srivenkateshwaraa
College of Engineering & Technology
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry)
13-A, Pondy - Villupuram Main Road, Ariyur, Puducherry - 605 102.
ASPIRE TO EXCEL
ISO 21001

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

EC T46 ELECTRONIC COMMUNICATION SYSTEM NOTES

II YEAR/ IV SEM

UNIT – I
AMPLITUDE MODULATION AND DEMODULATION

Need for modulation-Amplitude modulation-Frequency spectrum-Power relation-Different types of AM modulators-SSB and VSB generation-AM transmitters-Block diagram-Functions of each block-High level transmitter. Detection - Diode detectors-Synchronous detection. Receivers- different types AM receivers- Block diagram-choice of IF and oscillator frequencies-Tracking-Alignment-AVC, AFC-Communication receivers- AM characteristics.

Introduction to Communication:

Communication

- The transfer of information from one place to another.
- This should be done efficiently as possible, as much fidelity/reliability as possible and securely as possible.

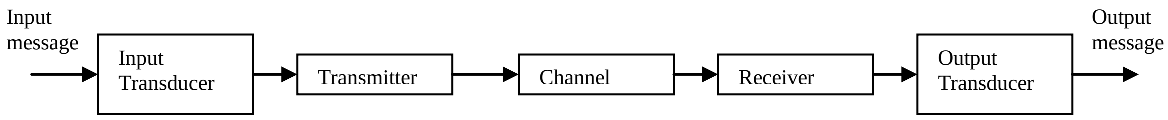
For example:

- Short distance: speech and graphical symbols
- Long distance: smoke signals, light beams, carrier pigeons, letter, telephone, e-mail, radio, tv, fax.

Communication System:

Components/subsystems act together to accomplish information transfer/exchange.

Elements of a Communication System:



Input Transducer:

- The message produced by a source must be converted by a transducer to a form suitable for the particular type of communication system.
- Example: In electrical communications, speech waves are converted by a microphone to voltage variation.

Transmitter:

- The transmitter processes the input signal to produce a signal suited to the characteristics of the transmission channel to transmit.
- Signal processing for transmission almost always involves modulation and may also include coding.
- In addition to modulation, other functions performed by the transmitter are amplification, filtering and coupling the modulated signal to the channel.

Channel:

- It is the medium of Propagation.
- The channel can have different forms: the atmosphere (or free space), coaxial cable, fiber optic, waveguide, etc.
- Here the signal undergoes some amount of degradation from noise, interference and distortion (resulting from band limiting and nonlinearities).

Receiver:

- The receiver function is to extract the desired signal from the received signal at the channel output and to convert it to a form suitable for the output transducer.
- Other functions performed by the receiver: amplification (the received signal may be extremely weak), demodulation and filtering.

Output Transducer:

- It converts the electrical signal at its input into the form desired by system user.
- Example: Loudspeaker, personal computer (PC), tape recorders.

Analog and Digital Communication System

- There are many kinds of information sources, which can be categorized into two distinct message categories, analog and digital.

- An **Analog** message is a physical quantity that varies with time, usually in a smooth and continuous fashion.
    - Examples of analog messages are the acoustic pressure produced when you speak, the angular position of an aircraft gyro, or the light intensity at some point in a television image.
    - Since information resides in a time-varying waveform, an analog communication system should deliver this waveform with a specified degree of **fidelity**.
- A **digital** message is an ordered sequence of symbols selected from a finite set of discrete elements.
    - Examples of digital messages are the letter printed on this page, a listing of hourly temperature readings or the keys you press on a computer keyboard.
    - Since the information resides in discrete symbols, a digital communication system should deliver these symbols with a specified degree of **accuracy** in a specified amount of time.

Comparisons of Digital and Analog Communication Systems

Digital Communication System	Analog Communication System
Advantage : -inexpensive digital circuits -privacy preserved (data encryption) -can merge diff. data (voice, video and data) and transmit over a common digital transmission system -error correction by coding	Disadvantages : -expensive analog components : L&C -no privacy -cannot merge data from diff. sources -no error correction capability
Disadvantages : -larger bandwidth -synchronization problem is relatively difficult	Advantages : -smaller bandwidth -synchronization problem is relatively easier

MODULATION

Modulation is the process whereby some characteristic of one wave is varied in accordance with some characteristic of another wave.

The basic types of modulation are,

- Amplitude modulation.
- Angular modulation (including the special cases of phase and frequency modulation).

i.e., **Modulation** is the process of putting information onto a high frequency carrier for transmission.

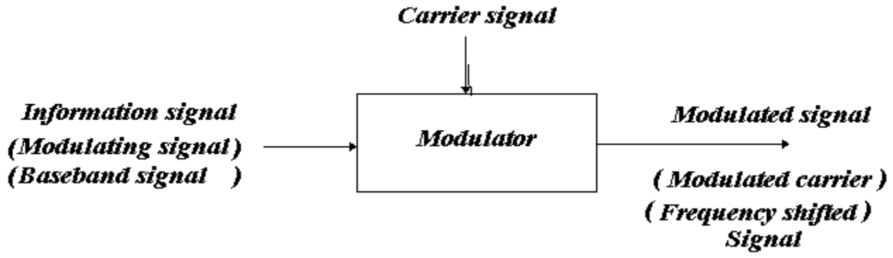


Fig. Process of Modulation

- The transmission takes place at the high frequency carrier, which has been modified to carry the lower frequency information.
- Once this information is received, the lower frequency information must be removed from the high-frequency carrier to retrieve the original message called **Demodulation.**

NEED FOR MODULATION:

- The frequency of the human voice range from about 20 to 30 kHz. If everyone transmitted those frequencies directly as radio waves interference would cause them to be inefficient.
- To overcome hardware limitations.
- Transmitting such lower frequencies require antennas with miles in wavelength.
- Modulation to reduce noise which result in optimization of Signal to Noise ratio (S/N)
- For multiplexing and frequency assignment.
- For efficient radio transmission.

CLASSIFICATION OF MODULATION:

By the nature of modulating signal

- Digital modulation: modulating signal is a digital signal.
  Examples: Amplitude Shift Keying (ASK), FSK, Phase Shift Keying (PSK)
- Analog modulation: modulating signal is an analog signal
  Examples: AM, FM, Phase Modulation (PM), Pulse Amplitude Modulation

By the nature of carrier

- Continuous wave (CW): modulation carrier is a sinusoidal wave
  Examples: AM, FM, ASK, FSK, PSK
- Pulse modulation, PM: Carrier is a train of pulses

Examples: Pulse amplitude modulation (PAM), Pulse width modulation (PWM), Pulse position modulation (PPM)

Basic Analog Modulation Methods

Consider the carrier signal below:

$$\mathbf{S_c(t) = A_c(t)\cos(2\pi f_c t + \theta)}$$

where

$A_c(t)$: carrier amplitude

f_c : carrier frequency

θ : carrier phase angle

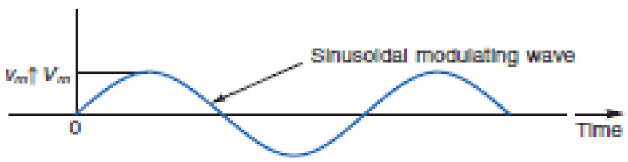
The above parameters may be varied for the purpose of transmitting information giving respectively the modulation methods; namely

- **Amplitude Modulation (AM),**
- **Frequency Modulation (FM),**
- **Phase Modulation (PM)**

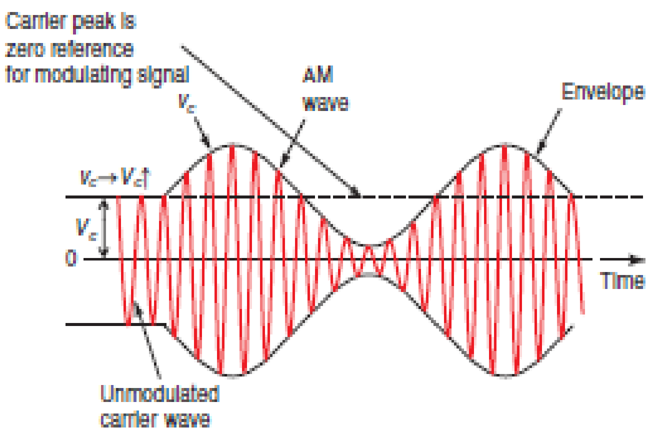
1. Changing of the amplitude produces **Amplitude Modulation signal**
   (i.e. the amplitude of the carrier waveform varies with the information signal.)
2. Changing of the frequency produces **Frequency Modulation signal**
3. Changing of the phase produces **Phase Modulation signal**

AMPLITUDE MODULATION:

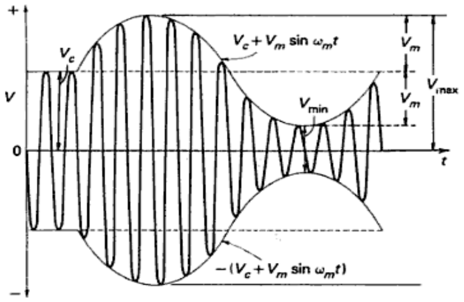
In the modulation process, the baseband voice, video, or digital signal modifies the amplitude of another, higher-frequency signal called the carrier. A sine wave carrier can be modified by the intelligence signal through amplitude modulation.



(a)



REPRESENTATION OF AM



Amplitude modulated signal (DSB-LC)

The above figure is amplitude modulated wave for one cycle of modulating sine wave given by the relation $A=V_c + V_m \sin \omega_m t$. The negative amplitude is given by $-A= - (V_c + V_m \sin \omega_m t)$.

The modulated wave extends between these two envelopes and has repetition rate equal to unmodulated carrier frequency.

V_{max} : is half the peak-to-peak value of the AM signal $V_{max(pk\text{-}pk)}/2$
V_{min} : is half the peak-to-peak value of the AM signal $V_{min(pk\text{-}pk)}/2$
V_m : is half the difference of V_{max} and V_{min} .
V_c : is half the sum of V_{max} and V_{min}.

From the figure

$$V_m = \frac{V_{max} - V_{min}}{2} \text{ and } \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\mathbf{(1)}$$

$$V_c = V_{max} - V_m$$

$$= V_{max} - \frac{V_{max} - V_{min}}{2} = \frac{V_{max} + V_{min}}{2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

The values for **V_{max}** and **V_{min}** can be obtained directly from the oscilloscope. The evaluation of the modulation index m can be achieved by invoking the following expression:

$$m = \frac{V_m}{V_c}$$

$$= \frac{V_{max} - V_{min}}{V_{max} + V_{min}}$$

POWER DISTRIBUTION IN (FULL) AM

The power in a sinusoidal signal is proportional to the square of its amplitude. The total transmitted power is the sum of the carrier power (P_C) and the power in the sidebands (P_{USB} and P_{LSB}). The total power in modulated wave will be

$$P_{total} = P_c + P_{USB} + P_{LSB} (\text{rms})$$

- Carrier power

$$P_c = \frac{V_{carr}^2}{R}$$

$$= \frac{(V_c / \sqrt{2})^2}{R} = \frac{{V_c}^2}{2R}$$

- Sideband power:

$$P_{USB} = P_{LSB} = \frac{V_{SB}^2}{R}$$

$$= \frac{\left(\frac{mV_c/2}{\sqrt{2}}\right)^2}{R} = \frac{m^2V_c^2}{8R}$$

$$= \frac{m^2}{4}.\frac{V_c{}^2}{2R}$$

- Total transmitted power:

$$P_{total} = P_c + P_{USB} + P_{LSB}(rms)$$

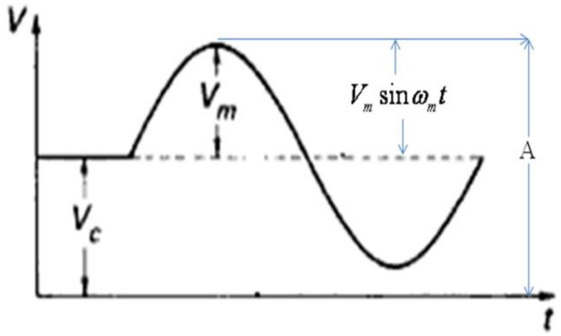
$$= .\frac{V_c{}^2}{2R} + \frac{m^2}{4}.\frac{V_c{}^2}{2R} + \frac{m^2}{4}.\frac{V_c{}^2}{2R}$$

$$= \frac{V_c{}^2}{2R}\left(1+\frac{m^2}{4}+\frac{m^2}{4}\right) = \frac{V_c{}^2}{2R}\left(1+\frac{m^2}{2}\right)$$

$$=P_c\left(1+\frac{m^2}{2}\right)$$

FREQUENCY SPECTRUM OF AM

The frequencies present in AM wave are the carrier frequency and the first pair of sideband frequencies where side band is defined by $f_{sb} = f_c \pm n\ f_m$ and in first pair n=1.



Amplitude of AM wave

From the figure we can write the amplitude of amplitude modulated wave $A = V_c + v_m$

$= V_c + V_m \sin \omega_m t$

We know that modulation index is given by $m = V_m / V_c$ rewriting the equation $V_m = m\ V_c$ substituting in above equation we get

$A = V_c + m\ V_c\ \sin \omega_m t$

$= V_c(1 + m \sin \omega_m t)$

The instantaneous voltage of resulting amplitude wave is given by $v = A \sin \theta$

$v = A \sin \omega_c t$

Substituting the value of A in other equation we get,

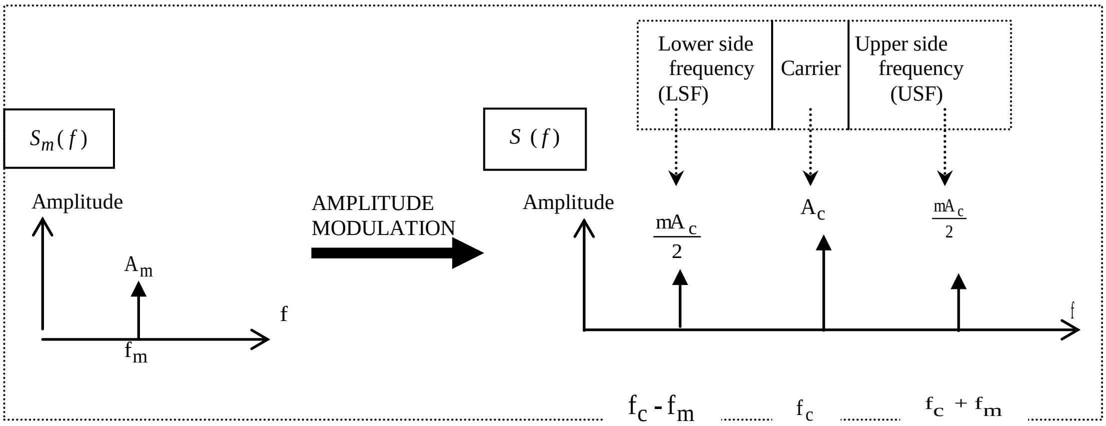
$v = V_c(1 + m \sin \omega_m t) \sin \omega_c t$

$v = V_c(\sin \omega_c t + m \sin \omega_m t . \sin \omega_c t)$

$$v = V_c \sin \omega_c t + \frac{mV_c}{2} \cos (\omega_c - \omega_m)t - \frac{mV_c}{2} \cos (\omega_c + \omega_m)t$$

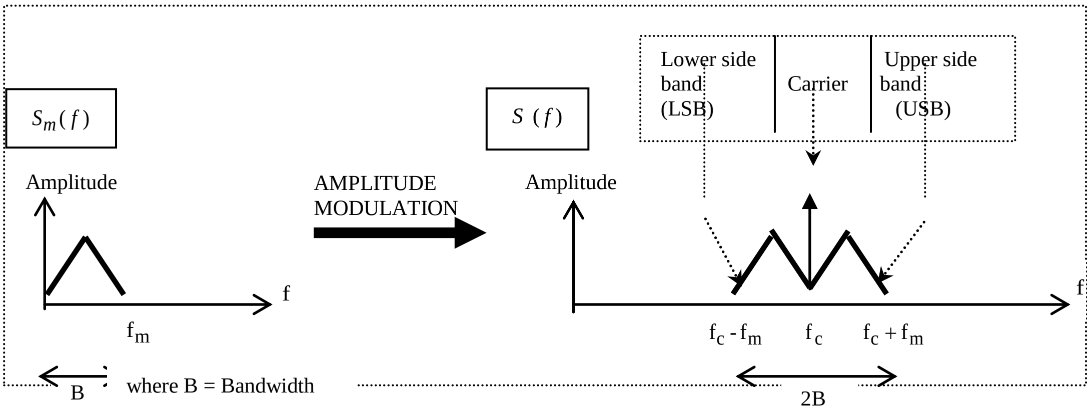
- The above equation contain three terms the first term is identical and it is a un-modulated carrier the other two additional terms produced are two side band the frequency of upper sideband (f_c+f_m)and lower sideband (f_c-f_m).
- The band width required is twice that of highest modulating frequency.
- When a carrier is first amplitude modulated, proportionality is made equal to unity and instantaneous modulating voltage variations are superimposed on to the carrier amplitude.

With single tone (one frequency) signals,



With complex signals (voice or music) or multi-tone signals.

When modulating signal (message signal) is multi-tone, AM signal becomes a band of frequencies. It is illustrated in the following figure.



The frequency spectrum of AM waveform contains **three parts**:

- A component at the carrier frequency (f_c)
- An upper sideband(USB), whose highest frequency component is at $f_c + f_m$
- A lower sideband(LSB), whose highest frequency component is at $f_c - f_m$
- The bandwidth of the modulated waveform is **twice** the information signal bandwidth.
- Because of two sidebands in frequency spectrum its often called Double Sideband with Large Carrier.(DSB-LC)
- Several baseband signals may be transmitted simultaneously on different carrier frequencies (multiplexing) provided the sidebands do no overlap.
- The information in baseband (information) signal is **duplicated** in **LSB** and **USB** and **carrier** conveys **no** information.

Modulation Index (m)

In preceding section, m is merely defined as a parameter, which determines amount of modulation. However, to establish a desirable AM communication link the depth of modulation should be less than 1.

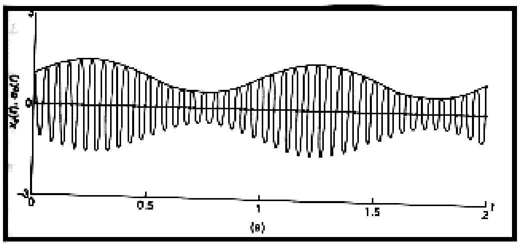
$$m \prec 1.0(100\%) .$$

This is important as to ensure successful retrieval of original transmitted information at receiver end. Note that by performing the demodulation process (reverse of modulation) the message signal is simply being traced out from the envelope of the modulated signal.

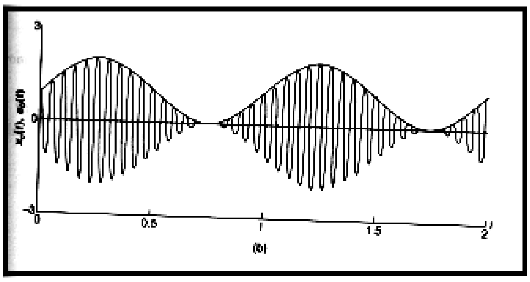
Thus, $m \succ 1.0(100\%)$, envelope distortion will occur and waveform is said to be **over modulated.** Under this circumstances, A_c is large enough, resulting the non-proportionality of $s(t)$ to $s_m(t)$ ----hence distortion of the desire message signal

Figures below show resulting AM signals when $m = 0.5\,(m \prec 1)$, $m = 1$ and $m = 1.5\,(m \succ 1)$

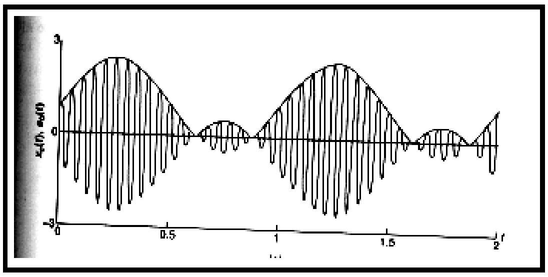
Modulation carrier and envelope detector outputs for various values of the modulation index



M<0.5



M=1



M = 1.5

- If the amplitude of the modulating signal is higher than the carrier amplitude, which in turn implies the modulation index $m \succ 1.0(100\%)$. This will cause severe distortion to the modulated signal.
- By ensuring the amplitude of $s_m(t)$ to be less than the carrier amplitude, message signal can comfortably be retrieved from the envelope waveform of $s(t)$. The ideal condition for amplitude modulation (AM) is when $m = 1$ also means $A_m = A_c$; this will give rise to the generation of the maximum message signal outputs at the receiver without distortion.

- The modulation index can be determined by measuring the actual values of the modulation voltage and the carrier voltage and computing the ratio.
- In practice, the modulation index of an AM signal can be computed from A_{max} and A_{min}.

The equation relates total power in amplitude modulated wave to un-modulated carrier power. It is used to determine the modulation index.

The maximum power in AM wave is P_t= 1.5 P_c when m =1. This is because it is the maximum power that relevant amplifier must be capable of handling without distortion.

Current calculations

Let Ic be the unmodulated carrier current and it be total or modulated current of AMtransmitter both being rms values. If R is the resistance in which these current flow, then

$$\frac{P_t}{P_c} = \frac{I_t^2 R}{I_c^2 R} = \frac{I_t^2}{I_c^2} = \left(1+\frac{m^2}{2}\right)$$

$$\frac{I_t}{I_c} = \sqrt{\left(1+\frac{m^2}{2}\right)}$$

$$or I_t = \sqrt{I_c\left(1+\frac{m^2}{2}\right)}$$

Modulation by several sine waves

Modulation of a carrier by several sine waves simultaneously is exceptional case there are 2 methods to calculate modulation index

1. Let V_1, V_2, V_3 etc …be simultaneous modulation voltages. Then total modulating voltage V_t will be equal to squre root of sum of squares of

$$V_t = \sqrt{V_1 + V_2 + V_3 + ...}$$

Dividing both sides by V_C, we get

$$\frac{V_t}{V_c} = \frac{\sqrt{V_1^2 + V_2^2 + V_3^2 + ...}}{V_c}$$

$$= \sqrt{\frac{V_1^2}{V_c^2}} + \sqrt{\frac{V_2^2}{V_c^2}} + \sqrt{\frac{V_3^2}{V_c^2}} ....$$

$$thatis$$

$$m_t = \sqrt{m_1^2 + m_2^2 + m_3^2 + ...}$$

2. To emphazise the total power in an am wave consists of carrier power and sideband power this yields

$$P_t = P_c\left(1+\frac{m^2}{2}\right) = P_c + \frac{P_c m^2}{2}$$

$$= P_c + P_{SB}$$

where P_{SB} is the total sideband power given by $P_{SB} + \frac{P_c m^2}{2}$

If several sine waves simultaneously modulate the carrier ,carrier power will be unaffected , but total sideband power will be sum of individual powers

$$P_{SB_T} = P_{SB_1} + P_{SB_2} + P_{SB_3} ....$$

Substituting gives

$$\frac{P_c m_t^2}{2} = \frac{P_c m_1^2}{2} + \frac{P_c m_2^2}{2} + \frac{P_c m_3^2}{2} + ...$$

$$m_t = \sqrt{m_1^2 + m_2^2 + m_3^2 + ...}$$

To calculate the modulation index, take the square root of sum of squares of individual modulation indices.

- Every transmitter is limited in the amount of power it can be used (more power means larger devices, too much power causes also interference with nearby stations).
- The receiver extracts the original information from the signal power that it receives.
- Greater the received power, easier it is to recover the desired signal.

The power in sidebands depends upon the value of modulation index. Greater percentage of modulation will yield a higher sideband power. Maximum power appears in the sidebands which is when the carrier is 100% (m=1) modulated. The power in each respective sideband, is given by

$$P_{USB} = P_{LSB} = \frac{P_c}{4}$$

This indicates that the power in each sideband is one-fourth, or 25 percent, of the carrier power. Since there are two sidebands, their power put together to give 50 percent of the carrier power.

DIFFERENT TYPES OF AM - MODULATORS:

There are various forms of Amplitude Modulation

1. Conventional Amplitude Modulation (Alternatively known as Full AM or Double Sideband Large carrier modulation(DSB-LC)
2. Double Sideband Suppressed carrier(DSB-SCS) modulation
3. Single Sideband (SSB) modulation
4. Vestigial Sideband (VSB) modulation

Double Sideband Large carrier modulation (DSB-LC) or Full AM

For the sake of simplicity we have let the phase of the carrier θ be zero (for AM no effect imposed on the phase but only with the amplitude)

Thus, the carrier signal above becomes

$$s_c(t) = A_c \cos(w_c t) \quad \text{where} w_c = 2\pi f_c$$

In the same way, a modulating signal (information signal) can also be expressed as $s_m(t)$.

(where $s_m(t) = A_m \cos w_m t$)

The amplitude-modulated wave can be expressed as,

$$s(t) = [A_c + s_m(t)]\cos(2\pi f_c t)$$

$$s(t) = [A_c + A_m \cos(2\pi f_m t)]\cos(2\pi f_c t)$$

The amplitude term of the AM signal $s(t)$ is

$$A = \left(A_c + A_m \cos(2\pi f_m t)\right)$$

$$= \left(A_c + mA_c \cos(2\pi f_m t)\right)$$

$$= A_c(1 + m\cos(2\pi f_m t))$$

Where, notation m in expression above is termed the modulation index.

Simply a measurement for the degree of modulation and bears the relationship of the ratio of A_m to A_c

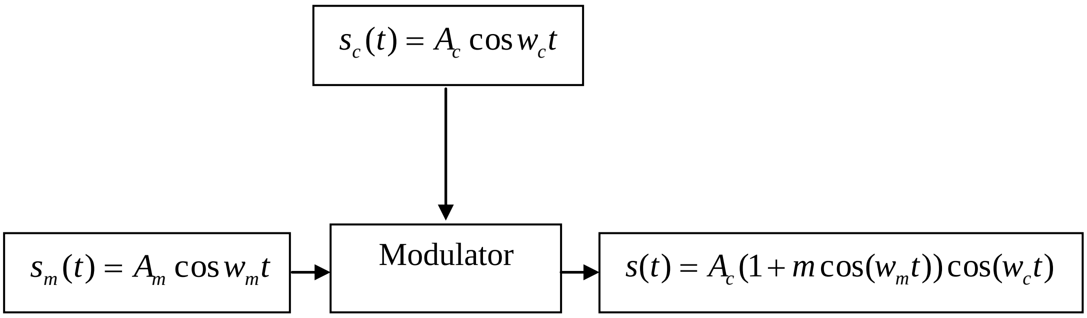
i.e. $m = \dfrac{A_m}{A_c}$

Therefore, The full AM signal may be written as

$$s(t) = A_c(1 + m\cos(w_m t))\cos(w_c t)$$

using: $\cos A \cos B = 1/2[\cos(A+B) + \cos(A-B)]$

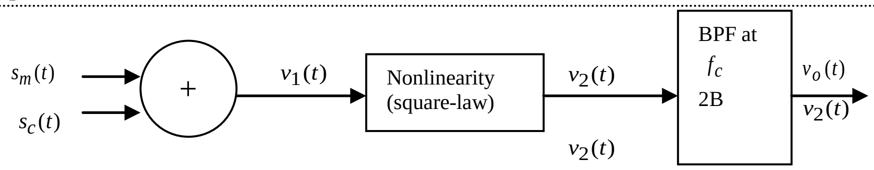
$$s(t) = \underbrace{A_c \cos w_c t}_{carrier} + \underbrace{\frac{mA_c}{2}\cos(w_c + w_m)}_{USB} + \underbrace{\frac{mA_c}{2}\cos(w_c - w_m)}_{LSB}$$



Both generation and detection require multiplication to be performed. The multiplication is achieved by using a network with a nonlinear characteristic. Nonlinear networks are not true multipliers because other components are produced and need to be filtered out.

Square-Law Modulator

Consists of a summer (summing the carrier and modulating signal), nonlinearity (square-law) block and a band pass filter (BPF) of bandwidth ($2B$) centered at f_c to extract the desired modulation products.



Square-Law Modulator

Square law of nonlinearity:

$$v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$$

where a_1 and a_2 are constants and $v_1(t)$ is the input voltage signal consists of the carrier plus the modulating signal

$$v_1(t) = s_c(t) + s_m(t) = A_c\cos(w_c t) + s_m(t)$$

Hence

$$v_2(t) = \underbrace{a_1 A_c(1 + \frac{2a_2}{a_1} s_m(t))\cos(w_c t)}_{\text{Desired AM signal}} + \underbrace{a_1 s_m(t) + a_2 s_m^2(t) + a_2 A_c^2 \cos^2 w_c t}_{\text{unwanted terms (removed by filtering)}}$$

By letting $a_1 = 1 \;\&\; a_2 = 1/2A_c$

$v_o = A_c\left(1 + m\cos(w_m t)\right)\cos(w_c t)$ ---Full AM signal

Double Side Band Suppressed Carrier (DSB-SC) Modulation

As noted early that carrier component in full AM or DSB-LC does no convey any information, it may be removed or suppressed during the modulation process to attain a higher power efficiency, hence Double Side Band Suppressed Carrier (DSB-SC) Modulation.

Consider the carrier

$$s_c(t) = A_c \cos(w_c t) \quad \text{where} \, w_c = 2\pi f_c$$

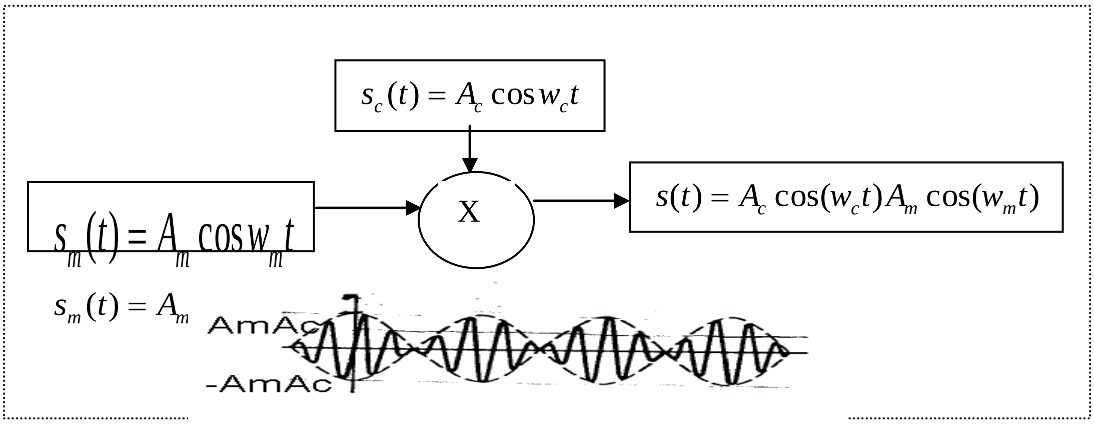
Modulated by a single sinusoidal signal

$$s_m(t) = A_m \cos w_m t \quad \text{where} \, \mathrm{w_m} = 2\pi f_m$$

The modulated signal is simply the product of these two

$$s\ (t) = A_c \cos(w_c t) A_m \cos(w_m t)$$

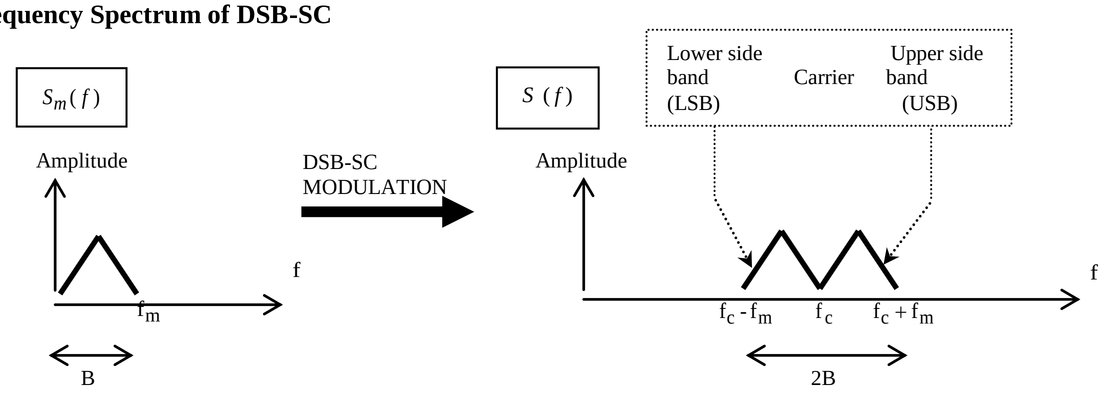
$$= \underbrace{\frac{A_m A_c}{2}\cos(w_c + w_m)t}_{USB} + \underbrace{\frac{A_m A_c}{2}\cos(w_c - w_m)t}_{LSB}$$



- The amplitude varies between the limits $\pm(A_m A_c)$

DSB-SC modulator & signal

Frequency Spectrum of DSB-SC



where B = Bandwidth

Frequency spectrum of DSB-SC

- All the transmitted power is contained in the two sidebands(no carrier present)
- The bandwidth is twice the modulating signal bandwidth.
- USB displays positive components of $s_m(t)$ and LSB displays negative components of $s_m(t)$.

The simplest method of generating a DSB-SC signal is merely to filter out the carrier portion of a full AM (or DSB-LC) waveform. Given carrier reference, modulation and demodulation (detection) can be implemented using product devices or balanced modulators.

SINGLE-SIDEBAND MODULATION

Single-sideband modulation (SSB) is a refinement of amplitude modulation that more efficiently uses electrical power and bandwidth. It is closely related to vestigial sideband modulation (VSB) (Amplitude modulation produces a modulated output signal that has twice the bandwidth of the original baseband signal. Single-sideband modulation avoids this bandwidth doubling, and the power wasted on a carrier, at the cost of somewhat increased device complexity.

SSB was also used over long distance telephone lines, as part of a technique known as frequency-division multiplexing (FDM). FDM was pioneered by telephone companies in the 1930s. This enabled many voice channels to be sent down a single physical circuit, for example in L-carrier. SSB allowed channels to be spaced (usually) just 4,000 Hz apart, while offering a speech bandwidth of nominally 300–3,400 Hz. Single-sideband suppressed-carrier (SSB-SC) is a telecommunication technique, which belongs to amplitude modulation class.

The information represented by the modulating signal is contained in both the upper and the lower sidebands. Since each modulating frequency f_c produces corresponding upper and lower side-frequencies $f_c + f_i$ and $f_c - f_i$

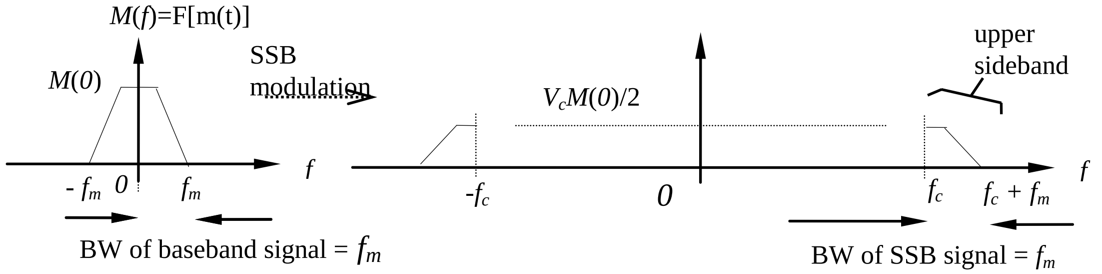
It is not necessary to transmit both side-bands. Either one can be suppressed at the transmitter without any loss of information.

Advantages

- Less transmitter power.
- Less bandwidth, one-half that of Double-Sideband (DSB).
- Less noise at the receiver.
- Size, weight and peak antenna voltage of single-sideband (SSB) transmitters is significantly less than that of a standard AM transmitter.

SINGLE-SIDEBAND MODULATION, SSB

When one (upper or lower) of the two sidebands of DSB-SC signal is removed before transmission, we get SSB modulation.



- **SSB = DSB - SB**

- **By transmitting only either upper side band or lower side band, the original baseband message can still be recovered**
- **Power transmitted,** $P_t = \frac{1}{2} P_{t(DSB-SC)} = \frac{1}{2} P_c P_m$

Note that conventional amplitude modulation (Full AM) and DSB-SC modulation require a transmission bandwidth equal to twice the information signal bandwidth (B = 2W). One half the transmission bandwidths is occupied by the upper sideband of the modulated signal. Whereas the other half is occupied by the lower sideband, the basic information is transmitted twice, once in each sideband. Since the sidebands are the sum and difference of the carrier and modulating signals, the information must be contained in both of them. There is absolutely no reason to transmit both sidebands in order to convey the information. One sideband may be suppressed. The remaining sideband is called a single-sideband suppressed carrier (SSSC or SSB) signal.

Advantages of SSB

i) The spectrum space occupied by the SSB signal is only half that of AM and DSB signals. This greatly conserves spectrum space and allows more signals to be transmitted in the same frequency range. It also means there should be less interference between signals.

ii) The second benefit is that all the power previously devoted to the carrier and other sideband can be channeled into the single sideband, thereby producing a stronger signal that should carry farther and be more reliably received at greater distances.

SUPPRESSION OF CARRIER

Effect of non linear resistance on added signals

The relationship voltage and current in a linear resistance is given by i=bv

where b is proportionality constant.

If the above equation refers to resistor, then b is conductance. If this is made to apply to collector current and base voltage of transistor, then I will be collector current. If amplifier operates in class A, there will be a dc component in collector current, which is not dependent on signal voltage at base.

$$i=a+bv$$

where, a is dc component of collector current ,b is transconductance of transistor

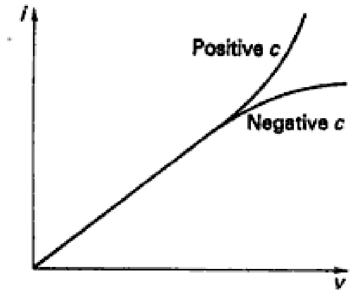
The three practical methods of SSB generation use the balanced modulator to suppress the carrier, but each uses a different method of removing the unwanted sideband.

All three systems will remove either the upper or lower sideband with equal ease, depending on the specific circuit arrangement. The non linear relation is most conveniently expressed by

$i=a+bv+cv^2+dv^3$+higher powers. Neglecting higher powers we get

$$i=a+bv+cv^2$$

c is coefficient of non linearity.



Non-linear resistance characteristics

The above equation is generally adequate in relating output current to input voltage of non linear resistance. It may be noted that gate voltage and drain current characteristic of FET of two voltages are applied simultaneously to gate then

$$i=a+b(v_1+v_2)+c(v_1+v_2)^2$$

$$=a+b(v_1+v_2)+c(v_1^2+v_2^2+2v_1v_2)$$

Let two input voltages be sinusoidal. Then v_1 is replaced by $V_1\sin\omega t$ and v_2 is replaced by $V_2\sin\rho t$ where ω and ρ are two angular velocities. Substituting in above equation we get

$$i = a+b(V_1\sin\omega t + V_2\sin\rho t) + c(V_1^2\sin^2\omega t + V_2^2\sin^2\rho t + 2V_1\sin\omega t\,.V_2\sin\rho t)$$

By applying trigonometric formula the equation is rewritten

$i = a+b(V_1\sin\omega t + V_2\sin\rho t) + 1/2\, V_1^2\, c\,(1-\cos 2\omega t) + 1/2\, V_2^2\, c\,(1-\cos 2\rho t) + c\, V_2V_1\,[\cos(\omega-\rho)t - \cos(\omega+\rho)t]$,

rearranging the equation we get ,

$$i = \underset{(I)}{(a+1/2\,c\,V_1^2+1/2\,c\,V_2^2)} + \underset{(II)}{b\,V_1\sin\omega t} + \underset{(III)}{b.\,V_2\sin\rho t} - \underset{(IV)}{(1/2\,V_1^2\,c\cos 2\omega t + 1/2\,V_2^2\,c\cos 2\rho t)}$$

$$+\underset{(V)}{c\,V_2V_1\cos(\omega-\rho)t} - \underset{(VI)}{c\,V_2V_1\cos(\omega+\rho)t}$$

If in above equation ω is carrier angular frequency and ρ is modulating angular frequency , term I is the dc component ,term II is the carrier, and term III is modulating signal, term IV is the harmonics of carrier and modulation, term V is the lower side band and term VI is the upper side band

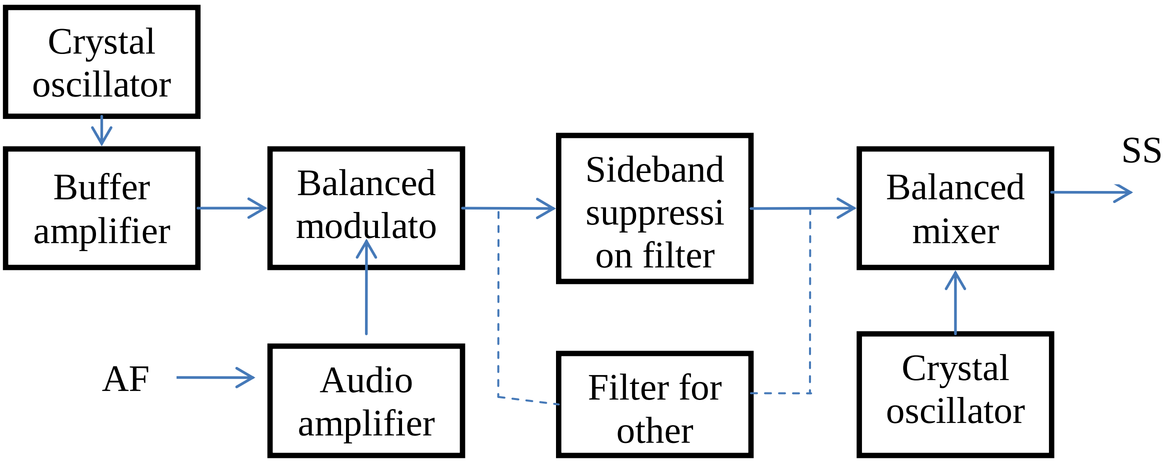
$$v_1{+}v_2T_2$$

It is a proof that

1. Harmonics and intermodulation occurs in audio and RF amplifiers.
2. Sum and difference frequencies are present in mixer.
3. Diode detector has audio frequencies in output
4. It is operation of beat frequency oscillator and product detector.
5. It is a part of proof that balanced modulator produces AM with carrier suppressed

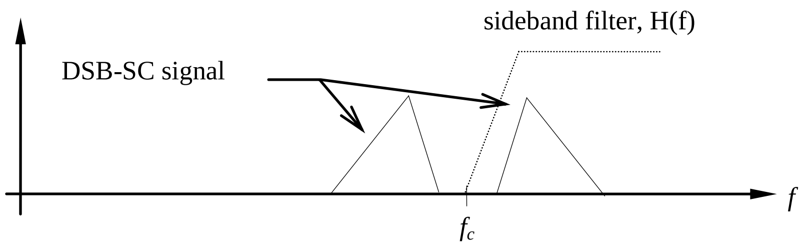
SUPPRESSION OF UNWANTED SIDEBAND

The filter system



Filter method of side suppression

SSB signal generation using filtering method



- The filter system is the simplest system of the three – after the balanced modulator the unwanted sideband is removed (actually heavily attenuated) by a filter. The filter may be LC, crystal, ceramic or mechanical, depending on the carrier frequency and other requirements.
- The key circuits in this transmitter are the balanced modulator and the sideband-suppression filter. Such a filter must have a flat bandpass and extremely high attenuation outside the bandpass. There is no limit on this; the higher the attenuation, the better.
- In radio communication systems, the frequency range used for voice is 300 to about 2800 Hz in most cases. If it is required to suppress the lower sideband and if the transmitting frequency is f, then the lowest frequency that this filter must pass without attenuation is f+300Hz, whereas the highest frequency that must be fully attenuated is f-300Hz.

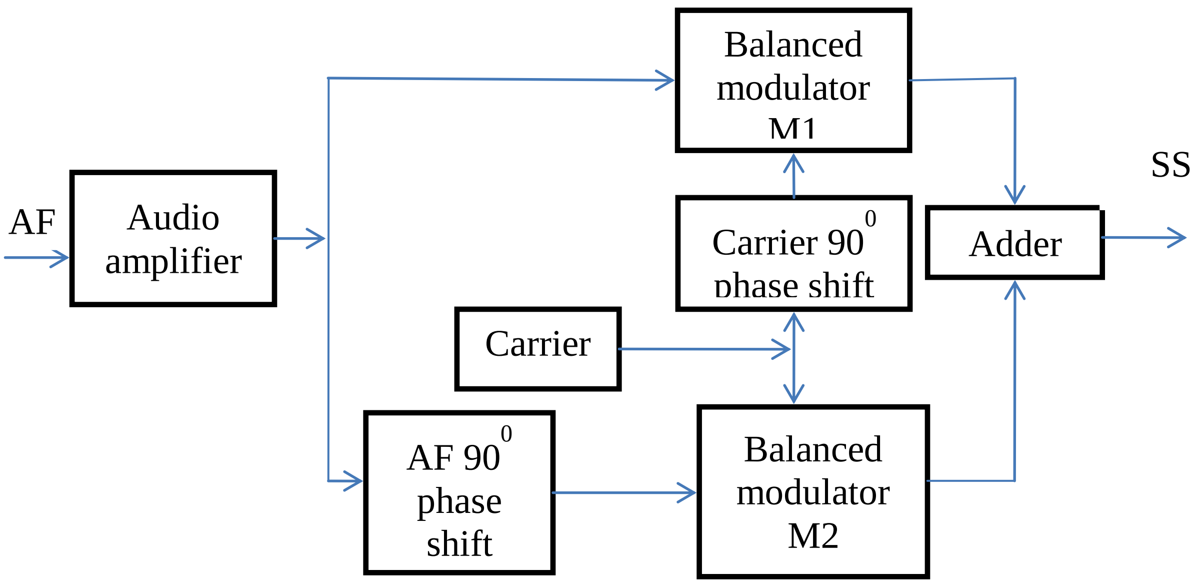
In other words, filter response must change from zero attenuation to full attenuation over a range of only 600 hertz. If the transmitting frequency is much above 10 MHz, this is not applicable. The situation become even worse if lower modulating frequencies are employed, such as the 50-Hz minimum in AM broadcasting. In order to obtain a filter response curve with skirts as steep as those suggested above, the Q of the tuned circuits used must be very high. As the transmitting frequency is raised, a situation is reached where necessary Q is so high that there is no practicable method of achieving it.

