Unit 2 Physical Layer

2.1 Data and Signals

Analog and Digital Data

- 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.
    - For example, an analog clock that has hour, minute, and second hands gives information in a continuous form; the movements of the hands are continuous.
    - On the other hand, a digital clock that reports the hours and the minutes will change suddenly from 8:05 to 8:06.
- When someone speaks, an analog wave is created in the air. This can be captured by a microphone and converted to an analog signal or sampled and converted to a digital signal.

Analog and Digital Data

- Digital data take on discrete values. For example, data are stored in computer memory in the form of 0s and 1s.
- They can be converted to a digital signal or modulated into an analog signal for transmission across a medium.

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

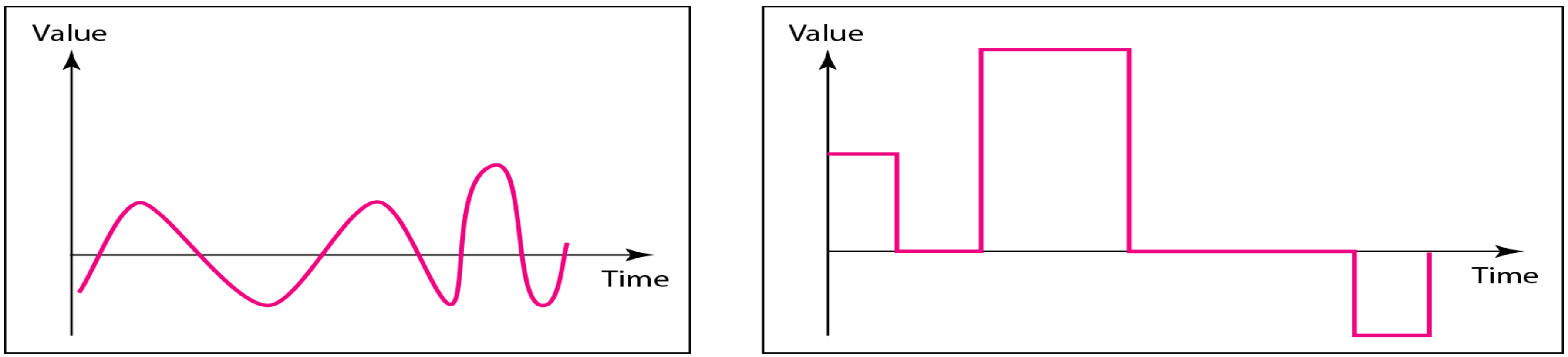
Analog and Digital Signals

- Signals can be either analog or digital. An analog signal has infinitely many levels of intensity over a period of time.
- As the wave moves from value A to value B, it passes through and includes an infinite number of values along its path.
- A digital signal, on the other hand, can have only a limited number of defined values.
- Although each value can be any number, it is often as simple as 1 and 0.

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.

Analog and Digital Signals

- Figure 3.1 illustrates an analog signal and a digital signal. The horizontal axis represents time. The vertical axis represents the value or strength of a signal.
- The curve representing the analog signal passes through an infinite number of points.
- The vertical lines of the digital signal, however, demonstrate the sudden jump that the signal makes from value to value.



a. Analog signal   b. Digital signal

Periodic and Nonperiodic Signals

- Both analog and digital signals can take one of two forms: periodic or nonperiodic (sometimes refer to as aperiodic, because the prefix a in Greek means "non").
- A periodic signal completes a pattern within a measurable time frame, called a period, and repeats that pattern over subsequent identical periods.
- The completion of one full pattern is called a cycle.
- A nonperiodic signal changes without exhibiting a pattern or cycle that repeats over time.
- Both analog and digital signals can be periodic or nonperiodic. In data communications, we commonly use periodic analog signals (because they need less bandwidth), and nonperiodic digital signals.

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

Periodic Analog 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.
- **Sine Wave**
- The sine wave is the most fundamental form of a periodic analog signal. When we visualize it as a simple oscillating curve, its change over the course of a cycle is smooth and consistent, a continuous, rolling flow.
- Figure 3.2 shows a sine wave. Each cycle consists of a single arc above the time axis followed by a single arc below it.

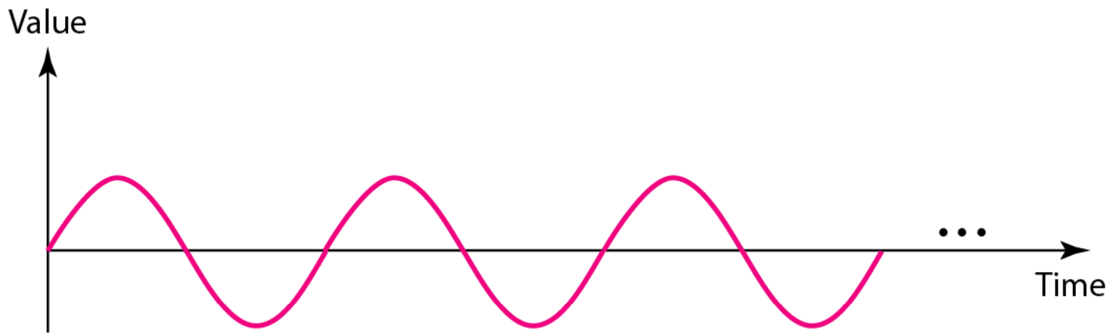


Figure 3.2 *A sine wave*

Periodic Analog Signals

- A sine wave can be represented by three parameters: the peak amplitude, the frequency, and the phase. These three parameters fully describe a sine wave.
- Peak Amplitude
- The peak amplitude of a signal is the absolute value of its highest intensity, proportional to the energy it carries. For electric signals, peak amplitude is normally measured in volts.
- Figure 3.3 shows two signals and their peak amplitudes.

