RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

Information Technology, IV-Semester

IT404 - Analog & Digital Comm.

Course Objectives

The study of communication systems starts with the concept of analog communication. In this course time and frequency representation of information is given. The objective of this course is to be familiar with the basic building blocks of communication systems such as modulator and demodulator. Different types of analog modulation techniques are given in this course.

Unit-I Signals and Systems: Block diagram of a communication system, signal-definition, types of signals continuous, discrete, deterministic, non-deterministic, periodic, non-periodic, energy, power, analog and digital signals. Electromagnetic Spectra, Standard signals- DC, sinusoidal, unit step, ramp, signum, rectangular pulse, impulse(delta) signal. System definition, classification of systems, linear, nonlinear, time variant, time invariant, causal, non causal, stable and unstable systems. Fourier transforms: Time domain and frequency domain representation of signal, Fourier Transform and its properties, conditions for existence, Transform of Gate, unit step, constant, impulse, sine and cosine wave. Shifting property of delta function, convolution, time and frequency convolution theorems.

Unit-II Amplitude modulation: Modulation, need of modulation, types of modulation techniques, amplitude modulation (DSB-FC), modulation index, frequency spectrum of AM wave, linear and over modulation, power relation in AM, transmission efficiency, modulation by a complex signal, bandwidth of AM, AM modulators, square law and switching modulator, advantages and disadvantages of AM. Demodulation of AM: Suppressed carrier amplitude modulation systems, DSB-SC, SSB-SC, VSB-SC systems, comparison of various amplitude modulation systems. Demodulation of AM, square law and envelope detector, synchronous detection of AM, Low and high power AM transmitters, AM receivers, TRF and superheterodyne receivers, sensitivity, selectivity and fidelity of receivers.

Unit-III Angle modulation: Introduction and types of angle modulation, frequency modulation, frequency deviation, modulation index, deviation ratio, bandwidth requirement of FM wave, types of FM. Phase modulation, difference between FM and PM, Direct and indirect method of FM generation, FM demodulators- slope detector, Foster seeley discriminator, ratio detector. Introduction to pulse modulation systems.

Unit-IV Sampling of signal, sampling theorem for low pass and Band pass signal, Pulse amplitude modulation (PAM), Time division, multiplexing (TDM). Channel Bandwidth for PAM-TDM signal Type of sampling instantaneous, Natural and flat top, Aperture effect, Introduction to pulse position and pulse duration modulations, Digital signal, Quantization, Quantization error, Pulse code modulation, signal to noise ratio, Companding, Data rate and Baud rate, Bit rate, multiplexed PCM signal, Differential PCM (DPCM), Delta Modulation (DM) and Adaptive Delta Modulation (ADM), comparison of various systems.

Unit-V Digital modulations techniques, Generation, detection, equation and Bandwidth of amplitude shift keying (ASK) Binary Phase Shift keying (BPSK), Differential phase shift keying (DPSK), offset and non offset quadrature phase shift keying (QPSK), M-Ary PSK, Binary frequency Shift Keying (BFSK), M-Ary FSK Quadrature Amplitude modulation (QAM).

Course Outcomes:
At the end of the course student will be able to :

1. Differentiate Analog and Digital Signal and types of signals.
2. Understand the communication of information over the communication channel.
3. Understand how information signal of low frequency can be transmitted with the help of modulation techniques over a long distance.
4. Differentiate different modulation techniques such as AM, SSB, DSB and FM.
5. Explain using block diagrams, modulation and demodulation techniques for digital signal and determine bandwidth requirement.

Reference Books:

1. Singh & Sapre, “Communication Systems”, TMH.
2. Taub Schilling, “Principles of Communication Systems”, TMH.
3. W. Tomasi “Electronic Communications Systems”, Pearson Education Pvt. Ltd.
4. Taub & shilling, “Communication Systems”, TMH.
5. Abhay Gandhi, “Analog and Digital Communication”, CENGAGE Learning.

List of Experiments:

1. AM Modulation and Demodulation (Envelope Detector)
2. Frequency modulation using reactance modulator.
3. Frequency modulation using varactor modulator.
4. Pulse Amplitude Modulation and Demodulation
5. Pre-emphasis and De-emphasis
6. Analog Multiplexing.
7. Amplitude Modulation using Pspice
8. Receiver characteristics (selectivity, sensitivity, fidelity).
9. Operation of foster-seeley loop detector.
10. Operation of ratio detector.

Unit-I: Signals and Systems: Block diagram of a communication system, signal-definition, types of signals continuous, discrete, deterministic, non-deterministic, periodic, non-periodic, energy, power, analog and digital signals. Electromagnetic Spectra, Standard signals- DC, sinusoidal, unit step, ramp, signum, rectangular pulse, impulse(delta) signal. System definition, classification of systems, linear, nonlinear, time variant, time invariant, causal, non causal, stable and unstable systems. Fourier transforms: Time domain and frequency domain representation of signal, Fourier Transform and its properties, conditions for existence, Transform of Gate, unit step, constant, impulse, sine and cosine wave. Shifting property of delta function, convolution, time and frequency convolution theorems.

1.1. Introduction to Communication System

Communication is a process whereby information is enclosed in a package and is channelled and imparted by a sender to a receiver via some medium. The receiver then decodes the message and gives the sender a feedback. So the basic elements of communication systems are:

- Transmitter: originates the signal
- Receiver: receives transmitted signal after it travels over the medium
- Medium: guides the signal from the transmitter to the receiver.

In a data transmission system, the transmission medium is the physical path between transmitter and receiver. For guided media, electromagnetic waves are guided along a solid medium, such as copper twisted pair, copper coaxial cable, and optical fibre. For unguided media, wireless transmission occurs through the atmosphere, outer space, or water.

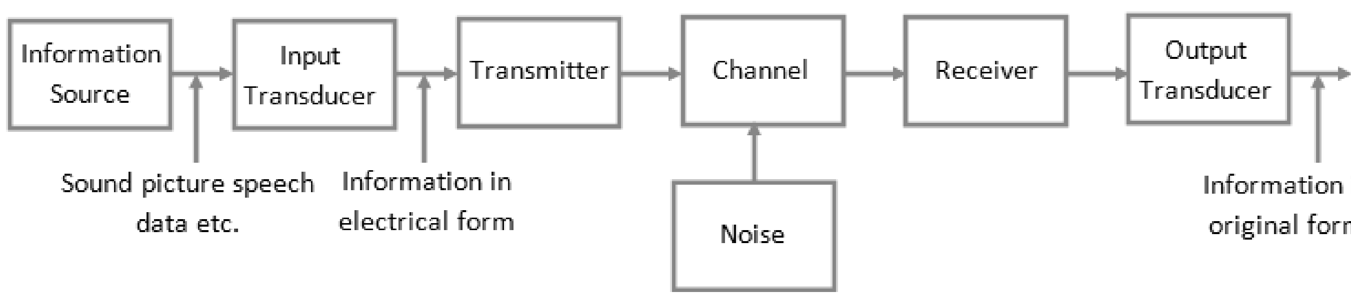


Figure1.1. Basic communication system

Elements of Communication system:

1. Information source: The message or information originates in the information source which has to be transmitted.

2. Input transducer: A transducer is the device which converts one form of energy to another form. In communication system transducer is usually required to convert the output of a source into an electrical signal that is suitable for transmission.

3. The Transmitter: The transmitter process the electrical signal into a form that is suitable for transmission through the transmission medium. The transmitter performs the signal processing of the message signal such as restriction of range of audio frequencies, amplification and modulation. All these processing are done to ease the transmission of the signal through the channel.

4. The Channel and the Noise: The communications channel is the physical medium that is used to send the signal from the transmitter to the receiver. In wireless transmission, the channel is usually the free space. On the other hand, telephone channels usually employ a variety of physical media, including wire lines, optical fibre cables, and wireless microwave radio. Whatever the physical medium for signal transmission, transmitted signal is corrupted in a random manner by noise.

5. The Receiver: The function of the receiver is to recover or reproduce the message signal contained in the received signal. If the message signal is transmitted by carrier modulation, the receiver performs *carrier demodulation* in order to extract the message from the sinusoidal carrier.

6. Output transducer: It is the final stage use to convert an electrical message signal into its original form.

1.2. Signal and its types:

A signal is a way of conveying information. Any time varying physical phenomenon that can convey information is called signal. Some examples of signals are human voice, electrocardiogram, sign language, videos etc. Technically - A signal is a function of one or more independent variables. a function of time, space, or another observation variable that conveys information

Signals are classified into the following categories:

- Continuous Time, Discrete Time Signals and Digital signals
- Deterministic and Non-deterministic Signals
- Even and Odd Signals
- Periodic and Aperiodic Signals
- Energy and Power Signals
- Real and Imaginary Signals

1. **Continuous Time Signal:** If the independent variable (t) is continuous, then the corresponding signal is continuous time signal. A finite, real-valued, smooth function x(t) of a variable *t* which usually represents time. Both *s* and *t* in X (t) are continuous. A continuous-time signal is a signal that can be defined at every instant of time. It is denoted by *x(t)*. Figure 1.2 shows continuous-time signal.

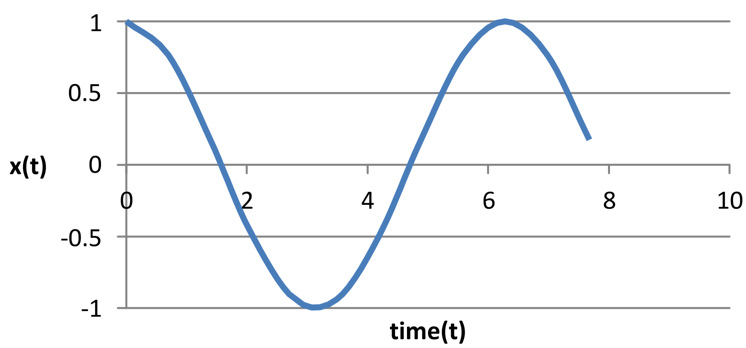


Figure 1.2.Continuous-time signal.

• Discrete Time Signal If the independent variable (t) takes on only discrete values, for example t = ±1, ±2, ±3, ... A *discrete-time* signal is a bounded, continuous-valued sequence x[n]. Alternately, it may be viewed as a continuous-valued function of a discrete index n. Discrete time signals can be obtained by sampling a continuous-time signal. It is denoted as *x(n)*.Figure 1.3 shows discrete-time signal.

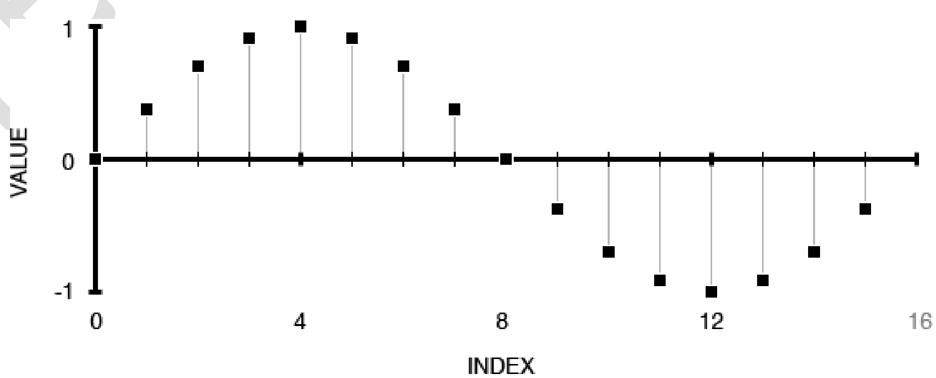


Figure 1.3.Discrete-time signal

Digital Signal: The signals that are discrete in time and quantized in amplitude are called digital signal. The term "digital signal" applies to the transmission of a sequence of values of a discrete-time signal in the form of some digits in the encoded form.

2. **Random (Non-deterministic) and Deterministic Signal:**

A random signal cannot be described by any mathematical function, where as a deterministic signal is one that can be described mathematically. A common example of random signal is noise. Random signal and deterministic signal are shown in the Figure 1.4 and 1.5 respectively. A signal is said to be non-deterministic if there is uncertainty with respect to its value at some instant of time. Non-deterministic signals are random in nature hence they are called random signals. Random signals cannot be described by a mathematical equation. They are modeled in probabilistic terms. A signal is said to be deterministic if there is no uncertainty with respect to its value at any instant of time.

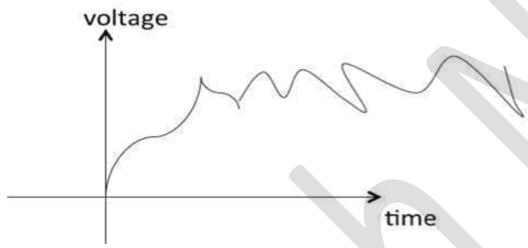


Figure 1.4.Random signal

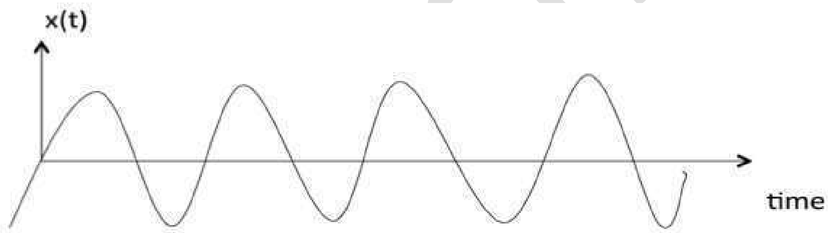


Figure 1.5.Deterministic signal

Even and Odd Signals

A signal is said to be even when it satisfies the condition x(t) = x(-t)

Example 1: t^2, t^4… cos t etc.

Let $x(t) = t^2$

$x(-t) = (-t)^2 = t^2 = x(t)$

t^2 is even function

Example 2: As shown in the following diagram, rectangle function x(t) = x(-t) so it is also even function.

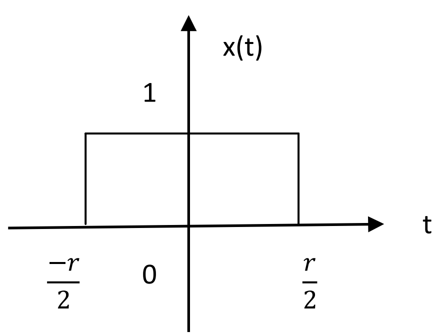


Figure 1.6.Even function

A signal is said to be odd when it satisfies the condition x(t) = -x(-t)

Example: t, t^3 ... And sin t

Let x(t) = sin t

x(-t) = sin(-t) = -sin t = -x(t)

∴, sin t is odd function.

Any function x(t) can be expressed as the sum of its even function $x_e(t)$ and odd function $x_o(t)$.

$x(t) = x_e(t) + x_0(t)$

where

$x_e(t) = ½[x(t) + x(-t)]$

$x_0(t) = ½[x(t) - x(-t)]$

A simple way of visualizing even and odd signal is to imagine that the ordinate *[x(t)]* axis is a mirror. For even signals, the part of *x(t)* for *t > 0* and the part of *x(t)* for *t < 0* are mirror images of each other. In case of an odd signal, the same two parts of the signals are negative mirror images of each other.

Periodic and Aperiodic Signals

A signal is said to be periodic if it satisfies the condition x(t) = x(t + T) or x(n) = x(n + N).

Where T = fundamental time period, 1/T = f = fundamental frequency.

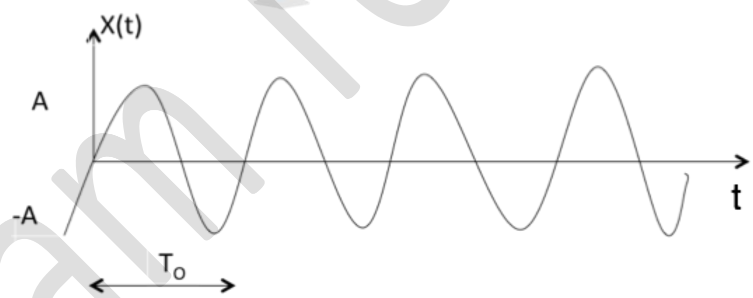


Figure 1.7.Periodic signal

The above signal will repeat for every time interval T_0 hence it is periodic with period T_0.

3. **Energy and Power Signals**

Energy and power for Continuous time signal

A signal is said to be energy signal when it has finite energy.

Energy E= $\int_{-\infty}^{\infty} |x(t)|^2 dt$

A signal is said to be power signal when it has finite power.

Power P = $\lim_{t\to\infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2\, dt$

NOTE: A signal cannot be both, energy and power simultaneously. Also, a signal may be neither energy nor power signal. If 0< E< ∞, then the signal x(t) is called an energy signal. However, there are signals where this condition is not satisfied. For such signals we consider the power. If 0< P< ∞, then the signal is called a power signal. Note that the power for an energy signal is zero (P = 0) and that the energy for a power signal is infinite (E = ∞). Some signals are neither energy nor power signals.

Power of energy signal = 0

Energy of power signal = ∞

Energy and power for discrete-time signal: The definition of signal energy and power for discrete signals is similar definitions for continuous signals.

Definition: The signal energy in the discrete-time signal x(n) is:

E = $\sum_{n=-\infty}^{\infty} |x(n)|^2$

The signal power in the signal x(n) is:

P = $\lim_{n\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x(n)|^2$

4. **Real and Imaginary Signals**

A signal is said to be real when it satisfies the condition x(t) = x*(t)
A signal is said to be imaginary when it satisfies the condition x(t) = -x*(t)
Example:
If x(t)= 3 then x*(t)=3*=3 here x(t) is a real signal.
If x(t)= 3j then x*(t)=3j* = -3j = -x(t) hence x(t) is a imaginary signal.
Note: For a real signal, imaginary part should be zero. Similarly for an imaginary signal, real part should be zero.

5. **Causal, Non-causal and Anti-causal Signal:**

Signal that are zero for all negative time, that type of signals are called causal signals, while the signals that are zero for all positive value of time are called anti-causal signal.

A non-causal signal is one that has non zero values in both positive and negative time. Causal, non-causal and anti-causal signals are shown below in the Figure 1.8(a), 1.8(b) and 1.8(c) respectively.

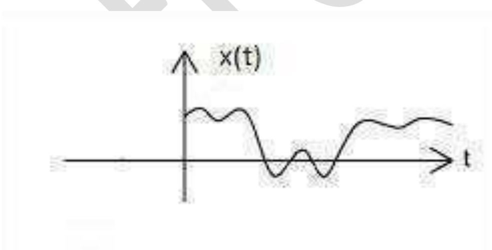


Fig.1.8(a) Causal signal

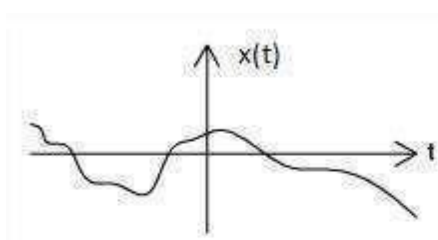


Fig.1.8(b) Non-causal signal

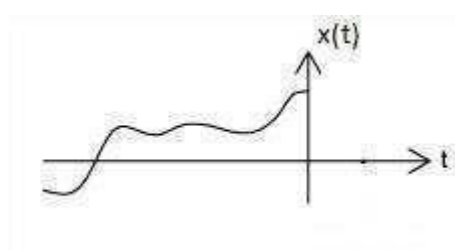


Fig.1.8(c) Anti-causal signal

Standard Signal:

1. **The DC Signal:** The "DC" or constant signal simply takes a constant value. In continuous time it would be represented as: s(t)=1. In discrete time, it would be s[n] =1. The number "1" may be replaced by any constant. The DC signal typically represents any constant offset from 0 in real-world signals. The analog DC signal has bounded amplitude and power and is smooth.
2. **Unit Step Function**
   The "unit step", also often referred to as a *Heaviside* function, is literally a step. It has 0 value until time 0, at which point, it abruptly switches to 1 from 0. The unit step represents events that change state, e.g. the switching on of a system, or of another signal. It is usually represented as u(t) in continuous time and u[n] in discrete time.

Unit step function is denoted by u(t). It is defined as

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Discrete time: $u[n] = \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases}$

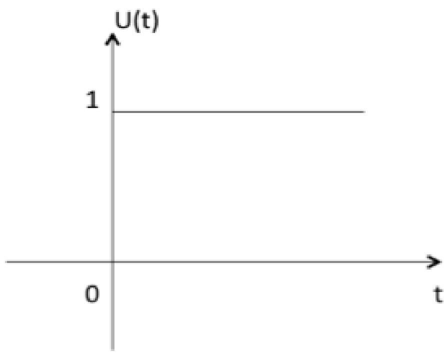


Figure 1.9(a) continuous time Unit step signal

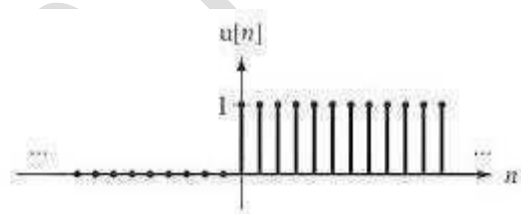


Figure 1.9(b) Discrete time Unit step signal

3. **Unit Impulse Function**

The unit impulse function also known as the Dirac Delta Function is denoted by δ(t) and it is defined as

$\int_{-\infty}^{\infty} \delta(t)\, dt = 1$ at t=0

$\delta(t) = 0$ t≠0

Consider a function

$$g(t) = \begin{cases} \frac{1}{W} & 0 < t < W \\ 0 & otherwise \end{cases}$$

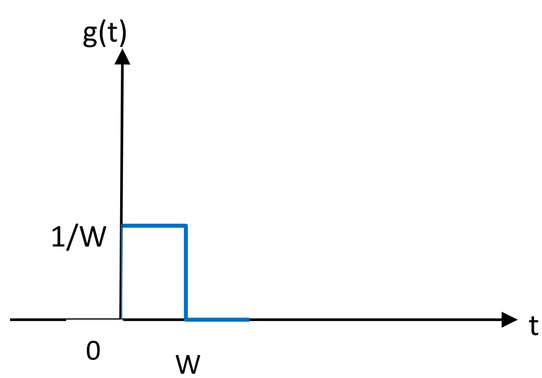


Figure 1.10.Gate function

One thing of note about g(t) is that

$\int_{0^-}^{W} g(t)\, dt = 1$:

The lower limit 0^- is a infinitesimally small amount less than zero. Now, suppose that the width w gets very small, indeed as small at 0^+, a number an infinitesimal amount bigger than zero. At that point, g(t) has become like the δ function, a very thin, very high spike at zero, such that

$\int_{-\infty}^{\infty} \delta(t)\,dt = \int_{0^-}^{0^+} \delta(t)\,dt = 1$

As w becomes very small the function g(t) turns into a function δ (t) indicated by the arrowed spike.

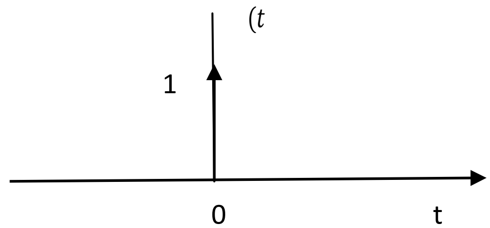


Figure 1.11.Unit Impulse signal

$\int_{-\infty}^{t} \delta(t)\,dt = u(t)$

$\delta(t) = \frac{du(t)}{dt}$

4. Ramp Signal

Ramp signal is denoted by r(t), and it is defined as r(t) = $\begin{cases} t & t \geqslant 0 \\ 0 & t < 0 \end{cases}$

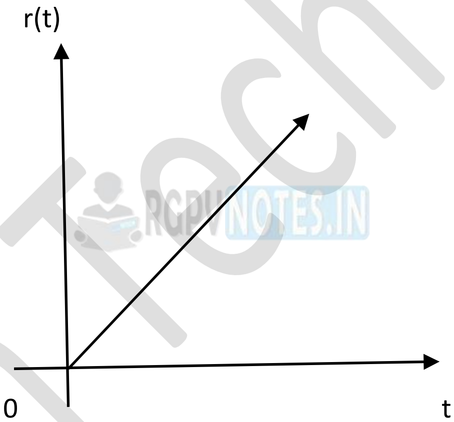


Figure 1.12.Ramp signal

∫u(t)=(t)

u(t)=dr(t)/dt

Area under unit ramp is unity.

5. Parabolic Signal

Parabolic signal can be defined as x(t) = $\begin{cases} \frac{t^2}{2} & t \geqslant 0 \\ 0 & t < 0 \end{cases}$

□u(t)dt=∫r(t)dt=∫tdt=t^2=parabolic signal

⇒u(t)=$d^2x(t)/dt^2$

⇒r(t)=dx(t)/dt

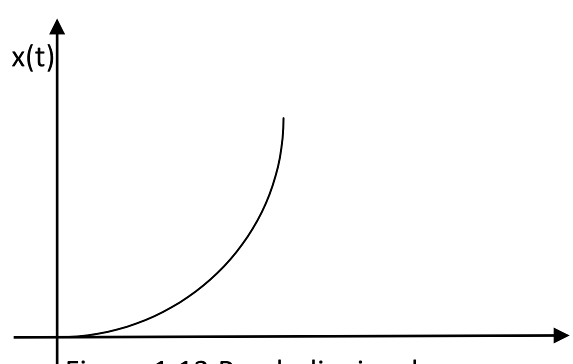


Figure 1.13.Parabolic signal

6. **Signum Function**

Signum function is denoted as sgn (t). It is defined as sgn (t) = $\begin{cases} 1 & t > 0 \\ 0 & t = 0 \\ -1 & t < 0 \end{cases}$

sgn(t) = 2u(t) – 1

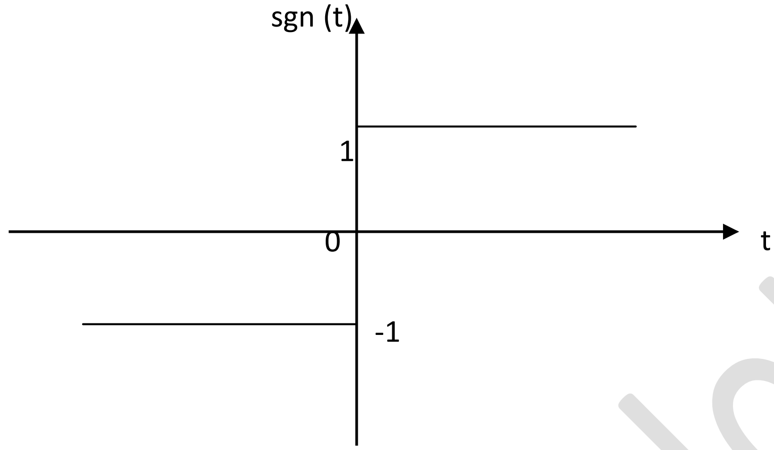


Figure 1.14.Signum signal

7. **Exponential Signal**

Exponential signal is in the form of x(t) = $e^{\alpha t}$.

The shape of exponential can be defined by α.

Case i: if α = 0 → x(t) = e^0 = 1

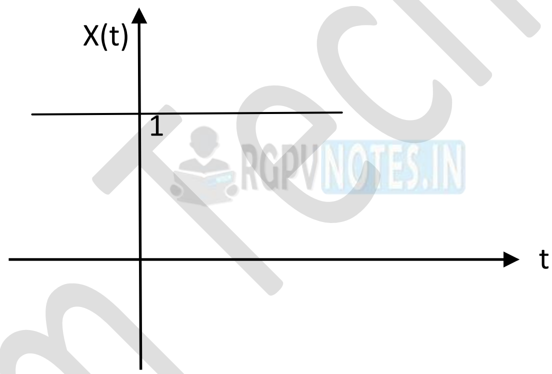


Figure 1.15(a) Exponential function at α=0

Case ii: if α < 0 i.e. -ve then x(t) = $e^{-\alpha t}$. The shape is called decaying exponential.

Case iii: if α > 0 i.e. +ve then x(t) = $e^{\alpha t}$. The shape is called raising exponential.

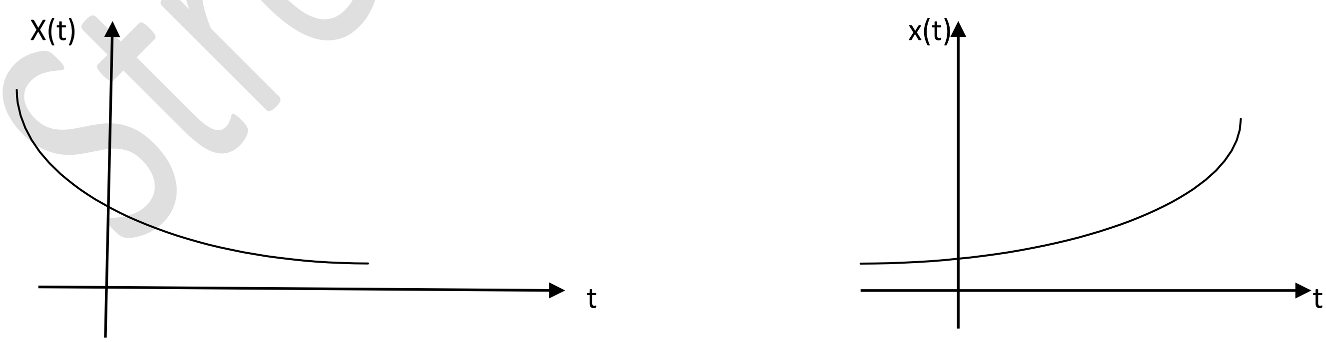


Figure 1.15(b).Exponential decaying and raising signals

8. Rectangular Signal

Let it be denoted as x(t) and it is defined as

$$x(t) = A \text{ rect } \left[\frac{r}{c}\right]$$

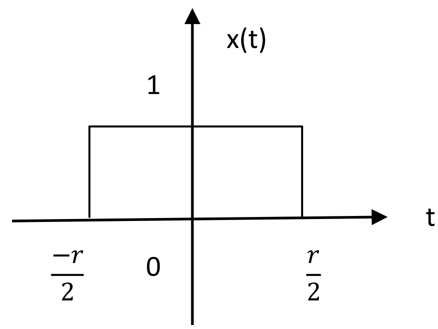


Figure 1.16.Rectangular function

9. Triangular Signal

Let it be denoted as x(t)= A $\left[1 - \frac{|t|}{T}\right]$

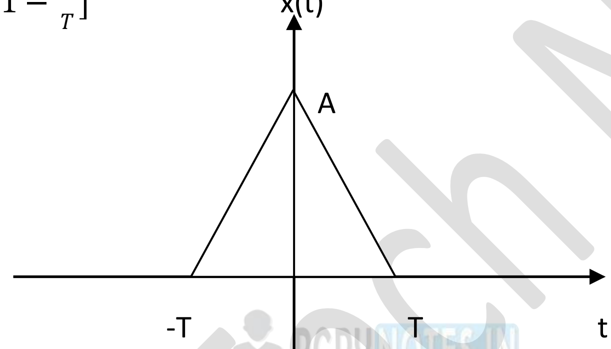


Figure 1.17.Triangular function

10. Sinusoidal Signal

Sinusoidal signal is in the form of x(t) = A cos ($w_0 \pm \phi$) or A sin($w_0 \pm \phi$)

Where $T_0 = 2\pi w_0$

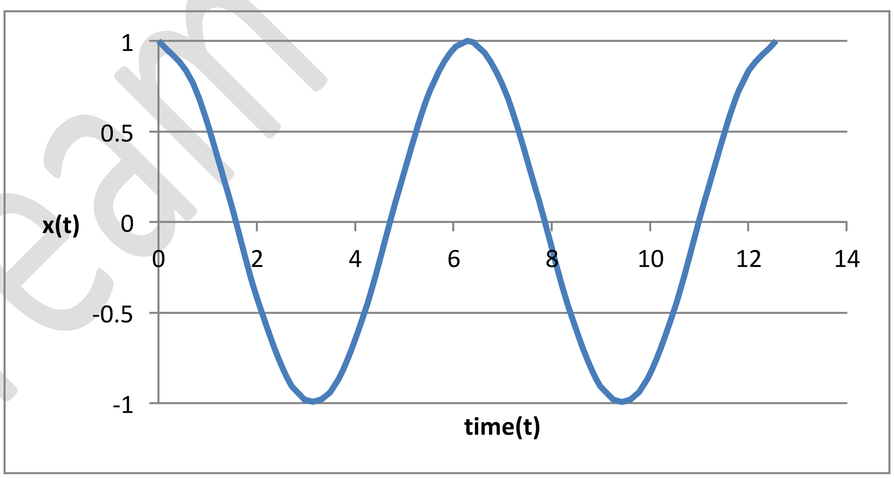


Figure 1.18.Cosinusoidal function

11. Sinc function:

It is denoted as sinc (t) and it is defined as

sinc(t)=sin (πt)/πt

=0 for t=±1, ±2 ,±3...

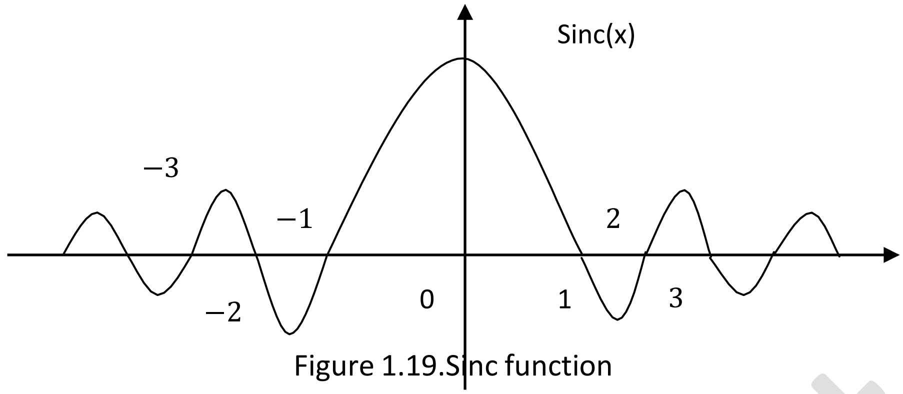


Figure 1.19.Sinc function

12. Sampling Function

It is denoted as sa(t) and it is defined as

Sa (t)=sin t/t

=0 for t=±π,±2π,±3π.....

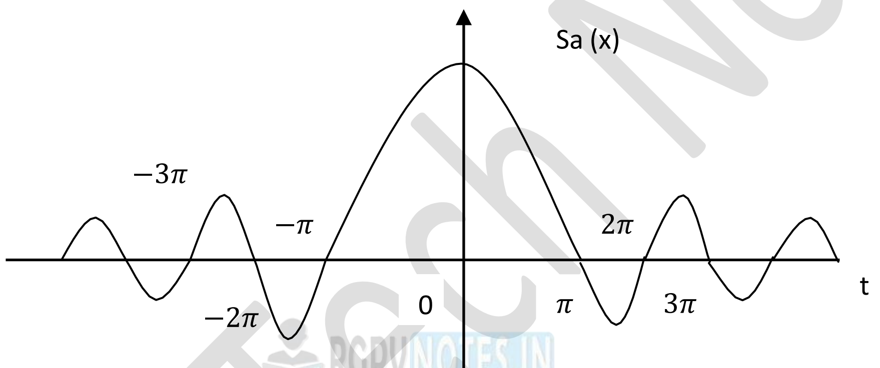


Figure 1.20.Sampling function

1.3. Electromagnetic Spectra: Electromagnetic spectrum ranges from dc to light. The lower radio frequencies are designated mainly by frequency. The optical ranges are referred by wavelength.

Signal parameters: Amplitude is the height of a wave. It is measured from a wave's midpoint to its peak. It is normally expressed in Volts (V). Frequency refers to the number of times a wave cycles past a given point each second. It is normally expressed in Hertz (Hz).Wave Length is the distance from the start to the end of a single wave cycle. It is typically expressed in meters.

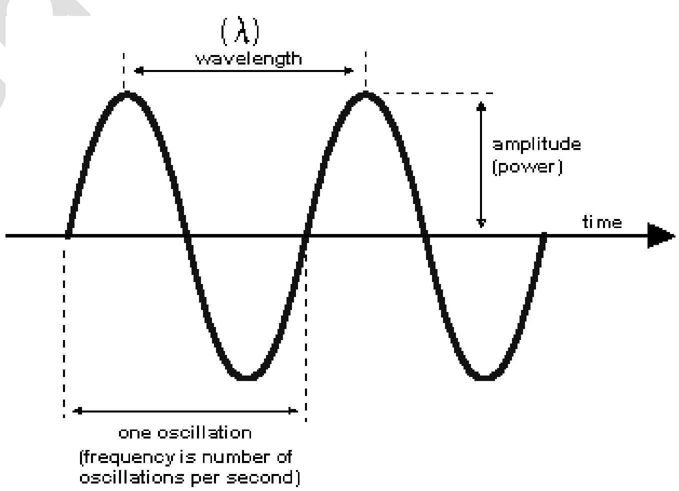


Figure 1.21.Signal Parameters

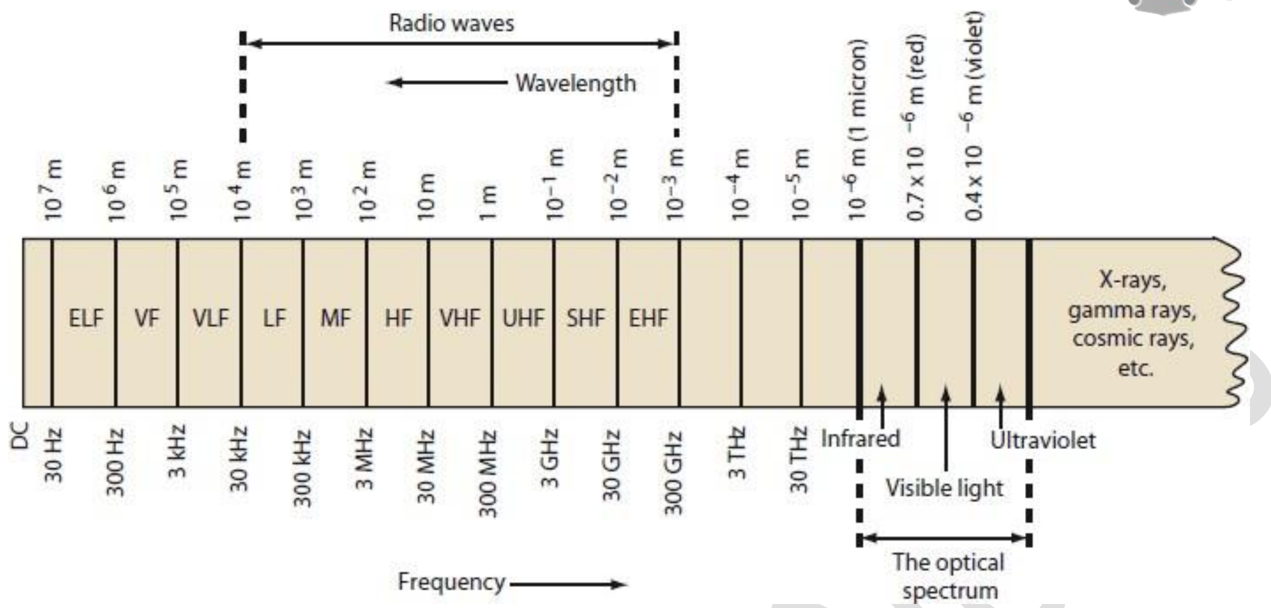


Figure 1.22.Electromagnetic spectra

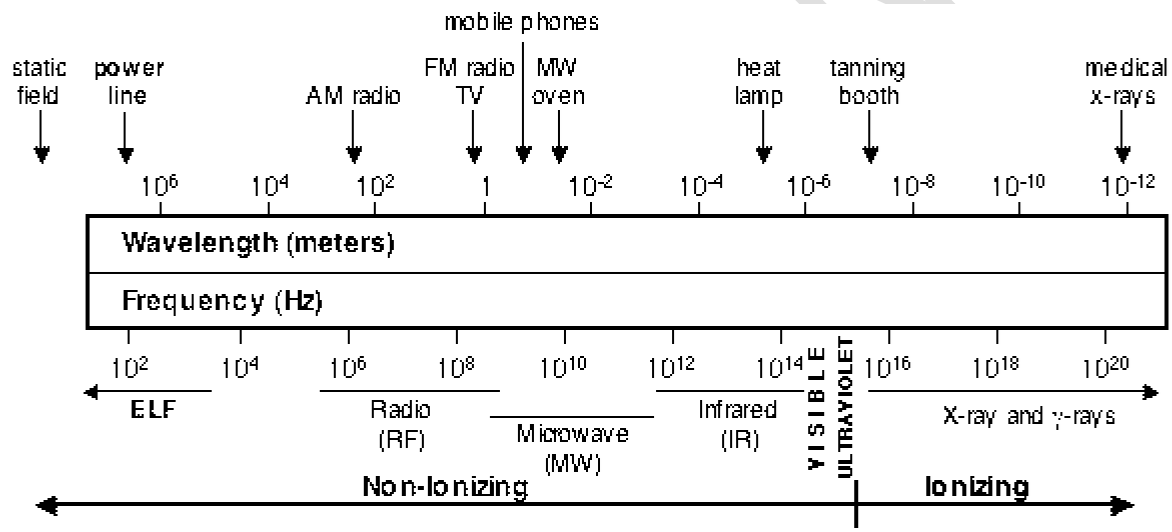


Figure 1.23.Electromagnetic Spectrum for communication systems

1.4. System definition and classification of systems:

System can be considered as a physical entity which manipulates one or more input signals applied to it. The system description specifies the transformation of the input signal to the output signal.

Systems are classified into the following categories:

- Linear and Non-linear Systems
- Time Variant and Time Invariant Systems
- Linear Time variant and Linear Time invariant systems
- Static and Dynamic Systems
- Causal and Non-causal Systems
- Stable and Unstable Systems

1. Liner and Non-liner Systems

A system is said to be linear when it satisfies superposition and homogeneity principles. Consider two systems with inputs as $x_1(t)$, $x_2(t)$, and outputs as $y_1(t)$, $y_2(t)$ respectively. Then, according to the superposition and homogenate principles,

Principle of homogeneity: $T[a_1*x_1(t)] = a_1*y_1$, $T[a_2*x_2(t)] = a_2*y_2$

Principle of superposition: $T[x_1(t)] + T[x_2(t)] = a_1*y_1+a_2*y_2$

Linearity: $T[a_1 x_1(t)] + T[a_2 x_2(t)] = a_1 y_1(t) + a_2 y_2(t)$

From the above expression, is clear that response of overall system is equal to response of individual system.

Example:

$y(t) = x^2(t)$

$y_1(t) = T[x_1(t)] = x_1^2(t)$

$y_2(t) = T[x_2(t)] = x_2^2(t)$

$T[a_1 x_1(t) + a_2 x_2(t)] = [a_1 x_1(t) + a_2 x_2(t)]^2$

Which is not equal to $a_1 y_1(t) + a_2 y_2(t)$. Hence the system is said to be non linear.

2. **Time Variant and Time Invariant Systems**

A system is said to be time variant if its input and output characteristics vary with time. If the system response to an input signal does not change with time such system is termed as time invariant system. The behavior and characteristics of time variant system are fixed over time.The condition for time invariant system is:

In time invariant systems if input is delayed by time t_0 the output will also gets delayed by t_0.

Mathematically it is specified as follows

$$\mathbf{y(t-t_0) = T[x(t-t_0)]}$$

For a discrete time invariant system the condition for time invariance can be formulated mathematically by replacing t as n*Ts is given as

$$\mathbf{y(n-n_0) = T[x(n-n_0)]}$$

3. **Linear Time variant (LTV) and Linear Time Invariant (LTI) Systems**

If a system is both linear and time variant, then it is called linear time variant (LTV) system.

If a system is both linear and time Invariant then that system is called liner time invariant (LTI) system.

4. **Static and Dynamic Systems**

Static system is memory-less whereas dynamic system is a memory system.

Example 1: $y(t) = 2x(t)$

For present value t=0, the system output is $y(0) = 2x(0)$. Here, the output is only dependent upon present input. Hence the system is memory less or static.

Example 2: $y(t) = 2x(t) + 3x(t-3)$

For present value t=0, the system output is $y(0) = 2x(0) + 3x(-3)$.

Here $x(-3)$ is past value for the present input for which the system requires memory to get this output. Hence, the system is a dynamic system.

5. **Causal and Non-Causal Systems**

A system is said to be causal if its output depends upon present and past inputs, and does not depend upon future input. For non causal system, the output depends upon future inputs also.

Example 1: y(n) = 2 x(t) + 3 x(t-3)

For present value t=1, the system output is y(1) = 2x(1) + 3x(-2).

Here, the system output only depends upon present and past inputs. Hence, the system is causal.

Example 2: y(n) = 2 x(t) + 3 x(t-3) + 6x(t + 3)

For present value t=1, the system output is y(1) = 2x(1) + 3x(-2) + 6x(4) Here, the system output depends upon future input. Hence the system is non-causal system.

6. **Stable and Unstable Systems**

The system is said to be stable only when the output is bounded for bounded input. For a bounded input, if the output is unbounded in the system then it is said to be unstable.

Note: For a bounded signal, amplitude is finite.

- **Example 1:** y (t) = x^2(t)

Let the input is u(t) (unit step bounded input) then the output y(t) = u2(t) = u(t) = bounded output.

Hence, the system is stable.

- **Example 2:** y (t) = $\int x(t)dt$

Let the input is u (t) (unit step bounded input) then the output y(t) = $\int u(t)dt$ = ramp signal (unbounded because amplitude of ramp is not finite it goes to infinite when t → infinite).

Hence, the system is unstable.

1.5. Time domain and frequency domain representation of signal

An electrical signal either, a voltage signal or a current signal can be represented in two forms: These two types of representations are as under:

i) Time Domain representation-: In time domain representation a signal is a time varying quantity as shown in Fig.1.24

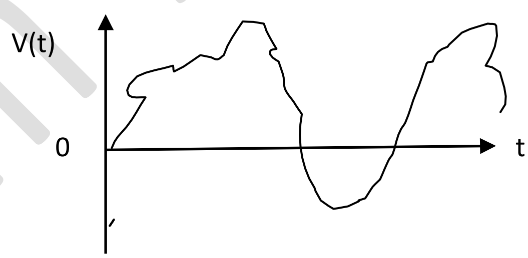


Fig 1.24 An arbitrary time domain signal

ii) Frequency Domain Representation: In frequency domain, a signal is represented by its frequency spectrum as shown in Fig 1.25

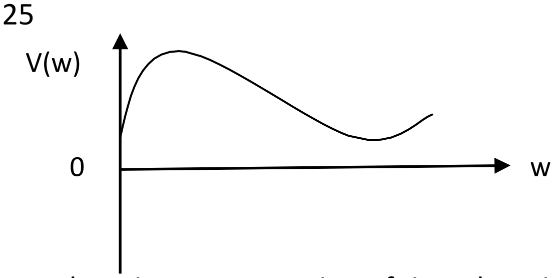


Fig 1.25 Frequency domain representation of time domain signal

1.6. Fourier Transform and its properties

Fourier Transform pair

Fourier transform may be expressed as

X(w)=F[x(t)]=$\int_{-\infty}^{\infty} x(t)\ e^{-jwt}dt$

In the above equation X(w) is called the Fourier transform of x(t). In other words X(w) is the frequency domain representation of time domain function x(t). This means that we are converting a time domain signal into its frequency domain representation with the help of fourier transform. Conversely if we want to convert frequency domain signal into corresponding time domain signal, we will have to take inverse fourier transform of frequency domain signal. Mathematically, Inverse fourier transform.

$$F^{-1}[X(w)] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(w)e^{jwt}dw$$

Example

Q.1 Find the fourier transform of a single-sided exponential function $e^{-at}u(t)$.

