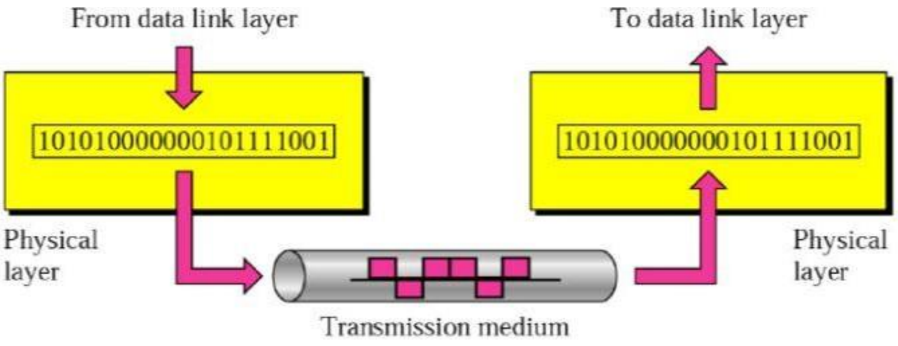
Chapter 2

Physical Layer

1. Physical Layer Services

The physical layer is the lowest layer in the OSI network model. It is mainly concerned with the transmission of raw bits over a communication channel. It uses the transmission media to move bits from one node to another in the network.

The physical layer provides services to the data link layer as well as takes care of the underlying physical communication medium.



Specific responsibilities of physical layer include:

- Bit-to-signal transformation (Signal Representation).
- Bit-rate control.
- Bit-synchronization.
- Multiplexing.
- Switching.

1.1. Bit-to-Signal Transformation (Signal Representation)

- The physical layer sends a stream of bits from one node to another using a transmission medium.
- For transmission, the data bits have to be represented by an appropriate signal, which can propagate through the medium.
- For example, if the medium is a copper wire, electromagnetic signals are used, and if it is a fiber optic cable, light signals are used.

1.2. Bit-Rate Control (Timing)

- The physical layer has to determine the rate of transmission.
- It has to take into consideration the upper limit of the transmission medium and control the rate of data flow.

1.3. Bit-Synchronization

- For the data to be correctly received and interpreted, bit timing is very crucial.
- Various encoding methods are used to synchronize the receiver to the sender.
- This layer also provides clocking mechanisms that control the sender and receiver.

1.4. Multiplexing

- This is the process of sharing a single physical medium among multiple transmissions.
- The physical layer uses different techniques like Frequency and Time Division Multiplexing for sharing the medium.

1.5. Switching

- Switching creates a temporary connection between two or more devices in the network.

2. Data and signals

Computer networks are used for sharing data. To be transmitted, data must be transformed into a signal that can be sent across the medium.

2.1. Definitions

- **Signal:** A signal is an electrical or electromagnetic current that is used for carrying data from one device or network to another. It is the key component behind data communication and networking. A signal can be either analog or digital.

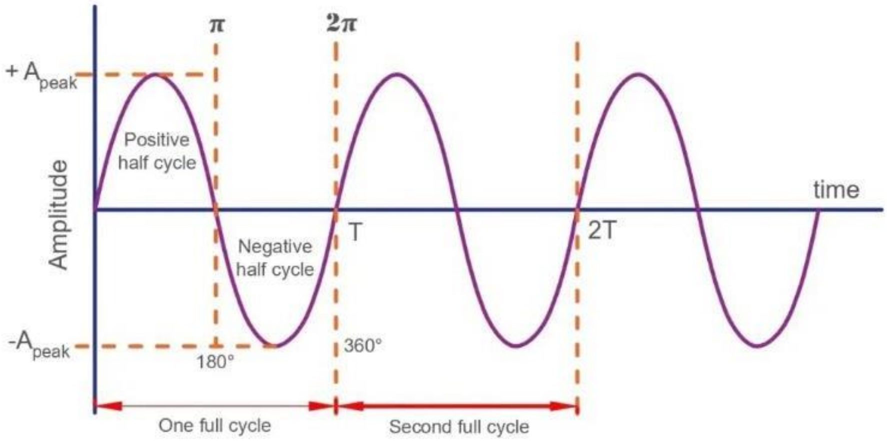
  The data as well as signals which represent the data can be of two types:
- **Analog data:** This refers to information that is continuous in nature. For example, human voice, data collected by sensors (temperature, pressure) are having continuous varying pattern. These are analog in nature.
- **Digital data:** This refers to information that is discrete in nature. It has well defined states. For example, computer data has only two discrete values 0 and 1. A digital clock has discrete states.
- **Analog signal:** An analog signal is a continuously varying signal over a period of time. Thus, an analog signal has infinite number of states. The signal curve passes over an infinite number of points. For example, telephone and television systems have used analog signals to transmit analog data.

The most basic analog signal is the sinusoidal wave, whose equation is as follows:

$$V=V_m \sin(\omega t+\Phi)$$

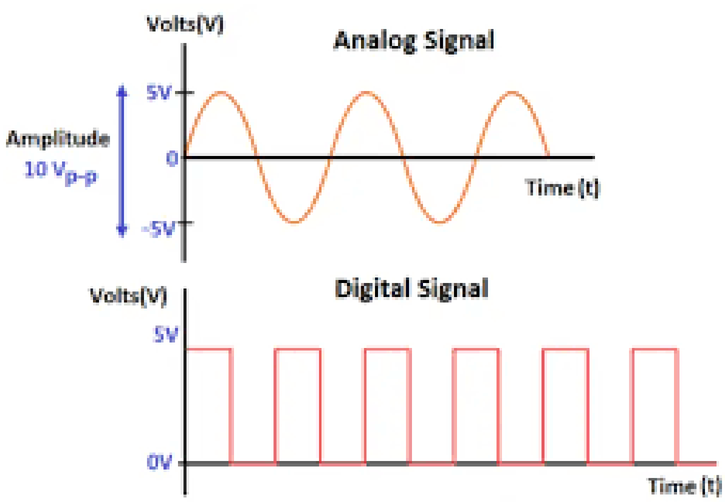
where:

- **V** is the instantaneous voltage at a given time.
- $\mathbf{V_m}$ is the peak voltage (maximum amplitude).
- **ω** is the angular frequency ($\omega=2\pi f$, where f is the frequency in Hz).
- **t** is time.
- **Φ** is the phase angle, which determines the starting position of the waveform.



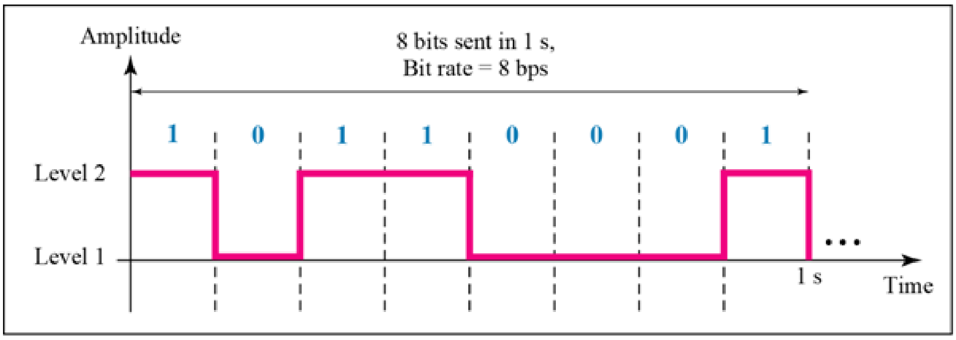
- **Digital signal:** A digital signal is a signal having finite number of states, i.e., it is a discrete signal. The signal maintains a constant level for some time and then changes to another level.

  The digital signal degrades quickly (due to attenuation), and if it is not frequently regenerated, it takes on a distorted form that the receiver cannot interpret. As a result, it can only be transmitted over short distances.



2.2. Bit Rate and Bit Length of Digital Signal

Consider transmission of binary data (digital) using a digital signal. Binary data has only bits - 0 and 1. If we use a digital signal with two levels - high and zero, we can transmit one bit per level. This is shown below.



a. A digital signal with two levels

Since a digital signal can have many discrete levels, more than one bit can be transmitted for each level. If a digital signal has **L** levels, the number of bits each level represents can be calculated as:

Bits per level = $\boldsymbol{\log_2 L}$ ***bits***

- **Example**: if the digital signal has 2 levels, each level can represent $\log_2 2 = 1$ bit. If L = 4, each level can represent 2 bits. For example, if the levels are L1, L2, L3 and L4, each level can represent two bits as follows:

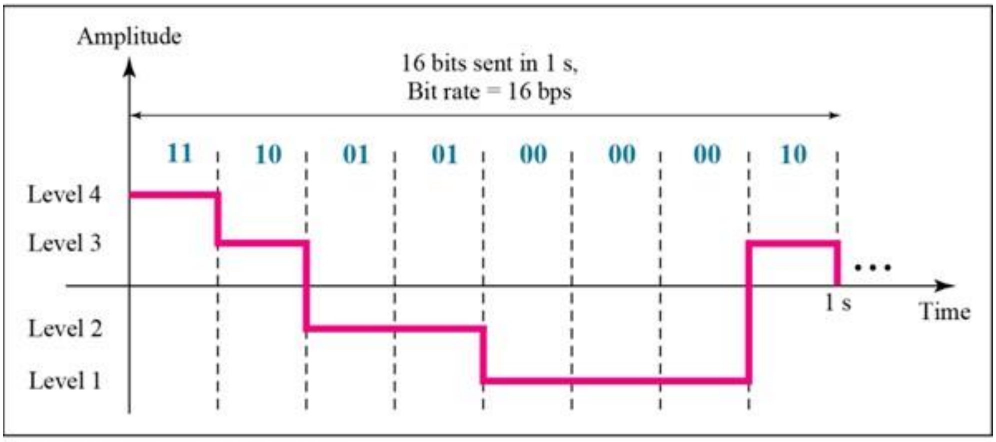
  L1:00

  L2:01

  L3:10

  L4:11

  If the data to be transmitted is 1110010100000010, it can be represented as follows:



b. A digital signal with four levels

- **Note:** If L is not a power of 2, consider the nearest higher power of 2 to calculate number of bits.

