UNIT I ANALOG COMMUNICATION

1.1 NOISE:

Noise is an unwanted electrical signal which gets added on a transmitted signal when it is travelling towards the receiver. Electrical noise is defined as any undesired electrical energy. For Example: In audio recording any unwanted electrical signals that fall within the audio frequency band of 0 khz to 15khz will interface with the music and therefore considered as noise.

Noise figure is a figure of merit and used to indicate how much the signal to noise ratio gets degraded as a signal passes through a series of circuits.

Noise can be divided into two general categories:

(i) Correlated Noise : Implies a relationship between the signal and the noise.
(ii) Uncorrelated Noise : It is present all the time whether there is a signal or not.

Uncorrelated Noise:

Uncorrelated can be divided into two general categories: (i) External noise and (ii) Internal noise.

1.2 EXTERNAL NOISE, INTERNAL NOISE:

External Noise:

It is a Noise generated outside the device or circuit. There are three primary sources of external noise.

(i) Atmospheric ,
(ii) Extra terrestrial and
(ii) Manmade noise.

Extraterrestrial Noise consists of electrical signals that originate from outside earths atmosphere and is therefore also called as deep space noise. This noise originates from the milky way , other galaxies and the sun.

Extraterrestrial noise is subdivided into two categories.: (i) Solar and (ii) Cosmic.

Internal Noise: It is the noise caused by electrical interference generated within a device or circuit.

There are three primary kinds of internally generated noise are:

(i) Thermal.
(ii) Shot ,Transits time.

1.3 INTRODUCTION TO COMMUNICATION SYSTEM:

Communication is the process of establishing connection (or link) between two points for information exchange.

The Science of Communications involving long distances is called Telecommunication (the world Tele standing for long distance)

The Two basic types of communication systems are

(i) Analog.
(ii) Digital.

In Analog Systems: Both the information and the carrier are analog signals.

In Digital Systems: The digital pulses are transferred between two or more points in a communication system.

Analog communication:

The modulation systems or techniques in which one of the characteristics of the carrier is changed in proportion with the instantaneous value of modulating signal is called analog communication system.

Advantages of Analog communications

- Transmitters and Receivers are simple
- Low bandwidth requirement
- FDM can be used

Disadvantages of analog communication

- Noise affects the signal quality
- It is not possible to separate noise and signal
- Repeaters can’t be used between transmitters and receivers
- Coding is not possible
- It is not suitable for the transmission of secret information

General Communication Systems:

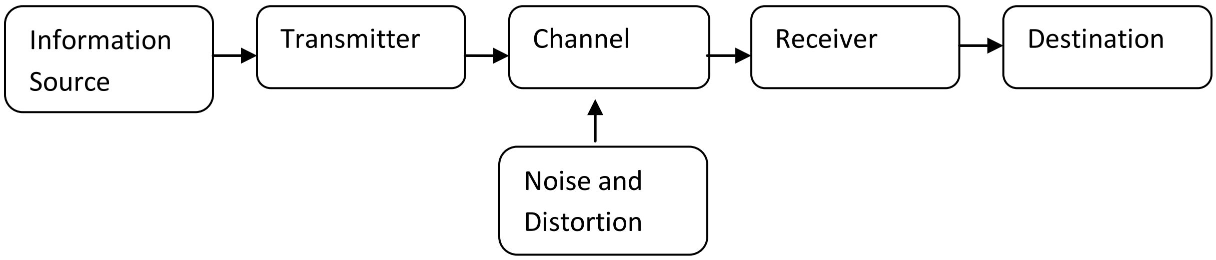


Fig: Block diagram of a general communication system

Drawbacks of Baseband Transmission (without Modulation)

- Excessively large antenna heights.
- Signals get mixed up.
- Short range of communication.
- Multiplexing is not possible.
- Poor quality of reception.

The above drawbacks can be overcome by means of modulation techniques:

1.4 Modulation, Types, Need for Modulation:

Modulation is the changing characteristics of the carrier signal with respect to the instantaneous change in message signal.

Needs for modulation: In order to carry the low frequency message signal to a longer distance, the high frequency carrier signal is combined with it.

a) Reduction in antenna height

b) Long distance communication

c) Ease of radiation

d) Multiplexing

e) Improve the quality of reception

f) Avoid mixing up of other signals

Frequency modulation: Frequency Modulation is the changing frequency of the carrier signal with respect to the instantaneous change in message signal.

Phase modulation: Phase Modulation is defined as changing the phase of the carrier signal with respect to the instantaneous change in message signal.

Deviation ratio: Deviation ratio is the worst case modulation index and is equal to the maximum peak frequency deviation divided by the maximum modulating signal frequency. Mathematically the deviation ratio is DR= f (max)/fm(max).

Amplitude modulation: Amplitude Modulation is defined as changing the amplitude of the carrier signal with respect to the instantaneous change in message signal.

Carson's rule: Carson's rule states that the bandwidth required to transmit an angle modulated wave as twice the sum of the peak frequency deviation and the highest modulating signal frequency. Mathematically carson's rule is B=2(f +fm) Hz.

Modulation index: It is defined as ratio of amplitude of the message signal to the amplitude of the carrier signal. m=Em/Ec.

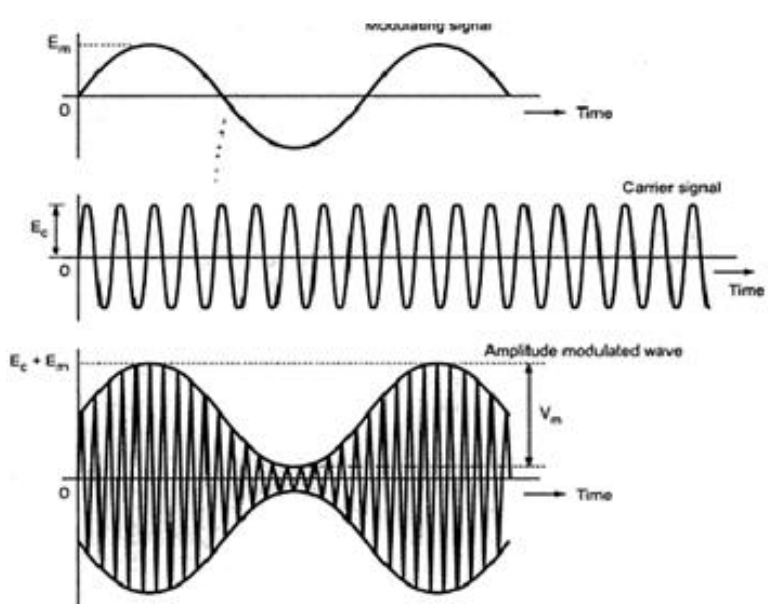
Percentage modulation: It is the percentage change in the amplitude of the output wave when the carrier is acted on by a modulating signal. M=(Em/Ec)*100

1.5 THEORY OF AMPLITUDE MODULATION

Amplitude Modulation

Amplitude Modulation is the changing the amplitude of the carrier signal with respect to the instantaneous change in message signal.

The amplitude modulated wave form, its envelope and its frequency spectrum and bandwidth. Fig (a) Sinosoidal modulation signal (b)High frequency carrier (c) AM signal.



Let us represent the modulating signal by e_m and it is given as,

$$e_m = E_m \sin \omega_m t \qquad \text{... (1.2.1)}$$

and carrier signal can be represented by e_c as,

$$e_c = E_c \sin \omega_c t \qquad \text{... (1.2.2)}$$

Here E_m is maximum amplitude of modulating signal

E_c is maximum amplitude of carrier signal

ω_m is frequency of modulating signal

and ω_c is frequency of carrier signal.

Using the above mathematical expressions for modulating and carrier signals, we can create a new mathematical expression for the complete modulated wave. It is given as,

$$E_{AM} = E_c + e_m$$

$$= E_c + E_m \sin \omega_m t \quad \text{by putting } e_m \text{ from equation (1.2.1)}$$

$\therefore$ The instantaneous value of the amplitude modulated wave can be given as,

$$e_{AM} = E_{AM} \sin \theta$$

$$= E_{AM} \sin \omega_c t$$

$$\therefore \quad \boxed{e_{AM} = (E_c + E_m \sin \omega_m t) \sin \omega_c t} \qquad \text{... (1.2.3)}$$

This is an equation of AM wave.

Modulation Index and Percent Modulation

The ratio of maximum amplitude of modulating signal to maximum amplitude of carrier signal is called modulation index. i.e.,

$$\boxed{\text{Modulation index, } m = \frac{E_m}{E_c}} \qquad \text{... (1.2.4)}$$

Value of E_m must be less than value of E_c to avoid any distortion in the modulated signal. Hence maximum value of modulation index will be equal to 1 when $E_m = E_c$. Minimum value will be zero. If modulation index is higher than 1, then it is called *over modulation*. Data is lost in such case. When modulation index is expressed in percentage, it is also called percentage modulation.

Calculation of modulation index from AM waveform :

Fig. 1.2.2 shows the AM waveform. This is also called time domain representation of AM signal.

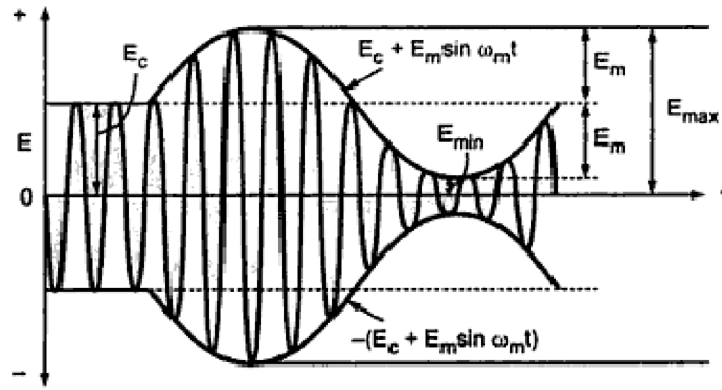


Fig. 1.2.2 AM wave

It is clear from the above signal that the modulating signal rides upon the carrier signal. From above figure we can write,

$$E_m = \frac{E_{max} - E_{min}}{2} \quad \text{... (1.2.5)}$$

and

$$E_c = E_{max} - E_m \quad \text{... (1.2.6)}$$

$$= E_{max} - \frac{E_{max} - E_{min}}{2} \text{ by putting for } E_m \text{ from equation (1.2.5)}$$

$$= \frac{E_{max} + E_{min}}{2} \quad \text{... (1.2.7)}$$

Taking the ratio of equation (1.2.5) and above equation,

$$m = \frac{E_m}{E_c} = \frac{\frac{E_{max} - E_{min}}{2}}{\frac{E_{max} + E_{min}}{2}}$$

$$\therefore \quad \boxed{m = \frac{E_{max} - E_{min}}{E_{max} + E_{min}}} \quad \text{... (1.2.8)}$$

This equation gives the technique of calculating modulation index from AM wave.

Frequency Spectrum and Bandwidth

The modulated carrier has new signals at different frequencies, called side frequencies or sidebands. They occur above and below the carrier frequency.

i.e.

$$f_{USB} = f_c + f_m$$

$$f_{LSB} = f_c - f_m$$

Here f_c is carrier frequency and

f_m is modulating signal frequency

f_{LSB} is lower sideband frequency

Consider the expression of AM wave given by equation (1.2.3), i.e.,

$$e_{AM} = (E_c + E_m \sin \omega_m t) \sin \omega_c t \qquad \text{... (1.2.9)}$$

We know that $m = \frac{E_m}{E_c}$ from equation (1.2.4). Hence we have $E_m = m E_c$. Putting this value of E_m in above equation we get,

$$\begin{aligned} e_{AM} &= (E_c + m E_c \sin \omega_m t) \sin \omega_c t \\ &= E_c (1 + m \sin \omega_m t) \sin \omega_c t \\ &= E_c \sin \omega_c t + m E_c \sin \omega_m t \sin \omega_c t \qquad \text{... (1.2.10)} \end{aligned}$$

We know that $\sin(A)\sin(B) = \frac{1}{2}\cos(A-B) - \frac{1}{2}\cos(A+B)$. Applying this result to last term in above equation we get,

$$\begin{aligned} e_{AM} &= E_c \sin \omega_c t + \frac{m E_c}{2} \cos(\omega_c - \omega_m) t \\ &\quad - \frac{m E_c}{2} \cos(\omega_c + \omega_m) t \qquad \text{... (1.2.11)} \end{aligned}$$

In the above equation, the first term represents unmodulated carrier, the second term represents lower sideband and last term represents upper sideband. Note that $\omega_c = 2\pi f_c$ and $\omega_m = 2\pi f_m$. Hence above equation can also be written as,

$$\begin{aligned} e_{AM} &= E_c \sin 2\pi f_c t + \frac{m E_c}{2} \cos 2\pi (f_c - f_m) t \\ &\quad - \frac{m E_c}{2} \cos 2\pi (f_c + f_m) t \qquad \text{... (1.2.12)} \\ &= E_c \sin 2\pi f_c t + \frac{m E_c}{2} \cos 2\pi f_{LSB} t + \frac{m E_c}{2} \cos 2\pi f_{USB} t \qquad \text{... (1.2.13)} \end{aligned}$$

From this equation we can prepare the frequency spectrum of AM wave as shown below in Fig. 1.2.3.

