1

Introduction

本书介绍作为数字通信系统分析和设计基础的基本原理。数字通信的研究内容包括数字形式的信息从产生该信息的信源到一个或多个目的地的传输问题。在通信系统分析和设计中，特别重要的是信息传输所通过物理信道的特征，信道特征一般影响通信系统基本组成部分的设计。下面将阐述通信系统的要素及其功能。

In this book, we present the basic principles that underlie the analysis and design of digital communication systems. The subject of digital communications involves the transmission of information in digital form from a source that generates the information to one or more destinations. Of particular importance in the analysis and design of communication systems are the characteristics of the physical channels through which the information is transmitted. The characteristics of the channel generally affect the design of the basic building blocks of the communication system. Below, we describe the elements of a communication system and their functions.

1.1
ELEMENTS OF A DIGITAL COMMUNICATION SYSTEM

Figure 1.1–1 illustrates the functional diagram and the basic elements of a digital communication system. The source output may be either an analog signal, such as an audio or video signal, or a digital signal, such as the output of a computer, that is discrete in time and has a finite number of output characters. In a digital communication system, the messages produced by the source are converted into a sequence of binary digits. Ideally, we should like to represent the source output (message) by as few binary digits as possible. In other words, we seek an efficient representation of the source output that results in little or no redundancy. The process of efficiently converting the output of either an analog or digital source into a sequence of binary digits is called *source encoding* or *data compression*.

The sequence of binary digits from the source encoder, which we call the *information sequence*, is passed to the *channel encoder*. The purpose of the channel encoder is to introduce, in a controlled manner, some redundancy in the binary information sequence that can be used at the receiver to overcome the effects of noise and interference encountered in the transmission of the signal through the channel. Thus, the added redundancy serves to increase the reliability of the received data and improves
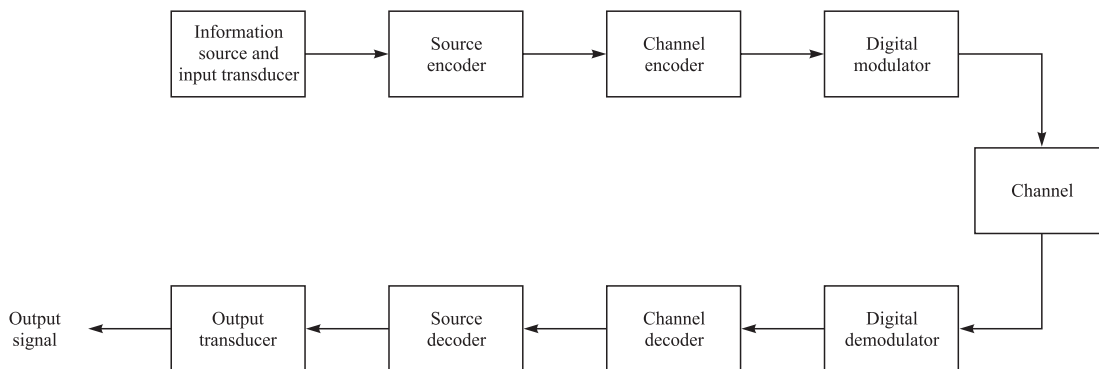
FIGURE 1.1–1
Basic elements of a digital communication system.

the fidelity of the received signal. In effect, redundancy in the information sequence aids the receiver in decoding the desired information sequence. For example, a (trivial) form of encoding of the binary information sequence is simply to repeat each binary digit m times, where m is some positive integer. More sophisticated (nontrivial) encoding involves taking k information bits at a time and mapping each k-bit sequence into a unique n-bit sequence, called a *code word*. The amount of redundancy introduced by encoding the data in this manner is measured by the ratio n/k. The reciprocal of this ratio, namely k/n, is called the rate of the code or, simply, the *code rate*.

The binary sequence at the output of the channel encoder is passed to the *digital modulator*, which serves as the interface to the communication channel. Since nearly all the communication channels encountered in practice are capable of transmitting electrical signals (waveforms), the primary purpose of the digital modulator is to map the binary information sequence into signal waveforms. To elaborate on this point, let us suppose that the coded information sequence is to be transmitted one bit at a time at some uniform rate R bits per second (bits/s). The digital modulator may simply map the binary digit 0 into a waveform $s_0(t)$ and the binary digit 1 into a waveform $s_1(t)$. In this manner, each bit from the channel encoder is transmitted separately. We call this *binary modulation*. Alternatively, the modulator may transmit b coded information bits at a time by using $M = 2^b$ distinct waveforms $s_i(t)$, $i = 0, 1, \ldots, M - 1$, one waveform for each of the 2^b possible b-bit sequences. We call this *M-ary modulation* ($M > 2$). Note that a new b-bit sequence enters the modulator every b/R seconds. Hence, when the channel bit rate R is fixed, the amount of time available to transmit one of the M waveforms corresponding to a b-bit sequence is b times the time period in a system that uses binary modulation.