Mechanical filters have been used at frequencies up to 500 kHz, and crystal or ceramic filters up to about 20 MHz. The mechanical filter seem to be best because

- Small size
- Good bandpass
- Very good attenuation
- Adequate upper frequency limit

- Crystal or ceramic filters may be cheaper but preferable for above 1 MHz. The disadvantages of these filters are their operating frequency is below usual transmitting frequencies.
- This is a reason for using a balanced mixer. In this mixer, frequency of crystal oscillator or synthesizer is added to SSB signal from the filter, frequency thus being raised to the value desired for transmission.
- If transmitting frequency is much higher than operating frequency of sideband filter, two stages of mixing will be required. It becomes too difficult to filter out unwanted frequencies in the output of mixer.
- The mixer is followed by a linear amplifier because the amplitude of SSB signal is variable. A class B RF amplifier is used because it is more efficient than class A amplifier.

Phase shift method



Single side band by phase shift method

- The phase shift method avoids filters and some of their inherent disadvantages, and instead makes use of two balanced modulators and two phase shifting networks.
- One of the modulators M1 receives the carrier voltage (shifted by 90^0) and the modulating voltage, whereas the other, M2, is fed the modulating voltage (shifted through 90^0) and the carrier voltage.
- Sometimes the modulating voltage phase shift is arranged slightly differently. It is made +45` for one of the balanced modulators and -45` for the other, but the result is the same.
- Both modulators produce an output consisting only of sidebands. It will be shown that both upper sidebands lead the input carrier voltage by 90^0.
- One of the lower sidebands leads the reference voltage by 90^0. The two lower sidebands are thus out of phase, and when combined in the adder, the cancel each other.
- The upper sidebands are in phase at the adder and therefore add, giving SSB in which lower sideband has been canceled. If it is taken for granted that the two balanced modulators are also balanced with respect to each other, then amplitudes may be ignored as they do not affect the result and it is fed with same sources.

Let $\sin \omega_c t$ be the carrier and $\sin \omega_m t$ be the modulation, balanced modulator M_1 will receive $\sin \omega_m t$ and $\sin (\omega_c t + 90^0)$ and balanced modulator M_2 will receive $\sin \omega_c t$ and $\sin (\omega_m t + 90^0)$.

The output of M_1 will contain sum and difference $v_1 = \sin \omega_m t . \sin (\omega_c t + 90^0)$

$$v_1 = \cos\left[\left(\omega_c t + 90^0\right) - \omega_m t\right] - \cos\left[\left(\omega_c t + 90^0\right) + \omega_m t\right]$$

$$= \cos\left(\omega_c\ t - \omega_m\ t + 90^0\right) - \omega_m\ t - \cos\left(\omega_c\ t + \omega_m\ t + 90^0\right) \text{ ---------------------------- (1)}$$

The output of M_2 will contain $v_1 = \cos\left[\omega_c t - \left(\omega_m\ t + 90^0\right)\right] - \cos\left[\omega_c t + \left(\omega_m\ t + 90^0\right)\right]$

$$= \cos\left(\omega_c\ t - \omega_m\ t - 90^0\right) - \cos\left(\omega_c\ t + \omega_m\ t + 90^0\right) \text{------------}$$

(2)

Adding 1 and 2 we get

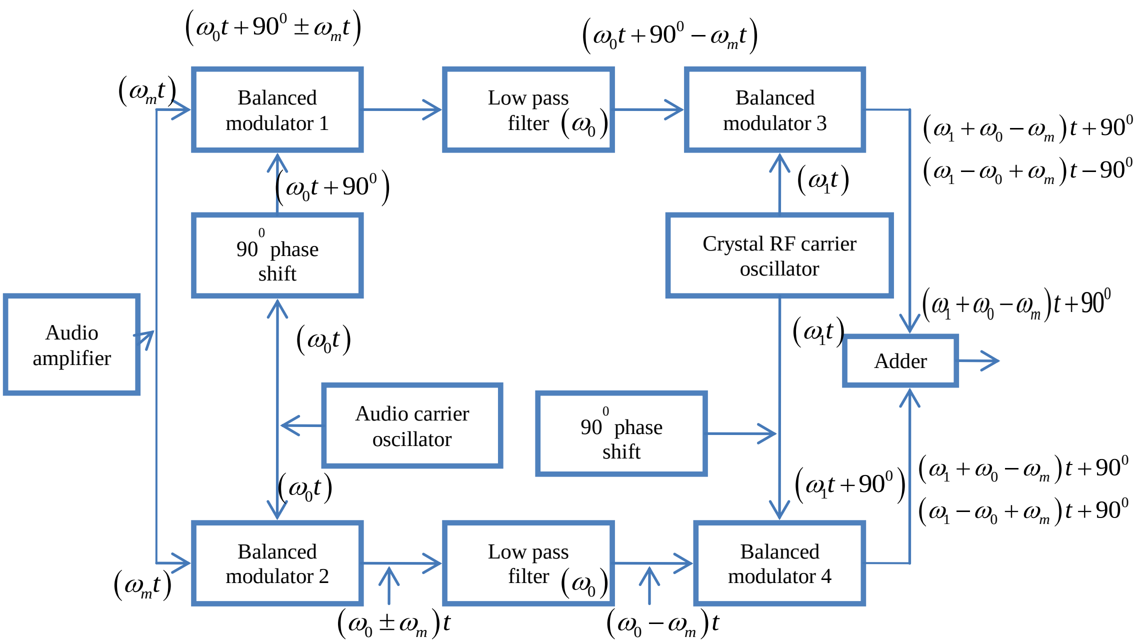
The output adder is $v_{o=} v_1 + v_2 = 2\cos\left(\omega_c\ t + \omega_m\ t + 90^0\right)$

The first term of first equation is 180^0 out of phase with first term of second equation. Thus one of the side band is cancelled in the adder. The system yields upper side band and it is reinforced.

THIRD METHOD

- The third method of generating single sideband suppressed carrier was developed by Weaver.

- It is similar to that of phase shift method but it differs in that modulating signal is first modulated on a low frequency carrier including phase shifts which is then modulated onto high frequency carrier.
- Balanced modulator 1and 2 have un-shifted modulating signal as inputs.
- Balanced modulator 1 takes low frequency subcarrier with a 90^0 phase shift introduced from oscillator signal.
- Balanced modulator 2 takes subcarrier signal directly from oscillator. Assuming unity magnitude and co-sinusoidal single frequency modulation



Single side band by third method

The output from balanced modulator 1becomes

$$\begin{aligned} e_{BM1} &= \cos\left(\omega_0 t + 90^0\right)\cos\omega_m t \\ &= \frac{1}{2}\left[\cos\left(\omega_0 t + \omega_m t + 90^0\right) + \cos\left(\omega_0 t - \omega_m t + 90^0\right)\right] \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

The output from balanced modulator 2 becomes

$$\begin{aligned} e_{BM2} &= \cos\omega_0 t\cos\omega_m t \\ &= \frac{1}{2}\left[\cos\left(\omega_0 t + \omega_m t\right) + \cos\left(\omega_0 t - \omega_m t\right)\right] \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

- Low pass filter with cut off frequency set at subcarrier frequency f_o removes first term from each of the above signal leaving difference term as input to BM 3 and BM 4. These signals are lower sideband on f_o.

- They are identical except the signal applied to BM 3 is shifted by $+90^0$ from that it is applied to BM 4.this process eliminates the need to provide a wideband 90^0 phase shifting network for base band signals. The high frequency oscillator signal at f_1 is applied to BM 3 but shifted by $+90^0$ before being applied BM 4.

  The output of BM 3 becomes

$$\begin{aligned} e_{BM3} &= \cos\omega_1 t\cos\left[\left(\omega_0 - \omega_m\right)t + 90^0\right] \\ &= \frac{1}{2}\left[\cos\left(\left(\omega_1 + \omega_0\right)t - \omega_m t + 90^0\right) + \cos\left(\left(\omega_1 - \omega_0\right)t + \omega_m t - 90^0\right)\right] \end{aligned} \quad \ldots\ldots\ldots (3)$$

The output of BM 4 becomes

$$\begin{aligned} e_{BM3} &= \cos\left(\omega_1 t + 90^0\right)\cos\left(\omega_0 - \omega_m\right)t \\ &= \frac{1}{2}\left[\cos\left(\left(\omega_1 + \omega_0\right)t - \omega_m t + 90^0\right) + \cos\left(\left(\omega_1 - \omega_0\right)t + \omega_m t + 90^0\right)\right] \end{aligned} \quad \ldots\ldots\ldots (4)$$

Output Spectra For Third Method a)LSB And b)USB

- The first terms of equation 3 and 4 are identical lower sidebands on an offset carrier frequency $f_c=f_1+f_0$.The second term are upper side band on an offset carrier at $f_c=f_1-f_0$.but are 180^0 out of phase with each other.
- The output of BM 3 and BM 4 are added in summing amplifier to produce final output. The first two terms get added and the second two get cancels leaving the output.

$$e_{out} = \cos\left(\left(\omega_1 + \omega_0 - \omega_m\right)t + 90^0\right)$$

If the output from BM 3 and BM 4 is inverted before input to adder, phasing becomes reverse such that first term, cancel and the second term add giving upper sideband on the carrier frequency.

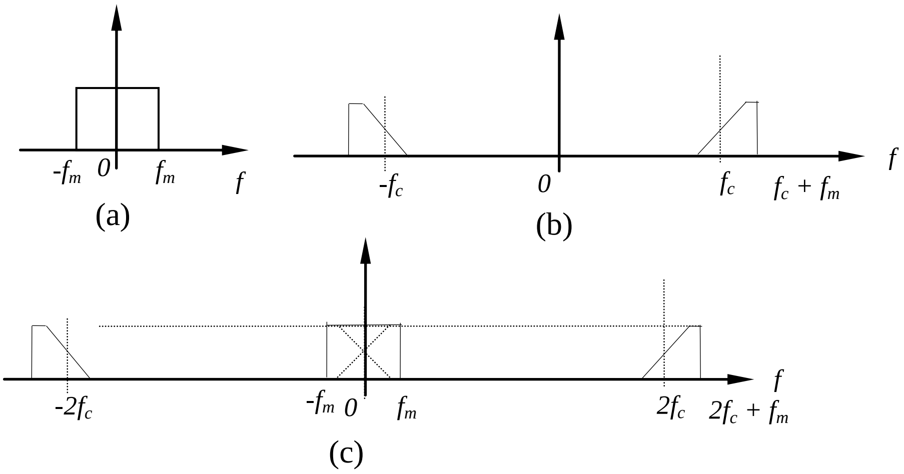
VESTIGIAL SIDEBAND (VSB)

1. If portion or one of the (upper or lower) sideband of AM signal is removed, we will get VSB signal.
   a) AM :

$$s_{AM}(t) = V_c\left[1 + m\cos\omega_m t\right]\cos\omega_c t$$
$$= V_c\cos\omega_c t + \frac{mV_c}{2}\left[\cos(\omega_c + \omega_m)t + \cos(\omega_c - \omega_m)t\right]$$

VSB Spectrum

(a) Message
(b) Modulated signal
(c) Frequency translated signal before LPF



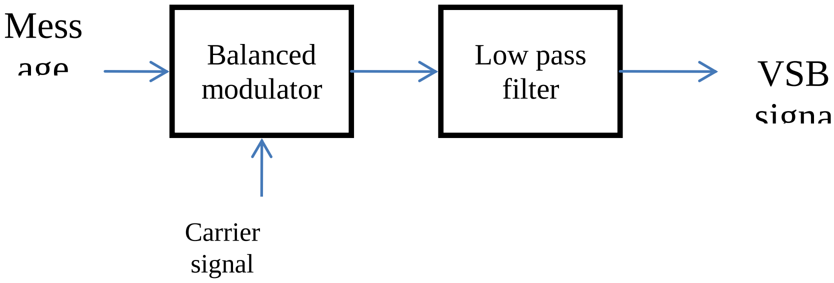
VSB: $s_{VSB}(t) = V_c\cos\omega_c t + \frac{mV_c}{2}\cos(\omega_c + \omega_m)t$

Average transmitted power, $P_t = P_c + \frac{1}{2}P_m P_c$

In general, $P_t = P_c + kP_m P_c$ where $0.5 < k < 1$

Example: suitable for signal with significant low-frequency contents

Generation of VSB

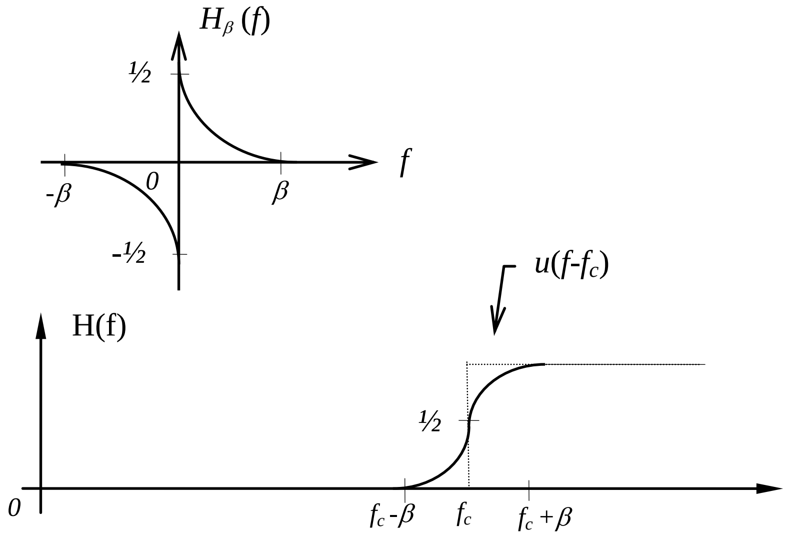


Generation of VSB signal

a) VSB is derived by filtering AM (with VSB filter) in such a fashion that one sideband is passed almost completely while just a trace, or vestige, of the other sideband is included.

b) VSB filter should have the following characteristics:

VSB Filter $H(f) = u(f - f_c) - H_\beta(f - f_c)$



i) odd symmetry about carrier frequency
ii) a relative response of ½ at f_c

VSB

SSB have good bandwidth efficiency, but practical SSB modulation systems have poor low frequency response. This is unpleasant when message (modulating) signal bandwidth is wide or where one cannot disregard the low frequency component. Whereas, DSB-SC works well for baseband signals with significant low frequency content but has a wider bandwidth and higher power than SSB.

VSB gets over this problem while retaining the advantages of SSB(compromise solution between SSB &DSB-SC). VSB relaxes the stringent sharp cutoff requirements of SSB by retaining a part (**vestige**) of the unwanted sideband in the transmitted signal.

- Offers a compromise between SSB and DSB-SC
- VSB has lower power less bandwidth than full AM and higher power and slightly greater bandwidth than SSB.
- VSB is standard for transmission of TV and similar signals(□TV has low frequency component)
- Bandwidth saving can be significant if modulating signals are of large bandwidth as in TV and wideband data signals.
- For example with TV the bandwidth of the modulating signal can extend up to 5.5MHz; with full AM the bandwidth required is 11MHz this is excessive from transmission bandwidth occupancy and cost points of view.

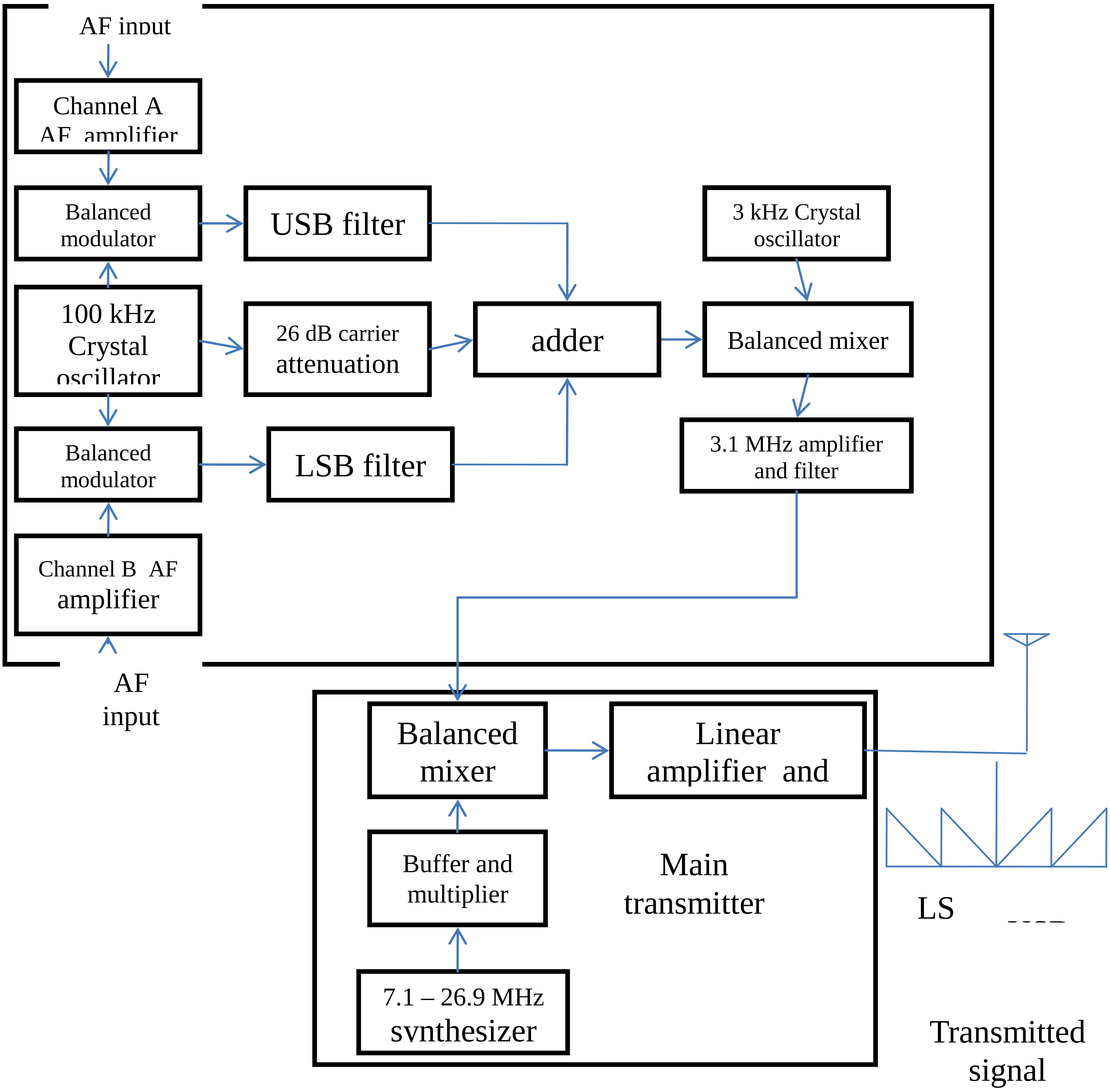
INDEPENDENT SIDEBAND

Independent sideband (ISB) is an AM single sideband mode which is used with some AM radio transmissions. Normally each sideband carries identical information, but ISB modulates two different input signals — one on the upper sideband, other on lower sideband. This is used in some kinds of AM stereo (sometimes known as the Kahn system).

ISB is a compromise between double sideband (DSB) and single sideband (SSB) — the other is vestigial sideband (VSB). If the sidebands are out of phase with each other, then phase modulation (PM) of the carrier occurs. AM and PM together then create quadrature amplitude modulation (QAM). ISB may or may not have the carrier suppressed.

Suppressed-carrier ISB was employed in point-to-point (usually overseas) radiotelephony and radio teletype by shortwave (HF). In military use, ISB usually referred to a close pair of FSK radio teletype channels which could be demodulated by a single receiver, and employed in fleet broadcast, point-to-point, and between larger vessels and shore stations on HF and UHF

ISB essentially consists of R3E with two SSB channels added to form tow sidebands around the reduced carrier .Each sideband is quite independent of the other. It can simultaneously convey a totally different transmission, to the extent that the upper sideband could be used for telephony while the lower sideband carries telegraphy.



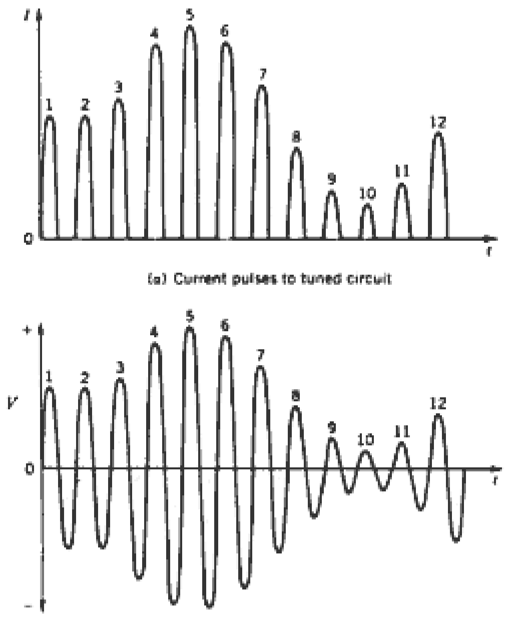
ISB Transmitter

Each 6-kHs channels is fed its own balanced modulator ,each balanced modulator also receiving the output of the 100-kHz crystal oscillator .The carrier is suppressed in the balanced modulator and the following filter, the main function of the filter still being the suppression of the unwanted side band ,as in all other SSB system .The difference here is that while one filter suppresses the lower sideband the other suppresses the upper sideband .Both outputs are then combined in the adder with the -26-dB carrier, so that a low –frequency ISB signal exists at this point , with a pilot carrier also present .Through mixing with the output of another crystal oscillator ,the frequency is then raised to the standard value of 3.1MHZ.Note the use of balanced mixers, to permits easier removal of unwanted frequencies by the output filter.

The signal now leaves the drive unit and enters the main transmitter. Its frequency is raised yet again, through mixing with the output of another crystal oscillator, or frequency synthesizer. This is done because the frequency range for such transmissions lies in the HF band from 3 to 30MHz.The resulting RF ISB signal is then amplified by linear amplifiers,

GENERATION OF AM

Flywheel effect



(a) Current pulses to tuned circuit

(b) Tuned-circuit AM voltage.

In order to generate AM wave it is necessarily merely to apply the series of current pulses to a tuned circuit. Each pulse, if it were the only one would initiate a damped oscillation in tuned circuit. The oscillations would have initial amplitude proportional to the size of current pulse and decay rate dependent on time constant of circuit. Since a train of pulse is fed into the tank circuit, each pulse will cause a complete sine wave proportional to size of next applied pulse At least 10times as many pulses per audio cycle is fed in to a practical circuit. the process is known as fly wheel of tuned circuit

It works best with a tuned circuit whose Q is not low.

The flywheel effect is the continuation of oscillations in an oscillator circuit after the control stimulus has been removed. This is usually caused by interacting inductive and capacitive elements in the oscillator. Circuits undergoing such oscillations are said to be fly wheeling.

The flywheel effect may be desirable, such as in phase-locked loops used in synchronous systems, or undesirable, such as in voltage-controlled oscillators.

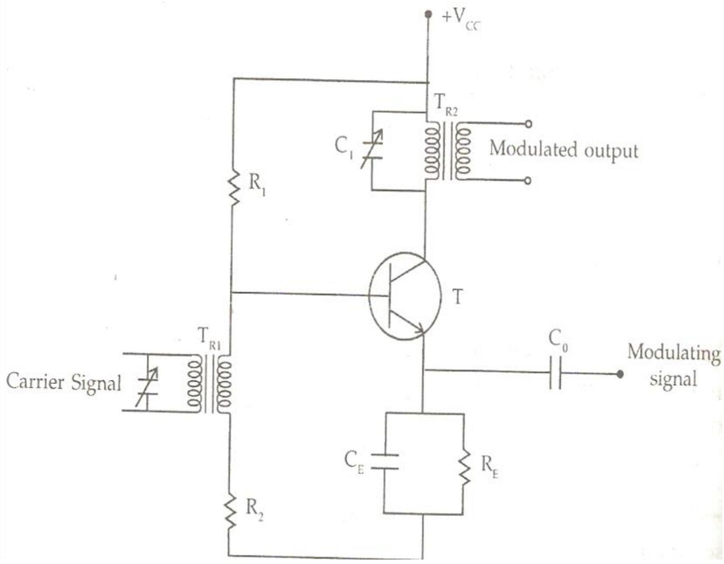
Flywheel effect is used in class C modulation where efficiency of modulation can be achieved as high as 90%

DIFFERENT TYPES OF MODULATOR CIRCUITS

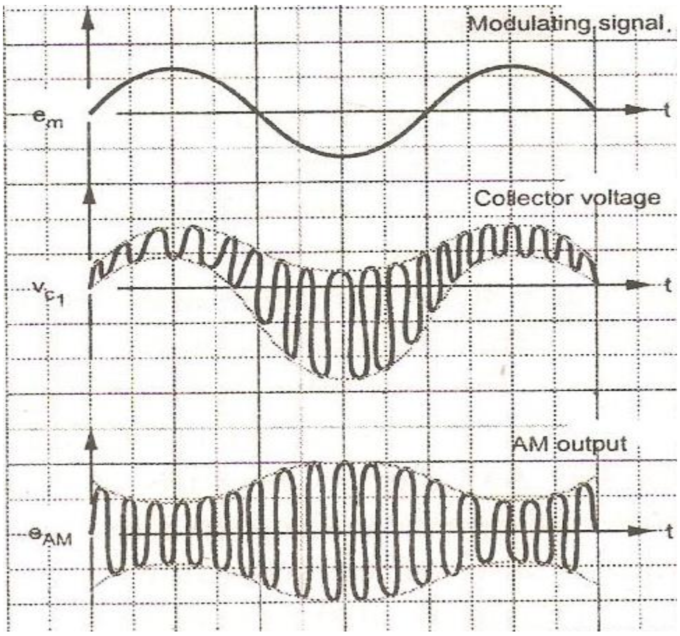
Emitter modulation

A small signal class A amplifier can be used to perform amplitude modulation. The amplifier must have two inputs one for carrier signal which is applied in base and the other is modulating signal which is applied in emitter and hence the name emitter modulation. When modulating signal is not applied, circuit operates in linear class A amplifier and the output is simply carrier amplified circuit by quiescent voltage gain. When modulating signal is applied amplifier operates non-linearly and the signal multiplication occurs. The modulating signal varies the gain of amplifier at sinusoidal rate equal to that of modulating signal.

The modulating signal is applied through isolation transformer T_1 to emitter of $Q_1$and carrier signal is applied to base. The modulating signal drives the circuit into both saturation and cutoff producing non linear amplification necessary for modulation to occur the collector waveform produces carrier, upper and lower frequencies as well as components at modulating signals frequency. In emitter modulation amplitude of output signal depends on amplitude of carrier and voltage gain of amplifier. The modulation index depends on the amplitude of modulating signal.



Circuit of emitter modulation



Waveform of emitter modulation

The voltage gain for emitter modulator is given by $A_v = A_q\left[1 + m\sin\omega_m t\right]$

where A_v amplifier voltage gain with modulation

A_q amplifier quiescent voltage gain without modulation

$\sin\omega_m t$ goes from a maximum value of +1 to minimum value of -1 the above equation reduces to $A_v = A_q\left[1 \pm m\right]$ where m is modulation index.At 100% modulation m=1 and the equation reduces to

$$A_{\max} = 2A_q$$
$$A_{\min} = 0$$

Advantages

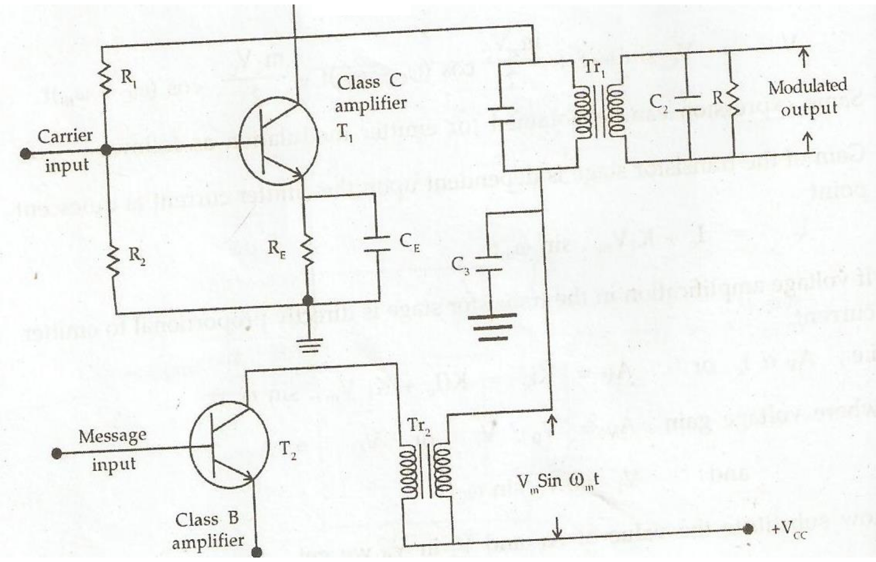
1. Less modulating signal power for to obtain modulation
2. Modulator circuit to be designed at low power

Disadvantages

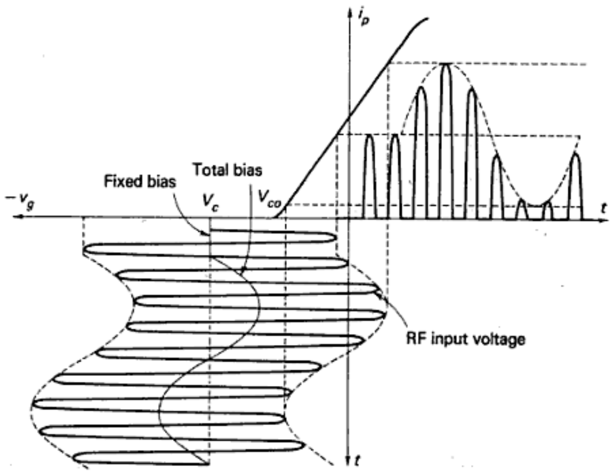
1. Amplifier operated in class A mode
2. Efficiency is low
3. Output power is very small

Collector modulation

The amplifier must have two inputs one for carrier signal which is applied in base and other is modulating signal which is applied in collector and hence the name collector modulation. Modulation takes place in collector, which is the output element of transistor. Therefore, if this is final of transmitter, it is a high level modulator. To achieve high power efficiency, medium and high power AM modulator generally operates at class C. A practical efficiency of 80% is possible. Class C amplifier with two inputs carrier (v_c) and modulating signal (v_m). Because the transistor is in biased class C, it operates non linear and is capable of non linear mixing (modulation).the RFC that acts as a short to dc and open to high frequencies. Therefore RFC isolates dc power supply from high frequency carrier and side frequencies.



Circuit of collector modulation



Waveform of collector modulation

When amplitude of carrier exceeds barrier potential of base emitter junction T_1 turns on and collector current flows. When amplitude of carrier drops barrier potential T_1 turns off and collector current ceases. Consequently Q_1, switches between saturation and cutoff controlled by carrier signal, collector current flows less than 180^0 of each carrier cycle, and class C operation is achieved. Each successive cycles of carrier turns on Q_1 for a instant of time and allows current to flow for short time producing a negative going waveform at collector.

Since Q_1 operating in non linear, collector waveform contains two original input frequencies (f_c and f_m) and their sum and difference frequencies ($f_c \pm f_m$). Because output waveform contains higher order harmonics and intermodulation distortion

Advantages

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

Disadvantages

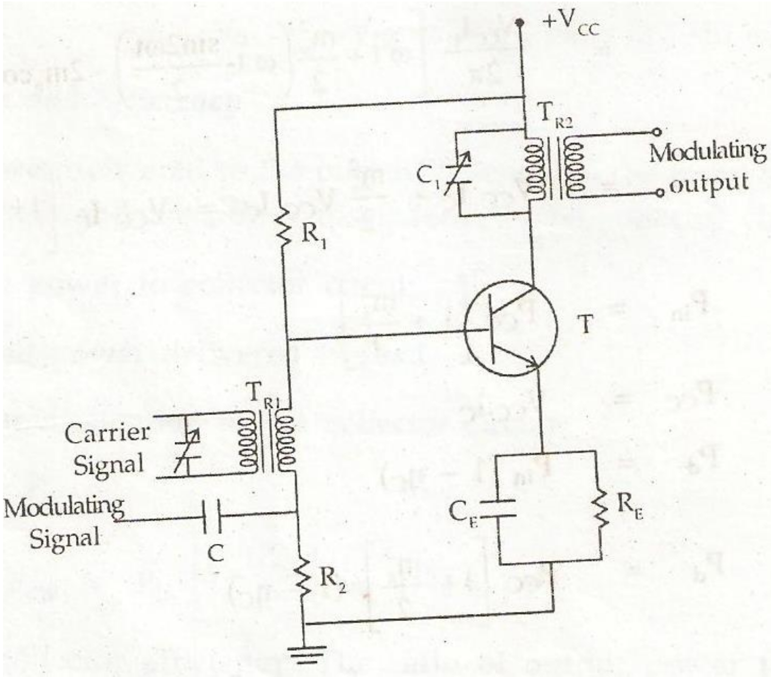
1. Large modulating power is required for high power applications.
2. Collector saturation prevents 100% percent modulation

Base modulation

As the name implies that, in this modulation, the modulating signal is injected into the base circuit of the transistor modulator to reduce the power level of the applied signal. A common emitter configuration is employed and the transistor is biased in class ‘C’mode, the resistors R_1 and R_2 provides potential divider biasing for the transistor through V_{cc} the resistor R_e and capacitor C_e acts as temperature stabilization elements

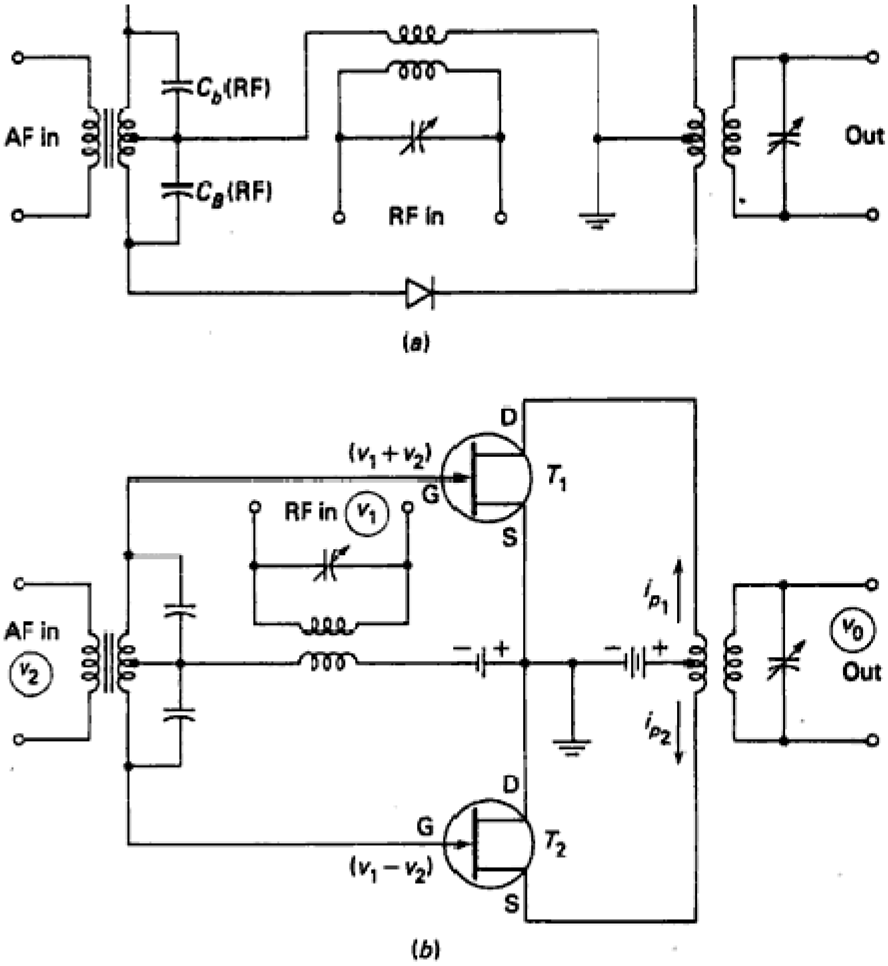
The message signal applied to the base circuit causes the amplitude of the carrier to vary between cutoff and saturation regions in order to produce the fully modulated output.

The gain of the circuit cannot be maintained as constant over the entire range of dynamic characteristics; hence the fully modulated output is not linearly modulated



Base modulation

1. **Balanced modulator**



Balanced modulator a) diode b) FET

The two circuits of balanced modulator utilize nonlinear principles. The modulation voltage v_2 is fed with push pull and carrier $v_1$is fed in parallel, to a pair of identical diodes or class A (transistor or FET) amplifiers. In FET circuit carrier voltage is applied to gates in phase whereas modulating voltage appears 180^0 out of phase since they are opposite ends of center tapped transformer. The modulated output currents of FET are combined in center tapped primary of push pull output transformer. They therefore subtract as indicated by direction of arrows. If the system is made symmetrical carrier frequency will be completely cancelled. No system is symmetrical in practice so the carrier is heavily suppressed rather than removed. The output of balanced modulator contains side band with miscellaneous components.

The input voltage will be v_1+v_2 at gate T_1 and v_1-v_2 at gate T_2. If perfect symmetry is assumed proportionality constant will be same for both FET and may be called a, b and c. The drain currents will be given by

$i_{d1}=a +b (v_1+v_2) +c (v_1+v_2)^2$

$=a+ b (v_1+v_2) +c (v_1^2+v_2^2+2 v_1v_2)$ ……………….. (1)

$i_{d2}=a +b (v_1+v_2) +c (v_1-v_2)^2$

$=a+ b (v_1+v_2) +c (v_1^2+v_2^2-2 v_1v_2)$ ……………….. (2)

The primary current is given by difference between individual drain currents $i_1= i_{d1}- i_{d2}=2bv_2+4cv_1v_2$

The carrier voltage v_1 is represented by $V_c \sin \omega_c t$ and modulating voltage v_2 is represented by $V_m \sin \omega_m t$ and substituting in above equation

$i_1= 2b V_m \sin \omega_m t +4c V_c \sin \omega_c t V_m \sin \omega_m t$

$=2b V_m \sin \omega_m t +4c V_c V_m ½[\cos (\omega_c - \omega_m)t - \cos (\omega_c + \omega_m)t]$

The output voltage v_0 is proportional to primary constant. Let it be α then

$v_0=\alpha i_1$

$=2\alpha b V_m \sin \omega_m t +2\alpha c V_c V_m [\cos (\omega_c - \omega_m)t - \cos (\omega_c + \omega_m)t]$

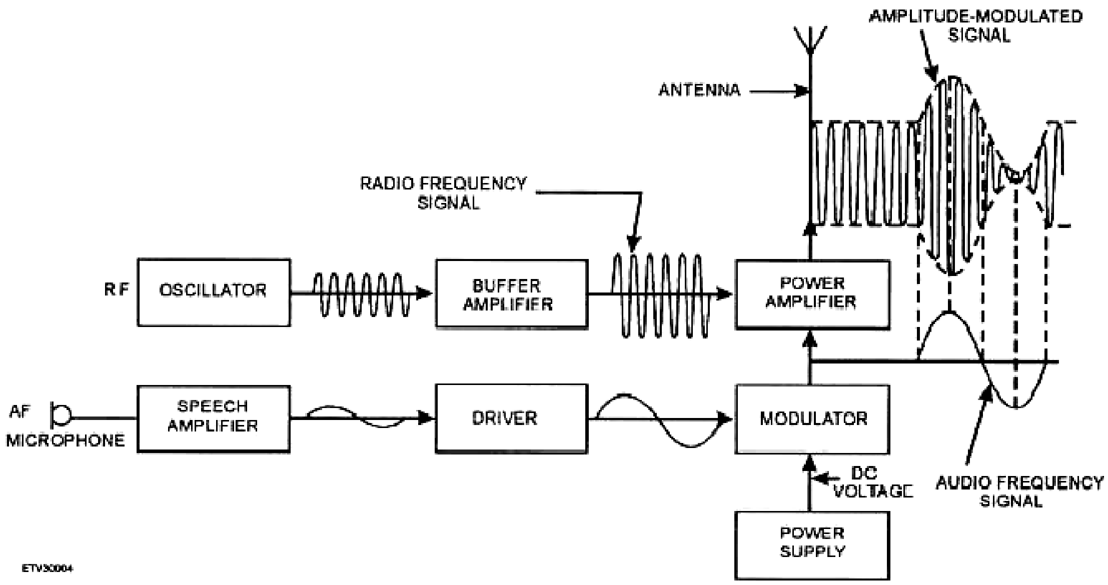
Simplifying we get $P=2\alpha b V_m$ and $Q=2\alpha c V_c V_m$ then

$v_0=P \sin \omega_m t + Q \cos (\omega_c - \omega_m)t - Q \cos (\omega_c + \omega_m)t$

Modulation Lower sideband upper sideband

The obtain equation contain modulating frequencies and two side bands the carriers cancels out each other.

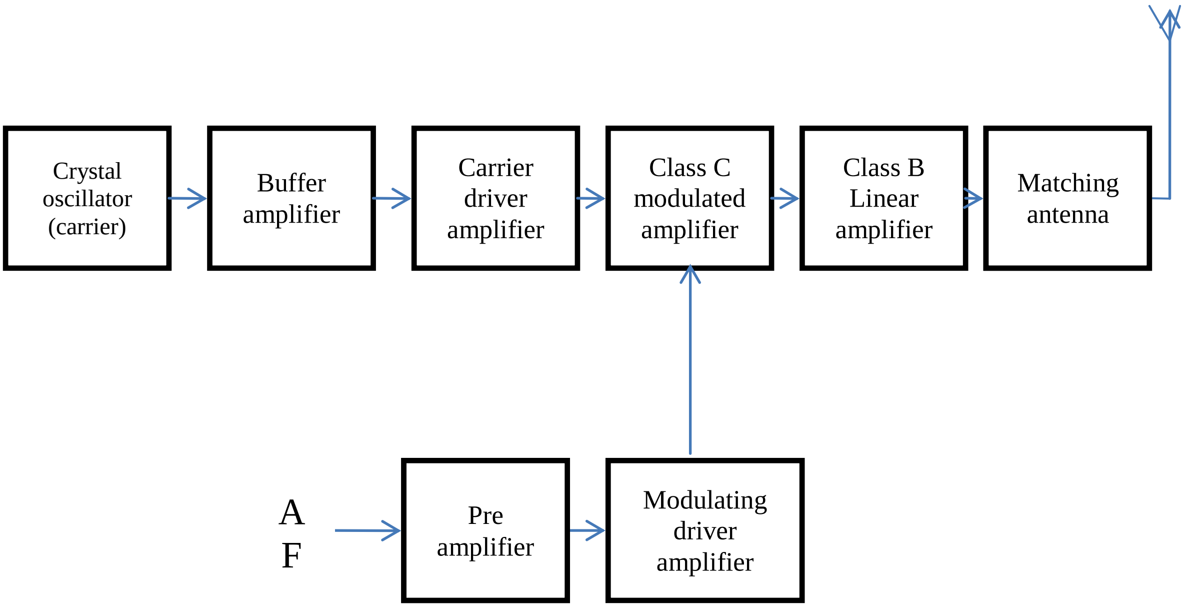
AM TRANSMITTERS



The block diagram is a simple AM transmitter. The microphone converts the audio frequency input to electrical energy. The driver and modulator amplify the audio signal to the level required to modulate the carrier fully. The signal is then applied to power amplifier. Power amplifier combines the RF carrier and the modulating signal to produce the AM signal for transmission. In SSB communications, carrier is suppressed (eliminated) and the sideband frequencies produced by the carrier are reduced to a minimum. This means no carrier is present in the transmitted signal. It is removed after the signal is modulated and reinserted at the receiver during demodulation. Since there is no carrier, all the energy is concentrated in the side- band(s)

LOW LEVEL TRANSMITTER

If modulation takes place at early stage of transmitter i.e., in between modulator and transmitter antenna one or more amplifier are connected because carrier power is low .A small audio stage is used to modulate a low power stage; the output of this stage is then amplified using a linear RF amplifier. This type of modulation is called **low level modulation.**



Low Level Modulation System

- Pre amplifier is a class A linear voltage amplifier. It is used to amplify modulating signal to a usable level with minimum distortion
- Driver amplifiers are used to further amplify carrier and modulating signals to adequate level to drive the modulator.
- Modulated waves powers are quite low. To increase power level, amplification is required.
- Class A or Class B power amplifier stages are required to amplify carrier as well as modulating signals. It provides sufficient bandwidth to accommodate frequencies, otherwise sideband would cut off.
- Matching network matches output impedance of power amplifier to transmission line and antenna to radiate maximum signal power.

Advantages

Advantage of using a linear RF amplifier is that the smaller early stages can be modulated, which requires a small audio amplifier to drive modulator.

Disadvantages

Disadvantage is amplifier chain is less efficient, because it has to be linear to preserve modulation. Hence Class C amplifiers cannot be employed.