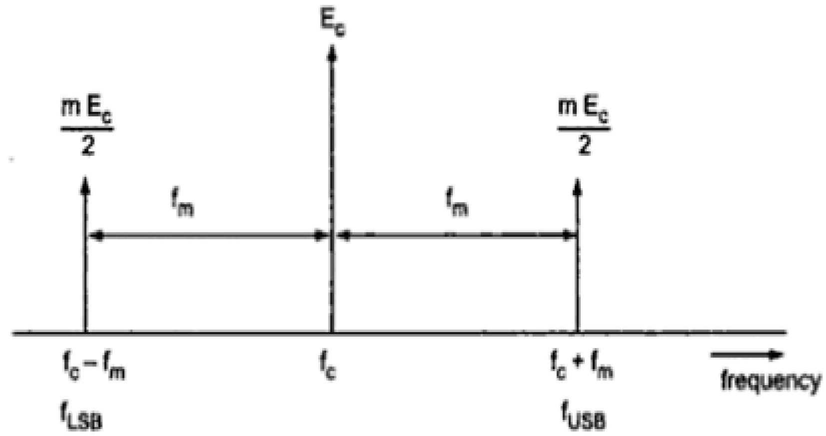


Fig:1.2.3 : Frequency domain Representation of AM Wave

This contains full carrier and both the sidebands, hence it is also called Double Sideband Full Carrier (DSBFC) system. We will be discussing this system, its modulation circuits and transmitters next, in this section.

We know that bandwidth of the signal can be obtained by taking the difference between highest and lowest frequencies. From above figure we can obtain bandwidth of AM wave as,

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

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

$$\therefore \quad \boxed{BW = 2 f_m} \quad \ldots (1.2.14)$$

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

<u>Amplitude Modulation of Power distribution</u>:

AM Power Distribution:

AM signal has three components : Unmodulated carrier, lower sideband and upper sideband. Hence total power of AM wave is the sum of carrier power P_c and powers in the two sidebands P_{USB} and P_{LSB}. i.e.,

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

$$= \frac{E_{carr}^2}{R} + \frac{E_{LSB}^2}{R} + \frac{E_{USB}^2}{R} \quad \ldots (1.2.15)$$

Here all the three voltages are rms values and R is characteristic impedence of antenna in which the power is dissipated. The carrier power is,

$$P_c = \frac{E_{carr}^2}{R} = \frac{\left(E_c / \sqrt{2}\right)^2}{R}$$

$$= \frac{E_c^2}{2R} \qquad \text{... (1.2.16)}$$

The power of upper and lower sidebands is same. i.e.,

$$P_{LSB} = P_{USB} = \frac{E_{SB}^2}{R}$$ Here E_{SB} is rms voltage of sidebands.

From equation (1.2.13) we know that the peak amplitude of both the sidebands is $\frac{m E_c}{2}$. Hence,

$$E_{SB} = \frac{m E_c / 2}{\sqrt{2}}$$

$$\therefore \qquad P_{LSB} = P_{USB} = \left(\frac{m E_c / 2}{\sqrt{2}}\right)^2 \times \frac{1}{R}$$

$$= \frac{m^2 E_c^2}{8R} \qquad \text{... (1.2.17)}$$

Hence the total power (equation 1.2.15) becomes,

$$P_{Total} = \frac{E_c^2}{2R} + \frac{m^2 E_c^2}{8R} + \frac{m^2 E_c^2}{8R}$$

$$= \frac{E_c^2}{2R}\left[1 + \frac{m^2}{4} + \frac{m^2}{4}\right]$$

$$\boxed{P_{Total} = P_c\left(1 + \frac{m^2}{2}\right)} \qquad \text{... (1.2.19)}$$

$$\frac{P_{Total}}{P_c} = 1 + \frac{m^2}{2} \qquad \text{... (1.2.20)}$$

This equation relates total power of AM wave to carrier power, Maximum Value of modulation index, m=1 to avoid distortion. At this value of modulation index, Ptotal = 1.5 Pc. From the above equation we have

$$\frac{m^2}{2} = \frac{P_{Total}}{P_c} - 1$$

$$\therefore \quad m = \sqrt{2\left(\frac{P_{total}}{P_c} - 1\right)}$$

Example Problems:

An audio frequency signal $10 \sin 2\pi \times 500\, t$ is used to amplitude modulate a carrier of $50 \sin 2\pi \times 10^5\, t$. Calculate

(i) Modulation index

(ii) Sideband frequencies

(iii) Amplitude of each sideband frequencies

(iv) Bandwidth required

(v) Total power delivered to the load of $600\,\Omega$.

Solution : (i) The given modulating signal is $e_m = 10 \sin 2\pi \times 500 t$. Hence, $E_m = 10$. The given carrier signal is $e_c = 50 \sin 2\pi \times 10^5\, t$, hence, $E_c = 50$. Therefore modulation index will be,

$$m = \frac{E_m}{E_c} = \frac{10}{50} = 0.2 \quad \text{or} \quad 20\%$$

(ii) From the given equations,

$$\omega_m = 2\pi \times 500, \qquad \text{Hence } f_m = 500\, Hz$$

And $$\omega_c = 2\pi \times 10^5, \qquad \text{Hence } f_c = 10^5\, Hz \text{ or } 100 \text{ kHz}$$

We know that $$f_{USB} = f_c + f_m = 100\, kHz + 500\, Hz = 100.5\, kHz$$

and $$P_{LSB} = f_c - f_m = 100\, kHz - 500\, Hz = 99.5\, kHz.$$

(iii) From equation (1.2.13) we know that the amplitudes of upper and lower sidebands is given as,

$$\text{Amplitude of upper and lower sidebands} = \frac{m E_c}{2} = \frac{0.2 \times 50}{2} = 5V$$

(iv) Bandwidth of AM wave is given by equation(1.2.10) as,

$$BW \text{ of } AM = 2f_m = 2 \times 500\ Hz = 1\ kHz$$

(v) Total power delivered to the load is given by equation (1.2.18) as

$$P_{total} = \frac{E_c^2}{2R}\left(1+\frac{m^2}{2}\right) = \frac{50^2}{2\times 600}\left(1+\frac{(0.2)^2}{2}\right)$$

$\therefore$ $= 2.125$ watts

Example : *A 400 W carrier is modulated to a depth of 80% calculate the totall power in the modulated wave.*

Solution : Here carrier power $P_c = 400\,W$ and $m = 0.8$

From equation (1.2.19) total power is,

$$P_{total} = P_c\left(1+\frac{m^2}{2}\right) = 400\left(1+\frac{(0.8)^2}{2}\right)$$

$$= 528\,W$$

Example : *A broadcast transmitter radiates 20 kW when the modulation percentage is 75. Calculate carrier power and power of each sideband.*

Solution : Here total power $P_{total} = 20{,}000\,W$ and $m = 0.75$

From equation (1.2.19) we have $P_{total} = P_c\left(1+\frac{m^2}{2}\right)$

$\therefore$ $$20{,}000 = P_c\left(1+\frac{(0.75)^2}{2}\right)$$

$\therefore$ $$P_c = 15.6\,kW$$

We know that $$P_{total} = P_c\left(1+\frac{m^2}{2}\right) = P_c + P_c\,\frac{m^2}{2}$$

The second term in above equation represents total sideband power. Hence power of one sideband will be,

$$P_{SB} = \left(P_c\,\frac{m^2}{2}\right)\times\frac{1}{2}$$

$$= 15.6\times\frac{(0.75)^2}{2}\times\frac{1}{2}$$

$$= 2.2\,kW$$

Thus $$P_{USB} = P_{LSB} = 2.2\,kW$$

1.6 FREQUENCY SPECTRUM AND BANDWIDTH REQUIREMENT OF ANGLE MODULATED WAVE.

Frequency Spectrum of Angle Modulated Waves

We know that AM contains only two sidebands per modulating frequency. But angle modulated signal contains large number of sidebands depending upon the modulation index. Since FM and PM have identical modulated waveforms, their frequency content is same. Consider the PM equation for spectrum analysis,

$$e(t) = E_c \sin[\omega_c t + m \cos\omega_m t]$$

Using Bessel functions, this equation can be expanded as,

$$\begin{aligned} e(t) = E_c \{ & J_0 \sin\omega_c t \\ & + J_1 [\sin(\omega_c + \omega_m)t - \sin(\omega_c - \omega_m)t] \\ & + J_2 [\sin(\omega_c + 2\omega_m)t + \sin(\omega_c - 2\omega_m)t] \\ & + J_3 [\sin(\omega_c + 3\omega_m)t + \sin(\omega_c - 3\omega_m)t] \\ & + J_4 [\sin(\omega_c + 4\omega_m)t - \sin(\omega_c - 4\omega_m)t] + \ldots \} \end{aligned}$$

Here $J_0, J_1, J_2 \ldots$ are the Bessel functions.

The values of Bessel functions depend upon modulation index m, They are listed in Table 2.1.1

x	n or Order																
(m)	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}	J_{16}
0.00	1.00	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.25	0.98	0.12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.5	0.94	0.24	0.03	—	—	—	—	—	—	—	—	—	—	—	—	—	—
1.0	0.77	0.44	0.11	0.02	—	—	—	—	—	—	—	—	—	—	—	—	—
1.5	0.51	0.56	0.23	0.06	0.01	—	—	—	—	—	—	—	—	—	—	—	—
2.0	0.22	0.58	0.35	0.13	0.03	—	—	—	—	—	—	—	—	—	—	—	—
2.5	−0.05	0.50	0.45	0.22	0.07	0.02	—	—	—	—	—	—	—	—	—	—	—
3.0	−0.26	0.34	0.49	0.31	0.13	0.04	0.01	—	—	—	—	—	—	—	—	—	—
4.0	−0.40	−0.07	0.36	0.43	0.28	0.13	0.05	0.02	—	—	—	—	—	—	—	—	—
5.0	−0.18	−0.33	0.05	0.36	0.39	0.26	0.13	0.05	0.02	—	—	—	—	—	—	—	—
6.0	0.15	−0.28	−0.24	0.11	0.36	0.36	0.25	0.13	0.06	0.02	—	—	—	—	—	—	—
7.0	0.30	0.00	−0.30	−0.17	0.16	0.35	0.34	0.23	0.13	0.06	0.02	—	—	—	—	—	—
8.0	0.17	0.23	−0.11	−0.29	−0.10	0.19	0.34	0.32	0.22	0.13	0.06	0.03	0.01	—	—	—	—
9.0	−0.09	0.24	0.14	−0.18	−0.27	−0.06	0.20	0.33	0.30	0.21	0.12	0.06	0.03		—	—	—
10.0	−0.25	0.04	0.25	0.06	−0.22	−0.23	−0.01	0.22	0.31	0.29	0.20	0.12	0.06	0.03	0.01	—	—
12.0	0.05	−0.22	−0.08	0.20	0.18	−0.07	−0.24	−0.17	0.05	0.23	0.30	0.27	0.20	0.12	0.07	0.03	0.01
15.0	−0.01	0.21	0.04	−0.19	−0.12	0.13	0.21	0.03	−0.17	−0.22	−0.09	0.10	0.24	0.28	0.25	0.18	0.12

It is clear from the above discussion that, angle modulated signal has infinite number of sidebands as well as carrier in the output. The sidebands are separated from the carrier by $f_m, 2f_m, 3f_m$..... etc. The frequency separation between successive sidebands is f_m. All the sidebands are symmetric around carrier frequency. The amplitudes of the sidebands are $E_c J_0, E_c J_1, E_c J_2, E_c J_3, E_c J_4$, ... and so on.

Bandwidth Requirement

The bandwidth requirement of angle modulated waveforms can be obtained depending upon modulation index. The modulation index can be classified as low (less than 1), medium (1 to 10) and high (greater than 10). The low index systems are called *narrowband FM*. For such systems the frequency spectrum resembles AM. Hence minimum bandwidth is given as,

$$BW = 2f_m \ Hz \qquad \text{... (2.1.19)}$$

For high index modulation, the minimum bandwidth is given as,

$$BW = 2\delta \qquad \text{... (2.1.20)}$$

The bandwidth can also be obtained using bessel table. i.e.,

$$BW = 2n f_m \qquad \text{... (2.1.21)}$$

Here 'n' is the number of significant sidebands obtained from bessel table.

Carson's Rule:

Carson's Rule gives appoximate minimum bandwidth of angle modulated signal as

$$\boxed{BW = 2\left[\delta + f_{m(\max)}\right] Hz} \qquad \text{... (2.1.22)}$$

Here $f_{m(\max)}$ is the maximum modulating frequency. As per Carson's rule, the bandwidth accomodates almost 98% of the total transmitted power.

1.7 THE CONCEPT OF ANGLE MODULATION AND ITS WAVEFORM, FREQUENCY AND PHASE MODULATION:

Angle Modulation

Definition

We know that amplitude, frequency or phase of the carrier can be varied by the modulating signal. Amplitude is varied in AM. *When frequency or phase of the carrier is varied by the modulating signal, then it is called angle modulation.* There are two types of angle modulation.

1. Frequency Modulation : When frequency of the carrier varies as per amplitude variations of modulating signal, then it is called Frequency Modulation (FM). Amplitude of the modulated carrier remains constant.

2. Phase Modulation : When phase of the carrier varies as per amplitude variations of modulating signal, then it is called Phase Modulation (PM). Amplitude of the modulated carrier remains constant.

The angle modulated wave is mathematically expressed as,

$$e(t) = E_c \sin[\omega_c t + \theta(t)]$$

Here $e(t)$ is angle modulated wave

E_c is peak amplitude of the carrier

ω_c carrier frequency

$\theta(t)$ instantaneous phase deviation

The phase deviation takes place in FM as well as PM. Hence phase is direct function of modulating signal. i.e.,

$$\theta(t) \propto e_m(t)$$

Here $e_m(t)$ is the modulating signal.

FM and PM Waveforms

Fig. 2.1.1 shows the waveforms of FM and PM.