The *communication channel* is the physical medium that is used to send the signal from the transmitter to the receiver. In wireless transmission, the channel may be the atmosphere (free space). On the other hand, telephone channels usually employ a variety of physical media, including wire lines, optical fiber cables, and wireless (microwave radio). Whatever the physical medium used for transmission of the information, the essential feature is that the transmitted signal is corrupted in a random manner by a

variety of possible mechanisms, such as additive *thermal noise* generated by electronic devices; man-made noise, e.g., automobile ignition noise; and atmospheric noise, e.g., electrical lightning discharges during thunderstorms.

At the receiving end of a digital communication system, the *digital demodulator* processes the channel-corrupted transmitted waveform and reduces the waveforms to a sequence of numbers that represent estimates of the transmitted data symbols (binary or *M*-ary). This sequence of numbers is passed to the channel decoder, which attempts to reconstruct the original information sequence from knowledge of the code used by the channel encoder and the redundancy contained in the received data.

A measure of how well the demodulator and decoder perform is the frequency with which errors occur in the decoded sequence. More precisely, the average probability of a bit-error at the output of the decoder is a measure of the performance of the demodulator–decoder combination. In general, the probability of error is a function of the code characteristics, the types of waveforms used to transmit the information over the channel, the transmitter power, the characteristics of the channel (i.e., the amount of noise, the nature of the interference), and the method of demodulation and decoding. These items and their effect on performance will be discussed in detail in subsequent chapters.

As a final step, when an analog output is desired, the source decoder accepts the output sequence from the channel decoder and, from knowledge of the source encoding method used, attempts to reconstruct the original signal from the source. Because of channel decoding errors and possible distortion introduced by the source encoder, and perhaps, the source decoder, the signal at the output of the source decoder is an approximation to the original source output. The difference or some function of the difference between the original signal and the reconstructed signal is a measure of the distortion introduced by the digital communication system.

1.2
COMMUNICATION CHANNELS AND THEIR CHARACTERISTICS

As indicated in the preceding discussion, the communication channel provides the connection between the transmitter and the receiver. The physical channel may be a pair of wires that carry the electrical signal, or an optical fiber that carries the information on a modulated light beam, or an underwater ocean channel in which the information is transmitted acoustically, or free space over which the information-bearing signal is radiated by use of an antenna. Other media that can be characterized as communication channels are data storage media, such as magnetic tape, magnetic disks, and optical disks.

One common problem in signal transmission through any channel is additive noise. In general, additive noise is generated internally by components such as resistors and solid-state devices used to implement the communication system. This is sometimes called *thermal noise*. Other sources of noise and interference may arise externally to the system, such as interference from other users of the channel. When such noise and interference occupy the same frequency band as the desired signal, their effect can be minimized by the proper design of the transmitted signal and its demodulator at

the receiver. Other types of signal degradations that may be encountered in transmission over the channel are signal attenuation, amplitude and phase distortion, and multipath distortion.

The effects of noise may be minimized by increasing the power in the transmitted signal. However, equipment and other practical constraints limit the power level in the transmitted signal. Another basic limitation is the available channel bandwidth. A bandwidth constraint is usually due to the physical limitations of the medium and the electronic components used to implement the transmitter and the receiver. These two limitations constrain the amount of data that can be transmitted reliably over any communication channel as we shall observe in later chapters. Below, we describe some of the important characteristics of several communication channels.

Wireline Channels

The telephone network makes extensive use of wire lines for voice signal transmission, as well as data and video transmission. Twisted-pair wire lines and coaxial cable are basically guided electromagnetic channels that provide relatively modest bandwidths. Telephone wire generally used to connect a customer to a central office has a bandwidth of several hundred kilohertz (kHz). On the other hand, coaxial cable has a usable bandwidth of several megahertz (MHz). Figure 1.2–1 illustrates the frequency range of guided electromagnetic channels, which include waveguides and optical fibers.

Signals transmitted through such channels are distorted in both amplitude and phase and further corrupted by additive noise. Twisted-pair wireline channels are also prone to crosstalk interference from physically adjacent channels. Because wireline channels carry a large percentage of our daily communications around the country and the world, much research has been performed on the characterization of their transmission properties and on methods for mitigating the amplitude and phase distortion encountered in signal transmission. In Chapter 9, we describe methods for designing optimum transmitted signals and their demodulation; in Chapter 10, we consider the design of channel equalizers that compensate for amplitude and phase distortion on these channels.

Fiber-Optic Channels

Optical fibers offer the communication system designer a channel bandwidth that is several orders of magnitude larger than coaxial cable channels. During the past two decades, optical fiber cables have been developed that have a relatively low signal attenuation, and highly reliable photonic devices have been developed for signal generation and signal detection. These technological advances have resulted in a rapid deployment of optical fiber channels, both in domestic telecommunication systems as well as for transcontinental communication. With the large bandwidth available on fiber-optic channels, it is possible for telephone companies to offer subscribers a wide array of telecommunication services, including voice, data, facsimile, and video.

