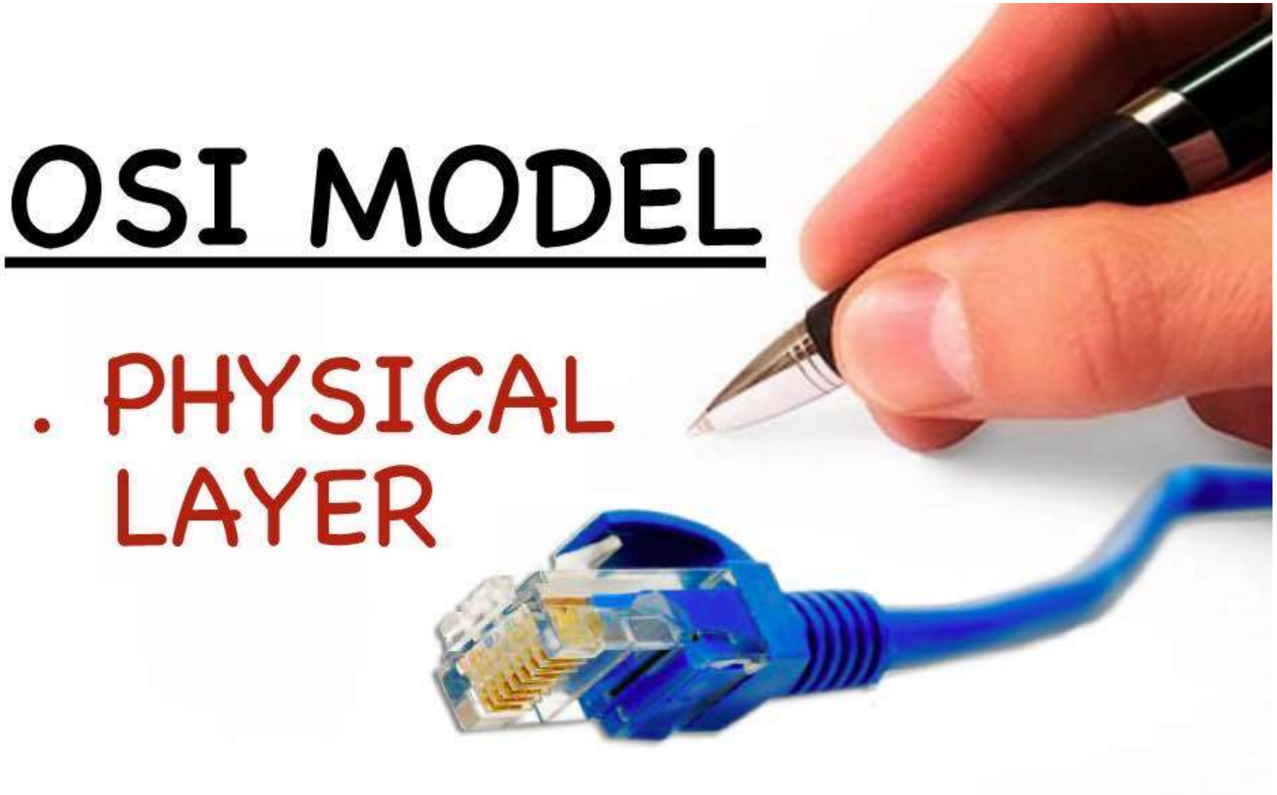
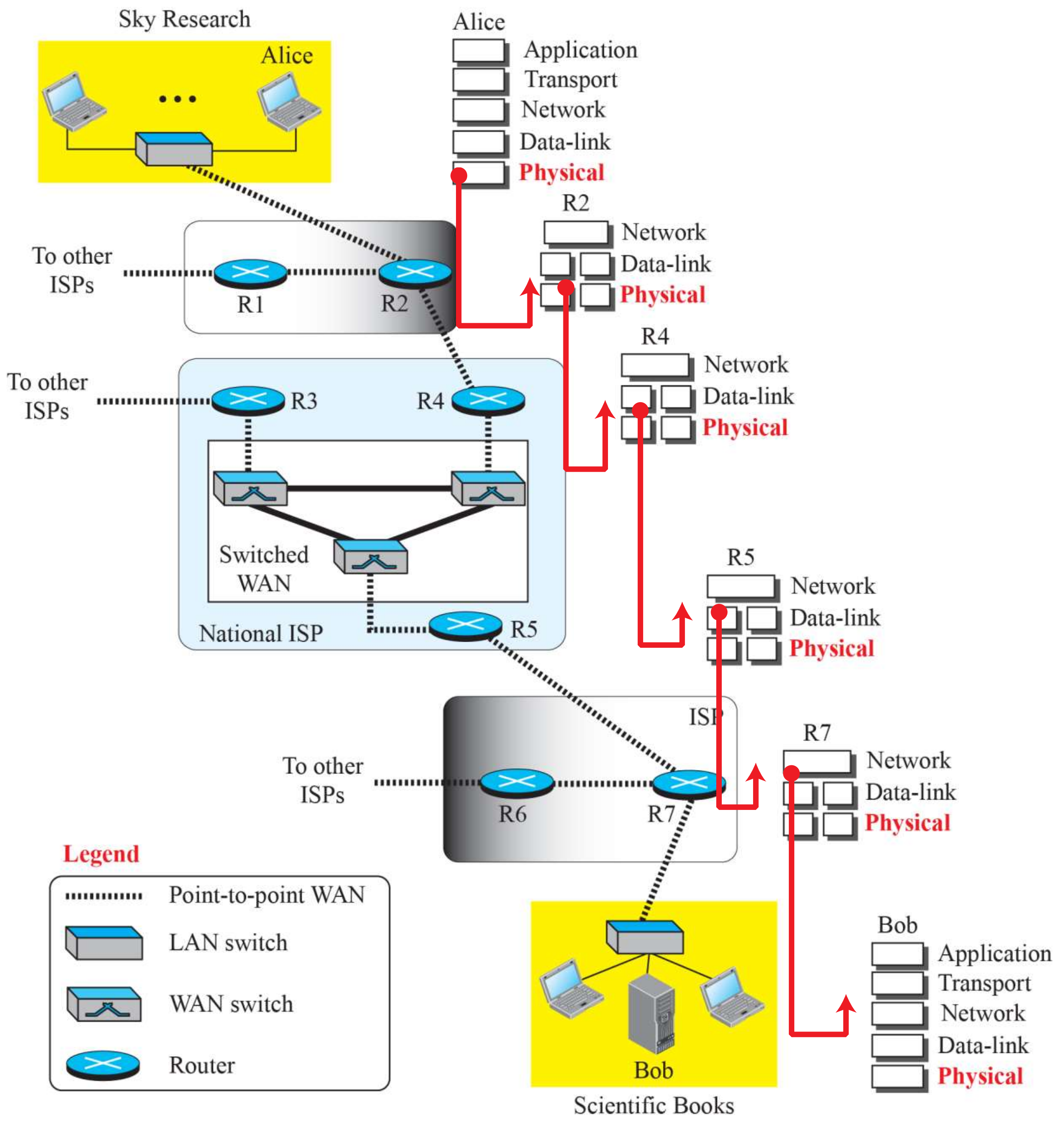
Chapter 3

Introduction To Physical Layer



Objective

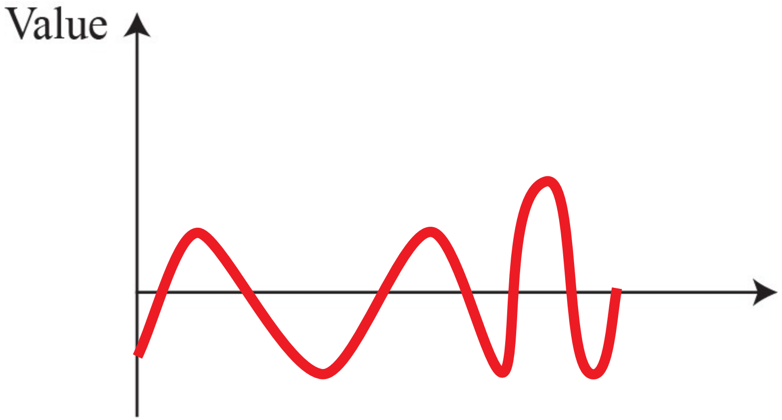
- Analog and digital.
- Periodic analog signals and non-periodic digital signals in data communication.
- How digital data can be sent using analog signals.
- Transmission impairment, data rate limit, and the performance of data transmission including bandwidth, throughput, latency, and jitter.