Solution: $e^{-at}u(t)$ is single sided function because her the main function e^{-at} is multiplied by unit step function u(t), then resulting signal will exist only for t>0.

u(t)= {1 for t>0

={0 for elsewhere

Now, given that x(t)= $e^{-at}u(t)$

X(w)=F[x(t)]=$\int_{-\infty}^{\infty} x(t)\ e^{-jwt}dt$

Or X(w)=$\int_{-\infty}^{\infty} e^{-at}\ u(t)\ e^{-jwt}dt$

$= \int_{0}^{\infty} e^{-at}\ e^{-jwt}dt$

$=\int_{0}^{\infty} e^{-t(a+jw)}dt$

$-\frac{-1}{(a+jw)}[e^{-\infty} - e\ ] = \ -\frac{-1}{(a+jw)}[0-1] = \frac{1}{(a+jw)}$

To obtain the above expression in the proper form we write

X(w)= $\frac{-1}{(a+jw)} * \frac{(a-jw)}{(a-jw)}$

X(w)= $\frac{(a-jw)}{(a^2+w^2)} = \frac{a}{(a^2+w^2)} - \frac{jw}{(a^2+w^2)}$

Obtaining the above expression in polar form

X(w)=$\frac{1}{\sqrt{a^2+w^2}}e^{-jtan^{-1}(\frac{\,}{a})}$

As we know that

X(w)=$|X(w)|e^{j\varphi(jw)}$

On comparision amplitude spectrum

$|X(w)|=\frac{1}{\sqrt{a^2+w^2}}$

$$\varphi(w) = -tan^{-1}(\frac{w}{a})$$

Properties of Continuous Time Fourier Transform (CTFT)

1. **Time Scaling Function**

Time scaling property states that the time compression of a signal results in its spectrum expansion and time expansion of the signal results in its spectral compression. Mathematically,

If $x(t) \longleftrightarrow X(w)$

Then, for any real constant a,

$x(at) \longleftrightarrow \frac{1}{|a|}X(\frac{w}{a})$

proof: The general expression for fourier transform is

$X(w)=F[x(t)]=\int_{-\infty}^{\infty} x(t)\ e^{-jwt}dt$

Now $F[x(at)]=\int_{-\infty}^{\infty} x(at)\ e^{-jwt}dt$

Putting

At=y

We have $dt=\frac{dy}{a}$

Case (i): When a is positive real constant

$F[x(at)]= \int_{-\infty}^{\infty} x(y)\, e^{-j(\frac{w}{a})y} \frac{dy}{a} = \frac{1}{a}\int_{-\infty}^{\infty} x(y)\ e^{-j(\frac{w}{a})y} = \frac{1}{a}X(\frac{w}{a})$

Case (ii): When a is negative real constant

$F[x(at)]= \frac{-1}{a}X(\frac{w}{a})$

Combining two cases, we have

$F[x(at)]= \frac{1}{|a|}X(\frac{w}{a})$ Or $x(at) \longleftrightarrow \frac{1}{|a|}X(\frac{w}{a})$

The function x(at) represents the function x(t) compressed in time domain by a factor a. Similarly, a function $X(\frac{w}{a})$ represents the function X(w) expanded in frequency domain by the same factor a.

2. **Linearity Property**

Linearity property states that fourier transform is linear. This means that

If $x_1(t) \longleftrightarrow X_1(w)$

And $x_2(t) \longleftrightarrow X_2(w)$

Then $a_1 x_1(t) + a_2 x_2(t) \longleftrightarrow a_1X_1(w) + a_2X_2(w)$

3. **Duality or Symmetry Property**

If $x(t) \longleftrightarrow X(w)$

Then $X(t) \longleftrightarrow 2\pi x(-w)$

Proof

The general expression for fourier transform is

$$F^{-1}[X(w)] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(w)e^{jwt}dw$$

Therefore,

$x(-t)= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(w)e^{-jwt}dw$

$2\pi x(-t)= \int_{-\infty}^{\infty} X(w)e^{-jwt}dw$

Since w is a dummy variable, interchanging the variable t and w we have

$2\pi x(-w)= \int_{-\infty}^{\infty} X(t)e^{-jwt}dw=F[X(t)]$

Or $F[X(t)]= 2\pi x(-w)$

Or X(t) $\longleftrightarrow$ 2πx(-w)
For an even function x(-w)=x(w)

Therefore , X(t) $\longleftrightarrow$ 2πx(w)

Example (1)
The fourier transform F[$e^{-t}u(t)$] is equal to $\frac{1}{1+\mathbf{j}2\pi\mathbf{f}}$. Therefore F[$\frac{1}{1+\mathbf{j}2\pi\mathbf{f}}$] is equal to

Solution:
Using Duality property of Fourier Transform, we have

If x(t) $\longleftrightarrow$ X(f)
Then X(t) $\longleftrightarrow$ x(-f)
Therefore,
$e^{-t}u(t) \longleftrightarrow \frac{1}{1+\mathbf{j}2\pi\mathbf{f}}$
Then $\frac{1}{1+j2\pi t} \longleftrightarrow e^{-\mathbf{f}}u(\mathbf{f})$

4. **Time Shifting property**

Time Shifting property states that a shift in the time domain by an amount b is equivalent to multiplication by e^{-jwb} in the frequency domain. This means that magnitude spectrum $|X(w)|$

Remains inchanged but phase spectrum θ(w) is changed by -wb.

If x(t) $\longleftrightarrow$ X(w)
Then X(t-b) $\longleftrightarrow$ X(w) e^{-jwb}
Proof: X(w)=F[x(t)]=$\int_{-\infty}^{\infty} x(t)\ e^{-jwt}dt$
And F[x(t-b)]= $\int_{-\infty}^{\infty} x(t-b)\ e^{-jwt}dt$
Putting t-b=y, so that dt=dy

F[x(t-b)]= $\int_{-\infty}^{\infty} x(y)\ e^{-jw(b+y)}dy = \int_{-\infty}^{\infty} x(y)\ e^{-jwb}e^{-jwy}\ dy$

Or F[x(t-b)]= $e^{-jwb} \int_{-\infty}^{\infty} x(y)\ e^{-jwy}\ dy$
Since y is a dummy variable, we have
F[x(t-b)]= e^{-jwb}X(w)=X(w) e^{-jwb}
Or x(t-b) $\longleftrightarrow$ X(w) e^{-jwb}

5. **Frequency Shifting Property**

Frequency shifting property states that the multiplication of function x(t) by e^{jw_0t} is equivalent to shifting its fourier transform X(w) in the positive direction by an amount w_0 . This means that the spectrum X(w) is translated by an amount c. hence this property is often called frequency translated theorem. Mathematically .

If x(t) $\longleftrightarrow$ X(w)
Then e^{jw_0t} x(t) $\longleftrightarrow$ X(w-w_0)
Proof: General expression for fourier transform is

$X(w)=F[x(t)]=\int_{-\infty}^{\infty} x(t)\ e^{-jwt}dt$

Now, $F[e^{jw_0t}x(t)] = \int_{-\infty}^{\infty} x(t)\ e^{jw_0t}e^{-jwt}dt$

Or $F[e^{jw_0t}x(t)] = \int_{-\infty}^{\infty} x(t)\ e^{-j(w-w_0)t}\,dt$

Or $F[e^{jw_0t}x(t)] = X(w-w_0)$

Or $e^{jw_0t}\ x(t) \longleftrightarrow X(w-w_0)$

6. **Time Differentiation Property**

The time differentiation property states that the differentiation of a function x(t) in the time domain is equivalent to multiplication of its fourier transform by a factor jw. Mathematically

If $x(t) \longleftrightarrow X(w)$

Then $\frac{dx(t)}{dt}$ x(t) $\longleftrightarrow$ jw X(w)

Proof: The general expression for fourier transform is

$$F^{-1}[X(w)] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(w)e^{jwt}dw$$

Taking differentiation, we have

$$\frac{dx(t)}{dt} = \frac{1}{2\pi}\frac{d}{dt}\left[\int_{-\infty}^{\infty} X(w)e^{jwt}dw\right]$$

Interchanging the order of differentiation and integration, we have

$$\frac{dx(t)}{dt} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d}{dt}[X(w)e^{jwt}]dw$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} jw\,X(w)e^{jwt}dw$$

Or $\frac{dx(t)}{dt} = F^{-1}[jwX(w)]$

Or $F[\frac{dx(t)}{dt}] = jwX(w)$

Or $\frac{dx(t)}{dt} \longleftrightarrow jwX(w)$ Hence proved

1.7 Transform of Gate

A gate function is rectangular pulse. Figure 1.3 shows gate function. The function or rectangular pulse shown in figure 1.3 is written as rect $(\frac{t}{c})$.

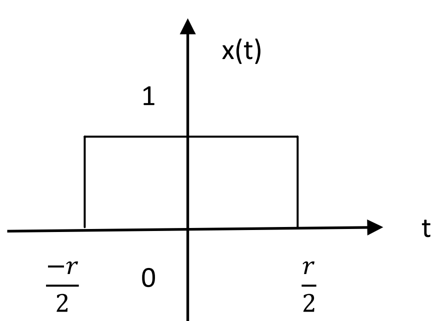


Fig2.3 A Gate Function

From the above figure it is clear that rect $(\frac{t}{c})$ represents a gate pulse of height or amplitude unity and width r.

$$x(t)= \text{rect}\left(\frac{t}{c}\right) = \begin{cases} 1 & \text{for } \frac{-c}{2} < t < \frac{c}{2} \\ 0 & \text{otherwise} \end{cases}$$

Sampling Function Or Interpolation Function Or Sinc function

The functions $\frac{sinx}{x}$ is the "sine over argument" and denoted by sinc(x). This function plays an important role in signal processing. It is also known as the filtering or interpolating function. Mathematically,

Sinc(x)=$\frac{sinx}{x}$

Or

Sa(x)= $\frac{sinx}{x}$

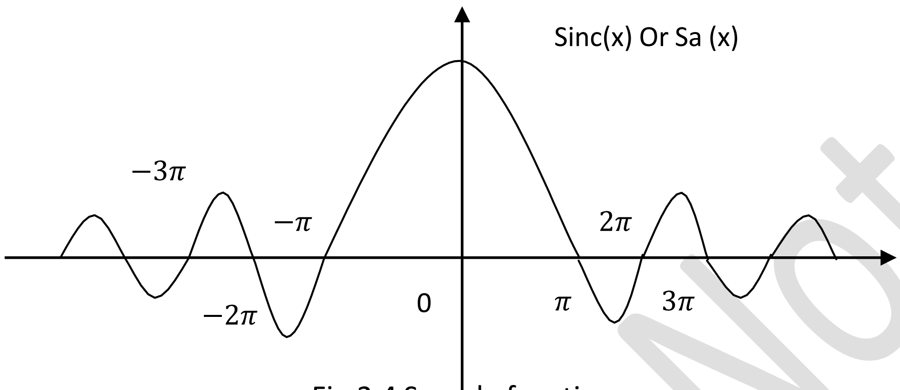


Fig.2.4 Sample function

From the figure, following points may be observed about the sampling function :

(i) Sa(x) or sinc(x) is an even function of x.

(ii) Sinc(x) =0 when sinx=0 except at x=0, where it is indeterminate. This means that sinc(x)=0 for x=$\pm$nπ , here n=$\pm 1, \pm 2$ ....

(iii) Sinc(x) is the product of oscillating signal sinx of period 2π and a decreasing function $\frac{1}{x}$ Therefore, sinc(x) exhibits sinusoidal oscillations of period 2π with amplitude decreasing continuously as 1/x.

Example 2: Find the fourier transform of the gate function shown in figure 1.5.

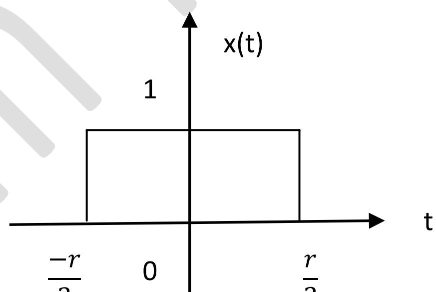


Fig.2.5 Gate function

Sol. x(t)= rect $(\frac{t}{c})$ =$\{1 \quad for\ \frac{-c}{2} < t < \frac{c}{2}\}$

$\{0 \quad otherwise\}$

X(w)=F[x(t)]=$\int_{-\infty}^{\infty} x(t)\, e^{-jwt}dt$

X(w)=F[x(t)]=$\int_{-\infty}^{\infty} \text{rect}\,(\frac{t}{c})\ e^{-jwt}\,dt$

$=\int_{\frac{-r}{2}}^{\frac{r}{2}} 1.e^{-jwt}dt= \left[\frac{-e^{-}}{jw}\right]^{2}_{\frac{-r}{2}}{}^{\frac{r}{2}}$

$=\frac{-1}{jw}[e^{-j\frac{r}{2}}-e^{j\frac{r}{2}}]$

$=\frac{1}{jw}[e^{j\frac{r}{2}}-e^{-j\frac{r}{2}}]$ --------(1)

We know that e^{j}=cos +jsin And

e^{-j}=cos –jsin

Hence 2cos=$e^{j}+e^{-j}$

2jsin=$e^{j}-e^{-j}$

Putting = $\frac{wc}{2}$, we get

$2j\sin\frac{wc}{2}=e^{j\frac{r}{2}}-e^{-j\frac{r}{2}}$ --------(2)

From (1) and (2)

$X(w)=\frac{1}{jw}[2j\sin\frac{wc}{2}]$

By multiplying and dividing the equation by r

$=\frac{2c}{jwc}[j\sin\frac{wc}{2}]$

$=\frac{c}{\frac{r}{2}}[\sin\ \frac{wc}{2}]$

$=r\ [\frac{\sin\frac{r}{2}}{\frac{r}{2}}]$

$=r\,\mathrm{sinc}(\frac{wc}{2})$

Now, since sinc(x)=0, when x=$\pm$nπ

Therefore, sinc($\frac{wc}{2}$) =0, when $\frac{wc}{2}=\pm n\pi$

Or $w=\frac{=\pm 2n\pi}{c}$

Figure 2.6 shows the plot of X(w) $\frac{2\pi}{c}$

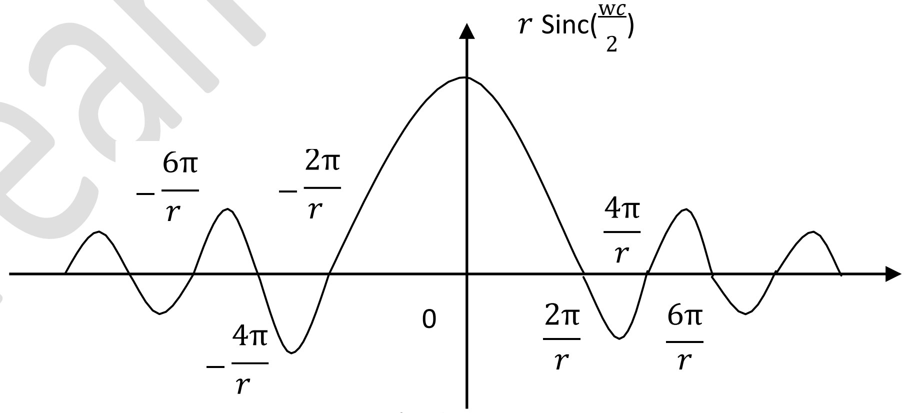


Fig. 2.6

1.8 Impulse Functions

Unit Impulse functions:

A unit impulse function was invented by P.A.M. Diarc and so it is also called as Delta function. It is denoted by $(t\ )$.

Mathematically,

(t)=0 , t$\neq$ 0

And, $\int_{-\infty}^{\infty}(t)$dt=1

Figure 1.6 shows the graphical representation of an unit impulse function. The following points may be observed about an unit-impulse function:

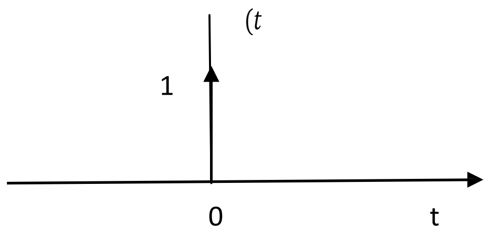


Fig.2.7 The Unit Impulse function

i) The width of pulse is zero. This means that pulse exist only at t=0.
ii) The height of the pulse goes to infinity
iii) The area under the pulse-curve is always is always unity.

Shifting Property of the Impulse function:

If we take the product of unit impulse function (t) and any given function x(t) which is continuous at t=0,thenthis product will provide the function x(t) existing only at t=0 since (t) exist only at t=0. Mathematically,

$\int_{-\infty}^{\infty}x(t)(t)$dt=x(0) $\int_{-\infty}^{\infty}(t)$dt=x(0).1=x(0)

The equation is also known as shifting or sampling property of the impulse function because the impulse shifts the value of x(t0 at t=0. This means that the value of x(t) has been sampled at t=0. to The shifting or sampling may be also done at any, instant t=t_0, if we define the impulse function at the instant. Mathematically,

$\int_{-\infty}^{\infty}x(t)(t-t_0)$dt=x($t_0$)

The above equation states that the product of a continuous function x(t) with an impulse function $(t-t_0)$ provides the sampled value of x(t) at t=t_0.

Q.1 Find the fourier transform of an impulse function x(t) = (t) Also draw the spectrumSol.

Expression of the fourier transform is given by

X(w)=F[x(t)]=$\int_{-\infty}^{\infty}x(t)\ e^{-jwt}$dt=$\int_{-\infty}^{\infty}(t)\ e^{-jwt}$dt

Using shifting property of impulse function

X(w)=[$e^{-jwt}]_{\text{at t=0}}$

X(w)=1

(t) $\longleftrightarrow$ 1

Hence the fourier transform of an impulse function is unity.

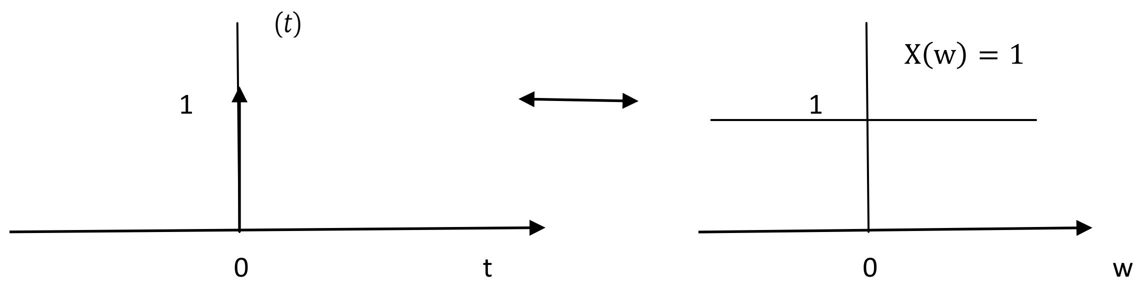


Fig 2.8 Time and frequency domain of impulse function

Figure 2.8 shows an unit impulse function and its fourier transform or spectrum. From the figure1.7 it is clear that an unit impulse contains the entire frequency components having identical magnitude. This means that the bandwidth of the unit impulse function is infinite. Also, since spectrum is real, only magnitude spectrum is required. The phase spectrum (w)=0, which means that all the frequency components are in the same phase.

Q.(2) Find the inverse fourier transform of (w) Solution.
Inverse fourier transform is expressed as

$$F^{-1}[\mathrm{X}(\mathrm{w})] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{X}(\mathrm{w})\mathrm{e}^{\mathrm{j}\mathrm{w}t}d\mathrm{w}$$

$$F^{-1}[(\mathrm{w})] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} (\mathrm{w})\mathrm{e}^{\mathrm{j}\mathrm{w}t}d\mathrm{w}$$

$F^{-1}[(\mathrm{w})] = \frac{1}{2\pi}[\mathrm{e}^{\mathrm{j}\mathrm{w}t}]_{\text{at w=0}}$

$F^{-1}[(\mathrm{w})] = \frac{1}{2\pi}[\mathrm{e}^{0}] = \frac{1}{2\pi}1 = \frac{1}{2\pi}$

$F[\frac{1}{2\pi}] = (\mathrm{w})$

$\frac{1}{2\pi} \longleftrightarrow (\mathrm{w})$

$1 \longleftrightarrow 2\pi(\mathrm{w})$

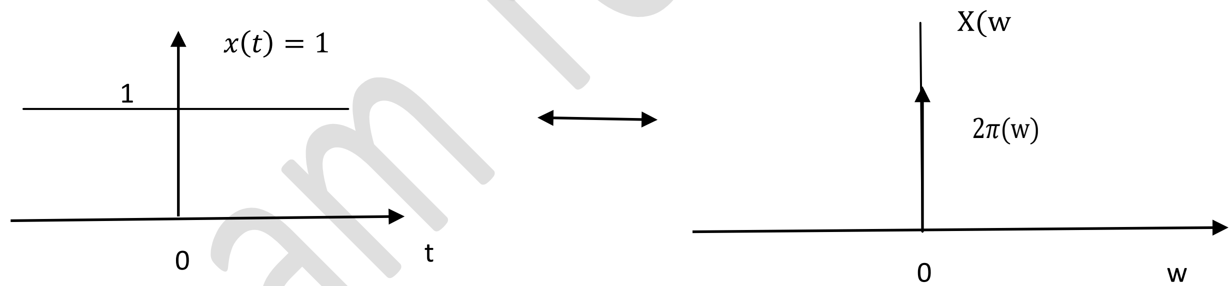


Fig.2.9 representation of Inverse fourier transform

This shows that the spectrum of a constant signal x(t)=1 an impulse function $2\pi(\mathrm{w})$. This can also be interpreted as that x(t) =1 is a d.c. signal which has single frequency. W=0(dc).

Q.(3) Find the inverse fourier transform of (w-w_0)
Solution. Inverse fourier transform is expressed as

$$F^{-1}[\mathrm{X}(\mathrm{w})] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{X}(\mathrm{w})\mathrm{e}^{\mathrm{j}\mathrm{w}t}d\mathrm{w}$$

Or

$$F^{-1}[(\mathrm{w}-\mathrm{w}_0)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} (\mathrm{w}-\mathrm{w}_0)\mathrm{e}^{\mathrm{j}\mathrm{w}t}d\mathrm{w}$$

Using shifting or sampling property of impulse function, we get

$F^{-1}[(\mathrm{w}-\mathrm{w}_0)] = \frac{1}{2\pi}[\mathrm{e}^{\mathrm{j}\mathrm{w}t}]_{\text{at } \mathrm{w}=\mathrm{w}_0}$

$F^{-1}[(w - w_0)] = \frac{1}{2\pi}[e^{jw_0t}]$

$F[\frac{1}{2\pi} e^{jw_0t}] = (w - w_0)$

$\frac{1}{2\pi} e^{jw_0t} \leftrightarrow (w - w_0)$

Or

$e^{jw_0t} \leftrightarrow 2\pi(w - w_0)$

The above expression shows that the spectrum of an overlasting exponential e^{jw_0t} is a single impulse at w=0. Similarly,

$e^{-jw_0t} \leftrightarrow 2\pi(w + w_0)$

1.9. Fourier Transform of Cosine wave

Q.4 Find the fourier transform of overlasting sinusoid $\cos w_0 t$.

Solution: We know that Euler's identity is given by

$e^{j}=\cos + j\sin$

And $e^{-j}=\cos - j\sin$

Hence $2\cos = e^{j}+e^{-j}$ Or

$\cos = \frac{e + e^{-}}{2}$

And

$2j\sin = e^{j} - e^{-j}$ Or

$\sin = \frac{e + e^{-}}{2j}$

Hence, $\cos w_0 t = \frac{e^{0} + e^{-0}}{2}$

We know that

$e^{jw_0t} \leftrightarrow 2\pi(w - w_0)$

And $e^{-jw_0t} \leftrightarrow 2\pi(w + w_0)$

So that $\cos w_0 t$ Or $\leftrightarrow \frac{1}{2}[2\pi(w - w_0) + 2\pi(w + w_0)]$

$\cos w_0 t \leftrightarrow [\pi(w - w_0) + \pi(w + w_0)]$

1.10. Fourier Transform of Periodic Function

Fourier transform of periodic function could also be found out. This means that Fourier transform may be used as a universal mathematical tool to analyze both periodic and non-periodic waveform over the entire interval. Let us find the fourier transform of periodic function x(t). x(t) may be expressed in terms of complex fourier series as

$x(t)=\sum_{n=-\infty}^{\infty} C_n e^{jnw_0t}$

Taking fourier transform of both the side

$F[x(t)]=F[\sum_{n=-\infty}^{\infty} C_n e^{jnw_0t}]$ $\sum_{n=-\infty}^{\infty} C_n . F[1. e^{jnw_0t}]$

Using frequency shifting shifting theorem, we can write

$F[1. e^{jnw_0t}]=2\pi\ (w - nw_0)$

Hence, $F[x(t)] = \sum_{n=-\infty}^{\infty} C_n 2\pi (w - nw_0) = 2\pi \sum_{n=-\infty}^{\infty} C_n (w - nw_0)$

Hence, the fourier transform of a periodic function consist of a train of equally spaced impulses. These impulses are located at the harmonic frequencies of the signal and the strength or area of each impulse is given by $2\pi C_n$.

1.11. Convolution

Convolution is a mathematical operation used to express the relation between input and output of an LTI system. It relates input, output and impulse response of an LTI system as

y(t)=x(t)∗h(t)

Where y (t) = output of LTI

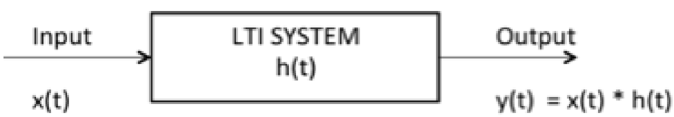
x (t) = input of LTI

h (t) = impulse response of LTI

There are two types of convolutions:

- Continuous convolution
- Discrete convolution

Continuous Convolution



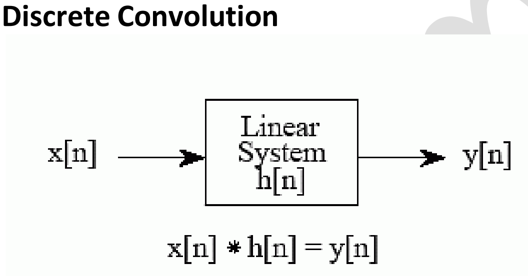
y(t)=x(t)∗h(t)

$$= \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$$

$$= \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau$$

A convolution is a mathematical operation that represents a signal passing through a LTI (Linear and Time-Invariant) system or filter.

Discrete Convolution



y(n)=x(n)∗h(n)

$$= \sum_{k=-\infty}^{\infty} x(k)h(n-k)$$

$$= \sum_{k=-\infty}^{\infty} x(n-k)h(k)$$

Delta function, symbolized by the Greek letter delta, δ[n] is a normalized impulse, that is, sample number zero has a value of one, while all other samples have a value of zero. For this reason, the delta function is called the unit impulse.

Impulse response is the signal that exits a system when a delta function (unit impulse) is the input. If two systems are different in any way, they will have different impulse responses. The input and output signals are often called x[n] and y[n], the impulse response is usually given the symbol, h[n]. Any impulse can be represented as a shifted and scaled delta function.

Properties of Convolution

1. Commutative Property: $x_1(t)*x_2(t)=x_2(t)*x_1(t)$
2. Distributive Property: $x_1(t)*[x_2(t)+x_3(t)]=[x_1(t)*x_2(t)]+[x_1(t)*x_3(t)]$
3. Associative Property: $x_1(t)*[x_2(t)*x_3(t)]=[x_1(t)*x_2(t)]*x_3(t)$
4. Shifting Property: $x_1(t)*x_2(t)=y(t)$
   $x_1(t)*x_2(t-t_0)=y(t-t_0)$
   $x_1(t-t_0)*x_2(t)=y(t-t_0)$
   $x_1(t-t_0)*x_2(t-t_1)=y(t-t_0-t_1)$
5. Convolution with Impulse: $x_1(t)*\delta(t)=x(t)$
   $x_1(t)*\delta(t-t_0)=x(t-t_0)$
6. Convolution of Unit Steps: $u(t)*u(t)=r(t)$
   $u(t-T_1)*u(t-T_2)=r(t-T_1-T_2)$
   $u(n)*u(n)=[n+1]u(n)$
7. Scaling Property: If $x(t)*h(t)=y(t)$
   then $x(at)*h(at)=\frac{1}{|a|}y(at)$
8. Differentiation of Output: If $y(t)=x(t)*h(t)$
   then $\frac{dy(t)}{dt}=\frac{dx(t)}{dt}*h(t)$

or

$$\frac{dy(t)}{dt}=x(t)*\frac{dh(t)}{dt}$$

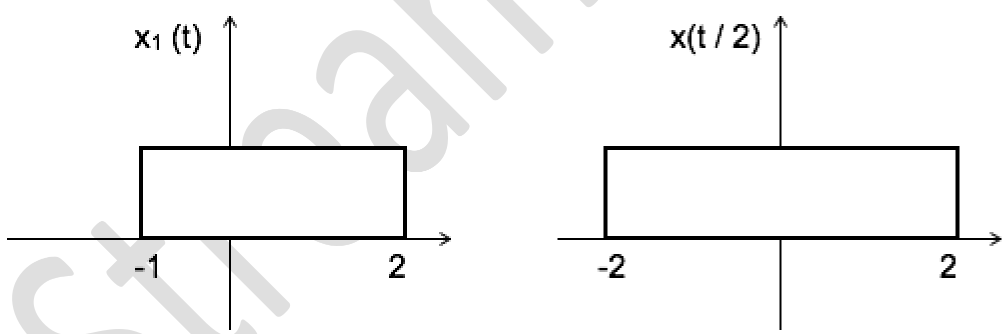
Note:

- Convolution of two causal sequences is causal.
- Convolution of two anti causal sequences is anti causal.
- Convolution of two unequal length rectangles results a trapezium.
- Convolution of two equal length rectangles results a triangle.
- A function convoluted itself is equal to integration of that function.

Limits of Convoluted Signal

If two signals are convoluted then the resulting convoluted signal has following range:

Sum of lower limits < t < sum of upper limits



Here, we have two rectangles of unequal length to convolute, which results a trapezium.

The range of convoluted signal is:

Sum of lower limits < t < sum of upper limits

$-1+-2<t<2+2$

$-3<t<4$

Hence the result is trapezium with period 7.

Area of Convoluted Signal

The area under convolved signal is given by $A_y=A_xA_h$

Where A_x = area under input signal
A_h = area under impulse response
A_y = area under output signal
Proof: y(t)= $= \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$

Take integration on both sides

$$\int y(t)dt = \int \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \;\; dt$$

$$= \int x(\tau)d\tau \int_{-\infty}^{\infty} h(t-\tau)dt$$

We know that area of any signal is the integration of that signal itself.
$\therefore A_y = A_x A_h$

DC Component

DC component of any signal is given by
DC component=area of the signal/period of the signal
NOTE:
1. The temporal output is the temporal input CONVOLVED with the Impulse Response Function.
2. The frequency domain output is the frequency domain input MULTIPLIED by the Transfer Function.
3. The frequency domain signal is the Fourier Transform of the temporal signal

Mathematically, it must be that the FT of a convolution is a product. $y(t) = x(t) * h(t) \xrightarrow{FT} Y(\omega) = X(\omega)H(\omega)$

1.12. Convolution Theorems:

Convolution of signals may be done either in time domain or frequency domain. So there are following two theorems of convolution associated with Fourier transforms:
1. Time convolution theorem
2. Frequency convolution theorem

Time convolution theorem: The time convolution theorem states that convolution in time domain is equivalent to multiplication of their spectra in frequency domain.
Mathematically, if

$x_1(t) \leftrightarrow X_1(\omega)$
And $x_2(t) \leftrightarrow X_2(\omega)$
Then $x_1(t) * x_2(t) \leftrightarrow X_1(\omega)X_2(\omega)$

Proof:

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} [x_1(t) * x3(t)]\, e^{-jwt}\, dt$$

We have $x_1(t) * x_2(t) = \int_{-\infty}^{\infty} [x_1(\tau)x3(t-\tau)]dr$

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} \{\int_{-\infty}^{\infty} [[x_1(\tau)x3(t-\tau)]\, d\tau\} e^{-jwt}\, dt$$

Interchanging the order of integration, we have

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau) \int_{-\infty}^{\infty} [x3(t-\tau)]e^{-jwt}\, dt]\, d\tau$$

Letting t-r = p, in the second integration, we have t=p+r and dt = dp

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau) \int_{-\infty}^{\infty} [x_2(p)\, e^{-jw(p+\tau)}dp]\, d\tau$$

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau) \int_{-\infty}^{\infty} [x^2(p)e^{-jwp}dp]e^{-jw\tau}\, d\tau$$

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau)\, X_2(w)\, e^{-jw\tau}\, d\tau$$

$$F[x_1(t) * x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau)\, e^{-jw\tau}\, d\tau\;\; X_2(w)$$

$x_1(t) * x_3(t) \leftrightarrow X_1(\omega) X_3(\omega)$

This is time convolution theorem

Frequency convolution theorem:

The frequency convolution theorem states that the multiplication of two functions in time domain is equivalent to convolution of their spectra in frequency domain.

Mathematically, if

$x_1(t) \leftrightarrow X_1(\omega)$

And $x_3(t) \leftrightarrow X_3(\omega)$

Then $x_1(t)\, x_3(t) \leftrightarrow 1/2\pi\, [X_1(\omega) * X_3(\omega)]$

Proof:

$F[x_1(t)\, x_3(t)] = \int_{-\infty}^{\infty} [x_1(t)\, x_2(t)]\, e^{-j\omega t} dt$

By definition of inverse Fourier transform

$$= \int_{-\infty}^{\infty} \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} X_1(\lambda) e^{j\lambda t} d\lambda\right] x_2(t)\, e^{-j\omega t} dt$$

Interchanging the order of integration, we get

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} X_1(\lambda) \left[\int_{-\infty}^{\infty} x_2(t)\, e^{-j\omega t} e^{j\lambda t} d\lambda\right] dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} X_1(\lambda) \left[\int_{-\infty}^{\infty} x_2(t)\, e^{-(\omega-\lambda)t} dt\right] d\lambda$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} X_1(\lambda)\, X_2(\omega-\lambda)\, d\lambda$$

$$= \frac{1}{2\pi}[X_1(\omega) * X_2(\omega)]$$

This is frequency convolution theorem in radian frequency in terms of frequency, we get

$F[x_1(t)\, x_3(t)] = X_1(f) * X_3(f)$

Parseval's Energy theorem:

For continuous time signals x(t) energy of the signal is expressed as:

$E = \int_{-\infty}^{\infty} x^2(t)\, dt$

According to Parseval's theorem

$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(jw)|^2 dw$

Let $x^*(t)$ is conjugate of x(t).

Therefore, $x(t) \cdot x^*(t) = |x(t)|^2$ .... (1)

On integrating the above equation with respect to t, we get

$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} x(t) \cdot x^*(t)\, dt$...........(2)

From the definition of inverse Fourier transform

$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(jw) e^{jwt} dw$ ....(3)

Taking conjugate to both sides

$x^*(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X^*(jw) e^{-jwt} dw$..............(4)

Putting the above value in equation (2)

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} x(t) \cdot \frac{1}{2\pi}\int_{-\infty}^{\infty} X^*(jw) e^{-jwt} dw\, dt$$

Changing the order of integration we get

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} X^*(jw) \int_{-\infty}^{\infty} x(t)\, e^{-jwt}\, dt\, dw$$

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} X^*(jw) X(jw)\, dw$$

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(jw)|^2 dt$$ Hence proved.

UNIT-II

Amplitude modulation: Modulation, need of modulation, types of modulation techniques, amplitude modulation (DSB-FC), modulation index, frequency spectrum of AM wave, linear and over modulation, power relation in AM, transmission efficiency, modulation by a complex signal, bandwidth of AM, AM modulators, square law and switching modulator, advantages and disadvantages of AM. Demodulation of AM: Suppressed carrier amplitude modulation systems, DSB-SC, SSB-SC, VSB-SC systems, comparison of various amplitude modulation systems. Demodulation of AM, square law and envelope detector, synchronous detection of AM, Low and high power AM transmitters, AM receivers, TRF and superheterodyne receivers, sensitivity, selectivity and fidelity of receivers.

2.1. INTRODUCTION:

Communication is a process whereby information is enclosed in a package and is channelled and imparted by a sender to a receiver via some medium. The receiver then decodes the message and gives the sender a feedback. So the basic elements of communication systems are:

- Transmitter: originates the signal
- Receiver: receives transmitted signal after it travels over the medium
- Medium: guides the signal from the transmitter to the receiver.

In a data transmission system, the transmission medium is the physical path between transmitter and receiver. For guided media, electromagnetic waves are guided along a solid medium, such as copper twisted pair, copper coaxial cable, and optical fibre. For unguided media, wireless transmission occurs through the atmosphere, outer space, or water.

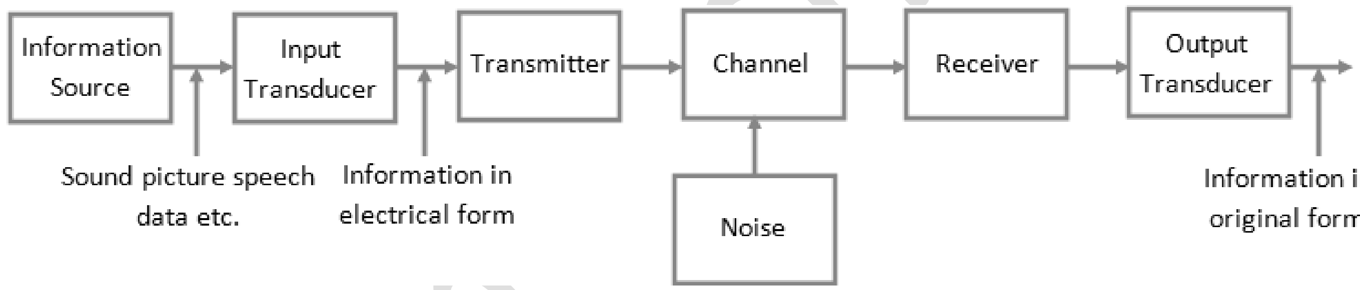


Figure2.1. Basic communication system

Elements of Communication system:

1. Information source: The message or information originates in the information source which has to be transmitted.

2. Input transducer: A transducer is the device which converts one form of energy to another form. In communication system transducer is usually required to convert the output of a source into an electrical signal that is suitable for transmission.

3. The Transmitter: The transmitter process the electrical signal into a form that is suitable for transmission through the transmission medium. The transmitter performs the signal processing of the message signal such as restriction of range of audio frequencies, amplification and modulation. All these processing are done to ease the transmission of the signal through the channel.

4. The Channel and the Noise: The communications channel is the physical medium that is used to send the signal from the transmitter to the receiver. In wireless transmission, the channel is usually the free space. On the other hand, telephone channels usually employ a variety of physical media, including wire lines, optical fibre cables, and wireless microwave radio. Whatever the physical medium for signal transmission, transmitted signal is corrupted in a random manner by noise.

5. The Receiver: The function of the receiver is to recover or reproduce the message signal contained in the received signal. If the message signal is transmitted by carrier modulation, the receiver performs *carrier demodulation* in order to extract the message from the sinusoidal carrier.
6. Output transducer: It is the final stage use to convert an electrical message signal into its original form.

2.2. Modulation:
Modulation is a technique used to convert a low frequency message signal to a higher frequency modulated signal using a higher frequency carrier.
Definition: Modulation is the process of changing the parameters of the carrier signal, in accordance with the instantaneous values of the modulating signal.
Signals in the Modulation Process:
1. Message or Modulating Signal
The signal which contains a message to be transmitted is called as a message signal. It is a baseband signal, which has to undergo the process of modulation, to get transmitted. Hence, it is also called as the modulating signal.
Baseband signal: Baseband refers to the original frequency range of a transmission signal before it is converted, or modulated, to a different frequency range.
2. Carrier Signal
The high frequency signal which has a certain phase, frequency, and amplitude but contains no information is called a carrier signal. It is an empty signal. It is just used to carry the signal to the receiver after modulation.
3. Modulated Signal
The resultant signal after the process of modulation is called as the modulated signal. This signal is a combination of the modulating signal and the carrier signal.
Signal Bandwidth:
The bandwidth of a signal represents the range of its frequency components. A complex signal is made of a range of frequencies called spectrum. The Bandwidth of a signal is calculated by subtracting the highest frequency component from the lowest frequency component.
Demodulation: It is the reverse process of modulation, which is used to get back the original message signal. Modulation is performed at the transmitting end whereas demodulation is performed at the receiving end.

2.3. Need for modulation:
The baseband signals are incompatible for direct transmission. When the signal is transmitted without modulation they cannot travel longer distances as low frequency signal get it attenuates, so its strength has to be increased by modulating with a high frequency carrier wave, which doesn’t affect the parameters of the modulating signal.
Modulation is needed to achieve the following basic needs:

1. Practicability of antennas: For the transmission of radio signals, the antenna height must be multiple of $\lambda/4$,where λ is the wavelength.
$\lambda = c/f$
Where c: is the velocity of light
f: is the frequency of the signal to be transmitted
The minimum antenna height required to transmit a baseband signal of f = 10 kHz is 7.5 Km.
The antenna of this height is practically impossible to install.
Now, let us consider a modulated signal at f = 1 MHz The minimum antenna height is 75 meters.
This antenna can be easily installed practically. Thus, modulation reduces the height of the antenna.

2. Avoids mixing of signals
If the baseband sound signals are transmitted without using the modulation by more than one transmitter,

then all the signals will be in the same frequency range i.e. 0 to 20 kHz. Therefore, all the signals get mixed together and a receiver cannot separate them from each other. If each baseband sound signal is used to modulate a different carrier then they will occupy different slots in the frequency domain i.e. through different channels. Thus, modulation avoids mixing of signals.

3. Multiplexing is possible: Multiplexing is a process in which two or more signals can be transmitted over the same communication channel simultaneously. If transmitted without modulation, the different message signals over a single channel will interfere with each other. So multiplexing helps in transmitting a number of messages simultaneously over a single channel which reduces cost of installation and maintenance of more channels.

4. Narrow banding: The frequency translation through modulation converts a wideband signal to a narrowband, which is termed as narrow banding.

Let us assume a system is radiating directly with the frequency range from 50 Hz to 10 kHz, the ratio of highest to lowest wavelength is 200. If antenna is designed for 50 Hz, it will be too long for 10 kHz and vice versa. But if signal is translated to higher frequency of 1 MHz range using modulation, then the ratio of lowest to highest frequency will be $\frac{10^6+50}{10^6+10^4} \approx 1$ and the same antenna will be suitable for the entire band.

5. Improves Quality of Reception

6. Increase the Range of Communication

2.4. Types of Modulation:

The types of modulations are broadly classified into continuous-wave modulation and pulse modulation.

Continuous-wave Modulation

In the continuous-wave modulation, a high frequency sine wave is used as a carrier wave. This is further divided into amplitude and angle modulation.

- If the amplitude of the high frequency carrier wave is varied in accordance with the instantaneous amplitude of the modulating signal, then such a technique is called as Amplitude Modulation.
- If the angle of the carrier wave is varied, in accordance with the instantaneous value of the modulating signal, then such a technique is called as Angle Modulation.

The angle modulation is further divided into frequency and phase modulation.

- If the frequency of the carrier wave is varied, in accordance with the instantaneous value of the modulating signal, then such a technique is called as Frequency Modulation.
- If the phase of the high frequency carrier wave is varied in accordance with the instantaneous value of the modulating signal, then such a technique is called as Phase Modulation.

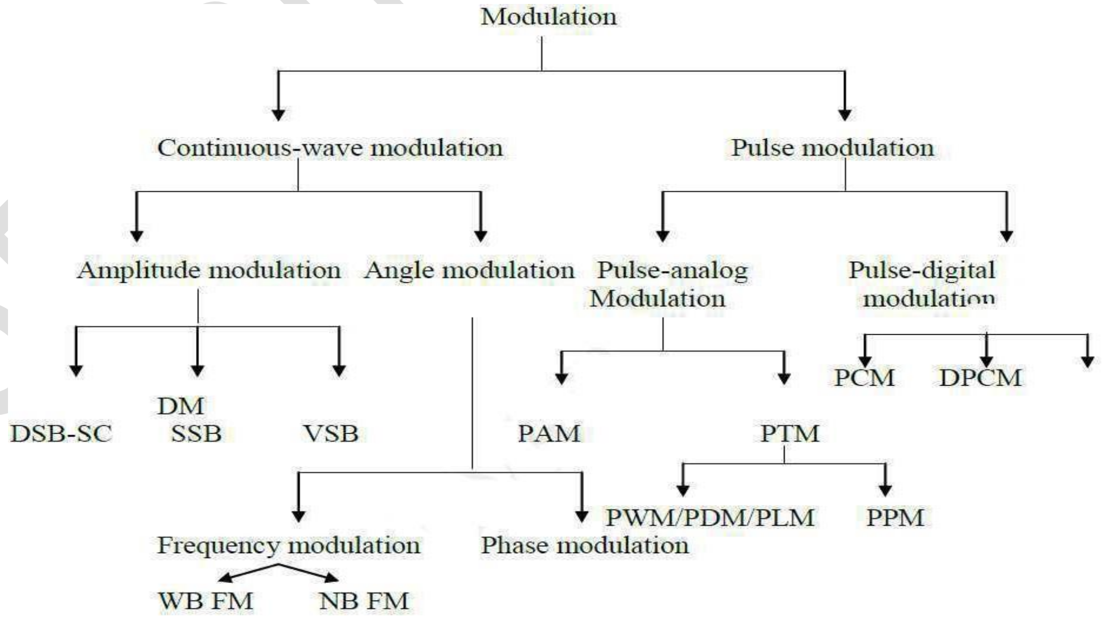


Figure 2.2 Types of modulation

Pulse Modulation

In Pulse modulation, a periodic sequence of rectangular pulses is used as a carrier wave. This is further divided into analog and digital modulation.

- In analog modulation technique, if the amplitude, duration or position of a pulse is varied in accordance with the instantaneous values of the baseband modulating signal, then such a technique is called as Pulse Amplitude Modulation (PAM) or Pulse Duration/Width Modulation (PDM/PWM), or Pulse Position Modulation (PPM).
- In digital modulation, the modulation technique used is Pulse Code Modulation (PCM) where the analog signal is converted into digital form of 1s and 0s. As the resultant is a coded pulse train, this is called as PCM. This is further developed as Delta Modulation (DM), which will be discussed in subsequent chapters. Hence, PCM is a technique where the analog signals are converted into a digital form.

2.5. Amplitude modulation

Definition:

The amplitude of the carrier signal varies in accordance with the instantaneous amplitude of the modulating signal i.e. the amplitude of the carrier signal containing no information varies as per the amplitude of the signal containing information, at each instant.

Mathematical expression:

Let m (t) is the baseband message and C (t) = $A_c \cos(\omega_c t)$ is called the carrier wave. The carrier frequency, f_c should be larger than the highest spectral component in m(t).

Consider a sinusoidal carrier signal C (t) is defined as

$$C(t) = A_c \cos(2\pi f_c t + \Phi)\, t$$

Where A_c= Amplitude of the carrier signal

f_c= frequency of the carrier signal

Φ = Phase angle.

For convenience, assume the phase angle of the carrier signal is zero. An amplitude-modulated (AM) wave, S(t) can be described as function of time is given by

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

Where the parameter k_a is a positive constant called the amplitude sensitivity of the modulator.

Let $e(t) = A_c|1 + k_a m(t)|$ is called the envelope of the AM signal. When f_c is large relative to the bandwidth of m(t), the envelope is a smooth signal that passes through the positive peaks of S(t) and it can be viewed as modulating the amplitude of the carrier wave in a way related to m(t) as shown in figure 2.3.

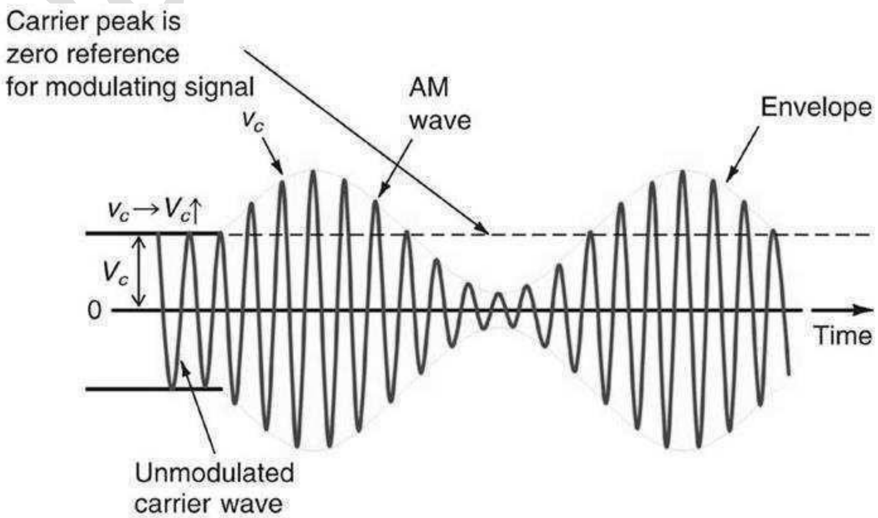


Figure 2.3.Amplitude modulation envelope in time domain

The amplitude modulated (AM) signal consists of both modulated carrier signal and un-modulated carrier signal. There are two requirements to maintain the envelope of AM signal is same as the shape of base band signal.