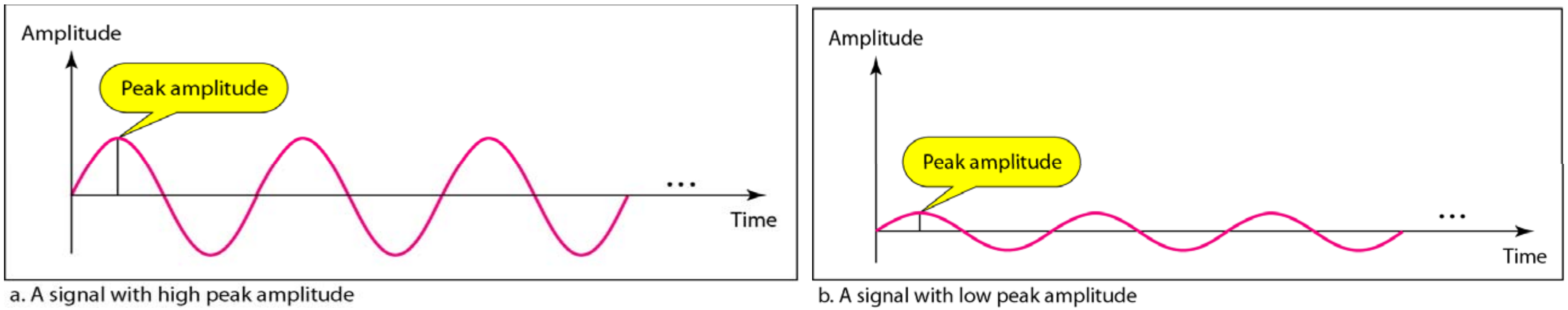


Figure 3.3 *Two signals with the same phase and frequency, but different amplitudes*

Periodic Analog Signals

- The power in your house can be represented by a sine wave with a peak amplitude of 230 to 250 V.
- Period and Frequency
- Period refers to the amount of time, in seconds, a signal needs to complete 1 cycle.
- Frequency refers to the number of periods in I s, expressed in Hertz (Hz).
- 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.
- Figure 3.4 shows two signals and their frequencies.

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

Frequency and period are the inverse of each other.

Periodic Analog Signals

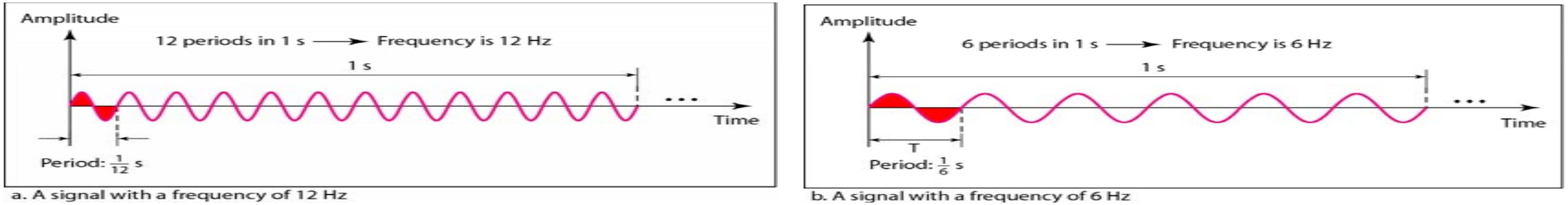


Figure 3.4 ***Two signals with the same amplitude and phase, but different frequencies***

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

The voltage of a battery is a constant; this constant value can be considered a sine wave, as we will see later. For example, the peak value of an AA battery is normally 1.5 V.

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 3.3

The power we use at home has a frequency of 60 Hz (50 Hz in Europe). The period of this sine wave can be determined as follows:

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

Example 3.4

Express a period of 100 ms in microseconds.

Solution

From Table 3.1 we find the equivalents of 1 ms (1 ms is 10^{-3} s) and 1 s (1 s is 10^6 μs). We make the following substitutions:

$$100 \text{ ms} = 100 \times 10^{-3} \text{ s} = 100 \times 10^{-3} \times 10^{6} \text{ μs} = 10^{2} \times 10^{-3} \times 10^{6} \text{ μs} = 10^{5} \text{ μs}$$

Example 3.5

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

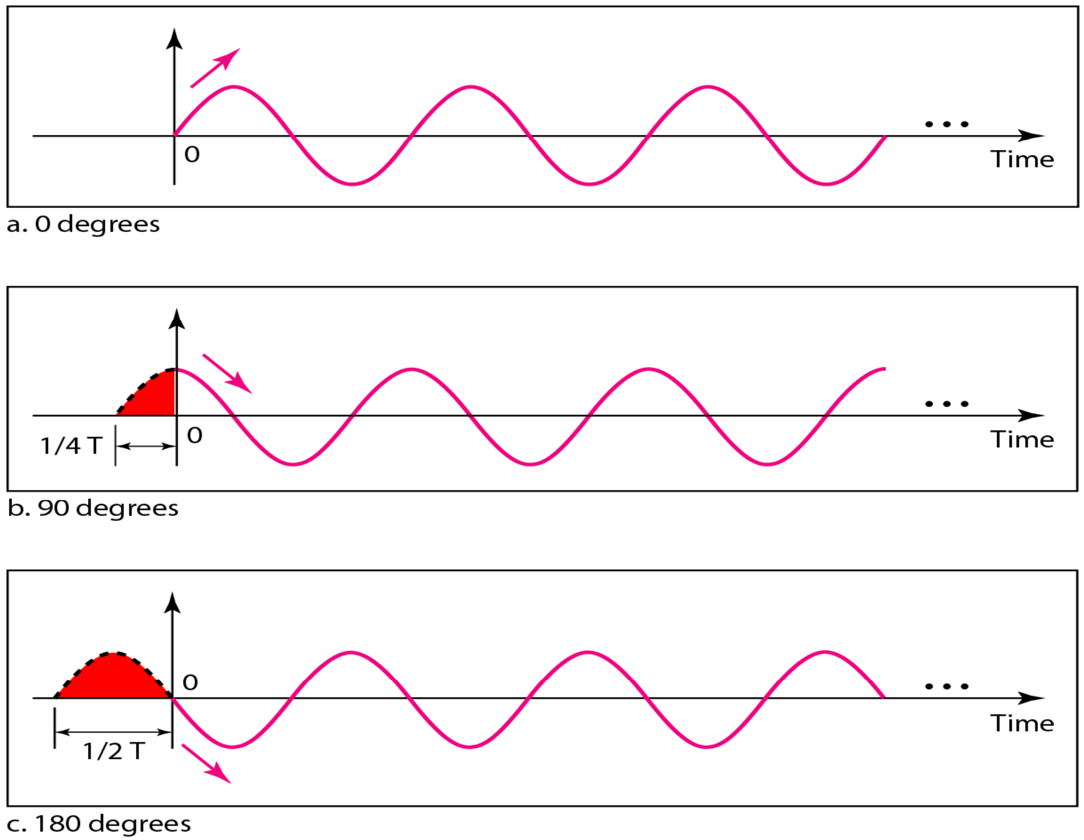
- frequency is the relationship of a signal to time and that the frequency of a wave is the number of cycles it completes in 1 s.
- But another way to look at frequency is as a measurement of the rate of change.
- Electromagnetic signals are oscillating waveforms; that is, they fluctuate continuously and predictably above and below a mean energy level.
- If a signal does not change at all, its frequency is zero.
- If a signal changes instantaneously, its frequency is (1/0) infinite.
- Phase describes the position of waveform relative to time 0.