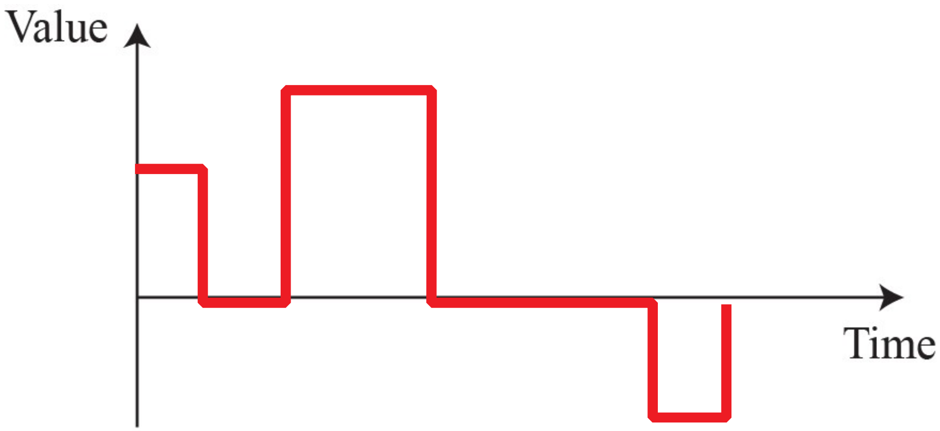
Communication at the physical layer

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.
- An analog signal has infinitely many levels of intensity over a period of time.
- A digital signal can have only a limited number of defined values; it is often as simple as 1 and 0.



a. Analog signal



b. Digital signal

Comparison of analog and digital signals

Periodic and Nonperiodic

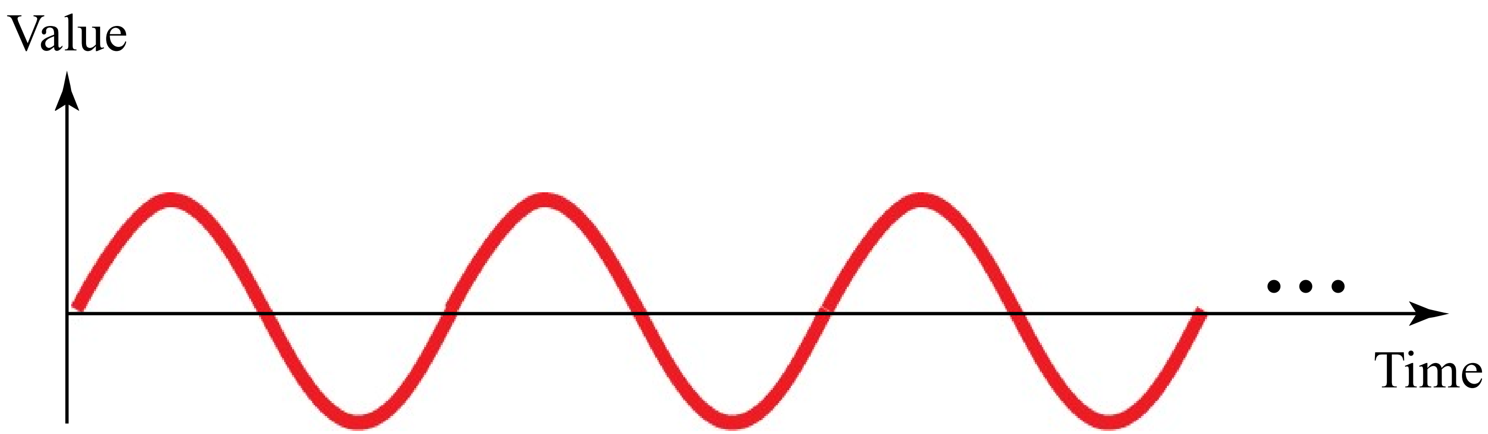
- A periodic signal completes a pattern within a period, and repeats that pattern over subsequent identical periods.
- The completion of one full pattern is called a **cycle**.
- A nonperiodic signal changes without exhibiting a pattern or cycle that repeats over time.

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

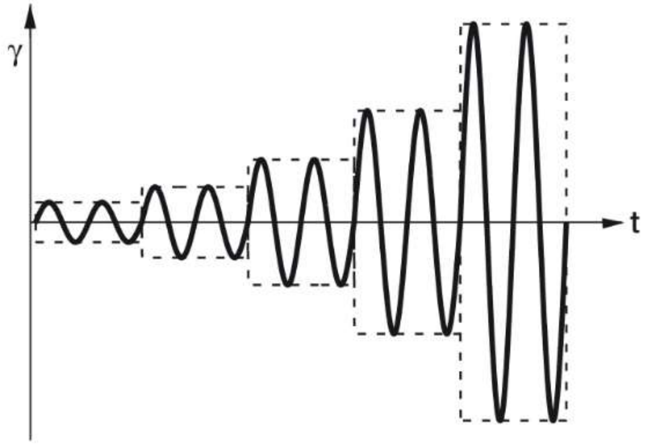
- The sine wave is the most fundamental form of a periodic analog signal. When we visualize it as a simple oscillating curve, its change over the course of a cycle is smooth and consistent, a continuous, rolling flow.

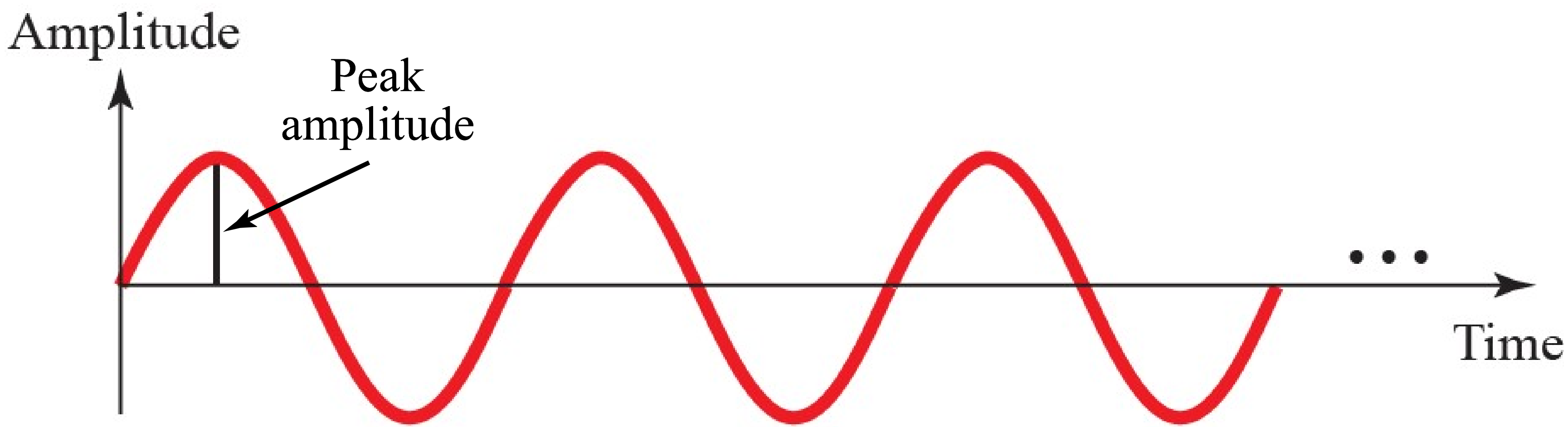


A sine wave

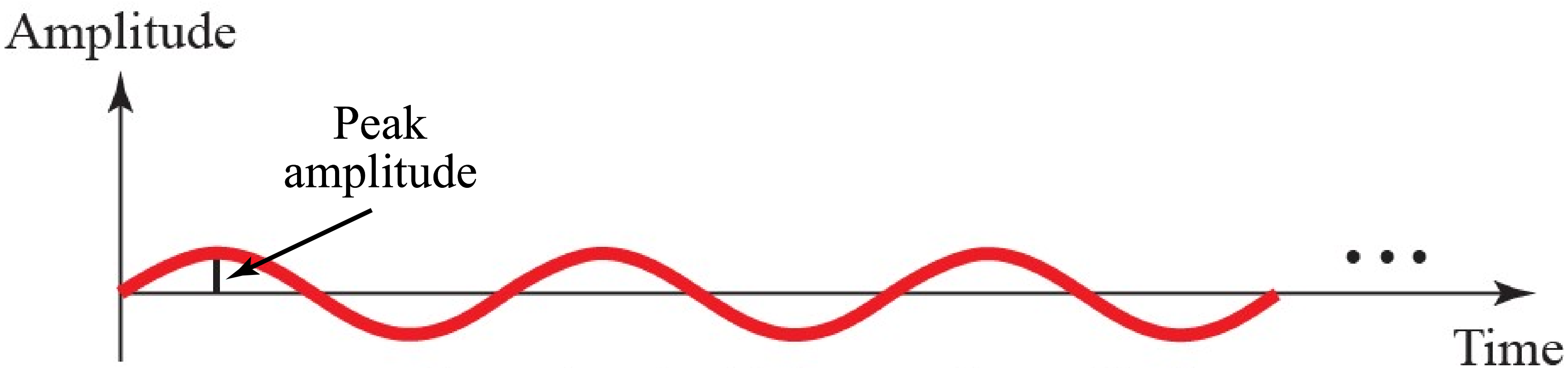
Amplitude

- The amplitude of a wave is a measure of how big its oscillation is.
- Amplitudes are always measured as positive numbers. That's because distance can only be greater than zero or equal to zero; negative distance does not exist.



