# 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.

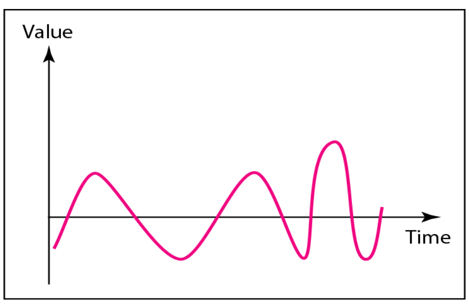
Analog data, such as the sounds made by a human voice, take on continuous values. When someone speaks, an analog wave is created in the air. This can be captured by a microphone and converted to an analog signal or sampled to a digital signal.

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.

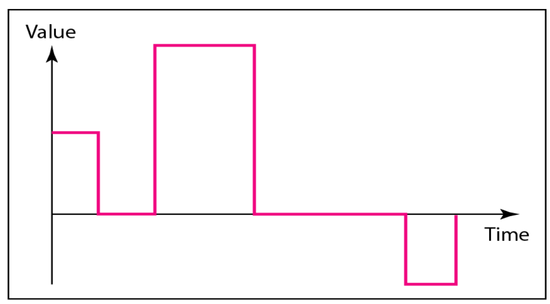
### Analog and Digital Signals

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**, 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.

The simplest way to show signals is by plotting them on a pair of perpendicular axes. The vertical axis represents the value or strength of a signal. The horizontal axis represents time.



a. Analog signal

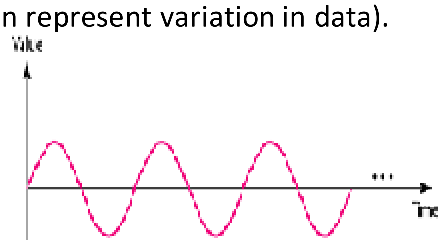


b. Digital signal

### Periodic and Non-periodic Signals

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 ***non-periodic signal*** changes without exhibiting a pattern or cycle that repeats over time.

Both analog and digital signals can be periodic or non-periodic. In data communications, we use periodic analog signals (because they need less bandwidth) and non-periodic digital signals (because they can represent variation in data).



Periodoic

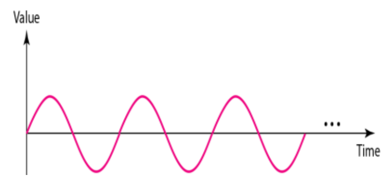
Non Periodic

## Periodic Analog Signals

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

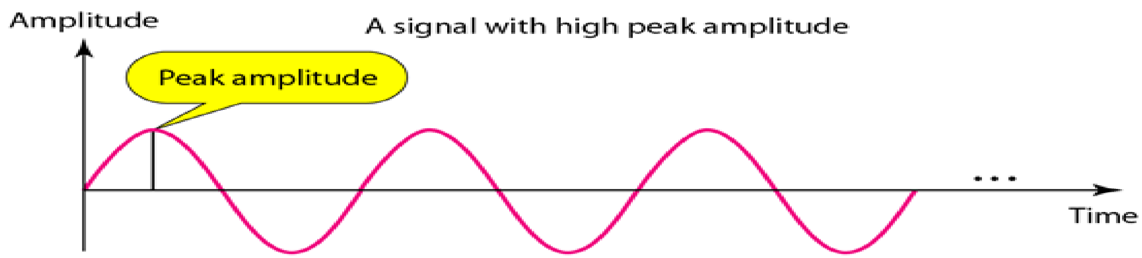
The sine wave is the most fundamental form of a periodic analog signal. When we visualize it as a simple oscillating curve, its change over the course of a cycle is smooth and consistent, a continuous, rolling flow.




A sine wave can be represented by 3 parameters: ***peak amplitude, frequency*** **and** ***phase****.*

***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.*

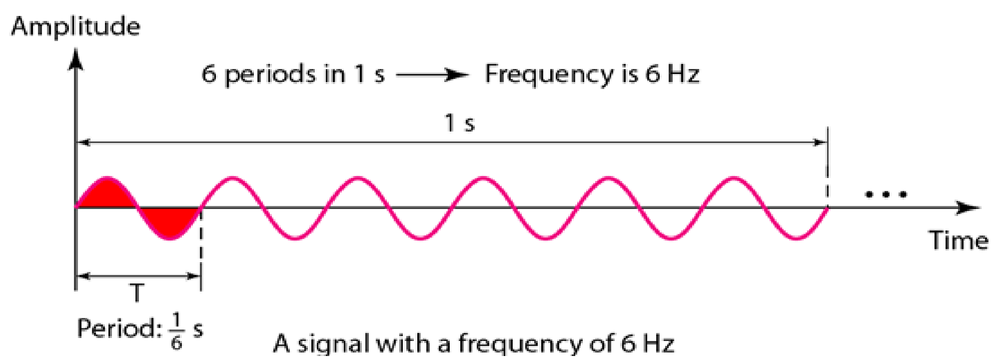



***Period and Frequency***

Period refers to the amount of time, in seconds, a signal needs to complete 1 cycle. Frequency refers to the number of periods in 1s. Period is the inverse of frequency, and frequency is the inverse of period, as the following formulas show.

