SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SECA1303 – ANALOG COMMUNICATION SYSTEMS

UNIT I AMPLITUDE MODULATION AND DEMODULATION

Need for modulation– Model of communication system and classification, Representation of AM – Modulation index and power calculation –Types of AM;DSB-FC: Collector and base modulation circuits, square law modulator- DSB-SC: Balanced modulator circuit using FET – SSB: Filter method and phase shift method – VSB, Comparison of various AM schemes-AM transmitter: Low level and high level Modulation. Demodulation –Envelope detector, Significance of RC time constant-Square law detector.

Introduction to communication systems:

- The word “communicate” means it is an act of passing on news, information, feeling etc.
- There are numerous ways for communication, i.e., Two people may communicate with each other through speech. [Gestures] or graphic symbol. But there is a lot of limitations, mainly [long] distance [For long distance communication]
- In Olden day’s people used so many kinds of communication method for Long distance [was accomplished by] such [means] as birds drum beats, smoke signals, [carrier pigeons] and light beams etc.
- Due to the invention of transistors, IC and other semiconductor devices, today the mode of long distance communication has been accomplished by electrical signals. This is because of the electrical signals can be transmitted over a much longer distance (theoretically any distance in the universe).
- Today, in our daily lives we are using powerful technologies that allow us to communicate with people around the world
- The tremendous growth information technology

Examples for Electronic Communication are

1. Telephone
2. Radio Broadcasting
3. Television Broadcasting
4. Radar communication
5. satellite communication: Cover the whole globe carrying voice, text data and images
6. Fax
7. Computer communication
8. Wireless communication

- This growth is possible due to high speed transmitter and receiver, development in VLSI technology, high speed microprocessors and high bandwidth transmission media.

What is Communication

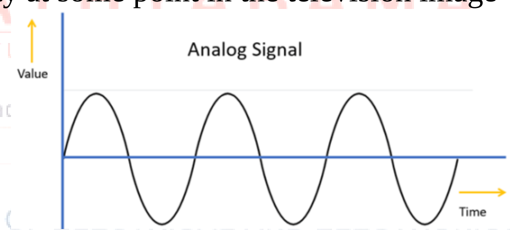
Electronic Communication is the process of establishing connection (or link) between two points (source and destination) for information exchange (or) It is the process of conveying or transferring messages such as sounds, words, pictures etc. from one point to another point (or) Communication refers to sending, receiving and processing of information by electronic means.

Message signal/ information signal/ baseband signal/ signal

- The basic function of communication system is to transmit a message or information signal from one place to another.
- The origin of the message is from some information source
- There are many kinds of information sources like machines as well as people.
- The message may be in the form of words, code symbols, music, picture etc.
- But we can classify the message signal into two categories
  Analog signal and Digital signal
- The nature of the information signal determine the nature and performance of a system

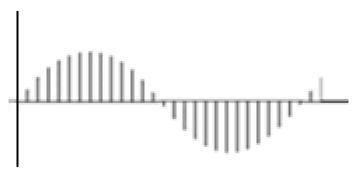
Analog message signal

- It is a physical quantity that varies with respect to time in a smooth and continuous manner.
- When a physical quantity is converted into its equivalent electrical quantity whose magnitude/strength varies with respect to time in a smooth and continuous manner.
  Examples
  1. An acoustic pressure is produced when we speak and it is converted into an equivalent electrical signal with the help of a transducer called microphone. This electrical quantity is varying with respect to time in a smooth and continuous manner.
  2. The signal intensity at some point in the television image



Digital message

- The amplitude of an electrical quantity varies with respect to time in a discrete manner not continuously or It is an ordered sequence of symbols (quantity from a set of discrete elements)
- Example: the keys when you press on a computer keyboard[the amount of information conveyed in any message is measured in bits]



In communication systems, this physical quantities are called as messages carry some sense or meaning and they are converted into equivalent electrical quantity called as message signal or modulating signal or base band signal. This electrical signal is transmitted to a distant place through a communication media (i.e) communication channel. At the receiving end electrical signal is reconverted back into original message.

Time/Frequency Domain Representation of Signals

Electrical signals have both time and frequency domain representations.

In the time domain, voltage or current is expressed as a function of time. Most people are relatively comfortable with time domain representations of signals. Signals measured on an oscilloscope are displayed in the time domain and digital information is often conveyed by a voltage as a function of time.

Time domain visualization provides information such as shape of the signal and variation in voltage with respect to time. But it did not provide complete information regarding frequency content of a signal.

Signals can also be represented by a magnitude and phase as a function of frequency. Signals that repeat periodically in time are represented by a discrete power spectrum. Frequency domain representations are particularly useful when analyzing linear systems. EMC and signal integrity engineers must be able to work with signals represented in both the time and frequency domains. Signal sources and interference are often defined in the time domain. However, system behavior and signal transformations are more convenient and intuitive when working in the frequency domain.

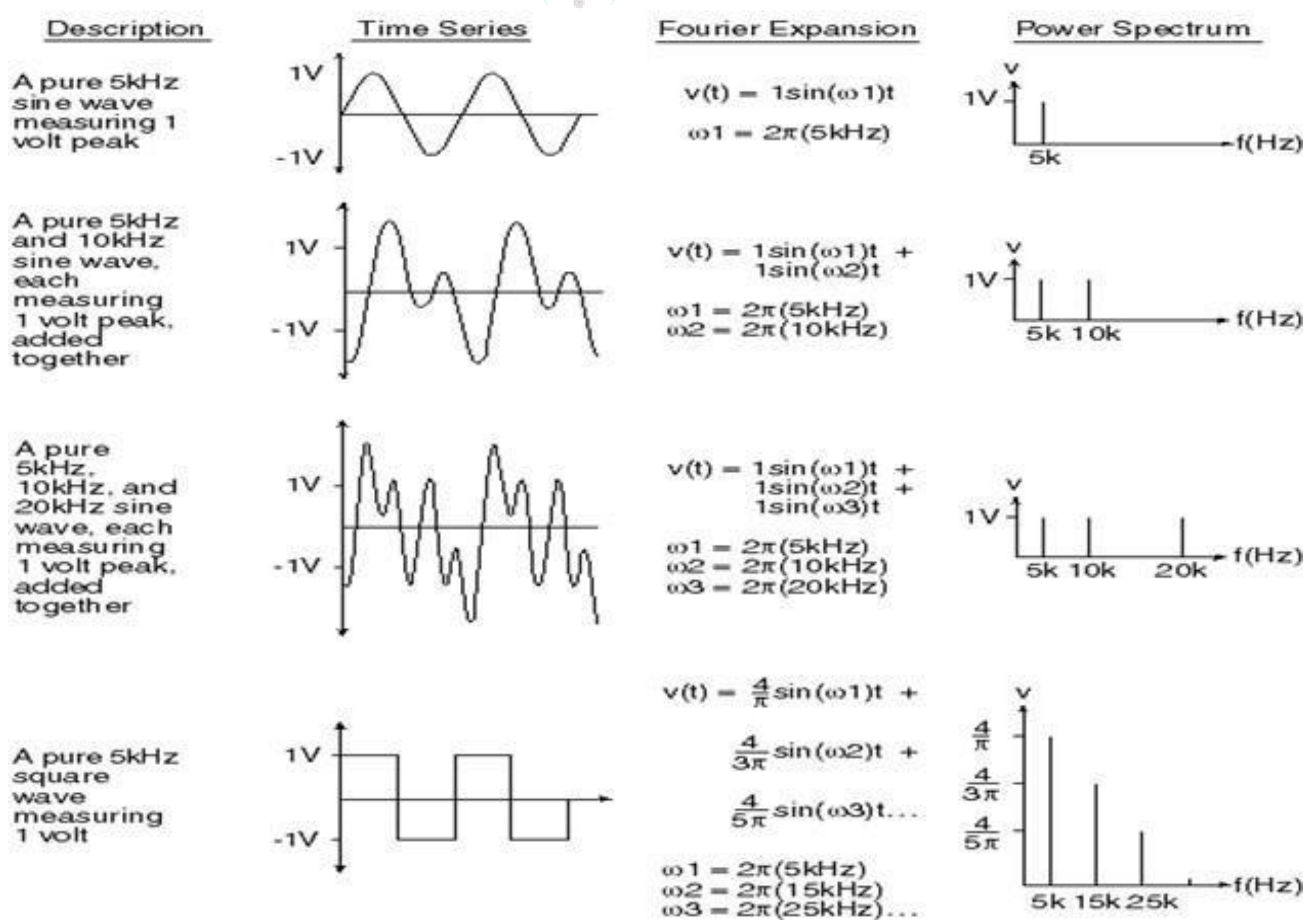


Figure Periodic signals in the time and frequency domain.

Various stages and block diagram of communication system or elements of a communication system or model of a communication system

Communication system is a group (collection) of subsystems such us transmission subsystem, channel and receiver subsystem. They are designed and assembled in a proper manner for sending and receiving information signals.

Transmission subsystem: It is a group or collection of individual components and circuits such as information source, input transducer, modulator circuits (mixer, carrier signal generator and band pass filter), amplifiers and transmitting antenna etc. They are designed and assembled in a proper manner to convert the data/ information/ any physical quantity (carries some meaning of sense) into a suitable form for the transmission over the channel

Receiver subsystem: It is also a group or collection of individual components and circuits such as receiving antenna, amplifiers, demodulator circuits, low pass filter, output transducer and etc. They are designed and assembled in a proper manner to reconstruct the original data/information from the received signal (from the channel)

The three essential components for communication system are

1. Transmitter: lt sends information. For example TV transmitting stations or radio transmitting stations are `senders', since they transmit information.
2. Receiver: It receives information. For example all TV sets and radios are receivers. They get information from transmitter
3. Communication Channel: This is the path through which the signal propagates from transmitter to receiver.

The basic components of a communication system are information source, input transducer, transmitter, communication channel, receiver, output transducer, and destination.

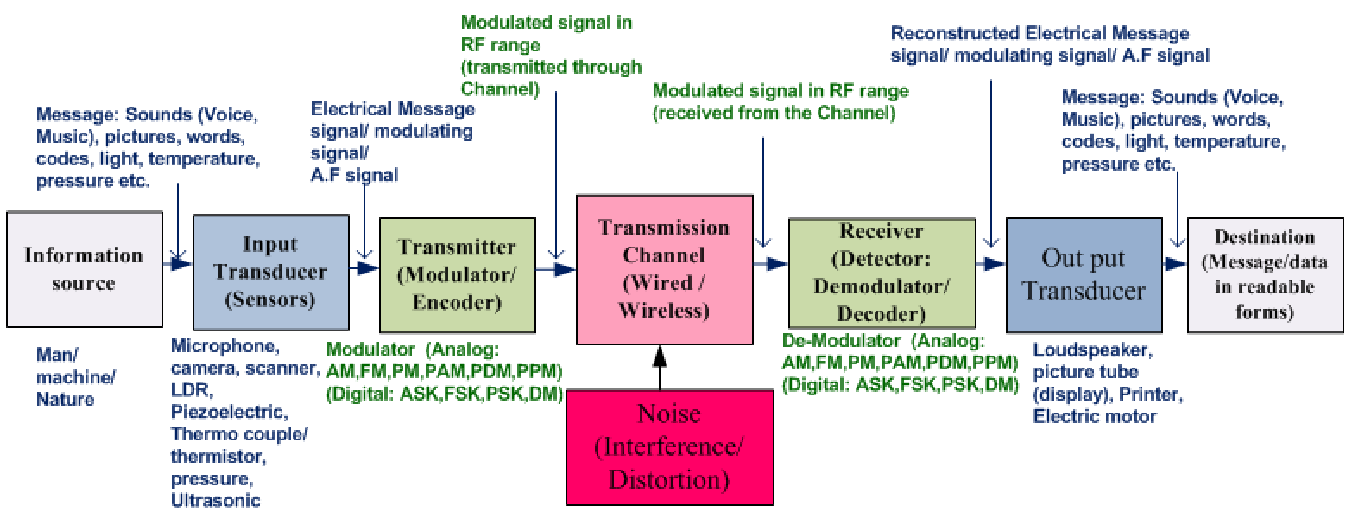


Figure: Basic Model of a Communication System

Information source

The message produced by the information source is not an electrical in nature. But it may be Sounds (Voice, Music), pictures, words, codes, light, temperature, pressure etc. So we need a transducer, which converts the original physical message into a time varying electrical signal (these signals are called baseband signals or message signal or modulating signal or Audio frequency signal) Similarly at the destination, another transducer is used to convert electrical signal into appropriate message.

<u>Transmitter or sender</u>

Its sends or transmits the information. Hear the message signal is converted into suitable form for the propagation over the communication medium, called modulation or encoding i.e. super imposing (placing) low frequency (A.F) message signal with high frequency (R.F) carrier signal. This is done by modulator circuits. [Since the message signal is low frequency and weak in nature, it cannot be transmitted over longer distance directly]

The output of the modulator circuit is called as modulated signal it can be transmitted any distance. For example TV broadcasting stations or Radio broadcasting (Sound) stations.

[Note: Explain need for modulation. It is discussed in the next topic]

<u>Receiver</u>

It receives information from the modulated signal available at the transmission channel; the main function of the receiver is to extract the original message signal from the degraded version of the transmitted signal (from the channel). During transmission, the transmitted signal is added with noise (disturbance)

Example: TV sets and Radio set are example for receivers.

<u>Transmission channel (medium)</u>

It is the path through which the signal propagates from transmitter to receiver. It may be a pair of wires, a coaxial cable, microwave links or radio waves or space. Every channel introduces some amount of Transmission loss or attenuation, so the signal power progressively decreases with increasing distance. This is reason that the transmitter signal is degraded.

Losses caused by

- Noise
- Electrical interference
- Distortion due to non-linearity
- Electromagnetic discharges such as lightning, power line discharge and etc.

<u>Need for Modulation</u>

The primary purpose of modulation in a communication system is to generate a modulated signal which is well suited to the characteristics of transmission medium.

The audio frequency message signal carrying information can't be transmitted for distance directly without modulation.

According to inverse square law, the fading of the signal directly proportional to the square of the distance it travels. So the signal strength decreases naturally. The attenuation of audio signal is more.

- A high audio frequency of 15 KHz may needs quarter wave antenna of dimension 5000m. It is impractical
- Audio signals (if transmitted directly) from various transmitters. May mix up inseparably (it is very difficult to separate ne from other. Hence to separate them, it is necessary to translate them all in to different portions of electromagnetic spectrum, employing modulation

- Modulation makes the receiver design simple.
- Without modulation wireless communication is impossible.

Simply we can easily conclude as,
1. To reduce the antenna height
2. To overcome hardware system limitations
3. To reduce the interference, noise & distortions made when we transmit the signals with nearly same frequency in the audio frequency range (20-20k) Hz.
4. To multiplex the more number of signals
5. To the assignment of channel frequency
6. To narrow banding the signal
7. To reduce the complexity of the transmission system
8. To increase the bandwidth of the signal

1. Reduction in height of the antenna

When free space is used as a communication media, messages are transmitted and received with the help of antennas.

For efficient and easy transmission/ reception of unmodulated signals, the transmitting and receiving antenna height should be more. In order to reduce height of the antenna, signal must be modulated.

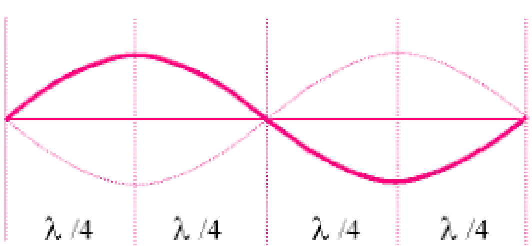
For example: In broadcast systems the maximum audio frequency transmitter from a radio station is 5 KHz. If the signals are to be transmitted without modulation, the size of antenna needed for an effective radiation would be of the order of the half of the wavelength, given as

Height of the antenna H = $\frac{\lambda}{2}$ or = $\frac{\lambda}{4}$

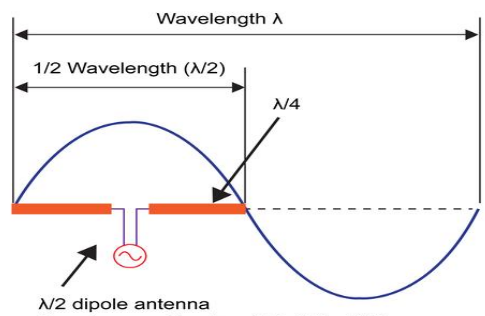
Where, $\boldsymbol{\lambda}$ - wavelength of the signal to be transmitted.

$\boldsymbol{\lambda} = \frac{c}{f}$, $\boldsymbol{f}$ - frequency of the signal to be transmitted & "c" – Velocity (speed) of light in space = $3 * 10^8$ m/s

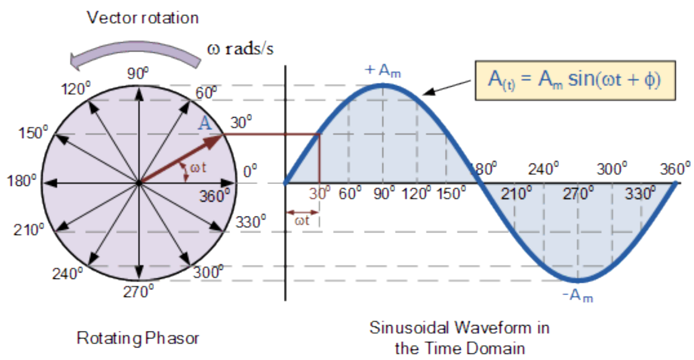
Note: To understand the relationship between frequency and wavelength of the signal to be transmitted with velocity of light, refer the figures given below.

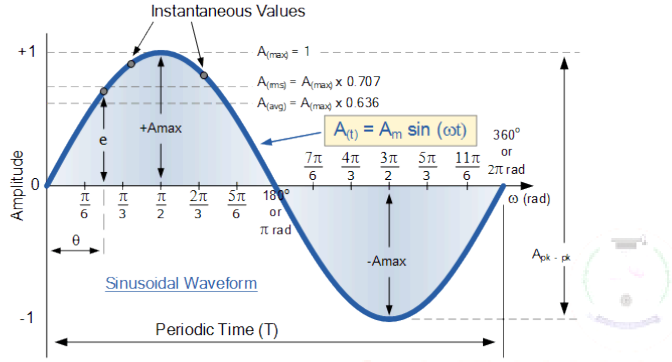


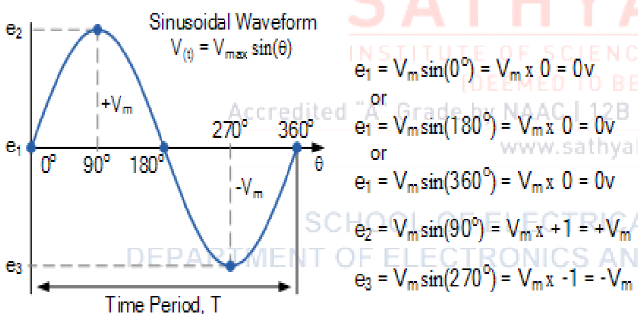
The distance between a node and the antinode next to it is λ/4.



λ/2 dipole antenna
An antenna with a length half that if the wave length resonates.







$$v = f\lambda \quad \text{or} \quad c = f\lambda$$

$$3\times10^8 ms^{-1} = \left(900\times10^6 Hz\right)\times\lambda$$

$$\lambda = \frac{3\times10^8}{900\times10^6} = \frac{3\times10^8}{9\times10^8}$$

$$= 0.333m \approx 0.33m$$

Height of the antenna $=\frac{\lambda}{2}$, $\lambda = \frac{c}{f}$, f- frequency of the signal to be transmitted that is 5 KHz

"**C**" – Velocity (speed) of light in space = 3* 10^8 m/s

Height of the antenna $\frac{\lambda}{2} = \frac{c}{2f} = \frac{3*10^8}{2*5*10^3} = \mathbf{0.3*10^5 = 30,000m = 30km}$

The vertical antenna of such a height can't be imagined or used practically. It is impossible to install.

So the height of the antenna must be reduced, this is possible by modulation process. After modulation process, low frequency of the signal to be transmitted is shifted in to high frequency of carrier signal (A.F into R.F)

Assume 5 KHz (low frequency of the signal) is shifted to 10 MHz (high frequency of carrier signal).

Now consider a modulated signal f=10 MHz . The minimum antenna height is given by

$$\frac{\lambda}{2} = \frac{c}{2f} = \frac{3*10^8}{2*10*10^6} = 0.15*10^2 m = 15m$$

This antenna height can be practically achieved.

Example 2

In broadcast systems the maximum audio frequency transmitter from a **radio station is 1 KHz.**

If the signals are to be transmitted without modulation, the size of antenna needed for an effective radiation would be of the order of the half of the wavelength, given as

$$\frac{\lambda}{2} = \frac{c}{2f} = \frac{3*10^8}{2*1*10^3} = 1.5*10^5 = 150,000m = 150km$$

The antenna of this height is practically impossible to install

Now consider a **modulated signal f=1MHz** . The minimum antenna height is given by

$$\frac{\lambda}{2} = \frac{c}{2f} = \frac{3*10^8}{2*1*10^6} = 1.5*10^2 m = 150m$$

This antenna height can be practically achieved.

2. **To overcome hardware system limitations**

Occasionally in signal processing applications, the frequency of the message signal to be processed does not match with the frequency range of the processing apparatus (like filters and amplifiers etc.).

Through the modulation process (frequency translation) the frequency of the message signal is brought into the equipment acceptable range of frequencies by selecting appropriate carrier signal.

3. **To reduce noise and interference**

If the transmission is carried at audio frequencies i.e. when the (Audio message frequency - AF) signals are transmitted directly, then **all the signals range from 20-20kHz from the different sources will get mixed up in the air with one another and it will not be possible to separate them,** so that we are going for modulation, by (after) modulation, signal frequencies are translated. (i.e.) the bandwidth of the translated signal is larger than that of the message signal. Thereby the effect of noise and interference is removed. For example: **Frequency modulation** and certain other types of **modulation have the property of suppressing both noise and interference.**

4. **Modulation for Multiplexing**

Multiplexing is the process of combining several message signals for simultaneous transmission over a common communication channel without any crosstalk or interference.

For sending each signal we need separate channel. Practically impossible (would create interference) for transmission and reception of multiple signals over a common communication channel without modulation. By doing modulation, different message signals are translated into different spectral locations, enabling the receiver to select the desired signal without cross talk or interference.

Ex: FDM (Channel bandwidth is shared by ‘n’ of signals but common time) and TDM (time is shared by ‘n’ of signals but common frequency – turns transmitting). Examples: FM or AM sound broadcasting, TV broadcasting, Data Telemetry and etc.

5. **For channel Frequency assignment**

Modulation allows several radio or television transmission stations to broadcast their programs simultaneously on different carrier frequencies and also it allows different receivers to be tuned to select different stations.

For example, when you tune a radio or television set to a particular station, i.e. you are selecting one of many signals from the common channel (free space) (Like from Chennai TV station or Trichy or Nagercoil station) being received at that same time. **Since each station has a different assigned carrier frequency**, and the voice signal (desired signal) can be separated from the others by filtering. This is possible only by modulation.

6. **Narrow banding**

Let us assume that the baseband signal in the broadcast system is radiated directly frequency range extending from 50Hz to 10 KHz. The ratio of the highest to lowest wavelength is 200. If an antenna is designed for 50Hz, it will be too large, for 10 KHz and vice versa. Hence we require a wideband antenna which can operate for band edge ratio of 200Hz which is practically impossible. however, if the audio signal if modulated or translated to Radio range of **1MHz** then the ratio of lowest to highest frequency will be which approximately unity, and the same antenna will be suitable for the band extending from $(\mathbf{10^6+50})\ \boldsymbol{to}\ (\mathbf{10^6+10^4})$. Thus modulation converts a wideband signal to a narrow band. This is called narrow banding.

7. **Increases the range of communication**

The modulation process increases the frequency of the signal to be transmitted. Hence increases the range of communication.

8. **Easy of radiation and Adjustment of bandwidth**

As two signals are translated into higher frequencies, it becomes relatively easier to design amplifier circuits as well as antenna systems at these increased frequencies.

Bandwidth of a modulated signal maybe made smaller or larger than the original signal

Signal to noise ratio in the receiver which is the function of the signal bandwidth can thus be improved by proper control of bandwidth at the modulating stage.

Types of Communication

Broadcast vs. Point-to-point

Broadcast: A method of sending a signal where multiple parties may hear a single sender. Radio stations are a good example of everyday life "Broadcast Network". In this example, you can see a single station is broadcasting a message to multiple locations that may or may not be able to hear it, and if they are able to hear it, may choose to listen or not.

Point-to-point: A method of communication where one "point" (person or entity) speaks to another entity.

Classification Based on Direction of Communication

Based on whether the system communicates only in one direction or otherwise, the communication systems are classified as under:

1. Simplex System
2. Half duplex System
3. Full duplex System

Simplex System

In these systems, the information is communicated in only one direction. For example, the radio or TV broadcasting system can only transmit, they cannot receive. Another example of simplex communication is the information transmitted by the telemetry system of a satellite to earth. The telemetry system transmits information about the physical status of the satellite such as its position or temperature.

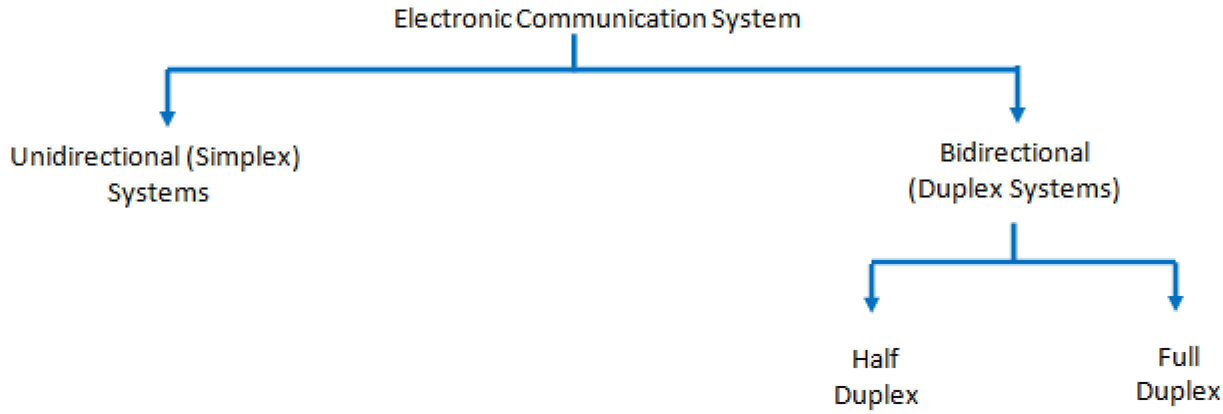
Half duplex System

These systems are bidirectional, i.e. they can transmit as well as receive but not simultaneously. At a time, these systems can either transmit or receive, for example, a transceiver or walky talky set. The direction of communication alternates. The radio communications such as those in military, fire fighting, citizen band (CB) and amateur radio are half duplex system.

Full duplex System

These are truly bidirectional systems as they allow the communication to take place in both the directions simultaneously. These systems can transmit as well as receive simultaneously. For example: the telephone systems. However, the bulk of electronic communications is two-way. The best example of full duplex communication system is telephone system. Classification of communication is shown in the figure below.

Fundamental limitations in a communication system



There are two major constraints when designing a communication system

1. Technological constraints
   - Hardware availability
   - Interaction and compatibility with existing systems
   - Economic factors
   - International and national regulating norms

These limitations can be solved theoretically

2. Fundamental physical limitations
   - Band width
   - Channel capacity or Information carrying capacity
   - Noise

These are the major significant limitations on the performance of a communication system

Bandwidth (signal bandwidth and channel bandwidth):
Bandwidth of an information signal is the difference between highest and lowest frequencies in the information.

Channel bandwidth: It is the difference between the highest and lowest frequencies, that the channel will allow passing through it (Pass band.

Channel band width must be greater than or equal to the bandwidth of the information.
Calculation of signal Bandwidth: For example for voice message signals,
The voice frequency ranges from 300Hz - 3400 Hz
In speech processing applications, (like telephone speech)
Low-frequency information below 300 Hz and high-frequency information above 3400 Hz is mostly severely distorted. Thus information (sense/meaning is available from300Hz to 3400HZ)
Voice/ speech Signal Bandwidth = H.F – L.F = 3400 Hz – 300 Hz = 3100Hz

Therefore channel must have a bandwidth equal or greater than 3100 Hz

If it is not possible to have sufficient bandwidth, signaling speed (transmission speed) will decrease, and as a result transmission time will increase.
The information-carrying capacity/channel capacity:
It represents the number of independent symbols that can be carried out through a system in a given unit of time. i.e, bits per second (bps).
According to Hartley law
Information capacity (I) = B * T
Where B - Bandwidth, T - Time taken for transmission
According to Shannon's law
Information capacity I = B log2 (1+ S/N)
Where B - Bandwidth, S/N - the ratio between Wanted signal power to noise (unwanted signal power)
Noise (unwanted message): Unwanted electrical signal that accompanies the message signal is referred to as noise.
Noise sources: (Internal or External)
1. Inadequate power supply
2. Switching transients
3. Internal circuit noise
4. Atmospheric disturbances
5. Extra-Terrestrial radiation etc.
We measure noise relative to an information signal in terms of signal to noise power ratio denoted by S/N.
For a good system
Signal (wanted) power to Noise (unwanted) power ratio at input and output are the same.

IEEE standard Electromagnetic (Radio) Frequency Spectrum allocated for various communication systems

AUDIO FREQUENCY SIGNALS (20HZ to 20 KHz): *A normal human can hear sound vibrations in the range of 20 Hz to 20 KHz. Signals that create such audible vibrations qualify as an audio signal.*
Human hearing and voice

Range is about 20 Hz to 20 kHz, most sensitive at 2 to 4 KHz.

Dynamic range (quietest to loudest) is about 96 dB

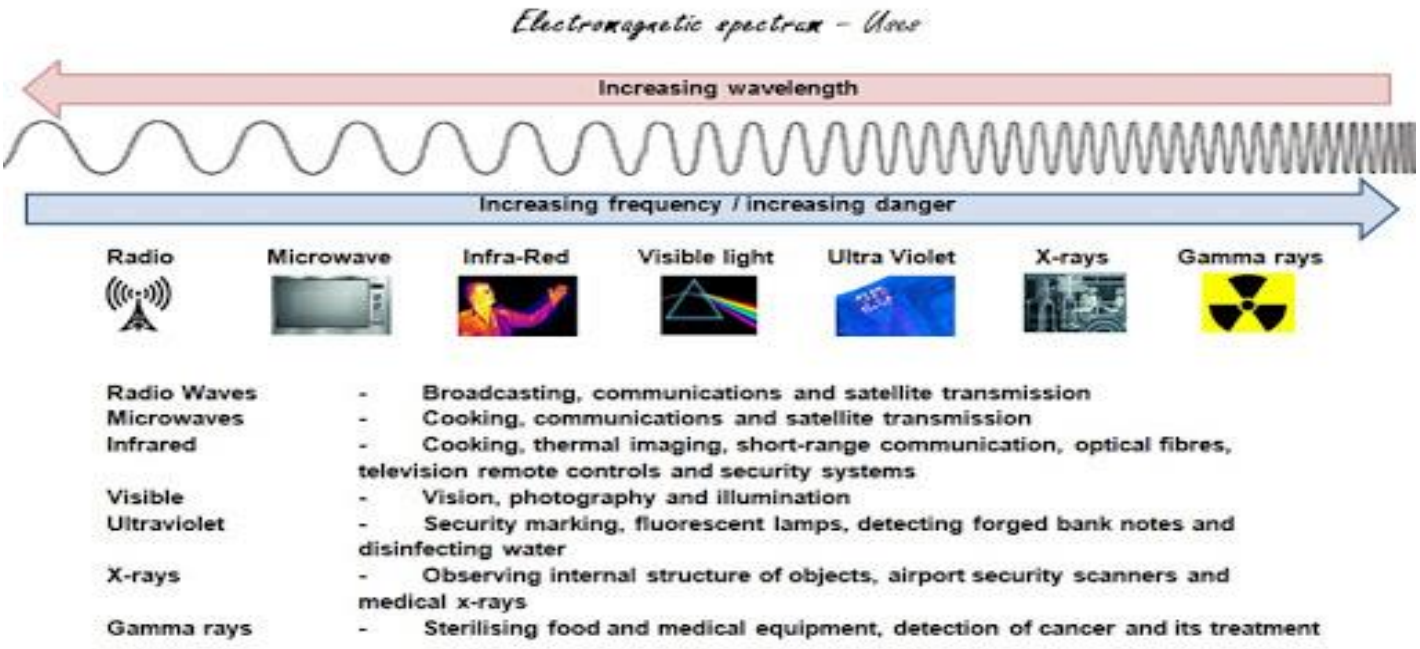
Normal voice range is about 500 Hz to 2 kHz

- Low frequencies are vowels and bass
- High frequencies are consonants

Generalized frequency ranges: (A.F)

Descriptive Unit	Sub-Woofer	Bass	Mid-Range	Tweeter
Frequency Range	10Hz to 100Hz	20Hz to 3KHz	1KHz to 10KHz	3KHz to 30KHz

Electromagnetic waves in this frequency range are called radio frequency bands or simply ‘radio waves’. All known transmission systems are operated in the RF spectrum range including analogue radio, aircraft navigation, marine radio, amateur radio, TV broadcasting, mobile networks and satellite systems.



IEEE standard Electromagnetic (Radio) Frequency Spectrum allocated for various communication systems

Band name	Abbreviation	ITU band number	Frequency	Wave length	Example Uses
Extremely low frequency	ELF	1	3 –30 Hz	100,000 – 10,000 km	Communication with submarines
Super low frequency	SLF	2	30 –300 Hz	10,000 – 1,000 km	Communication with submarines
Ultra low frequency	ULF	3	300 –3,000 Hz	1,000 –100 km	Submarine communication, communication within mines
Very low frequency	VLF	4	3 –30 kHz	100 –10 km	Navigation, time signals, submarine communication, wireless heart rate monitors, geophysics
Low frequency	LF	5	30 –300 kHz	10 –1 km	Navigation, time signals, AM long wave broadcasting (Europe and parts of Asia), RFID, amateur radio
Medium frequency	MF	6	300 –3,000 kHz	1,000 –100m	AM (medium-wave) broadcasts, amateur radio, avalanche beacons

High frequency	HF	7	3 –30 MHz	100 –10 m	Shortwave broadcasts, citizens band radio, amateur radio and over-the-horizon aviation communications, RFID, over-the-horizon radar, automatic link establishment (ALE) / near-vertical incidence sky wave (NVIS) radio communications, marine and mobile radio telephony
Very high frequency	VHF	8	30 –300 MHz	10 –1 m	FM, television broadcasts, line-of-sight ground-to-aircraft and aircraft-to-aircraft communications, land mobile and maritime mobile communications, amateur radio, weather radio
Ultra high frequency	UHF	9	300 –3,000 MHz	1– 0.1 m	Television broadcasts, microwave oven, microwave devices/communications, radio astronomy, mobile phones, wireless LAN, Bluetooth, ZigBee, GPS and two-way radios such as land mobile, FRS and GMRS radios, amateur radio, satellite radio, Remote control Systems, ADSB
Super high frequency	SHF	10	3 –30 GHz	100 –10 mm	Radio astronomy, microwave devices/communications, wireless LAN, DSRC, most modern radars, communications satellites, cable and satellite television broadcasting, DBS, amateur radio, satellite radio
Extremely high frequency	EHF	11	30 –300 GHz	10 –1 mm	Radio astronomy, high-frequency microwave radio relay, microwave remote sensing, amateur radio, directed-energy weapon, millimeter wave scanner, wireless LAN (802.11ad)
Terahertz or Tremendously high frequency	THz or THF	12	300 –3,000 GHz	1– 0.1 mm	Experimental medical imaging to replace X-rays, ultrafast molecular dynamics, condensed-matter physics, terahertz time-domain spectroscopy, terahertz computing/communications, remote sensing

Analog and Digital Communications

In analog communication systems, the message signals are transmitted in analog form itself. AM, FM and PM are common analog modulation schemes which uses sinusoidal carrier signal. In pulse modulation systems such as PAM, PWM and PPM, the carrier signal is a pulse train but the message signal is in analog form. Therefore PAM, PWM and PPM are also called as analog modulation schemes. They are generally not used for wireless communications.

In digital communication systems, the analog information is converted to digital binary data (ones and zeros) using analog to digital convertor ICs. Then the binary data is modulated with a sinusoidal carrier and transmitted. Amplitude shift keying (ASK), Frequency shift keying (FSK) and Phase shift keying (PSK) are some digital modulation schemes.

Modulation

Modulation is the process of placing or superimposing the low frequency message over the high frequency carrier signal make it suitable for transmission over long distance.

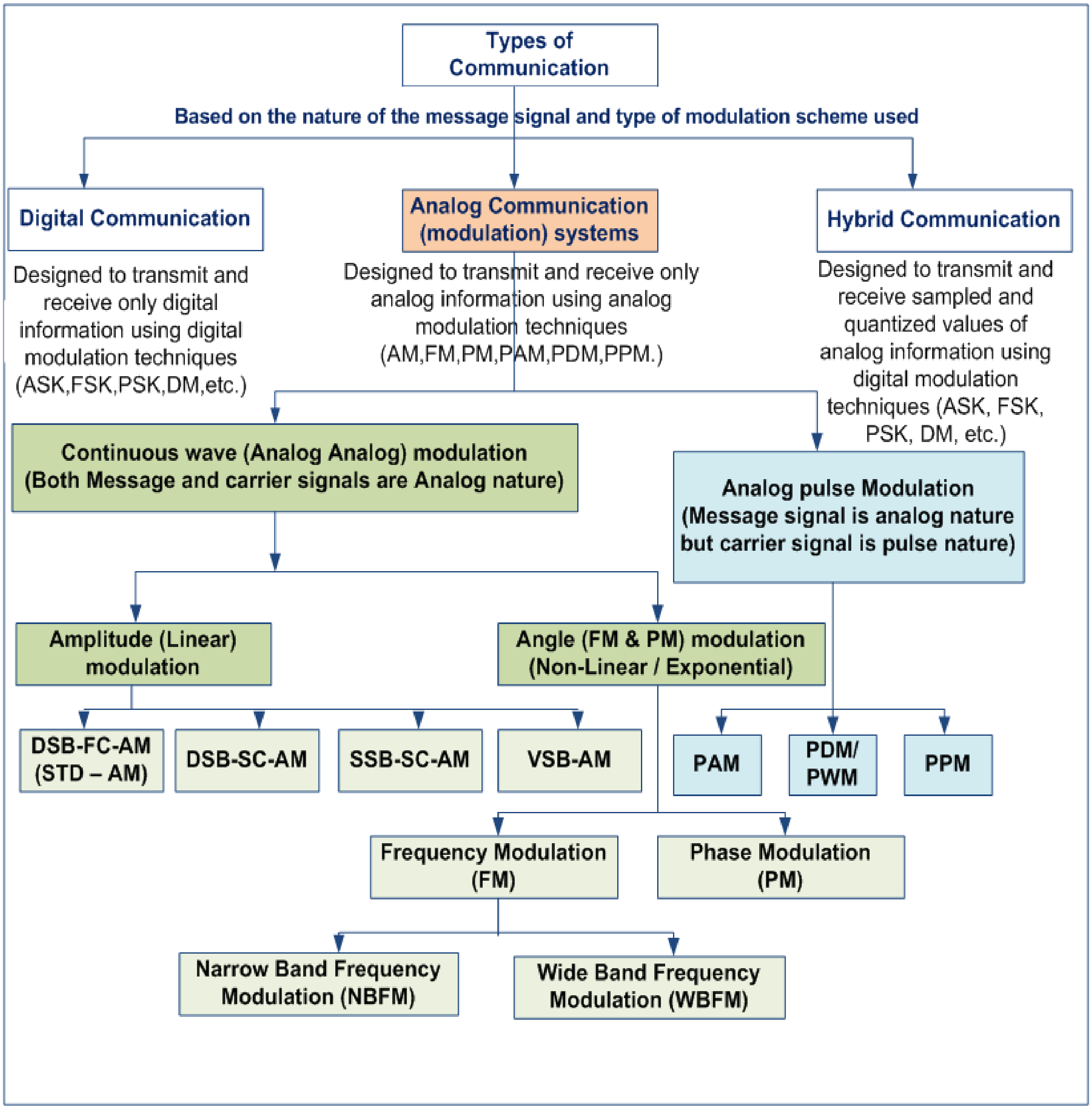
(Or)

Changing the carrier signal with respect to the message signal

Modulation is nothing but a process of changing any one of the characteristics (Amplitude or frequency or phase angle) of (high frequency) carrier signal in accordance with the instantaneous values (amplitude) of the message (information) signal.

Classification of modulation: Modulation can be classified as analog modulation and digital modulation based on the nature of message signal (analog and digital)

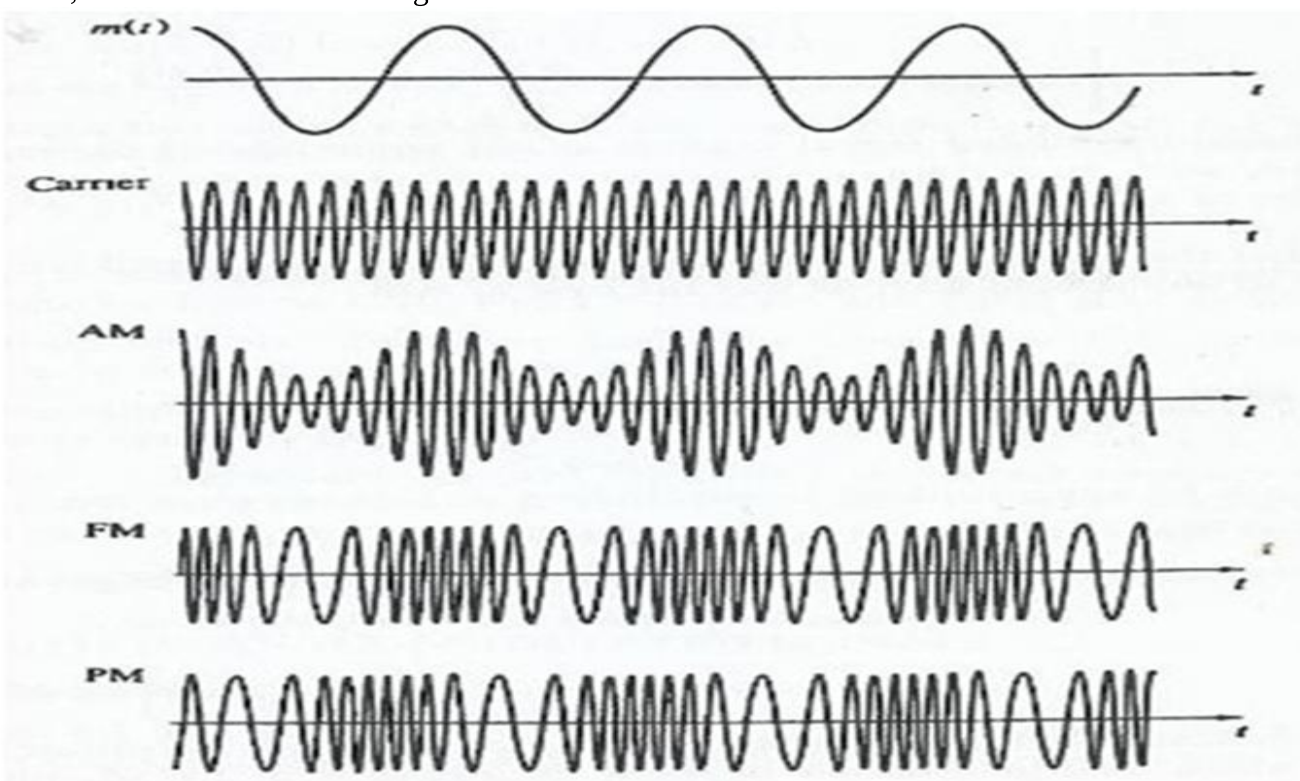
Types of modulation



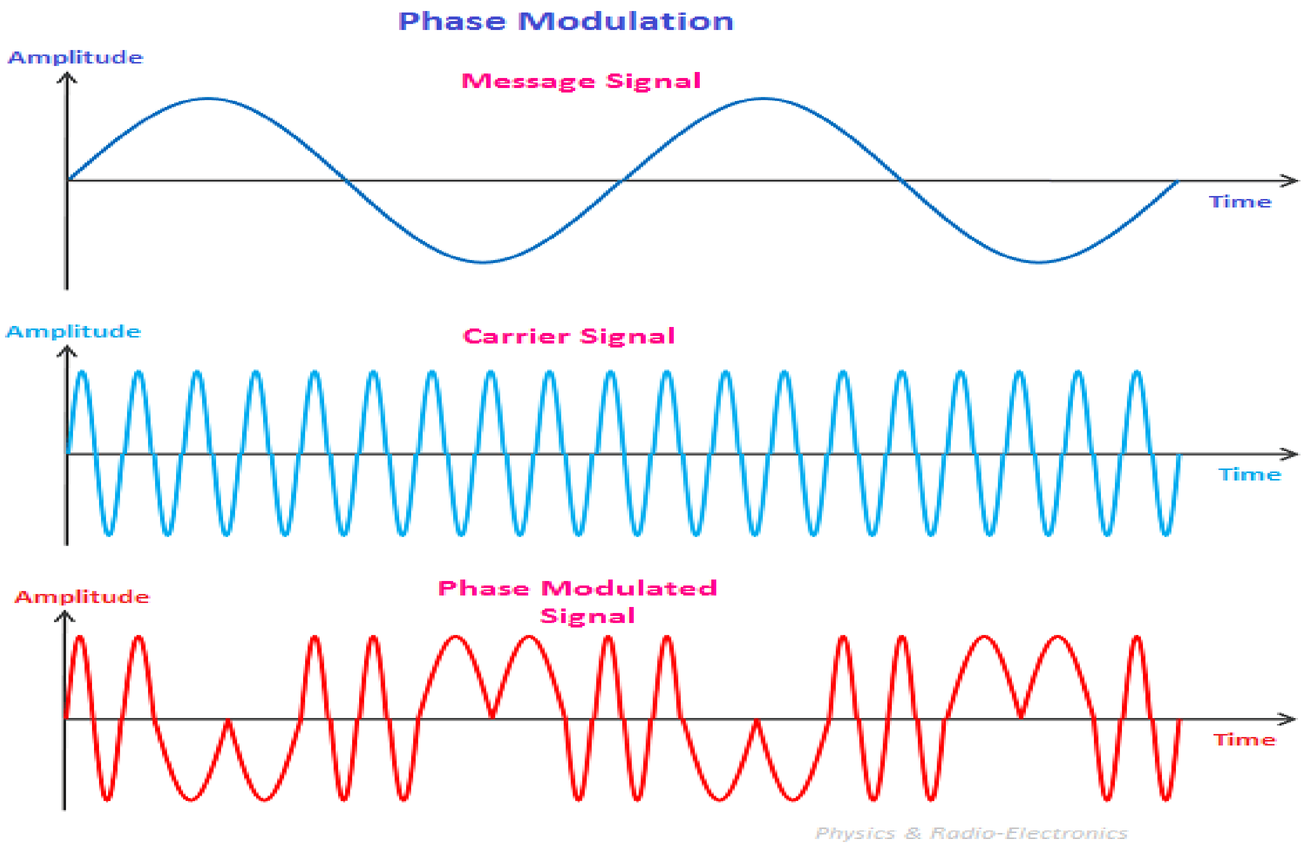
There are various types of modulation techniques used for transmitting information. If the carrier is sinusoidal, then its amplitude, frequency or phase is changed in accordance with the modulating signal to obtain AM, FM or PM respectively. These are continuous wave modulation systems.

Analog modulation can be pulsed modulation as well. Here the carrier is in the form of rectangular pulse. The amplitude, width or position of the carrier pulses is varied in accordance with the instantaneous value of modulating signal to obtain the PAM, PWM or PPM outputs.

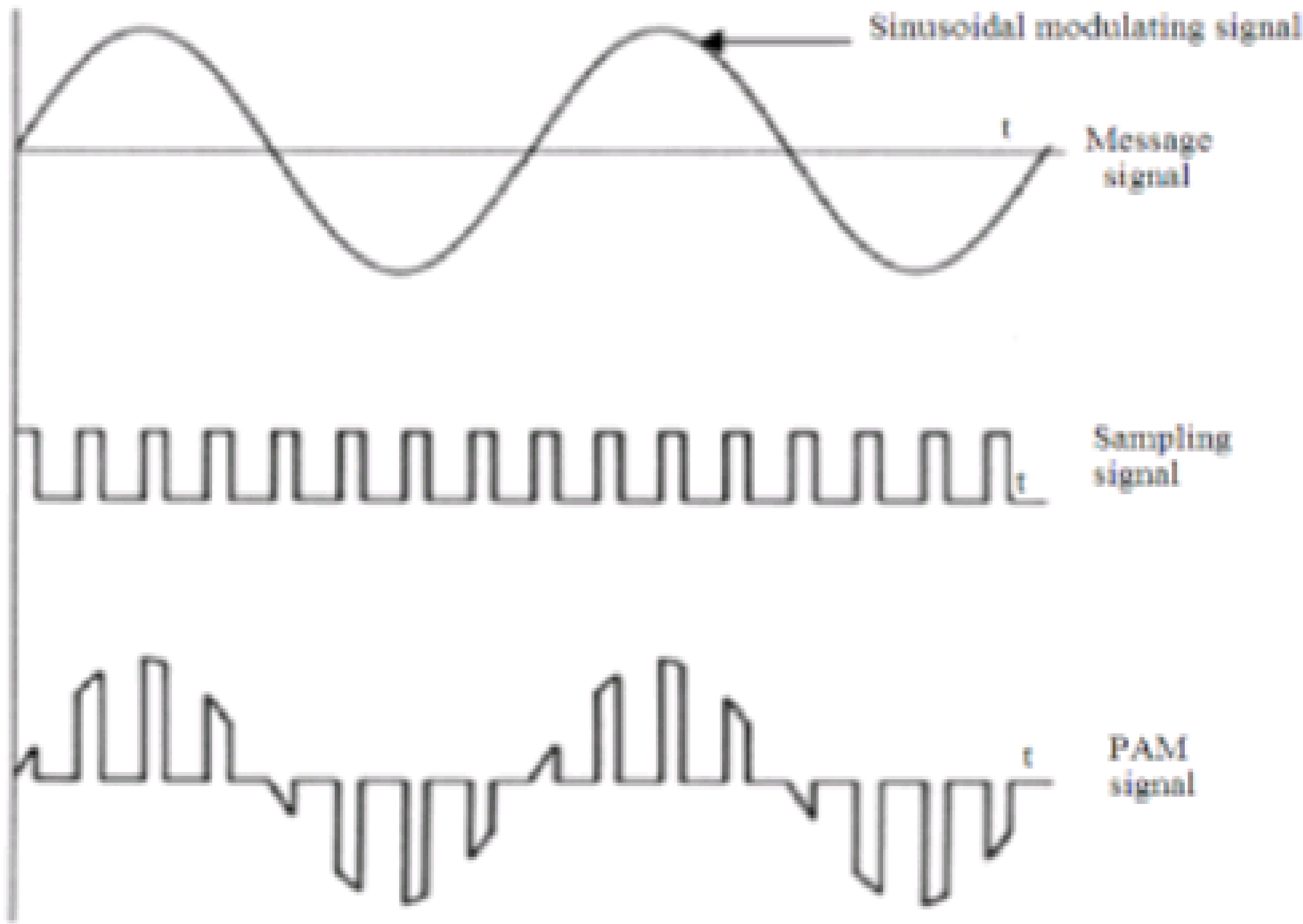
Some commonly used analog modulation techniques are outlined below in figure. AM, FM, PM, PAM, PWM and PPM are analog modulation schemes.

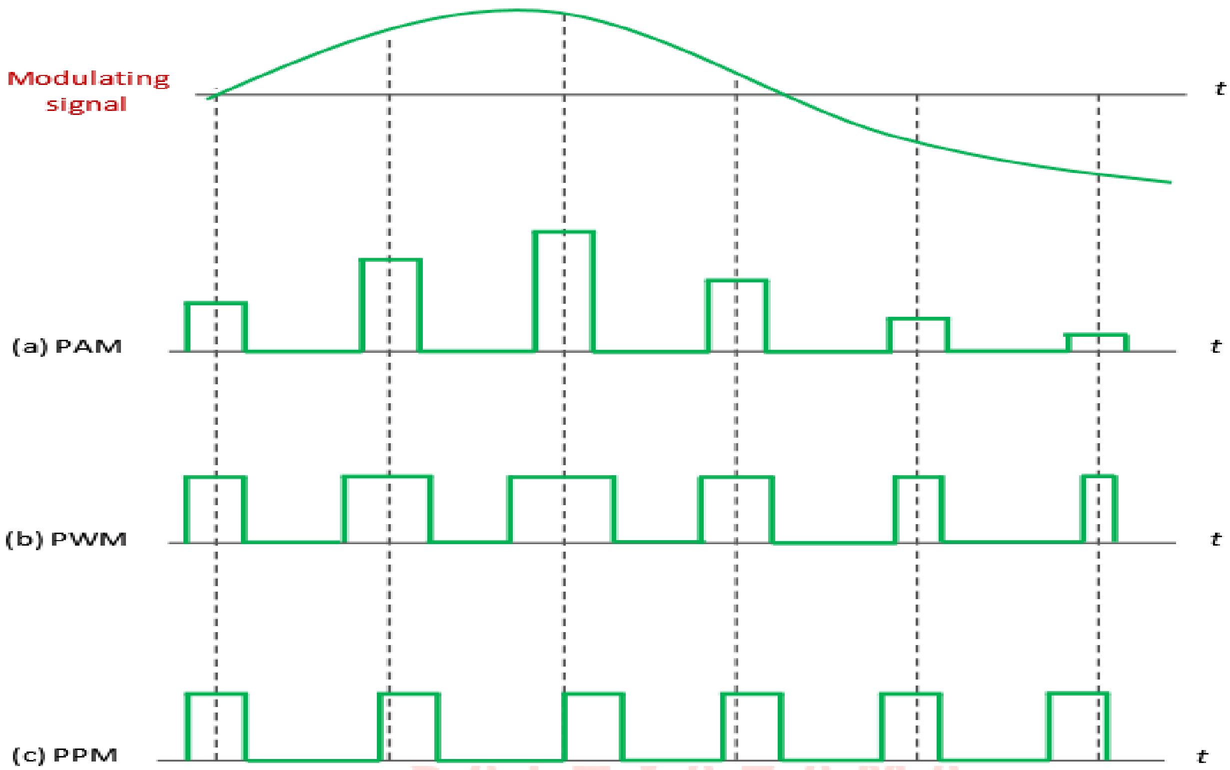


Graphical representation Analog Analog modulation (AM, FM and PM)


Phase Modulation
Amplitude
Message Signal
Time
Amplitude
Carrier Signal
Time
Amplitude
Phase Modulated
Signal
Time
Physics & Radio-Electronics

Analog Pulse Modulation: PAM, PDM and PPM





Types of Analog Modulation

There are two basic types of analog modulation:

Analog Analog modulation (Continuous Wave modulation): Amplitude modulation, Frequency modulation, and Phase modulation. In this case both message signal and carrier signal are analog nature.

Amplitude modulation

Amplitude Modulation is a process of varying amplitude of high frequency carrier signal in accordance with the instantaneous amplitude of the information signal and also the frequency and phase are kept constant.

Frequency modulation

Frequency Modulation is a process of varying **Frequency** of high frequency carrier signal in accordance with the instantaneous amplitude of the information signal and also the amplitude and phase are kept constant.

Phase modulation

Phase Modulation is a process of varying **Phase** of high frequency carrier signal in accordance with the instantaneous amplitude of the information signal and also the amplitude and frequency are kept constant.

Analog Pulse Modulation:

Pulse Amplitude Modulation (PAM): It is a type of analog pulse modulation in which amplitude of the carrier pulse is varied in accordance with instantaneous values of the message signal. But the duration or width and position of the carrier pulse remain constant.

Pulse Duration (Width) Modulation (PDM): It is a type of analog pulse modulation in which duration/ width of the carrier pulse is varied in accordance with instantaneous values of the message signal. But the amplitude and position of the carrier pulse remain constant

Pulse Position Modulation (PPM): It is a type of analog pulse modulation in which position of the carrier pulse is varied in accordance with instantaneous values of the message signal. But the amplitude and duration of the carrier pulse remain constant

<u>**Amplitude Modulation (AM)**</u>

(Draw the graphical representation of standard AM (Double Side Band with Carrier AM), Drive the mathematical expression, frequency spectrum, Band width, modulation index, total power and transmission power efficiency)

<u>Definition</u>:

Amplitude modulation is the process by which amplitude of the carrier signal is varied in accordance with the instantaneous value (amplitude) of the modulating signal, but frequency and phase remains constant.

Amplitude modulation is a relatively inexpensive, low-quality form of modulation that is used for commercial broadcasting of both audio and video signals.

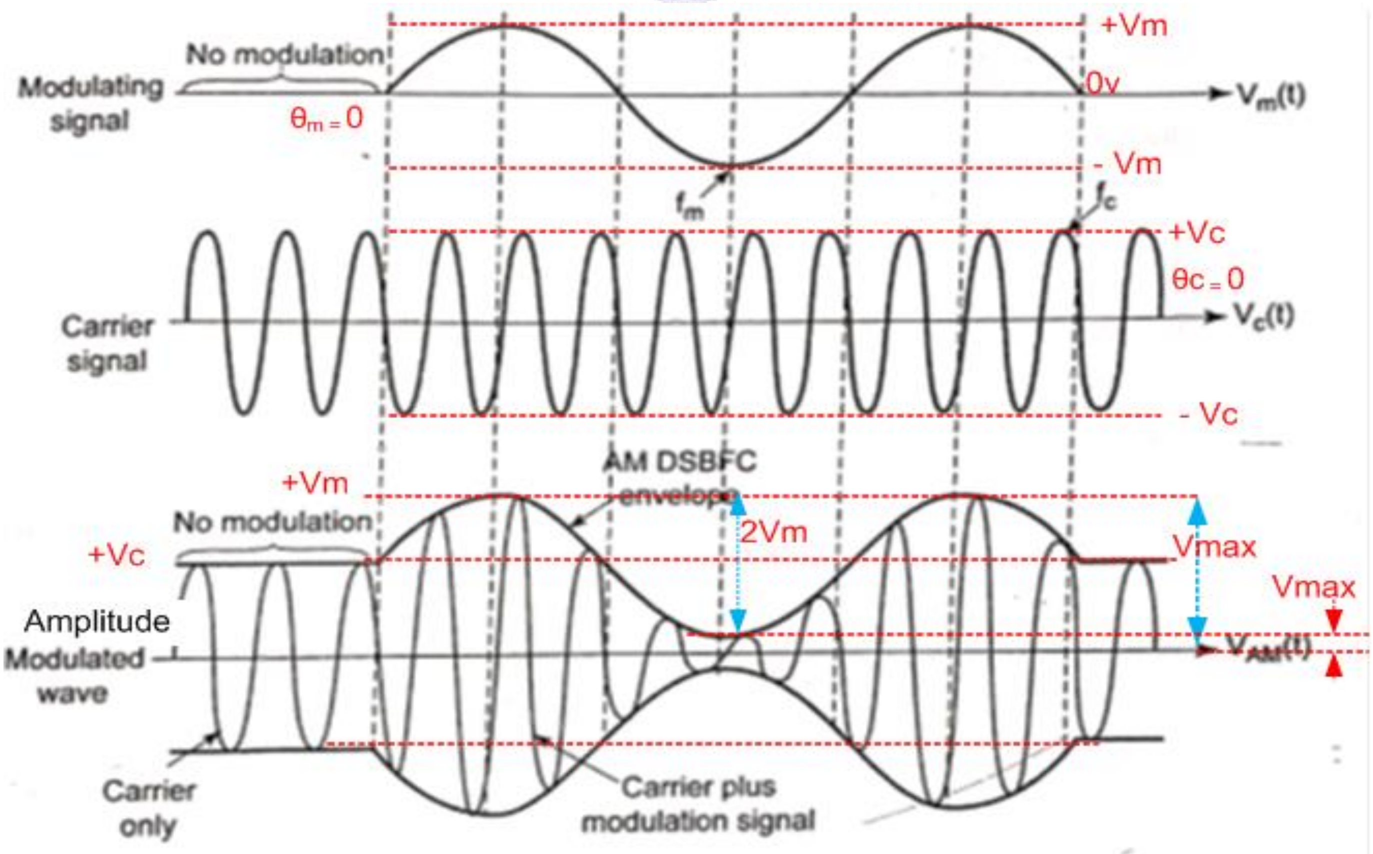


Figure Graphical representation of message signal, carrier signal and Amplitude modulated signal

Mathematical Representation of an AM wave:

Let the modulating (message/information) signal $V_m(t) = V_m \sin(\omega_{m} + \theta_m)t$ ………………….(1)
Where, V_m - (Max.) Amplitude of the modulating signal (volts),
$\omega_m = 2\pi f_m$ - Angular frequency of the modulating signal in radian
(or) f_m frequency of the modulating signal in Hertz i.e., Hz.
θ_m – Initial phase angle of the modulating signal in degree (θ_m is zero degree, it can be ignored)
Similarly
Let the Carrier signal (radio frequency) $V_c(t) = Vc \sin(\omega_c + \theta_c)t$…………………………………(2)
Where, V_c - (Max.) Amplitude of the carrier signal (volts).
$\omega_c = 2\pi f_c$ - Angular frequency of the carrier signal in radian
(or) f_c frequency of the carrier signal in Hertz i.e., Hz.
θ_c – Initial phase angle of the carrier signal in degree (θ_c is zero degree, it can be ignored)

From the graphical representation, we observe that the amplitude of carrier wave "Vc" remains constant (unaffected) when there is no modulation.
According to the definition of AM, the amplitude of the carrier (V_c) is changed with respect to the instantaneous values of the message signal i.e. $V_m(t)$...
Therefore we will get a new mathematical expression for a complete amplitude modulated signal.
$V_{AM}(t) = V_{AM} \operatorname{Sin} \omega_c t$
where V_{AM} is voltage or amplitude of the AM (Amplitude Modulated) signal

During Amplitude modulation, the amplitude of the carrier (V_c) is changed with respect to the instantaneous values (amplitude) of the message signal ($V_m(t)$). Therefore the (new) **amplitude or voltage of the AM signal** is given as $V_{AM} = V_c + V_m(t)$
Where Vc is carrier amplitude. Vm(t) is instantaneous values of the modulating signal.
The amplitude of the carrier signal is changed after modulation.
$_{AM} = V_c + v_m(t)$ …………(3)
Substitute equation (1) in equation (3)

$$V_{AM} = V_c + V_m \sin\omega_m t = V_c [1 + V_m/V_c \sin\omega_m t]$$

Substitute modulation Index (depth of modulation)

$$m_a = V_m / V_c$$

$$V_{AM} = V_c [1 + m_a \sin\omega_m t] \text{ …………………(4)}$$

Hence the resultant AM wave is given by

$$V_{AM}(t) = V_{AM} \sin\omega_c t \text{………………….(5)}$$

Substitute equation (4) in equation (5)

$$V_{AM}(t) = V_c [1 + m_a \sin\omega_m t] \sin\omega_c t$$

Or

$$V_{AM}(t) = V_c[1 + m_a \sin 2\pi f_m t] \sin 2\pi f_c t$$

[reason: $\omega_m = 2\pi f_m$ and $\omega_c = 2\pi f_c$]

Where, ma = Modulation index.
This expression represents the time domain representation of an AM signal.

AM ENVELOPE

- The shape of the modulated wave (AM) is called AM envelope which contains all the frequencies and is used to transfer the information through the system.
- An increase in the modulating signal amplitude causes the amplitude of the carrier to increase.
- Without modulating (i.e. absence of message) signal, the AM output waveform is simply the carrier signal (i.e. no change in the amplitude of the carrier signal)
- The repetition rate of the envelope is equal to the frequency of the modulating signal
- The shape of the envelope is identical to the shape of the modulating signal.

AM Frequency Spectrum and Bandwidth

An **AM modulator is a** nonlinear **device**. Therefore, nonlinear **mixing occurs**, and the output **AM envelope is a complex wave made up of a dc voltage, the carrier frequency, and the sum (f_c+ f_m) and difference (f_c- f_m) frequencies.**

Frequency Spectrum of AM:

After AM (mixing of message signal and carrier signal), the frequency components of the output AM (Amplitude Modulated) signal are determined as follows

The AM wave is given by

V_{AM} (t) = $V_c[1+ m_a \sin\omega_m t] \sin\omega_c t$

= $V_c \sin\omega_c t + m_a V_c \sin\omega_c t \sin\omega_m t$

We know that

$$\sin \omega_c t \sin\omega_m t = \frac{[\cos(\omega_c - \omega_m)t - \cos(\omega_c + \omega_m)t]}{2}$$

$$\text{VAM (t)} = \text{Vc} \sin\omega_c t + \frac{\text{ma Vc}}{2}[\cos(\omega_c - \omega_m)t - \cos(\omega_c + \omega_m)t]$$

$$\text{VAM (t)} = \text{Vc} \sin\omega_c t + \frac{\text{ma Vc}}{2}[\cos(\omega_c - \omega_m)t] - \frac{\text{ma Vc}}{2}\cos(\omega_c + \omega_m)t]$$

Assume $\omega_c t$ is "A "and $\omega_m t$ is "B"

Find out $\sin \omega_c t \sin \omega_m t$ = SIN A SIN B

We know that

2 SIN A SIN B = COS (A-B) − COS (A+B)

SIN A SIN B = $\frac{1}{2}$[COS (A-B) − COS (A+B)]

$\sin \omega_c t \sin \omega_m t = \frac{[\cos(\omega_m - \omega_c)t - \cos(\omega_m + \omega_c)t]}{2}$

COS (A+B) = COS A COS B – SIN A SIN B

COS (A-B) = COS A COS B + SIN A SIN B

SIN (A+B) = SIN A COS B + COS A SIN B

SIN (A-B) = SIN A COS B - COS A SIN B

Thus, Observations made from the AM wave equation are

(The expression for the AM wave shows that it consists of three terms)

(i) First term Vc sinωct represents the carrier signal which is same as the un-modulated carrier signal. (Observed that the carrier signal is present even after modulation process)

(ii) The second term $\frac{\text{ma Vc}}{2}[\cos(\omega_c - \omega_m)t]$ represents amplitude modulated term of Lower Side Band (LSB) with the (new) amplitude of $\frac{\text{ma Vc}}{2}$ at the frequency of ($\omega_c - \omega_m$ or fc −fm).

(iii) The third term $\frac{\text{ma Vc}}{2}[\cos(\omega_c + \omega_m)t]$ represents amplitude modulated term of Upper Side Band (USB) with the (new) amplitude of $\frac{\text{ma Vc}}{2}$ at the frequency of ($\omega_c + \omega_m$ or fc +fm).

The (-) sign associated with the USB represents a phase shift of 180 degree

Mathematical Representation of an AM wave:

Let the modulating signal $V_m(t) = V_m \sin\omega_m t$………………….(1)

Carrier signal $V_c(t)$= Vc $\sin\omega_c t$…………………………………(2)

Where,

V_c — Amplitude of the carrier signal (volts).

V_m — Amplitude of the modulating signal (volts).

The amplitude of the carrier signal is changed after modulation.

$V_{AM} = V_c + v_m\,(t)$ …………(3)

Substitute equation (1) in equation (3)

$V_{AM} = V_c + V_m \sin\omega_m t = V_c[1 + V_m/V_c \sin\omega_m t]$

Substitute modulation Index $m_a = V_m / V_c$

$V_{AM}\,(t) = V_c[1 + m_a \sin\omega_m t]$…………………(4)

Hence AM wave is given by

$V_{AM}\,(t) = V_{AM} \sin\omega_c t$…………………(5)

Substitute equation (4) in equation (5)

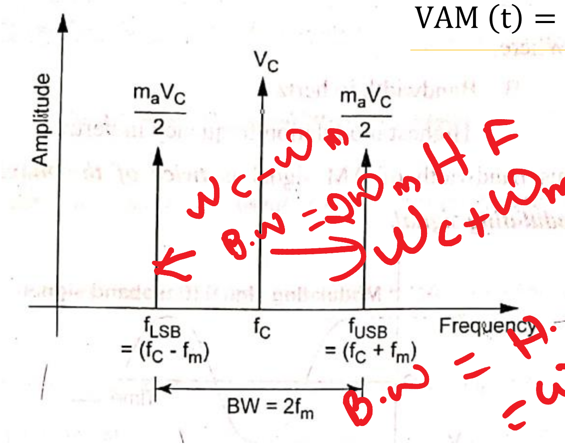
$V_{AM}\,(t) = V_c[1 + m_a \sin\omega_m t] \sin\omega_c t$ (Or) $V_{AM}\,(t) = V_c[1 + m_a \sin 2\pi f_m t] \sin 2\pi f_c t$

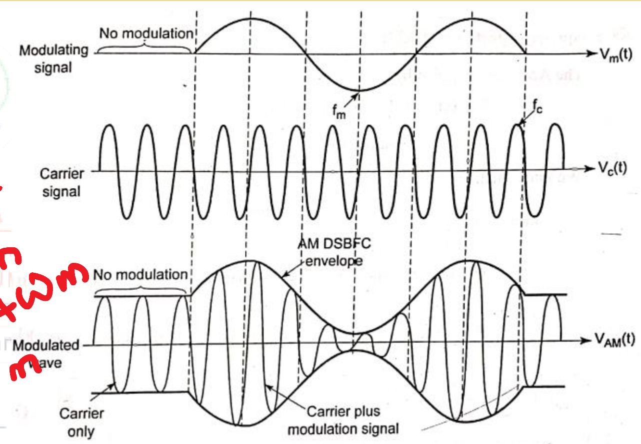
Where, ma = Modulation index.

This expression represents the time domain representation of an AM signal.

Frequency domain representation of AM wave

$$VAM\ (t) = Vc \sin\omega_c t + \frac{ma\ Vc}{2}[\cos(\omega_c - \omega_m)\,t] - \frac{ma\ Vc}{2} cos(\omega_c + \omega_m)\,t]$$





Sidebands:

- Whenever a carrier is modulated by an information signal, new signals at different frequencies are generated as part of the process. These new frequencies are called side frequencies or sidebands.
- The sidebands are occurs in the frequency spectrum directly above and below the carrier frequency.
- Assuming a carrier frequency of f_c and a modulating frequency of f_m. the upper sideband f_{USB} and lower sideband f_{LSB} are computed as follows: $f_{USB}=f_c+f_m$ and $f_{LSB}=f_c-f_m$

Bandwidth of AM:

The bandwidth of the AM signal is given by the subtraction of the highest and the lowest frequency-component in the frequency spectrum.

$B= f_{USB}- f_{LSB}=(f_c+f_m)-(f_c-f_m)=2^* f_m$

Where,

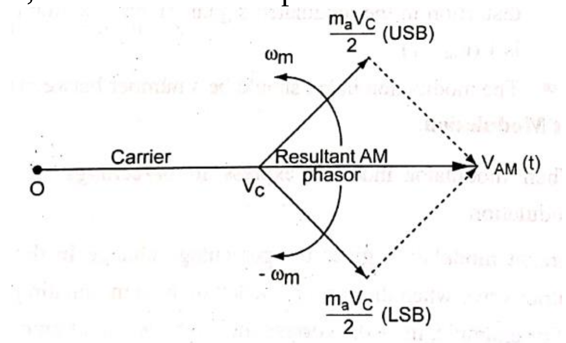
B - Bandwidth in hertz fm — Highest modulation frequency in hertz.

Thus bandwidth of AM signal is twice of the maximum frequency of modulating signal.

Phasor representation of AM

- The amplitude variation in an AM system can be explained with the help of a phasor diagram.
- The phasor for the upper sideband rotates anticlockwise at an angular frequency of ω_m.
- Similarly, the phasor for the lower sideband rotates clockwise at the same angular frequency ω_m.

- The upper side frequency rotates faster than the carrier ($\omega_m > \omega_c$), and the lower side frequency rotates slower ($\omega_m < \omega_c$).
- The resulting amplitude of the modulated wave at any instant is the vector sum of the two sideband phasors.
- Vc is carrier wave phasor, taken as reference phasor and the resulting phasor is $V_{AM(t)}$



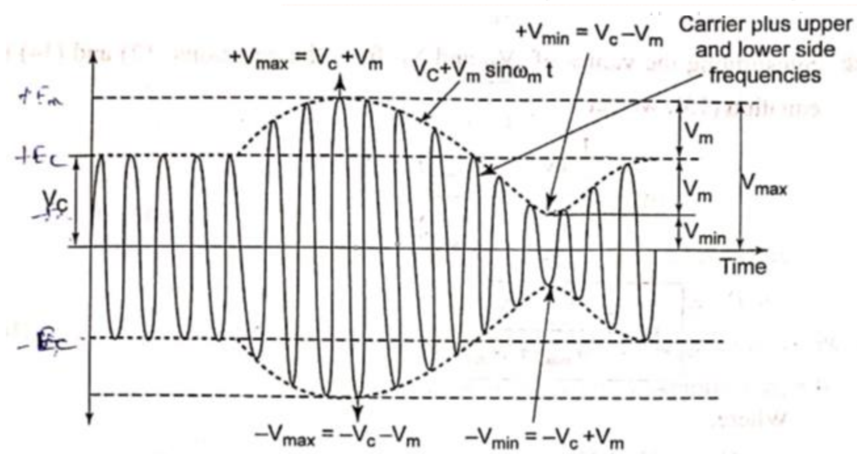
- The phasors for the carrier and the upper and lower side frequencies combine, sometimes in phase (adding) and sometimes out of phase (subtracting).

<u>Modulation Index and Percent Modulation or Coefficient of Modulation</u>

Modulation index is a term used to describe the amount of amplitude change (modulation) present in an AM waveform.

- In AM wave, the modulation index (ma) is defined as the ratio of maximum amplitude of modulating signal to maximum amplitude of carrier signal. $\mathrm{ma} = \frac{\mathrm{Vm}}{Vc}$

$\%\ \mathrm{modulation} = \frac{\mathrm{Vm}}{Vc} * 100$ (or) $\%\ \mathrm{modulation} = \mathrm{ma} * 100$



From the figure, we can write

$V_{max} = V_c + V_m$ & $V_{min} = V_c - V_m$

Then $2V_m = V_{max} - V_{min}$ $\mathrm{Vm} = \frac{\mathrm{Vmax-Vmin}}{2}$......

$V_c = V_{max} - V_m$ .....(2)

Substitute equation 1 in equation 2

$V_c = V_{max} - V_m = Vmax - \frac{\mathrm{Vmax-Vmin}}{2}$

$= Vmax - \frac{\mathrm{Vmax}}{2}\ \frac{\mathrm{Vmax}}{2}\ i.e. Vc = \frac{\mathrm{Vmax+Vmin}}{2}$.

Modulation Index

$$\mathrm{ma} = \frac{\mathrm{Vm}}{Vc} = \frac{\frac{1}{2}(\mathrm{Vmax-Vmin})}{\frac{1}{2}(\mathrm{Vmax+Vmin})}$$

$$= \frac{(\mathrm{Vmax-Vmin})}{(\mathrm{Vmax+Vmin})} \text{........(3)}$$

This is also called time domain representation of AM signal.

Modulation Index

$$\mathrm{ma} = \frac{\mathrm{Vm}}{Vc} = \frac{\frac{1}{2}(\mathrm{Vmax-Vmin})}{\frac{1}{2}(\mathrm{Vmax+Vmin})} = \frac{(\mathrm{Vmax-Vmin})}{(\mathrm{Vmax+Vmin})} \text{........(3)}$$

Where Vmax=Vc+Vm and Vmin=Vc-Vm

The modulation index is a number lying between 0 and l, and it is very often expressed as a percentage and called the percentage modulation.

$$\%\mathrm{ma} = \frac{(\mathrm{Vmax-Vmin})}{(\mathrm{Vmax+Vmin})} * 100 \text{........(4)}$$

Modulation Index for Multiple Modulating Frequencies:

When two or more modulating signals are modulated by a single carrier. Then the modulation index is given by,

$$ma = \sqrt{m1^2 + m2^2 + \cdots}$$

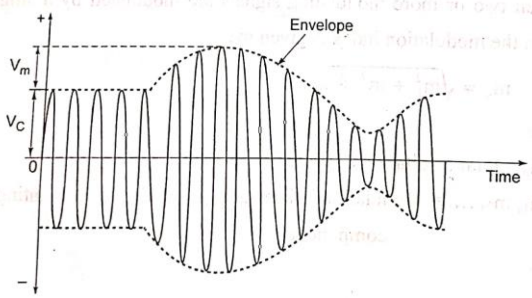
Where, ma = total resultant modulation index
ml, m2, .. = Modulation indices due to individual modulating components.

DEGREE OF MODULATION

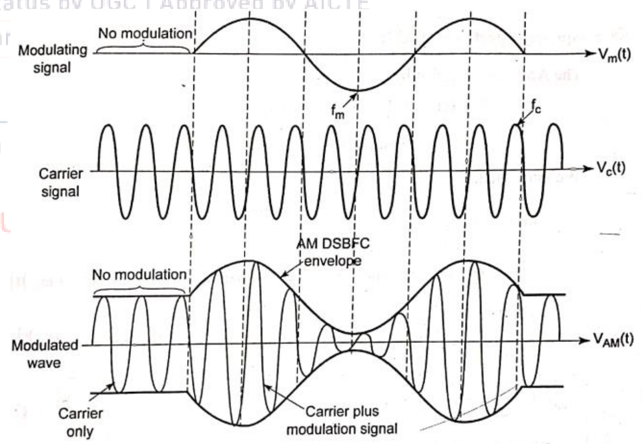
The modulating signals preserved in the envelope of amplitude modulated signal only if $V_m < V_c$ then m_a < l . Where.
V_m = Maximum amplitude of modulating signal.
V_c = Maximum amplitude of carrier signal.
In AM, there are three degrees of modulation are available. It depends upon the amplitude of the modulating signal relative to carrier amplitude. **(V_m<V_c, V_m=V_c and V_m>V_c)**

- Under modulation, **(V_m<V_c,)**
- Critical modulation **(V_m=V_c)**
- Over modulation **(V_m>V_c)**

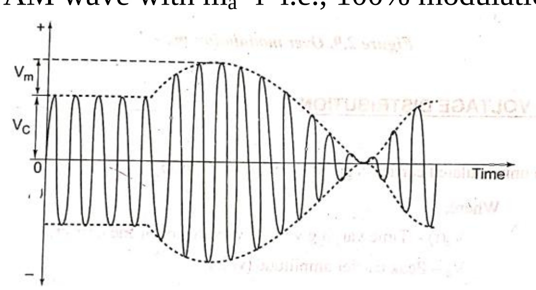
Under Modulation: When $V_m < V_c$ then m_a < l when Here the envelope of amplitude modulated signals does not reach the zero amplitude axis. Hence the message signal is fully preserved in the envelope of the AM wave.
An envelope detector can recover the message signal without any distortion.
AM wave with m_a < l

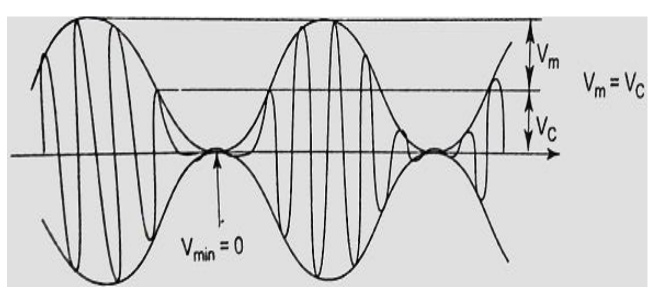


when $V_m < V_c$



Critical Modulation:
m_a = l when $V_m = V_c$
Here the envelope of the modulated signal just reaches the zero amplitude axis. The message signal remains preserved. An envelope detector can recover the message signal without any distortion.
AM wave with m_a=l i.e., 100% modulation $V_m = V_c$



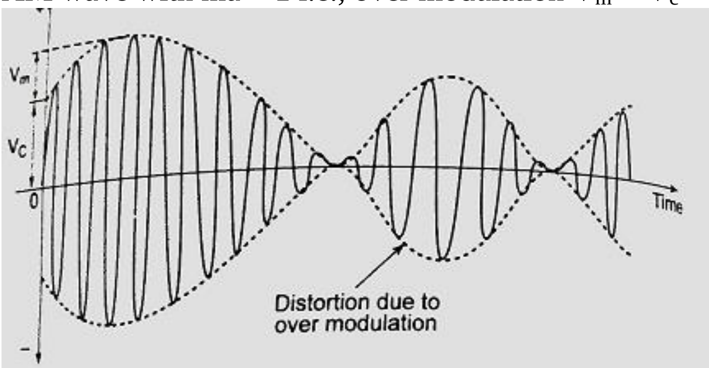


$V_{min} = 0$
$m_a = 1$

Over Modulation: ma > I when $V_m > V_c$

Here both positive and negative extensions of the modulating signals are cancelled (or) clipped out. The envelope of message signal are not same. Due to this envelope detector provides distorted message signal.

AM wave with ma > 1 i.e., over modulation $V_m > V_c$



AM POWER DISTRIBUTION

An AM wave consists of carrier and two sidebands.

The carrier component of the modulated wave has the same amplitude as the unmodulated carrier.

The modulated wave contains extra power in the two sideband components.

The amplitude of the sidebands depends on the modulation index 'ma'. Therefore the total power in the modulated wave will depend on the modulation index also.

The total power in the modulated wave will be

Pt=[carrier power]+ [power in LSB] + [power in USB]

$P_{total} = P_c + P_{LSB} + P_{USB}$

$Ptotal = \frac{(V_c)^2}{R} + \frac{(V_{LSB})^2}{R} + \frac{(V_{USB})^2}{R}$ (rms)

We know that Power = Current x Resistor
$P = V \times I = V \times V/R = V^2/R$
Where all three voltages are in RMS values, and R is the resistance (ex. Antenna resistance), in which the power is dissipated.

Carrier Power (Pc): The average power dissipated in a load by an unmodulated carrier is equal to the RMS carrier voltage squared divided by the load resistance.

We know that RMS value of Vc = $\frac{V_c}{\sqrt{2}}$ $Pc = \frac{(V_c)^2}{R} = \frac{\left(\frac{V_c}{\sqrt{2}}\right)^2}{R} = \frac{(V_c)^2}{2R}$ Where, Pc — Carrier power (watts)

Vc — Peak carrier voltage (volts) R — Load resistance (ohms).

Power in the Sidebands:

The sideband powers are expressed mathematically as

$P_{SB} = P_{USB} + P_{LSB} = \frac{(V_{LSB})^2}{R} + \frac{(V_{USB})^2}{R} = \frac{\left(\frac{m_a V_c}{2}\right)^2}{R} + \frac{\left(\frac{m_a V_c}{2}\right)^2}{R}$ where RMS value of Vc = $\frac{V_c}{\sqrt{2}}$

$$P_{SB} = \frac{\left(m_a \frac{V_c/\sqrt{2}}{2}\right)^2}{R} + \frac{\left(m_a \frac{V_c/\sqrt{2}}{2}\right)^2}{R} \quad Pt_{SB} = \frac{m_a^2 \frac{V_c^2}{8}}{R} + \frac{m_a^2 \frac{V_c^2}{8}}{R} = \frac{m_a^2 V_c^2}{8R} + \frac{m_a^2 V_c^2}{8R} = \frac{2m_a^2 V_c^2}{8R} = \frac{m_a^2 V_c^2}{4R}$$

$$Pt_{SB} = \frac{V_c^2}{2R}\left(\frac{m_a^2}{2}\right) = Pc\left(\frac{m_a^2}{2}\right)$$

Total power in AM wave:

$$P_t = Pc + P_{USB} + P_{LSB} = \frac{(V_c)^2}{2R} + \frac{V_c^2}{2R}\left(\frac{m_a^2}{4}\right) + \frac{V_c^2}{2R}\left(\frac{m_a^2}{4}\right) = Pc + Pc\left(\frac{m_a^2}{4}\right) + Pc\left(\frac{m_a^2}{4}\right)$$

$$= Pc[1 + \left(\frac{m_a^2}{4}\right) + \left(\frac{m_a^2}{4}\right)] = Pc\left[1 + \left(\frac{2m_a^2}{4}\right)\right] \quad Pt = Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]$$

Equation relates the total power in the amplitude modulated wave to the unmodulated carrier power with increases in the value of 'ma', the total power also increases.