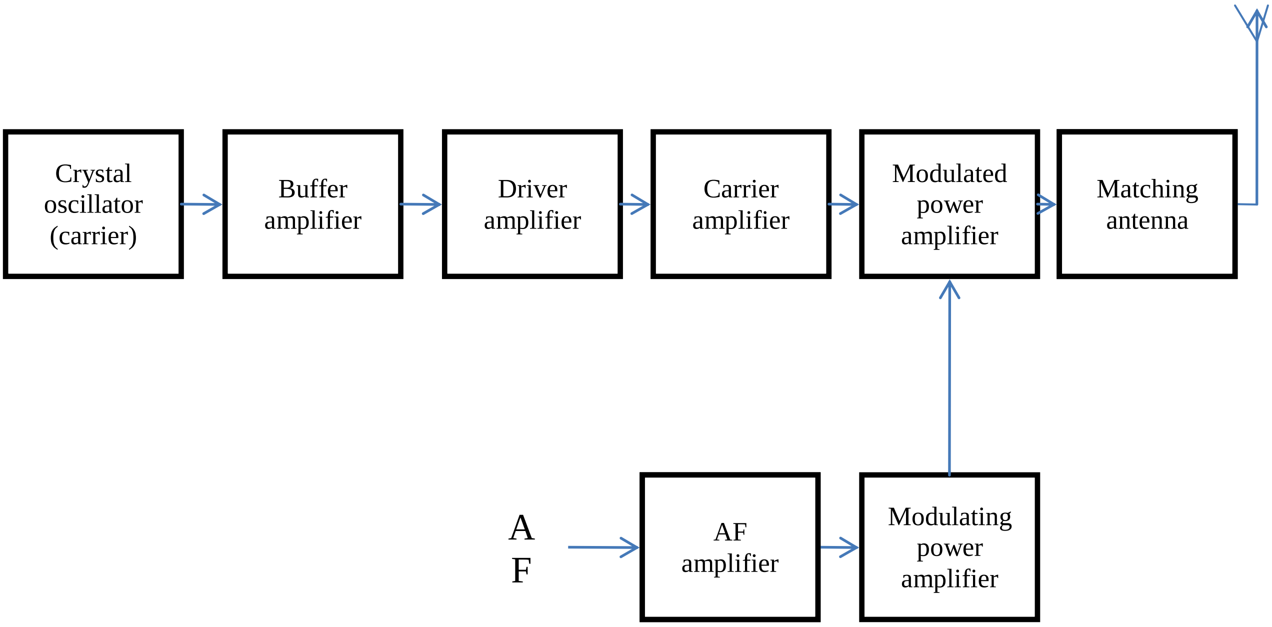
An approach which gets holdup the advantages of low-level modulation with the efficiency of a Class C power amplifier chain is to arrange a feedback system to compensate for the substantial distortion of the AM envelope. A simple detector at the transmitter output (which can be little more than a loosely coupled diode) recovers the audio signal, and this is used as negative feedback to the audio modulator stage. The overall chain then acts as a linear amplifier as far as the actual modulation is concerned, though the RF amplifier itself still retains Class C efficiency.

Applications

Remote control and walkie talkies wireless intercom and pagers AM radiotelephones.

HIGH LEVEL TRANSMITTER

If carrier and modulating signals are amplified to desired level before modulation (to produce 100%) such a system is said to be a high level modulation. With high level modulation, the modulation takes place at the final amplifier stage where the carrier signal is at its maximum



High Level Modulation System

- When a large modulated power is to be transmitted, low level modulation system is quite unsuitable because of low efficiency.
- The crystal oscillator generates carrier signal; driver amplifier amplifies carrier signal to desired level. Buffer amplifier isolates oscillator and driver amplifier to avoid loading effect.
- In this system both carrier and message signals are further amplified by power amplifiers before modulation takes place to raise the power to desired level
- Modulation which is usually carried out by using a collector modulated class ‘C’ amplifier, it provides three operations, such as modulator, final power amplifier and frequency converter. Frequency converter simply translates the low frequency message (AF) signal to R.F signal, which is fed into an antenna, it can radiate effectively into free space.
- In this transmitter power level of message signal and carrier signals are raised to the desired level before modulation thus modulated signal is directly connected to antenna through antenna matching network. The function of antenna matching network is providing proper impedance matching between antenna and modulated amplifier to radiate maximum power.
- Power conversion efficiency is higher .This system require less critical adjustments.

Advantages

One advantage of using class C amplifiers in a broadcast AM transmitter is that only the final stage needs to be modulated, and that all the earlier stages can be driven at a constant level. These class C stages will be able to generate the drive for the final stage for a smaller DC power input. However, in many designs in order to obtain better quality AM the penultimate RF stages will need to be subject to modulation as well as the final stage.

Advantages

Requires more power for modulation

A large audio amplifier will be needed for the modulation stage, at least equal to the power of the transmitter output itself. Traditionally the modulation is applied using an audio transformer, and this can be bulky. Direct coupling from the audio amplifier is also possible (known as a cascode arrangement), though this usually requires quite a high DC supply voltage (say 30 V or more), which is not suitable for mobile units.

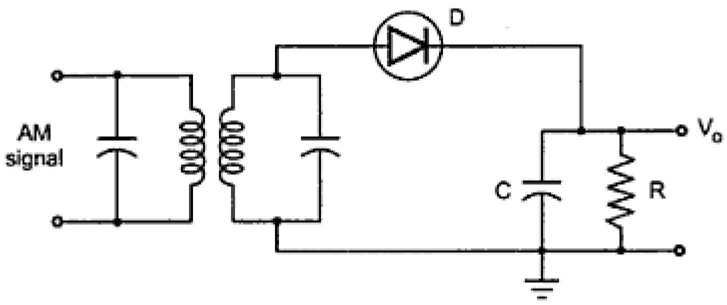
DIFFERNCE BETWEEN LOW LEVEL AND HIGH LEVEL MODULATION

LOW LEVEL MODULATION	HIGH LEVEL MODULATION
Modulated signals are amplified after modulation takes to raise the power level	Carrier signals and modulating signals are amplified before modulation takes to raise power level, so no amplification after modulation
Depth of modulation is less than 100%	Depth of modulation maximum100%
Low gain and efficiency	High gain and efficiency
Class B amplifier used to amplify modulated signal	Class C amplifier used. So no need to amplify after modulation
Base modulation is used.	Collector or emitter modulation is used.
Used for wireless intercom ,remote control, walkie talkie	Used for transmit radio and TV signals.

AM detector Circuits:

The detector is demodulator obtains the original modulation signal from the IF signal. Classified as coherent and Non coherent detectors.

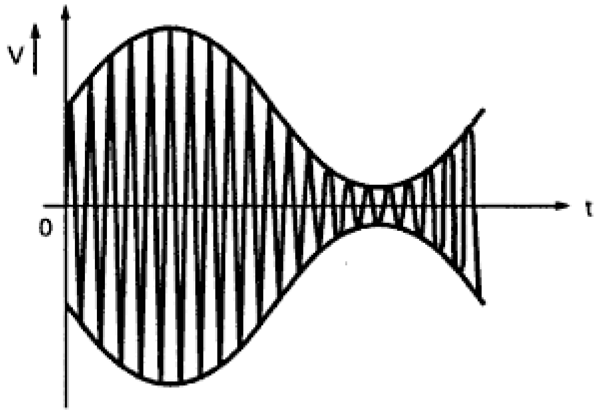
a. Diode detector or envelope detector



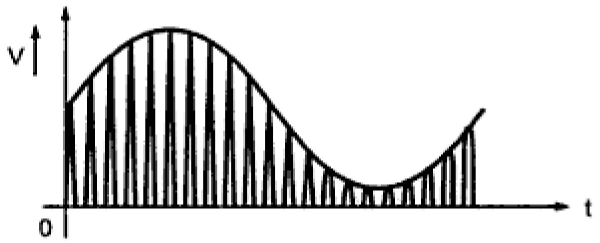
Envelope Detector

- The most commonly used AM detector is simple diode detector.
- The AM signal at fixed IF is applied to the transformer primary.
- The signal at secondary is half wave rectified by diode D. This diode is the detector diode.
- The resistance R is load resistance to rectifier and C is the filter capacitor.
- In the positive half cycle of AM signal, diode conducts and current flows through R, whereas in negative half cycle , the diode is reverse biased and no current flows.
- Therefore only positive half of the AM wave appears across resistance R as shown in fig.
- The capacitor across R provides low impedance at the carrier frequency and much higher impedance sat the modulation frequency.

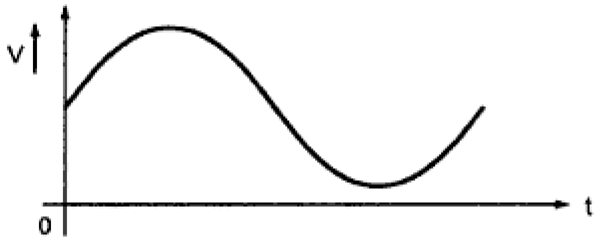
- Therefore capacitor reconstructs the original modulation signal as shown in fig and high frequency carrier is removed.



(a) AM Signal



(b) Current pulses through diode D



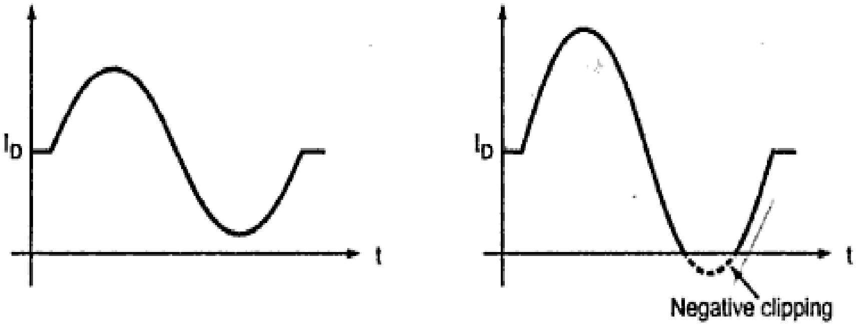
(c) Demodulating Signal

a) Negative peak clipping in diode detector

- This is the distortion that occurs in the output of diode detector because of unequal ac and dc load impedances of the diode. The modulation index is defined as Em/ Ec. Therefore it can also be defined as I_m/I_c with

$$I_m=E_m/Z_m \text{ and } I_c = E_c/R_c$$

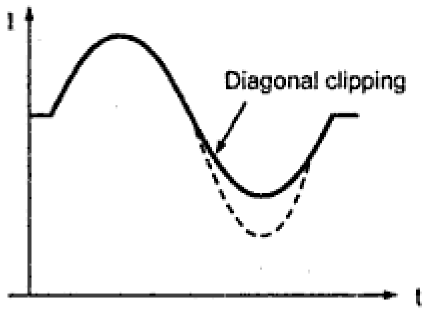
- Here Z_m is audio diode load impedance and Rc is the dc diode resistance.
- The audio load resistance of the diode is smaller than the dc resistance.
- Hence the AF current I_m is larger, in proportion to dc current.
- This makes the modulation index in the demodulated wave relatively higher than that of modulated wave applied at the detector input.
- This introduces the distortion due to over modulation in the detected signal for modulation index near 100% .
- This is illustrated in fig. In the figure observe that the negative peak of the detected signal takes place because of over modulation effect taking place in detector.



(a) Transmitted modulated wave (b) Received modulated wave

b. Diagonal clipping in diode detector:

.



Diagonal clipping

- As modulation frequency is increased, the diode ac load impedance, Z_m does not remain purely resistive.
- It does have reactive component also. At high modulation depths, the current changes so fast that the time constant of the load does not follow the changes.
- Hence the current decays slowly as shown in. the output voltage follows the discharge law of RC circuit. This introduces distortion in the detected signal and it is called diagonal peak clipping

<u>SYNCHRONOUS DETECTION:</u>

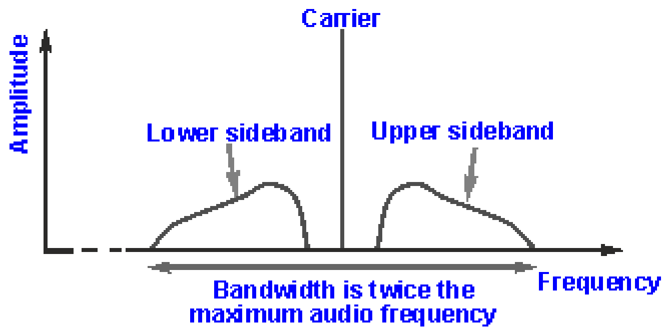
Synchronous forms of demodulation have inherent advantages over other forms of demodulation, although the additional levels of complexity mean that they are not always used.

Synchronous AM demodulation is generally reserved for higher performance radio receivers, although many integrated circuit technology means that it can incorporated into a chip with relative ease.

AM synchronous demodulation:

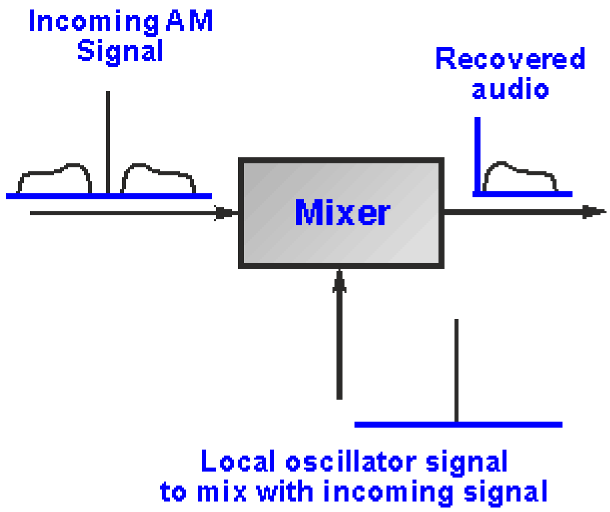
The simplest form of detection for an amplitude modulated signal utilizes a simple diode rectifier. To achieve improved performance a form of demodulation known as synchronous demodulation can be used.

- When looking at the synchronous demodulation of an AM signal, it is first useful to look at the spectrum of an amplitude modulated signal.
- It can be seen that it comprises a carrier with the two sidebands carrying he audio or other information spreading out either side. These two sidebands are reflections of each other.



Spectrum of an amplitude modulated, AM signal

The system uses an oscillator signal to mix with the incoming signal to convert it down to the baseband signal. If the local oscillator signal has exactly the same frequency as the carrier within the AM signal, this will appear as a DC component at the output - the DC level will depend on the phase between the carrier and the local oscillator. The sidebands of the AM signal will appear relative to zero frequency, i.e. as the original audio or other modulating signal.



Synchronous demodulation

Advantages and disadvantages of AM synchronous demodulation

There is a balance to be made between utilizing a simple diode detector and a synchronous detector. It is not always viable to incorporate an AM synchronous demodulator into a new design. Other formats may be more suitable.

ADVANTAGES & DISADVANTAGES OF AM SYNCHRONOUS DEMODULATORS	
ADVANTAGES	DISADVANTAGES
1. Increased linearity - lower levels of distortion. 2. Considerably less affected by selective fading experienced on the medium and short wave bands. 3. Improved sensitivity. 4. Improved signal to noise ratio.	1. Considerable additional complexity, although this is not such an important consideration if the synchronous detector can be included in an IC.

Methods of achieving AM synchronous demodulation

There are several ways in which synchronous AM detection can be achieved. They all have their own properties, but all follow the basic principle whereby a signal is injected into a mixer at the same frequency as the carrier to re-constitute the audio or other data.

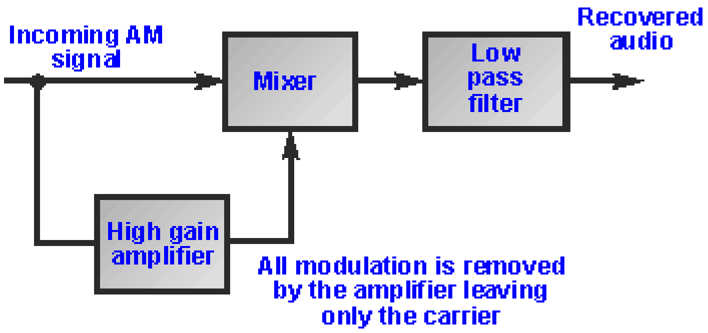
- **Filter method:** The most obvious method to create a local oscillator signal on the same frequency as the carrier is to use a fixed frequency filter that is tuned to remove only the carrier frequency. This can be phase shifted, 90°: and entered into the mixer. The phase shift will ensure the DC component at the output of the mixer is minimised.

  The drawback for this method is that the carrier has to be positioned exactly on the frequency of the filter for it to work. This means that the tuning has to be exact.

- **Phase locked loop:** Phase locked loops provide a convenient method of extracting the carrier. The phase locked loop will lock on to the carrier of the AM signal and the VCO output can be fed into the mixer with a 90°: phase shift as before.

- **Hard limiting amplifier:** Possible the cheapest and easiest option to implement is a hard limiting amplifier approach.

  This circuit operates using the principle that if the signal is hard limited, the output of the amplifier will not have any amplitude variations on it - hence it will only allow through the carrier without modulation. This is exactly what is required for the mixing process.



High gain limiting amplifier synchronous detector

AM synchronous detection performance

- A synchronous detector is more expensive to make than an ordinary diode detector when discrete components are used, although with integrated circuits being found in many receivers today there is little or no noticeable cost associated with its use as the circuitry is often included as part of an overall receiver IC.

- Synchronous detectors are used because they have several advantages over ordinary diode detectors. Firstly the level of distortion is less. This can be an advantage if a better level of quality is required but for many communications receivers this might not be a problem. Instead the main advantages lie in their ability to improve reception under adverse conditions, especially when selective fading occurs or when signal levels are low.

- Under conditions when the carrier level is reduced by selective fading, the receiver is able to re-insert its own signal on the carrier frequency ensuring that the effects of selective fading are removed. As a result the effects of selective fading can be removed to greatly enhance reception.

- The other advantage is an improved signal to noise ratio at low signal levels. As the demodulator is what is termed a coherent modulator it only sees the components of noise that are in phase with the local oscillator. Consequently the noise level is reduced and the signal to noise ratio is improved.
- Unfortunately synchronous detectors are only used in a limited number of receivers because of their increased complexity. Where they are used a noticeable improvement in receiver performance is seen and when choosing a receiver that will be used for short wave broadcast reception it is worth considering whether a synchronous detector is one of the facilities that is required.

RECEIVER

A **receiver** is an electronic circuit that receives its input from an antenna, uses electronic filters to separate a wanted radio signal from all other signals picked up by this antenna, amplifies it to a level suitable for further processing, and finally converts through demodulation and decoding the signal into a form usable for the consumer, such as sound, pictures, digital data, measurement values, navigational positions, etc

In consumer electronics, the terms radio and radio receiver are often used specifically for receivers designed for the sound signals transmitted by radio broadcasting services.

Types of radio receivers

Various types of radio receivers may include:

- Consumer audio and high fidelity audio receivers and AV receivers used by home stereo listeners and audio and home theatre system enthusiasts.
- Communications receivers, used as a component of a radio communication link, characterized by high stability and reliability of performance.
- Simple crystal radio receivers (also known as a crystal set) which operate using the power received from radio waves.
- Satellite television receivers, used to receive television programming from communication satellites in geosynchronous orbit.
- Specialized-use receivers such as telemetry receivers that allow the remote measurement and reporting of information.
- Measuring receivers (also: measurement receivers) are calibrated laboratory-grade devices that are used to measure the signal strength of broadcasting stations, the electromagnetic interference radiation emitted by electrical products, as well as to calibrate RF attenuators and signal generators.
- Scanners are specialized receivers that can automatically scan two or more discrete frequencies, stopping when they find a signal on one of them and then continuing to scan other frequencies when the initial transmission ceases. They are mainly used for monitoring VHF and UHF radio systems.
- Internet radio device

1. AM receiver

An AM receiver detects amplitude variations in the radio waves at a particular frequency and then amplifies changes in the signal voltage to drive a loudspeaker or earphones.

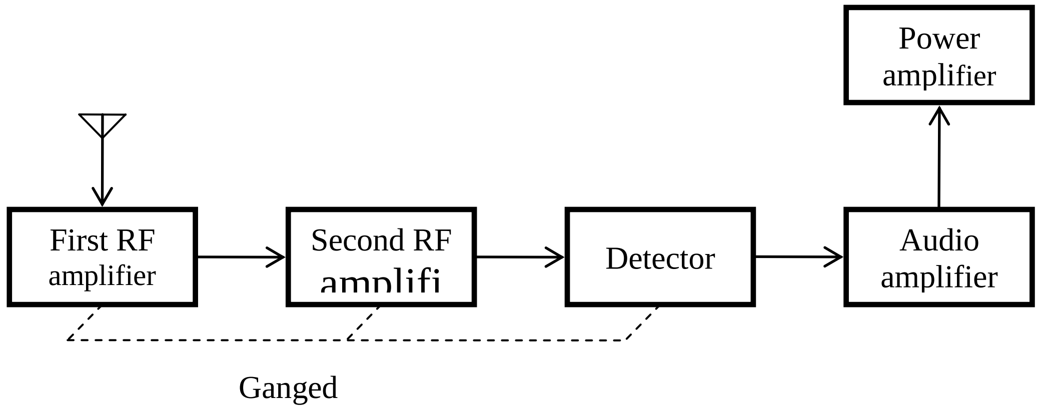
Basically there are two types of AM receiver

1. Simple receiver or Tuned radio frequency (TRF) receiver
2. Super heterodyne receiver

a) Simple receiver or Tuned radio frequency (TRF) receiver

- The TRF receiver was patented in 1916 by Ernst Alexanderson.

- The concept was that each stage would amplify the desired signal while reducing the interfering ones. The final stage was often simply a *grid-leak detector.
- A tuned radio frequency receiver (TRF receiver) is a radio receiver that is usually composed of several tuned radio frequency amplifiers followed by circuits to detect and amplify the audio signal.



9.Tuned Radio Frequency (TRF) Receiver

- A 3 stage TRF receiver includes a RF stage, a detector stage and an audio stage.
- Generally, 2 or 3 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 signal to a usable level. It can be difficult to operate because each stage must be individually tuned to the station's frequency.

Advantages

1. The significance of the term "tuned radio frequency" is very simple to implement when compared to superheterodyne receiver.
2. A tuned radio frequency receiver actually tunes the receiver on the true radio frequency whereas superheterodyne receiver, tunes the desired signal after conversion to an intermediate frequency.
3. It has high sensitivity
4. It is useful only to single channel low frequency oscillations.

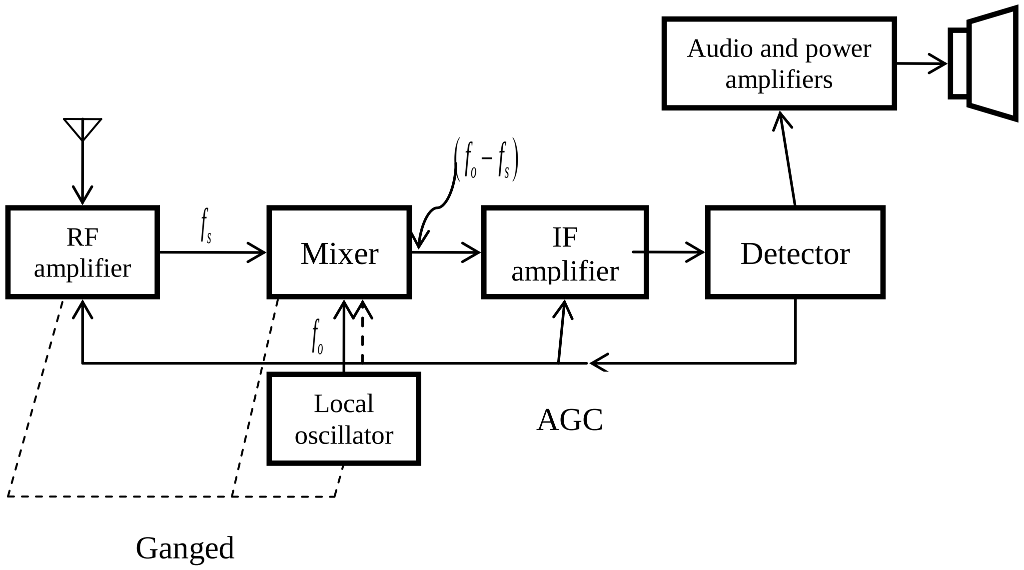
Disadvantages

1. 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 skin effect.
2. Second disadvantage is its instability due to large number of RF amplifiers all tuned to 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.
3. The third disadvantage is their gains are not uniform over a very wide frequency range because of the non-uniform L/C ratios of the transformer-coupled tank circuits in the RF amplifiers.

The problem of instability, insufficient adjacent frequency rejection and bandwidth variation can be overcome by superheterodyne receiver

b) Superheterodyne receiver

Superheterodyne receiver uses frequency mixing or heterodyning to convert a received signal to a fixed intermediate frequency, which can be more conveniently processed than original radio carrier frequency. TRF was replaced by the superheterodyne receiver invented by Edwin Armstrong. Virtually all modern radio and television receivers use the superheterodyne principle. A "heterodyne" refers to a beat or "difference" frequency produced when two or more radio frequency carrier waves are fed to a detector.



Superheterodyne receiver

RF amplifier
The antenna receives all the signals and gives it to wideband RF amplifier. It amplifies the signal in required range of frequencies hence it provides initial gain and selectivity. RF circuits are used to select wanted frequency and to reject interference such image frequency and to reduce receiver noise figure.

Mixer and local oscillator
The output of RF amplifier is given to mixer. Local oscillator output is applied to mixer. Let us assume that local oscillator frequency is f_o and signal frequency is f_s. The signal frequency f_s and local oscillator frequency f_o are mixed in the mixer in such a way that frequency difference $(f_o - f_s)$ is produced at the output of mixer. The difference $(f_o - f_s)$ is called intermediate frequency. The signal at this IF contains same modulation as incoming signal. During this stage, a class C non linear device processes the signals producing the sum, difference and originals.

IF amplifier
The output of mixer is given to IF amplifier. It has a very narrow bandwidth. Class A device capable of selecting frequencies between 0.455kHz ± 3kHz and rejecting all other frequencies. It consists of 2 or 3 transformers, each consisting pair of mutually coupled tuned circuits. With these circuits and specially chosen frequency it provides maximum gain and bandwidth required. Since its characteristics are independent, selectivity and sensitivity are uniform throughout is tuning range.

Automatic gain control
A part of output is taken from detector and it is applied to RF amplifier, Mixer and IF amplifier for controlling gains. This is known as Automatic gain control or AGC. It maintains constant output voltage level over a wide range of RF input signal levels.

Detector
It converts AM signals at IF to original modulating signal. Diode detector is commonly used for AM signals. The output of detector is low power audio or modulating signals.

Audio or power amplifier
The signal received from detector is weak and needs amplification. Normally audio amplifiers are used it possess one or more stages. The signal is amplified and given to speakers.

Envelope tracking is a technology associated with Radio Frequency, RF amplifier design.

When using envelope tracking technology the power supply voltage applied to the power amplifier is constantly adjusted to ensure that the amplifier is operating at peak efficiency for the given instantaneous output power requirements.

This brings many advantages to RF amplifier applications in many industry sectors and as a result it has been adopted in a wide variety of areas.

While the concept behind envelope tracking technology has been understand for many years, it has not been easy to achieve in a workable form until recently because of the difficulties of implementing it in a satisfactory format.

Advantages

1. Superheterodyne receivers have superior characteristics to simpler receiver types in frequency stability and selectivity. They offer better stability than Tuned radio frequency receivers (TRF).
2.
3. IF filters can give narrower pass bands at the same Q factor than an equivalent RF filter.
4. A fixed IF also allows the use of a crystal filter when exceptionally high selectivity is necessary.
5. Regenerative and super-regenerative receivers offer better sensitivity than a TRF receiver, but suffer from stability and selectivity problems.
6. A constant frequency difference is maintained between local oscillator and the RF circuits normally through capacitance tuning in which all capacitors are ganged and operated in unison.

Disadvantages

1. High-side and low-side injection
2. Image Frequency (f_{image})
3. Local oscillator radiation
4. Local oscillator sideband noise

RF amplifier efficiency improvement techniques

Over the years there have been several techniques, apart from envelope tracking that have been used to improve the efficiency performance of RF amplifiers. Each of these techniques has its advantages and disadvantages.

TECHNOLOGY	BASIS OF OPERATION	ADVANTAGES	DISADVANTAGES
Operate PA in compression	This scheme is only applicable for constant amplitude schemes such as FM and GMSK where there is no amplitude element in the modulation. Can use digital pre-distortion to provide linearity where some amplitude components are available.	• Easy operation where modulation and performance requirements permit. • Provides high efficiency while in compression	• Does not provide high efficiency for high peak to average power levels signals as amplifier will not run in compression all the time.
DC Tracking	The PA voltage is adjusted in line with the modulation level, often using protocol commanded power supply voltage level changes	• Relatively straightforward to implement. • Does not require particularly advanced techniques.	• Deals only with average power. • Does not address broadband issues. • Does not address instantaneous peaks and troughs.
Doherty	Amplifier consists of two	• Gives useful	• Requires combiner

TECHNOLOGY	BASIS OF OPERATION	ADVANTAGES	DISADVANTAGES
Amplifier	elements: main amplifier and a peaking amplifier that comes in to assist with peak power levels. The peaking amplifier is biased off for most of the time, only coming into operation for peaks.	increase in efficiency	for combining output of two amplifier sections. • Gain kinks can occur • Gain and phase vary with output power.
Envelope Tracking	Envelope tracking, has the PA supply voltage that tracks the RF envelope. This enables the amplifier to be run at a voltage that gives the optimum efficiency or other performance level at any instantaneous power level.	• Provides best performance at all power levels. • Permits broadband operation. • Provides additional advantages in terms of operation into mismatched loads, etc. .	• Envelope tracking requires very fast - high bandwidth - power supply. • Requires accurate envelope signal for power supply.

Each technique has its own advantages and can be used to provide significant benefits under certain scenarios. Envelope tracking is able to offer some significant benefits in terms of efficiency, and general operation for RF power amplifiers used in many applications.

Envelope tracking background

Before looking further into the technology in this envelope tracking tutorial, it is worth looking at the background and need for the technique.

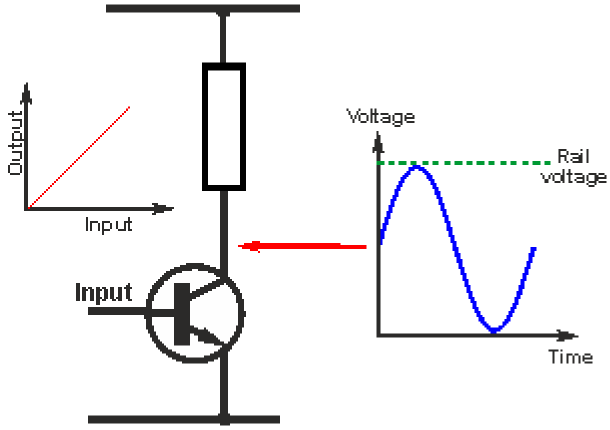
In most RF amplifier applications, the efficiency of the amplifier has an impact on the design, operation and efficiency of the overall system. Power supply requirements, RF amplifier capabilities and heat-sinks, battery life and many more elements are dependent upon the RF amplifier, especially when it is one of the main power users within a system.

For any RF amplifier power is supplied to the circuit, and a signal is produced. The output will always be less than the DC input power, the ratio of output to DC input being the efficiency.

The efficiency of an amplifier depends upon the shape of the waveform and the mode in which it is operating.

When operating in a linear mode, the output device must always be in conduction, with the output voltage rising and falling between the two limits.

When operating in this mode, often called Class A, the maximum theoretical efficiency that can be achieved is 50%. However in a real system the achieved levels are always below this.



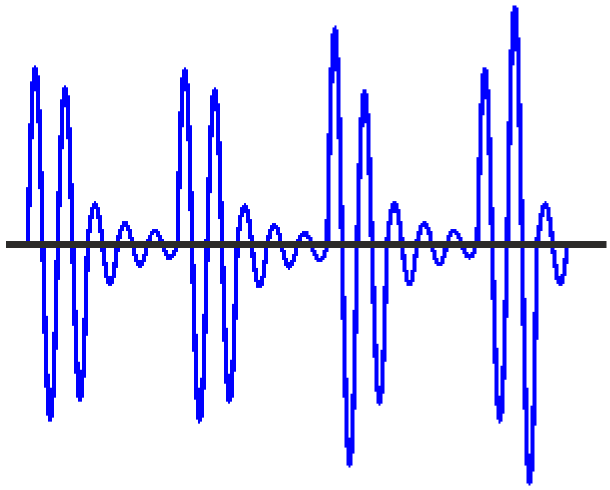
Linear operation of an amplifier

To achieve better efficiency levels, it is possible to drive the amplifier into compression. Much greater levels of efficiency can be achieved, and if a steady waveform, like FM is used, the only degradation of the signal is that additional harmonics of the fundamental carrier are generated and these can be filtered out using RF filters.

Unfortunately when modulation with an amplitude component is applied to a carrier this is distorted if it is passed through an amplifier that is run in compression.

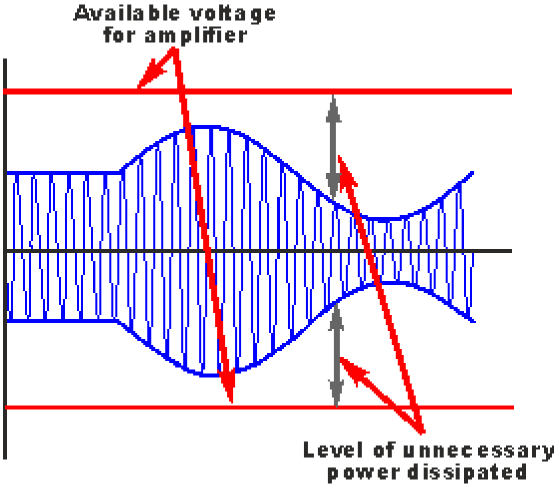
For data transmission systems that are used today like UMTS, HSPA, and 4G LTE, etc, the RF waveforms that are used incorporate an amplitude component in addition to the phase elements and therefore they require a linear amplifier.

The situation becomes worse as if the peak to average ratio is high, i.e. the waveform has higher peak levels when compared to the average because the amplifier has to be able to accommodate the peaks while still only running at a low average power level.



Waveform with high peak to average ratio

During the peaks, the amplifier requires the full voltage to be able to deliver the required power without running into compression, but during the periods of lower signal, this voltage is not required and means that power is dissipated in the device. The amplifier only requires a smaller voltage to deliver the lower levels of power and therefore running with the higher voltage all the time, unnecessarily wastes power.



High levels of power are dissipated if the full rail voltage is maintained

It can be seen that the power dissipated is proportional to the aea between the top of the RF envelope and the rail voltage. For low peak to average power ratio signals, this can be high.

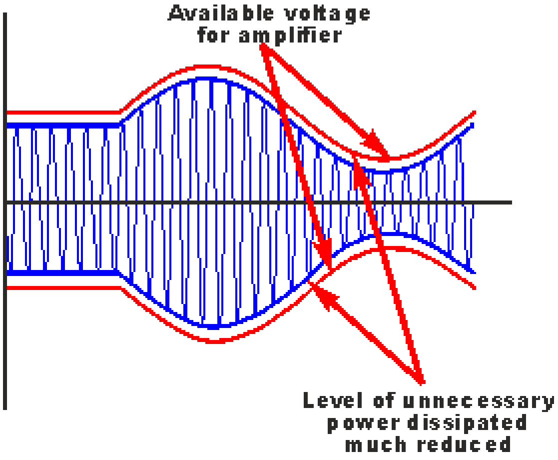
In addition to the modulation, many requirements placed on RF amplifiers used in modern applications such as cellular telecommunications still further reduce the efficiency levels - many amplifiers are required to operate over wide bands, or multiple frequency bands for example.

Envelope tracking: basic concept

In order to improve the efficiency levels of RF amplifiers, one of the approaches that can be used is to employ envelope tracking technology.

As the name indicates, envelope tracking employs a system whereby the amplitude envelope of the signal is tracked and utilised by the amplifier.

Using envelope tracking the power supply voltage applied to the power amplifier is constantly adjusted to ensure that the amplifier is operating at peak efficiency for the given instantaneous output power requirements.



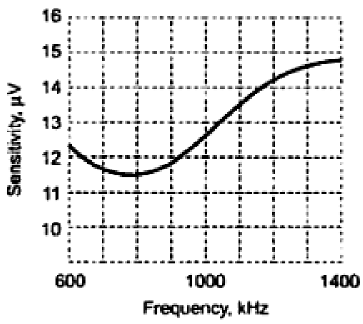
Envelope tracking concept

RECEIVER CHARACTERISTICS

Sensitivity, noise, selectivity, Image frequency rejection ratio and fidelity are important receiver characteristics.

Sensitivity

- The ability of a receiver to reproduce weak signals and amplify them is a function of the sensitivity. It is often defined in terms of voltage that must be applied to the receiver input terminals to give the standard output power, measured at the output terminals.
- As the gain of receiver is increased, sensitivity is also increased
- It is expressed usually in the microvolts or deciBel that must be applied to the antenna input terminals to give an established level of the output.
- The output may be an ac or dc voltage measured at the detector output or a power measurement (measured in decibels or watts) at the loudspeaker or headphone terminals.

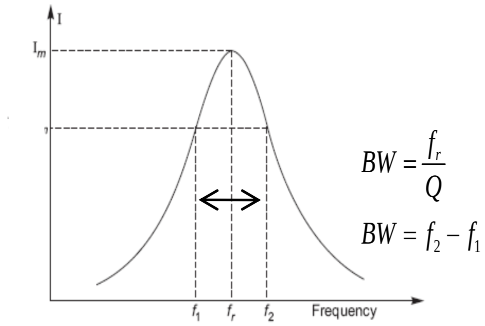


11. a Sensitivity curve of a typical receiver

Noise

- All receivers generate a certain amount of noise, which must be taken into account when measuring sensitivity.
- Receiver noise may originate from the atmosphere (lightning) or from internal components (transistors, tubes).
- Noise is the limiting factor of sensitivity.
- Sensitivity is the value of input carrier voltage (in microvolts) that must be applied from the signal generator to the receiver input to develop a specified output power.

Selectivity

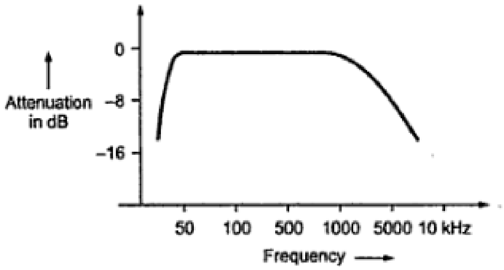


11. b Selectivity curve

- Selectivity is the ability of a receiver to decide on a signal of desired frequency and rejecting all others.
- Selectivity is the degree of distinction made by the receiver between the desired signal and unwanted signals. Better the ability of receiver to reject unwanted signals, better its selectivity.
- The selectivity of receiver is done partially by RF amplifier and mainly by IF amplifier.
- It shows the attenuation that receiver offers to signal at frequencies near which it is tuned.
- The selectivity depends on tuned LC circuits which is in RF and IF amplifiers where resonating frequency (f_r) and (Q) quality factor of circuits.
- Bandwidth should be narrow for better selectivity. Hence Q should be high.

- The degree of selection is determined by the sharpness of resonance to which the frequency-determining circuits have been tuned.

Fidelity

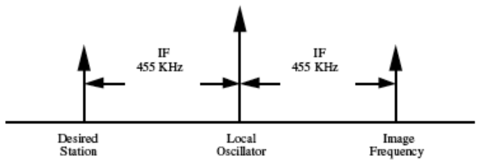


11. c Typical fidelity curve

- Fidelity of a receiver is its ability to accurately reproduce its output, signal that appears at its input.
- Ability of receiver to reproduce all range of modulating frequencies equally is called fidelity.
- Generally a fidelity curve must have a flat response over a wide range of frequencies.
- Fidelity requires wide band of frequencies to be amplified hence it requires more bandwidth of RF and IF stages required but results in poor selectivity.
- AM receivers are not good fidelity receivers, since band width is low
- Good fidelity requires broader band to amplify the outermost frequencies of the sidebands.

Image frequency rejection ratio

- It is the ratio of intermediate-frequency (f_i) signal level produced by desired input frequency to that is produced by the image frequency. The image rejection ratio is usually expressed in dB.
- When the image rejection ratio is measured, the input signal levels of the desired and image frequencies must be equal for the measurement.



$$f_{image} = f_s + 2f_{IF} \text{ if } f_o > f_s$$
$$f_{image} = f_s - 2f_{IF} \text{ if } f_s > f_o$$

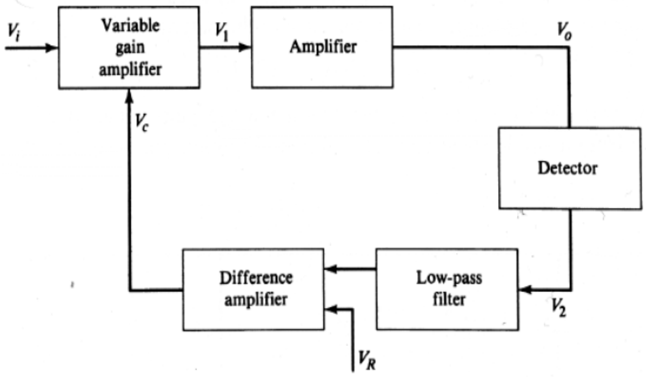
11. d Image frequency rejection

- The local oscillator frequency is made higher than signal frequency such that $f_o - f_s = f_i$. Here f_i is intermediate-frequency (IF) that is $f_o = f_s + f_i$.
- The IF stage passes only f_i.If the frequency $f_{si} = f_s + 2f_i$ appears at the input of mixer, then the mixer will produce difference equal to f_i. This is equal to IF.
- The frequency f_{si} is called image frequency and is defined as signal frequency plus twice the IF. This image frequency is converted in the IF range and it is amplified by IF amplifiers. This is the effect of two stations being received simultaneously.
- The image frequency rejection is done by tuned circuits in RF stage. It depends on the selectivity of the RF stage .The image rejection should be done before IF stages.
- It is the measure of ability of preselector to reject the image frequency. It is given by $IFRR = \sqrt{1+Q^2\rho^2}$. Here Q is the quality factor of preselector and $\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}}$

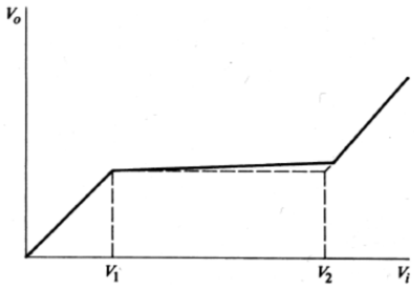
Automatic Gain control:

- The automatic gain AGC keeps the output signal level constant irrespective of the increase or decrease in the signal level at the input of receiver.

- The AGC circuit takes part of detected signal and derives a dc control voltage for RF, mixer and IF stages.
- This control voltage acts as negative feedback and controls the overall gain of these stages.
- The gain is varied such that output signal level is constant.
- There are two types of AGC: *simple AGC and delayed AGC.*
  - In the ***simple AGC***, the gain control is active at low output levels also, hence low AGC voltage decrease the overall gain and reduce the sensitivity of the receiver. Fig. Shows AGC characteristics. Observe the difference in characteristics due to no AGC and simple AGC. At low input signals. The output of receiver is low compared to that would have been with no AGC



AGC Block Diagram



AGC Ideal Transfer Function

Delayed AGC:

The disadvantage of automatic gain control, attenuating even the weak signal, is overcome by the use of delayed automatic gain control (DAGC). In the delayed AGC, the AGC bias is not applied till the output signal reaches the predetermined level. When the output signal tries to exceed this level, the AGC bias is applied and gain is reduced sot that output voltage remains at predetermined level. Hence this type of AGC is called *delayed AGC*. In fig observe that the characteristic of delayed AGC is very close to that of ideal AGC. Almost all of the receivers use delayed AGC.

