ANALOG & DIGITAL COMMUNICATION

LECTURE NOTES

B.TECH (II- YEAR, IV SEM)

(2019-20)

Prepared By: Dr. Deepak Bhatia

Department of Electronics Engineering

Rajasthan Technical University

Rawat Bhata Road, Kota-324010

Syllabus:

Analog and Digital Communication

(EC)

B. Tech. IV Semester **Max. Marks: 150**

3L+0T

Handout/ Lecture Plan:

4ECU5	DCC	Analog and Digital Communication	MM:150	3L:0T:0P	3 credit

Review of signals and systems, Frequency domain representation of signals, Principles of Amplitude Modulation Systems -DSB, SSB and VSB modulations. Angle Modulation, Representation of FM and PM signals, Spectral characteristics of angle modulated signals. **Total: 8 Lecture**
Review of probability and random process. Gaussian and white noise characteristics, Noise in amplitude modulation systems, Noise in Frequency modulation systems. Pre -emphasis and De -emphasis, Threshold effect in angle modulation. **Total: 7 Lecture**
Pulse modulation. Sampling process. Pulse Amplitude and Pulse code modulation (PCM), Differential pulse code modulation. Delta modulation, Noise considerations in PCM, Time Division multiplexing, Digital Multiplexers. **Total: 8 Lecture**
Elements of Detection Theory, Optimum detection of signals in noise, Coherent communication with waveforms - Probability of Error evaluations. Baseband Pulse Transmission - Inter symbol Interference and Nyquist criterion. Pass band Digital Modulation schemes - Phase Shift Keying, Frequency Shift Keying, Quadrature Amplitude Modulation, Continuous Phase Modulation and Minimum Shift Keying.

Total: 9 Lecture
Digital Modulation tradeoffs. Optimum demodulation of digital signals over band -limited channels -Maximum likelihood sequence detection (Viterbi receiver). Equalization Techniques. Synchronization and Carrier Recovery for Digital modulation. **Total: 8 Lecture**

Text/ Reference Books:

1.	Haykin S., "Communications Systems", John Wiley and Sons, 2001.
2.	Taub H. and Schilling D.L., "Principles of Communication Systems", Tata McGraw Hill, 2001.
3.	Proakis J. G. and Salehi M., "Communication Systems Engineering",Pearson Education, 2002.
4.	Wozencraf t J. M. and Jacobs I. M., ``Principles of Communication Engineering", John Wiley, 1965.
5.	Barry J. R., Lee E. A. and Messerschmitt D. G., ``Digital Communication", Kluwer Academic Publishers, 2004.
6.	Proakis J.G., ``Digital Communications", 4th Edition, McGraw Hill, 2000.

Course Outcome:

Course Code	Course Name	Course Outcome	Details
4ECU5	Analog and Digital Communication	CO 1	Analyze and compare different analog modulation schemes for their efficiency and bandwidth
		CO 2	Analyze the behavior of a communication system in presence of noise
		CO 3	Investigate pulsed modulation system and analyze their system performance
		CO 4	Analyze different digital modulation schemes and can compute the bit error performance
		CO 5	Design a communication system comprised of both analog and digital modulation techniques

CO-PO Mapping:

Subject	Course Out comes	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
4EC4 -02 Analog & Digital Communication	CO 1	3	3		3		1				1		
	CO 2	3	2		3		1						
	CO 3	3	2		3		2						
	CO 4	3	3		3		2				1		
	CO 5	3	2	3	3		3			2	2		

3: Strongly 2: Moderate 1: Weak

Content delivery method:

1. Chalk and Duster
2. PPT

Lecture Plan:

Lecture No.	Content to be taught
Lecture 1	Introduction to the COURSE
Lecture 2	Review of signals and systems, Frequency domain representation of signals
Lecture 3	Principles of Amplitude Modulation Systems - DSB, SSB and VSB modulations
Lecture 4	Principles of Amplitude Modulation Systems - DSB, SSB and VSB modulations
Lecture 5	Principles of Amplitude Modulation Systems - DSB, SSB and VSB modulations
Lecture 6	Angle Modulation, Representation of FM and PM signals
Lecture 7	Angle Modulation, Representation of FM and PM signals
Lecture 8	Spectral characteristics of angle modulated signals.
Lecture 9	Review of probability and random process
Lecture 10	Review of probability and random process
Lecture 11	Noise in amplitude modulation systems
Lecture 12	Noise in amplitude modulation systems
Lecture 13	Noise in Frequency modulation systems
Lecture 14	Pre-emphasis and De emphasis
Lecture 15	Threshold effect in angle modulation

Lecture 16	Pulse modulation. Sampling
Lecture 17	Pulse Amplitude and Pulse code modulation (PCM)
Lecture 18	Pulse Amplitude and Pulse code modulation (PCM)
Lecture 19	Differential pulse code modulation
Lecture 20	Delta modulation
Lecture 21	Noise considerations in PCM
Lecture 22	Time Division multiplexing, Digital Multiplexers
Lecture 23	Elements of Detection Theory
Lecture 24	Optimum detection of signals in noise
Lecture 25	Coherent communication with waveforms- Probability of Error evaluations
Lecture 26	Coherent communication with waveforms- Probability of Error evaluations
Lecture 27	Baseband Pulse Transmission - Inter symbol Interference and Nyquist criterion
Lecture 28	Baseband Pulse Transmission - Inter symbol Interference and Nyquist criterion
Lecture 29	Pass band Digital Modulation schemes
Lecture 30	Phase Shift Keying
Lecture 31	Frequency Shift Keying
Lecture 32	Quadrature Amplitude Modulation
Lecture 33	Continuous Phase Modulation and Minimum Shift Keying.

Lecture 34	Digital Modulation tradeoffs
Lecture 35	Optimum demodulation of digital signals over band-limited channels
Lecture 36	Optimum demodulation of digital signals over band-limited channels
Lecture 37	Maximum likelihood sequence detection (Viterbi receiver)
Lecture 38	Equalization Techniques
Lecture 39	Synchronization and Carrier Recovery for Digital modulation
Lecture 40	Synchronization and Carrier Recovery for Digital modulation

Assignments:

Assignment 1	Q1. Design Modulator and Demodulator of SSB-SC Modulation based on its mathematical expression.
	Q2. Derive the figure of merit in a) FM Receiver b) PM Receiver
	Q3. A Carrier signal c(t) = 20cos (2 n10 6t) is modulated by a message signal having three frequencies 5 KHz, 10 KHz & 20 KHz. The corresponding modulation indexes are 0.4, 0.5 & 0.6. Sketch the spectrum. Calculate bandwidth, power and efficiency.
Assignment 2	Q1. Derive the expression for probability of error in ASK, FSK and PSK systems and compare them.
	Q2. With block diagrams explain about DPCM & DM. also compare them.
	Q3. A message signal m(t) = 4 cos (2n10 3t) is sampled at nyquist rate and transmitted through a channel using 3 -bit PCM system. Calculate all the parameters of the PCM. If the sampled values are 3.8, 2.1, 0.5, - 1.7, -3.2 & -4 then determine the quantizer output, encoder output and quantization error per each sample. Sketch the transfer characteristics of the quantizer.

UNIT-I

Introduction to Communication System

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

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

The basic block diagram of a communication system is as shown in Fig.1.

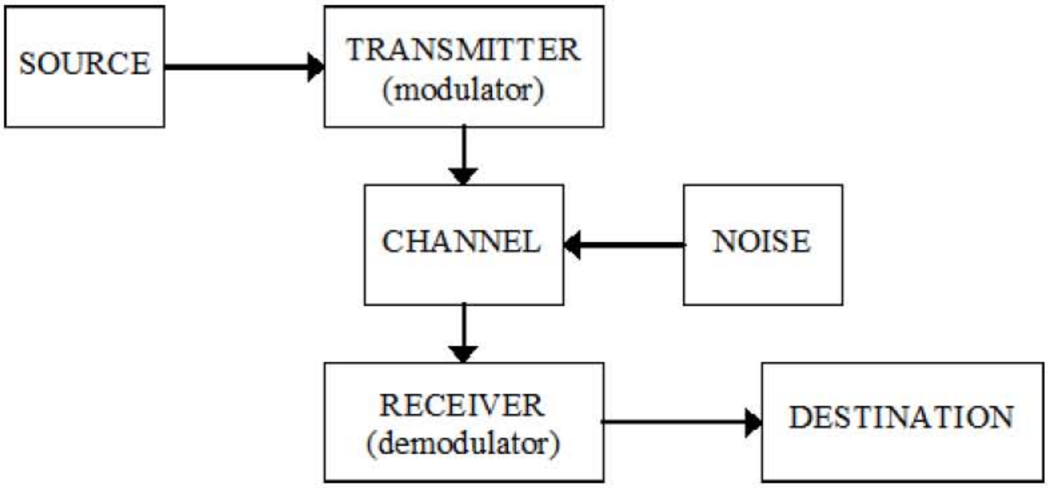


Fig.1 Block Diagram of a Communication System

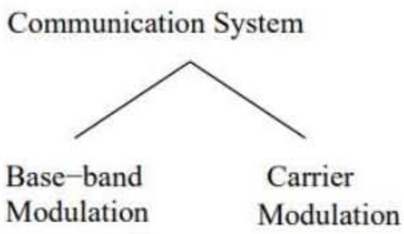
- **Transmitter:** Couples the message into the channel using high frequency signals.
- **Channel:** The medium used for transmission of signals
- **Modulation:** It is the process of shifting the frequency spectrum of a signal to a frequency range in which more efficient transmission can be achieved.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

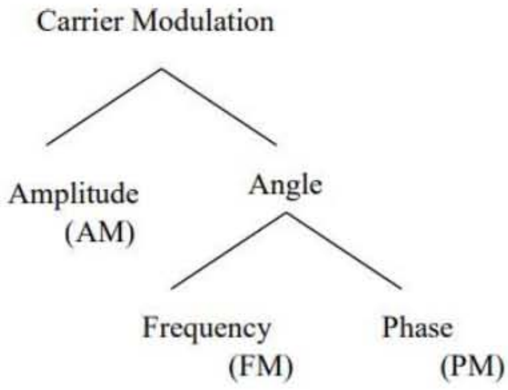
Modulation:

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

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

- Modulation helps in noise immunity, attenuation - depends on the physical medium. The below figure shows the different kinds of analog modulation schemes that are available





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

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

It involves two waveforms:

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

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

Need for Modulation:

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

Types of Modulation:

Three main types of modulations:

Analog Modulation

- **Amplitude modulation**

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

- **Angle modulation (Frequency modulation & Phase modulation)**

  Example: Narrow band frequency modulation (NBFM), Wideband frequency modulation (WBFM), Narrowband phase modulation (NBPM), Wideband phase modulation (NBPM)

Pulse Modulation

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

Digital Modulation

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

Frequency Division Multiplexing

Multiplexing is the name given to techniques, which allow more than one message to be transferred via the same communication channel. The channel in this context could be a transmission line, *e.g.* a twisted pair or co-axial cable, a radio system or a fiber optic system *etc*.

FDM is derived from AM techniques in which the signals occupy the same physical 'line' but in different frequency bands. Each signal occupies its own specific band of frequencies all the time, *i.e.* the messages share the channel **bandwidth**.

FDM – messages occupy **narrow** bandwidth – all the time.

Multiplexing requires that the signals be kept apart so that they do not interfere with each other, and thus they can be separated at the receiving end. This is accomplished by separating the signal either in frequency or time. The technique of separating the signals in frequency is referred to as frequency division multiplexing (FDM), whereas the technique of separating the signals in time is called time- division multiplexing (TDM).

Fig.2 shows the block diagram of FDM System. As shown in the figure, input message signals, assumed to be of the low-pass type are passed through input low-pass filters. This filtering action removes high-frequency components that do not contribute significantly to signal representation but may disturb other message signals that share the common channel.

The filtered message signals are then modulated with necessary carrier frequencies with the help of modulators. The most commonly used method of modulation in FDM is single sideband modulation, which requires a bandwidth that is approximately equal to that of original message signal. The band pass filters following the modulators are used to restrict the band of each modulated wave to its prescribed range. The outputs of BPF are combined in parallel which form the input to the common channel.

At the receiving end, BPF connected to the common channel separate the message signals on the frequency occupancy basis. Finally, the original message signals are recovered by individual demodulators.

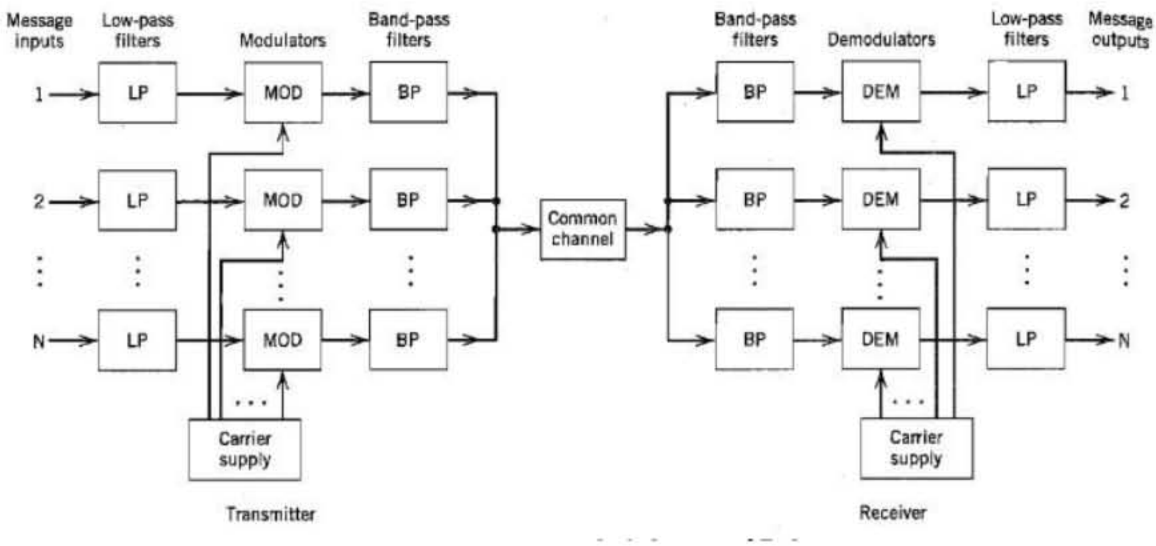


Fig.2. Frequency Division Multiplexing

Amplitude Modulation (AM)

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

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

Application of AM –

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

Various forms of Amplitude Modulation

- Conventional Amplitude Modulation (Alternatively known as Full AM or Double Sideband Large carrier modulation (DSBLC) /Double Sideband Full Carrier (DSBFC)
- Double Sideband Suppressed carrier (DSBSC) modulation
- Single Sideband (SSB) modulation
- Vestigial Sideband (VSB) modulation

Time Domain and Frequency Domain Description

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

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

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

Fourier Transform on both sides of the above equation

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

k_a is a constant called amplitude sensitivity.

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

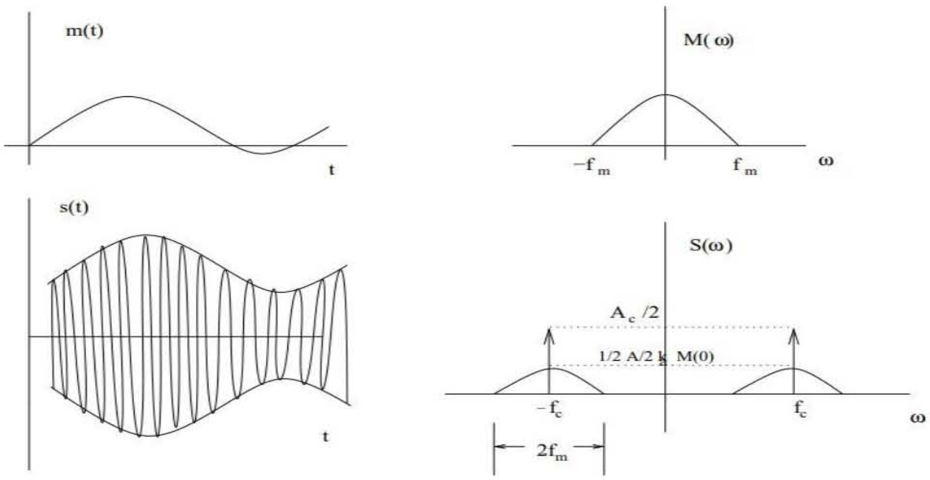


Fig.3. Amplitude modulation in time and frequency domain

Single Tone Modulation:

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

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

where Am is peak amplitude of the sinusoidal modulating wave

fm is the frequency of the sinusoidal modulating wave

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

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

m is the modulation index

Let A_{max} and A_{min} denote the maximum and minimum values of the envelope of the modulated wave. Then from the above equation we get

$$\frac{Amax}{Amin} = \frac{Ac(1+m)}{Ac(1-m)}$$

$$m = \frac{Amax - Amin}{Amax + Amin}$$

Expanding the equation (2), we get

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

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

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

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

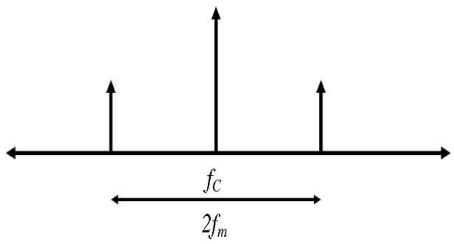


Fig.4. Frequency Domain characteristics of single tone AM

Power relations in AM waves:

Consider the expression for single tone/sinusoidal AM wave

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

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

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

$P_t = \frac{V_{car}^2}{R} + \frac{V_{LSB}^2}{R} + \frac{V_{USB}^2}{R}$ ------------ (2)

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

$$Pc = \frac{V_{car}^2}{R} = \frac{(^{A_c}/_{\sqrt{2}})^2}{R} = \frac{A_c^2}{2R}$$

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

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

Therefore total average power is given by

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

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

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

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

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

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

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

This ratio is called the efficiency of AM system Generation of AM waves:

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

Square Law Modulator

When the output of a device is not directly proportional to input throughout the operation, the device is said to be non-linear. The Input-Output relation of a non-linear device can be expressed as

$$V_O = a_0 + a_1 V_{in} + a_2 V_{in}^2 + a_3 V_{in}^3 + a_4 V_{in}^4 + \cdots \ldots \ldots \ldots \ldots \ldots$$

When the input is very small, the higher power terms can be neglected. Hence the output is approximately given by

$$V_O = a_0 + a_1 V_{in} + a_2 V_{in}^2$$

When the output is considered up to square of the input, the device is called a square law device and the square law modulator is as shown in the figure 5

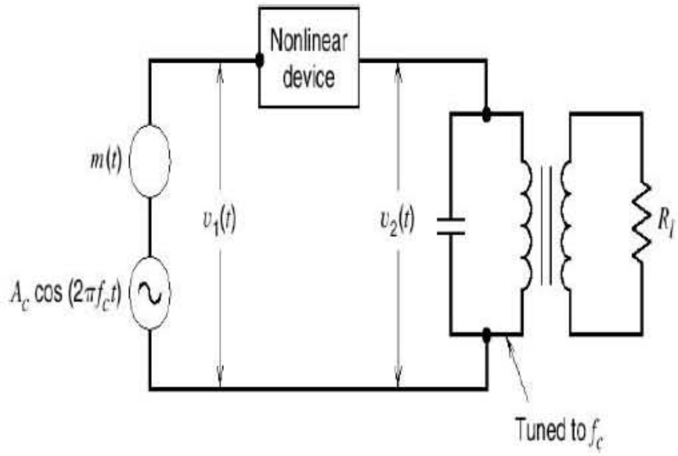


Fig.5. Square Law Modulator

Consider a non-linear device to which a carrier $c(t)=A_c\cos(2\pi f_c t)$ and an information signal m(t) are fed simultaneously as shown in figure 4. The total input to the device at any instant is

$$V_{in} = c(t) + m(t)$$

$$V_{in} = A_c \cos(2\pi f_c t) + m(t)$$

As the level of the input is very small, the output can be considered upto square of the input *i.e.,*

$$V_O = a_0 + a_1 V_{in} + a_2 V_{in}^2$$

$$V_O = a_0 + a_1[A_c \cos(2\pi f_c t) + m(t)] + a_2[A_c \cos(2\pi f_c t) + m(t)]^2$$

$$V_O = a_0 + a_1 A_c \cos(2\pi f_c t) + a_1 m(t)] + \frac{a_2 A_c^2}{2}(1 + cos4\pi f_c t) + a_2[m(t)]^2 + 2a_2 m(t) A_c \cos(2\pi f_c t)$$

$$V_O = a_0 + a_1 A_c \cos(2\pi f_c t) + a_1 m(t)] + \frac{a_2 A_c^2}{2}(cos4\pi f_c t) + \frac{a_2 A_c^2}{2} + a_2[m(t)]^2 + 2a_2 m(t) A_c \cos(2\pi f_c t)$$

Taking Fourier Transform on both sides , we get

$$V_o(f) = \left(a_0 + \frac{a_2 A_c^2}{2}\right)\delta(f) + \frac{a_1 A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] + a_1 M(f) + \frac{a_2 A_c^2}{4}[\delta(f - 2f_c) + \delta(f + 2f_c)] + a_2 M(f) + a_2 A_c\big[M\big(f - f_c + M(f + f_c)\big)\big]$$

Therefore the square law device output 0 V consists of the dc component at f = 0. The information signal ranging from 0 to W Hz and its second harmonics are signal at f_c and $2f_c$.

Frequency band centered at f_c with a deviation of ± W Hz.The required AM signal with a carrier frequency f_c can be separated using a BPF at the output of the square law device. The filter should have a lower cut off frequency ranging between 2W and $(f_c - W)$ and upper cut-off frequency between $(f_c + W)$ and $2f_c$.

Therefore the filter output is

$$s(t) = a_1 A_c \cos(2\pi f_c t) + 2a_2 A_c m(t) \cos(2\pi f_c t)$$

$$s(t) = a_1 A_c [1 + 2\frac{a_2}{a_1} m(t)] \cos(2\pi f_c t)$$

If $m(t) = A_m \cos(2\pi f_m t)$, we get

$$s(t) = a_1 A_c [1 + 2\frac{a_2}{a_1} A_m \cos(2\pi f_m t)] \cos(2\pi f_c t)$$

Comparing this equation with the standard representation of AM Signal,

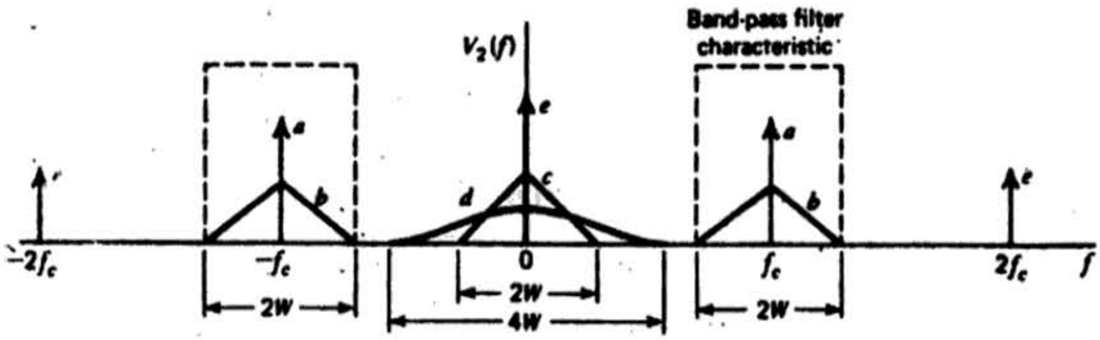
$$s(t) = A_c [1 + k_a m(t)] \cos(2\pi f_c t)$$

Therefore modulation index of the output signal is given by

$$m = 2\frac{a_2}{a_1} A_m$$

The output signal is free from distortion and attenuation only when $(f_c - W) > 2W$ or $f_c > 3W$

Spectrum is as shown below



Switching Modulator

Consider a semiconductor diode used as an ideal switch to which the carrier signal

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

and information signal $m(t)$ are applied simultaneously as shown in the Fig.6

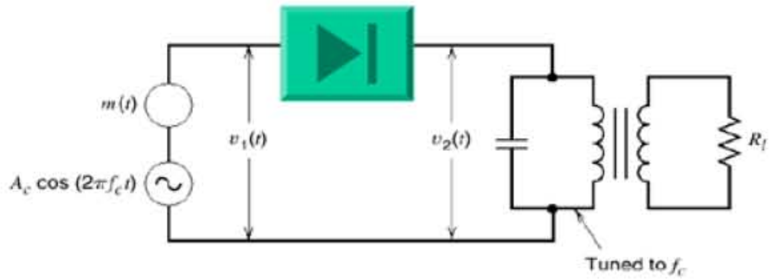


Fig.6. Switching Modulator

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

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

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

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

When c(t) is positive, $v_2 = v_1$ since the diode is forward biased. Similarly, when c(t) is negative, v2=0 since diode is reverse biased. Based upon above operation, switching response of the diode is periodic rectangular wave with an amplitude unity and is given by

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

$$p(t) = \frac{1}{2} + \frac{2}{\pi} \cos(2\pi f_c t) - \frac{2}{3\pi} \cos(6\pi f_c t) + \cdots \ldots \ldots$$

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

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

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

Applying the Fourier Transform we get

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

The diode output v_2 consists of a dc component at $f = 0$, information signal ranging from 0 to w Hz and infinite number of frequency bands centered at $f_c, 2f_c, 3f_c, 4f_c$ ………

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

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

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

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

Therefore modulation index,

$$m = \frac{4}{\pi}\frac{A_m}{A_c}$$

The output AM signal is free from distortions and attenuations only when

$$f_c - w > w \; or \; f_c > 2w$$

Detection of AM waves

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

Square Law Detector

Consider a non-linear device to which the AM signal s(t) is applied. When the level of s(t) is very small, output can be considered upto square of the input.

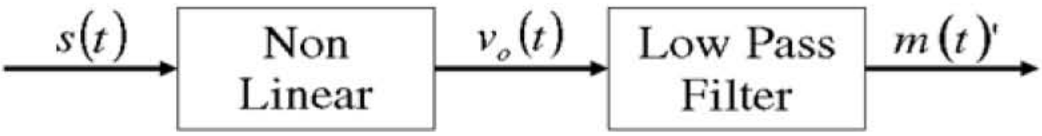


Fig .7. Demodulation of AM using square law device

Therefore

$$V_O = a_0 + a_1 V_{in} + a_2 V_{in}^2$$

If $m(t)$ is the information signal (0-wHz) and $c(t) = A_c \cos(2\pi f_c t)$ is the carrier, input AM signal to the non-linear device is given by

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

$$V_O = a_0 + a_1 s(t) + a_2 [s(t)]^2$$

$$V_O = a_0 + a_1 A_c \cos(2\pi f_c t) + a_1 A_c k_a m(t) \cos(2\pi f_c t) + a_2 [A_c \cos(2\pi f_c t) + A_c k_a m(t) \cos(2\pi f_c t)]^2$$

Applying Fourier Transform on both sides, we get

$$V_o(f) = \left[a_o + \frac{a_2 A_c^2}{2}\right]\delta(f) + \frac{a_1 A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] + \frac{a_1 A_c K_a}{2}[M(f - f_c) + M(f + f_c)] + \frac{a_2 A_c^2 K_a^2}{4}[M(f - 2f_c) + M(f + 2f_c)] + \frac{a_2 A_c^2 K_a^2}{2}\left[\underset{\pm 2W}{M(f)}\right] + \frac{a_2 A_c^2 K_a^2}{2}[M(f - 2f_c) + M(f + 2f_c)] + \frac{a_2 A_c^2}{4}[\delta(f - 2f_c) + \delta(f + 2f_c)] + a_2 A_c^2 K_a [M(f)]$$

The device output consists of a dc component at f=0, information signal ranging from 0-W Hz and its second harmonics and frequency bands centered at fc and 2fc. The required information can be separated using low pass filter with cut off frequency ranging between W and fc-w. The filter output is given by

$$m'(t) = \left(a_o + \frac{a_2 A_c^2}{2}\right) + a_2 A_c^2 K_a m(t) + \frac{a_2 A_c^2 K_a^2 m^2(t)}{2}$$

DC component + message signal + second harmonic

The dc component (first term) can be eliminated using a coupling capacitor or a transformer. The effect of second harmonics of information signal can be reduced by maintaining its level very low. When $m(t)$ is very low, the filter output is given by

$$m'(t) = a_2 A_c^2 k_a m(t)$$

When the information level is very low, the noise effect increases at the receiver, hence the system clarity is very low using square law demodulator.

Envelope Detector

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

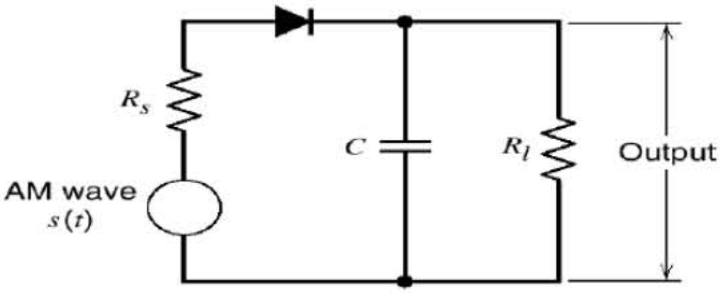


Fig.8. Envelope Detector

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

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

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

$$(r_f + R_s)C \ll \frac{1}{f_c}$$

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

That is the discharge time constant shall satisfy the condition,

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

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

Advantages and Disadvantages of AM: Advantages of AM:

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

Disadvantages:

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

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

DSBSC (Double Side Band Suppressed Carrier) modulation: In DSBC modulation, the modulated wave consists of only the upper and lower side bands. Transmitted power is saved through the suppression of the carrier wave, but the channel bandwidth requirement is the same as before.

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

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

DSB-SC MODULATION

DSB-SC Time domain and Frequency domain Description:

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

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

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

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

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

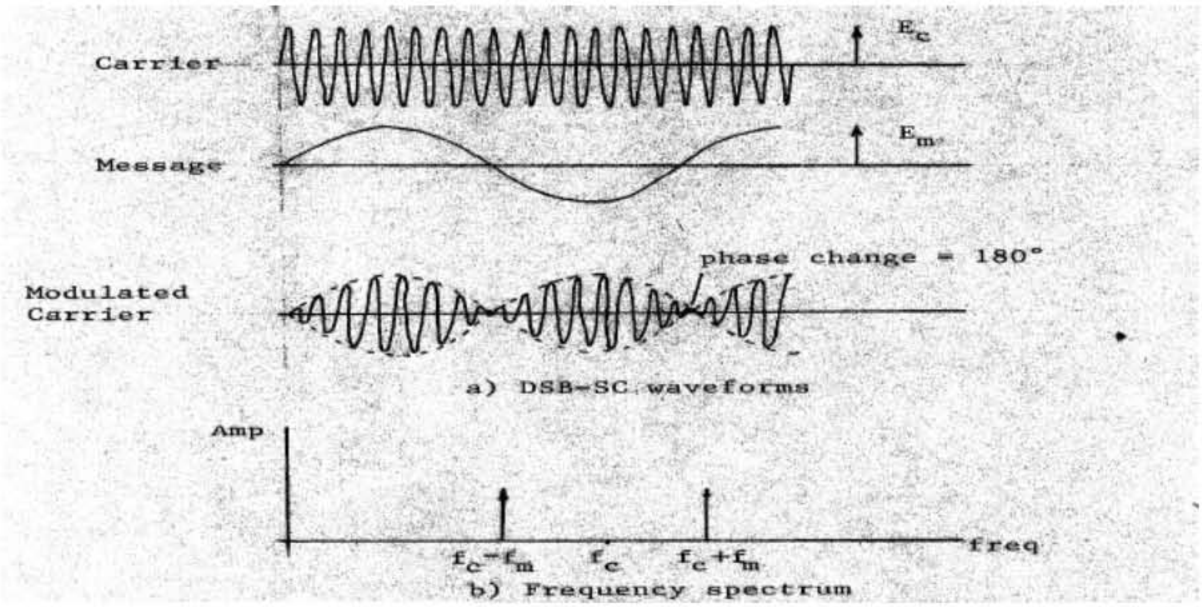


Fig.1. (a) DSB-SC waveform (b) DSB-SC Frequency Spectrum

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

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

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

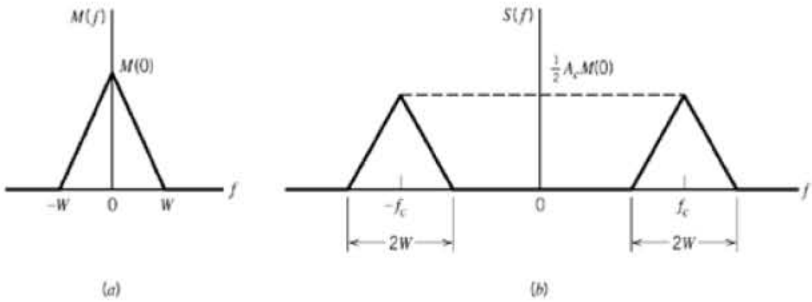


Fig.2. Message and the corresponding DSBSC spectrum

Generation of DSBSC Waves:

Balanced Modulator (Product Modulator)

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

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

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

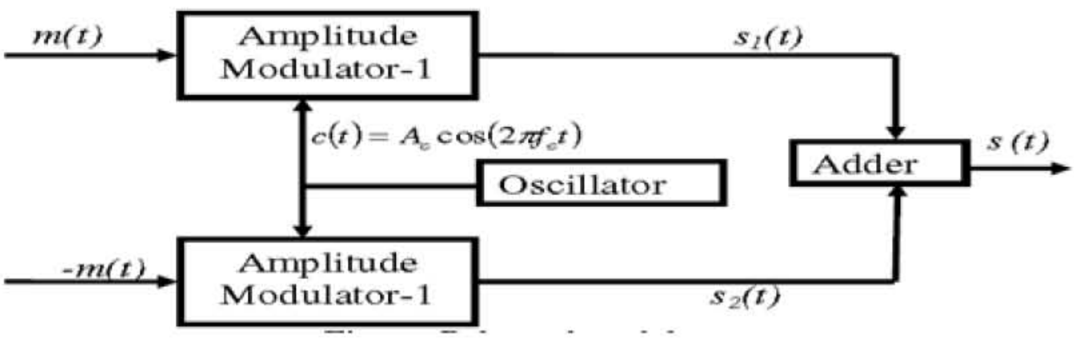


Fig.3. Balanced Modulator

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

$$s(t) = s_1(t) - s_2(t)$$

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

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

Ring Modulator

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

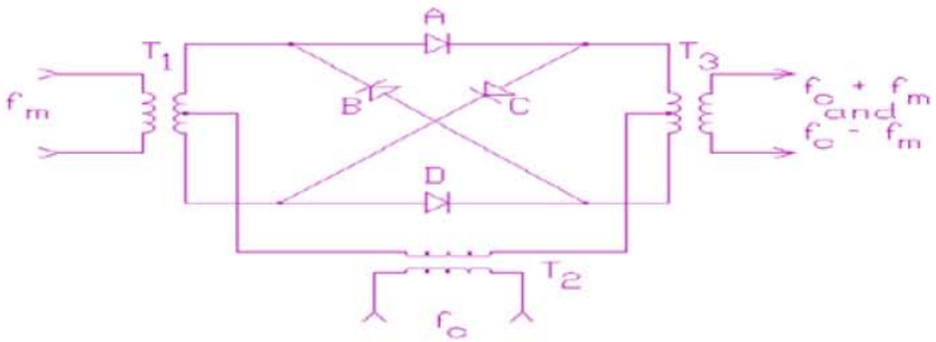


Fig.4. Ring Modulator

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

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

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

Therefore the ring modulator output is given by

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

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

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

Detection of DSB-SC waves:

Coherent Detection:

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

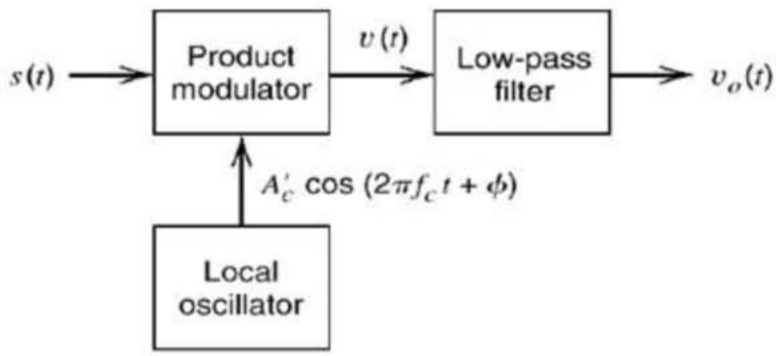


Fig.5 Coherent Detector

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

Let $A_c'\cos(2\pi f_c t + \phi)$ be the local oscillator signal, and$s(t) = A_c\cos(2\pi f_c t)\,m(t)$ be the DSBSC wave. Then the product modulator output $v(t)$ is given by

$$v(t) = A_c A_c' \cos(2\pi f_c t)\cos(2\pi f_c t + \phi)m(t)$$

$$v(t) = \frac{A_c A_c'}{4}\cos(4\pi f_c t + \emptyset)\,m(t) + \frac{A_c A_c'}{2}\cos(\emptyset)\,m(t)$$

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

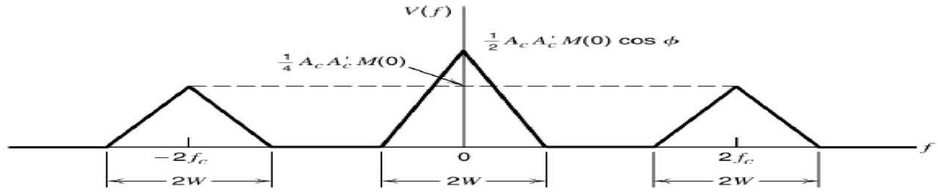


Fig.6.Spectrum of output of the product modulator

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

$$v_o(t) = \frac{A_c A_c'}{2}\cos(\emptyset)\, m(t)$$

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

Costas Receiver (Costas Loop):

Costas receiver is a synchronous receiver system, suitable for demodulating DSBSC waves. It consists of two coherent detectors supplied with the same input signal, that is the incoming DSBSC wave $s(t) = A_c\cos(2\pi f_c t)\, m(t)$ but with individual local oscillator signals that are in phase quadrature with respect to each other as shown in Fig.7

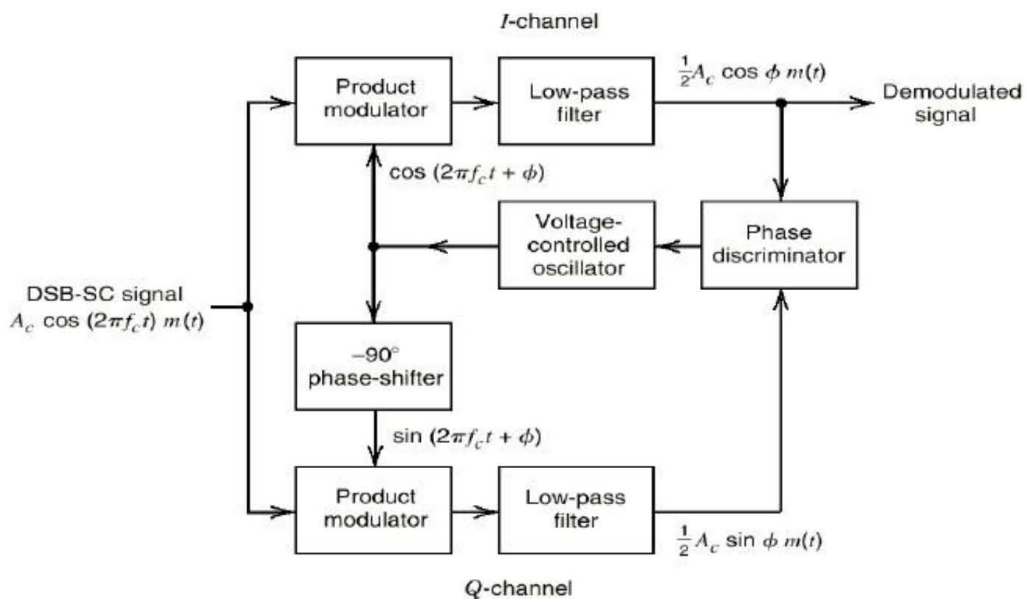


Fig.7. Costas Receiver

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

These two detector are coupled together to form a negative feedback system designed in such a way as to maintain the local oscillator synchronous with the carrier wave. Suppose

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

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

Radio Transmitters

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

Low-Level AM Transmitter:

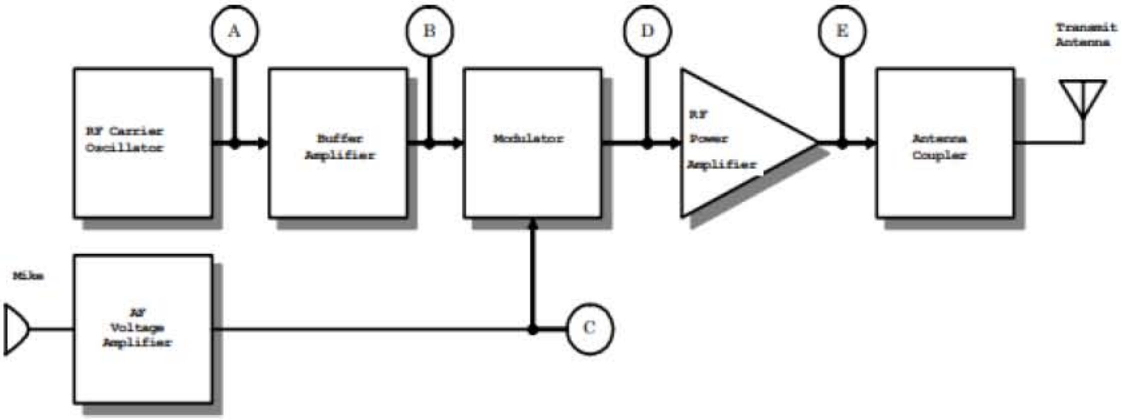


Fig.8. Low-Level AM Transmitter Block Diagram

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

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

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

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

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

AF Voltage Amplifier: In order for the modulator to function, it needs an information signal. A microphone is one way of developing the intelligence signal, however, it only produces a few millivolts of signal. This simply isn't enough to operate the modulator, so a voltage amplifier is used to boost the microphone's signal.

The signal level at the output of the AF voltage amplifier is usually at least 1 volt RMS; it is highly dependent upon the transmitter's design. Notice that the AF amplifier in the transmitter is only providing a voltage gain, and not necessarily a current gain for the microphone's signal. The power levels are quite small at the output of this amplifier; a few mW at best.

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

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

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

High-Level AM Transmitter:

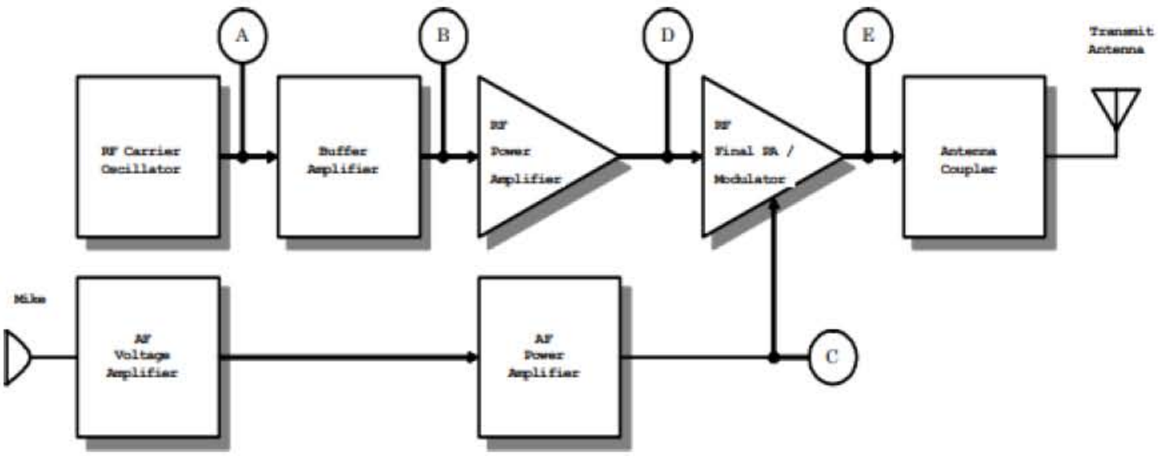


Fig.9. Low-Level AM Transmitter Block Diagram

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

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

Comparison:

Low Level Transmitters

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

High Level Transmitters

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

UNIT-II

Introduction to SSB-SC

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

Figure.1 : Spectrum of the DSBSC wave

Figure .2 : Spectrum of the SSBSC wave

Frequency Domain Description

Consider a message signal $m(t)$ with a spectrum $M(f)$ band limited to the interval $-w < f < w$ as shown in Fig.3, the DSBSC wave obtained by multiplexing $m(t)$ by the carrier wave $c(t) = A_c \cos(2\pi f_c t)$ and is also shown in figure 4. The upper side band is represented in duplicate by the frequencies above f_c and those below $-f_c$, and when only upper side band is transmitted; the resulting SSB modulated wave has the spectrum shown in figure 6. .

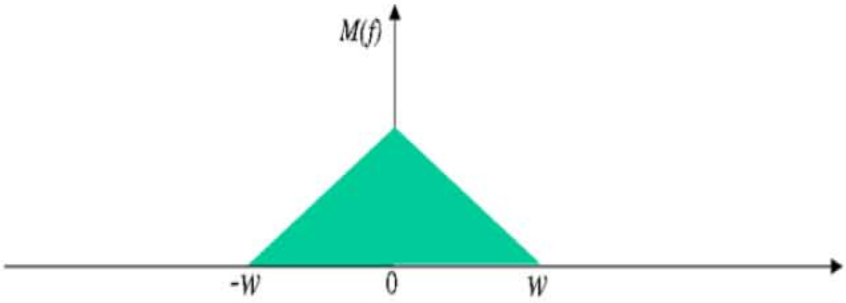


Figure 3. : Spectrum of message wave

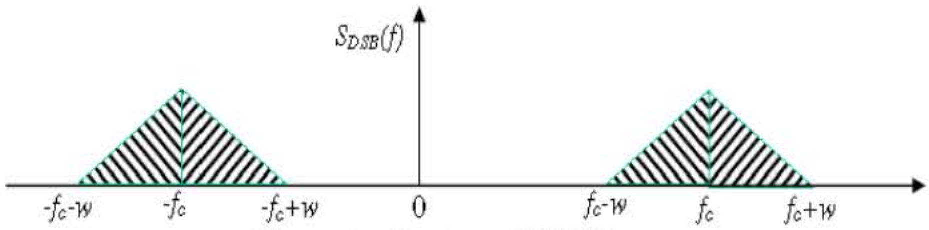


Figure .4 : Spectrum of DSBSC wave

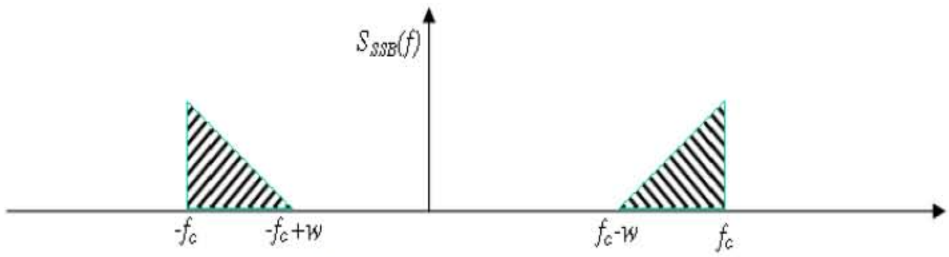


Figure.5 : Spectrum of SSBSC-LSB wave

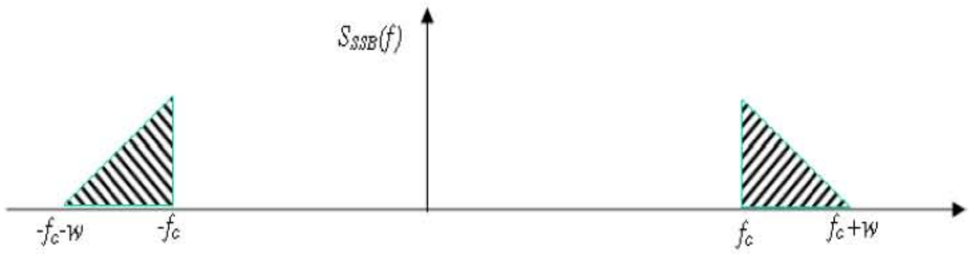


Figure .6 : Spectrum of SSBSC-USB wave

Similarly, the lower side band is represented in duplicate by the frequencies below fc and those above -fc and when only the lower side band is transmitted, the spectrum of the corresponding SSB modulated wave shown in figure 5.Thus the essential function of the SSB modulation is to translate the spectrum of the modulating wave, either with or without inversion, to a new location in the frequency domain.

The advantage of SSB modulation is reduced bandwidth and the elimination of high power carrier wave. The main disadvantage is the cost and complexity of its implementation.

