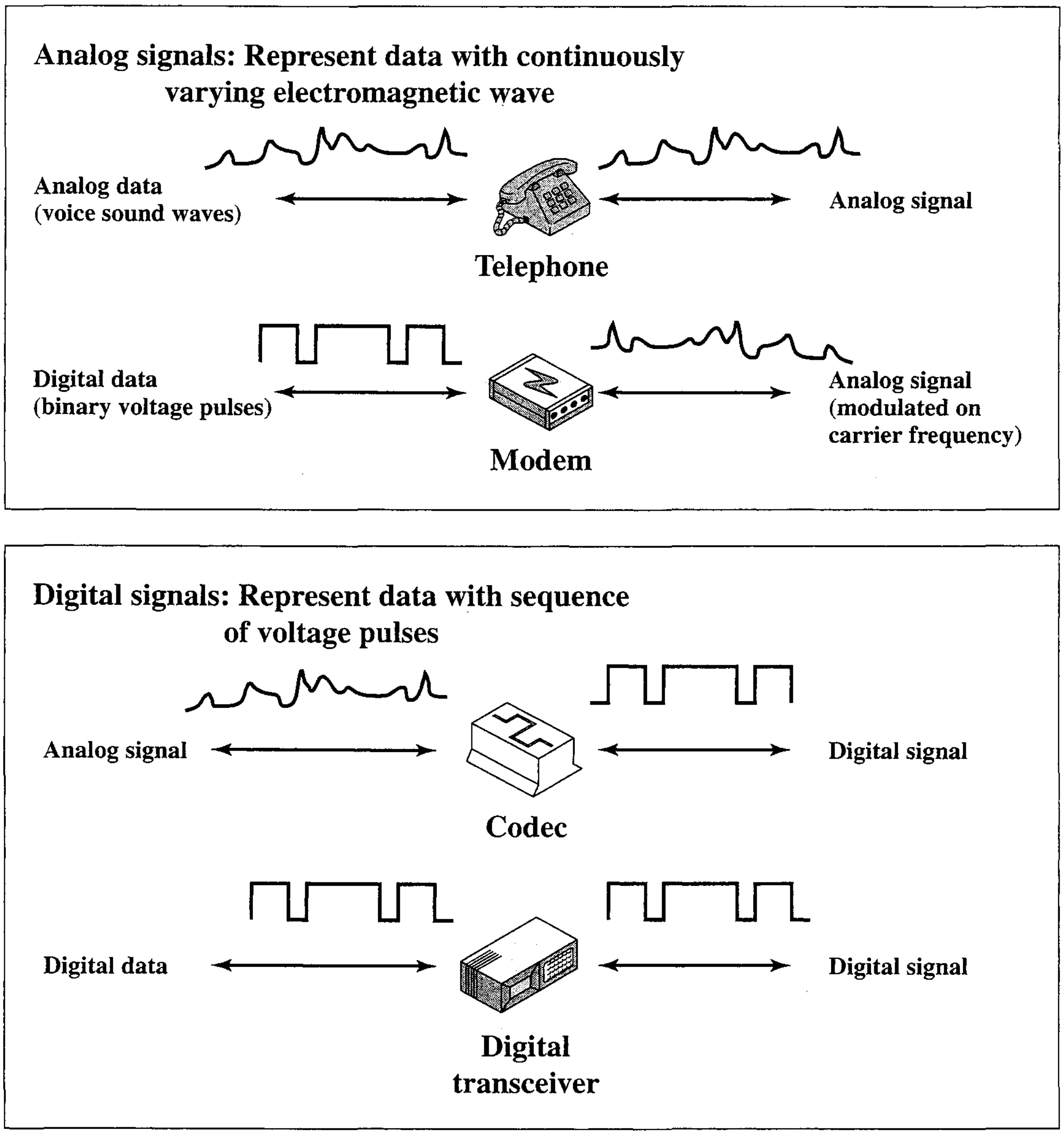


Figure 2.8 Analog and Digital Signaling of Analog and Digital Data

Finally, digital data can be represented directly, in binary form, by two voltage levels. To improve propagation characteristics, however, the binary data are often encoded into a more complex form of digital signal, as explained subsequently.

Each of the four combinations (Table 2.1a) just described is in widespread use. The reasons for choosing a particular combination for any given communications task vary. We list here some representative reasons:

- **Digital data, digital signal:** In general, the equipment for encoding digital data into a digital signal is less complex and less expensive than digital-to-analog equipment.
- **Analog data, digital signal:** Conversion of analog data to digital form permits the use of modern digital transmission and switching equipment for analog data.

Table 2.1 Analog and Digital Transmission

(a) Data and Signals

	Analog Signal	Digital Signal
Analog Data	Two alternatives: (1) signal occupies the same spectrum as the analog data; (2) analog data are encoded to occupy a different portion of spectrum.	Analog data are encoded using a codec to produce a digital bit stream.
Digital Data	Digital data are encoded using a modem to produce analog signal.	Two alternatives: (1) signal consists of two voltage levels to represent the two binary values; (2) digital data are encoded to produce a digital signal with desired properties.

(b) Treatment of Signals

	Analog Transmission	Digital Transmission
Analog Signal	Is propagated through amplifiers; same treatment whether signal is used to represent analog data or digital data.	Assumes that the analog signal represents digital data. Signal is propagated through repeaters; at each repeater, digital data are recovered from inbound signal and used to generate a new analog outbound signal.
Digital Signal	Not used	Digital signal represents a stream of 1s and 0s, which may represent digital data or may be an encoding of analog data. Signal is propagated through repeaters; at each repeater, stream of 1s and 0s is recovered from inbound signal and used to generate a new digital outbound signal.

- **Digital data, analog signal:** Some transmission media, such as optical fiber and satellite, will only propagate analog signals.
- **Analog data, analog signal:** Analog data are easily converted to an analog signal.

Analog and Digital Transmission

Both analog and digital signals may be transmitted on suitable transmission media. The way these signals are treated is a function of the transmission system. Table 2.1b summarizes the methods of data transmission. **Analog transmission** is a means of transmitting analog signals without regard to their content; the signals may represent analog data (e.g., voice) or digital data (e.g., data that pass through a modem). In either case, the analog signal will suffer attenuation that limits the length of the transmission link. To achieve longer distances, the analog transmission system includes amplifiers that boost the energy in the signal. Unfortunately, the amplifier also boosts the noise components. With amplifiers cascaded to achieve long distance, the signal becomes more and more distorted. For analog data, such as voice, quite a bit of distortion can be tolerated and the data remain intelligible. However, for digital data transmitted as analog signals, cascaded amplifiers will introduce errors.

Digital transmission, in contrast, is concerned with the content of the signal. We have mentioned that a digital signal can be propagated only a limited distance before attenuation endangers the integrity of the data. To achieve greater distances, repeaters are used. A repeater receives the digital signal, recovers the pattern of ones and zeros, and retransmits a new signal. Thus, the attenuation is overcome.

The same technique may be used with an analog signal if the signal carries digital data. At appropriately spaced points, the transmission system has retransmission devices rather than amplifiers. The retransmission device recovers the digital data from the analog signal and generates a new, clean analog signal. Thus, noise is not cumulative.