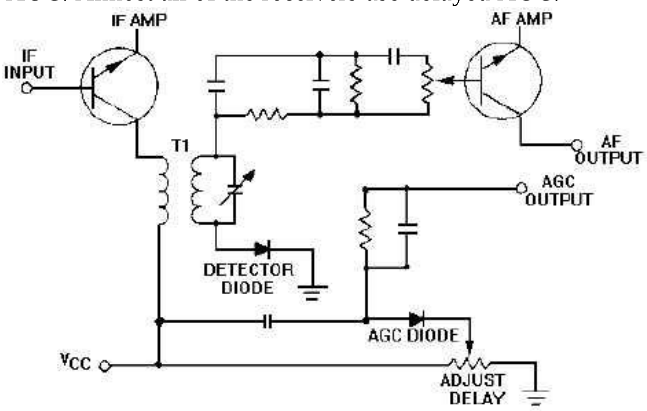


Fig. Delayed AGC Action

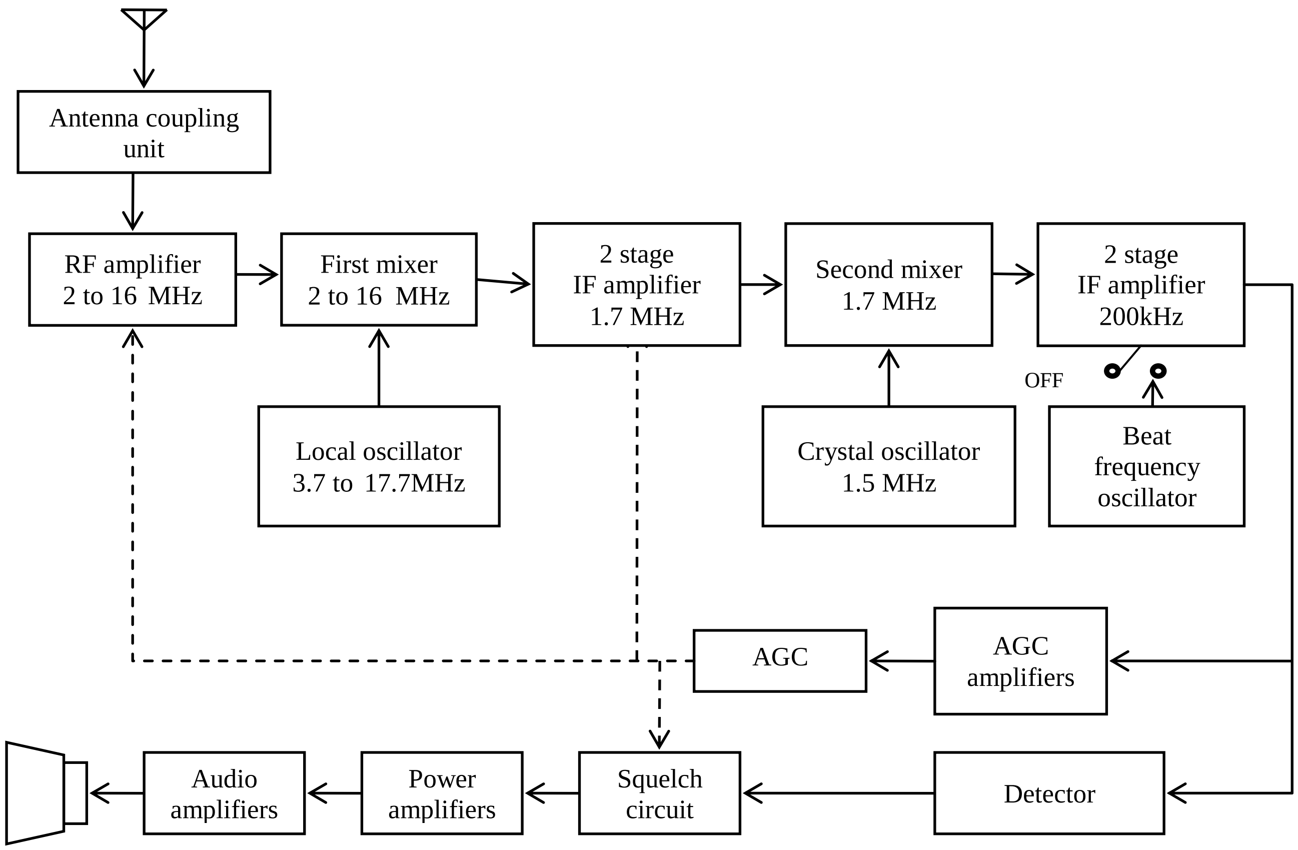
COMMUNICATION RECEIVER

Communication receivers are high quality, short wave, multipurpose, superheterodyne used to receive signals for communication. It can receive both high and low frequency signals, hence it has many applications. These receivers include special feature for noise free uninterrupted reception of signals and also use circuits to ensure accurate tuning. Some of the applications are:

1. Detection of signal strength from high frequency impendance bridges.
2. Measurement of signal strength and its frequency detection and display of different components of a high frequency FM wave.

It is an extension of superheterodyne receiver and has several addition features such as:

1. Band spread tuning
2. Double frequency conversion
3. Delayed AGC
4. Tuning calibration
5. Beat frequency oscillator(BFO)
6. Noise limiter
7. AFC
8. Tuning Indicator
9. Squelch circuit.



Communication Receiver

1. RF AMPLIFIER (or) TUNER:

The signal received by the antenna is applied to RF amplifier stage which can be tuned to any signal frequency over a frequency range of (2 MHz to 30 MHz). The amplifier signal is fed to the input of first mixer.

2. FREQUENCY CHANGER (or) CONVERTER:

The output of first mixer is the difference between incoming signal frequency and output of it is applied to the first IF amplifier. The IF is kept higher than usual IF used in broadcast receiver in order to obtain a best image rejection. The value of first IF employed in the circuit is 650 KHz.

After the amplification of 1st IF amplifier the signal is applied to second mixer which is designed to produce the difference frequency of 150 KHz. This is the second IF and is given to second If amplifier. Because of low frequency of this IF amplifier, a good adjacent channels selectivity is obtained between the control the band width of the receiver.

DETECTOR:

The output of second IF stage is fed to detector as well as to AGC. The AF signal output from the detector is passed onto AF amplifier where it is amplified and may be connected to a loud speaker, head phone or a telephone line. The DC components produced in the detector output are connected to squelch circuit and also to the tuning indicator. As long as the signal at the receiver input is present, DC component in the detector output keeps off the squelch circuit but when there is no incoming signal, the squelch circuit turns on in the absence DC bias from the detector and brings the audio amplifier to cut-off.

BAND SPREAD TUNING:

Generally, the frequency interval between adjacent stations in relation to the carrier frequency becomes small. The process of tuning of these stations then becomes tedious. Very slight movement of tuning control (capacitor) makes variations in the tuned circuit frequency. It is therefore advantageous to spread out the tuning range, so that considerably large movement of tuning results in only a small frequency change.

Band spread is generally obtained by anyone of the following two ways.

1. Using auxilary variable capacitors ganged together which forms a fine tune control. In this case the receiver is first tuned approxinmately by main tuning control and then the receiver is tuned accutrately with the help of fine tuning control.
2. By allocating full tuning band to a amall range of frequencies, that is instead of using less number of bands, a receiver may have more number of bands.

CRYSTAL FILTER:

Crystal filters are commonly employed to restrict the bandwidth of communication receivers so that interference signal may be wiped out from reception. They are used for obtaining different IF band widths in a typical communication receiver.

DOUBLE FREQUENCY CONVERSION:

The important feature of communication receiver is double frequency conversion i.e., there is more than one intermediate frequency. The first IF is high, generally several MHz and the second one is of few hundred KHz.

After the recovery of signal by RF amplifier, the signal is mixed with the output of a local oscillator to produce difference frequency of 650 KHz. The high intermediate frequency is then amplified by the high frequency IF amplifier and the output is fed to a second mixer and mixed with the output of a second local oscillator. The low second intermediate frequency is amplified by low frequency IF amplifier and is then detected.

Some of the main reasons for using double frequency conversion are:

1. The high first intermediate frequency separates the image frequency further away from the original frequency and therefore provides much better attenuation of it.
2. The low intermediate frequency provides sharp selectivity and hence provides good adjacent channel rejection.

The high intermediate frequency must come first to sufficiently reject the image frequency. Thus double frequency conversion provides a combination of higher image and adjacent frequency rejection.

DELAYED AGC:

The function of AGC is to reduce the gain of the receiver for strong signals. However even the weak signals get reduced, this could be avoided by use of delayed AGC. In this, the AGC bias is not applied until the signal strength has reached a predetermined level, after which it is applied , with normal AGC .As the signal strength rises , the receiver output also , raises , but relatively slightly , then the problem of reducing the gain of the receivers for weak signal has been avoided.

TUNNING CALIBRATION:

The facility uses a built in typical oscillator usually operating at 500-1000 KHz where output may be given to the input of the receiver by a operating switch. With the beat frequency oscillator in operation, whistles will now be heard at 500 or 1000 kHz intervals, especially crystal oscillator works into a resistive load, so as not attenuate harmonics of the fundamental frequency .the calibration of the receiver may now be corrected by adjustment of the pointer which must be moved independently of the gang.

BEAT-FREQUENCY OSCILLATOR (BFO):

1. In a normal receiver there is no provision for registering the presence and absence of a carrier, such as dots, dashes and spaces would produce no output whatever from the detector.
2. A communication receiver should be capable of receiving the phase modulated RF carrier i.e, morse code. To make the morse code audible, the receiver uses a beat frequency oscillator, which is a simple LC oscillator operating at frequencies 1Khz or 400Hz above or below the intermediate frequency.
3. Since signal is presented during a dot or dash in morse code, only the whistles are heard in the loud speaker.
4. To prevent interference, the BFO is switched OFF when normal operation is resumed.

NOISE LIMITER

Noise limiter in a communication receiver is really an impulse noise limiter whose function is to eliminate or atleast reduce the interfering noise pulses created by ignition systems and other electrical machinery of various types .In a common noise limiter a Diode is used in conjunction with a differentiating circuit. This circuit provides a negative voltage as a result of the noise impulse and this voltage applied to the detector which is thus cut-off.the detector thus remain cut-off for the duration of the noise pulses.

AUTOMATIC FREQUENCY CONTROL (AGC)

Its very difficult to tune accurately the design of a super heterodyne receiver to a station by inspecting its output. Since receiver has automatic gain control, it tends to maintain the receiver output constant in spite of variations in signal strength. Similarly a drift in a local oscillator frequency may cause mistuning of receiver but even under circumstances, automatic voltage will tend to help receiver output constant .These difficulties may be overcome by incorporating an AFC circuit into the receiver.

The circuit comprise of a discriminator which is tuned to the center frequency of the IF. When the receiver is correctly tuned the discriminator output is zero, but an incorrect tuning of the receiver results in corresponding DC bias being produced by the discriminator which is passed onto the reactance tube modulator for shifting the local oscillator frequency accordingly so as to produce the correct intermediate frequency. the presence of amplitude modulation does not affect the AFC circuit since the discriminator output is filtered by a low pass filter before being applied to the reactance tube.

TUNING INDICATOR:

This is an essential part of communication receiver and is employed in broadcast receiver in order to tune them correctly to any desired station. with automatic voltage control system in these receivers , it is virtually impossible to obtain correct tuning without a tuning indicator because of AVC reduces the variation in sound output, as tuning is done and the receiver may be so much off tuned that one of the sidebands be excessively cutoff causing a distorted output.

A simple DC voltmeter may be employed in a tuning indicator by connecting it to measure AVC bias. Maximum reading in this voltmeter will indicate accurate tuning this is used in communication receiver.

SQUELCH CIRCUIT:

While tuning from one station to another, or when the transmitting signal is absent, AVC increases the noise present in the receiver to disagreeable proportions. The circuit for suppressing this inter channel noises is called noise suppressor or squelch, or quieting system or muting system.

UNIT – II
FREQUENCY MODULATION AND DEMODULATION

Principle of frequency and phase modulation-Relation between FM and PM waves-Bandwidth of FM-Narrow band and wideband FM-Generation of FM wave-Direct and Indirect methods-FM transmitters-Block diagram-Function of each block. Detection –FM detectors-Slope detectors-Phase discriminators-Ratio detectors. Receivers- different types FM receivers - Block diagram.FM – Receiver characteristics.

PRINCIPLE OF FREQUENCY AND PHASE MODULATION

Angle Modulation

In angle modulation, the amplitude of the modulated carrier is held constant and either phase or time derivative of the phase of carrier is varied linearly with message signal m(t). Thus general angle-modulated signal is given by

$$x(t) = V_c \cos[2\pi f_c t + \theta(t)]$$

The quantity $(2\pi f_c t + \theta(t)) = \psi_i(t)$ is called the instantaneous phase of x(t), while the quantity θ(t) is called the phase deviation of x(t).

The instantaneous angular frequency of x (t), defined as rate of change of instantaneous phase and having units of radians per sec, is

$$\omega_i(t) = \frac{d\phi_i}{dt} = \frac{d(2\pi f_c t+\theta(t))}{dt} = 2\pi f_c + \frac{d\theta(t)}{dt}$$

The instantaneous frequency $f_i(t)$, having units of hertz (Hz), of x(t) is accordingly given by

$$f_i(t) = \frac{\omega_i(t)}{2\pi} = f_c + \frac{1}{2\pi}\frac{d\theta(t)}{dt}$$

The quantity dθ/dt is called the **angular frequency deviation**.

The two basic types of angle modulation are phase modulation (PM) and frequency modulation (FM).

Phase Modulation

Variation of θ(t) produces Phase Modulation

Phase modulation implies that θ (t) is proportional to the modulating signal. Thus

$$\theta(t)=2\pi k_p m(t)$$

where k_p is the deviation constant having units of volt^{-1}.

Frequency Modulation

Variation of dθ/dt produces Frequency Modulation

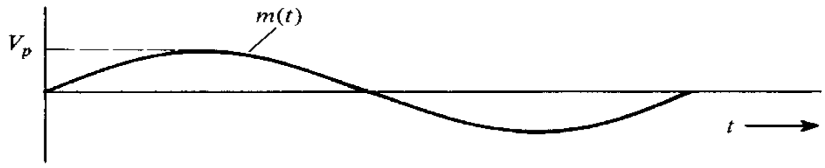
Frequency modulation implies that $\frac{d\theta}{dt}$ is proportional to the modulating signal.This yield

$$\frac{d\theta}{dt}=2\pi k_f m(t)$$

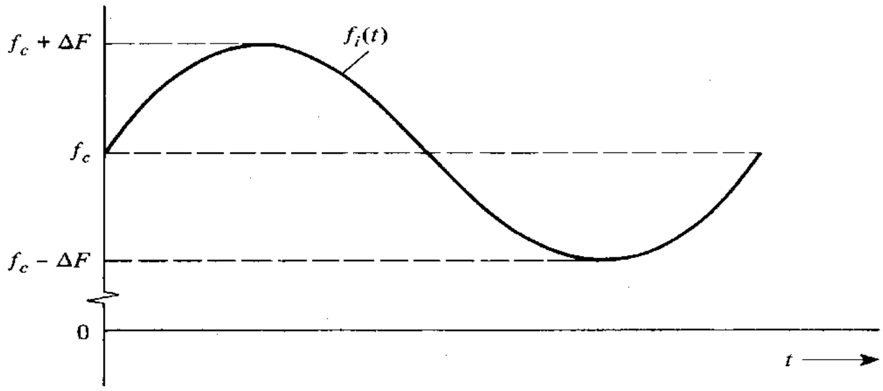
Thus, in FM the instantaneous frequency varies linearly with the message signal.

$$f_i=f_c+k_f m(t)$$

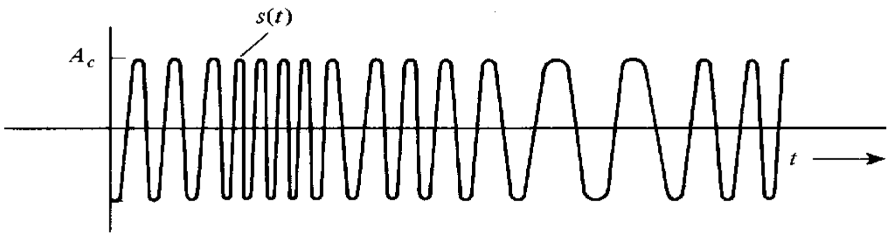
The term k_f, expressed in Hertz per unit of m(t), represents the **frequency sensitivity** of the FM signal.



(a) Sinusoidal Modulating Signal



(b) Instantaneous Frequency of the Corresponding FM Signal



(c) Corresponding FM Signal

The phase angle θ(t) of FM signal is given by $\theta(t)=2\pi k_f \int_0^t m(\tau)d\tau$

Frequency deviation

Consider a sinusoidal modulating information signal given by $m(t)=V_m\cos(2\pi f_m t)$.

The instantaneous frequency of the resulting FM signal equals $f_i(t)=f_c+k_f V_m \cos(2\pi f_m t)$.

The maximum change in instantaneous frequency f_i from the carrier frequency f_c, is known as frequency deviation *Δf.*

$$\Delta f=k_f V_m$$

The frequency deviation is a useful parameter for determining the bandwidth of the FM-signals.

The phase deviation of FM signal is given by $\theta(t)=\frac{\Delta f}{f_m}\sin(2\pi f_m t)$

The ratio of the frequency deviation to the modulation frequency f_m is called the **modulation index** of the FM signal. We denote it by $m_f = \frac{\Delta f}{f_m}$. m_f is measured in radians. It represents the phase deviation of the FM signal. In other words m_f represents the maximum departure of the angle θ(t) from the angle "$2\pi f_c$ t" of the carrier.

The FM signal is thus given by $x(t) = V_c \cos(2\pi f_c t + m_f \sin(2\pi f_m t))$, while PM signal is given by

$$x(t)=V_c \cos(2\pi f_c t+k_p v_m \cos(2\pi f_m t)).$$

Depending on the value of the modulation index β, we may distinguish two cases of frequency modulation - Narrow-Band FM and Wide-Band FM.

FREQUENCY MODULATION

Frequency modulation is defined as the modulation in which frequency of the carrier signal is varied in accordance with amplitude of the modulating signal keeping the amplitude of modulated carrier constant.

The shift in carrier frequency from its resting point compared to amplitude of modulating voltage is called deviation ratio

$$\text{Deviation ratio} = \frac{f_{dev}(\max)}{f_m(\max)} \quad \text{where}$$

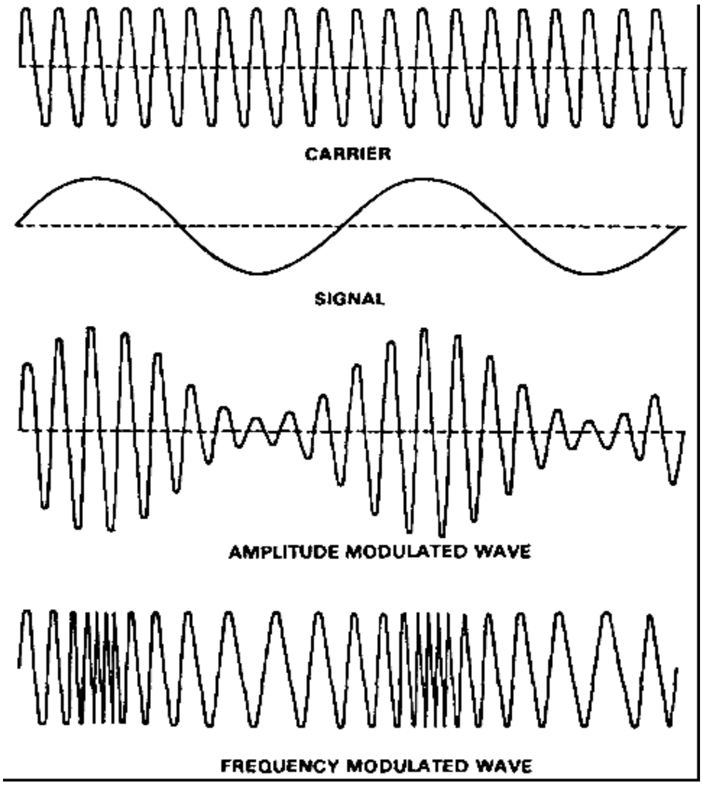
f_{dev}= frequency deviation of carrier
f_m =frequency of modulating voltage

The general equation of an unmodulated wave or carrier is given by $x = A\sin(\omega t+\theta)$

X= instantaneous value (voltage or current)
A= maximum amplitude
ω =angular velocity
θ = phase angle

By definition, amount by which carrier frequency is varied from its unmodulated wave called deviation is made proportional to the instantaneous amplitude of modulating voltage. The rate at which frequency deviation changes or takes place is equal to modulating frequency.

MATHEMATICAL REPRESENTATION OF FM



The instantaneous frequency of the resulting FM signal equals $f = f_c\left(1+kV_m\cos\omega_m t\right)$

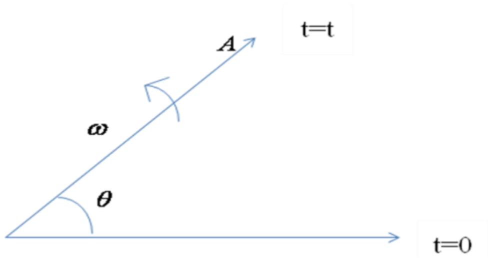
where f_c =unmodulated carrier frequency
k=proportionality constant
$V_m \cos\omega_m t$ =instantaneous modulating voltage

Maximum deviation of a signal occurs when cosine term has its maximum value ± 1. Under these condition instantaneous frequency becomes $f = f_c(1 \pm kV_m)$

The maximum deviation is given by $\delta = f_c kV_m$

Instantaneous amplitude of FM signal will be given by $v = A\sin\left[F\left(\omega_c, \omega_m\right)\right] = A\sin\theta$

where $\left[F\left(\omega_c, \omega_m\right)\right]$ is function of carrier and modulating frequencies



Frequency modulated vector

Consider the figure shown above to determine the value of θ which is traced by vector A in time t rotating at a constant angular velocity. In this instance angular velocity is constant. It is obtained by replacing f by angular velocity ω.

The instantaneous frequency of resulting FM signal now becomes $\omega = \omega_c\left(1+kV_m\cos\omega_m t\right)$

In order to find θ , ω must be integrated with respect to time.

$$\theta = \int \omega dt$$
$$= \int \omega_c \left(1 + kV_m \cos \omega_m t\right) dt$$
$$= \omega_c \int \left(1 + kV_m \cos \omega_m t\right) dt$$
$$= \omega_c \left(t + \frac{kV_m \sin \omega_m t}{\omega_m} \right)$$
$$= \left(\omega_c t + \frac{kV_m \omega_c \sin \omega_m t}{\omega_m} \right)$$

Now replacing ω by angular velocity f. the above equation becomes

$$= \left(\omega_c t + \frac{kV_m f_c \sin \omega_m t}{f_m} \right) \text{since } \delta = f_c kV_m$$
$$= \left(\omega_c t + \frac{\delta}{\omega_m} \sin \omega_m t \right)$$

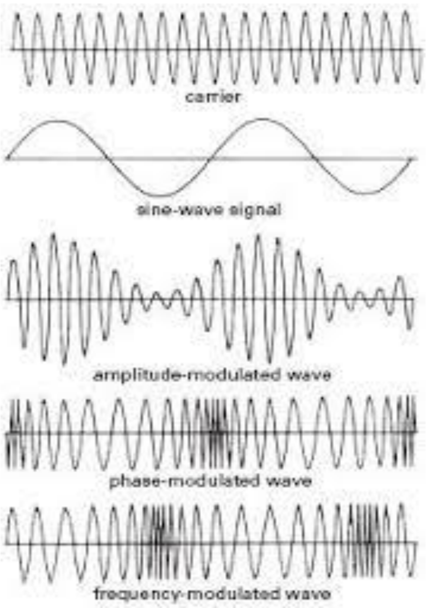
The instantaneous value of FM is given by $v = V_c \sin\left(\omega_c t + \frac{\delta}{\omega_m} \sin \omega_m t \right)$

The modulation index of FM is given by $m_f = \frac{\text{maximum frequency deviation}}{\text{modulating frequency}} = \frac{\delta}{f_m}$

Substituting in above equation we get $v = V_c \sin\left(\omega_c t + m_f \sin \omega_m t\right)$

If the modulating frequency decreases, maximum amplitude voltage remains constant, modulation index increases.

PHASE MODULATION



Comparison of AM, PM and FM

Phase modulation is defined as the modulation in which phase of the carrier signal is varied in accordance with amplitude of the modulating signal keeping amplitude and frequency of modulated carrier constant.

Let message signal be represented by $v_m(t)=V_m \cos \omega_m t$, carrier signal is represented by $v_c(t)=V_c \sin (\omega_c t+\theta)$ and the equation can also be rewritten as $v_c(t)=V_c \sin \phi$

V_m= Maximum amplitude of modulating signal
V_c= Maximum amplitude of carrier signal
ω_m= angular frequency of modulating signal
$\omega_c t$= angular frequency of carrier signal
θ = phase angle of carrier

Phase angle of carrier is varied in accordance with amplitude of modulating signal i.e.

$$\theta = K_p.v_m$$
$$= K_p.V_m \cos \omega_m t$$

where K_p =is phase deviation sensitivity

After phase modulation, instantaneous voltage is given by

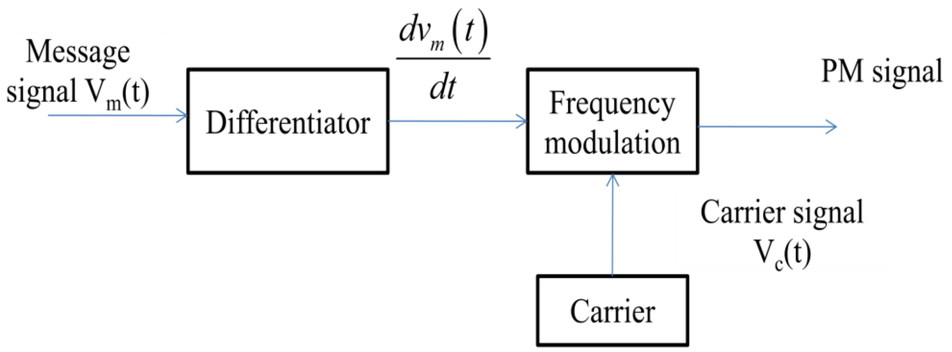
$$v = V_c \sin(\omega_c t + \theta)$$
$$= V_c \sin(\omega_c t + K_p V_m \cos \omega_m t)$$

$(\because m_{pm} = K_p V_m)$ Modulation index of phase modulation

$$v = V_c \sin(\omega_c t + m_p \cos \omega_m t)$$

RELATION BETWEEN FM AND PM

Conversion of FM to PM



Let message signal be represented by $v_m(t)=V_m \cos \omega_m t$, carrier signal is represented by $v_c(t)=V_c \sin (\omega_c t+\theta)$ and the equation can also be rewritten as $v_c(t)=V_c \sin \phi$

V_m= Maximum amplitude of modulating signal
V_c= Maximum amplitude of carrier signal
ω_m= angular frequency of modulating signal
$\omega_c t$=angular frequency of carrier signal
θ = phase angle of carrier
$\phi = (\omega_c t+\theta)$=total instantaneous phase angle of carrier

After differentiation the equation becomes $\frac{dv_m(t)}{dt} = -\omega_m.V_m Sin\omega_m t$

After frequency modulation, equation becomes
$$\omega_i = \omega_c + K[-\omega_m.V_m Sin\omega_m t]$$
$$= \omega_c - K\omega_m.V_m Sin\omega_m t$$

The instantaneous phase angle of frequency modulated signal is

$$\phi_i = \int \omega_i dt$$
$$= \int \left(\omega_c - K\omega_m.V_m Sin\omega_m t\right) dt$$
$$= \omega_c t + \frac{K\omega_m.V_m}{\omega_m} C\text{os}\,\omega_m t$$
$$= \omega_c t + K.V_m C\text{os}\,\omega_m t$$

The instantaneous voltage after modulation is given by

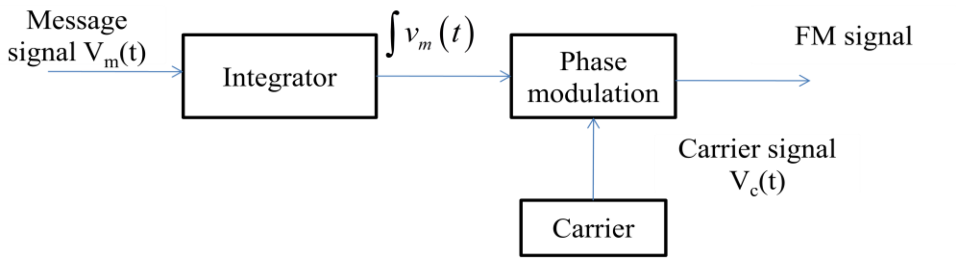
$$v_{pm}(t) = V_c Sin\phi_i$$
$$v_{pm}(t) = V_c Sin\left(\omega_c t + K.V_m C\text{os}\,\omega_m t\right)$$

By definition we know that

$$m_p = K.V_m$$
$$v_{pm}(t) = V_c Sin\left(\omega_c t + m_p C\text{os}\,\omega_m t\right)$$

The process of differentiation and integration are linear. The obtained equation is the expression for phase modulated wave.

Conversion of PM to FM



Let message signal be represented by $v_m(t)=V_m \cos \omega_m t$, carrier signal is represented by $v_c(t)=V_c \sin(\omega_c t+\theta)$ and the equation can also be rewritten as $v_c(t)=V_c \sin \phi$

V_m= Maximum amplitude of modulating signal
V_c= Maximum amplitude of carrier signal
ω_m= angular frequency of modulating signal
$\omega_c t$=angular frequency of carrier signal
$\phi=(\omega_c t+\theta)$ =total instantaneous phase angle of carrier

After integration the equation becomes $v_m(t) = \int V_m \cos\omega_m t.dt = \frac{V_m}{\omega_m} Sin\omega_m t$

After phase modulation, $\theta \,\alpha\, v_m(t)$

The equation becomes $\theta = K\, v_m(t)$

$$= \frac{KV_m}{\omega_m}.Sin\omega_m t$$

The instantaneous value of modulated voltage is $v_{fm}(t)=V_c \sin(\omega_c t+\theta)$

$$v_{fm}(t)=V_c \sin(\omega_c t+ \frac{KV_m}{\omega_m}.Sin\omega_m t\ \theta)$$

By definition we know that

$$m_f = \frac{K.V_m}{\omega_m} = \frac{\Delta f}{f_m} as\left[2\pi\Delta f = K.V_m\right]$$

$$v_{fm}\left(t\right) = V_c Sin\left(\omega_c t + m_f C\,os\,\omega_m t\right)$$

The obtained equation is the expression for frequency modulated wave.

SPECTRUM OF AN FM SIGNAL

Narrow-band Frequency Modulation

Sinusoidal Modulating Signal

For small values of β,

cos (β sin(2π f_m t) φ 1

sin (β sin(2π f_m t)) φ β sin(2π f_m t)

Thus the expression for FM signal can be expanded as

$$x(t)=V_c\cos(2\pi f_c t)-V_c\sin(2\pi f_c t)\ \beta\sin(2\pi f_m t)$$

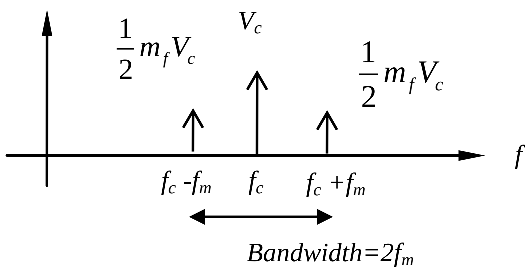
using $\cos(A+B)=\cos A\cos B-\sin A\sin B$

which may be written as follows

$$x(t)=V_c\cos(2\pi f_c t)+\frac{1}{2}\beta A_c\{\cos[2\pi(f_c+f_m)t]-\cos[2\pi(f_c-f_m)t]\}$$

using $\sin A\sin B=\frac{1}{2}[\cos(A-B)-\cos(A+B)]$.

Amplitude spectrum (single-sided plot)



Wide-band Frequency Modulation

The general expression for FM signal can be analyzed to give the spectral components of wide-band FM signal.

In order to compute the spectrum of an angle-modulated signal with a sinusoidal message signal, let

$$\theta(t)=\frac{\Delta f}{f_m}\ \sin(2\pi f_m t)$$

The corresponding FM signal

$$x(t)=V_c\cos(2\pi f_c t+\beta\sin(2\pi f_m t))$$

and may alternatively be written as

$$x(t)=V_c\mathrm{Re}\left(e^{j\omega_c t}e^{j\beta\sin 2\pi f_m t}\right)$$

Where Re(x) denotes the real part of x.

The parameter **β is known as the *modulation index*** and is the maximum value of phase deviation of the FM **signal.**

Consider the function z(t) given by

$$z(t)=e^{j\beta\sin 2\pi f_m t}$$

It is periodic with frequency f_m and can therefore be expanded in a Fourier series as follows.

$$z(t) = \sum_{-\infty}^{\infty} c_n \exp(j2\pi n f_m t)$$

The Fourier coefficient are given by

$$c_n = f_m \int_{-1/2f_m}^{1/2f_m} z(t) \exp(-j2\pi n f_m t) dt$$

$$c_n = f_m \int_{-1/2f_m}^{1/2f_m} \exp[j\beta \sin(2\pi f_m t) - j2\pi n f_m t) dt$$

Define a new variable $y = 2\pi f_m t$.

Hence we can rewrite c_n as

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin(y) - ny)] dy$$

The integral on the right hand side is a function of "n" and β and is known as the ***Bessel function* of the first kind of order n and argumentβ**. It is conventionally denoted by $J_n(\beta)$. That is,

$$J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin(y) - ny)] dy$$

Thus $c_n = J_n(\beta)$

and we get

$$z(t) = \sum_{-\infty}^{\infty} J_n(\beta) \exp(j2\pi n f_m t)$$

x(t) is accordingly given by

$$x(t) = V_c \cdot \operatorname{Re}\left[\sum_{n=-\infty}^{\infty} J_n(\beta) \quad \exp[j2\pi(f_c + nf_m)t] \right]$$

The discrete spectrum of x(t) is, therefore, given by

$$X(f) = \frac{V_c}{2} \sum_{-\infty}^{\infty} J_n(\beta) \quad [\delta(f - f_c - nf_m) + \delta(f + f_c + nf_m)]$$

Properties of Bessel functions

1. For n even, we have

   $J_n(\beta) = J_{-n}(\beta)$

   For n odd, we have

   $J_n(\beta) = (-1) J_{-n}(\beta)$

   Thus,

   $J_n(\beta) = (-1)^n J_{-n}(\beta)$

2. For small values of the modulation index β,

we have $J_0(\beta) \cong 1$

n	β = 0.05	β = 0.1	β = 0.2	β = 0.3	β = 0.5	β = 0.7	β = 1.0	β = 2.0	β = 3.0	β = 5.0	β = 7.0	β = 8.0	β = 10.0
0	0.999	0.998	0.990	0.978	0.938	0.881	0.765	0.224	−0.260	−0.178	0.300	0.172	−0.246
1	0.025	0.050	0.100	0.148	0.242	0.329	0.440	0.577	0.339	−0.328	−0.005	0.235	0.043
2		0.001	0.005	0.011	0.031	0.059	0.115	0.353	0.486	0.047	−0.301	−0.113	0.255
3				0.001	0.003	0.007	0.020	0.129	0.309	0.365	−0.168	−0.291	0.058
4						0.001	0.002	0.034	0.132	0.391	0.158	−0.105	−0.220
5								0.007	0.043	0.261	0.348	0.186	−0.234
6								0.001	0.011	0.131	0.339	0.338	−0.014
7									0.003	0.053	0.234	0.321	0.217
8										0.018	0.128	0.223	0.318
9										0.006	0.059	0.126	0.292
10										0.001	0.024	0.061	0.207
11											0.008	0.026	0.123
12											0.003	0.010	0.063
13											0.001	0.003	0.029
14												0.001	0.012
15													0.005
16													0.002
17													0.001

Table of Bessel functions

$J_1(\beta) \cong \beta/2$

$J_3(\beta) \cong 0$ for n > 2

3. $\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1$

- The first column gives the sideband number, while the first row gives the modulation index. The remaining columns indicate the amplitudes of the carrier and the various pairs of sidebands.
- Sidebands with relative magnitude of less than 0.001 have been eliminated.
- Some of the carrier and sideband amplitudes have negative signs. This means that the signal represented by that amplitude is simply shifted in phase 180° (phase inversion). As you can see, the spectrum of a FM signal varies considerably in bandwidth depending upon the value of the modulation index. The higher the modulation index, the wider the bandwidth of the FM signal.

Amplitude Spectrum

Typical example of the normalized amplitude spectrums i.e., $\frac{|S(f)|}{A/2}$ are shown below. With the increase in the modulation index, the carrier amplitude decreases while the amplitude of the various sidebands increases. With some values of modulation index, the carrier can disappear completely.

Power in angle-modulated signal

The power in an angle-modulated signal is easily computed

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

$$= V_c^2 / 2.$$

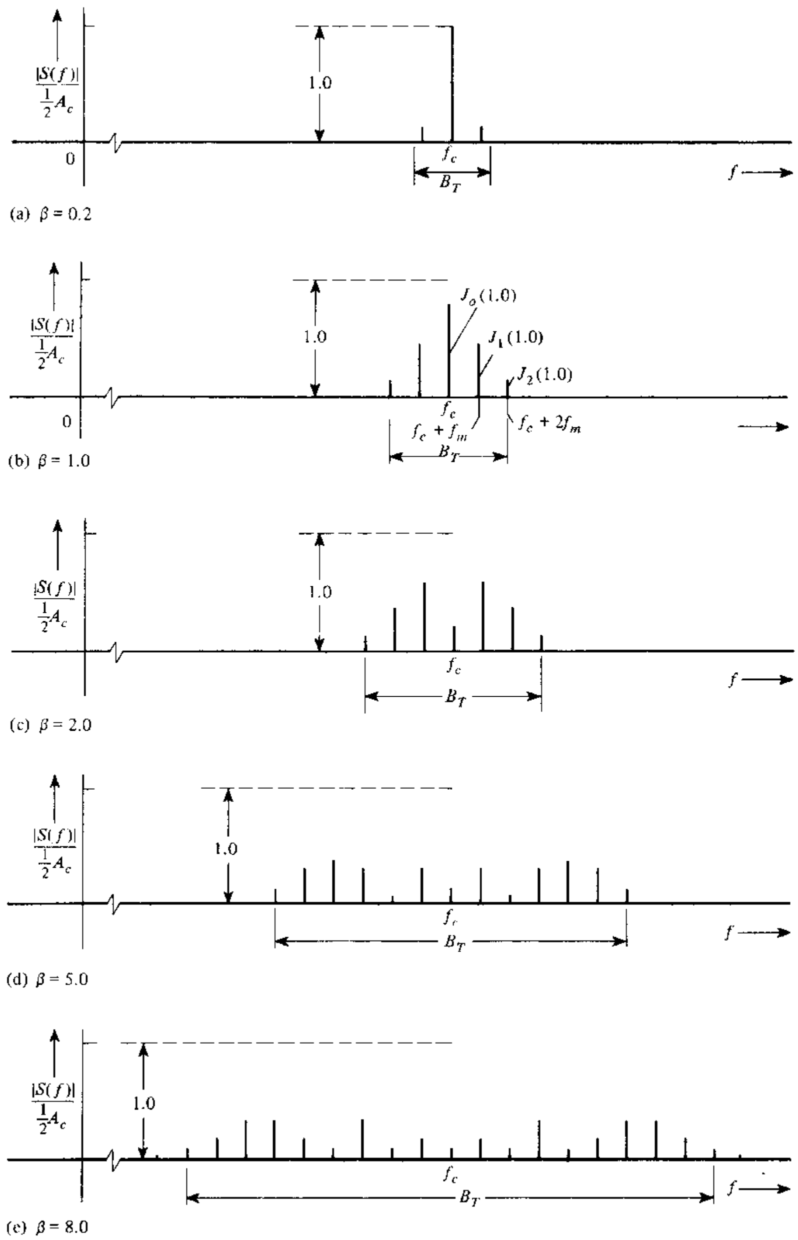
- Thus the power contained in the FM signal is independent of the message signal. This is an important difference between FM and AM.
- The time-average power of an FM signal may also be obtained from (where $\langle x(t) \rangle = \frac{1}{T} \int_T x(t)dt$ indicates average over duration of a period)

$$x(t) = A_c \cos(2\pi f_c t + \theta(t))$$

$$\langle x^2(t)\rangle = \langle V_c^{\,2}\cos^2(2\pi f_c t + \theta(t))\rangle$$
$$\langle x^2(t)\rangle = \frac{1}{2}V_c^{\,2} + \frac{1}{2}V_c^{\,2}\langle\cos 2(2\pi f_c t + \theta(t))\rangle$$
$$\langle x^2(t)\rangle = \frac{1}{2}V_c^{\,2}.$$

Thus *power*, $P = \frac{1}{2}V_c^{\,2}$.

Transmission Bandwidth of FM signals



Magnitude spectra for FM or PM with sinusoidal modulation for various modulation indexes

- Theoretically, a FM signal contains an infinite number of side frequencies so that the bandwidth required to transmit such signal is infinite.

- However, since the values of $J_n(\beta)$ become negligible for sufficiently large n, the bandwidth of an angle-modulated signal can be defined by considering only those terms that contain significant power.

In practice, the bandwidth of a FM signal can be determined by knowing the modulation index and using a Bessel function table (showing carrier and sideband amplitudes for different values of modulation index of FM signals).

Determine BW with Table of Bessel Functions

Example:

Assume that the modulation index is 2.

Referring to the table, we can see that this produces four significant pairs of sidebands. The bandwidth can then be determined with the simple formula

$$BW = 2\,N\,f_{max,}$$

where N is the number of significant sidebands. Using the example above and assuming a highest modulating frequency of 2.5 kHz, the bandwidth of the FM signal is

$$BW = 2\ (6)\ (2.5) = 30\ kHz$$

A FM signal with a modulation index of 2 and a highest modulating frequency of 2.5 kHz will then occupy a bandwidth of 30 kHz.

Determine BW with Carson's rule

An alternative way to calculate the bandwidth of a FM signal is to use Carson's rule. This rule takes into consideration only the power in the most significant sidebands whose amplitudes are greater than 2 percent of the carrier. These are the sidebands whose values are 0.02 or more.

Carson's rule is given by the expression

$$B_T = BW \cong 2\Delta f + 2 f_m$$

In this expression, Δf is the maximum frequency deviation, and f_m is the maximum modulating frequency.

We may thus define an approximate rule for the transmission bandwidth of an FM signal generated by a single of frequency f_m as follows:

$$B_T = BW \cong 2\Delta f + 2 f_m = 2\Delta f\,(1 + \frac{1}{\beta})$$

Example: Assuming a maximum frequency deviation of 5 kHz and a maximum modulating frequency of 2.5 kHz, the bandwidth would be

BW = 2 (5 kHz + 2.5 kHz)

= 2 (7.5 kHz) = 15 kHz.

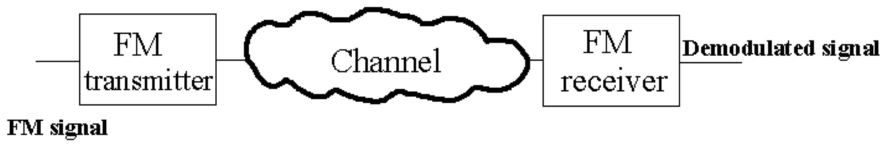
Comparing the bandwidth with that computed in the preceding example, you can see that Carson's rule gives a smaller bandwidth.

Noise in FM signal

- Noise is essentially amplitude variations. An FM signal, on the other hand, has constant carrier amplitude.
- Because of this, FM receivers contain limiter circuits that restrict the amplitude of received signal.
- Any amplitude variations occurring on the FM signal are effectively clipped off. This does not hurt the information content of the FM signal. Because of the clipping action of the limiter circuits, noise is almost completely eliminated.

Interference

- A major benefit of FM is that interfering signals on the same frequency will be effectively rejected.
- If the signal of one is more than twice the amplitude of the other, the stronger signal will "capture" the channel and will totally eliminate the weaker, interfering signal.



- This is known as the *capture effect* in FM.

- In FM, the capture effect allows the stronger signal to dominate while the weaker signal is eliminated. However, when the strengths of the two FM signals begin to be nearly the same, the capture effect may cause the signals to alternate in their domination of the frequency.

- At some time one signal will be stronger than the other, and it will capture the channel. At other times, the signal strengths will reverse, and the other signal will capture the channel.

- We may have experienced this effect yourself when listening to the FM radio in your car while driving on the highway.

- We may be listening to a strong station on a particular frequency, but as you drive, you move away from that station. At some point, you may begin to pick up the signal from another station on the same frequency.
- When the two signals are approximately the same amplitude, you will hear one station dominate and then the other as the signal amplitudes vary during your driving.
- However, at some point, the stronger signal will eventually dominate. In any case, once the strong signal dominates, the weaker is not heard at all on the channel.

- Despite the fact that FM has superior noise rejection qualities, noise still interferes with an FM signal. This is particularly true for the high-frequency components in the modulating signal.

- Since noise is primarily sharp spikes of energy, it contains a considerable number of harmonics and other high-frequency components.

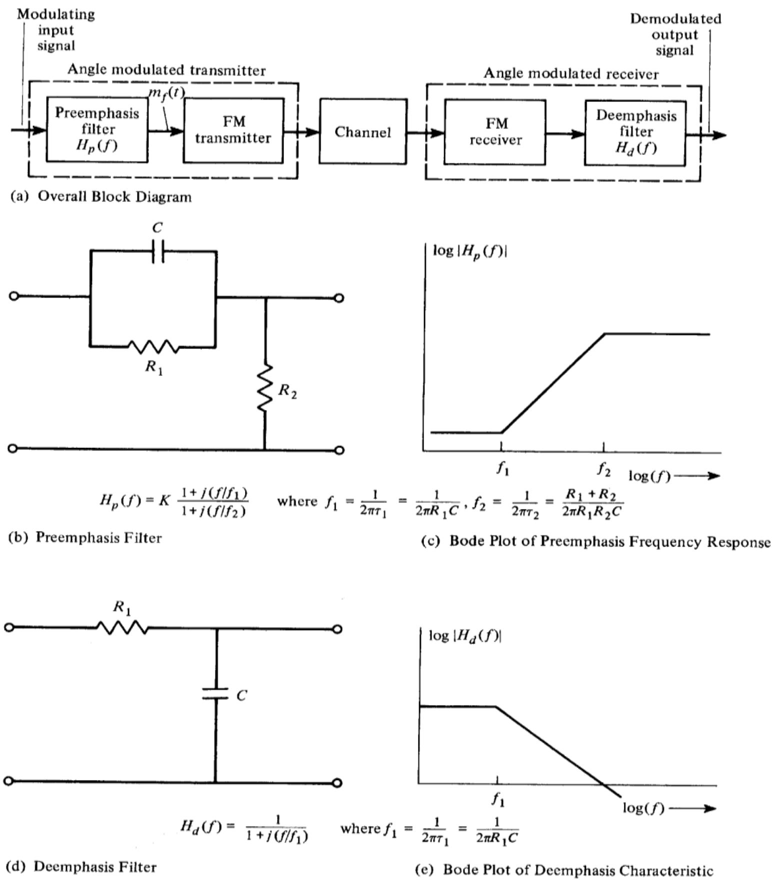
- These high frequencies can at times be larger in amplitude than the high-frequency content of the modulating signal.