1. The amplitude of the k_a m(t) is always less than unity i.e., $|k_a m(t)|<1$ for all 't'.
2. The carrier signal frequency f_c is far greater than the highest frequency component W of the message signal m (t) i.e., $f_c >> W$

Assume the message signal m (t) is band limited to the interval $-W < f < W$

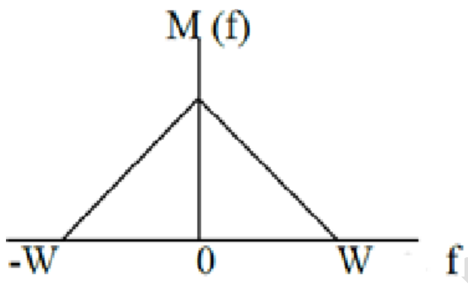


Fig. 2.4.Spectrum of message signal

The spectrum of AM is shown in fig. 3.5.The Fourier transform of AM signal S (t) is

$S(f) = \frac{Ac}{2}[\delta(f-fc)+ \delta(f+fc)]+ \frac{K_{ac}}{2}[M(f-fc)+ M(f+fc)]$

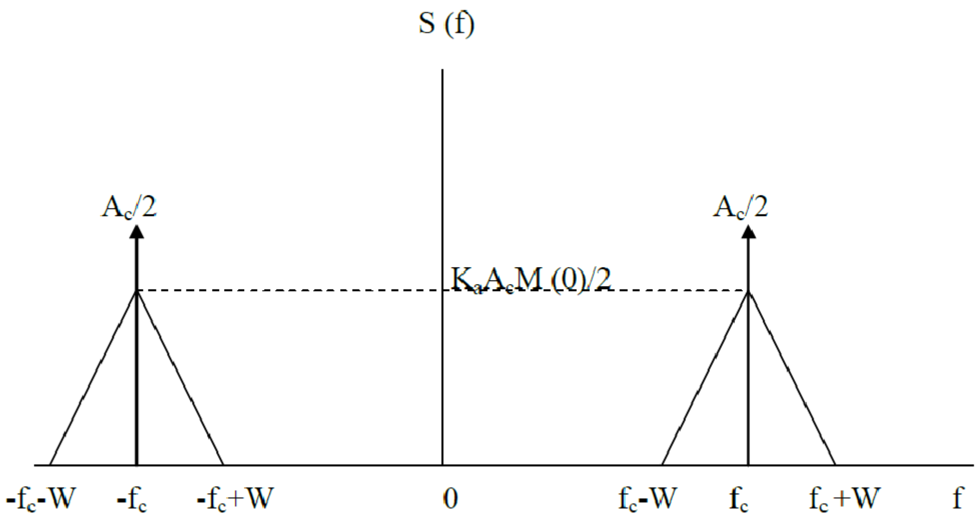


Fig.2.5. Spectrum of AM signal

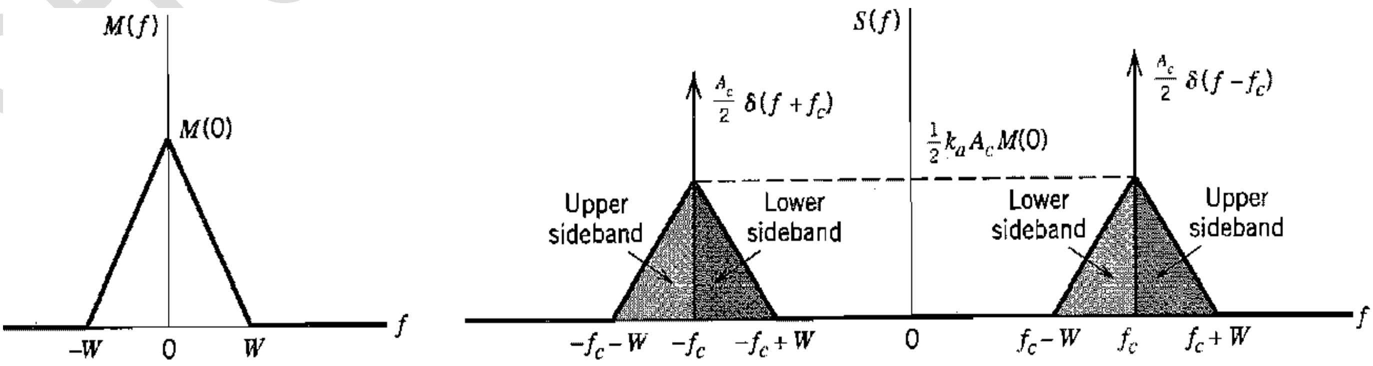


Fig.2.6. Spectrum of AM signal representing sidebands

The AM spectrum consists of two impulse functions which are located at f_c and $-f_c$ and weighted by $A_c/2$, two USBs, band of frequencies from f_c to f_c+W and band of frequencies from $-f_c-W$ to $-f_c$, and two LSBs, band of frequencies from f_c-W to f_c and $-f_c$ to $-f_c+W$.

The difference between highest frequency component and lowest frequency component is known as transmission bandwidth.
B = 2W

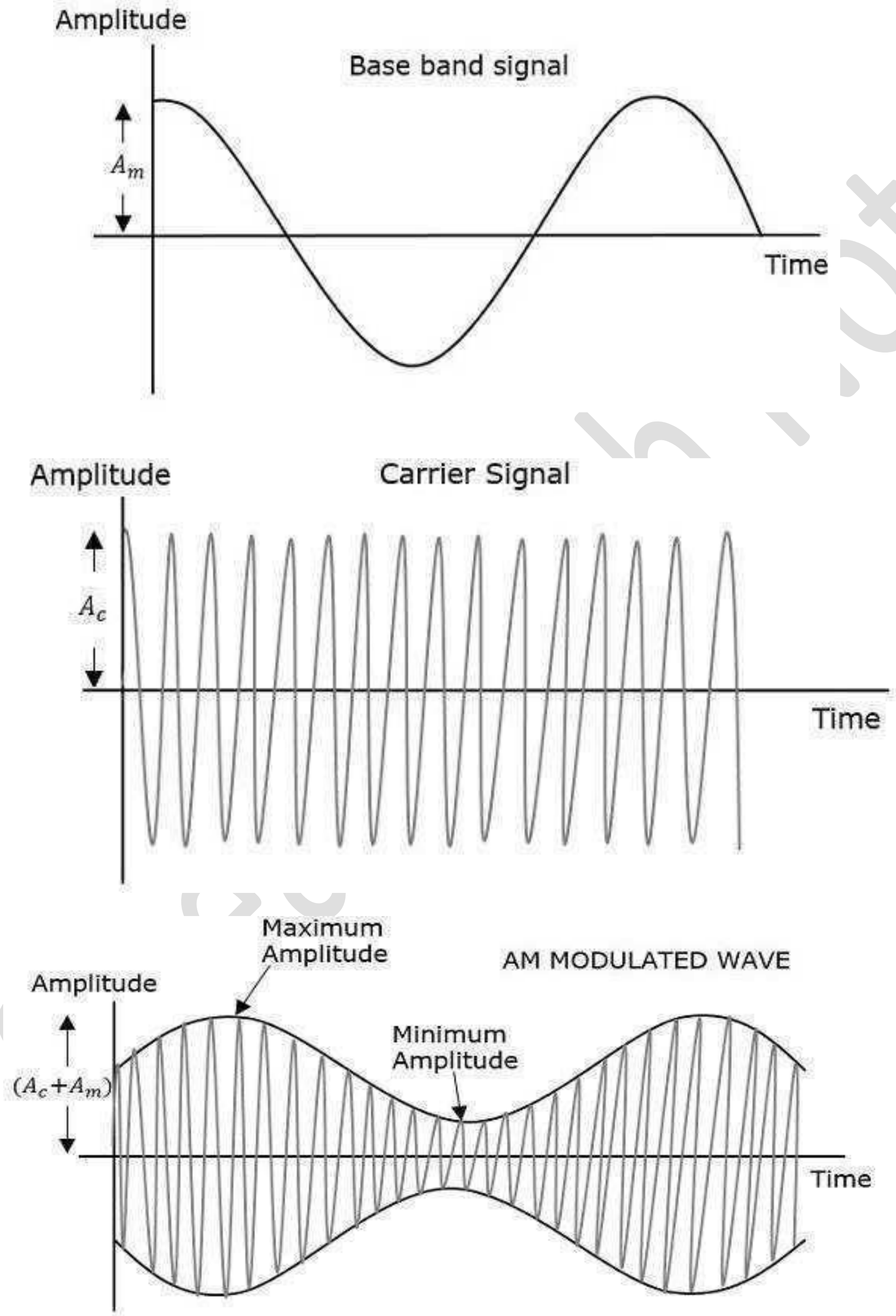


Figure 2.7.Amplitude modulation waveform in time domain

2.6. Single-tone modulation

In single-tone modulation modulating signal consists of only one frequency component where as in multi-tone modulation modulating signal consists of more than one frequency component.

Mathematical Expressions:

Following are the mathematical expressions for these waves.

Time-domain Representation of the Waves

Let the modulating signal be,

$$m(t) = A_m \cos 2\pi f_m t \qquad (2.1)$$

and the carrier signal be,

$$C(t) = A_c \cos(2\pi f_c t) \qquad (2.2)$$

Where,

A_m and A_c are the amplitude of the modulating signal and the carrier signal respectively.

f_m and f_c are the frequency of the modulating signal and the carrier signal respectively.

The equation for the overall modulated signal is obtained by multiplying the carrier and the modulating signal together.

$$S(t) = A_c [1 + k_a m(t)] \cos(2\pi f_c t) \qquad (2.3)$$

Substituting in the individual relationships for the carrier and modulating signal in equation (3.3), the overall signal becomes:

$$S(t) = Ac [1 + k_a A_m \cos 2\pi f_m t] \cos(2\pi f_c t)$$

Replace the term $k_a A_m$ by μ which is known as modulation index or modulation factor.

Or it can be written as

$$S(t) = [A_c + A_m \cos(2\pi f_m t)] \cos(2\pi f_c t) \qquad (2.4)$$

Modulation Index:

A carrier wave, after being modulated, if the modulated level is calculated, then such an attempt is called as Modulation Index or Modulation Depth. Modulation index can be defined as the measure of extent of amplitude variation about an un-modulated carrier.

Rearrange the Equation 4 as below.

$$S(t) = A_c[1 + (A_m/A_c) \cos(2\pi f_m t)] \cos(2\pi f_c t) \qquad (2.5)$$

$$S(t) = A_c[1 + \mu \cos(2\pi f_m t)] \cos(2\pi f_c t) \qquad (2.6)$$

Where, μ is Modulation index or Amplitude sensitivity of the modulator and it is equal to the ratio of A_m and A_c. Mathematically, we can write it as

$$\mu = (A_m/A_c)$$

Calculating the modulation index from AM envelope:

With reference to the figure 3.7 and 3.8, we can calculate the modulation index from the modulated waveform. We know that $\mu = (A_m/A_c)$

$$A_m = (A_{max} - A_{min})/2 \qquad (2.8)$$

$$A_c = A_{max} - A_m \qquad (2.9)$$

By substituting (3.8) equation in equation (3.9) we get

$$A_c = A_{max} - (A_{max} - A_{min})/2 \qquad (2.10)$$

By diving (3.8) and (3.10) equation we get

$$\mu = (A_m/A_c) = \frac{max - m}{max + mn} \qquad (2.11)$$

Where

A_{max} = maximum amplitude of the modulated carrier signal

A_{min} = minimum amplitude of the modulated carrier signal

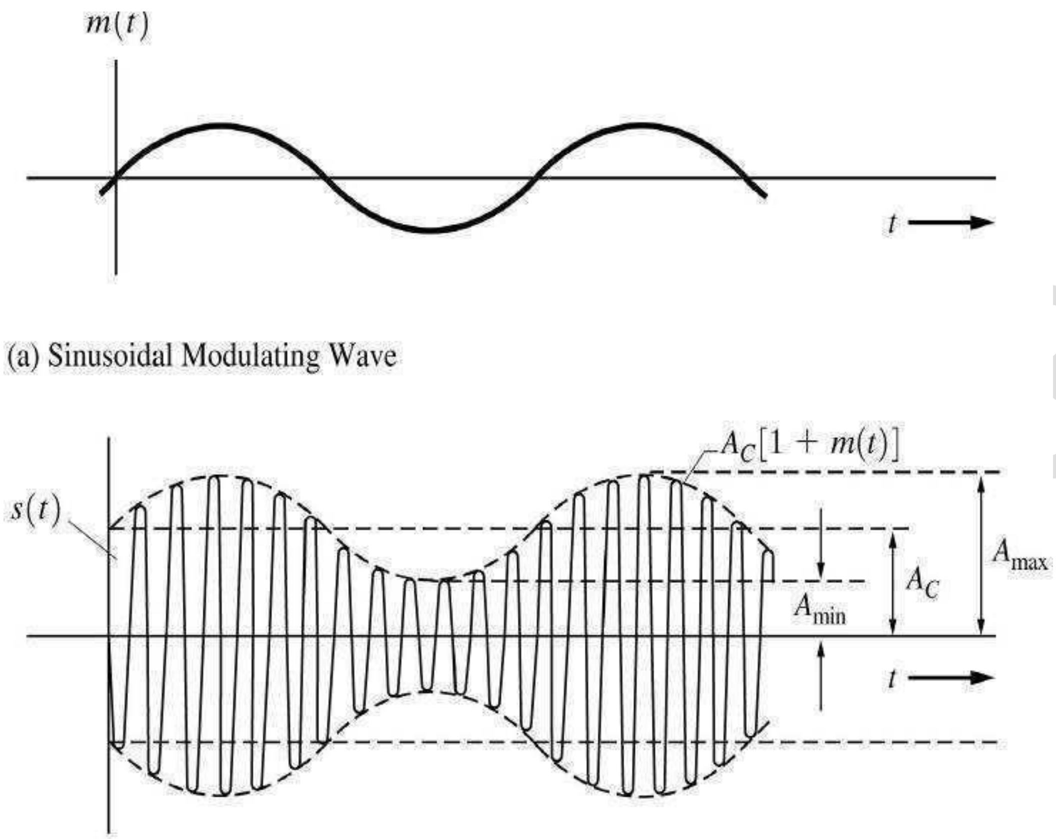


Figure 2.8.AM envelope

Modulation index μ has to be governed such that it is always less than unity; otherwise it results in a situation known as 'over-modulation' (μ >1). The over-modulation occurs, whenever the magnitude of the peak amplitude of the modulating signal exceeds the magnitude of the peak amplitude of the carrier signal. The signal gets distorted due to over modulation. Because of this limitation on'μ ', the system clarity is also limited. The AM waveforms for different values of modulation index m are as shown in figure 3.9.

If μ = 0 we haven't modulating wave, then no information is transmitted while engaging the channel with the carrier.

If μ= 1 we have the maximum of modulation. When the modulation index is 1, i.e. a modulation depth of 100%, the carrier level falls to zero and rise to twice its non-modulated level.
We are in optimal conditions if μ = 0.5.

If μ > 1 then we have strong crossover distortion. Any increase of the modulation index above 1.0, i.e. 100% modulation depth causes over-modulation. The carrier experiences 180° phase reversals where the carrier level would try to go below the zero point. These phase reversals give rise to additional sidebands resulting from the phase reversals (phase modulation) that extend out, in theory to infinity. This can cause interference to other users if not filtered.

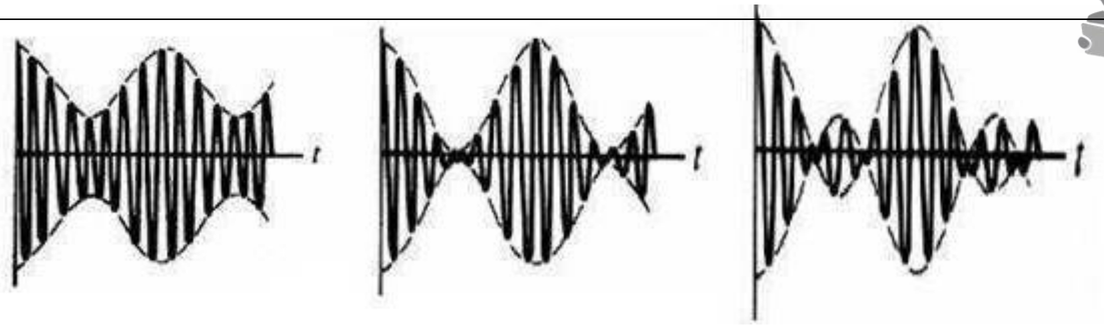


Figure 2.9.AM waveforms for different values of μ

$$S(t) = A_c \cos(2\pi f_c t) + A_c\mu/2[\cos 2\pi (f_c + f_m)t] + A_c\mu/2[\cos 2\pi (f_c - f_m)t] \quad (2.12)$$

- Looking at equation (3.12) we can say that 1st term represents un-modulated carrier and two additional terms represents two sidebands
- The frequency of the lower sideband (LSB) is $f_c - f_m$ and the frequency of the upper sideband (USB) is $f_c + f_m$

Fourier transform of S (t) is

$$S(f) = A_c/2[\delta(f-f_c) + \delta(f+f_c)] + A_c\mu/4[\delta(f-f_c-f_m) + \delta(f+f_c+f_m)] + A_c\mu/4[\delta(f-f_c+f_m) + \delta(f+f_c-f_m)] \quad (2.13)$$

Bandwidth of AM wave:

- We know bandwidth can be measured by subtracting lowest frequency of the signal from highest frequency of the signal
- For amplitude modulated wave it is given by

$BW = f_{USB} - f_{LSB}$

$= (f_c + f_m) - (f_c - f_m)$

$= 2 f_m$

Therefore the bandwidth required for the amplitude modulation is twice the frequency of the modulating signal.

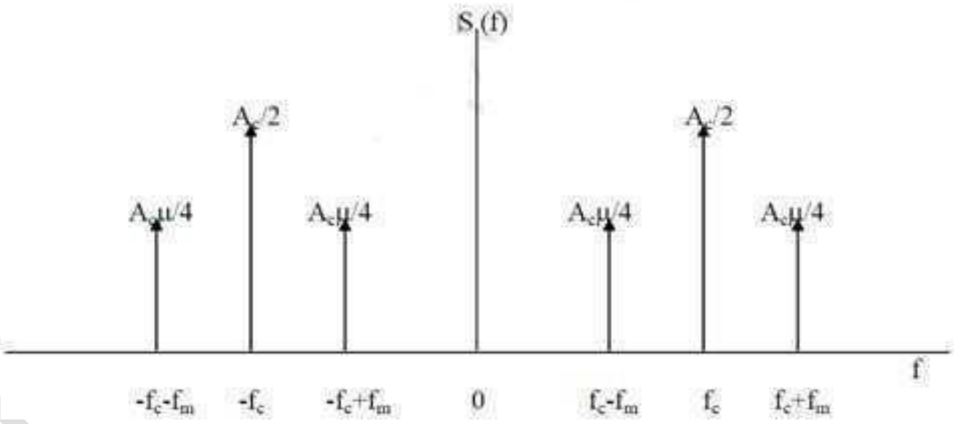


Figure 2.10. Spectrum of Single tone AM signal

Power calculations of single-tone AM signal:

The standard time domain equation for single-tone AM signal is given by equation 2.12

$$S(t) = A_c \cos(2\pi f_c t) + A_c\mu/2[\cos 2\pi (f_c + f_m)t] + A_c\mu/2[\cos 2\pi (f_c - f_m)t] \quad (2.12)$$

We have seen that AM wave has three components:

- Un-modulated carrier
- Lower sideband
- Upper sideband

Therefore the total power of AM wave is the sum of the carrier power Pc and Power in the two sidebands P_{USB} and $P_{LSB.}$ It is given as

Power of any signal is equal to the mean square value of the signal

Carrier power $P_c = A_c^2 /2$

Upper Side Band power $P_{USB} = A_c^2 \mu^2 /8$

Lower Side Band power $P_{LSB} = A_c^2\mu^2 /8$

Total power $P_T = P_c + P_{LSB} + P_{USB}$

Total power $P_T = A_c^2 /2 + A_c^2 \mu^2 /8 + A_c^2 \mu^2 /8$

$P_T = P_c [1+ \mu^2 /2]$

2.7. Multi-tone modulation:

In multi-tone modulation modulating signal consists of more than one frequency component where as in single-tone modulation modulating signal consists of only one frequency component.

Mathematical Expression

Let us consider that a carrier signal $A_c \mathrm{Cos}(2\pi f_c t)$ is modulated by a baseband or modulating signal m(t) which is expressed as :

$$m(t) = A_{m1} \mathrm{Cos}(2\pi f_{m1}t) + A_{m1} \mathrm{Cos}(2\pi f_{m2}t) \qquad (2.14)$$

We know that the general expression for AM wave is

$$S(t) = A_c \mathrm{Cos}(2\pi f_c t) + m(t) \mathrm{Cos}(2\pi f_c t)$$

Putting the value of x(t), we get

$$S(t) = A_c \mathrm{Cos}(2\pi f_c t) + [A_{m1} \mathrm{Cos}(2\pi f_{m1}t) + A_{m2} \mathrm{Cos}(2\pi f_{m2}t)] \mathrm{Cos}(2\pi f_c t) \qquad (2.15)$$

or it can be written as

$$S(t) = A_c [1 + K_a A_{m1} \mathrm{Cos}(2\pi f_{m1}t) + K_a A_{m2} \mathrm{Cos}(2\pi f_{m2}t)] \mathrm{Cos}(2\pi f_c t) \qquad (2.16)$$

Replace $K_a A_{m1}$ by μ_1 and $K_a A_{m2}$ by μ_2

So finally we get

$$S(t) = A_c \mathrm{Cos}(2\pi f_c t) + \frac{c\mu_1}{2} [\cos 2\pi (f_c + f_{m1})t] + \frac{c\mu_1}{2} [\cos 2\pi(f_c - f_{m1})t] + \frac{c\mu_2}{2} [\cos 2\pi (f_c + f_{m2})t] + \frac{c\mu_2}{2} [\mathrm{Cos}\, 2\pi (f_c - f_{m2})t] \qquad (2.17)$$

Power of multi-tone AM signal is given by:

$$P_T = P_c [1+ \mu_1^2 /2 + \mu_2^2 /2 + \ldots\ldots\ldots\ldots + \mu_n^2/2]$$

Where P_t = Total power

P_c = Carrier power

$$P_T = P_c [1+ \mu_t^2 /2]$$

Where $\mu_t = \sqrt{\mu_1^2 + \mu_2^2 + \cdots + \mu_n^2}$

Fourier transform of S(t) is

$$S(f) = \frac{c}{2} [\delta(f-f_c)+ \delta(f+f_c)] + \frac{c\mu_1}{4} [\delta(f-f_c-f_{m1})+ \delta(f+f_c+f_{m1})] + \frac{c\mu_1}{4} [\delta(f-f_c+f_{m1})+ \delta(f+f_c-f_{m1})] + \frac{c\mu_2}{4} [\delta(f-f_c-f_{m2})+ \delta(f+f_c+f_{m2})]+ \frac{c\mu_2}{4} [\delta(f-f_c+f_{m2})+ \delta(f+f_c-f_{m2})]$$

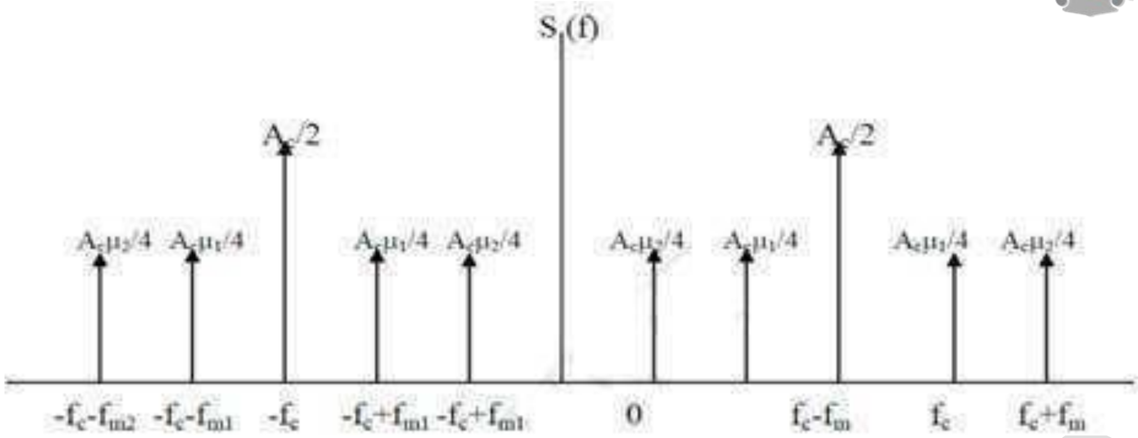


Figure 2.11. Spectrum of Multi tone AM signal

Transmission efficiency:
Transmission efficiency is defined as the ratio of total side band power to the total transmitted power. The yield of modulation is defined therefore as the ratio between the transmitted information signal strength content in one of the two side lines, divided by all the power you must transmit.

$$\eta = \frac{P_{LSB}+P_{USB}}{P_T}$$

$$\eta = \frac{\mu^2}{(2+\mu^2)} \text{ X } 100\ \% \qquad (2.18)$$

The transmission efficiency (η) of AM wave is defined as the percentage of total power contributed by side bands of the AM signal. The maximum transmission efficiency of an AM signal is 33.33%, i.e., only one third of the total transmitted power is carried by the side bands in an AM wave. The remaining two third of the total transmitted power gets wasted.

Advantages of Amplitude modulation:
Generation and detection of AM signals are very easy
It is very cheap to build, due to this reason it is most commonly used in AM radio broad casting

Disadvantages of Amplitude of modulation:
Amplitude modulation is wasteful of power
Amplitude modulation is wasteful of band width

2.8. Modulation by a complex signal
A complex carrier signal c(t), at a carrier frequency ω_c , is described mathematically as the complex exponential
$C(t) = e^{(\omega_c t+)}$
For convenience we choose the initial time so that the phase (δ) is zero. Then, if m(t) is the signal or information that is to be transmitted by the carrier, the signal m(t) is encoded onto the carrier by multiplying the carrier by m(t)
S(t) = m(t) c(t)
$S(t) = m(t)e^{(\omega_c t)}$
The carrier's amplitude is modulated by the signal m(t). Now we know that multiplication in the time domain is equivalent to convolution in the frequency domain. Thus, the Fourier transform of the signal s(t) is the convolution of the Fourier transforms of m(t) and c(t).
S(ω) = M(jω) * C(jω)
$$S(\omega) = \frac{1}{2\pi}\int_{-\infty} M(\omega)C((\omega-\omega_0))d\omega_0$$

Earlier we took the Fourier transform of a complex exponential and determined it is a delta function

$C(j\omega) = 2\pi\ \delta(\omega-\omega_c)$

and upon substitution into the convolution equation we obtain

$S(\omega) = M(j(\omega-\omega_c))$

Thus, as a result of modulation, the transform of the signal m(t) is shifted on the frequency axis by the carrier frequency. We can visualize the situation by considering the magnitude of M ($j\omega$). We suppose that the signal m(t) is a real function of time and that its frequency content is bounded by some maximum frequency ω_m . Hence, all of the signal power lies in the range $\pm\ \omega_m$, as depicted in the figure 2.12 below. The second figure depicts the delta function at ω_c and the third figure shows the result of amplitude modulation.

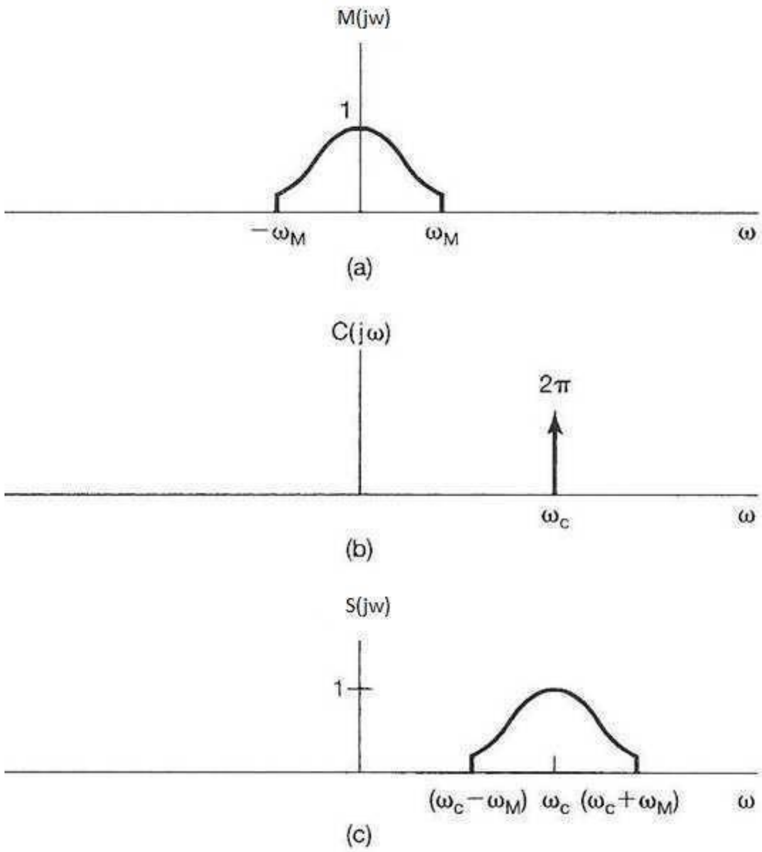


Figure 2.12 Complex AM spectrum

2.9. Generation of AM waves

The amplitude modulator is a circuit which generates amplitude modulated signal. In the process of modulation the frequency spectrum gets translated. The output of the modulator contains the frequencies which are different from those present in the input signal. The amplitude modulator therefore must be time varying linear systems such as switching or chopping circuit are a non linear time in varying system. The reason for this is that a linear time invariant system cannot produce new frequencies in its output. Here two methods for generating AM waves:

1. The square law are power law modulator
2. Switching modulator.

These two methods require non linear element as active device for generating AM signals. These two methods are use full in the low power generation of amplitude modulated waves.

Square-law modulator:

It consists of the following:

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

The modulating signal and carrier are connected in series with each other and their sum $V_1(t)$ is applied at the input of the non-linear device semi-conductor diodes and transistors are the most common nonlinear devices used for implementing square law modulators. The filtering requirement is usually satisfied by using a single or double tuned filters.

When a nonlinear element such as a diode is suitably biased and operated in a restricted portion of its characteristic curve, that is ,the signal applied to the diode is relatively weak, we find that transfer characteristic of diode-load resistor combination can be represented closely by a square law.

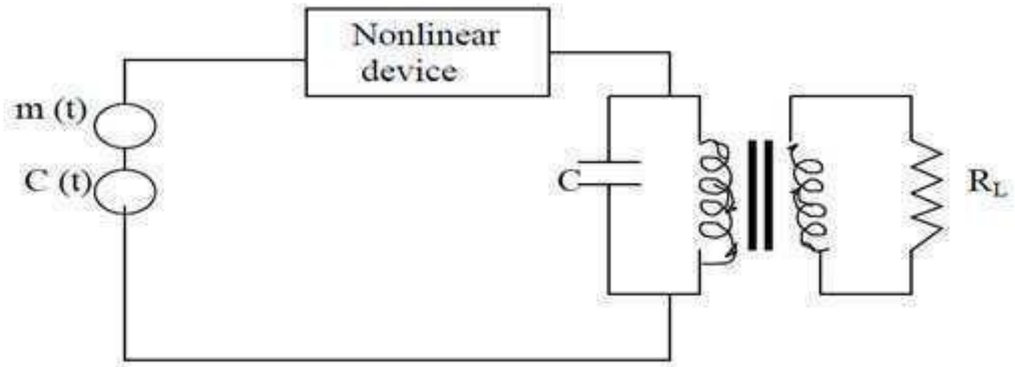


Figure 2.13.Square law modulator

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

$$V_0(t) = a_1 V_i(t) + a_2 V_i^2(t) \qquad (2.19)$$

Where a_1, a_2 are constants now, the input voltage $V_i(t)$ is the sum of both carrier and message signals

$$\text{i.e., } V_i(t) = A_c \text{Cos}(2\pi f_c t) + m(t) \qquad (2.20)$$

Substitute equation (2.20) in equation (2.19) we get

$$V_0(t) = a_1 A_c \text{Cos}(2\pi f_c t) + a_1 m(t) + a_2 [A_c \text{Cos}(2\pi f_c t) + m(t)]^2 \qquad (2.21)$$

$$V_0(t) = a_1 A_c \text{Cos}(2\pi f_c t) + a_1 m(t) + a_2 A_c^{\,2} \cos^2(2\pi f_c t) + a_2 m^2(t) + 2 a_2 A_c \text{Cos}(2\pi f_c t)\, m(t) \quad (2.22)$$

The five terms in the expression for $V_0(t)$ are as under :

Term 1: $a_1 m(t)$: Modulating Signal

Term 2: $a_1 A_c \text{Cos}(2\pi f_c t)$: Carrier Signal

Term 3: $a_2 m^2(t)$: Squared modulating Signal

Term 4: $2 a_2 A_c \text{Cos}(2\pi f_c t)\, m(t)$: AM wave with only sidebands

Term 5: $a_2 A_c^{\,2} \cos^2(2\pi f_c t) + a_2 m^2(t)$: Squared Carrier

Out of these five terms, terms 2 and 4 are useful whereas the remaining terms are not useful.

Let us combine terms 2, 4 and 1, 3, 5 as follows to get,

$$V_0(t) = \{a_1 m(t) + a_2 A_c^{\,2} \cos^2(2\pi f_c t) + a_2 m^2(t)\} + \{a_1 A_c \text{Cos}(2\pi f_c t) + 2 a_2 A_c \text{Cos}(2\pi f_c t)\, m(t)\} \quad (2.23)$$

Now design the tuned filter /Band pass filter with center frequency f_c and pass band frequency width 2W. We can remove the unwanted terms by passing this output voltage $V_0(t)$ through the band pass filter and finally we will get required AM signal.

$$V_0(t) = a_1 A_c \left[1 + 2\frac{a_2}{a_1} m(t)\right] \text{Cos}(2\pi f_c t) \qquad (2.24)$$

Where $K_a = 2\frac{a_2}{a_1}$

Assume the message signal m (t) is band limited to the interval $-W \le f \le W$

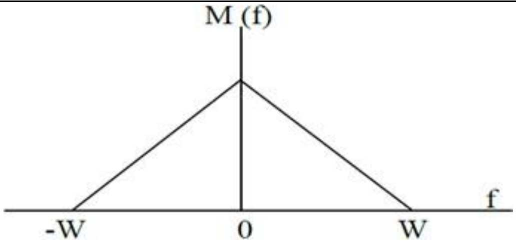


Figure 2.14.Spectrum of message signal

Spectrum of AM can represented a one shown in figure 2.15.The Fourier transform of output voltage V_O (t) is given by

$$V_O(f) = a_1 A_C/2[(f-f_c) + (f+f_c)] + a_2 A_C [M(f-f_c) + M(f+f_c)] \quad (2.25)$$

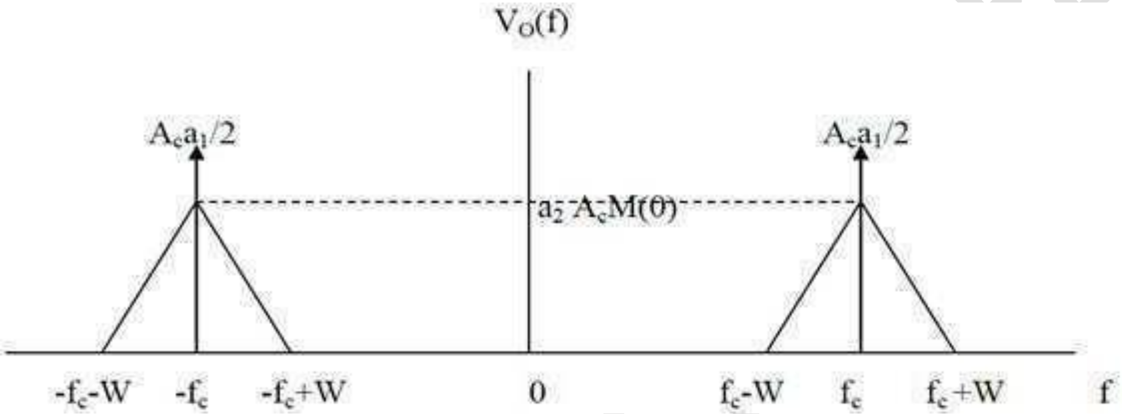


Figure 2.15.Spectrum of AM

The AM spectrum consists of two impulse functions which are located at f_c & $-f_c$ and weighted by $A_c a_1/2$ & $a_2 A_c/2$, two USBs, band of frequencies from f_c to f_c+W and band of frequencies from $-f_c$-W to $-f_c$, and two LSBs, band of frequencies from f_c-W to f_c & $-f_c$ to $-f_c$+W.

Switching Modulator:

In switching modulator the diode has to operate as an ideal switch as one shown in figure 2.16. Let the modulating and carrier signals be denoted as m(t) and $c(t)=A_c \cos(2\pi f_c t)$ respectively.

Working of circuit:

- The two signals i.e. modulating and carrier signals are applied as inputs to the summer (adder) block.
- Assume that carrier wave C(t) applied to the diode is large in amplitude, so that it swings right across the characteristic curve of the diode and also the diode acts as an ideal switch, that is, it presents zero impedance when it is forward-biased and infinite impedance when it is reverse-biased.
- We may thus approximate the transfer characteristic of the diode-load resistor combination by a piecewise-linear characteristic. Summer block produces an output, which is the addition of modulating and carrier signals.
- During the positive half cycle of the carrier signal i.e. if C (t)>0, the diode is forward biased, and then the diode acts as a closed switch. Now the output voltage V_o (t) is same as the input voltage V_i (t) .
- During the negative half cycle of the carrier signal i.e. if C (t) <0, the diode is reverse biased, and then the diode acts as an open switch. Now the output voltage V_O (t) is zero i.e. the output voltage varies periodically between the values input voltage V_i (t) and zero at a rate equal to the carrier frequency f_c.

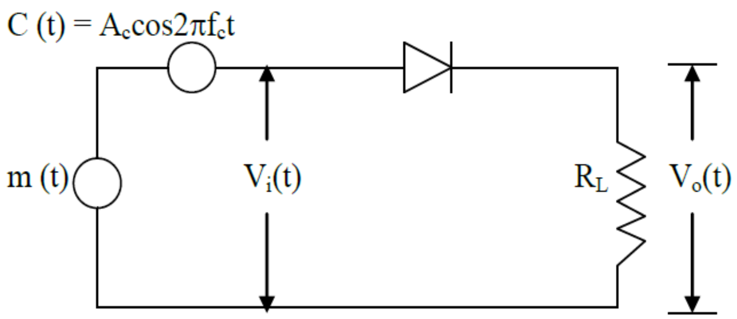


Figure 2.16.Switching modulator

Mathematically, we can write it as

The input voltage applied V_i (t) applied to the diode is the sum of both carrier and message signals.

$$V_i(t) = A_c \cos(2\pi f_c t) + m(t) \qquad (2.26)$$

$$V_o(t) = [A_c \cos(2\pi f_c t) + m(t)]\, g_P(t) \qquad (2.27)$$

Where $g_p(t)$ is the periodic pulse train with duty cycle one-half and period $T_c=1/f_c$ and which is given by

$$g_P(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos[2\pi f_c t(2n-1)] \qquad (2.28)$$

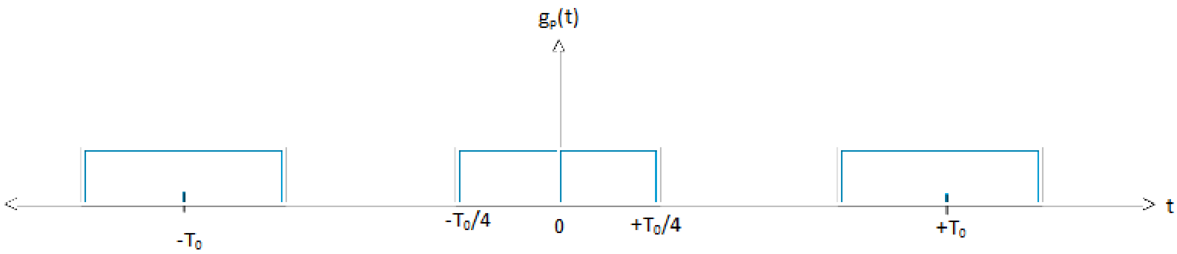


Figure 2.17 Pulse train

Substituting $g_p(t)$ into equation (2.27), we get

$$V_o(t) = \frac{1}{2} m(t) + \frac{1}{2} A_c \cos(2\pi f_c t) + \frac{2}{\pi} m(t)\cos(2\pi f c t) + \frac{2_c}{\pi} \cos^2(2\pi f_c t) \qquad (2.29)$$

The odd harmonics in this expression are unwanted, and therefore, are assumed to be eliminated. In this expression, the first and the fourth terms are unwanted terms whereas the second and third terms together represent the AM wave.

Combining the second and third terms together, we obtain

$$V_o(t) = \frac{c}{2}\left[1 + \frac{4}{\pi c} m(t)\right]\cos(2\pi f_c t) + \text{unwanted terms} \qquad (2.30)$$

This is the required expression for the AM wave with μ= [4/πE_c].

The unwanted terms can be eliminated using a band-pass filter (BPF). Now design the tuned filter /Band pass filter with center frequency f_c and pass band frequency width 2W.We can remove the unwanted terms by passing this output voltage V0(t) through the band pass filter and finally we will get required AM signal.

Assume the message signal m(t) is band limited to the interval –W ≤f ≤W as one shown in figure 2.18

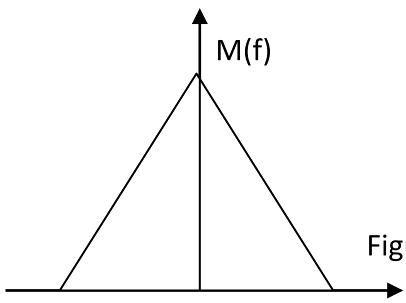


Figure 2.18 Spectrum of message signal

The spectrum of Am signal is shown in figure 2.19.The Fourier transform of output voltage $V_O(t)$ is given by

$V_O(f) = A_C/4[\delta(f-f_c) + \delta(f+fc)] + A_C/\pi\ [M(f-f_c) + M(f+f_c)]$ (2.31)

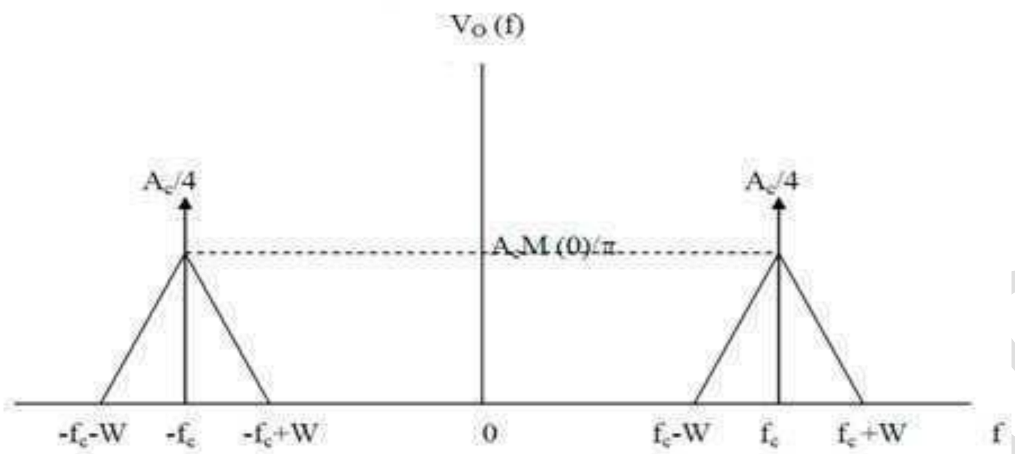


Figure 2.19. Spectrum of AM signal

The AM spectrum consists of two impulse functions which are located at fc & -fc and weighted by $A_c a_1/2$ & $a_2 A_c/2$, two USBs, band of frequencies from fc to fc +W and band of frequencies from -fc-W to –fc, and two LSBs, band of frequencies from fc-W to fc & -fc to -fc+W.

2.10. Advantages:

1. It is very simple to design and implement
2. It can be demodulated using a circuit consisting of very few components
3. AM receivers are very cheap as no specialised components are needed.
4. AM signal are reflected back to earth from ionosphere layer. Due to this fact, AM signals can reach far places which are thousands of miles from source. Hence AM radio has coverage wider compare to FM radio.

Disadvantage:

1. Due to large time constant, some distortion occurs which is known as diagonal clipping i.e., selection of time constant is somewhat difficult
2. The most natural as well as man-made radio noise are of AM type. The AM receivers do not have any means to reject this kind of noise.
3. Weak AM signals have low magnitude compare to strong signals. This requires AM receiver to have circuitry to compensate for signal level difference.
4. It is not efficient in terms of its use of bandwidth, requiring a bandwidth equal to twice that of the highest audio frequency

Application:

- **Broadcast transmissions:** AM is still widely used for broadcasting on the long, medium and short wave bands.

- **Air band radio:** VHF transmissions for many airborne applications still use AM. . It is used for ground to air radio communications as well as two way radio links for ground staff as well.

2.11. Suppressed carrier Amplitude modulation systems:

Objective: In full AM (DSB-AM), the carrier wave C (t) is completely independent of the message signal m(t), which means that the transmission of carrier wave represents a waste of power. This points to a shortcoming of amplitude modulation, that only a fraction of the total transmitted power is affected by m(t).Thus, the carrier signals and one of the two sidebands may be removed or attenuated so the resulting signals will require less transmitted power and will occupy less bandwidth, and yet perfectly acceptable communications will be possible.

2.12. Double Sideband-Suppressed Carrier (DSBSC) Modulation

Double sideband-suppressed (DSB-SC) modulation, in which the transmitted wave consists of only the upper and lower sidebands. Transmitted power is saved through the suppression of the carrier wave, but the channel bandwidth requirement is same as in AM that is twice the bandwidth of the message signal. In power calculation of AM signal, it has been observed that for single-tone sinusoidal modulation, the ratio of the total power and carrier power is

$\frac{P_t}{P_c} = [1 + \frac{\mu^2}{2}]$

$\frac{P_c}{P_t} = \frac{2}{3} \times 100\ \% = 67\%(\text{for } \mu = 1)$

So for 100% modulation that is μ = 1, about 67% of the total power is wasted for transmitting carrier which does not contain any information. So if carrier is suppressed, saving of two-third power may be achieved at 100% modulation.

Let m (t) be a band-limited baseband message signal with cutoff frequency W. The DSBSC-AM signal corresponding to m (t) consists of the product of both the message signal m (t) and the carrier signal C (t), as follows:

S (t) =C (t) m (t)

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

This is the same as AM except with the sinusoidal carrier component is eliminated.

The modulated signal S (t) undergoes a phase reversal whenever the message signal m (t) crosses zero. The envelope of a DSB-SC modulated signal is different from the message signal. The transmission bandwidth required by DSB-SC modulation can be seen from figure 2.21 which is same as that for amplitude modulation that is twice the bandwidth of the message signal 2W.

Assume that the message signal is band-limited to the interval $-W \le f \le W$.

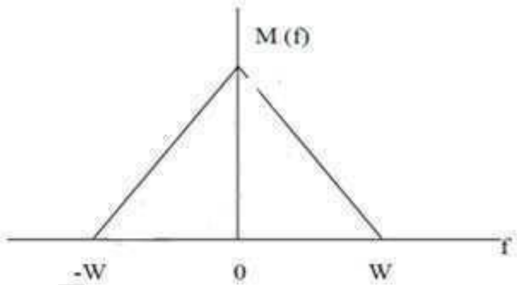


Figure 2.20 Spectrum of message signal

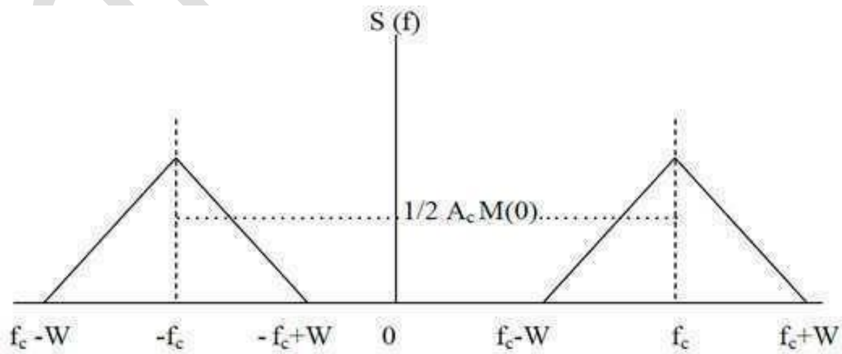


Figure 2.21.Spectrum of DSBSC signal

Single-tone modulation:

In single-tone modulation modulating signal consists of only one frequency component where as in multi-tone modulation modulating signal consists of more than one frequency components.