Generation of SSB wave:

Frequency discrimination method

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

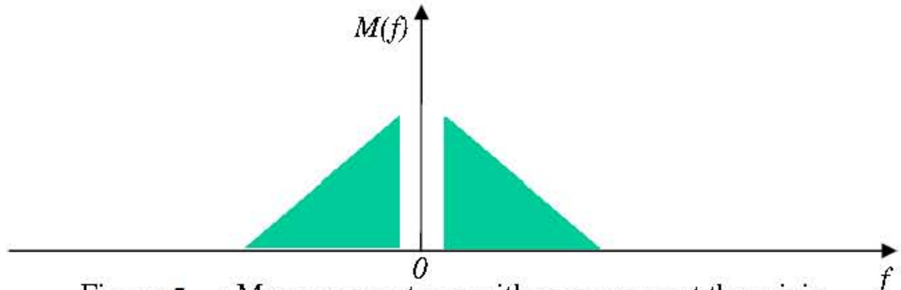


Figure .7 : Message spectrum with energy gap at the origin

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

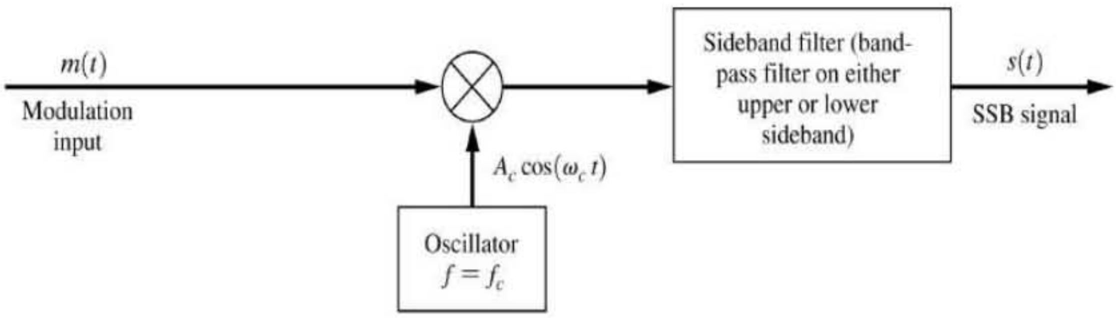


Figure .8 : Frequency discriminator to generate SSBSC wave

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

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

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

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

1. The pass band of the filter occupies the same frequency range as the spectrum of the desired SSB modulated wave.
2. The width of the guard band of the filter, separating the pass band from the stop band, where the unwanted sideband of the filter input lies, is twice the lowest frequency component of the message signal.

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

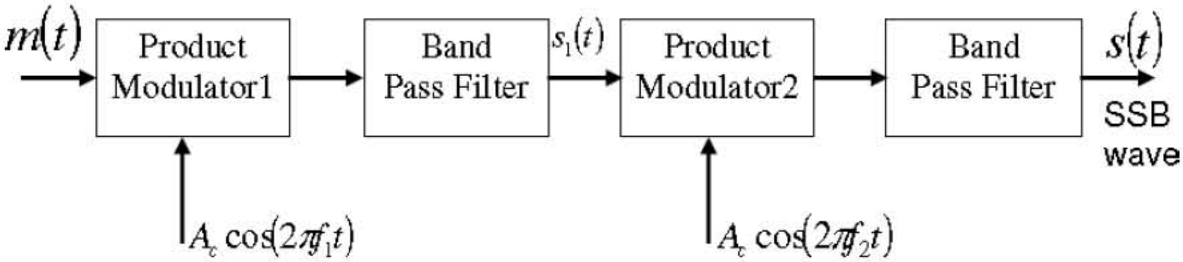


Figure .9 : Two stage frequency discriminator

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

Hilbert Transform & its Properties:

The Fourier transform is useful for evaluating the frequency content of an energy signal, or in a limiting case that of a power signal. It provides mathematical basis for analyzing and designing the frequency selective filters for the separation of signals on the basis of their frequency content. Another method of separating the signals is based on phase selectivity, which uses phase shifts between the appropriate signals (components) to achieve the desired separation. In case of a sinusoidal signal, the simplest phase shift of 180° is obtained by "Ideal transformer" (polarity reversal). When the phase angles of all the components of a given signal are shifted by 90°, the resulting function of time is called the "Hilbert transform" of the signal.

Consider an LTI system with transfer function defined by equation 1

$$H(f) = \begin{cases} -j, f > 0 \\ 0, f = 0 \\ j, f < 0 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(1)$$

and the Signum function given by

$$sgn(f) = \begin{cases} 1, f > 0 \\ 0, f = 0 \\ -1, f < 0 \end{cases}$$

The function $H(f)$ can be expressed using Signum function as given by equ (2)

$$H(f) = -jsgn(f) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(2)$$

We know that

$$e^{-j\pi/2} = -j,\ e^{j\pi/2} = j \ and \ e^{\pm j\theta} = \cos(\theta) \pm jsin(\theta)$$

Therefore,

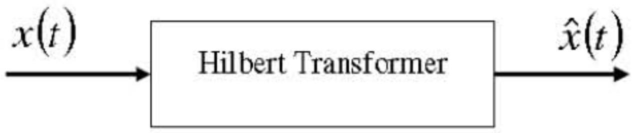
$$H(f) = \begin{cases} e^{-j\pi/2}, f > 0 \\ e^{j\pi/2}, f < 0 \end{cases}$$

Thus the magnitude $|H(f)| = 1$, for all f, and angle

$$\angle H(f) = \begin{cases} -\pi/2, f > 0 \\ \pi/2, f < 0 \end{cases}$$

The device which possesses such a property is called Hilbert transformer. Whenever a signal is applied to the Hilbert transformer, the amplitudes of all frequency components of the input signal remain unaffected. It produces a phase shift of -90° for all positive frequencies, while a phase shifts of 90° for all negative frequencies of the signal.

If x(t) is an input signal, then its Hilbert transformer is denoted by x^(t) and shown in the following diagram.



To find impulse response $h(t)$ of Hilbert transformer with transfer function $H(f)$. Consider the relation between Signum function and the unit step function.

$$sgn(t) = 2u(t) - 1 = x(t)$$

Differentiating both sides with respect to t,

$$\frac{d}{dt}\{x(t)\} = 2\delta(t)$$

Applying Fourier transform on both sides,

$$sgn(t) \leftrightarrow \frac{2}{j\omega} \rightarrow sgn(t) \leftrightarrow \frac{1}{j\pi f}$$

Applying duality property of Fourier transform,

$$-Sgn(f) \leftrightarrow \frac{1}{j\pi t}$$

We have,

$$H(f) = -jsgn(f)$$

$$H(f) \leftrightarrow \frac{1}{\pi t}$$

Therefore the impulse response $h(t)$ of an Hilbert transformer is given by the equation (3),

$$h(t) = \frac{1}{\pi t} \dots\dots\dots\dots (3)$$

Now consider any input x(t) to the Hilbert transformer, which is an LTI system. Let the impulse response of the Hilbert transformer is obtained by convolving the input x(t) and impulse response h(t) of the system.

$$\hat{x}(t) = x(t) * h(t)$$

$$\hat{x}(t) = x(t) * \frac{1}{\pi t}$$

$$\hat{x}(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{x(\tau)}{(t-\tau)} d\tau \dots\dots\dots\dots\dots\dots.(4)$$

This equation gives the Hilbert transform of $x(t)$.

The inverse Hilbert transform $x(t)$ is given by

$$x(t) = \frac{-1}{\pi}\int_{-\infty}^{+\infty} \frac{\hat{x}(\tau)}{(t-\tau)} d\tau \dots\dots\dots\dots\dots\dots.(5)$$

We have $\hat{x}(t) = x(t) * h(t)$

The Fourier transform $\hat{X}(f)$ of $\hat{x}(t)$ is given by

$$\hat{X}(f) = X(f)H(f)$$

$$\hat{X}(f) = -jsgn(f)X(f) \dots\dots\dots\dots\dots.(6)$$

Properties:

1. A signal $x(t)$ and its Hilbert transform $\hat{x}(t)$ have the same amplitude spectrum. The magnitude of $-jsgn(f)$ is equal to 1 for all frequencies *f*. Therefore *x(t)* and $\hat{x}(t)$ have the same amplitude spectrum. That is $|\hat{X}(f)| = |X(f)|\ for\ all\ f$
2. If $\hat{x}(t)$ is the Hilbert transform of *x(t),* then the Hilbert transform of $\hat{x}(t) is - x(t)$. To obtain its Hilbert transform of *x(t), x(t)* is passed through LTI system with a transfer function equal to $-jsgn(f)$. A double Hilbert transformation is equivalent to passing $x(t)$ through a cascade of two such devices. The overall transfer function of such a cascade is equal to

$$[-jsgn(f)]^2 = -1\ for\ all\ f$$

The resulting output is $-x(t)$. That is the Hilbert transform of $\hat{x}(t)$is equal to $-x(t)$.

Time Domain Description:

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

$$s(t) = s_I(t)\cos(2\pi f_c t) - s_Q(t) sin(2\pi f_c t) \dots\dots\dots\dots\dots\dots\dots\dots(1)$$

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

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

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

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

Where $-w \leq f \leq w$ defines the frequency band occupied by the message signal $m(t)$

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

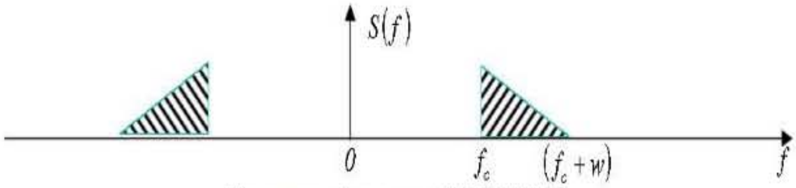


Figure 10 : Spectrum of SSBSC-USB

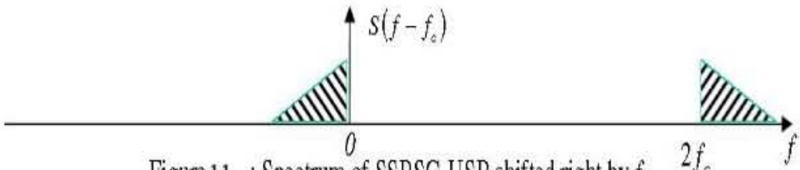


Figure 11 : Spectrum of SSBSC-USB shifted right by f_c

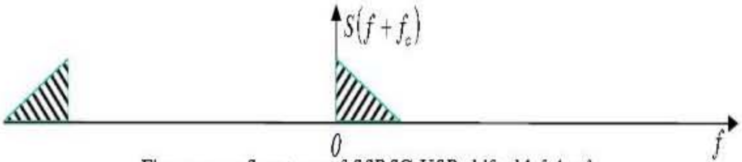


Figure 12 : Spectrum of SSBSC-USB shifted left by f_c

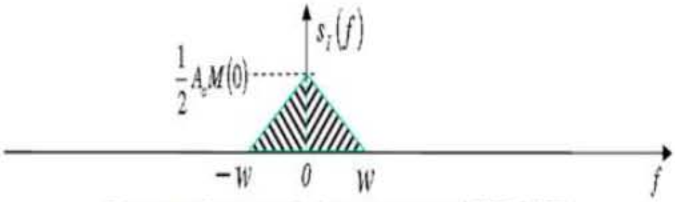


Figure 13 : Spectrum of in-phase component of SSBSC-USB

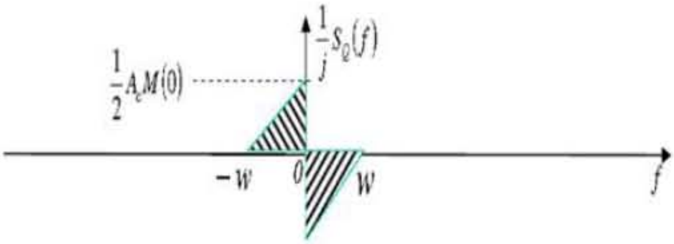


Figure 14 : Spectrum of quadrature component of SSBSC-USB

From the fig.13, it is found that

$$S_I(f) = \frac{1}{2}A_cM(f)$$

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

$$S_I(t) = \frac{1}{2}A_cm(t) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Now on the basis of fig.14, it is found that

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

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

Where $sgn(f)$ is the Signum Function

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

$$-jsgn(f)M(f) = \widehat{M}(f) \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Where $\widehat{M}(f)$ is the Fourier transform of the Hilbert transform of $m(t)$.

Hence, substituting equ.(6) in equ.(5), we get

$$S_Q(f) = \frac{1}{2}A_c\widehat{M}(f) \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

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

$$S_Q(t) = \frac{1}{2}A_c\widehat{m}(t) \ldots\ldots\ldots\ldots (8)$$

Therefore substituting equ.(4) and (8) in equ.(1), we find that the canonical representation of an SSB wave s(t) obtained by transmitting only the USB is given by equ (9)

$$S_U(t) = \frac{1}{2}A_c m(t)\cos(2\pi f_c t) - \frac{1}{2}A_c \hat{m}(t)\, sin(2\pi f_c t) \dots\dots\dots\dots\dots\dots\dots\dots\dots (9)$$

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

$$S_L(t) = \frac{1}{2}A_c m(t)\cos\left(2\pi f_c t\right) + \frac{1}{2}A_c \hat{m}(t)\, sin\left(2\pi f_c t\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots (10)$$

Phase discrimination method for generating SSB wave:

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

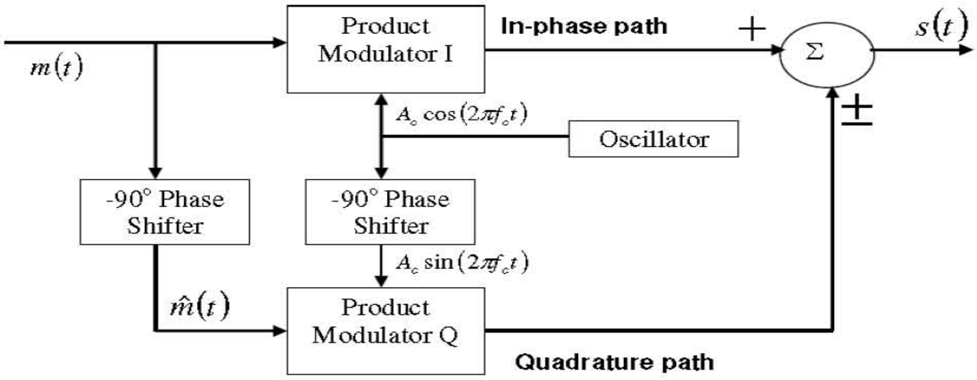


Figure 15 : Block diagram of phase discriminator

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

The Hilbert transform 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.

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

Demodulation of SSB Waves:

Demodulation of SSBSC wave using coherent detection is as shown in fig. 16. The SSB wave $s(t)$ together with a locally generated carrier $c(t) = A'_c \cos(2\pi f_c t + \emptyset)$ is applied to a product modulator and then low-pass filtering of the modulator output yields the message signal.

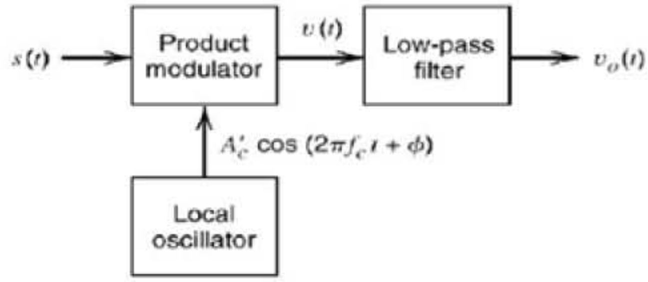


Fig.16. Coherent Detector of SSBSC

The product modulator output $v_0(t)$ is given by

$$v(t) = A'_c \cos(2\pi f_c t + \emptyset)\, s(t), Put\ \emptyset = 0$$

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

$$v(t) = \frac{1}{4} A_c m(t) + \frac{1}{4} A_c [m(t)\cos(4\pi f_c t) \pm \hat{m}(t)\sin(4\pi f_c t)] \ldots\ldots\ldots\ldots\ldots (1)$$

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

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.

Frequency Domain Description

The following Fig illustrates the spectrum of VSB modulated wave s (t) with respect to the message m (t) (band limited)

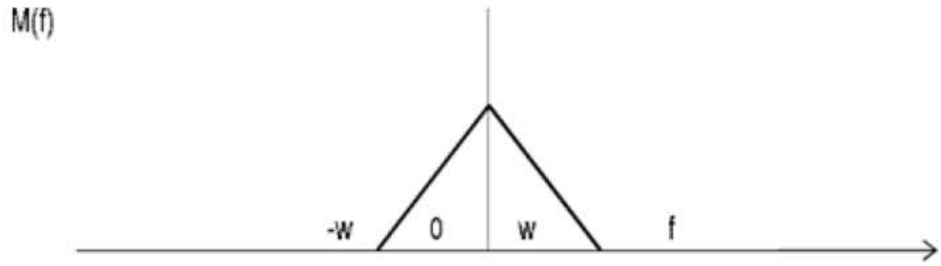


Fig.1. Spectrum of message signal

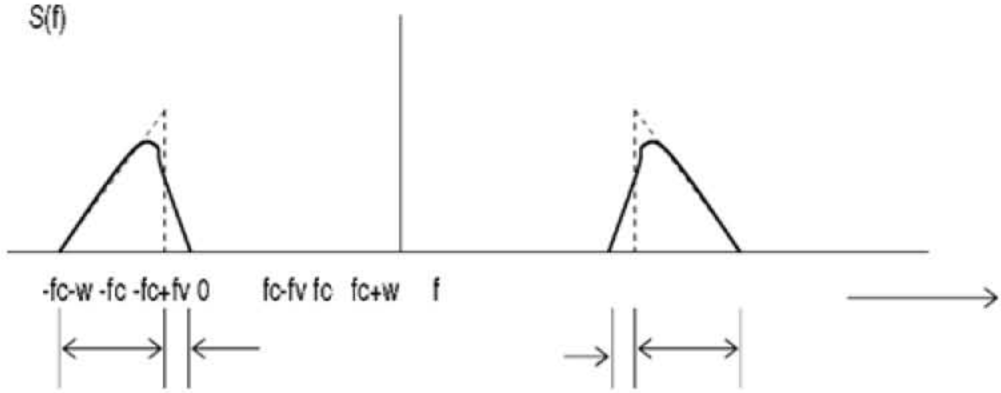


Fig.2. Spectrum of VSB wave containing the vestige of the LSB

Assume that the Lower side band is modified into the vestigial side band. The vestige of the lower sideband compensates for the amount removed from the upper sideband. The bandwidth required to send VSB wave is

$$B = w + f_v$$

where f_v is the width of the vestigial side band.

Similarly, if USB is modified into the VSB then,

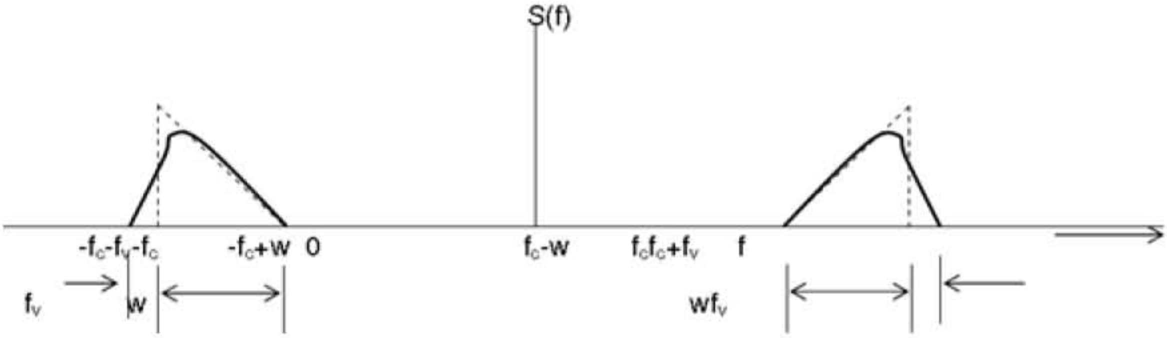


Fig.3. Spectrum of VSB wave containing the vestige of the USB

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 fv 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

VSB modulated wave is obtained by passing DSBSC through a sideband shaping filter as shown in below fig.

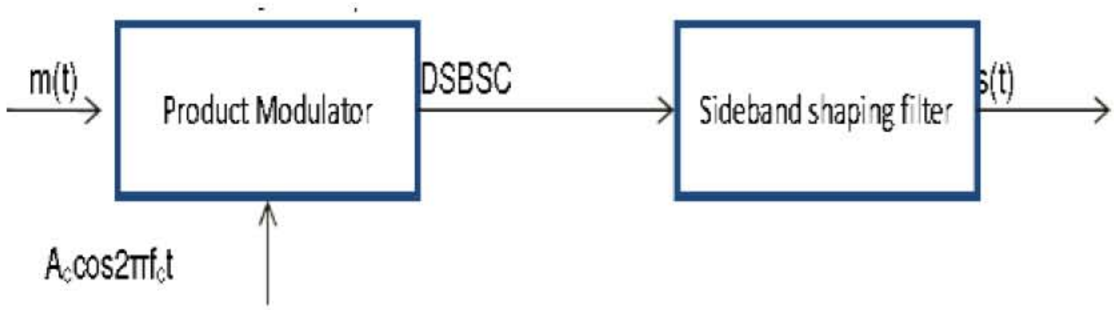


Fig.4. Block Diagram of VSB Modulator

The exact design of this filter depends on the spectrum of the VSB waves. The relation between filter transfer function H (f) and the spectrum of VSB waves is given by

$$S(f) = \frac{A_c}{2}[M(f - f_c) + M(f + f_c)]H(f) \dots\dots\dots\dots\dots\dots\dots (1)$$

Where M(f) is the spectrum of Message Signal. Now, we have to determine the specification for the filter transfer function H(f) It can be obtained by passing s(t) to a coherent detector and determining the necessary condition for undistorted version of the message signal m(t). Thus, s (t) is multiplied by a locally generated sinusoidal wave cos $(2\pi f_c t)$ which is synchronous with the carrier wave $A_c\cos(2\pi f_c t)$ in both frequency and phase, as in the below fig.,

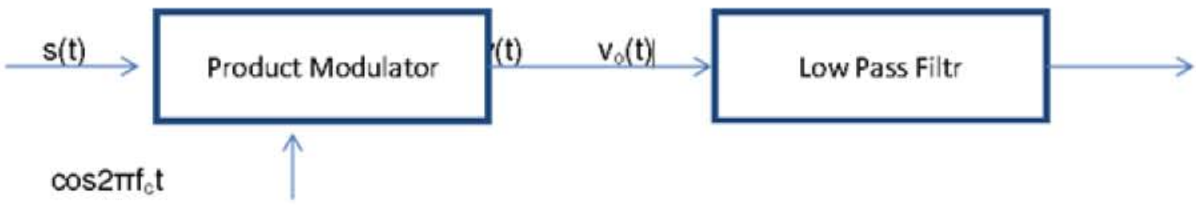


Fig.5 Block Diagram of VSB Demodulator

Then,

$$v(t) = s(t).\cos(2\pi f_c t) \dots\dots\dots\dots\dots\dots\dots\dots\dots(2)$$

In frequency domain equ (2) becomes,

$$V(f) = \frac{1}{2}[S(f - f_c) + S(f + f_c) \dots\dots\dots\dots\dots..(3)$$

Substituting equ.1 in equ. 3, we get

$$V(f) = \frac{1}{2}\left[{A_c}/{2}\,[M(f - 2f_c) + M(f)]H(f - f_c)\right]$$
$$+\frac{1}{2}\left[{A_c}/{2}\,[M(f + 2f_c) + M(f)]H(f + f_c)\right]$$

$$V(f) = \frac{A_c}{4}M(f)[H(f - f_c) + H(f + f_c)]$$
$$+\frac{A_c}{4}[M(f - 2f_c)H(f - f_c)$$
$$+ M(f + 2f_c)H(f + f_c)] \dots\dots\dots\dots\dots\dots..(4)$$

The spectrum of $V(f)$ as shown in fig below

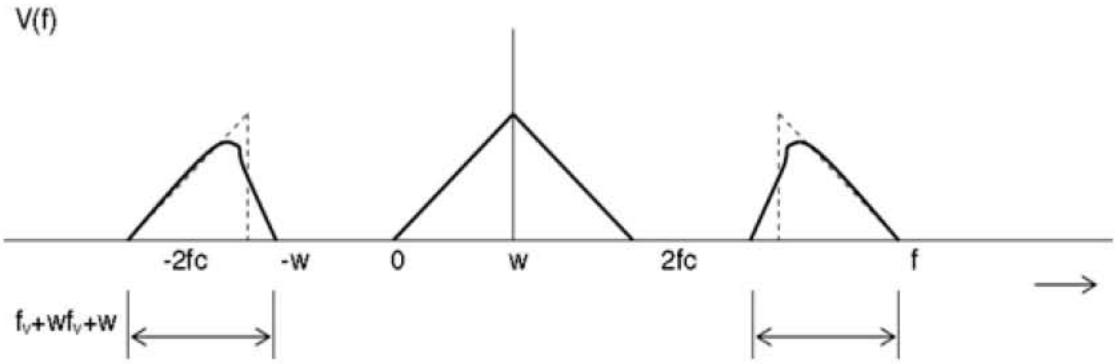


Fig.6. Spectrum of the product modulator output $v(t)$

Pass $v(t)$ to a LPF to eliminate VSB wave corresponding to $2f_c$

$$V_o(f) = \frac{A_c}{4} M(f)[H(f - f_c) + H(f + f_c)] \ldots\ldots\ldots\ldots\ldots (5)$$

The spectrum of $V_o(f)$ is in fig below,

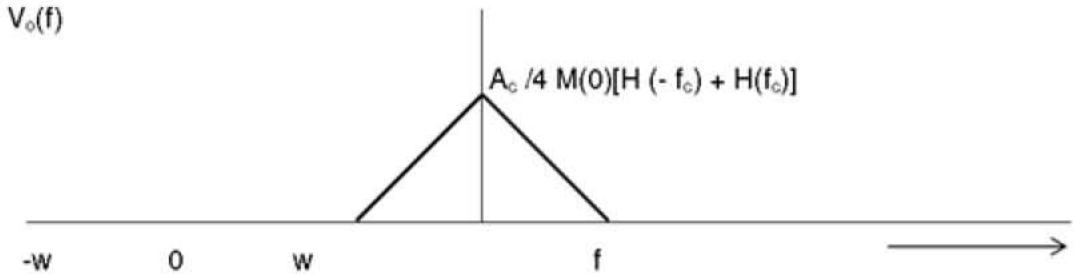


Fig.7. Spectrum of demodulated signal $v_o(t)$

For a distortion less reproduction of the original signal $m(t)$, $V_o(f)$ is to be a scaled version of $M(f)$. Therefore, the transfer function $H(f)$ must satisfy the condition

$$H(f - f_c) + H(f + f_c) = 2H(f_c) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Where $H(f_c)$ is a constant

Since $m(t)$ is a band limited signal, we need to satisfy eqn. (6) in the interval $-w \leq f \leq w$. The requirement of eqn. (6) is satisfied by using a filter whose transfer function is shown below.

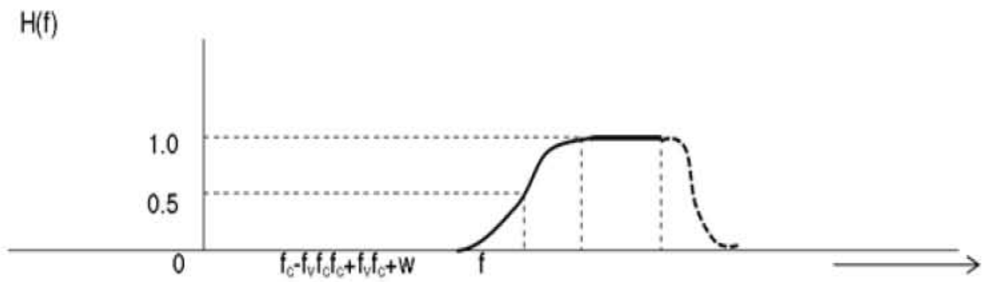


Fig.8. Frequency response of sideband shaping filter

Note: $H(f)$ is shown for positive frequencies only

The response is normalized so that $H(f)$ at f_c is 0.5. Inside this interval $f_c - f_v \leq f \leq f_c + f_v$ response exhibits odd symmetry *.i.e.,* sum of the values of $H(f)$ at any two frequencies equally displaced above & below is Unity

Similarly, the transfer function $H(f)$ of the filter for sending Lower sideband along with the vestige of the Upper sideband is shown in fig below,

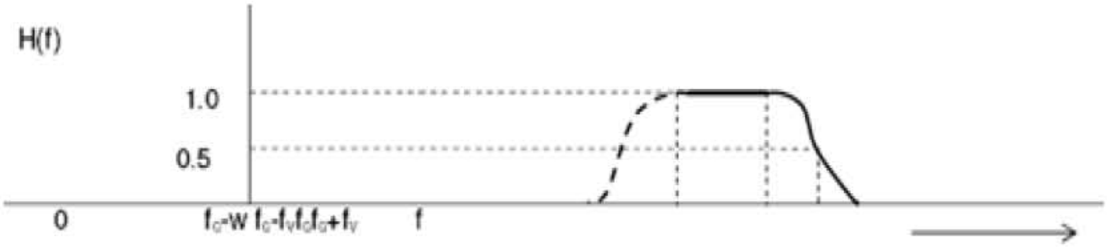


Fig.9. Frequency response of sideband shaping filter

Note: $H(f)$ is shown for positive frequencies only

Time Domain Description:

Time domain representation of VSB modulated wave, procedure is similar to SSB Modulated waves. Let s(t) denote a VSB modulated wave and assuming that s(t) containing Upper sideband along with the Vestige of the Lower sideband. VSB modulated wave s(t) is the output from Sideband shaping filter, whose input is DSBSC wave. The filter transfer function H(f) is of the form as in fig below,

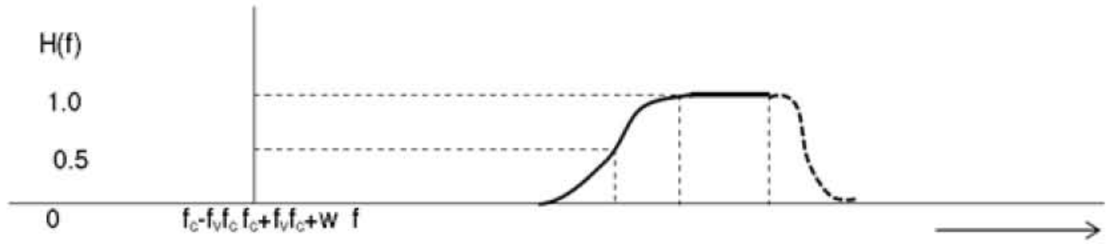


Fig.10. $H(f)$ of sideband shaping filter

The DSBSC modulated wave is

$$S_{DSBSC}(t) = A_c m(t) \cos(2\pi f_c t) \dots\dots\dots\dots\dots (1)$$

It is a band pass signal and has in-phase component only. Its low pass complex envelope is given by

$$\tilde{S}_{DSBSC}(t) = A_c m(t) \dots\dots\dots\dots\dots (2)$$

The VSB modulated wave is a band pass signal.

Let the low pass signal $\tilde{s}(t)$ denote the complex envelope of VSB wave $s(t)$, then

$$s(t) = Re[\tilde{s}(t) \exp(j2\pi f_c t)] \dots\dots\dots\dots\dots\dots\dots\dots (3)$$

To determine $\tilde{s}(t)$ we proceed as follows

1. The side band shaping filter transfer function $H(f)$ is replaced by its equivalent complex low pass transfer function denoted by $\tilde{H}(f)$ as shown in fig below

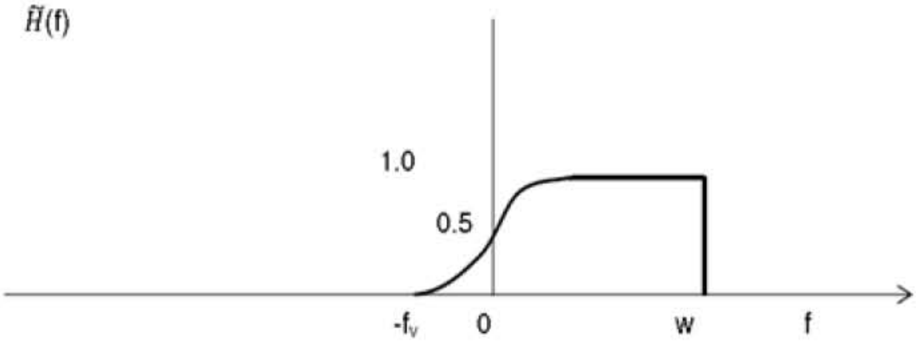


Fig.11. Low pass equivalent to $H(f)$

We may express $\tilde{H}(f)$ as the difference between two components $\tilde{H_u}(f)$ and $\tilde{H_v}(f)$ as

$$\tilde{H}(f) = \tilde{H_u}(f) - \tilde{H_v}(f) \ldots\ldots\ldots\ldots\ldots\ldots\ldots.(4)$$

These two components are considered individually as follows

i. The transfer function $\tilde{H_u}(f)$ pertains to a complex low pass filter equivalent to a ban pass filter design to reject the lower side band completely as

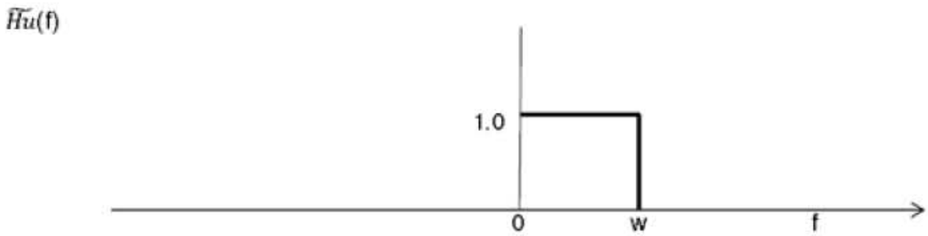


Fig.12. First component of $\tilde{H}(f)$

$$\tilde{H_u}(f) = \begin{cases} 1/2\,[1 + sgn(f)], & 0 < f < w \\ 0, & otherwise \end{cases} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

ii. The transfer function $\tilde{H_v}(f)$ accounts for the generation of vestige and removal of a corresponding portion from the upper side band.

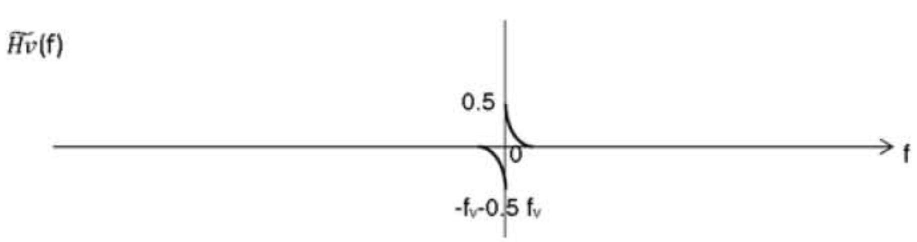


Fig.13. Second component of $\tilde{H}(f)$

Substitute eqn.5 in eqn. 4, we get,

$$\widetilde{H}(f) = \begin{cases} 1/2\,[1 + sgn(f) - 2\widetilde{H_v}(f)], & f_v < f < w \\ 0, & otherwise \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

The *sgn(f)* and $\widetilde{H_v}(f)$ are both odd functions of frequency. Hence both have purely imaginaryInverse Fourier transform (FT). Accordingly, the new transfer function is

$$H_Q(f) = 1/j\left[sgn(f) - 2\widetilde{H_v}(f)\right] \ldots\ldots\ldots\ldots\ldots\ldots.(7)$$

It has purely Inverse FT and $h_Q(t)$ *denote IFT of* $H_Q(f)$

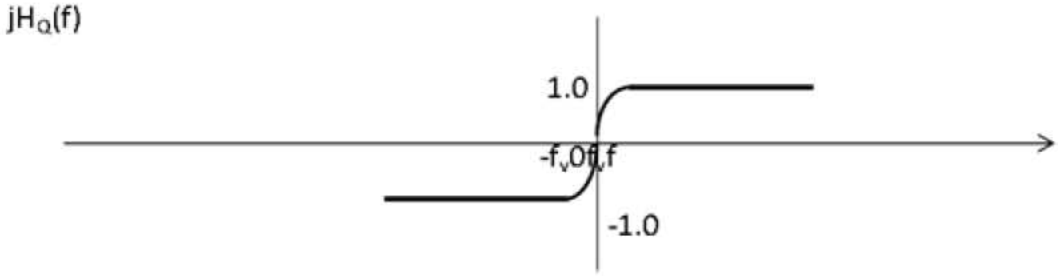


Fig.14. Transfer function of the filter $jH_Q(f)$

Rewrite eqn. (6) in terms of $H_Q(f)$ as

$$\widetilde{H}(f) = \begin{cases} 1/2\,[1 + jH_Q(f)], & f_v < f < w \\ 0, & otherwise \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots.(8)$$

2. The DSBSC modulated wave is replaced by its complex envelope as

$$\tilde{S}_{DSBSC}(f) = A_c M(f) \ldots\ldots\ldots\ldots\ldots.(9)$$

3. The desired complex envelope $\tilde{s}(t)$is determined by evaluating IFT of the product $\widetilde{H}(f)\tilde{S}_{DSBSC}(f)$. i.*e.*,

$$\tilde{S}(f) = \widetilde{H}(f)\tilde{S}_{DSBSC}(f) \ldots\ldots\ldots.(10)$$

$$\tilde{S}(f) = A_c/2\,[1 + jH_Q(f)]M(f) \ldots\ldots\ldots.(11)$$

Take IFT of eqn. 11, we get,

$$\tilde{s}(t) = A_c/2\,[m(t) + jm_Q(t)] \ldots\ldots\ldots\ldots\ldots.(12)$$

where $m_Q(t)$ is the response produced by passing the message through a low pass filter of impulse response $h_Q(t)$.

Substitute eqn. 12 in eqn.3 and simplify, we get

$$S(t) = {}^{A_c}/_2\, m(t)\cos(2\pi f_c t) - {}^{A_c}/_2\, m_Q(t)\sin(2\pi f_c t) \ldots\ldots\ldots\ldots\ldots..(13)$$

Where ${}^{A_c}/_2\, m(t) \ldots\ldots\ldots..in - phase\ component$

${}^{A_c}/_2\, m_Q(t) \ldots\ldots\ldots Quadrature\ component$

Note:

1. If vestigial side band is increased to full side band, VSB becomes DSCSB, i.e., $m_Q(t) = 0$.
2. If VSB is reduced to zero, VSB becomes SSB.i.e., $m_Q(t) = \hat{m}(t)$.Where the $\hat{m}$ (t) is the Hilbert transform of m(t)

Similarly if the VSB containing a vestige of the upper sideband, then s(t) is given by

$$S(t)= A_c/2\ m(t) \cos(2\pi f_c t) + A_c/2\ m_Q(t) \sin(2\pi f_c t)\text{------}(14)$$

Envelope detection of a VSB Wave plus Carrier

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

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

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

Where $k_a$is the modulation index; it determines the percentage modulation. When above signal $s(t)$ is passed through the envelope detector, the detector output is given by

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

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

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

Methods to reduce distortion

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

Comparison of AM Techniques:

S.No	Parameter	Standard AM	SSB	DSBSC	VSB
1	Power	High	Less	Medium	Less than DSBSC but greater than SSB
2	Bandwidth	$2f_m$	f_m	$2f_m$	f_m<BW< $2f_m$
3	Carrier Suppression	No	Yes	Yes	No
4	Receiver complexity	Simple	Complex	Complex	Simple
5	Modulation	Non-Linear	Linear	Linear	Linear
6	Sideband Suppression	No	One side band completely	No	One sideband suppressed parlty
7	Transmission efficiency	Minimum	Maximum	Moderate	Moderate
8	Application	Radio Communication	Point to Point Communication	Point to Point Communication	TV Broadcasting

Applications of different AM systems:

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

UNIT III
ANGLE MODULATION

- **Basic concepts**
- **Frequency Modulation**
- **Single tone frequency modulation**
- **Spectrum Analysis of Sinusoidal FM Wave**
- **Narrow band FM,Wide band FM,Constant Average Power**
- **Transmission bandwidth of FM Wave**
- **Generation of FM Waves:**
  - **Indirect FM,Direct FM: Varactor Diode and Reactance Modulator**
- **Detection of FM Waves:**
  - **Balanced Frequency discriminator, Zero crossing detector, Phase locked loop**
- **Comparison of FM & AM**
- **Pre-emphasis & de-emphasis**
- **FM Transmitter blockdiagram and explanation of each block**

Instantaneous Frequency

The frequency of a cosine function $x(t)$ that is given by

$$x(t) = \cos(\omega_c + \theta_0)$$

is equal to ω_c since it is a constant with respect to t, and the phase of the cosine is the constant θ_0. The angle of the cosine $\theta(t) = \omega_c t + \theta_0$ is a linear relationship with respect to t (a straight line with slope of ω_c and y–intercept of θ_0). However, for other sinusoidal functions, the frequency may itself be a function of time, and therefore, we should not think in terms of the constant frequency of the sinusoid but in terms of the INSTANTANEOUS frequency of the sinusoid since it is not constant for all t. Consider for example the following sinusoid

$$y(t) = \cos(\theta(t))$$

where $\theta(t)$is a function of time. The frequency of $y(t)$ in this case depends on the function of $\theta(t)$ and may itself be a function of time. The instantaneous frequency of $y(t)$ given above is defined as

$$\omega_i(t) = \frac{d}{dt}(\theta(t))$$

As a checkup for this definition, we know that the instantaneous frequency of $x(t)$ is equal to its frequency at all times (since the instantaneous frequency for that function is constant) and is equal to ω_c. Clearly this satisfies the definition of the instantaneous frequency since $\theta(t) = \omega_c t + \theta_0$ and therefore $\omega_i(t) = \omega_c$.

If we know the instantaneous frequency of some sinusoid from $-\infty$ to sometime t, we can find the angle of that sinusoid at time t using

$$\theta(t)) = \int_{-\infty}^{t} \omega(\alpha)d\alpha$$

Changing the angle $(\theta(t))$of some sinusoid is the bases for the two types of angle modulation: Phase and Frequency modulation techniques.

Phase Modulation (PM)

In this type of modulation, the phase of the carrier signal is directly changed by the message signal. The phase modulated signal will have the form

$$g_{pm}(t) = A\cos[\omega_C t + k_p(t)]$$

where A is a constant, ω_c is the carrier frequency, $m(t)$ is the message signal, and k_p is a parameter that specifies how much change in the angle occurs for every unit of change of $m(t)$. The phase and instantaneous frequency of this signal are

$$\theta_{pm}(t) = \omega_C t + k_p(t)$$

$$\omega_i(t) == \omega_C + k_p d/dt(m(t)$$

So, the frequency of a PM signal is proportional to the derivative of the message signal.

Frequency Modulation (FM)

This type of modulation changes the frequency of the carrier (not the phase as in PM) directly with the message signal. The FM modulated signal is

$$g_{fm}(t) = Acos[\omega_C t + k_f(t)\int_{-\infty}^{t} m(\alpha)d\alpha\,]$$

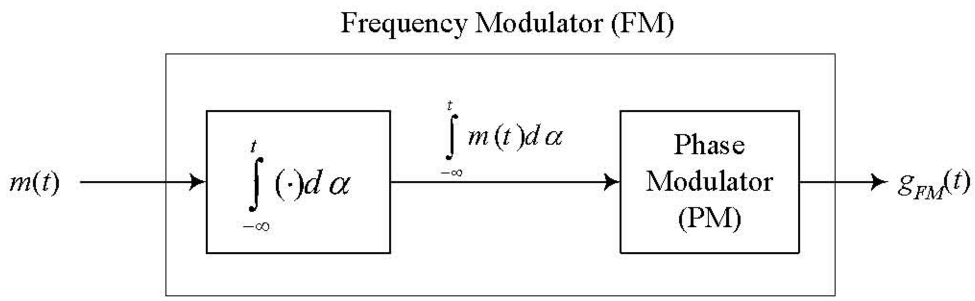
where k_f is a parameter that specifies how much change in the frequency occurs for every unit change of $m(t)$. The phase and instantaneous frequency of this FM are

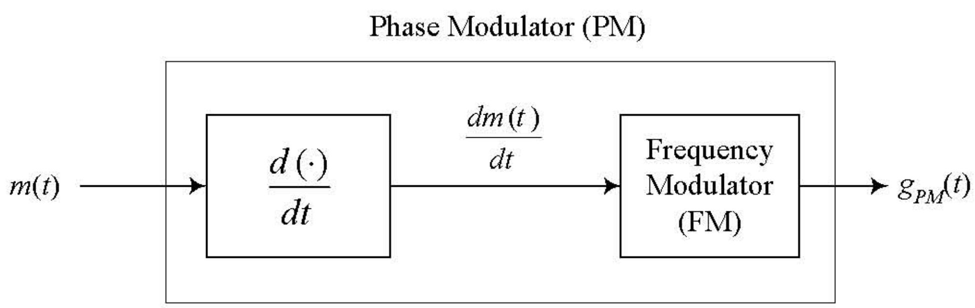
$$\theta_{fm}(t) = \omega_C t + k_f(t)\int_{-\infty}^{t} m(\alpha)d\alpha$$

$$\omega_i(t) = \omega_C t + k_f \frac{d}{dt}[\int_{-\infty}^{t} m(\alpha)d\alpha]$$

Relation between PM and FM

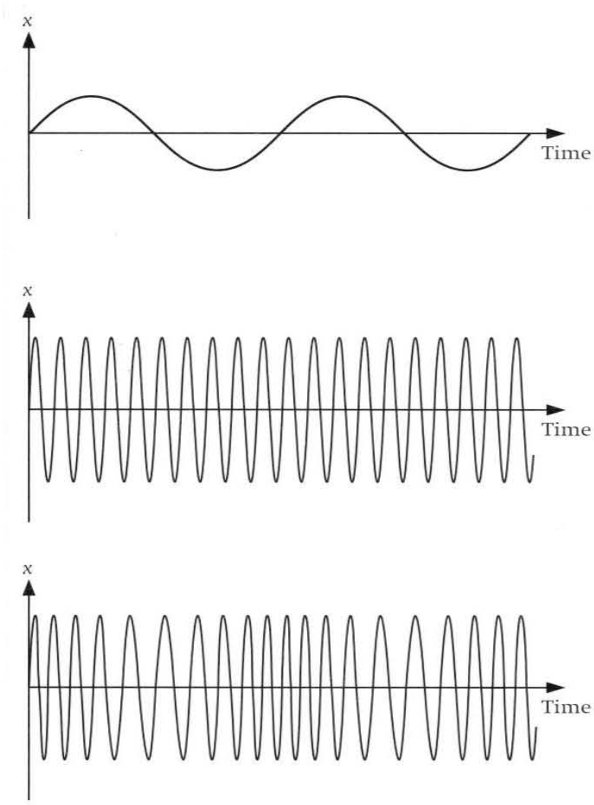
PM and FM are tightly related to each other. We see from the phase and frequency relations for PM and FM given above that replacing $m(t)$ in the PM signal with $\int_{-\infty}^{t} m(\alpha)d\alpha$ gives an FM signal and replacing $m(t)$ in the FM signal with $\frac{dm(t)}{dt}$ gives a PM signal. This is illustrated in the following block diagrams.





Frequency Modulation

In **F**requency **M**odulation (**FM**) the instantaneous value of the information signal controls the frequency of the carrier wave. This is illustrated in the following diagrams.



Notice that as the information signal increases, the frequency of the carrier increases, and as the information signal decreases, the frequency of the carrier decreases.

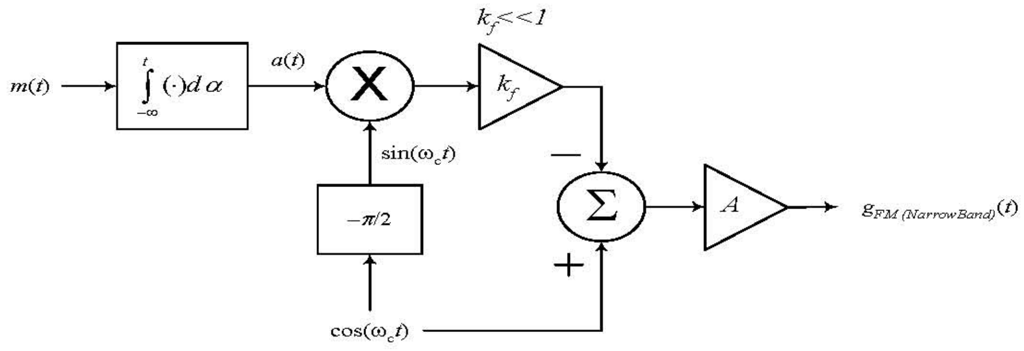
The frequency f_i of the information signal controls the rate at which the carrier frequency increases and decreases. As with AM, f_i must be less than f_c. The amplitude of the carrier remains constant throughout this process.

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.

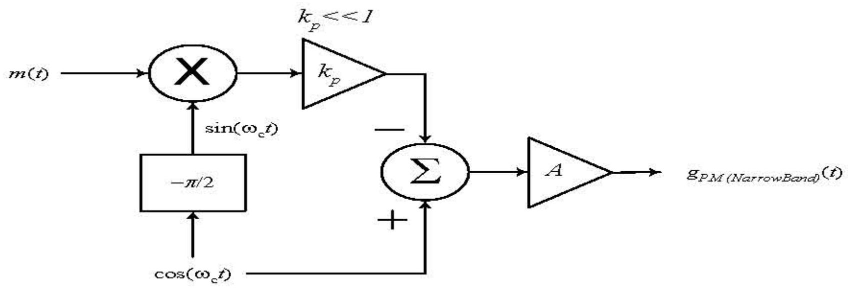
The maximum change that can occur to the carrier from its base value f_c is called the frequency deviation, and is given the symbol Δf_c. This sets the dynamic range (i.e. voltage range) of the transmission. The dynamic range is the ratio of the largest and smallest analogue information signals that can be transmitted.

Construction of Narrowband Frequency and Phase Modulators

The above approximations for narrowband FM and PM can be easily used to construct modulators for both types of signals



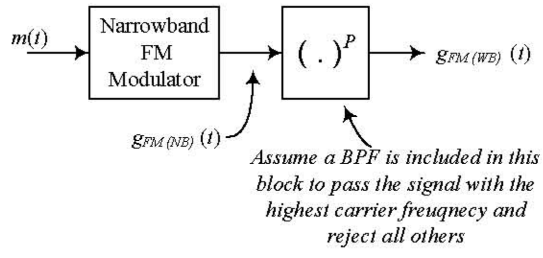
Narrowband FM Modulator



Narrowband PM Modulator

Generation of Wideband FM Signals

Consider the following block diagram



A narrowband FM signal can be generated easily using the block diagram of the narrowband FM modulator that was described in a previous lecture. The narrowband FM modulator generates a narrowband FM signal using simple components such as an integrator (an OpAmp), oscillators, multipliers, and adders. The generated narrowband FM signal can be converted to a wideband FM signal by simply passing it through a non–linear device with power *P*. Both the carrier frequency and the frequency deviation Δf of the narrowband signal are increased by a factor *P*. Sometimes, the desired increase in the carrier frequency and the desired increase in Δf are different. In this case, we increase Δf to the desired value and use a frequency shifter (multiplication by a sinusoid followed by a BPF) to change the carrier frequency to the desired value.