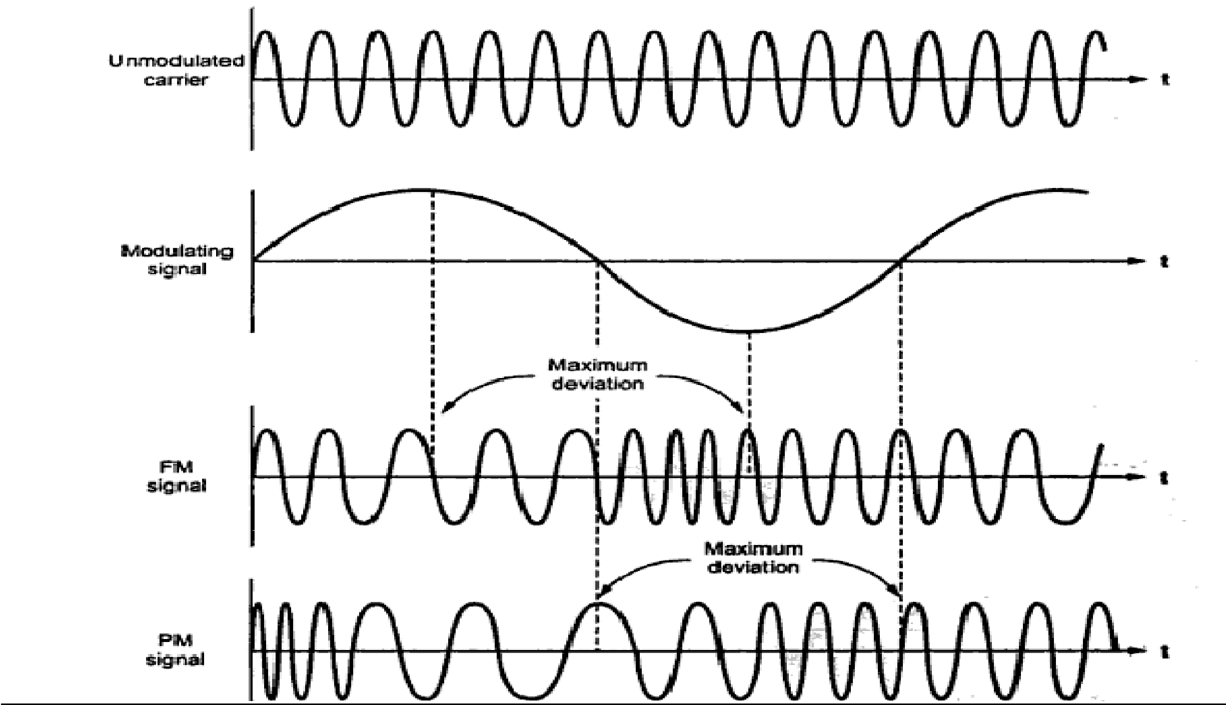
In this figure following observations can be noted :

(i) For FM signal, the maximum frequency deviation takes place when modulating signal is at positive and negative peaks.

(ii) For PM signal the maximum frequency deviation takes place near zero crossings of the modulating signal.

(iii) Both FM and PM waveforms are identical except the phase shift.

(iv) From modulated waveform it is difficult to know, whether the modulation is FM or PM.



The expression for frequency deviation, phase deviation and modulation index in angle modulation.

Phase Deviation, Modulation Index and Frequency Deviation

The FM signal, in general is expressed as,

$$e_{FM}(t) = E_c \sin[\omega_c t + m \sin \omega_m t] \quad \text{... (2.1.12)}$$

And the PM signal, in general is expressed as,

$$e_{PM}(t) = E_c \sin[\omega_c t + m \cos \omega_m (t)] \quad \text{... (2.1.13)}$$

In both the above equations, the term 'm' is called *modulation index*. Note that the term $m \sin \omega_m t$ in equation 2.1.12 and $m \cos \omega_m t$ in equation 2.1.13 indicates instantaneous phase deviation $\theta(t)$. Hence 'm' also indicates *maximum phase deviation*. In other words, modulation index can also be defined as maximum phase deviation.

Modulation index for PM :

Comparing equation 2.1.13 and equation 2.1.11, we find that,

$$\boxed{\text{Modulation index in PM : } m = k E_m \text{ rad}} \quad \text{... (2.1.14)}$$

Thus modulation index of PM signal is directly proportional to peak modulating voltage. And it's unit is radians.

Modulation index for FM :

Comparing equation 2.1.12 and equation 2.1.10 we find that,

$$m = \frac{k_1 E_m}{\omega_m} \qquad \text{... (2.1.15)}$$

Thus modulation index of FM is directly proportional to peak modulating voltage, but inversely proportional to modulating signal frequency.

Since $\omega_m = 2\pi f_m$ above equation becomes,

$$m = \frac{k_1 E_m}{2\pi f_m}$$

Here $\frac{k_1 E_m}{2\pi}$ is called *frequency deviation*. It is denoted by δ and its unit is Hz, i.e.,

$$\boxed{\textbf{Modulation index in FM :} \quad m = \frac{\delta}{f_m} = \frac{\text{Maximum frequency deviation}}{\text{Modulating frequency}}} \qquad \text{... (2.1.16)}$$

Thus modulation index of FM is unitless ratio. From above equation and equation 2.1.14, note that the modulation index is differently defined for FM and PM signals.

Percentage modulation :

For angle modulation, the percentage modulation is given as the ratio of actual frequency deviation to maximum allowable frequency deviation. i.e.,

$$\boxed{\%\ \text{modulation} = \frac{\text{Actual frequency deviation}}{\text{Maximum allowable frequency deviation}}} \qquad \text{... (2.1.17)}$$

UNIT II DIGITAL COMMUNICATION

2.1 INTRODUCTION TO DIGITAL COMMUNICATION:

Digital modulation : Digital Modulation is defined as changing the amplitude of the carrier signal with respect to the binary information or digital signal.

Bit rate is the number of bits transmitted during one second between the transmitter and receiver.

Baud rate is the rate of change of signal on transmission medium after encoding and modulation have occurred.

Bandwidth efficiency : Bandwidth efficiency is the ratio of the transmission bit rate to the minimum bandwidth required for a particular modulation

Advantages of Digital communications

- It has a better noise immunity
- Repeaters can be used between transmitters and receivers
- It becomes simpler and cheaper as compared to the analog communication

Disadvantages of Digital communications

- It requires a larger channel bandwidth
- Delta modulation needs synchronization incase of synchronous modulation

2.2 FREQUENCY SHIFT KEYING (FSK) , MINIMUM SHIFT KEYING (MSK)

Minimum Shift Keying (MSK) : The minimum frequency space that allows the 2 fsk representing symbols 0s and 1s. Thus CP (Continuous Phase) FSK signal with a deviation ratio if one half is defined as MSK.

Frequency Shift Keying (FSK) : Frequency Shift Keying is the as changing amplitude of the carrier signal with respect to the binary information or digital signal.

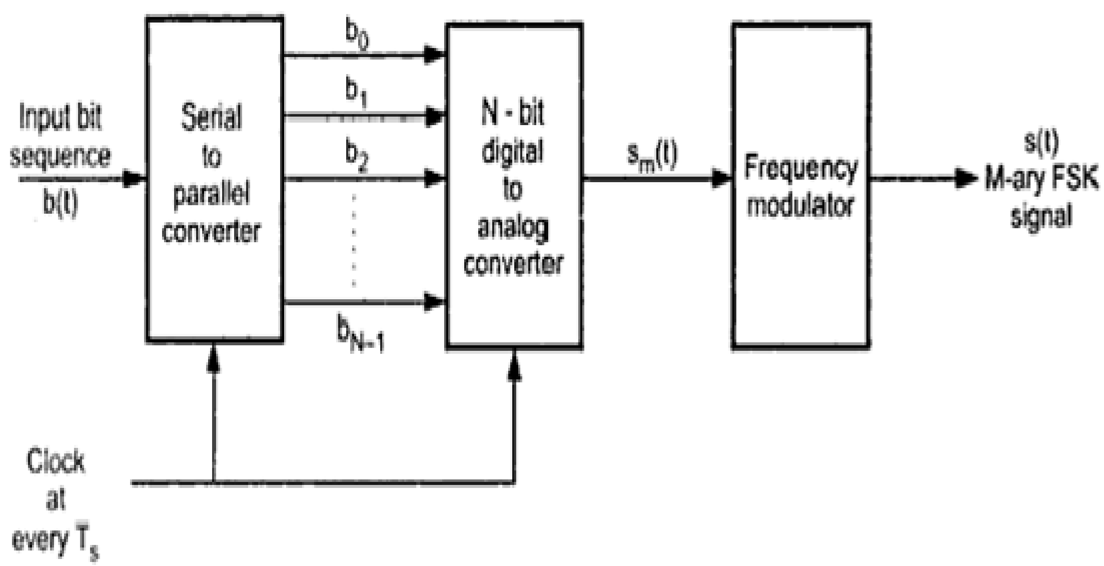
<u>The advantages of Minimum Shift Keying</u> :

MSK baseband waveform are smoother compared with QPSK MSK signals have continuous phase It does not have any amplitude variation

Frequency Shift Keying (FSK)

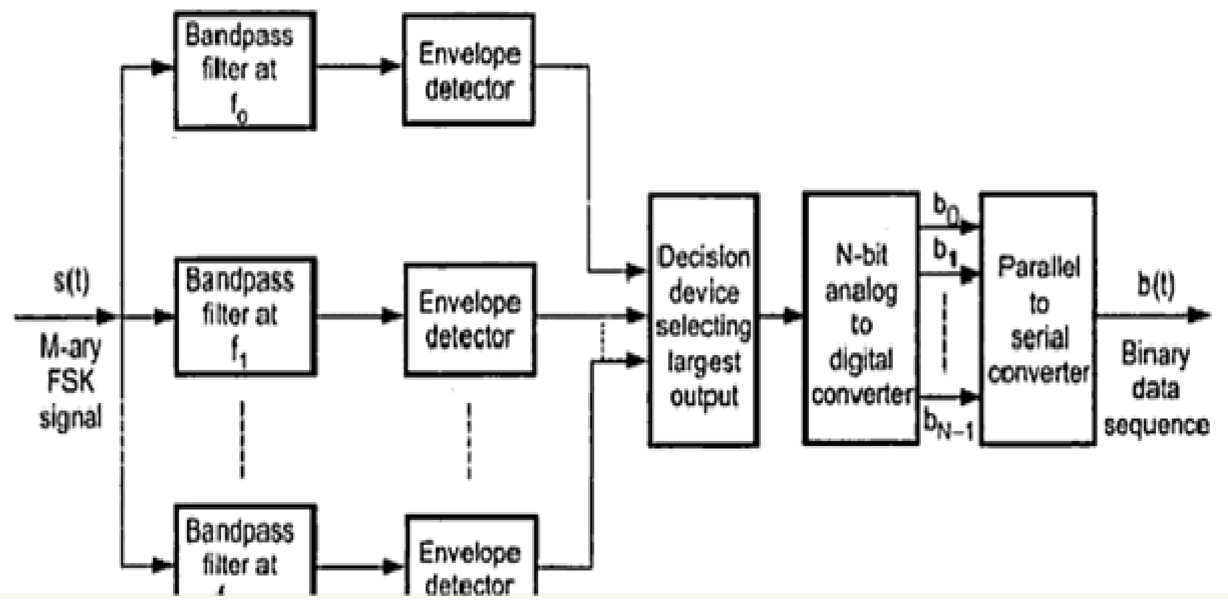
Transmitter

the M-ary FSK transmitter. The 'N' successive bits are presented in parallel to digital to analog converter. These 'N' bits forms a symbol at the output of digital to analog converter. There will be total $2^N = M$ possible symbols. The symbol is presented every $T_s = NT_b$ period. The output of digital to analog converter is given to a frequency modulator. Thus depending upon the value of symbol, the frequency modulator generates the output frequency. For every symbol, the frequency modulator produces different frequency output. This particular frequency signal remains at the output for one symbol duration. Thus for 'M' symbols, there are 'M' frequency signals at the output of modulator. Thus the transmitted frequencies are $f_0, f_1, f_2, \ldots f_{M-1}$ depending upon the input symbol to the modulator.



Receiver

The M-ary FSM signal is given to the set of 'M' bandpass filters. The center frequencies of those filters are $f_0, f_1, f_2, \ldots f_{M-1}$. These filters pass their particular frequency and alternate others. The envelope detectors outputs are applied to a decision device. The decision device produces its output depending upon the highest input. Depending upon the particular symbol, only one envelope detector will have higher output. The outputs of other detectors will be very low. The output of the decision device is given to 'N' bit analog to digital converter. The analog to digital converter output is the 'N' bit symbol in parallel. These bits are then converted to serial bit stream by parallel to serial converter. In some cases the bits appear in parallel. Then there is no need to use serial to parallel and parallel to serial converters.



2.3 AMPLITUDE SHIFT KEYING (ASK) PHASE SHIFT KEYING (PSK):

Amplitude Shift Keying (ASK) : Amplitude Shift Keying is the as changing amplitude of the carrier signal with respect to the binary information or digital signal.

Define Phase Shift Keying (PSK): Phase Shift Keying is the changing amplitude of the carrier signal with respect to the binary information or digital signal

Concept Of Amplitude Shift Keying In Detail:

The amplitude shift keying is also called on-off keying (OOK). This is the simplest digital modulation technique. The binary input data is converted to unipolar NRZ signal. A product modulator takes this NRZ signal and carrier signal. The output of the product modulator is the ASK signal, which can be expressed mathematically as,

$$v(t) = d \sin(2\pi f_c t) \qquad \text{... (2.3.1)}$$

Here f_c is the carrier frequency

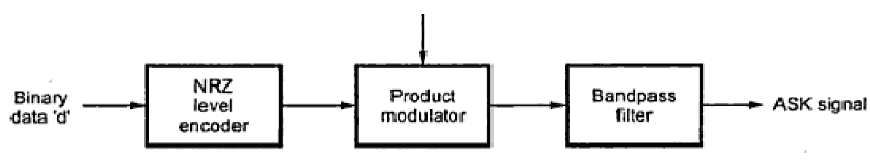
and d is the data bit, which is either 1 or 0.

