FIFTH EDITION

Modern Digital and Analog Communication Systems

B. P. LATHI | ZHI DING

OXFORD
UNIVERSITY PRESS

MODERN DIGITAL AND ANALOG COMMUNICATION SYSTEMS

THE OXFORD SERIES IN ELECTRICAL AND COMPUTER ENGINEERING

Adel S. Sedra, Series Editor

Allen and Holberg, *CMOS Analog Circuit Design, 3rd edition*
Boncelet, *Probability, Statistics, and Random Signals*
Bobrow, *Elementary Linear Circuit Analysis, 2nd edition*
Bobrow, *Fundamentals of Electrical Engineering, 2nd edition*
Campbell, *Fabrication Engineering at the Micro- and Nanoscale, 4th edition*
Chen, *Digital Signal Processing*
Chen, *Linear System Theory and Design, 4th edition*
Chen, *Signals and Systems, 3rd edition*
Comer, *Digital Logic and State Machine Design, 3rd edition*
Comer, *Microprocessor-Based System Design*
Cooper *and McGillem, Probabilistic Methods of Signal and System Analysis, 3rd edition*
Dimitrijev, *Principles of Semiconductor Device, 2nd edition*
Dimitrijev, *Understanding Semiconductor Devices*
Fortney, *Principles of Electronics: Analog & Digital*
Franco, *Electric Circuits Fundamentals*
Ghausi, *Electronic Devices and Circuits: Discrete and Integrated*
Guru and Hiziroğlu, *Electric Machinery and Transformers, 3rd edition*
Houts, *Signal Analysis in Linear Systems*
Jones, *Introduction to Optical Fiber Communication Systems*
Krein, *Elements of Power Electronics, 2nd Edition*
Kuo, *Digital Control Systems, 3rd edition*
Lathi and Green, *Linear Systems and Signals, 3rd edition*
Lathi and Ding, *Modern Digital and Analog Communication Systems, 5th edition*
Lathi, *Signal Processing and Linear Systems*
Martin, *Digital Integrated Circuit Design*
Miner, *Lines and Electromagnetic Fields for Engineers*
Mitra, *Signals and Systems*
Parhami, *Computer Architecture*
Parhami, *Computer Arithmetic, 2nd edition*
Roberts and Sedra, *SPICE, 2nd edition*
Roberts, Taenzler, and Burns, *An Introduction to Mixed-Signal IC Test and Measurement, 2nd edition*
Roulston, *An Introduction to the Physics of Semiconductor Devices*
Sadiku, *Elements of Electromagnetics, 7th edition*
Santina, Stubberud, and Hostetter, *Digital Control System Design, 2nd edition*
Sarma, *Introduction to Electrical Engineering*
Schaumann, Xiao, and Van Valkenburg, *Design of Analog Filters, 3rd edition*
Schwarz and Oldham, *Electrical Engineering: An Introduction, 2nd edition*
Sedra and Smith, *Microelectronic Circuits, 7th edition*
Stefani, Shahian, Savant, and Hostetter, *Design of Feedback Control Systems, 4th edition*
Tsividis, *Operation and Modeling of the MOS Transistor, 3rd edition*
Van Valkenburg, *Analog Filter Design*
Warner and Grung, *Semiconductor Device Electronics*
Wolovich, *Automatic Control Systems*
Yariv and Yeh, *Photonics: Optical Electronics in Modern Communications, 6th edition*
Żak, *Systems and Control*

MODERN DIGITAL AND ANALOG COMMUNICATION SYSTEMS

Fifth Edition

B. P. Lathi
Professor Emeritus
California State University—Sacramento

Zhi Ding
Professor
University of California—Davis

New York   Oxford
OXFORD UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

For titles covered by Section 112 of the US Higher Education Opportunity Act, please visit www.oup.com/us/he for the latest information about pricing and alternate formats.

Library of Congress Cataloging-in-Publication Data
Names: Lathi, B. P. (Bhagwandas Pannalal), author. | Ding, Zhi, 1962- author.
Title: Modern digital and analog communication systems / B.P. Lathi,
Professor Emeritus, California State University—Sacramento, Zhi Ding,
Professor, University of California—Davis.
Description: Fifth Edition. | New York : Oxford University Press, [2019] |
Series: Oxford series in electrical and computer engineering
Identifiers: LCCN 2017034966 | ISBN 9780190686840 (Hardcover)
Subjects: LCSH: Telecommunication systems. | Digital communications. |
Statistical communication theory.
Classification: LCC TK5101 .L333 2018 | DDC 621.382–dc23 LC record
available at https://lccn.loc.gov/2017034966

Printing number: 9 8 7 6 5 4 3 2 1
Printed by Edwards Brothers Malloy, United States of America

BRIEF TABLE OF CONTENTS

CONTENTS

APPENDICES

PREFACE

Since the publication of the fourth edition, we have continued to be astounded by the remarkable progress of digital revolution made possible by advanced telecommunication technologies. Within one decade, smartphones and smartphone applications have changed the lives of billions of people. Furthermore, there is little doubt that the next wave of digital revolution, likely centered around transformative technologies in machine learning, data mining, Internet of things, and artificial intelligence, shall continue to drive the development of novel communication systems and applications. It is therefore a good time for us to deliver a new edition of this textbook by integrating major new technological advances in communication systems. This fifth edition contains major updates to incorporate recent technological advances of telecommunications.