SINGLE-TONE FREQENCY MODULATION

Time-Domain Expression

Since the FM wave is a nonlinear function of the modulating wave, the frequency modulation is a nonlinear process. The analysis of nonlinear process is the difficult task. In this section, we will study single-tone frequency modulation in detail to simplify the analysis and to get thorough understanding about FM.

Let us consider a single-tone sinusoidal message signal defined by

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

The instantaneous frequency is then

$$f(t) = f_c + k_f A_m \cos(2\pi f_n t) = f_c + \Delta f \cos(2n f_n t)$$

where

$$\Delta f = k_f A_n$$

$$\begin{aligned}\theta(t) &= 2\pi f_c t + 2\pi k_f \int_0^t A_m \cos(2\pi f_m t)dt \\ &= 2\pi f_c t + 2\pi k_f \frac{A_m}{2\pi f_m} \sin(2\pi f_m t) \\ &= 2\pi f_c t + k_f \frac{A_m}{f_m} \sin(2\pi f_m t) \\ &= \frac{\Delta f}{f_m} \sin(2\pi f_m t) + 2\pi f_c t\end{aligned}$$

$$\theta(t) = 2\pi f_c t + \beta_f \sin(2\pi f_m t)$$

$$\text{Where } \beta_f = \frac{\Delta f}{f_m}$$

is the modulation index of the FM wave. Therefore, the single-tone FM wave is expressed by

$$s_{FM}(t) = A_c \cos[2\pi f_c t + \beta_f \sin(2\pi f_m t)]$$

This is the desired time-domain expression of the single-tone FM wave

Similarly, **single-tone phase modulated wave** may be determined from Eq.as

$$s_{PM}(t) = A_c \cos[2\text{n}f_c t + k_p A_n \cos(2\text{n}f_n t)]$$

$$\text{or, } s_{PM}(t) = A_c \cos[2\text{n}f_c t + \text{þ}_p \cos(2\text{n}f_n t)]$$

where

$$\text{þ}_p = k_p A_n$$

is the modulation index of the single-tone phase modulated wave.

The frequency deviation of the single-tone PM wave is

$$S_{FM}(t) = A_c \cos[2\pi f_c t + \beta_f \sin(2\pi f_m t)]$$

Spectral Analysis of Single-Tone FM Wave

The above Eq. can be rewritten as

$$s_{FM}(t) = \text{Re}\{A_c e^{j2\text{n}f_c t} e^{j\text{þ} \sin(2\text{n}f_n t)}\}$$

For simplicity, the modulation index of FM has been considered as þ instead of þ_f afterward. Since $\sin(2\text{n}f_n t)$ is periodic with fundamental period $T = 1/f_n$, the complex expontial $e^{j\text{þ} \sin(2\text{n}f_n t)}$ is also periodic with the same fundamental period. Therefore, this complex exponential can be expanded in Fourier series representation as

$$e^{j\beta \sin(2\pi f_m t)} = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_m t}$$

where the Fourier series coefficients c_n are obtained as

$$c_n = \frac{1}{T}\int_{-T/_2}^{T/_2} e^{j\beta \sin(2\pi f_m t)} e^{-j2\pi n f_m t} dt \tag{5.24}$$

Let $2\pi f_m t = x$, then Eq. (5.24) reduces to

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\beta \sin(x)} e^{-jnx} dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j(\beta \sin(x) - nx)} dx \tag{5.25}$$

The integral on the right-hand side is known as the n[th] order Bessel function of the first kind and is denoted by $J_n(\beta)$. Therefore, $c_n = J_n(\beta)$ and Eq. (4.23) can be written as

$$e^{j\beta \sin(2\pi f_m t)} = \sum_{n=-\infty}^{\infty} J_n(\beta) e^{j2\pi n f_m t} \tag{5.26}$$

By substituting Eq. (5.26) in Eq. (5.22), we get

$$\begin{aligned} s_{FM}(t) &= \mathrm{Re}\left\{A_c \sum_{n=-\infty}^{\infty} J_n(\beta) e^{j2\pi n f_m t}\, e^{j2\pi f_c t}\right\} \\ &= A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + n f_m)t] \end{aligned} \tag{5.27}$$

Taking Fourier transform of Eq. (5.27), we get

$$S(f) = \frac{1}{2} A_c \sum_{n=-\infty}^{\infty} J_n(\beta)[\delta(f - f_c - n f_m) + \delta(f + f_c + n f_m)] \tag{5.28}$$

From the spectral analysis we see that there is a carrier component and a number of side-frequencies around the carrier frequency at $\pm n f_m$.

The Bessel function may be expanded in a power series given by

$$J_n(\beta) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{1}{2}\beta\right)^{n+2k}}{k!\,(k+n)!} \tag{5.29}$$

Plots of Bessel function $J_n(\beta)$ versus modulation index β for $n = 0, 1, 2, 3, 4$ are shown in Figure 5.3.

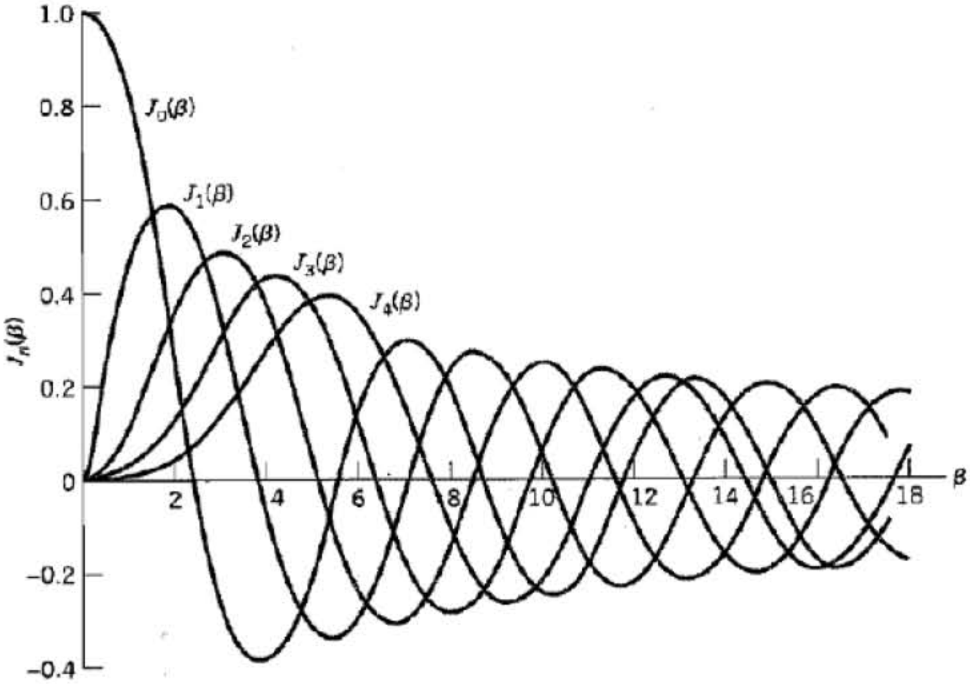


Figure 5.3 Plot of Bessel function as a function of modulation index.

Figure 5.3 shows that for any fixed value of n, the magnitude of $J_n(\beta)$ decreases as β increases. One property of Bessel function is that

$$J_{-n}(\beta) = \begin{cases} J_n(\beta), & n \text{ even} \\ -J_n(\beta), & n \text{ odd} \end{cases} \tag{5.30}$$

One more property of Bessel function is that

$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \tag{5.31}$$

(iii) The average power of the FM wave remains constant. To prove this, let us determine the average power of Eq. (5.27) which is equal to

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

Using Eq. (5.31), the average power P is now

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

TRANSMISSION BANDWIDTH OF FM WAVE

The transmission bandwidth of an FM wave depends on the modulation index þ. The modulation index, on the other hand, depends on the modulating amplitude and modulating frequency. It is almost impossible to determine the exact bandwidth of the FM wave. Rather, we use a rule-of-thumb expression for determining the FM bandwidth.

For single-tone frequency modulation, the approximated bandwidth is determined by the expression

$$B = 2(\Delta f + f_m) = 2(\beta + 1)f_m = 2\Delta f(1 + \frac{1}{\beta})$$

This expression is regarded as the Carson's rule. The FM bandwidth determined by this rule accommodates at least 98 % of the total power.

For an arbitrary message signal n(t) with bandwidth or maximum frequency W, the bandwidth of the corresponding FM wave may be determined by Carson's rule as

$$B = 2(\Delta f + W) = 2(D + 1)W = 2\Delta f(1 + \frac{1}{D})$$

GENERATION OF FM WAVES

FM waves are normally generated by two methods: indirect method and direct method.

Indirect Method (Armstrong Method) of FM Generation

In this method, narrow-band FM wave is generated first by using phase modulator and then the wideband FM with desired frequency deviation is obtained by using frequency multipliers.

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

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

$$\emptyset(t) = 2\pi k_f \int_0^t m(t)\, dt$$

$$s(t) = A_c \cos(2\pi f_c t) \cos[\emptyset(t)] - A_c \sin(2\pi f_c t) \sin[\emptyset(t)]$$

The above eq is the expression for narrow band FM wave
In this case

$$\cos[\emptyset(t)] \approx 1 \text{ and } \sin[\emptyset(t)] \approx \emptyset(t)$$

$$s(t) = A_c \cos(2\pi f_c t) - A_c \sin(2\pi f_c t)\emptyset(t)$$

$$\text{or, } s(t) = A_c \cos(2\pi f_c t) - 2\pi A_c k_f \sin(2\pi f_c t)\int_0^t m(t)dt$$

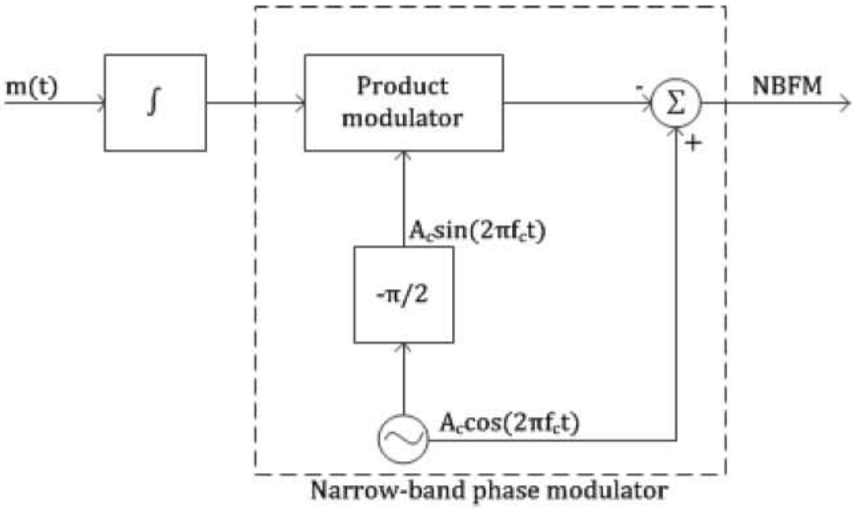


Fig: Narrowband FM Generator

The frequency deviation Δf is very small in narrow-band FM wave. To produce wideband FM, we have to increase the value of Δf to a desired level. This is achieved by means of one or multiple frequency multipliers. A frequency multiplier consists of a nonlinear device and a bandpass filter. The n^{th} order nonlinear device produces a dc component and n number of frequency modulated waves with carrier frequencies f_c, $2f_c$, … nf_c and frequency deviations Δf, $2\Delta f$, … $n\Delta f$, respectively. If we want an FM wave with frequency deviation of $6\Delta f$, then we may use a 6^{th} order nonlinear device or one 2^{nd} order and one 3^{rd} order nonlinear devices in cascade followed by a bandpass filter centered at $6f_c$. Normally, we may require very high value of frequency deviation. This automatically increases the carrier frequency by the same factor which may be higher than the required carrier frequency. We may shift the carrier frequency to the desired level by using mixer which does not change the frequency deviation.

The narrowband FM has some distortion due to the approximation made in deriving the expression of narrowband FM from the general expression. This produces some amplitude modulation in the narrowband FM which is removed by using a limiter in frequency multiplier.

Direct Method of FM Generation

In this method, the instantaneous frequency $f(t)$ of the carrier signal c(t) is varied directly with the instantaneous value of the modulating signal n(t). For this, an oscillator is used in which any one of the reactive components (either C or L) of the resonant network of the oscillator is varied linearly with n(t). We can use a varactor diode or a varicap as a voltage-variable capacitor whose capacitance solely depends on the reverse-bias voltage applied across it. To vary such capacitance linearly with n(t), we have to reverse-bias the diode by the fixed DC voltage and operate within a small linear portion of the capacitance-voltage characteristic curve. The unmodulated fixed capacitance C_0 is linearly varied by n(t) such that the resultant capacitance becomes

$$C(t) = C_0 - kn(t)$$

where the constant k is the sensitivity of the varactor diode (measured in capacitance per volt).

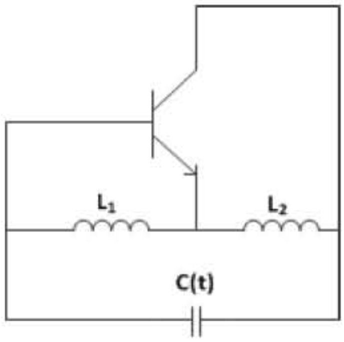


fig: Hartley oscillator for FM generation

The above figure shows the simplified diagram of the Hartley oscillator in which is implemented the above discussed scheme. The frequency of oscillation for such an oscillator is given

$$f(t) = \frac{1}{2\pi\sqrt{(L_1 + L_2)C(t)}}$$

$$f(t) = \frac{1}{2\pi\sqrt{(L_1 + L_2)(C_0 - km(t))}}$$

$$= \frac{1}{2\pi\sqrt{(L_1 + L_2)C_0}} \frac{1}{\sqrt{1 - \frac{km(t)}{C_0}}}$$

$$\text{or,} \qquad f(t) = f_c\left(1 - \frac{km(t)}{C_0}\right)^{-1/2}$$

where f_c is the unmodulated frequency of oscillation. Assuming,

$$\frac{km(t)}{C_0} \ll 1$$

we have from binomial expansion,

$$\left(1 - \frac{km(t)}{C_0}\right)^{-1/2} \approx 1 + \frac{km(t)}{2C_0}$$

$$f(t) \approx f_c\left(1 + \frac{km(t)}{2C_0}\right)$$

$$= f_c + \frac{kf_c m(t)}{2C_0}$$

$$\text{or, } f(t) = f_c + k_f m(t)$$

$$k_f = \frac{kf_c}{2C_0}$$

is the frequency sensitivity of the modulator. The Eq. (5.42) is the required expression for the instantaneous frequency of an FM wave. In this way, we can generate an FM wave by direct method.

Direct FM may be generated also by a device in which the inductance of the resonant circuit is linearly varied by a modulating signal n(t); in this case the modulating signal being the current.

The main advantage of the direct method is that it produces sufficiently high frequency deviation, thus requiring little frequency multiplication. But, it has poor frequency stability. A feedback scheme is used to stabilize the frequency in which the output frequency is compared with the constant frequency generated by highly stable crystal oscillator and the error signal is feedback to stabilize the frequency.

DEMODULATION OF FM WAVES

The process to extract the message signal from a frequency modulated wave is known as frequency demodulation. As the information in an FM wave is contained in its instantaneous frequency, the frequency demodulator has the task of changing frequency variations to amplitude variations. Frequency demodulation method is generally categorized into two types: direct method and indirect method. Under direct method category, we will discuss about limiter discriminator method and under indirect method, phase-locked loop (PLL) will be discussed.

Limiter Discriminator Method

Recalling the expression of FM signal,

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

In this method, extraction of n(t) from the above equation involves the three steps: amplitude limit, discrimination, and envelope detection.

A. Amplitude Limit

During propagation of the FM signal from transmitter to receiver the amplitude of the FM wave (supposed to be constant) may undergo changes due to fading and noise. Therefore, before further processing, the amplitude of the FM signal is limited to reduce the effect of fading and noise by using limiter as discussed in the section 5.9. The amplitude limitation will not affect the message signal as the amplitude of FM does not carry any information of the message signal.

B. Discrimination/ Differentiation

In this step we differentiate the FM signal as given by

$$\frac{ds(t)}{dt} = \frac{d}{dt}\left\{A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right]\right\}$$

$$= \frac{d\left\{A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right]\right\}}{d\left\{2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right\}} \frac{d\left\{2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right\}}{dt}$$

$$= -A_c[2\pi f_c + 2\pi k_f m(t)] \sin\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right]$$

Here both the amplitude and frequency of this signal are modulated.

In this case, the differentiator is nothing but a circuit that converts change in frequency into corresponding change in voltage or current as shown in Fig. 5.11. The ideal differentiator has transfer function

$$\mathrm{H(jw) = j2nf}$$

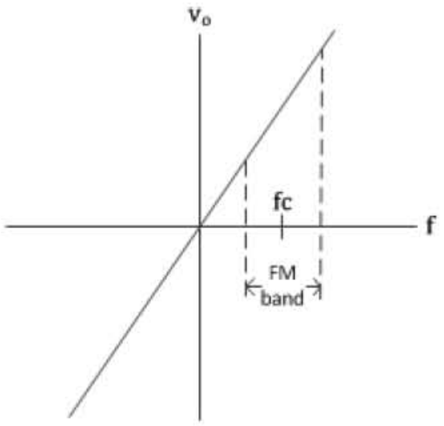


Figure : Transfer function of ideal differentiator.

Instead of ideal differentiator, any circuit can be used whose frequency response is linear for some band in positive slope. This method is known as slope detection. For this, linear segment with positive slope of RC high pass filter or LC tank circuit can be used. Figure 5.13 shows the use of an LC circuit as a differentiator. The drawback is the limited linear portion in the

slope of the tank circuit. This is not suitable for wideband FM where the peak frequency deviation is high.

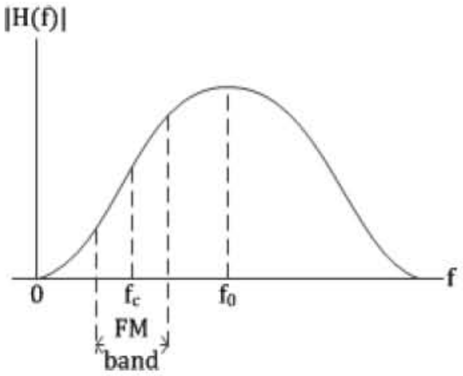


Figure : Use of LC tank circuit as a differentiator.

A better solution is the ratio or balanced slope detector in which two tank circuits tuned at $f_c + \Delta f$ and $f_c - \Delta f$ are used to extend the linear portion as shown in below figure.

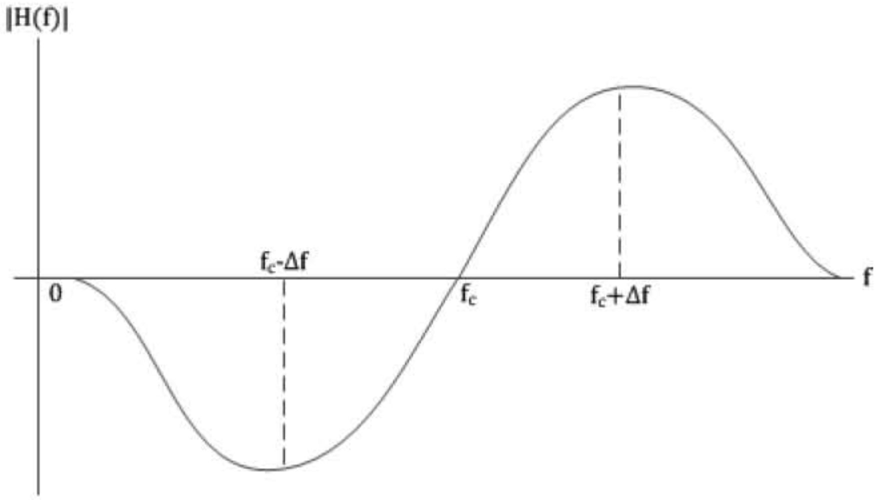


Figure : Frequency response of balanced slope detector.

Another detector called Foster-seely discriminator eliminates two tank circuits but still offer the same linear as the ratio detector.

C. Envelope Detection

The third step is to send the differentiated signal to the envelope detector to recover the message signal.

Phase-Lockd Loop (PLL) as FM Demodulator

A PLL consists of a multiplier, a loop filter, and a VCO connected together to form a feedback loop as shown in Fig. 5.15. Let the input signal be an FM wave as defined by

$$s(t) = A_c \cos[2\pi f_c t + \emptyset_1(t)]$$

$$\emptyset_1(t) = 2\pi k_f \int_0^t m(t)dt$$

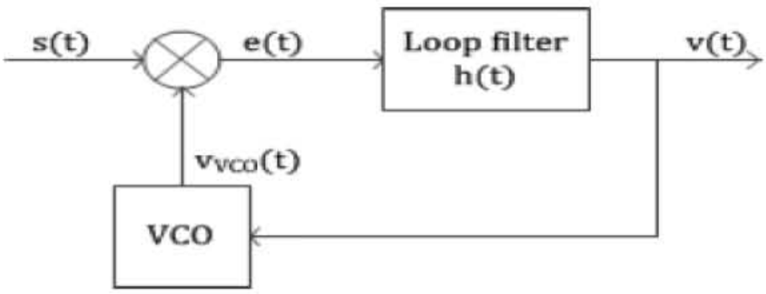


Fig: PLL Demodulator

Let the VCO output be defined by

$$v_{VCO}(t) = A_v \sin[2\pi f_c t + \emptyset_2(t)]$$

where

$$\emptyset_2(t) = 2\pi k_v \int_0^t v(t)dt$$

Here k_v is the frequency sensitivity of the VCO measured in hertz per volt. The multiplication of $s(t)$ and $v_{VCO}(t)$ results

$$s(t)v_{VCO}(t) = A_c \cos[2\pi f_c t + \emptyset_1(t)]\, A_v \sin[2\pi f_c t + \emptyset_2(t)]$$

$$= \frac{A_c A_v}{2} \sin[4\pi f_c t + \emptyset_1(t) + \emptyset_2(t)] + \frac{A_c A_v}{2} \sin[\emptyset_2(t) - \emptyset_1(t)]$$

The high-frequency component is removed by the low-pass filtering of the loop filter. Therefore, the input signal to the loop filter can be considered as

$$e(t) = \frac{A_c A_v}{2} \sin[\emptyset_2(t) - \emptyset_1(t)]$$

The difference $\emptyset_2(t) - \emptyset_1(t) = \emptyset_e(t)$ constitutes the phase error. Let us assume that the PLL is in phase lock so that the phase error is very small. Then,

$$\sin[\emptyset_2(t) - \emptyset_1(t)] \approx \emptyset_2(t) - \emptyset_1(t)$$

$$\emptyset_e(t) = 2\pi k_v \int_0^t v(t)dt - \emptyset_1(t)$$

$$e(t) = \frac{A_c A_v}{2} \emptyset_e(t)$$

Differentiating Eq. (5.48) with respect to time, we get

$$\frac{d\emptyset_e(t)}{dt} = 2\pi k_v v(t) - \frac{d\emptyset_1(t)}{dt}$$

Since

$$v(t) = e(t) * h(t) = \frac{A_c A_v}{2} [\emptyset_e(t) * h(t)]$$

Eq. (5.50) becomes

$$\frac{d\emptyset_e(t)}{dt} = 2\pi k_v \frac{A_c A_v}{2} [\emptyset_e(t) * h(t)] - \frac{d\emptyset_1(t)}{dt}$$

$$or, \pi k_v A_c A_v [\emptyset_e(t) * h(t)] - \frac{d\emptyset_e(t)}{dt} = \frac{d\emptyset_1(t)}{dt}$$

Taking Fourier transform of Eq. (5.52), we get

$$\pi k_v A_c A_v \emptyset_e(f) H(f) - j2\pi f \emptyset_e(f) = j2\pi f \emptyset_1(f)$$

$$or, \emptyset_e(f) = \frac{j2\pi f}{\pi k_v A_c A_v H(f) - j2\pi f} \emptyset_1(f)$$

$$or, \emptyset_e(f) = \frac{1}{\frac{\pi k_v A_c A_v}{j2\pi f} H(f) - 1} \emptyset_1(f)$$

Fourier transform of Eq. (5.51) is

$$V(f) = \frac{A_c A_v}{2} \emptyset_e(f) H(f)$$

Substituting Eq. (5.53) into (5.54), we get

$$V(f) = \frac{A_c A_v}{2} \frac{1}{\frac{\pi k_v A_c A_v}{j2\pi f} H(f) - 1} \emptyset_1(f) H(f)$$

We design $H(f)$ such that

$$\left| \frac{\pi k_v A_c A_v}{j2\pi f} H(f) \right| \gg 1$$

in the frequency band $|f| < W$ of the message signal.

78

$$\therefore V(f) = \frac{A_c A_v}{2} \frac{1}{\frac{\pi k_v A_c A_v}{j2\pi f} H(f)} \emptyset_1(f) H(f)$$

$$or, V(f) = \frac{1}{2\pi k_v} j2\pi f \emptyset_1(f)$$

Taking inverse Fourier transform of Eq. (4.56), we get

$$v(t) = \frac{1}{2\pi k_v} \frac{d\emptyset_1(t)}{dt}$$

$$= \frac{1}{2\pi k_v} \frac{d}{dt} \left\{ 2\pi k_f \int_0^t m(t) dt \right\}$$

$$= \frac{1}{2\pi k_v} 2\pi k_f m(t)$$

$$\therefore v(t) = \frac{k_f}{k_v} m(t)$$

Since the control voltage of the VCO is proportional to the message signal, v(t) is the demodulated signal.

We observe that the output of the loop filter with frequency response H(*f*) is the desired message signal. Hence the bandwidth of H(*f*) should be the same as the bandwidth W of the message signal. Consequently, the noise at the output of the loop filter is also limited to the bandwidth W. On the other hand, the output from the VCO is a wideband FM signal with an instantaneous frequency that follows the instantaneous frequency of the received FM signal.

PREEMPHASIS AND DEEMPHASIS NETWORK

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

The preemphasis network and its frequency response are shown in Figure 5.19 (a) and (b) respectively. Similarly, the counter part for deemphasis network is shown in Figure 5.20.

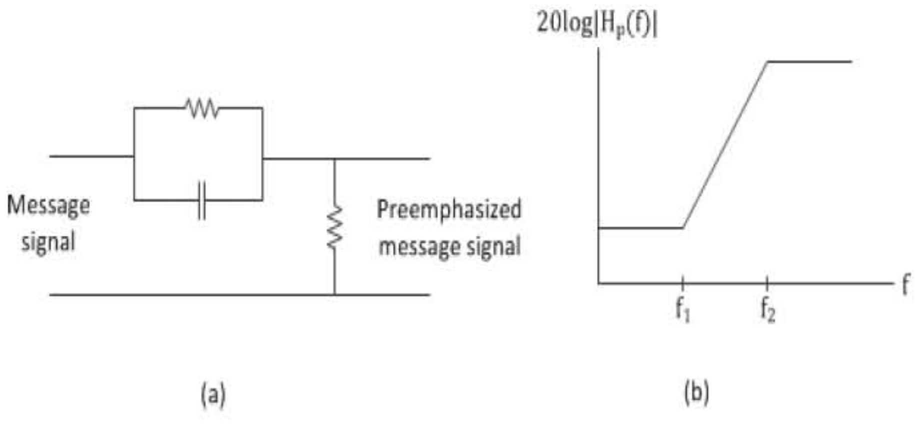


Figure ; (a) Pre-emphasis network. (b) Frequency response of pre-emphasis network.

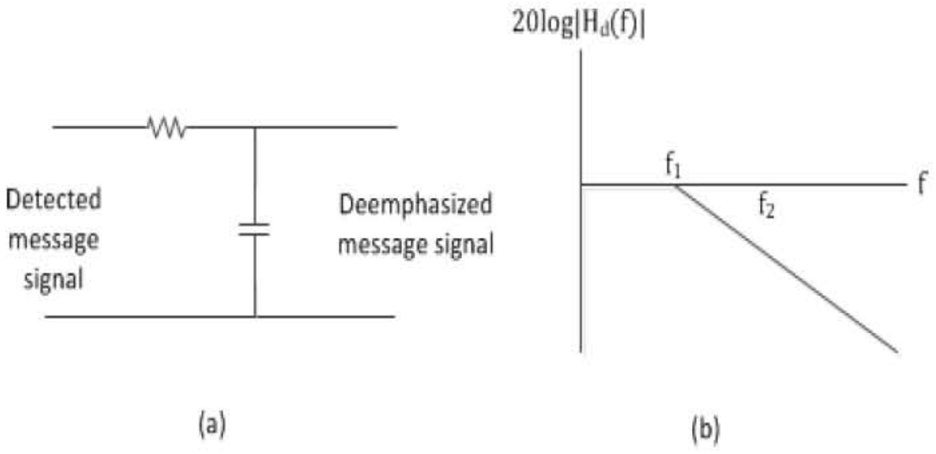


Figure (a) Deemphasis network. (b) Frequency response of Deemphasis network.

In FM broadcasting, f_1 and f_2 are normally chosen to be 2.1 kHz and 30 kHz respectively.

The frequency response of preemphasis network is

$$H_p(f) = \left(\frac{w_2}{w_1}\right)\frac{jw + w_1}{jw + w_2}$$

Here, $w = 2\pi f$ and $w_1 = 2\pi f_1$. For $w \ll w_1$,

$$H_p(f) \approx 1$$

And for $w_1 \ll w \ll w_2$,

$$H_p(f) \approx \frac{j2\pi f}{w_1}$$

So, the amplitude of frequency components less than 2.1 kHz are left unchanged and greater than that are increased proportional to f.

The frequency response of deemphasis network is

$$H_d(f) = \frac{w_1}{j2\pi f + w_1}$$

For $w \ll w_2$,

$$H_p(f) \approx \frac{j2\pi f + w_1}{w_1}$$

such that

$$H_p(f)H_d(f) \approx 1$$

over the baseband of 0 to 15 KHz.

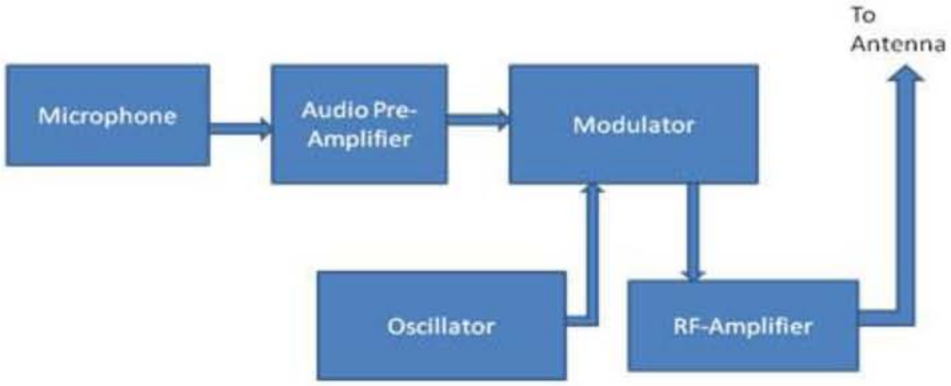
Comparison of AM and FM:

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

FM Transmitter

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

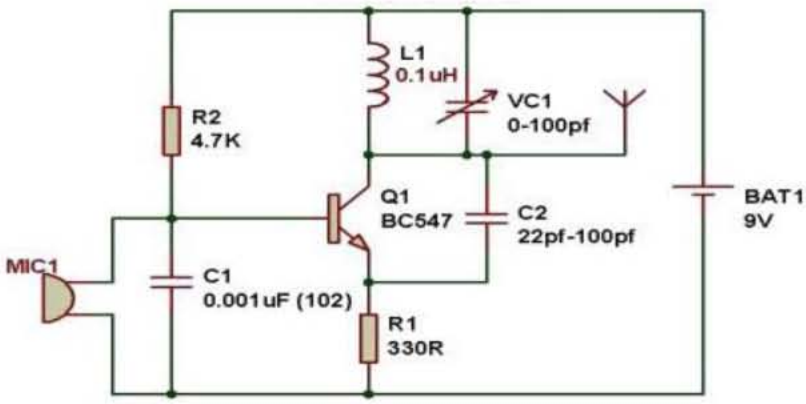
The FM transmitter is a low power transmitter and it uses FM waves for transmitting the sound, this transmitter transmits the audio signals through the carrier wave by the difference of frequency. The carrier wave frequency is equivalent to the audio signal of the amplitude and the FM transmitter produce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



Block Diagram of FM Transmitter

Worlng of FM Transmitter Circuit

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



FM Transmitter circuit

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

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

Application of Fm Transmitter

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

Advantages of the FM Transmitters

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

UNIT IV
NOISE

- Noise in communication System,
- White Noise
- Narrowband Noise –In phase and Quadrature phase components
- Noise Bandwidth
- Noise Figure
- Noise Temperature
- Noise in DSB& SSB System
- Noise in AM System
- Noise in Angle Modulation System
- Threshold effect in Angle Modulation System

Noise in communication system

A signal may be contaminated along the path by noise. Noise may be defined as any unwanted introduction of energy into the desired signal. In radio receivers, noise may produce "hiss" in the loudspeaker output. Noise is random and unpredictable.

Noise is produced both external and internal to the system. External noise includes atmospheric noise (e.g., from lightning), galactic noise (thermal radiation from cosmic bodies), and industrial noise (e.g., from motors, ignition). We can minimize or eliminate external noise by proper system design. On the other hand, internal noise is generated inside the system. It is resulted due to random motion of charged particles in resistors, conductors, and electronic devices. With proper system design, it can be minimized but never can be eliminated. This is the major constraint in the rate of telecommunications.

- Noise is unwanted signal that affects wanted signal
- Noise is random signal that exists in communication systems

Effect of noise

- Degrades system performance (Analog and digital)
- Receiver cannot distinguish signal from noise
- Efficiency of communication system reduces

Types of noise

- Thermal noise/white noise/Johnson noise or fluctuation noise
- Shot noise
- Noise temperature
- Quantization noise

Noise temperature

Equivalent noise temperature is not the physical temperature of amplifier, but a theoretical construct, that is an equivalent temperature that produces that amount of noise power

$$T_e = (F - 1)$$

White noise

One of the very important random processes is the *white noise* process. Noises in many practical situations are approximated by the white noise process. Most importantly, the white noise plays an important role in modelling of WSS signals.

A white noise process w(t) is a random process that has constant power spectral density at all frequencies. Thus

$$S_W(\omega) = \frac{N_0}{2} \qquad -\infty < \omega < \infty$$

where N_0 is a real constant and called the *intensity* of the white noise. The corresponding autocorrelation function is given by

$R_W(\tau) = \frac{N}{2}\delta(\tau)$ where $\delta(\tau)$ is the Dirac delta

The average power of white noise

$$P_{avg} = EW^2(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{N}{2}\, d\,\omega \rightarrow \infty$$

The autocorrelation function and the PSD of a white noise process is shown in Figure 1 below.

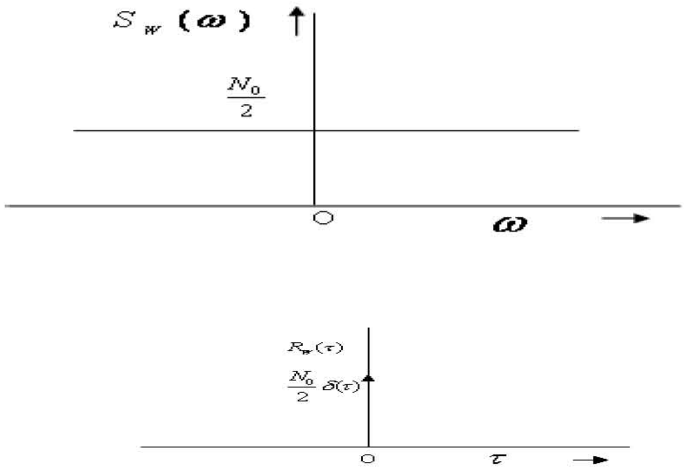


fig: auto correlation and psd of white noise

NARROWBAND NOISE (NBN)

In most communication systems, we are often dealing with band-pass filtering of signals. Wideband noise will be shaped into band limited noise. If the bandwidth of the band limited noise is relatively small compared to the carrier frequency, we refer to this as *narrowband noise*.

the narrowband noise is expressed as as

$$n(t)=x(t)\cos(2\pi f_c t)-y(t)\sin(2\pi f_c t)$$

where f_c is the carrier frequency within the band occupied by the noise. $x(t)$ and $y(t)$ are known as the *quadrature components* of the noise $n(t)$. The Hibert transform of $n(t)$ is

Proof.
The Fourier transform of $n(t)$ is

$$N(f)=\frac{1}{2}X(f-f_c)+\frac{1}{2}X(f+f_c)+\frac{1}{2}jY(f-f_c)-\frac{1}{2}jY(f+f_c)$$

Let $\hat{N}(f)$ be the Fourier transform of $\hat{n}(t)$. In the frequency domain, $\hat{N}(f) = N(f)[-j\ \mathrm{sgn}(f)]$. We simply multiply all positive frequency components of $N(f)$ by $-j$ and all negative frequency components of $N(f)$ by j. Thus

$$\hat{n}(t)=H[n(t)]=x(t)\sin(2\pi f_c t)+y(t)\cos(2\pi f_c t)$$

$$\hat{N}(f)=-j\frac{1}{2}X(f-f_c)+j\frac{1}{2}X(f+f_c)-j\frac{1}{2}jY(f-f_c)-j\frac{1}{2}jY(f+f_c)$$

$$\hat{N}(f)=-j\frac{1}{2}X(f-f_c)+j\frac{1}{2}X(f+f_c)+\frac{1}{2}Y(f-f_c)+\frac{1}{2}Y(f+f_c)$$

And the Inverse Fourier transform of $\hat{N}(f)$ is

$$\hat{n}(t)=x(t)\sin(2\pi f_c t)+y(t)\cos(2\pi f_c t)$$

The quadrature components x(t) and $y(t)$ can now be derived from equations

$x(t) = n(t)co2\pi f_c t + \hat{n}(t)\sin 2\pi f_c t$

and

$y(t) = n(t)\cos\ 2\pi f_c t - \hat{n}(t)\sin 2\pi f_c t$

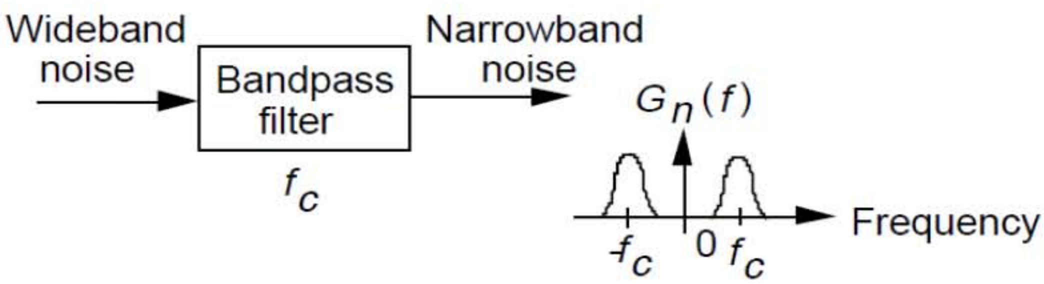


Fig: Generation of narrow band noise

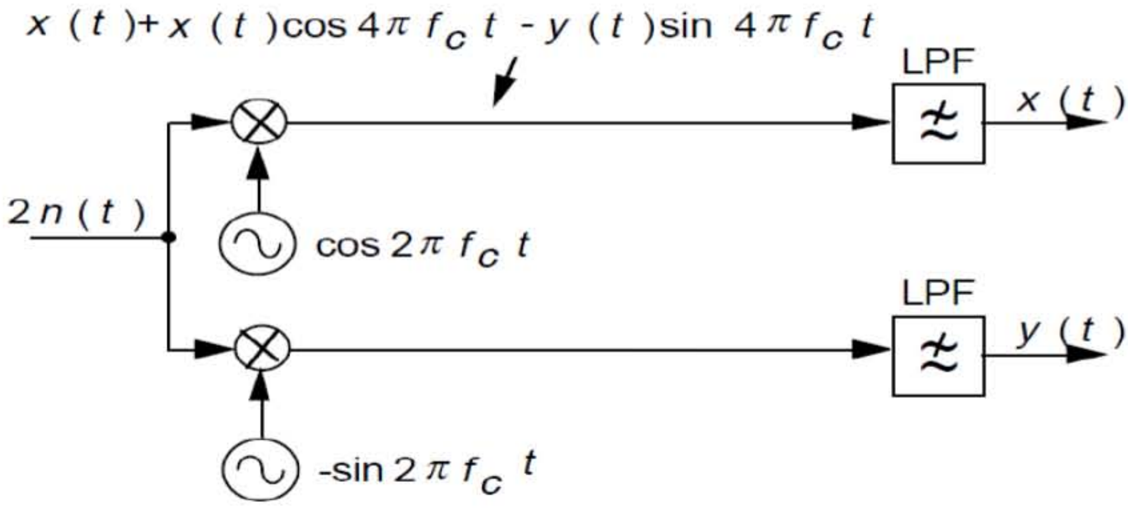


Fig: Generation of quadrature components of $n(t)$.

- Filters at the receiver have enough bandwidth to pass the desired signal but not too big to pass excess noise.
- Narrowband (NB) fc center frequency is much bigger that the bandwidth.
- Noise at the output of such filters is called narrowband noise (NBN).
- NBN has spectral concentrated about some mid-band frequency f_c

- The sample function of such NBN $n(t)$ appears as a sine wave of frequency f_C which modulates slowly in amplitude and phase

Input signal-to-noise ratio (SNR_I): is the ratio of the average power of modulated signal $s(t)$ to the average power of the filtered noise.

Output signal-to-noise ratio (SNR_O): is the ratio of the average power of demodulated message to the average power of the noise, both measured at the receiver output.

Channel signal-to-noise ratio (SNR_C): is the ratio of the average power of modulated signal $s(t)$ to the average power of the noise in the message bandwidth, both measured at the receiver input.

Noise figure

The Noise figure is the amount of noise power added by the electronic circuitry in the receiver to the thermal noise power from the input of the receiver. The thermal noise at the input to the receiver passes through to the demodulator. This noise is present in the receive channel and cannot be removed. The noise figure of circuits in the receiver such as amplifiers and mixers, adds additional noise to the receive channel. This raises the noise floor at the demodulator

$$\text{Noise figure} = \frac{signal\ to\ noise\ ratio\ at\ input}{signal\ to\ noise\ ratio\ at\ output}$$

Noise Bandwidth

A filter's equivalent noise bandwidth (ENBW) is defined as the bandwidth of a perfect rectangular filter that passes the same amount of power as the cumulative bandwidth of the channel selective filters in the receiver. At this point we would like to know the noise floor in our receiver, i.e. the noise power in the receiver intermediate frequency (IF) filter bandwidth that comes from kTB. Since the units of kTB are Watts/ Hz, calculate the noise floor in the channel bandwidth by multiplying the noise power in a 1 Hz bandwidth by the overall equivalent noise bandwidth in Hz.

NOISE IN DSB-SC SSYTEM:

Let the transmitted signal is

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

The received signal at the output of the receiver noise- limiting filter : Sum of this signal and filtered noise .A filtered noise process can be expressed in terms of its in-phase and quadrature components as

$$n(t) = A(t)\cos[2\pi f_c t + \theta(t)] = A(t)\cos\theta(t)\cos(2\pi f_c t) - A(t)\sin\theta(t)\sin(2\pi f_c t)$$
$$= n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t)$$

where $n_c(t)$ is in-phase component and $n_s(t)$ is quadrature component

Received signal (Adding the filtered noise to the modulated signal)

$$r(t)=u(t)+n(t)$$
$$=A_c m(t)\cos(2\pi f_c t)+ n_c(t)\cos(2\pi f_c t)- n_s(t)\sin(2\pi f_c t)$$

Demodulate the received signal by first multiplying r(t) by a locally generated sinusoid $\cos(2\pi f_c t + \emptyset)$, where □ is the phase of the sinusoid.Then passing the product signal through an ideal lowpass filter having a bandwidth W.

The multiplication of r(t) with $\cos(2\pi f_c t + \phi)$ yields

$$
\begin{aligned}
r(t)\cos(2\pi f_c t+\phi) \\
&= u(t)\cos(2\pi f_c t+\phi)+n(t)\cos(2\pi f_c t+\phi) \\
&= A_c m(t)\cos(2\pi f_c t)\cos(2\pi f_c t+\phi) \\
&\quad + n_c(t)\cos(2\pi f_c t)\cos(2\pi f_c t+\phi) - n_s(t)\sin(2\pi f_c t)\cos(2\pi f_c t+\phi) \\
&= \tfrac{1}{2}A_c m(t)\cos(\phi)+\tfrac{1}{2}A_c m(t)\cos(4\pi f_c t+\phi) \\
&\quad + \tfrac{1}{2}[n_c(t)\cos(\phi)+n_s(t)\sin(\phi)] + \tfrac{1}{2}[n_c(t)\cos(4\pi f_c t+\phi) - n_s(t)\sin(4\pi f_c t+\phi)]
\end{aligned}
$$

The low pass filter rejects the double frequency components and passes only the low pass components.

$$y(t)=\frac{1}{2} A_c m(t)\cos(\emptyset)+\frac{1}{2}[n_c(t)\cos(\emptyset)+ n_s(t)\sin(\emptyset)]$$

the effect of a phase difference between the received carrier and a locally generated carrier at the receiver is a drop equal to $cos^2(\emptyset)$ in the received signal power.

Phase-locked loop

The effect of a phase-locked loop is to generate phase of the received carrier at the receiver.

If a phase-locked loop is employed, then $\phi = 0$ and the demodulator is called a coherent or synchronous demodulator.

In our analysis in this section, we assume that we are employing a coherent demodulator.

With this assumption, we assume that $\phi = 0$

$$Y(t) = \frac{1}{2}[A_c m(t) + n_c(t)]$$

Therefore, at the receiver output, the message signal and the noise components are additive and we are able to define a meaningful SNR. The message signal power is given by

$$P_0 = \frac{A_C^2}{4} P_M$$

Power P_M is the content of the messagesignal

The noise power is given by

$$P_{n0} = \frac{1}{4} P_{nc} = \frac{1}{4} P_n$$

The power content of n(t) can be found by noting that it is the result of passing $n_W(t)$ through a filter with bandwidth B_c.Therefore, the power spectral density of n(t) is given by

$$S_n(f) = \begin{cases} \frac{N_0}{2} & |f - f_c| < W \\ 0 & otherwise \end{cases}$$

The noise power is

$$P_n = \int_{-\infty}^{\infty} S_n(f) df = \frac{N_0}{2} \times 4W = 2WN_0$$

Now we can find the output SNR as

$$\left(\frac{S}{N}\right)_0 = \frac{P_0}{P_{n_0}} = \frac{\frac{A_c^2}{4} P_M}{\frac{1}{4} 2WN_0} = \frac{A_c^2 P_M}{2WN_0}$$

In this case, the received signal power, given by

$$P_R = A_c^2 P_M / 2.$$

The output SNR for DSB-SC AM may be expressed as

$$\left(\frac{S}{N}\right)_{0_{DSB}} = \frac{P_R}{N_0 W}$$

which is identical to baseband SNR.

In DSB-SC AM, the output SNR is the same as the SNR for a baseband system. DSB-SC AM does not provide any SNR improvement over a simple baseband communication system.

NOISE IN SSB-SC SSYTEM:

Let SSB modulated signal is

$$u(t)=A_c m(t)\cos(2\pi f_c t)\pm\hat{m}(t)\sin(2\pi f_c t)$$

Input to the demodulator

$$\begin{aligned} r(t) &= A_c m(t)\cos(2\pi f_c t) \mp A_c\hat{m}(t)\sin(2\pi f_c t) + n(t) \\ &= A_c m(t)\cos(2\pi f_c t) \mp A_c\hat{m}(t)\sin(2\pi f_c t) + n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \\ &= [A_c m(t) + n_c(t)]\cos(2\pi f_c t) + [\mp A_c\hat{m}(t) - n_s(t)]\sin(2\pi f_c t) \end{aligned}$$

Assumption : Demodulation with an ideal phase reference.

Hence, the output of the lowpass filter is the in-phase component (with a coefficient of ½) of the preceding signal.

Parallel to our discussion of DSB, we have

$$P_o = \frac{A_c^2}{4}P_M$$

$$P_{n_0} = \frac{1}{4}P_{n_c} = \frac{1}{4}P_n$$

$$P_n = \int_{-\infty}^{\infty} S_n(f)\,df = \frac{N_0}{2}\times 2W = WN_0$$

$$\left(\frac{S}{N}\right)_0 = \frac{P_0}{P_{n_0}} = \frac{A_c^2 P_M}{WN_0}$$

$$P_R = P_U = A_c^2 P_M$$

$$\left(\frac{S}{N}\right)_{0_{SSB}} = \frac{P_R}{N_0 W} = \left(\frac{S}{N}\right)_b$$

The signal-to-noise ratio in an SSB system is equivalent to that of a DSB system.

Noise in Conventional AM

DSB AM signal : $u(t) = A_c[1 + am_n(t)]\cos(2\pi f_c t)$

Received signal at the input to the demodulator

$$\begin{aligned} r(t) &= A_c[1 + am_n(t)]\cos(2\pi f_c t) + n(t) \\ &= A_c[1 + am_n(t)]\cos(2\pi f_c t) + n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \\ &= [A_c[1 + am_n(t)] + n_c(t)]\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \end{aligned}$$

Where a is the modulation index

$m_n(t)$ is normalized so that its minimum value is -1

If a synchronous demodulator is employed, the situation is basically similar to the DSB case, except that we have $1 + am_n(t)$ instead of $m(t)$.

$$y(t)=\frac{1}{2}[A_c am_n(t)] + n_c(t)]$$

Received signal power

$$P_R = \frac{A_c^2}{2}\left[1 + a^2 P_{M_n}\right]$$

- Assumed that the message process is zero mean.