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.

Frequency

- A sine wave with a phase of 0° starts at time 0 with a zero amplitude. The amplitude is increasing.
- A sine wave with a phase of 90° starts at time 0 with a peak.
- A sine wave with a phase of 180° starts at time 0 with a zero amplitude.

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

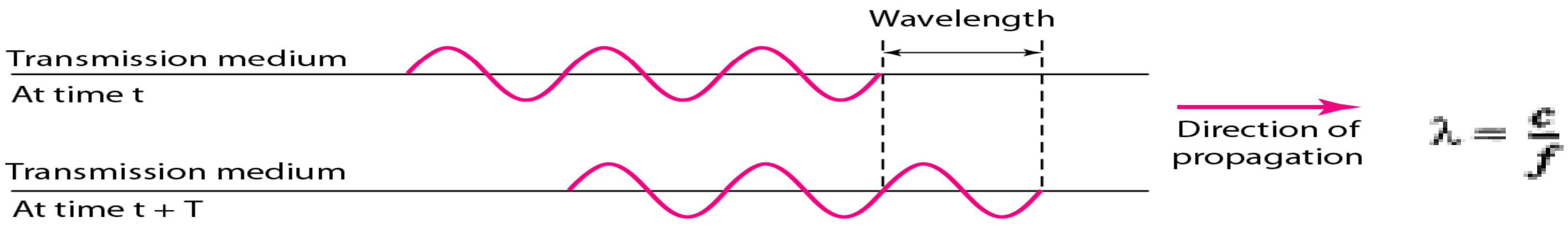


Phase

- Phase describes the position of waveform relative to time 0.
- Phase is measured in degrees or radians A phase shift of 360° corresponds to a shift of a complete period; a phase shift of 180° corresponds to a shift of one-half of a period; and a phase shift of 90° corresponds to a shift of one-quarter of a period (see Figure 3.5).

Wavelength and propagation speed

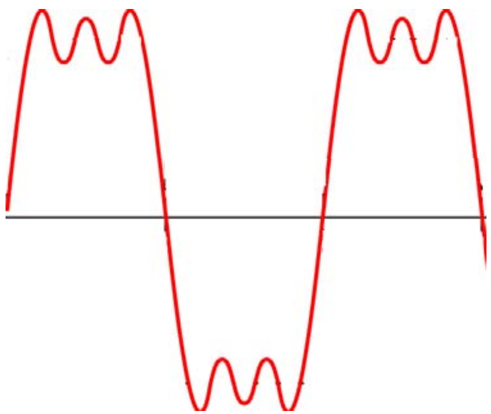
- Wavelength: the distance a simple signal can travel in one period.
- Propagation speed: the rate at which a signal or bit travels; measured by distance/second.
- Propagation time: the time required for a signal to travel from one point to another.
- Wavelength = propagation speed x period = propagation speed /frequency



- The propagation speed of electromagnetic signals depends on the medium and on the frequency of the signal. For example, in a vacuum, light is propagated with a speed of 3 x 108 m/s. That speed is lower in air and even lower in cable.
- The wavelength is normally measured in micrometers (microns) instead of meters. In a coaxial or fiber-optic cable, however, the wavelength is shorter (0.5 gm) because the propagation speed in the cable is decreased.

Composite signal

- 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.
- A composite signal is a combination of two or more simple sine waves with different frequency, phase and amplitude.
- According to Fourier analysis, any composite signal is a combination of simple sine waves with different frequencies, amplitudes, and phases.
- 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.