Fig. 2.3.1 (a) shows the block diagram of the ASK modulator. The binary data sequence 'd' is given to the NRZ level encoder. This NRZ level encoder converts the input binary sequence to the signal suitable for product modulator. The product modulator also accepts a sinusoidal carrier of frequency f_c. The output of the product modulator is passed through a bandpass filter for bandwidth limiting. The output of the bandpass filter is the ASK signal. This signal and other waveforms are shown in Fig. 2.3.1 (b). Observe that the ASK signal has on-off nature. In equation 2.3.1 when $d = 0$, $v(t) = 0$; i.e. no ASK signal. And when $d = 1$, $d = \sin(2\pi f_c t)$. The ASK is very sensitive to noise. It is used for very low bit rates less than around 100 *bps*. The only advantage of ASK is that it is very simple to implement.

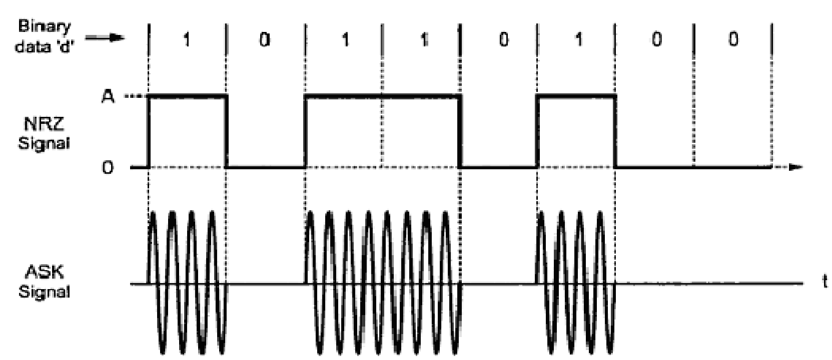
Baud rate

For ASK, the ASK waveform is changed at the bit rate. Hence Baud rate is given as,

$$\text{Baud rate} = f_b \qquad \text{... (2.3.2)}$$



(a) Block diagram of ASK modulator



(b) Waveforms

Fig. 2.3.1 Amplitude shift keying (ASK)

2.4 CONCEPT OF BINARY PHASE SHIFT KEYING TO BPSK, 8 PSK, 16 PSK,

CONCEPT OF BINARY PHASE SHIFT KEYING:

Binary Phase Shift Keying (BPSK)

Principle of BPSK

- In binary phase shift keying (BPSK), binary symbol '1' and '0' modulate the phase of the carrier. Let the carrier be,

$$s(t) = A \cos(2\pi f_0 t) \qquad \text{... (4.2.1)}$$

'A' represents peak value of sinusoidal carrier. In the standard 1Ω load register, the power dissipated will be,

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

$$\therefore \quad A = \sqrt{2P} \qquad \text{... (4.2.2)}$$

- When the symbol is changed, then the phase of the carrier is changed by 180 degrees (π radians).
- Consider for example,

$$\text{Symbol '1'} \Rightarrow s_1(t) = \sqrt{2P} \cos(2\pi f_0 t) \qquad \text{... (4.2.3)}$$

---

if next symbol is '0' then,

$$\text{Symbol '0'} \Rightarrow s_2(t) = \sqrt{2P} \cos(2\pi f_0 t + \pi) \qquad \text{... (4.2.4)}$$

Since $\cos(\theta + \pi) = -\cos\theta$, we can write above equation as,

$$s_2(t) = -\sqrt{2P} \cos(2\pi f_0 t) \qquad \text{... (4.2.5)}$$

With the above equation we can define BPSK signal combinely as,

$$\boxed{s(t) = b(t)\sqrt{2P} \cos(2\pi f_0 t)} \qquad \text{... (4.2.6)}$$

Here $b(t)$ = +1 when binary '1' is to be transmitted

= −1 when binary '0' is to be transmitted

Graphical Representation of BPSK Signal

Fig. 4.2.1 shows binary signal and its equivalent signal $b(t)$.

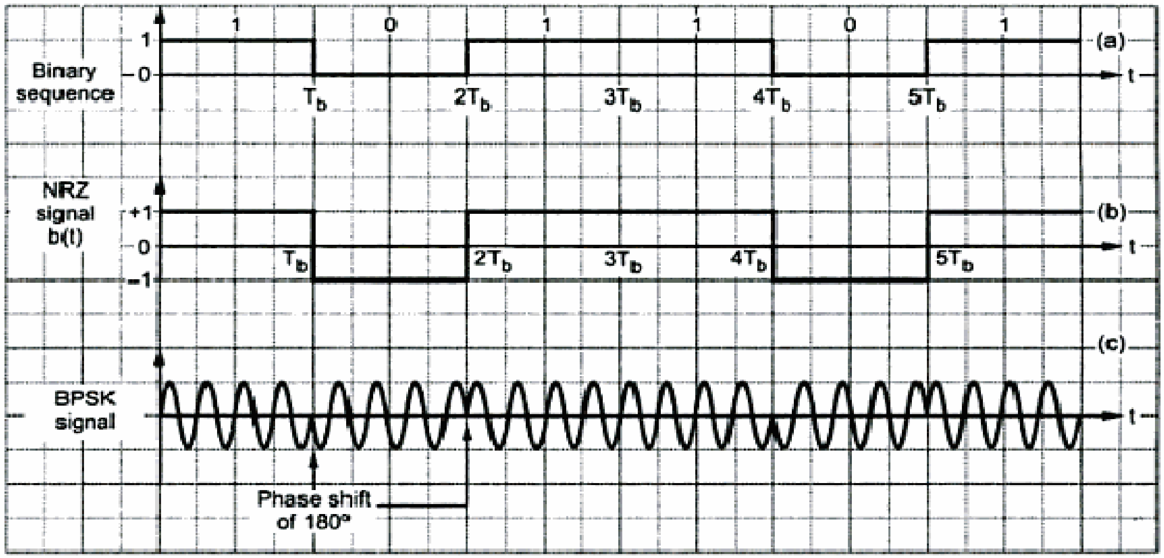


Fig. 4.2.1 (a) Binary sequence
(b) Its equivalent bipolar signal $b(t)$
(c) BPSK signal

As can be seen from Fig. 4.2.1 (b), the signal $b(t)$ is NRZ bipolar signal. This signal directly modulates carrier $\cos(2\pi f_0 t)$.

Generation and Reception of BPSK Signal

Generation of BPSK Signal

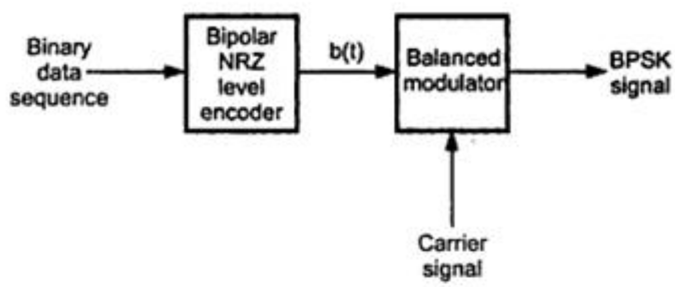


Fig. 4.2.2 BPSK generation scheme

- The BPSK signal can be generated by applying carrier signal to the balanced modulator.
- The baseband signal $b(t)$ is applied as a modulating signal to the balanced modulator. Fig. 4.2.2 shows the block diagram of BPSK signal generator.
- The NRZ level encoder converts the binary data sequence into bipolar NRZ signal.

Reception of BPSK Signal

Fig. 4.2.3 shows the block diagram of the scheme to recover baseband signal from BPSK signal. The transmitted BPSK signal is,

$$s(t) = b(t)\sqrt{2P}\cos(2\pi f_0 t)$$

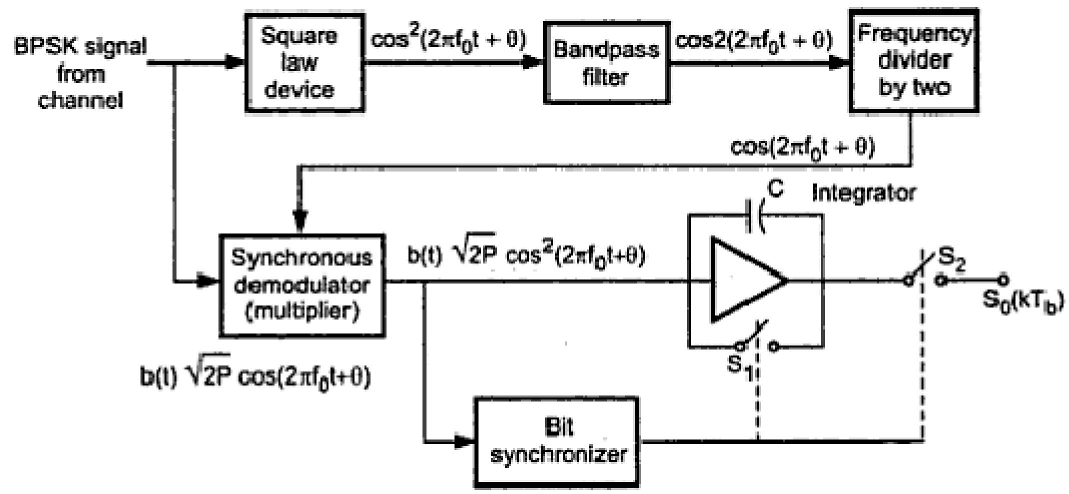


Fig. 4.2.3 Reception BPSK scheme

Operation of the receiver

1) **Phase shift in received signal** : This signal undergoes the phase change depending upon the time delay from transmitter to receiver. This phase change is normally fixed phase shift in the transmitted signal. Let the phase shift be θ. Therefore the signal at the input of the receiver is,

$$s(t) = b(t)\sqrt{2P}\cos(2\pi f_0 t+\theta) \quad \text{... (4.2.7)}$$

2) **Square law device** : Now from this received signal, a carrier is separated since this is coherent detection. As shown in the figure, the received signal is passed through a square law device. At the output of the square law device the signal will be,

$$\cos^2(2\pi f_0 t+\theta)$$

Note here that we have neglected the amplitude, because we are only interested in the carrier of the signal.

We know that,

$$\cos^2\theta = \frac{1+\cos 2\theta}{2}$$

$$\therefore \quad \cos^2(2\pi f_0 t+\theta) = \frac{1+\cos 2(2\pi f_0 t+\theta)}{2}$$

Bandwidth of BPSK Signal

The spectrum of the BPSK signal is centered around the carrier frequency f_0.

If $f_b = \frac{1}{T_b}$, then for BPSK the maximum frequency in the baseband signal will be

f_b see Fig. 4.2.6. In this figure the main lobe is centered around carrier frequency f_0 and extends from $f_0 - f_b$ to $f_0 + f_b$. Therefore Bandwidth of BPSK signal is,

$$BW = \text{Highest frequency} - \text{Lowest frequency in the main lobe}$$

$$= f_0 + f_b - (f_0 - f_b)$$

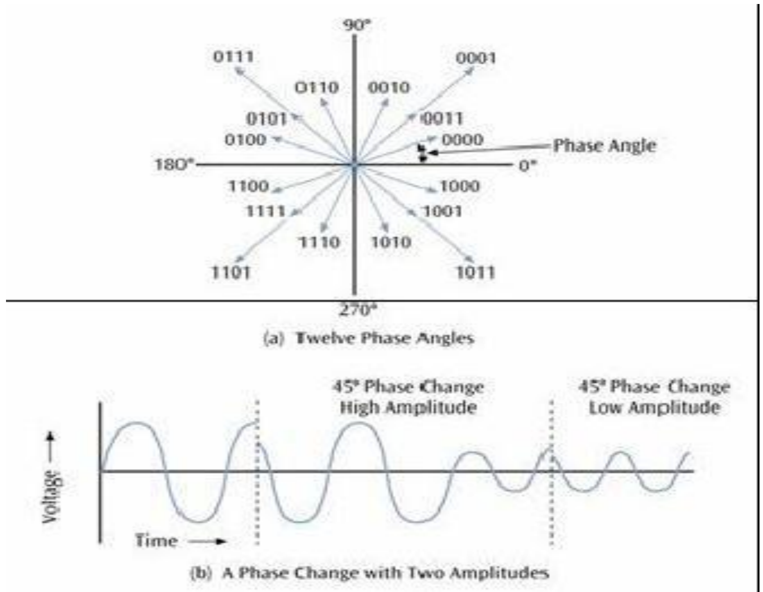
$$\therefore \quad \boxed{BW = 2f_b} \quad \text{... (4.2.21)}$$

Thus the minimum bandwidth of BPSK signal is equal to twice of the highest frequency contained in baseband signal.

2.5 QUADRATURE AMPLITUDE MODULATION:

Quadrature Amplitude modulation (QAM): QAM is a form of digital modulation similar to PSK except the digital information is contained in both the amplitude and phase of the transmitted carrier.

- QAM is a combination of ASK and PSKTwo different signals sent simultaneously on the same carrier frequency ie,M=4, 16, 32, 64, 128, 256 As an example of QAM, 12 different phases are combined with two different amplitudes.
- Since only 4 phase angles have 2 different amplitudes, there are a total of 16 combinations. With 16 signal combinations, each baud equals 4 bits of information (2 ^ 4 = 16).
- Combine ASK and PSK such that each signal corresponds to multiple bits.
- More phases than amplitudes. Minimum bandwidth requirement same as ASK or PSK.