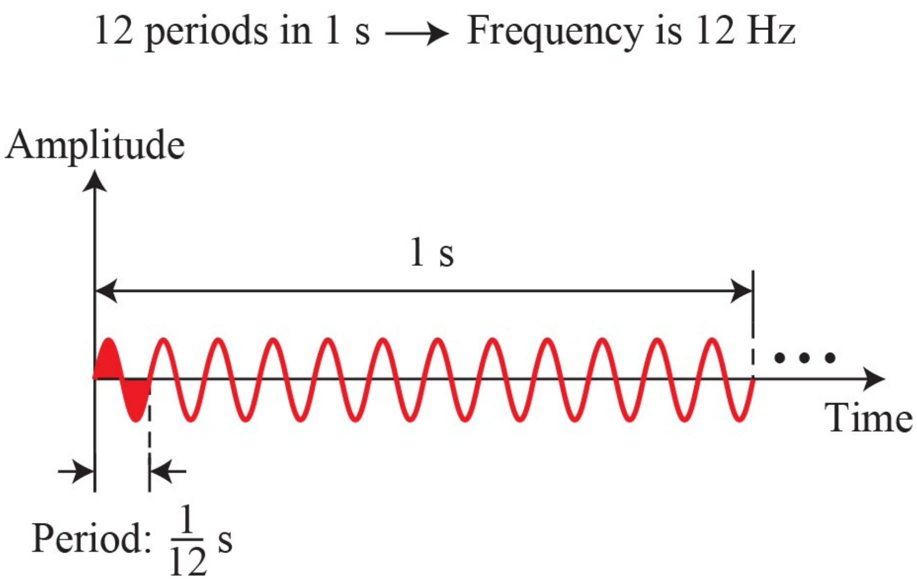


a. A signal with high peak amplitude

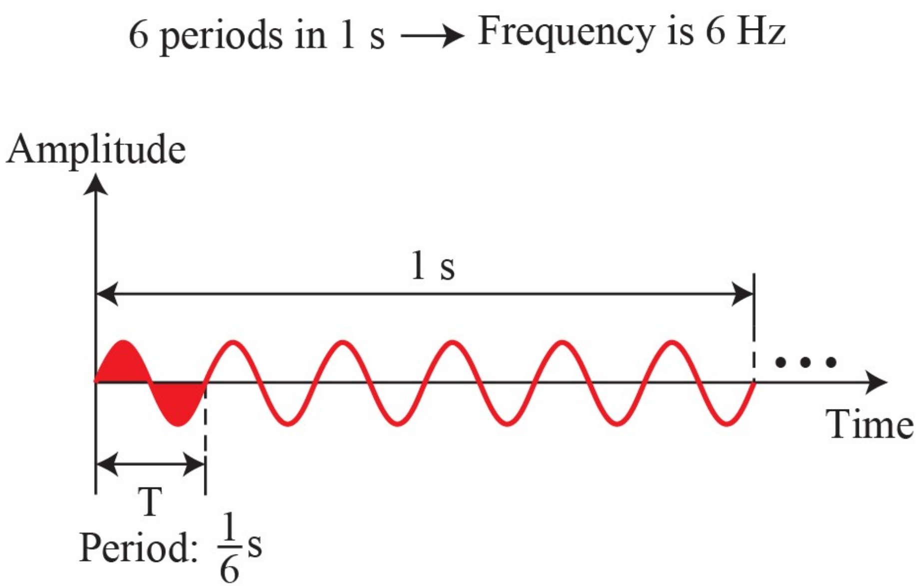


b. A signal with low peak amplitude

Two signals with two different amplitudes



a. A signal with a frequency of 12 Hz



b. A signal with a frequency of 6 Hz

Units of period and frequency

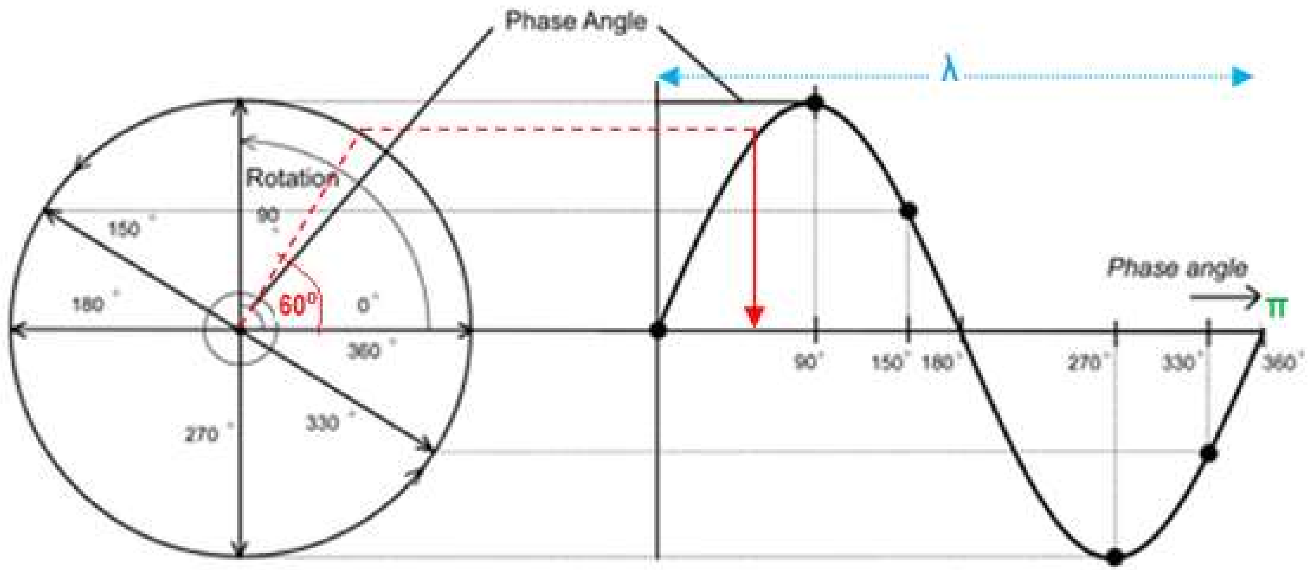
Period		*Frequency*	
Unit	*Equivalent*	*Unit*	*Equivalent*
Seconds (s)	1 s	Hertz (Hz)	1 Hz
Milliseconds (ms)	10^{-3} s	Kilohertz (kHz)	10^{3} Hz
Microseconds (μs)	10^{-6} s	Megahertz (MHz)	10^{6} Hz
Nanoseconds (ns)	10^{-9} s	Gigahertz (GHz)	10^{9} Hz
Picoseconds (ps)	10^{-12} s	Terahertz (THz)	10^{12}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

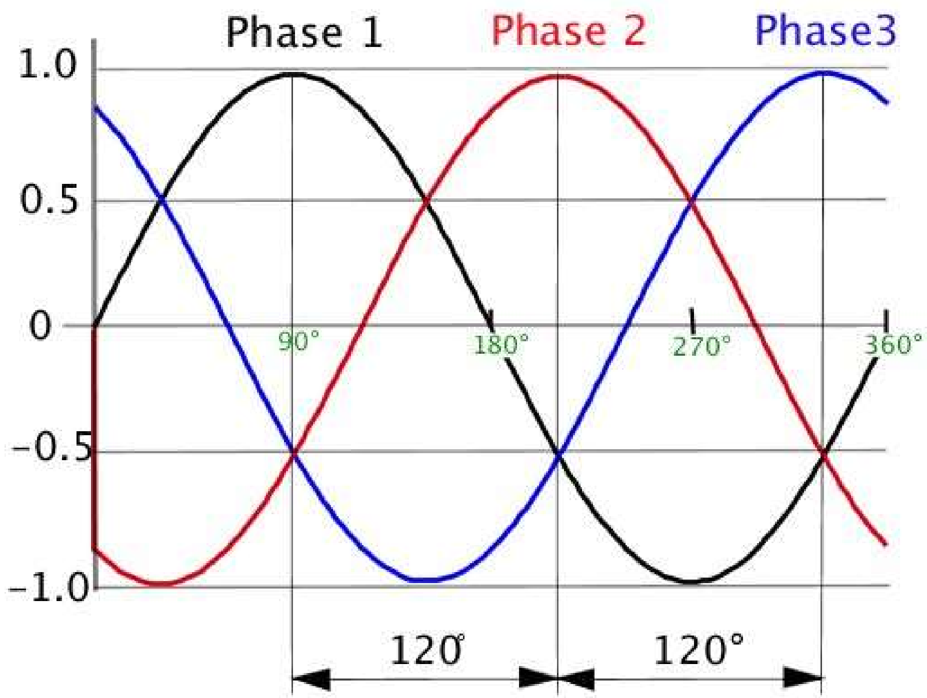
Phase

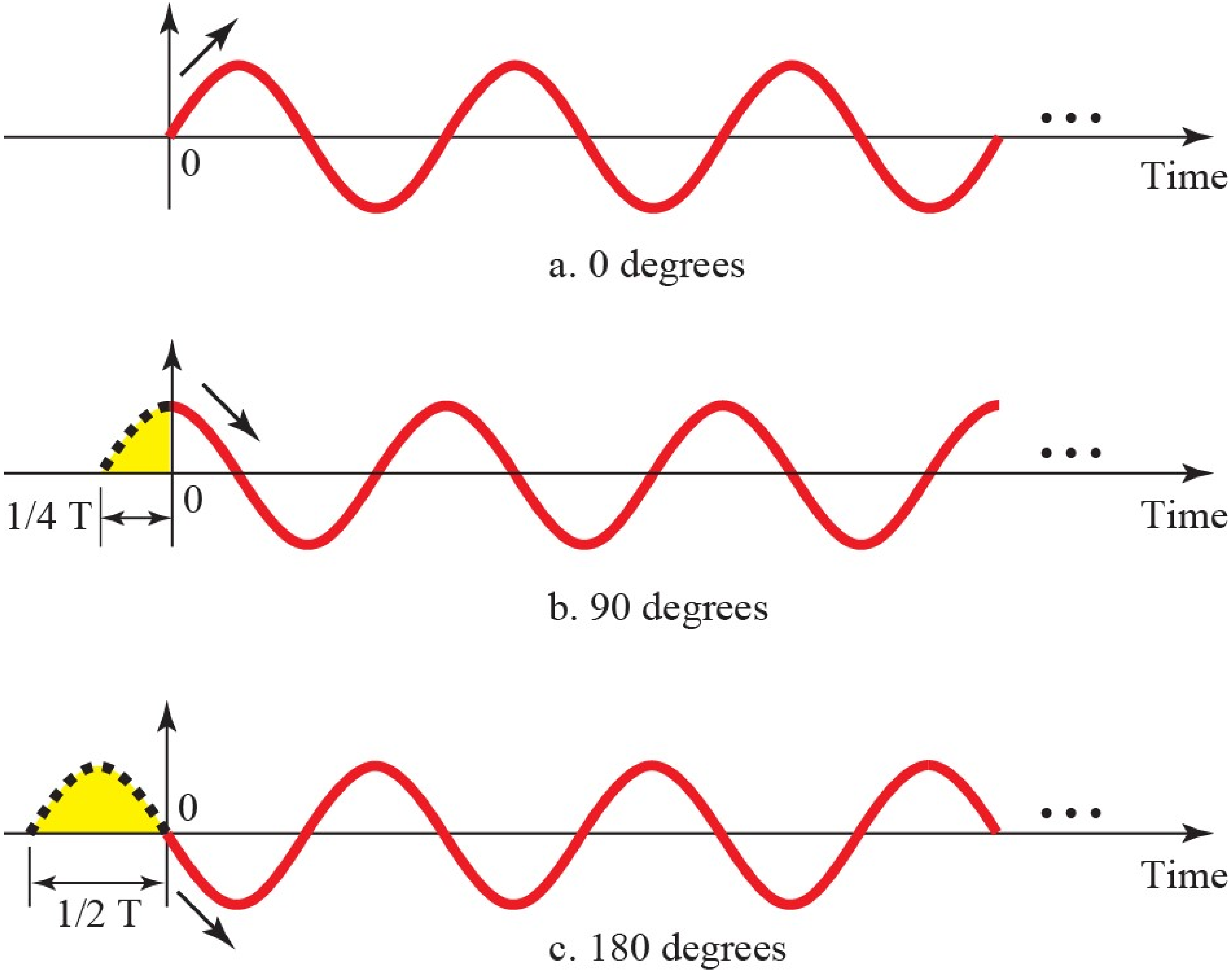
- The term phase, or phase shift, describes the position of the waveform relative to time 0.

