Chapter 1

Introduction

1-1 DATA COMMUNICATIONS

The term telecommunication means communication at a distance. The word data refers to information presented in whatever form is agreed upon by the parties creating and using the data. Data communications are the exchange of data between two devices via some form of transmission medium such as a wire cable.

Topics discussed in this section:

- **Components of a data communications system**
- **Data Flow**

Figure 1.1 *Components of a data communication system*

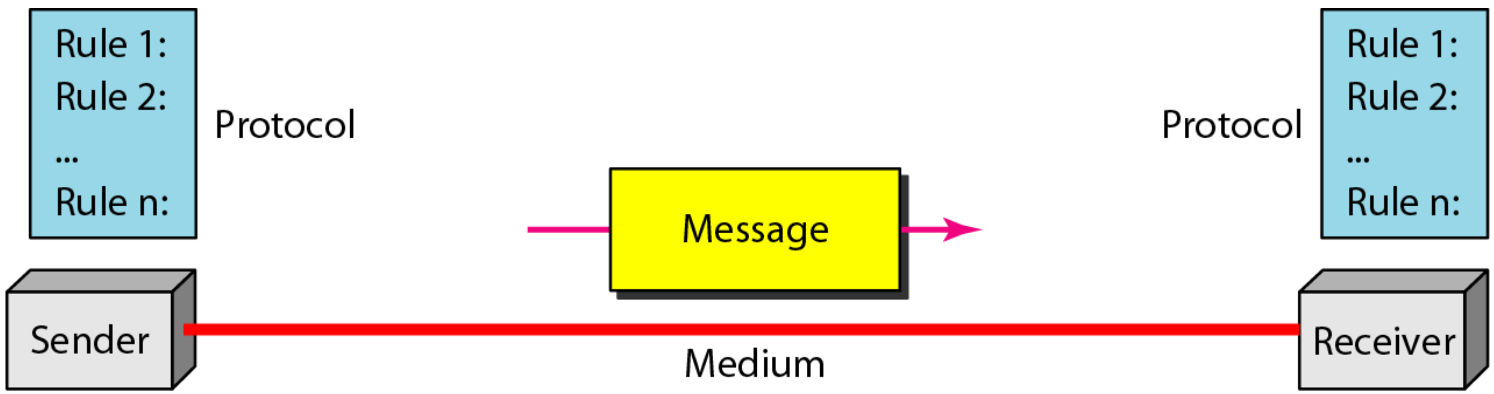
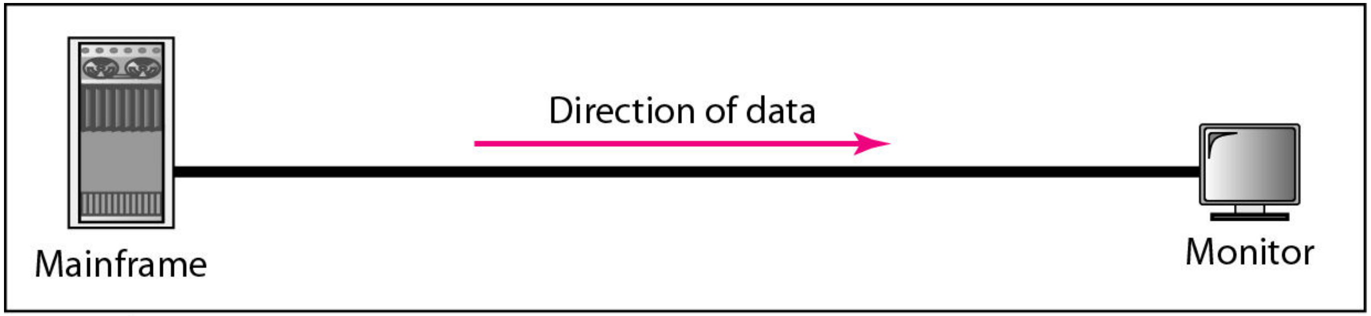
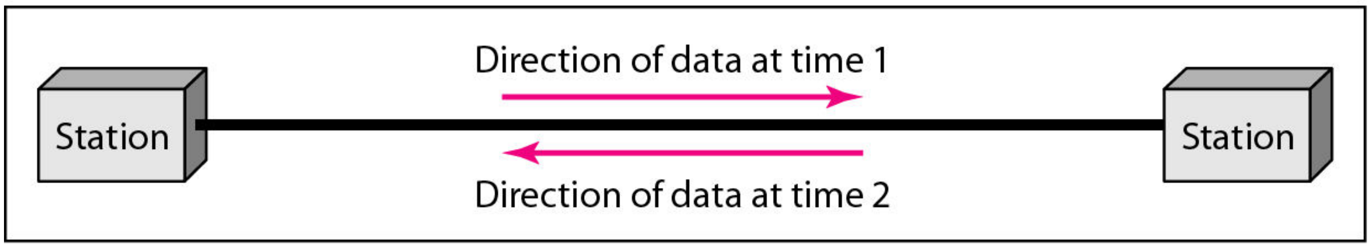


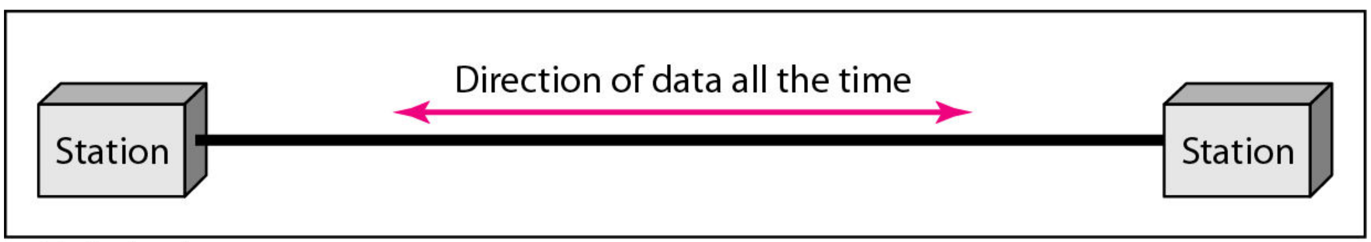
Figure 1.2 *Data flow (simplex, half-duplex, and full-duplex)*



a. Simplex



b. Half-duplex



c. Full-duplex

1-2 NETWORKS

A network is a set of devices (often referred to as nodes) connected by communication links. A node can be a computer, printer, or any other device capable of sending and/or receiving data generated by other nodes on the network. A link can be a cable, air, optical fiber, or any medium which can transport a signal carrying information.

Topics discussed in this section:

- **Network Criteria**
- **Physical Structures**
- **Categories of Networks**

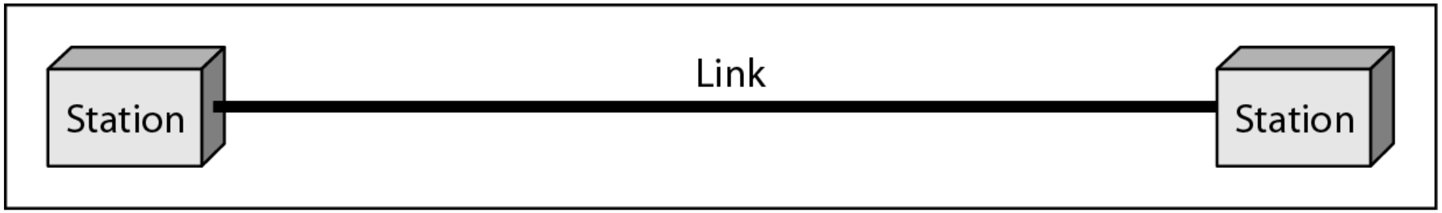
Network Criteria

- Performance
  - Depends on Network Elements
  - Measured in terms of Delay and Throughput
- Reliability
  - Failure rate of network components
  - Measured in terms of availability/robustness
- Security
  - Data protection against corruption/loss of data due to:
    - Errors
    - Malicious users

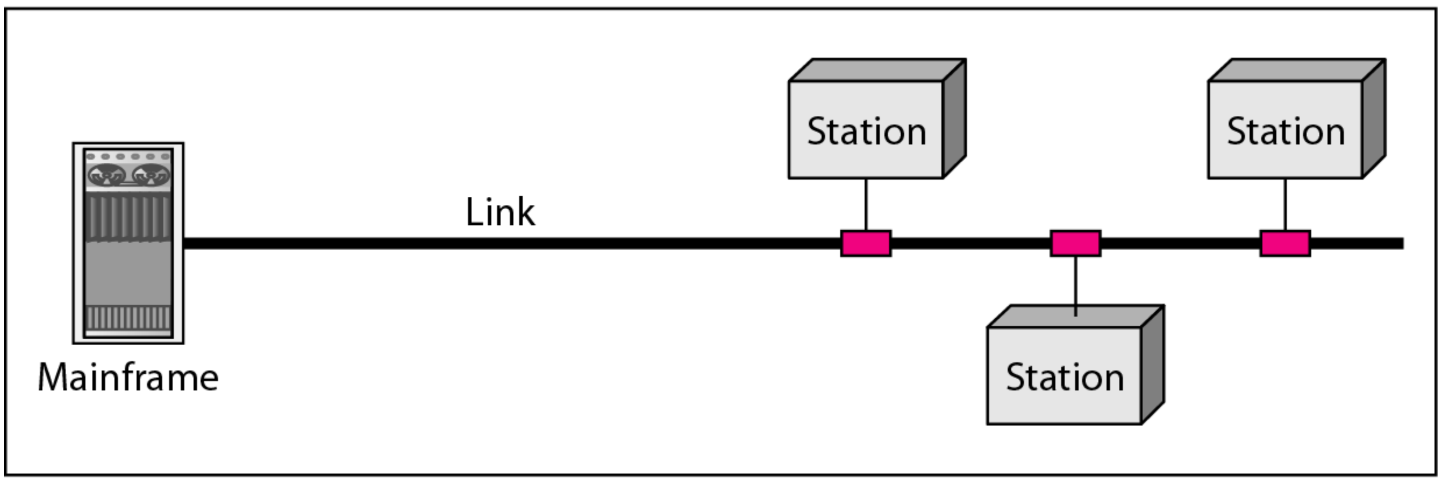
Physical Structures

- Type of Connection
  - Point to Point - single transmitter and receiver
  - Multipoint - multiple recipients of single transmission
- Physical Topology
  - Connection of devices
  - Type of transmission - unicast, mulitcast, broadcast

Figure 1.3 *Types of connections: point-to-point and multipoint*



a. Point-to-point



b. Multipoint

Figure 1.4 *Categories of topology*

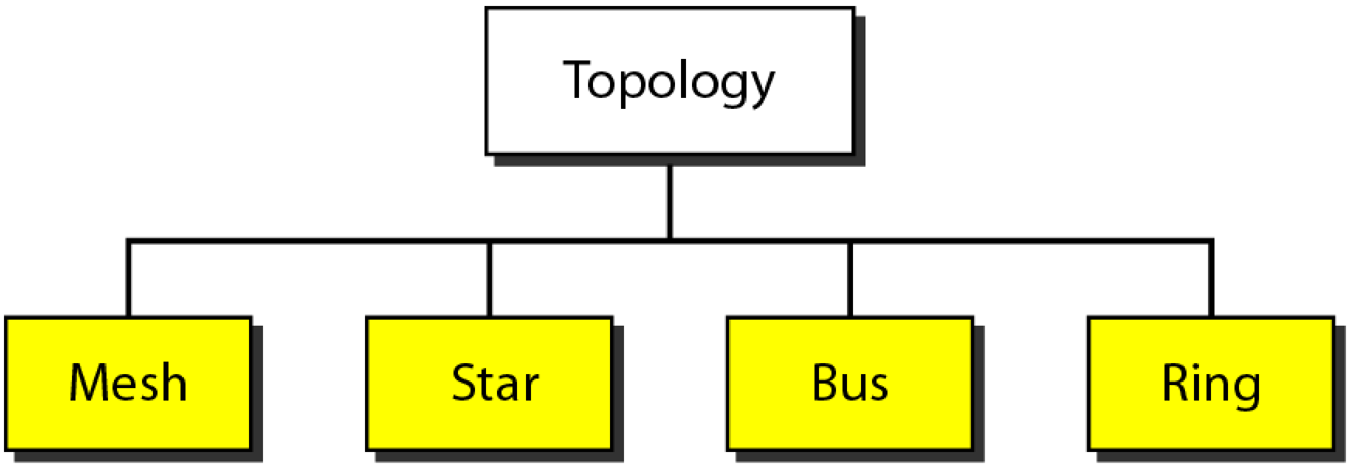


Figure 1.5 *A fully connected mesh topology (five devices)*

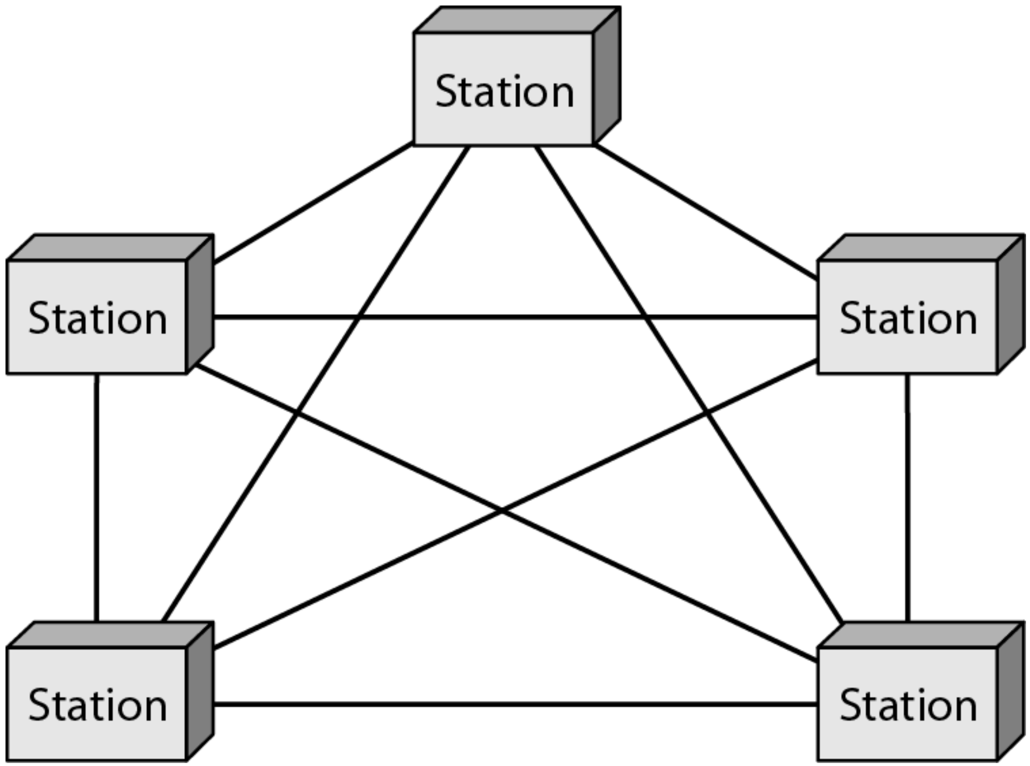


Figure 1.6 *A star topology connecting four stations*

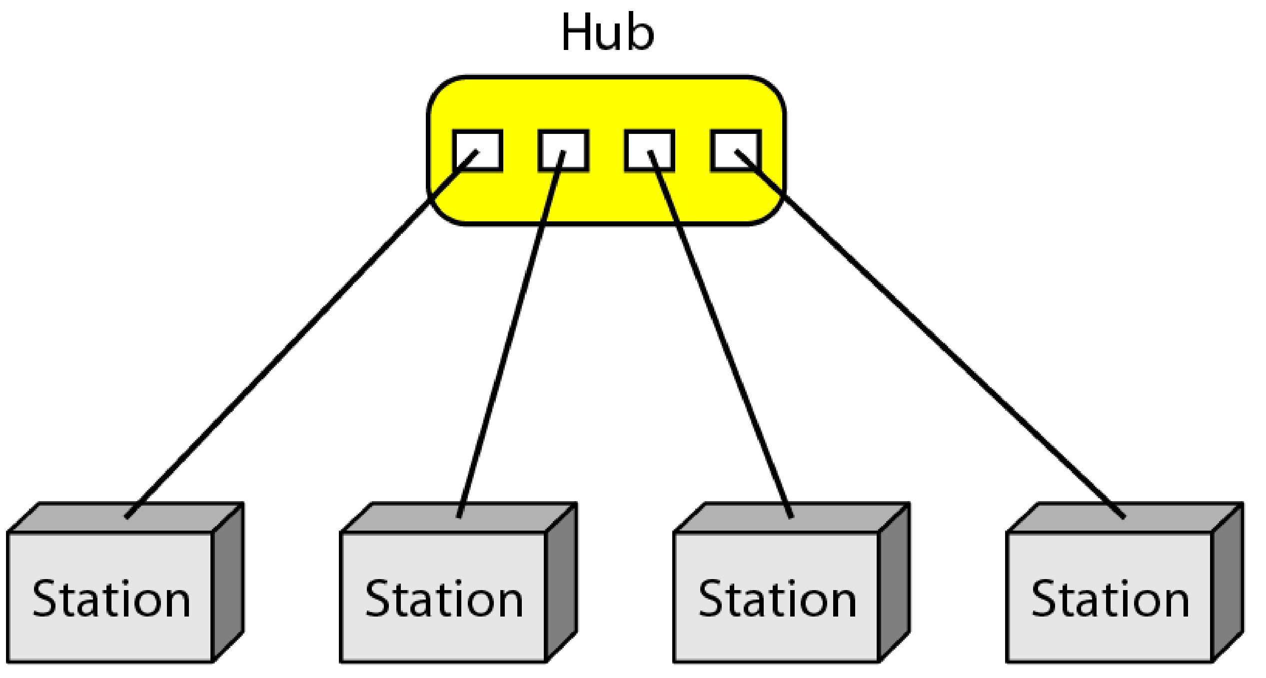


Figure 1.7 *A bus topology connecting three stations*

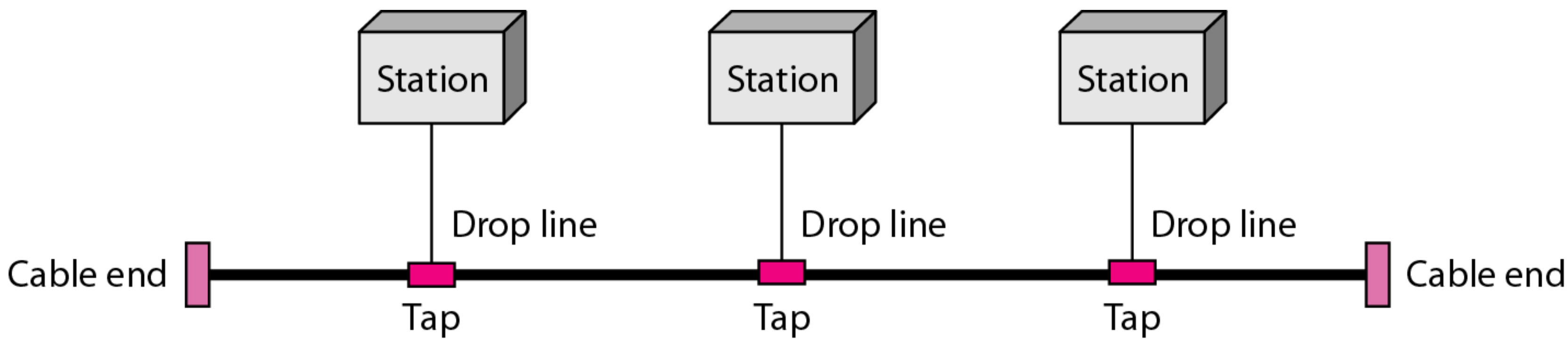


Figure 1.8 *A ring topology connecting six stations*

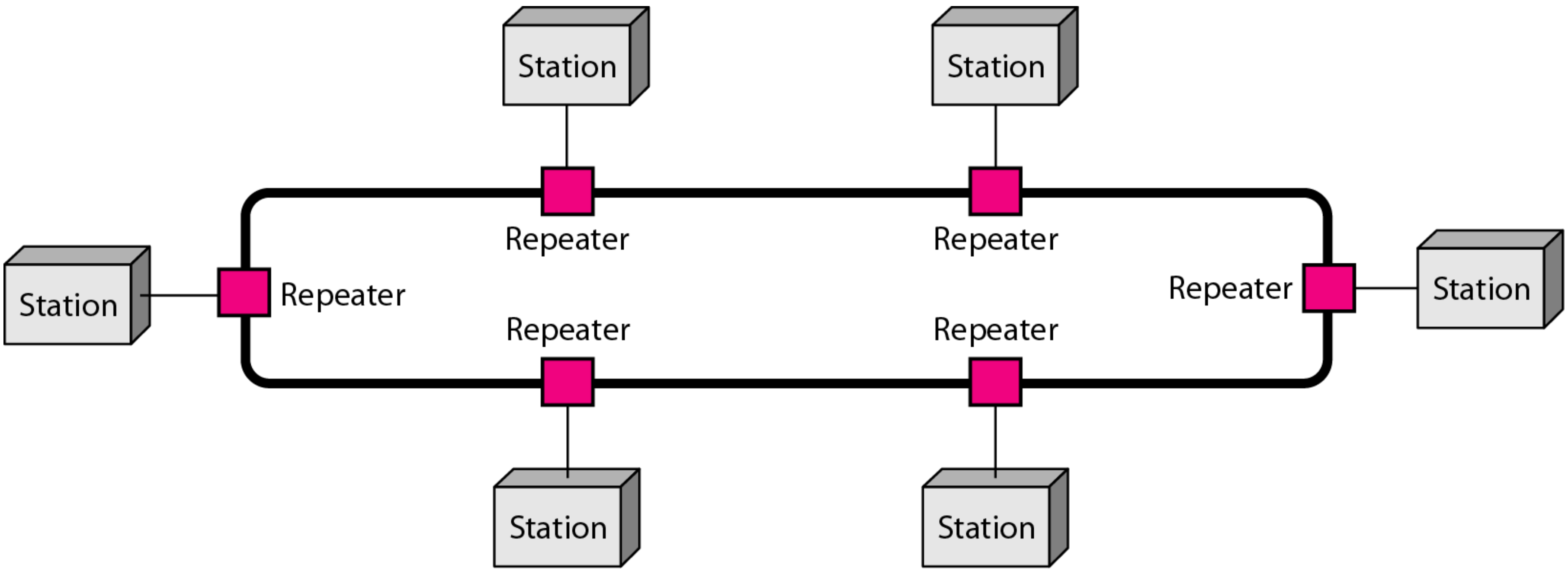
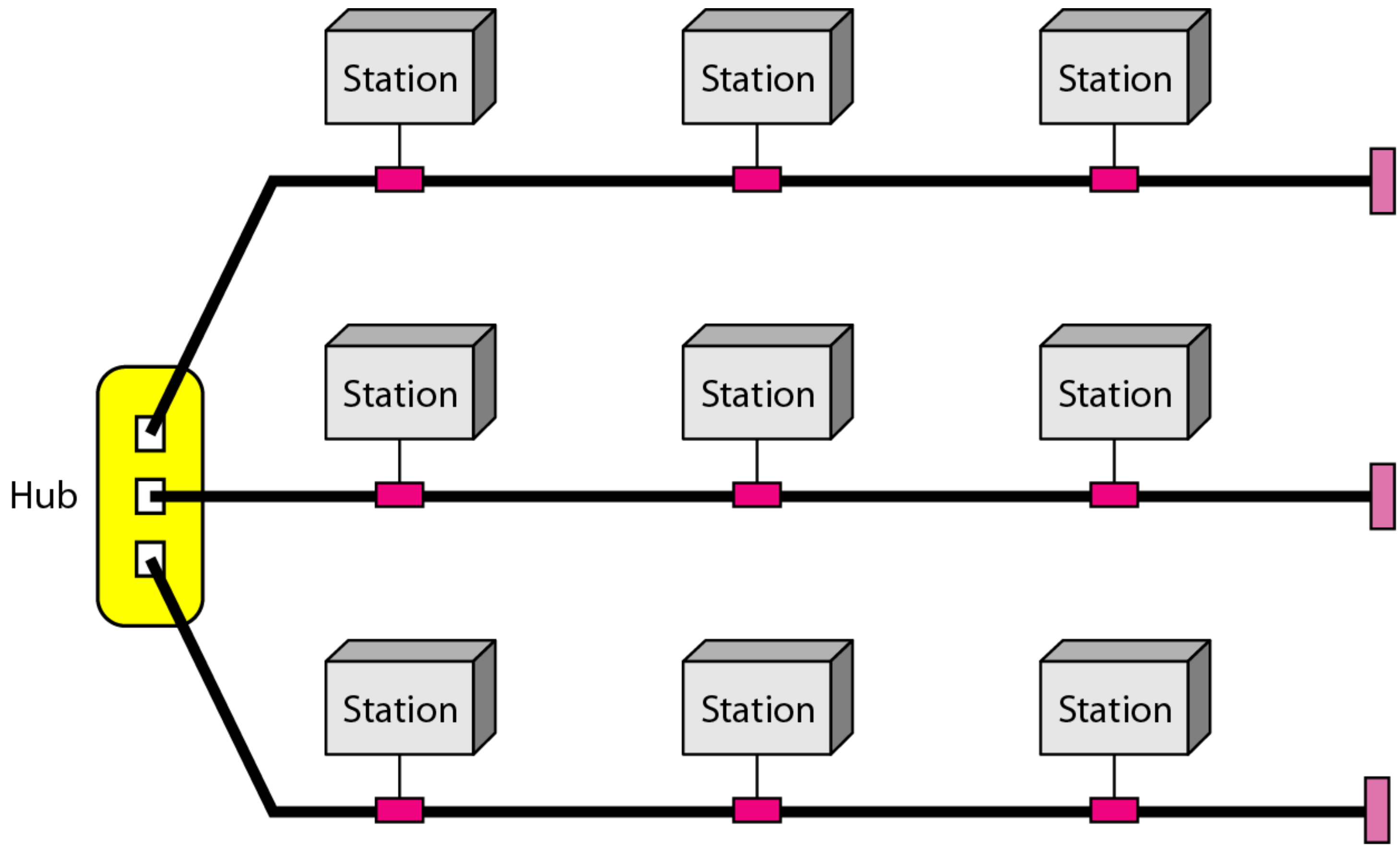


Figure 1.9 *A hybrid topology: a star backbone with three bus networks*



Categories of Networks

- Local Area Networks (LANs)
  - Short distances
  - Designed to provide local interconnectivity
- Wide Area Networks (WANs)
  - Long distances
  - Provide connectivity over large areas
- Metropolitan Area Networks (MANs)
  - Provide connectivity over areas such as a city, a campus

Figure 1.10 *An isolated LAN connecting 12 computers to a hub in a closet*

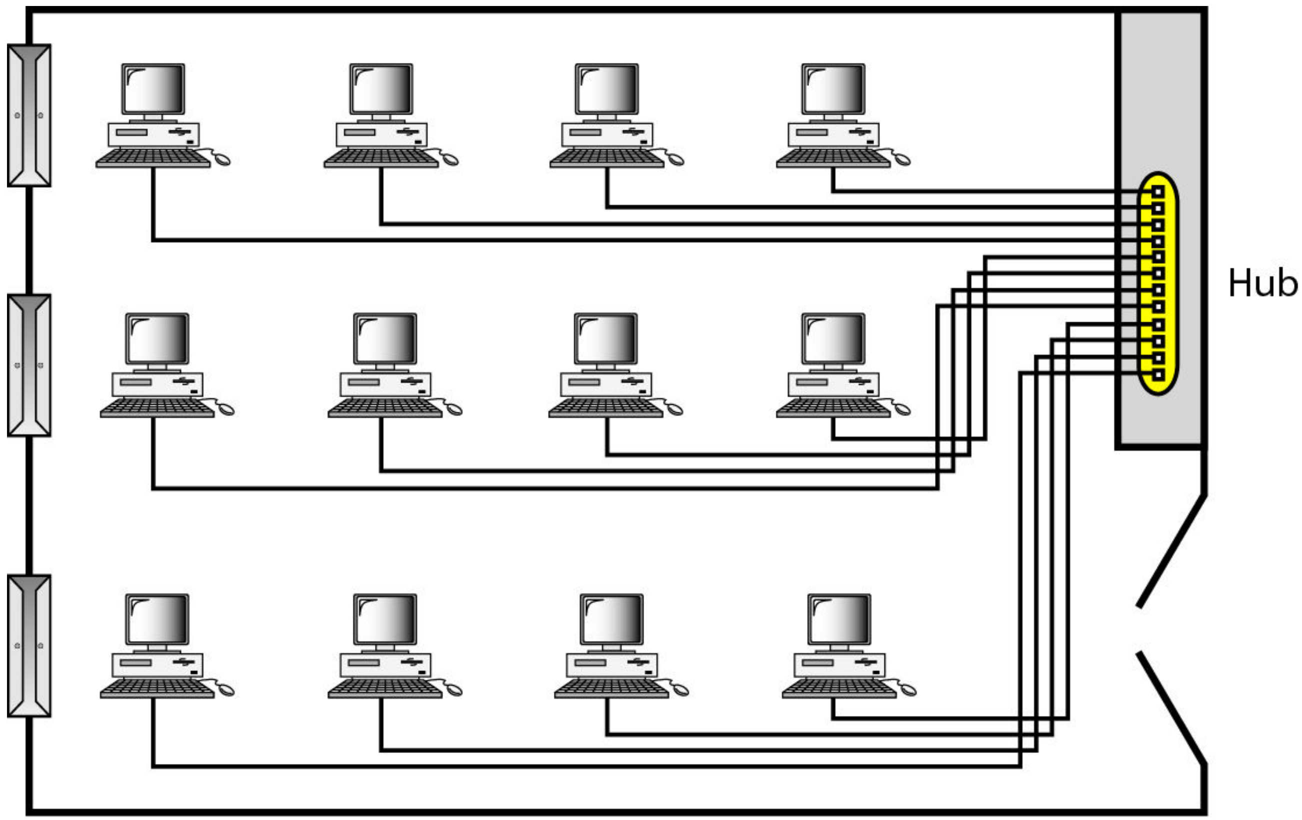
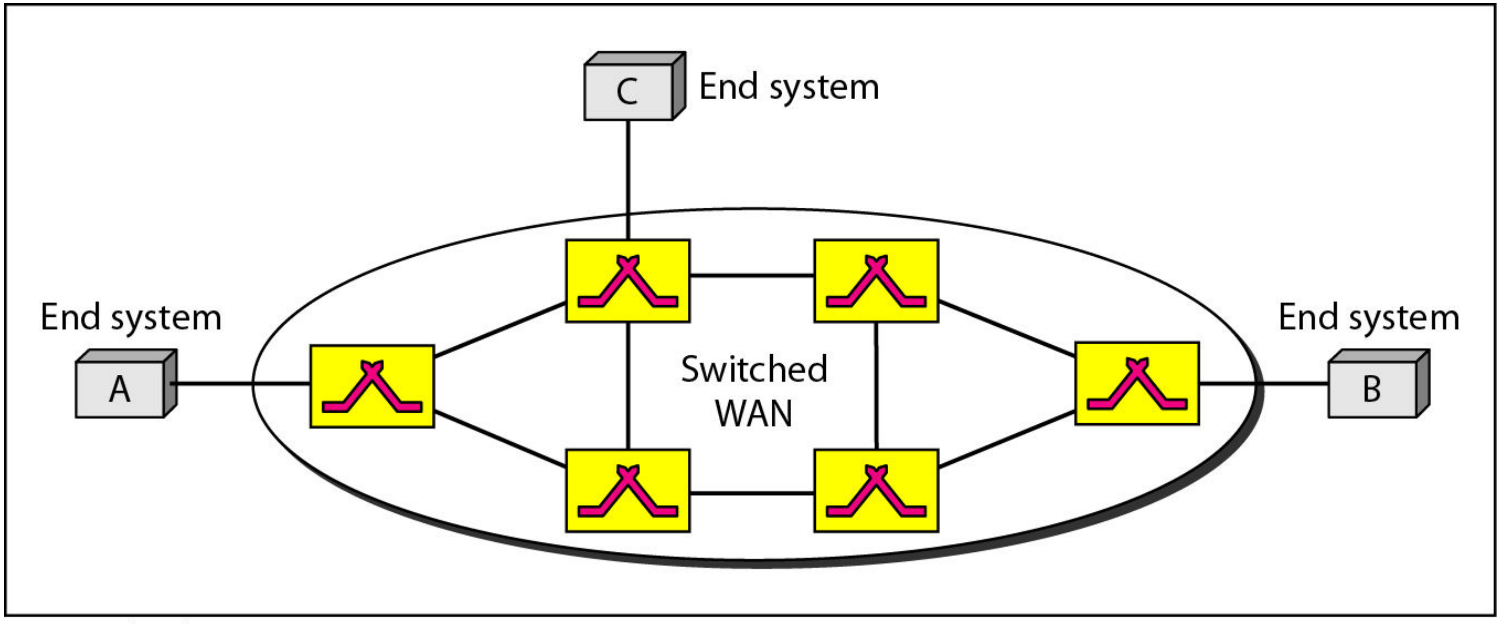
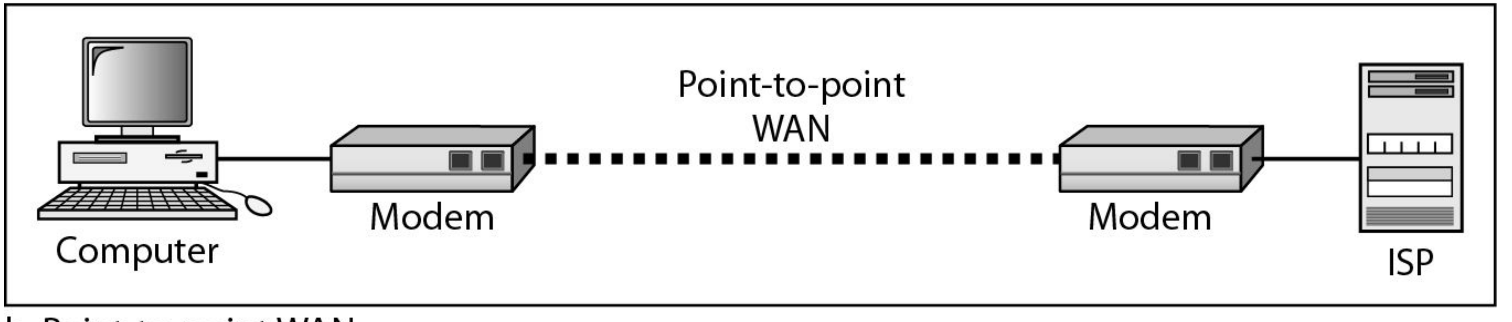


Figure 1.11 *WANs: a switched WAN and a point-to-point WAN*

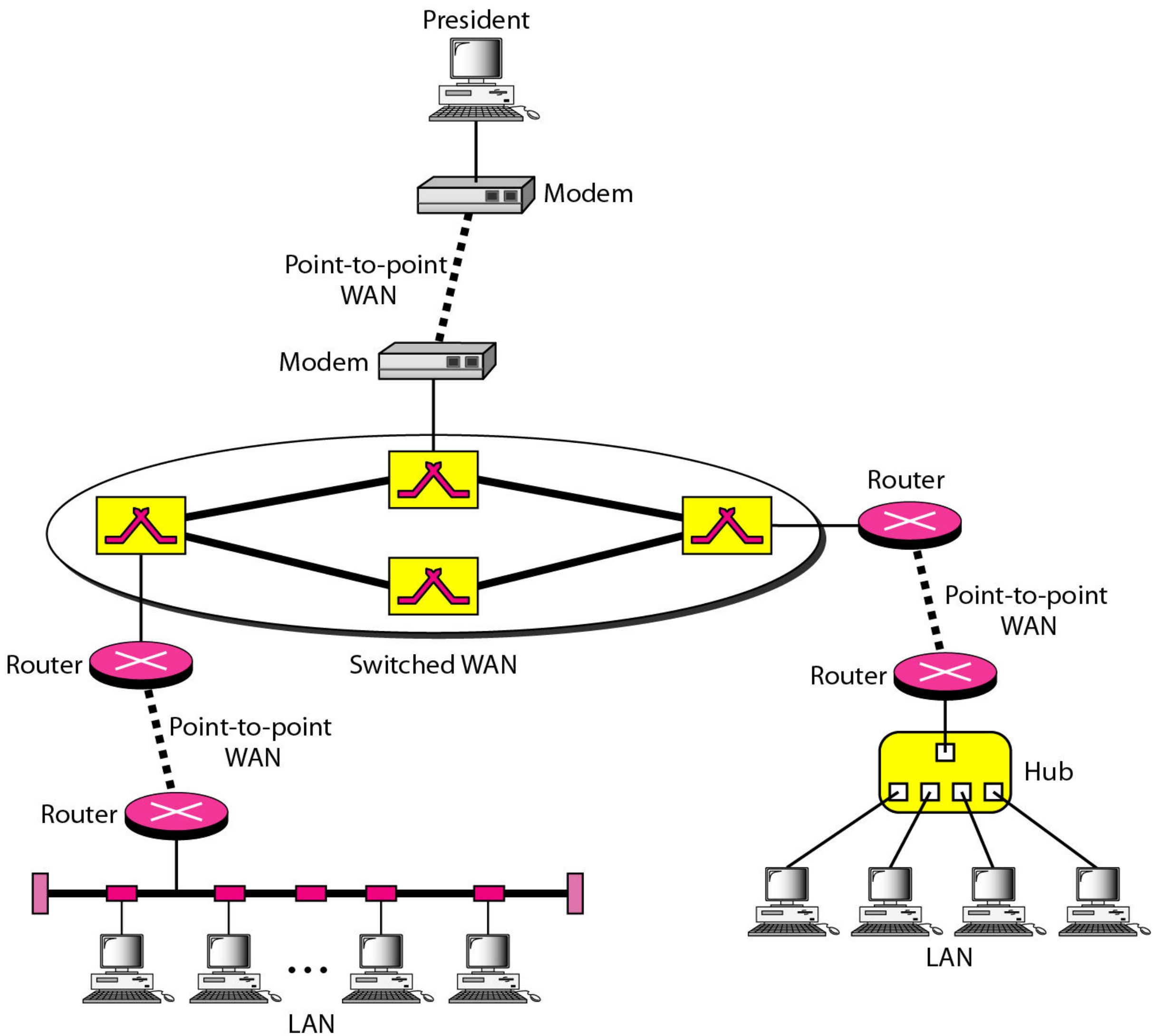


a. Switched WAN



b. Point-to-point WAN

Figure 1.12 *A heterogeneous network made of four WANs and two LANs*



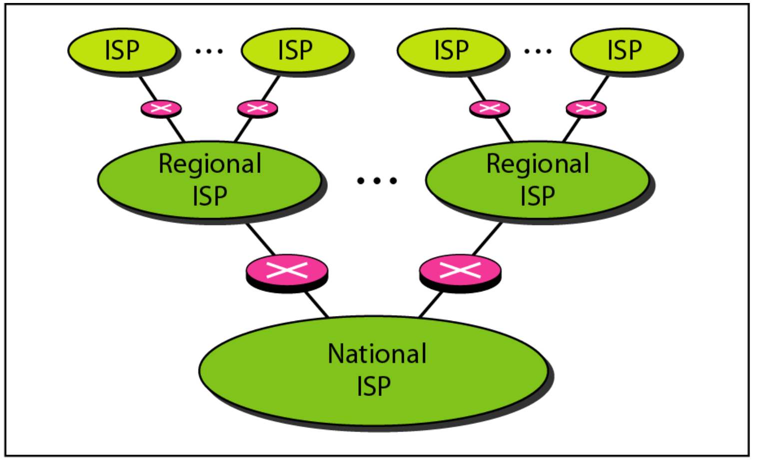
1-3 THE INTERNET

*The **Internet** has revolutionized many aspects of our daily lives. It has affected the way we do business as well as the way we spend our leisure time. The Internet is a communication system that has brought a wealth of information to our fingertips and organized it for our use.*

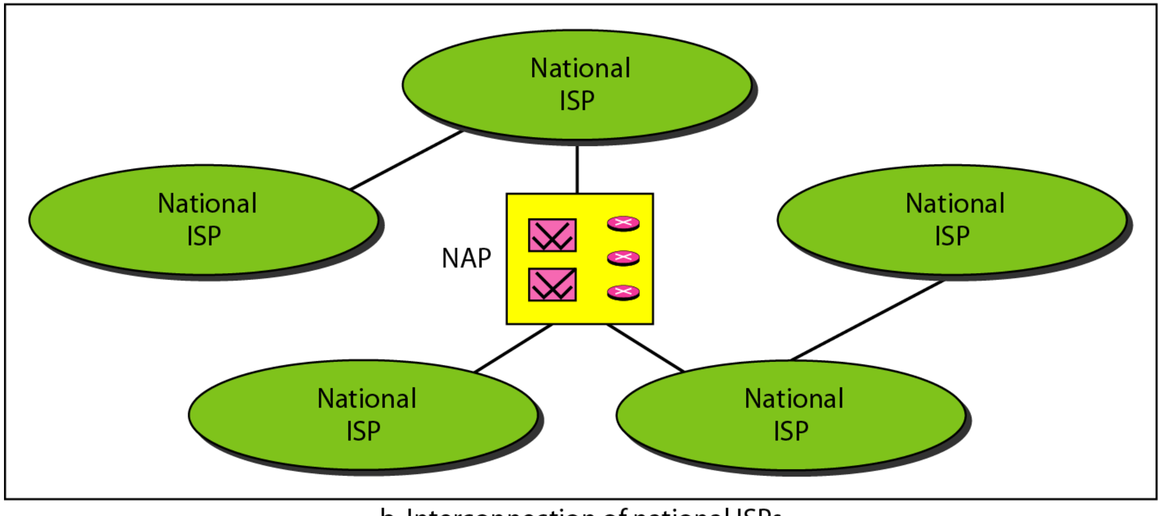
Topics discussed in this section:

Organization of the Internet
Internet Service Providers (ISPs)

Figure 1.13 *Hierarchical organization of the Internet*



a. Structure of a national ISP



b. Interconnection of national ISPs

1-4 PROTOCOLS

A protocol is synonymous with rule. It consists of a set of rules that govern data communications. It determines what is communicated, how it is communicated and when it is communicated. The key elements of a protocol are syntax, semantics and timing

Topics discussed in this section:

- **Syntax**
- **Semantics**
- **Timing**

Elements of a Protocol

- Syntax
  - Structure or format of the data
  - Indicates how to read the bits - field delineation
- Semantics
  - Interprets the meaning of the bits
  - Knows which fields define what action
- Timing
  - When data should be sent and what
  - Speed at which data should be sent or speed at which it is being received.

Chapter 2

Network Models

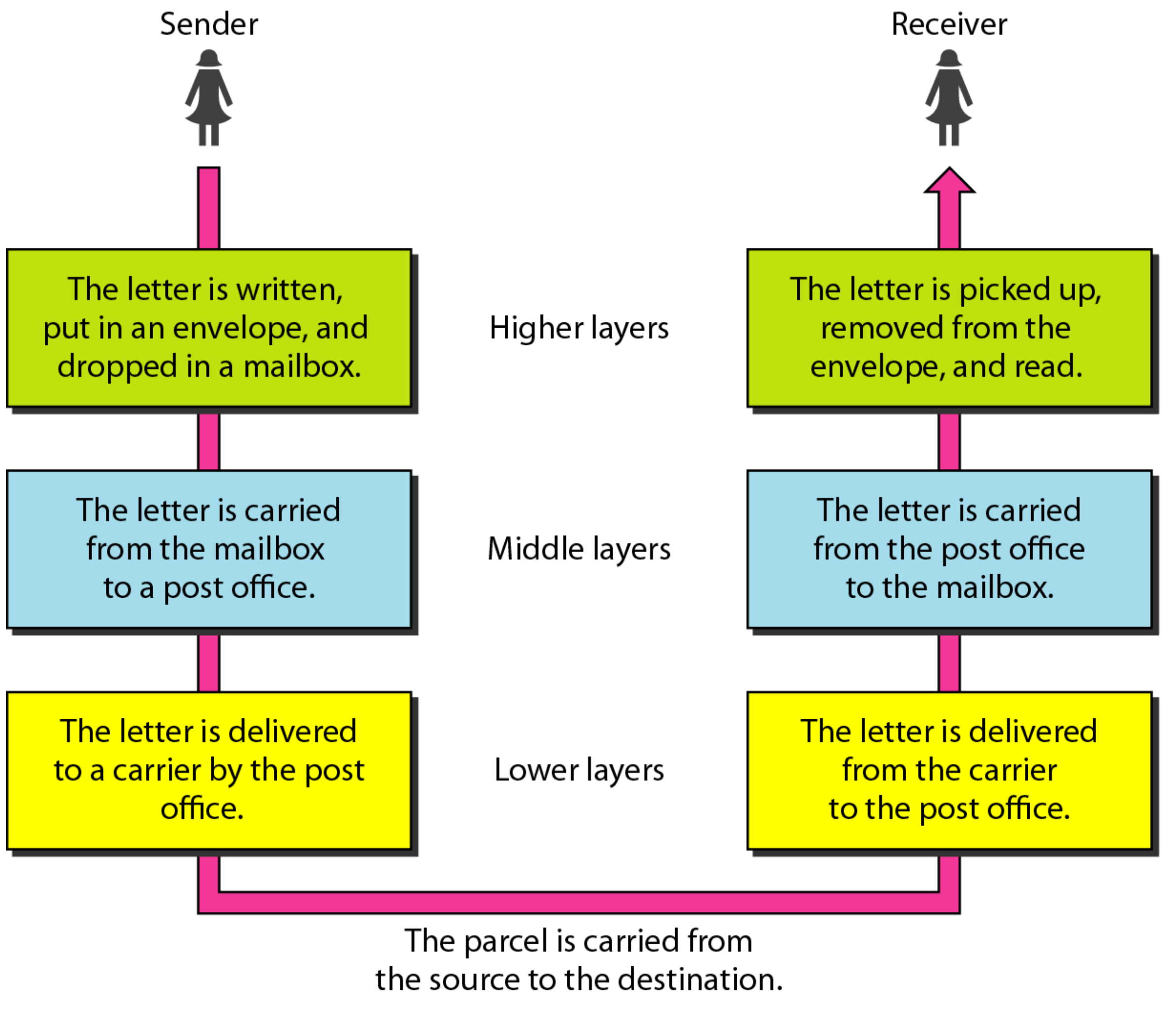
2-1 LAYERED TASKS

We use the concept of layers in our daily life. As an example, let us consider two friends who communicate through postal mail. The process of sending a letter to a friend would be complex if there were no services available from the post office.

Topics discussed in this section:

Sender, Receiver, and Carrier
Hierarchy

Figure 2.1 Tasks involved in sending a letter



2-2 THE OSI MODEL

Established in 1947, the International Standards Organization (ISO) is a multinational body dedicated to worldwide agreement on international standards. An ISO standard that covers all aspects of network communications is the Open Systems Interconnection (OSI) model. It was first introduced in the late 1970s.

Topics discussed in this section:

Layered Architecture
Peer-to-Peer Processes
Encapsulation

Note

ISO is the organization.
OSI is the model.

Figure 2.2 *Seven layers of the OSI model*

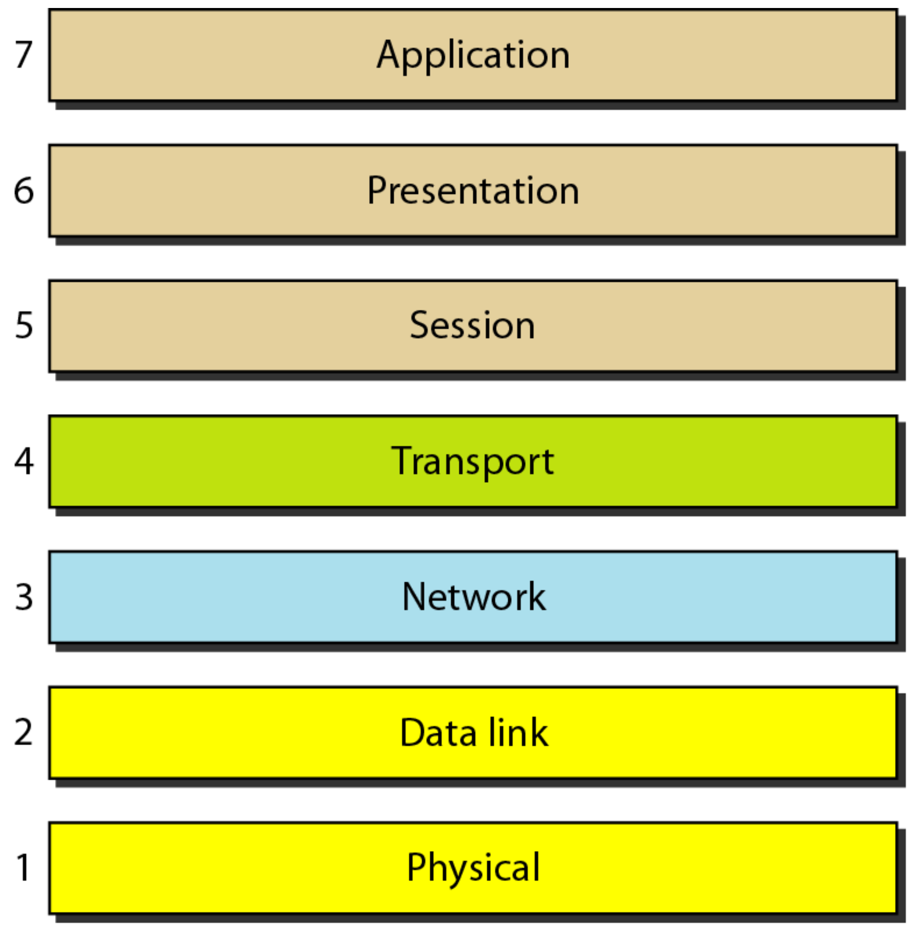


Figure 2.3 *The interaction between layers in the OSI model*

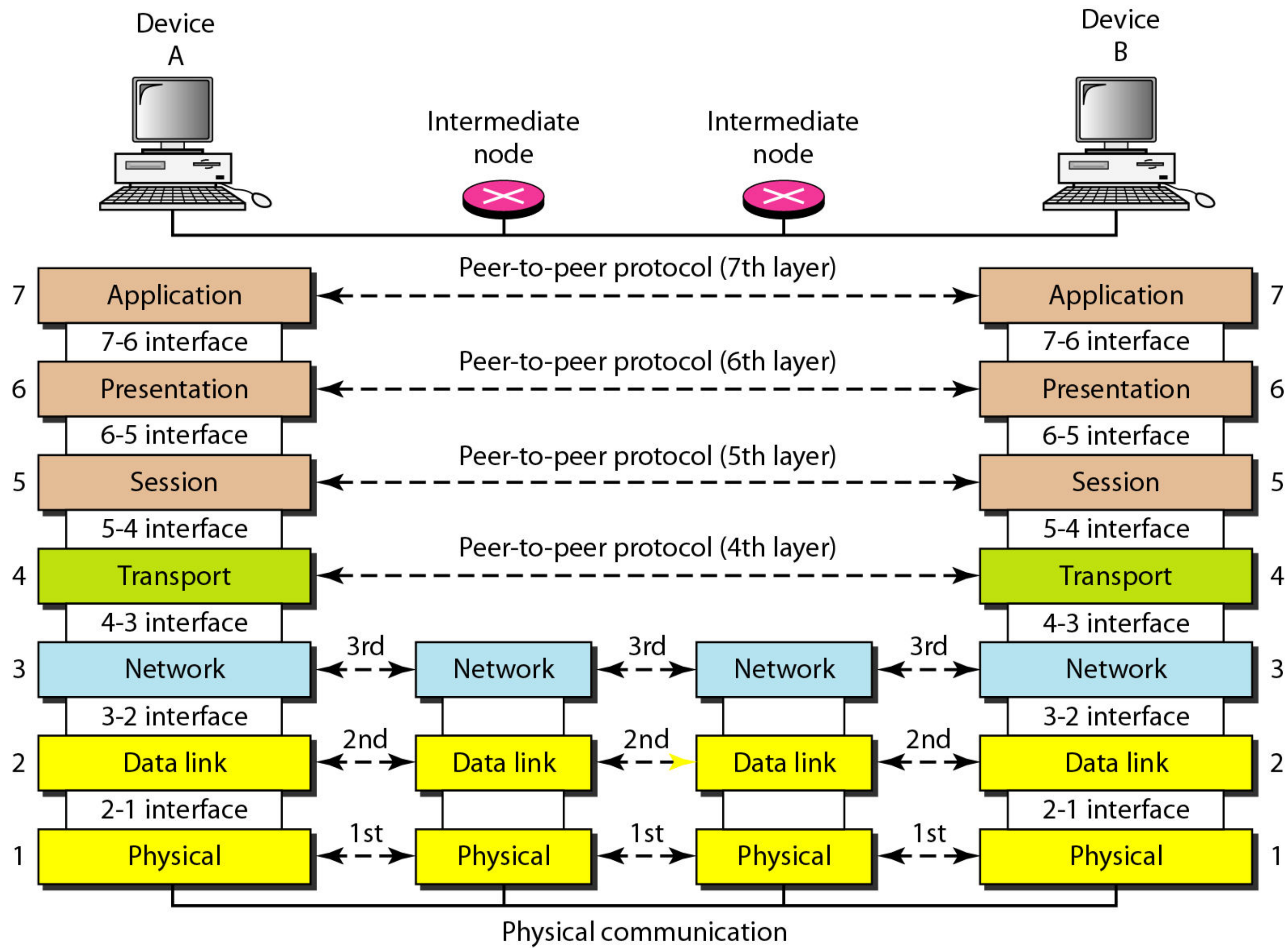
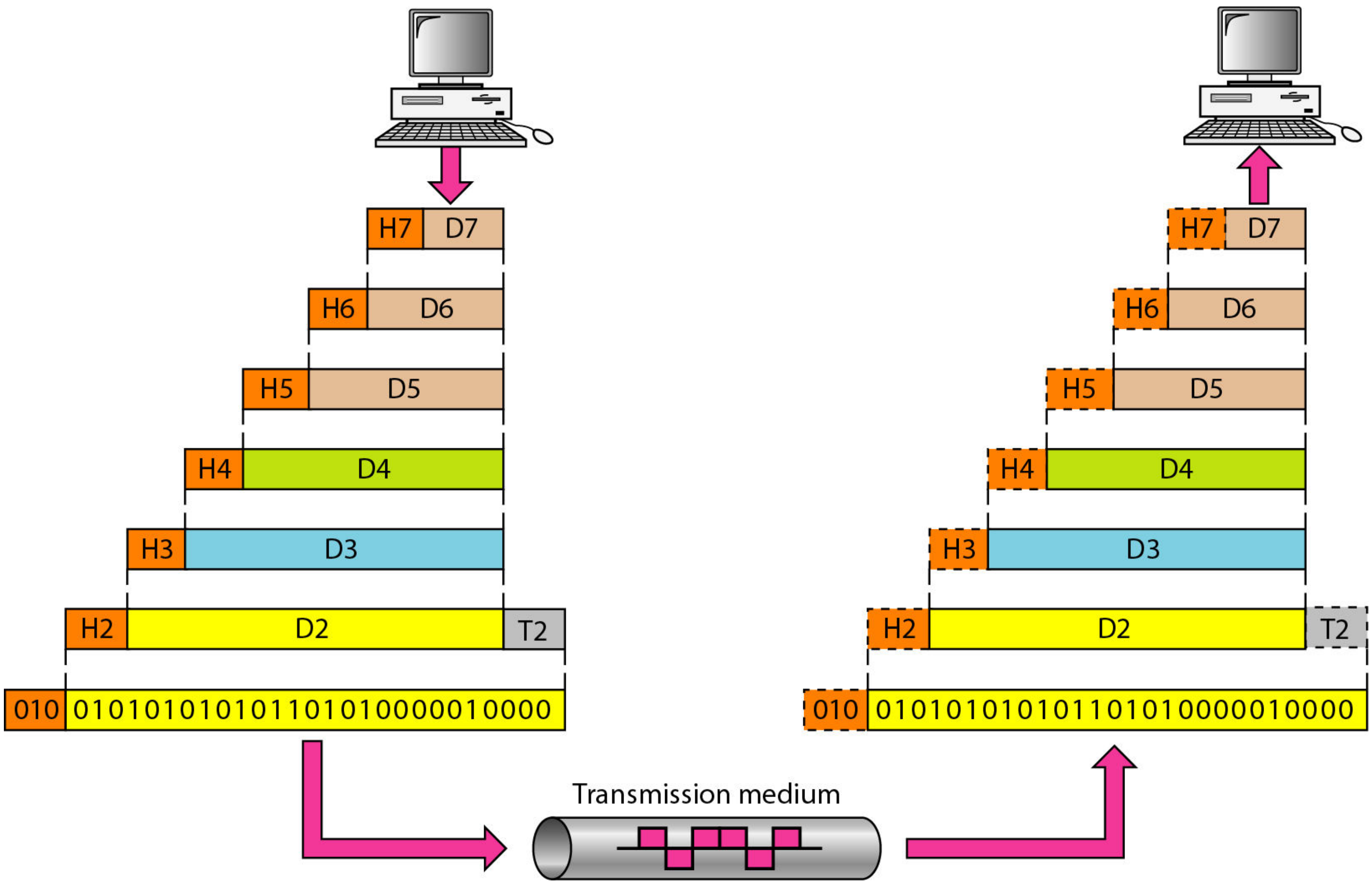


Figure 2.4 *An exchange using the OSI model*



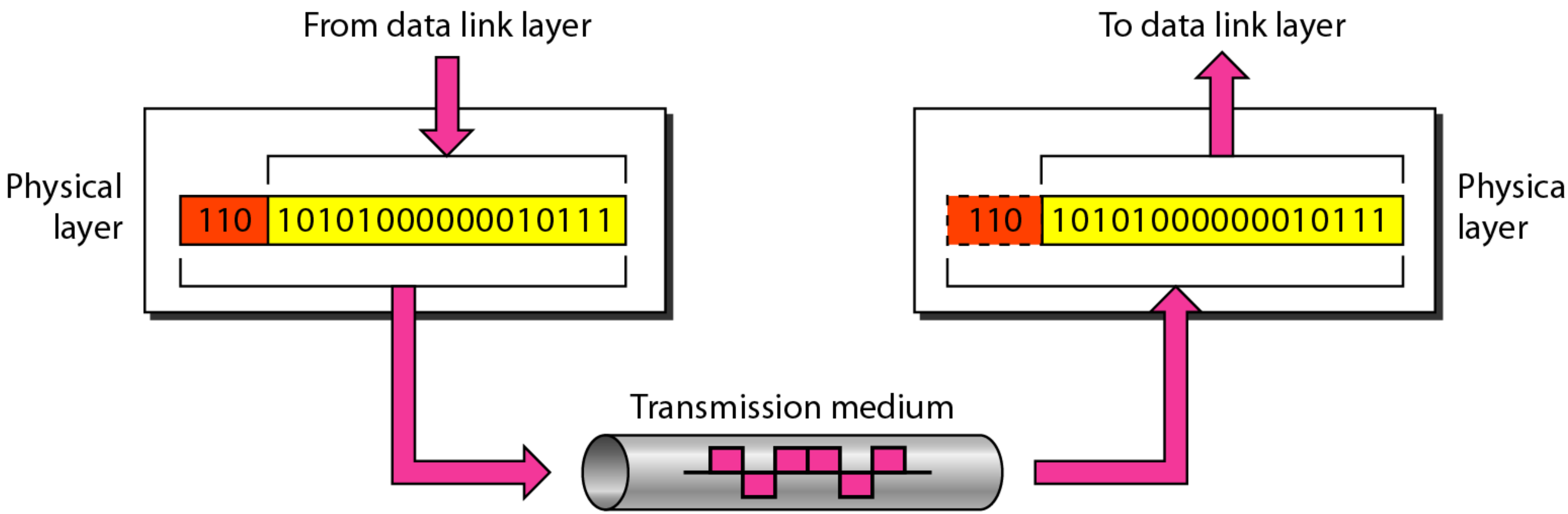
2-3 LAYERS IN THE OSI MODEL

In this section we briefly describe the functions of each layer in the OSI model.

Topics discussed in this section:

Physical Layer
Data Link Layer
Network Layer
Transport Layer
Session Layer
Presentation Layer
Application Layer

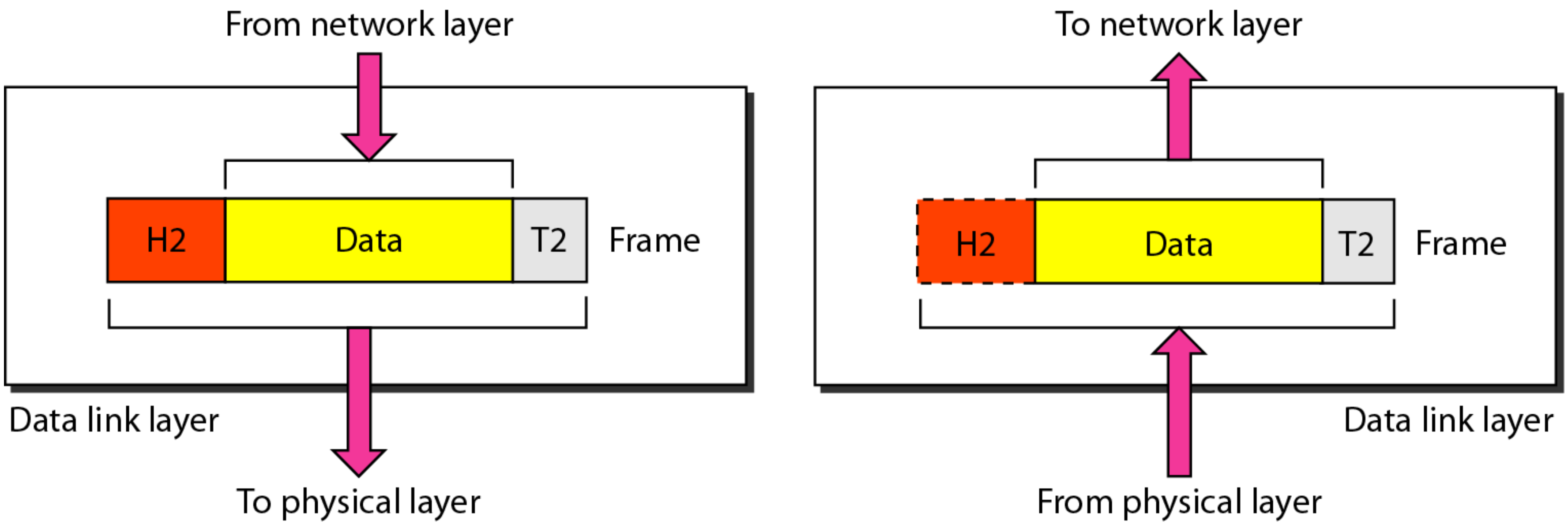
Figure 2.5 *Physical layer*



Note

The physical layer is responsible for movements of individual bits from one hop (node) to the next.

Figure 2.6 *Data link layer*



Note

The data link layer is responsible for moving frames from one hop (node) to the next.

Figure 2.7 *Hop-to-hop delivery*

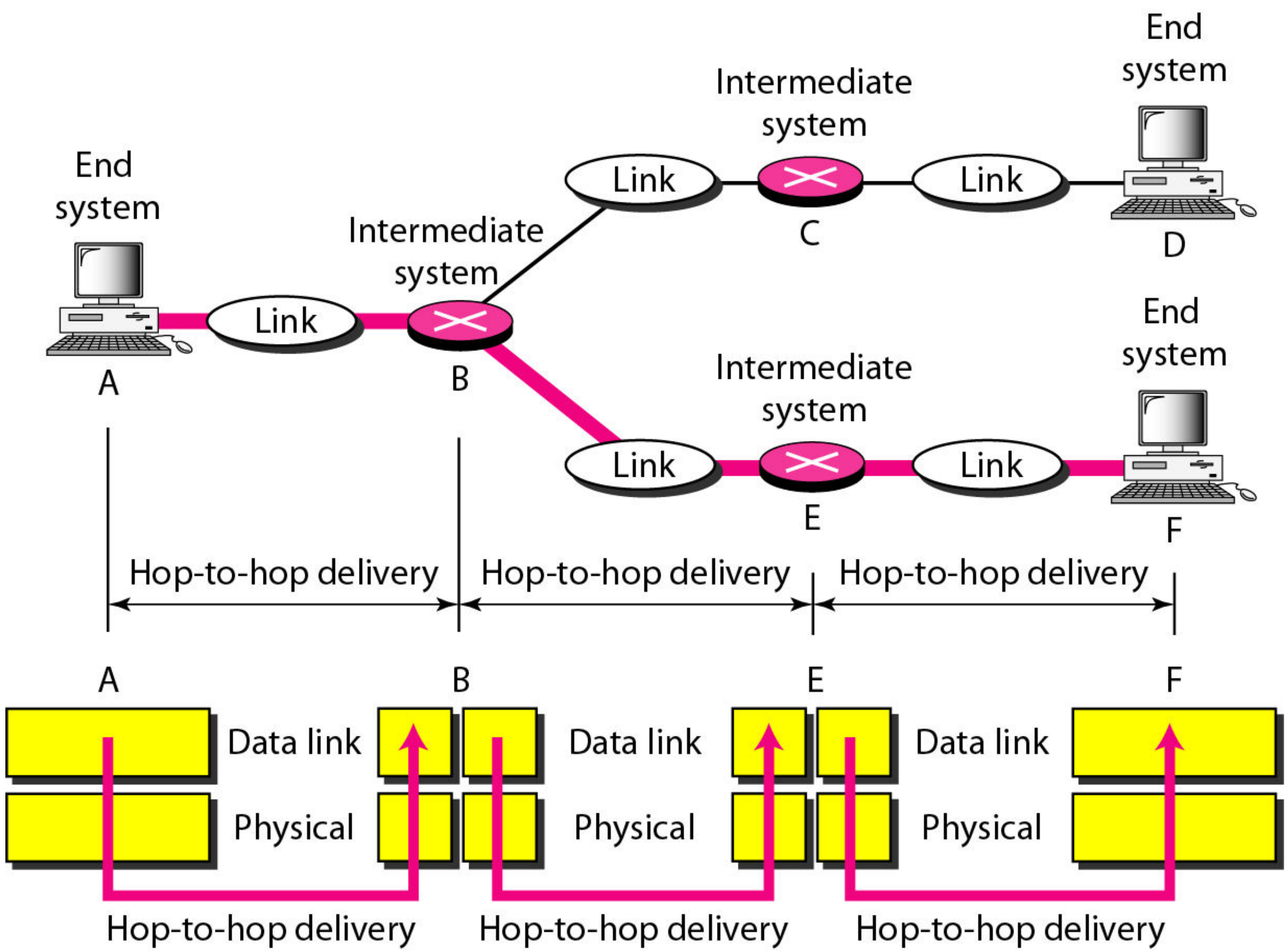
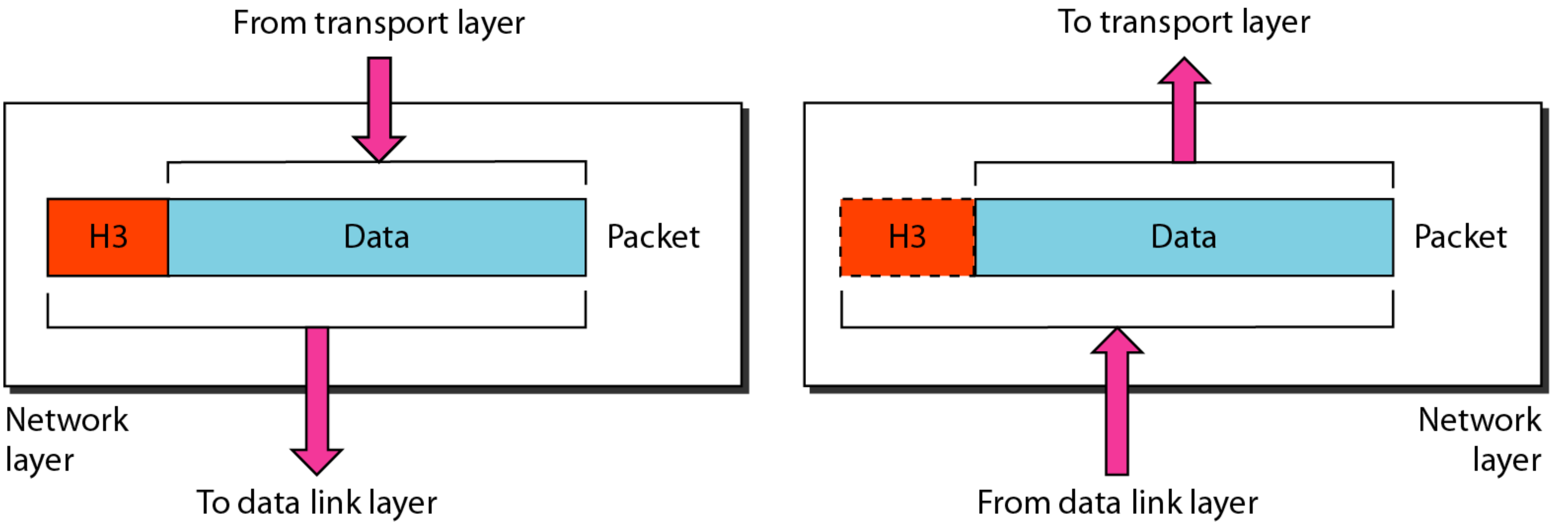


Figure 2.8 *Network layer*



Note

The network layer is responsible for the delivery of individual packets from the source host to the destination host.

Figure 2.9 *Source-to-destination delivery*

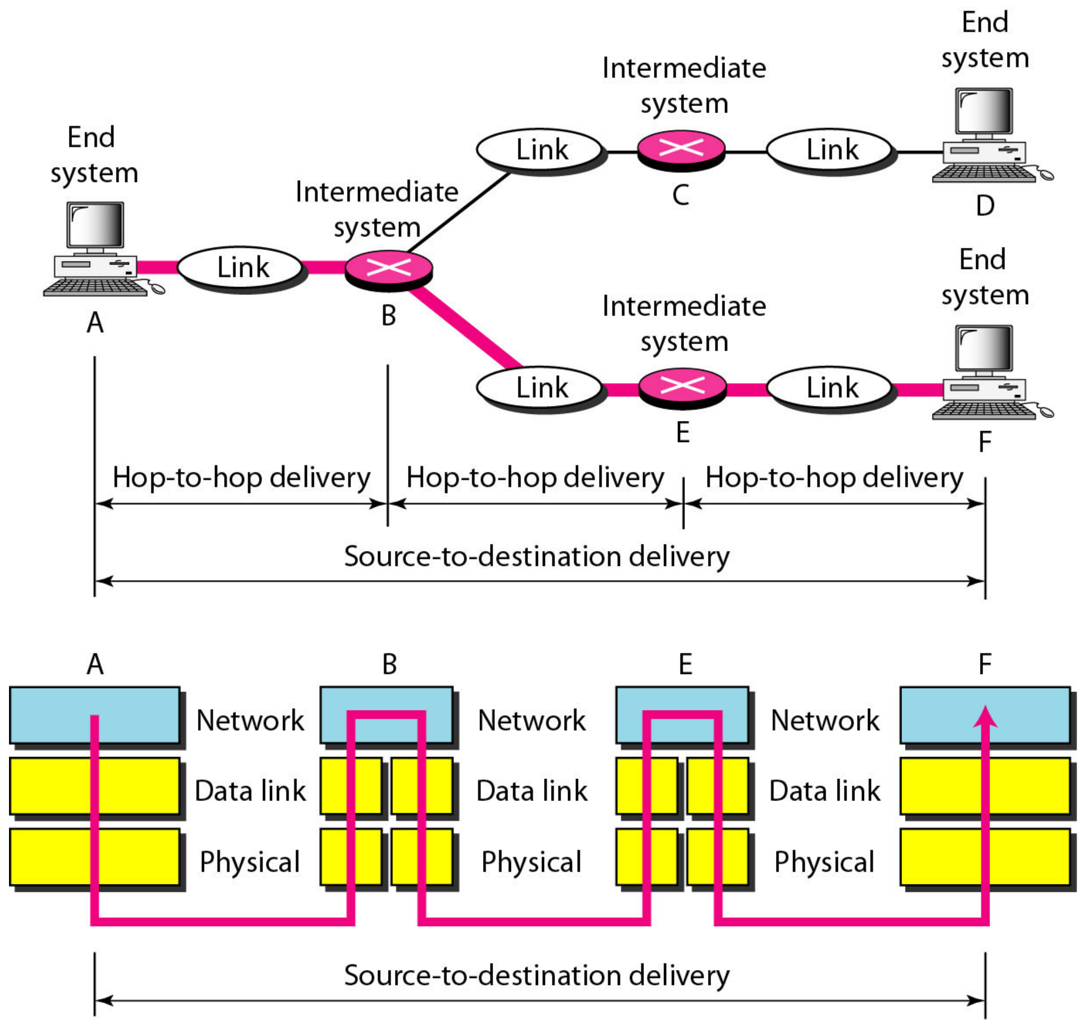
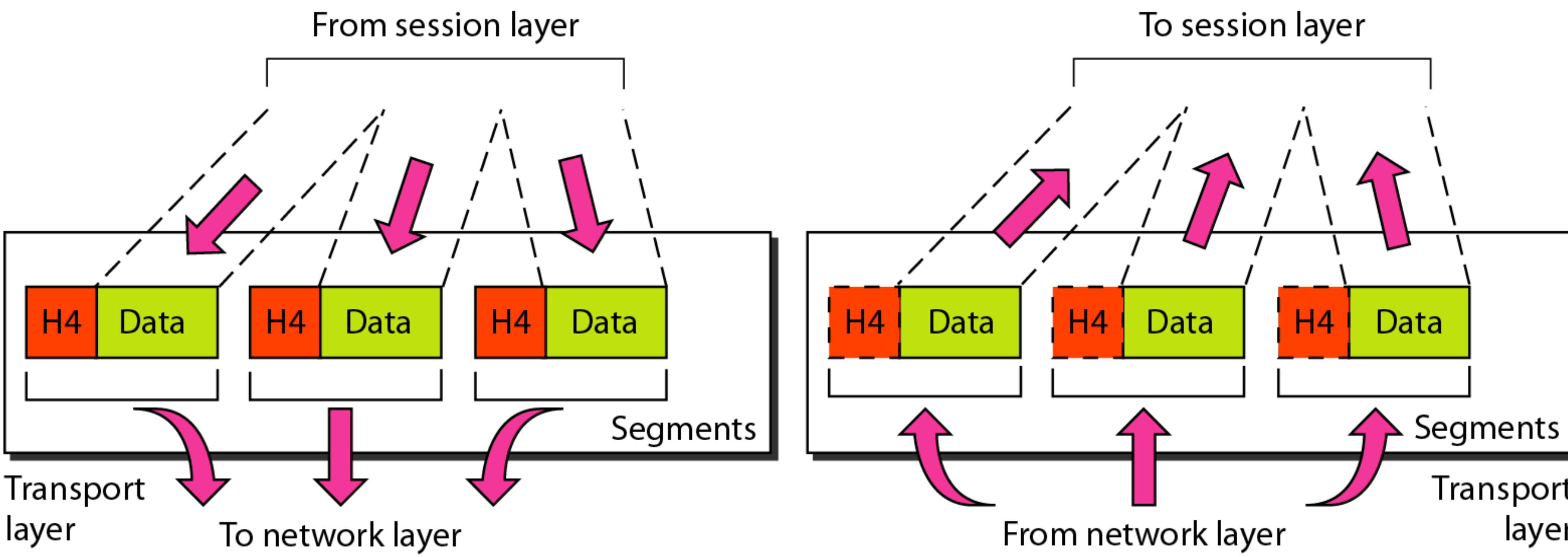


Figure 2.10 *Transport layer*



Note

The transport layer is responsible for the delivery of a message from one process to another.

Figure 2.11 *Reliable process-to-process delivery of a message*

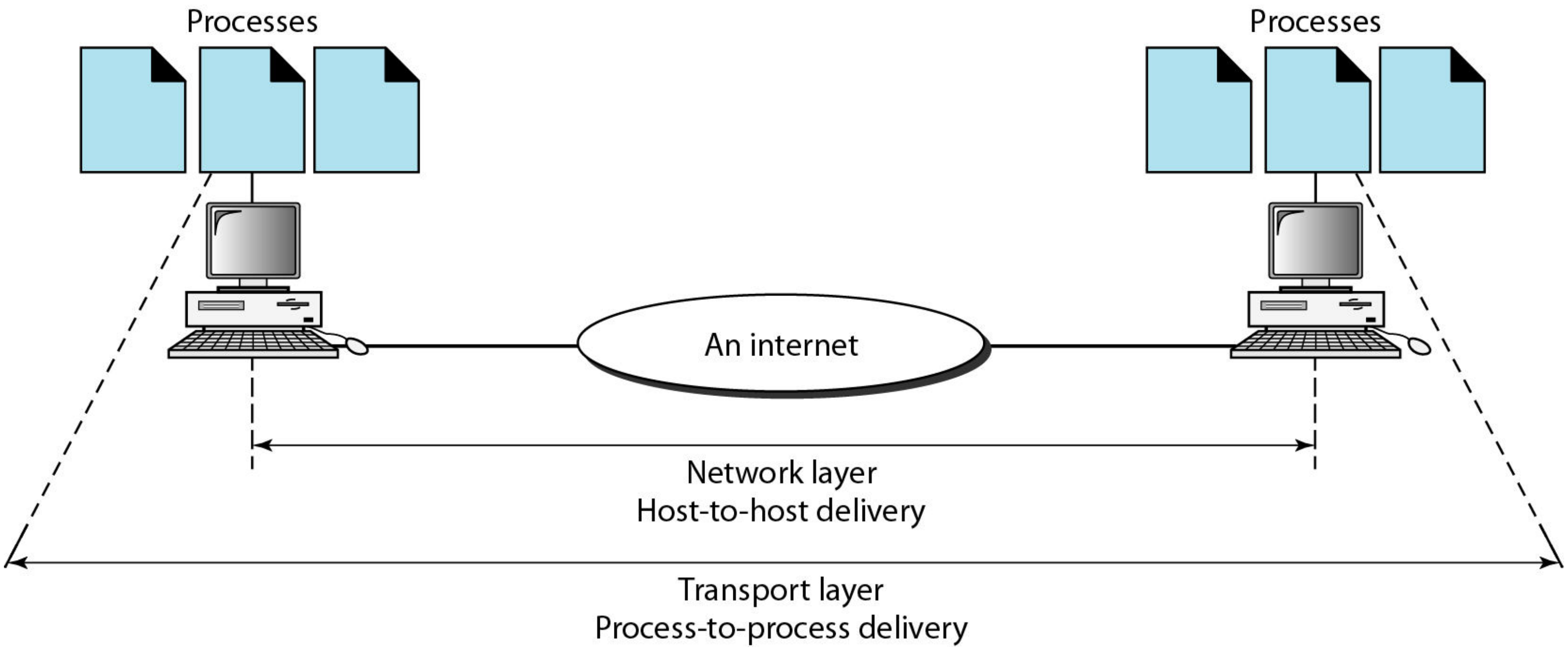
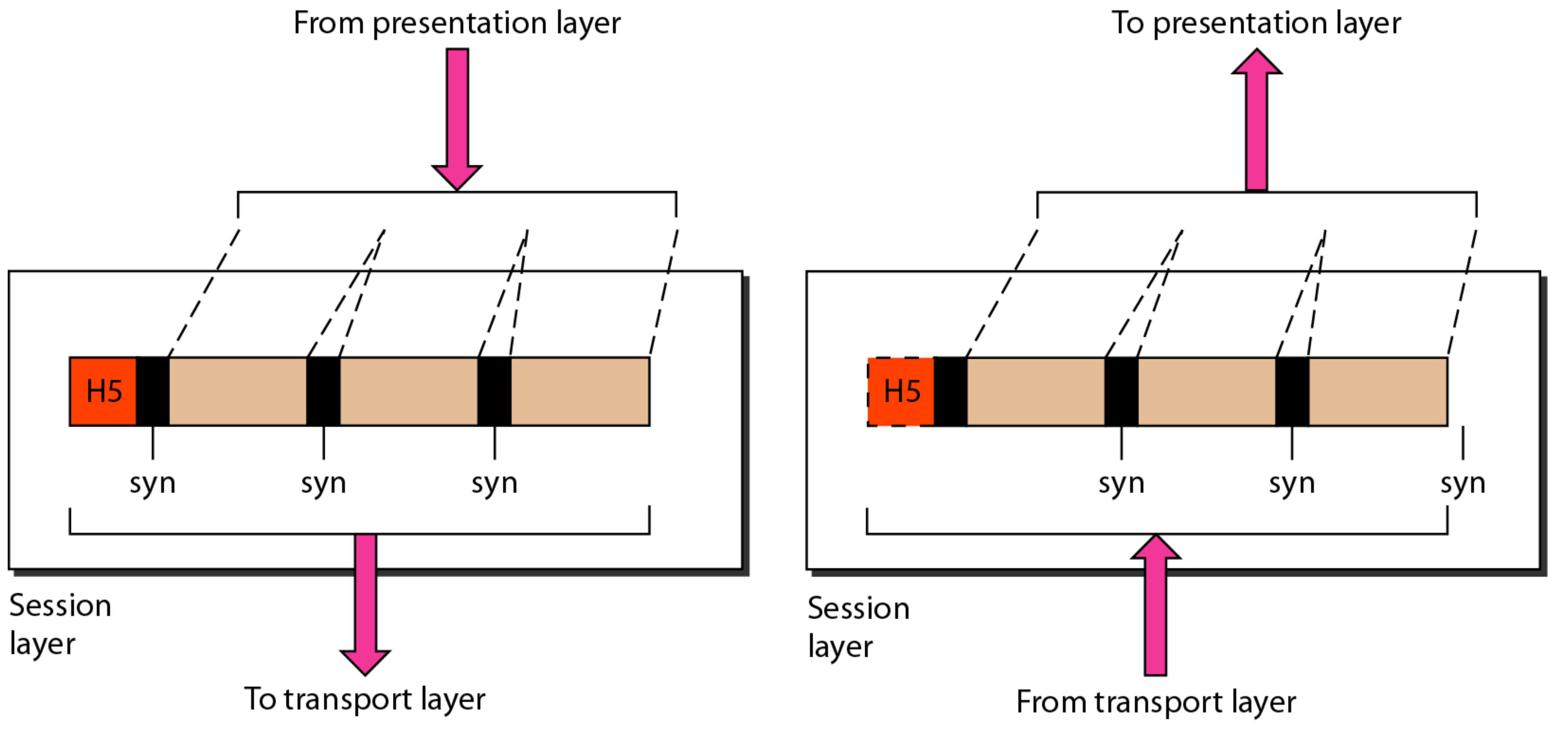


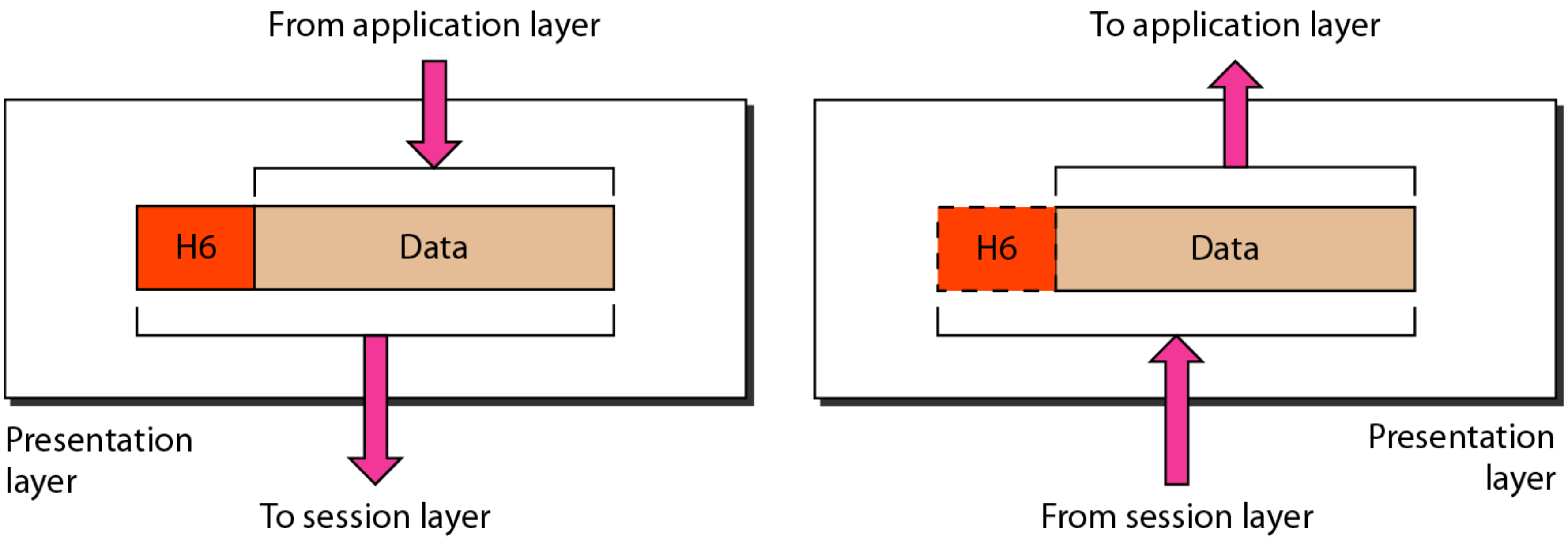
Figure 2.12 *Session layer*



Note

The session layer is responsible for dialog control and synchronization.