If ma = 1 for 100% modulation

Pt = 1.51 Pc. Modulation Index in terms of Pt and Pc: $Pt = Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right] \frac{Pt}{Pc} = \left[1 + \left(\frac{m_a^2}{2}\right)\right]$

$$\frac{m_a^2}{2} = \frac{Pt}{Pc} - 1 \quad m_a = \sqrt{2(\frac{Pt}{Pc} - 1)}$$

<u>Current calculations</u>

$Pt = Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]$, we know that $Pt = I_t^2 R$ and $Pc = I_c^2 R$

Where,

Pt — Total transmit power (watts) Pc — Carrier power (watts) It — Total transmit current (ampere)
Ic — Carrier current (amp) R — Antenna resistance (ohms)

$$\frac{Pt}{Pc} = \frac{I_t^2 R}{I_c^2 R} = \frac{I_t^2}{I_c^2} = 1 + \left(\frac{m_a^2}{2}\right) \quad \frac{I_t}{I_c} = \sqrt{1 + \left(\frac{m_a^2}{2}\right)} \quad I_t = I_c\sqrt{1 + \left(\frac{m_a^2}{2}\right)}$$

<u>Modulation Index in terms of Current:</u>

We know that $\frac{I_t^2}{I_c^2} = 1 + \left(\frac{m_a^2}{2}\right)$ $\quad \frac{m_a^2}{2} = \frac{I_t^2}{I_c^2} - 1 \quad m_a = \sqrt{2[\left(\frac{I_t^2}{I_c^2}\right) - 1]}$

<u>Multitone modulation (Modulation by Several Message Signals)</u>

More than two message signals are used to change the amplitude of the common carrier simultaneously

Example:In TV transmission picture and sound are the two messages used to modulate the single carrier

Likewise assume the first message signal $V_{m1}(t) = V_{m1} \sin\omega_{m1}t$ and second message signal $V_{m2}(t) = V_{m2} \sin\omega_{m2}t$ up to nth message signal $V_{mn}(t) = V_{mn} \sin\omega_{mn}t$

The total power of multi tone amplitude modulated signal is given as follows

We know that $Pt = Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]$ for only one message and carrier

If single carrier is simultaneously amplitude modulated with respect to $V_{m1}(t)$, $V_{m2}(t)$….. $V_{mn}(t)$ there fore the total power is sum of the power contributed by carrier and side bands produced by all the message signals.

Therefore $Pt = Pc\left[1 + \left(\frac{m_{ai}^2}{2}\right)\right]$ $_{i=1,2\ldots n}$

$$Pt = Pc\left[1 + \left(\frac{m_{a1}^2}{2} + \frac{m_{a2}^2}{2} \ldots\ldots \frac{m_{an}^2}{2}\right)\right]$$

$$Pc\left[1 + \left(\frac{m_{ai}^2}{2}\right)\right] = Pc\left[1 + \left(\frac{m_{a1}^2}{2} + \frac{m_{a2}^2}{2} \ldots\ldots \frac{m_{an}^2}{2}\right)\right]$$

then

$$m_{ai}^2 = m_{a1}^2 + m_{a2}^2 + \cdots . m_{an}^2 =$$

$$m_{ai} = \sqrt{m_{a1}^2 + m_{a2}^2 + \cdots m_{an}^2}$$

Transmission Efficiency (%):

The amount of useful message power present in AM wave is expressed by a term called transmission efficiency. The transmission efficiency of an AM wave is the "ratio of the transmitted power which contains the information (i.e., the total sideband power) to the total transmitted power"

Because AM wave expression contains three components such as carrier, USB and LSB. Carrier does not contains any information.

$$\%\eta = \frac{\textit{useful power}}{\textit{Total power}} = \frac{\textit{power in sideband}}{\textit{Total power}} * 100 = \frac{P_{\text{USB}} + P_{\text{LSB}}}{P_{\text{t}}} * 100$$

We know that $Pt_{\text{SB}} = P_{\text{USB}} + P_{\text{LSB}} = Pc\left(\frac{{\text{m}_\text{a}}^2}{2}\right)$ and $Pt = Pc\left[1 + \left(\frac{{\text{m}_\text{a}}^2}{2}\right)\right]$ $\%\eta = \frac{Pc\left(\frac{{\text{m}_\text{a}}^2}{2}\right)}{Pc\left[1+\left(\frac{{\text{m}_\text{a}}^2}{2}\right)\right]} * 100$

$\%\eta = \frac{\left(\frac{{\text{m}_\text{a}}^2}{2}\right)}{\left[1+\left(\frac{{\text{m}_\text{a}}^2}{2}\right)\right]} * 100 = \frac{{\text{m}_\text{a}}^2}{2+{\text{m}_\text{a}}^2} * 100$ If ma = I then %η=1/3*100=33.3%

Only 33.3% of energy is used and remaining power is wasted by the carrier information along with the sidebands.

The maximum transmission efficiency of the AM is 33.3%. This means, that only one-third of the total power is carried by the sidebands and the rest two-third is a waste and is transmitted only for a low cost reception system

Advantages:

- AM has the advantage of being usable with very simple modulators and demodulators.
- AM is a relatively inexpensive.
- AM wave can travel a long distance.

Applications:

- Low quality form of modulation that is used for commercial broadcasting of both audio and video signals
- Two-way mobile radio communications such as citizens band (CB) radio.
- Aircraft communications in the VHF frequency range.

Disadvantages:

- Poor performance in the presence of noise.
- Inefficient use of transmitter power.
- It needs larger bandwidth.

Types of Amplitude Modulation (AM)

1. Double Sideband with carrier (we will call it AM): This is the most widely used type of AM modulation. In fact, all radio channels in the AM band use this type of modulation.
2. Double Sideband Suppressed Carrier (DSBSC): This is the same as the AM modulation above but without the carrier.
3. Single Sideband (SSB): In this modulation, only half of the signal of the DSBSC is used.
4. Vestigial Sideband (VSB): This is a modification of the SSB to ease the generation and reception of the signal.

AM Modulators

- A device that accomplishes modulation, has the capability very one signal(carrier signal) in accordance with the variations of another signal(message signal) or Modulation is performed in a transmitted by a circuit is called modulator
- From the analysis of all types of amplitude modulation, modulated output signal is obtained by combining the low frequency message signal with the high frequency carrier signal.
- The process of amplitude translates the frequency spectrum of modulation of the information signal to produce the amplitude modulated signal, new proposed 100 to be generated. i.e the frequencies of the output signal of a modulator is different from the frequency of the input signal.
- The device that generates an amplitude modulated wave is called AM modulator
- Variety of modulator circuits which employs vacuum tubes or electronic devices(diode, transistor, BJT, FET) to produce amplitude modulated waves
- Depending upon the mode(linear and nonlinear) in which the device is or operated, amplitude modulator is divided into two types
    1. Linear Modulator or large scale modulator
    2. Non linear modulator are small signal modulator

Based on the power level at which modulation is carried out, we have to types

1. Low level modulation: Modulation is carried out the low power level
2. High level modulation: Modulation is carried out at a high power level

Linear modulators or large-scale modulators:

Device having linear V-I characteristics,i.e they are operated in linear region what is transfer characteristics is called linear modulators

They are divided into two types

1. Transistor modulator
2. Switching modulator

Non linear modulator or small signal modulator

They are divided into three types

1. Square Law modulator
2. Product modulator
3. Balanced modulator

Generation of Double side band Full carrier AM

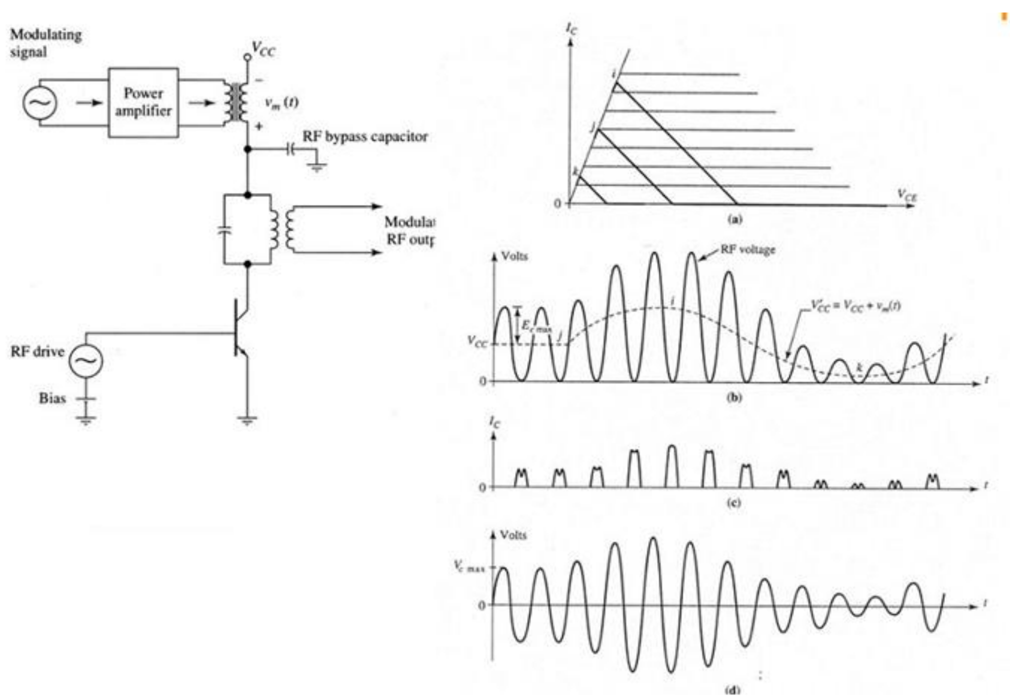
Transistor modulator

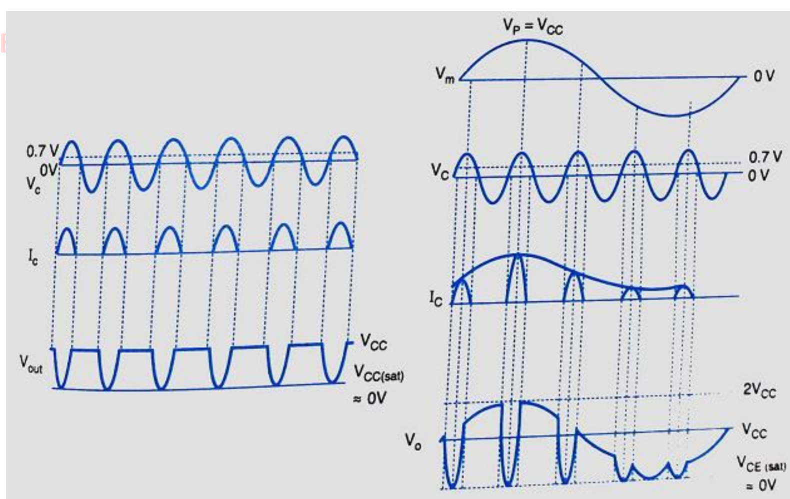
Modulation can be achieved in transistor RF power amplifier stages. The modulating signal can be conveniently supplied on any of the three terminals of the device, emitter base or collector. Accordingly the type of modulator will be called

1. Collector modulator
2. Base modulator
3. Emitter modulator

<u>BJT collector Modulator</u>

The diode modulator circuit doesn't provide amplification and hence it can be used for low power applications. However, amplifying devices like transistors and FET can be provided amplification and it can be used for high power applications each one of them can be used for generation of amplitude modulation by varying their gain parameters in accordance with the modulating signal. A very popular, so cute used for this purpose is the collector modulator





The modulating output can be obtained by making the voltage on the output electrode to vary according to the input modulating signal. The figure shows the collector modulator. The transistor is biased well beyond cut off so that it operates in in class C mode class. The class C mode is used because of its high efficiency. The RF drive is a carrier signal used for AM. The carrier amplitude such that it

drives transistor in conduction over part office cycle. It is applied to the base of the transistor. The modulating signal is passed through the power amplifier and applied to the collector through a low frequency transformer. This voltage is shown as Vm(t) in the figure. This modulating voltage is in series with the supply voltage Vcc.
Hence the collector voltage becomes Vcc=Vcc+Vm(t). The tuned LC circuit associated tuned transformer on the collector receives the AM signal. Because of modulating voltage, the net supply voltage of transistor changes according to the slope variation in Vm(t). Hence the RF carrier signal amplitude is also varied according to the variations in Vm(t). Thus AM signal is produced across the LC circuit at the collector.

<u>Circuit arrangement and operation:</u>
The modulation process starts in the collector, which is the final active stage of the transmitter.

- Carrier signal is applied to the base of the transistor T1 (operated as class C amplifier for higher efficiency)
- Vcc is the dc supply to Collector terminal for biasing.
- Modulating / message signal is directly applied to the collector true that class B amplifier with required amplification, after amplification of message signal it is applied to the collector in series with DC collector supply voltage Vcc.
- hence the Collector voltage becomes Vcc'i.e the Tuned circuit associated with the collector, receives the AM signal RF bypass capacitor prevents the carrier flowing through the output transformer T1

Operation without modulating signal (with reference of collector waveforms without modulating signal amplitude is zero)

- Amptitude of output signal constant which is equal to Vcc, in the absence of modulating signal.
- The reason is when the amplitude of the carrier exceeds the barrier potential(0.7v) of the emitter base Junction the (Vbe>0.7v) transistor T1 terms on and collector current flows which is equal to Vcc.
- When Vbe<0.7 i.e carrier signal voltage drops below 0.7v transistor T1 turns off and no collector current flows.
- Consequently transistor T1 switches between saturation(on) and cut off(off) controlled by carrier signal, electric current flows for less than 180 degree of each carrier cycle class C operation is achieved

- So that each successive cycle of the carrier T1 turns on current to flow producing negative going waveform at the collector

<u>Operation with modulating signal</u>

- When modulating signal appears across the modulation transformer is added with Vcc.
  the net voltage is Vcc+Vm(t)=Vcc' of transistor changes according to the slow variation in Vcc and vm(t)
- This slow variations in Vcc supply voltage changes the amplitude of the carrier voltage at the output of the modulated wave.
- The envelope of the output voltage is identical with the modulating voltage

The amplitude of the modulated voltage is

Vcc'= Vcc+Vm(t)

$$Vcc' = Vcc + Vm(t) = Vcc + Vm \sin \omega_m t = Vcc\left[1 + \frac{V_m}{V_c}(\sin \omega_m t)\right]$$

$$\boldsymbol{Vcc'} = \boldsymbol{Vcc}[\mathbf{1} + \mathbf{m_a}(\mathbf{sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t})]$$

The Instantaneous value of the modulated signal

$$\boldsymbol{Vcc'(t)} = \boldsymbol{Vcc'} = \boldsymbol{Vcc}[\mathbf{1} + \mathbf{m_a}(\mathbf{sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t})]\boldsymbol{sin\omega_c(t)}$$

<u>Power Calculation and efficiency</u>

Output modulated power i.e delivered power depends on the input power by Vcc supply voltage and the power dissipation in the collector current

$$P_{in} = P_{out} + P_d$$

P_{in} =Input power supplied to Collector circuit

P_{out} =Output power delivered to the load

P_d =Power dissipation in the collector circuit

$$P_d = P_{in} - P_{out} = P_{in}\left(1 - \frac{P_{out}}{P_{in}}\right)$$

Collector Efficiency $\eta_c = \frac{P_{out}}{P_{in}}$ then $P_d = P_{in}(1 - \eta_c)$

power input $P_{in} = \frac{1}{2\pi}\int_0^{2\pi} Vc.Ic\,dt = \frac{1}{2\pi}\int_0^{2\pi} Vcc'.Ic.dt$

Where

$$\boldsymbol{Vcc'} = \boldsymbol{Vcc}[\mathbf{1} + \mathbf{m_a}(\mathbf{sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t})]$$

and

$$\boldsymbol{Ic} = \boldsymbol{Vcc}[\mathbf{1} + \mathbf{m_a}(\mathbf{sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t})]$$

then $\mathbf{P_{in}} = \frac{\mathbf{1}}{\mathbf{2\pi}}\int_\mathbf{0}^{\mathbf{2\pi}} \boldsymbol{Vcc}[\mathbf{1} + \mathbf{m_a}(\mathbf{sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t})] * \boldsymbol{Ic}[\mathbf{1} + \mathbf{m_a}(\mathbf{sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t})]\,\boldsymbol{d\omega t}$

$$P_{in} = \frac{Vcc.Ic}{2\pi}\int_0^{2\pi}[1 + {m_a}^2 sin^2\omega_m t + 2m_a \sin \omega_m t]dwt$$

$$\boldsymbol{we\ \ know\ that\ sin^2}\boldsymbol{\omega}_\mathbf{m}\,\mathbf{t} = \frac{\mathbf{1 - cos\,2}\boldsymbol{\omega}_\mathbf{m}\,\mathbf{t}}{\mathbf{2}}$$

$$P_{in} = \frac{Vcc.Ic}{2\pi}\int_0^{2\pi}\left[1 + {m_a}^2\left[\frac{1 - \cos 2\omega_m t}{2}\right] + 2m_a \sin \omega_m t\right]dwt$$

$$\mathbf{P_{in}} = \frac{\mathbf{Vcc.Ic}}{\mathbf{2\pi}}\left[\boldsymbol{\omega}\mathbf{t} + \frac{\mathbf{m_a}^\mathbf{2}}{\mathbf{2}}\left[\boldsymbol{\omega}\mathbf{t} - \frac{\mathbf{sin\,2}\boldsymbol{\omega}_\mathbf{m}\,\mathbf{t}}{\mathbf{2}}\right] - \mathbf{2m_a\,sin}\,\boldsymbol{\omega}_\mathbf{m}\mathbf{t}\right]$$

$$\mathbf{P_{in}} = \frac{\mathbf{Vcc.Ic}}{\mathbf{2\pi}}\left[\mathbf{2\pi} + \frac{\mathbf{m_a}^\mathbf{2}}{\mathbf{2}}[\mathbf{2\pi - 0 - 2m_a + 2m_a}]\right] = \frac{\mathbf{Vcc.Ic}}{\mathbf{2\pi}}[\mathbf{2\pi}][\mathbf{1} + \frac{\mathbf{m_a}^\mathbf{2}}{\mathbf{2}}]$$

$$\mathbf{P_{in}} = \mathbf{Vcc.Ic} + \frac{\mathbf{m_a}^\mathbf{2}}{\mathbf{2}}\mathbf{Vcc.Ic} = \mathbf{Vcc.Ic}\left(\mathbf{1} + \frac{\mathbf{m_a}^\mathbf{2}}{\mathbf{2}}\right) = \boldsymbol{Pcc}\left(\mathbf{1} + \frac{\mathbf{m_a}^\mathbf{2}}{\mathbf{2}}\right)$$

We know that

$$P_d = P_{in}(1 - \eta_c) = Pcc\left(1 + \frac{m_a^2}{2}\right)(1 - \eta_c)$$

and

$$P_{out} = P_{in} * P_c = \eta_c Pcc\left(1 + \frac{m_a^2}{2}\right)$$

Advantages of collector modulator

1. Linearity is usually good.
2. Collector efficiency is high
3. Power output per transistor is usually high

Disadvantages

1. Large modulating power is required, then the modulating amplifier is the high power amplifier
2. Collector saturation prevents 100 under percent modulation from being achieved with just the collector being modulated.

Base modulator

Modulating signal is applied into the base of the transistor to reduce the power level
Common emitter configuration is used and it is biased into class c mode, the resistor R1 and R2 provides potential divider biasing for transistor through Vcc, the resistor RE and capacitor CE acts as temperature stabilization elements.

Operation

The message signal is applied to the base circuits based on the variations (instantaneous value) of its amplitude. Carrier amplitude is varied between cutoff and saturation regions in order to produce the fully modulated output.

The gain of the circuits cannot be maintained as constant over the entire range of its characteristics.
Hence the output is not linearly modulated

Mathematical analysis

Let the carrier signal $\boldsymbol{Vc(t) = Vc \sin \omega_c t}$

Message signal $\boldsymbol{Vm(t) = Vm \sin \omega_m t}$

Thus the total time varying base Bias Voltage is given by

$$\mathbf{V_{bias}}(\boldsymbol{t}) = \boldsymbol{Vcc + Vm(t)}$$

Base Bias Voltage with respect to time is

$$\mathbf{V_{bias}}(\boldsymbol{t}) = \boldsymbol{Vcc + Vm \sin \omega_m t} \ldots\ldots\ldots\ldots\ldots(\mathbf{1})$$

The total base to emitter voltage can be written as

$$\mathbf{V_{be}}(\boldsymbol{t}) = \mathbf{V_{bias}}(\boldsymbol{t}) + \mathbf{V_c}(\boldsymbol{t}) \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$\mathbf{V_{be}}(\boldsymbol{t}) = \boldsymbol{Vcc + Vm \sin \omega_m t + Vc \sin \omega_c t} \ldots\ldots\ldots\ldots(3)$$

The instantaneous value of current coursing through tank circuits at frequency $\boldsymbol{\omega}_c$ is given by where $\mathbf{I}_t = \boldsymbol{K}[\mathbf{V}_{bias}(\boldsymbol{t}) - \mathbf{V}_{be(0)}(\boldsymbol{t})]$

$\mathbf{V}_{be(0)}(\boldsymbol{t}) = \mathrm{V}_{be}$ ***for zero collector*** current or minimum forward voltage

$$\mathbf{i}_t(\boldsymbol{t}) = \sqrt{2}\,\boldsymbol{K}[\mathbf{V}_{bias}(\boldsymbol{t}) - \mathbf{V}_{be(0)}(\boldsymbol{t})]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

Substitute equation 1

$$\mathbf{i}_t(\boldsymbol{t}) = \sqrt{2}\,\boldsymbol{K}[\boldsymbol{Vcc} + \boldsymbol{Vm}\,\mathbf{sin}\,\boldsymbol{\omega}_m\mathbf{t} - \mathbf{V}_{be(0)}(\boldsymbol{t})]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

$$\mathbf{i}_t(\boldsymbol{t}) = \sqrt{2}\,\boldsymbol{K}[\boldsymbol{Vcc} - \mathbf{V}_{be(0)}(\boldsymbol{t}) + \boldsymbol{Vm}\,\mathbf{sin}\,\boldsymbol{\omega}_m\mathbf{t}]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

take out $\boldsymbol{Vcc} - \mathbf{V}_{be(0)}(\boldsymbol{t})$

$$\mathbf{i}_t(\boldsymbol{t}) = \sqrt{2}\,\boldsymbol{K}[\boldsymbol{Vcc} - \mathbf{V}_{be(0)}(\boldsymbol{t})][\mathbf{1} + \frac{\boldsymbol{Vm}\,\mathbf{sin}\,\boldsymbol{\omega}_m\mathbf{t}}{\boldsymbol{Vcc} - \mathbf{V}_{be(0)}(\boldsymbol{t})}]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

$$\mathbf{i}_t(\boldsymbol{t}) = \sqrt{2}\,\boldsymbol{K}[\boldsymbol{Vcc} - \mathbf{V}_{be(0)}(\boldsymbol{t})][\mathbf{1} + \mathbf{m_a}\,\mathbf{sin}\,\boldsymbol{\omega}_m\mathbf{t}]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

$$\mathbf{i}_t(\boldsymbol{t}) = \sqrt{2}\,\mathbf{I}_t[\mathbf{1} + \mathrm{m}_a\,\mathbf{sin}\,\boldsymbol{\omega}_m\mathbf{t}]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

Modulation index

$$\mathrm{m}_a = \frac{\boldsymbol{Vm}}{\boldsymbol{Vcc} - \mathbf{V}_{be(0)}(\boldsymbol{t})}$$

Instantaneous voltage of tank circuit

$$\mathbf{V_0}(\boldsymbol{t}) = \boldsymbol{Z}.\mathbf{i}_t(\boldsymbol{t}) = \boldsymbol{Z}.\sqrt{2}\,\mathbf{I}_t[\mathbf{1} + \mathrm{m}_a\,\mathbf{sin}\,\boldsymbol{\omega}_m\mathbf{t}]\,\mathbf{sin}\,\boldsymbol{\omega}_c\mathbf{t}$$

Z=Impedance of tank circuit

<u>Advantages</u>

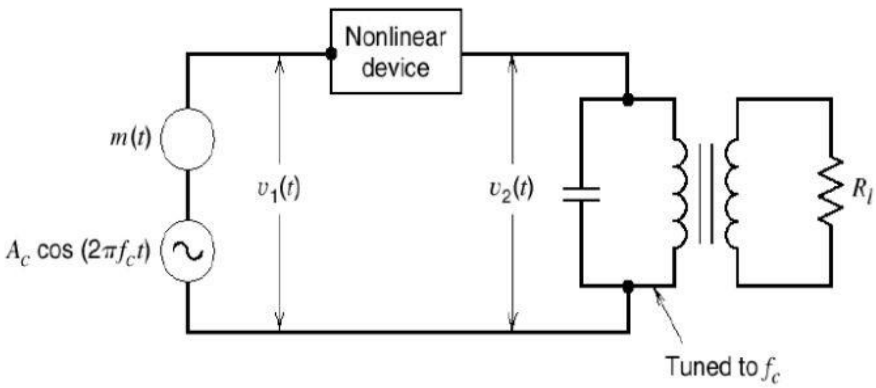
1. The amount of power required for the power supply is low as compared to collector modulation.
2. The power output and efficiency are comparatively low, since the modulated collector current peaks can be only about half as large as in the collector modulated circuit, the power output and efficiency suffer severely.
3. It is used in television transmission because it requires little power and can we power requirement of large bandwidth.
4. The Adjustment of the base modulated amplifier is more critical and the high degree of linearity is more difficult to obtain

<u>Disadvantages</u>

1. Linearity is very poor than collector modulator
2. The efficiency is less than collector modulator

Square Law Modulator

Figure shows the diagram of Square Law Modulator



It consists of the following:

1. A non-linear device
2. A band pass filter
3. A carrier source and modulating signal

The modulating signal and carrier are connected in series with each other and their sum $V_1(t)$ is applied at the input of the non-linear device, such as diode, transistor etc.

Thus,

$V_1(t) = x(t) + V_C cos(2\pi fct)$……………………….(1)

The input output relation for non-linear device is as under :

$V_2(t) = a_1 V_1(t) + b_1 V_1{}^2(t)$…………………………….(2)

where a_1 and b_1 are constants.

Now, substituting the expression (1) in (2), we get

$$V_2(t) = a[x(t) + V_C cos(2\pi fct)] + b[x(t) + V_C cos(2\pi fct)]^2$$

$$V_2(t) = ax(t) + aV_C cos(2\pi fct) + b[x^2(t) + 2x(t)cos(2\pi fct) + V_c{}^2 cos^2(2\pi fct)]$$

The five terms in the expression for $V_2(t)$ are as under $Term\ 1: ax(t): modulating\ \ signal$

$Term\ 2: aV_C cos(2\pi fct): carrier\ signal$

$Term\ 3: bx^2(t): squared\ modulating\ signal$

$Term\ 4: 2bx(t)cos(2\pi fct): AM\ wave\ with\ only\ sidebands$

$Term\ 5: bV_c{}^2 cos^2(2\pi fct): squared\ carrier$

Out of these five terms, terms 2 and 4 are useful whereas the remaining terms are not useful .
Let us club terms 2, 4 and 1, 3, 5 as follows to get ,

$$V_2(t) = ax(t) + bx^2(t) + bV_c^2 cos^2(2\pi fct) + aV_C cos(2\pi fct) + 2bx(t)cos(2\pi fct)$$

$$ax(t) + bx^2(t) + bV_c^2 cos^2(2\pi fct) = Unuseful\ terms$$
$$aV_C cos(2\pi fct) + 2bx(t)cos(2\pi fct) = useful\ terms$$

The LC tuned circuit acts as a band pass filter. This band pass filter eliminates the useless terms from the equation of $v_2(t)$.
Hence the output voltage $v_o(t)$ contains only the useful terms .

$$V_0(t) = aV_C cos(2\pi fct) + 2bx(t)cos(2\pi fct)$$

$$V_0(t) = [aV_C + 2bx(t)]cos(2\pi fct)$$

Therefore, $V_0(t) = aV_C[1 + \frac{2b}{a}x(t)]cos(2\pi fct)$…………………………(3)

Comparing this with the expression for standard AM wave i.e

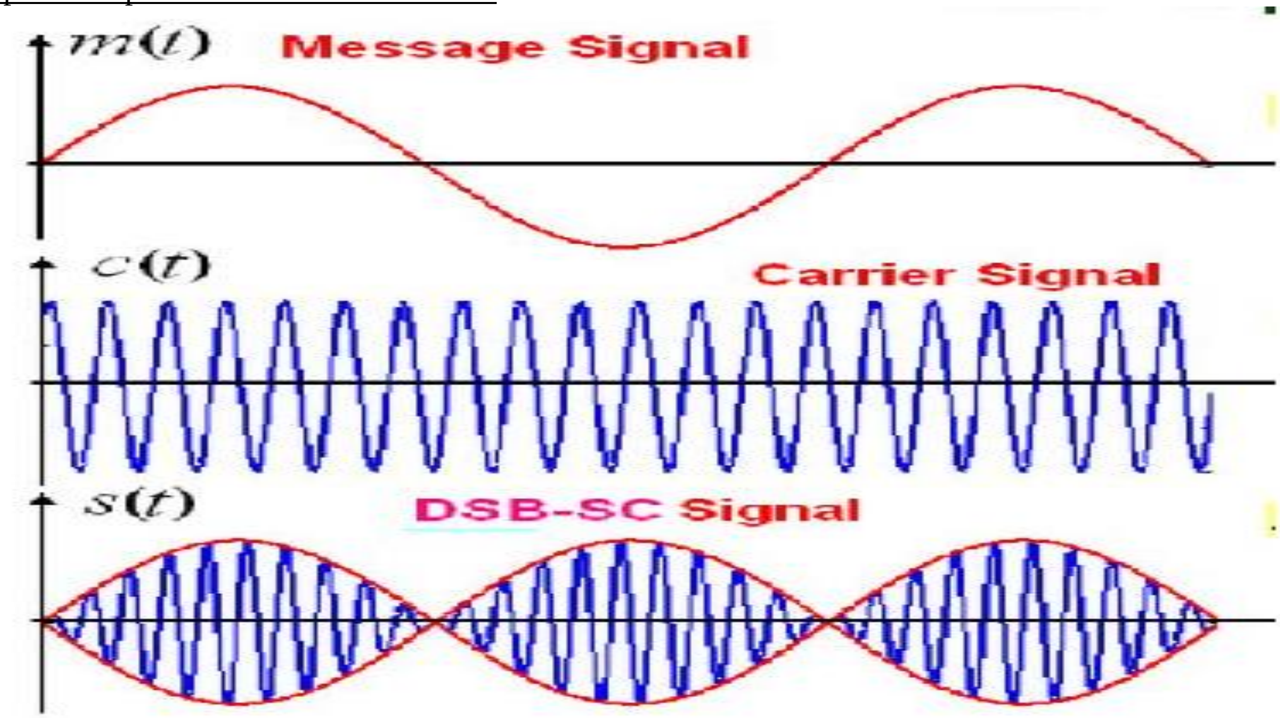
$$V_{AM}(t) = V_C[1 + m_a x(t)]cos(2\pi fct)$$

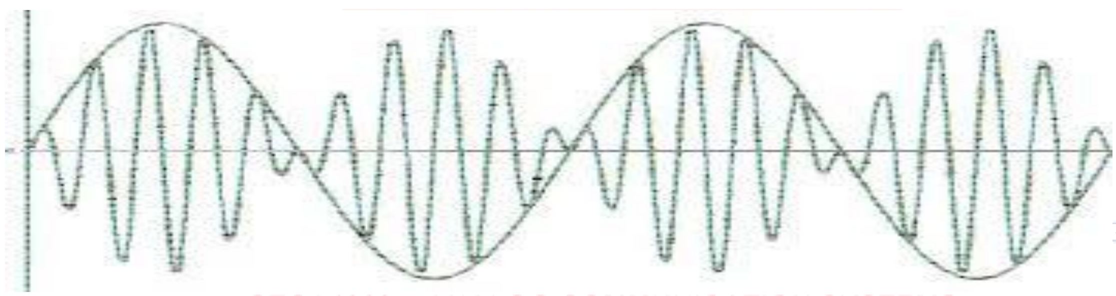
We find that the expression for $V_o(t)$ of equation (3) represents an AM wave with m = (2b/a) . Hence, the square law modulator produces an AM wave.

Double Sideband Suppressed Carrier (DSB-SC) Modulation

To improve the efficiency, and modulated carrier wave in modulated wave is suppressed which does not contain any Useful information, only the sidebands are transmitted. Such a modulation is called double sideband suppressed carrier (DSB-SC). This form of linear modulation is generated by using a product modulator that is when multiplying both messages signal and carrier signal, the resultant signal is DSB-SC signal.

Graphical Representation of DSB-SC AM





Mathematical representation of AM

Let modulating signal $V_m(t)=V_m \sin\omega_m t$ ……..(1)

V_m= Amplitude or voltage of message signal

ω_m= Frequency of message signal

Carrier signal $V_c(t)=V_c \sin\omega_c t$ ………………(2)

V_c= Amplitude or voltage of carrier signal

ω_c= frequency of carrier signal

Multiplying (1) and (2) by product modulator, modulated DSB-SC signal is generated

$V(t)= V_m(t)^* V_c(t)$

$= V_m \sin\omega_m t^* V_c \sin\omega_c t$

$= V_m V_c \sin\omega_m t \sin\omega_c t$

Apply 2sinA sin B=cos(A-B)-cos(A+B)

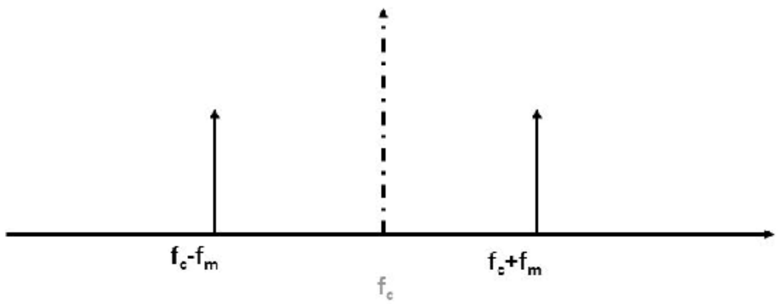
Then $V_{DSB\text{-}}sc(t) = \frac{V_m V_c}{2}[\cos(\omega_c - \omega_m)t - cos(\omega_c + \omega_m)t]$ ……(3)

We know that DSB with carrier AM wave modulated representation,

$$V_{AM}(t) = Vc\ \sin\omega_c t + \frac{V_m V_c}{2}[\cos(\omega_c - \omega_m)t - cos(\omega_c + \omega_m)t =$$

By comparing DSB with carrier and DSB-SC, $V_c \sin\omega_c t$ (ie) carrier wave is missing ;remaining two terms are same.

FREQUENCY SPECTRUM



Frequency spectrum double sideband suppressed carrier AM shows that carrier terms suppressed. It contains only two side band terms having the amplitude of ma * Vc / 2 , at the frequency of fc+fm and fc-fm.

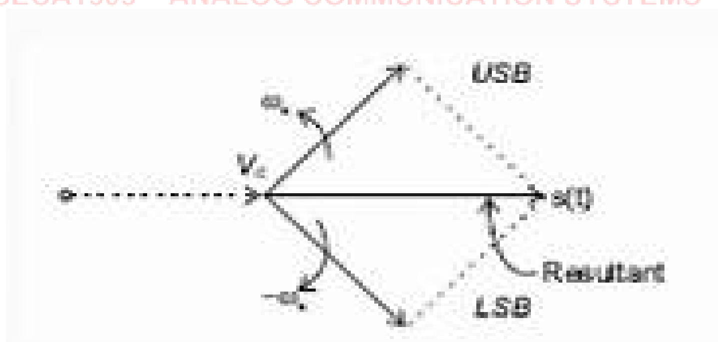
Bandwidth

Bandwidth= upper side band - lower side band

=fc+fm-(fc-fm)=

Bandwidth =2fm

Phasor diagram



Let us assume carrier Phasor is the reference phasor and it is directed in horizontal direction and it is denoted by dotted line, because it is suppressed after modulation.

- The USB term $m_a V_c/2\ \cos(\omega_c+\omega_m)t$ rotates at angular frequency of ω_m in anticlockwise direction and their LSB term $m_a V_c/2\ \cos(\omega_c-\omega_m)t$ rotates in clockwise direction with angular velocity ω_m.
- The resultant amplitude of the modulated wave at any point in the vector is the sum of the two side Bands

<u>Power calculation in DSB-SC AM</u>

In DSB-SC carrier is suppressed, then total power is only the sum of sideband power

$P_{t\ DSB\text{-}SC} = P_{USB} + P_{LSB} = \frac{\left(\frac{m_a V_c}{2}\right)^2}{R} + \frac{\left(\frac{m_a V_c}{2}\right)^2}{R}$

we know that RMS value of Vc = $\frac{V_c}{\sqrt{2}}$

$$Pt_{\text{DSB-SC}} = \frac{\left(m_a \frac{V_c/\sqrt{2}}{2}\right)^2}{R} + \frac{\left(m_a \frac{V_c/\sqrt{2}}{2}\right)^2}{R}$$

$$Pt_{\text{DSB-SC}} = \frac{m_a^2 \frac{V_c^2}{8}}{R} + \frac{m_a^2 \frac{V_c^2}{8}}{R} = \frac{m_a^2 V_c^2}{8R} + \frac{m_a^2 V_c^2}{8R} = \frac{2 m_a^2 V_c^2}{8R} = \frac{m_a^2 V_c^2}{4R}$$

$$Pt_{\text{DSB-SC}} = \frac{V_c^2}{2R}\left(\frac{m_a^2}{2}\right) = Pc\left(\frac{m_a^2}{2}\right)$$

Total power in both side bands are

$$Pt_{\text{DSB-SC}} = Pc\left(\frac{m_a^2}{2}\right)$$

Transmission Efficiency (%):

$$\%\eta = \frac{P_{t\ \text{DSB with carrier}} - P_{t\ \text{DSB- SC}}}{Total\ side\ band} * 100$$

We know that

$Pt_{DSB}\ with\ carrier = Total\ power\ Pt = Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]$

$$\%\eta = \frac{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right] - Pc\left(\frac{m_a^2}{2}\right)}{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]} * 100$$

$$\%\eta = \frac{Pc + Pc\left(\frac{m_a^2}{2}\right) - Pc\left(\frac{m_a^2}{2}\right)}{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]} * 100$$

$$\%\eta = \frac{1}{\left[1 + \left(\frac{m_a^2}{2}\right)\right]} * 100 = \frac{2}{2 + m_a^2} * 100$$

If ma = I then %η=2/3*100=66.7%

By suppressing carrier wave the efficiency or power saving increased to 66.7% But failed to conserve the band width.

<u>Generation of DSB-SC AM</u>

Balanced Modulator using FETs

A balanced modulator is a device that modifies a signal, usually in the form of an amplitude modulated (AM) radio signal. It takes the original signal that has both sidebands and a carrier signal, and then modulates it so that only the sideband signals come through the output modulator. This creates a balanced signal, as there is less noise because the carrier signal has been removed.

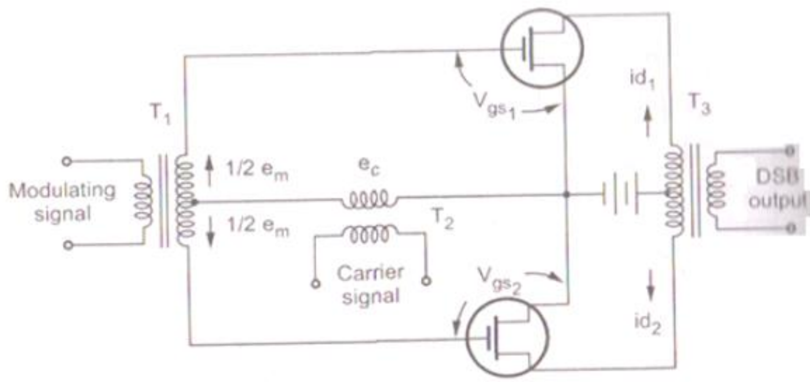


Fig: Balanced modulator using FETs

The balanced modulator can also be built using FETs. Figure shows the circuit diagram of balanced modulator using FETs. There are three transformers T1,T2 and T3. The carrier signal is applied to the center taps of the input transformer T1 and the output transformer T3 through the Transformer T1. The modulating signal is applied to the input transformer T1.The carrier signal is applied to the primary of transformer T2. This signal is further applied to two gates of FETs in phase through the secondary of T2. The modulating voltage appearing 180 degree out of phase at the gates, since these are the opposite ends of the center tapped transformer.
Consider that there is no modulating signal is applied. Then FET currents due to carrier signal are equal in amplitude but opposite in the directions. These opposite and equal currents are the primary of the output transformer cancel each other. Hence, no output is produced at the secondary of T3. Thus the carrier is suppressed.
When modulating signal is applied, the current id1 and id2 flow in the primary of T3 due to carrier signal as well as the modulating signal. The FET currents due to carrier are equal and opposite and cancel each other. Seems modulating signal is applied 180 degree out of phase at the gates, the FET currents due to modulating signal for equal but not opposite, hence do not cancel each other. Thus DSB output is produced by FET balanced modulator.

To prove that balanced modulator output produces DSB output
Now let us so mathematically that the FET balanced modulator really produces DSB output. The transfer curve id versus Vgs of a FET is almost parabolic and may be approximated by
.

$$i_{\mathrm{d}} = I_0 + aV_{\mathrm{gs}} + bV_{\mathrm{gs}}{}^2$$

Here I_0 is the current at zero Gate source voltage and a and b are constants. Since the drain currents i_{d1} and

i_{d2} flow in opposite direction in primary windings of output transformer T3. The effective primary current i_p is,

$$i_{\mathrm{p}} = i_{\mathrm{d1}} - i_{\mathrm{d2}} = aV_{\mathrm{gs1}} + bV_{\mathrm{gs1}}{}^2 - aV_{\mathrm{gs2}} - bV_{\mathrm{gs2}}{}^2$$

$$i_{\mathrm{p}} = a(V_{\mathrm{gs1}} - V_{\mathrm{gs2}}) + b(V_{\mathrm{gs1}}{}^2 - V_{\mathrm{gs2}}{}^2)$$

$i_{\mathrm{p}} = a(V_{\mathrm{gs1}} - V_{\mathrm{gs2}}) + b\,(V_{\mathrm{gs1}} + V_{\mathrm{gs2}})(V_{\mathrm{gs1}} - V_{\mathrm{gs2}})$……..(1)

Let us apply KVL to the input circuit

$$V_{\mathrm{gs1}} = \frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}}) \quad and\; V_{\mathrm{gs1}} = -\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}})$$

Putting these values in equation 1

$$i_{\mathrm{p}} = [(\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}}) - (-\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}})] + b[\,(\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}}) + (-\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}})][(\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}}) - (-\frac{1}{2}(e_{\mathrm{m}} + e_{\mathrm{c}})]$$

i_{p} =a e_m+2b e_c e_m

In the above equation e_m is the low frequency modulating signal. The output transformer T3 operates at carrier frequency, hence it will reject e_m. Because only the product term 2b e_c e_m of above equation. Thus

$$i_{\mathrm{p}} = 2\mathrm{b}e_{\mathrm{c}}e_m$$

We know that

$e_{\mathrm{c}} = V_{\mathrm{c}}\sin\omega_{\mathrm{c}}t$ and $e_{\mathrm{m}} = V_{\mathrm{m}}\sin\omega_{\mathrm{m}}t$

$$i_{\mathrm{p}} = 2\mathrm{b}V_{\mathrm{c}}\sin\omega_{\mathrm{c}}\mathrm{t}\, V_{\mathrm{m}}\sin\omega_{\mathrm{m}}\mathrm{t}$$

Here let us use

$$sin\omega_{\mathrm{c}}t\;sin\;\omega_{\mathrm{m}}t = \frac{[\cos(\omega_{\mathrm{c}} - \omega_{\mathrm{m}})t - cos(\omega_{\mathrm{c}} + \omega_{\mathrm{m}})t]}{2}$$

$$i_{\mathrm{p}} = \mathrm{b}V_{\mathrm{c}}\,V_{\mathrm{m}}\,[\cos(\omega_{\mathrm{c}} - \omega_{\mathrm{m}})t - cos(\omega_{\mathrm{c}} + \omega_{\mathrm{m}})t]$$

In the above equation $\omega_{\mathrm{c}} - \omega_{\mathrm{m}}$ represents LSB and $\omega_{\mathrm{c}} + \omega_{\mathrm{m}}$ represents USB. Thus the current flowing in the output transformer T3, produces only two side bands and no carrier. Thus proves that balanced modulator produces suppressor carrier DSB output.

Applications of DSBSC

- For transmitting stereo information in FM sound broadcast at VHF
- One important application of DSB is the transmission of color information in a TV signal. - CB radio - TV broadcasting - Air traffic control radios
- Garage door opens keyless remotes - DSB-SC is a technique used in electronic communication, most commonly for transmitting information via a radio carrier wave.

- DSB-SC used in stereo transmission of FM radio.
- Two way radio communications.

Advantages

- Lower power consumption
- The modulation system is simple
- DSB-SC is more efficient in transmitted power compared to DSB-FC
- Better signal to noise ratio

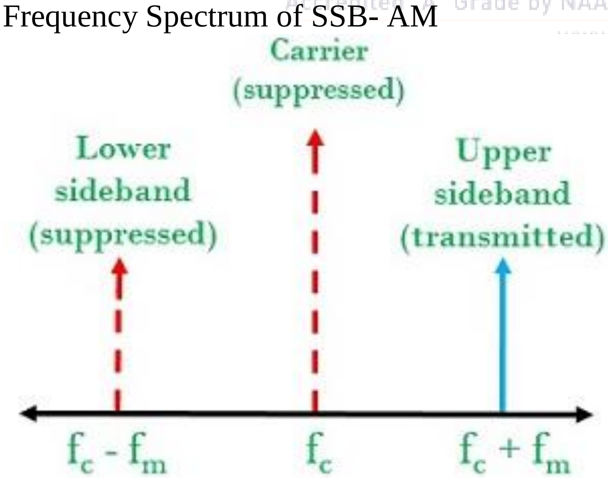
Disadvantages

- Even though the carrier is suppressed the bandwidth of DSB-FC remains same us DSB-FC
- Complex detection
- Less information about the carrier will be delivered to the receiver.
- Needs a coherent carrier detector at receiver

<u>Single-sideband modulation (SSB) (very important)</u>

Transmission in which only one sideband is transmitted is called **single- sideband transmission** or SSB. Carrier and one sideband are completely suppressed. The best way of thinking of SSB modulation is to first consider an amplitude modulated signal. This will have two frequency-shifted copies of the modulated signal (the lower one is frequency-inverted) on either side of the remaining carrier wave. These are known as sidebands: either upper sideband (USB) or less commonly lower sideband (LSB).

Frequency Spectrum of SSB- AM



When carrier and upper sideband are suppressed the bandwidth, reduced to half compared to DSB-FC and DSB-SC

BANDWIDTH= USB-LSB $=f_c-f_c+f_m =f_m$ [LSB= f_c and USB: f_c+f_m]

Power Calculation

We know that power in DSB-SC AM

$$Pt_{\text{DSB-SC}} = Pc\left(\frac{m_a^{\,2}}{2}\right)$$

In SSB-AM in total power s half of the power in side bands

$$Pt_{\text{SSB-SC}} = \frac{Pc\left(\frac{m_a^2}{2}\right)}{2} = Pc\left(\frac{m_a^2}{4}\right)$$

Transmission Efficiency (%):

$$\%\eta = \frac{P_{\text{t DSB with carrier}} - P_{\text{t SSB- SC}}}{Total\ side\ band} * 100$$

We know that

$$Pt_{\text{DSB}}\ with\ carrier = Total\ power\ Pt = Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]$$

$$\%\eta = \frac{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right] - Pc\left(\frac{m_a^2}{4}\right)}{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]} * 100$$

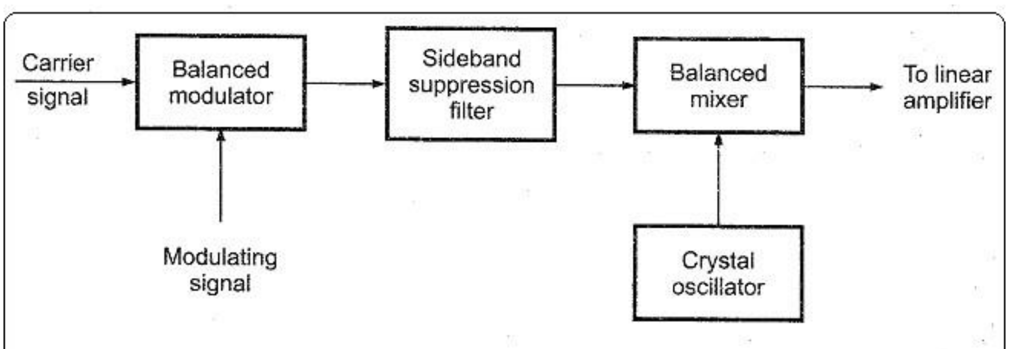
$$\%\eta = \frac{Pc + Pc\left(\frac{m_a^2}{2}\right) - Pc\left(\frac{m_a^2}{4}\right)}{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]} * 100$$

$$\%\eta = \frac{Pc + \frac{m_a^2}{4}}{Pc\left[1 + \left(\frac{m_a^2}{2}\right)\right]} * 100 = \frac{1+\frac{m_a^2}{4}}{1+\frac{m_a^2}{2}} * 100 = \frac{\frac{4+m_a^2}{4}}{\frac{2+m_a^2}{2}} = \frac{4+m_a^2}{4+2m_a^2}$$

If ma = I then %η=5/6*100=83.3%

By suppressing carrier wave and one of the side band, the efficiency or power saving increased to 66.3%

SSB Generation – Filter Method



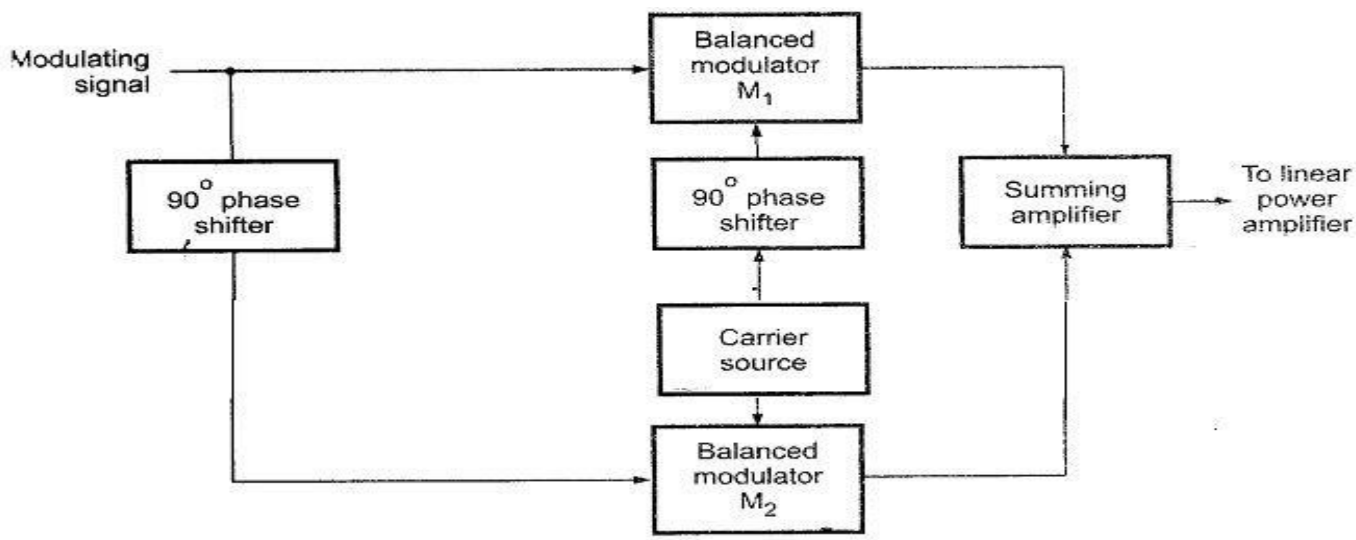
Filter method to suppress sideband

The figure shows the block diagram of the filter method to suppress one side band. The balanced modulator produces DSB output. This DSB signal contains both the side bands and it is given to the side band suppression filter to remove unwanted sideband. The filter must have a flat pass band and extremely high attenuation outside the pass band.

In order to have this type of response the Q of the tuned circuit must be very high. The required values of Q factor increases as the difference between modulating frequency and carrier frequency increases.

The Carrier frequency is usually same as the transmitter frequency. For highest transmitting frequencies required value of Q is so high that there is no practical of achieving it. In such situations, initial modulation is carried out at low frequency carrier say 100kHz by the balanced modulator. Then the filter suppresses one of the side bands.
The frequency of the SSB signal generated at the output of the filter is very low as compared to the transmitter frequency. The frequency is boosters up to the transmitter frequency by the balanced mixer and crystal oscillator. This process of frequency boosting is also called as up-conversion. The SSB the signal having frequency equal to the transmitter frequency is then amplified by the linear amplifiers.

SSB Generation – Phase Shift Method



Phase shift method to generate SSB signal

The figure shows a block diagram of a phase shift method to generate SSB. The carrier signal is shifted by 90° and applied to the balanced modulator M1. The modulating signal is also directly applied to this balanced modulator. The carrier signal directly applied to the balanced modulator M2. The modulating signal is phase shifted by 90° and applied to the balanced modulator M2. Both the modulators produce and output consisting of only side bands. The upper balance modulator M1 generates upper sideband and lower side band, but each one is shifted by +90°. Balanced modulator M2 generates the upper and lower side Bands, but the upper side band is shifted by +90°, whereas lower side band is shifted by -90°.

The balanced modulators or added by the summing amplifier. Both the modulators are phase shifted by +90°, and they are in phase and add to produce double amplitude signal. But lower side bands of the balanced modulators are (+90°,-90°) 180 degree out of phase and hence cancel each other. Thus the output of the summing amplifier contains only upper side band SSB signal. The carrier is already suppressed by balanced modulators.

Let us see mathematically, how the side bands add and cancel each other because of phase shift. Input to the balanced modulator M1 are $\sin\omega_m t$ and $\sin(\omega_c t + 90°)$. Hence the output of M1 will be

Output of M1= $\cos[(\omega_c t + 90°) - \omega_m t] .- \cos(\omega_c t + 90°) + \omega_m t]$.

$= \cos(\omega_c t - \omega_m t + 90°) - \cos(\omega_c t + \omega_m t + 90°)$

In the above equation observe the first term represents the lower side band with +90° phase shift and the second term represents the upper side band with +90° phase shift. Now inputs to the balanced modulator M2 are $\sin(\omega_m t + 90°)$ and $\sin\omega_c t$ and. Hence the output of M2 will be,

Output of M1= $\cos[\omega_c t - (\omega_m t + 90°)] - \cos[\omega_c t + (\omega_m t + 90°)]$.

$= \cos(\omega_c t - \omega_m t - 90°) - \cos(\omega_c t + \omega_m t + 90°)$

In the above equation, the first term represents lower sideband and has a phase shift of-90°. Similarly the second term represents the upper side band with the phase shift of + 90°. When signals of output M1 and M2 add in the summing amplifier, the lower side band cancel each other and they are out of phase. The second term adds since they have same phase shift of +90° that is in phase. Thus SSB is generated other output of summing amplifier

Uses of Single sideband modulation

- Single sideband modulation is widely used for two way radio communication.
- Single sideband modulation used for voice transmission

Advantages of Single sideband modulation

- As the carrier is not transmitted, this enables a 50% reduction in transmitter power level for the same level of information carrying signal
- As only one sideband is transmitted there is a further reduction in transmitter power.
- As only one sideband is transmitted the receiver bandwidth can be reduced by half. This improves the signal to noise ratio by a factor of two, i.e. 3 dB, because the narrower bandwidth used will allow through less noise and interference.

Disadvantages of Single sideband modulation

- Complex transmitter and receiver configurations are required and hard to demodulate.
- Amplitude modulation typically produces a modulated output signal that has twice the bandwidth of the modulating signal, with a significant power component at the center carrier frequency. Single-sideband modulation improves this, at the cost of extra complexity.

Vestigial Sideband in AM

Need for VSB or limitations of double sideband suppressor carrier and single sideband suppressed carrier

- Many message signals such as television video, facsimile(fax) and high speed data signals having large band width and significant low frequency content.
- Single sideband AM system is used to conserve the bandwidth but practical SSB-SC AM systems have poor low frequency response. i.e simply we can say that SSB-SC is well suited for transmission of voice signals, having no frequency components between 0 to few hundred Hertz. On the other hand, when signals contain frequency components at extremely low frequencies(as in between television and telegraph signals) single sideband suppressed carrier is not suitable for transmission of the signals. Because these low frequency components give rise to side bands of that translated or modulated signal and side bands are very close to carrier frequency, therefore it is very difficult to isolate or remove one side one from the other; the required filter must have a very sharp frequency characteristics.

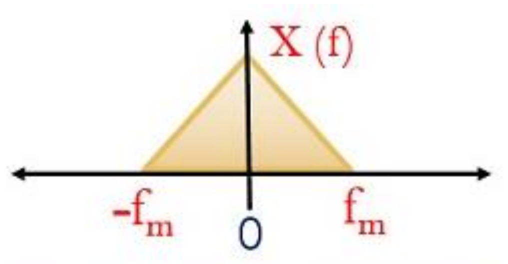
- Double sideband suppressed carrier is well suited for (low-frequency messages) messages with low frequency content, but the transmission bandwidth is twice that of single sideband suppressed carrier.
- Therefore a new modulation scheme has been introduced, that offers the best compromise between band width conservation, improved low frequency response and improved power efficiency, it is called as vestigial sideband modulation

- In VSB, instead of rejecting one side band completely (as in SSB-SC), gradual cut off one side band is allowed. This gradual cut is compensated by a vestige or portion of the other side band.

Frequency domain description

Frequency spectrum

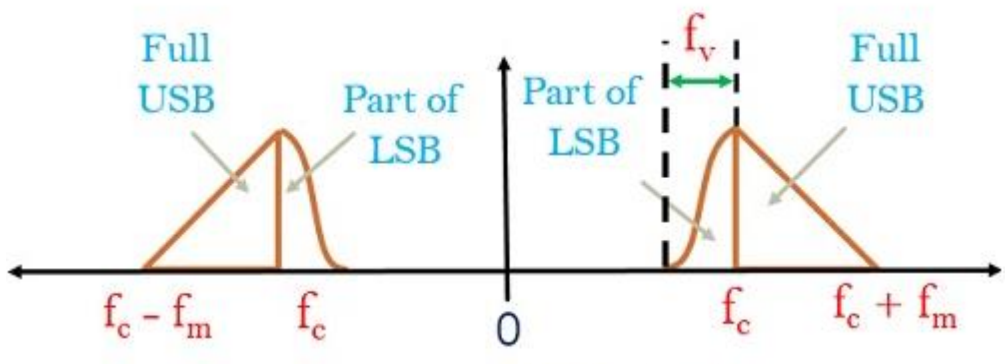
Spectrum of VSB is shown in figure. The spectrum of message signal x(t) has also been shown. In the frequency spectrum, it is assumed that the upper side band is transmitted as it is on the lower side band is modified into the vestigial sideband.



Spectrum of message signal

Transmission bandwidth

From the figure, it is evident that the transmission bandwidth of VSB modulated wave is given by,
B=(fm+fv)Hz
Where fm= message bandwidth
fc= width of the vestigial sideband



Spectrum of VSB signal

Advantages of VSB

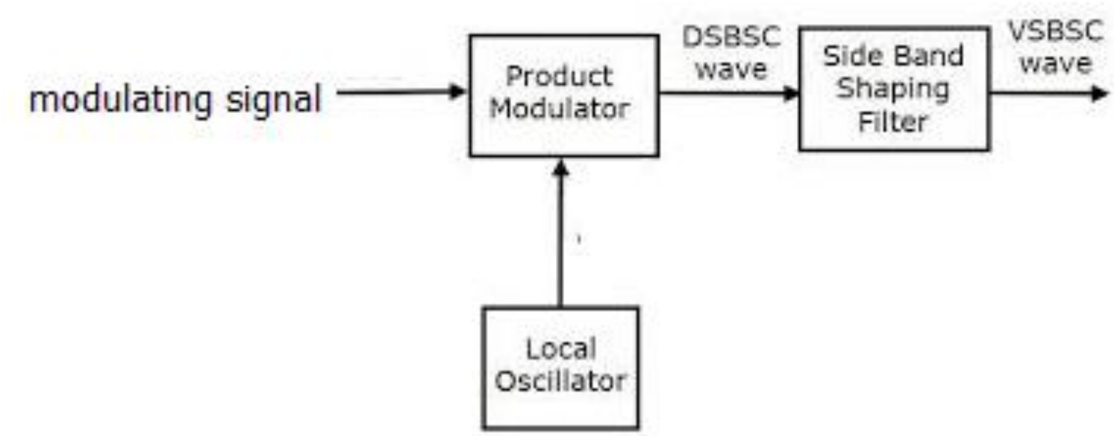
- From the above frequency spectrum, the VSB filter or a type of band pass filter required need not have sharp cut off, which is an advantage of VSB system, however as compared to SSB-SC to the bandwidth of VSB becomes larger although it remains much smaller than DSB-SC signal or VSB has band width greater than SSB-SC but less than DSB-SC system.
- Power transmission is greater than DSB-SC but less than SSB-SC system. (75%)
- No low-frequency component lost, hence it avoids phase distortion.
- It is used in TV for transmission of picture signals.

Applications of VSB

VSB modulation has become standard for the transmission of television signals. Because the video signals need a large bandwidth if transmitted using DSB-FC or DSB-SC.

GENERATION OF VSB MODULATED WAVE
Filter Method

The block diagram of VSB modulator is shown in figure. The modulating signal is applied to a product modulator. The output of the carrier oscillator is also applied to the input of the product modulator. The output of the product modulator is given by DSB-SC modulated wave. This DSB-SC signal is then applied to your side band shaping filter. The design of this filter depends on the desire spectrum of the VSB modulated signal. This filter will pass wanted side band and the vestige of the unwanted sideband.



<u>Mathematical Representation of VSB</u>

Let the message signals the sum of two sinusoids

$V_m(t)=V_{m1}\cos\omega_{m1}t+V_{m2}\cos\omega_{m2}t$

$\quad =V_{m1}\sin\omega_1 t+V_{m2}\cos\omega_2 t$

Assume VSB is used to for modulating TV signal contain audio and video signal, it may be represented by sum of two sinusoids

In the product modulator message and carrier signals are multiplied

$$V_{DSB}(t)=V_m(t)*V_c(t)$$

$$V_{DSB\text{-}SC}(t)=V_{m1}\cos\omega_{m1}t+V_{m2}\cos\omega_{m2}t*V_c\cos\omega_c t$$

$$V_{DSB\text{-}SC}(t)=V_{m1}V_c\cos\omega_{m1}t\cos\omega_c t+V_{m2}V_c\cos\omega_{m2}t\cos\omega_c t$$

Apply

$$\cos\omega_m t\cos\omega_c t=\frac{[\cos(\omega_c-\omega_m)t-\cos(\omega_c+\omega_m)t]}{2}$$

$$V_{DSB\text{-}SC}(t)=V_{m1}V_c\frac{[\cos(\omega_c-\omega_{m1})t-\cos(\omega_c+\omega_{m1})t]}{2}+V_{m2}V_c\frac{[\cos(\omega_c-\omega_{m2})t-\cos(\omega_c+\omega_{m2})t]}{2}$$

$$V_{DSB\text{-}SC}(t) = \frac{V_{m1}V_c}{2}[\cos(\omega_c - \omega_{m1})t - cos(\omega_c + \omega_{m1})t] + \frac{V_{m2}V_c}{2}[\cos(\omega_c - \omega_{m2})t - cos(\omega_c + \omega_{m2})t]$$

from this carrier wave is completely suppressed, but we have lower side band and upper side bands for both audio and video message envelopes. Now the output of DSB-SC Amis passed through VSB (side band shaping filter),the resultant VSB signal is

$$V_{VSB}(t) = \frac{V_1 V_c \in}{2}\cos(\omega_c - \omega_1)\,t + \frac{V_1 V_c}{2}(1-\in)\cos(\omega_c + \omega_1)\,t + \frac{V_2 V_c}{2}\cos(\omega_c + \omega_2)\,t$$

Where ∈=constant

From this equation carrier wave is fully suppressed

A portion of the lower sideband is suppressed (i.e) $\cos(\omega_c - \omega_2)\,t$ (one side band is suppressed completely)

A smart part or trace or vestige of upper side band is included (i.e) $\frac{V_1 V_c}{2}(1-\in)\cos(\omega_c + \omega_1)\,t$

Uses of vestigial sideband

- Used to transmit television (TV) Signal.

Advantages

- VSB is intended to save bandwidth over regular AM.
- Portions of one of the redundant sidebands are removed to form a vestigial **sideband** signal.
- The actual information is transmitted in the sidebands, rather than the carrier; both sidebands carry the same information. Because LSB and USB are essentially mirror images of each other, one can be discarded or used for a second channel or for diagnostic purposes.

Disadvantages

- VSB transmission is similar to *single-sideband* (SSB) transmission, in which one of the sidebands is completely removed. In VSB transmission, however, the second sideband is not completely removed, but is filtered to remove all but the desired range of frequencies.

Advantages of Amplitude Modulation:

- Because of amplitude modulation wavelength, AM signals can propagate longer distances.
- For amplitude modulation, we use simple and low cost circuit; we don't need any special equipment and complex circuits that are used in frequency modulation.
- The Amplitude modulation receiver will be wider when compared to the FM receiver. Because, atmospheric propagation is good for amplitude modulated signals.
- Bandwidths limit is also big advantage for Amplitude modulation, which doesn't have in frequency modulation.
- Transmitter and receiver are simple in Amplitude modulation. When we take a demodulation unit of AM receiver, it consists of RC filter and a diode which will demodulate the message signal or modulating signal from modulated AM signal, which is unlike in Frequency modulation.
- Zero crossing in Amplitude modulation is equidistant.

Disadvantages of Amplitude Modulation:

- Adding of noise for amplitude modulated signal will be more when compared to frequency modulated signals. Data loss is also more in amplitude modulation due to noise addition. Demodulators cannot reproduce the exact message signal or modulating signal due to noise.
- More power is required during modulation because Amplitude modulated signal frequency should be double than modulating signal or message signal frequency. Due to this reason more power is required for amplitude modulation.

- Sidebands are also transmitted during the transmission of carrier signal. More chances of getting different signal interfaces and adding of noise is more when compared to frequency modulation. Noise addition and signal interferences are less for frequency modulation. That is why Amplitude modulation is not used for broadcasting songs or music.

Applications of Amplitude Modulation:

- Used to carry message signals in early telephone lines.
- Used to transmit Morse code using radio and other communication systems.
- Used in Navy and Aviation for communications as AM signals can travel longer distances.
- Widely used in amateur radio.

AM Transmitter

Transmitters that transmit AM signals are known as AM transmitters. These transmitters are used in medium wave (MW) and short wave (SW) frequency bands for AM broadcast. The MW band has frequencies between 550 KHz and 1650 KHz, and the SW band has frequencies ranging from 3 MHz to 30 MHz. The two types of AM transmitters that are used based on their transmitting powers are:

· High Level
· Low Level

High level transmitters use high level modulation, and low level transmitters use low level modulation. The choice between the two modulation schemes depends on the transmitting power of the AM transmitter. In broadcast transmitters, where the transmitting power may be of the order of kilowatts, high level modulation is employed. In low power transmitters+-, where only a few watts of transmitting power are required, low level modulation is used.

High-Level and Low-Level Transmitters

Fig shows the block diagram of high-level and low-level transmitters. The basic difference between the two transmitters is the power amplification of the carrier and modulating signals.
Figure (a) shows the block diagram of high-level AM transmitter.

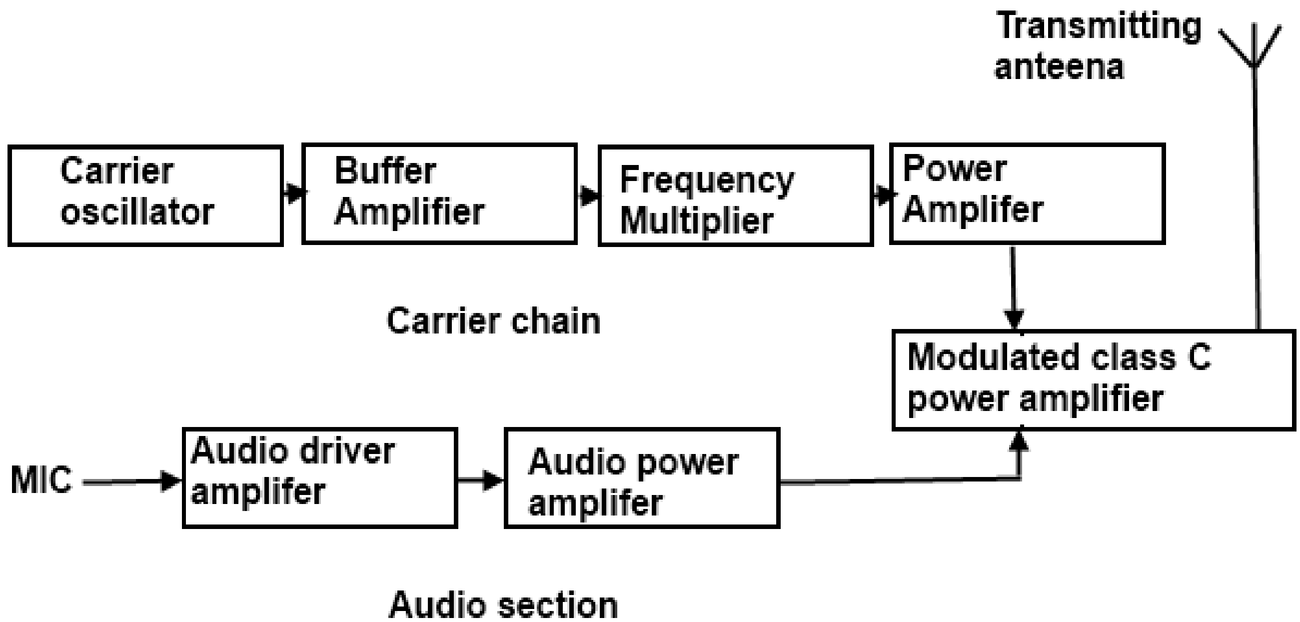


Figure (a) Block diagram of high level AM transmitter

In high-level transmission, the powers of the carrier and modulating signals are amplified before applying them to the modulator stage, as shown in figure (a). In low- level modulation, the powers of the two input signals of the modulator stage are not amplified. The required transmitting power is obtained from the last stage of the transmitter, the class C power amplifier. The various sections of the figure (a) are:

- Carrier oscillator
- Buffer amplifier
- Frequency multiplier
- Power amplifier
- Audio chain
- Modulated class C power amplifier

Carrier oscillator

The carrier oscillator generates the carrier signal, which lies in the RF range. The frequency of the carrier is always very high. Because it is very difficult to generate high frequencies with good frequency stability, the carrier oscillator generates a sub multiple with the required carrier frequency. This sub multiple frequency is multiplied by the frequency multiplier stage to get the required carrier frequency. Further, a crystal oscillator can be used in this stage to generate a low frequency carrier with the best frequency stability. The frequency multiplier stage then increases the frequency of the carrier to its required value.

Buffer Amplifier

The purpose of the buffer amplifier is twofold. It first matches the output impedance of the carrier oscillator with the input impedance of the frequency multiplier, the next stage of the carrier oscillator. It then isolates the carrier oscillator and frequency multiplier.
This is required so that the multiplier does not draw a large current from the carrier oscillator. If this occurs, the frequency of the carrier oscillator will not remain stable.

Frequency Multiplier

The sub-multiple frequency of the carrier signal, generated by the carrier oscillator , is now applied to the frequency multiplier through the buffer amplifier. This stage is also known as harmonic generator. The frequency multiplier generates higher harmonics of carrier oscillator frequency. The frequency multiplier is a tuned circuit that can be tuned to the requisite carrier frequency that is to be transmitted.

Power Amplifier

The power of the carrier signal is then amplified in the power amplifier stage. This is the basic requirement of a high-level transmitter. A class C power amplifier gives high power current pulses of the carrier signal at its output.

Audio Chain

The audio signal to be transmitted is obtained from the microphone, as shown in figure (a). The audio driver amplifier amplifies the voltage of this signal. This amplification is necessary to drive the audio power amplifier. Next, a class A or a class B power amplifier amplifies the power of the audio signal.

Modulated Class C Amplifier

This is the output stage of the transmitter. The modulating audio signal and the carrier signal, after power amplification, are applied to this modulating stage. The modulation takes place at this stage. The class C amplifier also amplifies the power of the AM signal to the reacquired transmitting power. This signal is finally passed to the antenna, which radiates the signal into space of transmission.

Figure shows the block diagram of a low-level AM transmitter.

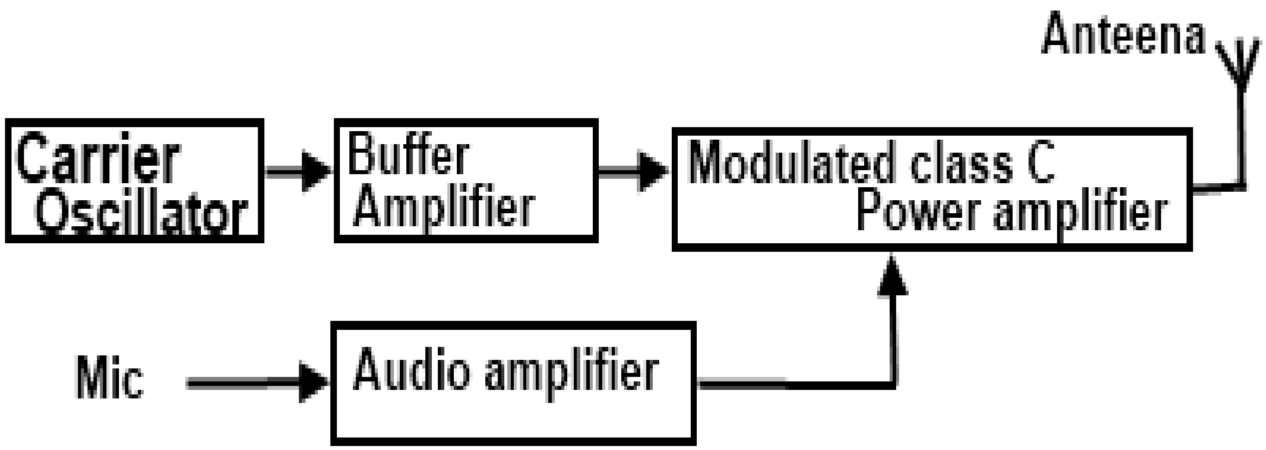


Figure (b) Block diagram of Low-level AM transmitter

The low-level AM transmitter shown in the figure (b) is similar to a high-level transmitter, except that the powers of the carrier and audio signals are not amplified. These two signals are directly applied to the modulated class C power amplifier.

Modulation takes place at the stage, and the power of the modulated signal is amplified to the required transmitting power level. The transmitting antenna then transmits the signal.

DETECTOR/DEMODULATORS

The process of separating or extracting the message signal from modulated (received) signal is called demodulation or detection. For amplitude modulation, the process of demodulation or detection can be accomplished very simply using a diode, or it may be achieved in other ways that provide more effective demodulation of the waveform.

Detection of AM waves can be done using

Diode detectors

- This is a simplest form of AM demodulator.
- It requires just a diode along with a capacitor to remove the high frequency components.
- It suffers from a number of disadvantages, but its performance is
- more than adequate for most applications including broadcast receivers where cost is a significant driver.

Synchronous detection

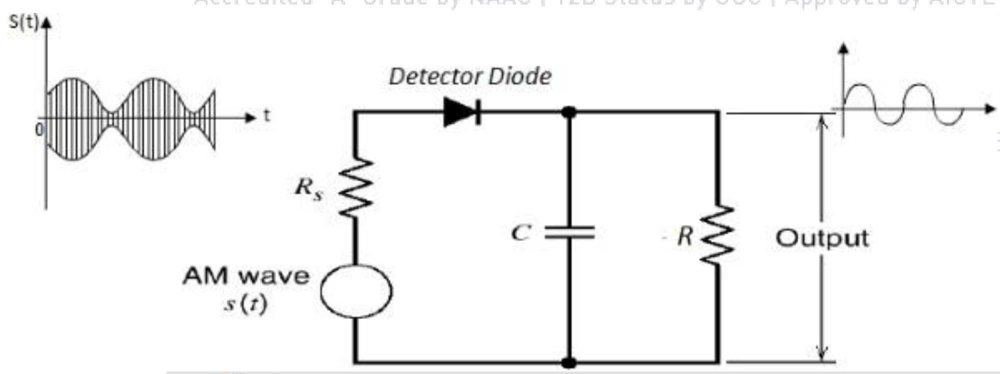
- This detector offers a higher level of performance
- The cost is high because of the use of more components. This means that it is only used in receivers where the levels of performance are paramount and can justify the additional component costs.

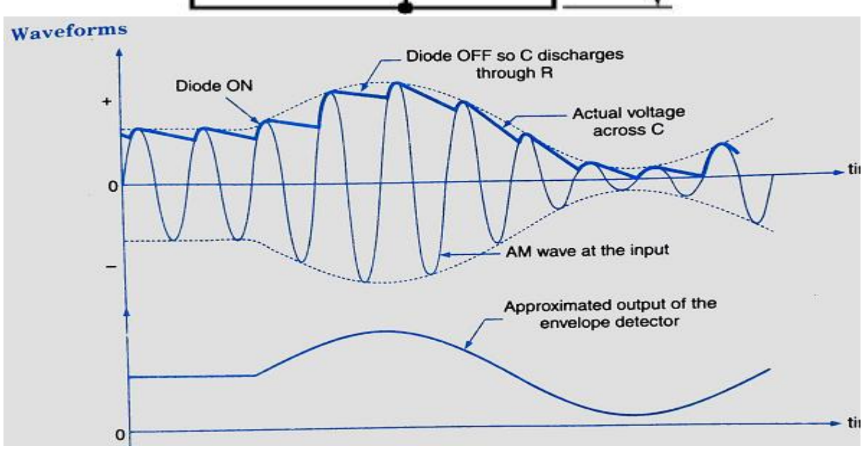
There are two types of AM detectors

1. Envelope detector
2. Square Law detectors or nonlinear detectors

Envelope Detector

The envelope demodulator is a simple and very efficient device which is suitable for the detection of a narrowband AM signal. A narrowband AM wave is the one in which the carrier frequency f_c is much higher as compared to the bandwidth of the modulating signal .





- Envelope detector is otherwise called as linear diode detector
- The simplest and most widely used amplitude Demodulator is the envelope detector
- Commonly used for detecting double sideband with carrier(standard AM) or vestigial sideband signals like speech ,music, video etc
- An envelope demodulator produces an output signal that follows the envelope of the input AM signal exactly.
- It is used in all the commercial AM radio receivers.
- The envelope demodulator consists of a diode and RC filter.
- A detector circuit whose output follows the peak of the modulated carrier, reproduce the modulating signal. such a detector may be termed as peak detector or envelope detector
- The diode is operating in a linear region of its characteristics can extract the envelope of AM wave. so that this detector is also called as linear diode detector

- The selected AM modulated signal is applied to the junction diode and then applied to the load impedance consisting of R and C network.
- Since the magnitude of applied AM is large, the operation takes place in linear region of transfer characteristics of diode.
- During the positive half cycle of the modulated signal the diode conducts(forward biased),while during the negative of cycle it does not conduct(diode is Reverse biased
- Let us assume first the capacitor is absent. The output waveform will be positive half cycle of modulated wave (rectified modulated signal) across resistor R.
- Now the capacitor C is introduced parallel to the resistor R i.e now the rectified output modulated signal is passed through RC network.
- Here the capacitor C is charged to the peak value of carrier voltage, for the positive half cycle of rectified wave.
- But for the negative half cycle, diode is Reverse biased and the carrier voltages disconnected from RC circuit. so the capacitor starts discharging through the resistor R with the time constant =RC
- If the time constant is properly chosen, the voltage across the capacitor does not fall appreciably during the small
- If the time constant is properly chosen, the voltage across the capacitor does not fall appreciably during small period of the negative half cycle and by that time next positive half cycle appears. this positive cycle further charges the capacitor through the peak value of the carrier voltage and process continuous
- thus for the voltage across the capacitor is same as the envelope of the modulated carrier, like spiky modulating signal due to charging and discharging of the capacitor

Significance of RC time constant

By keeping the time constant, RC large, the capacitor discharging a small that is negligible hence spikes can be reduced.
But the large values of RC create another problem called diagonal clipping. Hence we cannot increase it beyond the certain limit.

Choice of time constant RC

- It is desired to keep the time constant RC very high as compared to time of carrier dev in order to minimize spikes fluctuation in detected envelope. On the other hand, if it is kept too high the discharge curve becomes approximately horizontal. In that case, negative pics of detected envelope baby completely or partially missing. Therefore the recovered baseband signal is distorted at negative peaks. This type of distortion is called diagonal clipping.

Diagonal clipping maybe caused

- When the time constant of detector is not selected properly
- Increase in depth of modulation index make the envelope slope steeper than the discharge path of capacitor
- To avoid diagonal clipping proper value of time constant may be obtained as follows
- During the non-conducting period of diode, the voltage across RC combination at an instant “t” is given by

$$V_c(t) = V_0 e^{-t/RC}$$

- The slope of capacitor discharge or rate of fall may be obtained by differentiating $V_c(t)$ with respect to time t

$$\frac{dV_c(t)}{dt} = V_0 e^{-t/RC} * \frac{1}{RC} = \frac{-V_0}{RC} * e^{-t/RC}$$

$$\frac{dV_c(t)}{dt} = \frac{-V_c(t)}{RC}$$

- If the distortion is avoided, the decrease in capacitor voltage must follows envelope
- We know the envelope of the modulated voltage signal which is given by

$$V_{AM}(t) = V_c[1 + m_a \sin\omega_m t] \sin\omega_c t$$

V_{AM}=envelope voltage

$V_{AM} = V_c[1 + m_a \sin\omega_m t]$

The slope of the envelope is given by

$$\frac{dV_{AM}(t)}{dt} = \frac{d}{dt} V_c[1 + m_a sin\omega_m t]$$

$$= \frac{dV_c}{dt} + \frac{d}{dt} V_c m_a sin\omega_m t = 0 + K V_c m_a \omega_m cos\omega_m t$$

Where K=proportionality constant

- The diagonal clipping is avoided, rate of fall for slope of capacitor is algebraically greater than or equal to slope of envelope

$$\frac{dV_c(t)}{dt} \geq \frac{dV_{AM}}{dt}$$

$$\frac{-V_c(t)}{RC} \geq KV_c m_a \omega_m cos\omega_m t$$

Put $\frac{-V_c[1+m_a sin\omega_m(t)]}{RC} \geq KV_c m_a \omega_m cos\omega_m t$

$$\frac{-1}{RC} \geq \frac{-Km_a \omega_m cos\omega_m t}{[1+m_a sin\omega_m(t)]}$$

- The minimum value of RC, can be evaluated by differentiating write inside of the above equation and equating it to zero.