- If we think of the wave as something that can be shifted backward or forward along the time axis, phase describes the amount of that shift.

- It indicates the status of the first cycle.


Phase Angle
λ
Rotation
90
150
180
0
60°
360
270
330
Phase angle
π
90°
150°
180°
270°
330°
360°


Phase 1
Phase 2
Phase3
1.0
0.5
0
-0.5
-1.0
90°
180°
270°
360°
120°
120°

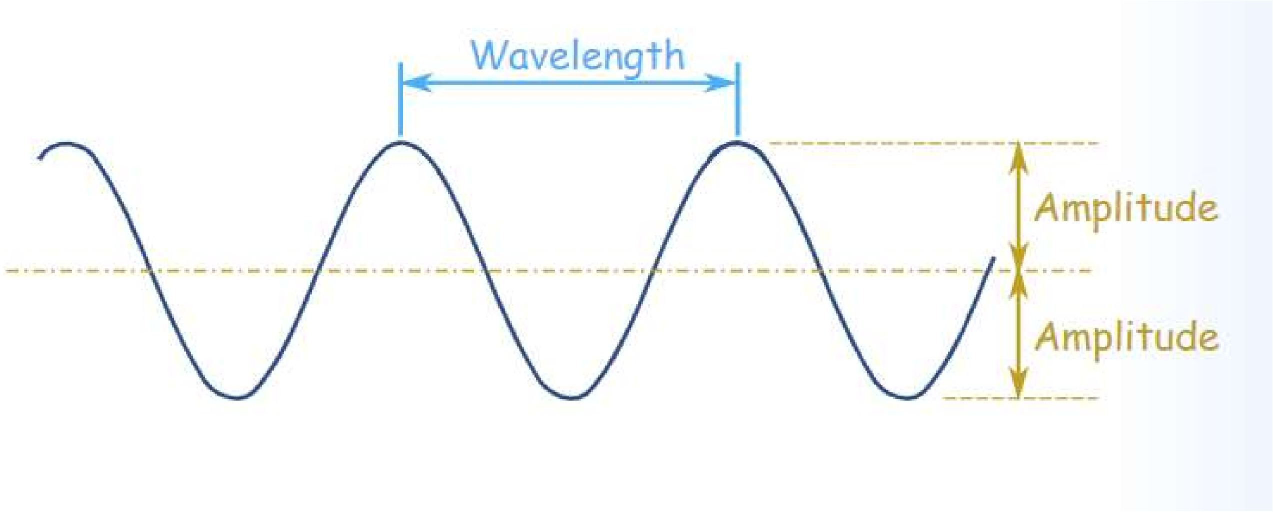


Three sine waves with different phases

Wavelength

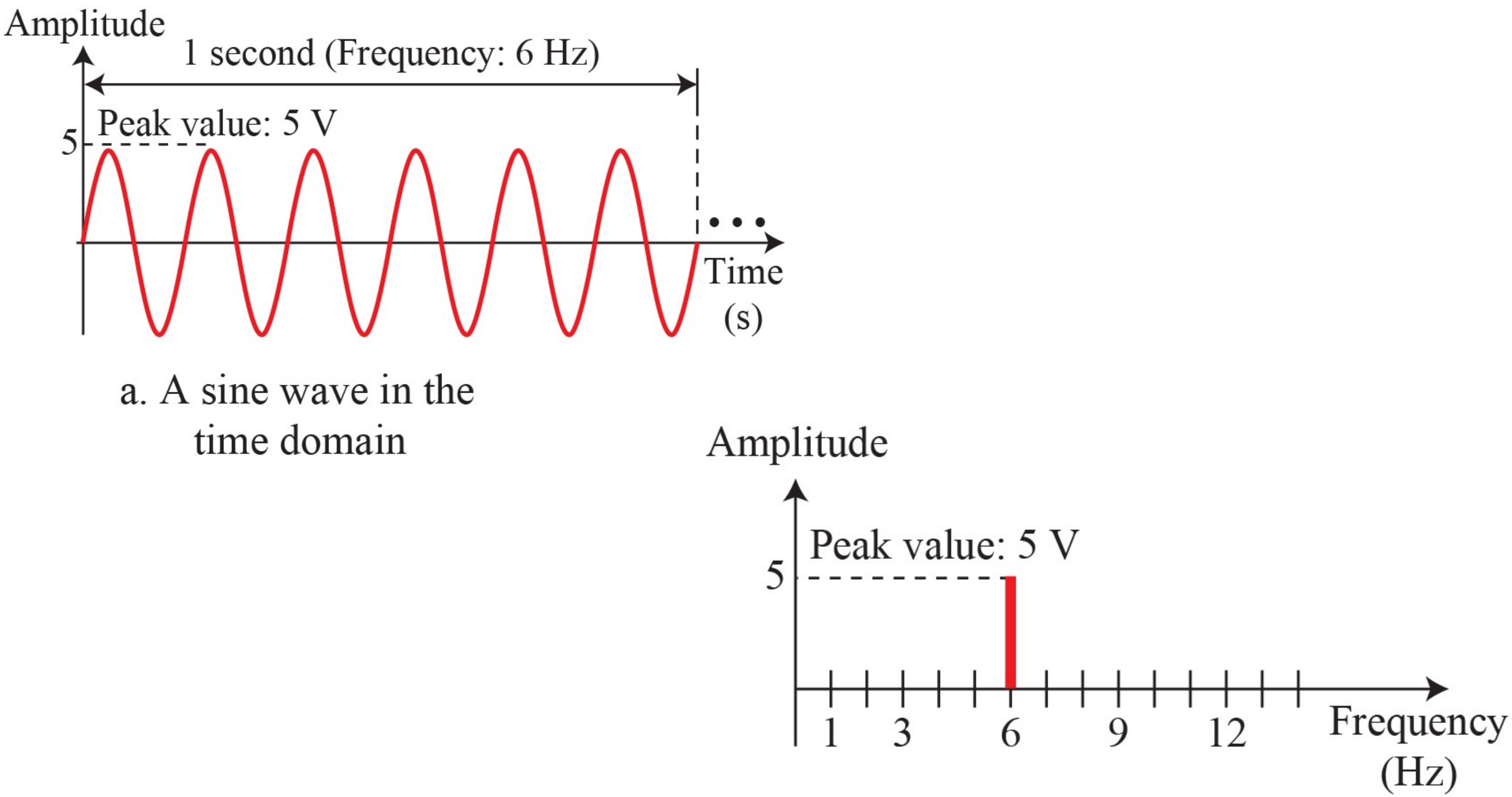
- Wavelength is another characteristic of a signal traveling through a transmission medium.

- Wavelength is the spatial period of a periodic wave; it binds the period to the propagation speed of the medium



Time and Frequency Domains

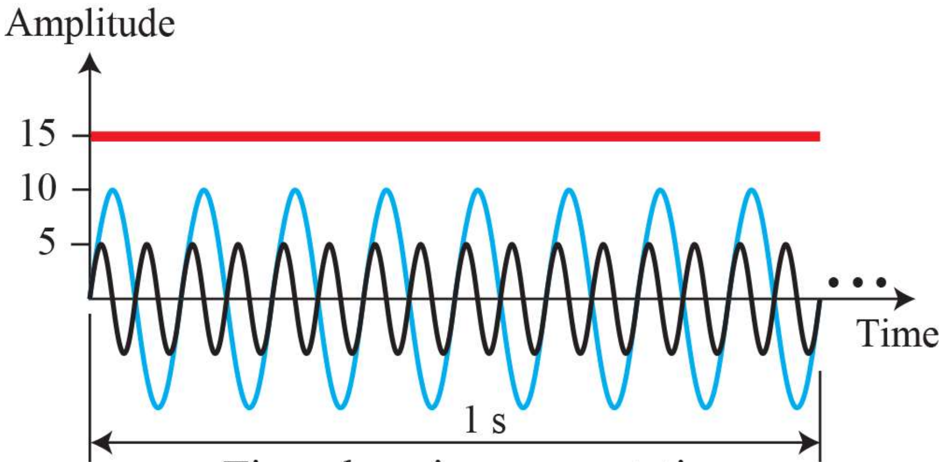
- A sine wave is comprehensively defined by its amplitude, frequency, and phase.
- We have been showing a sine wave by using what is called a time domain plot.
- The time-domain plot shows changes in signal amplitude with respect to time (it is an amplitude-versus-time plot).