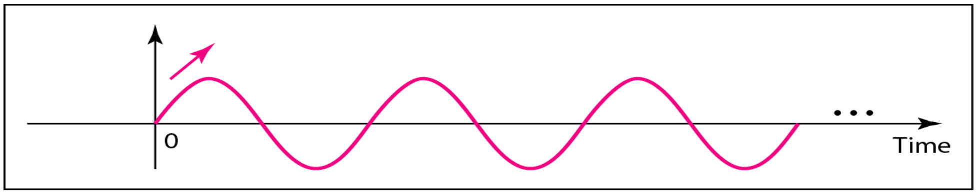
**f = 1/ T** and ***T*=1 / *f***

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.

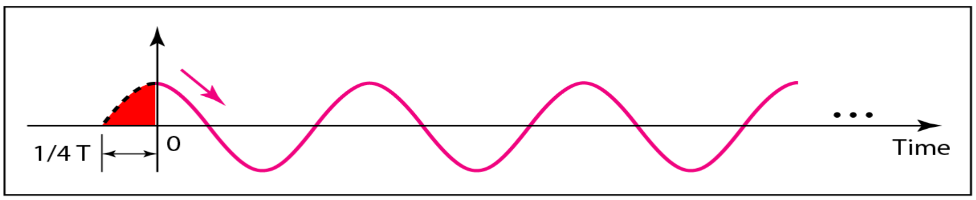


### Phase

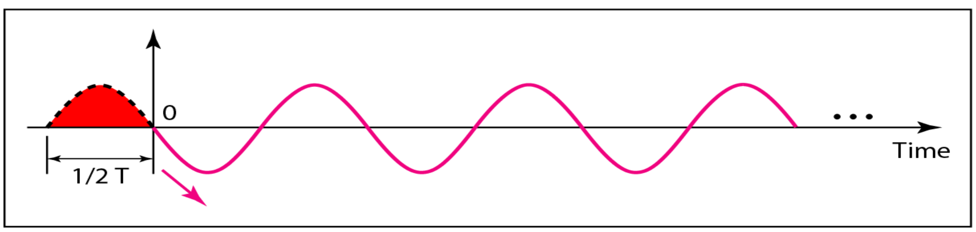
The term phase describes the position of the waveform relative to time 0.
Phase is measured in degrees or radians [360° is 2π rad; 1° is 2 π /360 rad, and 1 rad is 360/(2 π)].
A phase shift of 360° corresponds to a shift of a complete period; a phase shift of 180° corresponds to a shift of one-half of a period; and a phase shift of 90° corresponds to a shift of one-quarter of a period.




a. 0 degrees




b. 90 degrees




c. 180 degrees

### Wavelength

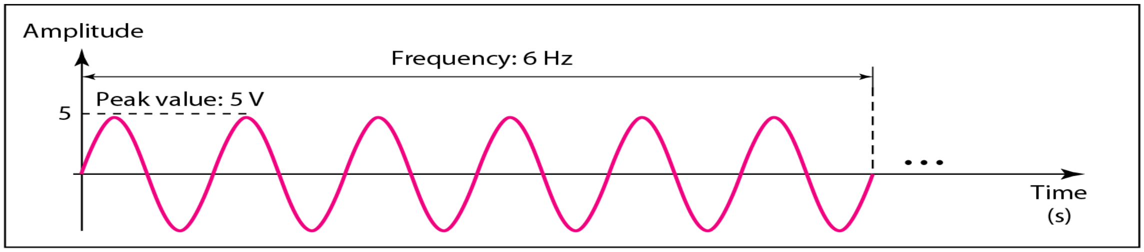
**Wavelength is the distance a simple signal can travel in one period.** The frequency of a signal is independent of the medium, the wavelength depends on both the frequency and the medium.
Wavelength can be calculated as follows:

**Wavelength = propagation speed x period =** $\frac{\textit{propagation speed}}{\textit{frequency}}$

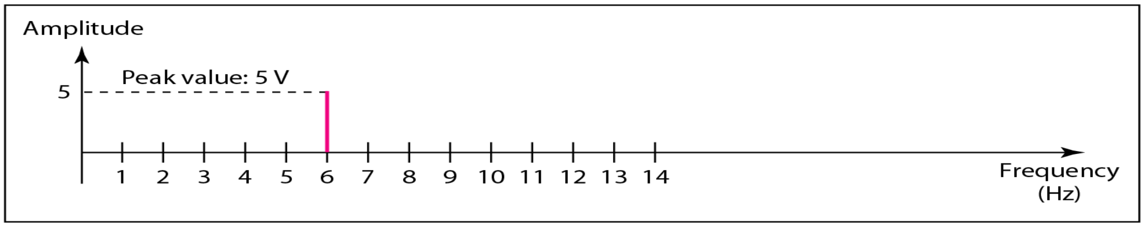
The propagation speed of electromagnetic signals depends on the medium and on the frequency of the signal. For example, in a vacuum, light is propagated with a speed of 3 x $10^8$ m/s.

### Time and Frequency Domains

The time-domain plot shows changes in signal amplitude with respect to time (it is an amplitude-versus-time plot). A frequency-domain plot is concerned with only the peak value and the frequency.
The advantage of the frequency domain is that we can immediately see the values of the frequency and peak amplitude. A complete sine wave is represented by one spike. The position of the spike shows the frequency; its height shows the peak amplitude.



