UNIT-I

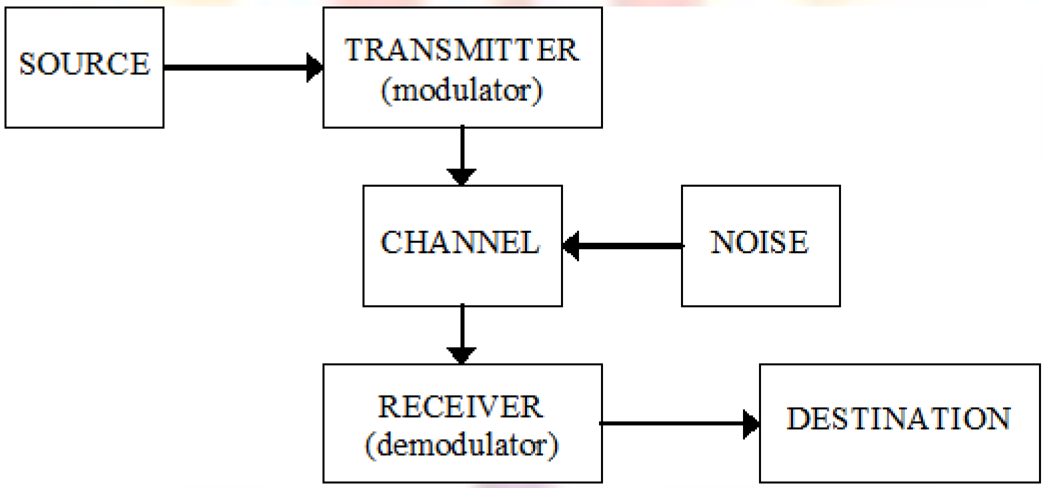
AMPLITUDE MODULATION

Introduction to Communication System

Communication is the process by which information is exchanged between individuals through a medium.

Communication can also be defined as the transfer of information from one point in space and time to another point.

The basic block diagram of a communication system is as follows.



- **Transmitter:** Couples the message into the channel using high frequency signals.
- **Channel:** The medium used for transmission of signals
- **Modulation:** It is the process of shifting the frequency spectrum of a signal to a frequency range in which more efficient transmission can be achieved.
- **Receiver:** Restores the signal to its original form.
- **Demodulation:** It is the process of shifting the frequency spectrum back to the original baseband frequency range and reconstructing the original form.

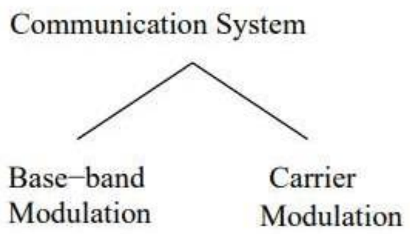
Modulation:

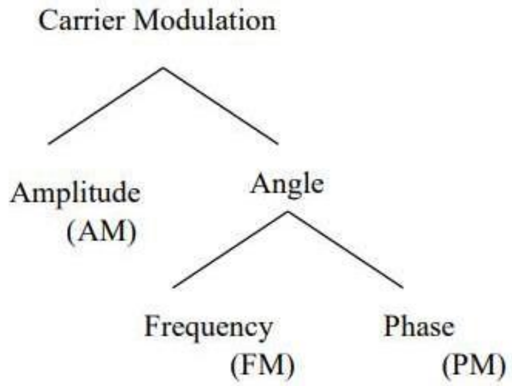
Modulation is a process that causes a shift in the range of frequencies in a signal.

• Signals that occupy the same range of frequencies can be separated.

• Modulation helps in noise immunity, attenuation - depends on the physical medium.

The below figure shows the different kinds of analog modulation schemes that are available





Modulation is operation performed at the transmitter to achieve efficient and reliable information transmission.

For analog modulation, it is frequency translation method caused by changing the appropriate quantity in a carrier signal.

It involves two waveforms:

- A modulating signal/baseband signal – represents the message.
- A carrier signal – depends on type of modulation.

•Once this information is received, the low frequency information must be removed from the high frequency carrier. •This process is known as "Demodulation".

Need for Modulation:

- Baseband signals are incompatible for direct transmission over the medium so, modulation is used to convey (baseband) signals from one place to another.
- Allows frequency translation:
  - Frequency Multiplexing
  - Reduce the antenna height
  - Avoids mixing of signals
  - Narrowbanding
- Efficient transmission
- Reduced noise and interference

Types of Modulation:

Three main types of modulations:

Analog Modulation

- **Amplitude modulation**
  Example: Double sideband with carrier (DSB-WC), Double- sideband suppressed carrier (DSB-SC), Single sideband suppressed carrier (SSB-SC), vestigial sideband (VSB)

- **Angle modulation (frequency modulationλ & phase modulation)**
  Example: Narrow band frequency modulation (NBFM), Widebandλ frequency modulation (WBFM), Narrowband phase modulation (NBPM), Wideband phase modulation (NBPM)

Pulse Modulation

- Carrier is a train of pulses
- Example: Pulse Amplitude Modulation (PAM), Pulse width modulation (PWM) , Pulse Position Modulation (PPM)

Digital Modulation

- Modulating signal is analog
  - Example: Pulse Code Modulation (PCM), Delta Modulationλ (DM), Adaptive Delta Modulation (ADM), Differential Pulse Code Modulation (DPCM), Adaptive Differential Pulse Code Modulation (ADPCM) etc.
- Modulating signal is digital (binary modulation)
  - Example: Amplitude shift keying (ASK), frequency Shift Keyingλ (FSK), Phase Shift Keying (PSK) etc

Amplitude Modulation (AM)

Amplitude Modulation is the process of changing the amplitude of a relatively high frequency carrier signal in accordance with the amplitude of the modulating signal (Information).

The carrier amplitude varied linearly by the modulating signal which usually consists of a range of audio frequencies. The frequency of the carrier is not affected.

- Application of AM - Radio broadcasting, TV pictures (video), facsimile transmission
- Frequency range for AM - 535 kHz – 1600 kHz
- Bandwidth - 10 kHz

Various forms of Amplitude Modulation

• Conventional Amplitude Modulation (Alternatively known as Full AM or Double Sideband Large carrier modulation (DSBLC) /Double Sideband Full Carrier (DSBFC)

• Double Sideband Suppressed carrier (DSBSC) modulation

• Single Sideband (SSB) modulation

• Vestigial Sideband (VSB) modulation

Time Domain and Frequency Domain Description

It is the process where, the amplitude of the carrier is varied proportional to that of the message signal.

Let m (t) be the base-band signal, m (t) ←→ M (ω) and c (t) be the carrier, $c(t) = A_c \cos(\omega_c t)$. fc is chosen such that fc >> W, where W is the maximum frequency component of m(t). The amplitude modulated signal is given by

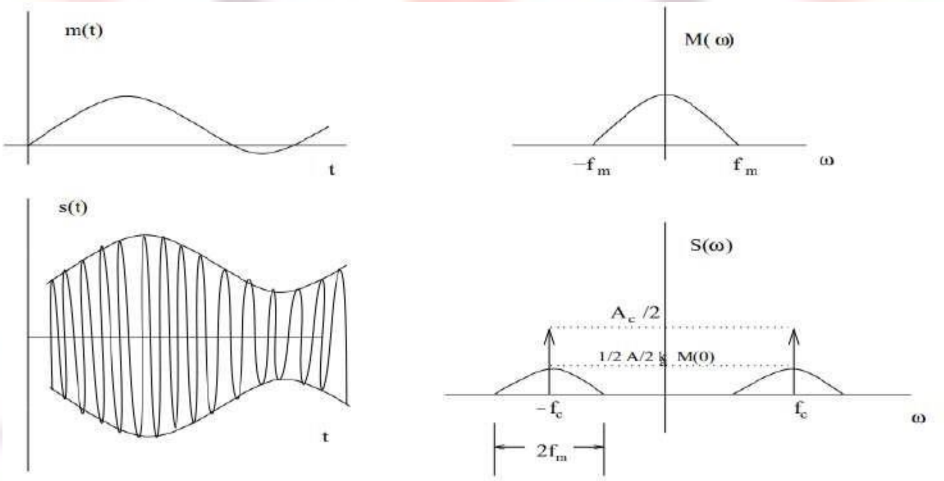
$$s(t) = Ac\,[1 + k_a m(t)]\cos(\omega c t)$$

Fourier Transform on both sides of the above equation

$$S(\omega) = \pi\, Ac/2\, (\delta(\omega - \omega c) + \delta(\omega + \omega c)) + k_a Ac/\,2\, (M(\omega - \omega c) + M(\omega + \omega c))$$

k_a is a constant called amplitude sensitivity.

$k_a m(t) < 1$ and it indicates percentage modulation.



Amplitude modulation in time and frequency domain

Single Tone Modulation:

Consider a modulating wave m(t) that consists of a single tone or single frequency component given by

$$m(t) = A_m \cos(2\pi f_m t) \quad \text{.......................(1)}$$

where A_m is peak amplitude of the sinusoidal modulating wave

f_m is the frequency of the sinusoidal modulating wave

Let A_c be the peak amplitude and f_c be the frequency of the high frequency carrier signal. Then the corresponding single-tone AM wave is given by

$$s(t) = A_c[1 + m\cos(2\pi f_m t)]Cos(2\pi f_c t) \quad \text{.......................(2)}$$

Let $A_{\max}$ and $A_{\min}$ denote the maximum and minimum values of the envelope of the modulated wave. Then from the above equation (2.12), we get

$$\frac{A_{\max}}{A_{\min}} = \frac{A_c(1+m)}{A_c(1-m)}$$

$$m = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}}$$

Expanding the equation (2), we get

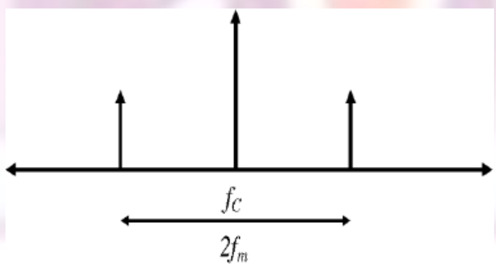
$$s(t) = A_c \cos(2\pi f_c t) + \frac{1}{2} m A_c \cos[2\pi(f_c + f_m)t] + \frac{1}{2} m A_c \cos[2\pi(f_c - f_m)t]$$

The Fourier transform of $s(t)$ is obtained as follows.

$$s(f) = \frac{1}{2} A_c[\delta(f - f_c) + \delta(f + f_c)] + \frac{1}{4} m A_c[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)]$$

$$+ \frac{1}{4} m A_c[\delta(f - f_c + f_m) + \delta(f + f_c - f_m)]$$

Thus the spectrum of an AM wave, for the special case of sinusoidal modulation consists of delta functions at $\pm f_c$, $f_c \pm f_m$, and $-f_c \pm f_m$. The spectrum for positive frequencies is as shown in figure



Frequency Domain characteristics of single tone AM

Power relations in AM waves:

Consider the expression for single tone/sinusoidal AM wave

$$s(t) = A_c Cos(2\pi f_c t) + \frac{1}{2} mA_c Cos[2\pi(f_c + f_m)t] + \frac{1}{2} mA_c Cos[2\pi(f_c - f_m)t] \quad \text{..............(1)}$$

This expression contains three components. They are carrier component, upper side band and lower side band. Therefore Average power of the AM wave is sum of these three components.

Therefore the total power in the amplitude modulated wave is given by

$$Pt = \frac{V^2{}_{car}}{R} + \frac{V^2{}_{LSB}}{R} + \frac{V^2{}_{USB}}{R} \quad \text{.....................(2)}$$

Where all the voltages are rms values and R is the resistance, in which the power is dissipated.

$$P_C = \frac{V^2{}_{car}}{R} = \frac{\left(A_c/\sqrt{2}\right)^2}{R} = \frac{A_c{}^2}{2R}$$

$$P_{LSB} = \frac{V_{LSB}{}^2}{R} = \left(\frac{mA_c}{2\sqrt{2}}\right)^2 \frac{1}{R} = \frac{m^2 A_c{}^2}{8R} = \frac{m^2}{4} P_c$$

$$P_{USB} = \frac{V_{USB}{}^2}{R} = \left(\frac{mA_c}{2\sqrt{2}}\right)^2 \frac{1}{R} = \frac{m^2 A_c{}^2}{8R} = \frac{m^2}{4} P_c$$

Therefore total average power is given by

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

$$P_t = P_c + \frac{m^2}{4} P_c + \frac{m^2}{4} P_c$$

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

$$P_t = P_c\left(1 + \frac{m^2}{2}\right) \quad \text{..............................(3)}$$

The ratio of total side band power to the total power in the modulated wave is given by

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

$$\frac{P_{SB}}{P_t} = \frac{m^2}{2+m^2} \qquad \text{.........................(4)}$$

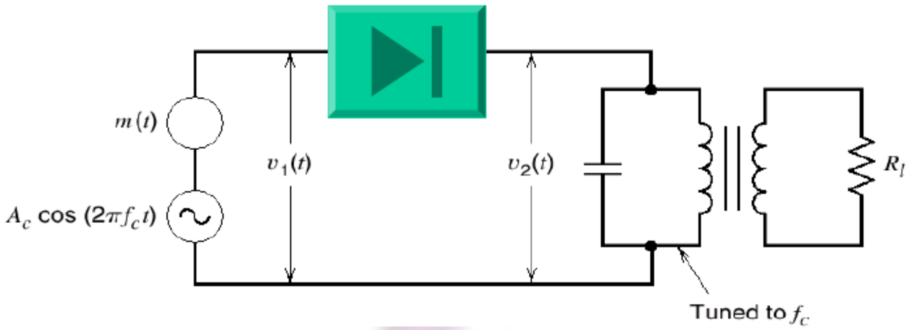
This ratio is called the efficiency of AM system

Generation of AM waves:

Two basic amplitude modulation principles are discussed. They are square law modulation and switching modulator.

Switching Modulator

Consider a semiconductor diode used as an ideal switch to which the carrier signal $c(t) = A_c \cos(2\pi f_c t)$ and information signal $m(t)$ are applied simultaneously as shown figure



Switching Modulator

The total input for the diode at any instant is given by

$$v_1 = c(t) + m(t)$$

$$v_1 = A_c \cos 2\pi f_c t + m(t)$$

When the peak amplitude of c(t) is maintained more than that of information signal, the operation is assumed to be dependent on only c(t) irrespective of m(t).

When c(t) is positive, v2=v1since the diode is forward biased. Similarly, when c(t) is negative, v2=0 since diode is reverse biased. Based upon above operation, switching response of the diode is periodic rectangular wave with an amplitude unity and is given by

$$p(t) = \frac{1}{2} + \frac{1}{\pi}\sum_{n=-\infty}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos(2\pi f_c t(2n-1))$$

$$p(t) = \frac{1}{2} + \underbrace{\frac{2}{\pi}}_{n=0,1}\cos(2\pi f_c t) - \underbrace{\frac{2}{3\pi}}_{n=-1,+2}\cos(6\pi f_c t) + -$$

Therefore the diode response V_o is a product of switching response *p(t)* and input v_1.

$$v_2 = v_1 * p(t)$$

$$V_2 = [A_c \cos 2\pi f_c t + m(t)]\left[\frac{1}{2} + \frac{2}{\pi}\cos 2\pi f_c t - \frac{2}{3\pi}\cos 6\pi f_c t + - + -\right]$$

Applying the Fourier Transform, we get

$$V_2(f) = \frac{A_c}{4}[\delta(f - f_c) + \delta(f + f_c)] + \frac{M(f)}{2} + \frac{A_c}{\pi}\delta(f)$$
$$+ \frac{A_c}{2\pi}[\delta(f - 2f_c) + \delta(f + 2f_c)] + \frac{1}{\pi}[M(f - f_c) + M(f + f_c)]$$
$$- \frac{A_c}{6\pi}[\delta(f - 4f_c) + \delta(f + 4f_c)] - \frac{A_c}{3\pi}[\delta(f - 2f_c) + \delta(f + 2f_c)]$$
$$- \frac{1}{3\pi}[M(f - 3f_c) + M(f + f_c)]$$

The diode output v_2 consists of

a dc component at $f = 0$.

Information signal ranging from 0 to *w* Hz and infinite number of frequency bands centered at f, $2f_c$, $3f_c$, $4f_c$, ---------

The required AM signal centred at fc can be separated using band pass filter. The lower cut off-frequency for the band pass filter should be between w and fc-w and the upper cut-off frequency between fc+w and 2fc. The filter output is given by the equation

$$S(t) = \frac{A_c}{2}\left[1 + \frac{4}{\pi}\frac{m(t)}{A_c}\right]\cos 2\pi f_c t$$

For a single tone information, let $m(t) = A_m \cos(2\pi f_m t)$

$$S(t) = \frac{A_c}{2}\left[1 + \frac{4}{\pi}\frac{A_m}{A_c}\cos 2\pi f_m t\right]\cos 2\pi f_c t$$

Therefore modulation index, $m = \frac{4}{\pi}\frac{A_m}{A_c}$

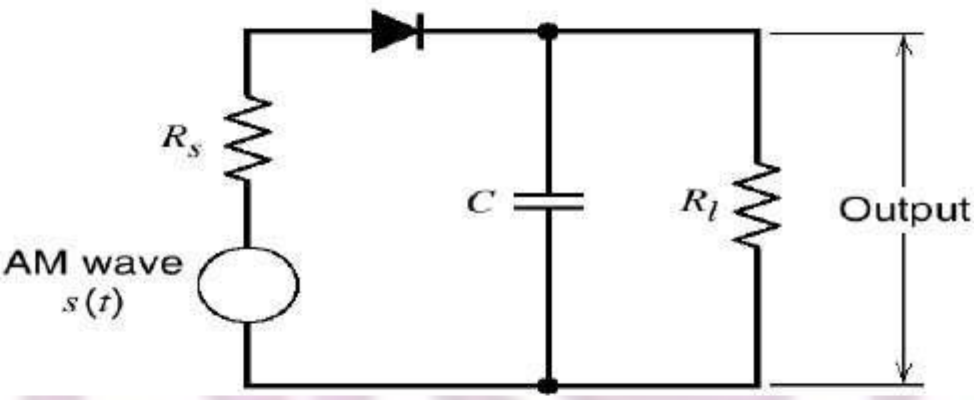
The output AM signal is free from distortions and attenuations only when $fc\text{-}w > w$ or $fc > 2w$.

Detection of AM waves

Demodulation is the process of recovering the information signal (base band) from the incoming modulated signal at the receiver. There are two methods, they are Square law Detector and Envelope Detector

Envelope Detector

It is a simple and highly effective system. This method is used in most of the commercial AM radio receivers. An envelope detector is as shown below.



Envelope Detector

During the positive half cycles of the input signals, the diode D is forward biased and the capacitor C charges up rapidly to the peak of the input signal. When the input signal falls below this value, the diode becomes reverse biased and the capacitor C discharges through the load resistor RL.

The discharge process continues until the next positive half cycle. When the input signal becomes greater than the voltage across the capacitor, the diode conducts again and the process is repeated.

The charge time constant (rf+Rs)C must be short compared with the carrier period, the capacitor charges rapidly and there by follows the applied voltage up to the positive peak when the diode is conducting.That is the charging time constant shall satisfy the condition,

$$\left(r_f + R_s\right)C << \frac{1}{f_c}$$

On the other hand, the discharging time-constant $R_L C$ must be long enough to ensure that the capacitor discharges slowly through the load resistor R_L between the positive peaks of the carrier wave, but not so long that the capacitor voltage will not discharge at the maximum rate of change of the modulating wave.

That is the discharge time constant shall satisfy the condition,

$$\frac{1}{f_c} << R_L C << \frac{1}{W}$$

Where 'W' is band width of the message signal. The result is that the capacitor voltage or detector output is nearly the same as the envelope of AM wave.

Advantages and Disadvantages of AM:

Advantages of AM:

- Generation and demodulation of AM wave are easy.
- AM systems are cost effective and easy to build.

Disadvantages:

- AM contains unwanted carrier component, hence it requires more transmission power.
- The transmission bandwidth is equal to twice the message bandwidth.

To overcome these limitations, the conventional AM system is modified at the cost of increased system complexity. Therefore, three types of modified AM systems are discussed.

DSBSC (Double Side Band Suppressed Carrier) modulation:

In DSBC modulation, the modulated wave consists of only the upper and lower side bands. Transmitted power is saved through the suppression of the carrier wave, but the channel bandwidth requirement is the same as before.

SSBSC (Single Side Band Suppressed Carrier) modulation: The SSBSC modulated wave consists of only the upper side band or lower side band. SSBSC is suited for transmission of voice signals. It is an optimum form of modulation in that it requires the minimum transmission power and minimum channel band width. Disadvantage is increased cost and complexity.

VSB (Vestigial Side Band) modulation: In VSB, one side band is completely passed and just a trace or vestige of the other side band is retained. The required channel bandwidth is therefore in excess of the message bandwidth by an amount equal to the width of the vestigial side band. This method is suitable for the transmission of wide band signals.

DSB-SC MODULATION

DSB-SC Time domain and Frequency domain Description:

DSBSC modulators make use of the multiplying action in which the modulating signal multiplies the carrier wave. In this system, the carrier component is eliminated and both upper and lower side bands are transmitted. As the carrier component is suppressed, the power required for transmission is less than that of AM.

If $m(t)$ is the message signal and $c(t) = A_c \cos(2\pi f_c t)$ is the carrier signal, then DSBSC modulated wave $s(t)$ is given by

$$s(t) = c(t)\, m(t)$$

$$s(t) = A_c \cos(2\pi f_c t) m(t)$$

Consequently, the modulated signal s(t) under goes a phase reversal , whenever the message signal m(t) crosses zero as shown below.

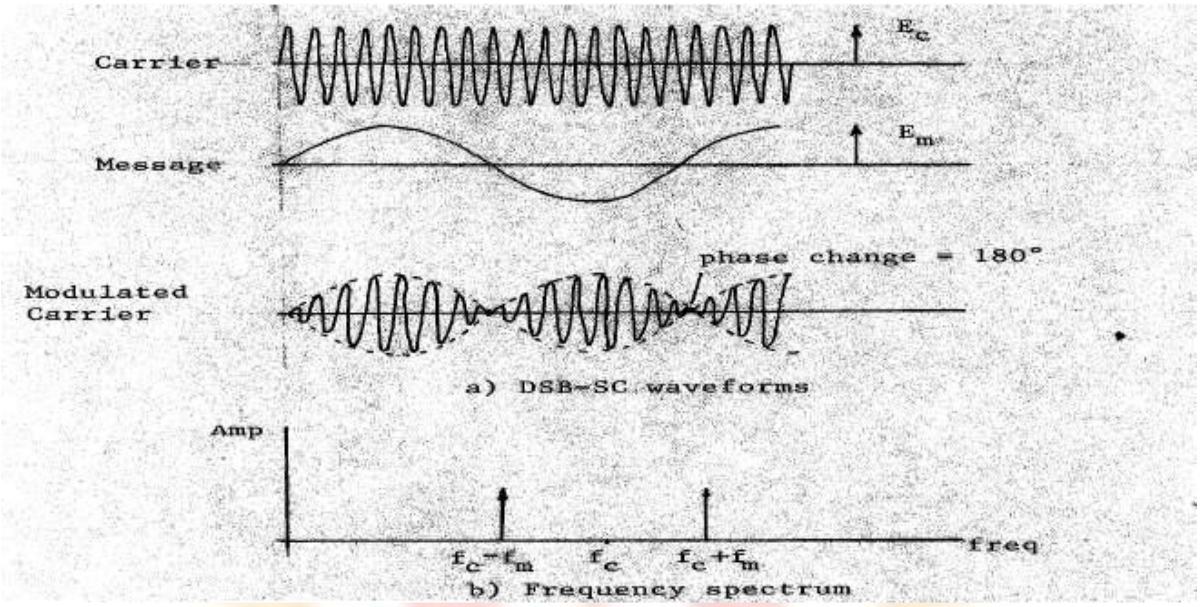


Fig.1. (a) DSB-SC waveform (b) DSB-SC Frequency Spectrum

The envelope of a DSBSC modulated signal is therefore different from the message signal and the Fourier transform of s(t) is given by

$$S(f) = \frac{A_c}{2}\left[M(f - f_c) + M(f + f_c)\right]$$

For the case when base band signal *m(t)* is limited to the interval $-W<f<W$ as shown in figure below, we find that the spectrum *S(f)* of the DSBSC wave *s(t)* is as illustrated below. Except for a change in scaling factor, the modulation process simply translates the spectrum of the base band signal by f_c. The transmission bandwidth required by DSBSC modulation is the same as that for AM.

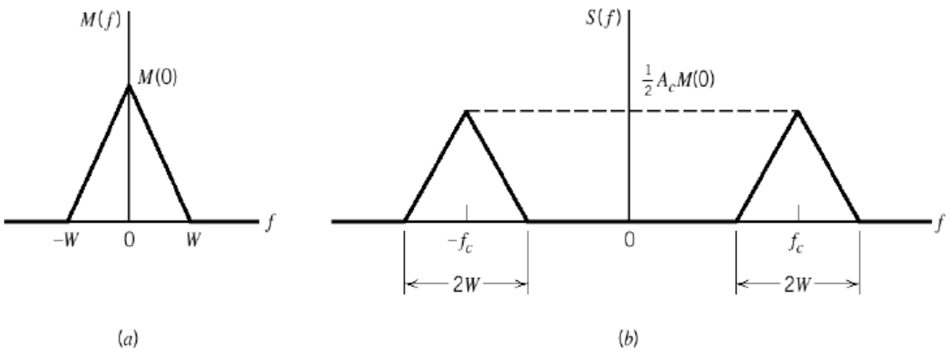


Figure: Message and the corresponding DSBSC spectrum

Generation of DSBSC Waves:

Balanced Modulator (Product Modulator)

A balanced modulator consists of two standard amplitude modulators arranged in a balanced configuration so as to suppress the carrier wave as shown in the following block diagram. It is assumed that the AM modulators are identical, except for the sign reversal of the modulating wave applied to the input of one of them. Thus, the output of the two modulators may be expressed as,

$$s_1(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t)$$
$$s_2(t) = A_c[1 - k_a m(t)]\cos(2\pi f_c t)$$

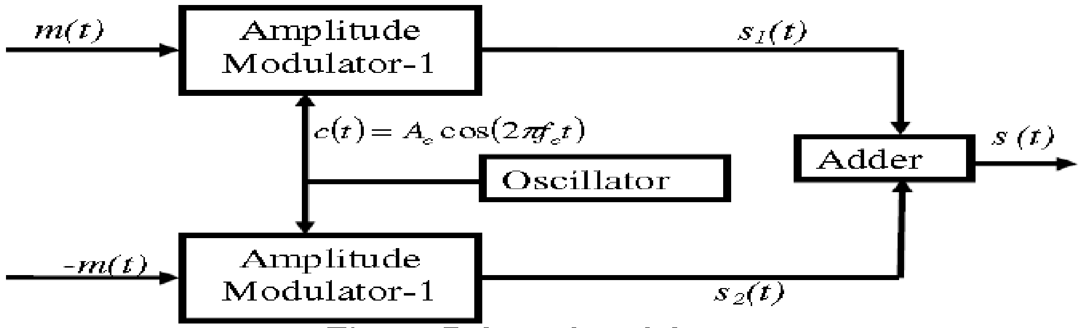


Fig 3 : Balanced modulator

Subtracting $s_2(t)$ from $s_1(t)$,

$$s(t) = s_1(t) - s_2(t)$$
$$s(t) = 2k_a m(t) A_c \cos(2\pi f_c t)$$

Hence, except for the scaling factor 2ka, the balanced modulator output is equal to the product of the modulating wave and the carrier.

Ring Modulator

Ring modulator is the most widely used product modulator for generating DSBSC wave and is shown below.

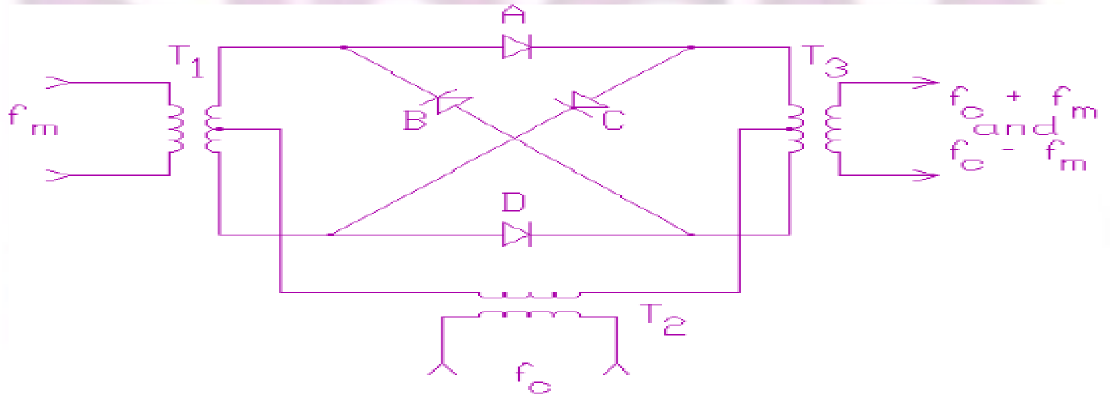


Fig.4 : Ring modulator

The four diodes form a ring in which they all point in the same direction. The diodes are controlled by square wave carrier c(t) of frequency fc, which is applied longitudinally by means of two center-tapped transformers. Assuming the diodes are ideal, when the carrier is positive, the outer diodes D1 and D2 are forward biased where as the inner diodes D3 and D4 are reverse biased, so that the modulator multiplies the base band signal m(t) by c(t). When the carrier is negative, the diodes D1 and D2 are reverse biased and D3 and D4 are forward, and the modulator multiplies the base band signal –m(t) by c(t).

Thus the ring modulator in its ideal form is a product modulator for square wave carrier and the base band signal m(t). The square wave carrier can be expanded using Fourier series as

$$c(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos(2\pi f_c t(2n-1))$$

Therefore the ring modulator out put is given by

$$s(t) = m(t)c(t)$$

$$s(t) = m(t)\left[\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos(2\pi f_c t(2n-1))\right]$$

From the above equation it is clear that output from the modulator consists entirely of modulation products. If the message signal m(t) is band limited to the frequency band − w < f < w, the output spectrum consists of side bands centred at fc.

Detection of DSB-SC waves:

Coherent Detection:

The message signal m(t) can be uniquely recovered from a DSBSC wave s(t) by first multiplying s(t) with a locally generated sinusoidal wave and then low pass filtering the product as shown.

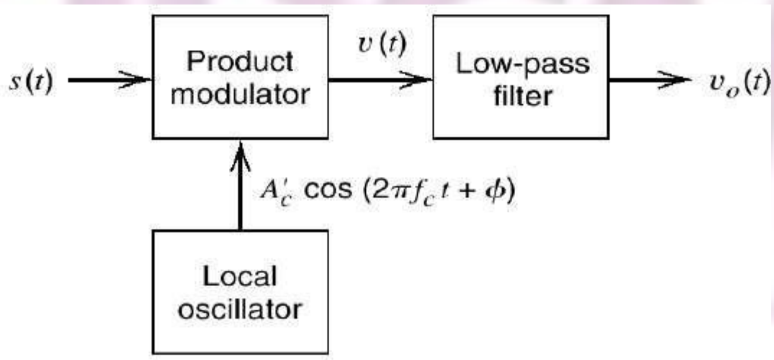


Fig.5 : Coherent detector

It is assumed that the local oscillator signal is exactly coherent or synchronized, in both frequency and phase, with the carrier wave c(t) used in the product modulator to generate s(t). This method of demodulation is known as coherent detection or synchronous detection.

Let $A_c^1\cos(2\pi f_c t+\phi)$ be the local oscillator signal, and $s(t)=A_c\cos(2\pi f_c t)m(t)$ be the DSBSC wave. Then the product modulator output $v(t)$ is given by

$$v(t)=A_cA_c^1\cos(2\pi f_c t)\cos(2\pi f_c t+\phi)m(t)$$

$$v(t)=\frac{A_cA_c^1}{4}\cos(4\pi f_c t+\phi)m(t)+\frac{A_cA_c^1}{2}\cos(\phi)m(t)$$

The first term in the above expression represents a DSBSC modulated signal with a carrier frequency $2f_c$, and the second term represents the scaled version of message signal. Assuming that the message signal is band limited to the interval $-w<f<w$, the spectrum of $v(t)$ is plotted as shown below.

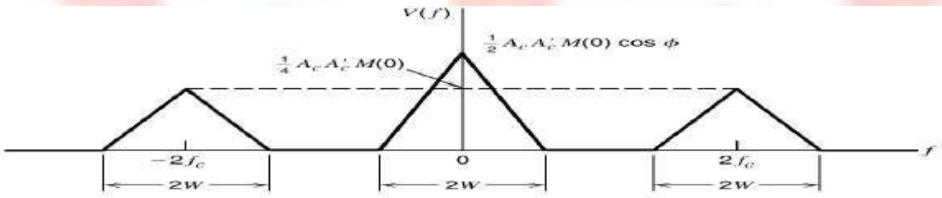


Fig.6.Spectrum of output of the product modulator

From the spectrum, it is clear that the unwanted component (first term in the expression) can be removed by the low-pass filter, provided that the cut-off frequency of the filter is greater than W but less than 2fc-W. The filter output is given by

$$v_o(t)=\frac{A_cA_c^1}{2}\cos(\phi)m(t)$$

The demodulated signal $v_o(t)$ is therefore proportional to m(t) when the phase error ϕ is constant.

Costas Receiver (Costas Loop):

Costas receiver is a synchronous receiver system, suitable for demodulating DSBSC waves. It consists of two coherent detectors supplied with the same input signal,

that is the incoming DSBSC wave $s(t)=A_c\cos(2\pi f_c t)m(t)$ but with individual local oscillator signals that are in phase quadrature with respect to each other as shown below.

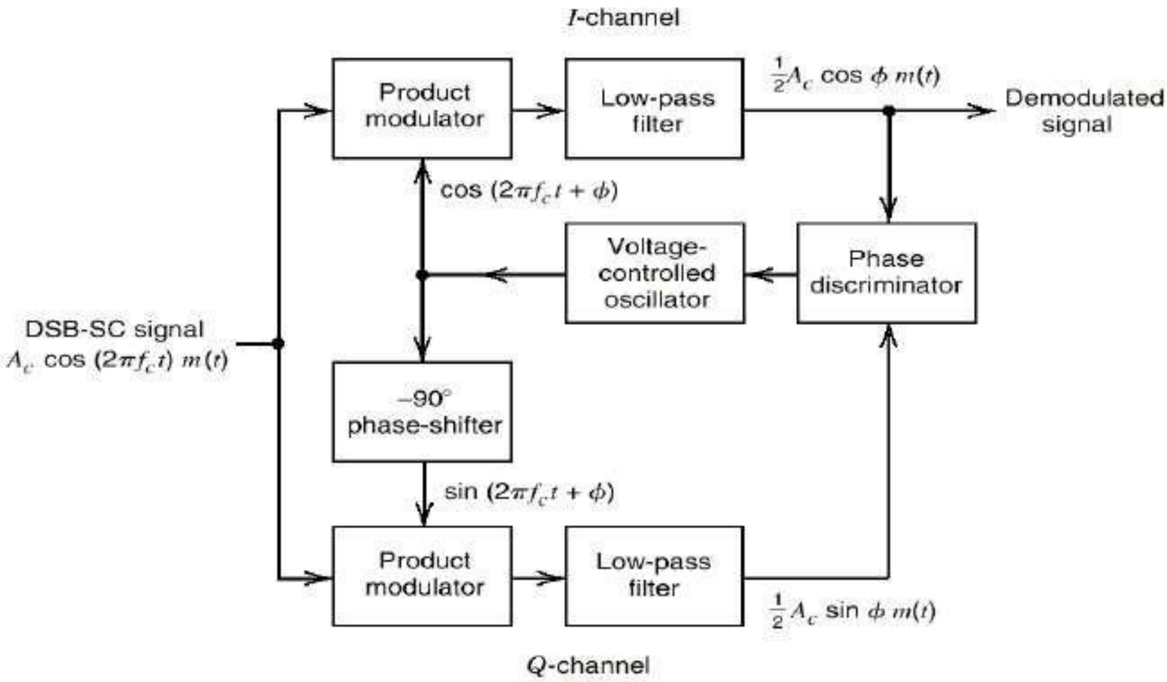


Fig.7. Costas Receiver

The frequency of the local oscillator is adjusted to be the same as the carrier frequency fc. The detector in the upper path is referred to as the in-phase coherent detector or I-channel, and that in the lower path is referred to as the quadrature-phase coherent detector or Q-channel.

These two detector are coupled together to form a negative feedback system designed in such a way as to maintain the local oscillator synchronous with the carrier wave. Suppose the local oscillator signal is of the same phase as the carrier $c(t) = A_c\cos(2\pi f_c t)$ wave used to generate the incoming DSBSC wave. Then we find that the I-channel output contains the desired demodulated signal m(t), where as the Q-channel output is zero due to quadrature null effect of the Q-channel. Suppose that the local oscillator phase drifts from its proper value by a small angle ϕ radians. The I-channel output will remain essentially unchanged, but there will be some signal appearing at the Q-channel output, which is proportional to $\sin(\phi) \approx \phi$ for small ϕ.

This Q-channel output will have same polarity as the I-channel output for one direction of local oscillator phase drift and opposite polarity for the opposite direction of local oscillator phase drift. Thus by combining the I-channel and Q-channel outputs in a phase discriminator (which consists of a multiplier followed by a LPF), a dc control signal is obtained that automatically corrects for the local phase errors in the voltage-controlled oscillator.

Introduction of SSB-SC

Standard AM and DSBSC require transmission bandwidth equal to twice the message bandwidth. In both the cases spectrum contains two side bands of width W Hz, each. But the upper and lower sides are uniquely related to each other by the virtue of their symmetry about the carrier frequency. That is, given the amplitude and phase spectra of either side band, the other can be uniquely determined. Thus if only one side band is transmitted, and if both the carrier and the other side band are suppressed at the transmitter, no information is lost. This kind of modulation is called SSBSC and spectral comparison between DSBSC and SSBSC is shown in the figures 1 and 2.

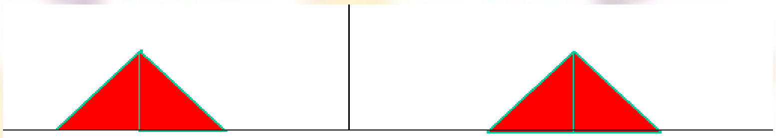


Figure.1 : Spectrum of the DSBSC wave

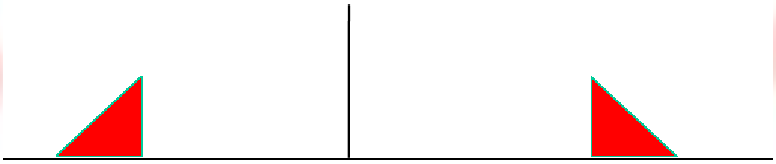


Figure .2 : Spectrum of the SSBSC wave

Frequency Domain Description

Consider a message signal *m(t)* with a spectrum *M(f)* band limited to the interval $-w < f < w$ as shown in figure 3 , the DSBSC wave obtained by multiplexing *m(t)* by the carrier wave $c(t) = A_c \cos(2\pi f_c t)$ and is also shown, in figure 4 . The upper side band is represented in duplicate by the frequencies above f_c and those below $-f_c$, and when only upper

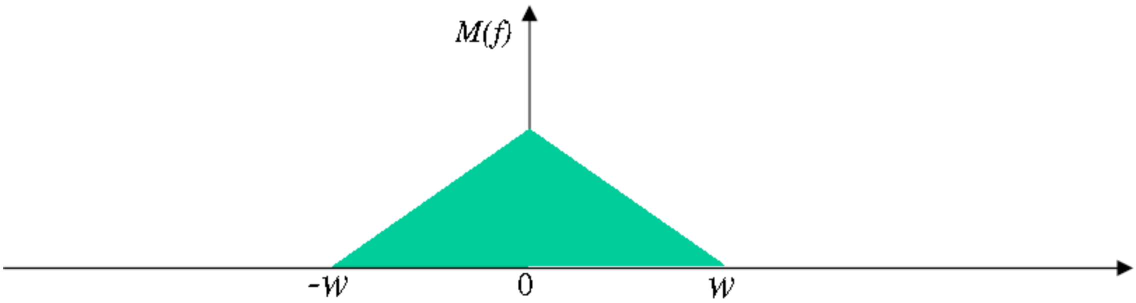


Figure 3. : Spectrum of message wave

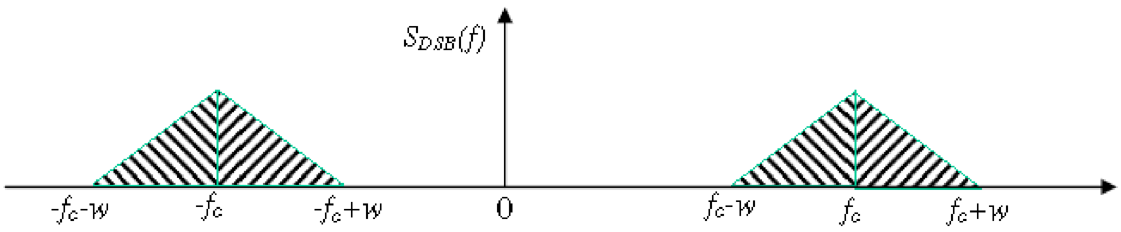


Figure .4 : Spectrum of DSBSC wave

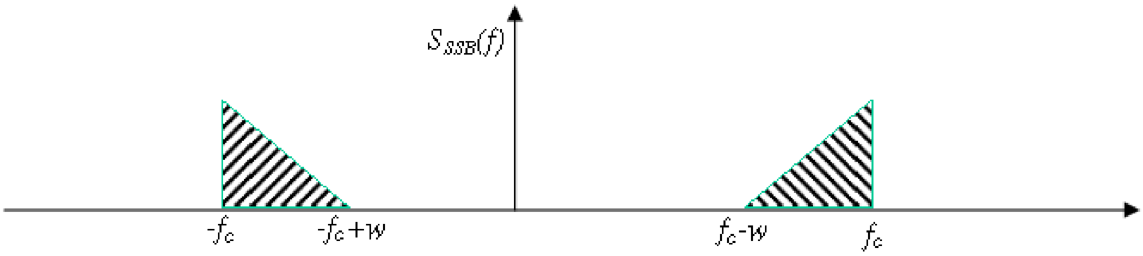


Figure.5 : Spectrum of SSBSC-LSB wave

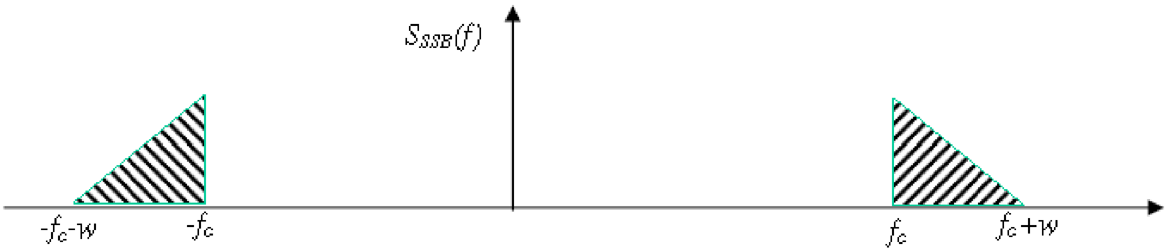


Figure .6 : Spectrum of SSBSC-USB wave

side band is transmitted; the resulting SSB modulated wave has the spectrum shown in figure 6. Similarly, the lower side band is represented in duplicate by the frequencies below fc and those above -fc and when only the lower side band is transmitted, the spectrum of the corresponding SSB modulated wave shown in figure 5.Thus the essential function of the SSB modulation is to translate the spectrum of the modulating wave, either with or without inversion, to a new location in the frequency domain. The advantage of SSB modulation is reduced bandwidth and the elimination of high power carrier wave. The main disadvantage is the cost and complexity of its implementation.