The standard time domain equation for the DSB-SC modulation is given by

S (t) $=A_c \cos(2\pi f_c t)$ m (t) (1)

m (t) $=A_m \cos(2\pi f_m t)$ (2)

Substitute equation (2) in equation (1) we will get

$S(t) = A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m t)$

$S(t) = \frac{A_c A_m}{2} [\cos 2\pi (f_c - f_m) t + \cos 2\pi (f_c + f_m) t]$ (3)

The Fourier transform of S (t) is

$S(f) = \frac{A_c A_m}{4} [\delta (f - f_c - f_m) + \delta (f + f_c + f_m)] + \frac{A_c A_m}{4} [\delta (f - f_c + f_m) + \delta (f + f_c + f_m)]$

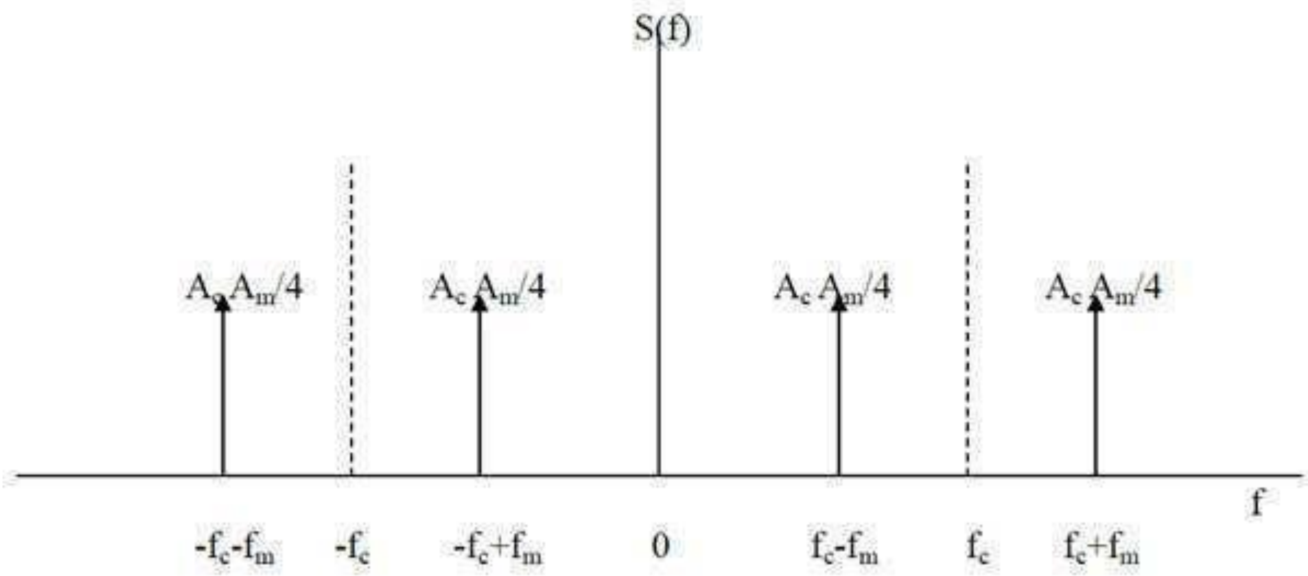


Figure 2.22.Spectrum of single tone DSBSC

Bandwidth:

The DSBSC modulated wave has only two frequencies. So, the maximum and minimum frequencies are f_c+f_m and f_c-f_m respectively.

$f_{max}=f_c+f_m$ and $f_{min}=f_c-f_m$

Substitute, f_{max} and f_{min} values in the bandwidth formula.

$BW=f_c+f_m-(f_c-f_m)$

$BW=2f_m$

Power calculations of DSB-SC waves:-

Consider the following equation of DSBSC modulated wave

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

Power of DSBSC wave is equal to the sum of powers of upper sideband and lower sideband frequency components.

$P_T=P_{USB}+P_{LSB}$

We know the standard formula for power of cosine signal is

$P = \frac{{V_{rms}}^2}{R}$

Average power delivered to a 1ohm resistor can be calculated as,

$P_{USB} = \left(\frac{A_m A_c}{2\sqrt{2}}\right)^2$

$P_{USB} = A_m^2 A_c^2/8$

Similarly; $P_{LSB} = \left(\frac{A_m A_c}{2\sqrt{2}}\right)^2 = A_m^2 A_c^2/8$

So total power $P_T = A_c^2 A_m^2/4$

$\frac{P_{USB}}{P_T} = \frac{P_{LSB}}{P_T} = \frac{A_m^2 A_c^2/8}{A_m^2 A_c^2/4}$ x 100 % = 50%

For the sinusoidal modulation, the average power in the lower or upper side-frequency with respect to the total power in the DSB-SC modulated wave is 50%.

Generation of DSB-SC waves:

The generation of a DSB-SC modulated wave consists simply of the product of the message signal m(t) and the carrier wave $A_c \cos(2\pi f_c t)$. Devices for achieving this requirement is called a product modulator. There are two methods to generate DSB-SC waves. They are:

- Balanced modulator
- Ring modulator

Balanced Modulator:

1. Balanced modulator consists of two identical AM modulators which are arranged in a balanced configuration in order to suppress the carrier signal. Hence, it is called as balanced modulator as shown in figure 4.4.
2. Assume that two AM modulators are identical, except for the sign reversal of the modulating signal applied to the input of one of the modulators.
3. The same carrier signal C (t) = A_c Cos(2πfct) is applied as one of the inputs to these two AM modulators.
4. The modulating signal m(t) is applied as another input to the upper AM modulator. Whereas, the modulating signal with opposite polarity, −m(t) is applied as another input to the lower AM modulator.

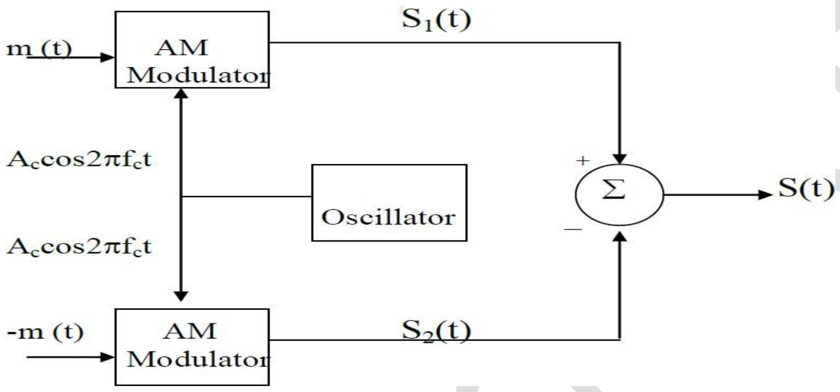


Figure 2.23.Balanced modulator

Mathematical analysis:

The outputs of the two AM modulators can be expressed as follows:

$S_1(t) = A_c[1+k_a m(t)] \cos 2\pi f_c t$

$S_2(t) = A_c[1- k_a m(t)] \cos 2\pi f_c t$

Subtracting $S_2(t)$ from $S_1(t)$, we obtain

$S(t) = S_1(t) – S_2(t)$

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

Hence, except for the scaling factor $2k_a$ the balanced modulator output is equal to product of the modulating signal and the carrier signal. The Fourier transform of S (t) is

$S(f) = k_a A_c [M(f-f_c) + M(f+f_c)]$

Assume that the message signal is band-limited to the interval $–W \le f \le W$ as shown in figure 2.24 and its DSB-SC modulated spectrum is shown in figure 4.6.

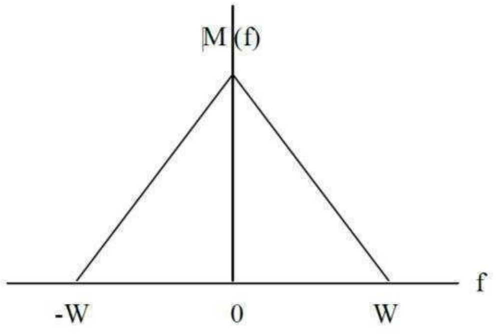


Figure 2.24.Spectrum of Baseband signal

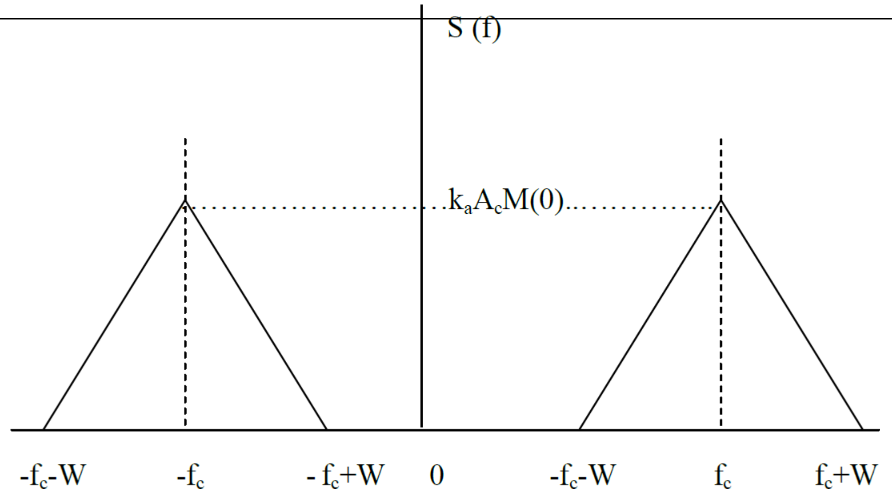


Figure 2.25.Spectrum of DSBSC wave

Ring modulator:

One of the most useful product modulator, for generating a DSBSC wave, is the ring modulator shown in figure 2.26.

1. In this diagram, the four diodes D1,D2,D3 and D4 are connected in the ring structure. Hence, this modulator is called as the ring modulator.
2. The diodes are controlled by a square-wave carrier C (t) of frequency f_c, which applied longitudinally by means of to center-tapped transformers. If the transformers are perfectly balanced and the diodes are identical, there is no leakage of the modulation frequency into the modulator output.
3. The message signal m(t) is applied to the input transformer. Whereas, the carrier signals C (t) is applied between the two centre-tapped transformers.
4. For positive half cycle of the carrier signal, the diodes D1 and D3 are switched ON and the other two diodes D2 and D4 are switched OFF. In this case, the message signal is multiplied by +1.
5. For negative half cycle of the carrier signal, the diodes D2 and D4 are switched ON and the other two diodes D1 and D3 are switched OFF. In this case, the message signal is multiplied by -1. This results in 180^0 phase shift in the resulting DSBSC wave.

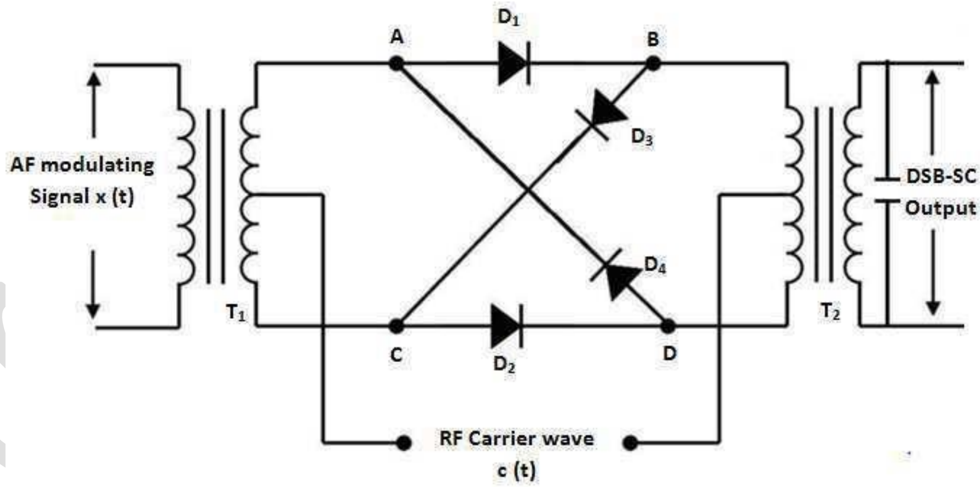


Figure 2.26.Ring modulator

Mathematical Analysis:

The square wave carrier c (t) can be represented by a Fourier series as follows:

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

= 4/π Cos(2πfct) + higher order harmonics(n=1)

Now, the Ring modulator output is the product of both message signal m (t) and carrier signal c (t).

S (t) =c (t) m (t)

$$S(t) == \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos 2\pi f_c t(2n-1)\, m(t) \qquad \text{For } n=1$$

$S(t) = 4/\pi \cos(2\pi f_c t)\, m(t)$

There is no output from the modulator at the carrier frequency i.e the modulator output consists of modulation products. The ring modulator is also called as a double-balanced modulator, because it is balanced with respect to both the message signal and the square wave carrier signal.

The Fourier transform of S (t) is

$S(f) = 2/\pi\, [M(f-fc) + M(f+fc)]$

Assume that the message signal is band-limited to the interval $-W \le f \le W$ as shown in figure 2.27 and its DSB-SC modulated spectrum in figure 2.28.

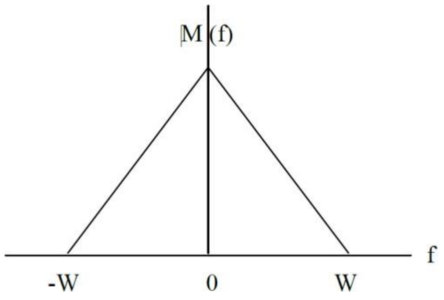


Figure 2.27. Spectrum of Baseband signal

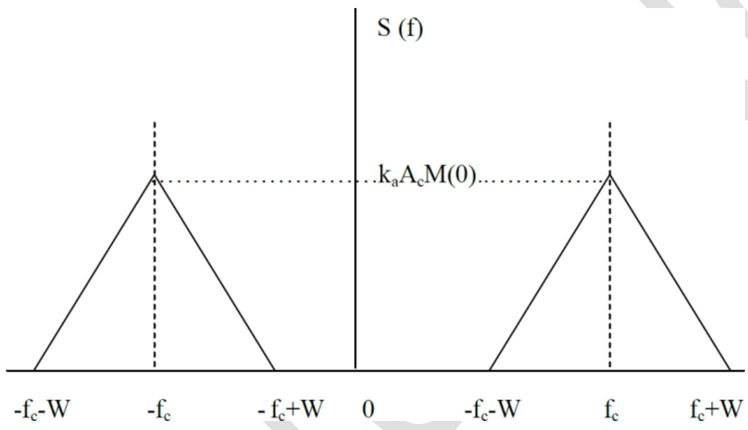


Figure 2.28.Spectrum of DSBSC wave

Coherent Detection of DSB-SC Waves:

The base band signal can be recovered from a DSB-SC signal by multiplying DSB-SC wave S (t) with a locally generated sinusoidal signal and then low pass filtering the product. It is assumed that local oscillator signal is coherent or synchronized, in both frequency and phase, with the carrier signal C (t) used in the product modulator to generate S (t). This method of demodulation is known as coherent detection or synchronous demodulation.

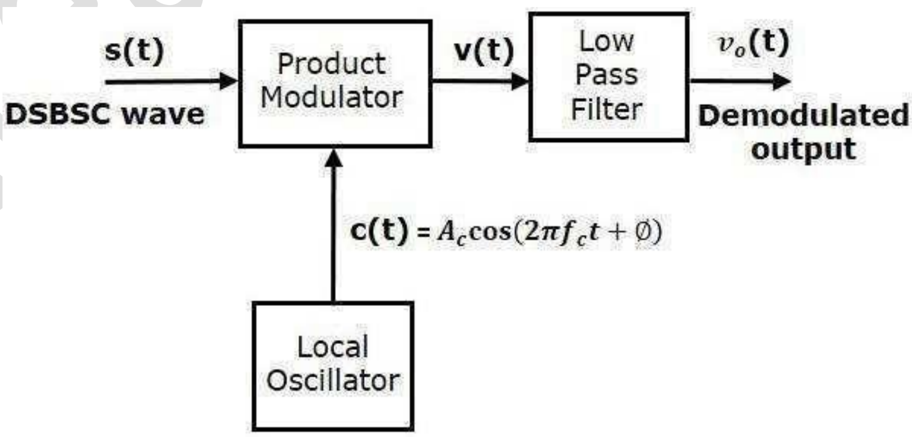


Figure 2.29.Coherent detection of DSB-SC signal

Analysis of coherent detection:

The product modulator produces the product of both input signal s(t) and local oscillator signal and the output of the product modulator is v (t).

$S(t) = A_c \cos(2\pi f c t)\, m(t)$

$C(t) = A_c \cos(2\pi f c t + \emptyset)$

$V(t) = C(t)\, S(t)$

$V(t) = A_c \cos(2\pi f_c t + \emptyset)\, S(t)$

$V(t) = A_c \cos(2\pi f_c t + \emptyset)\, A_c \cos(2\pi f_c t)\, m(t)$

$V(t) = A_c^2 \cos(2\pi f_c t + \emptyset) \cos(2\pi f_c t)\, m(t)$

$V(t) = \frac{A_c^2}{2} \cos\emptyset\ m(t) + \frac{A_c^2}{2} \cos(4\pi f_c t + \emptyset)\, m(t)$

In the above equation, the first term is the scaled version of the message signal. It can be extracted by passing the above signal through a low pass filter. Therefore, the output of low pass filter is

$V_o(t) = \frac{A_c^2}{2} \cos\emptyset\ m(t)$

The Fourier transform of $V_o(t)$ is

$V_O(f) = \frac{A_c^2}{2} \cos\emptyset\ M(f)$

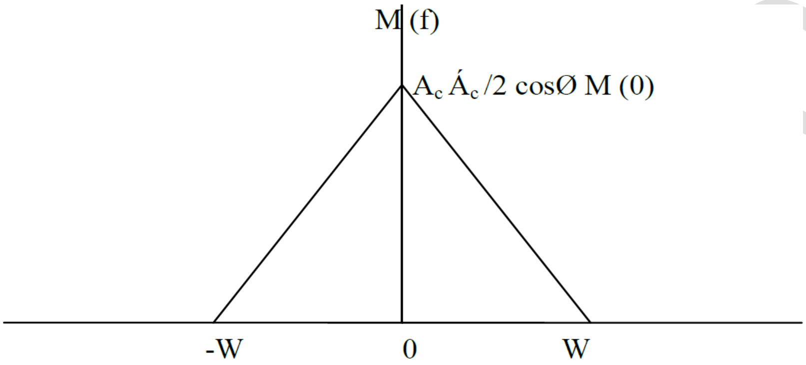


Figure 2.30.DSB-SC demodulated output

The demodulated signal is proportional to the message signal m (t) when the phase error is constant. The amplitude of this demodulated signal is maximum when $\emptyset=0$, the local oscillator signal and the carrier signal should be in phase, i.e., there should not be any phase difference between these two signals. The demodulated signal amplitude will be zero, when $\emptyset=\pm\pi/2$. This effect is called as **quadrature null effect**.

Costa's loop detection:

1. The receiver consists of two coherent detectors supplied with same DSB-SC wave while the other input for both product modulators is taken from Voltage Controlled Oscillator (VCO) with -90^0 phase shift to one of the product modulator as shown in figure 2.29.
2. The frequency of the local oscillator is adjusted to be the same as the carrier frequency f_c. The 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.
3. 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.
4. The output of product modulator is applied as an input of the lower low pass filter.
5. The output of lower Low pass filter has -90^0 phase difference with the output of the upper low pass filter. The outputs of these two low pass filters are applied as inputs of the phase discriminator. Based on the phase difference between these two signals, the phase discriminator produces a DC control signal.
6. This signal is applied as an input of VCO to correct the phase error in VCO output. Therefore, the carrier signal (used for DSBSC modulation) and the locally generated signal (VCO output) are in phase.

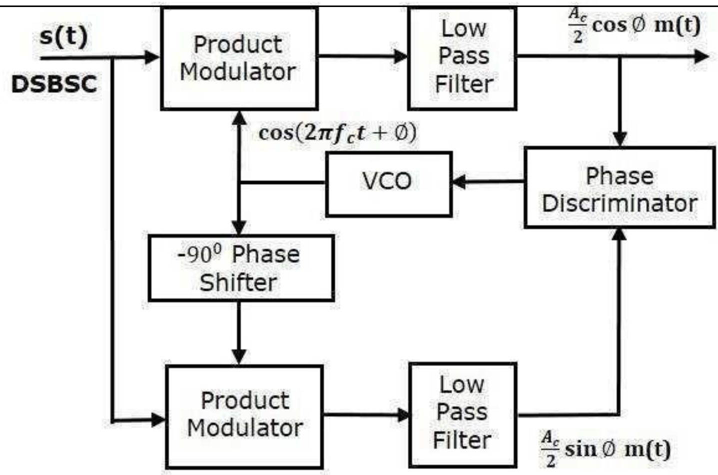


Figure 2.31.Costa's receiver

Mathematical Analysis: We know that the equation of DSBSC wave is

$S(t) = A_c \mathrm{Cos}(2\pi f_c t) m(t)$

Let the output of VCO be $c_1(t) = \mathrm{Cos}(2\pi f_c t + \phi)$

This output of VCO is applied as the carrier input of the upper product modulator. Hence, the output of the upper product modulator is

$v_1(t) = S(t)\, c_1(t)$

Substitute, S(t) and $c_1(t)$ values in the above equation.

$v_1(t) = A_c \mathrm{Cos}(2\pi f_c t)\, m(t)\, \mathrm{Cos}(2\pi f_c t + \phi)$

$v_1(t) = \frac{A_c}{2} \mathrm{Cos}\,\phi\; m(t) + \frac{A_c}{2} \mathrm{Cos}(4\pi f_c t + \phi)\, m(t)$

This signal is applied as an input of the upper low pass filter. The output of this low pass filter is

$v_{01}(t) = A_c^2 \cos\phi\; m(t)$

Therefore, the output of this low pass filter is the scaled version of the modulating signal.The output of -90^0 phase shifter is

$c_2(t) = \mathrm{Cos}(2\pi f_c t + \phi - 90^0) = \sin(2\pi f_c t + \phi)$

This signal is applied as the carrier input of the lower product modulator. The output of the lower product modulator is

$v_2(t) = S(t)\, c_2(t)$

Substitute, S(t) and $c_2(t)$ values in the above equation.

$v_2(t) = A_c \mathrm{Cos}(2\pi f_c t) m(t) \sin(2\pi f_c t + \phi)$

After simplifying, we will get $v_2(t)$ as

$v_2(t) = A_c^2 \sin\phi\, m(t) + A_c^2 \sin(4\pi f c t + \phi) m(t)$

This signal is applied as an input of the lower low pass filter. The output of this low pass filter is

$v_{02}(t) = A_c^2 \sin\phi\; m(t)$

The output of this Low pass filter has -90^0 phase difference with the output of the upper low pass filter.

2.13. Single Sideband Modulation

Single sideband modulation (SSB) is a form of amplitude modulation which uses only one sideband for a given message signal to provide the final signal. The process of suppressing one of the sidebands along with the carrier and transmitting a single sideband is called as Single Sideband Suppressed Carrier system or simply SSBSC.

SSB provides a considerably more efficient form of communication when compared to ordinary amplitude modulation in terms of the radio spectrum used a can be seen from figure 2.30, and also the power used to transmit the signal.

Depending on which half of DSB-SC signal is transmitted, there are two types of SSB modulation

1. Lower Side Band (LSB) Modulation
2. Upper Side Band (USB) Modulation

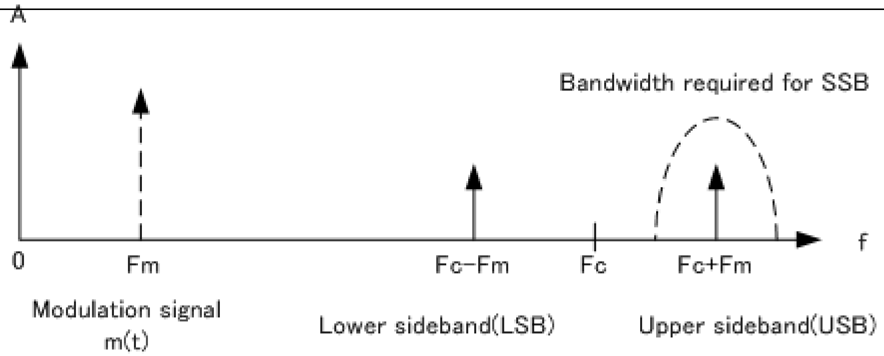


Figure 2.32.SSB-SC spectrum

Mathematical Expressions

Let us consider the mathematical expressions for the modulating and the carrier signals as follows

Modulating signal $m(t) = A_m \cos(2\pi f_m t)$

Carrier signal $c(t) = A_c \cos(2\pi f_c t)$

Mathematically, we can represent the equation of SSBSC wave as

$S(t) = \frac{mc}{2} \cos[2\pi(f_c + f_m)t]$ for the upper sideband

Or

$S(t) = \frac{mc}{2} \cos[2\pi(f_c - f_m)t]$ for the lower sideband

Bandwidth of SSBSC Wave

As can be seen in figure 2.33, the DSBSC modulated wave contains two sidebands and its bandwidth is $2f_m$. Since the SSBSC modulated wave contains only one sideband, its bandwidth is half of the bandwidth of DSBSC modulated wave. Therefore, the bandwidth of SSBSC modulated wave is f_m and it is equal to the frequency of the modulating signal.

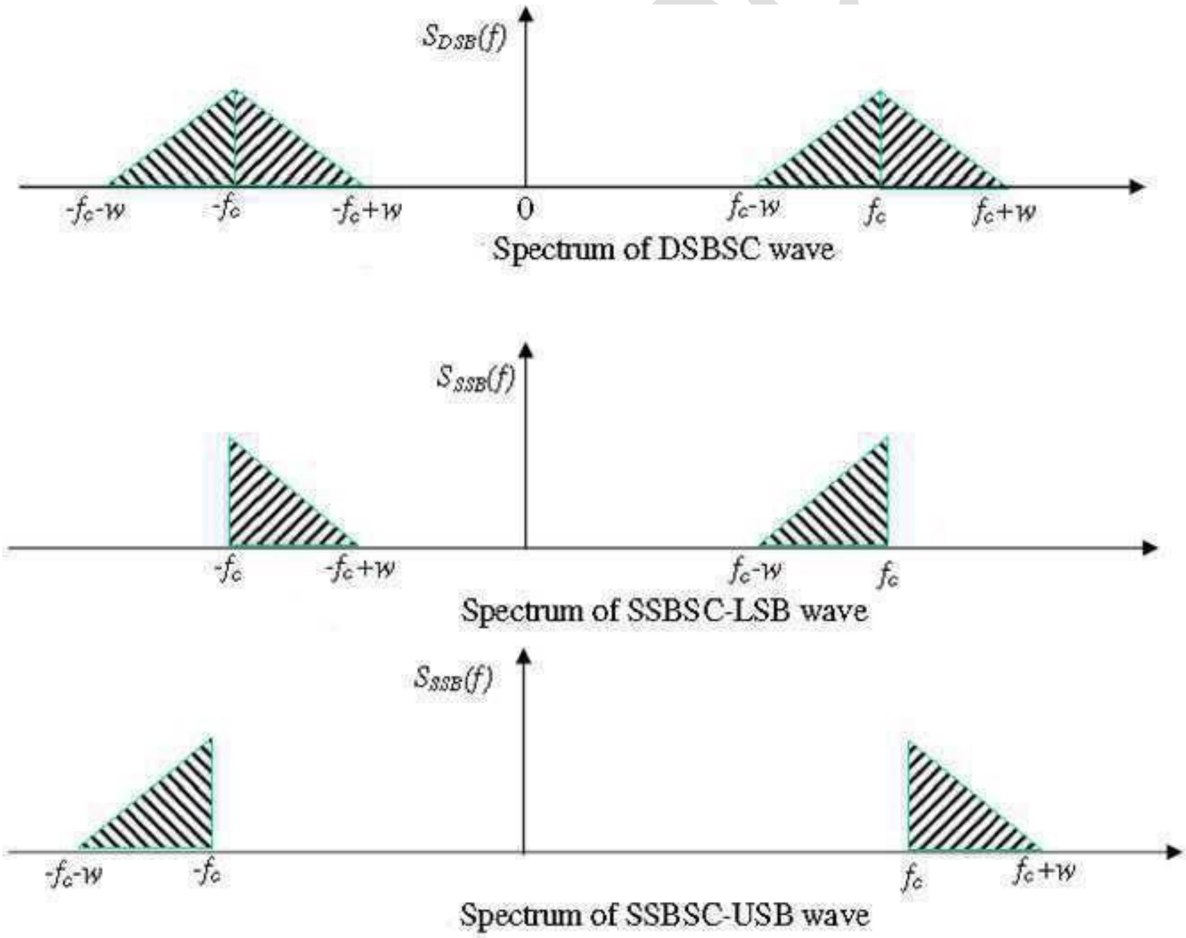


Figure 2.33 Spectrums of DSBSC and SSBSC

Power Calculations of SSBSC signal:

Consider the following equation of SSBSC modulated wave.

$s(t)=\frac{mc}{2}\cos[2\pi(f_c+f_m)t]$ for the upper sideband

Or

$s(t)=\frac{mc}{2}\cos[2\pi(f_c-f_m)t]$ for the lower sideband

Power of SSBSC wave is equal to the power of any one sideband frequency components.

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

We know that the standard formula for power of cosine signal is

$$P_{rms}=\frac{V^2}{R}=\frac{\left(\frac{V_m}{\sqrt{2}}\right)^2}{R}$$

In this case, the power of the upper sideband is

$$P_{USB}=\frac{(mc)^2}{8R}$$

Similarly, we will get the lower sideband power same as that of the upper side band power.

$$P_{LSB}=\frac{(mc)^2}{8R}$$

Therefore, the power of SSBSC wave for 1 ohm resistance is

$$P_t = P_{USB} = P_{LSB} = \frac{(mc)^2}{8}$$

Advantages

- Bandwidth or spectrum space occupied is lesser than AM and DSBSC waves.
- Transmission of more number of signals is allowed.
- Power is saved.
- High power signal can be transmitted.
- Less amount of noise is present.
- Signal fading is less likely to occur.

Disadvantages

- The generation and detection of SSBSC wave is a complex process.
- The quality of the signal gets affected unless the SSB transmitter and receiver have an excellent frequency stability.

Applications

- For power saving requirements and low bandwidth requirements.
- In land, air, and maritime mobile communications.
- In point-to-point communications.
- In radio communications.
- In television, telemetry, and radar communications.
- In military communications, such as amateur radio, etc.

2.14. Generation of SSB waves:

1. Frequency Discrimination Method

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 2.34. Application of this method requires that the message signal satisfies two conditions:

1. The message signal m(t) has low or no low-frequency content. M(ω) has a “hole” at zero-frequency Example: - speech, audio, music.

2. The highest frequency component W of the message signal m(t) is much less than the carrier frequency. 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.

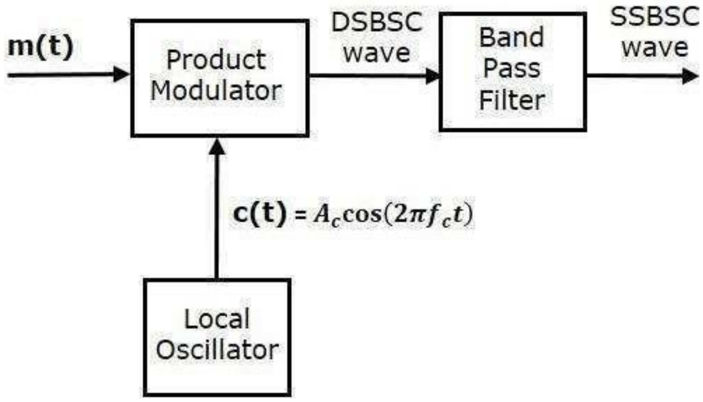


Figure 2.34.Filter method

2. Phase discrimination method

1. The phase discriminator consists of two product modulators I and Q, supplied with carrier waves in-phase quadrature to each other as shown in figure 2.35.
2. 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.
3. The Hilbert transform m^(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.
4. 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.

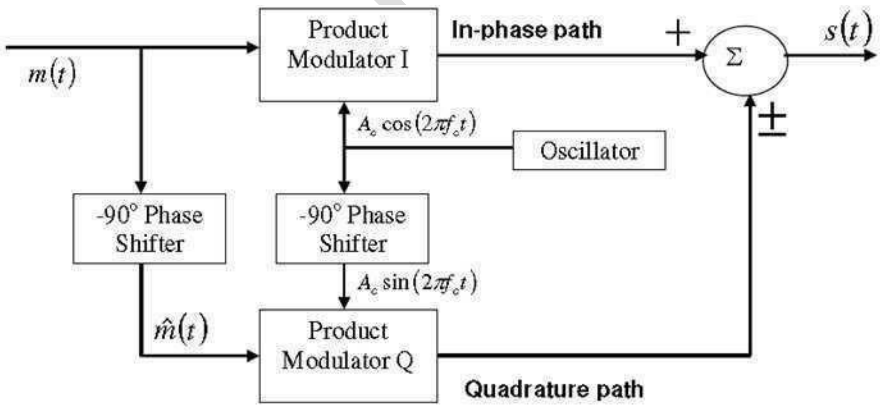


Figure 2.35.Phase discrimination method

Demodulation of SSB waves:

Coherent detection: It assumes perfect synchronization between the local carrier and that used in the transmitter both in frequency and phase. The carrier signal which is used for generating SSBSC wave is used to detect the message signal. Hence, this process of detection is called as coherent or synchronous detection. Following is the block diagram of coherent detector.

In this process, the message signal can be extracted from SSBSC wave by multiplying it with a coherent carrier and then the resulting signal is passed through a Low Pass Filter. The output of this filter is the desired message signal.

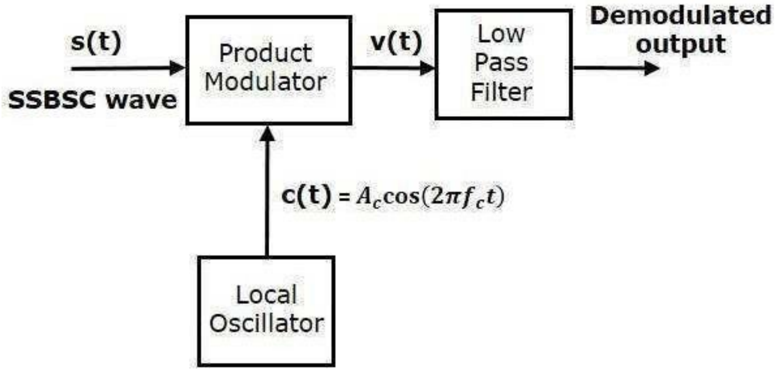


Figure 2.36.Coherent detection

Mathematical Analysis:

$S(t) = A_m A_c/2 \cos[2\pi(f_c - f_m)t]$

The output of the local oscillator is

$c(t) = A_c \cos(2\pi f_c t)$

From the figure, we can write the output of product modulator as

$v(t) = s(t)c(t)$

Substitute s(t) and c(t) values in the above equation

$$V(t) = \frac{A_m A_c}{2}\cos[2\pi(f_c + f_m)t]\; A_c\cos(2\pi f_c t)$$

$$V(t) = \frac{A_m A_c^2}{4}\cos(2\pi f_m t) + \frac{A_m A_c^2}{4}\cos[2\pi(2f_c - f_m)t]$$

In the above equation, the first term is the scaled version of the message signal the scaling factor is $\frac{A_c^2}{4}$. It can be extracted by passing the above signal through a low pass filter.

Therefore, the output of low pass filter is

$$V_0(t) = \frac{A_m A_c^2}{4}\cos(2\pi f_m t)$$

2.15. 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. The word “vestige” means “a part” from which, the name is derived.

VSBSC Modulation is the process, where a part of the signal called as vestige is modulated along with one sideband. The frequency spectrum of VSBSC wave is shown in the figure 2.37. Along with the upper sideband, a part of the lower sideband is also being transmitted in this technique. Similarly, we can transmit the lower sideband along with a part of the upper sideband.

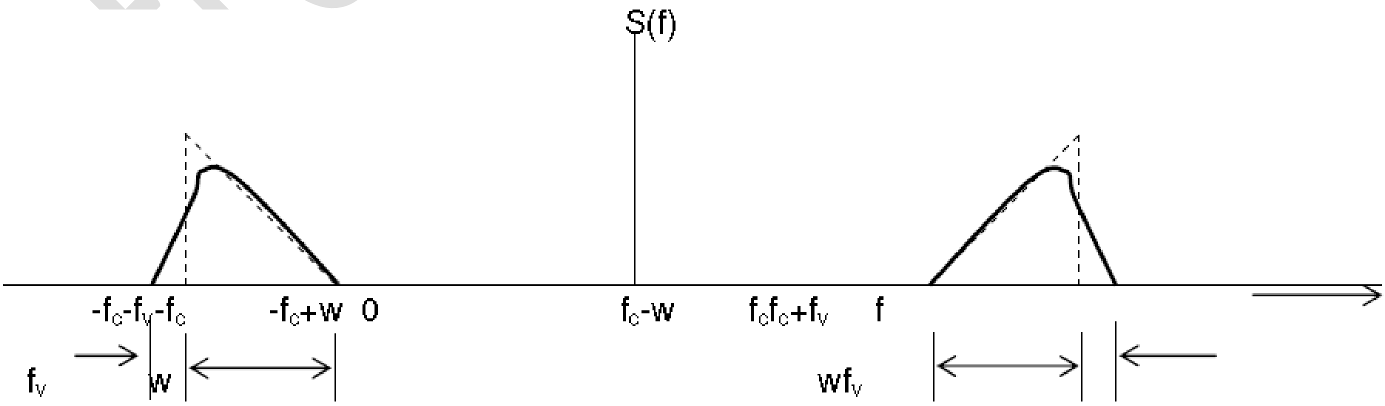


Figure 2.37.Spectrum of VSB containing vestige of USB

The vestige of the Upper sideband compensates for the amount removed from the Lower sideband. The bandwidth required to send VSB wave is

$B = w + f_v$

Where f_v is the width of the vestigial side band.

Therefore, VSB has the virtue of conserving bandwidth almost as efficiently as SSB modulation, while retaining the excellent low-frequency base band characteristics of DSBSC and it is standard for the transmission of TV signals.

Generation of VSB Modulated wave:

To generate a VSB modulated wave, we pass a DSBSC modulated wave through a sideband-shaping filter. The modulating signal m(t) is applied to a product modulator. The output of the local oscillator is also applied to the other input of the product modulator.

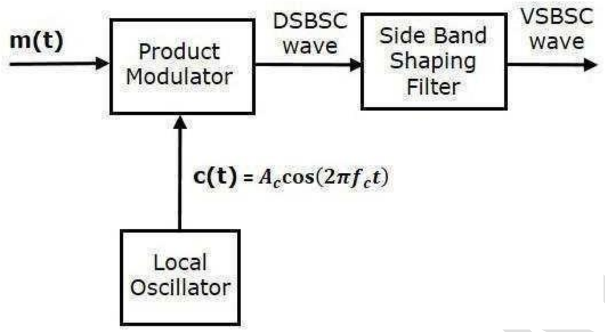


Figure 2.38.VSB modulator

Mathematical Analysis:

The output of the product modulator is then given by :

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

Apply Fourier transform on both sides

$P(f) = A_c/2[M(f-f_c)+M(f+f_c)]$

The above equation represents the equation of DSBSC frequency spectrum.

Let the transfer function of the sideband shaping filter be H(f). This filter has the input p(t) and the output is VSBSC modulated wave S(t).The Fourier transforms of p(t) and S(t) are P(f) and S(f) respectively.

$S(f)=P(f)H(f)$

Substitute P(f) in the above equation.

$S(f)=A_c/2[M(f-f_c)+M(f+f_c)]H(f)$

The above equation represents the equation of VSBSC frequency spectrum.

Demodulation of VSBSC

Demodulation of VSBSC wave is similar to the demodulation of SSBSC wave. Here, the same carrier signal which is used for generating VSBSC wave is used to detect the message signal. Hence, this process of detection is called as coherent or synchronous detection. The VSBSC demodulator is shown in the figure 2.39.In this process, the message signal can be extracted from VSBSC wave by multiplying it with a carrier, which is having the same frequency and the phase of the carrier used in VSBSC modulation. The resulting signal is then passed through a Low Pass Filter. The output of this filter is the desired message signal.

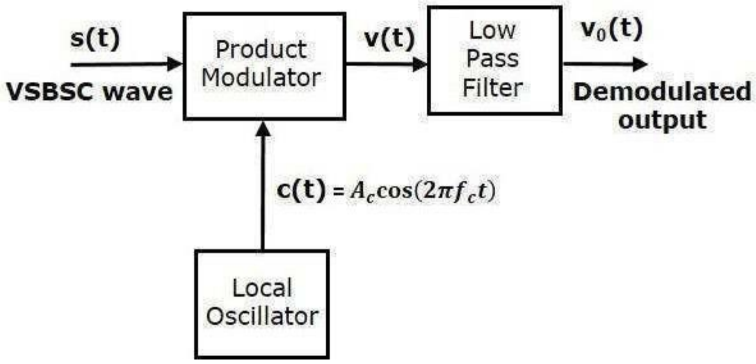


Figure 2.39.Demodulation of VSB-SC signal

Advantages of VSB

1. The main advantage of VSB modulation is the reduction in bandwidth. It is almost as efficient as the SSB.
2. Due to allowance of transmitting a part of lower sideband, the constraint on the filter has been relaxed. So practically, easy to design filters can be used.
3. It possesses good phase characteristics and makes the transmission of low frequency components possible.

Application of VSB

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

2.16. Comparison of amplitude modulation techniques:

- In commercial AM radio broadcast systems standard AM is used in preference to DSBSC or SSB modulation.
- Suppressed carrier modulation systems require the minimum transmitter power and minimum transmission bandwidth. Suppressed carrier systems are well suited for point –to-point communications.
- SSB is the preferred method of modulation for long-distance transmission of voice signals over metallic circuits, because it permits longer spacing between the repeaters.
- VSB modulation requires a transmission bandwidth that is intermediate between that required for SSB or DSBSC.
- DSBSC, SSB, and VSB are examples of linear modulation. In Commercial TV broadcasting; the VSB occupies a width of about 1.25MHz, or about one-quarter of a full sideband.
- In standard AM systems the sidebands are transmitted in full, accompanied by the carrier. Accordingly, demodulation is accomplished by using an envelope detector or square law detector. On the other hand in a suppressed carrier system the receiver is more complex because additional circuitry must be provided for purpose of carrier recovery.
- Suppressed carrier systems require less power to transmit as compared to AM systems thus making them less expensive.
- SSB modulation requires minimum transmitter power and maximum transmission band with for conveying a signal from one point to other thus SSB modulation is preferred.
- VSB modulation requires a transmission band width that is intermediate of SSB or DSBSC.
- In SSB and VSB modulation schemes the quadrature component is only to interfere with the in phase component so that power can be eliminated in one of the sidebands.

Parameter of comparison	**AM**	**DSB-SC**	**SSB-SC**	**VSB**
Carrier suppression	NA	Fully	Fully	NA
Sideband suppression	NA	NA	One sideband completely	One sideband suppressed partially
Bandwidth	$2f_m$	$2f_m$	f_m	f_m <BW>$2f_m$
Transmission efficiency	Minimum	Moderate	Maximum	moderate
Power requirement	More power is required for transmission	Power required is less than AM	Power required is less than Am and DSB-SC	Power required is less than DSB-SC but more than SSB-SC
Power saving (%)	0	66.67	83.33	Lies between DSB and SSB
Applications	Radio broadcasting	Radio broadcasting	Point to point mobile communication	TV

2.17. Demodulation of AM waves:

There are two methods to demodulate AM signals. They are:

1. Square-law detector
2. Envelope detector

Square-law detector:

Square-law detector is used to detect low level modulated signals (below 1v). A Square-law detector requires nonlinear element and a low pass filter for extracting the desired message signal. Semi-conductor diodes and transistors are the most common nonlinear devices used for implementing square law detectors as shown in figure 2.40. The filtering requirement is usually satisfied by using a single or double tuned filters.

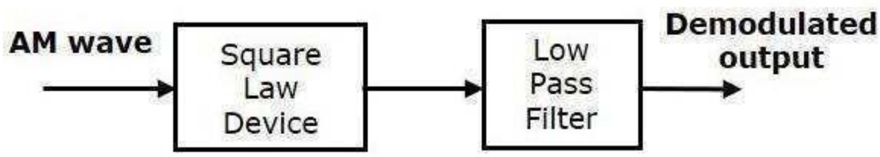


Figure 2.40.Square law detector

When a nonlinear element is suitably biased and operated in a restricted portion of its characteristic curve, we find that transfer characteristic of diode-load resistor combination can be represented closely by a square law :

$V_0(t) = a_1V_i(t) + a_2 V_i^2(t)$ (4)

Where a_1, a_2 are constants

Now, the input voltage $V_i(t)$ is the sum of both carrier and message signals

$V_i(t) = A_c [1+k_a m(t)] \cos 2\pi f_c t$ (5)

Substitute equation (5) in equation (4) we get

$V_0(t) = a_1A_c [1+k_am(t)] \cos 2\pi f_ct + 1/2\, a_2A_c^2 [1+2 k_a m(t) + k_a^2m^2(t)] [\cos 4\pi f_ct]$ (6)

Now design the low pass filter with cutoff frequency f is equal to the required message signal bandwidth. We can remove the unwanted terms by passing this output voltage $V_0(t)$ through the low pass filter and finally we will get required message signal.

$V_0(t) = A_c^2 a_2 m(t)$

The Fourier transform of output voltage $V_O(t)$ is given by

$V_O(f) = Ac^2 a_2 M(f)$

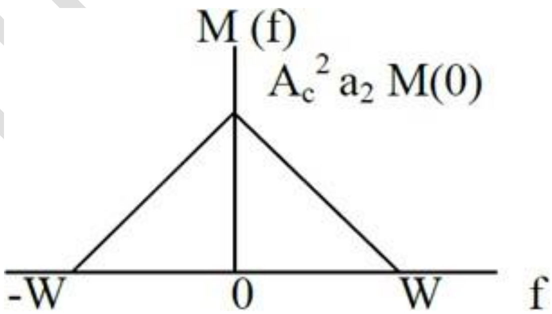


Figure 2.41.Spectrum of output signal

Envelope Detector:

Envelope detector is used to detect (demodulate) high level AM wave. Following figure 2.42 is the block diagram of the envelope detector. It is also based on the switching action or switching characteristics of a diode. It consists of a diode and a resistor-capacitor filter.

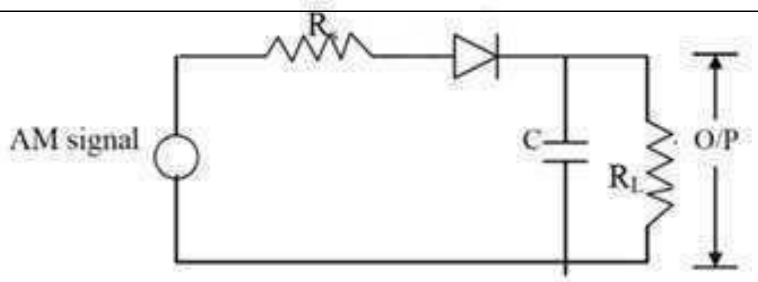


Figure 2.42.Envelope detector

The operation of the envelope detector is as follows.

1. On a positive half cycle of the AM signal, the diode is forward biased and the capacitor C charges up rapidly to the peak value of the input signal.
2. When the AM signal level falls below this value, the diode becomes reverse biased and the capacitor C discharges slowly through the load resistor R_L till the next positive cycle of AM signal.
3. When the input signal becomes greater than the voltage across the capacitor, the diode conducts again and the process is repeated.
4. The component values should be selected in such a way that the capacitor charges very quickly and discharges very slowly. As a result, we will get the capacitor voltage waveform same as that of the envelope of AM wave as shown in figure 2.43.

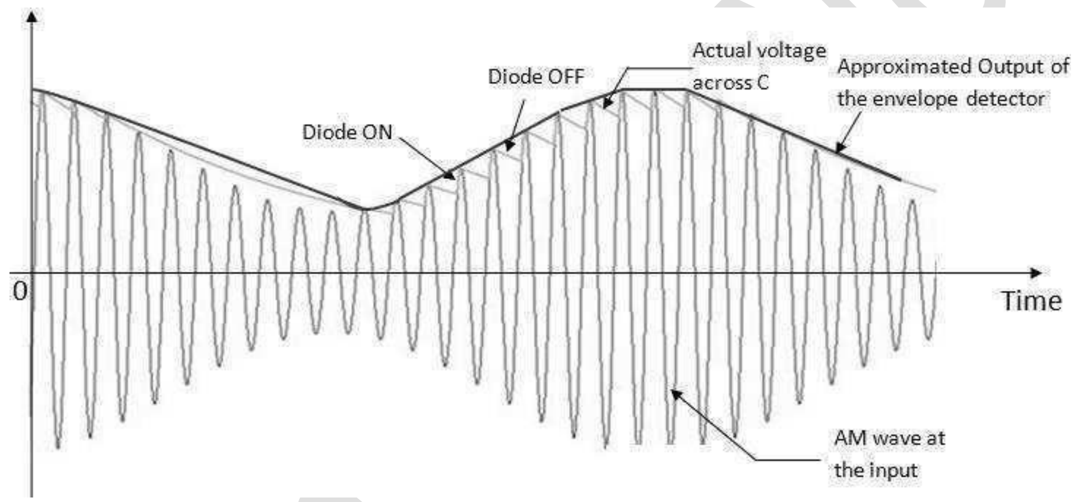


Figure 2.43.Input-output waveform for envelope detector

The charging time constant R_s C is very small when compared to the carrier period $1/f_c$, the capacitor C charges rapidly to the peak value of the signal.