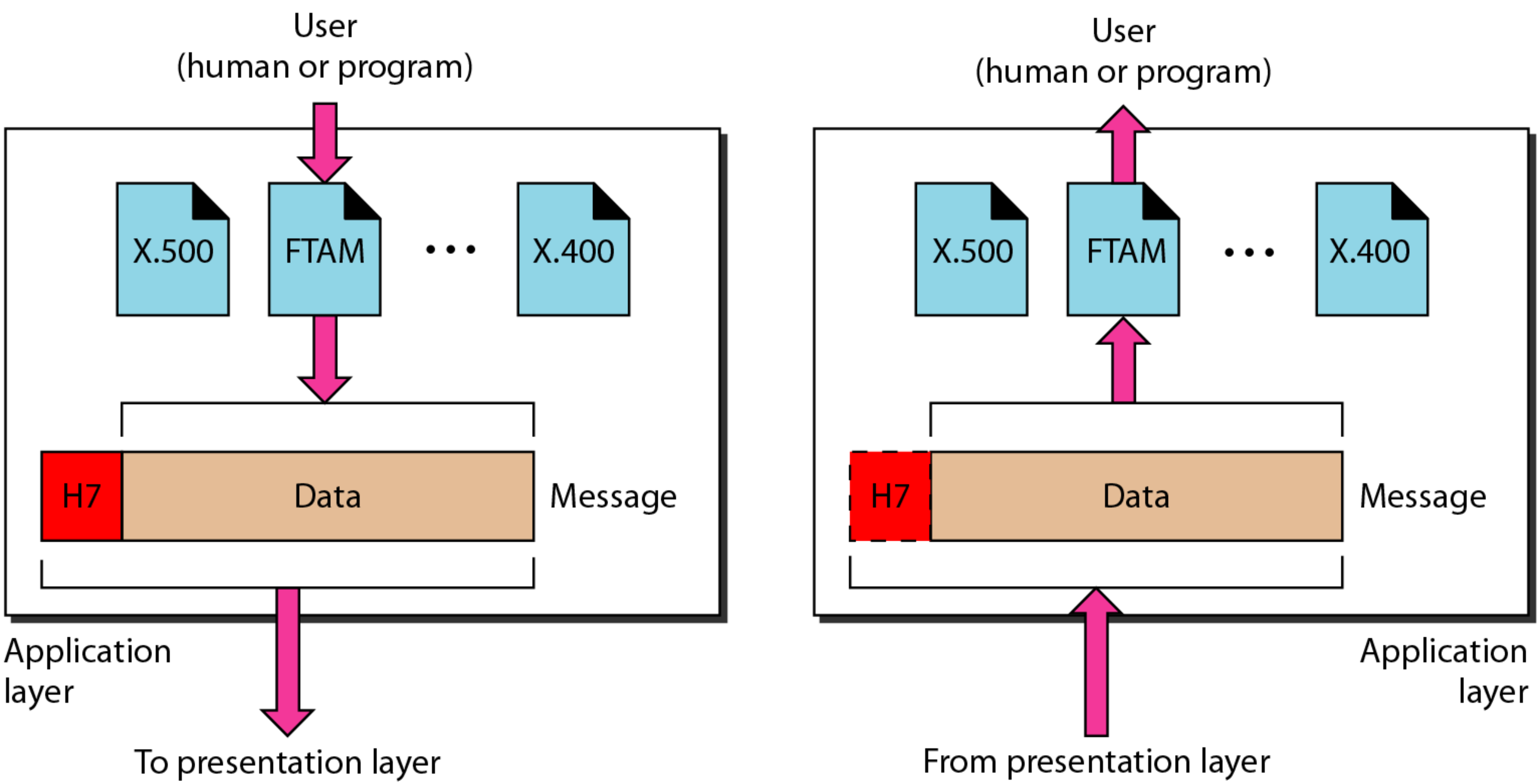
Figure 2.13 *Presentation layer*



Note

The presentation layer is responsible for translation, compression, and encryption.

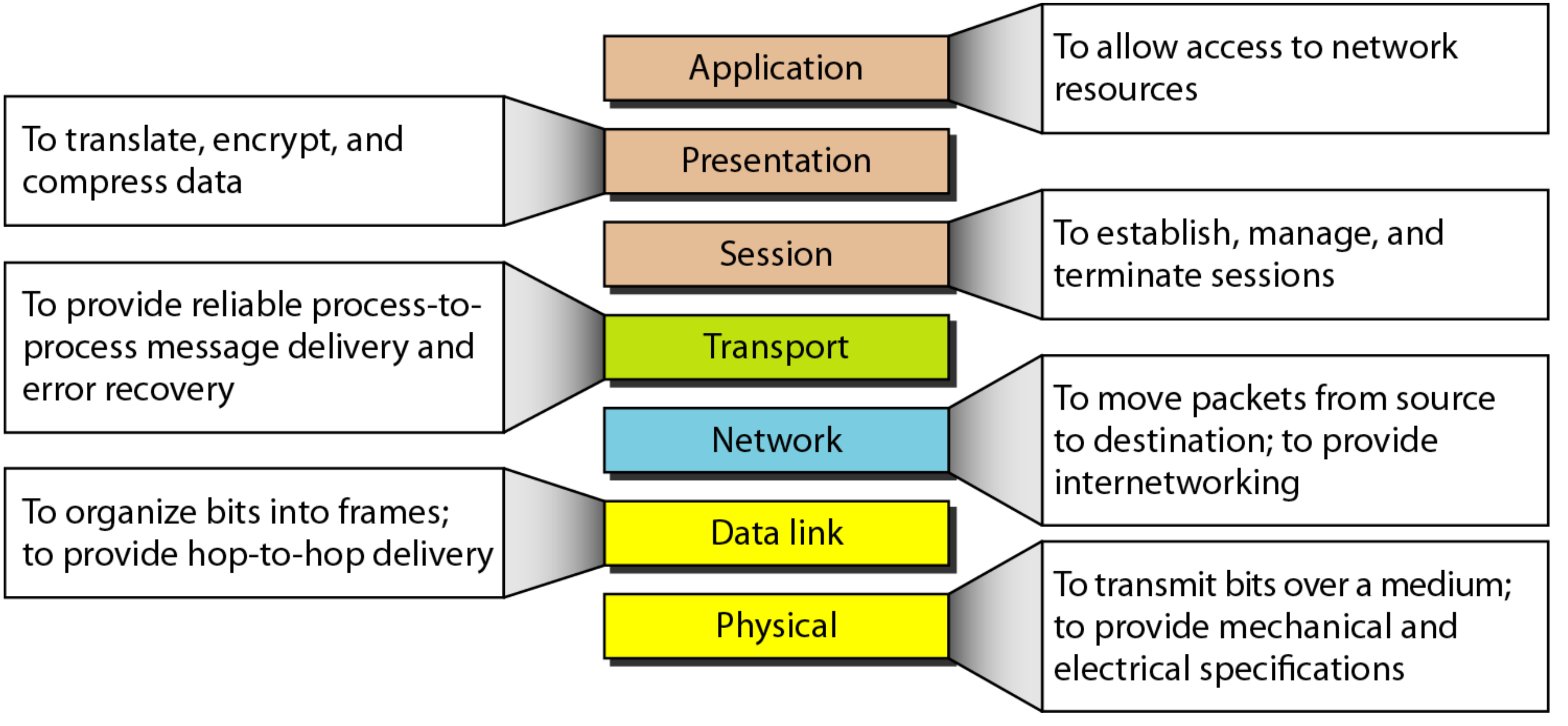
Figure 2.14 *Application layer*



Note

The application layer is responsible for providing services to the user.

Figure 2.15 *Summary of layers*



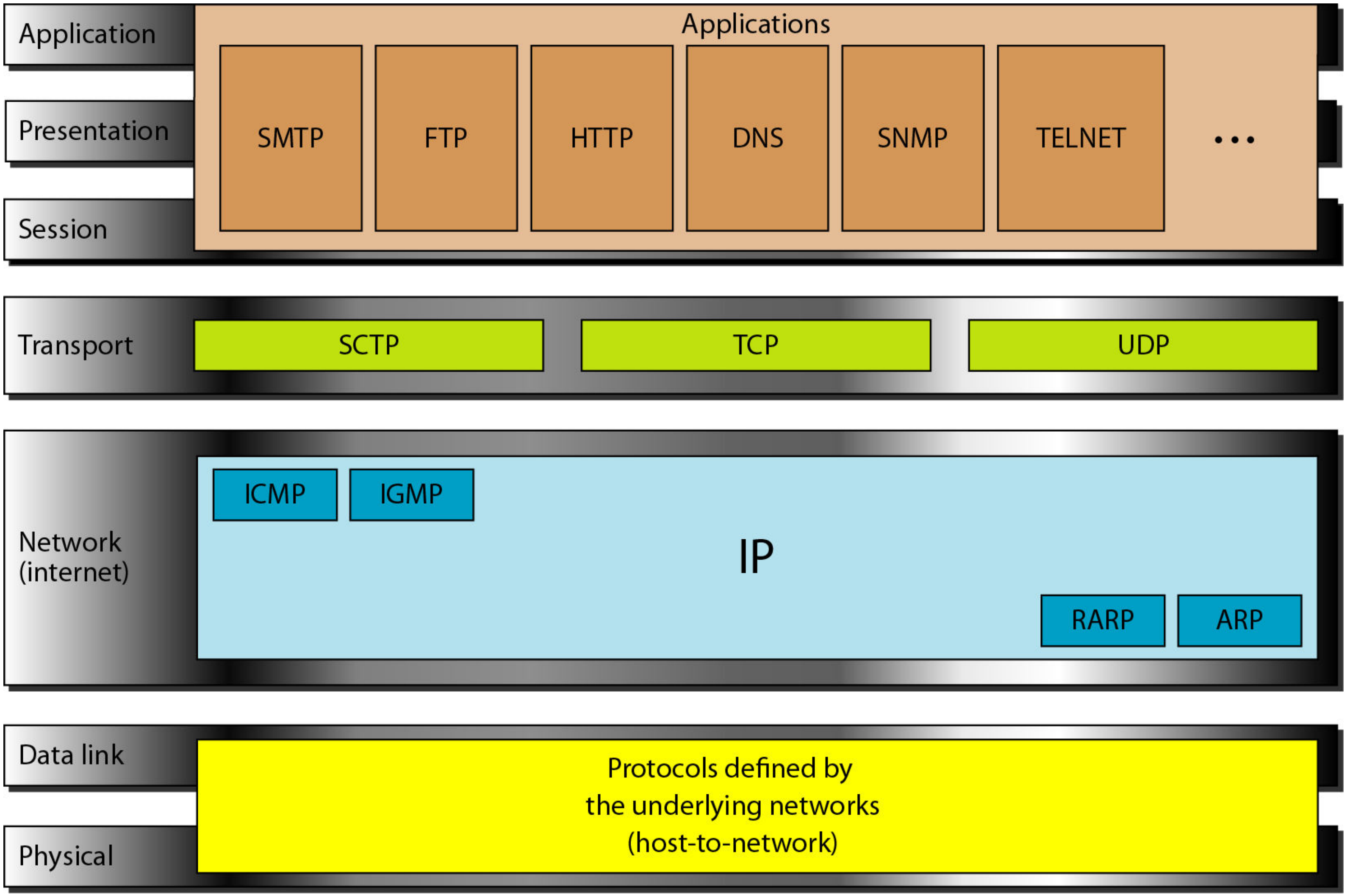
2-4 TCP/IP PROTOCOL SUITE

The layers in the TCP/IP protocol suite do not exactly match those in the OSI model. The original TCP/IP protocol suite was defined as having four layers: host-to-network, internet, transport, and application. However, when TCP/IP is compared to OSI, we can say that the TCP/IP protocol suite is made of five layers: physical, data link, network, transport, and application.

Topics discussed in this section:

Physical and Data Link Layers
Network Layer
Transport Layer
Application Layer

Figure 2.16 *TCP/IP and OSI model*



2-5 ADDRESSING

Four levels of addresses are used in an internet employing the TCP/IP protocols: physical, logical, port, and specific.

<u>Topics discussed in this section:</u>

Physical Addresses
Logical Addresses
Port Addresses
Specific Addresses

Figure 2.17 *Addresses in TCP/IP*

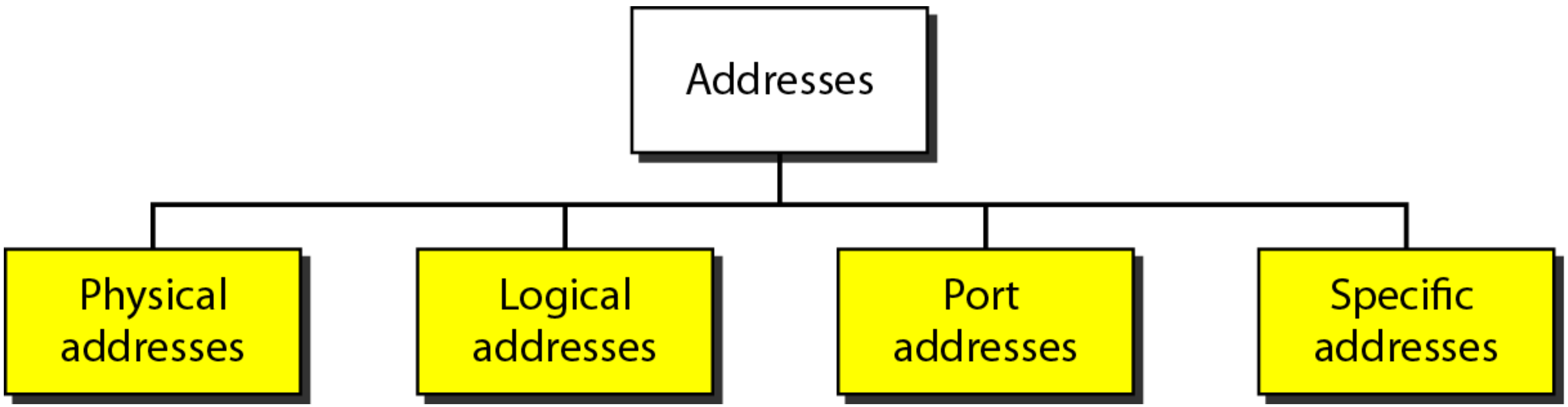
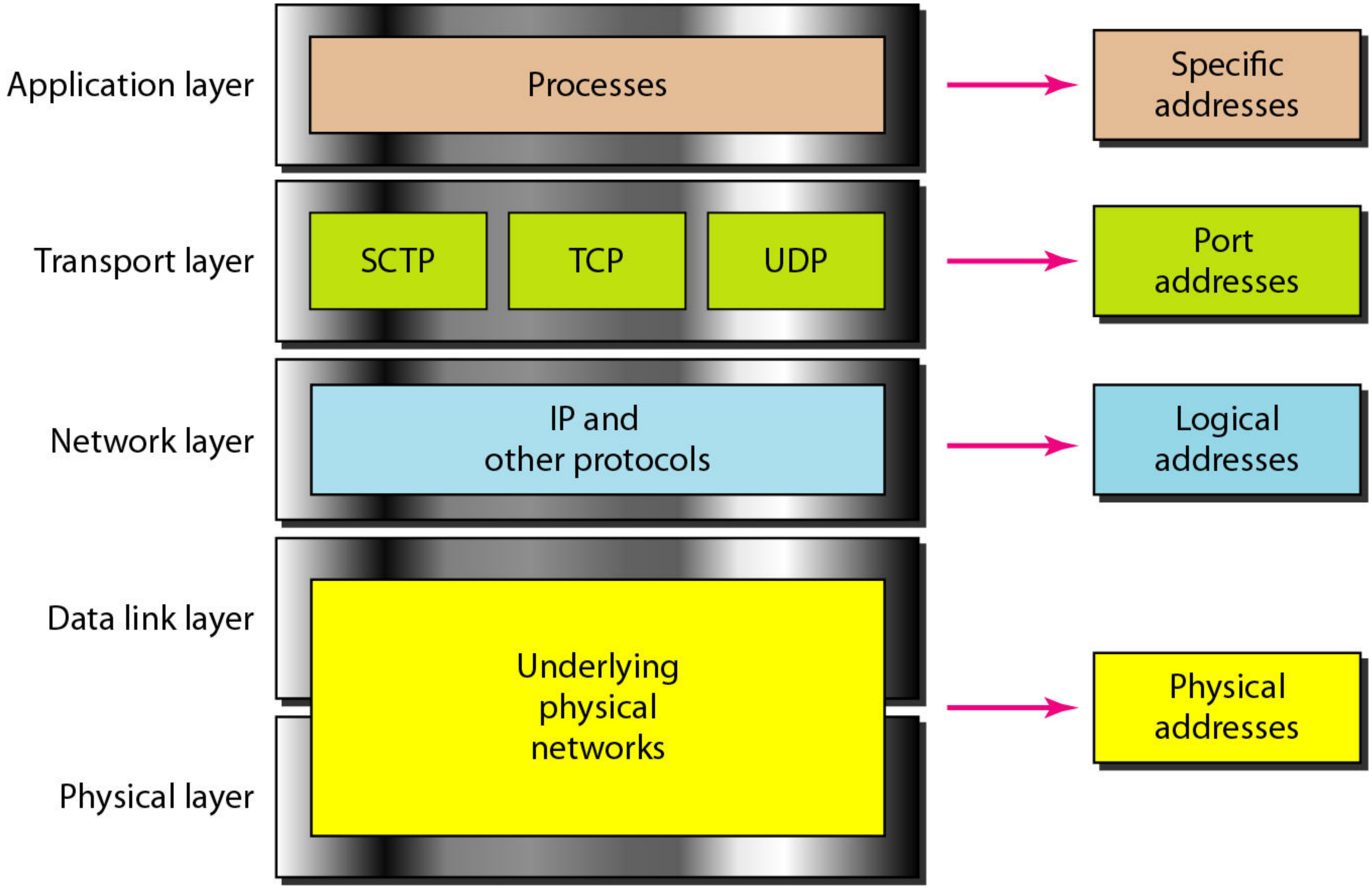


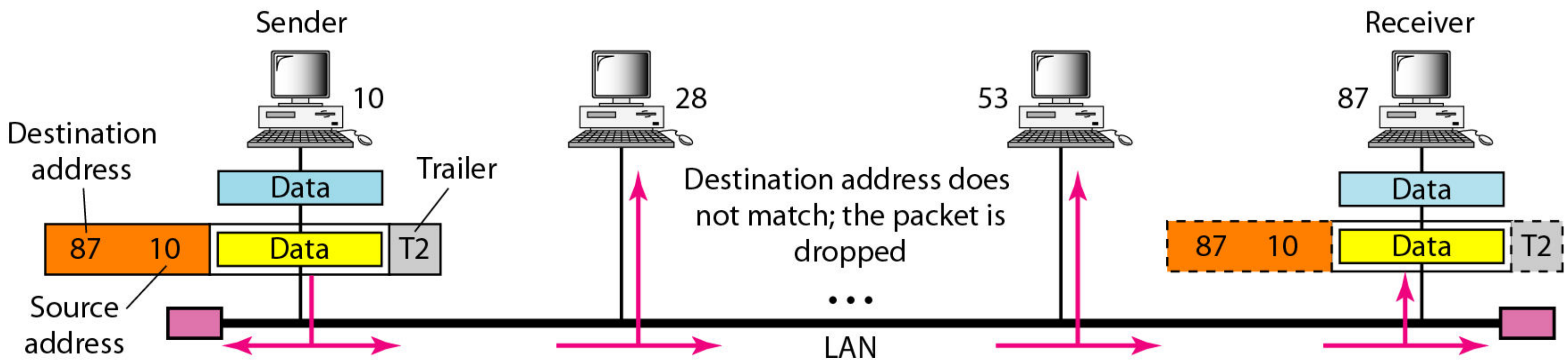
Figure 2.18 *Relationship of layers and addresses in TCP/IP*



Example 2.1

In Figure 2.19 a node with physical address 10 sends a frame to a node with physical address 87. The two nodes are connected by a link (bus topology LAN). As the figure shows, the computer with physical address 10 is the sender, and the computer with physical address 87 is the receiver.

Figure 2.19 *Physical addresses*



Example 2.2

Most local-area networks use a 48-bit (6-byte) physical address written as 12 hexadecimal digits; every byte (2 hexadecimal digits) is separated by a colon, as shown below:

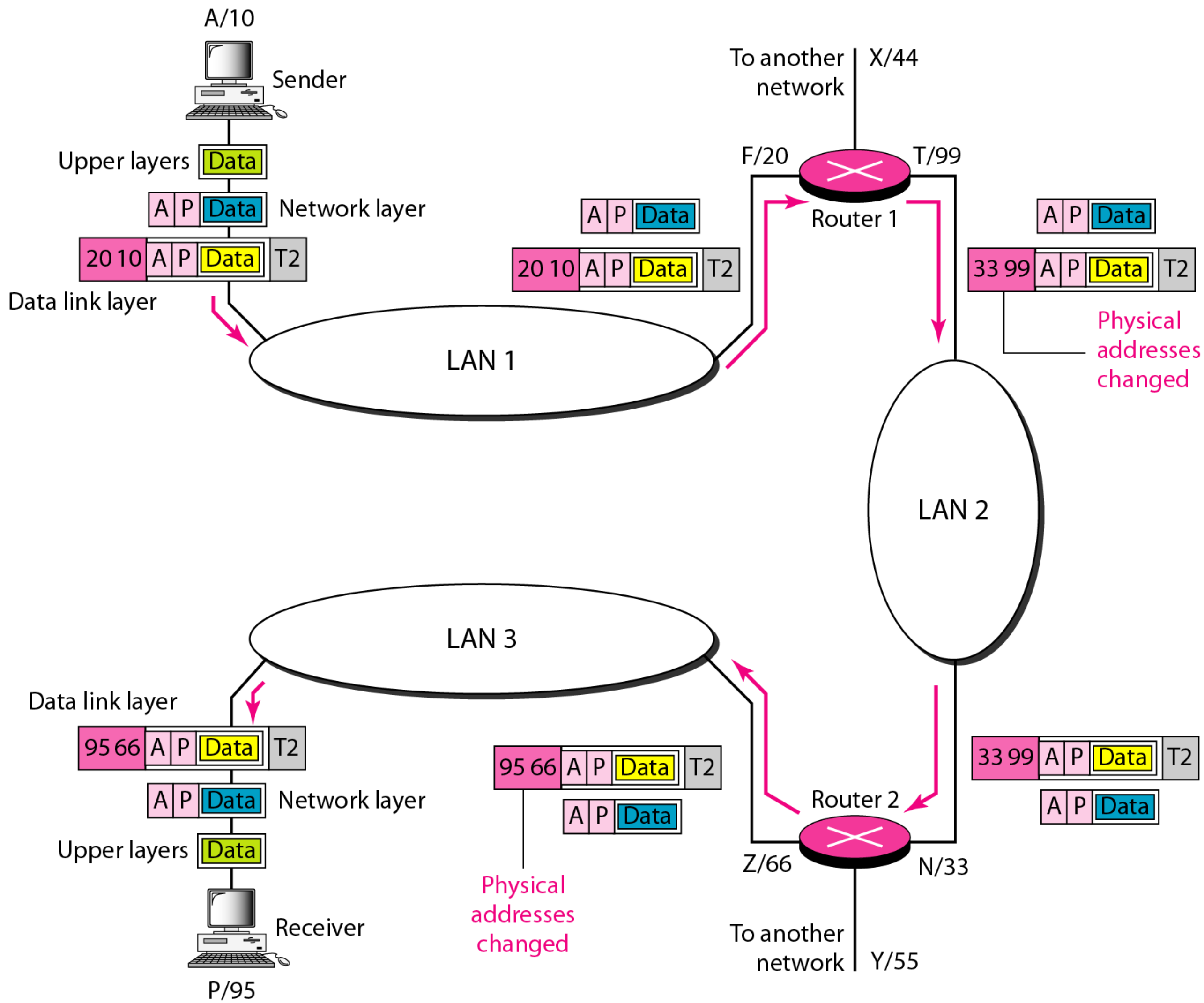
07:01:02:01:2C:4B

A 6-byte (12 hexadecimal digits) physical address.

Example 2.3

Figure 2.20 shows a part of an internet with two routers connecting three LANs. Each device (computer or router) has a pair of addresses (logical and physical) for each connection. In this case, each computer is connected to only one link and therefore has only one pair of addresses. Each router, however, is connected to three networks (only two are shown in the figure). So each router has three pairs of addresses, one for each connection.

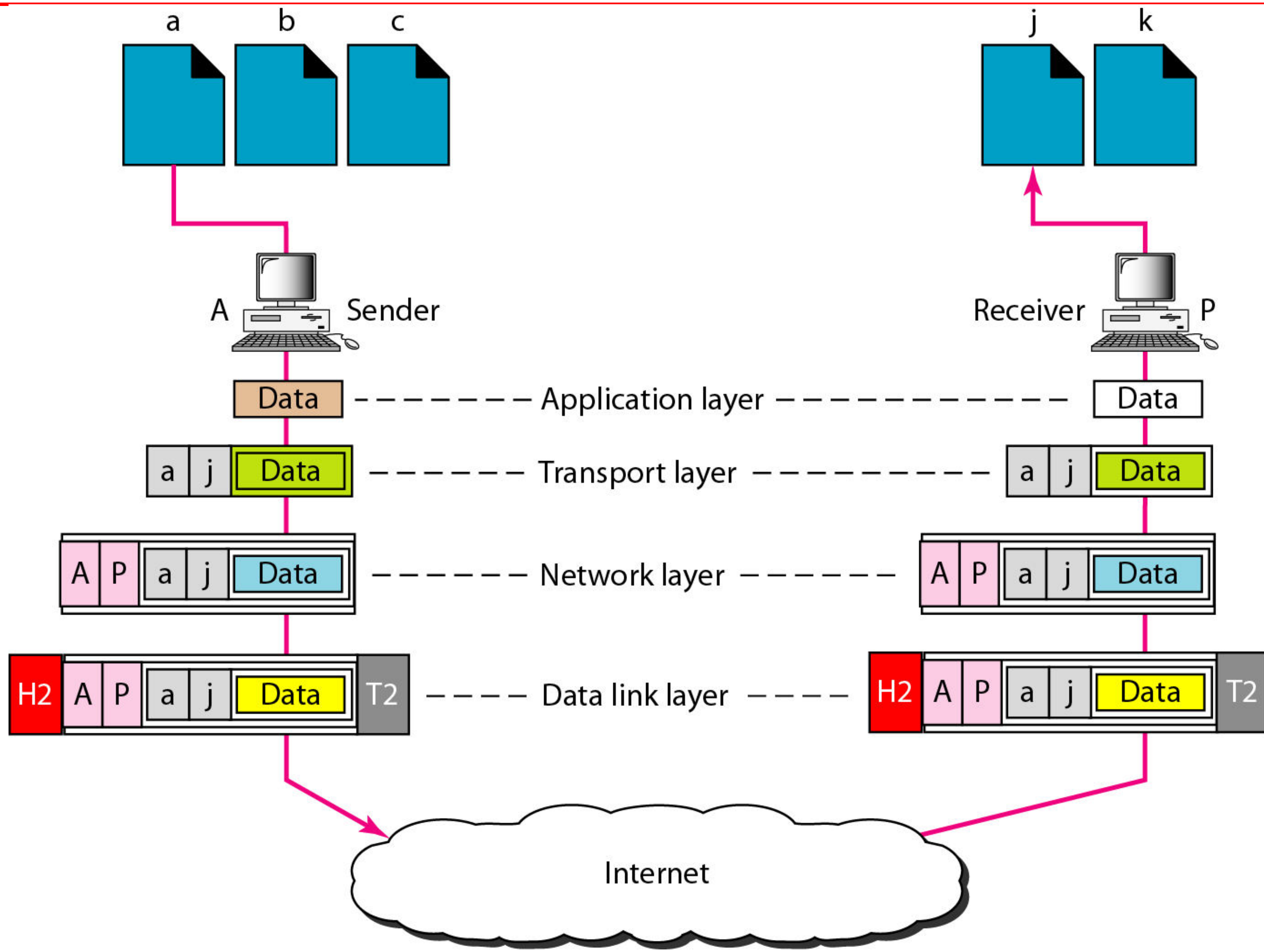
Figure 2.20 *IP addresses*



Example 2.4

Figure 2.21 shows two computers communicating via the Internet. The sending computer is running three processes at this time with port addresses a, b, and c. The receiving computer is running two processes at this time with port addresses j and k. Process a in the sending computer needs to communicate with process j in the receiving computer. Note that although physical addresses change from hop to hop, logical and port addresses remain the same from the source to destination.

Figure 2.21 *Port addresses*



Note

The physical addresses will change from hop to hop, but the logical addresses usually remain the same.

Example 2.5

A port address is a 16-bit address represented by one decimal number as shown.

753

A 16-bit port address represented as one single number.

Chapter 3

Data and Signals

Note

To be transmitted, data must be transformed to electromagnetic signals.

3-1 ANALOG AND DIGITAL

Data can be analog or digital. The term analog data refers to information that is continuous; digital data refers to information that has discrete states. Analog data take on continuous values. Digital data take on discrete values.

Topics discussed in this section:

- **Analog and Digital Data**
- **Analog and Digital Signals**
- **Periodic and Nonperiodic Signals**

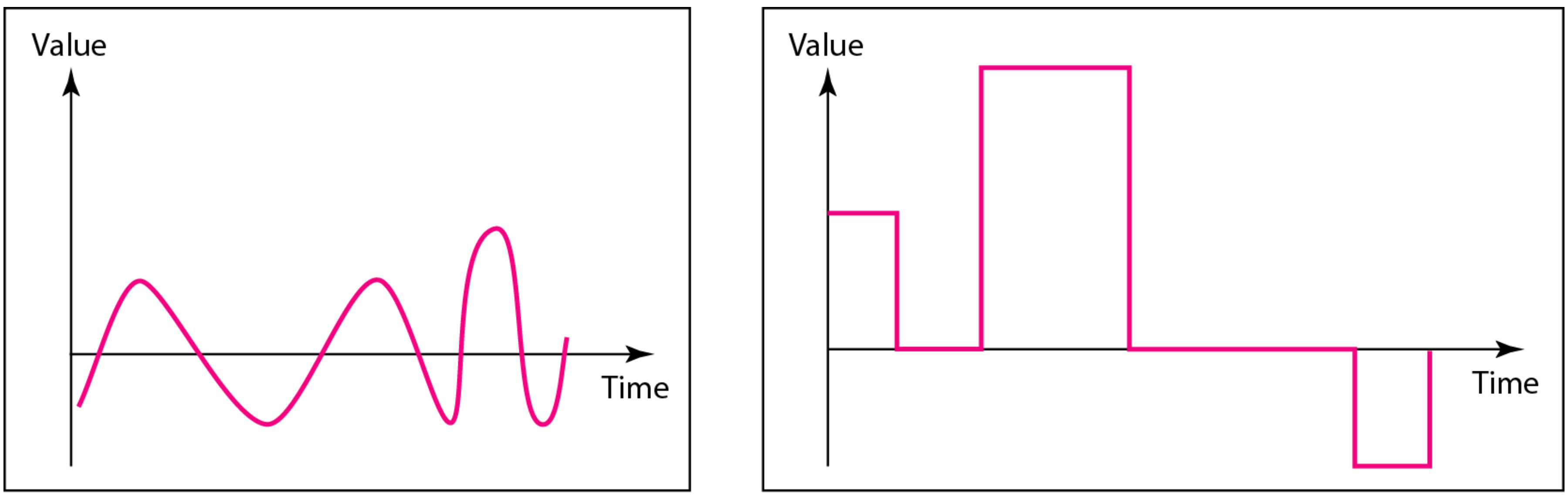
Analog and Digital Data

- Data can be analog or digital.
- Analog data are continuous and take continuous values.
- Digital data have discrete states and take discrete values.

Analog and Digital Signals

- Signals can be analog or digital.
- Analog signals can have an infinite number of values in a range.
- Digital signals can have only a limited number of values.

Figure 3.1 *Comparison of analog and digital signals*



a. Analog signal

b. Digital signal

3-2 PERIODIC ANALOG SIGNALS

In data communications, we commonly use periodic analog signals and nonperiodic digital signals.

Periodic analog signals can be classified as simple or composite. A simple periodic analog signal, a sine wave, cannot be decomposed into simpler signals. A composite periodic analog signal is composed of multiple sine waves.

Topics discussed in this section:

- **Sine Wave**
- **Wavelength**
- **Time and Frequency Domain**
- **Composite Signals**
- **Bandwidth**

Figure 3.2 *A sine wave*

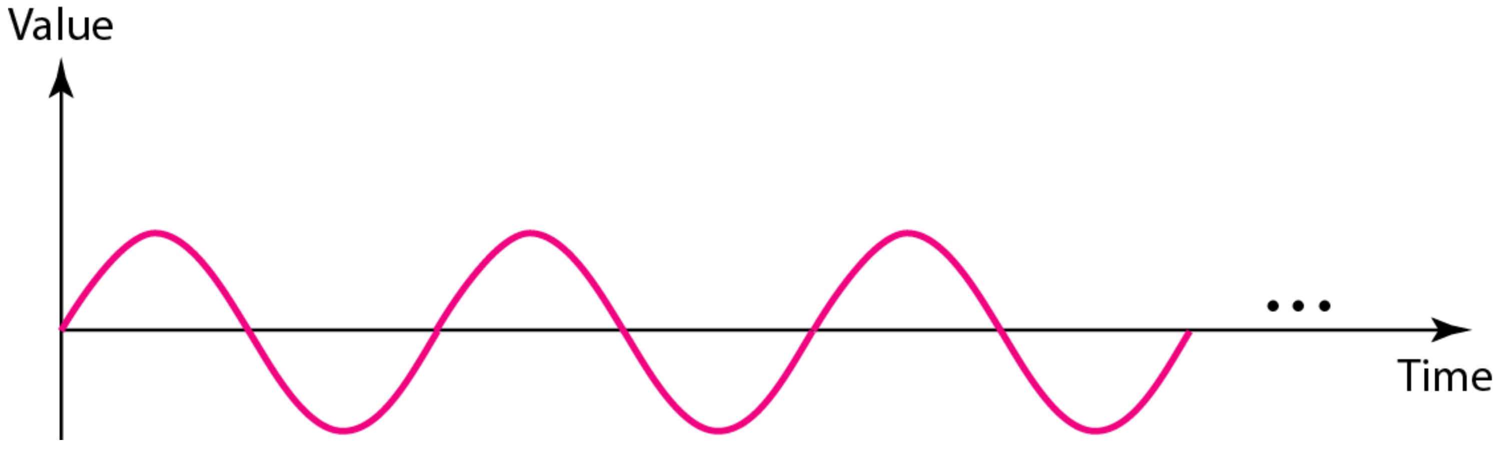
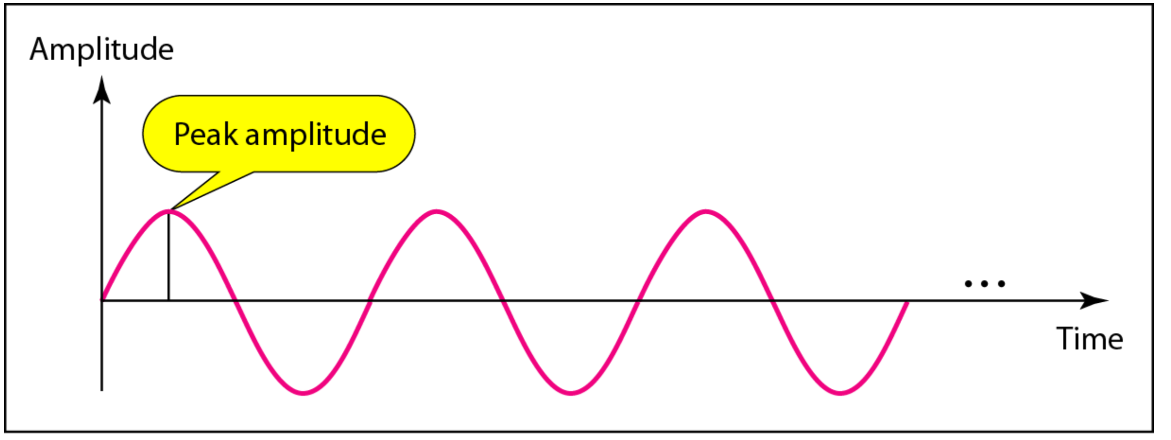
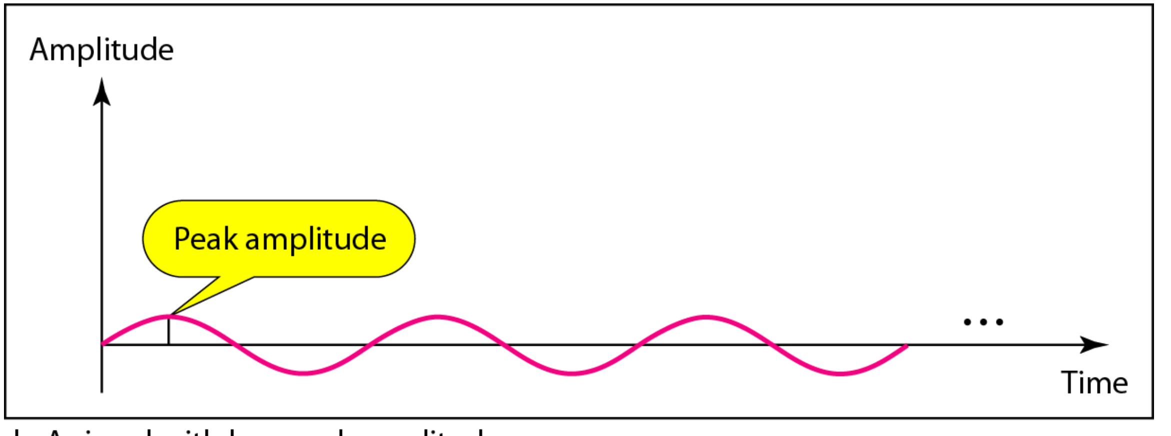


Figure 3.3 *Two signals with the same phase and frequency, but different amplitudes*



a. A signal with high peak amplitude



b. A signal with low peak amplitude

Note

Frequency and period are the inverse of each other.

$$f = \frac{1}{T} \qquad \text{and} \qquad T = \frac{1}{f}$$

Figure 3.4 *Two signals with the same amplitude and phase, but different frequencies*

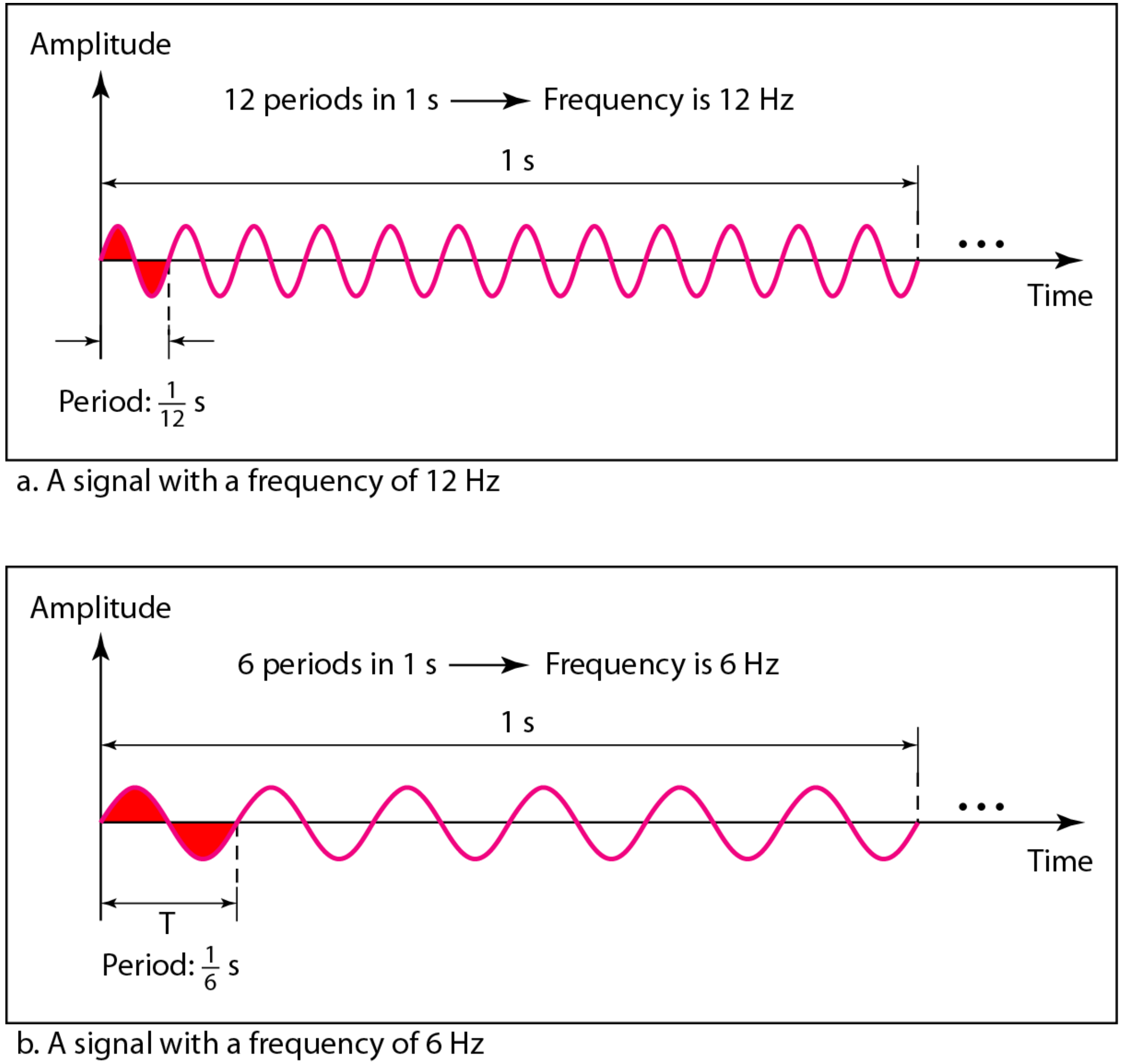


Table 3.1 *Units of period and frequency*

Unit	*Equivalent*	*Unit*	*Equivalent*
Seconds (s)	1 s	Hertz (Hz)	1 Hz
Milliseconds (ms)	10^{-3} s	Kilohertz (kHz)	10^{3} Hz
Microseconds (μs)	10^{-6} s	Megahertz (MHz)	10^{6} Hz
Nanoseconds (ns)	10^{-9} s	Gigahertz (GHz)	10^{9} Hz
Picoseconds (ps)	10^{-12} s	Terahertz (THz)	10^{12} Hz

Example 3.1

The power we use at home has a frequency of 60 Hz. The period of this sine wave can be determined as follows:

$$\boxed{T = \frac{1}{f} = \frac{1}{60} = 0.0166 \text{ s} = 0.0166 \times 10^3 \text{ ms} = 16.6 \text{ ms}}$$

Example 3.2

The period of a signal is 100 ms. What is its frequency in kilohertz?

Solution

First we change 100 ms to seconds, and then we calculate the frequency from the period (1 Hz = 10^{-3} kHz).

$$100 \text{ ms} = 100 \times 10^{-3} \text{ s} = 10^{-1} \text{ s}$$

$$f = \frac{1}{T} = \frac{1}{10^{-1}} \text{ Hz} = 10 \text{ Hz} = 10 \times 10^{-3} \text{ kHz} = 10^{-2} \text{ kHz}$$

Frequency

- Frequency is the rate of change with respect to time.
- Change in a short span of time means high frequency.
- Change over a long span of time means low frequency.

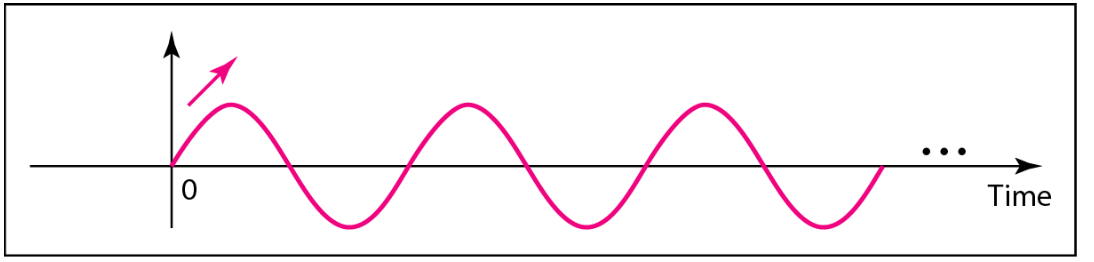
Note

If a signal does not change at all, its frequency is zero.
If a signal changes instantaneously, its frequency is infinite.

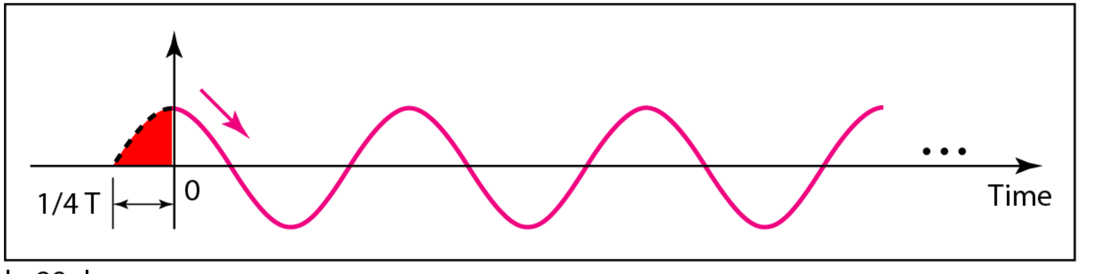
Note

Phase describes the position of the waveform relative to time 0.

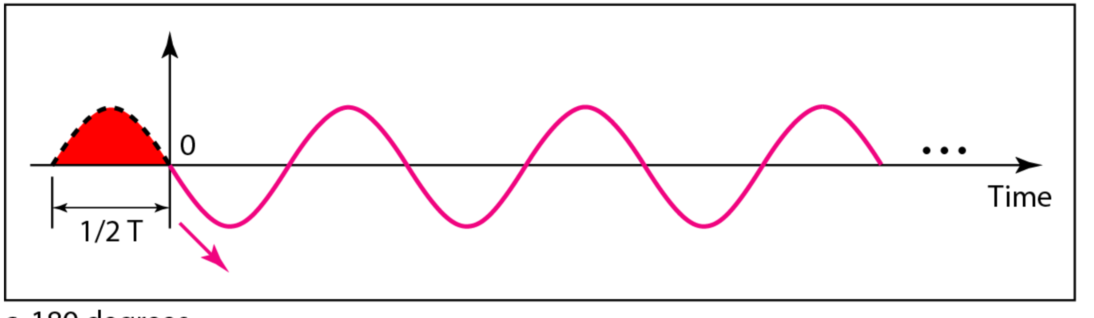
Figure 3.5 *Three sine waves with the same amplitude and frequency, but different phases*



a. 0 degrees



b. 90 degrees



c. 180 degrees

Example 3.3

A sine wave is offset 1/6 cycle with respect to time 0. What is its phase in degrees and radians?

Solution

We know that 1 complete cycle is 360°. Therefore, 1/6 cycle is

$$\boxed{\frac{1}{6} \times 360 = 60° = 60 \times \frac{2\pi}{360} \text{ rad} = \frac{\pi}{3} \text{ rad} = 1.046 \text{ rad}}$$

Figure 3.6 *Wavelength and period*

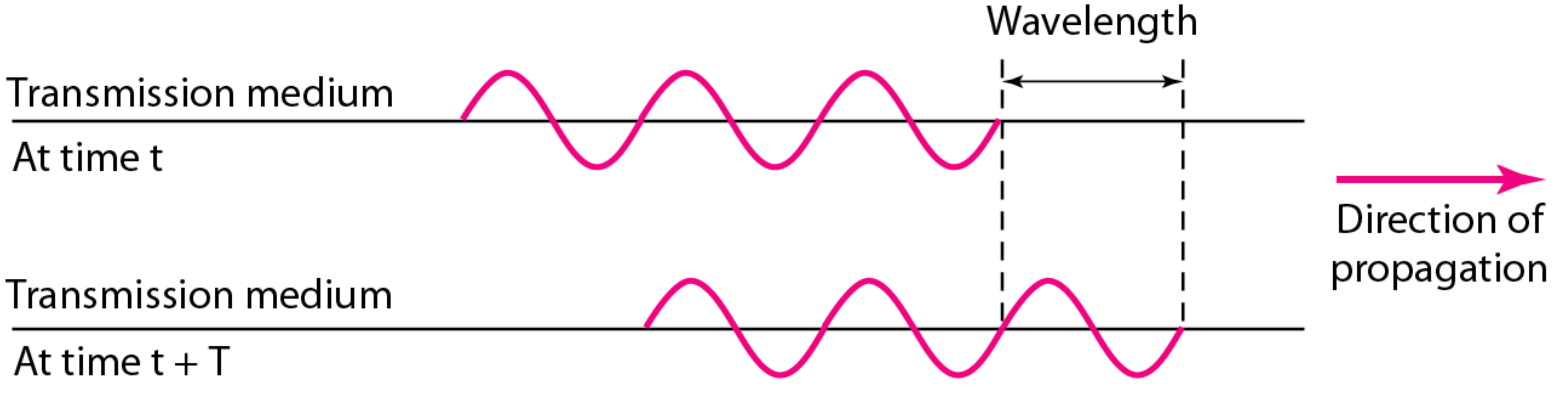
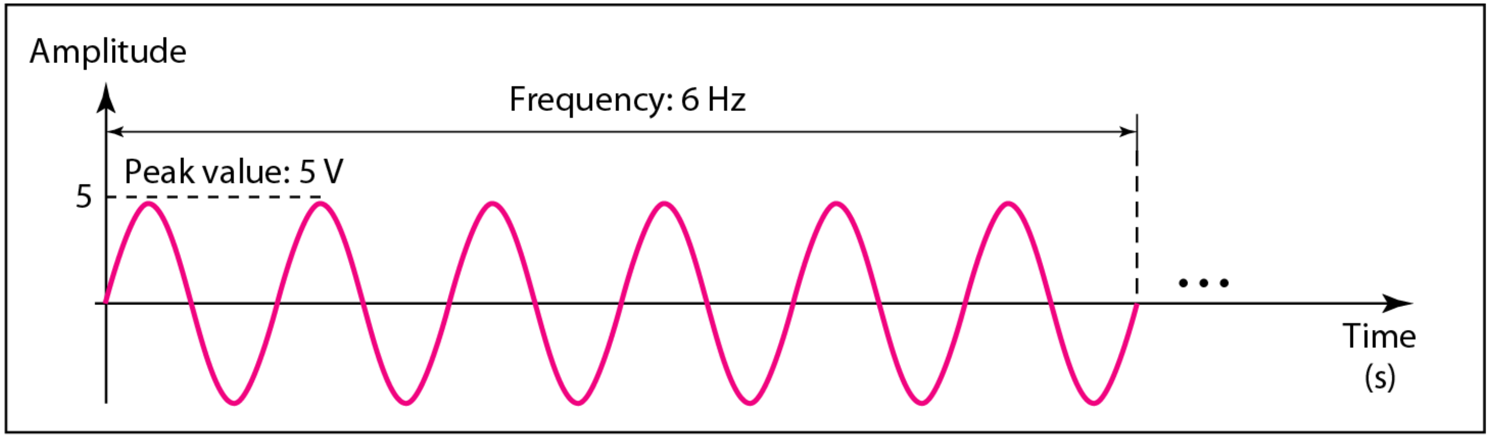
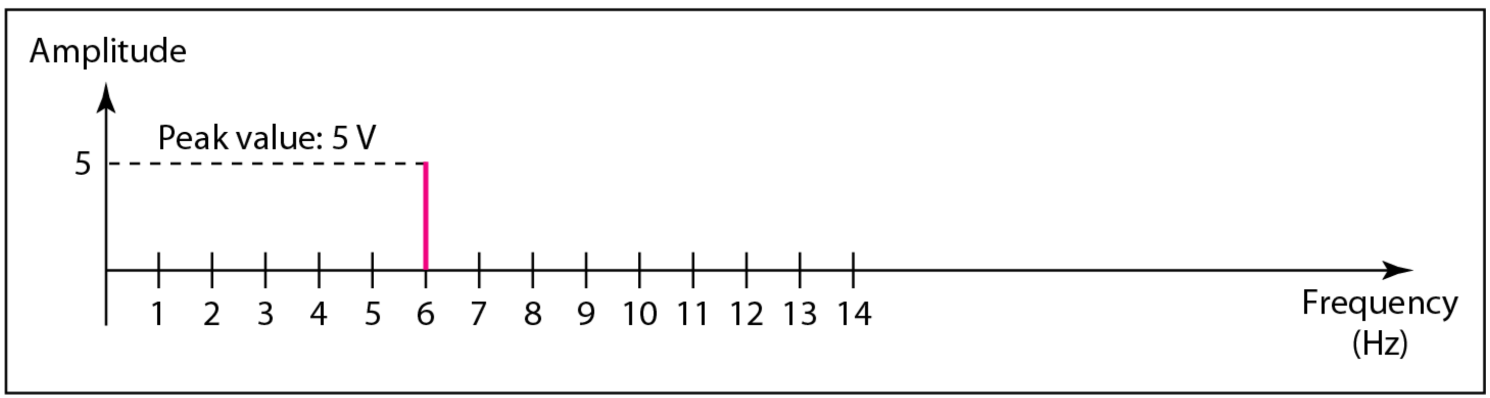


Figure 3.7 *The time-domain and frequency-domain plots of a sine wave*



a. A sine wave in the time domain (peak value: 5 V, frequency: 6 Hz)



b. The same sine wave in the frequency domain (peak value: 5 V, frequency: 6 Hz)

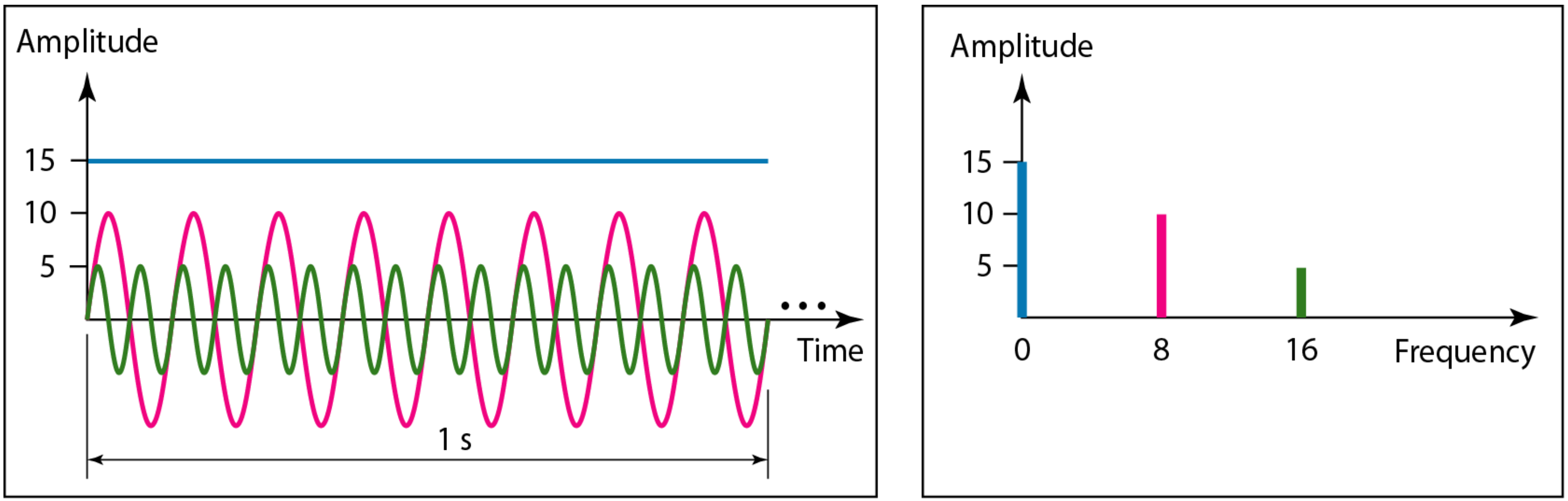
Note

A complete sine wave in the time domain can be represented by one single spike in the frequency domain.

Example 3.7

The frequency domain is more compact and useful when we are dealing with more than one sine wave. For example, Figure 3.8 shows three sine waves, each with different amplitude and frequency. All can be represented by three spikes in the frequency domain.

Figure 3.8 *The time domain and frequency domain of three sine waves*



a. Time-domain representation of three sine waves with frequencies 0, 8, and 16

b. Frequency-domain representation of the same three signals

Signals and Communication

- A single-frequency sine wave is not useful in data communications
- We need to send a composite signal, a signal made of many simple sine waves.
- According to Fourier analysis, any composite signal is a combination of simple sine waves with different frequencies, amplitudes, and phases.

Composite Signals and Periodicity

- If the composite signal is periodic, the decomposition gives a series of signals with discrete frequencies.
- If the composite signal is nonperiodic, the decomposition gives a combination of sine waves with continuous frequencies.

Example 3.4

Figure 3.9 shows a periodic composite signal with frequency f. This type of signal is not typical of those found in data communications. We can consider it to be three alarm systems, each with a different frequency. The analysis of this signal can give us a good understanding of how to decompose signals.

Figure 3.9 ***A composite periodic signal***

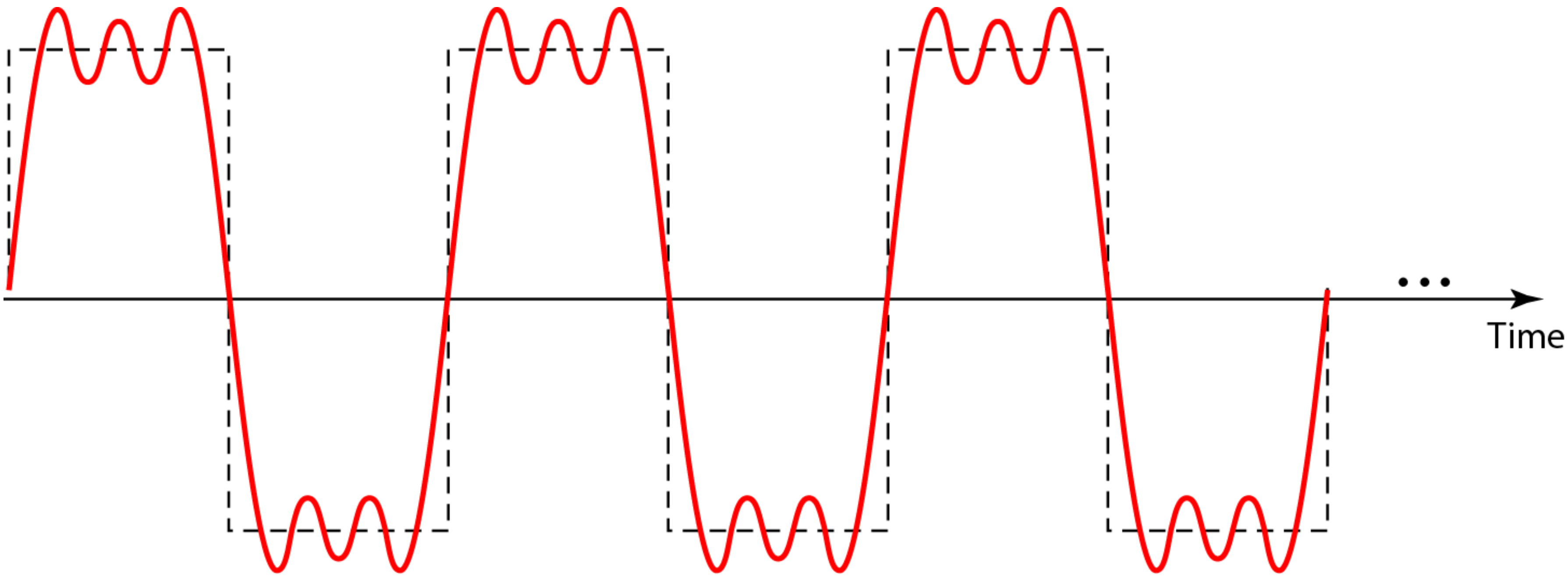
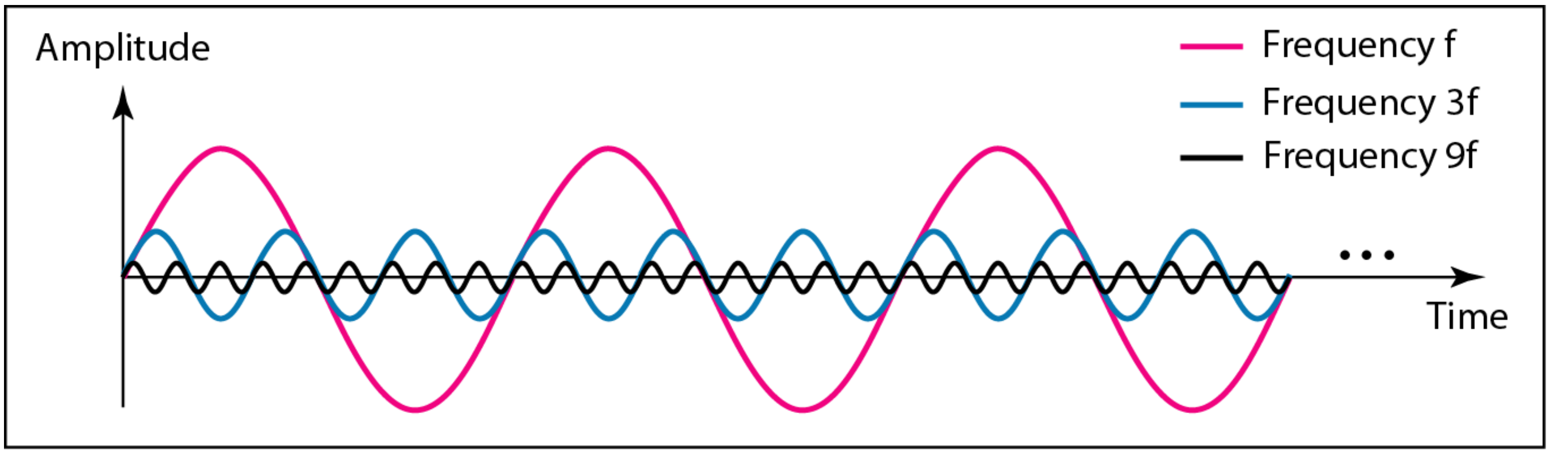
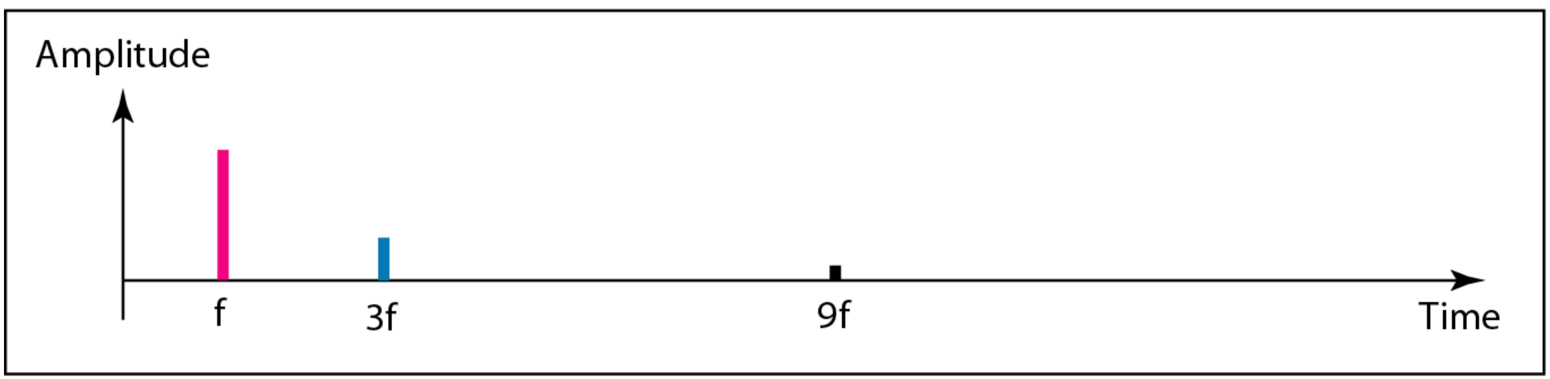


Figure 3.10 *Decomposition of a composite periodic signal in the time and frequency domains*



a. Time-domain decomposition of a composite signal

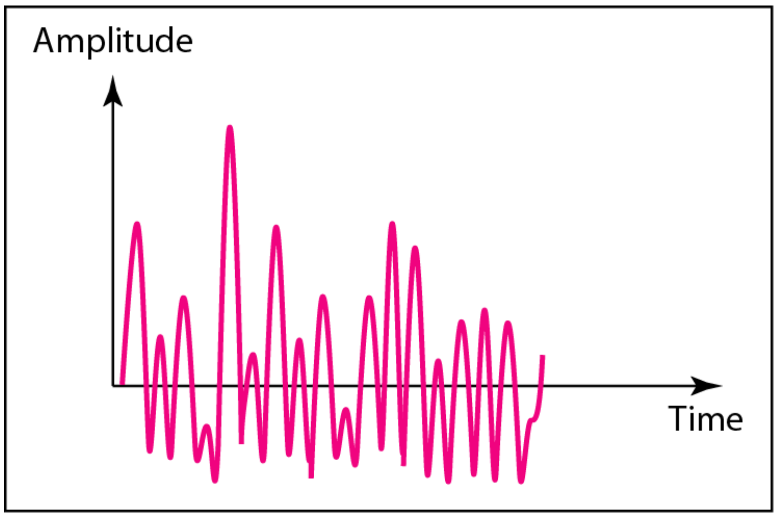


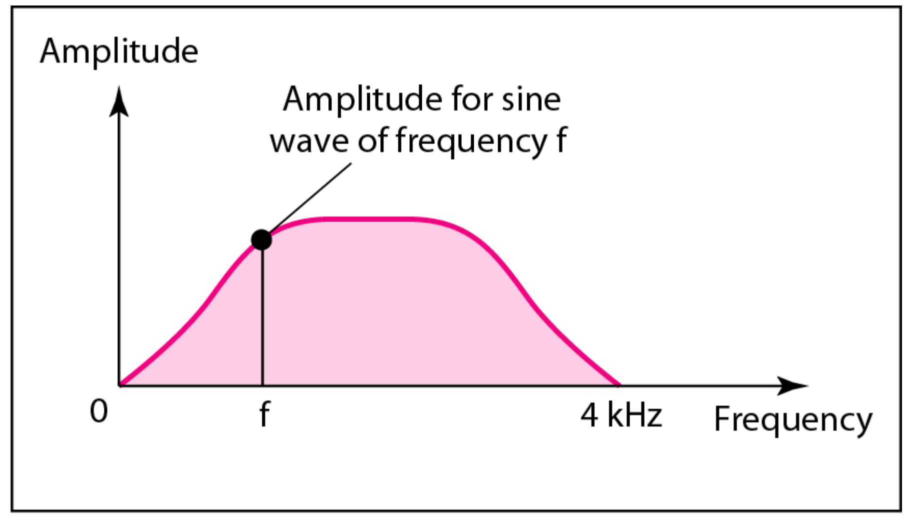
b. Frequency-domain decomposition of the composite signal

Example 3.5

Figure 3.11 shows a nonperiodic composite signal. It can be the signal created by a microphone or a telephone set when a word or two is pronounced. In this case, the composite signal cannot be periodic, because that implies that we are repeating the same word or words with exactly the same tone.

Figure 3.11 *The time and frequency domains of a nonperiodic signal*

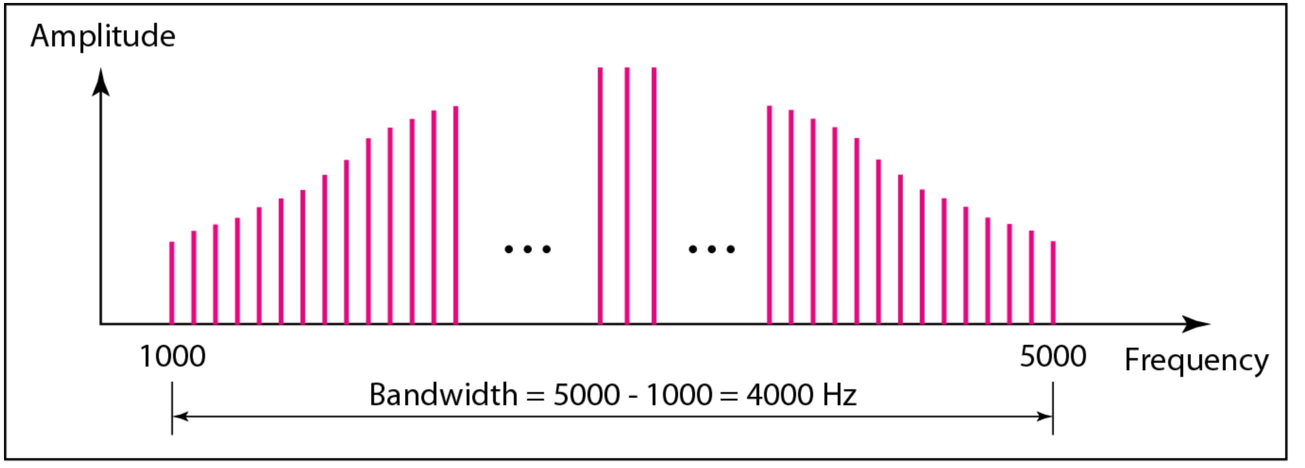




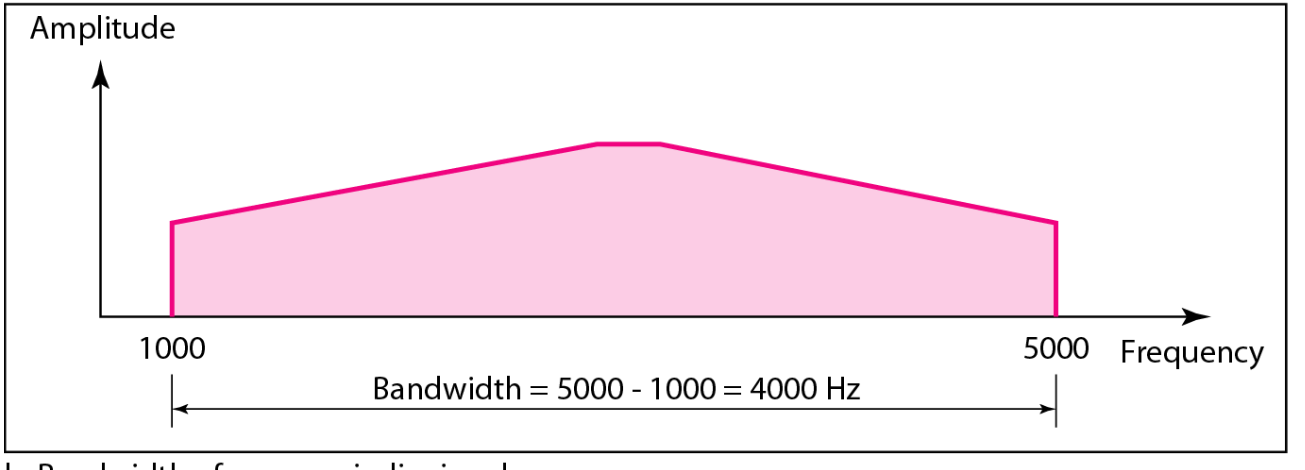
Bandwidth and Signal Frequency

- The bandwidth of a composite signal is the difference between the highest and the lowest frequencies contained in that signal.

Figure 3.12 *The bandwidth of periodic and nonperiodic composite signals*



a. Bandwidth of a periodic signal



b. Bandwidth of a nonperiodic signal

Example 3.6

If a periodic signal is decomposed into five sine waves with frequencies of 100, 300, 500, 700, and 900 Hz, what is its bandwidth? Draw the spectrum, assuming all components have a maximum amplitude of 10 V.

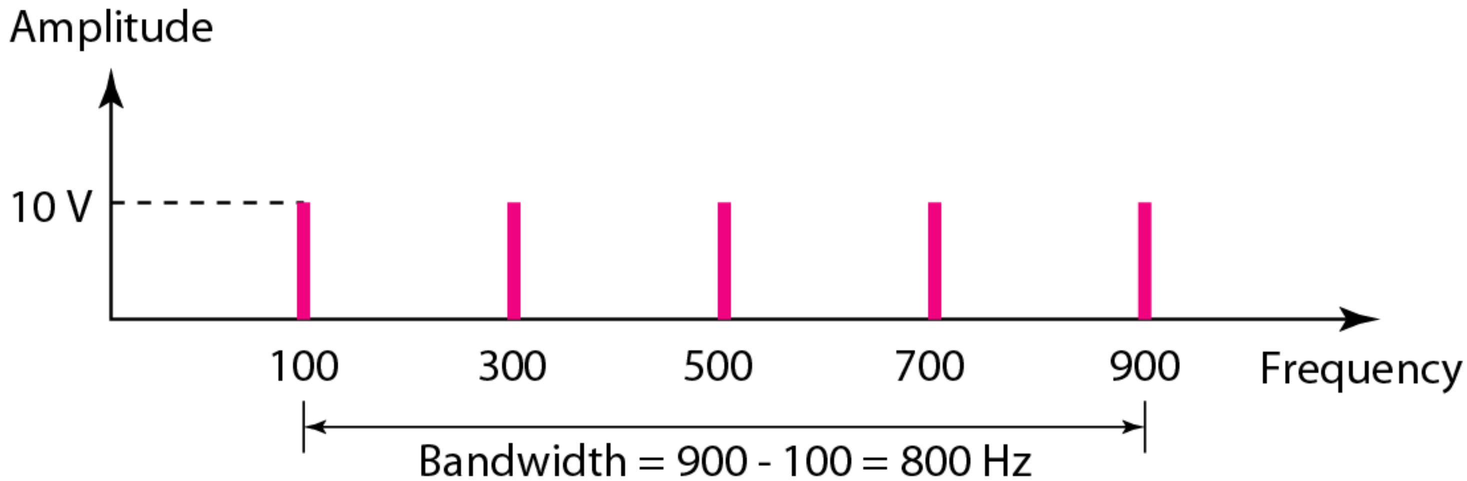
Solution

Let f_h be the highest frequency, f_l the lowest frequency, and B the bandwidth. Then

$$\boxed{B = f_h - f_l = 900 - 100 = 800 \text{ Hz}}$$

The spectrum has only five spikes, at 100, 300, 500, 700, and 900 Hz (see Figure 3.13).

Figure 3.13 *The bandwidth for Example 3.6*



Example 3.7

A periodic signal has a bandwidth of 20 Hz. The highest frequency is 60 Hz. What is the lowest frequency? Draw the spectrum if the signal contains all frequencies of the same amplitude.

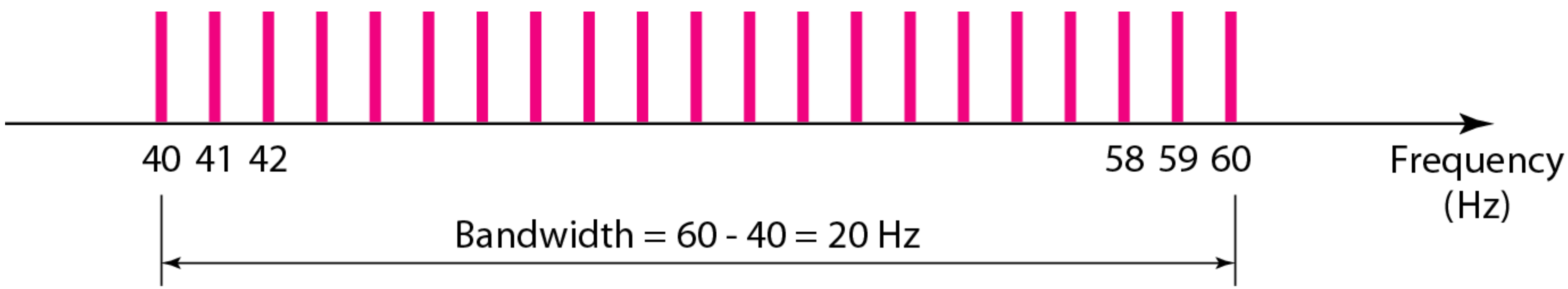
Solution

Let f_h be the highest frequency, f_l the lowest frequency, and B the bandwidth. Then

$$B = f_h - f_l \quad \Rightarrow \quad 20 = 60 - f_l \quad \Rightarrow \quad f_l = 60 - 20 = 40 \text{ Hz}$$

The spectrum contains all integer frequencies. We show this by a series of spikes (see Figure 3.14).

Figure 3.14 *The bandwidth for Example 3.7*



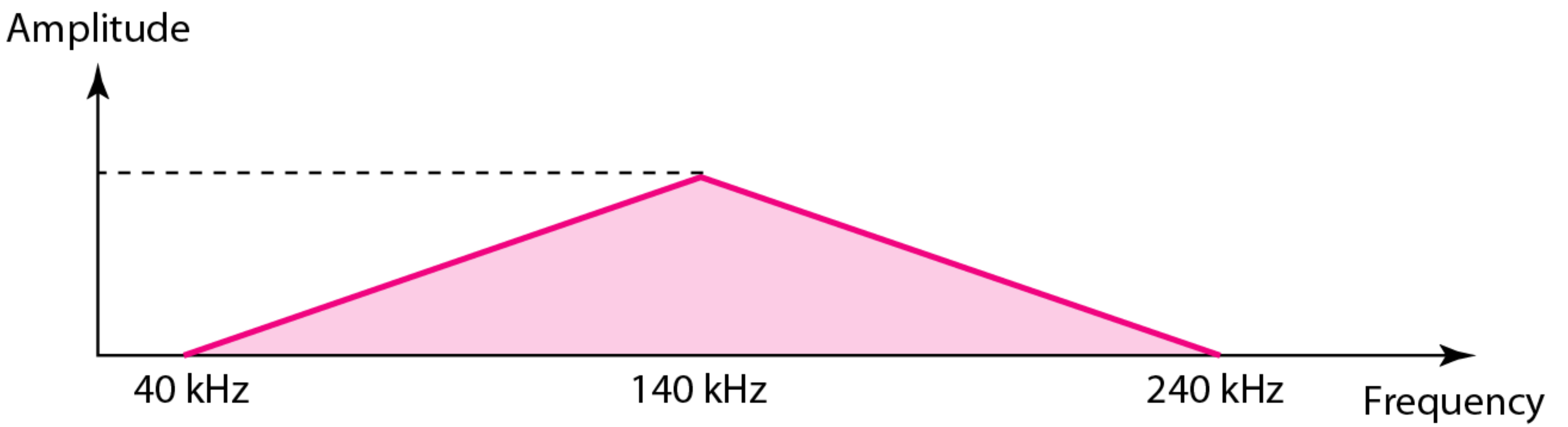
Example 3.8

A nonperiodic composite signal has a bandwidth of 200 kHz, with a middle frequency of 140 kHz and peak amplitude of 20 V. The two extreme frequencies have an amplitude of 0. Draw the frequency domain of the signal.

Solution

The lowest frequency must be at 40 kHz and the highest at 240 kHz. Figure 3.15 shows the frequency domain and the bandwidth.

Figure 3.15 *The bandwidth for Example 3.8*



Example 3.9

An example of a nonperiodic composite signal is the signal propagated by an AM radio station. In the United States, each AM radio station is assigned a 10-kHz bandwidth. The total bandwidth dedicated to AM radio ranges from 530 to 1700 kHz. We will show the rationale behind this 10-kHz bandwidth in Chapter 5.

Example 3.10

Another example of a nonperiodic composite signal is the signal propagated by an FM radio station. In the United States, each FM radio station is assigned a 200-kHz bandwidth. The total bandwidth dedicated to FM radio ranges from 88 to 108 MHz. We will show the rationale behind this 200-kHz bandwidth in Chapter 5.

Example 3.11

Another example of a nonperiodic composite signal is the signal received by an old-fashioned analog black-and-white TV. A TV screen is made up of pixels. If we assume a resolution of 525 × 700, we have 367,500 pixels per screen. If we scan the screen 30 times per second, this is 367,500 × 30 = 11,025,000 pixels per second. The worst-case scenario is alternating black and white pixels. We can send 2 pixels per cycle. Therefore, we need 11,025,000 / 2 = 5,512,500 cycles per second, or Hz. The bandwidth needed is 5.5125 MHz.

Fourier Analysis

Note

Fourier analysis is a tool that changes a time domain signal to a frequency domain signal and vice versa.

Fourier Series

- Every composite periodic signal can be represented with a series of sine and cosine functions.
- The functions are integral harmonics of the fundamental frequency “f” of the composite signal.
- Using the series we can decompose any periodic signal into its harmonics.

Fourier Series

Fourier series

$$s(t) = A_0 + \sum_{n=1}^{\infty} A_n \sin(2\pi n f t) + \sum_{n=1}^{\infty} B_n \cos(2\pi n f t)$$

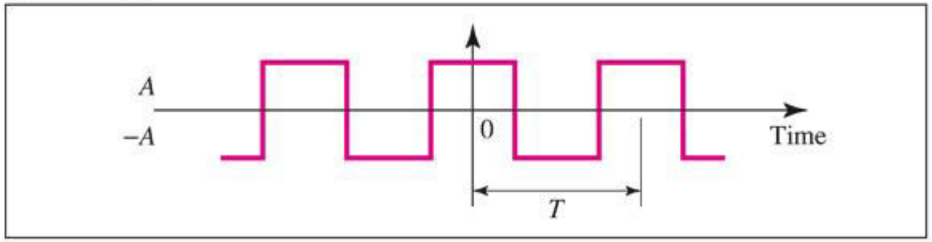
$$A_0 = \frac{1}{T}\int_0^T s(t)\, dt \qquad A_n = \frac{2}{T}\int_0^T s(t) \cos(2\pi n f t)\, dt$$

$$B_n = \frac{2}{T}\int_0^T s(t) \sin(2\pi n f t)\, dt$$

Coefficients

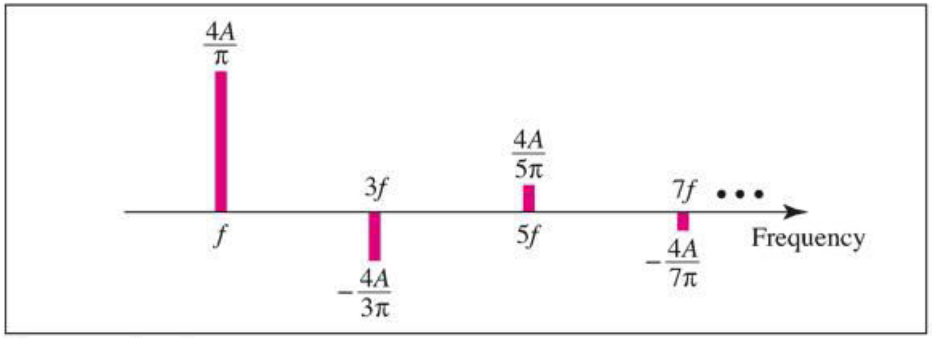
Examples of Signals and the Fourier Series Representation

Time domain



$$A_0 = 0 \qquad A_n = \begin{cases} \dfrac{4A}{n\pi} & \text{for } n = 1, 5, 9, \ldots \\ \dfrac{-4A}{n\pi} & \text{for } n = 3, 7, 11, \ldots \end{cases} \qquad B_n = 0$$

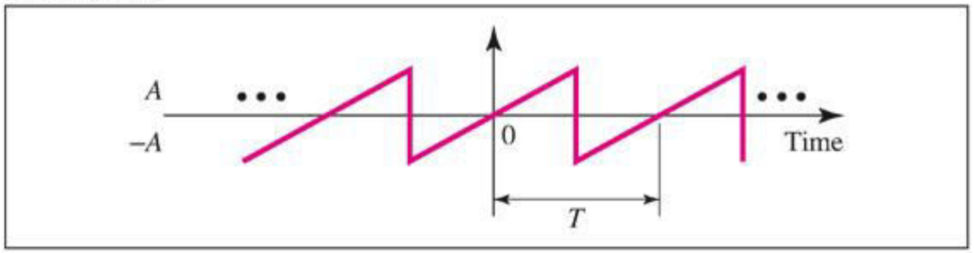
$$s(t) = \frac{4A}{\pi}\cos(2\pi f t) - \frac{4A}{3\pi}\cos(2\pi 3 f t) + \frac{4A}{5\pi}\cos(2\pi 5 f t) - \frac{4A}{7\pi}\cos(2\pi 7 f t) + \cdots$$



Frequency domain

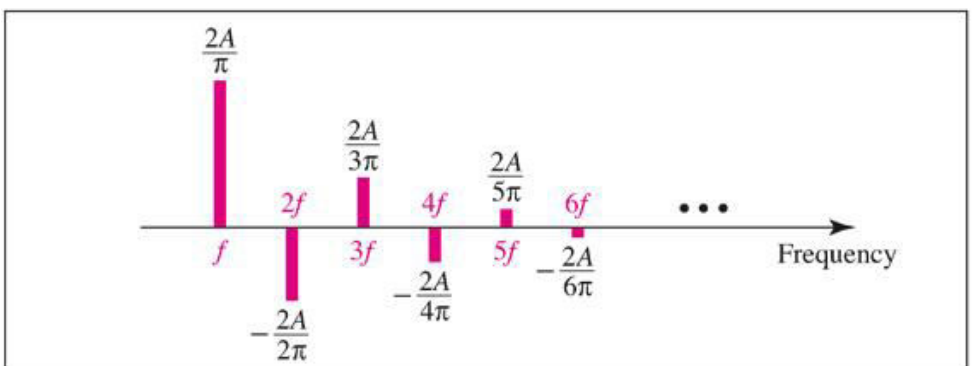
Sawtooth Signal

Time domain



$A_0 = 0 \qquad A_n = 0 \qquad B_n = \begin{cases} \dfrac{2A}{n\pi} & \text{for } n \text{ odd} \\ -\dfrac{2A}{n\pi} & \text{for } n \text{ even} \end{cases}$

$$s(t) = \frac{2A}{\pi}\sin(2\pi ft) - \frac{2A}{2\pi}\sin(2\pi 2ft) + \frac{2A}{3\pi}\sin(2\pi 3ft) - \frac{2A}{4\pi}\sin(2\pi 4ft) + \cdots$$

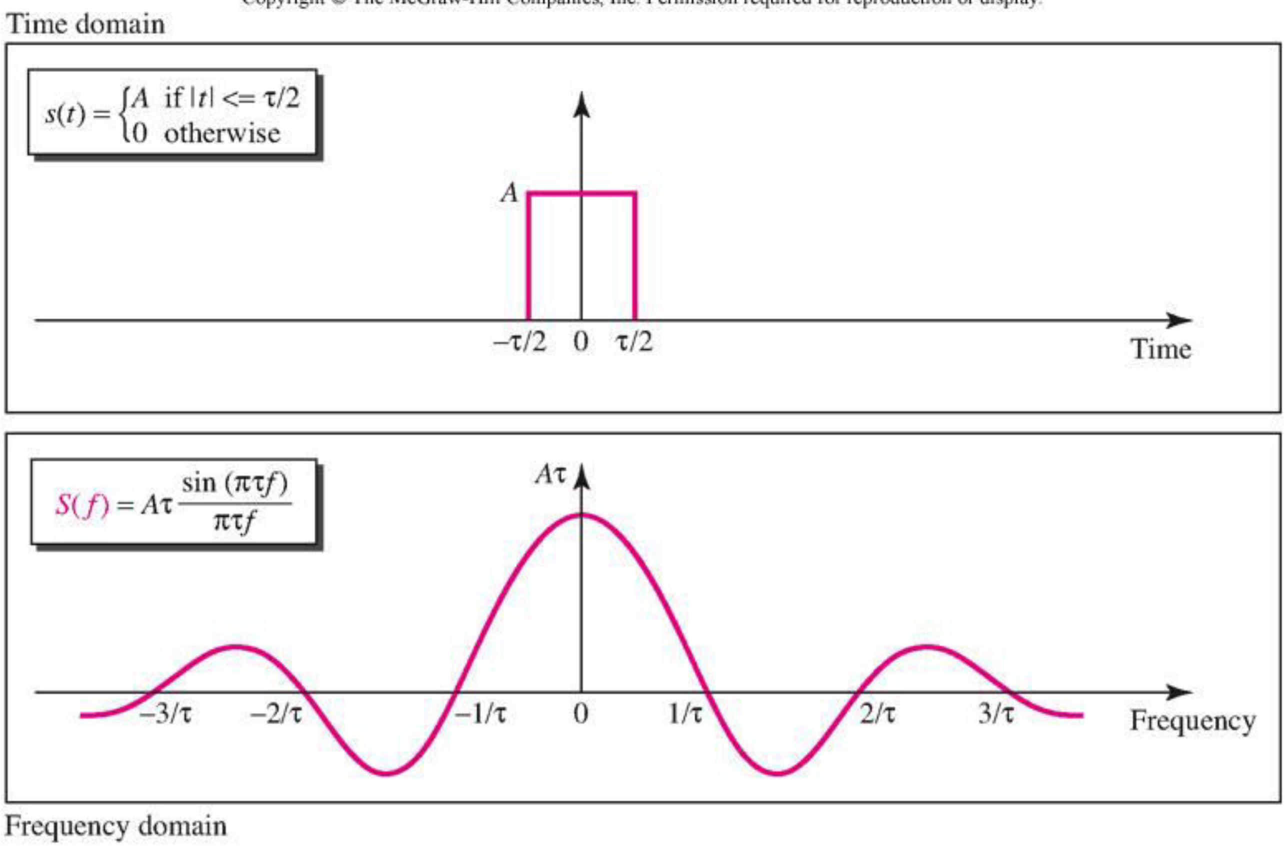


Frequency domain

Fourier Transform

- Fourier Transform gives the frequency domain of a nonperiodic time domain signal.