(a) Twelve Phase Angles

(b) A Phase Change with Two Amplitudes

2.6 QUADRATURE PHASE SHIFT KEYING (QPSK) TECHNIQUES AND ITS BLOCK DIAGRAM:

Advantages of QPSK

- Very good noise immunity
- Effective utilization of available bandwidth
- Low error probability

Very high bit rate data transmission

Quadrature Phase Shift Keying (QPSK)

Principle

- In communication systems we know that there are two main resources, i.e. transmission power and the channel bandwidth. The channel bandwidth depends upon the bit rate or signalling rate f_b. In digital bandpass transmission, a carrier is used for transmission. This carrier is transmitted over a channel.
- If two or more bits are combined in some symbols, then the signalling rate is reduced. Therefore the frequency of the carrier required is also reduced. This reduces the transmission channel bandwidth. Thus because of grouping of bits in symbols, the transmission channel bandwidth is reduced.
- In quadrature phase shift keying, two successive bits in the data sequence are grouped together. This reduces the bits rate of signalling rate (i.e. f_b) and hence reduces the bandwidth of the channel.

- In BPSK we know that when symbol changes the level, the phase of the carrier is changed by 180°. Since there were only two symbols in BPSK, the phase shift occurs in two levels only.
- In QPSK two successive bits are combined. This combination of two bits forms four distinct symbols. When the symbol is changed to next symbol the

---

Since $b_o(t)$ and $b_e(t)$ cannot change at the same time, the phase change in QPSK signal will be maximum $\pi / 2$. This is clear from Fig. 4.4.3.

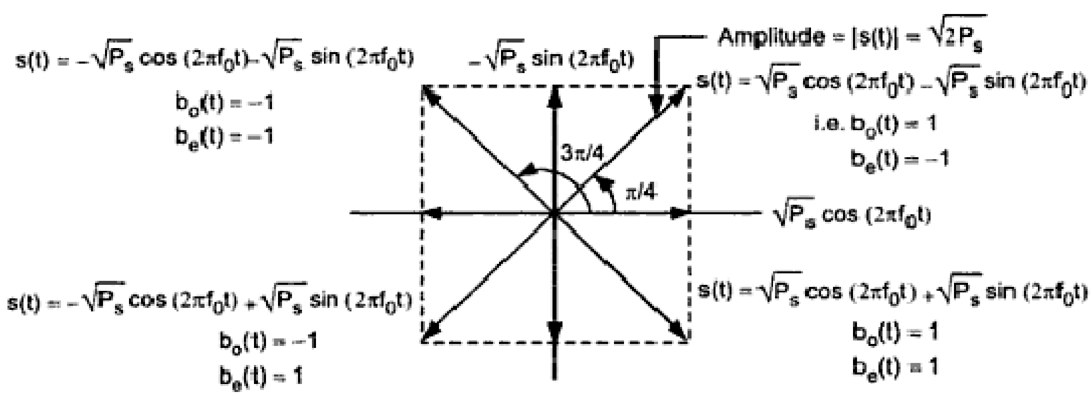
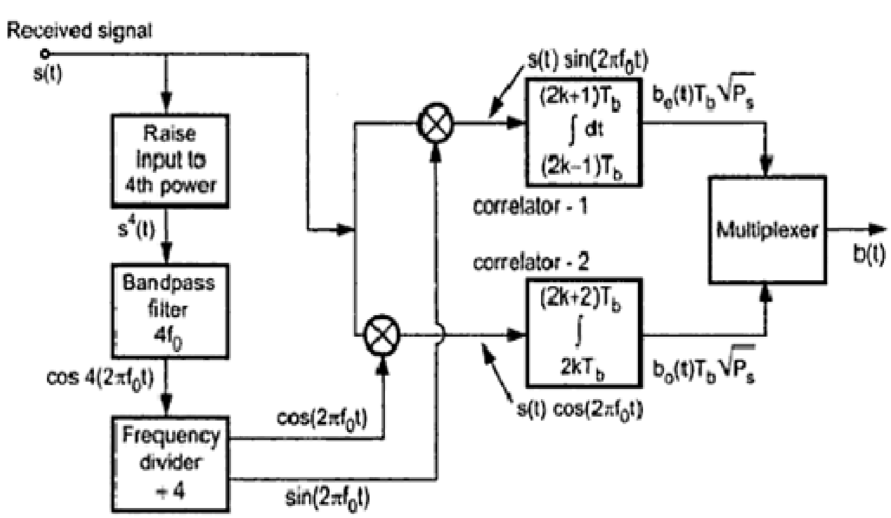


Fig. 4.4.3 Phasor diagram of QPSK signal

The QPSK Receiver



This is Synchronous reception. Therefore coherent carrier is to be recovered from the received signal s(t).

Operation

Step 1 : Isolation of carrier

The received signal $s(t)$ is first raised to its 4^{th} power, i.e. $s^4(t)$. Then it is passed through a bandpass filter centered around $4f_0$. The output of the bandpass filter is a coherent carrier of frequency $4f_0$. This is divided by 4 and it gives two coherent quadrature carriers $cos(2\pi f_0 t)$ and $sin(2\pi f_0 (t)$.

Step 2 : Synchronous detection

These coherent carriers are applied to two synchronous demodulators. These synchronous demodulators consist of multiplier and an integrator.

Step 3 : Integration over two bits interval

The incoming signal is applied to both the multipliers. The integrator integrates the product signal over two bit interval (i.e. $T_s = 2T_b$).

Step 4 : Sampling and multiplexing odd and even bit sequences

At the end of this period, the output of integrator is sampled. The outputs of the two integrators are sampled at the offset of one bit period, T_b. Hence the output of

Advantages of QPSK

QPSK has some definite advantages and disadvantages as compared to BPSK and DPSK.

Advantages :

1) For the same bit error rate, the bandwidth required by QPSK is reduced to half as compared to BPSK.
2) Because of reduced bandwidth, the information transmission rate of QPSK is higher.
3) Variation in OQPSK amplitude is not much. Hence carrier power almost remains constant.

Bandwidth Efficiency (Information Density)

Definition : It is the ratio of transmission bit rate to minimum required bandwidth.

i.e.,

$$\text{BW efficiency} = \frac{\text{Transmission rate(Bits / sec)}}{\text{Minimum bandwidth (cycles / sec)}}$$

$$= \frac{\text{Transmission rate}}{\text{Minimum bandwidth}} \text{ bits/cycle} \qquad \ldots(2.9.1)$$

- When bandwidth efficiency is normalized to 1-Hz bandwidth, it gives number of bits that can be propagated per hertz of bandwidth.
- Bandwidth efficiency is used to compare the performance of digital modulation techniques.

Compare binary PSK with QPSK.

BPSK QPSK

1. One bit forms a symbol. Two bits form a symbol.
2. Two possible symbols. Four possible symbols.
3. Minimum bandwidth is twice of fb. Minimum bandwidth is equal to fb.
4. Symbol duration = Tb. Symbol duration = 2Tb.

9. What are the advantages of M-ary signaling scheme?

1. M-ary signaling schemes transmit bits at a time.
2. Bandwidth requirement of M-ary signaling schemes is reduced.

The probability of error in M-Ary FSK as the value of m increases:

As the value of „M" increases, the Euclidean distance between the symbols reduces. Hence the symbols come closer to each other. This increases the probability of error in M-ary systems.

Correlative coding allows the signaling rate of 2B0 in the channel of bandwidth B0. This is made physically possible by allowing ISI in the transmitted signal in controlled manner. This ISI is known to the receiver. Hence effects of ISI are eliminated at the receiver. Correlative coding is implemented by duobinary signaling and modified duobinary signaling.

UNIT III DATA AND PULSE COMMUNICATION

3.1 DATA COMMUNICATION:

HISTORY OF DATA COMMUNICATION, STANDARDS ORGANIZATIONS FOR DATA COMMUNICATION.

Data communication can be defined as two personal computers connected through a Public Telecommunication Network (PTN). Point to point communication is the link between two stations A and B ie)., information is transferred between a main frame computer and a remote computer terminal.

A multipoint line configuration is one in which more than two specific devices share a single link. Morse code is used to send messages. A key which turned the carrier of a transmitter ON and OFF to produce the dots and dashes. These dots and dashes were detected at the receiver and it is converter back into letters and numbers makes the original message.

Forward Error Correction (FEC)

FEC, a receiver can use an error correcting code, which automatically correct certain errors without any retransmissions

In FEC, bits are added to the message before the transmission

Purpose of FEC code is to reduce the wasted time of retransmission

Data Communications History

- 1838: Samuel Morse & Alfred Veil Invent Morse CodeTelegraph System
- 1876: Alexander Graham Bell invented Telephone
- 1910:Howard Krum developed Start/Stop Synchronisation

History of Computing

- 1930: Development of ASCII Transmission Code
- 1945: Allied Governments develop the First Large Computer
- 1950: IBM releases its first computer IBM 710
- 1960: IBM releases the First Commercial Computer IBM 360

Main Contributors of Data Comm.

• Transmission Technology

• Packet Switching Technology

• Internet – 1967: ARPANET by Advanced Research Project Agency (ARPA) of U.S.

– 1975: TCP/IP protocol

• LAN Technology

– DIX-Ethernet & IEEE 802 Networks

• WAN – 1976: ISO releases HDLC & CCITT releases X.25 (PSPDN)

The Applications Of Data Communication:

- Used in Automatic Teller Machine (ATM) Internet
- Airline and Hotel reservation system
- Mass media
- NEWS network

Advantages And Disadvantages Of Parallel Communication:

Advantages:

Parallel transmission is speed

Used for short distance communication

Disadvantages:

Require more lines between source and destination

More cost

3.2 DATA COMMUNICATION CIRCUITS, DATA COMMUNICATION CODES.

NARROWBAND FM.

When the modulation index is less than 1, the angle modulated systems are called low index. The bandwidth requirement of low index systems is approximately twice of the modulating signal frequency. Therefore low index systems are called narrowband FM.

Reactance modulator is direct FM, where as Armstrong method is indirect FM. Armstrong method generates FM from PM. Hence crystal oscillators can be used in Armstrong method. Therefore frequency stability is better than reactance modulator.

In narrow band FM, the frequency deviation is very small. Hence the frequency spectrum consists of two major sidebands like AM. Other sidebands are negligible and hence they can be neglected. Therefore the bandwidth of narrowband FM is limited only to twice of the highest modulating frequency.

If the deviation in carrier frequency is large enough so that other sidebands cannot be neglected, then it is called wideband FM. The bandwidth of wideband FM is calculated as per Carson‟s rule.

FM has following advantages over AM.

i) The amplitude of FM is constant. It is independent of depth of modulation. Hence transmitter power remains constant in FM whereas it varies in AM.
ii) Since amplitude of FM constant, the noise interference is minimum in FM. Any noise superimposing amplitude can be removed with the help of amplitude limits. Whereas it is difficult to remove amplitude variations due to noise in AM.
iii) The depth of modulation has limitation in AM. But in FM the depth of modulation can be increased to any value by increasing the deviation. This does not cause any distortion in FM signal.
iv) Since guard bands are provided in FM, there is less possibility of adjacent channel interference.
v) Since space waves are used for FM, the radius of propagation is limited to line of sight. Hence it is possible to operate several independent transmitters on same frequency with minimum interference.
vi) Since FM uses UHF and VHF ranges, the noise interference is minimum compared to AM which uses MF and HF ranges.

Carson‟s rule:

Carson‟s rule of FM bandwidth is given as,

BW = 2(_ f+ fm (max))

Here _f is the maximum frequency deviation and fm (max)) is the maximum signal frequency.

In direct FM type of angle modulation, the frequency of the carrier is varied directly by the modulating signal. This means; an instantaneous frequency deviation is directly proportional to amplitude of the modulating signal.

In indirect FM type of angle modulation, FM is obtained by phase modulation of the carrier. This means, an instantaneous phase of the carrier directly proportional to amplitude of the modulating signal

Coherent (synchronous) detection: In coherent detection, the local carrier generated at the receiver is phase locked with the carrier at the transmitter. The detection is done by correlating received noisy signal and locally generated carrier. The coherent detection is a synchronous detection.

Non coherent (envelope) detection: This type of detection does not need receiver carrier to be phase locked with transmitter carrier. The advantage of such a system is that the system becomes simple, but the drawback is that error probability increases. The different digital modulation techniques are used for specific application areas. The choice is made such that the transmitted power and channel bandwidth are best exploited.

ASCII code

ASCII in ANSI X3.4

- Corresponding CCITT recommendation is IA5 (International Alphabet No.5)
- – ISO specification is ISO 646

Total 128 codes

- – 96 codes are graphic symbols (in Col. 2~7).
- 94 codes are printable
- And 2 codes viz. SPACE & DEL characters arenon printable
- 32 codes control symbols (Col. 0 & 1) All are non printable

EBCDIC code

- It is an 8-bit code with 256symbols
- No parity bit for error checking
- The graphic symbols are almostsame as ASCII
- Several differences in Controlcharacters as compared to ASCII

BAUDOT code

- It is a 5-bit code also known asITA2 (International TelegraphAlphabet No. 2).
- 32 codes are possible. With thehelp of Letter shift & Figure shift key same code is used torepresent two symbols.
- Maximum symbols in this code are58
- Used in Telegraphy/Telex

Antipodal Signals

In BPSK, the two symbols are transmitted with the help of following signals,

Symbol „1" s1 (t) = 2P cos (2_ f0 t)

Symbol „0" s2 (t) = 2P cos (2_ f0 t + _)Here observe that above two signals differ only in a relative phase shift of 1800. Such signals are called antipodal signals.

3.3 Error Detection and Correction Techniques.

Different Types Of Error Detection Techniques:

a. Redundancy

b. Echoplex

c. Exact count encoding

d. Parity

e. Check sum

f. Vertical Redundancy Check

g. Horizontal Redundancy Check

h. Cyclic Redundancy Check

Circumstances of M-ary signaling schemes are preferred over binary Schemes:

Binary schemes transmit only one bit at a time. M-ary schemes transmit log2 M bit at a time. When available channel bandwidth is less, then M-ary schemes are used. M ary schemes require less bandwidth compared to binary schemes. For example binary PSK requires a bandwidth of 2fb. But M-ary PSK requires a bandwidth of 2fb.