The transmitter or modulator in a fiber-optic communication system is a light source, either a light-emitting diode (LED) or a laser. Information is transmitted by varying (modulating) the intensity of the light source with the message signal. The light propagates through the fiber as a light wave and is amplified periodically (in the case of
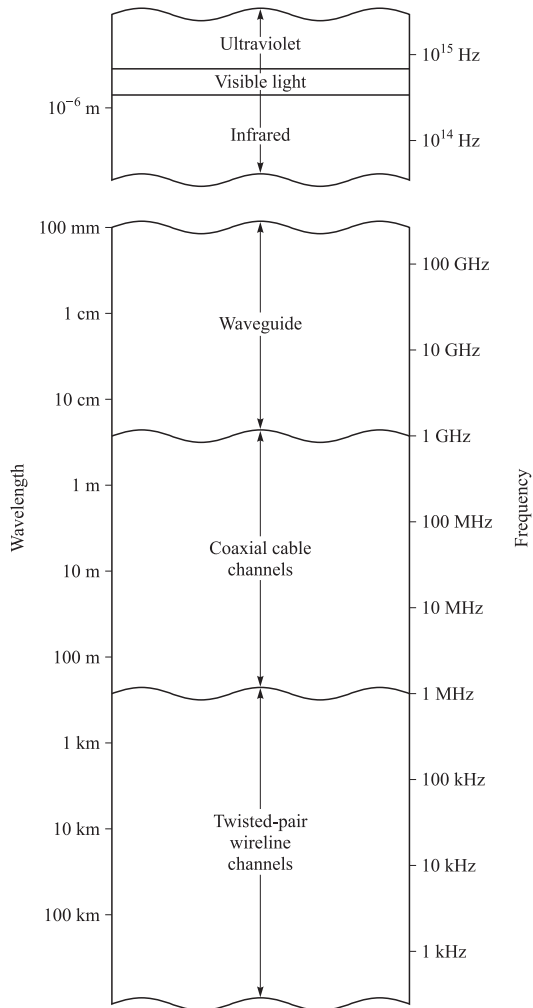
FIGURE 1.2–1
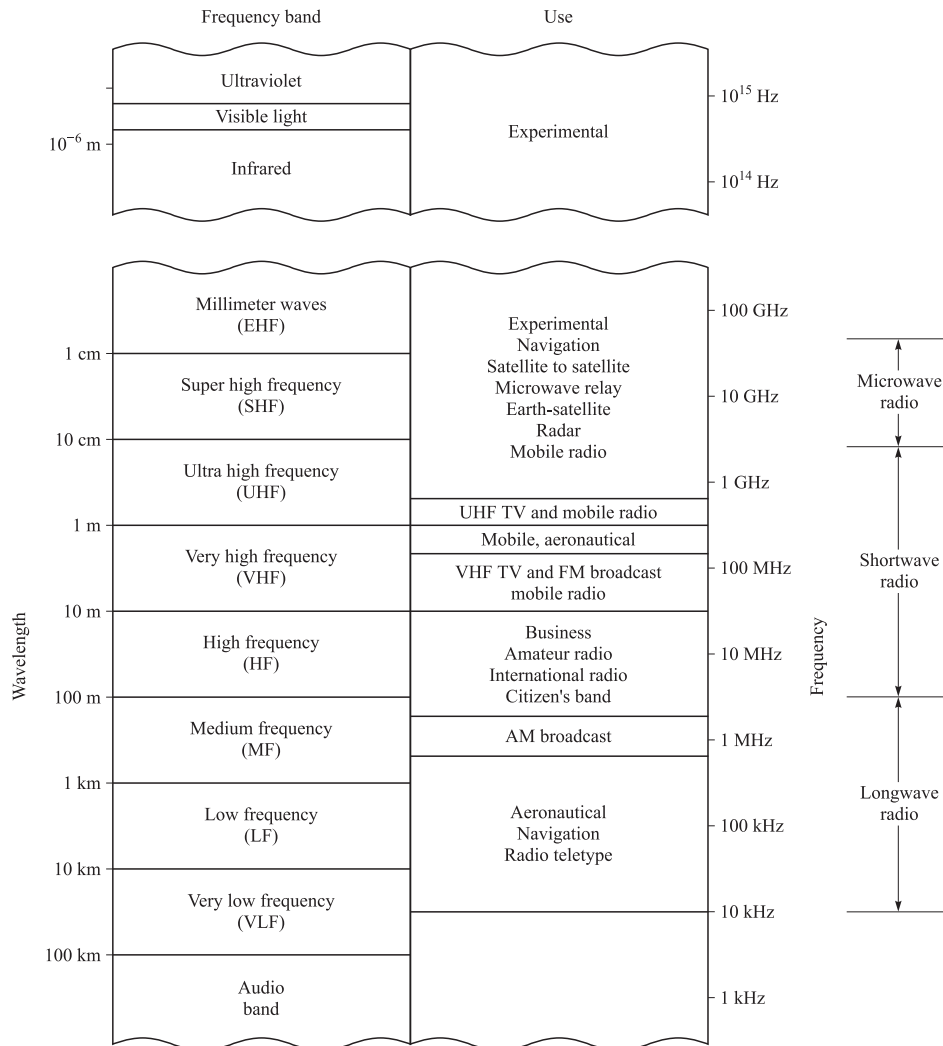Frequency range for guided wire channel.

digital transmission, it is detected and regenerated by repeaters) along the transmission path to compensate for signal attenuation. At the receiver, the light intensity is detected by a photodiode, whose output is an electrical signal that varies in direct proportion to the power of the light impinging on the photodiode. Sources of noise in fiber-optic channels are photodiodes and electronic amplifiers.