Example of a Fourier Transform



Inverse Fourier Transform

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft}\, dt$$

Fourier transform

$$s(t) = \int_{-\infty}^{\infty} S(f) e^{j2\pi ft}\, dt$$

Inverse Fourier transform

Time limited and Band limited Signals

- A time limited signal is a signal for which the amplitude $s(t) = 0$ for $t > T_1$ and $t < T_2$
- A band limited signal is a signal for which the amplitude $S(f) = 0$ for $f > F_1$ and $f < F_2$

Chapter 3

Data and Signals

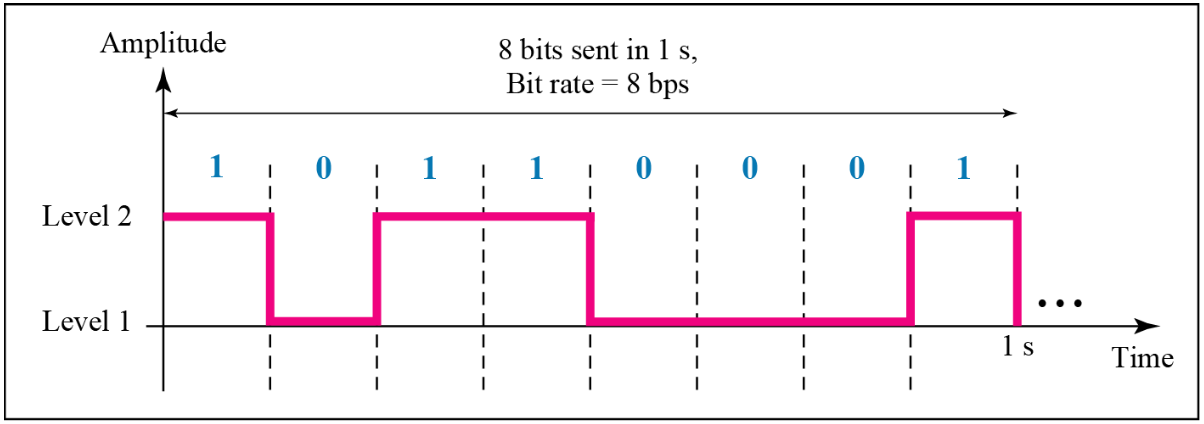
3-3 DIGITAL SIGNALS

In addition to being represented by an analog signal, information can also be represented by a digital signal. For example, a 1 can be encoded as a positive voltage and a 0 as zero voltage. A digital signal can have more than two levels. In this case, we can send more than 1 bit for each level.

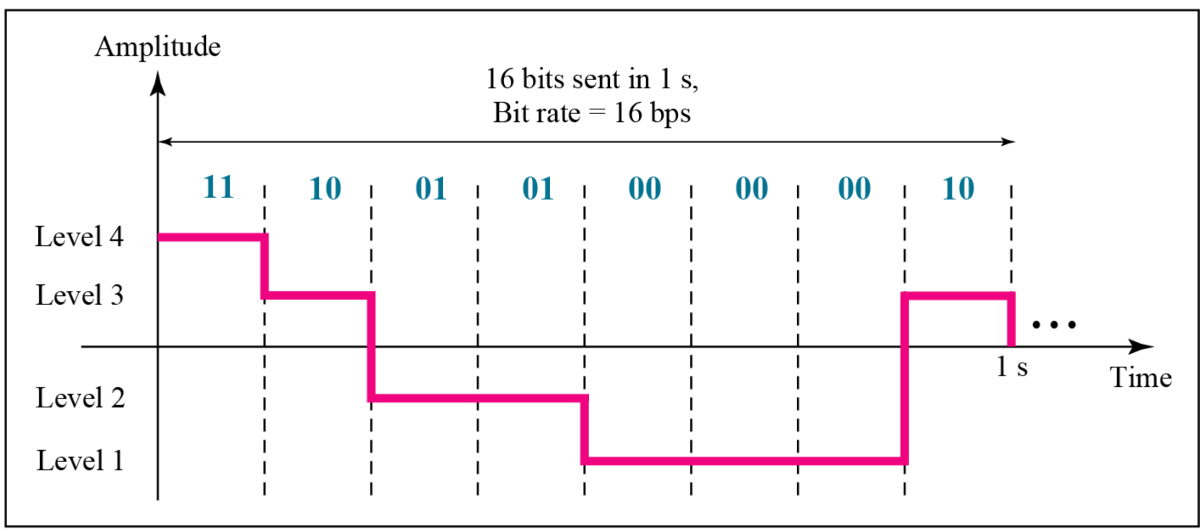
Topics discussed in this section:

- **Bit Rate**
- **Bit Length**
- **Digital Signal as a Composite Analog Signal**
- **Application Layer**

Figure 3.16 *Two digital signals: one with two signal levels and the other with four signal levels*



a. A digital signal with two levels



b. A digital signal with four levels

Example 3.16

A digital signal has eight levels. How many bits are needed per level? We calculate the number of bits from the formula

$$\boxed{\text{Number of bits per level} = \log_2 8 = 3}$$

Each signal level is represented by 3 bits.

Example 3.17

A digital signal has nine levels. How many bits are needed per level? We calculate the number of bits by using the formula. Each signal level is represented by 3.17 bits. However, this answer is not realistic. The number of bits sent per level needs to be an integer as well as a power of 2. For this example, 4 bits can represent one level.

Example 3.18

Assume we need to download text documents at the rate of 100 pages per sec. What is the required bit rate of the channel?

Solution

A page is an average of 24 lines with 80 characters in each line. If we assume that one character requires 8 bits (ascii), the bit rate is

$$\boxed{100 \times 24 \times 80 \times 8 = 1{,}636{,}000 \text{ bps} = 1.636 \text{ Mbps}}$$

Example 3.19

A digitized voice channel, as we will see in Chapter 4, is made by digitizing a 4-kHz bandwidth analog voice signal. We need to sample the signal at twice the highest frequency (two samples per hertz). We assume that each sample requires 8 bits. What is the required bit rate?

Solution

The bit rate can be calculated as

$$\boxed{2 \times 4000 \times 8 = 64{,}000 \text{ bps} = 64 \text{ kbps}}$$

Example 3.20

What is the bit rate for high-definition TV (HDTV)?

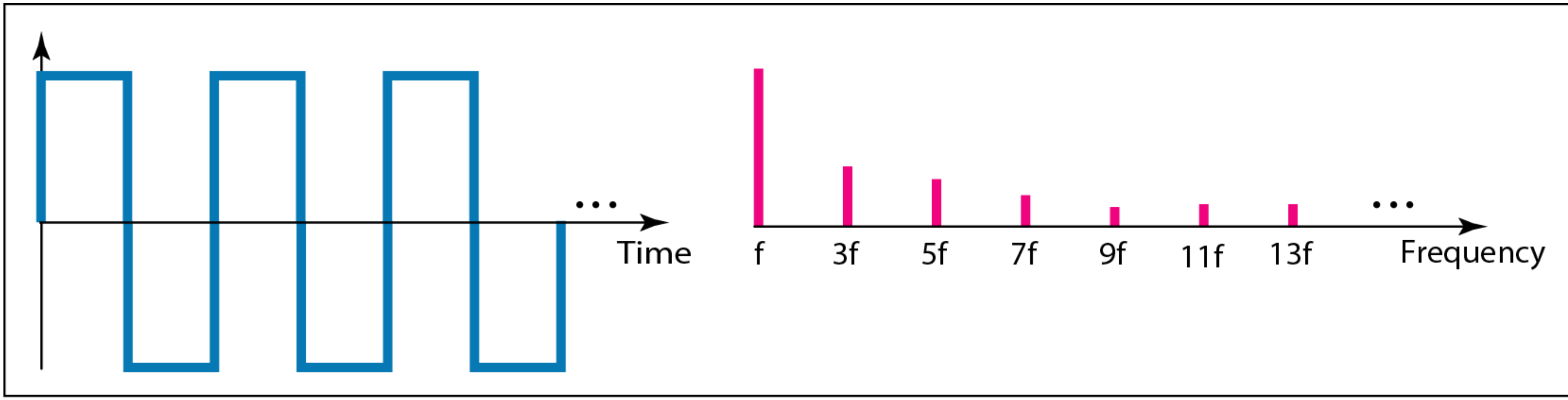
Solution

HDTV uses digital signals to broadcast high quality video signals. The HDTV screen is normally a ratio of 16 : 9. There are 1920 by 1080 pixels per screen, and the screen is renewed 30 times per second. Twenty-four bits represents one color pixel.

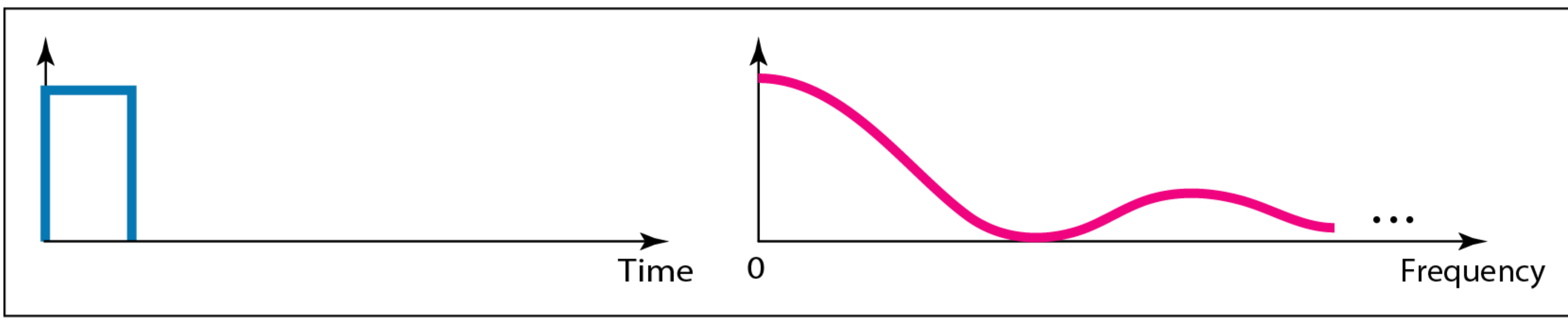
$$1920 \times 1080 \times 30 \times 24 = 1{,}492{,}992{,}000 \text{ or } 1.5 \text{ Gbps}$$

The TV stations reduce this rate to 20 to 40 Mbps through compression.

Figure 3.17 *The time and frequency domains of periodic and nonperiodic digital signals*

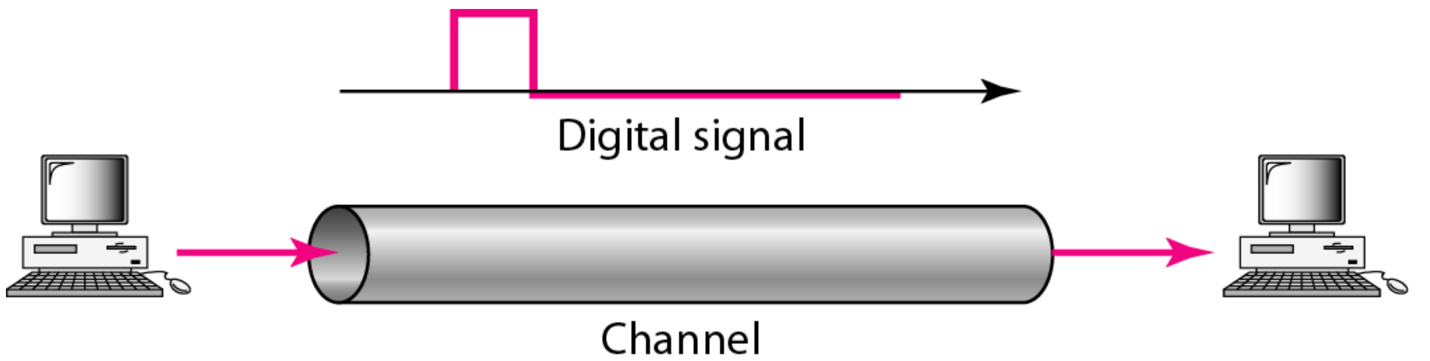


a. Time and frequency domains of periodic digital signal



b. Time and frequency domains of nonperiodic digital signal

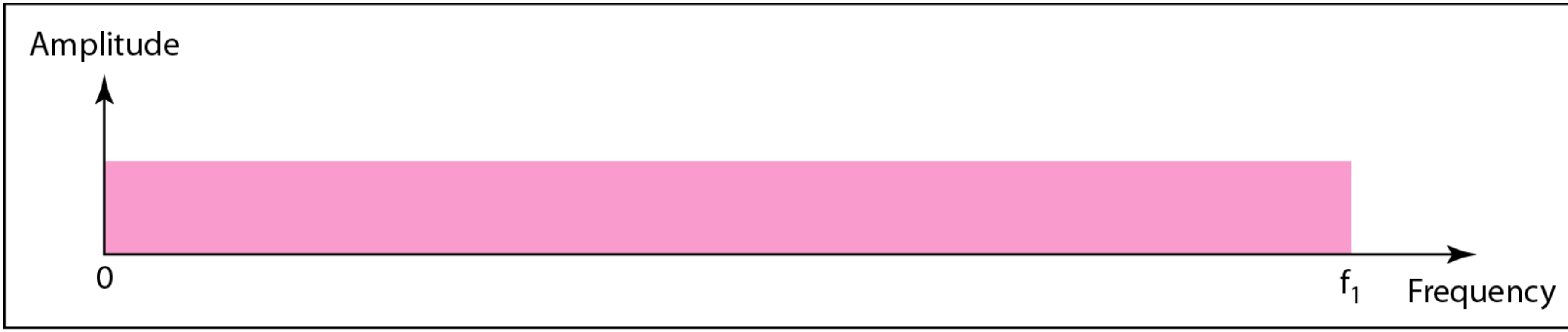
Figure 3.18 *Baseband transmission*



Note

A digital signal is a composite analog signal with an infinite bandwidth.

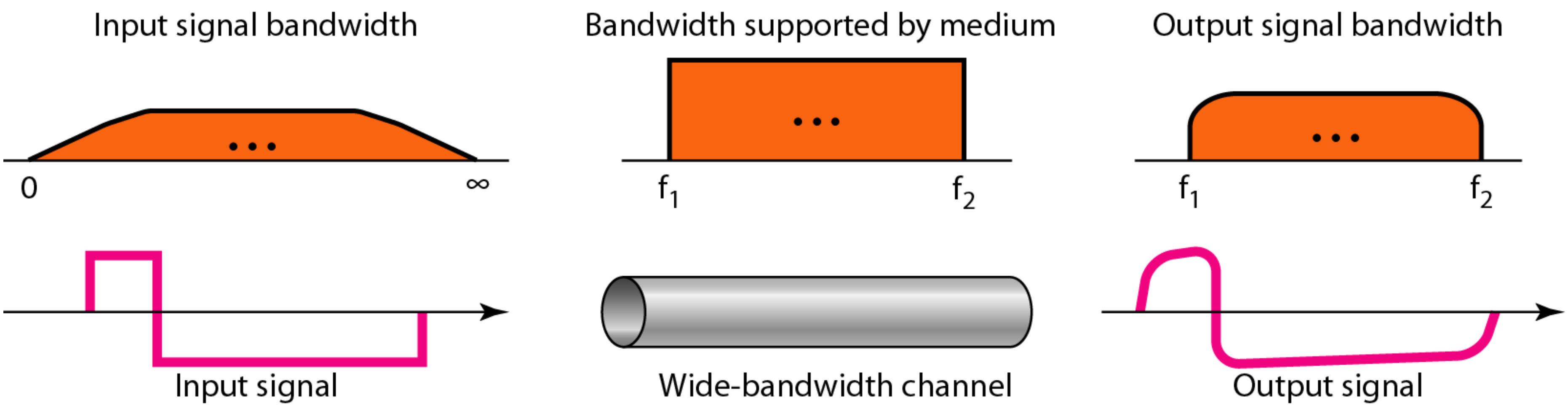
Figure 3.19 *Bandwidths of two low-pass channels*



a. Low-pass channel, wide bandwidth

b. Low-pass channel, narrow bandwidth

Figure 3.20 *Baseband transmission using a dedicated medium*



Note

Baseband transmission of a digital signal that preserves the shape of the digital signal is possible only if we have a low-pass channel with an infinite or very wide bandwidth.

Example 3.21

An example of a dedicated channel where the entire bandwidth of the medium is used as one single channel is a LAN. Almost every wired LAN today uses a dedicated channel for two stations communicating with each other. In a bus topology LAN with multipoint connections, only two stations can communicate with each other at each moment in time (timesharing); the other stations need to refrain from sending data. In a star topology LAN, the entire channel between each station and the hub is used for communication between these two entities.

Figure 3.21 *Rough approximation of a digital signal using the first harmonic for worst case*

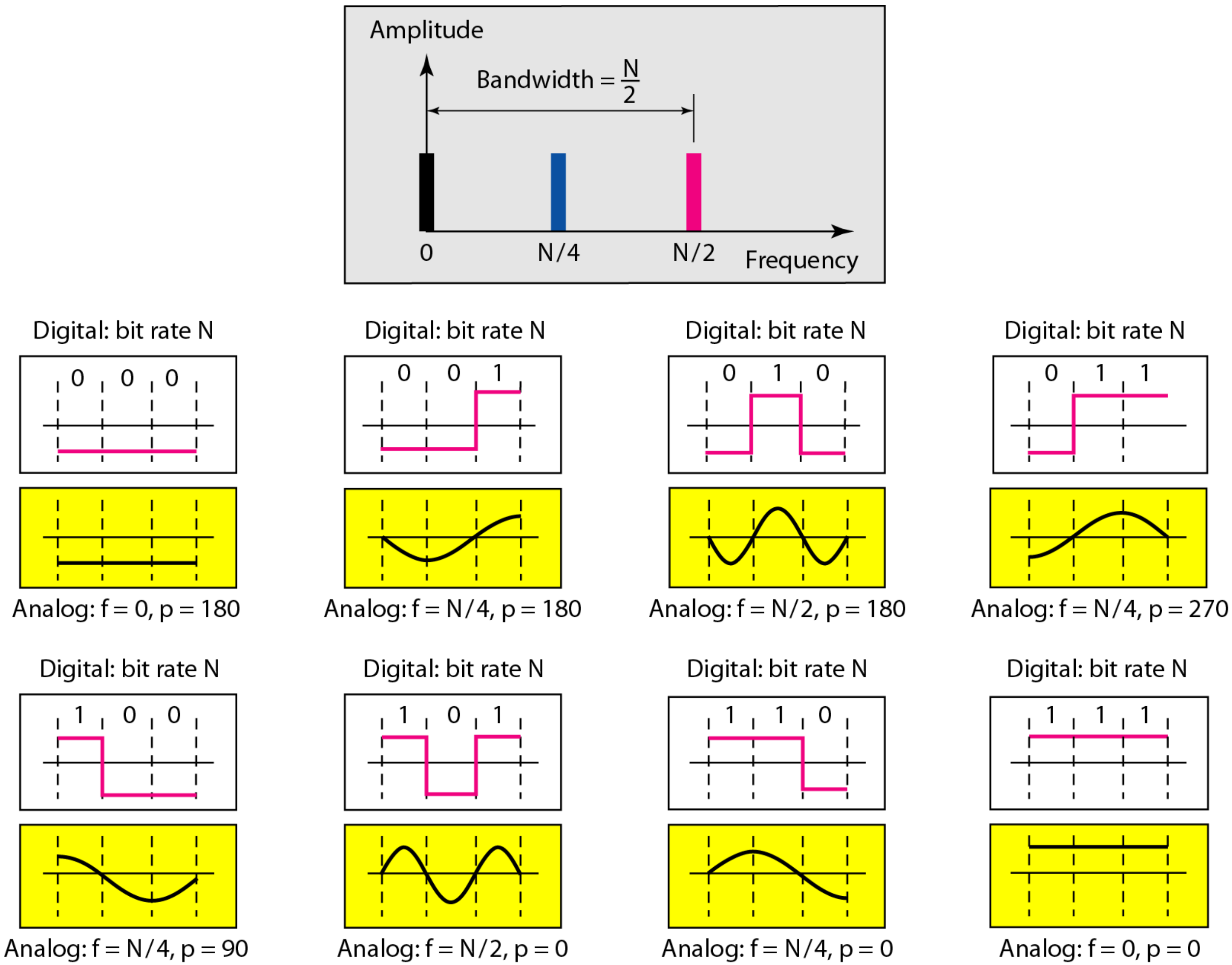
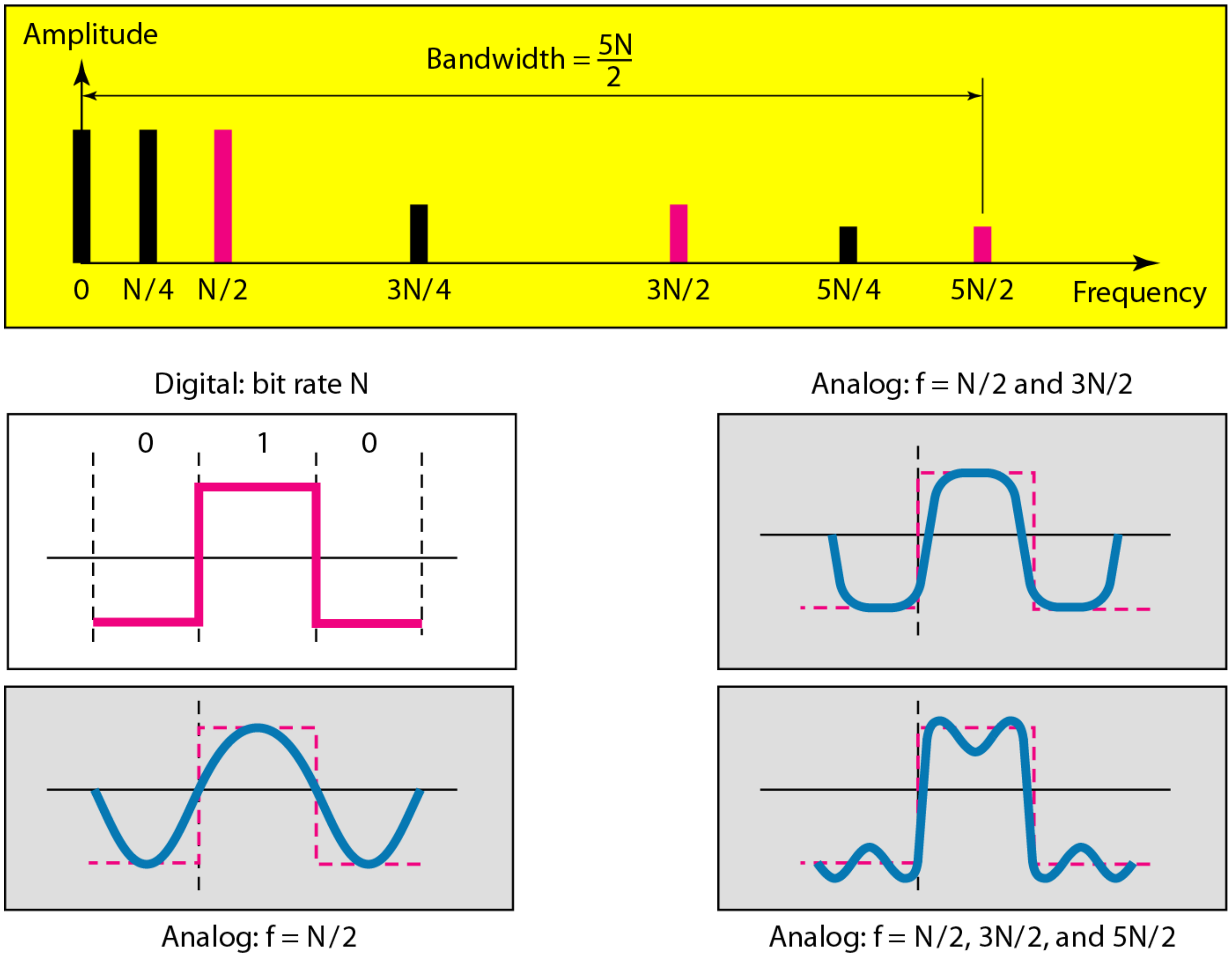


Figure 3.22 *Simulating a digital signal with first three harmonics*



Note

In baseband transmission, the required bandwidth is proportional to the bit rate; if we need to send bits faster, we need more bandwidth.

Table 3.2 ***Bandwidth requirements***

Bit Rate	*Harmonic 1*	*Harmonics 1, 3*	*Harmonics 1, 3, 5*
$n = 1$ kbps	$B = 500$ Hz	$B = 1.5$ kHz	$B = 2.5$ kHz
$n = 10$ kbps	$B = 5$ kHz	$B = 15$ kHz	$B = 25$ kHz
$n = 100$ kbps	$B = 50$ kHz	$B = 150$ kHz	$B = 250$ kHz

Example 3.22

What is the required bandwidth of a low-pass channel if we need to send 1 Mbps by using baseband transmission?

Solution

The answer depends on the accuracy desired.

a. The minimum bandwidth, is B = bit rate /2, or 500 kHz.

b. A better solution is to use the first and the third harmonics with $B = 3 \times 500$ kHz = 1.5 MHz.

c. Still a better solution is to use the first, third, and fifth harmonics with $B = 5 \times 500$ kHz = 2.5 MHz.

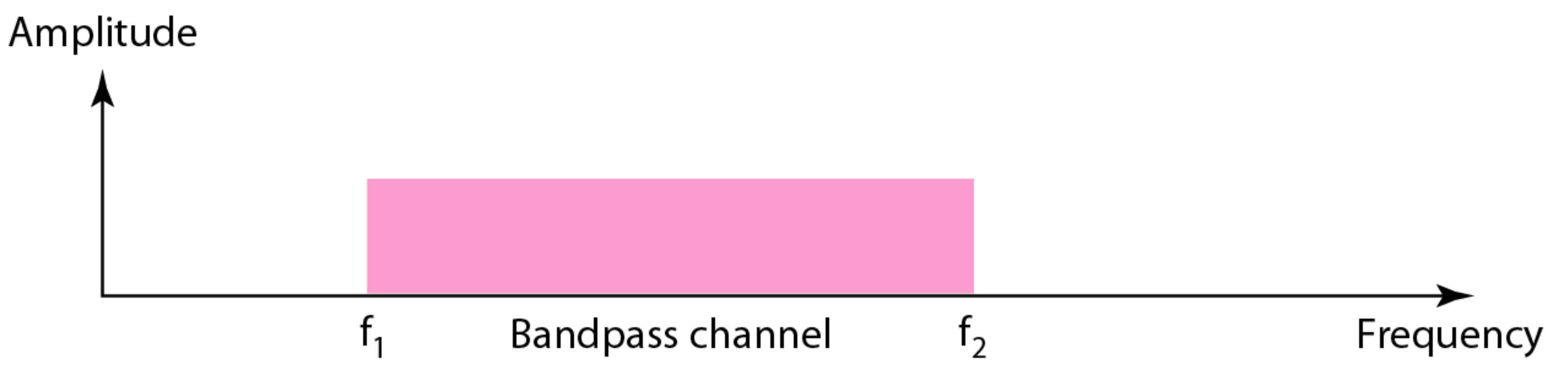
Example 3.22

We have a low-pass channel with bandwidth 100 kHz. What is the maximum bit rate of this channel?

Solution

The maximum bit rate can be achieved if we use the first harmonic. The bit rate is 2 times the available bandwidth, or 200 kbps.

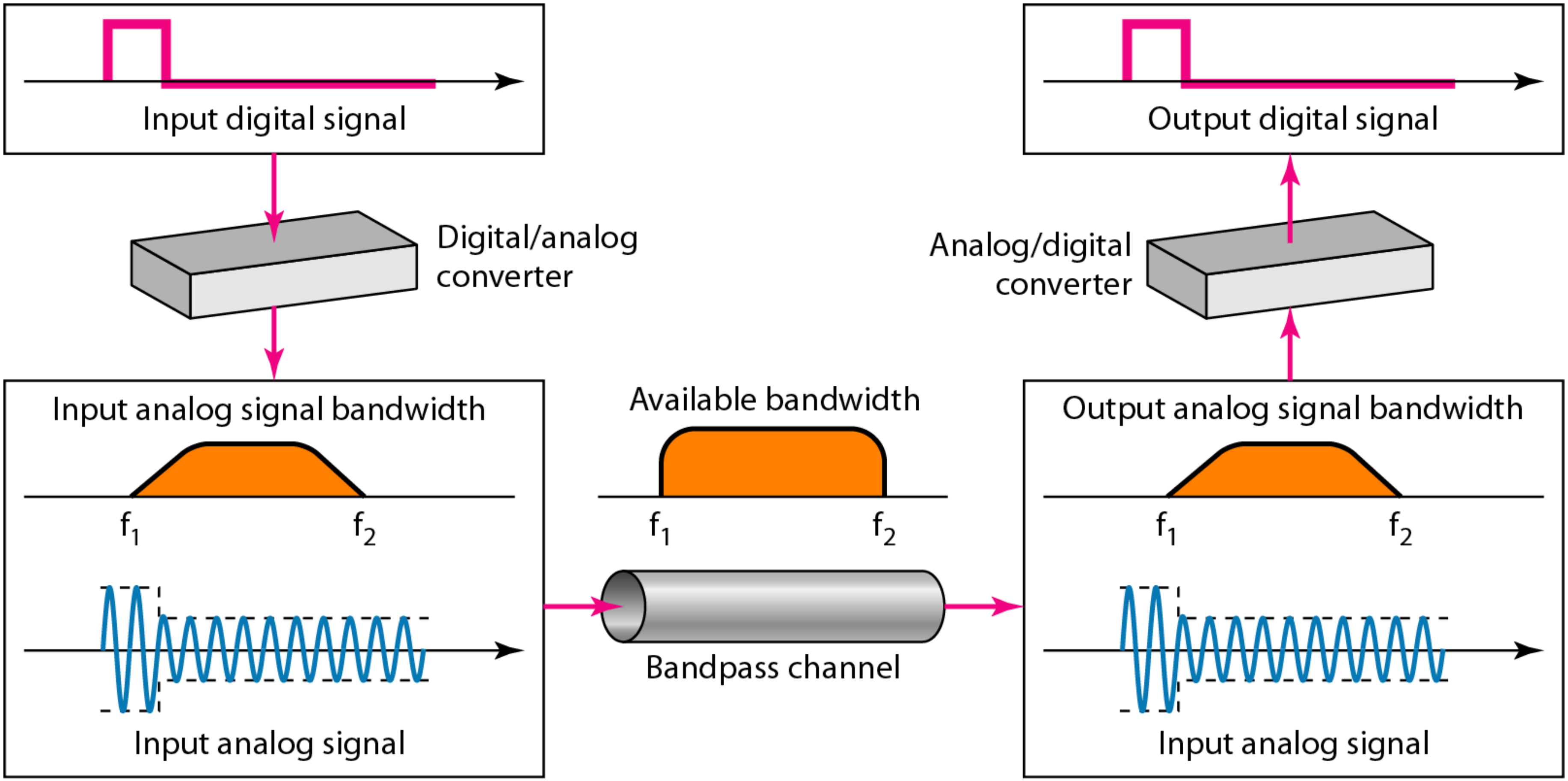
Figure 3.23 ***Bandwidth of a bandpass channel***



Note

If the available channel is a bandpass channel, we cannot send the digital signal directly to the channel; we need to convert the digital signal to an analog signal before transmission.

Figure 3.24 *Modulation of a digital signal for transmission on a bandpass channel*



Example 3.24

An example of broadband transmission using modulation is the sending of computer data through a telephone subscriber line, the line connecting a resident to the central telephone office. These lines are designed to carry voice with a limited bandwidth. The channel is considered a bandpass channel. We convert the digital signal from the computer to an analog signal, and send the analog signal. We can install two converters to change the digital signal to analog and vice versa at the receiving end. The converter, in this case, is called a modem which we discuss in detail in Chapter 5.

Example 3.25

A second example is the digital cellular telephone. For better reception, digital cellular phones convert the analog voice signal to a digital signal (see Chapter 16). Although the bandwidth allocated to a company providing digital cellular phone service is very wide, we still cannot send the digital signal without conversion. The reason is that we only have a bandpass channel available between caller and callee. We need to convert the digitized voice to a composite analog signal before sending.

Chapter 3

Data and Signals

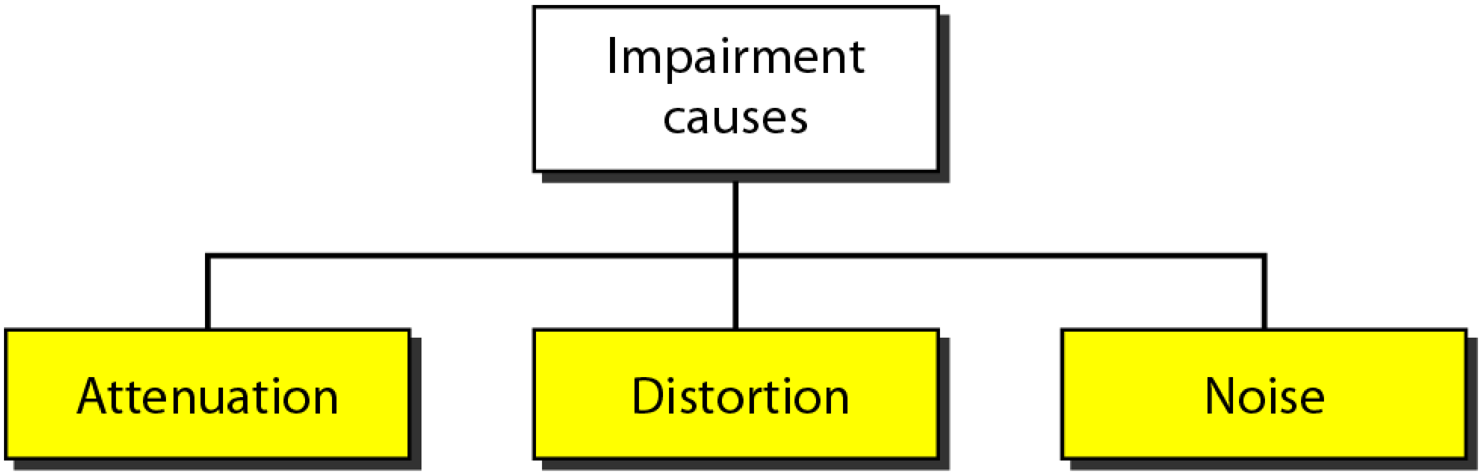
3-4 TRANSMISSION IMPAIRMENT

Signals travel through transmission media, which are not perfect. The imperfection causes signal impairment. This means that the signal at the beginning of the medium is not the same as the signal at the end of the medium. What is sent is not what is received. Three causes of impairment are attenuation, distortion, and noise.

Topics discussed in this section:

- **Attenuation**
- **Distortion**
- **Noise**

Figure 3.25 *Causes of impairment*



3.4

Attenuation

- Means loss of energy -> weaker signal
- When a signal travels through a medium it loses energy overcoming the resistance of the medium
- Amplifiers are used to compensate for this loss of energy by amplifying the signal.

Measurement of Attenuation

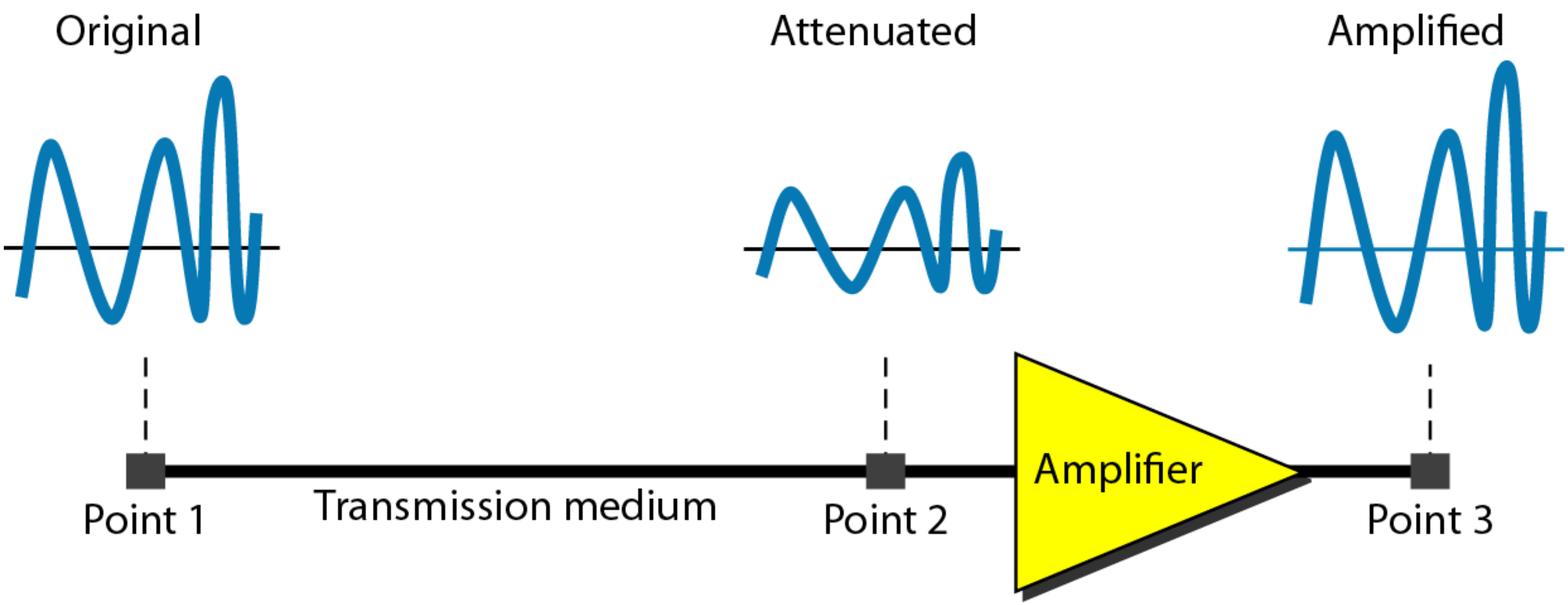
- To show the loss or gain of energy the unit "decibel" is used.

$$dB = 10\log_{10} P_2/P_1$$

P_1 - input signal

P_2 - output signal

Figure 3.26 *Attenuation*



Example 3.26

Suppose a signal travels through a transmission medium and its power is reduced to one-half. This means that P_2 is $(1/2)P_1$. In this case, the attenuation (loss of power) can be calculated as

$$10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{0.5 P_1}{P_1} = 10 \log_{10} 0.5 = 10(-0.3) = -3 \text{ dB}$$

A loss of 3 dB (–3 dB) is equivalent to losing one-half the power.

Example 3.27

A signal travels through an amplifier, and its power is increased 10 times. This means that $P_2 = 10P_1$. In this case, the amplification (gain of power) can be calculated as

$$10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{10P_1}{P_1}$$

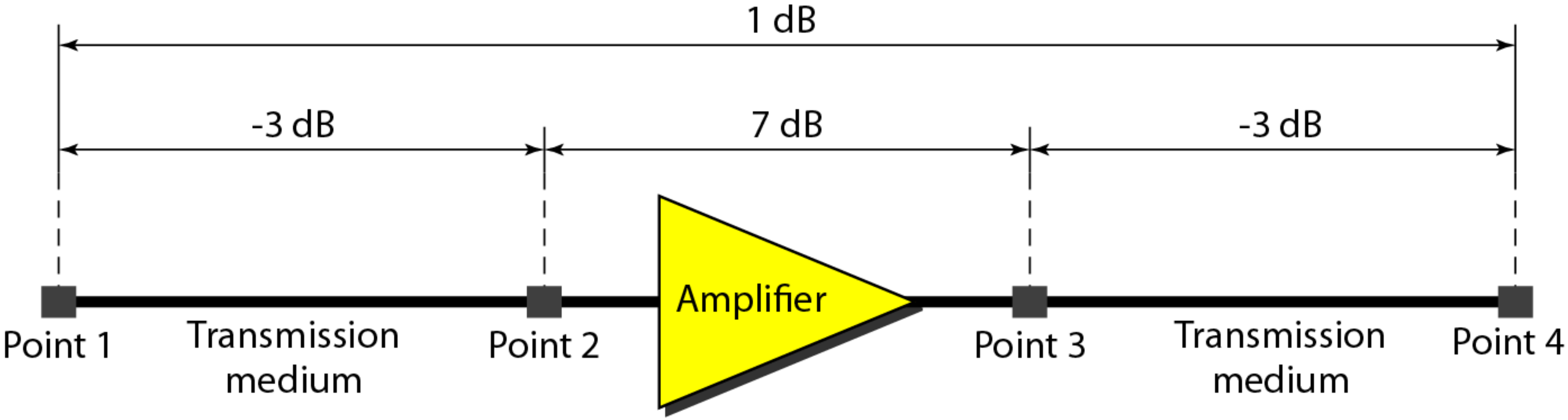
$$= 10 \log_{10} 10 = 10(1) = 10 \text{ dB}$$

Example 3.28

One reason that engineers use the decibel to measure the changes in the strength of a signal is that decibel numbers can be added (or subtracted) when we are measuring several points (cascading) instead of just two. In Figure 3.27 a signal travels from point 1 to point 4. In this case, the decibel value can be calculated as

$$\boxed{\text{dB} = -3 + 7 - 3 = +1}$$

Figure 3.27 *Decibels for Example 3.28*



Example 3.29

Sometimes the decibel is used to measure signal power in milliwatts. In this case, it is referred to as dB_m and is calculated as $dB_m = 10 \log10 P_m$, where P_m is the power in milliwatts. Calculate the power of a signal with $dB_m = -30$.

Solution

We can calculate the power in the signal as

$$\mathrm{dB_m} = 10 \log_{10} P_m = -30$$

$$\log_{10} P_m = -3 \qquad P_m = 10^{-3}\ \mathrm{mW}$$

Example 3.30

The loss in a cable is usually defined in decibels per kilometer (dB/km). If the signal at the beginning of a cable with −0.3 dB/km has a power of 2 mW, what is the power of the signal at 5 km?

Solution

The loss in the cable in decibels is 5 × (−0.3) = −1.5 dB. We can calculate the power as

$$\text{dB} = 10 \log_{10} \frac{P_2}{P_1} = -1.5$$

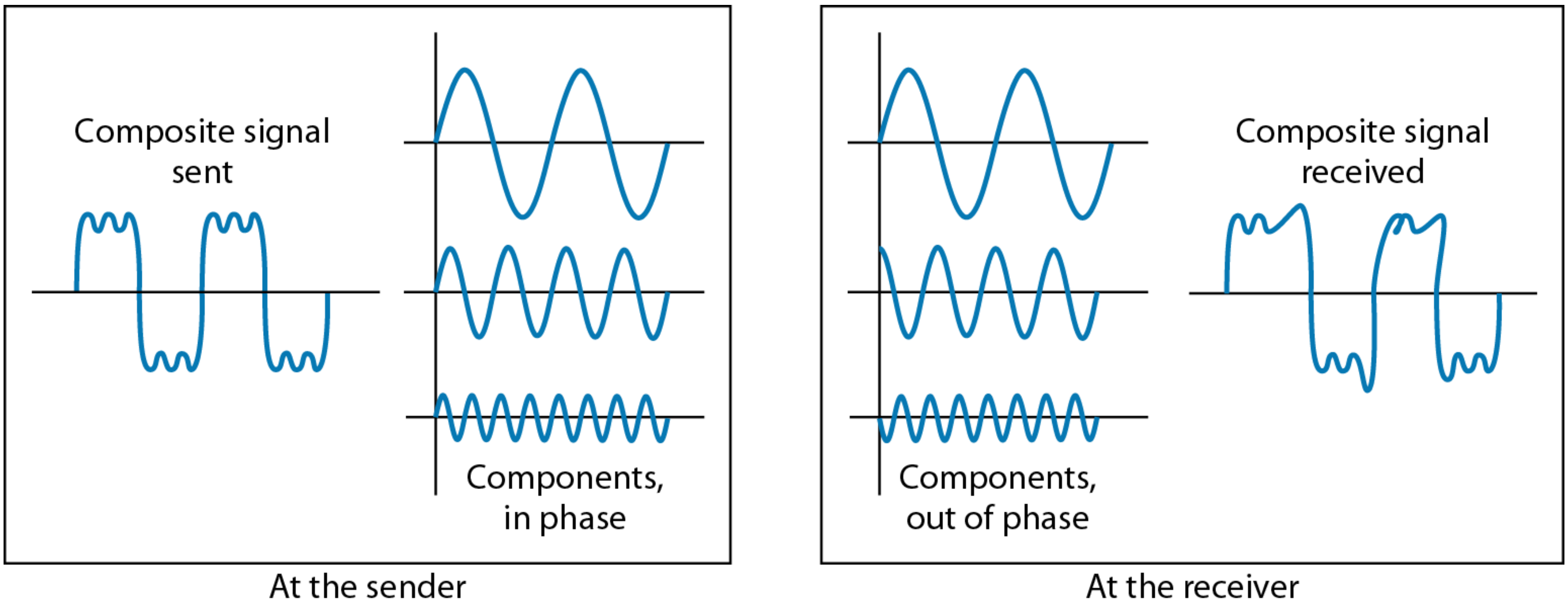
$$\frac{P_2}{P_1} = 10^{-0.15} = 0.71$$

$$P_2 = 0.71 P_1 = 0.7 \times 2 = 1.4 \text{ mW}$$

Distortion

- Means that the signal changes its form or shape
- Distortion occurs in composite signals
- Each frequency component has its own propagation speed traveling through a medium.
- The different components therefore arrive with different delays at the receiver.
- That means that the signals have different phases at the receiver than they did at the source.

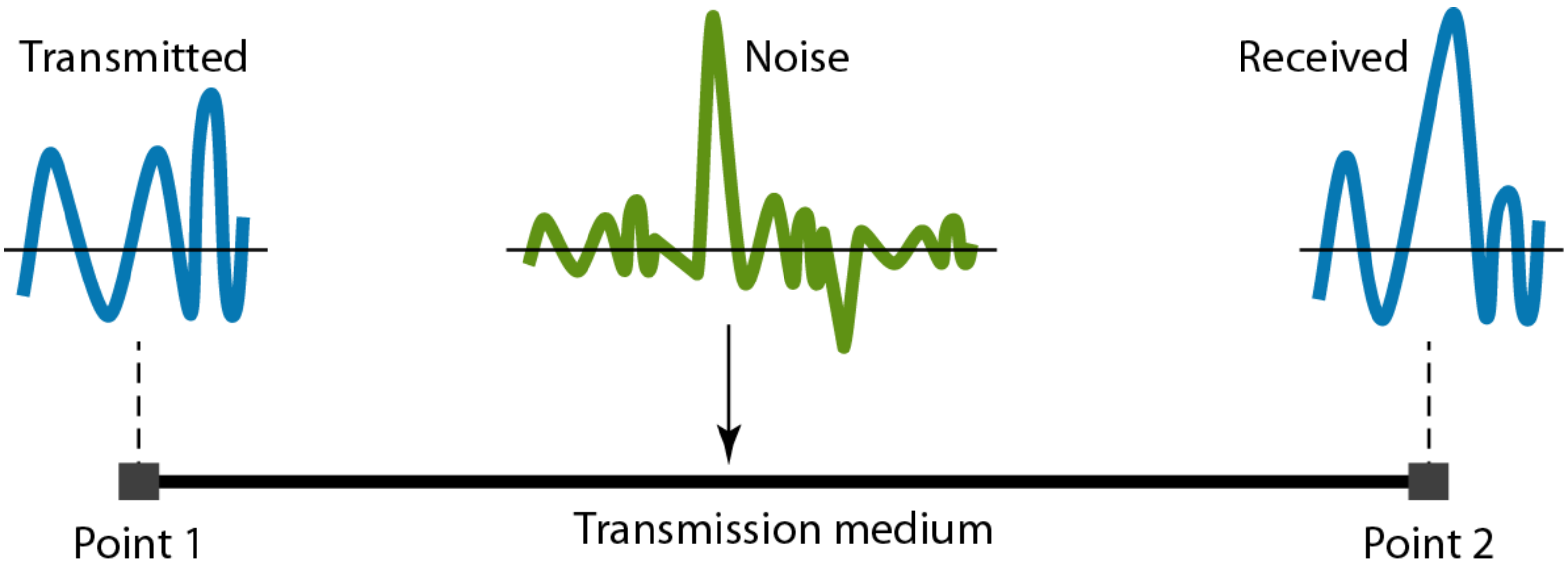
Figure 3.28 *Distortion*



Noise

- There are different types of noise
  - Thermal - random noise of electrons in the wire creates an extra signal
  - Induced - from motors and appliances, devices act are transmitter antenna and medium as receiving antenna.
  - Crosstalk - same as above but between two wires.
  - Impulse - Spikes that result from power lines, lighning, etc.

Figure 3.29 *Noise*



Signal to Noise Ratio (SNR)

- To measure the quality of a system the SNR is often used. It indicates the strength of the signal wrt the noise power in the system.
- It is the ratio between two powers.
- It is usually given in dB and referred to as SNR_{dB}.

Example 3.31

The power of a signal is 10 mW and the power of the noise is 1 μW; what are the values of SNR and SNR_{dB}?

Solution

The values of SNR and SNR_{dB} can be calculated as follows:

$$\text{SNR} = \frac{10{,}000\ \mu\text{W}}{1\ \text{mW}} = 10{,}000$$

$$\text{SNR}_{\text{dB}} = 10 \log_{10} 10{,}000 = 10 \log_{10} 10^4 = 40$$

Example 3.32

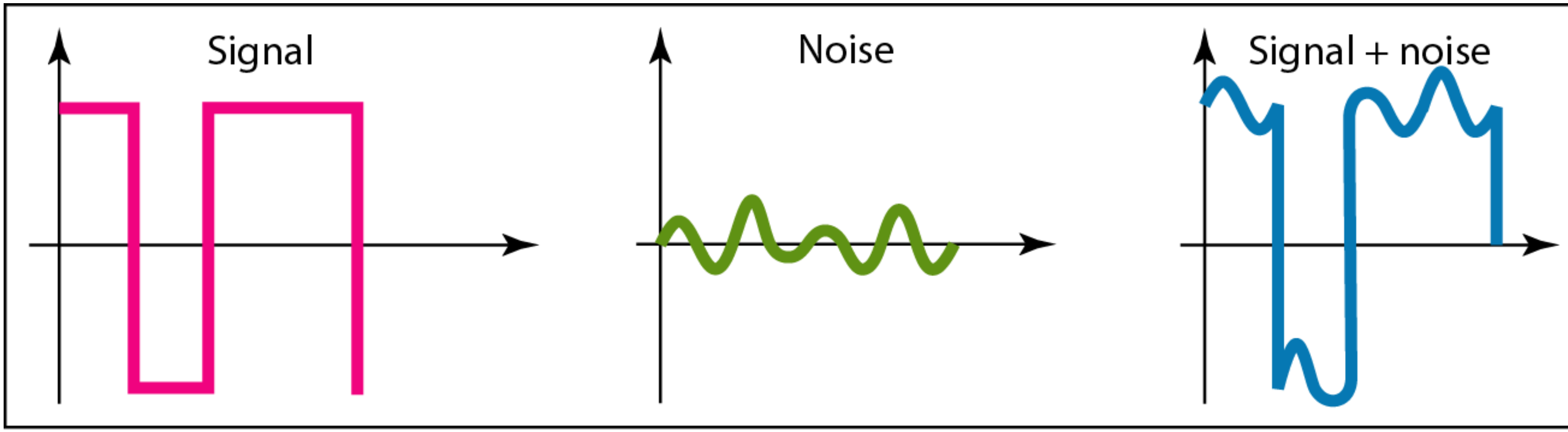
The values of SNR and SNR_{dB} for a noiseless channel are

$$\text{SNR} = \frac{\text{signal power}}{0} = \infty$$

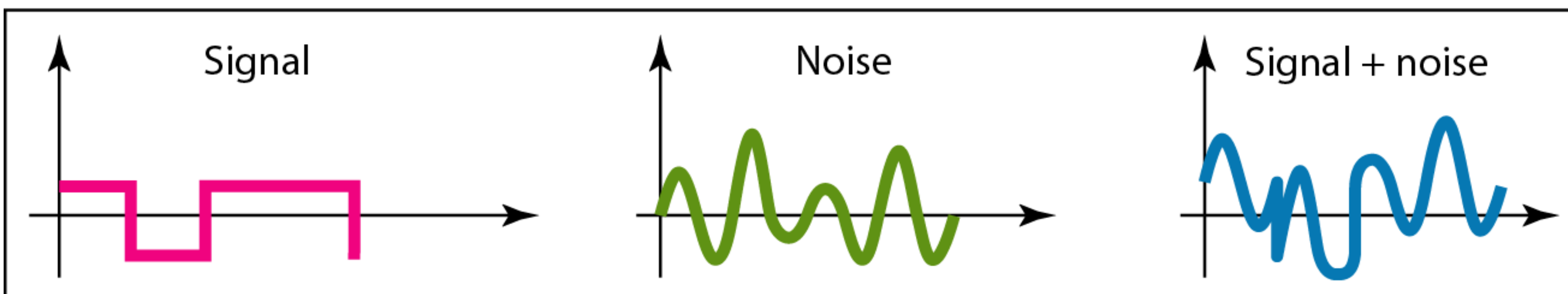
$$\text{SNR}_{\text{dB}} = 10 \log_{10} \infty = \infty$$

We can never achieve this ratio in real life; it is an ideal.

Figure 3.30 *Two cases of SNR: a high SNR and a low SNR*



a. Large SNR



b. Small SNR

3-5 DATA RATE LIMITS

A very important consideration in data communications is how fast we can send data, in bits per second, over a channel. Data rate depends on three factors:

1. ***The bandwidth available***
2. ***The level of the signals we use***
3. ***The quality of the channel (the level of noise)***

Topics discussed in this section:

- **Noiseless Channel: Nyquist Bit Rate**
- **Noisy Channel: Shannon Capacity**
- **Using Both Limits**

Note

Increasing the levels of a signal increases the probability of an error occurring, in other words it reduces the reliability of the system. Why??

Capacity of a System

- The bit rate of a system increases with an increase in the number of signal levels we use to denote a symbol.
- A symbol can consist of a single bit or "n" bits.
- The number of signal levels = 2^n.
- As the number of levels goes up, the spacing between level decreases -> increasing the probability of an error occurring in the presence of transmission impairments.

Nyquist Theorem

- Nyquist gives the upper bound for the bit rate of a transmission system by calculating the bit rate directly from the number of bits in a symbol (or signal levels) and the bandwidth of the system (assuming 2 symbols/per cycle and first harmonic).
- Nyquist theorem states that for a noiseless channel:

$$C = 2\,B\,\log_2 2^n$$

C= capacity in bps

B = bandwidth in Hz

Example 3.33

Does the Nyquist theorem bit rate agree with the intuitive bit rate described in baseband transmission?

Solution

They match when we have only two levels. We said, in baseband transmission, the bit rate is 2 times the bandwidth if we use only the first harmonic in the worst case. However, the Nyquist formula is more general than what we derived intuitively; it can be applied to baseband transmission and modulation. Also, it can be applied when we have two or more levels of signals.

Example 3.34

Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. The maximum bit rate can be calculated as

$$\boxed{\text{BitRate} = 2 \times 3000 \times \log_2 2 = 6000 \text{ bps}}$$

Example 3.35

Consider the same noiseless channel transmitting a signal with four signal levels (for each level, we send 2 bits). The maximum bit rate can be calculated as

$$\text{BitRate} = 2 \times 3000 \times \log_2 4 = 12{,}000 \text{ bps}$$

Example 3.36

We need to send 265 kbps over a noiseless channel with a bandwidth of 20 kHz. How many signal levels do we need?

Solution

We can use the Nyquist formula as shown:

$$265{,}000 = 2 \times 20{,}000 \times \log_2 L$$

$$\log_2 L = 6.625 \qquad L = 2^{6.625} = 98.7 \text{ levels}$$

Since this result is not a power of 2, we need to either increase the number of levels or reduce the bit rate. If we have 128 levels, the bit rate is 280 kbps. If we have 64 levels, the bit rate is 240 kbps.

Shannon's Theorem

- Shannon's theorem gives the capacity of a system in the presence of noise.

$$C = B \log_2(1 + SNR)$$

Example 3.37

Consider an extremely noisy channel in which the value of the signal-to-noise ratio is almost zero. In other words, the noise is so strong that the signal is faint. For this channel the capacity C is calculated as

$$\boxed{C = B \log_2 (1 + \mathrm{SNR}) = B \log_2 (1 + 0) = B \log_2 1 = B \times 0 = 0}$$

This means that the capacity of this channel is zero regardless of the bandwidth. In other words, we cannot receive any data through this channel.

Example 3.38

We can calculate the theoretical highest bit rate of a regular telephone line. A telephone line normally has a bandwidth of 3000. The signal-to-noise ratio is usually 3162. For this channel the capacity is calculated as

$$C = B \log_2(1 + \text{SNR}) = 3000 \log_2(1 + 3162) = 3000 \log_2 3163$$
$$= 3000 \times 11.62 = 34{,}860 \text{ bps}$$

This means that the highest bit rate for a telephone line is 34.860 kbps. If we want to send data faster than this, we can either increase the bandwidth of the line or improve the signal-to-noise ratio.

Example 3.39

The signal-to-noise ratio is often given in decibels. Assume that SNR_{dB} = 36 and the channel bandwidth is 2 MHz. The theoretical channel capacity can be calculated as

$$SNR_{dB} = 10 \log_{10} SNR \rightarrow SNR = 10^{SNR_{dB}/10} \rightarrow SNR = 10^{3.6} = 3981$$

$$C = B \log_2 (1 + SNR) = 2 \times 10^6 \times \log_2 3982 = 24 \text{ Mbps}$$

Example 3.40

For practical purposes, when the SNR is very high, we can assume that SNR + 1 is almost the same as SNR. In these cases, the theoretical channel capacity can be simplified to

$$\boxed{C = B \times \frac{\mathrm{SNR}_{\mathrm{dB}}}{3}}$$

For example, we can calculate the theoretical capacity of the previous example as

$$\boxed{C = 2\ \mathrm{MHz} \times \frac{36}{3} = 24\ \mathrm{Mbps}}$$

Example 3.41

We have a channel with a 1-MHz bandwidth. The SNR for this channel is 63. What are the appropriate bit rate and signal level?

Solution

First, we use the Shannon formula to find the upper limit.

$$\boxed{C = B \log_2 (1 + \text{SNR}) = 10^6 \log_2 (1 + 63) = 10^6 \log_2 64 = 6 \text{ Mbps}}$$

Example 3.41 (continued)

The Shannon formula gives us 6 Mbps, the upper limit. For better performance we choose something lower, 4 Mbps, for example. Then we use the Nyquist formula to find the number of signal levels.

$$\boxed{4 \text{ Mbps} = 2 \times 1 \text{ MHz} \times \log_2 L \quad \longrightarrow \quad L = 4}$$

Note

The Shannon capacity gives us the upper limit; the Nyquist formula tells us how many signal levels we need.

3-6 PERFORMANCE

One important issue in networking is the performance of the network—how good is it? We discuss quality of service, an overall measurement of network performance, in greater detail in Chapter 24. In this section, we introduce terms that we need for future chapters.

Topics discussed in this section:

- **Bandwidth - capacity of the system**
- **Throughput - no. of bits that can be pushed through**
- **Latency (Delay) - delay incurred by a bit from start to finish**
- **Bandwidth-Delay Product**

Note

In networking, we use the term bandwidth in two contexts.

- **The first, bandwidth in hertz, refers to the range of frequencies in a composite signal or the range of frequencies that a channel can pass.**
- **The second, bandwidth in bits per second, refers to the speed of bit transmission in a channel or link. Often referred to as Capacity.**

Example 3.42

The bandwidth of a subscriber line is 4 kHz for voice or data. The bandwidth of this line for data transmission can be up to 56,000 bps using a sophisticated modem to change the digital signal to analog.

Example 3.43

If the telephone company improves the quality of the line and increases the bandwidth to 8 kHz, we can send 112,000 bps by using the same technology as mentioned in Example 3.42.

Example 3.44

A network with bandwidth of 10 Mbps can pass only an average of 12,000 frames per minute with each frame carrying an average of 10,000 bits. What is the throughput of this network?

Solution

We can calculate the throughput as

$$\boxed{\text{Throughput} = \frac{12{,}000 \times 10{,}000}{60} = 2\ \text{Mbps}}$$

The throughput is almost one-fifth of the bandwidth in this case.

3.6

Propagation & Transmission delay

- Propagation speed - speed at which a bit travels though the medium from source to destination.
- Transmission speed - the speed at which all the bits in a message arrive at the destination. (difference in arrival time of first and last bit)

Propagation and Transmission Delay

- Propagation Delay = Distance/Propagation speed
- Transmission Delay = Message size/bandwidth bps
- Latency = Propagation delay + Transmission delay + Queueing time + Processing time

Example 3.45

What is the propagation time if the distance between the two points is 12,000 km? Assume the propagation speed to be 2.4 × 108 m/s in cable.

Solution

We can calculate the propagation time as

$$\text{Propagation time} = \frac{12{,}000 \times 1000}{2.4 \times 10^8} = 50 \text{ ms}$$

The example shows that a bit can go over the Atlantic Ocean in only 50 ms if there is a direct cable between the source and the destination.

Example 3.46

What are the propagation time and the transmission time for a 2.5-kbyte message (an e-mail) if the bandwidth of the network is 1 Gbps? Assume that the distance between the sender and the receiver is 12,000 km and that light travels at 2.4 × 108 m/s.

Solution

We can calculate the propagation and transmission time as shown on the next slide:

Example 3.46 (continued)

$$\text{Propagation time} = \frac{12{,}000 \times 1000}{2.4 \times 10^8} = 50 \text{ ms}$$

$$\text{Transmission time} = \frac{2500 \times 8}{10^9} = 0.020 \text{ ms}$$

Note that in this case, because the message is short and the bandwidth is high, the dominant factor is the propagation time, not the transmission time. The transmission time can be ignored.

Example 3.47

What are the propagation time and the transmission time for a 5-Mbyte message (an image) if the bandwidth of the network is 1 Mbps? Assume that the distance between the sender and the receiver is 12,000 km and that light travels at 2.4×10^8 m/s.

Solution

We can calculate the propagation and transmission times as shown on the next slide.

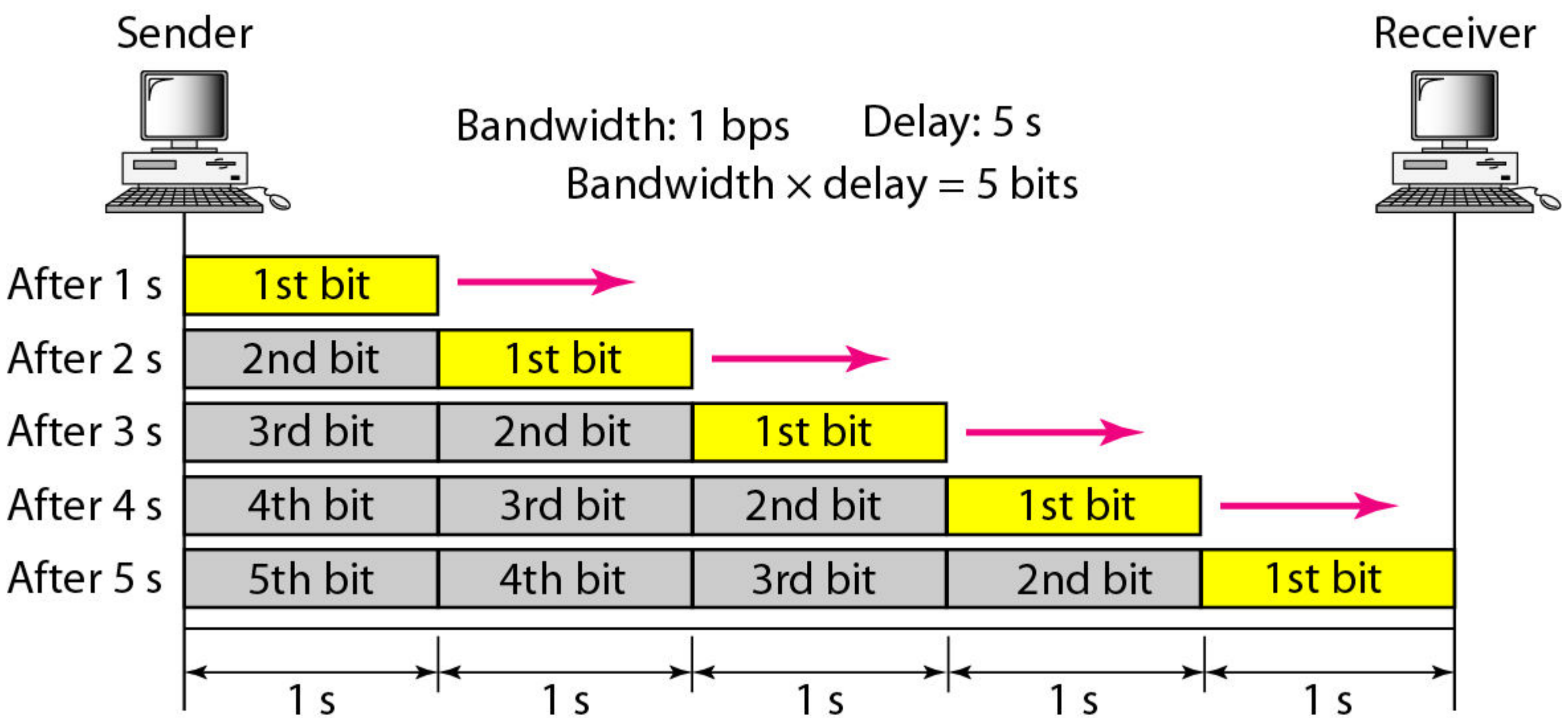
Example 3.47 (continued)

$$\text{Propagation time} = \frac{12{,}000 \times 1000}{2.4 \times 10^8} = 50 \text{ ms}$$

$$\text{Transmission time} = \frac{5{,}000{,}000 \times 8}{10^6} = 40 \text{ s}$$

Note that in this case, because the message is very long and the bandwidth is not very high, the dominant factor is the transmission time, not the propagation time. The propagation time can be ignored.

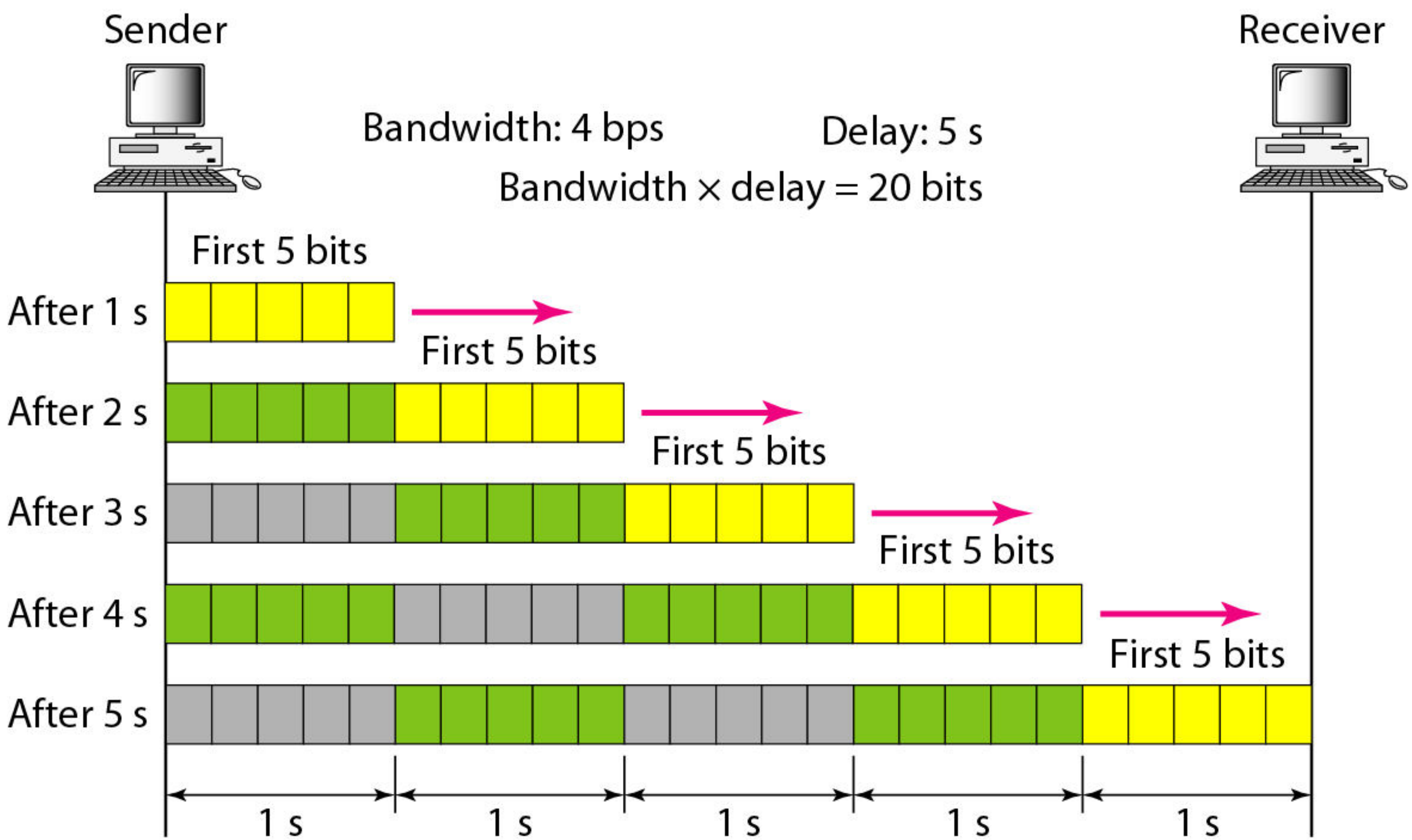
Figure 3.31 *Filling the link with bits for case 1*



Example 3.48

We can think about the link between two points as a pipe. The cross section of the pipe represents the bandwidth, and the length of the pipe represents the delay. We can say the volume of the pipe defines the bandwidth-delay product, as shown in Figure 3.33.

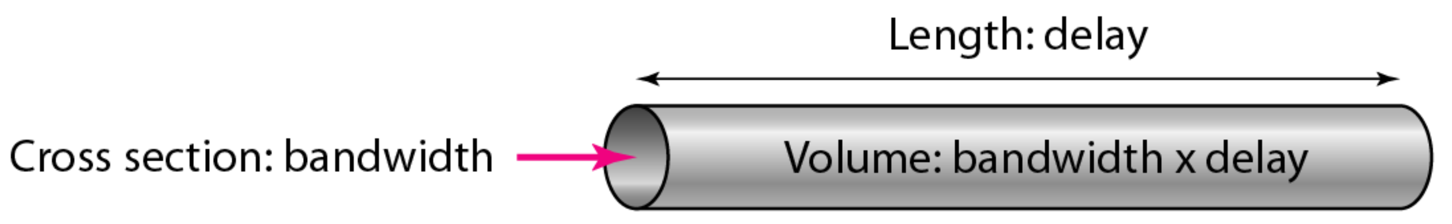
Figure 3.32 *Filling the link with bits in case 2*



Note

The bandwidth-delay product defines the number of bits that can fill the link.

Figure 3.33 *Concept of bandwidth-delay product*



Chapter 4

Digital Transmission

4-1 DIGITAL-TO-DIGITAL CONVERSION

In this section, we see how we can represent digital data by using digital signals. The conversion involves three techniques: line coding, block coding, and scrambling. Line coding is always needed; block coding and scrambling may or may not be needed.

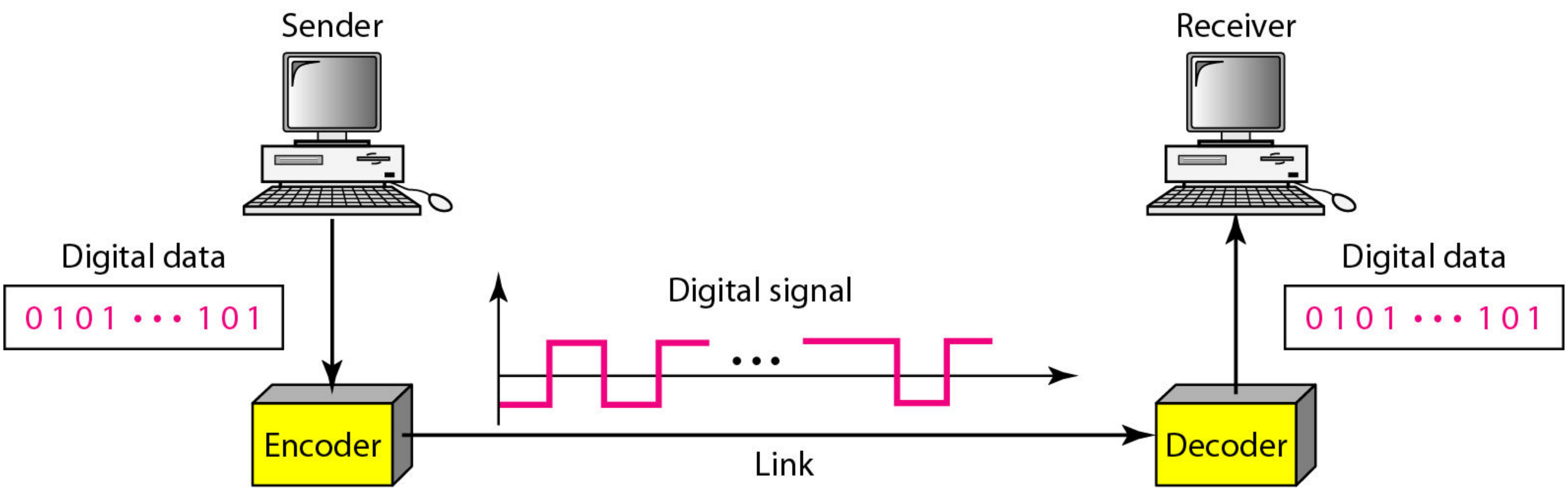
Topics discussed in this section:

- **Line Coding**
- **Line Coding Schemes**
- **Block Coding**
- **Scrambling**

Line Coding

- Converting a string of 1's and 0's (digital data) into a sequence of signals that denote the 1's and 0's.
- For example a high voltage level (+V) could represent a "1" and a low voltage level (0 or -V) could represent a "0".

Figure 4.1 *Line coding and decoding*



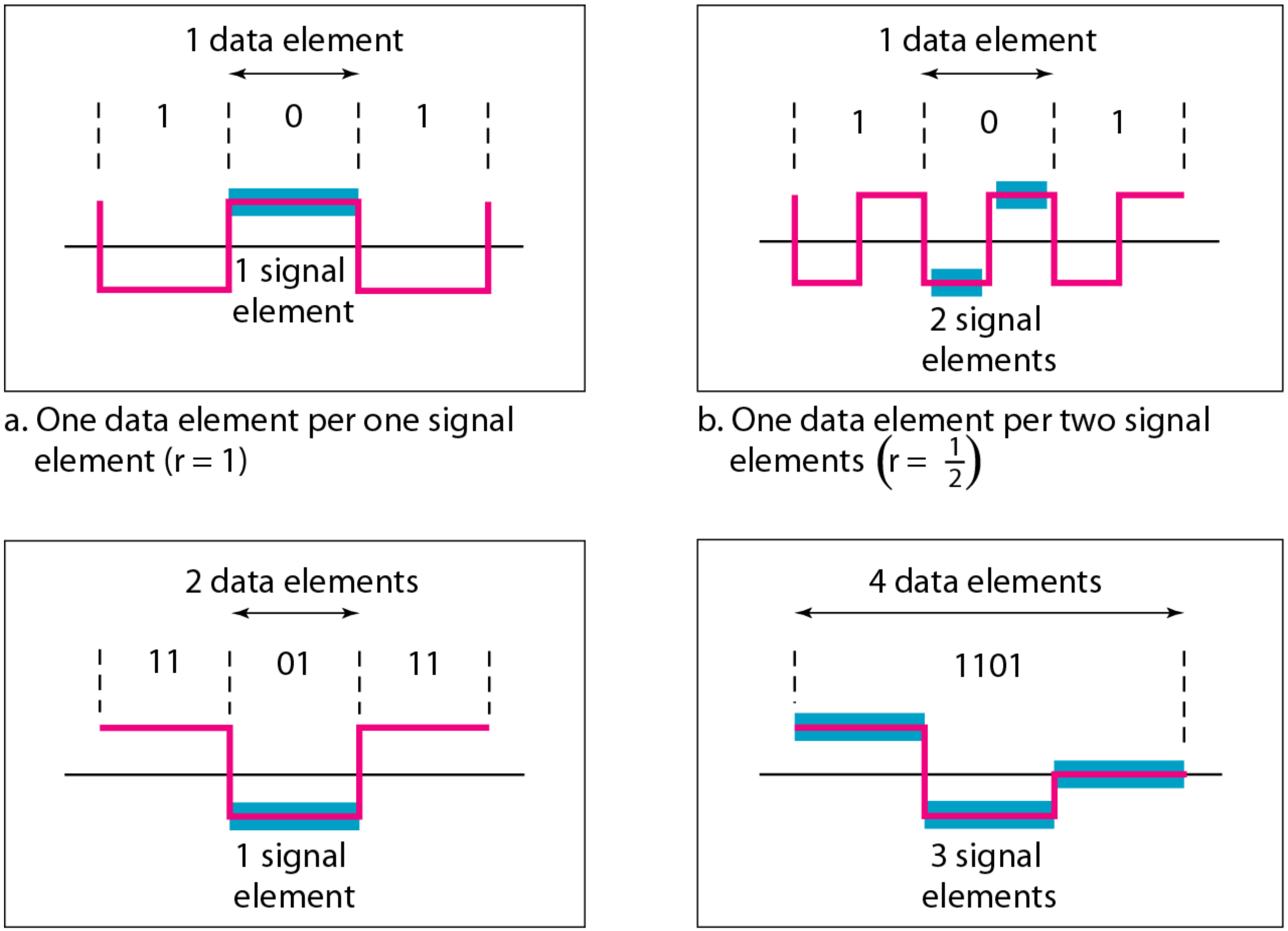
Mapping Data symbols onto Signal levels

- A data symbol (or element) can consist of a number of data bits:
  - 1 , 0 or
  - 11, 10, 01, ……
- A data symbol can be coded into a single signal element or multiple signal elements
  - 1 -> +V, 0 -> -V
  - 1 -> +V and -V, 0 -> -V and +V
- The ratio 'r' is the number of data elements carried by a signal element.

Relationship between data rate and signal rate

- The data rate defines the number of bits sent per sec - bps. It is often referred to the bit rate.
- The signal rate is the number of signal elements sent in a second and is measured in bauds. It is also referred to as the modulation rate.
- Goal is to increase the data rate whilst reducing the baud rate.

Figure 4.2 *Signal element versus data element*



a. One data element per one signal element (r = 1)

b. One data element per two signal elements $\left(r = \frac{1}{2}\right)$

c. Two data elements per one signal element (r = 2)

d. Four data elements per three signal elements $\left(r = \frac{4}{3}\right)$

Data rate and Baud rate

- The baud or signal rate can be expressed as:

  $$S = c \times N \times 1/r \text{ bauds}$$

  where N is data rate

  c is the case factor (worst, best & avg.)

  r is the ratio between data element & signal element

Example 4.1

A signal is carrying data in which one data element is encoded as one signal element (r = 1). If the bit rate is 100 kbps, what is the average value of the baud rate if c is between 0 and 1?

Solution

We assume that the average value of c is 1/2 . The baud rate is then

$$\boxed{S = c \times N \times \frac{1}{r} = \frac{1}{2} \times 100{,}000 \times \frac{1}{1} = 50{,}000 = 50 \text{ kbaud}}$$

Note

Although the actual bandwidth of a digital signal is infinite, the effective bandwidth is finite.

Example 4.2

The maximum data rate of a channel (see Chapter 3) is $N_{max} = 2 \times B \times \log_2 L$ (defined by the Nyquist formula). Does this agree with the previous formula for N_{max}?

Solution

A signal with L levels actually can carry $\log_2 L$ bits per level. If each level corresponds to one signal element and we assume the average case ($c = 1/2$), then we have

$$\boxed{N_{\max} = \frac{1}{c} \times B \times r = 2 \times B \times \log_2 L}$$

Considerations for choosing a good signal element referred to as line encoding

- Baseline wandering - a receiver will evaluate the average power of the received signal (called the baseline) and use that to determine the value of the incoming data elements. If the incoming signal does not vary over a long period of time, the baseline will drift and thus cause errors in detection of incoming data elements.
- A good line encoding scheme will prevent long runs of fixed amplitude.

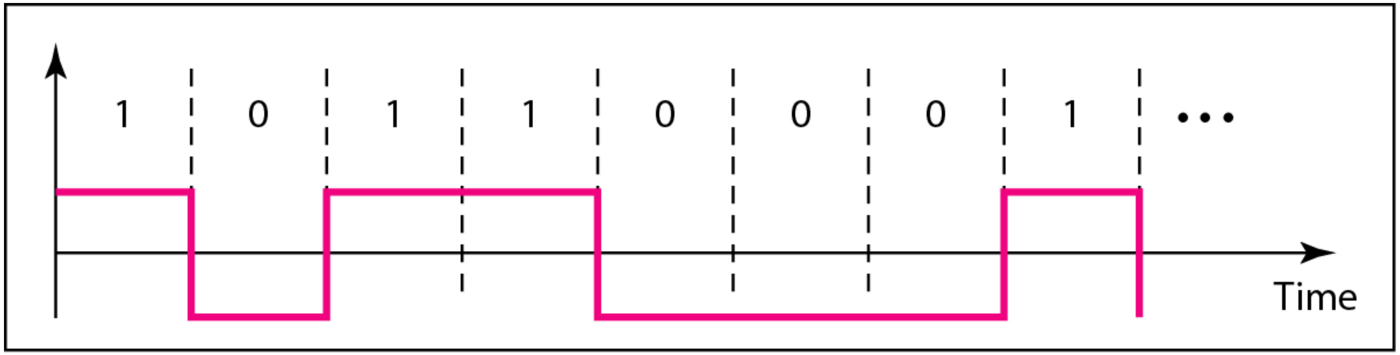
Line encoding C/Cs

- DC components - when the voltage level remains constant for long periods of time, there is an increase in the low frequencies of the signal. Most channels are bandpass and may not support the low frequencies.
- This will require the removal of the dc component of a transmitted signal.

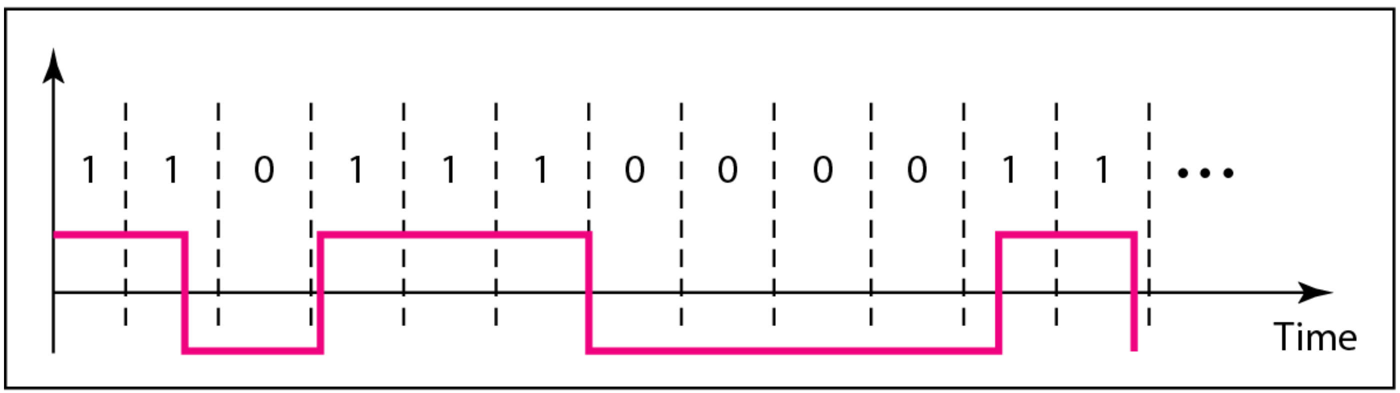
Line encoding C/Cs

- Self synchronization - the clocks at the sender and the receiver must have the same bit interval.
- If the receiver clock is faster or slower it will misinterpret the incoming bit stream.

Figure 4.3 *Effect of lack of synchronization*



a. Sent



b. Received

Example 4.3

In a digital transmission, the receiver clock is 0.1 percent faster than the sender clock. How many extra bits per second does the receiver receive if the data rate is 1 kbps? How many if the data rate is 1 Mbps?

Solution

At 1 kbps, the receiver receives 1001 bps instead of 1000 bps.

1000 bits sent	1001 bits received	1 extra bps

At 1 Mbps, the receiver receives 1,001,000 bps instead of 1,000,000 bps.

1,000,000 bits sent	1,001,000 bits received	1000 extra bps

Line encoding C/Cs

- Error detection - errors occur during transmission due to line impairments.
- Some codes are constructed such that when an error occurs it can be detected. For example: a particular signal transition is not part of the code. When it occurs, the receiver will know that a symbol error has occurred.