$where\ sin\omega_m(t) = -m_a and\ cos\omega_m(t) = \sqrt{1-m_a{}^2}$

$$\frac{-1}{RC} \geq \frac{m_a \omega_m \sqrt{1-m_a{}^2}}{[1-m_a{}^2]}$$

Time constant $RC \leq \frac{\sqrt{1-m_a{}^2}}{m_a \omega_m}$

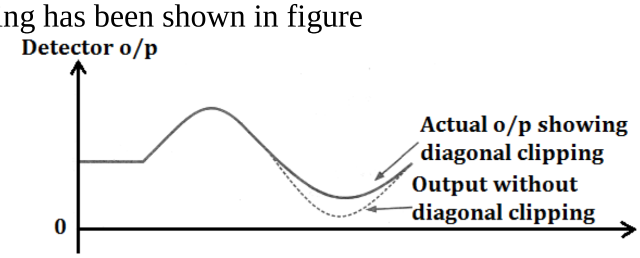
- This condition can be satisfied for avoiding discussions in the detected output
- From this the time constant RC cannot be kept too high or too low
- If RC is very low, discharging path during non-conducting period is almost vertical, resulting large fluctuations in the output voltage
- If RC is very high, discharge curve for path is almost horizontal and it then misses several pics of rectified output during negative peak

<u>Distortions in the envelope detector output</u>

- There are two types of distortions which can occur in the detector output. they are
  1. Diagonal clipping
  2. negative peak clipping

Diagonal clipping

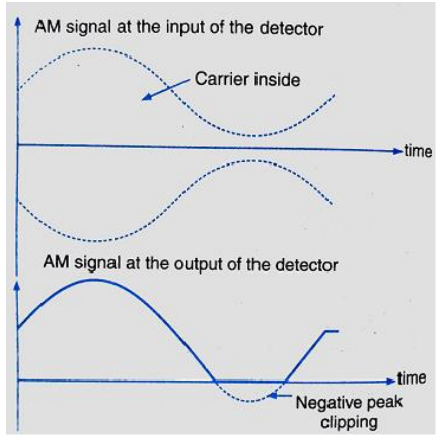
- This type of distortion occurs when the RC time constant of the load circuit is too long. Due to this, the RC circuit cannot follow the first changes in the modulating envelope. The diagonal clipping has been shown in figure



Negative peak clipping

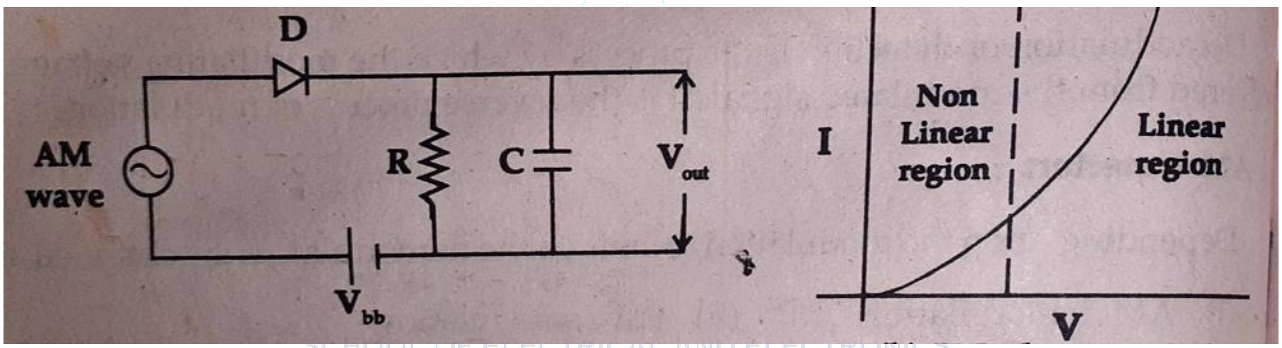
- This distortion occurs due to a fact that the modulation index on the output side of the deducted is higher than that on its input side. Hence, at a depth of modulation of the transmitted signal, the over modulation (more than 100% modulation) may take place at the output of the detector. The negative peak clipping will take place as a result of this over modulation

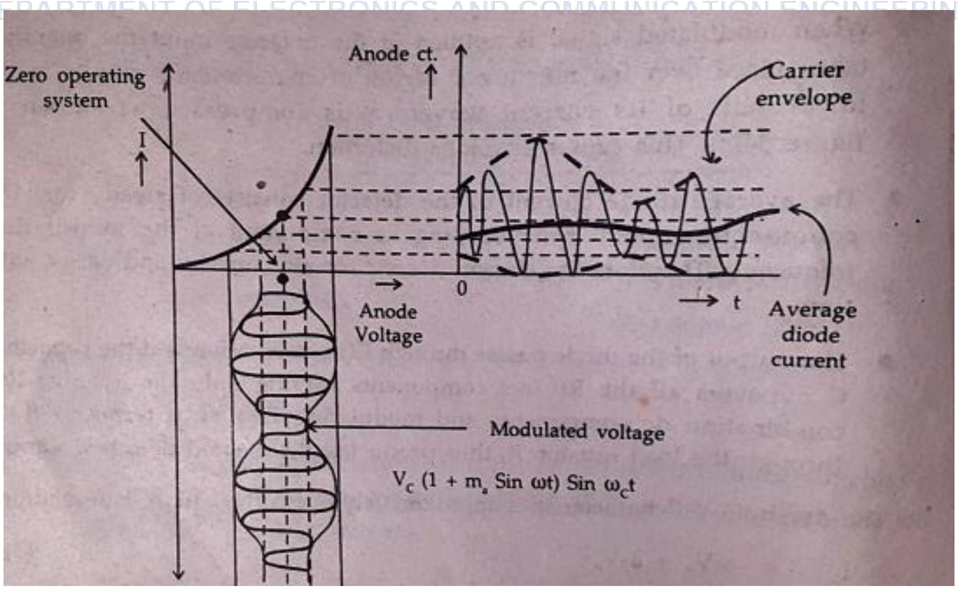
- The only way to reduce or eliminate the distortions is to choose the RC time constants properly



.

<u>Square Law Demodulator</u>





A device is said to be nonlinear its output this is not a linear function of the input amplitude that is the device should work in the nonlinear portion of its transfer characteristics.
A nonlinear detector thus operates in the nonlinear region of its transfer characteristics and is

suitable for small signals. The circuit is very similar to square Law modulator, the only difference being the filter circuit. In the detector the filter circuit is a low pass filter instead of a band pass filter used in modulator. A Simple square Law detector is shown in figure. A diode can be used as a Square Law detector if it is made to operate in the nonlinear portion of its dynamic V-I characteristics. In the diode detector, the input carrier voltage is of very small in magnitude. Thus the operating point will be on the nonlinear portion of its transfer characteristics. On the other hand, in linear diode circuits this operating point shifted towards the linear portion of characteristics because of large input carrier voltage and most of the operation takes place over in the linear region.
operation. When the modulating signal is applied at the detector input the operation takes place over the nonlinear region of characteristics and the lower half of its current waveform is compressed. This causes envelope distortion. The average diode current consists of steady or DC component and time-varying component at the modulation frequency. Therefore the current remained constant and vary with time. The capacitor C bypasses all the RF components leaving only the average DC components to flow through the load resistor R, producing the desired detected output.
The input output characteristics, i.e., transfer characteristics of a square law detector is nonlinear and it is expressed mathematically as under;
By non-linear square law relation
Diode current
$i_o(t) = a_1 V_a + a_2 V_a^2$……………………………….(1)
where a_1 and b_1 are constants and
V_a=anode voltage=biasing voltage + AM modulated voltage
$V_a = V_b + V_c[1 + m_a sin\omega_m t]sin\omega_c t$……………..(2)
Now, substituting the expression (1) in (2), we get

$$i_o(t) = a_1[V_b + V_c[1 + m_a sin\omega_m t]sin\omega_c t] + a_2[V_b + V_c[1 + m_a sin\omega_m t]sin\omega_c t]^2$$

$$i_o(t) = a_1 V_b + a_1 V_c sin\omega_c t[1 + m_a sin\omega_m t] + a_2 V_b^2 + a_2 V_c^2 sin^2\omega_c t[1 + m_a sin\omega_m t]^2 + 2a_2 V_b V_c sin\omega_c t[1 + m_a sin\omega_m t]$$

Put

$$sin^2\omega_c t = \frac{[1 - \cos 2\,\omega_c\,t]}{2}$$

$$\boldsymbol{i_o(t) = a_1 V_b + a_1 V_c sin\omega_c t[1 + m_a sin\omega_m t] + a_2 V_b^2 + a_2 V_c^2 \frac{[1 - \cos 2\,\omega_c\,t]}{2}[1^2 + m_a^2 sin^2\omega_m t + 2m_a sin\omega_m t] + 2a_2 V_b V_c sin\omega_c t[1 + m_a sin\omega_m t]}$$

$$\boldsymbol{i_o(t) = a_1 V_b + a_1 V_c sin\omega_c t[1 + m_a sin\omega_m t] + a_2 V_b^2 + \frac{a_2 V_c^2}{2} - \frac{a_2 V_c^2 \cos 2\,\omega_c\,t}{2}[1 + m_a sin\omega_m t]^2 + 2a_2 V_b V_c[1 + m_a sin\omega_m t]}$$

The (R.F) high frequency carrier terms are bypassed through the capacitor and the circuit is tuned to ω_m alone**.** Thus, output current contains ω_m and DC terms.

Therefore $\boldsymbol{i_o(t) = a_1 V_b + a_2 V_b^2 + \frac{a_2 V_c^2}{2} + \frac{a_2 V_c^2}{2} 2m_a sin\omega_m t}$

Thus the base band signal (original message signal) is recovered.

Reference Books

1. Dr. Sanjay Sharma, "Analog Communication Systems" 2013, January 2013
2. R.P Singh, S.D. Spare, "Communication Systems: Analog and Digital" ,Tata McGraw-Hill, 3rd Edition, July 2017.
3. Sanjeev Guptha, "Electronic Communication (Analog, Digital and Wireless)", Khanna publications, January 1997
4. B.P.Lathi, "Modern Digital and Analog Communication Systems", 3rd Edition, Oxford University Press, 2007.
5. Louis. E. Frenzel, "Communication Electronics Principles and applications", 3rd Edition, Tata McGraw Hill, 2002.
6. Louis. E. Frenzel, "Principles of Electronic Communication Systems", 3rd Edition, Tata McGraw Hill, July 2017.
7. Dennis Reddy and John Coolen, "Electronic Communications", 4th Edition, Prentice Hall Publishers, 1995.
8. Electronic Communications Systems by Wayne Tomasi. Pearson education India.
9. Electronic Communication Systems by Roy Blake By Cengage learning.
10. Communication Systems By Simon Haykins By Wiley India.
11. Theory and Problem Of Electronic Communication By Lloyd Temes and Mitchel E.Schulz(Second edition), McGraw Hill Publication.
12. Kennedy, "Electronic Communications Systems", 4th Edition, McGraw-Hill Publishers, 1992.
13. Deshpande, N.D, "Communication Electronics", Tata McGraw Hill Publishers,1989.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SECA1303 – ANALOG COMMUNICATION SYSTEMS

SECA1303 – ANALOG COMMUNICATION SYSTEMS

UNIT 2 ANGLE (FM AND PM) MODULATION AND DEMODULATION

Single tone FM: Mathematical representation, frequency spectrum and bandwidth- Multi-tone FM - NBFM and WBFM -Phase modulation (PM): Mathematical representation - Conversion: FM to PM and PM to FM – Comparison of AM, FM and PM- FM Generation: Direct method using Varactor diode and indirect method (Armstrong modulator) - Pre-emphasis – FM transmitter. FM Detector: Balanced slope detector, Foster Seeley frequency discriminator and Ratio detector - De- emphasis.

Introduction to Angle Modulation

Basic definitions:

The other type of modulation in continuous-wave modulation is **Angle Modulation**. Angle Modulation is the process in which the frequency or the phase of the carrier signal varies according to the message signal.

Features of angle modulation:

•It can provide a better discrimination (robustness) against noise and interference than AM

•This improvement is achieved at the expense of increased transmission bandwidth

•In case of angle modulation, channel bandwidth may be exchanged for improved noise performance

•Such trade- off is not possible with AM

The standard equation of the angle modulated wave is

$$\boxed{s(t) = A_c \cos\theta_i(t)}$$

Where, Ac - is the amplitude of the modulated wave, which is the same as the amplitude of the carrier signal. θi(t) is the angle of the modulated wave

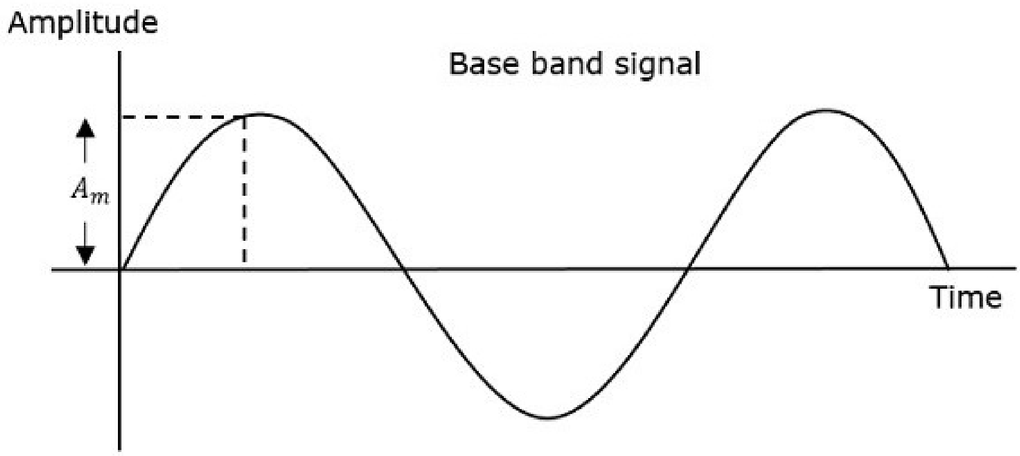
Angle modulation is further divided into frequency modulation and phase modulation.

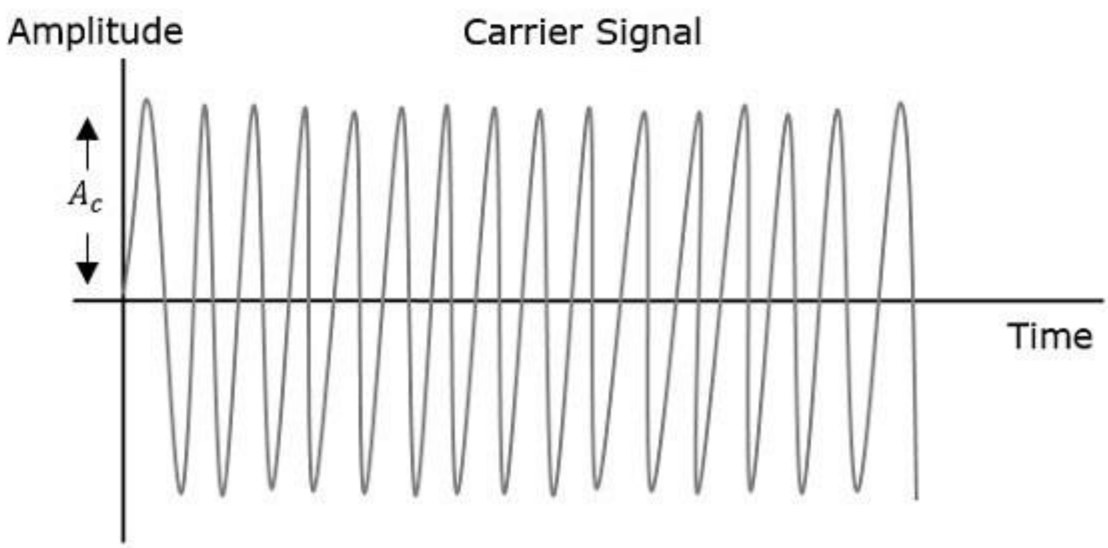
- **Frequency Modulation** is the process of varying the frequency of the carrier signal linearly with the message signal.
- **Phase Modulation** is the process of varying the phase of the carrier signal linearly with the message signal.

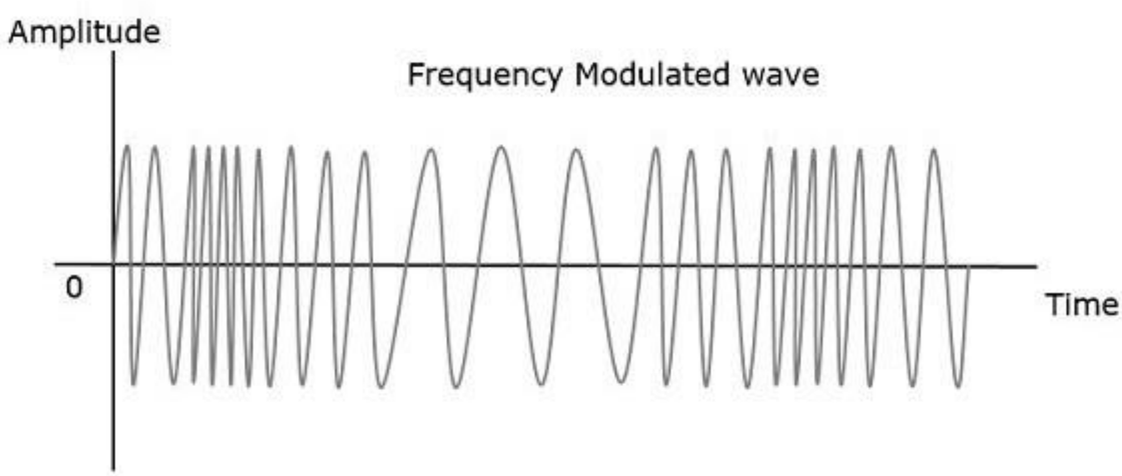
Now, let us discuss these in detail.

Frequency Modulation: **In amplitude modulation, the amplitude of the carrier signal varies. Whereas in "**Frequency Modulation" **,the frequency of the carrier signal varies in accordance with the instantaneous amplitude of the modulating signal.**

Hence, in frequency modulation, the amplitude and the phase of the carrier signal remains constant. This can be better understood by observing the following figures.







The frequency of the modulated wave increases, when the amplitude of the modulating or message wave increases. Similarly, the frequency of the modulated wave decreases, when the amplitude of the modulating signal decreases. Note that, the frequency of the modulated wave remains constant and it is equal to the frequency of the carrier signal, when the amplitude of the modulating signal is zero.

- There is yet another way of modulation namely the angle modulation in which the angle of the carrier wave changes in accordance with the signal
- In this method of modulation the amplitude of the carrier wave is maintained constant
- The advantage is it can show better discrimination against noise and interference than amplitude modulation

  Let $\theta_i(t)$ denote the angle of a modulated sinusoidal carrier
- $s(t) = A_c \cos[\theta_i(t)]$
- Where A_c is the carrier amplitude. A complete oscillation occurs whenever $\theta_i(t)$ changes by 2π radians. If $\theta_i(t)$ increases monotonically with time the average frequency in Hertz, over an interval from t to Δt

  $$f_{\Delta t}(t) = \frac{\theta_i(t+\Delta t) - \theta_i(t)}{2\pi\Delta t}$$
- We may thus define the instantaneous frequency of the angle-modulated signal s(t)

  $f_i = \lim_{\Delta t \to 0} f_{\Delta t}(t)$

  $$= \lim_{\Delta t \to 0} \frac{\theta_i(t+\Delta t) - \theta_i(t)}{2\pi\Delta t}$$

  $$= \frac{1}{2\pi}\frac{d\theta_i(t)}{dt}$$

- The angle modulated signal s(t) as a rotating phasor of length Ac and angle $\theta_i(t)$
- In simple case of unmodulated carrier the angle $\theta_i(t)$ is

  $\theta_i(t) = 2\pi fct + \emptyset c$

  And the corresponding phasor rotates with a constant angular velocity equal to $2\pi fc$
- There are infinite numbers of ways in which the angle $\theta_i(t)$ may be varied in some manner with the message signal. We consider only two methods phase modulation and frequency modulation

- MODULATION INDEX FOR FREQUENCY MODULATION

The frequency modulation index is the equivalent of the modulation index for AM, but obviously related to FM. In view of the differences between the two forms of modulation, the FM modulation index is measured in a different way.

The FM modulation index is equal to the ratio of the frequency deviation to the modulating frequency.

$$m = \frac{Frequency\ deviation}{Modulation\ frequency}$$

FM deviation ratio

Accordingly the FM deviation ratio can be defined as: the ratio of the maximum carrier frequency deviation to the highest audio modulating frequency.

$$m = \frac{Max\ frequency\ deviation}{Max\ modulation\ frequency}$$

There are two main classifications for frequency modulated signals and these can be related to the modulation index and deviation ratio.

- ***Wideband FM:*** Wideband FM is typical used for signals where the FM modulation index is above about 0.5. For these signals the sidebands beyond the first two terms are not insignificant. Broadcast FM stations use wide-band FM which enables them to transmit high quality audio, as well as other facilities like stereo, and other facilities like RDS, etc.

  The wide bandwidth of wide band FM is enables high quality broadcast transmissions to be made, combining a wide frequency response with low noise levels. Once the signal is sufficiently strong, the audio signal to noise ratio is very good.

  Sometimes high fidelity FM tuners may use a wide-band filter for strong signals to ensure the optimum fidelity and performance. Here the quieting effect of the strong signal will allow for wide-band reception and the full audio bandwidth. For for lower strength signals they may switch to a narrower filter to reduce the noise level, although this will result in the audio bandwidth being reduced. However on balance the narrower bandwidth will give a more pleasing sound when the received signal is low.

Narrowband FM: Narrow band FM, NBFM, is used for signals where the deviation is small enough that the terms in the Bessel function are small and the main sidebands are those appearing at ± modulation frequency. The sidebands further out are negligible.

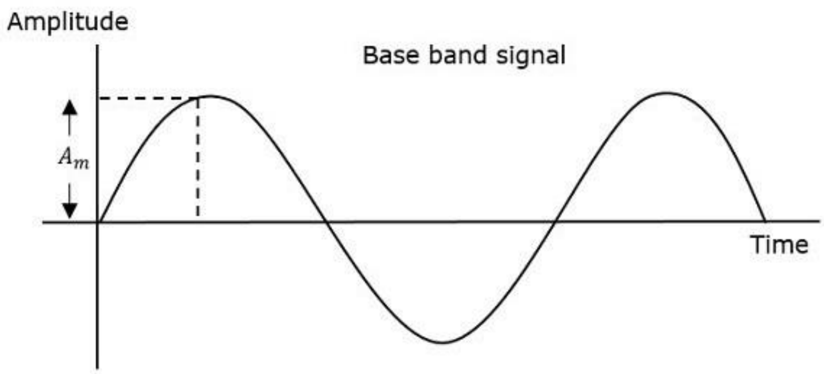
For NBFM, the FM modulation index must be less than 0.5, although a figure of 0.2 is often used. For NBFM the audio or data bandwidth is small, but this is acceptable for this type of communication.

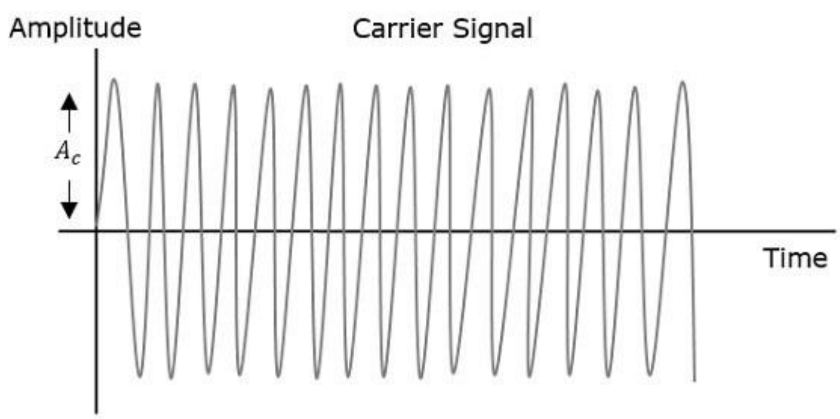
Narrowband FM is widely used for two way radio communications. Although digital technologies are taking over, NBFM is still widely used and very effective. Many two way radios or walkie talkies use NBFM, especially those which conform to the licence-free standards like PMR446 and FRS radio communications systems.

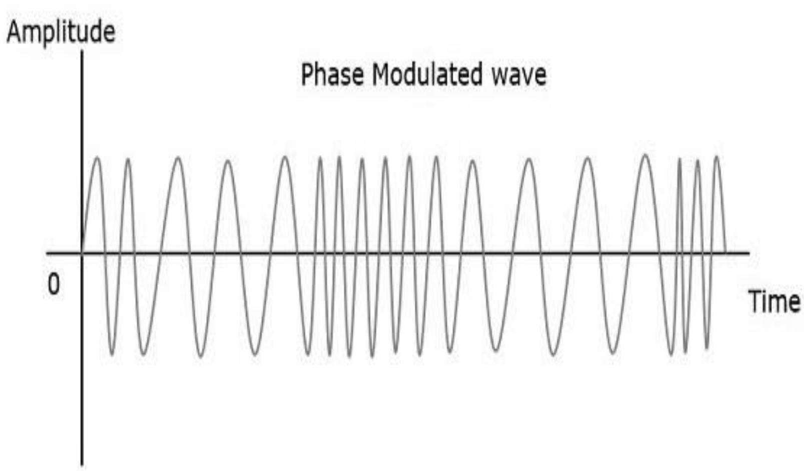
Phase Modulation

In frequency modulation, the frequency of the carrier varies. Whereas in "**Phase Modulation (PM)"** , the phase of the carrier signal varies in accordance with the instantaneous amplitude of the modulating signal.

So, in phase modulation, the amplitude and the frequency of the carrier signal remains constant. This can be better understood by observing the following figures.

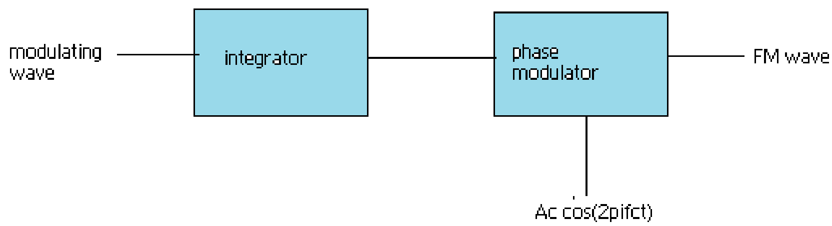






The phase of the modulated wave has got infinite points, where the phase shift in a wave can take place. The instantaneous amplitude of the modulating signal changes the phase of the carrier signal. When the amplitude is positive, the phase changes in one direction and if the amplitude is negative, the phase changes in the opposite direction.

Phase Modulation:



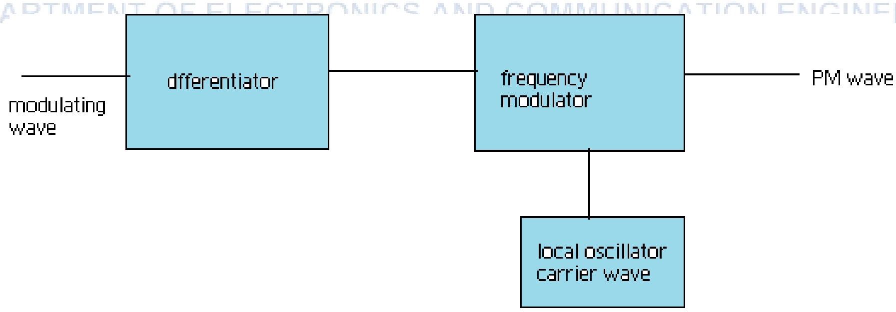
Phase modulation is that form of angle modulation in which the angle $\theta_i(t)$ is varied linearly with the message signal m(t)

$$\theta_i(t) = 2\pi fct + kpm(t)$$

- The term $2^{\pi f ct}$ represents the angle of the un-modulated carrier wave and constant Kp is the phase sensitivity of the modulator expressed in radian per volt
- We have assumed that the angle of the un-modulated carrier is zero at time t =0
- The phase –modulated signal s(t) is thud described in the time domain by

$$s(t) = A_c \cos 2\pi fct + Kpm(t)]$$

FREQUENCY MODULATION



Frequency modulation is that form of angle modulation in which the instantaneous frequency if(t) is varied linearly with the message signal m(t)

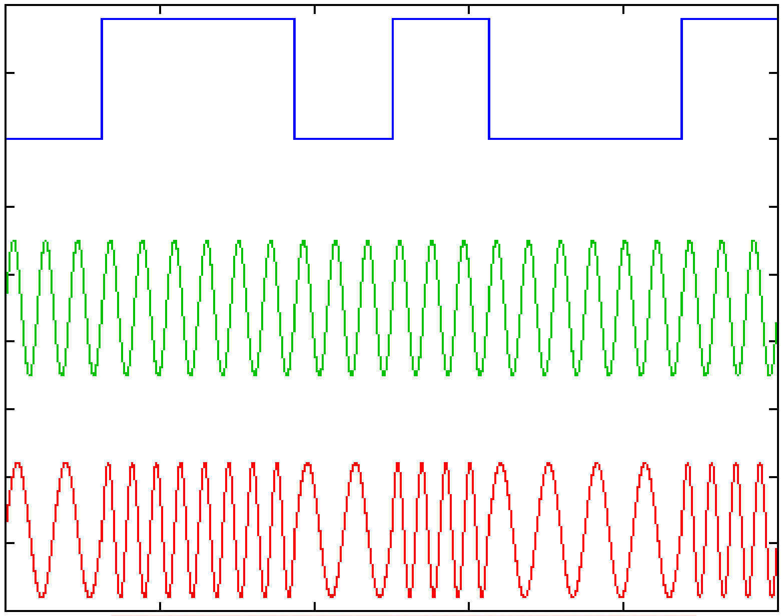
$$f_i = f_c + kfm(t)$$

$$\theta_i = 2\pi \int_0^t f_i(t)dt$$

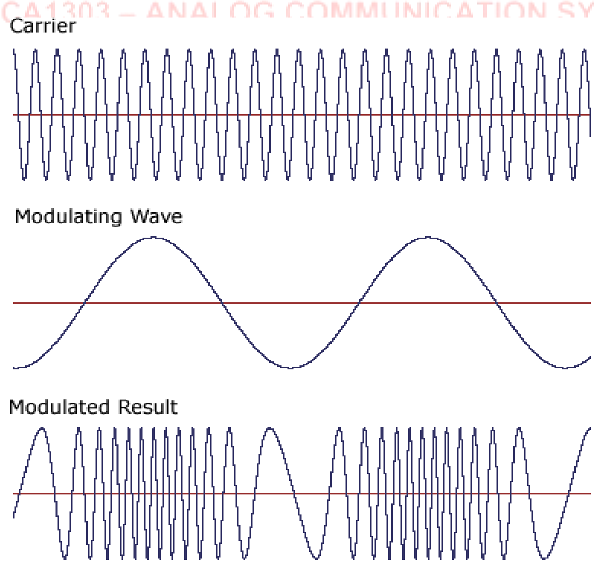
$$\theta_i = 2\pi fct + 2\pi kf \int_0^t m(t)dt$$

$$s(t) = A_c \cos[2\pi fct + 2\pi kf \int_0^t m(t)dt$$

Frequency of the carrier varies with the signal mathematically in the above equation



- Comparison between phase and frequency modulated equations frequency modulated signal is the same as phase modulation with the message signal integrated
- The variation of the frequency is discrete differing from the sinusoidal modulated wave where the frequency changes constantly
- There will be a phase discontinuity in case of phase modulation when message is a square wave
- There is a phase reversal in the phase modulation
- The visualization is easier

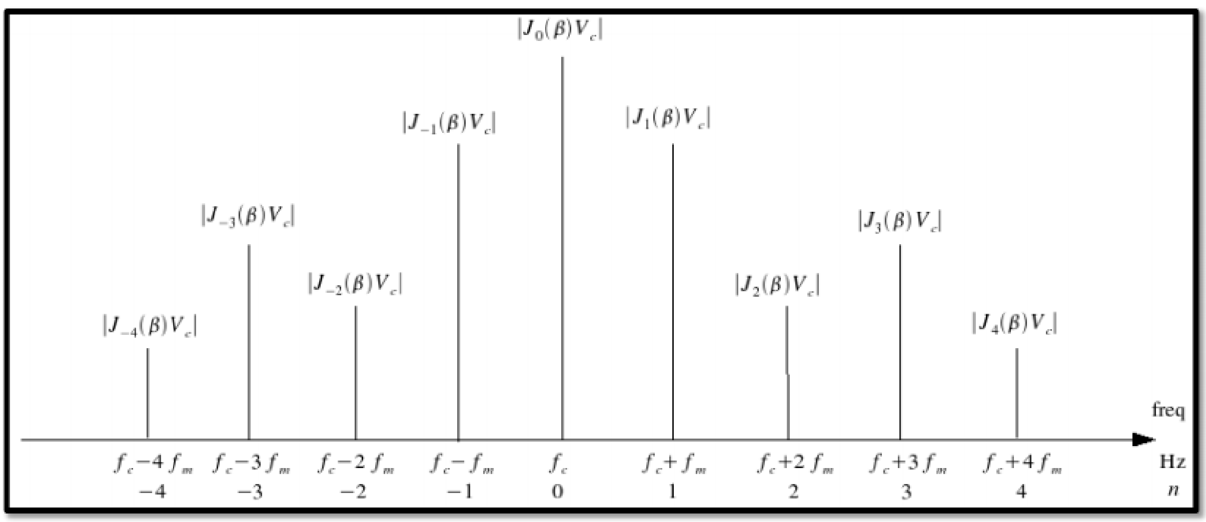


Note: The FM wave is a non-linear function of the modulating wave m(t)

TRANSMISSION BANDWIDTH OF FM

The number of significant sidebands ‘n’ produced in an FM waves can be obtained from the plot of Bessel function Jn(mf). • For n>mf, the values of Jn(mf) are negligible particularly when mf>>1. • Therefore the significant sidebands produced in wideband FM may be considered to be an integer approximately equal to mf • i.e., n approx. equal to mf>>1. • Frequency span of USB=LSB= nwm • Transmission bandwidth of FM wave is defined as the separation between the frequencies beyond which none of side frequencies is greater than 1% of the carrier amplitude obtained when modulation is removed. • B.W = 2 nwm rad/sec. where n= number of sidebands.

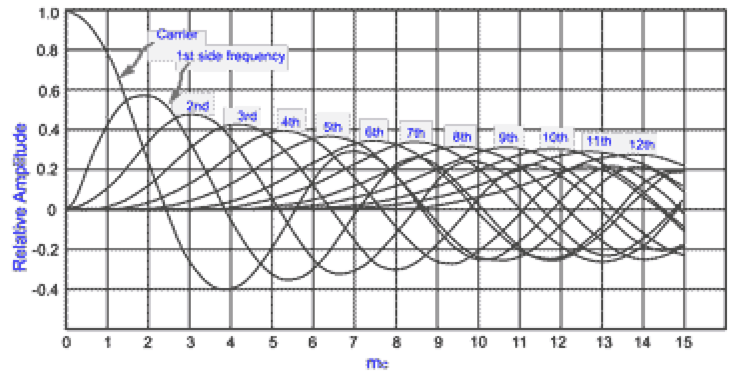
Thus the approx. B.W of a wide band FM system is given as Twice the frequency deviation. (for mf >>1) For smaller values of mf, B.W may be more than 2Dω.



FM SIGNAL SPECTRUM

Frequency modulation sidebands

The modulation of any carrier in any way produces sidebands. For amplitude modulated signals, the way in which these sidebands are created and their bandwidth and amplitude are quite straightforward. The situation for frequency modulated signals is rather different. The FM sidebands are dependent on both the level of deviation and the frequency of the modulation. In fact the total spectrum for a frequency modulated signal consists of the carrier plus an infinite number of sidebands spreading out on either side of the carrier at integral multiples of the modulating frequency. From the diagram it can be seen that the values for the levels of the sidebands rise and fall with varying values of deviation and modulating frequency.



Frequency modulation sideband levels

Carson's Rule for FM bandwidth

The bandwidth of an FM signal is not as straightforward to calculate as that of an AM signal.

A very useful rule of thumb used by many engineers to determine the bandwidth of an FM signal for radio broadcast and radio communications systems is known as Carson's Rule. This rule states that 98% of the signal power is contained within a bandwidth equal to the deviation frequency, plus the modulation frequency doubled. Carson's Rule can be expressed simply as a formula:

$BT=2(\Delta f+fm)$

Where: Δf = deviation

BT = total bandwidth (for 98% power)

fm = modulating frequency

To take the example of a typical broadcast FM signal that has a deviation of ±75kHz and a maximum modulation frequency of 15 kHz, the bandwidth of 98% of the power approximates to 2 (75 + 15) = 180kHz. To provide conveniently spaced channels 200 kHz is allowed for each station.The rule is also very useful when determining the bandwidth of many two way radio communications systems. These use narrow band FM, and it is particularly important that the sidebands do not cause interference to adjacent channels that may be occupied by other users.

Equations & calculation for FM sideband levels

Whilst it is very useful to have an understanding of the broad principles of the generation of sidebands within an FM signal, it is sometimes necessary to determine the levels mathematically.

The calculations are not nearly as simple as they are for amplitude modulated signals and they involve some long equations. It is for this reason that rules like Carson's rule are so useful as they provide workable approximations that are simple and straightforward to calculate, whist being sufficiently accurate for most radio communications applications.

The sideband levels can be calculated for a carrier modulated by a single sine wave using Bessel functions of the first kind as a function of modulation index.

The basic Bessel function equation is described below:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + \left(x^2 - \alpha^2\right) y = 0$$

Where:

α is an arbitrary complex number

In terms of the format of the equation, α and -α produce the same differential equation, but it is conventional to define different Bessel functions for these two values in such a way that the Bessel functions are mostly smooth functions of α.

Solving the Bessel equations to determine the levels of the individual sidebands can be quite complicated, but is ideal for solution using a computer.

By manipulating the mathematics, it is possible to solve the basic Bessel function equation and express it in the format:

$$e_{FM}(t) = E_C \left\{ \begin{array}{l} J_0(m_f) \sin \omega_c t - \\ J_1(m_f) [\sin(\omega_c - \omega_M) t - \sin(\omega_c + \omega_M) t] + \\ J_2(m_f) [\sin(\omega_c - 2\omega_M) t - \sin(\omega_c + 2\omega_M) t] - \\ J_3(m_f) [\sin(\omega_c - 3\omega_M) t - \sin(\omega_c + 3\omega_M) t] + \\ \cdots\cdots\cdots \end{array} \right\}$$

The way the series has expanded shows how the various sidebands are generated and how they extend out to infinity.

For large values of modulation index m_f, the FM wave ideally contains the carrier and an infinite number of sidebands located symmetrically around the carrier.

Such a FM wave has infinite bandwidth and hence called as wideband FM.

The modulation index of wideband FM is higher than 1.

The maximum permissible deviation is 75 kHz and it is used in the entertainment broadcasting applications such as FM radio, TV etc.

Frequency Spectrum of a Wideband FM wave

The expression for the wideband FM is complex since it is sine of sine function.

The only way to solve this equation is by using the Bessel functions. By using the Bessel functions the equation for wideband FM wave can be expanded as follows :

SINGLE TONE FREQUENCY MODULATION

Consider a single tone sinusoidal modulating wave

The instantaneous frequency is $$m(t) = A_m \cos(2\pi fmt)$$

$$f_i(t) = fc + k_f Am\, cos2\pi fmt$$
$$= fc + \Delta f\, cos2\pi fmt$$
$$\Delta f = k_f Am$$

Δf is called frequency deviation representing maximum departure of the instantaneous frequency from the carrier

The fundamental characteristics of an FM is that the frequency deviation is directly proportional to the amplitude of the base band signal and is independent of the modulation frequency Instantaneous angle is given by

$$\theta(t) = 2\pi \int_0^t fi(t)dt$$

$$= 2\pi fct + \frac{\Delta f}{fm} \sin 2\pi fmt$$

The modulation index of FM wave is given by

$$\beta = \frac{\Delta f}{fm}$$

SPECTRUM ANALYSIS OF SINUSOIDAL FM WAVE

$$s(t) = A_c \cos(2\pi fct) \cos[\beta \sin 2\,\pi fmt] - A_c \sin(2\pi fct) \sin[\beta \sin 2\,\pi fmt]$$

We can express this as in phase and quadrature components

$$s_i(t) = A_c \cos[\beta \sin 2\,\pi fmt]$$
$$s_q(t) = A_c \sin[\beta \sin 2\,\pi fmt]$$

Hence the complex envelope equals

$$s(t) = s_i(t) \quad s_q(t)$$

+

$$= A_c \exp\,[j\beta \sin 2\,\pi fmt]$$

And the FM signal is given by

$$s(t) = Re[s`(t) \exp(j2\pi fct)]$$

The complex envelope is a periodic sequence of time with the fundamental frequency equal to modulation frequency, we may therefore expand s`(t) in the form of a complex Fourier series as follows

$$s`(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi nfmt)$$

Where the complex Fourier co-efficient equals

$$c_n = fm\int_{-\frac{1}{2}fm}^{\frac{1}{2}fm} s`(t)\exp\left(-j2\pi nfmt\right)dt$$

$$= \frac{A_c}{2\pi}\int_{-\pi}^{\pi}\exp[j(\beta sinx - nx)]\,dx$$

$$x = 2\pi fmt$$

The integral on the right hand side is recognized as Bessel function of nth order and first kind and argument β, this function is commonly denoted by the symbol Jn(β)

$$Jn(\beta) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp[j\,(\beta\sin x - nx)]dx$$

Hence we can write $Cn = A_c J_n(\beta)$

$$s`(t) = A_c\sum_{n=-\infty}^{\infty} Jn(\beta)\exp(j2\pi nfmt)$$

$$s(t) = A_c Re[\sum_{n=-\infty}^{\infty} Jn(\beta)\exp(j2\pi(nfm + fc)t)]$$

On evaluation we get

$$s(t) = A_c\sum_{n=-\infty}^{\infty} Jn(\beta)cos(j2\pi(nfm + fc)t)]$$

Taking Fourier transform we obtain a discrete spectrum on both sides

$$S(j`) = A_c\sum_{n=-\infty}^{\infty} Jn(\beta)[\partial(f - fc - nfm) + \partial(f + fc + nfm)]$$

The plots of Bessel function shows us that for a fixed n $Jn(\beta)$ alternates between positive and negative values for increasing β and $|Jn(\beta)|$ approaches infinity. Note that for a fixed value of β

$$J_{-n}(\beta) = \begin{Bmatrix} J_n(\beta) & n\ even \\ -J_n(\beta) & n\ odd \end{Bmatrix}$$

We only need the positive values hence those are considered

The following are the properties of FM waves

PROPERTY 1:

For small values of the modulation index compared to one radian, the FM wave assumes a narrow band form and consisting essentially of a carrier, an upper side frequency and a lower side frequency component

This property follows from the fact that for small values of (β) we have

$$J_0(\beta) = 1$$

$$J_1(\beta) = \frac{(\beta)}{2}$$

$$J_n(\beta) = 0 \quad n > 1$$

These are the approximations assumed for

$\beta \leq 0.3$

The FM wave can be approximated as a sum of carrier an upper side frequency of amplitude and a lower side frequency component and phase shift equals to 180

PROPERTY 2:

For large values of modulation index compared to one radian the FM contains a carrier and an infinite number of side bands on either side located symmetrically around the carrier

Note that the amplitude of the carrier component in a wide band FM wave varies with the modulation index in accordance with $J_0(\beta)$

PROPERTY 3:

The envelope of an FM wave is constant so that the average power of such a wave dissipated in 1Ω resistor is also a constant

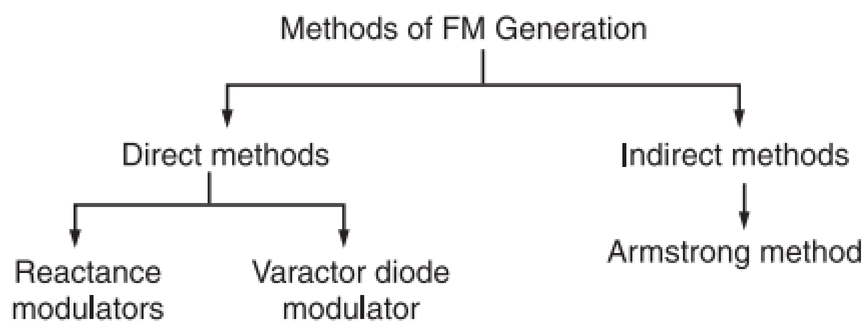
The average power is equal to the power of the carrier component it is given by

$$P = \frac{1}{2} A_c^2$$

The average power of a single tone FM wave s(t) may be expressed in the form of a corresponding series as

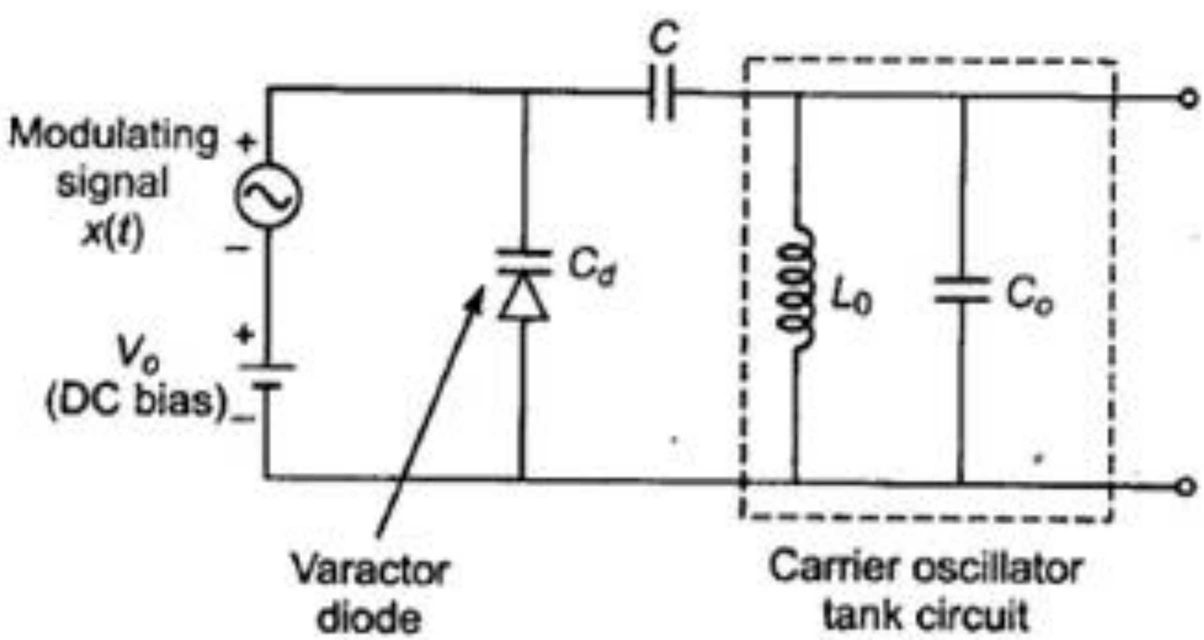
$$P = \frac{1}{2} A_c^2 \sum_{-\infty}^{\infty} J_n^2(\beta)$$

GENERATION OF FM WAVES:



There are two types
Direct method and indirect method
In the indirect method of producing a narrow band FM is generated then frequency multiplied to obtain wide band frequency modulated wave

VARACTOR DIODE MODULATOR



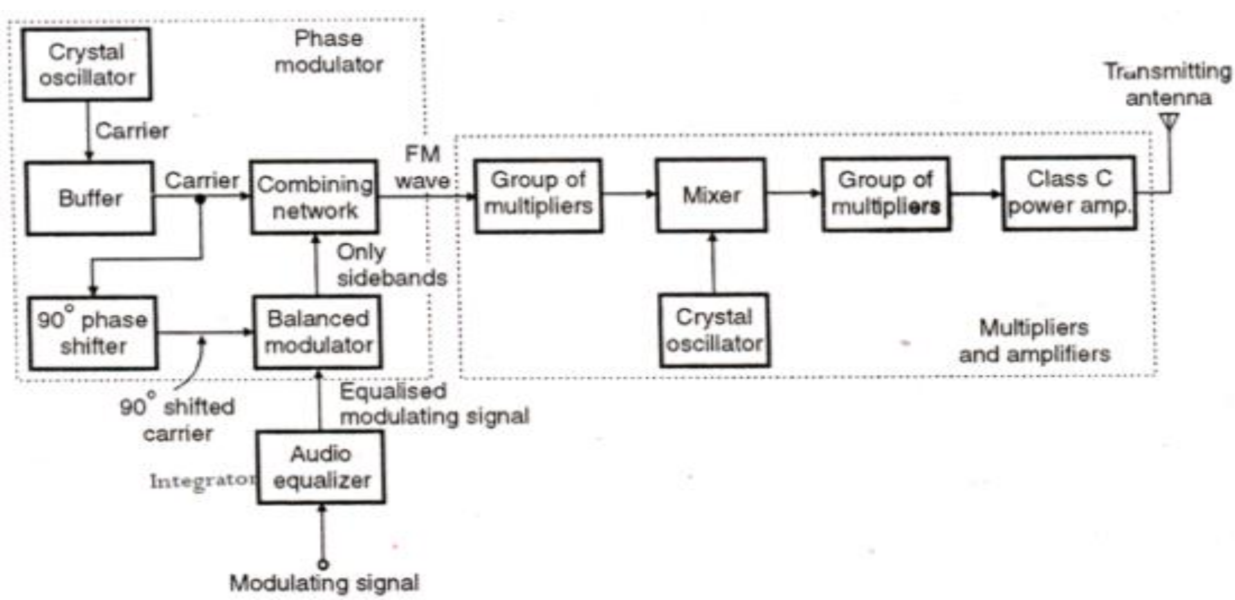
Varactor diode method for FM generation

A varactor diode is a semiconductor diode whose junction capacitance varies linearly with the applied bias andThe varactor diode must be reverse biased.

Working Operation

The varactor diode is reverse biased by the negative dc source $-V_b$.
The modulating AF voltage appears in series with the negative supply voltage. Hence, the voltage applied across the varactor diode varies in proportion with the modulating voltage. This will vary the junction capacitance of the varactor diode. The varactor diode appears in parallel with the oscillator tuned circuit. Hence the oscillator frequency will change with change in varactor diode capacitance and FM wave is produced. The RFC will connect the dc and modulating signal to the varactor diode but it offers a very high impedance at high oscillator frequency. Therefore, the oscillator circuit is isolated from the dc bias and modulating signal.

ARMSTRONG METHOD OF FM GENERATION



The crystal oscillator generates the carrier at low frequency typically at 1MHz. This is applied to the combining network and a 90° phase shifter.

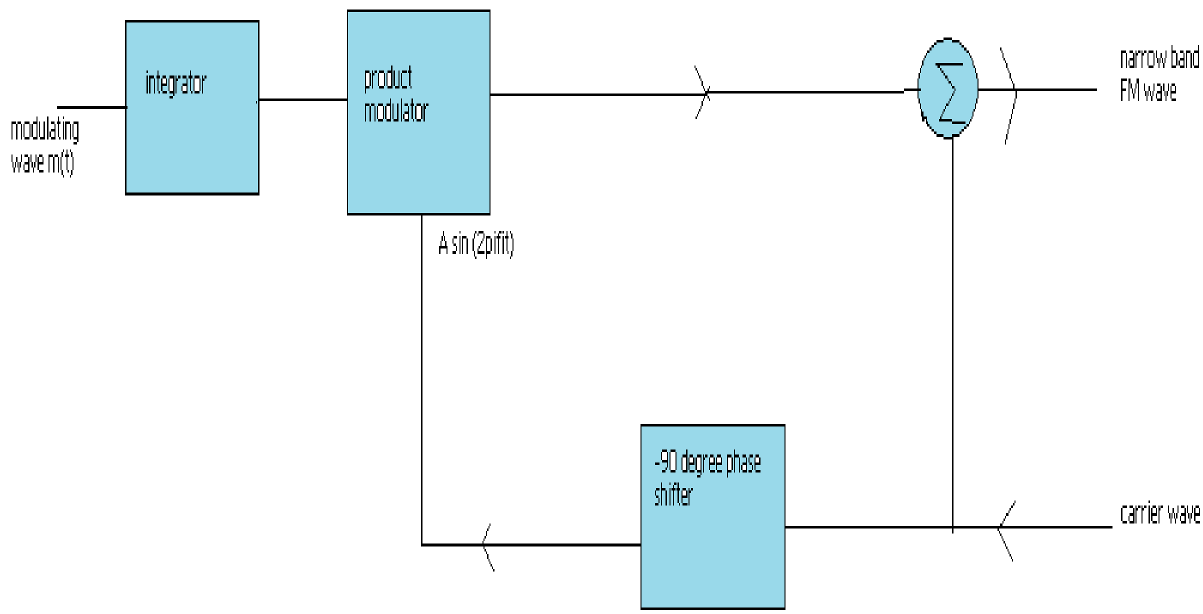
- The modulating signal is passed through an audio equalizer to boost the low modulating frequencies .The modulating signal is then applied to a balanced modulator.
- The balanced modulator produced two side bands such that their resultant is 90° phase shifted with respect to the un-modulated carrier.
- The un-modulated carrier and 90° phase shifted sidebands are added in the combining network.
- At the output of the combining network we get *FM* wave. This wave has a low carrier frequency *fc* and low value of the modulation index *mf*.

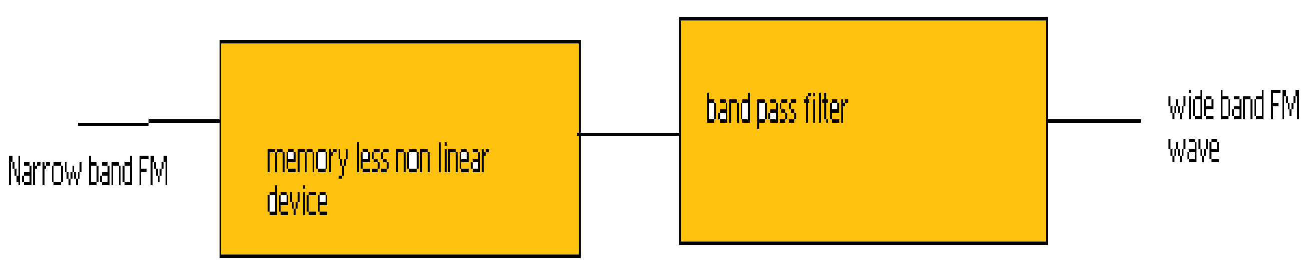
- *The carrier frequency and the modulation index are then raised by passing the FM wave through the first group of multipliers. The carrier frequency is then raised by using a mixer and then the fc and "mf", both are raised to required high values using the second group of multipliers.*

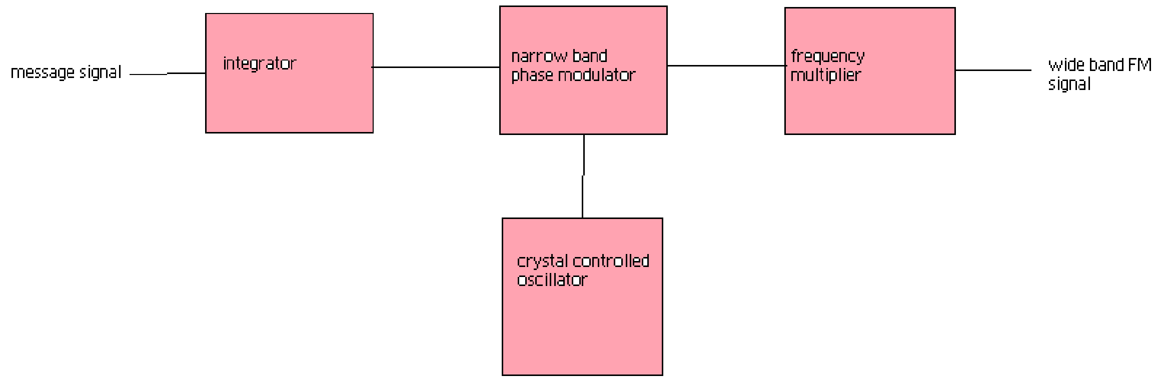
- *The FM signal with high fc and high mf is then passed through a class C power amplifier to raise the power level of the FM signal.*

- *The Armstrong method uses the phase modulation to generate frequency modulation. This method can be understood by dividing it into four parts as follows:*

INDIRECT METHOD:





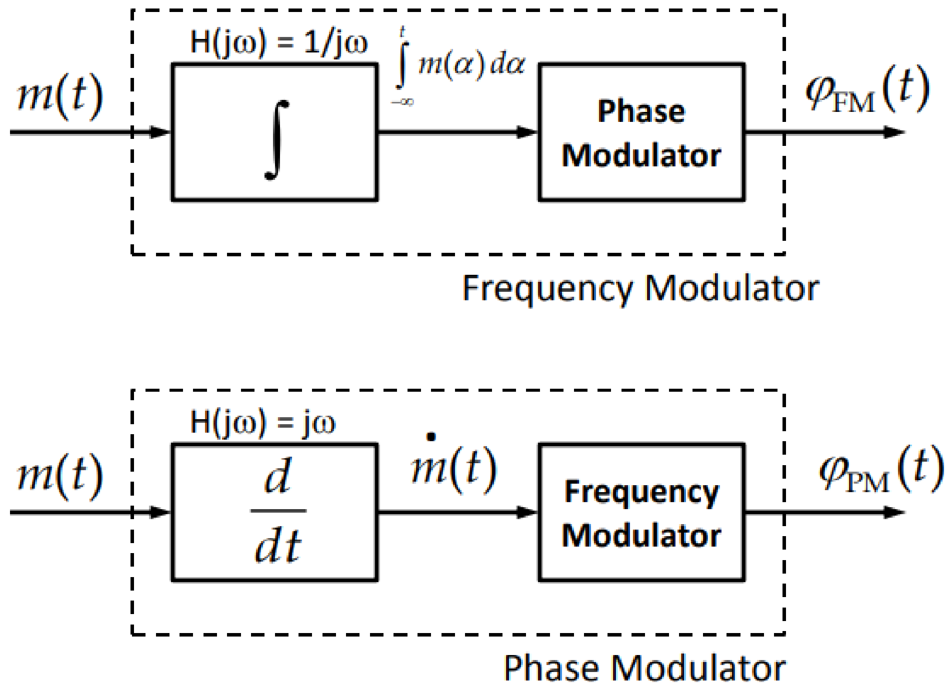


FM Modulation : The amplitude of the modulated carrier is held constant and the time derivative of the phase of the carrier is varied linearly with the information signal. Hence, NBFM signal can be generated using phase modulator circuit as shown.

To obtain WBFM signal, the output of the modulator circuit (NBFM) is fed into frequency multiplier circuit and the mixer circuit.

The function of the frequency multiplier is to increase the frequency deviation or modulation index so that WBFM can be generated. The instantaneous value of the carrier frequency is increased by N times.

CONVERSION OF FM TO PM AND PM TO FM



We require that H(jω) be a reversible (or invertible) operation so that m(t) is recoverable.

FM signals can be **generated** using either direct or indirect frequency modulation: Direct **FM** modulation can be achieved by directly feeding the message into the input of a voltage-controlled oscillator. For indirect **FM** modulation, the message signal is integrated to **generate** a phase-modulated signal.

Phase modulation (**PM**) is a modulation pattern for conditioning communication signals for transmission. It encodes a message signal as variations in the instantaneous phase of a carrier wave. Phase modulation is one of the two principal forms of angle modulation, together with frequency modulation.The phase of a carrier signal is modulated to follow the changing signal level (amplitude) of the message signal. The peak amplitude and the frequency of the carrier signal are maintained constant, but as the amplitude of the message signal changes, the phase of the carrier changes correspondingly. Phase modulation is widely used for transmitting radio waves and is an integral part of many digital transmission coding schemes that underlie a wide range of technologies like Wi-Fi, GSM and satellite television. PM is used for signal and waveform generation in digital synthesizers, such as the Yamaha DX7, to implement FM synthesis. A related type of sound synthesis called phase distortion is used in the Casio CZ synthesizers.

The change in phase, changes the frequency of the modulated wave. The frequency of the wave also changes the phase of the wave. ... Phase modulation is an indirect method of producing **FM**. The amount of frequency shift, **produced** by a phase modulator increases with the modulating frequency.

COMPARISON BETWEEN AM AND FM

Sr. No.	FM	AM
1	FM receivers are immune to noise	AM receivers are not immune to noise
2	It is possible to decrease noise by increasing deviation	This feature is absent in AM
3	Bandwidth is higher and depends on modulation index	Bandwidth is lower compared to AM but independent of modulation index
4	FM transmission and reception equipment are more complex	FM transmission and reception equipment are less complex
5	All transmitted power is useful	Carrier power and one sideband power is useless

COMPARISON OF FM AND PM

#	Frequency Modulation (FM)	Phase Modulation (PM)
1	Frequency deviation is proportional to modulating signal $m(t)$	Phase deviation is proportional to modulating signal $m(t)$
2	Noise immunity is superior to PM (and of course AM)	Noise immunity better than AM but not FM
3	Signal-to-noise ratio (SNR) is better than in PM	Signal-to-noise ratio (SNR) is not as good as in FM
4	FM is widely used for commercial broadcast radio (88 MHz to 108 MHz)	PM is primarily for some mobile radio services
5	Modulation index is proportional to modulating signal $m(t)$ as well as modulating frequency f_m	Modulation index is proportional to modulating signal $m(t)$

FM TRANSMITTERS

- The frequency modulated wave can be produced by two methods namely;
  - (i) Directly modulated FM transmitter.
  - (ii) Indirectly modulated FM transmitter.
- In a directly modulated FM transmitter the modulating system directly produce FM waves by varying the master oscillator frequency.
- Alternately, indirectly modulated FM transmitter generates the phase modulated signals. The PM wave is then converted into frequency modulated wave.
- The basic difference between the two circuits is that, the first circuit employs an LC circuit in master oscillator and its frequency is likely to change with changes in circuit parameters, the second circuit uses the crystal oscillator as a master oscillator, thus it gives a drift free frequency.
- As such, the first circuit employs some form of AFC, and it produces more frequency deviation and requires less number of frequency multiplier stages. The phase modulator circuit produces smaller frequency deviation and requires more number of frequency multiplier stages.

Direct FM Transmitter (Frequency modulated transmitter using reactance tube modulators)

- The transmitter employs a reactance tube modulator to produce a frequency deviation in proportion to the signal amplitude. Alternatively a varactor diode modulator may also be employed for this purpose.
- The resulting FM is passed through a number of frequency multiplier stages. These stages not only raise the centre frequency of the signal and the frequency deviation also is multiplied by the same factor.

- A part of the output of the frequency multiplier stages is passed to AFC as shown in figure . The purpose of this circuit is to make corrections in the centre frequency of the transmitter if any frequency drift takes place due to change in the circuit parameters.
- Signal from multiplier stages is mixed with local oscillator frequency, the output of mixer is the difference frequency and is fed to a discriminator which gives dc output according to frequency shift w.r.to centre frequency.

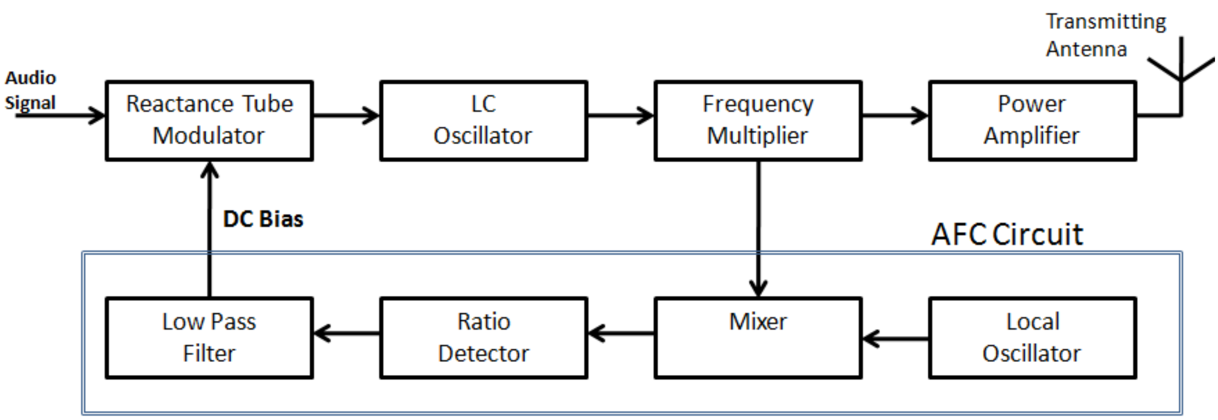


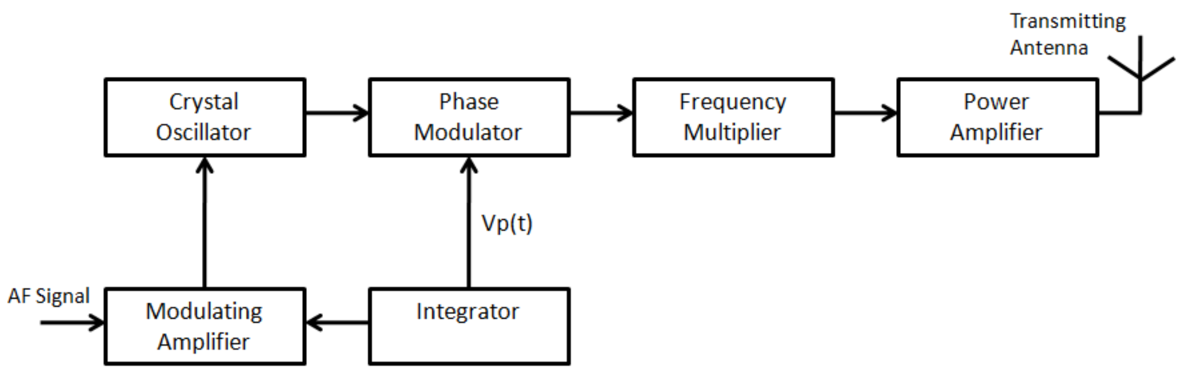
Fig. Direct method of FM generation

- When the frequency of the transmitter is exactly equal to centre frequency, discriminator output is zero and there is no dc correcting bias. Any positive or negative drift in the frequency produces a corresponding, correction bias at the discriminator which when applied to the reactance tube modulator brings LC master oscillator frequency back to its centre value.
- If any modulating signal is present, at the output of mixer, resulting AP signal produced at the discriminator output is not allowed to reach the reactance tube modulator because of low pass filter which has a cutoff lower than the AF signal
- The modulated wave is then amplified to the required power level by class 'C' power amplifier stages and then transmitted through antenna.

INDIRECT METHOD

(i) Indirect method FM transmitter [PM transmitter]

- The block diagram of indirect method of FM transmitter is shown in Figure. It consists of a crystal oscillator, the output of which is led to a phase modulator. The audio signal is integrated and also applied to phase modulator.
- The resultant wave at the phase modulator output is passed through several stages of frequency multipliers to obtain the desired frequency deviation and increase the centre frequency. The signal is amplified by the power amplifier stage to the required power level.



- To understand the circuit action, in the production of PM waves, assume an audio signal $V_m \sin\omega t$ at the input of the transmitter. The output of the, phase modulator is given by the integrator.

$$V_p(t) = \int V_m\, Sin\omega_m t\, dt = \frac{-V_m\, Cos\omega_m t}{\omega_m}$$

The phase shift produced by this signal at the modulator output is given by

$$\theta \propto V_p\,(t) \quad or \quad \theta = KV_p\,(t) = \frac{-KV_m\, Cos\omega_m t}{\omega_m}$$

We know that

$$\omega_f = \frac{d\theta}{dt} = -\frac{d}{dt}\left[\frac{KV_m\, Cos\,\omega_m t}{\omega_m}\right] = KV_m\, Sin\,\omega_m t$$

Frequency deviation Af = KVm.

Thus, the circuit results in frequency modulation with deviation proportional to the peak amplitude of the modulating signal.

Armstrong FM transmitter

- The carrier frequency is generated with the help of a crystal oscillator. This method utilizes a balanced modulator with audio signal and carrier signal with 90° phase shift as shown in figure.
- The balanced modulator output gives DSB-SC-AM. The frequency of the side bands is increased in a harmonic generator stages and fed to mixer stages the other input to this stage being the carrier signal after passing through another harmonic generator.
- The different frequency components at the mixer output are the carrier and side band frequencies. This output is again multiplied by a number of frequency multiplier stages, raised to the required power level by using power amplifier and then it is transmitted.

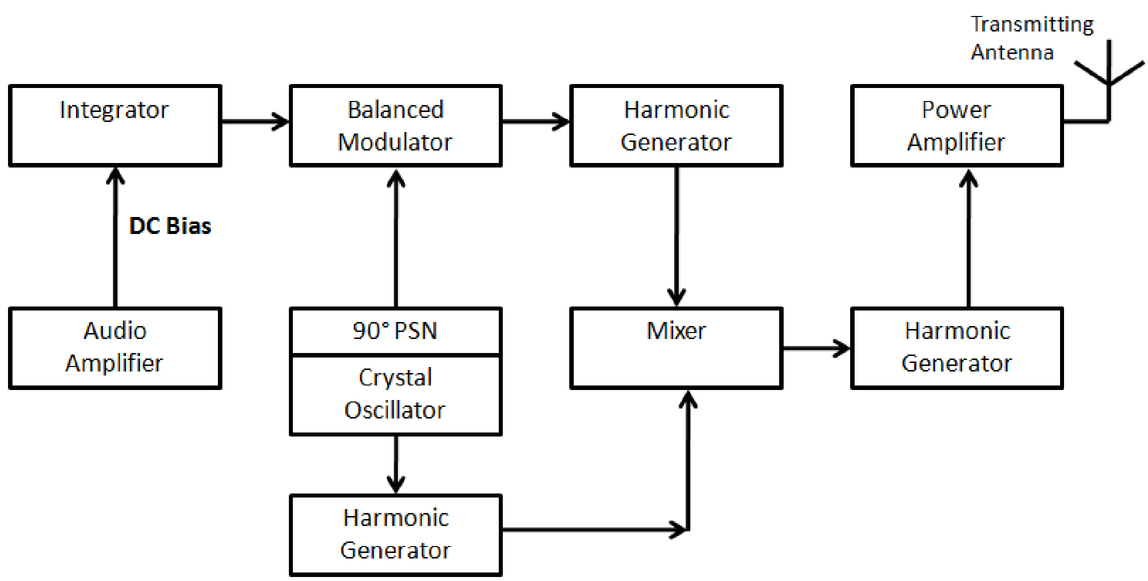


Fig 15 Armstrong method of FM transmitter

Let the modulating voltage $V_m(t) = V_m \, Sin \, \omega_m(t)$

The carrier voltage $V_c(t) = V_c \, Sin \, (\omega_c + \theta)t = V_c \, Cos \, \omega_c t \quad if \; \theta = 90°$

Audio input to the balanced modulator is

$$\int V_m \, Sin \, \omega_m(t) = \frac{V_m}{\omega_m} \mathrm{Cos} \, \omega_m t$$

Balanced modulator output is

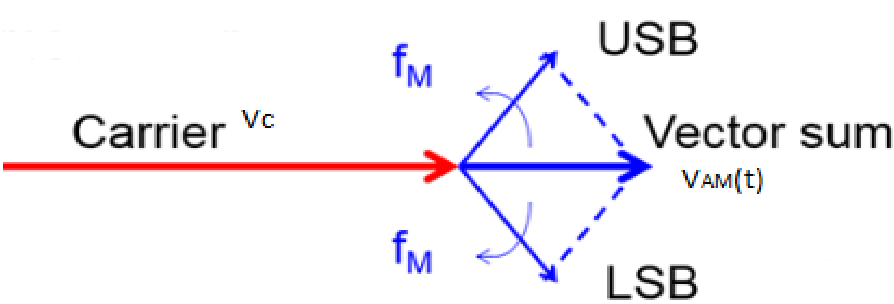
$$= V_c \, Cos \, \omega_c t * \frac{V_m}{\omega_m} \cos \omega_m t = \frac{V_c V_m}{\omega_m} Cos \, \omega_c t \, Cos \, \omega_m t$$

$$= \frac{V_c V_m}{2\omega_m} [Cos \, (\omega_c + \omega_m)t + Cos \, (\omega_c - \omega_m)t]$$

Output of the mixer

$$= V_c \, Sin \, \omega_c t + \frac{V_c V_m}{2\omega_m} [Cos \, (\omega_c + \omega_m)t + Cos \, (\omega_c - \omega_m)t]$$

Phasor diagram of the above equation (AM)



Peak phase deviation $\emptyset_p = \tan^{-1} \frac{V_c V_m}{2\omega_m}$

- The mixer output contains a phase deviation φp and it also contains an amplitude modulation component. However if the sideband components have very low amplitude, the amplitude modulation is negligible. The frequency deviation produced by the system equals.

$$\Delta f = \emptyset_P f_m = \frac{V_c V_m}{\omega_m} f_m = \frac{V_c V_m}{2\pi}$$

The above equation shows that the frequency deviation produced by the system is directly proportional to the magnitude of the modulating signal.

PRINCIPLE OF FM DETECTORS

- The process of extracting modulating signal from a frequency modulated carrier is known as frequency demodulation or detection. The electronic circuits that perform the demodulation process are called the FM Detectors.
- The FM detectors perform the detection in two steps.
  - (i) It converts the frequency modulated signal into its corresponding amplitude modulated signal by using frequency dependent circuits i.e., circuits whose output voltage depends on input frequency from which original modulating signal is detected, such circuits are called frequency discriminators.
  - (ii) The original modulating signal is recovered from this AM signal (converted from FM to AM in previous step) by using a linear diode envelope detector.

- FM discriminators are obtained by using a simple RC and LC combination. This discriminator suffers from threshold effect in the presence of excessive noise. The threshold can be improved by using pre-emphasis and de-emphasis circuits.
- FM detectors or discriminators can be divided into two types

  Slope detectors and phase discriminators

a. Slope detectors:

The principle of operation depends on the slope of the frequency response characteristics of a frequency selective network. Two main FM discriminators which used to tune resonant circuits comes under this categories as follows: Single tuned discriminator or Slope detector

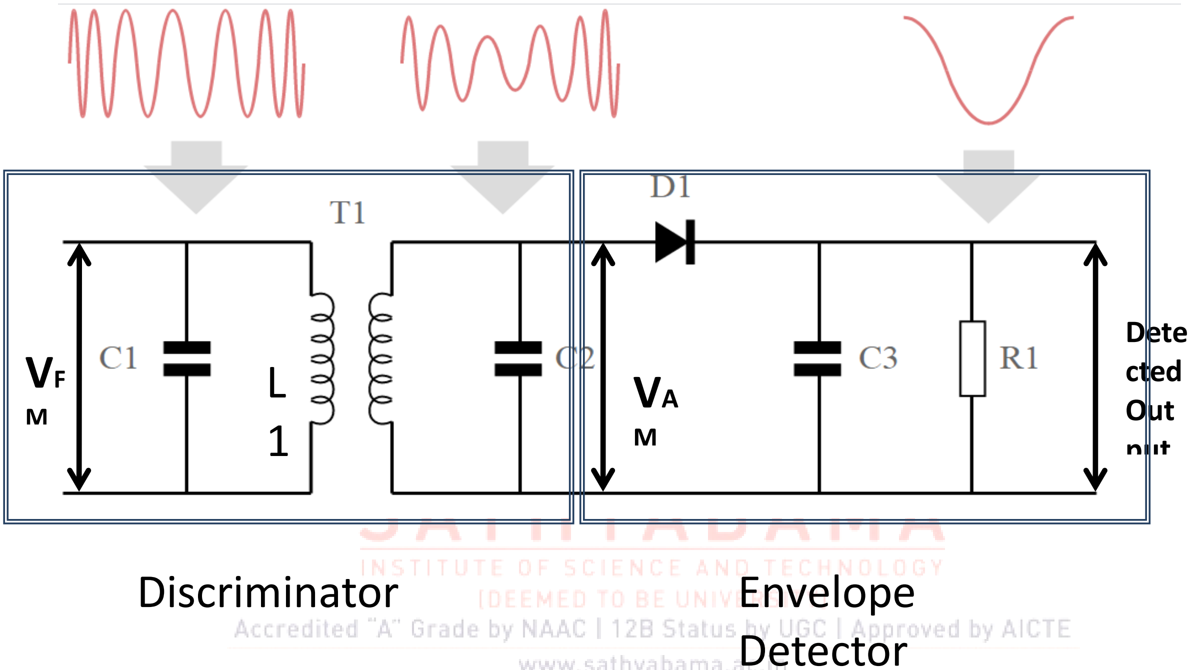
(i) Stagger tuned discriminator or Balanced Slope detector

b. Phase discriminators

There is two types of phase discriminators which are used to extract the original modulating signal from the frequency modulated wave. They are (i) Foster Seely discriminators

(i) Ratio detector

Slope detectors:



The figure (a) shows the circuit of slope detector

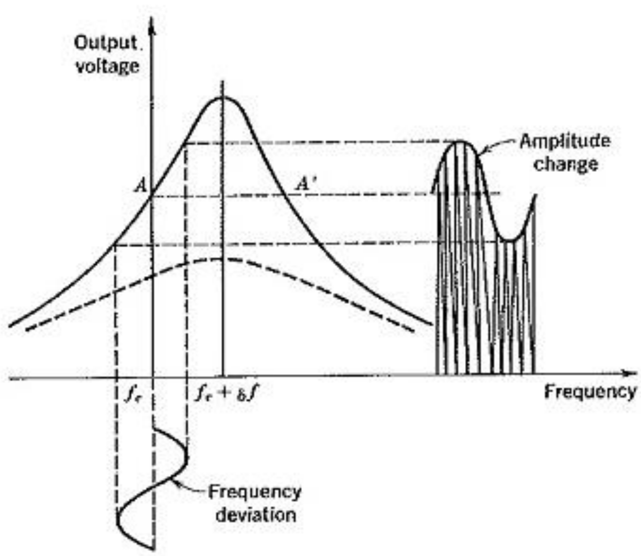


Figure (b) : Characteristics of Slope detector

- It consists of a parallel LC tuned circuit which acts as frequency discriminator.
- It is slightly detuned from the carrier frequency ωc.
- A low frequency deviation produces small amplitude variation, while a high frequency deviation produces large amplitude variation through this action the FM signal is changed to AM signal.
- Thus the AM signal is detected by a diode detector followed by the discriminator circuit.
- The frequency response of this detuned input is shown in figure (b). The slope of the characteristics curve is given as

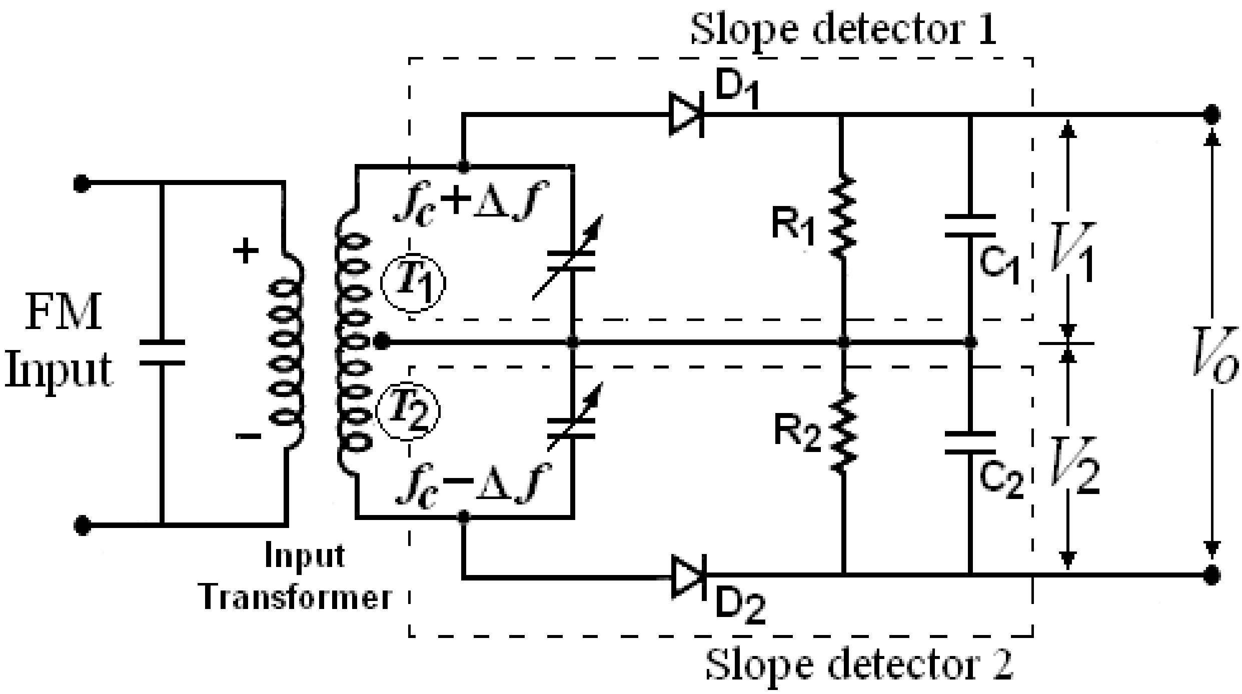
$$\theta = \frac{dV_{AM}}{d\omega}$$

- The advantage of this detector is its simplicity in construction and cheapness.
- The main disadvantages are:
  - (i) The non linear characteristic of the circuit causes a harmonic distortion. The non linearity is obvious from the fact that the slope is not same at every point of the characteristic curve.
  - (ii) It does not eliminate the amplitude variation and the output is sensitive to any amplitude variation in the input FM signal which is not a desirable feature.

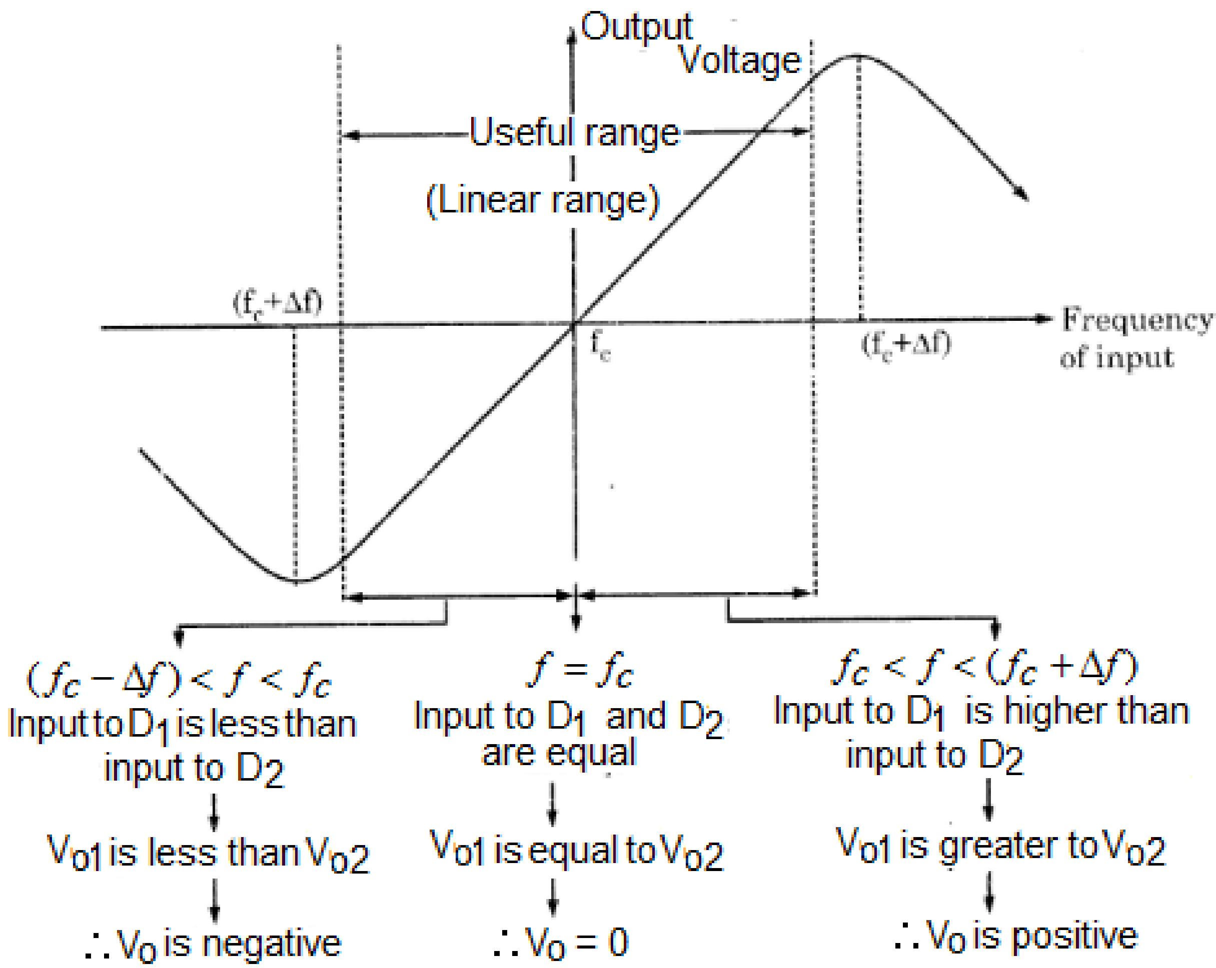
Balanced Slope detector

- The balanced slope detector, also known as Travis detector, is a combination of two slope detectors. They are connected to the opposite ends of the secondary of the transformer, hence fed 180 degree out of phase.
- The primary coil of the transformer is tuned at central frequency fc (carrier frequency) of the incoming signal. The upper half of the secondary coil T' is tuned at fc + δf and the lower half T" is tuned at fc – δf, where δf is higher in value than the largest frequency deviation in the incoming FM signal to make sure that the entire range of frequency variation in the incoming signal falls in the linear part of selectivity curve of the tuned circuit.