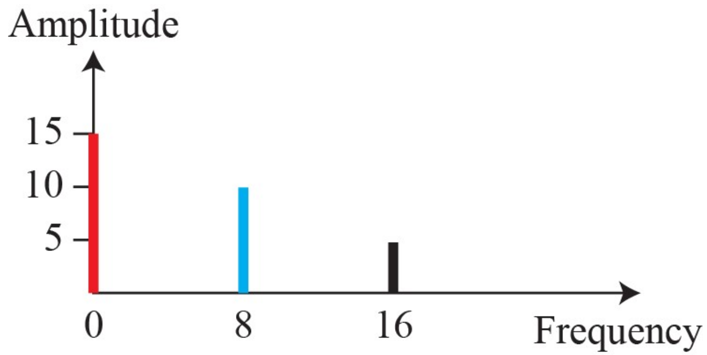
a. A sine wave in the time domain

b. The same sine wave in the frequency domain

The time-domain and frequency-domain plots of a sine wave



a. Time-domain representation

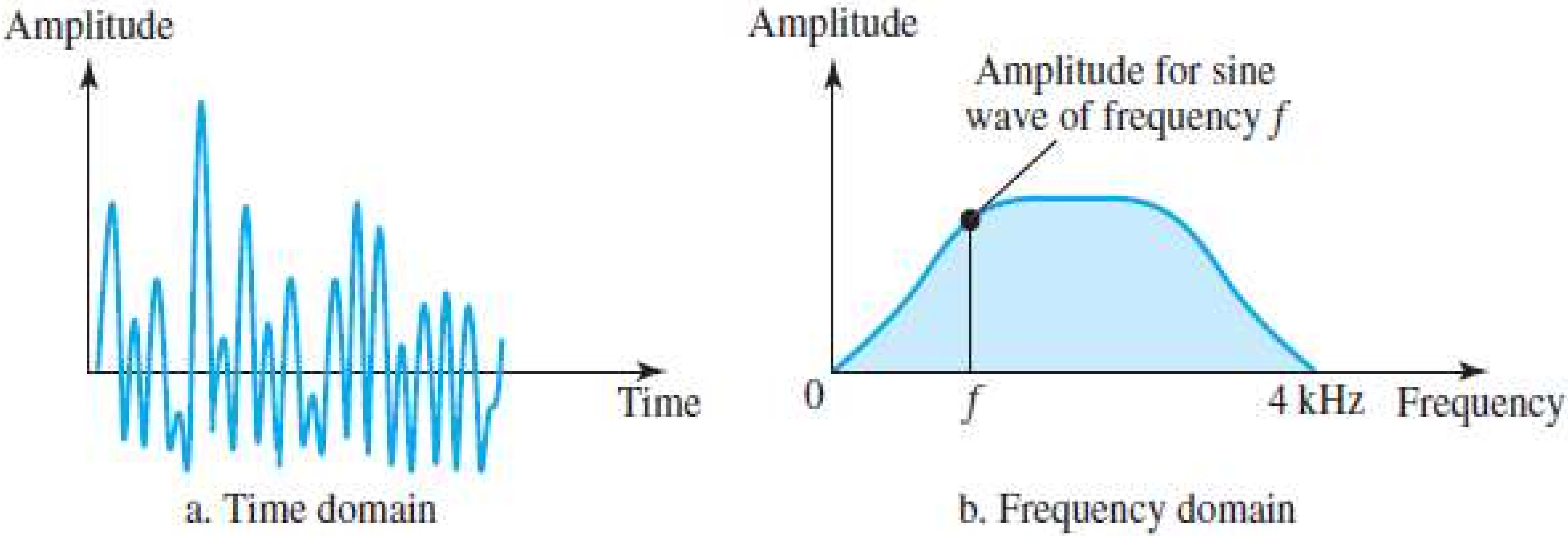


b. Frequency-domain representation

The time domain and frequency domain of three sine waves

Composite Signals

- A composite signal can be periodic or non-periodic.
- A periodic composite signal can be decomposed into a series of simple sine waves with discrete frequencies, frequencies that have integer values (1, 2, 3, and so on).
- A non-periodic composite signal can be decomposed into a combination of an infinite number of simple sine waves with continuous frequencies, frequencies that have real values.



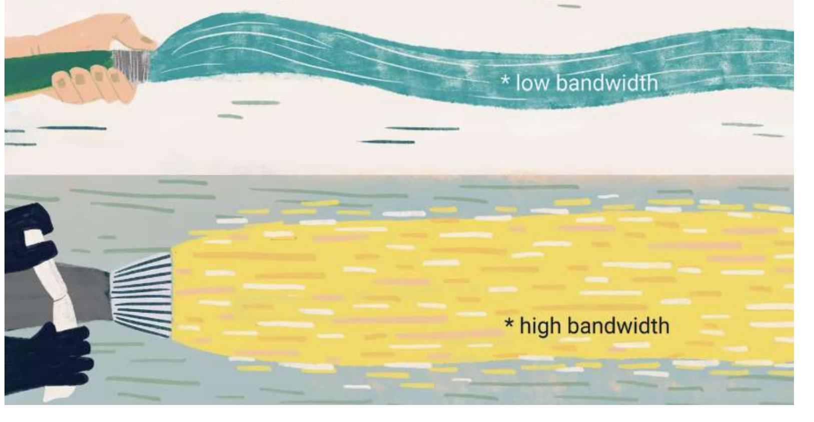
The time and frequency domains of a nonperiodic signal

Bandwidth

- The range of frequencies in a composite signal is its bandwidth.

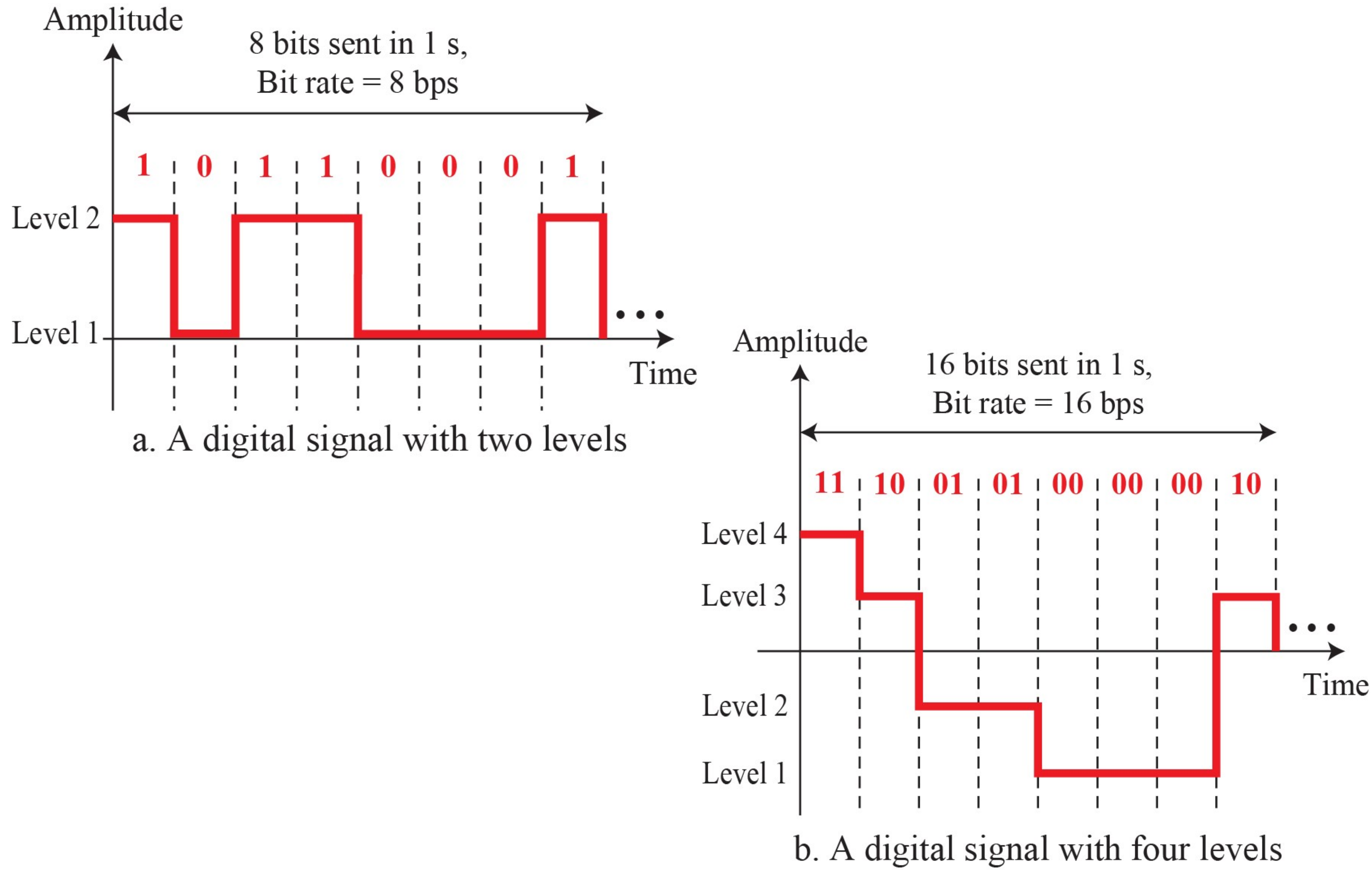
- The bandwidth is normally a difference between two numbers.

- If a composite signal contains frequencies between 1000 and 5000, its bandwidth is 5000 − 1000, or 4000.



DIGITAL SIGNALS

- In addition to being represented by an analog signal, information can also be represented by a digital signal.
- 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.