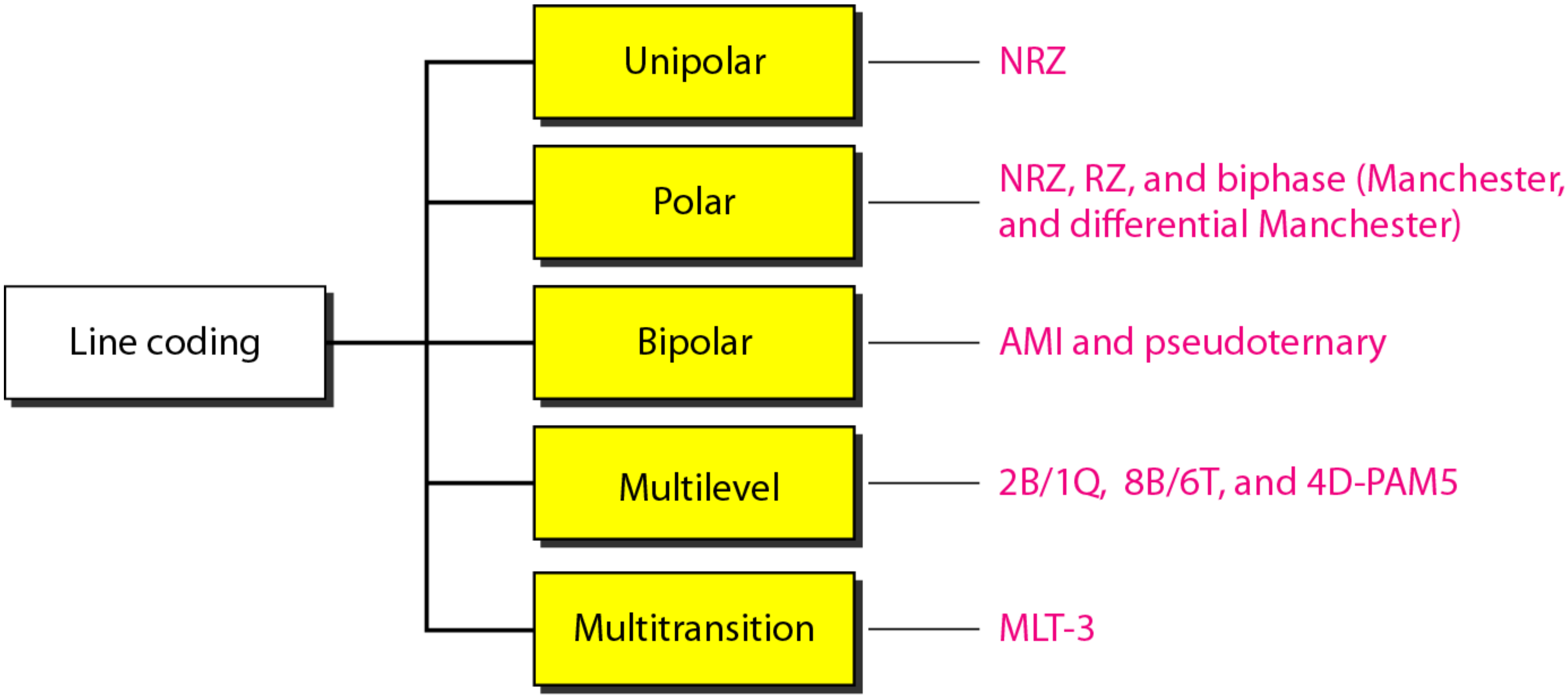
Line encoding C/Cs

- Noise and interference - there are line encoding techniques that make the transmitted signal “immune” to noise and interference.
- This means that the signal cannot be corrupted, it is stronger than error detection.

Line encoding C/Cs

- Complexity - the more robust and resilient the code, the more complex it is to implement and the price is often paid in baud rate or required bandwidth.

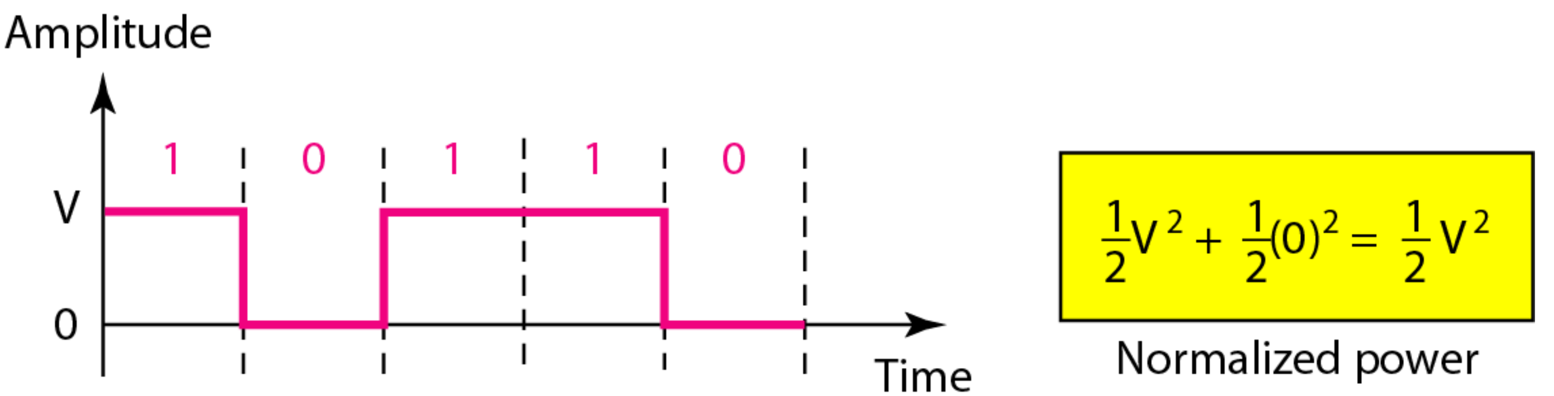
Figure 4.4 *Line coding schemes*



Unipolar

- All signal levels are on one side of the time axis - either above or below
- NRZ - Non Return to Zero scheme is an example of this code. The signal level does not return to zero during a symbol transmission.
- Scheme is prone to baseline wandering and DC components. It has no synchronization or any error detection. It is simple but costly in power consumption.

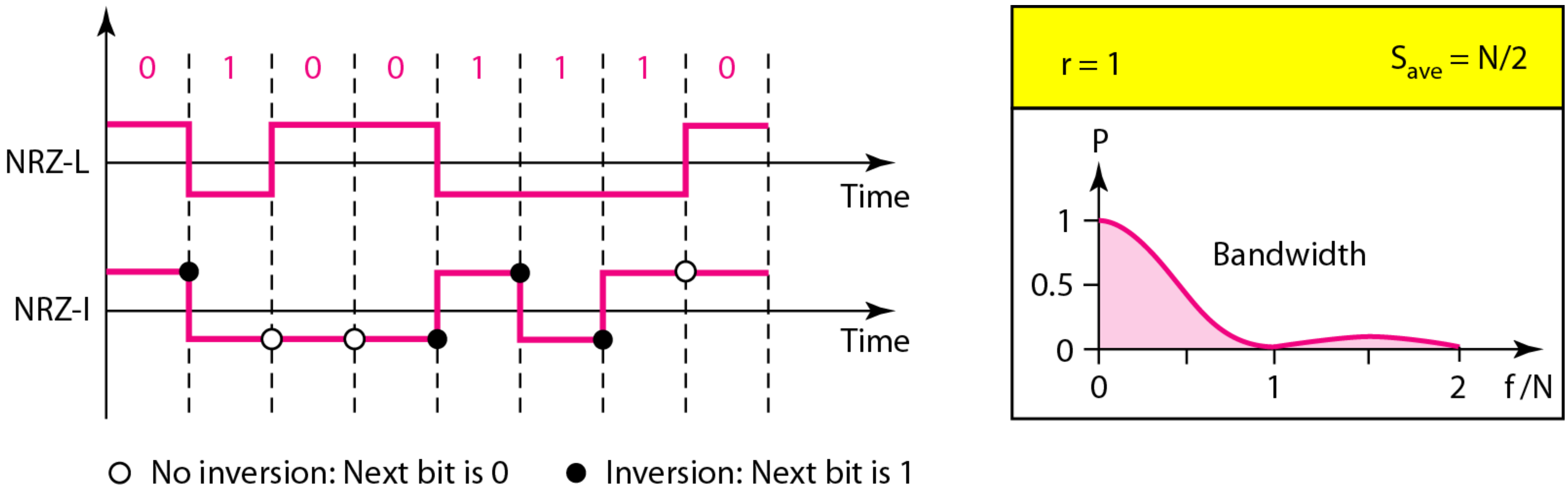
Figure 4.5 *Unipolar NRZ scheme*



Polar - NRZ

- The voltages are on both sides of the time axis.
- Polar NRZ scheme can be implemented with two voltages. E.g. +V for 1 and -V for 0.
- There are two versions:
  - NZR - Level (NRZ-L) - positive voltage for one symbol and negative for the other
  - NRZ - Inversion (NRZ-I) - the change or lack of change in polarity determines the value of a symbol. E.g. a "1" symbol inverts the polarity a "0" does not.

Figure 4.6 *Polar NRZ-L and NRZ-I schemes*



Note

In NRZ-L the level of the voltage determines the value of the bit. In NRZ-I the inversion or the lack of inversion determines the value of the bit.

Note

NRZ-L and NRZ-I both have an average signal rate of N/2 Bd.

Note

NRZ-L and NRZ-I both have a DC component problem and baseline wandering, it is worse for NRZ-L. Both have no self synchronization &no error detection. Both are relatively simple to implement.

Example 4.4

A system is using NRZ-I to transfer 1-Mbps data. What are the average signal rate and minimum bandwidth?

Solution

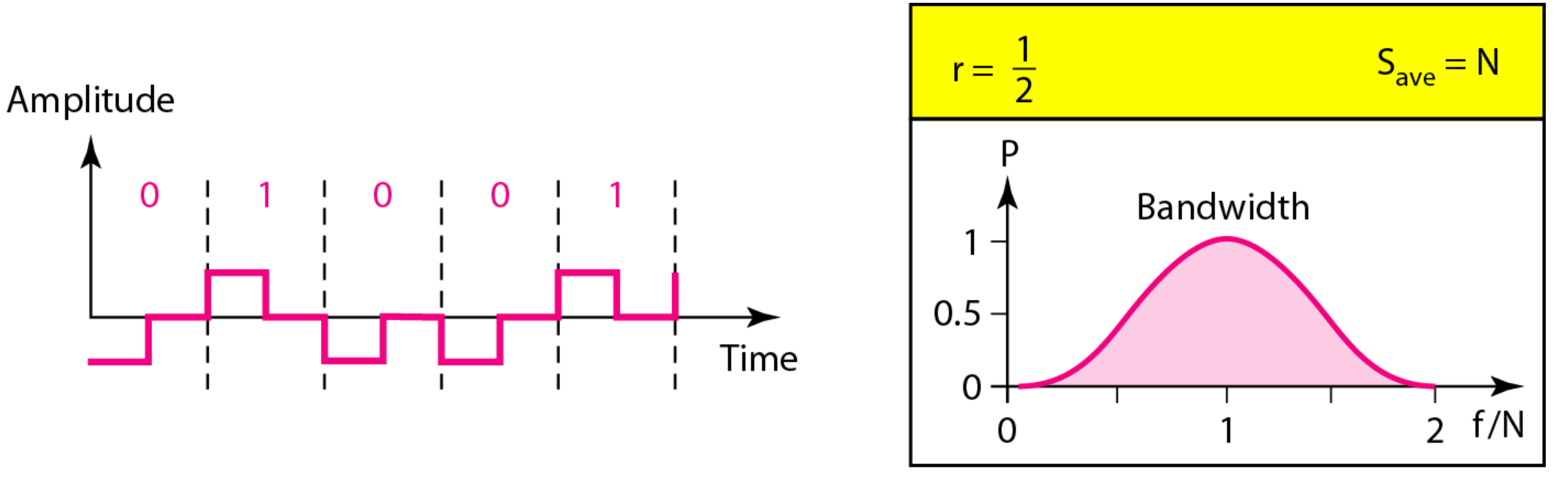
The average signal rate is S= c x N x R = 1/2 x N x 1 = 500 kbaud. The minimum bandwidth for this average baud rate is B_{min} = S = 500 kHz.

Note c = 1/2 for the avg. case as worst case is 1 and best case is 0

Polar - RZ

- The Return to Zero (RZ) scheme uses three voltage values. +, 0, -.
- Each symbol has a transition in the middle. Either from high to zero or from low to zero.
- This scheme has more signal transitions (two per symbol) and therefore requires a wider bandwidth.
- No DC components or baseline wandering.
- Self synchronization - transition indicates symbol value.
- More complex as it uses three voltage level. It has no error detection capability.

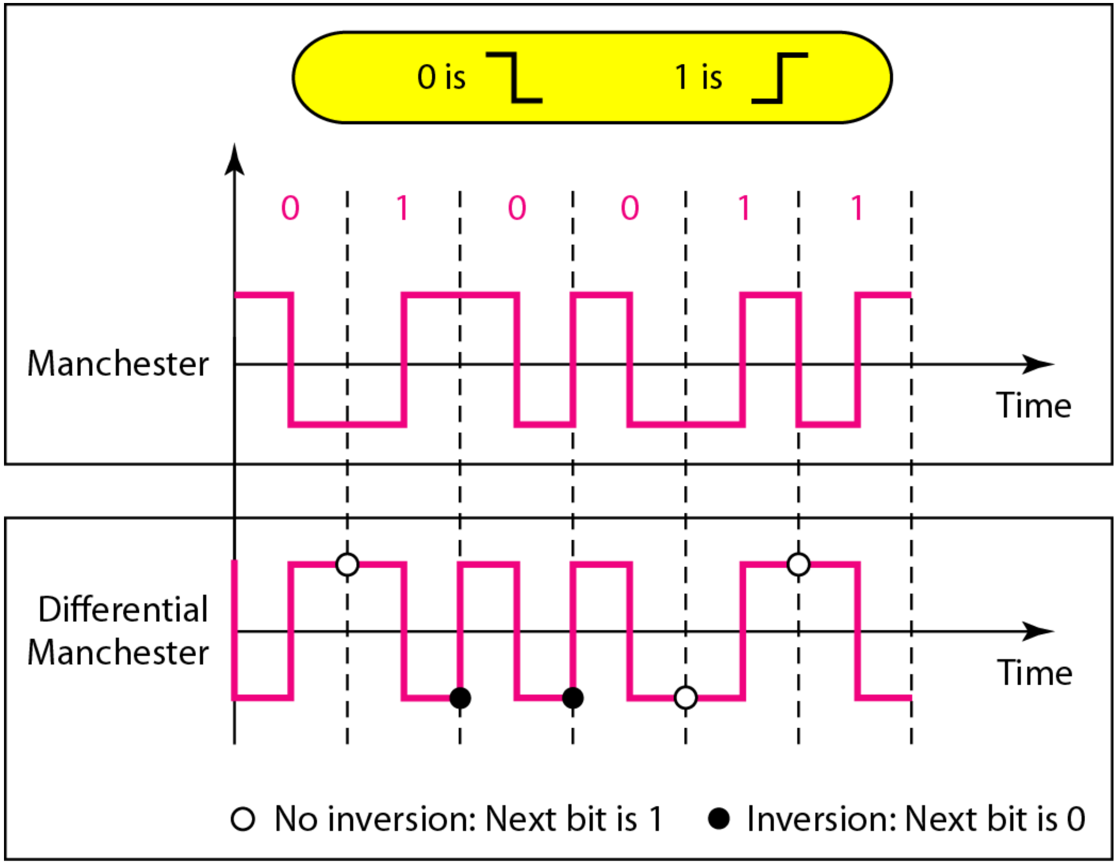
Figure 4.7 *Polar RZ scheme*

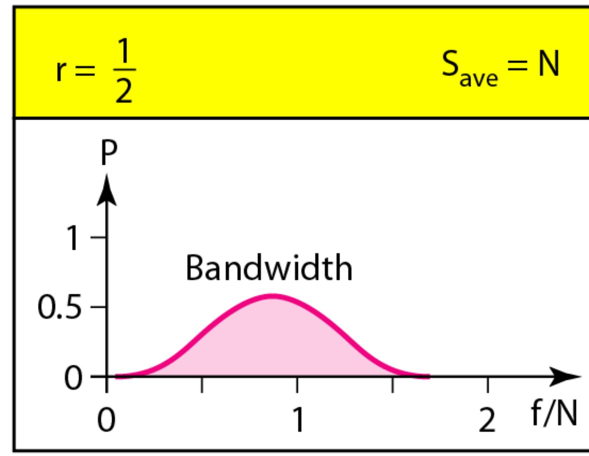


Polar - Biphase: Manchester and Differential Manchester

- Manchester coding consists of combining the NRZ-L and RZ schemes.
  - Every symbol has a level transition in the middle: from high to low or low to high. Uses only two voltage levels.
- Differential Manchester coding consists of combining the NRZ-I and RZ schemes.
  - Every symbol has a level transition in the middle. But the level at the beginning of the symbol is determined by the symbol value. One symbol causes a level change the other does not.

Figure 4.8 *Polar biphase: Manchester and differential Manchester schemes*





Note

In Manchester and differential Manchester encoding, the transition at the middle of the bit is used for synchronization.

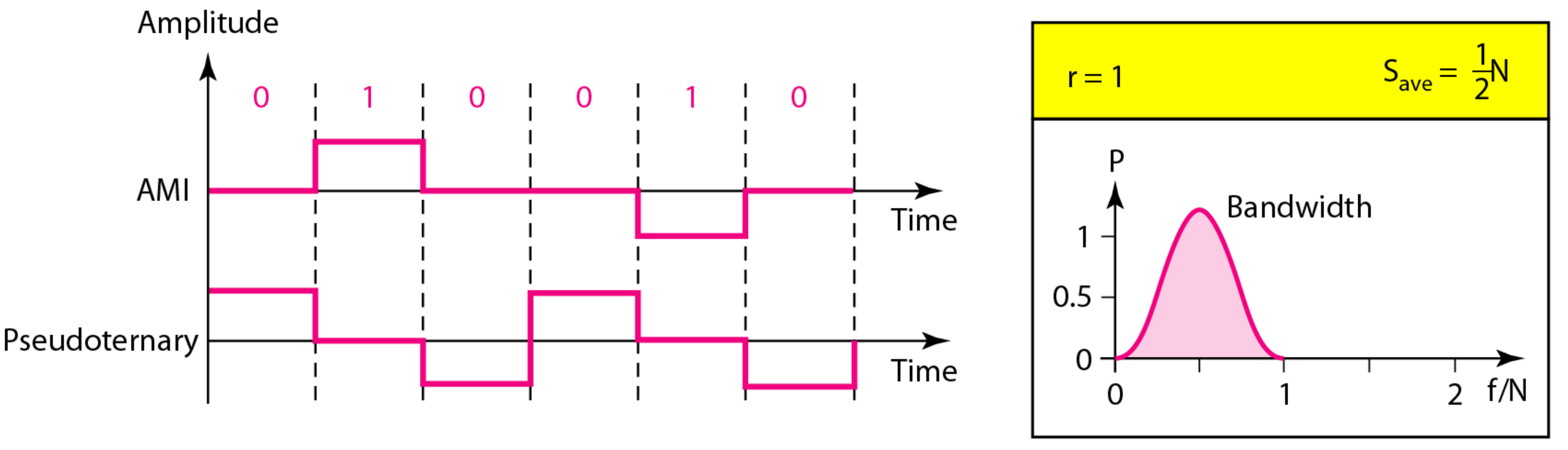
Note

The minimum bandwidth of Manchester and differential Manchester is 2 times that of NRZ. The is no DC component and no baseline wandering. None of these codes has error detection.

Bipolar - AMI and Pseudoternary

- Code uses 3 voltage levels: - +, 0, -, to represent the symbols (note not transitions to zero as in RZ).
- Voltage level for one symbol is at "0" and the other alternates between + & -.
- Bipolar Alternate Mark Inversion (AMI) - the "0" symbol is represented by zero voltage and the "1" symbol alternates between +V and -V.
- Pseudoternary is the reverse of AMI.

Figure 4.9 *Bipolar schemes: AMI and pseudoternary*



Bipolar C/Cs

- It is a better alternative to NRZ.
- Has no DC component or baseline wandering.
- Has no self synchronization because long runs of "0"s results in no signal transitions.
- No error detection.

Multilevel Schemes

- In these schemes we increase the number of data bits per symbol thereby increasing the bit rate.
- Since we are dealing with binary data we only have 2 types of data element a 1 or a 0.
- We can combine the 2 data elements into a pattern of "m" elements to create "2^m" symbols.
- If we have L signal levels, we can use "n" signal elements to create L^n signal elements.

Code C/Cs

- Now we have 2^m symbols and L^n signals.
- If $2^m > L^n$ then we cannot represent the data elements, we don't have enough signals.
- If $2^m = L^n$ then we have an exact mapping of one symbol on one signal.
- If $2^m < L^n$ then we have more signals than symbols and we can choose the signals that are more distinct to represent the symbols and therefore have better noise immunity and error detection as some signals are not valid.

Note

In *m*B*n*L schemes, a pattern of *m* data elements is encoded as a pattern of *n* signal elements in which $2^m \leq L^n$.

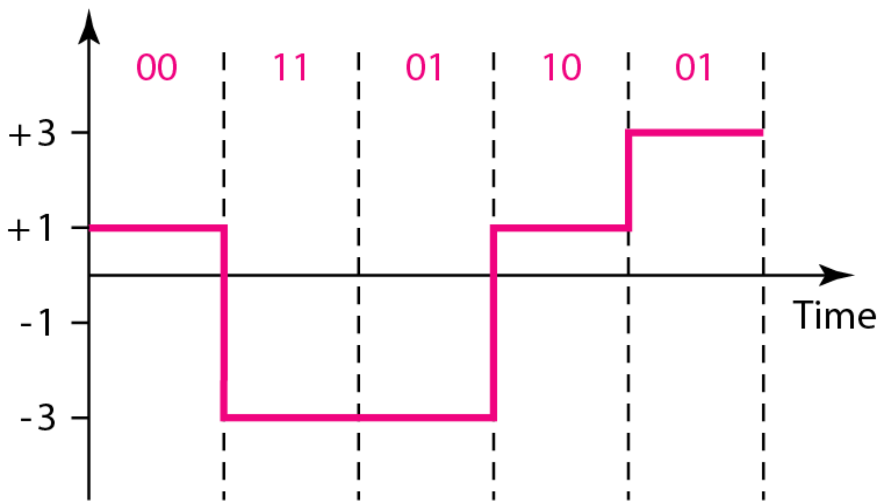
Representing Multilevel Codes

- We use the notation mBnL, where m is the length of the binary pattern, B represents binary data, n represents the length of the signal pattern and L the number of levels.
- L = B binary, L = T for 3 ternary, L = Q for 4 quaternary.

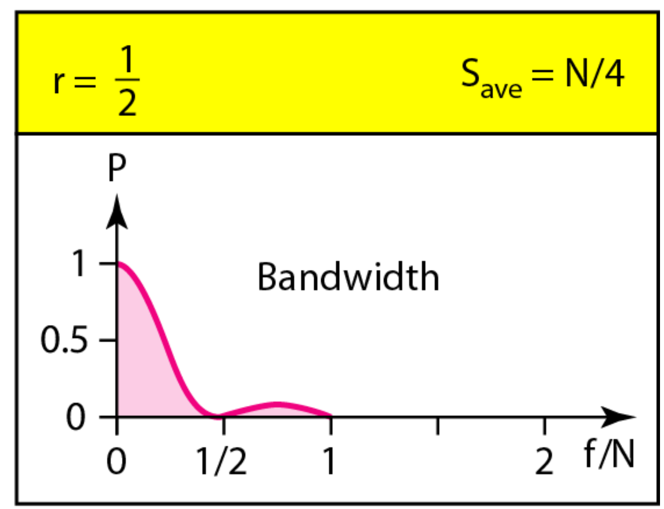
Figure 4.10 *Multilevel: 2B1Q scheme*

Next bits	Next level (Previous level: positive)	Next level (Previous level: negative)
00	+1	-1
01	+3	-3
10	-1	+1
11	-3	+3

Transition table



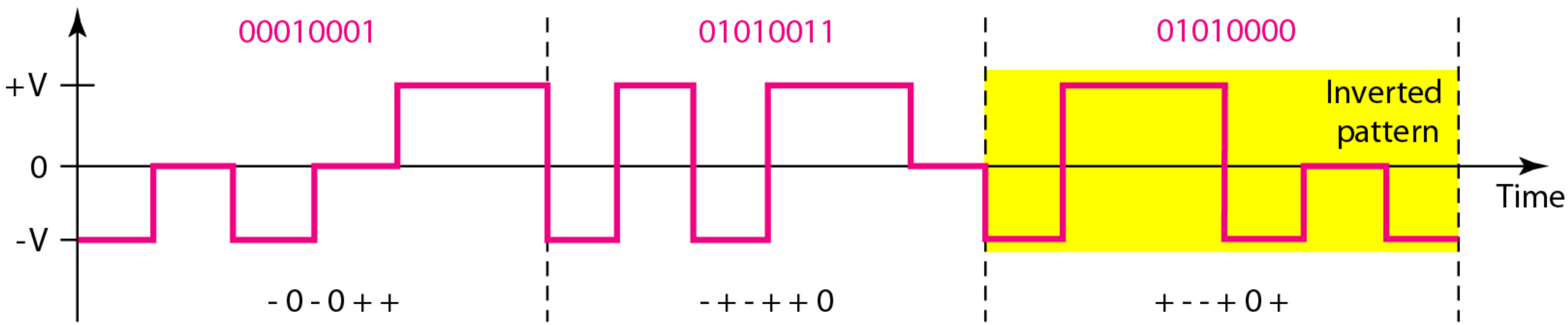
Assuming positive original level



Redundancy

- In the 2B1Q scheme we have no redundancy and we see that a DC component is present.
- If we use a code with redundancy we can decide to use only “0” or “+” weighted codes (more +’s than -’s in the signal element) and invert any code that would create a DC component. E.g. ‘+00++-’ -> ‘-00--+’
- Receiver will know when it receives a “-” weighted code that it should invert it as it doesn’t represent any valid symbol.

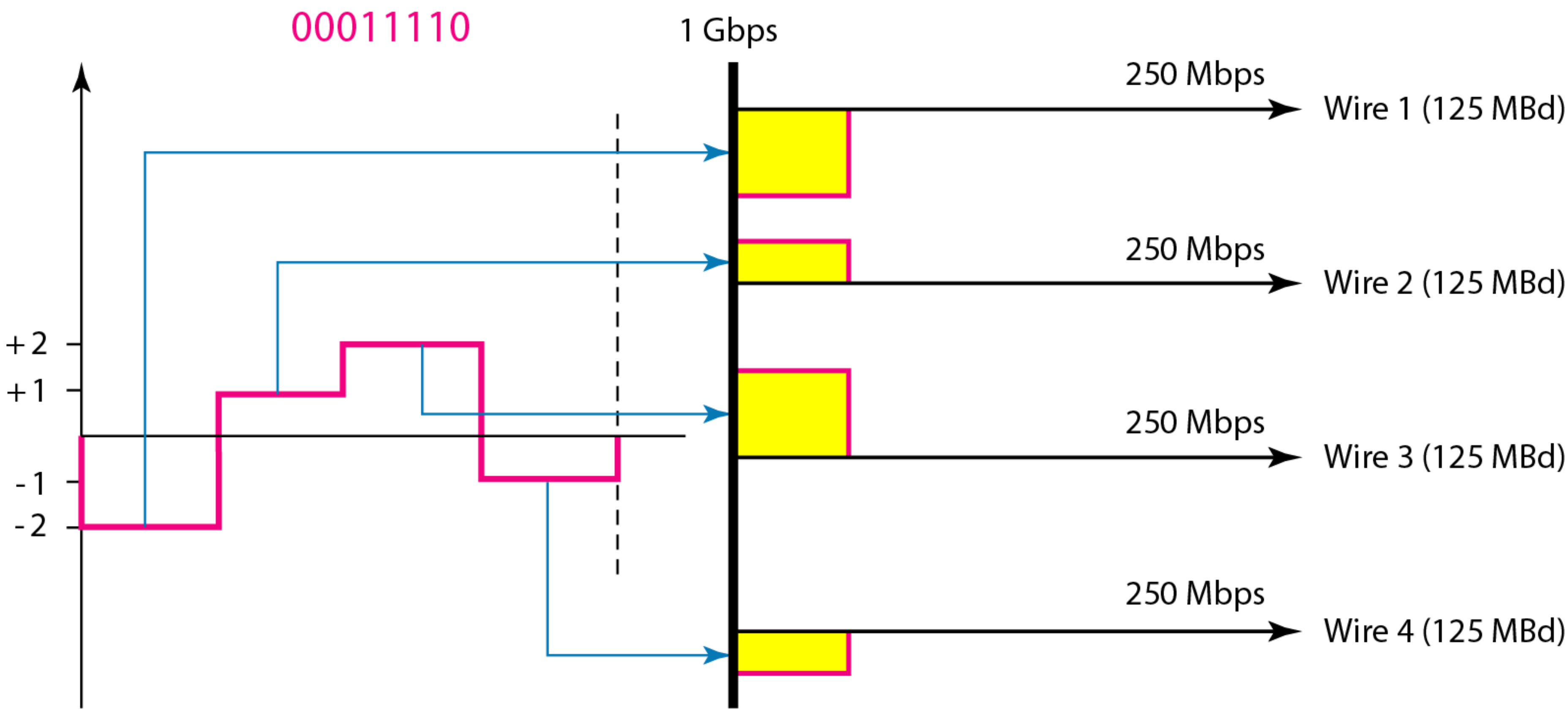
Figure 4.11 *Multilevel: 8B6T scheme*



Multilevel using multiple channels

- In some cases, we split the signal transmission up and distribute it over several links.
- The separate segments are transmitted simultaneously. This reduces the signalling rate per link -> lower bandwidth.
- This requires all bits for a code to be stored.
- xD: means that we use ‘x’ links
- YYYz: We use ‘z’ levels of modulation where YYY represents the type of modulation (e.g. pulse ampl. mod. PAM).
- Codes are represented as: xD-YYYz

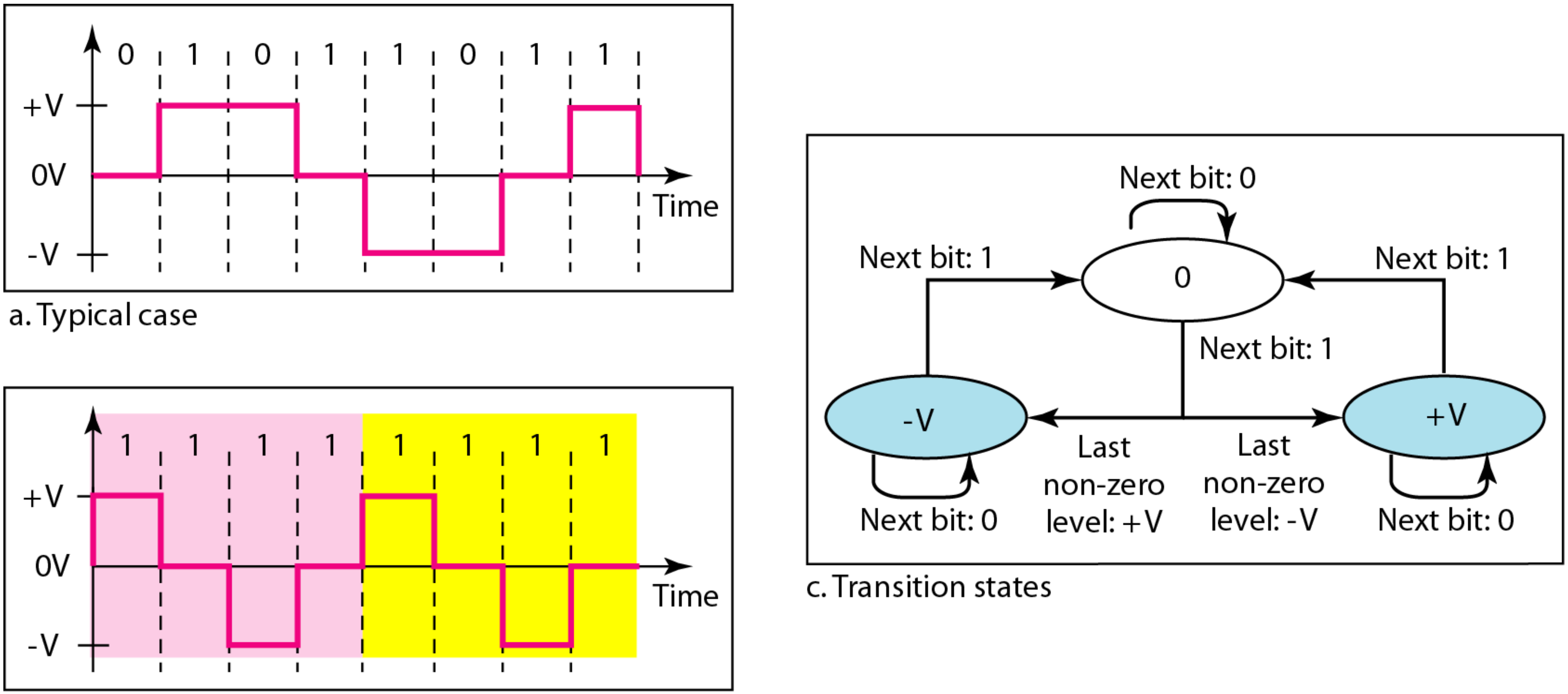
Figure 4.12 *Multilevel: 4D-PAM5 scheme*



Multitransition Coding

- Because of synchronization requirements we force transitions. This can result in very high bandwidth requirements -> more transitions than are bits (e.g. mid bit transition with inversion).
- Codes can be created that are differential at the bit level forcing transitions at bit boundaries. This results in a bandwidth requirement that is equivalent to the bit rate.
- In some instances, the bandwidth requirement may even be lower, due to repetitive patterns resulting in a periodic signal.

Figure 4.13 *Multitransition: MLT-3 scheme*



a. Typical case

b. Worse case

c. Transition states

MLT-3

- Signal rate is same as NRZ-I
- But because of the resulting bit pattern, we have a periodic signal for worst case bit pattern: 1111
- This can be approximated as an analog signal a frequency 1/4 the bit rate!

Table 4.1 ***Summary of line coding schemes***

Category	*Scheme*	*Bandwidth (average)*	*Characteristics*
Unipolar	NRZ	$B = N/2$	Costly, no self-synchronization if long 0s or 1s, DC
Unipolar	NRZ-L	$B = N/2$	No self-synchronization if long 0s or 1s, DC
	NRZ-I	$B = N/2$	No self-synchronization for long 0s, DC
	Biphase	$B = N$	Self-synchronization, no DC, high bandwidth
Bipolar	AMI	$B = N/2$	No self-synchronization for long 0s, DC
Multilevel	2B1Q	$B = N/4$	No self-synchronization for long same double bits
	8B6T	$B = 3N/4$	Self-synchronization, no DC
	4D-PAM5	$B = N/8$	Self-synchronization, no DC
Multiline	MLT-3	$B = N/3$	No self-synchronization for long 0s

Block Coding

- For a code to be capable of error detection, we need to add redundancy, i.e., extra bits to the data bits.
- Synchronization also requires redundancy - transitions are important in the signal flow and must occur frequently.
- Block coding is done in three steps: division, substitution and combination.
- It is distinguished from multilevel coding by use of the slash - xB/yB.
- The resulting bit stream prevents certain bit combinations that when used with line encoding would result in DC components or poor sync. quality.

Note

Block coding is normally referred to as *m*B/*n*B coding;
it replaces each *m*-bit group with an *n*-bit group.

Figure 4.14 *Block coding concept*

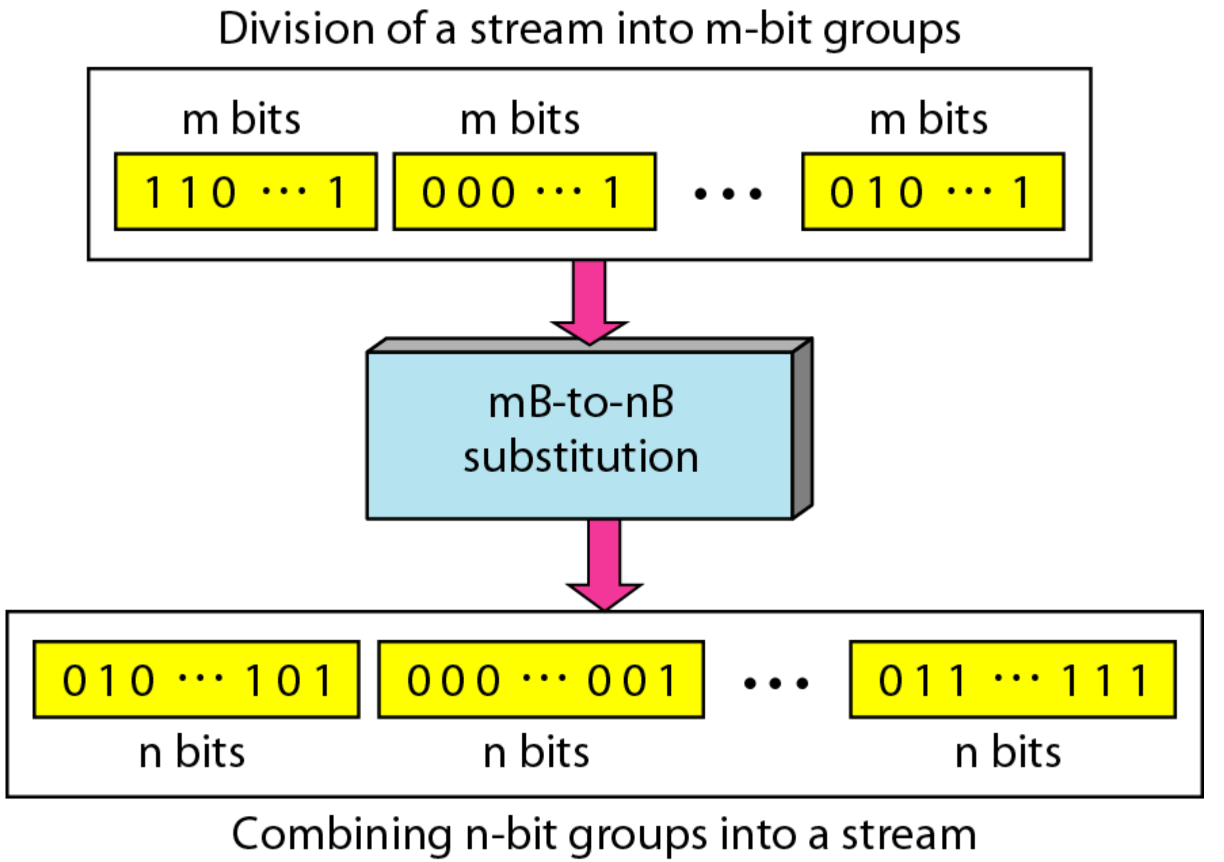


Figure 4.15 *Using block coding 4B/5B with NRZ-I line coding scheme*

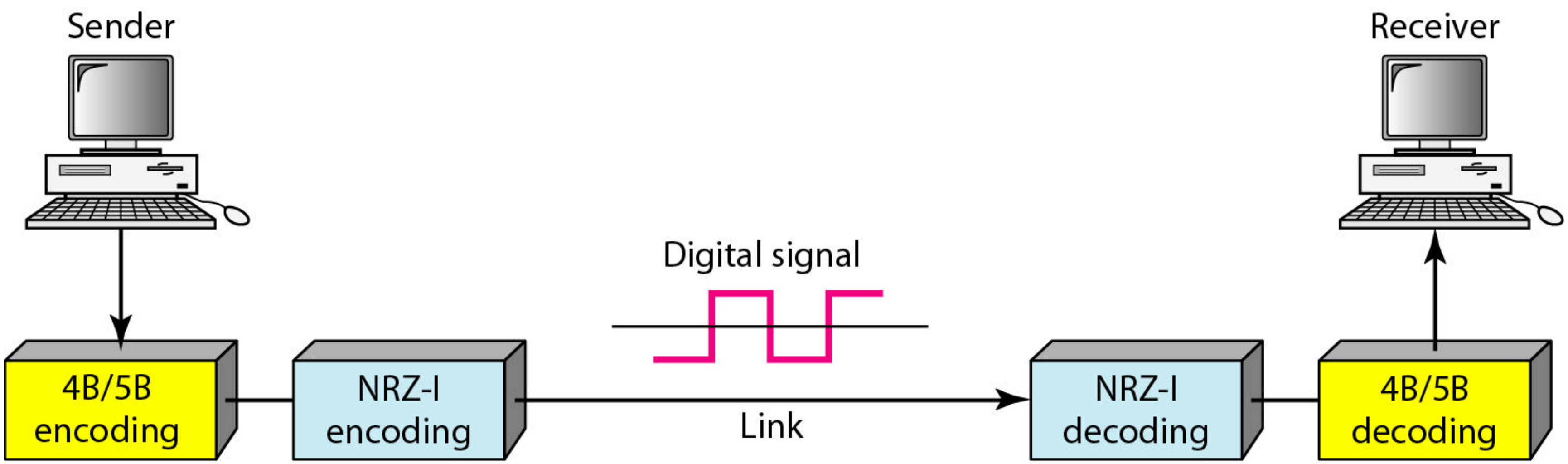
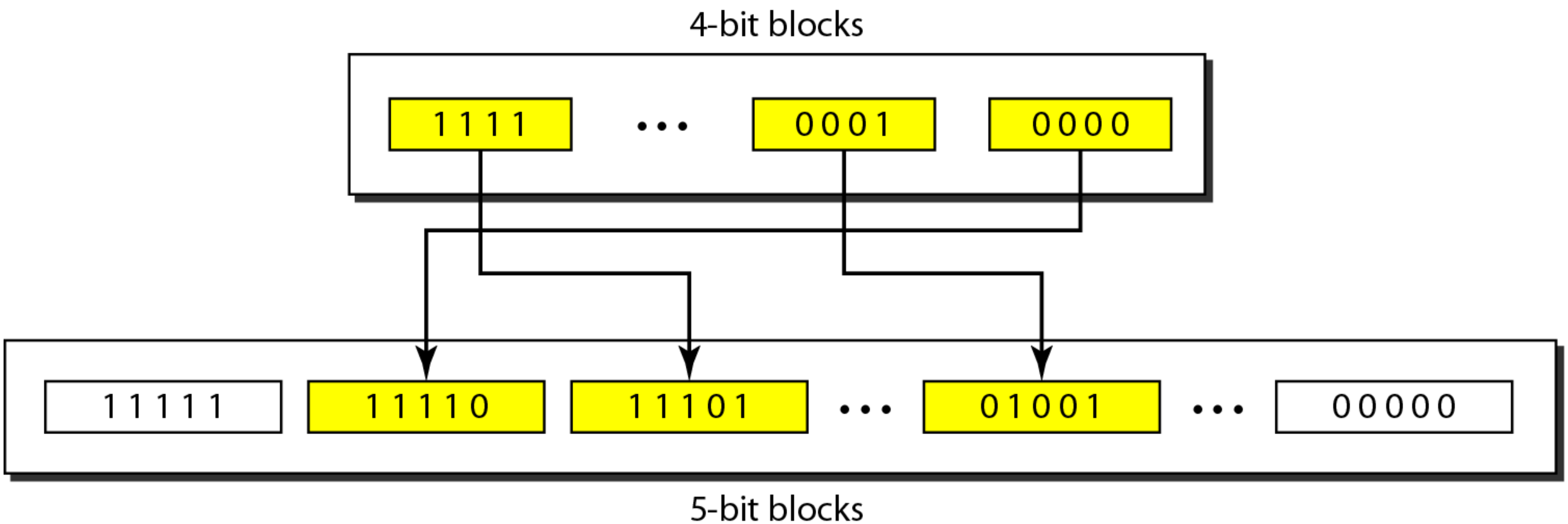


Table 4.2 *4B/5B mapping codes*

Data Sequence	*Encoded Sequence*	*Control Sequence*	*Encoded Sequence*
0000	11110	Q (Quiet)	00000
0001	01001	I (Idle)	11111
0010	10100	H (Halt)	00100
0011	10101	J (Start delimiter)	11000
0100	01010	K (Start delimiter)	10001
0101	01011	T (End delimiter)	01101
0110	01110	S (Set)	11001
0111	01111	R (Reset)	00111
1000	10010		
1001	10011		
1010	10110		
1011	10111		
1100	11010		
1101	11011		
1110	11100		
1111	11101		

Figure 4.16 *Substitution in 4B/5B block coding*



Redundancy

- A 4 bit data word can have 24 combinations.
- A 5 bit word can have 25=32 combinations.
- We therefore have 32 - 26 = 16 extra words.
- Some of the extra words are used for control/signalling purposes.

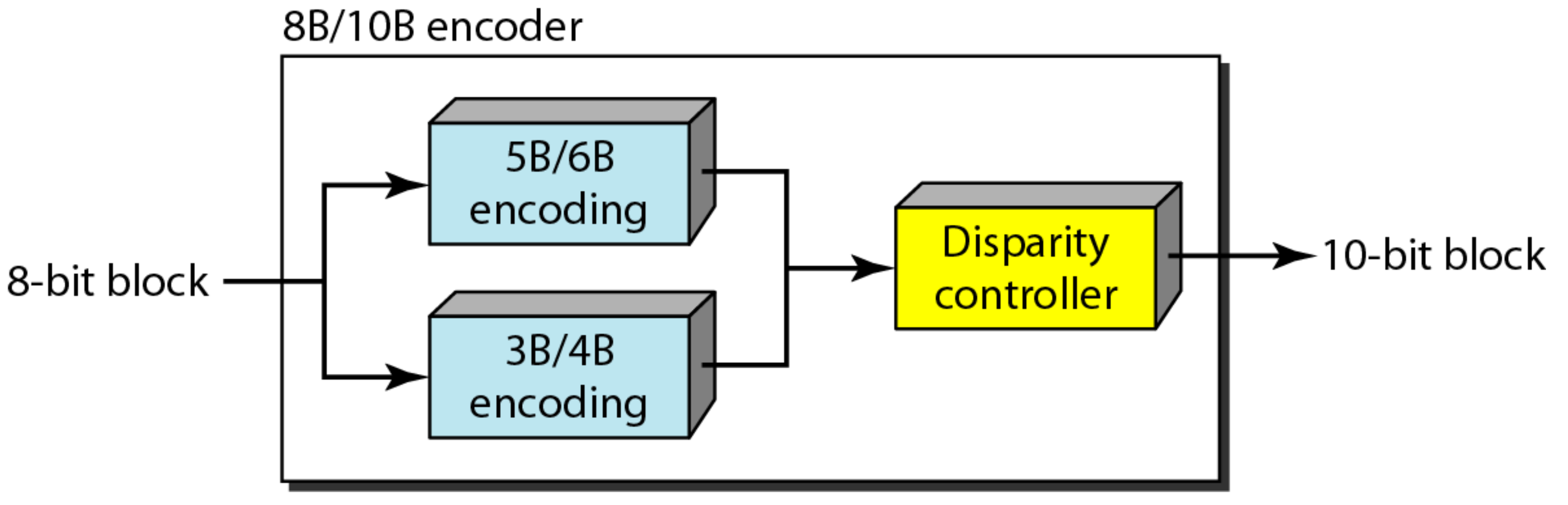
Example 4.5

We need to send data at a 1-Mbps rate. What is the minimum required bandwidth, using a combination of 4B/5B and NRZ-I or Manchester coding?

Solution

First 4B/5B block coding increases the bit rate to 1.25 Mbps. The minimum bandwidth using NRZ-I is N/2 or 625 kHz. The Manchester scheme needs a minimum bandwidth of 1.25 MHz. The first choice needs a lower bandwidth, but has a DC component problem; the second choice needs a higher bandwidth, but does not have a DC component problem.

Figure 4.17 *8B/10B block encoding*



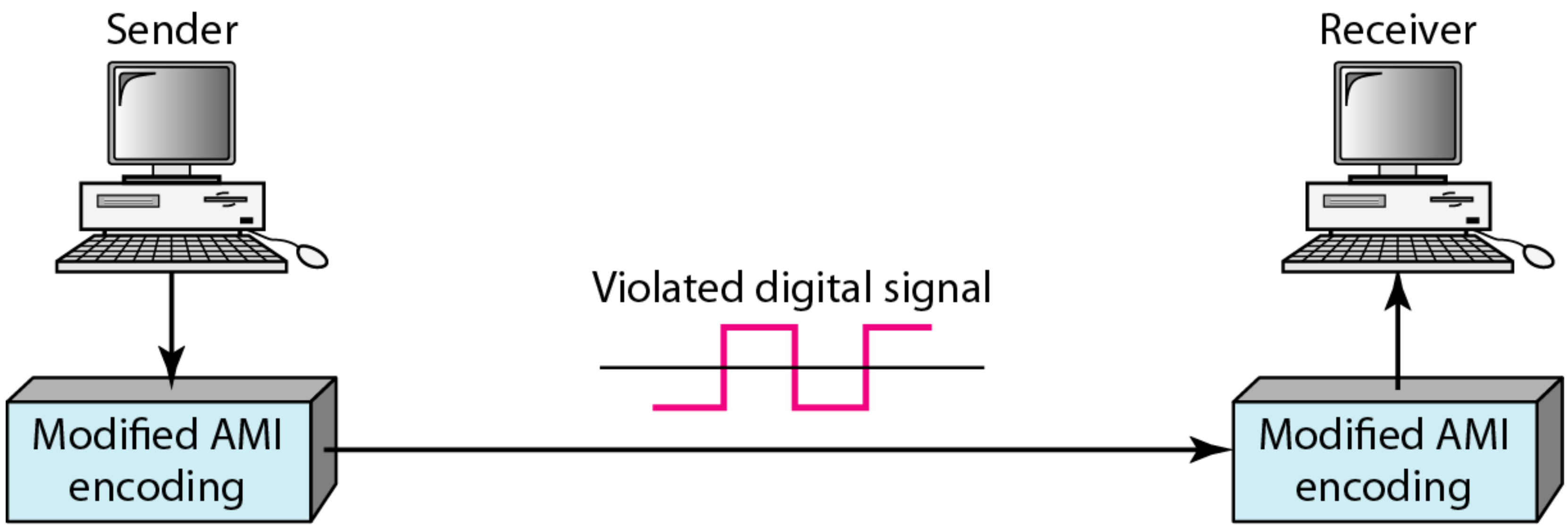
More bits - better error detection

- The 8B10B block code adds more redundant bits and can thereby choose code words that would prevent a long run of a voltage level that would cause DC components.

Scrambling

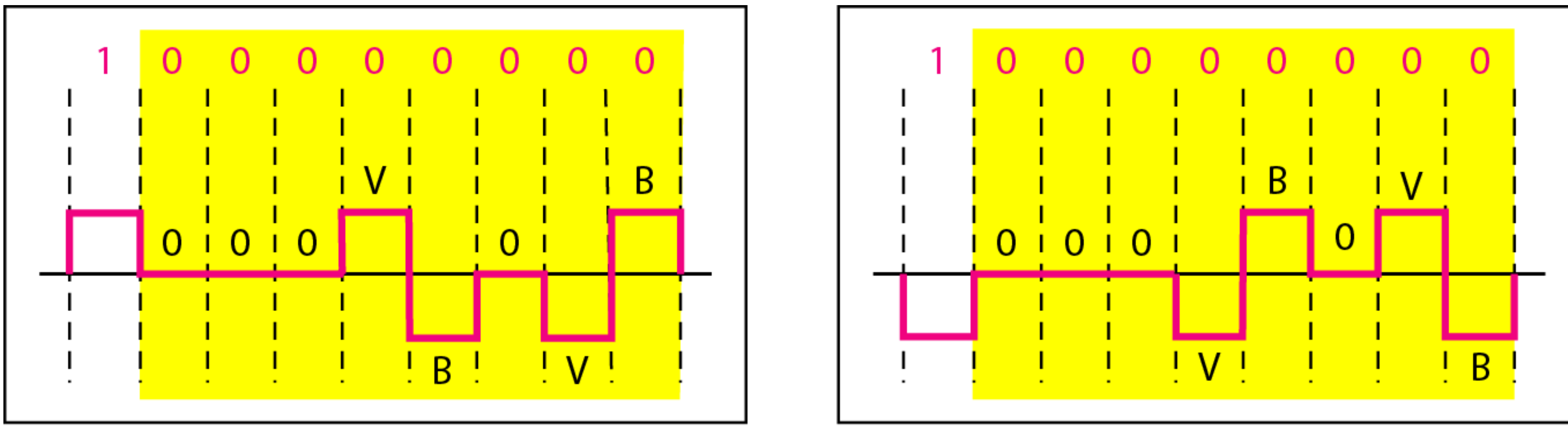
- The best code is one that does not increase the bandwidth for synchronization and has no DC components.
- Scrambling is a technique used to create a sequence of bits that has the required c/c’s for transmission - self clocking, no low frequencies, no wide bandwidth.
- It is implemented at the same time as encoding, the bit stream is created on the fly.
- It replaces ‘unfriendly’ runs of bits with a violation code that is easy to recognize and removes the unfriendly c/c.

Figure 4.18 *AMI used with scrambling*



For example: B8ZS substitutes eight consecutive zeros with 000VB0VB.
The V stands for violation, it violates the line encoding rule
B stands for bipolar, it implements the bipolar line encoding rule

Figure 4.19 *Two cases of B8ZS scrambling technique*



a. Previous level is positive.

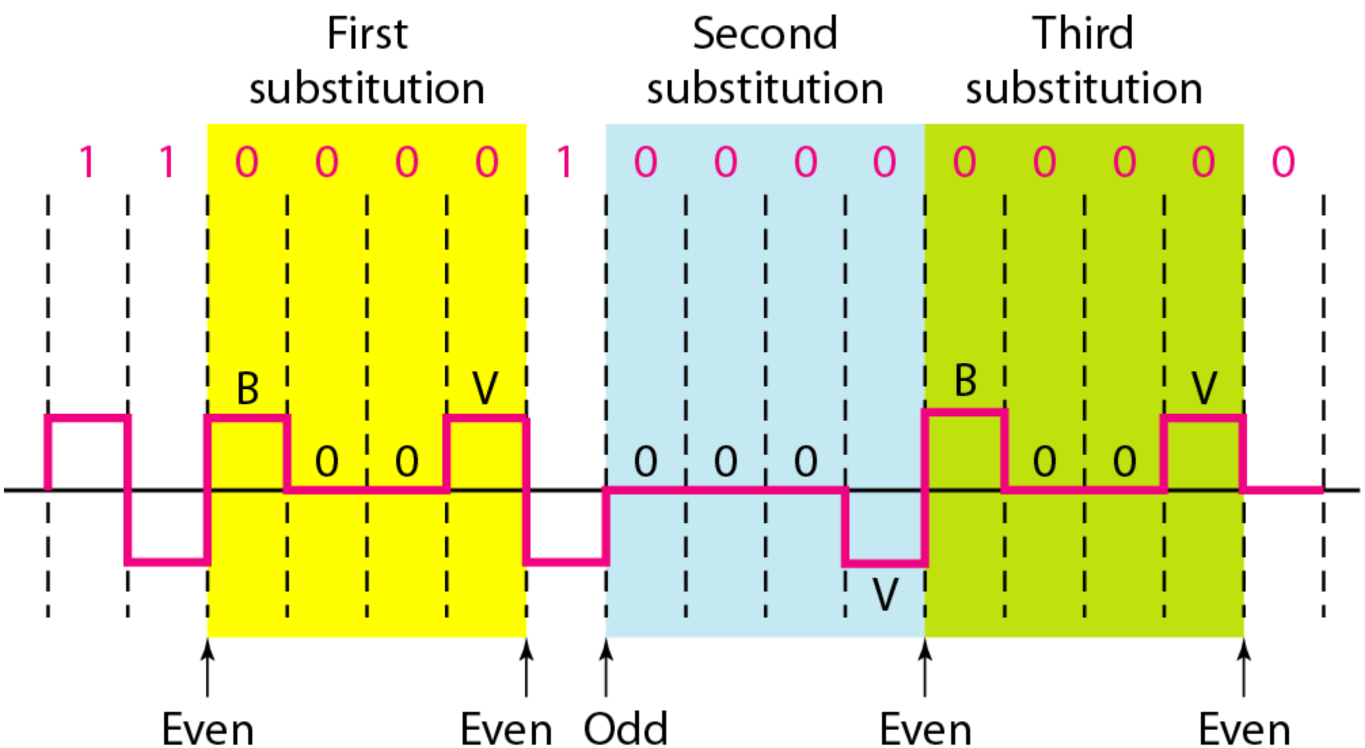
b. Previous level is negative.

HDB3 substitutes four consecutive zeros with 000V or B00V depending on the number of nonzero pulses after the last substitution.

If # of non zero pulses is even the substitution is B00V to make total # of non zero pulse even.

If # of non zero pulses is odd the substitution is 000V to make total # of non zero pulses even.

Figure 4.20 *Different situations in HDB3 scrambling technique*



Chapter 4

Digital Transmission

4-2 ANALOG-TO-DIGITAL CONVERSION

A digital signal is superior to an analog signal because it is more robust to noise and can easily be recovered, corrected and amplified. For this reason, the tendency today is to change an analog signal to digital data. In this section we describe two techniques, pulse code modulation and delta modulation.

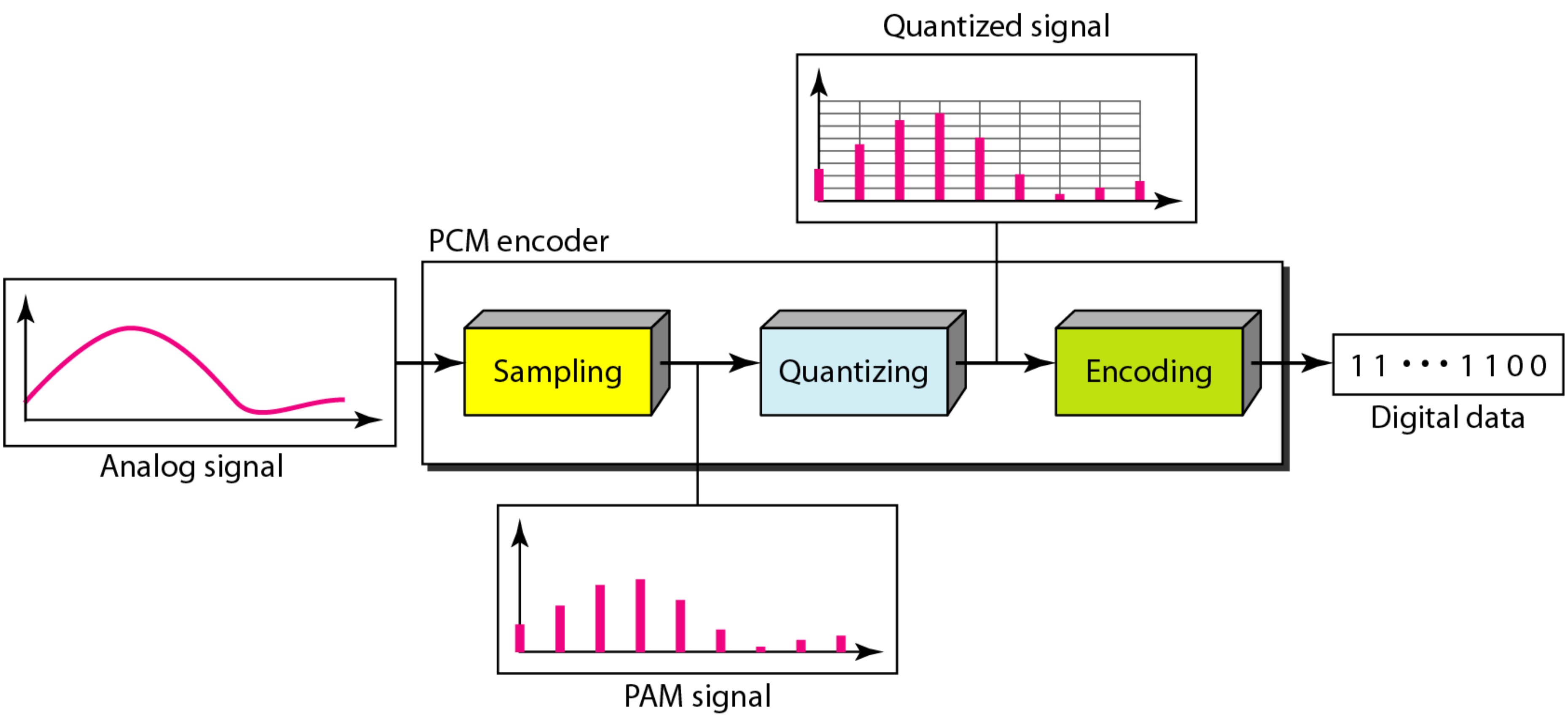
Topics discussed in this section:

- **Pulse Code Modulation (PCM)**
- **Delta Modulation (DM)**

PCM

- PCM consists of three steps to digitize an analog signal:
  1. Sampling
  2. Quantization
  3. Binary encoding
- Before we sample, we have to filter the signal to limit the maximum frequency of the signal as it affects the sampling rate.
- Filtering should ensure that we do not distort the signal, ie remove high frequency components that affect the signal shape.

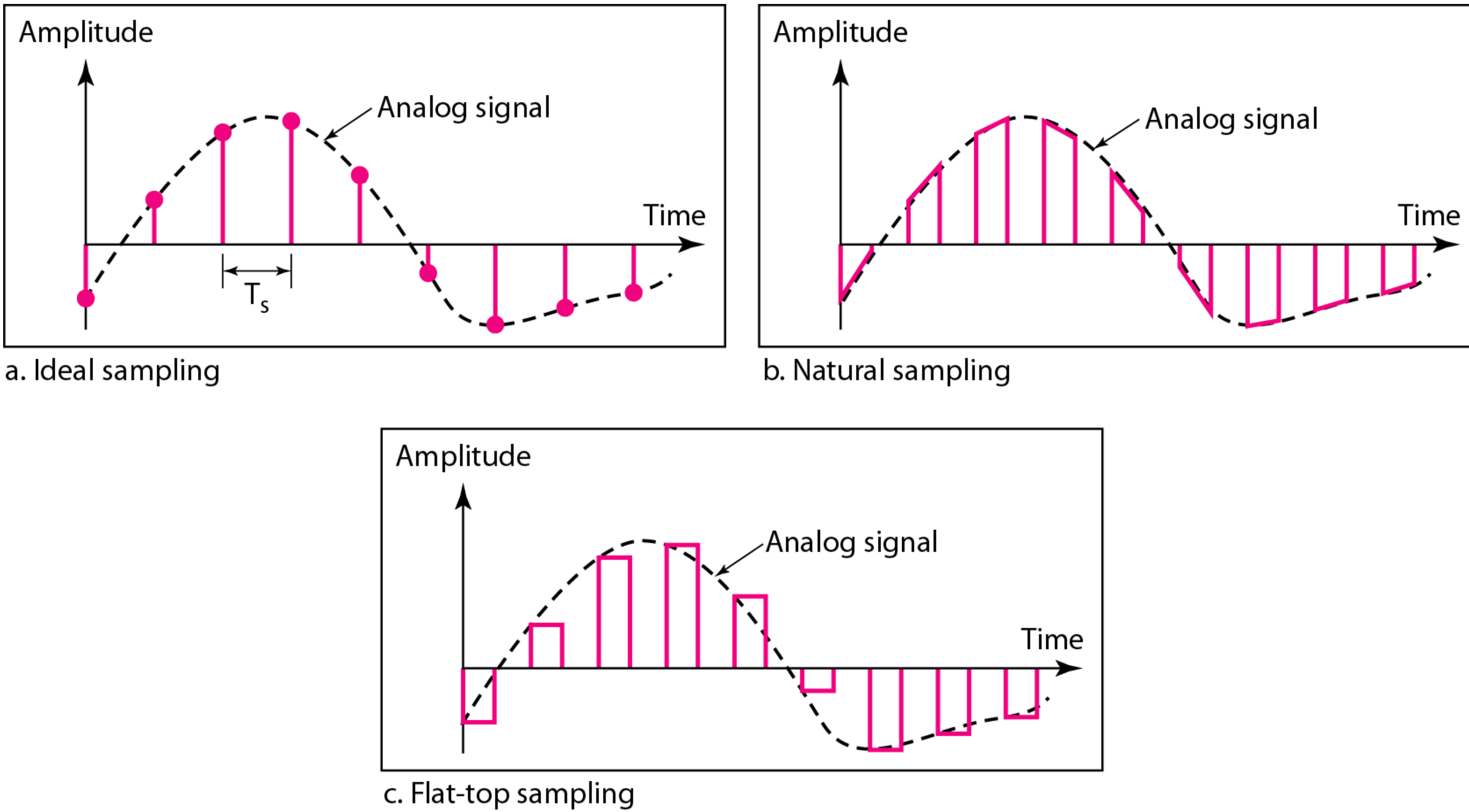
Figure 4.21 *Components of PCM encoder*



Sampling

- Analog signal is sampled every T_S secs.
- T_s is referred to as the sampling interval.
- $f_s = 1/T_s$ is called the sampling rate or sampling frequency.
- There are 3 sampling methods:
  - Ideal - an impulse at each sampling instant
  - Natural - a pulse of short width with varying amplitude
  - Flattop - sample and hold, like natural but with single amplitude value
- The process is referred to as pulse amplitude modulation PAM and the outcome is a signal with analog (non integer) values

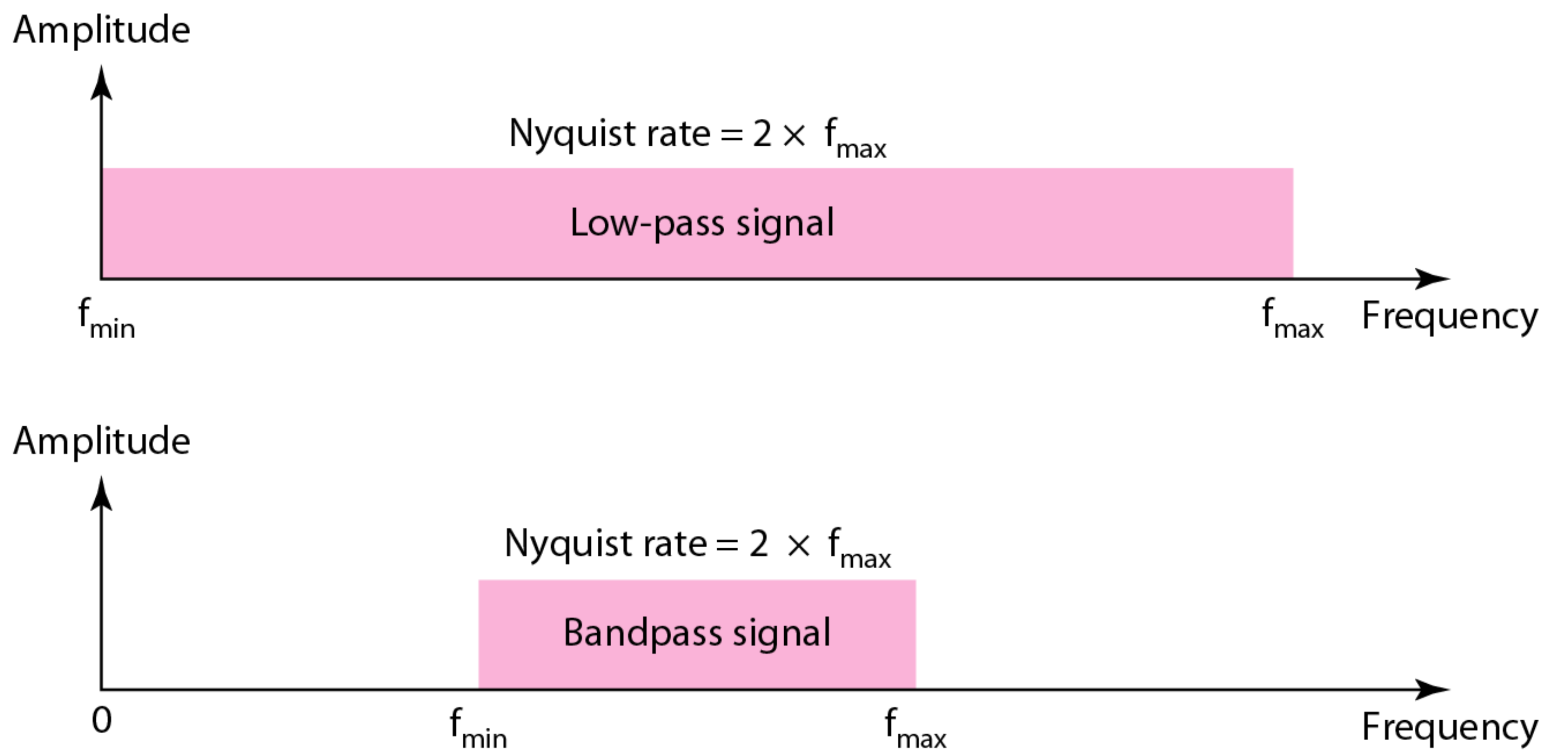
Figure 4.22 *Three different sampling methods for PCM*



Note

According to the Nyquist theorem, the sampling rate must be at least 2 times the highest frequency contained in the signal.

Figure 4.23 *Nyquist sampling rate for low-pass and bandpass signals*

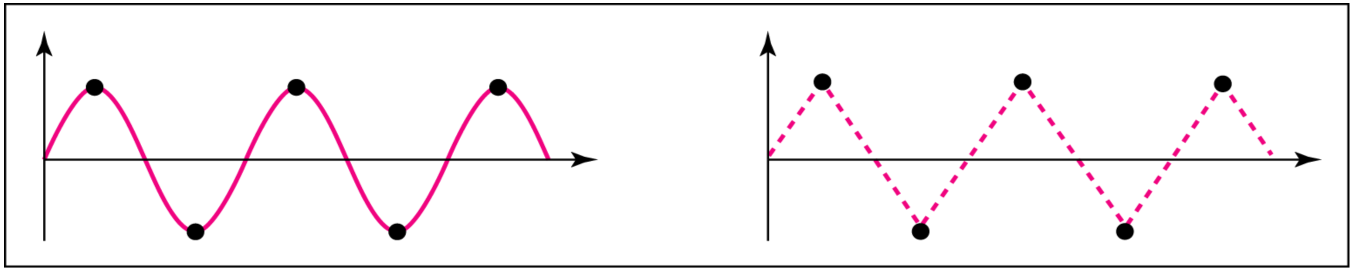


Example 4.6

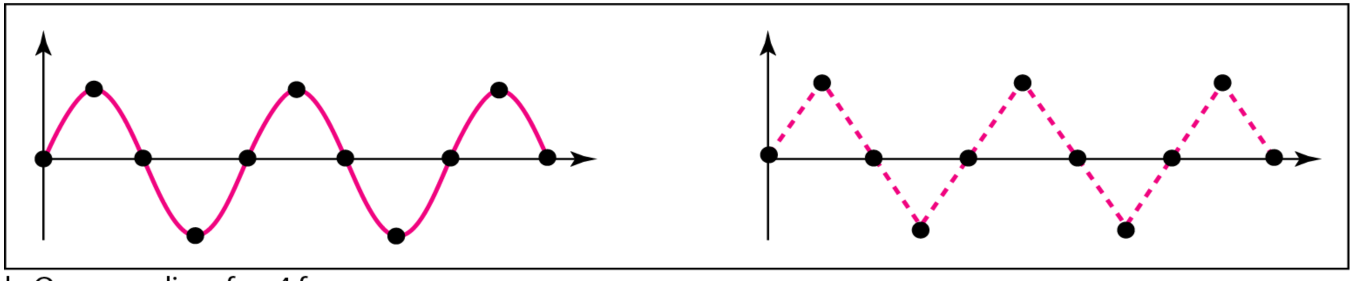
For an intuitive example of the Nyquist theorem, let us sample a simple sine wave at three sampling rates: $f_s = 4f$ (2 times the Nyquist rate), $f_s = 2f$ (Nyquist rate), and $f_s = f$ (one-half the Nyquist rate). Figure 4.24 shows the sampling and the subsequent recovery of the signal.

It can be seen that sampling at the Nyquist rate can create a good approximation of the original sine wave (part a). Oversampling in part b can also create the same approximation, but it is redundant and unnecessary. Sampling below the Nyquist rate (part c) does not produce a signal that looks like the original sine wave.

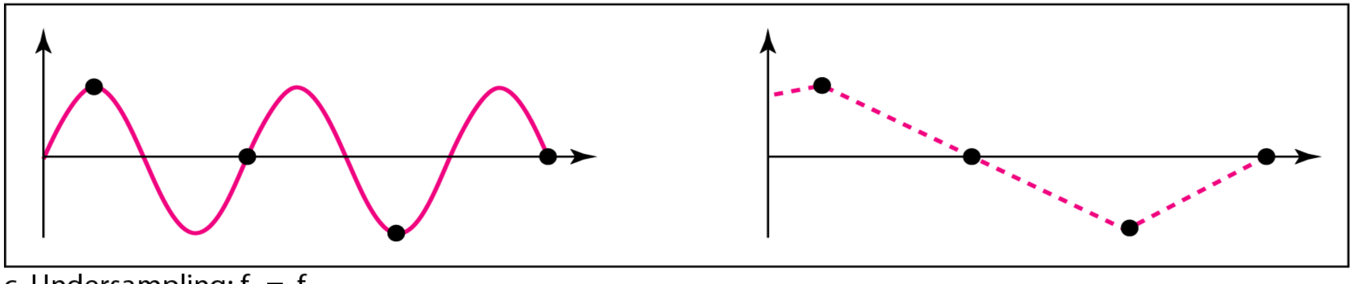
Figure 4.24 *Recovery of a sampled sine wave for different sampling rates*



a. Nyquist rate sampling: $f_s = 2\,f$



b. Oversampling: $f_s = 4\,f$

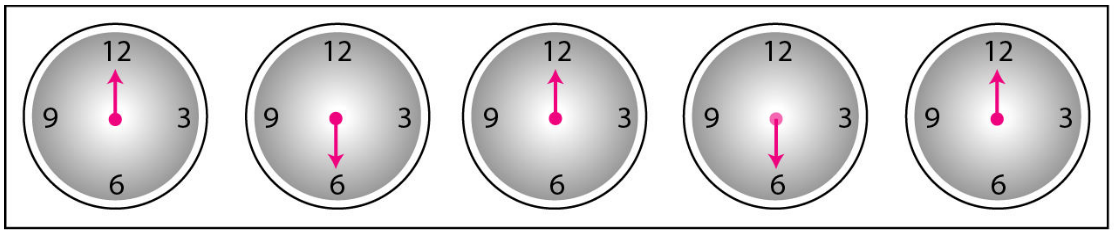


c. Undersampling: $f_s = f$

Example 4.7

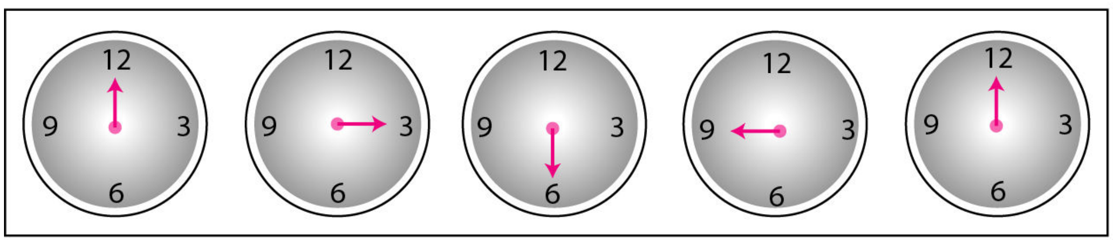
Consider the revolution of a hand of a clock. The second hand of a clock has a period of 60 s. According to the Nyquist theorem, we need to sample the hand every 30 s ($T_s = T$ or $f_s = 2f$). In Figure 4.25a, the sample points, in order, are 12, 6, 12, 6, 12, and 6. The receiver of the samples cannot tell if the clock is moving forward or backward. In part b, we sample at double the Nyquist rate (every 15 s). The sample points are 12, 3, 6, 9, and 12. The clock is moving forward. In part c, we sample below the Nyquist rate ($T_s = T$ or $f_s = f$). The sample points are 12, 9, 6, 3, and 12. Although the clock is moving forward, the receiver thinks that the clock is moving backward.

Figure 4.25 *Sampling of a clock with only one hand*



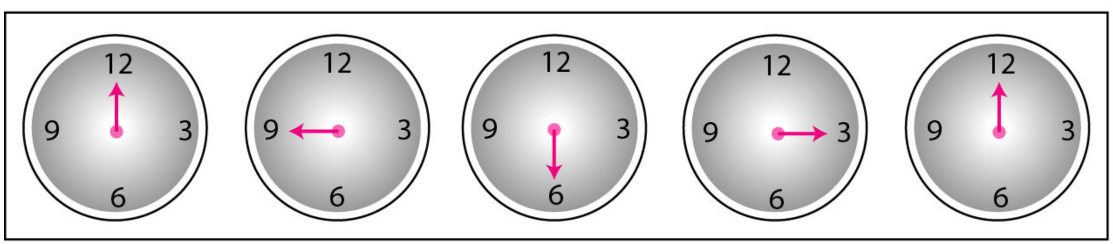
Samples can mean that the clock is moving either forward or backward. (12-6-12-6-12)

a. Sampling at Nyquist rate: $T_s = T\frac{1}{2}$



Samples show clock is moving forward. (12-3-6-9-12)

b. Oversampling (above Nyquist rate): $T_s = T\frac{1}{4}$



Samples show clock is moving backward. (12-9-6-3-12)

c. Undersampling (below Nyquist rate): $T_s = T\frac{3}{4}$

Example 4.8

An example related to Example 4.7 is the seemingly backward rotation of the wheels of a forward-moving car in a movie. This can be explained by under-sampling. A movie is filmed at 24 frames per second. If a wheel is rotating more than 12 times per second, the under-sampling creates the impression of a backward rotation.

Example 4.9

Telephone companies digitize voice by assuming a maximum frequency of 4000 Hz. The sampling rate therefore is 8000 samples per second.

Example 4.10

A complex low-pass signal has a bandwidth of 200 kHz. What is the minimum sampling rate for this signal?

Solution

The bandwidth of a low-pass signal is between 0 and f, where f is the maximum frequency in the signal. Therefore, we can sample this signal at 2 times the highest frequency (200 kHz). The sampling rate is therefore 400,000 samples per second.

Example 4.11

A complex bandpass signal has a bandwidth of 200 kHz. What is the minimum sampling rate for this signal?

Solution

We cannot find the minimum sampling rate in this case because we do not know where the bandwidth starts or ends. We do not know the maximum frequency in the signal.

Quantization

- Sampling results in a series of pulses of varying amplitude values ranging between two limits: a min and a max.
- The amplitude values are infinite between the two limits.
- We need to map the *infinite* amplitude values onto a finite set of known values.
- This is achieved by dividing the distance between min and max into L zones, each of height Δ.

$$\Delta = (\max - \min)/L$$

Quantization Levels

- The midpoint of each zone is assigned a value from 0 to L-1 (resulting in L values)
- Each sample falling in a zone is then approximated to the value of the midpoint.

Quantization Zones

- Assume we have a voltage signal with amplitutes V_{min}=-20V and V_{max}=+20V.
- We want to use L=8 quantization levels.
- Zone width Δ = (20 - -20)/8 = 5
- The 8 zones are: -20 to -15, -15 to -10, -10 to -5, -5 to 0, 0 to +5, +5 to +10, +10 to +15, +15 to +20
- The midpoints are: -17.5, -12.5, -7.5, -2.5, 2.5, 7.5, 12.5, 17.5

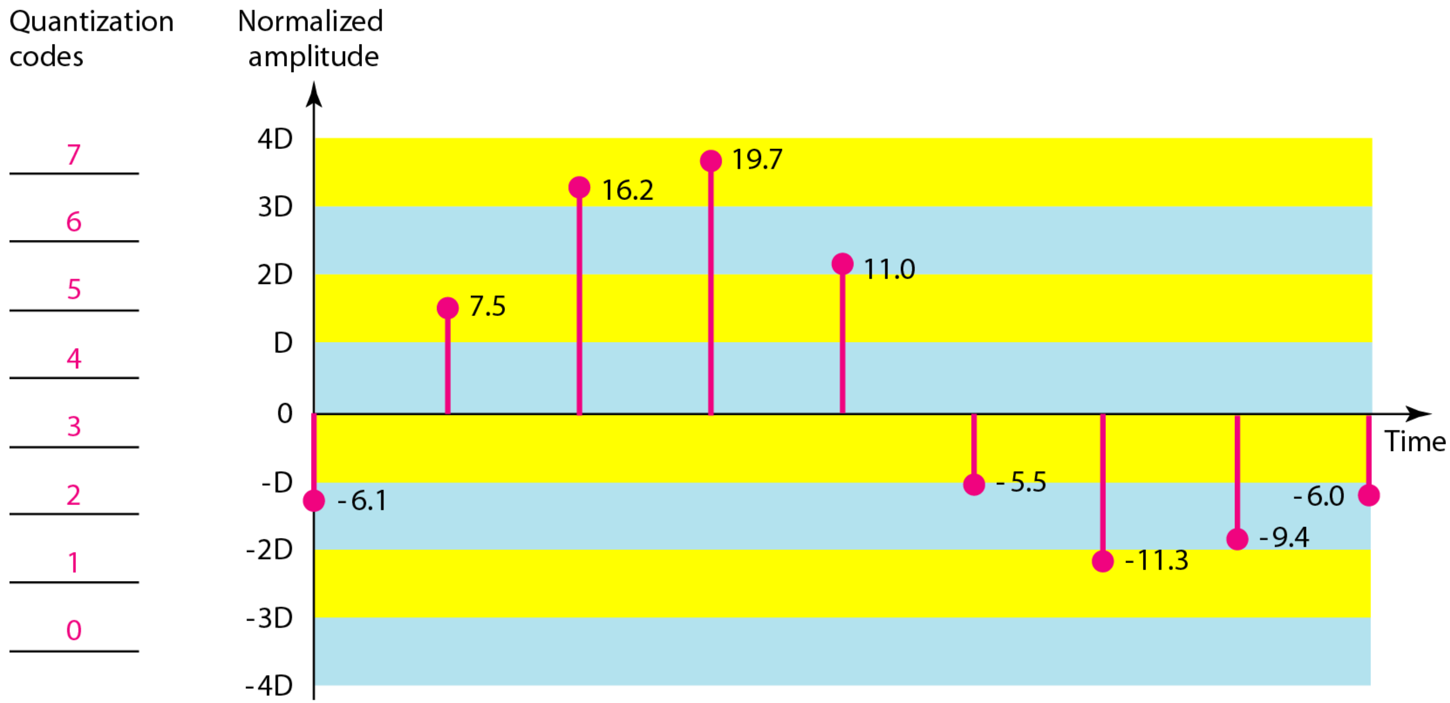
Assigning Codes to Zones

- Each zone is then assigned a binary code.
- The number of bits required to encode the zones, or the number of bits per sample as it is commonly referred to, is obtained as follows:

$$n_b = \log_2 L$$

- Given our example, $n_b = 3$
- The 8 zone (or level) codes are therefore: 000, 001, 010, 011, 100, 101, 110, and 111
- Assigning codes to zones:
  - 000 will refer to zone -20 to -15
  - 001 to zone -15 to -10, etc.

Figure 4.26 *Quantization and encoding of a sampled signal*



Normalized PAM values	-1.22	1.50	3.24	3.94	2.20	-1.10	-2.26	-1.88	-1.20
Normalized quantized values	-1.50	1.50	3.50	3.50	2.50	-1.50	-2.50	-1.50	-1.50
Normalized error	-0.38	0	+0.26	-0.44	+0.30	-0.40	-0.24	+0.38	-0.30
Quantization code	2	5	7	7	6	2	1	2	2
Encoded words	010	101	111	111	110	010	001	010	010

Quantization Error

- When a signal is quantized, we introduce an error - the coded signal is an approximation of the actual amplitude value.
- The difference between actual and coded value (midpoint) is referred to as the quantization error.
- The more zones, the smaller Δ which results in smaller errors.
- BUT, the more zones the more bits required to encode the samples -> higher bit rate

Quantization Error and SN_QR

- Signals with lower amplitude values will suffer more from quantization error as the error range: $\Delta/2$, is fixed for all signal levels.
- Non linear quantization is used to alleviate this problem. Goal is to keep SN_QR fixed for all sample values.
- Two approaches:
  - The quantization levels follow a logarithmic curve. Smaller Δ's at lower amplitudes and larger Δ's at higher amplitudes.
  - Companding: The sample values are compressed at the sender into logarithmic zones, and then expanded at the receiver. The zones are fixed in height.

Bit rate and bandwidth requirements of PCM

- The bit rate of a PCM signal can be calculated form the number of bits per sample x the sampling rate

$$\text{Bit rate} = n_b \times f_s$$

- The bandwidth required to transmit this signal depends on the type of line encoding used. Refer to previous section for discussion and formulas.
- A digitized signal will always need more bandwidth than the original analog signal. Price we pay for robustness and other features of digital transmission.

Example 4.14

We want to digitize the human voice. What is the bit rate, assuming 8 bits per sample?

Solution

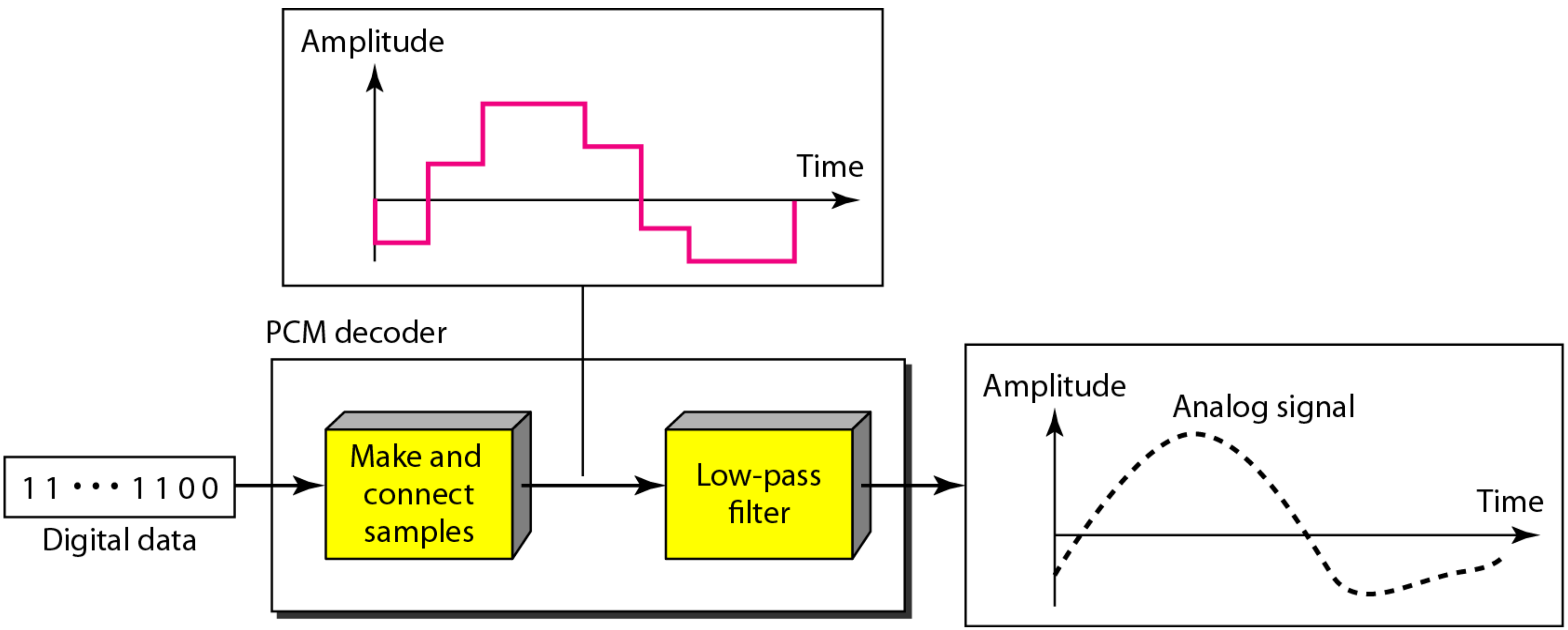
The human voice normally contains frequencies from 0 to 4000 Hz. So the sampling rate and bit rate are calculated as follows:

$$\text{Sampling rate} = 4000 \times 2 = 8000 \text{ samples/s}$$
$$\text{Bit rate} = 8000 \times 8 = 64{,}000 \text{ bps} = 64 \text{ kbps}$$

PCM Decoder

- To recover an analog signal from a digitized signal we follow the following steps:
  - We use a hold circuit that holds the amplitude value of a pulse till the next pulse arrives.
  - We pass this signal through a low pass filter with a cutoff frequency that is equal to the highest frequency in the pre-sampled signal.
- The higher the value of L, the less distorted a signal is recovered.