Generation of SSB wave:

Frequency discrimination method

Consider the generation of SSB modulated signal containing the upper side band only. From a practical point of view, the most severe requirement of SSB generation arises from the unwanted sideband, the nearest component of which is separated from the desired side band by twice the lowest frequency component of the message signal. It implies that, for the generation of an SSB wave to be possible, the message spectrum must have an energy gap centered at the origin as shown in figure 7. This requirement is naturally satisfied by voice signals, whose energy gap is about 600Hz wide.

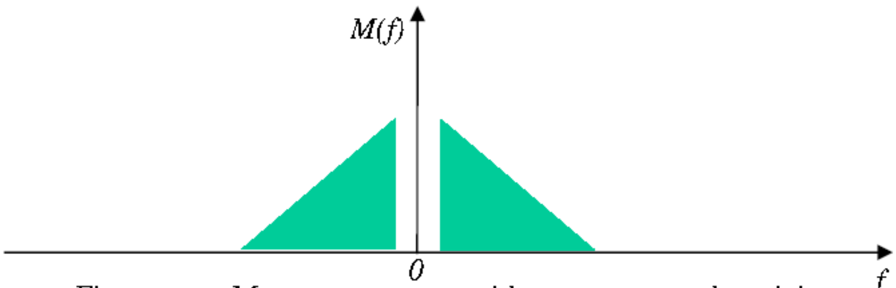


Figure .7 : Message spectrum with energy gap at the origin

The frequency discrimination or filter method of SSB generation consists of a product modulator, which produces DSBSC signal and a band-pass filter to extract the desired side band and reject the other and is shown in the figure 8.

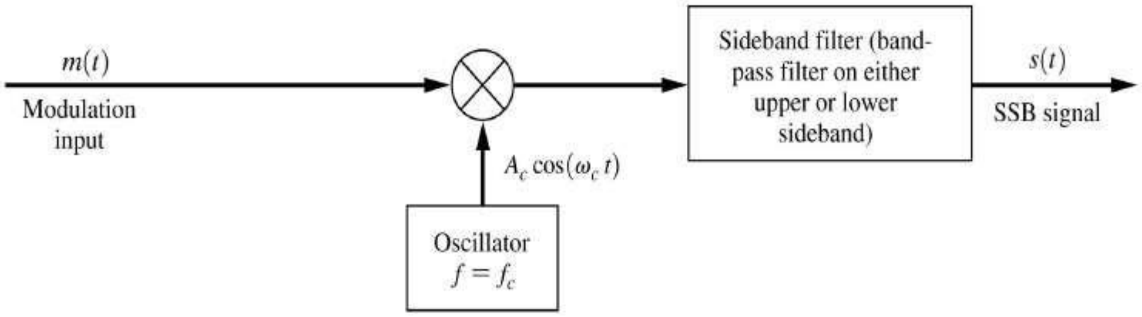


Figure .8 : Frequency discriminator to generate SSBSC wave

Application of this method requires that the message signal satisfies two conditions:
1. The message signal m(t) has no low-frequency content. Example: speech, audio, music.
2. The highest frequency component W of the message signal m(t) is much less than the carrier frequency fc.

Then, under these conditions, the desired side band will appear in a non-overlapping interval in the spectrum in such a way that it may be selected by an appropriate filter.

In designing the band pass filter, the following requirements should be satisfied:

1.The pass band of the filter occupies the same frequency range as the spectrum of the desired SSB modulated wave.

2. The width of the guard band of the filter, separating the pass band from the stop band, where the unwanted sideband of the filter input lies, is twice the lowest frequency component of the message signal.

When it is necessary to generate an SSB modulated wave occupying a frequency band that is much higher than that of the message signal, it becomes very difficult to design an appropriate filter that will pass the desired side band and reject the other. In such a situation it is necessary to resort to a multiple-modulation process so as to ease the filtering requirement. This approach is illustrated in the following figure 9 involving two stages of modulation.

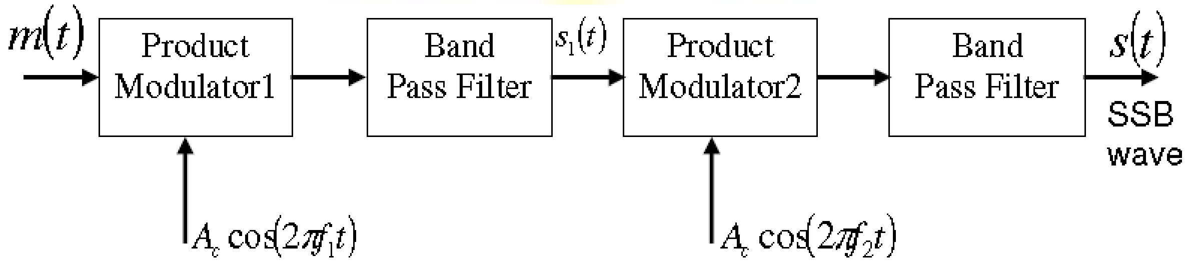


Figure .9 : Two stage frequency discriminator

The SSB modulated wave at the first filter output is used as the modulating wave for the second product modulator, which produces a DSBSC modulated wave with a spectrum that is symmetrically spaced about the second carrier frequency f2. The frequency separation between the side bands of this DSBSC modulated wave is effectively twice the first carrier frequency f1, thereby permitting the second filter to remove the unwanted side band.

Time Domain Description:

The time domain description of an SSB wave s(t) in the canonical form is given by the equation 1.

$$s(t) = s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t) \quad \text{-------------------- (1)}$$

where $S_I(t)$ is the in-phase component of the SSB wave and $S_Q(t)$ is its quadrature component. The in-phase component $S_I(t)$ except for a scaling factor, may be derived from $S(t)$ by first multiplying $S(t)$ by $\cos(2\pi f_c t)$ and then passing the product through a low-pass filter. Similarly, the quadrature component $S_Q(t)$, except for a scaling factor, may be derived from $s(t)$ by first multiplying $s(t)$ by $\sin(2\pi f_c t)$ and then passing the product through an identical filter.

The Fourier transformation of $S_I(t)$ and $S_Q(t)$ are related to that of SSB wave as follows, respectively.

$$S_I(f) = \begin{cases} S(f - f_c) + S(f + f_c), -w \le f \le w \\ 0, elsewhere \end{cases} \quad \text{------------------- (2)}$$

$$S_Q(f) = \begin{cases} j[S(f-f_c) - S(f+f_c)], -w \le f \le w \\ 0, elsewhere \end{cases} \quad \text{------------------ (3)}$$

where $-w < f < w$ defines the frequency band occupied by the message signal $m(t)$.

Consider the SSB wave that is obtained by transmitting only the upper side band, shown in figure 10 . Two frequency shifted spectras $S(f-f_c)$ and $S(f+f_c)$ are shown in figure 11 and figure 12 respectively. Therefore, from equations 2 and 3 , it follows that the corresponding spectra of the in- phase component $S_I(t)$ and the quadrature component $S_Q(t)$ are as shown in figure 13 and 14 respectively.

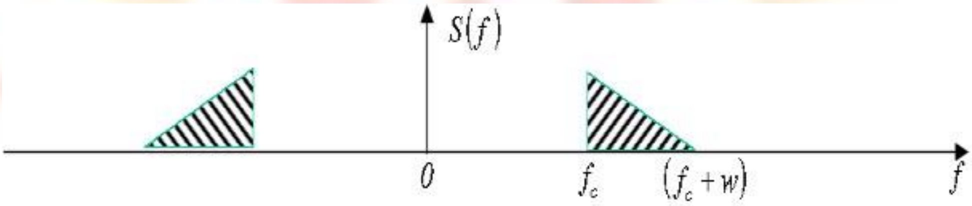


Figure 10 : Spectrum of SSBSC-USB

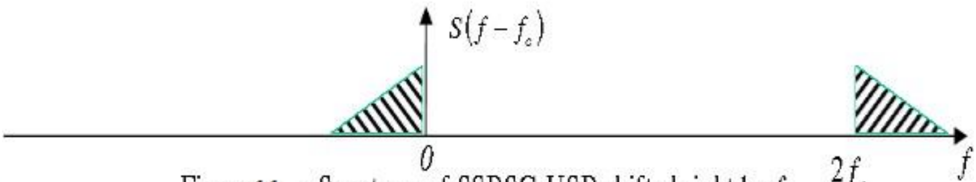


Figure 11 : Spectrum of SSBSC-USB shifted right by f_c

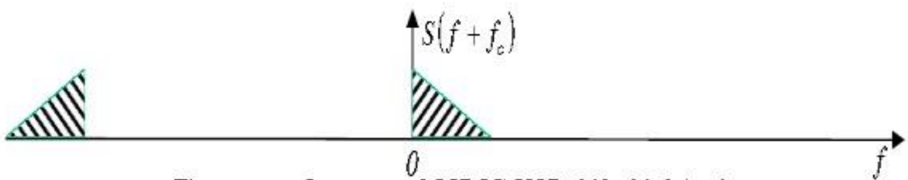


Figure 12 : Spectrum of SSBSC-USB shifted left by f_c

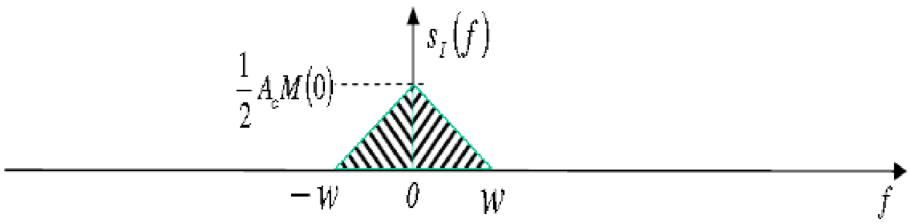


Figure 13 : Spectrum of in-phase component of SSBSC-USB

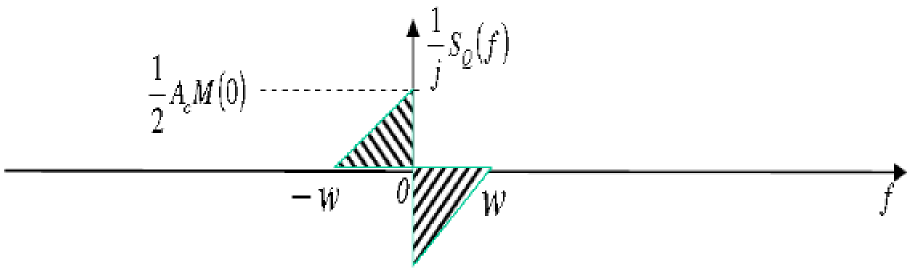


Figure 14 : Spectrum of quadrature component of SSBSC-USB

From the figure 13 , it is found that

$$s_I(f) = \frac{1}{2}A_c M(f)$$

where *M(f)* is the Fourier transform of the message signal *m(t)*. Accordingly in-phase component $S_I(t)$ is defined by equation 4

$$\boxed{s_I(t) = \frac{1}{2}A_c m(t)} \qquad \text{------------------ (4)}$$

Now on the basis of figure14 , it is found that

$$S_Q(f) = \begin{cases} \frac{-j}{2}A_c M(f), f > 0 \\ 0, f = 0 \\ \frac{j}{2}A_c M(f), f < 0 \end{cases}$$

$$S_Q(f) = \frac{-j}{2}A_c \operatorname{sgn}(f) M(f) \qquad \text{------------------ (5)}$$

where $\operatorname{sgn}(f)$ is the Signum function.

But from the discussions on Hilbert transforms, it is shown that

$$-j\,\mathrm{sgn}(f)M(f)=\hat{M}(f) \qquad \text{------------------ (6)}$$

where $\hat{M}(f)$ is the Fourier transform of the Hilbert transform of $m(t)$. Hence the substituting equation (6) in (5), we get

$$S_Q(f)=\frac{1}{2}A_c\hat{M}(f) \qquad \text{------------------------ (7)}$$

Therefore quadrature component $s_Q(t)$ is defined by equation 8

$$\boxed{s_Q(t)=\frac{1}{2}A_c\hat{m}(t)} \qquad \text{------------------------- (8)}$$

Therefore substituting equations (4) and (8) in equation in (1), we find that canonical representation of an SSB wave $s(t)$ obtained by transmitting only the upper side band is given by the equation 9

$$\boxed{s_U(t)=\frac{1}{2}A_c m(t)\cos(2\pi f_c t)-\frac{1}{2}A_c\hat{m}(t)\sin(2\pi f_c t)} \qquad \text{------------- (9)}$$

Following the same procedure, we can find the canonical representation for an SSB wave
s(t) obtained by transmitting only the lower side band is given by

$$s_L(t)=\frac{1}{2}A_c m(t)\cos(2\pi f_c t)+\frac{1}{2}A_c\hat{m}(t)\sin(2\pi f_c t) \qquad \text{-------------- (10)}$$

Phase discrimination method for generating SSB wave:

Time domain description of SSB modulation leads to another method of SSB generation using the equations 9 or 10. The block diagram of phase discriminator is as shown in figure 15.

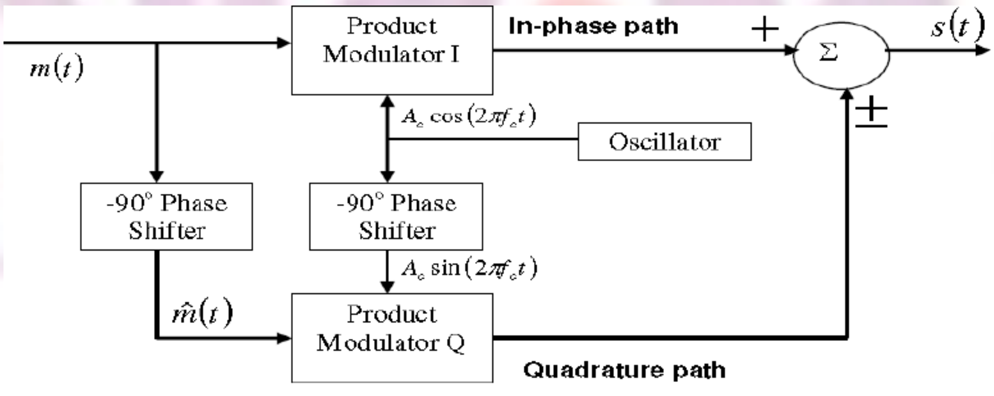


Figure 15 : Block diagram of phase discriminator

The phase discriminator consists of two product modulators I and Q, supplied with carrier waves in-phase quadrature to each other. The incoming base band signal m(t) is applied to product modulator I, producing a DSBSC modulated wave that contains reference phase sidebands symmetrically spaced about carrier frequency fc.

The Hilbert transform mˆ (t) of m (t) is applied to product modulator Q, producing a DSBSC modulated that contains side bands having identical amplitude spectra to those of modulator I, but with phase spectra such that vector addition or subtraction of the two modulator outputs results in cancellation of one set of side bands and reinforcement of the other set.

The use of a plus sign at the summing junction yields an SSB wave with only the lower side band, whereas the use of a minus sign yields an SSB wave with only the upper side band. This modulator circuit is called Hartley modulator.

Demodulation of SSB Waves:

Demodulation of SSBSC wave using coherent detection is as shown in 16 . The SSB wave $s(t)$ together with a locally generated carrier $c(t) = A_c^1 \cos(2\pi f_c^{\prime} t + \phi)$ is applied to a product modulator and then low-pass filtering of the modulator output yields the message signal.

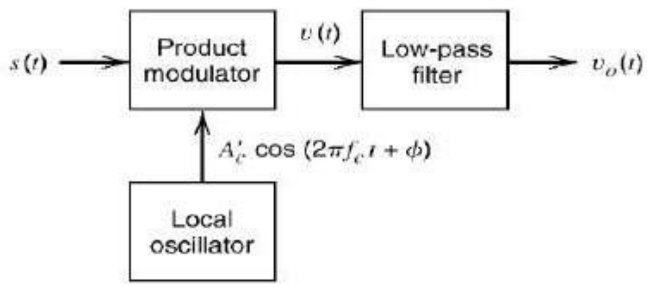


Figure 16 : Block diagram of coherent detector for SSBSC

The product modulator output $v(t)$ is given by

$$v(t) = A_c^1 \cos(2\pi f_c t + \phi)s(t) \quad , \quad \text{Put } \phi = 0$$

$$v(t) = \frac{1}{2} A_c \cos(2\pi f_c t)\left[m(t)\cos(2\pi f_c t) \pm \hat{m}(t)\sin(2\pi f_c t)\right]$$

$$v(t) = \frac{1}{4} A_c m(t) + \frac{1}{4} A_c\left[m(t)\cos(4\pi f_c t) \pm \hat{m}(t)\sin(4\pi f_c t)\right] \quad \text{..............(1)}$$

The first term in the above equation 1 is desired message signal. The other term represents an SSB wave with a carrier frequency of $2f_c$ as such; it is an unwanted component, which is removed by low-pass filter.

Introduction to Vestigial Side Band Modulation

Vestigial sideband is a type of Amplitude modulation in which one side band is completely passed along with trace or tail or vestige of the other side band. VSB is a compromise between SSB and DSBSC modulation. In SSB, we send only one side band, the Bandwidth required to send SSB wave is w. SSB is not appropriate way of modulation when the message signal contains significant components at extremely low frequencies. To overcome this VSB is used.

Vestigial Side Band (VSB) modulation is another form of an amplitude-modulated signal in which a part of the unwanted sideband (called as vestige, hence the name vestigial sideband) is allowed to appear at the output of VSB transmission system.

The AM signal is passed through a sideband filter before the transmission of SSB signal. The design of sideband filter can be simplified to a greater extent if a part of the other sideband is also passed through it. However, in this process the bandwidth of VSB system is slightly increased.

Generation of VSB Modulated Signal

VSB signal is generated by first generating a DSB-SC signal and then passing it through a sideband filter which will pass the wanted sideband and a part of unwanted sideband. Thus, VSB is so called because a vestige is added to SSB spectrum.

The below figure depicts functional block diagram of generating VSB modulated signal

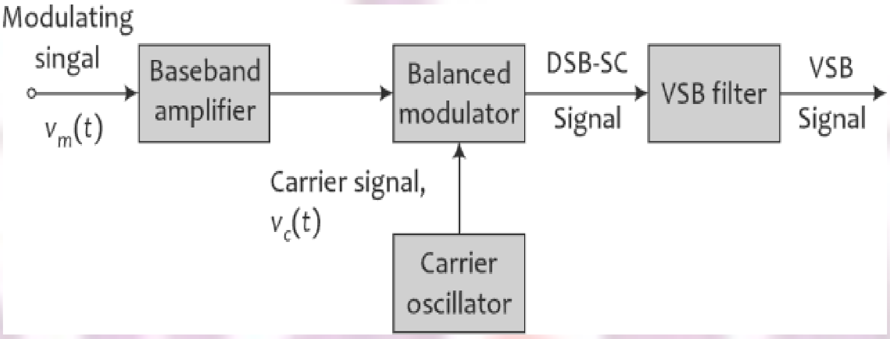


Figure: Generation of VSB Modulated Signal

A VSB-modulated signal is generated using the frequency discrimination method, in which firstly a DSB-SC modulated signal is generated and then passed through a sideband-suppression filter. This type of filter is a specially-designed bandpass filter that distinguishes VSB modulation from SSB modulation.the cutoff portion of the frequency response of this filter around the carrier frequency exhibits odd symmetry, that is, (fc-fv)≤|f|≤(fc+fv).

Accordingly the bandwidth of the VSB signal is given as

$$BW = (f_m + f_v) \text{ Hz}$$

Where fm is the bandwidth of the modulating signal or USB, and fv is the bandwidth of vestigial sideband (VSB)

Time domain description of VSB Signal

Mathematically, the VSB modulated signal can be described in the time-domain as

$$s(t)= m(t)\ Ac \cos(2\pi fct) \pm m_Q(t)\ Ac \sin(2\pi fct)$$

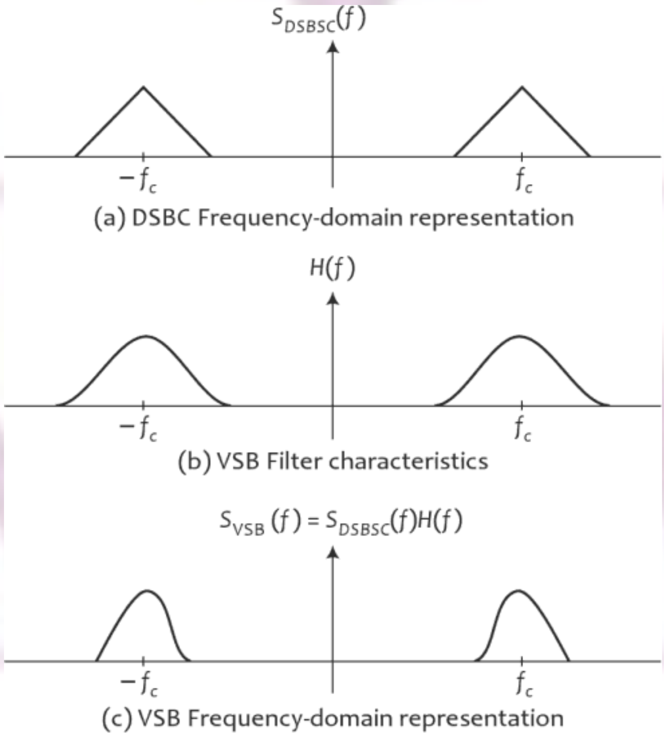
where m(t) is the modulating signal, $m_Q(t)$ is the component of m(t) obtained by passing the message signal through a vestigial filter, Ac cos(2πfct) is the carrier signal, and Ac sin(2πfct) is the 90^o phase shift version of the carrier signal.

The ± sign in the expression corresponds to the transmission of a vestige of the upper-sideband and lower-sideband respectively. The Quadrature component is required to partially reduce power in one of the sidebands of the modulated wave s(t) and retain a vestige of the other sideband as required.

Frequency domain representation of VSB Signal

Since VSB modulated signal includes a vestige (or trace) of the second sideband, only a part of the second sideband is retained instead of completely eliminating it. Therefore, VSB signal can be generated from DSB signal followed by VSB filter which is a practical filter.

The below figure shows the DSB signal spectrum, the VSB filter characteristics, and the resulting output VSB modulated signal spectrum.



(a) DSBC Frequency-domain representation
(b) VSB Filter characteristics
(c) VSB Frequency-domain representation

Bandwidth Consideration in TV Signals

An important application of VSB modulation technique is in broadcast television. In commercial TV broadcasting system, there is a basic need to conserve bandwidth.

- The upper-sideband of the video carrier signal is transmitted upto 4MHz without any attenuation.
- The lower-sideband of the video carrier signal is transmitted without any attenuation over the range 0.75 MHz (Double side band transmission) and is entirely attenuated at 1.25MHz (single sideband transmission) and the transition is made from one o another between 0.75MHz and 1.25 MHz (thus the name vestige sideband)
- The audio signal which accompanies the video signal is transmitted by frequency modulation method using a carrier signal located 4.5 MHz above the video-carrier signal.
- The audio signal is frequency modulated on a separate carrier signal with a frequency deviation of 25 KHz. With an audio bandwidth of 10 KHz, the deviation ratio is 2.5 and an FM bandwidth of approximately 70 KHz.
- The frequency range of 100 KHz is allowed on each side of the audio-carrier signal for the audio sidebands.
- One sideband of the video-modulated signal is attenuated so that it does not interfere with the lower- sideband of the audio carrier.

Advantages of VSB Modulation

VSB transmission system has several advantages which include

- Use of simple filter design
- Less bandwidth as compared to that of DSBSC signal
- As efficient as SSB
- Possibility of transmission of low frequency components of modulating signals

Facts to Know

VSB is mainly used as a standard modulation technique for transmission of video signals in TV signals in commercial television broadcasting because the modulating video signal has large bandwidth and high speed data transmission

Envelope detection of a VSB Wave plus Carrier

To make demodulation of VSB wave possible by an envelope detector at the receiving end it is necessary to transmit a sizeable carrier together with the modulated wave. The scaled expression of VSB wave by factor k_a with the carrier component $A_c \cos(2\pi f_c t)$ can be given by

$$s(t) = A_c \cos(2\pi f_c t) + \frac{A_c}{2} k_a\, m(t) \cos(2\pi f_c t)$$

$$- \frac{A_c}{2} k_a\, m_Q(t) \sin(2\pi f_c t)$$

$$= A_c \left[1 + \frac{k_a}{2} m(t)\right] \cos(2\pi f_c t) - \frac{A_c k_a}{2} m_Q(t) \sin(2\pi f_c t) \quad \text{.................(1)}$$

where k_a is the modulation index; it determines the percentage modulation.

When above signal s(t) is passed through the envelope detector, the detector output is given by,

$$a(t) = A_c \left[\left(1 + \frac{k_a}{2} m(t)\right)^2 + \left(\frac{k_a}{2} m_Q(t)\right)^2\right]^{1/2}$$

$$= A_c \left[1 + \frac{k_a}{2} m(t)\right] \left(1 + \left[\frac{\frac{1}{2} k_a\, m_Q(t)}{1 + \frac{1}{2} k_a\, m(t)}\right]^2\right)^{1/2} \quad \text{.................(2)}$$

The detector output is distorted by the quadrature component $m_Q(t)$ as indicated by equation (2).

Methods to reduce distortion

- Distortion can be reduced by reducing percentage modulation, k_a.
- Distortion can be reduced by reducing $m_Q(t)$ by increasing the width of the vestigial sideband.

Comparison of AM Techniques:

Sr. No.	Parameter	Standard AM	SSB	DSBSC	VSB
1	Power	High	Less	Medium	Less than DSBSC but greater than SSB
2	Bandwidth	$2 f_m$	f_m	$2 f_m$	$f_m < B_w < 2 f_m$
3	Carrier supression	No	Yes	Yes	No
4	Receiver complexity	Simple	Complex	Complex	Simple
5	Application	Radio communication	Point to point communication preferred for long distance transmission.	Point to point communication	Television broadcasting
6	Modulation type	Non linear	Linear	Linear	Linear
7	Sideband suppression	No	One sided completely	No	One sideband suppressed partly
8	Transmission efficiency	Minimum	Maximum	Moderate	Moderate

Applications of different AM systems:

- Amplitude Modulation: AM radio, Short wave radio broadcast
- DSB-SC: Data Modems, Color TV's color signals.
- SSB: Telephone
- VSB: TV picture signals

UNIT-II

ANGLE MODULATION

Introduction

There are two forms of angle modulation that may be distinguished – phase modulation and frequency modulation

Basic Definitions: Phase Modulation (PM) and Frequency Modulation (FM)

Let $\theta_i(t)$ denote the angle of modulated sinusoidal carrier, which is a function of the message. The resulting angle-modulated wave is expressed as

$$s(t) = A_c cos[\theta_i(t)] \ldots\ldots\ldots\ldots\ldots (1)$$

Where A_c is the carrier amplitude. A complete oscillation occurs whenever $\theta_i(t)$ changes by 2π radians. If $\theta_i(t)$ increases monotonically with time, the average frequency in Hz, over an interval from t to $t+\Delta t$, is given by

$$f_{\Delta t}(t) = \frac{\theta_i(t+\Delta t) - \theta_i(t)}{2\pi\Delta t} \ldots\ldots\ldots\ldots\ldots (2)$$

Thus the instantaneous frequency of the angle-modulated wave $s(t)$ is defined as

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

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

$$f_i(t) = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt} \ldots\ldots\ldots\ldots\ldots (3)$$

Thus, according to equation (1), the angle modulated wave s(t) is interpreted as a rotating Phasor of length Ac and angle $\theta_i(t)$. The angular velocity of such a Phasor is $d\theta_i(t)/dt$, in accordance with equ (3).In the simple case of an unmodulated carrier, the angle $\theta_i(t)$ is

$$\theta_i(t) = 2\pi f_c t + \emptyset_c$$

And the corresponding Phasor rotates with a constant angular velocity equal to $2\pi f_c$.The constant ϕ_c is the value of $\theta_i(t)$ at t=0.

There are an infinite number of ways in which the angle $\theta_i(t)$ may be varied in some manner with the baseband signal.

But the 2 commonly used methods are **Phase modulation** and **Frequency modulation.**

Phase Modulation (PM) is that form of angle modulation in which the angle $\theta_i(t)$ is varied linearly with the baseband signal m(t), as shown by

$$\theta_i(t) = 2\pi f_c t + k_p m(t) \ldots\ldots\ldots\ldots\ldots\ldots..(4)$$

The term $2\pi f_c t$ represents the angle of the unmodulated carrier, and the constant k_p represents the ***phase sensitivity*** of the modulator, expressed in radians per volt.

The phase-modulated wave *s(t)* is thus described in time domain by

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)] \ldots\ldots\ldots\ldots\ldots (5)$$

Frequency Modulation (FM) is that form of angle modulation in which the instantaneous frequency $f_i(t)$ is varied linearly with the baseband signal *m(t)*, as shown by

$$f_i(t) = f_c + k_f m(t) \ldots\ldots\ldots\ldots\ldots\ldots.(6)$$

The term f_c represents the frequency of the unmodulated carrier, and the constant k_f represents the ***frequency sensitivity*** of the modulator, expressed in hertz per volt.

Integrating equ.(6) with respect to time and multiplying the result by 2π, we get

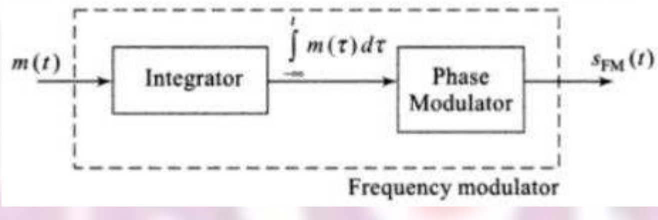
$$\theta_i(t) = 2\pi f_c t + 2\pi k_f \int_0^t m(t)\ dt \ldots\ldots\ldots\ldots\ldots.(7)$$

Where, for convenience it is assumed that the angle of the unmodulated carrier wave is zero at t=0. The frequency modulated wave is therefore described in the time domain by

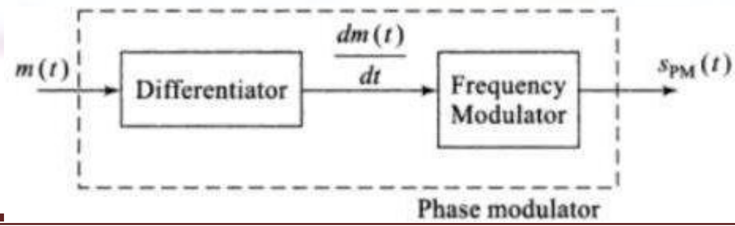
$$s(t) = A_c \cos[\,2\pi f_c t + 2\pi k_f \int_0^t m(t)\ dt] \ldots\ldots\ldots\ldots\ldots\ldots..(8)$$

Relationship between PM and FM

Comparing equ (5) with (8) reveals that an FM wave may be regarded as a PM wave in which the modulating wave is $\int_0^t m(t)dt$ in place of *m(t)*.



A PM wave can be generated by first differentiating *m(t)* and then using the result as the input to a frequency modulator.



Thus the properties of PM wave can be deduced from those of FM waves and vice versa

Single tone Frequency modulation

Consider a sinusoidal modulating wave defined by

$$\boldsymbol{m(t) = A_m \cos(2\pi f_m t)} \ldots\ldots\ldots\ldots\ldots (\mathbf{1})$$

The instantaneous frequency of the resulting FM wave is

$$\boldsymbol{f_i(t) = f_c + k_f A_m \cos(2\pi f_m t)}$$

$$\boldsymbol{f_i(t) = f_c + \Delta f \cos(2\pi f_m t)} \ldots\ldots\ldots\ldots..(\mathbf{2})$$

Where $\boldsymbol{\Delta f = k_f A_m}$ .... $(\mathbf{3})$

The quantity Δf is called the ***frequency deviation***, representing the maximum departure of the instantaneous frequency of the FM wave from the carrier frequency f_c.

Fundamental characteristic of an FM wave is that the frequency deviation Δf is proportional to the amplitude of the modulating wave, and is independent of the modulation frequency.

Using equation (2), the angle $\theta_i(t)$ of the FM wave is obtained as

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

$$\boldsymbol{\theta_i(t) = 2\pi f_c t + \frac{\Delta f}{f_m} sin(2\pi f_m t)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (\mathbf{4})$$

The ratio of the frequency deviation Δf *to* the modulation frequency f_m is commonly called the ***modulation index*** of the FM wave. Modulation index is denoted by $\boldsymbol{\beta}$ and is given as

$$\mathrm{Q} = \boldsymbol{\frac{\Delta f}{f_m}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots.(\mathbf{5})$$

And

$$\boldsymbol{\theta_i(t) = 2\pi f_c t + \mathrm{Q} sin(2\pi f_m t)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (\mathbf{6})$$

In equation (6) the parameter $\boldsymbol{\beta}$ represents the phase deviation of the FM wave, that is, the maximum departure of the angle $\theta_i(t)$ from the angle $\boldsymbol{2\pi f_c t}$ of the unmodulated carrier.

The FM wave itself is given by

$$\boldsymbol{s(t) = A_c \cos[2\pi f_c t + \mathrm{Q} sin(2\pi f_m t)]} \ldots\ldots\ldots\ldots\ldots\ldots..(\mathbf{7})$$

Depending on the value of modulation index β, we may distinguish two cases of frequency modulation. Narrow-band FM for which β is small and Wide-band FM for which β is large, both compared to one radian.

Narrow-Band Frequency modulation

Consider the Single tone FM wave

$$s(t) = A_c \cos[2\pi f_c t + \text{Q}sin(2\pi f_m t)] \quad \ldots\ldots\ldots\ldots(1)$$

Expanding this relation we get

$$s(t) = A_c \cos(2\pi f_c t)\cos[\beta \sin(2\pi f_m t)] - A_c \sin(2\pi f_c t)\sin[\beta \sin(2\pi f_m t)] \quad \textbf{.......(2)}$$

Assuming that the modulation index β is small compared to one radian, we may use the following approximations:

$$\cos[\beta \sin(2\pi f_m t)] \simeq 1$$

and

$$\sin[\beta \sin(2\pi f_m t)] \simeq \beta \sin(2\pi f_m t)$$

Hence, Equation (2) simplifies to

$$s(t) \simeq A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t)\sin(2\pi f_m t) \quad \textbf{..................(3)}$$

Equation (3) defines the approximate form of a narrowband FM signal produced by a sinusoidal modulating signal $A_m \cos(2\pi f_m t)$. From this representation we deduce the modulator shown in block diagram form in Figure . This modulator involves splitting the carrier wave $A_c \cos(2\pi f_c t)$ into two paths. One path is direct; the other path contains a -90 degree phase-shifting network and a product modulator, the combination of which generates a DSB-SC modulated signal. The difference between these two signals produces a narrowband FM signal, but with some distortion.

Ideally, an FM signal has a constant envelope and, for the case of a sinusoidal modulating signal of frequency f_m, the angle $\theta_i(t)$ is also sinusoidal with the same frequency.

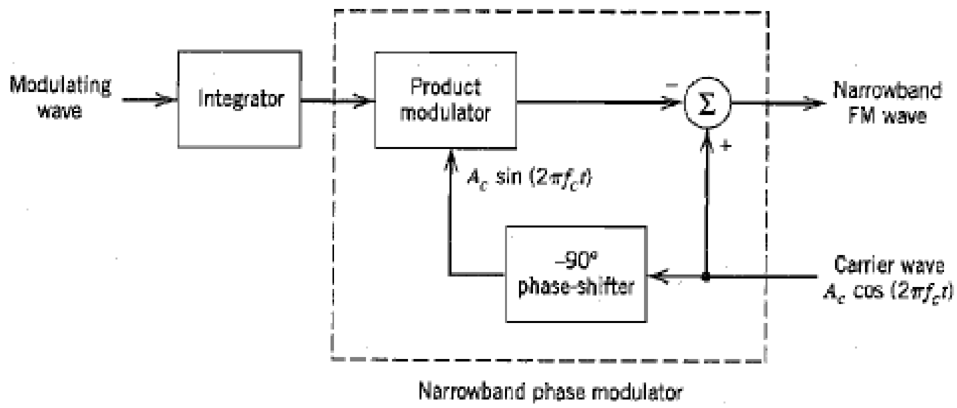


Figure Block diagram of a method for generating a narrowband FM signal.

But the modulated signal produced by the narrowband modulator of Figure differs from this ideal condition in two fundamental respects:

1. The envelope contains a *residual* amplitude modulation and, therefore, varies with time.
2. For a sinusoidal modulating wave, the angle $\theta_i(t)$ contains *harmonic distortion* in the form of third- and higher-order harmonics of the modulation frequency f_m.

However, by restricting the modulation index to $\beta \leq 0.3$ radians, the effects of residual AM and harmonic PM are limited to negligible levels.

Returning to Equation (3), we may expand it as follows:

$$s(t) = A_c \cos(2\pi f_c t) + \frac{1}{2}\beta A_c\{\cos[2\pi(f_c + f_m)t] - \cos[2\pi(f_c - f_m)t]\} \text{ ....(4)}$$

This expression is somewhat similar to the corresponding one defining an AM signal, which is as follows:

$$s_{AM}(t) = A_c \cos(2\pi f_c t) + \frac{1}{2}\mu A_c\{\cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t]\} \text{ ....(5)}$$

where μ is the modulation factor of the AM signal. Comparing Equations (4) and (5), we see that in the case of sinusoidal modulation, the basic difference between an AM signal and a narrowband FM signal is that the algebraic sign of the lower side frequency in the narrowband FM is reversed. Thus, a narrowband FM signal requires essentially the same transmission bandwidth (i.e., $2f_m$) as the AM signal.

We may represent the narrowband FM signal with a phasor diagram as shown in Figure a , where we have used the carrier phasor as reference. We see that the resultant

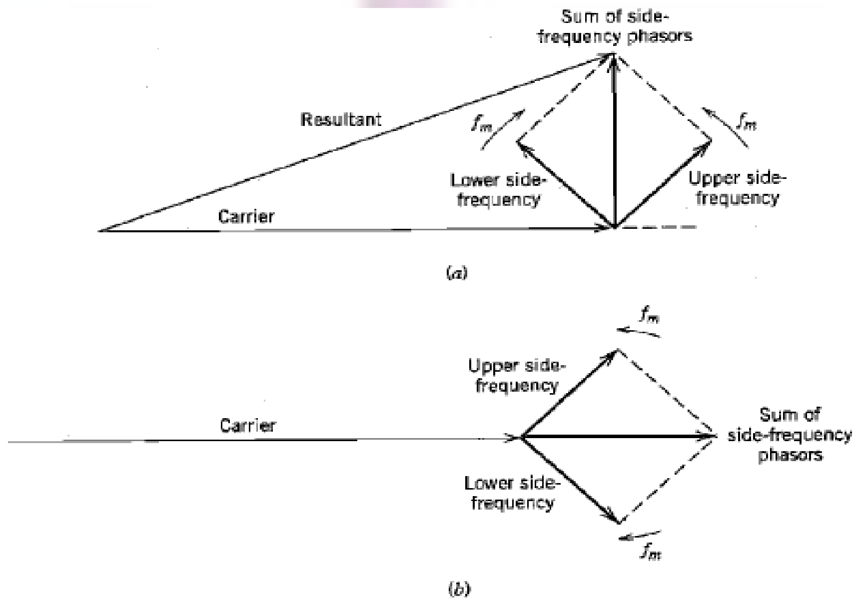


FIGURE A phasor comparison of narrowband FM and AM waves for sinusoidal modulation. (*a*) Narrowband FM wave. (*b*) AM wave.

of the two side-frequency phasors is always at right angles to the carrier phasor. The effect of this is to produce a resultant phasor representing the narrowband FM signal that is approximately of the same amplitude as the carrier phasor, but out of phase with respect to it. This phasor diagram should be contrasted with that of Figure *(b)* , representing an AM signal. In this latter case we see that the resultant phasor representing the AM signal has an amplitude that is different from that of the carrier phasor but always in phase with it.

Wide band frequency Modulation

The spectrum of the signle-tone FM wave of equation

$$\boldsymbol{s(t) = A_c \cos[2\pi f_c t + \mathrm{Q} sin(2\pi f_m t)]} \; \dots\dots\dots\dots (\mathbf{1})$$

For an arbitrary vale of the modulation index Q is to be determined.