$R_s C << 1/f_c$

Where R_s = internal resistance of the voltage source, C = capacitor, f_c = carrier frequency

The discharging time constant R_L C is very large when compared to the charging time constant i.e., the capacitor discharges slowly through the load resistor.

i.e., $1/fc << R_L C << 1/W$

Where R_L = load resistance value, W = message signal bandwidth

Distortions in the Envelope Demodulator Output

There are two types of distortions which can occur in the detector output such as:

1. Diagonal clipping
2. Negative peak clipping

Diagonal Clipping: This type of distortion occurs when the RC time constant of the load circuit is too long. Due to this, the RC circuit cannot follow the fast changes in the modulating envelope.

Negative peak clopping: This distortion occurs due to a fact that the modulation index on the output side of the detector is higher than that on its input side. Hence, at higher depth of modulation of the transmitted signal, the over-modulation may takes place at the output of the detector. The negative peak clipping will take place as a result of this over-modulation as shown in figure 2.44.

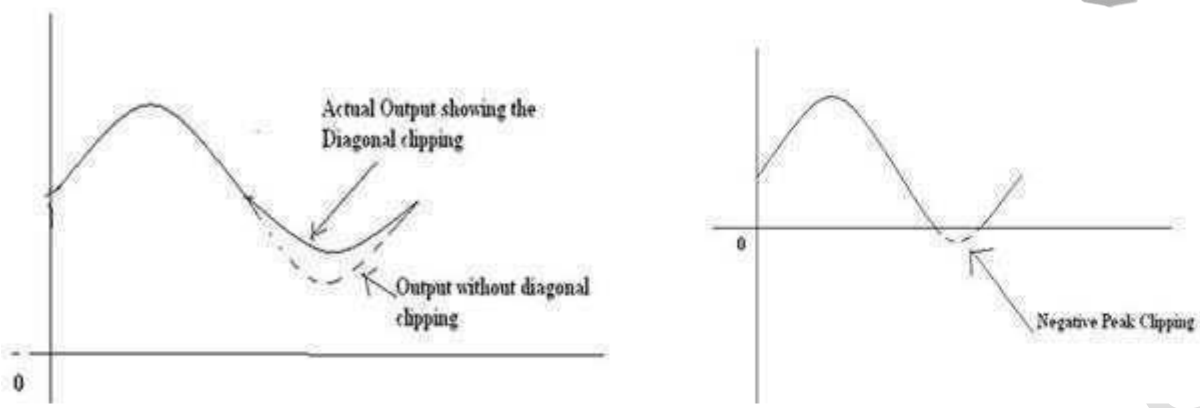


Figure 2.44.Ditortion in output of envelope detector

4.7. Low and high power AM transmitters:

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

1. High Level
2. Low Level

The basic difference between the two transmitters is the power amplification of the carrier and modulating signals. High level transmitters use high level modulation, and low level transmitters use low level modulation. In broadcast transmitters, where the transmitting power may be of the order of kilowatts, high level modulation is employed. In low power transmitters, where only a few watts of transmitting power are required, low level modulation is used.

High-Level Transmitters:

In high-level transmission, the powers of the carrier and modulating signals are amplified before applying them to the modulator stage, as shown in figure 2.45.

1. **Carrier oscillator:** The carrier oscillator generates the carrier signal, which lies in the RF range. The frequency of the carrier is always very high. Because it is very difficult to generate high frequencies with good frequency stability, the carrier oscillator generates a sub multiple with the required carrier frequency. This sub multiple frequency is multiplied by the frequency multiplier stage to get the required carrier frequency.
2. **Buffer Amplifier:** The purpose of the buffer amplifier is to match the output impedance of the carrier oscillator with the input impedance of the frequency multiplier, the next stage of the carrier oscillator. It then isolates the carrier oscillator and frequency multiplier.
3. **Frequency Multiplier:** The sub-multiple frequency of the carrier signal, generated by the carrier oscillator, is now applied to the frequency multiplier through the buffer amplifier. This stage is also known as harmonic generator. The frequency multiplier generates higher harmonics of carrier oscillator frequency.
4. **Power Amplifier:** The power of the carrier signal is then amplified in the power amplifier stage. This is the basic requirement of a high-level transmitter. A class C power amplifier gives high power current pulses of the carrier signal at its output.
5. **Audio Chain:** The audio signal to be transmitted is obtained from the microphone. The audio driver amplifier amplifies the voltage of this signal. This amplification is necessary to drive the audio power amplifier. Next, a class A or a class B power amplifier amplifies the power of the audio signal.
6. **Modulated Class C Amplifier:** This is the output stage of the transmitter. The modulating audio signal and the carrier signal, after power amplification, are applied to this modulating stage. The modulation takes place at this stage. The class C amplifier also amplifies the power of the AM signal to the required transmitting power. This signal is finally passed to the antenna, which radiates the signal into space of transmission.

Low-Level Transmitters: In low-level modulation, the powers of the two input signals of the modulator stage are not amplified. The required transmitting power is obtained from the last stage of the transmitter, the class C power amplifier. The low-level AM transmitter shown in the figure 2.45 is similar to a high-level transmitter, except that the powers of the carrier and audio signals are not amplified. These two signals are directly applied to the modulated class C power amplifier. Modulation takes place at the stage, and the power of the modulated signal is amplified to the required transmitting power level. The transmitting antenna then transmits the signal.

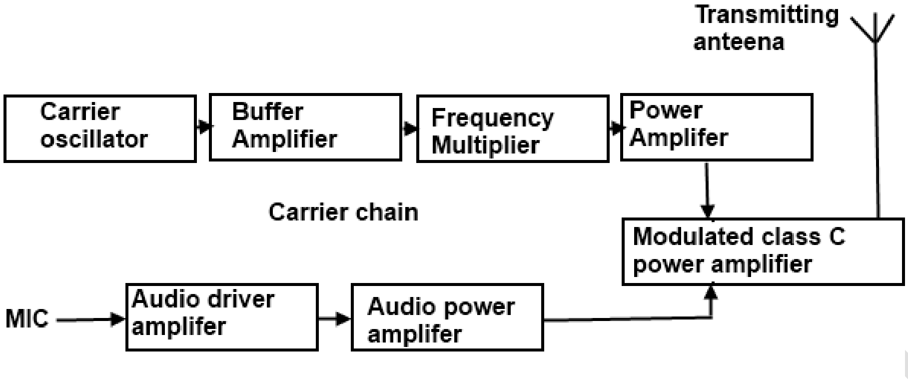


Figure 2.45.High level AM transmitter

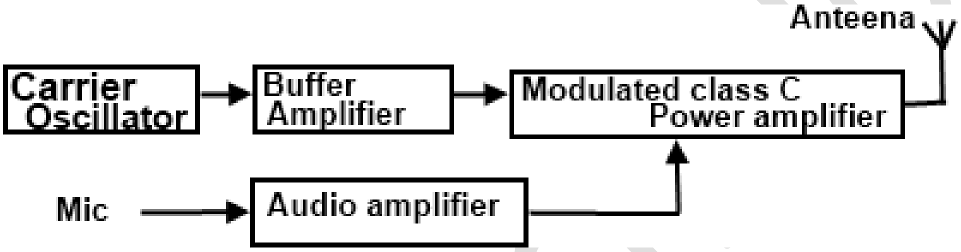


Figure 2.46.Low level AM transmitter

Coupling of Output Stage and Antenna

The output stage of the modulated class C power amplifier feeds the signal to the transmitting antenna. To transfer maximum power from the output stage to the antenna it is necessary that the impedance of the two sections match. For this, a matching network is required. The matching between the two should be perfect at all transmitting frequencies.

2.19. AM Receiver:

Radio receivers amplify and tune the radio signals. The receiver picks up the signals from the airwaves, and converts them to the original message signal. The radio signal that is transmitted into the air contains a carrier wave that is much higher in frequency than message siganl.

2.20. Tuned Radio Frequency Receiver:

A TRF receiver amplifies and tunes the raw radio signal as present in the air waves by means of an RF (radio frequency) amplifier. Some receivers will have as many 4 or 5 stages of RF amplification before the carrier signal is stripped away leaving only the audio portion of the signal. The process of removing the carrier signal is done by the detector circuit of a radio receiver. Afterwards the final process is amplifying the audio signal to a level strong enough to drive a speaker.

Typically a TRF receiver would consist of three main sections:

- **Tuned radio frequency stages:** The tuner circuit is an LC circuit, which is also called as resonant or tank circuit. It selects the frequency, desired by the AM receiver. This consisted of one or more amplifying and tuning stages. In a TRF receiver a series of loosely coupled tuned circuits are used to increase selectivity.
- **Signal detector:** The detector enabled the audio from the amplitude modulation signal to be extracted. It uses envelope detection.
- **Audio amplifier**: This is the power amplifier stage, which is used to amplify the detected audio signal. The processed signal is strengthened to be effective. This signal is passed on to the loudspeaker to get the original sound signal.

Drawbacks :

1. Instability
2. Poor selectivity at high frequencies
3. Bandwidth variation over the tuning range
4. Insufficient adjacent frequency rejection
5. In TRF receiver, amplification is not constant over the tuning range.

2.21. Superheterodyne AM Receiver:

In super heterodyne receiver the incoming RF signal is combined with local oscillator signal frequency through a mixer and converted into signal of lower fixed frequency known as intermediate frequency. It consists of RF section, frequency converter, IF amplifier, detector, audio amplifier.

RF section:

- RF section mainly consists of a tuneable filter and an amplifier which picks up the desired station by tuning the filter to the exact frequency band.
- The signal at the antenna has lower signal noise found anywhere in the receiver.
- Then RF amplifier provides gain to increase signal to noise ratio (SNR).

Frequency converter:

- It converts the carrier frequency f_c to a fixed IF frequency of 455 KHz.
- A constant frequency difference should be maintained between the local oscillator signal and incoming RF signal frequency. (Through capacitor tuning in which the capacitance are together and operated by a common knob.)
- For this purpose it uses local oscillator whose frequency f_{co} is exactly 455 KHz above the incoming carrier frequency f_c and $f_{co}=f_c+455$

IF amplifier:

- The intermediate frequency generated from the mixer/converter is amplified by IF amplifier. After the IF amplifier the signal is applied at the demodulator which extract the original modulated signal.
- The reason for translating all stations to a fixed carrier frequency of 455 KHz is to obtain adequate selectivity. The characteristics of the IF amplifier are not dependent on the incoming frequency to which the receiver is tuned. The selectivity and sensitivity of super-heterodyne receiver are quite uniform throughout its tuning range.
- The main function of the RF section is image frequency suppression. The mixer or converter output consists of components of difference between the incoming f_c and the local oscillator f_{co}.
- Audio amplifier: Once demodulated, the recovered audio is applied to an audio amplifier block to be amplified by a power amplifier to the required level for loudspeakers or headphones.

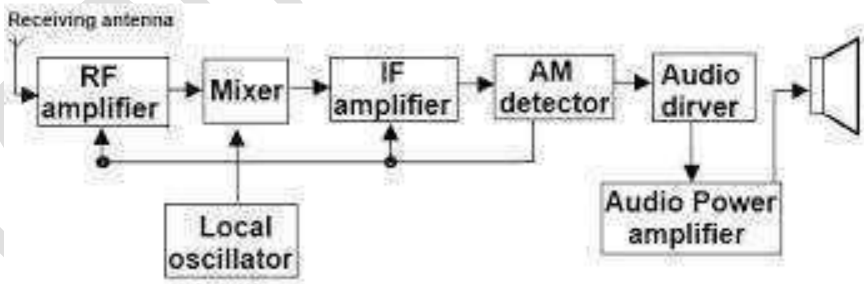


Figure 2.47.Superheterodyne AM receiver

Local oscillator frequency

At design level there are two choices for the local oscillator frequency:

$f_{LO} = f_{RF} + f_{IF}$ (high-side injection) or $f_{LO} = f_{RF} - f_{IF}$ (low-side injection)

Usually for medium wave AM receivers the frequency of the oscillator is higher than the desired RF frequency ($f_{LO} = f_{RF} + f_{IF}$).

Image frequency

When the receiver demodulates the incoming desired signal at f_{RF}, unfortunately it demodulates down to IF also an unwanted signal at $f_{RF}+2f_{IF}$.This frequency is called image frequency

To reduce the design complexity of the receivers the IF frequency is chosen in such a way that the signal at

$f_{image} = f_{RF} + 2f_{IF}$ can be rejected by a simple tuneable RF band pass filter such as a tank circuit with a variable capacitor.

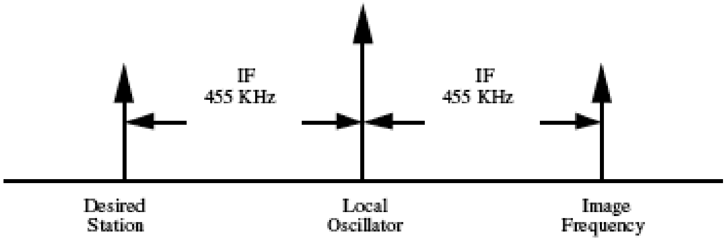


Figure 2.48.Concept of image frequency

2.22. Terminologies of Receiver:

1. Selectivity:

Selectivity is the measure of the ability of a radio receiver to select a particular frequency or particular band of frequencies and rejecting all other unwanted frequencies. The receiver selectivity performance determines the level of interference that may be experienced. It is the ability to reject unwanted signals. The signal bandwidth should be narrow for better selectivity.

The selectivity can be aimed at rejecting signals that may reach the receiver output in a variety of ways.

- **Adjacent channel selectivity:** Adjacent channel selectivity of the form of selectivity that rejects signals on nearby frequencies.
- **Image rejection selectivity:** When using a super heterodyne radio, it is possible for the image frequency to reach the final stages of the receiver. Rejecting these signals is important as they can cause significant levels of interference. The selectivity required to remove these signals is contained within the radio frequency stages of the radio.
- Image frequency rejection ratio is the ratio of gain at the signal frequency to the gain at the image frequency.Image frequency rejection ratio (α) is given by:

$\alpha = \sqrt{^2\rho^2}$

Where $\rho = f_{IF}/f_{RF} - f_{RF}/f_{IF}$; Q is the quality factor of the tuned circuit

2. Sensitivity:

The ability of the radio receiver to pick up the required level of radio signals will enable it to operate more effectively within its application.

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

3. Fidelity

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

UNIT 3

Syllabus Angle modulation: Introduction and types of angle modulation, frequency modulation, frequency deviation, modulation index, deviation ratio, bandwidth requirement of FM wave, types of FM. Phase modulation, difference between FM and PM, Direct and indirect method of FM generation, FM demodulators- slope detector, Foster seeley discriminator, ratio detector. Introduction to pulse modulation systems.

3.1. Introduction: In Frequency Modulation (FM) the instantaneous value of the information signal (message) controls the frequency of the carrier wave. This is illustrated in the following diagrams.

Notice that as the Amplitude of information signal increases, the frequency of the carrier increases, and as the Amplitude of information signal decreases, the frequency of the carrier decreases.
The frequency f_i of the information signal controls the rate at which the carrier frequency increases and decreases. As with AM, f_i must be very much less than f_c. The amplitude of the carrier remains constant throughout this process.
When the information voltage reaches its maximum value then the change in frequency of the carrier will have also reached its maximum deviation above the nominal value. Similarly when the information reaches a minimum the carrier will be at its lowest frequency below the nominal carrier frequency value. When the information signal is zero, then no deviation of the carrier will occur.
The maximum change that can occur to the carrier from its base value f_c is called the **frequency deviation**, and is given the symbol Δf_c. This sets the *dynamic range* (i.e. voltage range) of the transmission. The *dynamic range* is the ratio of the largest and smallest analogue information signals that can be transmitted.

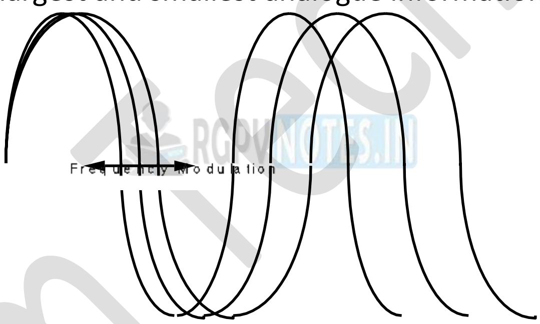


Figure 3.1: Frequency deviation illustration

Notice that frequency modulation looks very much like phase modulation. They are in fact very similar, and many textbooks refer to them both as angle modulation.

3.2. Concept of Frequency Modulation:

Frequency modulation: It is the form of angle modulation in which instantaneous frequency $f_i(t)$ is varied linearly with the information signal m(t)

$$f_i(t)=f_c+k_f\,m(t) \quad \text{.......(1)}$$

where f_c –un-modulated carrier, k_f –Frequency sensitivity of the modulator, m(t)-Information signal
Integrating above equation with respect to time limit 0 to t and multiplying with 2π

$$2\pi\int f_i(t)\,dt = 2\pi f_c\int dt + 2\pi K_f\int m(t)\,dt$$

$$\Theta_i(t) = 2\pi f_c\int dt + 2\pi K_f\int m(t)\,dt$$

$$s(t)=A_c\cos(\Theta_i(t))$$

$$s(t)= A_c\cos(2\pi f_c t + 2\pi K_f\int m(t)\,dt) \quad \text{.......(2)}$$

Phase modulation:

It is that form of Angle modulation in which angle $\phi_i(t)$ is varied linearly with the base band signal m(t) as as shown by

$$\phi_i(t) = K_p m(t)$$

$$S(t)=A_c\cos(\omega_i(t)+\phi_i(t))$$

$$S(t)= A_c\cos(2\pi f_c t+K_p m(t)) \quad \text{.......(3)}$$

3.3. Relationship between PM and FM

PM and FM are closely related in the sense that the net effect of both is variation in total phase angle. In PM, phase angle varies linearly with m(t) where in FM phase angle varies linearly with the integral of m(t). In other words, we can get FM by using PM, provided that at first, the modulating signal is integrated, and then applied to the phase modulator. The converse is also true, i.e. we can generate a PM wave using frequency modulator provided that m(t) is first differentiated and then applied to the frequency modulator.

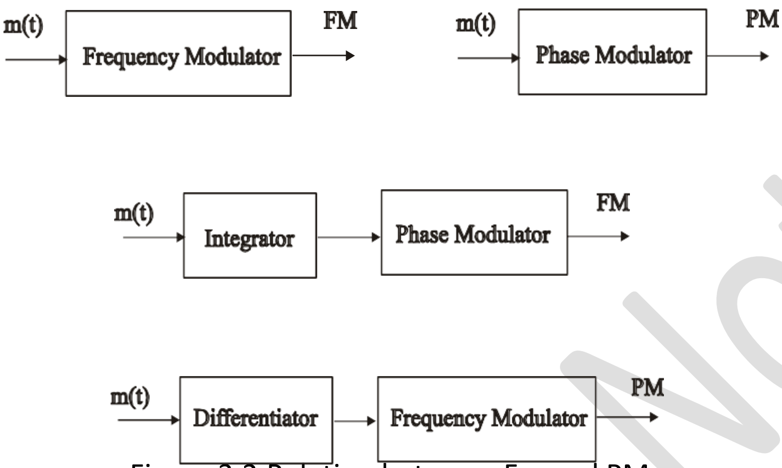


Figure 3.2.Relation between Fm and PM

Recall that a general sinusoid is of the form: $e_c = \sin(\omega_c t + \theta)$

Frequency modulation involves deviating a carrier frequency by some amount. If a sine wave was used to frequency modulate a carrier, the mathematical expression would be:

$$\omega_i = \omega_c + \Delta\omega \sin \omega_m t$$

Where

$\omega_i =$ instantaneous frequency

$\omega_c =$ carrier frequency

$\Delta\omega =$ carrier deviation

$\omega_m =$ modulation frequency

This expression shows a signal varying sinusoidally about some average frequency. However, we cannot simply substitute expression in the general equation for a sinusoid. This is because the sine operator acts upon angles, not frequency. Therefore, we must define the instantaneous frequency in terms of angles. It should be noted that the amplitude of the modulation signal governs the amount of carrier deviation, while the modulation frequency governs the rate of carrier deviation.

The term $\frac{d\theta}{dt}$ is an angular velocity and it is related to frequency and angle by the following relationship:

$$\omega = 2\pi f = \frac{d\theta}{dt}.$$

To find the angle, we must integrate ω with respect to time, we obtain:

$$\int \omega dt = \theta$$

We can now find the instantaneous angle associated with an instantaneous frequency:

$$\theta = \int \omega_i \, dt = \int (\omega_c + \Delta\omega \sin \omega_m t) dt$$

$$= \omega_c t - \frac{\Delta\omega}{\omega_m} \cos \omega_m t = \omega_c t - \frac{\Delta f}{f_m} \cos \omega_m t$$

This angle can now be substituted into the general carrier signal to define FM:

$$e_{fm} = \sin\left(\omega_c t - \frac{\Delta f}{f_m}\cos\omega_m t\right) \qquad \text{.......................................(3)}$$

All FM transmissions are governed by a modulation index, β, which controls the dynamic range of the information being carried in the transmission.

$$\beta = \frac{\Delta f_c}{f_i}$$

3.4. Tone modulation:

Tone modulation is special case when message is sinusoidal as $m(t) = A_m \cos\omega_m t$

For Phase Modulation equation become $s_{PM}(t) = A\cos[\omega_c t + \theta_0 + k_{PM} m(t)]$

$= A\cos[\omega_c t + \theta_0 + k_{PM} A_m \cos\omega_m t]$

$= A\cos[\omega_c t + \theta_0 + m_p \cos\omega_m t]$

where $m_p = k_{PM} A_m$ is the **phase modulation index**, representing the maximum phase deviation $\Delta\theta$.

3.5. Frequency Modulation

$$s_{FM}(t) = A\cos[\omega_c t + \theta_0 + k_{FM}\int m(t)dt] = A\cos[\omega_c t + \theta_0 + k_{FM}\int A_m \cos\omega_m t dt]$$

$$= A\cos[\omega_c t + \theta_0 + \frac{k_{FM} A_m}{\omega_m}\sin\omega_m t] = A\cos[\omega_c t + \theta_0 + m_f \sin\omega_m t]$$

where $\beta = m_f = k_{FM}A_m / \omega_m = \Delta\omega / \omega_m$, i.e. the ratio of frequency deviation to the modulating frequency, is called the **frequency modulation index**.

$\beta = \Delta\omega / \omega_m$..(4)

The relationship between phase deviation and frequency deviation in FM is given by

$\Delta\theta = \beta = \Delta\omega / \omega_m$ ....(5)

3.6. Types of frequency modulation

The bandwidth of an FM signal depends on the frequency deviation. When the deviation is high, the bandwidth will be large, and vice-versa. According to the equation $\Delta\omega = k_{FM}|m(t)|_{max}$, for a given m(t), the frequency deviation, and hence the bandwidth, will depend on frequency sensitivity k_{FM}. Thus, depending on the value of k_{FM} (or $\Delta\omega$) we can divide FM into two categories: *narrowband FM* and *wideband FM*.

3.6.1. Narrowband FM(NBFM (β<<1))

When k_{FM} is small, the bandwidth of FM is narrow, this type of FM is called narrowband FM.

Since when $x << 1$, $\cos x \approx 1$, $\sin x \approx x$, we have

$$s_{NBFM}(t) = A\cos[\omega_c t + \theta_0 + k_{FM}\int m(t)dt]$$

$$= A\cos(\omega_c t + \theta_0)\cos[k_{FM}\int m(t)dt] - A\sin(\omega_c t + \theta_0)\sin[k_{FM}\int m(t)dt]$$

$$\approx A\cos[\omega_c t + \theta_0] - Ak_{FM}\int m(t)dt \sin[\omega_c t + \theta_0]$$

Narrowband modulation methods:

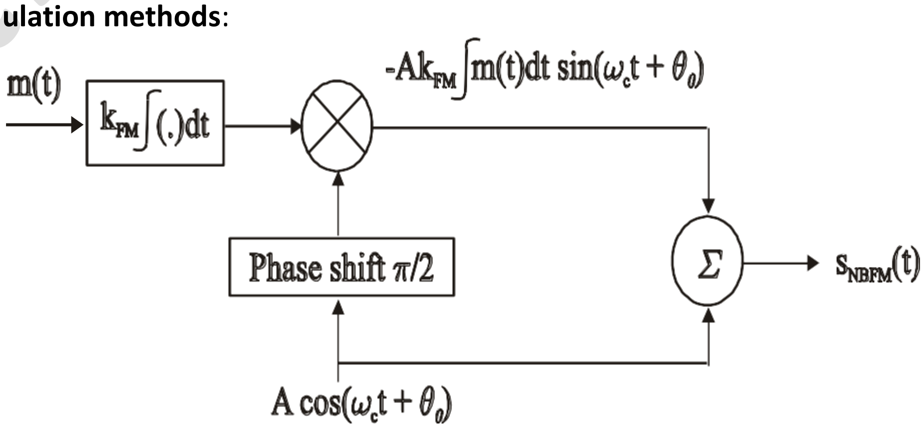


Figure 3.3.NBFM generation

Equation of narrowband frequency modulation (Tone modulation)

The message signal $m(t) = A_m \cos\omega_m t$

Signal waveform (assume $\theta_0 = 0$ for simplicity)

$$s_{NBFM}(t) = A\cos\omega_c t - Ak_{FM}\int m(t)dt \sin\omega_c t = A\cos\omega_c t - Ak_{FM}\int A_m \cos\omega_m t dt \sin\omega_c t$$

$$= A\cos\omega_c t - Ak_{FM}\frac{A_m}{\omega_m}\sin\omega_m t \sin\omega_c t = A\cos\omega_c t - Am_f \sin\omega_m t \sin\omega_c t$$

$$= A\cos\omega_c t + \frac{1}{2}Am_f \cos(\omega_c + \omega_m)t - \frac{1}{2}Am_f \cos(\omega_c - \omega_m)t$$

where $\beta = k_{FM}A_m / \omega_m$ is the FM modulation index.

Signal spectrum $S_{NBFM}(\omega) = \pi A[\delta(\omega - \omega_c) + \delta(\omega + \omega_c)] +$
$(1/2)\pi Am_f[\delta(\omega - \omega_c - \omega_m) + \delta(\omega + \omega_c + \omega_m) - \delta(\omega - \omega_c + \omega_m) - \delta(\omega + \omega_c - \omega_m)]$

Narrowband FM demodulation method

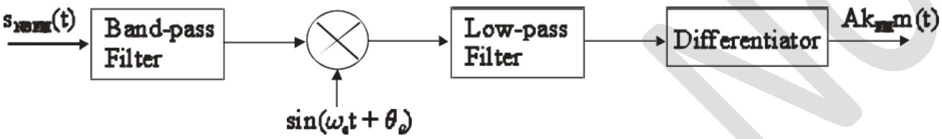


Figure 3.4.NBFM demodulation

3.6.2. Wideband FM (WBFM (β>>1))

When k_{FM} is large, the bandwidth of FM is wide, this type of FM is called *wideband FM*.

It is usually very difficult to analyze a general FM signal, we will restrict our analysis to the wideband FM with sinusoidal signal.

The message signal $m(t) = A_m \cos\omega_m t$

Signal waveform (assume $\theta_0 = 0$ for simplicity)

$$s_{FM}(t) = A\cos[\omega_c t + Ak_{FM}\int m(t)dt] = A\cos[\omega_c t + Ak_{FM}\int A_m \cos\omega_m t dt]$$

$$= A\cos[\omega_c t + m_f \sin\omega_m t] = A\cos\omega_c t \cos(m_f \sin\omega_m t) - A\sin\omega_c t \sin(m_f \sin\omega_m t)$$

$\cos(m_f \sin\omega_m t)$ and $\sin(m_f \sin\omega_m t)$ can be expressed in Fourier series

$$\cos(m_f \sin\omega_m t) = J_0(m_f) + 2\sum_{n=1}^{\infty} J_{2n}(m_f)\cos 2n\omega_m t$$

$$\sin(m_f \sin\omega_m t) = 2\sum_{n=1}^{\infty} J_{2n-1}(m_f)\cos(2n-1)\omega_m t$$

where $J_n(m_f) = \sum_{m=0}^{\infty}\frac{(-1)^m(\frac{m_f}{2})^{2m+n}}{m!(m+n)!}$

is the **Bessel function** of the first kind. Thus,

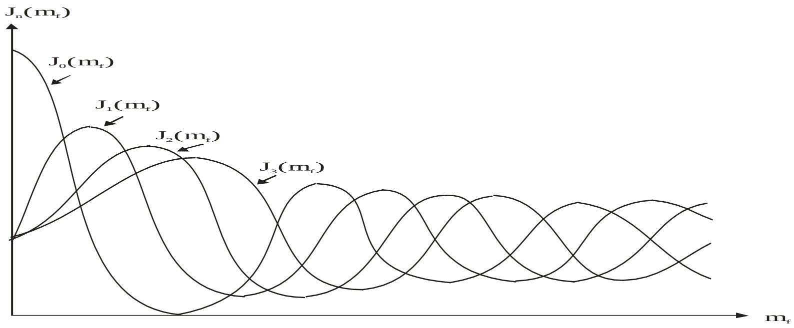
$$s_{FM}(t) = A\cos\omega_c t[J_0(m_f) + 2\sum_{n=1}^{\infty} J_{2n}(m_f)\cos 2n\omega_m t] - A\sin\omega_c t[2\sum_{n=1}^{\infty} J_{2n-1}(m_f)\cos(2n-1)\omega_m t]$$

by using $\cos\alpha\cos\beta = (1/2)[\cos(\alpha + \beta) + \cos(\alpha - \beta)]$,
$\sin\alpha\sin\beta = (1/2)[\cos(\alpha - \beta) - \cos(\alpha + \beta)]$
and the property of Bessel function $J_{-n}(m_f) = (-1)^n J_n(\beta_{FM})$

$s_{FM}(t)$ can be written in the *Bessel function* form

$$s_{FM}(t) = A\sum_{n=-\infty}^{\infty} J_n(m_f)\cos[(\omega_c + n\omega_m)t]$$

The spectrum of $s_{FM}(t)$ $S_{FM}(\omega) = A\pi\sum_{n=-\infty}^{\infty} J_n(m_f)[\delta(\omega - \omega_c - n\omega_m) + \delta(\omega + \omega_c + n\omega_m)]$



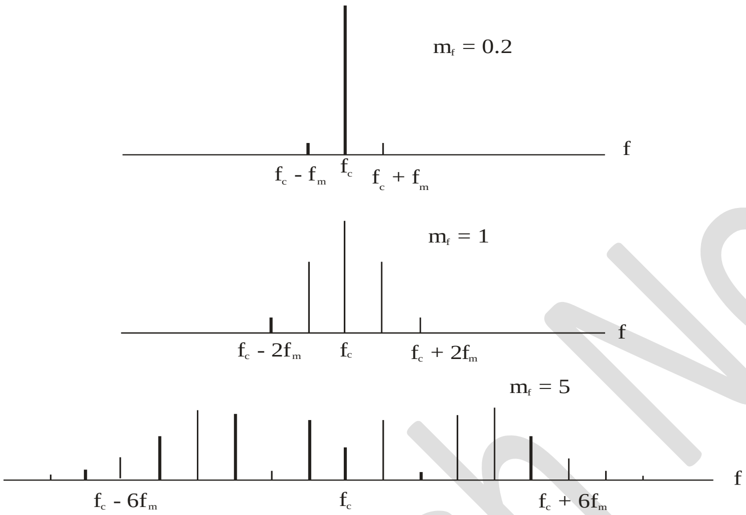


Figure 3.5.Typical plots of $S_{FM}(\omega)$ for different β.

The following observation can be made

- The carrier term $\cos\omega_c t$ has a magnitude of $J_0(m_f)$. The maximum value of $J_0(m_f)$ is 1 when $m_f = 0$, which is equivalent to no modulation.
- Theoretically infinitely number of sidebands are produced, and the amplitude of each sideband is decided by the corresponding *Bessel function* $J_n(m_f)$. The presence of infinite number of sidebands makes the ideal bandwidth of the FM signal infinite.
- When *m_f is small*, there are *few sideband frequencies of large amplitude* and, when *m_f is large*, there are *many sideband frequencies but with smaller amplitudes*. Hence, in practice, to determine the bandwidth, it is only necessary to consider a finite number of *significant sideband* components.
- Thus, the sidebands with small amplitudes can be ignored. The sidebands having amplitudes more than or equal to 1% of the carrier amplitude are known as *significant sidebands*. They are finite in number.

3.7. Bandwidth of a sinusoidally modulated FM signal

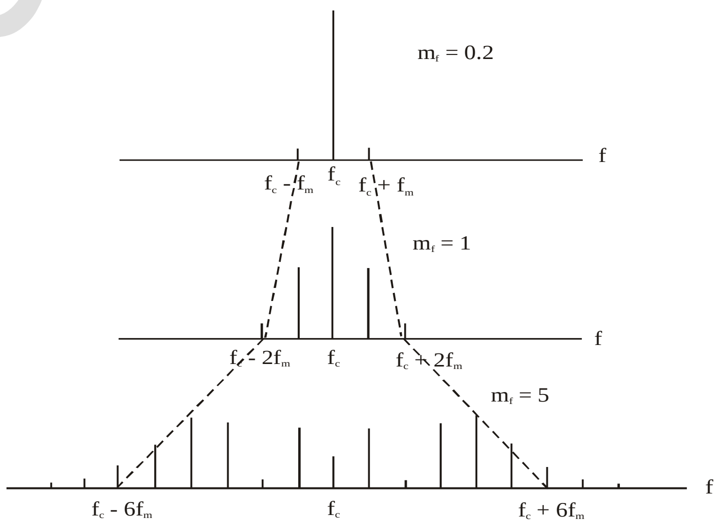


Figure 3.6.Spectrum of FM with different values of m_f

How many sidebands are significant in the FM?

$J_n(m_f)$ diminishes rapidly for $n > m_f$, particularly as m_f becomes large, the number of significant sideband is $m_f + 1$, i.e. $W_{FM} = 2(m_f + 1)\omega_m = 2(\Delta\omega + \omega_m)$

or

$$B_{FM} = 2(m_f + 1)f_m = 2(\Delta f + f_m)$$

Expressed in words, *the bandwidth is twice the sum of the maximum frequency deviation and the modulating frequency*. This rule for bandwidth is called **Carson's rule**.

Bandwidth of FM signal with arbitrary modulating signals

$$W_{FM} = 2(D_{FM} + 1)\omega_m = 2(\Delta\omega + \omega_m)$$

or $$B_{FM} = 2(D_{FM} + 1)f_m = 2(\Delta f + f_m)$$

Where D_{FM} is the **frequency deviation ratio** defined by

$$D_{FM} = \frac{\Delta\omega}{\omega_m} = \frac{\Delta f}{f_m}$$

3.8. Power of the FM signal

Since the amplitude of FM remains unchanged, the power of FM signal is the same as that of unmodulated carrier.

$$P_{FM} = \overline{s^2_{FM}(t)} = A^2 \overline{\sum_{n=-\infty}^{\infty} [J_n(m_f)\cos(\omega_c + n\omega_m)t]^2} = \frac{A^2}{2}\sum_{n=-\infty}^{\infty} J_n^2(m_f) = \frac{A^2}{2}$$

where we used one of the properties of Bessel function $\sum_{n=-\infty}^{\infty} J_n^2(m_f) = 1$

The total power is independent of the FM modulation process, since the power is related to signal amplitude, which is constant for FM, and not necessarily dependent on the signal's phase.

Comparison between WBFM and NBFM

S. No	WBFM	NBFM
I.	Modulating index is greater than1	Modulation index is less than 1
II.	Frequency deviation =75 KHz.	Frequency deviation 5 Khz.
III.	Modulating frequency range	Modulation frequency =3 Khz
iV.	From 30 Hz-15 Khz. Bandwidth 15 times NBFM.	Bandwidth =2 FM
V.	Noise is more suppressed.	Less suppressing of noise
Vi.	Use: Entertainment and broadcasting	Use: Mobile communication.

3.9. Wideband modulation methods

There are two methods for generating wideband FM signals: direct and indirect methods.

Direct method, voltage-controlled oscillator (VCO)

The direct method depends on varying the frequency of an oscillator linearly with m(t) for FM.

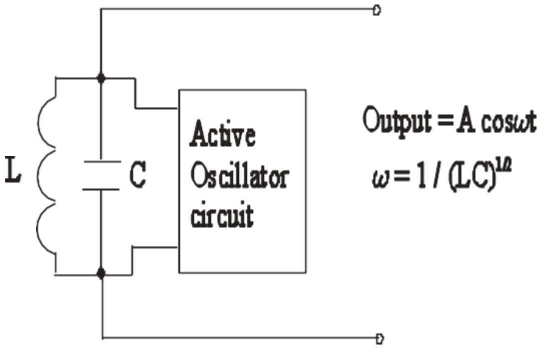


Figure 3.7.Direct method of FM generation

In the VCO, the modulating signal varies the voltage across the capacitor, as a consequence, the capacitance changes and causes a corresponding change in the oscillator frequency, i.e

$C = C_0 + \Delta C = C_0 + k_0 m(t)$

where k_0 is a constant.

Assume that $$\omega_0 = \frac{1}{\sqrt{LC_0}}$$

then $$\omega = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{LC_0(1+\frac{\Delta C}{C_0})}} = \frac{1}{\sqrt{LC_0}}\frac{1}{\sqrt{1+\frac{\Delta C}{C_0}}} \approx \omega_0(1-\frac{\Delta C}{2C_0}) = \omega_0(1-\frac{k_0 m(t)}{2C_0}) = \omega_0 - \frac{\omega_0 k_0}{2C_0}m(t) = \omega_0 - km(t)$$

where $k = \omega_0 k_0 / 2C_0$ is a constant, and the result: $(1 + x)^{-1/2} = 1 - x/2$, when x is small, is used.

Indirect or multiplication method

The indirect method depends on first generating a narrow FM signal and then using a multiplication technique whereby the deviation ratio can be raised to a large value.

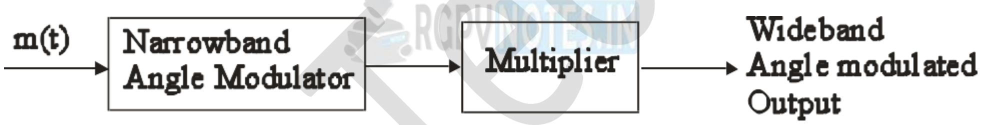


Figure 3.8.WBFM signal using NBFM

The multiplier is a device that multiplies the instantaneous frequency of its input waveform by a factor N.

3.10. Wideband FM demodulation method

Apply $s_{FM}(t)$ to a differentiator, the output is

$$\frac{ds_{FM}(t)}{dt} = \frac{d\{A\cos[\omega_c t + k_{FM}\int m(t)dt]\}}{dt} = -A[\omega_c + k_{FM}m(t)]\sin[\omega_c t + k_{FM}\int m(t)dt]$$

which is similar to a standard AM signal with small deviation ratio. The deviation ratio is small, since usually $\Delta\omega = k_{FM}|m(t)|_{max} << \omega_c$. The response of an envelope detector becomes

$$A[\omega_c + k_{FM}m(t)]$$

Blocking the dc term $A\omega_c$, the output is $s_o(t) = Ak_{FM}m(t)$

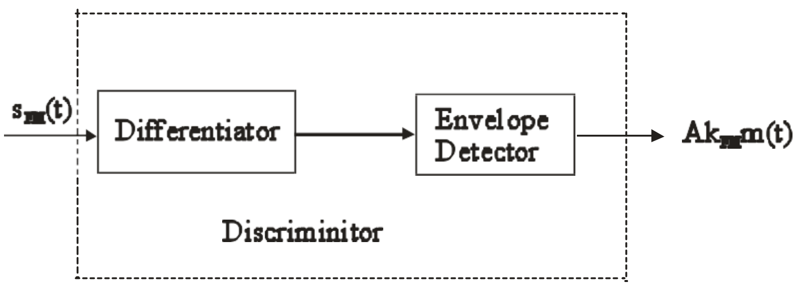


Figure 3.9.WBFM demodulator

The FM detector extracts a modulating signal from a frequency modulated carrier in two steps:

1. It converts the frequency modulated (FM) signal into a corresponding amplitude modulated (AM) signal by using frequency dependent circuits whose output voltage depends on input frequency. Such circuits are called *frequency discriminators*.
2. The original modulating signal m(t) is recovered from this AM signal by using a linear diode envelope detector.

3.11. Comparison between AM and FM

1. Noise performance: Wideband FM has better noise performance than AM. The greater the bandwidth, better is the noise performance. Narrowband FM has a noise performance equivalent to AM.

2. Channel bandwidth
   The wideband FM has a larger bandwidth as compared to AM because wideband FM produces a larger number of sidebands. In a typical broadcast system, each channel bandwidth in AM is 15kHz, whereas, in FM, it is 150kHz. Therefore, FM has a disadvantage over AM.

The modulation index β, is the ratio of the frequency deviation, Δf_c, to the maximum information frequency, f_i, as shown below:

The diagrams opposite show examples of how the modulation index affects the FM output, for a simple sinusoidal information signal of fixed frequency. The carrier signal has a frequency of ten times that of the information signal.

The first graph shows the information signal, the second shows the unmodulated carrier.

This graph shows the frequency modulated carrier when the modulation index = 3.

This graph shows the frequency modulated carrier when the modulation index = 5.

This graph shows the frequency modulated carrier when the modulation index = 7.

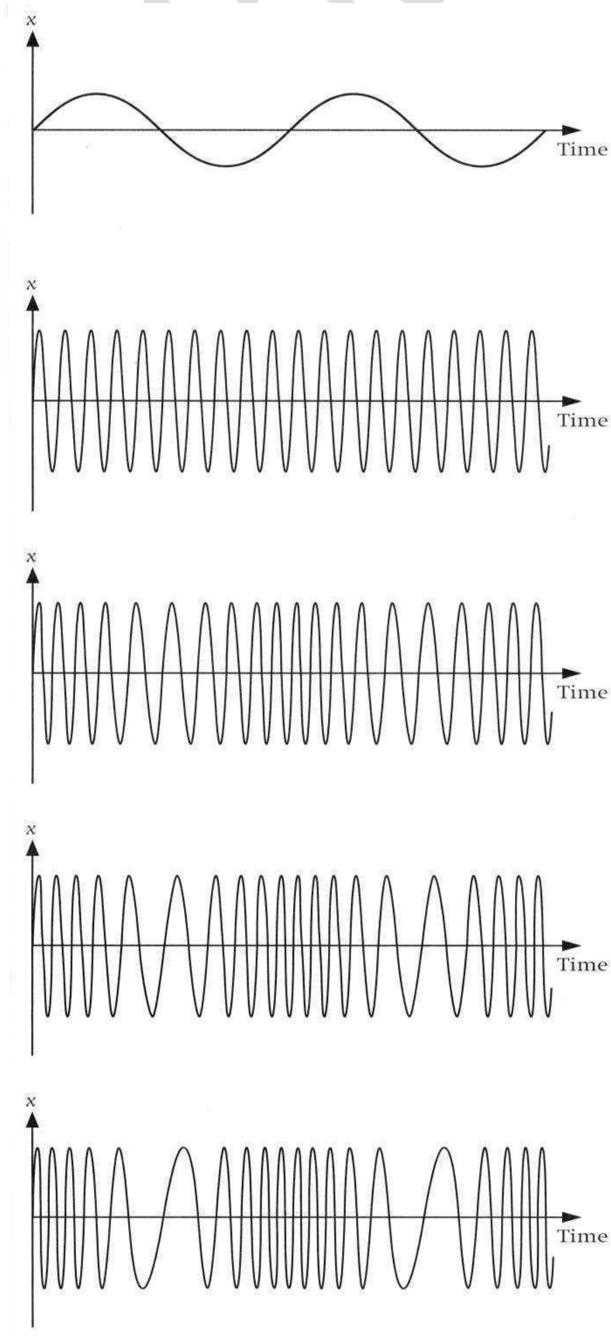


Figure 3.10.FM with different Modulation indices

As the modulation index increases you should notice that the peaks of the high frequency get closer together and low frequency get further apart. For the same information signal therefore, the carrier signal has a higher maximum frequency.

The FM modulation index is defined as the ratio of carrier deviation to modulation frequency:

As a result, the FM equation is generally written as:

$$e_{fm} = \sin\left(\omega_c t - m_{fm} \cos\omega_m t\right)$$

This is a very complex expression and it is not readily apparent what the sidebands of this signal are like. The solution to this problem requires knowledge of Bessel's functions of the first kind and order *p*. In open form, it resembles:

$$J_p(x) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{x}{2}\right)^{2k+p}}{k!(k+p)!}$$

$$J_p(x) = \text{magnitude of frequency component}$$

where $p = \text{side frequency number}$

$x = \text{modulation index}$

As a point of interest, Bessel's functions are a solution to the following equation:

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

Bessel's functions occur in the theory of cylindrical and spherical waves, much like sine waves occur in the theory of plane waves. It turns out that FM generates an infinite number of side frequencies. Each frequency is an integer multiple of the modulation signal. It should be noted that the amplitude of the higher order sided frequencies drops off quickly. It is also interesting to note that the amplitude of the carrier signal is also a function of the modulation index. Under some conditions, the amplitude of the carrier frequency can actually go to zero. This does not mean that the signal disappears, but rather that all of the broadcast energy is redistributed to the side frequencies.

3.12. Generation of frequency Modulation

3.12.1. Direct FM

(a) Reactance Modulator

The reactance modulator is a voltage controlled capacitor and is used to vary an oscillator's frequency or phase. A simplified circuit resembles:

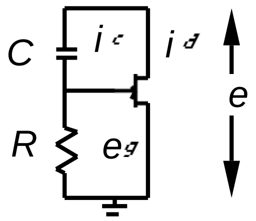


Figure 3.11.Reactance modulator

Since the gate does not draw an appreciable amount of current, applying Ohm's law in the RC branch results in:

$$e_g = i_C$$

$$i_C = \frac{e}{R - jX_C}$$

$$\therefore e_g = \frac{e}{R - jX_C} R$$

The JFET drain current is given by:

$$i_d = g_m e_g = g_m \frac{e}{R - jX_C} R$$

where g_m is the trans-conductance.

The impedance as seen from the drain to ground is given by:

$$Z = \frac{e}{i_d} = e\frac{1}{g_m}\frac{R - jX_C}{e}\frac{1}{R} = \frac{1}{g_m} - j\frac{X_C}{g_m R}$$

Since trans-conductance is normally very large, the impedance reduces to:

$$Z \approx -j\frac{X_C}{g_m R} = \frac{-j}{2\pi f C g_m R}$$

The term in the denominator can be thought of as an equivalent capacitance:

$$C_{eq} = C g_m R$$

Then

$$Z = \frac{-j}{2\pi f C_{eq}}$$

Since the equivalent capacitance is larger than the original capacitor, we have created a capacitance amplifier. Because the value of this capacitance is a function of applied voltage, we actually have a voltage controlled capacitor. This device can be used to control an oscillator frequency, thus producing FM.

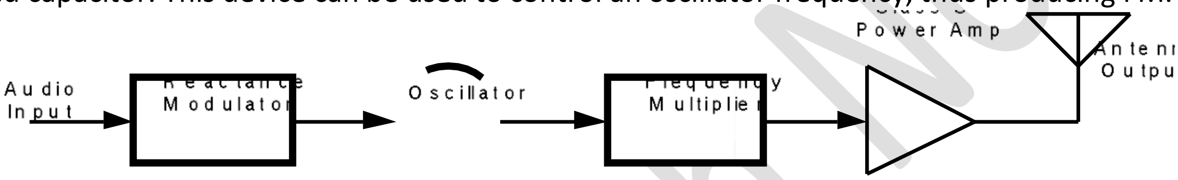


Figure 3.12.FM transmitter

Varactor Diodes

The capacitance of a varactor diode is a function of its' reverse bias voltage.

$$C_d = \frac{C_0}{\sqrt{1 + 2|V_R|}}$$

Where C_0 is the diode capacitance at zero bias, and V_R is the reverse bias voltage. A typical response is:

3.12.2. Indirect Method:

Because crystal oscillators are so stable, it is desirable to use them in modulator circuits. However, their extreme stability makes it difficult to modulate their frequency.