### Wireless Electromagnetic Channels

In wireless communication systems, electromagnetic energy is coupled to the propagation medium by an antenna which serves as the radiator. The physical size and the configuration of the antenna depend primarily on the frequency of operation. To obtain efficient radiation of electromagnetic energy, the antenna must be longer than

$\frac{1}{10}$ of the wavelength. Consequently, a radio station transmitting in the amplitude-modulated (AM) frequency band, say at $f_c = 1$ MHz [corresponding to a wavelength of $\lambda = c/f_c = 300$ meters (m)], requires an antenna of at least 30 m. Other important characteristics and attributes of antennas for wireless transmission are described in Chapter 4.

Figure 1.2–2 illustrates the various frequency bands of the electromagnetic spectrum. The mode of propagation of electromagnetic waves in the atmosphere and in

FIGURE 1.2–2
Frequency range for wireless electromagnetic channels. [*Adapted from Carlson* (1975), 2nd edition, © *McGraw-Hill Book Company Co. Reprinted with permission of the publisher.*]
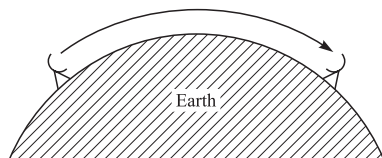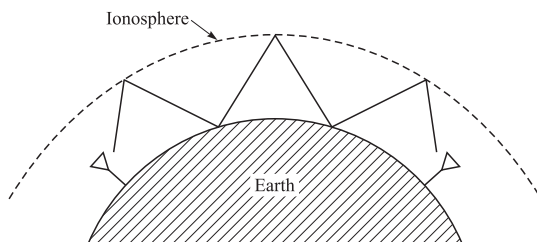
FIGURE 1.2–3
Illustration of ground-wave propagation.

free space may be subdivided into three categories, namely, ground-wave propagation, sky-wave propagation, and line-of-sight (LOS) propagation. In the very low frequency (VLF) and audio frequency bands, where the wavelengths exceed 10 km, the earth and the ionosphere act as a waveguide for electromagnetic wave propagation. In these frequency ranges, communication signals practically propagate around the globe. For this reason, these frequency bands are primarily used to provide navigational aids from shore to ships around the world. The channel bandwidths available in these frequency bands are relatively small (usually 1–10 percent of the center frequency), and hence the information that is transmitted through these channels is of relatively slow speed and generally confined to digital transmission. A dominant type of noise at these frequencies is generated from thunderstorm activity around the globe, especially in tropical regions. Interference results from the many users of these frequency bands.

Ground-wave propagation, as illustrated in Figure 1.2–3, is the dominant mode of propagation for frequencies in the medium frequency (MF) band (0.3–3 MHz). This is the frequency band used for AM broadcasting and maritime radio broadcasting. In AM broadcasting, the range with ground-wave propagation of even the more powerful radio stations is limited to about 150 km. Atmospheric noise, man-made noise, and thermal noise from electronic components at the receiver are dominant disturbances for signal transmission in the MF band.

Sky-wave propagation, as illustrated in Figure 1.2–4, results from transmitted signals being reflected (bent or refracted) from the ionosphere, which consists of several layers of charged particles ranging in altitude from 50 to 400 km above the surface of the earth. During the daytime hours, the heating of the lower atmosphere by the sun causes the formation of the lower layers at altitudes below 120 km. These lower layers, especially the D-layer, serve to absorb frequencies below 2 MHz, thus severely limiting sky-wave propagation of AM radio broadcast. However, during the nighttime hours, the electron density in the lower layers of the ionosphere drops sharply and the frequency absorption that occurs during the daytime is significantly reduced. As a consequence, powerful AM radio broadcast stations can propagate over large distances via sky wave over the F-layer of the ionosphere, which ranges from 140 to 400 km above the surface of the earth.

FIGURE 1.2–4
Illustration of sky-wave propagation.

A frequently occurring problem with electromagnetic wave propagation via sky wave in the high frequency (HF) range is *signal multipath*. Signal multipath occurs when the transmitted signal arrives at the receiver via multiple propagation paths at different delays. It generally results in intersymbol interference in a digital communication system. Moreover, the signal components arriving via different propagation paths may add destructively, resulting in a phenomenon called *signal fading*, which most people have experienced when listening to a distant radio station at night when sky wave is the dominant propagation mode. Additive noise in the HF range is a combination of atmospheric noise and thermal noise.