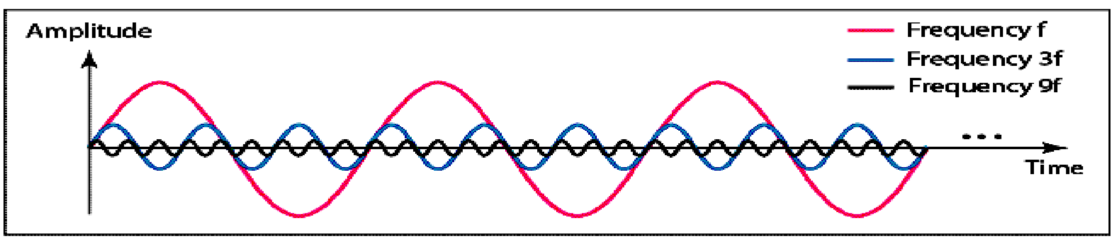

a. A sine wave in the time domain (peak value: 5 V, frequency: 6 Hz)



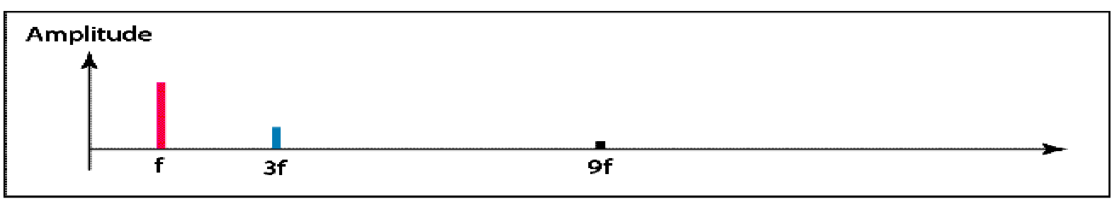

b. The same sine wave in the frequency domain (peak value: 5 V, frequency: 6 Hz)

### Composite Signal

Composite signal is actually a combination of simple sine wave with different frequencies, amplitude and phases. It can be a periodic or non-periodic. A periodic composite signal can be decomposed into a series of simple sine waves with discrete frequencies. A non-periodic composite signal can be decomposed into a combination of an infinite number of simple sine waves with continuous frequencies.




a. Time-domain decomposition of a composite signal




b. Frequency-domain decomposition of the composite signal

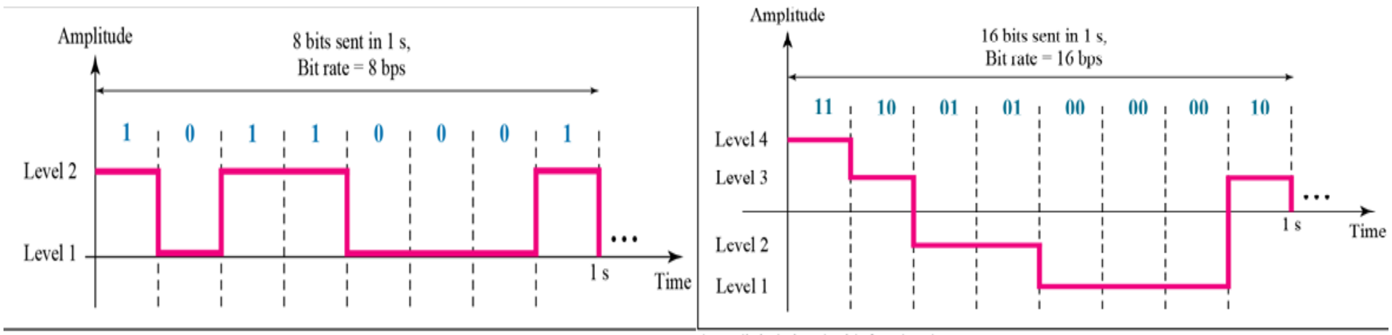
### Bandwidth

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

# DIGITAL SIGNALS

**Bit Rate**

The bit rate is the number of bits sent in 1s, expressed in bits per second (bps).




a. A digital signal with two levels

b. A digital signal with four levels

**Bit Length**

The bit length is the distance one bit occupies on the transmission medium.

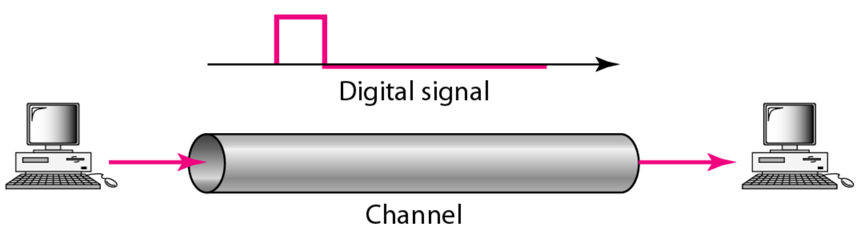
$$Bit\ length\ = propagation\ speed\ x\ bit\ duration$$

## Transmission of Digital Signals

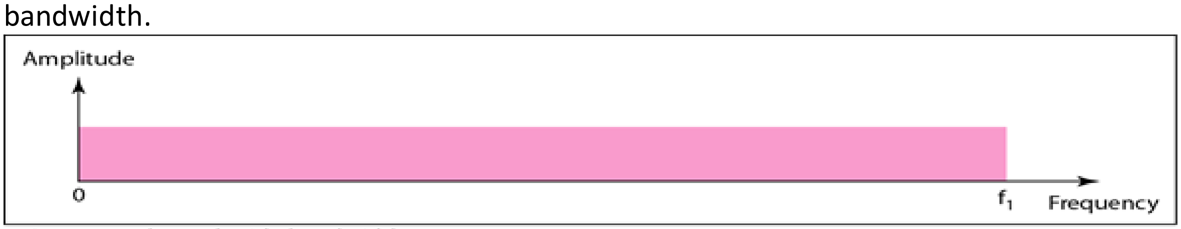
We can transmit a digital signal by using one of two different approaches: **baseband transmission** or **broadband transmission** (using modulation).