Figure 4.27 *Components of a PCM decoder*



Example 4.15

We have a low-pass analog signal of 4 kHz. If we send the analog signal, we need a channel with a minimum bandwidth of 4 kHz. If we digitize the signal and send 8 bits per sample, we need a channel with a minimum bandwidth of 8 × 4 kHz = 32 kHz.

Delta Modulation

- This scheme sends only the difference between pulses, if the pulse at time t_{n+1} is higher in amplitude value than the pulse at time t_n, then a single bit, say a "1", is used to indicate the positive value.
- If the pulse is lower in value, resulting in a negative value, a "0" is used.
- This scheme works well for small changes in signal values between samples.
- If changes in amplitude are large, this will result in large errors.

Figure 4.28 *The process of delta modulation*

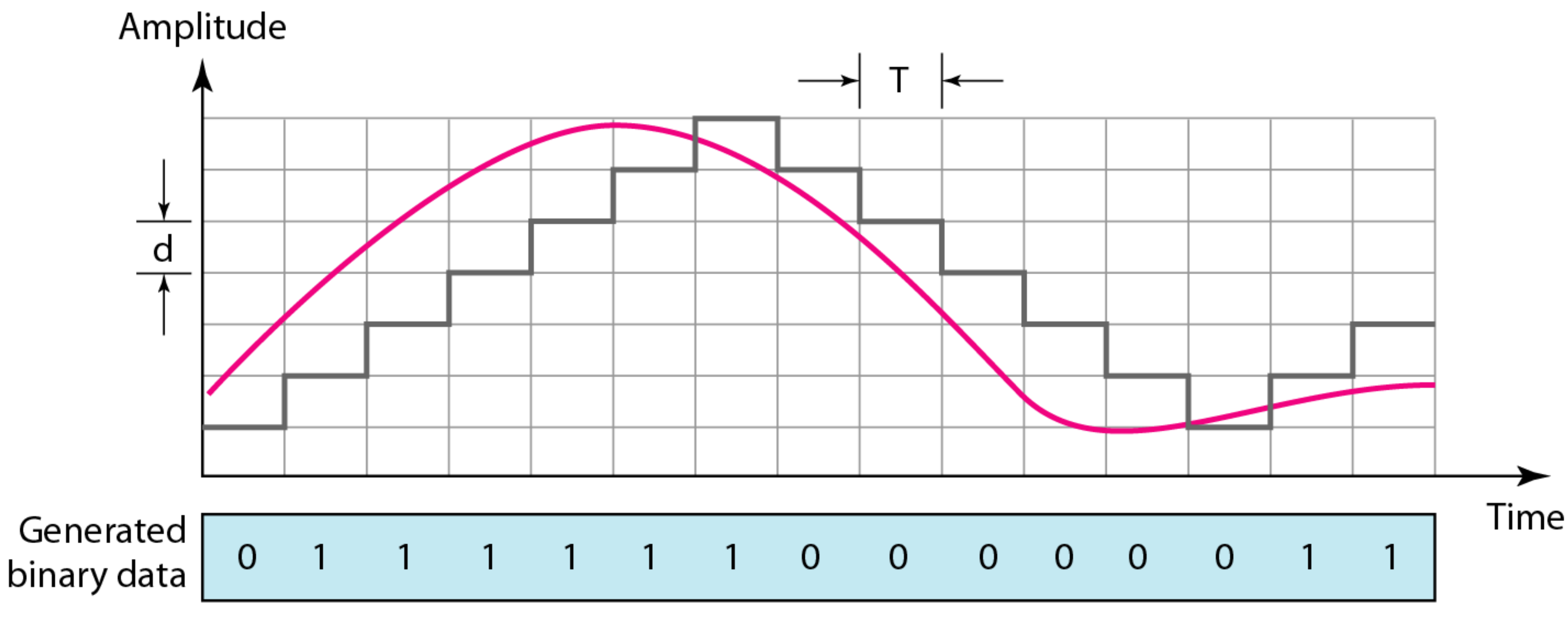


Figure 4.29 *Delta modulation components*

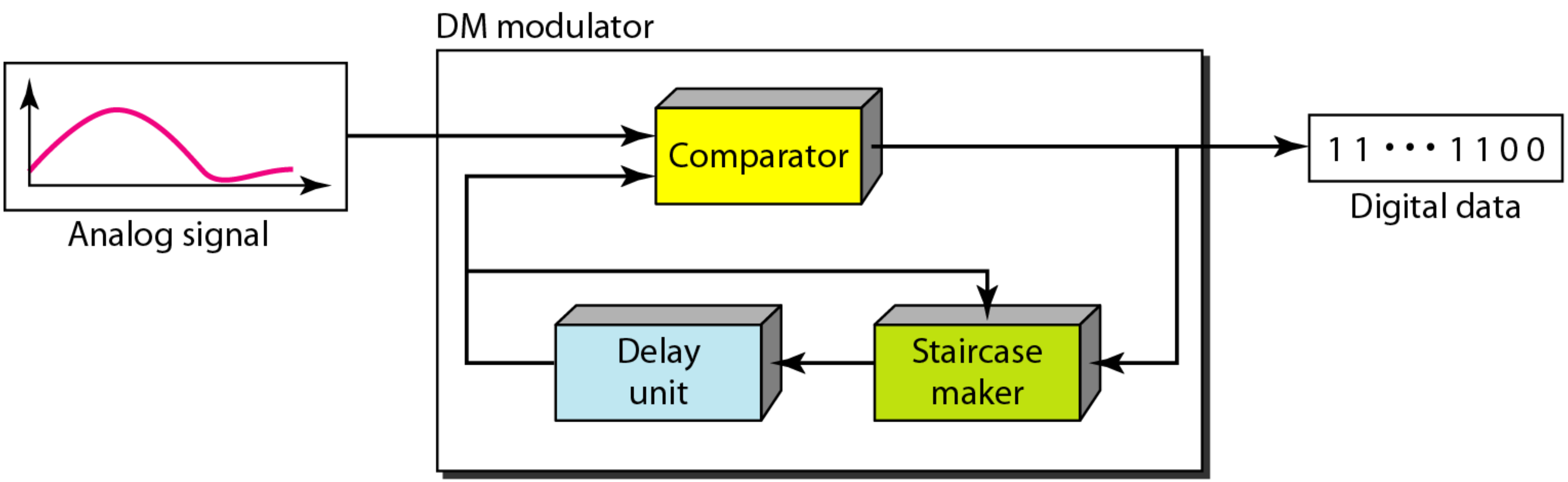
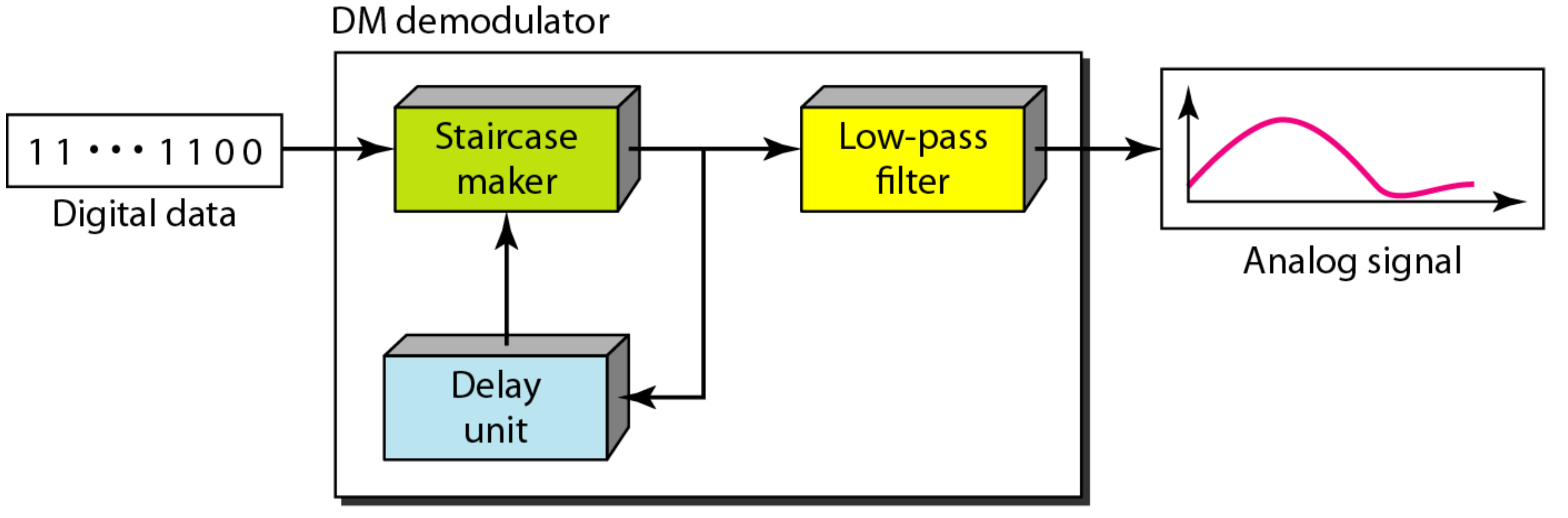


Figure 4.30 *Delta demodulation components*



Delta PCM (DPCM)

- Instead of using one bit to indicate positive and negative differences, we can use more bits -> quantization of the difference.
- Each bit code is used to represent the value of the difference.
- The more bits the more levels -> the higher the accuracy.

4-3 TRANSMISSION MODES

The transmission of binary data across a link can be accomplished in either parallel or serial mode. In parallel mode, multiple bits are sent with each clock tick. In serial mode, 1 bit is sent with each clock tick. While there is only one way to send parallel data, there are three subclasses of serial transmission: asynchronous, synchronous, and isochronous.

Topics discussed in this section:

- **Parallel Transmission**
- **Serial Transmission**

Figure 4.31 *Data transmission and modes*

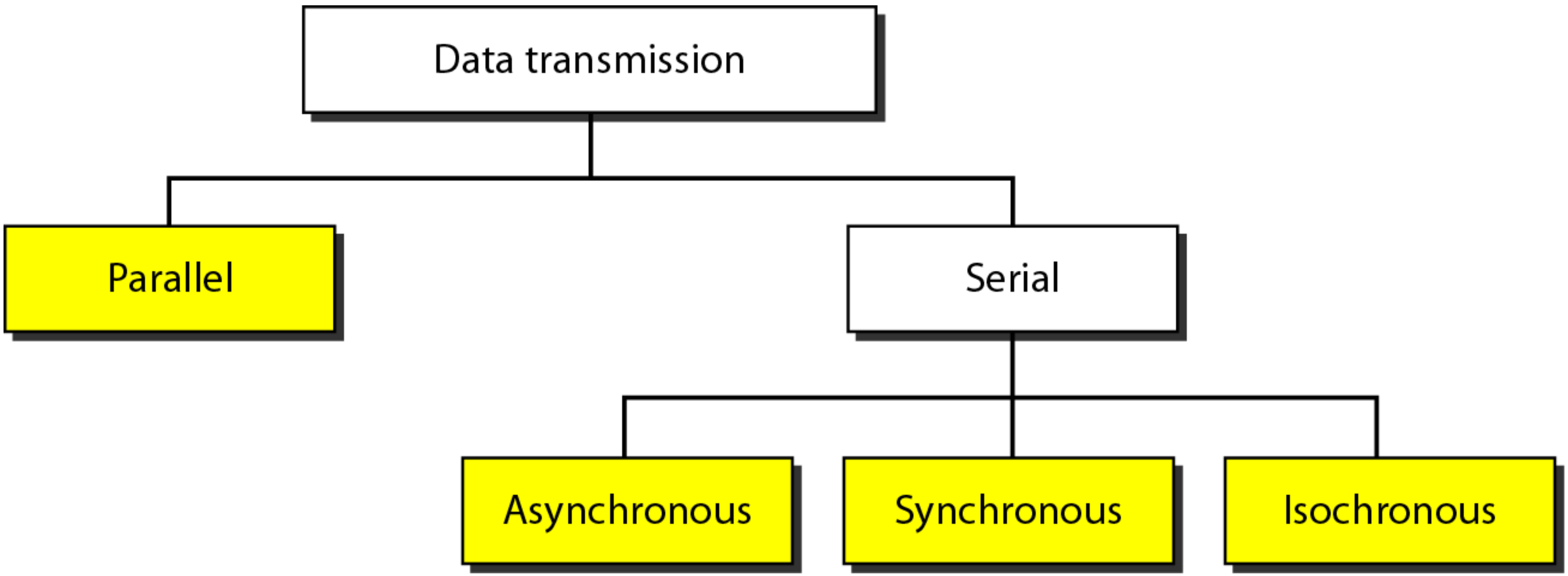


Figure 4.32 *Parallel transmission*

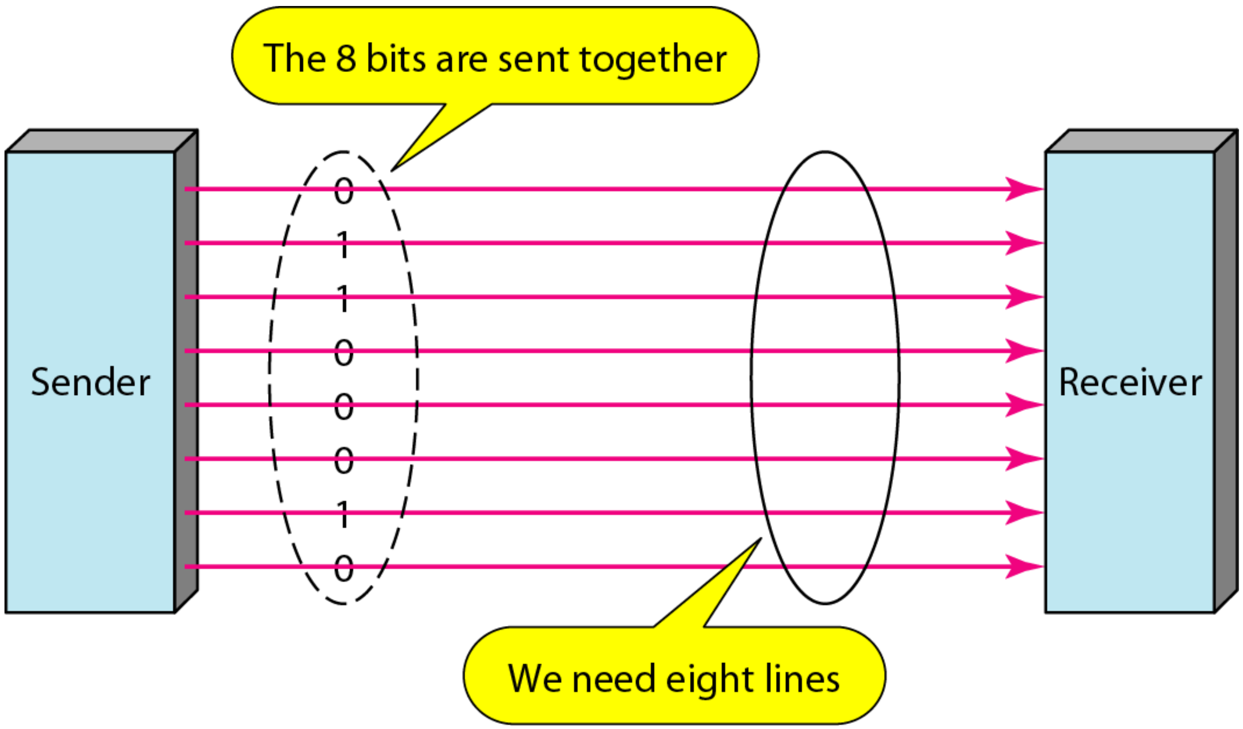
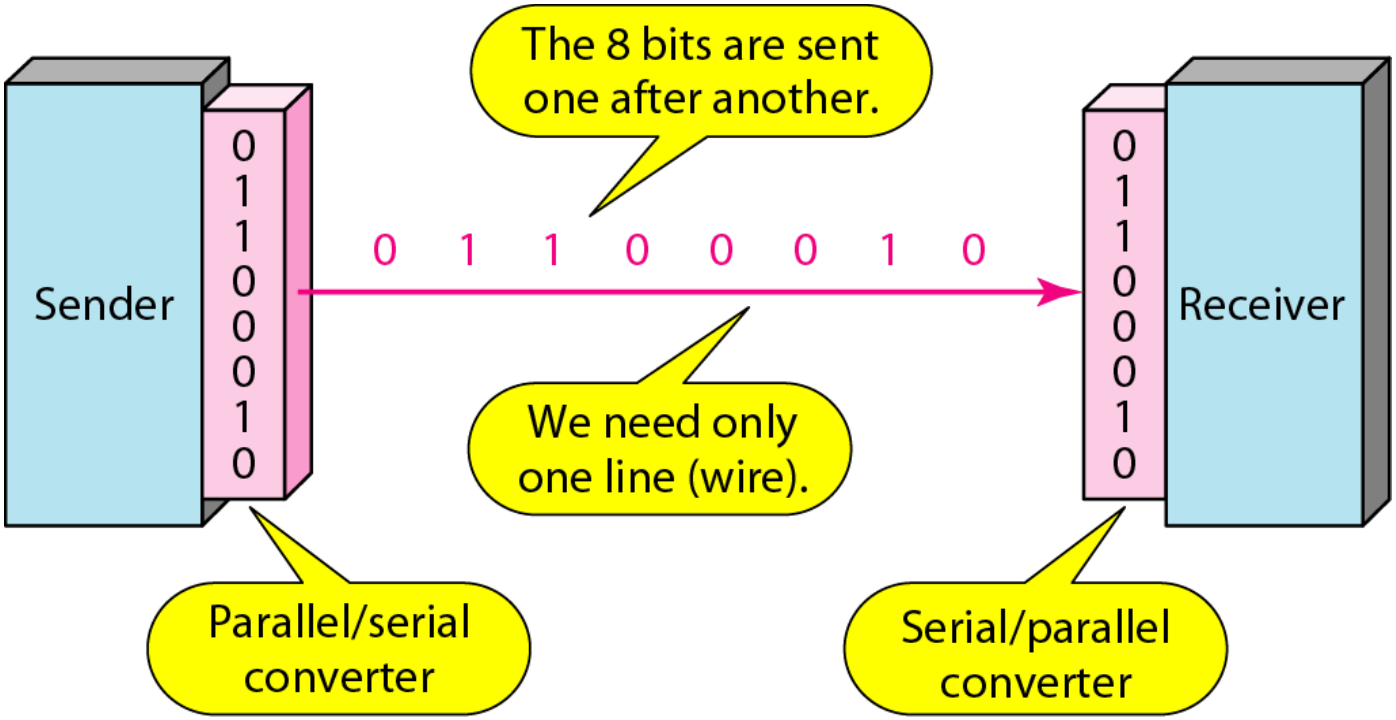


Figure 4.33 *Serial transmission*



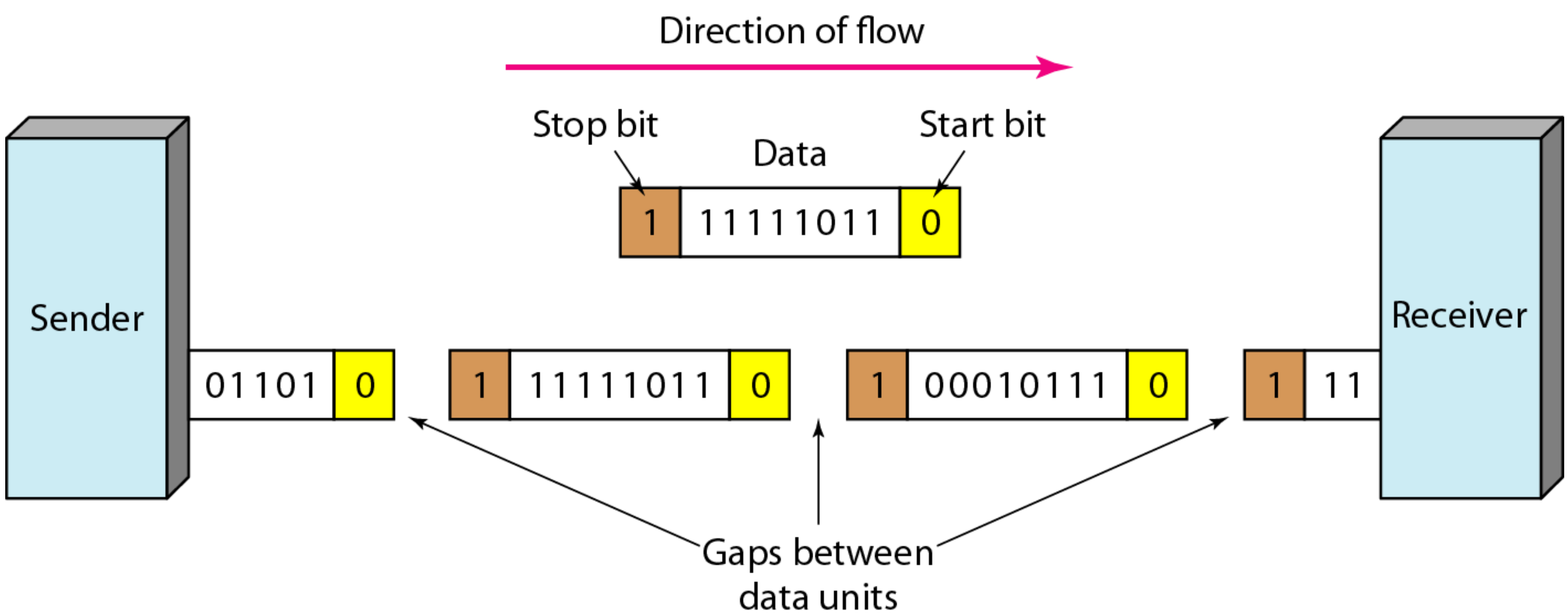
Note

In asynchronous transmission, we send 1 start bit (0) at the beginning and 1 or more stop bits (1s) at the end of each byte. There may be a gap between each byte.

Note

Asynchronous here means "asynchronous at the byte level," but the bits are still synchronized; their durations are the same.

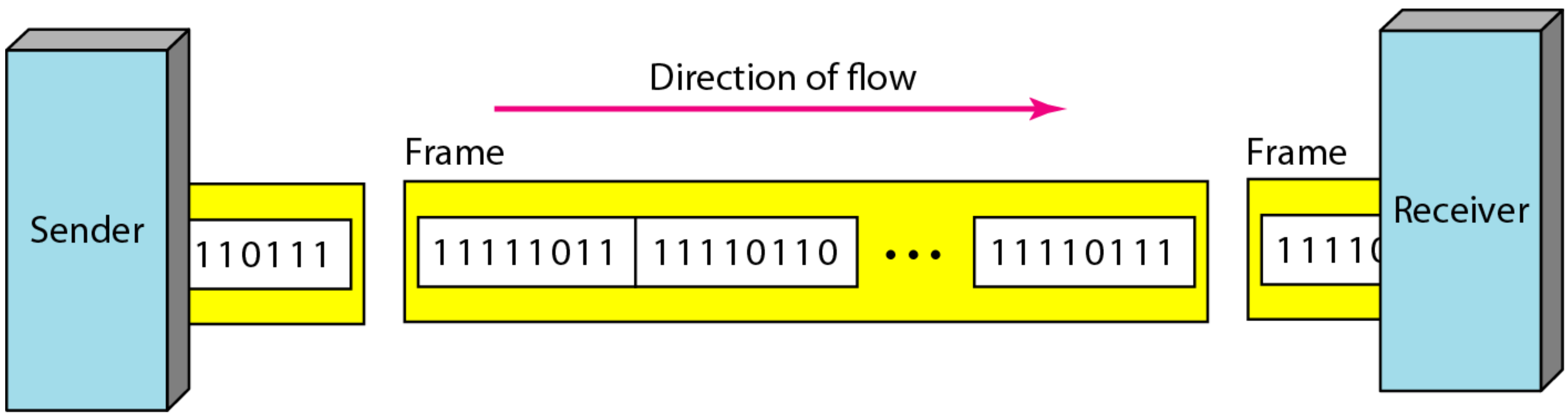
Figure 4.34 *Asynchronous transmission*



Note

In synchronous transmission, we send bits one after another without start or stop bits or gaps. It is the responsibility of the receiver to group the bits. The bits are usually sent as bytes and many bytes are grouped in a frame. A frame is identified with a start and an end byte.

Figure 4.35 *Synchronous transmission*



Isochronous

- In isochronous transmission we cannot have uneven gaps between frames.
- Transmission of bits is fixed with equal gaps.

Chapter 5

Analog Transmission

5-1 DIGITAL-TO-ANALOG CONVERSION

Digital-to-analog ***conversion is the process of changing one of the characteristics of an analog signal based on the information in digital data.***

Topics discussed in this section:

- **Aspects of Digital-to-Analog Conversion**
- **Amplitude Shift Keying**
- **Frequency Shift Keying**
- **Phase Shift Keying**
- **Quadrature Amplitude Modulation**

Digital to Analog Conversion

- Digital data needs to be carried on an analog signal.
- A carrier signal (frequency f_c) performs the function of transporting the digital data in an analog waveform.
- The analog carrier signal is manipulated to uniquely identify the digital data being carried.

Figure 5.1 *Digital-to-analog conversion*

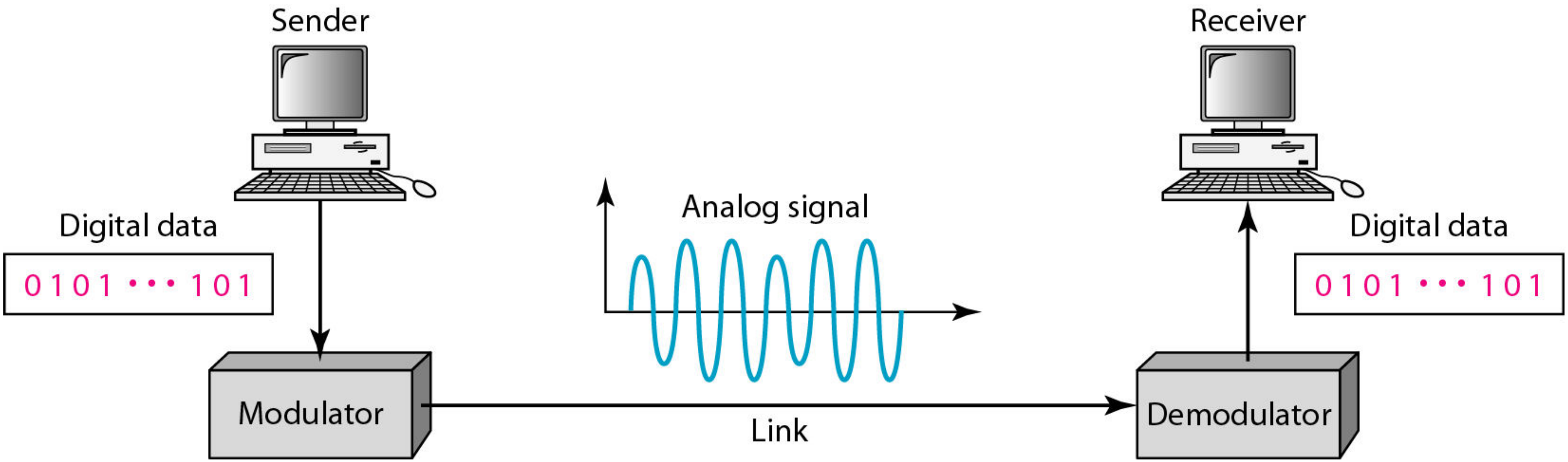
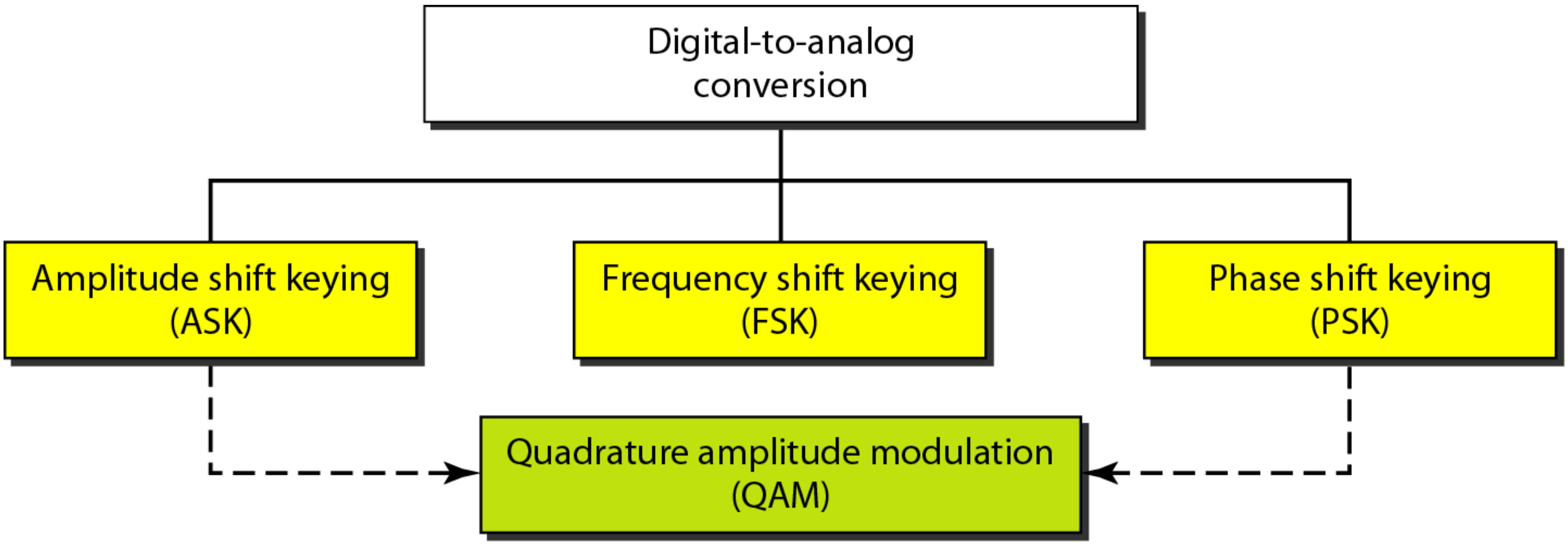


Figure 5.2 *Types of digital-to-analog conversion*



Note

Bit rate, N, is the number of bits per second (bps). Baud rate is the number of signal
elements per second (bauds).
In the analog transmission of digital data, the signal or baud rate is less than or equal to the bit rate.
S=Nx1/r bauds
Where r is the number of data bits per signal element.

Example 5.1

An analog signal carries 4 bits per signal element. If 1000 signal elements are sent per second, find the bit rate.

Solution

In this case, r = 4, S = 1000, and N is unknown. We can find the value of N from

$$S = N \times \frac{1}{r} \qquad \text{or} \qquad N = S \times r = 1000 \times 4 = 4000 \text{ bps}$$

Example 5.2

An analog signal has a bit rate of 8000 bps and a baud rate of 1000 baud. How many data elements are carried by each signal element? How many signal elements do we need?

Solution

In this example, S = 1000, N = 8000, and r and L are unknown. We find first the value of r and then the value of L.

$$S = N \times \frac{1}{r} \quad \rightarrow \quad r = \frac{N}{S} = \frac{8000}{1000} = 8 \text{ bits/baud}$$

$$r = \log_2 L \quad \rightarrow \quad L = 2^r = 2^8 = 256$$

Amplitude Shift Keying (ASK)

- ASK is implemented by changing the amplitude of a carrier signal to reflect amplitude levels in the digital signal.
- For example: a digital "1" could not affect the signal, whereas a digital "0" would, by making it zero.
- The line encoding will determine the values of the analog waveform to reflect the digital data being carried.

Bandwidth of ASK

- The bandwidth B of ASK is proportional to the signal rate S.

$$B = (1+d)S$$

- "d" is due to modulation and filtering, lies between 0 and 1.

Figure 5.3 *Binary amplitude shift keying*

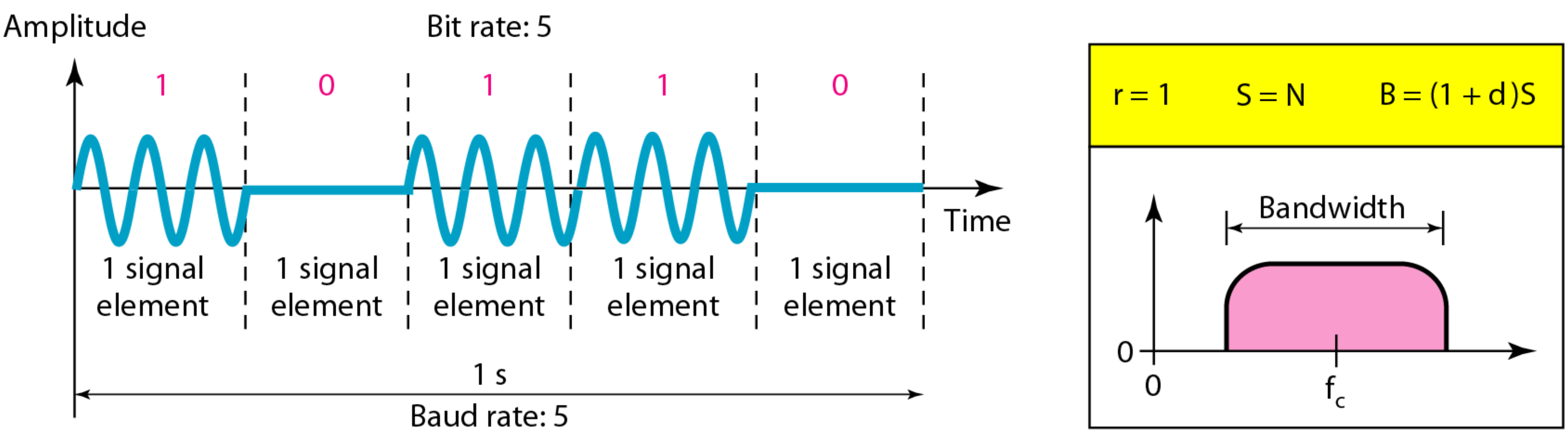
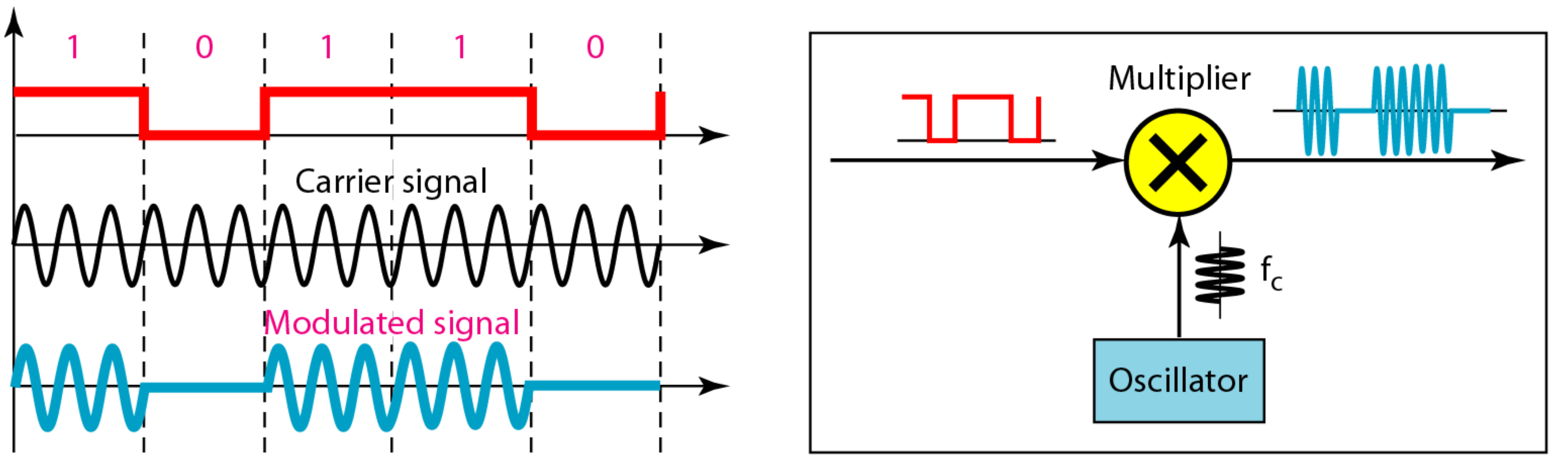


Figure 5.4 *Implementation of binary ASK*



Example 5.3

We have an available bandwidth of 100 kHz which spans from 200 to 300 kHz. What are the carrier frequency and the bit rate if we modulated our data by using ASK with d = 1?

Solution

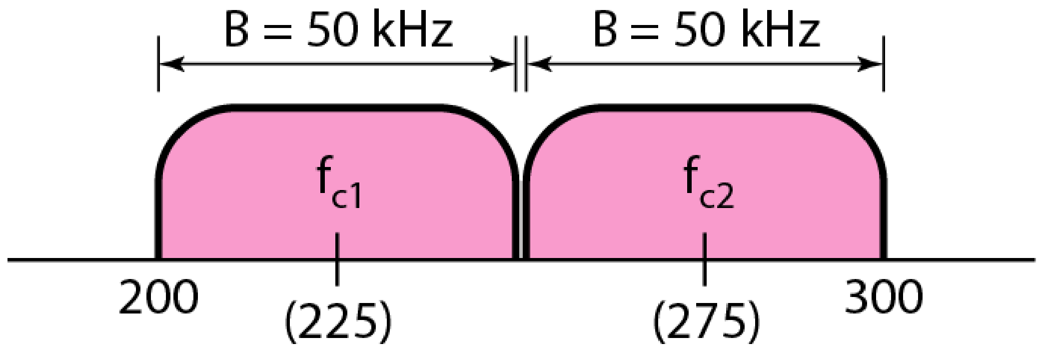
The middle of the bandwidth is located at 250 kHz. This means that our carrier frequency can be at f_c = 250 kHz. We can use the formula for bandwidth to find the bit rate (with d = 1 and r = 1).

$$B = (1 + d) \times S = 2 \times N \times \frac{1}{r} = 2 \times N = 100 \text{ kHz} \quad \longrightarrow \quad N = 50 \text{ kbps}$$

Example 5.4

In data communications, we normally use full-duplex links with communication in both directions. We need to divide the bandwidth into two with two carrier frequencies, as shown in Figure 5.5. The figure shows the positions of two carrier frequencies and the bandwidths. The available bandwidth for each direction is now 50 kHz, which leaves us with a data rate of 25 kbps in each direction.

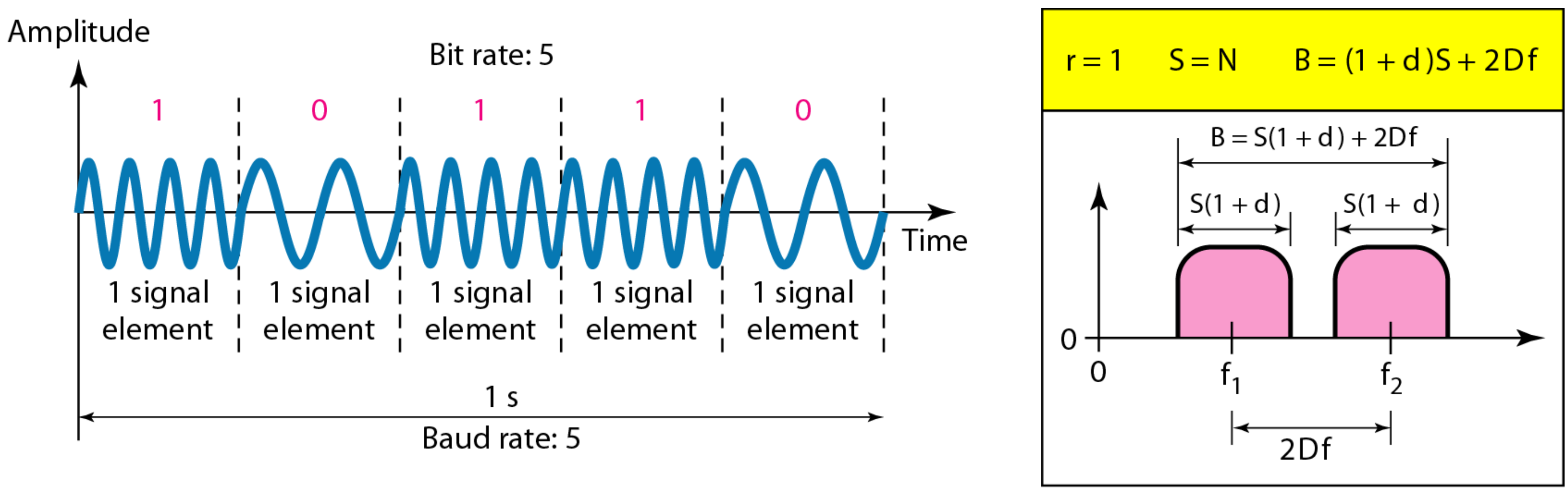
Figure 5.5 *Bandwidth of full-duplex ASK used in Example 5.4*



Frequency Shift Keying

- The digital data stream changes the frequency of the carrier signal, f_c.
- For example, a “1” could be represented by $f_1=f_c+\Delta f$, and a “0” could be represented by $f_2=f_c-\Delta f$.

Figure 5.6 *Binary frequency shift keying*



Bandwidth of FSK

- If the difference between the two frequencies (f_1 and f_2) is $2\Delta f$, then the required BW B will be:

$$B = (1+d)xS + 2\Delta f$$

Example 5.5

We have an available bandwidth of 100 kHz which spans from 200 to 300 kHz. What should be the carrier frequency and the bit rate if we modulated our data by using FSK with d = 1?

Solution

This problem is similar to Example 5.3, but we are modulating by using FSK. The midpoint of the band is at 250 kHz. We choose 2Δf to be 50 kHz; this means

$$B = (1 + d) \times S + 2\Delta f = 100 \quad \longrightarrow \quad 2S = 50 \text{ kHz} \quad S = 25 \text{ kbaud} \quad N = 25 \text{ kbps}$$

Coherent and Non Coherent

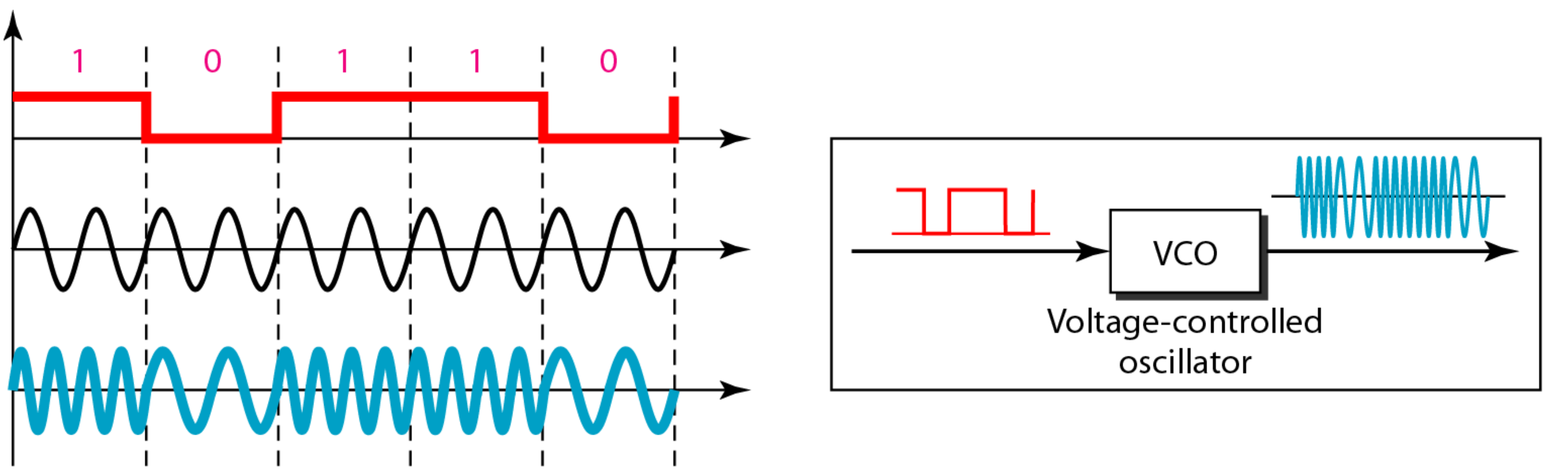
- In a non-coherent FSK scheme, when we change from one frequency to the other, we do not adhere to the current phase of the signal.
- In coherent FSK, the switch from one frequency signal to the other only occurs at the same phase in the signal.

Multi level FSK

- Similarly to ASK, FSK can use multiple bits per signal element.
- That means we need to provision for multiple frequencies, each one to represent a group of data bits.
- The bandwidth for FSK can be higher

$$B = (1+d)xS + (L-1)/2\Delta f = LxS$$

Figure 5.7 *Bandwidth of MFSK used in Example 5.6*



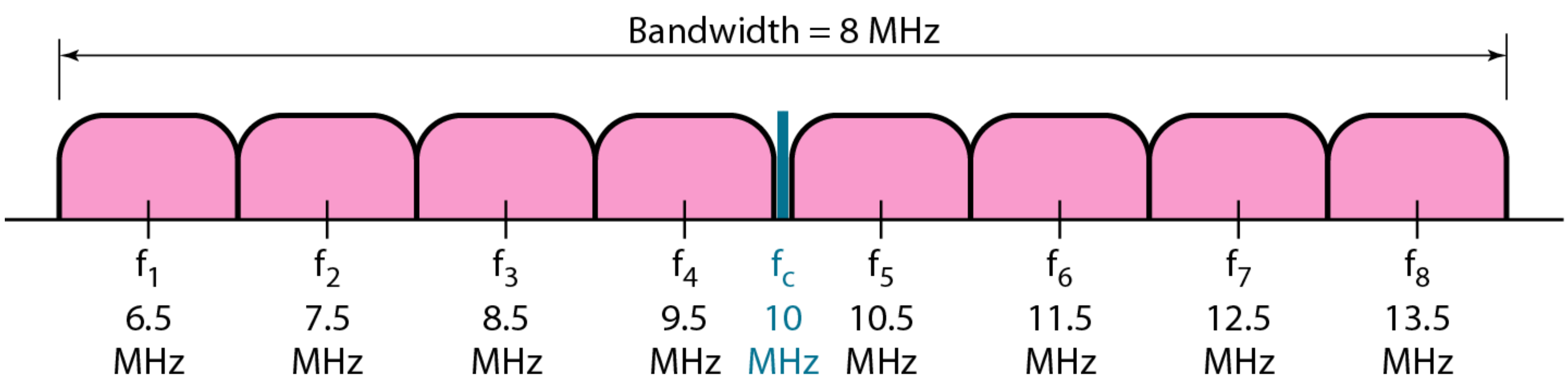
Example 5.6

We need to send data 3 bits at a time at a bit rate of 3 Mbps. The carrier frequency is 10 MHz. Calculate the number of levels (different frequencies), the baud rate, and the bandwidth.

Solution

We can have $L = 2^3 = 8$. The baud rate is $S = 3$ Mbps/3 = 1 Mbaud. This means that the carrier frequencies must be 1 MHz apart ($2\Delta f = 1$ MHz). The bandwidth is $B = 8 \times 1M = 8M$. Figure 5.8 shows the allocation of frequencies and bandwidth.

Figure 5.8 ***Bandwidth of MFSK used in Example 5.6***



Phase Shift Keyeing

- We vary the phase shift of the carrier signal to represent digital data.
- The bandwidth requirement, B is:

$$B = (1+d)xS$$

- PSK is much more robust than ASK as it is not that vulnerable to noise, which changes amplitude of the signal.

Figure 5.9 *Binary phase shift keying*

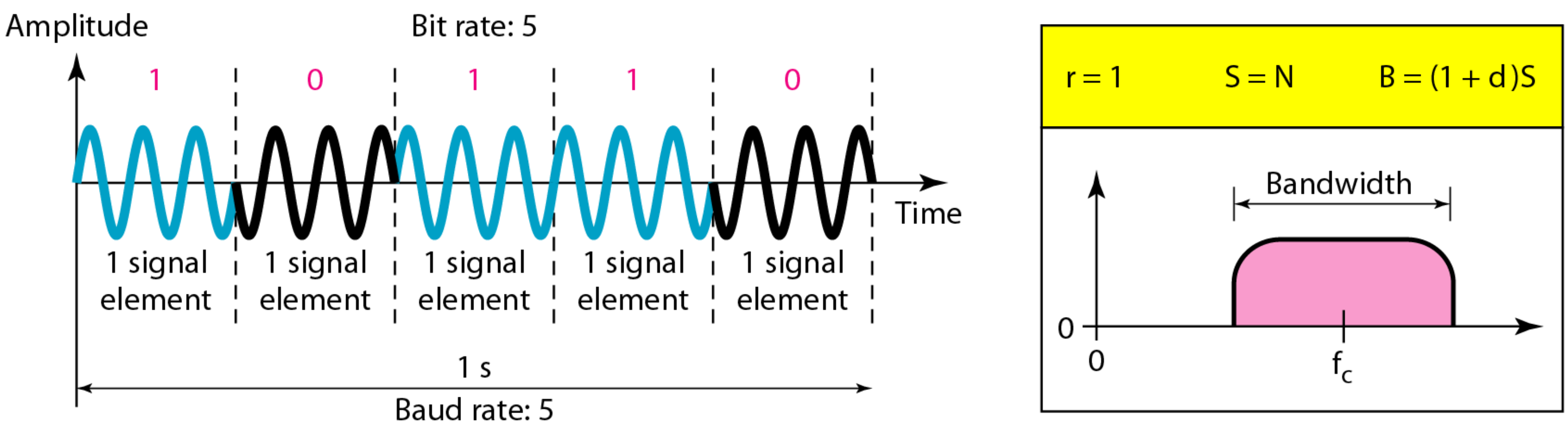
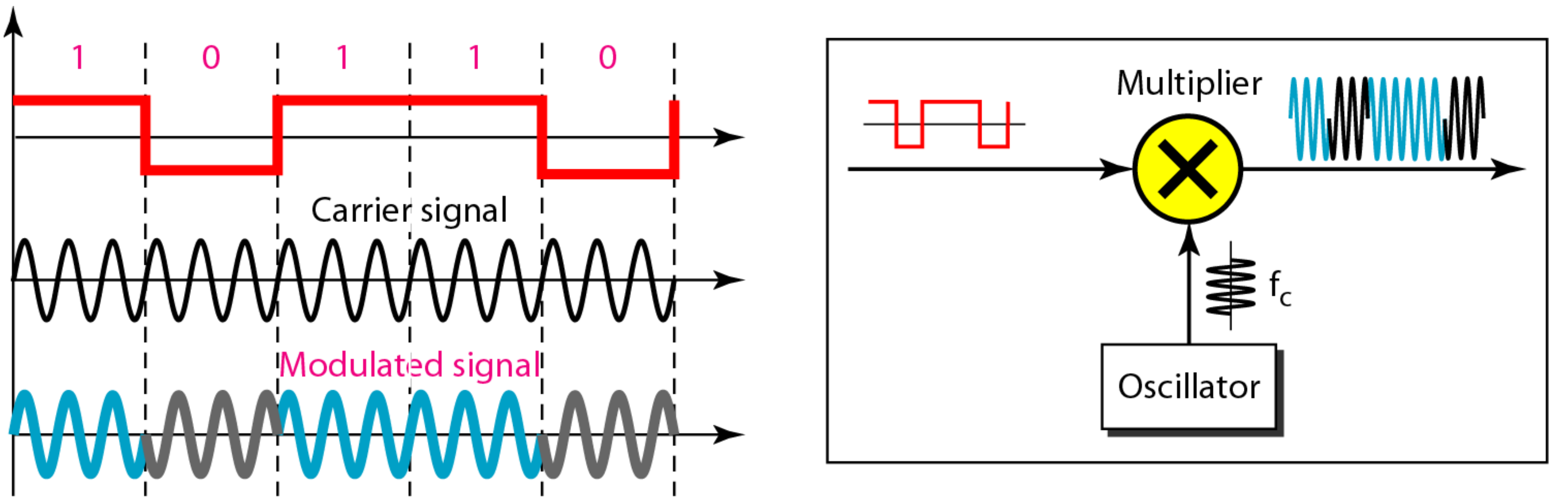


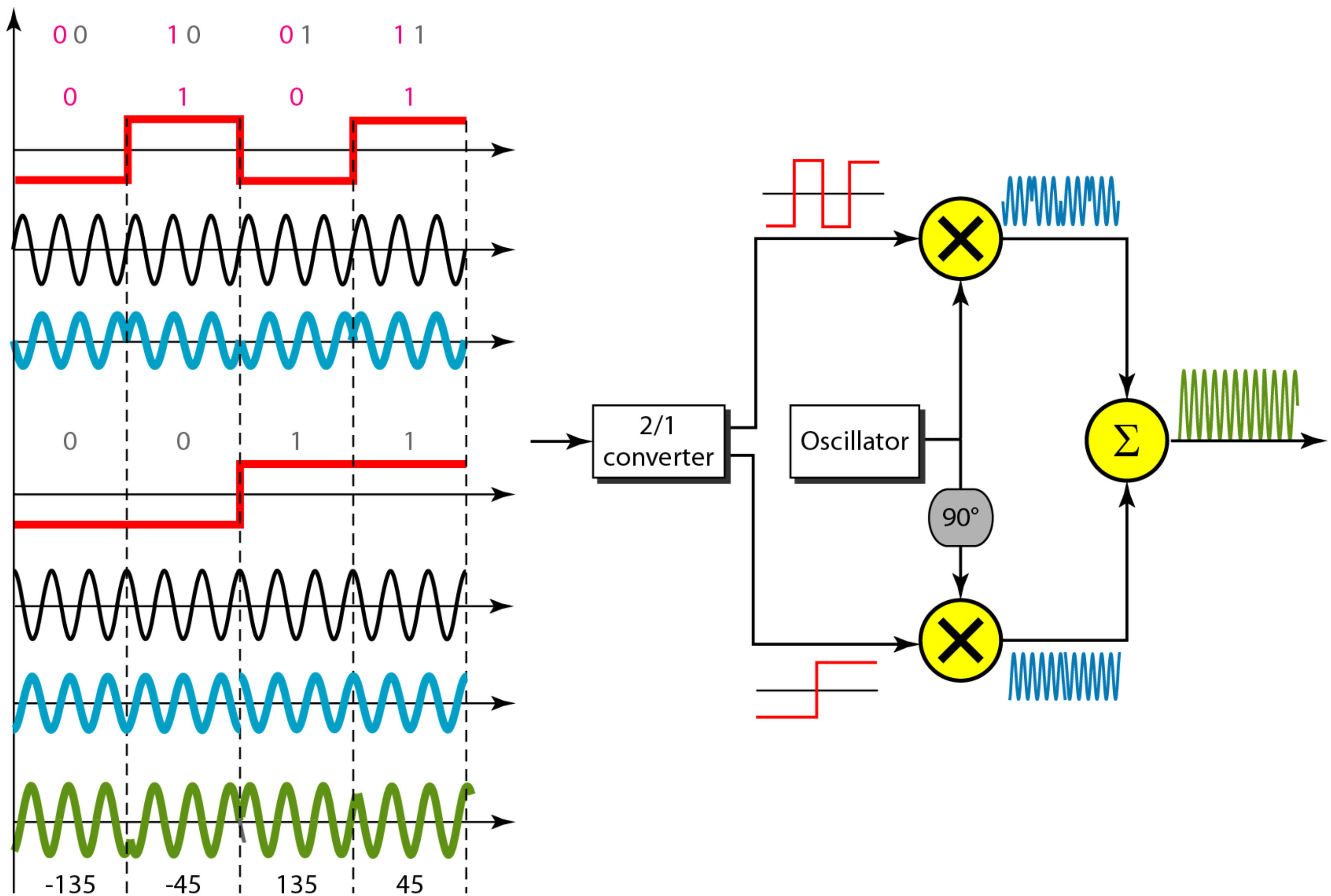
Figure 5.10 *Implementation of BASK*



Quadrature PSK

- To increase the bit rate, we can code 2 or more bits onto one signal element.
- In QPSK, we parallelize the bit stream so that every two incoming bits are split up and PSK a carrier frequency. One carrier frequency is phase shifted 90^{o} from the other - in quadrature.
- The two PSKed signals are then added to produce one of 4 signal elements. L = 4 here.

Figure 5.11 *QPSK and its implementation*



Example 5.7

Find the bandwidth for a signal transmitting at 12 Mbps for QPSK. The value of d = 0.

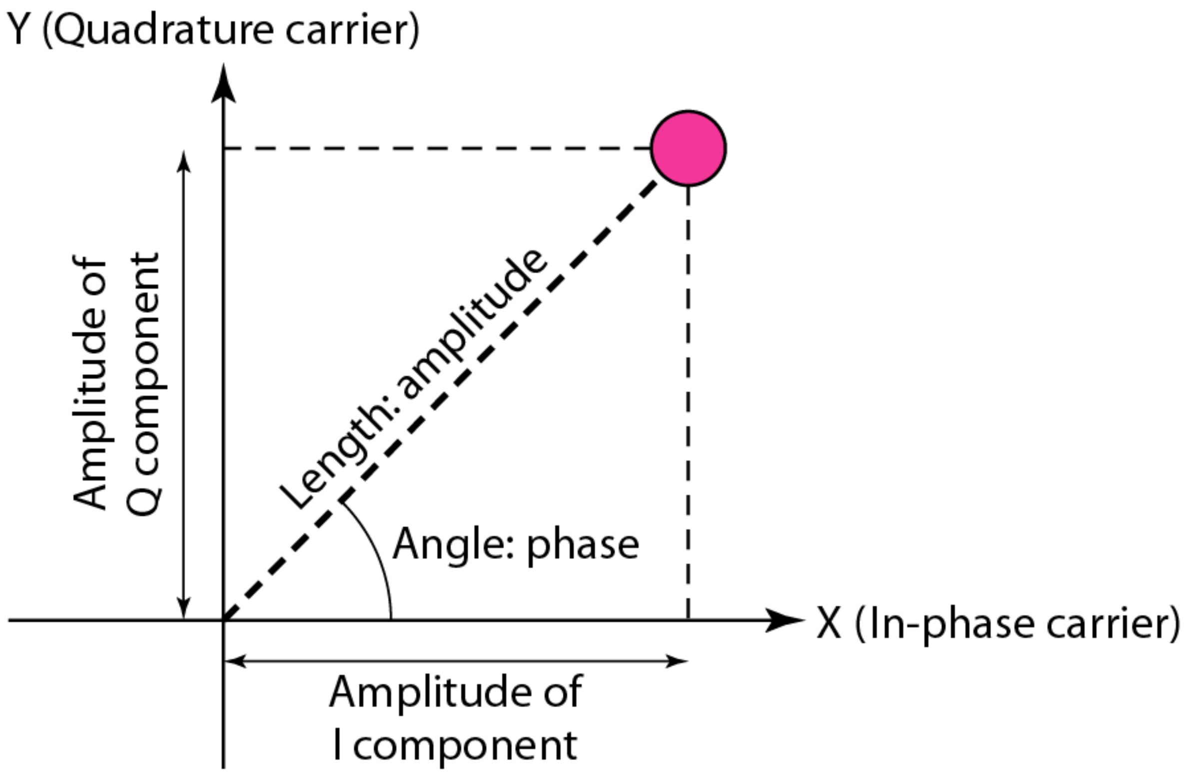
Solution

For QPSK, 2 bits is carried by one signal element. This means that r = 2. So the signal rate (baud rate) is S = N × (1/r) = 6 Mbaud. With a value of d = 0, we have B = S = 6 MHz.

Constellation Diagrams

- A constellation diagram helps us to define the amplitude and phase of a signal when we are using two carriers, one in quadrature of the other.
- The X-axis represents the in-phase carrier and the Y-axis represents quadrature carrier.

Figure 5.12 *Concept of a constellation diagram*



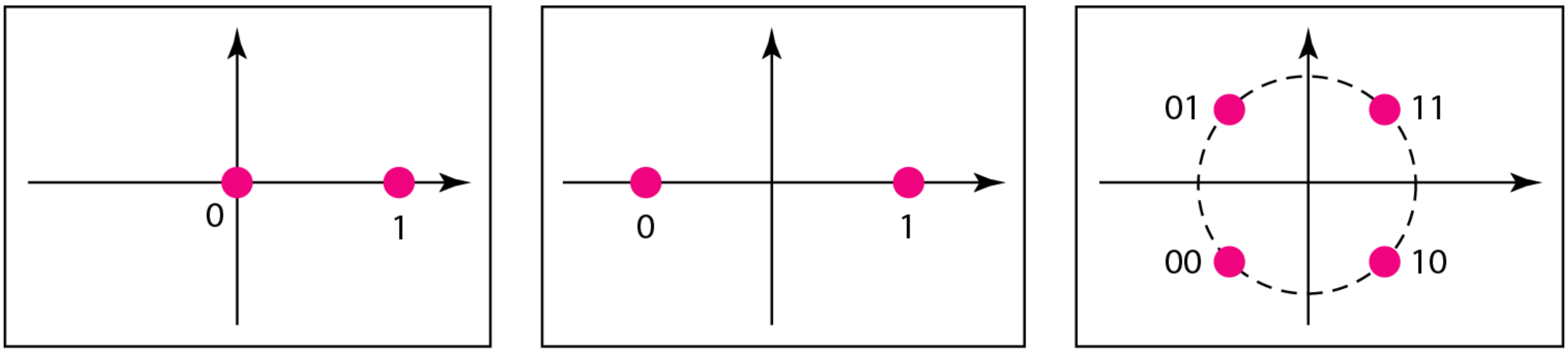
Example 5.8

Show the constellation diagrams for an ASK (OOK), BPSK, and QPSK signals.

Solution

Figure 5.13 shows the three constellation diagrams.

Figure 5.13 *Three constellation diagrams*



a. ASK (OOK)

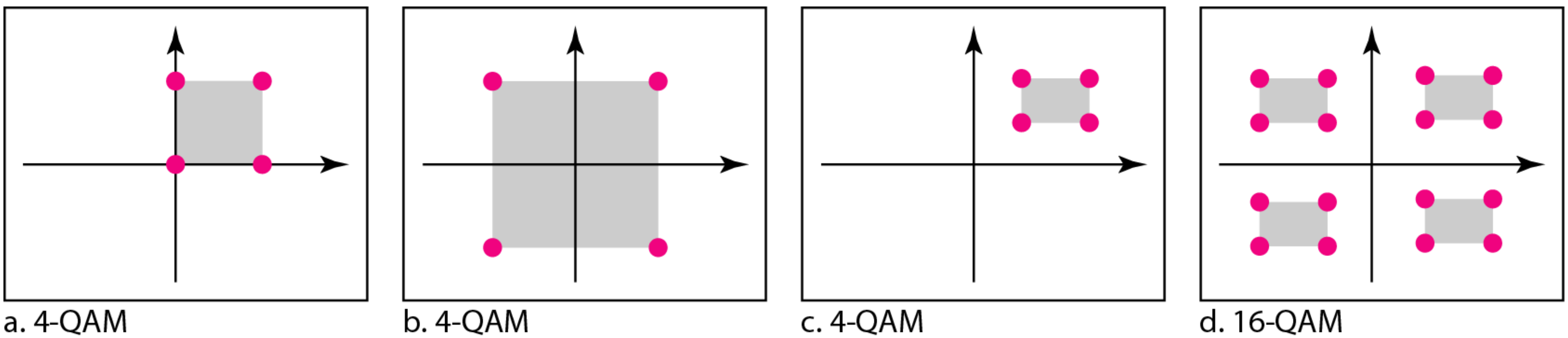
b. BPSK

c. QPSK

Note

Quadrature amplitude modulation is a combination of ASK and PSK.

Figure 5.14 *Constellation diagrams for some QAMs*



Chapter 5

Analog Transmission

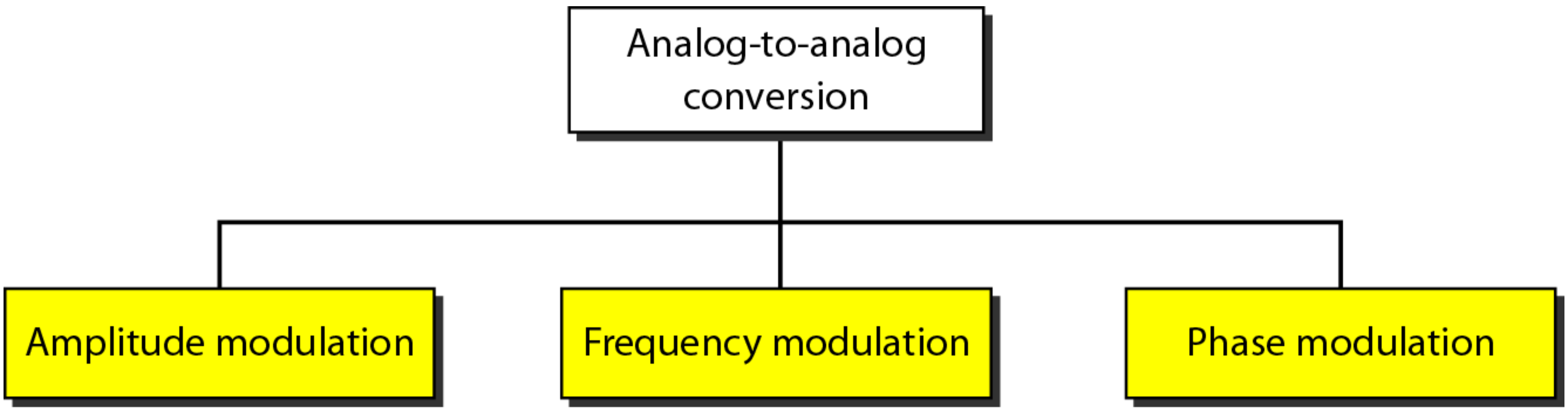
5-2 ANALOG AND DIGITAL

Analog-to-analog conversion is the representation of analog information by an analog signal. One may ask why we need to modulate an analog signal; it is already analog. Modulation is needed if the medium is bandpass in nature or if only a bandpass channel is available to us.

Topics discussed in this section:

- **Amplitude Modulation**
- **Frequency Modulation**
- **Phase Modulation**

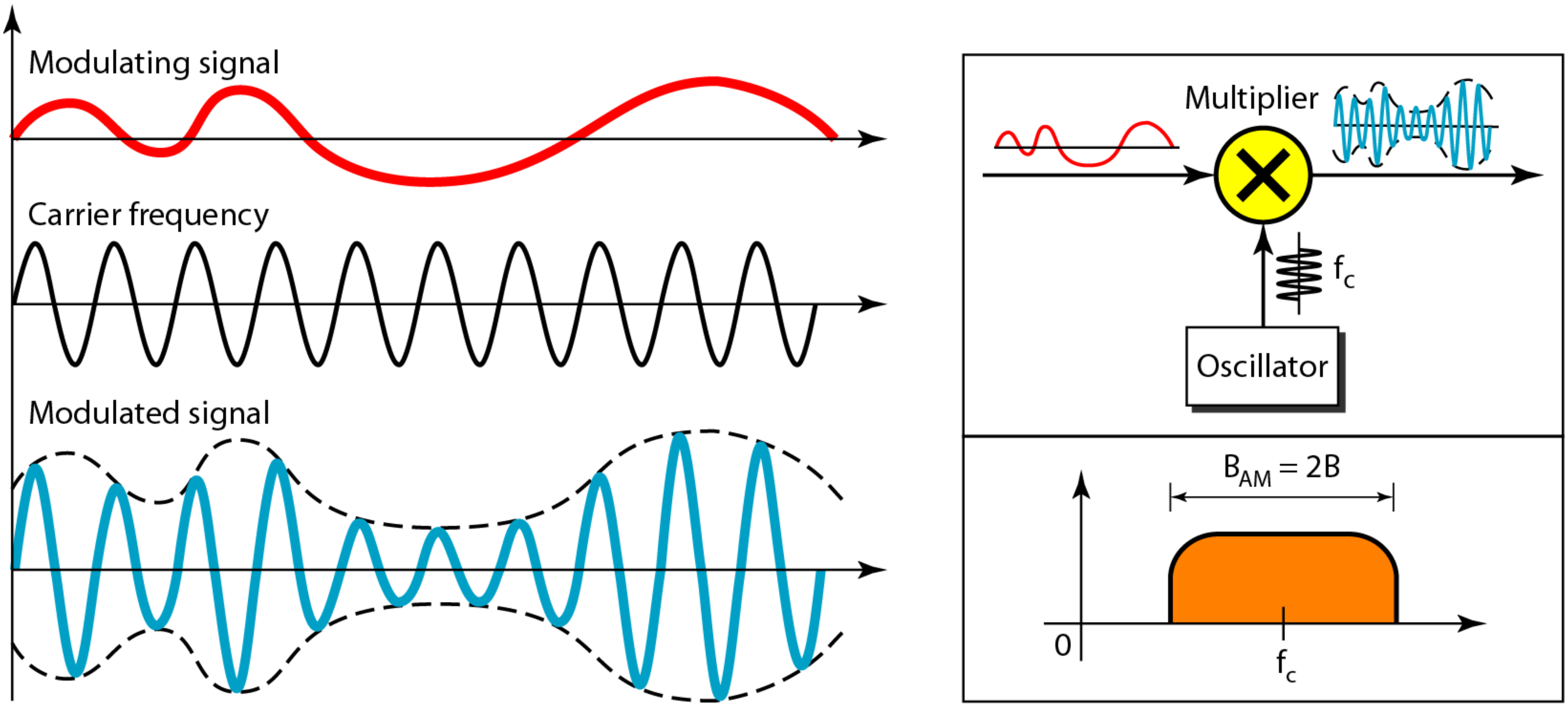
Figure 5.15 *Types of analog-to-analog modulation*



Amplitude Modulation

- A carrier signal is modulated only in amplitude value
- The modulating signal is the envelope of the carrier
- The required bandwidth is 2B, where B is the bandwidth of the modulating signal
- Since on both sides of the carrier freq. f_c, the spectrum is identical, we can discard one half, thus requiring a smaller bandwidth for transmission.

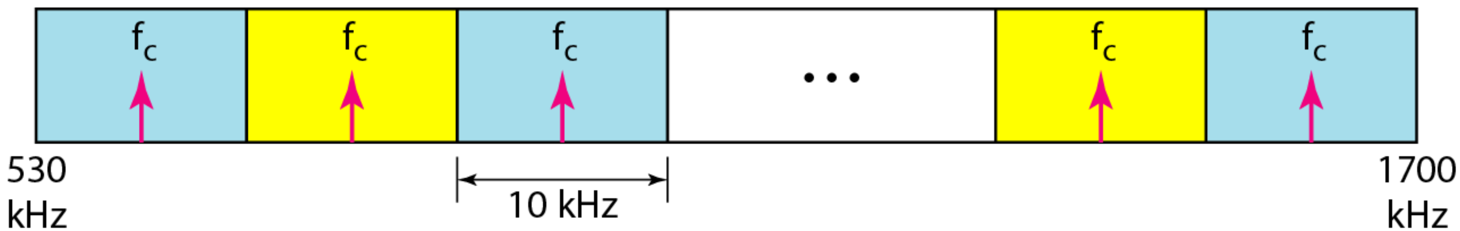
Figure 5.16 *Amplitude modulation*



Note

The total bandwidth required for AM can be determined from the bandwidth of the audio signal: $B_{AM} = 2B$.

Figure 5.17 *AM band allocation*



Frequency Modulation

- The modulating signal changes the freq. f_c of the carrier signal
- The bandwidth for FM is high
- It is approx. 10x the signal frequency

Note

The total bandwidth required for FM can be determined from the bandwidth of the audio signal: $B_{FM} = 2(1 + \beta)B$. Where β is usually 4.

Figure 5.18 *Frequency modulation*

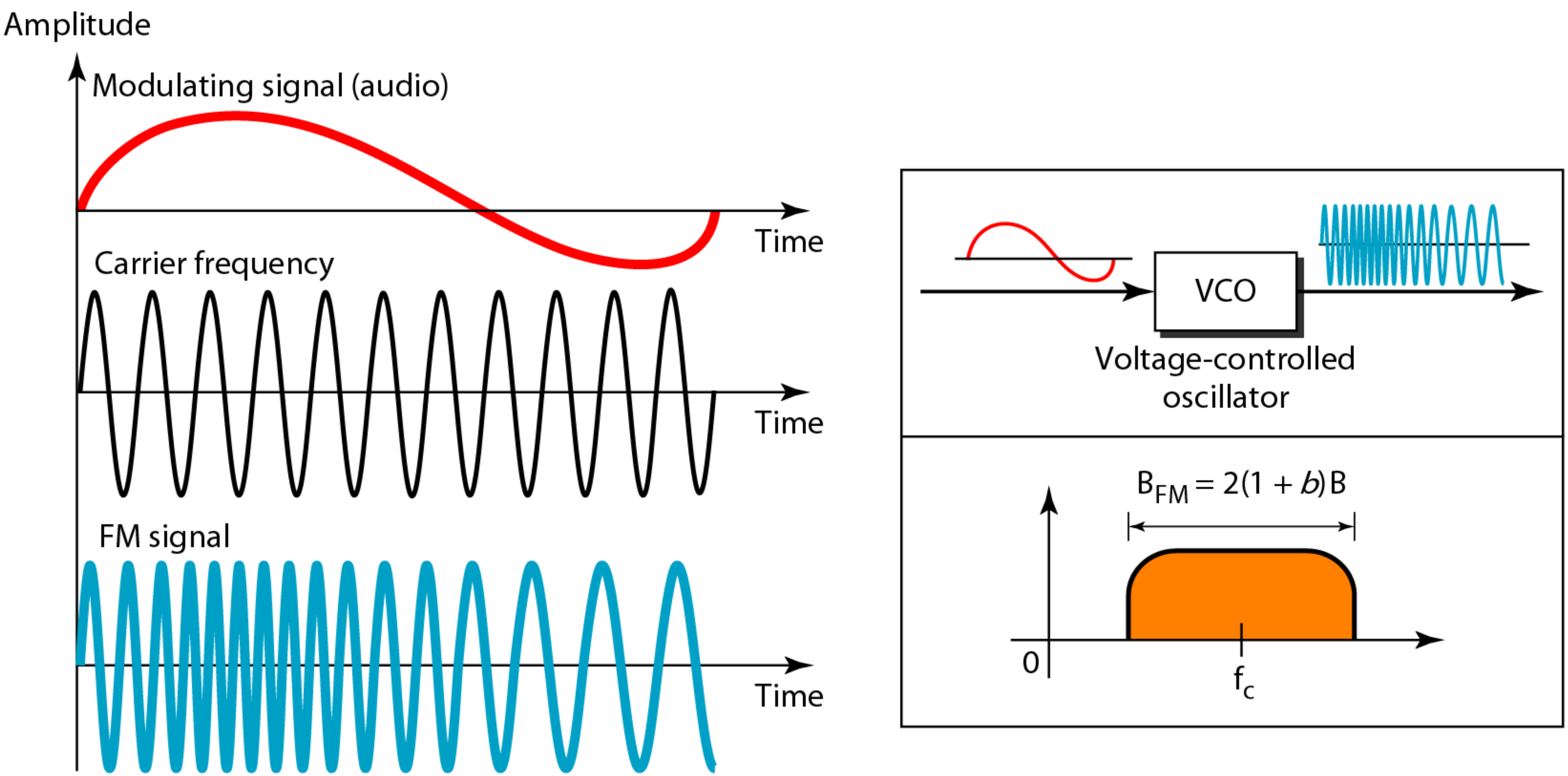
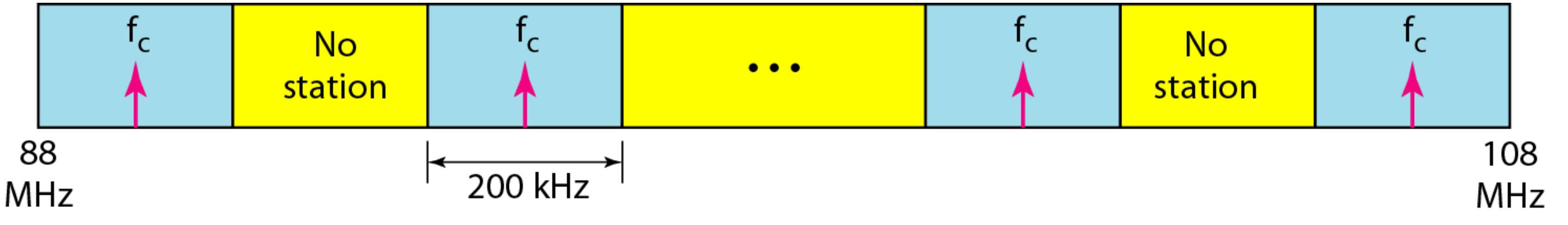


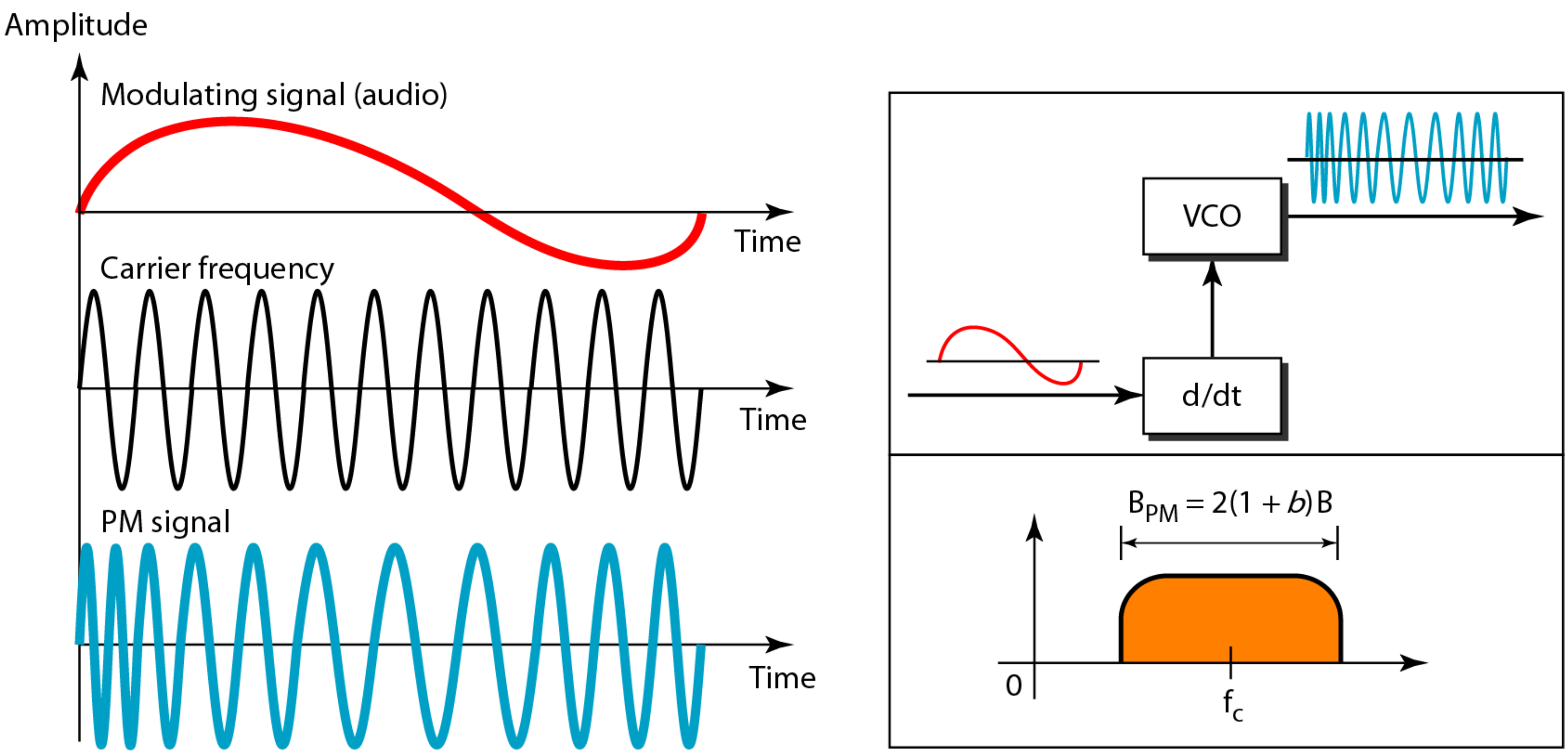
Figure 5.19 *FM band allocation*



Phase Modulation (PM)

- The modulating signal only changes the phase of the carrier signal.
- The phase change manifests itself as a frequency change but the instantaneous frequency change is proportional to the derivative of the amplitude.
- The bandwidth is higher than for AM.

Figure 5.20 *Phase modulation*



Note

The total bandwidth required for PM can be determined from the bandwidth and maximum amplitude of the modulating signal:

$$B_{PM} = 2(1 + \beta)B.$$

Where $\beta = 2$ most often.

Chapter 6

Bandwidth Utilization: Multiplexing and Spreading

Note

Bandwidth utilization is the wise use of available bandwidth to achieve specific goals.

Efficiency can be achieved by multiplexing; i.e., sharing of the bandwidth between multiple users.

6-1 MULTIPLEXING

Whenever the bandwidth of a medium linking two devices is greater than the bandwidth needs of the devices, the link can be shared. Multiplexing is the set of techniques that allows the (simultaneous) transmission of multiple signals across a single data link. As data and telecommunications use increases, so does traffic.