- **Bit Rate:** The bit rate of a digital signal is the number of bits transmitted per second. It is expressed as bps. The bit rate also depends on the channel bandwidth.
  For example, if the digital signal shown in figure-a is transmitted in 1 second, the bit rate is 8 bps and for figure-b, it is 16 bps.
- **Bit Length:** The bit length is the amount of distance one bit occupies on the transmission medium. It depends on the bit duration i.e. time for which one bit is transmitted and the propagation speed i.e. the speed at which the bit travels on the medium.

  Bit length = bit duration x propagation speed

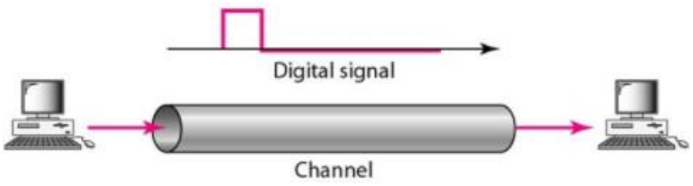
3. Baseband and Broadband Transmission

To transmit a digital signal over a transmission medium, two mechanisms are used:

1. Baseband transmission
2. Broadband transmission

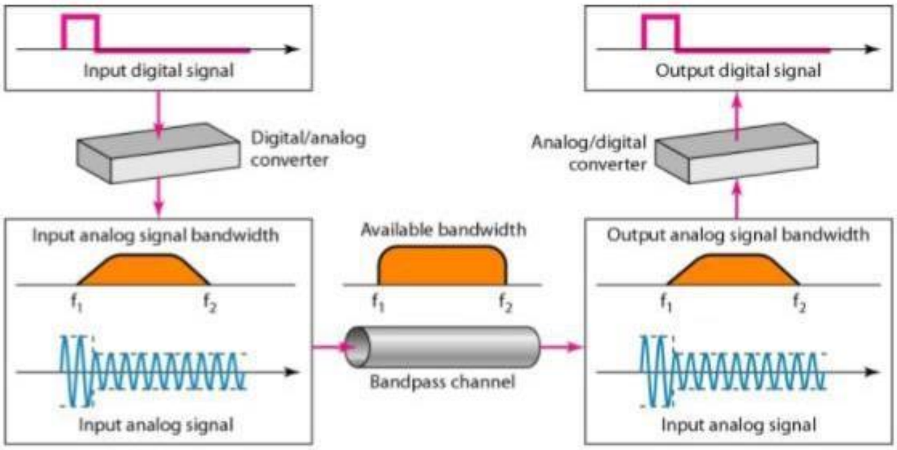
3.1. Baseband transmission

In this method, the digital signal is transmitted over the communication link without converting the digital signal to analog. In this case, only one signal can occupy the channel at a time. Baseband transmission requires that we have a low-pass channel, a channel with a bandwidth that starts from zero.



3.2. Broadband transmission

In this method, the digital signal is converted to analog using modulation. A high frequency carrier signal is modulated according to the digital signal and transmitted. Thus, several signals can be transmitted at the same time in different frequency bands (Frequency multiplexing). At the receiver, the analog signal is reconverted back to digital. Broadband transmission requires a **bandpass channel** - a channel with a bandwidth that does not start from zero.



4. Transmission Impairment

Communication channels are not perfect. When a signal travels through a communication medium, the signal undergoes several impairment, i.e., the quality of the signal does not remain the same but weakens. There are three causes of signal impairment: attenuation, distortion, and noise.

4.1. Attenuation

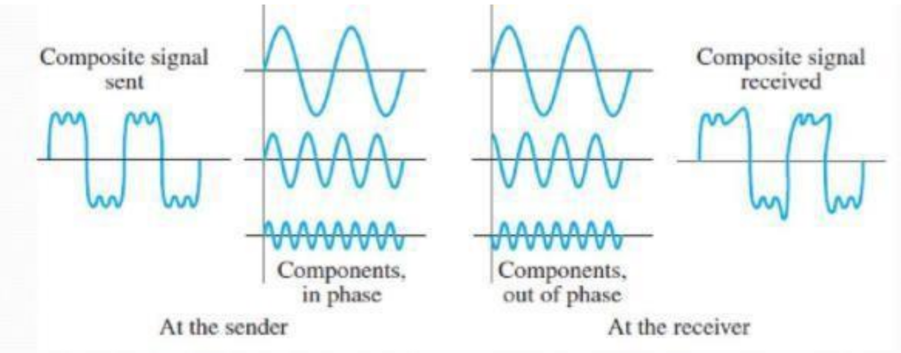
Attenuation means a loss of signal energy. As a signal travels through a communication medium, it loses its energy. This is because the signal has to overcome the resistance of the medium. The amount of attenuation depends on the medium characteristics and the distance traveled.

To compensate for this loss, amplifiers are used to amplify the signal.

Attenuation is measured in **decibels (dB).**

4.2. Distortion

Distortion means that the signal changes its shape or form. This happens because different frequency components of a composite signal have different propagation delays and different phase shifts as they travel. Hence, they do not arrive at the same time at the receiver due to which the received signal does not have the same shape as the original.



4.3. Noise

Noise is nothing but other signals near the medium which corrupt the data signal. This is an important cause of impairment. There are several types of noise:

- **Thermal noise**: This is caused due to the random motion of molecules (electrons) in the medium.
- **Crosstalk**: This is caused due to the interference of signals in channels or wires close to each other.
- **Induced noise**: This comes from sources such as motors and other appliances like microwaves, etc.
- **Impulse noise**: This is a spike which is a signal having a very high power but for a very short time. It is caused by high power lines, lightning and fluctuations in voltages.

Measure of noise is the Signal to Noise Ratio (SNR):

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

SNR is an important characteristic of the transmission system. **More the SNR, better is the quality of transmitted signal.** It is often represented in decibels as:

$$SNR_{dB} = 10 \log_{10} SNR$$

5. Data Rate Limits

Data rate can be defined as how fast can we send the data, in bits per second, over a channel. Maximum Data Rate (Channel Capacity) is the tight upper bound on the rate at which information can be reliably transmitted over a communication channel.

There are two theoretical formulas to calculate the data rate:

- Nyquist for a noiseless channel.
- Shannon for a noisy channel.

5.1. Noiseless Channel: Nyquist Bit Rate

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

$$BitRate = 2 * W * \log_2 (L)$$

In the above equation, **W** 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.

Bandwidth is a fixed quantity, so it cannot be changed. Hence, the data rate is directly proportional to the number of signal levels.

Increasing the levels of a signal may reduce the reliability of the system.

5.2. Noisy Channel: Shannon Capacity

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

$$Capacity = W * \log_2 (1 + SNR)$$

In the above equation, **W** 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.

Bandwidth is a fixed quantity, so it cannot be changed. Hence, the channel capacity is directly proportional to the power of the signal.

6. Performance of the Network

A computer network consists of several communicating devices, connecting devices and communication links. All these factors have an important impact on the performance of the network. The important criteria which are used to measure performance of a network are:

1. Bandwidth.
2. Throughput.
3. Latency (Delay).
4. Bandwidth –Delay Product.
5. Jitter.

- **Bandwidth:** 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.
  - 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. For example, we can say the bandwidth of a subscriber telephone line is 4 kHz.

    It is calculated as the difference between the lowest and highest frequency:

$$Bandwidth_{Hz} = f_h - f_l$$

  - The second, bandwidth in bits per second, refers to the speed of bit transmission in a channel or link. For example, one can say the bandwidth of a Fast Ethernet network (or the links in this network) is a maximum of 100 Mbps. This means that this network can send 100 Mbps.

- **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.