### *Baseband Transmission*

Baseband transmission means sending a digital signal over a channel without changing the digital signal to an analog signal. Baseband transmission requires that we have a low-pass channel, a channel with a bandwidth that starts from zero. This is the case if we have a dedicated medium with a bandwidth constituting only one channel. For example, the entire bandwidth of a cable connecting two computers is one single channel.




There are two low-pass channels: one with a narrow bandwidth and the other with a wide bandwidth.



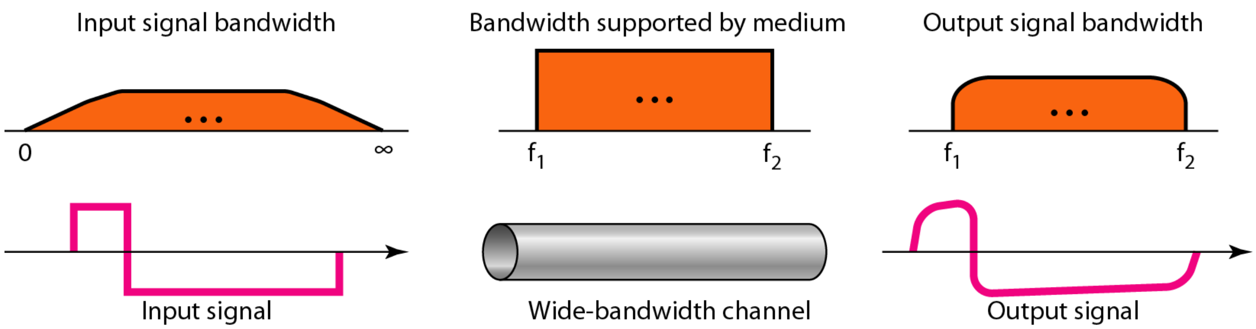

a. Low-pass channel, wide bandwidth


b. Low-pass channel, narrow bandwidth

Let us study two cases of a baseband communication: a low-pass channel with a wide bandwidth and one with a limited bandwidth.

***Case 1: Low-Pass Channel with Wide Bandwidth***

If we want to preserve the exact form of a non-periodic digital signal, we need to send the entire spectrum, the continuous range of frequencies between zero and infinity. This is possible if we have a dedicated medium with an infinite bandwidth between the sender and receiver that preserves the exact amplitude of the composite signal. The amplitudes of the frequencies at the border of the bandwidth are so small that they can be ignored. This means that if we have a medium, such as a coaxial cable or fiber optic, with a very wide bandwidth, two stations can communicate by using digital signals with very good accuracy.