Now we can derive the output SNR as

$$\left(\frac{S}{N}\right)_{0_{AM}} = \frac{\frac{1}{4}A_c^2 a^2 P_{M_n}}{\frac{1}{4}P_{n_c}} = \frac{A_c^2 a^2 P_{M_n}}{2N_0 W} = \frac{a^2 P_{M_n}}{1+a^2 P_{M_n}} \frac{\frac{A_c^2}{2}\left[1+a^2 P_{M_n}\right]}{N_0 W}$$

$$= \frac{a^2 P_{M_n}}{1+a^2 P_{M_n}} \frac{P_R}{N_0 W} = \frac{a^2 P_{M_n}}{1+a^2 P_{M_n}} \left(\frac{S}{N}\right)_b = \eta \left(\frac{S}{N}\right)_b$$

- η denotes the modulation efficiency
- Since $a^2 P_{M_n} < 1 + a^2 P_{M_n}$, the SNR in conventional AM is always smaller than the SNR in a baseband system.

- In practical applications, the modulation index a is in the range of 0.8-0.9.
- Power content of the normalized message process depends on the message source.
- Speech signals : Large dynamic range, P_M is about 0.1.
- The overall loss in SNR, when compared to a baseband system, is a factor of 0.075 or equivalent to a loss of 11 dB.

The reason for this loss is that a large part of the transmitter power is used to send the carrier component of the modulated signal and not the desired signal. To analyze the envelope-detector performance in the presence of noise, we must use certain approximations.

This is a result of the nonlinear structure of an envelope detector, which makes an exact analysis difficult

In this case, the demodulator detects the envelope of the received signal and the noise process. The input to the envelope detector is

$$r(t)= [A_c[1 + am_n(t)] + n_c(t)]\ )\ \cos(2\pi f_c t)- n_s(t)\sin(2\pi f_c t)$$

Therefore, the envelope of $r(t)$ is given by

$$V_r(t) = \sqrt{[A_c[1+am_n(t)]+n_c(t)]^2 + n_s^2(t)}$$

Now we assume that the signal component in $r(t)$ is much stronger than the noise component. Then

$$P\left(n_c(t) << A_c[1+am_n(t)]\right) \approx 1$$

Therefore, we have a high probability that

$$V_r(t) \approx A_c[1+am_n(t)] + n_c(t)$$

After removing the DC component, we obtain

$$r(t) = A_c a m_n(t) + n_c(t)$$

which is basically the same as y*(t)* for the synchronous demodulation without the ½ coefficient.

This coefficient, of course, has no effect on the final SNR. So we conclude that, under the assumption of high SNR at the receiver input, the performance of synchronous and envelope demodulators is the same.

However, if the preceding assumption is not true, that is, if we assume that, at the receiver input, the noise power is much stronger than the signal power, Then

$$\begin{aligned} V_r(t) &= \sqrt{\left[A_c[1+am_n(t)] + n_c(t)\right]^2 + n_s^2(t)} \\ &= \sqrt{A_c^2[1+am_n(t)]^2 + n_c^2(t) + n_s^2(t) + 2A_c n_c(t)[1+am_n(t)]} \\ &\xrightarrow{a} \sqrt{\left(n_c^2(t)+n_s^2(t)\right)\left[1+\frac{2A_c n_c(t)}{n_c^2(t)+n_s^2(t)}\left(1+am_n(t)\right)\right]} \\ &\xrightarrow{b} V_n(t)\left[1+\frac{A_c n_c(t)}{V_n^2(t)}\left(1+am_n(t)\right)\right] \\ &= V_n(t) + \frac{A_c n_c(t)}{V_n(t)}\left(1+am_n(t)\right) \end{aligned}$$

(a) : $A_c^2[1+am_n(t)]^2$ is small compared with the other components

(b) : $\sqrt{n_c^2(t)+n_s^2(t)} = V_n(t)$;the envelope of the noise process

Use the approximation

$\sqrt{1+\varepsilon} \approx 1+\frac{\varepsilon}{2}$, *for small* ε, where $\varepsilon = \dfrac{2A_c n_c(t)}{n_c^2(t)+n_s^2(t)}\left(1+am_n(t)\right)$

Then $V_r(t) = V_n(t) + \dfrac{A_c n_c(t)}{V_n(t)}\left(1+am_n(t)\right)$

We observe that, at the demodulator output, the signal and the noise components are no longer additive. In fact, the signal component is multiplied by noise and is no longer distinguishable. In this case, no meaningful SNR can be defined. We say that this system is operating below the threshold.

The subject of threshold and its effect on the performance of a communication system will be covered in more detail when we discuss the noise performance in angle modulation.

Effect of threshold in angle modulation system:

FM THRESHOLD EFFECT FM threshold is usually defined as a Carrier-to-Noise ratio at which demodulated Signal-to-Noise ratio falls 1dB below the linear relationship . This is the effect produced in an FM receiver when noise limits the desired information signal. It occurs at about 10 dB, as earlier stated in 5 the introduction, which is at a point where the FM signal-to-Noise improvement is measured. Below the FM threshold point, the noise signal (whose amplitude and phase are randomly varying) may instantaneously have amplitude greater than that of the wanted signal. When this happens, the noise will produce a sudden change in the phase of the FM demodulator output. In an audio system, this sudden phase change makes a “click”. In video applications the term “click noise” is used to describe short horizontal black and white lines that appear randomly over a picture

An important aspect of analogue FM satellite systems is FM threshold effect. In FM systems where the signal level is well above noise received carrier-to-noise ratio and demodulated signal-to-noise ratio are related by:

$$S/N = 3\beta^2 C/N \qquad \text{Eqn 9}$$

where S/N = signal-to-noise ratio at output of FM demodulator
β = FM deviation ratio or modulation index = ($\Delta f/B$)
C/N = carrier-to-noise ratio at input of FM demodulator
Δf = peak deviation
B = basebandwidth of signal being modulated

The expression however does not apply when the carrier-to-noise ratio decreases below a certain point. Below this critical point the signal-to-noise ratio decreases significantly. This is known as the FM threshold effect (FM threshold is usually defined as the carrier-to-noise ratio at which the demodulated signal-to-noise ratio fall 1 dB below the linear relationship given in Eqn 9. It generally is considered to occur at about 10 dB).

Below the FM threshold point the noise signal (whose amplitude and phase are randomly varying), may instantaneously have an amplitude greater than that of the wanted signal. When this happens the noise will produce a sudden change in the phase of the FM demodulator output. In an audio system this sudden phase change makes a "click". In video applications the term "click noise" is used to describe short horizontal black and white lines that appear randomly over a picture, because satellite communications systems are power limited they usually operate with only a small design margin above the FM threshold point (perhaps a few dB). Because of this circuit designers have tried to devise techniques to delay the onset of the FM threshold effect. These devices are generally known as FM threshold extension demodulators. Techniques such as FM feedback, phase locked loops and frequency locked loops are used to achieve this effect. By such techniques the onset of FM threshold effects can be delayed till the C/N ratio is around 7 dB.

Noise in Angle Modulated Systems

Like AM, noise performance of angle modulated systems is characterized by parameter γ

$$\gamma_{FM} = \frac{3}{2}\beta^2$$

If it is compared with AM

$$\frac{\gamma_{FM}}{\gamma_{AM}} = \frac{1}{2}\left(\frac{\omega_{FM}}{\omega_{AM}}\right)^2$$

Note if bandwidth ratio is increased by a factor 2.Then

$\frac{\gamma_{FM}}{\gamma_{AM}}$ Increases by a factor 4

This exchange of bandwidth and noise performance is an important feature of FM

$$\text{Figure of merit } (\gamma) = \frac{SNR_O}{SNR_C}$$

CW- Modulation System	SNR_O	SNR_C	Figure of merit	Figure of merit (single tone)
DSB-SC	$\frac{C^2A_c^2P}{2WN_0}$	$\frac{C^2A_c^2P}{2WN_0}$	1	1
SSB	$\frac{C^2A_c^2P}{4WN_0}$	$\frac{C^2A_c^2P}{4WN_0}$	1	1
AM	$\frac{A_c^2k_a^2P}{2WN_0}$	$\frac{A_c^2(1+k_a^2P)}{2WN_0}$	$\approx \frac{k_a^2P}{1+k_a^2P} < 1$	$\frac{\mu^2}{2+\mu^2}$
FM	$\frac{3A_c^2k_f^2P}{2N_0W^3}$	$\frac{A_c^2}{2WN_0}$	$\frac{3k_f^2P}{W^2}$	$\frac{3}{2}\beta^2$

P is the average power of the message signal.

C^2 is a constant that ensures that the ration is dimensionless.

W is the message bandwidth.

A_c is the amplitude of the carrier signal.

k_a is the amplitude sensitivity of the modulator.

$\mu = k_a A_m$ and A_m is the amplitude sinusoidal wave

$\beta = \frac{\Delta f}{W}$ is the modulation index.

k_f is the frequency sensitivity of the modulator.

Δf is the frequency deviation.

UNIT-V

Receivers

Introduction:

In radio communications, a radio receiver is an electronic device that receives radio waves and converts the information carried by them to a usable form.The antenna intercepts radio waves (electromagnetic waves) and converts them to tiny alternating currents which are applied to the receiver and the receiver extracts the desired information. The receiver uses electronic filters to separate the desired radio frequency signal from all the other signals picked up by the antenna, an electronic amplifier to increase the power of the signal for further processing, and finally recovers the desired information through demodulation. The information produced by the receiver may be in the form of sound, moving images (television), or data. Radio receivers are very widely used in modern technology, as components of communications, broadcasting, remote control, and wireless networking systems.

Types of Receivers:

The component technology, and in particular semiconductor technology has surged forwards enabling much higher levels of performance to be achieved in a much smaller space.Many of the different radio receiver types have been around for many years. Some of them are

1. **Tuned radio frequency, TRF**:

This type of radio receiver was one of the first that was used. The very first radio receivers of this type simply consisted of a tuned circuit and a detector. Crystal sets were early forms of TRF radios. Later amplifiers were added to boost the signal level, both at the radio frequencies and audio frequencies. There were several problems with this form of receiver. The main one was the lack of selectivity. Gain / sensitivity were also a use.

2 Regenerative receiver:

The regenerative radio receiver significantly improved the levels of gain and selectivity obtainable. It used positive feedback and ran at the point just before oscillation occurred. In this way a significant multiplication in the level of "Q" of the tuned circuit was gained. Also major improvements in gain were obtained this way.

3. Super regenerative receiver:

The super regenerative radio receiver takes the concept of regeneration a stage further. Using a second lower frequency oscillation within the same stage, this second oscillation quenches or interrupts the oscillation of the main regeneration – typically at frequencies of around 25 kHz or so above the audio range. In this way the main regeneration can be run so that the stage is effectively in oscillation where it provides very much higher levels of gain. Using the second quench oscillation, the effects of running the stage in oscillation are not apparent to the listener, although it does emit spurious signals which can cause interference locally.

4.Super heterodyne receiver:

The super heterodyne form of radio receiver was developed to provide additional levels of selectivity. It uses the heterodyne or mixing process to convert signals done to a fixed intermediate frequency. Changing the frequency of the local oscillator effectively tunes the radio.

5. Direct conversion receiver:

This type of radio format converts the signal directly down to the baseband frequency. Initially it was used for AM, Morse (CW) and SSB transmissions. Now it is widely used for digital communications where IQ demodulators are used to take advantage of the variety of phase shift keying, PSK, and quadrature amplitude modulation, QAM signals.

Many of these different types of radio receiver are in widespread use today. Each type of radio has its own characteristics that lend its use to particular applications.

Tuned Radio Frequency Receiver(TRF)

The tuned radio frequency receiver is one in which the tuning or selectivity is provided at the radio frequency stages.Tuning is provided by a tuned coil / capacitor combination, and then the signal is presented to a simple crystal or diode detector where the amplitude modulated signal is recovered. This is then passed straight to the headphones.

The tuned radio frequency receiver was used in the early days of wireless technology but it is rarely used today as other techniques offering much better performance are available.

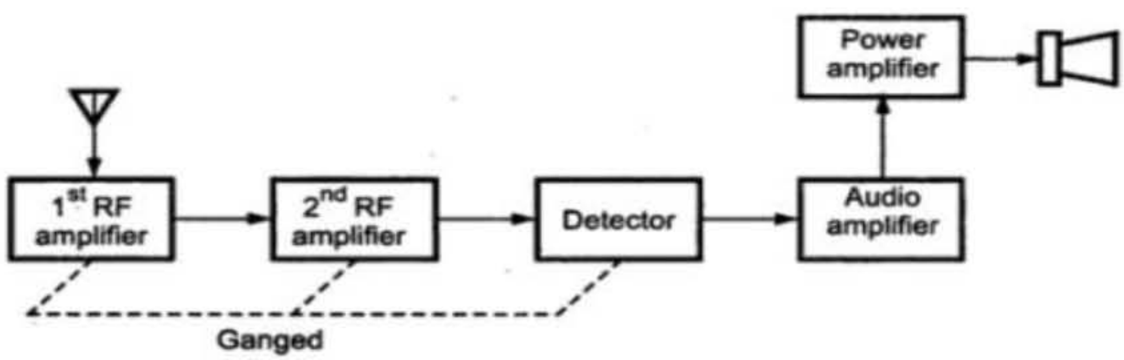


Fig.Block diagram of Tuned radio frequency receiver

Operation:

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

The amplified signal is then demodulated using detector to recover the modulating signal. The recovered signal is amplified further by the audio amplifier followed by power amplifier which provides sufficient gain to operate a loud speaker. The TRF receivers suffered from number of annoying problems.

The tuned radio frequency receiver was popular in the 1920s as it provided sufficient gain and selectivity for the receiving the broadcast stations of the day. However tuning is difficult in which as each stage in the early radios needed to be adjusted separately. The TRF receiver has largely been disregarded in recent years. Other receiver topologies offer far better levels of performance, and with integrated circuit technology, the additional circuitry of other types of receiver is not an issue.Later ganged tuning capacitors were introduced, but by this time the superheterodyne receiver was becoming more widespread.

Disadvantages of TRF receiver

• Poor selectivity and low sensitivity in proportion to the number of tuned amplifiers used.
• Selectivity requires narrow bandwidth, and narrow bandwidth at a high radio frequency implies high Q or many filter sections.
• An additional problem for the TRF receiver is tuning different frequencies. All the tuned circuits need to tune together to the same frequency or track very closely. Another problem is to keep the narrow bandwidth tuning. Keeping several tuned circuits aligned is difficult.
• The bandwidth of a tuned circuit doesn’t remain constant and increases with the frequency increase.
• The need to have all RF stages track one another
• Instability due to large number of RF stages.
• Received bandwidth increases with frequency (varies with center frequency)
• Gain is non-uniform over a wide range of frequencies

Superheterodyne Receiver:

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

The basic concept and theory behind the superheterodyne radio involves the process of mixing. This enables signals to be translated from one frequency to another. The input frequency is often referred to as the RF input, whilst the locally generated oscillator signal is referred to as the local oscillator, and the output frequency is called the intermediate frequency as it is between the RF and audiofrequencies.

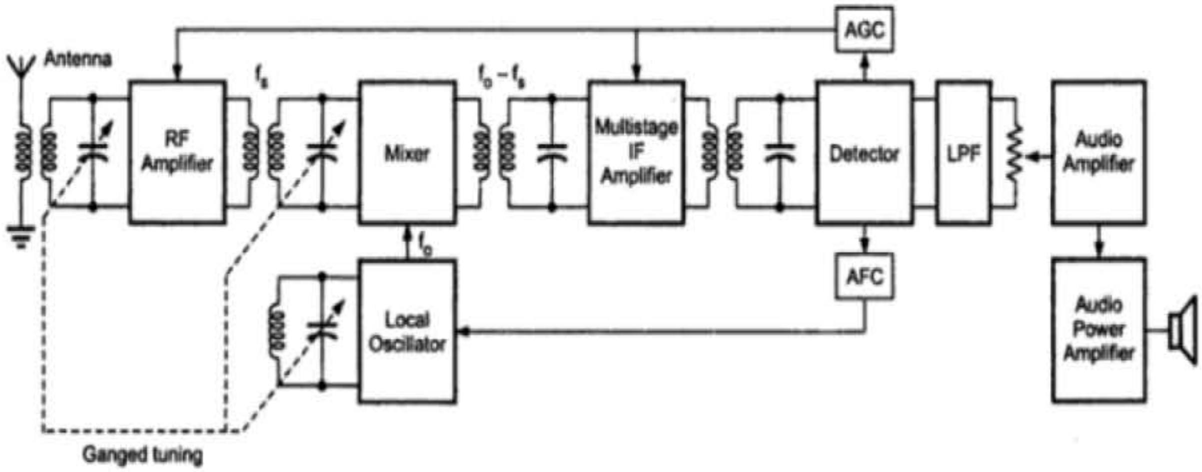


Fig .Block diagram of a Superheterodyne receiver

Operation:

Signals enter the receiver from the antenna and are applied to the RF amplifier where they are tuned to remove the image signal and also reduce the general level of unwanted signals on other frequencies that are not required.

The signals are then applied to the mixer along with the local oscillator where the wanted signal is converted down to the intermediate frequency. Here significant levels of amplification are applied and the signals are filtered. This filtering selects signals on one channel against those on the next. It is much larger than that employed in the front end.The advantage of the IF filter as opposed to RF filtering is that the filter can be designed for a fixed frequency. This allows for much better tuning. Variable filters are never able to provide the same level of selectivity that can be provided by fixed frequency ones.

Once filtered the next block in the superheterodyne receiver is the demodulator. This could be for amplitude modulation, single sideband, frequency modulation, or indeed any form of modulation. It is also possible to switch different demodulators in according to the mode being received.

The final element in the superheterodyne receiver block diagram is shown as an audio amplifier, although this could be any form of circuit block that is used to process or amplified the demodulated signal.

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

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

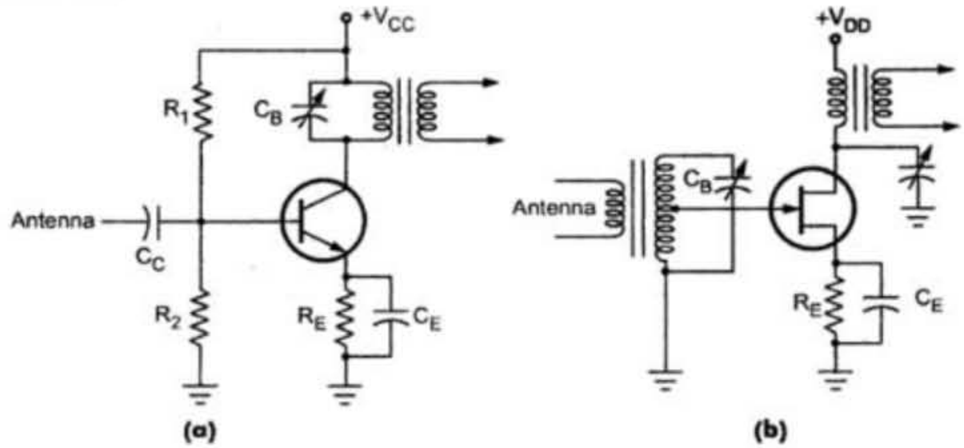
Advantages of the superheterodyne receiver

• IF stage permits use at very high frequencies.
• Because many components operate at the fixed IF, they can be optimized.
• Less expensive.
• Better selectivity
• Improved circuit stability.
• Uniform gain over a wide range of frequencies

Receiver Sections:

RF tuning & amplification:

RF amplifier provides initial gain and selectivity. RF amplifier is a simple class A circuit. This RF stage within the overall block diagram for the receiver provides initial tuning to remove the image signal. If noise performance for the receiver is important, then this stage will be designed for optimum noise performance. This RF amplifier circuit block will also increase the signal level so that the noise introduced by later stages is at a lower level in comparison to the wanted signal. A typical bipolar circuit as (a) and FET circuit as (b) is shown below.



Values of resistors R1 and R2 in the bipolar circuit are adjusted such that the amplifier works as a class A amplifier.The antenna is connected through coupling capacitor to the base of the transistor.This makes the circuit very broad band as the transistor will amplify virtually any signal picked up by the antenna. The collector is tuned with a parallel resonant circuit to provide the initial selectivity for the mixer input.

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

Local oscillator:

The local oscillator circuit block can take a variety of forms. Early receivers used free running local oscillators. Today most receivers use frequency synthesizers, normally based around phase locked loops. These provide much greater levels of stability and enable frequencies to be programmed in a variety of ways.

Mixer or Frequency Changer or Converter:

Real-life mixers produce a variety of other undesired outputs, including noise and they may also suffer overload when very strong signals are present.
Although very basic non-linear devices can actually perform a basic RF mixing or multiplication process, the performance will be far from the ideal, and where good receiver performance is required, the specification of the RF mixer must be match this expectation.

The basic process of RF mixing or multiplication where the incoming RF signal and a local oscillator are mixed or multiplied together to produce signals at the sum and difference frequencies is key to the whole operation of the superheterodyne receiver.

There are a number of considerations when looking at the receiver design and topology with respect to the RF mixer.There are many different forms of mixer that can be used, and the choice of the type depends very much upon the receiver and the anticipated performance.

Separately Excited Mixer

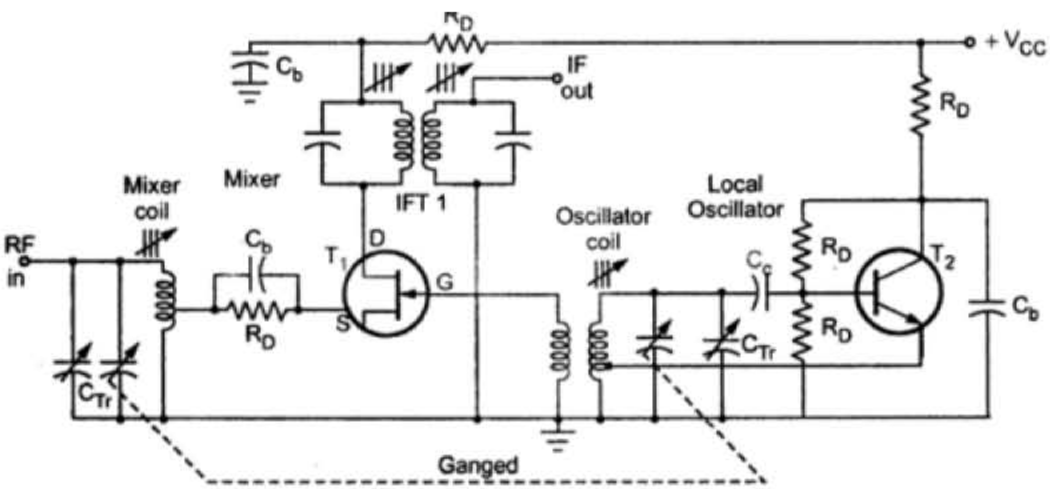


Fig. Separately Excited Mixer

In Separately excited mixer, one device acts as a mixer while the other supplies the necessary oscillations. Bipolar transistor T2 forms the Hartley oscillator circuit and oscillates with local frequency. FET T1 is a mixer whose gate s fed with the output of local oscillator and its bias is adjusted. The local oscillator varies the gate bias of the FET to vary its transconductance resulting intermediate frequency at the output. Output is taken through double tuned transformer in the drain of the mixer and fed to the IF amplifier.

Self Excited Mixer

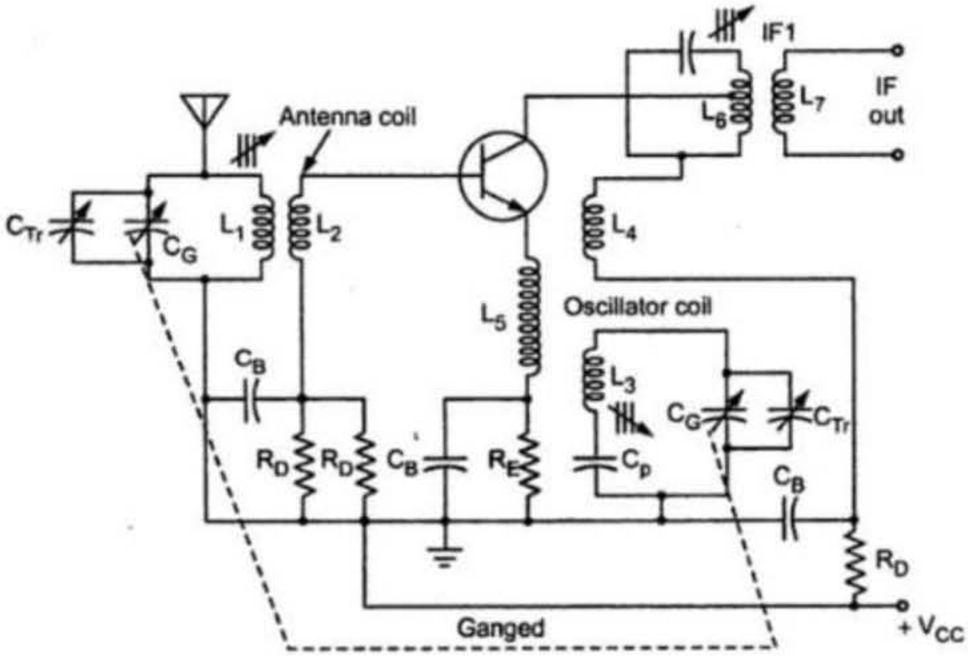


Fig: Self Excited Mixer

Self Excited circuit oscillates and the transconductance of the transistor is varied in a nonlinear manner at the local oscillator gate. This variable transconductance is used by the transistor to amplify the incoming RF signal.

IF amplifier & filter:

Tracking

The Superheterodyne receiver has number of tunable circuits which must all be tuned correctly if any given station is to be received. The ganged tuning is employed which mechanically couples all tuning circuits so that only one tuning control is required. Usually there are three tuned circuits: Antenna or RF tuned circuit, Mixer tuned circuit and local oscillator tuned circuit.

All these circuits must be tuned to get proper RF input and to get IF frequency at the output of the mixer. The process of tuning circuit to get the desired output is called Tracking. Tracking error will result in incorrect frequency being fed to the IF amplifier and these must be avoided.

To avoid tracking errors,ganged capacitors with identical sections are used.

A different value of inductance and capacitors called trimmers and padders are used to adjust the capacitance of the oscillator to the proper range. Common methods used for tracking are

- Padder Tracking
- Trimmer Tracking
- Three-Point Tracking

Intermediate IF amplifier:

Figure shows two Stage IF Amplifier. Two stages are transformer coupled and all IF transformers are single tuned i.e, tuned for single frequency.

IF amplifiers are tuned voltage amplifiers which istuned for the fixed frequency. Its function is to amplify only tuned frequency signal and reject all others .Most of the receiver gain is provided by the IF amplifiers and the required gain is obtained usually by two or more stages of IF amplifiers are required.

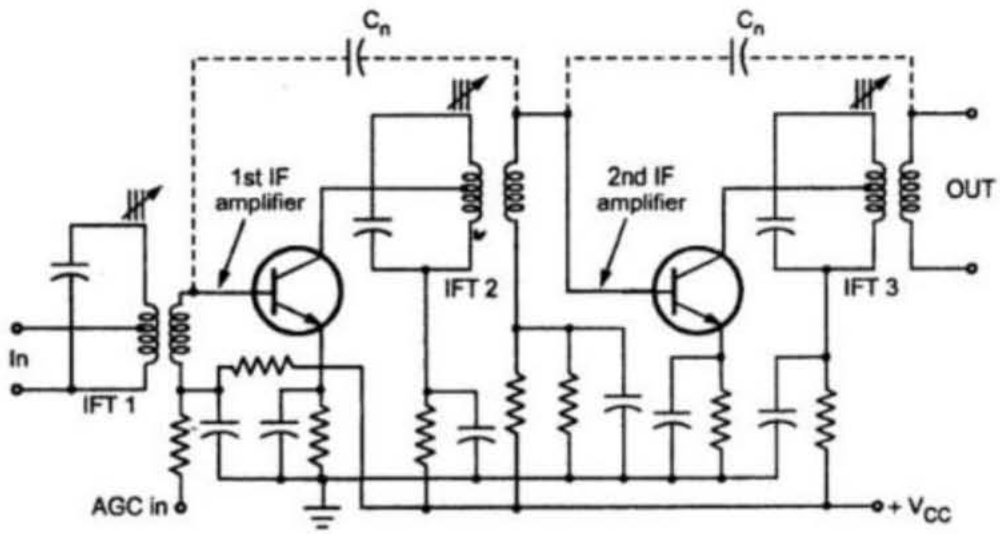


Fig. Two Stage IF Amplifier

Automatic Gain Control,AGC:

An automatic gain control is incorporated into most superheterodyne radios. Its function is to reduce the gain for strong signals so that the audio level is maintained for amplitude sensitive forms of modulation, and also to prevent overloading. It is a system in which the overall gain of a radio receiver is varied automatically with the variations in the strength of the receiver signal to maintain the output substantially constant.

When the average signal level increases, the size of the AGC bias increases and the gain is deceased. When there is no signal, there is a minimum AGC bias and the amplifiers produce maximum gain. There are two types of AGC circuits. They are Simple AGC and Delayed AGC.

Simple AGC

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

In the circuit, the dc bias produced by half wave rectifier is used to control the gain of RF or IF amplifier. The time constant of the filter is kept at least 10 times longer than the period of the lowest modulation frequency received. If the time constant is kept longer, it will give better filtering. The recovered signal is then passes through capacitor to remove dc. The resulting ac signal is further amplified and applied to the loud speaker.

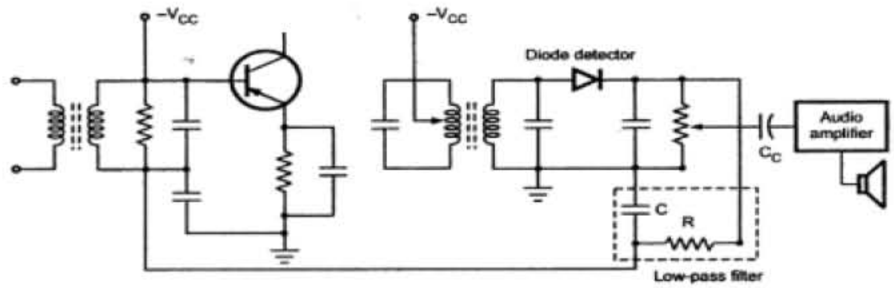


Fig. Simple AGC circuit

Delayed AGC

In simple AGC, the unwanted weak signals (noise signals) are amplified with high gain. To avoid this, in delayed AGC circuits, AGC bias is not applied to amplifiers until signal strength has reached a predetermined level, after which AGC bias is applied as with simple AGC, but more strongly.

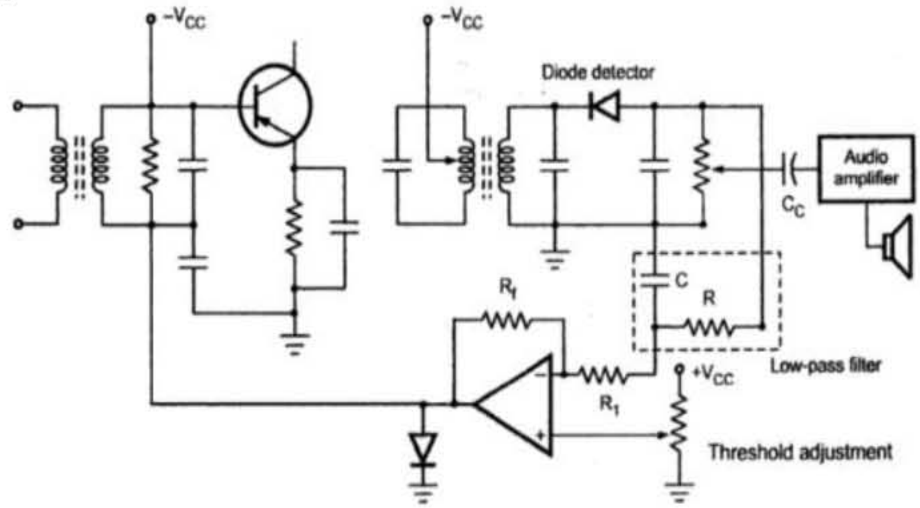


Fig. Delayed AGC circuit

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

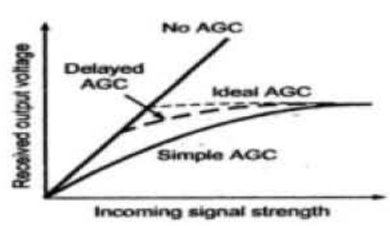


Fig. Response of receiver with various AGC

Demodulator:

The superheterodyne receiver block diagram only shows one demodulator, but in reality radios may have one or more demodulators dependent upon the type of signals being receiver.

Audio amplifier:

Once demodulated, the recovered audio is applied to an audio amplifier block to be amplified to the required level for loudspeakers or headphones. Alternatively the recovered modulation may be used for other applications whereupon it is processed in the required way by a specific circuit block.

Receiver Characteristics

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

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

Selectivity

The ability of the receiver to select the wanted signals among the various incoming signals is termed as Selectivity. It rejects the other signals at closely lying frequencies. Selectivity of a receiver changes with incoming signal frequency and are poorer at high frequencies.

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

Sensitivity

The sensitivity of a radio receiver is its ability to amplify weak signals. It is often defined in terms of the voltage that must be applied to the receiver input terminals to give a standard output power, measured at the output terminals.The most important factors determining the sensitivity of a superheterodyne receiver are the gain of the IF amplifier(s) and that of the RF amplifier .The more gain that a receiver has, the smaller the input signal necessary to produce the desired output power. Therefore sensitivity is a primary function of the overall receiver gain. Good communication receiver has a sensitivity of 0.2 to 1 μV

Fidelity

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

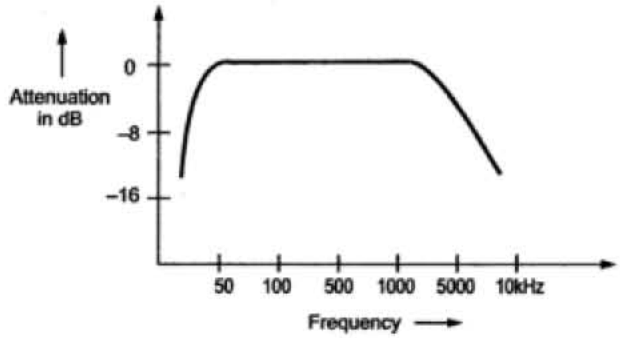


Fig.3. Typical Fidelity curve

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

Image frequency and its Rejection

In a standard broadcast receiver the local oscillator frequency is made higher than the incoming signal frequency for reasons that will become apparent .It is made equal at all times to the signal frequency plus the intermediate frequency.Thus $f0=fs+fi$ or $f0=fs-fi$, no matter what the signal frequency may be. When f0 and fs are mixed, the difference frequency, which is one of the by-products, equal to fi is passed and amplified by the IF stage.If a frequency fsi manages to reach the mixer, such that $fsi=fo+fi$, that is , $fsi=fs+2fi$ then this frequency will also produce fi when mixed with f0.

Unfortunately, this spurious intermediate-frequency signal will also be amplified by the IF stage and will therefore provide interference.This has the effect of two stations being received simultaneously and is naturally undesirable.The term fsi is called image frequency and is defined as the signal frequency plus twice the intermediate frequency.

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

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

where

Q=loaded Qoftuned circuit

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

Image rejection depends on the front-end selectivity of the receiver and must be achieved before the IF stage.Once the spurious frequency enters the first IF amplifier, it becomes impossible to remove it from the wanted signal.

Double spotting

This is well-known phenomenon, which manifests itself by the picking up of the same shortwave station at two nearby points on the receiver dial.It is caused by poor front-end selectivity, i.e., inadequate image-frequency rejection .That is to say, the front end of the receiver does not select different adjacent signals very well , but the IF stage takes care of eliminating almost all of them .

As a matter of interest, double spotting may be used to calculate the intermediate frequency of an unknown receiver, since the spurious point on the dial is precisely 2f, below the correct frequency .An improvement in image-frequency rejection will produce a corresponding reduction in double spotting.

FM Receiver

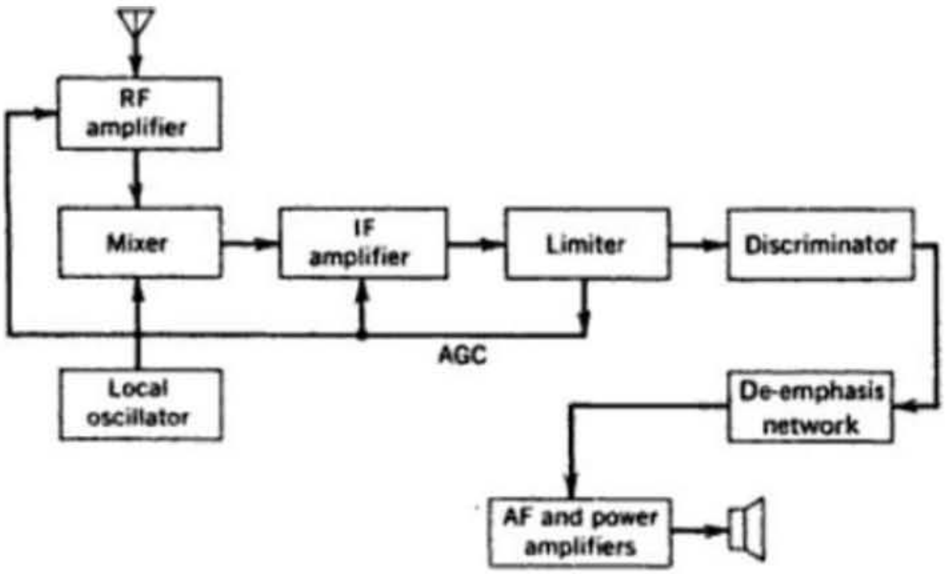


Fig. Block diagram of FM receiver

FM receiver is a Superheterodyne receiver and the basic block diagram is same as AM receiver. The basic differences are as follows.

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

Comparison with AM Receiver

Different stages of FM receiver are explained below.

RF Amplifier

Since FM signal has a larger bandwidth it is likely to encounter more noise. Hence to reduce the noise figure of the receiver, an RF amplifier is used. The RF amplifier stage matches the antenna to the receiver. For this purpose and to avoid neutralization, grounded-base or grounded gate circuits are employed for this stage. Both circuits have low input impedance, suitable for matching with antenna impedance and nether require neutralization.

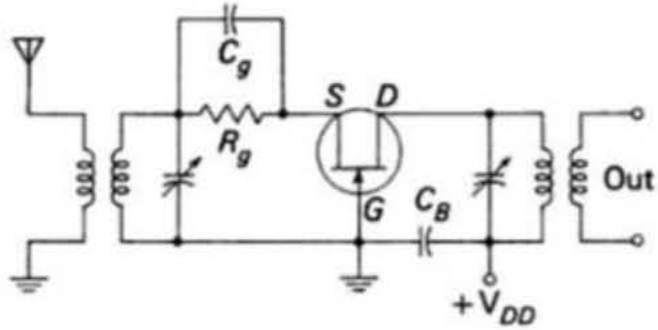


Fig. Grounded-gate FET RF amplifier

In figure, since gate terminal is grounded, the input and output sides are isolated for RF purposes. There is no feedback and hence no instability. Therefore circuit does not require neutralization. The low impedance of the FET amplifier is matched to antenna through a tuned RF transformer. Both the input and output tank circuits are tuned to carrier frequency.

Oscillators and Mixers

The oscillator circuit may be Clapp and Colpitts which is suited in VHF operation .Tracking is not normally much of a problem in FM broadcast receivers. This is because the tuning frequency is only 1.25:1.much less than in AM broadcasting.

The mixer stage uses a tuned circuit as its load. The circuit is tuned to Intermediate frequency of 10.7 MHz and hence selects the difference between incoming carrier frequency and locally generated oscillator frequency.

Intermediate Frequency and IF amplifier

The types and operation do not differ much from their AM counterparts. But the intermediate frequency and bandwidth required are far higher than in AM broadcast receivers. For receivers operating in the 88 to 108 MHz band is an IF of 10.7 MHz and a bandwidth of 200 KHz. Due to large bandwidth, gain per stage may be low. Two IF amplifier stages are often provided, in which case the shrinkage of bandwidth as stages are cascaded must be taken into account.

Amplitude Limiter:

Frequency modulation is developed to provide a communication system which is less noise sensitive than AM system.The most of the nose accompanying the desired signal accompanies it as AM. Thus the intelligence is contained in the frequency variation of signal. So at the receiver remove all the amplitude variations without loss of the information content of the desired signal. To remove the amplitude variations of the signal is the main function of the limiter. At the output of the limiter stage, a constant amplitude signal is obtained even though the amplitude of input signal may be varying.

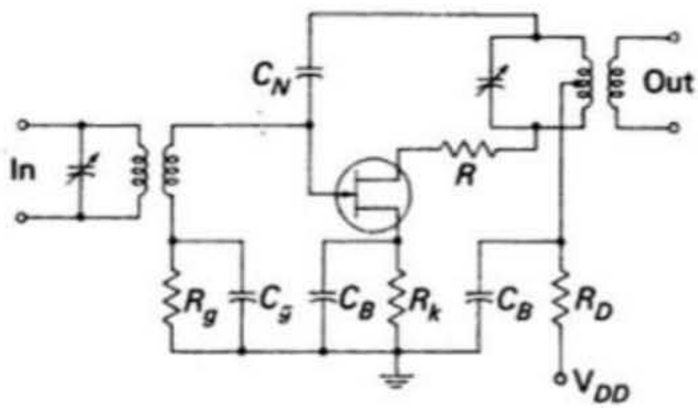


Fig. Amplitude Limiter

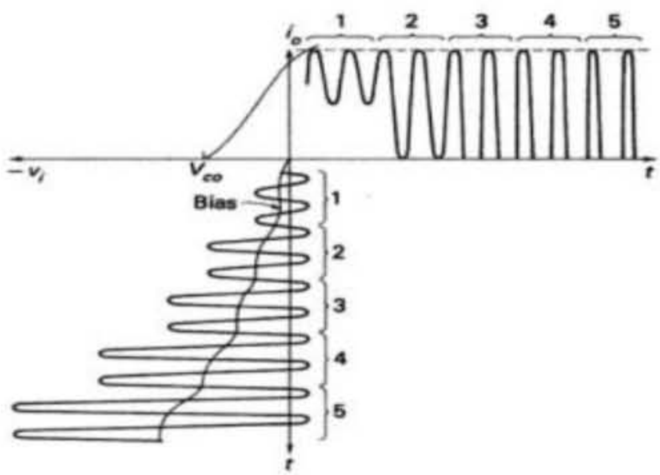


Fig. Amplitude limiter transfer characteristics

Limiter is basically a clipper circuit which clips off the undesired amplitude variations of the input signal. The input signal provides the bias for the FET circuit. Negative bias increases as input increases and hence it lowers the gain of the amplifier for high amplitude of the input signal and output voltage remains constant.

The basic function of the limiter is flat topping(Squaring off) the upper and lower extremities of the signal.

Although the signal is distorted, it makes no difference as far as FM is considered, since the information is contained in frequency variation and not in amplitude variation.

Use of AGC and Double Limiting

Sometimes it is quite practicable that average input signal amplitude may lie outside the limiting range. As a result, further limiting becomes necessary. The solution for this is the use of double limiter consisting two amplitude limiters in cascade. This gives satisfactory limiting range.

An alternative to the used of second limiter is automatic gain control. The AGC ensures, by reducing the gain for higher signal strengths, that the signal applied to the limiter is within the limiting range of the limiter. This also prevents the overloading of the last IF amplifier stage.

PULSE MODULATION

Introduction:

Pulse Modulation

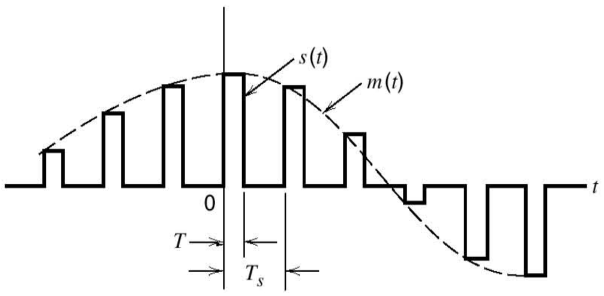
- Carrier is a train of pulses
- Example: Pulse Amplitude Modulation (PAM), Pulse width modulation (PWM) , Pulse Position Modulation (PPM)

Types of Pulse Modulation:

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

Pulse Amplitude Modulation:

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



PAM Generation:

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

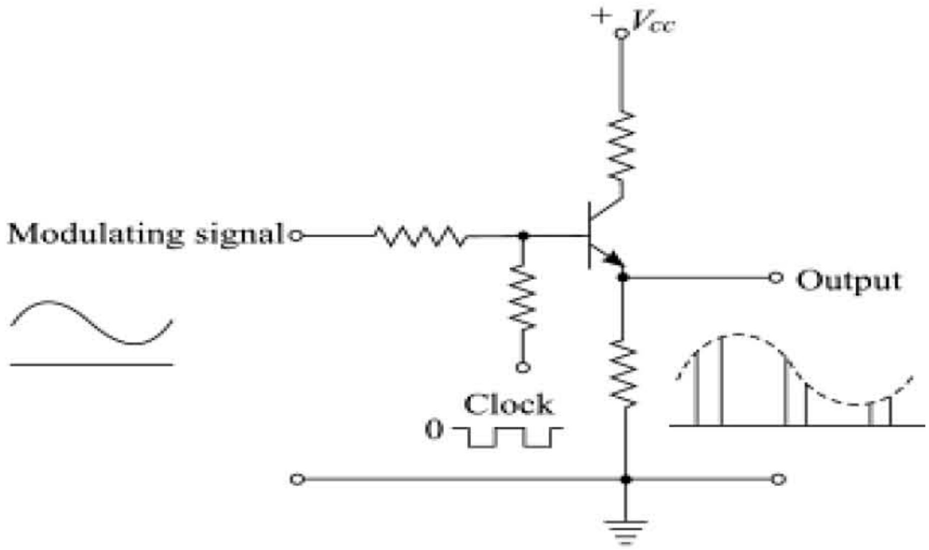


Fig.12. PAM Modulator

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

PAM Demodulator:

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

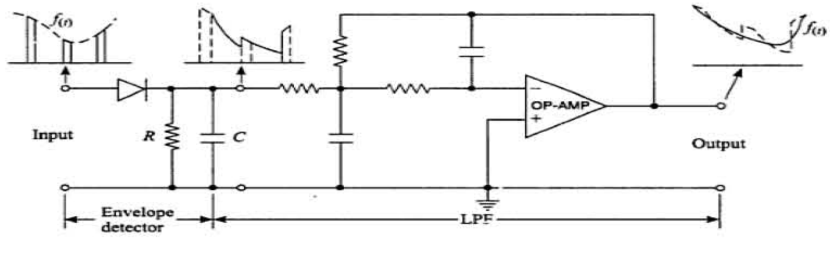
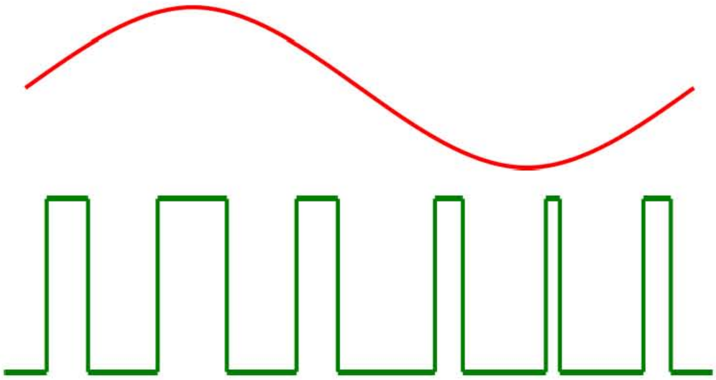


Fig.13. PAM Demodulator

Pulse Width Modulation:

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



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

Pulse Position Modulation:

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

PPM & PWM Modulator:

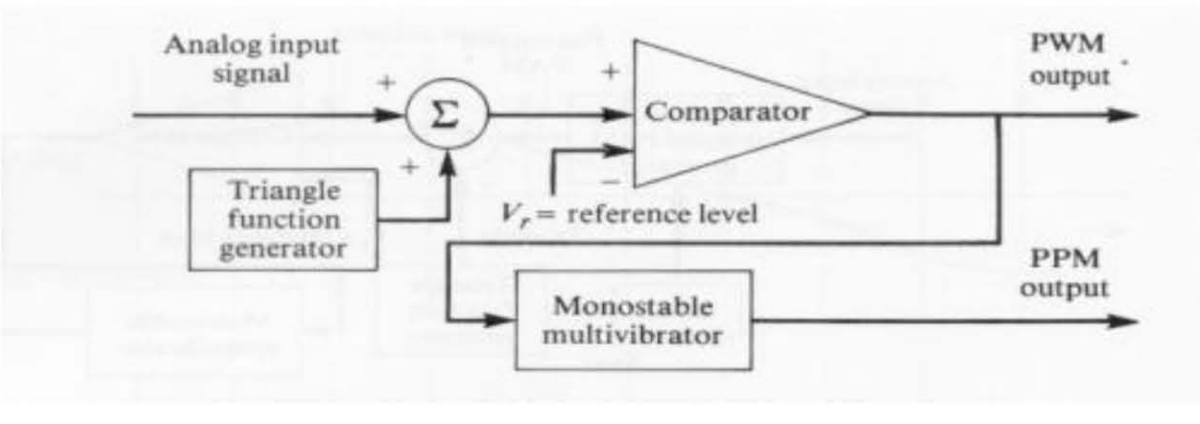


Fig.14. PWM & PPM Modulator

- The PPM signal can be generated from PWM signal.
- The PWM pulses obtained at the comparator output are applied to a mono stable multi vibrator which is negative edge triggered.