- **Delay (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

  - **Propagation time** is the time required for a single bit to travel from the source to destination. It is calculated as:

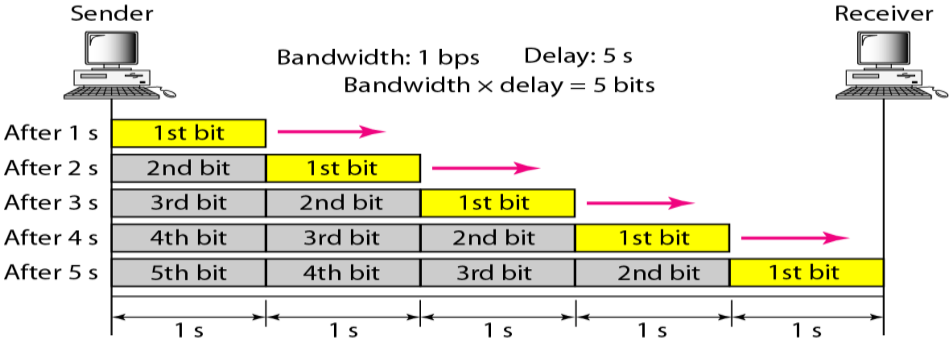
    Propagation time = Distance/Propagation speed

  - **Transmission time** A time based on how long it takes to send the signal down the transmission line. It depends on the message size and the bandwidth.

    Transmission time = Message Size/Bandwidth

  - **Queuing time** is the time taken by intermediate or end devices to store the entire message before it is processed. It depends on the network traffic and the device characteristics.
  - **Processing time** is the time taken by a device to process the message. The processing includes inspecting message headers to read source and destination address, calculate checksum to identify errors, etc.

- **Bandwidth-Delay Product:** The bandwidth-delay product defines the number of bits that can fill the communication link. This is used to calculate how many bits can be sent at a time before the acknowledgement is received so that the channel does not remain idle.



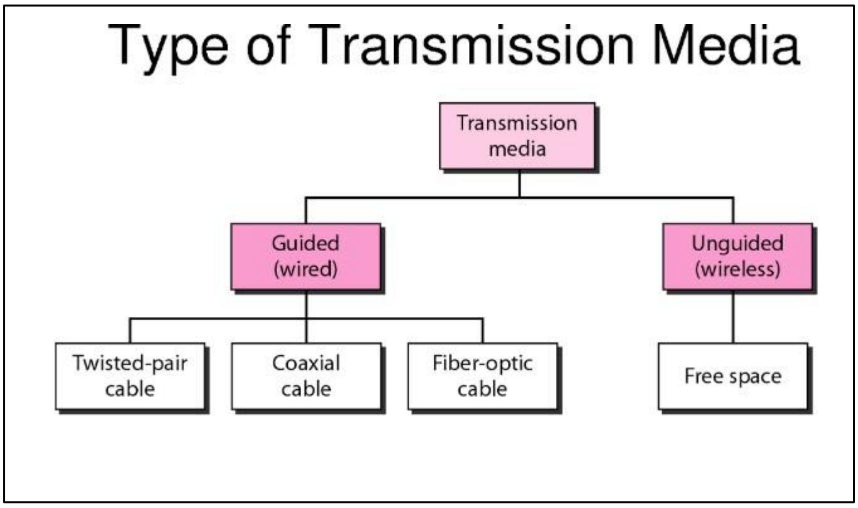
- ***Jitter:*** Latency refers to the delay in the message reaching the receiver. Jitter is the variation in this delay.

7. Transmission Media

A transmission medium can be broadly defined as anything that can carry information from a source to a destination. For example, the transmission medium for two people having a dinner conversation is the air. For a written message, the transmission medium might be a mail carrier, a truck, or an airplane.

In data communications, the definition of the information and the transmission medium is more specific. The transmission medium is usually free space, metallic cable, or fiber-optic cable. The information is usually a signal that is the result of a conversion of data from another form.

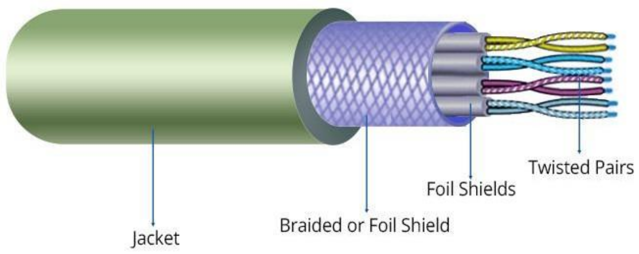
7.1. Classification of Transmission Media



7.1.1. Guided Media

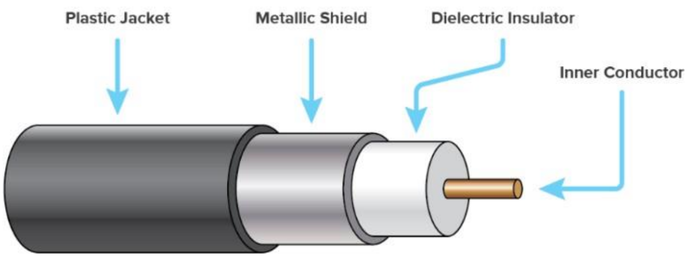
Guided media, which are those that provide a channel from one device to another, include twisted-pair cable, coaxial cable, and fiber-optic cable. A signal traveling along any of these media is directed and contained by the physical limits of the medium. Twisted-pair and coaxial cable use metallic (copper) conductors that accept and transport signals in the form of electric current. Optical fiber is a cable that accepts and transports signals in the form of light.

- **Twisted-Pair Cable:** A twisted pair consists of two conductors (normally copper), each with its own plastic insulation, twisted together, as shown in Figure below.



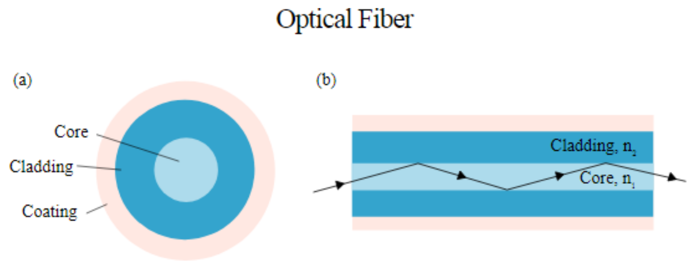
- **Applications:** Twisted-pair cables are used in telephone lines to provide voice and data channels. The DSL (Digital Subscriber Line) lines that are used by the telephone companies to provide high-data-rate connections also use the high-bandwidth capability of unshielded twisted-pair cables. Local-area networks also use twisted-pair cables.

- **Coaxial Cable:** Coaxial cable (or coax) carries signals of higher frequency ranges than those in twisted pair cable, in part because the two media are constructed quite differently. Instead of having two wires, coax has a central core conductor of solid or stranded wire (usually copper) enclosed in an insulating sheath, which is, in turn, encased in an outer conductor of metal foil, braid, or a combination of the two. The outer metallic wrapping serves both as a shield against noise and as the second conductor, which completes the circuit. This outer conductor is also enclosed in an insulating sheath, and the whole cable is protected by a plastic cover (see Figure below).



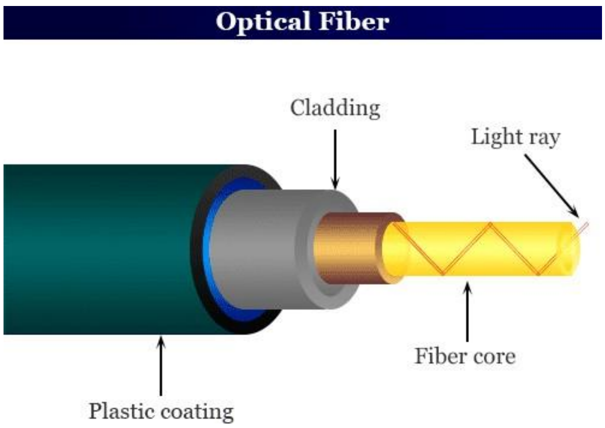
- **Applications:** Coaxial cable was widely used in analog telephone networks. Later it was used in digital telephone networks. However, coaxial cable in telephone networks has largely been replaced today with fiber-optic cable. Cable TV networks also use coaxial cables. However, cable TV providers replaced most of the media with fiber-optic cable. Another common application of coaxial cable is in traditional Ethernet LANs.

- **Fiber Optic Cable:** A fiber-optic cable is made of glass or plastic and transmits signals in the form of light. Optical fibers use reflection to guide light through a channel. A glass or plastic core is surrounded by a cladding of less dense glass or plastic. The difference in density of the two materials must be such that a beam of light moving through the core is reflected off the cladding instead of being refracted into it. See Figure below.



(a) Diagram of a bare fiber in the surrounding. (b) propagation of light by total internal reflection inside of an optical fiber.

- **Cable Composition:** Figure below shows the composition of a typical fiber-optic cable. The outer jacket is made of either PVC or Teflon. Inside the jacket are Kevlar strands to strengthen the cable. Kevlar is a strong material used in the fabrication of bulletproof vests. Below the Kevlar is another plastic coating to cushion the fiber. The fiber is at the center of the cable, and it consists of cladding and core.