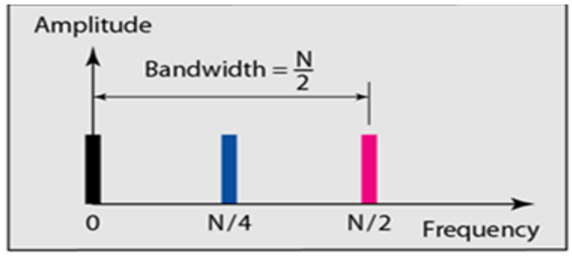


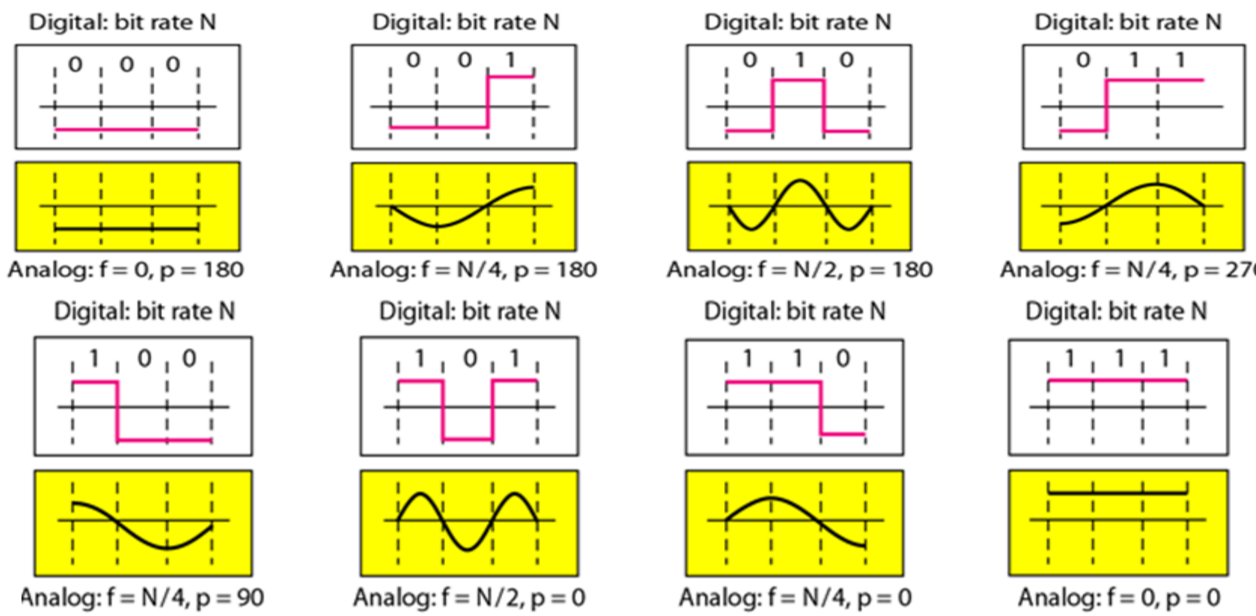

In the fig above, although the output signal is not an exact replica of the original signal, the data can still be deduced from the received signal.

***Case 2: Low-Pass Channel with Limited Bandwidth***

In a low-pass channel with limited bandwidth, we approximate the digital signal with an analog signal. The level of approximation depends on the bandwidth available.

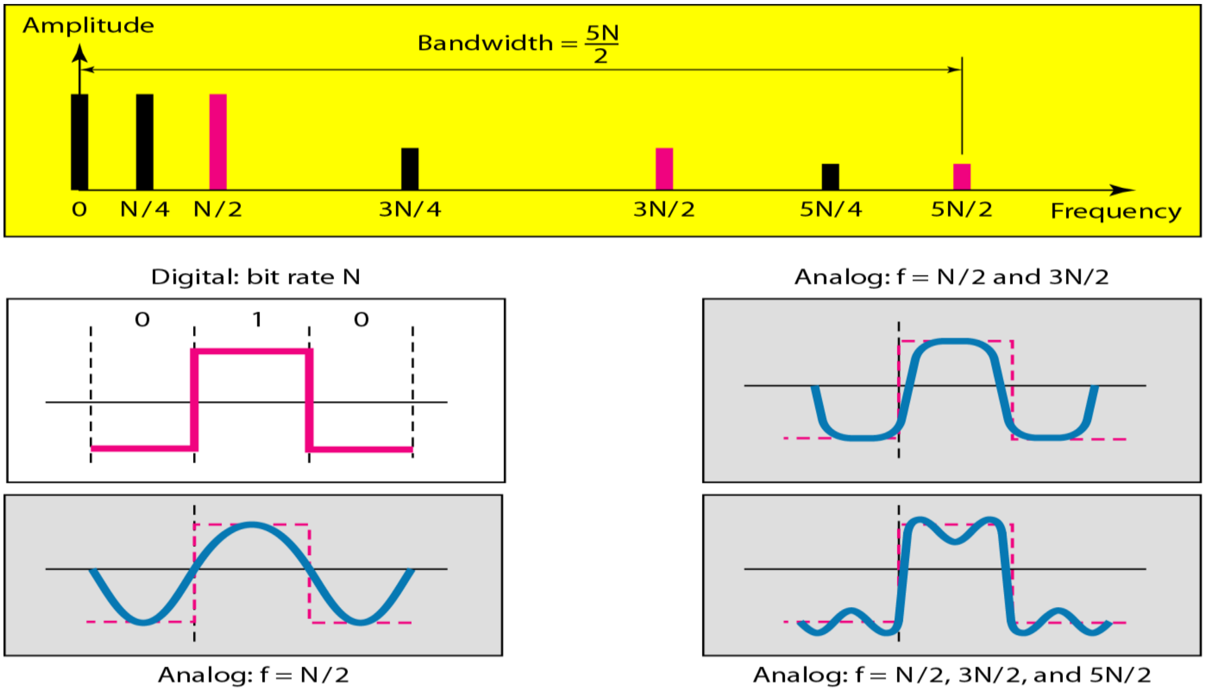
***i. Rough Approximation:*** Let us assume that we have a digital signal of bit rate *N.* If we want to send analog signals to roughly simulate this signal, we need to consider the worst case, a maximum number of changes in the digital signal. This happens when the signal carries the sequence 01010101 ... or 10101010... Let 1 be the positive peak value and 0 be the negative peak value. We send 2 bits in each cycle; the frequency of the analog signal is *N/2 (maximum).* As an example of this concept, let us see how a digital signal with a 3-bit pattern can be simulated by using analog signals. Figure below shows the idea. The two similar cases (000 and 111) are simulated with a signal with frequency $f$ =0 and a phase of 180° for 000 and a phase of 0° for 111. The two worst cases (010 and 101) are simulated with an analog signal with frequency f= N/2 and phases of 180° and 0°. The other four cases can only be simulated with an analog signal with f = N/4 and phases of 180°, 270°, 90°, and 0°. In other words, we need a channel that can handle frequencies 0, N/4, and N/2. This rough approximation is referred to as using the first harmonic (N/2) frequency.