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

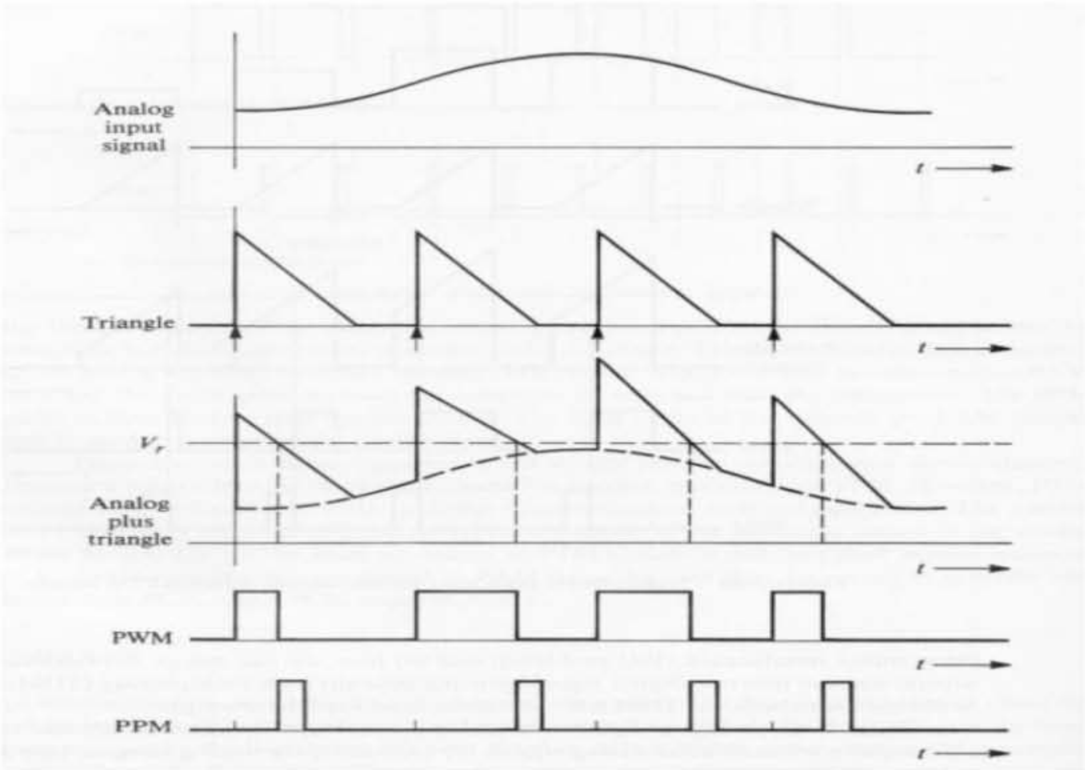


Fig.15. PWM & PPM Modulation waveforms

PWM Demodulator:

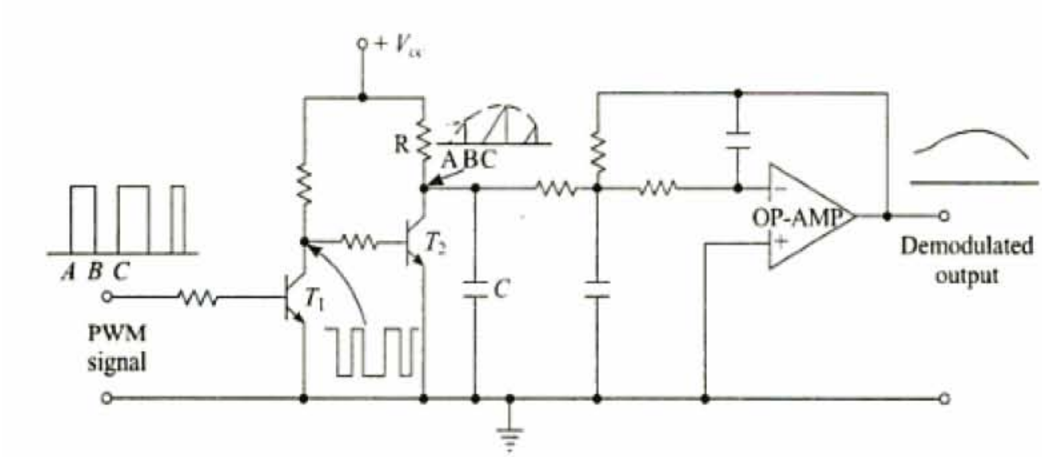


Fig.16. PWM Demodulator

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

PPM Demodulator:

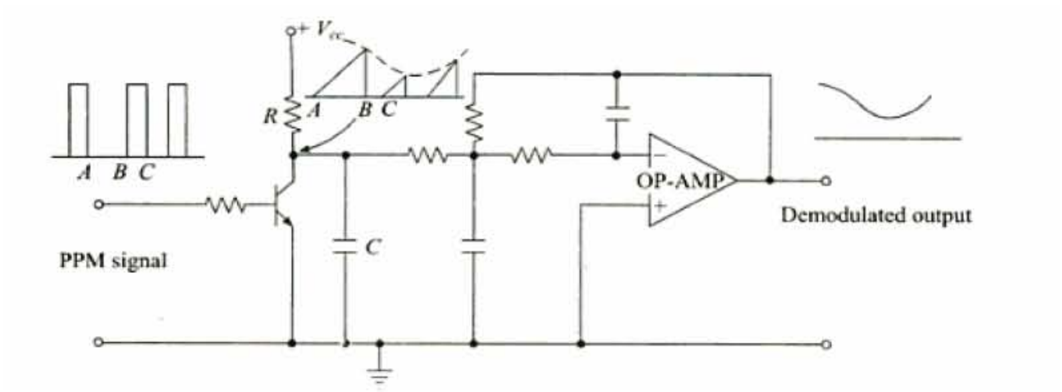


Fig.17. PPM Demodulator

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

UNIT-1

Digital Pulse Modulation

Elements of Digital Communication Systems:

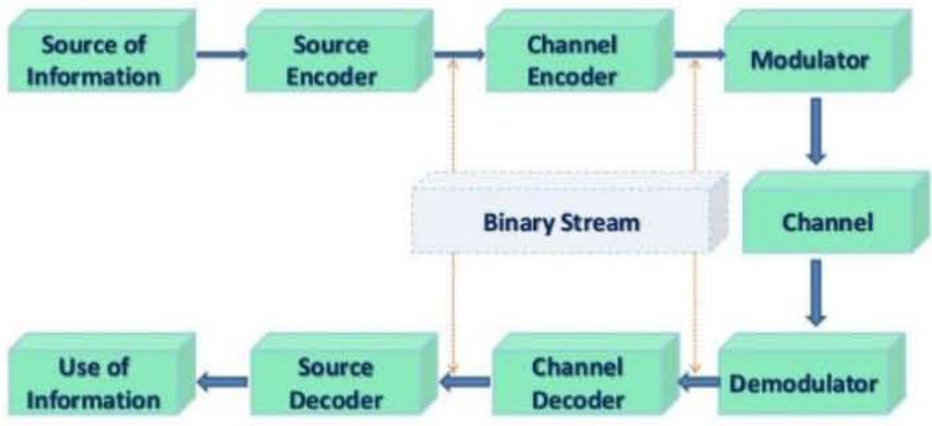


Fig. 1 Elements of Digital Communication Systems

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

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

2. **Source Encoder:**

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

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

3. **Channel Encoder:**

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

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

4. **Digital Modulator:**

   The binary sequence is passed to digital modulator which in turns convert the sequence into electric signals so that we can transmit them on channel (we will see channel later). The digital modulator maps the binary sequences into signal wave forms , for example if we represent 1 by sin x and 0 by cos x then we will transmit sin x for 1 and cos x for 0. (a case similar to BPSK)

5. **Channel:**

   The communication channel is the physical medium that is used for transmitting signals from transmitter to receiver. In wireless system, this channel consists of atmosphere , for traditional telephony, this channel is wired , there are optical channels, under water acoustic channels etc.We further discriminate this channels on the basis of their property and characteristics, like AWGN channel etc.

6. **Digital Demodulator:**

   The digital demodulator processes the channel corrupted transmitted waveform and reduces the waveform to the sequence of numbers that represents estimates of the transmitted data symbols.

7. **Channel Decoder:**

   This sequence of numbers then passed through the channel decoder which attempts to reconstruct the original information sequence from the knowledge of the code used by the channel encoder and the redundancy contained in the received data

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

8. **Source Decoder:**

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

9. **Output Transducer:**

Finally we get the desired signal in desired format analog or digital.

Advantages of digital communication:

- **Can withstand channel noise and distortion much better as long as the noise and the distortion are within limits.**
- **Regenerative repeaters prevent accumulation of noise along the path.**
- **Digital hardware implementation is flexible.**
- **Digital signals can be coded to yield extremely low error rates, high fidelity and well as privacy.**
- **Digital communication is inherently more efficient than analog in realizing the exchange of SNR for bandwidth.**
- **It is easier and more efficient to multiplex several digital signals.**
- **Digital signal storage is relatively easy and inexpensive.**
- **Reproduction with digital messages is extremely reliable without deterioration.**
- **The cost of digital hardware continues to halve every two or three years, while performance or capacity doubles over the same time period.**

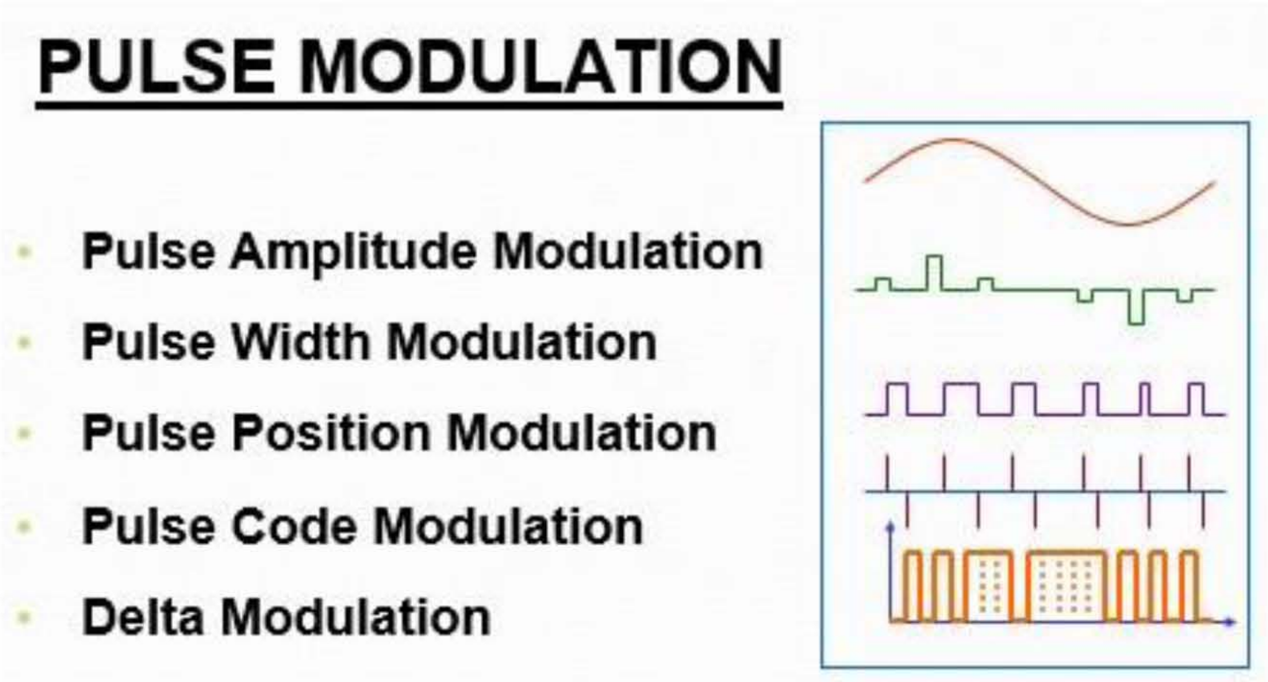
Disadvantages

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

Introduction to Pulse Modulation

What is the need for Pulse Modulation?

- Many Signals in Modern Communication Systems are digital
- Also, analog signals are transmitted digitally.
- Reduced distortion and improvement in signal to noise ratios.
- PAM, PWM, PPM, PCM and DM.
- In CW modulation schemes some parameter of modulated wave varies continuously with message.
- In Analog pulse modulation some parameter of each pulse is modulated by a particular sample value of the message.
- Pulse modulation is of two types
  - Analog Pulse Modulation
    - Pulse Amplitude Modulation (PAM)
    - Pulse width Modulation (PWM)
    - Pulse Position Modulation (PPM)
  - Digital Pulse Modulation
    - Pulse code Modulation (PCM)
    - Delta Modulation (DM)



Pulse Code Modulation:

Three steps involved in conversion of analog signal to digital signal

- Sampling
- Quantization
- Binary encoding

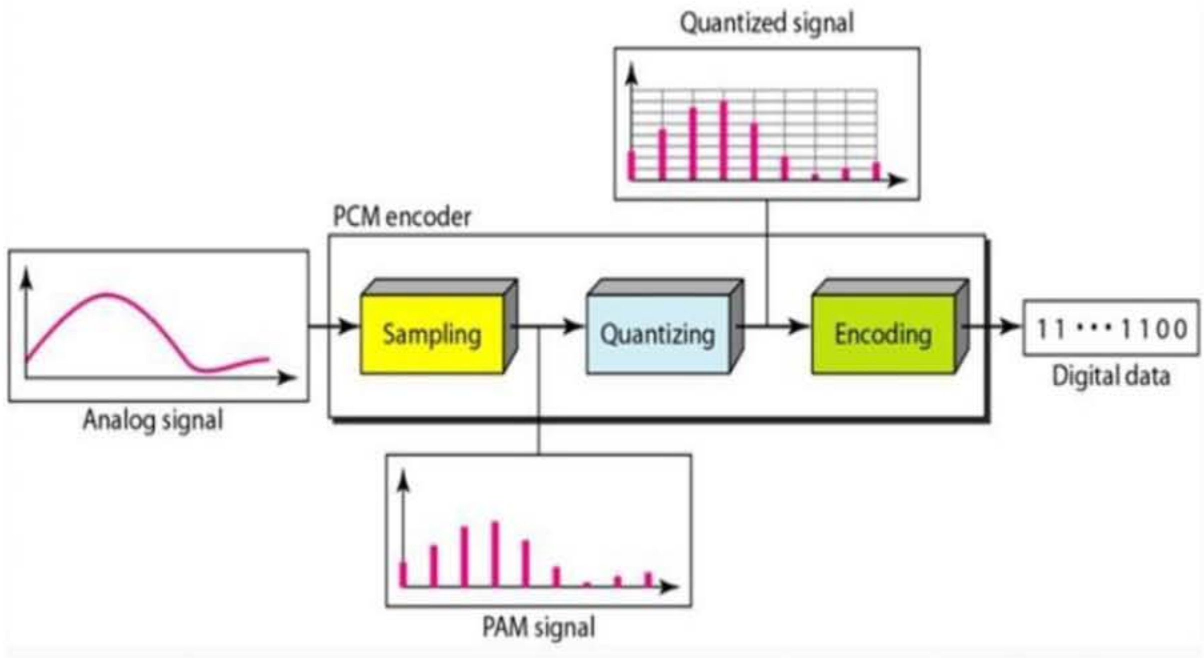


Fig. 2 Conversion of Analog Signal to Digital Signal

Note: Before sampling the signal is filtered to limit bandwidth.

<u>Elements of PCM System</u>:

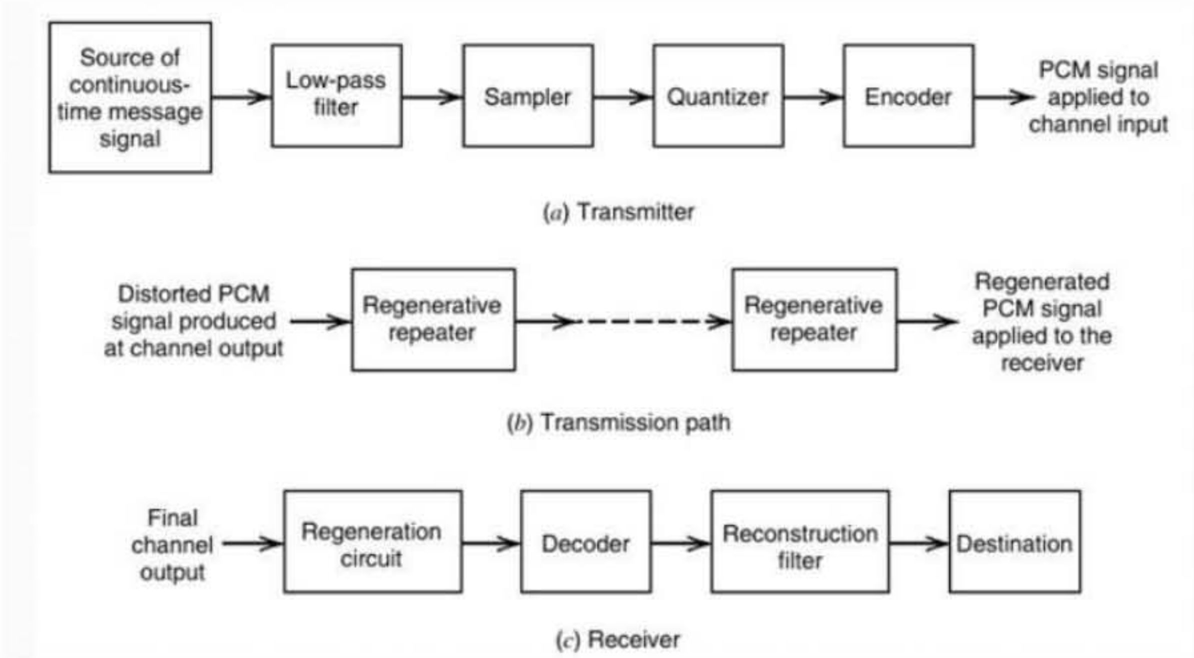


Fig. 3 Elements of PCM System

<u>Sampling:</u>

- Process of converting analog signal into discrete signal.
- Sampling is common in all pulse modulation techniques

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

Sampling methods:

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

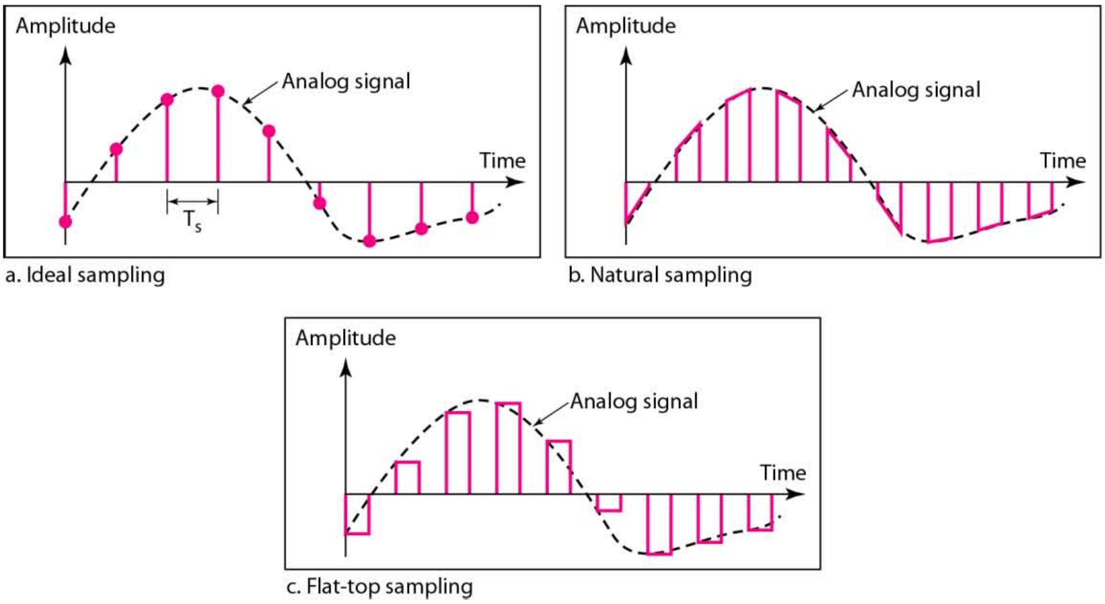


Fig. 4 Types of Sampling

Sampling of band-pass Signals:

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

  Min. sampling rate =2×$Bandwidth$

  =2×2f_m=4f_m

- Range of minimum sampling frequencies is in the range of 2×BW to 4×BW

Instantaneous Sampling or Impulse Sampling:

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

Natural sampling:

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

Flat top sampling:

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

Sampling Theorem:

Statement of sampling theorem

> 1) ***A band limited signal of finite energy, which has no frequency components higher than W Hertz, is completely described by specifying the values of the signal at instants of time separated by*** $\frac{1}{2W}$ ***seconds and***
>
> 2) ***A band limited signal of finite energy, which has no frequency components higher than W Hertz, may be completely recovered from the knowledge of its samples taken at the rate of 2W samples per second.***

The first part of above statement tells about sampling of the signal and second part tells about reconstruction of the signal. Above statement can be combined and stated alternately as follows :

A continuous time signal can be completely represented in its samples and recovered back if the sampling frequency is twice of the highest frequency content of the signal. i.e.,

$$f_s \geq 2W$$

Here f_s is the sampling frequency and

W is the higher frequency content

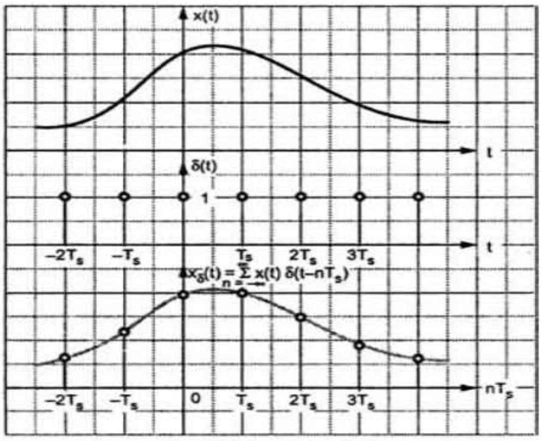


Fig. 5 CT and its DT signal

Proof of sampling theorem

There are two parts : (I) Representation of $x(t)$ in terms of its samples

(II) Reconstruction of $x(t)$ from its samples.

Part I : Representation of $x(t)$ in its samples $x(nT_s)$

> **Step 1 :** Define $x_\delta(t)$
>
> **Step 2 :** Fourier transform of $x_\delta(t)$ i.e. $X_\delta(f)$
>
> **Step 3 :** Relation between $X(f)$ and $X_\delta(f)$
>
> **Step 4 :** Relation between $x(t)$ and $x(nT_s)$

Step 1 : Define $x_\delta(t)$

The sampled signal $x_\delta(t)$ is given as,

$$x_\delta(t) = \sum_{n=-\infty}^{\infty} x(t)\,\delta(t-nT_s) \qquad \dots 1$$

Here observe that $x_\delta(t)$ is the product of x_δ and impulse train $\delta(t)$ as shown in above fig In the above equation $\delta(t-nT_s)$ indicates the samples placed at $\pm T_s$, $\pm 2T_s$, $\pm 3T_s$ ... and so on.

Step 2 : FT of $x_\delta(t)$ i.e. $X_\delta(f)$

Taking FT of equation (1.3.1).

$$X_\delta(f) = \text{FT}\left\{\sum_{n=-\infty}^{\infty} x(t)\,\delta(t-nT_s)\right\}$$

$$= \text{FT}\ \{\text{Product of } x(t) \text{ and impulse train}\}$$

We know that FT of product in time domain becomes convolution in frequency domain. i.e.,

$$X_\delta(f) = \text{FT}\ \{x(t)\} * FT\ \{\delta(t-nT_s)\} \qquad \text{.......................} \quad 2$$

By definitions, $x(t) \xleftrightarrow{FT} X(f)$ and

$$\delta(t-nT_s) \xleftrightarrow{FT} f_s \sum_{n=-\infty}^{\infty} \delta(f-nf_s)$$

Hence equation (1.3.2) becomes,

$$X_\delta(f) = X(f) * f_s \sum_{n=-\infty}^{\infty} \delta(f-nf_s)$$

Since convolution is linear,

$$X_\delta(f) = f_s \sum_{n=-\infty}^{\infty} X(f) * \delta(f-nf_s)$$

$$= f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad \textit{By shifting property of impulse function}$$

$$= \cdots f_s\, X(f-2f_s) + f_s\, X(f-f_s) + f_s\, X(f) + f_s\, X(f-f_s) + f_s\, X(f-2f_s) + \cdots$$

Comments

(i) The RHS of above equation shows that X(f) is placed at $\pm f_s, \pm 2f_s, \pm 3f_s, \cdots$

(ii) This means X(f) is periodic in f_s.

(iii) If sampling frequency is $f_s = 2W$, then the spectrums $X(f)$ just touch each other.

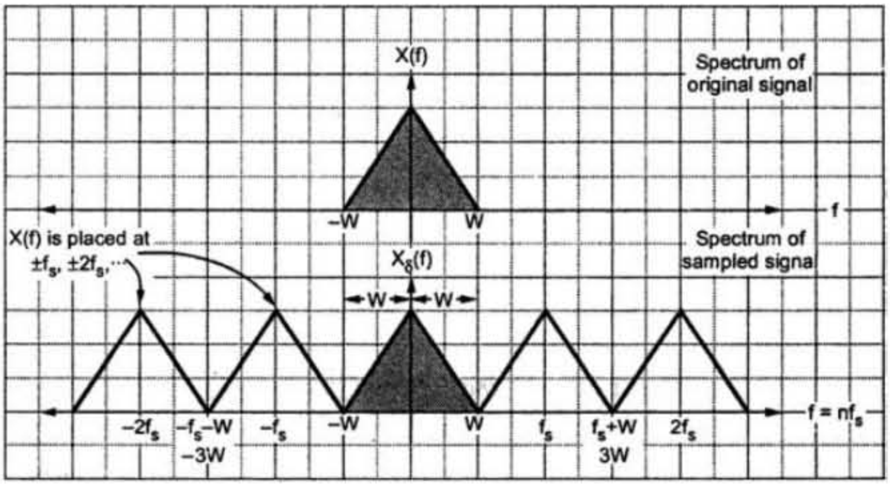


Fig. 6 **Spectrum of original signal and sampled signal**

Step 3 : Relation between $X(f)$ and $X_\delta(f)$

Important assumption : Let us assume that $f_s = 2W$, then as per above diagram.

$$X_\delta(f) = f_s\, X(f) \qquad \text{for } -W \le f \le W \text{ and } f_s = 2W$$

or
$$X(f) = \frac{1}{f_s} X_\delta(f) \qquad \dots \quad 3$$

Step 4 : Relation between $x(t)$ and $x(nT_s)$

DTFT is,
$$X(\Omega) = \sum_{n=-\infty}^{\infty} x(n)\, e^{-j\Omega n}$$

∴
$$X(f) = \sum_{n=-\infty}^{\infty} x(n)\, e^{-j2\pi f n} \qquad \dots \quad 4$$

In above equation 'f' is the frequency of DT signal. If we replace $X(f)$ by $X_\delta(f)$, then 'f' becomes frequency of CT signal. i.e.,

$$X_\delta(f) = \sum_{n=-\infty}^{\infty} x(n) e^{-j2\pi \frac{f}{f_s} n}$$

In above equation 'f' is frequency of CT signal. And $\frac{f}{f_s}$ = Frequency of DT signal in equation 4 Since $x(n) = x(nT_s)$, i.e. samples of $x(t)$, then we have,

$$X_\delta(f) = \sum_{n=-\infty}^{\infty} x(nT_s) e^{-j2\pi f n T_s} \text{ since } \frac{1}{f_s} = T_s$$

Putting above expression in equation 3 ,

$$X(f) = \frac{1}{f_s} \sum_{n=-\infty}^{\infty} x(nT_s) e^{-j2\pi f n T_s}$$

Inverse Fourier Transform (IFT) of above equation gives $x(t)$ i.e.,

$$x(t) = IFT\left\{\frac{1}{f_s} \sum_{n=-\infty}^{\infty} x(nT_s) e^{-j2\pi f n T_s}\right\} \quad \ldots \quad 5$$

Comments :

i) Here $x(t)$ is represented completely in terms of $x(nT_s)$.

ii) Above equation holds for $f_s = 2W$. This means if the samples are taken at the rate of 2W or higher, $x(t)$ is completely represented by its samples.

iii) First part of the sampling theorem is proved by above two comments.

Part II : Reconstruction of $x(t)$ from its samples

Step 1 : Take inverse Fourier transform of $X(f)$ which is in terms of $X_\delta(f)$.

Step 2 : Show that $x(t)$ is obtained back with the help of interpolation function.

Step 1 : The IFT of equation 5 becomes,

$$x(t) = \int_{-\infty}^{\infty} \left\{\frac{1}{f_s} \sum_{n=-\infty}^{\infty} x(nT_s) e^{-j2\pi f n T_s}\right\} e^{j2\pi f t}\, df$$

Here the integration can be taken from $-W \le f \le W$. Since $X(f) = \frac{1}{f_s} X_\delta(f)$ for $-W \le f \le W$. (See Fig. 6).

$$\therefore \qquad x(t) = \int_{-W}^{W} \frac{1}{f_s} \sum_{n=-\infty}^{\infty} x(nT_s) e^{-j2\pi f n T_s} \cdot e^{j2\pi f t}\, df$$

Interchanging the order of summation and integration,

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \frac{1}{f_s} \int_{-W}^{W} e^{j2\pi f(t-nT_s)}\, df$$

$$= \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \frac{1}{f_s} \cdot \left[\frac{e^{j2\pi f(t-nT_s)}}{j2\pi(t-nT_s)} \right]_{-W}^{W}$$

$$= \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \frac{1}{f_s} \left\{ \frac{e^{j2\pi W(t-nT_s)} - e^{-j2\pi W(t-nT_s)}}{j2\pi(t-nT_s)} \right\}$$

$$= \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \frac{1}{f_s} \cdot \frac{\sin 2\pi W(t-nT_s)}{\pi(t-nT_s)}$$

$$= \sum_{n=-\infty}^{\infty} x(nT_s) \frac{\sin \pi(2Wt - 2WnT_s)}{\pi(f_s t - f_s nT_s)}$$

Here $f_s = 2W$, hence $T_s = \frac{1}{f_s} = \frac{1}{2W}$. Simplifying above equation,

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \frac{\sin \pi(2Wt-n)}{\pi(2Wt-n)}$$

$$= \sum_{n=-\infty}^{\infty} x(nT_s) \operatorname{sinc}(2Wt-n) \qquad \text{since } \frac{\sin \pi\theta}{\pi\theta} = \operatorname{sinc}\theta \qquad \ldots \quad 6$$

Step 2 : Let us interpret the above equation. Expanding we get,

$$x(t) = \cdots + x(-2T_s)\operatorname{sinc}(2Wt+2) + x(-T_s)\operatorname{sinc}(2Wt+1) + x(0)\operatorname{sinc}(2Wt) + x(T_s)\operatorname{sinc}(2Wt-1) + \cdots$$

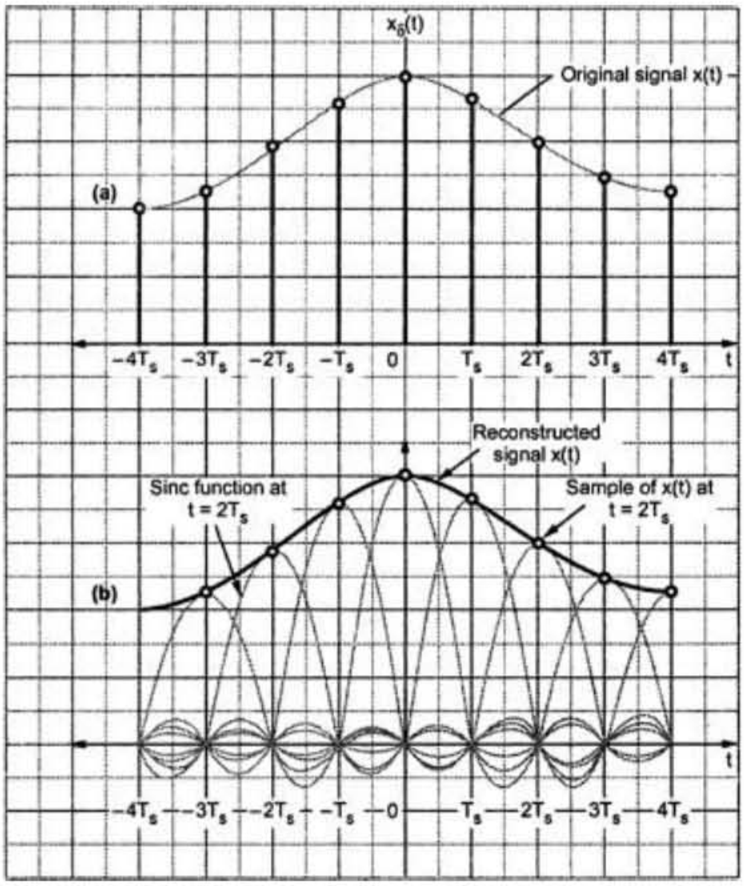


Fig. 7 (a) Sampled version of signal x(t)
(b) Reconstruction of x(t) from its samples

Comments :

i) The samples $x(nT_s)$ are weighted by sinc functions.

ii) The sinc function is the interpolating function. Fig. 7 shows, how $x(t)$ is interpolated.

Step 3 : Reconstruction of $x(t)$ by lowpass filter

When the interpolated signal of equation 6 is passed through the lowpass filter of bandwidth $-W \leq f \leq W$, then the reconstructed waveform shown in above Fig. 7 (b) is obtained. The individual sinc functions are interpolated to get smooth $x(t)$.

PCM Generator:

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

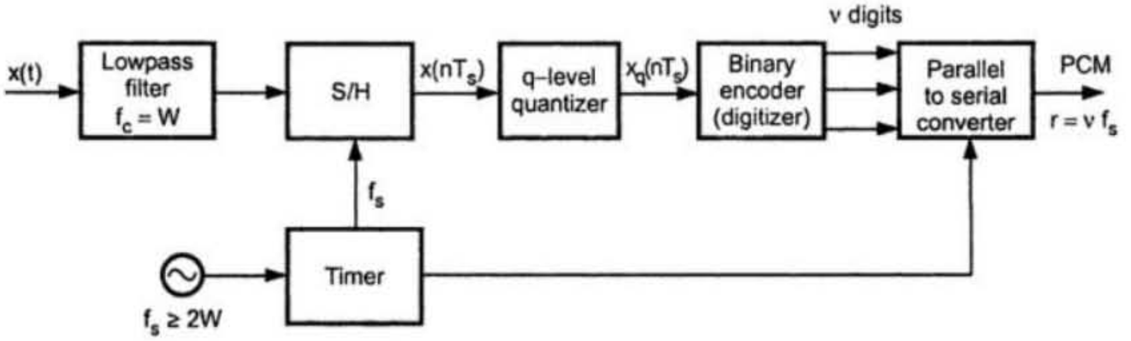


Fig. 8 PCM generator

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

$$f_s \geq 2W$$

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

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

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

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

Transmission BW in PCM:

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

$$q = 2^v \quad \ldots \quad 1$$

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

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

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

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

We know that, Number of samples per second = f_s

∴ Number of bits per second is given by,

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

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

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

Here $f_s \geq 2W$.

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

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

PCM Receiver:

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

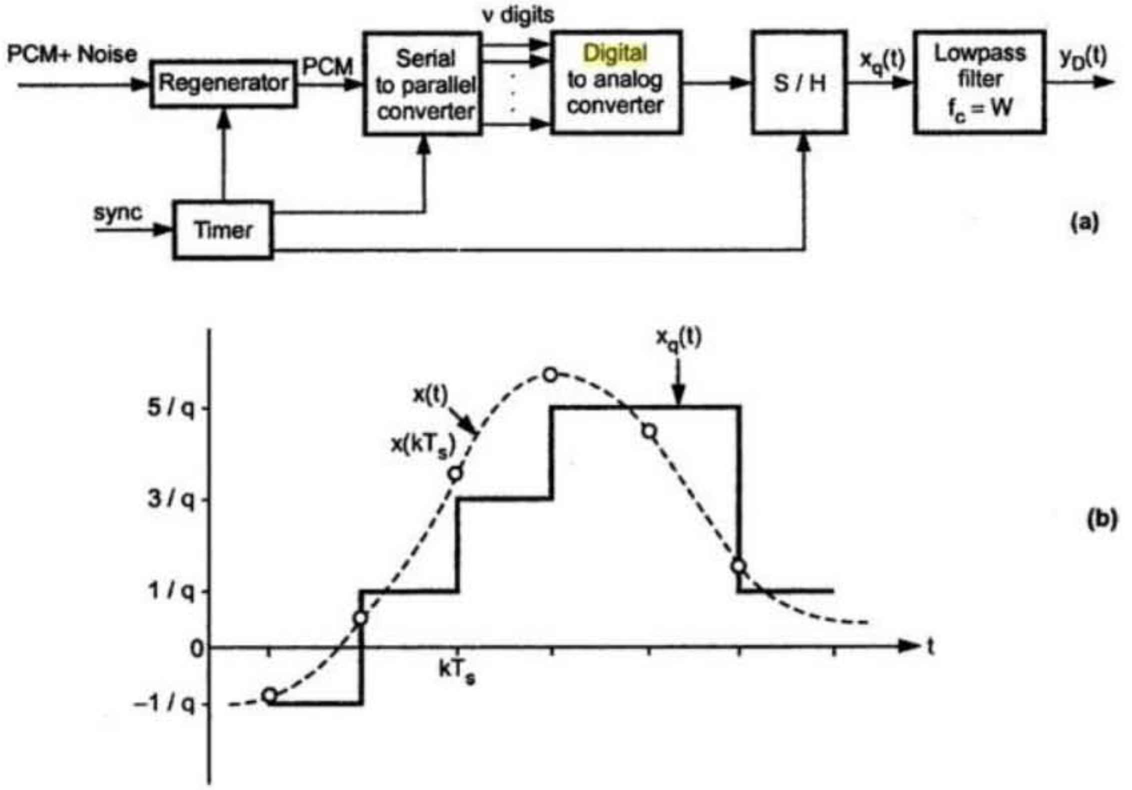


Fig. 9 (a) PCM receiver
(b) Reconstructed waveform

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

Quantization

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

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

Uniform Quantization:

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

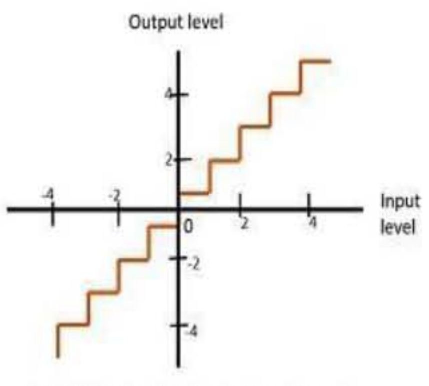


Fig 1 : Mid-Rise type Uniform Quantization

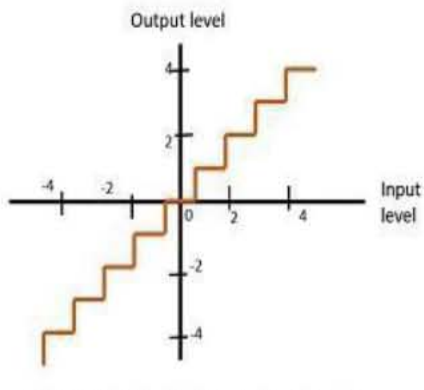


Fig 2 : Mid-Tread type Uniform Quantization

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

Quantization Noise and Signal to Noise ratio in PCM System:

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

Step 1 : Quantization Error

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

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

Step 2 : Step size

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

Therefore the total amplitude range becomes,

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

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

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

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

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

Therefore step size will be,

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

Step 3 : Pdf of Quantization error

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

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

i.e.

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

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

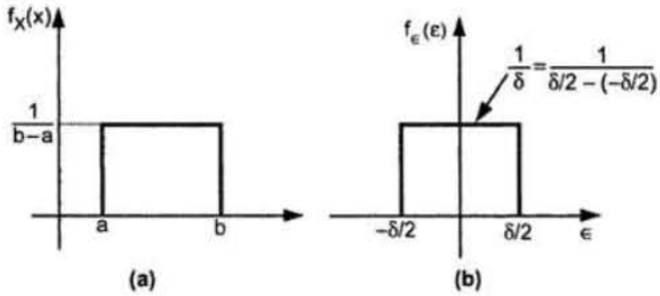


Fig. 10 **(a) Uniform distribution**

(b) Uniform distribution for quantization error

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

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

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

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

Step 4 : Noise Power

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

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

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

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

The noise power is given as,

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

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

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

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

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

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

From equation 9 we can write above equation as,

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

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

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

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

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

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

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

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

Thus we have,

Normalized noise power
or Quantization noise power $= \frac{\delta^2}{12}$ **; For linear quantization.**
or Quantization error (in terms of power)

... (16)

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

signal to quantization noise ratio is given as,

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

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

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

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

Putting this value in equation (3) we have,

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

Putting this value in equation 1.8.30 we get,

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

Let normalized signal power be denoted as 'P'.

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

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

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

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

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

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

Then signal to quantization noise ratio will be,

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

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

$$P \leq 1 \quad \text{... (23)}$$

Then the signal to noise ratio is given as,

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

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

Expressing the signal to noise ratio in decibels,

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

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

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

Thus,

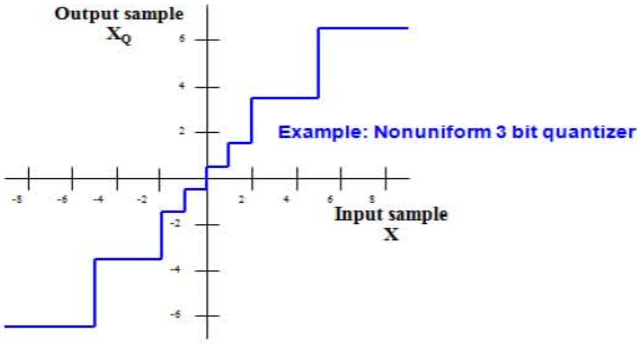
Signal to Quantization noise ratio

for normalized values of power : $\left(\frac{S}{N}\right) dB \leq (4.8 + 6v) \, dB$

'P' and amplitude of input $x(t)$ ... (25)

Non-Uniform Quantization:

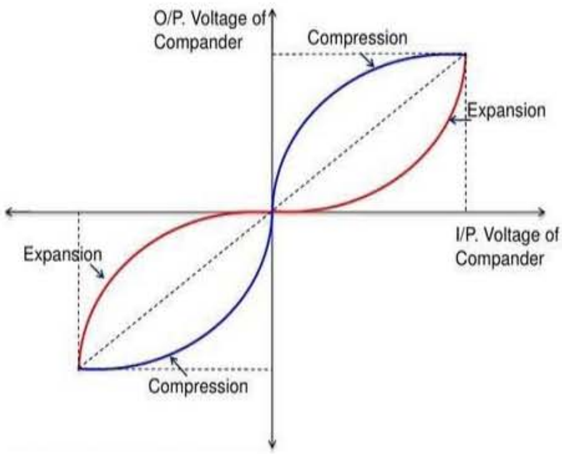
In non-uniform quantization, the step size is not fixed. It varies according to certain law or as per input signal amplitude. The following fig shows the characteristics of Non uniform quantizer.

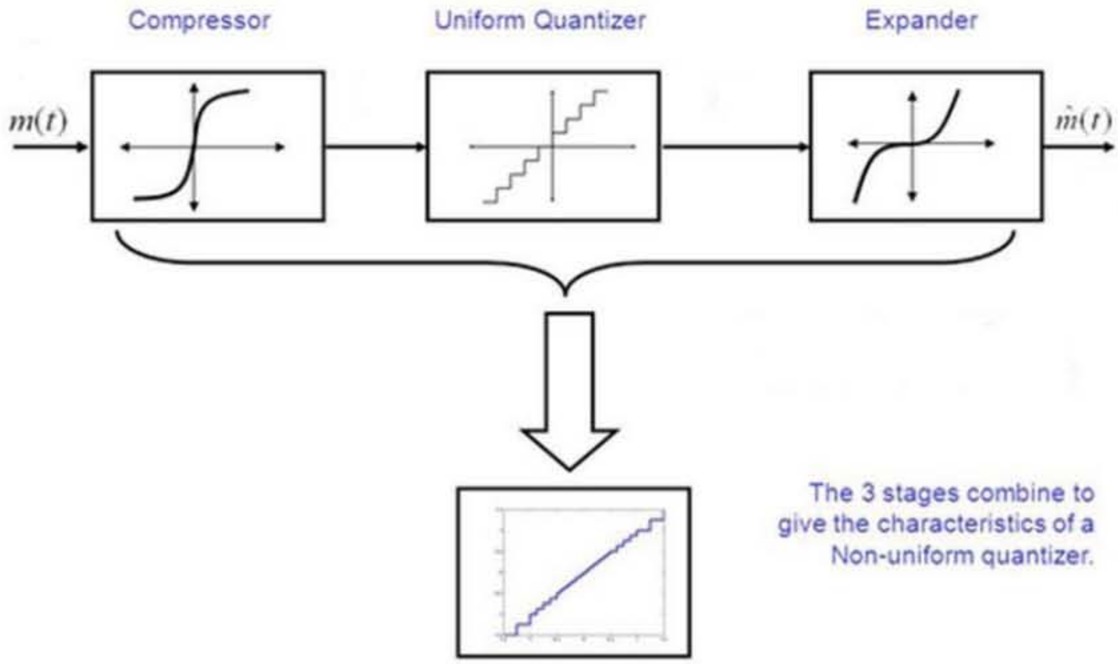


In this figure observe that step size is small at low input signal levels. Hence quantization error is also small at these inputs. Therefore signal to quantization noise power ratio is improved at low signal levels. Stepsize is higher at high input levels. Hence signal to noise power ratio remains almost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

Companding PCM System:

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





μ - Law Companding for Speech Signals

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

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

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

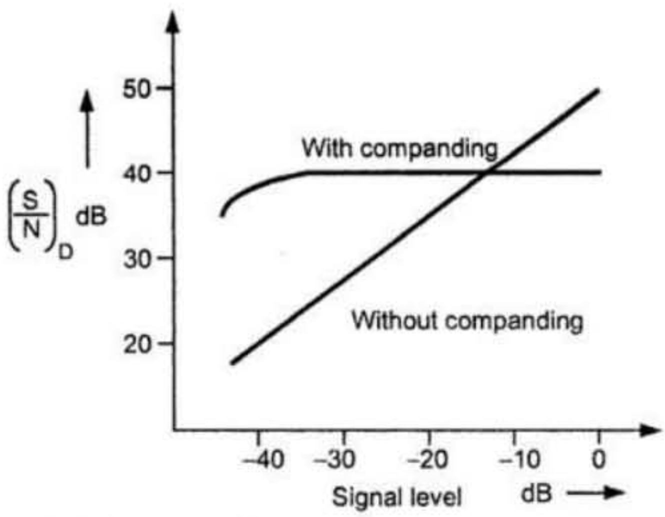


Fig. 11 PCM performance with μ - law companding

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

A-Law for Companding

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

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

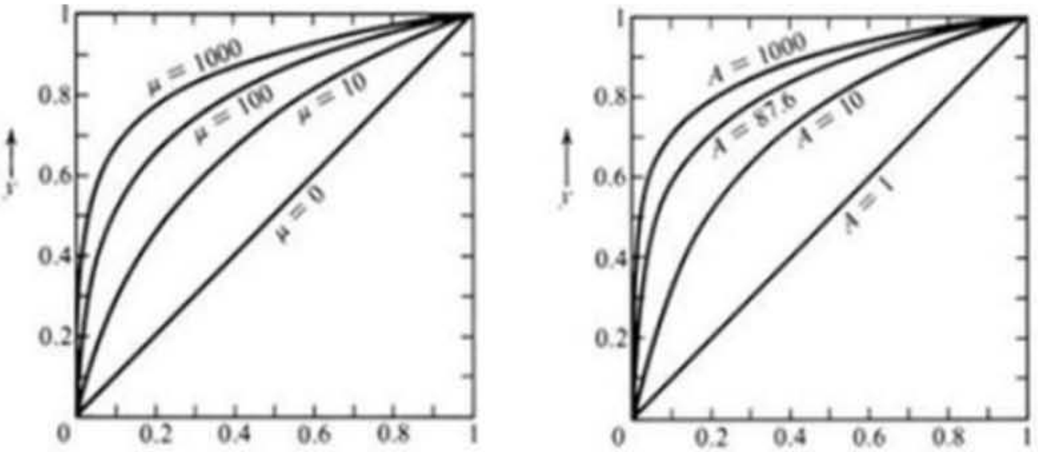
When A = 1, we get uniform quantization. The practical value for A is 87.56. Both A-law and μ-law companding is used for PCM telephone systems.