Topics discussed in this section:

- **Frequency-Division Multiplexing**
- **Wavelength-Division Multiplexing**
- **Synchronous Time-Division Multiplexing**
- **Statistical Time-Division Multiplexing**

Figure 6.1 *Dividing a link into channels*

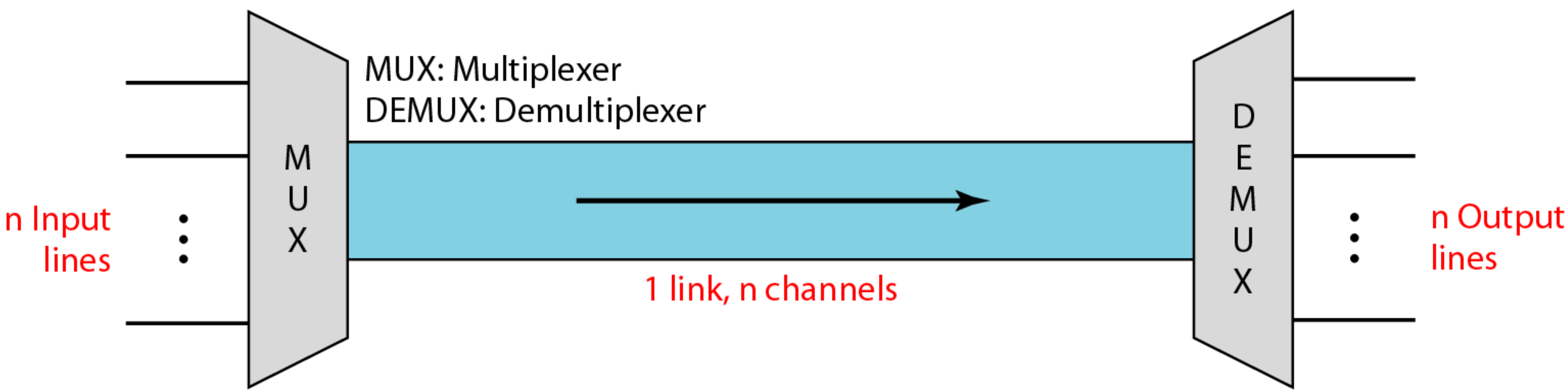


Figure 6.2 *Categories of multiplexing*

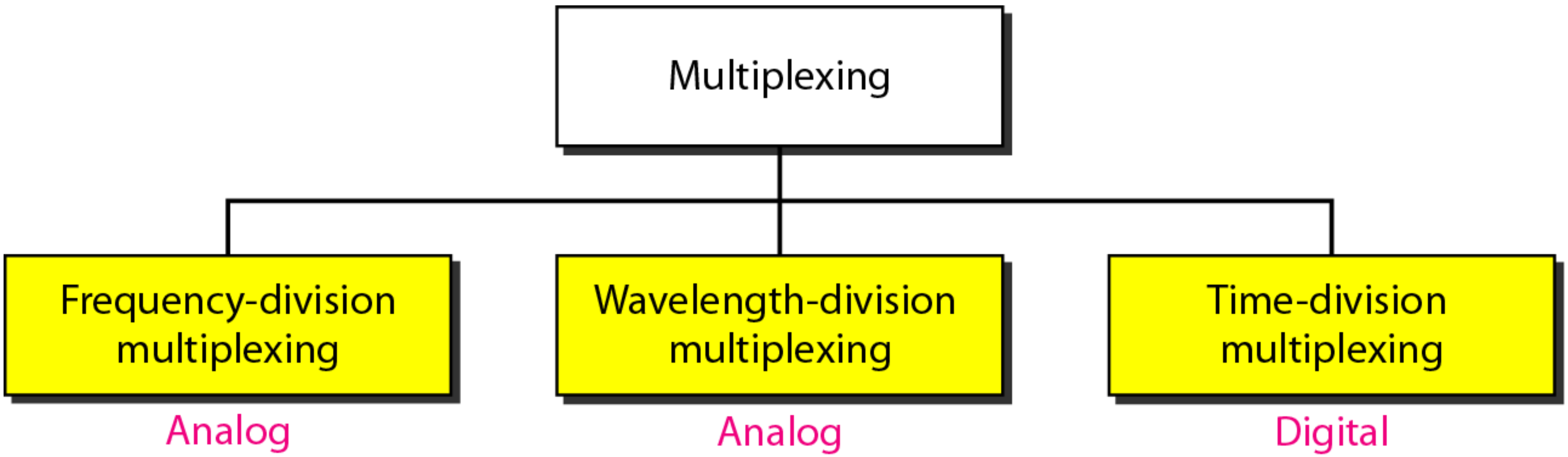
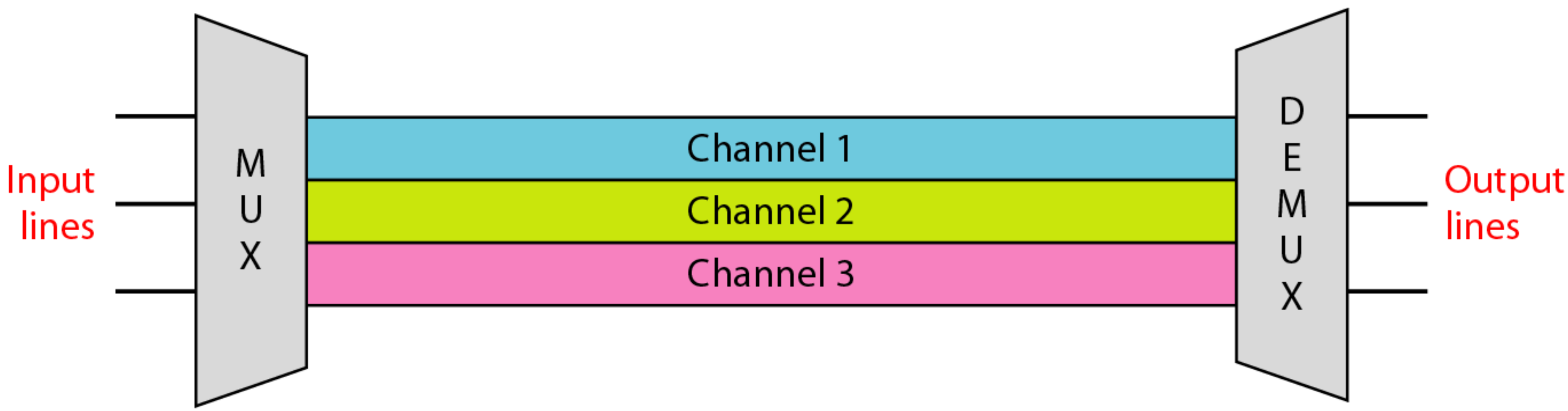


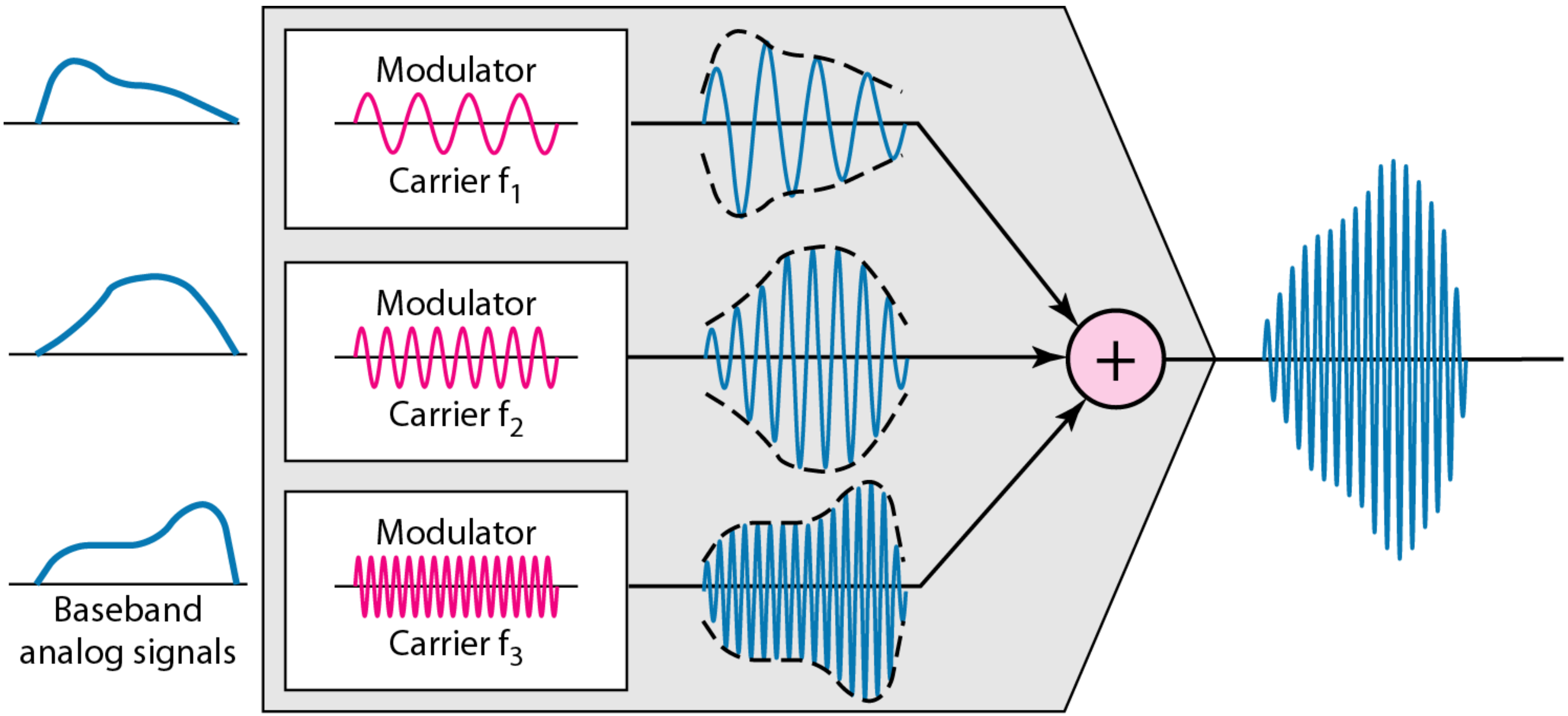
Figure 6.3 ***Frequency-division multiplexing (FDM)***



Note

FDM is an analog multiplexing technique that combines analog signals.
It uses the concept of modulation discussed in Ch 5.

Figure 6.4 *FDM process*



FM

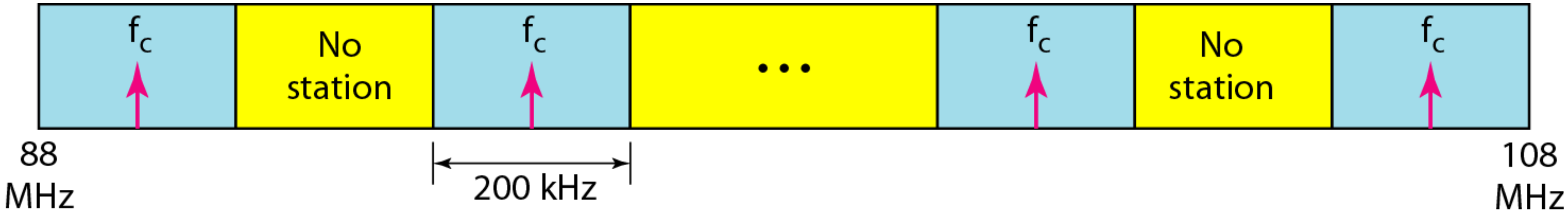
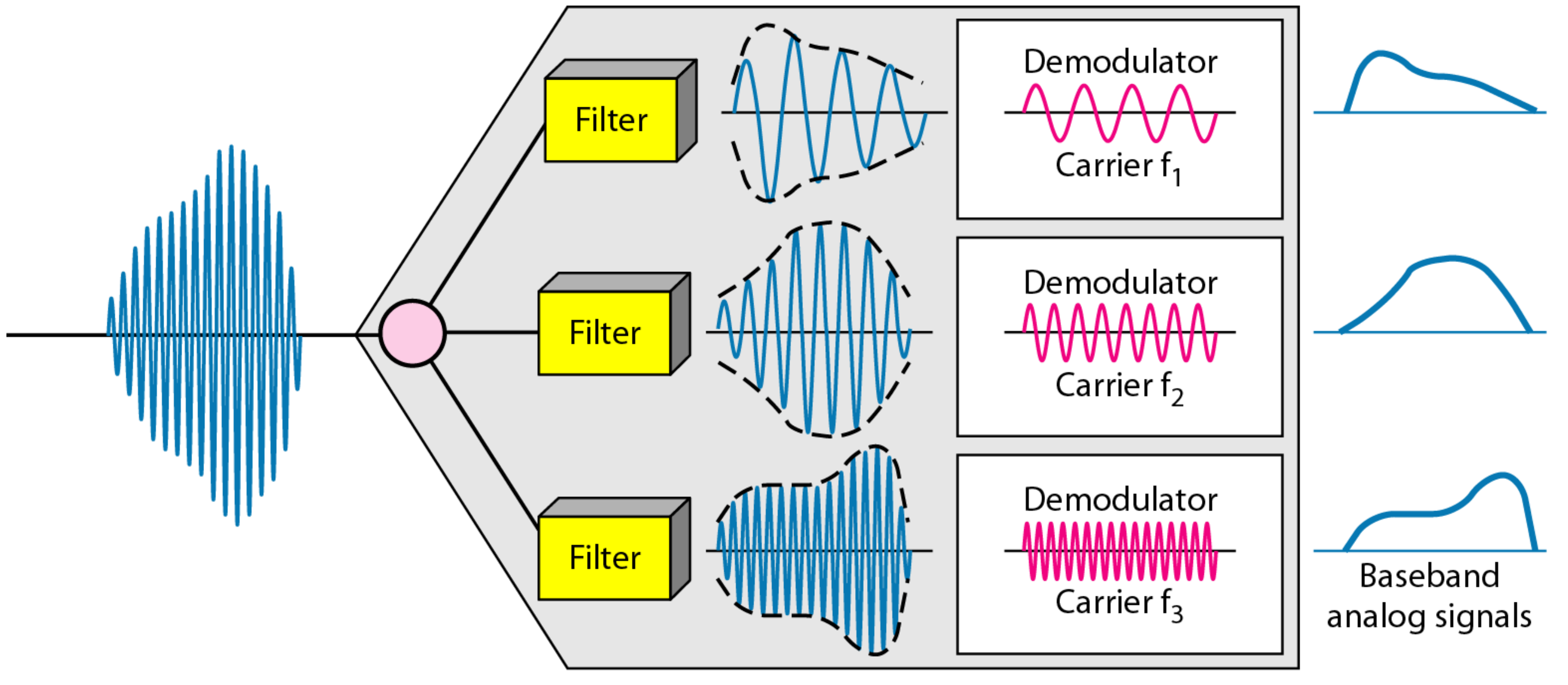


Figure 6.5 *FDM demultiplexing example*



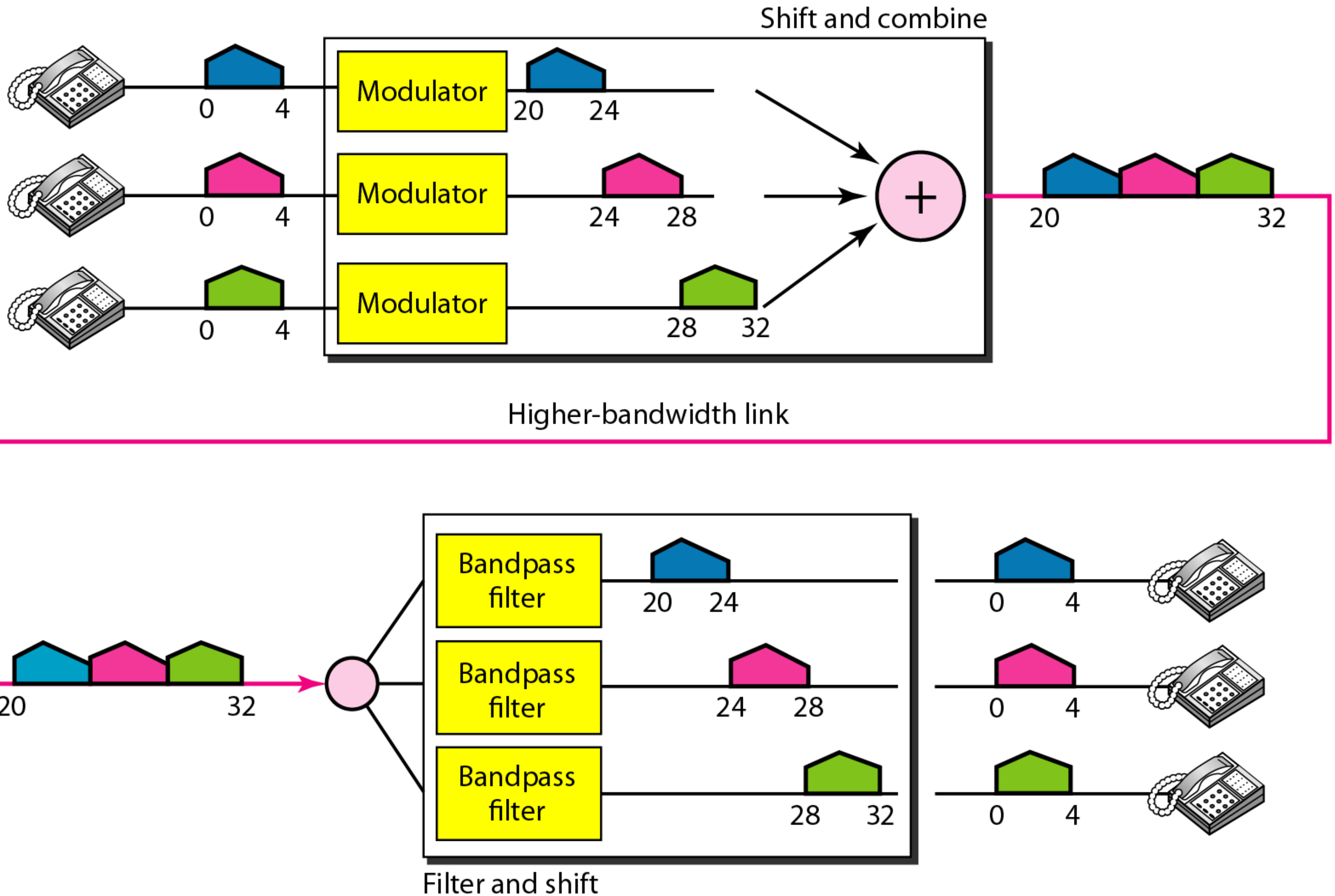
Example 6.1

Assume that a voice channel occupies a bandwidth of 4 kHz. We need to combine three voice channels into a link with a bandwidth of 12 kHz, from 20 to 32 kHz. Show the configuration, using the frequency domain. Assume there are no guard bands.

Solution

We shift (modulate) each of the three voice channels to a different bandwidth, as shown in Figure 6.6. We use the 20- to 24-kHz bandwidth for the first channel, the 24- to 28-kHz bandwidth for the second channel, and the 28- to 32-kHz bandwidth for the third one. Then we combine them as shown in Figure 6.6.

Figure 6.6 *Example 6.1*



Example 6.2

Five channels, each with a 100-kHz bandwidth, are to be multiplexed together. What is the minimum bandwidth of the link if there is a need for a guard band of 10 kHz between the channels to prevent interference?

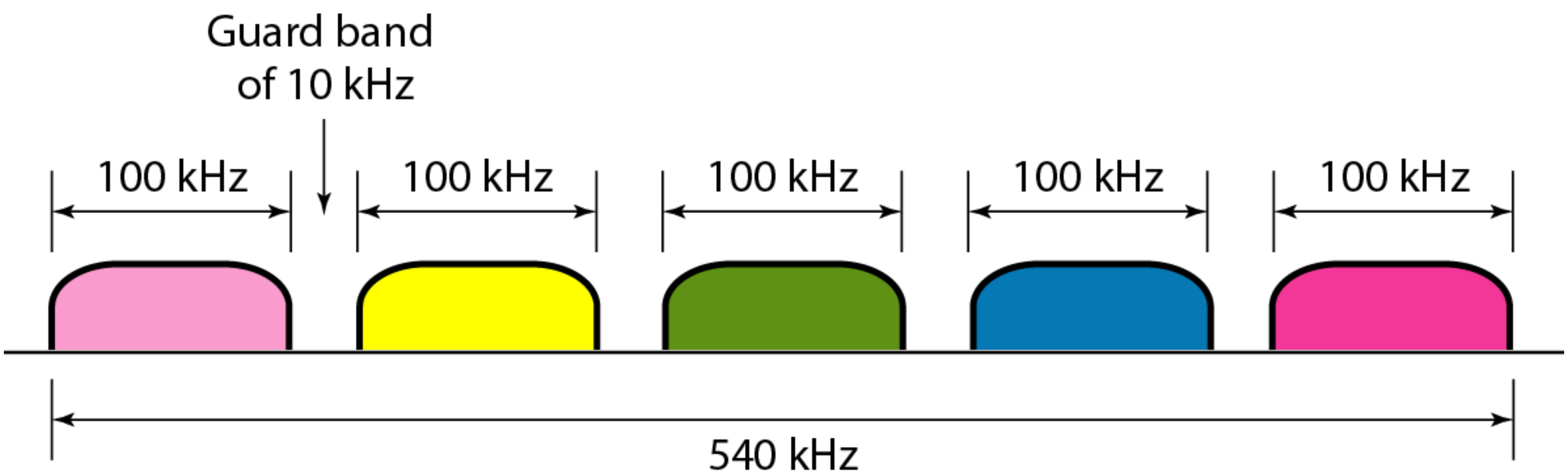
Solution

For five channels, we need at least four guard bands. This means that the required bandwidth is at least

$$5 \times 100 + 4 \times 10 = 540 \text{ kHz},$$

as shown in Figure 6.7.

Figure 6.7 *Example 6.2*



Example 6.3

Four data channels (digital), each transmitting at 1 Mbps, use a satellite channel of 1 MHz. Design an appropriate configuration, using FDM.

Solution

The satellite channel is analog. We divide it into four channels, each channel having 1M/4=250-kHz bandwidth.

Each digital channel of 1 Mbps must be transmitted over a 250KHz channel. Assuming no noise we can use Nyquist to get:

$C = 1Mbps = 2x250K \; x \; \log_2 L$ -> $L = 4$ or $n = 2$ bits/signal element.

One solution is 4-QAM modulation. In Figure 6.8 we show a possible configuration with $L = 16$.

Figure 6.8 *Example 6.3*

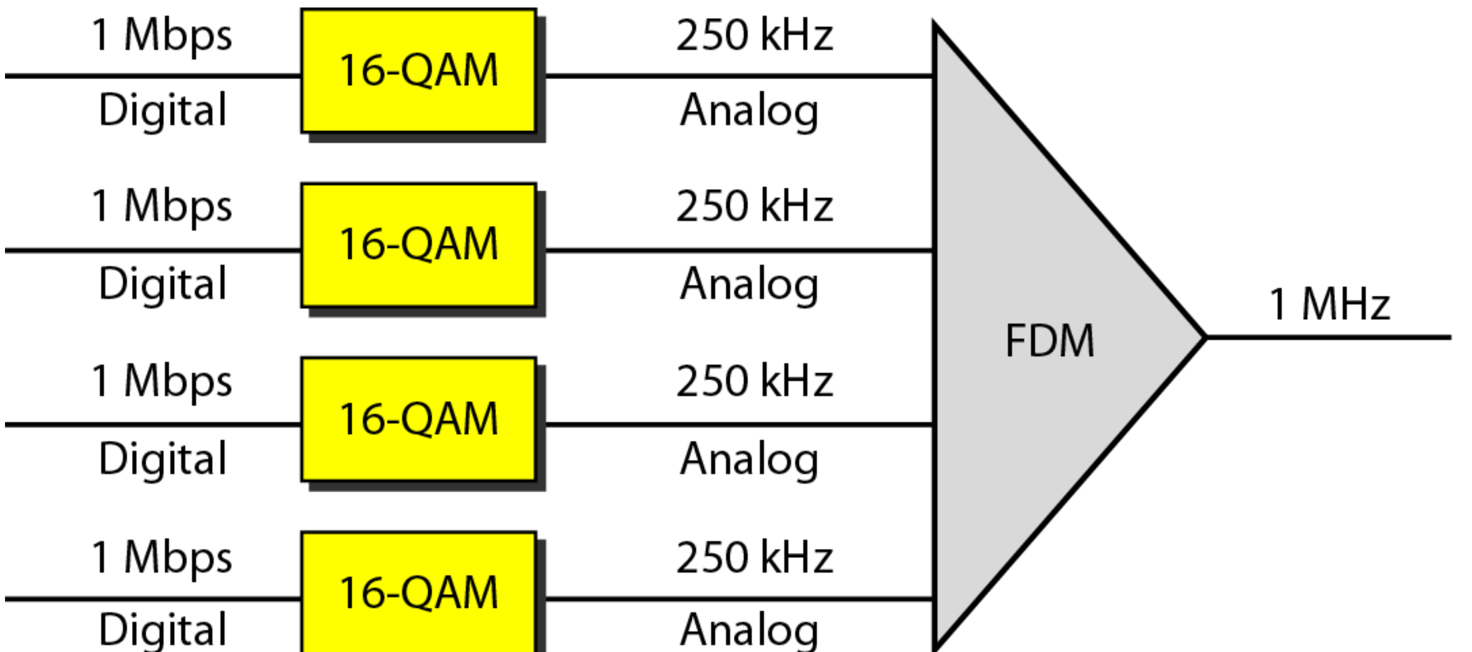
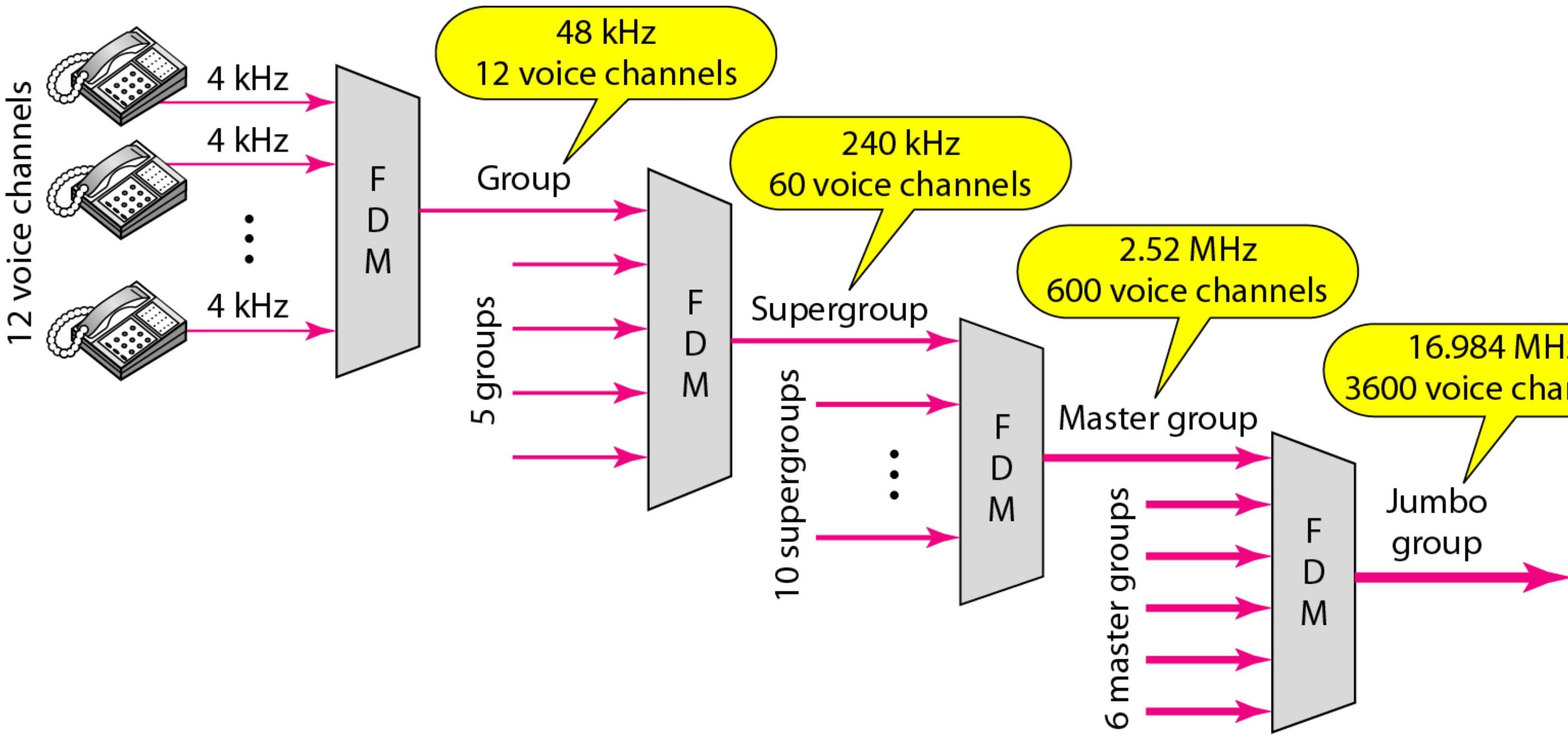


Figure 6.9 *Analog hierarchy*



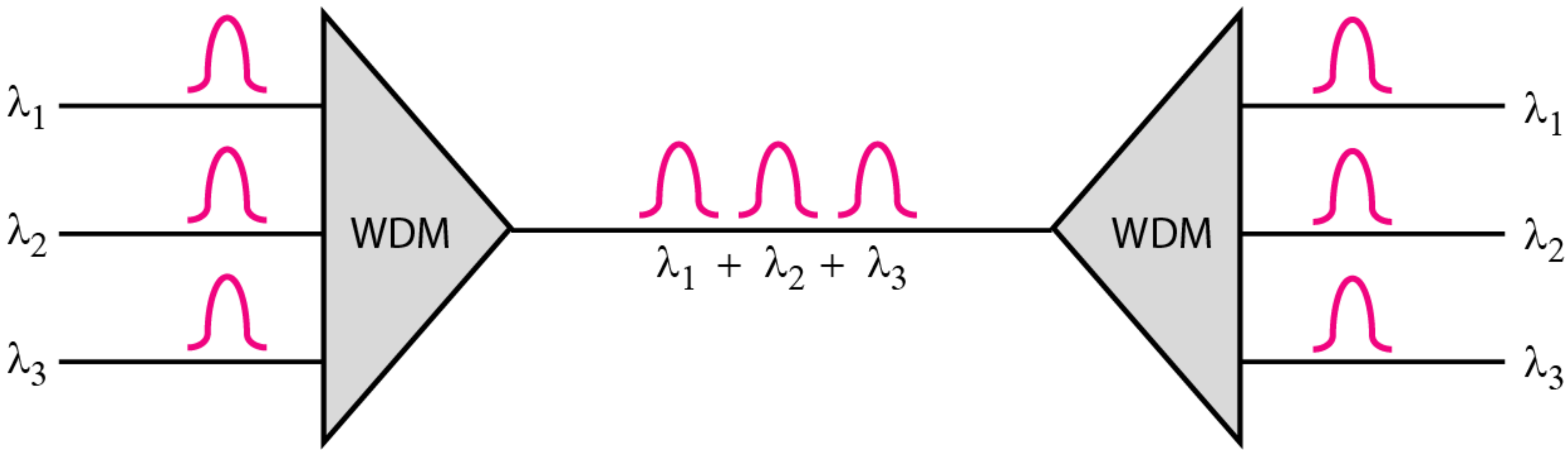
Example 6.4

The Advanced Mobile Phone System (AMPS) uses two bands. The first band of 824 to 849 MHz is used for sending, and 869 to 894 MHz is used for receiving. Each user has a bandwidth of 30 kHz in each direction. How many people can use their cellular phones simultaneously?

Solution

Each band is 25 MHz. If we divide 25 MHz by 30 kHz, we get 833.33. In reality, the band is divided into 832 channels. Of these, 42 channels are used for control, which means only 790 channels are available for cellular phone users.

Figure 6.10 *Wavelength-division multiplexing (WDM)*



Note

WDM is an analog multiplexing technique to combine optical signals.

Figure 6.11 *Prisms in wavelength-division multiplexing and demultiplexing*

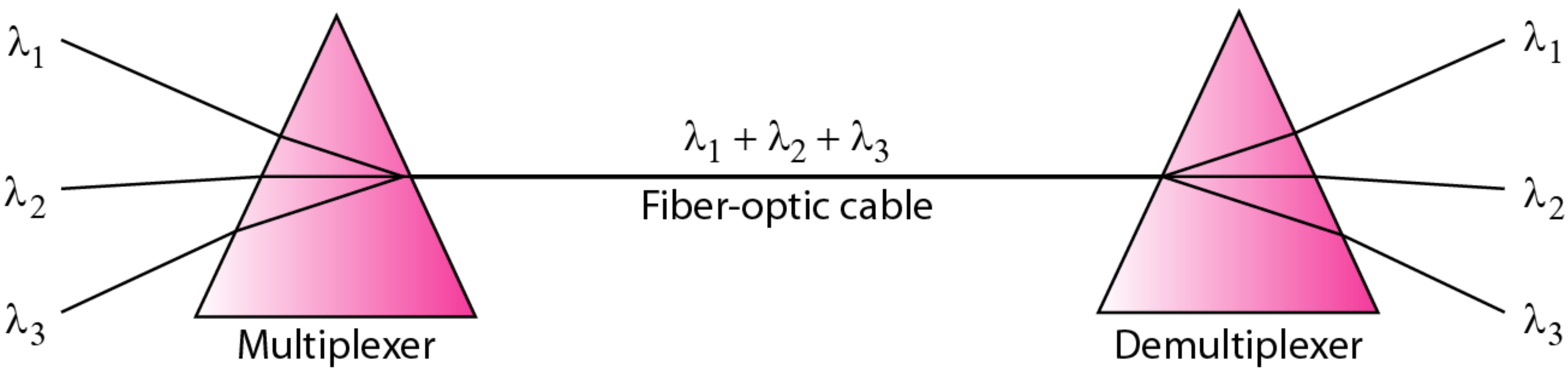
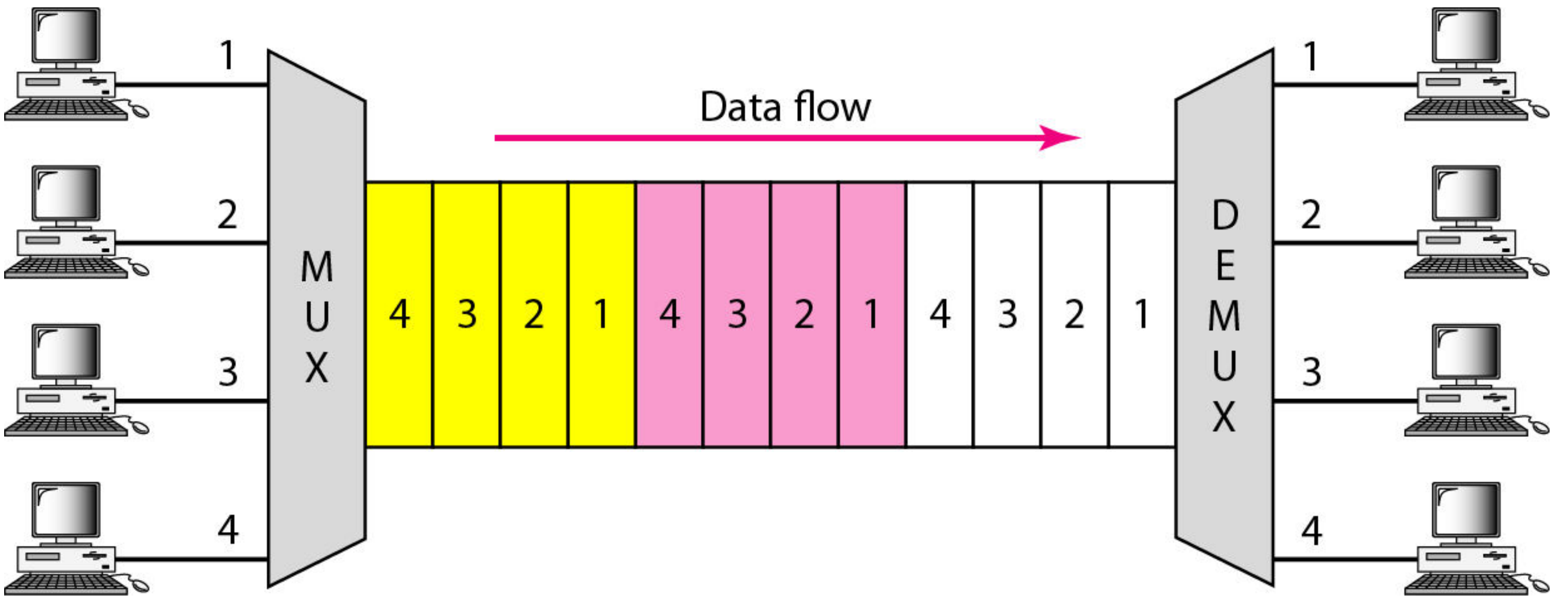


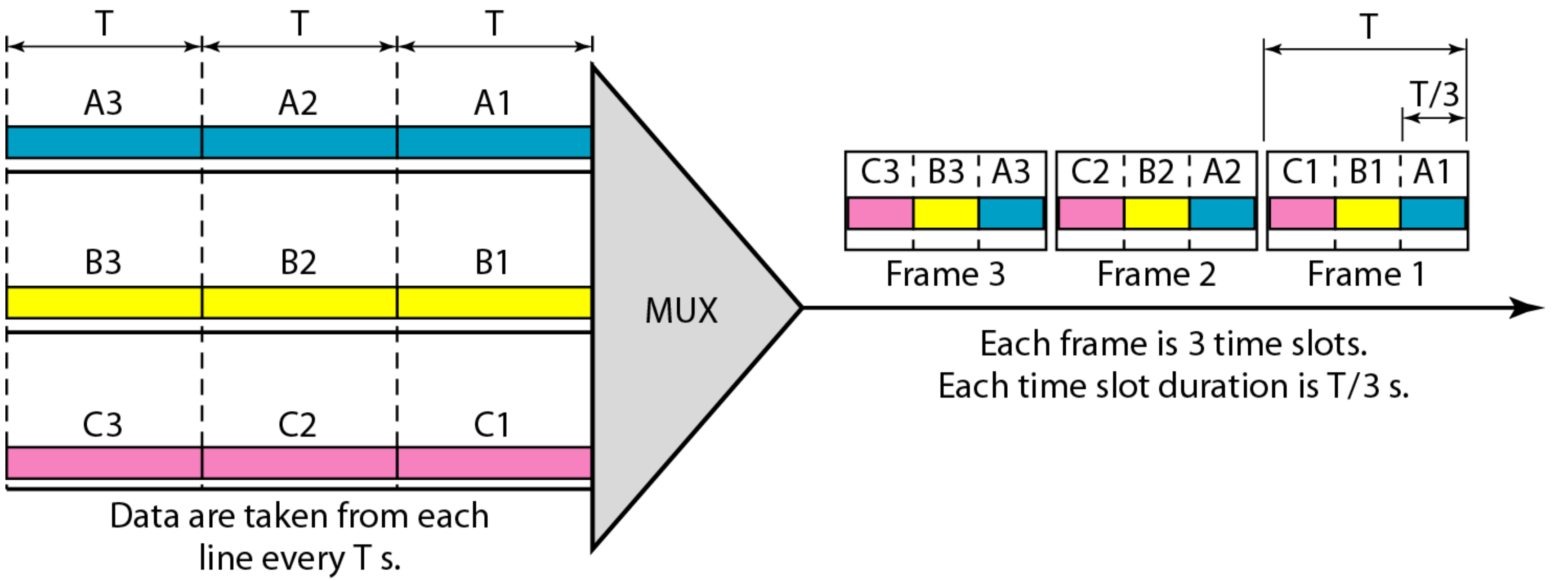
Figure 6.12 Time Division Multiplexing (*TDM*)



Note

TDM is a digital multiplexing technique for combining several low-rate digital channels into one high-rate one.

Figure 6.13 *Synchronous time-division multiplexing*



Note

In synchronous TDM, the data rate of the link is *n* times faster, and the unit duration is *n* times shorter.

Example 6.5

In Figure 6.13, the data rate for each one of the 3 input connection is 1 kbps. If 1 bit at a time is multiplexed (a unit is 1 bit), what is the duration of (a) each input slot, (b) each output slot, and (c) each frame?

Solution

We can answer the questions as follows:

a. The data rate of each input connection is 1 kbps. This means that the bit duration is 1/1000 s or 1 ms. The duration of the input time slot is 1 ms (same as bit duration).

Example 6.5 (continued)

b. ***The duration of each output time slot is one-third of the input time slot. This means that the duration of the output time slot is 1/3 ms.***

c. ***Each frame carries three output time slots. So the duration of a frame is 3 × 1/3 ms, or 1 ms.***

Note: The duration of a frame is the same as the duration of an input unit.

Example 6.6

Figure 6.14 shows synchronous TDM with 4 1Mbps data stream inputs and one data stream for the output. The unit of data is 1 bit. Find (a) the input bit duration, (b) the output bit duration, (c) the output bit rate, and (d) the output frame rate.

Solution

We can answer the questions as follows:

a. The input bit duration is the inverse of the bit rate: 1/1 Mbps = 1 μs.

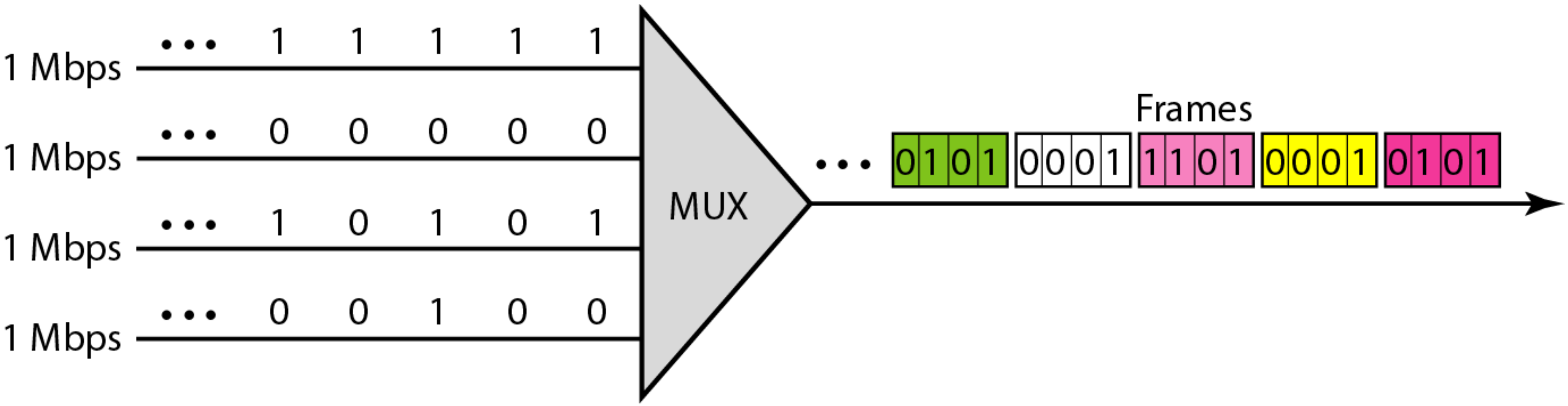
b. The output bit duration is one-fourth of the input bit duration, or ¼ μs.

Example 6.6 (continued)

c. *The output bit rate is the inverse of the output bit duration or $1/(4\mu s)$ or 4 Mbps. This can also be deduced from the fact that the output rate is 4 times as fast as any input rate; so the output rate = 4×1 Mbps = 4 Mbps.*

d. *The frame rate is always the same as any input rate. So the frame rate is 1,000,000 frames per second. Because we are sending 4 bits in each frame, we can verify the result of the previous question by multiplying the frame rate by the number of bits per frame.*

Figure 6.14 *Example 6.6*



Example 6.7

Four 1-kbps connections are multiplexed together. A unit is 1 bit. Find (a) the duration of 1 bit before multiplexing, (b) the transmission rate of the link, (c) the duration of a time slot, and (d) the duration of a frame.

Solution

We can answer the questions as follows:

a. The duration of 1 bit before multiplexing is 1 / 1 kbps, or 0.001 s (1 ms).

b. The rate of the link is 4 times the rate of a connection, or 4 kbps.

Example 6.7 (continued)

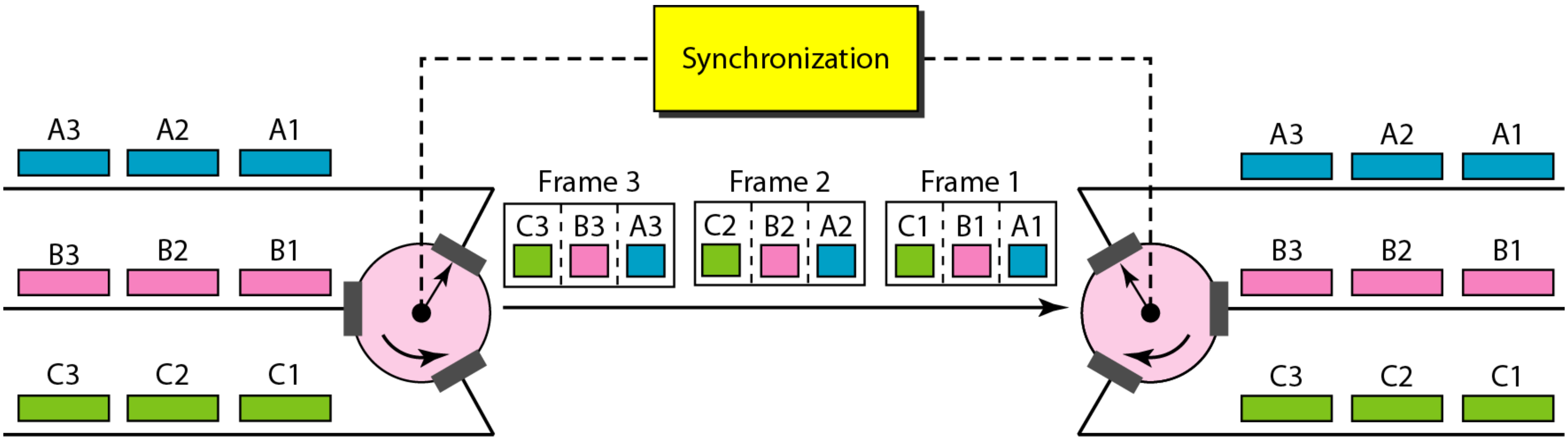
c. ***The duration of each time slot is one-fourth of the duration of each bit before multiplexing, or 1/4 ms or 250 μs. Note that we can also calculate this from the data rate of the link, 4 kbps. The bit duration is the inverse of the data rate, or 1/4 kbps or 250 μs.***

d. ***The duration of a frame is always the same as the duration of a unit before multiplexing, or 1 ms. We can also calculate this in another way. Each frame in this case has four time slots. So the duration of a frame is 4 times 250 μs, or 1 ms.***

Interleaving

- The process of taking a group of bits from each input line for multiplexing is called interleaving.
- We interleave bits (1 - n) from each input onto one output.

Figure 6.15 *Interleaving*



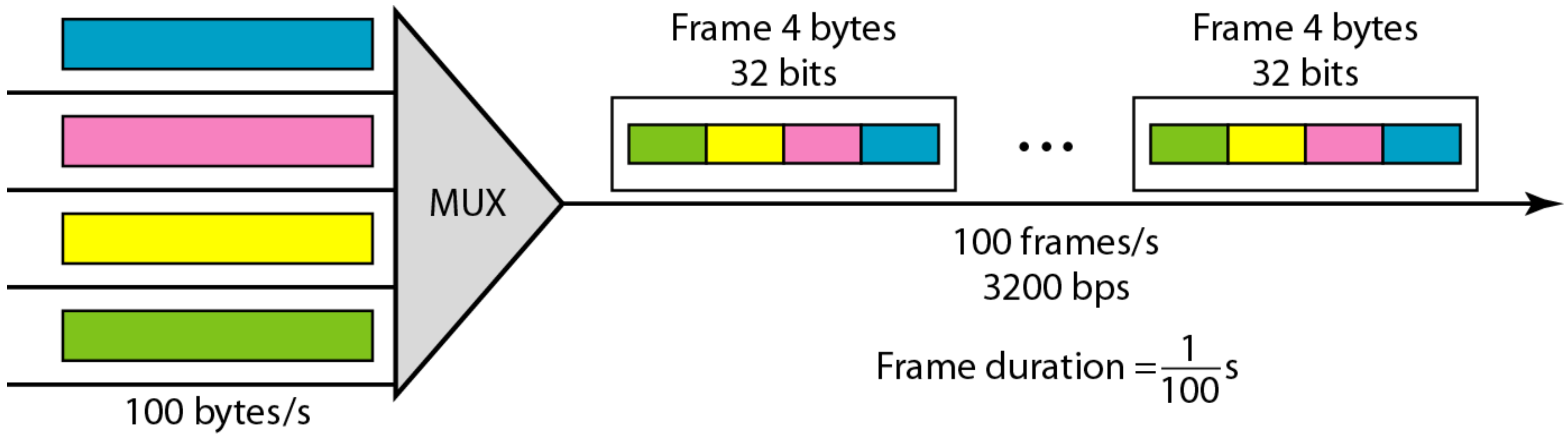
Example 6.8

Four channels are multiplexed using TDM. If each channel sends 100 bytes /s and we multiplex 1 byte per channel, show the frame traveling on the link, the size of the frame, the duration of a frame, the frame rate, and the bit rate for the link.

Solution

The multiplexer is shown in Figure 6.16. Each frame carries 1 byte from each channel; the size of each frame, therefore, is 4 bytes, or 32 bits. Because each channel is sending 100 bytes/s and a frame carries 1 byte from each channel, the frame rate must be 100 frames per second. The bit rate is 100 × 32, or 3200 bps.

Figure 6.16 *Example 6.8*



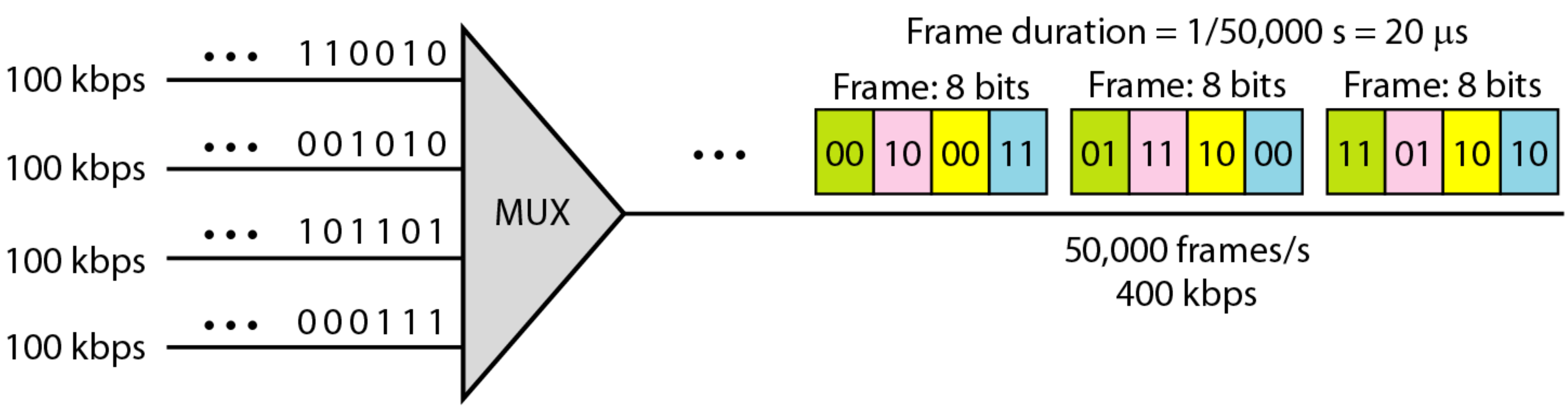
Example 6.9

A multiplexer combines four 100-kbps channels using a time slot of 2 bits. Show the output with four arbitrary inputs. What is the frame rate? What is the frame duration? What is the bit rate? What is the bit duration?

Solution

Figure 6.17 shows the output (4x100kbps) for four arbitrary inputs. The link carries 400K/(2x4)=50,000 2x4=8bit frames per second. The frame duration is therefore 1/50,000 s or 20 μs. The bit duration on the output link is 1/400,000 s, or 2.5 μs.

Figure 6.17 *Example 6.9*



Data Rate Management

- Not all input links maybe have the same data rate.
- Some links maybe slower. There maybe several different input link speeds
- There are three strategies that can be used to overcome the data rate mismatch: multilevel, multislot and pulse stuffing

Data rate matching

- Multilevel: used when the data rate of the input links are multiples of each other.
- Multislot: used when there is a GCD between the data rates. The higher bit rate channels are allocated more slots per frame, and the output frame rate is a multiple of each input link.
- Pulse Stuffing: used when there is no GCD between the links. The slowest speed link will be brought up to the speed of the other links by bit insertion, this is called pulse stuffing.

Figure 6.19 *Multilevel multiplexing*

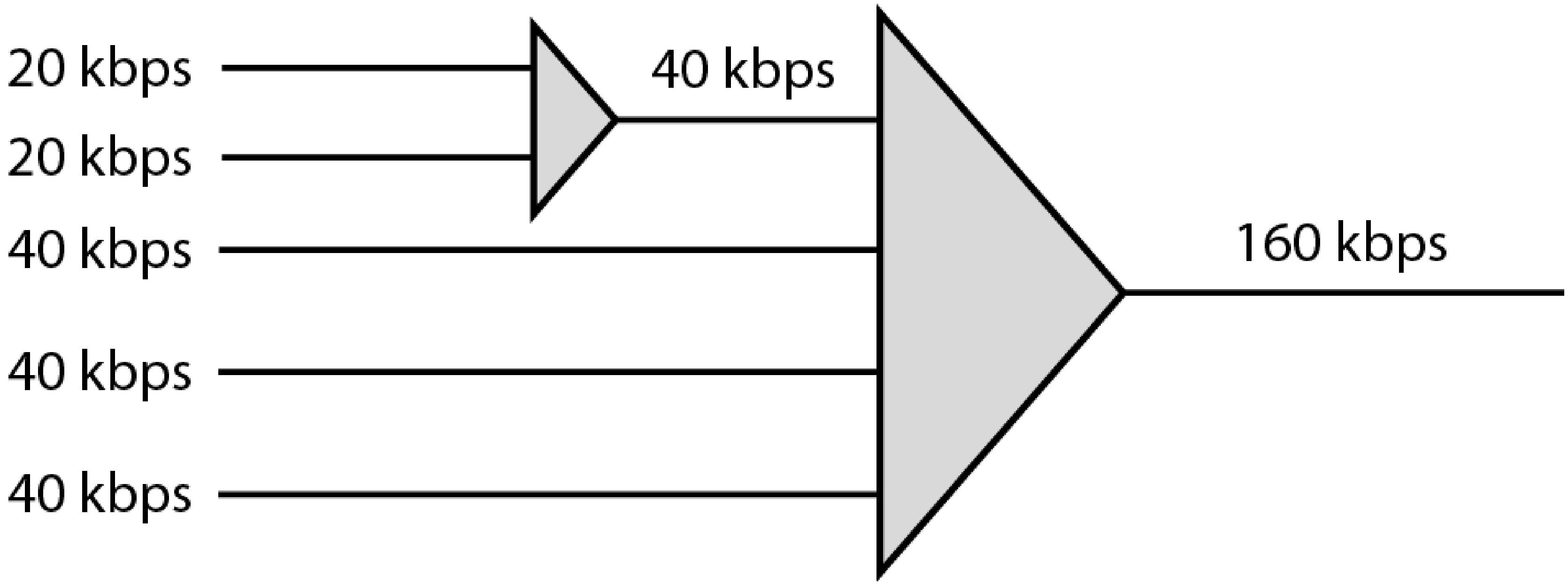


Figure 6.20 *Multiple-slot multiplexing*

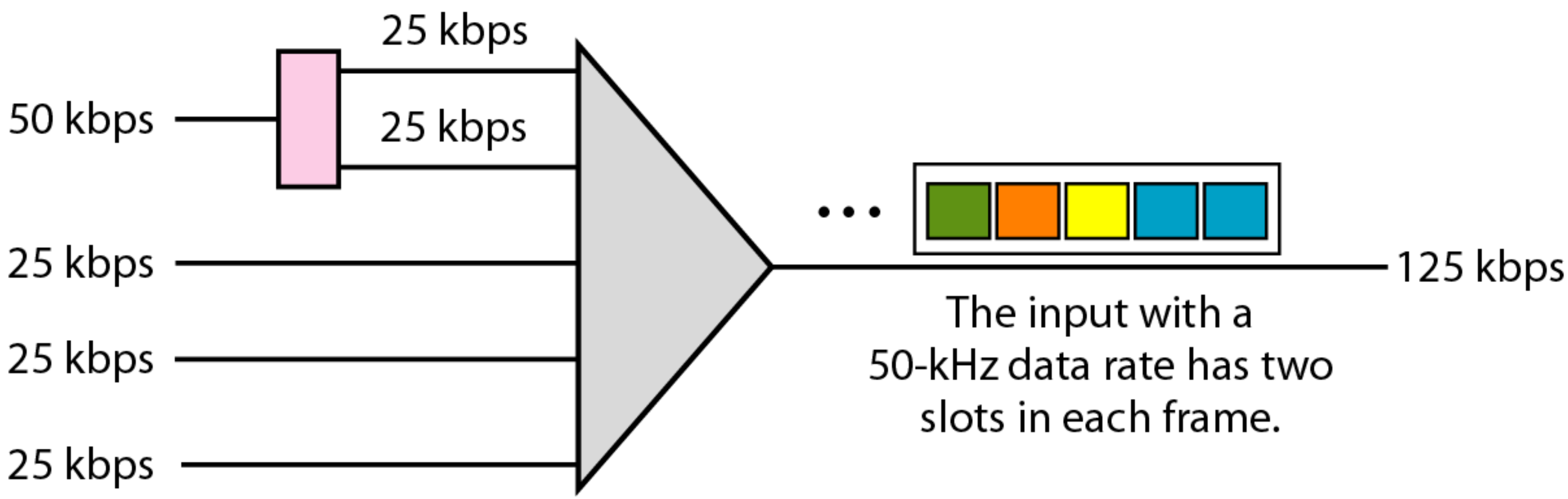
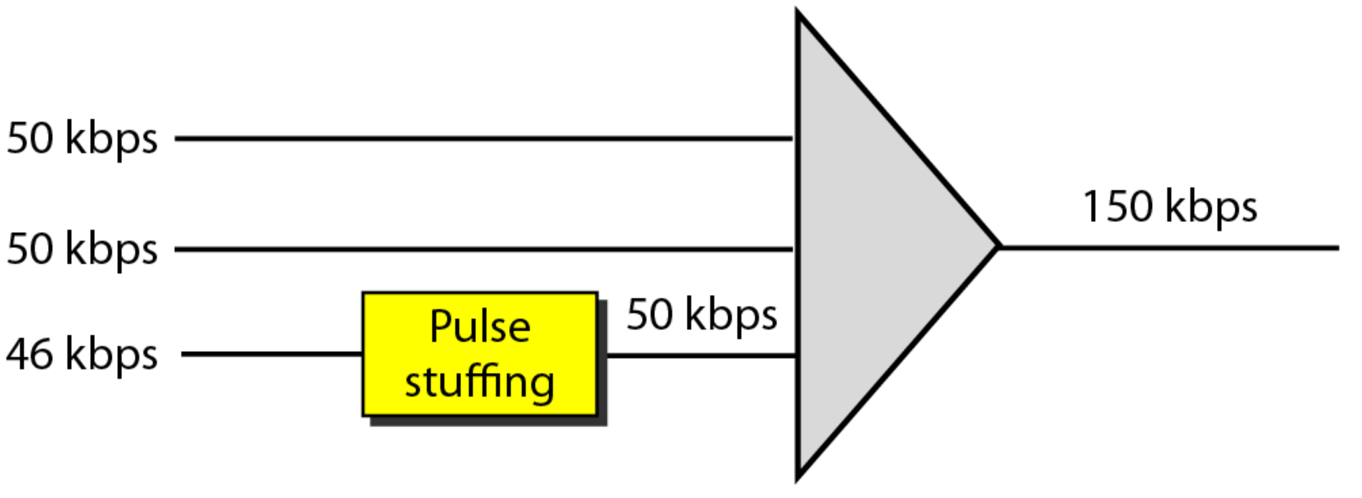


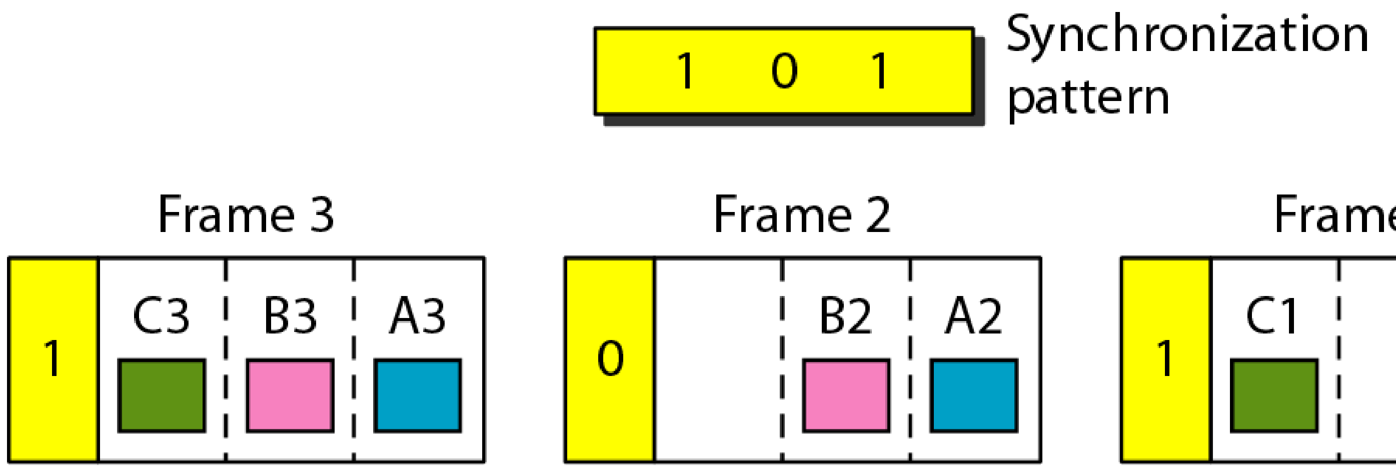
Figure 6.21 *Pulse stuffing*



Synchronization

- To ensure that the receiver correctly reads the incoming bits, i.e., knows the incoming bit boundaries to interpret a "1" and a "0", a known bit pattern is used between the frames.
- The receiver looks for the anticipated bit and starts counting bits till the end of the frame.
- Then it starts over again with the reception of another known bit.
- These bits (or bit patterns) are called synchronization bit(s).
- They are part of the overhead of transmission.

Figure 6.22 *Framing bits*



Example 6.10

We have four sources, each creating 250 8-bit characters per second. If the interleaved unit is a character and 1 synchronizing bit is added to each frame, find (a) the data rate of each source, (b) the duration of each character in each source, (c) the frame rate, (d) the duration of each frame, (e) the number of bits in each frame, and (f) the data rate of the link.

Solution

We can answer the questions as follows:

a. The data rate of each source is 250 × 8 = 2000 bps = 2 kbps.

Example 6.10 (continued)

b. *Each source sends 250 characters per second; therefore, the duration of a character is 1/250 s, or 4 ms.*

c. *Each frame has one character from each source, which means the link needs to send 250 frames per second to keep the transmission rate of each source.*

d. *The duration of each frame is 1/250 s, or 4 ms. Note that the duration of each frame is the same as the duration of each character coming from each source.*

e. *Each frame carries 4 characters and 1 extra synchronizing bit. This means that each frame is* $4 \times 8 + 1 = 33$ *bits.*

Example 6.11

Two channels, one with a bit rate of 100 kbps and another with a bit rate of 200 kbps, are to be multiplexed. How this can be achieved? What is the frame rate? What is the frame duration? What is the bit rate of the link?

Solution

We can allocate one slot to the first channel and two slots to the second channel. Each frame carries 3 bits. The frame rate is 100,000 frames per second because it carries 1 bit from the first channel. The bit rate is 100,000 frames/s × 3 bits per frame, or 300 kbps.

Figure 6.23 *Digital hierarchy*

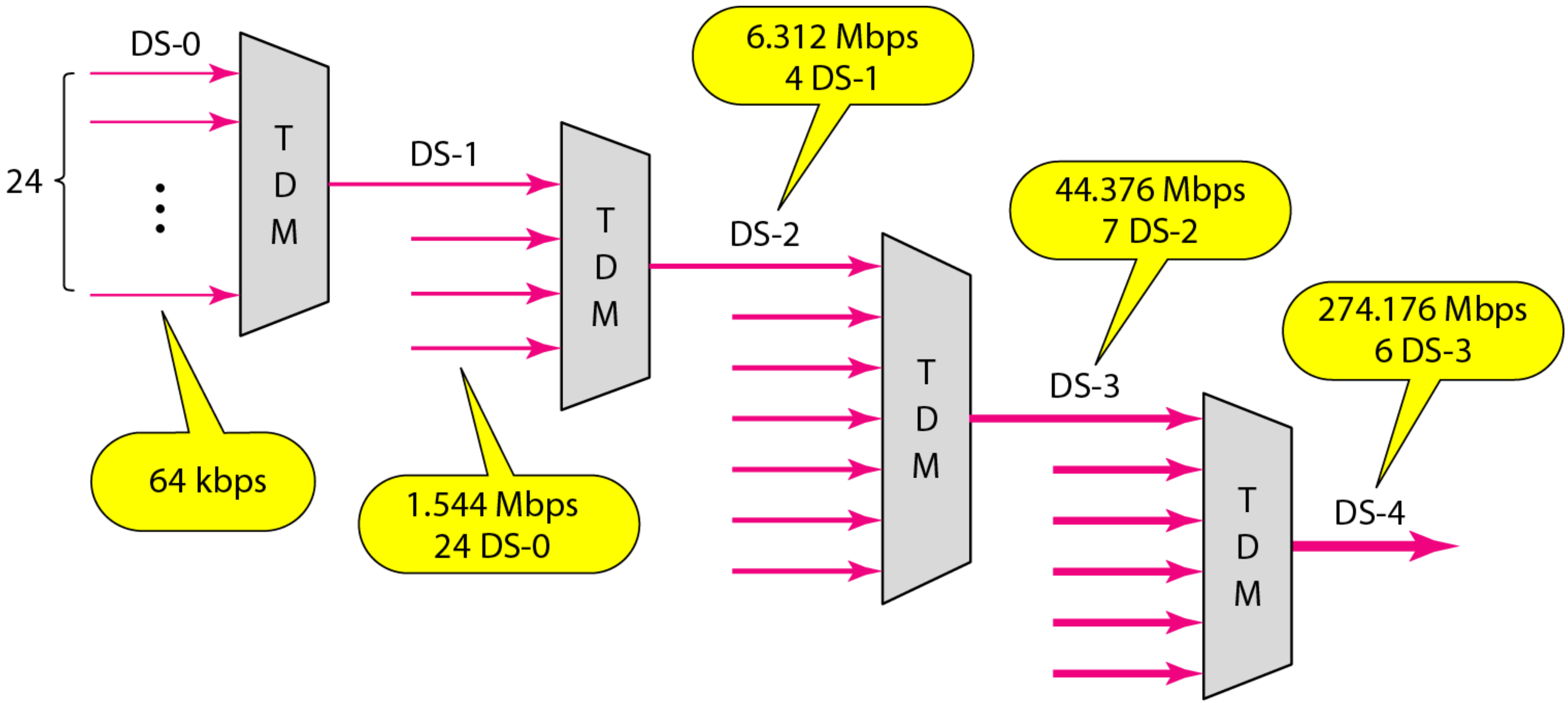


Table 6.1 *DS and T line rates*

Service	*Line*	*Rate (Mbps)*	*Voice Channels*
DS-1	T-1	1.544	24
DS-2	T-2	6.312	96
DS-3	T-3	44.736	672
DS-4	T-4	274.176	4032

Figure 6.24 *T-1 line for multiplexing telephone lines*

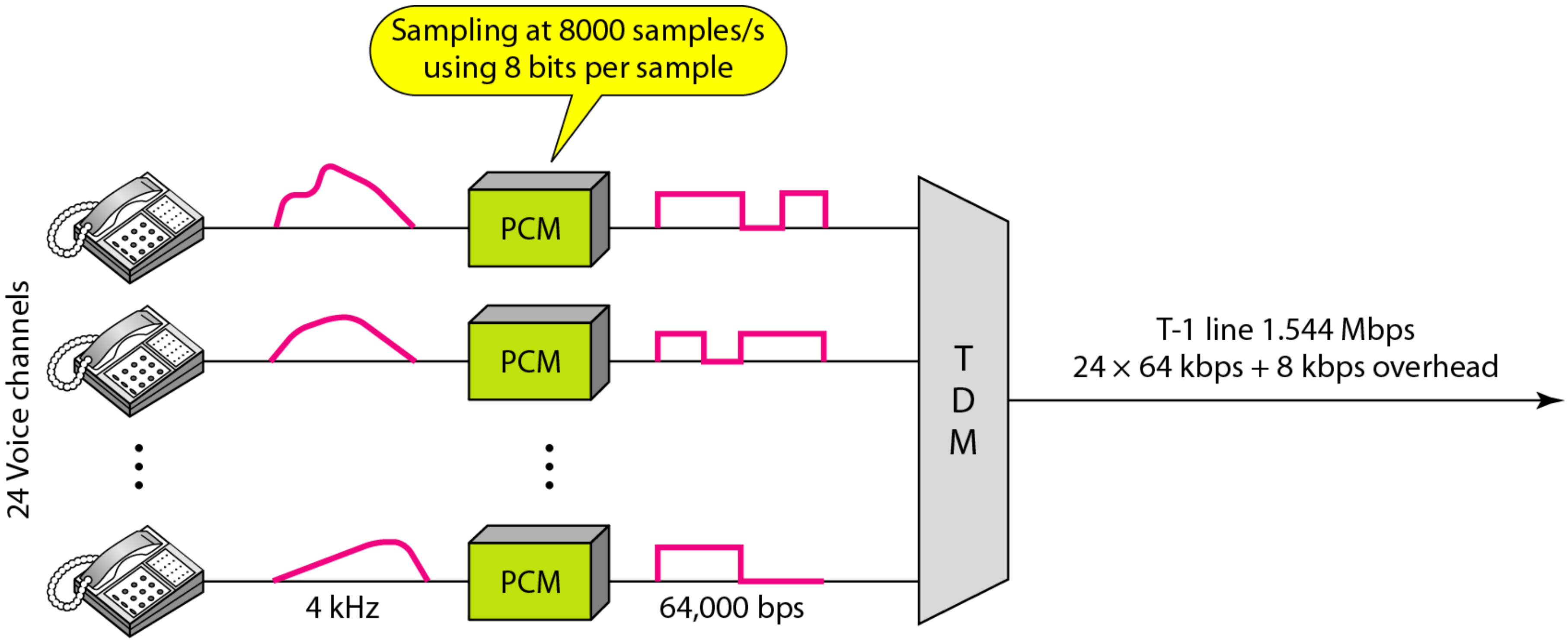


Figure 6.25 *T-1 frame structure*

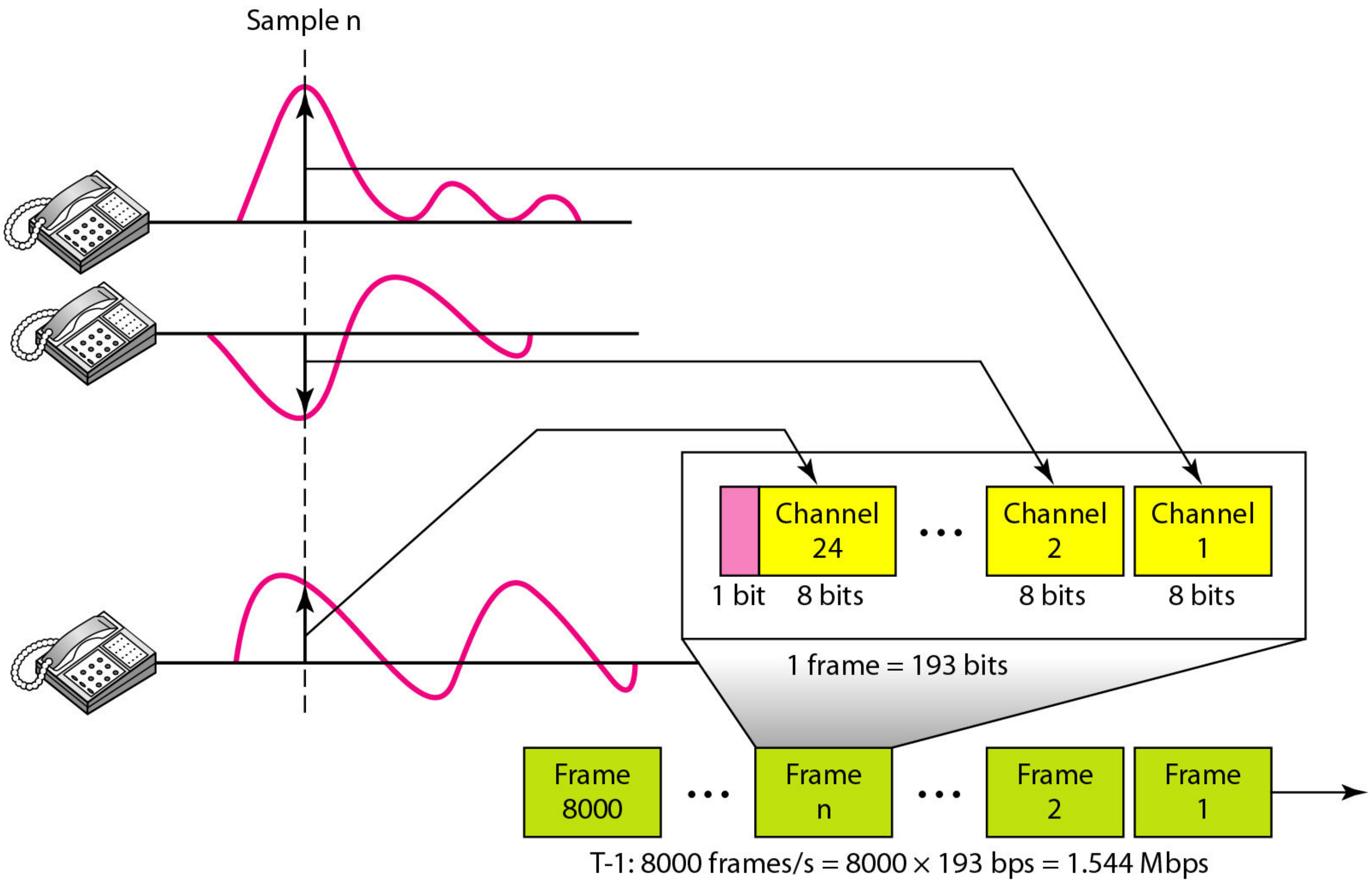


Table 6.2 ***E line rates***

Line	*Rate (Mbps)*	*Voice Channels*
E-1	2.048	30
E-2	8.448	120
E-3	34.368	480
E-4	139.264	1920

Inefficient use of Bandwidth

- Sometimes an input link may have no data to transmit.
- When that happens, one or more slots on the output link will go unused.
- That is wasteful of bandwidth.

Figure 6.18 *Empty slots*

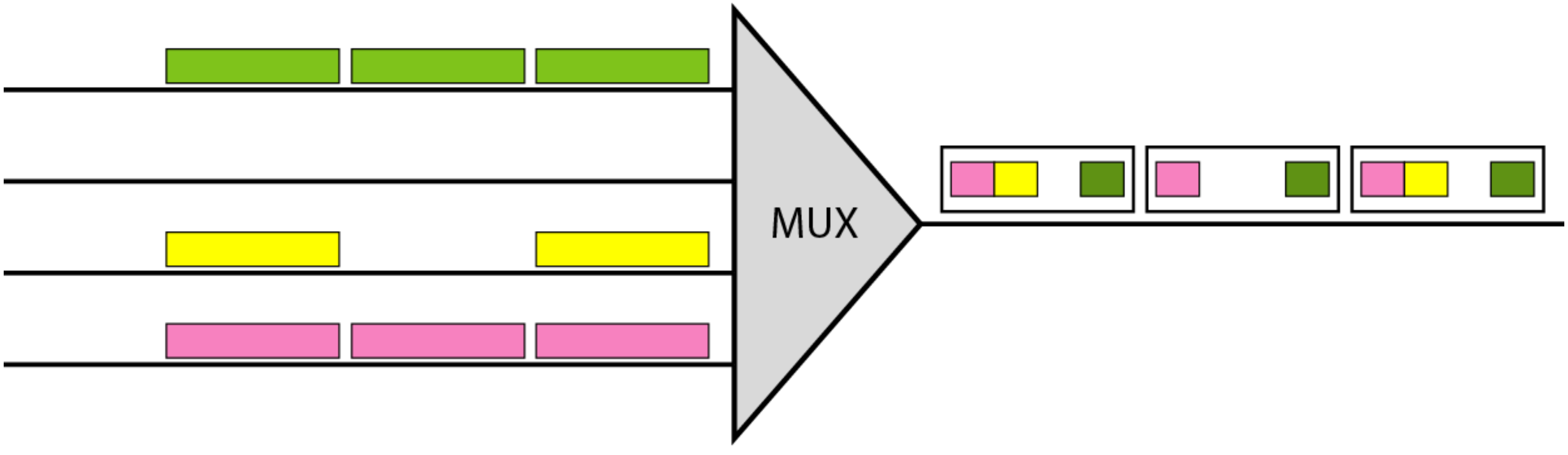
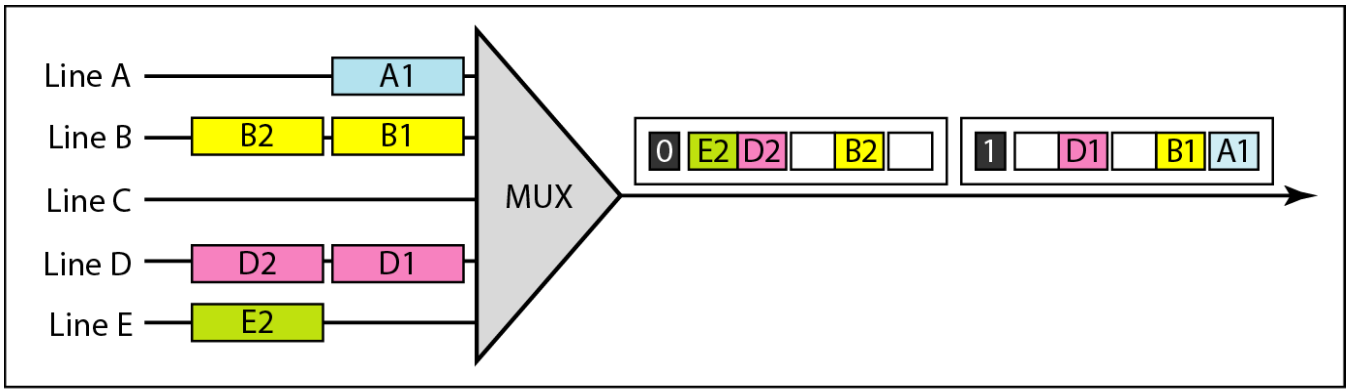
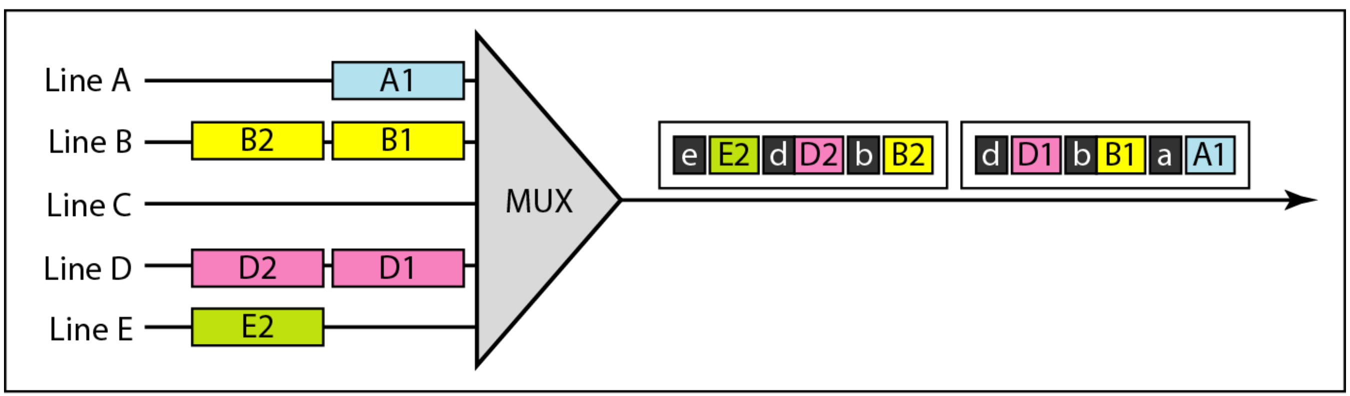


Figure 6.26 *TDM slot comparison*



a. Synchronous TDM



b. Statistical TDM

Chapter 6

Bandwidth Utilization: Multiplexing and Spreading

6-1 SPREAD SPECTRUM

In spread spectrum (SS), we combine signals from different sources to fit into a larger bandwidth, but our goals are to prevent eavesdropping and jamming. To achieve these goals, spread spectrum techniques add redundancy.

Topics discussed in this section:

- **Frequency Hopping Spread Spectrum (FHSS)**
- **Direct Sequence Spread Spectrum (DSSS)**

Spread Spectrum

- A signal that occupies a bandwidth of B, is spread out to occupy a bandwidth of B_{ss}
- All signals are spread to occupy the same bandwidth B_{ss}
- Signals are spread with different codes so that they can be separated at the receivers.
- Signals can be spread in the frequency domain or in the time domain.