2.3 CHANNEL CAPACITY

A variety of impairments can distort or corrupt a signal. A common impairment is noise, which is any unwanted signal that combines with and hence distorts the signal intended for transmission and reception. Noise and other impairments are discussed in Chapter 5. For the purposes of this section, we simply need to know that noise is something that degrades signal quality. For digital data, the question that then arises is to what extent these impairments limit the data rate that can be achieved. The maximum rate at which data can be transmitted over a given communication path, or channel, under given conditions is referred to as the **channel capacity**.

There are four concepts here that we are trying to relate to one another:

- **Data rate:** This is the rate, in bits per second (bps), at which data can be communicated.
- **Bandwidth:** This is the bandwidth of the transmitted signal as constrained by the transmitter and the nature of the transmission medium, expressed in cycles per second, or Hertz.
- **Noise:** For this discussion, we are concerned with the average level of noise over the communications path.
- **Error rate:** This is the rate at which errors occur, where an error is the reception of a 1 when a 0 was transmitted or the reception of a 0 when a 1 was transmitted.

The problem we are addressing is this: Communications facilities are expensive and, in general, the greater the bandwidth of a facility, the greater the cost. Furthermore, all transmission channels of any practical interest are of limited bandwidth. The limitations arise from the physical properties of the transmission medium or from deliberate limitations at the transmitter on the bandwidth to prevent interference from other sources. Accordingly, we would like to make as efficient use as possible of a given bandwidth. For digital data, this means that we would like to get as high a data rate as possible at a particular limit of error rate for a given bandwidth. The main constraint on achieving this efficiency is noise.

Nyquist Bandwidth

To begin, let us consider the case of a channel that is noise free. In this environment, the limitation on data rate is simply the bandwidth of the signal. A formulation of this limitation, due to Nyquist, states that if the rate of signal transmission is $2B$, then a signal with frequencies no greater than B is sufficient to carry the signal rate. The converse is also true: Given a bandwidth of B, the highest signal rate that can be carried is $2B$. This limitation is due to the effect of intersymbol interference, such as is produced by delay distortion.[3] The result is useful in the development of digital-to-analog encoding schemes.

Note that in the preceding paragraph, we referred to signal rate. If the signals to be transmitted are binary (take on only two values), then the data rate that can be supported by B Hz is $2B$ bps. As an example, consider a voice channel being used, via modem, to transmit digital data. Assume a bandwidth of 3100 Hz. Then the capacity, C, of the channel is $2B = 6200$ bps. However, as we shall see in Chapter 6, signals with more than two levels can be used; that is, each signal element can represent more than one bit. For example, if four possible voltage levels are used as signals, then each signal element can represent two bits. With multilevel signaling, the Nyquist formulation becomes

$$C = 2B \log_2 M$$

where M is the number of discrete signal elements or voltage levels. Thus, for $M = 8$, a value used with some modems, a bandwidth of $B = 3100$ Hz yields a capacity $C = 18{,}600$ bps.

So, for a given bandwidth, the data rate can be increased by increasing the number of different signal elements. However, this places an increased burden on the receiver: Instead of distinguishing one of two possible signal elements during each signal time, it must distinguish one of M possible signals. Noise and other impairments on the transmission line will limit the practical value of M.

Shannon Capacity Formula

Nyquist's formula indicates that, all other things being equal, doubling the bandwidth doubles the data rate. Now consider the relationship among data rate, noise, and error rate. The presence of noise can corrupt one or more bits. If the data rate is increased, then the bits become "shorter" in time, so that more bits are affected by a given pattern of noise. Thus, at a given noise level, the higher the data rate, the higher the error rate.

Figure 2.9 is an example of the effect of noise on a digital signal. Here the noise consists of a relatively modest level of background noise plus occasional larger spikes of noise. The digital data can be recovered from the signal by sampling the received waveform once per bit time. As can be seen, the noise is occasionally sufficient to change a 1 to a 0 or a 0 to a 1.

All of these concepts can be tied together neatly in a formula developed by the mathematician Claude Shannon. As we have just illustrated, the higher the data rate, the more damage that unwanted noise can do. For a given level of noise, we would expect that a greater signal strength would improve the ability to receive data correctly in the presence of noise. The key parameter involved in this reasoning is

[3]Delay distortion of a signal occurs when the propagation delay for the transmission medium is not constant over the frequency range of the signal.