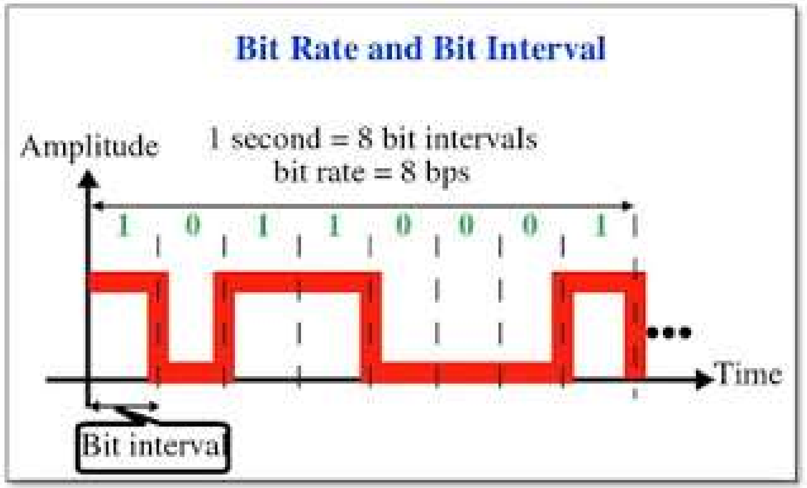
Two digital signals: one with two signal levels and the other with four signal levels

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 1s, in **bits per second (bps).**



Bit Length

- We discussed the concept of the wavelength for an analog signal: the distance one cycle occupies on the transmission medium.

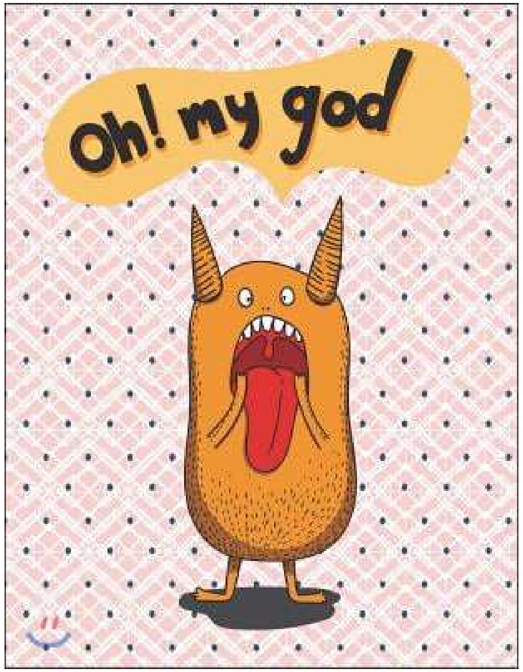
- We can define something similar for a **digital signal: the bit length**. The bit length is the distance one bit occupies on the transmission medium.

$$\text{Bit length} = \text{propagation speed} \times \text{bit duration}$$

Digital As Composite Analog

- Based on Fourier analysis,………

bla~~ bla~~ bla~~ bla~~.



TRANSMISSION IMPAIRMENT

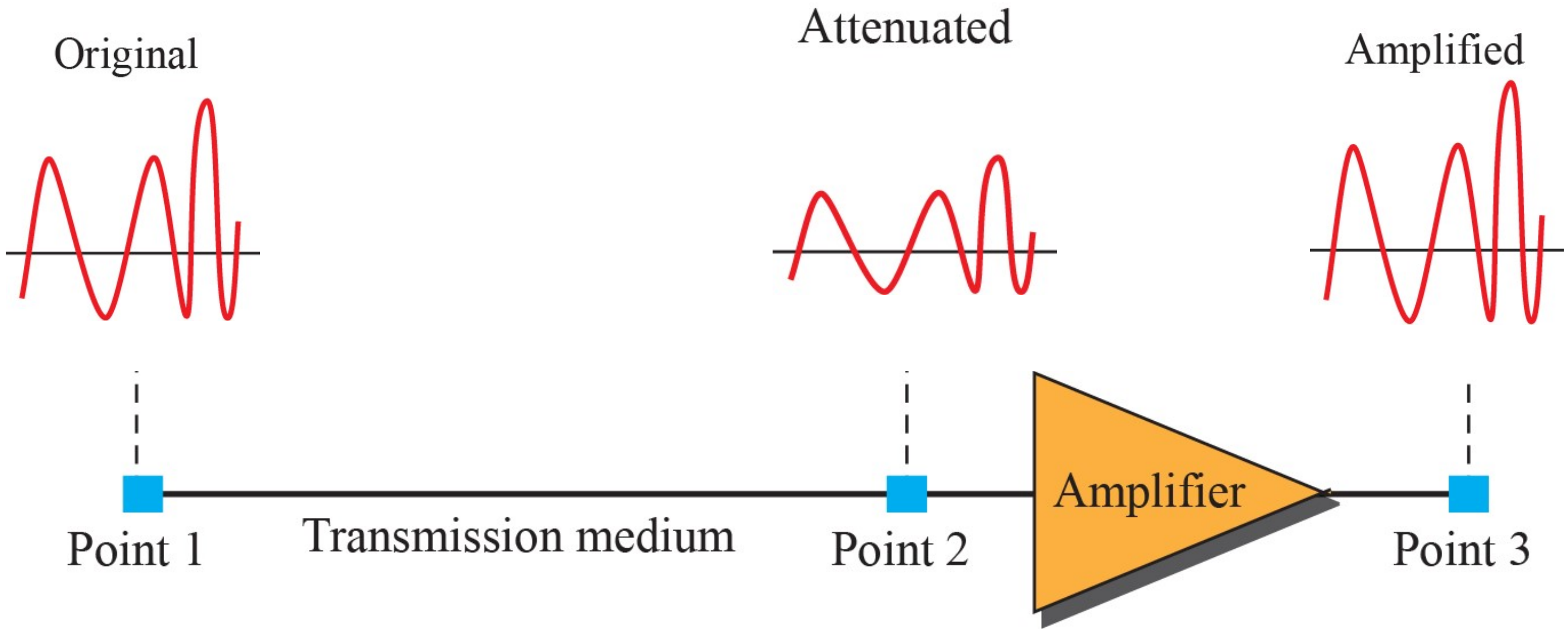
- Signals travel through transmission media, which are not perfect. The imperfection causes **signal impairment.**

- 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.**

Attenuation

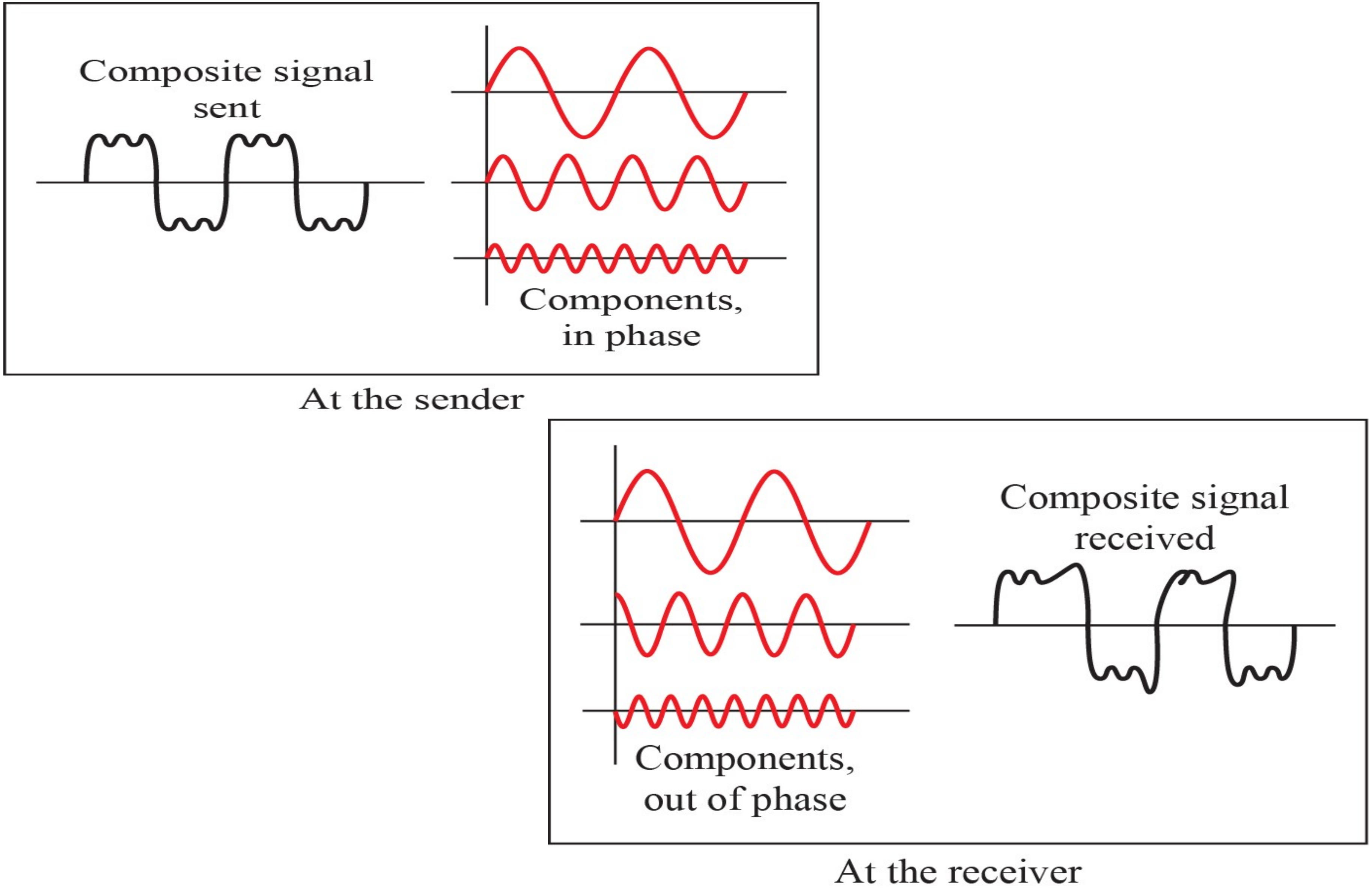
- Attenuation means a loss of energy. When a signal, simple or composite, travels through a medium, it loses some of its energy in overcoming the resistance of the medium.
- That is why a wire carrying electric signals gets warm. Some of the electrical energy in the signal is converted to heat.
- To compensate for this loss, **amplifiers** are used to amplify the signal.