Figure 6.27 *Spread spectrum*

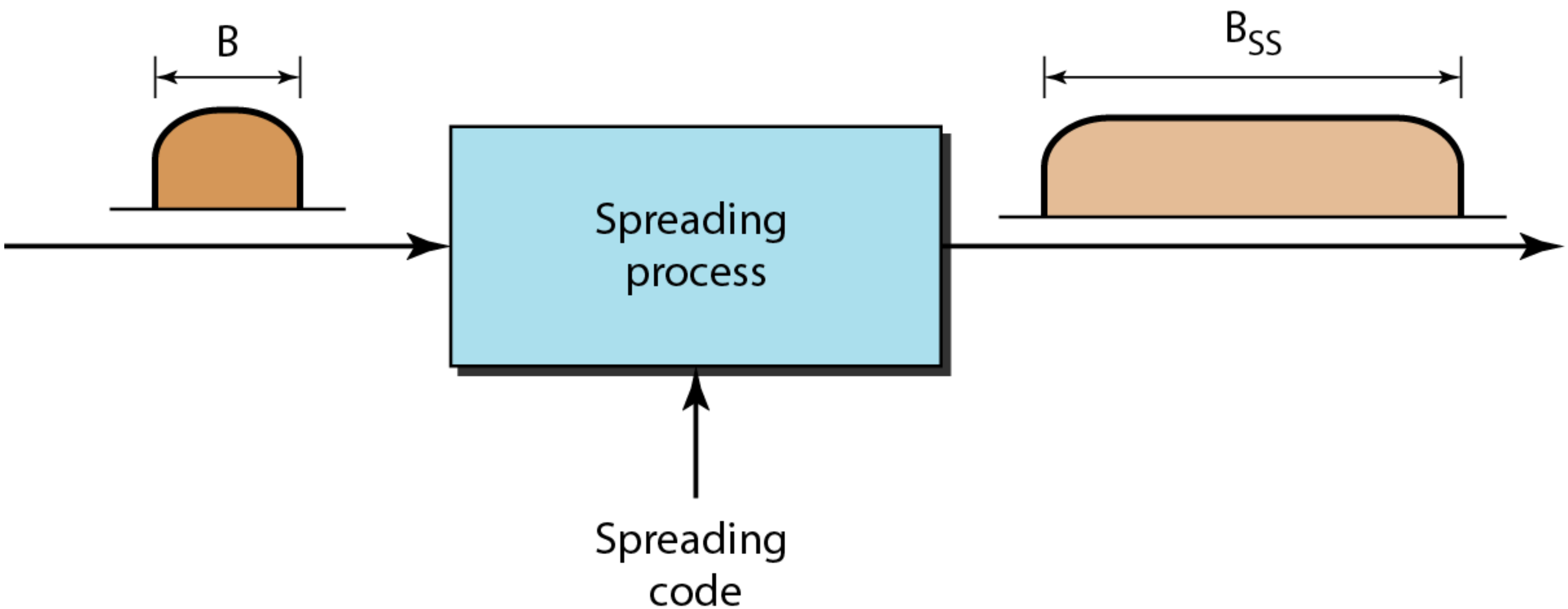


Figure 6.28 *Frequency hopping spread spectrum (FHSS)*

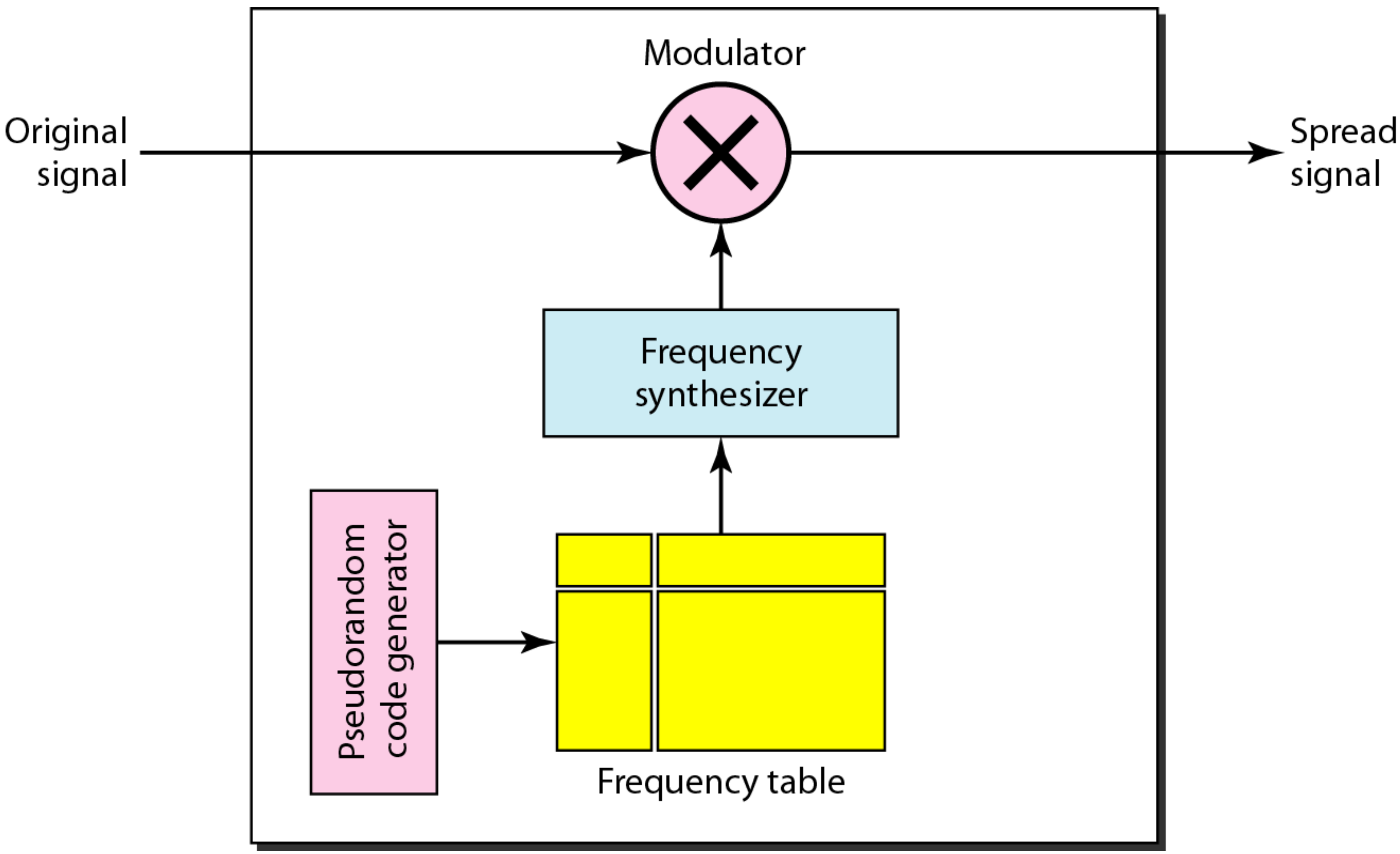


Figure 6.29 ***Frequency selection in FHSS***

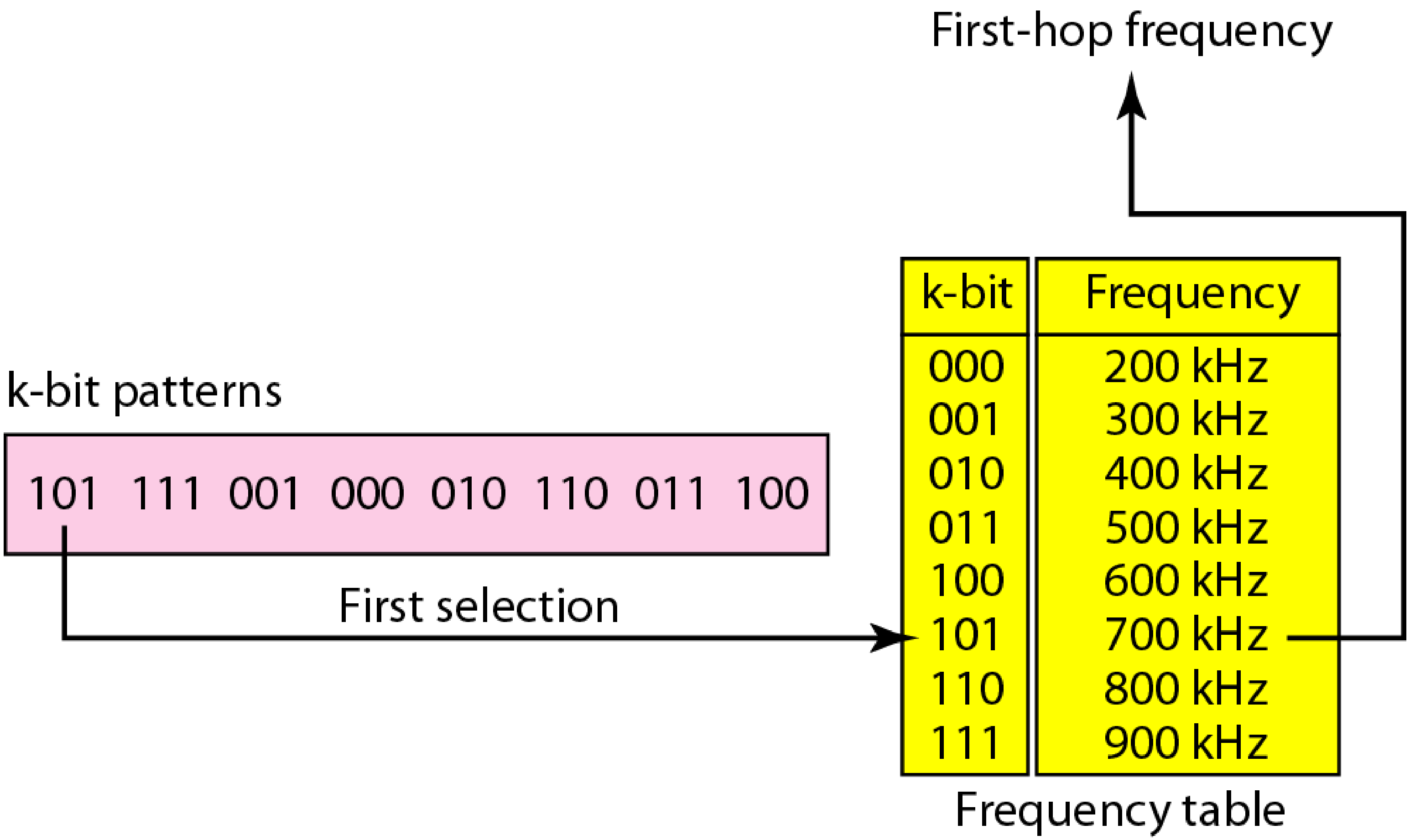


Figure 6.30 *FHSS cycles*

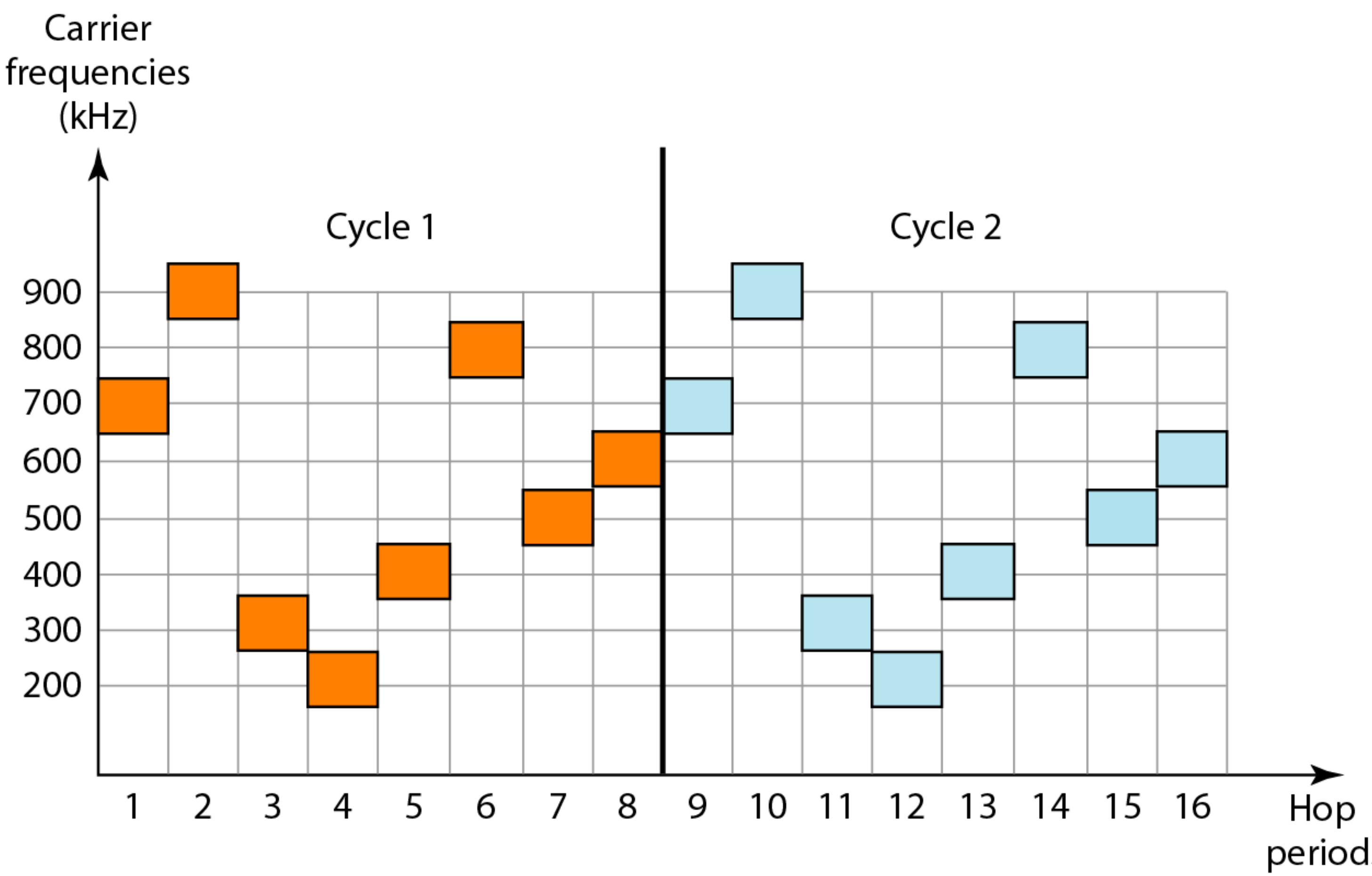


Figure 6.31 *Bandwidth sharing*

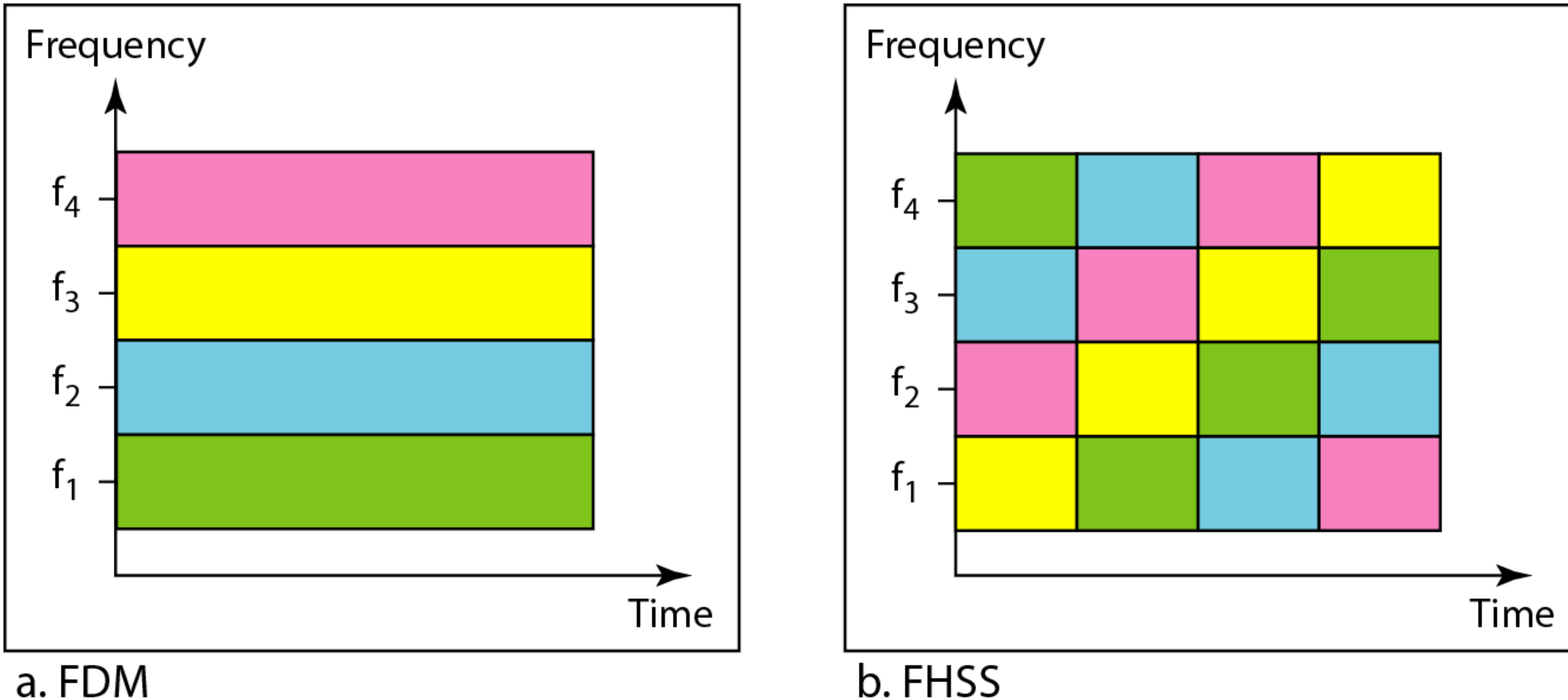


Figure 6.32 *DSSS*

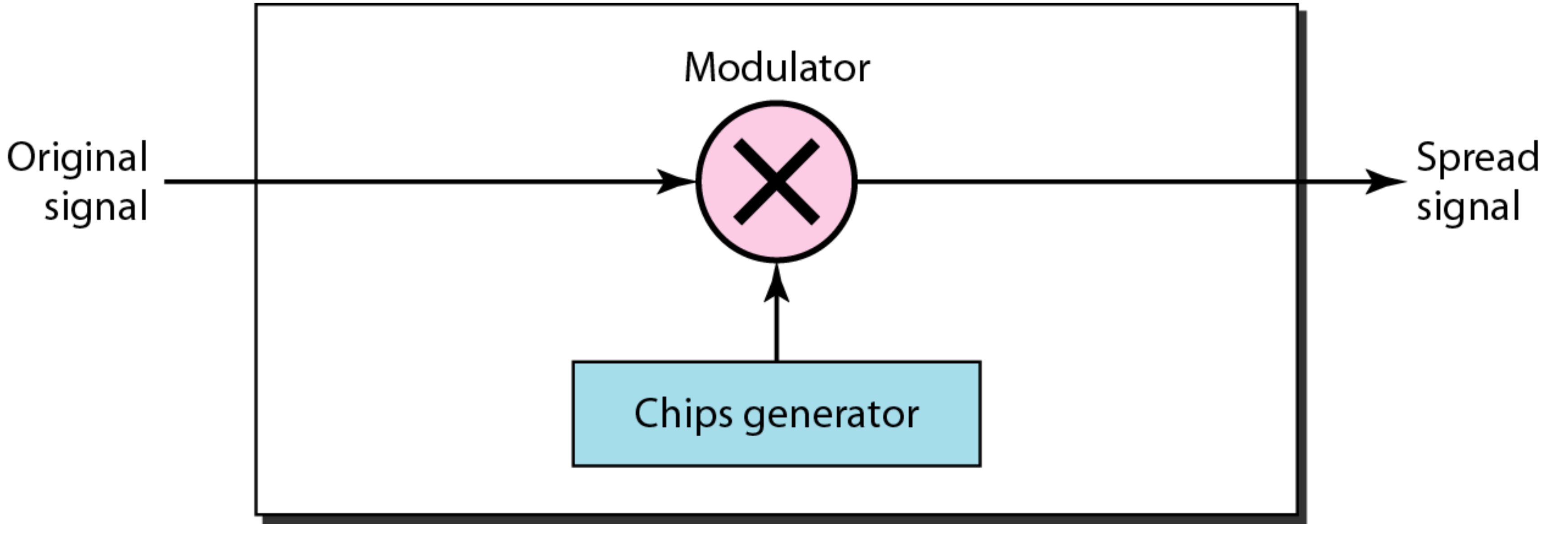
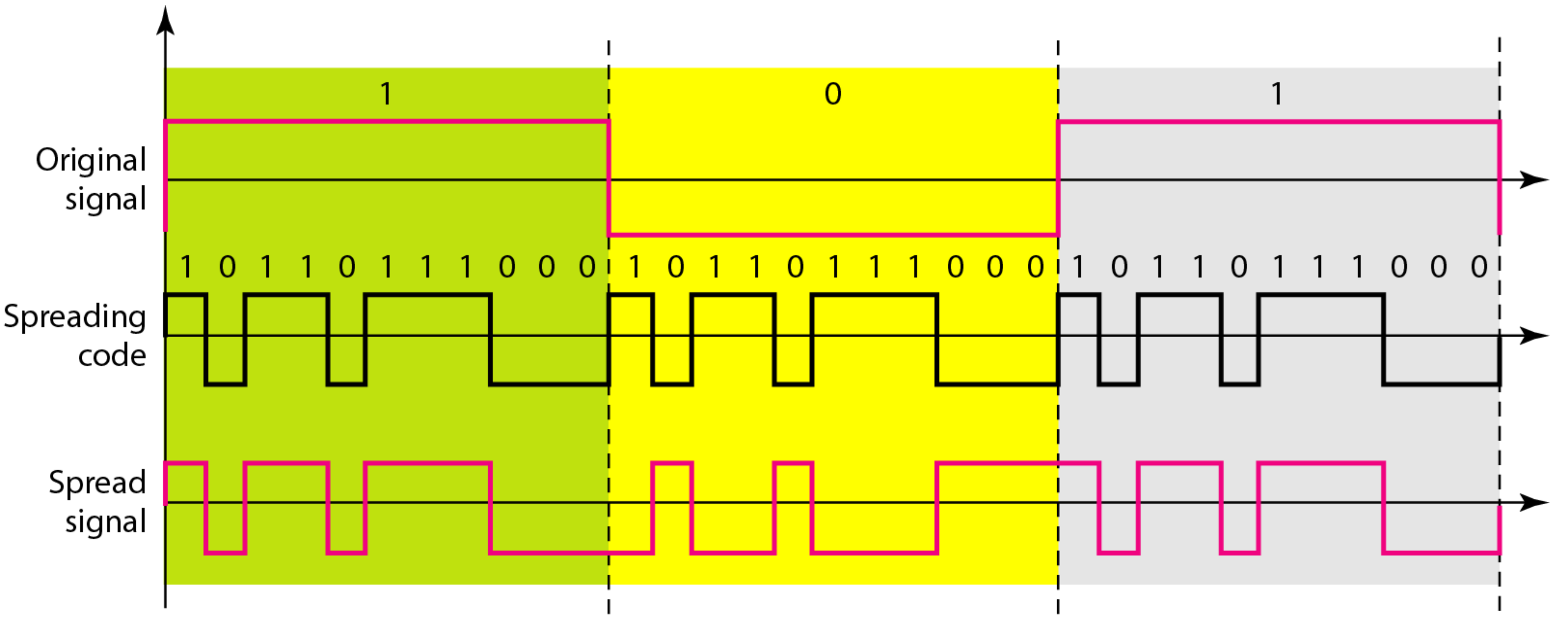


Figure 6.33 *DSSS example*



Chapter 7

Transmission Media

Figure 7.1 *Transmission medium and physical layer*

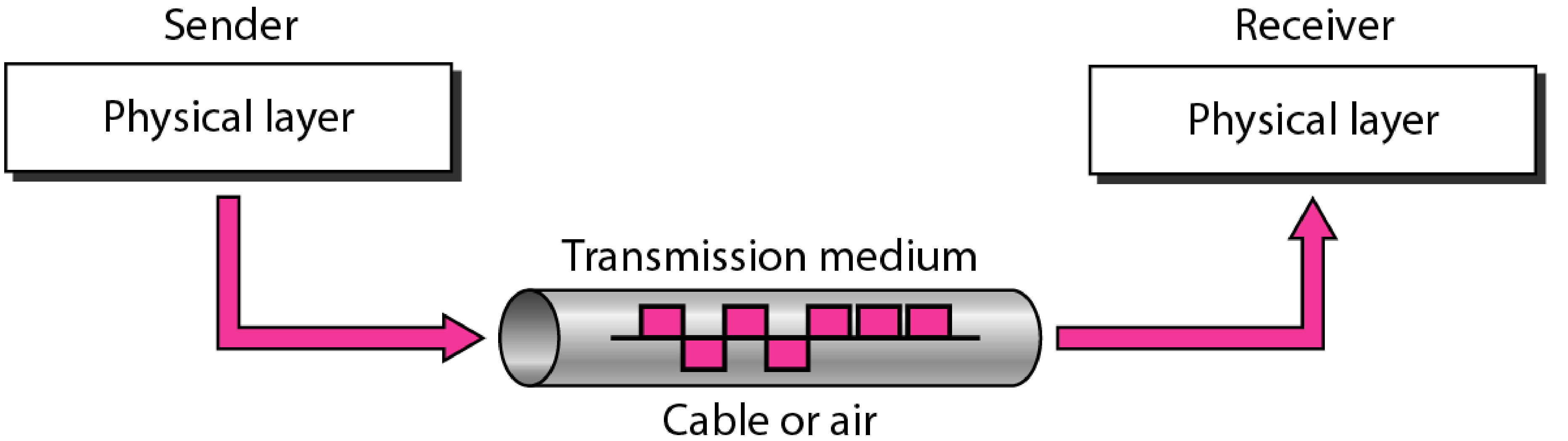
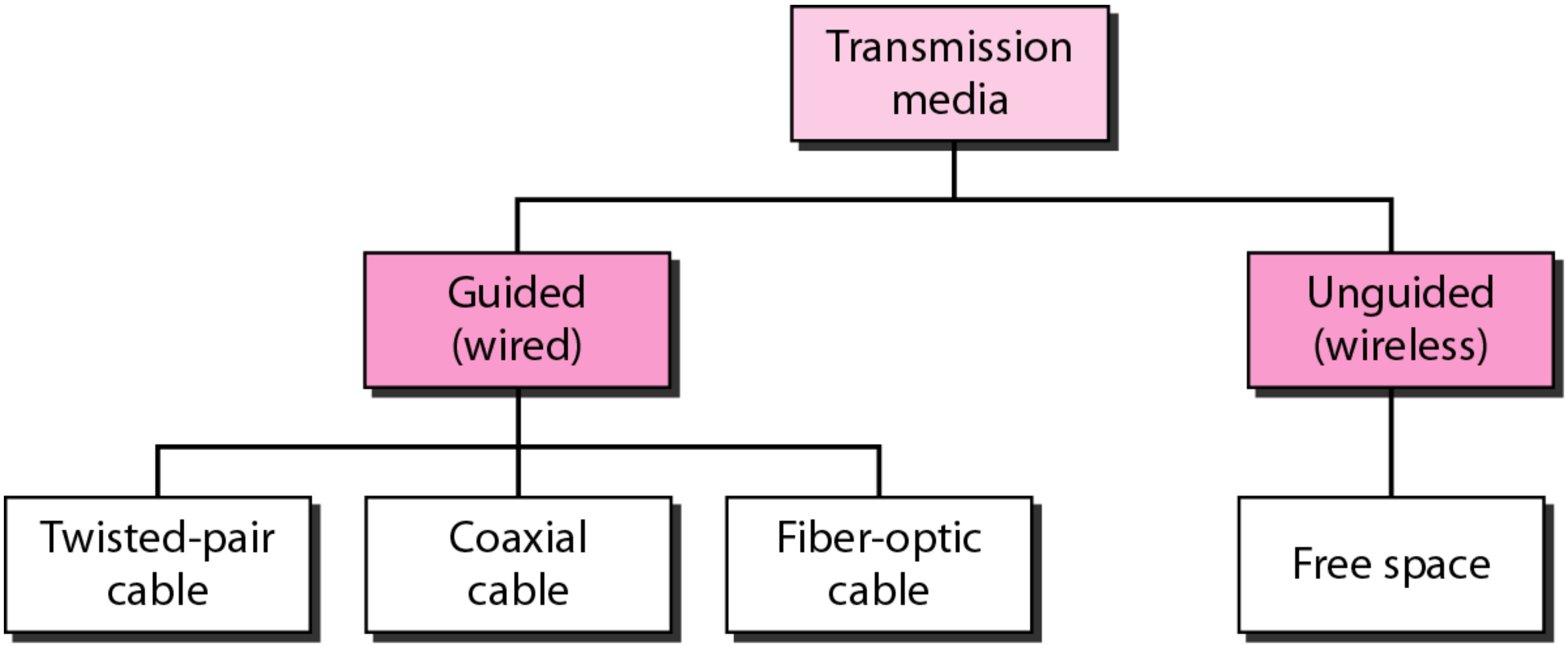


Figure 7.2 *Classes of transmission media*



7-1 GUIDED MEDIA

Guided media, which are those that provide a conduit from one device to another, include twisted-pair cable, coaxial cable, and fiber-optic cable.

Topics discussed in this section:

Twisted-Pair Cable
Coaxial Cable
Fiber-Optic Cable

Figure 7.3 *Twisted-pair cable*

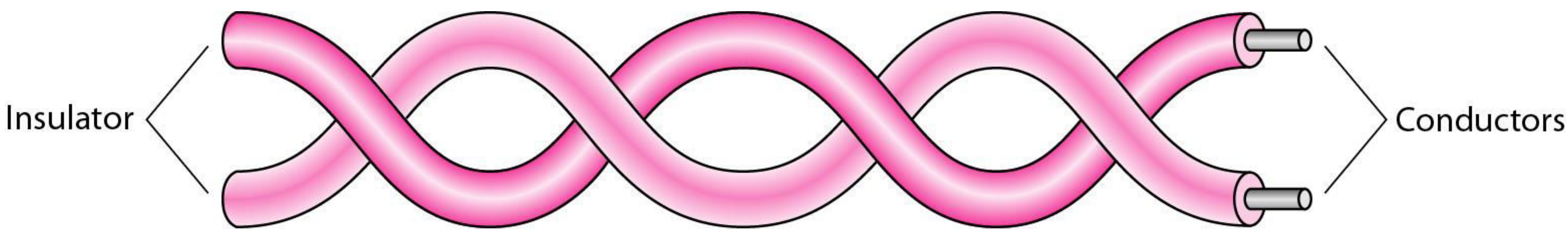
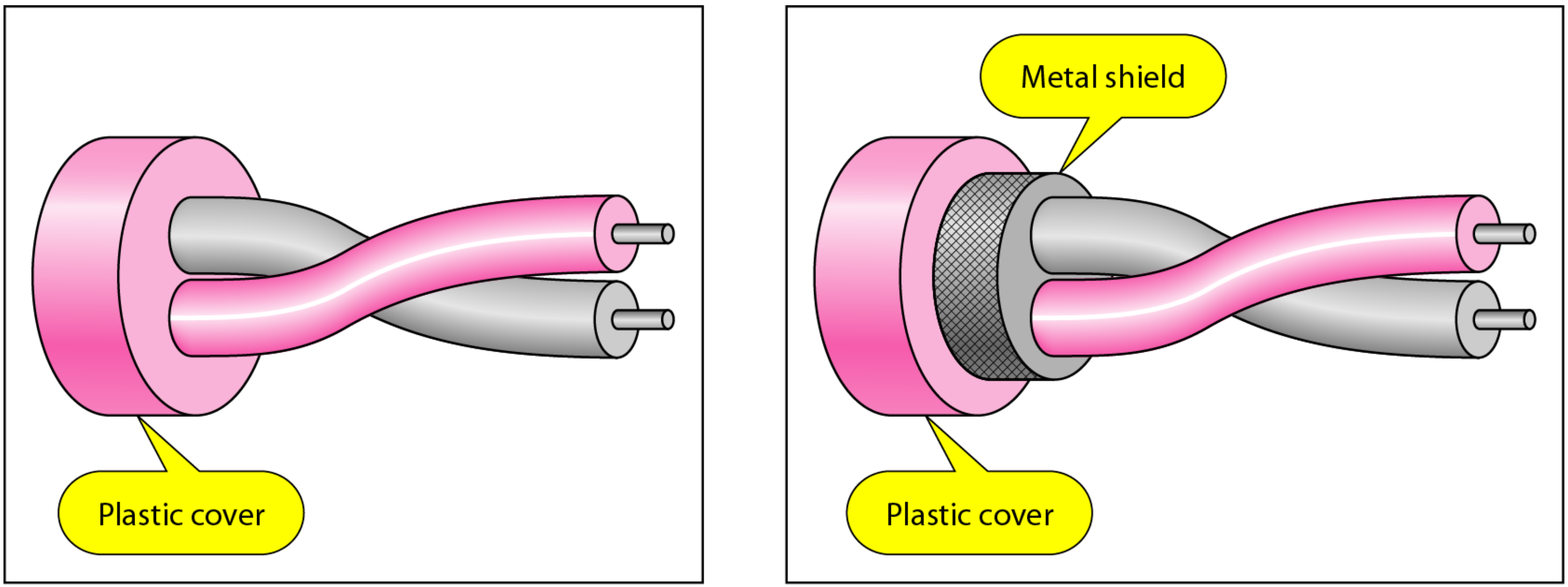


Figure 7.4 *UTP and STP cables*



a. UTP b. STP

Table 7.1 ***Categories of unshielded twisted-pair cables***

Category	*Specification*	*Data Rate (Mbps)*	*Use*
1	Unshielded twisted-pair used in telephone	< 0.1	Telephone
2	Unshielded twisted-pair originally used in T-lines	2	T-1 lines
3	Improved CAT 2 used in LANs	10	LANs
4	Improved CAT 3 used in Token Ring networks	20	LANs
5	Cable wire is normally 24 AWG with a jacket and outside sheath	100	LANs
5E	An extension to category 5 that includes extra features to minimize the crosstalk and electromagnetic interference	125	LANs
6	A new category with matched components coming from the same manufacturer. The cable must be tested at a 200-Mbps data rate.	200	LANs
7	Sometimes called SSTP (shielded screen twisted-pair). Each pair is individually wrapped in a helical metallic foil followed by a metallic foil shield in addition to the outside sheath. The shield decreases the effect of crosstalk and increases the data rate.	600	LANs

Figure 7.5 *UTP connector*

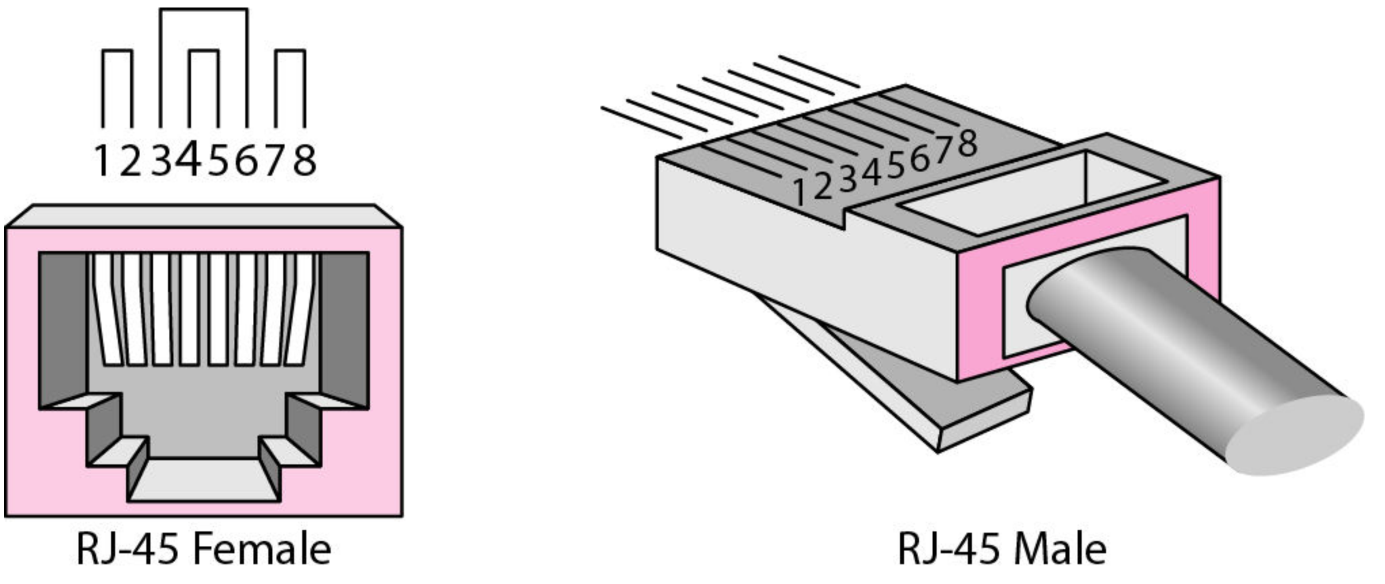


Figure 7.6 *UTP performance*

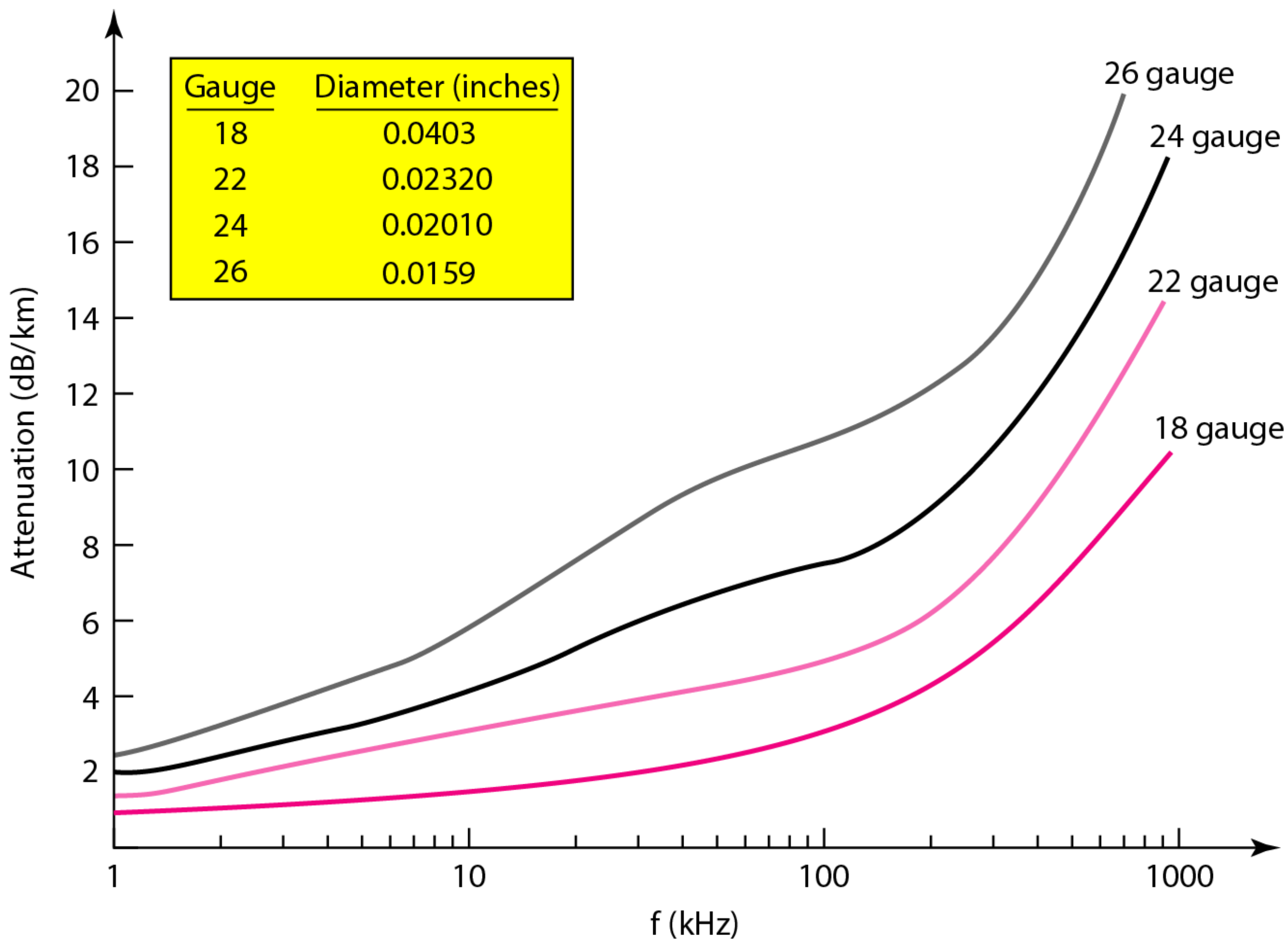


Figure 7.7 *Coaxial cable*

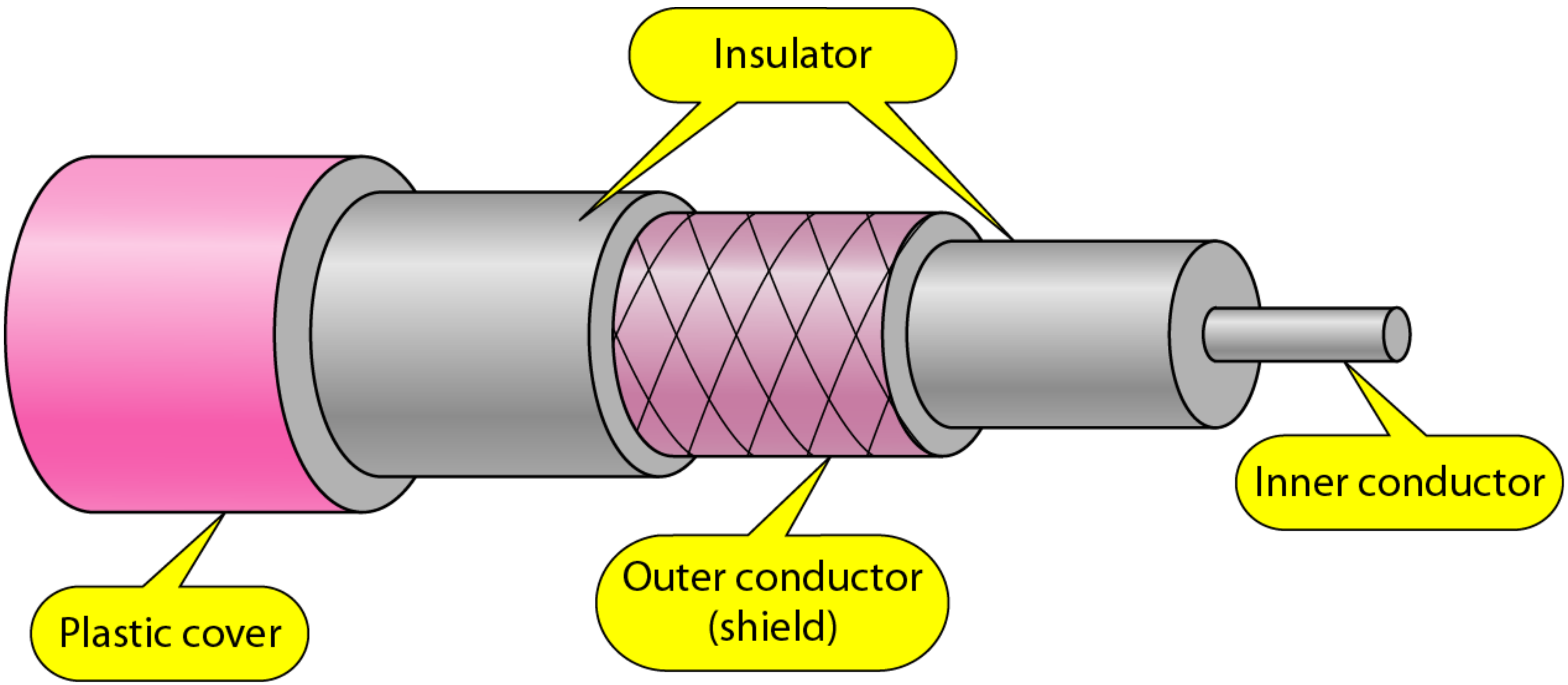


Table 7.2 ***Categories of coaxial cables***

Category	*Impedance*	*Use*
RG-59	75 Ω	Cable TV
RG-58	50 Ω	Thin Ethernet
RG-11	50 Ω	Thick Ethernet

Figure 7.8 *BNC connectors*

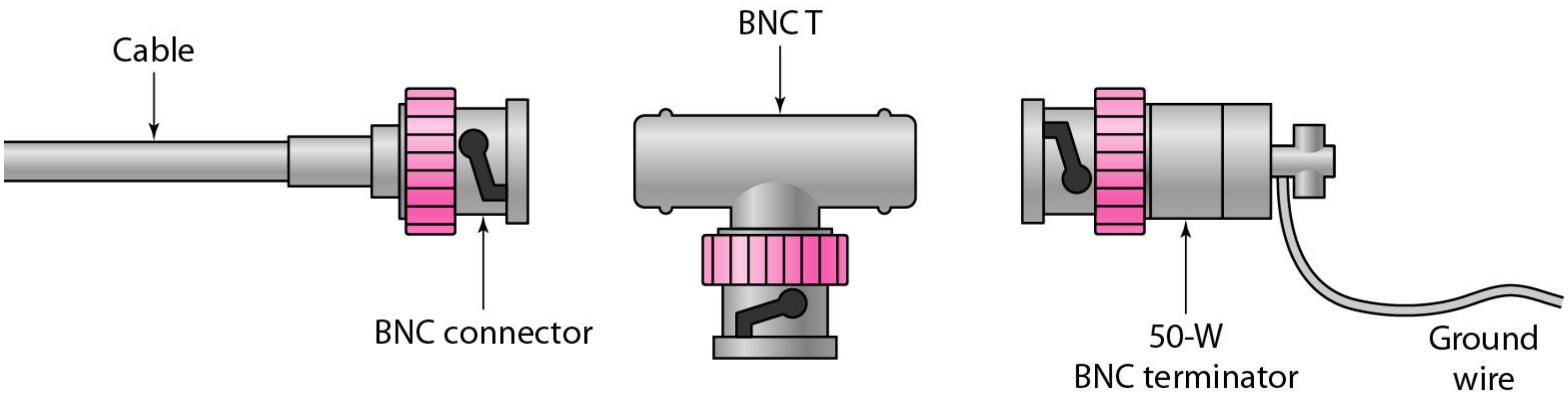


Figure 7.9 *Coaxial cable performance*

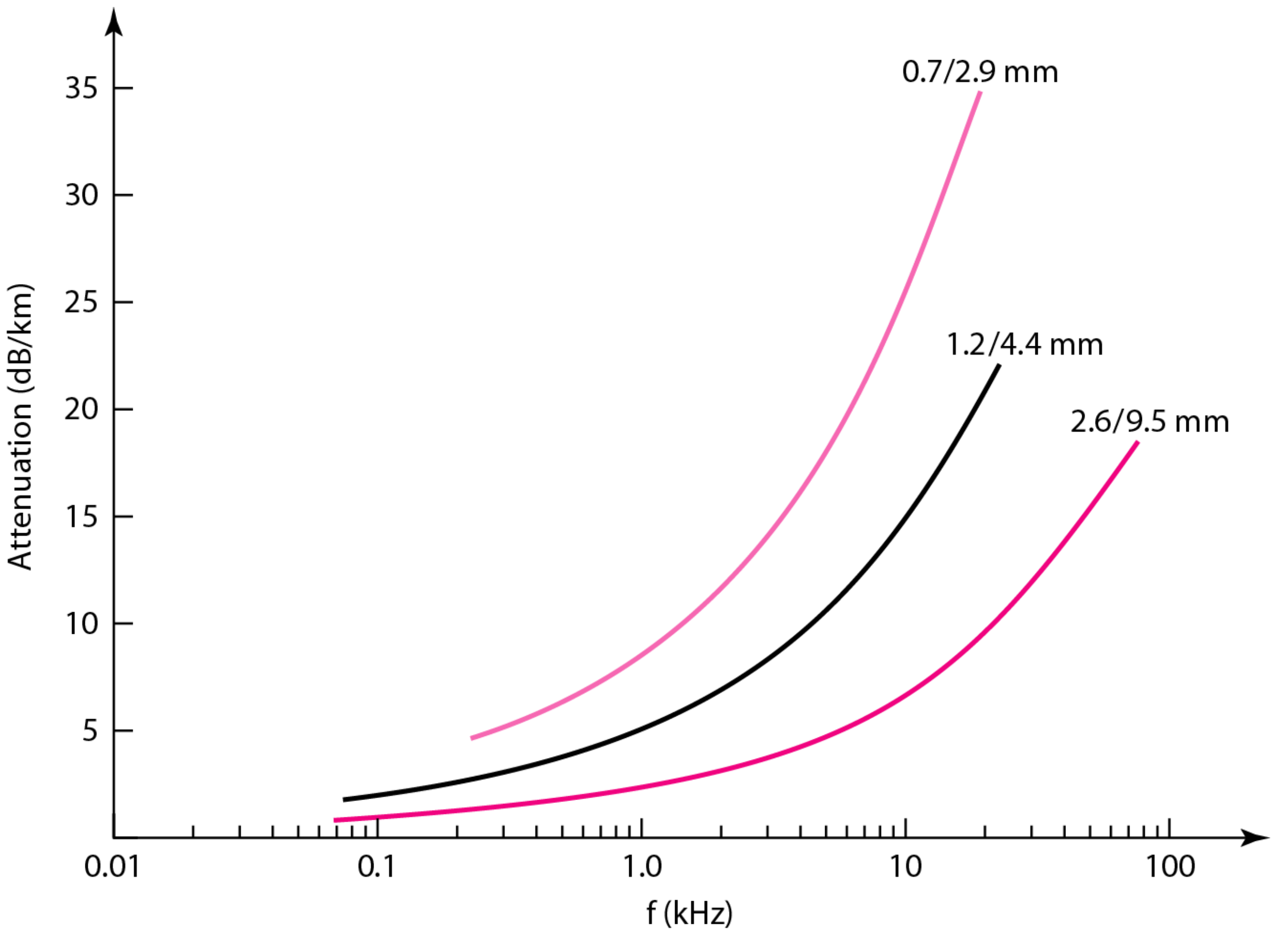


Figure 7.10 Fiber optics: *Bending of light ray*

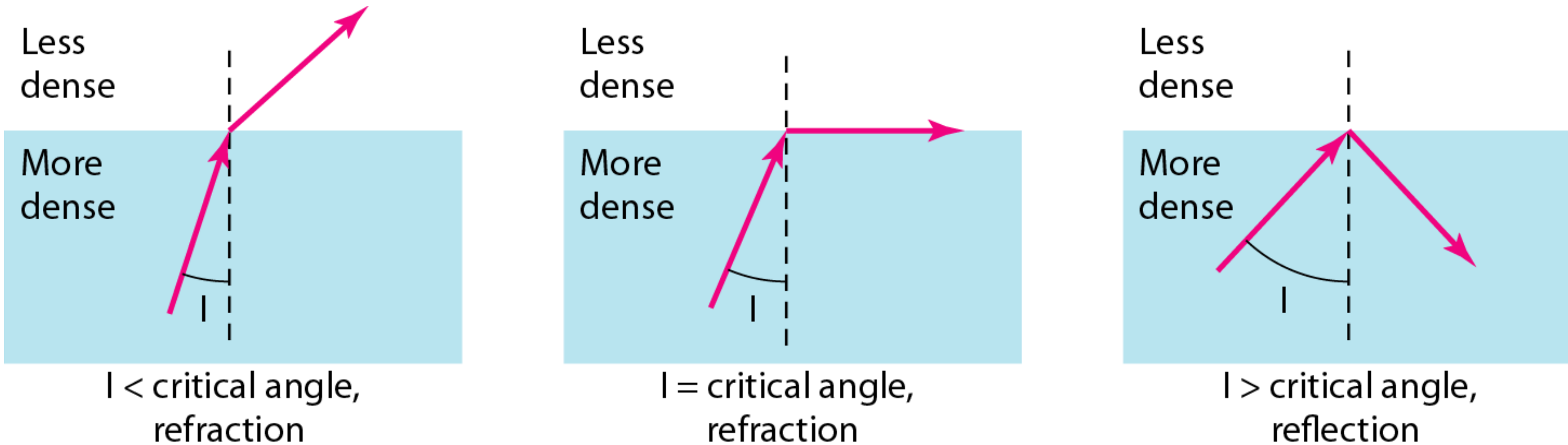


Figure 7.11 *Optical fiber*

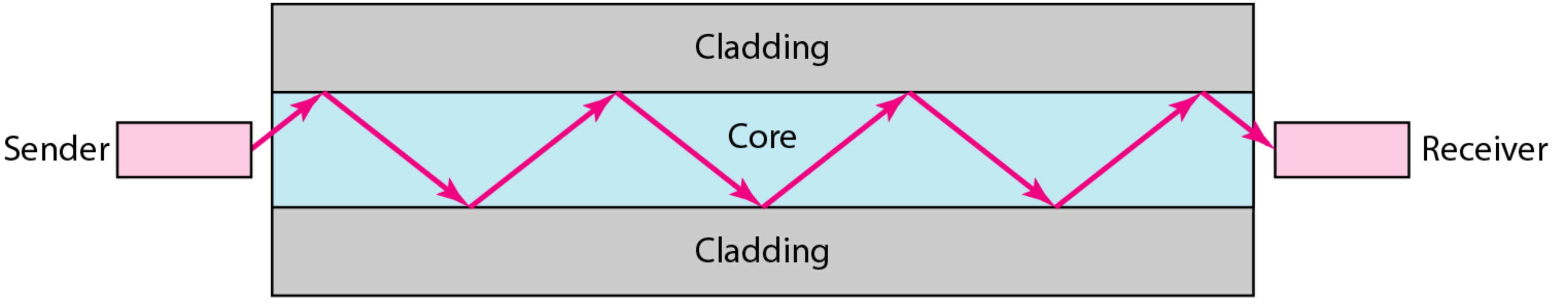


Figure 7.12 *Propagation modes*

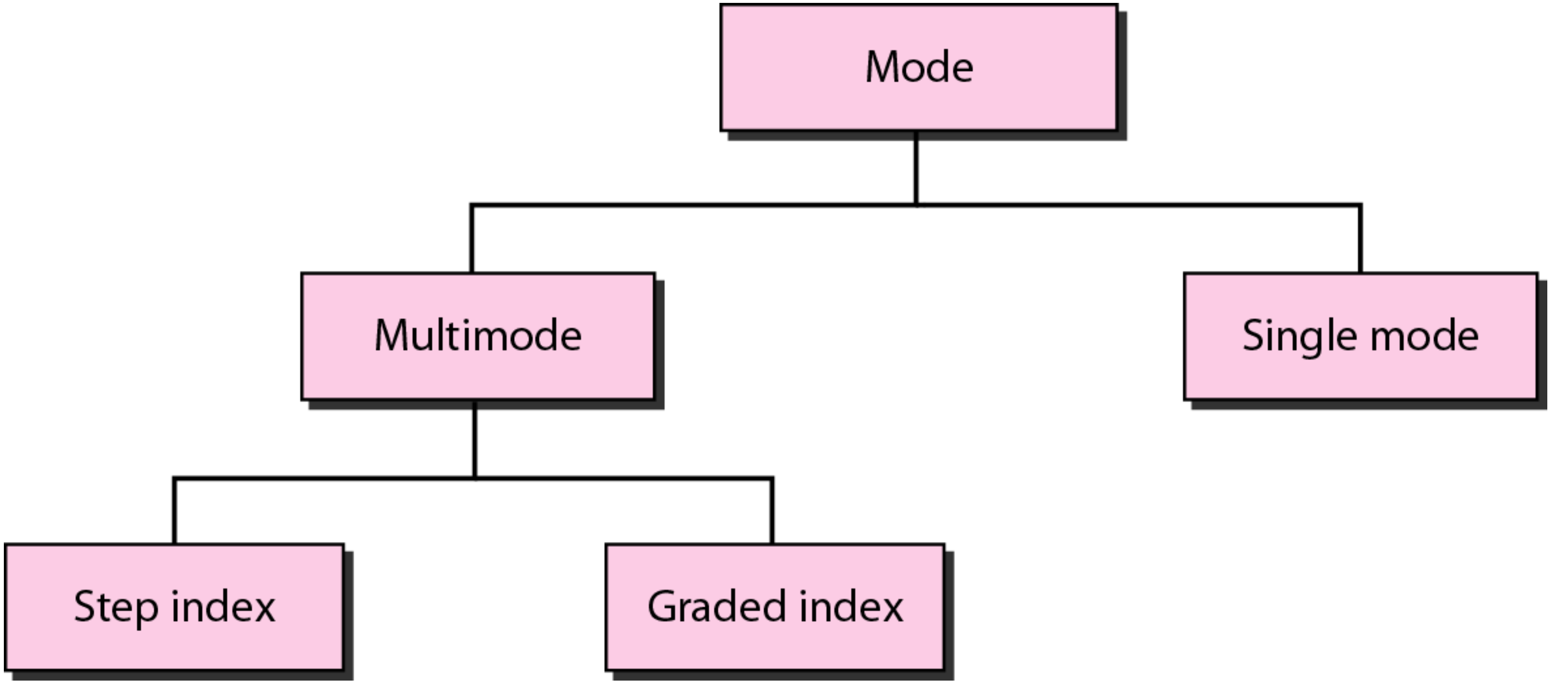


Figure 7.13 *Modes*

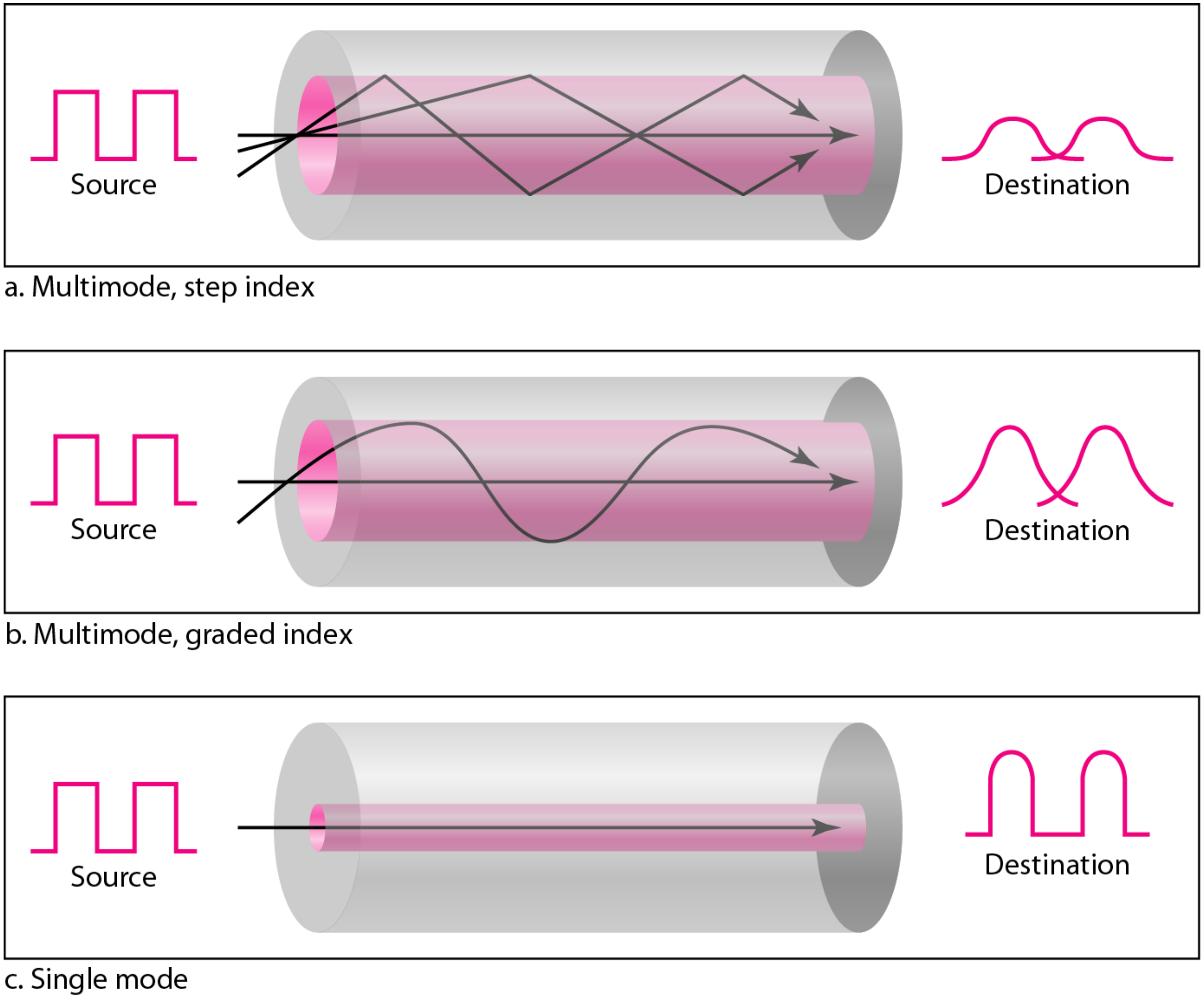


Table 7.3 ***Fiber types***

Type	*Core (μm)*	*Cladding (μm)*	*Mode*
50/125	50.0	125	Multimode, graded index
62.5/125	62.5	125	Multimode, graded index
100/125	100.0	125	Multimode, graded index
7/125	7.0	125	Single mode

Figure 7.14 *Fiber construction*

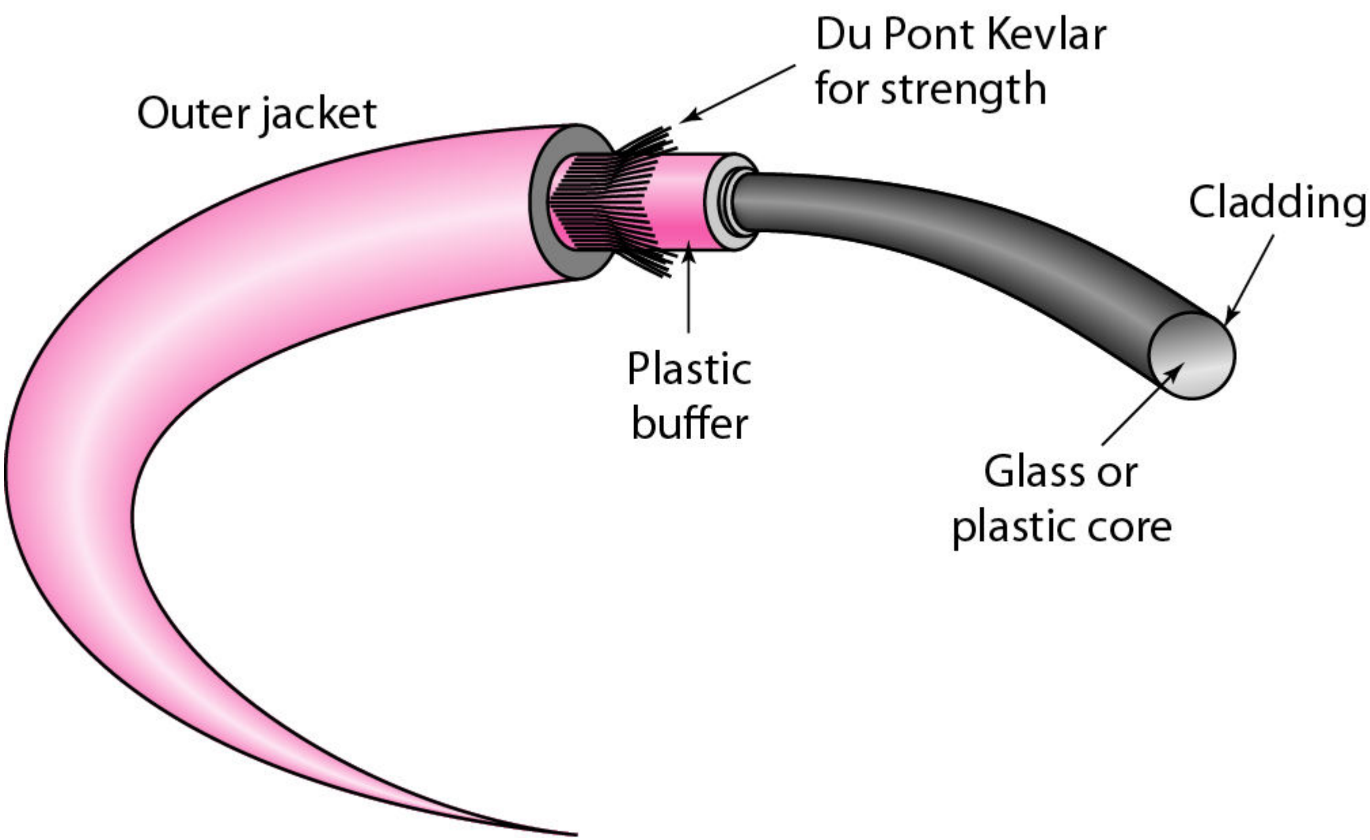


Figure 7.15 *Fiber-optic cable connectors*

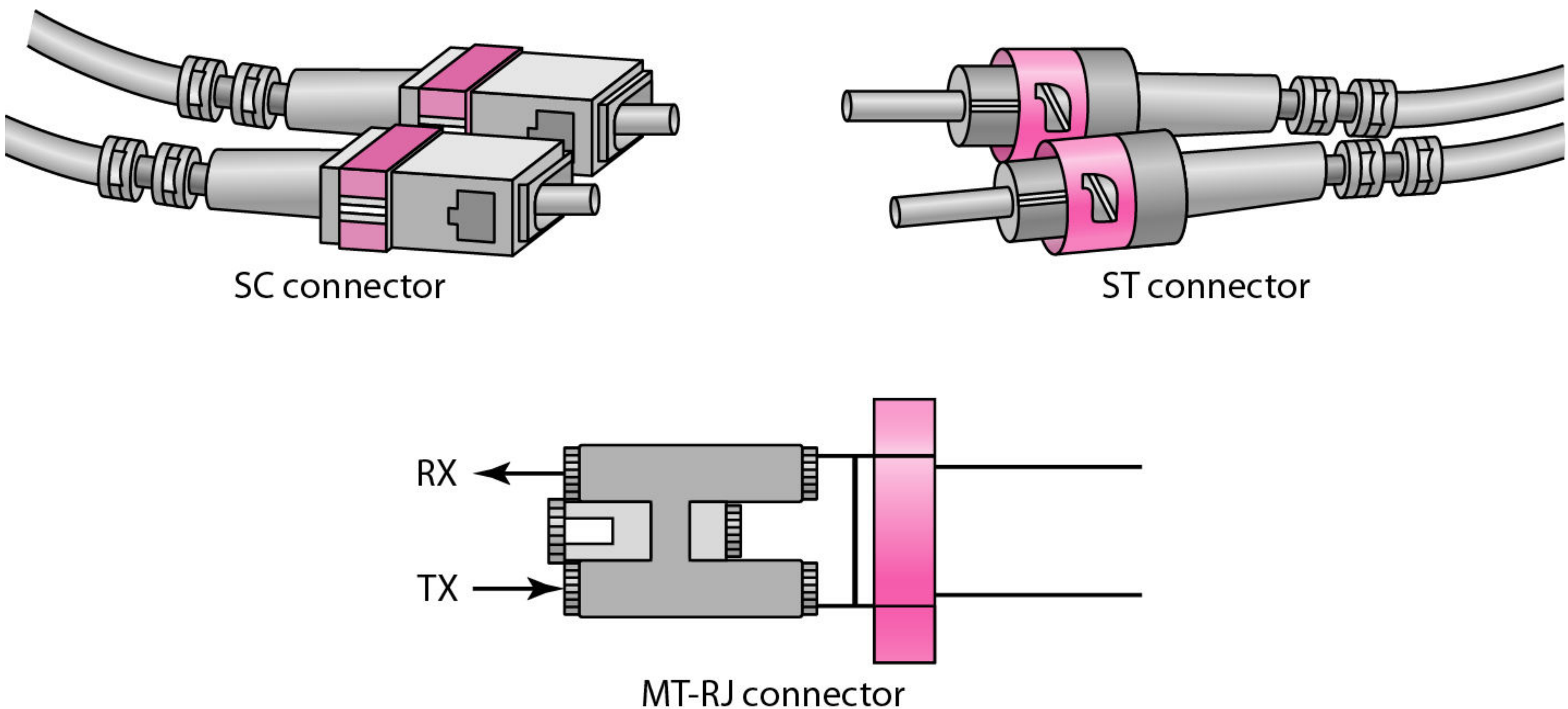
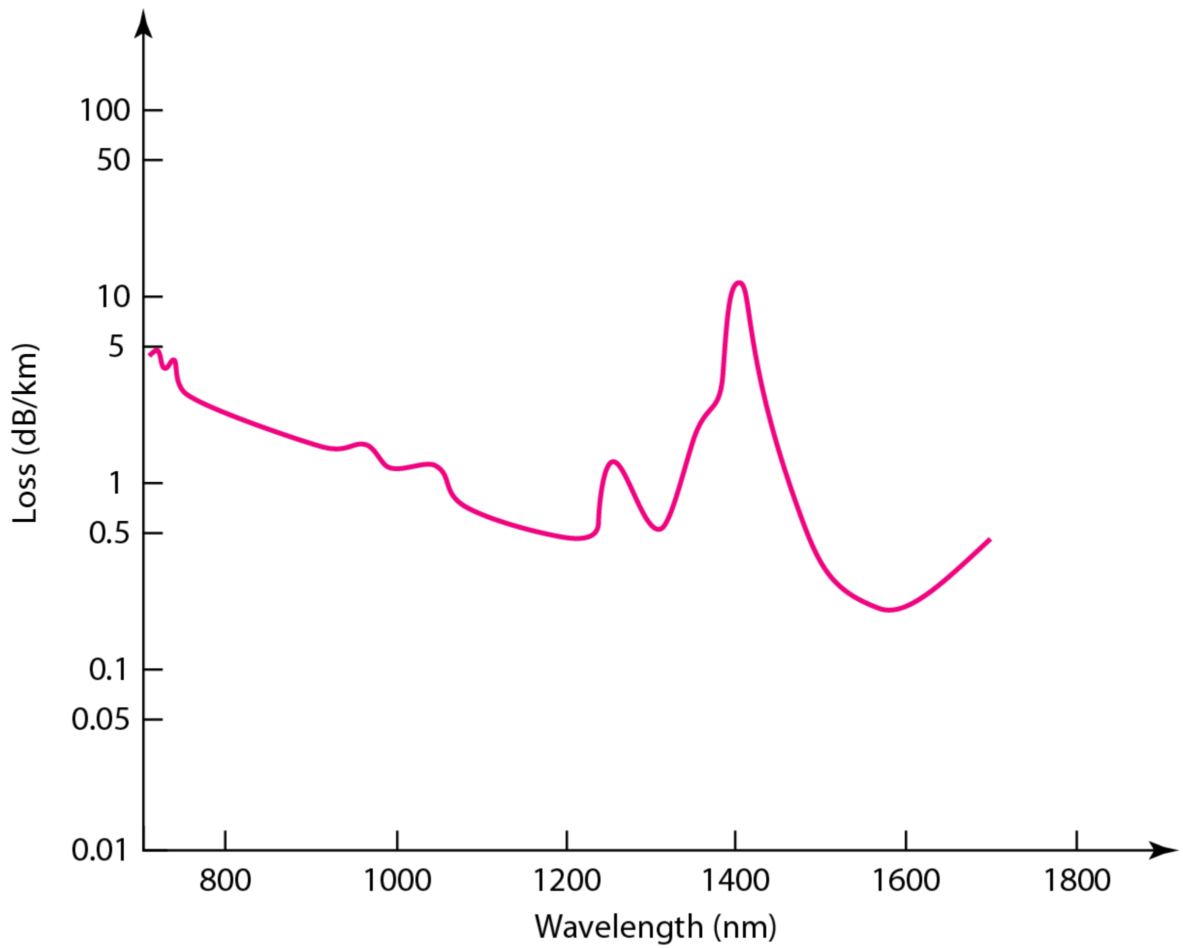


Figure 7.16 *Optical fiber performance*



7-2 UNGUIDED MEDIA: WIRELESS

Unguided media transport electromagnetic waves without using a physical conductor. This type of communication is often referred to as wireless communication.

Topics discussed in this section:

Radio Waves
Microwaves
Infrared

Figure 7.17 *Electromagnetic spectrum for wireless communication*

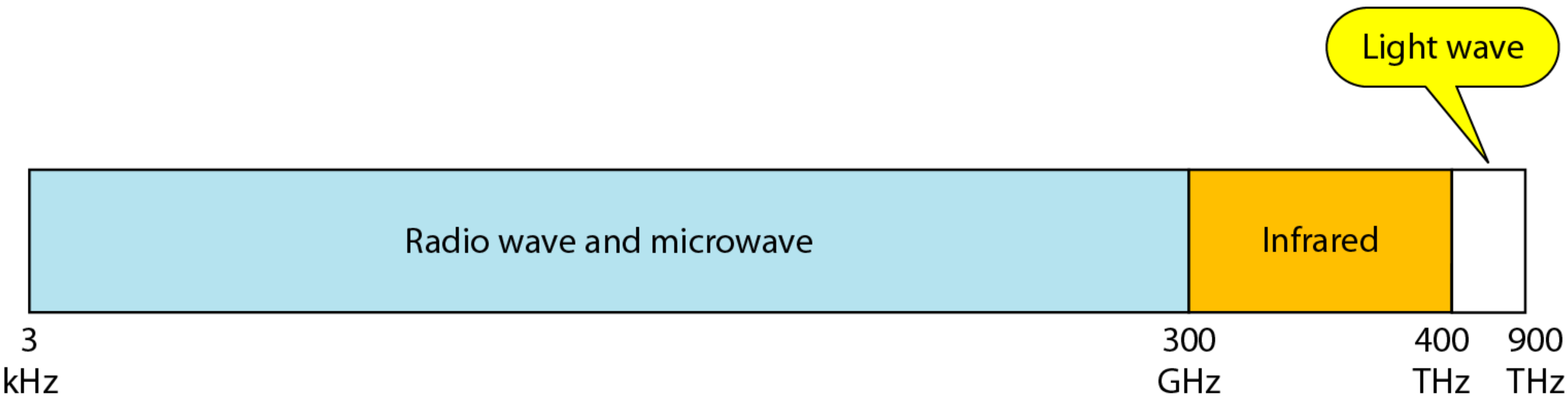


Figure 7.18 ***Propagation methods***

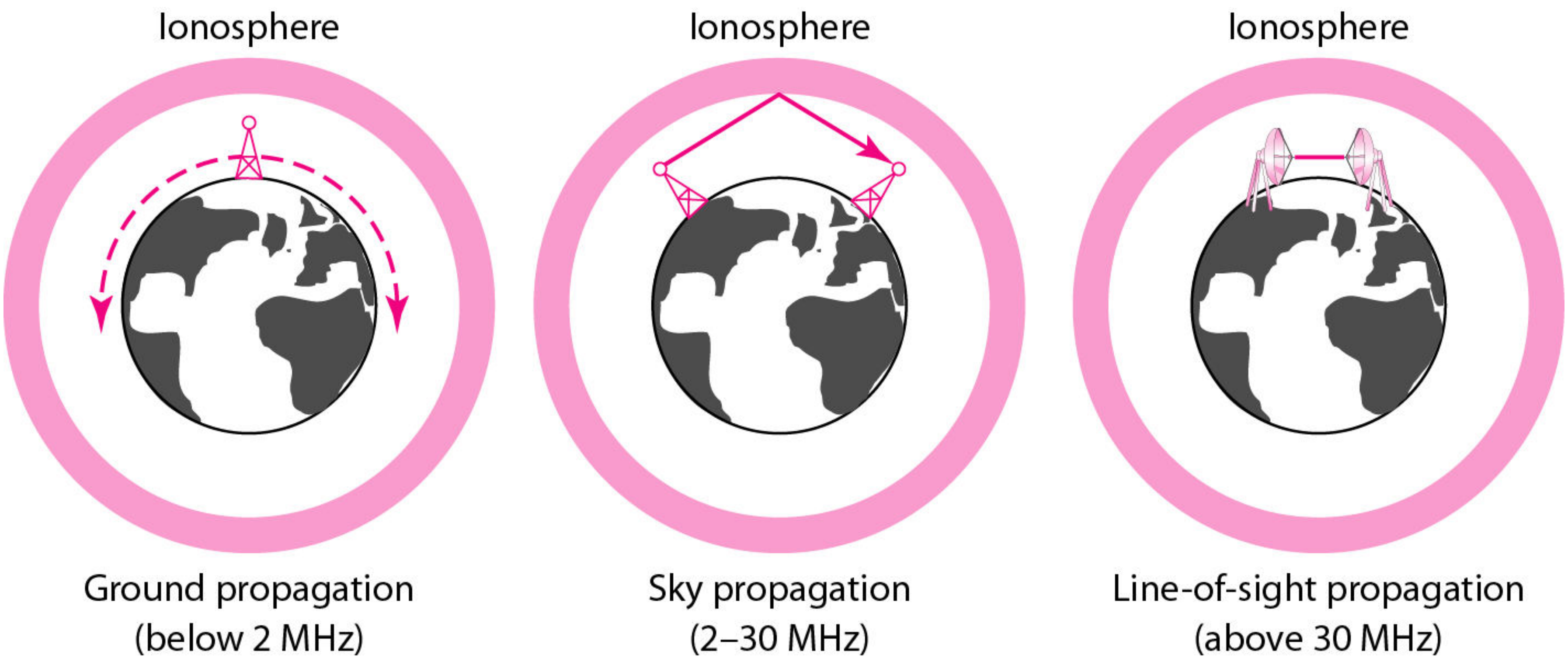
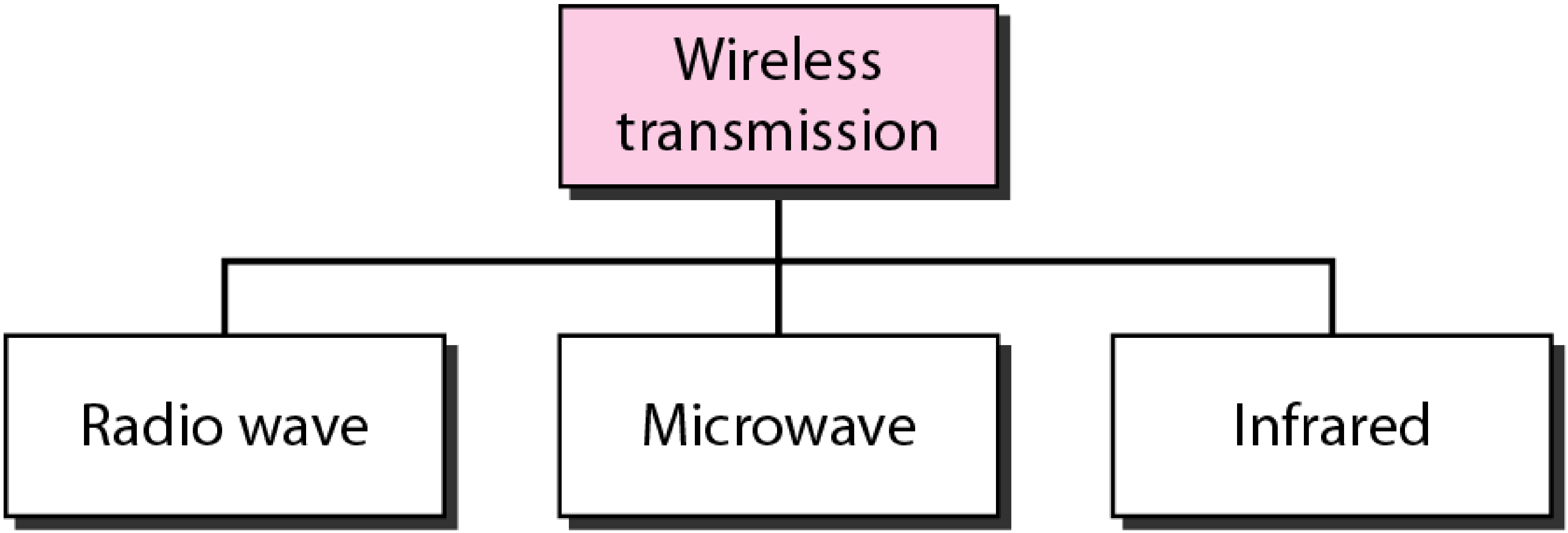


Table 7.4 *Bands*

Band	*Range*	*Propagation*	*Application*
VLF (very low frequency)	3–30 kHz	Ground	Long-range radio navigation
LF (low frequency)	30–300 kHz	Ground	Radio beacons and navigational locators
MF (middle frequency)	300 kHz–3 MHz	Sky	AM radio
HF (high frequency)	3–30 MHz	Sky	Citizens band (CB), ship/aircraft communication
VHF (very high frequency)	30–300 MHz	Sky and line-of-sight	VHF TV, FM radio
UHF (ultrahigh frequency)	300 MHz–3 GHz	Line-of-sight	UHF TV, cellular phones, paging, satellite
SHF (superhigh frequency)	3–30 GHz	Line-of-sight	Satellite communication
EHF (extremely high frequency)	30–300 GHz	Line-of-sight	Radar, satellite

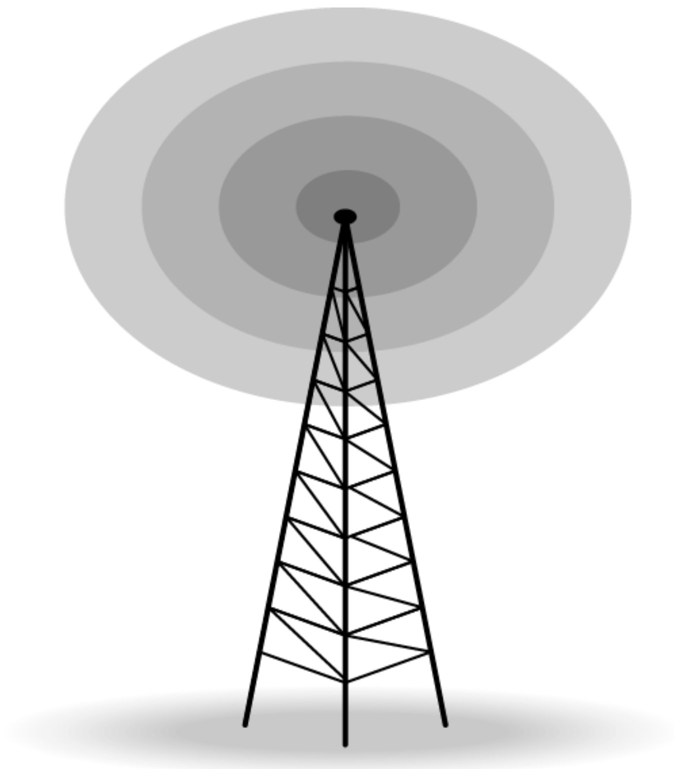
Figure 7.19 *Wireless transmission waves*



Note

**Radio waves are used for multicast communications, such as radio and television, and paging systems. They can penetrate through walls.
Highly regulated. Use omni directional antennas**

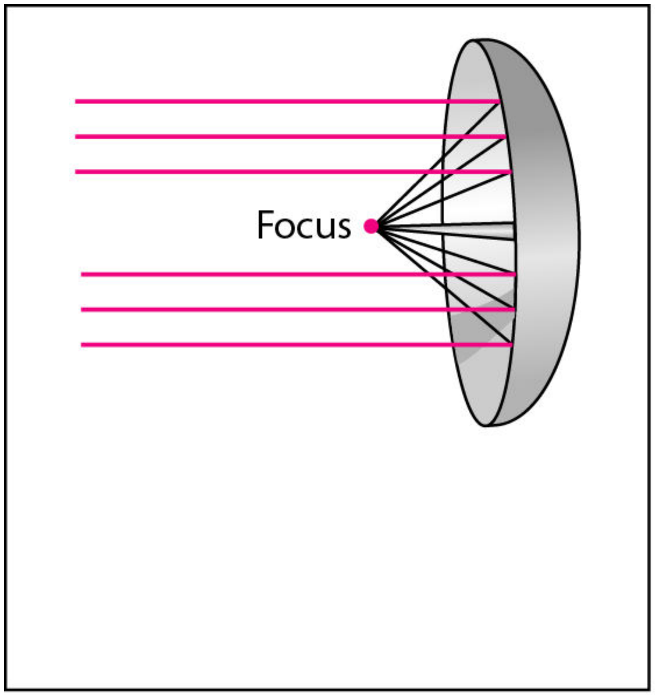
Figure 7.20 *Omnidirectional antenna*



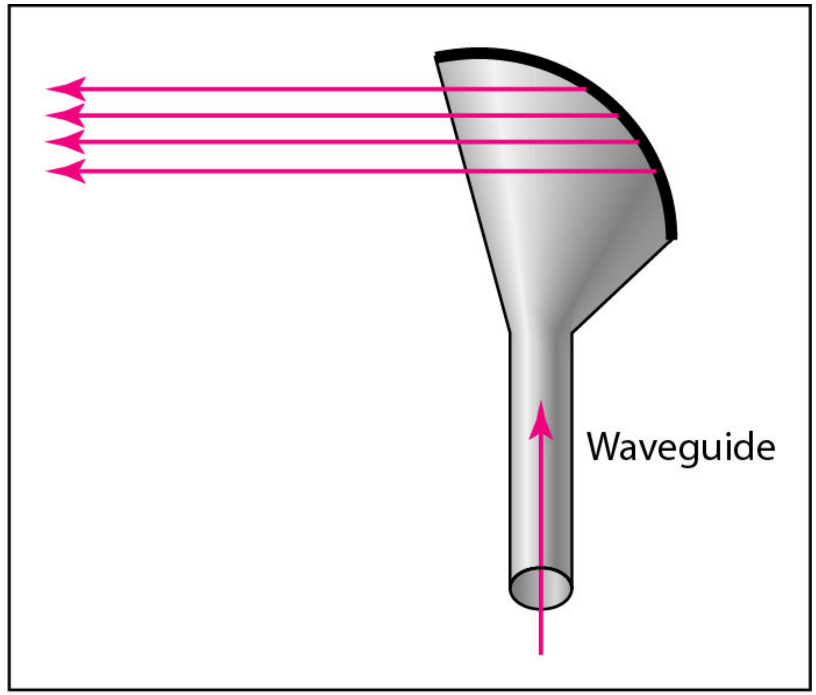
Note

Microwaves are used for unicast communication such as cellular telephones, satellite networks, and wireless LANs.
Higher frequency ranges cannot penetrate walls.
Use directional antennas - point to point line of sight communications.

Figure 7.21 *Unidirectional antennas*



a. Dish antenna



b. Horn antenna

Note

Infrared signals can be used for short-range communication in a closed area using line-of-sight propagation.

Wireless Channels

- Are subject to a lot more errors than guided media channels.
- Interference is one cause for errors, can be circumvented with high SNR.
- The higher the SNR the less capacity is available for transmission due to the broadcast nature of the channel.
- Channel also subject to fading and no coverage holes.

Chapter 10

Error Detection and Correction

Note

Data can be corrupted during transmission.

Some applications require that errors be detected and corrected.

10-1 INTRODUCTION

Let us first discuss some issues related, directly or indirectly, to error detection and correction.

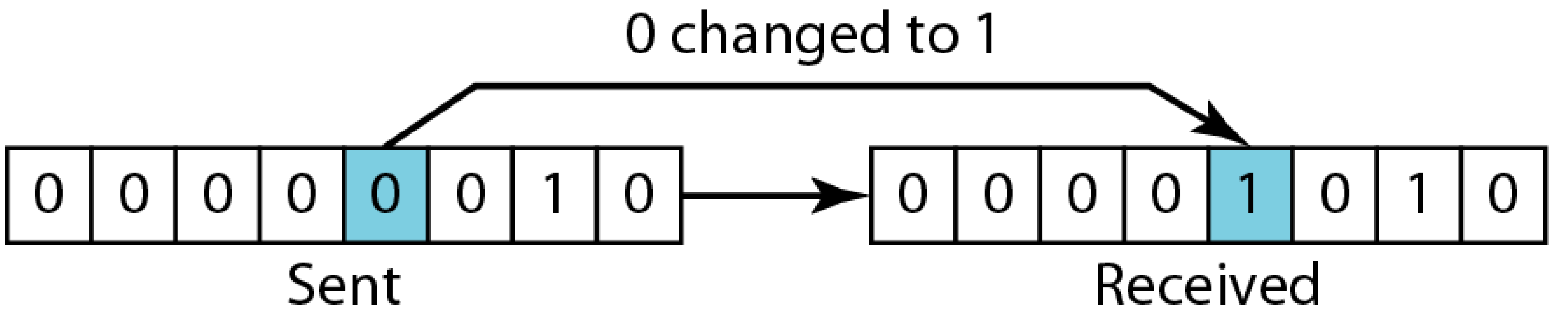
Topics discussed in this section:

Types of Errors
Redundancy
Detection Versus Correction
Forward Error Correction Versus Retransmission
Coding
Modular Arithmetic

Note

In a single-bit error, only 1 bit in the data unit has changed.

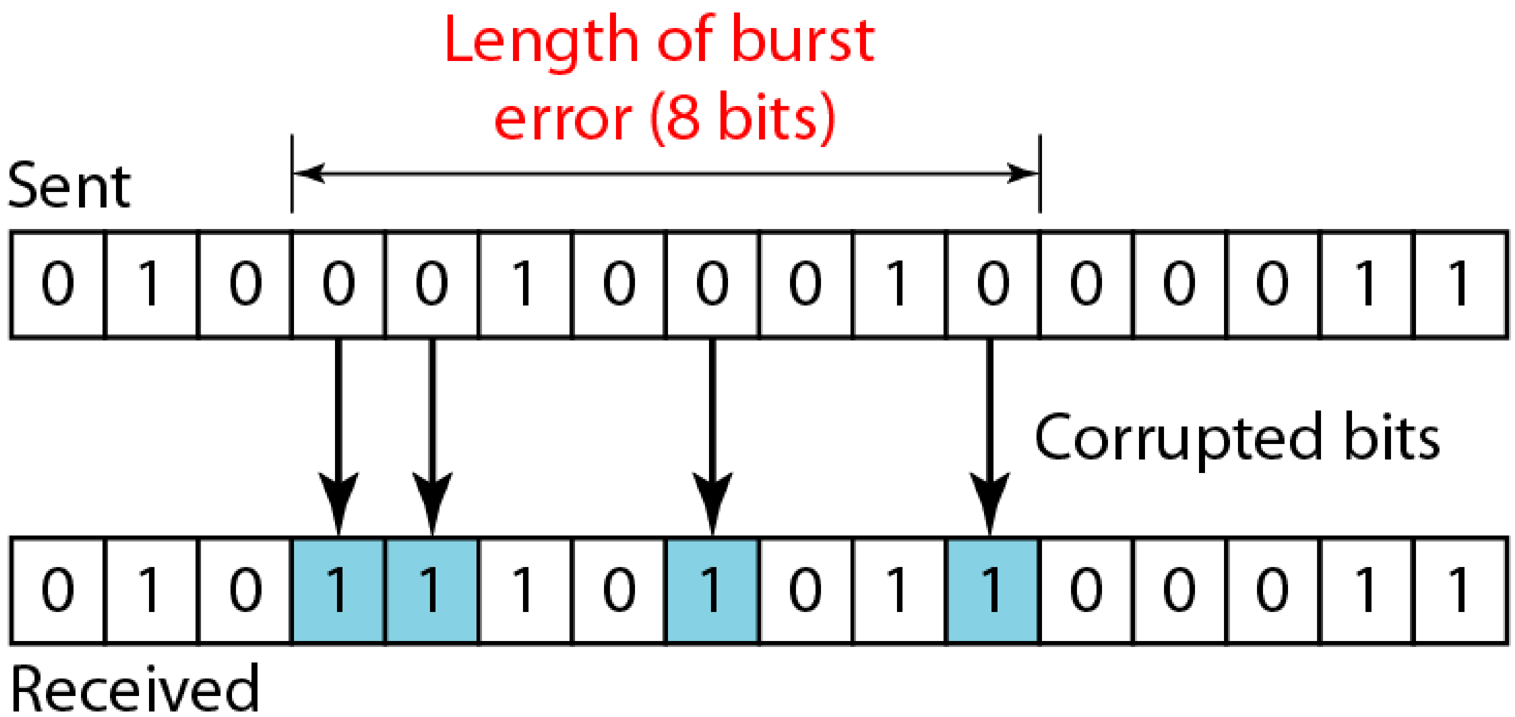
Figure 10.1 *Single-bit error*



Note

A burst error means that 2 or more bits in the data unit have changed.

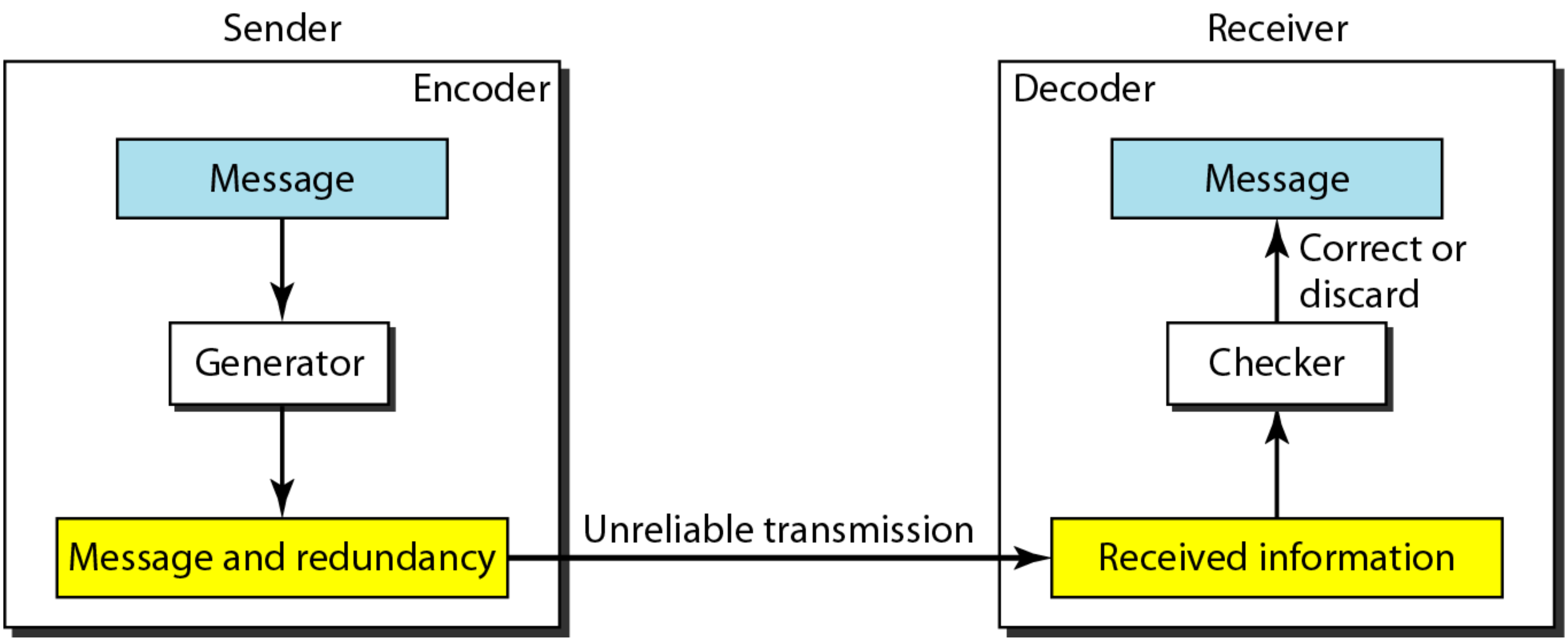
Figure 10.2 *Burst error of length 8*



Note

To detect or correct errors, we need to send extra (redundant) bits with data.

Figure 10.3 *The structure of encoder and decoder*



Note

In this book, we concentrate on block codes; we leave convolution codes to advanced texts.

Note

In modulo-N arithmetic, we use only the integers in the range 0 to N −1, inclusive.