- **Applications:** Fiber-optic cable is often found in backbone networks because its wide bandwidth is cost-effective. Some cable TV companies use a combination of optical fiber and coaxial cable, thus creating a hybrid network. Optical fiber provides the backbone structure while coaxial cable provides the connection to the user premises. Local-area network such as Fast Ethernet uses fiber-optic cable.
- **Advantages of Optical Fiber:**

  Fiber-optic cable has several advantages over metallic cable (twisted pair or coaxial):

  - *Higher bandwidth*. Fiber-optic cable can support dramatically higher bandwidths (and hence data rates) than either twisted-pair or coaxial cable.
  - *Less signal attenuation.* Fiber-optic transmission distance is significantly greater than that of other guided media. A signal can run for 50 km without requiring regeneration. We need repeaters every 5 km for coaxial or twisted-pair cable.

- *Immunity to electromagnetic interference.* Electromagnetic noise cannot affect fiber-optic cables.
- *Resistance to corrosive materials.* Glass is more resistant to corrosive materials than copper.
- *Light weight.* Fiber-optic cables are much lighter than copper cables.
- *Greater immunity to tapping.* Fiber-optic cables are more immune to tapping than copper cables. Copper cables create antenna effects that can easily be tapped.

➢ **Disadvantages of Optical Fiber:**

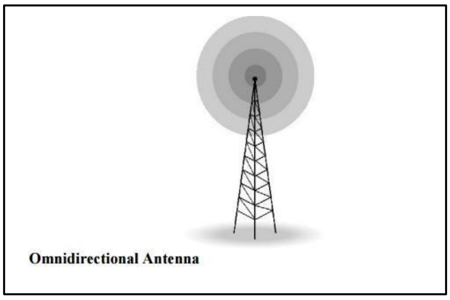
There are some disadvantages in the use of optical fiber:

- *Installation and maintenance.* Fiber-optic cable is a relatively new technology. Its installation and maintenance require expertise that is not yet available everywhere.
- *Unidirectional light propagation.* Propagation of light is unidirectional. If we need bidirectional communication, two fibers are needed. The cable and the interfaces are relatively more expensive than those of other guided media. If the demand for bandwidth is not high, often the use of optical fiber cannot be justified.

7.1.2. Unguided media: Wireless

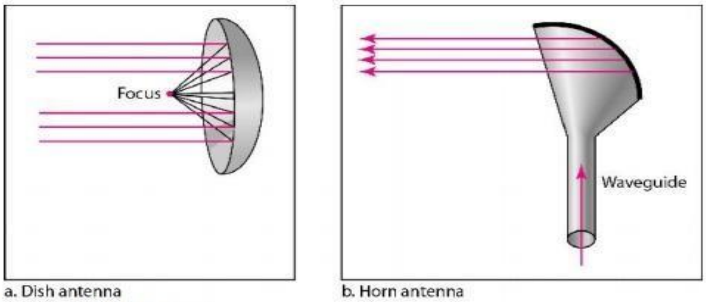
Unguided media transport electromagnetic waves without using a physical conductor. This type of communication is often referred to as wireless communication. Signals are normally broadcast through free space and thus are available to anyone who has a device capable of receiving them.

- **Radio Waves:** Waves ranging in frequencies between 3 kHz and 1 GHz are called radio waves. Radio waves, for the most part, are omnidirectional. When an antenna transmits radio waves, they are propagated in all directions. This means that the sending and receiving antennas do not have to be aligned. A sending antenna sends waves that can be received by any receiving antenna. The omnidirectional property has a disadvantage, too. The radio waves transmitted by one antenna are susceptible to interference by another antenna that may send signals using the same frequency or band.



- **Applications:** The omnidirectional characteristics of radio waves make them useful for multicasting, in which there is one sender but many receivers. AM and FM radio, television, maritime radio, cordless phones are examples of multicasting.

- **Microwaves:** Electromagnetic waves having frequencies between 1 and 300 GHz are called microwaves. Microwaves are unidirectional. When an antenna transmits microwave waves, they can be narrowly focused. This means that the sending and receiving antennas need to be aligned. The unidirectional property has an obvious advantage. A pair of antennas can be aligned without interfering with another pair of aligned antennas.

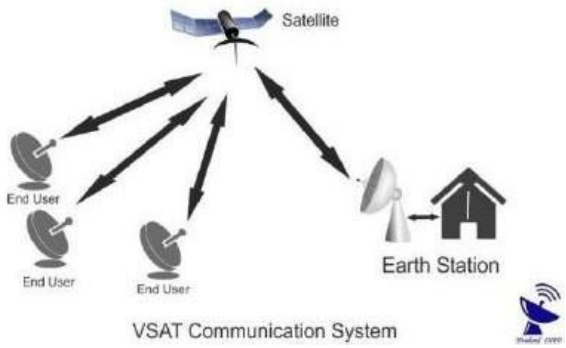


Unidirectional Antennas

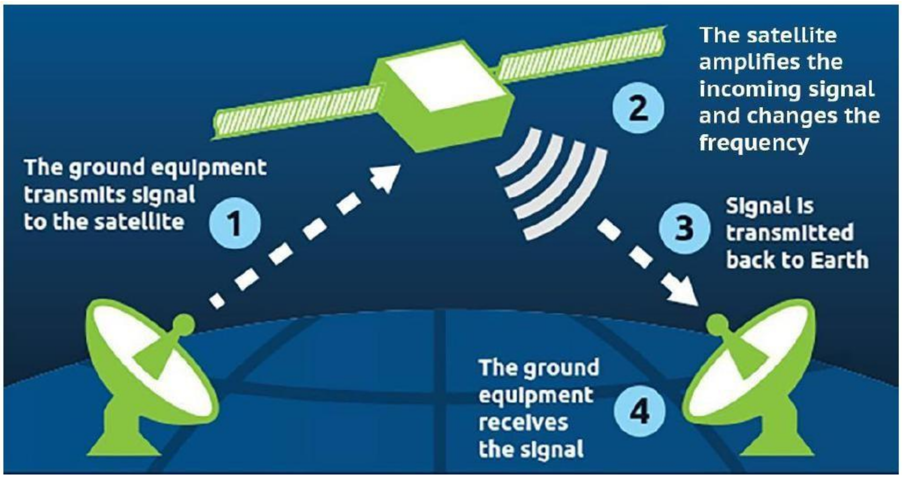
- **Applications:** Microwaves, due to their unidirectional properties, are very useful when unicast (one-to-one) communication is needed between the sender and the receiver. They are used in cellular phones, satellite networks and wireless LANs.

- **Infrared:** Infrared waves, with frequencies from 300 GHz to 400 THz, can be used for short-range communication. Infrared waves, having high frequencies, cannot penetrate walls. This advantageous characteristic prevents interference between one system and another; a short-range communication system in one room cannot be affected by another system in the next room.
  - **Applications:** There is standards for using these signals for communication between devices such as keyboards, mice, PCs, and printers.

- **VSAT:** A very small aperture terminal (VSAT) is a small telecommunication earth station that receives and transmits real-time data via satellite. VSAT is a satellite communications system that serves home and business users. A VSAT end user needs a box that interfaces between the user's computer and an outside antenna with a transceiver. The transceiver receives or sends a signal to a satellite transponder in the sky. The satellite sends and receives signals from an earth station computer that acts as a hub for the system. For one end user to communicate with another, each transmission has to first go to the hub station which retransmits it via the satellite to the other end user's VSAT.



- **Satellite:** A communications satellite is an artificial satellite that relays and amplifies radio telecommunications signals via a transponder; it creates a communication channel between a source transmitter and a receiver at different locations on Earth. Communications satellites are used for television, telephone, radio, internet, and military applications. The purpose of communications satellites is to relay the signal around the curve of the Earth allowing communication between widely separated geographical points.

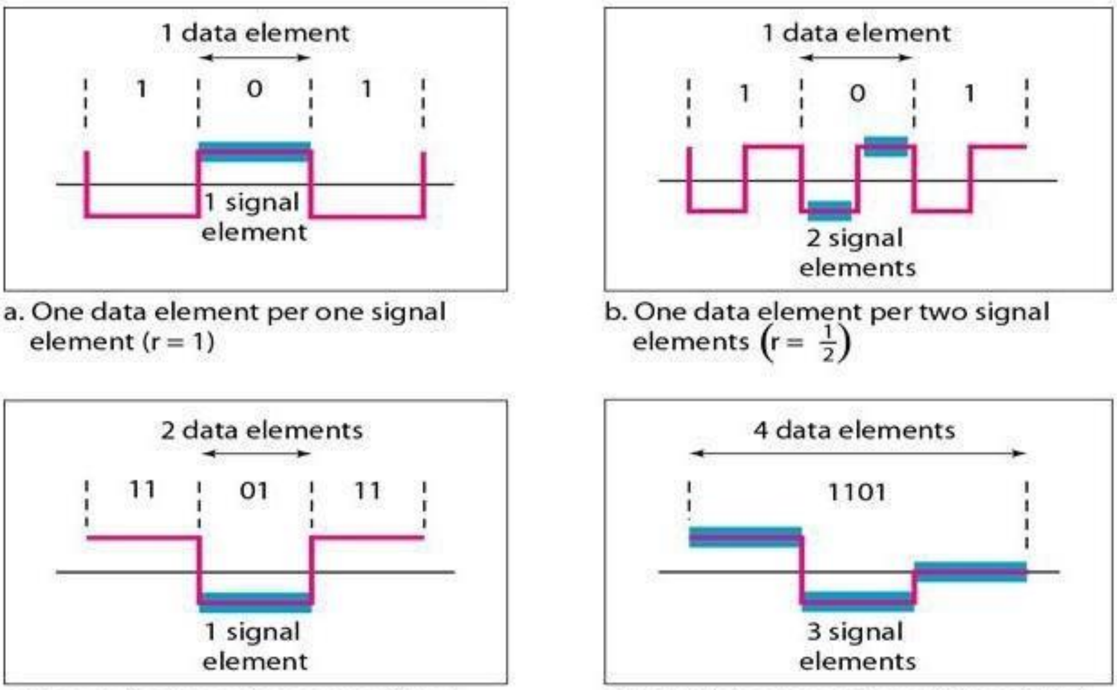


8. Line Coding

Line coding is a process of converting a series of bits into a digital signal. Several encoding schemes have been developed. This is done so that the receiver can clearly determine the start, stop or the middle of each bit in order to correctly receive the data.