- When the input frequency is equal to fc, the voltage across T' is Vo Fig a. A similar condition exists across T" at this frequency producing voltage Vo' which happens to be equal to Vo as fc lies as much away from fc + δf as it is from fc – δf.



- Hence the voltages applied to the two diodes are equal leading to equal but opposite currents across the resistors R1 and R2.
- So, the output voltage will be zero as it is the difference of these two voltages.
- When the input frequency is higher than the carrier frequency fc, the voltage across T' is V1 (Fig a) and the voltage across T" at this frequency is V1 ' (Fig b).
- As can be seen from the above figs, V1 > V1 '. The current in the diode D1 is greater than that in D2 leading to positive output voltage for fi >fc
- When the input frequency is lower than the carrier frequency fc, the voltage across T' is V2 (Fig a) and voltage across T" is V2 ' (Fig b). As can be seen from the above figs, V2



- The output voltage will be positive or negative depending on which side of fc the input frequency happens to lie.
- If the input frequency goes outside the prescribed range, the output will start falling. The S-shaped frequency response shown in the above fig. is obtained.
- The main disadvantage of Balanced modulator is to

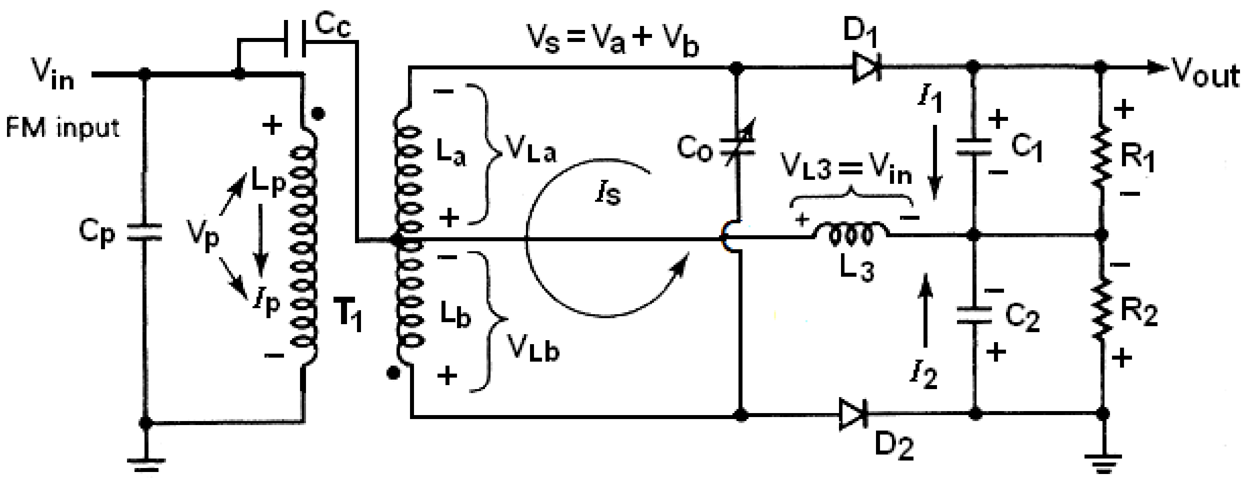
  Manage three resonant frequencies in the primary and secondary of the transformer.

  Though linearity in frequency response is better than that of slope detector, it is not good enough.

  Amplitude limiting is not provided.

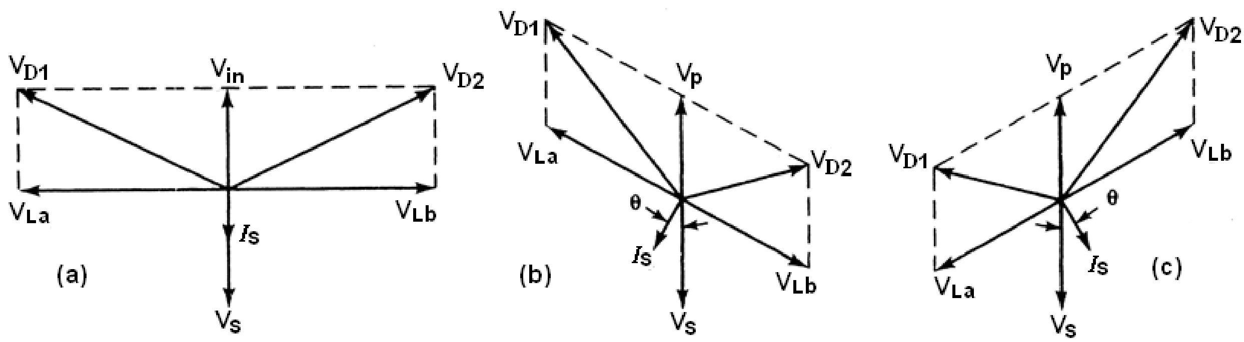
Foster-Seeley Discriminator (Phase Discriminator)

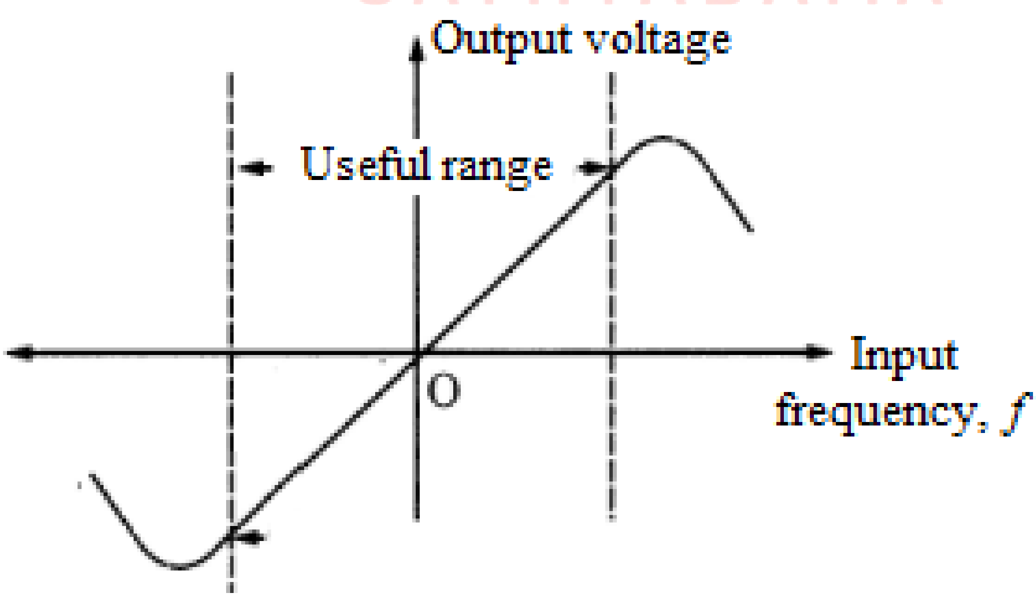
- It is also known as the PHASE-SHIFT DISCRIMINATOR.
- It uses a double-tuned RF transformer to convert frequency variations in the received fm signal to amplitude variations.
- These amplitude variations are then rectified and filtered to provide a dc output voltage. This voltage varies in both amplitude and polarity as the input signal varies in frequency.



- Fig. shows a typical Foster-Seeley discriminator. The primary tank circuit consists of C1 and L1. C2 and L2 form the secondary tank circuit. Both tank circuits are tuned to the center frequency of the incoming fm signal.
- Choke L3 is the dc return path for diode rectifiers D1 and D2. Resistors R3 and R4 are the load resistors and are bypassed by C3 and C4 to remove rf.
- To obtain the different phased signals a connection is made to the primary side of the transformer using a capacitor, and this is taken to the centre tap of the transformer. This gives a signal that is 90° out of phase.
- When an un-modulated carrier is applied at the centre frequency, both diodes conduct, to produce equal and opposite voltages across their respective load resistors. These voltages cancel each one another out at the output so that no voltage is present.
- As the carrier moves off to one side of the centre frequency the balance condition is destroyed, and one diode conducts more than the other. This results in the voltage across one of the resistors being larger than the other, and a resulting voltage at the output corresponding to the modulation on the incoming signal.

- The choke is required in the circuit to ensure that no RF signals appear at the output. The capacitors C1 and C2 provide a similar filtering function.
- The operation of the Foster-Seeley discriminator can best be explained using vector diagrams that show phase relationships between the voltages and currents in the circuit. Let's look at the phase relationships when the input frequency is equal to the center frequency of the resonant tank circuit.





- The output voltage is 0 when the input frequency is equal to the carrier frequency (FR).
- When the input frequency rises above the center frequency, the output increases in the positive direction. When the input frequency drops below the center frequency, the output increases in the negative direction.

- The output of the Foster-Seeley discriminator is affected not only by the input frequency, but also to a certain extent by the input amplitude. Therefore, using limiter stages before the detector is necessary.

Advantages of Foster-Seeley FM discriminator:

- Offers good level of performance and reasonable linearity.
- Simple to construct using discrete components.
- Provides higher output than the ratio detector
- Provides a more linear output, i.e. lower distortion than the ratio detector

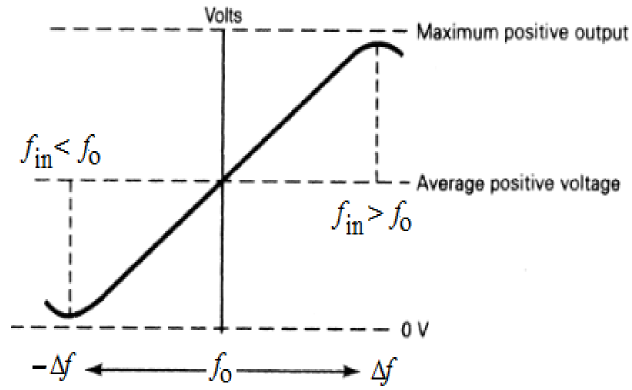
Disadvantages of Foster-Seeley FM discriminator:

- Does not easily lend itself to being incorporated within an integrated circuit.
- High cost of transformer.
- Narrower bandwidth than the ratio detector

Ratio Detector

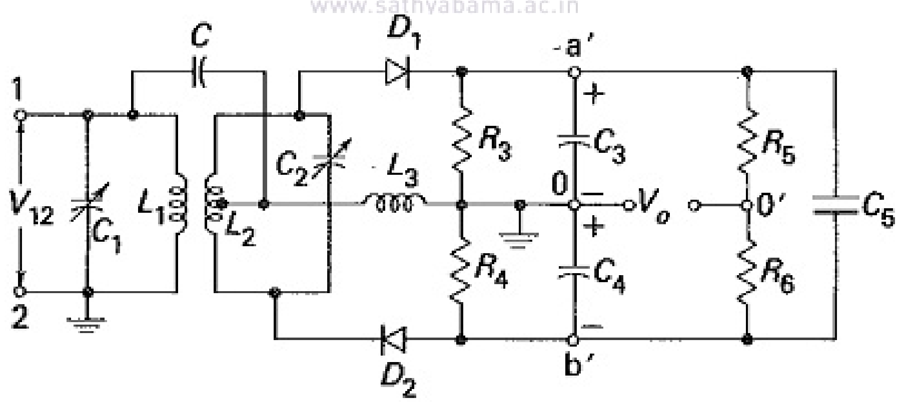
In the Foster-Seeley discriminator, changes in the magnitude of the input signal will give rise to amplitude changes in the resulting output voltage. This makes prior limiting necessary. It is possible to modify the discriminator circuit to provide limiting, so that the amplitude limiter may be dispensed with. A circuit so modified is called a Ratio Detector Circuit.

As we now, the sum $V_{ao} + V_{bo}$ remains constant, although the difference varies because of changes in input frequency. This assumption is not completely true. Deviation from this ideal does not result in undue distortion in the Ratio Detector Circuit, although some distortion is undoubtedly introduced. It follows that any variations in the magnitude of this sum voltage can be considered spurious here. Their suppression will lead to a discriminator which is unaffected by the amplitude of the incoming signal. It will therefore not react to noise amplitude or spurious amplitude modulation.



(b)

It now remains to ensure that the sum voltage is kept constant. Unfortunately, this cannot be accomplished in the phase discriminator, and the circuit must be modified. This has been done in Figure 6-41, which presents the Ratio Detector Circuit in its basic. form. This is used to show how the circuit is derived from the discriminator and to explain its operation. It is seen that three important changes have been made: one of the diodes has been reversed, a large capacitor (C_5) has been placed across what used to be the output, and the output now is taken from elsewhere.



Operation:

With diode D_2 reversed, o is now positive with respect to b', so that $V_{a'b'}$ is now a sum voltage, rather than the difference it was in the discriminator. It is now possible to connect a large capacitor between a' and b' to keep this sum voltage constant. Once C_5 has been connected, it is obvious that $V_{a'b'}$ is no longer the output voltage; thus the output voltage is now taken between o and o′. It is now necessary to ground one of these two points, and o happens to be the more convenient, as will be seen when dealing with practical Ratio Detector Circuit.

Bearing in mind that in practice $R_5 = R_6$, V_o is calculated as follows:

The output voltage V_o is taken across the terminal t_1, t_2.

$$V_o = V_{t_1,t_2} = |V_{O2}| - \left|\frac{V_R}{2}\right|$$

$$V_R = V_{O1} + V_{O2}$$

$$V_o = V_{O2} - \frac{V_{O1} + V_{O2}}{2}$$

$$V_o = \frac{V_{O2} - V_{O1}}{2}$$

Equation shows that the ratio detector output voltage is equal to half the difference between the output voltages from the individual diodes. Thus (as in the phase discriminator) the output voltage is proportional to the difference between the individual output voltages. The Ratio Detector Circuit therefore behaves identically to the discriminator for input frequency changes. The S curve of Figure 6-40 applies equally to both circuits.

Amplitude limiting by the ratio detector:

It is thus established that the ratio detector behaves in the same way as the phase discriminator when input frequency varies (but input voltage remains constant). The next step is to explain how the Ratio Detector Circuit reacts to amplitude changes. If the input voltage V_{12} is constant and has been so for some time, C_5 has been able to charge up to the potential existing between a’ and b’. Since this is a dc voltage if V_{12} is constant, there will be no current either flowing in to charge the capacitor or flowing out to discharge it. In other words, the input impedance of C_5 is infinite. The total load impedance for the two diodes is therefore the sum of R_3 and R_4, since these are in practice much smaller than R_5 and R_6.

If V_{12} tries to increase, C_5 will tend to oppose any rise in V_o. The way in which it does this is not, however, merely to have a fairly long time constant, although this is certainly part of the operation. As soon as the input voltage tries to rise, extra diode current flows, but the excess current flows into the capacitor C_5, charging it. The voltage $V_{a'b'}$ remains constant at first because it is not possible for the voltage across a capacitor to change instantaneously. The situation now is that the current in the diodes load has risen, but the voltage across the load has not changed. The conclusion is that the load impedance has decreased. The secondary of the ratio detector transformer is more heavily damped, the Q falls, and so does the gain of the amplifier driving the Ratio Detector Circuit. This neatly counteracts the initial rise in input voltage.

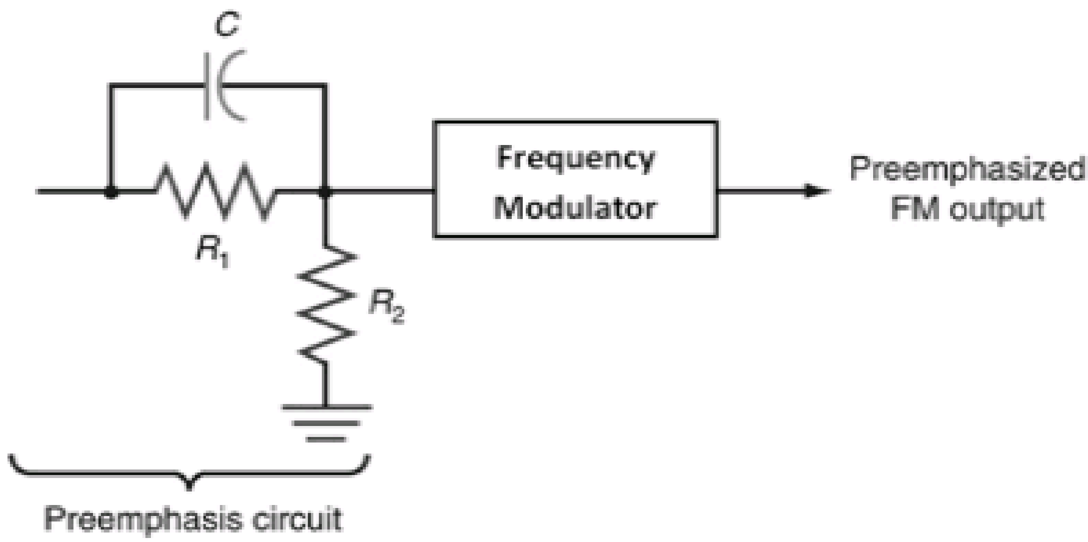
Should the input voltage fall, the diode current will fall, but the load voltage will not, at first, because of the presence of the capacitor. The effect is that of an increased diode load impedance; the diode current has fallen, but the load voltage has remained constant. Accordingly, damping is reduced, and the gain of the driving amplifier rises, this time counteracting an initial fall in the input voltage. The ratio detector provides what is known as diode variable damping. We have here a system of varying the gain of an amplifier by changing the damping of its tuned circuit. This maintains a constant output voltage despite changes in the amplitude of the input.

Performance Comparison of FM Demodulators

S.No.	Parameter of Comparison	Balanced Slope detector	Foster-Seeley (Phase) discriminator	Ratio Detector
(i)	Alignment/tuning	Critical as three circuits are to be tuned at different frequencies	Not Critical	Not Critical
(ii)	Output characteristics depends on	Primary and secondary frequency relationship	Primary and secondary phase relation.	Primary and secondary phase relation.
(iii)	Linearity of output characteristics	Poor	Very good	Good
(iv)	Amplitude limiting	Not providing inherently	Not Provided inherently	Provided by the ratio detector.
(v)	Amplifications	Not used in practice	FM radio, satellite station receiver etc.	TV receiver sound section, narrow band FM receivers.

Pre-emphasis:

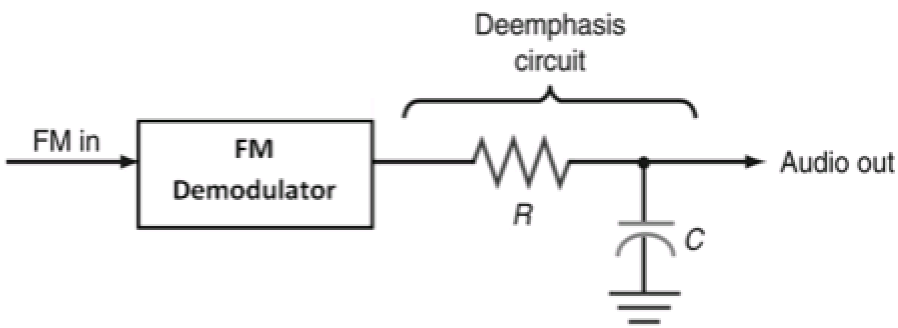
The noise suppression ability of FM decreases with the increase in the frequencies. Thus increasing the relative strength or amplitude of the high frequency components of the message signal before modulation is termed as Pre-emphasis. The Figure below shows the circuit of pre-emphasis.



At the transmitter, the modulating signal is passed through a simple network which amplifies the high frequency components more than the low-frequency components. This pre-emphasis circuit increases the energy content of the higher-frequency signals so that they will tend to become stronger than the high frequency noise components. This improves the signal to noise ratio and increases intelligibility and fidelity.

De-emphasis:

In the de-emphasis circuit, by reducing the amplitude level of the received high frequency signal by the same amount as the increase in pre-emphasis is termed as De-emphasis. The Fig. below shows the circuit of de-emphasis.



- The pre-emphasis process is done at the transmitter side, while the de-emphasis process is done at the receiver side.
- Thus a high frequency modulating signal is emphasized or boosted in amplitude in transmitter before modulation. To compensate for this boost, the high frequencies are attenuated or de-emphasized in the receiver after the demodulation has been performed. Due to pre-emphasis and de-emphasis, the S/N ratio at the output of receiver is maintained constant.
- The de-emphasis process ensures that the high frequencies are returned to their original relative level before amplification.
- Pre-emphasis circuit is a high pass filter or differentiator which allows high frequencies to pass, whereas de-emphasis circuit is a low pass filter or integrator which allows only low frequencies to pass.

Reference Books

1. Dr. Sanjay Sharma, “Analog Communication Systems” 2013, January 2013
2. R.P Singh, S.D. Spare, “Communication Systems: Analog and Digital” ,Tata McGraw-Hill, 3rd Edition, July 2017.
3. Sanjeev Guptha, “Electronic Communication (Analog, Digital and Wireless)”, Khanna publications, January 1997
4. B.P.Lathi, “Modern Digital and Analog Communication Systems”, 3rd Edition, Oxford University Press, 2007.
5. Louis. E. Frenzel, “Communication Electronics Principles and applications”, 3rd Edition, Tata McGraw Hill, 2002.
6. Louis. E. Frenzel, “Principles of Electronic Communication Systems”, 3rd Edition, Tata McGraw Hill, July 2017.
7. Dennis Reddy and John Coolen, "Electronic Communications", 4th Edition, Prentice Hall Publishers, 1995.
8. Electronic Communications Systems by Wayne Tomasi. Pearson education India.
9. Electronic Communication Systems by Roy Blake By Cengage learning.
10. Communication Systems By Simon Haykins By Wiley India.
11. Theory and Problem Of Electronic Communication By Lloyd Temes and Mitchel E.Schulz(Second edition), McGraw Hill Publication.
12. Kennedy, "Electronic Communications Systems", 4th Edition, McGraw-Hill Publishers, 1992.
13. Deshpande, N.D, “Communication Electronics”, Tata McGraw Hill Publishers,1989.
14. R.Gulati, “Monochrome and color television”, 2nd Edition, New age international 2005

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SECA1303 – ANALOG COMMUNICATION SYTEMS

UNIT 3 ANALOG PULSE MODULATION AND MULTIPLEXING

Sampling theorem – Types of sampling-Concepts of PAM, PWM, PPM and PCM - Modulators and demodulators. Types of Multiplexing: Frequency Division Multiplexing, Time Division Multiplexing and Quadrature Multiplexing - Comparison of multiplexing.

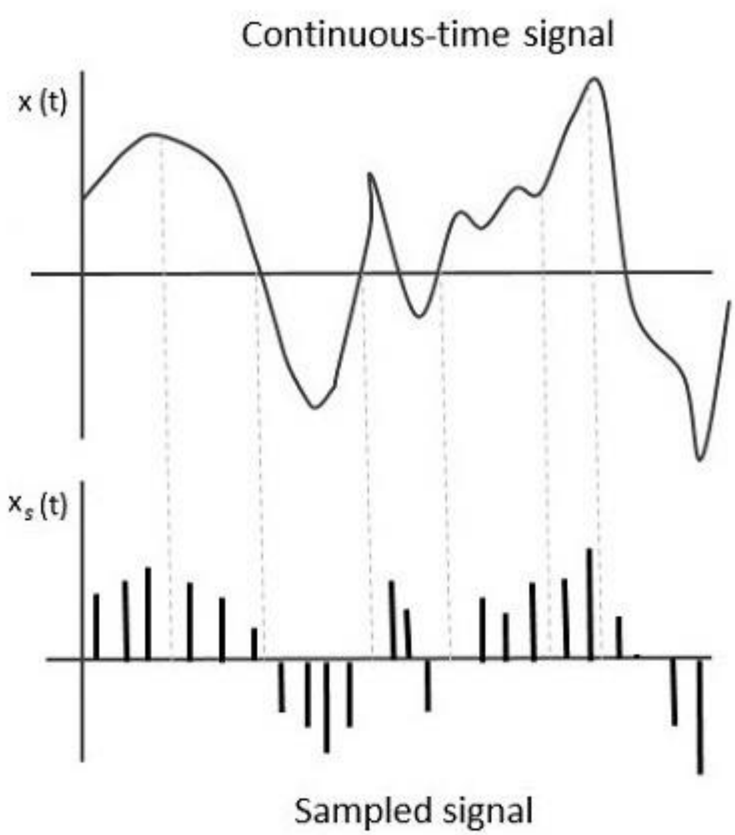
INTRODUCTION

SAMPLING

These pulse modulation techniques deal with discrete signals. So, now let us see how to convert a continuous time signal into a discrete one.

The process of converting continuous time signals into equivalent discrete time signals, can be termed as **Sampling**. A certain instant of data is continually sampled in the sampling process.

The following figure shows a continuous-time signal **x(t)** and the corresponding sampled signal **x_s(t)**. When **x(t)** is multiplied by a periodic impulse train, the sampled signal **x_s(t)** is obtained.



A **sampling signal** is a periodic train of pulses, having unit amplitude, sampled at equal intervals of time T_s, which is called as **sampling time**. This data is transmitted at the time instants T_s and the carrier signal is transmitted at the remaining time.

SAMPLING THEOREM

The sampling rate should be such that the data in the message signal should neither be lost nor it should get over-lapped. The **sampling theorem** states that, "a signal can be exactly reproduced if it is sampled at the rate f_s, which is greater than or equal to twice the maximum frequency of the given signal **W**."

- It states that a continuous time signal can be recovered from its discrete samples if and only if the sampling frequency is greater than or equal to twice the highest frequency of the continuous time signal.

$$\boxed{f_s \geq 2 f_m}$$

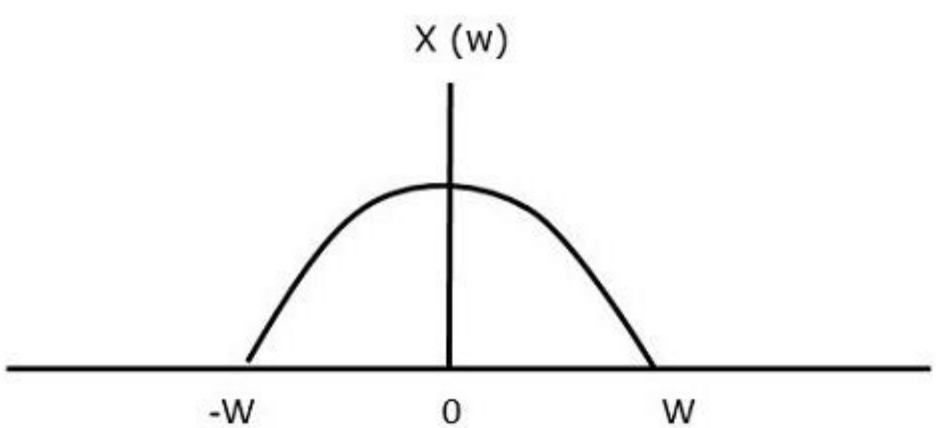
fs - Sampling frequency;

fm-maximum frequency of the message signal

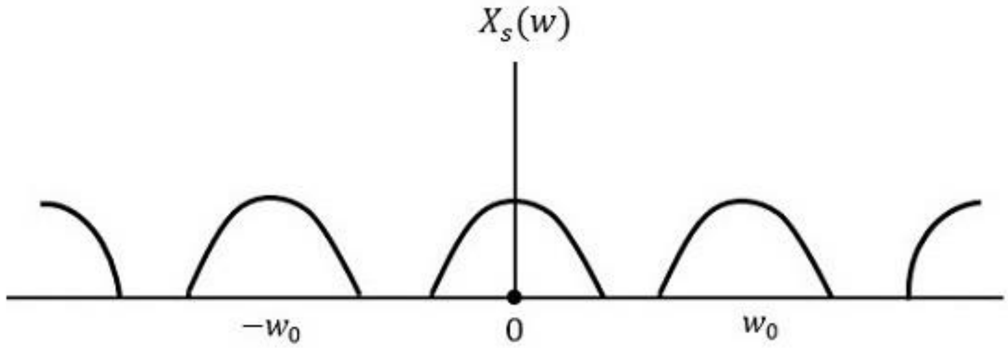
If the sampling rate is equal to twice the maximum frequency of the given signal W, then it is called as **Nyquist rate**.

The sampling theorem, which is also called as **Nyquist theorem**, delivers the theory of sufficient sample rate in terms of bandwidth for the class of functions that are bandlimited.

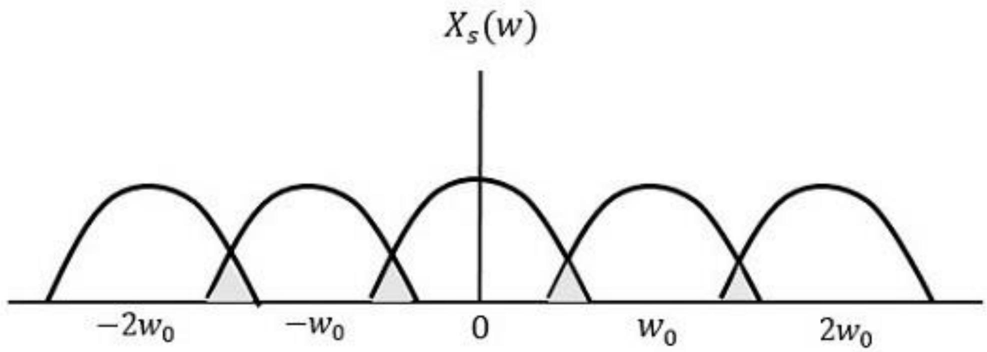
For continuous-time signal **x(t)**, which is band-limited in the frequency domain is represented as shown in the following figure.



If the signal is sampled above Nyquist rate, then the original signal can be recovered. The following figure explains a signal, if sampled at a higher rate than **2w** in the frequency domain.

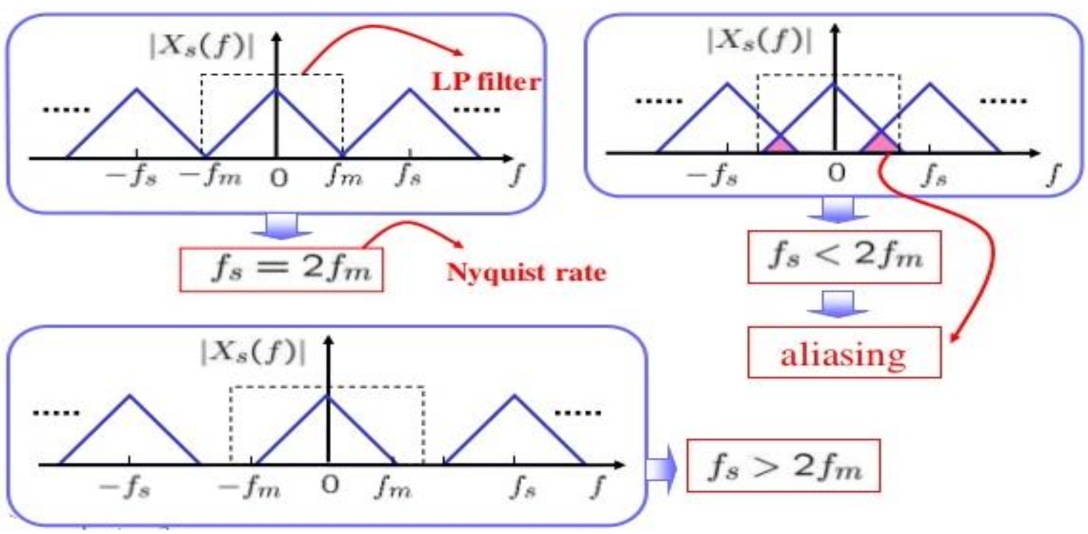


If the same signal is sampled at a rate less than **2w**, then the sampled signal would look like the following figure.



ALIASING EFFECT

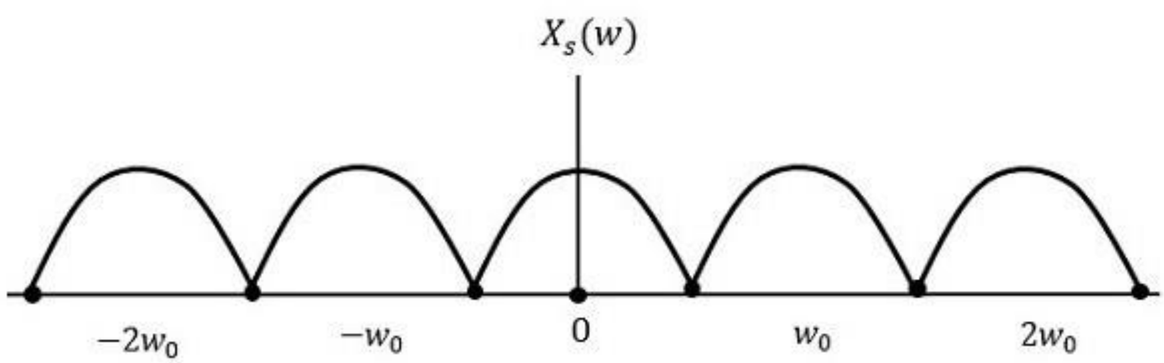
We can observe from the above pattern that there is over-lapping of information, which leads to mixing up and loss of information. This unwanted phenomenon of over-lapping is called as **Aliasing**.



- If fs<2fm, low pass filtered signal contains some high frequency components along with message signal due to spectral overlapping.
- The presence of high frequency signal in the reconstructed signal causes distortion. This is called as **Aliasing effect**.

Aliasing can be referred to as “the phenomenon of a high-frequency component in the spectrum of a signal, taking on the identity of a low-frequency component in the spectrum of its sampled version.”

Hence, the sampling rate of the signal is chosen to be as Nyquist rate. If the sampling rate is equal to twice the highest frequency of the given signal **W**, then the sampled signal would look like the following figure.



In this case, the signal can be recovered without any loss. Hence, this is a good sampling rate.

TYPES OF SAMPLING

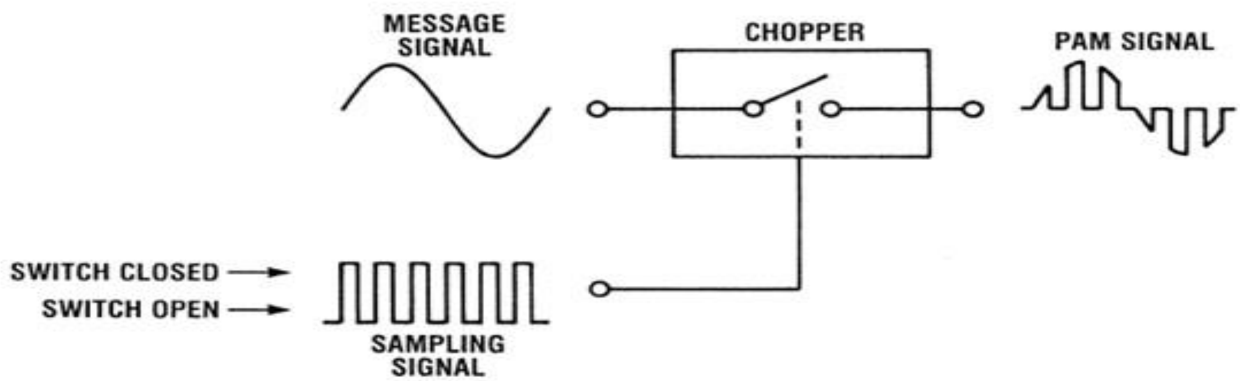
There are basically three types of Sampling techniques, namely:

1. Natural Sampling
2. Flat-top Sampling
3. Ideal Sampling

1. Natural Sampling:

Natural Sampling is a practical method of sampling in which pulse have finite width equal to τ. Sampling is done in accordance with the carrier signal which is digital in nature.

Natural Sampled Waveform

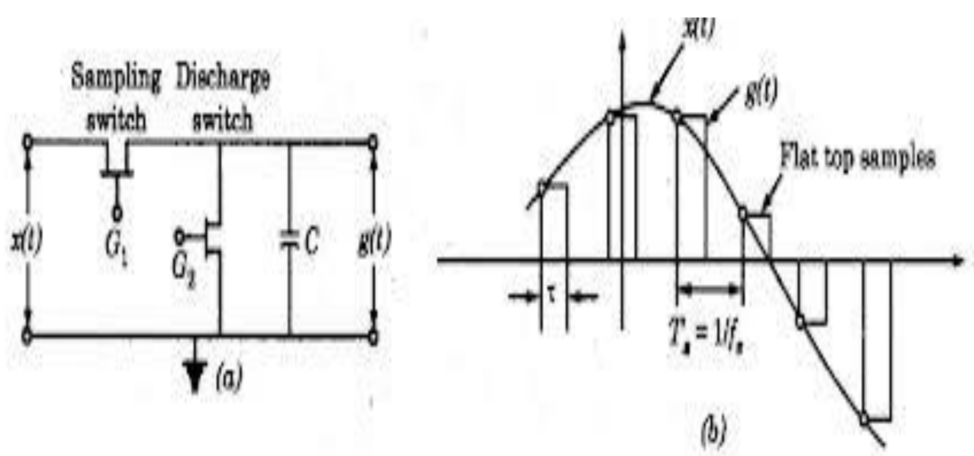


Functional Diagram of Natural Sampler

With the help of functional diagram of a Natural sampler, a sampled signal g(t) is obtained by multiplication of sampling function c(t) and the input signal x(t).

2. Flat Top Sampling:

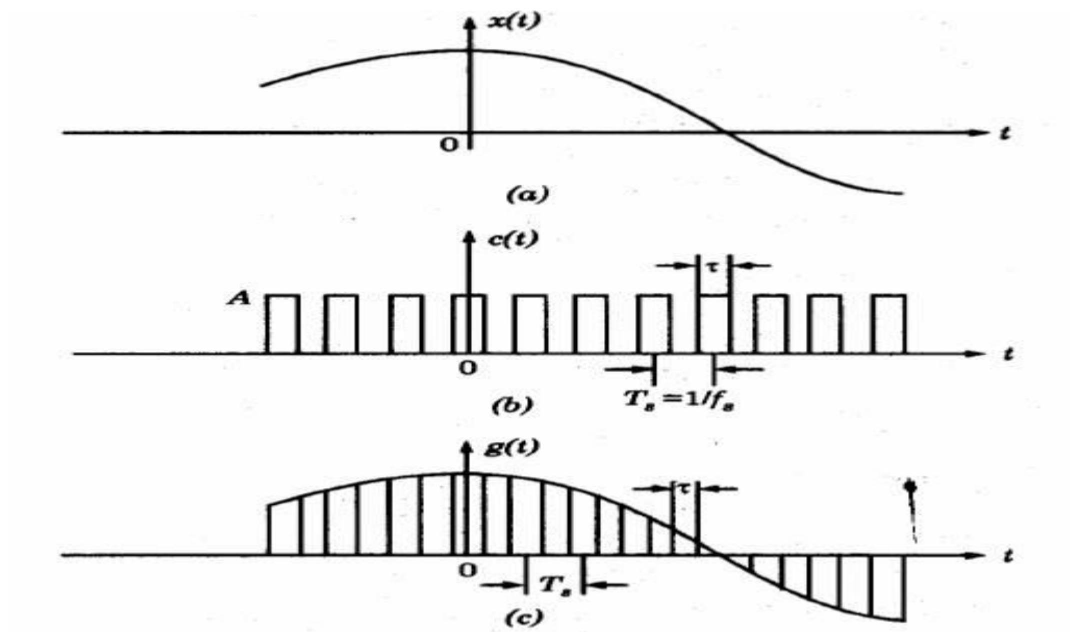
Flat top sampling is like natural sampling i.e; practical in nature. In comparison to natural sampling flat top sampling can be easily obtained. In this sampling techniques, the top of the samples remains constant and is equal to the instantaneous value of the message signal x(t) at the start of sampling process. Sample and hold circuit are used in this type of sampling.



Block Diagram and Waveform

Figure(a), shows functional diagram of a sample hold circuit which is used to generate fat top samples. Figure(b), shows the general waveform of the flat top samples. It can be observed that only starting edge of the pulse represent the instantaneous value of the message signal x(t).

3. Ideal Sampling:



Ideal Sampling Wave form

Ideal Sampling is also known as Instantaneous sampling or Impulse Sampling. Train of impulse is used as a carrier signal for ideal sampling. In this sampling technique the sampling function is a train of impulses and the principle used is known as multiplication principle.

Here,

Figure (a), represent message signal or input signal or signal to be sampled.

PULSE MODULATION

- In Pulse Modulation, amplitude or width or position of a carrier pulse signal is varied in accordance with instantaneous value of the message signal.
- It gives rise to three forms of pulse modulation techniques:
  - Pulse amplitude Modulation (PAM)
  - Pulse width Modulation (PWM)
  - Pulse Position Modulation (PPM)

PULSE AMPLITUDE MODULATION

Pulse Amplitude Modulation (PAM) is an analog modulating scheme in which the amplitude of the pulse carrier varies proportional to the instantaneous amplitude of the message signal.

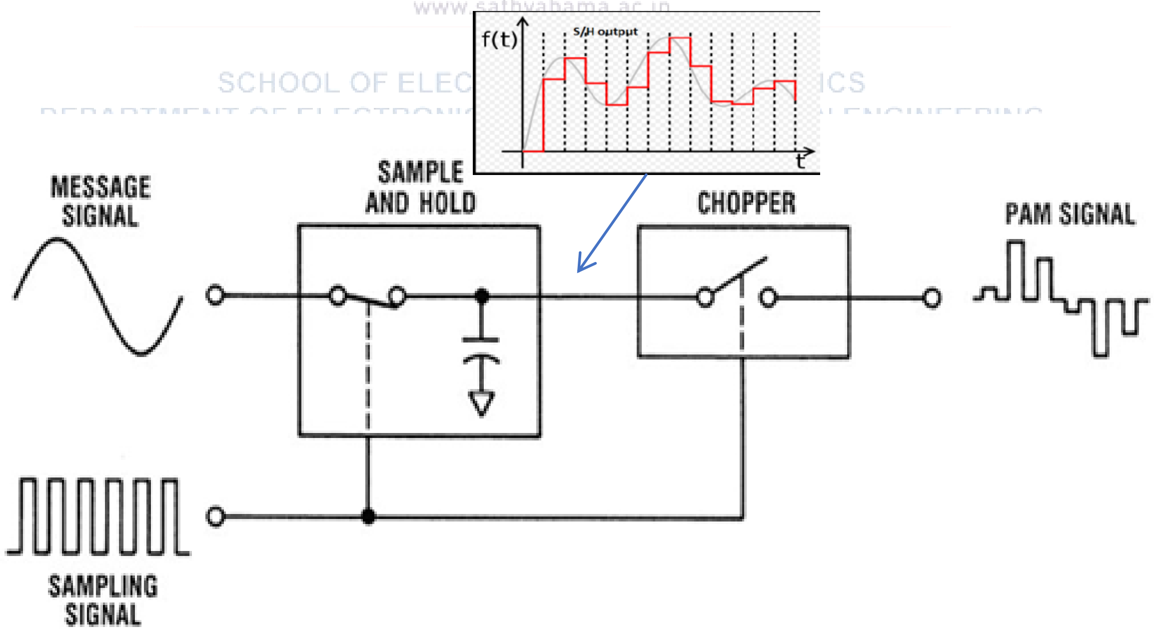
The pulse amplitude modulated signal, will follow the amplitude of the original signal, as the signal traces out the path of the whole wave. In natural PAM, a signal sampled at the Nyquist rate is reconstructed, by passing it through an efficient **Low Pass Frequency (LPF)** with exact cutoff frequency.

PAM MODULATION

PAM generation involves two steps:

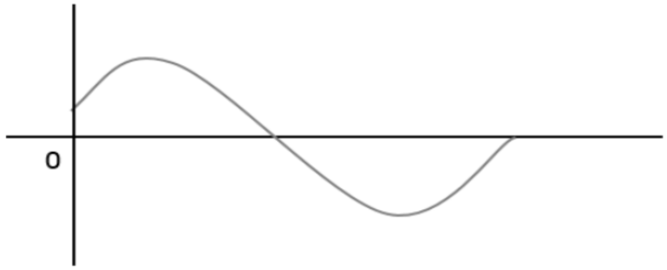
i. The input signal is sampled using a narrow pulse and the sampled value is held constant until the next sample using a capacitor. It is called as **Sample and Hold (S/H)**output.

The S/H output is again sampled using a switch to produce flat top sampled output.

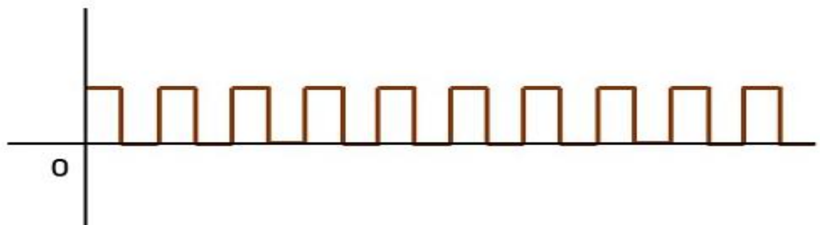


Flat-top sampling – generation of PAM signals.

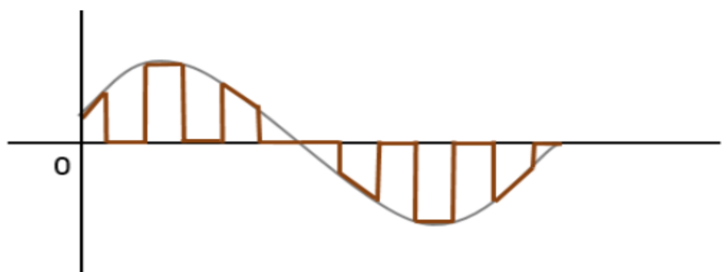
The following figures explain the Pulse Amplitude Modulation.



Modulating signal

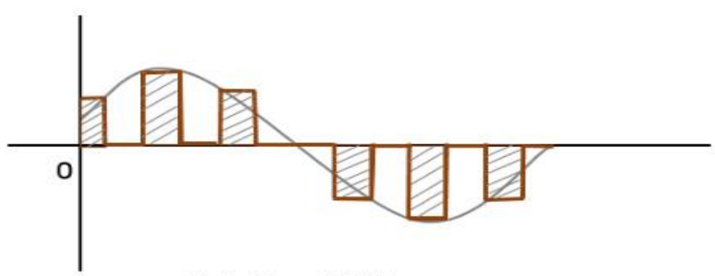


Carrier Pulse train



Natural PAM

Though the PAM signal is passed through an LPF, it cannot recover the signal without distortion. Hence to avoid this noise, flat-top sampling is done as shown in the following figure.

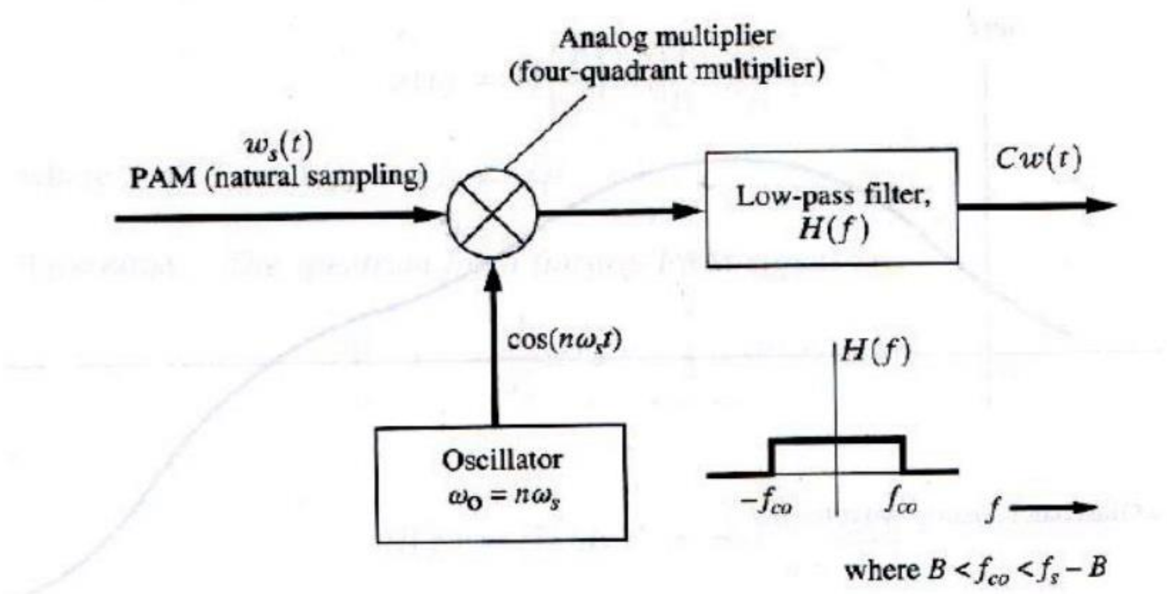


Flat-Top PAM

Flat-top sampling is the process in which sampled signal can be represented in pulses for which the amplitude of the signal cannot be changed with respect to the analog signal, to be sampled. The tops of amplitude remain flat. This process simplifies the circuit design.

PAM DEMODULATION

For the demodulation of the PAM signal, the PAM signal is fed to the low pass filter. The low pass filter eliminates the high-frequency ripples and generates the demodulated signal. This signal is then applied to the inverting amplifier to amplify its signal level to have the demodulated output with almost equal amplitude with the modulating signal.



PAM Demodulator circuit

PULSE WIDTH MODULATION

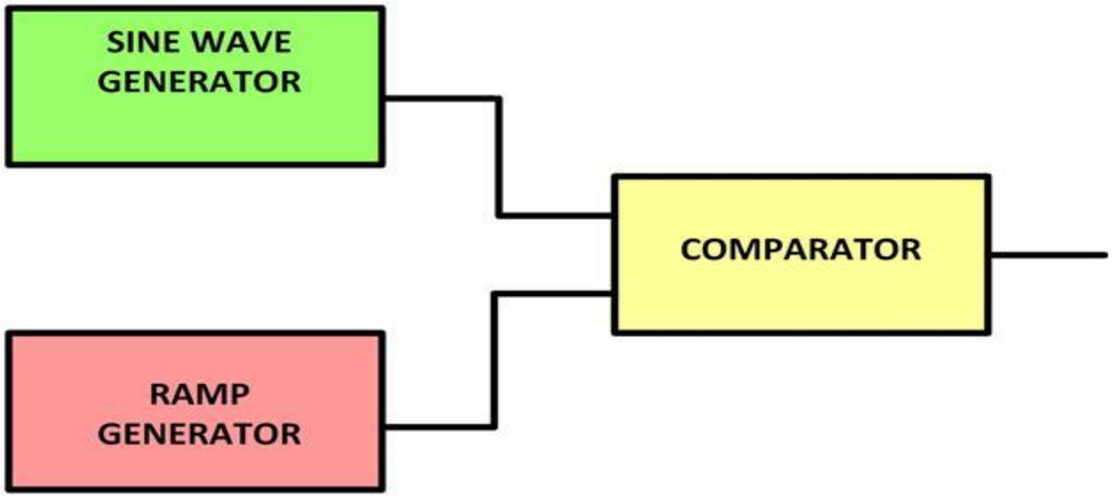
Pulse Width Modulation (PWM) or **Pulse Duration Modulation (PDM)** or **Pulse Time Modulation (PTM)** is an analog modulating scheme in which the duration or width or time of the pulse carrier varies proportional to the instantaneous amplitude of the message signal.

PWM Modulator:

- In PWM, amplitude and frequency of the pulse remains constant.
- PWM signal can be obtained by comparing the message signal with a saw tooth signal.
- A comparator circuit will set its output to High level(+Vcc) if input1 is greater than input2 and set its output to low level(zero) if input2 is grater than input1 .

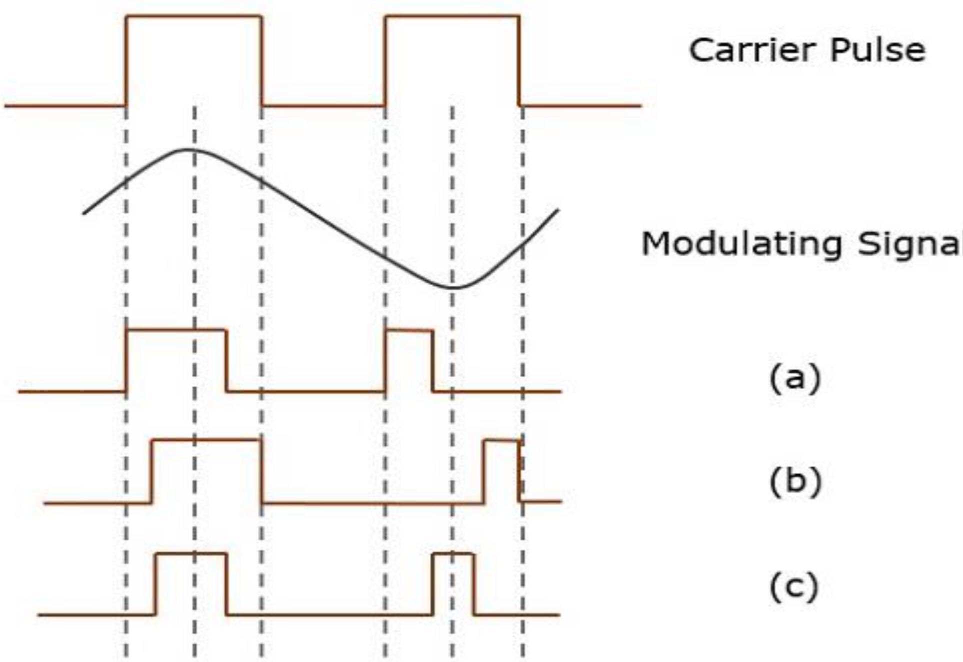
- A monostable multivibrator sets its output to High level for each trailing edge of the input and then the output automatically resets to zero after a predetermined time.

The width of the pulse varies in this method, but the amplitude of the signal remains constant. Amplitude limiters are used to make the amplitude of the signal constant. These circuits clip off the amplitude, to a desired level and hence the noise is limited.



PWM Modulator Block Diagram

The following figures explain the types of Pulse Width Modulations.

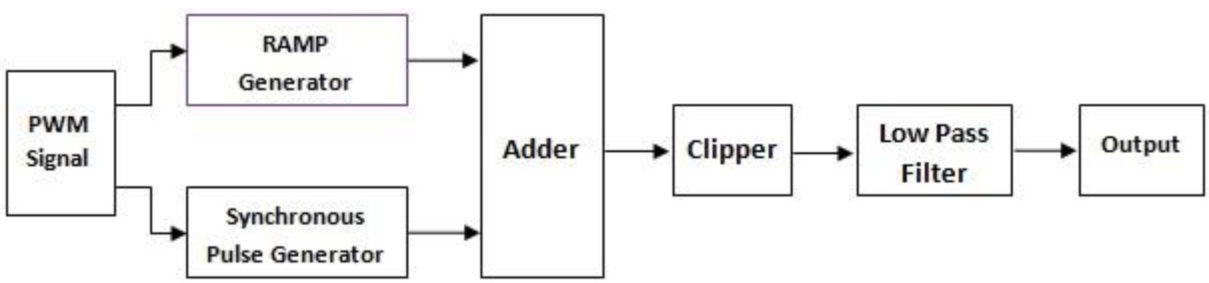


There are three variations of PWM. They are −

- The leading edge of the pulse being constant, the trailing edge varies according to the message signal.
- The trailing edge of the pulse being constant, the leading edge varies according to the message signal.
- The center of the pulse being constant, the leading edge and the trailing edge varies according to the message signal.

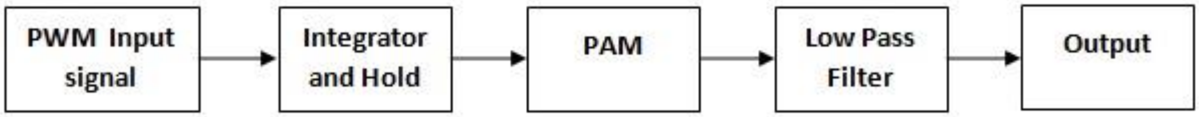
PWM DEMODULATION

There are two common techniques used for pulse-width demodulation. One method is that the PWM signal must first be converted to a pulse-amplitude modulation (PAM) signal and then passed through a low-pass filter. The PWM signal is applied to an integrator and hold circuit. When the positive edge of pulse appears, the integrator generates ramp output whose magnitude is proportional to the pulse width. After the negative edge, the hold circuit maintains the peak ramp voltage for a given period and then forces the output voltage to zero. The waveform is the sum of a sequence of constant-amplitude and constant-width pulse generated by demodulator. This signal is then applied to the input of clipping circuit, which cuts off the portion of signal below the threshold voltage and outputs the reminder. Therefore, the output of clipping circuit is a PAM signal whose amplitude is proportional to the width of PWM signal. Finally, the PAM signal passes through a simple low-pass filter and the original audio signal is obtained.



Block diagram of PWM Demodulation.

The other technique for demodulating a PWM signal is consist of a product detector and a low-pass filter. The PWM and the carrier signals are connected to the inputs of a product detector, and then a sequence of pulses having the width inversely proportional to the width of PWM pulse presents at output. When the Va signal passes through the low-pass filter, a demodulated signal is obtained.



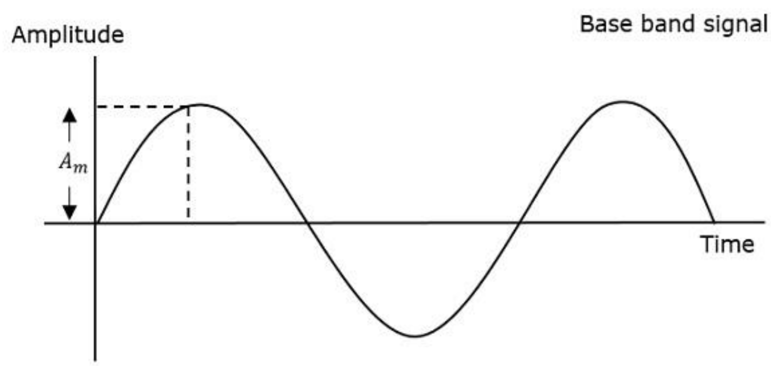
Block diagram of another scheme for PWM Demodulation.

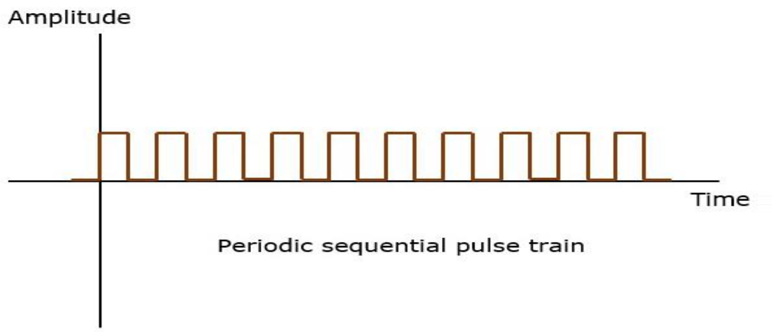
For any pulse wave modulation, before modulating, the original continuous type signal must be sampled and the sampling rate of the sampling signal cannot be low, or else the recovered signal will cause distortion. The sampling rate depends on the sampling theorem which the sampling theorem is defined as: for any pulse wave modulation system, if the sampling rate excesses double or more times of the maximum frequency of the signal, then the distortion level of the data recovery at the receiver will be the minimum. For example, the frequency range of the audio signal is 40 Hz ~ 4 kHz, then the sampling signal frequency of the pulse wave modulation must be at least 8 kHz, therefore, the sampling error can be reduced to minimum.

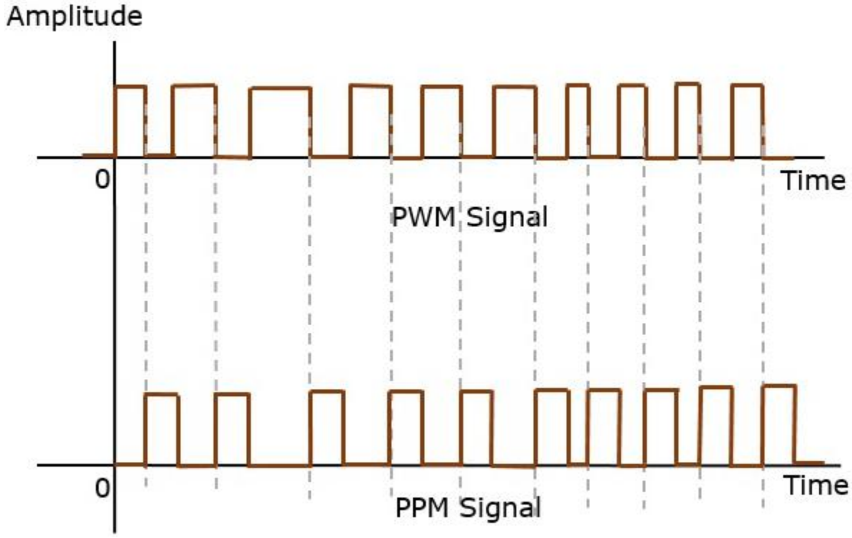
PULSE POSITION MODULATION

Pulse Position Modulation (PPM) is an analog modulating scheme in which the amplitude and width of the pulses are kept constant, while the position of each pulse, with reference to the position of a reference pulse varies according to the instantaneous sampled value of the message signal.

The transmitter has to send synchronizing pulses (or simply sync pulses) to keep the transmitter and receiver in synchronism. These sync pulses help maintain the position of the pulses. The following figures explain the Pulse Position Modulation.



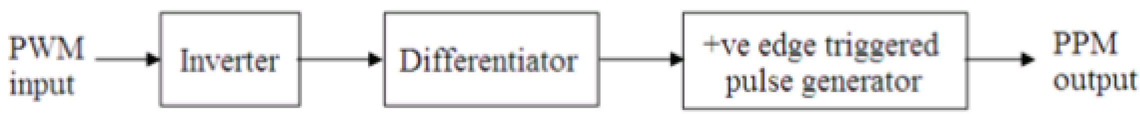




Pulse position modulation is done in accordance with the pulse width modulated signal. Each trailing of the pulse width modulated signal becomes the starting point for pulses in PPM signal. Hence, the position of these pulses is proportional to the width of the PWM pulses.

GENERATION (MODULATION) OF PPM

- PPM signal can be generated with the help of PWM as shown in Fig7 below.



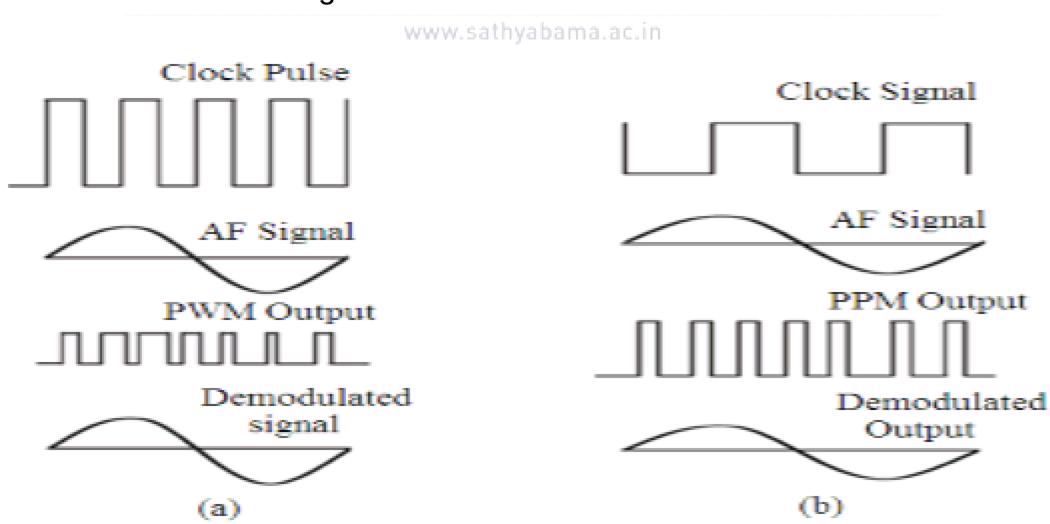
PPM generation from PWM

- The PWM signal generated above is sent to an inverter which reverses the polarity of the pulses.
- This is then followed by a differentiator which generates +ve spikes for PWM signal going from High to Low and -ve spikes for Low to High transistion. The spikes generated are shown in the fourth waveform of Fig8.

- These spikes are then fed to the positive edge triggered pulse generator which generates fixed width pulses when a +ve spike appears, coinciding with the falling edge of the PWM signal.
- Thus PPM signal is generated at the output which is shown in the fifth waveform of Fig8.where pulse position carry the message information.

DEMODULATION OF PPM

- For PWM demodulation, put a ramp at the +ve edge which will stop at the arrival of –ve egde.
- The ramp will attain different heights in each cycle since the widths are different and the heights attained are directly proportional to the pulse width and in turn the amplitude of the message signal.
- This is then passed through a low pass filter where it will follow the envelop i.e. the message signal, which produces the demodulated signal at the output.
- For PPM demodulation, ramp is used which starts at the +ve edge of the one pulse and stops at the +ve edge of the next pulse.
- Thus the height of the generated ramp is determined by the delay between the pulses which indirectly follows the amplitude of the modulating signal.
- This is then passed through a low pass filter which filters the envelop information as the demodulated signal.



Modulation and Demodulation of (a) PWM and (b) PPM

COMPARISON BETWEEN PAM, PWM, AND PPM

The comparison between the above modulation processes is presented in a single table.

PAM	PWM	PPM
Amplitude is varied	Width is varied	Position is varied
Bandwidth depends on the width of the pulse	Bandwidth depends on the rise time of the pulse	Bandwidth depends on the rise time of the pulse
Instantaneous transmitter power varies with the amplitude of the pulses	Instantaneous transmitter power varies with the amplitude and width of the pulses	Instantaneous transmitter power remains constant with the width of the pulses
System complexity is high	System complexity is low	System complexity is low
Noise interference is high	Noise interference is low	Noise interference is low
It is similar to amplitude modulation	It is similar to frequency modulation	It is similar to phase modulation

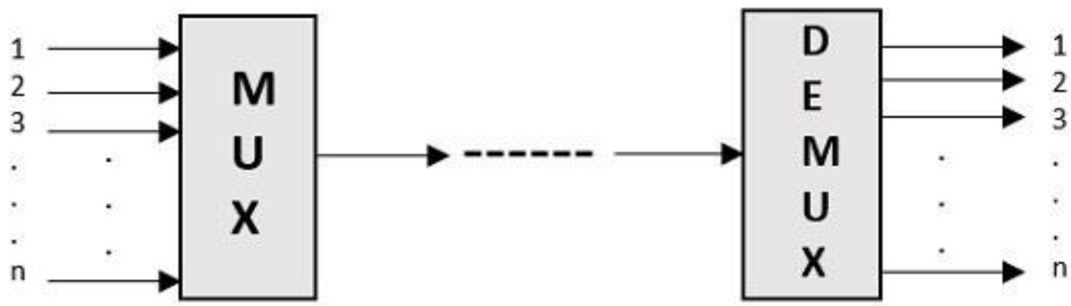
MULTIPLEXING

Multiplexing is the process of combining multiple signals into one signal, over a shared medium. If the analog signals are multiplexed, then it is called as **analog multiplexing**. Similarly, if the digital signals are multiplexed, then it is called as **digital multiplexing**.

Multiplexing was first developed in telephony. A number of signals were combined to send through a single cable. The process of multiplexing divides a communication channel into several number of logical channels, allotting each one for a different message signal or a data stream to be transferred. The device that does multiplexing can be called as **Multiplexer** or **MUX**.

The reverse process, i.e., extracting the number of channels from one, which is done at the receiver is called as **de-multiplexing**. The device that does de-multiplexing can be called as **de-multiplexer** or **DEMUX**.

The following figure illustrates the concept of MUX and DEMUX. Their primary use is in the field of communications.



Multiplexing and Demultiplexing

TYPES OF MULTIPLEXERS

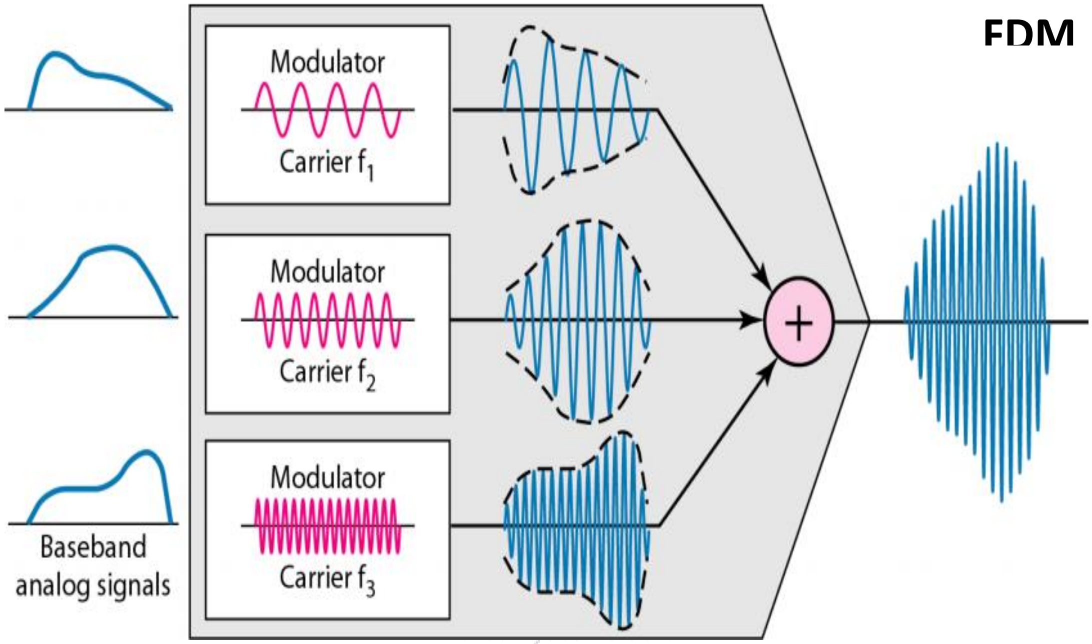
There are mainly two types of multiplexers, namely analog and digital. They are further divided into

- Frequency Division Multiplexing (FDM),
- Time Division Multiplexing (TDM).
- Quadrature carrier multiplexing (QCM)

FREQUENCY DIVISION MULTIPLEXING (FDM)

- FDM is an analog multiplexing technique that combines different signals by modulating each analog signal with a different carrier frequency.
- In FDM, the total bandwidth is divided to a set of frequency bands that do not overlap. Each of these bands is a carrier of a different signal that is generated and modulated by one of the sending devices. The frequency bands are separated from one another by strips of unused frequencies called the guard bands, to prevent overlapping of signals.
- The modulated signals are combined together using a multiplexer (MUX) in the sending end. The combined signal is transmitted over the communication channel, thus allowing multiple independent data streams to be transmitted simultaneously. Example
- The following diagram conceptually represents multiplexing using FDM. It has 4 frequency bands, each of which can carry signal from 1 sender to 1 receiver.

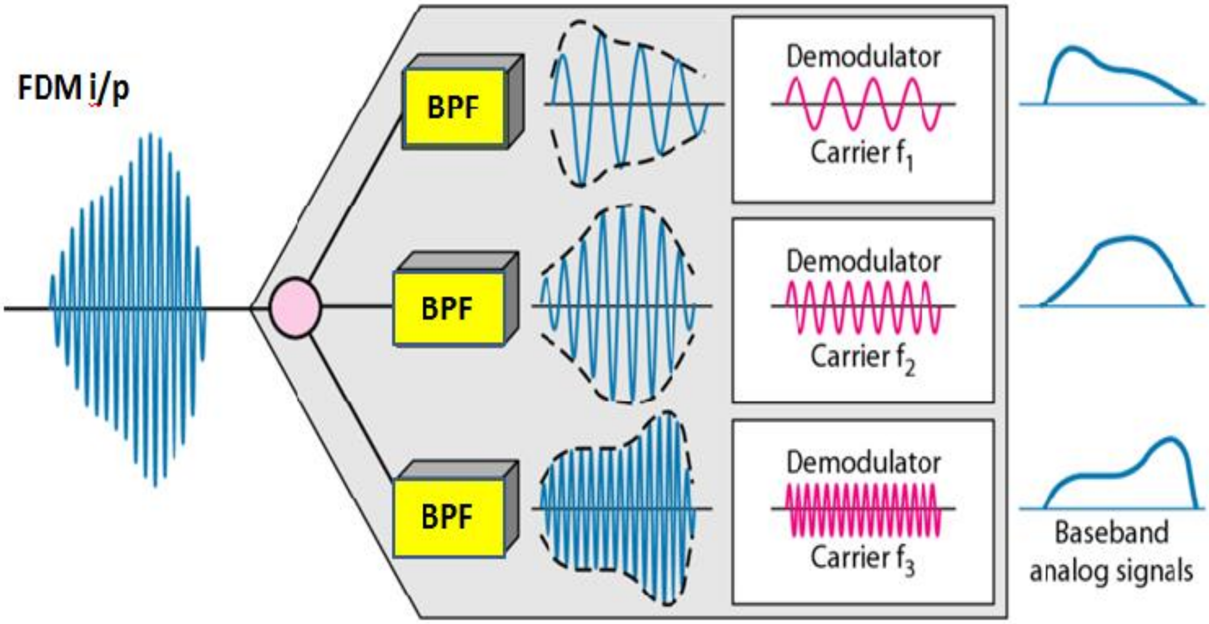
Each of the 4 senders is allocated a frequency band. The four frequency bands are multiplexed and sent via the communication channel. At the receiving end, a demultiplexer regenerates the original four signals as outputs.



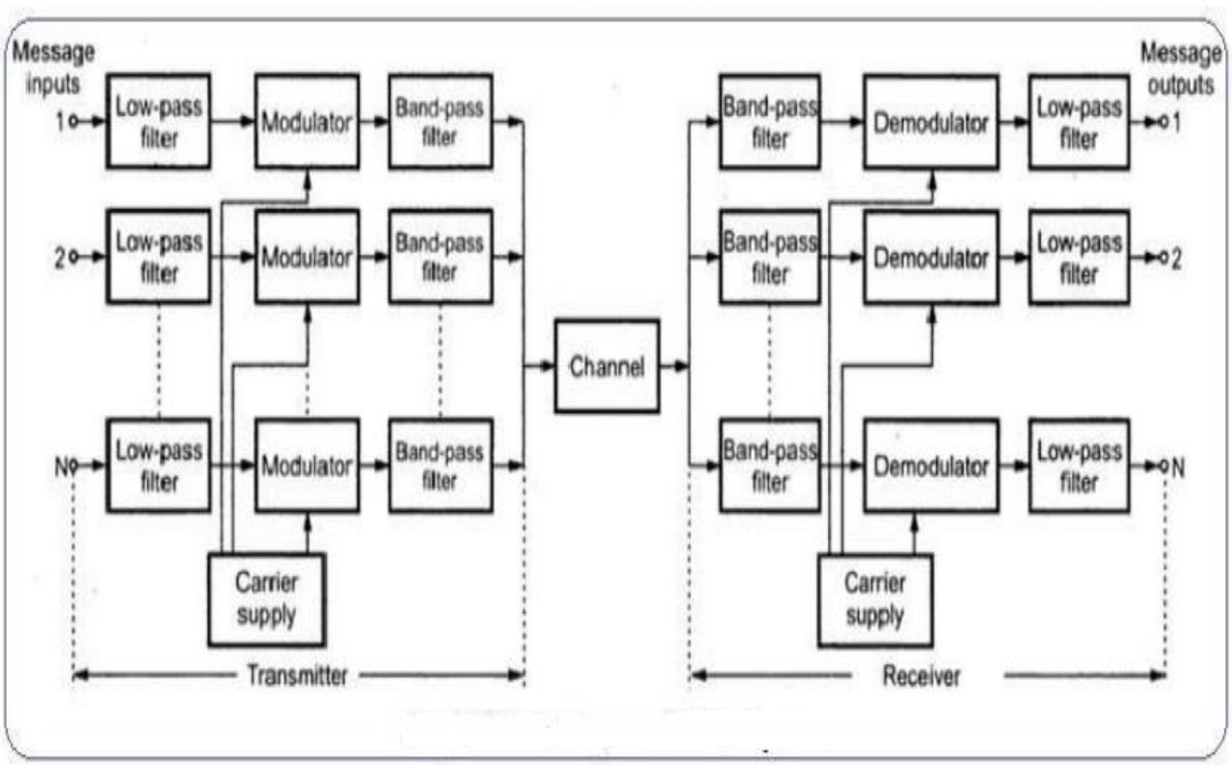
Demultiplexing FDM signal

Demultiplexing is the processing of recovering the individual baseband signals from the multiplexed signal.

- At the receiving end, the individual signals are extracted from the combined signal by the process of demultiplexing (DEMUX).

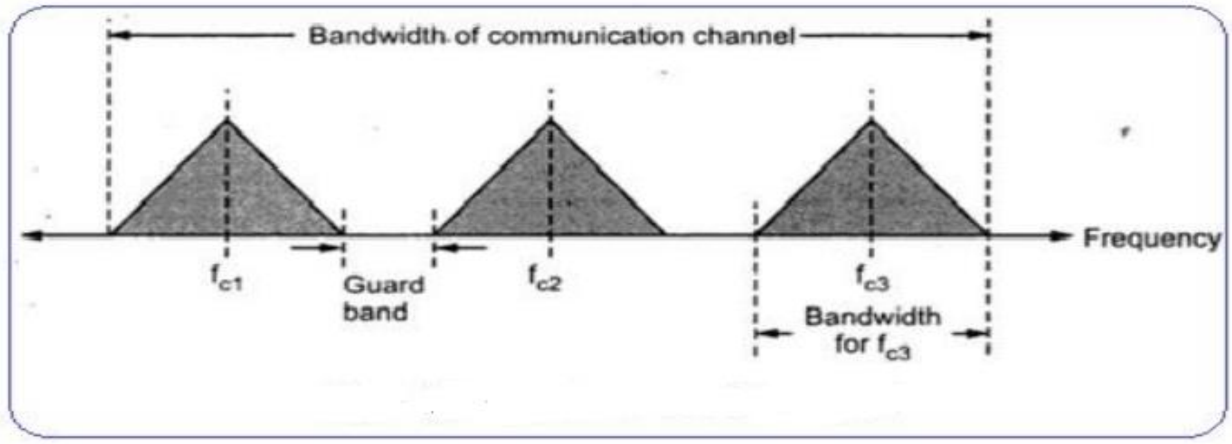


Block Diagram of FDM receiver System



Spectrum of FDM

- Spectrum of FDM signal shows that each subcarrier modulated signal is separated by a **small frequency band** to prevent inter-channel interference or cross talk. These unused frequency band between each successive channel are known as **guard bands**.If the channels are very close to one other, it leads to inter-channel cross talk.



TIME DIVISION MULTIPLEXING

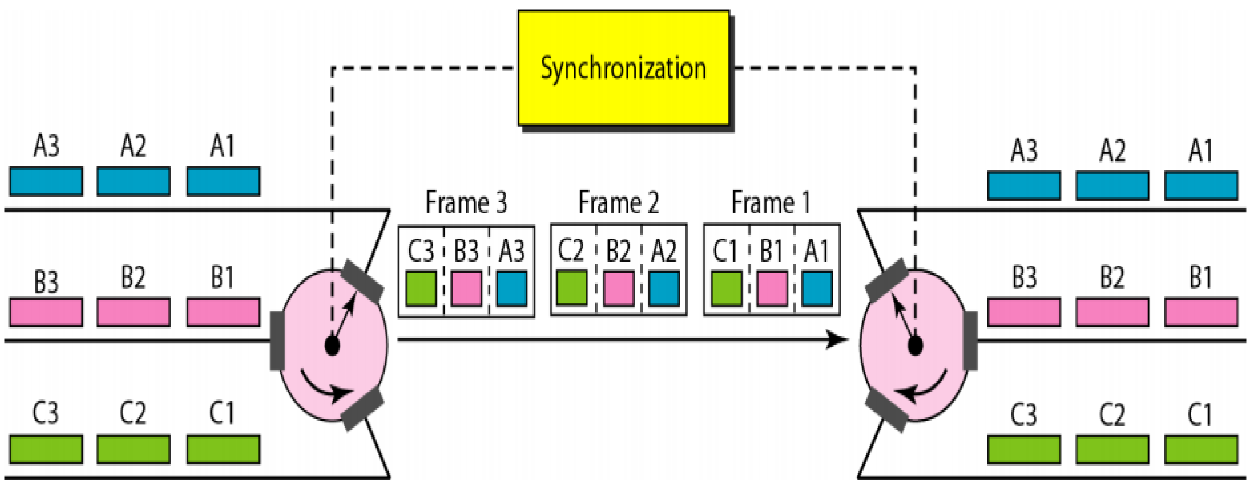
TDM is a multiplexing technique in which each signal is assigned a different time slot for transmission. TDM requires synchronization between the switching unit at Transmitter and Receiver.

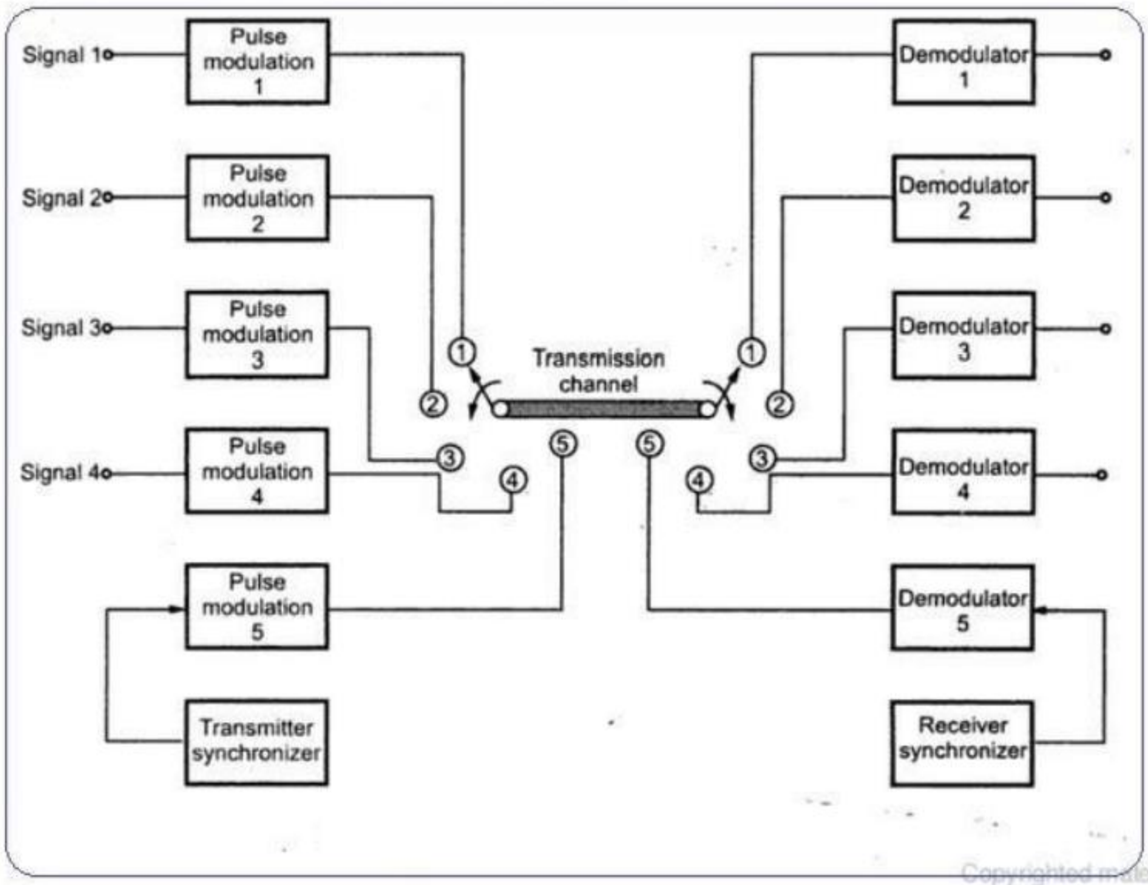
Time division multiplexing (FDM) is a technique of multiplexing, where the users are allowed the total available bandwidth on time sharing basis. Here the time domain is divided into several recurrent slots of fixed length, and each signal is allotted a time slot on a round-robin basis.