As engineering students become more and more aware of the important role that communication systems play in the modern society, they are increasingly motivated to learn through experimenting with solid, illustrative examples. To captivate students' attention and stimulate their imaginations, this new edition places strong emphasis on connecting fundamental concepts of communication theory to their daily experiences of communication technologies. We provide highly relevant information on the operation and features of wireless cellular systems, Wi-Fi network access, and broadband Internet services, among others.

Major Revisions and Additions

A number of major changes are motivated by the need to emphasize the fundamentals of digital communication systems that have permeated our daily lives. Instead of traditional approaches that disproportionally drill on the basics of analog modulation and demodulation, this new edition shifts the major focus onto the theory and practice of the broadly deployed digital communication systems. Specifically, after introducing the important tools of Fourier analysis in Chapter 2 and Chapter 3, only a single chapter (Chapter 4) is devoted to the analog amplitude and angle modulations. The authors expect most students to be far more interested in digital systems that they use daily and to be highly motivated to master the state-of-the-art digital communication technologies in order to contribute to future waves of the digital revolution.

One of the major goals in writing this new edition is to make learning a gratifying or at least a less intimidating experience for students by presenting the subject in a clear, understandable, and logically organized manner. To enhance interactive learning, this new edition has updated a number of computer-based experimental practices that are closely tied to the fundamental concepts and examples in the main text. Students can further strengthen their understanding and test their own designs through numerical experimentation based on the newly included computer assignment problems following each major chapter.

Every effort has been made to deliver insights—rather than just derivations—as well as heuristic explanations of theoretical results wherever possible. Many examples are provided for further clarification of abstract results. Even a partial success in achieving this stated goal would make all our efforts worthwhile.

Reorganization

A torrent of technological advances has nurtured a new generation of students extremely interested in learning about the new technologies and their implementations. These students are eager to understand how and where they may be able to make contributions as future innovators. Such strong motivation must be encouraged and leveraged. This new edition will enable instructors either to cover the topics themselves or to assign reading materials that will allow students to acquire relevant information. The new edition achieves these goals by stressing the digital aspects of the text and by incorporating the most commonly known wireless and wireline digital technologies.

With respect to organization, the fifth edition begins with a traditional review of signal and linear system fundamentals before proceeding to the core communication topics of analog and digital modulations. We then present the fundamental tools of probability theory and random processes to be used in the design and analysis of digital communications in the second part of the text. After covering the fundamentals of digital communication systems, the final two chapters provide an overview of information theory and the fundamentals of forward error correction codes.

Ideally, to cover the major subjects in this text with sufficient technical depth would require a sequence of two courses: one on the basic operations of communication systems and one on the analysis of modern communication systems under noise and other distortions. The former relies heavily on deterministic analytical tools such as Fourier series, Fourier transform, and the sampling theorem, while the latter relies on tools from probability and random processes to tackle the unpredictable aspects of message signals and noises. In today's academic environment, however, with so many competing courses and topics, it may be difficult to fit two basic courses on communications into a typical electrical or computer engineering curriculum. Some universities do require a course in probability and random processes as a prerequisite. In that case, it is possible to cover both areas reasonably well in a one-semester course. This book is designed for adoption in both cases regardless of whether a probability prerequisite is available. It can be used as a one-semester course in which the deterministic aspects of communication systems are emphasized with mild consideration of the effects of noise and interference. It can also be used for a course that deals with both the deterministic and the probabilistic aspects of communication systems. The book is self-contained, by providing all the necessary background in probabilities and random processes. It is important to note that if both deterministic and probabilistic aspects of communications are to be covered in one semester, it is highly desirable for students to have a solid background in probabilities.

Chapter 1 presents a panoramic view of communication systems by explaining important concepts of communication theory qualitatively and heuristically. Building on this momentum, students are motivated to study the signal analysis tools in Chapters 2 and 3, which describe a signal as a vector, and view the Fourier spectrum as a way of representing a signal in a well-known signal space. Chapter 4 discusses the traditional analog modulation and demodulation systems. Some instructors may feel that in this digital age, analog modulation should be removed altogether. We hold the view that modulation is not so much a method of communication as a basic tool of signal processing and transformation; it will always be needed, not only in the area of communication (digital or analog), but also in many other areas of engineering. Hence, fully neglecting modulation could prove to be shortsighted.

Chapter 5 serves as the fundamental bridge that connects analog and digital communication systems by covering the process of analog-to-digital (A/D) conversion for a variety of applications that include speech and video signals. Chapter 6 utilizes deterministic signal analysis tools to present the principles and techniques of digital modulations. It further introduces the concept of channel distortion and presents equalization as an effective means of distortion compensation.

Chapters 7 and 8 provide the essential background on theories of probability and random processes, tools that are essential to the performance analysis of digital communication systems. Every effort was made to motivate students and to guide them through these chapters by providing applications to communications problems wherever possible. Chapter 9 teaches the analysis and the design of digital communication systems in the presence of additive channel noise. It derives the optimum receiver structure based on the principle of minimizing error probability in signal detection. Chapter 10 focuses on the interference resilient spread spectrum communication systems. Chapter 11 presents various practical techniques that can be used to combat typical channel distortions. One major emphasis is on the popular OFDM (orthogonal frequency division modulation) that has found broad applications in state-of-the-art systems ranging from 4G-LTE cellular systems, IEEE 802.11a/g/n Wi-Fi networks, to DSL broadband services. Chapter 12 provides many fundamental concepts of information theory, including the basic principles of multiple-input–multiple-output (MIMO) technology that continues to gain practical acceptance and popularity. Finally, the principal and key practical aspects of error control coding are given in Chapter 13.