An FM wave produced by a sinusoidal modulating wave as in equation (1) is by itself nonperiodic, unless the carrier frequency f_c is an integral multiple of the modualtion frequency f_m. Rewriting the equation in the form

$$\begin{aligned} s(t) &= \mathrm{Re}[A_c \exp(j2\pi f_c t + j\beta \sin(2\pi f_m t))] \\ &= \mathrm{Re}[\tilde{s}(t) \exp(j2\pi f_c t)] \qquad \mathbf{..........(2)} \end{aligned}$$

where $\tilde{s}(t)$ is the *complex envelope* of the FM signal $s(t)$, defined by

$$\tilde{s}(t) = A_c \exp[j\beta \sin(2\pi f_m t)] \qquad \mathbf{...............(3)}$$

$\tilde{s}(t)$ is periodic function of time,with a fundamental frequency equal to the modulation frequency f_m. $\tilde{s}(t)$ in the form of complex Fourier series is as follows

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_m t) \qquad \mathbf{................(4)}$$

where the complex Fourier coefficient c_n is defined by

$$\begin{aligned} c_n &= f_m \int_{-1/2f_m}^{1/2f_m} \tilde{s}(t) \exp(-j2\pi n f_m t)\, dt \\ &= f_m A_c \int_{-1/2f_m}^{1/2f_m} \exp[j\beta \sin(2\pi f_m t) - j2\pi n f_m t]\, dt \qquad \mathbf{............(5)} \end{aligned}$$

Define a new variable:

$$x = 2\pi f_m t \qquad \mathbf{.....................(6)}$$

Hence, we may rewrite Equation (**5**) in the new form

$$c_n = \frac{A_c}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin x - nx)]\, dx \qquad \mathbf{..............(7)}$$

The integral on the RHS of equation (7) is recognizedasthe nth order Bessel Function of the first kind and argument Q. This function is commonly denoted by the symbol J_n(Q), that is

$$J_n(\beta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp[j(\beta \sin x - nx)]\, dx \qquad ....(8)$$

Accordingly, we may reduce Equation (7) to

$$c_n = A_c J_n(\beta) \qquad ....(9)$$

Substituting Equation (9) in (5), we get, in terms of the Bessel function $J_n(\beta)$, the following expansion for the complex envelope of the FM signal:

$$\tilde{s}(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \exp(j2\pi n f_m t) \qquad ....(10)$$

Next, substituting Equation (10) in (2), we get

$$s(t) = A_c \cdot \mathrm{Re}\left[\sum_{n=-\infty}^{\infty} J_n(\beta) \exp[j2\pi(f_c + nf_m)t]\right] \qquad ....(11)$$

Interchanging the order of summation and evaluation of the real part in the right-hand side of Equation (11), we finally get

$$s(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + nf_m)t] \qquad ....(12)$$

Equ. (12) is the Fourier series representation of the single-tone FM wave *s(t)* for an arbitrary value of Q.

The discrete spectrum of *s(t)* is obtained by taking the Fourier transform of both sides of equation (12); thus

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta)[\delta(f - f_c - nf_m) + \delta(f + f_c + nf_m)] \;.............(13)$$

In the figure below, we have plotted the Bessel function J_n(Q) versus the modulation index Q for different positive integer value of n.

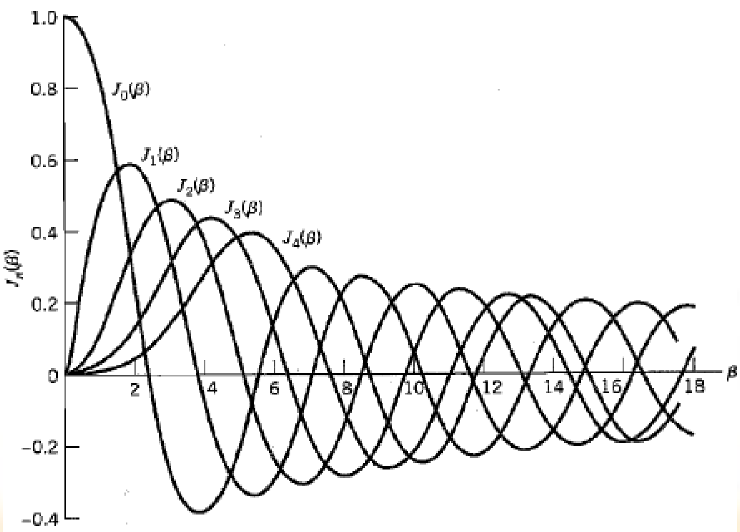


FIGURE Plots of Bessel functions of the first kind for varying order.

Properties of Bessel Function

1. $J_n(\beta) = (-1)^n J_{-n}(\beta)$ for all n, both positive and negative ....**(14)**

2. For small values of the modulation index β, we have

$$\left.\begin{aligned} J_0(\beta) &\simeq 1 \\ J_1(\beta) &\simeq \frac{\beta}{2} \\ J_n(\beta) &\simeq 0, \qquad n > 2 \end{aligned}\right\} \qquad \text{...............(15)}$$

3.
$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \qquad \text{...................(16)}$$

Thus using equations (13) through (16) and the curves in the above figure, following observations are made

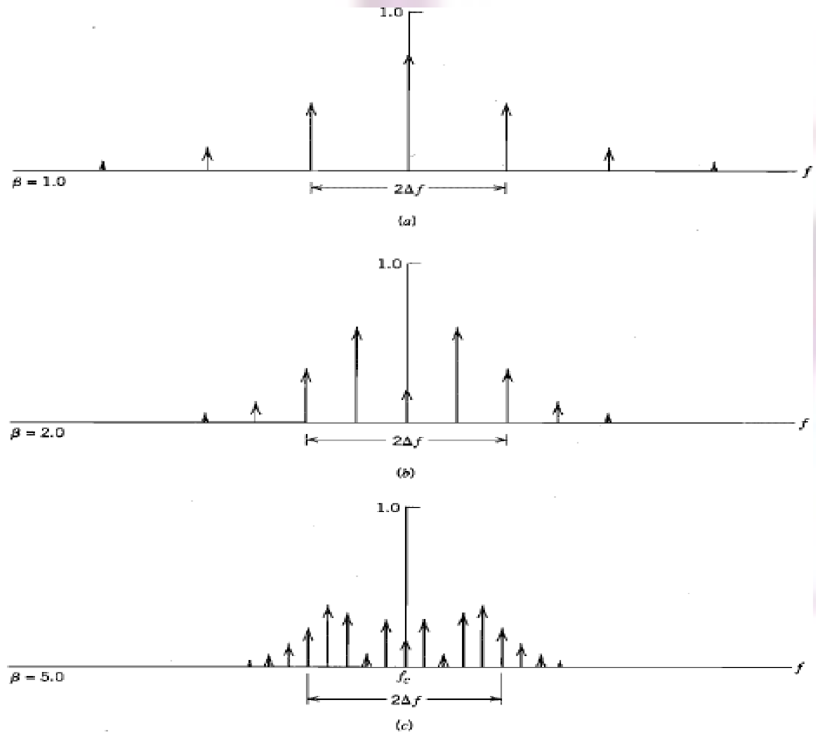
1. The spectrum of an FM signal contains a carrier component and an infinite set of side frequencies located symmetrically on either side of the carrier at frequency separations of $f_m, 2f_m, 3f_m, \cdots$. In this respect, the result is unlike that which prevails in an AM system, since in an AM system a sinusoidal modulating signal gives rise to only one pair of side frequencies.
2. For the special case of β small compared with unity, only the Bessel coefficients $J_0(\beta)$ and $J_1(\beta)$ have significant values, so that the FM signal is effectively composed of a carrier and a single pair of side frequencies at $f_c \pm f_m$. This situation corresponds to the special case of narrowband FM that was considered earlier.
3. The amplitude of the carrier component varies with β according to $J_0(\beta)$. That is, unlike an AM signal, the amplitude of the carrier component of an FM signal is dependent on the modulation index β. The physical explanation for this property is that the envelope of an FM signal is constant, so that the average power of such a signal developed across a 1-ohm resistor is also constant, as shown by

$$P = \frac{1}{2} A_c^2 \qquad \text{..................(17)}$$

When the carrier is modulated to generate the FM signal, the power in the side frequencies may appear only at the expense of the power originally in the carrier, thereby making the amplitude of the carrier component dependent on β. Note that the average power of an FM signal may also be determined from Equation (**12**), obtaining

$$P = \frac{1}{2} A_c^2 \sum_{n=-\infty}^{\infty} J_n^2(\beta) \qquad \text{......................(18)}$$

Spectrum Analysis of Sinusoidal FM Wave using Bessel functions



The above figure shows the Discrete amplitude spectra of an FM signal, normalized with respect to the carrier amplitude, for the case of sinusoidal modulation of varying frequency and fixed amplitude. Only the spectra for positive frequencies are shown.

Transmission Bandwidth of FM waves

In theory, an FM signal contains an infinite number of side frequencies so that the bandwidth required to transmit such a signal is similarly infinite in extent. In practice, however, we find that the FM signal is effectively limited to a finite number of significant side frequencies compatible with a specified amount of distortion. We may therefore specify an effective bandwidth required for the transmission of an FM signal. Consider first the

case of an FM signal generated by a single-tone modulating wave of frequency f_m. In such an FM signal, the side frequencies that are separated from the carrier frequency f_c by an amount greater than the frequency deviation Δf decrease rapidly toward zero, so that the bandwidth always exceeds the total frequency excursion, but nevertheless is limited. Specifically, for large values of the modulation index β, the bandwidth approaches, and is only slightly greater than, the total frequency excursion $2\Delta f$ in accordance with the situation shown in Figure 2.25. On the other hand, for small values of the modulation index β, the spectrum of the FM signal is effectively limited to the carrier frequency f_c and one pair of side frequencies at $f_c \pm f_m$, so that the bandwidth approaches $2f_m$. We may thus define an approximate rule for the transmission bandwidth of an FM signal generated by a single-tone modulating signal of frequency f_m as follows:

$$B_T \simeq 2\Delta f + 2f_m = 2\Delta f\left(1 + \frac{1}{\beta}\right)$$

This relation is known as **Carson's rule.**

Generation of FM Signal

Direct methods for FM generation

Reactance modulator:

The direct method of FM generation using the reactance modulator involves providing a voltage-variable reactance across the tank circuit of an oscillator. Though the varactor diode modulator can be called a reactance modulator, the term is generally applied to those modulators in which an active device is made to behave as a variable reactance. In Fig. we show the basic FET capacitive reactance modulator.

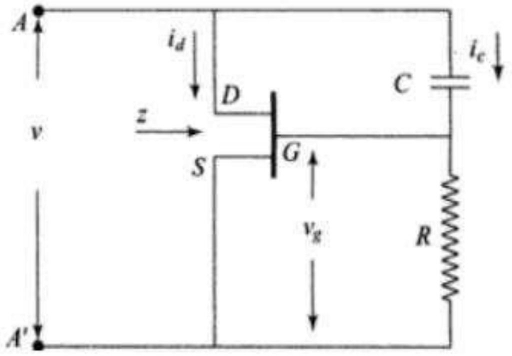


Fig. *FET capacitive reactance modulator*

Under certain conditions, the impedance z, across terminals AA' is almost entirely reactive. The circuit can be made either inductive or capacitive by a simple component change and its reactance can be shown to be proportional to the transconductance of the device, which in turn can be made to depend on the variations in the gate bias. The circuit impedance is: $z = \frac{v}{i_d}$, where v is the voltage applied across AA' and i_d is the resulting drain current. For the impedance across AA' to be capacitive, the following conditions have to be met:

- The bias current i_c must be negligible compared to the drain current i_d.
- The drain-to-gate impedance (X_c) must be greater than gate-to-source impedance (R here) by at least a factor of five.

$$v_g = i_c R = \frac{vR}{(R - jX_c)}$$

Drain current: $i_d = g_m v_g = \dfrac{g_m vR}{(R - jX_c)}$

Since $X_c \gg R$, we have:

$$z = \frac{v}{i_d} \approx -\frac{jX_c}{g_m R}$$

This is clearly a capacitive reactance with equivalent impedance:

$$X_{eq} = \frac{X_c}{g_m R} = \frac{1}{2\pi f C g_m R} = \frac{1}{2\pi f C_{eq}}$$

Hence, under these assumptions, the impedance looking into AA' is a pure reactance given by:

$$C_{eq} = g_m CR$$

Since C_{eq} depends on the transconductance g_m, it can be varied with bias voltage.

Now $X_c \gg R$. Let $X_c = nR$ at the carrier frequency. Then, $\dfrac{1}{\omega C} = nR$.

Therefore, $$C = \frac{1}{\omega n R} = \frac{1}{2\pi f n R}$$

$$C_{eq} = g_m C R = \frac{g_m}{2\pi f n}$$

What would have happened if the positions of C and R are interchanged and if $R \gg X_c$?

$$v_g = i_c R = \left[\frac{v(-jX_c)}{(R - jX_c)}\right]$$

Drain current: $i_d = g_m v_g = \left[\frac{g_m v(-jX_c)}{(R - jX_c)}\right]$

$$z = \frac{v}{i_d} = \frac{(R - jX_c)}{(-jX_c\, g_m)}$$

$$= \left(\frac{1}{X_c\, g_m}\right)[X_c + jR] \approx \frac{jR}{(X_c\, g_m)}$$

Clearly, the impedance is inductive and can be written as:

$$X_{eq} = \frac{R}{X_c\, g_m} = \frac{(2\pi f C R)}{g_m} = 2\pi f L_{eq}, \text{ where } L_{eq} = \frac{CR}{g_m}$$

Thus, the FET reactance modulator behaves as a three-terminal reactive element (either inductive or capacitive) that may be connected across the tank circuit of the oscillator to be frequency modulated. The reactance appears between the drain and the source and its value may be controlled by a signal at the third terminal, i.e., the gate.

A disadvantage of the above methods of FM generation is that they do not provide carrier frequency stability. For attaining the required order of carrier frequency stability in the 88-108 MHz range used for FM transmission, it is necessary to use crystal oscillators. However, the high Q of crystal oscillators permits direct modulation only in some narrowband applications[2]. For wideband FM generation using crystal oscillators, the indirect method is adopted.

Indirect Method for WBFM Generation (ARMSTRONG'S Method):

In this method, an NBFM signal is generated using an integrator and a phase modulator. The NBFM signal is then converted to WBFM using a frequency multiplier. Let us first consider the principle behind a frequency multiplier.

A frequency multiplier is an amplifier whose output signal frequency is an integer multiple of the input frequency. If the input to the frequency multiplier is $A\cos\theta(t)$ then the output is $A\cos[n\theta(t)]$, where n is an integer. This can be achieved by feeding a signal frequency that is rich in harmonic distortion (e.g. from a class C amplifier) into an LC tank circuit tuned to n times the input frequency. This arrangement results in the n^{th} harmonic being the only significant output. For larger multiplication factors, a cascade of doublers and triplers are used. For example $n = 1000$ could be approximated by a cascade of ten doublers ($2^{10} = 1024$).

Effect of frequency multiplication on a NBFM signal

Consider single tone NBFM:

$s_{NBFM}(t) = A_c\cos(\omega_c' t + \beta' \sin\omega_m t)$ where $\beta' < 1$ rad and ω_c' is a stable sub-carrier frequency.

When applied to a frequency multiplier, the NBFM signal is converted to a WBFM signal given by:

$$s_{FM}(t) = A_c\cos[n(\omega_c' t + \beta' \sin\omega_m t)]$$
$$= A_c\cos[n\omega_c' t + n\beta' \sin\omega_m t]$$
$$s_{FM}(t) = A_c\cos[\omega_c t + \beta\sin\omega_m t]$$

where $\omega_c = n\omega_c'$ is the final carrier frequency and $\beta = n\beta'$ is the final β.

Note that ω_m is unaffected by frequency multiplication.

The maximum frequency deviation of the NBFM signal also gets multiplied by a factor of n since $\beta = n\beta'$ implies $(\Delta\omega/\omega_m) = n(\Delta\omega'/\omega_m)$ or $\Delta\omega = n\,\Delta\omega'$.

Detection of FM Signal

Balanced Slope Detector

Balanced slope detector. Figure **(a)** shows the schematic diagram for a *balanced slope detector.* A single-ended slope detector is a tuned-circuit frequency discriminator, and a balanced slope detector is simply two single-ended slope detectors connected in parallel and fed 180° out of phase. The phase inversion is accomplished by center tapping the tuned secondary windings of transformer T_1. In Figure **(a),** the tuned circuits (L_a, C_a, and L_b, C_b) perform the FM-to-AM conversion, and the balanced peak detectors (D_1, C_1, R_1, and D_2, C_2, R_2) remove the information from the AM envelope. The top tuned circuit (L_a and C_a) is tuned to a frequency (f_a) that is above the IF center frequency (f_o) by approximately $1.33 \times \Delta f$ (for the FM broadcast band, this is approximately 1.33×75 kHz = 100 kHz). The lower tuned circuit (L_b and C_b) is tuned to a frequency (f_b) that is below the IF center frequency by an equal amount.

Circuit operation is quite simple. The output voltage from each tuned circuit is proportional to the input frequency, and each output is rectified by its respective peak detector. Therefore, the closer the input frequency is to the tank-circuit resonant frequency, the greater the tank-circuit output voltage. The IF center frequency falls exactly halfway between the resonant frequencies of the two tuned circuits. Therefore, at the IF center frequency, the output voltages from the two tuned circuits are equal in amplitude but opposite in polarity. Consequently, the rectified output voltage across $R1$ and $R2$,

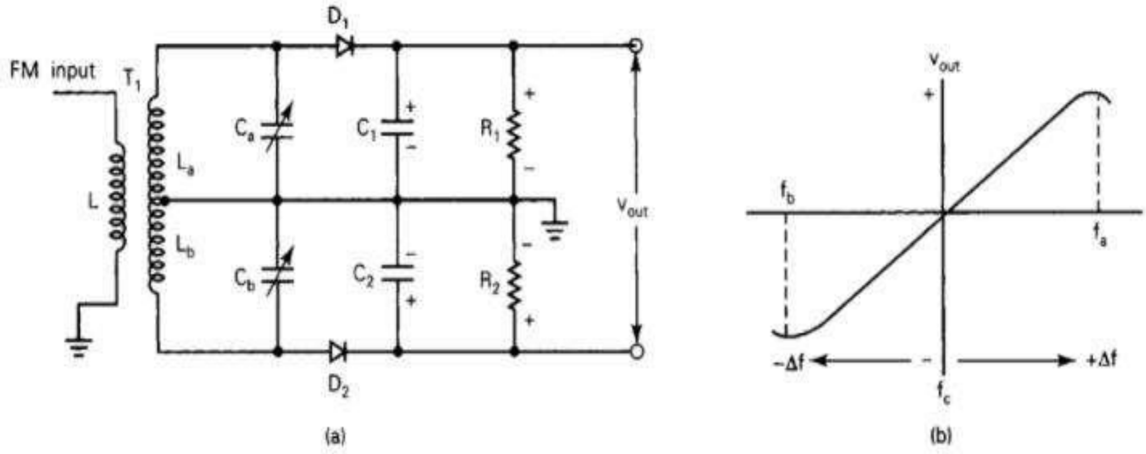


FIGURE Balanced slope detector: (a) schematic diagram; (b) voltage-versus-frequency response curve

when added, produce a differential output voltage V_{out} 5 0 V. When the IF deviates above resonance, the top tuned circuit produces a higher output voltage than the lower tank circuit, and Vout goes positive. When the IF deviates below resonance, the output voltage from the lower tank circuit is larger than the output voltage from the upper tank circuit, and V_{out} goes negative. The output-versus-frequency response curve is shown in Figure **(b).**

Although the slope detector is probably the simplest FM detector, it has several inherent disadvantages, which include poor linearity, difficulty in tuning, and lack of provisions for limiting. Because limiting is not provided, a slope detector produces an output voltage that is proportional to amplitude, as well as frequency variations in the input signal and, consequently, must be preceded by a separate limiter stage. A balanced slope detector is aligned by injecting a frequency equal to the IF center frequency and tuning C_a and C_b for 0 V at the output. Then frequencies equal to f_a and f_b are alternately injected while C_a and C_b are tuned for maximum and equal output voltages with opposite polarities.

Phase Locked Loop

Basically, the *phase-locked loop* consists of three major components: a *multiplier*, a *loop filter*, and a *voltage-controlled oscillator* (VCO) connected together in the form of a feedback system, as shown in Figure below. The VCO is a sinusoidal generator whose frequency is determined by a voltage applied to it from an external source. In effect, any frequency modulator may serve as a VCO. We assume that initially we have adjusted the VCO so that when the control voltage is zero, two conditions are satisfied:

1. The frequency of the VCO is precisely set at the unmodulated carrier frequency f_c.
2. The VCO output has a 90 degree phase-shift with respect to the unmodulated carrier wave.

Suppose then that the input signal applied to the phase-locked loop is an FM signal defined by

$$s(t) = A_c \sin[2\pi f_c t + \phi_1(t)]$$

where A_c is the carrier amplitude. With a modulating signal $m(t)$, the angle $\phi_1(t)$ is related to $m(t)$ by the integral

$$\phi_1(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau$$

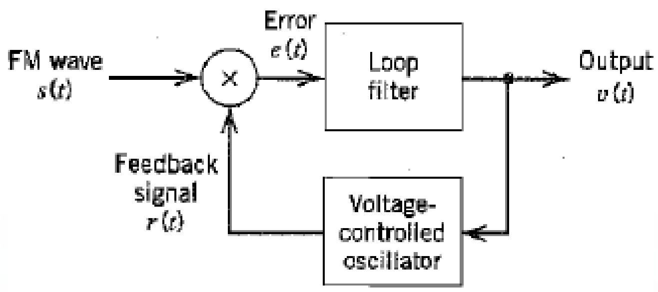


FIGURE Phase-locked loop.

where k_f is the frequency sensitivity of the frequency modulator. Let the VCO output in the phase-locked loop be defined by

$$r(t) = A_v \cos[2\pi f_c t + \phi_2(t)]$$

where A_v is the amplitude. With a control voltage $v(t)$ applied to the VCO input, the angle $\phi_2(t)$ is related to $v(t)$ by the integral

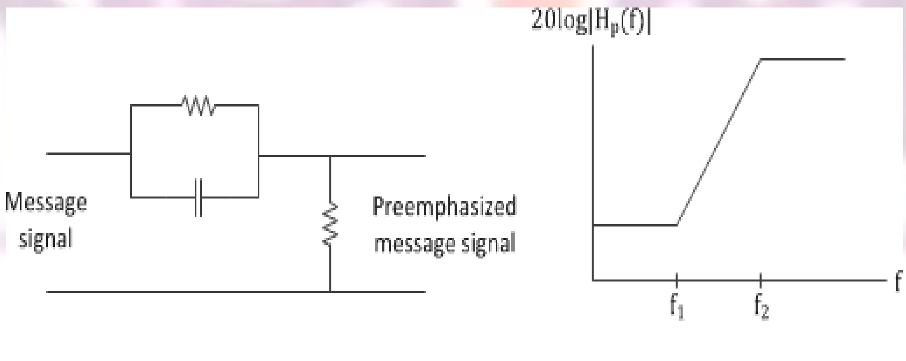
$$\phi_2(t) = 2\pi k_v \int_0^t v(\tau)\, d\tau$$

where k_v is the frequency sensitivity of the VCO, measured in Hertz per volt. The object of the phase-locked loop is to generate a VCO output $r(t)$ that has the same phase angle (except for the fixed difference of 90 degrees) as the input FM signal $s(t)$. The time-varying phase angle $\phi_1(t)$ characterizing $s(t)$ may be due to modulation by a message signal $m(t)$, in which case we wish to recover $\phi_1(t)$ and thereby produce an estimate of $m(t)$. In other applications of the phase-locked loop, the time-varying phase angle $\phi_1(t)$ of the incoming signal $s(t)$ may be an unwanted phase shift caused by fluctuations in the communication channel; in this latter case, we wish to *track* $\phi_1(t)$ so as to produce a signal with the same phase angle for the purpose of coherent detection (synchronous demodulation).

PRE-EMPHASIS AND DE-EMPHASIS NETWORKS

In FM, the noise increases linearly with frequency. By this, the higher frequency components of message signal are badly affected by the noise. To solve this problem, we can use a pre-emphasis filter of transfer function $H_p(f)$ at the transmitter to boost the higher frequency components before modulation. Similarly, at the receiver, the de-emphasis filter of transfer function $H_d(f)$can be used after demodulator to attenuate the higher frequency components thereby restoring the original message signal.

The pre-emphasis network and its frequency response are shown in Figure (a) and (b) respectively. Similarly, the counter part for de-emphasis network is shown in Figure below.



(a) (b)

Figure (a) Pre-emphasis network. (b) Frequency response of pre-emphasis network.

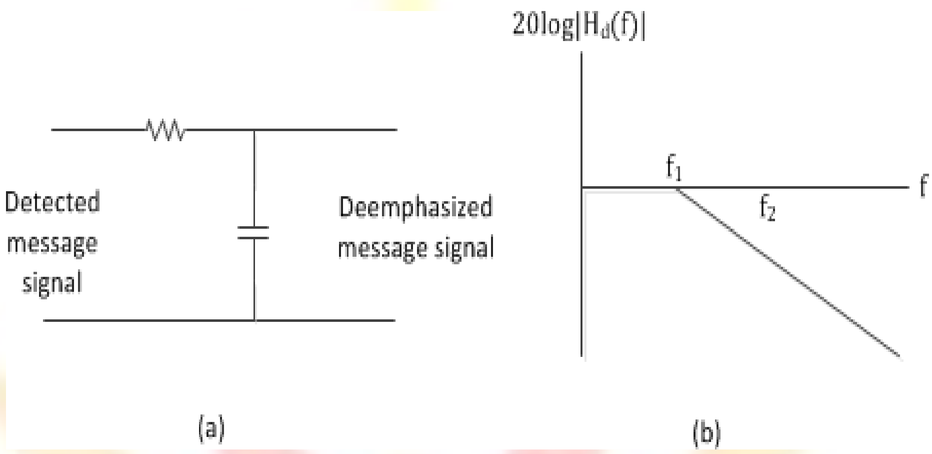


Figure (a) De-emphasis network. (b) Frequency response of De-emphasis network.

Comparison of AM and FM

S.NO	AMPLITUDE MODULATION	FREQUENCY MODULATION
1.	Band width is very small which is one of the biggest advantage	It requires much wider channel (7 to 15 times) as compared to AM.
2.	The amplitude of AM signal varies depending on modulation index.	The amplitude of FM signal is constant and independent of depth of the modulation.
3.	Area of reception is large	The area of reception is small since it is limited to line of sight.
4.	Transmitters are relatively simple & cheap.	Transmitters are complex and hence expensive.
5.	The average power in modulated wave is greater than carrier power. This added power is provided by modulating source.	The average power in frequency modulated wave is same as contained in un-modulated wave.
6.	More susceptible to noise interference and has low signal to noise ratio, it is more difficult to eliminate effects of noise.	Noise can be easily minimized amplitude variations can be eliminated by using limiter.
7.	It is not possible to operate without interference.	It is possible to operate several independent transmitters on same frequency.
8.	The maximum value of modulation index = 1, otherwise over-modulation would result in distortions.	No restriction is placed on modulation index.

UNIT-III

TRANSMITTERS AND RECEIVERS

Radio Transmitters

There are two approaches in generating an AM signal. These are known as low and high level modulation. They're easy to identify: A low level AM transmitter performs the process of modulation near the beginning of the transmitter. A high level transmitter performs the modulation step last, at the last or "final" amplifier stage in the transmitter. Each method has advantages and disadvantages, and both are in common use.

Low-Level AM Transmitter:

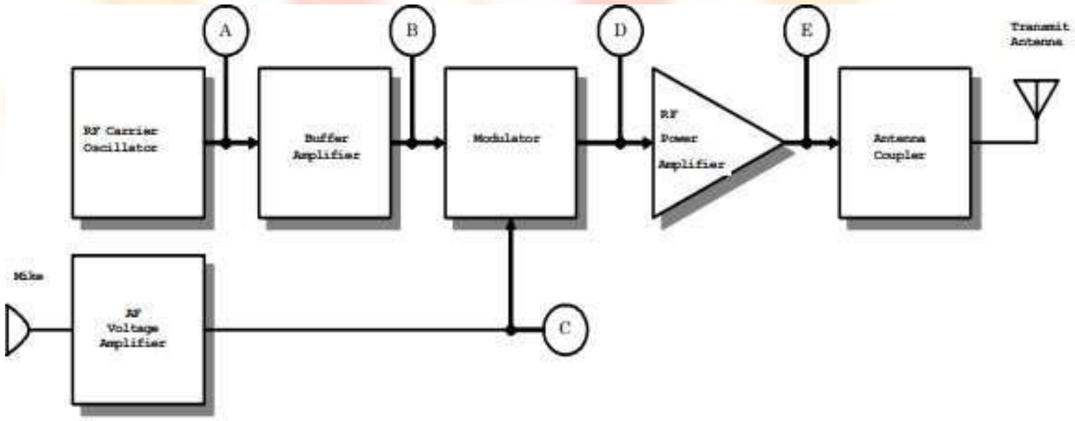


Fig.1. Low-Level AM Transmitter Block Diagram

There are two signal paths in the transmitter, audio frequency (AF) and radio frequency (RF). The RF signal is created in the RF carrier oscillator. At test point A the oscillator's output signal is present. The output of the carrier oscillator is a fairly small AC voltage, perhaps 200 to 400 mV RMS. The oscillator is a critical stage in any transmitter. It must produce an accurate and steady frequency. Every radio station is assigned a different carrier frequency. The dial (or display) of a receiver displays the carrier frequency. If the oscillator drifts off frequency, the receiver will be unable to receive the transmitted signal without being readjusted. Worse yet, if the oscillator drifts onto the frequency being used by another radio station, interference will occur. Two circuit techniques are commonly used to stabilize the oscillator, buffering and voltage regulation.

The buffer amplifier has something to do with buffering or protecting the oscillator. An oscillator is a little like an engine (with the speed of the engine being similar to the oscillator's frequency). If the load on the engine is increased (the engine is asked to do more work), the engine will respond by slowing down. An oscillator acts in a very similar fashion. If the current drawn from the oscillator's output is increased or decreased, the oscillator may speed up or slow down slightly.

Buffer amplifier is a relatively low-gain amplifier that follows the oscillator. It has a constant input impedance (resistance). Therefore, it always draws the same amount of current from the oscillator. This helps to prevent "pulling" of the oscillator frequency. The buffer amplifier is needed because of what's happening "downstream" of the oscillator. Right after this stage is the modulator. Because the modulator is a nonlinear amplifier, it may not have a constant input resistance -- especially when information is passing into it. But since there is a buffer amplifier between the oscillator and modulator, the oscillator sees a steady load resistance, regardless of what the modulator stage is doing.

Voltage Regulation: An oscillator can also be pulled off frequency if its power supply voltage isn't held constant. In most transmitters, the supply voltage to the oscillator is regulated at a constant value. The regulated voltage value is often between 5 and 9 volts; zener diodes and three-terminal regulator ICs are commonly used voltage regulators. Voltage regulation is especially important when a transmitter is being powered by batteries or an automobile's electrical system. As a battery discharges, its terminal voltage falls. The DC supply voltage in a car can be anywhere between 12 and 16 volts, depending on engine RPM and other electrical load conditions within the vehicle.

Modulator: The stabilized RF carrier signal feeds one input of the modulator stage. The modulator is a variable-gain (nonlinear) amplifier. To work, it must have an RF carrier signal and an AF information signal. In a low-level transmitter, the power levels are low in the oscillator, buffer, and modulator stages; typically, the modulator output is around 10 mW (700 mV RMS into 50 ohms) or less.

AF Voltage Amplifier: In order for the modulator to function, it needs an information signal. A microphone is one way of developing the intelligence signal, however, it only produces a few millivolts of signal. This simply isn't enough to operate the modulator, so a voltage amplifier is used to boost the microphone's signal. The signal level at the output of the AF voltage amplifier is usually at least 1 volt RMS; it is highly dependent upon the transmitter's design. Notice that the AF amplifier in the transmitter is only providing a voltage gain, and not necessarily a current gain for the microphone's signal. The power levels are quite small at the output of this amplifier; a few mW at best.

RF Power Amplifier: At test point D the modulator has created an AM signal by impressing the information signal from test point C onto the stabilized carrier signal from test point B at the buffer amplifier output. This signal (test point D) is a complete AM signal, but has only a few milliwatts of power. The RF power amplifier is normally built with several stages. These stages increase both the voltage and current of the AM signal. We say that power amplification occurs when a circuit provides a current gain. In order to accurately amplify the tiny AM signal from the modulator, the RF power amplifier stages must be linear. You might recall that amplifiers are divided up into "classes," according to the conduction angle of the active device within. Class A and class B amplifiers are considered to be linear amplifiers, so the RF power amplifier stages will normally be constructed using one or both of these type of amplifiers. Therefore, the signal at test point E looks just like that of test point D; it's just much bigger in voltage and current.

Antenna Coupler: The antenna coupler is usually part of the last or final RF power amplifier, and as such, is not really a separate active stage. It performs no amplification, and has no active devices. It performs two important jobs: Impedance matching and filtering. For an RF power amplifier to function correctly, it must be supplied with a load resistance equal to that for which it was designed.

The antenna coupler also acts as a low-pass filter. This filtering reduces the amplitude of harmonic energies that may be present in the power amplifier's output. (All amplifiers generate harmonic distortion, even "linear" ones.) For example, the transmitter may be tuned to operate on 1000 kHz. Because of small nonlinearities in the amplifiers of the transmitter, the transmitter will also produce harmonic energies on 2000 kHz (2nd harmonic), 3000 kHz (3rd harmonic), and so on. Because a low-pass filter passes the fundamental frequency (1000 kHz) and rejects the harmonics, we say that harmonic attenuation has taken place.

High-Level AM Transmitter:

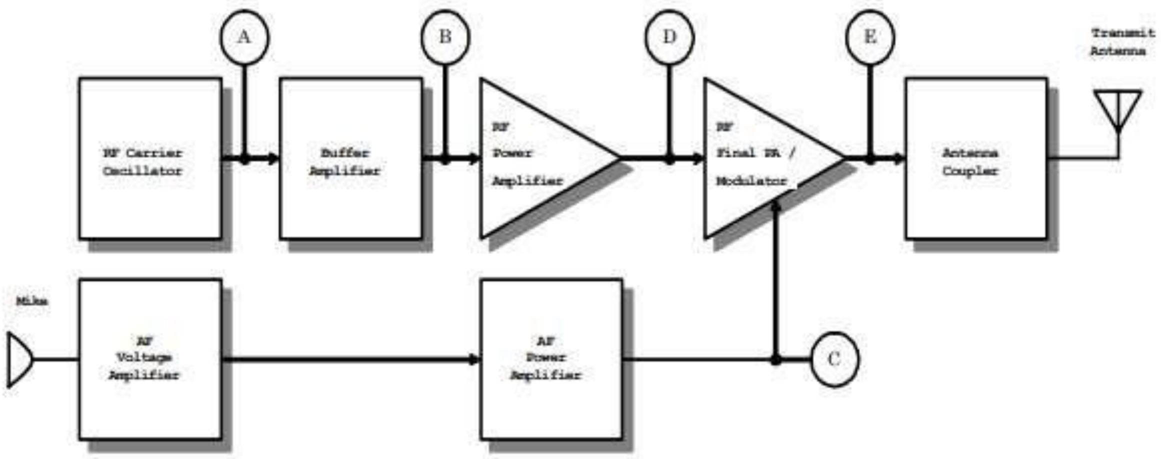


Fig.2. High-Level AM Transmitter Block Diagram

The high-level transmitter of Figure 9 is very similar to the low-level unit. The RF section begins just like the low-level transmitter; there is an oscillator and buffer amplifier. The difference in the high level transmitter is where the modulation takes place. Instead of adding modulation immediately after buffering, this type of transmitter amplifies the unmodulated RF carrier signal first. Thus, the signals at points A, B, and D in Figure 9 all look like unmodulated RF carrier waves. The only difference is that they become bigger in voltage and current as they approach test point D.

The modulation process in a high-level transmitter takes place in the last or final power amplifier. Because of this, an additional audio amplifier section is needed. In order to modulate an amplifier that is running at power levels of several watts (or more), comparable power levels of information are required. Thus, an audio power amplifier is required. The final power amplifier does double-duty in a high-level transmitter. First, it provides power gain for the RF carrier signal, just like the RF power amplifier did in the low-level transmitter. In addition to providing power gain, the final PA also performs the task of modulation. The final power amplifier in a high-level transmitter usually operates in class C, which is a highly nonlinear amplifier class.

Comparison:

Low Level Transmitters

- Can produce any kind of modulation; AM, FM, or PM.
- Require linear RF power amplifiers, which reduce DC efficiency and increases production costs.

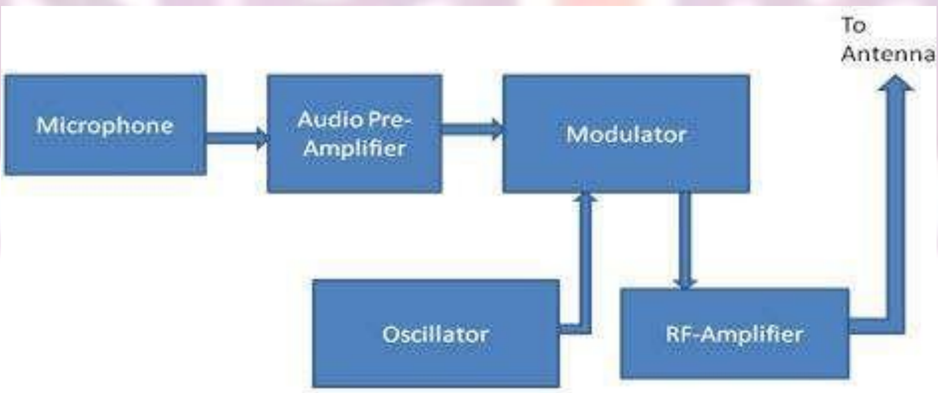
High Level Transmitters

- Have better DC efficiency than low-level transmitters, and are very well suited for battery operation.
- Are restricted to generating AM modulation only.

FM Transmitter

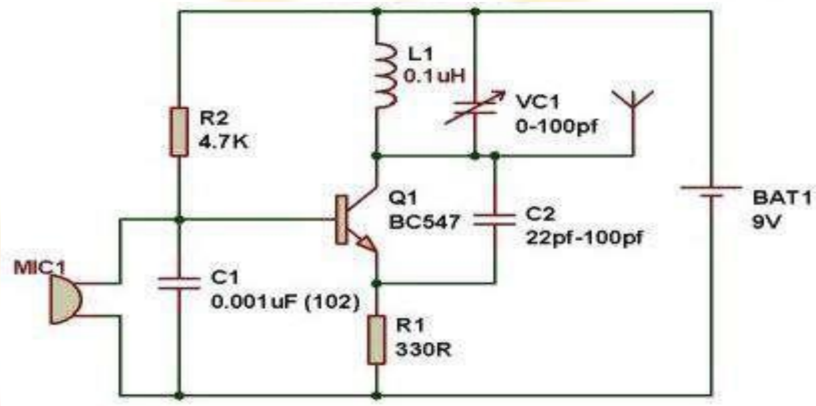
The FM transmitter is a single transistor circuit. In the telecommunication, the frequency modulation (FM)transfers the information by varying the frequency of carrier wave according to the message signal. Generally, the FM transmitter uses VHF radio frequencies of 87.5 to 108.0 MHz to transmit & receive the FM signal. This transmitter accomplishes the most excellent range with less power. The performance and working of the wireless audio transmitter circuit is depends on the induction coil & variable capacitor. This article will explain about the working of the FM transmitter circuit with its applications.

The FM transmitter is a low power transmitter and it uses FM waves for transmitting the sound, this transmitter transmits the audio signals through the carrier wave by the difference of frequency. The carrier wave frequency is equivalent to the audio signal of the amplitude and the FM transmitter produce VHF band of 88 to 108MHZ.Plese follow the below link for: Know all About Power Amplifiers for FM Transmitter



Working of FM Transmitter Circuit

The following circuit diagram shows the FM transmitter circuit and the required electrical and electronic components for this circuit is the power supply of 9V, resistor, capacitor, trimmer capacitor, inductor, mic, transmitter, and antenna. Let us consider the microphone to understand the sound signals and inside the mic there is a presence of capacitive sensor. It produces according to the vibration to the change of air pressure and the AC signal.



The formation of the oscillating tank circuit can be done through the transistor of 2N3904 by using the inductor and variable capacitor. The transistor used in this circuit is an NPN transistor used for general purpose amplification. If the current is passed at the inductor L1 and variable capacitor then the tank circuit will oscillate at the resonant carrier frequency of the FM modulation. The negative feedback will be the capacitor C2 to the oscillating tank circuit.

To generate the radio frequency carrier waves the FM transmitter circuit requires an oscillator. The tank circuit is derived from the LC circuit to store the energy for oscillations. The input audio signal from the mic penetrated to the base of the transistor, which **modulates the LC tank circuit** carrier frequency in FM format. The variable capacitor is used to change the resonant frequency for fine modification to the FM frequency band. The modulated signal from the antenna is radiated as radio waves at the FM frequency band and the antenna is nothing but copper wire of 20cm long and 24 gauge. In this circuit the length of the antenna should be significant and here you can use the 25-27 inches long copper wire of the antenna.

Application of FM Transmitter

- The FM transmitters are used in the homes like sound systems in halls to fill the sound with the audio source.
- These are also used in the cars and fitness centres.
- The correctional facilities have used in the FM transmitters to reduce the prison noise in common areas.

Advantages of the FM Transmitters

- The FM transmitters are easy to use and the price is low
- The efficiency of the transmitter is very high
- It has a large operating range
- This transmitter will reject the noise signal from an amplitude variation.

Receivers

Introduction to Radio Receivers:

In radio communications, a **radio receiver** (**receiver** or simply **radio**) is an electronic device that receives **radio** waves and converts the information carried by them to a usable form.

Types of Receivers:

The TRF (Tuned Radio Frequency) Receiver and Superheterodyne Receiver are the two main configurations of the receivers, they have real practical or commercial significance. Most of the present day receivers use superheterodyne configuration. But the TRF receivers are simple and easy to understand.

Tuned Radio Frequency Receiver:

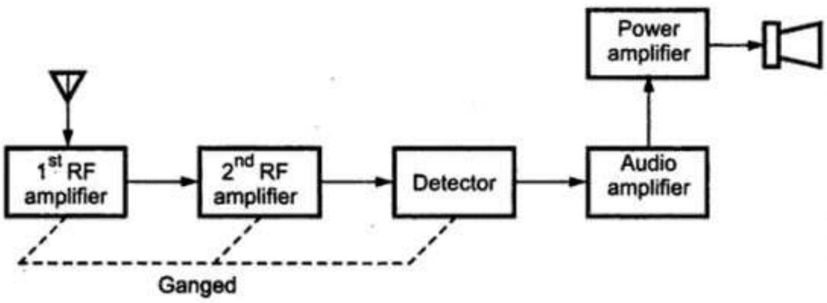


Fig. 1 shows the block diagram of TRF receiver.

It consists of two or three stages of RF amplifiers, detector, audio amplifier and power amplifier. The RF amplifier stages placed between the antenna and detector are used to increase the strength of the received signal before it is applied to the detector. These RF amplifiers are tuned to fix frequency, amplify the desired band of frequencies. Therefore, they provide amplification for selected band of frequencies and rejection for all others. As selection and amplification process is carried out in two or three stages and each stage must amplify the same band of frequencies, the ganged tuning is provided.