Here, are some characteristics of line encoding schemes:

- **Signal element vs data element:** A data element is the smallest entity that represents piece of information. A signal element is the shortest unit in a digital signal that carries data elements. More than one data elements can be carried in a single signal element and more than one signal element can carry a single data element. **r** is the number of data elements carried by a signal element.



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

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

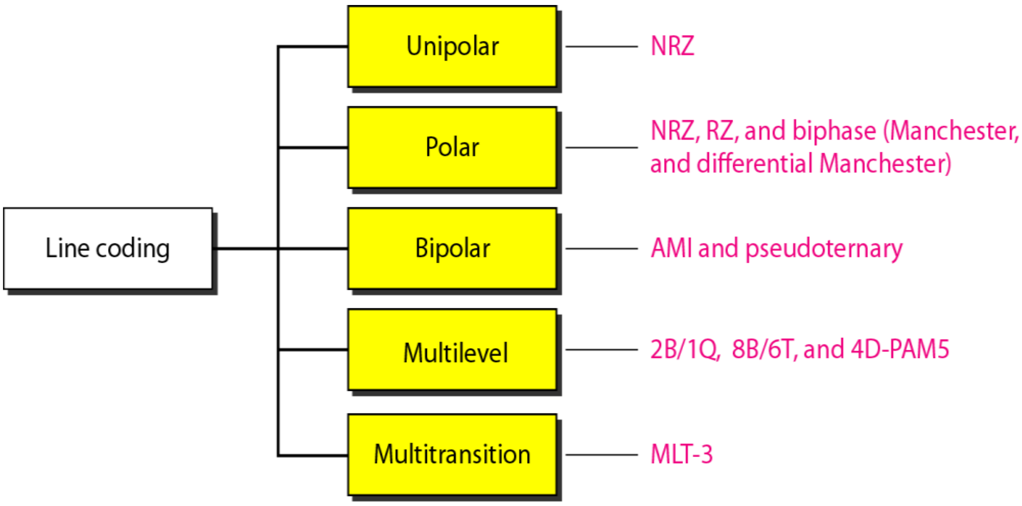
c. Two data elements per one signal element (r = 2)

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

- **Signal rate vs data rate:** Signal rate (**baud rate**) represents the number of signal elements in the digital signal sent per second. The data rate (bit rate) is the number of data elements sent per second. The goal of data communication is to increase the bit rate while keeping the baud rate low.
- **Self synchronization:** A long sequence of 0's or l's may cause the receiver to go out of sync and hence identify data bits wrongly. The encoding scheme must include information by having transitions in the beginning, middle or end of the pulse so that the receiver can synchronize itself.

8.1. Encoding Schemes

Binary encoding schemes fall into five broad categories as shown in the figure below. These are Unipolar, Polar, Bipolar, Multilevel and Multitransition.



8.1.1. Unipolar

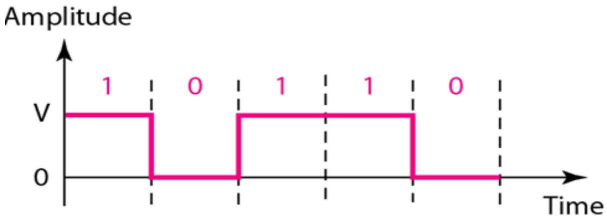
This is the simplest encoding scheme. In Unipolar encoding, only one polarity is used; i.e. all signal levels are on one side of the time axis - above or below.

- **NRZ - Non Return to Zero:** NRZ means the signal does not return to zero in the middle of the bit. In this scheme,

  Bit 1 is represented by a positive voltage

  Bit 0 is represented by zero voltage signal

  - **Example:** Binary data – 10110



Unipolar NRZ

8.1.2. Polar

Polar encoding uses two voltage levels, positive and negative on both sides of the zero axis.

The types of polar encoding are:

i. NRZ - Non Return to Zero

1. NRZ-L (NRZ- Level)
2. NRZ-I (NRZ- Invert)

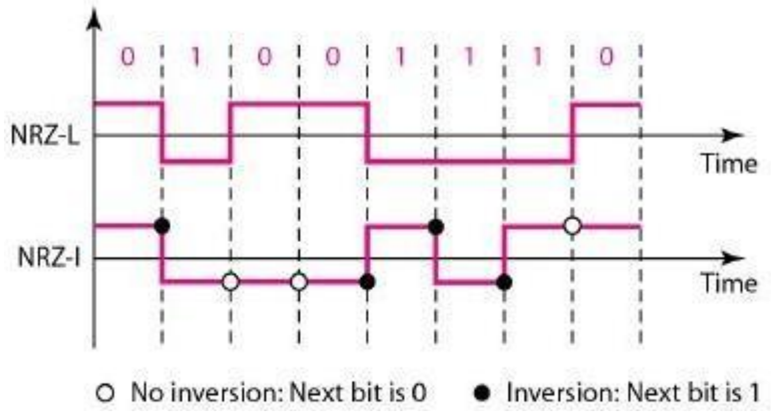
ii. RZ- Return to Zero

iii. Manchester

iv. Differential Manchester

- **NRZ (Non Return to Zero):** Here, the signal level is always either positive or negative but never zero. It has two variations:
  - **NRZ-L (NRZ- Level):** where 0 is usually a positive voltage and 1 is a negative voltage. However, for a long series of l's or 0's it can lead to synchronization problems.
  - **NRZ-I (NRZ- Invert):** In this method, fixed voltage levels are not assigned to represent 0 or 1 but an inversion of existing voltage level represents a 1 and non-inversion of level represents 0.

- **Example:** Binary data 01001110

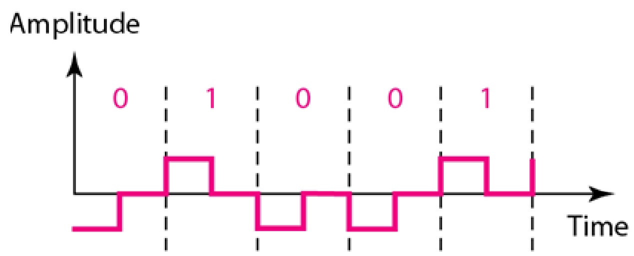


Polar NRZ-L and NRZ-I

- **RZ (Return to Zero):** In the above schemes, a long string of 0's or l's can make the receiver lose synchronization. To ensure synchronization, there should be a signal change for each bit.

  RZ uses three levels - positive, negative and zero. The RZ code requires the signal to return to 0 in the middle of its bit period i.e. for every bit, there has to be a transition in the middle. Bit 1 is represented as Positive to zero and 0 is represented as Negative to zero signal.

  - **Example:** Binary data 01001



Polar RZ

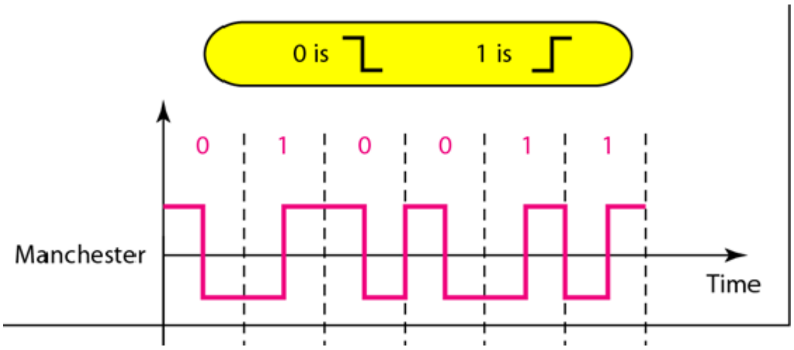
- **Manchester:** This code also requires a transition in the middle of each bit interval.