The Advantages And Disadvantages Of Series Communication

Advantages:

Number of transmission lines is less Used for long distance communication Low cost

Disadvantages:

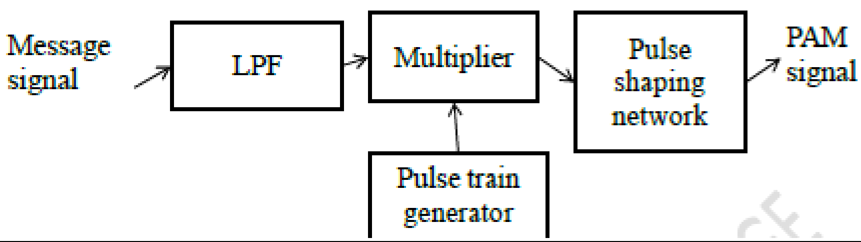
Speed is low

3.4 PULSE COMMUNICATION:

PAM Modulator

The amplitude of a carrier pulse is altered in accordance to that of amplitude of message signal to make it accommodate the information signal.

- Message signal is transmitted to LPF
- LPF performs bandlimiting
- Band limited signal is then sampled at the multiplier.
- Multiplier samples with the help of pulse train generator
- Pulse train generator produces the pulse train
- The multiplication of message signal and pulse train produces PAM



signal

PAM DEMODULATOR:

PAM SIGNAL --→ RECONSTRUCTION FILTER ------→ RECONSTRUCTED PAM SIGNAL

Pulse width modulation

In PWM system, the message signals are used to vary the duration of carrier pulse. The message signal may vary either the trailing edge or leading edge or both of the carrier pulses n order to accommodate the intelligence of information system.

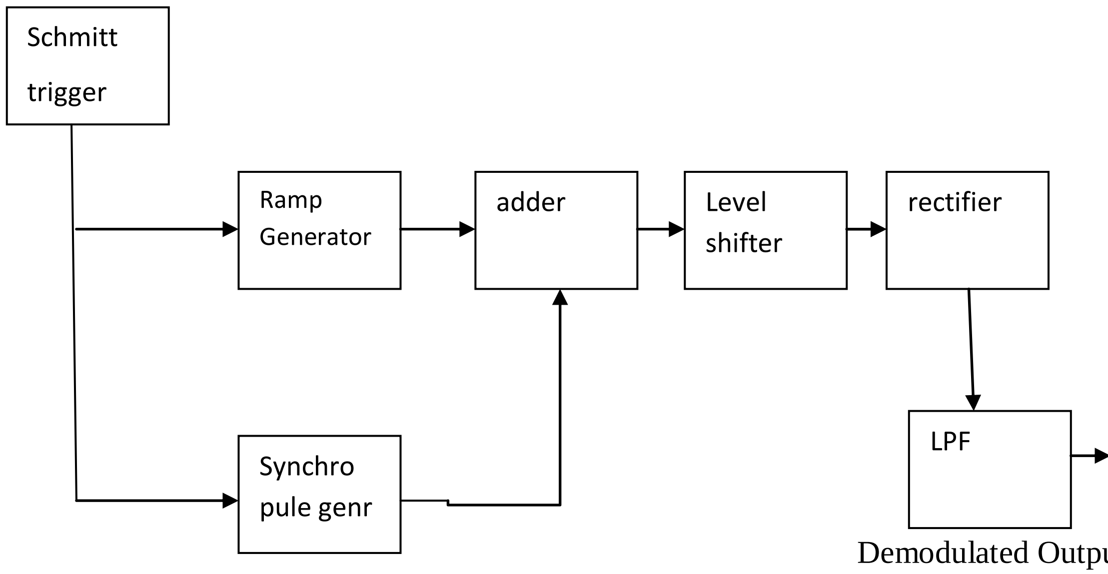
Width of pulse is proportional to the amplitude of the modulating signal.The amplitude and position of the pulse remains unchanged.

PWM Modulator

- It is basically a monostablemultivibrator with message signal applied at the control voltage input.

• Externally applied modulating signal changes the control voltage and hence the threshold voltage level
• The time period required to charge the capacitor upto threshold level changes giving pulse modulated signal

PWM demodulator

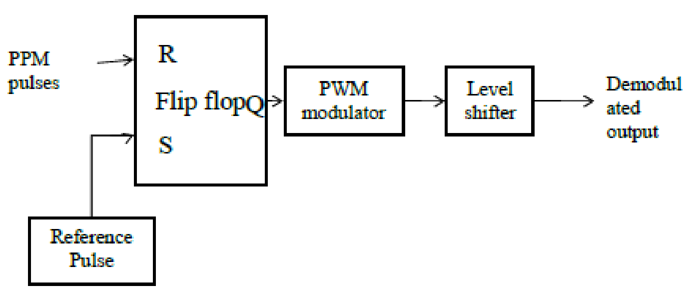


Demodulated Output.

PPM Modulator

• Sawtooth generator generates sawtooth signal of frequency which is applied to inverting input of comparator
• Modulating signal is applied to the non-inverting input of comparator
• When the value of message signal is higher than value of sawtooth ,then the output is high
• When the value of message signal is lower than value of sawtooth ,then the output is high.

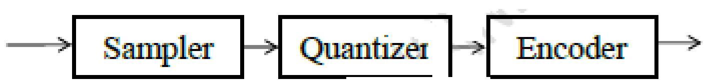
PPM demodulator



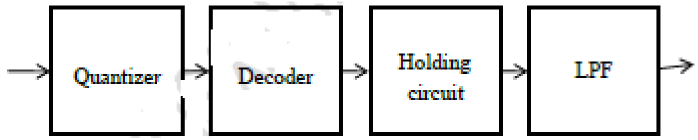
PULSE CODE MODULATION:

Pulse code modulation refers a form of source coding. It is a form of digital modulation techniques in which the code refers a binary word that represent digital data.With PCM, the pulses are of fixed length and fixed amplitude.

Block Diagram of Transmitter



Block Diagram of Receiver



Pulse position modulation

The position of a carrier pulse is altered in accordance with information contained in sampled waveform.

Sampling rate

The sampling rate fs must be atleast two times the highest frequency component of the original signal to be accurately represented fs>=2fm

Baseband signal receiver.

A baseband signal receiver increases the signal to noise at the instant of sampling. This reduces the probability of error. The baseband signal receiver is also called optimum receiver.

Matched filter.

The matched filter is a baseband signal receiver, which works in presence of white Gaussian noise. The impulse response of the matched filter is matched to the shape of the input signal.

The impulse response of matched filter

Impulse response is given as,

h (t) = [2k/ N0]{x1 (T − t)}

Here T is the period of sampling x1 (t) and x2 (t) are the two signals used for transmission.

The value of maximum signal to noise ratio of the matched filter

Maximum signal to noise ratio of the matched filter is the ratio of energy of the signal to psd of white noise.

Correlator: It is the coherent receiver. It correlates the received noisy signal f (t) with the locally generated replica of the known signal x (t). Its output is given as,

r (t) = 0_T f (t) x(t) dt

Matched filter and correlator are functionally same.

The advantages of QPSK as compared to BPSK

1. For the same bit error rate, the bandwidth required by QPSK Is reduced to half as compared to BPSK.
2. Because of reduced bandwidth, the information transmission rate of QPSK is higher.
3. Variation in QPSK amplitude is not much. Hence carrier power almost remains constant

------

UNIT IV SOURCE AND ERROR CONTROL CODING

4.1 ENTROPY, SOURCE ENCODING THEOREM

Source encoding theorem

The discrete memory less source of entropy H(X), the average code word length (L) for any distortion less source encoding is bounded.

Code redundancy is the measure of redundancy bits in the encoded message sequence.

Mutual information is the amount of information transferred when Xi is transmitted and Yi is received. It is represented by I(Xi,Yi) .The average mutual information is defined as the amount of source information gain per received symbol.

A block code of length n and 2k code words is calleda linear (n, k) code if and only if its 2k code words form a k-dimensional subspace of the vector space of all the n-tuples over the field GF(2). The message occurring frequently can be assigned short code words, whereas message which occur rarely are assigned long code word, such coding is called variable length coding.

The efficient representation of data generated by a discrete source is known as source encoding. This device that performs this representation is called source encoder.

The types of error control method

Error detection and retransmission

Error detection and correction

Channel capacity is defined as the maximum of the mutual information that may be transmitted through the channel.

The needs for encoding

To improve the efficiency of communication

To improve the transmission quality.

The entropy of a source is a measure of the average amount of information per source symbol in a long message. Channel coding theorem is applied for discrete memory less additive white gaussian noise channels.

The advantages of Shannon fano coding

Reduced bandwidth

Reduced noise

It can be used for error detection and correction.

The objectives of cyclic codes :

Encoding and syndrome calculations can be easily implemented by using simple shift register with feedback connection It is possible to design codes having useful error correction properties

Source Coding Theorem

The **source coding theorem** shows that (in the limit, as the length of a stream of independent and identically-distributed random variable (data tends to infinity) it is impossible to compress the data such that the code rate (average number of bits per symbol) is less than the Shannon entropy of the source, without it being virtually certain that information will be lost. However it is possible to get the code rate arbitrarily close to the Shannon entropy, with negligible probability of loss.

Proof: Source coding theorem

Given X is an i.i.d. source, its time series X_1, ..., X_n is i.i.d. with entropy $H(X)$ in the discrete-valued case and differential entropy in the continuous-valued case. The Source coding theorem states that for any $\varepsilon > 0$ for any rate larger than the entropy of the source, there is large enough n and an encoder that takes n i.i.d. repetition of the source, $X^{1:n}$, and maps it to $n(H(X) + \varepsilon)$ binary bits such that the source symbols $X^{1:n}$ are recoverable from the binary bits with probability at least $1 - \varepsilon$.

Proof of Achievability. Fix some $\varepsilon > 0$, and let

$$p(x_1, \ldots, x_n) = \Pr\left[X_1 = x_1, \cdots, X_n = x_n\right].$$

The typical set, *Aεn*, is defined as follows:

$$A_n^\varepsilon = \left\{ (x_1, \cdots, x_n) \ : \ \left| -\frac{1}{n} \log p(x_1, \cdots, x_n) - H_n(X) \right| < \varepsilon \right\}.$$

The Asymptotic Equipartition Property (AEP) shows that for large enough *n*, the probability that a sequence generated by the source lies in the typical set, *Aεn*, as defined approaches one. In particular there for large enough *n*, $P(A_n^\varepsilon) > 1 - \varepsilon$(See AEP for a proof):

The definition of typical sets implies that those sequences that lie in the typical set satisfy:

$$2^{-n(H(X)+\varepsilon)} \leq p\left(x_1, \cdots, x_n\right) \leq 2^{-n(H(X)-\varepsilon)}$$

Note that:

- The probability of a sequence from *X* being drawn from *Aε n* is greater than 1 − *ε*.
- $|A_n^\varepsilon| \leq 2^{n(H(X)+\varepsilon)}$since the probability of the whole set *Aε n* is at most one.
- $|A_n^\varepsilon| \geq (1-\varepsilon)2^{n(H(X)-\varepsilon)}$. For the proof, use the upper bound on the probability of each term in typical set and the lower bound on the probability of the whole set *Aε n*.

Since $|A_n^\varepsilon| \leq 2^{n(H(X)+\varepsilon)}, n.(H(X)+\varepsilon)$bits are enough to point to any string in this set.

The encoding algorithm: The encoder checks if the input sequence lies within the typical set; if yes, it outputs the index of the input sequence within the typical set; if not, the encoder outputs an arbitrary *n*(*H*(*X*) + *ε*) digit number. As long as the input sequence lies within the typical set (with probability at least 1 − *ε*), the encoder doesn't make any error. So, the probability of error of the encoder is bounded above by *ε*.

Proof of Converse. The converse is proved by showing that any set of size smaller than *Aε*
n (in the sense of exponent) would cover a set of probability bounded away from 1.

Proof: Source coding theorem for symbol codes

For $1 \leq i \leq n$ let s_i denote the word length of each possible x_i. Define $q_i = a^{-s_i}/C$, where C is chosen so that $q_1 + ... + q_n = 1$. Then

$$\begin{aligned} H(X) &= -\sum_{i=1}^{n} p_i \log_2 p_i \\ &\leq -\sum_{i=1}^{n} p_i \log_2 q_i \\ &= -\sum_{i=1}^{n} p_i \log_2 a^{-s_i} + \sum_{i=1}^{n} p_i \log_2 C \\ &= -\sum_{i=1}^{n} p_i \log_2 a^{-s_i} + \log_2 C \\ &\leq -\sum_{i=1}^{n} -s_i p_i \log_2 a \\ &\leq \mathbb{E} S \log_2 a \end{aligned}$$

where the second line follows from Gibbs' inequality and the fifth line follows from Kraft's inequality:

$$C = \sum_{i=1}^{n} a^{-s_i} \leq 1$$

so $\log C \leq 0$.

For the second inequality we may set

$$s_i = \lceil -\log_a p_i \rceil$$

so that

$$-\log_a p_i \leq s_i < -\log_a p_i + 1$$

and so

$$a^{-s_i} \leq p_i$$

and

$$\sum a^{-s_i} \leq \sum p_i = 1$$

and so by Kraft's inequality there exists a prefix-free code having those word lengths. Thus the minimal S satisfies

$$\begin{aligned} \mathbb{E} S &= \sum p_i s_i \\ &< \sum p_i \left(-\log_a p_i + 1\right) \\ &= \sum -p_i \frac{\log_2 p_i}{\log_2 a} + 1 \\ &= \frac{H(X)}{\log_2 a} + 1 \end{aligned}$$

4.2 Techniques used for compression of information.

Shannon Fano Coding Techniques

In the field of data compression, **Shannon-Fano coding** is a suboptimal technique for constructing a prefix code based on a set of symbols and their probabilities (estimated or measured).