Course Adoption

With a combined teaching experience of over 60 years, we have taught communication classes under both quarter and semester systems in several major universities. On the other hand, the students' personal experiences with communication systems have continued to multiply, from a simple radio set in the 1960s, to the turn of the twenty-first century, with its easy access to Wi-Fi, cellular devices, satellite radio, and home Internet services. Hence, more and more students are interested in learning how familiar electronic gadgets work. With this important need and our past experiences in mind, we revised the fifth edition of this text to fit well within several different curriculum configurations. In all cases, basic coverage should always teach the fundamentals of analog and digital communications (Chapters 1–6).

Option A: One-Semester Course (without strong probability background)

In many existing curricula, undergraduate students are only exposed to very simple probability tools before they study communications. This occurs often because the students were required to take an introductory statistical course disconnected from engineering science. This text is well suited to students of such a background. Chapters 1–6 deliver a comprehensive coverage of modern digital and analog communication systems for average undergraduate engineering students. Such a course can be taught within one semester (in approximately 45 instructional hours). Under the premise that each student has built a solid background in Fourier analysis via a prerequisite class on *signals and systems*, most of the first three chapters can be treated as a review in a single week. The rest of the semester can be fully devoted to teaching Chapters 4—6, with selective coverage on the practical systems of Chapters 10 and 11 to broaden students' communication background.

Option B: One-Semester Course (with a strong probability background)

For students who have built a strong background on probability theory, a much more extensive coverage of digital communications can be achieved within one semester. A rigorous probability class can be taught within the context of signal and system analysis (cf. Cooper and McGillem, *Probabilistic Methods of Signal and System Analysis*, 3rd ed., Oxford University Press, 1998). Under this scenario, in addition to Chapters 1–6, Chapter 9 and part of Chapters 10–11 can also be taught in one semester, provided that the students have a solid probability background that permits covering Chapter 7 and Chapter 8 in a handful of hours. Students completing such a course would be well prepared to enter the telecommunications industry or to continue graduate studies.

Option C: Two-Semester Series (without a separate probability course)

The entire text can be thoroughly covered in two semesters for a curriculum that does not have any prior probability course. In other words, for a two-course series, the goal is to teach both communication systems and fundamentals of probabilities. In an era of many competing courses in a typical engineering curriculum, it is hard to set aside two-semester courses for communications alone. In this case, it would be desirable to fold probability theory into the two communication courses. Thus, for two-semester courses, the coverage can be as follows:

- 1st semester: Chapters 1–6 (Signals and Communication Systems)
- 2nd semester: Chapters 7–12 (Modern Digital Communication Systems)

Option D: One-Quarter Course (with a strong probability background)

In a quarter system, students must have prior exposure to probability and statistics at a rigorous level (cf. Cooper and McGillem, *Probabilistic Methods of Signal and System Analysis*, 3rd ed., Oxford University Press, 1998). They must also have solid knowledge of Fourier analysis (covered in Chapters 2 and 3). Within a quarter, the class can teach the basics of analog and digital communication systems (Chapters 3–6), analysis of digital communication systems (Chapter 9), and spread spectrum communications (Chapter 10).

Option E: One-Quarter Course (without a strong probability background)

In the rare case of students who come in without much probability knowledge, it is important to impart basic knowledge of communication systems. It is wise not to attempt to analyze digital communication systems. Instead, the basic coverage without prior knowledge of probability can be achieved by teaching the operations of analog and digital systems (Chapters 1–6) and a high-level discussion of spread spectrum wireless systems (Chapter 10).

Option F: Two-Quarter Series (with basic probability background)

Unlike a one-quarter course, a two-quarter series can be well designed to teach most of the important materials on communication systems and their analysis. The entire text can be extensively taught in two quarters for a curriculum that has some preliminary coverage of Fourier analysis and probabilities. Essentially treating Chapters 2, 3, and 7 partly as information review, the coverage can be as follows:

- 1st quarter: Chapters 1–8 (Communication Systems and Analysis)
- 2nd quarter: Chapters 9–12 (Digital Communication Systems)

MATLAB and Experiments

Since many institutions no longer have hardware communication laboratories, we provide MATLAB-based communication tests and design exercises to enhance the interactive learning experience. Students will be able to design systems and modify their parameters to evaluate the overall effects on the performance of communication systems through signal displays and bit error rate measurement. The students will acquire first-hand knowledge of how to design and test communication systems. To assist the instructors, computer assignment problems are suggested for most chapters in this edition.

Acknowledgments

First, the authors would like to thank all the students and teaching assistants they have worked with over the many years of teaching. This edition would not have been possible without much feedback from, and many discussions with, our students. The authors thank all the reviewers for providing invaluable inputs to improve the text. Finally, the authors also wish to thank many fellow instructors for their helpful comments regarding the last edition.