- This causes a form of frequency distortion that can make the signal unintelligible.

To overcome this problem Most FM system uses a technique known as Pre-emphasis.

Pre-emphasis and De-emphasis

- At the transmitter the modulating signal is passing through a simple network which amplifies the high frequency component more the low-frequency component
- The simplest form of such circuit is a simple high pass filter
- The pre-emphasis circuit increases the energy of the higher content of the higher-frequency signals so that will tend to become stronger than the high-frequency noise component.
- This improves the signal-to-noise ratio.
- To return the frequency response to its normal level, a de-emphasis circuit is used at the receiver.
- The de-emphasis circuit provides a normal frequency response.
- The combined effect of pre-emphasis and de-emphasis is to increase the high-frequency components during the transmission so that they will be stronger and not masked by noise



Angle modulated system with pre emphasis and de emphasis

Advantages in FM over AM

1. Amplitude of frequency –modulated wave is constant. It is thus independent of modulation depth, whereas in AM modulation depth governs the transmitted power. This means that, in FM transmitters, low –level modulation may be used but all the subsequent can be class c and therefore more efficient. Since all these amplifiers will handle constant power, they need not be capable of man-aging up to four times the average power, as they must in AM. All transmitted power in FM is useful, whereas in AM most of it is in the transmitted carrier, which contains no useful information.
2. FM receivers can be fitted with amplitude limiters to remove the amplitude variations caused by noise this make FM reception a good deal more immune to noise than AM reception.
3. It is possible to reduce noise still further by increasing the deviation. AM does not have this feature, since it is not possible to exceed 100 percent modulation without causing severe distortion.
4. Commercial FM broadcasts have a number of advantages due to better planning and provide guard band between FM stations, so that there is less adjacent-channel interference than in AM
5. FM broadcasts operate in the upper VHF and UHF frequency ranges, at which there happens to be less noise than in the MF and HF ranges occupied by AM broadcasts
6. At FM broadcast frequencies, space wave is used for propagation, so that radius of operation is slightly more than line of sight. It is thus possible to operate several independent transmitters on same frequency with less interference than would be possible in AM.

Disadvantages in FM

1. A much wider channel is required in FM up to 10 times as large as that needed by AM
2. FM transmitter and receiver tend to be more complex.
3. Since reception is limited to line of sight, area of reception for FM is much smaller than AM.

FM Signal Generation

If either the capacitance or inductance of an LC oscillator tank is varied, frequency modulation of some form will result. If this variation can be made directly proportional to the voltage supplied by the modulation circuits, true FM will be obtained.

If such a system is used, a voltage-variable reactance is placed across the tank. When the modulating voltage is zero, the variable resistance will have its average value.

They are two basic methods of generating frequency-Modulated signals

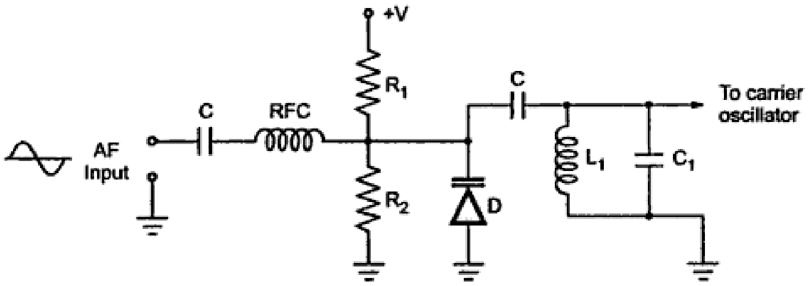
1. Direct Method

$$f_i = f_c + k_f m(t)$$

- In a direct FM system the instantaneous frequency is directly varied with the information signal.
- To vary the frequency of the carrier is to use an Oscillator whose resonant frequency is determined by components that can be varied.
- The oscillator frequency is thus changed by the modulating signal amplitude.

For example, an electronic Oscillator has an output frequency that depends on energy-storage devices. There are a wide variety of oscillators whose frequencies depend on a particular capacitor value. By varying the capacitor value, the frequency of oscillation varies. If the capacitor variations are controlled by *m (t)*, the result is an FM waveform

VARACTOR DIODE MODULATOR



Varactor diode FM generation

- FM modulator which is widely used circuitry uses a voltage-variable capacitor (VARACTOR). All diodes exhibit small junction capacitance in reverse biased condition.
- The varactor diode is designed to optimize this characteristic.
- The varactor is a simple diode, or PN junction that is designed to have a certain amount of capacitance between junctions.
- The capacitance of a varactor, as with regular capacitors, is determined by the area of the capacitor plates and the distance between the plates.
- The depletion region in the varactor is dielectric and is located between the P and N elements, which serve as the plates.
- Capacitance is varied in varactor by varying reverse bias which controls the thickness of depletion region.
- The varactor is so designed that change in capacitance is linear with change in applied voltage. This is a special design characteristic of varactor diode.
- It must not be forward biased because it cannot tolerate much current flow. Proper circuit design prevents the application of forward bias.
- The varactor diodes provide reverse junction capacitance in the range of 1 to 200 pF. L_1 and C_1 form the tank circuit of the carrier oscillator. Either R_1 or R_2 is made variable so that center frequency can be adjusted over a narrow range.
- The radio frequency choke (RFC) has high reactance at carrier frequency to prevent carrier signal from getting into modulating signal.
- An RF signal is applied as input results in the following actions:
  (1) On positive alternation, reverse bias increases and the dielectric (depletion region) width increases. This decreases capacitance which increases the frequency of the oscillator.
  (2) On the negative alternation, reverse bias decreases, this result in a decrease in oscillator frequency.

The frequency of LC oscillator changes due to temperature effects. Hence crystal oscillator is used in FM generators to provide frequency stability.

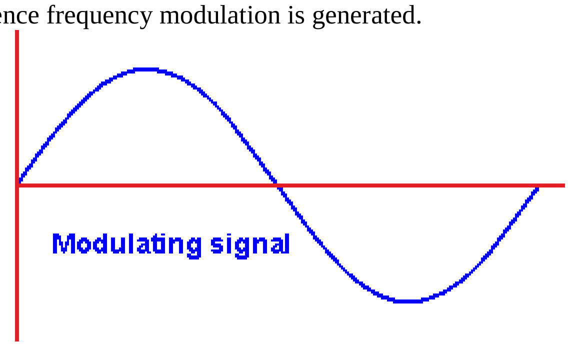
- The capacitance C_d of varactor diode is expressed as $C_d = \frac{k}{\sqrt{v_D}} = k\left(v_D\right)^{-1/2}$

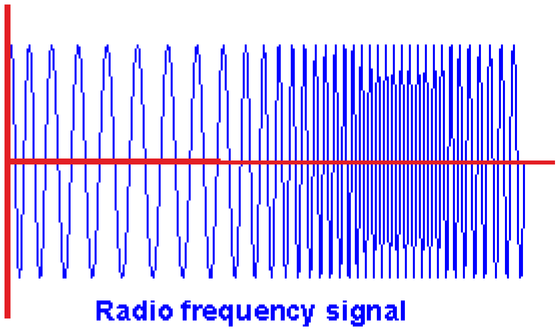
v_D = *total instantaneous voltage across varactor diode*
K = *proportionality constant*

- It is given by $v_D = V_o + x(t)$
- The oscillation of frequency is given by $\omega_c = \frac{1}{\sqrt{LC}}$
- The total capacitance of oscillator tank circuit will be $C_o + C_d$
- The instantaneous frequency of oscillation is given by $\omega_i = \frac{1}{\sqrt{L_o(C_o + C_d)}}$

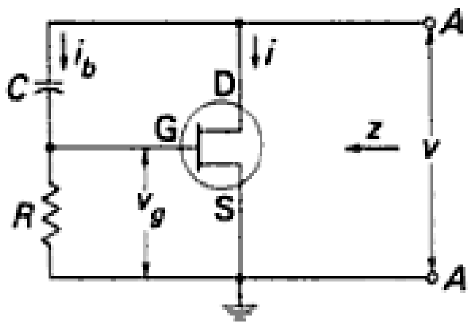
Substituting C_d in above equation we get $\omega_i = \frac{1}{\sqrt{L_o(C_o + k\left(v_D\right)^{-1/2})}}$

The instantaneous frequency ω_i of FM signal depends on v_D which in turn depends upon value of modulating signal x(t). Thus the instantaneous oscillator frequency ω_i also depends on baseband or modulating signal and hence frequency modulation is generated.





REACTANCE TUBE MODULATOR



- The circuit is the three-terminal reactance that may be connected across the tank circuit of the oscillator to be frequency-modulated.

- The value of this reactance is proportional to the transconductance of the device, which can be made to depend on the gate bias and its variations.

Theory of reactance modulators:

- In order to determine z, a voltage v is applied to the terminals A-A between which the impedance is to be measured, and the resulting current i is calculated.
- The applied voltage is then divided by this current, giving the impedance seen when looking into the terminals. In order for this impedance to be a pure reactance (it is capacitive here), two requirements must be fulfilled.
- The first is that the bias network current i_b must be negligible compared to the drain current. The impedance of the bias network must be large enough to be ignored.
- The second requirement is that the drain-to-gate impedance (X_C here) must be greater than the gate-to-source impedance (R in this case), preferably by more than 5:1.
- The following analysis may then be applied:

$$v_g = i_b R = \frac{R}{R - jX_c}$$

- The FET drain current is $i = g_m v_g = \frac{g_m R_v}{R - jX_c} = \frac{1}{g_m}\left(1 - \frac{jX_c}{R}\right)$
- If $X_c \gg R$ in above equation will reduce to $z = -\frac{jX_c}{R}$
- The impedance seen at terminal A-A is

$$\begin{aligned} z &= \frac{v}{i} \\ &= v \div \frac{g_m R_v}{R - jX_c} \\ &= \frac{1}{g_m}\left(1 - \frac{jX_c}{R}\right) \end{aligned}$$

- The impedance is quite clearly a capacitive reactance

$$\begin{aligned} X_{eq} &= \frac{X_c}{g_m R} \\ &= \frac{1}{2\pi f g_m RC} \\ &= \frac{1}{2\pi f C_{eq}} \end{aligned}$$

- From above equation, input impedance of device at A-A is a pure reactance is given by

$$C_{eq} = g_m RC$$

The following should be noted from the above equation

1. This equivalent capacitance depends on the device transconductance and can therefore be varied with bias voltage.
2. Capacitance can be originally adjusted to any value, within reason, by varying R and C.
3. The expression g_m RC has correct dimensions of capacitance; R, is measured in ohms, and g_m, measured in Siemens (s), cancel each other's dimensions, leaving C as required.
4. It was stated earlier that the gate-to-drain impedance must be much larger than the gate-to-source impedance. If X_C/R had not been much greater than unity, z would have had a resistive component as well.

$$X_c = \frac{1}{\omega C} = nR$$

$$C = \frac{1}{\omega nR} = \frac{1}{2\pi fnR}$$

Substituting the values of C and X_c in C_{eq} we get

$$C_{eq} = g_m RC$$

$$= \frac{g_m R}{2\pi fnR}$$

$$= \frac{g_m}{2\pi fn}$$

2. Indirect Method

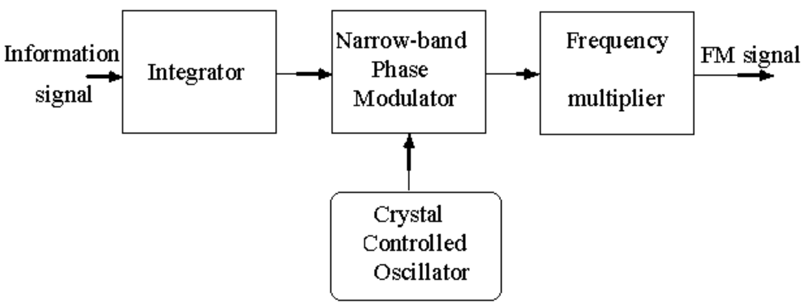
$$x(t) = V_c \cos[2\pi f_c t + \theta(t)]$$

$$\theta(t) = 2\pi k_p m(t)$$

$$\theta(t) = 2\pi k_f \int_0^t m(\tau)d\tau$$

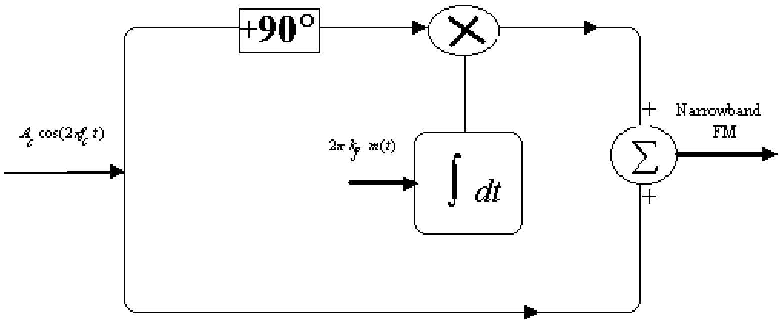
- Angle modulation includes frequency modulation FM and phase modulation PM. FM and PM are interrelated; one cannot change without the other changing.
- The information signal frequency also deviates the carrier frequency in PM.
- Phase modulation produces frequency modulation. Since the amount of phase shift is varying, the effect is as if the frequency is changed.
- Since FM is produced by PM, the later is referred to as indirect FM.
- The information signal is first integrated and then used to phase modulate a crystal-controlled oscillator, which provides frequency stability.
- In order to minimize the distortion in the phase modulator, the modulation index is kept small, thereby is resulting in a ***narrow-band FM-signal***
- The narrow-band FM signal is multiplied in frequency by means of frequency multiplier so as to produce the desired wide-band FM signal.
- The frequency multiplier is used to perform narrow band to wideband conversion.
- The frequency deviation of this new waveform is "*M*" times that of the old, while the rate at which the instantaneous frequency varies has not changed

Block diagram of the indirect method of generating a wide-band FM-signal



For high enough values of M, frequency multiplication changes narrowband FM into wideband FM. It also moves the carrier frequency, but the carrier has no effect on whether an FM wave is narrowband or wideband

Modulator for narrowband FM



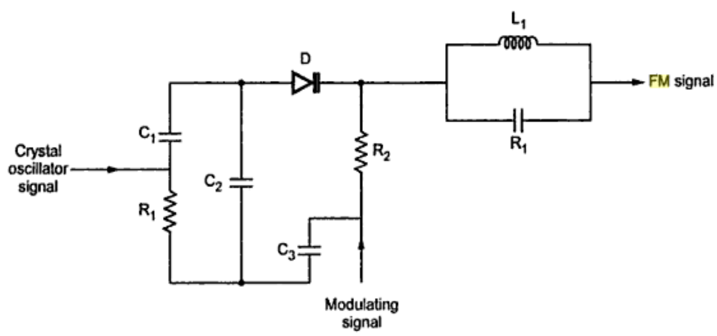
Narrowband FM: $x(t) = V_C \cos(2\pi f_C t) - V_C \sin(2\pi f_C t)\ \beta \sin(2\pi f_m t)$

VARACTOR DIODE MODULATOR

- The circuit consists of varactor diode D_1 in series with tuned L_1R_1 network.
- The complete series and parallel network is in series with crystal oscillator.
- The modulating signal is applied to varactor diode. The capacitance of varactor diode is changed by modulating signal.
- This changes the phase angle of complete network and creates the phase shift in carrier signal from crystal oscillator.
- The instantaneous phase shift is directly proportional to instantaneous amplitude of modulating signal.

The advantage is crystal oscillator is isolated from modulator so frequency stability will be more.

Disadvantages are Capacitance and voltage characteristic is non linear which results distortion in modulated waveform so amplitude of modulating signal should be kept low.

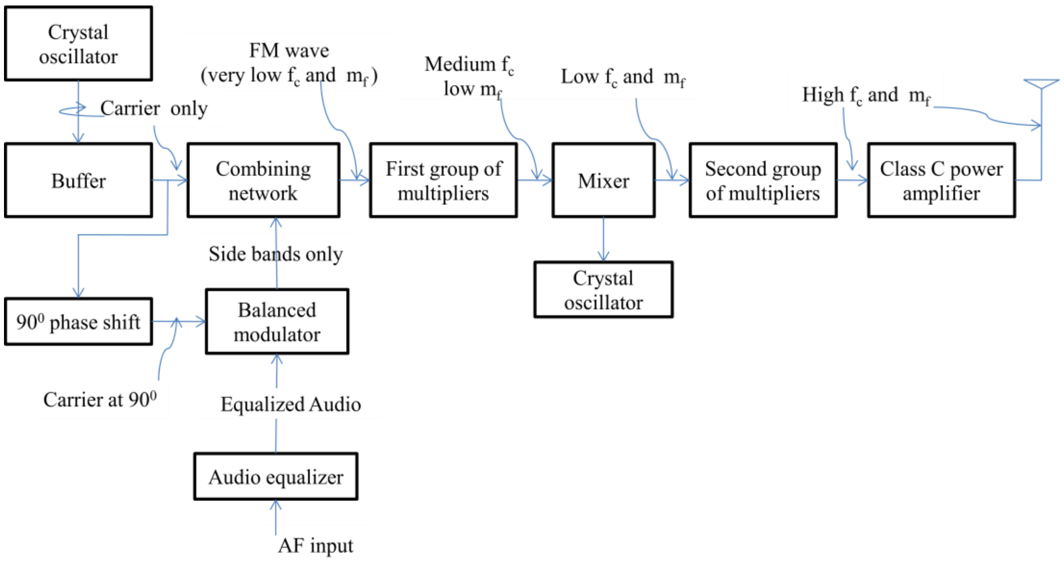


Varactor diode FM generation

ARMSTRONG METHOD

- The block diagram of an Armstrong system is shown.
- The system terminates at the output of the combining network; the remaining blocks are included to show how wideband FM might be obtained.
- The effect of mixing on an FM signal is to change the center frequency only, whereas the effect of frequency multiplication is to multiply center frequency and deviation equally.
- The carrier of the amplitude-modulated signal has been removed so that only the two sidebands are added to the unmodulated voltage.
- This has been accomplished by the balanced modulator, and the addition takes place in the combining network.
- The resultant of the two sideband voltages will always be in quadrature with the carrier voltage.
- As the modulation increases, so will the phase deviation, and hence phase modulation has been obtained.
- The resultant voltage coming from the combining network is phase-modulated, but there is also a little amplitude modulation present.

- The AM is no problem since it can be removed with an amplitude limiter.



Armstrong Method

- The output of amplitude limiter is phase modulation. Since frequency modulation is the requirement, modulating voltage will have to be equalized before it enters the balanced modulator.
- If a frequency-modulated signal $f_c \pm \delta$ is fed to a frequency doubler, the output signal will contain twice each input frequency.
- For the extreme frequencies here, this will be $2f_c$ - 2δ and $2f_c + 2\delta$.
- The frequency deviation has quite clearly doubled to $\pm 2\delta$, with the result that the modulation index has also doubled.
- In this fashion, both center frequency and deviation may be increased by the same factor or, if frequency division should be used, reduced by same factor.
- When a frequency-modulated wave is mixed, the resulting output contains difference frequencies (among others).
- The original signal might again be $f_c \pm \delta$. When mixed with a frequency f_0, it will yield $f_c - f_0 - \delta$ and $f_c - f_0 + \delta$ as the two extreme frequencies in its output.
- It is seen that the FM signal has been translated to a lower center frequency $f_c - f_0$, but the maximum deviation has remained a $\pm$ δ.
- It is possible to reduce (or increase, if desired) the center frequency of an FM signal without affecting the maximum deviation.
- Since modulating frequency has obviously remained constant in the two cases treated, modulation index will be affected in the same manner as the deviation.
- It will thus be multiplied together with center frequency or unaffected by mixing.
- Also it is possible to raise the modulation index without affecting the center frequency by multiplying both by 9 and mixing the result with a frequency eight times the original frequency.
- The difference will be equal to the initial frequency, but the modulation index will have been multiplied ninefold.

FM Transmitter

Depending upon the modulation there are 2 types

- i) Direct method ii) indirect method

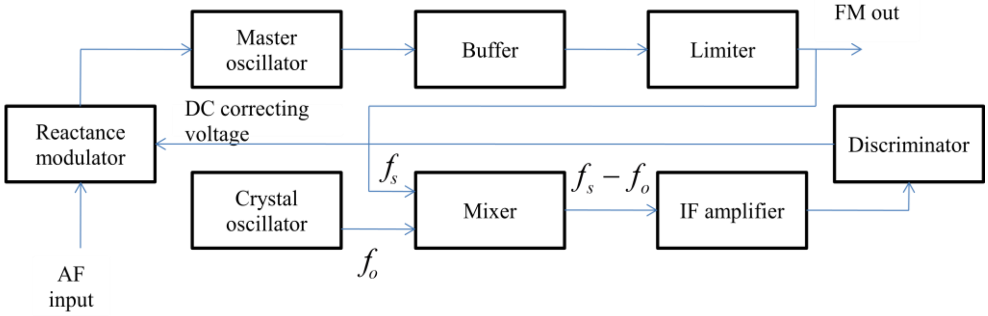
i) Direct method

- The transmitter produces FM signal whose frequency deviation is directly proportional to amplitude of modulating signal.
- Therefore carrier oscillator frequency is directly deviated. For this purpose crystal oscillator cannot be used since their frequency cannot be varied significantly.

ii) Indirect method

- The transmitter produces FM signal whose phase deviation is directly proportional to amplitude modulating signal.
- The frequency of carrier oscillator is not directly varied by modulating signal. Crystal oscillator are used in Indirect method to maintain stability.

A typical AFC Transmitter

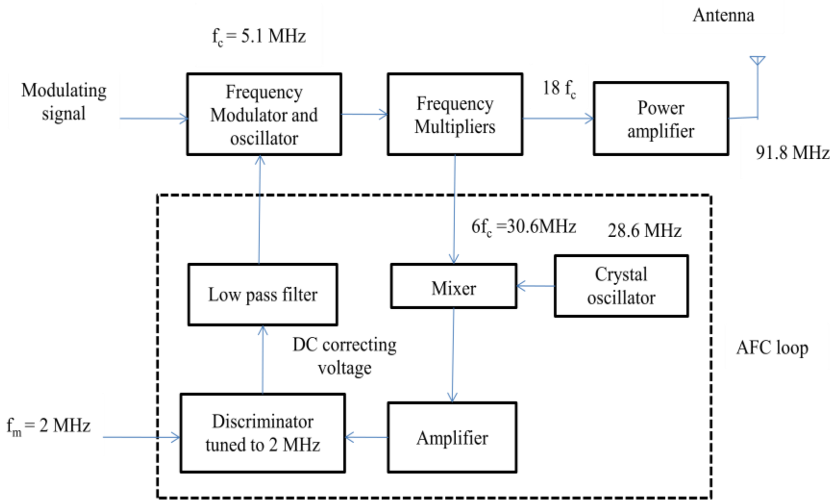


AFC TRANSMITTER

- AFC circuits are used to control frequency of an oscillator by some external signal.
- Basically, this type circuit does two things:
- It senses difference between actual oscillator frequency and frequency that is desired and produces a control voltage proportional to the difference; it also uses the control voltage to change the oscillator to desired frequency.
- AFC circuits are used to control frequency of sinusoidal oscillators and nonsinusoidal oscillators.
- AFC circuits are used in radio receivers, FM transmitters, and frequency synthesizers to maintain frequency stability.
- The discriminator reacts only to small changes in carrier frequency but not to frequency deviations in the carrier.
- If the frequency of carrier increases higher frequency is fed to mixer for which other input frequency is from a stable crystal oscillator.
- A high frequency will be fed to discriminator comparatively. Since the discriminator is fed tuned to correct frequency difference will exist between LC oscillator and crystal oscillator.
- Its input frequency is higher and discriminator will develop a positive d.c. voltage.
- This voltage is applied to reactance modulator whose transconductance is increased by positive voltage.
- It increases equivalent capacitance of reactance modulator thereby decreasing oscillator frequency. The increase in carrier frequency is thus lowered and brought to correct value.
- The correcting d.c. voltage developed by discriminator is fed to reactance modulator connected across tank circuit of oscillator and used for AFC purposes.
- In directly modulated FM transmitters, many times frequency modulation is carried at lower frequency and with smaller frequency deviation.
- Then the frequency modulated wave is passed through frequency multiplier circuit to achieve desired carrier frequency and desired frequency deviation.

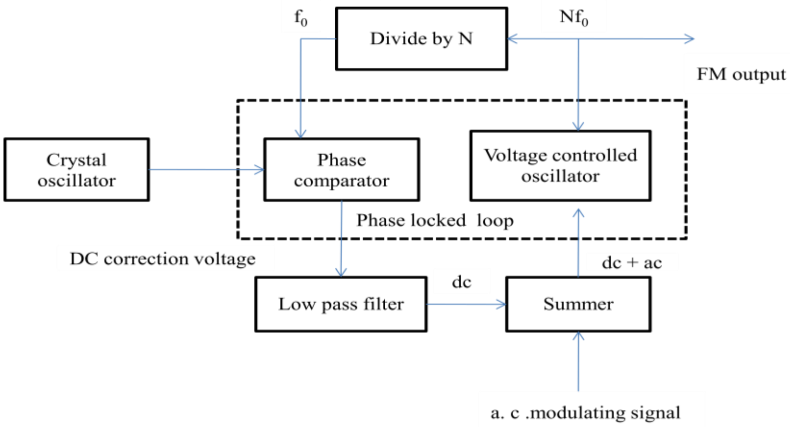
DIRECT FM TRANSMITTERS

Crossby Direct FM transmitters



- The modulating frequency is given to frequency modulator and oscillator.
- The frequency modulator can be reactance modulator or voltage controlled oscillator.
- The frequency of unmodulated carrier is f_c=5.1 MHz Its frequency is multiplied by 18 to generate transmitted frequency of 91.8 MHz. It's the center frequency of FM signal.
- The AFC loop is used to maintain the center frequency of unmodulated carrier stable.
- The multiplier output given to mixer is $6f_c$ = 30.6 MHz .
- The crystal reference oscillator generates 28.6 MHz .
- The mixer generates 2 MHz difference of these two frequencies and gives it to discriminator through amplifier.
- The discriminator is tuned to 2 MHz .If there is difference in mixer output frequency, discriminator generates DC correction voltage.
- If the multiplier frequency is exactly $6f_c$ = 30.6 MHz, no frequency correction is required and DC correction voltage becomes zero. However the frequency of $6f_c$ contains FM.
- This means there is frequency deviation of $6f_c$ depending on modulating signal. Hence DC correction voltage also contains corresponding variation.
- Therefore DC correction voltage is passed through low pass filter to remove the effect of modulation. Such filtered voltage is then used for frequency correction.

PHASE LOCKED LOOP Direct FM transmitter

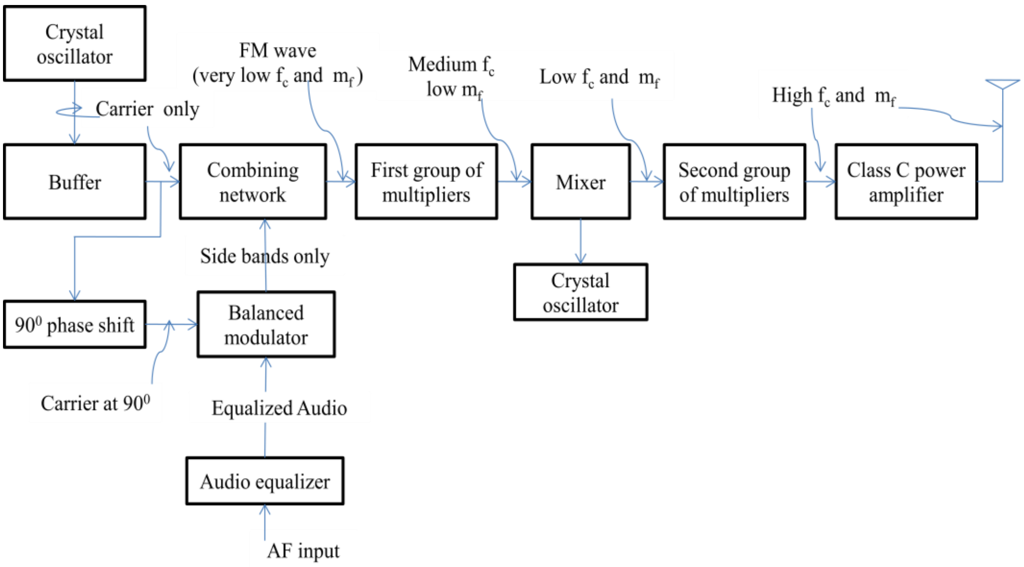


- These transmitters are used to generate high index wideband FM signal.
- When both inputs frequencies of phase comparator are same they are locked to each other. Under this condition phase comparator output is zero.
- It is passed through low pass filter to summer.
- The summer has modulating signal as another input.
- This modulating signal is used to control output frequency of VCO.
- Thus the output of VCO is FM signal whose frequency depends on modulating signal.
- The output of VCO is divided by N and given to phase comparator.
- If there is any shift in center frequency of VCO phase comparator generates DC correction voltage which is given to summer through low pass filter.

- This correction voltage adds to modulating signal voltage and corrects the VCO output.
- The low pass filter is used to remove rapid changes in correction voltage due to frequency variation in FM signal.

INDIRECT METHOD

ARMSTRONG FM TRANSMITTER



- Let message signal be represented by $v_m(t)=V_m \cos \omega_m t$,
- The carrier signal is represented by $v_c(t)=V_c \sin \phi$ (or) $v_c(t)=V_c \sin (\omega_c t+\theta)$

  V_m= *Maximum amplitude of modulating signal*
  V_c= *Maximum amplitude of carrier signal*
  ω_m= *angular frequency of modulating signal*
  $\omega_c t$= *angular frequency of carrier signal*
  $\phi=(\omega_c t+\theta)$=*total instantaneous phase angle of carrier*

- As the carrier is given a phase shift of 90^0 the equation becomes $v_c(t)=V_c \cos \omega_c t$ if $\theta= 90^0$
  Input to balanced modulator is

$$= \int V_m Sin\omega_m t.dt$$

$$= \frac{V_m}{\omega_m}\cos\omega_m t$$

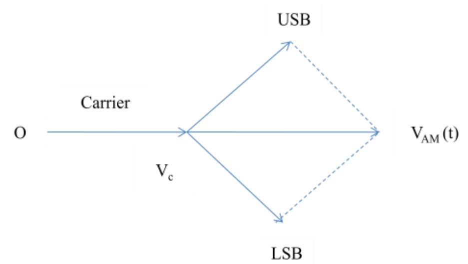
- Balanced modulator output

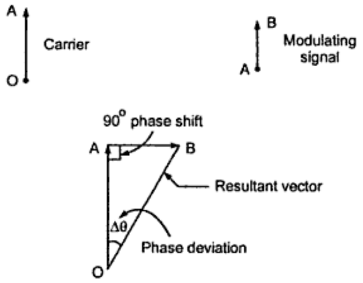
$$= \frac{V_m V_c}{\omega_m}\cos\omega_m t\cos\omega_c t$$

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

- Output of the mixer $= V_c \sin\omega_c t + \dfrac{V_m V_c}{2\omega_m}\left[\cos(\omega_c+\omega_m)t+\cos(\omega_c-\omega_m)t\right]$

The phasor diagram for above equation is


USB
Carrier
O
$V_{AM}(t)$
V_c
LSB

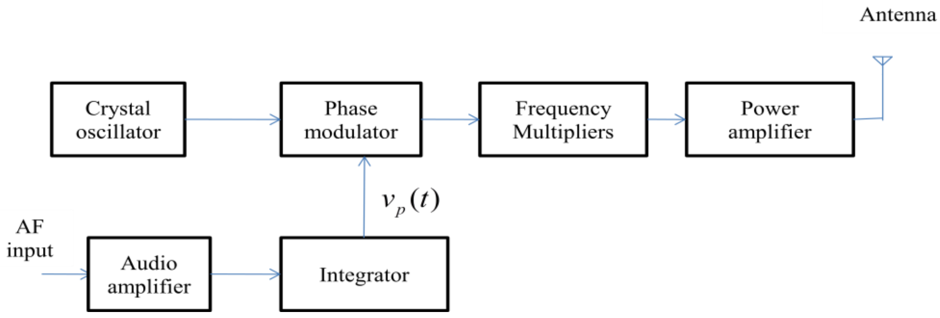

A
Carrier
O
B
Modulating signal
A
90^o phase shift
A
B
Resultant vector
$\Delta\theta$
Phase deviation
O

- Peak phase deviation $\phi_p = \tan^{-1}\left(\frac{V_m V_c}{2\omega_m}\right)$
- The mixer output contains phase deviation ϕ_p and also contains amplitude modulation components however if side bands components have very low amplitude, the amplitude modulation is negligible.
- The frequency deviation produced by the system equals.

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

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

INDIRECT METHOD FM TRANSMITER



- Let us assume that signal given to the transmitter be $V_m Sin\omega_m t$
- The input to the phase modulator given by integrator is

$$v_p(t) = \int V_m Sin\omega_m t.dt$$
$$= \frac{V_m \cos\omega_m t}{\omega_m}$$

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

$$\theta \propto v_p(t)$$
$$or$$
$$\theta = K v_p(t)$$
$$= \frac{K V_m \cos\omega_m t}{\omega_m}$$

We know that

$$\omega_i = \frac{d\theta}{dt}$$
$$= -\frac{d}{dt}\left(\frac{K V_m \cos\omega_m t}{\omega_m}\right)$$
$$= K V_m Sin\omega_m t$$

- Frequency deviation $\Delta f = K V_m$
- The derivation results in frequency modulation with deviation proportional to amplitude of modulating signal

AMPLITUDE MODULATION	FREQUENCY MODULATION
Amplitude of carrier is varied in accordance to instantaneous amplitude of modulating signal	Frequency of carrier is varied in accordance to instantaneous amplitude of modulating signal
It has carrier LSB and USB and so bandwidth is finite	It has infinite side band and so bandwidth is infinite
Transmitted power is used by carrier and so	All transmitted power is useful and so efficiency

efficiency is low.	is high.
Noise interference is more	Comparatively less
Amplitude of AM wave depends on modulation index and transmitted power is varied accordingly	Amplitude of FM wave is independent on modulation index and hence transmitted power is constant.
Modulation should be less than 1	No limitation
Used in broadcasting medium frequency and high frequency.	Used in broadcasting very high frequency and ultra high frequency.
Adjacent channel interference is more.	Adjacent channel interference is less.
Simple to generate	Complicated process
Amplitude of modulated signal is varied but frequency of carrier remains constant.	Amplitude of modulated signal remains constant but frequency of carrier is varied.
AM has poor fidelity due to narrow band width.	FM has better fidelity due to large band width.
Transmission equipment is simple	Transmission equipment is complex

PHASE MODULATION	**FREQUENCY MODULATION**
Phase of carrier is varied in accordance to instantaneous amplitude of modulating signal	Frequency of carrier is varied in accordance to instantaneous amplitude of modulating signal
It has two side bands and so bandwidth is finite	It has infinite side band and so bandwidth is infinite
Maximum phase duration depends on amplitude of modulating voltage.	Maximum frequency duration depends on amplitude of modulating voltage and frequency.
Modulation index is low.	Modulation index is high.
Modulation index remains constant.	Modulation index is increased if modulation frequency is decreased.

Wideband FM	Narrowband FM
Modulation index is greater than 1	Modulation index is less than 1
Frequency deviation 75 KHz	Frequency deviation 5 KHz
Modulating frequency range is 30 Hz to 15 KHz	Modulating frequency range is 3KHz
Bandwidth is 15 times NBFM	Bandwidth twice of modulating frequency.
Used in broadcasting.	Used in mobile communication.

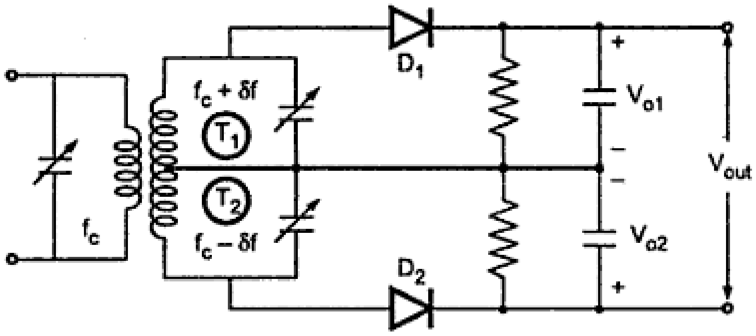
Direct FM	Indirect FM
In direct FM system, frequency of carrier is varied by modulating signal.	In indirect system, FM is obtained by phase modulation of carrier.
Instantaneous frequency is directly proportional by modulating signal.	Instantaneous phase of carrier is directly proportional by modulating signal.
FM generated is not stable	FM generated is stable
Crystal oscillator is not employed	To maintain stability crystal oscillator is employed
AFC is utilized	No need of AFC
The transmitter produces FM signal whose frequency deviation is directly proportional to amplitude of modulating signal	The transmitter produces FM signal whose phase deviation is directly proportional to amplitude modulating signal.
Carrier oscillator frequency is directly deviated by modulating signal.	Frequency of carrier oscillator is not directly varied by modulating signal.

FM demodulators:

- The FM receivers are also super heterodyne receivers. But they have different types of demodulators or detectors. FM receivers have amplitude limiters which are absent in AM receivers.
- The AGC system of FM receiver is different than that that of AM receivers. RF amplifiers, mixers, local oscillators IF amplifiers, audio amplifiers etc, all are present in FM receivers.
- The detection of FM is totally different compared to AM.
- The FM detector should be able to produce the signal whose amplitude is proportional to the deviation in the frequency of FM signal. Thus the job of FM detector is almost similar to frequency to voltage converter. Here we will discuss these types of FM detectors. *Phase discriminator and ratio detector.*

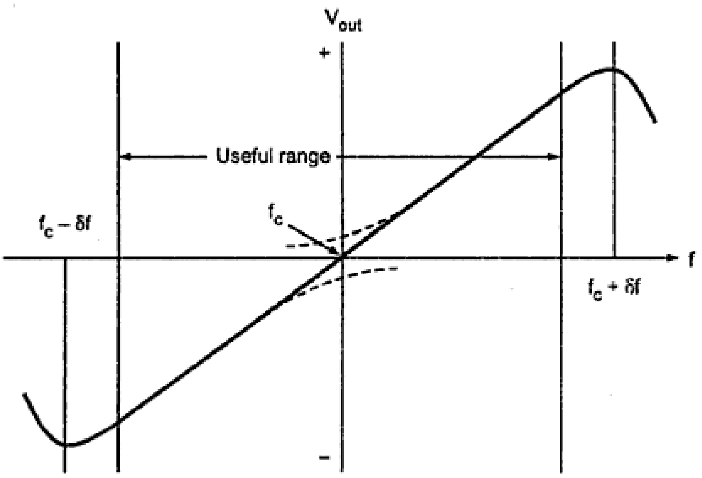
a. Round-Travis detector or balanced slope detector (Frequency Discriminator):

- The circuit shows the balanced slope detector.
- It consists of two identical circuits connected back to back.
- The FM signal is applied to the tuned LC circuit. Two tuned LC circuits are connected in series.
- The inductance of this secondary tuned LC circuit is coupled with the inductance of the primary (or input side) LC circuit. Thus it forms a tuned transformer.
- In fig , the upper tuned circuit is shown as T_c and lower tuned circuit is shown as T_c.
- The input side LC circuit is tuned to f_c, carrier frequency, T1 is tuned to fc+ δf, which represents highest frequency. And lower LC circuit R2 is tuned to fc-δf, which represents the minimum frequency of FM signal .
- The input FM signal is coupled to T1 and T2 180º out of phase.
- The secondary side tuned circuits (T1 and T2) are connected to diodes D1 and D2 with RC loads.
- The total output V_{out} is equal to difference between Vo1 and Vo2, since they subtract.
- The fig shows the characteristic of the balanced slope detector. It shows V_{out} with respect to input frequency.



Balanced Slope Detector

- When the input frequency is equal to fc, both T1 and T2 produce the same voltage.
- Hence V_{o1} and V_{o2} are identical and they subtract each other. Therefore V_{out} is Zero .
- This is shown in the fig , when the input frequency I fc+ δf, the upper circuit T1 produces maximum voltage since it is tuned to this frequency (i.e. fc+ δf) whereas lower circuit T2 is tuned to fc- δf. Which is quite away from fc+ δf.
- Hence T2 produces minimum voltage, hence the output V01 is maximum where V02 is minimum. Therefore Vout- Vo1- Vo2 is maximum positive for fc+ δf.
- When input frequency is fc- δf, the lower circuit T2 produces maximum signal since it is tuned to it.
- But upper circuit T1 produces minimum signal. Hence rectified outputs Vo2 is maximum and Vo1 is minimum, Therefore output Vout=Vo1-Vo2 is maximum negative for fc- δf. This is shown in fig.
- For the other frequencies of input, the output Vout is produced according to the characteristic shown in fig. For example if input frequency tries to increase above fc then Vo1 will be greater than Vo2 and output Vout will be positive, it is desirable that the characteristic shown in fig. should be linear between f_c- δf and f_c+ δf , then only proper detection will take place.
- The linearity of the characteristic depends upon alignment of tuning circuits and coupling characteristics of the tuned coils.



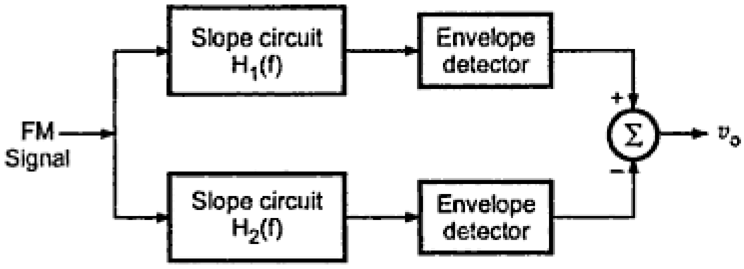
Characteristic of Balanced Slope Detector or S curve

Disadvantages:

1. Amplitude limiting cannot be provided
2. Linearity is not sufficient when compared to slope detector
3. Difficult to align because of three different frequency to which various tuned circuits are to be tuned
4. Tuned circuits is not purely band limited and hence low pass filter of envelope detector introduces distortion

Frequency discriminator:

The fig shows the block diagram of a frequency discriminator which is based on the principle of slope detection. There are two slope circuits. Their transfer functions are as follows:



Frequency discriminator:

$$H_1(f) = H_2(-f) = \left\{ j2\pi a \left(f - f_c + \frac{BW}{2} \right) for f_c - \frac{BW}{2} \le f \le f_c + \frac{BW}{2} \right.$$

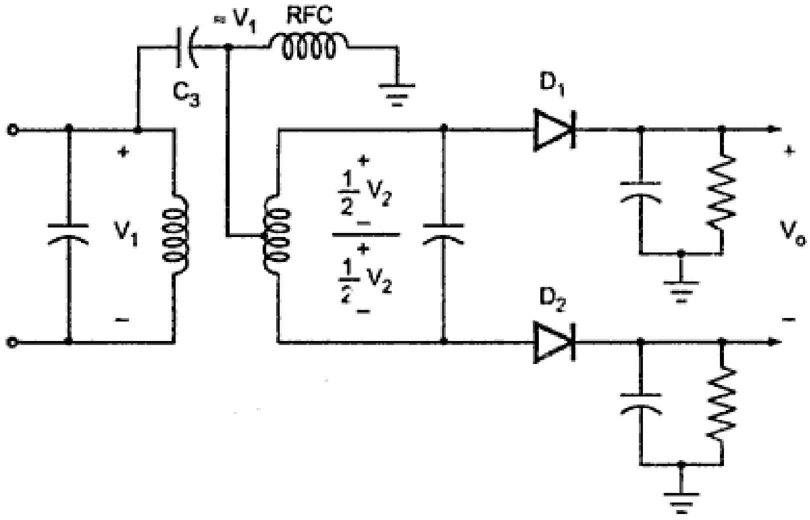
- When the signal is passed through slope circuits , its amplitude as well as frequency varies as per amplitude of modulating signal $e_m(t)$
- This signal is then passed through envelop detectors. It recovers amplitude variations.
- The outputs of two envelope detector (one responds to frequency variation above f_c and other responds to frequency variations below f_c)s are finally subtracted to give final detected output.

b. Foster-seeley discriminator (Phase discriminator or Center-tuned discriminator)- Phase shift is a function of frequency.