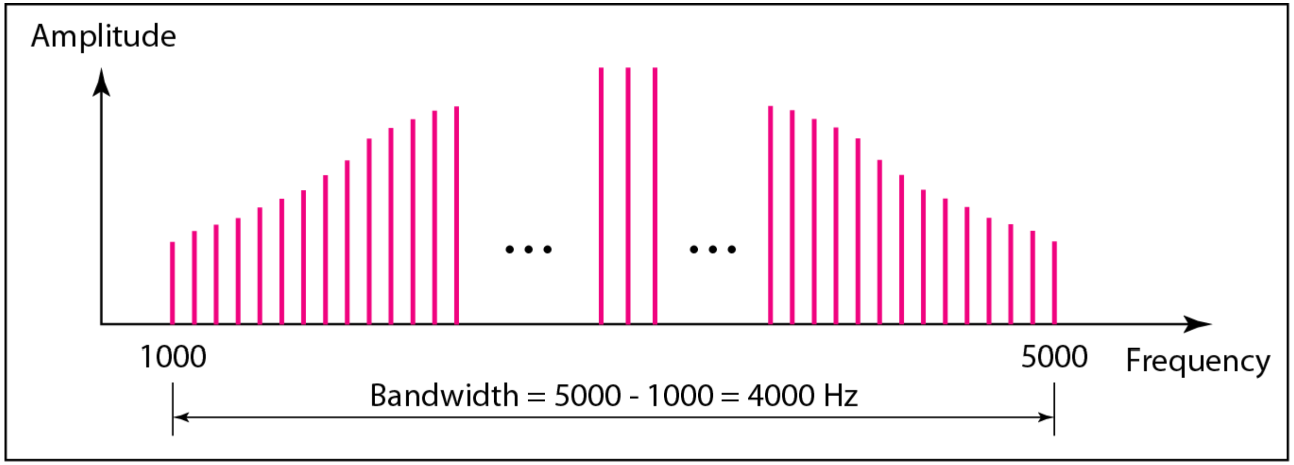


Bandwidth

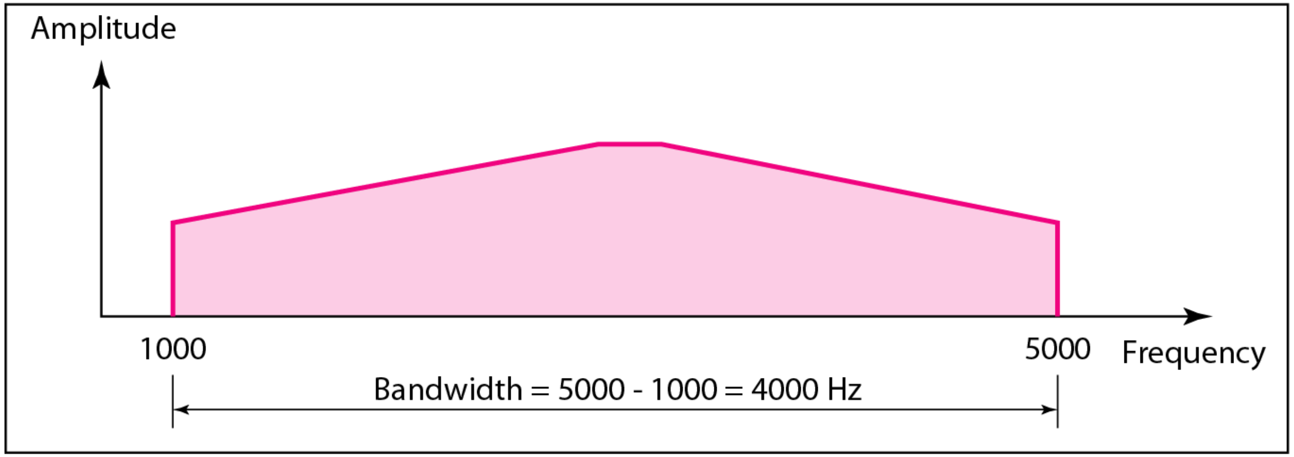
- The range of frequencies contained in a composite signal is its bandwidth. The bandwidth is normally a difference between two numbers.
- For example, if a composite signal contains frequencies between 1000 and 5000, its bandwidth is 5000 - 1000, or 4000.
- Figure 3.12 shows the concept of bandwidth. The figure depicts two composite signals, one periodic and the other nonperiodic.
- The bandwidth of the periodic signal contains all integer frequencies between 1000 and 5000 (1000, 100 l, 1002, ...).
- The bandwidth of the nonperiodic signals has the same range, but the frequencies are continuous.