B. P. Lathi, Carmichael, California, USA
Zhi Ding, Davis, California, USA

1 INTRODUCTION

Let's face it. Our world has been totally transformed by recent advances in communication and information technologies. Specifically in the past 20 years, we have witnessed an explosive growth of communication applications ranging from Internet to Bluetooth hand-free devices. In particular, smartphones and smartphone applications have made information technologies and Internet fully accessible to people of every age group on every continent almost ubiquitously. In less than a decade, wireless communication technologies have completely transformed the world economy and people's lives in more ways than imaginable at the beginning of this millennium. Globally, it is quite difficult to find an individual in any part of the world today that has not been touched by new communication technologies ranging from e-commerce to online social media. This book teaches the basic principles of communication systems based on electrical signals.

Before modern times, messages were carried by runners, homing pigeons, lights, and smoke signals. These schemes were adequate for the distances and "data rates" of the age. In most parts of the world, these modes of communication have been superseded by electrical communication systems,* which can transmit signals over vast distances (even to distant planets and galaxies) and at the speed of light.

Modern electronic communication systems are more dependable and more economical, often playing key roles in improving productivity and energy efficiency. Increasingly, businesses are conducted electronically, saving both time and energy over traditional means. Ubiquitous communication allows real-time management and coordination of project participants from around the globe. E-mail is rapidly replacing the more costly and slower "snail mail." E-commerce has also drastically reduced costs and delays associated with marketing and transactions, allowing customers to be much better informed about new products and to complete online transactions with a click. Traditional media outlets such as television, radio, and newspapers have also been rapidly evolving in recent years to cope with and better utilize new communication and networking technologies. Furthermore, communication technologies have been, and will always be, playing an important role in current and future waves of remarkable technological advances in artificial intelligence, data mining, and machine learning.

The goal of this textbook is to provide the fundamental technical knowledge needed by future-generation communication engineers and technologists for designing even more efficient and more powerful communication systems of tomorrow. Critically, one major objective of this book is to answer the question: How do communication systems work? That is, how can we access information remotely using small devices such as a smartphone? Being able to answer this question is essential to designing better communication systems for the future.

* With the exception of the postal service.

1.1 COMMUNICATION SYSTEMS

Figure 1.1 presents three familiar communication scenarios: a wireline telephone-to-cellular phone connection, a TV broadcasting system, and a computer network. Because of the numerous examples of communication systems in existence, it would be unwise to attempt to study the details of all kinds of communication systems in this book. Instead, the most efficient and effective way to learn is by studying the major functional blocks common to practically all communication systems. This way, we are not merely learning the mechanics of those existing systems under study. More importantly, we can acquire the basic knowledge needed to design and analyze new systems never encountered in a textbook. To begin, it is essential to establish a typical communication system model as shown in Fig. 1.2. The key components of a communication system are as follows.