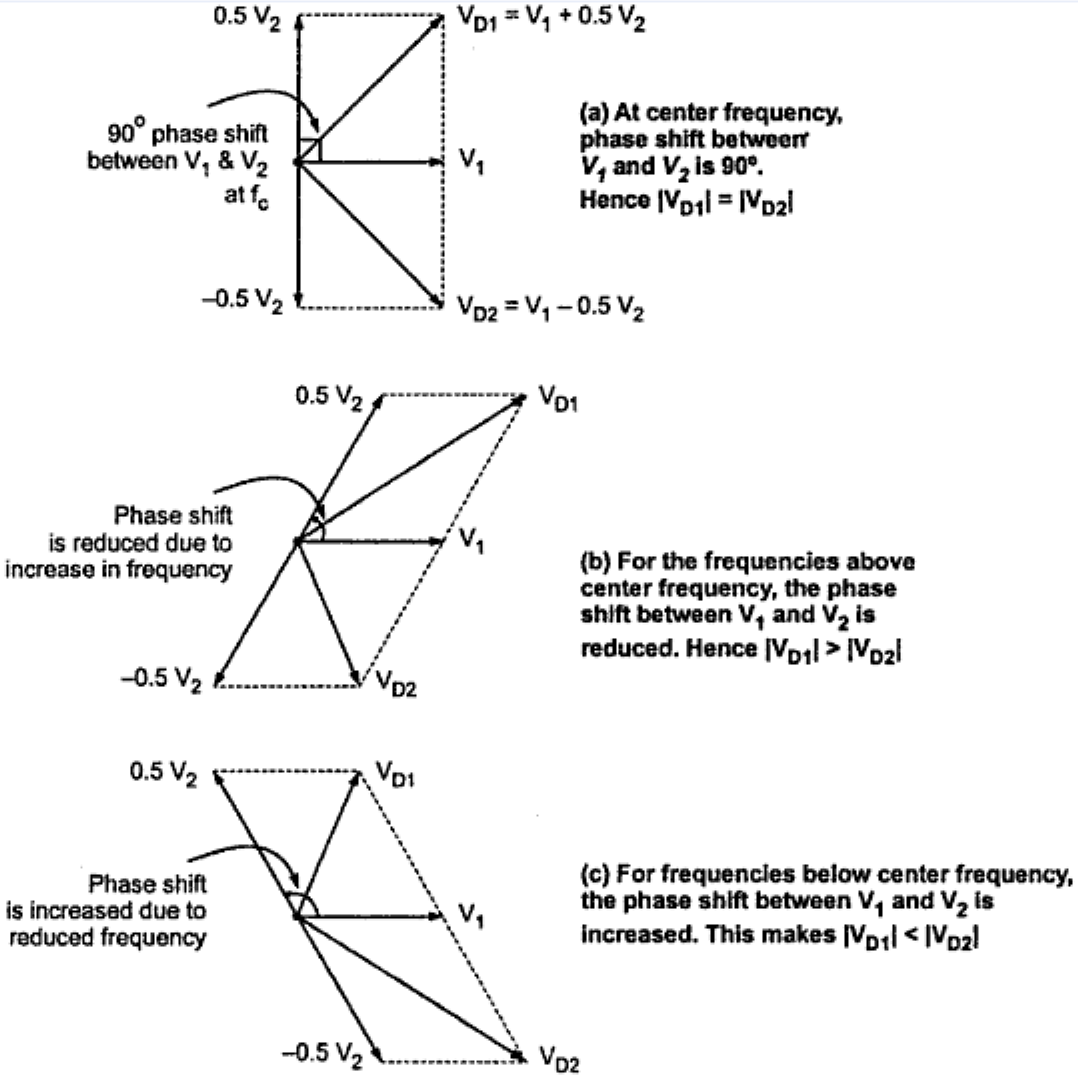
- The phase shift between the primary and secondary voltages of the tuned transformer is a function of frequency.
- It can be shown that the secondary voltage lags primary voltage by 90º at the carrier center frequency.
- This carrier frequency (fc) is the resonance or tuned frequency of the transformer.
- Foster-Seeley discriminator utilizes this principle for FM detection.
- Fig shows the circuit diagram of basic Foster-Seeley discriminator. In the figure observe that primary voltage is coupled through C3 and RFC to the center tap on the secondary.
- The capacitor C3 passes all the frequencies of FM, thus the voltage V1 is generated across RFC. RFC offers high impedance to frequencies of FM .the voltage V1 thus appears across RFC center tap of secondary and ground also.
- The voltage of secondary is V2 and equally divided across upper half and lower half of the secondary coil.
- Fig shows the generator equivalent circuit of Foster-Seeley discriminator. In this fig observe that the voltage across diode D1 is $V_{D1}= V_1 + 0.5V_2$.

The output of upper rectifier is Vo1 and lower rectifier is Vo2. The net output Vo = Vo1 –Vo2. Since $V_{o1} \approx |V_{D1}|$ and $V_{o2} \approx |V_{D2}|$, output $V_o \approx |V_{D1}| - |V_{D2}|$ .thus the net output depends upon the difference between magnitude of V_{D1} and V_{D2}.

- At the frequency, both V_{D1} an V_{D2} will be equal, since V2 will have 90º phase shift with V1.
- Fig shows how V_{D1} and V_{D2} are generated from V_1 and V_2 . In fig vector addition is shown and it shows that $|V_{D1}| = |V_{D2}|$ .Hence the net output of the discriminator will be zero.
- Now consider the situation when input frequency increases above f_c.
- Hence the phase shift between V_1 and V_2 reduces. Therefore $|V_{D1}|$ is greater than $|V_{D2}|$.
- This is shown by vector addition in fig. hence the net output $V_o = |V_{D1}| - |V_{D2}|$ will be positive.
- Thus the increase in frequency increases output voltage. Now consider the situation when frequency reduces below f_c. This makes $|V_{D1}|$ less then $|V_{D2}|$.



- This is shown in the fig. Hence the output $V_o = |V_{D1}| - |V_{D2}|$ will be negative.
- Thus the Foster-Seeley discrimination produces output depending upon the phase shift.
- The linearity of the output depends upon the linearity between frequency and induced phose shift. The characteristic of the Foster- Seeley discriminator (i.e. S-curve) is similar to that show in the fig. with more linearity in the operation.



Advantages:

i. The phase discriminator is much easier to align than the balanced slope detector
ii. There are now only tuned circuits, and both are tuned to same frequency
iii. Linearity is better, because the device relies less upon the frequency response and more on the primary secondary phase relation which is quite linear

The only disadvantage is that is needs a separate amplitude limiting circuit.

Ratio detector:

- In phase discriminator circuits, or foster Seeley discriminator, the changes in the magnitude of the input signal will give rise to amplitude change in output voltage. This makes prior limiting necessary.

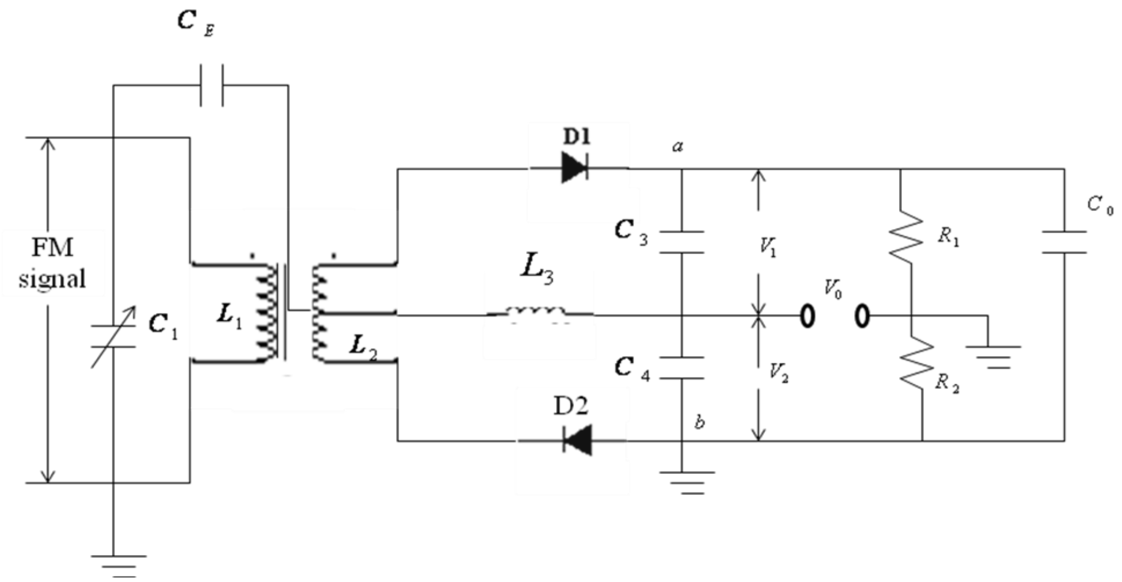
- It is possible to modify the discriminator circuit to provide limiting, so that the amplitude limiter may be dispensed with. A circuit so modified is called a ratio detector.

- Large sum of V_{a0}+ V_{bo} remains constant, although the difference varies because of changes in input frequency. However deviation from this idea does not result in undue distortion in the ratio detector , although some distortion is undoubtedly introduced.

- Therefore any variation in the magnitude of this sum voltage can be considered spurious , accordingly their suppression will lead to a discriminator which is unaffected by the amplitude of the incoming signal

- Form the fig it is seen that three important changes have been made :

    - One of the diodes has been reversed
    - A large capacitor has been placed across the output
    - The output is taken in between resistor and capacitor network



Operation:

- With diode D_2 reversed, O is now positive wrt. B, So V_{ab} is the sum voltage rather than the difference it was in the discriminator. It is now possible to connect large capacitor between a and b to keep this sum voltage constant. once C_o is connected, then Vab is not the output voltage in this case and output is taken across O and O’

- The capacitor charges to a voltage V =(V1+V2). Normally R_1 and R_2 are made equal and the voltage at the junction of R_1 and R_2 measured with respect to the bottom point equal $(V_1+V_2)/2$. This voltage remains fixed because of large value of capacitor C_o and resulting high time constant

  - When signal frequency equals f_c , input to both the diodes will be equal . As C_1 always kept equal to C_2 , V_1 and V_2 are equal . The potential difference across the output terminals equal zero.

  - When signal frequency is greater than f_c ($f_{in}> f_c$) , input for D_1 exceeds input for D_2 . As a result V_1 is increased to $V_1+\Delta V$ and V_2 becomes $V_2-\Delta V$

  - For signal frequencies smaller than $f_c(f_{in}<f_c)$, V_1 (Voltage across D_1) is reduced to $V_1 – \Delta V$ and voltage through D_2 increases to $V_1 + \Delta V$.

  - The potential at the junction of R_1R_2 is still constant and equals $(V_1+V_2)/2$, while the potential at the junction of C_1C_2 measured with respect to bottom becomes $V_2+ \Delta V$. Therefore the output voltage.

$$V_o=(V_1+V_2)/2 \quad - (V_1 - \Delta V) \qquad\qquad \text{assume } V_1 =V_2$$
$$V_o =\Delta V$$

The above cases show that the output voltage follows deviation in the frequency of the signal. Hence this circuit will translate frequency modulated signal into original modulating signal in the same way as Foster Seeley discriminator

Amplitude limiting by ratio detector:

The main advantage of using ratio detector is that there is no need for separate amplitude limiting circuitry. This limiting action is achieved with the help of capacitor C_o connected across R_1 and R_2. It is explained below, how the detector reacts to amplitude changes.

- If the input voltage V_{12} is constant, Co will charge up to potential existing between a and b. this will be a dc voltage and no current will flow through the capacitor i.e. ten put impedance of C_o is infinite. The total impedance for two diodes is therefore the sum of R_1 and R_2 .the output will remain the same

- Consider an increase in input due to noise. In order that the output voltage follows the input, the capacitor C_o must get charged to that level. To charge itself to the increased input level, the capacitor C_0 draws charging current form the input resonant circuit thereby loading this circuit to a great extent , As a result the magnification factor Q , of the circuit is lowered reducing the input signal level.

- If there is a decrease in the input signal level the capacitor must discharge to the input level. Thus there is a discharging current through R_1 and R_2 . the current drawn from the input circuit is reduced, this reduces loading upon the input resonant circuit causing its magnification factor Q to increase . As a result , the input signal level is increased. Thus the ration detector output remains free form amplitude fluctuations in the signal input and converts frequency changes into amplitude changes.

- It is essential to realize that AGC is necessary in a receiver which incorporates a ratio detector.

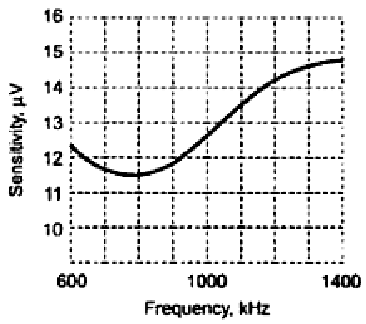
- In TV receivers, this AGC voltage is derived from the video detector, which is an AM detector and a more convenient source of AGC. In FM receivers AGC is obtainable from the ratio detector itself.

RECEIVER CHARACTERISTICS

Sensitivity, noise, selectivity, Image frequency rejection ratio and fidelity are important receiver characteristics.

Sensitivity

- The ability of a receiver to reproduce weak signals and amplify them is a function of the sensitivity. It is often defined in terms of voltage that must be applied to the receiver input terminals to give the standard output power, measured at the output terminals.
- As the gain of receiver is increased, sensitivity is also increased
- It is expressed usually in the microvolts or deciBel that must be applied to the antenna input terminals to give an established level of the output.
- The output may be an ac or dc voltage measured at the detector output or a power measurement (measured in decibels or watts) at the loudspeaker or headphone terminals.

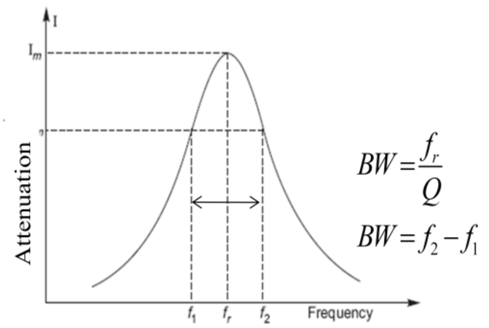


11. a Sensitivity curve of a typical receiver

Noise

- All receivers generate a certain amount of noise, which must be taken into account when measuring sensitivity.
- Receiver noise may originate from the atmosphere (lightning) or from internal components (transistors, tubes).
- Noise is the limiting factor of sensitivity.
- Sensitivity is the value of input carrier voltage (in microvolts) that must be applied from the signal generator to the receiver input to develop a specified output power.

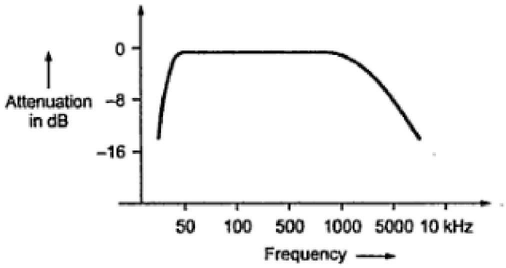
Selectivity



11. b Selectivity curve

- Selectivity is the ability of a receiver to decide on a signal of desired frequency and rejecting all others.
- Selectivity is the degree of distinction made by the receiver between the desired signal and unwanted signals. Better the ability of receiver to reject unwanted signals, better its selectivity.
- The selectivity of receiver is done partially by RF amplifier and mainly by IF amplifier.
- It shows the attenuation that receiver offers to signal at frequencies near which it is tuned.
- The selectivity depends on tuned LC circuits which is in RF and IF amplifiers where resonating frequency (f_r) and (Q) quality factor of circuits.
- Bandwidth should be narrow for better selectivity. Hence Q should be high.
- The degree of selection is determined by the sharpness of resonance to which the frequency-determining circuits have been tuned.

Fidelity

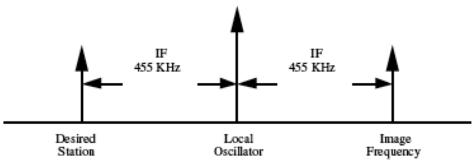


11. c Typical fidelity curve

- Fidelity of a receiver is its ability to accurately reproduce its output, signal that appears at its input.
- Ability of receiver to reproduce all range of modulating frequencies equally is called fidelity.
- Generally a fidelity curve must have a flat response over a wide range of frequencies.
- Fidelity requires wide band of frequencies to be amplified hence it requires more bandwidth of RF and IF stages required but results in poor selectivity.
- AM receivers are not good fidelity receivers, since band width is low
- Good fidelity requires broader band to amplify the outermost frequencies of the sidebands.

Image frequency rejection ratio

- It is the ratio of intermediate-frequency (f_i) signal level produced by desired input frequency to that is produced by the image frequency. The image rejection ratio is usually expressed in dB.
- When the image rejection ratio is measured, the input signal levels of the desired and image frequencies must be equal for the measurement.



$$f_{image} = f_s + 2f_{IF} \text{ if } f_o > f_s$$
$$f_{image} = f_s - 2f_{IF} \text{ if } f_s > f_o$$

11. d Image frequency rejection

- The local oscillator frequency is made higher than signal frequency such that $f_o - f_s = f_i$. Here f_i is intermediate-frequency (IF) that is $f_o = f_s + f_i$.
- The IF stage passes only f_i.If the frequency $f_{si} = f_s + 2f_i$ appears at the input of mixer, then the mixer will produce difference equal to f_i. This is equal to IF.
- The frequency f_{si} is called image frequency and is defined as signal frequency plus twice the IF. This image frequency is converted in the IF range and it is amplified by IF amplifiers. This is the effect of two stations being received simultaneously.
- The image frequency rejection is done by tuned circuits in RF stage. It depends on the selectivity of the RF stage .The image rejection should be done before IF stages.
- It is the measure of ability of preselector to reject the image frequency. It is given by $IFRR = \sqrt{1 + Q^2\rho^2}$. Here Q is the quality factor of preselector and $\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}}$

UNIT-III
NOISE

External and internal noise-Noise figure and noise temperature- AWGN. Noise performance of AM system- Introduction –Destination SNR of a Base band system-Model for linear modulation system-S/N for SSB-SC, DSB-SC and AM systems. Noise performance of frequency modulated system. Pre-emphasis and De-emphasis-Threshold effect in FM.

Noise:

- Noise may be defined as any unwanted introduction of energy tending to interfere with proper reception and reproduction of transmitted signals.
- The word noise means any unwanted sound.
- In both analog and digital electronics, noise or signal noise is an unwanted perturbation to a wanted signal.
- In signal processing or computing it can be considered unwanted data without meaning; that is, data that is not being used to transmit a signal, but is simply produced as an unwanted by product of other activities.
- In pulse communication system noise may produce unwanted pulse or perhaps cancel the unwanted ones. It may cause serious mathematical errors.
- Noise can limit the range of system.
- It affects the sensitivity of receivers, by placing a limit on weakest signal that can be amplified. It even forces to reduce the band width of a system
- Noise can block, distort, change or interfere with the meaning of a message in both human and electronic communication.

Noise Characteristics

- Electrical noise is one of the basic limitations on the performance of a communication system
- Electrical noise may be defined as any unwanted voltage or current that ultimately ends up appearing at the receiver output
- Noise signals at their point of origin are generally very small, for example, at the micro-volt level.

Classification: (Based on the source)**:**

- **External noise**
- **Internal noise**

1. **External Noise:**

   The noise present in a received radio signal, introduced by the transmitting medium, is termed *external noise.*

- External noise sources are either natural (such as solar noise, galactic noise, and atmospheric noise) or man-made (which include industrial noise, electric motors, arc welders, switches, broadcast communication systems, mobile phones, etc.).
- External noise or line noise is the noise injected into the signal when the signal is propagated over the medium that connects the transmitter to the receiver.

- There is nothing the designer can do to control this noise. Instead the designer must take it into account this noise when designing the receiver.
- With careful design, correct circuit board layout, proper choice of components and correct screening of RF/IF stages, the correct use of heat sinks etc., the effects of internal noise on the transmitted and received traffic can be controlled. Once the signal is transmitted down the medium, it comes under the influences of external sources of noise that cannot be controlled.

External noise is **caused** by:

- Static interference (sand storms/dust storms).
- Radiation (sun spot activity, cosmic radiation, manmade radiation).
- Crosstalk (interference from adjacent pairs).
- Mains induction (interference from adjacent power routes).
- Voltaic interference (volcanic activity).

i. Atmospheric noise:

- It is caused by naturally occurring disturbances in the earth's atmosphere; lightning discharges being the most prominent contributors.
- Their location is time-variable, depending on time of the day, season of the year, weather, altitude, and geographical latitude.
- They are more frequently encountered in the equatorial region than at temperate latitudes and above. However, the electromagnetic waves produced by thunderstorms propagate at thousands of kilometers via ionosphere sky wave.
- In the time domain, this noise is characterized by large spikes against a back ground of short random pulses. Its frequency spectrum extends up to 20MHz and the spectral density is proportional to 1/f. Consequently, it mainly affects long-range navigation systems (maritime radio), terrestrial radio broadcasting stations (LW, MW, and SW) and to a considerably lesser extent, FM and TV reception.
- This kind of noise is encountered in rain, snow, hail, and dust storms in the vicinity of the receiving antenna. Its frequency spectrum peaks below 10MHz. It can be reduced by eliminating sharp metallic points from the antenna and its surroundings, and by providing paths to drain static charges that build up on an antenna and in its vicinity during storms.

ii. Extra Terrestrial Noise

- **Galactic Noise:**
  - It is defined as noise at radiofrequencies caused by various celestial objects in outer space or disturbances that originate outside the Earth or its atmosphere.
  - Galactic noise sources can be grouped into two classes: discrete sources and distributed sources.
  - Depending on emission mechanisms, distributed noise sources are thermal or non thermal. So-called thermal noise sources are associated with random encounters of electrons and ions in gas clouds, mostly ionized hydrogen.
  - Non thermal noise sources (also called synchrotron radiation) involve electrons moving in magnetic fields. This is a general galactic phenomenon, encountered even in interstellar space.

- As a general rule, cosmic radio noise covers frequency range from 15MHz up to 100 GHz, with pre dominance between 40MHz and 250 MHz .
- It is observed that it reaches a maximum when the receiving antenna points towards center of Milky Way.

➢ **Solar Noise:**

- The sun is the most powerful noise source, with its temperature of about 6000◦C and its proximity to Earth.
- Its energy is radiated in a continuous mode, and frequency spectrum mainly covers range from several MHz up to several GHz. However, during sunspot and solar-flare activity these values are considerably higher. Also belonging to this category are quasars, which emit copious quantities of energy (usually as powerful radio waves), often in a continuous mode, although the signals arriving at Earth are perceptible only by sensitive radio telescopes.

iii. Industrial noise:

Noise is generated in two distinct areas:

- In the ionized gas column, this presents a small but fluctuating resistance when the light is on.
- When the temperature changes and abruptly breaks the current flowing through an inductor. A voltage spike occurs, which is used to trigger the discharge; however, this spike is also a source of interference for nearby systems.

2. Internal Noise:

The noise already present at the receiving antenna (external noise) has another component added to it before it reaches the output. This is the noise introduced by the receiver itself. It is termed internal noise.

- The receiver's major noise contribution occurs in its very first stage of amplification.
- It is there that the desired signal is at its lowest level, and noise injected at that point will be at its largest value in proportion to the intelligence signal.

Types of noise generated by electronic circuits:

Thermal Noise:

➢ This type of noise arises due to random motion of free electrons in conducting medium such as resistor.

➢ Each free electron inside a resistor is in motion due to its thermal energy.

➢ The path of electron motion is random and zigzag due to collisions with the lattice structure.

➢ The net effect of the motion of all electrons constitutes an electric current flowing through resistor.

➢ It causes rate of arrival of electrons at either end of a resistor to vary randomly, and thereby varies resistor's potential difference. That is, direction of current flow is random and has a zero mean value.

➢ Resistors and resistance within all electronic devices are constantly producing noise voltage, V_n.

$$
\begin{aligned}
P_n &= \frac{V^2}{R_L} \\
&= \frac{V^2}{R} = \frac{(V_n/2)^2}{R} \\
&= \frac{V_n{}^2}{4R} \\
V_n{}^2 &= 4RP_n \\
&= 4RKT\partial f \\
V_n &= \sqrt{4RKT\partial f}
\end{aligned}
$$

Since, it is dependent on temperature, it is referred to as thermal noise.

- It has been determined through measurement that the normalized two-sided power spectral density, psd. of thermal noise in electrical communication systems can be assumed to be

  k -Boltzmann's constant (1.38×10^{-23} J/K)

  R -is the resistance generating the noise

  T -Resistor temperature in degrees Kelvin (K)

. ∂f Bandwidth of a system

Shot Noise

- ✓ The major contributor of noise in semiconductor devices such as diodes and transistors is the shot noise.
- ✓ It is due to random diffusion of minority carriers and random generation and recombination of hole-electron pairs.
- ✓ In semiconductor devices although current being continuous, it is actually a discrete phenomenon.
- ✓ In fact, current occurs in discrete pulses each time an electron moves across an observation point.

- ✓ Short noise is the variation of current around the average value.
- ✓ The psd of shot noise is approximately flat in the range of frequencies Shot noise and thermal noise
- ✓ are additive.

- ✓ Noise that has a flat power spectral density over a wide range of frequencies (theoretically infinite) is called white noise, in analogy to white light.

Transient Noise Pulses:

- The input current to a transistor flows from the emitter to the base. After crossing the barrier it divides between the base terminal and the collector terminal.
- This base current is also subject to random fluctuations and contributes to the overall noise produced by the transistor. The noise is produced at the base junction.

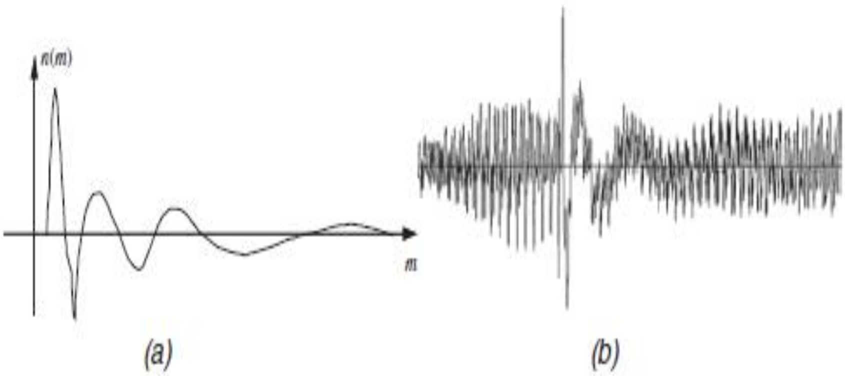


Figure (a) A scratch pulse and music from a gramophone record. (b) Averaged profile of a gramophone record scratch pulse.

- Transient noise pulses, observed in most communication systems, are bursts of noise, or long clicks caused by interference or damage to signals during storage or transmission.
- Transient noise pulses often consist of a relatively short sharp initial pulse followed by decaying low-frequency oscillations.
- The initial pulse is usually due to some external or internal impulsive interference, whereas the oscillations are often due to the resonance of the communication channel excited by the initial pulse, and may be considered as the response of the channel to the initial pulse.
- The noise pulse is shaped by the channel characteristics, and may be considered as the channel pulse response.
- Scratch noise pulses are acoustic manifestations of the response of the stylus and the associated electro-mechanical playback system to a sharp physical discontinuity on the recording medium.
- A typical scratch pulse waveform often exhibits two distinct regions:
    (1) Initial high-amplitude pulse response of the playback system to the physical discontinuity on the record medium
    (2) Decaying oscillations that cause additive distortion. The initial pulse is relatively short and has duration on the order of 1–5 ms, whereas the oscillatory tail has a longer duration and may last up to 50 ms or more.

MISCELLANEOUS NOISE

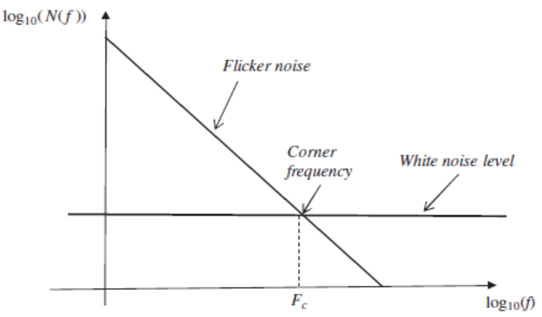
Flicker noise

- Flicker noise is an electronic device noise that occurs due to the random fluctuations of electron flow that accompanies direct current in electronic devices such as transistors and vacuum tubes.
- It results from a variety of effects, such as crystal surface defects in semiconductors, impurities in a conductive channel, generation and recombination noise in a transistor due to base current etc.
- Flicker noise is more prominent in field effect transistors (FETs) and bulky carbon resistors.

$$F_c \propto \frac{I_e T_j}{f}$$

- Where I_e is value of emitter current , f is frequency and T_j is junction temperature.

- A characteristic parameter of the flicker noise is the corner frequency F_c defined as frequency at which magnitude spectrum of flicker noise is equal to, and crosses, that of white noise; for frequencies above F_c flicker noise goes below white noise.



Spectrum of flicker noise

- In electronic devices, flicker noise is usually a low-frequency phenomenon; as the higher frequencies are drown by white noise from other sources.
- However, in oscillators, the low-frequency noise can be modulated and shifted to higher frequencies by the oscillator carrier. Since flicker noise is related to direct current, it may be kept low if the level of direct current is kept low, for example in resistors at low current levels thermal noise will be the predominant effect.

Resistance

Fluctuations in the conductivity of the semiconductor material provide a noise source which is inversely proportional to frequency. This noise is known as current or excess noise. It is usually negligible above 10 kHz and for some transistors above 1 kHz.

Noise in Mixers

Mixers are much noisier than amplifiers using identical devices. The high values if noise in mixers is caused by two effect.

- Conversion transconductance of a mixer is much lower than transconductance of amplifiers
- If image frequency rejection is inadequate, noise associated with image frequency will also be accepted.

Quantifying the noise

- The presence of noise degrades the performance of analog and digital communications. The extent to which noise affects performance of communication systems is measured by output signal to noise power ratio or SNR (for analog communication systems) and probability of error (for digital communication systems).

- The signal quality at input of receiver is characterized by input signal to noise ratio. Because of noise sources within receiver, which is introduced during the filtering and amplification processes, the SNR at the output of the receiver will be lower than at the input of the receiver.

- This degradation in the signal quality is characterized in terms of noise equivalent bandwidth, effective noise temperature, and noise figure.

SIGNAL TO NOISE RATIO

- Although digital transmission is used extensively today, many of the media used require an analogous signal.
- This means that the digital information must be changed into an analogue format. A typical example is radio and microwave transmission.
- To a large extent frequency modulation is used on radio and micro-wave systems.
- The criterion used to determine the performance of any analogue system is the signal-to-noise ratio (SNR).
- Both signal and noise power must be measured at the same and equivalent points in a system, and within same system bandwidth.
- If the signal and noise are measured across the same impedance, then the SNR can be obtained by calculating the square of amplitude ratio: This is determined by means of the following equation:

$$\frac{S}{N} = \frac{V_S^2 / R}{V_n^2 / R} = \left(V_S^2 / V_n^2 \right)$$

- Signal-to-noise ratio (SNR or S/N) is defined as the ratio of a signal power to the noise power corrupting the signal.
- A ratio higher than 1:1 indicates more signal than noise. In non-technical terms, signal-to-noise ratio compares the level of a desired signal (such as music) to the level of background noise

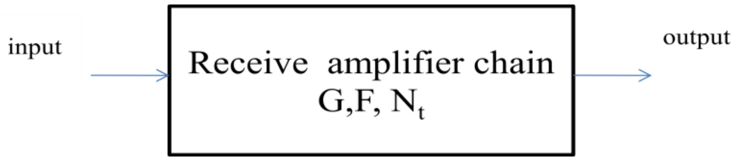
  Very often the SNR is expressed in dB

$$SNR_{dB} = 10\log_{10}\left(\frac{S}{N}\right) dB$$

- When using digital storage the number of bits of each value determines the maximum signal-to-noise ratio. In this case the noise is the error signal caused by the quantization of the signal, taking place in the analog-to-digital conversion.
- The noise level is non-linear and signal-dependent; different calculations exist for different signal models. The noise is modeled as an analog error signal being summed with the signal before quantization ("additive noise").

- The modulation error ratio (MER) is a measure of the SNR in a digitally modulated signal. Like SNR, MER can be expressed in dB.

NOISE FIGURE OR NOISE FACTOR



where G = gain

F = noise figure

N_t = internally generated noise power

- A measure which is used extensively on amplifier chain at front end of a receiver is noise figure or noise factor. The noise figure is a measure of the noisiness of input stage.
- In other words it determines the effect that the internal noise, produced by that stage, has on the received signal. As will be shown below the more internal noise produced by a stage the worse the output signal-to-noise ratio becomes.
- Hence the internal noise must be controlled, especially for the input stage of the receiver where the received signal strength is already weak and a large amount of noise accompanies the signal due to the transmission medium.

The noise factor is a measure of the degradation of the output signal, and is determined by

$$F=SNR_{in} /SNR_{out}$$

Here equations are in ratio not in dB now $SNR_{in=}S_i/N_i$ and $SNR_{out}= S_o/N_o$

$$S_0=GS_i \text{ and } N_o =G(N_i+N_{ai})$$

$$F=(S_i/N_i) / (G\, S_i /(N_i+N_{ai}))$$

where S_i= signal power at input

N_i= noise power at input

G= amplifier gain as ratio

N_{ai}=amplifier noise referred to input of amplifier

All the internal noise generated in the amplifier is assumed to be generated by an external source at the input to the amplifier and the amplifier then becomes a noise-free device. The externally generated noise is now referred to as N_{ai}. Therefore

$$F= (N_i + N_{ai})/ N_i$$

This is then expressed as $F=1+ (N_{ai}/ N_i)$

It can be seen that if the amplifier is noise free then $N_{ai} = 0$ and the noise figure F would be equal to unity. This means that $SNR_{in} = SNR_{out}$. An alternative derivation for an expression of Nr in terms of the noise figure is given $F= (S_i/N_i) / (G\, S_i /(N_i+N_{ai}))$

But $G(N_i+N_{ai}) = N_o$

$F= (S_i/N_i) \times (N_o /GS_i) = N_o / GN_i$

$N_o= GN_i F$

We have $(GN_i+GN_{ai}) = N_o$ but $GN_{ai} = N_r$

Thus $GN_i+N_r = N_o$

$F= (S_i/N_i) \times ((GN_i+GN_{ai}) /GS_i))$

$GFN_i = GN_i +N_r$

$N_r =GFN_i - GN_i$

$= G N_i (F– 1)$

- Essentially the measurement assesses the amount of noise each part of the system or the system as a whole introduces.
- This could be the radio receiver or an RF amplifier for example. If the system were perfect then no noise would be added to signal when it passed through the system and the signal to noise ratio would be same at output as at the input.
- As we know this is not the case and some noise is always added. This means that SNR at the output is worse than the SNR ratio at the input. In fact the noise figure is simply the comparison of the SNR at the input and the output of the circuit.
- A figure known as the noise factor can be derived simply by taking the SNR at the input and dividing it by the SNR at the output. As the SNR at the output will always be worse, i.e. lower, this means that the noise factor is always greater than one.

NOISE TEMPERATURE

The concept of noise figure is not always the most convenient measure of noise in dealing UHF and microwave low noise antennas, receivers or devices. This may be illustrated by noise power equation

$$P_t = kT\delta f$$
$$P_t = P_1 + P_2$$
$$kT_t\delta f = kT_1\delta f + kT_2\delta f$$
$$T_t = T_1 + T_2$$

Where P_1 and P_2 are individual noise power and P_t is their sum and T_1 and T_2 are individual noise power and T_t is total noise temperature.

- Another advantage of the use of noise temperature for low level is that it shows greater variations for any given noise level change than does the noise figure.
- Similarly, T_{eq} the equivalent noise temperature may also be utilized in defining the equivalent noise temperature of a receiver or amplifier. It is assumed that $R'_{eq} = R_a$. If this

is to lead to correct value of noise output power, then obviously R'_{eq} must be at a temperature other than the standard one at which all the components are assumed to be. By using the equation

$$F = 1 + \frac{R'_{eq}}{R_a}$$

$$= 1 + \frac{kT_{eq}\delta f}{kT_0\delta f} \cdot \frac{R'_{eq}}{R_a}$$

$$= 1 + \frac{T_{eq}}{T_0}$$

where $R'_{eq} = R_a$, R'_{eq} is the equivalent noise resistance and R_a is the antenna resistance T_o is the initial temperature in Kelvin and T_{eq} equivalent noise temperature of amplifier or receiver whose noise figure is F. It is a ratio and not expressed in decibel. T_{eq} may be influenced by actual ambient temperature of amplifier or receiver finally the equation becomes

$$T_0 F = T_0 + T_{eq}$$

$$T_{eq} = T_0(F + 1)$$

From the above equation equivalent noise temperature can be calculated by knowing noise figure.

UNIT-IV
RADAR

Basic principles of RADAR system- Range equation- Pulse radar system- Basic Pulsed radar system-antennas and scanning-display methods-pulsed radar system- MIT radar-radar beacons. CW Doppler Radar- FM CW Radar –Phased array radars-Planar array radar.

Introduction

- Radar is an electromagnetic system for the detection and location of objects.
- It operates by transmitting a particular type of waveform, a pulse-modulated sine wave for example, and detects the nature of the echo signal.
- Radar is used to extend the capability of one's senses for observing the environment, especially the sense of vision.
- The value of radar lies not in being a substitute for the eye, but in doing what the eye cannot do-Radar cannot resolve detail as well as the eye, nor is it capable of recognizing the "colour" of objects to the degree of sophistication of which the eye is capable.
- However, radar can be designed to see through those conditions impervious to normal human vision, such as darkness, haze, fog, rain, and snow.
- In addition, radar has the advantage of being able to measure the distance or range to the object. This is probably its most important attribute.
- An elementary form of radar consists of a transmitting antenna emitting electromagnetic radiation generated by an oscillator of some sort, a receiving antenna, and an energy-detecting device, or receiver.
- A portion of the transmitted signal is intercepted by a reflecting object (target) and is reradiated in all directions. It is the energy reradiated in the back direction that is of prime interest to the radar.
- The receiving antenna collects the returned energy and delivers it to a receiver, where it is processed to detect the presence of the target and to extract its location and relative velocity.
- The distance to the target is determined by measuring the time taken for the radar signal to travel to the target and back.
- The direction, or angular position, of the target may be determined from the direction of arrival of the reflected wave-front.
- The usual method of measuring the direction of arrival is with narrow antenna beams. If relative motion exists between target and radar, the shift in the carrier frequency of the reflected wave (Doppler Effect) is a measure of the target's relative (radial) velocity and may be used to distinguish moving targets from stationary objects.
- In radars which continuously track the movement of a target, a continuous indication of the rate of change of target position is also available.
- The name radar reflects the emphasis placed by the early experimenters on a device to detect the presence of a target and measure its range. Radar is a contraction of the words radio detection and ranging. It was first developed as a detection device to warn of the approach of hostile aircraft and for directing antiaircraft weapons.
- Although well-designed modern radar can usually extract more information from the target signal than merely range, the measurement of range is still one of radar's most important functions. There seem to be no other competitive techniques which can measure range as well or as rapidly as can radar.
- The most common radar waveform is a train of narrow, rectangular-shape pulses modulating a sinewave carrier. The distance, or range, to the target is determined by

measuring the time T_R taken by the pulse to travel to the target and return. Since electromagnetic energy propagates at the speed of light c = 3 x 10^8 m/s, the range R is

$$R = \frac{cT_R}{2}$$

- The factor 2 appears in the denominator because of the two-way propagation of radar. With the range in kilometers or nautical miles, and TR in microseconds, Eq. (1.1) becomes

$$R(km) = 0.15T_R(\mu s) or R(nmi) = 0.081T_R(\mu s)$$

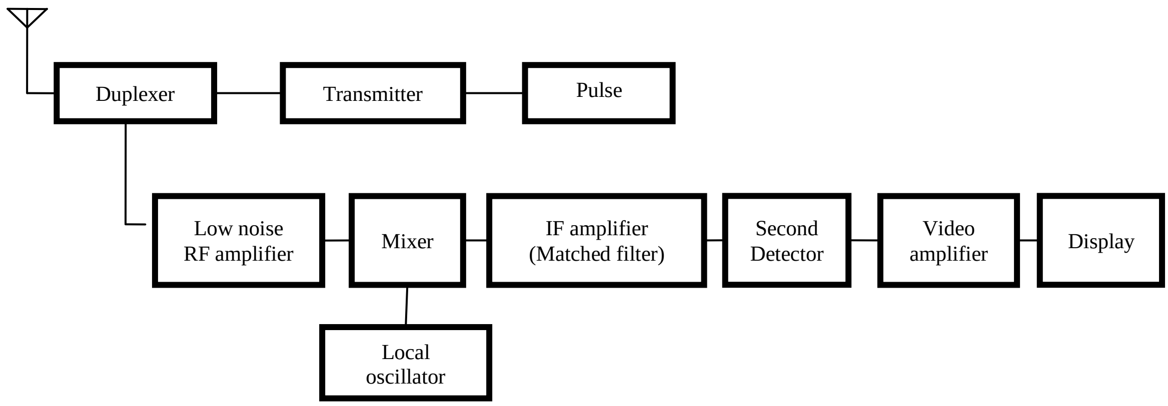
- Each microsecond of round-trip travel time corresponds to a distance of 0.081 nautical mile, 0.093 statute mile, 150 meters, 164 yards, or 492 feet.
- Once the transmitted pulse is emitted by the radar, a sufficient length of time must elapse to allow any echo signals to return and be detected before the next pulse may be transmitted.
- Therefore the rate at which the pulses may be transmitted is determined by the longest range at which targets are expected. If the pulse repetition frequency is too high, echo signals from some targets might arrive after the transmission of the next pulse, and ambiguities in measuring range might result.
- Echoes that arrive after the transmission of the next pulse are called second-time-around (or multiple-time-around) echoes. Such an echo would appear to be at a much shorter range than the actual and could be misleading if it were not known to be a second-time-around echo.
- The range beyond which targets appear as second-time-around echoes is called the maximum unambiguous range and is

$$R_{unamb} = \frac{c}{2f_p}$$

where f, = pulse repetition frequency, in Hz. A plot of the maximum unambiguous range as a function of pulse repetition frequency is shown in Fig. 1.1.

- Although the typical radar transmits a simple pulse-modulated waveform, there are a number of other suitable modulations that might be used. The pulse carrier might be frequency- or phase-modulated to permit the echo signals to be compressed in time after reception.
- This achieves the benefits of high range-resolution without the need to resort to a short pulse. The technique of using a long, modulated pulse to obtain the resolution of a short pulse, but with the energy of a long pulse, is known as pulse compression.
- Continuous waveforms (CW) also can be used by taking advantage of the doppler frequency shift to separate the received echo from the transmitted signal and the echoes from stationary clutter.
- Unmodulated CW waveforms do not measure range, but a range measurement can be made by applying either frequency- or phase-modulation.

Radar Block Diagram and Operation



Block Diagram of Pulse radar

- ✓ The operation of typical pulse radar may be described with the aid of the block diagram shown in Figure 1.
- ✓ The transmitter may be an oscillator, such as a magnetron, that is "pulsed" (turned on and off) by the modulator to generate a repetitive train of pulses.
- ✓ The magnetron has probably been the most widely used of the various microwave generators for radar.
- ✓ Typical radar for the detection of aircraft at ranges of 100 or 200 nmi might employ a peak power of the order of a megawatt, an average power of several kilowatts, a pulse width of several microseconds, and a pulse repetition frequency of several hundred pulses per second.
- ✓ The waveform generated by the transmitter travels via a transmission line to the antenna, where it is radiated into space. A single antenna is generally used for both transmitting and receiving.
- ✓ The receiver must be protected from damage caused by the high power of the transmitter. This is the function of the duplexer. The duplexer also serves to channel the returned echo signals to the receiver and not to the transmitter. The duplexer might consist of two gas-discharge devices, one known as a TR (transmit-receive) and the other an ATR (anti-transmnit-receive).
- ✓ The TR protects the receiver during transmission and the ATR directs the echo signal to the receiver during reception. Solid-state ferrite circulators and receiver protectors with gas-plasma TR devices and/or diode limiters are also employed as duplexers.
- ✓ The receiver is usually of the superheterodyne type. The first stage might be a low-noise RF amplifier, such as a parametric amplifier or a low-noise transistor.
- ✓ However, it is not always desirable to employ a low-noise first stage in radar. The receiver input can simply be the mixer stage, especially in military radars that must operate in a noisy environment.
- ✓ Although a receiver with a low-noise front-end will be more sensitive, the mixer input can have greater dynamic range, less susceptibility to overload, and less vulnerability to electronic interference.
- ✓ The mixer and local oscillator (LO) convert the RF signal to an intermediate frequency (IF). A " typical " IF amplifier for an air-surveillance radar might have a center frequency of 30 or 60 MHz and a bandwidth of the order of one megahertz.
- ✓ The IF amplifier should be designed as a matched filter; i.e., its frequency-response function H(f) should maximize the peak-signal-to-mean-noise-power ratio at the output. This occurs when the magnitude of the frequency-response function H(f) is equal to the magnitude of the echo signal spectrum S(f), and the phase spectrum of the matched filter is the negative of the phase spectrum of the echo signal.

- In a radar whose signal waveform approximates a rectangular pulse, the conventional IF filter bandpass characteristic approximates a matched filter when the product of the IF bandwidth B and the pulse width r is of the order of unity, that is, Bt 1.
- After maximizing the signal-to-noise ratio in the IF amplifier, the pulse modulation is extracted by the second detector and amplified by the video amplifier to a level where it can be properly displayed, usually on a cathode-ray tube (CRT).
- Timing signals are also supplied to the indicator to provide the range zero. Angle information is obtained from the pointing direction of the antenna.
- The most common form of cathode-ray tube display is the plan position indicator, or PPI (Fig. 1.3a), which maps in polar coordinates the location of the target in azimuth and range.
- This is an intensity-modulated display in which the amplitude of the receiver output modulates the electron-beam intensity (z axis) as the electron beam is made to sweep outward from the center of the tube.
- The beam rotates in angle in response to the antenna position. A B-scope display is similar to the PPI except that it utilizes rectangular, rather than polar, coordinates to display range vs. angle. Both the B-scope and the PPI, being intensity modulated, have limited dynamic range.
- Another form of display is the A-scope, shown in Fig. 1.3b, which plots target .amplitude (y axis) vs. range (x axis), for some fixed direction. This is a deflection-modulated display. It is more suited for tracking-radar application than for surveillance radar properly displayed, usually on a cathode-ray tube (CRT).
- Timing signals are also supplied to the indicator to provide the range zero. Angle information is obtained from the pointing direction of the antenna.
- The most common form of cathode-ray tube display is the plan position indicator, or PPI (Fig. 1.3a), which maps in polar coordinates the location of the target in azimuth and range. This is an intensity-modulated display in which the amplitude of the receiver output modulates the electron-beam intensity (z axis) as the electron beam is made to sweep outward from the center of the tube.
- The beam rotates in angle in response to the antenna position. A B-scope display is similar to the PPI except that it utilizes rectangular, rather than polar, coordinates to display range vs. angle.
- Both the B-scope and the PPI, being intensity modulated, have limited dynamic range. Another form of display is the A-scope, shown in Fig. 1.3b, which plots target amplitude (y axis) vs. range (x axis), for some fixed direction. This is a deflection modulated display. It is more suited for tracking-radar application than for surveillance radar.