The amplified signal is then demodulated using detector to recover the modulating signal. The recovered signal is amplified further by the audio amplifier followed by power amplifier which provides sufficient gain to operate a loudspeaker. The TRF receivers suffered from number of annoying problems. These are listed in the next section.

Problems in TRF Receivers:

1. Tracking of Tuned Circuit

In a receiver, tuned circuits are made variable so that they can be set to the frequency of the desired signal. In most of the receivers, the capacitors in the tuned circuits are made variable. These capacitors are 'ganged' between the stages so that they all can be changed simultaneously when the tuning knob is rotated. To have perfect tuning the capacitor values between the stages must be exactly same but this is not the case. The differences in the capacitors cause the resonant frequency of each tuned circuit to be slightly different, thereby increasing the pass band.

2. Instability

As high gain is achieved at one frequency by a multistage amplifier, there are more chances of positive feedback (of getting back the small part of output of the last stage at the input to the first with the correct polarity) through some stray path, resulting in oscillations. These oscillations are unavoidable at high frequencies.

3. Variable Bandwidth

TRF receivers suffer from a variation in bandwidth over the tuning range. Consider a medium wave receiver required to tune over 535 kHz to 1640 kHz and it provides the necessary bandwidth of 10 kHz at 535 kHz. Let us calculate Q of this circuit.

$$Q = \frac{f}{\text{Bandwidth}} = \frac{535\,\text{kHz}}{10\,\text{kHz}} = 53.5$$

Now consider the frequency at the other end of the broadcast band, i.e. 1640 kHz. At 1640 kHz, Q of the coil should be 164 (1640 kHz / 10 kHz). However, in practice due to various losses depending on frequency, we will not set so large increase in Q. Let us assume that at 1640 kHz frequency Q is increased to value 100 instead of 164. With this Q of the tuned circuit bandwidth can be calculated as follows

$$\text{Bandwidth} = \frac{f}{Q} = \frac{1640\,\text{kHz}}{100} = 16.4\text{ kHz}$$

We know, necessary bandwidth is 10 kHz. This increase in bandwidth of tuned circuit, pick up the adjacent stations along with station it is tuned for, providing insufficient adjacent frequency rejection. In other words we can say that in TRF receivers the bandwidth of the tuned circuit varies over the frequency range, resulting in poor selectivity of the receiver.

Because of the problems of tracking, instability and bandwidth variation, the TRF receivers have almost been replaced by superheterodyne receivers.

Superheterodyne Receivers

To solve basic problems of TRF receivers, in these receivers, first all the incoming RF frequencies are converted to a fix lower frequency called **intermediate frequency** (IF). Then this fix intermediate frequency is amplified and detected to reproduce the original information. Since the characteristics of the IF amplifier are independent of the frequency to which the receiver is tuned, the selectivity and sensitivity of superheterodyne receivers are fairly uniform throughout its tuning range.

Mixer circuit is used to produce the frequency translation of the incoming signal down to the IF. The incoming signals are mixed with the local oscillator frequency signal in such a way that a constant frequency difference is maintained between the local oscillator and the incoming signals. This is achieved by using ganged tuning capacitors.

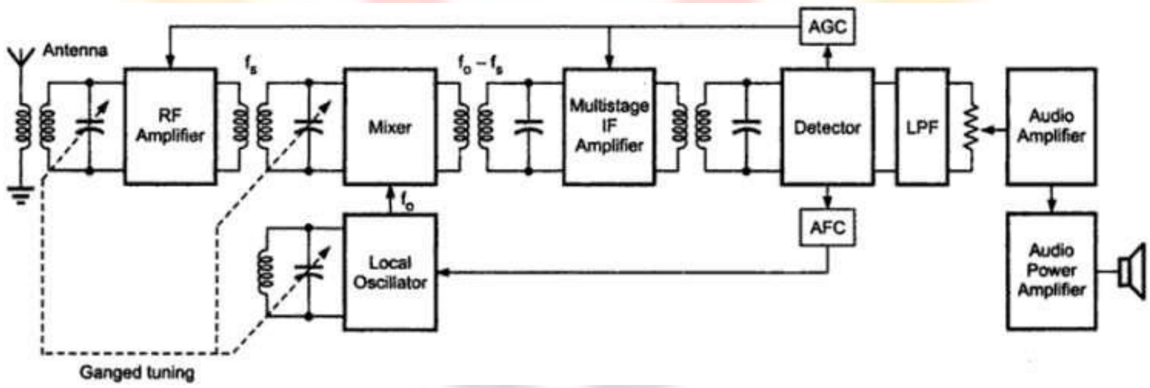


Fig.2. Block diagram of Super heterodyne Receiver.

Fig. 2 shows the block diagram of superheterodyne receiver. As shown in the Fig. 2 antenna picks up the weak radio signal and feeds it to the RF amplifier. The RF amplifier provides some initial gain and selectivity. The output of the RF amplifier is applied to the input of the mixer. The mixer also receives an input from local oscillator.

The output of the mixer circuit is difference frequency ($f_o - f_s$) commonly known as IF (Intermediate Frequency). The signal at this intermediate frequency contains the same modulation as the original carrier. This signal is amplified by one or more IF amplifier stages, and most of the receiver gain is obtained in these IF stages.

The highly amplified IF signal is applied to detector circuits to recover the original modulating information. Finally, the output of detector circuit is fed to audio and power amplifier which provides a sufficient gain to operate a speaker.

Another important circuit in the superheterodyne receiver are AGC and AFC circuit. AGC is used to maintain a constant output voltage level over a wide range of RF input signal levels.

It derives the dc bias voltage from the output of detector which is proportional to the amplitude of the received signal. This dc bias voltage is feedback to the IF amplifiers, and sometimes to the RF amplifier, to control the gain of the receiver. As a result, it provides a constant output voltage level over a wide range of RF input signal levels. AFC circuit generates AFC signal which is used to adjust and stabilize the frequency of the local oscillator.

Characteristics of Radio Receiver:

The performance of the radio receiver can be measured in terms of following receiver characteristics

- Selectivity
- Sensitivity
- Fidelity
- Image frequency and its rejection
- Double spotting

Selectivity

Selectivity refers to the ability of a receiver to select a signal of a desired frequency while reject all others. Selectivity in a receiver is obtained by using tuned circuits. These are LC circuits tuned to resonate at a desired signal frequency. The Q of these tuned circuits determines the selectivity. Selectivity shows the attenuation that the receiver offers to signals at frequencies near to the one to which it is tuned. A good receiver isolates the desired signal in the RF spectrum and eliminate all other signals.

Recall that Q is the ratio of inductive reactance to resistance ($Q = X_L/R$), and we know that bandwidth of the tuned circuit is given by

$$B_w = \frac{f_r}{Q}$$

where f_r is the resonant frequency. The bandwidth of a tuned circuit is measure of the selectivity. Narrower the bandwidth the better selectivity. To have narrower bandwidth and better selectivity the Q of the tuned circuit must be high.

Sensitivity

The sensitivity of a communication receiver refers to the receivers ability to pick up weak signals, and amplify it. It is often defined in terms of the voltage that must be applied to the receiver input terminals to give a standard output power, measured at the output terminals. The more gain that a receiver has, the smaller the input signal necessary to produce desired output power. Therefore, sensitivity is a primary function of the overall receiver gain. It is often expressed in microvolts or in decibels. The sensitivity of receiver mostly depends on the gain of the IF amplifiers. Good communication receiver has sensitivity of 0.2 to 1μV.

Fidelity

Fidelity refers to the ability of the receiver to reproduce all the modulating frequencies equally. Fig. 3 shows the typical fidelity curve for radio receiver.

The fidelity at the lower modulating frequencies is determined by the low frequency response of the IF amplifier and the fidelity at the higher modulating frequencies is determined by the high frequency response of the IF amplifier. Fidelity is difficult to obtain in AM receiver because good fidelity requires more bandwidth of IF amplifier resulting in poor selectivity.

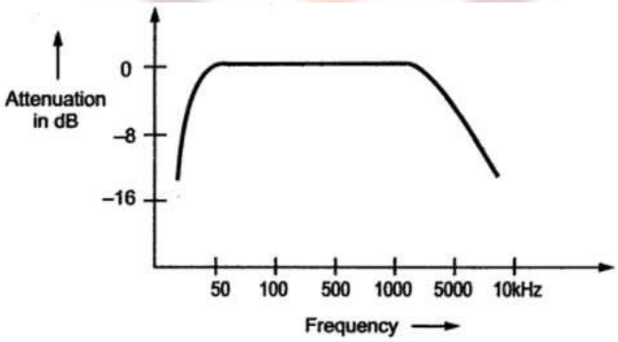


Fig.3. Typical Fidelity curve

Image Frequency and its Rejection

In standard broadcast receiver the local oscillator frequency is made higher than the signal frequency by an amount equal to intermediate frequency (IF). Therefore $f_o = f_s + f_i$. When f_o and f_s are mixed, the difference frequency, which is one of the by products, is equal to f_i only f_i is passed and amplified by the IF stage.

If a frequency f_{si} ($f_o + f_i$), i.e. $f_{si} = f_s + 2\,f_i$, appears at the input of the mixer then it will produce the sum and difference frequencies regardless of the inputs. Therefore, the mixer output will be the difference frequency at the IF value. The terms f_{si} is called the **image** frequency and is defined as the signal frequency plus twice the intermediate frequency. Unfortunately, this image frequency signal is also amplified by the IF amplifiers resulting in interference. This has the effect of two stations being received simultaneously and is naturally undesirable.

The rejection of an image frequency by a single tuned circuit is the ratio of the gain at the signal frequency to the gain at the image frequency. It is given by

$$\alpha = \sqrt{1 + Q^2 \rho^2}$$

where $\rho = \dfrac{f_{si}}{f_s} - \dfrac{f_s}{f_{si}}$

and Q = loaded Q of tuned circuit

If the receiver has an RF stage, then there are two tuned circuits, both tuned to f_s ; the rejection of each will be calculated by the same formula, and the total rejection will be the product of two.

The image rejection depends on the selectivity of the RF amplifier and tuned circuits and must be achieved before the IF stage. Once the spurious frequency enters the first IF amplifier, it becomes impossible to remove it from the wanted signal.

Double Spotting

The phenomenon of double spotting occurs at higher frequencies due to poor front end selectivity of the receivers. In this, receiver picks up same short-wave station at two nearby points on the receiver dial.

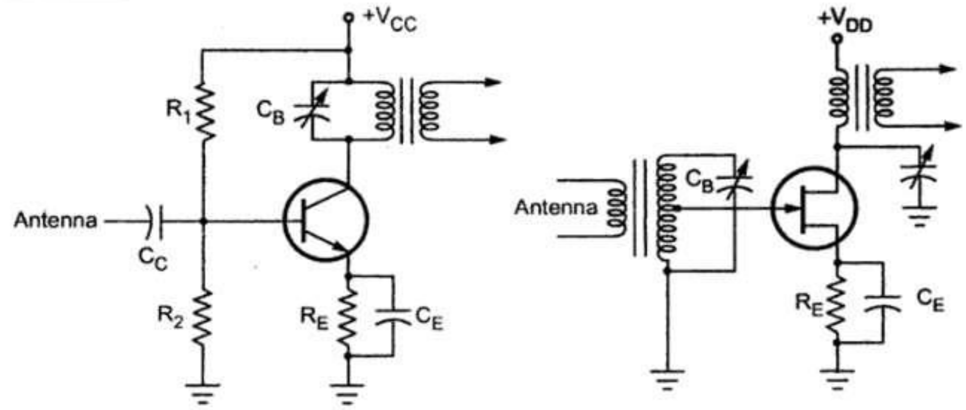
When the receiver is tuned across the band, a strong signal appears to be at two different frequencies, once at the desired frequency and again when the receiver is tuned to 2 times IF below the desired frequency. In this second case, the signal becomes the image, reduced in strength by the image rejection, thus making it appear that the signal is located at two frequencies in the band.

Blocks in Super heterodyne Receiver:

- **Basic principle**
  - Mixing
  - Intermediate frequency of 455 KHz
  - Ganged tuning
- **RF section**
  - Tuning circuits – reject interference and reduce noise figure
  - Wide band RF amplifier
- **Local Oscillator**
  - 995 KHz to 2105 KHz
  - Tracking
- **IF amplifier**
  - Very narrow band width Class A amplifier – selects 455 KHz only
  - Provides much of the gain
  - Double tuned circuits
- **Detector**
  - RF is filtered to ground

1. RF Amplifier:

RF amplifier provides initial gain and selectivity. Fig. 4 shows the RF amplifier circuits. It is a tuned circuit followed by an amplifier The RF amplifier is usually a simple class A circuit. A typical bipolar circuit is shown in Fig. 4. (a) , and a typical FET circuit is shown in Fig. 4. (b).



(a) (b)

The values of resistors R_1 and R_2 in the bipolar circuit are adjusted such that the amplifier works as class A amplifier. The antenna is connected through coupling capacitor to the base of the transistor. This makes the circuit very broad band as the transistor will amplify virtually any signal picked up by the antenna. However the collector is tuned with a parallel resonant circuit to provide the initial selectivity for the mixer input.

The FET circuit shown in Fig. 4 (b) is more effective than the transistor circuit. Their high input impedance minimizes the loading on tuned circuits, thereby permitting the Q of the circuit to be higher and selectivity to be sharper.

The receiver having an RF amplifier stage has following advantages :

1. It provides greater gain, i.e. better sensitivity.
2. It improves image-frequency rejection.
3. It improves signal to noise ratio.
4. It improves rejection of adjacent unwanted signals, providing better selectivity.
5. It provides better coupling of the receiver to the antenna.
6. It prevents spurious frequencies from entering the mixer and heterodyning there to produce an interfering frequency equal to the IF from the desired signal.
7. It also prevents reradiation of the local oscillator through the antenna of the receiver.

2. Mixer

The frequency converter is a nonlinear resistance having two sets of input terminals and one set of output terminal. The two inputs to the frequency converter are the input signal along with any modulation and the input from a local oscillator (LO). The output contains several frequencies including the difference between the input frequencies. The difference frequency is called intermediate frequency and output circuit of the mixer is tuned for the intermediate frequency.

Separately Excited Mixer:

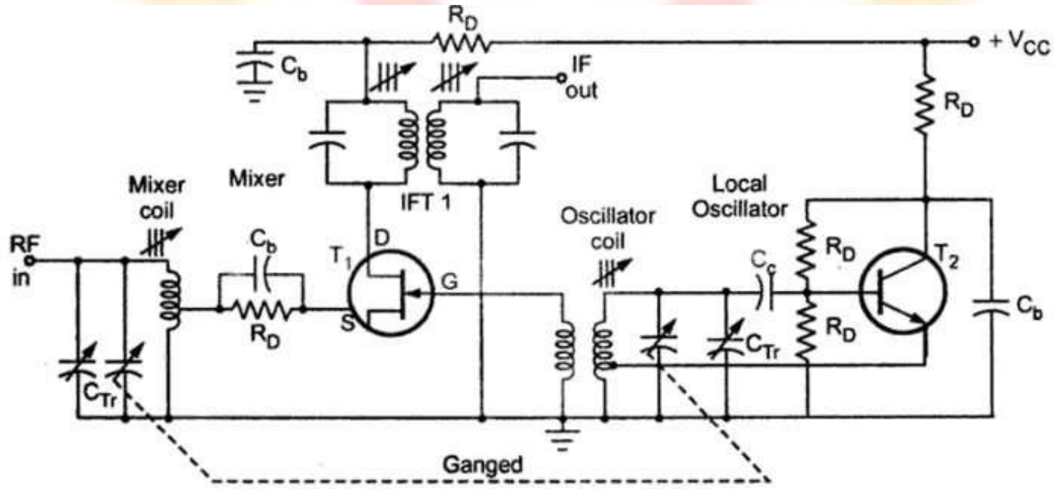


Fig.5 Separately Excited FET Mixer

Fig. 5 shows the separately excited mixer using FET. Here, one device acts as a mixer while the other supplies the necessary oscillations. The bipolar transistor T_2, forms the Hartley oscillator circuit. It oscillates with local frequency (f_o). FET T_1, is a mixer, whose gate is fed with the output of local oscillator and its bias is adjusted such that it operates in a nonlinear portion of its characteristic. The local oscillator varies the gate bias of the FET to vary its transconductance in a nonlinear manner, resulting intermediate frequency (IF) at the output. The output is taken through double tuned transformer in the drain of the mixer and fed to the IF amplifier. The ganged tuning capacitor allows simultaneous tuning of mixer and local oscillator.

The C_{Tr}, a small trimmer capacitors across each of the tuning capacitors are used for fine adjustments.

Self Excited Mixer:

It is possible to combine the function of the mixer and local oscillator in one circuit. The circuit is commonly known as self excited mixer. Fig. 6 shows self excited bipolar transistor mixer. The circuit oscillates and the transconductance of the transistor is varied in a nonlinear manner at the local oscillator rate. This variable transconductance (g_m) is used by the transistor to amplify the incoming RF signal.

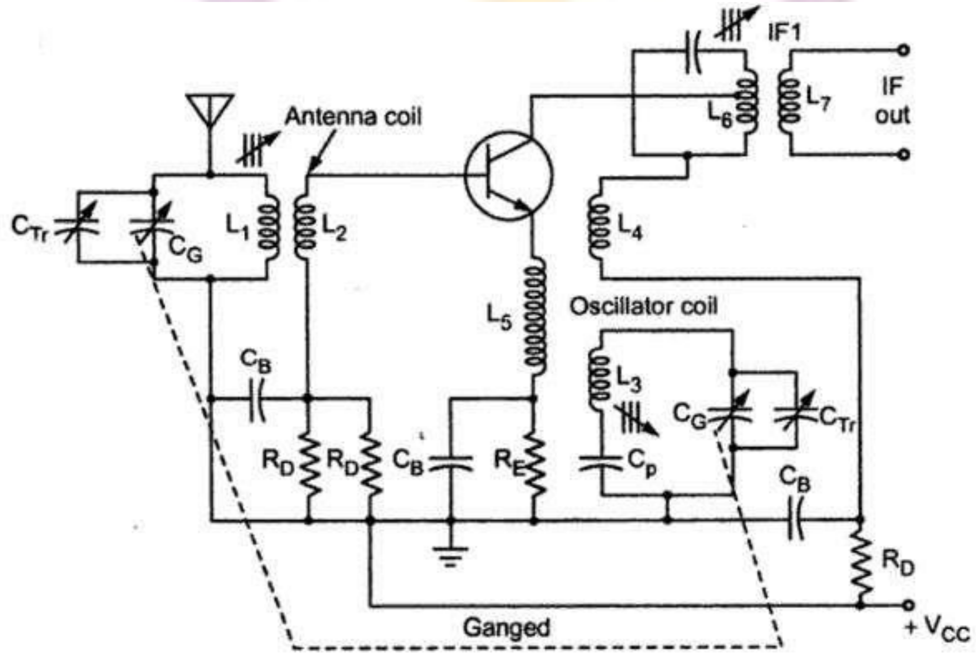


Fig.6. Self Excited Mixer

3. Tracking

The superheterodyne receiver has number of tunable circuits which must all be tuned correctly if any given station is to be received. The ganged tuning is employed to do this work, which mechanically couples all tuning circuits so that only one tuning control or dial is required. Usually, there are three tuned circuits : Antenna or RF tuned circuit, mixer tuned circuit and local oscillator tuned circuit. All these circuits must be tuned to get proper RF input and to get IF frequency at the output of mixer. The process of tuning circuits to get the desired output is called **Tracking**. Any error that exists in the frequency difference will result in an incorrect frequency being fed to the IF amplifier. Such errors are known as **'Tracking Errors'** and these must be avoided.

To avoid tracking errors standard capacitors are not used, and ganged capacitors with identical sections are used. A different value of inductance and special extra capacitors called trimmers and padders are used to adjust the capacitance of the oscillator to the proper range. There are three common methods used for tracking. These are

- Padder tracking
- Trimmer tracking

4. Local Oscillator

In shortwave broadcasting, the operating limit for receivers is 36 MHz. For such operating limit local oscillators such as Armstrong, Hartley, Colpitts, Clapp or ultra-audion are used. The Colpitts, Clapp and ultra-audion oscillators are used at the top of the operating limit, whereas Hartley oscillator is used for frequencies below 120 MHz. All these oscillators are LC oscillators and each employs only one tuned circuit to defermine its frequency. When higher frequency stability of local oscillator is required, the circuits like AFC (Automatic Frequency Control) are used.

5. IF Amplifier

IF amplifiers are tuned voltage amplifiers tuned for the fixed frequency. Its important function is to amplify only tuned frequency signal and reject all others. As we know, most of the receiver gain is provided by the IF amplifiers, to obtain required gain, usually two or more stages of IF amplifiers are required.

Fig. 7 shows the two stage IF amplifier. Two stages are transformer coupled and all IF transformers are single tuned, i.e. tuned for single frequency.

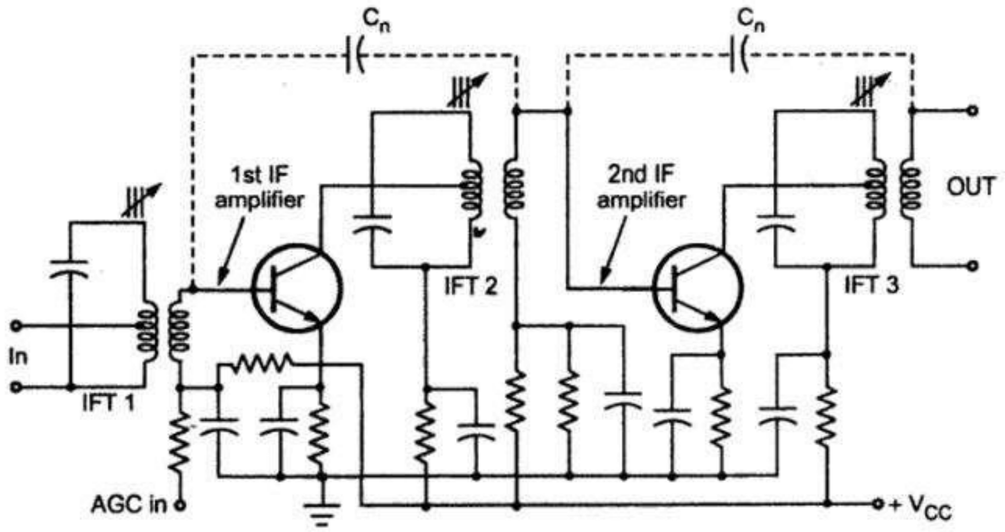


Fig.7 Two Stage IF Amplifier

Choice of Intermediate Frequency:

Selection of the intermediate frequency depends on various factors. While choosing the intermediate frequency it is necessary to consider following factors.

1. Very high intermediate frequency will result in poor selectivity and poor adjacent channel rejection.
2. A high value of intermediate frequency increases tracking difficulties.
3. At low values of intermediate frequency, image frequency rejection is poor.
4. At very low values of intermediate frequency, selectivity is too sharp. Cutting off the sidebands.
5. At very low IF, the frequency stability of the local oscillator must be correspondingly high because any frequency drift is now a larger proportion of the low IF than of a high IF.
6. The IF must not fall in the tuning range of the receiver, otherwise instability will occur and heterodyne whistles will be heard, making it impossible to tune to the frequency band immediately adjacent to the intermediate frequency.

With the above considerations the standard broadcast AM receivers [tuning to 540 to 1650 kHz] use an IF within the 438 kHz to 465 kHz range. The 465 kHz IF is most commonly used.

6. Automatic Gain Control

Automatic Gain Control is a system by means of which the overall gain of a radio receiver is varied automatically with the variations in the strength of the receiver signal, to maintain the output substantially constant. AGC circuitry derives the dc bias voltage from the output of the detector. It applies this derived dc bias voltage to a selected number of RF, IF and mixer stages to control their gains. When the average signal level increases, the size of the AGC bias increases, and the gain of the controlled stages decreases. When there is no signal, there is a minimum AGC bias, and the amplifiers produce maximum gain. There are two types of AGC circuits in use : Simple AGC and Delayed AGC

Simple AGC

In simple AGC receivers the AGC bias starts to increase as soon as the received signal level exceeds the background noise level. As a result receiver gain starts falling down, reducing the sensitivity of the receiver.

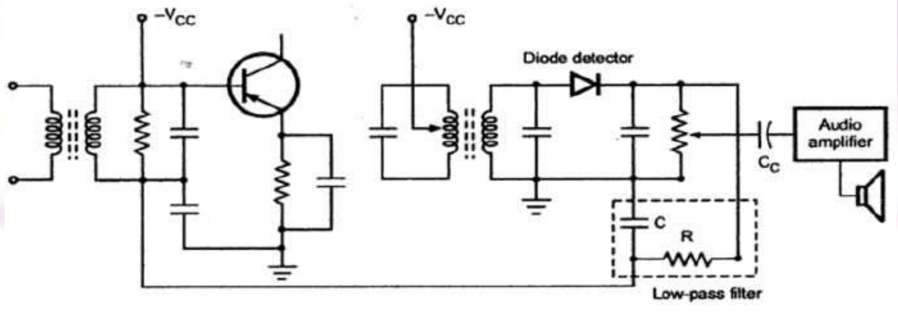


Fig.8. Simple AGC circuit

Fig. 8 shows the simple AGC circuit. In this circuit, dc bias produced by half wave rectifier as a AM detector, is used to control the gain of RF or IF amplifier. Before application of this voltage to the base of the RF and / or IF stage amplifier the audio signal is removed by the lowpass filter. The time constant of the filter is kept at least 10 times longer than the period of the lowest modulation frequency received. If the time constant is kept longer, it will give better filtering, but it will cause an annoying delay in the application of the AGC control when tuning from one signal to another. The recovered signal is then passed through C_C to remove the dc. The resulting ac signal is further amplified and applied to the loudspeaker.

Delayed AGC

Simple AGC is clearly an improvement over no AGC at all. Unfortunately, in simple AGC circuit, the unwanted weak signals(noise signals) are amplified with high gain. To

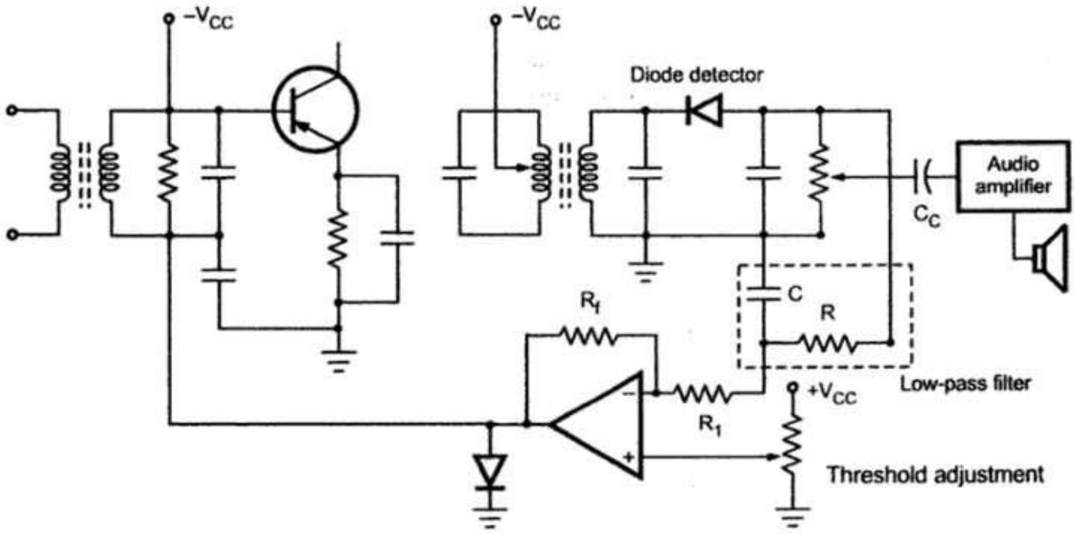


Fig.9. Delayed AGC circuit

avoid this, in delayed AGC circuits, AGC bias is not applied to amplifiers until signal strength has reached a predetermined level, after which AGC bias is applied as with simple AGC, but more strongly.

Here, AGC output is applied to the difference amplifier. It gives negative dc AGC only when AGC output generated by diode detector is above certain dc threshold voltage. This threshold voltage can be adjusted by adjusting the voltage at the positive input of the operational amplifier.

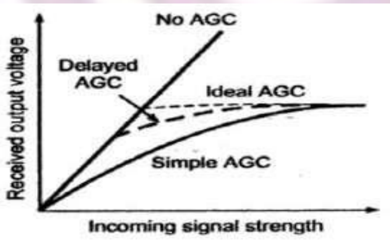


Fig.10. Response of receiver with various AGC circuits.

FM Receiver:

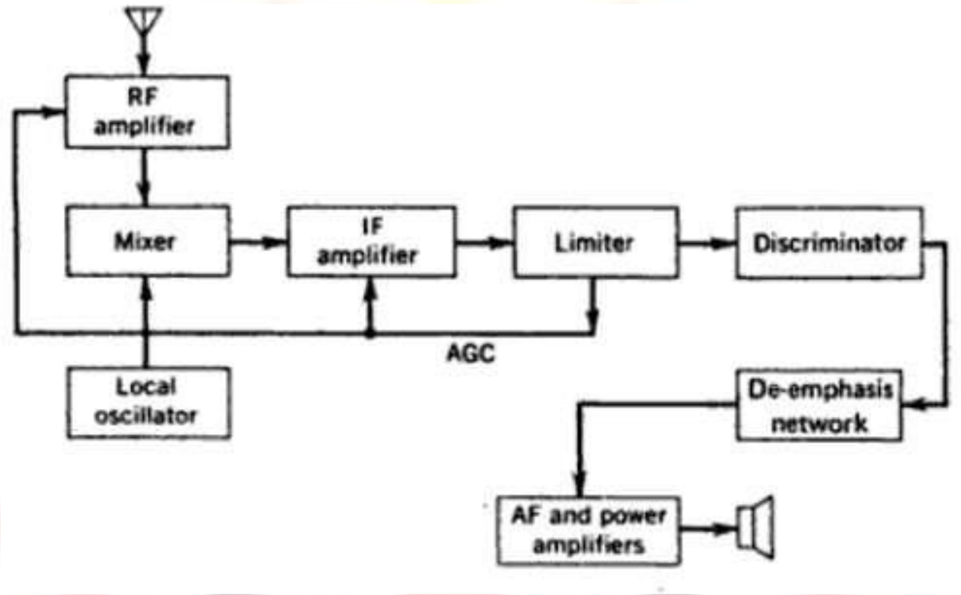


Fig.11. FM Receiver Block diagram

The FM receiver is a superheterodyne receiver, and the block diagram of Figure 11 shows just how similar it is to an AM receiver. The basic differences are as follows:

1. Generally much higher operating frequencies in FM
2. Need for limiting and de-emphasis in FM
3. Totally different methods of demodulation
4. Different methods of obtaining AGC

Comparisons with AM Receivers

A number of sections of the FM receiver correspond exactly to those of other receivers already discussed. The same criteria apply in the selection of the intermediate frequency, and IF amplifiers are basically similar. A number of concepts have very similar meanings so that only the differences and special applications need be pointed out.

RF amplifiers An RF amplifier is always used in an FM receiver. Its main purpose is to reduce the noise figure, which could otherwise be a problem because of the large bandwidths needed for FM. It is also required to match the input impedance of the receiver to that of the antenna. To meet the second requirement, grounded gate (or base) or cascode amplifiers are employed. Both types have the property of low input impedance and matching the antenna, while neither requires neutralization. This is because the input electrode is grounded on either type of amplifier, effectively isolating input from output. A typical FET grounded-gate RF amplifier is shown in Figure
It has all the good points mentioned and the added features of low distortion and simple operation.

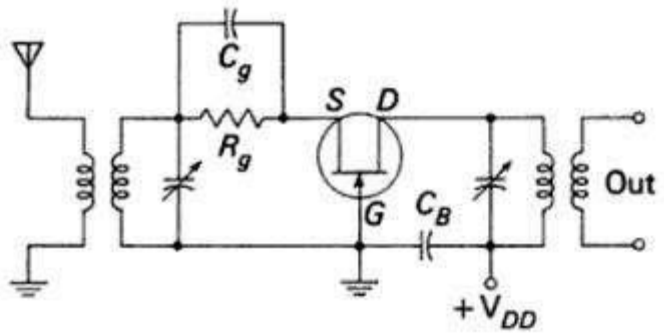


FIGURE **Grounded-gate FET RF amplifier.**

Oscillators and mixers The oscillator circuit takes any of the usual forms, with the Colpitts and Clapp predominant, being suited to VHF operation. Tracking is not normally much of a problem in FM broadcast receivers. This is because the tuning frequency range is only 1.25:1, much less than in AM broadcasting.

A very satisfactory arrangement for the front end of an FM receiver consists of FETs for the RF amplifier and mixer, and a bipolar transistor oscillator. As implied by this statement, separately excited oscillators are normally used

Intermediate frequency and IF amplifiers Again, the types and operation do not differ much from their AM counterparts. It is worth noting, however, that the intermediate frequency and the bandwidth required are far higher than in AM broadcast receivers. Typical figures for receivers operating in the 88- to 108-MHz band are an IF of 10.7 MHz and a bandwidth of 200 kHz. As a consequence of the large bandwidth, gain per stage may be low. Two IF amplifier stages are often provided, in which case the shrinkage of bandwidth as stages are cascaded must be taken into account.

Double limiter

A double limiter consists of two amplitude limiters in cascade, an arrangement that increases the limiting range very satisfactorily. Numerical values given to illustrate limiter performance showed an output voltage (all values peak-to-peak, as before) of 5 V for any input within the 0.4- to 4-V range, above which output gradually decreases. It is quite possible that an output of 0.6 V is not reached until the input to the first limiter is about 20 V. If the range of the second limiter is 0.6 to 6 V, it follows that all voltages between 0.4 and 20 V fed to the double limiter will be limited. The use of the double limiter is seen to have increased the limiting range quite considerably.

Automatic gain control (AGC)

A suitable alternative to the second limiter is automatic gain control. This is to ensure that the signal fed to the limiter is within its limiting range, regardless of the input signal strength, and also to prevent overloading of the last IF amplifier. If the limiter used has leak-type bias, then this bias voltage will vary in proportion to the input voltage (as shown in Figure 6-31) and may therefore be used for AGC. Sometimes a separate AGC detector is used, which takes part of the output of the last IF amplifier and rectifies and filters it in the usual manner.

Amplitude Limiter:

In order to make full use of the advantages offered by FM, a demodulator must be preceded by an amplitude limiter, on the grounds that any amplitude changes in the signal fed to the FM demodulator are spurious

They must therefore be removed if distortion is to be avoided. The point is significant, since most FM demodulators react to amplitude changes as well as frequency changes. The limiter is a form of clipping device, a circuit whose output tends to remain constant despite changes in the input signal. Most limiters behave in this fashion, provided that the input voltage remains within a certain range. The common type of limiter uses two separate electrical effects to provide a relatively constant output. There are leak-type bias and early (collector) saturation.

Operation of the amplitude limiter Figure shows a typical FET amplitude limiter. Examination of the dc conditions shows that the drain supply voltage has been dropped through resistor R_D. Also, the bias on the gate is leak-type bias supplied by the parallel $R_g - C_g$ combination. Finally, the FET is shown neutralized by means of capacitor C_N, in consideration of the high frequency of operation.

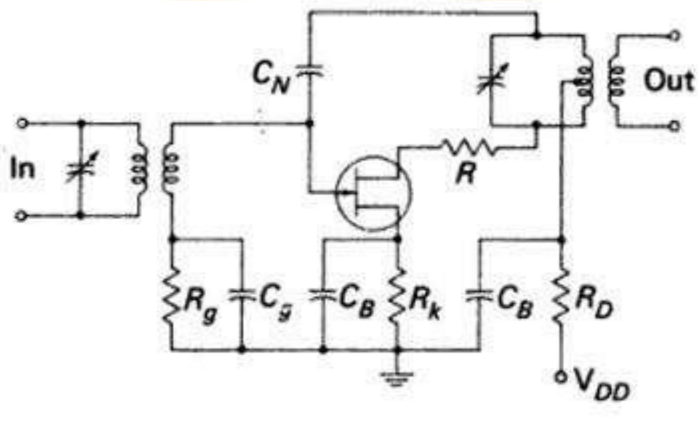


FIGURE **Amplitude limiter.**

Leak-type bias provides limiting, as shown in Figure . When input signal voltage rises, current flows in the $R_g - C_g$ bias circuit, and a negative voltage is developed across the capacitor. It is seen that the bias on the FET is increased in proportion to the size of the input voltage. As a result, the gain of the amplifier is lowered, and the output voltage tends to remain constant.

Although some limiting is achieved by this process, it is insufficient by itself, the action just described would occur only with rather large input voltages. To overcome this, early saturation of the output current is used, achieved by means of a low drain supply voltage. This is the reason for the drain dropping resistor of Figure . The supply voltage for a limiter is typically one-half of the normal dc drain voltage. The result of early saturation is to ensure limiting for conveniently low input voltages.

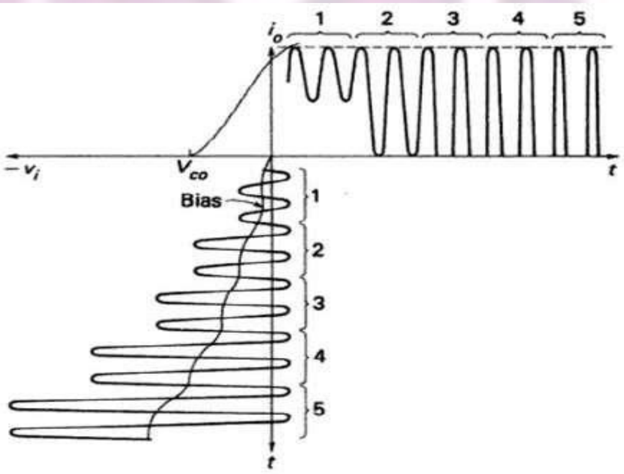


FIGURE **Amplitude limiter transfer characteristic.**

It is possible for the gate-drain section to become forward-biased under saturation conditions, causing a short circuit between input and output. To avert this, a resistance of a few hundred ohms is placed between the drain and its tank. This is *R* of Figure

UNIT-IV

PULSE MODULATION

Introduction:

Pulse Modulation

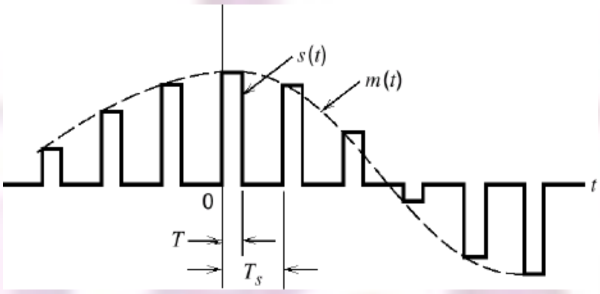
- Carrier is a train of pulses
- Example: Pulse Amplitude Modulation (PAM), Pulse width modulation (PWM) , Pulse Position Modulation (PPM)

Types of Pulse Modulation:

- The immediate result of sampling is a pulse-amplitude modulation (PAM) signal
- PAM is an analog scheme in which the amplitude of the pulse is proportional to the amplitude of the signal at the instant of sampling
- Another analog pulse-forming technique is known as **pulse-duration modulation (PDM).** This is also known as **pulse-width modulation (PWM)**
- **Pulse-position modulation** is closely related to PDM

Pulse Amplitude Modulation:

In PAM, amplitude of pulses is varied in accordance with instantaneous value of modulating signal.



PAM Generation:

The carrier is in the form of narrow pulses having frequency fc. The uniform sampling takes place in multiplier to generate PAM signal. Samples are placed Ts sec away from each other.

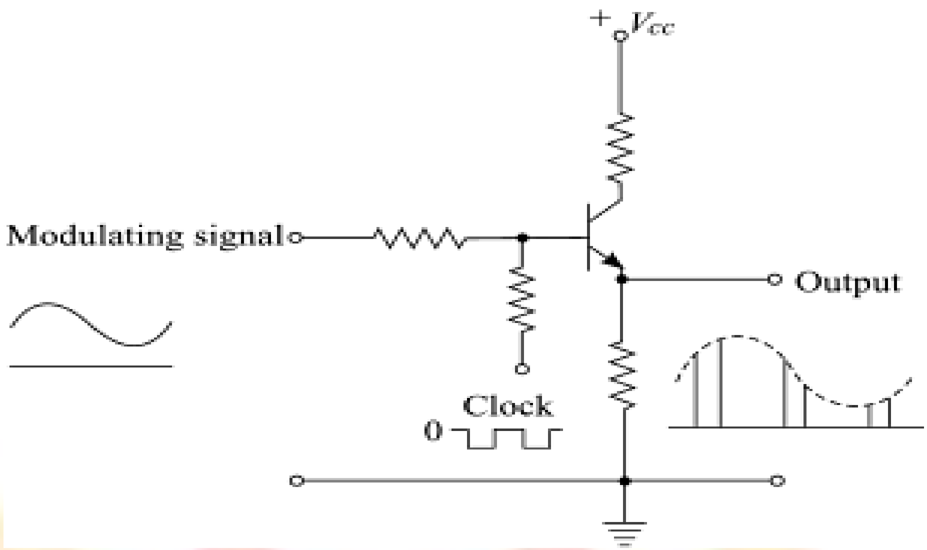


Figure PAM Modulator

- The circuit is simple emitter follower.
- In the absence of the clock signal, the output follows input.
- The modulating signal is applied as the input signal.
- Another input to the base of the transistor is the clock signal.
- The frequency of the clock signal is made equal to the desired carrier pulse train frequency.
- The amplitude of the clock signal is chosen the high level is at ground level(0v) and low level at some negative voltage sufficient to bring the transistor in cutoff region.
- When clock is high, circuit operates as emitter follower and the output follows in the input modulating signal.
- When clock signal is low, transistor is cutoff and output is zero.
- Thus the output is the desired PAM signal.