Attenuation and amplification

Distortion

- Distortion means that the signal changes its form or shape. 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.


Composite signal sent
Components, in phase
At the sender
Components, out of phase
Composite signal received
At the receiver

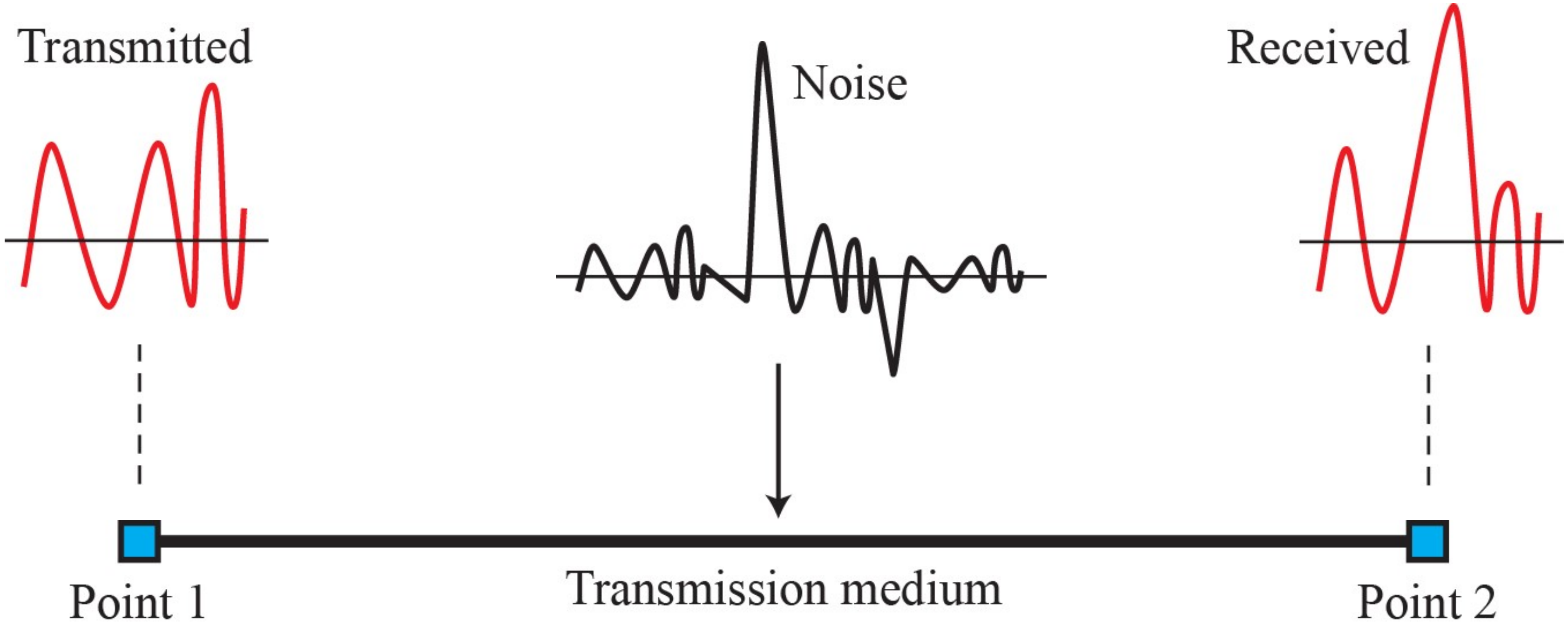
Distortion

Noise

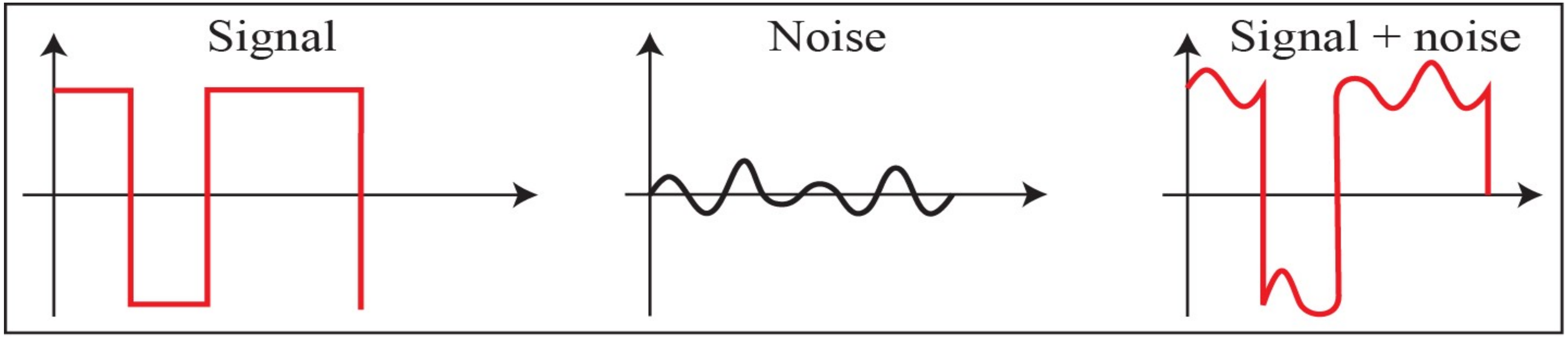
- Noise is another cause of impairment. Several types of noise, such as thermal noise, induced noise, crosstalk, and impulse noise, may corrupt the signal.

- Thermal noise is the random motion of electrons in a wire, which creates an extra signal not originally sent by the transmitter.

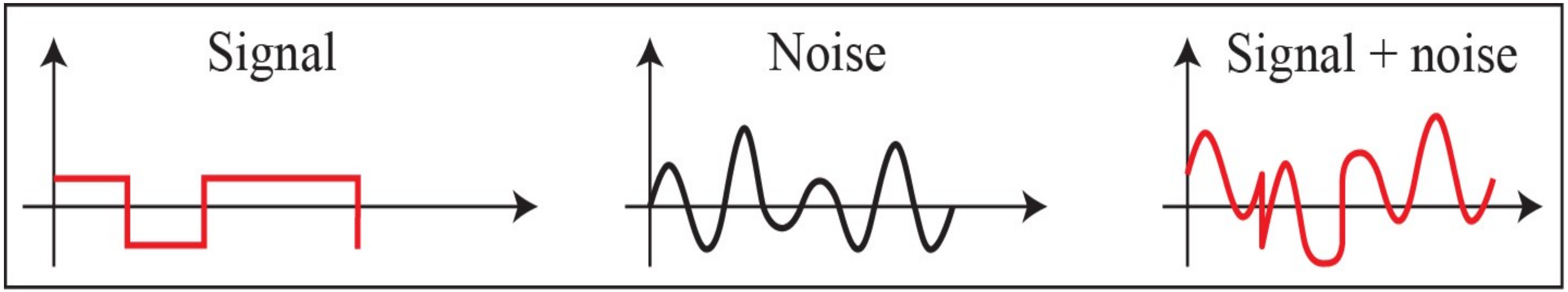
- Induced noise comes from sources such as motors. Crosstalk is the effect of one wire on the other.


Transmitted
Noise
Received
Point 1
Transmission medium
Point 2

Noise



a. High SNR



b. Low SNR

Two cases of SNR: a high SNR and a low 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.

- Two theoretical formulas were developed to calculate the data rate: one by Nyquist for a noiseless channel, another by Shannon for a noisy channel.

- For a noiseless channel, the Nyquist bit rate formula defines the theoretical maximum bit rate.

$$\text{BitRate} = 2 \times \text{bandwidth} \times \log_2 L$$

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:

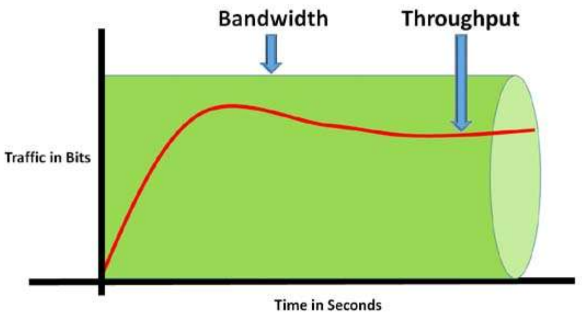
$$\text{Capacity} = \text{bandwidth} \times \log_2(1 + \text{SNR})$$

PERFORMANCE

- One important issue in networking is the performance of the network - how good is it?
- One characteristic that measures network performance is bandwidth.
- However, the term can be used in two different contexts with two different measuring values: bandwidth in hertz and bandwidth in bits per second.

Throughput

- The throughput is a measure of how fast we can actually send data through a network.
- Although, bandwidth in bits per second and throughput seem the same, they are different.
- A link may have a bandwidth of B bps, but we can only send T bps through this link with T always less than B.