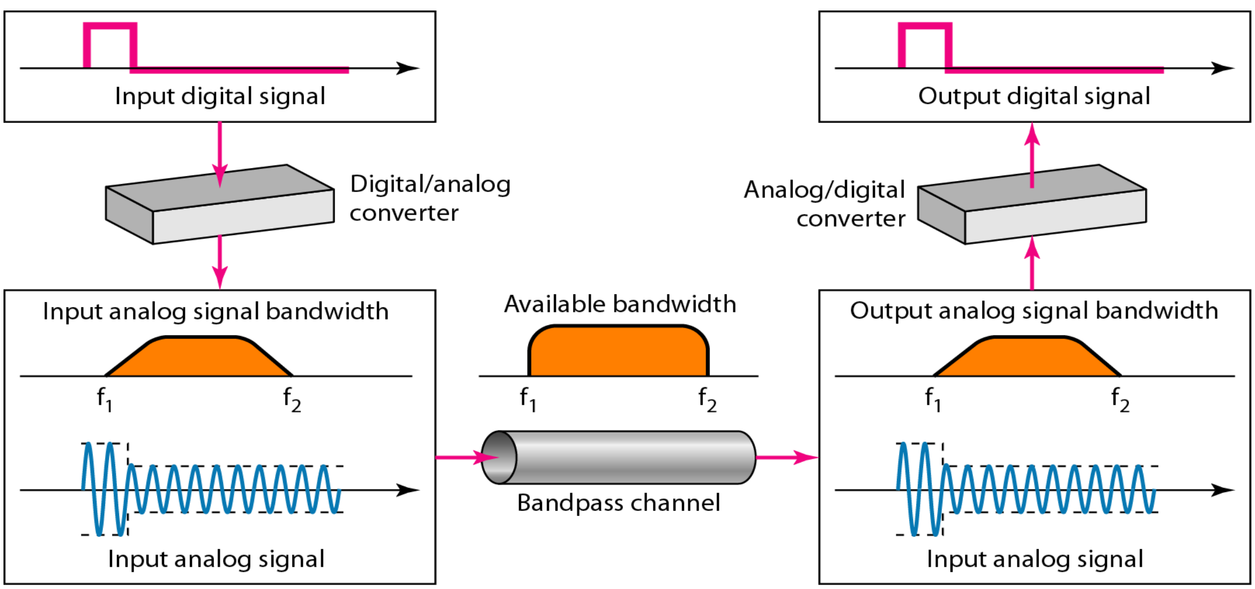

***ii. Better Approximation:*** To make the shape of the analog signal look more like that of a digital signal, we need to add more harmonics of the frequencies. We need to increase the bandwidth. We can increase the bandwidth to 3N/2, 5N/2, 7N/2, and so on. We use the first, third, and fifth harmonics. The required bandwidth is now 5N/2, the difference between the lowest frequency 0 and the highest frequency 5N/2. In baseband transmission, the required bandwidth is proportional to the bit rate; if we need to send bits faster, we need more bandwidth. Table below shows how much bandwidth we need to send data at different rates.

| *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 |




***Broadband Transmission (Using Modulation)***

Broadband transmission or modulation means changing the digital signal to an analog signal for transmission. Modulation allows us to use a bandpass channel - a channel with a bandwidth that does not start from zero. In broadband transmission, a digital signal is converted to a composite analog signal. A single-frequency analog signal (called a carrier) is used; the amplitude of the carrier has been changed to look like the digital signal. At the receiver, the received analog signal is converted to digital, and the result is a replica of what has been sent.

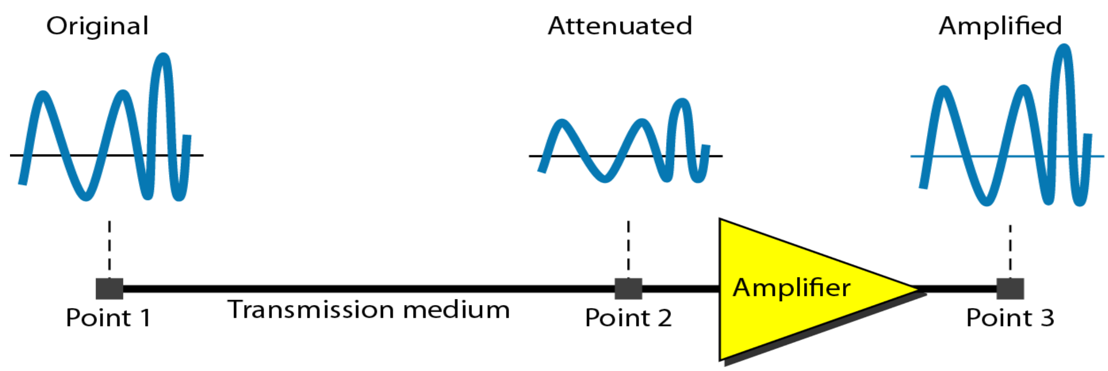


# TRANSMISSION IMPAIRMENT

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. 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, 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.




*Decibel*

To show that a signal has lost or gained strength, engineers use the unit of the decibel. The decibel (dB) measures the relative strengths of two signals or one signal at two different points. Note that the decibel is negative if a signal is attenuated and positive if a signal is amplified.