RADAR RANGE EQUATION

One of the simpler equations of radar theory is the radar range equation. Although it is one of the simpler equations, ironically, it is an equation that few radar analysts understand and many radar analysts misuse.

The problem lies not with the equation itself but with the various terms that make-up the equation. It is my belief that if one really understands the radar range equation one will have a very solid foundation in the fundamentals of radar theory.

Because of the difficulties associated with using and understanding the radar range equation we will devote considerable class time to it and to the things it impacts, like detection theory, matched filters and the ambiguity function.

BASIC RADAR RANGE EQUATION

One form of the basic radar range equation is

$$SNR = \frac{P_S}{P_N} = \frac{P_T G_T G_R \lambda^2 \sigma}{(4\pi)^3 R^4 k T_0 B F_n L}$$

- ✓ Where,
  - SNR is termed the signal-to-noise ratio and has the units of watts/watt, or w/w.
  - P_N is the signal power at some point in the radar receiver – usually at the output of the matched filter or the signal processor. It has the units of watts (w).
  - P_T is the noise power at the same point that is specified and has the units of watts.
  - P_T is termed the peak transmit power and is the average power when the radar is transmitting a signal. It can be specified at the output of the transmitter or at some other point like the output of the antenna feed. G_T is the directive gain of the transmit antenna and has the units of w/w.
  - G_R is the directive gain of the receive antenna and has the units of w/w.
  - λ is the radar wavelength and has the units of meters (m).
  - σ is the target radar cross-section or RCS and has the units of square meters.
  - R is the range from the radar to the target and has the units of meters.

Classification of RADARS:

The various radar systems can be classified as below:

1. Pulse radar
2. Moving target indication (MTI) radar
3. Beacons radar
4. CW radars
5. Tracking radar
6. Laser radar

PULSE RADAR

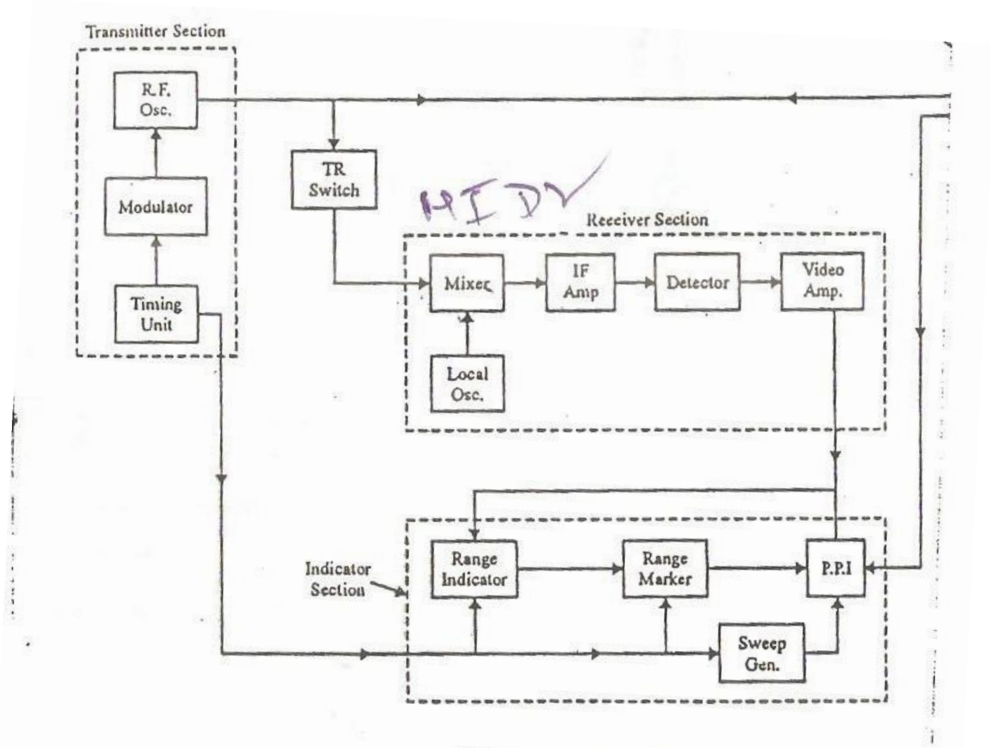


Figure 16.6 Pulse Radar System

1. Transmitter:

- In this section, a high power oscillator (usually a magnetron) generates RF oscillators.
- The voltage pulses are formed in a modulator which is applied in its plate cathode circuit.
- The pulse operation is controlled by timer unit,which is basically a multivibrator.
- The trigger pulses obtained from the timer unit are also applied to the indicator section for synchronization .
- The RF energy of the oscillator is transmitted through an antenna which is specially shapped and designed parabolic reflector .
- The TR switch which periodically makes contact with transmitter and receiver employs a gas discharge tube .
- The high power pulses from oscillator break down the gap in the tube and short circuits the line to the receiver.

2. Reciever:

- The radar reciever is a superheterodyne receiver.The intermediate frequency (IF) is kept usually 30 to 60 MHz. A diode is used as a dectector.
- The video amplifier raises the signal amplitude to the required level to be applied in the indicators.

3. Indicators:

- The indicators panel of radar is provied with the various measuring instruments and indicators e.g. Pulse position indicator (PPI).
- It is provided also with a sweep generator for range measurements.

DESCRIPTION OF MAIN COMPONENTS

(i) T.R Switch (duplexer):

- A TR switch or duplexer is circuit which enables the same antenna to be used for both transmission and reception without any interference.
- In radar, generally Branch type Duplexer is used.It has two switches (TR and anti TR) which alternately connect transmitter and receiver with the antenna.
- Cold cathode tubes are generally used for these switches.

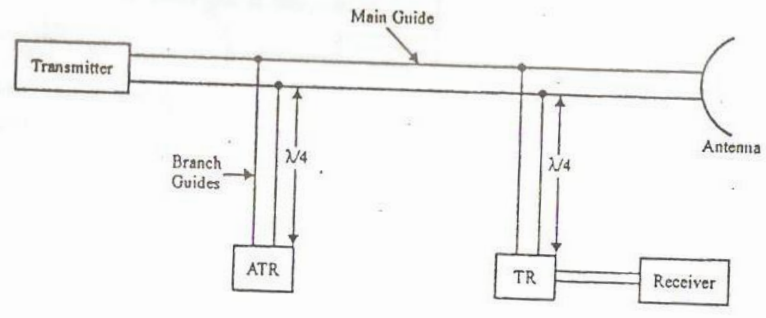


Figure 16.7 Duplexer

Operation:

- A main guide connects the transmitter to the antenna. With this main guide, TR and ATR (anti TR) switches are connected through the branch guides of quarter wave lengths.
- When the transmitter produces RF pulses, both switches get short circuited. The short circuiting of TR switch prevents the RF power entering into transmitter .
- The guide leading through TR switch to the receiver becomes continuous, thus the echo reflected back from the antenna can go to the receiver.

(ii)Modulator:

- Modulator is a component of radar transmitter. The function of modulator is to switch on and off the output tube as and when required .
- There are two types of modulators in use:

  (a) Line pulsing switch modulators
  (b) Active switch modulator

Line pulsing modulator:

- In this anode of the output tube is modulated directly by system that generates and provides large pulses of supply voltage. This is achieved by slowly charging and discharging a transmission line.
- But practically a Pulse Forming Net work (PFNW) replaces the line. As shown in figure the PFNW behaves as a transmission line for frequencies below $f=1/\sqrt{LC}$ where L and C are inductance and capacitance respectively.For switch , a thyratron or a SCR may employed.

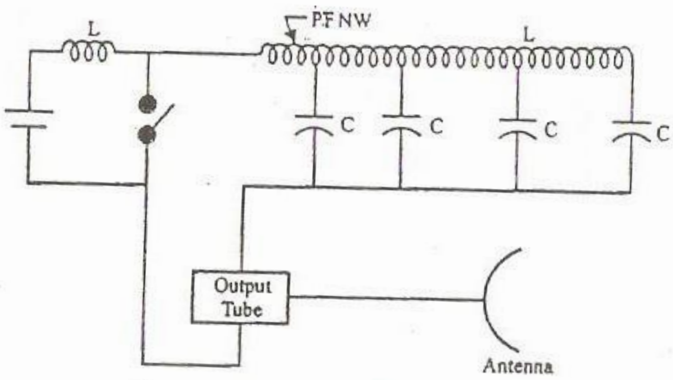


Figure 16.8 Line pulsating Modulator

(b) Active switch modulator:

- The line pulsing modulator cannot give variable pulse length,therefore active switch modulator may be used.
- In this,pulses are generated at low power level and later on amplified by an amplifier , but this modulator is less efficient , more complex and bulky.

(iii) Antenna:

- The majority of radars use dipole or horn type parabolic reflector as antenna .
- The beam width in vertical direction is worst than in the horizontal direction , but this is immaterial in ground to ground or even air to ground radars.
- It has the advantage of having small size and weight , reduced wind load,morever they need small motors.

The various shapes used for radar antenna are shown in the figure:

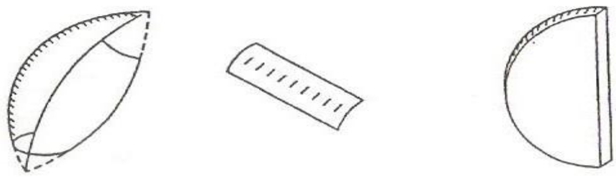


Figure 16.9 Radar Antennas

MTI RADAR

- In radar system , only moving targets are displayed on the screen ,and the echo obtained from the stationary targets is not displayed at all. The MTI radar employs Doppler's effect explained below.

DOPPLER EFFECT

- "The apparent frequency of an electromagnetic (i.e.,light,sound) wave depends on the relative motion of source and the observer."
- If source and the observer are moving away from each other the apparent frequency will decrease and if they are moving towards each other , the apparent frequency will increase.
- In case of radar , involving a moving target ,the signal undergoes a change to the doppler effect. The target acts as a source of the reflected waves. Now we have a moving source and a stationary observer . If the target is moving towards a radar , the over all effect is doubled.

- The doppler effect is observed only in radial motions and not in the tangential motions. If target is rotating , the radar can distinguish its leading edge from its trailing edge . The rotation of the plane Venus has been measured by this effect as due to dense cloud cover over this plant, this cannot be dont by telescope.

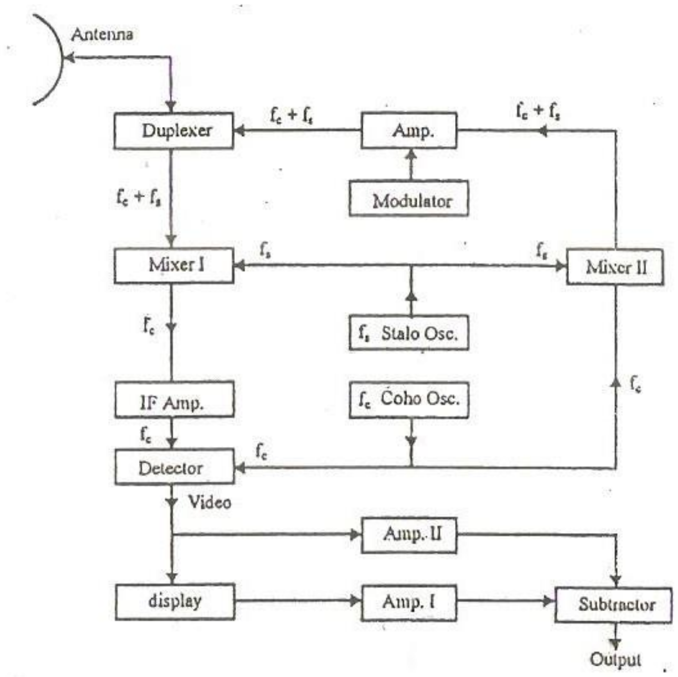


Figure 16.10 MTI Radar System

- Basically, a MTI radar system compares a set of received echoes with those received during previous sweep.
- The echoes whose phase remains constant are cancelled out as they are obtained from station targets but the echoes obtained from the moving targets show a phase difference due to Doppler Effect hence not cancelled. It also helps in the detection of moving targets whose echoes are very small to long distance than those of nearby stationary targets.
- The transmitted frequency is the sum of the outputs of oscillator .The first oscillator is called Stalo (stable) and other a coho (coherent). The mixers I and II are identical and both use the stalo thus relations of their inputs are preserved in their outputs also. The output of the I.F. amplifier and signal of coho oscillator is fed to a detector circuit.
- The coho oscillator is used to generate RF signal as well as the signal for the detector. So the output of the detector is phase sensitive , an output will be obtatined for all fixed and moving target.
- The phase difference between transmitted and received signals will be constant for fixed targets it will be changing for moving targets as per the Doppler effect.

BLIND SPEEDS IN MTI RADARS

- In an MTI radar, the phase difference between transmitted pulse and the echo remains constant from pulse to pulse in case of stationary targets.
- The phase difference, however varies in case of moving targets . That is , why MTI radar can distinguish between stationary and moving targets.
- If the targets happen happen to be in moving with a velocity such that its velocity component along the radar axis results in phase difference of 2π radians or its integral multiple between successive of pulses to radar , the targets appear as stationary . These target velocities are said to be Blind speed. If the target velocity is such that it moves a distance equal to $n.\lambda/2$, where n is the integer and λ is the transmitted wave length along the radar axis , the phase difference will be $n.2\pi$ radians,

- blind speed is expressed by

$$V_b = \frac{n\lambda/2}{1/fr} = \frac{n\lambda \text{fr}}{2}$$

- When n=1 the lowest blind speed = λfr/2.Where,n is the integer λ= wave length transmitted and fr = pulse repetition frequency

BEACON RADAR (RESPONDERS)

- The beacon is a small radar set , consisting of receiver , transmitter and an all directional antenna when another radar transmits a coded signal to a beacon is interrogated, the beacon responds by sending back its own coded pulses .
- A beacon is called responder and it may reside in same frequency as interrogated by the interrogating radar or in the special beacon frequency which case a separate receiver is required by the interrogating radar .
- One of the functions of beacon is to identify . For example a beacon may be installed on such as an aircraft and it will transmit a coded pulse when interrogated. These pulses appear of PPI of the Interrogating radar and inform it about identify of the target. When used as such it is called Identification Friend or Foe (IFF).
- An IFF operates satisfactorily with naval targets . It however if large number of fast planes : some friends and some foes are involved.
- Another application of beacon is that it can work as a "light house". A beacon light house can operate over much larger distances.

CW RADARS

Two types of continuous wave (CW) radars are described below:

(a) Unmodulated Continuous Wave (CW) Radars

- A CW radar sends continuous waves (sine wave) rather than pulses. It also uses the Doppler effect to detect the frequency change caused by moving target and displays it as a velocity of the target in kilometre per hour.
- Un-modulated CW radar has been shown in figure . Since the transmission is continuous , a duplexer is not needed. Seperate antennas for transmission and reception have been shown.
- A small portion of transmitter output is mixed with local oscillator and is sum is fed to the mixer. This also receives the shifted signal from the receiving antenna and produces an output difference frequency called Doppler frequency.

- The output to the mixer is amplified , demodulated and signal from the second detector gives the Doppler frequency , after amplification the output is given to the indicator.

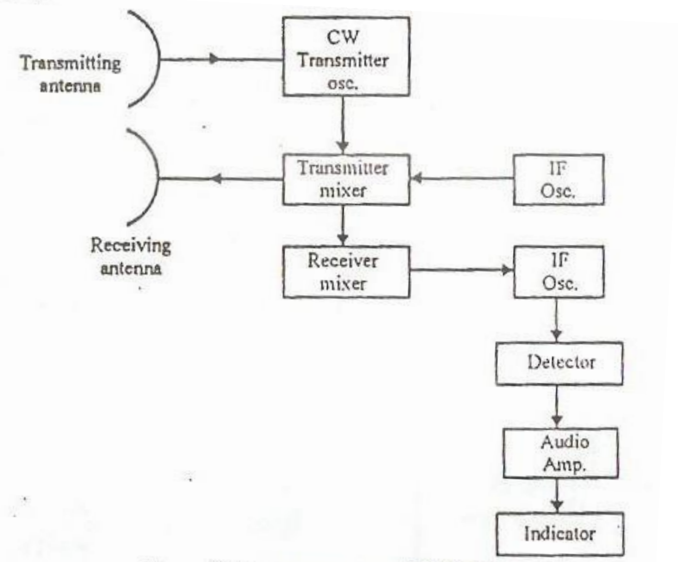


Figure 16.11 Unmodulated CW Radar

Advantages :

(i)Un-modulated CW radar gives accurate measurement of relative velocities , using low transmitter power and low power consumption.

(ii)This radar is comparatively of lesser size than the widely used pulse radars.

(iii)It is unaffected by stationary targets.

(iv)The receiver of the radar is ON all the time, it can be operated even at zero range.

(v)It can find the direction of the target in addition to its speed.

Disadvantages:

(i) It has limitations of its maximum range , which is very smaller than of pulse radar.

(ii) It is incapable to indicate the range of target , it can show only its velocity.

Applications:

(i)It is used in aircraft navigation.

(ii) It can measure speed of missiles, automobiles etc.

(b) Frequency Modulated CW Radar (Altimeter)

- The greatest limitation of unmodulated CW radars was that they could not measure the range .
- If the transmitted signal is frequency modulated the same radar will able to measure the range, because then it will able to distinguish one cycle from another .
- Using FM will require an increase in BW of the system .
- In this way , a FMCW radar is an improved form of a CW unmodulated radar. As the main application of a FMCW radar is to measure the height of an aircraft or plane, it is generally known as an Altimeter. The figure shows the block diagram of FMCW radar being used as an Altimeter.

eing used as an altimeter.

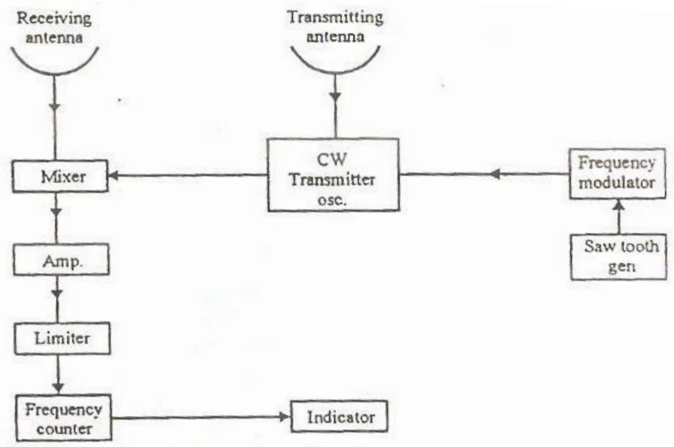


Figure 16.12 Modulated CW Radar

- As an altimeter , if the target is stationary wrt to a plane whose height is to be measured, a frequency difference proportional to the height of the plane will exit between transmitted and the received signals , because the signal now being received was at a time, when the frequency was different .
- If the rate of change of frequency w.r.t time due to FM process is known , the difference in the time between sent and the received signals may be found and thus the gieght of the plane above earth can be easily calculated.
- Thus output of the mixer which produces the frequency difference can be amplified and fed to the frequency counter and then to the indicator, whose output is calibrated in kilometres.
- As an altimeter, FMCW radar is preferred to the pulse radar , as the pulse radar , as the FMCW radar as no limit on the minimum range, as short heights are involved in this measurement .
- Further simple, low power and small equipment may be used with small antenna . A typical FMCW radar uses a 1-2 W transmitter power easily obtained from a diode and has a range of about 10 km or more.

TRACKING RADAR

- The tracking radar does the job of continuously tracking a moving target. It is usually a ground base system used to tack the air born targets.
- The tracking radar antenna sends out the very narrow beam , whose beam width could be anywhere between fraction of degree to one degree both in azimuth and elevation to get the requisite resolution for tracking purpose .
- One can visualize that it is imperative to use a search radar and acquire the target with comparatively much larger bandwidth before a track action is initiated.

LASER RADAR

- A laser radar uses an optical beam instead of microwaves, in other words, the e,m, energy transmitted lies in the optical spectrum in laser radars where as in microwave radar, it lies in the microwave region .
- In laser radar, the frequencies involved are high. It is possible to generate radar pulses as narrow as a fraction of picoseconds.

PHASED ARRAYRADAR

A phased array is a group of antennas whose effective (summed) radiation pattern can be altered by phasing the signals of the individual elements.

By varying the phasing of the different elements, the radiation pattern can be modified to be maximized / suppressed in given directions, within limits determined by,

- The radiation pattern of the elements,
- The size of the array, and
- The configuration of the array.

SOME BENEFITS OF PHASED ARRAYS

- Does not require moving a large structure around the sky forpointing. (Less infrastructure)
- Fast steering. (Pulse-to-pulse)
- Distributed, solid-state transmitters as opposed to single RFsources. (Less warm-up time, no need for complex feed system,elimination of single-point failures)

These features allow for:

- Remote operations
- Graceful degradation / continual operations
- Impact on ionospheric research:
- Elimination of some time-space ambiguities
- Ability to "zoom-in" in time
- Long durations runs (e.g., IPY)
- Non-ionospheric scientific benefits
- Radio astronomy - affordable way to achieve spatial resolutions of a few arc minutes or better
- Aperture real-estate - directly associated with cost of system. E.g., consider a square kilometer dish versus square kilometer array.
- Non-scientific benefits
- Conformity of a phased array to the "skin" of a vehicle/ aircraft
- Surveillance/tracking - can both survey and track 1000s of objects
- Communication/downlink - small satellites

PLANAR ARRAY RADAR

A fixed delay is established between horizontal arrays in the elevation plane. As the frequency is changed, the phase front across the aperture tends to tilt, with the result that the beam is moved in elevation. The differing frequencies cause each successive beam to be elevated slightly more than previous beams. A 27.5 degree elevation is scanned by the radar.

ADVANTAGES:

- Each beam group has full transmitter peak power, full antenna gain and full antenna sidelobe performance.
- The use of frequency changes allows economical, simple and reliable inertialess elevation scanning.

UNIT-V
TELEVISION

Introduction of Television-Television systems and standards-Black and white Transmission-black and white reception-colour transmission and reception- Digital Television-Digital TV receivers – colour receiver of new generation- EDTV- HDTV- colour receiver of the future. Introduction to modern TV cameras, LCD and plasma displays.

INTRODUCTION

- The aim of a television system is to extend the sense of sight beyond its natural limits and to transmitsound associated with the scene.
- The picture signal is generated by a TV camera and sound signal by amicrophone.
- In the 625 line CCIR monochrome and PAL-B colour TV systems adopted by India, thepicture 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 toobtain video (picture) and audio (sound) signals.
- The video signal after large amplification drives thepicture tube to reconstruct the televised picture on the receiver screen.
- Similarly, the audio signal isamplified and fed to the loudspeaker to produce sound output associated with the scene.

PICTURE TRANSMISSION

- The picture information is 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 amethod known as ‘scanning’ where conversion of optical information to electrical form is carried outelement 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 signalpath.

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 intocorresponding electrical signal, the amplitude of which varies in accordance with variations of brightness.
- Fig. 1.1 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.
- Thephotolayer 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 photolayer changes accordingly.
- An electron beam is used to pick-uppicture information now available on the target plate in terms of varying resistance at each point.

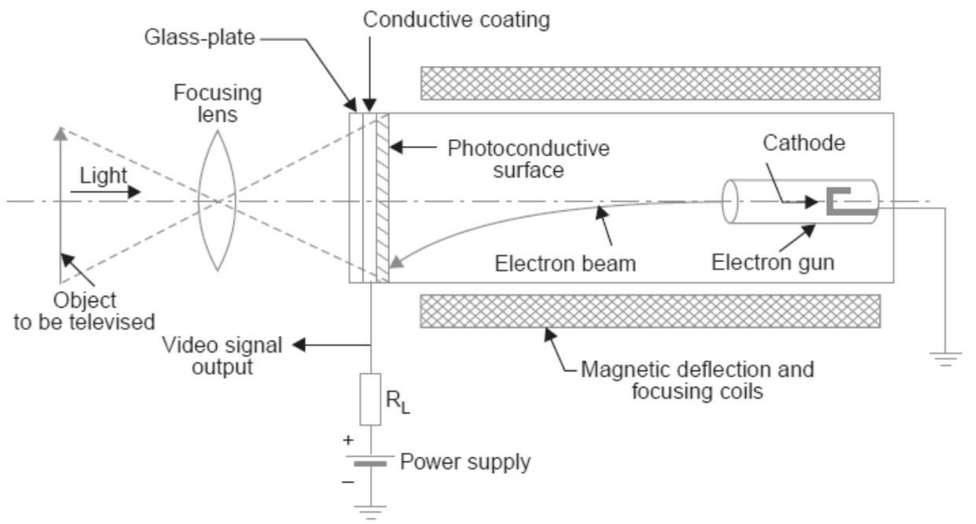


Fig. 1.1. Simplified cross-sectional view of a Vidicon camera tube and associated components.

- 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. 1.2).

- To achieve this, the deflecting coils are fed separately from two sweep oscillators whichcontinuously 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.
- Theother coil is used to deflect the beam from top to bottom at a uniform rate and for its quick retrace backto the top of the plate to start thisprocess over again.

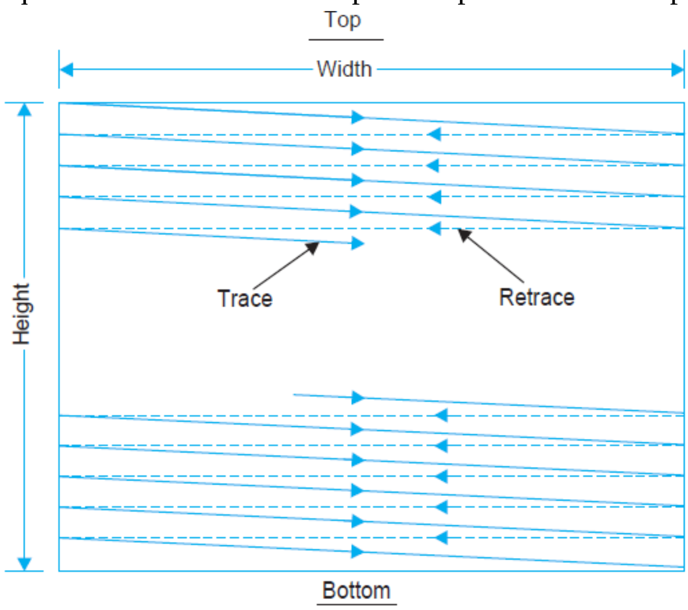


Fig. 1.2. Path of scanning beam in covering picture area.

- Two simultaneous motions are thus given to thebeam, one from left to right across the target plate and the other from top to bottom thereby coveringentire area on which electrical image of the picture is available.
- As the beam moves from element toelement, 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 voltageappears across resistance RLand this corresponds to optical information of the picture.
- If the scanning beam moves at such a rate that any portion of the scene content does not havetime to change perceptibly in the time required for one complete scan of the image, the resultant electricalsignal contains true information existing in the picture during the time of scan.
- The desired informationis now in the form of a signal varying with time and scanning may thus be identified as a particularprocess which permits conversion of information existing in space and time co-ordinates into timevariations only.

- The electrical information thus obtained from the TV camera tube is generally referredto as video signal (video is Latin for 'see').

Colour Pictures

- It is possible to create any colour including white by additive mixing of red, green and bluecolour lights in suitable proportions.
- For example, yellow can be obtained by mixing red and green colour lights in intensity ratio of 30 : 59.
- Similarly, light reflected from any colour picture element can be synthesised (broken up) into red, green and blue colour light constituents.
- This forms the basis ofcolour television where Red (R), Green (G) and Blue (B) colours are called primary colours and thoseformed by mixing any two of the three primaries as complementary colours.
- A colour camera, theelements of which are shown in Fig. 1.3, is used to develop signal voltages proportional to the intensityof each primary colour light.

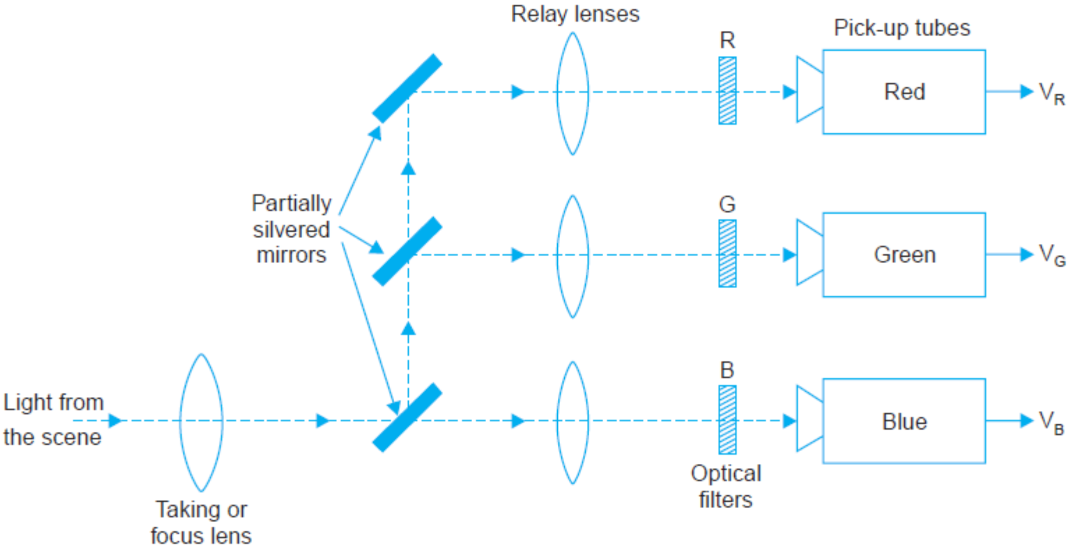


Fig. 1.3. Simplified block diagram of a colour camera.

- It contains three camera tubes (vidicons) where each pick-up tube receives light of only oneprimary colour. Light from the scene falls on the focus lens and through that on special mirrors.
- Colour filters that receive reflected light via relay lenses split it into R, G and B colour lights.
- Thus, each vidicon receives a single colour light and develops a voltage proportional to the intensity ofone of the primary colours.
- If any primary colour is not present in any part of the picture, thecorresponding vidicon does not develop any output when that picture area is scanned.
- The electronbeams of all the three camera tubes are kept in step (synchronism) by deflecting them horizontally andvertically from common driving sources.
- Any colour light has a certain intensity of brightness.

- Therefore, light reflected from any colourelement 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 proportionsof VR, VG and VB (30:59:11).
- This is the same as would be developed by a monochrome (blackand white) camera when made to scan the same colour scene. *i.e.*, the luminance (Y) signal is alsotransmitted alongwith colour information and used at picture tube in the receiver for reconstructing thecolour picture withbrightness levels as in the televised picture.

TELEVISION TRANSMITTER

- An oversimplified block diagram of a monochrome TV transmitter is shown in Fig. 1.4.
- The luminancesignal from the camera is amplified and synchronizing pulses added before feeding it to the modulatingamplifier.
- 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 continuouswave (CW) sine wave output is given large amplification before feeding to the power amplifierwhere its amplitude is made to vary (AM) in accordance with the modulating signal received fromthe modulating amplifier.
- The modulated output is combined (see Fig. 1.4) with the frequencymodulated (FM) sound signal in the combining network and then fed to the transmitting antenna forradiation.

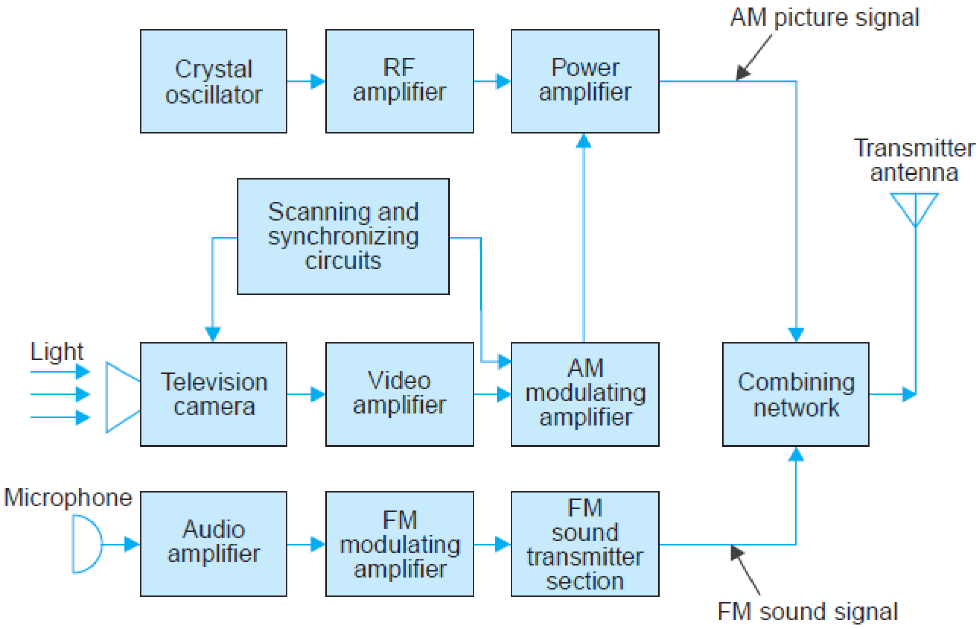


Fig. 1.4. Elementary block diagram of a monochrome television transmitter.

Colour Transmitter

- A colour TV transmitter is essentially the same as the monochrome transmitter except for theadditional need that colour (chroma) information is also to be transmitted.
- Any colour system is madecompatible with the corresponding monochrome system. Compatibility means that the colour TV signalmust produce a normal black and white picture on a monochrome receiver and a colour receiver mustbe able to produce a normal black and white picture from a monochrome TV signal.
- For this, theluminance (brightness) signal is transmitted in a colour system in the same way as in the monochromesystem and with the same bandwidth.
- However, to ensure compatibility, the colour camera outputs aremodified to obtain (B-Y) and (R-Y) signals.
- These are modulated on the colour sub-carrier, the value ofwhich is so chosen that on combining with the luminance signal, the sidebands of the two do notinterfere with each other *i.e.,* the luminance and colour signals are correctly interleaved.
- A colour syncsignal called ‘colour burst’ is also transmitted for correct reproduction of colours.

Sound Transmission

- There is no difference in sound transmission between monochrome and colour TV systems.
- Themicrophone converts the sound associated with the picture being televised into proportionate electricalsignal, 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 signalfrom the microphone after amplification is frequency modulated, employing the assigned carrierfrequency.
- In FM, the amplitude of carrier signal is held constant, whereas its frequency is varied inaccordance with amplitude variations of the modulating signal.
- As shown in Fig. 1.4, output of thesound FM transmitter is finally combined with the AM picture transmitter output, through a combiningnetwork, and fed to a common antenna for radiation of energy in the form of electromagnetic waves.

TELEVISION RECEIVER

- A simplified block diagram of a black and white TV receiver is shown in Fig. 1.5.
- The receivingantenna intercepts radiated RF signals and the tuner selects desired channel’s frequency band andconverts 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.

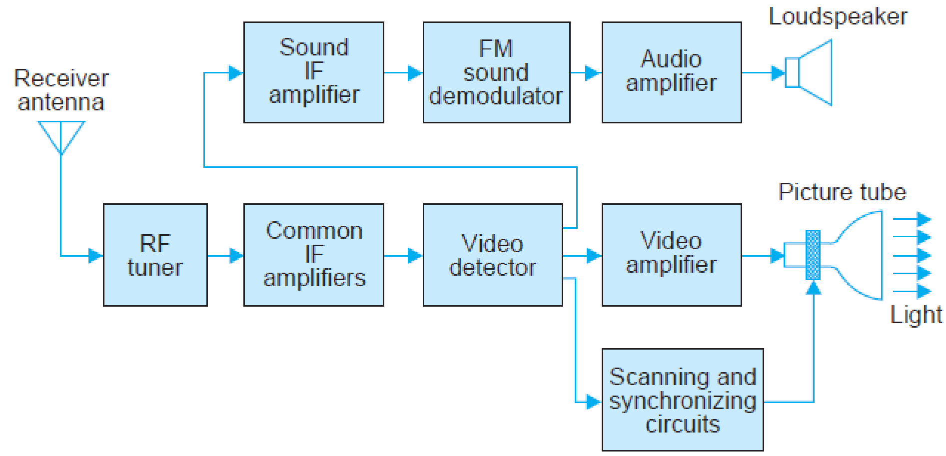


Fig. 1.5. Simplified block diagram of a black and white TV receiver.

- The picture tube shown in Fig. 1.6 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 thefluorescent screen.
- When the electron beam strikes the screen, light is emitted.
- The beam is deflectedby a pair of deflecting coils mounted on the neck of picture tube in the same way as the beam of cameratube scans the target plate.
- The amplitudes of currents in the horizontal and vertical deflecting coils areso adjusted that the entire screen, called raster, gets illuminated because of the fast rate of scanning.

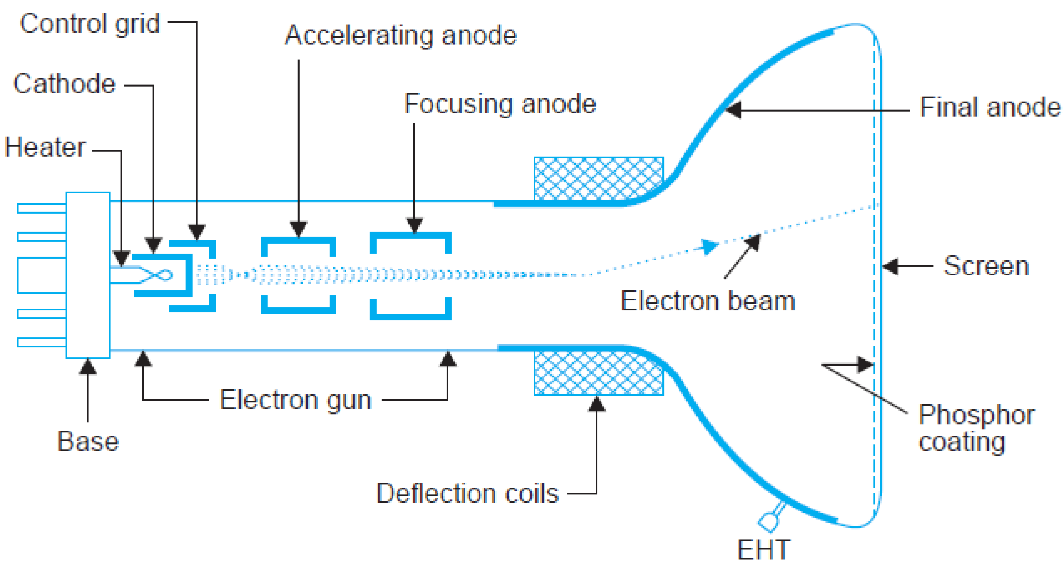


Fig. 1.6. Elements of a picture tube.

- The video signal is fed to the grid or cathode of picture tube.
- When the varying signal voltagemakes the control grid less negative, the beam current is increased, making the spot of light on thescreen brighter.

- More negative grid voltage reduces brightness. If the grid voltage is negative enough tocut-off the electron beam current at the picture tube, there will be no light. This state corresponds toblack.
- Thus the videosignal illuminates the fluorescent screen from white to black through variousshades of grey depending on its amplitude at any instant.
- This corresponds to brightness changesencountered by the electron beam of the camera tube while scanning picture details element by element.
- The rate at which the spot of light moves is so fast that the eye is unable to follow it and so a completepicture 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 detectorsection 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 audiooutput from the FM detector is given due amplification before feeding it to the loudspeaker.

Colour Receiver

- A colour receiver is similar to the black and white receiver as shown in Fig. 1.7.
- The maindifference between the two is the need of a colour or chroma subsystem.
- It accepts only the coloursignal and processes it to recover (B-Y) and (R-Y) signals.
- These are combined with the Y signal toobtain VR, VG and VB signals as developed by the camera at the transmitting end. VG becomes availableas it is contained in the Y signal.
- The three colour signals are fed after sufficient amplification to thecolour picture tube to produce a colour picture on its screen.
- As shown in Fig. 1.7, the colour picture tube has three guns corresponding to the three pick-up tubes in the colour 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 colour phosphor separately on the fluorescent screen.
- The eye then integrates the red, green and blue colourinformations and their luminance to perceive actual colour and brightness of the picture being televised.
- The sound signal is decoded in the same way as in a monochrome receiver.

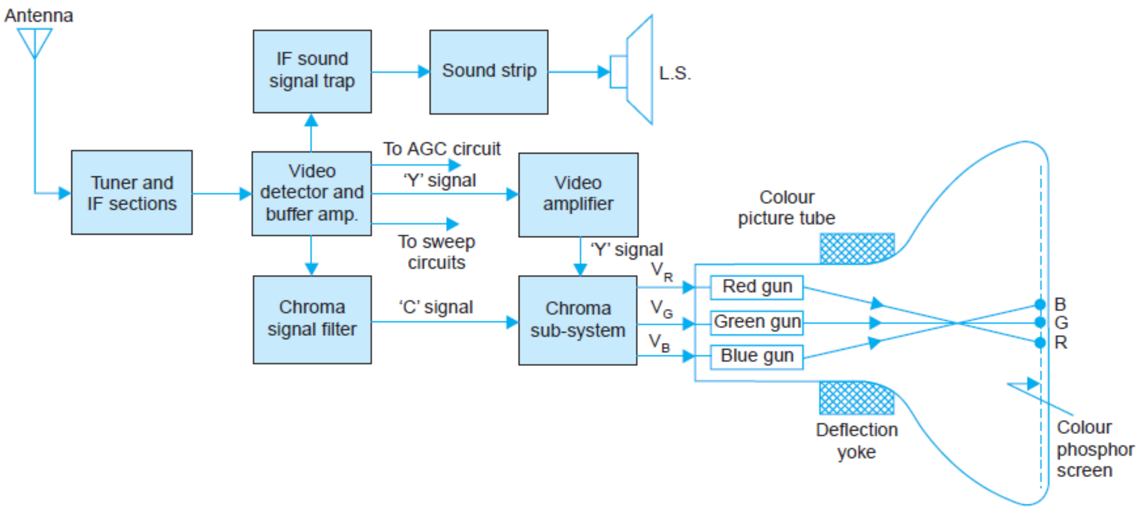


Fig. 1.7. An oversimplified block diagram of a colour receiver.

SYNCHRONIZATION

- It is essential that the same co-ordinates be scanned at any instant both at the camera tube target plateand at the raster of picture tube, otherwise, the picture details would split and get distorted.
- To ensureperfect synchronization between the scene being televised and the picture produced on the raster,synchronizing pulses are transmitted during the retrace, *i.e.*, fly-back intervals of horizontal and verticalmotions of the camera scanning beam.
- Thus, in addition to carrying picture details, the radiated signalat the transmitter also contains synchronizing pulses.
- These pulses which are distinct for horizontal andvertical motion control, are processed at the receiver and fed to the picture tube sweep circuitry thusensuring that the receiver picture tube beam is in step with the transmitter camera tube beam.
- As stated earlier, in a colour TV system additional sync pulses called **colour burst** are transmittedalong with horizontal sync pulses.
- These are separated at the input of chroma section and used tosynchronize the colour demodulator carrier generator.
- This ensures correct reproduction of colours inthe 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 desiredchannel.

- 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 variesbeam intensity of the picture tube and is set for optimum average brightness of the picture.
- The contrastcontrol is actually gain control of the video amplifier. This can be varied to obtain desired contrastbetween white and black contents of the reproduced picture.
- The volume and tone controls form part ofthe audio amplifier in sound section, and are used for setting volume and tonal quality of the soundoutput from the loudspeaker.
- In colour receivers there is an additional control called ‘colour’ or ‘saturation’ control. It is used tovary intensity or amount of colours in the reproduced picture.
- In modern colour receivers that employ integratedcircuits in most sections of the receiver, the hold control is not necessary and hence usually not provided.