Sky-wave ionospheric propagation ceases to exist at frequencies above approximately 30 MHz, which is the end of the HF band. However, it is possible to have ionospheric scatter propagation at frequencies in the range 30–60 MHz, resulting from signal scattering from the lower ionosphere. It is also possible to communicate over distances of several hundred miles by use of tropospheric scattering at frequencies in the range 40–300 MHz. Troposcatter results from signal scattering due to particles in the atmosphere at altitudes of 10 miles or less. Generally, ionospheric scatter and tropospheric scatter involve large signal propagation losses and require a large amount of transmitter power and relatively large antennas.

Frequencies above 30 MHz propagate through the ionosphere with relatively little loss and make satellite and extraterrestrial communications possible. Hence, at frequencies in the very high frequency (VHF) band and higher, the dominant mode of electromagnetic propagation is LOS propagation. For terrestrial communication systems, this means that the transmitter and receiver antennas must be in direct LOS with relatively little or no obstruction. For this reason, television stations transmitting in the VHF and ultra high frequency (UHF) bands mount their antennas on high towers to achieve a broad coverage area.

In general, the coverage area for LOS propagation is limited by the curvature of the earth. If the transmitting antenna is mounted at a height h m above the surface of the earth, the distance to the radio horizon, assuming no physical obstructions such as mountains, is approximately $d = \sqrt{15h}$ km. For example, a television antenna mounted on a tower of 300 m in height provides a coverage of approximately 67 km. As another example, microwave radio relay systems used extensively for telephone and video transmission at frequencies above 1 gigahertz (GHz) have antennas mounted on tall towers or on the top of tall buildings.

The dominant noise limiting the performance of a communication system in VHF and UHF ranges is thermal noise generated in the receiver front end and cosmic noise picked up by the antenna. At frequencies in the super high frequency (SHF) band above 10 GHz, atmospheric conditions play a major role in signal propagation. For example, at 10 GHz, the attenuation ranges from about 0.003 decibel per kilometer (dB/km) in light rain to about 0.3 dB/km in heavy rain. At 100 GHz, the attenuation ranges from about 0.1 dB/km in light rain to about 6 dB/km in heavy rain. Hence, in this frequency range, heavy rain introduces extremely high propagation losses that can result in service outages (total breakdown in the communication system).

At frequencies above the extremely high frequency (EHF) band, we have the infrared and visible light regions of the electromagnetic spectrum, which can be used to provide LOS optical communication in free space. To date, these frequency bands

have been used in experimental communication systems, such as satellite-to-satellite links.

Underwater Acoustic Channels

Over the past few decades, ocean exploration activity has been steadily increasing. Coupled with this increase is the need to transmit data, collected by sensors placed under water, to the surface of the ocean. From there, it is possible to relay the data via a satellite to a data collection center.

Electromagnetic waves do not propagate over long distances under water except at extremely low frequencies. However, the transmission of signals at such low frequencies is prohibitively expensive because of the large and powerful transmitters required. The attenuation of electromagnetic waves in water can be expressed in terms of the *skin depth*, which is the distance a signal is attenuated by $1/e$. For seawater, the skin depth $\delta = 250/\sqrt{f}$, where f is expressed in Hz and δ is in m. For example, at 10 kHz, the skin depth is 2.5 m. In contrast, acoustic signals propagate over distances of tens and even hundreds of kilometers.

An underwater acoustic channel is characterized as a multipath channel due to signal reflections from the surface and the bottom of the sea. Because of wave motion, the signal multipath components undergo time-varying propagation delays that result in signal fading. In addition, there is frequency-dependent attenuation, which is approximately proportional to the square of the signal frequency. The sound velocity is nominally about 1500 m/s, but the actual value will vary either above or below the nominal value depending on the depth at which the signal propagates.

Ambient ocean acoustic noise is caused by shrimp, fish, and various mammals. Near harbors, there is also man-made acoustic noise in addition to the ambient noise. In spite of this hostile environment, it is possible to design and implement efficient and highly reliable underwater acoustic communication systems for transmitting digital signals over large distances.

Storage Channels

Information storage and retrieval systems constitute a very significant part of data-handling activities on a daily basis. Magnetic tape, including digital audiotape and videotape, magnetic disks used for storing large amounts of computer data, optical disks used for computer data storage, and compact disks are examples of data storage systems that can be characterized as communication channels. The process of storing data on a magnetic tape or a magnetic or optical disk is equivalent to transmitting a signal over a telephone or a radio channel. The readback process and the signal processing involved in storage systems to recover the stored information are equivalent to the functions performed by a receiver in a telephone or radio communication system to recover the transmitted information.

Additive noise generated by the electronic components and interference from adjacent tracks is generally present in the readback signal of a storage system, just as is the case in a telephone or a radio communication system.

The amount of data that can be stored is generally limited by the size of the disk or tape and the density (number of bits stored per square inch) that can be achieved by

the write/read electronic systems and heads. For example, a packing density of 10^9 bits per square inch has been demonstrated in magnetic disk storage systems. The speed at which data can be written on a disk or tape and the speed at which it can be read back are also limited by the associated mechanical and electrical subsystems that constitute an information storage system.