Signal to Noise Ratio of Companded PCM

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

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

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



Differential Pulse Code Modulation (DPCM):

Redundant Information in PCM:

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

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

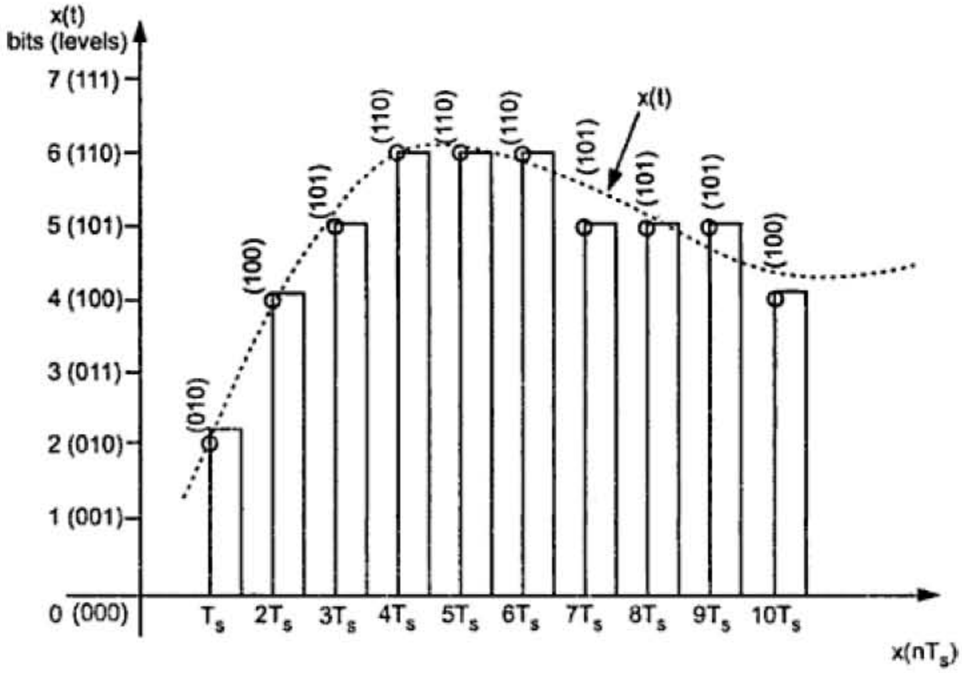


Fig. Redundant information in PCM

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

Principle of DPCM

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

DPCM Transmitter

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

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

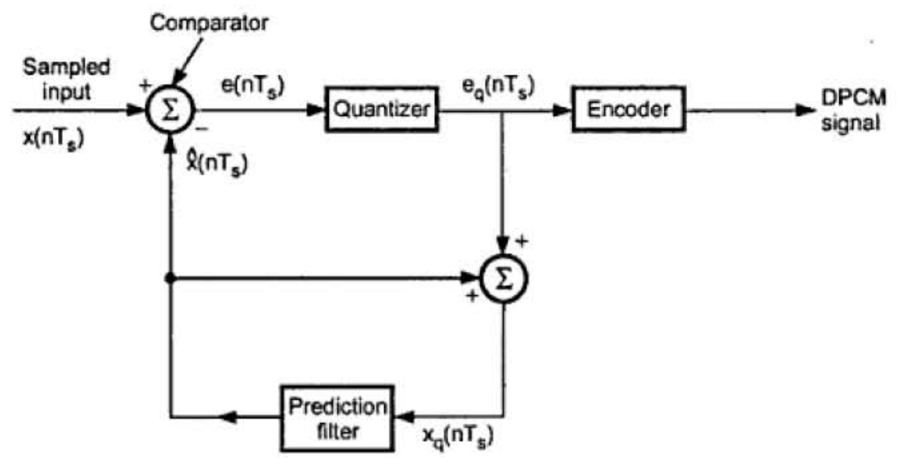


Fig. Differential pulse code modulation transmitter

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

The quantizer output can be written as,

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

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

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

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

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

Equation 1 is written as,

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

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

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

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

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

Reconstruction of DPCM Signal

Fig. shows the block diagram of DPCM receiver.

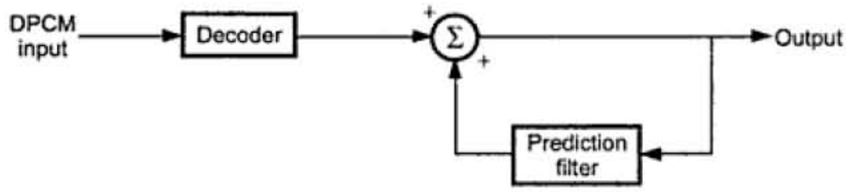


Fig. DPCM receiver

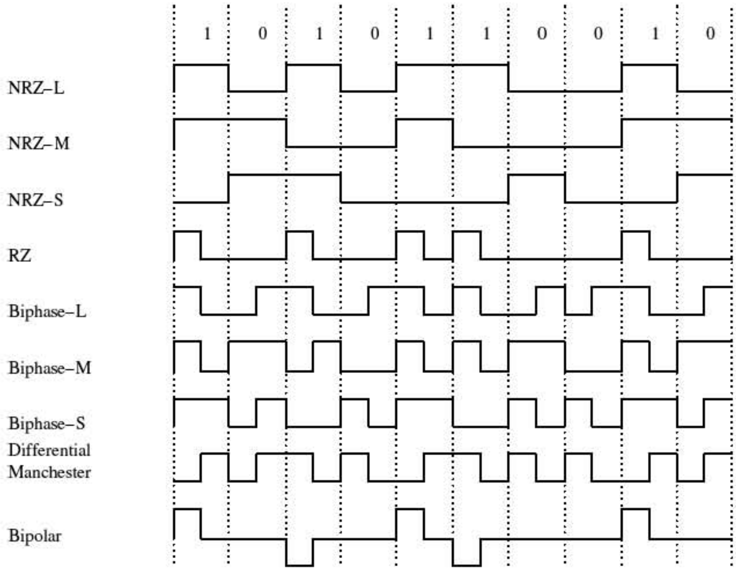
The decoder first reconstructs the quantized error signal from incoming binary signal. The prediction filter output and quantized error signals are summed up to give the quantized version of the original signal. Thus the signal at the receiver differs from actual signal by quantization error $q(nT_s)$,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

Line Coding:

In telecommunication, a line code is a code chosen for use within a communications system for transmitting a digital signal down a transmission line. Line coding is often used for digital data transport.

The waveform pattern of voltage or current used to represent the 1s and 0s of a digital signal on a transmission link is called **line** encoding. The common types of

line encoding are unipolar, polar, bipolar and Manchester encoding. **Line codes** are used commonly in computer communication networks over short distances.



Signal	Comments
NRZ–L	Non-return to zero level. This is the standard positive logic signal format used in digital circuits. 1 forces a high level 0 forces a low level
NRZ–M	Non return to zero mark 1 forces a transition 0 does nothing
NRZ–S	Non return to zero space 1 does nothing 0 forces a transition
RZ	Return to zero 1 goes high for half the bit period 0 does nothing
Biphase–L	Manchester. Two consecutive bits of the same type force a transition at the beginning of a bit period. 1 forces a negative transition in the middle of the bit 0 forces a positive transition in the middle of the bit
Biphase–M	There is always a transition at the beginning of a bit period. 1 forces a transition in the middle of the bit 0 does nothing
Biphase–S	There is always a transition at the beginning of a bit period. 1 does nothing 0 forces a transition in the middle of the bit
Differential Manchester	There is always a transition in the middle of a bit period. 1 does nothing 0 forces a transition at the beginning of the bit
Bipolar	The positive and negative pulses alternate. 1 forces a positive or negative pulse for half the bit period 0 does nothing

Time Division Multiplexing:

The sampling theorem provides the basis for transmitting the information contained in a band-limited message signal $m(t)$ as a sequence of samples of $m(t)$ taken uniformly at a rate that is usually slightly higher than the Nyquist rate. An important feature of the sampling process is a *conservation of time*. That is, the transmission of the message samples engages the communication channel for only a fraction of the sampling interval on a periodic basis, and in this way some of the time interval between adjacent samples is cleared for use by other independent message sources on a time-shared basis. We thereby obtain a *time-division multiplex* (TDM) *system*, which enables the joint utilization of a common communication channel by a plurality of independent message sources without mutual interference among them.

The concept of TDM is illustrated by the block diagram shown in Figure Each input message signal is first restricted in bandwidth by a low-pass anti-aliasing filter to remove the frequencies that are nonessential to an adequate signal representation. The low-pass filter outputs are then applied to a *commutator*, which is usually implemented using electronic switching circuitry. The function of the commutator is twofold: (1) to take a narrow sample of each of the N input messages at a rate f_s that is slightly higher than $2W$, where W is the cutoff frequency of the anti-aliasing filter, and (2) to sequentially interleave these N samples inside the sampling interval T_s. Indeed, this latter function is the essence of the time-division multiplexing operation. Following the commutation process, the multiplexed signal is applied to a *pulse modulator*, the purpose of which is to transform the multiplexed signal into a form suitable for transmission over the common channel. It is clear that the use of time-division multiplexing introduces a bandwidth expansion factor N, because the scheme must squeeze N samples derived from N independent message sources into a time slot equal to one sampling interval. At the receiving end of the system, the received signal is applied to a *pulse demodulator*, which performs the reverse operation of the pulse modulator. The narrow samples produced at the pulse demodulator output are distributed to the appropriate low-pass reconstruction filters by means of a *decommutator*, which operates in *synchronism* with the commutator in the transmitter. This synchronization is essential for a satisfactory operation of the system.

The way this synchronization is implemented depends naturally on the method of pulse modulation used to transmit the multiplexed sequence of samples.

The TDM system is highly sensitive to dispersion in the common channel, that is, to variations of amplitude with frequency or lack of proportionality of phase with frequency. Accordingly, accurate equalization of both magnitude and phase responses of the channel is necessary to ensure a satisfactory operation of the system;

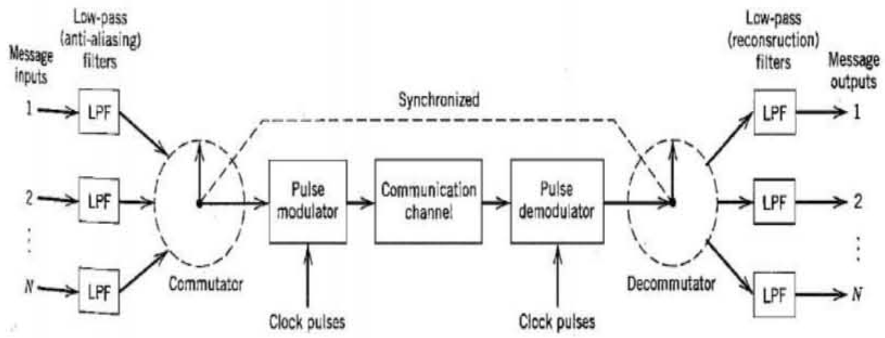


FIGURE Block diagram of TDM system.

TDM is immune to nonlinearities in the channel as a source of crosstalk. The reason for this behaviour is that different message signals are not simultaneously applied to the channel.

Introduction to Delta Modulation

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

Delta Modulation

Operating Principle of DM

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

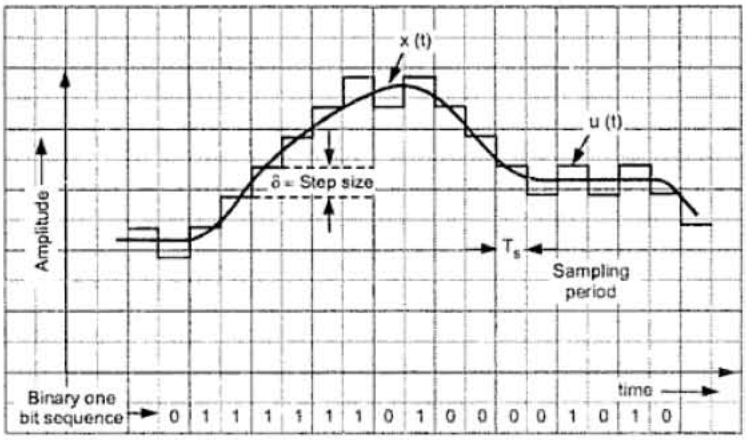


Fig. Delta modulation waveform

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

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

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

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

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

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

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

= Last sample approximation of staircase waveform.

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

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

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

$$b(nT_s) = +\delta \quad \text{if} \quad x(nT_s) \geq \hat{x}(nT_s)$$

$$= -\delta \quad \text{if} \quad x(nT_s) < \hat{x}(nT_s) \quad \dots (4)$$

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

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

T_s = Sampling interval.

DM Transmitter

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

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

$$u(nT_s) = u(nT_s - T_s) + [\pm\delta] \quad \text{or}$$

$$= u[(n-1)T_s] + b(nT_s) \quad \ldots (5)$$

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

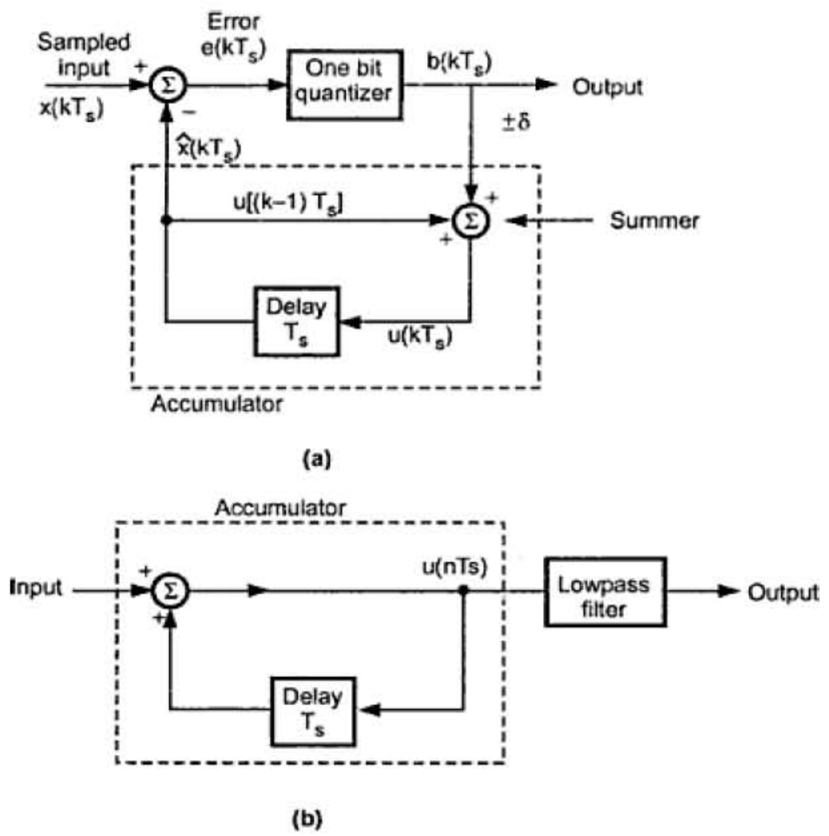


Fig. (a) Delta modulation transmitter and (b) Delta modulation receiver

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

DM Receiver

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

Advantages and Disadvantages of Delta Modulation

Advantages of Delta Modulation

The delta modulation has following advantages over PCM,

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

Disadvantages of Delta Modulation

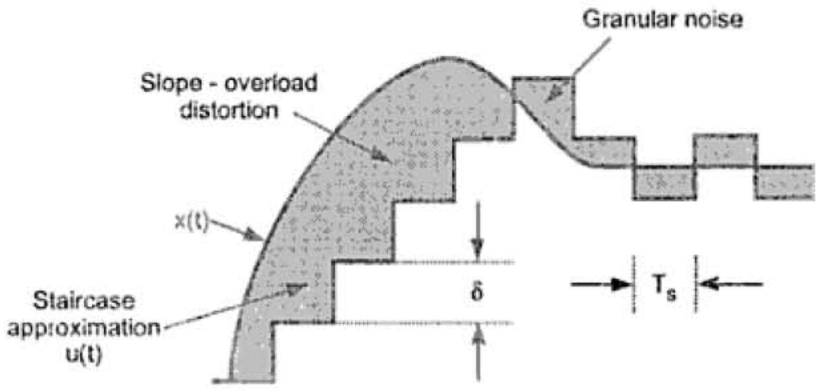


Fig. Quantization errors in delta modulation

The delta modulation has two drawbacks -

Slope Overload Distortion (Startup Error)

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

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

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

Granular Noise (Hunting)

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

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

Adaptive Delta Modulation

Operating Principle

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

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

Transmitter and Receiver

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

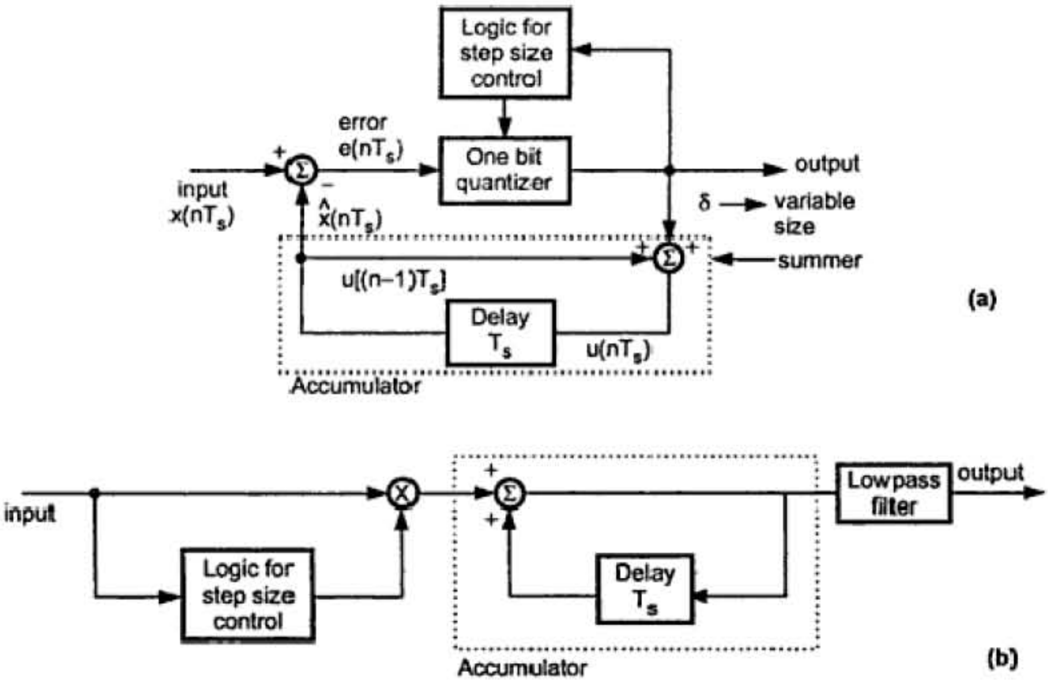


Fig. Adaptive delta modulator (a) Transmitter (b) Receiver

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

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

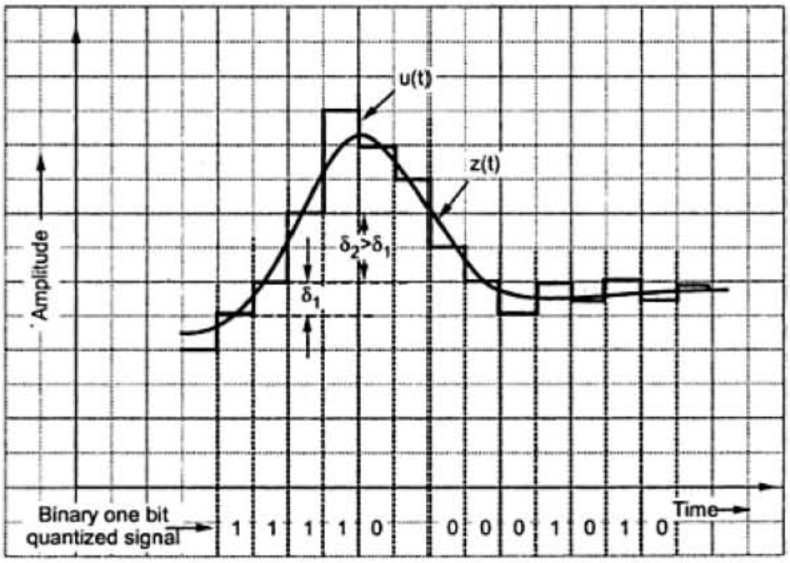


Fig. Waveforms of adaptive delta modulation

Advantages of Adaptive Delta Modulation

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

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

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

Condition for Slope overload distortion occurrence:

Slope overload distortion will occur if

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

where T_s is the sampling period.

Let the sine wave be represented as,

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

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

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

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

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

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

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

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

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

or

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

Expression for Signal to Quantization Noise power ratio for Delta Modulation:

To obtain signal power :

slope overload distortion will not occur if

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

Here A_m is peak amplitude of sinusoided signal

δ is the step size

f_m is the signal frequency and

T_s is the sampling period.

From above equation, the maximum signal amplitude will be,

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

Signal power is given as,

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

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

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

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

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

Putting for A_m from equation 1

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

This is an expression for signal power in delta modulation.

(ii) To obtain noise power

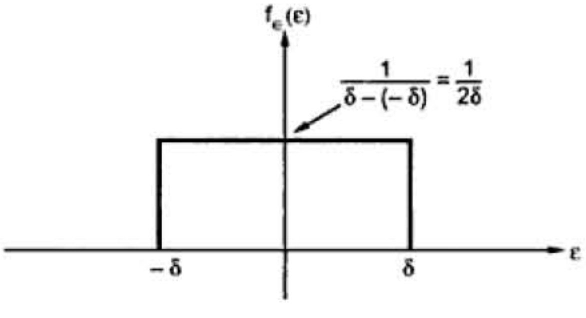


Fig. Uniform distribution of quantization error

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

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

The noise power is given as,

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

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

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

mean square value is given as,

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

From equation 3

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

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

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

Hence noise power will be,

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

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

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

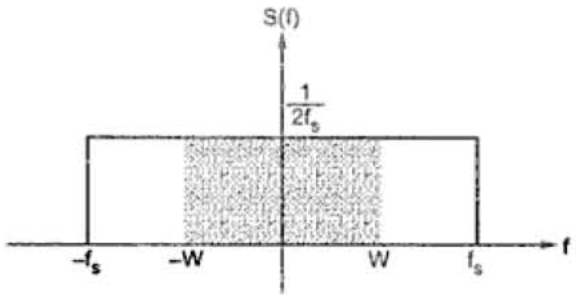


Fig. PSD of noise

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

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

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

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

(iii) To obtain signal to noise power ratio

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

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

From equation 2. and equation 6

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

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

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

UNIT-2

DIGITAL MODULATION TECHNIQUES

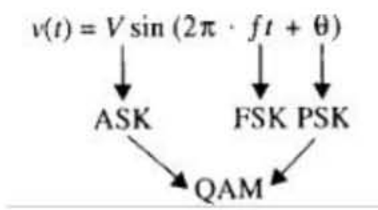
Digital Modulation provides more information capacity, high data security, quicker system availability with great quality communication. Hence, digital modulation techniques have a greater demand, for their capacity to convey larger amounts of data than analog ones.

There are many types of digital modulation techniques and we can even use a combination of these techniques as well. In this chapter, we will be discussing the most prominent digital modulation techniques.

if the information signal is digital and the amplitude (lV of the carrier is varied proportional to the information signal, a digitally modulated signal called amplitude shift keying (ASK) is produced.

If the frequency (f) is varied proportional to the information signal, frequency shift keying (FSK) is produced, and if the phase of the carrier (0) is varied proportional to the information signal,

phase shift keying (PSK) is produced. If both the amplitude and the phase are varied proportional to the information signal, quadrature amplitude modulation (QAM) results. ASK, FSK, PSK, and QAM are all forms of digital modulation:



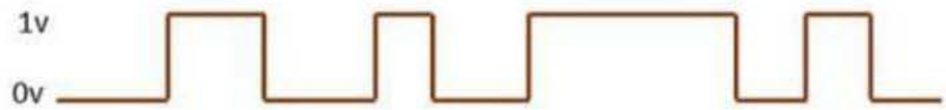
a simplified block diagram for a digital modulation system.

Amplitude Shift Keying

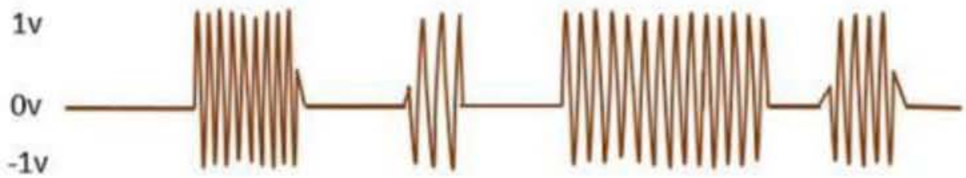
The amplitude of the resultant output depends upon the input data whether it should be a zero level or a variation of positive and negative, depending upon the carrier frequency.

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.

Following is the diagram for ASK modulated waveform along with its input.



Input binary sequence



ASK Modulated output wave

Any modulated signal has a high frequency carrier. The binary signal when ASK is modulated, gives a zero value for LOW input and gives the carrier output for HIGH input.
Mathematically, amplitude-shift keying is

$$v_{(ask)}(t) = [1 + v_m(t)]\left[\frac{A}{2}\cos(\omega_c t)\right]$$

where vask(t) = amplitude-shift keying wave
vm(t) = digital information (modulating) signal (volts)
A/2 = unmodulated carrier amplitude (volts)
ωc= analog carrier radian frequency (radians per second, 2πfct)

In above Equation, the modulating signal [vm(t)] is a normalized binary waveform, where + 1 V = logic 1 and -1 V = logic 0. Therefore, for a logic 1 input, vm(t) = + 1 V, Equation 2.12 reduces to

$$v_{(ask)}(t) = [1 + 1]\left[\frac{A}{2}\cos(\omega_c t)\right]$$

$$= A\cos(\omega_c t)$$

Mathematically, amplitude-shift keying is (2.12) where vask(t) = amplitude-shift keying wave vm(t) = digital information (modulating) signal (volts) A/2 = unmodulated carrier amplitude (volts) ωc= analog carrier radian frequency (radians per second, 2πfct) In Equation 2.12, the modulating signal [vm(t)] is a normalized binary waveform, where + 1 V = logic 1 and -1 V = logic 0. Therefore, for a logic 1 input, vm(t) = + 1 V, Equation 2.12 reduces to and for a logic 0 input, vm(t) = -1 V,Equation reduces to

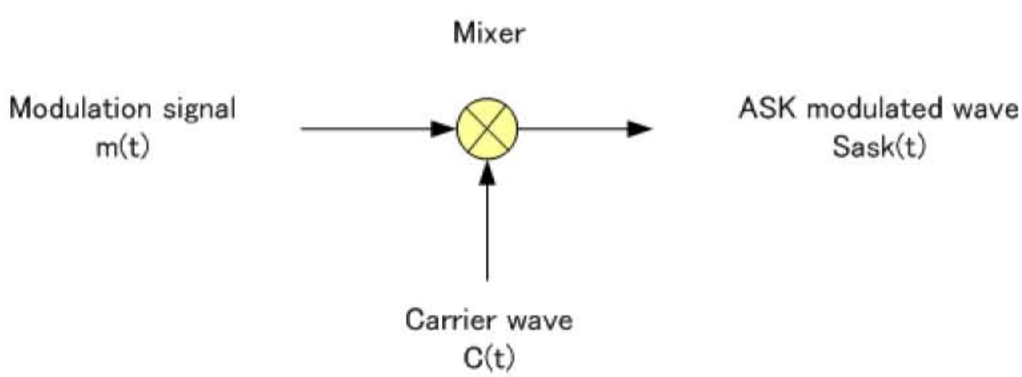
$$v_{(ask)}(t) = [1 - 1]\left[\frac{A}{2}\cos(\omega_c t)\right]$$

Thus, the modulated wave vask(t),is either A cos(ωct) or 0. Hence, the carrier is either "on “or "off," which is why amplitude-shift keying is sometimes referred to as on-off keying (OOK).
it can be seen that for every change in the input binary data stream, there is one change in the ASK waveform, and the time of one bit (tb) equals the time of one analog signaling element (t,).
B = fb/1 = fb   baud = fb/1 = fb

Example :
Determine the baud and minimum bandwidth necessary to pass a 10 kbps binary signal using amplitude shift keying. 10Solution For ASK, N = 1, and the baud and minimum bandwidth are determined from Equations 2.11 and 2.10, respectively:

B = 10,000 / 1 = 10,000
baud = 10, 000 /1 = 10,000

The use of amplitude-modulated analog carriers to transport digital information is a relatively low-quality, low-cost type of digital modulation and, therefore, is seldom used except for very low-speed telemetry circuits.
ASK TRANSMITTER:



The input binary sequence is applied to the product modulator. The product modulator amplitude modulates the sinusoidal carrier .it passes the carrier when input bit is '1' .it blocks the carrier when input bit is '0.'

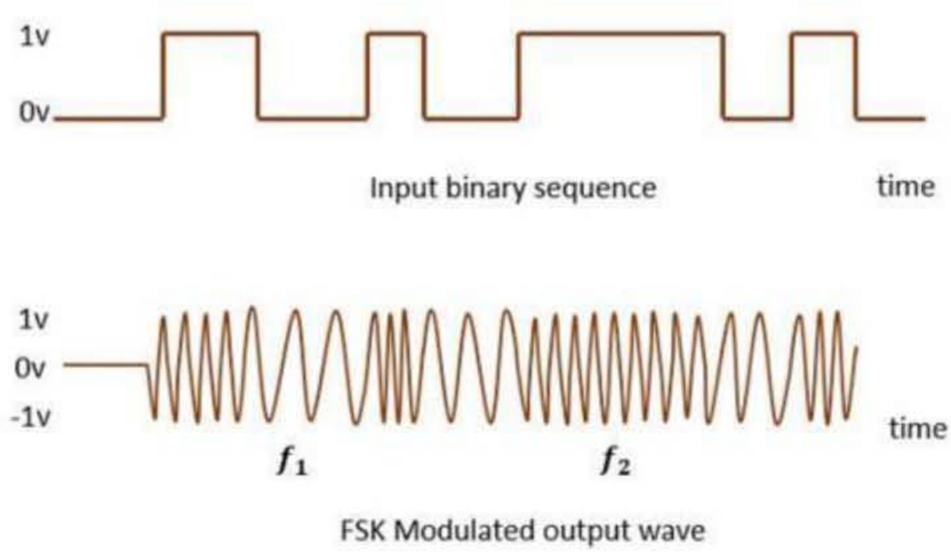
Coherent ASK DETECTOR:

FREQUENCYSHIFT KEYING

The frequency of the output signal will be either high or low, depending upon the input data applied.

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

Following is the diagram for FSK modulated waveform along with its input.



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**.

FSK is a form of constant-amplitude angle modulation similar to standard frequency modulation (FM) except the modulating signal is a binary signal that varies between two discrete voltage levels rather than a continuously changing analog waveform.Consequently, FSK is sometimes called *binary FSK* (BFSK). The general expression for FSK is where

$$v_{fsk}(t) = V_c \cos\{2\pi[f_c + v_m(t)\,\Delta f]t\}$$

$v_{fsk}(t)$ = binary FSK waveform

V_c = peak analog carrier amplitude (volts)

f_c = analog carrier center frequency (hertz)

f=peak change (shift)in the analog carrier frequency(hertz)

$v_m(t)$ = binary input (modulating) signal (volts)

From Equation 2.13, it can be seen that the peak shift in the carrier frequency (f) is proportional to the amplitude of the binary input signal ($v_m[t]$), and the direction of the shift is determined by the polarity.

The modulating signal is a normalized binary waveform where a logic 1 = + 1 V and a logic 0 = -1 V. Thus, for a logic 1 input, $v_m(t)$ = + 1, Equation 2.13 can be rewritten as

$$v_{fsk}(t) = V_c \cos[2\pi(f_c + \Delta f)t]$$

For a logic 0 input, $v_m(t)$ = -1, Equation becomes

$$v_{fsk}(t) = V_c \cos[2\pi(f_c - \Delta f)t]$$

With binary FSK, the carrier center frequency (f_c) is shifted (deviated) up and down in the frequency domain by the binary input signal as shown in Figure 2-3.

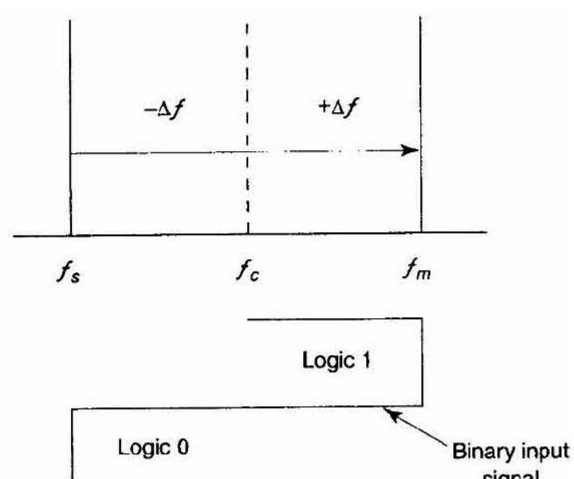


FIGURE: FSK in the frequency domain

As the binary input signal changes from a logic 0 to a logic 1 and vice versa, the output frequency shifts between two frequencies: a mark, or logic 1 frequency (f_m), and a space, or logic 0 frequency (f_s). The mark and space frequencies are separated from the carrier frequency by the peak frequency deviation (f) and from each other by 2 f.

Frequency deviation is illustrated in Figure 2-3 and expressed mathematically as

$$f = |f_m - f_s| / 2 \qquad (2.14)$$

where f = frequency deviation (hertz)

$|f_m - f_s|$ = absolute difference between the mark and space frequencies (hertz)

Figure 2-4a shows in the time domain the binary input to an FSK modulator and the corresponding FSK output.

When the binary input (f_b) changes from a logic 1 to a logic 0 and vice versa, the FSK output frequency shifts from a mark (f_m) to a space (f_s) frequency and vice versa.

In Figure 2-4a, the mark frequency is the higher frequency (f_c + f) and the space frequency is the lower frequency (f_c - f), although this relationship could be just the opposite.

Figure 2-4b shows the truth table for a binary FSK modulator. The truth table shows the input and output possibilities for a given digital modulation scheme.

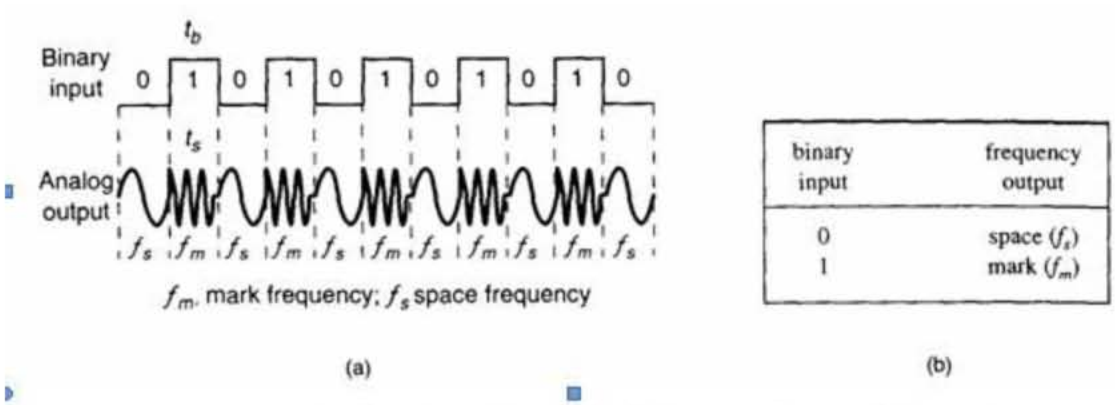


FIGURE 2-4 FSK in the time domain: (a) waveform: (b) truth table

FSK Bit Rate, Baud, and Bandwidth

In Figure 2-4a, it can be seen that the time of one bit *(t_b)* is the same as the time the FSK output is a mark of space frequency *(t_s)*. Thus, the bit time equals the time of an FSK signaling element, and the bit rate equals the baud.

The baud for binary FSK can also be determined by substituting N = 1 in Equation 2.11:

baud = f_b / 1 = f_b

The minimum bandwidth for FSK is given as

$$B= |(f_s - f_b) - (f_m - f_b)|$$

$$=|(f_s - f_m)| + 2f_b$$

and since $|(f_s - f_m)|$ equals 2 f, the minimum bandwidth can be approximated as

$$B= 2(f + f_b) \qquad (2.15)$$

where

B= minimum Nyquist bandwidth (hertz)

f= frequency deviation $|(f_m - f_s)|$ (hertz)

f_b = input bit rate (bps)

Example 2-2

Determine (a) the peak frequency deviation, (b) minimum bandwidth, and (c) baud for a binary FSK signal with a mark frequency of 49 kHz, a space frequency of 51 kHz, and an input bit rate of 2 kbps.

Solution

a. The peak frequency deviation is determined from Equation 2.14:

$$f= |149kHz - 51\ kHz| / 2 =1\ kHz$$

b. The minimum bandwidth is determined from Equation 2.15:

$$B = 2(100+ 2000)$$

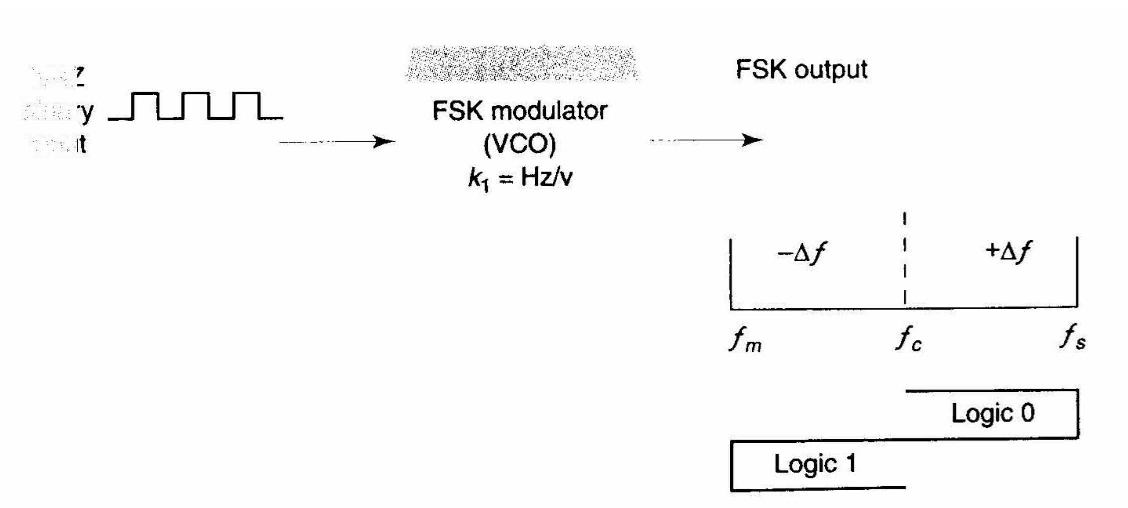
$$=6\ kHz$$

c. For FSK, $N = 1$, and the baud is determined from Equation 2.11 as

$$\text{baud} = 2000 / 1 = 2000$$

FSK TRANSMITTER:

Figure 2-6 shows a simplified binary FSK modulator, which is very similar to a conventional FM modulator and is very often a voltage-controlled oscillator (VCO).The center frequency (f_c) is chosen such that it falls halfway between the mark and space frequencies.



A logic 1 input shifts the VCO output to the mark frequency, and a logic 0 input shifts the VCO output to the space frequency. Consequently, as the binary input signal changes back and forth between logic 1 and logic 0 conditions, the VCO output shifts or deviates back and forth between the mark and space frequencies.

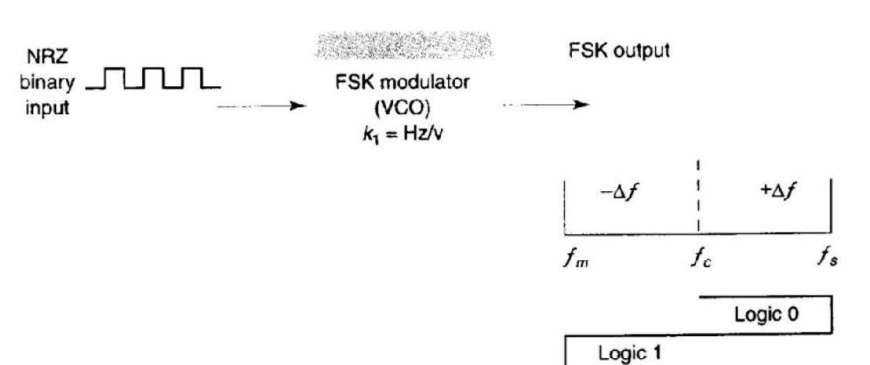


FIGURE 2-6 FSK modulator

A VCO-FSK modulator can be operated in the sweep mode where the peak frequency deviation is simply the product of the binary input voltage and the deviation sensitivity of the VCO.

With the sweep mode of modulation, the frequency deviation is expressed mathematically as

$$f = v_m(t)k_l \quad (2\text{-}19)$$

$v_m(t)$ = peak binary modulating-signal voltage (volts)

k_l = deviation sensitivity (hertz per volt).

FSK Receiver

FSK demodulation is quite simple with a circuit such as the one shown in Figure 2-7.

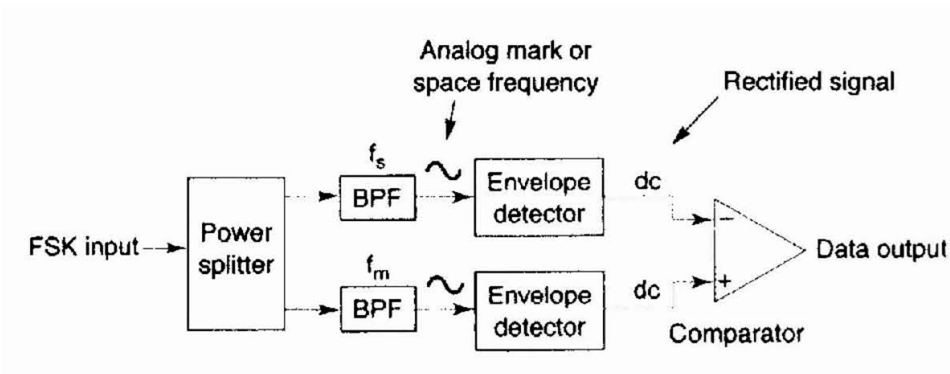


FIGURE 2-7 Noncoherent FSK demodulator

The FSK input signal is simultaneously applied to the inputs of both bandpass filters (BPFs) through a power splitter.The respective filter passes only the mark or only the space frequency on to its respective envelope detector.The envelope detectors, in turn, indicate the total power in each passband, and the comparator responds to the largest of the two powers.This type of FSK detection is referred to as noncoherent detection.

Figure 2-8 shows the block diagram for a coherent FSK receiver.The incoming FSK signal is multiplied by a recovered carrier signal that has the exact same frequency and phase as the transmitter reference.

However, the two transmitted frequencies (the mark and space frequencies) are not generally continuous; it is not practical to reproduce a local reference that is coherent with both of them. Consequently, coherent FSK detection is seldom used.

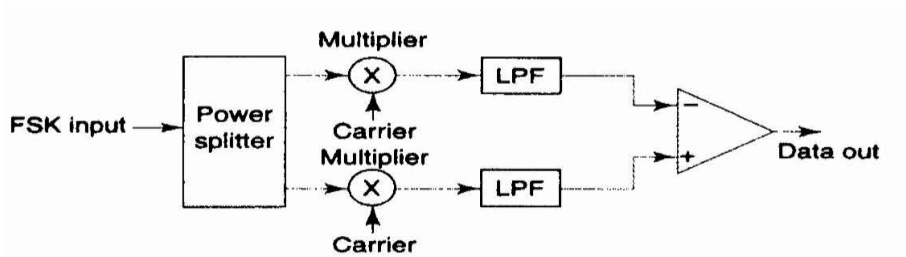
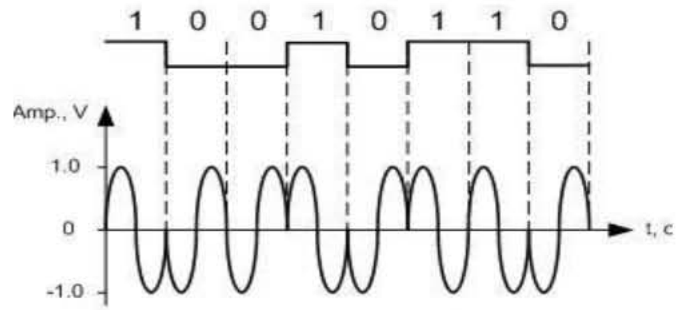


FIGURE 2-8 Coherent FSK demodulator

PHASESHIFT KEYING:

The phase of the output signal gets shifted depending upon the input. These are mainly of two types, namely BPSK and QPSK, according to the number of phase shifts. The other one is DPSK which changes the phase according to the previous value.



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.

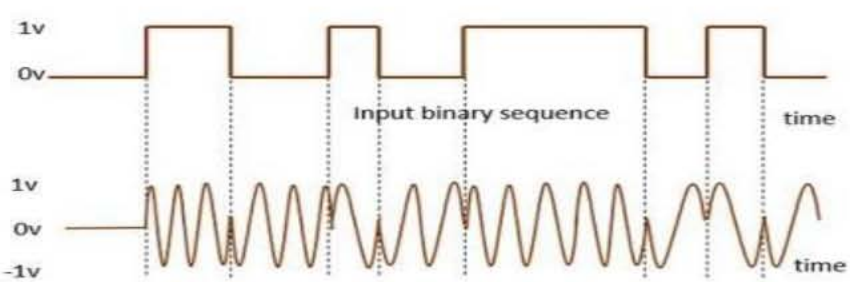
PSK is of two types, depending upon the phases the signal gets shifted. They are −

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°.

BPSK is basically a DSB-SC (Double Sideband Suppressed Carrier) modulation scheme, for message being the digital information.

Following is the image of BPSK Modulated output wave along with its input.



BPSK Modulated output wave

Binary Phase-Shift Keying

The simplest form of PSK is ***binary phase-shift keying*** (BPSK), where N = 1 and *M* = *2*.Therefore, with BPSK, two phases ($2^1 = 2$) are possible for the carrier.One phase represents a logic 1, and the other phase represents a logic 0. As the input digital signal changes state (i.e., from a 1 to a 0 or from a 0 to a 1), the phase of the output carrier shifts between two angles that are separated by 180°.

Hence, other names for BPSK are *phase reversal keying* (PRK) and *biphase modulation.* BPSK is a form of square-wave modulation of a *continuous wave (CW)* signal.

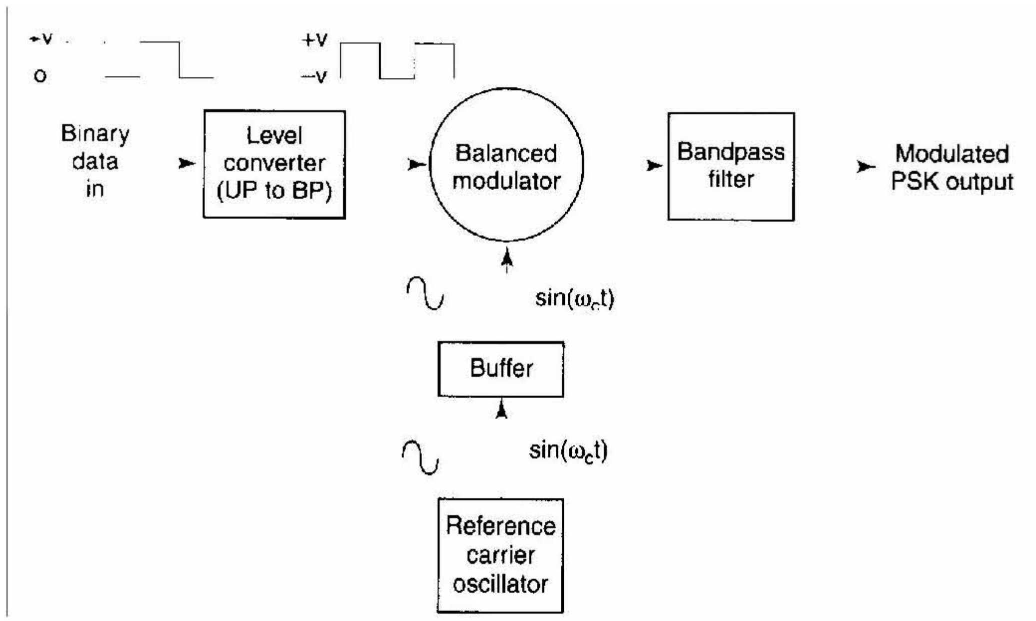


FIGURE 2-12 BPSK transmitter

BPSK TRANSMITTER:

Figure 2-12 shows a simplified block diagram of a BPSK transmitter. The balanced modulator acts as a phase reversing switch. Depending on the logic condition of the digital input, the carrier is transferred to the output either in phase or 180° out of phase with the reference carrier oscillator.

Figure 2-13 shows the schematic diagram of a balanced ring modulator. The balanced modulator has two inputs: a carrier that is in phase with the reference oscillator and the binary digital data. For the balanced modulator to operate properly, the digital input voltage must be much greater than the peak carrier voltage.

This ensures that the digital input controls the on/off state of diodes D1 to D4. If the binary input is a logic 1(positive voltage), diodes D 1 and D2 are forward biased and on, while diodes D3 and D4 are reverse biased and off (Figure 2-13b). With the polarities shown, the carrier voltage is developed across transformer T2 in phase with the carrier voltage across T

1. Consequently, the output signal is in phase with the reference oscillator.
If the binary input is a logic 0 (negative voltage), diodes Dl and D2 are reverse biased and off, while diodes D3 and D4 are forward biased and on (Figure 9-13c). As a result, the carrier voltage is developed across transformer T2 180° out of phase with the carrier voltage across T 1.