Fortunately, it is possible to vary the phase of a crystal oscillator. However, in order to use this as an FM source, the relationship between frequency and phase needs to be reexamined.

Frequency is the rate of change of angle, its first derivative:

$$\omega = \frac{d}{dt}\phi$$

The instantaneous phase angle is comprised of two components, the number of times the signal has gone through its cycle, and its starting point or offset:

$$\phi(t) = \underbrace{\omega_c t}_{\text{rotating angle}} + \underbrace{\theta}_{\text{offset angle}}$$

The instantaneous frequency is therefore:

$$\omega_i = \frac{d}{dt}\phi(t) = \frac{d}{dt}\left[\omega_c t + \theta\right] = \omega_c + \frac{d}{dt}\theta$$

From this we observe that the instantaneous frequency of a signal is its un-modulated frequency plus a change. This is equivalent to frequency modulation. Therefore we may write:

$$\omega_c + \frac{d}{dt}\theta = \omega_c + \omega_{eq}$$

$$\omega_{eq} = \frac{d}{dt}\theta$$

$$f_{eq} = \frac{1}{2\pi}\frac{d}{dt}\theta$$

This means that the output of a phase modulator is proportional to the equivalent frequency modulation.
If the angle is proportional to the amplitude of a modulation signal $\theta = k e_m$ Then:

$$f_{eq} = \frac{1}{2\pi}\frac{d}{dt} k e_m$$

and by integrating the modulation signal prior to modulation, we obtain:

$$f_{eq} = \frac{1}{2\pi}\frac{d}{dt}\int k e_m \, dt = \frac{k}{2\pi} e_m$$

This means that the equivalent frequency modulation is directly proportional to the amplitude of a phase modulation signal if the modulation signal is integrated first.
This indirect modulation scheme is the heart of the Armstrong modulator.

3.13. Demodulation / Detector of Frequency Modulation:

3.13.1. Phase Detector [Foster-Seeley]

The Foster-Seeley detector converts the incoming frequency variation to an equivalent phase variation and then to an equivalent amplitude variation. This is accomplished by using the phase angle shift which occurs between the primary and secondary of a transformer tuned circuit.
It is important to recognize that the signal on the primary side gets to the secondary through two distinctly different paths:

- Through the transformer via the primary winding
- bypassing the primary winding and directly into the secondary center tap

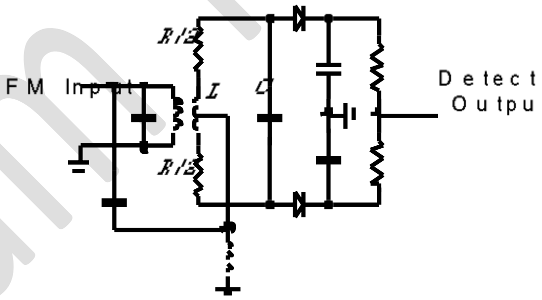


Figure 3.13.Foster seeley detector

The voltage appearing on the secondary side of the transformer is given by:

$$e_s \approx e_p k \sqrt{\frac{L_p}{L_s}}$$

e_p = primary voltage

k = coupling coefficient

L_p = primary inductance

L_s = secondary inductance

This is applied to the series resonant circuit in the transformer secondary winding. The impedance of this resonant circuit is given by:

$$Z = R + j\left(\omega L - \frac{1}{\omega C}\right) = R\left[1 + j\left(\frac{\omega L}{R} - \frac{1}{\omega RC}\right)\right]$$

$$= R\left[1 + j\left(\frac{\omega}{\omega_o}\frac{\omega_o L}{R} - \frac{\omega_o}{\omega}\frac{1}{\omega_o RC}\right)\right]$$

Where ω_o is the resonant frequency

Since, $Q = \frac{\omega_o L}{R} = \frac{1}{\omega_o RC}$, the impedance can be written as:

$$Z = R\left[1 + j\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)Q\right]$$

Defining a new parameter: $Y = \frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}$, we obtain:

$$Z = R\left[1 + jYQ\right]$$

The impedance phase angle is given by: $\varphi = \tan^{-1}(YQ)$.

It is interesting to observe what happens to this angle when the input frequency varies.

Let the input frequency be of the form: $\omega = \omega_o + \Delta\omega$, then:

$$Y = \frac{\omega_o + \Delta\omega}{\omega_o} - \frac{\omega_o}{\omega_o + \Delta\omega} = \frac{2\omega_o \Delta\omega + \Delta\omega^2}{\omega_o^2 + \omega_o \Delta\omega}$$

But if the deviation is much smaller than the carrier: $\Delta\omega << \omega_o$, then:

$$Y \approx \frac{2\Delta\omega}{\omega_o}$$

Notice that the parameter *Y* varies directly with deviation. For small angle changes $\varphi \approx \tan^{1}\varphi$. This means that the impedance phase angle varies directly with the frequency deviation. This in turn causes a variation in the currents and voltages in the secondary.

The output of the transformer consists of the vector sum of two components:

- The phase shifted signal passing through the transformer
- The un-shifted signal which has bypassed the transformer

The combination of these two signals results in amplitude variations which are directly proportional to the frequency deviation. This AM signal is then detected through a standard envelope detector.

This circuit is quite sensitive however; any amplitude variations in the signal caused by varying signal strength are also detected.

3.13.2. Ratio Detector

This circuit is a slight modification of the Foster-Seeley detector:

- One diode is reversed
- The output is taken from the combined loads

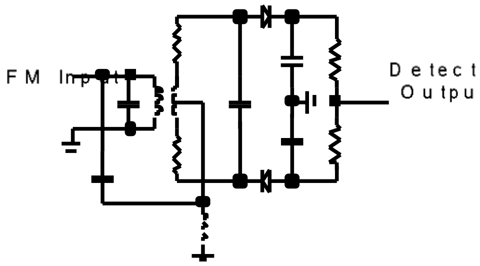


Figure 3.14.Ratio detector

The limiting action of the detector has been improved and variations in signal strength are not as noticeable. This also implies that it is less sensitive.

3.13.3. Phase Locked Loop

Although phase locked loops can be implemented using analog or digital circuitry, the following discussion will be limited to linear circuits since they are much easier to analyze.

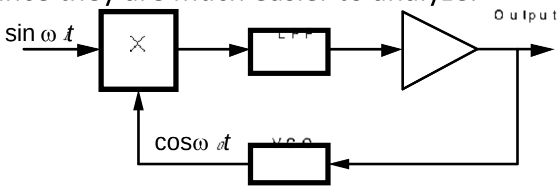


Figure 3.16.Phase lock loop

The loop achieves lock in two steps. First it acquires frequency lock, and then it acquires phase lock.
If the input signal and VCO output are completely different, the output of the multiplier is given by:

$$V_{mult} = \sin(\omega_i t)\cos(\omega_o t) = \frac{1}{2}\sin(\omega_i + \omega_o)t + \frac{1}{2}\sin(\omega_i - \omega_o)t$$

The low pass filter removes the high frequency component before passing the signal to the output amplifier. The output, which also is the input to the VCO, is of the form:

$$V_{output} = \sin(\omega_i - \omega_o)t$$

Notice that if $\omega_i > \omega_o$, the output is positive, but if $\omega_i < \omega_o$, the output is negative.
This change in polarity can be seen by observing the sine function near the origin:

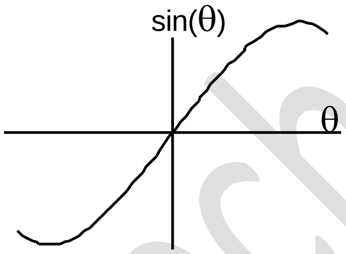


Figure 3.17.sinusoidal variation

This voltage can be used as an error signal to drive the VCO until $\omega_i = \omega_o$, and frequency lock is achieved.
The two signals however, are not necessarily in phase at this point.
Once frequency lock occurs, the multiplier output becomes:

$$V_{mult} = \sin(\omega_i t)\cos(\omega_i t - \varphi) = \frac{1}{2}\sin(2\omega_i - \varphi)t + \frac{1}{2}\sin(\varphi)t$$

Again the low pass filter removes the high frequency component and the output is of the form:

$$V_{output} = \sin(\varphi)t$$

Again if $\varphi > 0$ the output is positive; but if $\varphi < 0$ the output is negative. This error signal is used to drive the VCO until $\varphi = 0$, and phase lock is achieved. Notice that at lock, the incoming signal and the VCO output are in quadrature.
A PLL can be used after the IF amplifier in a radio to reproduce the original modulation signal. This allows the free running VCO oscillator frequency to be preset, thus making it easier to acquire lock.

3.14. FM Spectrum:

When the amplitude of the frequency components of FM waveform are plotted as a function of frequency, the resulting spectrum is much more complicated than that of the AM waveform This is because there are now multiple frequencies present in the FM signal.
Theoretically, an FM spectrum has an infinite number of sidebands, spaced at multiples of f_i above and below the carrier frequency f_c. However the size and significance of these sidebands is very dependent on the modulation index, β. If β<1, then the spectrum looks like this:

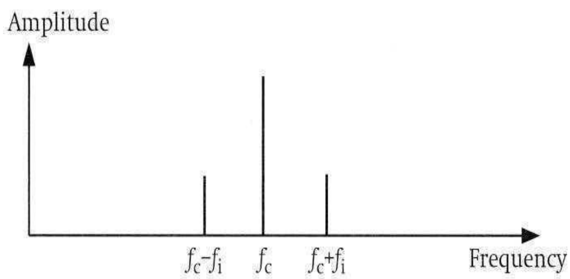


Figure 3.18.FM spectrum for β<1

From the spectrum above it can be seen that there are only two significant sidebands, and thus the spectrum looks very similar to that for an AM carrier.
Ifβ=1, then the spectrum looks like this:

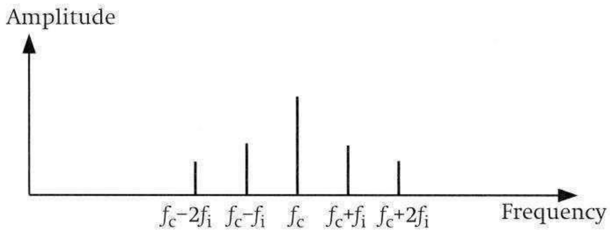


Figure 3.19.FM spectrum for β =1

From the spectrum above we can see that the number of significant sidebands has increased to four.

Ifβ=3, then the spectrum looks like this:

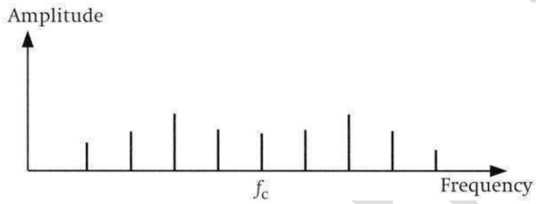


Figure 3.20.FM spectrum for β=3

From the spectrum above we can see that the number of significant sidebands has increased to eight. It can be deduced that the number of significant sidebands in an FM transmission is given by2(β+1). The implication for the bandwidth of an FM signal should now be coming clear. The practical bandwidth is going to be given by the number of significant sidebands multiplied by the width of each sideband (i.e. f_i).

$$\begin{aligned} Bandwidth_{FM} &= 2(\beta + 1) f_i \\ &= 2\left(\frac{\Delta f_c}{f_i} + 1\right) f_i \\ &= 2(\Delta f_c + f_i) \end{aligned}$$

The bandwidth of an FM waveform is therefore twice the sum of the frequency deviation and the maximum frequency in the information.

3.15. Carson rule:

It states that the bandwidth required transmitting an angle modulated wave as twice the sum of the peak frequency deviation and the highest modulating signal frequency.

Band Width = 2 [$\Delta f + f_{m(max)}$] Hz

Δf = frequency deviation in Hz

$f_{m(max)}$ = highest modulating signal frequency in Hz

Additional Points to remember.

- An FM transmission is a constant power wave, regardless of the information signal or modulation index, β because it is operated at constant amplitude with symmetrical changes in frequency.
- As β increases, the relative amplitude of the carrier component decreases and may become much smaller than the amplitudes of the individual sidebands. The effect of this is that a much greater proportion of the transmitted power is in the sidebands (rather than in the carrier), which is more efficient than AM.

3.15.1. Determination of Bandwidth for FM Radio

FM radio uses a modulation index, $\beta > 1$, and this is called **wideband FM.** As its name suggests the bandwidth is much larger than AM.

In national radio broadcasts using FM, the frequency deviation of the carrier Δf_c, is chosen to be 75 kHz, and the information baseband is the high fidelity range 20 Hz to 15 kHz.
Thus the modulation index, β is 5 (i.e. 75 kHz + 15 kHz), and such a broadcast requires an FM signal bandwidth given by:

$$Bandwidth_{FM\ \ Radio} = 2(\Delta f_c + f_{i(\max)})$$

$$= 2(75+15)$$

$$= 180kHz$$

3.16. Advantages of FM over AM

a) The amplitude of FM is constant. Hence transmitter power remains constant in FM where as it varies in AM.
b) Since amplitude of FM is constant, the noise interference is minimum in FM. Any noise superimposing on modulated carrier can be removed with the help of amplitude limiter.
c) The depth of modulation has limitation in AM. But in FM, the depth of modulation can be increased to any value.
d) Since guard bands are provided in FM, there is less possibility of adjacent channel interference.
e) Since space waves are used for FM, the radius of propagation is limited to line of sight(LOS) . Hence it is possible to operate several independent transmitters on same frequency with minimum interference.
f) Since FM uses UHF and VHF ranges, the noise interference is minimum compared to AM which uses MF and HF ranges.

3.17. Introduction to pulse modulation systems:

Pulse modulation is "the process in which signal is transmitted by pulses (i.e., discontinuous signals) with a special technique". The pulse modulation is classified as analog pulse modulation and digital pulse modulation. The analog pulse modulation is again classified as,
1. Pulse amplitude modulation
2. Pulse width modulation and
3. Pulse position modulation

3.17.1. Pulse Amplitude Modulation

In Pulse Amplitude Modulation (PAM) technique, the amplitude of the pulse carrier varies, which is proportional to the instantaneous amplitude of the message signal. The width and positions of the pulses are constants in this modulation.

3.17.2. Pulse Width Modulation

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

3.17.3. Pulse Position Modulation

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

Analog & Digital communication
IT-404

UNIT 4

Syllabus: Sampling of signal, sampling theorem for low pass and Band pass signal, Pulse amplitude modulation (PAM), and Time division multiplexing (TDM). Channel Bandwidth for PAM-TDM signal Type of sampling instantaneous, Natural and flat top, Aperture effect, Introduction to pulse position and pulse duration modulations, Digital signal, Quantization, Quantization error, Pulse code modulation, signal to noise ratio, Companding, Data rate and Baud rate, Bit rate, multiplexed PCM signal, Differential PCM (DPCM), Delta Modulation (DM) and Adaptive Delta Modulation (ADM), comparison of various systems.

4.1. Sampling of signal: The signals we use in the real world, such as our voices, are called "analog" signals. To process these signals in computers, we need to convert the signals to "digital" form. While an analog signal is continuous in both time and amplitude, a digital signal is discrete in both time and amplitude. To convert a signal from continuous time to discrete time, a process called sampling is used. The value of the signal is measured at certain intervals in time. Each measurement is referred to as a sample.

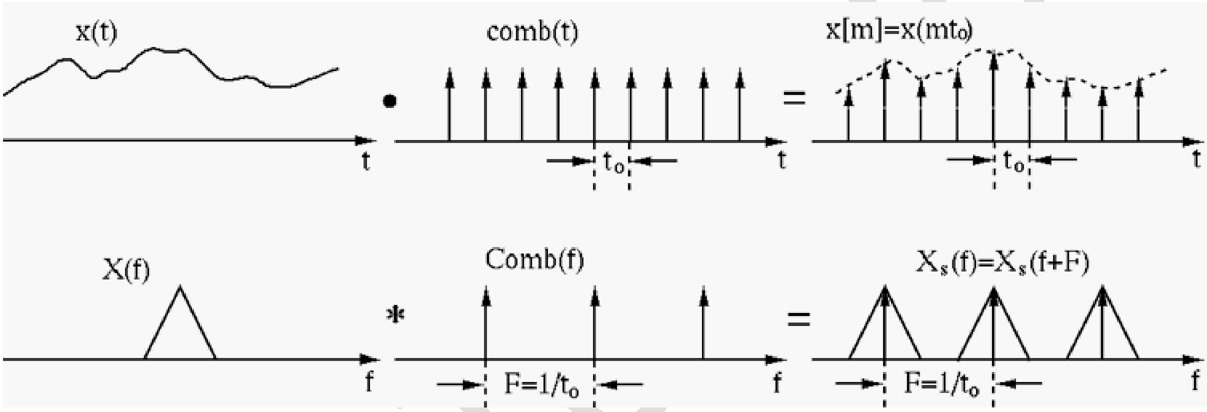


Figure.4.1. Sampling of signal and its frequency spectrum

A discrete-time signal can be obtained by uniformly sampling a continuous-time signal at $t_n = nT_S$, i.e.,x[n] = $x[nT_s]$. The values x[n] are samples of x(t), the time interval between samples is T_s, the sampling rate is $f_s = 1/T_s$. A system which performs the sampling operation is called a continuous-to-discrete (C-to-D) converter.

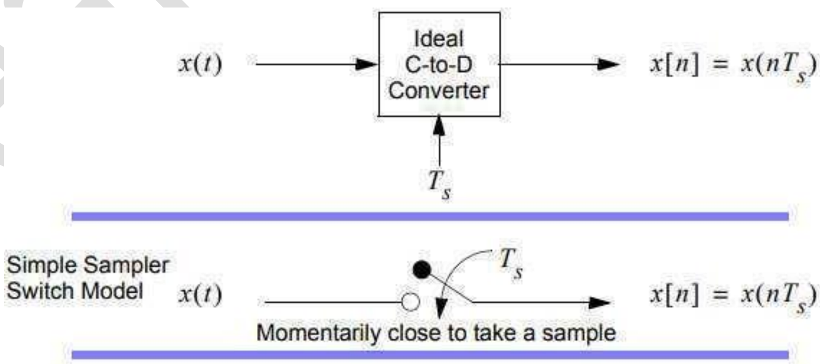


Figure.4.2.Sampler as switch

4.2. Sampling theorem for low pass signal

The lowpass sampling theorem states that, "a signal whose spectrum is band limited to B Hz, can be reproduced exactly from its samples if it is sampled at the rate fs which is greater than twice the maximum frequency ω of the signal to be sampled." Therefore minimum sampling frequency is

$$f_s = 2B\ Hz$$

If we have a continuous time signal x(t). The spectrum of x(t) is a band limited to f_m Hz i.e. the spectrum of x(t) is zero for $|\omega|>\omega_m$.

Sampling theorem statement 2: A continuous-time signal x(t) with frequencies no higher than f_m (Hz) can be reconstructed EXACTLY from its samples x[n] = x(nTs), if the samples are taken at a rate fs = 1/Ts that is greater than $2f_m$.

Sampling of input signal x(t) can be obtained by multiplying x(t) with an impulse train δ(t) of period T_s. The output of multiplier is a discrete signal called sampled signal which is represented with y(t) in the following diagrams:

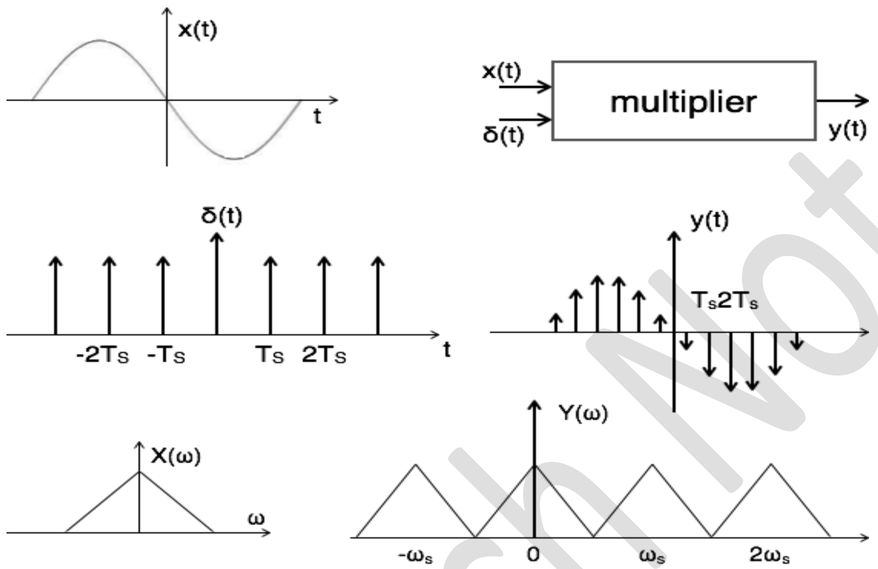


Figure.4.3. Sampling of signal x(t)

Here, you can observe that the sampled signal takes the period of impulse. The process of sampling can be understood as under.

The sampled signal is given by

$$y(t) = x(t).(t) = x(t).T_s(t)$$
$$= \sum_n x(nT_s).(t-nT_s)$$

The impulse train $T_s(t)$ is a periodic signal of period T_s, hence it can be expressed as a Fourier series

$$T_s(t) = \frac{1}{T_s}[1 + 2\, cos\omega_s t + \; 2\, cos2\omega_s t + \; 2\, cos3\omega_s t + \cdots]$$
$$\text{Where } \omega_s = \frac{2\pi}{T_s} = 2\pi f_s$$

Substitute (t)in equation 1.

$$\rightarrow y(t) \; = \; x(t).\,(t)$$
$$y(t) = \frac{1}{T_s}[x(t) + 2x(t)\, cos\omega_s t + 2x(t)\, cos2\omega_s t + \; 2\, x(t) cos3\omega_s t + \cdots]$$

Now to find the $Y(\omega)$, we have to take the fourier transform of both the sides,

$$Y(\omega) = \frac{1}{T_s}[X(\omega) + X(\omega - \omega_s) + X(\omega + \omega_s) + X(\omega - 2\omega_s) + X(\omega + 2\omega_s) + \cdots]$$
$$Y(\omega) = \frac{1}{T_s}\sum_{n=-\infty}^{\infty} X(\omega - n\omega_s) \text{ Where } n = 0, \pm 1, \pm 2, \ldots$$

The Fourier spectrum $Y(\omega)$ is shown in figure. Now If we have to recover $x(t)$ from $y(t)$, we should be able to recover $X(\omega)$ from $Y(\omega)$, and it is possible only if there is no overlapping between the successive cycles of $Y(\omega)$, and for this condition

$$f_s > 2B\; Hz$$

Therefore the sampling interval

$$T_s < \frac{1}{2B}$$

Therefore as long as the sampling frequency f_s is greater than $2B$, $Y(\omega)$, will consist of non overlapping repetitions of $X(\omega)$, and $x(t)$ can be recovered from $y(t)$ by passing $y(t)$ by an ideal low pass filter with cut off frequency B Hz.

4.3. Sampling theorem for Band pass signal

In case of band pass signals, the spectrum of band pass signal X[ω] = 0 for the frequencies outside the range $f_1 \leq f \leq f_2$. The frequency f_1 is always greater than zero. Plus, there is no aliasing effect when $f_s > 2f_2$. But it has two disadvantages:

- The sampling rate is large in proportion with f_2. This has practical limitations.
- The sampled signal spectrum has spectral gaps.

To overcome this, the band pass theorem states that the input signal x(t) can be converted into its samples and can be recovered back without distortion when sampling frequency $f_s < 2f_2$.

4.4. Nyquist criteria and Signal reconstruction of sampled signal:

According to the Nyquist criterion, in order to recover the original signal from its samples, the minimum sampling frequency rate must be twice of the bandwidth of the message signal.

This minimum sampling rate f_s=2B required to recover x(t) from its samples y(t) is called the Nyquist rate and the corresponding sampling interval is called Nyquist Interval for y(t).

Effect of Sampling Rate on Y(ω)

Possibility of sampled frequency spectrum with different conditions is given by the following diagrams:

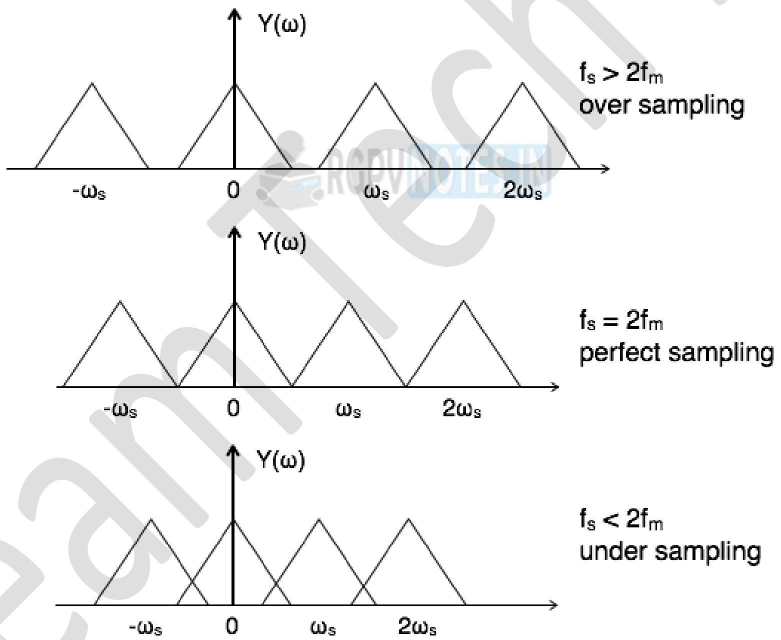


Figure.4.4. Sampling Conditions

Therefore, if we are sampling below the Nyquist rate then the recovery of the signal will not be possible.

4.5. Aliasing – Aliasing refers to the phenomenon of a high frequency component in the spectrum of a signal seemingly taking on the identity of a lower frequency in the spectrum of its sampled version (under sampled version of the message signal) Corrective measures for aliasing effects

1. Prior to sampling, a low-pass anti-aliasing filter is used to attenuate those high-frequency components of the signal that are not essential to the information being conveyed by the signal.

2. The filtered signal is sampled at a rate higher than the Nyquist rate.

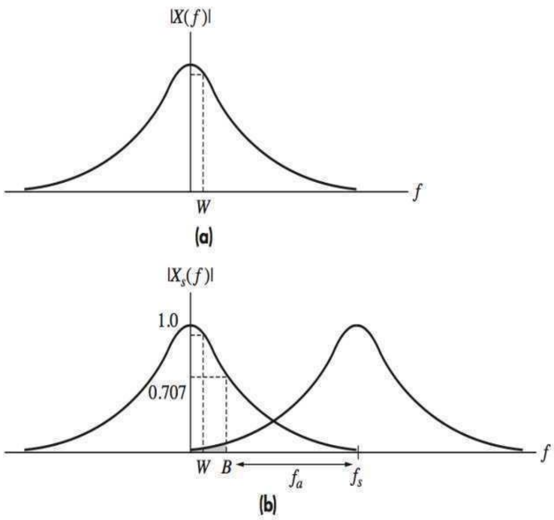
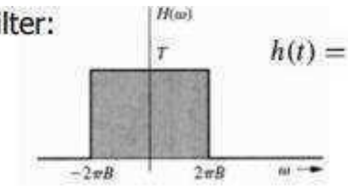


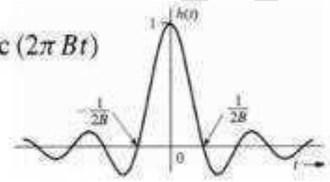
Figure.4.5. Aliasing

Ideal signal reconstruction:

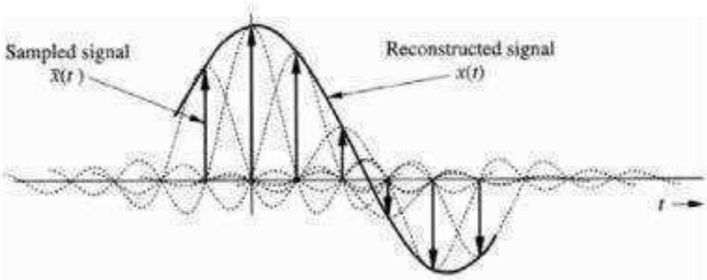
- Use ideal lowpass filter:



$h(t) = \text{sinc}\,(2\pi Bt)$



- That's why the sinc function is also known as the **interpolation** function:



$$x(t) = \sum_n x(nT)h(t - nT)$$

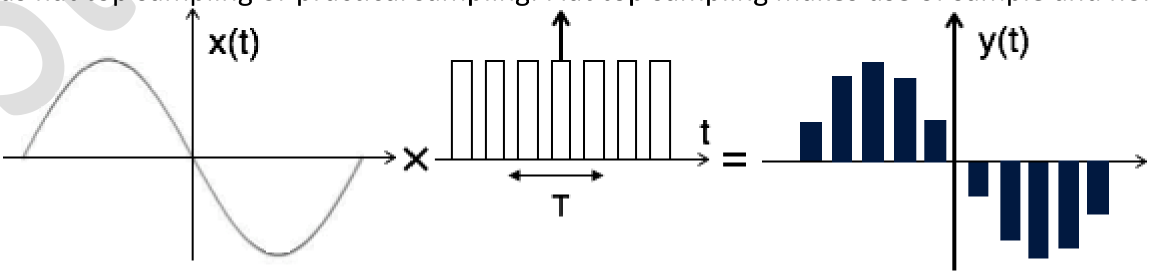
$$= \sum_n x(nT)\,\text{sinc}\,(2\pi Bt - n\pi)$$

Figure.4.6. Ideal reconstruction system

4.5. Types of sampling:

4.5.1. Flat Top Sampling

During transmission, noise is introduced at top of the transmission pulse which can be easily removed if the pulse is in the form of flat top. Here, the top of the samples are flat i.e. they have constant amplitude. Hence, it is called as flat top sampling or practical sampling. Flat top sampling makes use of sample and hold circuit.



Figur.4.7. Flat Top Sampling

Theoretically, the sampled signal can be obtained by convolution of rectangular pulse p(t) with ideally sampled signal say $y_\delta(t)$ as shown in the diagram:

4.5.2. Natural Sampling

Natural sampling is similar to impulse sampling, except the impulse train is replaced by pulse train of period T. i.e. you multiply input signal x(t) to pulse train $\sum_{n=-\infty}^{\infty} P(t-nT)$ as shown below

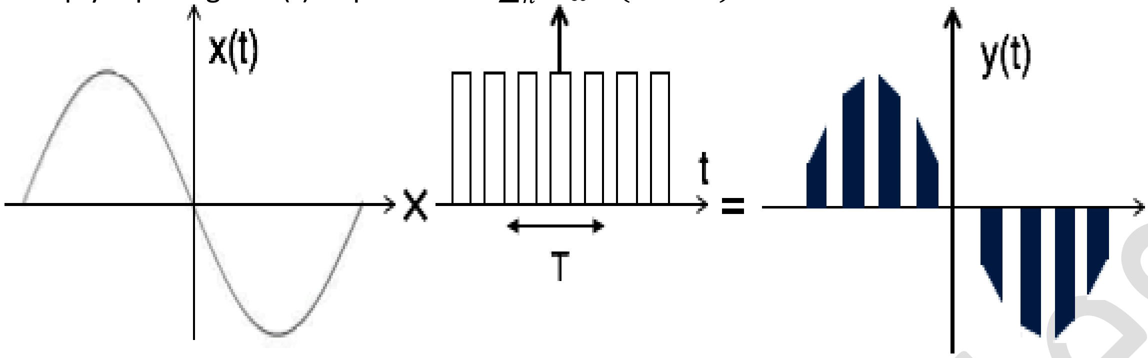


Figure.4.8.Natural Sampling

4.5.3. Impulse Sampling

Impulse sampling can be performed by multiplying input signal x(t) with impulse train Σδ(t−nT) of period 'T'. Here, the amplitude of impulse changes with respect to amplitude of input signal x(t). The output of sampler is given by

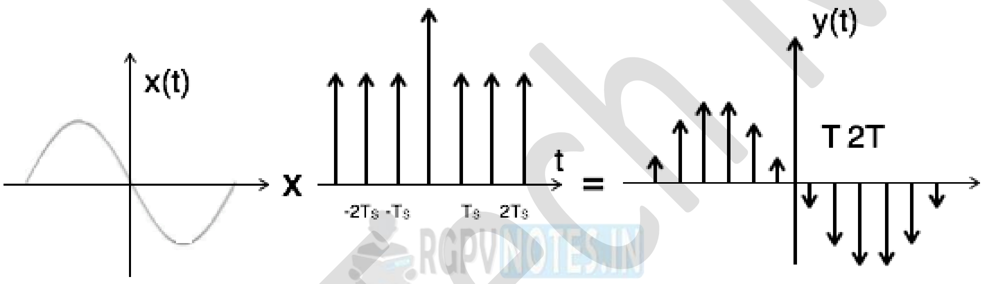


Figure.4.9.Impulse Sampling

y(t)=x(t)×y(t)=x(t)× impulse train

=x(t)×$\sum_{n=-\infty}^{\infty} (t-nT)$

y(t)=yδ(t)= $\sum_{n=-\infty}^{\infty} x(nt)\,(t-nT)$

To get the spectrum of sampled signal, consider Fourier transform of equation on both sides

Y(ω)=1$\frac{1}{T}\sum_{n=-\infty}^{\infty} X(\omega - n\omega_s)$

This is called ideal sampling or impulse sampling. You cannot use this practically because pulse width cannot be zero and the generation of impulse train is not possible practically.

4.6. Time Division Multiplexing:

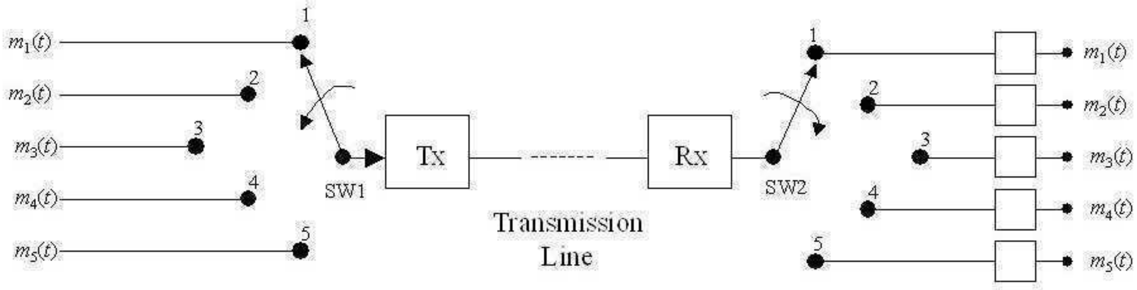
In TDM, the data flow of each input stream is divided into units. One unit may be 1 bit, 1 byte, or a block of few bytes. Each input unit is allotted an input time slot. One input unit corresponds to one output unit and is allotted an output time slot. During transmission, one unit of each of the input streams is allotted one-time slot, periodically, in a sequence, on a rotational basis. TDM allows each channel the full band width of the transmission medium whenever its signal is transmitted, although each channel is not continuously on the system.

Advantages of TDM

– high reliability and efficient operation as the circuitry required is digital.

– Relatively small inter channel cross-talk arising from nonlinearities in the amplifiers that handle the signals in the transmitter and receiver.

Disadvantages of TDM – timing jitter



Switches SW1 and SW2 rotate in synchronism, and in effect sample each message input in a sequence $m_1(t)$, $m_2(t)$, $m_3(t)$, $m_4(t)$, $m_5(t)$, $m_1(t)$, $m_2(t)$,…

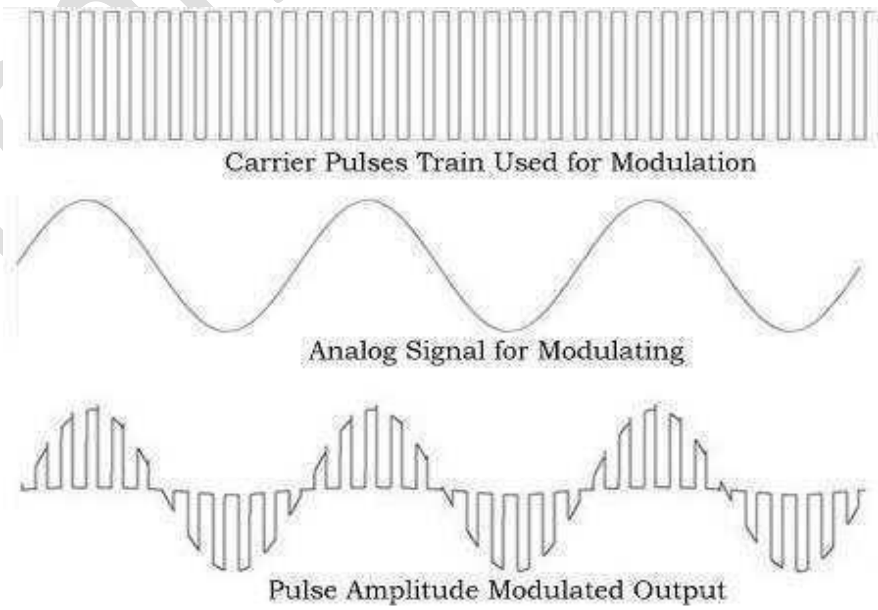
Figure.4.10. TDM system

4.7. Pulse Amplitude Modulation:

In pulse amplitude modulation, the amplitude of regular interval of periodic pulses or electromagnetic pulses is varied in proposition to the sample of modulating signal or message signal. This is an analog type of modulation. In the pulse amplitude modulation, the message signal is sampled at regular periodic or time intervals and this each sample is made proportional to the magnitude of the message signal. These sample pulses can be transmitted directly using wired media or we can use a carrier signal for transmitting through wireless. There are two types of sampling techniques for transmitting messages using pulse amplitude modulation, they are

- **FLAT TOP PAM:** The amplitude of each pulse is directly proportional to instantaneous modulating signal amplitude at the time of pulse occurrence and then keeps the amplitude of the pulse for the rest of the half cycle.
- **Natural PAM:** The amplitude of each pulse is directly proportional to the instantaneous modulating signal amplitude at the time of pulse occurrence and then follows the amplitude of the modulating signal for the rest of the half cycle.

Flat top PAM is the best for transmission because we can easily remove the noise and we can also easily recognize the noise. When we compare the difference between the flat top PAM and natural PAM, flat top PAM principle of sampling uses sample and hold circuit. In natural principle of sampling, noise interference is minimum. But in flat top PAM noise interference maximum. Flat top PAM and natural PAM are practical and sampling rate satisfies the sampling criteria.



Figur.4.11. Pulse Amplitude Modulation

4.8. PAM TDM system:

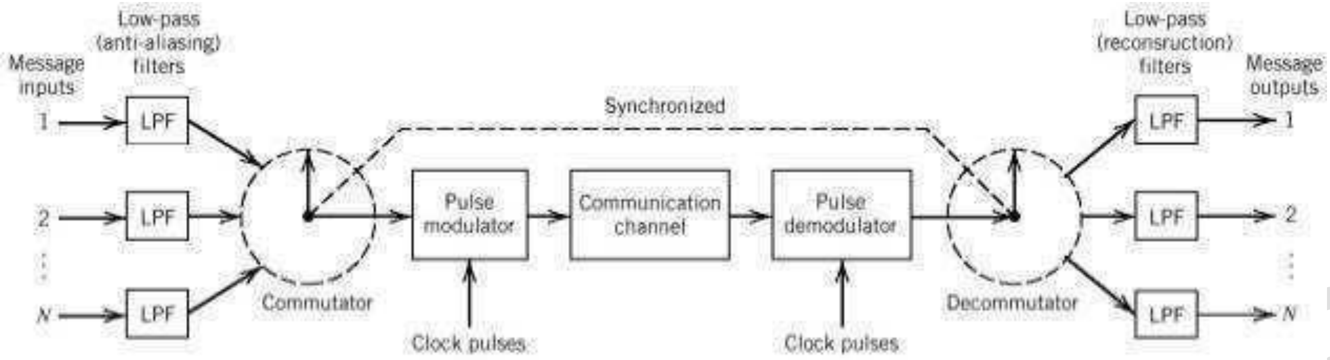


Figure.4.12. PAM TDM conceptual circuit

4.9. Pulse Duration Modulation (PDM) or Pulse Width Modulation (PWM):

It is a type of analog modulation. In pulse width modulation or pulse duration modulation, the width of the pulse carrier is varied in accordance with the sample values of message signal or modulating signal or modulating voltage. In pulse width modulation, the amplitude is made constant and width of pulse and position of pulse is made proportional to the amplitude of the signal. We can vary the pulse width in three ways:

1. By keeping the leading edge constant and vary the pulse width with respect to leading edge
2. By keeping the tailing constant.
3. By keeping the center of the pulse constant.

We can generate pulse width using different circuitry. In practical, we use 555 Timer which is the best way for generating the pulse width modulation signals. By configuring the 555 timer as mono stable or a stable multi vibrator, we can generate the PWM signals. We can use PIC, 8051, AVR, ARM, etc. microcontrollers to generate the PWM signals. PWM signal generation has n number of ways. In demodulation, we need PWM detector and its related circuitry for demodulating the PWM signal. The waveforms are shown in the figure.

4.10. Pulse Position Modulation (PPM): In the pulse position modulation, the position of each pulse in a signal by taking the reference signal is varied according to the sample value of message or modulating signal instantaneously. In the pulse position modulation, width and amplitude is kept constant. It is a technique that uses pulses of the same breath and height but is displaced in time from some base position according to the amplitude of the signal at the time of sampling. The position of the pulse is 1:1 which is propositional to the width of the pulse and also propositional to the instantaneous amplitude of sampled modulating signal. The position of pulse position modulation is easy when compared to other modulation. It requires pulse width generator and mono stable multi vibrator.

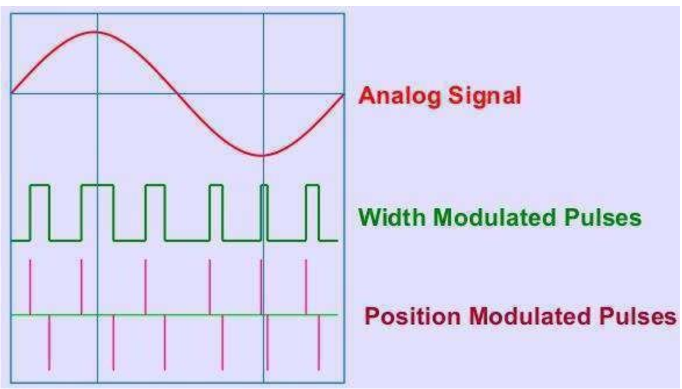


Figure.4.13. PWM and PPM

Pulse width generator is used for generating pulse width modulation signal which will help to trigger the mono stable multi vibrator; here trial edge of the PWM signal is used for triggering the mono stable multi vibrator. After triggering the mono stable multi vibrator, PWM signal is converted into pulse position modulation signal. For demodulation, it requires reference pulse generator, flip-flop and pulse width modulation demodulator.

4.11. Quantization:

In the process of quantization we create a new signal $m_q(t)$, which is an approximation to $m(t)$. The quantized signal $m_q(t)$, has the great merit that it is separable from the additive noise.

The operation of quantization is represented in figure 4.14. Here we have a signal m(t), whose amplitude varies in the range from V_H to V_L as shown in the figure.

We have divided the total range in to M equal intervals each of size S, called the step size and given by

$$S = \Delta = \frac{(V_H - V_L)}{M}$$

In our example M=8. In the centre of each of this step we located quantization levels $m0, m1, m2, \ldots m7$. The $m_q(t)$ is generated in the following manner-

Whenever the signal $m(t)$ is in the range Δ_0, the signal $m_q(t)$ maintains a constant level m_0, whenever the signal $m(t)$ is in the range Δ_1, the signal $m_q(t)$ maintains a constant level m_1 and so on. Hence the signal $m_q(t)$ will found all times to one of the levels $m0, m1, m2, \ldots m7$. The transition in $m_q(t)$ from $m0$ to $m1$ is made abruptly when $m(t)$ passes the transition level L_{01}, which is mid way between m_0 and m_1 and so on.

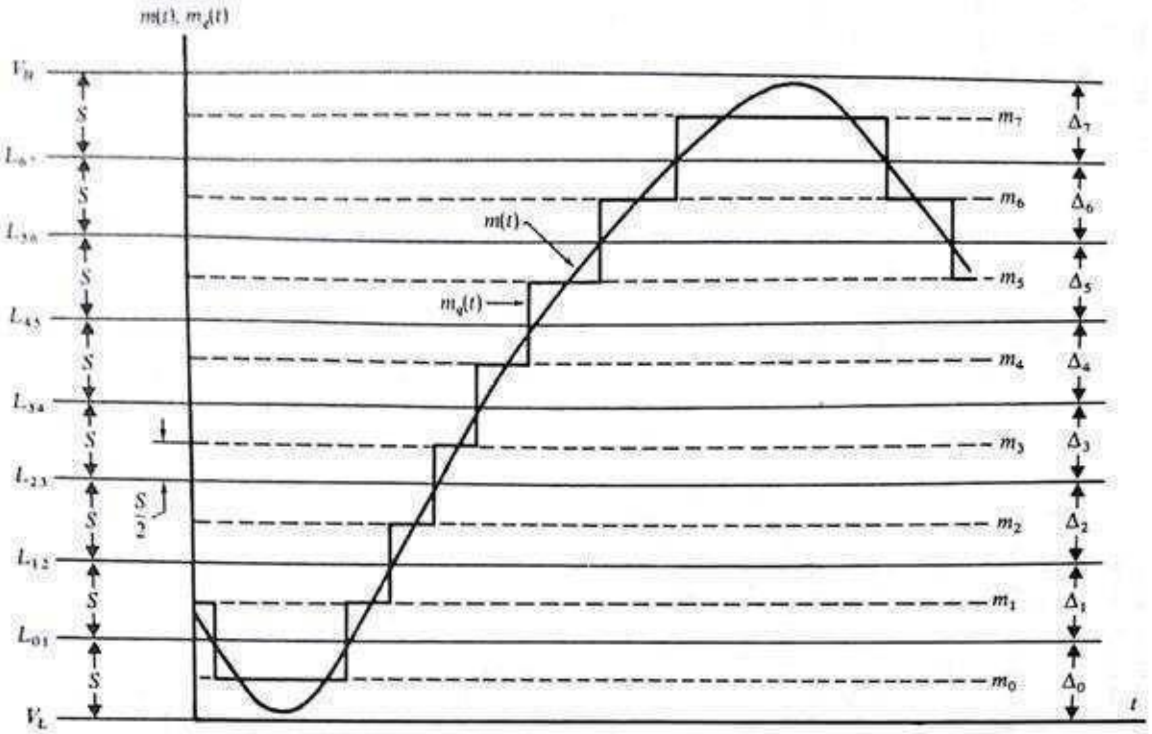


Figure.4.14. Quantization Process

Using quantization of signals, the effect of noise can be reduced significantly. The difference between $m(t)$ and $m_q(t)$ can be regarded as noise and is called quantization noise.

$$quantzaton\ nose = m(t) - m_q(t)$$

Also the quantized signal and original signal differs from one another in a ransom manner. This difference or error due to quantization process is called quantization error and is given by

$$e = m(t) - m_k$$

When $m(t)$ happens to be close to quantization level m_k, quantizer output will be m_k.