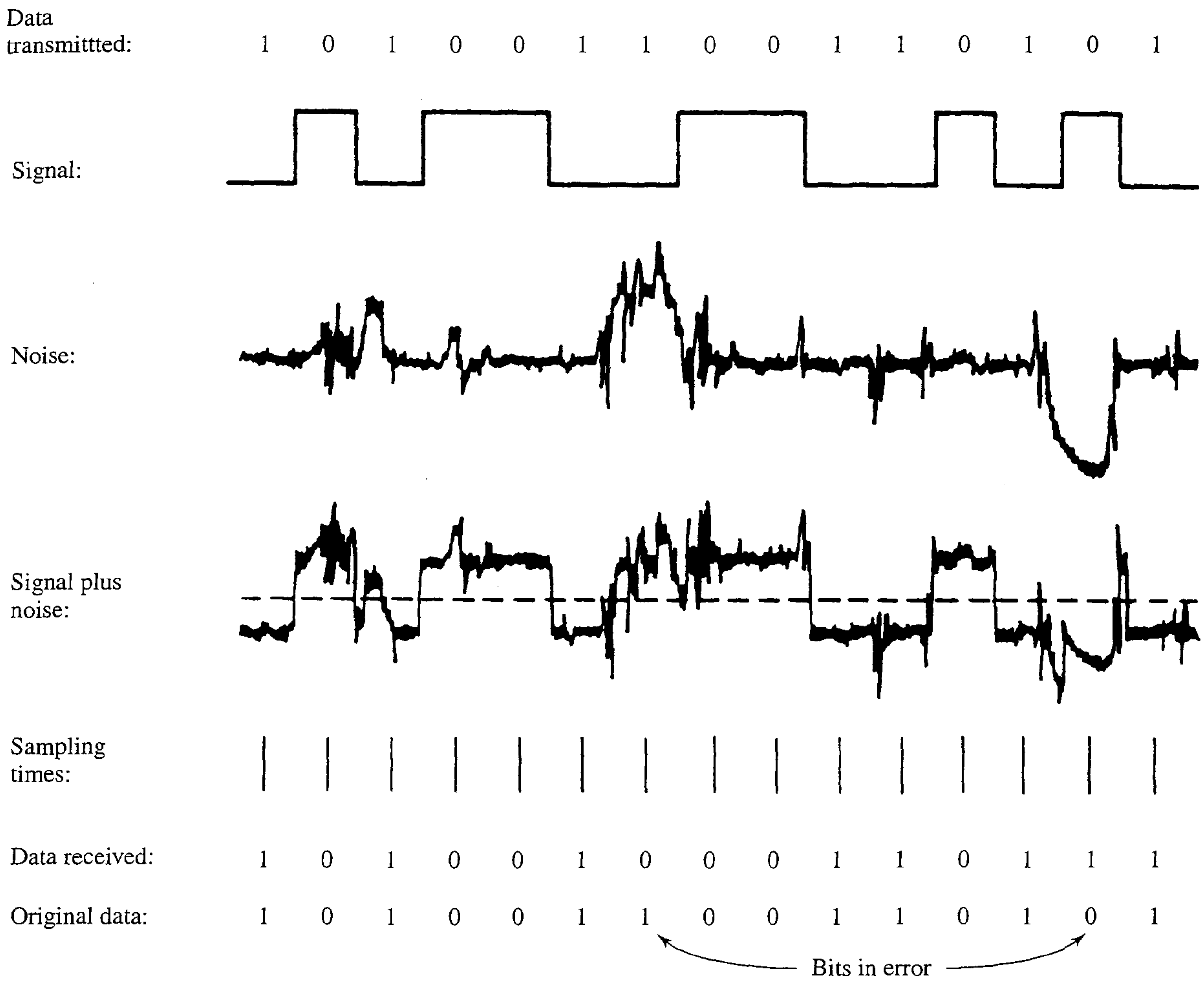


Figure 2.9 Effect of Noise on a Digital Signal

the signal-to-noise ratio (SNR, or S/N),[4] which is the ratio of the power in a signal to the power contained in the noise that is present at a particular point in the transmission. Typically, this ratio is measured at a receiver, because it is at this point that an attempt is made to process the signal and eliminate the unwanted noise. For convenience, this ratio is often reported in decibels:

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \frac{\text{signal power}}{\text{noise power}}$$

This expresses the amount, in decibels, that the intended signal exceeds the noise level. A high SNR will mean a high-quality signal.

[4]Some of the literature uses SNR; others use S/N. Also, in some cases the dimensionless quantity is referred to as SNR or S/N and the quantity in decibels is referred to as $\mathrm{SNR}_{\mathrm{db}}$ or $(\mathrm{S/N})_{\mathrm{db}}$. Others use just SNR or S/N to mean the dB quantity. This text uses SNR and $\mathrm{SNR}_{\mathrm{db}}$.

The signal-to-noise ratio is important in the transmission of digital data because it sets the upper bound on the achievable data rate. Shannon's result is that the maximum channel capacity, in bits per second, obeys the equation

$$C = B \log_2(1 + \text{SNR})$$

where C is the capacity of the channel in bits per second and B is the bandwidth of the channel in Hertz. The Shannon formula represents the theoretical maximum that can be achieved. In practice, however, only much lower rates are achieved. One reason for this is that the formula assumes white noise (thermal noise). Impulse noise is not accounted for, nor are attenuation distortion or delay distortion. Various types of noise and distortion are discussed in Chapter 5.

The capacity indicated in the preceding equation is referred to as the error-free capacity. Shannon proved that if the actual information rate on a channel is less than the error-free capacity, then it is theoretically possible to use a suitable signal code to achieve error-free transmission through the channel. Shannon's theorem unfortunately does not suggest a means for finding such codes, but it does provide a yardstick by which the performance of practical communication schemes may be measured.

Several other observations concerning the preceding equation may be instructive. For a given level of noise, it would appear that the data rate could be increased by increasing either signal strength or bandwidth. However, as the signal strength increases, so do the effects of nonlinearities in the system, leading to an increase in intermodulation noise. Note also that, because noise is assumed to be white, the wider the bandwidth, the more noise is admitted to the system. Thus, as B increases, SNR decreases.

Example 2.1 Let us consider an example that relates the Nyquist and Shannon formulations. Suppose that the spectrum of a channel is between 3 MHz and 4 MHz and $\text{SNR}_{\text{dB}} = 24$ dB. Then

$$B = 4 \text{ MHz} - 3 \text{ MHz} = 1 \text{ MHz}$$

$$\text{SNR}_{\text{dB}} = 24 \text{ dB} = 10 \log_{10}(\text{SNR})$$

$$\text{SNR} = 251$$

Using Shannon's formula,

$$C = 10^6 \times \log_2(1 + 251) \approx 10^6 \times 8 = 8 \text{ Mbps}$$

This is a theoretical limit and, as we have said, is unlikely to be reached. But assume we can achieve the limit. Based on Nyquist's formula, how many signaling levels are required? We have

$$C = 2B \log_2 M$$

$$8 \times 10^6 = 2 \times (10^6) \times \log_2 M$$

$$4 = \log_2 M$$

$$M = 16$$

2.4 TRANSMISSION MEDIA

In a data transmission system, the **transmission medium** is the physical path between transmitter and receiver. Transmission media can be classified as guided or unguided. In both cases, communication is in the form of electromagnetic waves. With **guided media**, the waves are guided along a solid medium, such as copper twisted pair, copper coaxial cable, or optical fiber. The atmosphere and outer space are examples of **unguided media**, which provide a means of transmitting electromagnetic signals but do not guide them; this form of transmission is usually referred to as **wireless transmission**.

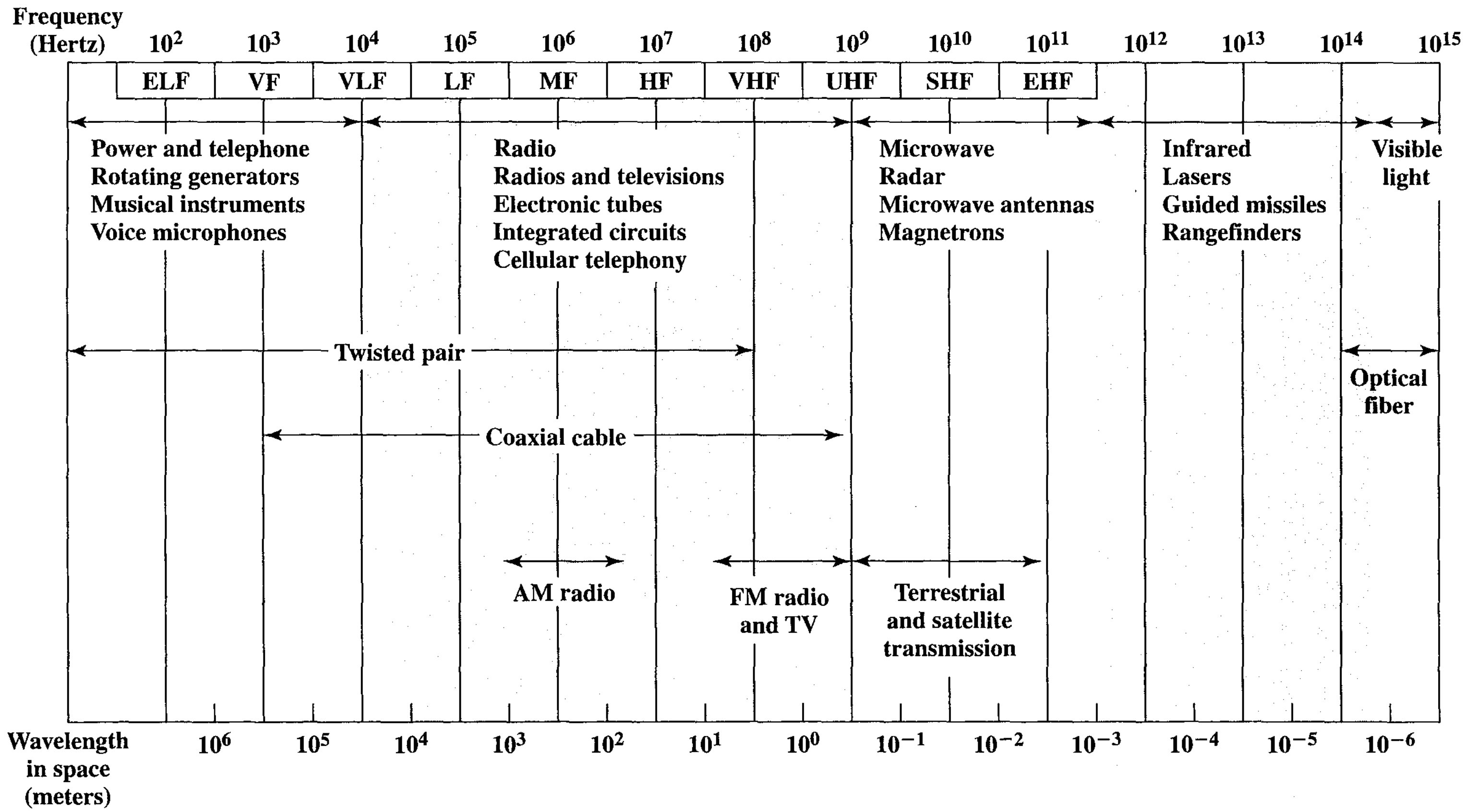
The characteristics and quality of a data transmission are determined both by the characteristics of the medium and the characteristics of the signal. In the case of guided media, the medium itself is usually more important in determining the limitations of transmission. For unguided media, the bandwidth of the signal produced by the transmitting antenna is usually more important than the medium in determining transmission characteristics. One key property of signals transmitted by antenna is directionality. In general, signals at lower frequencies are omnidirectional; that is, the signal propagates in all directions from the antenna. At higher frequencies, it is possible to focus the signal into a directional beam.