Bandwidth

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



a. Bandwidth of a periodic signal

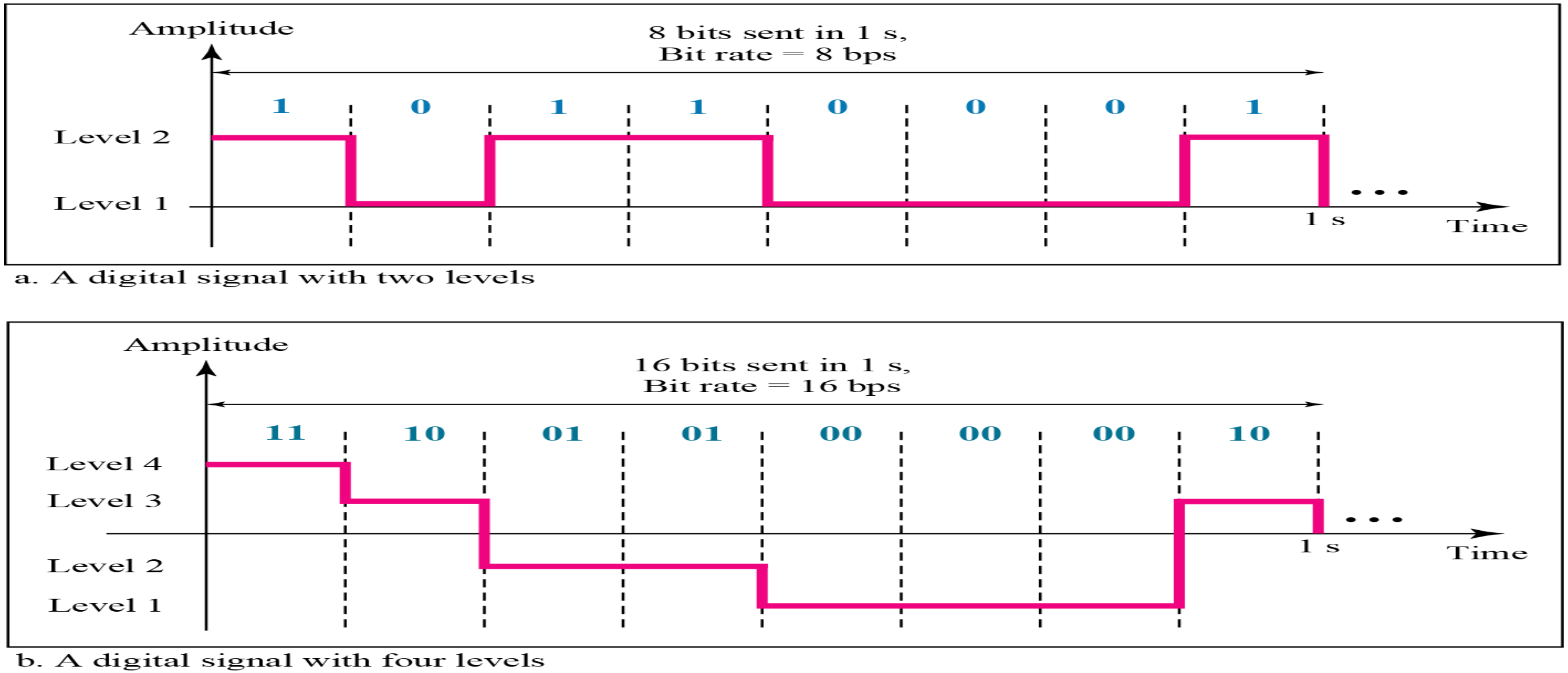


b. Bandwidth of a nonperiodic signal

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.

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



Bit Rate

- Most digital signals are nonperiodic, and thus period and frequency are not appropriate characteristics.
- Another term-bit rate (instead of frequency)-is used to describe digital signals.
- The bit rate is the number of bits sent in ls, expressed in bits per second (bps). Figure 3.16 shows the bit rate for two signals.

Example 3.18

Assume we need to download text documents at the rate of 100 pages per minute. 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, the bit rate is

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

Bit Length

- We discussed the concept of the wavelength for an analog signal: the distance one cycle occupies on the transmission medium.
- We can define something similar for a digital signal: the bit length. The bit length is the distance one bit occupies on the transmission medium.

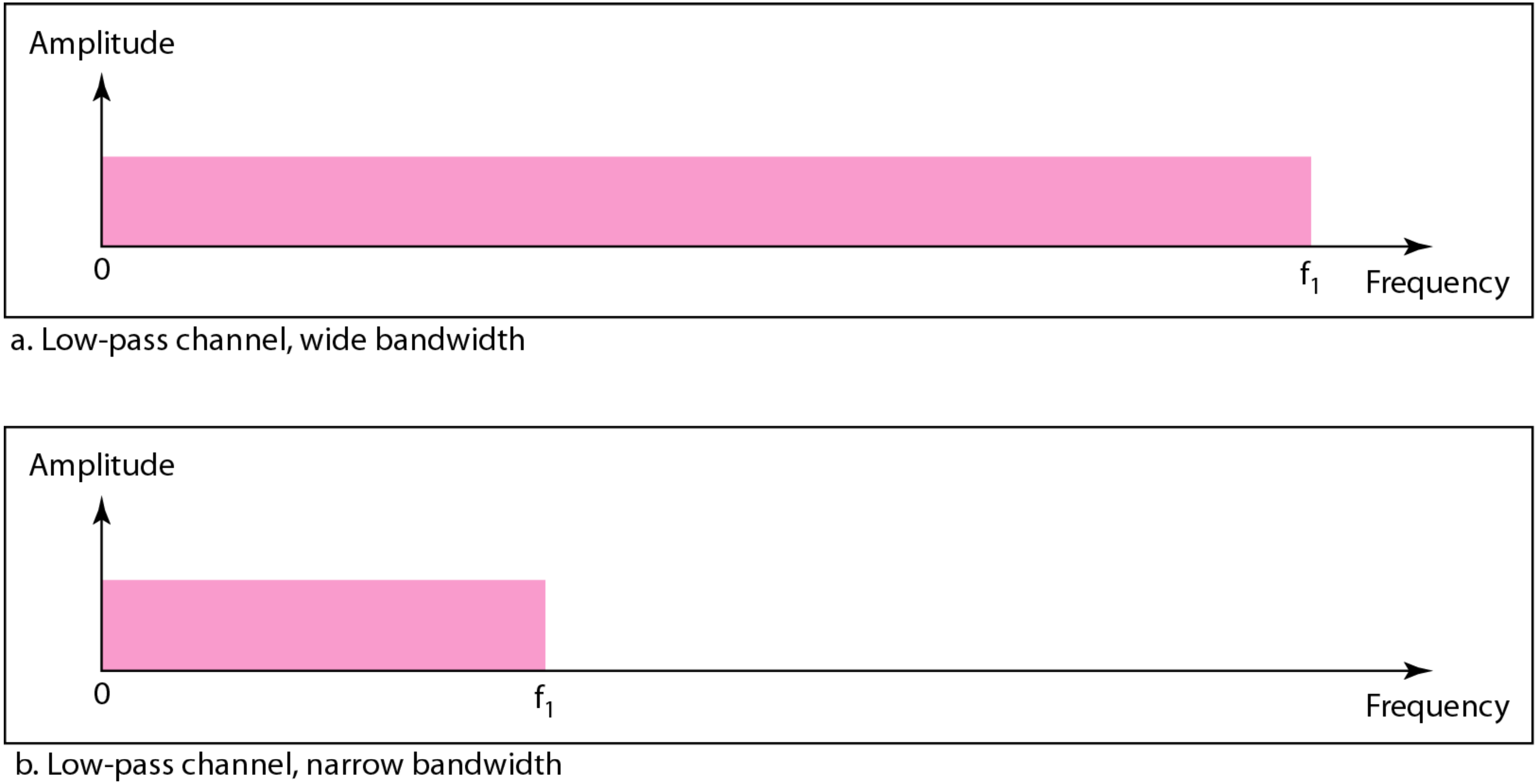
Bit length = propagation speed * bit duration

Digital Signal as a Composite Analog Signal

- Based on Fourier analysis, a digital signal is a composite analog signal.
- The bandwidth is infinite, but the periodic signal has discrete frequencies while the nonperiodic signal has continuous frequencies.