  Negative-to-positive transition represents 1

  Positive-to-negative represents 0

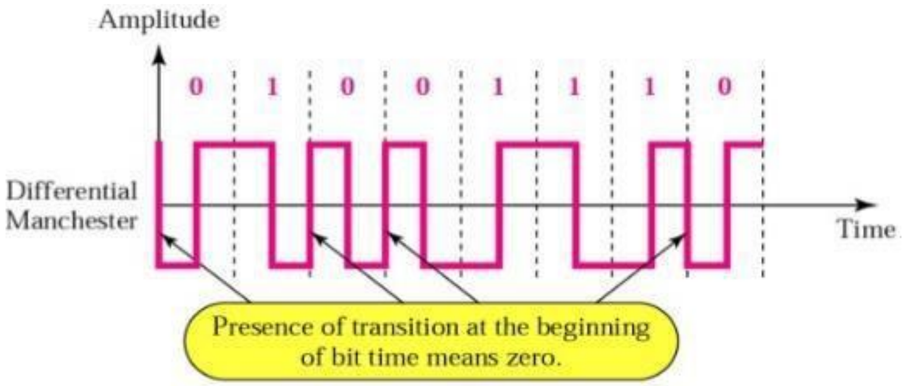
  Thus, by using just two levels of amplitude, it can achieve the same level of synchronization as RZ code.

- **Example:** Binary data 010011



Manchester encoding

- **Differential Manchester:** This combines RZ and NRZ-I.
  A transition at the beginning of the next bit means 0 and no transition means 1. Here, a transition in the middle is used for synchronization but the presence or absence of a transition at the beginning identifies the bit.
  - Example: Data: 01001110



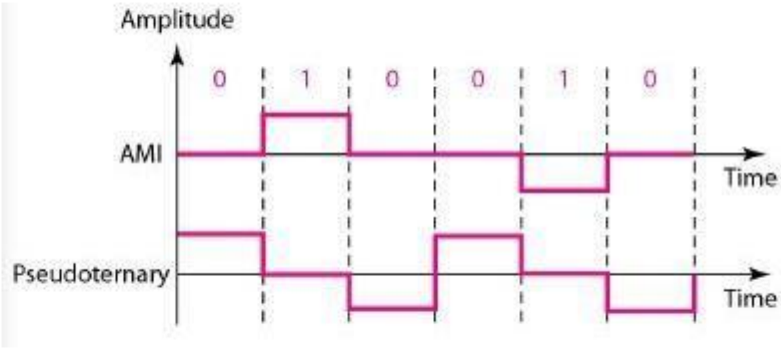
Differential Manchester encoding

8.1.3. Bipolar

In bipolar encoding, we use three levels: positive, zero, and negative.

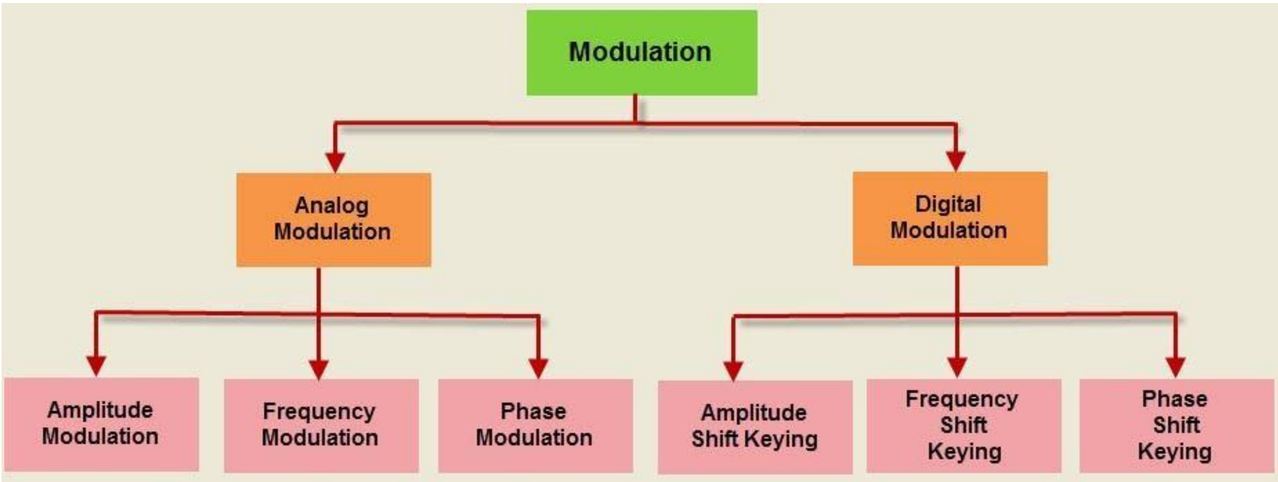
- **Alternate Mark Inversion (AMI)**: is a bipolar encoding system where neutral (zero) voltage represents binary 0 and alternating positive and negative voltages represents binary 1. With this line encoding it is the alternating voltages that determine the binary ones.
- **Pseudoternary**: Bit 1 is encoded as a zero voltage and the bit 0 is encoded as alternating positive and negative voltages i.e., opposite of AMI scheme.

- **Example:** Data: 010010



9. Modulation

Modulation is a process of suppressing low frequency information signal on a high frequency carrier signal. Basically, the modulation is of following two types: Analog Modulation and Digital Modulation.



Categories of modulation

9.1. Digital Modulation Techniques

Digital modulation is defined as the modulation process in which discrete signals are used for modulating carrier waves and it is used for removing noise from the waves. Digital modulation methods can be considered as digital-to-analog conversion.

There are many types of digital modulation techniques and also their combinations, depending upon the need. Of them all, we will discuss the prominent ones.

There are three types of digital modulation:

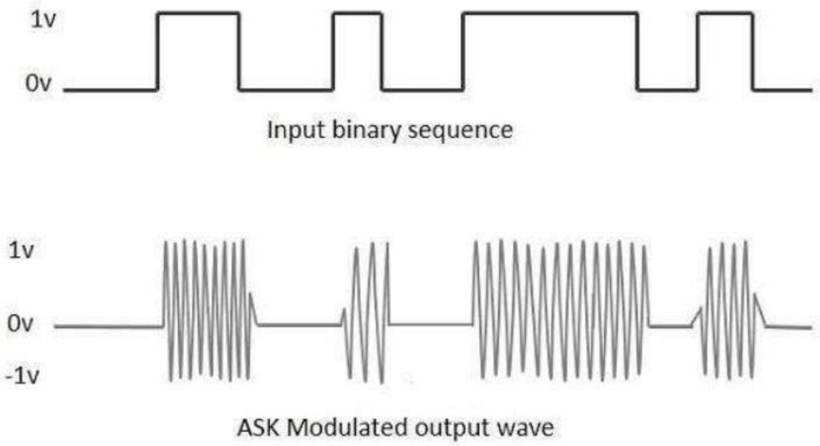
- Amplitude shift Keying (ASK).
- Frequency shift Keying (FSK).
- Phase shift Keying (PSK).

9.1.1. ASK – Amplitude Shift Keying

Amplitude Shift Keying (ASK) is a type of Amplitude Modulation which represents the binary data in the form of variations in the amplitude of a signal.

Any modulated signal has a high frequency carrier. The binary signal when ASK modulated, gives a zero value for Low input while it gives the carrier output for High input.

The following figure represents ASK modulated waveform along with its input.

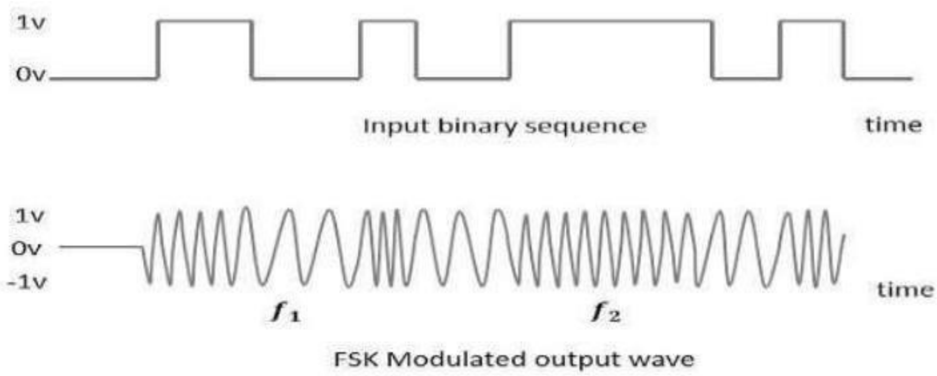


9.1.2. Frequency Shift Keying (FSK)

Frequency Shift Keying (FSK) is the digital modulation technique in which the frequency of the carrier signal varies according to the digital signal changes. FSK is a scheme of frequency modulation.

The output of a FSK modulated wave is high in frequency for a binary High input and is low in frequency for a binary Low input. The binary **1s** and **0s** are called Mark and Space frequencies.

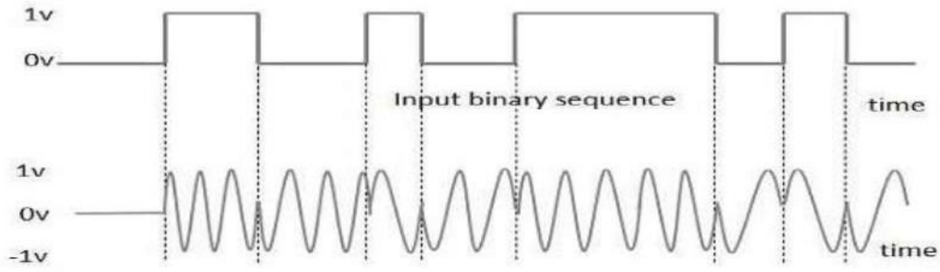
The following image is the diagrammatic representation of FSK modulated waveform along with its input.