Figure 2.10 depicts the electromagnetic spectrum and indicates the frequencies at which various guided media and unguided transmission techniques operate. In the remainder of this section, we provide a brief overview of unguided, or wireless, media.

For unguided media, transmission and reception are achieved by means of an antenna. For transmission, the antenna radiates electromagnetic energy into the medium (usually air), and for reception, the antenna picks up electromagnetic waves from the surrounding medium. There are basically two types of configurations for wireless transmission: directional and omnidirectional. For the directional configuration, the transmitting antenna puts out a focused electromagnetic beam; the transmitting and receiving antennas must therefore be carefully aligned. In the omnidirectional case, the transmitted signal spreads out in all directions and can be received by many antennas.

Three general ranges of frequencies are of interest in our discussion of wireless transmission. Frequencies in the range of about 1 GHz (gigahertz = 10^9 Hz) to 100 GHz are referred to as **microwave frequencies**. At these frequencies, highly directional beams are possible, and microwave is quite suitable for point-to-point transmission. Microwave is also used for satellite communications. Frequencies in the range 30 MHz to 1 GHz are suitable for omnidirectional applications. We refer to this range as the radio range.

Another important frequency range, for local applications, is the infrared portion of the spectrum. This covers, roughly, from 3×10^{11} to 2×10^{14} Hz. Infrared is useful in local point-to-point and multipoint applications within confined areas, such as a single room.



ELF = Extremely low frequency
VF = Voice frequency
VLF = Very low frequency
LF = Low frequency
MF = Medium frequency
HF = High frequency
VHF = Very high frequency
UHF = Ultrahigh frequency
SHF = Superhigh frequency
EHF = Extremely high frequency

Figure 2.10 Electromagnetic Spectrum for Telecommunications

Terrestrial Microwave

Physical Description The most common type of microwave antenna is the parabolic "dish." A typical size is about 3 m in diameter. The antenna is fixed rigidly and focuses a narrow beam to achieve line-of-sight transmission to the receiving antenna. Microwave antennas are usually located at substantial heights above ground level to extend the range between antennas and to be able to transmit over intervening obstacles. To achieve long-distance transmission, a series of microwave relay towers is used, and point-to-point microwave links are strung together over the desired distance.

Applications A primary use for terrestrial microwave systems is in long-haul telecommunications service, as an alternative to coaxial cable or optical fiber. The microwave facility requires far fewer amplifiers or repeaters than coaxial cable over the same distance but requires line-of-sight transmission. Microwave is commonly used for both voice and television transmission.

Another increasingly common use of microwave is for short point-to-point links between buildings. This can be used for closed-circuit TV or as a data link between local area networks. Short-haul microwave can also be used for the so-called bypass application. A business can establish a microwave link to a long-distance telecommunications facility in the same city, bypassing the local telephone company.

Two other important uses of microwave are examined in some detail in Part Three: cellular systems and fixed wireless access.

Transmission Characteristics Microwave transmission covers a substantial portion of the electromagnetic spectrum. Common frequencies used for transmission are in the range 2 to 40 GHz. The higher the frequency used, the higher the potential bandwidth and therefore the higher the potential data rate. Table 2.2 indicates bandwidth and data rate for some typical systems.

As with any transmission system, a main source of loss is attenuation. For microwave (and radio frequencies), the loss can be expressed as

$$L = 10 \log \left(\frac{4 \pi d}{\lambda} \right)^2 \text{ dB} \tag{2.2}$$

where d is the distance and λ is the wavelength, in the same units. Thus, loss varies as the square of the distance. In contrast, for twisted pair and coaxial cable, loss varies exponentially with distance (linear in decibels). Thus repeaters or amplifiers may be placed farther apart for microwave systems—10 to 100 km is typical. Attenuation is increased with rainfall. The effects of rainfall become especially noticeable above 10 GHz. Another source of impairment is interference. With the growing popularity of

Table 2.2 Typical Digital Microwave Performance

Band (GHz)	Bandwidth (MHz)	Data Rate (Mbps)
2	7	12
6	30	90
11	40	135
18	220	274

microwave, transmission areas overlap and interference is always a danger. Thus the assignment of frequency bands is strictly regulated.

The most common bands for long-haul telecommunications are the 4-GHz to 6-GHz bands. With increasing congestion at these frequencies, the 11-GHz band is now coming into use. The 12-GHz band is used as a component of cable TV systems. Microwave links are used to provide TV signals to local CATV installations; the signals are then distributed to individual subscribers via coaxial cable. Higher-frequency microwave is being used for short point-to-point links between buildings; typically, the 22-GHz band is used. The higher microwave frequencies are less useful for longer distances because of increased attenuation but are quite adequate for shorter distances. In addition, at the higher frequencies, the antennas are smaller and cheaper.

Satellite Microwave

Physical Description A communication satellite is, in effect, a microwave relay station. It is used to link two or more ground-based microwave transmitter/receivers, known as earth stations, or ground stations. The satellite receives transmissions on one frequency band (uplink), amplifies or repeats the signal, and transmits it on another frequency (downlink). A single orbiting satellite will operate on a number of frequency bands, called *transponder channels*, or simply *transponders*.

Applications The communication satellite is a technological revolution as important as fiber optics. The following are among the most important applications for satellites:

- Television distribution
- Long-distance telephone transmission
- Private business networks

Because of their broadcast nature, satellites are well suited to television distribution and are being used extensively in the United States and throughout the world for this purpose. In its traditional use, a network provides programming from a central location. Programs are transmitted to the satellite and then broadcast down to a number of stations, which then distribute the programs to individual viewers. One network, the Public Broadcasting Service (PBS), distributes its television programming almost exclusively by the use of satellite channels. Other commercial networks also make substantial use of satellite, and cable television systems are receiving an ever-increasing proportion of their programming from satellites. The most recent application of satellite technology to television distribution is direct broadcast satellite (DBS), in which satellite video signals are transmitted directly to the home user. The dropping cost and size of receiving antennas have made DBS economically feasible, and DBS is now commonplace.