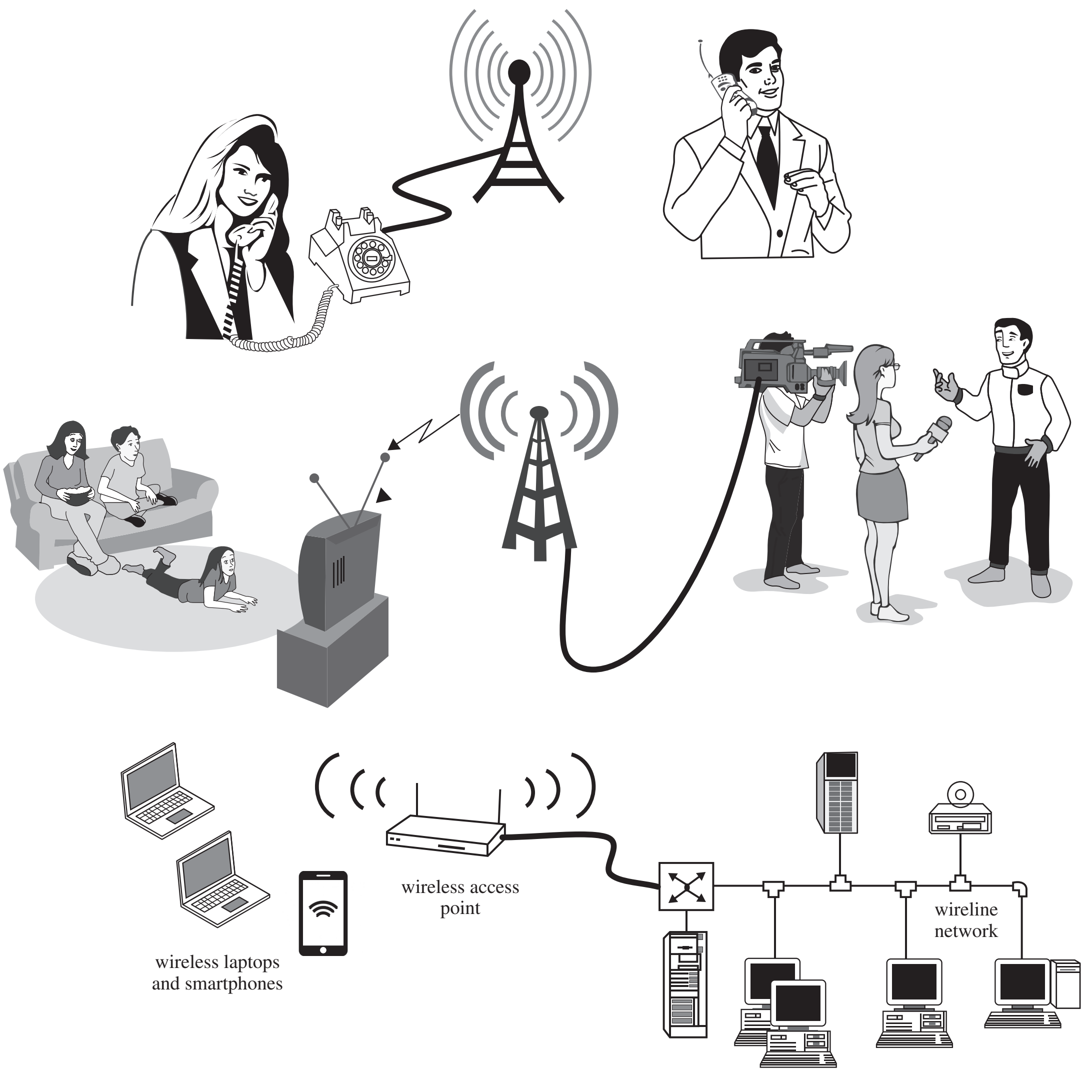


Figure 1.1 Some examples of communication systems.

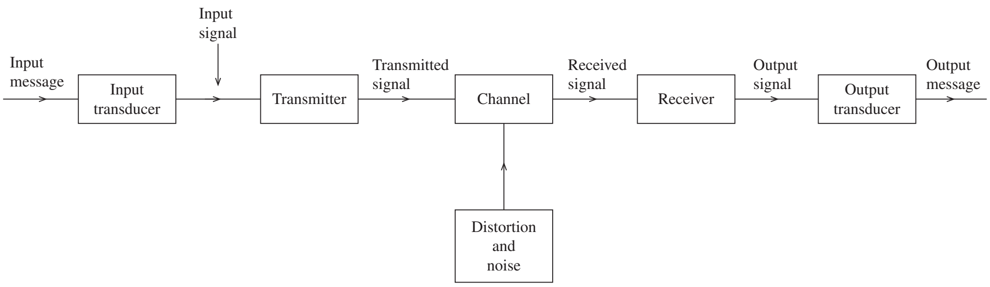


Figure 1.2 Communication system.

The **source** originates a message, such as a human voice, a television picture, an e-mail message, or data. If the data is nonelectric (e.g., human voice, e-mail text, a scene), it must be converted by an **input transducer** into an electric waveform referred to as the **message signal** through physical devices such as a microphone, a computer keyboard, or a charge-coupled device (CCD) camera.

The **transmitter** transforms the input (message) signal into an appropriate form for efficient transmission. The transmitter may consist of one or more subsystems: an analog-to-digital (A/D) converter, an encoder, and a modulator. Similarly, the receiver may consist of a demodulator, a decoder, and a digital-to-analog (D/A) converter.

The **channel** is a medium of choice that can convey the electric signals at the transmitter output over a distance. A typical channel can be a pair of twisted copper wires (e.g., in telephone and DSL), coaxial cable (e.g. in television and Internet), an optical fiber, or a radio cellular link. Additionally, a channel can also be a point-to-point connection in a mesh of interconnected channels that form a communication network.

The **receiver** reprocesses the signal received from the channel by reversing the signal transformation made at the transmitter and removing the distortions caused by the channel. The receiver output is passed to the **output transducer**, which converts the electric signal to its original form—the message.

The **destination** is the unit where the message transmission terminates.

1.2 DESIGN CHALLENGES: CHANNEL DISTORTIONS AND NOISES

A channel is a physical medium that behaves practically like an imperfect filter that generally attenuates the signal and distorts the transmitted waveforms. The channel attenuation depends on the distance the signals must travel between the transmitter and the receiver, varying from mild to severe. Signal waveforms are further distorted because of physical phenomena such as frequency-dependent electronics, multipath effects, and Doppler shift. For example, a *frequency-selective* channel causes different amounts of attenuation and phase shift to different frequency components within the input signal. A short rectangular pulse can be rounded or “spread out” during transmission over a lowpass channel. These types of distortion, called **linear distortion**, can be partly corrected at the receiver by an equalizer with gain and phase characteristics complementary to those of the channel. Channels may also cause **nonlinear distortion** through attenuation that varies with the signal amplitude. Such distortions can also be partly mitigated by a complementary equalizer at the receiver. Channel distortions, if known, can also be precompensated by transmitters using channel-dependent predistortions.

In a practical environment, signals passing through communication channels not only experience channel distortions but also are corrupted along the path by interfering signals and disturbances lumped under the broad term **noise**. These interfering signals are often random and unpredictable from sources both external and internal. External noise includes interference signals transmitted on nearby channels, human-made noise generated by faulty switch contacts of electric equipment, automobile ignition radiation, fluorescent lights or natural noise from lightning, microwave ovens, and cellphone emissions, as well as electric storms and solar or intergalactic radiation. With proper care in system designs, external noise can be minimized or even eliminated in some cases. Internal noise results from thermal motion of charged particles in conductors, random emission, and diffusion or recombination of charged carriers in electronic devices. Proper care can mitigate the effect of internal noise but can never fully eliminate it. Noise is one of the underlying factors that limit the rate of telecommunications.

Thus in practical communication systems, the channel distorts the signal, and noise accumulates along the path. Worse yet, the signal strength attenuates while the noise level remains steady regardless of the distance from the transmitter. Thus, the signal quality would continuously degrade along the length of the channel. Amplification of the received signal to make up for the attenuation is ineffective because the noise will be amplified by the same proportion, and the quality remains, at best, unchanged.* These are the key challenges that we must face in designing modern communication systems.