Figure 3.19 *Bandwidths of two low-pass channels*

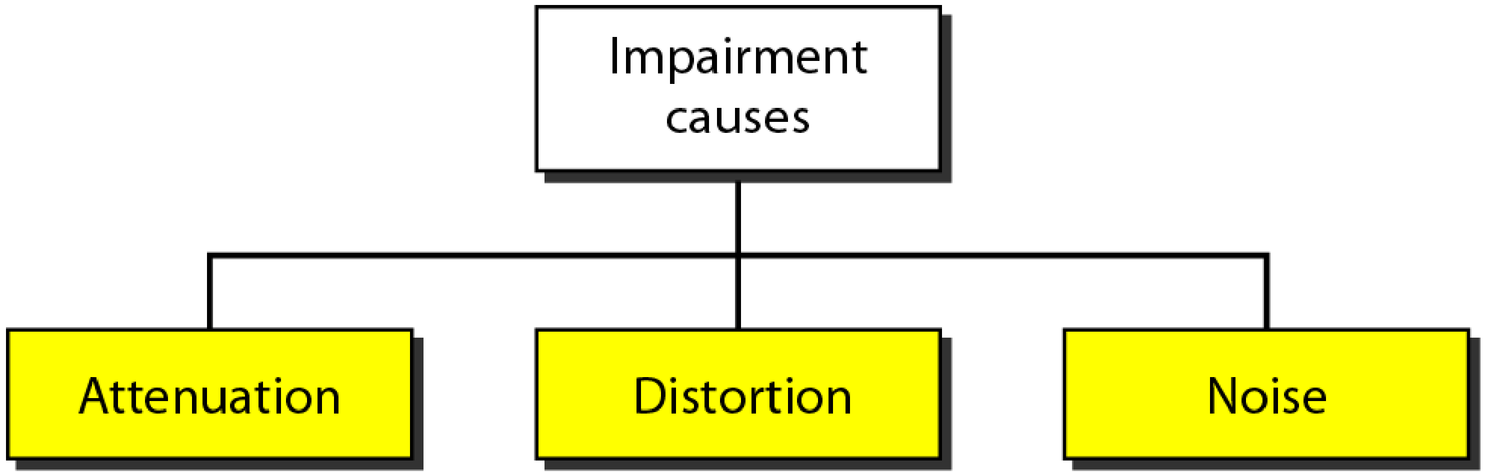
- A low-pass channel with infinite band-width is ideal, but we cannot have such a channel in real life. However, we can get close.



TRANSMISSION IMPAIRMENT

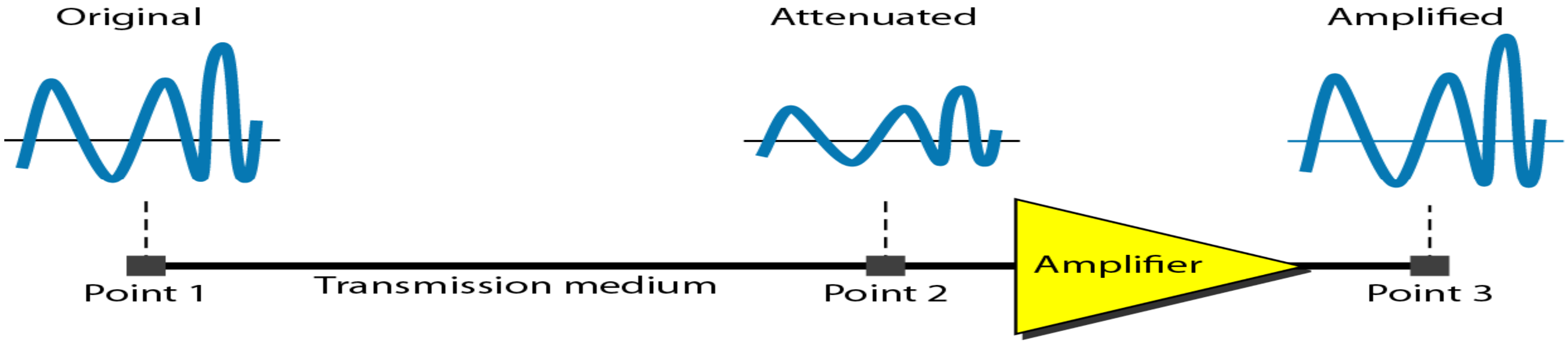
- Signals travel through transmission media, which are not perfect. The imperfection causes signal impairment. This means that the signal at the beginning of the medium is not the same as the signal at the end of the medium.
- What is sent is not what is received. Three causes of impairment are attenuation,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

Figure 3.25 ***Causes of impairment***



- ***Attenuation*** means a loss of energy. When a signal, simple or composite, travels through a medium, it loses some of its energy in overcoming the resistance of the medium.
- That is why a wire carrying electric signals gets warm, if not hot, after a while. Some of the electrical energy in the signal is converted to heat.
- To compensate for this loss, amplifiers are used to amplify the signal. Figure 3.26 shows the effect of attenuation and amplification.

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 P2 is (1/2)P1. 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 P2 = 10P1 . In this case, the amplification (gain of power) can be calculated as:

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

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

Example 3.29

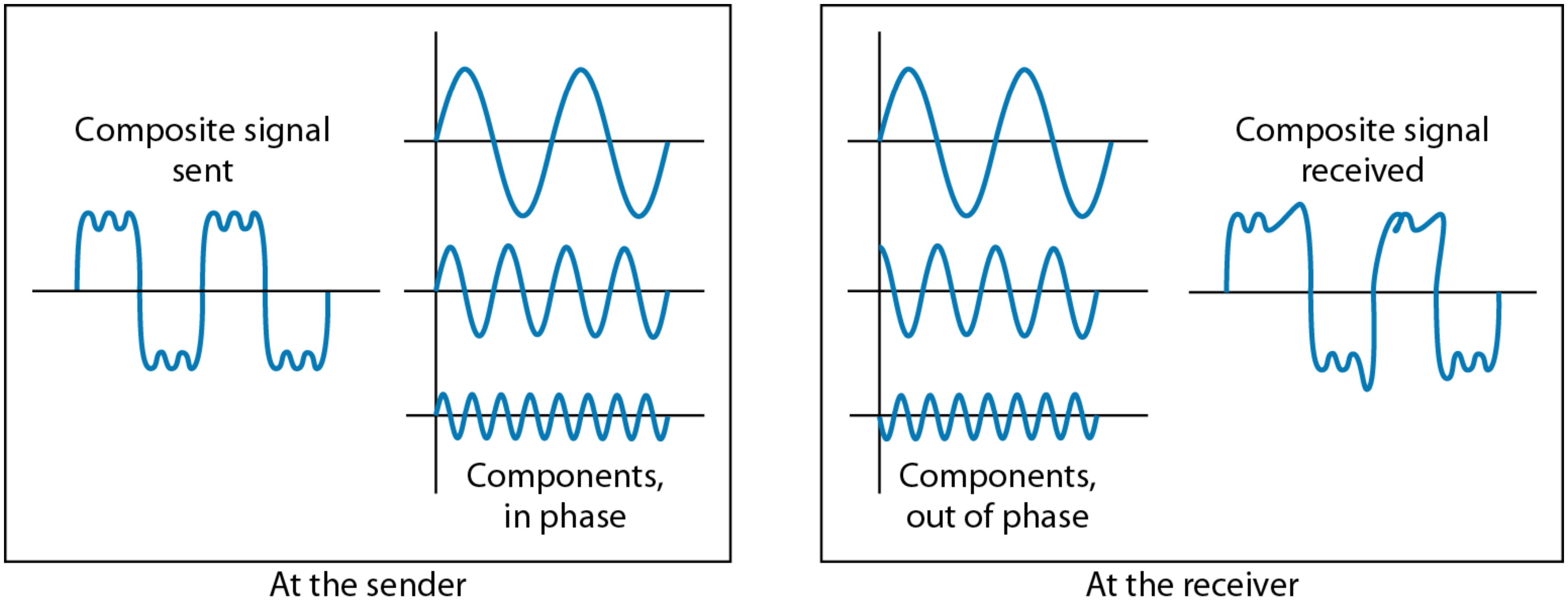
- Sometimes the decibel is used to measure signal power in milliwatts. In this case, it is referred to as dBm and is calculated as dBm = 10 log10 Pm , where Pm is the power in milliwatts. Calculate the power of a signal with dBm = −30.
- Solution
- We can calculate the power in the signal as

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

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

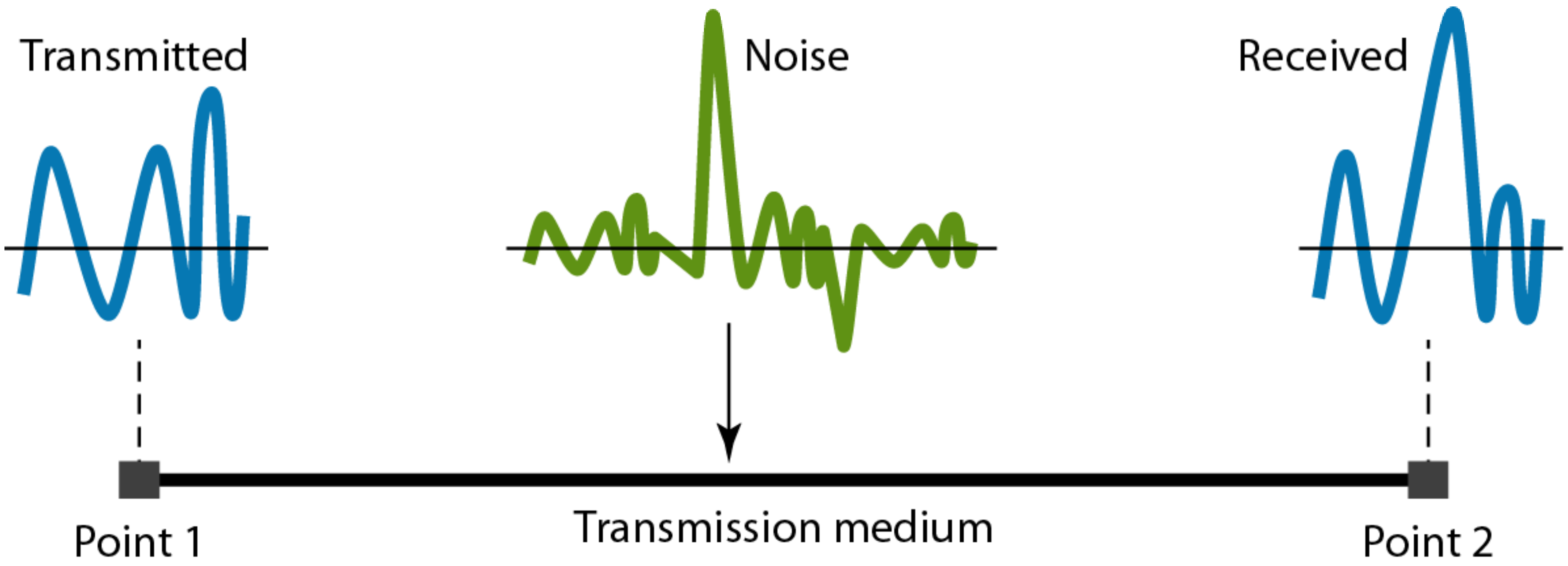
- ***Distortion*** means that the signal changes its form or shape. A digital signal is a composite analog signal with an infinite bandwidth. Distortion can occur in a composite signal made of different frequencies.
- Each signal component has its own propagation speed through a medium and, therefore, its own delay in arriving at the final destination.
- Differences in delay may create a difference in phase if the delay is not exactly the same as the period duration. In other words, signal components at the receiver have phases different from what they had at the sender.
- The shape of the composite signal is therefore not the same.