We can define the variable Δv to be the height of the each of the L levels of the quantizer as shown above. This gives a value of Δv equal to

$$\Delta v = \frac{2m_k}{L}$$

Therefore, for a set of quantizers with the same m_k, the larger the number of levels of a quantizer, the smaller the size of each quantization interval, and for a set of quantizers with the same number of quantization intervals, the larger m_p is the larger the quantization interval length to accommodate all the quantization range. The process of transforming sampled amplitude values of a message signal into a discrete amplitude value is referred to as Quantization. The quantization Process has a two-fold effect:

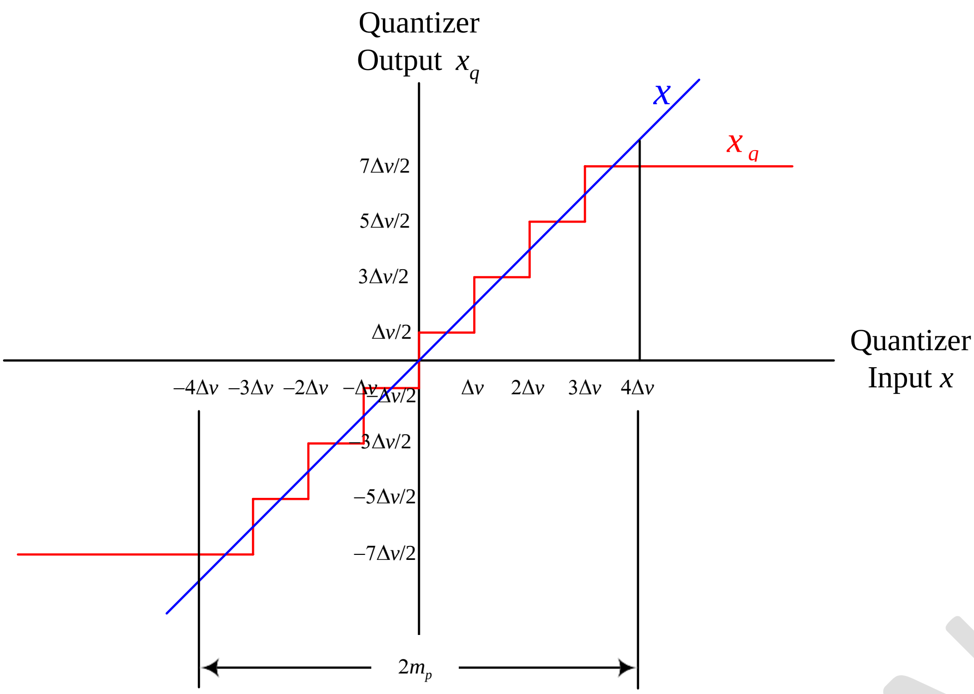
1. The peak-to-peak range of the input sample values is subdivided into a finite set of decision levels or decision thresholds that are aligned with the risers of the staircase, and
2. The output is assigned a discrete value selected from a finite set of representation levels that are aligned with the treads of the staircase.

A quantizer is memory less in that the quantizer output is determined only by the value of a corresponding input sample, independently of earlier analog samples applied to the input.

4.11.1 Quantization error:

Both sampling and quantization results 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 two adjacent representation levels is called a quantum or step-size.

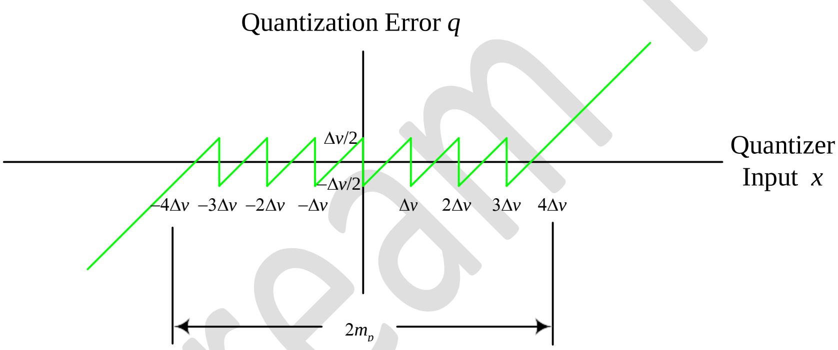
Now if we look at the input output characteristics of the quantizer, it will be similar to the red line in the following figure. Note that as long as the input is within the quantization range of the quantizer, the output of the quantizer represented by the red line follows the input of the quantizer. When the input of the quantizer exceeds the range of $-m_p$ to m_p, the output of the quantizer starts to deviate from the input and the quantization error (difference between an input and the corresponding output sample) increases significantly.



Now let us define the quantization error represented by the difference between the input sample and the corresponding output sample to be q, or

$$q = x - x_q$$

Plotting this quantization error versus the input signal of a quantizer is seen next. Notice that the plot of the quantization error is obtained by taking the difference between the blow and red lines in the above figure.

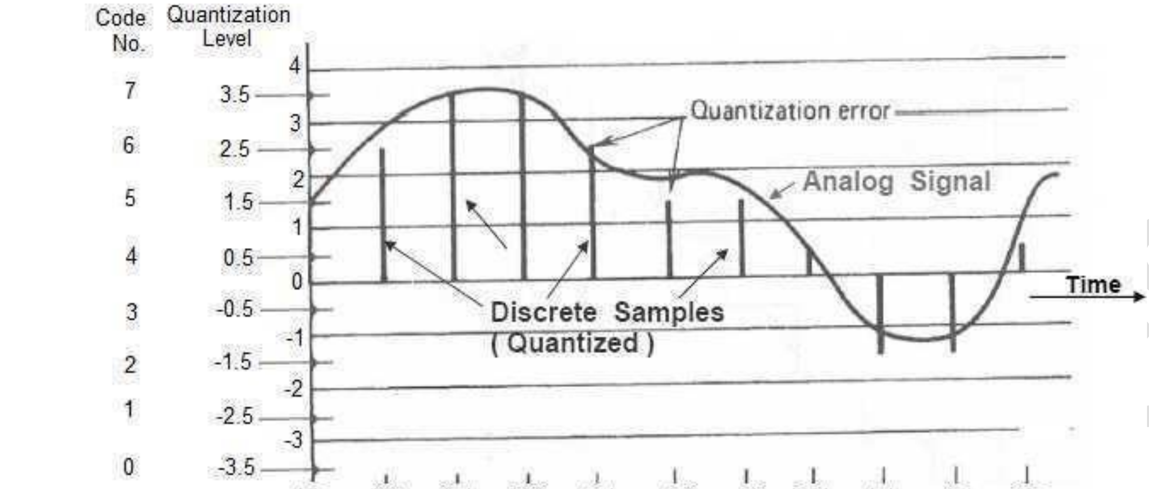


It is seen from this figure that the quantization error of any sample is restricted between $-\Delta v/2$ and $\Delta v/2$ except when the input signal exceeds the range of quantization of $-m_p$ to m_p.

4.12. Pulse Code Modulation (PCM) Transmitter:

A signal which is to be quantized before transmission is sampled as well. The quantization is used to reduce the effect of noise and the sampling allows us to do the time division multiplexing. The combined operation of sampling and quantization generate a quantized PAM waveform i.e. a train of pulses whose amplitude is restricted to a number of discrete levels.

Rather than transmitting the sampled values itself, we may represent each quantization level by a code number and transmit the code number. Most frequently the code number is converted in to binary equivalent before transmission. Then the digits of the binary representation of the code are transmitted as pulses. This system of transmission is called binary Pulse Code Modulation. The whole process can be understood by the following diagram.



Sample Value	1.5	2.8	3.4	3.3	2.4	1.8	1.6	0.4	-1.4	-1.4	0.5
Nearest Quantization Level	1.5	2.5	3.5	3.5	2.5	1.5	1.5	0.5	-1.5	-1.5	0.5
Code No.	5	6	7	7	6	5	5	4	2	2	4
Binary Representation	101	110	111	111	110	101	101	100	010	010	100

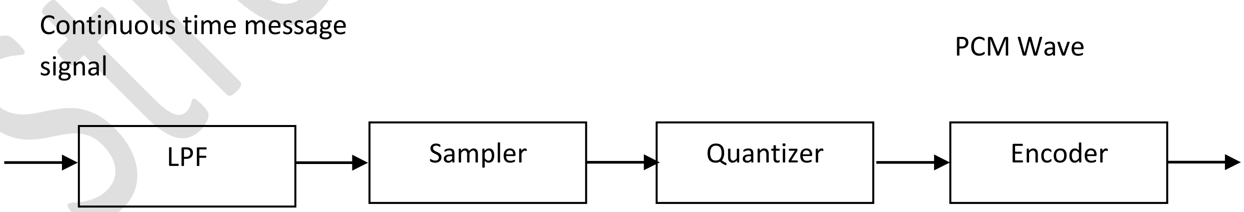
Figure.4.15. PCM Process

Basic Blocks:

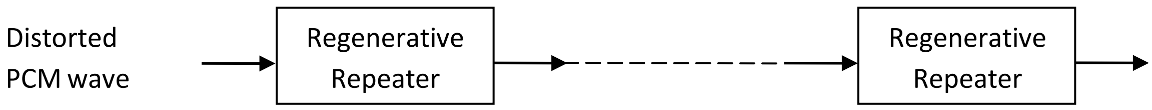
1. Anti aliasing Filter, 2. Sampler, 3. Quantizer, 4. Encoder

The block diagram of a PCM transmitter is shown in figure 4.16. An anti-aliasing filter is basically a filter used to ensure that the input signal to sampler is free from the unwanted frequency components. For most of the applications these are low-pass filters. It removes the frequency components of the signal which are above the cutoff frequency of the filter. The cutoff frequency of the filter is chosen such it is very close to the highest frequency component of the signal.

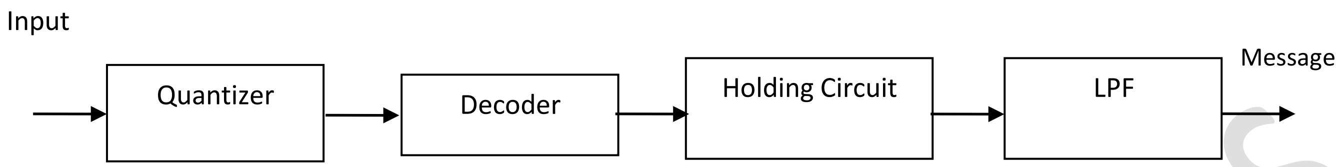
The message signal is sampled at the Nyquist rate by the sampler. The sampled pulses are then quantized by the quantizer. The encoder encodes these quantized pulses in to binary equivalent, which are then transmitted over the channel. During the channel the regenerative repeaters are used to maintain the signal to noise ratio.



(A) Transmitter



(b) Transmission Path



(c) Receiver

Figure.4.16. PCM System Basic Block Diagram

Figure (c) shows the receiver. The first block is again the quantizer, but this quantizer is different from the transmitter quantizer as it has to take the decision regarding the presence or absence of the pulse only. Thus there are only two quantization levels. The output of the quantizer goes to the decoder which is an D/A converter that performs the inverse operation of the encoder. The decoder output is a sequence of quantized pulses. The original signal is reconstructed in the holding circuit and the LPF.

4.13. Types of Quantizer: Uniform and Non-uniform quantization

4.13.1. Uniform Quantizer: In Uniform type, the quantization levels are uniformly spaced, where as in non-uniform type the spacing between the levels will be unequal and mostly the relation is logarithmic.

Types of Uniform Quantizer: (based on I/P - O/P Characteristics)

1. Mid-Rise type Quantizer
2. Mid-Tread type Quantizer

In the stair case like graph, the origin lies the middle of the tread portion in Mid –Tread type where as the origin lies in the middle of the rise portion in the Mid-Rise type.

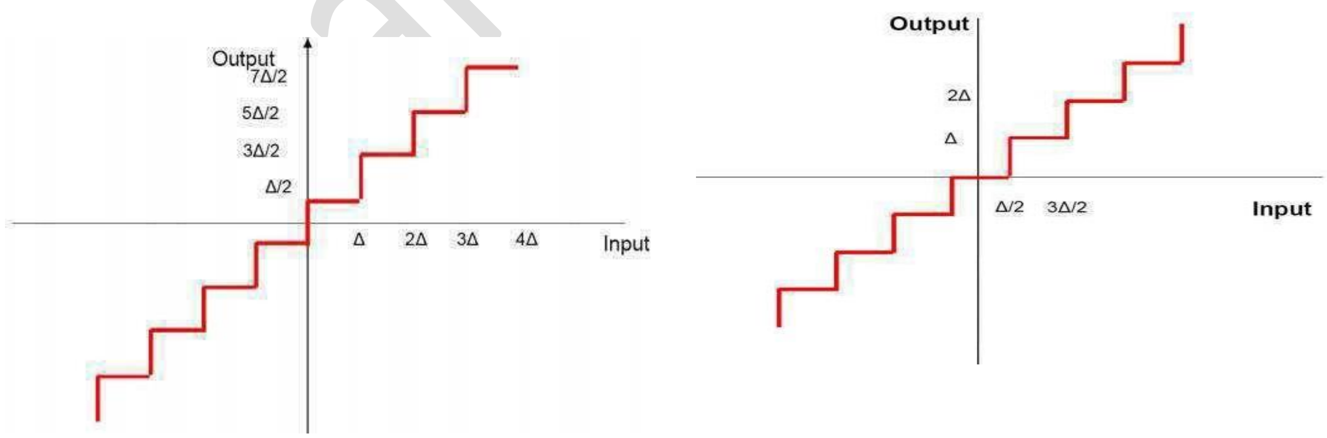


Figure.4.17 (a) Mid – Rise type: Quantization levels: even number (b) Mid – tread type :Quantization levels: odd number

4.13.2. Non-Uniform Quantizer Companding:

The word Companding is a combination of Compressing and Expanding, which means that it does both. This is a non-linear technique used in PCM which compresses the data at the transmitter and expands the same data at the receiver. The effects of noise and crosstalk are reduced by using this technique. In Non - Uniform Quantizer the step size varies. The use of a non – uniform quantizer is equivalent to passing the baseband signal through a compressor and then applying the compressed signal to a uniform quantizer. The resultant signal is then transmitted.

Figure.4.18. Non-uniform quantizer

At the receiver, a device with a characteristic complementary to the compressor called Expander is used to restore the signal samples to their correct relative level. The Compressor and expander take together constitute a Compander.

- The compressor will compress the dynamic range of the signal so that high dynamic range signal can be passed through components of low dynamic range capability, the uniform quantizer will undergo the quantization process of the compressed signal and the lastly the expander will undergo expansion and invert the compression function to reconstruct the original signal.
- The expander has complementary characteristics as that of compressor so that the compressor input is equal to expander output in order to reproduce the signal at the receiver.
- The Figure.4.19 below illustrates the input-output characteristics and curves of the companding process, and it can be seen that companding has linear characteristics.

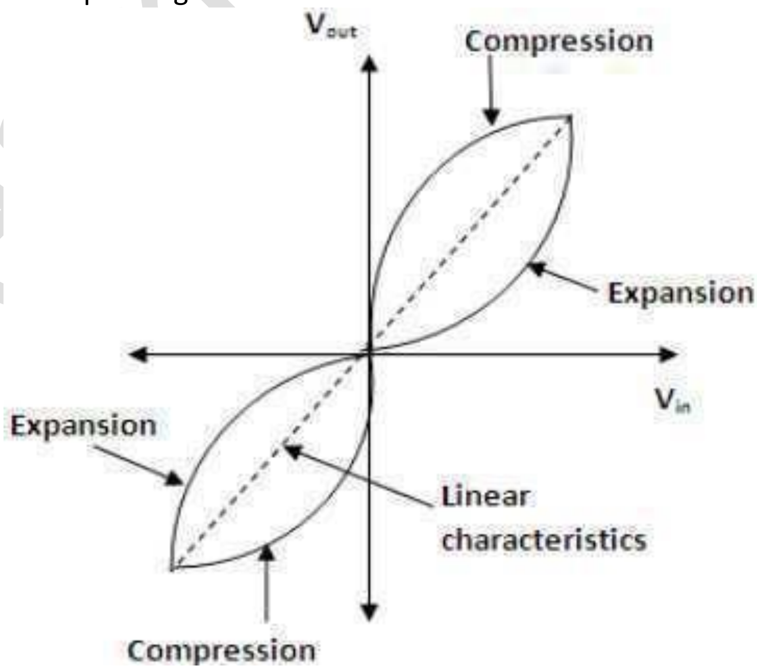


Figure.4.19. Input-output characteristics of Companding

Advantages of Non- Uniform Quantization:

1. Higher average signal to quantization noise power ratio than the uniform quantizer when the signal is non uniform which is the case in many practical situations.

2. RMS value of the quantizer noise power of a non – uniform quantizer is substantially proportional to the sampled value and hence the effect of the quantizer noise is reduced.

There are two types of Companding techniques.

- μ-law: North America and Japan. For $\mu = 255$ (for 8-bit codes),

$$y = \text{sgn}(x)\frac{1}{\ln(1+\mu)}\ln(1+\mu|x|), \quad (0 < x < 1)$$

- A-law: Europe, rest of world.

$$y = \begin{cases} \text{sgn}(x)\dfrac{A|x|}{1+\ln(A)} & |x| < \dfrac{1}{A} \\ \text{sgn}(x)\dfrac{1+\ln(A|x|)}{1+\ln(A)} & \dfrac{1}{A} < |x| < 1 \end{cases}$$

The standard value is $A = 87.7$.

For both laws, the input to the compressor is

$$x = \frac{m(t)}{m_p}$$

where $-m_p \le m(t) \le m_p$.

A-law Companding Technique

- Uniform quantization is achieved at A = 1, where the characteristic curve is linear and there is no compression.
- A-law has mid-rise at the origin. Hence, it contains a non-zero value.
- A-law companding is used for PCM telephone systems.
- A-law is used in many parts of the world.

µ-law Companding Technique

- Uniform quantization is achieved at µ = 0, where the characteristic curve is linear and there is no compression.
- µ-law has mid-tread at the origin. Hence, it contains a zero value.
- µ-law companding is used for speech and music signals.
- µ-law is used in North America and Japan

µ-law provides slightly larger dynamic range than A-law. A-law has smaller proportional distortion for small signals. A-law is used for international connections if at least one country uses it.

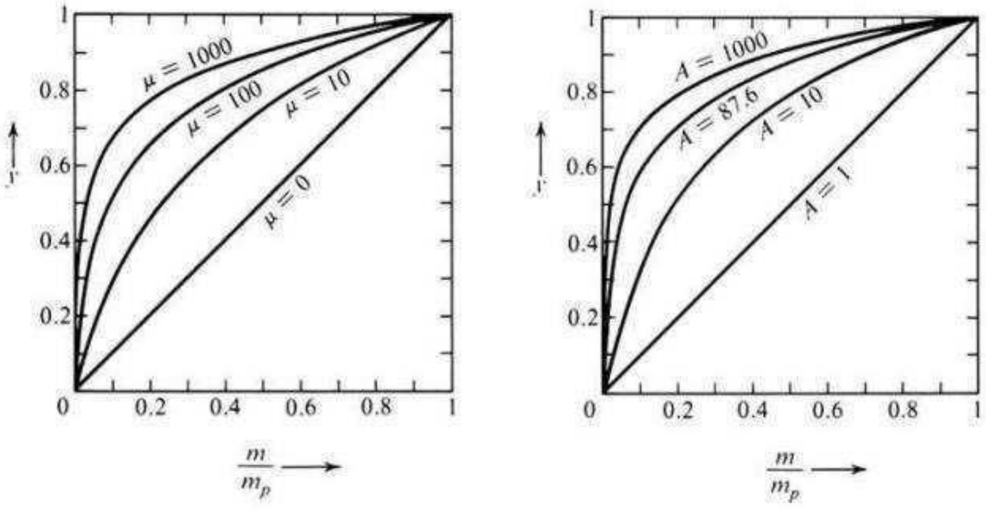


Figure.4.20. (a) μ-law curve (b) A-law curve

4.14. PCM Signal to noise ratio:

The signal to quantization noise ratio is given as

$$\text{SNR} = \frac{S}{N} = \frac{\textit{Normalzed snal poer}}{\textit{Normalzed nose poer}}$$

$$= \frac{\textit{Normalzed sgnal power}}{\frac{\Delta^2}{12}}$$

The number of quantization value is equal to: $q=2^{v}$

$$\Delta = \frac{2X_{\max}}{2^{v}}$$

$$\frac{S}{N_q} = \frac{\text{Normalized signal power}}{\left[\frac{2X_{\max}}{2^{v}}\right]^2 * \frac{1}{12}}$$

Let the normalized signal power is equal to P then the signal to quantization noise will be given by:

$$\boxed{\frac{S}{N_q} = \frac{3P * 2^{2v}}{X_{\max}^{2}}}$$

4.15. Data rate: Bit rate is typically seen in terms of the actual data rate.

Bit Rate: The speed of the data is expressed in bits per second (bits/s or bps). The data rate R is a function of the duration of the bit or bit time (T_B). $R = 1/T_B$ .Rate is also called channel capacity C.

Baud rate: Baud rate refers to the number of signal or symbol changes that occur per second. A symbol is one of several voltage, frequency, or phase changes. If N is the number of bits per symbol, then the number of required symbols is $S = 2^N$. Thus, the gross bit rate is:

R = baud rate x $\log_2 S$ = baud rate x 3.32 $\log_{10} S$

Multiplexed PCM signal:

When a large number of PCM signals are to be transmitted over a common channel, multiplexing of these PCM signals is required. Figure shows the basic time division multiplexing scheme, called as the PCM multiplexed digital system. This system has been designed to accommodate N voice and each signal is band limited to F_m kHz, and the sampling is done at a standard rate of $2F_m$ kHz. This is Nyquist rate. The sampling is done by the commutator switch SW1. These voice signals are selected one by one and connected to a PCM transmitter by the commutator switch SW1. Each sampled signal is then applied to the PCM transmitter which converts it into a digital signal by the process of A to D conversion and companding. The resulting digital waveform is transmitted over a co-axial cable. Periodically, after every constant distance, the PCM-TDM signal is regenerated by amplifiers called "repeaters". They eliminate the distortion introduced by the channel and remove the superimposed noise and regenerate a clean PCM-TDM signal at their output. This ensures that the received signal is free from the distortions and noise. At the destination the signal is companded, decoded and demultiplexed, using a PCM receiver. The PCM receiver output is connected to different low pass filters via commutator switch SW2. Synchronization between the transmitter and receiver commutator SW1 and SW2 is essential in order to ensure proper communication.

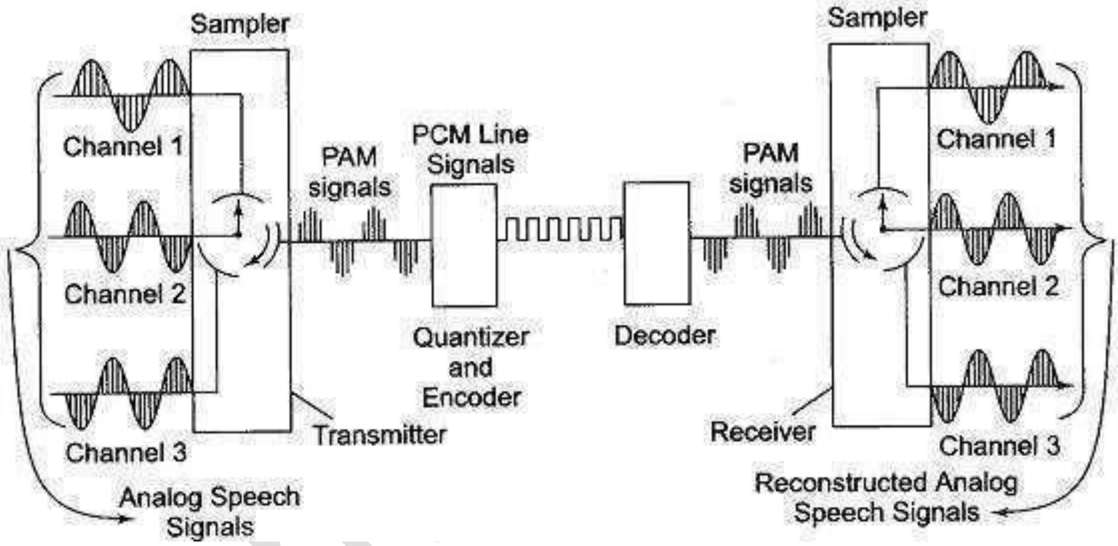


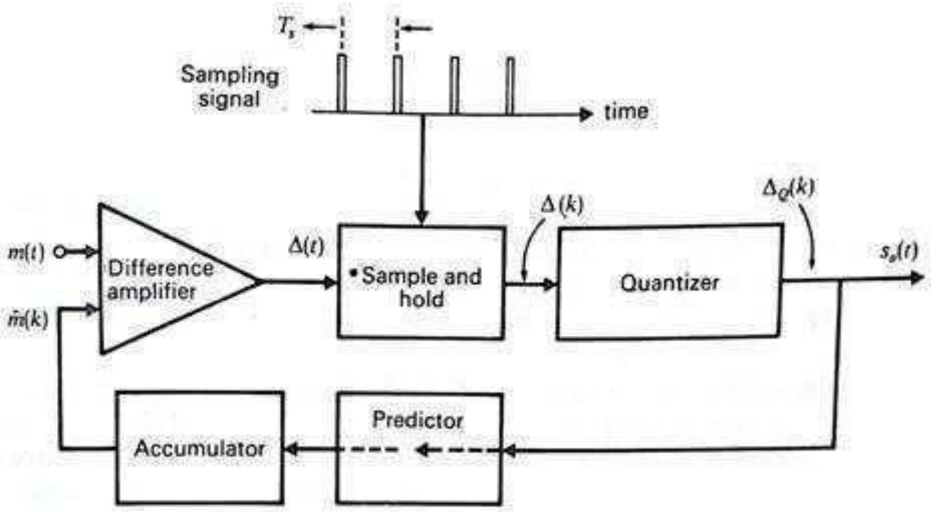
Figure.4.20. Multiplexed PCM TDM system

4.16. Differential Pulse Code Modulation (DPCM):

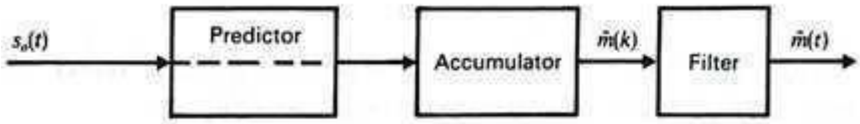
In DPCM instead of transmitting the sampled values itself at each sampling time; we can transmit the difference between the two successive samples. If such changes are transmitted then at the receiving end we can generate a waveform identical to the m(t) by simply adding up these changes.

The DPCM has the special merit that when these differences are transmitted by PCM. The differences $m(k) - m(k-1)$ will be smaller than the sample values them and fewer levels will be required to quantize $m(k)$, and corresponding fewer bits will be needed to encode the signal. The basic principle of DPCM is shown in figure.

The receiver consists of an accumulator which adds-up the receiver quantized differences $\Delta(k)$ and a filter which smoothes out the quantization noise. The output of accumulator is the signal approximation $\hat{m}(k)$ which becomes $\hat{m}(t)$ at the filter output.



(a) Transmitter



(b) Receiver

Figure.4.21. Differential PCM

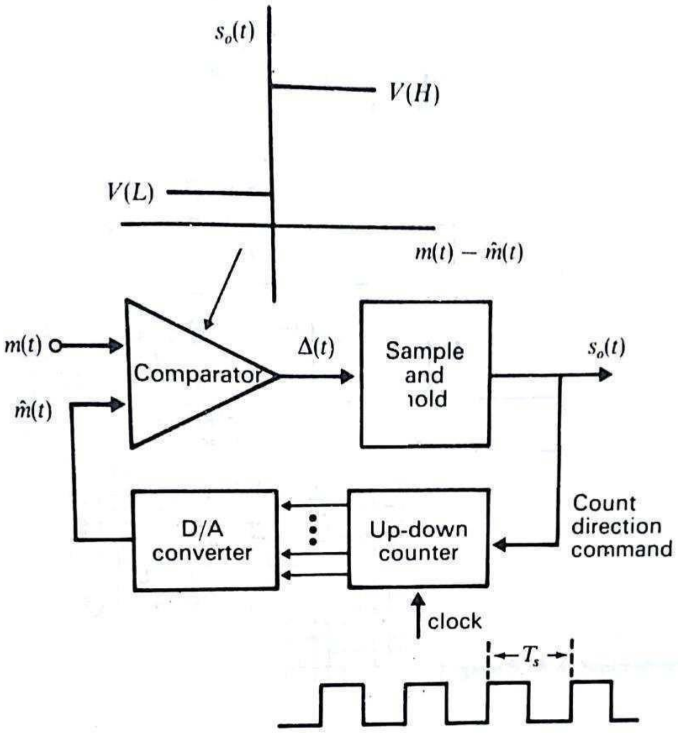
At the transmitter we need to know whether the $\hat{m}(t)$ is larger or smaller than $m(t)$ and by how much amount. We may than determine whether the next difference $\Delta(k)$ needs to be positive or negative and of what amplitude in order to bring $\hat{m}(t)$ as close as possible to $m(t)$. For this reason we have a duplicate accumulator at transmitter.

At each sampling time the transmitter difference amplifier compares $m(t)$ and $\hat{m}(t)$, and the sample and hold circuit holds the result of that comparison $\Delta(t)$, for the duration of interval between sampling times. The quantizer generates the signal $S_0(t) = \Delta(k)$ both for the transmission to the receiver and to provide the input to the receiver accumulator in the transmitter. The basic limitation of the DPCM scheme is that the transmitted differences are quantized and are of limited values.

4.17. Delta Modulation (DM):

Delta Modulation is a DPCM scheme in which the difference signal $\Delta(t)$ is encoded into just a single bit. The single bit providing just for two possibilities is used to increase or decrease the estimate $\hat{m}(t)[m_q(t)]$. The baseband signal $m(t)$ and its quantized approximation $\hat{m}(t)$ are applied as input to a comparator. The comparator has one fixed output V(H) when $m(t) > m_q(t)$ and a difference output V(L) when $m(t) < m_q(t)$. Ideally the transition between V(H) and V(L) is arbitrarily abrupt as $m(t) - m_q(t)$ passes through zero. The up-down counter increments or decrements its count by 1 at each active edge of the clock waveform. The count direction i.e. incrementing or decrementing is determined by the voltage levels at the "Count direction command" input to the counter. When this binary input is at level V(H), the counter counts up and when this binary input is at level V(L), the counter counts down.

The digital output of the counter is converted into analog quantized approximation $m_q(t)$ by a D/A converter. The waveforms for the delta modulator is shown in figure (b), assuming that the active clock edge is falling edge. The Linear Delta Modulator is shown in figure (a).



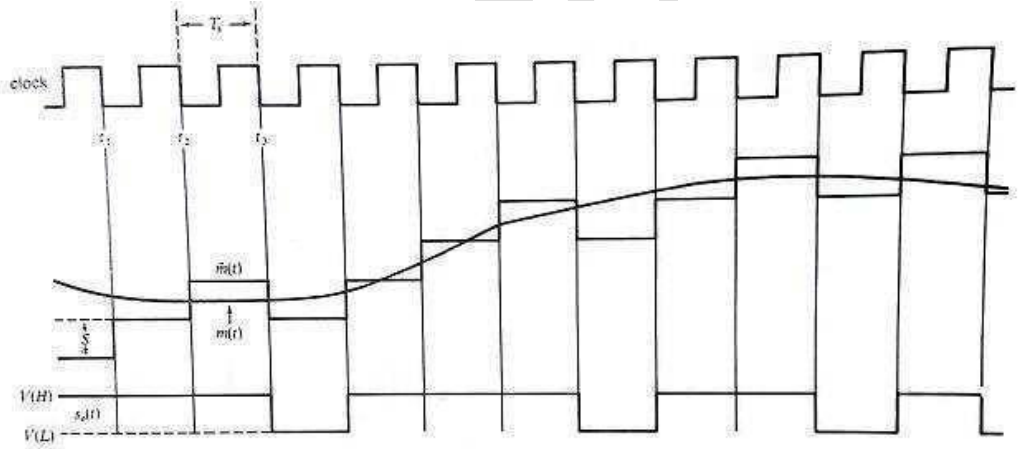
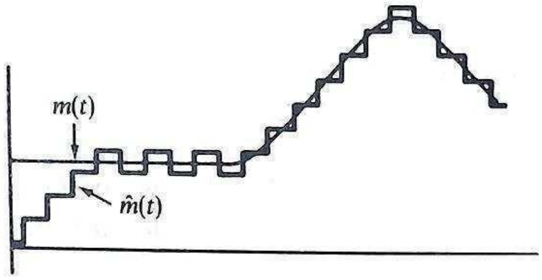


Figure.4.22. (a) Delta Modulator (b) The response of the delta modulator to a baseband signal m(t)

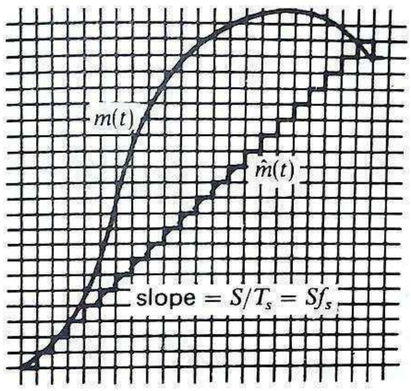
4.17.1. Delta Modulation and the slope overload error.

It may be noted that at startup there is a brief interval when the quantized signal may be a poor approximation to the baseband signal as shown in figure (a).

The initial large discrepancy between $m(t)$ and $m_q(t)$ and stepwise approach of $m_q(t)$ to $m(t)$ is shown in figure (b).



(a) Startup response of DM



(b) Slope Overload in a linear DM

Figure.4.23. Response of DM and slope overload distortion

It should be noted that when $m_q(t)$ has caught up $m(t)$ and even though $m(t)$ remains constant, $m_q(t)$ hunts, swinging up and down to $m(t)$.

4.17.2. Slope Overload distortion

The excessive disparity between $m(t)$ and $m_q(t)$ is described as a slope overload error and occurs whenever $m(t)$ has a slope larger than the slope S/T_s which can be sustained by the waveform $m_q(t)$. The slope overload as shown in figure (b) is developed due to the small size of S. To overcome the overload we have to increase the sampling rate above the rate initially selected to satisfy the Nyquist criterion. The sampling rate f_s must satisfy the following condition

$$sf_s = 2f_s = \pi f A$$

4.17.3. Features of DM: Following are some of the features of delta modulation.

- An over-sampled input is taken to make full use of the signal correlation.
- The quantization design is simple.
- The input sequence is much higher than the Nyquist rate.
- The quality is moderate.
- The design of the modulator and the demodulator is simple.
- The stair-case approximation of output waveform.
- The step-size is very small, i.e., Δ (delta).
- The bit rate can be decided by the user.
- This involves simpler implementation.

4.17.4. Advantages of DM Over DPCM

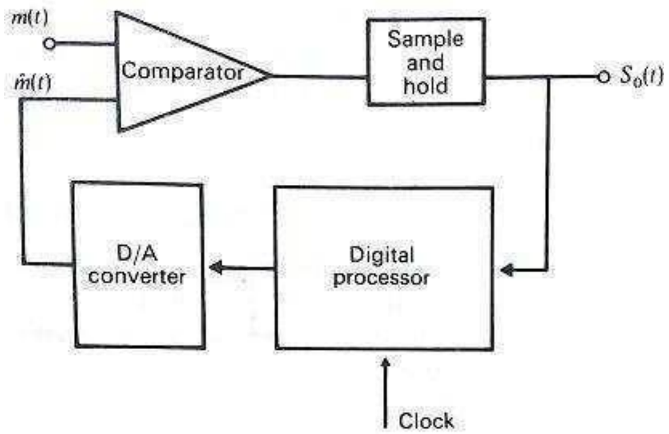
- 1-bit quantizer
- Very easy design of the modulator and the demodulator. However, there exists some noise in DM.
- Slope Over load distortion (when Δ is small) & Granular noise (when Δ is large)

4.18. Adaptive Delta Modulation (ADM).

In digital modulation, we have come across certain problem of determining the step-size, which influences the quality of the output wave.

A larger step-size is needed in the step slope of modulating signal and a smaller step size is needed where the message has a small slope. The minute details get missed in the process. So, it would be better if we can control the adjustment of step-size, according to our requirement in order to obtain the sampling in a desired fashion. This is the concept of Adaptive Delta Modulation.

Following is the



a: Adaptive Delta Modulation (ADM)

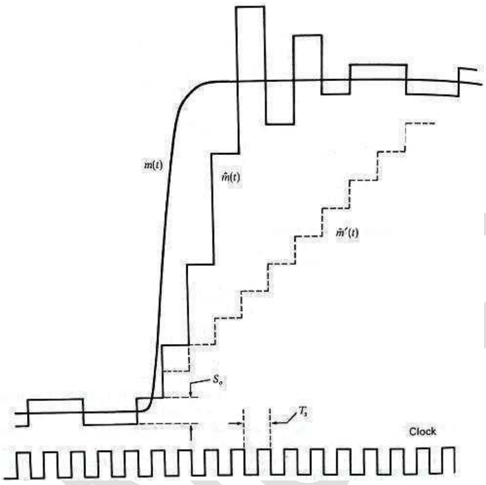


Figure b: Waveforms comparin response of DM and ADM

Figure.4.24. Block diagram of Adaptive delta modulator

The step size S is not of fixed size but it is always a multiple of basic step size S_0. The basic step size S_0 is either added or subtracted by the accumulator as required to move $m_q(t)$ more close to $m(t)$. If the direction of the step at the clock edge K is same as at edge K-1, then the processor increases the step size by an amount S_0. If the directions are opposite then the processor decreases the magnitude of the step by S_0.

In figure (a), the output $S_0(t)$ is called $\mathrm{e}(k)$, which represents the error i.e. the discrepancy between the $m(t)$ and $m_q(t)$, and it is either V(H) or V(L).

$\mathrm{e}(k) = +1, f\ m(t) > m_q(t)$ immediately before Kth edge,

$\mathrm{e}(k) = -1, f\ m(t) < m_q(t)$ immediately before Kth edge,

The features of ADM are shown in figure (b) As long as the condition $m(t) > m_q(t)$ persists the jumps in $m_q(t)$ becomes larger, that's why $m_q(t)$ catches up with $m(t)$ sooner than in the case of linear DM, as shown by $m'_q(t)$.

On the other hand, when the response to the large slope in $m(t)$, $m_q(t)$ develops large jumps and large number of clock cycles are required for these jumps to settle down. Therefore the ADM system reduces the slope overload but it increases the quantization error. Also when $m(t)$ is constant $m_q(t)$ oscillates about $m(t)$ but the oscillation frequency is half of the clock frequency.

4.18. Comparison of various systems:

S.NO	Parameter of Comparison	Pulse Code Modulation (PCM)	Delta Modulation (DM)	Adaptive Delta Modulation (ADM)	Differential Pulse Code Modulation (DPCM)
1.	Number of bits	It can use 4,8, or 16 bits per sample.	It uses only one bit for one sample	It uses only one bit for one sample	Bits can be more than one but are less than PCM.
2.	Levels and step size	The number of levels depends on number of bits. Level size is fixed.	Step size is kept fixed and cannot be varied.	According to the signal variation, step size varies.	Number of levels is fixed.
3.	Quantization error and distortion	Quantization error depends on number of levels used.	Slope overload distortion and granular noise are present.	Quantization noise is present but other errors are absent.	Slope overload distortion and quantization noise is present.
4.	Transmission bandwidth	Highest bandwidth is required since numbers of bits are high.	Lowest bandwidth is required.	Lowest bandwidth is required.	Bandwidth required is less than PCM.
5.	Feedback	There is no feedback in transmitter or receiver.	Feedback exists in transmitter.	Feedback exists.	Feedback exists.
6.	Complexity of Implementation	System is complex.	Simple	Simple	Simple

Unit-5

Digital modulations techniques, Generation, detection, equation and Bandwidth of amplitude shift keying (ASK) Binary Phase Shift keying (BPSK), Differential phase shift keying (DPSK), offset and non offset quadrature phase shift keying (QPSK), M-Ary PSK, Binary frequency Shift Keying (BFSK), M-Ary FSK Quadrature Amplitude modulation (QAM)

5.1. Digital Modulation

Digital Modulation provides more information capacity, high data security, quicker system availability with great quality communication. Hence, digital modulation techniques have a greater demand, for their capacity to convey larger amounts of data than analog modulation techniques.

There are many types of digital modulation techniques and also their combinations, as listed below.

ASK – Amplitude Shift Keying

The amplitude of the resultant output depends upon the input data whether it should be a zero level or a variation of positive and negative, depending upon the carrier frequency.

FSK – Frequency Shift Keying

The frequency of the output signal will be either high or low, depending upon the input data applied.

PSK – Phase Shift Keying

The phase of the output signal gets shifted depending upon the input. These are mainly of two types, namely Binary Phase Shift Keying (BPSK) and Quadrature Phase Shift Keying (QPSK), according to the number of phase shifts. The other one is Differential Phase Shift Keying (DPSK) which changes the phase according to the previous value.

M-ary Encoding

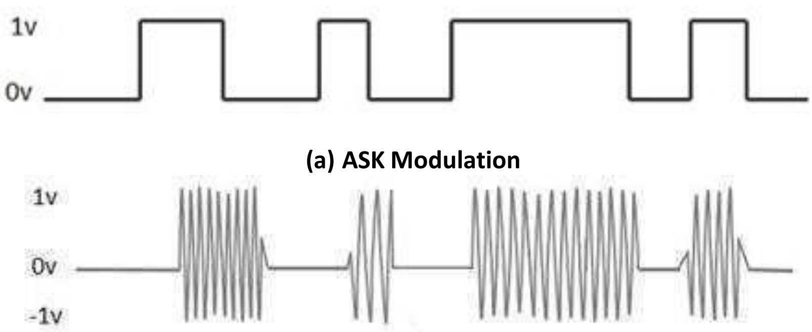
M-ary Encoding techniques are the methods where more than two bits are made to transmit simultaneously on a single signal. This helps in the reduction of bandwidth.

The types of M-ary techniques are M-ary ASK, M-ary FSK & M-ary PSK.

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

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

The figure 3.1.1 represents ASK modulated waveform along with its input.



(b) ASK Modulated Wave

Figure 5.1. ASK Modulation

To find the process of obtaining this ASK modulated wave, let us learn about the working of the ASK modulator.

5.2.1. ASK Modulator

The ASK modulator block diagram comprises of the carrier signal generator, the binary sequence from the message signal and the band-limited filter. Following is the block diagram of the ASK Modulator.

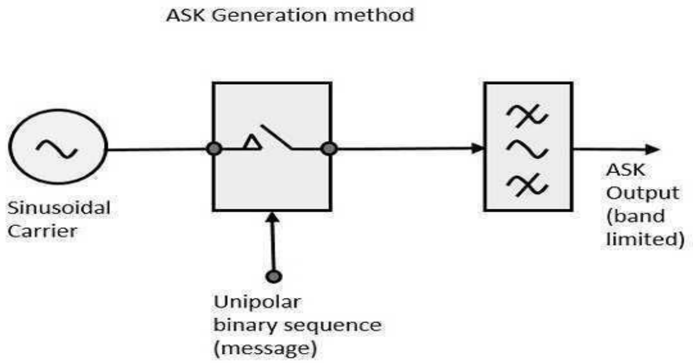


Figure 5.2. ASK Modulator

The carrier generator sends a continuous high-frequency carrier. The binary sequence from the message signal makes the unipolar input to be either High or Low. The high signal closes the switch, allowing a carrier wave. Hence, the output will be the carrier signal at high input. When there is low input, the switch opens, allowing no voltage to appear. Hence, the output will be low.
The band-limiting filter, shapes the pulse depending upon the amplitude and phase characteristics of the band-limiting filter or the pulse-shaping filter.

5.2.2. ASK Demodulator

There are two types of ASK Demodulation techniques. They are –

- Asynchronous ASK Demodulation/detection
- Synchronous ASK Demodulation/detection

The clock frequency at the transmitter when matches with the clock frequency at the receiver, it is known as a Synchronous method, as the frequency gets synchronized. Otherwise, it is known as Asynchronous.

5.2.3. Asynchronous ASK Demodulator

The Asynchronous ASK detector consists of a half-wave rectifier, a low pass filter, and a comparator. Following is the block diagram for the same.

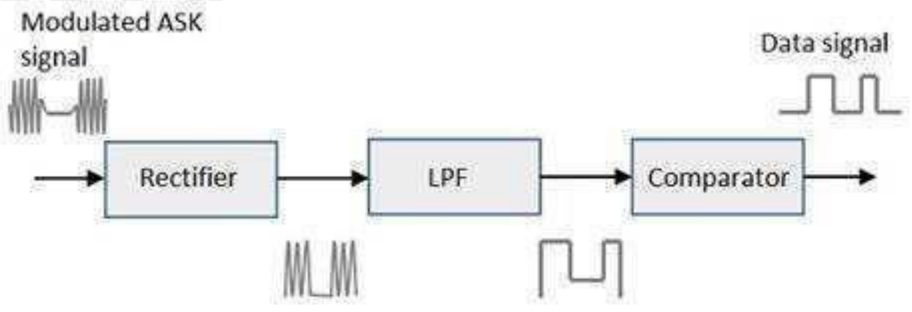


Figure 5.3. ASK Demodulator

The modulated ASK signal is given to the half-wave rectifier, which delivers a positive half output. The low pass filter suppresses the higher frequencies and gives an envelope detected output from which the comparator delivers a digital output.

5.2.4. Synchronous ASK Demodulator

Synchronous ASK detector consists of a Square law detector, low pass filter, a comparator, and a voltage limiter. Following is the block diagram for the same.

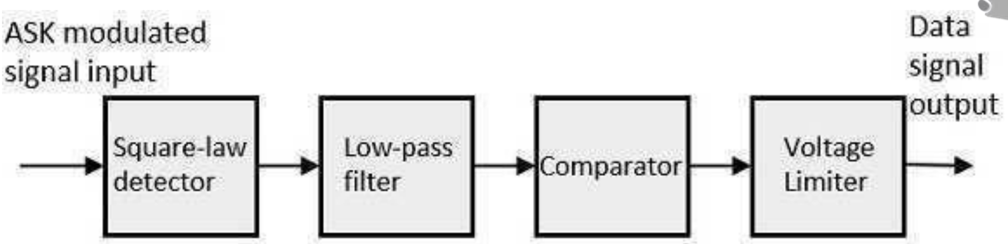


Figure 5.4. Synchronous ASK Demodulator

The ASK modulated input signal is given to the Square law detector. A square law detector is one whose output voltage is proportional to the square of the amplitude modulated input voltage. The low pass filter minimizes the higher frequencies. The comparator and the voltage limiter help to get a clean digital output.

5.3. Frequency Shift Keying

Frequency Shift Keying (FSK) is the digital modulation technique in which the frequency of the carrier signal varies according to the digital signal changes. FSK is a scheme of frequency modulation. The output of a FSK modulated wave is high in frequency for a binary High input and is low in frequency for a binary Low input. The binary 1s and 0s are called Mark and Space frequencies. The following image is the diagrammatic representation of FSK modulated waveform along with its input.

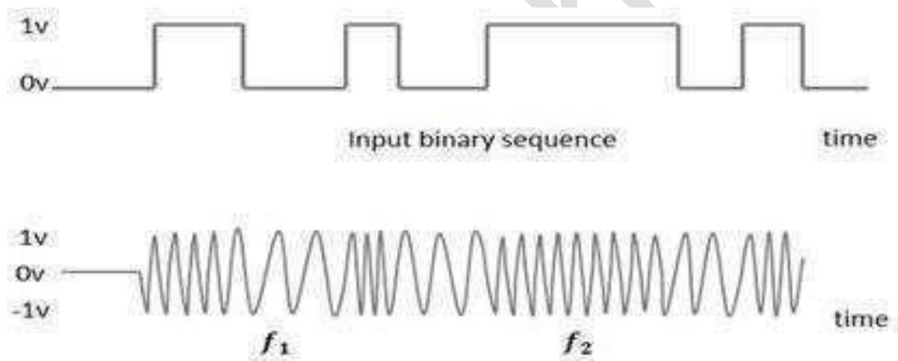


Figure 5.5. Frequency Shift Keying (FSK)

To find the process of obtaining this FSK modulated wave, let us know about the working of a FSK modulator.

5.3.1. FSK Modulator

Binary Frequency Shift Keying (BFSK)

In binary frequency-shift keying (BFSK) the binary data waveform d(t) generates a binary signal

$$v_{BFSK}(t) = \sqrt{2P_s}\cos[\omega t + d(t)\Omega t] \qquad \text{...3.7.1}$$

Here d(t) = + 1 or -1 corresponding to the logic levels 1and 0 of the data waveform. The transmitted signal is of amplitude $\sqrt{2P_s}$ and is either

$$v_{BFSK}(t) = S_H(t) = \sqrt{2P_s}\cos(\omega + \Omega)\,t \qquad \text{...3.7.2}$$

$$v_{BFSK}(t) = S_L(t) = \sqrt{2P_s}\cos(\omega - \Omega)\,t \qquad \text{...3.7.3}$$

and thus has an angular frequency $\omega_o + \Omega$ or $\omega_o - \Omega$ with Ω a constant offset from the nominal carrier frequency ω_o. We shall call the higher frequency $\omega_H(= \omega_o + \Omega)$ and the lower frequency $\omega_L.(= \omega_o - \Omega)$. We may conceive that the BFSK signal is generated in the manner indicated in Fig. 3.7.1. Two balanced modulators are used, one with carrier ω_H and one with carrier ω_L. The voltage values of $P_H(t)$ and of $P_L(t)$ are related to the voltage values of d(t) in the following manner

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

Thus when d(t) changes from +1 to -1 P_H changes from 1 to 0 and P_L from 0 to 1. At any time either P_H or P_L is 1 but not both so that the generated signal is either at angular frequency ω_H or at ω_L.