1.3 MESSAGE SOURCES

Messages in communication systems can be either digital or analog. Digital messages are ordered combinations of finite symbols or codewords. For example, printed English consists of 26 letters, 10 numbers, a space, and several punctuation marks. Thus, a text document written in English is a digital message constructed from the ASCII keyboard of 128 symbols. Analog messages, on the other hand, are characterized by signals whose values vary over a continuous range and are defined for a continuous range of time. For example, the temperature or the atmospheric pressure of a certain location over time can vary over a continuous range and can assume an (uncountably) infinite number of possible values. An analog message typically has a limited range of amplitude and power. A digital message typically contains M symbols and is called an **M-ary** message.

The difference between digital and analog messages can be subtle. For example, the text in a speech is a digital message, since it is made up from a finite vocabulary in a language. However, the actual recorded voice from a human speaker reading the text is an analog waveform whose amplitude varies over a continuous range. Similarly, a musical note is a digital message, consisting of a finite number of musical symbols. The same musical note, when played by a musician, becomes an audio waveform that is an analog signal.

1.3.1 The Digital Revolution in Communications

It is no secret to even a casual observer that every time one looks at the latest electronic communication products, another newer and better "digital technology" is displacing the old analog technology. Between 1990 and 2015, cellular networks completed their transformation from

---

* Actually, amplification may further deteriorate the signal because of additional amplifier noise.

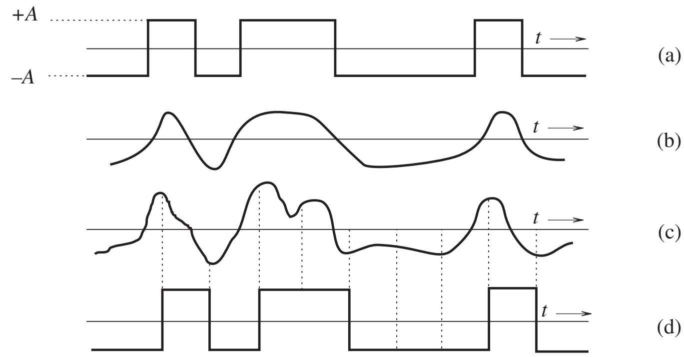


Figure 1.3
(a) Transmitted signal.
(b) Received distorted signal (without noise).
(c) Received distorted signal (with noise).
(d) Regenerated signal (delayed).

the first-generation analog AMPS to the current third-generation (UMTS, CDMA2000) and fourth-generation (i.e., 4G-LTE) digital offsprings. Most visibly in every household, digital video technology (DVD) and Blu-ray have made the analog VHS cassette systems obsolete. Digital iPod and MP3 players have totally vanquished the once popular audio-cassette players in consumer electronics. The global conversion to digital television is now nearly complete in driving out the last analog holdout of color television. This begs the question: Why are digital technologies superior? The answer has to do with both economics and quality. The case for economics is made by the ease of adopting versatile, powerful, and inexpensive high-speed digital microprocessors. But more importantly at the quality level, one prominent feature of digital communications is the enhanced immunity of digital signals to noise and interferences.

Digital messages are transmitted as a finite set of electrical waveforms. In other words, a digital message is generated from a finite alphabet, while each character in the alphabet can be represented by one waveform or a sequential combination of such waveforms. For example, in sending messages via Morse code, a dash can be transmitted by an electrical pulse of amplitude A and a dot can be transmitted by a pulse of negative amplitude $-A$ (Fig. 1.3a). In an M-ary case, M distinct electrical pulses (or waveforms) are used; each of the M pulses represents one of the M possible symbols. Once transmitted, the receiver must extract the message from a distorted and noisy signal at the channel output. Message extraction is often easier from digital signals than from analog signals because the digital decision must belong to the finite-sized alphabet. Consider a binary case: two symbols are encoded as rectangular pulses of amplitudes A and $-A$. The only decision at the receiver is to select between two possible pulses received; the fine details of the pulse shape are not an issue. A finite alphabet leads to noise and interference immunity. The receiver's decision can be made with reasonable certainty even if the pulses have suffered from modest distortion and noise (Fig. 1.3). The digital message in Fig. 1.3a is distorted by the channel, as shown in Fig. 1.3b. Yet, if the distortion is not too large, we can recover the data without error because we only need to make a simple binary decision: Is the received pulse positive or negative? Figure 1.3c shows the same data with channel distortion and noise. Here again, the data can be recovered correctly as long as the distortion and the noise are within limits. In contrast, the waveform shape itself in an analog message carries the needed information, and even a slight distortion or interference in the waveform will show up in the received signal. Clearly, a digital communication system is more rugged than an analog communication system in the sense that it can better withstand noise and distortion (as long as they are within a limit).

A typical distorted binary signal with noise acquired over the channel is shown in Fig. 1.3c. If A is sufficiently large in comparison to typical noise amplitudes, the receiver can still correctly distinguish between the two pulses. The pulse amplitude is typically 5 to 10 times the average noise amplitude. For such a high signal-to-noise ratio (SNR) the probability of error at the receiver is less than 10^{-6}; that is, on the average, the receiver will make fewer than one error per million pulses. The effect of random channel noise and distortion is thus practically eliminated.