In TDM, the data flow of each input stream is divided into units. One unit may be 1 bit, 1 byte, or a block of few bytes. Each input unit is allotted an input time slot. One input unit corresponds to one output unit and is allotted an output time slot. During transmission, one unit of each of the input streams is allotted one-time slot, periodically, in a sequence, on a rotational basis. This system is popularly called round-robin system.

Example

Consider a system having four input streams, A, B, C and D. Each of the data streams is divided into units which are allocated time slots in the round – robin manner. Hence, the time slot 1 is allotted to A, slot 2 is allotted to B, slot 3 is allotted to C, slot 4 is allotted to D, slot 5 is allocated to A again, and this goes on till the data in all the streams are transmitted





Block diagram of TDM system

Merits and Demerits of TDM

Advantages:

- Time division multiplexing circuitry is not complex.
- Problem of cross talk is not severe.
- Full available channel bandwidth can be utilized for each channel.

Disadvantage:

- Synchronization is required in time division multiplexing.

Applications:

- It used in ISDN (Integrated Services Digital Network) telephone lines.
- It is used in PSTN (public switched telephone network).

QUADRATURE CARRIER MULTIPLEXING

- A **Quadrature Carrier Multiplexing (QCM) or Quadrature Amplitude Modulation (QAM)** method enables two DSBSC modulated waves to occupy the same transmission band width and yet the two message signals can be separated at the receiver.

- Quadrature amplitude modulation (QAM) is modulation techniques that we can utilize in analog modulation concept and digital modulation concept. Depending upon the input signal form we can use it in either analog or digital modulation schemes. In QAM, we can modulate two individual signals and transmitted to the receiver level. And by using the two input signals, the channel bandwidth also increases. QAM can able to transmit two message signals over the same channel. This QAM technique also is known as "quadrature carrier multiplexing"
- The transmitter involves the use of two separate product modulators that are supplied with two carrier waves of the same frequency but differing in phase by 90 degrees.

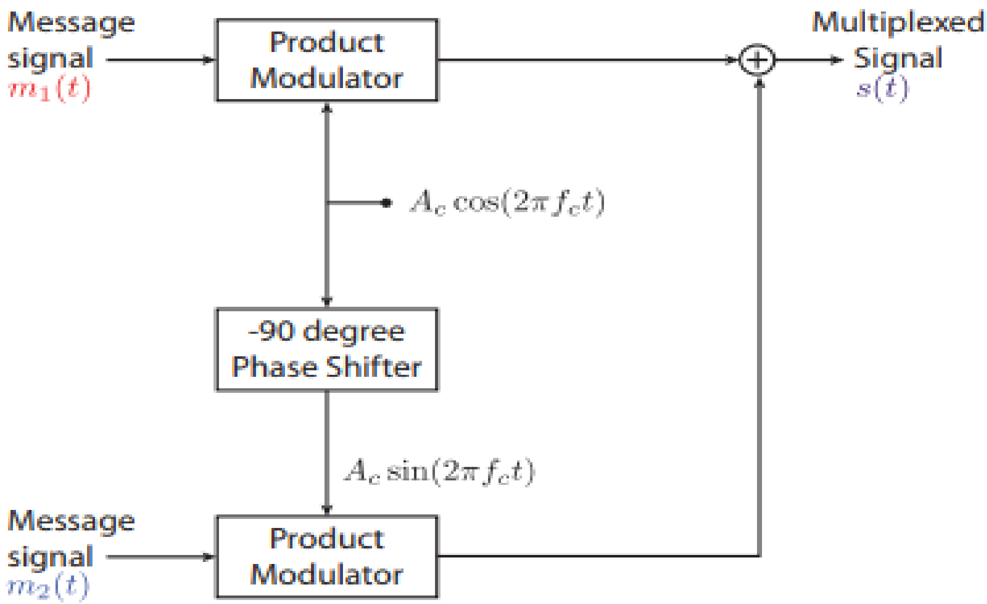
- The multiplexed signal s(t) consists of the sum of the two product modulator outputs given by :

$$s(t) = A_c m_1(t)\cos(2\pi f_c t) + A_c m_2(t)\sin(2\pi f_c t)$$

QCM/QAM Transmitter and Receiver

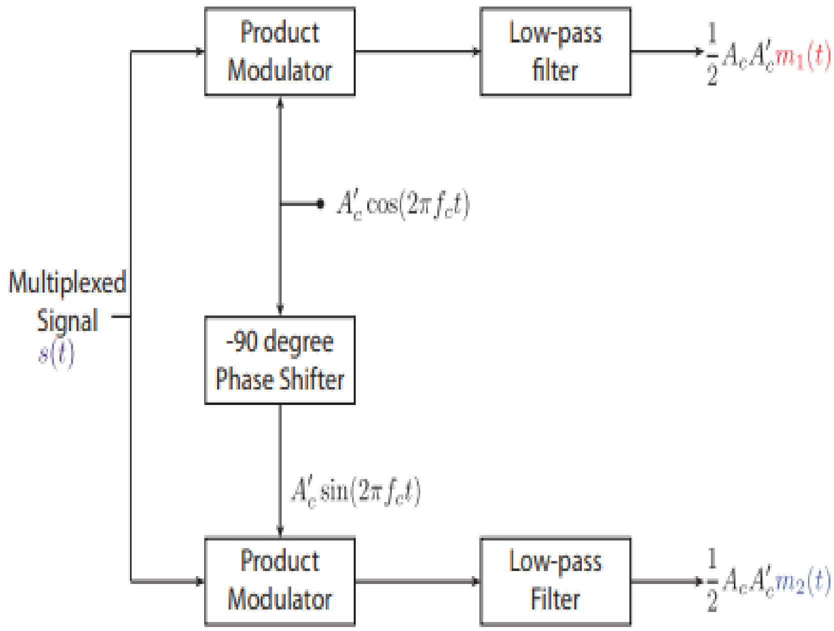
QCM/QAM Transmitter

"In the QAM transmitter, the above section i.e., product modulator1 and local oscillator are called the in-phase channel and product modulator2 and local oscillator are called a quadrature channel. Both output signals of the in-phase channel and quadrature channel are summed so the resultant output will be QAM."



QCM/QAM Receiver

At the receiver level, the QAM signal is forwarded from the upper channel of receiver and lower channel, and the resultant signals of product modulators are forwarded from LPF1 and LPF2. These LPF's are fixed to the cut off frequencies of input 1 and input 2 signals. Then the filtered outputs are the recovered original signals.



QDM/QCM/QAM Receiver

APPLICATIONS OF QCM/QAM

- Used in both analog and digital communication systems.
- Used in Satellite Communications, Cellular mobile communication, Color television for mixing color information, etc.,
- The applications of QAM are mostly observed in radio communications and data delivery applications systems.
- QAM technique has wide applications in the radio communications field because, as the increment of the data rate there is the chance of noise increment but this QAM technique is not affected by noise interference hence there is an easy mode of signal transmission can be possible with this QAM.
- QAM has wide applications in transmitting digital signals like digital cable television and in internet services.
- In cellular technology, wireless device technology quadrature amplitude modulation is preferred.

QAM ADVANTAGES AND DISADVANTAGES

Although QAM appears to increase the efficiency of transmission for radio communications systems by utilising both amplitude and phase variations, it has a number of drawbacks.

The first is that it is more susceptible to noise because the states are closer together so that a lower level of noise is needed to move the signal to a different decision point.

Receivers for use with phase or frequency modulation are both able to use limiting amplifiers that are able to remove any amplitude noise and thereby improve the noise reliance. This is not the case with QAM.

The second limitation is also associated with the amplitude component of the signal.

When a phase or frequency modulated signal is amplified in a radio transmitter, there is no need to use linear amplifiers, whereas when using QAM that contains an amplitude component, linearity must be maintained.

Unfortunately linear amplifiers are less efficient and consume more power, and this makes them less attractive for mobile applications.

COMPARISON OF MULTIPLEXING TECHNIQUES

Sr no.	FDM	TDM
1.	The signals which are to be multiplexed are added in the time domain . But they occupy different slots in the frequency domain .	The signals which are to be multiplexed can occupy the entire bandwidth in the time domain .
2.	FDM is usually preferred for the analog signals .	TDM is preferred for the digital signals .
3.	Synchronization is not required .	Synchronization is required .
4.	The FDM requires a complex circuitry at Tx and Rx .	TDM circuitry is not very complex .
5.	FDM suffers from the problem of crosstalk due to imperfect BPF .	In TDM the problem of crosstalk is not severe .
6.	Due to bandwidth fading in the Tx medium , all the FDM channels are affected .	Due to fading only a few TDM channels will be affected .
7.	Due to slow narrowband fading taking place in the transmission channel may be affected in FDM .	Due to slow narrowband fading all the TDM channels may get wiped out .

Reference Books

1. Dr. Sanjay Sharma, “Analog Communication Systems” 2013, January 2013
2. R.P Singh, S.D. Spare, “Communication Systems: Analog and Digital” ,Tata McGraw-Hill, 3rd Edition, July 2017.
3. Sanjeev Guptha, “Electronic Communication (Analog, Digital and Wireless)”, Khanna publications, January 1997
4. B.P.Lathi, “Modern Digital and Analog Communication Systems”, 3rd Edition, Oxford University Press, 2007.
5. Louis. E. Frenzel, “Communication Electronics Principles and applications”, 3rd Edition, Tata McGraw Hill, 2002.
6. Louis. E. Frenzel, “Principles of Electronic Communication Systems”, 3rd Edition, Tata McGraw Hill, July 2017.
7. Dennis Reddy and John Coolen, "Electronic Communications", 4th Edition, Prentice Hall Publishers, 1995.
8. Electronic Communications Systems by Wayne Tomasi. Pearson education India.
9. Electronic Communication Systems by Roy Blake By Cengage learning.
10. Communication Systems By Simon Haykins By Wiley India.
11. Theory and Problem Of Electronic Communication By Lloyd Temes and Mitchel E.Schulz(Second edition), McGraw Hill Publication.
12. Kennedy, "Electronic Communications Systems", 4th Edition, McGraw-Hill Publishers, 1992.
13. Deshpande, N.D, “Communication Electronics”, Tata McGraw Hill Publishers,1989.
14. R.Gulati, “Monochrome and color television”, 2nd Edition, New age international 2005

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SECA1303 – ANALOG COMMUNICATION SYSTEMS

UNIT 4 ANALOG RECEIVERS AND NOISE

AM Receivers: TRF receivers -Super heterodyne receivers - FM Receivers: FM stereo broadcast receivers - AFC - Capture effect, FM threshold effect. Communication Receivers: Sensitivity, fidelity and selectivity - Squelch circuit - Beat frequency Oscillator-Types of Noise- Noise factor and noise temperature for cascaded amplifier (Friis formula)- Noise in AM and FM systems.

I. Receiver

A typical radio communication system from a broadcasting station consists of a transmitter. The broadcasting station is allocated with a unique RF carrier wave along with a well defined channel width. The transmitter transmits the modulated carrier into space through an antenna. These wave propagate through space. Elsewhere in a remote location there exists a receiver which receives the modulated carrier through the receiving antenna with the help of a tuning circuit. The receiver demodulates the modulated carrier and converts it into speech or intelligence.

A radio receiver completes a communications system. Without a receiver, a transmitter is useless! Receivers come in many different forms. They can be designed to receive voice, digital data, and many other kinds of signals. Receivers of all types share many common features.

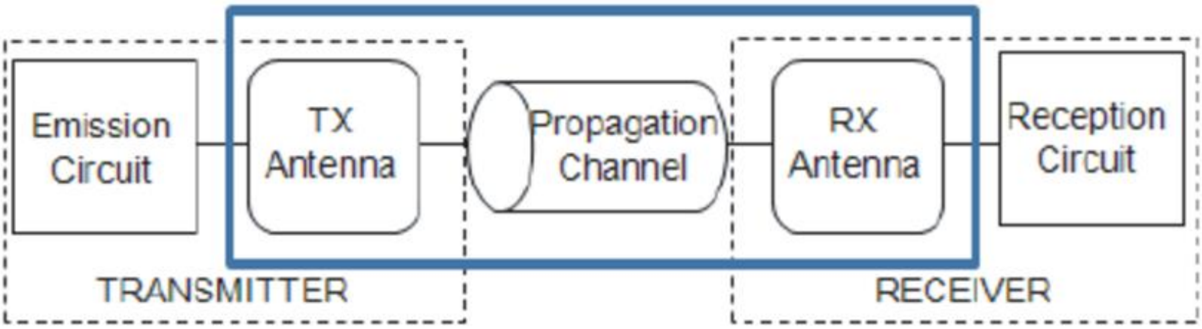


Fig. Block diagram of a typical radio system consisting a transmitter and receiver

Transmission : The radio transmitter consists of a transducer which converts speech or intelligence into audio frequency electrical signals. These amplified AF signals modulate the radio frequency carrier. The modulator performs the task of modulation. The modulated RF carrier is then amplified and transmitted through an antenna.

Reception : The radio receiver consists of an antenna connected to a tuning circuit. The received modulated RF carrier is amplified and then passed through the demodulator to extract the AF signals. The AF signal is then amplified and fed to transducer which converts it into speech or intelligence.

Requirements of a Receiver

AM receiver receives AM wave and demodulates it by using the envelope detector. Similarly, FM receiver receives FM wave and demodulates it by using the Frequency Discrimination method. Following are the requirements of both AM and FM receiver.

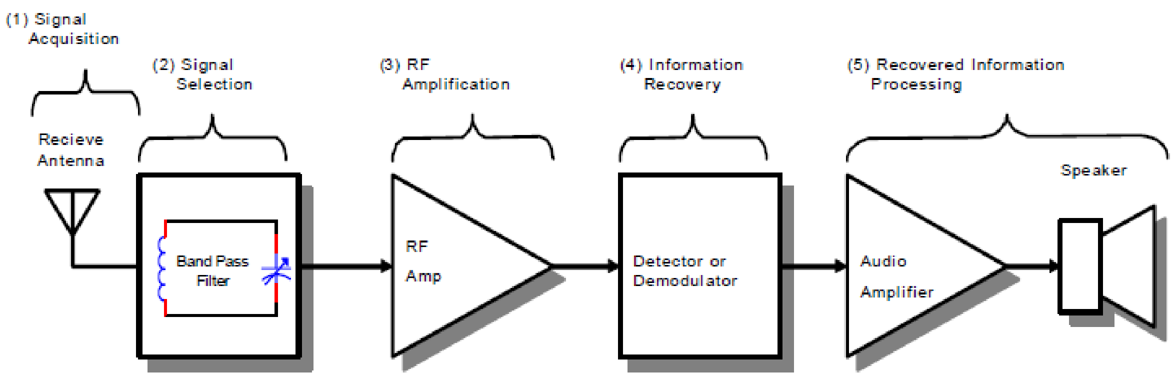
- It should be cost-effective.

- It should receive the corresponding modulated waves.
- The receiver should be able to tune and amplify the desired station.
- It should have an ability to reject the unwanted stations.
- Demodulation has to be done to all the station signals, irrespective of the carrier signal frequency.

For these requirements to be fulfilled, the tuner circuit and the mixer circuit should be very effective. The procedure of RF mixing is an interesting phenomenon.

II. AM Receiver (Receiver Operation)

The process of receiving a radio signal can be broken down into a series of five steps. Not every receiver will perform every step, but most do. Figure shows this in block diagram form.



➢ ***Signal Acquisition*: ** To *acquire* a signal means to get it. Radio signals are in the form of electromagnetic energy traveling through space at the speed of light. In order for a radio signal to be useful in an electronic circuit, it must first be converted back into an electrical signal. This is the job of the *antenna*.

➢ ***Signal Selection*: ** There are thousands of radio signals in the air at any instant in time. An antenna combines many of them in its electrical output to a receiver. Reception of more than one signal at a time would be annoying to the listener. It would be like listening in a crowded room. How can one signal be extracted from the pile? Right --every radio transmitter uses a different *carrier frequency*. The receiver's *bandpass filter*

is tuned to the frequency of the radio station we wish to receive. Ideally, *only* thedesired carrier will get through this filter. In reality, there are problems with thisapproach; filters are not perfect, and interfering signals can get through.

➢ ***RF Amplification*: ** The distance between a radio transmitter and receiver can be very small, or many miles. The transmitted power can be a fraction of a watt, or millions of watts. In general, the signal received at a receiver's antenna is very small. At a receiving antenna, the amplitude of a "strong" received signal is usually 100 µV or less. Many receivers must deal with signals less than 1 µV in size. Before such small signals can be processed, they must be amplified.

➢ ***Information Recovery*: **The actions in the first three steps resulted in reproduction of the *modulated carrier wave* that was sent from the transmitter. The modulated carrierwave holds the information; in order to recover the information, we use a *detector* or*demodulator* circuit. Both words have the same meaning. When we detect a signal, weare extracting the information from the modulated carrier wave. The information issaved and used, and the carrier portion of the wave is discarded.

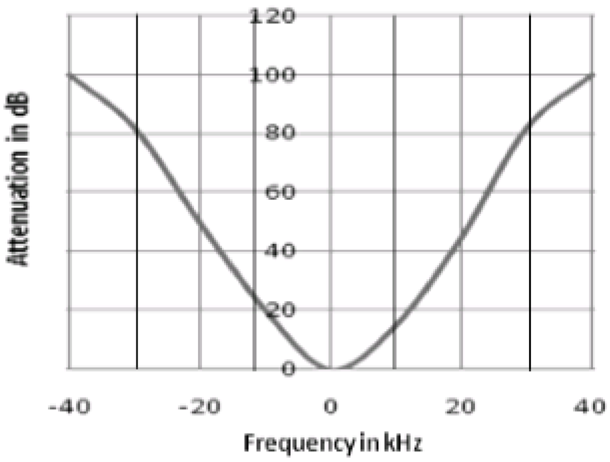
➢ ***Recovered Information Processing*: **This is a general way of saying that we'll be doing something useful with the information the detector extracted. The type of receiver will determine what needs to be done with the information. In a radio receiver, the detected information is an *audio* signal with insufficient voltage and current to drive a loudspeaker. Therefore, the last stage in a radio receiver is an audio power amplifier which provides the voltage and current needed to operate the loudspeaker. For example, a *television* receiver differs from a radio receiver only in how the detected information (a *video signal* in analog TV, or a *data signal* in digital TV) is processed.

III. Receiver Characteristics (Parameters)

Understanding receiver characteristics is mandatory in determining operational condition and for comparing receivers. Important receiver characteristics are Selectivity, Fidelity, Sensitivity, Noise, Bandwidth Improvement Factor, Dynamic Range, Insertion Loss and Double Spotting

Selectivity

- The selectivity of an AM receiver is defined as its ability to accept or select the desired band of frequency and reject all other unwanted frequencies which can be interfering signals.
- Adjacent channel rejection of the receiver can be obtained from the selectivity parameter.
- Response of IF section, mixer and RF section considerably contribute towards selectivity.
- The signal bandwidth should be narrow for better selectivity.
- Graphically selectivity can be represented as a curve shown in Fig. below, which depicts the attenuation offered to the unwanted signals around the tuned frequency.



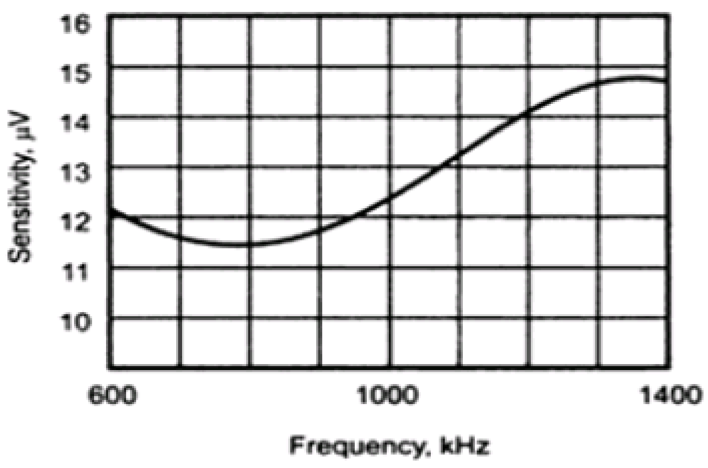
Selectivity curve

Fidelity

- Fidelity of a receiver is its ability to reproduce the exact replica of the transmitted signals at the receiver output.
- For better fidelity, the amplifier must pass high bandwidth signals to amplify the frequencies of the outermost sidebands, while for better selectivity the signal should have narrow bandwidth. Thus a trade off is made between selectivity and fidelity.
- Low frequency response of IF amplifier determines fidelity at the lower modulating frequencies while high frequency response of the IF amplifier determines fidelity at the higher modulating frequencies.

Sensitivity

- Sensitivity of a receiver is its ability to identify and amplify weak signals at the receiver output.
- It is often defined in terms of voltage that must be applied to the input terminals of the receiver to produce a standard output power which is measured at the output terminals.
- The higher value of receiver gain ensures smaller input signal necessary to produce the desired output power.
- Thus a receiver with good sensitivity will detect minimum RF signal at the input and still produce utilizable demodulated signal.
- Sensitivity is also known as receiver threshold.
- It is expressed in microvolts or decibels.
- Sensitivity of the receiver mostly depends on the gain of IF amplifier.
- It can be improved by reducing the noise level and bandwidth of the receiver.
- Sensitivity can be graphically represented as a curve shown in Fig. Below, which depicts that sensitivity varies over the tuning band.



Sensitivity curve

Noise:

All receivers generate noise. Noise is the limiting factor on the minimum usable signal that the receiver can process and still produce a usable output. Ex-pressed in decibels, it is an indication of the degree to which a circuit deviates from the ideal; a noise figure of0 decibels is ideal.

Bandwidth Improvement Factor:

One way of reducing the noise level is by reducing the bandwidth of the signal. There is limitation for reducing the bandwidth to make sure information is not lost. As RF bandwidth at the input of the receiver is higher than the IF bandwidth at the output of the receiver, reducing the RF bandwidth to IF bandwidth ratio effectively reducing the noise figure of the receiver, thus reducing the noise.

Dynamic Range:

The minimum input level necessary to discern a signal and the input that will overdrive the receiver and produce distortion. Minimum receive level is a function of front-end noise, noise figure and the desired signal quality. Input that produce distortion is a function of the net gain of the receiver. 1 dB compression point is used for the upper limit for usefulness.

Insertion Loss :

Loss occur when a signal enters the input of the receiver. Parameters associated with the frequencies that fall within the passband of a filter. Defined as the ratio of the power transferred to the load with a filter in the circuit to the power transferred to the load without a filter.

Double spotting

- Double spotting is a condition where the same desired signal is detected at two nearby points on the receiver tuning dial.
- One point is the desired point while the other is called the spurious or image point.
- It can be used to determine the IF of an unknown receiver.
- Poor front-end selectivity and inadequate image frequency rejection leads to double spotting.
- Double spotting is undesirable since the strong signal might mask and overpower the weak signal at the spurious point in the frequency spectrum.
- Double spotting can be counter acted by improving the selectivity of RF amplifier and increasing the value of IF.
- Consider an incoming strong signal of 1000 kHz and local oscillator tuned at 1455 kHz. Thus a signal of 455 kHz is produced at the output of the mixer which is the IF frequency.
    - Now consider the same signal but with 545kHz tuned local oscillator. Again we get 455 kHz signal at the output.
    - Therefore, the same 1000 kHz signal will appear at 1455 kHz as well as 545 kHz on the receiver dial and the image will not get rejected. This is known as Double spotting phenomenon.
- It is also known as Adjacent channel selectivity.

IV. Types of Receiver:

There are two basic types of radio receivers: coherent and noncoherent.

- With a coherent, or synchronous, receiver, the frequencies generated in the receiver and used for demodulation are synchronized to oscillator frequencies generated in the transmitter (the receiver must have some means of recovering the received carrier and synchronizing to it).
- With noncoherent, or asynchronous, receivers, either no frequencies are generated in the receiver or the frequencies used for demodulation are completely independent from the transmitter's carrier frequency. Noncoherent detection is often called envelope detection because the information is recovered from the received waveform by detecting the shape of the modulated envelope.
- Example for Non Coherent receiver is Tuned Radio-Frequency Receiver.

V. AM Receiver Example

The tuned radio frequency (TRF) receiver:

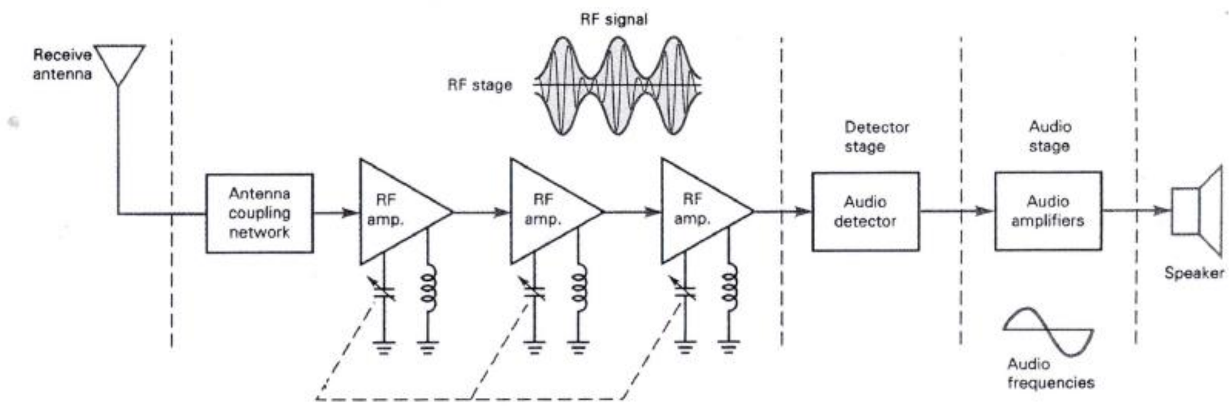
It was one of the earliest types of AM receivers. TRF receivers are probably the simplest designed radio receiver available today; however, they have several shortcomings that limit their use to special applications. The following figure shows the block diagram of a three-stage TRF receiver that includes an RF stage, a detector stage, and an audio stage. Generally, two or three RF amplifiers are required to filter and amplify the received signal to a level sufficient to drive the detector stage. The detector converts RF signals directly to information, and the audio stage amplifies the information signals to a usable level.

The Tuned Radio Frequency Receiver is a simple "logical" receiver. A person with just a little knowledge of communications would probably expect all radio receivers to have this form. The virtues of this type, which is now not used except as a fixed-frequency receiver in special applications, are its simplicity and high sensitivity. It must also be mentioned that when the Tuned Radio Frequency Receiver was first introduced, it was a great improvement on the types used previously mainly crystal, regenerative and superregenerative receivers.

Two or perhaps three RF amplifiers, all tuning together, employed to select and amplify the incoming frequency and simultaneously to reject all others. After the signal was amplified to a suitable level, it was demodulated (detected) and fed to the loudspeaker after being passed through the appropriate audio amplifying stages. Such receivers were simple to design and align at broadcast frequencies (535 to 1640 kHz), but they presented difficulties at higher frequencies. This was mainly because of the instability associated with high gain being achieved at one frequency by a multistage amplifier. If such an amplifier has a gain of 40,000, all that is needed is 1/40,000 of the output of the last stage (positive feedback) to find itself back at the input to the first stage, and oscillations will occur, at the frequency at which the polarity of this spurious feedback is positive.

Such conditions are almost unavoidable at high frequencies and are certainly not conducive to good receiver operation.

Although TRF receivers are simple and have a relatively high sensitivity, they have three distinct disadvantages that limit their usefulness to single-channel, low-frequency applications.



The primary disadvantage is their bandwidth is inconsistent and varies with center frequency when tuned over a wide range of input frequencies. This is caused by a phenomenon called the skin effect.At radio frequencies, current flow is limited to the outermost area of a conductor; thus, the higher the frequency, the smaller the effective area and the greater the resistance. Consequently, the quality factor (Q = XL/R) of the tank circuits remains relatively constant over a wide range of frequencies, causing the bandwidth (f/Q) to increase with frequency. As a result, the selectivity of the input filter changes over any appreciable range of input frequencies. If the bandwidth is set to the desired value for low-frequency RF signals, it will be excessive for high-frequency signals.

The second disadvantage of TRF receivers is instability due to the large number of RF amplifiers all tuned to the same center frequency. High-frequency, multistage amplifiers are susceptible to breaking into oscillations. This problem can be reduced somewhat by tuning each amplifier to a slightly different frequency, slightly above or below the desired center frequency. This technique is called stagger tuning.

The third disadvantage of TRF receivers is their gains are not uniform over a very wide frequency range because of the nonuniform L/C ratios of the transformer-coupled tank circuits in the RF amplifiers. With the development of the superheterodyne receiver, TRF receivers are seldom used except for special-purpose, single-station receivers.

Superheterodyne AM Receiver:

The nonuniform selectivity of the TRF led to the development of the superheterodyne receiver near the end of World War I. Although the quality of the superheterodyne receiver has improved greatly since its original design, its basic configuration has not changed much, and it is still used today for a wide variety of radio communications services. The superheterodyne receiver has remained in use because its gain, selectivity, and sensitivity characteristics are superior to those of other receiver configurations. Heterodyne means to mix two frequencies together in a nonlinear device or to translate one frequency to another using nonlinear mixing.

Heterodyning

Necessity of heterodyning: Superheterodyne AM receiver works on the principle of heterodyning action.

The necessity of heterodyning action is due to the following reasons.

1. It is difficult to design a RF amplifier with high gain and high band width.

2. It is relatively easier to design a high gain IF amplifier having uniform gain over a narrow band of comparatively lower intermediate frequencies (IF).

3. Hence it is necessary to convert the Radio frequencies to Intermediate Frequencies for efficient processing.

Heterodyning Action: Heterodyning action is a process of combining two ac signals of different frequencies in-order to obtain signals of new frequencies. A circuit called mixer or converter is used for heterodyning two signals. If f1 and f2 are the two frequencies combined, then heterodyning results in two components

1. The sum component with frequency f1+f2 which is filtered out using a bandpass filter.

2. The difference component with frequency f1- f2 is retained and processed.

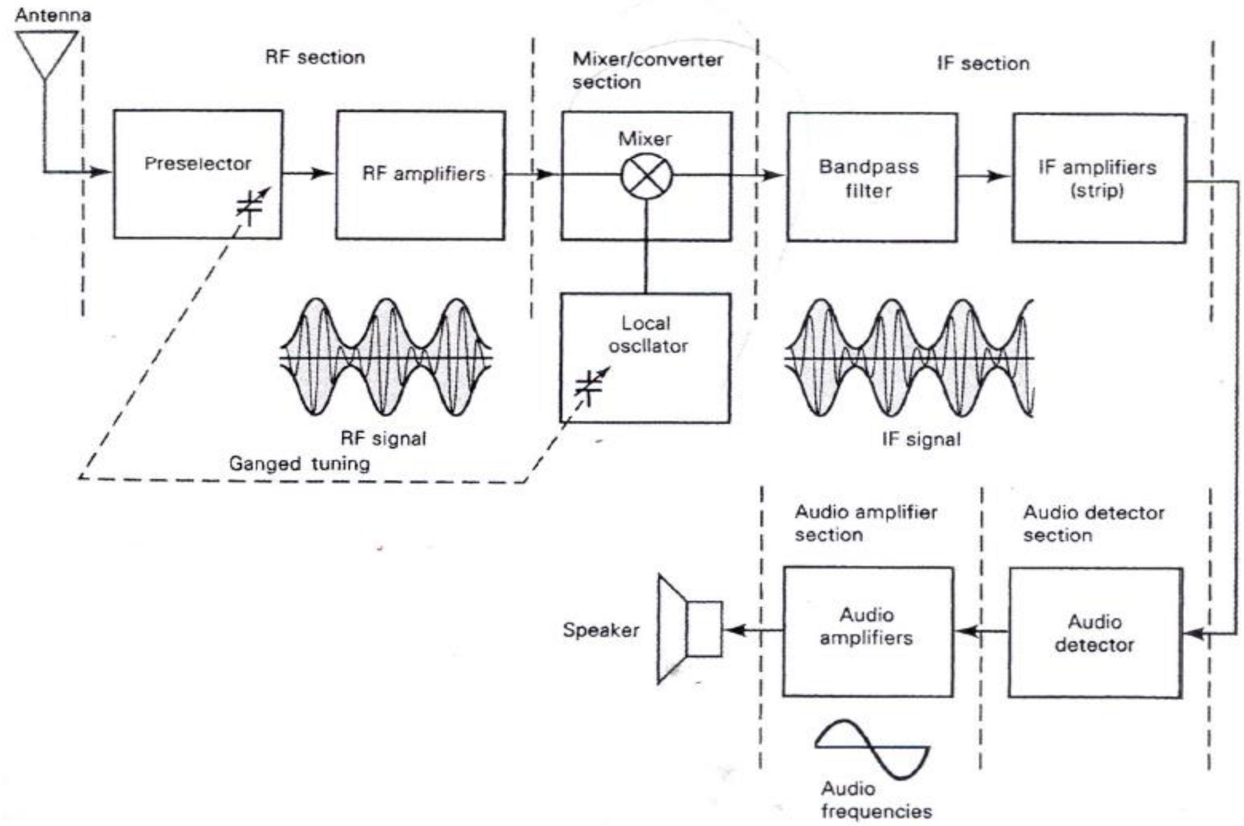
In case of superheterodyne receiver the RF carrier fc is heterodyned with a higher RF local signal fs (From Local Oscillator or BFO) so that the output difference component (fs- fc) is always of frequency 455kHz.

A block diagram of a noncoherent superheterodyne receiver is shown in Figure.

Essentially, there are five sections to a superheterodyne receiver: the RF section, the mixer/converter section, the IF section audio detector section, and the audio amplifier section.

RF section: The RF section generally consists of a preselector and an amplifier stage. They can be separate circuits or a single combined circuit. The preselector is a broad bandpass filter with an adjustable centre frequency that is tuned to the desired carrier frequency The primary purpose of the preselector is to provide enough initial band limiting to prevent unwanted radio frequency, called the image frequency, from entering the receiver (frequency is explained later in this section). The preselector also reduces the noise bandwidth the receiver and provides the initial step toward reducing the overall receiver bandwidth minimum bandwidth required to pass the information signals. The RF amplifier determine sensitivity of the receiver (i.e., sets the signal threshold). Also, because the RF amplifier first active device encountered by a received signal, it is the primary contributor of not therefore, a predominant factor in determining the noise figure for the receiver. A receiver has one or more RF amplifiers, or it may not have any, depending on the desired sensitivity: Several advantages of including RF amplifiers in a receiver are as follows:

1. Greater gain, thus better sensitivity 2. Improved image-frequency rejection 3. Better signal-to-noise ratio 4. Better selectivity



Mixer/converter Section: The mixer/converter section includes frequency oscillator stage (commonly called a local oscillator) and a mixer/convert (commonly called the first detector). This stage is a nonlinear device and its purpose is to convert radio frequencies to inters frequencies (RF-to-IF frequency translation). Heterodyning takes place in the mixer and radio frequencies are down-converted to intermediate frequencies. Although the and sideband frequencies are translated from RF to IF, the shape of the envelope the same and, therefore, the original information contained in the envelope remain changed. It is important to note that although the carrier and upper and lower side frequencies change frequency, the bandwidth is unchanged by the heterodyning process. The common intermediate frequency used in AM broadcast-band receivers is 455 kHz

IF section: The IF section consists of a series of IF amplifiers and the filters and is often called the IF strip. Most of the receiver gain and selectivity is act the IF section. The IF center frequency and bandwidth are constant for all stations. So that their frequency is less than any of the RF signals to be received. The IF is lower in frequency than the RF because it is easier and less expensive to construct hi stable amplifiers for the low-frequency signals. Also, low-frequency IF amplifiers likely to oscillate than their RF counterparts. Therefore, it is not uncommon to see a with five or six IF amplifiers and a single RF amplifier or possibly no RF amplification.

Detector section: The purpose of the detector section is to convert signals back to the original source information. The detector is generally called detector or the second detector in a broadcast-band receiver because the inform signals are audio frequencies. The detector can be as simple as a single diode or as a phase-locked loop or balanced demodulator.

Audio amplifier section. The audio section comprises several audio amplifiers and one or more speakers. The number of amplifiers used depend audio signal power desired.

VI. FM Receiver

A radio or FM receiver is an electronic device that receives radio waves and converts the information carried by them to a usable form. An antenna is used to catch the desired frequency waves. The receiver uses electronic filters to separate the desired radio frequency signal from all the other signals picked up by the antenna, an electronic amplifier to increase the power of the signal for further processing, and finally recovers the desired information through demodulation.

Of the radio waves, FM is the most popular one. Frequency modulation is widely used for FM radio broadcasting. It is also used in telemetry, radar, seismic prospecting, and monitoring newborns for seizures via EEG, two-way radio systems, music synthesis, magnetic tape-recording systems and some video-transmission systems. An advantage of frequency modulation is that it has a larger signal-to-noise ratio and therefore rejects radio frequency interference better than an equal power amplitude modulation (AM) signal.

FM frequency ranges

Frequency modulation is used in a radio broadcast in the 88-108MHz VHF band. This bandwidth range is marked as FM on the band scales of radio receivers, and the devices that are able to receive such signals are called FM receivers. The FM radio transmitter has a 200kHz wide channel. The maximum audio frequency transmitted in FM is 15 kHz as compared to 4.5 kHz in AM. This allows a much larger range of frequencies to be transferred in FM and thus the quality of FM transmission is significantly higher than of AM transmission.

Stereo VHF FM Broadcast

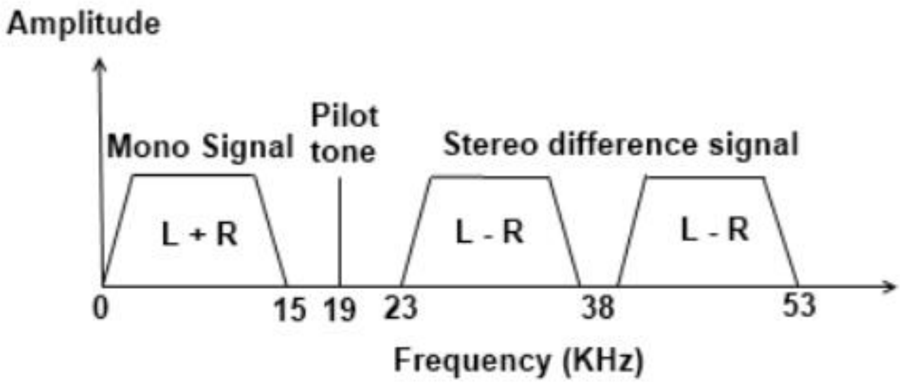
In recent years’ stereo transmission has become an accepted part of VHF FM transmissions. The system that is used maintains compatibility with mono only receivers without any noticeable degradation in performance. The system that is used is quite straightforward.

A stereo signal consists of two channels that can be labelled L and R, (Left and Right), providing one channel for each of the two speakers that are needed. An ordinary mono signal consists of the summation of the two channels, i.e. L + R, and this can be transmitted in the normal way. If a signal containing the difference between the left and right channels, i.e. L - R is transmitted then it is possible to reconstitute the left only and right only signals. Adding the sum and difference signals, i.e. (L + R) + (L - R) gives 2L, i.e. the left signal, and subtracting the two signal, i.e. (L + R) - (L - R) gives 2R, i.e. the right signal. This can be achieved relatively simply by adding and subtracting the two signals electronically. It only remains to find a method of transmitting the stereo difference signal in a way that does not affect any mono receivers.

This is achieved by transmitting the difference signal above the audio range. It is amplitude modulated onto a 38 kHz subcarrier. Both the upper and lower sidebands are retained, but the 38 kHz subcarrier itself is suppressed to give a double sideband signal above the normal audio bandwidth as shown below. This whole of the baseband is used to frequency modulate the final radio frequency carrier. It is the baseband signal that is regenerated after the signal is demodulated in the receiver.

To regenerate the 38 kHz subcarrier, a 19 kHz pilot tone is transmitted. The frequency of this is doubled in the receiver to give the required 38 kHz signal to demodulate the double sideband stereo difference signal.

The presence of the pilot tone is also used to detect whether a stereo signal is being transmitted. If it is not present, the stereo reconstituting circuitry is turned off. However, when it is present the stereo signal can be reconstituted.



The modulating, (base-band) signal for a stero VHF F.M transmission

Stereo Broadcast Transmission Sidebands

Comparison Stereo amplifier with Mono amplifier

Sl.No	Stereo Amplifier	Mono Amplifier
1	As the name suggest, the basic stereophonic system has two separate channels after the pre amplifier stage.	There is only one channel after pre amplifier stage.
2	Stereophoic sound is created by two independent audio channels to provide sense of direction	Sense of direction is not pronounced as in stereophonic amplifier.
3	Consists of two preamplifiers power amplifiers and loud speakers	Consists of one preamplifiers tone and volume controls, power amplifier.
4	Stereophony creates the impression of sound heard from various directions as in natural hearing	In monophonic or 'mono' sound, the audio is in the form of one channel, often centered in the sound field.
5	Less signal to noise ratio	Better than 50 dB is the S/N ratio.
6	Non-linear distortion occurs.	Non-linear distortion not more than input/output.
7	Equalizers are not used	Contains equalizer circuit.
8	Two-way cross-over network with gain control exist.	Three-way cross-over network exist.

STEREO SIGNAL GENERATOR (Transmitter)

To generate the stereo signal, a system similar to that shown in Fig. 8.5 is used. The left and right signals enter the encoder where they are passed through a circuit to add the required pre-emphasis. After this they are passed into a matrix circuit. This adds and subtracts the two signals to provide the L + R and L - R signals. The L + R signal is passed straight into the final summation circuit to be transmitted as the ordinary mono audio. The difference L - R signal is passed into a balanced modulator to give the double sideband suppressed carrier signal centred on 38 kHz. This is passed into the final summation circuit as the stereo difference signal. The other signal entering the balanced modulator is a 38 kHz signal which has been obtained by doubling the frequency of the 19 kHz pilot tone. The pilot tone itself is also passed into the final summation circuit. The final modulating signal consisting of the L + R mono signal, 19 kHz pilot tone, and the L - R difference signal based around 38 kHz is then used to frequency modulate the radio frequency carrier before being transmitted.

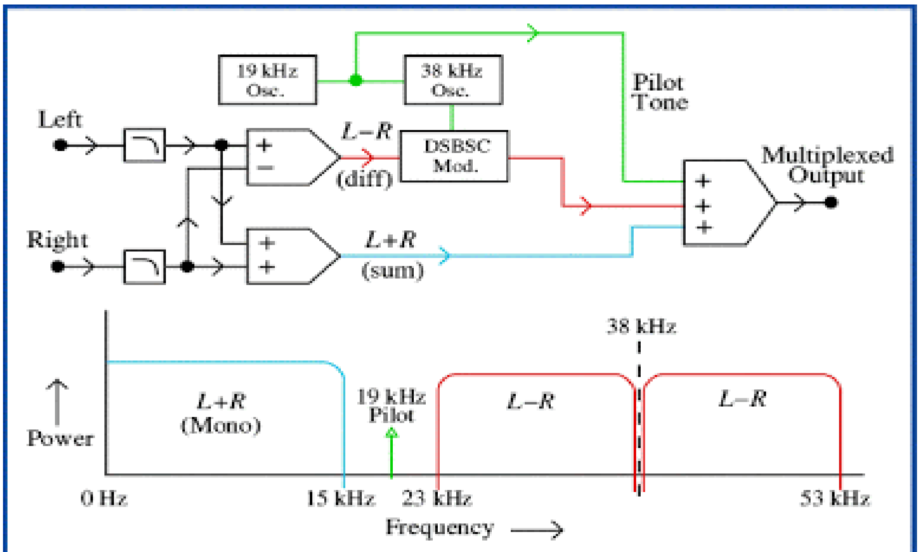
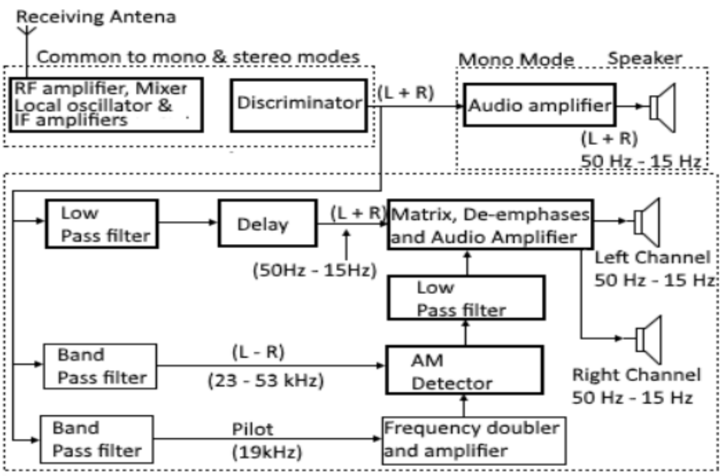


Figure FM Stereo Multiplex signal.

FM Stereo Receiver:

Reception of a stereo signal is very much the reverse of the transmission. A mono radio receiving a stereo transmission will only respond to the L + R signal. The other components being above 15 kHz are above the audio range, and in any case they will be suppressed by the de-emphasis circuitry.

For stereo receivers the baseband signal consisting of the stereo sum signal (L+R) and the difference signal (L-R) centered around 38 kHz and the pilot 19kHz tone are obtained directly from the FM demodulator. The decoder then extracts the Left only and Right only signals.



Stereo FM Receiver Block Diagram

A stereo FM receiver has three major sections:

- Mono mode
- Stereophonic mode
- Section common to both mono and stereo modes

The section that is common to both mono and stereo modes is a standard FM receiver that recovers the modulating signal. The output of this section is routed to the remaining two sections. The output consists of both the left and right channel marked as (L + R) in Figure. This output is applied to the mono section and the speaker produces audio signals monophonic mode.

The stereo section is more complicated. It uses three filters to extract (L + R) and (L — R) signals and the pilot-carrier from the discriminator output. The (L + R) signal is obtained from the low-pass filter, which contains frequencies between 50 Hz and 15 kHz. This signal delayed for a fixed time before applying it to the matrix and the de-emphasis network. This is done to simultaneously get the (L + R) and (L - R) signals at the matrix. The matrix network separates the left (L) and right (R) channels. These are then de-emphasized and amplified by the audio amplifiers and are given to their respective speakers.

A band pass filter is used to extract the (L - R) signal varying between 23-53 kHz. It is a double-side band (DSB) signal. This signal is applied to an AM detector to demodulate. The transmitter uses a 38 kHz carrier signal to get a DSB-SC signal from the (L - R) signal. Thus, at the receiver, a carrier of 38 kHz is required to demodulate the received (L - R) signal.

The pilot carrier of 19 kHz is extracted using another band pass filter. This pilot carrier is given to the frequency doubler, which doubles its frequency to 38 kHz. After amplification of this The AM detector detects the (L - R) signal, which is carrier, it is applied to the AM detector matrix. As some time is taken for the (L —R) signal to demodulate, the (L + R) signal is delayed so that both (L + R) and (L — R) reach the matrix at the same time.

Trouble shooting FM Receiver

To locate, the faulty stage in the receiver, use the signal injection method and analyze the waveforms on the oscilloscope or examine the audio from the speaker. Before starting the fault-finding exercise in the receiver, consult the service manual. The manual will provide the voltages at the various key points or the circuit and the details or the various signals and their waveforms at the key points.

You may start injecting the signal from the last stage for a dead receiver. However, ensure that the dc power supply and the speaker are not faulty. Then start injecting the signal at the input of each stage and examine the audio output from the speaker. The faulty stage will not pass the signal. Once the faulty stage is identified, use the usual troubleshooting techniques to check to faulty components.

VII. Automatic Gain Control (AGC)

- Automatic gain control (AGC) is a mechanism wherein the overall gain of the radio receiver is automatically varied according to the changing strength of the received signal. This is done to maintain the output at a constant level.
- AGC facilitates tuning to varying signal strength stations providing a constant output. AGC smoothens the amplitude variations of the input signal and the gain control does not have to be recalibrated every time the receiver is tuned from station to station.
- An AGC which is not designed correctly can lead to considerable distortion to a smooth signal.
- There are two types of AGC circuits:

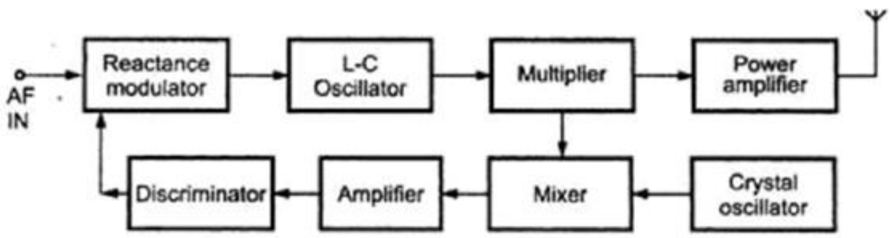
i. Simple AGC: the gain control mechanism is active for high as well as low value of carrier voltage.

ii. Delayed AGC: AGC bias is not applied to the amplifiers until signal strength crosses a predetermined level, after which AGC bias is applied.

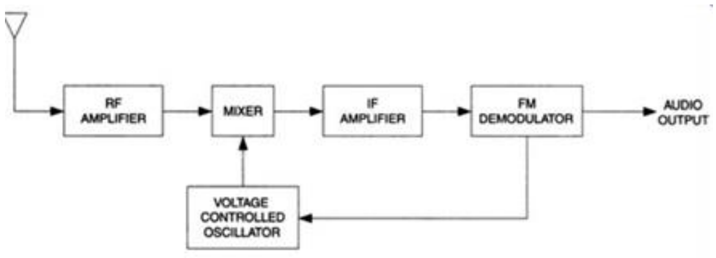
VIII. Automatic Frequency Control (AFC)

Where the main oscillator is only an L-C oscillator then it is not capable of keeping the carrier frequency stable. The carrier frequency can be kept stable by using an automatic frequency control AFC circuit. The radar receiver requires a limited tuning range to compensate for transmitter and local oscillator frequency changes because of variations in temperature and loading. Microwave radar receivers usually use automatic frequency control (AFC) for this purpose.

It is the heart of an AFC circuit is a frequency-sensitive device, such as the phase discriminator, which produces a dc voltage whose amplitude and polarity are proportional to the amount and direction of the local oscillator frequency error. This dc control voltage is then used to vary, automatically, the bias on a variable-reactance device, whose output capacitance is thus changed. This variable capacitance appears across the (first) local oscillator coil, and the frequency of this variable-frequency oscillator (VFO), is automatically kept from drifting with temperature, line voltage changes or component aging. An Automatic Frequency Control Block Diagram system is shown in Figure



The discriminator reacts only to small changes in the carrier frequency but not to the frequency deviations in the carrier (since it is too fast). Suppose frequency of the carrier increases. This higher frequency is fed to the mixer for which the other input frequency is from the stable crystal oscillator. A somewhat higher frequency will be fed to the discriminator. Since the discriminator is tuned to the correct frequency difference which should exist between the LC oscillator and crystal oscillator, and its input frequency is now somewhat higher, the discriminator will develop a positive dc voltage. This voltage is applied to the reactance modulator whose transconductance is increased by the positive voltage developed by the discriminator. This increases the equivalent capacitance of the reactance modulator thereby decreasing the oscillator frequency. The frequency increase in the carrier frequency is thus lowered and brought to the correct value. The correcting dc voltage developed by the discriminator may be fed to a varactor diode connected across the tank circuit of the oscillator and be used for AFC purposes.



Automatic frequency control circuit.

Automatic frequency control is used in some FM receivers to keep the tuning frequency stable. This type of circuit is confined largely to older, analog FM receivers, because modem digital frequency synthesizers prevent the kind of drift and outright frequency shifting common in analog receivers. A typical modern AFC system is shown in above figure. The AFC output of the FM demodulator is 0 V when the input signal frequency is directly on the resonant frequency of the demodulator, but positive or negative voltage depending whether the frequency is above or below resonance. This voltage is fed back to a voltage controlled oscillator circuit, which feeds the mixer. in practice, a varactor (Van- able capacitance diode) is used to shunt the 1-C or quartz crystal in the oscillator, and this varactor converts the oscillator into a VCO which then takes care of the frequency correction or restoration.

IX. Capture Effect

A phenomenon, associated with FM reception, in which only the stronger of two signals at or near the same frequency will be demodulated

- The complete suppression of the weaker signal occurs at the receiver limiter, where it is treated as noise and rejected.
- When both signals are nearly equal in strength, or are fading independently, the receiver may switch from one to the other.

In the frequency modulation, the signal can be affected by another frequency modulated signal whose frequency content is close to the carrier frequency of the desired FM wave. The receiver may lock such an interference signal and suppress the desired FM wave when interference signal is stronger than the desired signal. When the strength of the desired signal and interference signal are nearly equal, the receiver fluctuates back and forth between them, i.e., receiver locks interference signal for some times and desired signal for some time and this goes on randomly.

- This phenomenon is known as the capture effect.

X. Threshold Effect

Threshold effect in AM Receiver:

- When the carrier to noise ratio reduces below certain value, the message information is lost. The performance of envelope detector deteriorates rapidly and it has no proportion to carrier to noise ratio. This is called threshold effect.
- Every nonlinear receiver exhibits threshold effect. Coherent receivers do not have threshold effect.
- The detector output does not depend only on message signal m(t), rather it is the function of noise also.
- When the noise is higher compared to signal, the noise dominates the performance of the receiver.
- Let us express noise in terms of envelope and phase components.

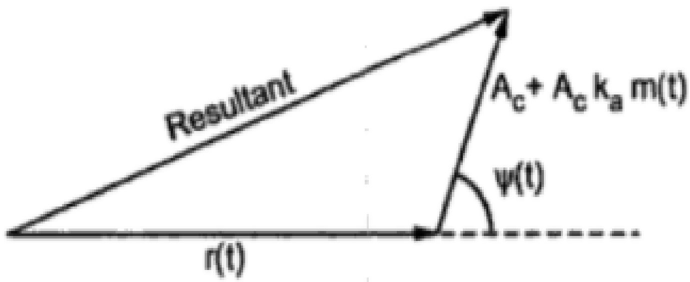
$$n(t) = r(t)\cos(2\,\pi\, f_c\, t + \psi(t)\,)$$

- Here r(t) is magnitude of noise and ψ(t) is the phase of noise.
- The signal plus noise will be,

$$x(t) = s(t) + n(t)$$

$$x(t) = A_c\,[1 + k_a m(t)]cos2\pi f_c t + r(t)cos[2\pi f_c t + \psi(t)]$$

$$x(t) = [A_c + A_c k_a m(t)]cos2\pi f_c t + r(t)cos[2\pi f_c t + \psi(t)]$$



- In this figure observe that the phasor r(t) is added to phasor $[A_c + A_c k_a m(t)]$.
- As per equations the angle between these phasors is ψ. Since Ac is very small compared to r(t), we can approximately express the resultant y(t) as

$$y(t) = r(t) + A_c\cos\psi(t) + A_c k_a m(t)\cos\psi(t)$$

- The above equation shows that output of envelope detector is the function of noise components as well as message.
- Output is not strictly proportional to message signal.

Threshold effect in FM Receiver:

- As the carrier to noise ratio is reduced, clicks are heard in the receiver output.

- As the carrier to noise ratio reduces further, crackling, or sputtering sound appears at the receiver output.
- Near the breaking point the theoretically calculated output signal to noise ratio becomes large, but its actual value is very small. This phenomenon is called threshold effect.
- Definition of threshold It is the minimum carrier to noise ratio yielding an FM improvement which is not significantly deteriorated from the value predicted by the usual signal to noise formula assuming small noise.
- Consider that the carrier is unmodulated. The signal at the output of FM discriminator is represented as,

$$x(t) = s(t) + n(t)$$

Here s(t) = Ac cos(2πfct) with no modulation, since φ(t) =0

Putting for s(t) and n (t) in above equation,

$$x(t) = A_c \cos(2\pi f_c t) + n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t)$$
$$x(t) = [A_c + n_c(t)]\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t)$$

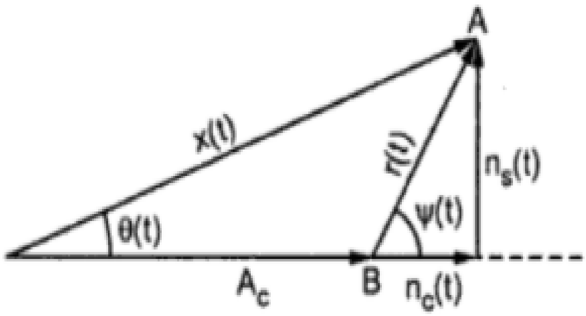


Fig. Phasor representation of above equation

- The amplitude and phases of $n_c(t)$ and $n_s(t)$ change randomly with time. The point A wanders randomly around point B.
- The angle θ(t) varies approximately from $n_s(t)/Ac$ to multiples of 2π. Whenever θ(t) changes by ±2π, clicks are produced in the discriminator output.
- As the carrier to noise ratio is decreased further, the clicks per unit time increase. The threshold is said to occurred when these clicks are very large.
- The signal and noise processes at the output of the demodulator are completely mixed in a single process by a complicated non-linear functional.

XI. Squelch circuit

The term "squelch " means suppress or crush completely. The concept is applied in receiver where in audio part is turned off completely until RF signal is appears at the input.

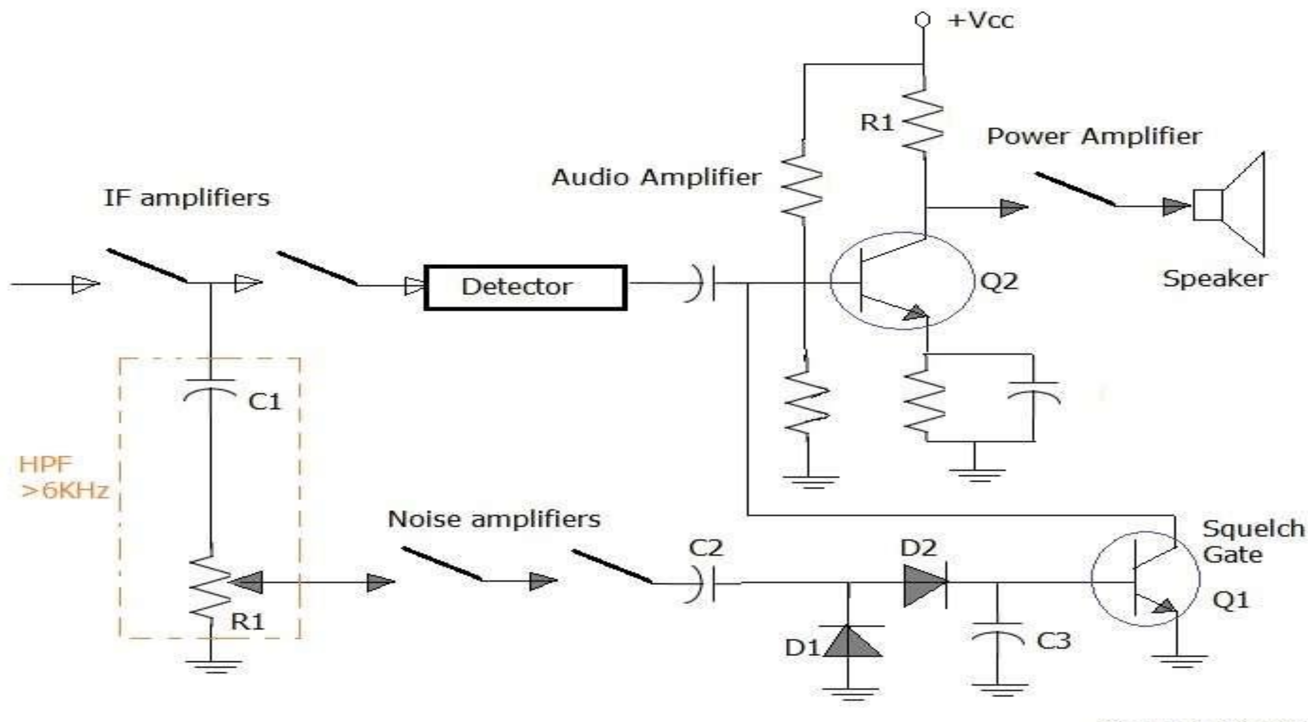
As we know if receiver in the mobile phone is kept on continuously, it will receive noise in the absence of call. This noise level is usually high and will also get amplified and generates annoying noise. Most of the people do not like to listen such noises. Hence squelch circuit has been developed which turns on the receiver only when useful call is present at the input and keeps the receiver off when noise signal is present. Let us see how squelch circuit functions as explained below.

Function of Squelch circuit

As depicted in the figure, there are two main parts in a squelch circuit viz. high pass filter and audio amplifier. There are two transistors Q1 and Q2. Q1 acts as squelch gate which drives the base of the Q2 transistor.

As we know most of the audio frequencies are below 4KHz. The high pass filter is designed to pass the frequencies above these audio frequency range. Hence when audio is present, no squelch voltage is developed at the receiver output. Here Q1 transistor is cut off and hence Q2 transistor gets biased as usual. As a result audio signal is passed through the audio power amplifier and to the speaker.

Let see how squelch circuit functions when noise is present. As we know most of noise signal is of high frequency, it is amplified by two transistor stages and later rectified into DC control voltage by rectifier-voltage double circuit(made of D1, D2, C2, C3). This rectifier output will drive Q1 transistor to saturation region. The base current of Q2 transistor is shunted away from Q1. Hence no audio amplification takes place and receiver will be quiet. This squelch circuit is widely used in communication receiver. It is also known as mute circuit.

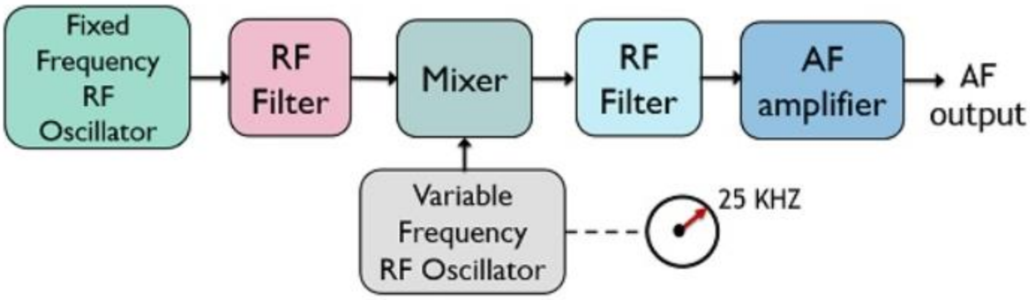


Squelch Circuit

XII. The beat frequency oscillator (BFO):

It is used to produce a variable frequency output in the audio-frequency (AF) range. BFO is used when the need comes to cover a very large frequency range with a single dial rotation as it can produce a very large frequency range with a single dial rotation.

The figure below represents the block diagram of a BFO:



Block Diagram of Beat Frequency Oscillator

The BFO has mainly two RF oscillators. One of the oscillator gives a fixed frequency and the other one produces variable frequency. The variable frequency will be slightly different from the fixed frequency. The fixed and variable frequency outputs are fed to a heterodyne or mixer device. The sum and difference terms of frequencies f1 and f2 are obtained as the output of the mixer. It is so arranged that the difference terms of frequencies f1 and f2 lies in the audio-frequency range. All the RF components, leaving only the audio-frequency difference component, are removed in the RF filter. Audio-frequency output is then amplified in the AF amplifier.

The practical value of a beat frequency oscillator arises from the fact that a small or moderate percentage variation in the frequency of one of the individual oscillators (such as can be had by the rotation of the shaft controlling a variable tuning capacitor) varies the beat or difference output continuously from a few Hz to throughout the entire audio-frequency range. At the same time, the amplitude of the difference frequency output is largely constant as frequency is varied.

Frequency stability of the individual oscillators is important, because a slight change in their relative frequency would cause a relatively large change in the difference frequency. To minimize the drift of the difference frequency with time, the individual oscillators should have high inherent stability with respect to variations in temperature and to supply voltage variations.

It should be noted that the two RF oscillators are completely isolated from each other. If there is any sort of coupling between them, they will synchronize when the difference is small. Hence, low values of difference frequencies are impossible to be obtained, and in addition cause interaction between the oscillators that result in a highly distorted wave shape. To reduce distortion in the output, one of the voltages applied to the mixer (preferably the one derived from the fixed frequency oscillator) should be considerably smaller than the voltage derived from the other oscillator, and preferably free from harmonics.

XIII. NOISE

- The term noise is used to designate unwanted signals that tend to disturb the transmission and processing of signals in communication systems.
- Electrical noise is defined as any undesirable electrical energy that falls within the passband of the signal.

- In telecommunication, noise is described as the electrical disturbance that gives rise to audible noise in a system.
- In video systems as White flecks on TV picture when the signal received is weak. Picture in such case is referred to as noisy picture.
- Correlation implies a relationship between the signal and noise Therefore, correlated noise exists only when a signal is present.
- Uncorrelated noise, it is present all the time even if the signal is not present

Noise Sources and types:

- Noise is random, undesirable electrical energy that enters the communication system via the communication medium and interferes with the transmitted image.
- There are many potential sources of noise in communication system.

The sources of noise may be

(i) External noise and (ii) Internal noise

External Noise:

This noise is generated outside the device or circuit. There are number of external sources of noise. These are grouped into three categories:

(a) Atmospheric noise

(b) Extra-terrestrial (or galactic) noise

(c) Man-made noise (or industrial)

Atmospheric Noise:

- The noise generated due to the electrical disturbance within the earth's atmosphere is called atmospheric noise. This type of noise creates strange, sounds like sputtering, cracking, etc., in short wave receivers.
- These electrical impulses are random in nature. Hence the energy is spread over the complete frequency spectrum used for radio communication.
- Atmospheric noise is commonly called static electricity. It is caused by lightning and thunderstorms (local or distant). The static is likely to be more severe but less frequent if the storm is local. It is in the form of impulses that spread energy throughout a wide range of frequencies.
- The magnitude of this energy is inversely proportional to the frequency. Hence, at frequencies above 80 MHz, atmospheric noise is less relevant.
- The atmospheric noise interferes more with reception of radio than that of Television. The reason for this noise becoming less severe at frequencies above 30 MHz is due to line of sight (LOS) propagation.
- The nature of the mechanism generating this noise is such that very little of it is created in the very high frequency (VHF) range and above.

Extra-terrestrial Noise:

It consists of electrical signals that originate from outside the earth's atmosphere and is, therefore sometimes called deep space noise. It originates from the milky way, other galaxies, and the sun.

Extra-terrestrial noise is classified into two groups

(1) Solar noise (2) Cosmic noise

Solar Noise:

- This noise is generated directly from the sun's heat. This is due to the radiation from the sun, which is of two types. One takes place in a situation called quiet condition, when relatively constant radiation intensity exists.
- In such a condition, the radiation, radiates over a very broad frequency spectrum, which includes the frequencies we use for communication.
- The other type of radiation arises due to the sun spot activity and solar flare-up. This high-intensity radiation occurs approximately every 11 years.
- Although this high intensity noise comet from a limited portion of the sun's surface, it may still be of the order of magnitude greater than that received during the period of quiet sun.

Cosmic noise

- The sources of COSMIC noise are continuously distributed throughout the galaxies since distant stars are also suns and have high temperatures they radiate RF noise in the same manner as the sun.
- These sources are located much farther away from the sun and their noise intensity is relatively small. The noise received is called **black body noise** and is distributed evenly throughout the sky.
- We also receive noise from the center of our own galaxy (The Milky Way) from other distant galaxies and from other virtual paint sources such as quasars and pulsars.
- Space noise is observable at frequencies in the range from about 8 MHz to 1.43 GHz. Apart from the man - made noise, it is the strongest component over the range of about 20 to 120 MHz.
- The noise below 20 MHz does not penetrate much through the ionosphere but it disappears at frequencies is excess of 1.5 GHz.

Man - made Noise (or Industrial Noise):

- Thin is simply the noise generated by the human race. It lies between the frequencies of 1 MHz and 600 MHz. The intensity of this noise overtakes the noise created by any other source, internal or external to the receiver.
- The predominant sources of manmade noise are automobile and aircraft ignition, electric motor, switching equipment, leakage from high voltage lines, and a multimode of other heavy electric machines
- Fluorescent lights are another powerful source, of this noise and should avoided near very sensitive receivers. This noise is a impulsive in nature and contains a wide range of frequencies that propagate through space in the, same manner as radio waves.
- They are most intense in densely populated and industrial areas. Hence, it is called industrial noise.

- The nature of industrial noise is so variable that it is difficult to analyze it on any basis other than static. But the received noise increases as the receiver bandwidth increases.

Internal Noise:

- This is the noise generated within a device or circuit. It can be generated by any of the active or passive devices found in a receiver. Such noise is distributed randomly over the entire radio spectrum.
- Random noise power is proportional to the bandwidth over which it measured.
- This noise arises from spontaneous fluctuations of current or voltage in electrical circuits
- This type of noise represents a basic limitation on the transmission or detection of signals in communication systems involving the use of electronic devices.
- Various types of internal noise are

  (a) Thermal noise or white noise or Johnson noise.

  (b) Shot noise

  (c) Transit time noise

  (d) Miscellaneous noise or Flicker noise.

Thermal noise (Johnson Noise)

- This type of noise is generated by all resistances. Thermal noise is the electrical noise arising from the random motion of electrons in a conductor due to thermal agitation.
- Conductors contain a large number of free electrons and ions strongly bound by molecular forces. The ions, vibrate randomly about their normal (average) positions, however this vibration being the function of the temperature.
- Continuous collisions between the electrons and the vibrating ions take Place. Thus there is a continuous transfer of energy between the ions and electrons. This is the source of resistance in a conductor.
- The movement of free electrons constitutes a current which is purely random in nature and over a long time average zero. This random motion of the electrons which give rise to noise voltage called thermal noise or Johnson noise.
- The average or mean noise voltage across a conductor is zero, but the root mean square (RMS) value is finite. This RMS value of the noise voltage proportional to the resistance of the conductor, its absolute temperature. This noise is also called white noise. It contains all spectral frequencies equally on an average. When there is no FM signal this noise can be heard as a sound 'hiss".
- The analysis of thermal noise is based on the kinetic theory that is temperature of particles is a way of expressing its internal kinetic energy. Johnson (By experiment) proved that the thermal noise power is proportional to the product of the bandwidth (B) and the temperature (T).

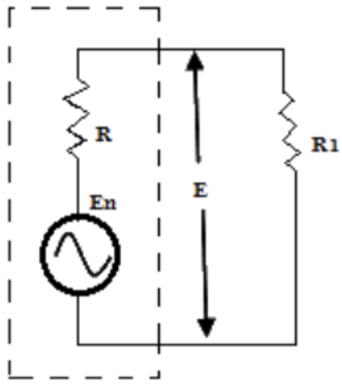
$$P_n \propto TB = kTB -------------- \boldsymbol{A}$$

where

- ✓ P_n- noise power in watt
- ✓ k - Boltzmann's constant (1.38 x 10^{-23} J/K.)

- ✓ B - Bandwidth in Hz
- ✓ T - absolute temperature in kelvin

From equation **A** noise equivalent circuit can be drawn as follow



From figure if $R = R_1$, $E_n = E/2$

$$Pn = \frac{E^2}{R_1} = \frac{E^2}{R} = \frac{(E_n/2)^2}{R} = \frac{{E_n}^2}{4R}$$

$${E_n}^2 = 4RP_n = 4RKTB$$

$$E_n = \sqrt{4RKTB}$$

Short Noise:

- ➢ Any direct current crossing a potential barrier in a random fashion results in shot noise. This occurs because the carrier holes and electrons do not cross the barrier simultaneously, but with the random distribution in the timing for each carrier. This gives rise to a random component of current super imposed on the steady current.
- ➢ Shot noise arises in electronic devices such as diodes the transistors because the discrete nature of current flow in these devices.
- ➢ Example: Ina photodetector circuit a current pulse is generated every time an electron is emitted by the cathode due to incident light from a source of constant intensity.

Transit Time noise:

- ➢ This is noise due to transit time and it occurs in transistors
- ➢ Transit time is the duration of time that it takes for current as a hole or election to move from the input to output (or if the time taken by an electron to travel from the emitter to the collector in a transistor it becomes significant to the period of signal being amplified)
- ➢ The transit time noise in a transistor is determined by the carrier mobility bias voltage and transistor construction.

- If transit delays are excessive at high frequencies the device may add more noise than amplification to the signal causing frequency distortion. At low frequencies this distortion is negligible.
- The transit time shows up as a kind of random noise within the device this is directly proportional to the frequency of operation.

Miscellaneous Noise or Flicker Noise:

- At low frequencies. that is. at low audio frequencies of a few kHz component of noise appears whose spectral density increases as the frequency decreases. This is known as Flicker noise. Otherwise called as 1/f noise or pink noise
- It is proportional to the emitter current and junction temperature This noise is inversely proportional to the frequency. Hence it may be neglected frequencies above 500 Hz.
- The mean square voltage will be proportional to the square of the direct current This flicker noise limits the sensitivity of microwave diode mixers used for Doppler radar system. This is because, although the input frequencies to the mixer are in microwave range, the Doppler frequency output is in the low audio frequency range.
- Partition Noise occurs whenever the current has to divide between two or more electrodes. The reason for this is the random fluctuations in the division. Hence, a diode is less noisy than a transistor or an FET as it has less junctions.
- Burst noise or popcorn noise arises at low frequencies and in transistors. This noise appears as a series of bursts at two or more levels. When present in an audio system, it produces popping sound.

XIV. Noise Figure

- Noise figure is a figure of merit and used to indicate how much S/N ratio gets degraded as a signal passes through a circuit or series of circuit.
- It is the ratio of the input signal-to-noise power ratio to the output signal-to-noise power ratio expressed in decibels.

Noise Figure. NF (dB) = 10 $\log_{10}$ F
where F- is the noise factor.

The noise factor F of an amplifier, or any other network is defined as.

$$Noise\ Factor, F = \frac{\text{input signal} - \text{to} - \text{noise power ratio}}{\text{output signal} - \text{to} - \text{noise power ratio}}$$

$$F = \frac{[P_{si}/P_{Ni}]}{[P_{so}/P_{No}]} = \frac{P_{si}}{P_{Ni}} \times \frac{P_{No}}{P_{so}}$$

- Example: an amplifier with a noise figure of 3 dB means that the signal-to-noise ratio at the output is 3 dB less than it was at the input. If an amplifier or circuit is perfectly noiseless and adds no additional noise to the signal the S/N ratio at the output will equal to S/N ratio at the input. Hence, for a noiseless circuit, the noise Factor F is 1 and its noise Figure NF = 0dB.

- Let us consider the relation between the noise factor F and the available output noise power G. In many cases, the noise factor depends upon frequency and it is calculated at one single frequency where it is known as spot noise Factor Fav.
- An average value of the noise factor can be calculated over a given frequency range

$$\text{The available power gain G} = \frac{Signal\ power\ at\ output}{Signal\ power\ at\ input} = \frac{P_{so}}{P_{si}}$$

$$There\ fore\ F = \frac{P_{si}}{P_{Ni}} \times \frac{P_{No}}{P_{so}} = \frac{P_{si}}{P_{so}} \times \frac{P_{No}}{P_{Ni}} = \frac{P_{N0}}{P_{Ni}} \times \frac{1}{G}$$

$$P_{Ni}\ F\ G = P_{No} = FGP_{Ni}$$

Since most of the noise introduced by an amplifier is thermal in nature thermal noise power at input

Thermal noise power, P_{Ni}= kTB

Using the above formula

Noise power at output P_{No} = F G kTB

Where

- ✓ F is noise factor
- ✓ G – gain at amplifier
- ✓ k - Boltzmann's constant (1.38 x 10^{-23} J/K.)
- ✓ B - Bandwidth in Hz
- ✓ T - absolute temperature in kelvin

Since the noise figure is always greater than unity. For improvement of the noise figure of a receiver the active device used in the receiver should have low noise. The equation can be written as

$$F = \frac{P_{si}}{P_{Ni}} \times \frac{P_{No}}{P_{So}} = \frac{P_{si}}{kTB} \times \frac{P_{No}}{GP_{si}}$$

$$F = \frac{P_{No}}{k\ TBGP_{si}}$$

Then the total noise power at the input is $\frac{P_{No}}{G}$

$$i.e., Total\ P_{ni} = \frac{P_{No}}{G} = F\ k\ T\ B$$

Noise factor is a measured parameter and will be usually specified for a given amplifier or network and specified in decibels. Then it is called **Noise Figure.**

The source contributes the available power kTB and hence the amplifier contributes noise P_{na} given by

$$P_{na} = Total\ P_{ni} - \ P_{Ni}\ due\ to\ source$$

$$P_{na} = FkTB - kTB = (F-1)k\,TB$$

The fraction of total available noise contributes by the amplifier is

$$= \frac{(F-1)k\,TB}{FkTB} = \frac{(F-1)}{F}$$

XV. Noise Temperature

The available noise power is directly proportional to temperature and it is independent of value of resistance. This power specified in terms of temperatures is called noise temperature

Noise power due to amplifier, having a is factor F, is

$$P_{na} = (F - 1)\, k\, TB$$

If T_e represents equivalent noise temperature representing noise power, then

$$P_{na} = k\, T_e B$$

$$(F - 1)\, k\, TB = k\, T_e B$$

$$(F - 1)\, T = T_e$$

$$T_e = (F - 1)\, T$$

This indicates that the relation between T_e and F. The noise figure F is measured under matched input conditions and with the noise source at temperature T.

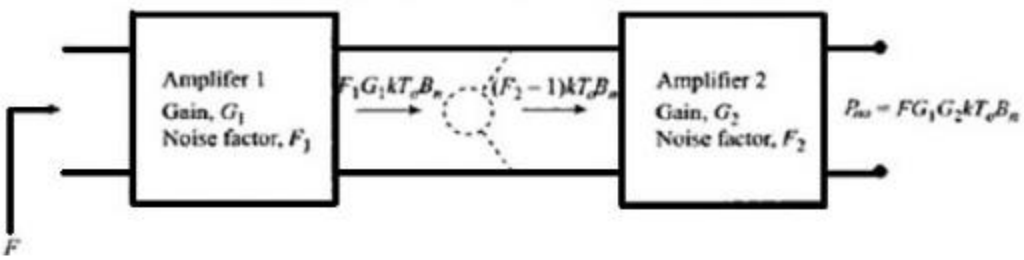
The temperature T is taken as 'room temperature" for convenient. The unit is degree kelvin K.

Noise temperature is a better measure for low noise devices such as low noise amplifier used in satellite receiving system, while noise factor is a better measure for the main receiving system.

XVI. Noise in cascaded systems

Consider two amplifier connected in cascade, the power gain G_1. and its noise factor is F_1; the other amplifier has the power gain G_2 and its noise factor is F_2.

Assume, the devices are matched and the noise figure F_2 of the second network is defined assuming an input noise power N_1



Ainput of the first amplifier, we have a noise power N_1 contributed by the source, plus an equivalent noise power $(F1-1)N_1$ contributed by the network itself. Therefore the output noise power from the first network $F_1 N_1 G_1$.

Added to this noise power at the input of the second amplifier we have the equivalent extra power $(F_2 - 1) N_1$ contributed by the second amplifier network itself. Therefore, the output noise power from this second amplifier is $F_1 G_1 N_1 G_2 + (F_2 -1) N_1 G_2$.

Consider, the **noise figure F** as the ratio of the actual output noise power to the output noise power assuming the amplifier network to be noiseless

Therefore, Express the overall noise figure of the cascade connection if figure.

$$F = \frac{\mathrm{F_1G_1N_1G_2} + (\mathrm{F_2} - 1)\mathrm{N_1G_2}}{\mathrm{N_1G_1G_2}}$$

$$F = \frac{\mathrm{N_1G_2}[\mathrm{F_1G_1} + (\mathrm{F_2} - 1)]}{\mathrm{N_1G_1G_2}}$$

$$F = \frac{[\mathrm{F_1G_1} + (\mathrm{F_2} - 1)]}{\mathrm{G_1}} = \mathrm{F_1} + \frac{(\mathrm{F_2} - 1)}{\mathrm{G_1}}$$

$$F = \mathrm{F_1} + \frac{(\mathrm{F_2} - 1)}{\mathrm{G_1}}$$

Hence, this equation can be generalized for more number of stages, say N and F is given as

$$F = \mathrm{F_1} + \frac{(\mathrm{F_2} - 1)}{\mathrm{G_1}} + \frac{(\mathrm{F_3} - 1)}{\mathrm{G_1G_2}} + \frac{(\mathrm{F_4} - 1)}{\mathrm{G_1G_2G_3}} + \cdots \ldots \quad I$$

Where F1, F2, F3 ……. are individual noise figure

G1, G2, G3 …... are the available power gains

If the first stage of the cascade connection has a high gain, the overall noise figure. F is denominated by the noise figure of the first stage.

We may express the overall equivalent noise temperature of the cascade connections of any number of noisy two port networks as follows:

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2} + \frac{T_4}{G_1G_2G_3} \ldots \ldots \ldots \ldots \qquad II$$

Where

T_1, T_2, T_3 – are the equivalent noise temperature of the individual stages.

G_1, G_2, G_3 – are the available power gains

The equations I and II are known as **Friis formula**. If the gain $G_1$of the first stage is high, the equivalent noise temperature $T_e$is dominated by that of the first stage. Hence we have to be taken case of the first stage noise is as minimum as possible.

Signal to Noise Ratio

Signal-to-Noise Ratio (SNR) is the ratio of the signal power to noise power. The higher the value of SNR, the greater will be the quality of the received output.

Signal-to-Noise Ratio at different points can be calculated using the following formulas.

$$\textbf{Input SNR} = (SNR)_I = \frac{\text{Average power of modulating signal}}{\text{Average power of noise at input}}$$

$$\textbf{Output SNR} = (SNR)_O = \frac{\text{Average power of demodulated signal}}{\text{Average power of noise at output}}$$

$$\textbf{Channel SNR} = (SNR)_C = \frac{\text{Average power of modulated signal}}{\text{Average power of noise in message bandwidth}}$$

Figure of Merit

The ratio of output SNR and input SNR can be termed as **Figure of Merit**. It is denoted by **F**. It describes the performance of a device.

$$F = \frac{(SNR)_O}{(SNR)_I}$$

Figure of merit of a receiver is

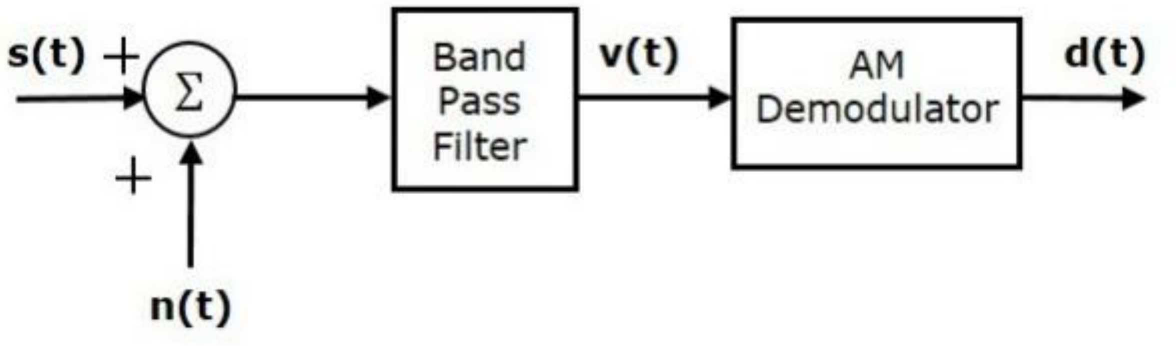
$$F = \frac{(SNR)_O}{(SNR)_C}$$

It is so because for a receiver, the channel is the input.