Satellite transmission is also used for point-to-point trunks between telephone exchange offices in public telephone networks. It is the optimum medium for high-usage international trunks and is competitive with terrestrial systems for many long-distance intranational links.

Finally, there are a number of business data applications for satellite. The satellite provider can divide the total capacity into a number of channels and lease these channels to individual business users. A user equipped with antennas at a number of sites can use a satellite channel for a private network. Traditionally, such applications have been quite expensive and limited to larger organizations with high-volume requirements.

Transmission Characteristics The optimum frequency range for satellite transmission is in the range 1 to 10 GHz. Below 1 GHz, there is significant noise from natural sources, including galactic, solar, and atmospheric noise, and human-made interference from various electronic devices. Above 10 GHz, the signal is severely attenuated by atmospheric absorption and precipitation.

Most satellites providing point-to-point service today use a frequency bandwidth in the range 5.925 to 6.425 GHz for transmission from earth to satellite (uplink) and a bandwidth in the range 3.7 to 4.2 GHz for transmission from satellite to earth (downlink). This combination is referred to as the 4/6-GHz band. Note that the uplink and downlink frequencies differ. For continuous operation without interference, a satellite cannot transmit and receive on the same frequency. Thus signals received from a ground station on one frequency must be transmitted back on another.

The 4/6-GHz band is within the optimum zone of 1 to 10 GHz but has become saturated. Other frequencies in that range are unavailable because of sources of interference operating at those frequencies, usually terrestrial microwave. Therefore, the 12/14-GHz band has been developed (uplink: 14 to 14.5 GHz; downlink: 11.7 to 12.2 GHz). At this frequency band, attenuation problems must be overcome. However, smaller and cheaper earth-station receivers can be used. It is anticipated that this band will also saturate, and use is projected for the 20/30-GHz band (uplink: 27.5 to 30.0 GHz; downlink: 17.7 to 20.2 GHz). This band experiences even greater attenuation problems but will allow greater bandwidth (2500 MHz versus 500 MHz) and even smaller and cheaper receivers.

Several properties of satellite communication should be noted. First, because of the long distances involved, there is a propagation delay of about a quarter second from transmission from one earth station to reception by another earth station. This delay is noticeable in ordinary telephone conversations. It also introduces problems in the areas of error control and flow control, which we discuss in later chapters. Second, satellite microwave is inherently a broadcast facility. Many stations can transmit to the satellite, and a transmission from a satellite can be received by many stations.

Broadcast Radio

Physical Description The principal difference between broadcast radio and microwave is that the former is omnidirectional and the latter is directional. Thus broadcast radio does not require dish-shaped antennas, and the antennas need not be rigidly mounted to a precise alignment.

Applications *Radio* is a general term used to encompass frequencies in the range of 3 kHz to 300 GHz. We are using the informal term *broadcast radio* to cover the VHF and part of the UHF band: 30 MHz to 1 GHz. This range covers FM radio and UHF and VHF television. This range is also used for a number of data networking applications.

Transmission Characteristics The range 30 MHz to 1 GHz is an effective one for broadcast communications. Unlike the case for lower-frequency electromagnetic waves, the ionosphere is transparent to radio waves above 30 MHz. Thus transmission is limited to the line of sight, and distant transmitters will not interfere with each other due to reflection from the atmosphere. Unlike the higher frequencies of the microwave region, broadcast radio waves are less sensitive to attenuation from rainfall.

As with microwave, the amount of attenuation due to distance for radio obeys Equation (2.2), namely $10 \log\left(\frac{4\pi d}{\lambda}\right)^2$ dB. Because of the longer wavelength, radio waves suffer relatively less attenuation.

A prime source of impairment for broadcast radio waves is multipath interference. Reflection from land, water, and natural or human-made objects can create multiple paths between antennas. This effect is frequently evident when TV reception displays multiple images as an airplane passes by.

Infrared

Infrared communications is achieved using transmitters/receivers (transceivers) that modulate noncoherent infrared light. Transceivers must be within the line of sight of each other either directly or via reflection from a light-colored surface such as the ceiling of a room.

One important difference between infrared and microwave transmission is that the former does not penetrate walls. Thus the security and interference problems encountered in microwave systems are not present. Furthermore, there is no frequency allocation issue with infrared, because no licensing is required.

2.5 MULTIPLEXING

In both local and wide area communications, it is almost always the case that the capacity of the transmission medium exceeds the capacity required for the transmission of a single signal. To make efficient use of the transmission system, it is desirable to carry multiple signals on a single medium. This is referred to as *multiplexing*.

Figure 2.11 depicts the multiplexing function in its simplest form. There are *n* inputs to a multiplexer. The multiplexer is connected by a single data link to a demultiplexer. The link is able to carry *n* separate channels of data. The multiplexer

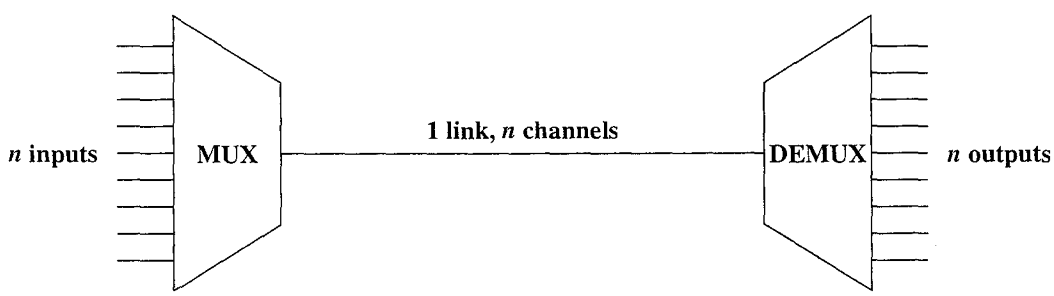


Figure 2.11 Multiplexing

combines (multiplexes) data from the *n* input lines and transmits over a higher-capacity data link. The demultiplexer accepts the multiplexed data stream, separates (demultiplexes) the data according to channel, and delivers them to the appropriate output lines.

The widespread use of multiplexing in data communications can be explained by the following:

1. The higher the data rate, the more cost effective the transmission facility. That is, for a given application and over a given distance, the cost per kbps declines with an increase in the data rate of the transmission facility. Similarly, the cost of transmission and receiving equipment, per kbps, declines with increasing data rate.
2. Most individual data communicating devices require relatively modest data rate support. For example, for most client/server applications, a data rate of up to 64 kbps is often more than adequate.