In Shannon-Fano coding, the symbols are arranged in order from most probable to least probable, and then divided into two sets whose total probabilities are as close as possible to being equal.All symbols then have the first digits of their codes assigned; symbols in the first set receive "0" and symbols in the second set receive "1". As long as any sets with more than one member remain, the same process is repeated on those sets, to determine successive digits of their codes. When a set has been reduced to one symbol, of course, this means the symbol's code is complete and will not form the prefix of any other symbol's code.

The algorithm works, and it produces fairly efficient variable-length encodings; when the two smaller sets produced by a partitioning are in fact of equal probability, the one bit of information used to distinguish them is used most efficiently. Unfortunately, Shannon-Fano does not always produce optimal prefix codes; the set of probabilities {0.35, 0.17, 0.17, 0.16, 0.15} is an example of one that will be assigned

Shannon-Fano Algorithm

A Shannon-Fano tree is built according to a specification designed to define an effective code table. The actual algorithm is simple:

1. For a given list of symbols, develop a corresponding list of probabilities or frequency counts so that each symbol's relative frequency of occurrence is known.
2. Sort the lists of symbols according to frequency, with the most frequently occurring symbols at the left and the least common at the right.

1. Divide the list into two parts, with the total frequency counts of the left half being as close to the total of the right as possible.

2. The left half of the list is assigned the binary digit 0, and the right half is assigned the digit 1. This means that the codes for the symbols in the first half will all start with 0, and the codes in the second half will all start with 1.
3. Recursively apply the steps 3 and 4 to each of the two halves, subdividing groups and adding bits to the codes until each symbol has become a corresponding code leaf on the tree.

4.3 Huffmann Coding Techniques

Huffman coding is an entropy encoding algorithm used for lossless data compression. The term refers to the use of a variable-length code table for encoding a source symbol (such as a character in a file)

Huffman coding uses a specific method for choosing the representation for each symbol, resulting in a prefix code (sometimes called "prefix-free codes") (that is, the bit string representing some particular symbol is never a prefix of the bit string representing any other symbol) that expresses the most common characters using shorter strings of bits than are used for less common source symbols. Huffman was able to design the most efficient compression method *of this type*: no other mapping of individual source symbols to unique strings of bits will produce a smaller average output size when the actual symbol frequencies agree with those used to create the code.

Although Huffman coding is optimal for a symbol-by-symbol coding (i.e. a stream of unrelated symbols) with a known input probability distribution, its optimality can sometimes accidentally be over-stated. For example, arithmetic coding and LZW coding often have better compression capability.

Given

A set of symbols and their weights (usually proportional to probabilities).

Find

A prefix-free binary code (a set of codewords) with minimum expected codeword length (equivalently, a tree with minimum weighted path length).

Input.

Alphabet , which is the symbol alphabet of size *n*.

Set , which is the set of the (positive) symbol weights (usually proportional to probabilities), i.e. .

Output.

Code , which is the set of (binary) codewords, where c_i is the codeword for .

Goal.

Let be the weighted path length of code *C*. Condition: for any code .

For any code that is *biunique*, meaning that the code is ***uniquely decodable***, the sum of the probability budgets across all symbols is always less than or equal to one. In this example, the sum is strictly equal to one; as a result, the code is termed a *complete* code. If this is not the case, you can always derive an equivalent code by adding extra symbols (with associated null probabilities), to make the code complete while keeping it *biunique*.

In general, a Huffman code need not be unique, but it is always one of the codes minimizing $L(C)$.

4.4 MUTUAL INFORMATION

On an average we require ***H*(*X*)** bits of information to specify one input symbol. However, if we are allowed to observe the output symbol produced by that input, we require, then, only ***H* (*X*|*Y*)** bits of information to specify the input symbol. Accordingly, we come to the conclusion, that on an average, observation of a single output provides with [***H*(*X*) – *H* (*X*|*Y*)**] bits of information. This difference is called ‘ ***Mutual Information***’ or ‘ ***Transinformation***’ of the channel, denoted by ***I*(*X, Y*).** Thus:

$$I(X, Y) \triangleq H(X) - H(X|Y)$$

Notice that in spite of the variations in the source probabilities, $p(x_k)$ (may be due to noise in the channel), certain probabilistic information regarding the state of the input is available, once the conditional probability $p(x_k \mid y_j)$ is computed at the receiver end. The difference between the initial uncertainty of the source symbol x_k, i.e. $\log 1/p(x_k)$ and the final uncertainty about the same source symbol x_k, after receiving y_j, **i.e.** $\log 1/p(x_k \mid y_j)$ is the information gained through the channel. This difference we call as the mutual information between the symbols $\boldsymbol{x_k}$ and $\boldsymbol{y_j}$. Thus

$$I(x_k, y_j) = \log \frac{1}{p(x_k)} - \log \frac{1}{p(x_k | y_j)}$$

$$= \log \frac{p(x_k | y_j)}{p(x_k)}$$

Or $$I(x_k, y_j) = \log \frac{p(x_k . y_j)}{p(x_k).p(y_j)}$$

$$I(x_k) = I(x_k, x_k) = \log \frac{p(x_k | x_k)}{p(x_k)} = \log \frac{1}{p(x_k)}$$

This is the definition with which we started our discussion on informationtheory.Accordingly ***I (x*** $\boldsymbol{I(x_k)}$ is also referred to as 'Self Information'.

It is clear from Eq (3.21*b*) that, as $\dfrac{p(x_k, y_j)}{p(x_k)} = p(y_j | x_k)$,

$$I(x_k, y_j) = \log \frac{p(y_j | x_k)}{p(y_j)} = \log \frac{1}{p(y_j)} - \log \frac{1}{p(y_j | x_k)}$$

Or $$I(x_k, y_j) = I(y_j) - I(y_j | x_k)$$

Eq (4.22) simply means that "the Mutual information ' is symmetrical with respect to its arguments.i.e.

$$I(x_k, y_j) = I(y_j, x_k)$$

Averaging Eq. (4.21*b*) over all admissible characters x_k and y_j, we obtain the average information gain of the receiver:

$$I(X, Y) = E\{I(x_k, y_j)\}$$

$$= \sum_k \sum_j I(x_k, y_j).\, p(x_k, y_j)$$

$$= \sum_k \sum_j p(x_k, y_j).\log \frac{p(x_k, y_j)}{p(x_k)p(y_j)}$$

1) $$I(X, Y) = \sum_k \sum_j p(x_k, y_j) \left[\log \frac{1}{p(x_k)} - \log \frac{1}{p(x_k \mid y_j)}\right] = H(X) - H(X \mid Y)$$

2) $$I(X, Y) = \sum_k \sum_j p(x_k, y_j) \left[\log \frac{1}{p(y_j)} - \log \frac{1}{p(y_j \mid x_k)}\right] = H(Y) - H(Y \mid X)$$

3) $$I(X,Y) = \sum_k \sum_j p(x_k, y_j) \log \frac{1}{p(x_k)} + \sum_k \sum_j p(x_k, y_j) \log \frac{1}{p(y_j)} - \sum_K \sum_J p(x_k, y_j) \log \frac{1}{p(x_k\, y_j)}$$

Or $I(X, Y) = H(X) + H(Y) - H(X, Y)$

Further, we conclude that, " ***even though for a particular received symbol, y_j, H(X) – H(X | Yj) may be negative, when all the admissible output symbols are covered, the average***

mutual information is always non- negative". That is to say, we cannot loose information on an average by observing the output of a channel. An easy method, of remembering the various relationships, is given in Fig 4.2.Althogh the diagram resembles a Venn-diagram, it is not, and the diagram is only a tool to remember the relationships. That is all. You cannot use this diagram for proving any result.

The entropy of X is represented by the circle on the left and that of Y by the circle on the right. The overlap between the two circles (dark gray) is the mutual information so that the remaining (light gray) portions of $H(X)$ and $H(Y)$ represent respective equivocations. Thus we have

$$H(X \mid Y) = H(X) - I(X, Y) \text{ and } H(Y \mid X) = H(Y) - I(X, Y)$$

The joint entropy $H(X,Y)$ is the sum of $H(X)$ and $H(Y)$ except for the fact that the overlap is added twice so that

$$H(X, Y) = H(X) + H(Y) - I(X, Y)$$

Also observe $H(X, Y) = H(X) + H(Y|X)$

$$= H(Y) + H(X|Y)$$

For the **JPM** given by $I(X, Y)$ **= 0.760751505 bits / sym**

4.5 ENTROPY:

The different conditional entropies, Joint and Conditional Entropies:

It is clear that all the probabilities encountered in a two dimensional communication system could be derived from the **JPM**. While we can compare the **JPM**, therefore, to the impedance or admittance matrices of an ***n***-port electric network in giving a unique description of the system under consideration, notice that the **JPM** in general, need not necessarily be a square matrix and even if it is so, it need not be symmetric.

We define the following entropies, which can be directly computed from the **JPM**.

$$H(X, Y) = p(x_1, y_1) \log \frac{1}{p(x_1, y_1)} + p(x_1, y_2) \log \frac{1}{p(x_1, y_2)} + \ldots + p(x_1, y_n) \log \frac{1}{p(x_1, y_n)}$$

$$+ p(x_2, y_1) \log \frac{1}{p(x_2, y_1)} + p(x_2, y_2) \log \frac{1}{p(x_2, y_2)} + \ldots + p(x_2, y_n) \log \frac{1}{p(x_1, y_1)}$$

$$+ \ldots\ p(x_m, y_1) \log \frac{1}{p(x_m, y_1)} + p(x_m, y_2) \log \frac{1}{p(x_m, y_2)} + \ldots\ p(x_m, y_n) \log \frac{1}{p(x_m, y_n)}$$ or

$$H(X, Y) = \sum_{k=1}^{m}\sum_{j=1}^{n} p(x_k, y_j) \log \frac{1}{p(x_k, y_j)}$$

$$H(X) = \sum_{k=1}^{m} p(x_k) \log \frac{1}{p(x_k)}$$

Using Eq (4.6) <u>only for the multiplication term</u>, this equation can be re-written as:

$$H(X) = \sum_{k=1}^{m}\sum_{j=1}^{n} p(x_k, y_j) \log \frac{1}{p(x_k)}$$

Similarly, $H(Y) = \sum^{n}\sum^{m} p(x_k, y_j) \log \frac{1}{p(y_j)}$

Next, from the definition of the conditional probability we have:

$$P\{X = x_k \mid Y = y_j\} = \frac{P\{X = x_k, Y = y_j\}}{}$$

$$P\{Y = y_j\}$$

i.e., $p(x_k \mid y_j) = p(x_k, y_j) / p(y_j)$

Then $\displaystyle\sum_{k=1}^{m} p(x_k \mid y_j) = \frac{1}{p(y_j)} \sum_{k=1}^{m} p(x_k, y_j) = \frac{1}{p(y_j)} . p(y_j) = 1$

Thus, the set $[X \mid y_j] = \{x_1 \mid y_j, x_2 \mid y_j \ldots x_m \mid y_j\}$; $P[X \mid y_j] = \{p(x_1 \mid y_j), p(x_2 \mid y_j) \ldots p(x_m \mid y_j)\}$, forms a complete finite scheme and an entropy function may therefore be defined for this scheme as below:

$$H(X \mid y_j) = \sum_{k=1}^{m} p(x_k \mid y_j) \log \frac{1}{p(x_k \mid y_j)}.$$

Taking the average of the above entropy function for all admissible characters received, we have the average " ***conditional Entropy" or "Equivocation":***

$$H(X \mid Y) = E\{H(X \mid y_j)\}_j$$

$$= \sum_{j=1}^{n} p(y_j)\, H(X \mid y_j)$$

$$= \sum_{j=1}^{n} p(y_j) \sum_{k=1}^{M} p(x_k \mid y_j) \log \frac{1}{p(x_k \mid y_j)}$$

Or $\displaystyle H(X \mid Y) = \sum^{n}\sum^{m} p(x_k, y_j) \log \frac{1}{p(x_k \mid y_j)}$

$j=1\ k=1$

specifies the " ***Equivocation*** ". It specifies the average amount of information n eeded to specify an input character provided we are allowed to make an observation of the output produced by that input. Similarly one can define the conditional entropy ***H*(*Y* | *X*)** by:

$$\boldsymbol{H(Y\,|\,X) = \sum_{k=1}^{m}\sum_{j=1}^{n} p(x_k, y_j)\log\frac{1}{p(y_j\,|\,x_k)}}$$

Observe that the manipulations, ' ***The entropy you want is simply the double summation of joint probability multiplied by logarithm of the reciprocal of the probability of interest'*** . For example, if you want joint entropy, then the probability of interest will be joint probability. If you want source entropy, probability of interest will be the source probability. If you want the equivocation or conditional entropy, ***H*** **(*X*|*Y*)** then probability of interest will be the conditional probability $\boldsymbol{p\,(x_K\,|y_j)}$ and so on**.**

All the five entropies so defined are all inter-related. For example, We have

$$\boldsymbol{H(Y\,|\,X)} = \sum_k\sum_j p(x_k, y_j)\log\frac{1}{p(y_j\,|\,x_k)}$$

$$\text{Since,}\quad \frac{1}{p(y_j\,|\,x_k)} = \frac{p(x_k)}{p(x_k, y_j)}$$

We can straight away write:

$$\boldsymbol{H\,(Y|X)} = \sum\sum p(x_k, y_j)\,log\frac{1}{p(y_j\,|\,x_k)} - \sum\sum p(x_k, y_j)\,log\frac{1}{p(x_k)}$$

$k \quad j \qquad\qquad k \quad j$

Or $\boldsymbol{H(Y \mid X) = H(X, Y) - H(X)}$

That is: $\boldsymbol{H(X, Y) = H(X) + H(Y \mid X)}$

Similarly, you can show: $\boldsymbol{H(X, Y) = H(Y) + H(X \mid Y)}$

Consider $\boldsymbol{H(X) - H(X \mid Y)}$**.** We have:

$$H(X) - H(X \mid Y) = \sum_k \sum_j p(x_k, y_j) \left[\log \frac{1}{p(x_k)} \square \log \frac{1}{p(x_k \mid y_j)} \right]$$

$$= \sum_k \sum_j p(x_k, y_j) \log \frac{p(x_k, y_j)}{p(x_k) \cdot p(y_j)}$$

Using the logarithm inequality derived earlier, you can write the above equation as:

$$H(X) - H(X \mid Y) = \log e \sum_k \sum_j p(x_k, y_j) \ln \frac{p(x_k, y_j)}{p(x_k) \cdot p(y_j)}$$

$$\geq \log e \sum_k \sum_j p(x_k, y_j) \left(1 - \frac{p(x_k) \cdot p(y_j)}{p(x_k, y_j)} \right)$$

$$\geq \log e \left(\sum_k \sum_j p(x_k, y_j) - \sum_{kj} \sum p(xk).p(yj) \right)$$

$$\geq \log e \left(\sum_k \sum_j p(x_k, y_j) - \sum_k p(xk).\sum_j p(yj) \right) \geq 0$$

Because $\sum_k \sum_j p(x_k, y_j) \square \sum_k p(x_k) \square \sum_j p(y_j) = 1$. Thus it follows that:

$$H(X) \geq H(X \mid Y)$$

Similarly, $\boldsymbol{H(Y) \square H(Y \mid X)}$

Equality in holds iffy P $\boldsymbol{(x_k, y_j) = p(x_k).p(y_j)}$; ***i***.e., **if and only if input symbols and output symbols are statistically independent of each other.**

Binary Symmetric Channels (BSC):

The channel is called a 'Binary Symmetric Channel' or (***BSC***). It is one of the most common and widely used channels. The channel diagram of a ***BSC*** is shown in Fig 3.4. Here ' ***p***' is called the error probability.