9.1.3. Phase Shift Keying (PSK)

Phase Shift Keying (PSK) is the digital modulation technique in which the phase of the carrier signal is changed by varying the sine and cosine inputs at a particular time.

Following is the diagrammatic representation of BPSK Modulated output wave along with its given input.



BPSK Modulated output wave

10. Digitization (Demodulation)

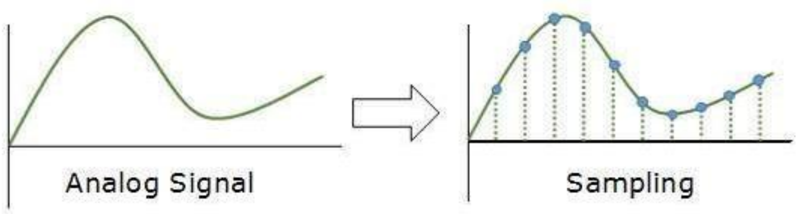
Digitization is the process of converting an analog (continuous) signal into a digital (discrete) signal by capturing values at regular time intervals. At specific moments, a sample of the analog signal is taken to obtain its digital equivalent. The analog-to-digital conversion technique is known as PCM (Pulse Code Modulation) or MIC (Modulation by Impulse and Coding). This process consists of three main steps: sampling, quantization, and encoding.

10.1. Sampling

Sampling involves extracting a certain number of fractions of the signal (samples). The rate at which these samples are captured is called the sampling frequency (or rate), expressed in Hertz (Hz). For example, an audio CD contains data sampled at 44.1 kHz, meaning 44,100 samples per second.

Sampling reduces the signal to a sequence of discrete points, which has two main consequences:

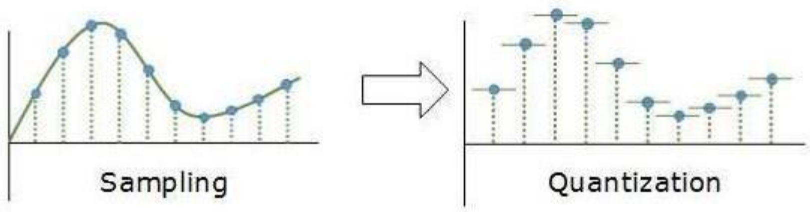
- Only the information present at the capture points is recorded.
- All other information is lost.



- **Shannon's Theorem (Digitization):** Shannon states that the minimum sampling frequency must be at least twice the maximum frequency of the signal being sampled.
  - **Example:** For a speech signal, most information is contained in a 4000 Hz bandwidth. A sampling rate of 8000 Hz is sufficient, meaning a sample is taken every 125 µs.
  - **Note:**
    - Sampling at a lower frequency results in a poor-quality reconstructed signal.
    - Sampling at a higher frequency increases the amount of data to be transmitted without significantly improving quality.

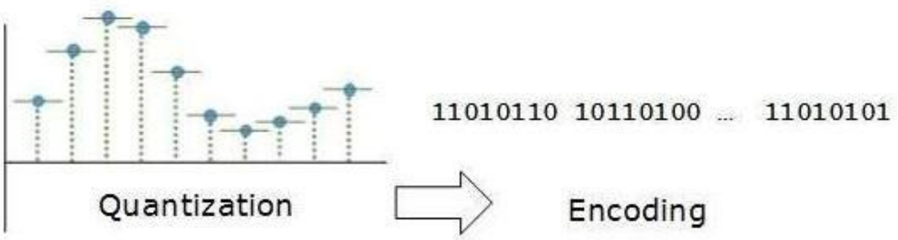
10.2. Quantization

Quantization assigns each sample a value from a predefined range of possible values.



10.3. Encoding

Each sample (quantized value) is then encoded into a sequence of bits. 8 bits are required to encode different values accurately.

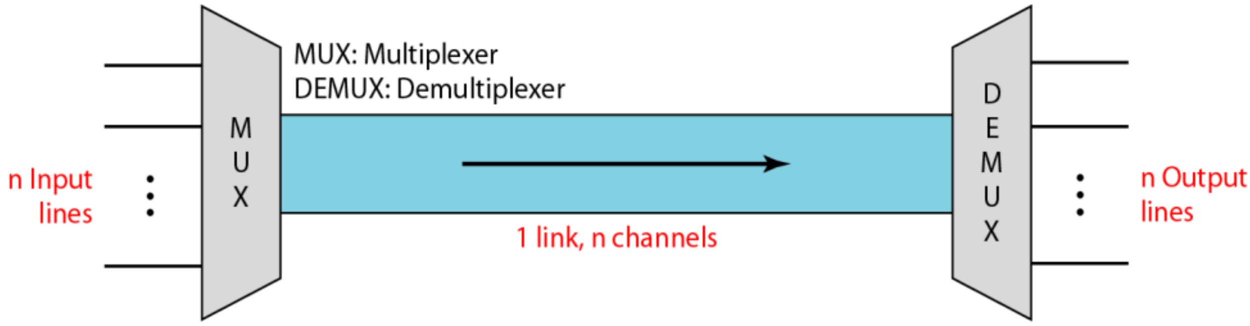


11. Multiplexing

It is not possible to set up a separate connection between each pair of communicating devices. In such cases, a single communication link may be shared among multiple devices.

This communication link is usually of a high capacity.

Multiplexing is a technique which allows simultaneous transmission of multiple signals across a single communication channel. High bandwidth media such as fiber optic cables and satellite links have the capacity to accommodate several signals. Here, multiplexing schemes are used so that the bandwidth can be efficiently shared among multiple simultaneous transmissions. The following figure depicts multiplexing.



Multiplexing

A multiplexer takes several inputs, which are transmitted over a high capacity signal link. At the receiving end, the demultiplexer separates the data and delivers it to appropriate output line. Two basic multiplexing schemes are used:

1. Frequency Division Multiplexing (FDM)
2. Time Division Multiplexing (TDM)

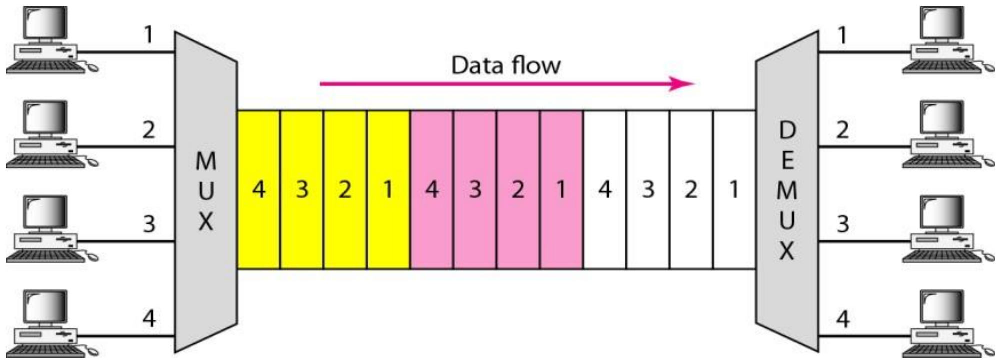
11.1. Frequency Division Multiplexing

- Here, the frequency spectrum of the communication channel is divided among several logical channels, each channel carrying one signal.
- A station is assigned one channel and it has exclusive possession of that frequency band as long as it wants to communicate. This scheme is used when the bandwidth of the transmission medium exceeds the required bandwidth of the signal to be transmitted.
- A guard band is a small unused frequency channel between two channels to prevent interference.
    - **Example:** Radio Broadcasting is a common example where we receive multiple stations. Each station is assigned a frequency range.

- **Advantages:**
    - Multiple stations can share the same communication channel simultaneously.
    - Each station gets full use of the channel for the duration of the communication.
- **Disadvantages:**
    - Since the channel bandwidth is divided into smaller channels, each station gets only a small bandwidth for communication.
    - If there are many stations, it may not be possible to assign a channel to each station. Moreover, this will reduce the transmission bandwidth for each channel.

11.2. Time Division Multiplexing

- In TDM, time is divided into short periods called time slots.
- Each user is assigned a time slot or time slice for communication in a round robin fashion.
- During a time slot, the user can use the entire channel capacity for transmission.
- In the figure, portion of the signals 1, 2, 3 & 4 occupy the link sequentially.



Time Division Multiplexing

- **Example:** Radio transmission where the stations are transmitted using FDM but each station is subdivided into two logical stations – one for music and other for advertising which are transmitted one after the other using TDM.

- **Advantages:**
  - Each station can use the entire bandwidth for the duration of a time slice.
  - Stations can be assigned priorities and assigned more time slots.
- **Disadvantages:**
  - If some station does not have data to send, the time slot is wasted and the channel is underutilized.
  - If the number of stations is very large, a station has to wait for a long time for its time slot.