Latency

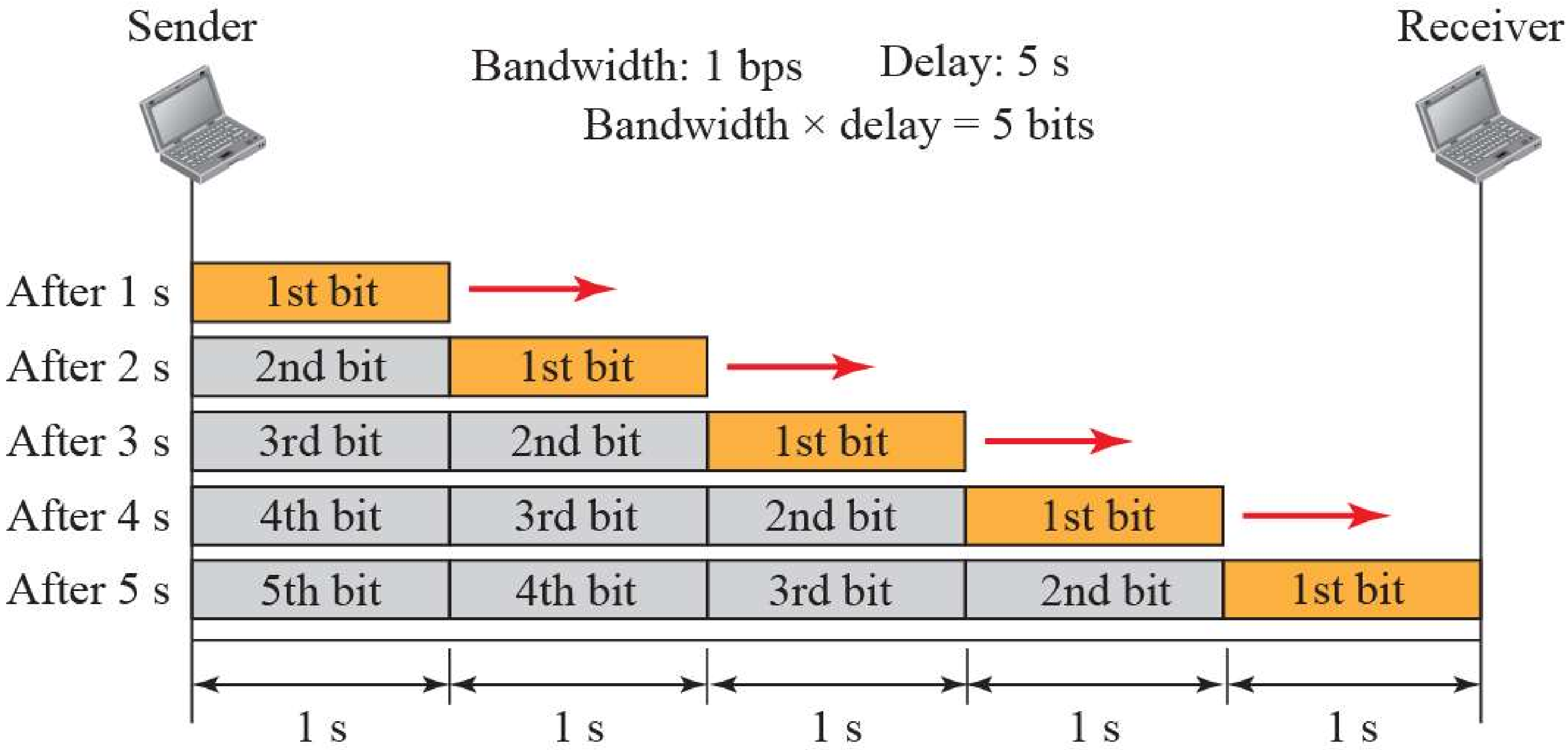
- The latency or delay defines how long it takes for an entire message to completely arrive at the destination from the time the first bit is sent out from the source.

- We can say that latency is made of four components: propagation time, transmission time, queuing time and processing delay.

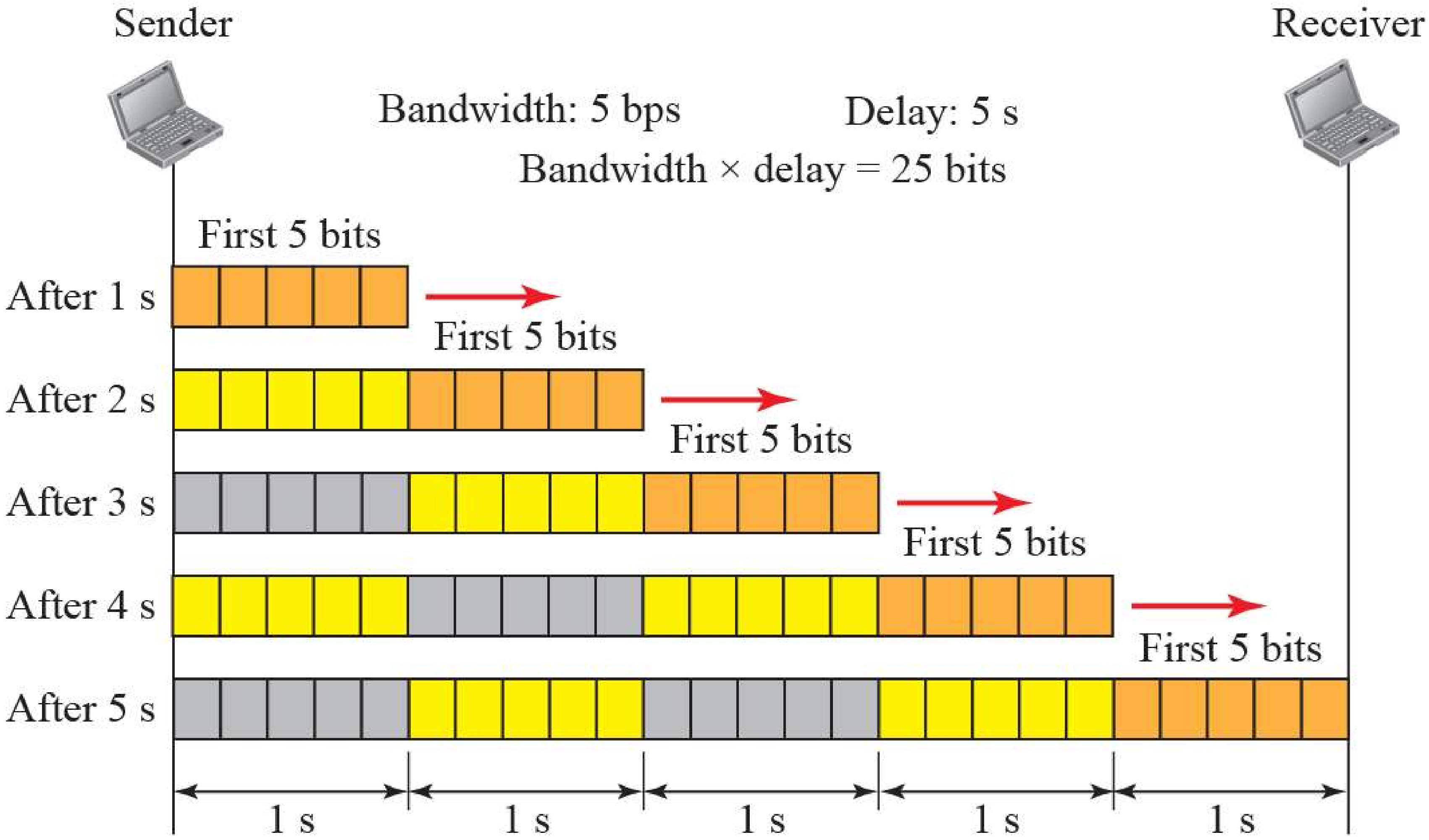
Latency = propagation time + transmission time + queuing time + processing delay

Bandwidth-Delay Product

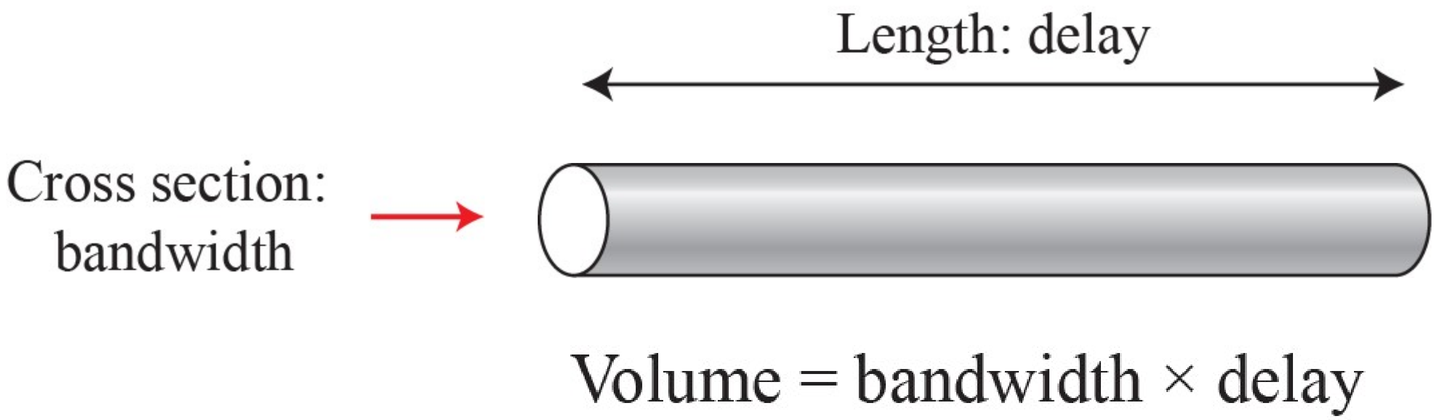
- Bandwidth and delay are two performance metrics of a link.
- What is very important in data communications is the product of the two, the bandwidth-delay product.



Filling the links with bits for Case 1



Filling the pipe with bits for Case 2



Concept of bandwidth-delay product

Jitter

- Another performance issue that is related to delay is **jitter**.
- We can roughly say that jitter is a problem if different packets of data encounter different delays and the application using the data at the receiver site is time-sensitive (audio and video data, for example).

Jitter

- If the delay for the first packet is 20 ms, for the second is 45 ms, and for the third is 40 ms, then the real-time application that uses the packets endures jitter.