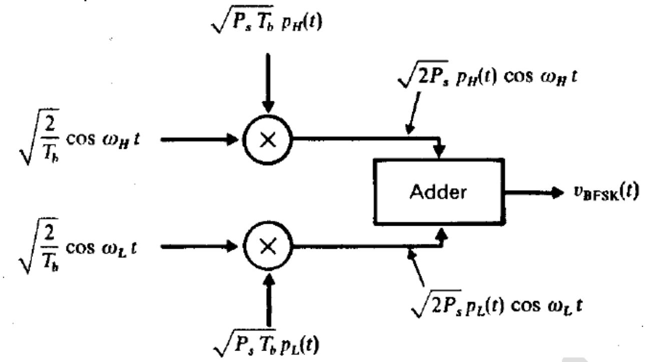


Figure 5.6. A representation of a manner in which a BFSK signal can be generated.

The two oscillators, producing a higher and a lower frequency signals, are connected to a switch along with an internal clock. To avoid the abrupt phase discontinuities of the output waveform during the transmission of the message, a clock is applied to both the oscillators, internally. The binary input sequence is applied to the transmitter so as to choose the frequencies according to the binary input.

5.3.2. FSK Demodulator

There are different methods for demodulating a FSK wave. The main methods of FSK detection are asynchronous detector and synchronous detector. The synchronous detector is a coherent one, while asynchronous detector is a non-coherent one.

5.3.3. Asynchronous FSK Detector

The block diagram of Asynchronous FSK detector consists of two band pass filters, two envelope detectors, and a decision circuit. Following is the diagrammatic representation.

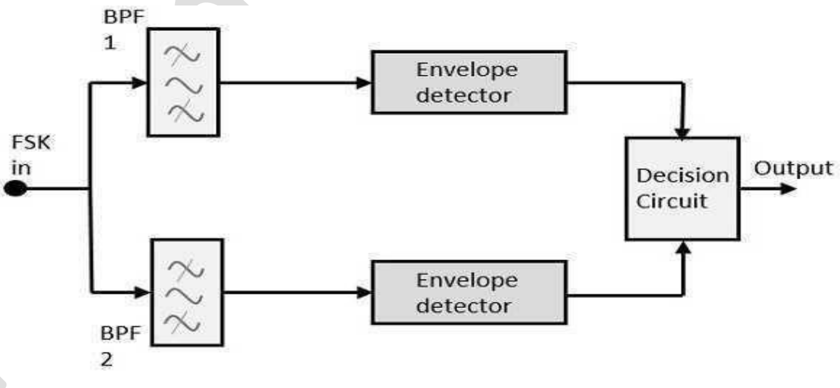


Figure 5.7. Asynchronous FSK Detector

The FSK signal is passed through the two Band Pass Filters (BPFs), tuned to Space and Mark frequencies. The output from these two Band Pass Filters looks like ASK signal; which is then applied to the envelope detector. The signal in each envelope detector is modulated asynchronously.

The decision circuit chooses which output is more likely and selects it from any one of the envelope detectors. It also re-shapes the waveform to a rectangular one.

5.3.4. Synchronous FSK Detector

The block diagram of Synchronous FSK detector consists of two mixers with local oscillator circuits, two band pass filters and a decision circuit. Following is the diagrammatic representation.

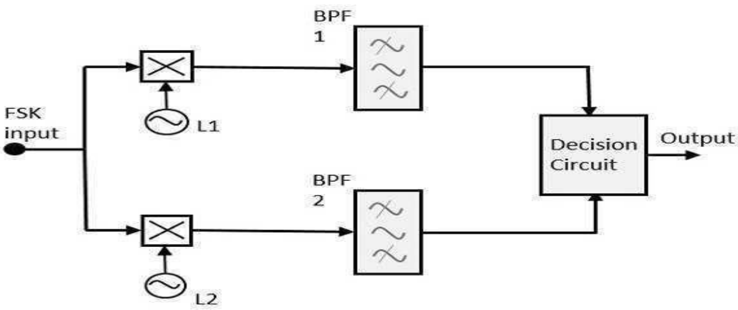


Figure 5.8. Synchronous FSK Detector

The FSK signal input is given to the two mixers with local oscillator circuits. These two are connected to two band pass filters. These combinations act as demodulators and the decision circuit chooses which output is more likely and selects it from any one of the detectors. The two signals have a minimum frequency separation.

For both of the demodulators, the bandwidth of each of them depends on their bit rate. This synchronous demodulator is a bit complex than asynchronous type demodulators.

5.3.5. Geometrical Representation of Orthogonal BFSK

In M-ary phase-shift keying and in quadrature-amplitude shift keying, any signal could be represented as $C_1u_1(t) + C_2u_2(t)$. There $u_1(t)$ and $u_2(t)$ are the orthonormal vectors in signal space, that is, $u1(t) = \sqrt{\frac{2}{T_s}}.\cos(\omega\ t)$and $u2(t) = \sqrt{\frac{2}{T_s}}.\sin(\omega\ \ t)$.

The functions u1 and u2 are orthonormal over the symbol interval T_S. And, if the symbol is a single bit, $T_S = T_b$. The coefficients C1 and C2 are constants. The normalized energies associated with $C_1u_1(t)$ and with $C_2u_2(t)$are respectively C_1^2 and C_2^2 and the total signal energy is $C_1^2 + C_2^2$.

In the present case of BFSK it is appropriate that the orthogonality should result from a special selection of the frequencies of the unit vectors. Accordingly, with m and n integers, let us establish unit vectors

$$u_1(t) = \sqrt{\frac{2}{T_b}}\cos 2\pi m f_b t \qquad \text{...1}$$

and

$$u_2(t) = \sqrt{\frac{2}{T_b}}\cos 2\pi n f_b t \qquad \text{...2}$$

Where $f_b=1/T_b$. The vectors U1 and U2 are the m^{th} and n^{th} harmonics of the (fundamental) frequency f_b. As we are aware, from the principles of Fourier analysis, different harmonics (m ± n) are orthogonal over the interval of the fundamental period $T_b = 1/f_b$.

If now the frequencies f_H and f_L in a BFSK system are selected to be (assuming m > n)

$$f_H = mf_b \qquad \text{...3}$$

and

$$f_L = nf_b \qquad \text{...4}$$

Then corresponding signal vectors are

$$S_H(t) = \sqrt{E_b}u_1(t) \qquad \text{...5}$$

and

$$S_L(t) = \sqrt{E_b}u_2(t) \qquad \text{...6}$$

The signal space representation of these signals is shown in Fig. 3.7.4. The signals, like the unit vectors are orthogonal. The distance between signal end points is therefore

$$d = \sqrt{2E_b}$$

Note that this distance is considerably smaller than the distance separating end points of BPSK signals, which are antipodal.

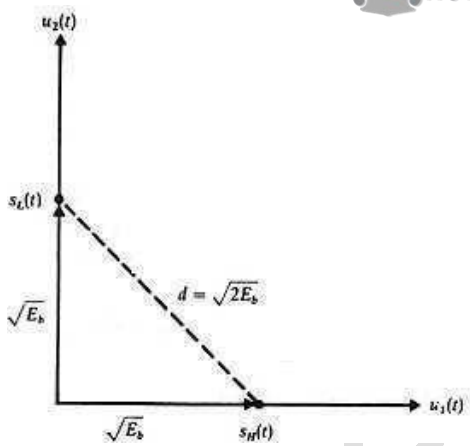


Figure 5.9. Signal Space representation of BFSK

5.4. Phase Shift Keying (PSK)

Phase Shift Keying (PSK) is the digital modulation technique in which the phase of the carrier signal is changed by varying the sine and cosine inputs at a particular time. PSK technique is widely used for wireless LANs, bio-metric, contactless operations, along with RFID and Bluetooth communications.
PSK is of two types, depending upon the phases the signal gets shifted. They are −

1. **Binary Phase Shift Keying (BPSK):** This is also called as 2-phase PSK or Phase Reversal Keying. In this technique, the sine wave carrier takes two phase reversals such as 0° and 180°.
2. **Quadrature Phase Shift Keying (QPSK):** This is the phase shift keying technique, in which the sine wave carrier takes four phase reversals such as 0°, 90°, 180°, and 270°.

5.4.1. BPSK Modulator:

The block diagram of Binary Phase Shift Keying consists of the balance modulator which has the carrier sine wave as one input and the binary sequence as the other input. Following is the diagrammatic representation.

BPSK Modulator

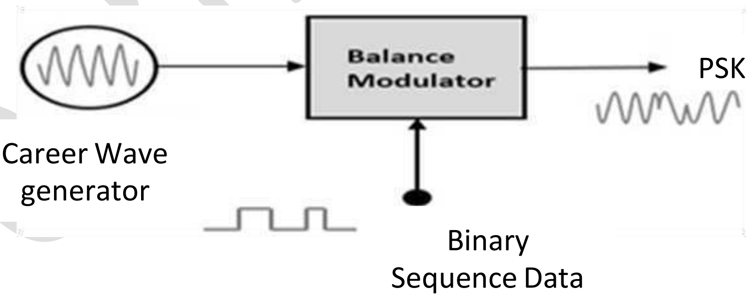


Figure 5.10 BPSK Modulator

The modulation of BPSK is done using a balance modulator, which multiplies the two signals applied at the input. For a zero binary input, the phase will be 0° and for a high input, the phase reversal is of 180°.
Following is the diagrammatic representation of BPSK Modulated output wave along with its given input.

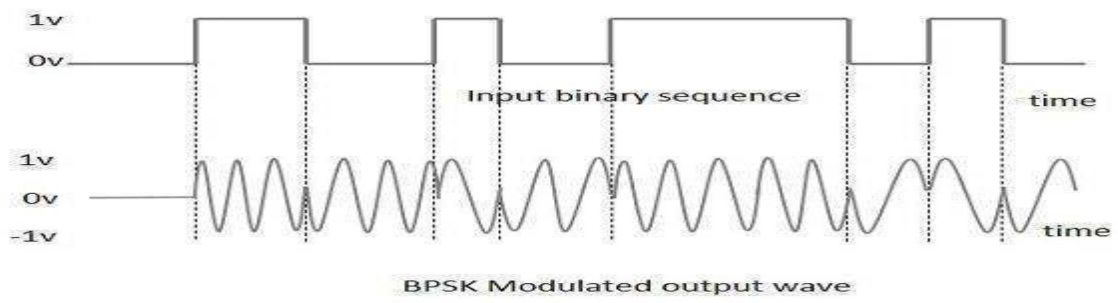


Figure 5.11. BPSK Modulated Waveform

The output sine wave of the modulator will be the direct input carrier or the inverted (180° phase shifted) input carrier, which is a function of the data signal.

5.4.2. BPSK Demodulator

The block diagram of BPSK demodulator consists of a mixer with local oscillator circuit, a band pass filter, a two-input detector circuit. The diagram is as follows.

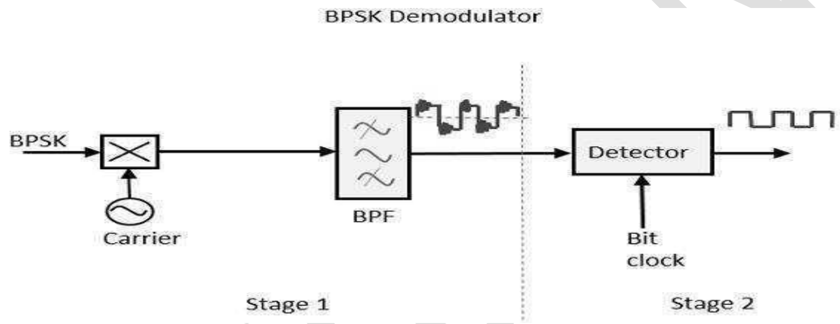


Figure 5.12 BPSK Demodulator

By recovering the band-limited message signal, with the help of the mixer circuit and the band pass filter, the first stage of demodulation gets completed. The base band signal which is band limited is obtained and this signal is used to regenerate the binary message bit stream.
In the next stage of demodulation, the bit clock rate is needed at the detector circuit to produce the original binary message signal. If the bit rate is a sub-multiple of the carrier frequency, then the bit clock regeneration is simplified. To make the circuit easily understandable, a decision-making circuit may also be inserted at the 2nd stage of detection.

5.5. Quadrature Phase Shift Keying (QPSK):

The Quadrature Phase Shift Keying (QPSK) is a variation of BPSK, and it is also a Double Side Band Suppressed Carrier (DSBSC) modulation scheme, which sends two bits of digital information at a time, called as digits.

Instead of the conversion of digital bits into a series of digital stream, it converts them into bit pairs. This decreases the data bit rate to half, which allows space for the other users.

5.5.1. QPSK Modulator

The QPSK Modulator uses a bit-splitter, two multipliers with local oscillator, a 2-bit serial to parallel converter, and a summer circuit. Following is the block diagram for the same.

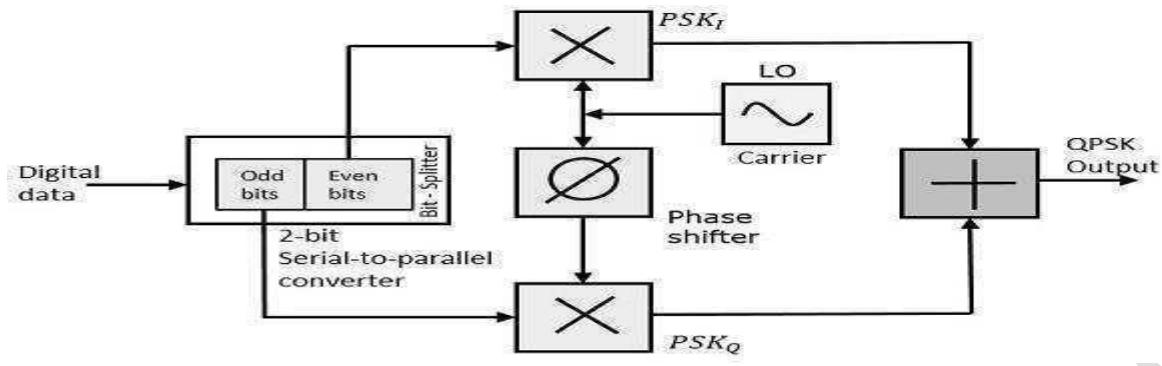


Figure 5.13 QPSK Modulator

At the modulator's input, the message signal's even bits (i.e., 2nd bit, 4th bit, 6th bit, etc.) and odd bits (i.e., 1st bit, 3rd bit, 5th bit, etc.) are separated by the bits splitter and are multiplied with the same carrier to generate odd BPSK (called as PSK_I) and even BPSK (called as PSK_Q). The PSK_Q signal is anyhow phase shifted by 90° before being modulated.

The QPSK waveform for two-bits input is as follows, which shows the modulated result for different instances of binary inputs.

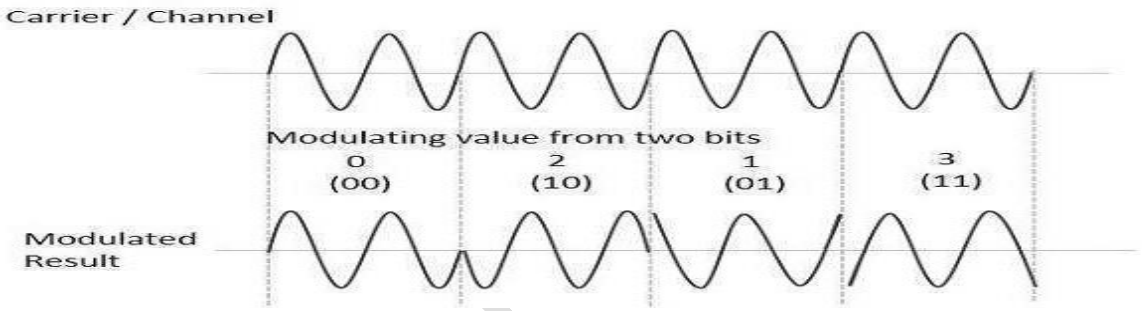


Figure 5.14 QPSK Waveforms

5.5.2. QPSK Demodulator

The QPSK Demodulator uses two product demodulator circuits with local oscillator, two band pass filters, two integrator circuits, and a 2-bit parallel to serial converter. Following is the diagram for the same.

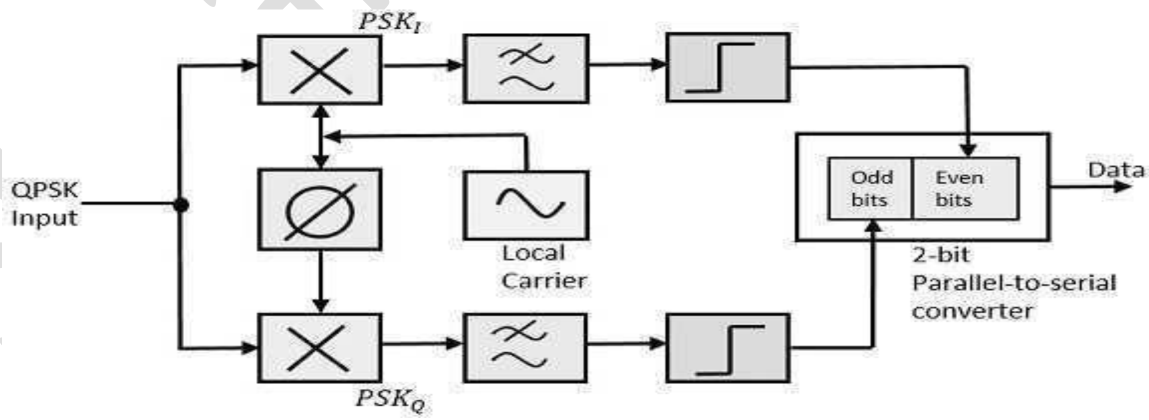


Figure 5.15 QPSK Demodulator

The two product detectors at the input of demodulator simultaneously demodulate the two BPSK signals. The pair of bits is recovered here from the original data. These signals after processing, are passed to the parallel to serial converter.

5.5.3. Phasor Diagram:

When $b_0 = 1$ the signal $s_o(t) = \sqrt{P_s}.\ \sin(\omega t)$, and $s_o(t) = -\sqrt{P_s}.\ \sin(\omega t)$ when $b_o = -1$. Correspondingly, for $b_e(t) = \pm 1$, $s_e(t) = \pm.\sqrt{P_s}.\ (t)\ \cos(\omega t)$ These four signals have been represented as phasors in Fig. 3.4.4.4. They are in mutual phase quadrature. Also drawn are the phasors representing the four possible output signals $v_m(t) = so(t) + Se(t)$. These four possible output signals have equal amplitude $\sqrt{2P_s}$ and are in phase quadrature; they have been identified by their corresponding values of b_o and b_e. At the end of each bit interval (i.e., after each time T_b) either b_o, or b_e can change, but both cannot change at the same time. Consequently, the QPSK system shown in Fig. 3.4.4.3 is called offset or staggered QPSK and abbreviated OQPSK. After each time T_b, the transmitted signal, if it changes, changes phase by 90° rather than by 180° as in BPSK.

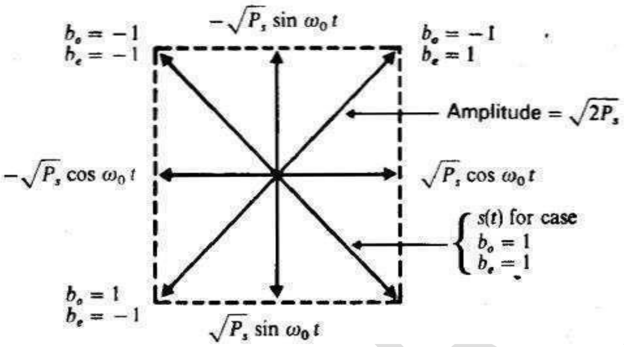


Figure 5.16. Phasor diagrams for the sinusoids

5.5.4. Non-offset QPSK

Suppose that in Fig. 3.4.4.3 we introduce an additional flip-flop before either the odd or even flip-flop. Let this added flip-flop be driven by the clock which runs at the rate f_b. Then one or the other bit streams, odd or even, will be delayed by one bit interval. As a result, we shall find that two bits which occur in time sequence (i.e., serially) in the input bit stream b(t) will appear at the same time (i.e., in parallel) at the outputs of the odd and even flip-flops. In this case $b_e(t)$ and $b_o(t)$ can change at the same time, after each time $2T_b$, and there can be a phase change of 180° in the output signal. There is no difference, in principle, between a staggered and non-staggered system.

In practice, there is often a significant difference between QPSK and OQPSK. At each transition time, T" for OQPSK and $2T_b$ for QPSK, one bit for OQPSK and perhaps two bits for QPSK change from 1V to -1V or -1V to 1V. Now the bits be(t) and bo(t) can, not change instantaneously and, in changing, must pass through zero and dwell in that neighborhood at least briefly. Hence there will be brief variations in the amplitude of the transmitted waveform. These variations will be more pronounced in QPSK than in OQPSK since in the first case both be(t) and bo(t) may be zero simultaneously so that the signal amplitude may actually be reduced to zero temporarily.

5.6. Differential Phase Shift Keying (DPSK)

In Differential Phase Shift Keying (DPSK) the phase of the modulated signal is shifted relative to the previous signal element. No reference signal is considered here. The signal phase follows the high or low state of the previous element. This DPSK technique doesn't need a reference oscillator.

The following figure represents the model waveform of DPSK.

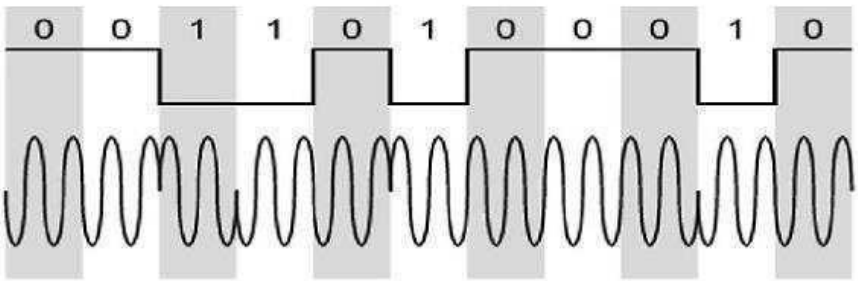


Figure 5.17 Differential Phase Shift Keying (DPSK)

It is seen from the above figure that, if the data bit is Low i.e., 0, then the phase of the signal is not reversed, but continued as it was. If the data is a High i.e., 1, then the phase of the signal is reversed, as with NRZI, invert on 1 (a form of differential encoding).
If we observe the above waveform, we can say that the High state represents an M in the modulating signal and the Low state represents a W in the modulating signal.

5.6.1. DPSK Modulator

DPSK is a technique of BPSK, in which there is no reference phase signal. Here, the transmitted signal itself can be used as a reference signal. Following is the diagram of DPSK Modulator.

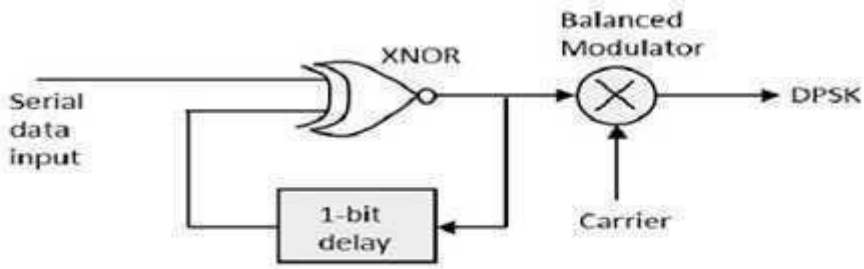


Figure 5.18 DPSK Modulator

DPSK encodes two distinct signals, i.e., the carrier and the modulating signal with 180° phase shift each. The serial data input is given to the XNOR gate and the output is again fed back to the other input through 1-bit delay. The output of the XNOR gate along with the carrier signal is given to the balance modulator, to produce the DPSK modulated signal.

5.6.2. DPSK Demodulator

In DPSK demodulator, the phase of the reversed bit is compared with the phase of the previous bit. Following is the block diagram of DPSK demodulator.

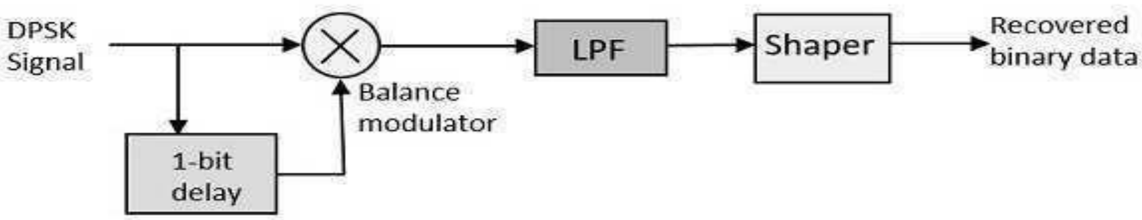


Figure 5.19 DPSK Demodulator

From the above figure, it is evident that the balance modulator is given the DPSK signal along with 1-bit delay input. That signal is made to confine to lower frequencies with the help of LPF. Then it is passed to a shaper circuit, which is a comparator or a Schmitt trigger circuit, to recover the original binary data as the output.

5.7. M-ary Equation

The word binary represents two bits. M represents a digit that corresponds to the number of conditions, levels, or combinations possible for a given number of binary variables.
This is the type of digital modulation technique used for data transmission in which instead of one bit, two or more bits are transmitted at a time. As a single signal is used for multiple bit transmission, the channel bandwidth is reduced.
If a digital signal is given under four conditions, such as voltage levels, frequencies, phases, and amplitude, then M = 4. The number of bits necessary to produce a given number of conditions is expressed mathematically as $N=\log 2M$. Where N is the number of bits necessary M is the number of conditions, levels, or combinations possible with N bits.
The above equation can be re-arranged as

$$2N=M$$

For example, with two bits, $2^2 = 4$ conditions are possible.

5.7.1. Types of M-ary Techniques

In general, Multi-level (M-ary) modulation techniques are used in digital communications as the digital inputs with more than two modulation levels are allowed on the transmitter's input. Hence, these techniques are bandwidth efficient. There are many M-ary modulation techniques. Some of these techniques, modulate one parameter of the carrier signal, such as amplitude, phase, and frequency.

5.7.2. M-ary ASK

This is called M-ary Amplitude Shift Keying (M-ASK) or M-ary Pulse Amplitude Modulation (PAM).
The amplitude of the carrier signal, takes on M different levels.

Representation of M-ary ASK

$S_m(t) = A_m \cos(2\pi f_c t) \qquad A_m \in (2m - 1 - M)\Delta, m = 1,2, \ldots\ldots M$ and $0 \le t \le T_s$

Some prominent features of M-ary ASK are –

- This method is also used in PAM.
- Its implementation is simple.
- M-ary ASK is susceptible to noise and distortion.

5.4. M-ARY FSK

An M-ary FSK communications system is shown in Fig. 5.19. It is an obvious extension of a binary FSK system. At the transmitter an N-bit symbol is presented each T_S, to an N-bit D/A converter. The converter output is applied to a frequency modulator, i.e., a piece of hardware which generates a carrier waveform whose frequency is determined by the modulating waveform. The transmitted signal, for the duration of the symbol interval, is of frequency f_0 or f_1 ...or f_{M-1} with $M = 2^N$. At the receiver, the incoming signal is applied to M paralleled band pass filters each followed by an envelope detector. The band pass filters have center frequencies $f_0, f_1, \ldots, f_{M-1}$. The envelope detectors apply their outputs to a device which determines which of the detector indications is the largest and transmits that envelope output to an N-bit AID converter.
The probability of error is minimized by selecting frequencies $f_0, f_1, \ldots, f_{M-1}$ so that the M signals are mutually orthogonal. One commonly employed arrangement simply provides that the carrier frequency be successive even harmonics of the symbol frequency $f_S=1/T_S$. Thus the lowest frequency, say f_0, is $f_0 = k f_S$, while $f_1 = (k + 1) f_S$, $f_2 = (k + 2) f_S$ etc. In this case, the spectral density patterns of the individual possible transmitted signals overlap in the manner shown in Fig. 5.20. We observe that to pass M-ary FSK the required spectral range is

$$B = 2Mf_S \qquad \ldots 1$$

Since $f_S = f_b/N$ and $M=2^N$, we have

$$B = 2^{N+1} f_b/N \qquad \ldots 2$$

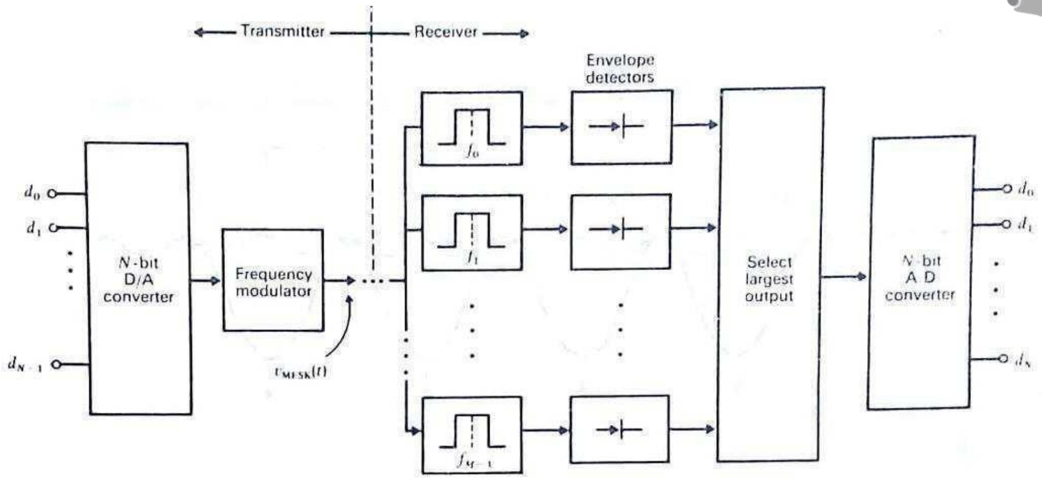


Figure 5.20 An M-ARY Communication System

Note that M-ary FSK requires a considerably increased bandwidth in comparison with M-ary PSK. However, as we shall see, the probability of error for M-ary FSK decreases as M increases, while for M-ary PSK, the probability of error increases with M.

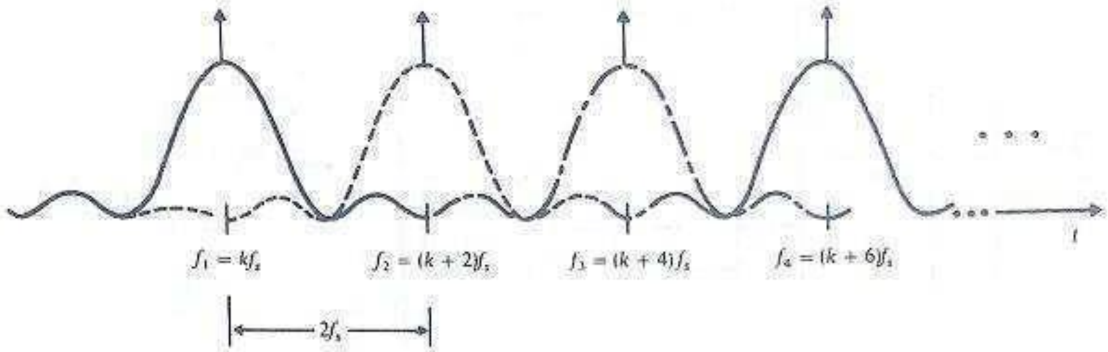


Figure 5.21. Power Spectral Density of an M-ARY FSK (Four Frequencies are shown)

Geometrical Representation of an M-ARY FSK

The case of M-ary orthogonal FSK signals is shown in figure 5.21. We simply conceive of a coordinate system with M mutually orthogonal coordinate axes. The square of the length of the signal vector is the normalized signal energy. When the frequencies are selected to generate orthogonal signals.

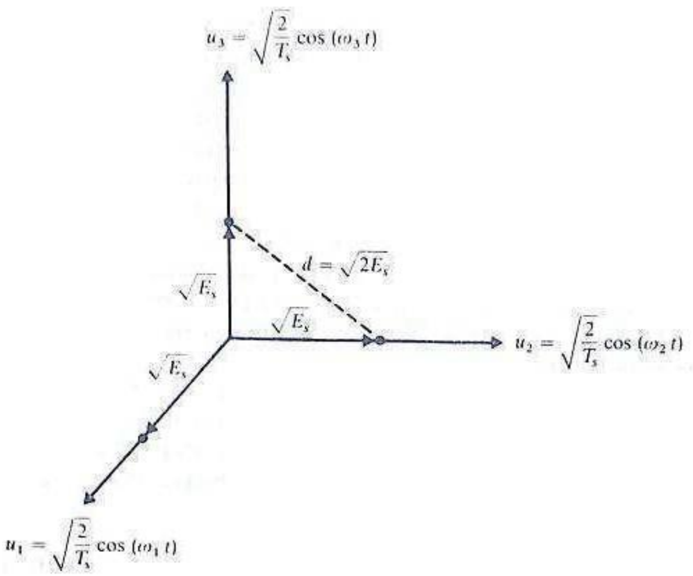


Figure 5.22 Geometrical representation of orthogonal M-ary FSK (M = 3)

Note that this value of d is greater than the values of d calculated for M-ary PSK with the exception of the cases M = 2 and M = 4. It is also greater than d in the case of 16-QASK.

$$d = \sqrt{2E_s} = \sqrt{2NE_b}$$

5.7.3. M-ary FSK

This is called as M-ary Frequency Shift Keying (M-ary FSK).

The frequency of the carrier signal, takes on M different levels.

Representation of M-ary FSK

$$S(t) = \sqrt{\frac{2E_s}{T_s}} \cos\left(\frac{\pi}{T_s}(n_c +)t\right) \qquad 0 \leq t \leq T_s \qquad = 1,2, \dots \dots M$$

Where $fc = nc/\,2Ts$

for some fixed integer n.

Some prominent features of M-ary FSK are −

- Not susceptible to noise as much as ASK.
- The transmitted M number of signals are equal in energy and duration.
- The signals are separated by $12Ts$
- Hz making the signals orthogonal to each other.
- Since M signals are orthogonal, there is no crowding in the signal space.
- The bandwidth efficiency of M-ary FSK decreases and the power efficiency increases with the increase in M.

5.7.3. M-ary PSK

This is called as M-ary Phase Shift Keying (M-ary PSK).

The phase of the carrier signal, takes on M different levels.

Representation of M-ary PSK

$$S\ (t) = \sqrt{\frac{2E}{T}} \cos(\omega_{0t} + \ t) \qquad 0 \leq t \leq T\ and \qquad = 1,2, \dots \dots M$$

$$(t) = \frac{2\pi}{M} \text{ where } \quad = 1,2, \dots \dots M$$

Some prominent features of M-ary PSK are −

- The envelope is constant with more phase possibilities.
- This method was used during the early days of space communication.
- Better performance than ASK and FSK.
- Minimal phase estimation error at the receiver.
- The bandwidth efficiency of M-ary PSK decreases and the power efficiency increases with the increase in M.

In BPSK we transmit each bit individually. Depending on whether b(t) is logic 0 or logic 1, we transmit one or another of a sinusoid for the bit time T_b, the sinusoids differing in phase by 2π/2 = 180^0. In QPSK we lump together two bits. Depending on which of the four two-bit words develops, we transmit one or another of four sinusoids of duration $2T_b$ the sinusoids differing in phase by amount 2π/4 = 90°. The scheme can be extended. Let us lump together N bits so that in this N-bit symbol, extending over the time NT_b, there are 2N = M possible symbols. Now let us represent the symbols by sinusoids of duration $NT_b = T_s$ which differ from one another by the phase 2 π / M. Hardware to accomplish such M-ary communication is available.

Thus in M-ary PSK the waveforms used to identify the symbols are

$$v_m(t) = \sqrt{2P_s} \cos(\omega t + \emptyset_m) \text{ (m=0, 1, ..., M-1)} \qquad \dots 1$$

Where phase angle is given by

$$\emptyset_m = (2m + 1)\frac{\pi}{M} \qquad \dots 2$$

The waveforms of Eq. are represented by the dots in Fig. 5.1 in a signal space in which the coordinate axes are the orthonormal waveforms $u1(t) = \sqrt{2/T_s}\cos(\omega t)$ and $u2(t) = \sqrt{2/T_s}\sin(\omega t)$. The distance of each dot from the origin is $\sqrt{E_s} = \sqrt{P_s T_s}$

From Eq. (1) we have

$$v_m(t) = (\sqrt{2P_s}\cos\emptyset_m)\cos(\omega t) - (\sqrt{2P_s}\sin\emptyset_m)\sin(\omega t) \quad ...3$$

Defining p_e and p_o by

$$p = \sqrt{2P_s}\cos\emptyset_m \quad ...4$$

$$p = \sqrt{2P_s}\sin\emptyset_m \quad ...5$$

Equation 3 becomes

$$v_m(t) = p\cos(\omega t) - p\sin(\omega t) \quad ...6$$

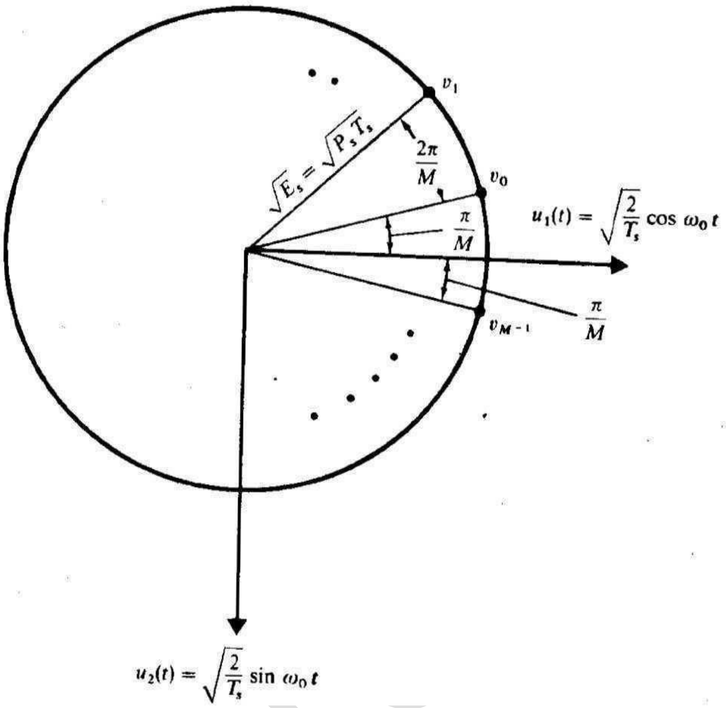


Figure 5.23 Graphical representation of M-ary PSK Signals

M-ary Transmitter and Receiver

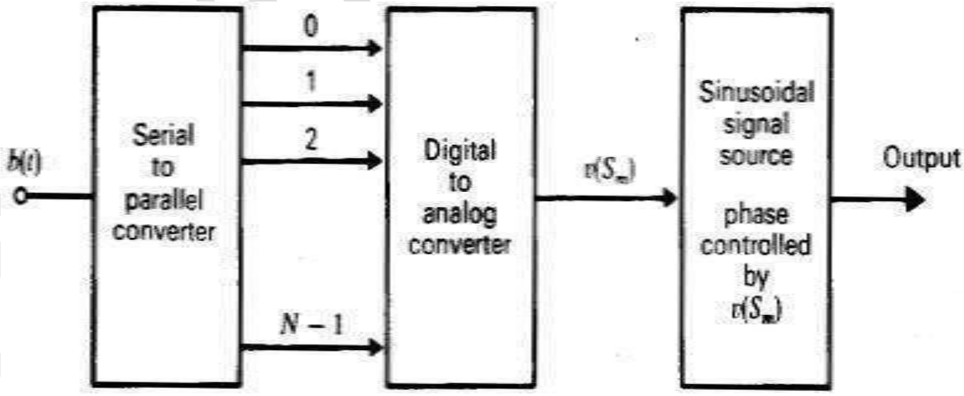


Figure.5.24. M Ary Transmitter

The transmitter, the bit stream b(t) is applied to a serial-to-parallel converter. This converter has facility for storing the N bits of a symbol. The N bits have been presented serially, that is, in time sequence, one after another. These N bits, having been assembled, are then presented all at once on N output lines of the converter, that is they are presented in parallel. The converter output remains unchanging for the duration NTb of a symbol during which time the converter is assembling a new group of N bits. Each symbol time the converter output is updated.

The converter output is applied to a D/A converter. This D/A converter generates an output voltage which assumes one of 2N = M different values in a one to-one correspondence to the M possible symbols applied to its input. That is, the D/A output is a voltage v(Sm) which depends on the symbol Sm (m = 0, 1,... ,M - 1).

Finally v(Sm) is applied as a control input to a special type of constant amplitude sinusoidal signal source whose phase 4>m is determined by v(Sm). Altogether, then, the output is a fixed amplitude, sinusoidal waveform, whose phase has a one-to-one correspondence to the assembled N-bit symbol. The phase can change once per symbol time.

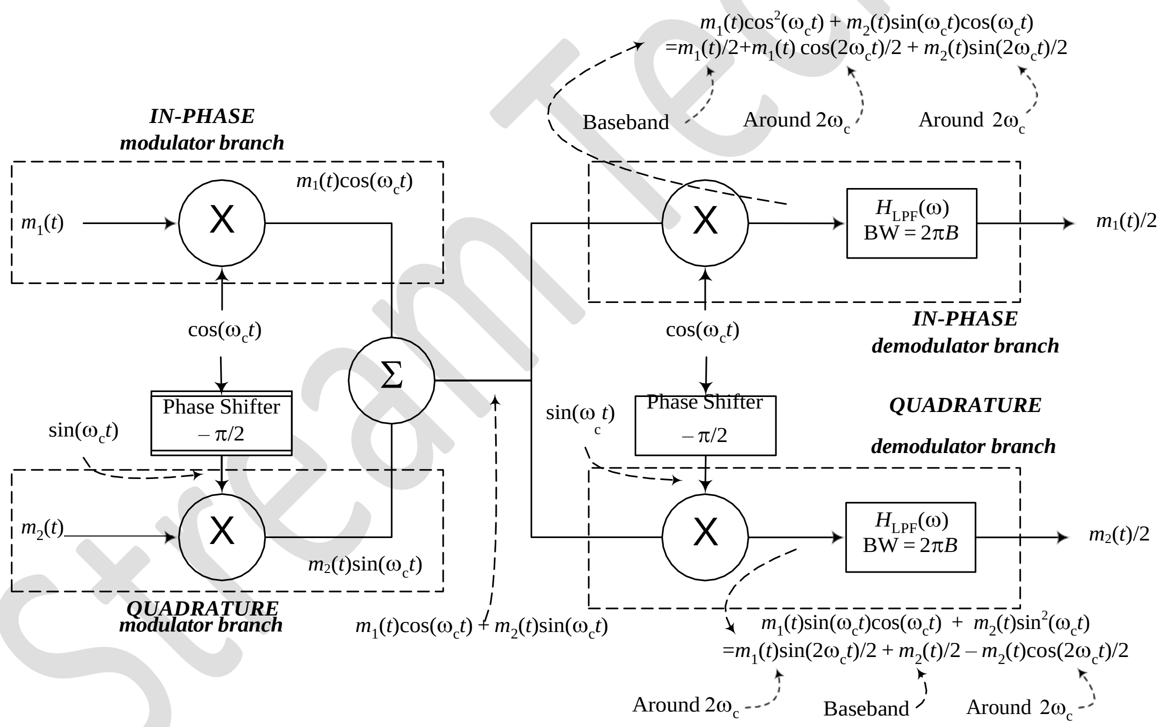
5.9. Quadrature Amplitude Modulation (QAM):

Quadrature Amplitude Modulation, QAM utilises both amplitude and phase components to provide a form of modulation that is able to provide high levels of spectrum usage efficiency. QAM, quadrature amplitude modulation has been used for some analogue transmissions. QAM is a signal in which two carriers shifted in phase by 90 degrees (i.e. sine and cosine) are modulated and combined. As a result of their 90° phase difference they are in quadrature and this gives rise to the name. Often one signal is called the In-phase or "I" signal, and the other is the quadrature or "Q" signal.

The resultant overall signal consisting of the combination of both I and Q carriers contains of both amplitude and phase variations. In view of the fact that both amplitude and phase variations are present it may also be considered as a mixture of amplitude and phase modulation.

The basic way in which a QAM signal can be generated is to generate two signals that are 90° out of phase with each other and then sum them. This will generate a signal that is the sum of both waves, which has certain amplitude resulting from the sum of both signals and a phase which again is dependent upon the sum of the signals.

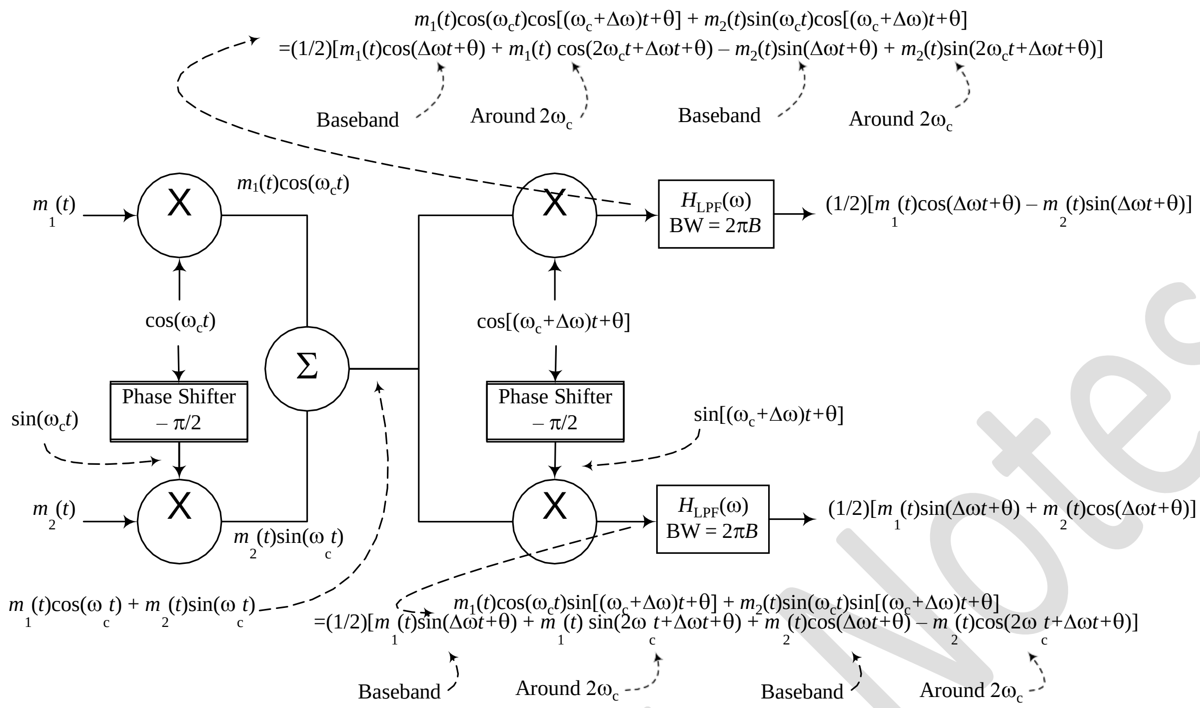
Consider the following block diagram of a Quadrature Amplitude Modulation (QAM) and Demodulation system:



QAM Modulator/Demodulator

Figure.5.24. QAM Modulator and demodulator

The modulator/demodulator system shown above clearly is able to modulate and demodulate two different signals without any interference. However, if the generation of the carrier at the demodulator had even small phase or frequency errors, the demodulated signals will interfere at the outputs. The following figure illustrate what happens when the carrier at the demodulator has a small frequency error $\Delta\omega$ (must be a small value much less than ω_c) and/or a small phase error θ.



QAM Modulator/Demodulator with Demodulator Carrier Phase and/or Frequency Error

If the carrier at the receiver has a small frequency error $\Delta\omega$ (but a phase error θ=0), we see that the two output signal become

$$r_1(t) = \frac{1}{2}\left[m_1(t)\cos(\Delta\omega t) - m_2(t)\sin(\Delta\omega t)\right]$$
$$r_2(t) = \frac{1}{2}\left[m_1(t)\sin(\Delta\omega t) + m_2(t)\cos(\Delta\omega t)\right]$$

Clearly, in this case, the output signals are not purely either of the two message signals but a combination. The ratio of message 1 to message 2 at the different outputs changes as a sinusoid with a frequency equal to the frequency error $\Delta\omega$.

If the carrier at the receiver has a phase error θ (but a frequency error $\Delta\omega$ = 0), we see that the two output signal become

$$r_1(t) = \frac{1}{2}\left[m_1(t)\cos(\theta) - m_2(t)\sin(\theta)\right]$$
$$r_2(t) = \frac{1}{2}\left[m_1(t)\sin(\theta) + m_2(t)\cos(\theta)\right]$$