PAM Demodulator:

- The PAM demodulator circuit which is just an envelope detector followed by a second order op-amp low pass filter (to have good filtering characteristics) is as shown below

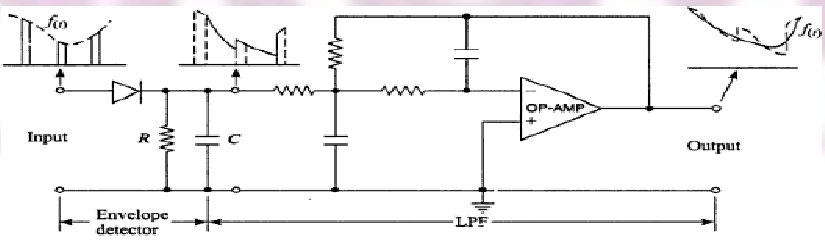
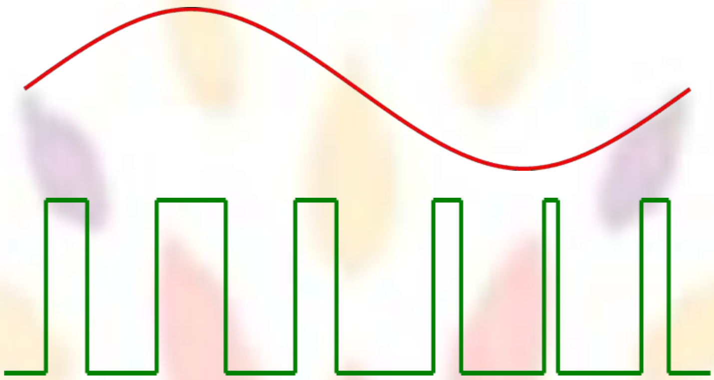


Figure PAM Demodulator

Pulse Width Modulation:

- In this type, the amplitude is maintained constant but the width of each pulse is varied in accordance with instantaneous value of the analog signal.



- In PWM information is contained in width variation. This is similar to FM.
- In pulse width modulation (PWM), the width of each pulse is made directly proportional to the amplitude of the information signal.

Pulse Position Modulation:

- In this type, the sampled waveform has fixed amplitude and width whereas the position of each pulse is varied as per instantaneous value of the analog signal.
- PPM signal is further modification of a PWM signal.

PPM & PWM Modulator:

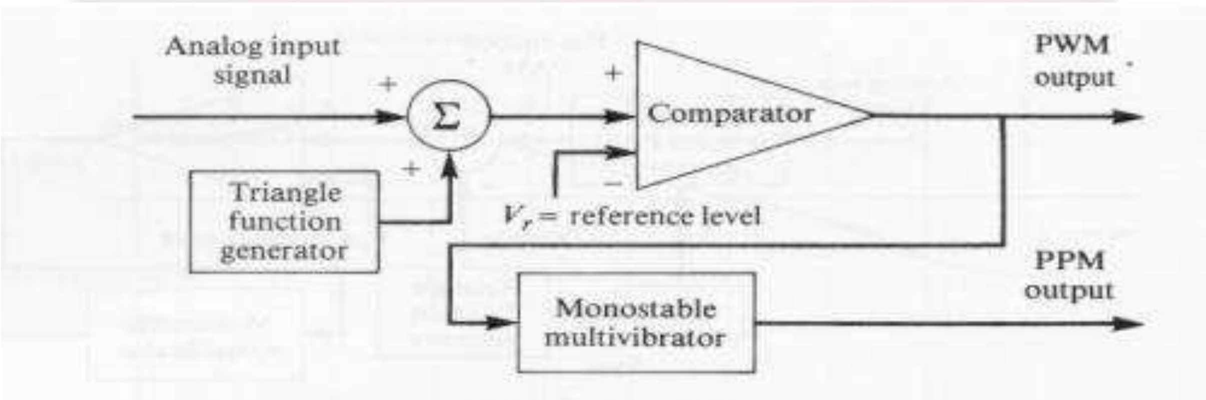


Figure PWM & PPM Modulator

- The PPM signal can be generated from PWM signal.
- The PWM pulses obtained at the comparator output are applied to a mono stable multi

vibrator which is negative edge triggered.

- Hence for each trailing edge of PWM signal, the monostable output goes high. It remains high for a fixed time decided by its RC components.
- Thus as the trailing edges of the PWM signal keeps shifting in proportion with the modulating signal, the PPM pulses also keep shifting.
- Therefore all the PPM pulses have the same amplitude and width. The information is conveyed via changing position of pulses.

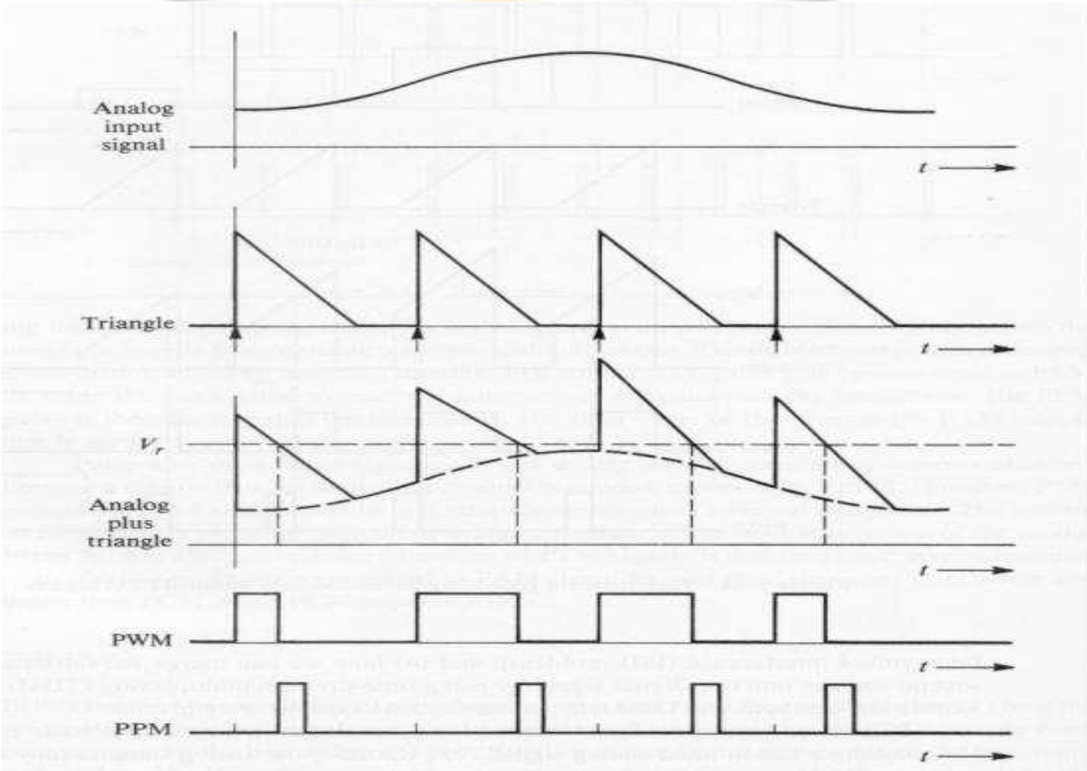


Figure PWM & PPM Modulation waveforms

PWM Demodulator:

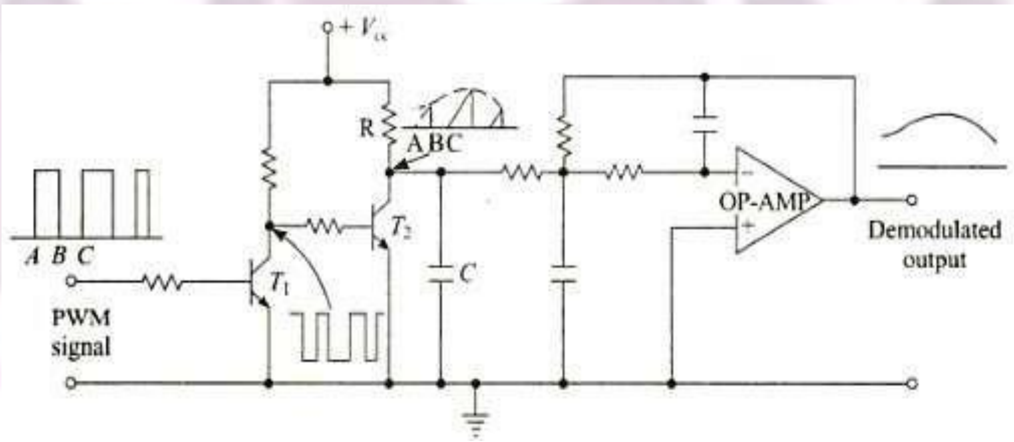


Figure PWM Demodulator

- Transistor T1 works as an inverter.
- During time interval A-B when the PWM signal is high the input to transistor T2 is low.
- Therefore, during this time interval T2 is cut-off and capacitor C is charged through an R-C combination.
- During time interval B-C when PWM signal is low, the input to transistor T2 is high, and it gets saturated.
- The capacitor C discharges rapidly through T2. The collector voltage of T2 during B-C is low.
- Thus, the waveform at the collector of T2is similar to saw-tooth waveform whose envelope is the modulating signal.
- Passing it through 2nd order op-amp Low Pass Filter, gives demodulated signal.

PPM Demodulator:

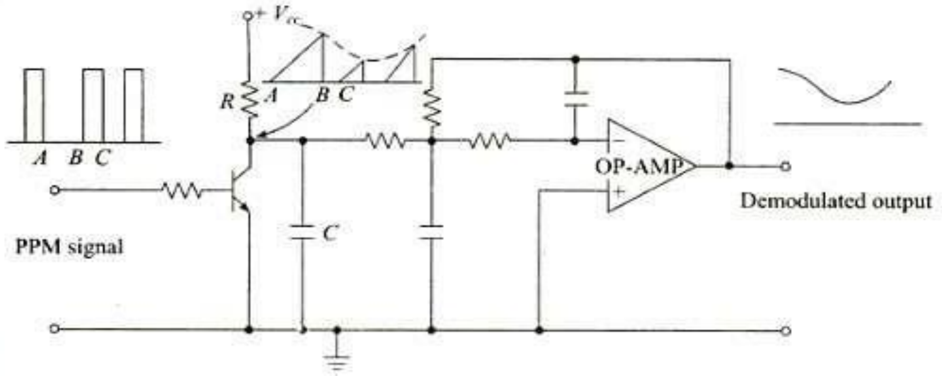


Figure PPM Demodulator

- The gaps between the pulses of a PPM signal contain the information regarding the modulating signal.
- During gap A-B between the pulses the transistor is cut-off and the capacitor C gets charged through R-C combination.
- During the pulse duration B-C the capacitor discharges through transistor and the collector voltage becomes low.
- Thus, waveform across collector is saw-tooth waveform whose envelope is the modulating signal.
- Passing it through 2nd order op-amp Low Pass Filter, gives demodulated signal.

Multiplexing

Multiplexing is the set of techniques that allows the simultaneous transmission of multiple signals across a single common communications channel.

Multiplexing is the transmission of analog or digital information from one or more sources to one or more destination over the same transmission link.

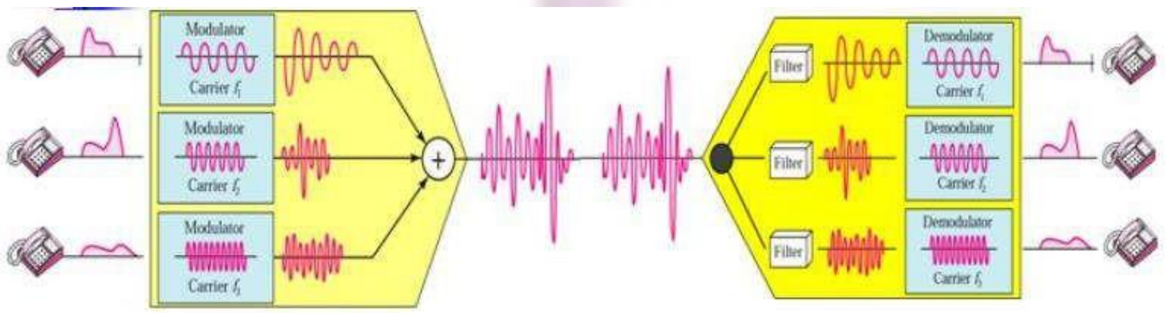
Although transmissions occur on the same transmitting medium, they do not necessarily occupy the same bandwidth or even occur at the same time.

Frequency Division Multiplexing

Frequency division multiplexing (FDM) is a technique of multiplexing which means combining more than one signal over a shared medium. In FDM, signals of different frequencies are combined for concurrent transmission.

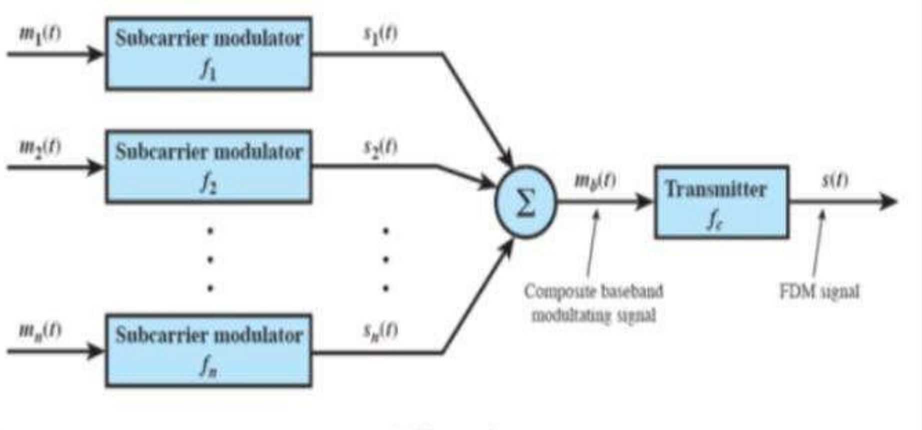
In FDM, the total bandwidth is divided to a set of frequency bands that do not overlap. Each of these bands is a carrier of a different signal that is generated and modulated by one of the sending devices. The frequency bands are separated from one another by strips of unused frequencies called the guard bands, to prevent overlapping of signals.

The modulated signals are combined together using a multiplexer (MUX) in the sending end. The combined signal is transmitted over the communication channel, thus allowing multiple independent data streams to be transmitted simultaneously. At the receiving end, the individual signals are extracted from the combined signal by the process of demultiplexing (DEMUX).



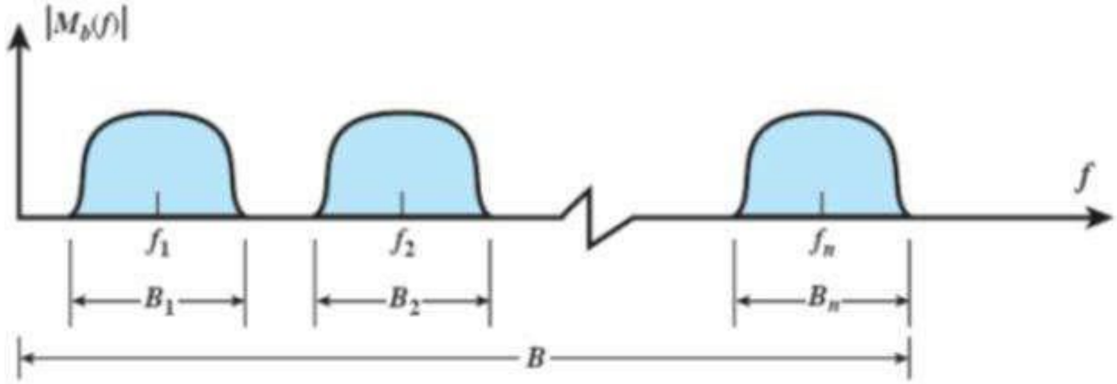
FDM system Transmitter

- Analog or digital inputs: m_i (t); i = 1,2, ... n
- Each input modulates a subcarrier of frequency fi; i=1, 2, ... n
- Signals are summed to produce a composite baseband signal denoted as $m_b(t)$
- f_i is chosen such that there is no overlap.



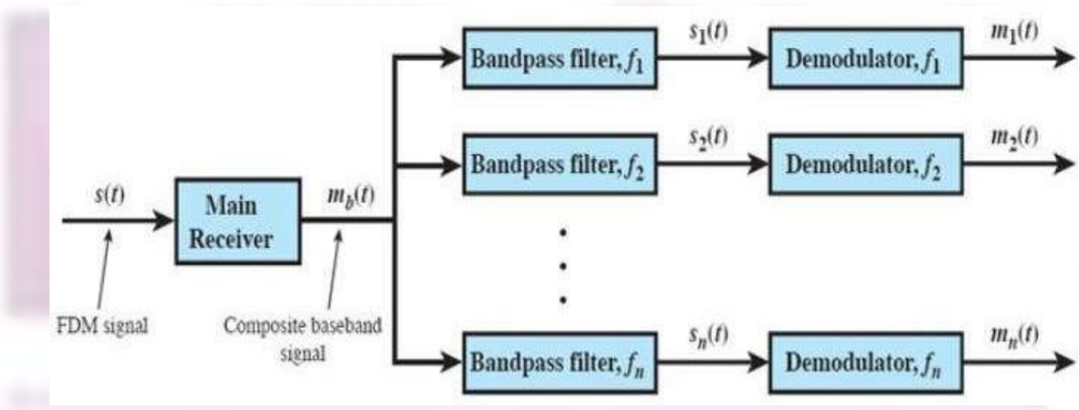
(a) Transmitter

Spectrum of composite baseband modulating signal



FDM system Receiver

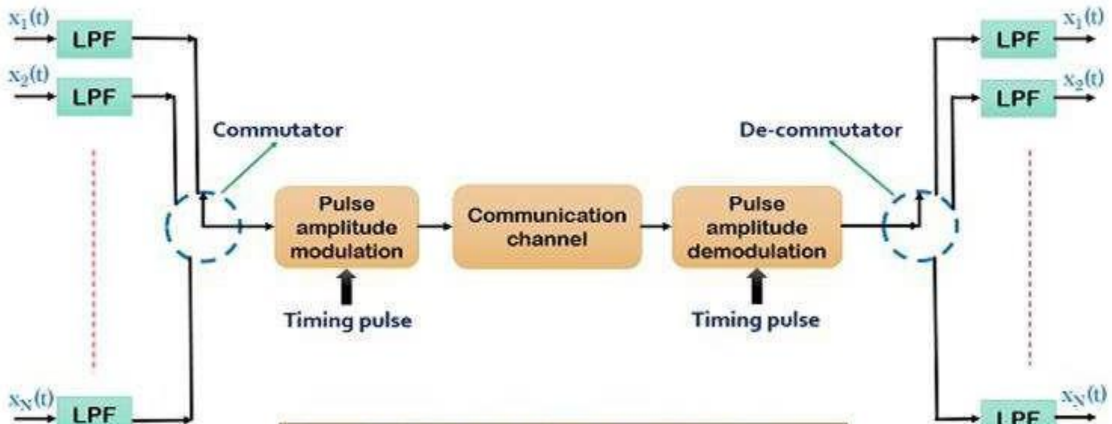
- The Composite base band signal $m_b(t)$ is passed through n band pass filters with response centred on f_i
- Each $s_i(t)$ component is demodulated to recover the original analog/digital data.



Time Division Multiplexing

TDM technique combines time-domain samples from different message signals (sampled at same rate) and transmits them together across the same channel.

The multiplexing is performed using a commutator (switch). At the receiver a decommutator (switch) is used in synchronism with the commutator to demultiplex the data.



Block diagram of PAM/TDM system

The input signals, all band limited to fm (max) by the LPFs are sequentially sampled at the transmitter by a commutator.

The Switch makes one complete revolution in Ts,(1/fs) extracting one sample from each input. Hence the output is a PAM waveform containing the individual message sampled periodically interlaced in time.

A set of pulses consisting of one sample from each input signal is called a frame.

At the receiver the de-commutator separates the samples and distributes them to a bank of LPFs, which in turn reconstruct the original messages.

Synchronizing is provided to keep the de-commutator in step with the commutator.

Elements of Digital Communication Systems

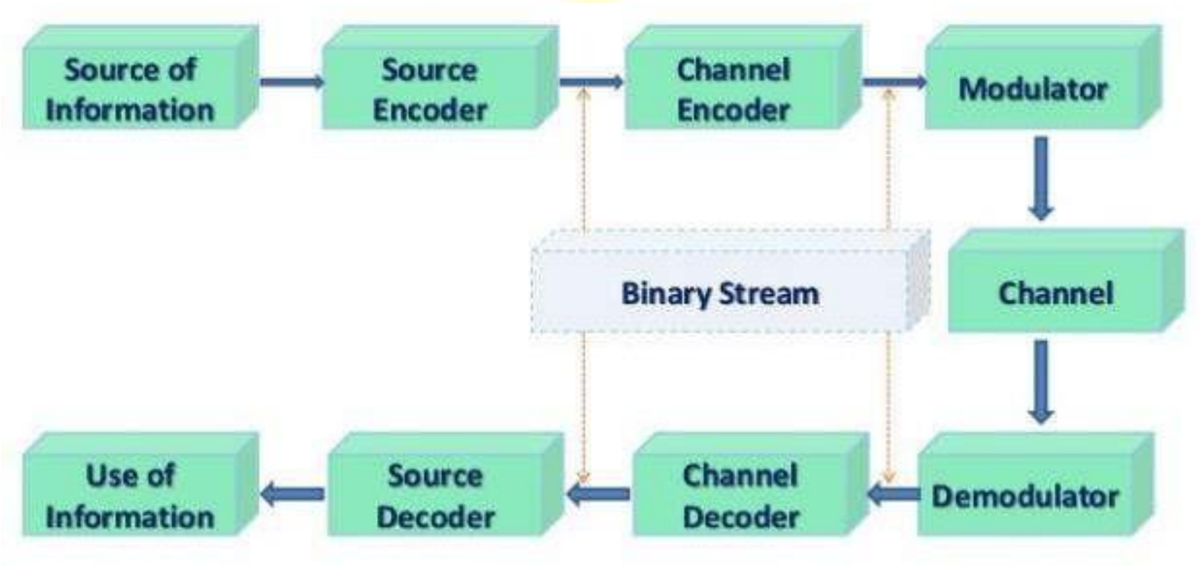


Figure Elements of Digital Communication Systems

1. **Information Source and Input Transducer:**

   The source of information can be analog or digital, e.g. analog: audio or video signal, digital: like teletype signal. In digital communication the signal produced by this source is converted into digital signal which consists of 1′s and 0′s. For this we need a source encoder.

2. **Source Encoder:**

   In digital communication we convert the signal from source into digital signal as mentioned above. The point to remember is we should like to use as few binary digits as possible to represent the signal. In such a way this efficient representation of the source output results in little or no redundancy. This sequence of binary digits is called ***information sequence.***

   Source Encoding or Data Compression: the process of efficiently converting the output of whether analog or digital source into a sequence of binary digits is known as source encoding.

3. **Channel Encoder:**

   The information sequence is passed through the channel encoder. The purpose of the channel encoder is to introduce, in controlled manner, some redundancy in the binary information sequence that can be used at the receiver to overcome the effects of noise and interference encountered in the transmission on the signal through the channel.

   For example take k bits of the information sequence and map that k bits to unique n bit sequence called code word. The amount of redundancy introduced is measured by the ratio *n/k* and the reciprocal of this ratio (k/n) is known as *rate of code or code rate.*

4. **Digital Modulator:**
   The binary sequence is passed to digital modulator which in turns convert the sequence into electric signals so that we can transmit them on channel (we will see channel later). The digital modulator maps the binary sequences into signal wave forms , for example if we represent 1 by sin x and 0 by cos x then we will transmit sin x for 1 and cos x for 0. (a case similar to BPSK)
5. **Channel:**
   The communication channel is the physical medium that is used for transmitting signals from transmitter to receiver. In wireless system, this channel consists of atmosphere , for traditional telephony, this channel is wired , there are optical channels, under water acoustic channels etc.We further discriminate this channels on the basis of their property and characteristics, like AWGN channel etc.
6. **Digital Demodulator:**
   The digital demodulator processes the channel corrupted transmitted waveform and reduces the waveform to the sequence of numbers that represents estimates of the transmitted data symbols.
7. **Channel Decoder:**
   This sequence of numbers then passed through the channel decoder which attempts to reconstruct the original information sequence from the knowledge of the code used by the channel encoder and the redundancy contained in the received data

Note: The average probability of a bit error at the output of the decoder is a measure of the performance of the demodulator – decoder combination.

8. **Source Decoder:**
   At the end, if an analog signal is desired then source decoder tries to decode the sequence from the knowledge of the encoding algorithm. And which results in the approximate replica of the input at the transmitter end.

9. **Output Transducer:**
   Finally we get the desired signal in desired format analog or digital.

Advantages of digital communication

- Can **withstand channel noise and distortion** much better as long as the noise and the distortion are within limits.
- **Regenerative repeaters** prevent accumulation of noise along the path.
- Digital **hardware implementation is flexible.**
- Digital signals **can be coded** to yield extremely **low error rates**, **high fidelity** and well as **privacy**.
- Digital communication is inherently more efficient than analog in realizing the exchange of SNR for bandwidth.

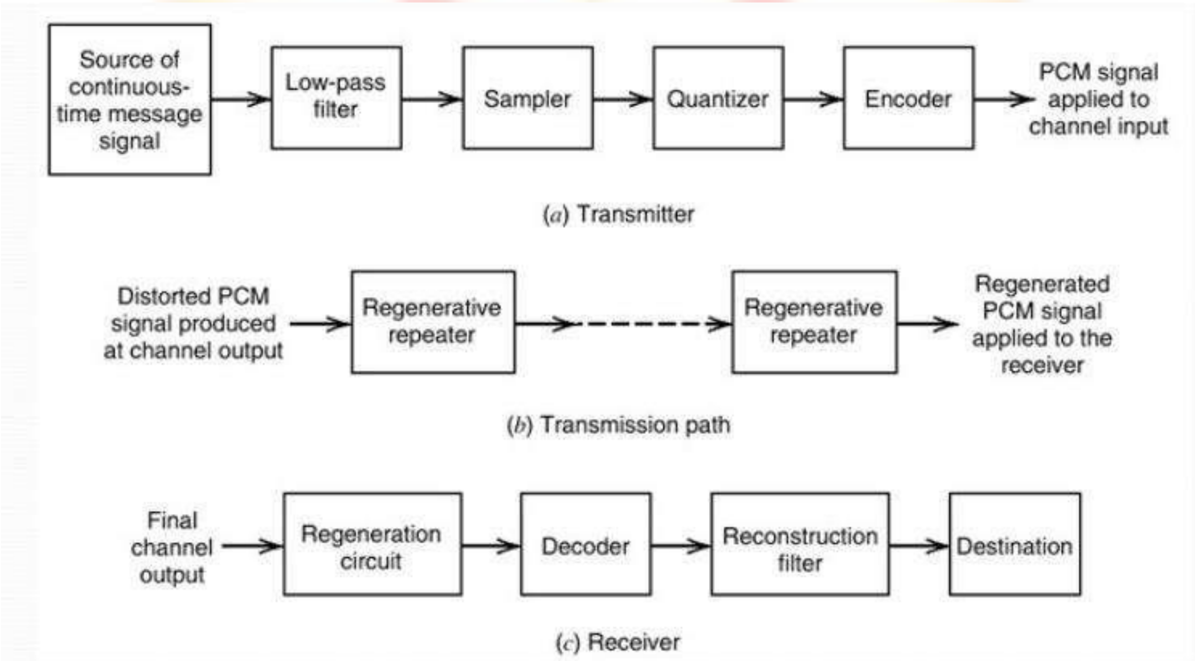
- It is easier and more **efficient to multiplex** several digital signals.

- Digital signal **storage is relatively easy** and **inexpensive.**
- **Reproduction** with digital messages is extremely reliable **without deterioation.**
- The **cost** of digital hardware continues to halve every two or three years while **performance or capacity doubles** over the same time period.

Disadvantages

- TDM digital transmission is not compatible with FDM
- A Digital system requires large bandwidth.

Elements of PCM System



Sampling:

- Process of converting analog signal into discrete signal.
- Sampling is common in all pulse modulation techniques
- The signal is sampled at regular intervals such that each sample is proportional to amplitude of signal at that instant
- Analog signal is sampled every T_s $Secs$, called sampling interval. f_s=1/T_S is called sampling rate or sampling frequency.
- f_s=2f_m is Min. sampling rate called **Nyquist rate.** Sampled spectrum (ω) is repeating periodically without overlapping.
- Original spectrum is centered at ω=0 and having bandwidth of ω_m. Spectrum can be recovered by passing through low pass filter with cut-off ω_m.
- For f_s<2f_m sampled spectrum will overlap and cannot be recovered back. This is called **aliasing.**

Sampling methods:

- Ideal – An impulse at each sampling instant.
- Natural – A pulse of Short width with varying amplitude.
- Flat Top – Uses sample and hold, like natural but with single amplitude value.

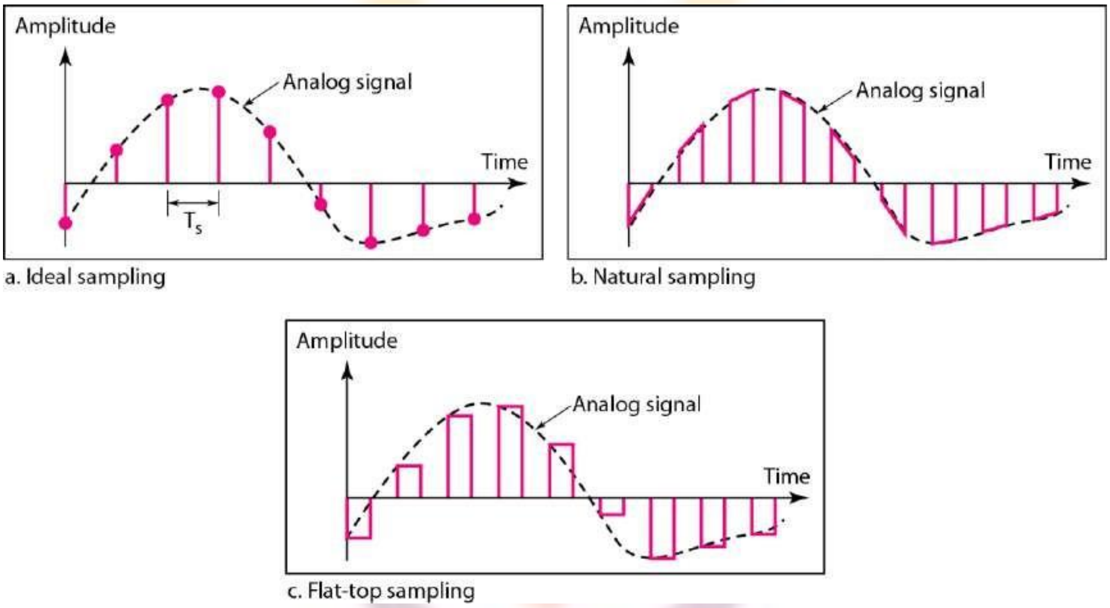


Fig. 4 Types of Sampling

Sampling of band-pass Signals:

- A band-pass signal of bandwidth $2f_m$ can be completely recovered from its samples.

  Min. sampling rate $=2\times Bandwidth$

  $=2\times 2f_m=4f_m$

- Range of minimum sampling frequencies is in the range of $2\times BW$ *to* $4\times BW$

Instantaneous Sampling or Impulse Sampling:

- Sampling function is train of spectrum remains constant impulses throughout frequency range. It is not practical.

Natural sampling:

- The spectrum is weighted by a **sinc** function.
- Amplitude of high frequency components reduces.

Flat top sampling:

- Here top of the samples remains constant.
- In the spectrum high frequency components are attenuated due sinc pulse roll off. This is known as **Aperture effect.**
- If pulse width increases aperture effect is more i.e. more attenuation of high frequency components.

PCM Generator

The pulse code modulator technique samples the input signal $x(t)$ at frequency $f_s \geq 2W$. This sampled 'Variable amplitude' pulse is then digitized by the analog to digital converter. The parallel bits obtained are converted to a serial bit stream. Fig. 8 shows the PCM generator.

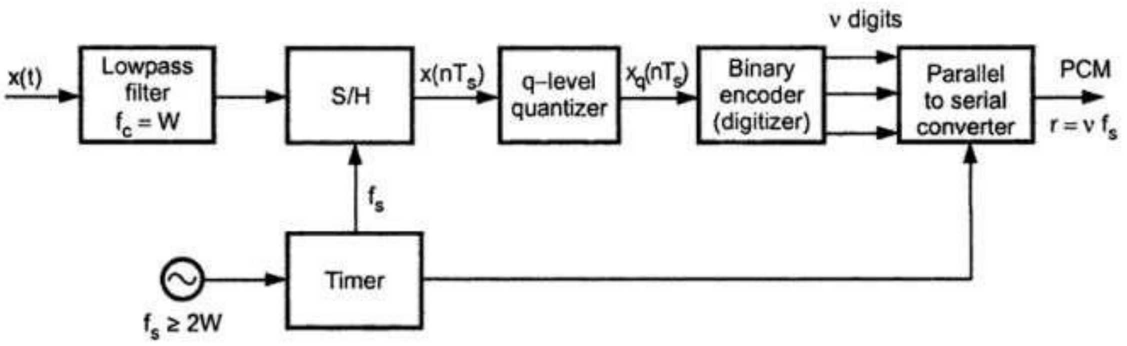


Fig. 8 PCM generator

In the PCM generator of above figure, the signal $x(t)$ is first passed through the lowpass filter of cutoff frequency 'W' Hz. This lowpass filter blocks all the frequency components above 'W' Hz. Thus $x(t)$ is bandlimited to 'W' Hz. The sample and hold circuit then samples this signal at the rate of f_s. Sampling frequency f_s is selected sufficiently above Nyquist rate to avoid aliasing i.e.,

$$f_s \geq 2W$$

In Fig. 8 output of sample and hold is called $x(nT_s)$. This $x(nT_s)$ is discrete in time and continuous in amplitude. A q-level quantizer compares input $x(nT_s)$ with its fixed digital levels. It assigns any one of the digital level to $x(nT_s)$ with its fixed digital levels. It then assigns any one of the digital level to $x(nT_s)$ which results in minimum distortion or error. This error is called *quantization error*. Thus output of quantizer is a digital level called $x_q(nT_s)$.

Now coming back to our discussion of PCM generation, the quantized signal level $x_q(nT_s)$ is given to binary encoder. This encoder converts input signal to 'v' digits binary word. Thus $x_q(nT_s)$ is converted to 'V' binary bits. The encoder is also called digitizer.

It is not possible to transmit each bit of the binary word separately on transmission line. Therefore '*v*' binary digits are converted to serial bit stream to generate single baseband signal. In a parallel to serial converter, normally a shift register does this job. The output of PCM generator is thus a single baseband signal of binary bits.

An oscillator generates the clocks for sample and hold an parallel to serial converter. In the pulse code modulation generator discussed above ; sample and hold, quantizer and encoder combinely form an analog to digital converter.

Transmission BW in PCM

Let the quantizer use 'v' number of binary digits to represent each level. Then the number of levels that can be represented by 'v' digits will be,

$$q = 2^v \qquad \dots \quad 1$$

Here 'q' represents total number of digital levels of q-level quantizer.

For example if $v = 3$ bits, then total number of levels will be,

$$q = 2^3 = 8 \text{ levels}$$

Each sample is converted to 'v' binary bits. i.e. Number of bits per sample $= v$

We know that, Number of samples per second $= f_s$

$\therefore$ Number of bits per second is given by,

$$\begin{aligned}(\text{Number of bits per second}) &= (\text{Number of bits per samples}) \\ &\quad \times (\text{Number of samples per second}) \\ &= v \text{ bits per sample} \times f_s \text{ samples per second} \qquad \dots \quad 2\end{aligned}$$

The number of bits per second is also called signaling rate of PCM and is denoted by 'r' i.e.,

$$\boxed{\textbf{Signaling rate in PCM : } r = v f_s} \qquad \dots \quad 3$$

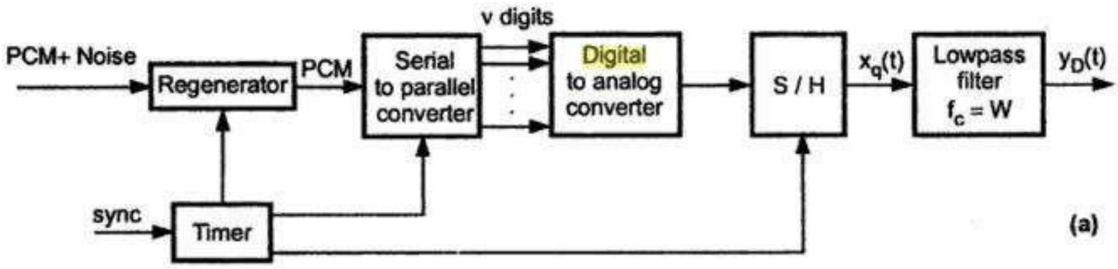
Here $f_s \geq 2W$.

Bandwidth needed for PCM transmission will be given by half of the signaling rate i.e.,

$$\text{Transmission Bandwidth of PCM : } \begin{cases} B_T \geq \frac{1}{2} r & \dots \quad 4 \\ B_T \geq \frac{1}{2} v f_s \quad \text{Since } f_s \geq 2W & \dots \quad 5 \\ B_T \geq v W & \dots \quad 6 \end{cases}$$

PCM Receiver

Fig. 9 (a) shows the block diagram of PCM receiver and Fig. 9 (b) shows the reconstructed signal. The regenerator at the start of PCM receiver reshapes the pulses and removes the noise. This signal is then converted to parallel digital words for each sample.



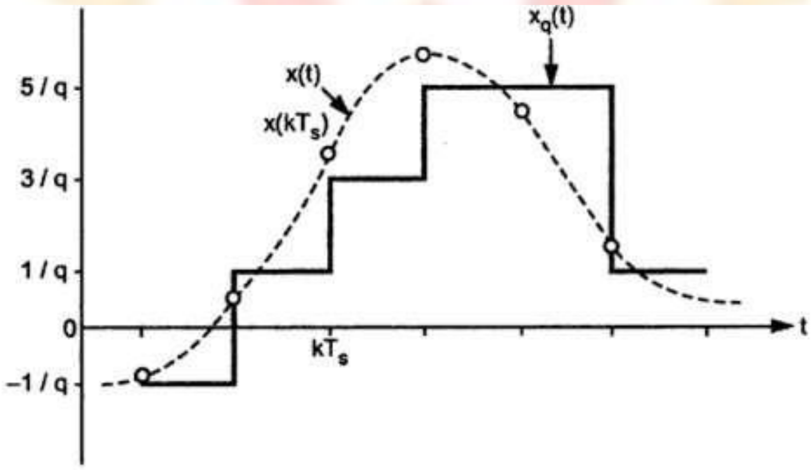


Fig. 9 (a) PCM receiver
(b) Reconstructed waveform

The digital word is converted to its analog value $x_q(t)$ along with sample and hold. This signal, at the output of S/H is passed through lowpass reconstruction filter to get $y_D(t)$. As shown in reconstructed signal of Fig. 9 (b), it is impossible to reconstruct exact original signal $x(t)$ because of permanent quantization error introduced during quantization at the transmitter. This quantization error can be reduced by increasing the binary levels. This is equivalent to increasing binary digits (bits) per sample. But increasing bits 'v' increases the signaling rate as well as transmission bandwidth as we have seen in equation 3 and equation 6. Therefore the choice of these parameters is made, such that noise due to quantization error (called as quantization noise) is in tolerable limits.

<u>Quantization</u>

- The quantizing of an analog signal is done by discretizing the signal with a number of quantization levels.

- **Quantization** is representing the sampled values of the amplitude by a finite set of levels, which means converting a continuous-amplitude sample into a discrete-time signal
- Both sampling and quantization result in the loss of information.
- The quality of a Quantizer output depends upon the number of quantization levels used.
- The discrete amplitudes of the quantized output are called as **representation levels** or **reconstruction levels**.
- The spacing between the two adjacent representation levels is called a **quantum** or **step-size**.
- There are two types of Quantization
  - Uniform Quantization
  - Non-uniform Quantization.
- The type of quantization in which the quantization levels are uniformly spaced is termed as a **Uniform Quantization**.
- The type of quantization in which the quantization levels are unequal and mostly the relation between them is logarithmic, is termed as a **Non-uniform Quantization.**

Uniform Quantization:

- There are two types of uniform quantization.
  - Mid-Rise type
  - Mid-Tread type.
- The following figures represent the two types of uniform quantization.

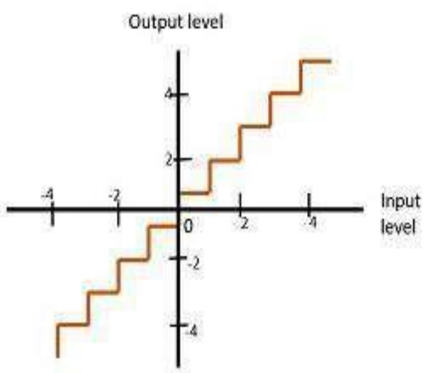


Fig 1 : Mid-Rise type Uniform Quantization

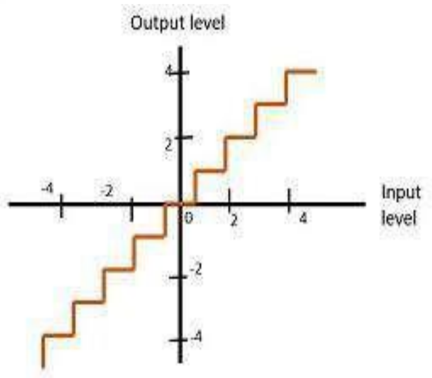


Fig 2 : Mid-Tread type Uniform Quantization

- The **Mid-Rise** type is so called because the origin lies in the middle of a raising part of the stair-case like graph. The quantization levels in this type are even in number.
- The **Mid-tread** type is so called because the origin lies in the middle of a tread of the stair-case like graph. The quantization levels in this type are odd in number.
- Both the mid-rise and mid-tread type of uniform quantizer is symmetric about the origin.

Quantization Noise and Signal to Noise ratio in PCM System

Derivation of Quantization Error/Noise or Noise Power for Uniform (Linear) Quantization

Step 1 : Quantization Error

Because of quantization, inherent errors are introduced in the signal. This error is called *quantization error*. We have defined quantization error as,

$$\varepsilon = x_q(nT_s) - x(nT_s) \qquad \text{..........(1)}$$

Step 2 : Step size

Let an input $x(nT_s)$ be of continuous amplitude in the range $-x_{max}$ to $+x_{max}$.