1.3.2 Distortionless Regeneration of Digital Signals

One main reason for the superior quality of digital systems over analog ones is the viability of signal **regeneration** by repeaters and relay nodes. When directly communicating over a long distance, transmitted signals can be severely attenuated and distorted. For digital pulse signals used in digital communications, repeater nodes can be placed along the communication path at distances short enough to ensure that noise and distortion effects are minor such that digital pulses can be detected with high accuracy. At each repeater or relay node, the incoming digital pulses are detected such that new, “clean” pulses are regenerated for transmission to the next node along the path. This process prevents the accumulation of noise and distortion along the path by cleaning up the pulses at regular path intervals. We can thus transmit messages over longer distances with greater accuracy. There has been widespread application of distortionless regeneration by repeaters in long-haul communication systems or by nodes in a large (possibly heterogeneous) network. The same argument applies when making copies of digital content.

In analog systems, however, signals and noise within the same bandwidth cannot be separated. Repeaters in analog systems are basically filters plus amplifiers and are not “regenerative.” It is therefore impossible to avoid in-band accumulation of noise and distortion along the path. As a result, the distortion and the noise interference can accumulate over the entire long-distance path as a signal traverses through the network. To compound the problem, the signal is also attenuated continuously over the transmission path. Thus, with increasing distance the signal becomes weaker, whereas more distortions and the noise accumulate to greater strength. Ultimately, the signal, weakened and overwhelmed by the cumulative distortions and noises, is buried beyond recognition. Amplification offers little help, since it enhances both the signal and the noise equally. Consequently, the distance over which an analog message can be successfully received is limited by the first transmitter power. Despite these limitations, analog communication is simpler and was used widely and successfully in the past for short- to medium-range communications. In modern times, however, almost all new communication systems being installed are digital, although a small number of old analog communication technologies are still in use, such as those for AM and FM radio broadcasting.

1.3.3 Analog-to-Digital (A/D) Conversion for Digital Communications

Despite the differences between analog and digital messages, digital communication systems can carry analog messages by first converting analog signals to digital signals. A key device in electronics, the analog-to-digital (A/D) converter, enables digital communication systems to convey analog source signals such as audio and video. Generally, analog signals are continuous in time and in range; that is, they have values at every time instant, and their values can be anywhere within the range. On the other hand, digital signals exist only at discrete points of time, and they can take on only finite values. A/D conversion can never be 100% accurate. Fortunately, since human perception does not require infinite accuracy, A/D conversion can effectively capture necessary information from the analog source for digital signal transmission.

Two steps take place in A/D conversion: a continuous time signal is first *sampled* into a discrete time signal, whose continuous amplitude is then *quantized* into a discrete level signal. First, the frequency spectrum of a signal indicates relative strengths of various frequency components. The **sampling theorem** (Chapter 5) states that if the highest frequency in the signal spectrum is B (in hertz), the signal can be reconstructed from its discrete samples, taken uniformly at a rate above $2B$ samples per second. This means that to preserve the information from a continuous-time signal, we only need to transmit its samples. However, the sample values are still not digital because they lie in a continuous dynamic range. Here, the second step of **quantization** comes to the rescue. Through quantization, each sample is approximated, or "rounded off," to the nearest quantized level. Since human perception has only limited sensitivity, quantization with sufficient granularity does not compromise the signal quality.

A quantizer partitions the signal range into L intervals. Each sample amplitude is approximated by the midpoint of the interval in which the sample value falls. Each sample is now represented by one of the L numbers. The information is thus digitized. Hence, after the two steps of sampling and quantizing, the A/D conversion is completed. The quantized signal is an approximation of the original one. We can improve the accuracy of the quantized signal to any desired level by increasing the number of levels L.

1.3.4 Pulse-Coded Modulation—A Digital Representation

Once the A/D conversion is over, the original analog message is represented by a sequence of samples, each of which takes on one of the L preset quantization levels. The transmission of this quantized sequence is the task of digital communication systems. For this reason, signal waveforms must be used to represent the quantized sample sequence in the transmission process. Similarly, a digital storage device would also need to represent the samples as signal waveforms. *Pulse-coded modulation* (PCM) is a very simple and yet common mechanism for this purpose.

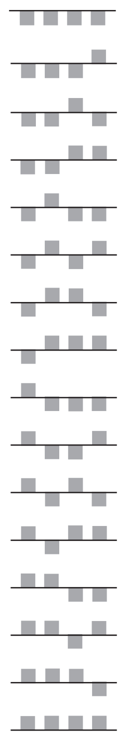
First, one information *bit* refers to one *binary digit* of **1** or **0**. The idea of PCM is to represent each quantized sample by an ordered combination of two basic pulses: $p_1(t)$ representing **1** and $p_0(t)$ representing **0**. Because each of the L possible sample values can be written as a bit string of length $\log_2 L$, each sample can therefore also be mapped into a short pulse sequence to represent $\log_2 L$ bits. For example, if $L = 16$, then, each quantized level can be described uniquely by 4 bits. If we use two basic pulses $p_1(t) = A$ and $p_0(t) = -A$, respectively, to represent 1 and 0 for each bit, then a sequence of four such pulses gives $2 \times 2 \times 2 \times 2 = 16$ distinct patterns, as shown in Fig. 1.4. We can assign one pattern to each of the 16 quantized values to be transmitted. Each quantized sample is now coded into a sequence of four binary pulses. This is the principle of PCM transmission, where signaling is carried out by means of only two basic pulses (or symbols). The binary case is of great practical importance because of its simplicity and ease of detection. Much of today's digital communication is binary.*

Although PCM was invented by P. M. Rainey in 1926 and rediscovered by A. H. Reeves in 1939, it was not until the early 1960s that the Bell System installed the first communication link using PCM for digital voice transmission. The cost and size of vacuum tube circuits

* An intermediate case exists where we use four basic pulses (quaternary pulses) of amplitudes $\pm A$ and $\pm 3A$. A sequence of two quaternary pulses can form $4 \times 4 = 16$ distinct levels of values.