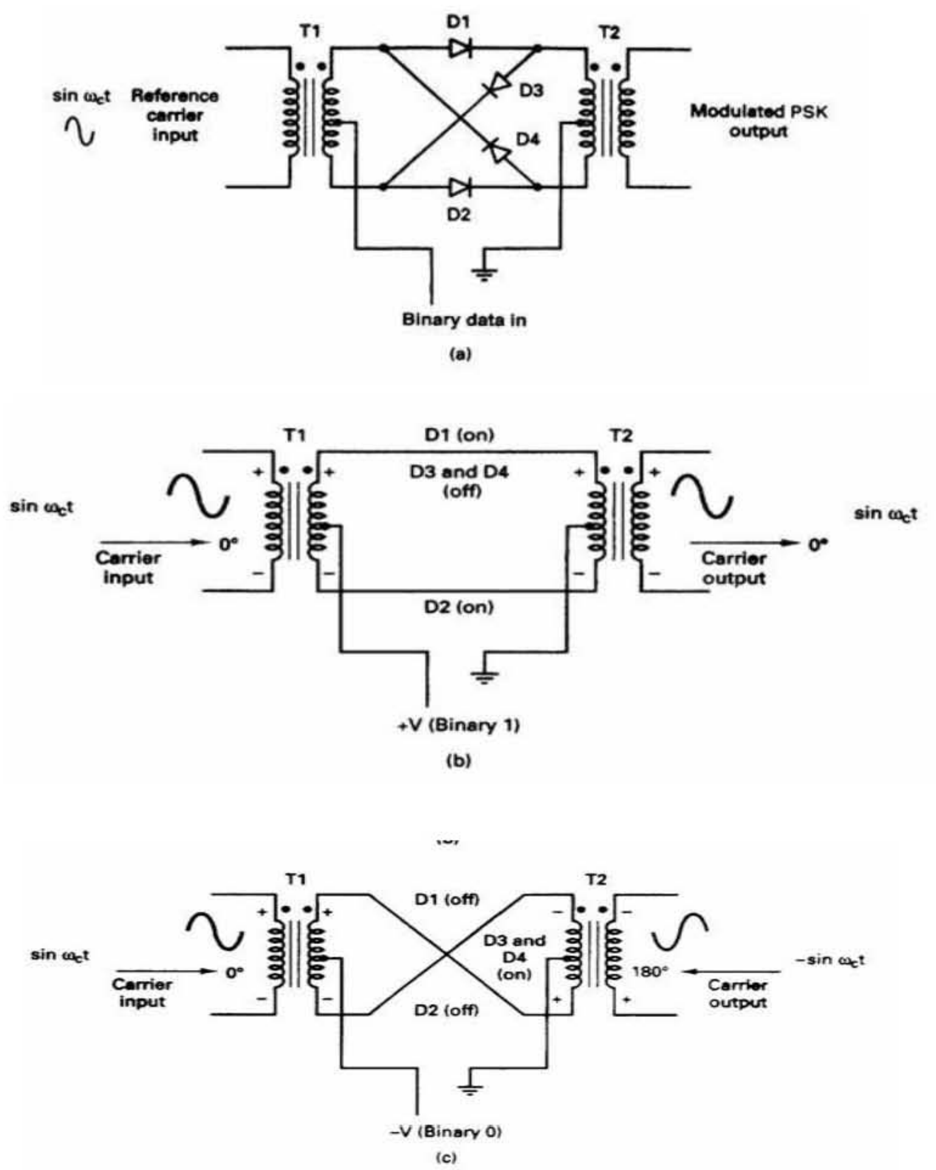
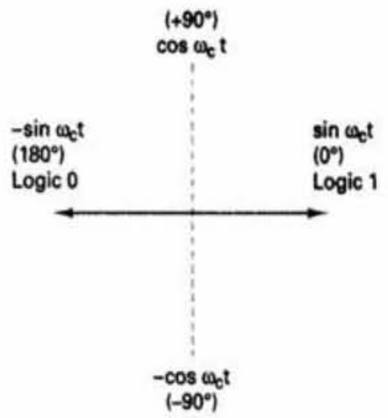


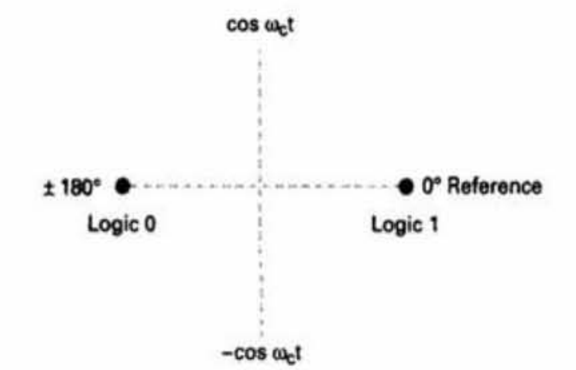
FIGURE 9-13 (a) Balanced ring modulator; (b) logic 1 input; (c) logic 0 input

Binary input	Output phase
Logic 0	180°
Logic 1	0°

(a)



(b)



(c)

FIGURE 2-14 BPSK modulator: (a) truth table; (b) phasor diagram; (c) constellation diagram

BANDWIDTH CONSIDERATIONS OF BPSK:

In a BPSK modulator. the carrier input signal is multiplied by the binary data.

If + 1 V is assigned to a logic 1 and -1 V is assigned to a logic 0, the input carrier (sin $\omega_c t$) is multiplied by either a + or - 1 .

The output signal is either + 1 sin $\omega_c t$ or -1 sin $\omega_c t$ the first represents a signal that is *in phase* with the reference oscillator, the latter a signal that is 180° out of phase with the reference oscillator.Each time the input logic condition changes, the output phase changes.

Mathematically, the output of a BPSK modulator is proportional to

$$\text{BPSK output} = [\sin(2\pi f_a t)] \times [\sin(2\pi f_c t)] \qquad (2.20)$$

where

f_a = maximum fundamental frequency of binary input (hertz)

f_c = reference carrier frequency (hertz)

Solving for the trig identity for the product of two sine functions,

$0.5\cos[2\pi(f_c - f_a)t] - 0.5\cos[2\pi(f_c + f_a)t]$

Thus, the minimum double-sided Nyquist bandwidth *(B) is*

$$\begin{array}{l} f_c + f_a \\ -(f_c + f_a) \end{array} \quad \text{or} \quad \begin{array}{r} f_c + f_a \\ \underline{-f_c + f_a} \\ 2f_a \end{array}$$

and because $f_a = f_b / 2$, where f_b = input bit rate,

where *B is* the minimum double-sided Nyquist bandwidth.

Figure 2-15 shows the output phase-versus-time relationship for a BPSK waveform. Logic 1 input produces an analog output signal with a 0° phase angle, and a logic 0 input produces an analog output signal with a 180° phase angle.

As the binary input shifts between a logic 1 and a logic 0 condition and vice versa, the phase of the BPSK waveform shifts between 0° and 180°, respectively.

BPSK signaling element *(t_s)* is equal to the time of one information bit *(t_b),* which indicates that the bit rate equals the baud.

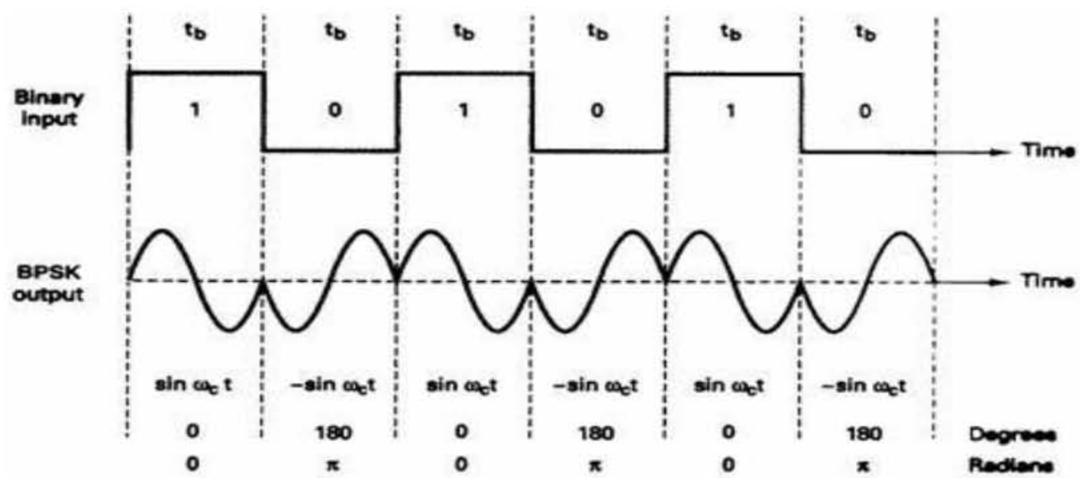


FIGURE 2-15 Output phase-versus-time relationship for a BPSK modulator

Example:

For a BPSK modulator with a carrier frequency of 70 MHz and an input bit rate of 10 Mbps, determine the maximum and minimum upper and lower side frequencies, draw the output spectrum, de-termine the minimum Nyquist bandwidth, and calculate the baud..

Solution

Substituting into Equation 2-20 yields

$$\text{output} = [\sin(2\pi f_a t)] \times [\sin(2\pi f_c t)];\ f_a = f_b / 2 = 5\text{ MHz}$$

$$= [\sin 2\pi(5\text{MHz})t)] \times [\sin 2\pi(70\text{MHz})t)]$$

$$= 0.5\cos[2\pi(70\text{MHz} - 5\text{MHz})t] - 0.5\cos[2\pi(70\text{MHz} + 5\text{MHz})t]$$

lower side frequency upper side frequency

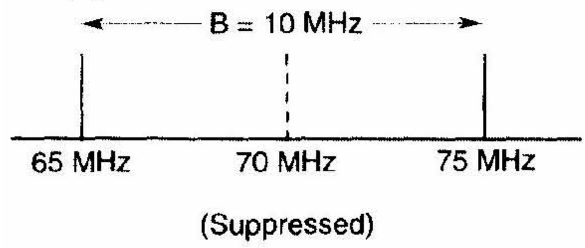
Minimum lower side frequency (LSF):

LSF=70MHz - 5MHz = 65MHz

Maximum upper side frequency (USF):

USF = 70 MHz + 5 MHz = 75 MHz

Therefore, the output spectrum for the worst-case binary input conditions *is* as follows: The minimum Nyquist bandwidth *(B) is*



B = 75 MHz - 65 MHz = 10 MHz

and the baud = f_b or 10 megabaud.

BPSK receiver:.

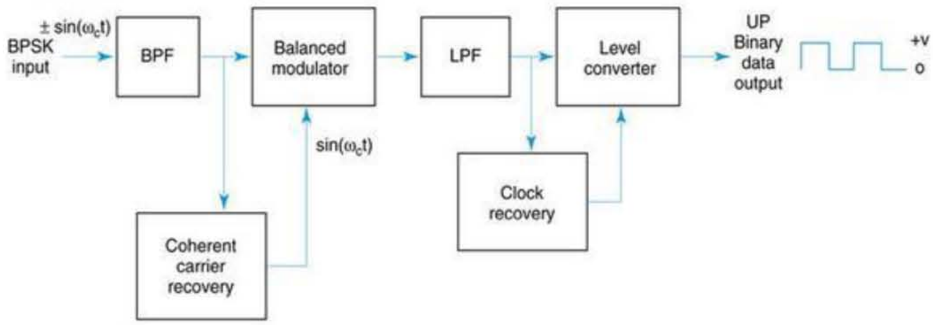
Figure 2-16 shows the block diagram of a BPSK receiver.

The input signal maybe+ sin $\omega_c t$ or - sin $\omega_c t$ .The coherent carrier recovery circuit detects and regenerates a carrier signal that is both frequency and phase coherent with the original transmit carrier.

The balanced modulator is a product detector; the output is the product d the two inputs (the BPSK signal and the recovered carrier).

The low-pass filter (LPF) operates the recovered binary data from the complex demodulated signal.

FIGURE 2-16 Block diagram of a BPSK receiver



Mathematically, the demodulation process is as follows.

For a BPSK input signal of + sin $\omega_c t$ (logic 1), the output of the balanced modulator is

$$\text{output} = (\sin \omega_c t)(\sin \omega_c t) = \sin^2\omega_c t \qquad (2.21)$$

or

$$\sin^2\omega_c t = 0.5(1 - \cos 2\omega_c t) = 0.5 - 0.5\cos 2\omega_c t$$

filtered out

leaving output = + 0.5 V = logic 1

It can be seen that the output of the balanced modulator contains a positive voltage (+[1/2]V) and a cosine wave at twice the carrier frequency (2 $\omega_c t$).

The LPF has a cutoff frequency much lower than 2 $\omega_c t$, and, thus, blocks the second harmonic of the carrier and passes only the positive constant component. A positive voltage represents a demodulated logic 1.

For a BPSK input signal of -sin $\omega_c t$ (logic 0), the output of the balanced modulator is

$$\text{output} = (-\sin \omega_c t)(\sin \omega_c t) = \sin^2\omega_c t$$

or

$$\sin^2\omega_c t = -0.5(1 - \cos 2\omega_c t) = 0.5 + 0.5\cos 2\omega_c t$$

filtered out

leaving

output = - 0.5 V = logic 0

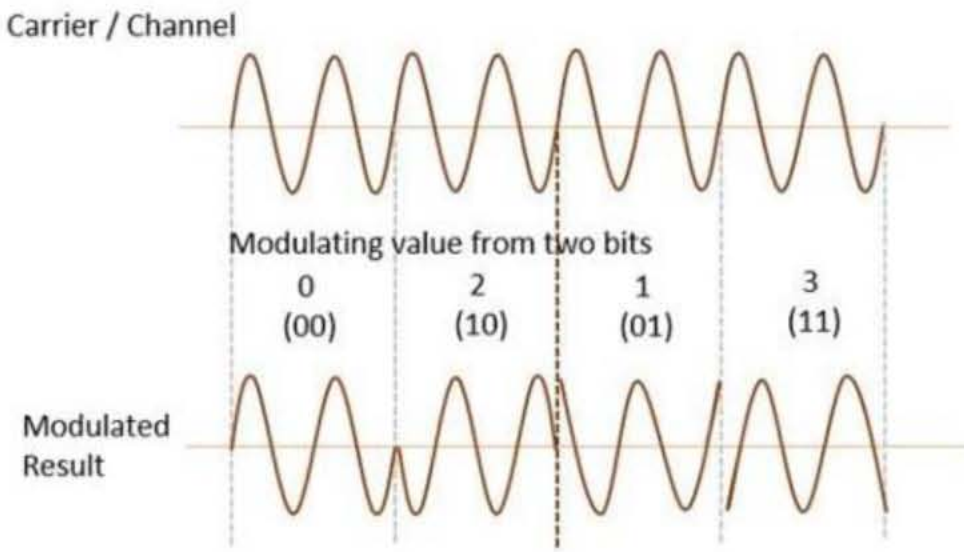
The output of the balanced modulator contains a negative voltage (-[l/2]V) and a cosine wave at twice the carrier frequency ($2\omega_c t$).

Again, the LPF blocks the second harmonic of the carrier and passes only the negative constant component. A negative voltage represents a demodulated logic 0.

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°.

If this kind of techniques are further extended, PSK can be done by eight or sixteen values also, depending upon the requirement. The following figure represents the QPSK waveform for two bits input, which shows the modulated result for different instances of binary inputs.



QPSK is a variation of BPSK, and it is also a DSB-SC (Double Sideband Suppressed Carrier) modulation scheme, which sends two bits of digital information at a time, called as **bigits**.

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.

QPSK transmitter.

A block diagram of a QPSK modulator is shown in Figure 2-17Two bits (a dibit) are clocked into the bit splitter. After both bits have been serially inputted, they are simultaneously parallel outputted.

The I bit modulates a carrier that is in phase with the reference oscillator (hence the name "I" for "in phase" channel), and theQ bit modulate, a carrier that is 90° out of phase.

For a logic 1 = + 1 V and a logic 0= - 1 V, two phases are possible at the output of the I balanced modulator (+sin $\omega_c t$ and - sin $\omega_c t$), and two phases are possible at the output of the Q balanced modulator (+cos $\omega_c t$), and (-cos $\omega_c t$).

When the linear summer combines the two quadrature (90° out of phase) signals, there are four possible resultant phasors given by these expressions: + sin $\omega_c t$ + cos $\omega_c t$, + sin $\omega_c t$ - cos $\omega_c t$, -sin $\omega_c t$ + cos $\omega_c t$, and -sin $\omega_c t$ - cos $\omega_c t$.

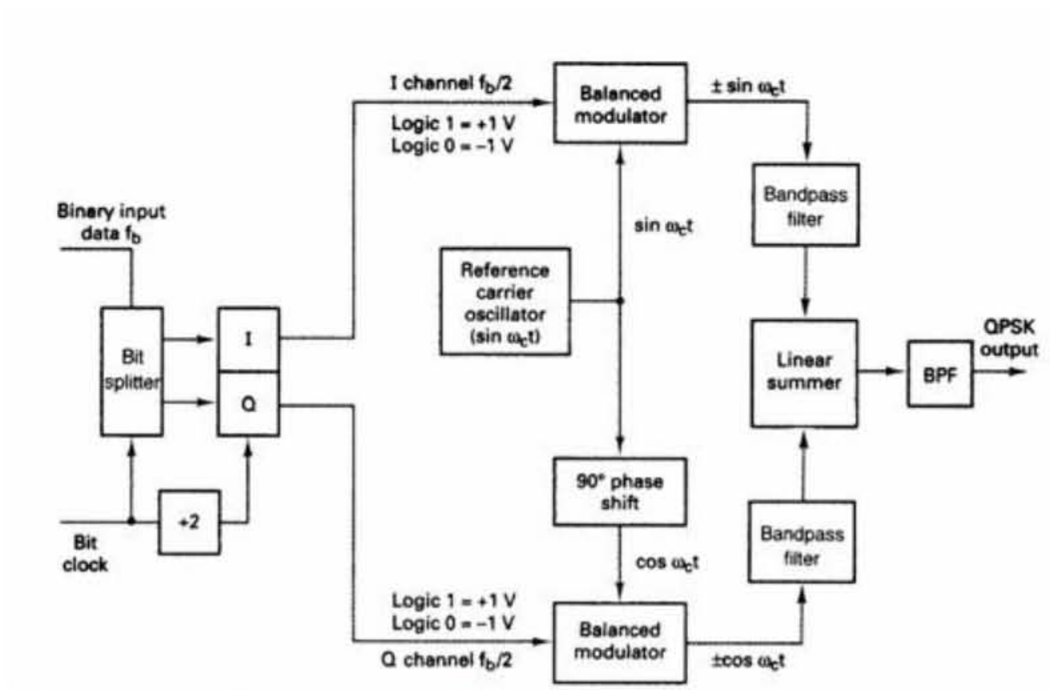


FIGURE 2-17 QPSK modulator

Example:

For the QPSK modulator shown in Figure 2-17, construct the truthtable, phasor diagram, and constellation diagram.

Solution

For a binary data input of Q = O and I= 0, the two inputs to the Ibalanced modulator are -1 and sin $\omega_c t$, and the two inputs to the Q balanced modulator are -1 and cos $\omega_c t$.

Consequently, the outputs are

I balanced modulator =(-1)(sin $\omega_c t$) = -1 sin $\omega_c t$

Q balanced modulator =(-1)(cos $\omega_c t$) = -1 cos $\omega_c t$ and the output of the linear summer is
-1 cos $\omega_c t$ - 1 sin $\omega_c t$ = 1.414 sin($\omega_c t$ - 135°)

For the remaining dibit codes (01, 10, and 11), the procedure is the same. The results are shown in Figure 2-18a.

Binary input		QPSK output phase
Q	I	
0	0	-135°
0	1	-45°
1	0	+135°
1	1	+45°

(a)

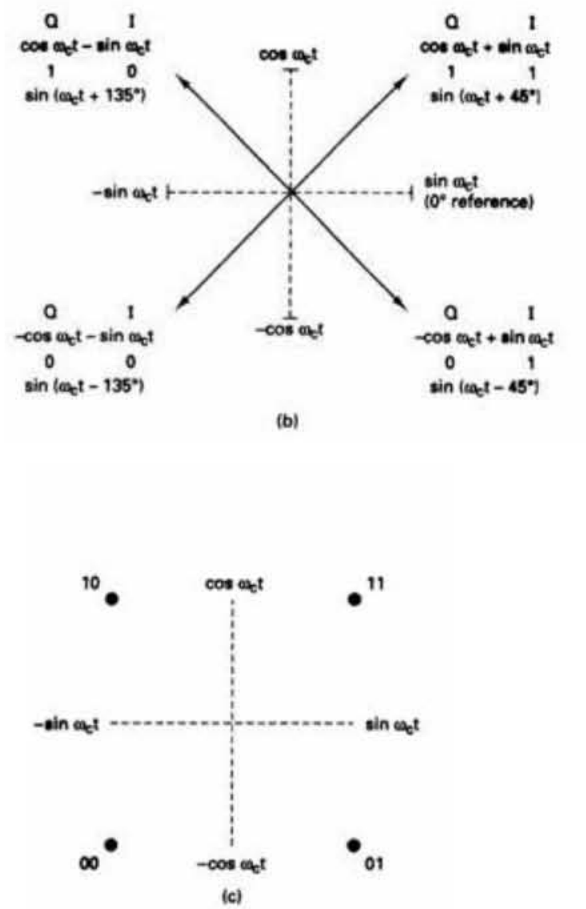


FIGURE 2-18 QPSK modulator: (a) truth table; (b) phasor diagram; (c) constellation diagram

In Figures 2-18b and c, it can be seen that with QPSK each of the four possible output phasors has exactly the same amplitude. Therefore, the binary information must be encoded entirely in the phase of the output signal

Figure 2-18b, it can be seen that the angular separation between any two adjacent phasors in QPSK is 90°.Therefore, a QPSK signal can undergo almost a+45° or -45° shift in phase during transmission and still retain the correct encoded information when demodulated at the receiver.

Figure 2-19 shows the output phase-versus-time relationship for a QPSK modulator.

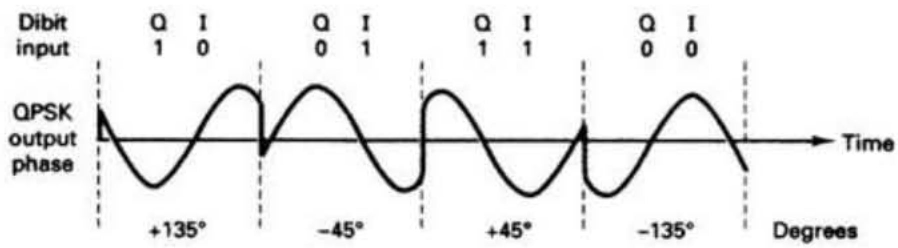


FIGURE 2-19 Output phase-versus-time relationship for a PSK modulator

Bandwidth considerations of QPSK

With QPSK, because the input data are divided into two channels, the bit rate in either the I or the Q channel is equal to one-half of the input data rate ($f_b/2$) (one-half of $f_b/2 = f_b/4$).

QPSK RECEIVER:

The block diagram of a QPSK receiver is shown in Figure 2-21

The power splitter directs the input QPSK signal to the I and Q product detectors and the carrier recovery circuit. The carrier recovery circuit reproduces the original transmit carrier oscillator signal. The recovered carrier must be frequency and phase coherent with the transmit reference carrier. The QPSK signal is demodulated in the I and Q product detectors, which generate the original I and Q data bits. The outputs of the product detectors are fed to the bit combining circuit, where they are converted from parallel I and Q data channels to a single binary output data stream. The incoming QPSK signal may be any one of the four possible output phases shown in Figure 2-18. To illustrate the demodulation process, let the incoming QPSK signal be $-\sin \omega_c t + \cos \omega_c t$. Mathematically, the demodulation process is as follows.

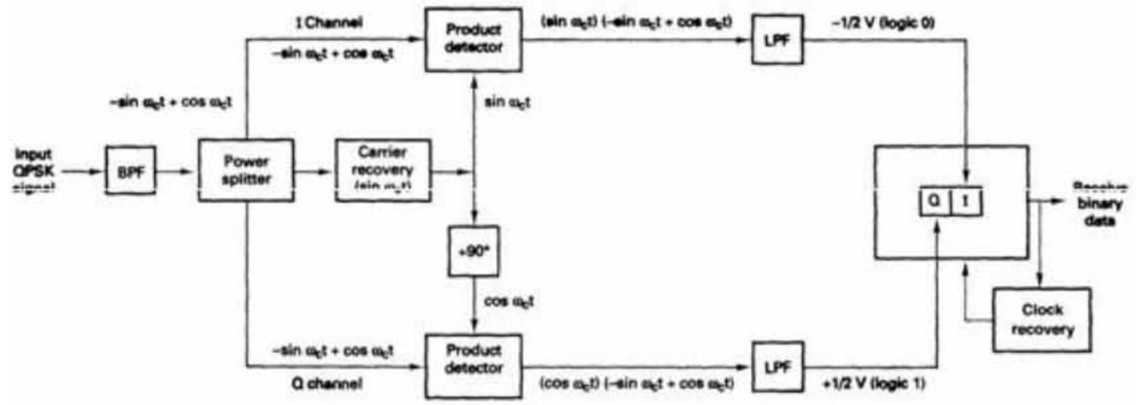


FIGURE 2-21 QPSK receiver

The receive QPSK signal ($-\sin \omega_c t + \cos \omega_c t$) is one of the inputs to the I product detector. The other input is the recovered carrier ($\sin \omega_c t$). The output of the I product detector is

$$\mathrm{I} = \underbrace{(-\sin \omega_c t + \cos \omega_c t)}_{\text{QPSK input signal}} \underbrace{(\sin \omega_c t)}_{\text{carrier}}$$

$$= (-\sin \omega_c t)(\sin \omega_c t) + (\cos \omega_c t)(\sin \omega_c t)$$

$$= -\sin^2 \omega_c t + (\cos \omega_c t)(\sin \omega_c t)$$

$$= -\frac{1}{2}(1 - \cos 2\omega_c t) + \frac{1}{2}\sin(\omega_c + \omega_c)t + \frac{1}{2}\sin(\omega_c - \omega_c)t$$

$$\mathrm{I} = -\frac{1}{2} + \underbrace{\frac{1}{2}\cos 2\omega_c t}_{\text{(filtered out)}} + \frac{1}{2}\sin 2\omega_c t + \underbrace{\frac{1}{2}\sin 0}_{\text{(equals 0)}}$$

$$= -\frac{1}{2}\mathrm{V}\ (\text{logic } 0) \qquad (2.23)$$

Again, the receive QPSK signal ($-\sin \omega_c t + \cos \omega_c t$) is one of the inputs to the Q product detector. The other input is the recovered carrier shifted 90° in phase ($\cos \omega_c t$). The output of the Q product detector is

$$Q = \underbrace{(-\sin \omega_c t + \cos \omega_c t)}_{\text{QPSK input signal}} \underbrace{(\cos \omega_c t)}_{\text{carrier}}$$

$$= \cos^2 \omega_c t - (\sin \omega_c t)(\cos \omega_c t)$$

$$= \frac{1}{2}(1 + \cos 2\omega_c t) - \frac{1}{2}\sin(\omega_c + \omega_c)t - \frac{1}{2}\sin(\omega_c - \omega_c)t$$

$$Q = \frac{1}{2} + \frac{1}{2}\cos 2\omega_c t - \underbrace{\frac{1}{2}\sin 2\omega_c t}_{\text{(filtered out)}} - \underbrace{\frac{1}{2}\sin 0}_{\text{(equals 0)}}$$

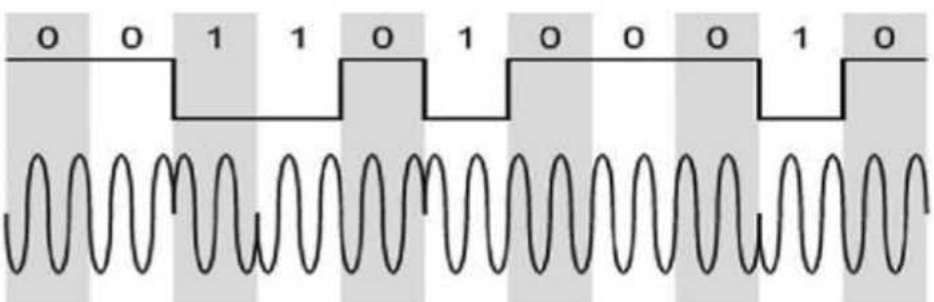
$$= \frac{1}{2}V(\text{logic } 1) \qquad (2.24)$$

The demodulated I and Q bits (0 and 1, respectively) correspond to the constellation diagram and truth table for the QPSK modulator shown in Figure 2-18.

DIFFERENTIAL PHASE SHIFT KEYING (DPSK):

In DPSK (Differential Phase Shift Keying) 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.



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 is continued as it was. If the data is 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.

The word binary represents two-bits. **M** simply 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.

DBPSK TRANSMITTER.:

Figure 2-37a shows a simplified block diagram of a ***differential binary phase-shift keying*** (DBPSK) transmitter. An incoming information bit is XNORed with the preceding bit prior to entering the BPSK modulator (balanced modulator).

For the first data bit, there is no preceding bit with which to compare it. Therefore, an initial reference bit is assumed. Figure 2-37b shows the relationship between the input data, the XNOR output data, and the phase at the output of the balanced modulator. If the initial reference bit is assumed a logic 1, the output from the XNOR circuit is simply the complement of that shown.

In Figure 2-37b, the first data bit is XNORed with the reference bit. If they are the same, the XNOR output is a logic 1; if they are different, the XNOR output is a logic 0. The balanced modulator operates the same as a conventional BPSK modulator; a logic I produces $+\sin \omega_c t$ at the output, and A logic 0 produces $-\sin \omega_c t$ at the output.

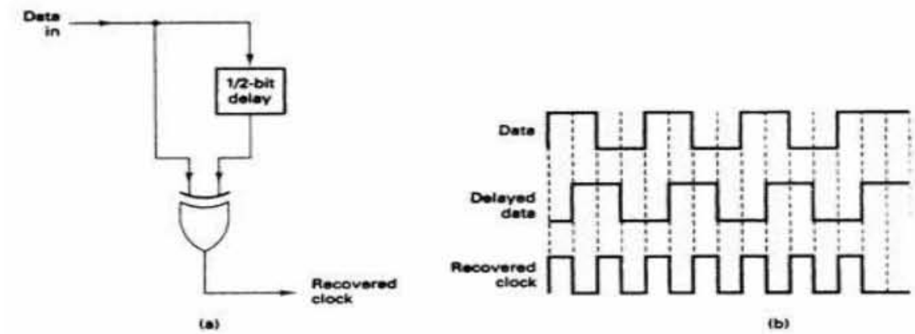


FIGURE 9-40 (a) Clock recovery circuit; (b) timing diagram

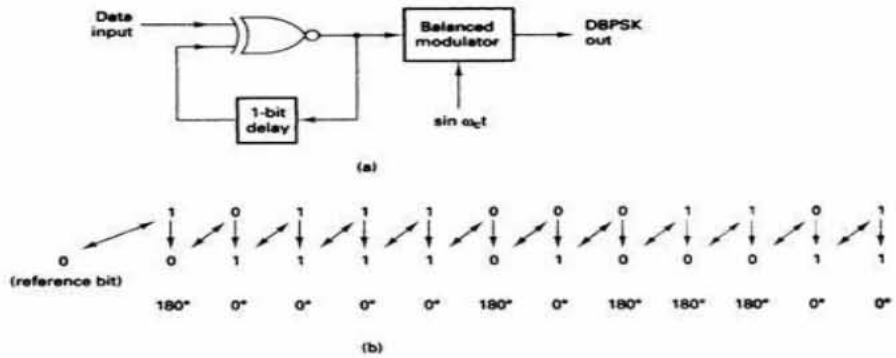


FIGURE 2-37 DBPSK modulator (a) block diagram (b) timing diagram

BPSK RECEIVER:

Figure 9-38 shows the block diagram and timing sequence for a DBPSK receiver. The received signal is delayed by one bit time, then compared with the next signaling element in the balanced modulator. If they are the same. J logic 1(+ voltage) is generated. If they are different, a logic 0 (- voltage) is generated. [f the reference phase is incorrectly assumed, only the first demodulated bit is in error. Differential encoding can be implemented with higher-than-binary digital modulation schemes, although the differential algorithms are much more complicated than for DBPS K.

The primary advantage of DBPSK is the simplicity with which it can be implemented. With DBPSK, no carrier recovery circuit is needed. A disadvantage of DBPSK is, that it requires between 1 dB and 3 dB more signal-to-noise ratio to achieve the same bit error rate as that of absolute PSK.

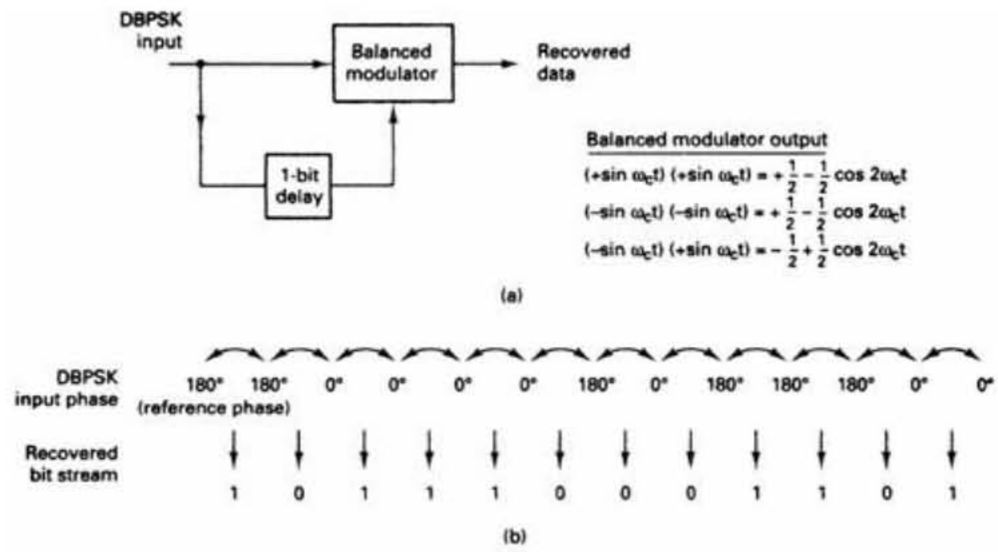


FIGURE 2-38 DBPSK demodulator: (a) block diagram; (b) timing sequence

COHERENT RECEPTION OF FSK:

The coherent demodulator for the coherent FSK signal falls in the general form of coherent demodulators described in Appendix B. The demodulator can be implemented with two correlators as shown in Figure 3.5, where the two reference signals are cos(27r f t) and cos(27r fit). They must be synchronized with the received signal. The receiver is optimum in the sense that it minimizes the error probability for equally likely binary signals. Even though the receiver is rigorously derived in Appendix B, some heuristic explanation here may help understand its operation. When s 1 (t) is transmitted, the upper correlator yields a signal 1 with a positive signal component and a noise component. However, the lower correlator output 12, due to the signals' orthogonality, has only a noise component. Thus the output of the summer is most likely above zero, and the threshold detector will most likely produce a 1. When s2(t) is transmitted, opposite things happen to the two correlators and the threshold detector will most likely produce a 0. However, due to the noise nature that its values range from -00 to m, occasionally the noise amplitude might overpower the signal amplitude, and then detection errors will happen. An alternative to Figure 3.5 is to use just one correlator with the reference signal cos (27r f t) - cos(2s f2t) (Figure 3.6). The correlator in Figure

can be replaced by a matched filter that matches cos(27r fit) - cos(27r f2t) (Figure 3.7). All implementations are equivalent in terms of error performance (see Appendix B). Assuming an AWGN channel, the received signal is

$$r(t) = s_i(t) + n(t), \qquad i = 1, 2$$

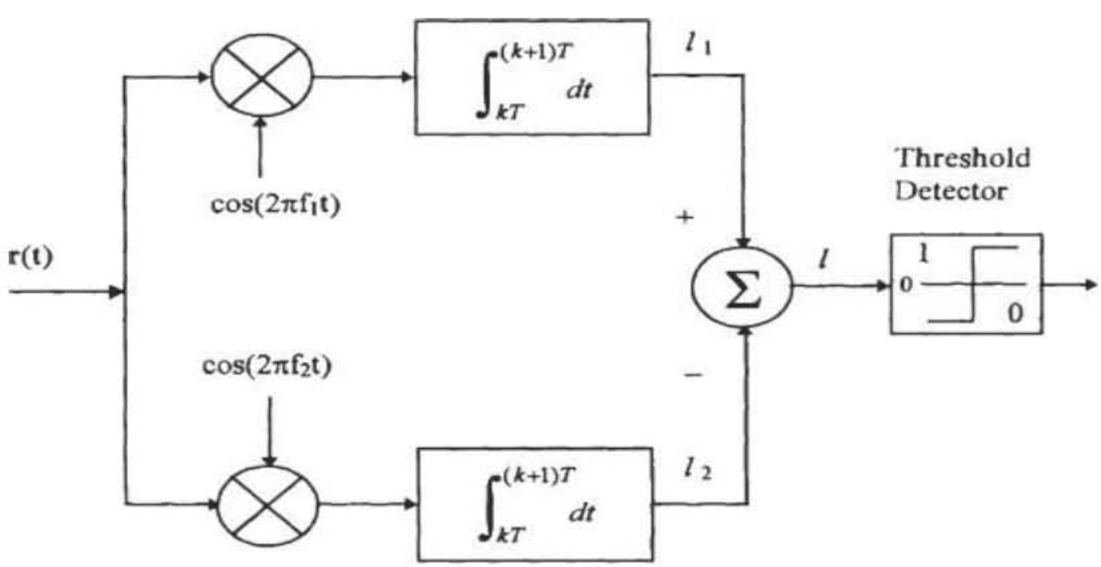
where n(t) is the additive white Gaussian noise with zero mean and a two-sided power spectral density A',/2. From (B.33) the bit error probability for any equally likely binary signals is

$$P_b = Q\left(\sqrt{\frac{E_1 + E_2 - 2\rho_{12}\sqrt{E_1 E_2}}{2N_o}}\right)$$

where No/2 is the two-sided power spectral density of the additive white Gaussian noise. For Sunde's FSK signals El = Ez = Eb, pI2 = 0 (orthogonal). thus the error probability is

$$P_b = Q\left(\sqrt{\frac{E_b}{N_o}}\right)$$

where Eb = A2T/2 is the average bit energy of the FSK signal. The above Pb is plotted in Figure 3.8 where Pb of noncoherently demodulated FSK, whose expression will be given shortly, is also plotted for comparison.



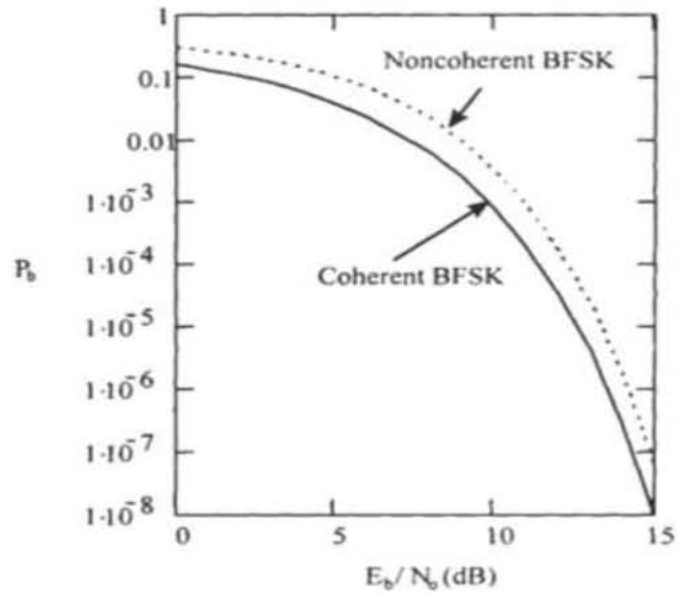


Figure: Pb of coherently and non-coherently demodulated FSK signal.

NONCOHERENT DEMODULATION AND ERROR PERFORMANCE:

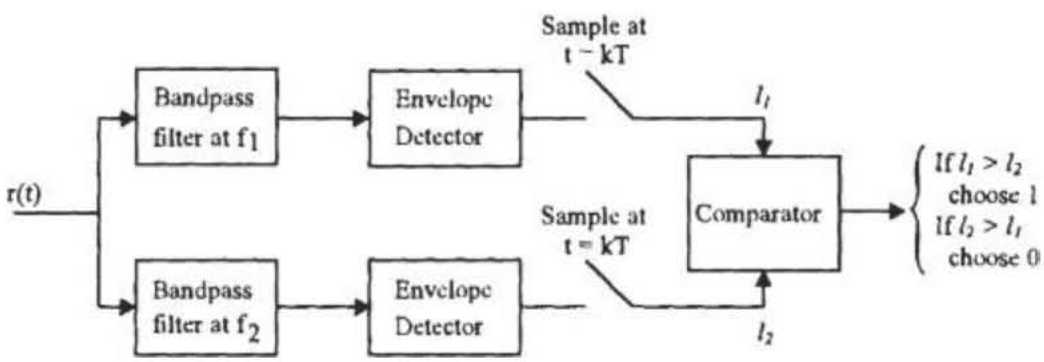
Coherently FSK signals can be noncoherently demodulated to avoid the carrier recovery. Noncoherently generated FSK can only be noncoherently demodulated. We refer to both cases as noncoherent FSK. In both cases the demodulation problem becomes a problem of detecting signals with unknown phases. In Appendix B we have shown that the optimum receiver is a quadrature receiver. It can be implemented using correlators or equivalently, matched filters. Here we assume that the binary noncoherent FSK signals are equally likely and with equal energies. Under these assumptions, the demodulator using correlators is shown in Figure 3.9. Again, like in the coherent case, the optimality of the receiver has been rigorously proved (Appendix B). However, we can easily understand its operation by some heuristic argument as follows. The received signal (ignoring noise for the moment) with an unknown phase can be written as

$$\begin{aligned} s_i(t.\theta) &= A\cos(2\pi f_i t + \theta), \qquad i = 1, 2 \\ &= A\cos\theta\cos 2\pi f_i t - A\sin\theta\sin 2\pi f_i t \end{aligned}$$

The signal consists of an in phase component A cos 8 cos 27r f t and a quadrature component A sin 8 sin 2x f,t sin 0. Thus the signal is partially correlated with cos 2s fit and partiah'y correlated with sin 27r fit. Therefore we use two correlators to collect the signal energy in these two parts. The outputs of the in phase and quadrature correlators will be cos 19 and sin 8, respectively. Depending on the value of the unknown phase 8, these two outputs could be anything in (- 5, y). Fortunately the squared sum of these two signals is not dependent on the unknown phase. That is

$$(\frac{AT}{2}\cos\theta)^2 + (\frac{AT}{2}\sin\theta)^2 = \frac{A^2T^2}{2}$$

This quantity is actually the mean value of the statistics I? when signal si (t) is transmitted and noise is taken into consideration. When si (t) is not transmitted the mean value of 1: is 0. The comparator decides which signal is sent by checking these I?. The matched filter equivalence to Figure 3.9 is shown in Figure 3.10 which has the same error performance. For implementation simplicity we can replace the matched filters by bandpass filters centered at f and fi, respectively (Figure 3.1 1). However, if the bandpass filters are not matched to the FSK signals, degradation to



various extents will result. The bit error probability can be derived using the correlator demodulator (Appendix B). Here we further assume that the FSK signals are orthogonal, then from Appendix B the error probability is

$$P_b = \frac{1}{2}e^{-E_b/2N_o}$$

PART-2

DATATRANSMISSION

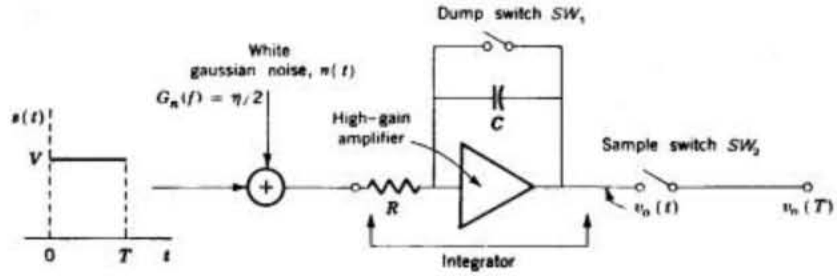
BASE BAND SIGNAL RECEIVER:

Consider that a binary encoded signal consists of a time sequence of voltage levels +V or –V. if there is a guard interval between the bits, the signal forms a sequence of positive and negative pulses. in either case there is no particular interest in preserving the waveform of the signal after reception .we are interested only in knowing within each bit interval whether the transmitted voltage was +V or –V. With noise present, the receives signal and noise together will yield sample values generally different from ±V. In this case, what deduction shall we make from the sample value concerning the transmitted bit?

Suppose that the noise is gaussian and therefore the noise voltage has a probability density which is entirely symmetrical with respect to zero volts. Then the probability that the noise has increased the sample value is the same as the probability that the noise has decreased the sample value. It then seems entirely reasonable that we can do no better than to assume that if the sample value is positive the transmitted level was $+V$, and if the sample value is negative the transmitted level was $-V$. It is, of course, possible that at the sampling time the noise voltage may be of magnitude larger than V and of a polarity opposite to the polarity assigned to the transmitted bit. In this case an error will be made as indicated in Fig. 11.1-1. Here the transmitted bit is represented by the voltage $+V$ which is sustained over an interval T from t_1 to t_2. Noise has been superimposed on the level $+V$ so that the voltage v represents the received signal and noise. If now the sampling should happen to take place at a time $t = t_1 + \Delta t$, an error will have been made.

We can reduce the probability of error by processing the received signal plus noise in such a manner that we are then able to find a sample time where the sample voltage due to the signal is emphasized relative to the sample voltage due to the noise. Such a processer (receiver) is shown in Fig. 11.1-2. The signal input during a bit interval is indicated. As a matter of convenience we have set $t = 0$ at the beginning of the interval. The waveform of the signal $s(t)$ before $t = 0$ and after $t = T$ has not been indicated since, as will appear, the operation of the receiver during each bit interval is independent of the waveform during past and future bit intervals.

The signal $s(t)$ with added white gaussian noise $n(t)$ of power spectral density $\eta/2$ is presented to an integrator. At time $t = 0+$ we require that capacitor C be uncharged. Such a discharged condition may be ensured by a brief closing of switch SW_1 at time $t = 0-$, thus relieving C of any charge it may have acquired during the previous interval. The sample is taken at the output of the integrator by closing this sampling switch SW_2. This sample is taken at the end of the bit interval, at $t = T$. The signal processing indicated in Fig. 11.1-2 is described by the phrase *integrate and dump*, the term *dump* referring to the abrupt discharge of the capacitor after each sampling.



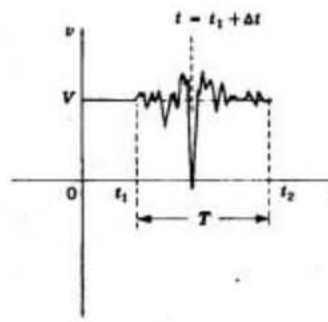


Figure 11.1-2 A receiver for a binary coded signal.

Peak Signal to RMS Noise Output Voltage Ratio

The integrator yields an output which is the integral of its input multiplied by $1/RC$. Using $\tau = RC$, we have

$$v_o(T) = \frac{1}{\tau}\int_0^T [s(t) + n(t)]\, dt = \frac{1}{\tau}\int_0^T s(t)\, dt + \frac{1}{\tau}\int_0^T n(t)\, dt \qquad (11.1\text{-}1)$$

The sample voltage due to the signal is

$$s_o(T) = \frac{1}{\tau}\int_0^T V\, dt = \frac{VT}{\tau} \qquad (11.1\text{-}2)$$

The sample voltage due to the noise is

$$n_o(T) = \frac{1}{\tau}\int_0^T n(t)\, dt \qquad (11.1\text{-}3)$$

This noise-sampling voltage $n_o(T)$ is a gaussian random variable in contrast with $n(t)$, which is a gaussian random process.

The variance of $n_o(T)$ was found in Sec. 7.9 [see Eq. (7.9-17)] to be

$$\sigma_o^2 = \overline{n_o^2(T)} = \frac{\eta T}{2\tau^2} \qquad (11.1\text{-}4)$$

and, as noted in Sec. 7.3, $n_o(T)$ has a gaussian probability density.

The output of the integrator, before the sampling switch, is $v_o(t) = s_o(t) + n_o(t)$. As shown in Fig. 11.1-3a, the signal output $s_o(t)$ is a ramp, in each bit interval, of duration T. At the end of the interval the ramp attains the voltage $s_o(T)$ which is $+VT/\tau$ or $-VT/\tau$, depending on whether the bit is a 1 or a 0. At the end of each interval the switch SW_1 in Fig. 11.1-2 closes momentarily to discharge the capacitor so that $s_o(t)$ drops to zero. The noise $n_o(t)$, shown in Fig. 11.1-3b, also starts each interval with $n_o(0) = 0$ and has the random value $n_o(T)$ at the end of each interval. The sampling switch SW_2 closes briefly just before the closing of SW_1 and hence reads the voltage

$$v_o(T) = s_o(T) + n_o(T) \qquad (11.1\text{-}5)$$

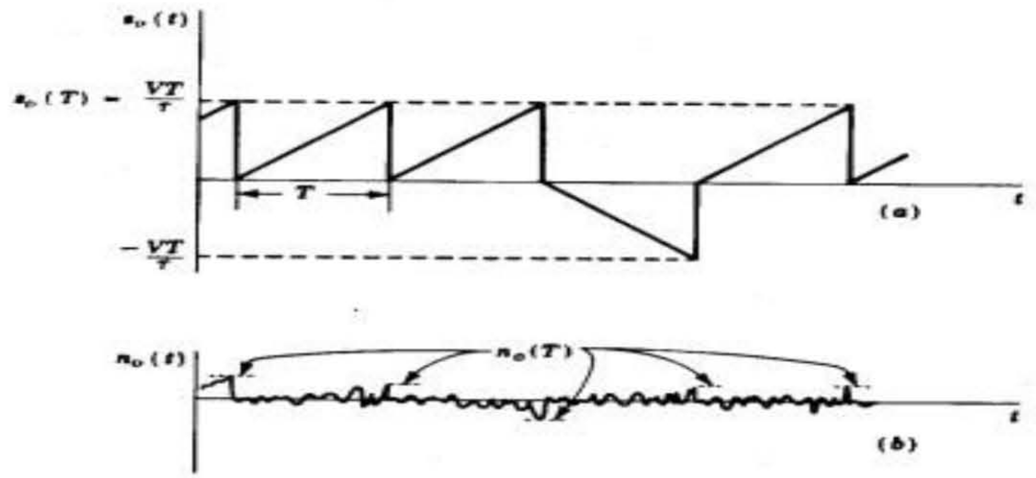


Figure 11.1-3 (*a*) The signal output and (*b*) the noise output of the integrator of Fig. 11.1-2.

We would naturally like the output signal voltage to be as large as possible in comparison with the noise voltage. Hence a figure of merit of interest is the signal-to-noise ratio

$$\frac{[s_o(T)]^2}{[n_o(T)]^2} = \frac{2}{\eta} V^2 T \qquad (11.1\text{-}6)$$

This result is calculated from Eqs. (11.1-2) and (11.1-4). Note that the signal-to-noise ratio increases with increasing bit duration T and that it depends on $V^2 T$ which is the normalized energy of the bit signal. Therefore, a bit represented by a narrow, high amplitude signal and one by a wide, low amplitude signal are equally effective, provided $V^2 T$ is kept constant.