Therefore the total amplitude range becomes,

$$\text{Total amplitude range} = x_{max} - (-x_{max}) = 2x_{max} \qquad \text{..........(2)}$$

If this amplitude range is divided into 'q' levels of quantizer, then the step size 'δ' is given as,

$$\delta = \frac{x_{max} - (-x_{max})}{q}$$

$$= \frac{2x_{max}}{q} \qquad \text{..........(3)}$$

If signal $x(t)$ is normalized to minimum and maximum values equal to 1, then

$$x_{max} = 1, \quad -x_{max} = -1 \qquad \text{..........(4)}$$

Therefore step size will be,

$$\delta = \frac{2}{q} \quad \text{(for normalized signal)} \qquad \text{..........(5)}$$

Step 3 : Pdf of Quantization error

If step size 'δ' is sufficiently small, then it is reasonable to assume that the quantization error 'ε' will be uniformly distributed random variable. The maximum quantization error is given by

$$\varepsilon_{max} = \left|\frac{\delta}{2}\right| \qquad \text{..........(6)}$$

i.e.

$$-\frac{\delta}{2} \geq \varepsilon_{max} \geq \frac{\delta}{2} \qquad \text{..........(7)}$$

Thus over the interval $\left(-\frac{\delta}{2}, \frac{\delta}{2}\right)$ quantization error is uniformly distributed random variable.

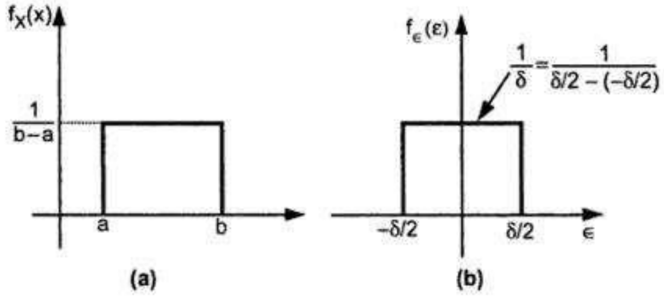


Fig. 10 (a) Uniform distribution
(b) Uniform distribution for quantization error

In above figure, a random variable is said to be uniformly distributed over an interval (a, b). Then PDF of 'X' is given by, (from equation of Uniform PDF).

$$f_X(x) = \begin{cases} 0 & \text{for} \quad x \leq a \\ \dfrac{1}{b-a} & \text{for} \quad a < x \leq b \\ 0 & \text{for} \quad x > b \end{cases} \qquad \text{.......................(8)}$$

Thus with the help of above equation we can define the probability density function for quantization error 'ε' as,

$$f_\varepsilon(\varepsilon) = \begin{cases} 0 & \text{for} \quad \varepsilon \leq \dfrac{\delta}{2} \\ \dfrac{1}{\delta} & \text{for} \quad -\dfrac{\delta}{2} < \varepsilon \leq \dfrac{\delta}{2} \\ 0 & \text{for} \quad \varepsilon > \dfrac{\delta}{2} \end{cases} \qquad \text{..................................(9)}$$

Step 4 : Noise Power

quantization error 'ε' has zero average value. That is mean 'm_ε' of the quantization error is zero.

The signal to quantization noise ratio of the quantizer is defined as,

$$\frac{S}{N} = \frac{\text{Signal power (normalized)}}{\text{Noise power (normalized)}} \quad \cdots 10$$

If type of signal at input i.e., $x(t)$ is known, then it is possible to calculate signal power.

The noise power is given as,

$$\text{Noise power} = \frac{V_{noise}^2}{R} \quad \cdots (11)$$

Here V_{noise}^2 is the mean square value of noise voltage. Since noise is defined by random variable 'ε' and PDF $f_\varepsilon(\varepsilon)$, its mean square value is given as,

$$\text{mean square value} = E[\varepsilon^2] = \bar{\varepsilon}^2 \quad \cdots (12)$$

The mean square value of a random variable 'X' is given as,

$$\bar{X}^2 = E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx \quad \text{By definition} \quad \cdots (13)$$

Here $$E[\varepsilon^2] = \int_{-\infty}^{\infty} \varepsilon^2 f_\varepsilon(\varepsilon)\, d\varepsilon \quad \cdots (14)$$

From equation 9 we can write above equation as,

$$E[\varepsilon^2] = \int_{-\delta/2}^{\delta/2} \varepsilon^2 \times \frac{1}{\delta}\, d\varepsilon$$

$$= \frac{1}{\delta}\left[\frac{\varepsilon^3}{3}\right]_{-\delta/2}^{\delta/2} = \frac{1}{\delta}\left[\frac{(\delta/2)^3}{3} + \frac{(\delta/2)^3}{3}\right]$$

$$= \frac{1}{3\delta}\left[\frac{\delta^3}{8} + \frac{\delta^3}{8}\right] = \frac{\delta^2}{12} \quad \cdots (15)$$

∴ From equation 1.8.25, the mean square value of noise voltage is,

$$V_{noise}^2 = \text{mean square value} = \frac{\delta^2}{12}$$

When load resistance, $R=1$ ohm, then the noise power is normalized i.e.,

$$\text{Noise power (normalized)} = \frac{V_{noise}^2}{1} \qquad \text{[with } R=1 \text{ in equation 11]}$$

$$= \frac{\delta^2/12}{1} = \frac{\delta^2}{12}$$

Thus we have,

Normalized noise power

or Quantization noise power $= \dfrac{\delta^2}{12}$ **; For linear quantization.**

or Quantization error (in terms of power) ... (16)

Derivation of Maximum Signal to Quantization Noise Ratio for Linear Quantization:

signal to quantization noise ratio is given as,

$$\frac{S}{N} = \frac{\text{Normalized signal power}}{\text{Normalized noise power}}$$

$$= \frac{\text{Normalized signal power}}{(\delta^2/12)} \qquad \text{... (17)}$$

The number of bits 'v' and quantization levels 'q' are related as,

$$q = 2^v \qquad \text{... (18)}$$

Putting this value in equation (3) we have,

$$\delta = \frac{2x_{\max}}{2^v} \qquad \text{... (19)}$$

Putting this value in equation 1.8.30 we get,

$$\frac{S}{N} = \frac{\text{Normalized signal power}}{\left(\frac{2x_{\max}}{2^v}\right)^2 + 12}$$

Let normalized signal power be denoted as 'P'.

$$\frac{S}{N} = \frac{P}{\frac{4x_{\max}^2}{2^{2v}} \times \frac{1}{12}} = \frac{3P}{x_{\max}^2} \cdot 2^{2v}$$

This is the required relation for maximum signal to quantization noise ratio. Thus,

$$\boxed{\textbf{Maximum signal to quantization noise ratio : } \frac{S}{N} = \frac{3P}{x_{max}^2} \cdot 2^{2v}} \quad \cdots (20)$$

This equation shows that signal to noise power ratio of quantizer increases exponentially with increasing bits per sample.

If we assume that input $x(t)$ is normalized, i.e.,

$$x_{max} = 1 \quad \cdots (21)$$

Then signal to quantization noise ratio will be,

$$\frac{S}{N} = 3 \times 2^{2v} \times P \quad \cdots (22)$$

If the destination signal power 'P' is normalized, i.e.,

$$P \leq 1 \quad \cdots (23)$$

Then the signal to noise ratio is given as,

$$\frac{S}{N} \leq 3 \times 2^{2v} \quad \cdots (24)$$

Since $x_{max} = 1$ and $P \leq 1$, the signal to noise ratio given by above equation is normalized.

Expressing the signal to noise ratio in decibels,

$$\left(\frac{S}{N}\right) dB = 10 \log_{10} \left(\frac{S}{N}\right) dB \quad \text{since power ratio.}$$

$$\leq 10 \log_{10} [3 \times 2^{2v}]$$

$$\leq (4.8 + 6v) \, dB$$

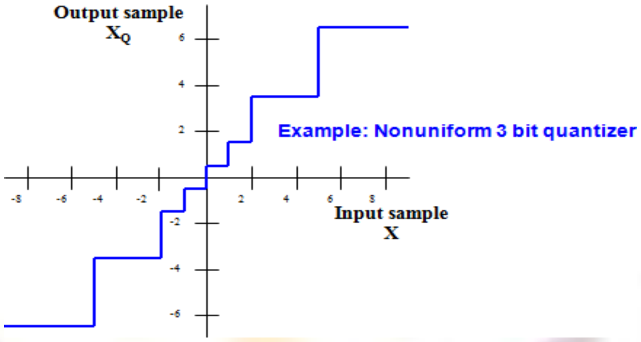
Thus,

Signal to Quantization noise ratio
for normalized values of power : $\left(\frac{S}{N}\right) dB \leq (4.8 + 6v) \, dB$
'P' and amplitude of input $x(t)$

$\cdots (25)$

Non-Uniform Quantization:

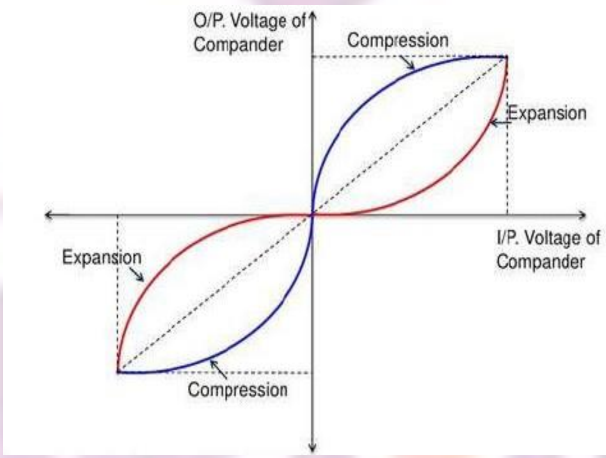
In non-uniform quantization, the step size is not fixed. It varies according to certain law or as per input signal amplitude. The following fig shows the characteristics of Non uniform quantizer.

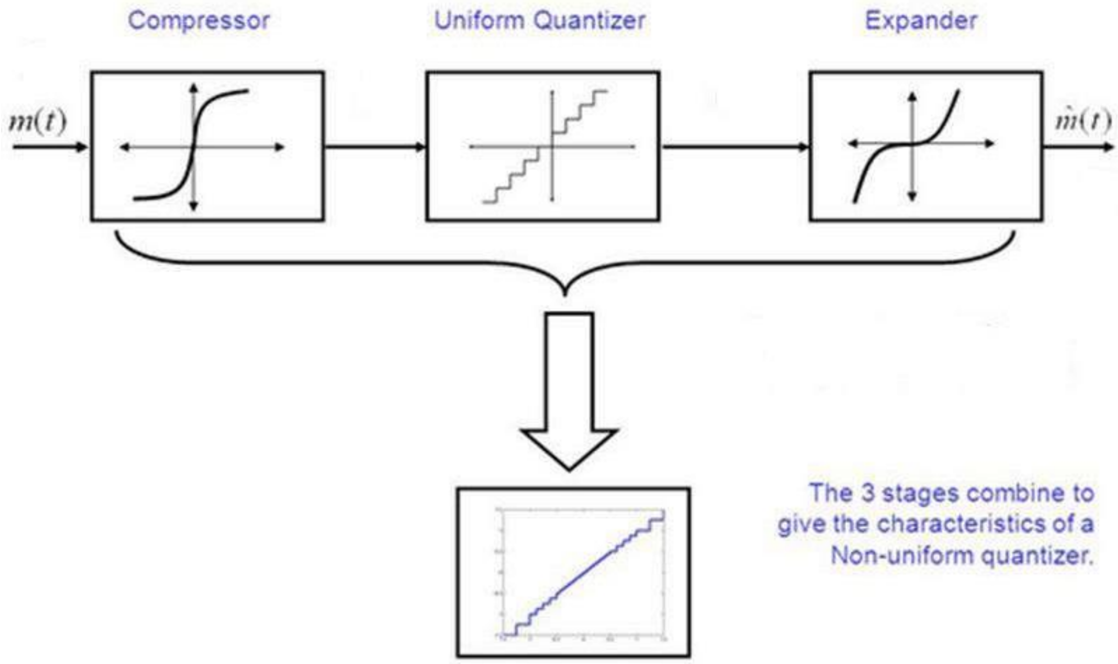


In this figure observe that step size is small at low input signal levels. Hence quantization error is also small at these inputs. Therefore signal to quantization noise power ratio is improved at low signal levels. Stepsize is higher at high input levels. Hence signal to noise power ratio remains almost same throughout the dynamic range of quantizer.

Companding PCM System

- Non-uniform quantizers are difficult to make and expensive.
- An alternative is to first pass the speech signal through nonlinearity before quantizing with a uniform quantizer.
- The nonlinearity causes the signal amplitude to be ***compressed***.
  - The input to the quantizer will have a more uniform distribution.
- At the receiver, the signal is ***expanded*** by an inverse to the nonlinearity.
- The process of compressing and expanding is called ***Companding***.





μ - Law Companding for Speech Signals

Normally for speech and music signals a μ - law compression is used. This compression is defined by the following equation,

$$Z(x) = (\text{Sgn } x) \frac{\ln(1+\mu|x|)}{\ln(1+\mu)} \quad |x| \leq 1 \qquad \cdots (1)$$

Below Fig shows the variation of signal to noise ratio with respect to signal level without companding and with companding.

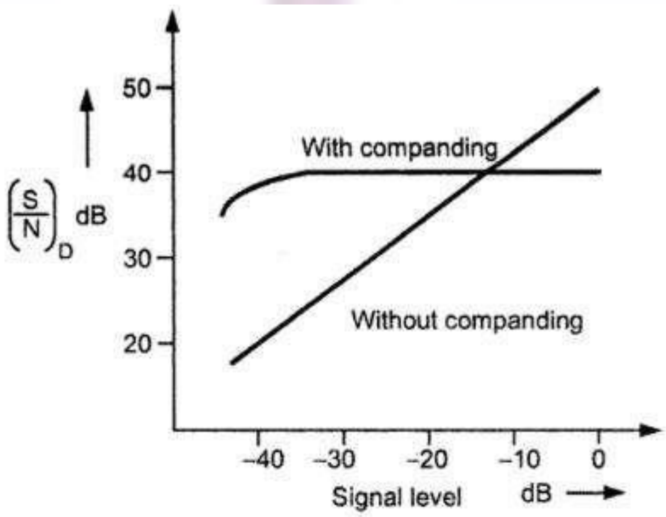


Fig. 11 PCM performance with μ - law companding

It can be observed from above figure that signal to noise ratio of PCM remains almost constant with companding.

A-Law for Companding

The A law provides piecewise compressor characteristic. It has linear segment for low level inputs and logarithmic segment for high level inputs. It is defined as,

$$Z(x) = \begin{cases} \dfrac{A|x|}{1+\ln A} & \text{for } 0 \le |x| \le \dfrac{1}{A} \\ \dfrac{1+\ln(A|x|)}{1+\ln A} & \text{for } \dfrac{1}{A} \le |x| \le 1 \end{cases} \quad \cdots (2)$$

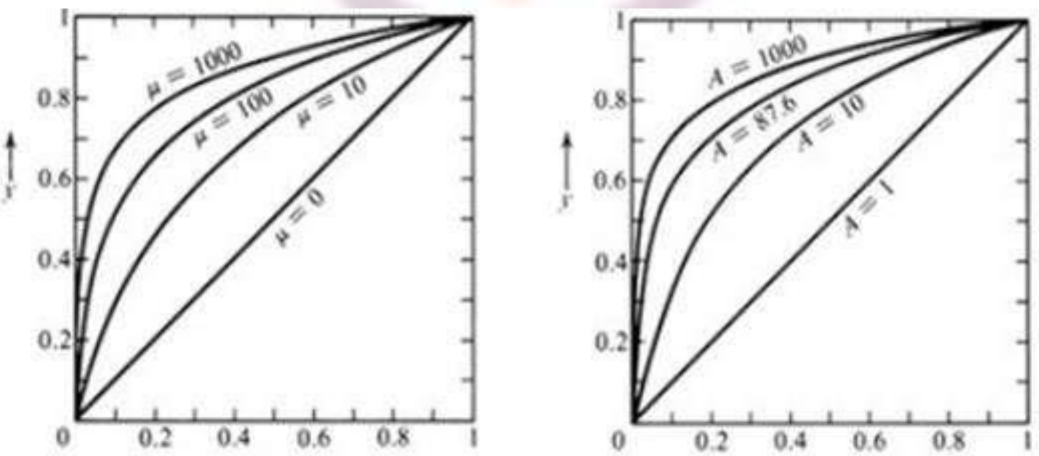
When A = 1, we get uniform quantization. The practical value for A is 87.56. Both A-law and μ-law companding is used for PCM telephone systems.

Signal to Noise Ratio of Companded PCM

The signal to noise ratio of companded PCM is given as,

$$\frac{S}{N} = \frac{3q^2}{[\ln(1+\mu)]^2} \quad \cdots (3)$$

Here $q = 2^v$ is number of quantization levels.



Differential Pulse Code Modulation (DPCM)

Redundant Information in PCM

The samples of a signal are highly corrected with each other. This is because any signal does not change fast. That is its value from present sample to next sample does not differ by large amount. The adjacent samples of the signal carry the same information with little difference. When these samples are encoded by standard PCM system, the resulting encoded signal contains redundant information.

Fig. shows a continuous time signal $x(t)$ by dotted line. This signal is sampled by flat top sampling at intervals $T_s, 2T_s, 3T_s \ldots nT_s$. The sampling frequency is selected to be higher than nyquist rate. The samples are encoded by using 3 bit (7 levels) PCM. The sample is quantized to the nearest digital level as shown by small

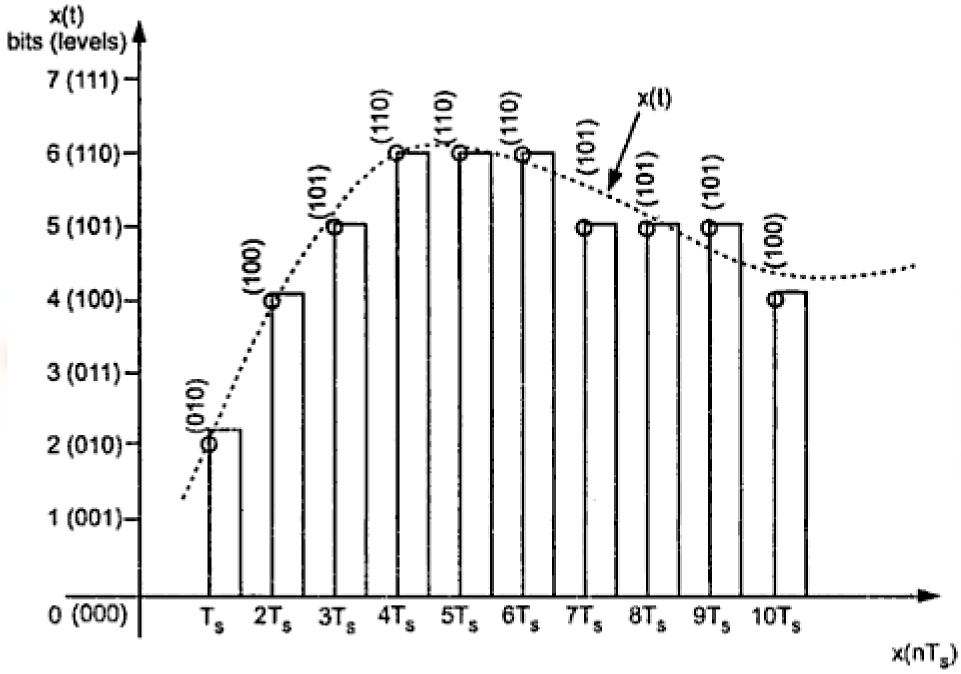


Fig. Redundant information in PCM

circles in the diagram. The encoded binary value of each sample is written on the top of the samples. We can see from Fig. that the samples taken at $4T_s, 5T_s$ and $6T_s$ are encoded to same value of (110). This information can be carried only by one sample. But three samples are carrying the same information means it is redundant. Consider another example of samples taken at $9T_s$ and $10T_s$. The difference between these samples is only due to last bit and first two bits are redundant, since they do not change.

Principle of DPCM

If this redundancy is reduced, then overall bit rate will decrease and number of bits required to transmit one sample will also be reduced. This type of digital pulse modulation scheme is called Differential Pulse Code Modulation.

DPCM Transmitter

The differential pulse code modulation works on the principle of prediction. The value of the present sample is predicted from the past samples. The prediction may not be exact but it is very close to the actual sample value. Fig. shows the transmitter of Differential Pulse Code Modulation (DPCM) system. The sampled signal is denoted by $x(nT_s)$ and the predicted signal is denoted by $\hat{x}(nT_s)$. The comparator finds out the difference between the actual sample value $x(nT_s)$ and predicted sample value $\hat{x}(nT_s)$. This is called error and it is denoted by $e(nT_s)$. It can be defined as,

$$e(nT_s) = x(nT_s) - \hat{x}(nT_s) \qquad \text{..................(1)}$$

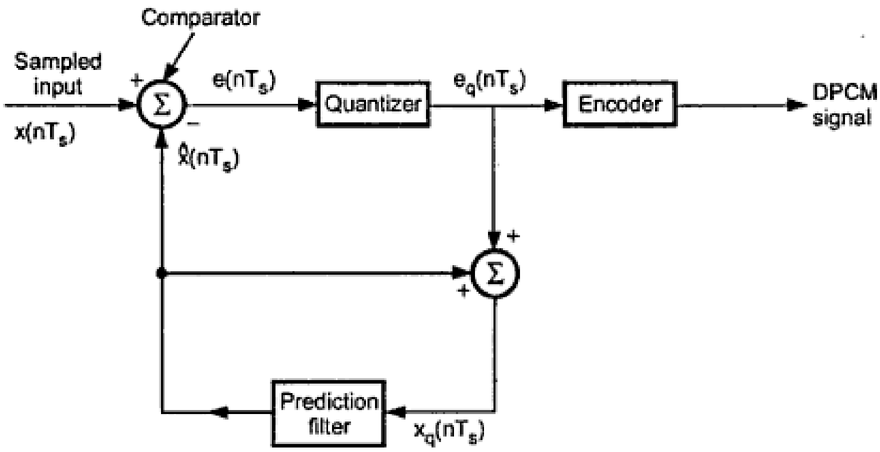


Fig. Differential pulse code modulation transmitter

Thus error is the difference between unquantized input sample $x(nT_s)$ and prediction of it $\hat{x}(nT_s)$. The predicted value is produced by using a prediction filter. The quantizer output signal $e_q(nT_s)$ and previous prediction is added and given as input to the prediction filter. This signal is called $x_q(nT_s)$. This makes the prediction more and more close to the actual sampled signal. We can see that the quantized error signal $e_q(nT_s)$ is very small and can be encoded by using small number of bits. Thus number of bits per sample are reduced in DPCM.

The quantizer output can be written as,

$$e_q(nT_s) = e(nT_s) + q(nT_s) \qquad \text{.................(2)}$$

Here $q(nT_s)$ is the quantization error. As shown in Fig. the prediction filter input $x_q(nT_s)$ is obtained by sum $\hat{x}(nT_s)$ and quantizer output i.e.,

$$x_q(nT_s) = \hat{x}(nT_s) + e_q(nT_s) \qquad \text{..................(3)}$$

Putting the value of $e_q(nT_s)$ from equation 2 in the above equation we get,

$$x_q(nT_s) = \hat{x}(nT_s) + e(nT_s) + q(nT_s) \qquad \text{................(4)}$$

Equation 1 is written as,

$$e(nT_s) = x(nT_s) - \hat{x}(nT_s)$$

$$\therefore \quad e(nT_s) + \hat{x}(nT_s) = x(nT_s) \qquad \text{...............(5)}$$

$\therefore$ Putting the value of $e(nT_s) + \hat{x}(nT_s)$ from above equation into equation 4 we get,

$$x_q(nT_s) = x(nT_s) + q(nT_s) \qquad \text{..................(6)}$$

Thus the quantized version of the signal $x_q(nT_s)$ is the sum of original sample value and quantization error $q(nT_s)$. The quantization error can be positive or negative. Thus equation 6 does not depend on the prediction filter characteristics.

Reconstruction of DPCM Signal

Fig. shows the block diagram of DPCM receiver.

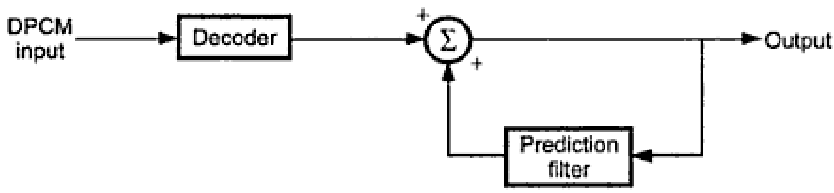


Fig. DPCM receiver

The decoder first reconstructs the quantized error signal from incoming binary signal. The prediction filter output and quantized error signals are summed up to give the quantized version of the original signal. Thus the signal at the receiver differs from actual signal by quantization error $q(nT_s)$, which is introduced permanently in the reconstructed signal.

Introduction to Delta Modulation

PCM transmits all the bits which are used to code the sample. Hence signaling rate and transmission channel bandwidth are large in PCM. To overcome this problem Delta Modulation is used.

Delta Modulation

Operating Principle of DM

Delta modulation transmits only one bit per sample. That is the present sample value is compared with the previous sample value and the indication,whether the amplitude is increased or decreased is sent. Input signal $x(t)$ is approximated to step signal by the delta modulator. This step size is fixed. The difference between the input signal $x(t)$ and staircase approximated signal confined to two levels, i.e. $+\delta$ and $-\delta$. If the difference is positive, then approximated signal is increased by one step i.e. 'δ'. If the difference is negative, then approximated signal is reduced by 'δ'. When the step is reduced, '0' is transmitted and if the step is increased, '1' is transmitted. Thus for each sample, only one binary bit is transmitted. Fig. shows the analog signal $x(t)$ and its staircase approximated signal by the delta modulator.

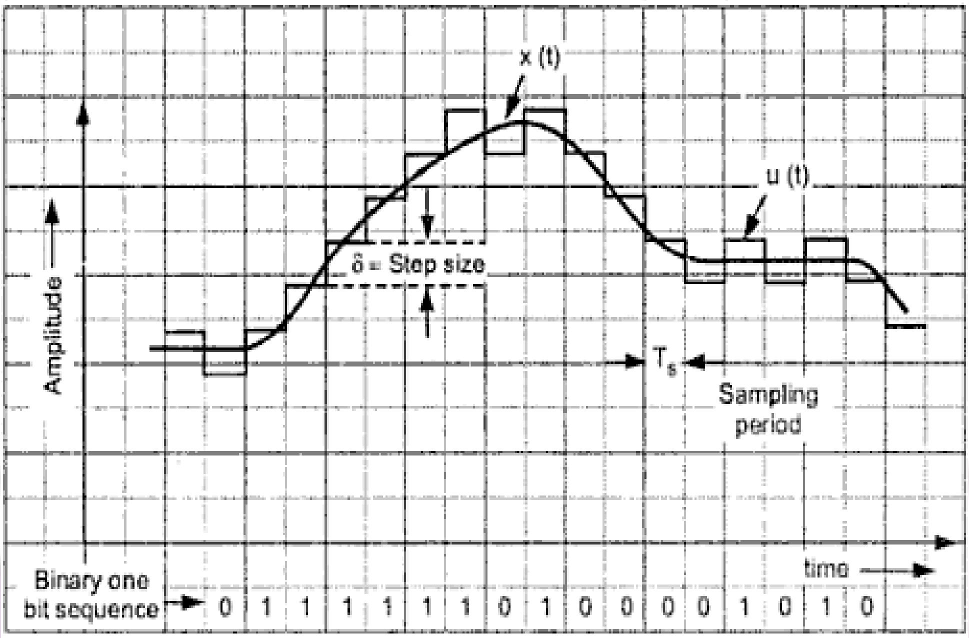


Fig. Delta modulation waveform

The principle of delta modulation can be explained by the following set of equations. The error between the sampled value of $x(t)$ and last approximated sample is given as,

$$e(nT_s) = x(nT_s) - \hat{x}(nT_s) \qquad \dots (\quad 1)$$

Here, $e(nT_s)$ = Error at present sample

$x(nT_s)$ = Sampled signal of $x(t)$

$\hat{x}(nT_s)$ = Last sample approximation of the staircase waveform.

We can call $u(nT_s)$ as the present sample approximation of staircase output.

Then, $u[(n-1)T_s] = \hat{x}(nT_s) \qquad \dots (\quad 2)$

= Last sample approximation of staircase waveform.

Let the quantity $b(nT_s)$ be defined as,

$$b(nT_s) = \delta \, sgn[e(nT_s)] \qquad \dots (\quad 3)$$

That is depending on the sign of error $e(nT_s)$ the sign of step size δ will be decided. In other words,

$$\begin{aligned} b(nT_s) &= +\delta \quad \text{if} \quad x(nT_s) \geq \hat{x}(nT_s) \\ &= -\delta \quad \text{if} \quad x(nT_s) < \hat{x}(nT_s) \qquad \dots (\quad 4) \end{aligned}$$

If $b(nT_s) = +\delta$; binary '1' is transmitted

and if $b(nT_s) = -\delta$; binary '0' is transmitted.

T_s = Sampling interval.

DM Transmitter

Fig. (a) shows the transmitter based on equations 3 to 5.

The summer in the accumulator adds quantizer output ($\pm\delta$) with the previous sample approximation. This gives present sample approximation. i.e.,

$$\begin{aligned} u(nT_s) &= u(nT_s - T_s) + [\pm\delta] \quad \text{or} \\ &= u[(n-1)T_s] + b(nT_s) \qquad \dots (\quad 5) \end{aligned}$$

The previous sample approximation $u[(n-1)T_s]$ is restored by delaying one sample period T_s. The sampled input signal $x(nT_s)$ and staircase approximated signal $\hat{x}(nT_s)$ are subtracted to get error signal $e(nT_s)$.

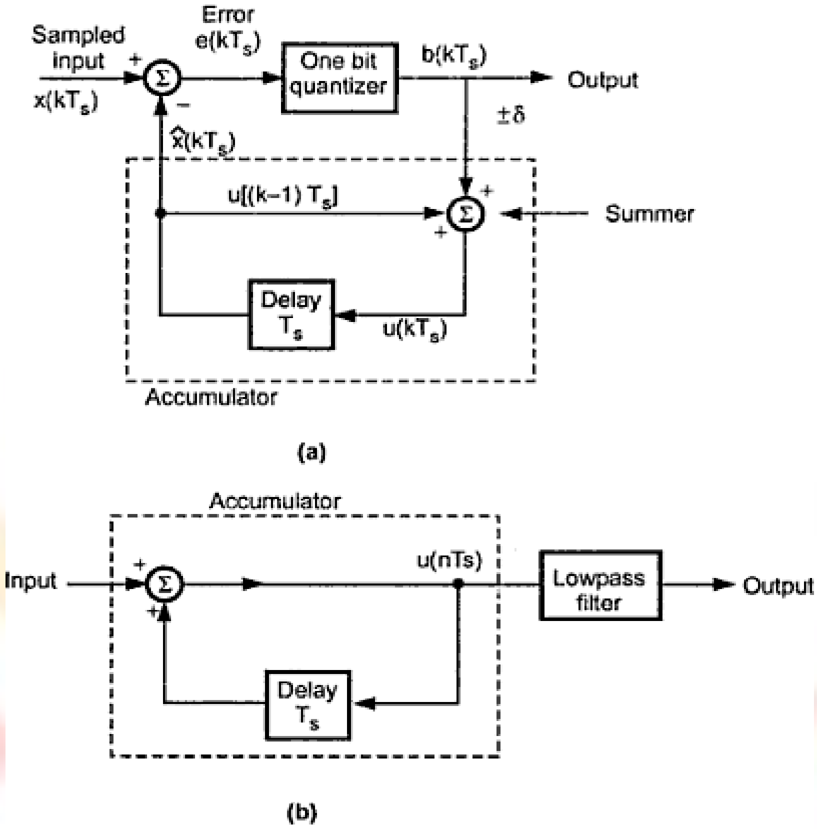


Fig. (a) Delta modulation transmitter and (b) Delta modulation receiver

Depending on the sign of $e(nT_s)$ one bit quantizer produces an output step of $+\delta$ or $-\delta$. If the step size is $+\delta$, then binary '1' is transmitted and if it is $-\delta$, then binary '0' is transmitted.

DM Receiver

At the receiver shown in Fig. (b), the accumulator and low-pass filter are used. The accumulator generates the staircase approximated signal output and is delayed by one sampling period T_s. It is then added to the input signal. If input is binary '1' then it adds $+\delta$ step to the previous output (which is delayed). If input is binary '0' then one step 'δ' is subtracted from the delayed signal. The low-pass filter has the cutoff frequency equal to highest frequency in $x(t)$. This filter smoothen the staircase signal to reconstruct $x(t)$.

Advantages and Disadvantages of Delta Modulation

Advantages of Delta Modulation

The delta modulation has following advantages over PCM,

1. Delta modulation transmits only one bit for one sample. Thus the signaling rate and transmission channel bandwidth is quite small for delta modulation.
2. The transmitter and receiver implementation is very much simple for delta modulation. There is no analog to digital converter involved in delta modulation.

Disadvantages of Delta Modulation

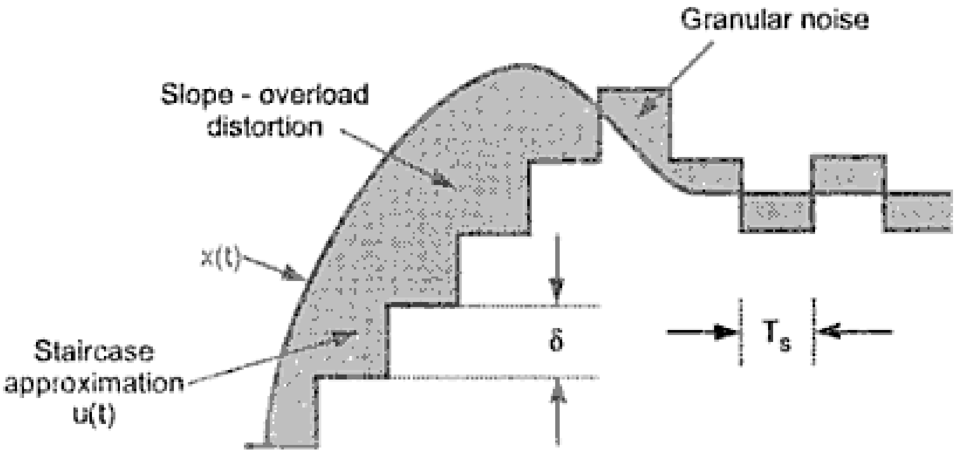


Fig. Quantization errors in delta modulation

The delta modulation has two drawbacks -

Slope Overload Distortion (Startup Error)

This distortion arises because of the large dynamic range of the input signal.

As can be seen from Fig. the rate of rise of input signal $x(t)$ is so high that the staircase signal cannot approximate it, the step size 'δ' becomes too small for staircase signal $u(t)$ to follow the steep segment of $x(t)$. Thus there is a large error between the staircase approximated signal and the original input signal $x(t)$. This error is called *slope overload distortion*. To reduce this error, the step size should be increased when slope of signal of $x(t)$ is high.

Since the step size of delta modulator remains fixed, its maximum or minimum slopes occur along straight lines. Therefore this modulator is also called Linear Delta Modulator (LDM).

Granular Noise (Hunting)

Granular noise occurs when the step size is too large compared to small variations in the input signal. That is for very small variations in the input signal, the staircase signal is changed by large amount (δ) because of large step size. Fig shows that when the input signal is almost flat, the staircase signal $u(t)$ keeps on oscillating by $\pm\delta$ around the signal. The error between the input and approximated signal is called *granular noise.* The solution to this problem is to make step size small.

Thus large step size is required to accommodate wide dynamic range of the input signal (to reduce slope overload distortion) and small steps are required to reduce granular noise. Adaptive delta modulation is the modification to overcome these errors.

Adaptive Delta Modulation

Operating Principle

To overcome the quantization errors due to slope overload and granular noise, the step size (δ) is made adaptive to variations in the input signal $x(t)$. Particularly in the steep segment of the signal $x(t)$, the step size is increased. When the input is varying slowly, the step size is reduced. Then the method is called *Adaptive Delta Modulation (ADM).*

The adaptive delta modulators can take continuous changes in step size or discrete changes in step size.

Transmitter and Receiver

Fig. (a) shows the transmitter and (b) shows receiver of adaptive delta modulator. The logic for step size control is added in the diagram. The step size increases or decreases according to certain rule depending on one bit quantizer output.

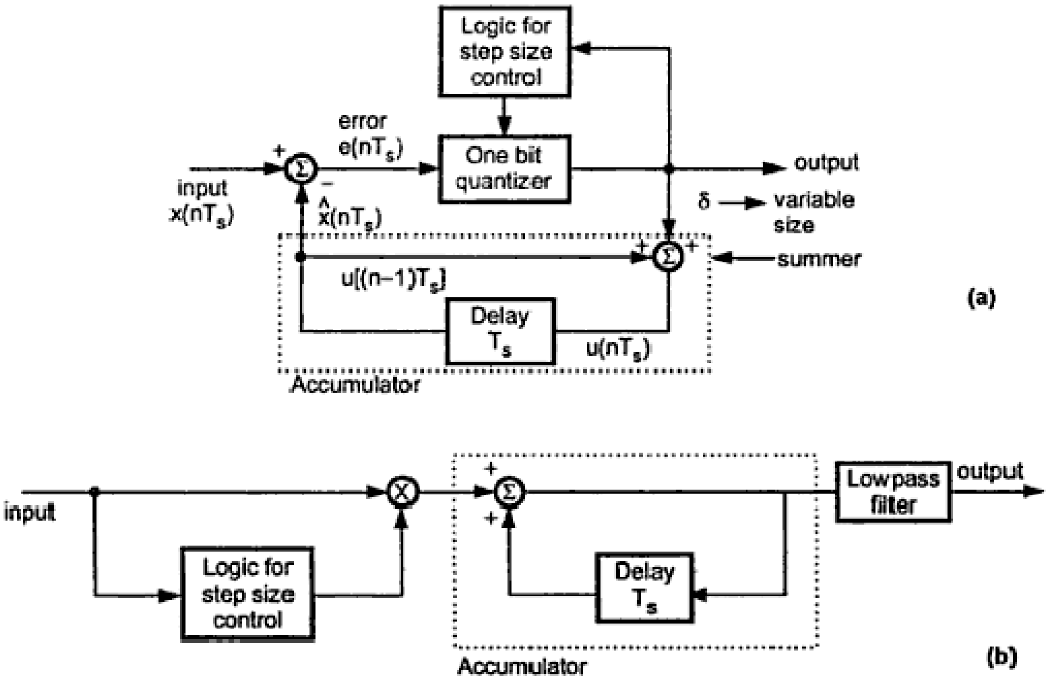


Fig. Adaptive delta modulator (a) Transmitter (b) Receiver

For example if one bit quantizer output is high (1), then step size may be doubled for next sample. If one bit quantizer output is low, then step size may be reduced by one step. Fig. shows the waveforms of adaptive delta modulator and sequence of bits transmitted.

In the receiver of adaptive delta modulator shown in Fig. (b) the first part generates the step size from each incoming bit. Exactly the same process is followed as that in transmitter. The previous input and present input decides the step size. It is then given to an accumulator which builds up staircase waveform. The low-pass filter then smoothens out the staircase waveform to reconstruct the smooth signal.

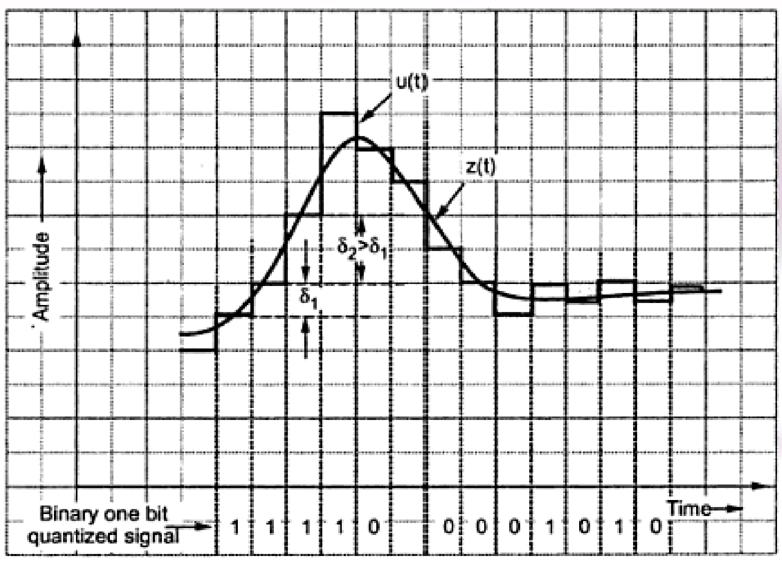


Fig. Waveforms of adaptive delta modulation

Advantages of Adaptive Delta Modulation

Adaptive delta modulation has certain advantages over delta modulation. i.e.,

1. The signal to noise ratio is better than ordinary delta modulation because of the reduction in slope overload distortion and granular noise.
2. Because of the variable step size, the dynamic range of ADM is wide.
3. Utilization of bandwidth is better than delta modulation.

Plus other advantages of delta modulation are, only one bit per sample is required and simplicity of implementation of transmitter and receiver.

Condition for Slope overload distortion occurrence

Slope overload distortion will occur if

$$A_m > \frac{\delta}{2\pi f_m T_s}$$

where T_s is the sampling period.

Let the sine wave be represented as,

$$x(t) = A_m \sin(2\pi f_m t)$$

Slope of $x(t)$ will be maximum when derivative of $x(t)$ with respect to 't' will be maximum. The maximum slope of delta modulator is given

$$\text{Max. slope} = \frac{\text{Step size}}{\text{Sampling period}}$$

$$= \frac{\delta}{T_s} \qquad \text{.............(1)}$$

Slope overload distortion will take place if slope of sine wave is greater than slope of delta modulator i.e.

$$\max\left|\frac{d}{dt} x(t)\right| > \frac{\delta}{T_s}$$

$$\max\left|\frac{d}{dt} A_m \sin(2\pi f_m t)\right| > \frac{\delta}{T_s}$$

$$\max\left|A_m 2\pi f_m \cos(2\pi f_m t)\right| > \frac{\delta}{T_s}$$

$$A_m 2\pi f_m > \frac{\delta}{T_s}$$

or $$\boxed{A_m > \frac{\delta}{2\pi f_m T_s}} \qquad \text{...............(2)}$$

Expression for Signal to Quantization Noise power ratio for Delta Modulation

To obtain signal power :

slope overload distortion will not occur if

$$A_m \le \frac{\delta}{2\pi f_m T_s}$$

Here A_m is peak amplitude of sinusoided signal

δ is the step size

f_m is the signal frequency and

T_s is the sampling period.