Figure 1.4 Example of PCM encoding.

Digit	Binary equivalent	Pulse code waveform
0	**0000**	
1	**0001**	
2	**0010**	
3	**0011**	
4	**0100**	
5	**0101**	
6	**0110**	
7	**0111**	
8	**1000**	
9	**1001**	
10	**1010**	
11	**1011**	
12	**1100**	
13	**1101**	
14	**1110**	
15	**1111**	



were the chief impediments to PCM in the early days before the discovery of semiconductor devices. It was the transistor that made PCM practical.

From all these discussions on PCM, we arrive at a rather interesting (and to a certain extent not obvious) conclusion—that every possible communication can be carried on with a minimum of two symbols. Thus, merely by using a proper sequence of a wink of the eye, one can convey any message, be it a conversation, a book, a movie, or an opera. Every possible detail (such as various shades of colors of the objects and tones of the voice, etc.) that is reproducible on a movie screen or on the high-definition color television can be conveyed with no less accuracy, merely by winks of an eye*.

1.4 CHANNEL EFFECT, SIGNAL-TO-NOISE RATIO, AND CAPACITY

In designing communication systems, it is vital to understand and analyze important factors such as the channel and signal characteristics, the relative noise strength, the maximum number of bits that can be sent over a channel per second, and, ultimately, the signal quality.

* Of course, to convey the information in a movie or a television program in real time, the winking would have to be at an inhumanly high rate. For example, the HDTV signal is represented by 19 million bits (winks) per second.

1.4.1 Signal Bandwidth and Power

In a given communication system, the fundamental parameters and physical limitations that control the connection's rate and quality are the channel bandwidth B and the signal power P_s. Their precise and quantitative relationships will be discussed in Chapter 12. Here, we shall demonstrate these relationships qualitatively.

The **bandwidth** of a channel is the range of frequencies that it can carry with reasonable fidelity. For example, if a channel can carry with reasonable fidelity a signal whose frequency components vary from 0 Hz (dc) up to a maximum of 5000 Hz (5 kHz), the channel bandwidth B is 5 kHz. Likewise, each signal also has a bandwidth that measures the maximum range of its frequency components. The faster a signal changes, the higher its maximum frequency is, and the larger its bandwidth is. Signals rich in content with quick changes (such as those for battle scenes in a video) have larger bandwidth than signals that are dull and vary slowly (such as those for a daytime soap opera or a video of sleeping lions). A signal transmission is likely successful over a channel if the channel bandwidth exceeds the signal bandwidth.

To understand the role of B, consider the possibility of increasing the speed of information transmission by compressing the signal in time. Compressing a signal in time by a factor of 2 allows it to be transmitted in half the time, and the transmission speed (rate) doubles. Time compression by a factor of 2, however, causes the signal to "wiggle" twice as fast, implying that the frequencies of its components are doubled. Many people have had firsthand experience of this effect when playing a piece of audiotape twice as fast, making the voices of normal people sound like the high-pitched speech of cartoon characters. Now, to transmit this compressed signal without distortion, the channel bandwidth must also be doubled. Thus, the rate of information transmission that a channel can successfully carry is directly proportional to B. More generally, if a channel of bandwidth B can transmit N pulses per second, then to transmit KN pulses per second by means of the same technology, we need a channel of bandwidth KB. To reiterate, the number of pulses per second that can be transmitted over a channel is directly proportional to its bandwidth B.

The **signal power** P_s plays a dual role in information transmission. First, P_s is related to the quality of transmission. Increasing P_s strengthens the signal pulse and suppresses the effect of channel noise and interference. In fact, the quality of either analog or digital communication systems varies with the SNR. In any event, a certain minimum SNR at the receiver is necessary for successful communication. Thus, a larger signal power P_s allows the system to maintain a minimum SNR over a longer distance, thereby enabling successful communication over a longer span. The second role of the signal power is less obvious, although equally important. From the information theory point of view, the channel bandwidth B and the signal power P_s are, to some extent, exchangeable; that is, to maintain a given rate and accuracy of information transmission, we can trade P_s for B, and vice versa. Thus, one may use less B if one is willing to increase P_s, or one may reduce P_s if one is given bigger B. The rigorous proof of this will be provided in Chapter 12.

In short, the two primary resources in communication are the bandwidth and the transmit power. Facing a specific communication channel, one resource may be more valuable than the other, and the communication scheme should be designed accordingly. A typical telephone channel, for example, has a limited bandwidth (3 kHz), but the transmit power is less restrictive. On the other hand, in deep-space explorations, huge bandwidth is available but the transmit power is severely limited. Hence, the communication solutions in the two cases are radically different.