It is instructive to note that the integrator filters the signal and the noise such that the signal voltage increases linearly with time, while the standard deviation (rms value) of the noise increases more slowly, as $\sqrt{T}$. Thus, the integrator enhances the signal relative to the noise, and this enhancement increases with time as shown in Eq. (11.1-6).

PROBABILITY OF ERROR

Since the function of a receiver of a data transmission is to ditinguish the bit 1 from the bit 0 in the presence of noise, a most important charcteristic is the probability that an error will be made in such a determination.we now calculate this error probabilty P_e for the integrate and dump receiver of fig 11.1-2

We have seen that the probability density of the noise sample $n_o(T)$ is gaussian and hence appears as in Fig. 11.2-1. The density is therefore given by

$$f[n_o(T)] = \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}} \tag{11.2-1}$$

where σ_o^2, the variance, is $\sigma_o^2 \equiv \overline{n_o^2(T)}$ given by Eq. (11.1-4). Suppose, then, that during some bit interval the input-signal voltage is held at, say, $-V$. Then, at the sample time, the signal sample voltage is $s_o(T) = -VT/\tau$, while the noise sample is $n_o(T)$. If $n_o(T)$ is positive and larger in magnitude than VT/τ, the total sample voltage $v_o(T) = s_o(T) + n_o(T)$ will be positive. Such a positive sample voltage will result in an error, since as noted earlier, we have instructed the receiver to interpret such a positive sample voltage to mean that the signal voltage was $+V$ during the bit interval. The probability of such a misinterpretation, that is, the probability that $n_o(T) > VT/\tau$, is given by the area of the shaded region in Fig. 11.2-1. The probability of error is, using Eq. (11.2-1),

$$P_e = \int_{VT/\tau}^{\infty} f[n_o(T)]\, dn_o(T) = \int_{VT/\tau}^{\infty} \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}}\, dn_o(T) \tag{11.2-2}$$

Defining $x \equiv n_o(T)/\sqrt{2}\sigma_o$, and using Eq. (11.1-4), Eq. (11.2-2) may be rewritten as

$$P_e = \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{x=V\sqrt{T/\eta}}^{\infty} e^{-x^2}\, dx$$

$$= \frac{1}{2}\,\text{erfc}\left(V\sqrt{\frac{T}{\eta}}\right) = \frac{1}{2}\,\text{erfc}\left(\frac{V^2T}{\eta}\right)^{1/2} = \frac{1}{2}\,\text{erfc}\left(\frac{E_s}{\eta}\right)^{1/2} \tag{11.2-3}$$

in which $E_s = V^2T$ is the signal energy of a bit.

If the signal voltage were held instead at $+V$ during some bit interval, then it is clear from the symmetry of the situation that the probability of error would again be given by P_e in Eq. (11.2-3). Hence Eq. (11.2-3) gives P_e quite generally.

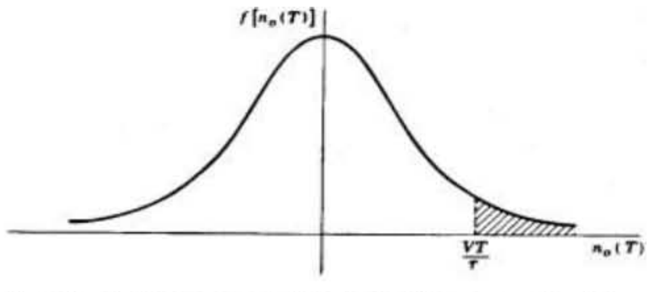


Figure 11.2-1 The gaussian probability density of the noise sample $n_o(T)$.

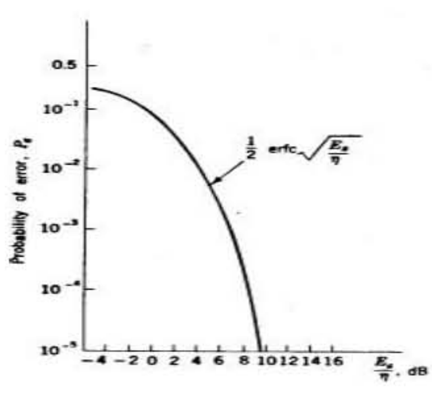


Figure 11.2-2 Variation of P_e versus E_s/η.

The probability of error p_e, as given in eq.(11.2-3),is plotted in fig.11.2-2.note that p_e decreases rapidly as E_s/η increases. The maximum value of p_e is ½.thus ,even if the signal is entirely lost in the noise so that any determination of the receiver is a sheer guess, the receiver cannot bi wrong more than half the time on the average.

THE OPTIMUM FILTER:

In the receiver system of Fig 11.1-2, the signal was passed through a filter(integrator),so that at the sampling time the signal voltage might be emphasized in comparison with the noise voltage. We are naturally led to risk whether the integrator is the optimum filter for the purpose of minimizing the probability of error. We shall find that the received signal contemplated in system of fig 11.1-2 the integrator is indeed the optimum filter. However, before returning specifically to the integrator receiver.

We assume that the received signal is a binary waveform. One binary digit is represented by a signal waveformS_1 (t) which persists for time T, while the4 other bit is represented by the waveform S_2(t) which also lasts for an interval T. For example, in the transmission at baseband, as shown in fig 11.1-2 S_1(t)=+V; for other modulation systems, different waveforms are transmitted. for example for PSK signaling , $S_1(t)=A\cos w_0 t$ and $S_2(t)=-A\cos w_0 t$;while for FSK, $S_1(t)=A\cos(w_{0+\Omega})t$.

As shown in Fig. 11.3-1 the input, which is $s_1(t)$ or $s_2(t)$, is corrupted by the addition of noise $n(t)$. The noise is gaussian and has a spectral density $G(f)$. [In most cases of interest the noise is white, so that $G(f) = \eta/2$. However, we shall assume the more general possibility, since it introduces no complication to do so.] The signal and noise are filtered and then sampled at the end of each bit interval. The output sample is either $v_o(T) = s_{o1}(T) + n_o(T)$ or $v_o(T) = s_{o2}(T) + n_o(T)$. We assume that immediately after each sample, every energy-storing element in the filter has been discharged.

We have already considered in Sec. 2.22, the matter of signal determination in the presence of noise. Thus, we note that in the absence of noise the output sample would be $v_o(T) = s_{o1}(T)$ or $s_{o2}(T)$. When noise is present we have shown that to minimize the probability of error one should assume that $s_1(t)$ has been transmitted if $v_o(T)$ is *closer* to $s_{o1}(T)$ than to $s_{o2}(T)$. Similarly, we assume $s_2(t)$ has been transmitted if $v_o(T)$ is *closer* to $s_{o2}(T)$. The decision boundary is therefore midway between $s_{o1}(T)$ and $s_{o2}(T)$. For example, in the baseband system of Fig. 11.1-2, where $s_{o1}(T) = VT/\tau$ and $s_{o2}(T) = -VT/\tau$, the decision boundary is $v_o(T) = 0$. In general, we shall take the decision boundary to be

$$v_o(T) = \frac{s_{o1}(T) + s_{o2}(T)}{2} \qquad (11.3\text{-}1)$$

The probability of error for this general case may be deduced as an extension of the considerations used in the baseband case. Suppose that $s_{o1}(T) > s_{o2}(T)$ and that $s_2(t)$ was transmitted. If, at the sampling time, the noise $n_o(T)$ is positive and larger in magnitude than the voltage difference $\frac{1}{2}[s_{o1}(T) + s_{o2}(T)] - s_{o2}(T)$, an error will have been made. That is, an error [we decide that $s_1(t)$ is transmitted rather than $s_2(t)$] will result if

$$n_o(T) \geq \frac{s_{o1}(T) - s_{o2}(T)}{2} \qquad (11.3\text{-}2)$$

Hence probability of error is

$$P_e = \int_{[s_{o1}(T) - s_{o2}(T)]/2}^{\infty} \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}}\, dn_o(T) \qquad (11.3\text{-}3)$$

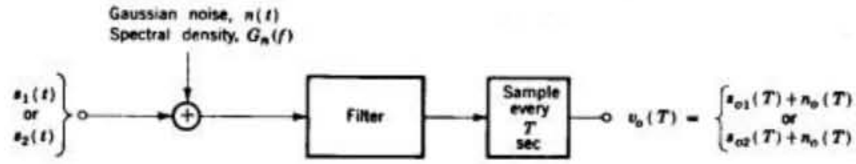


Figure 11.3-1 A receiver for binary coded signalling.

If we make the substitution $x \equiv n_o(T)/\sqrt{2}\sigma_o$, Eq. (11.3-3) becomes

$$P_e = \frac{1}{2}\frac{2}{\sqrt{\pi}} \int_{[s_{o1}(T) - s_{o2}(T)]/2\sqrt{2}\sigma_o}^{\infty} e^{-x^2}\, dx \qquad (11.3\text{-}4a)$$

$$P_e = \frac{1}{2}\,\text{erfc}\left[\frac{s_{o1}(T) - s_{o2}(T)}{2\sqrt{2}\sigma_o}\right] \qquad (11.3\text{-}4b)$$

Note that for the case $s_{o1}(T) = VT/\tau$ and $s_{o2}(T) = -VT/\tau$, and, using Eq. (11.1-4), Eq. (11.3-4*b*) reduces to Eq. (11.2-3) as expected.

The complementary error function is a monotonically decreasing function of its argument. (See Fig. 11.2-2.) Hence, as is to be anticipated, P_e decreases as the difference $s_{o1}(T) - s_{o2}(T)$ becomes larger and as the rms noise voltage σ_o becomes smaller. The optimum filter, then, is the filter which maximizes the ratio

$$\gamma = \frac{s_{o1}(T) - s_{o2}(T)}{\sigma_o} \qquad (11.3\text{-}5)$$

We now calculate the transfer function $H(f)$ of this optimum filter. As a matter of mathematical convenience we shall actually maximize γ^2 rather than γ.

Calculation of the Optimum-Filter Transfer Function $H(f)$

The fundamental requirement we make of a binary encoded data receiver is that it distinguishes the voltages $s_1(t) + n(t)$ and $s_2(t) + n(t)$. We have seen that the ability of the receiver to do so depends on how large a particular receiver can make γ. It is important to note that γ is proportional not to $s_1(t)$ nor to $s_2(t)$, but rather to the *difference* between them. For example, in the baseband system we represented the signals by voltage levels $+V$ and $-V$. But clearly, if our only interest was in distinguishing levels, we would do just as well to use $+2$ volts and 0 volt, or $+8$ volts and $+6$ volts, etc. (The $+V$ and $-V$ levels, however, have the advantage of requiring the least average power to be transmitted.) Hence, while $s_1(t)$ or $s_2(t)$ is the received signal, the signal which is to be compared with the noise, i.e., the signal which is relevant in all our error-probability calculations, is the difference signal

$$p(t) \equiv s_1(t) - s_2(t) \qquad (11.3\text{-}6)$$

Thus, for the purpose of calculating the minimum error probability, we shall assume that the input signal to the optimum filter is $p(t)$. The corresponding *output signal* of the filter is then

$$p_o(t) \equiv s_{o1}(t) - s_{o2}(t) \qquad (11.3\text{-}7)$$

We shall let $P(f)$ and $P_o(f)$ be the Fourier transforms, respectively, of $p(t)$ and $p_o(t)$.

If $H(f)$ is the transfer function of the filter,

$$P_o(f) = H(f)P(f) \tag{11.3-8}$$

and

$$p_o(T) = \int_{-\infty}^{\infty} P_o(f)e^{j2\pi fT}\, df = \int_{-\infty}^{\infty} H(f)P(f)e^{j2\pi fT}\, df \tag{11.3-9}$$

The input noise to the optimum filter is $n(t)$. The output noise is $n_o(t)$ which has a power spectral density $G_{n_o}(f)$ and is related to the power spectral density of the input noise $G_n(f)$ by

$$G_{n_o}(f) = |H(f)|^2 G_n(f) \tag{11.3-10}$$

Using Parseval's theorem (Eq. 1.13-5), we find that the normalized output noise power, i.e., the noise variance σ_o^2, is

$$\sigma_o^2 = \int_{-\infty}^{\infty} G_{n_o}(f)\, df = \int_{-\infty}^{\infty} |H(f)|^2 G_n(f)\, df \tag{11.3-11}$$

From Eqs. (11.3-9) and (11.3-11) we now find that

$$\gamma^2 = \frac{p_o^2(T)}{\sigma_o^2} = \frac{|\int_{-\infty}^{\infty} H(f)P(f)e^{j2\pi Tf}\, df|^2}{\int_{-\infty}^{\infty} |H(f)|^2 G_n(f)\, df} \tag{11.3-12}$$

Equation (11.3-12) is unaltered by the inclusion or deletion of the absolute value sign in the numerator since the quantity within the magnitude sign $p_o(T)$ is a positive real number. The sign has been included, however, in order to allow further development of the equation through the use of the *Schwarz inequality*.

The *Schwarz inequality* states that given arbitrary complex functions $X(f)$ and $Y(f)$ of a common variable f, then

$$\left|\int_{-\infty}^{\infty} X(f)Y(f)\, df\right|^2 \leq \int_{-\infty}^{\infty} |X(f)|^2\, df \int_{-\infty}^{\infty} |Y(f)|^2\, df \tag{11.3-13}$$

The equal sign applies when

$$X(f) = KY^*(f) \tag{11.3-14}$$

where K is an arbitrary constant and $Y^*(f)$ is the complex conjugate of $Y(f)$.

We now apply the Schwarz inequality to Eq. (11.3-12) by making the identification

$$X(f) \equiv \sqrt{G_n(f)}\, H(f) \tag{11.3-15}$$

and

$$Y(f) \equiv \frac{1}{\sqrt{G_n(f)}} P(f)e^{j2\pi Tf} \tag{11.3-16}$$

Using Eqs. (11.3-15) and (11.3-16) and using the Schwarz inequality, Eq. (11.3-13), we may rewrite Eq. (11.3-12) as

$$\frac{p_o^2(T)}{\sigma_o^2} = \frac{|\int_{-\infty}^{\infty} X(f)Y(f)\, df|^2}{\int_{-\infty}^{\infty} |X(f)|^2\, df} \leq \int_{-\infty}^{\infty} |Y(f)|^2\, df \tag{11.3-17}$$

or, using Eq. (11.3-16),

$$\frac{p_o^2(T)}{\sigma_o^2} \le \int_{-\infty}^{\infty} |Y(f)|^2\, df = \int_{-\infty}^{\infty} \frac{|P(f)|^2}{G_n(f)}\, df \tag{11.3-18}$$

The ratio $p_o^2(T)/\sigma_o^2$ will attain its maximum value when the equal sign in Eq. (11.3-18) may be employed as is the case when $X(f) = KY^*(f)$. We then find from Eqs. (11.3-15) and (11.3-16) that the optimum filter which yields such a maximum ratio $p_o^2(T)/\sigma_o^2$ has a transfer function

$$H(f) = K\, \frac{P^*(f)}{G_n(f)}\, e^{-j2\pi fT} \tag{11.3-19}$$

Correspondingly, the maximum ratio is, from Eq. (11.3-18),

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{max} = \int_{-\infty}^{\infty} \frac{|P(f)|^2}{G_n(f)}\, df \tag{11.3-20}$$

In succeeding sections we shall have occasion to apply Eqs. (11.13-19) and (11.13-20) to a number of cases of interest.

11.4 WHITE NOISE: THE MATCHED FILTER

An optimum filter which yields a maximum ratio $p_o^2(T)/\sigma_o^2$ is called a *matched filter* when the input noise is *white*. In this case $G_n(f) = \eta/2$, and Eq. (11.3-19) becomes

$$H(f) = K\, \frac{P^*(f)}{\eta/2}\, e^{-j2\pi fT} \tag{11.4-1}$$

The impulsive response of this filter, i.e., the response of the filter to a unit strength impulse applied at $t = 0$, is

$$h(t) = \mathscr{F}^{-1}[H(f)] = \frac{2K}{\eta} \int_{-\infty}^{\infty} P^*(f) e^{-j2\pi fT} e^{j2\pi ft}\, df \tag{11.4-2a}$$

$$= \frac{2K}{\eta} \int_{-\infty}^{\infty} P^*(f) e^{j2\pi f(t-T)}\, df \tag{11.4-2b}$$

A physically realizable filter will have an impulse response which is real, i.e., not complex. Therefore $h(t) = h^*(t)$. Replacing the right-hand member of Eq. (11.4-2*b*) by its complex conjugate, an operation which leaves the equation unaltered, we have

$$h(t) = \frac{2K}{\eta} \int_{-\infty}^{\infty} P(f) e^{j2\pi f(T-t)}\, df \tag{11.4-3a}$$

$$= \frac{2K}{\eta}\, p(T-t) \tag{11.4-3b}$$

Finally, since $p(t) \equiv s_1(t) - s_2(t)$ [see Eq. (11.3-6)], we have

$$h(t) = \frac{2K}{\eta}\, [s_1(T-t) - s_2(T-t)] \tag{11.4-4}$$

The significance of these results for the matched filter may be more readily appreciated by applying them to a specific example. Consider then, as in Fig. 11.4-1*a*, that $s_1(t)$ is a triangular waveform of duration T, while $s_2(t)$, as shown in Fig. 11.4-1*b*, is of identical form except of reversed polarity. Then $p(t)$ is as shown in Fig. 11.4-1*c*, and $p(-t)$ appears in Fig. 11.4-1*d*. The waveform $p(-t)$ is the waveform $p(t)$ rotated around the axis $t = 0$. Finally, the waveform $p(T - t)$ called for as the impulse response of the filter in Eq. (11.4-3*b*) is this rotated waveform $p(-t)$ translated in the positive t direction by amount T. This last translation ensures that $h(t) = 0$ for $t < 0$ as is required for a *causal* filter.

In general, the impulsive response of the matched filter consists of p(t) rotated about t=0 and then delayed long enough(i.e., a time T) to make the filter realizable. We may note in passing, that any additional delay that a filter might introduce would in no way interfere with the performance of the filter ,for both signal and noise would be delayed by the same amount, and at the sampling time (which would need similarity to be delayed)the ratio of signal to noise would remain unaltered.

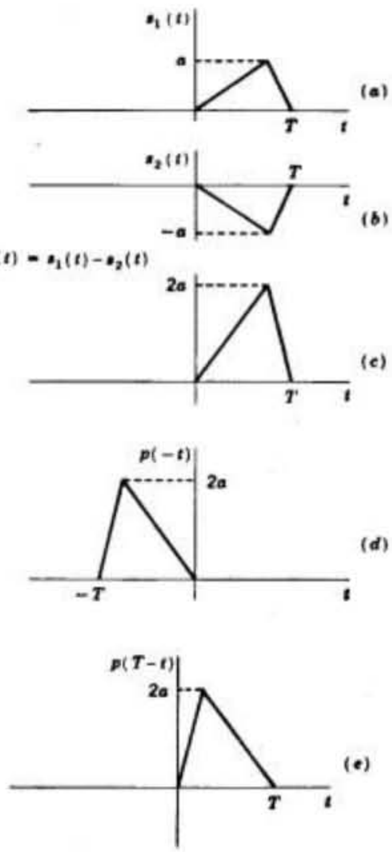


Figure 11.4-1 The signals (*a*) $s_1(t)$, (*b*) $s_2(t)$, and (*c*) $p(t) = s_1(t) - s_2(t)$. (*d*) $p(t)$ rotated about the axis $t = 0$. (*e*) The waveform in (*d*) translated to the right by amount T.

11.5 PROBABILITY OF ERROR OF THE MATCHED FILTER

The probability of error which results when employing a matched filter, may be found by evaluating the maximum signal-to-noise ratio $[p_o^2(T)/\sigma_o^2]_{max}$ given by Eq. (11.3-20). With $G_n(f) = \eta/2$, Eq. (11.3-20) becomes

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{max} = \frac{2}{\eta}\int_{-\infty}^{\infty} |P(f)|^2\, df \qquad (11.5\text{-}1)$$

From parseval's theorem we have

$$\int_{-\infty}^{\infty} |P(f)|^2 \, df = \int_{-\infty}^{\infty} p^2(t) \, dt = \int_0^T p^2(t) \, dt \qquad (11.5\text{-}2)$$

In the last integral in Eq. (11.5-2), the limits take account of the fact that $p(t)$ persists for only a time T. With $p(t) = s_1(t) - s_2(t)$, and using Eq. (11.5-2), we may write Eq. (11.5-1) as

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = \frac{2}{\eta}\int_0^T [s_1(t) - s_2(t)]^2 \, dt \qquad (11.5\text{-}3a)$$

$$= \frac{2}{\eta}\left[\int_0^T s_1^2(t) \, dt + \int_0^T s_2^2(t) \, dt - 2\int_0^T s_1(t)s_2(t) \, dt \right. \qquad (11.5\text{-}3b)$$

$$= \frac{2}{\eta}(E_{s1} + E_{s2} - 2E_{s12}) \qquad (11.5\text{-}3c)$$

Here E_{s1} and E_{s2} are the energies, respectively, in $s_1(t)$ and $s_2(t)$, while E_{s12} is the energy due to the correlation between $s_1(t)$ and $s_2(t)$.

Suppose that we have selected $s_1(t)$, and let $s_1(t)$ have an energy E_{s1}. Then it can be shown that if $s_2(t)$ is to have the *same energy*, the optimum choice of $s_2(t)$ is

$$s_2(t) = -s_1(t) \qquad (11.5\text{-}4)$$

The choice is optimum in that it yields a maximum output signal $p_o^2(T)$ for a given signal energy. Letting $s_2(t) = -s_1(t)$, we find

$$E_{s1} = E_{s2} = -E_{s12} \equiv E_s$$

and Eq. (11.5-3c) becomes

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = \frac{8E_s}{\eta} \qquad (11.5\text{-}5)$$

Rewriting Eq. (11.3-4b) using $p_o(T) = s_{o1}(T) - s_{o2}(T)$, we have

$$P_e = \frac{1}{2}\operatorname{erfc}\left[\frac{p_o(T)}{2\sqrt{2}\,\sigma_o}\right] = \frac{1}{2}\operatorname{erfc}\left[\frac{p_o^2(T)}{8\sigma_o^2}\right]^{1/2} \qquad (11.5\text{-}6)$$

Combining Eq. (11.5-6) with (11.5-5), we find that the minimum error probability $(P_e)_{\min}$ corresponding to a maximum value of $p_o^2(T)/\sigma_o^2$ is

$$(P_e)_{\min} = \frac{1}{2}\operatorname{erfc}\left\{\frac{1}{8}\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max}\right\}^{1/2} \qquad (11.5\text{-}7)$$

$$= \frac{1}{2}\operatorname{erfc}\left(\frac{E_s}{\eta}\right)^{1/2} \qquad (11.5\text{-}8)$$

We note that Eq. (11.5-8) establishes more generally the idea that the error probability depends only on the signal energy and not on the signal waveshape. Previously we had established this point only for signals which had constant voltage levels.

We note also that Eq. (11.5-8) gives $(P_e)_{min}$ for the case of the matched filter and when $s_1(t) = -s_2(t)$. In Sec. 11.2 we considered the case when $s_1(t) = +V$ and $s_2(t) = -V$ and the filter employed was an integrator. There we found [Eq. (11.2-3)] that the result for P_e was identical with $(P_e)_{min}$ given in Eq. (11.5-8). This agreement leads us to suspect that for an input signal where $s_1(t) = +V$ and $s_2(t) = -V$, the integrator is the matched filter. Such is indeed the case. For when we have

$$s_1(t) = V \qquad 0 \le t \le T \tag{11.5-9a}$$

$$s_2(t) = -V \qquad 0 \le t \le T \tag{11.5-9b}$$

the impulse response of the matched filter is, from Eq. (11.4-4),

$$h(t) = \frac{2K}{\eta}[s_1(T-t) - s_2(T-t)] \tag{11.5-10}$$

The quantity $s_1(T-t) - s_2(T-t)$ is a pulse of amplitude $2V$ extending from $t = 0$ to $t = T$ and may be rewritten, with $u(t)$ the unit step,

$$h(t) = \frac{2K}{\eta}(2V)[u(t) - u(t-T)] \tag{11.5-11}$$

The constant factor of proportionality $4KV/\eta$ in the expression for $h(t)$ (that is, the gain of the filter) has no effect on the probability of error since the gain affects signal and noise alike. We may therefore select the coefficient K in Eq. (11.5-11) so that $4KV/\eta = 1$. Then the inverse transform of $h(t)$, that is, the transfer function of the filter, becomes, with s the Laplace transform variable,

$$H(s) = \frac{1}{s} - \frac{e^{-sT}}{s} \tag{11.5-12}$$

The first term in Eq. (11.5-12) represents an integration beginning at $t = 0$, while the second term represents an integration with reversed polarity beginning at $t = T$. The overall response of the matched filter is an integration from $t = 0$ to $t = T$ and a zero response thereafter. In a physical system, as already described, we achieve the effect of a zero response after $t = T$ by sampling at $t = T$, so that so far as the determination of one bit is concerned we ignore the response after $t = T$.

COHERENT RECEPTION: CORRELATION:

We discuss now an alternative type of receiving system which, as we shall see, is identical in performance with the matched filter receiver. Again, as shown in Fig. 11.6-1, the input is a binary data waveform $s_1(t)$ or $s_2(t)$ corrupted by noise $n(t)$. The bit length is T. The received signal plus noise $v_i(t)$ is multiplied by a locally generated waveform $s_1(t) - s_2(t)$. The output of the multiplier is passed through an integrator whose output is sampled at $t = T$. As before, immediately after each sampling, at the beginning of each new bit interval, all energy-storing elements in the integrator are discharged. This type of receiver is called a *correlator*, since we are *correlating* the received signal and noise with the waveform $s_1(t) - s_2(t)$.

The output signal and noise of the correlator shown in Fig. 11.6-1 are

$$s_o(T) = \frac{1}{\tau}\int_0^T s_i(t)[s_1(t) - s_2(t)]\, dt \tag{11.6-1}$$

$$n_o(T) = \frac{1}{\tau}\int_0^T n(t)[s_1(t) - s_2(t)]\, dt$$

(11.6-2)

Where $s_1(t)$ is either $s_1(t)$ or $s_2(t)$,and wthere π is the constant of the integrator(i.e.,the integrator output is $1/\pi$ times the integral of its input).we now compare these outputs with the matched filter outputs.

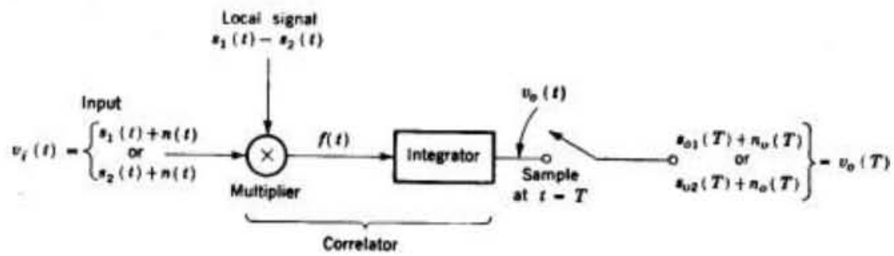


Fig:11.6-1 Coherent system of signal reception

If h(t) is the impulsive response of the matched filter ,then the output of the matched filter *$v_0(t)$* can be found using the convolution integral. we have

$$v_o(t) = \int_{-\infty}^{\infty} v_i(\lambda)h(t-\lambda)\, d\lambda = \int_0^T v_i(\lambda)h(t-\lambda)\, d\lambda$$

(11.6-3)

The limits on the integral have been charged to 0 and T since we are interested in the filter response to a bit which extends only over that interval. Using Eq.(11.4-4) which gives h(t) for the matched filter, we have

$$h(t) = \frac{2K}{\eta}[s_1(T-t) - s_2(T-t)]$$

(11.6-4)

so that $$h(t-\lambda) = \frac{2K}{\eta}[s_1(T-t+\lambda) - s_2(T-t+\lambda)]$$

(11.6-5)

sub 11.6-5 in 11.6-3

$$v_o(t) = \frac{2K}{\eta}\int_0^T v_i(\lambda)[s_1(T-t+\lambda) - s_2(T-t+\lambda)]\, d\lambda$$

(11.6-6)

Since $v_i(\lambda) = s_i(\lambda) + n(\lambda)$, and $v_o(t) = s_o(t) + n_o(t)$, setting $t = T$ yields

$$s_o(T) = \frac{2K}{\eta}\int_0^T s_i(\lambda)[s_1(\lambda) - s_2(\lambda)]\, d\lambda$$

(11.6-7)

where $s_i(\lambda)$ is equal to $s_1(\lambda)$ or $s_2(\lambda)$. Similarly we find that

$$n_o(T) = \frac{2K}{\eta}\int_0^T n(\lambda)[s_1(\lambda) - s_2(\lambda)]\, d\lambda$$

(11.6-8)

Thus $s_0(T)$ and $n_0(T)$, as calculated from eqs.(11.6-1) and (11.6-2) for the correlation receiver, and as calculated from eqs.(11.6-7) and (11.6-8) for the matched filter receiver, are identical .hence the performances of the two systems are identical. The matched filter and the correlator are not simply two distinct, independent techniques which happens to yield the same result. In fact they are two techniques of synthesizing the optimum filter h(t)

5.13.1 Error Probability of ASK

In Amplitude Shift Keying (ASK), some number of carrier cycles are transmitted to send '1' and no signal is transmitted for binary '0'. Thus,

$$\text{Binary '1'} \Rightarrow x_1(t) = \sqrt{2P_s}\cos(2\pi f_0 t) \text{ and}$$

$$\text{Binary '0'} \Rightarrow x_2(t) = 0 \text{ (i.e. no signal)} \quad \text{... (5.13.1)}$$

Here P_s is the normalized power of the signal in 1Ω load. i.e. power $P_s = \frac{A^2}{2}$. Hence $A = \sqrt{2P_s}$. Therefore in above equation for $x_1(t)$ amplitude 'A' is replaced by $\sqrt{2P_s}$.

We know that the probability of error of the optimum filter is given as,

$$P_e = \frac{1}{2} erfc\left\{\frac{x_{01}(T) - x_{02}(T)}{2\sqrt{2}\sigma}\right\} \quad \text{... (5.13.2)}$$

$$\text{Here } \left[\frac{x_{01}(T) - x_{02}(T)}{\sigma}\right]^2_{max} = \int_{-\infty}^{\infty} \frac{|X(f)|^2}{S_{ni}(f)} df$$

The above equations can be applied to matched filter when we consider white Gaussian noise. The power spectral density of white Gaussian noise is given as,

$$S_{ni}(f) = \frac{N_0}{2}$$

Putting this value of $S_{ni}(f)$ in above equations we get,

$$\left[\frac{x_{01}(T) - x_{02}(T)}{\sigma}\right]^2_{max} = \int_{-\infty}^{\infty} \frac{|X(f)|^2}{\frac{N_0}{2}} df$$

$$= \frac{2}{N_0}\int_{-\infty}^{\infty} |X(f)|^2 df \quad \text{... (5.13.3)}$$

Parseval's power theorem states that,

$$\int_{-\infty}^{\infty} |X(f)|^2 df = \int_{-\infty}^{\infty} x^2(t)\, dt$$

Hence equation 5.13.3 becomes,

$$\left[\frac{x_{01}(T) - x_{02}(T)}{\sigma}\right]^2_{max} = \frac{2}{N_0}\int_{-\infty}^{\infty} x^2(t)\, dt$$

We know that $x(t)$ is present from 0 to T. Hence limits in above equation can be changed as follows :

$$\left[\frac{x_{01}(T)-x_{02}(T)}{\sigma}\right]^2_{max} = \frac{2}{N_0}\int_0^T x^2(t)\,dt \quad \text{... (5.13.4)}$$

We know that $x(t) = x_1(t) - x_2(t)$. For ASK $x_2(t)$ is zero, hence $x(t) = x_1(t)$. Hence above equation becomes,

$$\left[\frac{x_{01}(T)-x_{02}(T)}{\sigma}\right]^2_{max} = \frac{2}{N_0}\int_0^T x_1^2(t)\,dt$$

Putting equation of $x_1(t)$ from equation 5.13.1 in above equation we get,

$$\left[\frac{x_{01}(T)-x_{02}(T)}{\sigma}\right]^2_{max} = \frac{2}{N_0}\int_0^T \left[\sqrt{2P_s}\cos(2\pi f_0 t)\right]^2 dt$$

$$= \frac{4P_s}{N_0}\int_0^T \cos^2(2\pi f_0 t)\,dt$$

We know that $\cos^2\theta = \frac{1+\cos 2\theta}{2}$. Here applying this formula to $\cos^2(2\pi f_0 t)$ we get,

$$\left[\frac{x_{01}(T)-x_{02}(T)}{\sigma}\right]^2_{max} = \frac{4P_s}{N_0}\int_0^T \frac{1+\cos 4\pi f_0 t}{2}\,dt$$

$$= \frac{4P_s}{N_0}\cdot\frac{1}{2}\left\{\int_0^T dt + \int_0^T \cos 4\pi f_0 t\,dt\right\}$$

$$= \frac{2P_s}{N_0}\left\{[t]_0^T + \left[\frac{\sin 4\pi f_0 t}{4\pi f_0}\right]_0^T\right\}$$

$$= \frac{2P_s}{N_0}\left\{T + \frac{\sin 4\pi f_0 T}{4\pi f_0}\right\} \quad \text{... (5.13.5)}$$

We know that T is the bit period and in this one bit period, the carrier has integer number of cycles. Thus the product $f_0 T$ is an integer. This is illustrated in Fig. 5.13.1

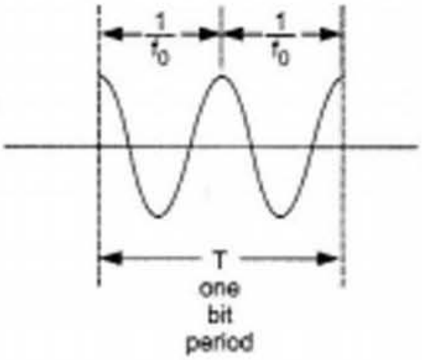


Fig. 5.13.1 In one bit period T, the carrier completes its two cycles. The carrier has frequency f_0. From figure we can write,

$$T = \frac{1}{f_0} + \frac{1}{f_0}$$

i.e. $$T = \frac{2}{f_0}$$

$\therefore$ $f_0 T = 2$ (integer no. of cycles)

As shown in above figure, the carrier completes two cycles in one bit duration. Hence

$$f_0 T = 2$$

Therefore, in general if carrier completes 'k' number of cycles, then,

$$f_0 T = k \qquad \text{(Here } k \text{ is an integer)}$$

Therefore the sine term in equation 5.13.5 becomes, $\sin 4\pi k$ and k is integer.

For all integer values of k, $\sin 4\pi k = 0$. Hence equation 5.13.5 becomes,

$$\left[\frac{x_{01}(T) - x_{02}(T)}{\sigma}\right]^2_{max} = \frac{2P_s T}{N_0} \qquad \text{... (5.13.6)}$$

$$\therefore \quad \left[\frac{x_{01}(T) - x_{02}(T)}{\sigma}\right]_{max} = \sqrt{\frac{2P_s T}{N_0}} \qquad \text{... (5.13.7)}$$

Putting this value in equation 5.13.2 we get error probability of ASK using matched filter detection as,

$$P_e = \frac{1}{2} erfc\left\{\frac{1}{2\sqrt{2}} \cdot \sqrt{\frac{2P_s T}{N_0}}\right\} = \frac{1}{2} erfc \sqrt{\frac{P_s T}{4N_0}}$$

Here $P_s T = E$, i.e. energy of one bit hence above equation becomes,

$$\textbf{Error probability of ASK : } P_e = \frac{1}{2} erfc \sqrt{\frac{E}{4N_0}} \qquad \text{... (5.13.8)}$$

This is the expression for error probability of ASK using matched filter detection.

Error Probability of Binary FSK

The observation vector **x** has two elements x_1 and x_2 that are defined by, respectively,

$$x_1 = \int_0^{T_b} x(t)\phi_1(t)\, dt \tag{6.92}$$

$$x_2 = \int_0^{T_b} x(t)\phi_2(t)\, dt \tag{6.93}$$

where $x(t)$ is the received signal, the form of which depends on which symbol was transmitted. Given that symbol 1 was transmitted, $x(t)$ equals $s_1(t) + w(t)$, where $w(t)$ is the sample function of a white Gaussian noise process of zero mean and power spectral density $N_0/2$. If, on the other hand, symbol 0 was transmitted, $x(t)$ equals $s_2(t) + w(t)$.

Now, applying the decision rule of Equation (5.59), we find that the observation space is partitioned into two decision regions, labeled Z_1 and Z_2 in Figure 6.25. The decision boundary, separating region Z_1 from region Z_2 is the perpendicular bisector of the line joining the two message points. The receiver decides in favor of symbol 1 if the received signal point represented by the observation vector **x** falls inside region Z_1. This occurs when $x_1 > x_2$. If, on the other hand, we have $x_1 < x_2$, the received signal point falls inside region Z_2, and the receiver decides in favor of symbol 0. On the decision boundary, we have $x_1 = x_2$, in which case the receiver makes a random guess in favor of symbol 1 or 0.

Define a new Gaussian random variable Y whose sample value y is equal to the difference between x_1 and x_2; that is,

$$y = x_1 - x_2 \tag{6.94}$$

The mean value of the random variable Y depends on which binary symbol was transmitted. Given that symbol 1 was transmitted, the Gaussian random variables X_1 and X_2, whose sample values are denoted by x_1 and x_2, have mean values equal to $\sqrt{E_b}$ and zero, respectively. Correspondingly, the conditional mean of the random variable Y, given that symbol 1 was transmitted, is

$$\begin{aligned} E[Y|1] &= E[X_1|1] - E[X_2|1] \\ &= +\sqrt{E_b} \end{aligned} \tag{6.95}$$

On the other hand, given that symbol 0 was transmitted, the random variables X_1 and X_2 have mean values equal to zero and $\sqrt{E_b}$, respectively. Correspondingly, the conditional mean of the random variable Y, given that symbol 0 was transmitted, is

$$E[Y|0] = E[X_1|0] - E[X_2|0] = -\sqrt{E_b} \tag{6.96}$$

The variance of the random variable Y is independent of which binary symbol was transmitted. Since the random variables X_1 and X_2 are statistically independent, each with a variance equal to $N_0/2$, it follows that

$$\text{var}[Y] = \text{var}[X_1] + \text{var}[X_2] = N_0 \tag{6.97}$$

Suppose we know that symbol 0 was transmitted. The conditional probability density function of the random variable Y is then given by

$$f_Y(y|0) = \frac{1}{\sqrt{2\pi N_0}} \exp\left[-\frac{(y + \sqrt{E_b})^2}{2N_0}\right] \tag{6.98}$$

Since the condition $x_1 > x_2$, or equivalently, $y > 0$, corresponds to the receiver making a decision in favor of symbol 1, we deduce that the conditional probability of error, given that symbol 0 was transmitted, is

$$\begin{aligned} p_{10} &= P(y > 0|\text{symbol 0 was sent}) \\ &= \int_0^\infty f_Y(y|0)\,dy \\ &= \frac{1}{\sqrt{2\pi N_0}} \int_0^\infty \exp\left[-\frac{(y + \sqrt{E_b})^2}{2N_0}\right] dy \end{aligned} \tag{6.99}$$

$$\frac{y + \sqrt{E_b}}{\sqrt{2N_0}} = z \tag{6.100}$$

Then, changing the variable of integration from y to z, we may rewrite Equation (6.99) as follows:

$$\begin{aligned} p_{10} &= \frac{1}{\sqrt{\pi}} \int_{\sqrt{E_b/2N_0}}^\infty \exp(-z^2)\,dz \\ &= \frac{1}{2}\,\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) \end{aligned} \tag{6.101}$$

Similarly, we may show the p_{01}, the conditional probability of error given that symbol 1 was transmitted, has the same value as in Equation (6.101). Accordingly, averaging p_{10} and p_{01}, we find that the *average probability of bit error* or, equivalently, the *bit error rate for coherent binary FSK* is (assuming equiprobable symbols)

$$P_e = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) \tag{6.102}$$

Comparing Equations (6.20) and (6.102), we see that, in a coherent binary FSK system, we have to double the *bit energy-to-noise density ratio*, E_b/N_0, to maintain the same bit error rate as in a coherent binary PSK system. This result is in perfect accord with the signal-space diagrams of Figures 6.3 and 6.25, where we see that in a binary PSK system the Euclidean distance between the two message points is equal to $2\sqrt{E_b}$, whereas in a binary FSK system the corresponding distance is $\sqrt{2E_b}$. For a prescribed E_b, the minimum distance d_{min} in binary PSK is therefore $\sqrt{2}$ times that in binary FSK. Recall from Chapter 5 that the probability of error decreases exponentially as d^2_{min}, hence the difference between the formulas of Equations (6.20) and (6.102).

Error Probability of QPSK

In a coherent QPSK system, the received signal $x(t)$ is defined by

$$x(t) = s_i(t) + w(t), \qquad \begin{cases} 0 \le t \le T \\ i = 1, 2, 3, 4 \end{cases} \tag{6.28}$$

where $w(t)$ is the sample function of a white Gaussian noise process of zero mean and power spectral density $N_0/2$. Correspondingly, the observation vector $\mathbf{x}$ has two elements, x_1 and x_2, defined by

$$\begin{aligned} x_1 &= \int_0^T x(t)\phi_1(t)\,dt \\ &= \sqrt{E}\cos\left[(2i-1)\frac{\pi}{4}\right] + w_1 \\ &= \pm\sqrt{\frac{E}{2}} + w_1 \end{aligned} \tag{6.29}$$

$$
\begin{aligned}
x_2 &= \int_0^T x(t)\phi_2(t)\, dt \\
&= -\sqrt{E}\sin\left[(2i-1)\frac{\pi}{4}\right] + w_2 \\
&= \mp\sqrt{\frac{E}{2}} + w_2
\end{aligned}
\tag{6.30}
$$

Thus the observable elements x_1 and x_2 are sample values of independent Gaussian random variables with mean values equal to $\pm\sqrt{E/2}$ and $\mp\sqrt{E/2}$, respectively, and with a common variance equal to $N_0/2$.

The decision rule is now simply to decide that $s_1(t)$ was transmitted if the received signal point associated with the observation vector $\mathbf{x}$ falls inside region Z_1, decide that $s_2(t)$ was transmitted if the received signal point falls inside region Z_2, and so on. An erroneous decision will be made if, for example, signal $s_4(t)$ is transmitted but the noise $w(t)$ is such that the received signal point falls outside region Z_4.

To calculate the average probability of symbol error, we note from Equation (6.24) that a coherent QPSK system is in fact equivalent to two coherent binary PSK systems working in parallel and using two carriers that are in phase quadrature; this is merely a statement of the quadrature-carrier multiplexing property of coherent QPSK. The in-phase channel output x_1 and the quadrature channel output x_2 (i.e., the two elements of the observation vector $\mathbf{x}$) may be viewed as the individual outputs of the two coherent binary PSK systems. Thus, according to Equations (6.29) and (6.30), these two binary PSK systems may be characterized as follows:

- The signal energy per bit is $E/2$.
- The noise spectral density is $N_0/2$.

Hence, using Equation (6.20) for the average probability of bit error of a coherent binary PSK system, we may now state that the average probability of bit error in *each* channel of the coherent QPSK system is

$$
\begin{aligned}
P' &= \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E/2}{N_0}}\right) \\
&= \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right)
\end{aligned}
\tag{6.31}
$$

Another important point to note is that the bit errors in the in-phase and quadrature channels of the coherent QPSK system are statistically independent. The in-phase channel makes a decision on one of the two bits constituting a symbol (dibit) of the QPSK signal, and the quadrature channel takes care of the other bit. Accordingly, the *average probability of a correct decision* resulting from the combined action of the two channels working together is

$$
\begin{aligned}
P_c &= (1 - P')^2 \\
&= \left[1 - \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right)\right]^2 \\
&= 1 - \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) + \frac{1}{4}\operatorname{erfc}^2\left(\sqrt{\frac{E}{2N_0}}\right)
\end{aligned}
\tag{6.32}
$$

The average probability of symbol error for coherent QPSK is therefore

$$
\begin{aligned}
P_e &= 1 - P_c \\
&= \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) - \frac{1}{4}\operatorname{erfc}^2\left(\sqrt{\frac{E}{2N_0}}\right)
\end{aligned}
\tag{6.33}
$$

In the region where $(E/2N_0) \gg 1$, we may ignore the quadratic term on the right-hand side of Equation (6.33), so we approximate the formula for the average probability of symbol error for coherent QPSK as

$$P_e \simeq \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) \tag{6.34}$$

The formula of Equation (6.34) may also be derived in another insightful way, using the signal-space diagram of Figure 6.6. Since the four message points of this diagram are circularly symmetric with respect to the origin, we may apply Equation (5.92), reproduced here in the form

$$P_e \le \frac{1}{2}\sum_{\substack{k=1 \\ k \ne i}}^{4} \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \qquad \text{for all } i \tag{6.35}$$

Consider, for example, message point m_1 (corresponding to dibit 10) chosen as the transmitted message point. The message points m_2 and m_4 (corresponding to dibits 00 and 11) are the *closest* to m_1. From Figure 6.6 we readily find that m_1 is equidistant from m_2 and m_4 in a Euclidean sense, as shown by

$$d_{12} = d_{14} = \sqrt{2E}$$

Assuming that E/N_0 is large enough to ignore the contribution of the most distant message point m_3 (corresponding to dibit 01) relative to m_1, we find that the use of Equation (6.35) yields an approximate expression for P_e that is the same as Equation (6.34). Note that in mistaking either m_2 or m_4 for m_1, a single bit error is made; on the other hand, in mistaking m_3 for m_1, two bit errors are made. For a high enough E/N_0, the likelihood of both bits of a symbol being in error is much less than a single bit, which is a further justification for ignoring m_3 in calculating P_e when m_1 is sent.

In a QPSK system, we note that since there are two bits per symbol, the transmitted signal energy per symbol is twice the signal energy per bit, as shown by

$$E = 2E_b \tag{6.36}$$

Thus expressing the average probability of symbol error in terms of the ratio E_b/N_0, we may write

$$P_e \simeq \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{6.37}$$

With Gray encoding used for the incoming symbols, we find from Equations (6.31) and (6.36) that the *bit error rate* of QPSK is exactly

$$\text{BER} = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{6.38}$$

We may therefore state that a coherent QPSK system achieves the same average probability of bit error as a coherent binary PSK system for the same bit rate and the same E_b/N_0, but uses only half the channel bandwidth. Stated in a different way, for the same E_b/N_0 and therefore the same average probability of bit error, a coherent QPSK system transmits information at twice the bit rate of a coherent binary PSK system for the same channel bandwidth. For a prescribed performance, QPSK uses channel bandwidth better than binary PSK, which explains the preferred use of QPSK over binary PSK in practice.

ERROR PROBABILITY OF BINARY PSK:

To realize a rule for making a decision in favor of symbol 1 or symbol 0,we partition the signal space into two regions:

- The set of points closest to message point 1 at $+\sqrt{E_b}$.
- The set of points closest to message point 2 at $-\sqrt{E_b}$.

This is accomplished by constructing the midpoint of the line joining these two message points, and then marking off the appropriate decision regions. In Figure 6.3 these decision regions are marked Z_1 and Z_2, according to the message point around which they are constructed.

The decision rule is now simply to decide that signal $s_1(t)$ (i.e., binary symbol 1) was transmitted if the received signal point falls in region Z_1, and decide that signal $s_2(t)$ (i.e., binary symbol 0) was transmitted if the received signal point falls in region Z_2. Two kinds of erroneous decisions may, however, be made. Signal $s_2(t)$ is transmitted, but the noise is such that the received signal point falls inside region Z_1 and so the receiver decides in favor of signal $s_1(t)$. Alternatively, signal $s_1(t)$ is transmitted, but the noise is such that the received signal point falls inside region Z_2 and so the receiver decides in favor of signal $s_2(t)$.

To calculate the probability of making an error of the first kind, we note from Figure 6.3 that the decision region associated with symbol 1 or signal $s_1(t)$ is described by

$$Z_1 \colon 0 < x_1 < \infty$$

where the observable element x_1 is related to the received signal $x(t)$ by

$$x_1 = \int_0^{T_b} x(t)\phi_1(t)\, dt \tag{6.15}$$

The conditional probability density function of random variable X_1, given that symbol 0 [i.e., signal $s_2(t)$] was transmitted, is defined by

$$\begin{aligned} f_{X_1}(x_1|0) &= \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_1 - s_{21})^2\right] \\ &= \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_1 + \sqrt{E_b})^2\right] \end{aligned} \tag{6.16}$$

The conditional probability of the receiver deciding in favor of symbol 1, given that symbol 0 was transmitted, is therefore

$$\begin{aligned} p_{10} &= \int_0^{\infty} f_{X_1}(x_1|0)\, dx_1 \\ &= \frac{1}{\sqrt{\pi N_0}} \int_0^{\infty} \exp\left[-\frac{1}{N_0}(x_1 + \sqrt{E_b})^2\right] dx_1 \end{aligned} \tag{6.17}$$

Putting

$$z = \frac{1}{\sqrt{N_0}}(x_1 + \sqrt{E_b}) \tag{6.18}$$

and changing the variable of integration from x_1 to z, we may rewrite Equation (6.17) in the compact form

$$\begin{aligned} p_{10} &= \frac{1}{\sqrt{\pi}} \int_{\sqrt{E_b/N_0}}^{\infty} \exp(-z^2)\, dz \\ &= \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \end{aligned} \tag{6.19}$$

where erfc(·) is the complementary error function.

Thus, averaging the conditional error probabilities p_{10} and p_{01}, we find that the *average probability of symbol error* or, equivalently, the *bit error rate for coherent binary PSK* is (assuming equiprobable symbols)

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{6.20}$$

As we increase the transmitted signal energy per bit, E_b, for a specified noise spectral density N_0, the message points corresponding to symbols 1 and 0 move further apart, and the average probability of error P_e is correspondingly reduced in accordance with Equation (6.20), which is intuitively satisfying.