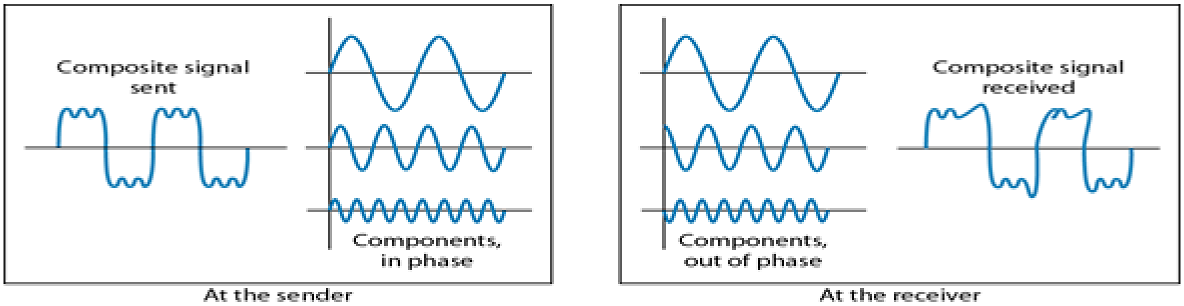
$$dB = 10 \log_{10} \frac{P_2}{P_1}$$

Variables *P1* and *P2* are the powers of a signal at points 1 and 2, respectively. We know that power is proportional to the square of the voltage, the formula in terms of voltage is

$$dB = 20 \log_{10} \frac{V_2}{V_1}$$

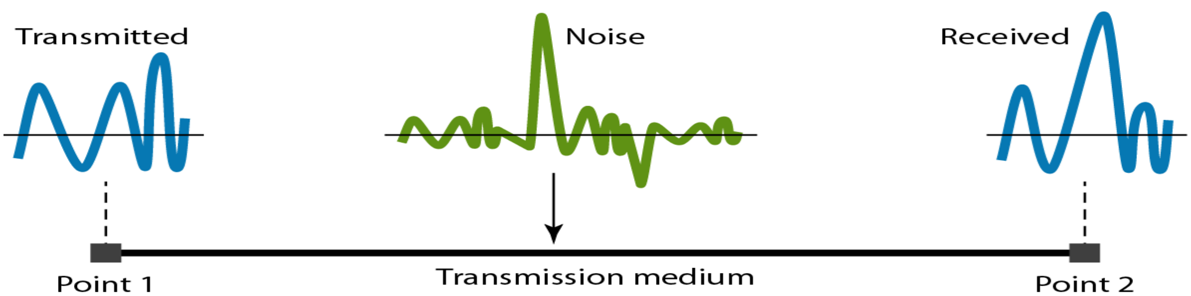
### *Distortion*

**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. 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.




### *Noise*

**Noise** is unwanted signal that corrupt the original signal. Types of noise are thermal noise, induced noise, crosstalk, and impulse noise. **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.



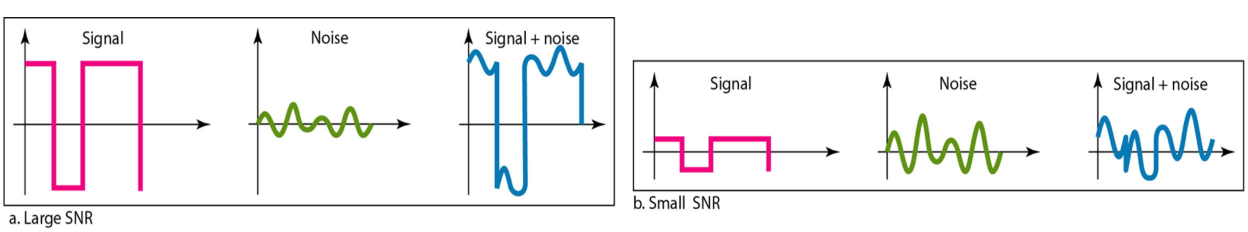

### *Signal-to-Noise Ratio (SNR)*

The signal-to-noise ratio is defined as

$$SNR = \frac{\text{average signal power}}{\text{average noise power}}$$

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.

SNR in decibels, is defined as $\mathbf{SNR_{db} = 10\log_{10} SNR}$



# DATA RATE LIMITS

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)

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

## Noiseless Channel: Nyquist Bit Rate

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

$$\boldsymbol{BitRate} = \mathbf{2} \times \boldsymbol{bandwidth\, x} \ \mathbf{log_2}\, \boldsymbol{L}$$

In this formula, *bandwidth* is the bandwidth of the channel, *L* is the number of signal levels used to represent data, and *BitRate* is the bit rate in bits per second.
According to the formula, given a specific bandwidth, we can have any bit rate we want by increasing the number of signal levels. Although the idea is theoretically correct, practically there is a limit. When we increase the number of signal levels, we impose a burden on the receiver. If the number of levels in a signal is just 2, the receiver can easily distinguish between a 0 and a 1. If the level of a signal is 64, the receiver must be very sophisticated to distinguish between 64 different levels. In other words, increasing the levels of a signal reduces the reliability of the system.

## Noisy Channel: Shannon Capacity

In reality, we cannot have a noiseless channel; the channel is always noisy. Shannon capacity, to determine the theoretical highest data rate for a noisy channel:

$$\boldsymbol{Capacity} = \boldsymbol{bandwidth} \times \boldsymbol{log}\,\mathbf{log_2}(\mathbf{1} + \boldsymbol{SNR})$$

In this formula, *bandwidth* is the bandwidth of the channel, SNR is the signal to noise ratio, and *capacity* is the capacity of the channel in bits per second. Note that in this formula there is no indication of the signal level, which means that no matter how many levels we have, we cannot achieve a data rate higher than the capacity of the channel.

# 3.6 PERFORMANCE

## Bandwidth

One characteristic that measures network performance is bandwidth. The term can be used in two different contexts with two different measuring values: bandwidth in hertz and bandwidth in bits per second.