Channel coding and modulation are essential components of a well-designed digital magnetic or optical storage system. In the readback process, the signal is demodulated and the added redundancy introduced by the channel encoder is used to correct errors in the readback signal.

■ 1.3
MATHEMATICAL MODELS FOR COMMUNICATION CHANNELS

In the design of communication systems for transmitting information through physical channels, we find it convenient to construct mathematical models that reflect the most important characteristics of the transmission medium. Then, the mathematical model for the channel is used in the design of the channel encoder and modulator at the transmitter and the demodulator and channel decoder at the receiver. Below, we provide a brief description of the channel models that are frequently used to characterize many of the physical channels that we encounter in practice.
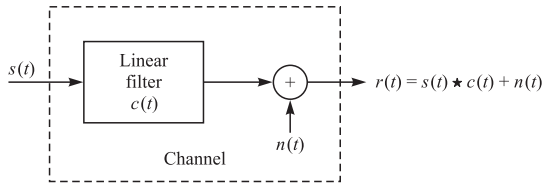
The Additive Noise Channel

The simplest mathematical model for a communication channel is the additive noise channel, illustrated in Figure 1.3–1. In this model, the transmitted signal $s(t)$ is corrupted by an additive random noise process $n(t)$. Physically, the additive noise process may arise from electronic components and amplifiers at the receiver of the communication system or from interference encountered in transmission (as in the case of radio signal transmission).

If the noise is introduced primarily by electronic components and amplifiers at the receiver, it may be characterized as thermal noise. This type of noise is characterized statistically as a *Gaussian noise process*. Hence, the resulting mathematical model for the channel is usually called the *additive Gaussian noise channel*. Because this channel model applies to a broad class of physical communication channels and because of its mathematical tractability, this is the predominant channel model used in our communication system analysis and design. Channel attenuation is easily incorporated into the model. When the signal undergoes attenuation in transmission through the
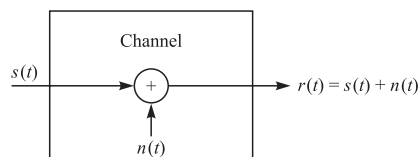
FIGURE 1.3–1
The additive noise channel.

FIGURE 1.3–2
The linear filter channel with additive noise.

channel, the received signal is

$$r(t) = \alpha s(t) + n(t) \tag{1.3–1}$$

where α is the attenuation factor.

The Linear Filter Channel

In some physical channels, such as wireline telephone channels, filters are used to en-
sure that the transmitted signals do not exceed specified bandwidth limitations and thus
do not interfere with one another. Such channels are generally characterized mathemat-
ically as linear filter channels with additive noise, as illustrated in Figure 1.3–2. Hence,
if the channel input is the signal $s(t)$, the channel output is the signal

$$
\begin{aligned}
r(t) &= s(t) \star c(t) + n(t) \\
&= \int_{-\infty}^{\infty} c(\tau) s(t - \tau)\, d\tau + n(t)
\end{aligned}
\tag{1.3–2}
$$

where $c(t)$ is the impulse response of the linear filter and $\star$ denotes convolution.

The Linear Time-Variant Filter Channel

Physical channels such as underwater acoustic channels and ionospheric radio chan-
nels that result in time-variant multipath propagation of the transmitted signal may be
characterized mathematically as time-variant linear filters. Such linear filters are charac-
terized by a time-variant channel impulse response $c(\tau; t)$, where $c(\tau; t)$ is the response
of the channel at time t due to an impulse applied at time $t - \tau$. Thus, τ represents the
"age" (elapsed-time) variable. The linear time-variant filter channel with additive noise
is illustrated in Figure 1.3–3. For an input signal $s(t)$, the channel output signal is

$$
\begin{aligned}
r(t) &= s(t) \star c(\tau; t) + n(t) \\
&= \int_{-\infty}^{\infty} c(\tau; t) s(t - \tau)\, d\tau + n(t)
\end{aligned}
\tag{1.3–3}
$$
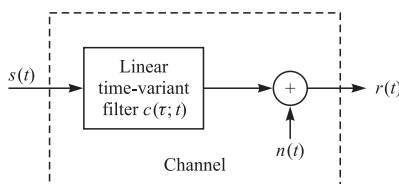
FIGURE 1.3–3
Linear time-variant filter channel with additive noise.

A good model for multipath signal propagation through physical channels, such as the ionosphere (at frequencies below 30 MHz) and mobile cellular radio channels, is a special case of (1.3–3) in which the time-variant impulse response has the form

$$c(\tau; t) = \sum_{k=1}^{L} a_k(t)\delta(\tau - \tau_k) \qquad (1.3\text{–}4)$$

where the $\{a_k(t)\}$ represents the possibly time-variant attenuation factors for the L multipath propagation paths and $\{\tau_k\}$ are the corresponding time delays. If (1.3–4) is substituted into (1.3–3), the received signal has the form

$$r(t) = \sum_{k=1}^{L} a_k(t)s(t - \tau_k) + n(t) \qquad (1.3\text{–}5)$$

Hence, the received signal consists of L multipath components, where the kth component is attenuated by $a_k(t)$ and delayed by τ_k.