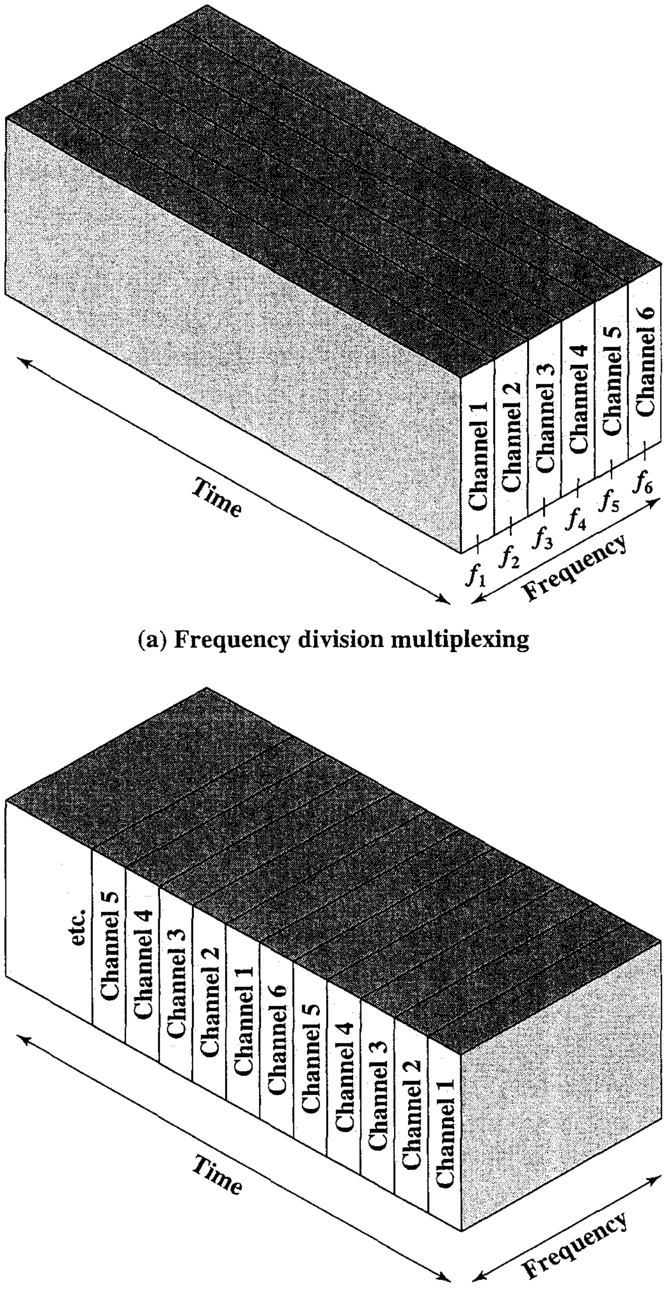
The preceding statements were phrased in terms of data communicating devices. Similar statements apply to voice communications. That is, the greater the capacity of a transmission facility, in terms of voice channels, the less the cost per individual voice channel, and the capacity required for a single voice channel is modest.

Two techniques for multiplexing in telecommunications networks are in common use: **frequency division multiplexing (FDM)** and **time division multiplexing (TDM)**.

FDM takes advantage of the fact that the useful bandwidth of the medium exceeds the required bandwidth of a given signal. A number of signals can be carried simultaneously if each signal is modulated onto a different carrier frequency and the carrier frequencies are sufficiently separated so that the bandwidths of the signals do not overlap. Figure 2.12a depicts a simple case. Six signal sources are fed into a multiplexer that modulates each signal onto a different frequency ($f_1, \ldots, f_6$). Each signal requires a certain bandwidth centered on its carrier frequency, referred to as a **channel**. To prevent interference, the channels are separated by **guard bands**, which are unused portions of the spectrum (not shown in the figure).

An example is the multiplexing of voice signals. We mentioned that the useful spectrum for voice is 300 to 3400 Hz. Thus, a bandwidth of 4 kHz is adequate to carry the voice signal and provide a guard band. For both North America (AT&T standard) and internationally (International Telecommunication Union Telecommunication Standardization Sector [ITU-T] standard), a standard voice multiplexing scheme is twelve 4-kHz voice channels from 60 to 108 kHz. For higher-capacity links, both AT&T and ITU-T define larger groupings of 4-kHz channels.

TDM takes advantage of the fact that the achievable bit rate (sometimes, unfortunately, called bandwidth) of the medium exceeds the required data rate of a digital signal. Multiple digital signals can be carried on a single transmission path by interleaving portions of each signal in time. The interleaving can be at the bit level or in blocks of bytes or larger quantities. For example, the multiplexer in Figure 2.12b has six inputs that might each be, say, 9.6 kbps. A single line with a capacity of 57.6 kbps could accommodate all six sources. Analogously to FDM, the sequence of time slots dedicated to a particular source is called a channel. One cycle of time slots (one per source) is called a frame.



(a) Frequency division multiplexing

(b) Time division multiplexing

Figure 2.12 FDM and TDM

The TDM scheme depicted in Figure 2.12b is also known as synchronous TDM, referring to the fact that time slots are preassigned and fixed. Hence the timing of transmission from the various sources is synchronized. In contrast, asynchronous TDM allows time on the medium to be allocated dynamically. Unless otherwise noted, the term *TDM* will be used to mean synchronous TDM.

A generic depiction of a synchronous TDM system is provided in Figure 2.13. A number of signals $[m_i(t), i = 1, n]$ are to be multiplexed onto the same transmission medium. The signals carry digital data and are generally digital signals. The incoming data from each source are briefly buffered. Each buffer is typically one bit