From above equation, the maximum signal amplitude will be,

$$A_m = \frac{\delta}{2\pi f_m T_s} \qquad \text{..................(1)}$$

Signal power is given as,

$$P = \frac{V^2}{R}$$

Here V is the rms value of the signal. Here $V = \frac{A_m}{\sqrt{2}}$. Hence above equation becomes,

$$P = \left(\frac{A_m}{\sqrt{2}}\right)^2 / R$$

Normalized signal power is obtained by taking R = 1. Hence,

$$P = \frac{A_m^2}{2}$$

Putting for A_m from equation 1

$$P = \frac{\delta^2}{8\pi^2 f_m^2 T_s^2} \qquad \text{.................(2)}$$

This is an expression for signal power in delta modulation.

(ii) To obtain noise power

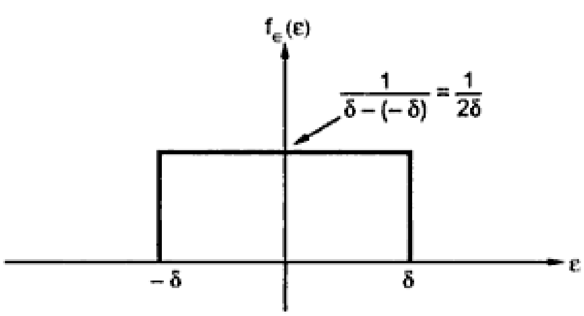


Fig. Uniform distribution of quantization error

We know that the maximum quantization error in delta modulation is equal to step size 'δ'. Let the quantization error be uniformly distributed over an interval $[-\delta, \delta]$. This is shown in Fig. From this figure the PDF of quantization error can be expressed as,

$$f_{\epsilon}(\varepsilon) = \begin{cases} 0 & for \quad \varepsilon < \delta \\ \frac{1}{2\delta} & for \quad -\delta < \varepsilon < \delta \\ 0 & for \quad \varepsilon > \delta \end{cases} \qquad ....(3)$$

The noise power is given as,

$$\text{Noise power} = \frac{V_{noise}^2}{R}$$

Here V_{noise}^2 is the mean square value of noise voltage. Since noise is defined by random variable 'ε' and PDF $f_{\epsilon}(\varepsilon)$, its mean square value is given as,

$$\text{mean square value} = E[\varepsilon^2] = \overline{\varepsilon^2}$$

mean square value is given as,

$$E[\varepsilon^2] = \int_{-\infty}^{\infty} \varepsilon^2 f_{\epsilon}(\varepsilon)\, d\varepsilon$$

From equation 3

$$E[\varepsilon^2] = \int_{-\delta}^{\delta} \varepsilon^2 \cdot \frac{1}{2\delta}\, d\varepsilon$$

$$= \frac{1}{2\delta}\left[\frac{\varepsilon^3}{3}\right]_{-\delta}^{\delta}$$

$$= \frac{1}{2\delta}\left[\frac{\delta^3}{3} + \frac{\delta^3}{3}\right] = \frac{\delta^2}{3} \qquad ....(4)$$

Hence noise power will be,

$$\text{noise power} = \left(\frac{\delta^2}{3}\right) / R$$

Normalized noise power can be obtained with R = 1. Hence,

$$\text{noise power} = \frac{\delta^2}{3} \qquad ....(5)$$

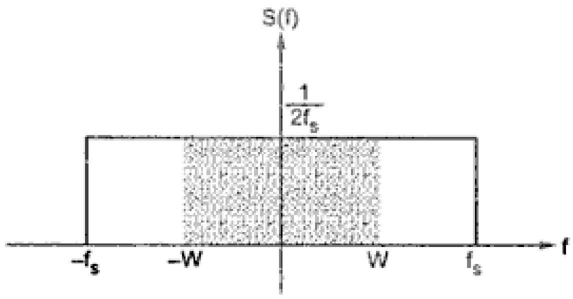


Fig. PSD of noise

This noise power is uniformly distributed over $-f_s$ to f_s range. This is illustrated in Fig. At the output of delta modulator receiver there is lowpass reconstruction filter whose cutoff frequency is 'W'. This cutoff frequency is equal to highest signal frequency. The reconstruction filter passes part of the noise power at the output as Fig. From the geometry of Fig. , output noise power will be,

$$\text{Output noise power} = \frac{W}{f_s} \times \text{noise power} = \frac{W}{f_s} \times \frac{\delta^2}{3}$$

We know that $f_s = \frac{1}{T_s}$, hence above equation becomes,

$$\text{Output noise power} = \frac{WT_s\delta^2}{3} \quad \text{.................(6)}$$

(iii) To obtain signal to noise power ratio

Signal to noise power ratio at the output of delta modulation receiver is given as,

$$\frac{S}{N} = \frac{\textit{Normalized signal power}}{\textit{Normalized noise power}}$$

From equation 2. and equation 6

$$\frac{S}{N} = \frac{\frac{\delta^2}{8\pi^2 f_m^2 T_s^2}}{\frac{WT_s\delta^2}{3}}$$

$$\boxed{\frac{S}{N} = \frac{3}{8\pi^2 W f_m^2 T_s^3}} \quad \text{..............(7)}$$

This is an expression for signal to noise power ratio in delta modulation.

UNIT-V

DIGITAL MODULATION TECHNIQUES

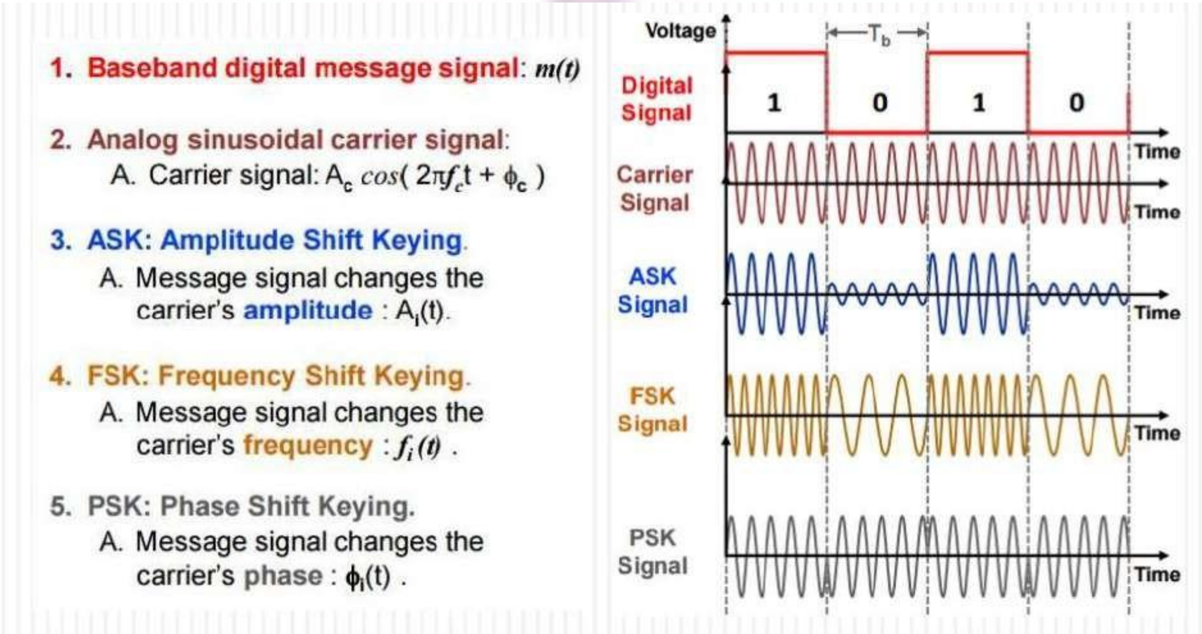
Introduction

- There are basically two types of transmission of Digital Signals
- **Baseband data transmission**
  - The digital data is transmitted over the channel directly. There is no carrier or any modulation. Suitable for transmission over short distances.
- **Pass band data transmission**
  - The digital data modulates high frequency sinusoidal carrier. Suitable for transmission over longer distances.

Types of Pass band Modulation

- The digital data can modulate phase, frequency or amplitude of carrier. This gives rise to three basic techniques:
- **Phase Shift Keying (PSK):** The digital data modulates the phase of the carrier.
- **Frequency Shift Keying(FSK):** The digital data modulates the frequency of the carrier.
- **Amplitude Shift Keying (ASK):** The digital modulates the amplitude of the carrier.

Digital Modulation Techniques



Types of Reception for Pass band Transmission

- Two Types of methods for detection of pass band signals
- **Coherent (Synchronous) Detection:** The local carrier generated at the receiver is phase locked with the carrier at the transmitter. Hence called Synchronous Detection.
- **Non Coherent (Envelope) Detection:** The receiver carrier need not be phase locked with the transmitter carrier. It is called Envelope detection. It is simple but it has higher probability of error.

Requirements of Pass band Transmission Scheme

- Maximum Data transmission rate
- Minimum Probability of symbol error
- Minimum Transmitted power
- Minimum Channel Bandwidth
- Maximum resistance to interfering signals
- Minimum circuit complexity

Advantages of Pass band Transmission over Baseband transmission

- Long Distance Transmission
- Analog Channels, can be used for Transmission
- Multiplexing techniques can be used for BW conservation.
- Problems such as ISI and crosstalk are absent
- Pass band transmission can take place over wireless channels also.

Introduction

- In digital modulation, an analog carrier signal is modulated by a discrete signal.
- Digital modulation can be considered as digital-to-analog and the corresponding demodulation is considered as analog-to-digital conversion.
- In Digital communications, the modulating wave consists of binary data and the carrier is sinusoidal wave.

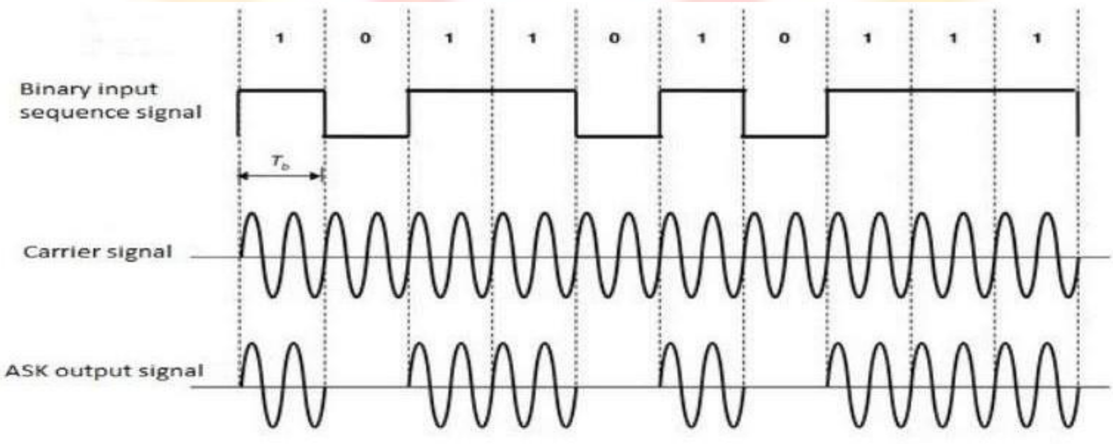
Amplitude Shift Keying (On-Off Keying)

- In this there is only one unit energy carrier and it is switched on or off depending upon the Binary sequence.

ASK waveform may be represented as

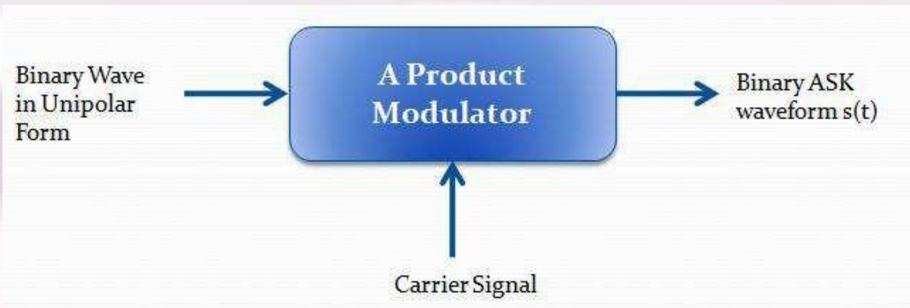
$$s(t) = \begin{cases} A\cos(2\pi f_c t)\ ; When\ bit\ transmitted\ is\ '1' \\ 0 \qquad\qquad\quad ; When\ bit\ transmitted\ is\ '0' \end{cases}$$

- Signal s(t) contains some complete cycles of carrier frequency (fc).
- Hence the ASK waveform looks like an On-Off of the signal. Therefore it is also known as the On-Off Keying(OOK)



Generation of ASK Signal

- ASK signal may be generated by simply applying the incoming binary data and the sinusoidal carrier to the 2 inputs of a product modulator.



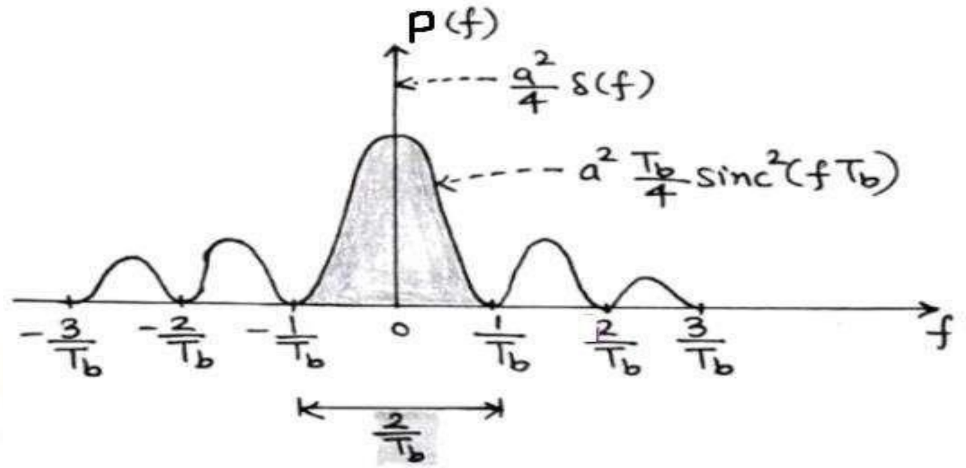
- The resulting output will be the ASK waveform.
- Modulation causes the shift of the baseband signal spectrum.

Power Spectral Density (PSD) of Unipolar NRZ:

- The PSD of Unipolar NRZ is given by equ

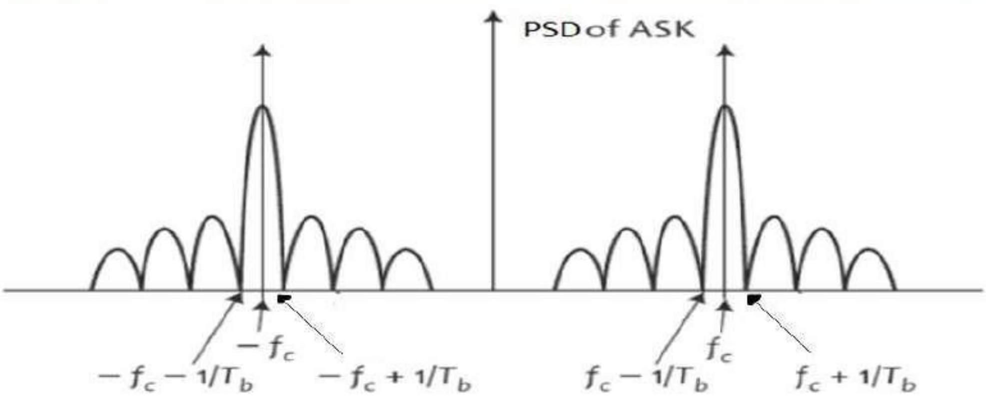
$$P(f) = \frac{a^2 T_b}{4} sinc^2(fT_b) + \frac{a^2}{4}\delta(f)$$

- PSD of Unipolar NRZ is as shown below



Power Spectral Density (PSD) of ASK

- The PSD of ASK signal is same as that of a baseband on-off signal but shifted in the frequency domain by $\pm f_c$
- It may be noted that 2 impulses occur at $\pm f_c$
- The spectrum of ASK shows that it has infinite bandwidth.
- Bandwidth is defined as the BW of an ideal band pass filter centred at fc whose output contains about 95% of the total average power content of the ASK signal.



Power Spectral Density (PSD) Waveform for an ASK Signal

- According to this criterion the Bandwidth of ASK signal is approximately $3/T_b$.

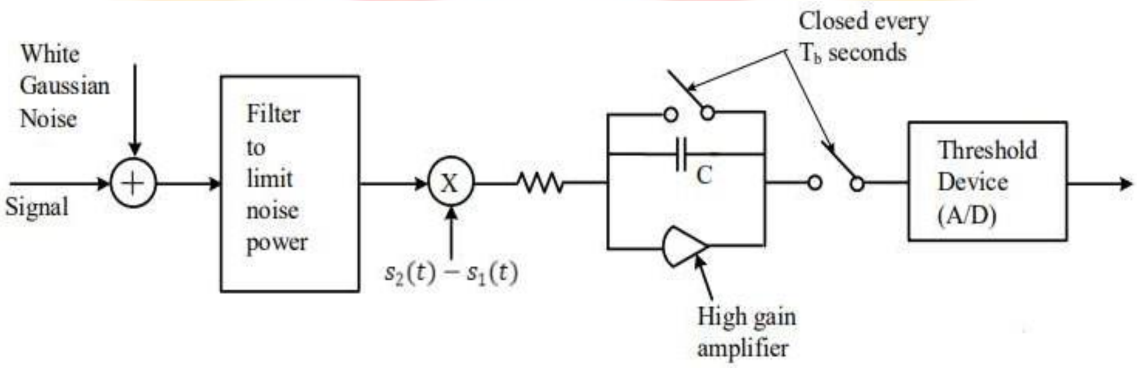
Demodulation of ASK

Coherent Detection of ASK (Integrate and Dump):

- The input to the receiver consists of an ASK signal that is corrupted by AWGN.
- The receiver integrates the product of the signal plus noise & a copy of the noise free signal over one signal interval.
- Assume that the local signal

$$s_2(t) - s_1(t) = A\cos(2\pi f_c t)$$

is carefully synchronized with the frequency & phase of the carrier received.



- Output of integrator is compared against a set threshold and at the end of each signalling interval the receiver makes the decision about which of the 2 signals s1(t) or s2(t) was present at its input during the signalling interval.
- Errors might occur in the demodulation process because of noise.
- Assume

$$s_1(t) = 0$$

$$s_2(t) = A\cos(2\pi f_c t)$$

$$s_2(t) - s_1(t) = A\cos(2\pi f_c t)$$

- The signalling components of the receiver output at the end of the signalling interval are

$$s_{01}(kT_b) = \int_0^{T_b} s_1(t)[s_2(t) - s_1(t)]dt = 0$$

$$s_{02}(kT_b) = \int_0^{T_b} s_2(t)[s_2(t) - s_1(t)]dt = \frac{A^2}{2}T_b$$

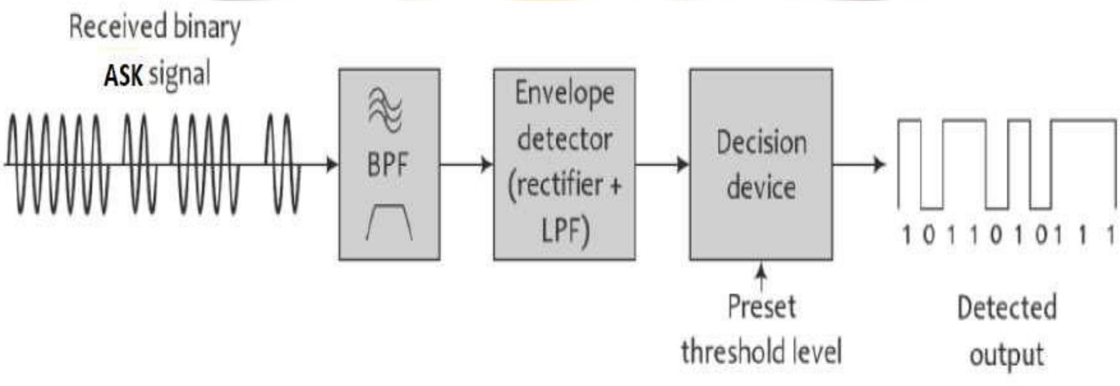
- The optimum threshold setting in the receiver is

$$V_{th} = \frac{s_{o1}(kT_b) + s_{o2}(kT_b)}{2} = \frac{A^2}{4} T_b$$

- The receiver decodes the kth transmitted bit as 1 if the output at the kth signalling interval is greater than Vth , as a '0' otherwise.

Non Coherent ASK detection

This scheme involves detection in the form of 'rectifier' & 'low pass filter'.



Block Diagram of Noncoherent ASK Demodulator

- Input to the receiver is

$$v(t) = s(t) + n_i(t)$$

Where

$$v(t) = \begin{cases} A\cos(\omega_c t) + n_i(t) \,; When\ b_k = 1 \\ n_i(t) \qquad\qquad\qquad ; When\ b_k = 0 \end{cases}$$

- $n_i(t)$ represents represents AWGN with zero mean at the receiver input.

Now if the BPF is assumed to have BW of $2/T_b$ centred at fc , then it passes the signal component without much distortion.

- The filter output will be

$$Y(t) = A_k \cos(\omega_c t) + n(t)$$

- Where A_k=A when the k^{th} transmitted bit b_k=1, and A_k=0, when b_k=0
- The above equ can be written in envelope and phase form as

$$Y(t) = R(t)\cos(\omega_c t + \theta(t))$$

$$R(t) = \sqrt{[A_k + n_c(t)]^2 + [n_s(t)]^2}$$

Where $n_c(t)$ and $n_s(t)$ are the quadrature components of narrow band noise

Advantages and Disadvantages of ASK

- **Advantages**
  - Simple to design, easy to generate and detect.
  - Requires low Bandwidth
  - Requires less energy to transmit the binary data.
- **Disadvantages**
  - Susceptible to sudden amplitude variations due to noise and interference.

Applications of ASK

- Mostly used for very low-speed data rate (upto 1200bps) requirements on voice grade lines in telemetry applications.
- Used to transmit digital data over optical fibre for LED –based optical transmitters.
- Wireless infrared transmissions using a directed beam or diffuse light in wireless LANs applications.

Frequency Shift Keying

- In Binary FSK, the frequency of the carrier is shifted according to the binary symbol. Phase unaffected.
- That is there are 2 different frequency signals according to binary symbols.
- Let there be a frequency shift by Ω.
- If b(t)=1, then

$$s_H(t) = \sqrt{2P_s}\cos(2\pi f_c + \Omega)\, t$$

b(t)=0, then

$$s_L(t) = \sqrt{2P_s}\cos(2\pi f_c - \Omega)\, t$$

- Hence there is increase or decrease in frequency by Ω.

Conversion table for BFSK representation

b(t) Input	d(t)	$P_H(t)$	$P_L(t)$
1	+1V	+1V	0V
0	-1V	0V	+1V

- FSK equ can be written as

$$s(t) = \sqrt{2P_s}\cos[(2\pi f_c + d(t)\Omega)t]$$

- Hence if symbol '1' is to be transmitted then the carrier frequency will be

$$f_c + \frac{\Omega}{2\pi}$$

- If the symbol '0' is to be transmitted then the carrier frequency will be

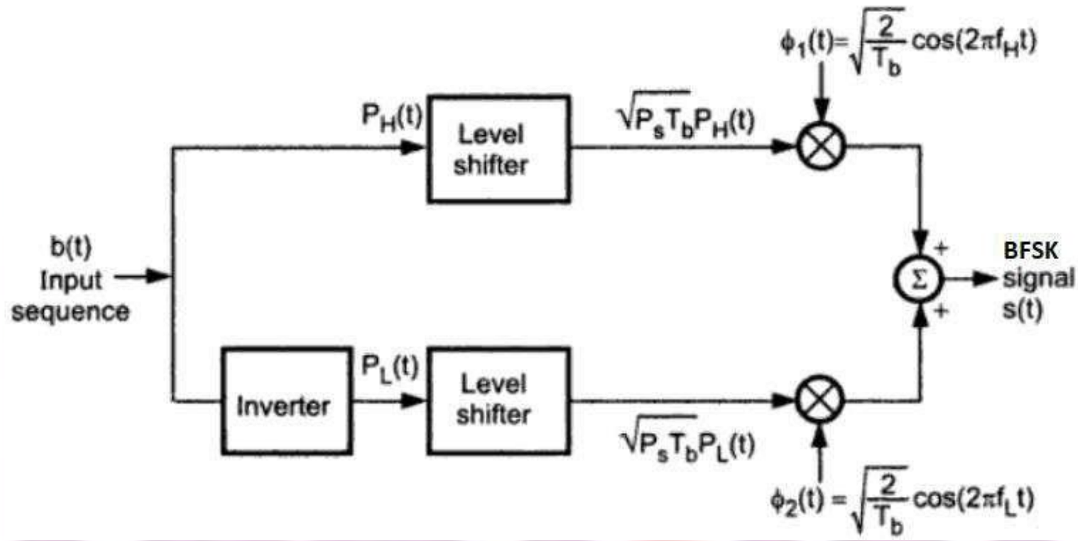
$$f_c - \frac{\Omega}{2\pi}$$

- Thus

$$f_H = f_c + \frac{\Omega}{2\pi} \dots \dots symbol\ '1'$$

$$f_L = f_c - \frac{\Omega}{2\pi} \dots \dots symbol\ '0'$$

Generation of BFSK



- $P_H(t)$ is same as b(t) and $P_L(t)$ is inverted version of b(t)
- $P_H(t)$ and $P_L(t)$ are Unipolar signals.
- The level shifter converts '+1' to $\sqrt{PsTb}$ and the zero level is unaffected.
- Further there are product modulators after the level shifters.

- The two carrier signals $\varphi_1(t)$ or $\varphi_2(t)$ are used which are orthogonal to each other. $f_H - f_L = 2f_b$
- The adder then adds the 2 signals coming from the multipliers, but outputs from the multipliers are not possible at the same time. This is because $P_H(t)$ and $P_L(t)$ are complementary to each other.

PSD of BFSK Signal

- BFSK signal s(t) can be written as

$$s(t) = \sqrt{2P_s}\, P_H(t)\cos(2\pi f_H t) + \sqrt{2P_s}\, P_L(t)\cos(2\pi f_L t)$$

- Let us convert those coefficients in bipolar form as follows

$$P_H(t) = \frac{1}{2} + \frac{1}{2}P_H'(t)$$

$$P_L(t) = \frac{1}{2} + \frac{1}{2}P_L'(t)$$

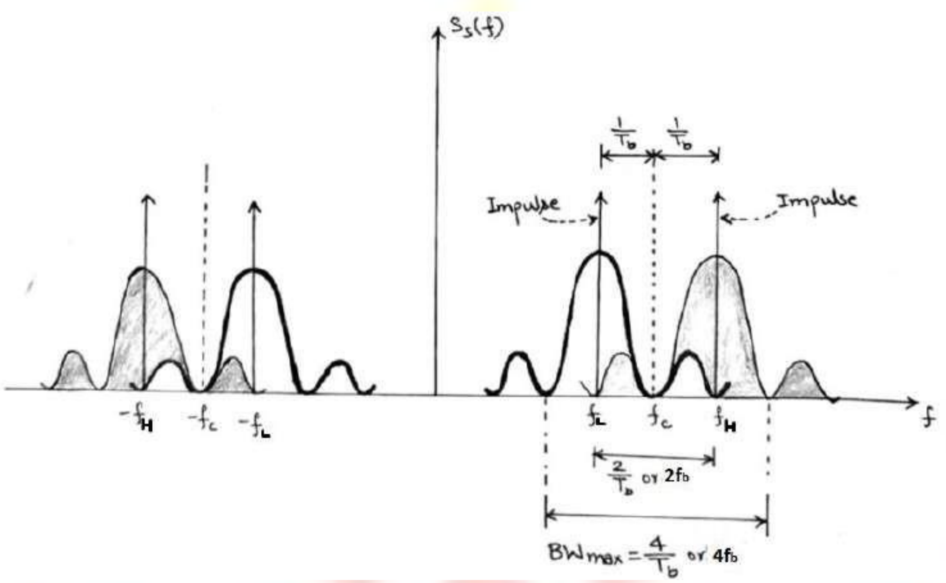
- Here $P_H'(t)$ and $P_L'(t)$ will be bipolar, alternating between +1 and -1, and complementary.
- Now s(t) can be written as

$$s(t) = \sqrt{2P_s}\,[\frac{1}{2} + \frac{1}{2}P_H'(t)]\cos(2\pi f_H t) + \sqrt{2P_s}\,[\frac{1}{2} + \frac{1}{2}P_L'(t)]\cos(2\pi f_L t)$$

- The below equation is used to find the PSD of BFSK Signal

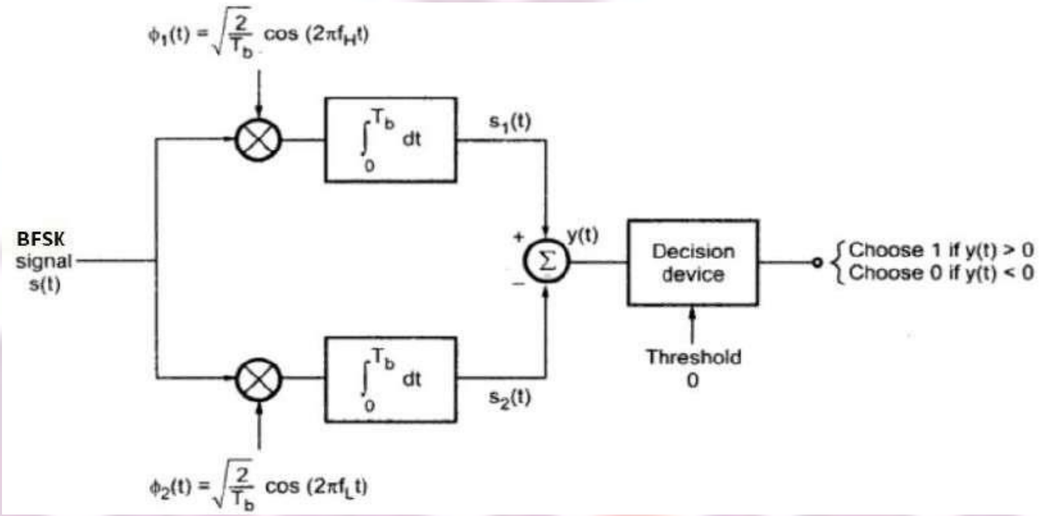
$$s(t) = \sqrt{\frac{P_s}{2}}\cos(2\pi f_H t) + \sqrt{\frac{P_s}{2}}\cos(2\pi f_L t) + \sqrt{\frac{P_s}{2}}\,P_H'(t)\cos(2\pi f_H t) + \sqrt{\frac{P_s}{2}}\,P_L'(t)\cos(2\pi f_L t)$$

PSD and BW of BFSK Signal



Coherent Detection of BFSK Signal

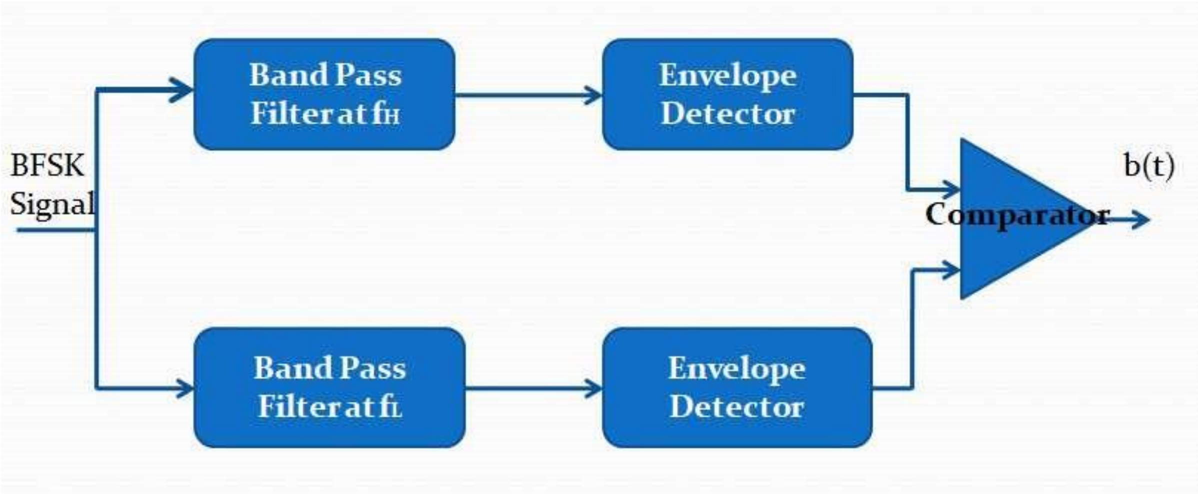
The incoming FSK signal is multiplied by a recovered carrier signal that has the exact same frequency and phase as the transmitter reference. However, the two transmitted frequencies (the mark and space frequencies) are not generally continuous; it is not practical to reproduce a local reference that is coherent with both of them. Consequently, coherent FSK detection is seldom used.



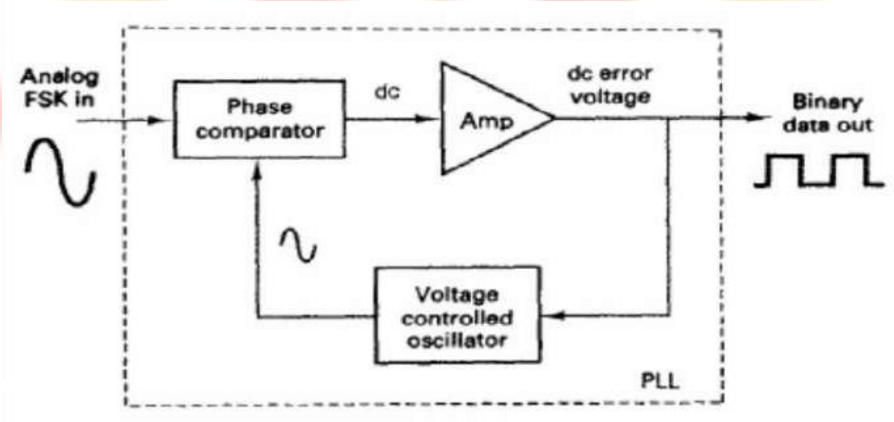
Non-coherent Detection of BFSK Signal

The FSK input signal is simultaneously applied to the inputs of both band pass filters (BPFs) through a power splitter. The respective filter passes only the mark or only the space frequency on to its respective envelope detector. The envelope detectors, in turn, indicate the total power in each pass band, and the comparator responds to the larger of the two powers.

This type of FSK detection is referred to as non coherent detection.



Detection of BFSK signal using PLL



Principle : The error voltage in PLL is proportional to difference between the phase or frequency of the input signal and VCO frequency.

Block diagram and operation

- Fig. shows the block diagram of FSK detection using PLL. The free running frequency of the voltage controlled oscillator (VCO) is kept in between F_H and F_L.
- When F_H is transmitted, the error voltage becomes positive, and hence binary output goes high. It remains high as long as F_H is transmitted for a bit period.
- When F_L is transmitted, the error voltage becomes negative and hence binary goes low. It remains low as long as F_L is transmitted for a bit period.

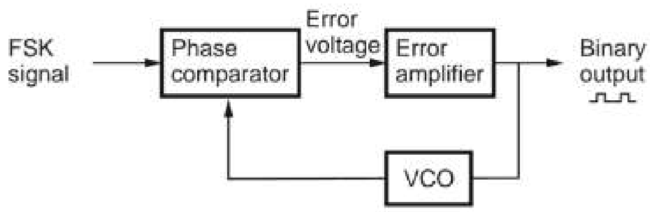


Fig. FSK detection using PLL

- The VCO generates free running frequency between F_H and F_L. Hence phase comparator detects the difference between VCO frequency and F_H or F_L.
- PLL detector is a non coherent type of detector.

Advantages and Disadvantages of FSK

- **Advantages**
  - It is less susceptible to errors than ASK.
  - Better noise immunity than ASK.
  - Peak frequency offset is constant and always at its maximum.
  - The highest fundamental frequency is equal to half the information bit rate.
  - Relatively easy to implement.
- **Disadvantages**
  - Not efficient in terms of transmission bandwidth requirement
  - It has poorer error performance than PSK or QAM.

Applications of FSK

- Used in low-speed modems (up to 1200bps) over analog voice-band telephone lines.
- Finds applications in pager systems, HF radio tele-type transmission systems, and LANs using coaxial cables.

Binary Phase Shift Keying

- Principle of BPSK
- In BPSK the binary symbol '1' and '0' modulate the phase of the carrier.

Let the carrier be

$$s(t) = Acos(2\pi f_c t)$$

'A' represents peak of the sinusoidal carrier

$$A = \sqrt{2P_s}$$

When the symbol is changed, then phase of the carrier is changed by 180^o

Consider, for symbol '1'

$$s_1(t) = \sqrt{2P_s}\cos(2\pi f_c t)$$

- For symbol '0'

$$s_2(t) = \sqrt{2P_s}\cos\,(2\pi f_c t + \pi)$$

- Therefore

$$s_2(t) = -\sqrt{2P_s}\cos\,(2\pi f_c t)$$

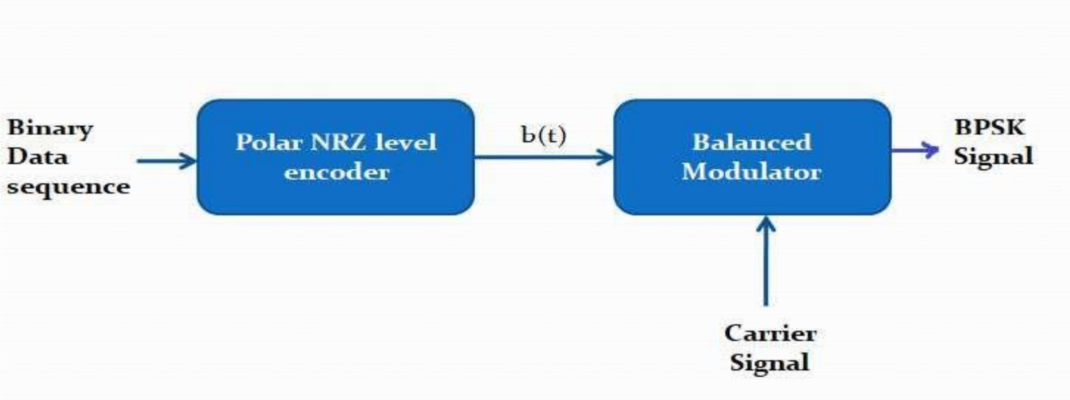
- Which implies

$$s(t) = b(t)\sqrt{2P_s}\cos\,(2\pi f_c t)$$

Where b(t)=+1; for symbol '1'

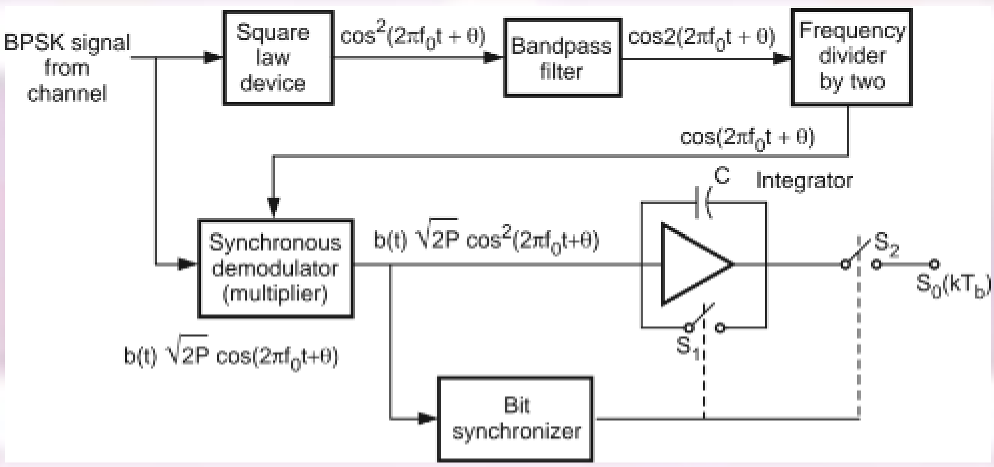
b(t)=-1; for symbol '0'

Generation of BPSK



Coherent Reception of BPSK Signal

Operation of the receiver



- **Phase shift in the received signal:**

$$s(t) = b(t)\sqrt{2P_s}\cos\,(2\pi f_c t + \theta)$$

- The signal undergoes the phase change depending upon the time delay from transmitter to receiver. Let the phase shift be

- **Square Law device:**
  - From the received signal carrier is separated. Since it is coherent detection.
- The output of square law device

$$\cos^2(2\pi f_c t + \theta) = \frac{1}{2} + \frac{\cos 2(2\pi f_c t + \theta)}{2}$$

- **Band Pass Filter:**
  - The signal is passed through band pass filter with centre frequency 2fc.
  - BPF removes DC level and its output is

$$\cos 2(2\pi f_c t + \theta)$$

- **Frequency Divider:**
  - The signal is passed through a frequency divider by 2.
  - Therefore at the output of the frequency divider we get the carrier signal whose frequency is fc *i.e.,*

$$\cos(2\pi f_c t + \theta)$$

- **Synchronous Demodulator:**
  - The synchronous demodulator multiplies the input signal & the recovered carrier.
- Therefore at the output of multiplier we get

$$\begin{aligned} b(t)\sqrt{2P}\cos(2\pi f_0 t+\theta)\times\cos(2\pi f_0 t+\theta) &= b(t)\sqrt{2P}\cos^2(2\pi f_0 t+\theta) \\ &= b(t)\sqrt{2P}\times\frac{1}{2}[1+\cos 2(2\pi f_0 t+\theta)] \\ &= b(t)\sqrt{\frac{P}{2}}[1+\cos 2(2\pi f_0 t+\theta)] \end{aligned}$$

- **Bit synchronizer and integrator:**
  - The above signal is applied to the bit synchronizer & integrator.
  - The integrator integrates the signal over one bit period. The bit synchronizer takes care of starting and end times of a bit.
  - At the end of the bit duration the bit synchronizer closes switch s2 temporarily connecting the output of integrator to the decision device.

- Synchronizer then opens s2 and closes s1 temporarily to reset the integrator.

- **Output of integrator:** In the k^{th} bit interval the output can be written as

$$s_o(kT_b) = b(kT_b)\sqrt{\frac{P}{2}}\int_{(k-1)T_b}^{kT_b}[1+\cos 2(2\pi f_0 t+\theta)]\,dt$$

The above equation gives the output of an interval for k^{th} bit. Therefore integration is performed from $(k-1)T_b$ to kT_b. Here T_b is the one bit period.