BASIC FACTORS AFFECTING TELEVISION TRANSMISSION & RECEPTION

1.Gross structure
2. Image continuity
3. Number of scanning lines
4. Flicker
5. Fine structure
6. Tone gradation

1. Gross Structure

- Frame adopted is rectangle with Aspect Ratio (Width/Height) = 4/3

Reasons:

1. Most of the motion occurs in horizontal plane
2. Eyes can view more easily and comfortably
3. For enabling direct television transmission of film programs without wastage of any film area - Motion pictures use a rectangular frame with width/height ratio of 4/3 – so adopted this aspect ratio in TV

Requirements:

1. Aspect ratio of the size of the picture produced on the receiver screen and the picture being televised must be the same.
    - achieved by setting the magnitude of the current in the deflection coils to correct values both at the TV camera and the receiving picture tube
2. Same coordinated should be scanned at any instant both by the camera tube beam and the picture tube beam
    - achieved by transmitting synchronizing pulses along with the picture information

2. Image Continuity

- To create an illusion of continuity we make use of persistence of vision : sensation produced when nerves of the eye’s retina are stimulated by incident light does not cease immediately after the light is removed but persists for about 1/16th of a second

- Scanning rate is made greater than 16 per second i.e. number of pictures shown per second is more than 16 – hence our eye can able to integrate the changing levels of brightness in the scene
- Present day motion pictures – 24 still pictures of the scene are taken per second and projected on the screen at the same rate

Scanning in Television systems

- The scene is scanned rapidly both in the horizontal and vertical directions simultaneously
- Frame repetition rate is 25 per second

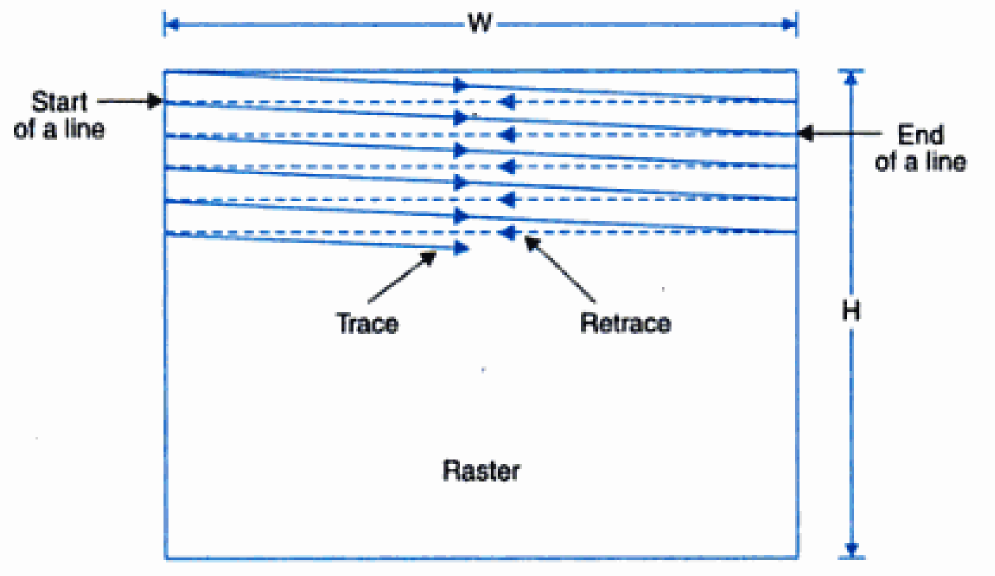


Fig. 2.1 (a) Path of scanning beam in covering picture area (Raster).

Horizontal scanning

- Linear rise of current in the deflection coils deflects the beam across the screen with a continuous uniform motion for the trace from left to right .
- At the peak of the rise, the saw tooth wave reverses its direction and decreases rapidly to its initial value, producing the retrace or flyback

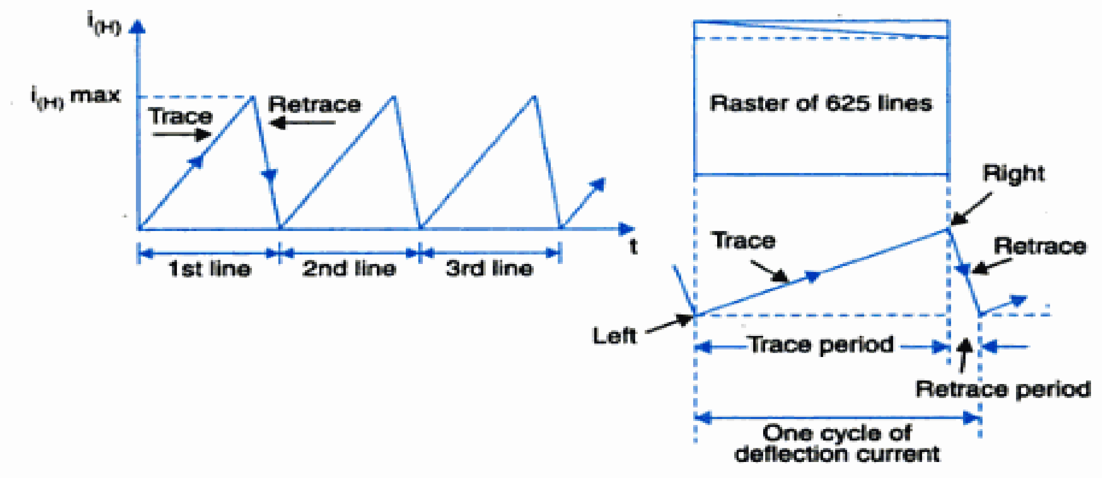
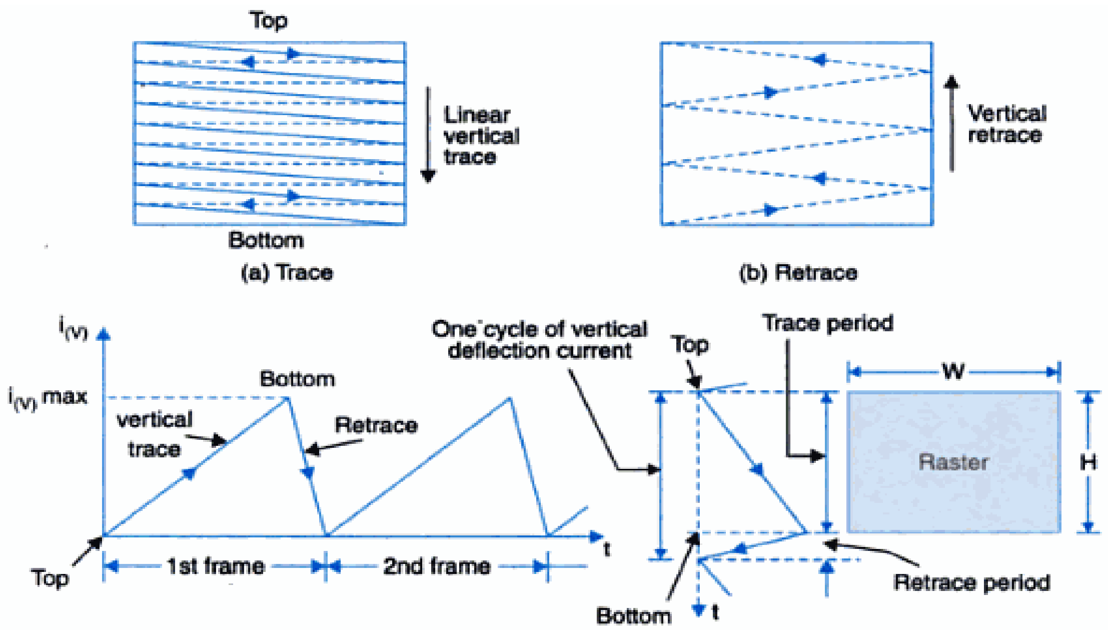


Fig. 2.1 (*b*) Waveform of current in the horizontal deflection coils producing linear (constant velocity) scanning in the horizontal direction.

Vertical Scanning

- Saw tooth current in the vertical deflection coils moves the electron beam from top to bottom of the raster at uniform speed while the electron beam is being deflected horizontally



(c) Current waveform in vertical deflection coils

- Because of the motion in the scene being televised, the information or brightness at the top of the target plate or picture tube screen normally changes by the time the beam returns to the top to recommence the whole process. This information is picked up during the next scanning cycle and the whole process is repeated 25 times to cause an illusion of continuity
- During the horizontal and vertical retrace intervals, the scanning beams at the camera tube and the picture tube are blanked and no picture information is either picked up or reproduced

- Synchronizing pulses are transmitted during this period – resulting in distortionless reproduction of the picture details

3. Number of scanning lines

- Most scenes have brightness gradations in the vertical directions
- The ability of the scanning beam to allow reproduction of electrical signals according to these variations and the capability of the human eye to resolve these distinctly (while viewing) depends on the total number of lines employed for scanning
- Number of scanning lines is judged by considering the bar pattern as shown where alternate lines are black and white
- If the thickness of the scanning beam is equal to the width of each black and white bar and the number of scanning lines is chosen equal to the number of bars, then the electrical information corresponding to the brightness of each bar will be correctly

reproduced duringthe scanning process

- Greater the number of lines, better will be the resolution

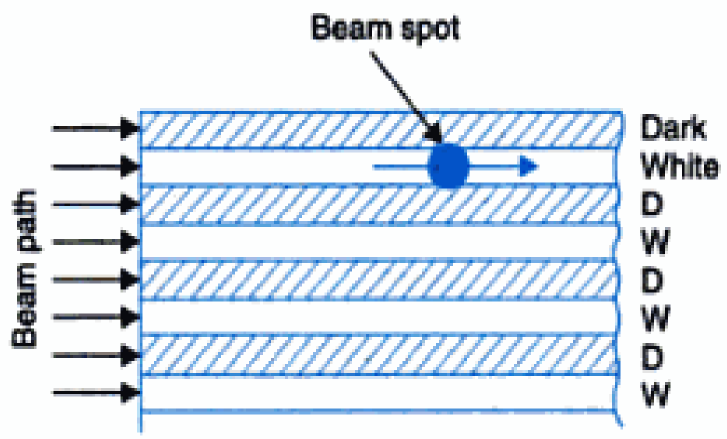


Fig. 2.3 (*a*) Scanning spot perfectly aligned with black and white lines.

- However the total number of lines is limited by the resolving capability of the human eye at the minimum viewing distance

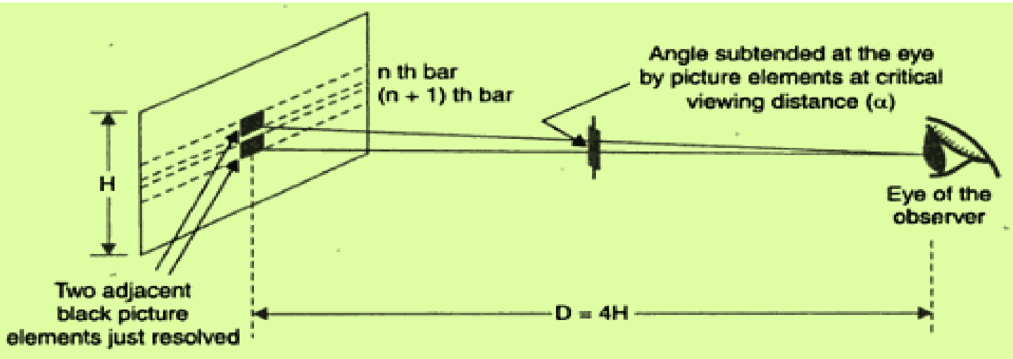


Fig. 2.3 (*b*) Critical viewing distance as determined by the ability of the eye to resolve two separate picture elements.

The maximum number of alternate light and dark elements (lines) which can be resolved by the eye is given by

$$N_v = \frac{1}{\alpha\rho}$$

where N_v = total number of lines (elements) to be resolved in the vertical direction, α = minimum resolving angle of the eye expressed in radians, and $\rho = D/H$ = viewing-distance/picture height.

- With reasonable brightness variation and a minimum viewing distance of 4 times the picture height (D/H = 4), the angle that any two adjacent elements must subtend at the eye for distinct resolution is approximately one minute (1/60 degree)
- Substituting the above values

$$N_v = \frac{1}{(\pi / 180 \times 1/60) \times 4} \approx 860$$

- In practice, the picture elements are not arranged as equally spaced elements but have random distribution of black, grey and white depending on the nature of the picture details.
- Analysis and tests suggests that about 70% of the total line get separately scanned in the vertical direction and the remaining 30% get merged with other elements due to the beam spot falling equally on two consecutive lines (as shown in figure)

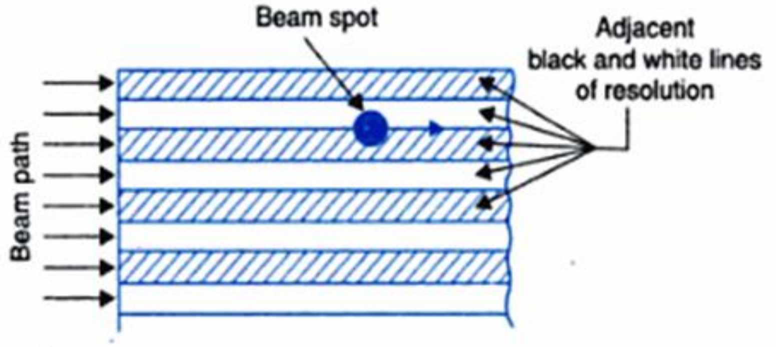


Fig. 2.3 (*c*) Scanning beam focused on the junction of black and white lines.

- Thus the effective number of lines distinctly resolved

  Nr = Nv × K

  where K is the resolution factor whose value lies between 0.65 & 0.75
- Assuming the value of k = 0.7;

  Nr = 860 × 0.7 = 602

Other factors influencing the choice of total number of lines

- Improvement in resolution is not very significant with line numbers > 500
- Channel bandwidth increases with the increase in number of lines
  - cost of the system increases
  - reduces the number of channels in a given VHF/UHFtransmission band

NB: As a compromise between quality and cost, the total number of lines (inclusive of those lost during vertical retrace) has been chosen to be 625 in the 625-B monochrome TV system.

4. Flicker

- 25 frames per second in television picture is not rapid enough to allow the brightness of one picture or frame to blend smoothly into the next during the time when the screen is blanked between successive frames
- Produces ***flicker***
- *Eliminated in motion pictures by showing each picture twice – ie 48 views of the scene per second – still the same 24 picture frames per second*
- *ie increased blanking rate*

Interlaced scanning

- In Television pictures *50 vertical scans per second* – to reduce flicker
- Downward rate of travel of scanning electron beam is increased
- Alternate lines get scanned instead of every successive lines
- Here total number of lines are divided into two groups called 'fields'
- ie each field is scanned alternatively – called as *interlaced scanning* – reduces flicker

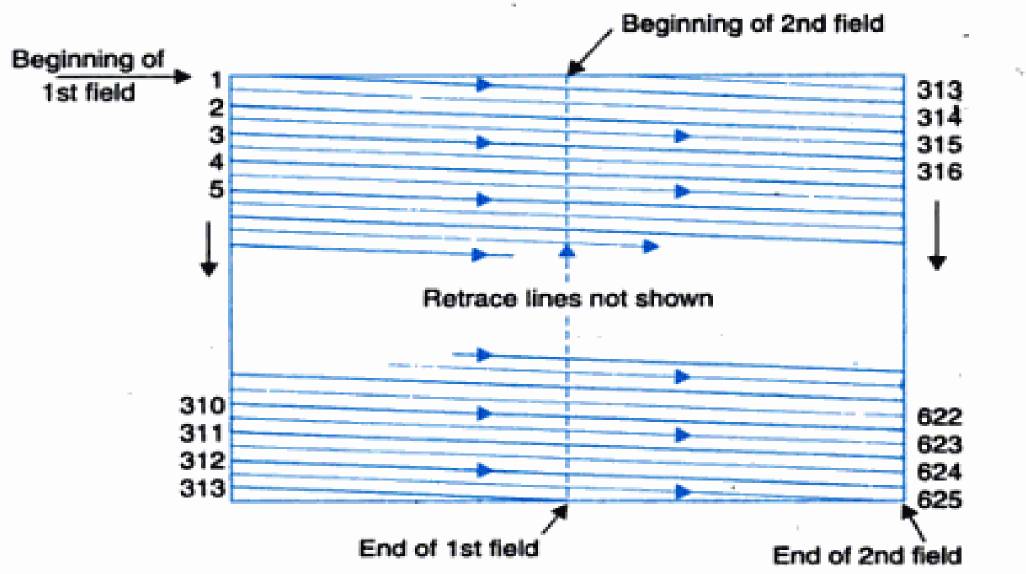


Fig. 2.4 Principle of interlaced scanning. Note that the vertical retrace time has been assumed to be zero.

- 625 lines of each frame /picture are divided into sets of 312.5 lines
- Each set is scanned alternatively
- Horizontal sweep oscillator is made to work at a frequency of 15625 Hz ($312.5 \times 50 = 15625$)
- Vertical sweep circuit run at a frequency of 50 Hz instead of 25 Hz
- Reduces the undesired effects of hum due to pickup from mains

Scanning periods

- Normal duration of the horizontal line is 64μs (1/15625 = 64μs)

active line period = 52µs
line blanking period = 12µs

- Normal duration of the vertical trace is 20 ms (1/50 = 20 ms)

  18.720 ms - for bringing the beam from top to bottom
  1.280 ms – to commence the next cycle

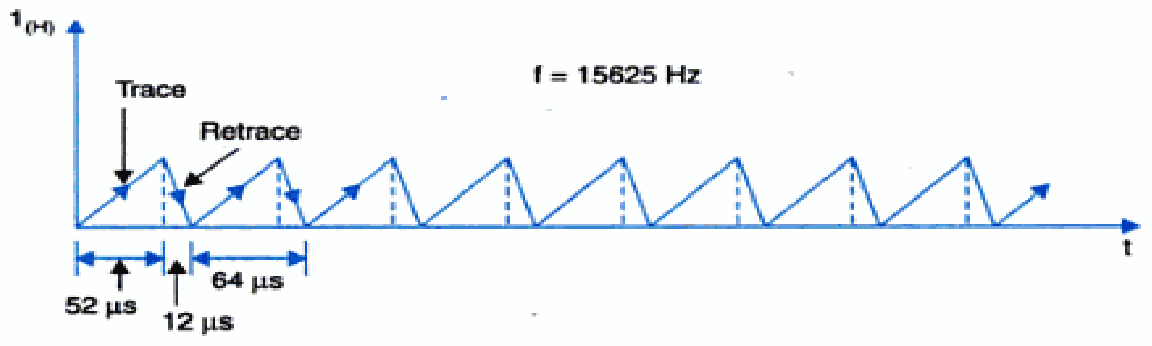


Fig. 2.5 (a) Horizontal deflection current.

- 20 horizontal lines could be scanned during each vertical retrace interval
- Thus 40 scanning lines are lost per frame
- Now active number of scanning lines Na = 625-40 = 585

5. Fine Structure

- The ability of the image reproducing system to represent the fine structure of an object – resolving power or resolution

  Vertical resolution Vr = Na × K; Na-active number of lines
- K-resolution factor
- Assuming K=0.69; Vr = 585 × 0.69 = 400 lines

  Horizontal resolution

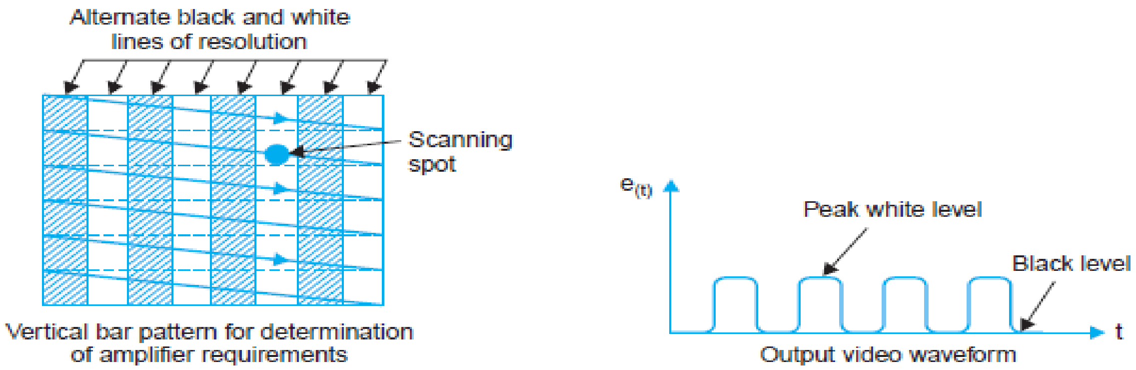


Fig. 2.7 (a) Determination of horizontal resolution.

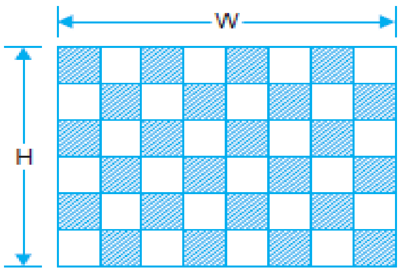


Fig. 2.7(b) Chess-board pattern for studying vertical and horizontal resolution.

time duration t_h of one square wave cycle is equal to

$$t_h = \frac{\text{active period of each horizontal line}}{\text{number of cycles}}$$

$$= \frac{52 \times 10^{-6}}{267} \text{ seconds}$$

$\therefore$ thg frequency of the periodic wave

$$f_h = \frac{1}{t_h} = \frac{267 \times 10^6}{52} = 5 \text{ MHz}$$

Therefore, the highest approximate modulating frequency 'f_h' that the 625 line television system must be capable of handling for successful transmission and reception of picture details is

$$f_h = \frac{\text{No. of active lines} \times \text{aspect ratio} \times \text{resolution factor}}{2 \times \text{time duration of one active line}}$$

$$= \frac{585 \times 4/3 \times 0.69}{2 \times 52 \times 10^{-6}}$$

$$\approx 5 \text{ MHz}$$

Interlace error

Usually the interlace ratio is 2:1

- Any error in the scanning timings and sequence wouldreduce the quality of the reproduced picture

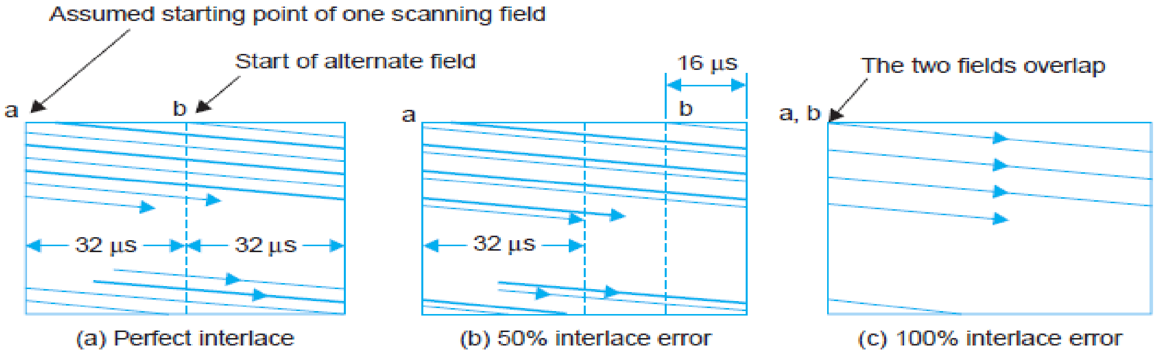


Fig. 2.9 Examples of interlace error.

$$\text{percentage interlace error} = \frac{48-32}{32} \times 100 = 50\%$$

For interlaced scanning, the total number of lines in any TV system must be odd

6. Tonal Gradation
Factors that affect the tonal quality of the reproduced picture are

a) Contrast
b) Contrast ratio
c) Viewing distance

COMPARISON OF VARIOUS TV SYSTEM
CCIR 625 B monochrome system – most parts of Europe and India. BW: 5MHz, Resolution factor:0.69, Line frequency: 15625

625 line – Britan - BW: 5.5MHz, Resolution factor: 0.73

819 line – France. BW:10.4MHz (improved vertical and horizontal resolution)

525 line – America. Frame frequency:30, line frequency: 15750, BW: 4MHz(ie lesser horizontal resolution)
Note: Greater the no. of lines, better the vertical resolution
Greater the BW, better the horizontal resolution

DIGITAL TELEVISION

Digital television (**DTV**) is the transmission of audio and video by digitally processed and multiplexed signal, in contrast to the totally analog and channel separated signals used by analog television. Digital TV can support more than one program in the same channel bandwidth. It is an innovative service that represents the first significant evolution in television technology.

There are four different widely used digital television terrestrial broadcasting standards (DTTB):

- Terrestrial Integrated Services Digital Broadcasting (ISDB-T) is a system designed to provide good reception to fix receivers and also portable or mobile receivers. It utilizes OFDM and two-dimensional interleaving. It supports hierarchical transmission of up to three layers and usesMPEG-2 video and Advanced Audio Coding. ISDB-T International is an adaptation of this standard using H.264/MPEG-4 AVC.
- Digital Video Broadcasting-Terrestrial (DVB-T) uses coded orthogonal frequency-division multiplexing (OFDM) modulation and supports hierarchical transmission
- Advanced Television System Committee (ATSC) uses eight-level vestigial sideband (8VSB) for terrestrial broadcasting.
- Digital Terrestrial Multimedia Broadcasting (DTMB) adopts time-domain synchronous (TDS) OFDM technology with a pseudo-random signal frame to serve as the guard interval (GI) of the OFDM block and the training symbol.

FORMATS AND BANDWIDTH:

- Digital television supports many different picture formats defined by the broadcast television systems which are a combination of size, aspect ratio(width to height ratio).
- With digital terrestrial television (DTT) broadcasting, the range of formats can be broadly divided into two categories: high definition television(HDTV) for the transmission of high-definition video and standard-definition television (SDTV).
- One of several different HDTV formats that can be transmitted over DTV is: 1280 × 720 pixels in progressive scan mode (abbreviated *720p*) or 1920 × 1080 pixels in interlaced video mode. Each of these uses a 16:9 aspect ratio. (Some televisions are capable of receiving an HD resolution of 1920 × 1080 at a 60 Hz progressive scan frame rate — known as 1080p.) HDTV cannot be transmitted over current analog television channels because of channel capacity issues.
- Standard definition TV (SDTV), by comparison, may use one of several different formats taking the form of various aspect ratios depending on the technology used in the country of broadcast. For 4:3 aspect-ratio broadcasts, the 640 × 480 format is used in NTSC countries, while 720 × 576 is used in PAL countries. For 16:9 broadcasts, the 720 × 480 format is used in NTSC countries, while 720 × 576 is used in PAL countries. However, broadcasters may choose to reduce these resolutions to save bandwidth (e.g., many DVB-T channels in the United Kingdom use a horizontal resolution of 544 or 704 pixels per line).

- Each commercial broadcasting terrestrial television DTV channel in North America is permitted to be broadcast at a bit rate up to 19 megabits per second. However, the broadcaster does not need to use this entire bandwidth for just one broadcast channel. Instead the broadcast can use the channel to include PSIP and can also subdivide across several video subchannels (aka feeds) of varying quality and compression rates, including non-video datacasting services that allow one-way high-bandwidth streaming of data to computers like National Datacast.
- A broadcaster may opt to use a standard-definition (SDTV) digital signal instead of an HDTV signal, because current convention allows the bandwidth of a DTV channel (or "multiplex") to be subdivided into multiple digital sub-channels, (similar to what most FM radio stations offer with HD Radio), providing multiple feeds of entirely different television programming on the same channel. This ability to provide either a single HDTV feed or multiple lower-resolution feeds is often referred to as distributing one's "bit budget" ormulticasting. This can sometimes be arranged automatically, using a statistical multiplexer (or "stat-mux"). With some implementations, image resolution may be less directly limited by bandwidth; for example in DVB-T, broadcasters can choose from several different modulation schemes, giving them the option to reduce the transmission bitrate and make reception easier for more distant or mobile viewers.

RECEIVING DIGITAL SIGNAL:

- There are several different ways to receive digital television. One of the oldest means of receiving DTV (and TV in general) is using an antenna (known as an *aerial* in some countries). This way is known as Digital terrestrial television (DTT). With DTT, viewers are limited to whatever channels the antenna picks up. Signal quality will also vary. Regardless of what sales ads tried to leave the public to believe, there is no such thing as a specialized DTV antenna. ANY Over the Air antenna that worked for analog TV should work for Digital TV (BUT DTV signal levels are lower thus requiring actually a bigger antenna with more gain unless you are visually close to the transmitting towers).
- Other ways have been devised to receive digital television. Among the most familiar to people are digital cable and digital satellite. In some countries where transmissions of TV signals are normally achieved by microwaves, digital MMDS is used. Other standards, such as Digital multimedia broadcasting (DMB) and DVB-H, have been devised to allow handheld devices such as mobile phones to receive TV signals. Another way is IPTV, that is receiving TV via Internet Protocol, relying on Digital Subscriber Line (DSL) or optical cable line. Finally, an alternative way is to receive digital TV signals via the open Internet. For example, there is P2P (peer-to-peer) Internet television software that can be used to watch TV on a computer.

- Some signals carry encryption and specify use conditions (such as "may not be recorded" or "may not be viewed on displays larger than 1 m in diagonal measure") backed up with the force of law under the WIPO Copyright Treaty and national legislation implementing it, such as the U.S. Digital Millennium Copyright Act. Access to encrypted channels can be controlled by a removable smart card, for example via the Common Interface (DVB-CI) standard for Europe and via Point Of Deployment (POD) for IS or named differentlyCableCard.

ENHANCED-DEFINITION TELEVISION:

Enhanced-definition television,or **extended-definition television** (**EDTV**) is an American Consumer Electronics Association (CEA) marketing shorthand term for certaindigital television (DTV) formats and devices. Specifically, this term defines formats that deliver a picture superior to that of standard-definition television (SDTV) but not as detailed as high-definition television (HDTV).

HIGH-DEFINITION TELEVISION (HDTV)

High-definition television (**HDTV**) provides a resolution that is substantially higher than that of standard-definition television.

HDTV may be transmitted in various formats:

- 1080p: 1920×1080p: 2,073,600 pixels (~2.07 megapixels) per frame
- 1080i: 1920×1080i: 1,036,800 pixels (~1.04 MP) per field or 2,073,600 pixels (~2.07 MP) per frame
- A non-standard CEA resolution exists in some countries such as 1440×1080i: 777,600 pixels (~0.78 MP) per field or 1,555,200 pixels (~1.56 MP) per frame
- 720p: 1280×720p: 921,600 pixels (~0.92 MP) per frame

The letter "p" here stands for progressive scan while "i" indicates interlaced.

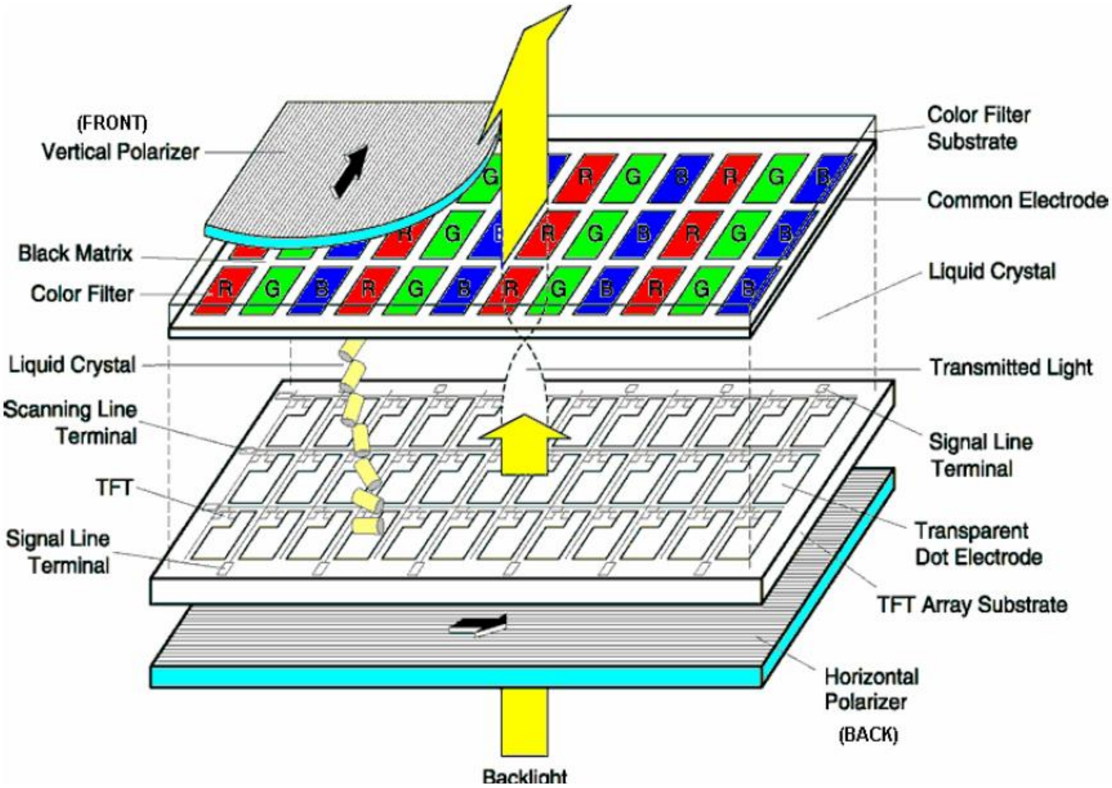
When transmitted at two megapixels per frame, HDTV provides about five times as many pixels as SD (standard-definition television).

LCD AND PLASMA DISPLAYS

Image production – LCD

- LCD and plasma displays incorporate different fixed matrix technologies that provide superior clarity and definition when used at their native resolution.
- How the two displays produce an image, however, is quite different.
- An active matrix LCD’s light source is normally generated by small fluorescent bulbs.
- The light from these bulbs is diffused to create uniform light across the back polarizer, which allows light to go in through only one direction.

Figure 1. Active matrix LCD structure diagram.



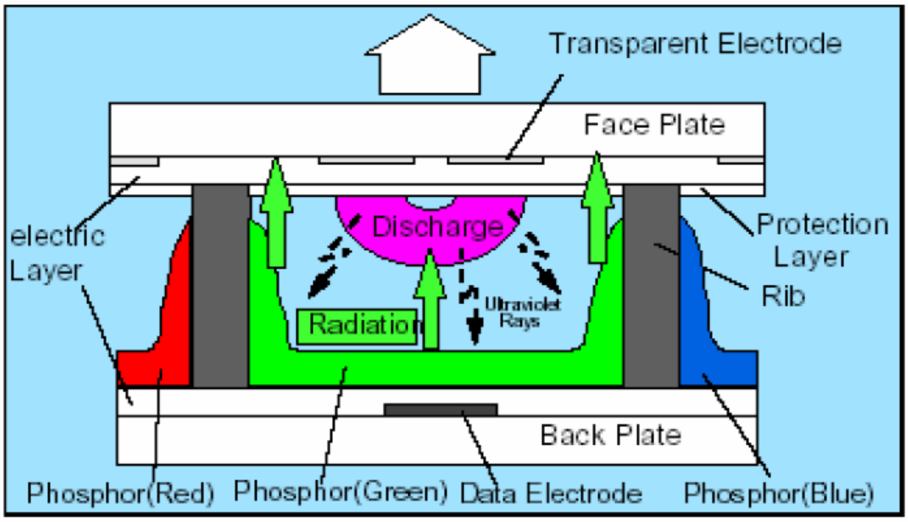
- Individual cells are turned "on" and "off" by applying a small electric charge to the thin film transistor (TFT), located in each sub pixel.
- This small charge causes the liquid crystal to "twist," allowing light to be passed through the color filters and the front polarizer.
- If the LCD cell is not turned on, then the light doesn't pass through the front polarizer, which is perpendicular in respect to the back polarizer.

PLASMA DISPLAYS

- The very first prototype for a plasma display monitor was invented in July 1964 at the University of Illinois by professors Donald Bitzer and Gene Slottow, and then graduate student Robert Willson.
- A plasma display is composed of two parallel sheets of glass, which enclose a mixture of discharged gases composed of helium, neon or xenon.
- Dike-like barriers, or ribs, keep the glass plates parallel and separate.
- Groups of electrodes sit at right angles between the panes, forming rectangular compartments, or cells, between the glass sheets.
- Phosphors embedded within each cell individually emit red, green or blue light and collectively create a single color pixel when excited.
- Selectively applying voltages to the electrodes causes them to generate a discharge in the panel's dielectric layer and on its protective surface.
- This generates ultraviolet light that excites the phosphors, stimulating them to emit light.
- This principle of operation is very similar to that of a fluorescent lamp.

- In this sense, it is possible to think of a plasma display as a screen incorporating thousands of miniature fluorescent lights of different color.

Figure 2. Plasma display structure diagram.



- Selectively applying voltages to the electrodes causes them to generate a discharge in the panel’s dielectric layer and on its protective surface.
- This generates ultraviolet light that excites the phosphors, stimulating them to emit light.
- This principle of operation is very similar to that of a fluorescent lamp.
- In this sense, it is possible to think of a plasma display as a screen incorporating thousands of miniature fluorescent lights of different color.

	Typical 40” LCD	Typical 42” Plasma
Brightness (full screen white)	450 cd/m²	100 cd/m²
Contrast Ratio (full screen white to full screen black)	600:1	200:1

- The specifications for LCD and plasma displays are measured differently for brightness and contrast.
- LCDs measure brightness according to the Video Electronics Standards Association (VESA) Flat Panel Display Measurements (FPDM) Standard Version 2.0 (June 1, 2001), by using a full screen white pattern.
- Contrast ratio according to the VESA standard is measured as the difference between full screen white and full screen black in a dark room.
- Plasma display specifications are measured in the same way a cathode ray tube (CRT) would be measured.
- The brightness is specified by using a peak value rather than a typical value.

- This is done by generating a small white square on the screen, concentrating all of the display's energy in this small area.
- The contrast ratio is calculated as the difference between the small white area and the black area surrounding it.

Figure 3. Comparison of LCD and plasma brightness measurements.

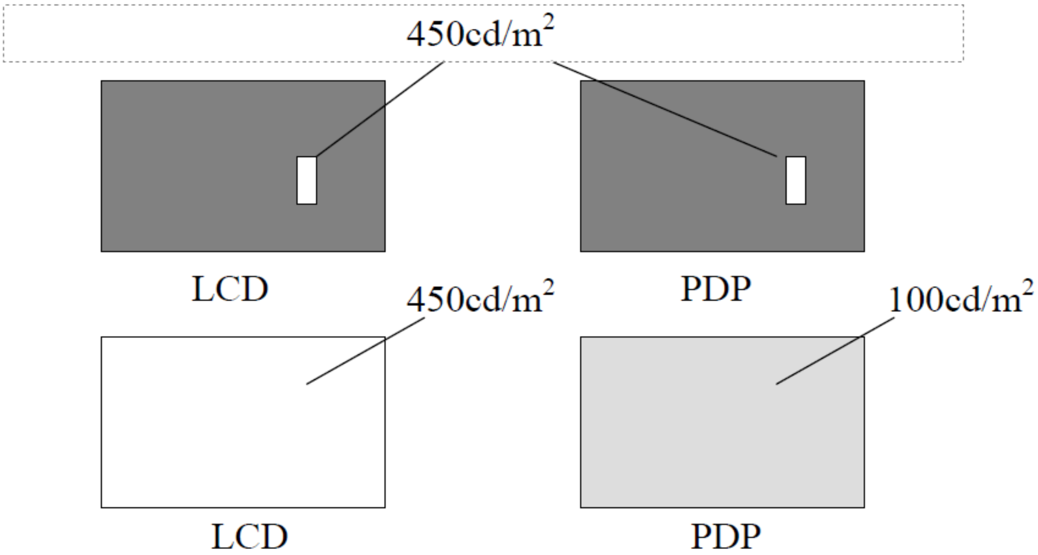


Figure 4. Comparison of LCD (l) and plasma (r) contrast ratio measurements.

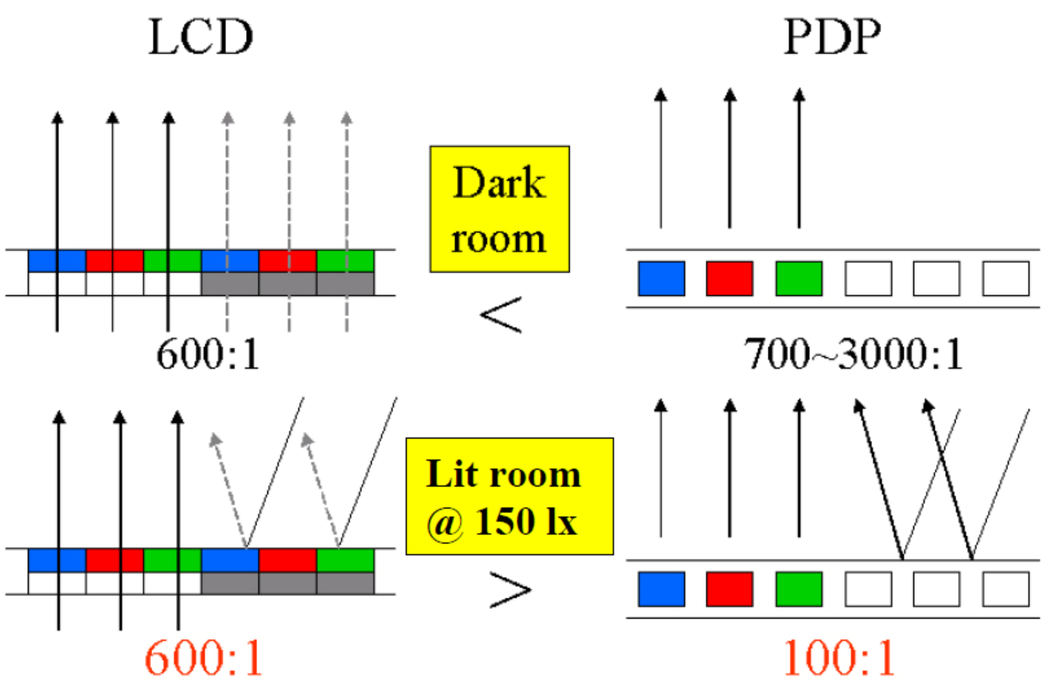


Figure 5. Comparison of LCD and plasma contrast ratios and brightnesses in certain operating environments.

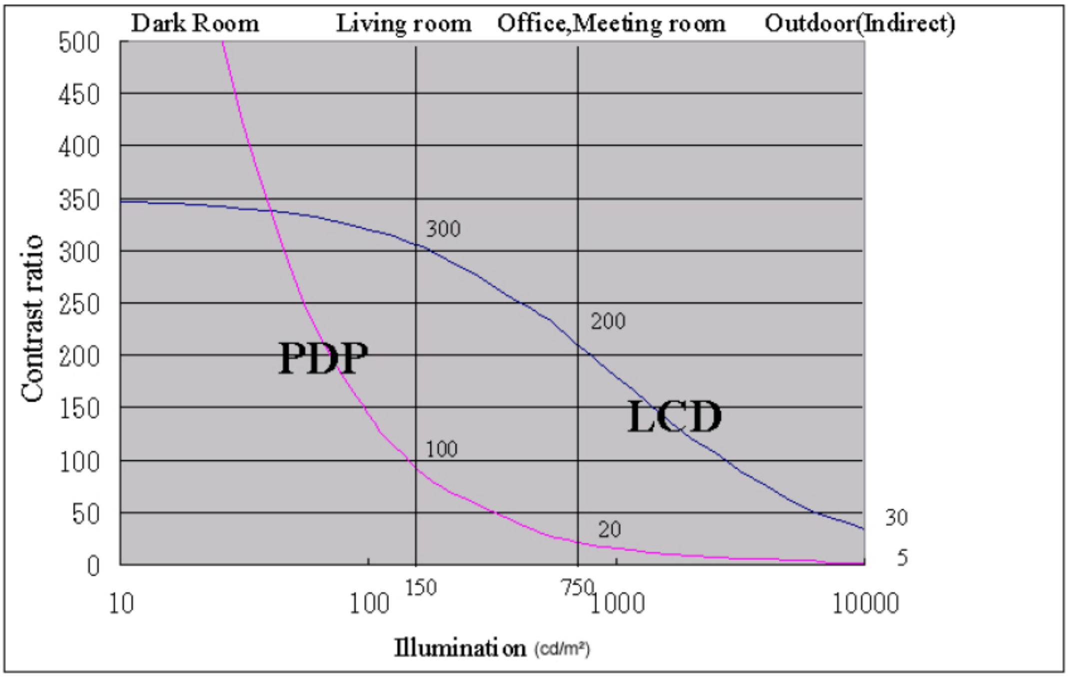


Figure 5 shows that LCD technology surpasses plasma technology in all environments except dark rooms.

- LCD displays also perform better than plasma displays under higher ambient-lit conditions (Fig. 4).
- An LCD display absorbs ambient light while a plasma display reflects it.
- The contrast of both displays is reduced under these conditions, but it is reduced to a much higher degree for the plasma.

Plasma display advantages and disadvantages

- **Advantages**
- Picture quality
  - Capable of producing deeper blacks allowing for superior contrast ratio
  - Wider viewing angles than those of LCD; images do not suffer from degradation at high angles like LCDs
  - Less visible motion blur, thanks in large part to very high refresh rates and a faster response time, contributing to superior performance when displaying content with significant amounts of rapid motion (though newer LCD screens have similar refresh rates, but that also introduces the soap opera effect).

Disadvantages

- Picture quality
  - Earlier generation displays were more susceptible to screen burn-in and image retention, recent models have a pixel orbiter that moves the entire picture faster than is noticeable to the human eye, which reduces the effect of burn-in but does not prevent it.
  - Earlier generation displays (circa 2006 and prior) had phosphors that lost luminosity over time, resulting in gradual decline of absolute image

brightness (newer models may be less susceptible to this, having advertised life spans exceeding 100,000 hours, far longer than older CRT technology)
  - Screen-door effects are noticeable on screen sizes smaller than 127 cm (50 in); the effect is more visible at shorter viewing distances.
- Other
  - Use more electricity, on average, than an LCD TV.
  - Do not work as well at high altitudes due to pressure differential between the gases inside the screen and the air pressure at altitude. It may cause a buzzing noise. Manufacturers rate their screens to indicate the altitude parameters.
  - For those who wish to listen to AM radio, or are amateur radio operators (hams) or shortwave listeners (SWL), the radio frequency interference (RFI) from these devices can be irritating or disabling.
  - Due to the strong infrared emissions inherent with the technology, standard IR repeater systems cannot be used in the viewing room. A more expensive "plasma compatible" sensor must be used.
- Contrast ratio is the difference between the brightest and darkest parts of an image, measured in discrete steps, at any given moment.
- Generally, the higher the contrast ratio, the more realistic the image is (though the "realism" of an image depends on many factors including color accuracy, luminance linearity, and spatial linearity.