The three mathematical models described above adequately characterize the great majority of the physical channels encountered in practice. These three channel models are used in this text for the analysis and design of communication systems.

■ 1.4
A HISTORICAL PERSPECTIVE IN THE DEVELOPMENT OF DIGITAL COMMUNICATIONS

It is remarkable that the earliest form of electrical communication, namely *telegraphy*, was a digital communication system. The electric telegraph was developed by Samuel Morse and was demonstrated in 1837. Morse devised the variable-length binary code in which letters of the English alphabet are represented by a sequence of dots and dashes (code words). In this code, more frequently occurring letters are represented by short code words, while letters occurring less frequently are represented by longer code words. Thus, the *Morse code* was the precursor of the variable-length source coding methods described in Chapter 6.

Nearly 40 years later, in 1875, Emile Baudot devised a code for telegraphy in which every letter was encoded into fixed-length binary code words of length 5. In the *Baudot code*, binary code elements are of equal length and designated as mark and space.

Although Morse is responsible for the development of the first electrical digital communication system (telegraphy), the beginnings of what we now regard as modern digital communications stem from the work of Nyquist (1924), who investigated the problem of determining the maximum signaling rate that can be used over a telegraph channel of a given bandwidth without intersymbol interference. He formulated a model of a telegraph system in which a transmitted signal has the general form

$$s(t) = \sum_n a_n g(t - nT) \qquad (1.4\text{–}1)$$

where $g(t)$ represents a basic pulse shape and $\{a_n\}$ is the binary data sequence of $\{\pm 1\}$ transmitted at a rate of $1/T$ bits/s. Nyquist set out to determine the optimum pulse shape that was band-limited to W Hz and maximized the bit rate under the constraint that the pulse caused no intersymbol interference at the sampling time k/T, $k = 0, \pm 1, \pm 2, \ldots$. His studies led him to conclude that the maximum pulse rate is $2W$ pulses/s. This rate is now called the *Nyquist rate*. Moreover, this pulse rate can be achieved by using the pulses $g(t) = (\sin 2\pi Wt)/2\pi Wt$. This pulse shape allows recovery of the data without intersymbol interference at the sampling instants. Nyquist's result is equivalent to a version of the sampling theorem for band-limited signals, which was later stated precisely by Shannon (1948b). The sampling theorem states that a signal of bandwidth W can be reconstructed from samples taken at the Nyquist rate of $2W$ samples/s using the interpolation formula

$$s(t) = \sum_n s\left(\frac{n}{2W}\right) \frac{\sin[2\pi W(t - n/2W)]}{2\pi W(t - n/2W)} \tag{1.4–2}$$

In light of Nyquist's work, Hartley (1928) considered the issue of the amount of data that can be transmitted reliably over a band-limited channel when multiple amplitude levels are used. Because of the presence of noise and other interference, Hartley postulated that the receiver can reliably estimate the received signal amplitude to some accuracy, say A_δ. This investigation led Hartley to conclude that there is a maximum data rate that can be communicated reliably over a band-limited channel when the maximum signal amplitude is limited to $A_{\max}$ (fixed power constraint) and the amplitude resolution is A_δ.

Another significant advance in the development of communications was the work of Kolmogorov (1939) and Wiener (1942), who considered the problem of estimating a desired signal waveform $s(t)$ in the presence of additive noise $n(t)$, based on observation of the received signal $r(t) = s(t) + n(t)$. This problem arises in signal demodulation. Kolmogorov and Wiener determined the linear filter whose output is the best mean-square approximation to the desired signal $s(t)$. The resulting filter is called the *optimum linear (Kolmogorov–Wiener) filter*.

Hartley's and Nyquist's results on the maximum transmission rate of digital information were precursors to the work of Shannon (1948a,b), who established the mathematical foundations for information transmission and derived the fundamental limits for digital communication systems. In his pioneering work, Shannon formulated the basic problem of reliable transmission of information in statistical terms, using probabilistic models for information sources and communication channels. Based on such a statistical formulation, he adopted a logarithmic measure for the information content of a source. He also demonstrated that the effect of a transmitter power constraint, a bandwidth constraint, and additive noise can be associated with the channel and incorporated into a single parameter, called the *channel capacity*. For example, in the case of an additive white (spectrally flat) Gaussian noise interference, an ideal band-limited channel of bandwidth W has a capacity C given by

$$C = W \log_2\left(1 + \frac{P}{W N_0}\right) \qquad \text{bits/s} \tag{1.4–3}$$

where P is the average transmitted power and N_0 is the power spectral density of the additive noise. The significance of the channel capacity is as follows: If the information rate R from the source is less than $C(R < C)$, then it is theoretically possible to achieve reliable (error-free) transmission through the channel by appropriate coding. On the other hand, if $R > C$, reliable transmission is not possible regardless of the amount of signal processing performed at the transmitter and receiver. Thus, Shannon established basic limits on communication of information and gave birth to a new field that is now called *information theory*.