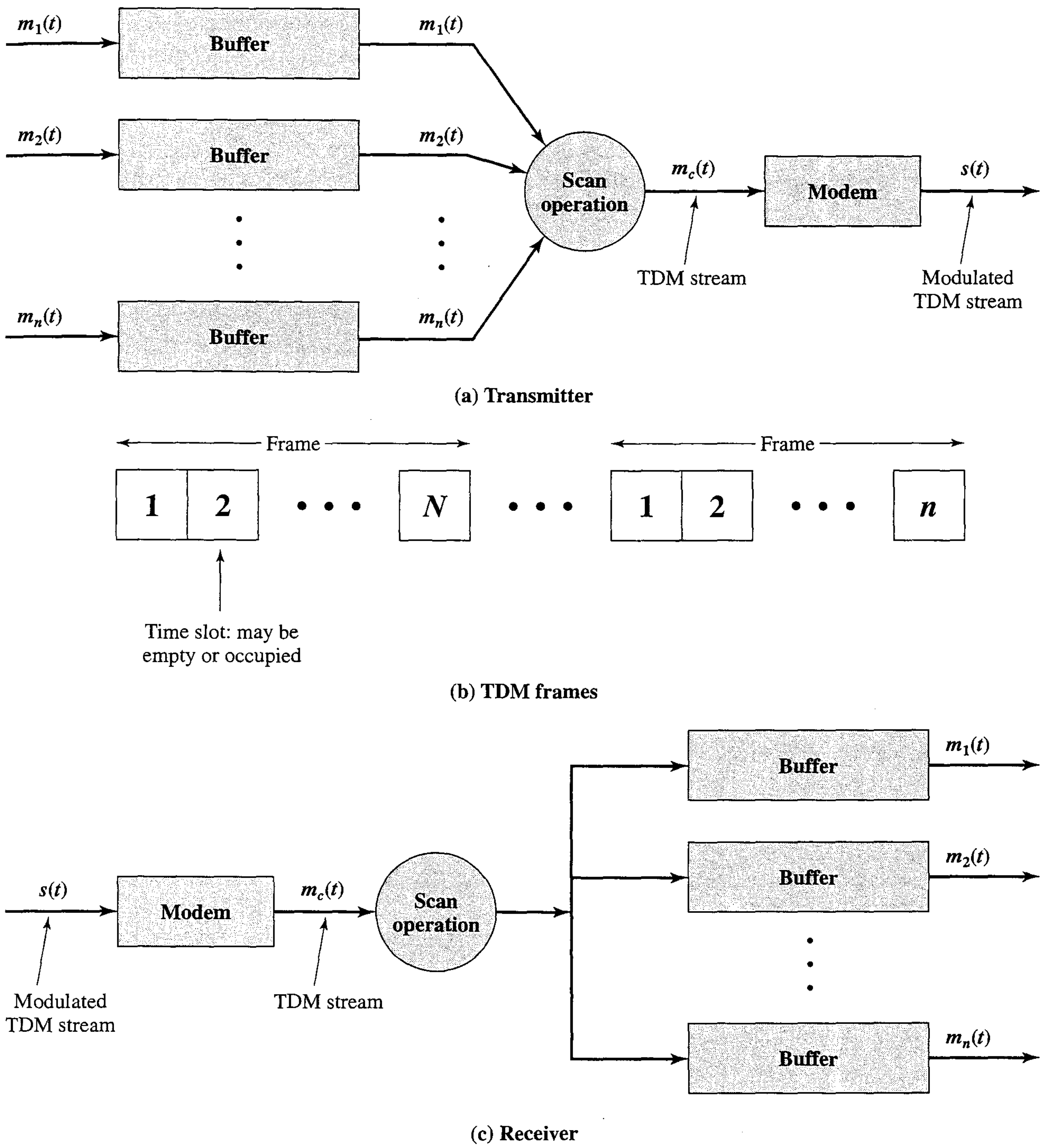


Figure 2.13 Synchronous TDM System

or one character in length. The buffers are scanned sequentially to form a composite digital data stream $m_c(t)$. The scan operation is sufficiently rapid so that each buffer is emptied before more data can arrive. Thus, the data rate of $m_c(t)$ must at least equal the sum of the data rates of the $m_i(t)$. The digital signal $m_c(t)$ may be transmitted directly or passed through a modem so that an analog signal is transmitted. In either case, transmission is typically synchronous.

The transmitted data may have a format something like Figure 2.13b. The data are organized into frames. Each frame contains a cycle of time slots. In each frame, one or more slots is dedicated to each data source. The sequence of slots dedicated to one source, from frame to frame, is called a channel. The slot length equals the transmitter buffer length, typically a bit or a byte (character).

The byte-interleaving technique is used with asynchronous and synchronous sources. Each time slot contains one character of data. Typically, the start and stop bits of each character are eliminated before transmission and reinserted by the receiver, thus improving efficiency. The bit-interleaving technique is used with synchronous sources and may also be used with asynchronous sources. Each time slot contains just one bit.

At the receiver, the interleaved data are demultiplexed and routed to the appropriate destination buffer. For each input source $m_i(t)$, there is an identical output source that will receive the input data at the same rate at which it was generated.

Synchronous TDM is called synchronous not because synchronous transmission is used but because the time slots are preassigned to sources and fixed. The time slots for each source are transmitted whether or not the source has data to send. This is, of course, also the case with FDM. In both cases, capacity is wasted to achieve simplicity of implementation. Even when fixed assignment is used, however, it is possible for a synchronous TDM device to handle sources of different data rates. For example, the slowest input device could be assigned one slot per cycle, while faster devices are assigned multiple slots per cycle.

One example of TDM is the standard scheme used for transmitting PCM voice data, known in AT&T parlance as T1 carrier. Data are taken from each source, one sample (7 bits) at a time. An eighth bit is added for signaling and supervisory functions. For T1, 24 sources are multiplexed, so there are $8 \times 24 = 192$ bits of data and control signals per frame. One final bit is added for establishing and maintaining synchronization. Thus a frame consists of 193 bits and contains one 7-bit sample per source. Since sources must be sampled 8000 times per second, the required data rate is $8000 \times 193 = 1.544$ Mbps. As with voice FDM, higher data rates are defined for larger groupings.

TDM is not limited to digital signals. Analog signals can also be interleaved in time. Also, with analog signals, a combination of TDM and FDM is possible. A transmission system can be frequency divided into a number of channels, each of which is further divided via TDM.

2.6 RECOMMENDED READING AND WEB SITES

[STAL04] covers all of the topics in this chapter in greater detail. [FREE99] is also a readable and rigorous treatment of the topics of this chapter. A thorough treatment of both analog and digital communication is provided in [COUC01].

COUC01 Couch, L. *Digital and Analog Communication Systems.* Upper Saddle River, NJ: Prentice Hall, 2001.

FREE99 Freeman, R. *Fundamentals of Telecommunications.* New York: Wiley, 1999.

STAL04 Stallings, W. *Data and Computer Communications, Seventh Edition.* Upper Saddle River: NJ: Prentice Hall, 2004.

Recommended Web sites:

- **Visualization tools:** The book Web site has links to a number of sites with resources that will help you to visualize the concepts of Section 2.1. These are useful learning tools.

2.7 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

Key Terms

analog data	frequency domain	satellite microwave
analog signal	fundamental frequency	spectrum
analog transmission	guided media	synchronous TDM
aperiodic	infrared	terrestrial microwave
bandwidth	microwave	time division multiplexing (TDM)
broadcast radio	multiplexing	time domain
channel capacity	noise	transmission media
decibel (dB)	peak amplitude	unguided media
digital data	period	wavelength
digital signal	periodic	wireless
digital transmission	phase	
frequency	radio	
frequency division multiplexing (FDM)		

Review Questions

2.1 Differentiate between an analog and a digital electromagnetic signal.
2.2 What are three important characteristics of a periodic signal?
2.3 How many radians are there in a complete circle of 360 degrees?
2.4 What is the relationship between the wavelength and frequency of a sine wave?
2.5 What is the relationship between a signal's spectrum and its bandwidth?
2.6 What is attenuation?
2.7 Define channel capacity.
2.8 What key factors affect channel capacity?
2.9 Differentiate between guided media and unguided media.
2.10 What are some major advantages and disadvantages of microwave transmission?
2.11 What is direct broadcast satellite (DBS)?
2.12 Why must a satellite have distinct uplink and downlink frequencies?
2.13 Indicate some significant differences between broadcast radio and microwave.
2.14 Why is multiplexing so cost-effective?
2.15 How is interference avoided by using frequency division multiplexing?
2.16 Explain how synchronous time division multiplexing (TDM) works.

Problems

2.1 A signal has a fundamental frequency of 1000 Hz. What is its period?

2.2 Express the following in the simplest form you can:

a. $\sin(2\pi ft - \pi) + \sin(2\pi ft + \pi)$

b. $\sin 2\pi ft + \sin(2\pi ft - \pi)$

2.3 Sound may be modeled as sinusoidal functions. Compare the wavelength and relative frequency of musical notes. Use 330 m/s as the speed of sound and the following frequencies for the musical scale.

Note	C	D	E	F	G	A	B	C
Frequency	264	297	330	352	396	440	495	528

2.4 If the solid curve in Figure 2.14 represents $\sin(2\pi t)$, what does the dotted curve represent? That is, the dotted curve can be written in the form $A \sin(2\pi ft + \phi)$; what are A, f, and ϕ?