- We can write the above equation as,

$$s_o(kT_b) = b(kT_b)\sqrt{\frac{P}{2}}\left[\int_{(k-1)T_b}^{kT_b} 1\,dt + \int_{(k-1)T_b}^{kT_b}\cos 2(2\pi f_0 t+\theta)\,dt\right]$$

Here $\int_{(k-1)T_b}^{kT_b}\cos 2(2\pi f_0 t+\theta)\,dt = 0$, because average value of sinusoidal waveform is zero if integration is performed over full cycles. Therefore we can write above equation as,

$$s_o(kT_b) = b(kT_b)\sqrt{\frac{P}{2}}\int_{(k-1)T_b}^{kT_b} 1\,dt$$

$$= b(kT_b)\sqrt{\frac{P}{2}}\,[t]_{(k-1)T_b}^{kT_b}$$

$$= b(kT_b)\sqrt{\frac{P}{2}}\{kT_b-(k-1)T_b\}$$

$$\therefore \quad \boxed{s_o(kT_b) = b(kT_b)\sqrt{\frac{P}{2}}\,T_b}$$

This equation shows that the output of the receiver depends on input i.e.

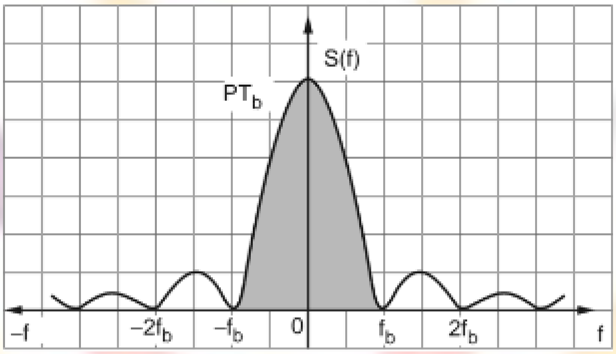
$$s_o(kT_b) \propto b(kT_b)$$

Depending upon the value of $b(kT_b)$, the output $s_0(kT_b)$ is generated in the receiver.

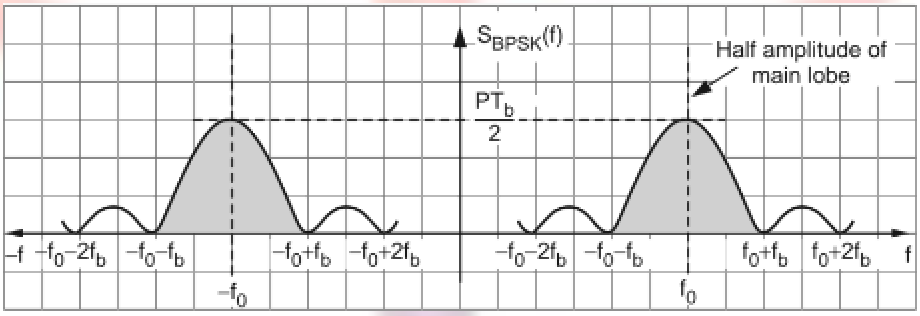
The signal is then given to the decision device, which decides whether transmitted symbol was zero or one.

PSD of BPSK

- PSD of polar NRZ baseband signal b(t)=+1 or -1



Power spectral density of BPSK signal



Inter channel interference and ISI

- Inter channel interference avoided by filtering.
- Because of filtering phase distortion takes place resulting in ISI.
- ISI can be reduced to some extent by using equalizers at the receiver.
- Equalizers have reverse effect to the filters adverse effects.

Differential Phase Shift Keying

- DPSK is an alternative form of digital modulation where the binary input information is contained in the difference between two successive signalling elements rather than the absolute phase.
- With DPSK, it is not necessary to recover a phase coherent carrier.
- Instead a received signalling element is delayed by one signalling element time slot and then compared with the next received signalling element.

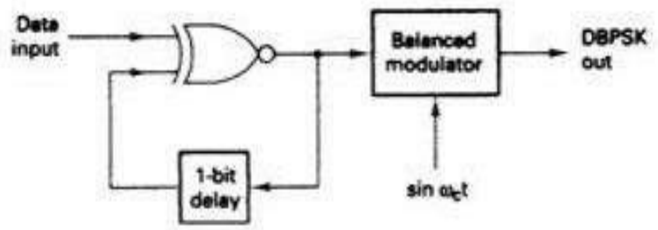
- The difference in phase of two signalling elements determines the logic condition of the data.

DPSK Transmitter

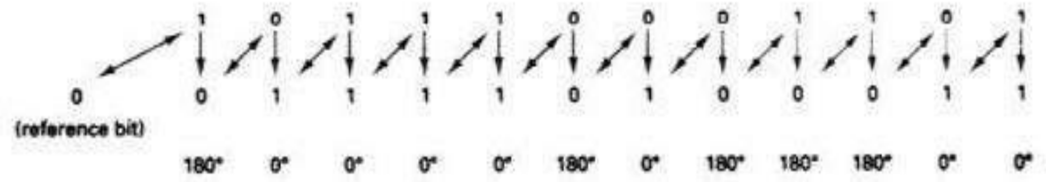
The figure (a) below shows a simplified block diagram of a *differential binary phase-shift keying* (DBPSK) transmitter. An incoming information bit is XNORed with the preceding bit prior to entering the BPSK modulator (balanced modulator).

For the first data bit, there is no preceding bit with which to compare it. Therefore, an initial reference bit is assumed. Figure (b) shows the relationship between the input data, the XNOR output data, and the phase at the output of the balanced modulator. If the initial reference bit is assumed logic 1, the output from the XNOR circuit is simply the complement of that shown.

In Figure b, the first data bit is XNORed with the reference bit. If they are the same, the XNOR output is logic 1; if they are different, the XNOR output is logic 0. The balanced modulator operates the same as a conventional BPSK modulator; a logic I produces +sin ωct at the output, and A logic 0 produces –sin ωct at the output.



(a)

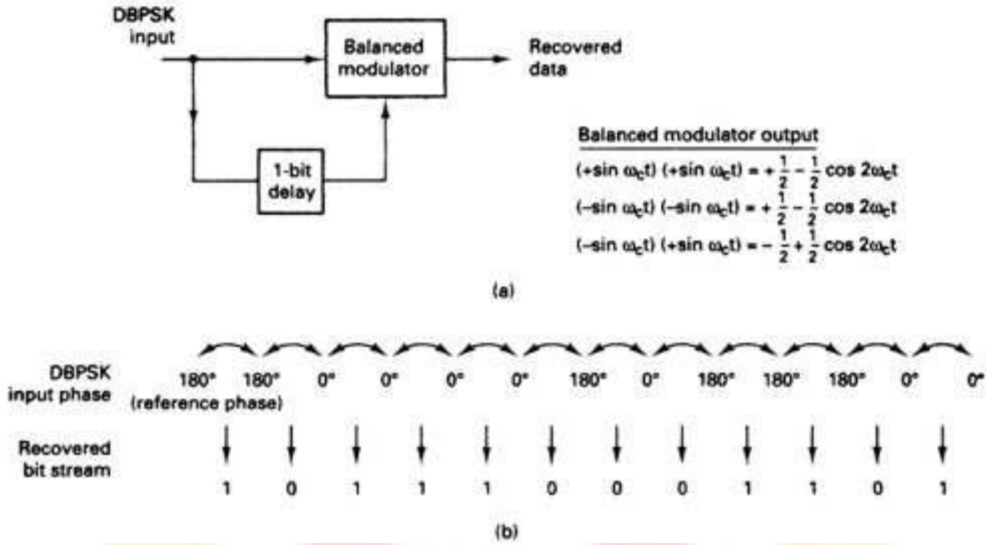


(b)

DPSK Receiver

The figure below shows the block diagram and timing sequence for a DBPSK receiver. The received signal is delayed by one bit time, then compared with the next signalling element in the balanced modulator. If they are the same, a logic 1(+ voltage) is generated. If they are different, a logic 0 (- voltage) is generated.

If the reference phase is incorrectly assumed, only the first demodulated bit is in error. Differential encoding can be implemented with higher-than-binary digital modulation schemes, although the differential algorithms are much more complicated than for DBPSK.



Advantages

- Simplicity of circuit.
- No carrier recovery circuit needed.
- BW requirement of DPSK (fb) is reduced compared to that of BPSK (2fb).

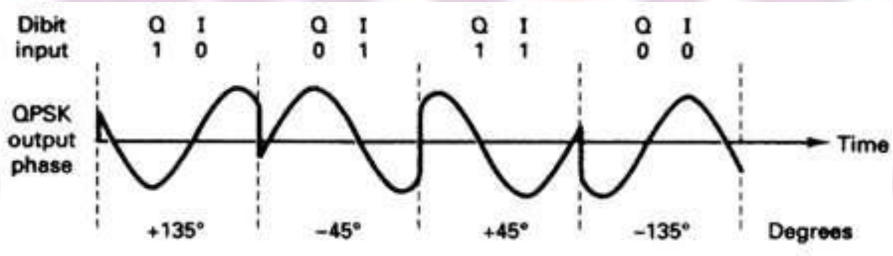
Disadvantages

- Disadvantage of DBPSK is, that it requires between 1 dB and 3 dB more signal-to-noise ratio to achieve the same bit error rate as that of absolute PSK.

Quadrature Phase Shift Keying (QPSK)

This is the phase shift keying technique, in which the sine wave carrier takes four phase reversals such as 45°, 135°, -45°, and -135°.

If these kinds of techniques are further extended, PSK can be done by eight or sixteen values also, depending upon the requirement. The following figure represents the QPSK waveform for two bits input, which shows the modulated result for different instances of binary inputs.

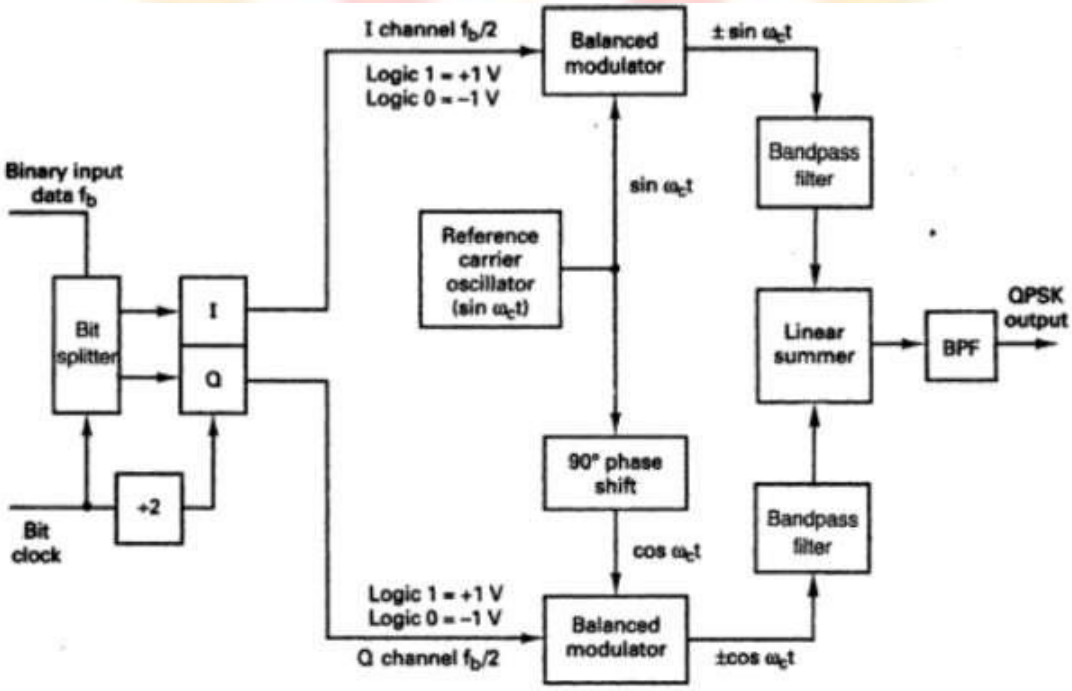


QPSK transmitter

A block diagram of a QPSK modulator is shown in Figure below. Two bits (a dibit) are clocked into the bit splitter. After both bits have been serially inputted, they are simultaneously parallel outputted.

The I bit modulates a carrier that is in phase with the reference oscillator (hence the name "I" for "in phase" channel), and the Q bit modulate, a carrier that is 90° out of phase. For a logic 1 = + 1 V and a logic 0= - 1 V, two phases are possible at the output of the I balanced modulator (+sin ωct and - sin ωct), and two phases are possible at the output of the Q balanced modulator (+cos ωct), and (-cos ωct).

When the linear summer combines the two quadrature (90° out of phase) signals, there are four possible resultant phasors given by these expressions: + sin ωct + cos ωct, + sin ωct - cos ωct, -sin ωct + cos ωct, and -sin ωct - cos ωct.



Example:

For the QPSK modulator shown in the above figure, construct the truth table, Phasor diagram, and constellation diagram.

Solution:

For a binary data input of Q = 0 and I= 0, the two inputs to the I balanced modulator are -1 and sin ωct, and the two inputs to the Q balanced modulator are -1 and cos ωct.

Consequently, the outputs are

I balanced modulator =(-1)(sin ωct) = -1 sin ωct

Q balanced modulator =(-1)(cos ωct) = -1 cos ωct and the output of the linear summer is

$$-1 \cos \omega ct - 1 \sin \omega ct = 1.414 \sin (\omega ct - 135°)$$

For the remaining dibit codes (01, 10, and 11), the procedure is the same. The results are shown in figure below.

Binary input Q	Binary input I	QPSK output phase
0	0	−135°
0	1	−45°
1	0	+135°
1	1	+45°

(a)

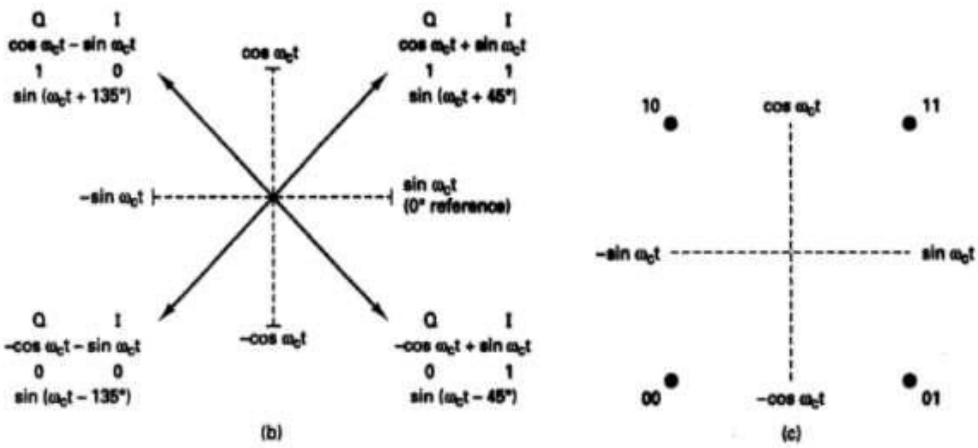
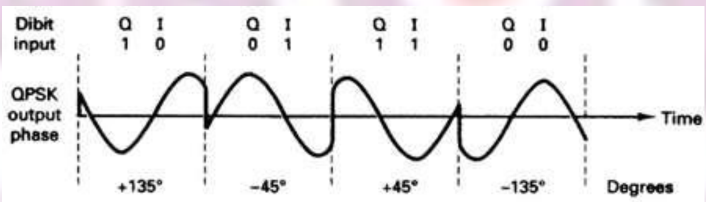


Figure QPSK modulator: (a) truth table; (b) Phasor diagram; (c) constellation diagram

In above figures b and c, it can be seen that with QPSK each of the four possible outputs Phasor has exactly the same amplitude. Therefore, the binary information must be encoded entirely in the phase of the output signal

In figure b, it can be seen that the angular separation between any two adjacent Phasor in QPSK is 90°.Therefore, a QPSK signal can undergo almost a+45° or -45° shift in phase during transmission and still retain the correct encoded information when demodulated at the receiver.

The figure below shows the output phase-versus-time relationship for a QPSK modulator.



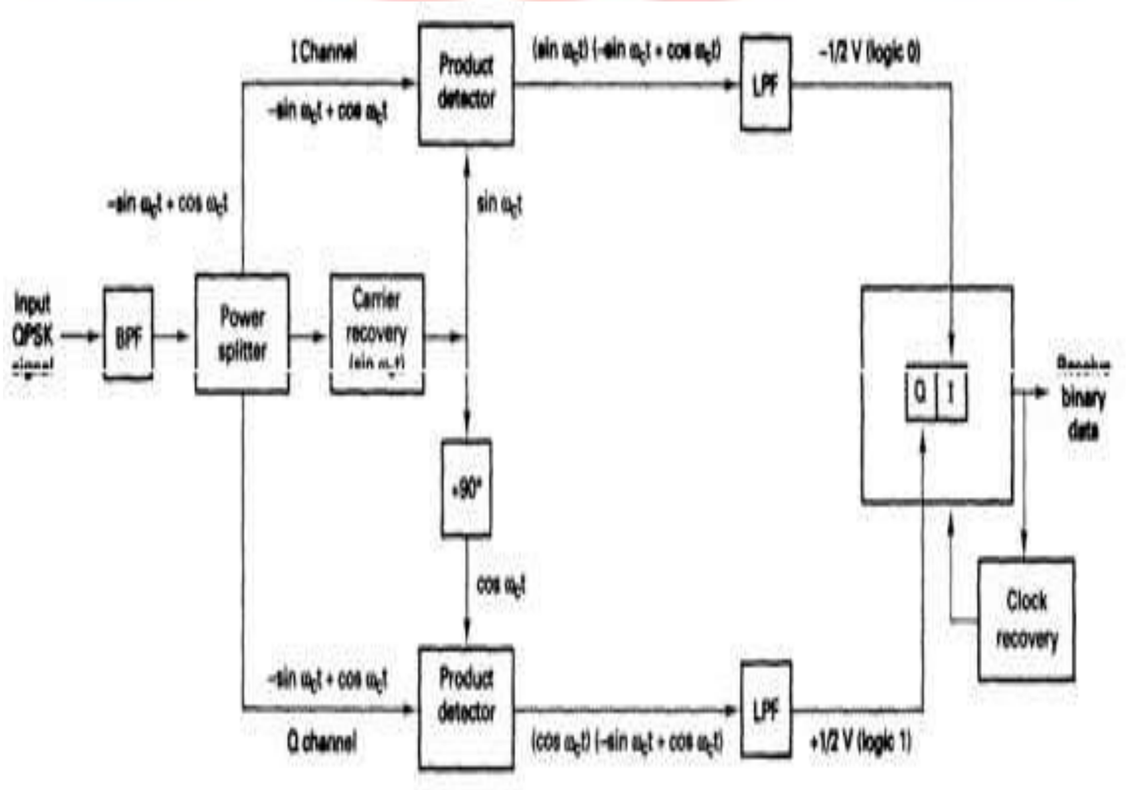
QPSK Receiver:

The power splitter directs the input QPSK signal to the I and Q product detectors and the carrier recovery circuit. The carrier recovery circuit reproduces the original transmit carrier oscillator signal.

The recovered carrier must be frequency and phase coherent with the transmit reference carrier. The QPSK signal is demodulated in I and Q product detectors, which generate the original I and Q data bits. The outputs of the product detectors are fed to the bit combining circuit, where they are converted from parallel I and Q data channels to a single binary output data stream.

The incoming QPSK signal may be any one of the four possible output phases shown in above figures.

To illustrate the demodulation process, let the incoming QPSK signal is -sin ωct + cos ωct. Mathematically, the demodulation process is as follows.



The received QPSK signal (-sin ωct + cos ωct) is one of the inputs to I product detector. The other input is the recovered carrier (sin ωct). The output of the I product detector is

$$\mathrm{I} = \underbrace{(-\sin \omega_c t + \cos \omega_c t)}_{\text{QPSK input signal}} \underbrace{(\sin \omega_c t)}_{\text{carrier}}$$

$$= (-\sin \omega_c t)(\sin \omega_c t) + (\cos \omega_c t)(\sin \omega_c t)$$

$$= -\sin^2 \omega_c t + (\cos \omega_c t)(\sin \omega_c t)$$

$$= -\frac{1}{2}(1 - \cos 2\omega_c t) + \frac{1}{2}\sin(\omega_c + \omega_c)t + \frac{1}{2}\sin(\omega_c - \omega_c)t$$

$$\mathrm{I} = -\frac{1}{2} + \underbrace{\frac{1}{2}\cos 2\omega_c t}_{\text{(filtered out)}} + \frac{1}{2}\sin 2\omega_c t + \underbrace{\frac{1}{2}\sin 0}_{\text{(equals 0)}}$$

$$= -\frac{1}{2}\mathrm{V} \text{ (logic 0)}$$

Again, the received QPSK signal (-sin ωct + cos ωct) is one of the inputs to the Q product detector. The other input is the recovered carrier shifted 90° in phase (cos ωct). The output of the Q product detector is

$$\mathrm{Q} = \underbrace{(-\sin \omega_c t + \cos \omega_c t)}_{\text{QPSK input signal}} \underbrace{(\cos \omega_c t)}_{\text{carrier}}$$

$$= \cos^2 \omega_c t - (\sin \omega_c t)(\cos \omega_c t)$$

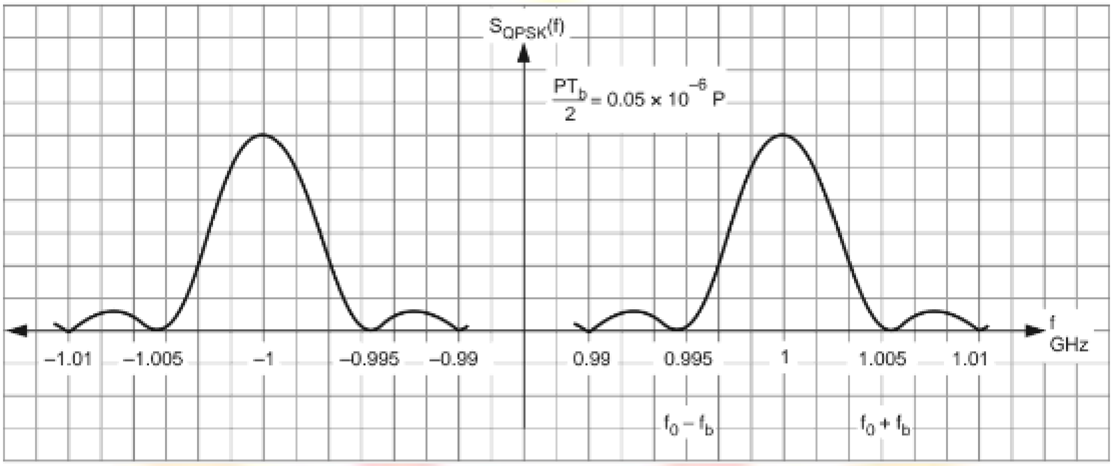
$$= \frac{1}{2}(1 + \cos 2\omega_c t) - \frac{1}{2}\sin(\omega_c + \omega_c)t - \frac{1}{2}\sin(\omega_c - \omega_c)t$$

$$\mathrm{Q} = \frac{1}{2} + \underbrace{\frac{1}{2}\cos 2\omega_c t}_{\text{(filtered out)}} - \frac{1}{2}\sin 2\omega_c t - \underbrace{\frac{1}{2}\sin 0}_{\text{(equals 0)}}$$

$$= \frac{1}{2}\mathrm{V} \text{(logic 1)}$$

The demodulated I and Q bits (0 and 1, respectively) correspond to the constellation diagram and truth table for the QPSK modulator shown in Figure.

Power spectral Density of QPSK



OPTIMAL RECEPTION OF DIGITAL SIGNALS

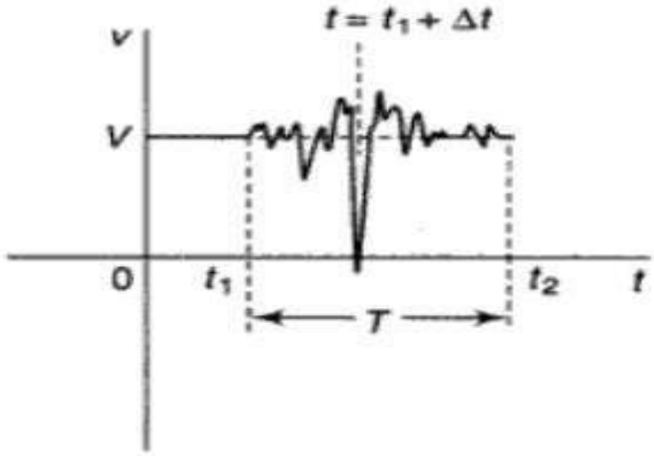
Introduction

- Digital data can be transmitted directly or as is usually the case, by modulating a carrier.
- The received signal is corrupted by noise and hence there is a finite probability that the receiver will make an error in determining within each time interval, whether a 1 or 0 is transmitted.

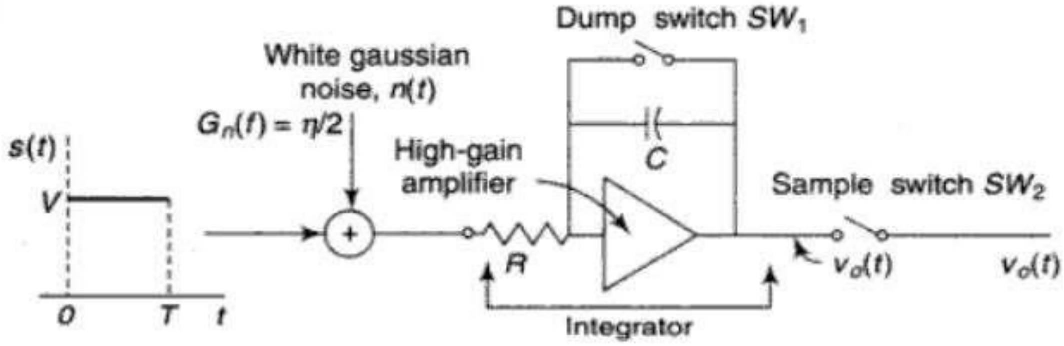
Therefore, an optimal receiver design with an objective to reduce error probability in reception is must.

Baseband Signal Receiver

- Consider that a binary encoded signal consists of a time sequence of voltage levels +V or –V
- With noise present, the received signal and noise together will yield sample values generally different from ± V.
- **Assumption:** Noise is Gaussian and therefore the noise voltage has probability density which is entirely symmetrical with respect to zero volts.
- Probability that noise has increased the sample value is same as the probability that the noise has decreased the sample value.
- If sample value is positive the transmitted level was +V, and if the sample value is negative the transmitted level was –V.
- It is possible that at the sampling time the noise voltage may be of magnitude larger than V and of a polarity opposite to the polarity assigned to the transmitted bit.



- The probability of error can be reduced by processing the received signal plus noise in such a manner that we are then able to find sample time where the sample voltage due to the signal is emphasized relative to the sample voltage due to the noise.



A receiver for a binary coded signal.

Operation of Baseband signal Receiver

- The operation of the receiver during each bit interval is independent of the waveform during past and future bit intervals.
- Signal s(t) and white Gaussian noise n(t) of PSD η/2 is presented to an integrator.
- At time t=0+ we require that capacitor C be uncharged which is ensured by a brief closing of the switch sw1 at time t=0-, thus relieving C of any charge it may have acquired during the previous interval.
- Sample is taken at the output of the integrator by closing this sampling switch sw2.
- This sample is taken at the end of the bit interval at t=T.
- Signal processing is described by the phrase integrate and dump.
- Dump- refers to the abrupt discharge of the capacitor after each sampling.

Peak Signal to RMS Noise output Voltage Ratio

- The integrator yields an output which is the integral of its input multiplied by 1/RC. Using τ=RC, we have

$$v_o(T) = \frac{1}{\tau}\int_0^T [s(t) + n(t)]\, dt = \frac{1}{\tau}\int_0^T s(t)\, dt + \frac{1}{\tau}\int_0^T n(t)\, dt$$

The sample voltage due to the signal is

$$s_o(T) = \frac{1}{\tau}\int_0^T V\, dt = \frac{VT}{\tau}$$

The sample voltage due to the noise is

$$n_o(T) = \frac{1}{\tau}\int_0^T n(t)\, dt$$

This noise-sampling voltage $n_o(T)$ is a gaussian random variable in contrast with $n(t)$, which is a gaussian random process.

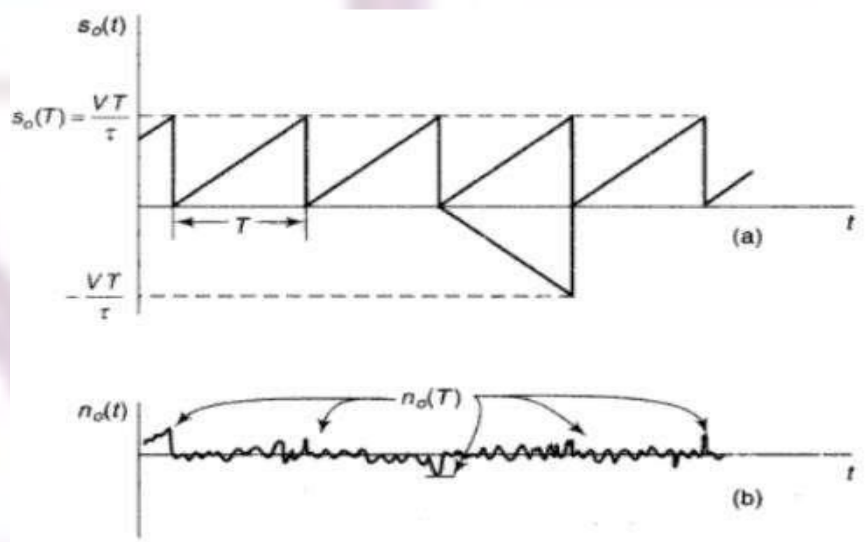
The variance of $n_o(T)$

$$\sigma_o^2 = \overline{n_o^2(T)} = \frac{\eta T}{2\tau^2}$$

- $n_o(T)$ has a gaussian probability density.
- The output of the integrator, before the sampling switch is

$$\boldsymbol{v_o(t) = s_o(t) + n_o(t)}$$

The signal output $s_o(t)$ is a ramp, in each bit interval of duration T. At the end of the interval the ramp attains the voltage $s_o(T)$ which is $+VT/\tau$ or VT/τ, depending on whether the bit is 1 or 0



- At the end of each interval the switch SW1 closes momentarily to discharge the capacitor so that $s_o(t)$ drops to zero.

- The noise $n_o(t)$ also starts each interval with $n_o(0)=0$ and has the random value $n_o(T)$ at the end of each interval.
- The sampling switch SW2 closes briefly just before the closing of SW1 and hence reads the voltage

$$\boldsymbol{v_o(T) = s_o(T) + n_o(T)}$$

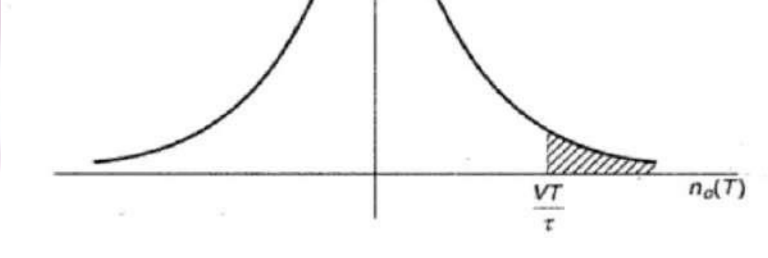
- The output signal voltage to be as large as possible in comparison with the noise voltage.
- Hence a figure of merit of interest is the signal-to-noise ratio

$$\frac{[s_o(T)]^2}{[n_o(T)]^2} = \frac{2}{\eta} V^2 T$$

- SNR increases with increasing bit duration T and it depends on V^2T which is the normalized energy of the bit signal.
- The integrator filters the signal and noise such that the signal voltage increases linearly with time, while the standard deviation (rms value) of the noise increases more slowly, as .
- Thus the integrator enhances the signal relative to the noise, and this enhancement increases with time.

Probability of Error (Pe) of Integrate-and-dump receiver

- **Function of a receiver:** To distinguish the bit 1 from the bit 0 in the presence of noise.
- A most important characteristic is the probability that an error will be made in such a determination.
- The probability density of the noise sample no(T) is Gaussian and hence appears as follows



- The density is therefore given by

$$f[n_o(T)] = \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}}$$

- Where $\boldsymbol{\sigma_o^2}$, the variance, is $\boldsymbol{\sigma_o^2 = \overline{n_o^2(T)}}$ given by equ.

$$\sigma_0^2 = \frac{\eta T}{2\tau^2}$$

- Suppose that during some bit interval the input signal voltage is held at, say –V.
- Then, at the sample time, the signal sample voltage is

$$s_0(T) = -\frac{VT}{\tau}$$

while the noise sample is $n_o(T)$

- If n_o (T) is positive and larger in magnitude than $VT/\boldsymbol{\tau}$, the total sample voltage

$$\boldsymbol{v_o(T) = s_o(T) + n_o(T)}$$

will be positive. Such a positive sample voltage will result in an error.

- The probability of such a misinterpretation, that is, the probability that

$$n_0(T) > \frac{VT}{\tau}$$

is given by the area of the shaded region in the figure.

- The probability of error is given by

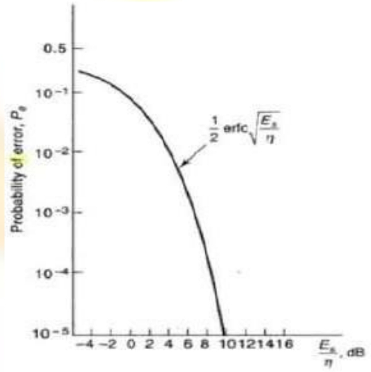
$$P_e = \int_{VT/\tau}^{\infty} f[n_o(T)]\, dn_o(T) = \int_{VT/\tau}^{\infty} \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}}\, dn_o(T)$$

- Defining, $\boldsymbol{x \equiv n_o(T)/\sqrt{2}\sigma_o}$, and using above equ.

$$P_e = \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{x=V\sqrt{T/\eta}}^{\infty} e^{-x^2}\, dx = \frac{1}{2}\operatorname{erfc}\left(V\sqrt{\frac{T}{\eta}}\right) = \frac{1}{2}\operatorname{erfc}\left(\frac{V^2T}{\eta}\right)^{1/2} = \frac{1}{2}\operatorname{erfc}\left(\frac{E_s}{\eta}\right)^{1/2}$$

- In which $E_s = V^2T$, is the signal energy of a bit.
- If the signal voltage were held instead at +V during some bit interval, then it is clear from the symmetry of the situation that the probability of error would again be given by Pe in the above equ.

- The probability of error Pe is plotted below



- Pe decreases rapidly as E_s/η increases. The maximum value of Pe is ½.
- Thus, even if the signal is entirely lost in the noise so that any determination of the receiver is a sheer guess, the receiver cannot be wrong more than half the time on the average.

Optimum Receiver for both Baseband and Pass band

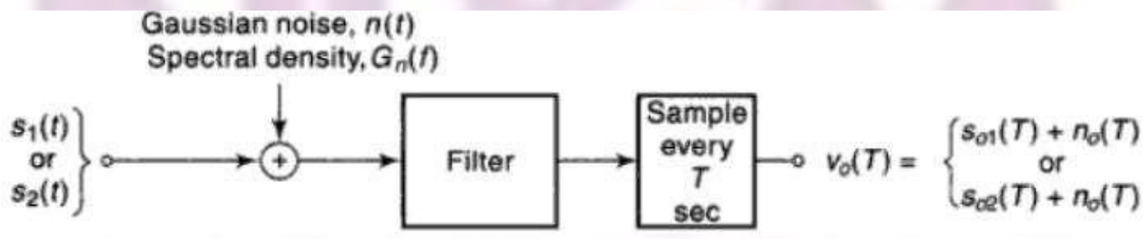
- Assume that the received signal is a binary waveform.
- One binary bit is represented by a signal waveform $s_1(t)$ which persists for time T, while the other bit is represented by the waveform $s_2(t)$which also lasts for an interval.
- In the case of transmission of baseband, $s_1(t)=+V$, while $s_2(t)=-V$.
- For other modulation systems, different waveforms are transmitted. Example:

PSK signalling: $s_1(t) = A\cos(\omega ct)$

$s_2(t) = -A\cos(\omega ct)$

FSK signalling: $s_1(t)= A\cos(\omega c+\Omega)t$

$s_2(t)= A\cos(\omega c-\Omega)t$



- As shown in the above figure, the input, which s1(t) or $s_2(t)$, is corrupted by the addition of noise n(t).
- The noise is Gaussian and has a spectral density G(f).

- The signal & noise are filtered and then sampled at the end of each bit interval.
- The output sample is either $V_o(T)=s_{01}(T)+n_0(T)$

(or)

$V_o(T)=s_{02}(T)+n_0(T)$

- **Assumption:** Immediately after each sample, every energy storing element in the filter will be discharged.
- In the absence of noise the output sample would be $V_0(T)=s_{01}(T)$ or $s_{02}(T)$.
- When noise is present to minimize the probability of error one should assume that $s_1(t)$ has been transmitted if $V_0(T)$ is closer to $s_{01}(T)$ than to $s_{02}(T)$, similarly it is assumed $s_2(t)$ has been transmitted if $V_0(T)$ is closer to $s_{02}(T)$.
- Decision boundary is midway between $s_{01}(T)$ and $s_{02}(T)$.
- Example: For the Integrate and Dump system, where

$$s_{01}(T) = \frac{VT}{\tau} \quad \text{and} \quad s_{02}(T) = -\frac{VT}{\tau}$$

the decision boundary is Vo(T)=0.

- Decision boundary is

$$V_o(T) = \frac{s_{01}(T) + s_{02}(T)}{2}$$

- Example: Suppose that $s_{01}(T) > s_{02}(T)$ and that $s_2(t)$ was transmitted.
- If at the sampling time, the noise $n_0(T)$ is positive and larger in magnitude than the voltage difference,

$\frac{1}{2}[s_{01}(T) + s_{02}(T)] - s_{02}(T)$ an error will have been made .

- That is , an error will result if

$$n_o(T) \geq \frac{s_{01}(T) - s_{02}(T)}{2}$$

- Hence the probability of error is

$$P_e = \int_{[s_{o1}(T)-s_{o2}(T)]/2}^{\infty} \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}}\, dn_o(T)$$

Substituting, $\boldsymbol{x} \equiv \boldsymbol{n_o(T)}/\sqrt{2}\boldsymbol{\sigma_o}$, then

$$P_e = \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{\frac{[s_{01}(T)-s_{02}(T)]}{2\sqrt{2}\sigma_0}}^{\infty} e^{-x^2}dx$$

$$P_e = \frac{1}{2}\, erfc\left[\frac{s_{01}(T)-s_{02}(T)}{2\sqrt{2}\sigma_0}\right]$$

- The complimentary error function is monotonically decreasing function of its argument.
- Pe decreases as the difference $s_{01}(T)$-$s_{02}(T)$ becomes larger and as the rms noise voltage σo becomes smaller.
- The optimum filter, then, is the filter which maximizes the ratio

$$\gamma = \left[\frac{s_{01}(T)-s_{02}(T)}{\sigma_0}\right]$$

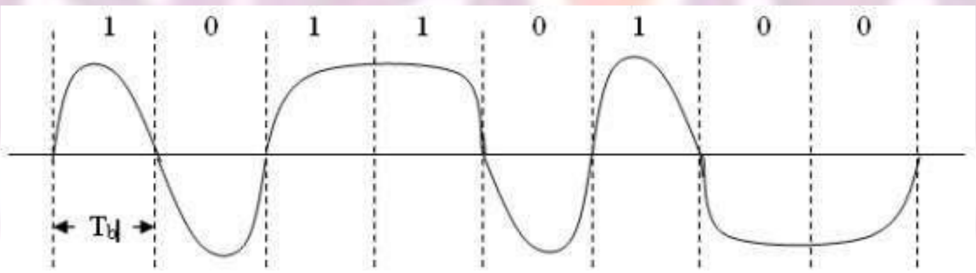
Eye Diagrams/Eye Patterns

The quality of digital transmission systems are evaluated using the bit error rate. Degradation of quality occurs in each process modulation, transmission, and detection.

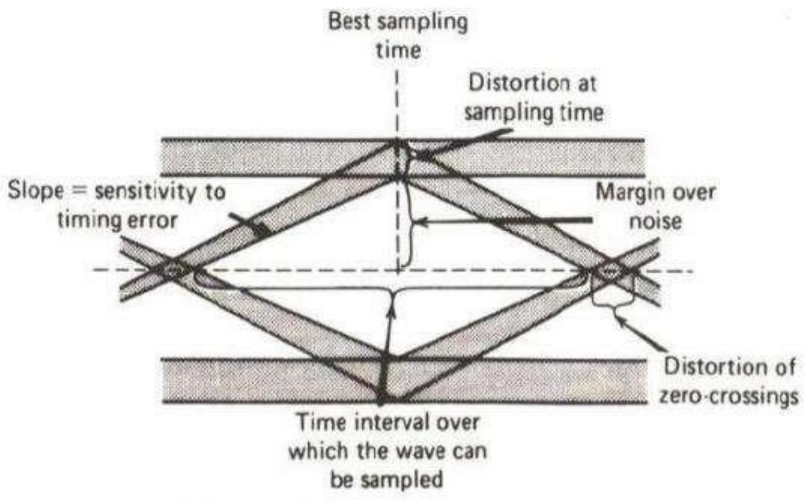
The eye pattern is experimental method that contains all the information concerning the degradation of quality. Therefore, careful analysis of the eye pattern is important in analyzing the degradation mechanism.

Eye patterns can be observed using an oscilloscope. The received wave is applied to the vertical deflection plates of an oscilloscope and the saw tooth wave at a rate equal to transmitted symbol rate is applied to the horizontal deflection plates, resulting display is eye pattern as it resembles human eye.

• The interior region of eye pattern is called **eye opening.**



We get superposition of successive symbol intervals to produce eye pattern as shown below.



Interpretation of eye pattern

The width of the eye opening defines the time interval over which the received wave can be sampled without error from ISI

The optimum sampling time corresponds to the maximum eye opening

The height of the eye opening at a specified sampling time is a measure of the margin over channel noise.

The sensitivity of the system to timing error is determined by the rate of closure of the eye as the sampling time is varied.

Any non linear transmission distortion would reveal itself in an asymmetric or squinted eye. When the effected of ISI is excessive, traces from the upper portion of the eye pattern cross traces from lower portion with the result that the eye is completely closed.