Another important contribution to the field of digital communication is the work of Kotelnikov (1947), who provided a coherent analysis of the various digital communication systems based on a geometrical approach. Kotelnikov's approach was later expanded by Wozencraft and Jacobs (1965).

Following Shannon's publications came the classic work of Hamming (1950) on error-detecting and error-correcting codes to combat the detrimental effects of channel noise. Hamming's work stimulated many researchers in the years that followed, and a variety of new and powerful codes were discovered, many of which are used today in the implementation of modern communication systems.

The increase in demand for data transmission during the last four decades, coupled with the development of more sophisticated integrated circuits, has led to the development of very efficient and more reliable digital communication systems. In the course of these developments, Shannon's original results and the generalization of his results on maximum transmission limits over a channel and on bounds on the performance achieved have served as benchmarks for any given communication system design. The theoretical limits derived by Shannon and other researchers that contributed to the development of information theory serve as an ultimate goal in the continuing efforts to design and develop more efficient digital communication systems.

There have been many new advances in the area of digital communications following the early work of Shannon, Kotelnikov, and Hamming. Some of the most notable advances are the following:

- The development of new block codes by Muller (1954), Reed (1954), Reed and Solomon (1960), Bose and Ray-Chaudhuri (1960a,b), and Goppa (1970, 1971).
- The development of concatenated codes by Forney (1966a).
- The development of computationally efficient decoding of Bose–Chaudhuri-Hocquenghem (BCH) codes, e.g., the Berlekamp–Massey algorithm (see Chien, 1964; Berlekamp, 1968).
- The development of convolutional codes and decoding algorithms by Wozencraft and Reiffen (1961), Fano (1963), Zigangirov (1966), Jelinek (1969), Forney (1970b, 1972, 1974), and Viterbi (1967, 1971).
- The development of trellis-coded modulation by Ungerboeck (1982), Forney et al. (1984), Wei (1987), and others.
- The development of efficient source encodings algorithms for data compression, such as those devised by Ziv and Lempel (1977, 1978), and Linde et al. (1980).
- The development of low-density parity check (LDPC) codes and the sum-product decoding algorithm by Gallager (1963).
- The development of turbo codes and iterative decoding by Berrou et al. (1993).

■ 1.5
OVERVIEW OF THE BOOK

Chapter 2 presents a review of deterministic and random signal analysis. Our primary objectives in this chapter are to review basic notions in the theory of probability and random variables and to establish some necessary notation.

Chapters 3 through 5 treat the geometric representation of various digital modulation signals, their demodulation, their error rate performance in additive, white Gaussian noise (AWGN) channels, and methods for synchronizing the receiver to the received signal waveforms.

Chapters 6 to 8 treat the topics of source coding, channel coding and decoding, and basic information theoretic limits on channel capacity, source information rates, and channel coding rates.

The design of efficient modulators and demodulators for linear filter channels with distortion is treated in Chapters 9 and 10. Channel equalization methods are described for mitigating the effects of channel distortion.

Chapter 11 is focused on multichannel and multicarrier communication systems, their efficient implementation, and their performance in AWGN channels.

Chapter 12 presents an introduction to direct sequence and frequency hopped spread spectrum signals and systems and an evaluation of their performance under worst-case interference conditions.

The design of signals and coding techniques for digital communication through fading multipath channels is the focus of Chapters 13 and 14. This material is especially relevant to the design and development of wireless communication systems.

Chapter 15 treats the use of multiple transmit and receive antennas for improving the performance of wireless communication systems through signal diversity and increasing the data rate via spatial multiplexing. The capacity of multiple antenna systems is evaluated and space-time codes are described for use in multiple antenna communication systems.

Chapter 16 of this book presents an introduction to multiuser communication systems and multiple access methods. We consider detection algorithms for uplink transmission in which multiple users transmit data to a common receiver (a base station) and evaluate their performance. We also present algorithms for suppressing multiple access interference in a broadcast communication system in which a transmitter employing multiple antennas transmits different data sequences simultaneously to different users.

■ 1.6
BIBLIOGRAPHICAL NOTES AND REFERENCES

There are several historical treatments regarding the development of radio and telecommunications during the past century. These may be found in the books by McMahon (1984), Millman (1984), and Ryder and Fink (1984). We have already cited the classical works of Nyquist (1924), Hartley (1928), Kotelnikov (1947), Shannon (1948), and

Hamming (1950), as well as some of the more important advances that have occurred in the field since 1950. The collected papers by Shannon have been published by IEEE Press in a book edited by Sloane and Wyner (1993) and previously in Russia in a book edited by Dobrushin and Lupanov (1963). Other collected works published by the IEEE Press that might be of interest to the reader are *Key Papers in the Development of Coding Theory*, edited by Berlekamp (1974), and *Key Papers in the Development of Information Theory*, edited by Slepian (1974).