2.5 Decompose the signal $(1 + 0.1 \cos 5t)\cos 100t$ into a linear combination of sinusoidal function, and find the amplitude, frequency, and phase of each component. *Hint:* Use the identity for $\cos a \cos b$.

2.6 Find the period of the function $f(t) = (10 \cos t)^2$.

2.7 Consider two periodic functions $f_1(t)$ and $f_2(t)$, with periods T_1 and T_2, respectively. Is it always the case that the function $f(t) = f_1(t) + f_2(t)$ is periodic? If so, demonstrate this fact. If not, under what conditions is $f(t)$ periodic?

2.8 Figure 2.5 shows the effect of eliminating higher-harmonic components of a square wave and retaining only a few lower harmonic components. What would the signal look like in the opposite case; that is, retaining all higher harmonics and eliminating a few lower harmonics?

2.9 What is the channel capacity for a teleprinter channel with a 300-Hz bandwidth and a signal-to-noise ratio of 3 dB?

2.10 A digital signaling system is required to operate at 9600 bps.

a. If a signal element encodes a 4-bit word, what is the minimum required bandwidth of the channel?

b. Repeat part (a) for the case of 8-bit words.

2.11 Study the works of Shannon and Nyquist on channel capacity. Each places an upper limit on the bit rate of a channel based on two different approaches. How are the two related?

2.12 Given the narrow (usable) audio bandwidth of a telephone transmission facility, a nominal SNR of 56dB (400,000), and a distortion level of <0.2%,

a. What is the theoretical maximum channel capacity (Kbps) of traditional telephone lines?

b. What is the actual maximum channel capacity?

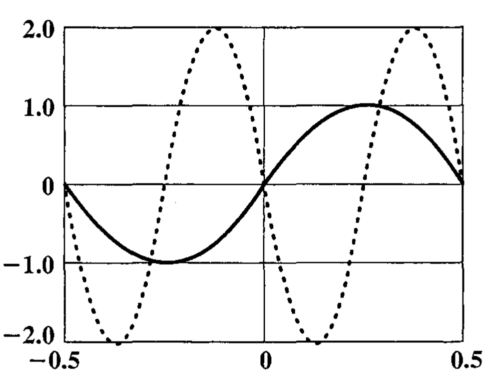


Figure 2.14 Figure for Problem 2.4

2.13 Given a channel with an intended capacity of 20 Mbps, the bandwidth of the channel is 3 MHz. What signal-to-noise ratio is required to achieve this capacity?

2.14 Show that doubling the transmission frequency or doubling the distance between transmitting antenna and receiving antenna attenuates the power received by 6 dB.

2.15 Fill in the missing elements in the following table of approximate power ratios for various dB levels.

Decibels	1	2	3	4	5	6	7	8	9	10
Losses			0.5							0.1
Gains			2							10

2.16 If an amplifier has a 30 dB voltage gain, what voltage ratio does the gain represent?

2.17 An amplifier has an output of 20 W. What is its output in dBW?

APPENDIX 2A DECIBELS AND SIGNAL STRENGTH

An important parameter in any transmission system is the signal strength. As a signal propagates along a transmission medium, there will be a loss, or *attenuation*, of signal strength. To compensate, amplifiers may be inserted at various points to impart a gain in signal strength.

It is customary to express gains, losses, and relative levels in decibels because

- Signal strength often falls off exponentially, so loss is easily expressed in terms of the decibel, which is a logarithmic unit.
- The net gain or loss in a cascaded transmission path can be calculated with simple addition and subtraction.

The decibel is a measure of the ratio between two signal levels. The decibel gain is given by

$$G_{dB} = 10 \log_{10} \frac{P_{out}}{P_{in}}$$

where

G_{dB} = gain, in decibels
P_{in} = input power level
P_{out} = output power level
$\log_{10}$ = logarithm to the base 10 (from now on, we will simply use log to mean $\log_{10}$)

Table 2.3 shows the relationship between decibel values and powers of 10.

Table 2.3 Decibel Values

Power Ratio	dB	Power Ratio	dB
10^1	10	10^{-1}	−10
10^2	20	10^{-2}	−20
10^3	30	10^{-3}	−30
10^4	40	10^{-4}	−40
10^5	50	10^{-5}	−50
10^6	60	10^{-6}	−60

There is some inconsistency in the literature over the use of the terms *gain* and *loss*. If the value of G_{dB} is positive, this represents an actual gain in power. For example, a gain of 3 dB means that the power has doubled. If the value of G_{dB} is negative, this represents an actual loss in power. For example a gain of −3 dB means that the power has halved, and this is a loss of power. Normally, this is expressed by saying there is a loss of 3 dB. However, some of the literature would say that this is a loss of −3 dB. It makes more sense to say that a negative gain corresponds to a positive loss. Therefore, we define a decibel loss as

$$L_{dB} = -10 \log_{10}\frac{P_{out}}{P_{in}} = 10 \log_{10}\frac{P_{in}}{P_{out}} \tag{2.3}$$

Example 2.2 If a signal with a power level of 10 mW is inserted onto a transmission line and the measured power some distance away is 5 mW, the loss can be expressed as $L_{dB} = 10 \log(10/5) = 10(0.3) = 3$ dB.

Note that the decibel is a measure of relative, not absolute, difference. A loss from 1000 mW to 500 mW is also a loss of 3 dB.

The decibel is also used to measure the difference in voltage, taking into account that power is proportional to the square of the voltage:

$$P = \frac{V^2}{R}$$

where

P = power dissipated across resistance R

V = voltage across resistance R

Thus

$$L_{dB} = 10 \log\frac{P_{in}}{P_{out}} = 10 \log\frac{V_{in}^2/R}{V_{out}^2/R} = 20 \log\frac{V_{in}}{V_{out}}$$

Example 2.3 Decibels are useful in determining the gain or loss over a series of transmission elements. Consider a series in which the input is at a power level of 4 mW, the first element is a transmission line with a 12-dB loss (−12 dB gain), the second element is an amplifier with a 35-dB gain, and the third element is a transmission line with a 10-dB loss. The net gain is $(-12 + 35 - 10) = 13$ dB. To calculate the output power P_{out}:

$$G_{dB} = 13 = 10 \log(P_{out}/4\text{ mW})$$
$$P_{out} = 4 \times 10^{1.3}\text{ mW} = 79.8\text{ mW}$$

Decibel values refer to relative magnitudes or changes in magnitude, not to an absolute level. It is convenient to refer to an absolute level of power or voltage in decibels so that gains and losses with reference to an initial signal level may be calculated easily. The **dBW (decibel-Watt)** is used extensively in microwave applications. The value of 1 W is selected as a reference and defined to be 0 dBW. The absolute decibel level of power in dBW is defined as

$$\text{Power}_{dBW} = 10 \log\frac{\text{Power}_W}{1\text{ W}}$$

Example 2.4 A power of 1000 W is 30 dBW, and a power of 1 mW is −30 dBW.

Another common unit is the **dBm (decibel-milliWatt)**, which uses 1 mW as the reference. Thus 0 dBm = 1 mW. The formula is

$$\text{Power}_{\text{dBm}} = 10 \log \frac{\text{Power}_{\text{mW}}}{1 \text{ mW}}$$

Note the following relationships:

$$+30 \text{ dBm} = 0 \text{ dBW}$$
$$0 \text{ dBm} = -30 \text{ dBW}$$