12. Switching

When there are multiple devices, how to have one-to-one communication between two communicating devices?

One solution is to have a mesh topology or star topology. However, for very large networks, this is highly impractical. A multipoint connection like bus is also not possible due to the distances involved.

A better solution is switching. A switched network consists of a series of interlinked nodes called switches which are capable of creating temporary connections between two or more devices.

The types of switching are:

1. Circuit Switching
2. Message Switching
3. Packet Switching

12.1. Circuit Switching

This is a method in which an end-to-end path or circuit is established between communicating machines. The machines have exclusive use of this path until connection is released. This involves three phases:

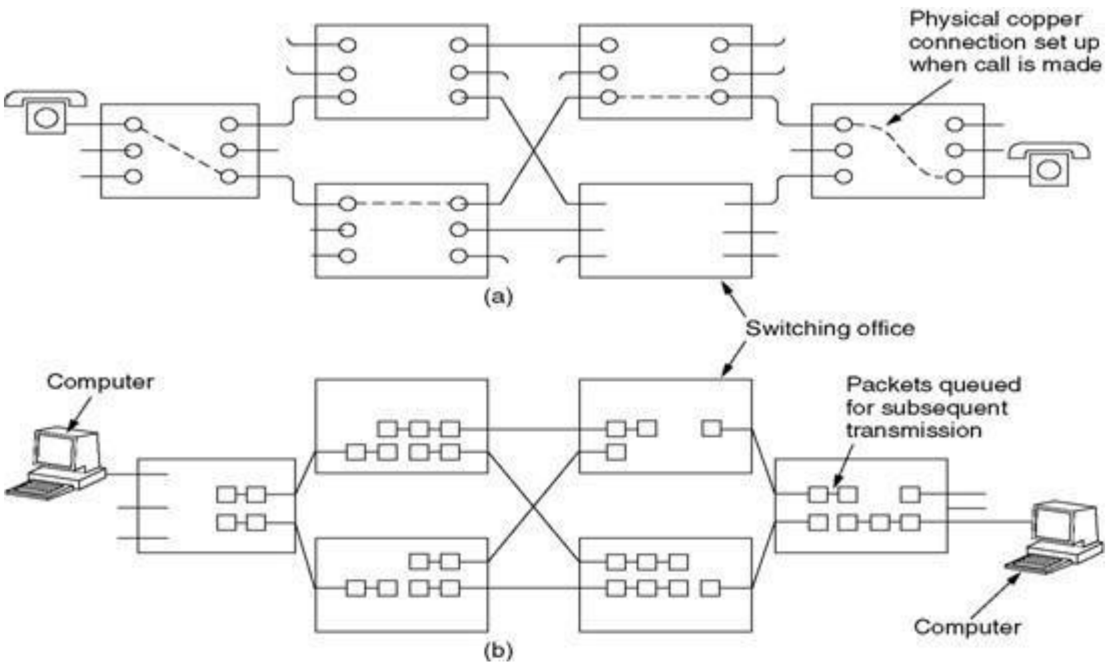
i. **Circuit Establishment:** Before any data transfer can begin, an end-to-end path has to be set up. Some setup time is required.

ii. **Data Transfer:** Once the circuit is established, analog or digital data can be transmitted depending upon the nature of the network. The only delay now is the propagation delay. There is no danger of congestion or a busy signal.

iii. **Circuit Disconnection:** The circuit can be released by either of the connected stations after data transfer takes place.

An example of circuit switching is the telephone system where a physical circuit is setup between the source and destination machine.

The schematic of circuit switching is shown in figure (a).



Here, the six rectangles represent switching offices, each having 3 incoming and 3 outgoing lines. When a call is placed, a physical connection is established between the incoming line and one of the output lines (shown by dotted lines).

- **Advantage:** This method transfers data in real time with the only delay being in circuit setup and propagation delay.
- **Disadvantages:**
  - Circuit switching can be inefficient. Even if no data is being transferred, the channel remains dedicated for the duration of the connection.
  - Problems occur when two or more stations attempt to establish a link with the same station. Such situations have to be resolved.

12.2. Message Switching

- In message switching, a physical path is not set up between the sender and the receiver.
- When the sender has a block of data to send, it is sent to intermediate switching station which store it and sends it to the appropriate station when an output line is free.
- This mechanism is referred to as 'Store and Forward' method.
- Each block is received as a whole checked for errors and retransmitted. Thus, a block may 'visit' several switching stations, before reaching the destination.
- An example of message switching is the e-mail system which is now replaced by packet switching.
  - **Advantages:**
    - No circuit has to be set-up in advance.
    - The sender can send data whenever it wants to and does not need to check the status of the receiver-whether it is busy or idle.
  - **Disadvantages:**
    - The system implementing message switching requires sufficiently large data buffers to hold the messages.
    - A single block may tie up a line for a long time, thereby causing delay to the other messages queued up on that link.
    - If there's a lot of traffic on the network, the delay will be very high thereby, reducing throughput.
    - Complicated routing algorithms are required.

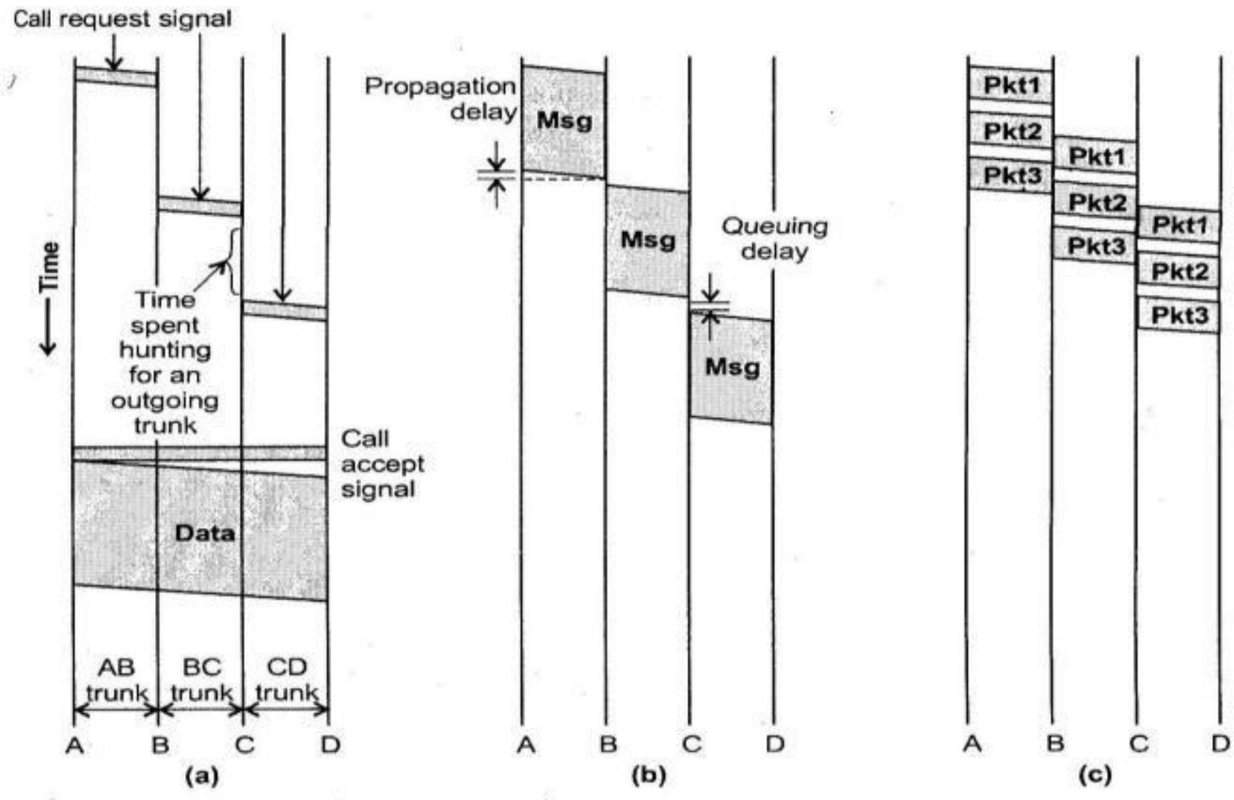
12.3. Packet Switching

- In packet switching, a message is broken into several small blocks called “packets".
- A limit is placed on the maximum block size, thereby, making it easier to store packets and route them through the network.

- Each packet contains control information including source and destination address and are routed independent of the other packets to the same destination, i.e. two packets for the same destination may be sent via different paths.
- Thus, it is possible that the packets arrive out of order. So, some identification scheme has to be employed.

- **Advantages:**
  - Call setup phase is avoided.
  - This is more flexible. Thus, if congestion develops in one part of the network, the packets can be routed via different paths.
- **Disadvantages:**
  - There is no guarantee that packets will be delivered.
  - Requires more overheads since each packet has to carry a lot of control information.
  - The packets may arrive out of order.

The difference between Circuit, Message and Packet switching is illustrated in figure:



Timing of Events

(a) Circuit Switching (b) Message Switching (c) Packet Switching