### *Bandwidth in Hertz*

Bandwidth in hertz is the range of frequencies contained in a composite signal or the range of frequencies a channel can pass. For example, we can say the bandwidth of a subscriber telephone line is 4 kHz.

***Bandwidth in Bits per Seconds***

The term *bandwidth* can also refer to the number of bits per second that a channel, a link, or even a network can transmit. For example, one can say the bandwidth of a Fast Ethernet network is a maximum of 100 Mbps. This means that this network can send 100 Mbps.

***Relationship***

There is an explicit relationship between the bandwidth in hertz and bandwidth in bits per seconds. Basically, an increase in bandwidth in hertz means an increase in bandwidth in bits per second. The relationship depends on whether we have baseband transmission or transmission with modulation.

## Throughput

The throughput is a measure of how fast we can actually send data through a network. Although, at first glance, 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.* In other words, the bandwidth is a potential measurement of a link; the throughput is an actual measurement of how fast we can send data. For example, we may have a link with a bandwidth of 1 Mbps, but the devices connected to the end of the link may handle only 200 kbps. This means that we cannot send more than 200 kbps through this link.

## Latency (Delay)

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.

$$\boldsymbol{Latency = propagation\ time + transmission\ time + queuing\ time + processing\ delay}$$

***Propagation Time***

Propagation time measures the time required for a bit to travel from the source to the destination. The propagation time is calculated by dividing the distance by the propagation speed.

$$\boldsymbol{Propagation\ time = \frac{Distance}{Propagation\ speed}}$$

***Transmission Time***

In data communications we don't send just 1 bit, we send a message. The first bit may take a time equal to the propagation time to reach its destination; the last bit also may take the same amount of time. However, there is a time between the **first bit leaving the sender and the last bit arriving at the receiver**. The time required for transmission of a message depends on the size of the message and the bandwidth of the channel.

$$\boldsymbol{Transmission\ time = \frac{Message\ size}{Bandwidth}}$$

***Queuing Time***

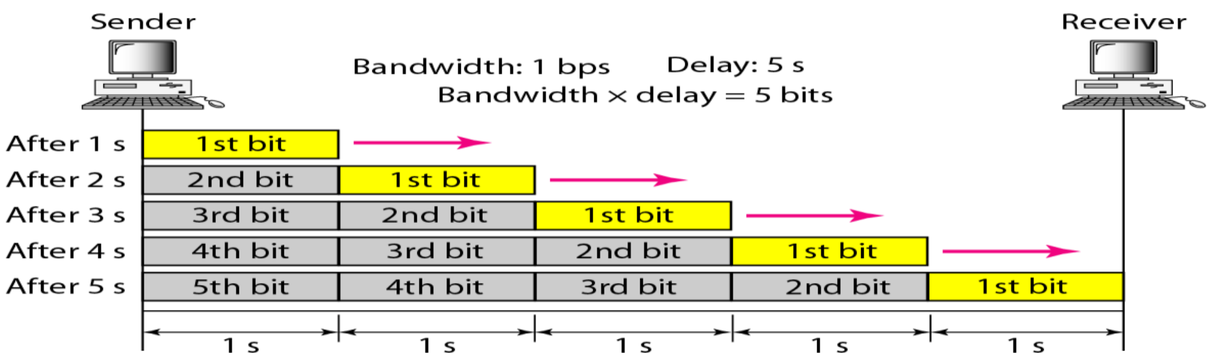
The third component in latency is the queuing time, the time needed for each intermediate or end device to hold the message before it can be processed. The queuing time is not a fixed factor; it changes with the load imposed on the network. When there is heavy traffic on the network, the queuing time increases. An intermediate device, such as a router, queues the arrived messages and processes them one by one.

***Processing Delay***

Time required for processing the message at the receiver side.

**Bandwidth-Delay Product**

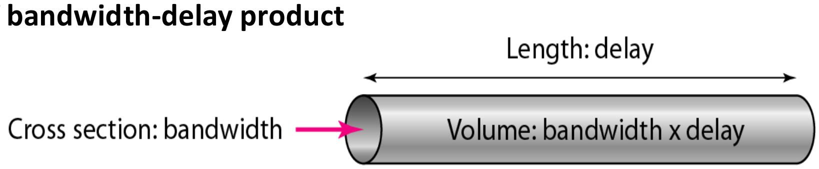
Bandwidth and delay are two performance metrics of a link. However, what is very important in data communications is the product of the two, the bandwidth-delay product.




Let us assume that we have a link with a bandwidth of 1 bps. We also assume that the delay of the link is 5s. We want to see what the bandwidth-delay product means in this case. Looking at figure, we can say that this product 1 x 5 is the maximum number of bits that can fill the link. There can be no more than 5 bits at any time on the link.

The above case show that the product of bandwidth and delay is the number of bits that can fill the link. This measurement is important if we need to send data in bursts and wait for the acknowledgment of each burst before sending the next one. To use the maximum capability of the link, we need to make the size of our burst 2 times the product of bandwidth and delay; we need to fill up the full-duplex channel (two directions). The sender should send a burst of data of (2 x bandwidth x delay) bits. The sender then waits for receiver acknowledgment for part of the burst before sending another burst.

**Concept of bandwidth-delay product**




*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 above.*

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