XVII. Noise in AM Receivers

SNR Calculations in AM System (DSB FC)

Consider the following receiver model of AM system to analyze noise.



We know that the Amplitude Modulated (AM) wave is

$$s(t) = A_c [1 + k_a m(t)] \cos(2\pi f_c t)$$

$$\Rightarrow s(t) = A_c \cos(2\pi f_c t) + A_c k_a m(t) \cos(2\pi f_c t)$$

Average power of AM wave is

$$P_s = \left(\frac{A_c}{\sqrt{2}}\right)^2 + \left(\frac{A_c k_a m(t)}{\sqrt{2}}\right)^2 = \frac{A_c{}^2}{2} + \frac{A_c{}^2 k_a{}^2 P}{2}$$

$$\Rightarrow P_s = \frac{A_c{}^2 \left(1 + k_a{}^2 P\right)}{2}$$

Average power of noise in the message Bandwidth is

$$P_{nc} = W N_0$$

Substitute the values in channel SNR formula

$$(SNR)_{C,AM} = \frac{Average\ Power\ of\ AM\ Wave}{Average\ Power\ of\ noise\ in\ message\ bandwidth}$$

$$\Rightarrow (SNR)_{C,AM} = \frac{A_c{}^2 \left(1 + k_a{}^2\right) P}{2WN_0}$$

Where,

- **P** is the power of the message signal= $A_m{}^2$ **/ 2**
- **W** is the message bandwidth

Assume the band pass noise is mixed with AM wave in the channel as shown in the above figure. This combination is applied at the input of AM demodulator. Hence, the input of AM demodulator is.

$$v(t) = s(t) + n(t)$$

$$\Rightarrow v(t) = A_c \left[1 + k_a m(t)\right] \cos(2\pi f_c t) +$$
$$\left[n_1(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t)\right]$$

$$\Rightarrow v(t) = \left[A_c + A_c k_a m(t) + n_1(t)\right] \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t)$$

Where $n_I(t)$ and $n_Q(t)$ are in phase and quadrature phase components of noise.

The output of AM demodulator is nothing but the envelope of the above signal

$$d(t) = \sqrt{\left[A_c + A_c K_a m(t) + n_I(t)\right]^2 + \left(n_Q(t)\right)^2}$$

$$\Rightarrow d(t) \approx A_c + A_c k_a m(t) + n_1(t)$$

Average power of the demodulated signal is

$$P_m = \left(\frac{A_c k_a m(t)}{\sqrt{2}}\right)^2 = \frac{A_c{}^2 k_a{}^2 P}{2}$$

Average power of noise at the output is

$$P_n o = WN_0$$

Substitute, these values in **output SNR** formula.

$$(SNR)_{O,AM} = \frac{Average\ Power\ of\ demodulated\ signal}{Average\ Power\ of\ noise\ at\ Output}$$

$$\Rightarrow (SNR)_{O,AM} = \frac{A_c{}^2 k_a{}^2 P}{2WN_0}$$

Substitute, the values in **Figure of merit** of AM receiver formula.

$$F = \frac{(SNR)_{O,AM}}{(SNR)_{C,AM}}$$

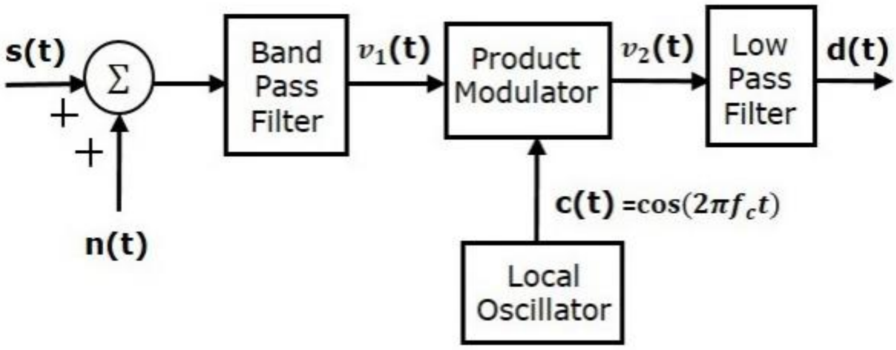
$$\Rightarrow F = \left(\frac{A_c^2 k_a^2 P}{2WN_0}\right) / \left(\frac{A_c{}^2 \left(1 + k_a{}^2\right) P}{2WN_0}\right)$$

$$\Rightarrow F = \frac{K_a{}^2 P}{1 + K_a{}^2 P}$$

Therefore, the Figure of merit of AM receiver is less than one.

SNR Calculations in DSBSC System

Consider the following receiver model of DSBSC system to analyze noise.



We know that the DSBSC modulated wave is

$$s\left(t\right) = A_c m\left(t\right)\cos(2\pi f_c t)$$

Average power of DSBSC modulated wave is

$$P_s = \left(\frac{A_c m\left(t\right)}{\sqrt{2}}\right)^2 = \frac{A_c{}^2 P}{2}$$

Average power of noise in the message bandwidth is

$$P_{nc} = WN_0$$

Substitute, these values in **channel SNR** formula.

$$(SNR)_{C,DSBSC} = \frac{Average\ Power\ of\ DSBSC\ modulated\ wave}{Average\ Power\ of\ noise\ in\ message\ bandwidth}$$

$$\Rightarrow (SNR)_{C,DSBSC} = \frac{{A_c}^2 P}{2WN_0}$$

Assume the band pass noise is mixed with DSBSC modulated wave in the channel as shown in the above figure. This combination is applied as one of the input to the product modulator. Hence, the input of this product modulator is

$$v_1(t) = s(t) + n(t)$$

$$\Rightarrow v_1(t) = A_c m(t)\cos(2\pi f_c t) + [n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)]$$

$$\Rightarrow v_1(t) = [A_c m(t) + n_I(t)]\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)$$

Local oscillator generates the carrier signal $c(t)=\cos(2\pi f_c t)$. This signal is applied as another input to the product modulator. Therefore, the product modulator produces an output, which is the product of v1(t) and c(t).

$$v_2(t) = v_1(t)\, c(t)$$

Substitute, v1(t) and c(t) values in the above equation.

$$\Rightarrow v_2(t) = ([A_c m(t) + n_I(t)]\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t))\cos(2\pi f_c t)$$

$$\Rightarrow v_2(t) = [A_c m(t) + n_I(t)]\cos^2(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)\cos(2\pi f_c t)$$

$$\Rightarrow v_2(t) = [A_c m(t) + n_I(t)]\left(\frac{1+\cos(4\pi f_c t)}{2}\right) - n_Q(t)\frac{\sin(4\pi f_c t)}{2}$$

When the above signal is applied as an input to low pass filter, we will get the output of low pass filter as

$$d(t) = \frac{[A_c m(t) + n_I(t)]}{2}$$

Average power of the demodulated signal is

$$P_m = \left(\frac{A_c m(t)}{2\sqrt{2}}\right)^2 = \frac{{A_c}^2 P}{8}$$

Average power of noise at the output is

$$P_{no} = \frac{WN_0}{4}$$

Substitute, these values in **output SNR** formula.

$$(SNR)_{O,DSBSC} = \frac{Average\ Power\ of\ demodulated\ signal}{Average\ Power\ of\ noise\ at\ Output}$$

$$\Rightarrow (SNR)_{O,DSBSC} = \left(\frac{{A_c}^2 P}{8}\right) / \left(\frac{WN_0}{4}\right) = \frac{{A_c}^2 P}{2WN_0}$$

Substitute, the values in **Figure of merit** of DSBSC receiver formula.

$$F = \frac{(SNR)_{O,DSBSC}}{(SNR)_{C,DSBSC}}$$

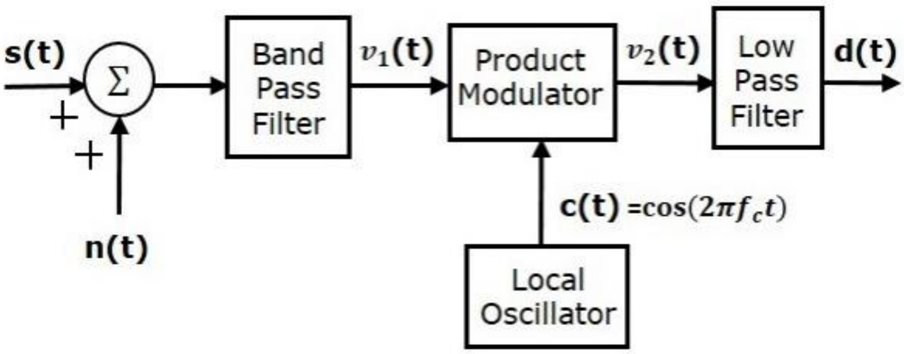
$$\Rightarrow F = \left(\frac{{A_c}^2 P}{2WN_0}\right) / \left(\frac{{A_c}^2 P}{2WN_0}\right)$$

$$\Rightarrow F = 1$$

Therefore, the Figure of merit of DSBSC receiver is 1.

SNR Calculations in SSBSC System

Consider the following receiver model of SSBSC system to analyze noise.



We know that the SSBSC modulated wave having lower sideband is

$$s(t) = \frac{A_m A_c}{2} \cos[2\pi (f_c - f_m) t]$$

Average power of SSBSC modulated wave is

$$P_s = \left(\frac{A_m A_c}{2\sqrt{2}}\right)^2 = \frac{A_m{}^2 A_c{}^2}{8}$$

Average power of noise in the message bandwidth is

$$P_{nc} = W N_0$$

Substitute, these values in **channel SNR** formula.

$$(SNR)_{C,SSBSC} = \frac{Average\ Power\ of\ SSBSC\ modulated\ wave}{Average\ Power\ of\ noise\ in\ message\ bandwidth}$$

$$\Rightarrow (SNR)_{C,SSBSC} = \frac{A_m{}^2 A_c{}^2}{8 W N_0}$$

Assume the band pass noise is mixed with SSBSC modulated wave in the channel as shown in the above figure. This combination is applied as one of the input to the product modulator. Hence, the input of this product modulator is

$$v_1(t) = s(t) + n(t)$$

$$v_1(t) = \frac{A_m A_c}{2} \cos[2\pi(f_c - f_m)t] + n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)$$

The local oscillator generates the carrier signal c(t)=cos(2πf$_c$t). This signal is applied as another input to the product modulator. Therefore, the product modulator produces an output, which is the product of v1(t) and c(t).

$$v_2(t) = v_1(t)\, c(t)$$

Substitute, v_1(t) and c(t) values in the above equation.

$$\Rightarrow v_2(t) = \left(\frac{A_m A_c}{2} \cos[2\pi(f_c - f_m)t] + n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)\right)\cos(2\pi f_c t)$$

$$\Rightarrow v_2(t) = \frac{A_m A_c}{2} \cos[2\pi(f_c - f_m)t]\cos(2\pi f_c t) + n_I(t)\cos^2(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)\cos(2\pi f_c t)$$

$$\Rightarrow v_2(t) = \frac{A_m A_c}{4} \{\cos[2\pi(2f_c - f_m)t] + \cos(2\pi f_m t)\} + n_I(t)\left(\frac{1+\cos(4\pi f_c t)}{2}\right) - n_Q(t)\frac{\sin(4\pi f_c t)}{2}$$

When the above signal is applied as an input to low pass filter, we will get the output of low pass filter as

$$d(t) = \frac{A_m A_c}{2} \cos(2\pi f_m t) + \frac{n_I(t)}{2}$$

Average power of the demodulated signal is

$$P_{no} = \frac{WN_0}{4}$$

Substitute, these values in **output SNR** formula

$$(SNR)_{O,SSBSC} = \frac{Average\ Power\ of\ demodulated\ signal}{Average\ Power\ of\ noise\ at\ output}$$

$$\Rightarrow (SNR)_{O,SSBSC} = \left(\frac{A_m{}^2 A_c{}^2}{32}\right) / \left(\frac{WN_0}{4}\right) = \frac{A_m{}^2 A_c{}^2}{8WN_0}$$

Substitute, the values in **Figure of merit** of SSBSC receiver formula

$$F = \frac{(SNR)_{O,SSBSC}}{(SNR)_{C,SSBSC}}$$

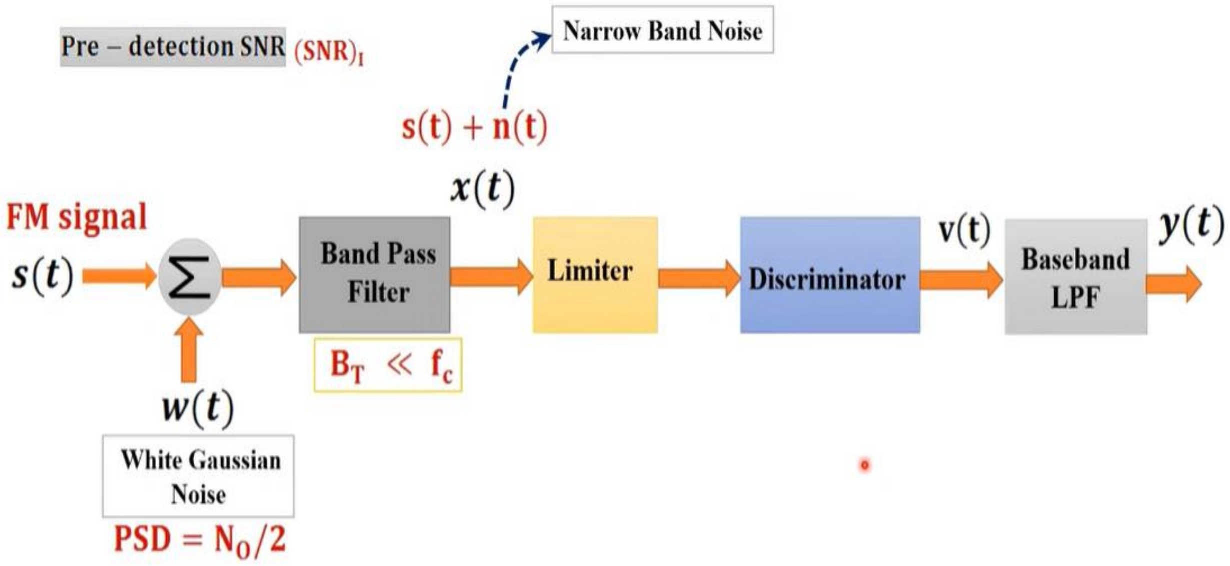
$$F = \left(\frac{A_m{}^2 A_c{}^2}{8WN_0}\right) / \left(\frac{A_m{}^2 A_c{}^2}{8WN_0}\right)$$

$$F = 1$$

Therefore, the Figure of merit of SSBSC receiver is 1.

XVIII. Noise in FM System

In an FM system, the message information is transmitted by variations of the instantaneous frequency of a sinusoidal carrier wave, and its amplitude is maintained constant. Therefore any variation on its amplitude is removed by the limiter. The receiver model is given by



The Noise w(t) is modeled as white Gaussian noise of zero mean and power spectral density No/2. The received FM signal s(t) has a carrier frequency f_c and transmission bandwidth B_T, such that only a negligible amount of power lies outside the frequency band $f_c \pm B_T/2$ for positive frequencies. The band-pass filter has a mid-band frequency f_c and bandwidth B_T and therefore passes the FM signal essentially without distortion. Ordinary, B_T is small compared with the mid-band frequency f_c so that we may use the narrowband representation for n(t), the filtered version of receiver noise w(t), in terms of its in-phase and quadrature components.

The limiter is used to remove amplitude variations by clipping the modulated wave at the filter output almost to the zero axis. The resulting rectangular wave is rounded off by another bandpass filter that is an integral part of the limiter, thereby suppressing harmonics of the carrier frequency. The filter output is again sinusoidal, with an amplitude that is practically independent of the carrier amplitude at the receiver input.

The discriminator consists of two components:

A **slope network** or differentiator with a purely imaginary transfer function that varies linearly with frequency. It produces a hybrid-modulated wave in which both amplitude and frequency vary in accordance with the message signal.

An **envelope detector** that recovers the amplitude variation and thus reproduces the message signal.

The slope network and envelope detector are usually implemented as integral parts of a single physical unit.

The post-detection filter, labeled "baseband low-pass filter," has a bandwidth that is just large enough to accommodate the highest frequency component of the message signal. This filter removes the out-of-band components of the noise at the discriminator output and thereby keeps the effect of the output noise to a minimum.

Pre- Determined SNR $(SNR)_I$

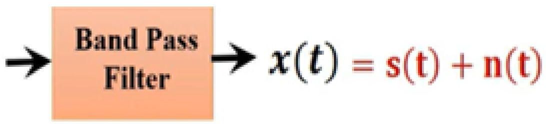
We know that FM modulated Signal general equation is

$$S(t) = A_c cos\left[2\pi f_c t + 2\pi K_f \int_0^t m(\tau)d\tau\right]$$

$$(\mathrm{SNR})_\mathrm{I} = \frac{\text{Carrer Power}}{\text{Noise Power}} = \frac{\mathrm{A_c^2}/2}{\mathrm{N_0 B_T}}$$

$$(\mathrm{SNR})_\mathrm{I} = \frac{\mathrm{A_c^2}}{\mathrm{2N_0 B_T}}$$

Post- Detection SNR $(SNR)_O$



n(t) = Narrow band Noise $n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t)$

Polar representation of n(t) is $n(t) = r(t)cos[\cos(2\pi f_c t) + \varphi_n(t)]$

Where r(t) – Envelope of noise and $\varphi_n(t)$ - Phase of Noise

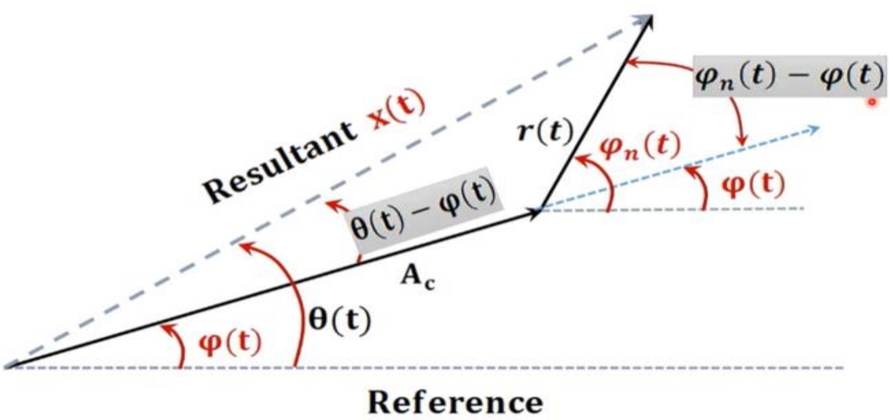
We can write this as follows

$$r(t) = \sqrt{n_I^2(t) + n_Q^2(t)}$$

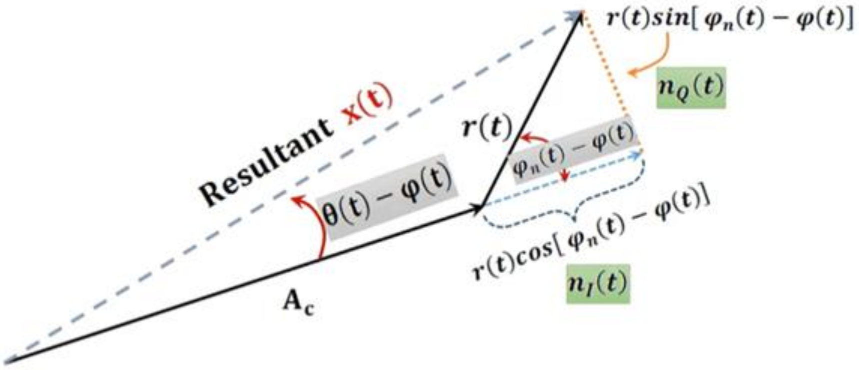
$$\varphi_n(t) = \tan^{-1}\left[\frac{n_Q(t)}{n_I(t)}\right]$$

This φ is Uniformly distributed between (0 to 2π) radian. Phasor diagram of output of Band Pass Filter is ie x(t) is

$$x(t) = \mathrm{A_c cos[2\pi f_c t + \varphi(t)]} + r(t)cos[2\pi f_c t + \varphi_n(t)]$$



Now using this phase diagram, we can represent inphase and quadrature phase noise component. For this we have to draw the perpendicular line from the tip of r(t) on to the extended A_C tip



$$[\theta(t)-\varphi(t)] = tan^{-1}\frac{r(t)sin[\varphi_n(t)-\varphi(t)]}{A_c + r(t)cos[\varphi_n(t)-\varphi(t)]}$$

Now for the sake of simplicity and calculations

Assume : A_C ≈ Very large (Amplitude of Unmodulated career)

Such that $\frac{\text{Carrer}}{\text{Noise}} \gg 1$

To find noise performance accurately we need to find Time dependent angle first

$$[\theta(t) - \varphi(t)] \cong \tan^{-1}\left[\frac{r(t)sin[\varphi_n(t)-\varphi(t)\,]}{A_c}\right]$$

Since Ac is very large the frection in above equation is very small <<1

Under this condition $\tan^{-1}$ can be written as

$$\tan^{-1}\left[\frac{r(t)sin[\varphi_n(t)-\varphi(t)\,]}{A_c}\right] \cong \left[\frac{r(t)sin[\varphi_n(t)-\varphi(t)\,]}{A_c}\right]$$

Now θ(t) can be written as

$$\theta(t) \cong \varphi(t) + \frac{r(t)sin[\varphi_n(t)-\varphi(t)\,]}{A_c}$$

$$\cong \varphi(t) + \frac{n_Q(t)}{A_c}$$

Where $n_Q(t) = r(t)sin[\varphi_n(t)-\varphi(t)\,]$

We know that $\varphi(t) = 2\pi K_f \int_0^t m(\tau)d\tau$

Now filtered signal x in vector form with phase angle φ(t) passed through discriminator. Here for our convenience we assume the ideal discriminator which function as equal to differentator.

The output of the discriminator is

$$v(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt}\ with\ the\ intervel\ (0\ to\ 2\pi)$$

Which is distributed over the interval 0 to 2π

$$v(t) = \frac{1}{2\pi}\frac{d}{dt}\left\{2\pi K_f\int_0^t m(\tau)d\tau + \frac{r(t)sin[\varphi_n(t) - \varphi(t)\,]}{A_c}\right\}$$

Here the first term both consist of integration and differentiation also 2π will get cancel so only K_f. The second term is the derivative part of quadrature noise.

$$v(t) = K_f m(t) + n_d(t)$$

Therefore the output of discriminator becomes

<u>Derivative Noise $n_d(t)$</u>

$$n_d(t) = \frac{1}{2\pi}\frac{d}{dt}\left[n_Q(t)\right]$$

$$n_d(t) = \frac{1}{2\pi}\frac{d}{dt}\left[\frac{r(t)sin[\varphi_n(t) - \varphi(t)\,]}{A_c}\right] = \frac{1}{2\pi A_c}\frac{d}{dt}[r(t)sin[\varphi_n(t) - \varphi(t)\,]$$

$$v(t) = K_f m(t) + \boxed{n_d(t)} \dashrightarrow Additive\ noise$$

The derivative noise is the additive noise and can be determined using the quadrature component of additive noise

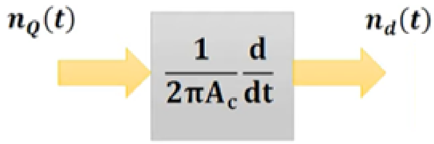
Now with the help of discriminator output we can find post detection SNR i.e output side SNR

<u>Post- Detection SNR $(SNR)_O$</u>

$$v(t) = K_f m(t) + n_d(t)$$

Where $K_f m(t) = Average\ Signal\ power = {K_f}^2 P$

$$n_d(t) = \frac{1}{2\pi}\frac{d}{dt}\left[n_Q(t)\right] \;--> Average\ Noise\ pow$$

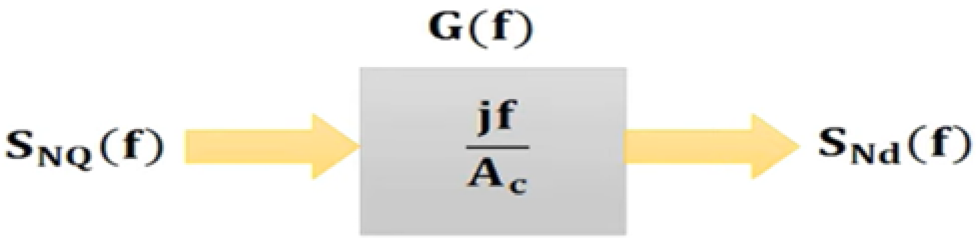


Phase Discriminator

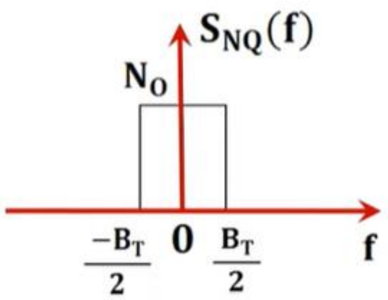
$$\frac{1}{2\pi A_c}\frac{d}{dt} \overset{F.T}{\longleftrightarrow} \frac{j2\pi f}{2\pi A_c} = \frac{jf}{A_c} \;\; ----\rightarrow G(f) \; is\; the\; transfer\; function$$

$$PSD\; of\; n_Q(t) \;\equiv\; S_{NQ}\,(f)$$

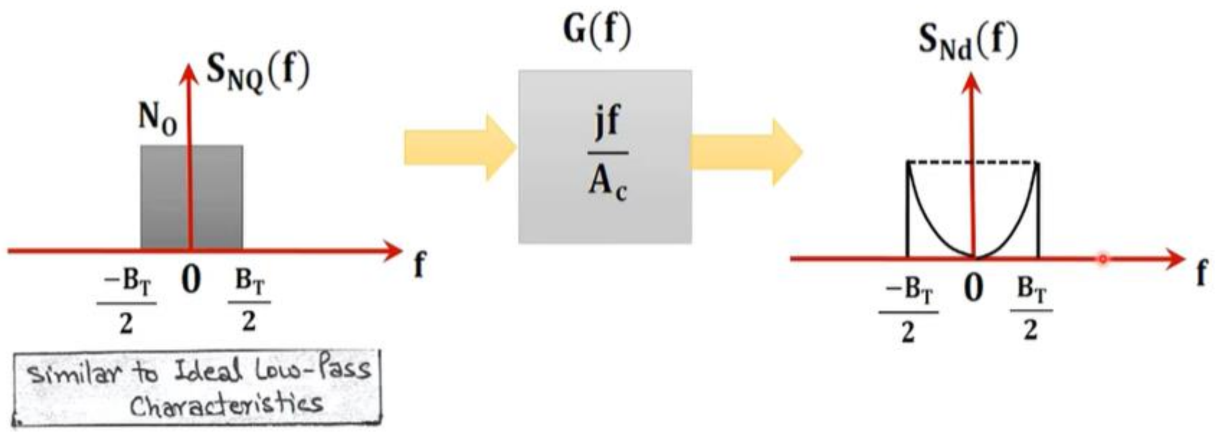
$$PSD\; of\; n_d(t) \;\equiv\; S_{Nd}\,(f)$$



$$S_{Nd}\,(f) \equiv\; |G(f)|^2\, S_{NQ}\,(f) = \frac{f^2}{A_c^2}$$



$$S_{Nd}(f) = \begin{cases} \dfrac{f^2}{A_c^2}\, N_O & ;\; |f| \leq \dfrac{B_T}{2} \\ 0 & ;\; \text{Otherwise} \end{cases}$$



$$S_{No}(f) = \begin{cases} \frac{f^2}{A_c^2} N_0 & ; |f| \le w \\ 0 & ; \text{Otherwise} \end{cases}$$

Average Power of output Noise

$$\equiv \int_{-w}^{w} S_{No}(f) \equiv \frac{N_o}{A_c^2} \int_{-w}^{w} f^2 df \equiv \frac{N_o}{A_c^2} \left[\frac{f^3}{3}\right]_{-w}^{w}$$

$$S_{No}(f)_{avg} = \frac{2N_o W^3}{3A_c^2}$$

Therefore **Post- Detection SNR (SNR)$_O$**

$$v(t) = K_f m(t) + n_d(t)$$

Where $\quad K_f m(t) = Average\ Signal\ power = {K_f}^2 P$

$$n_d(t) = \frac{2N_o W^3}{3A_c^2}$$

$$(SNR)_o^{FM} = \frac{{K_f}^2 P}{\frac{2N_o W^3}{3A_c^2}} = \frac{3A_c^2 {K_f}^2 P}{2N_o W^3}$$

$\rightarrow$ $Non\ linear\ dependence\ on\ both\ freq.sensitivity\ (K_f)\&\ message\ Bandwidth\ (w)$

Figure of Merit

Input signal for FM receiver is

$$S(t) = A_c cos\left[2\pi f_c t + 2\pi K_f \int_0^t m(\tau)d\tau\right]$$

$$(SNR)_{ref} = \frac{\mathrm{A_c^2}}{2\mathrm{N_0}W}$$

$$(SNR)_o^{FM} = \frac{3A_c^2 {K_f}^2 P}{2N_o W^3}$$

$$FOM = \frac{(SNR)_o^{FM}}{(SNR)_{ref}} = \frac{\frac{3A_c^2 {K_f}^2 P}{2N_o W^3}}{\frac{\mathrm{A_c^2}}{2\mathrm{N_0}W}} = \frac{3{K_f}^2 P}{W}$$

Deviation Ratio :

$$\text{Deviation Ratio} = \frac{Freq.deviation}{Message\ Bandwidth}$$

$$D = \frac{\Delta f}{W}$$

$$\Delta f = DW$$

Carson's rule

$$B_T = 2(\Delta f + f_m)$$

$$B_T = 2(DW + W)$$

$$B_T = 2(K_f\sqrt{P} + W)$$

$$B_T \cong 2\left(K_f\sqrt{P}\right) for\ wideband\ FM$$

$$\frac{B_T}{2} \cong \left(K_f\sqrt{P}\right)$$

Using above equation we can write

$$FOM = \frac{3{K_f}^2 P}{W} = 3\left(\frac{{K_f}^2 P}{W}\right) \cong \frac{3}{4}\left(\frac{B_T}{w}\right)^2$$

B_T improves noise performance accordance with the square law

Reference Books

1. Dr. Sanjay Sharma, “Analog Communication Systems” 2013, January 2013
2. R.P Singh, S.D. Spare, “Communication Systems: Analog and Digital” ,Tata McGraw-Hill, 3rd Edition, July 2017.
3. Sanjeev Guptha, “Electronic Communication (Analog, Digital and Wireless)”, Khanna publications, January 1997
4. B.P.Lathi, “Modern Digital and Analog Communication Systems”, 3rd Edition, Oxford University Press, 2007.
5. Louis. E. Frenzel, “Communication Electronics Principles and applications”, 3rd Edition, Tata McGraw Hill, 2002.
6. Louis. E. Frenzel, “Principles of Electronic Communication Systems”, 3rd Edition, Tata McGraw Hill, July 2017.
7. Dennis Reddy and John Coolen, "Electronic Communications", 4th Edition, Prentice Hall Publishers, 1995.
8. Electronic Communications Systems by Wayne Tomasi. Pearson education India.
9. Electronic Communication Systems by Roy Blake By Cengage learning.
10. Communication Systems By Simon Haykins By Wiley India.
11. Theory and Problem Of Electronic Communication By Lloyd Temes and Mitchel E.Schulz(Second edition), McGraw Hill Publication.
12. Kennedy, "Electronic Communications Systems", 4th Edition, McGraw-Hill Publishers, 1992.
13. Deshpande, N.D, “Communication Electronics”, Tata McGraw Hill Publishers,1989.
14. R.Gulati, “Monochrome and color television”, 2nd Edition, New age international 2005

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SECA1303 – ANALOG COMMUNICATION SYSTEMS

UNIT 5 - COMMUNICATION SYSTEMS

Telephone Systems - Electronics Telephone and cellular Telephone System - Fax-Television Systems – Scanning - camera tube and Transmitter, Picture tube and Receiver, CCTV and Set top box.

Introduction: Telephone Systems

Telecommunication means —communications at a distance. Tele in Greek means at a distance Electrical communications by wire, radio, or light (fiber optics).

Telephones

The telephone system was designed for full-duplex analog communication of voice signals. Today, this system is still primarily used for voice, but it employs mostly digital techniques, not only in signal transmission but also in control operations. The telephone system permits any telephone to connect with any other telephone in the world.

A telephone line or telephone circuit (or just line or circuit within the industry) is a single-user circuit on a telephone communication system. This is the physical wire or other signaling medium connecting the user's telephone apparatus to the telecommunications network, and usually also implies a single telephone number for billing purposes reserved for that user. Telephone lines are used to deliver landline telephone service and Digital subscriber line (DSL) phone cable service to the premises.

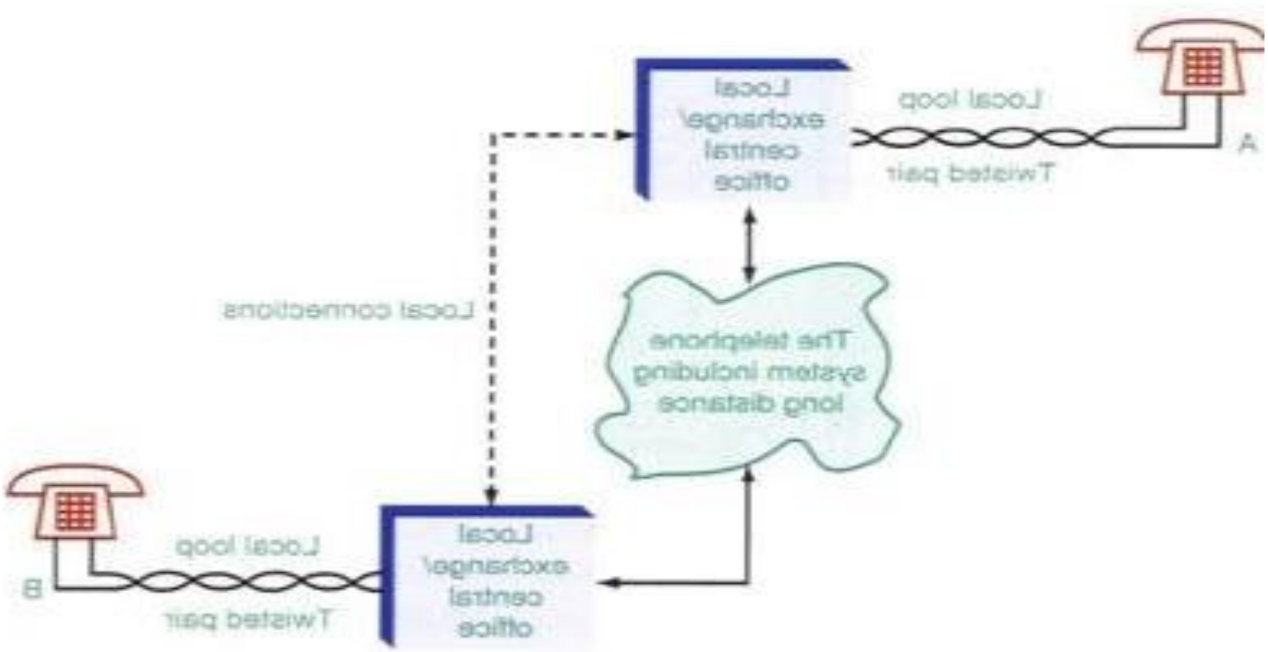
The Local loop:

The Local Loop in a telephone network (sometimes referred to as the "last mile" of the network) is the bit that connects home to local telephone exchange. It refers literally to the copper cables that run from home to the telephone exchange. Standard telephones are connected to the telephone system by way of a two-wire, twisted-pair cable that terminates at the local exchange or central office. As many as 10,000 telephone lines can be connected to a single central office. The two-wire, twisted-pair connection between the telephone and central office is referred to as the local loop or subscriber loop. The circuits in the telephone and at the central office form a complete electric circuit, or loop.

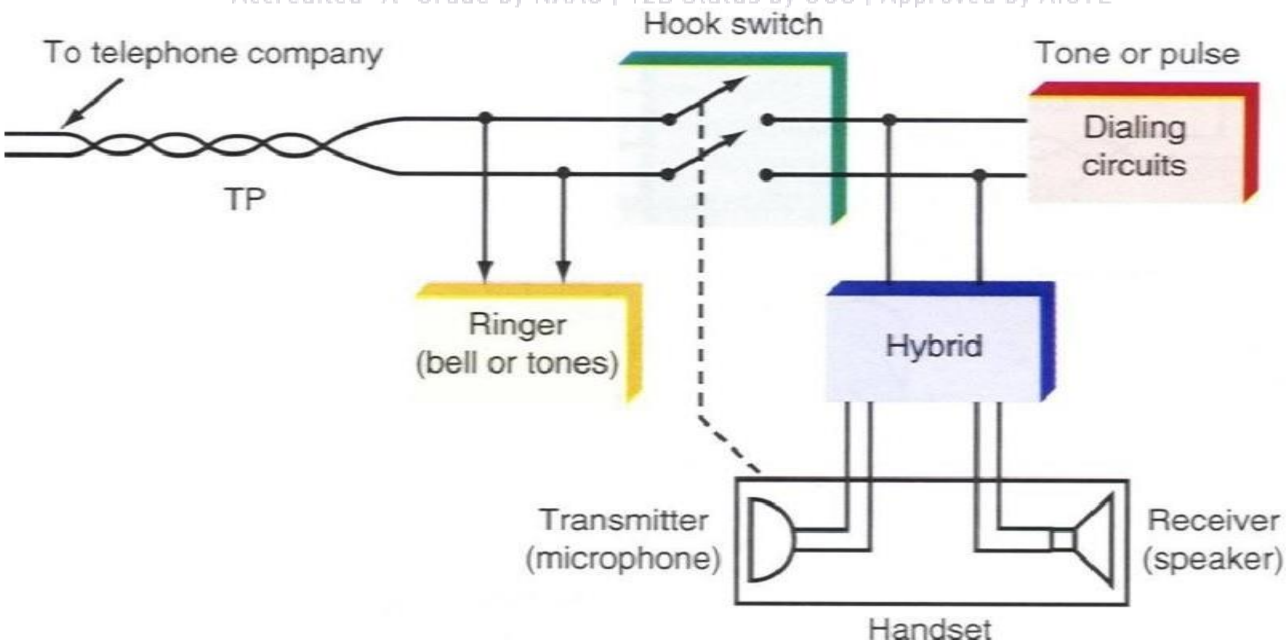
A basic telephone or telephone set (figure shown below) is an analog baseband transceiver. It consists of the following:

The ringer is either a bell or electronic oscillator connected to a speaker**. A switch hook** is a double-pole mechanical switch that is usually controlled by a mechanism actuated by the telephone handset. The **dialing circuits** provide a way for entering the telephone number to be called. Most telephones use the **dual-tone multi frequency (DTMF) system**. The handset contains a **microphone for the transmitter and a speaker or receiver**.
A combination of 350 Hz and 440 Hz sine waves sent to the Telephone from the **central office** (CO) indicating that the network is ready to receive calling instructions. The **hybrid is a special transformer** used to convert signals from the four wires from the transmitter and receiver into a signal suitable for a single two-line pair to the **local loop.**

Basic Telephone System:



Basic telephone set



It also contains a ringer and a dialing mechanism. Overall, the telephone set fulfills the following basic functions.

The receive mode provides:

1. *An incoming signal that rings a bell or produces an audio tone indicating that a call is being received*
2. *A signal to the telephone system indicating that the signal has been answered*
3. Transducers to convert voice to electric signals and electric signals to voice

The transmit mode:

1. Indicates to the telephone system that a call is to be made when the handset is lifted.
2. Indicates that the telephone system is ready to use by generating a signal called the *dial tone*
3. Provides a way of transmitting the telephone number to be called to the telephone system
4. Receives an indication that the call is being made by receiving a ringing tone
5. Provides a means of receiving a special tone indicating that the called line is busy
6. Provides a means of signaling the telephone system that the call is complete

All telephone sets provide these basic functions. Some of the more advanced electronic telephones have other features such as multiple line selection, hold, speaker phone, call waiting, and caller ID.
Figure shows a basic block diagram of a telephone set. The function of each block is described below. Detailed circuits for each of the blocks and their operation are described later when the standard and electronic telephones are discussed in detail.

Ringer:
The *ringer* is either a bell or an electronic oscillator connected to a speaker. It is continuously connected to the twisted pair of the local loop back to the central office. When an incoming call is received, a signal from the central office causes the bell or ringer to produce a tone.
Switch Hook. A *switch hook* is a double-pole mechanical switch that is usually controlled by a mechanism actuated by the telephone handset. When the handset is on the hook,‖ the hook switch is open, thereby isolating all the telephone circuitry from the central office local loop. When a call is to be made or to be received, the handset is taken off the hook. This closes the switch and connects the telephone circuitry to the local loop. The direct current from the central office is then connected to the telephone, closing its circuits to operate.

Dialing Circuits. The *dialing circuits* provide a way for entering the telephone number to be called. In older telephones, a pulse dialing system was used. A rotary dial connected to a switch produced a number of on/off pulses corresponding to the digit dialed. These on/off pulses formed a simple binary code for signaling the central office.

In most modern telephones, a tone dialing system is used. Known as the *Dual-Tone Multi Frequency (DTMF) system,* this dialing method uses a number of pushbuttons that generate pairs of audio tones that indicate the digits called.

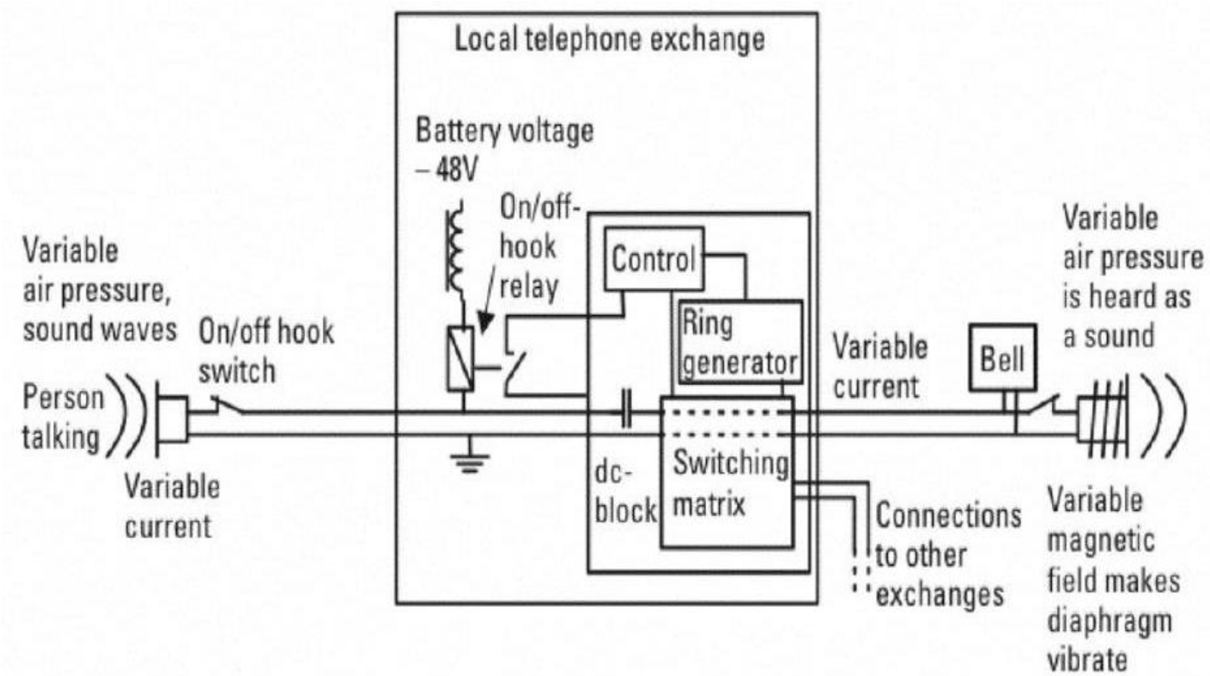
Whether pulse dialing or tone dialing is used, circuits in the central office recognize the signals and make the proper connections to the dialed telephone.

Handset. This unit contains a microphone for the transmitter and a speaker or receiver.

When speak into the transmitter, it generates an electric signal representing the voice. When a received electric voice signal occurs on the line, the receiver translates it to sound waves. The transmitter and receiver are independent units, and each has two wires connecting to the telephone circuit. Both connect to a special device known as the hybrid.

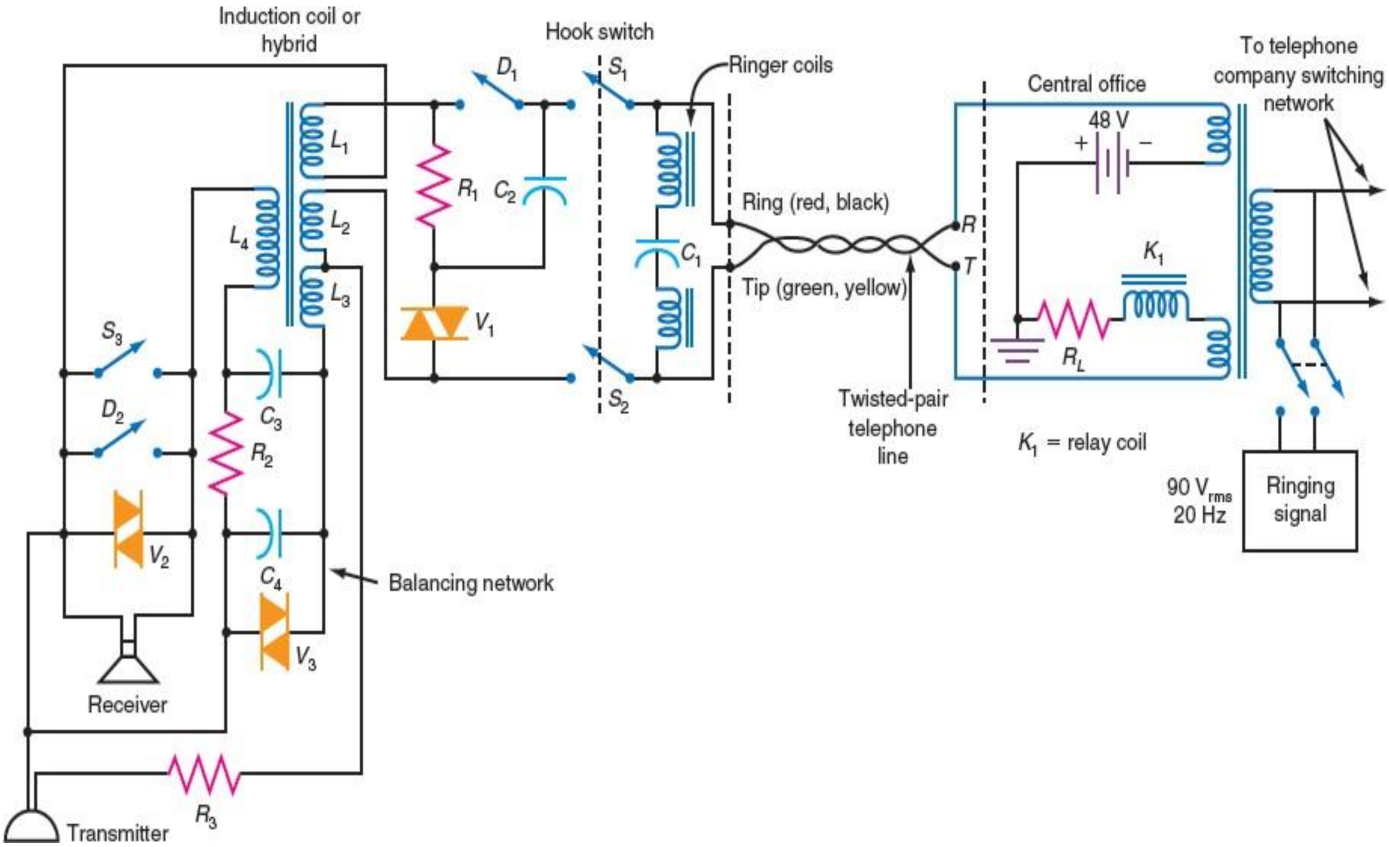
Hybrid. The *hybrid circuit* is a special transformer used to convert signals from the four wires from the transmitter and receiver to a signal suitable for a single two-line pair to the local loop. The hybrid permits *full duplex,* i.e., simultaneous send and receive, analog communication on the two-wire line. The hybrid also provides a side tone from the transmitter to the receiver so that the speaker can hear her or his voice in the receiver. This feedback permits automatic voice-level adjustment.

Conventional Telephone system:
There are some components associated with telephone systems that deserve special consideration.



Standard Telephone and Local loop:

Figure shows that the schematic diagram of a conventional telephone and the local loop connections back to the central office.



The circuitry at the central office is discussed in greater detail later. For now,note that the central office applies a dc voltage over the twisted-pair line to the telephone. This dc voltage is approximately 248 V with respect to ground in the open-circuit condition. When a subscriber picks up the telephone, the switch hook closes, connecting the circuitry to the telephone line. The load represented by the telephone circuitry causes current to low in the local loop and the voltage inside the telephone to drop to approximately 5 to 6 V.

The amount of current l owing in the local loop depends upon a number of factors. The dc voltage supplied by the central office may not be exactly 248 V. It can, in fact, vary many volts above or below the 48-V normal value. Figure shows, the central office also inserts some resistance *RL* to limit the total current low if a short circuit occurs on the line. This resistance can range from about 350 to 800 V. In Fig., the total resistance is approximately 400 V.

The resistance of the telephone itself also varies over a relatively wide range. It can be as low as 100 V and as high as 400 V, depending upon the circuitry. The resistance varies because of the resistance of the transmitter element and because of the variable resistors called *varistors* used in the circuit to provide automatic adjustment of line level.

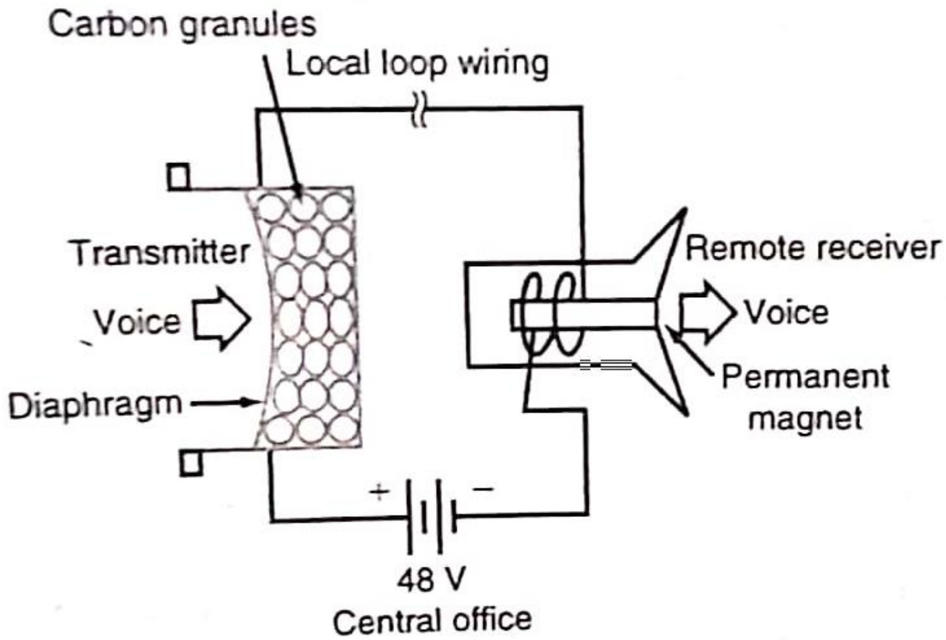
The local loop resistance depends considerably on the length of the twisted pair between the telephone and the central office. Although the resistance of copper wire in the twisted pair is relatively low, the length of the wire between the telephone and the central office can be many miles long. Thus the resistance of the local loop can be anywhere from 1000 to 1800 V, depending upon the distance. The local loop length can vary from a few thousand feet up to about 18,000 ft.

Finally, the frequency response of the local loop is approximately 300 to 3400 Hz. This is sufficient to pass voice frequencies that produce full intelligibility. An unloaded twisted pair has an upper cutoff frequency of about 4000 Hz. But this cutoff varies considerably depending upon the overall length of the cable. When long runs of cable are used, special loading coils are inserted into the line to compensate for excessive roll-off at the higher frequencies.

The two wires used to connect telephones are labeled *tip* and *ring.* These designations refer to the plug used to connect telephones to one another at the central office. At one time, large groups of telephone operators at the central office used plugs and jacks at a switchboard to connect one telephone to another manually.
The tip wire is green and is usually connected to ground; the ring wire is red. Many telephone cables into a home or an office also contain a second twisted pair if a separate telephone line is to be installed. These wires are usually color-coded black and yellow. Black and yellow correspond to ring and tip, respectively, where yellow is ground. Other color combinations are used in telephone wiring.

Ringer. circuitry connected directly to the tip and ring local loop wires is the ringer. The ringer in most older telephones is an electromechanical bell. A pair of electromagnetic coils is used to operate a small hammer that alternately strikes two small metallic bells. When an incoming call is received, a voltage from the central office operates the electromagnetic coils, which in turn operate the hammer to ring the bells. The bells make the familiar tone produced by most standard telephones. In Figure the ringing coils are connected in series with a capacitor *C*1. This allows the ac ringing voltage to be applied to the coils but blocks the 48 V of direct current, thus minimizing the current drain on the 48 V of power supplied at the central office. The ringing voltage supplied by the central office is a sine wave of approximately 90Vrms at a frequency of about 20 Hz. These are the nominal values, because the actual ringing voltage can vary from approximately 80 to 100 Vrms with a frequency somewhere in the 15- to 30-Hz range. This ac signal is supplied by a generator at the central office.
The ringing voltage is applied in series with the 248-V dc signal from the central office power supply. The ringing signal is connected to the local loop line by way of a transformer *T*1. The transformer couples the ringing signal into its secondary winding where it appears in series with the 48-V dc supply voltage.
The standard ringing sequence is shown in Figure. In U.S. telephones, the ringing voltage occurs for 1 s followed by a 3-s interval. Telephones in other parts of the world use different ringing sequences. For example, in the United Kingdom, the standard ring sequence is a higher-frequency tone occurring more frequently, and it consists of two ringing pulses 400 ms long, separated by 200 ms. This is followed by a 2-s interval of quiet before the tone sequence repeats.

transmitter and receiver in a telephone.



Transmitter:

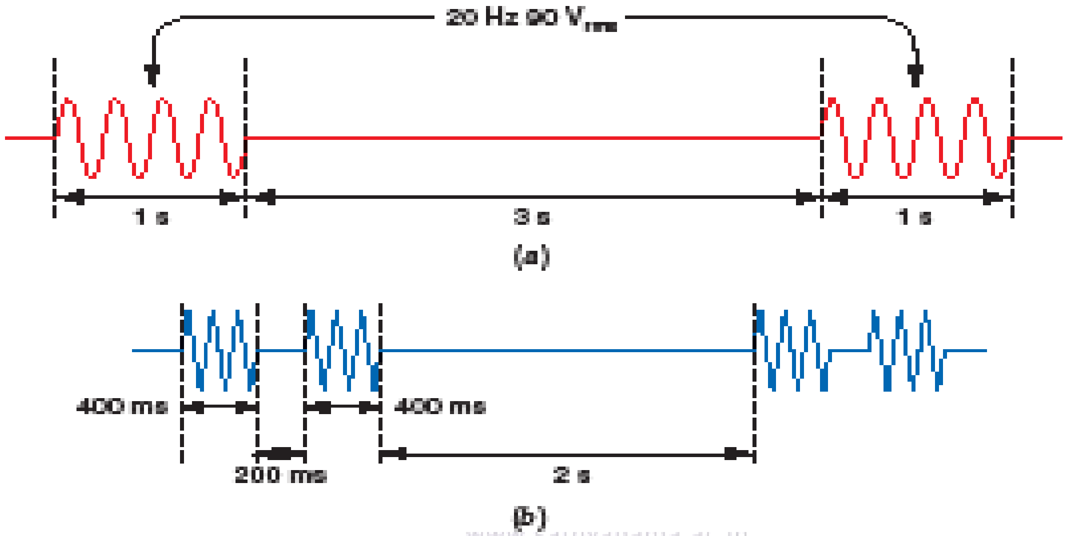
The transmitter is the microphone into which speak during a telephone call. In a standard telephone, this microphone uses a carbon element that effectively translates acoustical vibrations into resistance changes. The resistance changes, in turn, produce current variations in the local loop representing the speaker's voice. A dc voltage must be applied to the transmitter so that current flows through it during operation. The 48 V from the central office is used in this case to operate the transmitter. The resulting ac voice signal produced on the telephone line is approximately 1 to 2 Vrms.

Receiver. The receiver, or earpiece, is basically a small permanent-magnet speaker. A thin metallic diaphragm is physically attached to a coil that rests inside a permanent magnet. Whenever a voice signal comes down a telephone line, it develops a current in the receiver coil. The coil produces a magnetic field that interacts with the permanent-magnet field. The result is vibration of the diaphragm in the receiver, which converts the electric signal to the acoustic energy that supplies the voice to the ear. As it comes in over the local loop lines, the voice signal has an amplitude of approximately 0.5 to 1 Vrms.

Hybrid. The hybrid is a transformer like device that is used to simultaneously transmit and receive on a single pair of wires. The hybrid, which is also sometimes referred to as an *induction coil,* is really several transformers combined into a single unit. The windings on the transformers are connected in such a way that signals produced by the transmitter are put on the two-wire local loop but do not occur in the receiver. In the same way, the transformer windings permit a signal to be sent to the receiver, but the resulting voltage is not applied to the transmitter.

In practice, the hybrid windings are set up so that a small amount of the voice signal produced by the transmitter does occur in the receiver. This provides feedback to the speaker so that she or he may speak with normal loudness. The feedback from the transmitter to the receiver is referred to as the *side tone.* If the side tone were not provided, there would be no signal in the receiver and the person speaking would have the sensation that the telephone line was dead. By hearing his or her own voice in the receiver at a moderate level, the caller can speak at a normal level. Without the side tone, the speaker tends to speak more loudly, which is unnecessary.

Telephone ringing sequence. (*a*) United States and Europe. (*b*) United Kingdom.



Automatic Voice Level Adjustment: Because of the wide variation in the different loop lengths of the two telephones connected to each other, the circuit resistances will vary considerably, thereby causing a wide variation in the transmitted and received voice signal levels. All telephones contain some type of component or circuit that provides *automatic voice level adjustment* so that the signal levels are approximately the same regardless of the loop lengths. In the standard telephone, this automatic loop length adjustment is handled by components called *varistors.* These are labeled *V*1, *V*2, and *V*3 in Figure.

A varistor is a nonlinear resistance element whose resistance changes depending upon the amount of current passing through it. When the current passing through the varistor increases, its resistance decreases. A decrease in current causes the resistance to increase.

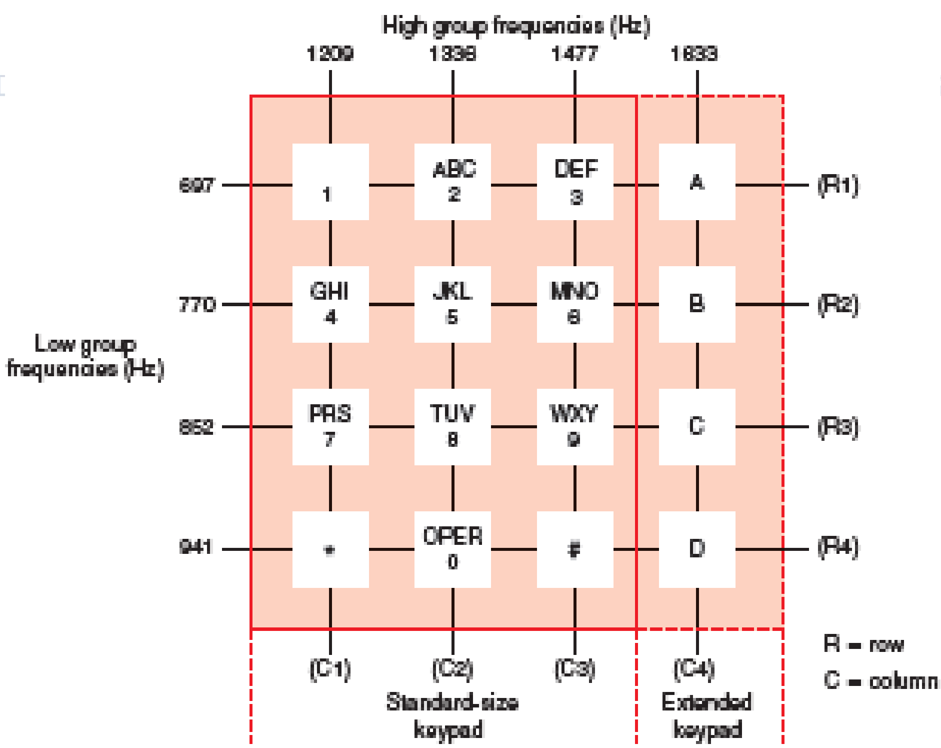
The varistors are usually connected across the line. In Figure, varistor *V*1 is connected in series with resistor *R*1. This varistor automatically shunts some of the current away from the transmitter and the receiver. If the loop is long, the current will be relatively low and the voltage at the telephone will be low. This causes the resistance of the varistor to increase, thus shunting less current away from the transmitter and receiver. On short local loops, the current will be high and the voltage at the telephone will be high. This causes the varistor resistance to decrease; thus more current is shunted away from the transmitter and receiver. The result is a relatively constant level of transmitted or received speech.Note that a second varistor *V*3 is used in the balancing network. The balancing network (*C*3, *C*4, *R*2) works in conjunction with the hybrid to provide the side tone discussed earlier. The varistor adjusts the level of the side tone automatically.

Pulse Dialing. The term *dialing* is used to describe the process of entering a telephone number to be called. In older telephones, a rotary dial was used. In more modern telephones, pushbuttons that generate electronic tones are used fordialing.

The use of a rotary dialing mechanism produces what is known as *pulse dialing.* Rotating the dial and releasing it cause a switch contact to open and close at a fixed rate, producing current pulses in the local loop. These current pulses are detected by the central office and used to operate the switches that connect the dialing telephone to the called telephone. While most telephone companies still support pulse dialing, most dial phones have been long retired. Pulse dialing is no longer widely used.

Tone Dialing. Although some dial telephones are still in use and all central offices can accommodate them, most modern telephones use a dialing system known as *Touch- Tone.* It uses pairs of audio tones to create signals representing the numbers to be dialed. This dialing system is referred to as the ***dual-tone multifrequency (DTMF)*** *system.* A typical DTMF keyboard on a telephone is shown in Figure. Most telephones use a standard keypad with 12 buttons or switches for the numbers 0 through 9 and the special symbols * and #. The DTMF system also accommodates four additional keys for special applications.
In Figure numbers represent audio frequencies associated with each row and column of pushbuttons. For example, the upper horizontal row containing the keys for 1, 2, and 3 is labeled 697, which means that when any one of these three keys is depressed, a sine wave of 697 Hz is produced. Each of the four horizontal rows produces a different frequency. The horizontal rows generate what is generally known as the *low group of frequencies.*

DTMF keypad.



A higher group of frequencies is associated with the vertical columns of keys. For example, the keys for the numbers 2, 5, 8, and 0 produce a frequency of 1336 Hz when depressed.
If the number 2 is depressed, two sine waves are generated simultaneously, one at 697 Hz and the other at 1336 Hz. These two tones are linearly mixed. This combination produces a unique sound and is easily detected and recognized at the central office as the signal representing the dialed digit 2. The tolerance on the generated frequencies is usually within (plus or minus1.5 percent.)

Typical IC Electronic Telephone

The major components of a typical electronic telephone circuit are shown in Figure. Most of the functions are implemented with circuits contained within a single IC. In the Figure, note that the Touch Tone keypad drives a DTMF tone generator circuit. An external crystal or ceramic resonator provides an accurate frequency reference for generating the dual dialing tones.
The tone ringer is driven by the 20-Hz ringing signal from the phone line and drives a piezoelectric sound element. The IC also contains a built-in line voltage regulator. It takes the dc voltage from the local loop and stabilizes it to provide a constant voltage to the internal electronic circuits. An external zener diode and transistor provide bias to the electret microphone.
The internal speech network contains a number of amplifiers and related circuits that fully duplicate the function of a hybrid in a standard telephone. This IC also contains a microcomputer interface. The box labeled MPU is a single-chip *microprocessing unit.* Although it is not necessary to use a microprocessor, if automatic dialing and other functions are implemented, this circuit is capable of accommodating them.

Finally, note the bridge rectifier and hook switch circuit. The twisted pair from the local loop is connected to the tip and ring. Both the 48-V dc and 20-Hz ring voltages will be applied to this bridge rectifier. For direct current, the bridge rectifier provides polarity protection for the circuit, ensuring that the bridge output voltage is always positive. When the ac ringing voltage is applied,the bridge rectifier it into a pulsating dc voltage. The hook switch is shown with the telephone on the hook or in the up position. Thus the dc voltage is not connected to the circuit at this time.
However, the ac ringing voltage will be coupled through the resistor and capacitor to the bridge, where it will be rectified and applied to the two zener diodes *D*1 and *D*2 that drive the tone ringer circuit.

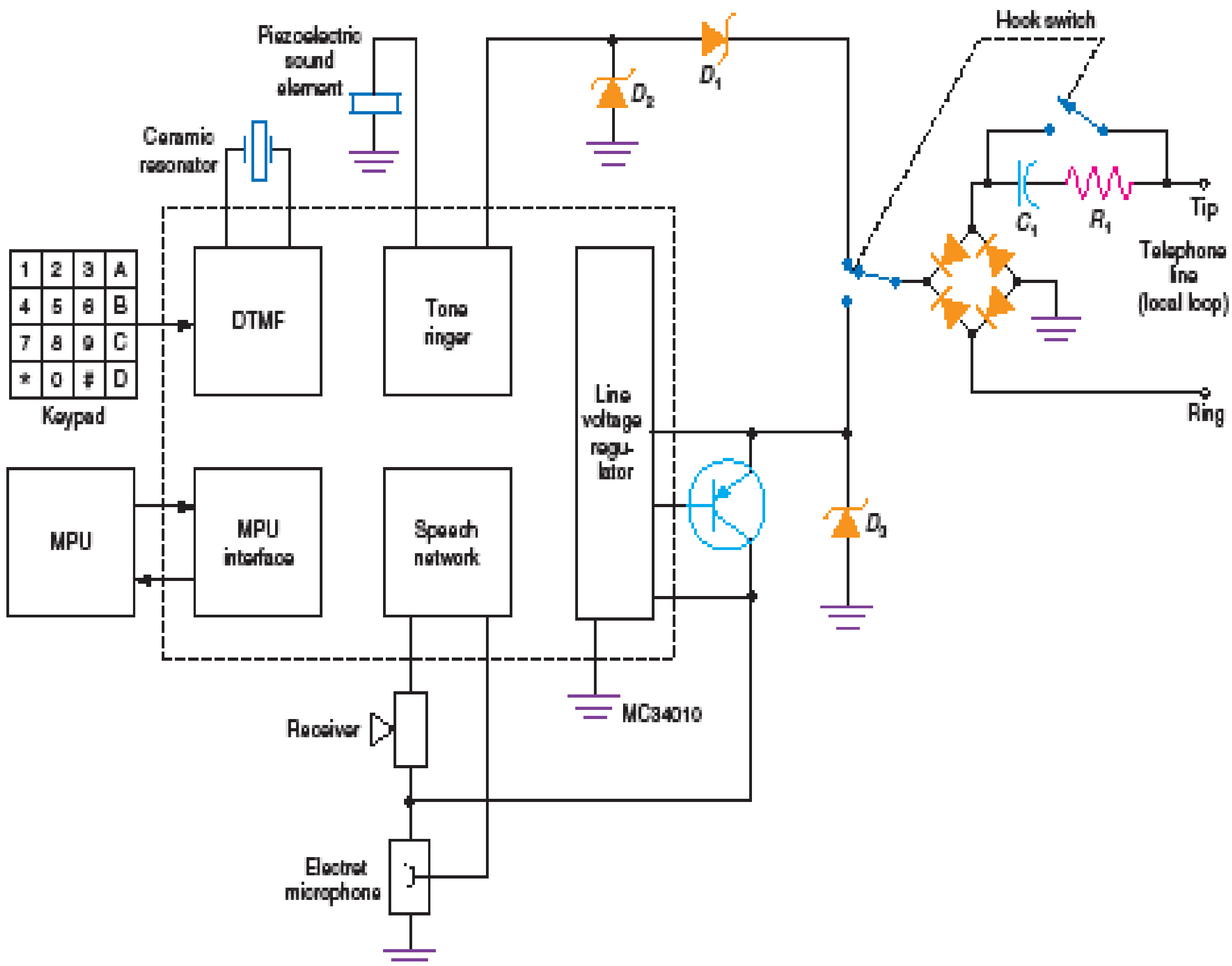
When the telephone is taken off the hook, the hook switch closes, providing a dc path around the resistor and capacitor *R*1 and *C*1. The path to the tone ringer is broken, and the output of the bridge rectii er is connected to zener diode *D*3 and the line voltage regulator.
Thus the circuits inside the IC are powered up, and calls may be received or made.
Microprocessor Control. All modern electronic telephones contain a built-in microcontroller. Like any microcontroller, it consists of the CPU, a ROM in which a control program is stored, a small amount of random access read-write memory, and I/O circuits. The microcontroller, usually a single-chip IC, may be directly connected to the telephone IC, or some type of intermediate interface circuit may be used.

The functions performed by the microcomputer include operating the keyboard and any LCD display, if present. Some other functions involve storing telephone numbers and automatically redialing. Many advanced telephones have the capability of storing 10 or more commonly called numbers. The user puts the telephone into a program mode and uses the Touch Tone keypad to enter the most frequently dialed numbers. These are stored in the microcontroller's RAM. To automatically dial one of the numbers, the user depresses a pushbutton on the front of the telephone. This may be one of the Touch Tone pushbuttons, or it may be a separate set of pushbuttons provided for the purpose. When one of the push buttons is depressed, the microcontroller supplies a preprogrammed set of binary codes to the DTMF circuitry in the telephone IC. Thus the number is automatically dialed. Other features implemented by the microcontroller are caller ID and an answering machine.

Voice Mail. Previously called an *answering machine,* this feature is implemented on most electronic phones. The microcontroller automatically answers the call after a preprogrammed number of rings and saves the voice message. In older answering machines, the message was recorded on a tape cassette. But in modern phones, the voice message is digitized, compressed, and then stored in a small l ash ROM ready for replay. The outgoing message is also stored there.

Caller ID. *Caller ID,* also known as the *calling line identification service,* is a feature that is now widely implemented on most electronic telephones.

Single-chip electronic telephone.

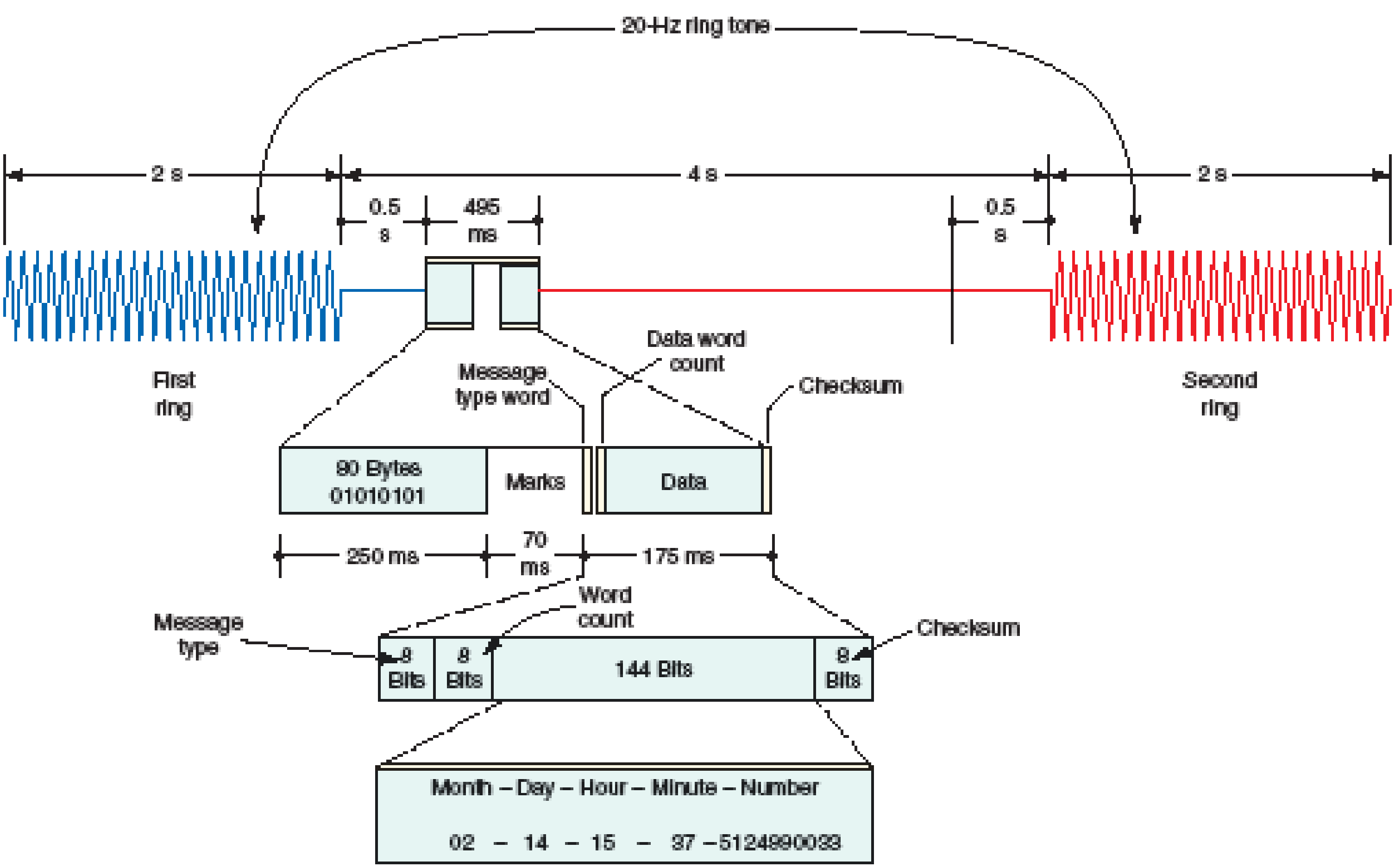


.
With this feature, any calling number will be displayed on an LCD readout when the phone is ringing. This allows to identify the caller. The caller ID service sends a digitized version of the calling number to phone during the first and second rings. The data transmitted includes the date, time, and calling number. Data is transmitted by FSK, where a binary 1 (mark) is a 1200-Hz tone and a binary 0 (space) is a 2200- Hz tone. The data rate is 1200 bps.

There are two message formats in use, the *single-data message format (SDMF)* and the *multiple-data message format (MDMF).* The SDMF is illustrated in Figure. One- half second after the first ring, 80 bytes of alternating 0s and 1s (hex 05) is transmitted for 250 ms followed by 70 ms of mark symbols. These two signals provide initialization and synchronization of the caller ID circuitry in the phone. This is followed by 1 byte describing the message type. This is usually a binary 4 (00000100), indicating the SDMF.

This is followed by a byte containing the message length, usually the number of digits in the calling number. Next the data is transmitted. This is the date, time, and the 10-digit phone number transmitted as ASCII bytes with the least significant digit first. The data format is 2 digits for the month, 2 digits for the day, 2 digits for the hour (military time), 2 digits for the minutes, and up to 10 digits for the calling number. For example, if the date is February 14, the time is 3:37 p.m., and the calling number is 512-499-0033, the data sequence would be 021415375124990033. The initial byte in the message is the checksum that is used for error detection. The checksum is the 2s complement sum (XOR) of all the data bytes not including the initialization and sync signals.

The caller ID transmission format.



If the calling number is outside the calling area, the system will display an O on the LCD rather than the calling number. Furthermore, a caller may also have his or her number blocked. This can be done by setting it up with the service provider in advance or by dialing *67 prior to making the call. This will cause a P to be displayed on the LCD instead of the calling number. A more advanced data format is the MDMF. It is similar to the SDMF but includes an extra field for the name of the calling party plus additional identification bytes.

Subscriber Interface (Optional)

Most telephones are connected to a local central office by way of the two-line, twisted pair local loop cable. The central office contains all the equipment that operates the telephone and connects it to the telephone system that makes the connection to any other telephone.
Each telephone connected to the central office is provided with a group of basic circuits that power the telephone and provide all the basic functions, such as ringing, dial tone, and dialing supervision. These circuits are collectively referred to as the *subscriber interface* or the *subscriber line interface circuit (SLIC).* In older central office systems, the subscriber interface circuits used discrete components. Today, most functions of the subscriber line interface are implemented by one or perhaps two integrated circuits plus supporting equipment. The subscriber line interface is also referred to as the *line side interface.*
The SLIC provides seven basic functions generally referred to as *BORSCHT.*

The (telephone) line interface is often referred to as a BORSCHT circuit.

This acronym describes the functional requirements of a standard telephone line interface. The tip and ring leads of the telephone set are wired through some protection devices to the line interface located in the peripheral module. This interface must perform the following functions:

B-Battery feed,
O- Over voltage protection,
R- Ringing, S Supervision & Signaling,
C – Coding, H- Hybrid and T - test

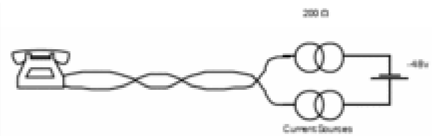
Many of these functions can be integrated into a single IC, often called a SLIC chip (subscriber line interface chip). SLICs have been available for the PBX market for over a decade. Recently however, they have also become available for the central office environment as well.

B - Battery Feed

Most domestic appliances are powered from an electric utility grid. The notable exception to this is the telephone. This is because the telephone should still operate in the event of a power failure. Indeed, the telephone is vital in case of disaster or emergency.
The telephone office provides a nominal -48 volt dc feed to power the phone. This magnitude is considered the maximum safe dc operating potential. It would not be in the telephone company's best interest to provide a dc voltage, which could electrocute its customers, or it's own employees. A negative potential was chosen to reduce corrosive action on buried cables.
Multi-function telephones cannot always be powered from the telephone exchange and

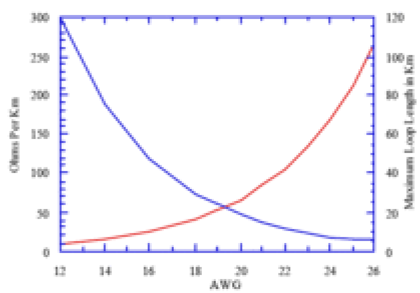
often require an alternate power source. For this reason, sophisticated line interfaces such as ISDN SAA interfaces have a ‗fail to POTS‘ mode. If the electric power fails, the complex phone cannot function to full capacity. The telephone exchange can sense the local power outage through the telephone loop and switches to POTS only service.
The POTS loop requires a nominal -48 v at 20 – 100 ma dc to maintain a voice and signaling path. The earpiece in the handset does not require biasing, but the carbon microphone does. Subscriber signaling is performed by temporarily placing a short circuit on the loop thus changing the loop current, which is then sensed at the central office.

There are several ways to provide loop current, the simplest being a resistor in series with a battery.



Another way to provide loop current is by an electronic current source.

Although this method is quite complex, it has become quite popular with the advent of high voltage bipolar technology. One of the more difficult requirements to meet is the 60-dB longitudinal line balance requirement. To achieve this, the impedance to ground on each side of the loop, must match within 0.1%. This is easy to do with laser trimmed thick film resistors, but a bit tricky with current sources.



A standard telephone requires a minimum of about 20 ma. This means that the maximum possible loop resistance is about 2000 Ω. In actual practice, the loop is generally limited to 1250 W. The maximum loop length is determined by the wire gauge.

O - Over-voltage Protection

The two major types of over-voltage that can occur are lightning strikes and power line contact. In both cases, the circuit must either recover or fail-safe. Under no circumstances can a surge be allowed to propagate further into the system, or create a fire.
Initial surge protection is provided at the MDF by gas tubes and/or carbon blocks, which arc if the applied voltage exceeds a few hundred volts. Since these devices take a finite time to respond, high-speed diodes are also used at the line circuit inputs.

R - Ringing

Ringing is often provided by means of a dedicated ringing generator that is connected onto the loop by means of a relay. It is possible to generate ringing voltages at the line interface if the current generators have a high enough voltage source available to them. Or alternately, a switching converter with step up capability can be place on the interface.

S - Supervision & Signaling

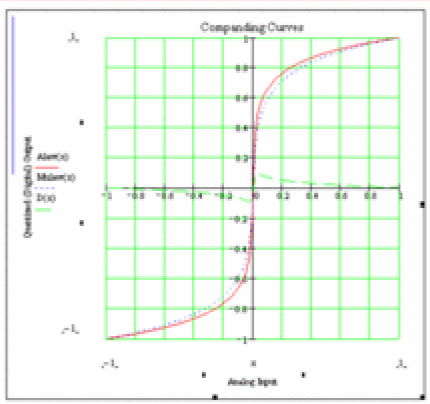
The central office must supervise the loop in order to identify customer requests for service. A request for service is initiated by going off-hook. This simply draws loop current from the CO. Loop current at the far-end is monitored during ringing to enable the CO to disconnect the ringing generator when the phone is answered. The office continues to monitor the loop current at both ends of the connection throughout the call, to determine when the call is terminated by hanging up.

Signaling is a way to inform the CO what the customer wants. The two basic signaling methods used in customer loops are dial pulse and touch-tone. It is interesting to note that preferred customer loop signaling method in analog exchanges is digital, while the preferred method in digital exchanges is analog.

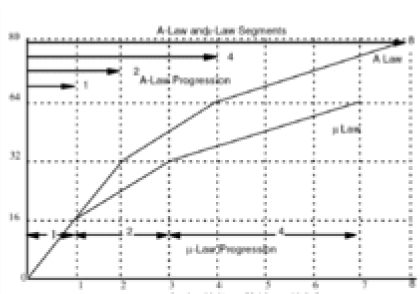
MF Signaling Tones

Two tones are used to perform the signaling function to eliminate the possibility that speech be interpreted as a signal. At one time DTMF decoders were costly and bulky devices located in a common equipment bay, but today with the advent of LSI technology, this function can be performed on a chip. An example is the Mitel MT8865 DTMF filter, and MT8860 DTMF decoder Positions 11 to 14 are not presently being used.

C - Coding



Telecommunications signals are seldom linearly encoded, but rather are companded (a combination of compression & expansion). This allows for a more uniform S/N ratio over the entire range of signal sizes. Without companding, a 12 bit linear encoding scheme would be needed to obtain the same S/N ratio at low volume levels. It also reduces the noise and crosstalk levels at the receiver.

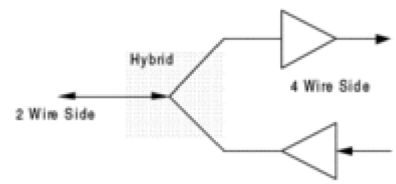


Since the highest frequency passed is about 3.4 kHz, a great deal of ingenuity is required to

pass data at 4.8, 9.6 kbps or even higher. Note that these are well above the Nyquist rate but considerably below the Shannon-Hartley limit.
All modern telephone systems today employ codecs in the BORSCHT interface to digitize the incoming analog signals. It is ironic that although the telephone system has been updated to digital technology, the telephone set and loop has remained analog.

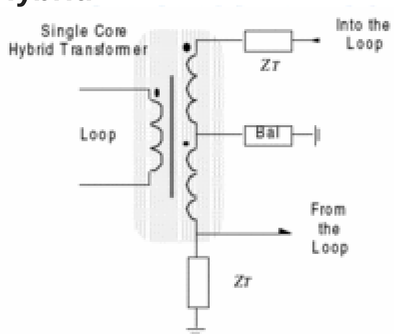
By international agreement, all voice codecs use an 8 kHz sampling rate. Since each transmitted sample is 8 bits long, the analog voice signal is encoded into a 64 kbps binary steam. This rate determines the basic channel data rate of most other digital communications systems.
By bypassing the codec, it is possible to send 64 kbps customer data through the telephone system. However, because of old style signaling schemes still in use, digital data rates are often limited to 56 kbps.