Figure 10.4 *XORing of two single bits or two words*

$0 \oplus 0 = 0 \qquad 1 \oplus 1 = 0$

a. Two bits are the same, the result is 0.

$0 \oplus 1 = 1 \qquad 1 \oplus 0 = 1$

b. Two bits are different, the result is 1.

$$\begin{array}{cccccc} & 1 & 0 & 1 & 1 & 0 \\ \oplus & 1 & 1 & 1 & 0 & 0 \\ \hline & 0 & 1 & 0 & 1 & 0 \end{array}$$

c. Result of XORing two patterns

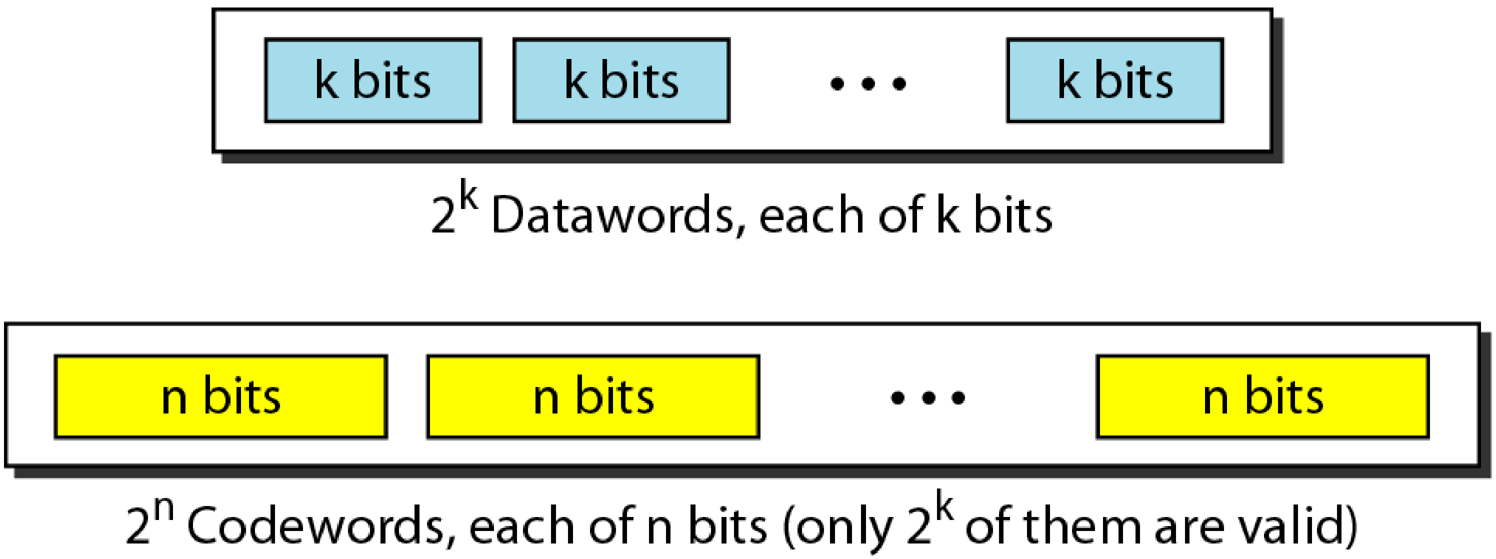
10-2 BLOCK CODING

In block coding, we divide our message into blocks, each of k bits, called datawords. We add r redundant bits to each block to make the length n = k + r. The resulting n-bit blocks are called codewords.

Topics discussed in this section:

Error Detection
Error Correction
Hamming Distance
Minimum Hamming Distance

Figure 10.5 *Datawords and codewords in block coding*



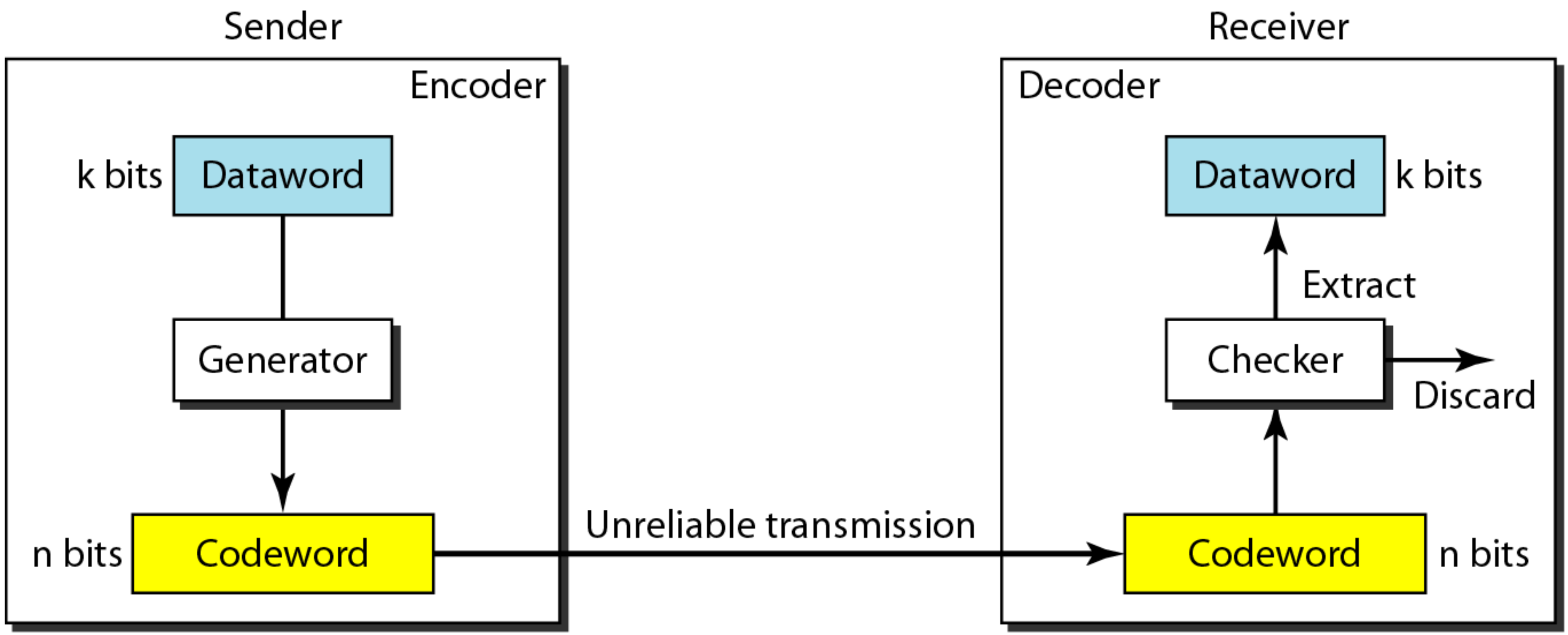
Example 10.1

The 4B/5B block coding discussed in Chapter 4 is a good example of this type of coding. In this coding scheme, $k = 4$ and $n = 5$. As we saw, we have $2^k = 16$ datawords and $2^n = 32$ codewords. We saw that 16 out of 32 codewords are used for message transfer and the rest are either used for other purposes or unused.

Error Detection

- Enough redundancy is added to detect an error.
- The receiver knows an error occurred but does not know which bit(s) is(are) in error.
- Has less overhead than error correction.

Figure 10.6 *Process of error detection in block coding*



Example 10.2

Let us assume that k = 2 and n = 3. Table 10.1 shows the list of datawords and codewords. Later, we will see how to derive a codeword from a dataword.

Assume the sender encodes the dataword 01 as 011 and sends it to the receiver. Consider the following cases:

1. *The receiver receives 011. It is a valid codeword. The receiver extracts the dataword 01 from it.*

Example 10.2 (continued)

2. ***The codeword is corrupted during transmission, and 111 is received. This is not a valid codeword and is discarded.***

3. ***The codeword is corrupted during transmission, and 000 is received. This is a valid codeword. The receiver incorrectly extracts the dataword 00. Two corrupted bits have made the error undetectable.***

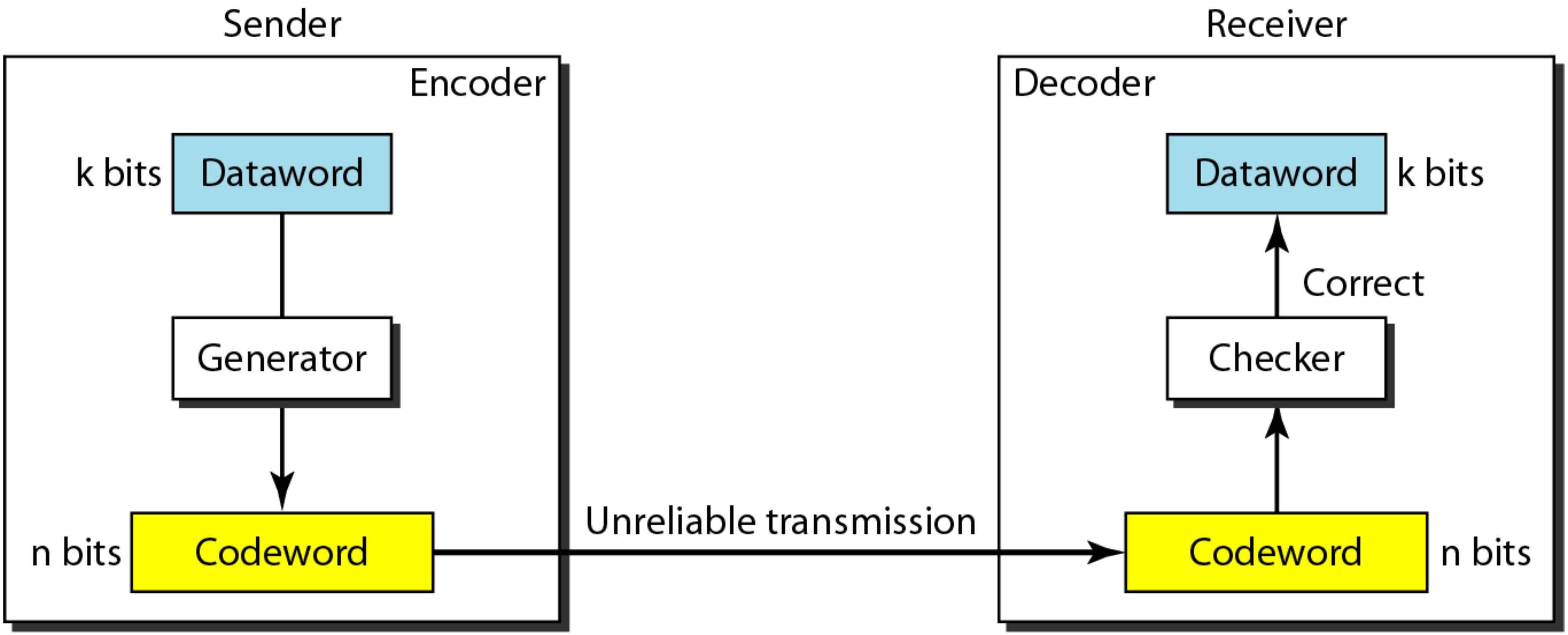
Table 10.1 ***A code for error detection (Example 10.2)***

Datawords	*Codewords*
00	000
01	011
10	101
11	110

Note

An error-detecting code can detect only the types of errors for which it is designed; other types of errors may remain undetected.

Figure 10.7 *Structure of encoder and decoder in error correction*



Example 10.3

Let us add more redundant bits to Example 10.2 to see if the receiver can correct an error without knowing what was actually sent. We add 3 redundant bits to the 2-bit dataword to make 5-bit codewords. Table 10.2 shows the datawords and codewords. Assume the dataword is 01. The sender creates the codeword 01011. The codeword is corrupted during transmission, and 01001 is received. First, the receiver finds that the received codeword is not in the table. This means an error has occurred. The receiver, assuming that there is only 1 bit corrupted, uses the following strategy to guess the correct dataword.

Example 10.3 (continued)

1. *Comparing the received codeword with the first codeword in the table (01001 versus 00000), the receiver decides that the first codeword is not the one that was sent because there are two different bits.*

2. *By the same reasoning, the original codeword cannot be the third or fourth one in the table.*

3. *The original codeword must be the second one in the table because this is the only one that differs from the received codeword by 1 bit. The receiver replaces 01001 with 01011 and consults the table to find the dataword 01.*

Table 10.2 ***A code for error correction (Example 10.3)***

Dataword	*Codeword*
00	00000
01	01011
10	10101
11	11110

Chapter 10

Error Detection and Correction

Note

The Hamming distance between two words is the number of differences between corresponding bits.

Example 10.4

Let us find the Hamming distance between two pairs of words.

1. The Hamming distance d(000, 011) is 2 because

$$\boxed{000 \oplus 011 \text{ is } 011 \text{ (two 1s)}}$$

2. The Hamming distance d(10101, 11110) is 3 because

$$\boxed{10101 \oplus 11110 \text{ is } 01011 \text{ (three 1s)}}$$

Note

The minimum Hamming distance is the smallest Hamming distance between all possible pairs in a set of words.

Example 10.5

Find the minimum Hamming distance of the coding scheme in Table 10.1.

Solution

We first find all Hamming distances.

$d(000, 011) = 2$ $d(000, 101) = 2$ $d(000, 110) = 2$ $d(011, 101) = 2$
$d(011, 110) = 2$ $d(101, 110) = 2$

The d_{min} in this case is 2.

Example 10.6

Find the minimum Hamming distance of the coding scheme in Table 10.2.

Solution

We first find all the Hamming distances.

$d(00000, 01011) = 3$ $d(00000, 10101) = 3$ $d(00000, 11110) = 4$
$d(01011, 10101) = 4$ $d(01011, 11110) = 3$ $d(10101, 11110) = 3$

The d_{min} in this case is 3.

Note

To guarantee the detection of up to s errors in all cases, the minimum Hamming distance in a block code must be $d_{min} = s + 1$.

Example 10.7

The minimum Hamming distance for our first code scheme (Table 10.1) is 2. This code guarantees detection of only a single error. For example, if the third codeword (101) is sent and one error occurs, the received codeword does not match any valid codeword. If two errors occur, however, the received codeword may match a valid codeword and the errors are not detected.

Example 10.8

Our second block code scheme (Table 10.2) has d_{min} = 3. This code can detect up to two errors. Again, we see that when any of the valid codewords is sent, two errors create a codeword which is not in the table of valid codewords. The receiver cannot be fooled.

However, some combinations of three errors change a valid codeword to another valid codeword. The receiver accepts the received codeword and the errors are undetected.

Figure 10.8 ***Geometric concept for finding d_{min} in error detection***

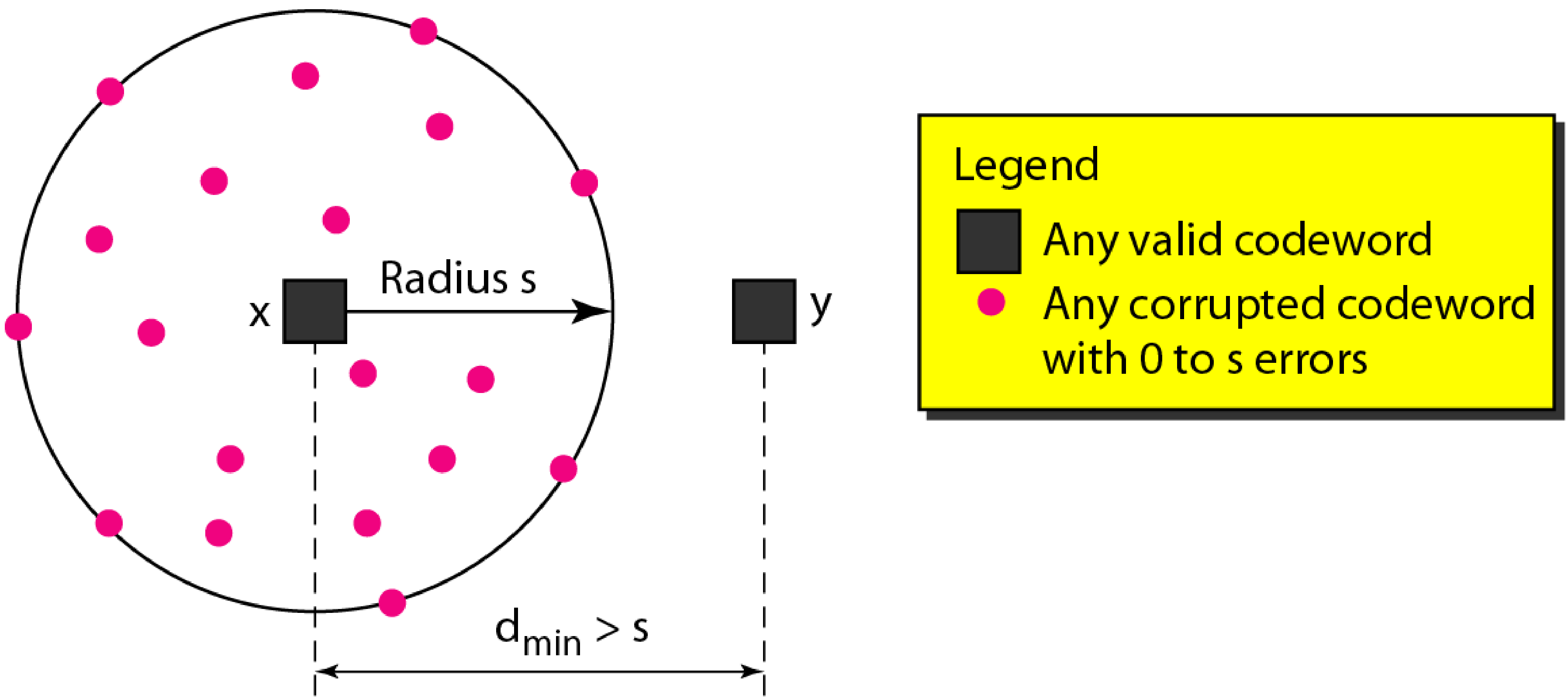
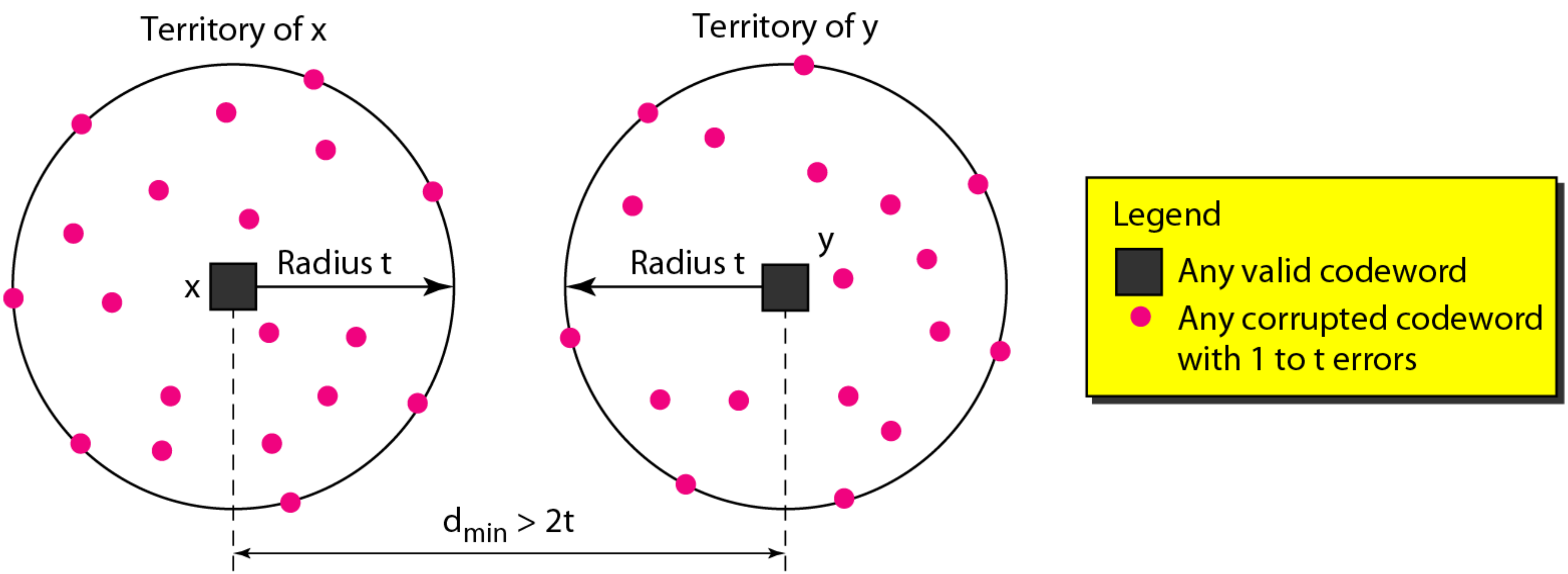


Figure 10.9 ***Geometric concept for finding d_{min} in error correction***



Note

To guarantee correction of up to t errors in all cases, the minimum Hamming distance in a block code must be $d_{min} = 2t + 1$.

Example 10.9

A code scheme has a Hamming distance $d_{min} = 4$. What is the error detection and correction capability of this scheme?

Solution

This code guarantees the detection of up to three errors (s = 3), but it can correct up to one error. In other words, if this code is used for error correction, part of its capability is wasted. Error correction codes need to have an odd minimum distance (3, 5, 7, . . .).

10-3 LINEAR BLOCK CODES

Almost all block codes used today belong to a subset called linear block codes. A linear block code is a code in which the exclusive OR (addition modulo-2) of two valid codewords creates another valid codeword.

Topics discussed in this section:

Minimum Distance for Linear Block Codes
Some Linear Block Codes

Note

In a linear block code, the exclusive OR (XOR) of any two valid codewords creates another valid codeword.

Example 10.10

Let us see if the two codes we defined in Table 10.1 and Table 10.2 belong to the class of linear block codes.

1. ***The scheme in Table 10.1 is a linear block code because the result of XORing any codeword with any other codeword is a valid codeword. For example, the XORing of the second and third codewords creates the fourth one.***

2. ***The scheme in Table 10.2 is also a linear block code. We can create all four codewords by XORing two other codewords.***

Example 10.11

In our first code (Table 10.1), the numbers of 1s in the nonzero codewords are 2, 2, and 2. So the minimum Hamming distance is d_{min} = 2. In our second code (Table 10.2), the numbers of 1s in the nonzero codewords are 3, 3, and 4. So in this code we have d_{min} = 3.

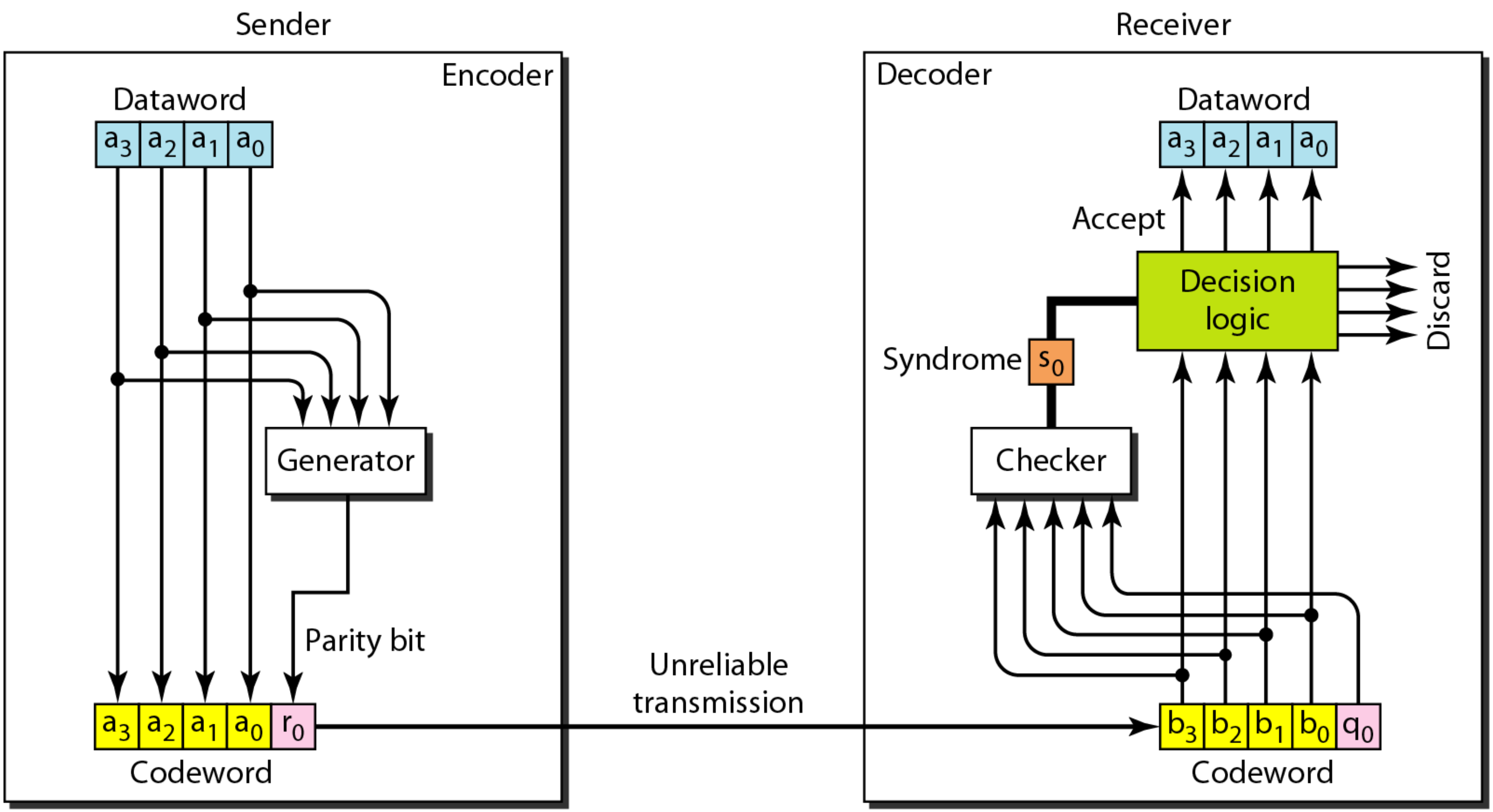
Note

A simple parity-check code is a single-bit error-detecting code in which $n = k + 1$ with $d_{min} = 2$. Even parity (ensures that a codeword has an even number of 1's) and odd parity (ensures that there are an odd number of 1's in the codeword)

Table 10.3 ***Simple parity-check code C(5, 4)***

Datawords	*Codewords*	*Datawords*	*Codewords*
0000	00000	1000	10001
0001	00011	1001	10010
0010	00101	1010	10100
0011	00110	1011	10111
0100	01001	1100	11000
0101	01010	1101	11011
0110	01100	1110	11101
0111	01111	1111	11110

Figure 10.10 *Encoder and decoder for simple parity-check code*



Example 10.12

Let us look at some transmission scenarios. Assume the sender sends the dataword 1011. The codeword created from this dataword is 10111, which is sent to the receiver. We examine five cases:

1. ***No error occurs; the received codeword is 10111. The syndrome is 0. The dataword 1011 is created.***
2. ***One single-bit error changes a_1. The received codeword is 10011. The syndrome is 1. No dataword is created.***
3. ***One single-bit error changes r_0. The received codeword is 10110. The syndrome is 1. No dataword is created.***

Example 10.12 (continued)

4. An error changes r_0 and a second error changes a_3. The received codeword is 00110. The syndrome is 0. The dataword 0011 is created at the receiver. Note that here the dataword is wrongly created due to the syndrome value.

5. Three bits—a_3, a_2, and a_1—are changed by errors. The received codeword is 01011. The syndrome is 1. The dataword is not created. This shows that the simple parity check, guaranteed to detect one single error, can also find any odd number of errors.

Note

A simple parity-check code can detect an odd number of errors.

Note

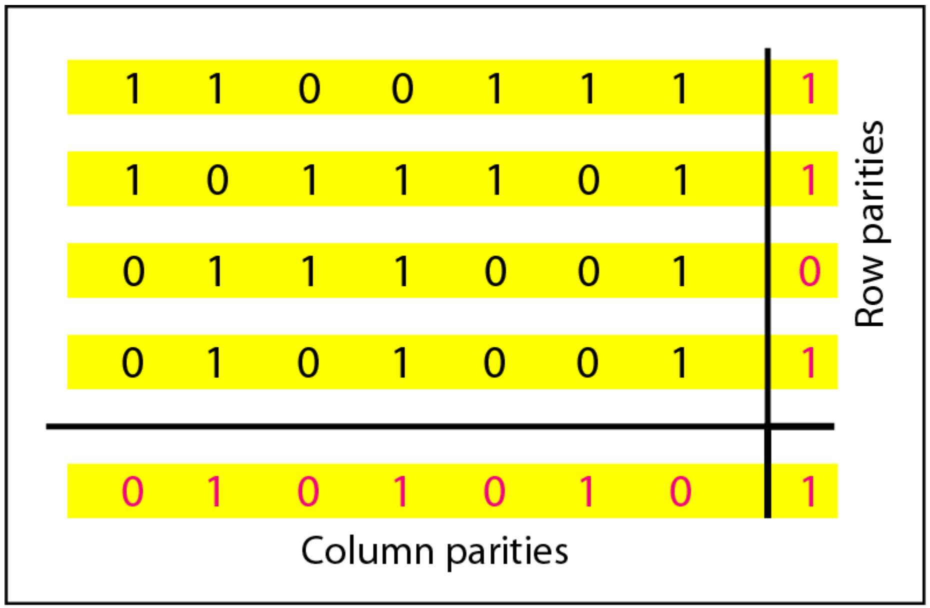
All Hamming codes discussed in this book have d_{min} = 3 (2 bit error detection and single bit error correction).

A codeword consists of n bits of which k are data bits and r are check bits.

Let m = r, then we have: $n = 2^m - 1$

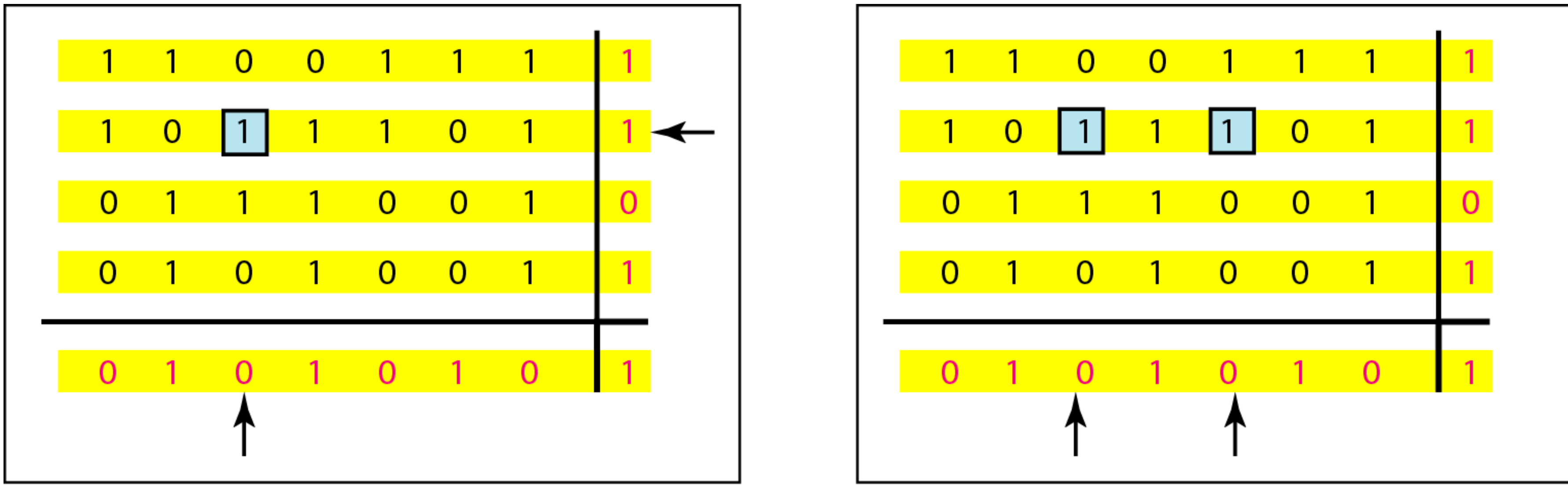
and k = n-m

Figure 10.11 *Two-dimensional parity-check code*



a. Design of row and column parities

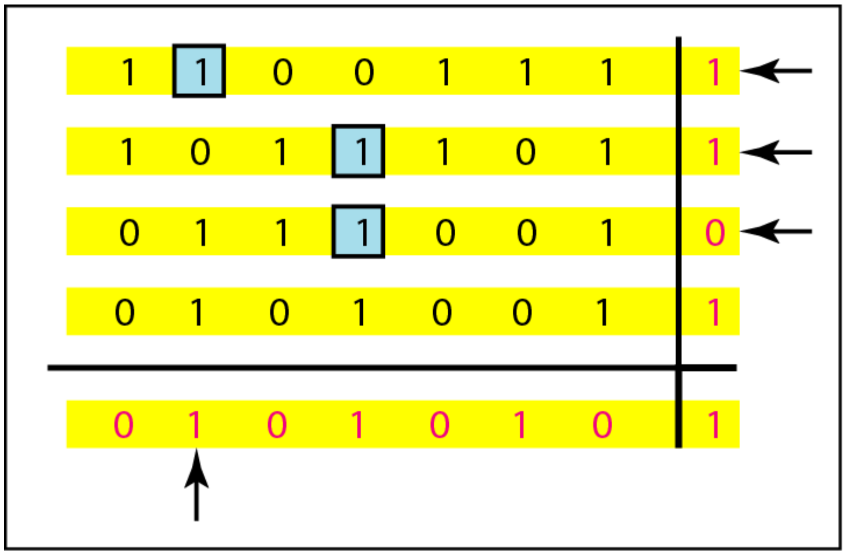
Figure 10.11 *Two-dimensional parity-check code*



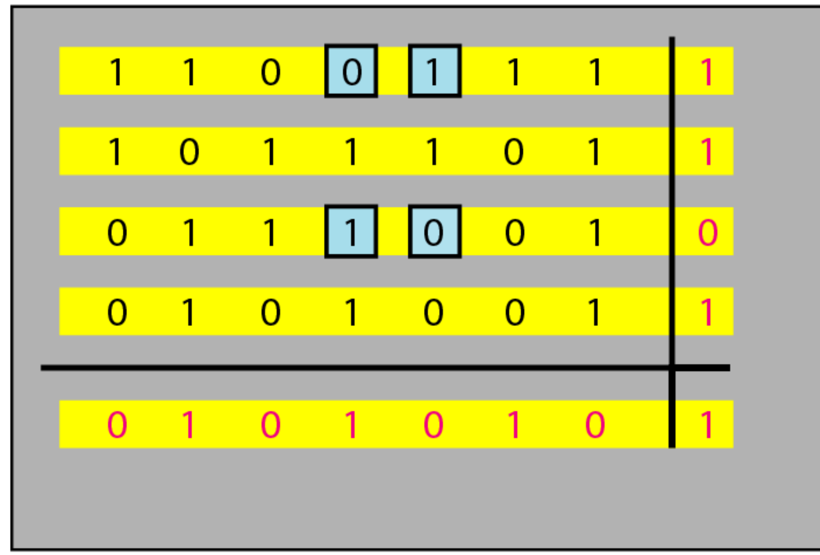
b. One error affects two parities

c. Two errors affect two parities

Figure 10.11 *Two-dimensional parity-check code*



d. Three errors affect four parities



e. Four errors cannot be detected

Table 10.4 ***Hamming code C(7, 4) - n=7, k = 4***

Datawords	*Codewords*	*Datawords*	*Codewords*
0000	0000000	1000	1000110
0001	0001101	1001	1001011
0010	0010111	1010	1010001
0011	0011010	1011	1011100
0100	0100011	1100	1100101
0101	0101110	1101	1101000
0110	0110100	1110	1110010
0111	0111001	1111	1111111

Calculating the parity bits at the transmitter
:

Modulo 2 arithmetic:

$r_0 = a_2 + a_1 + a_0$
$r_1 = a_3 + a_2 + a_1$
$r_2 = a_1 + a_0 + a_3$

Calculating the syndrome at the receiver:

$s_0 = b_2 + b_1 + b_0$
$s_1 = b_3 + b_2 + b_1$
$s_2 = b_1 + b_0 + b_3$

Figure 10.12 *The structure of the encoder and decoder for a Hamming code*

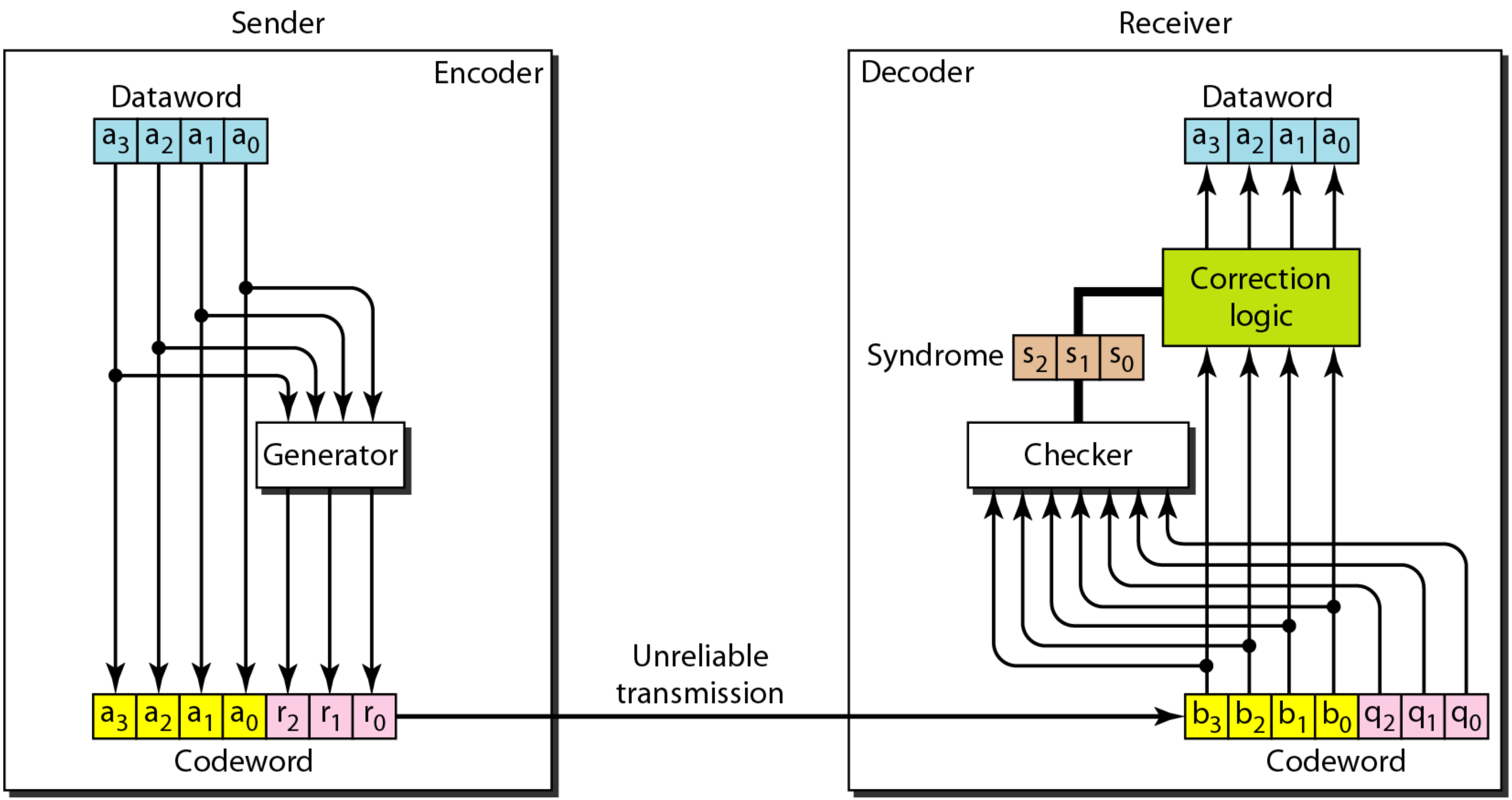


Table 10.5 ***Logical decision made by the correction logic analyzer***

Syndrome	000	001	010	011	100	101	110	111
Error	None	q_0	q_1	b_2	q_2	b_0	b_3	b_1

Example 10.13

Let us trace the path of three datawords from the sender to the destination:

1. ***The dataword 0100 becomes the codeword 0100011. The codeword 0100011 is received. The syndrome is 000, the final dataword is 0100.***
2. ***The dataword 0111 becomes the codeword 0111001. The received codeword is: 0011001. The syndrome is 011. After flipping b_2 (changing the 1 to 0), the final dataword is 0111.***
3. ***The dataword 1101 becomes the codeword 1101000. The syndrome is 101. After flipping b_0, we get 0000, the wrong dataword. This shows that our code cannot correct two errors.***

Example 10.14

We need a dataword of at least 7 bits. Calculate values of k and n that satisfy this requirement.

Solution

We need to make $k = n - m$ greater than or equal to 7, or $2m - 1 - m \geq 7$.

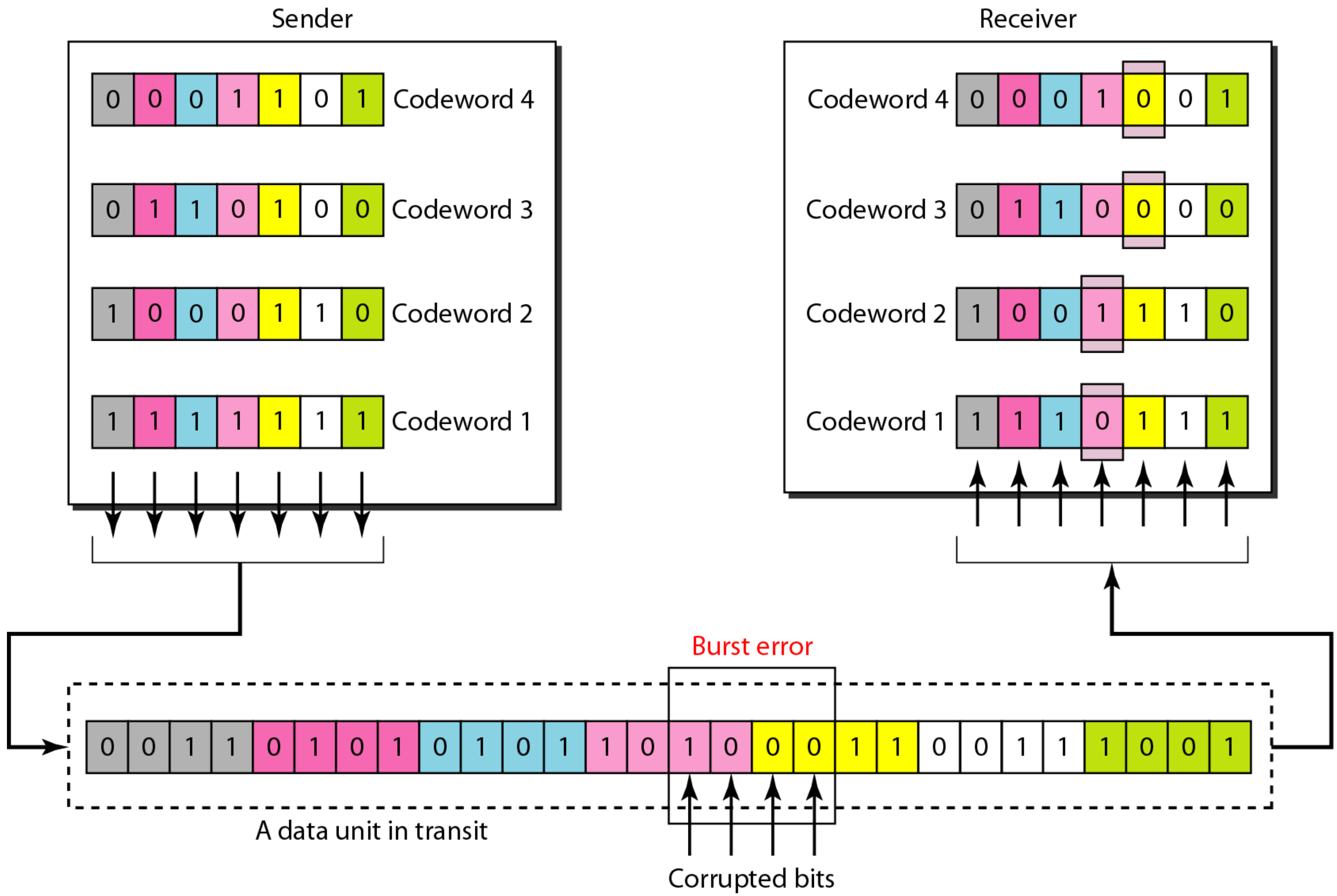
1. *If we set $m = 3$, the result is $n = 2^3 - 1 = 7$ and $k = 7 - 3$, or 4, which is < 7.*
2. *If we set $m = 4$, then $n = 2^4 - 1 = 15$ and $k = 15 - 4 = 11$, which satisfies the condition $k > 7$. So the code is*

$$\boxed{C(15, 11)}$$

Burst Errors

- Burst errors are very common, in particular in wireless environments where a fade will affect a group of bits in transit. The length of the burst is dependent on the duration of the fade.
- One way to counter burst errors, is to break up a transmission into shorter words and create a block (one word per row), then have a parity check per word.
- The words are then sent column by column. When a burst error occurs, it will affect 1 bit in several words as the transmission is read back into the block format and each word is checked individually.

Figure 10.13 *Burst error correction using Hamming code*



10-4 CYCLIC CODES

Cyclic codes ***are special linear block codes with one extra property. In a cyclic code, if a codeword is cyclically shifted (rotated), the result is another codeword.***

Topics discussed in this section:

Cyclic Redundancy Check
Hardware Implementation
Polynomials
Cyclic Code Analysis
Advantages of Cyclic Codes
Other Cyclic Codes

Table 10.6 *A CRC code with C(7, 4)*

Dataword	*Codeword*	*Dataword*	*Codeword*
0000	0000000	1000	1000101
0001	0001011	1001	1001110
0010	0010110	1010	1010011
0011	0011101	1011	1011000
0100	0100111	1100	1100010
0101	0101100	1101	1101001
0110	0110001	1110	1110100
0111	0111010	1111	1111111

Figure 10.14 *CRC encoder and decoder*

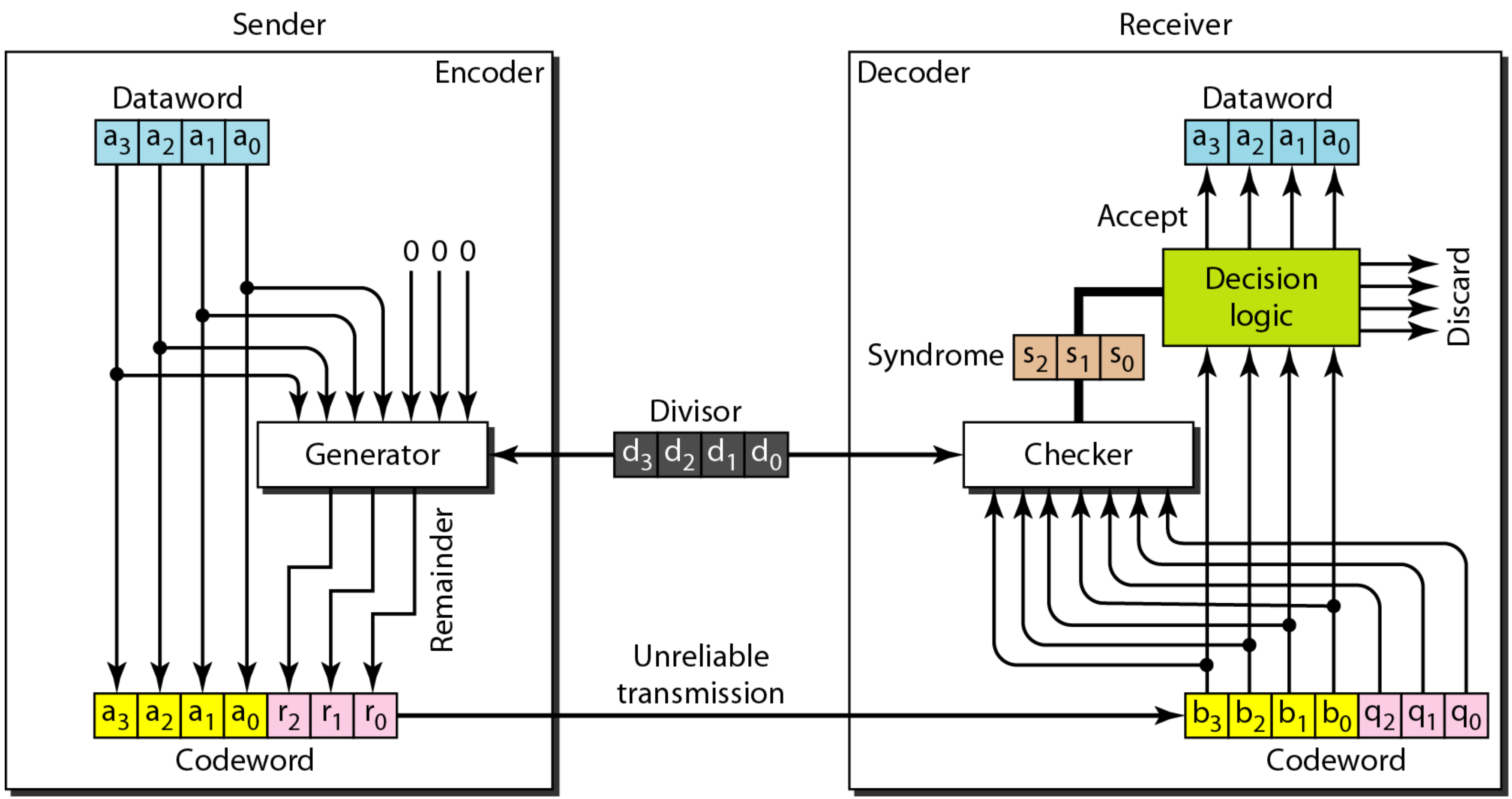


Figure 10.15 *Division in CRC encoder*

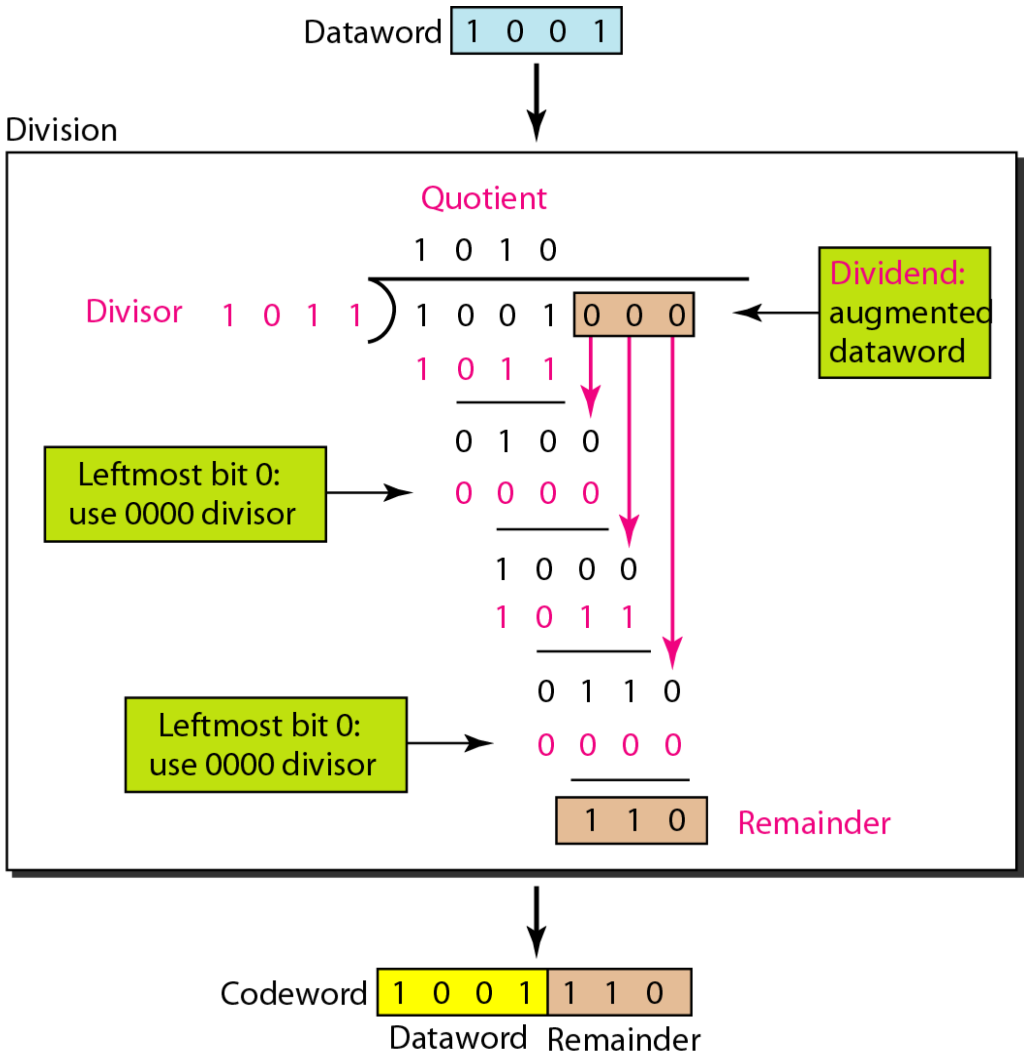


Figure 10.16 *Division in the CRC decoder for two cases*

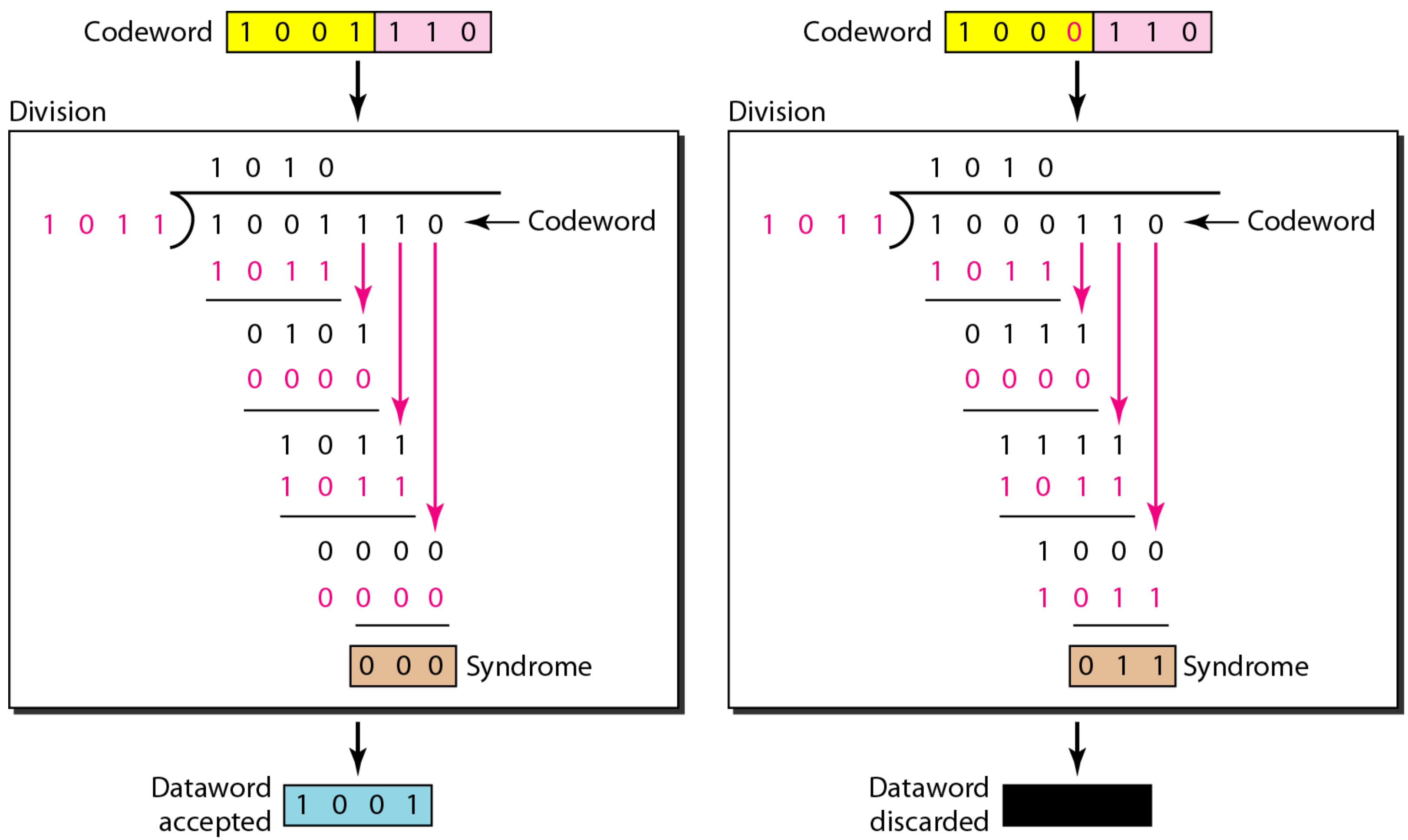


Figure 10.17 *Hardwired design of the divisor in CRC*

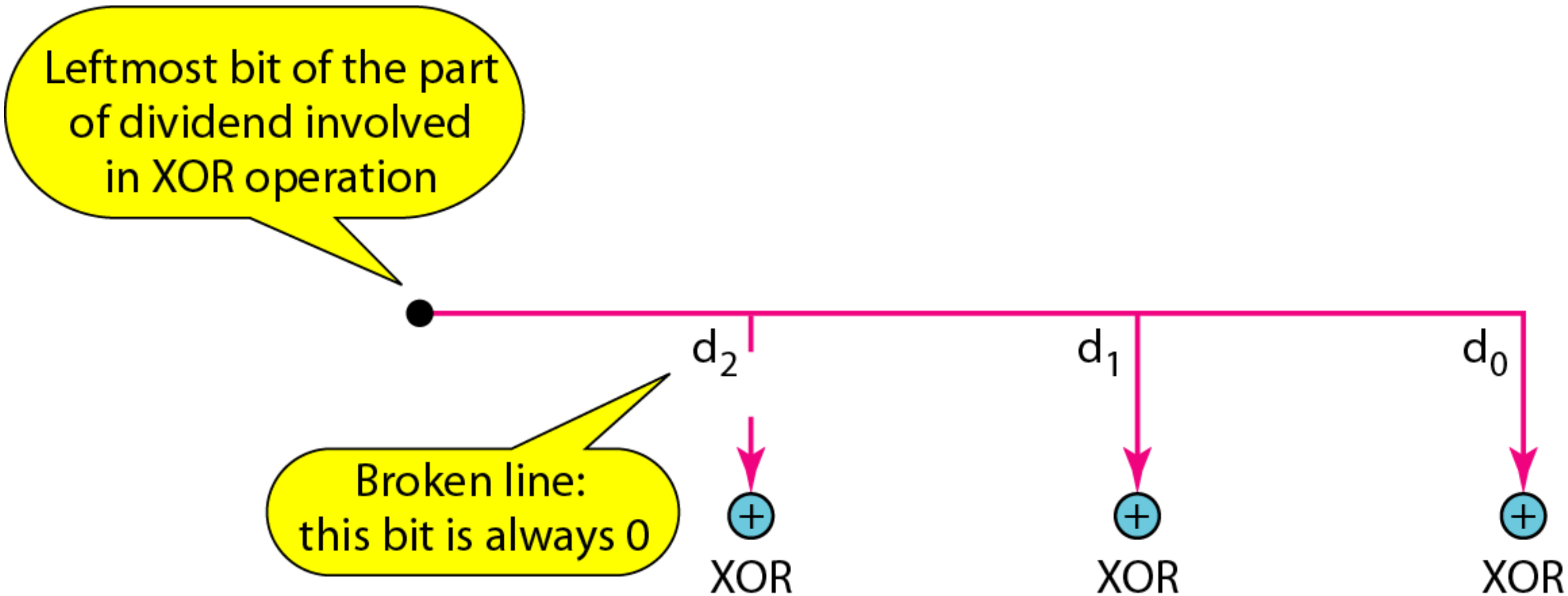


Figure 10.18 *Simulation of division in CRC encoder*

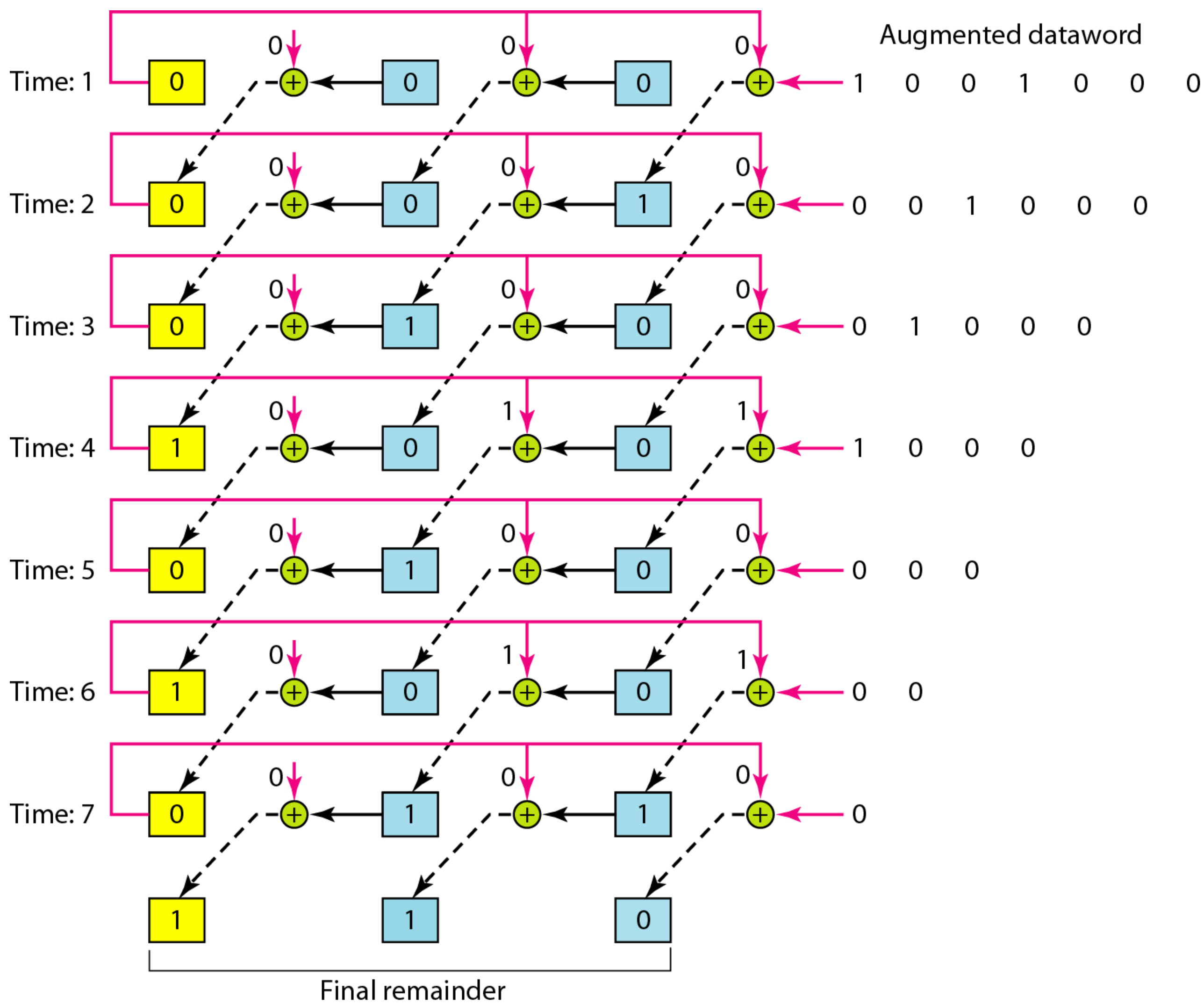


Figure 10.19 ***The CRC encoder design using shift registers***

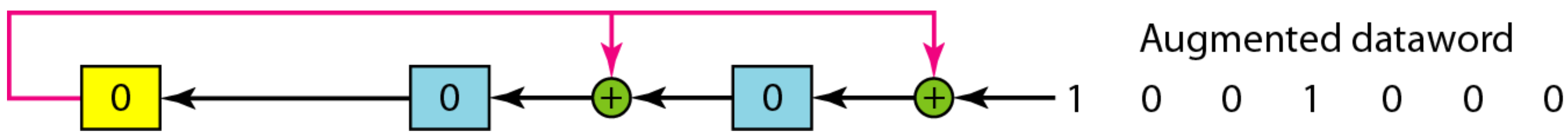
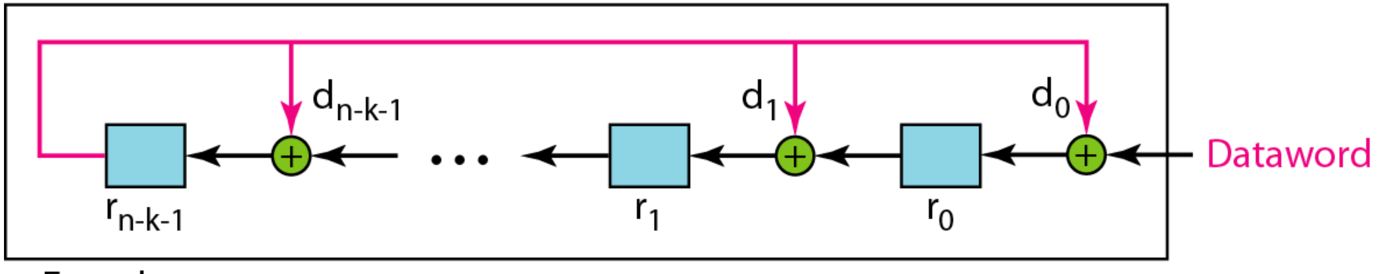


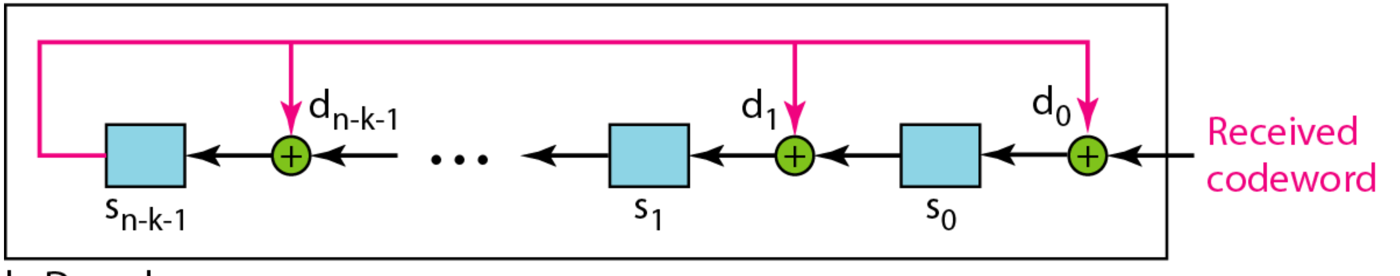
Figure 10.20 *General design of encoder and decoder of a CRC code*

Note:

The divisor line and XOR are missing if the corresponding bit in the divisor is 0.



a. Encoder



b. Decoder

Using Polynomials

- We can use a polynomial to represent a binary word.
- Each bit from right to left is mapped onto a power term.
- The rightmost bit represents the "0" power term. The bit next to it the "1" power term, etc.
- If the bit is of value zero, the power term is deleted from the expression.

Figure 10.21 *A polynomial to represent a binary word*

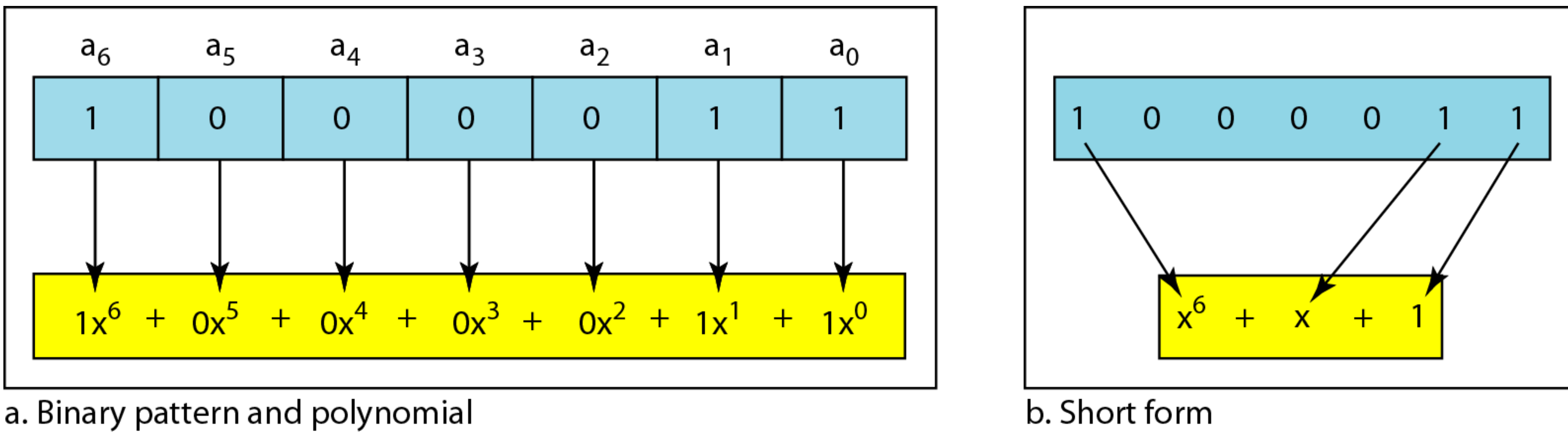
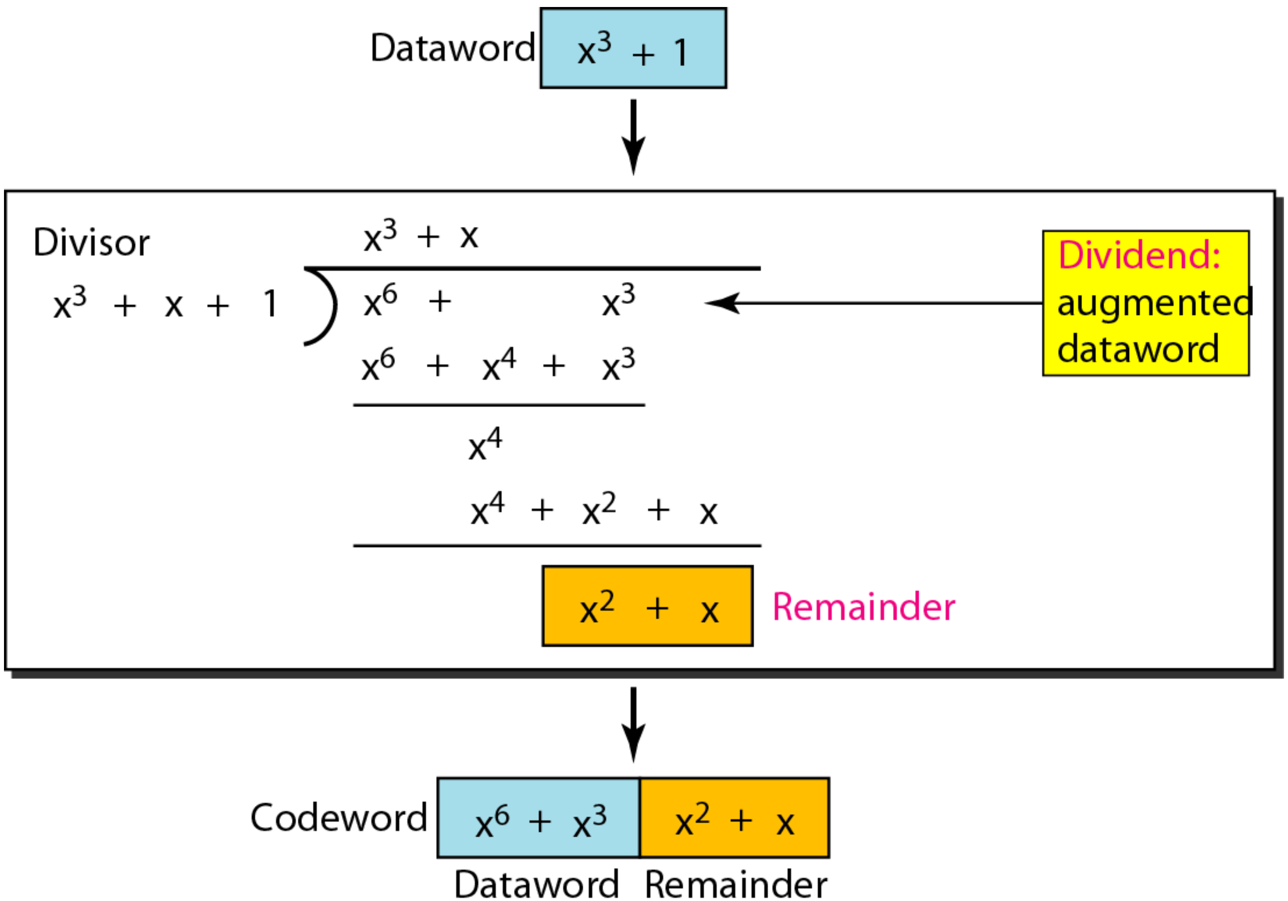


Figure 10.22 *CRC division using polynomials*



Note

The divisor in a cyclic code is normally called the generator polynomial or simply the generator.

Note

In a cyclic code,
If $s(x) \neq 0$, one or more bits is corrupted.
If $s(x) = 0$, either

a. No bit is corrupted. or
b. Some bits are corrupted, but the decoder failed to detect them.

Note

In a cyclic code, those *e*(x) errors that are divisible by *g*(x) are not caught.

Received codeword (c(x) + e(x))/g(x) = c(x)/g(x) + e(x)/gx

The first part is by definition divisible the second part will determine the error. If "0" conclusion -> no error occurred.

Note: that could mean that an error went undetected.

Note

If the generator has more than one term and the coefficient of x^0 is 1, all single errors can be caught.

Example 10.15

Which of the following g(x) values guarantees that a single-bit error is caught? For each case, what is the error that cannot be caught?

a. $x + 1$ **b.** x^3 **c.** 1

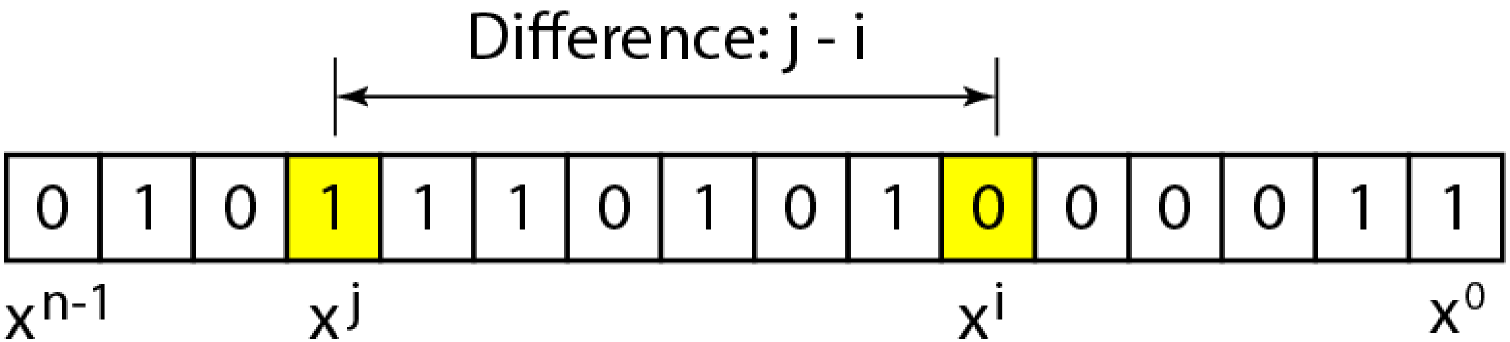
Solution

a. *No x^i can be divisible by $x + 1$. Any single-bit error can be caught.*

b. *If i is equal to or greater than 3, x^i is divisible by g(x). All single-bit errors in positions 1 to 3 are caught.*

c. *All values of i make x^i divisible by g(x). No single-bit error can be caught. This g(x) is useless.*

Figure 10.23 ***Representation of two isolated single-bit errors using polynomials***



Note

If a generator cannot divide $x^t + 1$ (t between 0 and n – 1), then all isolated double errors can be detected.

Example 10.16

Find the status of the following generators related to two isolated, single-bit errors.

a. $x + 1$ ***b.*** $x^4 + 1$ ***c.*** $x^7 + x^6 + 1$ ***d.*** $x^{15} + x^{14} + 1$

Solution

a. ***This is a very poor choice for a generator. Any two errors next to each other cannot be detected.***

b. ***This generator cannot detect two errors that are four positions apart.***

c. ***This is a good choice for this purpose.***

d. ***This polynomial cannot divide*** $x^t + 1$ ***if*** t ***is less than 32,768. A codeword with two isolated errors up to 32,768 bits apart can be detected by this generator.***

Note

A generator that contains a factor of $x + 1$ can detect all odd-numbered errors.

Note

- **All burst errors with $L \le r$ will be detected.**
- **All burst errors with $L = r + 1$ will be detected with probability $1 - (1/2)^{r-1}$.**
- **All burst errors with $L > r + 1$ will be detected with probability $1 - (1/2)^{r}$.**

Example 10.17

Find the suitability of the following generators in relation to burst errors of different lengths.

a. $x^6 + 1$ ***b.*** $x^{18} + x^7 + x + 1$ ***c.*** $x^{32} + x^{23} + x^7 + 1$

Solution

a. ***This generator can detect all burst errors with a length less than or equal to 6 bits; 3 out of 100 burst errors with length 7 will slip by; 16 out of 1000 burst errors of length 8 or more will slip by.***

Example 10.17 (continued)

b. This generator can detect all burst errors with a length less than or equal to 18 bits; 8 out of 1 million burst errors with length 19 will slip by; 4 out of 1 million burst errors of length 20 or more will slip by.

c. This generator can detect all burst errors with a length less than or equal to 32 bits; 5 out of 10 billion burst errors with length 33 will slip by; 3 out of 10 billion burst errors of length 34 or more will slip by.

Note

A good polynomial generator needs to have the following characteristics:

1. It should have at least two terms.
2. The coefficient of the term x^0 should be 1.
3. It should not divide $x^t + 1$, for t between 2 and $n - 1$.
4. It should have the factor $x + 1$.

Table 10.7 ***Standard polynomials***

Name	*Polynomial*	*Application*
CRC-8	$x^8 + x^2 + x + 1$	ATM header
CRC-10	$x^{10} + x^9 + x^5 + x^4 + x^2 + 1$	ATM AAL
CRC-16	$x^{16} + x^{12} + x^5 + 1$	HDLC
CRC-32	$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$	LANs

10-5 CHECKSUM

The last error detection method we discuss here is called the checksum. The checksum is used in the Internet by several protocols although not at the data link layer. However, we briefly discuss it here to complete our discussion on error checking

Topics discussed in this section:

Idea
One's Complement
Internet Checksum

Example 10.18

Suppose our data is a list of five 4-bit numbers that we want to send to a destination. In addition to sending these numbers, we send the sum of the numbers. For example, if the set of numbers is (7, 11, 12, 0, 6), we send (7, 11, 12, 0, 6, 36), where 36 is the sum of the original numbers. The receiver adds the five numbers and compares the result with the sum. If the two are the same, the receiver assumes no error, accepts the five numbers, and discards the sum. Otherwise, there is an error somewhere and the data are not accepted.

Example 10.19

We can make the job of the receiver easier if we send the negative (complement) of the sum, called the checksum. In this case, we send (7, 11, 12, 0, 6, −36). The receiver can add all the numbers received (including the checksum). If the result is 0, it assumes no error; otherwise, there is an error.

Example 10.20

How can we represent the number 21 in one's complement arithmetic using only four bits?

Solution

The number 21 in binary is 10101 (it needs five bits). We can wrap the leftmost bit and add it to the four rightmost bits. We have (0101 + 1) = 0110 or 6.

Example 10.21

How can we represent the number −6 in one's complement arithmetic using only four bits?

Solution

In one's complement arithmetic, the negative or complement of a number is found by inverting all bits. Positive 6 is 0110; negative 6 is 1001. If we consider only unsigned numbers, this is 9. In other words, the complement of 6 is 9. Another way to find the complement of a number in one's complement arithmetic is to subtract the number from $2^n - 1$ ($16 - 1$ in this case).

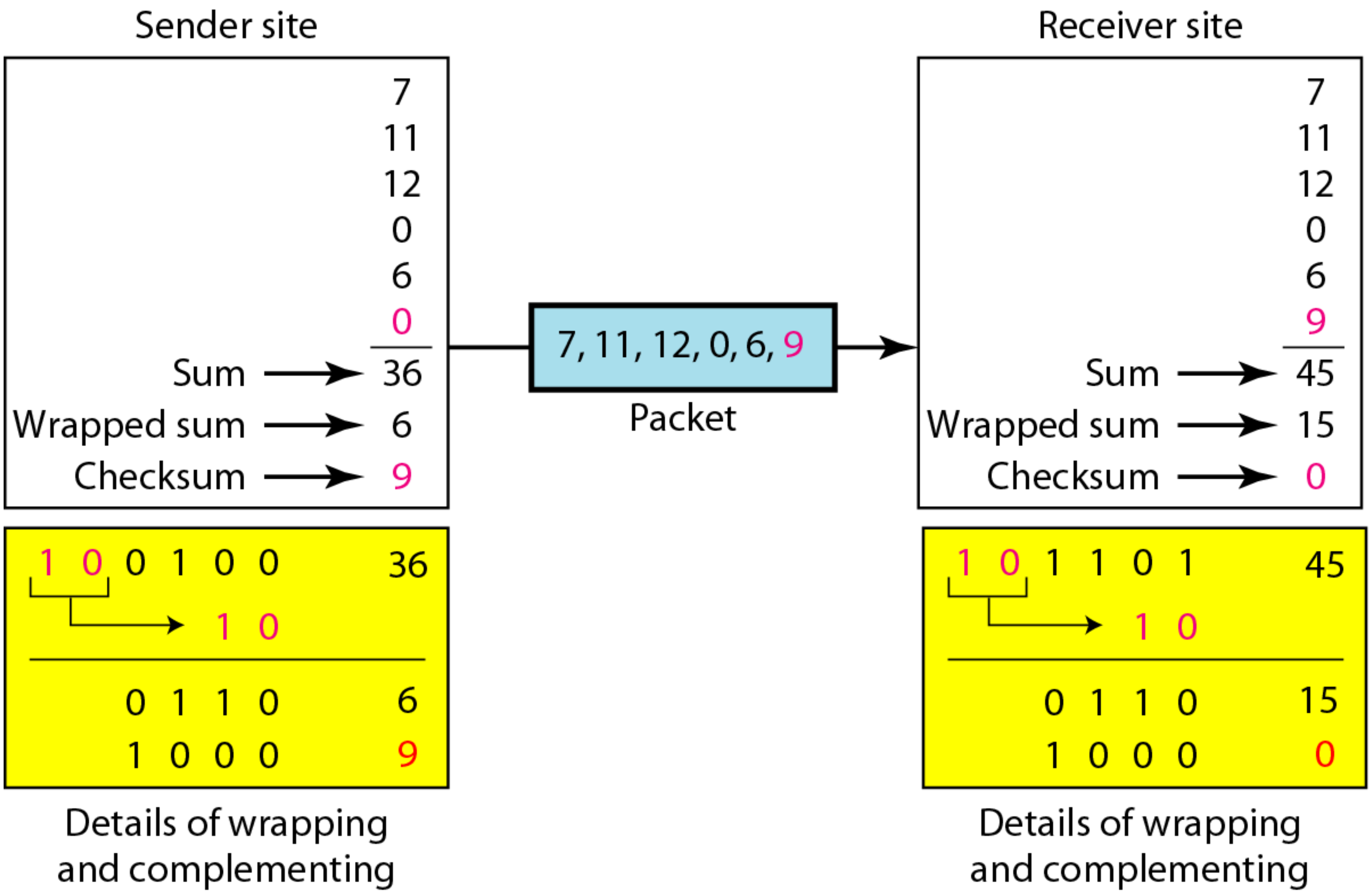
Example 10.22

Let us redo Exercise 10.19 using one's complement arithmetic. Figure 10.24 shows the process at the sender and at the receiver. The sender initializes the checksum to 0 and adds all data items and the checksum (the checksum is considered as one data item and is shown in color). The result is 36. However, 36 cannot be expressed in 4 bits. The extra two bits are wrapped and added with the sum to create the wrapped sum value 6. In the figure, we have shown the details in binary. The sum is then complemented, resulting in the checksum value 9 (15 − 6 = 9). The sender now sends six data items to the receiver including the checksum 9.

Example 10.22 (continued)

The receiver follows the same procedure as the sender. It adds all data items (including the checksum); the result is 45. The sum is wrapped and becomes 15. The wrapped sum is complemented and becomes 0. Since the value of the checksum is 0, this means that the data is not corrupted. The receiver drops the checksum and keeps the other data items. If the checksum is not zero, the entire packet is dropped.

Figure 10.24 *Example 10.22*



Note

Sender site:

1. The message is divided into 16-bit words.
2. The value of the checksum word is set to 0.
3. All words including the checksum are added using one's complement addition.
4. The sum is complemented and becomes the checksum.
5. The checksum is sent with the data.

Note

Receiver site:

1. The message (including checksum) is divided into 16-bit words.
2. All words are added using one's complement addition.
3. The sum is complemented and becomes the new checksum.
4. If the value of checksum is 0, the message is accepted; otherwise, it is rejected.

Example 10.23

Let us calculate the checksum for a text of 8 characters ("Forouzan"). The text needs to be divided into 2-byte (16-bit) words. We use ASCII (see Appendix A) to change each byte to a 2-digit hexadecimal number. For example, F is represented as 0x46 and o is represented as 0x6F. Figure 10.25 shows how the checksum is calculated at the sender and receiver sites. In part a of the figure, the value of partial sum for the first column is 0x36. We keep the rightmost digit (6) and insert the leftmost digit (3) as the carry in the second column. The process is repeated for each column. Note that if there is any corruption, the checksum recalculated by the receiver is not all 0s. We leave this an exercise.

Figure 10.25 *Example 10.23*

1	0	1	3	Carries
4	6	6	F	(Fo)
7	2	6	7	(ro)
7	5	7	A	(uz)
6	1	6	E	(an)
0	0	0	0	Checksum (initial)
8	F	C	6	Sum (partial)
			1	
8	F	C	7	Sum
7	0	3	8	Checksum (to send)

a. Checksum at the sender site

1	0	1	3	Carries
4	6	6	F	(Fo)
7	2	6	7	(ro)
7	5	7	A	(uz)
6	1	6	E	(an)
7	0	3	8	Checksum (received)
F	F	F	E	Sum (partial)
			1	
8	F	C	7	Sum
0	0	0	0	Checksum (new)

a. Checksum at the receiver site