Figure 3.28 *Distortion*



- ***Noise*** is another cause of impairment. Several types of noise may corrupt the signal.
- Thermal noise is the random motion of electrons in a wire which creates an extra signal not originally sent by the transmitter.
- Induced noise comes from sources such as motors and appliances. These devices act as a sending antenna, and the transmission medium acts as the receiving antenna.
- Crosstalk is the effect of one wire on the other. One wire acts as a sending antenna and the other as the receiving antenna.
- Impulse noise is a spike (a signal with high energy in a very short time) that comes from power lines, lightning, and so on.

Figure 3.29
Noise



- Signal-to-Noise Ratio (SNR)
- To find the theoretical bit rate limit, we need to know the ratio of the signal power to the noise power. The signal-to-noise ratio is defined as
- SNR = average signal power \average noise power
- We need to consider the average signal power and the average noise power because these may change with time.
- SNR is actually the ratio of what is wanted (signal) to what is not wanted (noise).
- A high SNR means the signal is less corrupted by noise; a low SNR means the signal is more corrupted by noise. Because SNR is the ratio of two powers, it is often described in decibel units, SNRdB, defined as $\mathbf{SNR = 10 \log_{10} SNR}$

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:

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

$$\mathrm{SNR_{dB}} = 10 \log_{10} 10{,}000 = 10 \log_{10} 10^4 = 40$$

Example 3.32

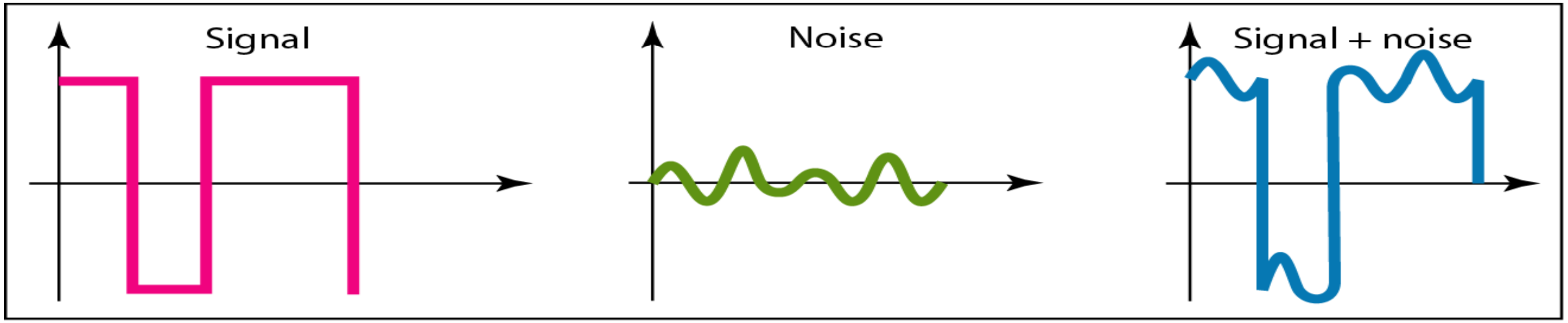
The values of SNR and SNR_{dB} for a noiseless channel are

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

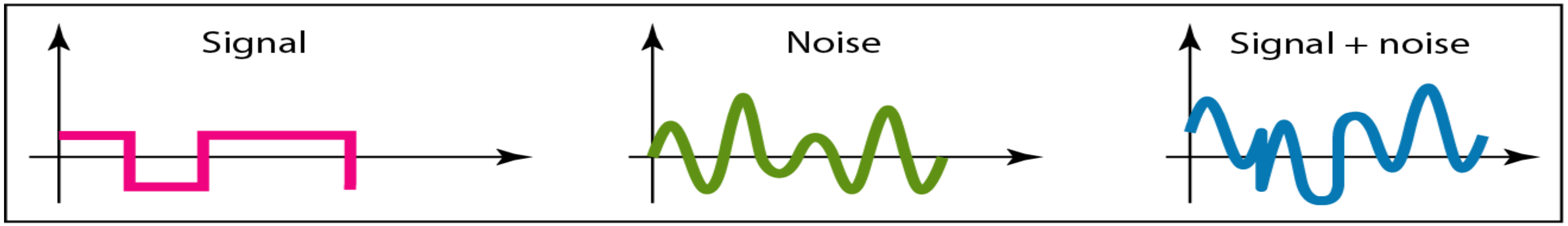
$$\text{SNR}_{\text{dB}} = 10 \log_{10} \infty = \infty$$

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

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)

- Increasing the levels of a signal may reduce the reliability of the system.
- Two theoretical formulas were developed to calculate the data rate: one by Nyquist for a noiseless channel, another by Shannon for a noisy channel. The Shannon capacity gives us the upper limit; the Nyquist formula tells us how many signal levels we need.
- Noiseless channel: the Nyquist bit rate formula defines the theoretical maximum bit rate. $\mathbf{BitRate = 2 * bandwidth * \log_2 L}$

- Noisy Channel: Shannon Capacity : In reality, we cannot have a noiseless channel; the channel is always noisy. In 1944, Claude Shannon introduced a formula, called the Shannon capacity, to determine the theoretical highest data rate for a noisy channel:

$$\mathbf{Capacity = bandwidth * \log_2 (1 + SNR)}$$

PERFORMANCE

- One important issue in networking is the performance of the network—how good is it? The performance is measured by

Bandwidth; Throughput; Latency (Delay); Bandwidth-Delay Product

Bandwidth: 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.
- An increase in bandwidth in hertz means an increase in bandwidth in bits per second.

Throughput and Latency

- The throughput is a measure of how fast we can actually send data through a network. Bandwidth and throughput is not same. For example a link with bandwidth 1 Mbps but can handle only 200 kbps so throughput will be 200 kbps.
- The latency or delay defines how long it takes for an entire message to completely arrive at the destination from the first bit is sent out from the source.
- Latency= propagation time+ transmission time + queuing time + processing time
- Propagation time measures the time required for a bit to travel from the source to the destination.
- Transmission time is the time between the first bit leaving the sender and the last bit arriving at the receiver.
- Queuing time is the required time for each intermediate or end device to hold the message before it can be processed.