H – Hybrid



A diplexer performs a bi-directional 2-wire to 4-wire conversion. It allows two unidirectional electrical paths to be combined into a single bi-directional one, and vice versa. It is advantageous to separate transmit and receive portions of the signal since it is easier to make unidirectional amplifiers, filters, and logic devices.
One of the simplest ways to create an audio band hybrid is to use a transformer hybrid.

Single Core Transformer Hybrid

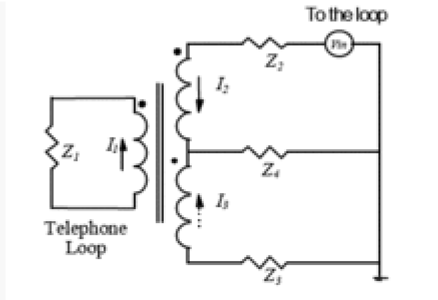


There are several ways to split transmit and receive paths, the simplest method uses a single core hybrid transformer.
The basic defining transformer equations are:

$$\frac{V_1}{n_1} = \frac{V_2}{n_2} = \frac{V_3}{n_3} = \cdots \qquad and \qquad I_1 n_1 + I_2 n_2 + I_3 n_3 + \cdots = 0$$

For a single core hybrid with a center-tapped secondary, the impedance relationships for proper operation (conjugate matching) are:

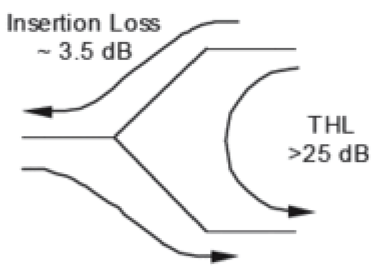
$$for \qquad n_2 = n_3 \qquad Z_2 = Z_3 = 2Z_4 \qquad and \qquad Z_1 = \left(\frac{n_2}{n_1}\right)^2 Z_4$$



$$let \qquad n_1 = n_2 = n_3 \qquad \therefore I_3 = -I_1 - I_2$$

$$I_3 = -(-I)_1 - I_2 \qquad and\,if \qquad |I_1| = |I_2| \qquad then \qquad I_3 = 0$$

Note what happens if the transformer is driven from one of the secondary windings: But I_1 and I_2 flow in the opposite directions, therefor:
This last requirement can be satisfied by adjusting the impedances Z_1 - Z_4 to make the currents equal. From this observe that signals injected into any port emerge only at adjacent ports but not at the opposite one.



In a properly balanced single core hybrid the typical throughput or insertion loss is about 3.5 dB and the THL (trans hybrid loss) is about 25 dB.

Double Core Hybrid

When properly balanced, a 2-core network can achieve a THL of 50 dB while the insertion loss remains at about 3.5 dB. It has better performance than the single core device, but is bulkier and more expensive.

Balancing Networks

All telecom equipment is tested and characterized against standard impedance terminations. These impedances are based on line surveys and are approximate equivalent circuit representations of the outside cabling plant.

T – Testing

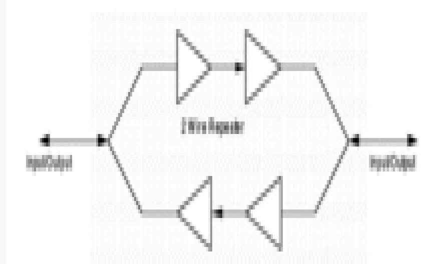
In order to maintain a high degree of service (99.999%), the equipment must be capable of detecting and repairing faults before the customer is even aware that there may be a problem.

As a result, a separate test buss and access relay is provided on a line interface. Tests may be performed in a bridged mode or with the loop and line card disconnected from each other.

Testing can be done in three basic directions:

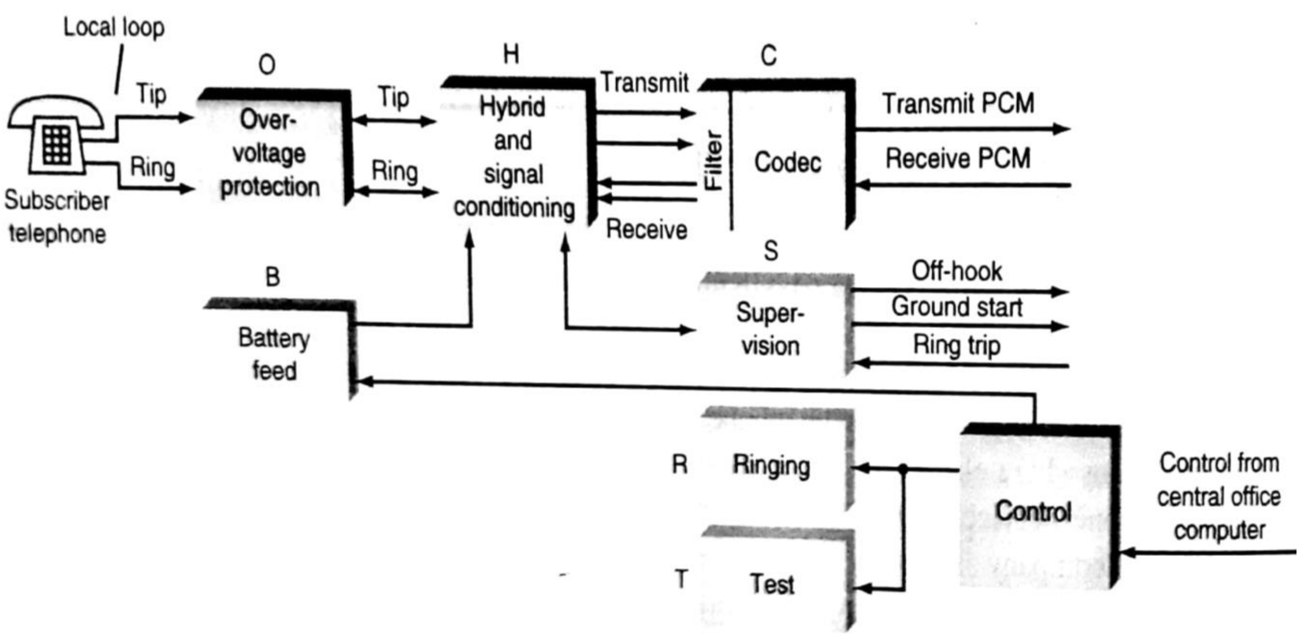
- From the line interface looking out towards the subscriber loop
- From the loop connection looking into the line card
- From the central office side of the line card .These tests are generally automated and are conducted late at night when there is little chance that the customer will request service, thus interrupting the test. Some of the scheduled tests may include:
- Transmit and receive levels
- Transmit and receive frequency response
- Insertion loss
- Trans-hybrid loss
- Quantization distortion
- Aliasing distortion Some other tests that may be performed when commissioning a line or when a complaint is lodged, include:
- Impulse noise test
- C-message noise
- Longitudinal balance

Repeaters



By placing two hybrids back to back, it is possible to create a bidirectional amplifier or repeater. The total gain in the 4-wire path within the repeater must not exceed the combined trans hybrid loss of the transformers. If this happens, the circuit will oscillate or sing. The total gain in the 4-wire path within the repeater must not exceed the combined trans hybrid loss of the transformers. If this happens, the circuit will oscillate or sing.

: 18-17 BORSCHT functions in the subscriber line interface at the central office.



Facsimile Machine

It scans the contents of a document (as an image, not text) to create electronic signals. Scanning is done electronically and the scanned signal is converted into a binary signal. These signals are then sent to the destination (another FAX machine) in an orderly manner using telephone lines. At the destination, the signals are reconverted into a replica of the original document. Note that FAX provides image of a static document unlike the image provided by television of objects that might be dynamic. Today's modern fax machine is a high-tech electro- optical machine. Digital transmission with standard modem techniques is used.

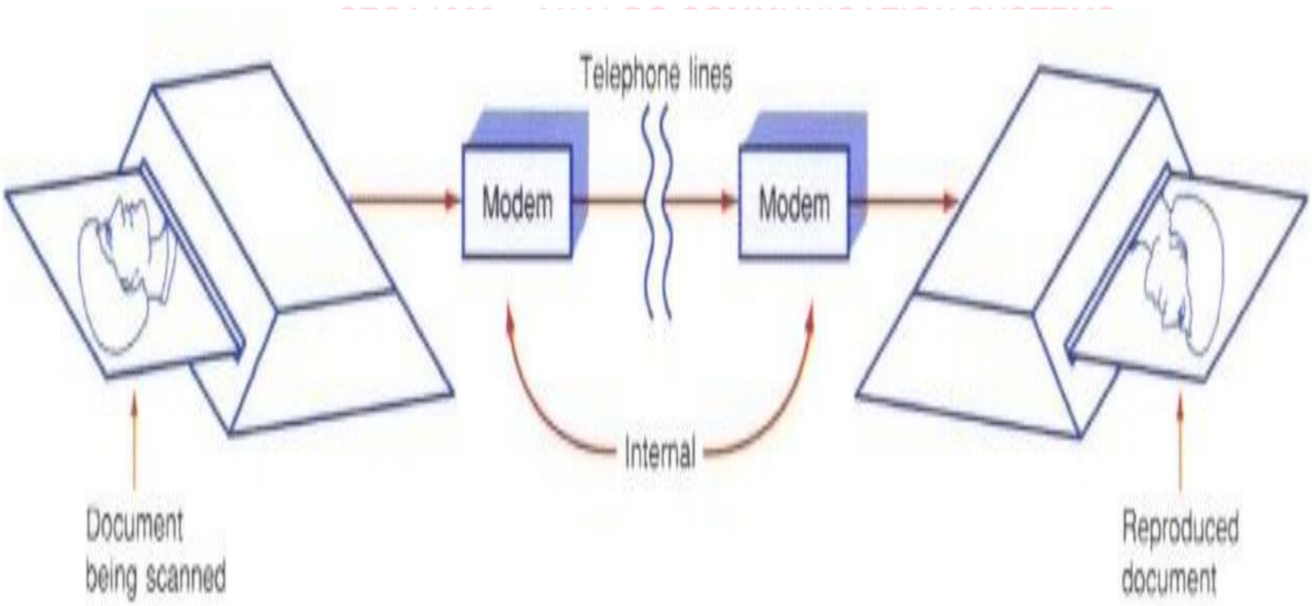
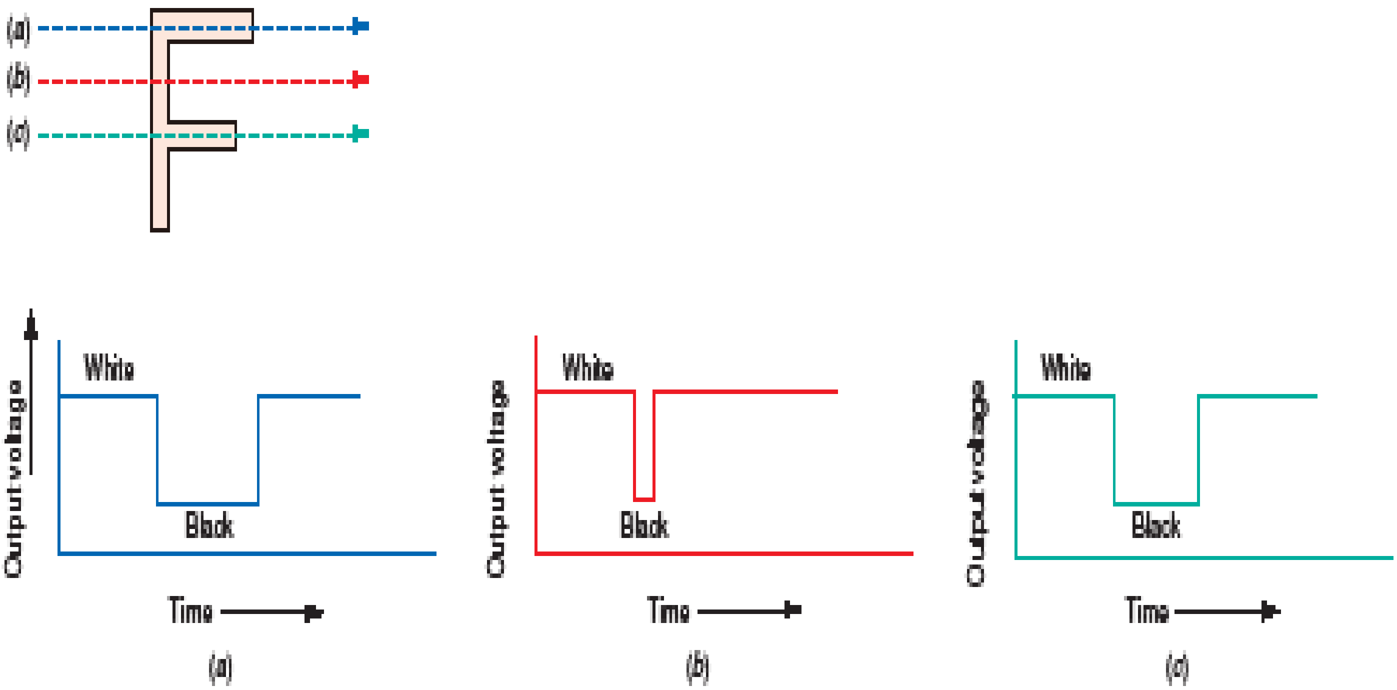


Figure shows how a printed letter might have been scanned. Assume that the letter F is black on a white background. The output of a photo detector as it scans across line *a* is shown in Fig. (*a*). The output voltage is high for white and low for black. The output of the photo detector is also shown for scan lines *b* and *c*. The output of the photo detector is used to modulate a carrier, and the resulting signal is put on the telephone line. The resolution of the transmission is determined by the number of scan lines per vertical inch. The greater

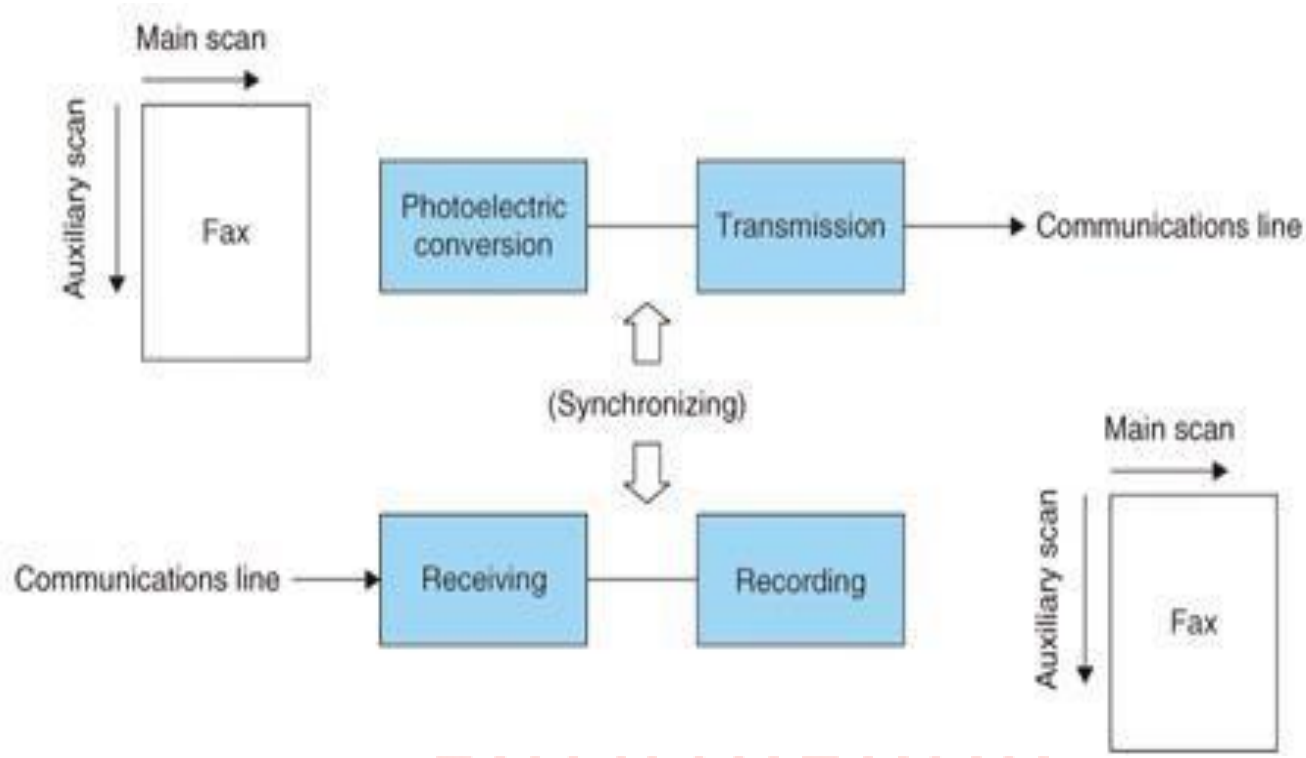
Output of a photosensitive detector during different scans.



the number of lines scanned, the inner the detail transmitted and the higher the quality of reproduction. Older systems had a resolution of 96 lines per inch (LPI), and the new systems have 200 LPI.

Facsimile Operation

A facsimile (commonly referred to as a fax) is a production of an exact copy of a document by electronic scanning, and the subsequent transmission of the resulting data. Faxes are transmitted over ordinary phone lines using fax machines. In a typical fax transmission, the document to be faxed is placed in the document feeder of a fax machine and the telephone number of the destination fax machine is dialed. In a very short time, a replica of the document is received at the destination fax machine. By their very nature, faxes can contain any information that appears in written form. As such, faxes will often contain information that is personal, or otherwise confidential.

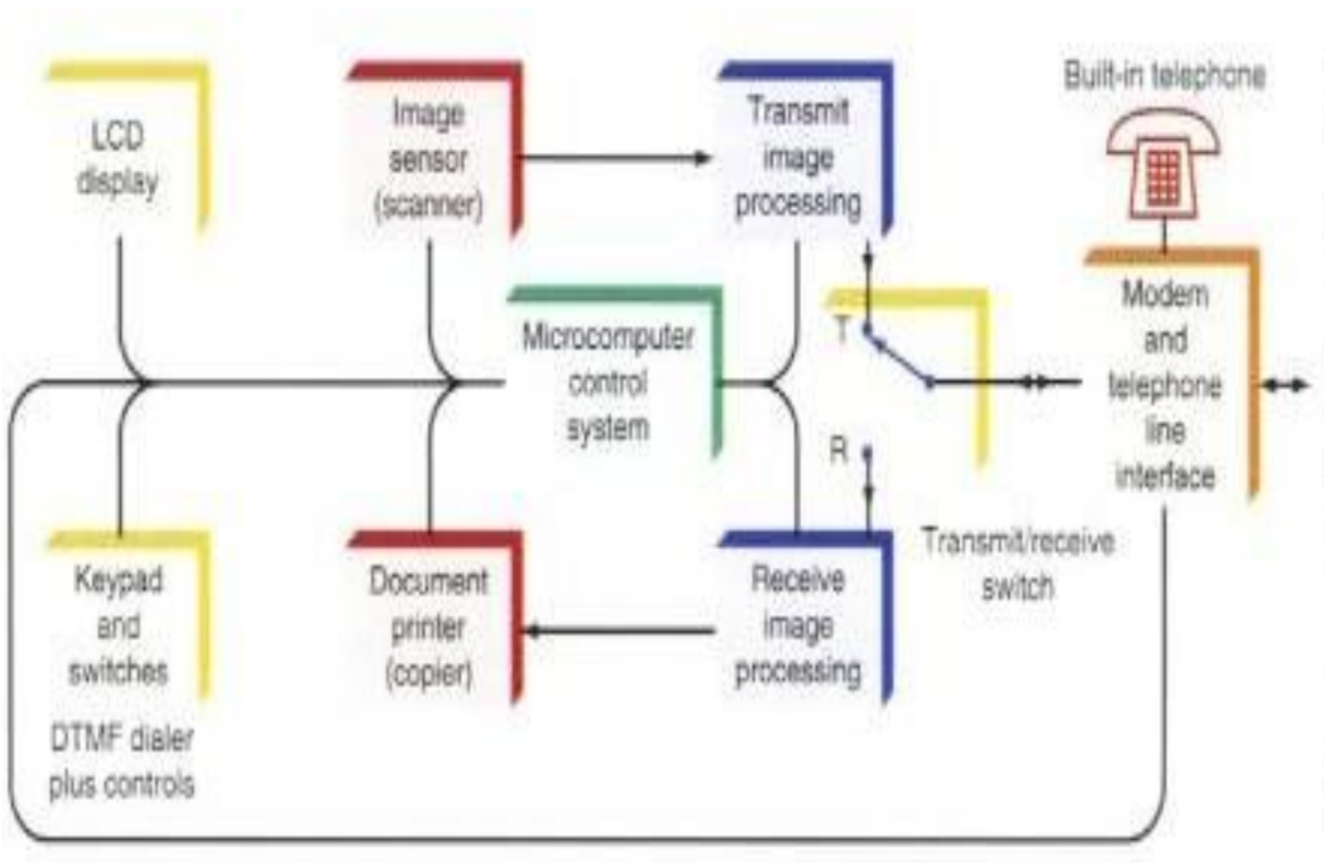


The process begins with an image scanner that converts the document into hundreds of horizontal scan lines. Many techniques are used, but they all incorporate a photo-(light) sensitive device to convert light variations along one scanned line into an electric voltage. The resulting signal is then processed in various ways to make the data smaller and faster to transmit. The resulting signal is sent to a modem where it modulates a carrier set to the middle of the telephone voice spectrum bandwidth. The signal is then transmitted to the receiving fax machine over the public-switched telephone network. The receiving machine's modem demodulates the signal that is then processed to recover the original data.

Modem fax machine

A device can attach to a personal computer that enables to transmit and receive electronic documents as faxes. A fax modem is like a regular modem except that it is designed to transmit documents to a fax machine or to another fax modem. Some, but not all, fax modems do double duty as regular modems. As with regular modems, fax modems can be either internal or external. Internal fax modems are often called fax boards. Documents sent through a fax modem must already be in an electronic form (that is, in a disk file), and the documents receive are likewise stored in files on disk. To create fax documents from images on paper, need an optical scanner. Fax modems come with communications software similar to communications software for regular modems. This software can give the fax modem many capabilities that are not available withstand-alone fax machines. For example, broadcast a fax document to several sites at once.
Figure shows a block diagram of a modern fax machine. The transmission process begins

with an image scanner that converts the document to hundreds of horizontal scan lines. Many different techniques are used, but they all incorporate a photo- (light-) sensitive device to convert light variations along one scanned line into an electrical voltage.
The resulting signal is then processed in various ways to make the data smaller and thus faster to transmit. The resulting signal is sent to a modem where it modulates a carrier set to the middle of the telephone voice spectrum bandwidth. The signal is then transmitted to the receiving fax machine over the public switched telephone network. The receiving fax machine‘s modem demodulates the signal that is then processed to recover the original data. The data is decompressed and then sent to a printer, which reproduces the document. Because all fax machines can transmit as well as receive, they are referred to as *transceivers.* The transmission is half duplex because only one machine may transmit or receive at a time.



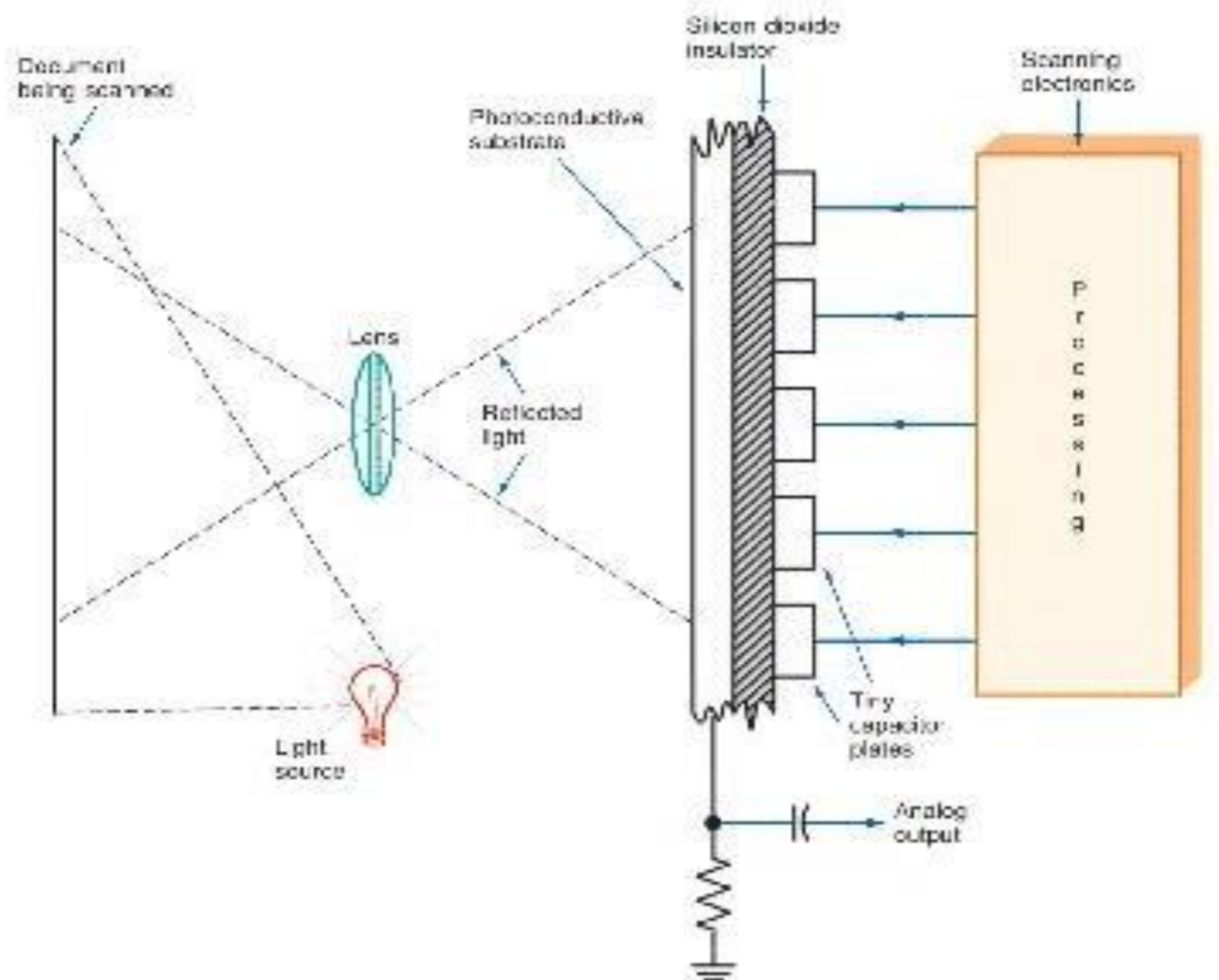
Most fax machines have a built-in telephone, and the printer can also be used as a copy machine. An embedded microcomputer handles all control and operation, including paper handling.

Most fax machines use **charged coupled devices (CCDs) for scanning. A CCD** is a light-sensitive semiconductor device that converts varying light amplitudes into an electrical signal.
Data compression is a digital data processing technique that looks for redundancy in the transmitted signal. . Every fax machine contains a built-in modem that is similar to a conventional data modem for computers.

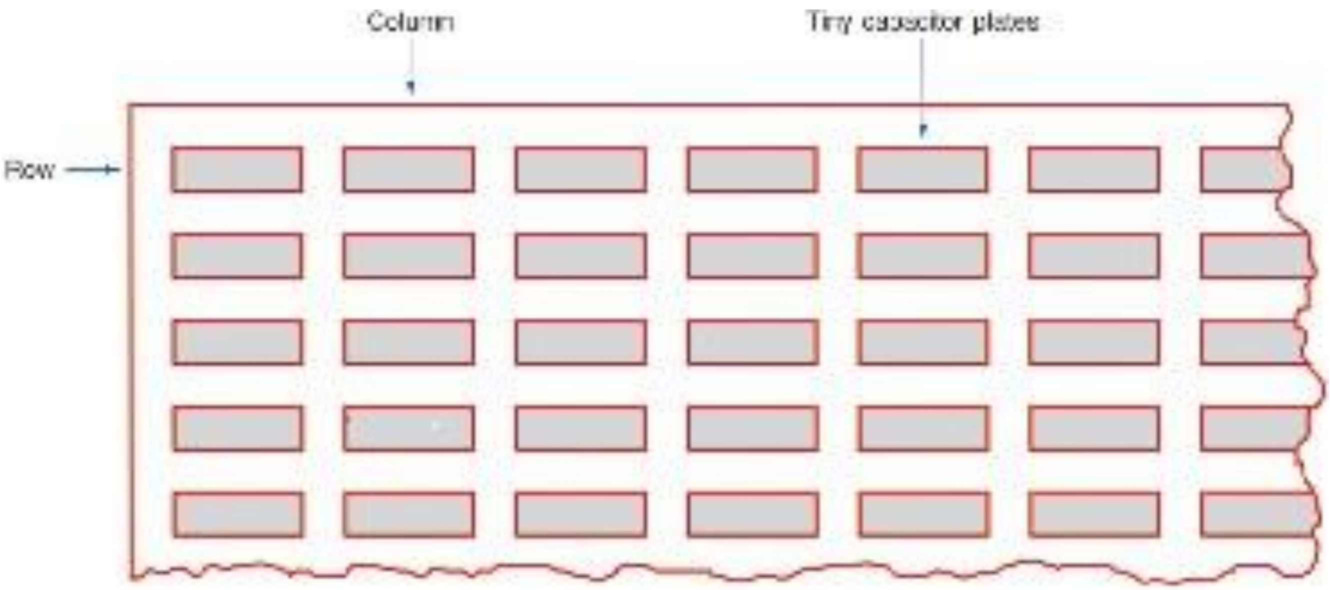
Image Processing

Most fax machines use *charge-coupled devices (CCDs)* for scanning. A CCD is a light sensitive semiconductor device that converts varying light amplitudes to an electric signal. The

A charge-coupled device is used to scan documents in modern fax machines. (*a*) Cross section. (*b*) Detail of capacitor matrix.



typical CCD is made up of many tiny reverse-biased diodes that act as capacitors, which are manufactured in a matrix on a silicon chip The base forms one large plate of a capacitor that is



electrically separated by a dielectric from many thousands of tiny capacitor plates, as shown in figure. When the CCD is exposed to light, the CCD capacitors charge to a value proportional to the light intensity. The capacitors are then scanned or sampled electronically to determine their charge. This creates an analog output signal that accurately depicts the image focused on the CCD.

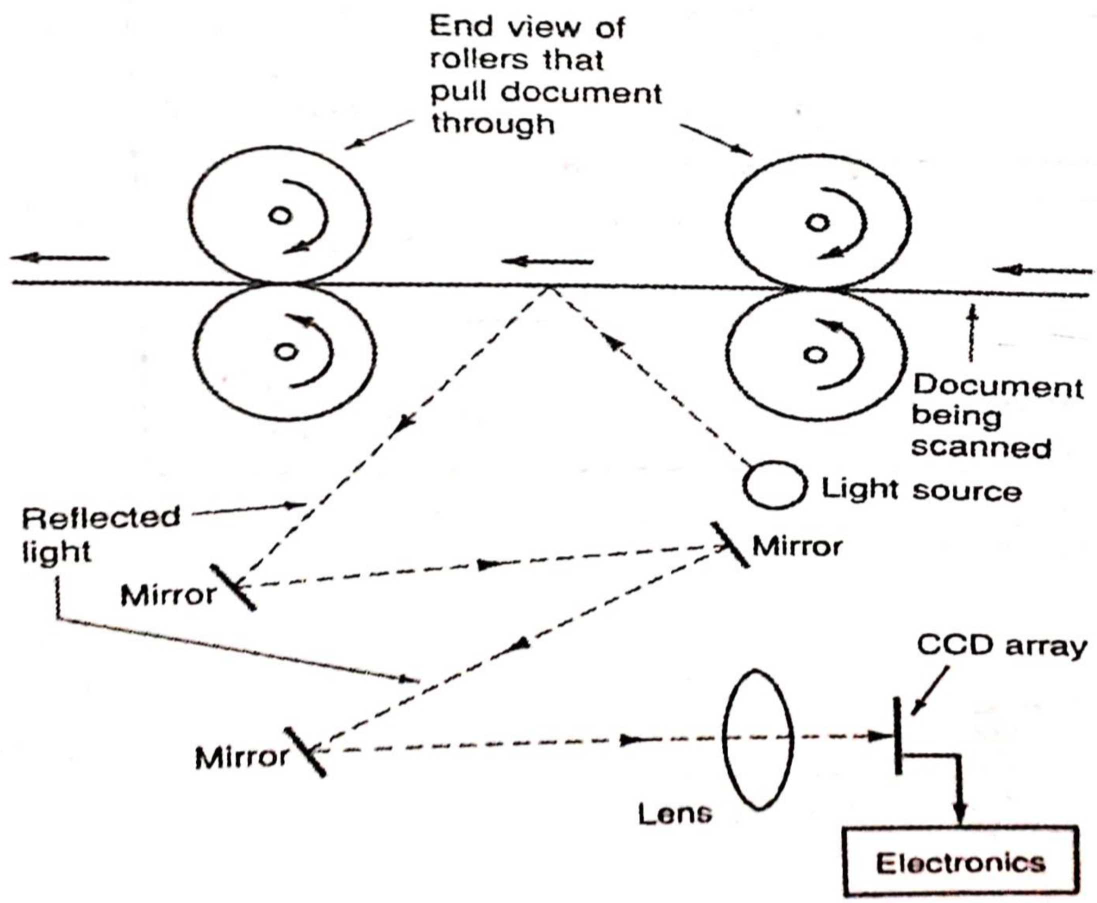
A CCD is actually a device that breaks up any scene or picture into *individual picture elements,* or *pixels.* The greater the number of CCD capacitors, or pixels, the higher the resolution and the more faithfully a scene, photograph, or document can be reproduced. CCDs are available with a matrix of many thousands of pixels, thereby permitting very high-resolution picture transmission. CCDs are widely used in modern video cameras in place of the more delicate and more expensive vidicon tubes. In the video camera (camcorder), the lens focuses the entire scene on a CCD matrix. This same approach is used in some fax machines. In one type of fax machine, the document to be transmitted is placed face down as it might be in a copy machine. The document is then illuminated with brilliant light from a xenon or fluorescent bulb. A lens system focuses the reflected light on a CCD. The CCD is then scanned, and the resulting output is an analog signal whose amplitude is proportional to the amplitude of the reflected light.

In most desktop fax machines, the entire document is not focused on a single CCD. Instead, only a narrow portion of the document is lighted and examined as it is moved through the fax machine with rollers. A complex system of mirrors is used to focus the lighted area on the CCD.

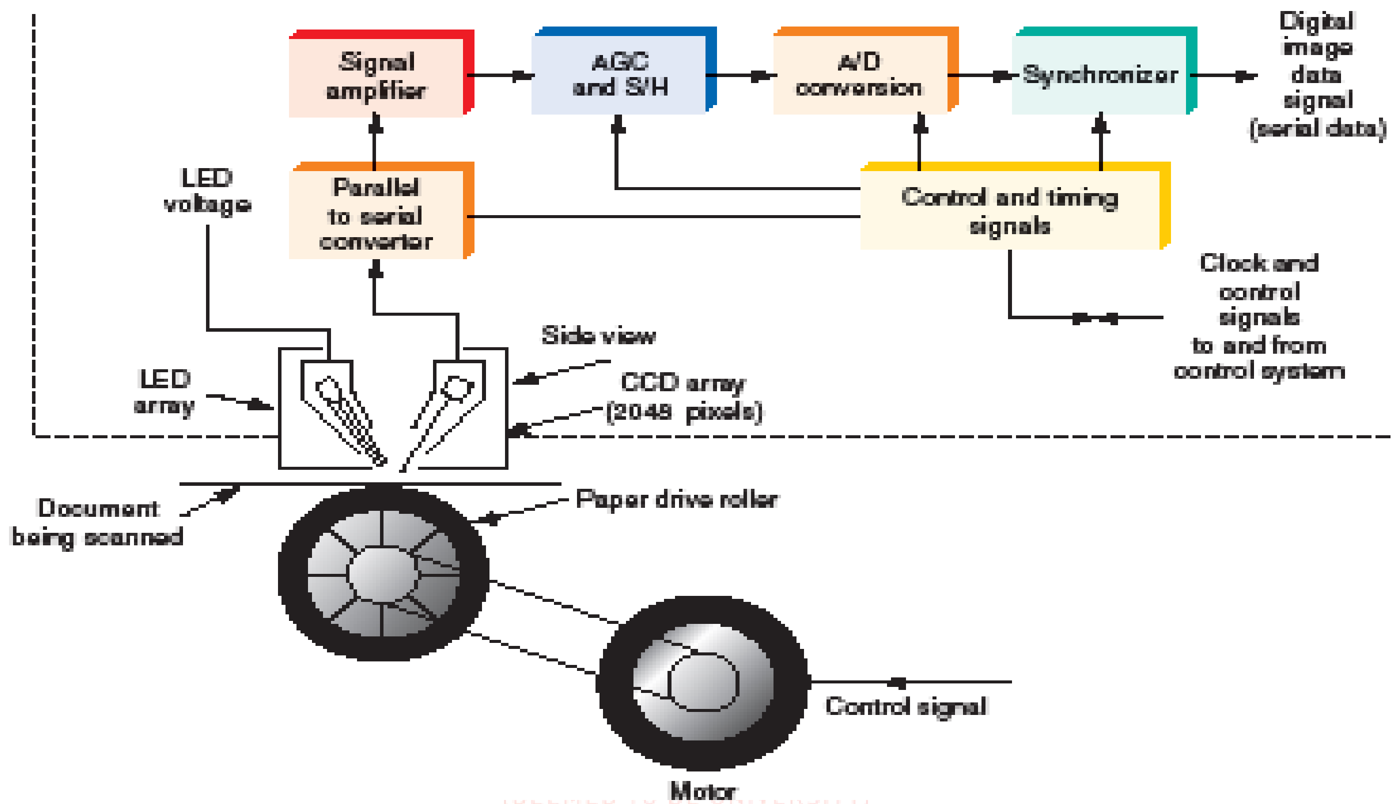
The more modern fax machines use another type of scanning mechanism that does not use lenses. The scanning mechanism is an assembly made up of an LED array and a CCD array. These are arranged so that the entire width of a standard 81⁄2 3 11 in page is scanned simultaneously one line at a time. The LED array illuminates a narrow portion of the document. The reflected light is picked up by the CCD scanner. A typical scanner has 2048 light sensors forming one scan line. Fig. 18-19 shows a side view of the scanning mechanism. The 2048 pixels of light are converted to voltages proportional to the light variations on one scanned line. These voltages are converted from a parallel format to a serial voltage signal. The resulting analog signal is amplified and sent to an AGC circuit and an S/H amplifier. The signal is then sent to an A/D converter where the light signals are translated to binary data words for transmission.

An enormous amount of data is generated by scanning one page of a document. A typical 81⁄2 3 11 in page represents about 40,000 bytes of data. This can be shortened by a factor of 10 or more with *data compression techniques.* Furthermore, because of the narrow bandwidth of telephone lines, data rates are limited. That is why it takes so long to transmit one page of data. Developments in high-speed modems have helped reduce the transmission time, but the most important developments are data compression techniques that reduce the overall amount of data, which significantly decreases the transmission time and telephone charges.

Scanning mechanism in a fax machine.



LED/CCD scanner mechanism in a modern fax machine.



Data Compression ***Data compression*** **is a digital data processing technique that looks for redundancy in the transmitted signal. White space or continuous segments of the page that are the same shade produce continuous strings of data words that are the same.**

These can be eliminated and transmitted as a special digital code that is significantly faster to transmit. Other forms of data compression use various mathematical algorithms to reduce the amount of data to be transmitted.

The data compression is carried out by a *digital signal processing (DSP) chip.* This is a high-speed microprocessor with embedded ROM containing the compression program. The digital data from the A/D converter is passed through the DSP chip, from which comes a significantly shorter string of data that represents the scanned image. This is what is transmitted, and in far less time than the original data could be transmitted. At the receiving end, the demodulated signal is decompressed. Again, this is done through a DSP chip especially programmed for this function. The original data signal is recovered and sent to the printer.

Modems

Every fax machine contains a built-in *modem* that is similar to a conventional data modem for computers. These modems are optimized for fax transmission and reception. And they follow international standards so that any fax machine can communicate with any other fax machine. A number of different modulation schemes are used in fax systems. Analog fax systems use AM or FM.

Digital fax uses PSK or QAM. To ensure compatibility between fax machines of different manufacturers, *facsimile standards* have been developed for speed, modulation methods, and resolution by the *International Telegraph and Telephone Consultative Committee,* better known by its French abbreviation, *CCITT.* The CCITT is now known as the *ITU-T,* or *International Telecommunications Union.* The ITU-T fax standards are divided into four groups:

1. ***Group 1 (G1 or GI):*** Analog transmission using frequency modulation where white is 1300 Hz and black is 2100 Hz. Most North American equipment uses 1500 Hz for white and 2300 Hz for black. The scanning resolution is 96 lines per inch (LPI). Average transmission speed is 6 minutes per page (81⁄2 3 11 in or A4 metric size, which is slightly longer than 11 in).
2. ***Group 2 (G2 or GII):*** Analog transmission using FM or vestigial sideband AM. The vestigial sideband AM uses a 2100-Hz carrier. The lower sideband and part of the upper sideband are transmitted. Resolution is 96 LPI. Transmission speed is 3 min or less for an 81⁄2 3 11 in or A4 page.
3. ***Group 3 (G3 or GIII):*** Digital transmission using PCM black and white only or upto 32 shades of gray. PSK or QAM to achieve transmission speeds of up to 9600 Bd. Resolution‘s 200 LPI. Transmission speed is less than 1 minute per page, with 15 to 30 s being typical.
4. ***Group 4 (G4 or GIV):*** Digital transmission, 56 kbps, resolution up to 400 LPI, and speed of transmission less than 5 s. The older G1 and G2 machines are no longer used. The most common configuration is group 3. Most G3 machines can also read the G2 format.

The G4 machines are not yet widely used. They are designed to use digital transmission only with no modem over very wideband dedicated digital-grade telephone lines. Both G3 and G4 formats also employ digital data compression methods that shorten the binary data stream considerably, thereby speeding up page transmission. This is important because shorter transmission times cut long- distance telephone charges and reduce operating costs.

Fax Machine Operation

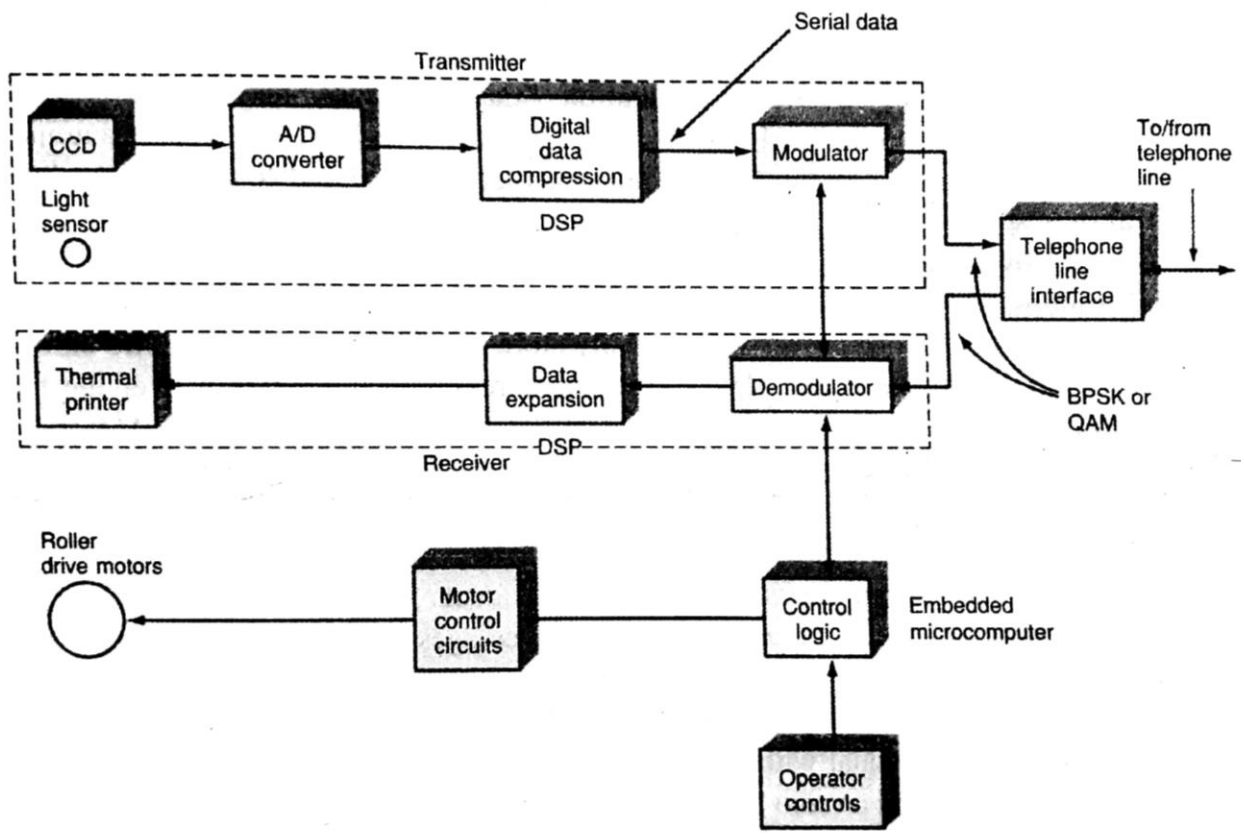
Figure shows simplified block diagram of the transmitting circuits in a modern G3 fax transceiver. The analog output from the CCD array is serialized and fed to an A/D converter that translates the continuously varying light intensity into a stream of binary numbers. Sixteen gray scale values between white and black are typical. The binary data is sent to a DSP digital data compression circuit as described earlier. The binary output in serial data format is used to modulate a carrier that is transmitted over the telephone lines. The techniques are similar to those employed in modems.

Speeds of 2400/4800 and 7200/9600 Bd are common Most systems use some form of PSK or QAM to achieve very high data rates on voice-grade lines. In the receiving portion of the fax machine, the received signal is demodulated and then sent to DSP circuits, where the data compression is removed and the binary signals are restored to their original form. The signal is then applied to a printing mechanism. The most common fax printer today is an ink jet printer like those popularly used with PCs. In the high-priced machines, laser scanning of an electro sensitive drum, similar to the drum used in laser printers, produces output copies by using the proven techniques of xerography. The control logic in Figure is usually an embedded microcomputer. Besides all the internal control functions it implements, it is used for –handshaking‖ between the two machines that will communicate. This ensures compatibility. Handshaking is usually carried out by exchanging different audio tones.

The called machine responds with tones designating its capability. The calling machine compares this to its own standards and then either initiates the transmission or terminates it because of incompatibility. If the transmission proceeds, the calling machine sends synchronizing signals to ensure that both machines start at the same time.

The called machine acknowledges the receipt of the sync signal, and transmission begins. All the protocols for establishing communication and sending and receiving the data are standardized by the ITU-T. Transmission is half duplex. As improvements have been made in picture resolution quality, transmission speed, and cost, facsimile machines have become much more popular. The units can be easily attached with standard RJ-11 modular connectors to any telephone system. In most business applications, the fax machine is typically dedicated to a single line. Most fax machines feature fully automatic operation with microprocessor-based control. A document can be sent to a fax machine automatically. The sending machine simply dials the receiving machine and initiates the transmission. The receiving machine answers the initial call and then reproduces the document before hanging up.

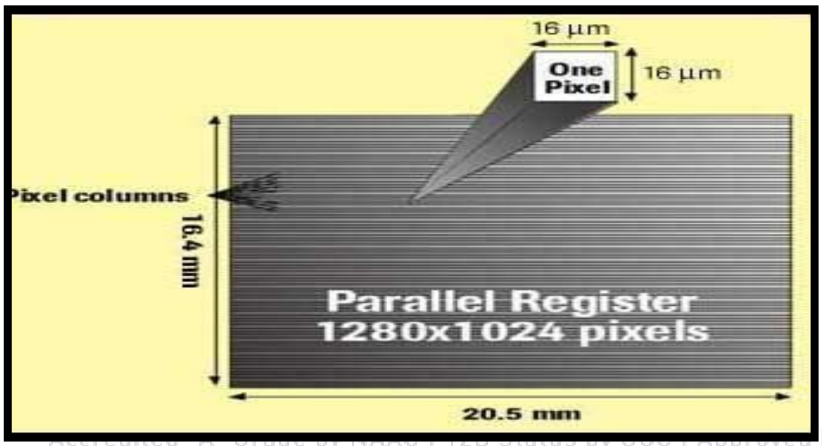
18–26 Block diagram of a facsimile machine.



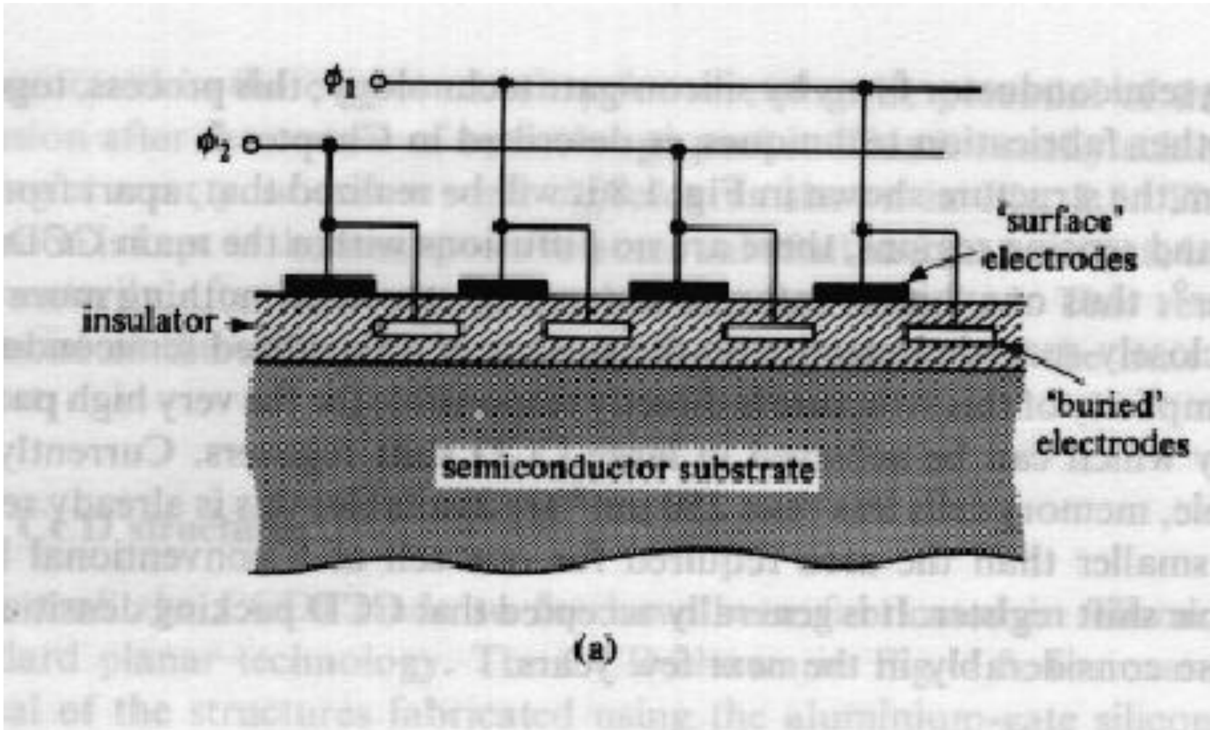
Most fax machines have a built-in telephone and are designed to share a single line with conventional voice transmission. The built-in telephone usually features Touch- Tone dialing and number memory plus automatic redial and other modern telephone features. Most fax machines also have automatic send and receive features for fully unattended operation. Fax machines are slowly fading away as technology changes. Today, most computer printers incorporate a scanner and a printer. The fax function including a data-only telephone with RJ-11 connection is built into the printer. A scanned documents the digitized and sent using the fax procedures described earlier.

CCD:

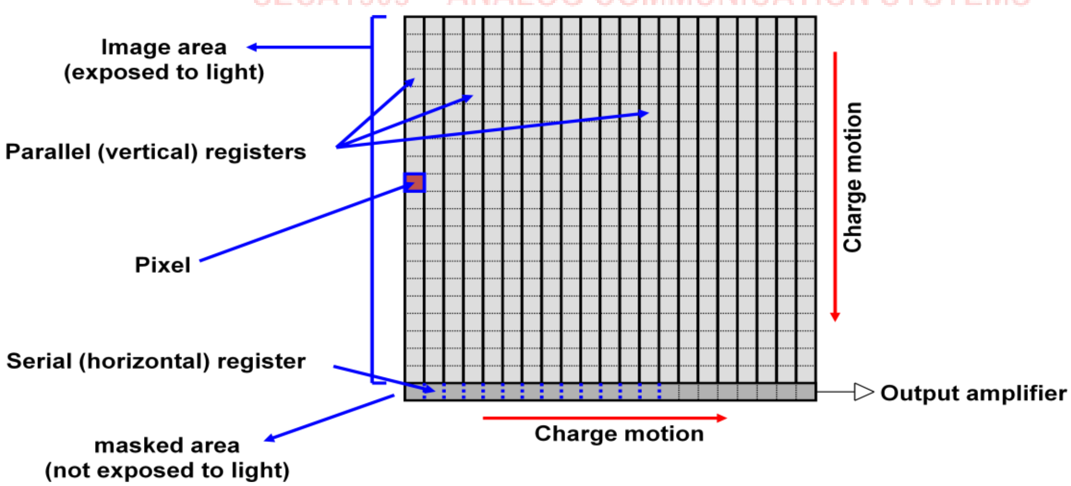
An analog shift register, that enables the transportation of analog signals (electric charges) through successive stages (pixels) controlled by a clock signal.



The CCD - comprised of many individual signal capture units (photo sites, capacitors, pixels)



A CCD chip is a metal oxide semiconductor (MOS) device. This means that its base, which is constructed of a material which is a good conductor under certain conditions, is topped with a layer of a metal oxide. In the case of the CCD, usually silicon is used as the base material and silicon dioxide is used as the coating. The final, top layer is also made of silicon – polysilicon.

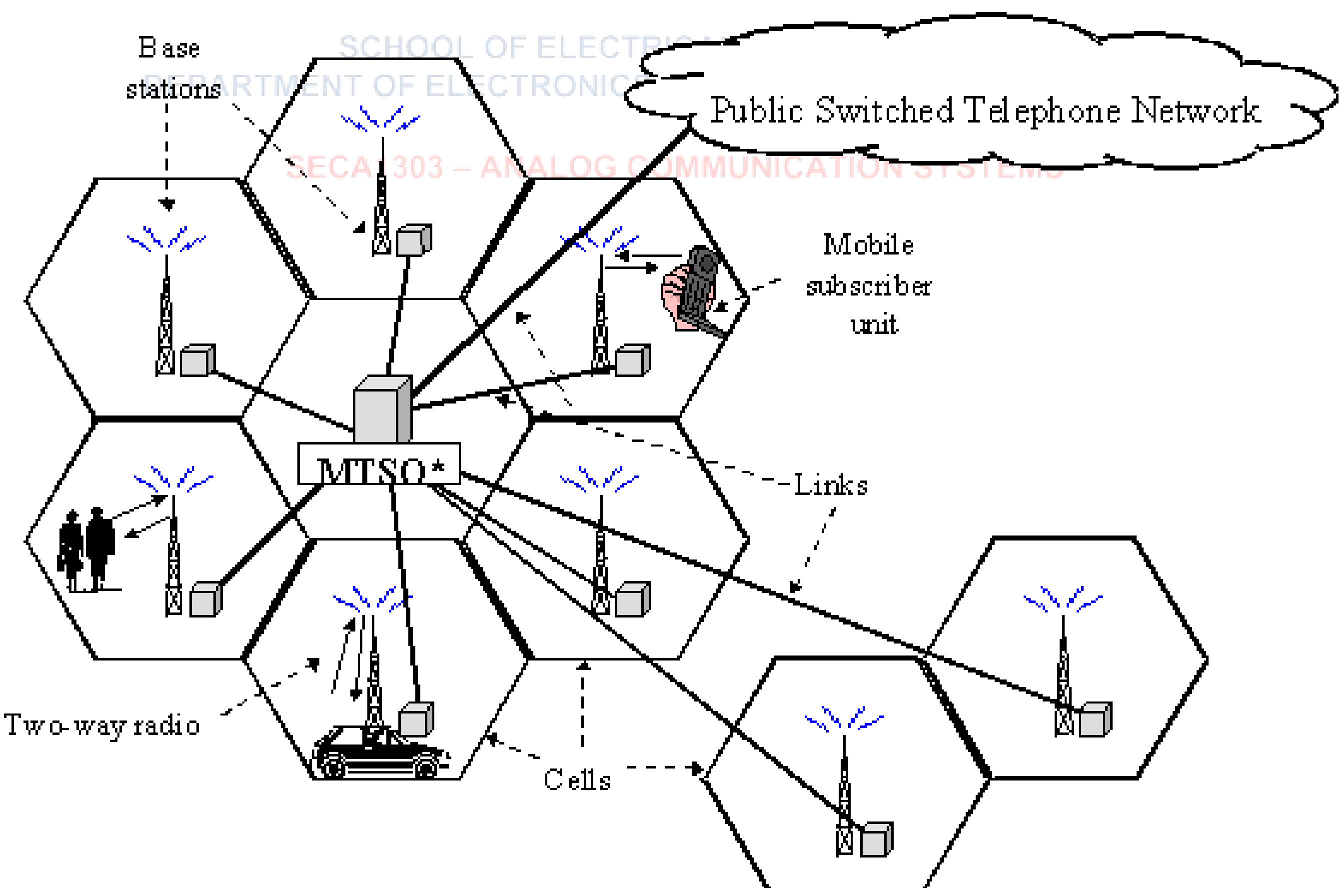


Fundamentally, a charge coupled device (CCD) is an integrated circuit etched onto a silicon surface forming light sensitive elements called pixels. Photons incident on this surface generate charge that can be read by electronics and turned into a digital copy of the light patterns falling on the device. CCDs come in a wide variety of sizes and types and are used in many applications from cell phone cameras to high-end scientific applications.

Primary uses of CCD:
Memory,
Delaying samples of analog signals
In an array of photoelectric light sensors (image sensors)
Digital Photography
Astronomy
Sensors
Electron Microscopy
Medical Fluoroscopy
Optical and UV Spectroscopy

Cellular Telephone system

In a cellular network, cells are generally organized in groups of seven to form a cluster. There is a ―cell site‖ or ‑ base station‖ at the centre of each cell, which houses the transmitter/receiver antennae and switching equipment. The size of a cell depends on the density of subscribers in an area: for instance, in a densely populated area, the capacity of the network can be improved by reducing the size of a cell or by adding more overlapping cells. This increases the number of channels available without increasing the actual number of frequencies being used. All base stations of each cell are connected to a central point, called the Mobile Switching Office (MSO), either by fixed lines or microwave. The MSO is generally connected to the PSTN (Public Switched Telephone Network):

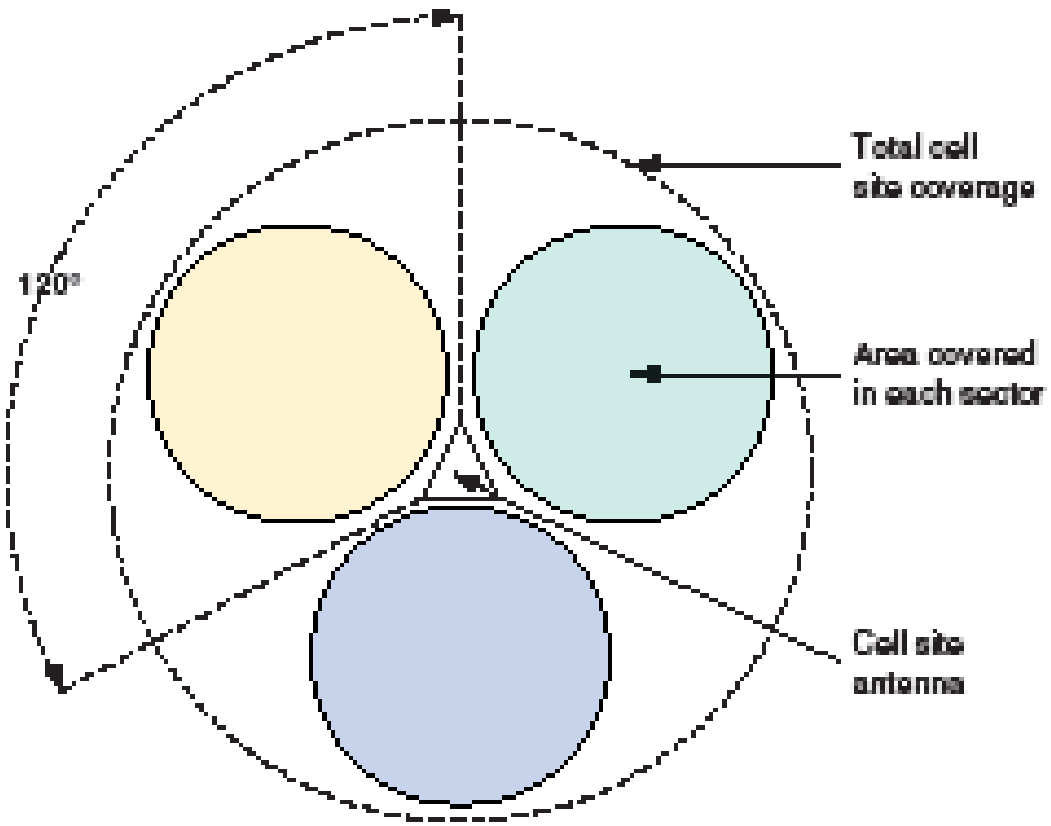


*MTSO= Mobile Telephone Switching Office

Cellular technology allows the - hand-offǁ of subscribers from one cell to another as they travel around. This is the key feature which allows the mobility of users. A computer constantly tracks mobile subscribers of units within a cell, and when a user reaches the border of a call, the computer automatically hands-off the call and the call is assigned a new channel in a different cell.

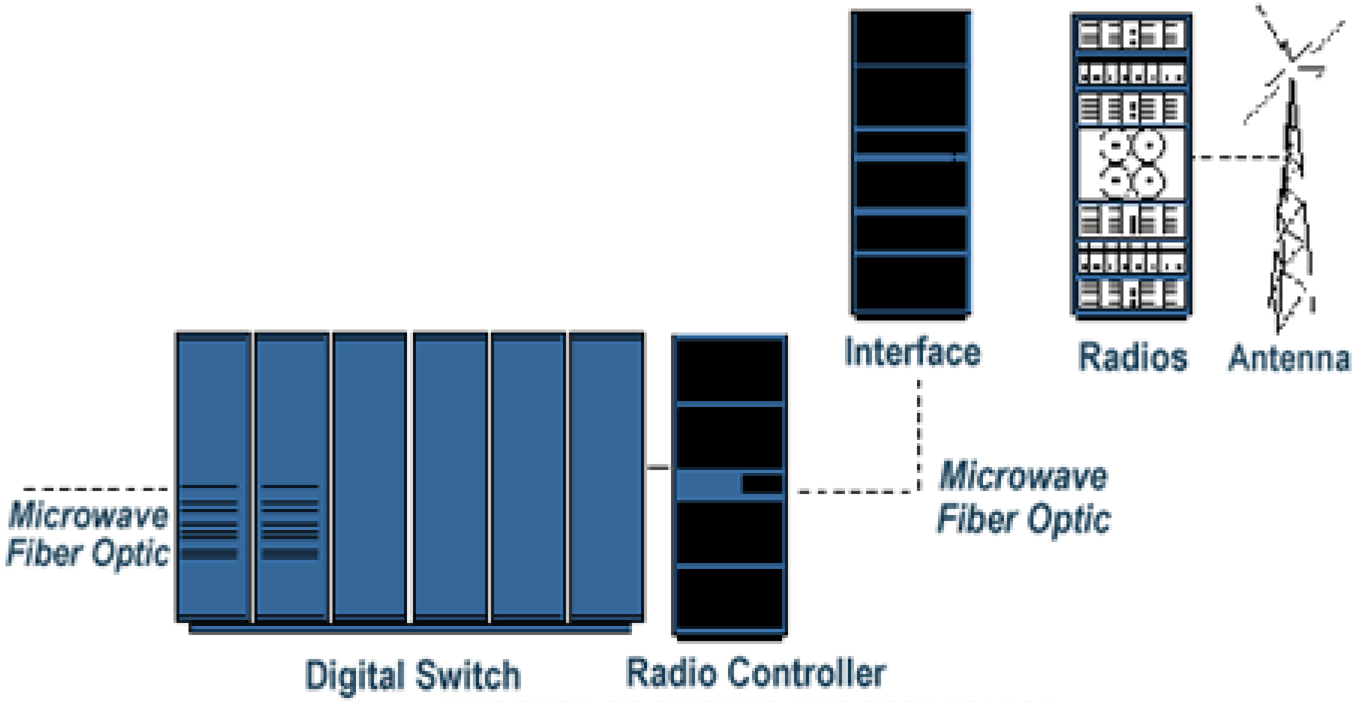
International roaming arrangements govern the subscriber‘s ability to make and receive calls the home network‘s coverage area.

Horizontal antenna radiation pattern of a common cell site showing 120° sectors that permit frequency reuse.



A cellular system comprises of the following basic components: Mobile Stations (MS): Mobile handsets, which is used by an user to communicate with another user Cell: Each cellular service area is divided into small regions called cell (5 to 20 Km). Base Stations (BS): Each cell contains an antenna, which is controlled by a small office. Mobile Switching Center (MSC): Each base station is controlled by a switching office, called mobile switching center.

Digital cellular Telephone system:



While the description of the analog telephone system provides an accurate overview of the principles of current telephone systems, it is a fact that most telephone calls today are really digital telephone calls.

In a digital telephone system, the two ends of the call are analog, and the middle section is digital. Conversions from analog to digital (A/D), and back to analog (D/A), are made in such a way that it is essentially impossible to determine that they were made at all.

Although the analog telephone system is gradually being converted to digital, the input and output of the system still remains analog because the eventual use is for humans that are able to process analog information .

At present, most telephone calls are analog from the telephone at home to the first switching office, so the A/D and D/A conversion is made at this office

In the future, as telephone systems become all digital, this conversion from A/D and from D/A will be made within the telephone set at home. The A/D conversion process was explained in the previous lectures- The voice signal- an analog waveform was sampled at a sampling frequency, and quantized to a number of levels. These values were then assigned binary codes to complete the conversion process from analog to digital

The D/A process was also explained briefly. The bits were decoded into their quantized values, and a waveform similar to the original analog waveform was obtained

For voice, remember that the standard sampling frequency is 8000Hz.

The standard number of quantization levels for audio signals is 256, requiring 8 bits.

So, the bit rate for a digital telephone call is: 8,000x8=64,000 bits per second (64 Kbps)

This is the bit rate that would reach the central office if the A/D conversion was being done inside the telephone at home Since many calls arrive at the central office, they can all be combined, and switched to another center to be routed to the destination Combining many channels and sending them simultaneously through a single transmission line is called multiplexing.

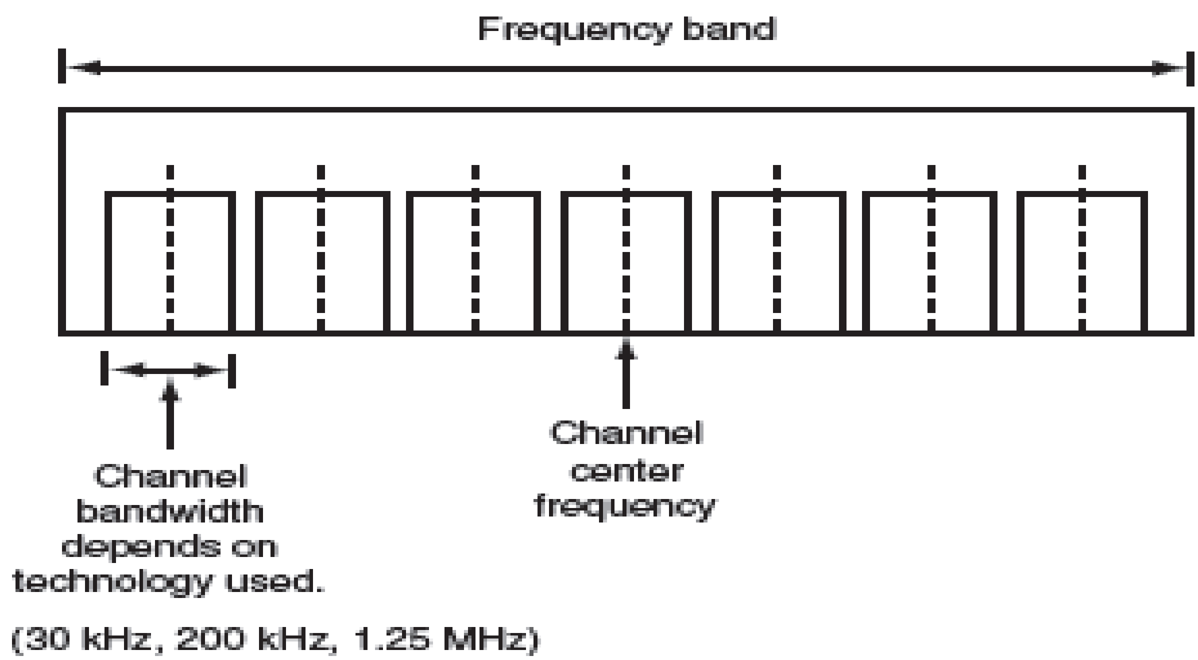
Multiple Access (optional)

Multiple access refers to how the subscribers are allocated to the assigned frequency spectrum. Access methods are the ways in which many users share a limited amount of spectrum.

The techniques include frequency reuse, frequency-division multiple access (FDMA), time-division multiple access (TDMA), code-division multiple access (CDMA), and spatial-division multiple access (SDMA).

Frequency Reuse: In frequency reuse, individual frequency bands are shared by multiple base stations and users. This is possible by ensuring that one subscriber or base station does not interfere with any others. This is achieved by controlling such factors as transmission power, base station spacing, and antenna height and radiation patterns. With low-power and lower-height antennas, the range of a signal is restricted to only a mile or so. Furthermore, most base stations use sectorized antennas with 1208 radiation patterns that transmit and receive over only a portion of the area they cover.In any given city, the same frequencies are used over and over simply by keeping cell site base stations isolated from one another.

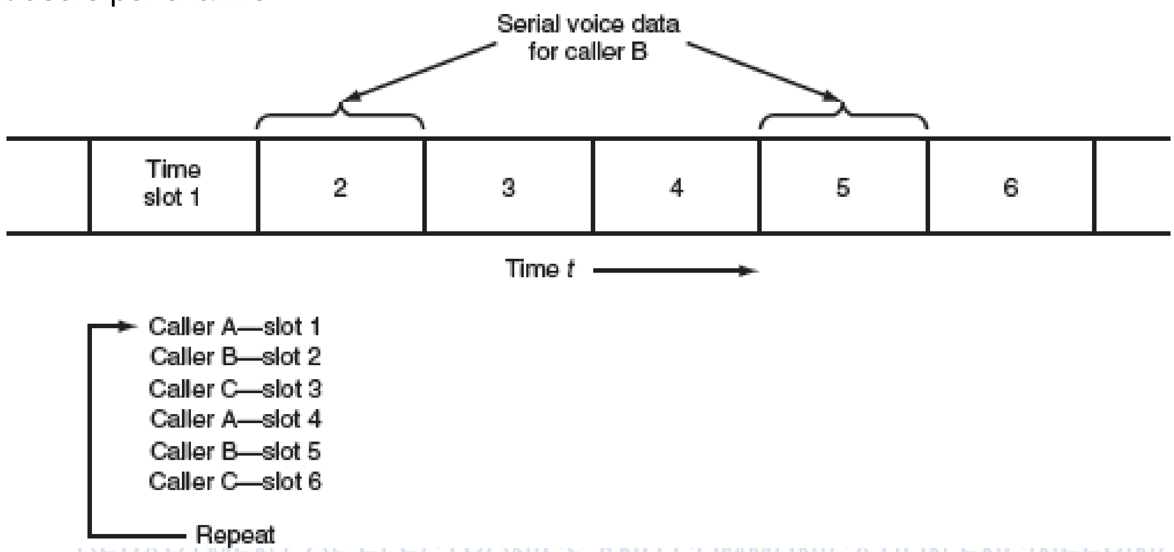
Frequency-division multiple-access (FDMA) spectrum.



Frequency-Division Multiple Access.

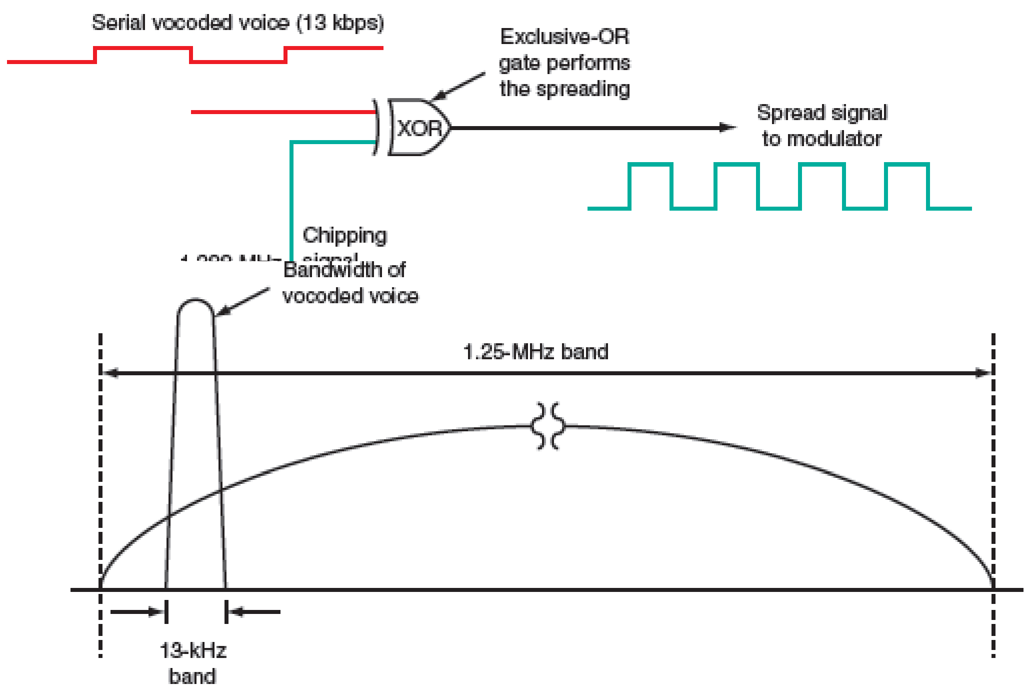
FDMA systems are like frequency- division multiplexing in that they allow many users to share a block of spectrum by simply dividing it up into many smaller channels. Each channel of a band is given an assigned number or is designated by the center frequency of the channel. One subscriber is assigned to each channel. Typical channel widths are 30 kHz, 200 kHz, 1.25 MHz, and 5 MHz. There are usually two similar bands, one for uplink and the other for downlink.

Time-Division Multiple Access: TDMA relies on digital signals and operates on a single channel. Multiple users use different time slots. Because the audio signal is sampled at a rapid rate, the data words can be interleaved into different time slots, Of the two common TDMA systems in use, one allows three users per frequency channel and the other allows eight users per channel.



Code-Division Multiple Access:

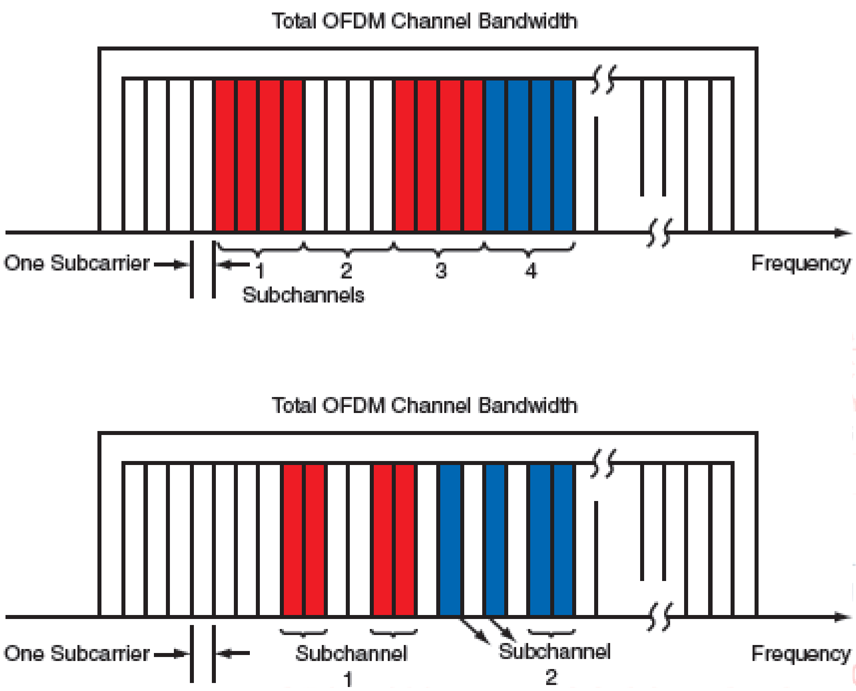
Code-division multiple access (CDMA). (*a*) Spreading the signal. (*b*) Resulting bandwidth.



CDMA is just another name for spread spectrum. A high percentage of cell phone systems use *direct sequence spread spectrum (DSSS)*. Here the digital audio signals are encoded in a circuit called a vocoder to produce a 13-kbps serial digital compressed voice signal. It is then combined with a higher-frequency chipping signal. One system uses a 1.288-Mbps chipping signal to encode the audio, spreading the signal over a 1.25-MHz channel. See Fig. 20-8. With unique coding, up to 64 subscribers can share a 1.25-MHz channel. A similar technique is used with the wideband CDMA system of third-generation cellphones. A 3.84-Mbps chipping rate is used in a 5-MHz channel to accommodate multiple users.

Orthogonal Frequency Division Multiplexing Access (OFDMA).

The Concept of OFDMA.



OFDMA is the access method used with OFDM. OFDM uses hundreds, even thousands, of subcarriers in a wideband channel. This large number of subcarriers can be subdivided Into smaller groups, and each group can be assigned to an individual user. In this way, many users can use the wideband channel assigned to the OFDM signal.

TELEVISION

The aim of a television system is to extend the sense of sight beyond its natural limits and to transmit sound associated with the scene. The picture signal is generated by a TV camera and sound signal by a microphone. In the 625 lines CCIR monochrome and PAL-B color TV systems adopted by India, the picture signal is amplitude modulated and sound signal frequency modulated before transmission. The two carrier frequencies are suitably spaced and their modulation products radiated through a common antenna. As in radio communication, each television station is allotted different carrier frequencies to enable selection of desired station at the receiving end. The TV receiver has tuned circuits in its input section called 'tuner'. It selects desired channel signal out of the many picked up by the antenna. The selected RF band is converted to a common fixed IF band for convenience of providing large amplification to it. The amplified IF signals are detected to obtain video (picture) and audio (sound) signals. The video signal after large amplification drives the picture tube to reconstruct the televised picture on the receiver screen. Similarly, the audio signal is amplified and fed to the loudspeaker to produce sound output associated with the scene.

PICTURE TRANSMISSION

The picture information is an optical in character and may be thought of as an assemblage of a large number of tiny areas representing picture details. These elementary areas into which picture details may be broken up are known as 'picture elements' or 'pixels', which when viewed together represent visual information of the scene. Thus, at any instant there are almost an infinite number of pieces of information that need to be picked up simultaneously for transmitting picture details. However, simultaneous pick-up is not practicable because it is not feasible to provide a separate signal path (channel) for the signal obtained from each picture element. In practice, this problem is solved by a method known as 'scanning' where conversion of optical information to electrical form is carried out element by element, one at a time and in a sequential manner to cover the entire picture. Besides, scanning is done at a very fast rate and repeated a large number of times per second to create an illusion (impression at the eye) of simultaneous reception from all the elements, though using only one signal path. Black and White Pictures In a monochrome (black and white) picture, each element is either bright, some shade of grey or dark.

A TV camera, the heart of which is a camera tube, is used to convert this optical information into corresponding electrical signal, the amplitude of which varies in accordance with variations of brightness. Fig. 3.1 shows very elementary details of one type of camera tube (vidicon) and associated components to illustrate the principle. An optical image of the scene to be transmitted is focused by a lens assembly on the rectangular glass face-plate of the camera tube. The inner side of the glass face-plate has a transparent conductive coating on which is laid a very thin layer of

Photoconductive material. The photo layer has very high resistance when no light falls on it, but decreases depending on the intensity of light falling on it. Thus, depending on light intensity variations in the focused optical image, the conductivity of each element of photo layer changes accordingly. An electron beam is used to pick-up picture information now available on the target plate in terms of varying resistance at each point.

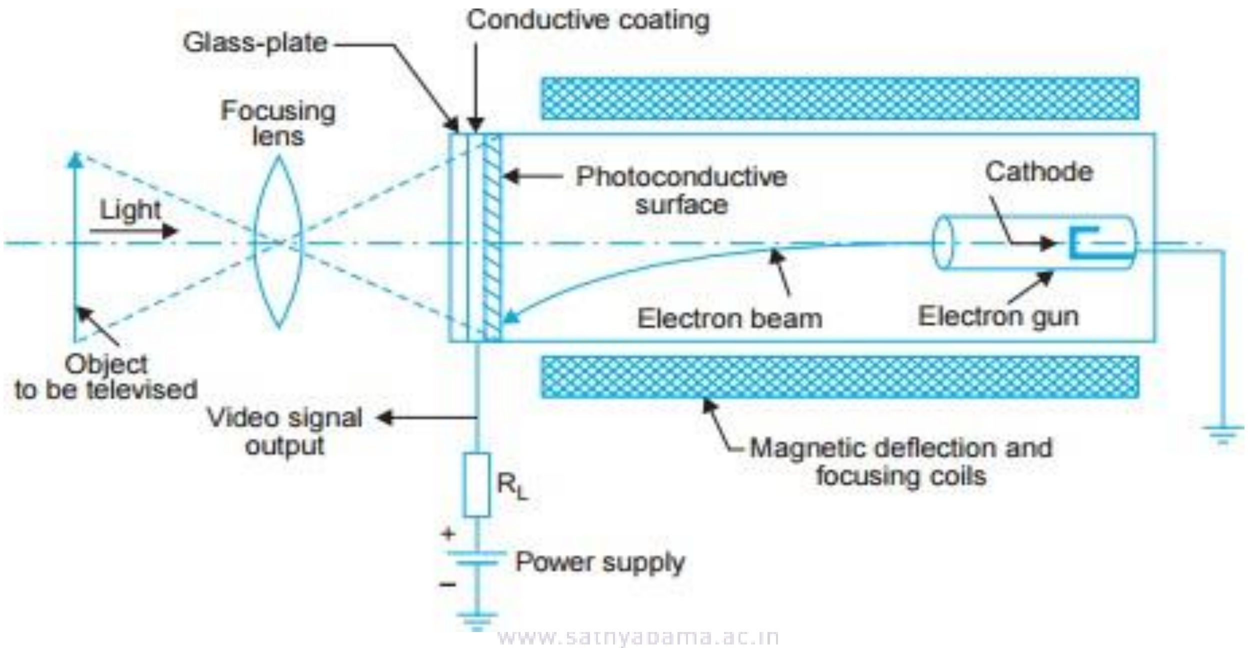


Fig. 3.1. Simplified cross-sectional view of a (Vidicon) camera tube and associated components.

The beam is formed by an electron gun in the TV camera tube. On its way to the inner side of glass face-plate, it is deflected by a pair of deflecting coils mounted on the glass envelope and kept mutually perpendicular to each other to achieve scanning of the entire target area. Scanning is done in the same way as one reads a written page to cover all the words in one line and all the lines on the page (see Fig. 3.2). To achieve this, the deflecting coils are fed separately from two sweep oscillators which continuously generate suitable waveform voltages, each operating at a different desired frequency. Magnetic deflection caused by the current in one coil gives horizontal motion to the beam from left to right at uniform rate and then brings it quickly to the left side to commence trace of the next line. The other coil is used to deflect the beam from top to bottom at a uniform rate and for its quick retrace back to the top of the plate to start this process over again. Two simultaneous motions are thus given to the beam, one from left to right across the target plate and the other from top to bottom thereby covering entire area on which electrical image of the picture is available. As the beam moves from element to element, it encounters a different resistance across the target-plate, depending on the resistance of photoconductive coating. The result is a flow of current which varies in magnitude as the elements are scanned. This current passes through a load resistance RL, connected to the conductive coating on one side and to a dc supply source on the other. Depending on the magnitude of current, a varying voltage appears across resistance RL and this corresponds to optical information of the picture.

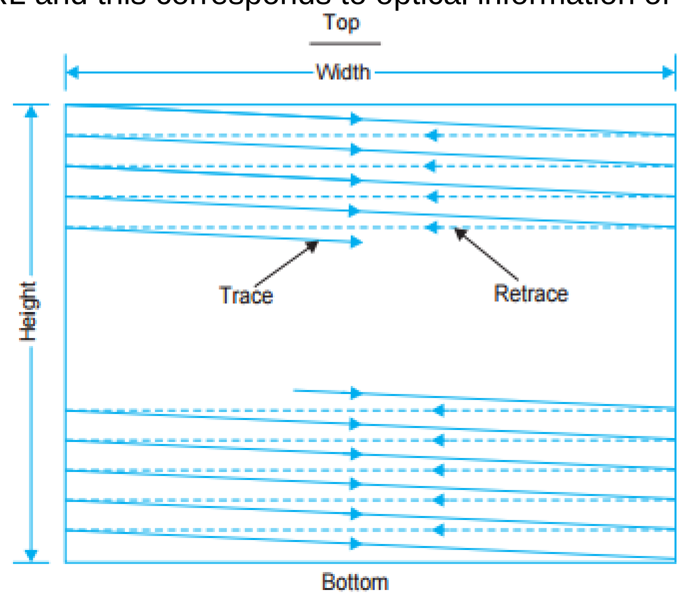


Fig. 3.2. Path of scanning beam in covering picture area.

If the scanning beam moves at such a rate that any portion of the scene content does not have time to change perceptibly in the time required for one complete scan of the image, the resultant electrical signal contains true information existing in the picture during the time of scan. The desired information is now in the form of a signal varying with time and scanning may thus be identified as a particular process which permits conversion of information existing in space and time co-ordinates into time variations only. The electrical information thus obtained from the TV camera tube is generally referred to as video signal (video is Latin for 'see').

COLOUR PICTURES

It is possible to create any color including white by additive mixing of red, green and blue color lights in suitable proportions. For example, yellow can be obtained by mixing red and green color lights in intensity ratio of 30 : 59. Similarly, light reflected from any color picture element can be synthesized (broken up) into red, green and blue color light constituents. This forms the basis of color television where Red (R), Green (G) and Blue (B) colors are called primary colors and those formed by mixing any two of the three primaries as complementary colors. A color camera, the elements of which are shown in Fig. 3.3, is used to develop signal voltages proportional to the intensity of each primary color light.

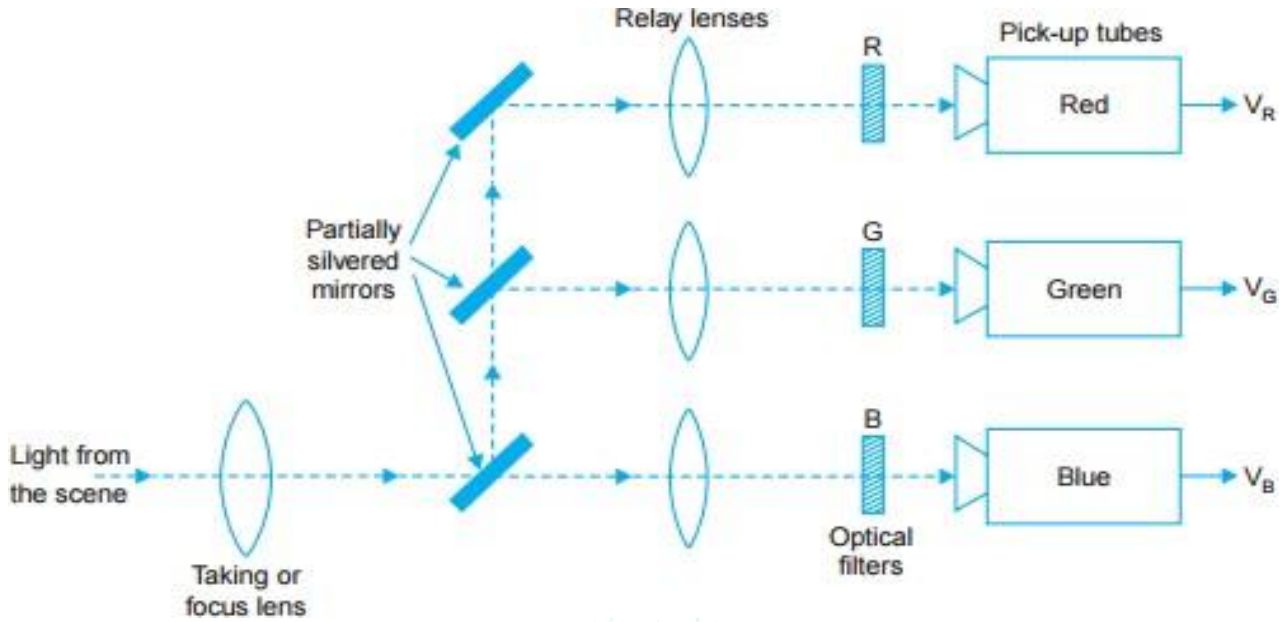


Fig. 3.3. Simplified block diagram of a color camera

It contains three camera tubes (vidicons) where each pick-up tube receives light of only one primary color. Light from the scene falls on the focus lens and pass through that on special mirrors. Color filters that receive reflected light via relay lenses split it into R, G and B color lights. Thus, each vidicon receives a single color light and develops a voltage proportional to the intensity of one of the primary colors. If any primary color is not present in any part of the picture, the corresponding vidicon does not develop any output when that picture area is scanned. The electron beams of all the three camera tubes are kept in step (synchronism) by deflecting them horizontally and vertically from common driving sources. Any color light has a certain intensity of brightness. Therefore, light reflected from any color element of a picture also carries information about its brightness called luminance. A signal voltage (Y) proportional to luminance at various parts of the picture is obtained by adding definite proportions of V_R, V_G and V_B (30:59:11). This then is the same as would be developed by a monochrome (black and white) camera when made to scan the same color scene. This i.e., the luminance (Y) signal is also transmitted along with color information and used at picture tube in the receiver for reconstructing the color picture with brightness levels as in the televised picture.

TELEVISION TRANSMITTER

An oversimplified block diagram of a monochrome TV transmitter is shown in Fig. 3.4. The luminance signal from the camera is amplified and synchronizing pulses added before feeding it to the modulating amplifier. Synchronizing pulses are transmitted to keep the camera and picture tube beams in step. The allotted picture carrier frequency is generated by a crystal controlled oscillator. The continuous wave (CW) sine wave output is given large amplification before feeding to the power amplifier where its amplitude is made to vary (AM) in accordance with the modulating signal received from the modulating amplifier. The modulated output is combined (see Fig. 3.4) with the frequency modulated (FM) sound signal in the combining network and then fed to the transmitting antenna for radiation.

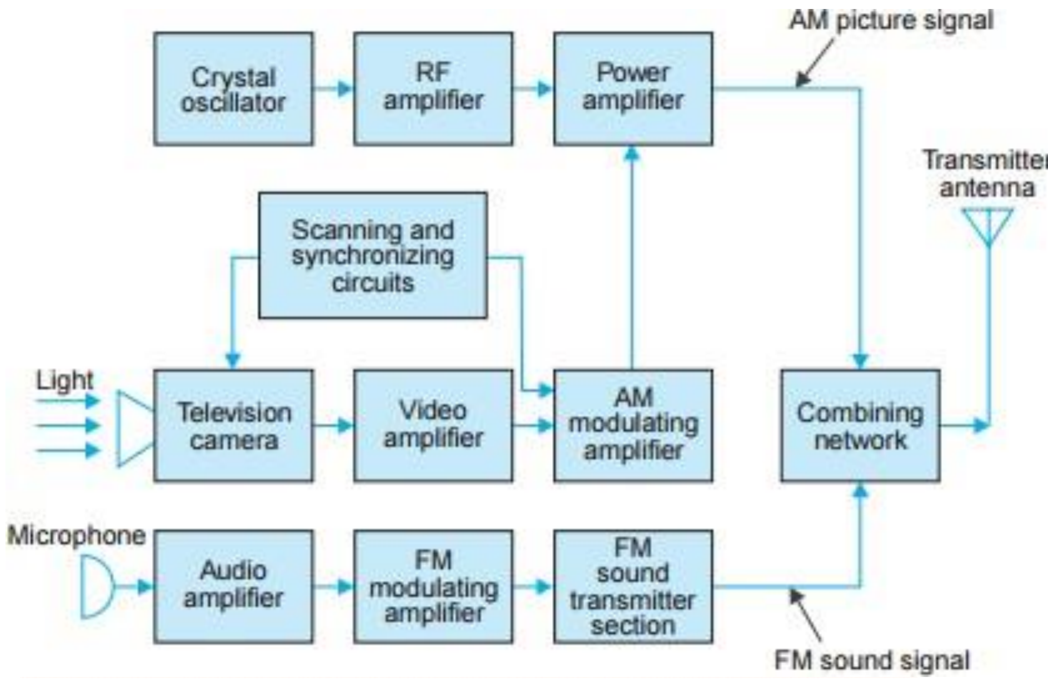
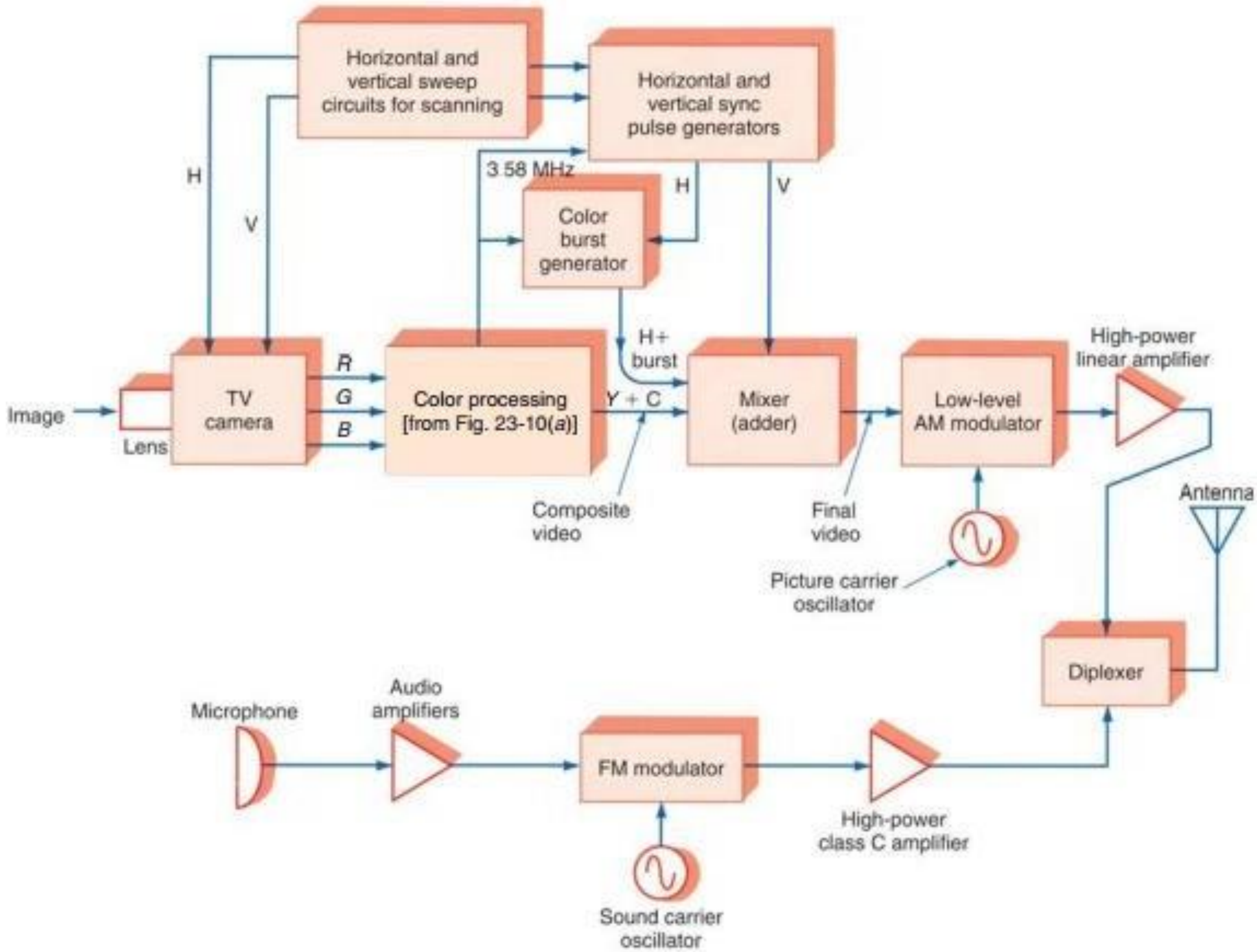


Fig. 3.4. Elementary block diagram of a monochrome television transmitter.

COLOUR TRANSMITTER

A color TV transmitter is essentially the same as the monochrome transmitter except for the additional need that color (Chroma) information is also to be transmitted. Any color system is made compatible with the corresponding monochrome system. Compatibility means that the color TV signal must produce a normal black and white picture on a monochrome receiver and a color receiver must be able to produce a normal black and white picture from a monochrome TV signal. For this, the luminance (brightness) signal is transmitted in a color system in the same way as in the monochrome system and with the same bandwidth. However, to ensure compatibility, the color camera outputs are modified to obtain (B-Y) and (R-Y) signals. These are modulated on the color sub-carrier, the value of which is so chosen that on combining with the luminance signal, the sidebands of the two do not interfere with each other i.e., the luminance and color signals are correctly interleaved. A color sync signal called 'color burst' is also transmitted for correct reproduction of colors.



Block diagram of TV transmitter is shown in the above figure. Note the Sweep and sync circuits that create the scanning signals for vidicon camera tube or CCDs as well as generate the sync signals that are transmitted along the video and color signals. The sync signals, luminance Y, and color signals are added to form the final video signal that is used to modulate the carrier. Low- level AM is used. The final AM signal is amplified by very high- power linear amplifiers and sent to the antenna via diplexer, which is a set of sharp band pass filters that pass the transmitter signal to the antenna but prevent signals from getting back into the sound transmitter.

At the same time, the voice or sound signals frequency - modulated carrier that is amplified by Class C amplifiers and fed to the same antenna by way of the diplexer. The resulting VHF or UHF signal travels by line-of-Sight propagation to the antenna and receiver

SOUND TRANSMISSION

There is no difference in sound transmission between monochrome and colour TV systems. The microphone converts the sound associated with the picture being televised into proportionate electrical signal, which is normally a voltage. This electrical output, regardless of the complexity of its waveform, is a single valued function of time and so needs a single channel for its transmission. The audio signal from the microphone after amplification is frequency modulated, employing the assigned carrier frequency. In FM, the amplitude of carrier signal is held constant, whereas its frequency is varied in accordance with amplitude variations of the modulating signal. As shown in Fig. 3.4, output of the sound FM transmitter is finally combined with the AM picture transmitter output, through a combining network, and fed to a common antenna for radiation of energy in the form of electromagnetic waves.

TELEVISION RECEIVER

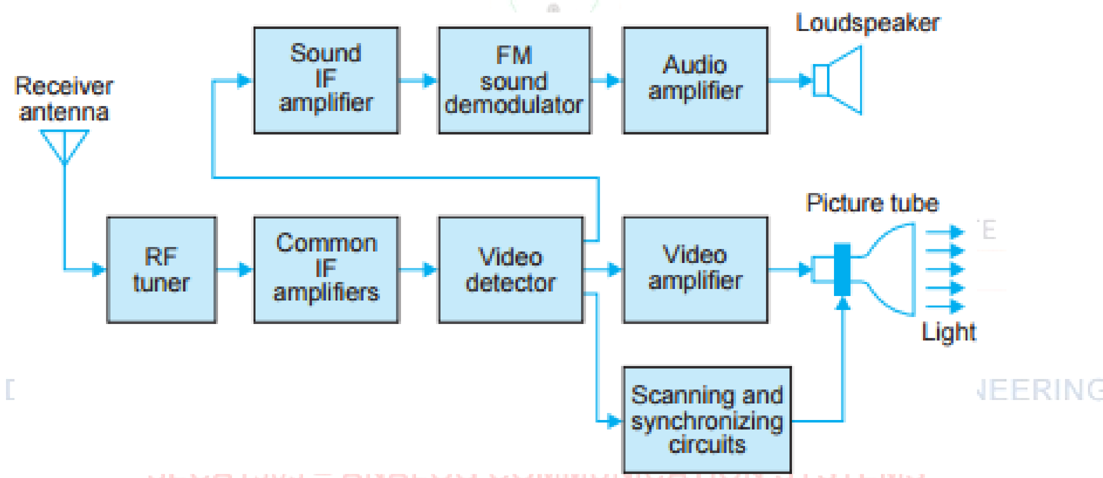


Fig. 3.5. Simplified block diagram of a black and white TV receiver

A simplified block diagram of a black and white TV receiver is shown in Fig. 3.5. The receiving antenna intercepts radiated RF signals and the tuner selects desired channel's frequency band and converts it to the common IF band of frequencies. The receiver employs two or three stages of intermediate frequency (IF) amplifiers. The output from the last IF stage is demodulated to recover the video signal. This signal that carries picture information is amplified and coupled to the picture tube which converts the electrical signal back into picture elements of the same degree of black and white.

The picture tube shown in Fig. 3.6 is very similar to the cathode-ray tube used in an oscilloscope. The glass envelope contains an electron-gun structure that produces a beam of electrons aimed at the fluorescent screen. When the electron beam strikes the screen, light is emitted. The beam is deflected by a pair of deflecting coils mounted on the neck of picture tube in the same way as the beam of camera tube scans the target plate. The amplitudes of currents in the horizontal and vertical deflecting coils are so adjusted that the entire screen, called raster, gets illuminated because of the fast rate of scanning.

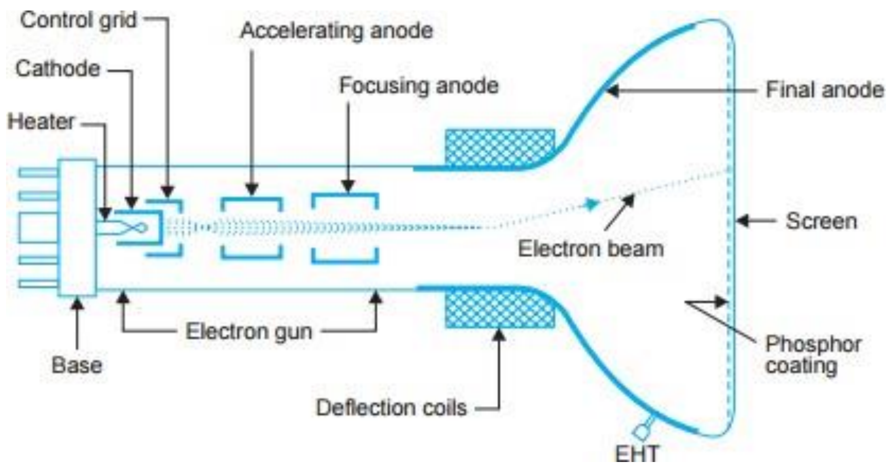


Fig. 3.6. Elements of a picture tube

The video signal is fed to the grid or cathode of picture tube. When the varying signal voltage makes the control grid less negative, the beam current is increased, making the spot of light on the screen brighter. More negative grid voltage reduces brightness. If the grid voltage is negative enough to cut-off the electron beam current at the picture tube, there will be no light. This state corresponds to black. Thus the video signal illuminates the fluorescent screen from white to black through various shades of grey depending on its amplitude at any instant. This corresponds to brightness changes encountered by the electron beam of the camera tube while scanning picture details element by element. The rate at which the spot of light moving very fast that the eye is not able to follow it and so a complete picture is seen because of storage capability of the human eye.

SOUND RECEPTION

The path of sound signal is common with the picture signal from antenna to video detector section of the receiver. Here the two signals are separated and fed to their respective channels. The frequency modulated audio signal is demodulated after at least one stage of amplification. The audio output from the FM detector is given due amplification before feeding it to the loudspeaker.

COLOUR RECEIVER

A color receiver is similar to the black and white receiver as shown in Fig. 3.7. The main difference between the two is the need of a color or Chroma subsystem. It accepts only the colour signal and processes it to recover (B-Y) and (R-Y) signals. These are combined with the Y signal to obtain VR, VG and VB signals as developed by the camera at the transmitting end. VG becomes available as it is contained in the Y signal. The three color signals are fed after sufficient amplification to the colour picture tube to produce a color picture on its screen.

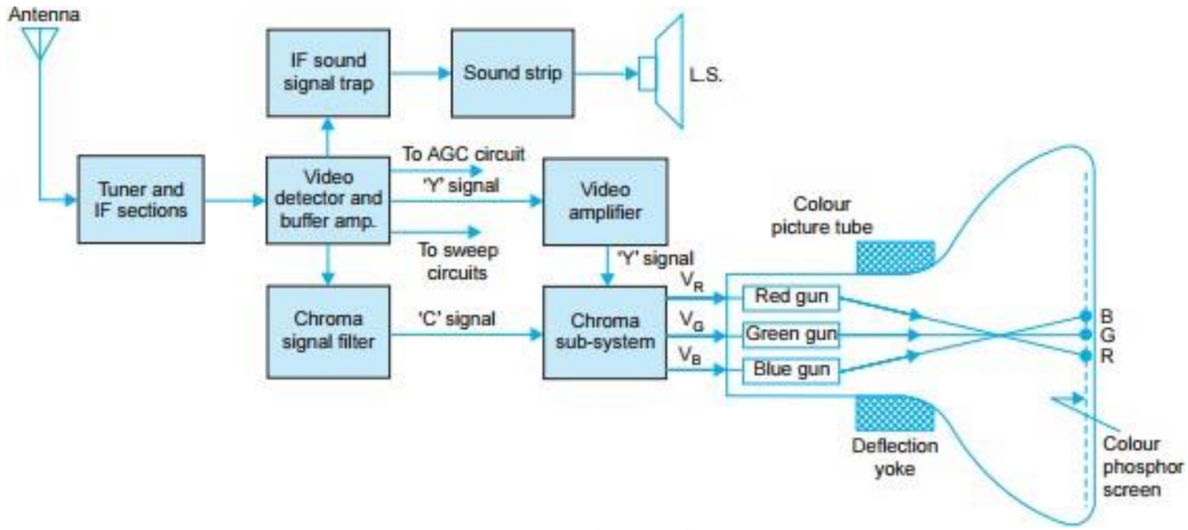


Fig. 3.7. An oversimplified block diagram of a colour receiver.

As shown in Fig. 3.7, the color picture tube has three guns corresponding to the three pick-up tubes in the color camera. The screen of this tube has red, green and blue phosphors arranged in alternate stripes. Each gun produces an electron beam to illuminate corresponding color phosphor separately on the fluorescent screen. The eye then integrates the red, green and blue color information and their luminance to perceive actual color and brightness of the picture being televised. The sound signal is decoded in the same way as in a monochrome receiver.

SYNCHRONIZATION

It is essential that the same co-ordinates be scanned at any instant both at the camera tube target plate and at the raster of picture tube, otherwise, the picture details would split and get distorted. To ensure perfect synchronization between the scenes being televised and the picture produced on the raster, synchronizing pulses are transmitted during the retrace, i.e., fly-back intervals of horizontal and vertical motions of the camera scanning beam. Thus, in addition to carrying picture details, the radiated signal at the transmitter also contains synchronizing pulses. These pulses which are distinct for horizontal and vertical motion control are processed at the receiver and fed to the picture tube sweep circuitry thus ensuring that the receiver picture tube beam is in step with the transmitter camera tube beam.

As stated earlier, in a color TV system additional sync pulses called color burst are transmitted along with horizontal sync pulses. These are separated at the input of Chroma section and used to synchronize the color demodulator carrier generator. This ensures correct reproduction of colors in the otherwise black and white picture.

RECEIVER CONTROLS

Most black and white receivers have on their front panel (i) channel selector, (ii) fine tuning, (iii) brightness, (iv) contrast, (v) horizontal hold and (vi) volume controls besides an ON-OFF switch. Some receivers also provide a tone control. The channel selector switch is used for selecting the desired channel. The fine tuning control is provided for obtaining best picture details in the selected channel. The hold control is used to get a steady picture in case it rolls up or down. The brightness control varies beam intensity of the picture tube and is set for optimum average brightness of the picture. The contrast control is actually gain control of the video amplifier. This can be varied to obtain desired contrast between white and black contents of the reproduced picture. The volume and tone controls form part of the audio amplifier in sound section, and are used for setting volume and tonal quality of the sound output from the loudspeaker.

In color receivers there is an additional control called ‘color’ or ‘saturation’ control. It is used to vary intensity or amount of colors in the reproduced picture. In modern color receivers that employ integrated circuits in most sections of the receiver, the hold control is not necessary and hence usually not provided.

INTRODUCTION: CCTV SYSTEMS

CCTV systems provide surveillance capabilities used in the protection of people, assets, and systems. A CCTV system serves mainly as a security force multiplier, providing surveillance for a larger area, more of the time, than would be feasible with security personnel alone. CCTV systems are often used to support comprehensive security systems by incorporating video coverage and security alarms for barriers, intrusion detection, and access control. For example, a CCTV system can provide the means to assess an alarm generated by an intrusion detection system and record the event. A CCTV system links a camera to a video monitor using a direct transmission system. This differs from broadcast television where the signal is transmitted over the air and viewed with a television. New approaches within the CCTV industry are moving towards more open architecture and transmission methods versus the closed circuit, hard-wired connection systems of the past. CCTV systems have many components with a variety of functions, features, and specifications. Key components include cameras, lenses, data distribution, power, and lighting, among others. CCTV technologies continuously undergo feature refinements to improve performance in areas such as digital equipment options, data storage, component miniaturization, wireless communications, and automated image analysis. The components, configuration options, and features available in today's CCTV market create a complex set of purchasing options. It is the intent of this handbook to provide information on the capabilities and limitations of CCTV components that will aid an agency procuring a new CCTV system or upgrading an existing one.

CCTV (closed-circuit television) is a TV system in which signals are not publicly distributed but are monitored, primarily for surveillance and security purposes.

CCTV relies on strategic placement of cameras, and observation of the camera's input on monitors somewhere. Because the cameras communicate with monitors and/or video recorders across private coaxial cable runs or wireless communication links, they gain the designation "closed-circuit" to indicate that access to their content is limited by design only to those able to see it.

Older CCTV systems used small, low-resolution black and white monitors with no interactive capabilities. Modern CCTV displays can be color, high-resolution displays and can include the ability to zoom in on an image or track something (or someone) among their features. Talk CCTV allows an overseer to speak to people within range of the camera's associated speakers.

COMPONENTS OF CCTV SYSTEMS:

CCTV uses components that are directly connected to generate, transmit, display, and store video data. A CCTV system can be as simple as a camera purchased from a retail electronics store connected to a video monitor. However, larger systems operated by professional security personnel are comprised of a number of components falling into several basic categories:• Cameras; • Lenses; • Housings and mounts; • Monitors; • Switchers and multiplexers; and • Video recorders. Many features exist within each of these categories that can satisfy an agency's operational requirements in the most challenging environments. The most complex CCTV systems may incorporate hundreds of cameras and sensors integrated into one overall security network. Figure 3-1 provides a CCTV component diagram example

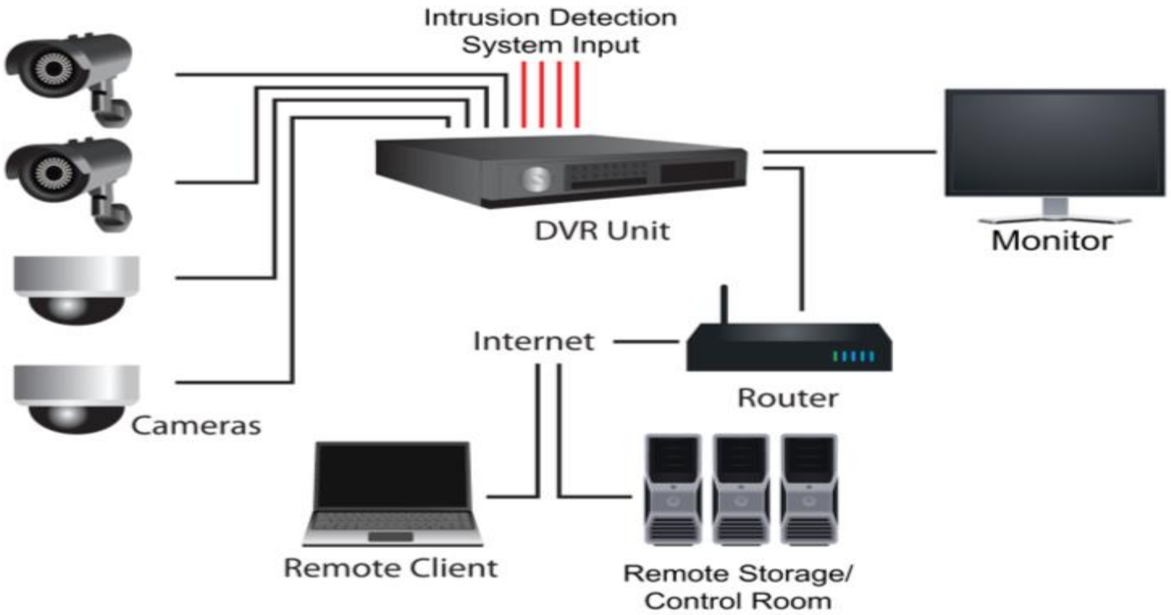


Figure 3-1. CCTV Component Diagram Example

Closed-Circuit TeleVision is a special application in which camera signals are made available only to a limited number of monitors or receivers. The particular type of link used depends on the distance between the two locations, the number and dispersion of receivers and mobility of either camera or receiver. The figure illustrates various link arrangements which are often used. The simplest link is a cable where video signal from the camera is connected directly through a cable to the receiver. A Television monitor, which is a receiver, without RF and IF circuits, is only required for reception in such a link arrangement. About one volt peak to peak signal is secured by the monitor.

Since the video signal is normally delivered via cables and even when transmitted, it is over the limit region and for restricted use, CCTV need not follow television broadcast standards.

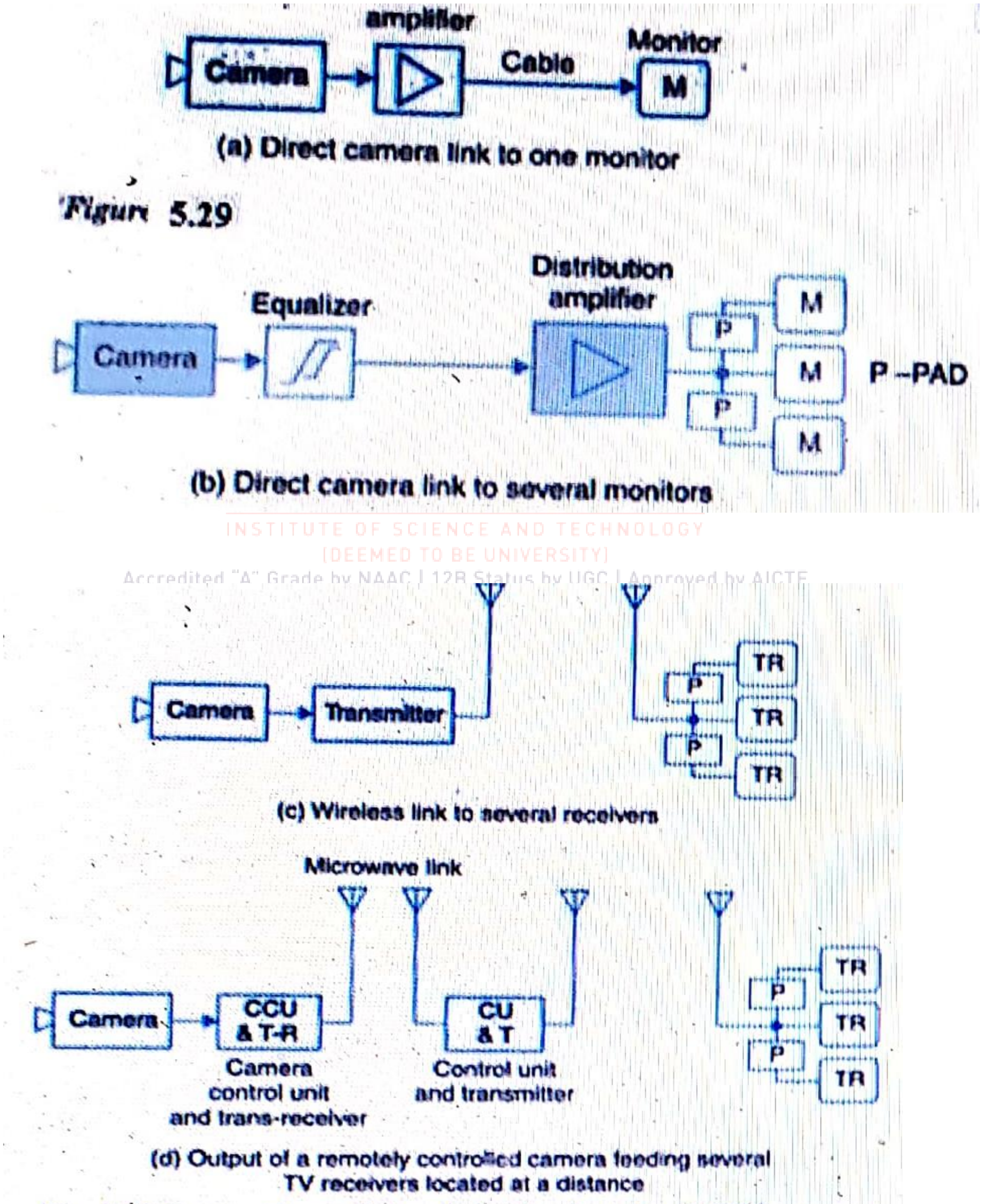


Fig. 5.29 Commonly used closed circuit television (CCTV) systems

CCTV Component Diagram Example: Most new CCTV systems maximize the advantages of digital technologies by utilizing electronic databases, compact components, and wireless transmission techniques. With larger quantities of data being collected, it is essential that the system be capable of retaining data in accordance with the organization's policies and procedures.

Cameras: Cameras are an essential component of any CCTV system. Matching the right CCTV camera to a particular application is increasingly complex due to rapid technological developments and a greater range of applications.

Effective camera selection requires detailed knowledge of the camera, application, supporting architecture, and host environment.

All CCTV cameras include three basic elements:

• **Image sensor**–Converts light (photons) into electronic signals;

• **Lens–**Gathers light reflected from a subject and focuses the light on the image sensor; and

• **Image processing circuitry**–Organizes, optimizes, and transmits video signals. The type of camera best suited for a CCTV system depends on the operational environment and how it will integrate into the system.

The answers to the following questions may help determine the best camera type:

• What is the desired image quality?

• What size is the desired field of view (FOV)?

• How much lighting is available?

• Will the camera be installed indoors or outdoors?

• Will the video be monitored on a full time basis?

• How will the video be transmitted? • Will the camera be exposed to extreme conditions?

There are many types of **cameras** designed to perform under specific environmental conditions but cameras can be grouped into two primary categories: fixed and pan-tilt-zoom (PTZ). Fixed cameras are intended to constantly view a single scene, while PTZ cameras are motor driven and can pan left or right, tilt up or down, and zoom in or out to instantly customize the view as needed. A combination of fixed and PTZ cameras are often used to provide the required surveillance coverage.

Lenses: The lens on a CCTV camera is the first element in the imaging chain, which consists of the lens, camera, transmission system, image management and analysis software, and monitor. The lens focuses light or IR energy onto the imaging sensor. A lens's role is to deliver an undistorted, evenly focused, accurate image to the imaging sensor. Systems that require superior quality images start with lenses engineered to produce a high-quality image for the imaging sensor. Other components of the imaging chain cannot compensate for an inferior lens.

Figure. Representative CCTV Lens

Variables to consider when selecting a lens include the distance required to clearly focus on objects, FOV, size of the camera's image sensor, and lighting conditions. Lenses are identified by their focal length, usually stated in millimeters; largest aperture, usually stated as an f-number; and the size of the image sensor for which it was designed.

Types of Lenses: Lenses are available in three basic types: fixed focal length, varifocal (variable focal length), and zoom. The focal length of a lens is the distance between the optical center of the lens and the image plane. The lens focal length and the image sensor size determine the camera's FOV.

Housing and Mounts: Part of designing and installing an effective CCTV system includes selection of the camera housings and mounts. Selecting CCTV housing and mounting hardware is directly related to the operational system requirements, which are developed during the design and procurement phases of a CCTV installation project. In any application, the housing and mounting hardware is selected on the basis of several criteria: • Environmental conditions, which include operating temperatures and weather conditions, such as humidity and corrosion; • Architectural considerations, which are important to the aesthetics of the hardware and can affect the architectural design or change the value of the property; and • Installation and other special considerations that match the installed materials to the system's intended use and planned maintenance. The following hardware and mounting options are briefly described for comparison with system requirements.

Video Monitors: The function of monitors is to display video images for viewing. The selection of monitors is as important to the quality of the image as the selection of cameras, lenses, and other components in the imaging chain. The video monitor market offers a number of choices, such as liquid crystal displays (LCDs) and LED displays, various sizes, and other features. The requirements of the system will determine the type of monitor for each application. This section details some of the many features and considerations for monitor selection.

Table 3-3. CCTV Monitor Technology Comparisons

Type	Pros	Cons*
CRT	• Good overall picture quality • Robust technology • Low-cost option that is compatible with most existing CCTV equipment	• High power consumption • High heat generation • High space requirements • Manufacturers have largely discontinued • Burn-in tendency (when consistent image is permanently etched onto screen)
LCD	• Compact and relatively light • Low power consumption and electromagnetic radiation • Wide range of screen sizes available • No flicker • Less burn-in. Long life expectancy (up to 50,000 hours)	• Limitations reproducing true black colors • Restricted viewing angle • Low image contrast • Time lag of pixels can lead to "smearing" effect
OLED	• Thinner, lightweight, low-power consumption, and better dynamic contrast ratio • Brighter display and wide color range	• More expensive than other monitor technologies • Shorter lifespan
Plasma	• Wider viewing angles than LCD • Higher black levels • Higher contrast ratio (the difference between the brightest and the darkest part of the image)	• Fragile • High power consumption • High heat generation • Susceptible to burn-in

*Many of the disadvantages listed are being addressed by the industry in newer products.

Table 3-4. Video Transmission Signal Display Types

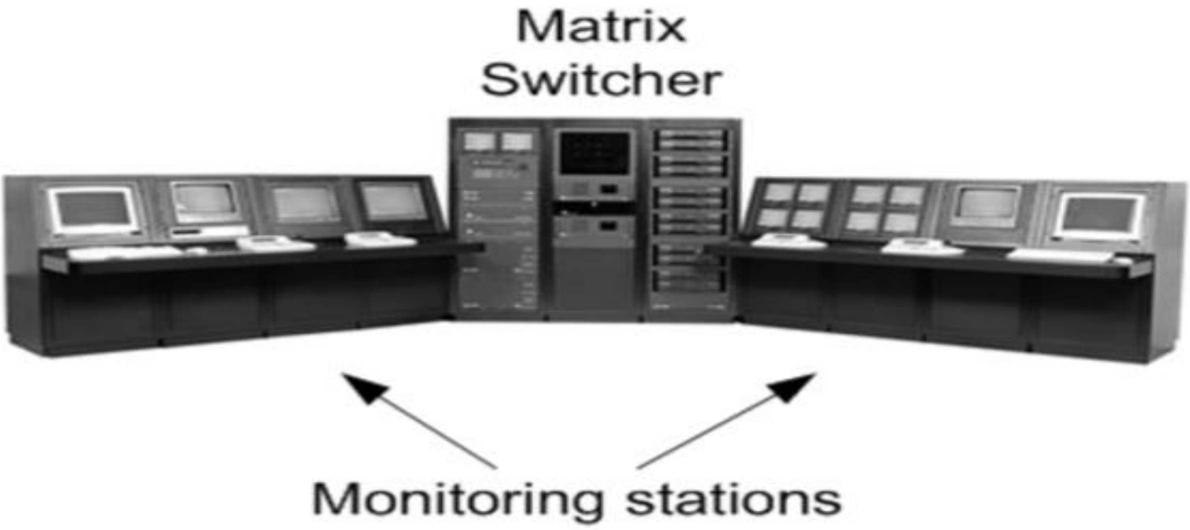
Video Transmission Signal	Analog	Digital	Digital	Digital	Digital
Signal Type	NTSC	Standard Definition	Standard Definition	High Definition	High Definition
Video Quality	Standard TV	Good Picture	Better Picture	Best Possible Picture	Best Possible Picture
Aspect Ratio	4:3	4:3	4:3 or 16:9	16:9	16:9
Maximum Resolution (Digital Pixels)	480i	480i (480x640)	480p (480x704)	720p (720x1080)	1080i/1080p (1080x1920)

Switchers and Multiplexers: In CCTV systems that have more cameras than monitors and recording devices, switchers and multiplexers are used to route the video signal. Switchers are simpler in concept than multiplexers. They can be set manually or automatically to send analog or digital video to a monitor or a recorder. Some switchers can send frames or fields from several cameras to a recorder in a sequential manner, recording a frame or field from each camera in sequence. Multiplexers, invented in the 1980s, have capabilities not available in switchers.

Multiplexers receive the analog video from several cameras and digitize the signal. Multiplexers can be programmed to prioritize the video from the different cameras according to rules. Cameras covering alarmed areas in an integrated security system may be prioritized so that their images are shown on a monitor and all frames recorded. Many multiplexers have imbedded motion detection and analysis software to support the recording or displaying of an image only when the software detects movement or some other phenomenon. Many multiplexers can be used in networks controlled by computer systems. This flexibility, when combined with digital storage media instead of tape storage systems, blurs the distinction between multiplexers and other components like digital video recorders (DVRs) and network video recorders (NVRs). DVRs and NVRs not only perform the functions of multiplexers, but also include integral hard drive storage so that video is recorded as compressed digital data. Section 3.6 provides additional information on DVRs and NVRs. Fully digital video systems, using network cameras, do not use switchers or multiplexers. The cameras send compressed digital video data directly to DVRs, or to monitors, over an Ethernet or other electronic data network. Digital imaging and digital storage devices are becoming standard, but switchers and multiplexers offer a low cost, easy-to-use alternative. The primary advantage of using switchers and multiplexers in analog CCTV systems is the ability to route the video signal to multiple output devices. Section 4 addresses transmission and storage of video using IP networks. This section on switchers and multiplexers provides information about more traditional CCTV systems, which may use analog or digital components.

As CCTV systems are built with greater numbers of cameras and monitors, switchers have become more powerful and versatile. Microprocessor-based switchers have a host of features such as: • Camera and lens control; • On-screen text; • Password protection for programming; • Partitioning of video for selected users; • Interface capabilities with additional alarm and relay panels; • Remote viewing and control over IP networks; • Macro programming and event timers; • Integrated color bar generators for setting up monitors; and • Networking for several switchers.

CCTV systems with large numbers of components need microprocessor-based switchers that can handle large numbers of video inputs and outputs. These are referred to as matrix switchers.



Video Recorders: Recording capability is essential for assessment, investigation, and evidence collection. Video recording has transformed from tape-based systems to digital hard drive systems. While some systems still use tape, the popularity of digital video has driven the demand for recorders with hard drive storage. Traditional analog CCTV systems in which video is recorded to video cassette recorders (VCR) are rare in today's environment and have rapidly been taken over by DVRs and NVRs on IP networks. Many types of recorders are being used today, but this section will focus on digital recording equipment as well as removable media and some emerging technologies.

Digital Video Recorders: A CCTV system may send digital or analog video to the recording system. A DVR receiving analog video takes two fields of the analog signal and builds one image, which is then digitized and compressed. If the video going to the DVR is digital, it is normally compressed to save storage space. Various data compression methods can be used that offer varying degrees of performance, quality, and storage economy. DVRs can include a variety of features and capabilities such as: • On-board software, such as video analytics; • Image protection/authentication techniques; • Ports for additional recording capabilities; • Internal hard drive for video storage; • Ability to easily search for and locate events; • Ability to record one or more camera inputs while performing video analytics; • Removable hard drive for archiving purposes; and • Ability to transfer data to expandable storage systems called

Redundant Array of Independent Disks (RAID) to free up recording space. DVRs may be classified as simplex, duplex, or triplex. Simplex DVRs cannot record while searching and viewing recorded images. Duplex systems can record while searching. Triplex DVR systems allow the operator to view recorded and live video while recording continues.

Data Compression Methods: The type of compression selected dictates the quality and amount of video data that can be stored on a DVR. Therefore, it is important to understand the differences in compression technologies. The compression aggressiveness affects how much video data can be stored on a hard drive recorder and the quality of the images during playback. If the compression is too aggressive and too many pixels are removed, faces or objects may not be recognizable. The following example illustrates the amount of video data that accumulates over time and why compression techniques are important. A DVR will require the following storage capacity for 2 hours of uncompressed data (1 megabyte [MB] of data per image). 1 MB of data x 30 images per second x 60 seconds x 60 minutes x 2 hours =216,000 MB or 216 gigabytes (GBs) This data storage requirement can quickly overwhelm the hard drive space of a DVR, especially when multiple cameras are being used. The number of images per second can be reduced to allow a smaller storage requirement, but the example above provides a good illustration of why compression techniques are needed.

Transmission: The transmission system is an important component of the CCTV imaging chain that sends and receives video signals between the cameras, the processing system (i.e., DVRs, NVRs, and multiplexers) and the monitoring system (i.e. the display). Transmitting a strong video signal with low noise is vital to producing a high-quality image on the monitor. Many problems associated with the quality of a CCTV system signal are attributable to the transmission system. Many types of video transmission technologies are available today. High-quality components are needed to produce a high-quality result. The distance between a camera, monitor, and storage system is one of the most important criteria for deciding which means of transmission to use. IP-based systems are quickly gaining popularity as digital formats are becoming more common within CCTV systems. Other selection factors include installation costs, existing infrastructure, and availability of power. The options described below are available when determining the best suited transmission strategy. Any copper conductor (coaxial cable, twisted pair, etc.) exposed to an outdoor environment is susceptible to noise and lightning strikes. Lightning protection is an essential added expense and could degrade the video transmission if improperly installed and maintained.

Wired Transmission: Wired CCTV systems use cables to connect cameras to other CCTV components. Wired transmission can provide good quality video images with fewer instances of interference because cables are shielded. Cameras can be located far away from recording or monitoring equipment. Three types of wired CCTV systems are commonly used today: coaxial cables, UTP cables, and fiber optic cables. Transmission over a public telephone network is not advisable for CCTV transmission due to the cyber security issues related to an open network; however, it is still used in some CCTV environments.

Coaxial cable is the most common method of transmitting video signals from the camera to the monitor or other CCTV components.

UTP Wire: In some cases, running coaxial from a camera to a monitoring location is not practical and existing telephone wire can be used. For example, many buildings contain abandoned telephone lines, known as UTP, which can be used for a video system. This has several advantages: overall cost savings, low susceptibility to EMI or induction, no ground loop concerns, and ease of use. Also, telephone wire is smaller and much lighter than coaxial cable. It should be noted that using abandoned telephone wire beyond a facility's boundaries may require an approval process and service agreement with the telephone company.

Fiber optic cable is lightweight and made up of a single spun glass or plastic fiber or a group of such fibers encased in a protective covering. It has a broad bandwidth, making it ideal for carrying video signals. Fiber optic cable can be used in runs up to 6 miles without amplification. The video signal coming from the camera must first go through a fiber transmitter which converts electrical signals to light impulses. A fiber receiver at the other end is required for conversion back into electrical signals. Fiber optic cable is immune to RFI and EMI. In addition, grounding is not an issue with fiber optics and the cable is less susceptible, if not immune, to lightning strikes. Furthermore, in systems designed with top-of-the-line components, fiber optic cable has high cost to performance ratios. A single strand of single mode fiber can carry 32 channels of analog video. In low-end systems, the expense of fiber optic cable may not be warranted. Fiber optic cable requires extremely precise installation as the most minor damage to the cable or sharp bends can cause a major degradation of the signal.

Telephone Network Another option for wired transmission of video signals is the telephone network. Although standard voice grade telephone lines do not have enough capacity to handle real-time full motion video, they still have value in specialized CCTV applications. However, telephone lines are not recommended when the security of the video is a concern due to the cyber security vulnerabilities.

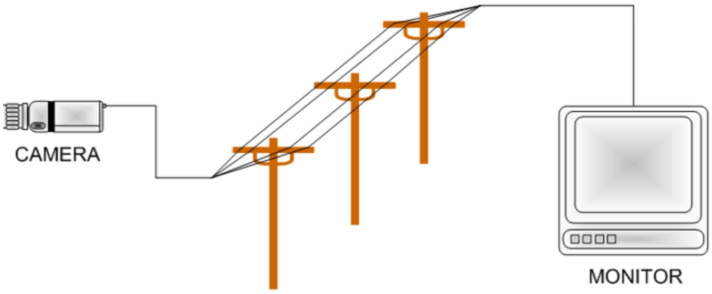


Figure 4-2. Telephone Network Example

Microwave transmission is a wireless, line-of-sight transmission medium with many similarities to laser and LED systems. Depending on its specific configuration, a microwave system can transmit video, audio, and data. Some factors to consider when exploring microwave transmission are: • Transmissions from 100 feet to 20 miles are possible, and longer distances can be achieved with microwave repeaters; • Microwave transmitters may require FCC licensing; • Microwave transmission tolerates adverse weather and obstructions better than laser transmission. However, very heavy rainfall and very dense fog can affect microwave transmissions; • Microwave systems emit at low energies, which are typically less than 1 milliwatt per square centimeter (mW/cm²). In comparison the minimum safe radiation exposure level established by the American National Standards Institute is 10 mW/cm²; and • Microwave systems typically are more expensive than LED or laser systems.

Microwaves can pass through glass; therefore, mounting a system indoors to maintain an aesthetically pleasing building exterior may be feasible. The receiver and transmitter require careful alignment for optimal results. Since the signal can weaken over a long distance, it is important to consider the distance and performance requirements carefully in the system design. Shiny surfaces, such as windows or water that are aligned parallel to the beam, may reflect energy in the outer portions of the beam toward the receiver and degrade the video signal.

IP Network Transmission IP-based systems have emerged as an attractive alternative to other technologies, due in large part to their ability to achieve high-performance video capabilities at a low cost. The industry has found ways to implement IP-based systems that use existing cameras, cables, and other equipment. However, organizations planning and designing new systems should consider IP-based technology. This section addresses the basic parameters of an IP-based network system.

Internet Protocol Network System Overview: IP-based CCTV systems are designed to provide the ability to monitor, record, and stream video over a network to computers or other equipment. The system can use existing local area networks (LANs), wide area networks (WANs), and/or wireless LANs (WLANs) to save on installation costs. However, for added security, an organization could install its own private area network (PAN) cabling and support hardware. Power over Ethernet (PoE) technology is also an option within an IP-based system to increase savings and reliability. PoE enables various networked devices to receive power and data through one standard cable, which can be a significant cost savings when designing CCTV systems. A simple IP-based CCTV system, such as the one seen in Figure 4-5, consists of a network camera (although analog cameras can be used with additional equipment), a network switch, and a PC for viewing, storing, and analyzing data and managing the CCTV system.

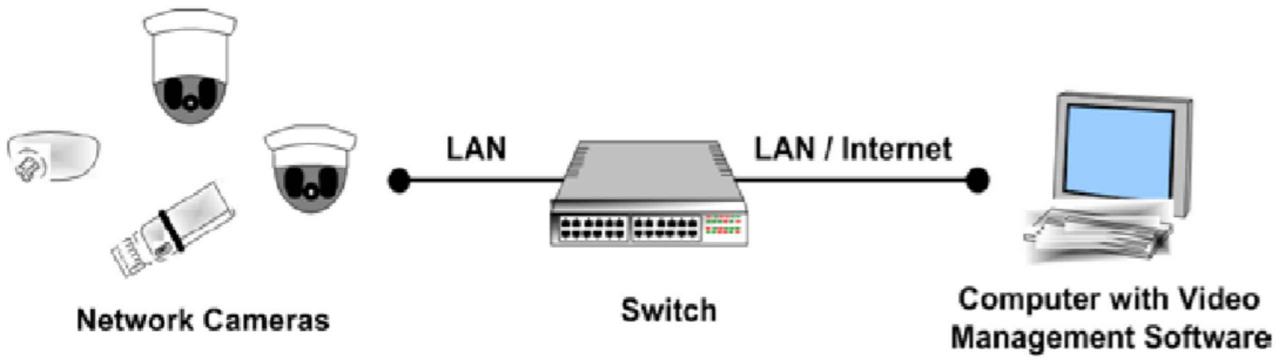


Figure 4-5. IP-Based CCTV System

Traditional analog-based CCTV systems require dedicated point-to-point cabling from each camera to the recording and/or viewing locations. In an IP-based CCTV system, video is digitized at the camera and can then be transmitted over the IP network to virtually any location around the world. Most analog systems are traditionally unidirectional, whereas network based systems are bidirectional, easier to integrate into larger systems, and highly scalable. Network cameras and other devices can not only send audio/video, but can also send other data like text or short message service (SMS) messages to users as well as receive audio and data (which can activate alarms, door entries, and external alarms). In addition, IP-based systems have the ability to interface and communicate with multiple parallel applications (e.g., motion detection or license plate readers).

Benefits of IP-Based Systems: Digital systems in general have a variety of advantages over analog systems such as ease of use, advanced search capabilities, simultaneous record and playback, improved image quality, and efficient compression and storage options. IP-based systems also provide many benefits that include: • Remote accessibility; • High image quality; • Future integration with digital technologies; • Flexibility; • Scalability; and • Cost-effective transmission.

IP-Based System Components: The flexibility of IP-based systems is attributed to the variety of configurations and types of components compatible with IP technology. Since the number of possible custom configurations is so vast, the following list is just a sample of the type of components compatible with IP-based systems. Cameras–Both IP network cameras and analog cameras can be used in an IP-based system. Video Encoders–When using analog cameras, a video encoder or video server needs to be connected to the analog cameras to convert the video to a digital format. The encoder then sends the data over an IP network. Network Switches–Switches allow CCTV devices to communicate with each other and share information. Networks–A network can be small or extensive, wired or wireless or a combination thereof. The most common approach taken by organizations is to use LANs or WANs. Network bandwidth capacity can be increased by adding switches and routers. Wireless networks are a good option when traditional wired networks are too costly or difficult to install. Power over Ethernet–PoE is an option for using a wired network to distribute both data and power. PC with Web Browser–PCs can access live and recorded video over the Internet as needed.

PC with Video Management Software–PCs can record and store video from cameras, as well as view live and recorded video as needed. Additionally, video management software can support video being accessed over smartphones or tablets. Storage Devices–Video transmitted through an IP system can be stored on a server, a network device such as direct attached storage (DAS), storage area networks (SAN), network attached storage (NAS), or a PC hard disk. These storage devices are discussed further in Section 5. Mobile Devices–IP-based systems can be easily configured to facilitate access to video via the Internet from smartphones, laptops, and other mobile devices

Cyber Security : The confidentiality, integrity, and availability of data are critical for any organization. CCTV systems, especially IP-based systems, present a cyber-security risk because their video images and critical operational surveillance data is transferred and stored on a network. Protecting information should be a high priority in security planning. Cyber security is a large and complex issue that extends far beyond the implications of a CCTV system. However, issues such as hardware and software control measures, network control measures, and network security should be considered by any agency incorporating an IP-based CCTV system.

VIDEO STORAGE: A CCTV system needs to be designed and configured so that it retains the necessary quantity and quality of video data. CCTV systems must also be equipped with appropriate export and archiving capabilities. An organization must clearly determine the purpose of the video that is being collected and understand how it may be used. Equally important is establishing the image resolution, image rate, and the number of days of recording that will be stored by the system. These factors will influence the use, access, recall, and storage requirements of a CCTV system.

Media Storage : Many organizations use write-once, read-many (WORM) media for long-term storage needs due to its secure and cost effective features. Current WORM technologies include optical discs such as CDs and DVDs, while older systems may use magnetic disks or tape. A disadvantage of using WORM media is that record management can be cumbersome. For example, a CD can be destroyed, damaged, or easily removed from its environment, which is not the case with data stored on servers. The storage capacity of WORM media is also a concern. It may take considerable time to copy all video data required for long-term archiving. In contrast, a secure server with appropriate disk storage offers a central, searchable repository of video images, which can be easily accessed, recalled, and viewed by authorized personnel. Servers also enable data to be migrated automatically and suffer no loss within a RAID system. RAID storage allows images to be distributed across multiple hard disk drives to protect against a single point of failure. RAID systems conduct integrity checks and perform repairs from the parity disk if data integrity has been compromised. WORM devices have minimal to no data recovery capability if they are damaged.

Scalable Network Storage : Data storage in CCTV systems is changing rapidly and has been influenced greatly by IP-based systems requiring efficient and cost effective storage. The market offers various network storage options for IP-based systems and hybrid systems, which incorporate both analog and IP technology and communications protocols. Most organizations with a sizeable CCTV system will require network storage beyond local DVR storage capabilities. Network storage involves a physical separation of storage media from the end user. For example, storage media located within a recording device (e.g., a hard drive or DVD) has limited capacity, but network storage is independent from the recording device and offers greater scalability for the large storage demands of video images.

Interface Protocols : Connecting servers to storage devices is typically accomplished through use of small computer systems interface (SCSI) protocol. Within CCTV systems, SCSI is the most dominantly used storage interface. With the increase in available IP technologies and the need for interoperable and open architecture, the iSCSI protocol incorporates new capabilities to access voice, video and data from multiple types of network storage devices and make it available across an IP network. The “i” in iSCSI stands for Internet and in the simplest terms, iSCSI combines the SCSI storage capabilities with the transmission control protocol/internet protocol (TCP/IP). Network storage device options include direct attached storage, storage area network, and network attached storage. These are discussed below.

Direct Attached Storage: DAS is considered an older technology that was developed as a stand-alone mechanism to connect hosts to storage devices through a direct, one-to-one SCSI attachment. Adding storage and servers to a DAS system to meet demands can result in a proliferation of server and storage islands. In a DAS environment, storage sharing is limited because of its direct affiliation to the servers. DAS is still used today in CCTV systems, but external storage solutions are usually better options for CCTV video than fixed DAS storage. As LANs gained popularity, the server attached storage (SAS) was developed as an alternative to DAS in order to achieve a distributed approach via a LAN.

Storage Area Network: A SAN consists of communications infrastructure and management layers that ensure secure and robust data transfer. Storage appliances within the SAN contain data blocks. Operators access the data blocks when needed over the network. A SAN is a dedicated, high-performance network, typically using fiber channel technology, as seen in Figure 5-1. Fiber channel protocol and interconnect technology provide high performance transfers of block data. SANs are generally used to connect numerous storage devices such as DVRs and NVRs to one or more centralized, shared storage systems. An increasing number of CCTV components are being designed to connect directly to SANs via iSCSI.

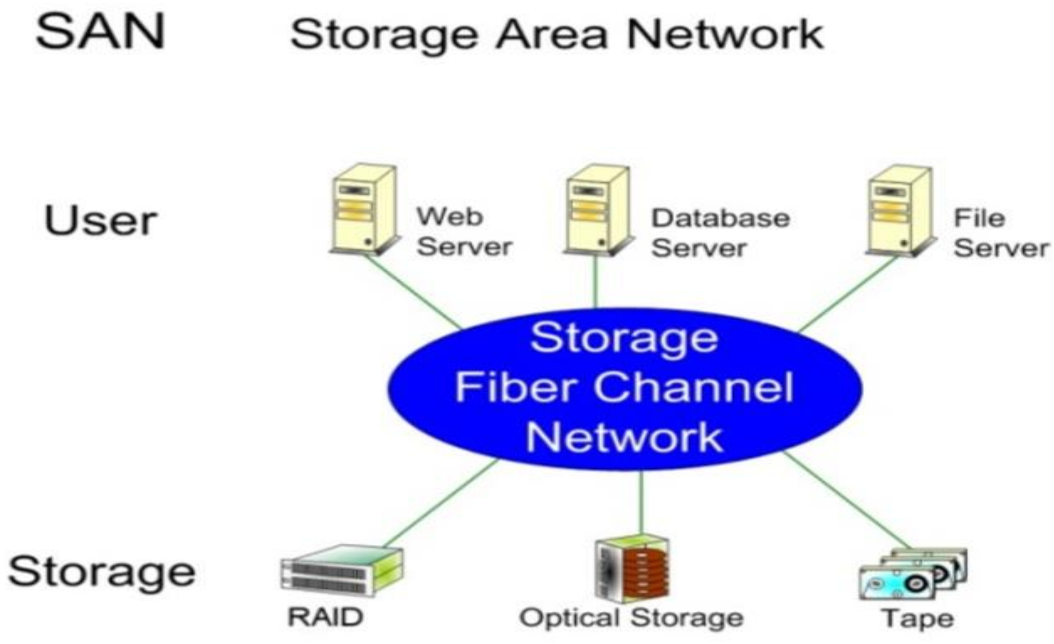


Figure 5-1. Storage Area Network

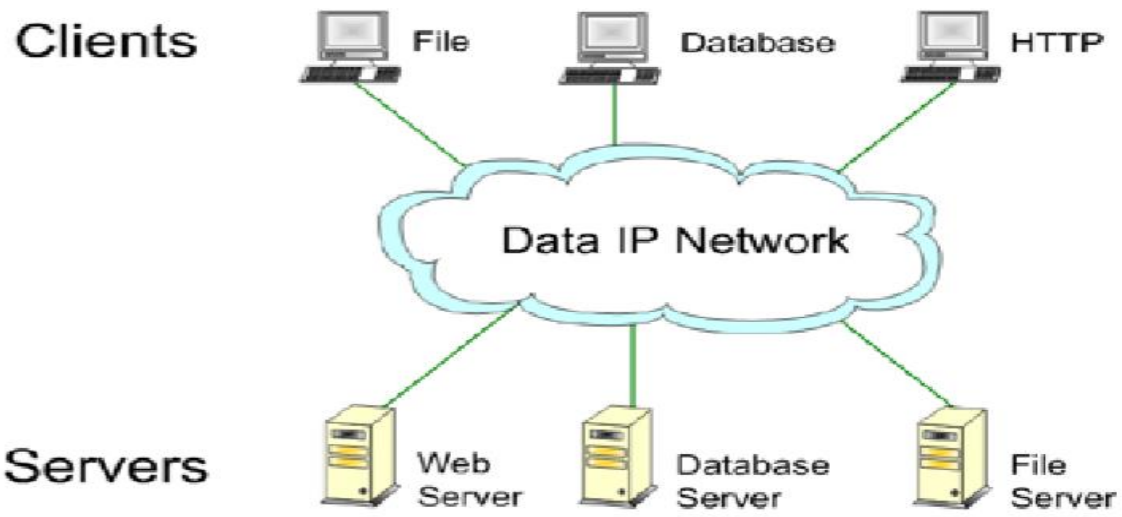


Figure 5-2. Network Attached Storage

NAS systems: This system record and access data in file format and consist of an engine that retrieves files from one or more storage devices. With NAS technology, servers maintain file systems on their local storage, and clients can access files at servers over a network via LAN or WAN technology, typically using Ethernet. NAS protocol is typically TCP/IP based, like the example shown in Figure 5-2. NAS is considerably less expensive than DAS and SAN; however, many considerations for the entire CCTV system will need to be evaluated to ensure NAS is compatible with other system components.

Video analytics : **Video analytics** software is often referred to as automated video surveillance, intelligent video, smart video, or video motion detection. The capabilities of video analytics are very beneficial within a CCTV system. Video analytics uses computer algorithms to monitor real-time video captured by CCTV cameras to enhance security surveillance of people, vehicles, objects, and their associated behaviors within a camera's view. Video analytics can help organizations become more efficient by automating part of the monitoring process and averting the time-consuming and tedious process of reviewing extensive quantities of stored video. Video analytics systems can be used to identify suspicious activity in airports, train stations, seaports, and any other high traffic areas. A common application of video analytics is constant monitoring of surveillance video to provide an alert to security officers on events, such as an unauthorized intrusion in progress or a suspicious individual loitering in the parking lot. Some video analytics systems include license plate recognition (LPR), which provides law enforcement and security personnel with an automated tool to identify vehicles from the information on their license plates. Analytics applications also include traffic and tollbooth monitoring, facility and border surveillance, building and parking lot security, and identifying vehicles of interest. Additional information on video analytics is included in the Video Analytics Systems Market Survey Report available in the SAVER section of the DHS S&T website

SYSTEM INTEGRATION: The complexity and sheer number of components, software applications, inputs, outputs, transmission infrastructure, processing and storage devices, and customized settings of CCTV systems provide a wide range of possible configurations to meet an organization's requirements. Integrating CCTV components requires thoughtful planning when new elements are brought online to achieve new capabilities or improve performance. Organizations benefit from system integration by removing stove-piped systems and their associated operations and maintenance costs. Integration will also improve data communications among complementary systems as well as efficiency. The training requirements alone to ensure operators understand all the features and functions of non-interoperable systems would be overwhelming. Alternatively, greater uniformity and interoperability among devices and communications architecture improves an organization's ability to maintain a CCTV system. Organizations focused on implementing integrated solutions will likely have fewer instances of compatibility issues and system failures.

Systems Approach: Organizations should strive to have all security systems and their subsystems linked together to ensure the system's components work together as a whole.

Achieving systems integration is both a conceptual and a logistical challenge. Figure 7-1 illustrates the many different layers involved in an integrated security system

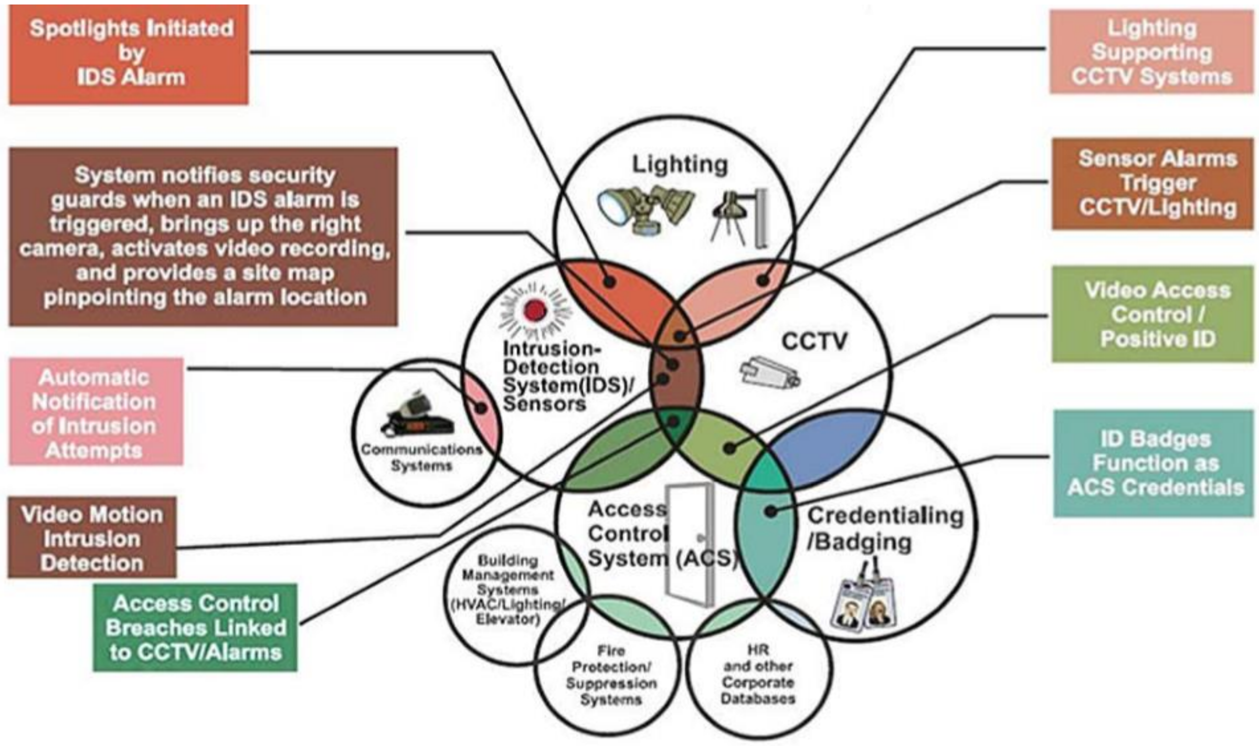


Figure 7-1. Integrated Security System

Integrating CCTV Components: Newly designed CCTV systems have an advantage over existing systems because they can be designed from start to finish with current technology components from manufacturers that are easy to integrate. When selecting CCTV devices, organizations should consider future needs and requirements, such as the potential for expansion, scalability, integration, and upgrading.

Some additional technology considerations include: • Ability to use a consistent hardware platform throughout the enterprise; • Off the shelf software and equipment, not proprietary; • Compatibility for data collection and storage; • Advanced software graphical user interfaces (GUIs) to integrate controls and displays; • Ability to create single security user profiles used by multiple security applications; and • Vendor support to facilitate, test, and commission system integration.

Other Considerations: An organization should consider system integration during the project planning and design phase. This applies to various types of projects that may impact the CCTV system including: • New CCTV systems; • Acquisition or expansion of existing structures into an existing CCTV system; • Newly designed and built structures; • Upgrades to associated parallel systems; • Adding new technologies into an existing CCTV system; and • Expansion of cameras (renaming cameras across stove-piped platforms could be a problem).

EMERGING TECHNOLOGY: The CCTV industry is part of an ever-changing IT environment. New capabilities in any one of the digital components, Internet, or telecommunications industries will eventually be applicable to CCTV systems. Entities investing in new CCTV systems or upgrades are incorporating these new technologies as they become available. While it is unclear how new IT developments will be addressed by the security surveillance industry in the future, the popularity of IP-based CCTV systems in recent years illustrates how the market may respond to similar innovations. This section explores how CCTV systems may be influenced by new developments and evolving expectations from users to improve existing equipment and software. It also discusses several emerging IT trends that could factor into future CCTV products and platform integration capabilities.

Digital Technologies: Newer CCTV systems are being built entirely with digital components, from the cameras to the recording devices, and no longer require conversion back and forth between analog and digitalvideo signals. Computers and digital recording devices are replacing tape-based storage systems, multiplexers, and switching systems. Since any kind of information can be digitized, CCTV systems can integrate with almost any other information handling system. They can be programmed to process, analyze, display, and store data from other media and from other surveillance sensors. CCTV can be blended into a facility's intrusion detection, access control, and alarm systems so that information from all devices is displayed to security personnel on the same displays, images, and maps. As a result, security monitoring personnel have access to multimedia presentations that merge video with radar, laser radar, sonar, intrusion detection alarms, satellite mapping, and imaging to create integrated visual situation images.

Improvements to Existing Technology: Many existing CCTV systems can be updated with current technology that incorporates new features to meet evolving consumer needs and expectations. Cameras will be made smaller and lighter and will consume less power. CCTV components are being designed to handle increased data and file sizes. Video compression algorithms are becoming more efficient, and network traffic management solutions are improving capabilities to store, retain, archive, and recall video as needed. Vendors may also focus on their product's capabilities to effectively integrate CCTV system hardware and software.

Major IT Trends: The major trends in CCTV systems are primarily related to developments in the information industry and the need for products and components to adopt new technology environments and capabilities. As new innovations related to digital formats and business-based IT solutions are realized, the CCTV industry will see a significant number of new capabilities. While some new technologies will involve hardware, the majority of emerging technologies will likely be focused on the interpretation of video through software, data storage management, and system integration.

APPLICATIONS: CCTV is commonly used for a variety of purposes, including:

- Maintaining perimeter security in medium- to high-secure areas and installations.
- Observing behavior of incarcerated inmates and potentially dangerous patients in medical facilities.
- Traffic monitoring.
- Overseeing locations that would be hazardous to a human, for example, highly radioactive or toxic industrial environments.
- Building and grounds security.
- Obtaining a visual record of activities in situations where it is necessary to maintain proper security or access controls (for example, in a diamond cutting or sorting operation; in banks, casinos, or airports).

CCTV is finding increasing use in law-enforcement, for everything from traffic observation (and automated ticketing) to observation of high-crime areas or neighborhoods. Such use of CCTV technology has fueled privacy concerns in many parts of the world.

There are numerous applications of CCTV and few are briefly discussed here.

Education: One instructor may lecture to a large number of students sitting at different locations. Similarly e close-ups of demonstration experiments and other aids can be shown on monitors during these lectures.

Medicine: Several monitors and camera units can be installed to observe seriously ill patients in Intensive Care units. In medical Institutions, operations when performed that can be shown to medical students without their gathering around the operation table.

Business: Television camera can be installed at a different location in big departmental stores to keep an eye over customers and sales personnel.

Surveillance: In banks, Railway yards ports, traffic points and several other similar locations, closed circuit TV can be effectively used for surveillance.

Industry: In Industry CCTV have applications in remote inspection of materials observance of nuclear reactions and other such phenomena would have been Impossible without television. Similarly television has played a great role in the scanning of Earth's surface and probing of other planets.

Home: In home, a CCTV monitor finds its application in seeing the caller before opening the door.

Aerospace and Oceanography: Here a wireless link is used between the transmitter and receiver. In some applications camera is remotely controlled over a microwave radio link. As shown in the above figure, for Aerospace and Oceanography a carrier is used for transmitting the signal and a complete receiver is than necessary for perception.

SET TOP BOX

It is an interactive device which integrates the video and audio decoding capabilities of television with a multimedia application execution environment. It provides a user friendly interface offering personalized multimedia services and regular cable TV service.

In other words, A **set-top box** (**STB**), also colloquially known as a cable box, is an information appliance device that generally contains a TV-tuner input and displays output to a television set and an external source of signal, turning the source signal into content in a form that can then be displayed on the television screen or other display device. They are used in cable television, satellite television, and over-the-air television systems as well as other uses. A computer that connects to your television, allows you to use telephone line or cable connection do you browse the Internet and exchange electronic mail on your television. Set-top boxes can also enhance source signal quality.

Difference between STB and multimedia computer

Multimedia computer	STB
Expensive and versatile device	Inexpensive limited functionality device primarily targeted at entertainment
Might be equipped with floppy or a hard disk drive	Not likely to be so equipped
Can execute a wide variety of application programs	Limited scope of applications that can be executed

Professional set-top boxes are referred to as IRDs or integrated receiver/decoders (An **integrated receiver/decoder** (**IRD**) is an electronic device used to pick up a radio-frequency signal and convert digital information transmitted in it) in the professional broadcast audio/video industry. They are designed for more robust field handling and rack mounting environments. IRDs are capable of outputting uncompressed serial digital interface signals.

Hybrid set-top boxes, such as those used for Smart TV programming, enable viewers to access multiple TV delivery methods (including terrestrial, cable, internet, and satellite); like IPTV boxes, they include video on demand, time-shifting TV, Internet applications, video telephony, surveillance, gaming, shopping, TV-centric electronic program guides, and e-government. By integrating varying delivery streams, hybrids (sometimes known as "TV-centric") enable pay-TV operators more flexible application deployment, which decreases the cost of launching new services, increases speed to market, and limits disruption for consumers.

As examples, Hybrid Broadcast Broadband TV (HbbTV) set-top boxes allow traditional TV broadcasts, whether from terrestrial (DTT), satellite, or cable providers, to be brought together with video delivered over the Internet and personal multimedia content. Advanced Digital Broadcast (ADB) launched its first hybrid DTT/IPTV set-top box in 2005, which provided Telefónica with the digital TV platform for its Movistar TV service by the end of that year, In 2009, ADB provided Europe's first three-way hybrid digital TV platform to Polish digital satellite operator n, which enables subscribers to view integrated content whether delivered via satellite, terrestrial, or internet.

Digital video networks(DVN)

- Digital television advantage.
- Digital video delivery requirements
- High bandwidth.
- Fibre cable, satellite, broadcast and computer networks can be used.
- Traditional cable networks analogue systems and their drawbacks.
- Satellite and terrestrial networks.
- Internet suitability for interactive network and drawback.

DVN need to deliver a high bandwidth stream into consumer homes and a low bandwidth communication layer for interaction between the STB user and the service provider.

- Cable operators approach for DVN.
- Need to extend unidirectional coaxial networking with a communication path from subscriber home to cable service gateways that can control the content being sent to subscriber.
- Advantage of video compression technology.
- Constraints with cable networks.
- Telephone companies approach.
- Advantage with telephone companies.
- Already have set up for P2P and technology for control and service gateways to manage wide area switched star networks.

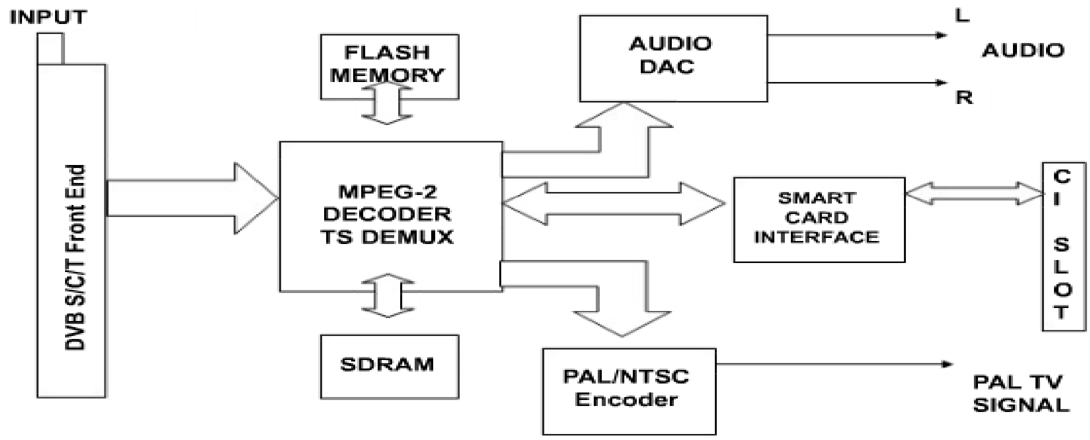
Drawback:

- Low bandwidth between head end equipment and consumer's home.

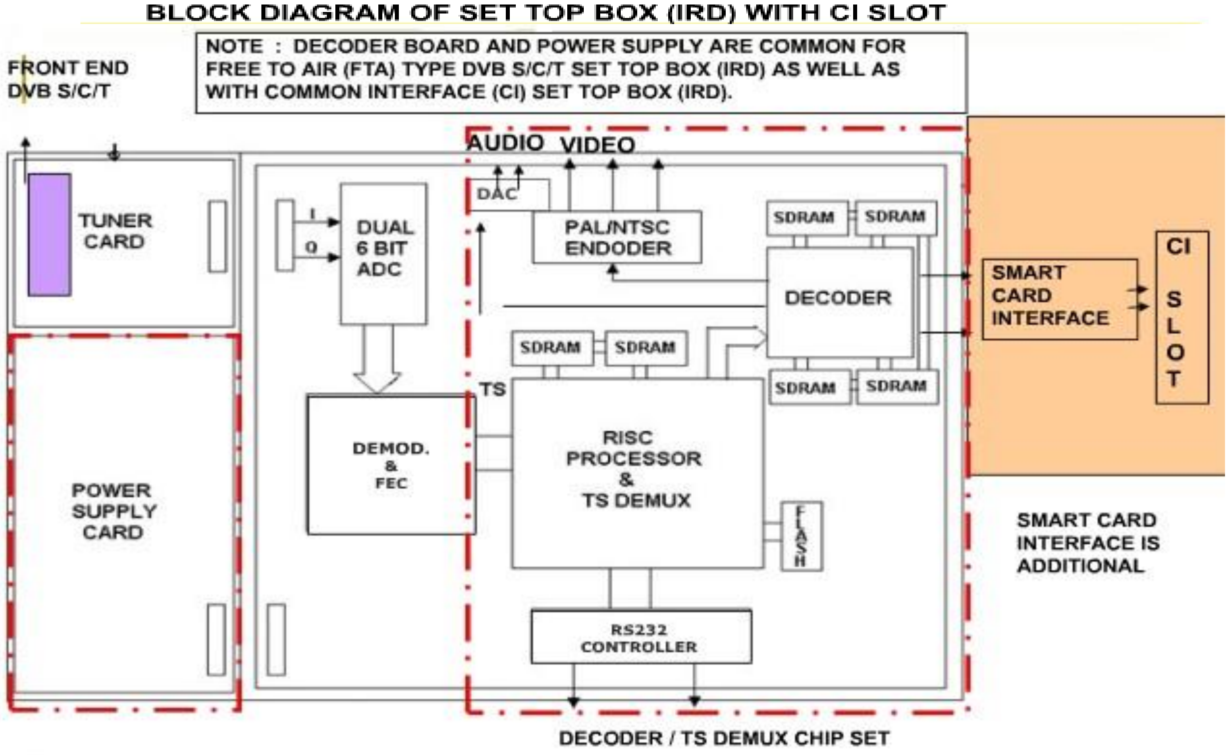
Set Top Box Hardware Architecture

- Requirements based on functionality.
- Should have MPEG-2 sub-system to decode MPEG-2 video and audio.

- Should enable user to download custom applications and execute them on the STB which requires a general purpose microprocessor in the STB with an architecture that supports control of the different devices.



BLOCK DIAGRAM OF SET TOP BOX (IRD) WITH CI SLOT



DETAILED BLOCK DIAGRAM OF A SET TOP BOX (IRD) WITH CI SLOT

STB software architecture

For personalized interactive services, a set top box should be addressable in the video dial-tone network, thus providing P2P communication between a video and information provider and a user. It should also have the ability to download client applications and execute them locally.

The control software on the set top box ties together all the hardware components into a functioning unit.

- STB software has two parts.
- System software which provides the DAVID application programming interface.
- Application software that provides cable TV functionality or some other personalized multimedia service.
- DAVID system software includes.
- Operating system (os-9) kernel.
- Device drivers.
- File manager.

Reference Books

1. Dr. Sanjay Sharma, “Analog Communication Systems” 2013, January 2013
2. R.P Singh, S.D. Spare, “Communication Systems: Analog and Digital” ,Tata McGraw-Hill, 3rd Edition, July 2017.
3. Sanjeev Guptha, “Electronic Communication (Analog, Digital and Wireless)”, Khanna publications, January 1997
4. B.P.Lathi, “Modern Digital and Analog Communication Systems”, 3rd Edition, Oxford University Press, 2007.
5. Louis. E. Frenzel, “Communication Electronics Principles and applications”, 3rd Edition, Tata McGraw Hill, 2002.
6. Louis. E. Frenzel, “Principles of Electronic Communication Systems”, 3rd Edition, Tata McGraw Hill, July 2017.
7. Dennis Reddy and John Coolen, "Electronic Communications", 4th Edition, Prentice Hall Publishers, 1995.
8. Electronic Communications Systems by Wayne Tomasi. Pearson education India.
9. Electronic Communication Systems by Roy Blake By Cengage learning.
10. Communication Systems By Simon Haykins By Wiley India.
11. Theory and Problem Of Electronic Communication By Lloyd Temes and Mitchel E.Schulz(Second edition), McGraw Hill Publication.
12. Kennedy, "Electronic Communications Systems", 4th Edition, McGraw-Hill Publishers, 1992.
13. Deshpande, N.D, “Communication Electronics”, Tata McGraw Hill Publishers,1989.
14. R.Gulati, “Monochrome and color television”, 2nd Edition, New age international 2005
15. https://www.dhs.gov/sites/default/files/publications/CCTV-Tech-HBK_0713-508.pdf
16. http://www.althos.com/tutorial/TV-Set-Top-Box-STB-tutorial-applications.html