For this channel we have:

$$H(Y \mid X) = p \log \frac{1}{p} + q \log \frac{1}{q} = H(p)$$

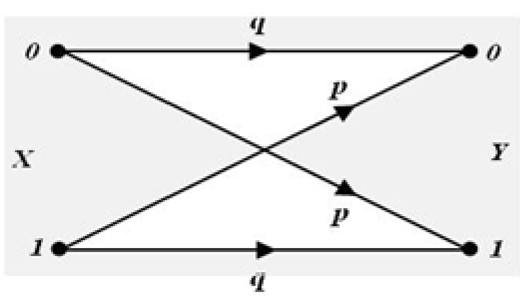
$$H(Y) = [p - \alpha(p - q)] \log \frac{1}{[p - \alpha(p - q)]} + [q + \alpha(p - q)] \log \frac{1}{[q + \alpha(p - q)]}$$

***I*(*X*, *Y*) = *H*(*Y*) – *H* (*Y*|*X*)** and the channel capacity is:

$$C = 1 + p \log p + q \log q$$

This occurs when ***α = 0.5 i.e. P(X=0) = P(X=1) = 0.5***

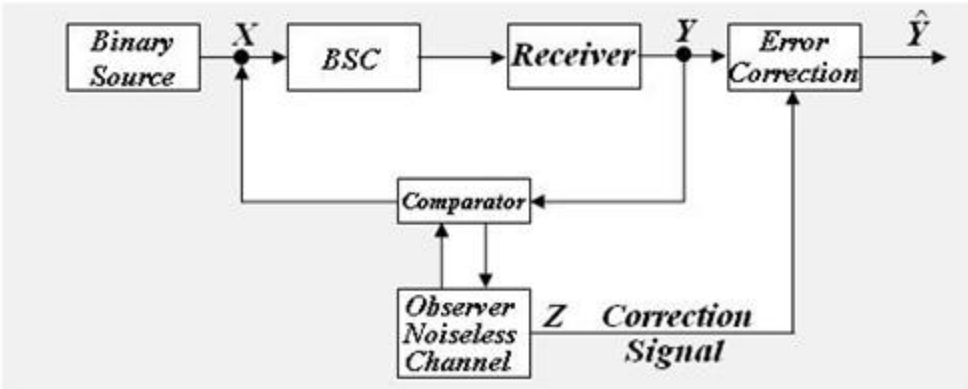
In this case it is interesting to note that the equivocation, ***H* (*X*|*Y*) =*H* (*Y*|*X*).**



$$P(Y|X) = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & q & p \\ 1 & q & p \end{array}$$

$p + q = 1$
$P(X=0) = \alpha$
$P(X=1) = 1 - \alpha$

An interesting interpretation of the equivocation may be given if consider an idealized communication system with the above symmetric channel.



The observer is a noiseless channel that compares the transmitted and the received symbols. Whenever there is an error a ‘ ***1***’ is sent to the receiver as a correction signal and appropriate correction is effected. When there is no error the observer transmits a ‘ ***0***’ indicating no change. Thus the observer supplies additional information to the receiver, thus compensating for the noise in the channel. Let us compute this additional information .With ***P (X=0) = P (X=1) = 0.5***, we have:

Probability of sending a ‘1’ = Probability of error in the channel .

Probability of error = P (Y=1|X=0).P(X=0) + P (Y=0|X=1).P(X=1) = p × 0.5 + p × 0.5 = p

$\therefore$ ***Probability of no error = 1 – p = q***

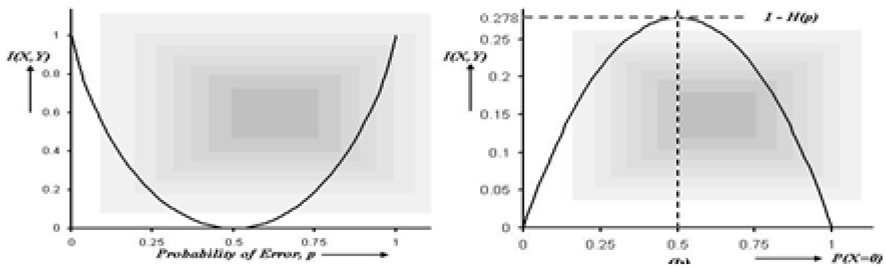
Thus we have P ***(Z = 1) = p*** and ***P (Z = 0) =q***

Accordingly, additional amount of information supplied is:

$$p\log\frac{1}{p} + q\log\frac{1}{q} = H(X|Y) = H(Y|X)$$

Thus the additional information supplied by the observer is exactly equal to the equivocation of the source. Observe that if ‘ ***p***’ and ‘ ***q***’ are interchanged in the channel matrix, the trans - information of the channel remains unaltered. The variation of the mutual information with the probability of error is

shown in Fig 3.6(*a*) for ***P (X=0) = P (X=1) = 0.5***. In Fig 4.6(*b*) is shown the dependence of the mutual information on the source probabilities.



Binary Erasure Channels (*BEC*):

BEC is one of the important types of channels used in digital communications. Observe that whenever an error occurs, the symbol will be received as ‘ ***y***’ and no decision will be made about the information but an immediate request will be made for retransmission, rejecting what have been received (***ARQ*** techniques), thus ensuring ***100%*** correct data recovery. Notice that this channel also is a symmetric channel and we have with ***P(X = 0) =α, P(X = 1) = 1 - α.***

$$H(Y|X) = p\log\frac{1}{p} + q\log\frac{1}{q}$$

$$H(X) = \alpha\log\frac{1}{\alpha} + (1-\alpha)\log\frac{1}{(1-\alpha)}$$

The ***JPM*** is obtained by multiplying first row of ***P (Y|X)*** by α and second row by ***(1– α).*** We get:

$$P(X,Y) = \begin{bmatrix} q\alpha & p\alpha & 0 \\ 0 & p(1-\alpha) & q(1-\alpha) \end{bmatrix}$$

Adding column wise we get: $P(Y) = [q\alpha, p, q(1-\alpha)]$

From which the CPM $P(X|Y)$ is computed as:

$$P(X|Y) = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & (1-\alpha) & 1 \end{bmatrix}$$

$$\therefore H(X|Y) = -q\alpha \log 1 - p\alpha \log \frac{1}{\alpha} - p(1-\alpha)\log \frac{1}{(1-\alpha)} - (1-\alpha) q \log 1$$

$$= pH(X)$$

$$\therefore I(X,Y) = H(X) - H(X|Y) = (1-p)\,H(X) = q\,H(X)$$

$$\therefore C = Max\ I(X,Y) = q \text{ bits / symbol.}$$

In this particular case, use of the equation ***I(X, Y) = H(Y) – H(Y | X)*** will not be correct, as ***H(Y)*** involves ‘ ***y***’ and the information given by ‘ ***y***’ is rejected at the receiver.

4.6 Channel Capacity theorem

Shannon’s theorem: on channel capacity(“coding Theo rem”)

It is possible, in principle, to device a means where by a communication system will transmit information with an arbitrary small probability of error, provided that the information rate R(=$r\times I$ (X,Y),where r is the symbol rate) is less than or equal to a rate ‘ C’ called “channel capacity”.

The technique used to achieve this objective is called coding. To put the matter more formally, the theorem is split into two parts and we have the following statements.

Positive statement:

“ *Given a source of **M** equally likely messages, with **M>>1**, which is generating information at a rate **R**, and a channel with a capacity **C.** If **R ≤ C**, then there exists a coding technique such that the output of the source may be transmitted with a probability of error of receiving the message that can be made arbitrarily small*”.

This theorem indicates that for $\boldsymbol{R \le C}$ transmission may be accomplished without error even in the presence of noise. The situation is analogous to an electric circuit that comprises of only pure capacitors and pure inductors. In such a circuit there is no loss of energy at all as the reactors have the property of storing energy rather than dissipating.

Negative statement:

" *Given the source of **M** equally likely messages with **M>>1**, which is generating information at a rate **R** and a channel with capacity **C**. Then, if **R>C**, then the probability of error of receiving the message is close to unity for every set of **M** transmitted symbols*".

This theorem shows that if the information rate ***R*** exceeds a specified value ***C***, the error probability will increase towards unity as ***M*** increases. Also, in general, increase in the complexity of the coding results in an increase in the probability of error. Notice that the situation is analogous to an electric network that is made up of pure resistors. In such a circuit, whatever energy is supplied, it will be dissipated in the form of heat and thus is a "lossy network".

You can interpret in this way: Information is poured in to your communication channel. You should receive this without any loss. Situation is similar to pouring water into a tumbler. Once the tumbler is full, further pouring results in an over flow. You cannot pour water more than your tumbler can hold. Over flow is the loss.

Shannon defines " ***C***" the channel capacity of a communication channel a s the maximum value of Transinformation, ***I*(*X, Y*):**

$$C \underset{-}{=} \Delta \; Max\; I(X, Y) = Max\,[H(X) - H(Y|X)]$$

The maximization in Eq (4.28) is with respect to all possible sets of probabilities that could be assigned to the input symbols. Recall the maximum power transfer theorem: 'In any network, maximum power will be delivered to the load only when the load and the source are properly matched'. The device used for this matching purpose, we shall call a "transducer ". For example, in a radio receiver, for optimum response, the impedance of the loud speaker will be matched to the impedance of the output power amplifier, through an output transformer.

This theorem is also known as "The Channel Coding Theorem" (Noisy Coding Theorem). It may be stated in a different form as below:

$$R \leq C \text{ or } r_s\, H(S) \leq r_c\, I(X,Y)_{Max} \text{ or} \{ H(S)/T_s\} \leq \{ I(X,Y)_{Max}/T_c\}$$

"If a discrete memoryless source with an alphabet 'S' has an entropy H(S) and produces symbols every 'T_s' seconds; and a discrete memoryless channel has a capacity $I(X,Y)_{Max}$ and is used once every T_c seconds; then if

$$\frac{H(S)}{T_s} \square \frac{I(X,Y)_{Max}}{T_c}$$

There exists a coding scheme for which the source output can be transmitted over the channel and be reconstructed with an arbitrarily small probability of error. The parameter C/T_c is called the critical rate. When this condition is satisfied with the equality sign, the system is said to be signaling at the critical rate.

Conversely, if $\frac{H(S)}{T_s} \square \frac{I(X,Y)_{Max}}{T_c}$, it is not possible to transmit information over the channel and reconstruct it with an arbitrarily small probability of error

A communication channel, is more frequently, described by specifying the source probabilities ***P(X)*** & the conditional probabilities ***P (Y|X)*** rather than specifying the **JPM**. The **CPM**, ***P (Y|X),*** is usually referred to as the '***noise characteristic***' of the channel. Therefore unless otherwise specified, we shall understand that the description of the channel, by a matrix or by a 'Channel diagram' always refers to **CPM, *P (Y|X)***. Thus, in a discrete communication channel with pre-specified noise characteristics (i.e. with a given transition probability matrix, $\boldsymbol{P(Y|X)}$) the rate of information transmission depends on the source that drives the channel. Then, the maximum rate corresponds to a proper matching of the source and the channel. This ideal characterization of the source depends in turn on the transition probability characteristics of the given channel.

Bandwidth-Efficiency: Shannon Limit:

In practical channels, the noise power spectral density $\boldsymbol{N_0}$ is generally constant. If $\boldsymbol{E_b}$ is the transmitted energy per bit, then we may express the average transmitted power as:

$$\boldsymbol{S = E_b C}$$

($\boldsymbol{C/B}$) is the "bandwidth efficiency" of the syste m. If $\boldsymbol{C/B} = 1$, then it follows that $\boldsymbol{E_b = N_0}$.This implies that the signal power equals the noise power. Suppose, $\boldsymbol{B = B_0}$ for which, $\boldsymbol{S = N}$, then Eq. (5.59) can be modified as:

That is, " *the maximum signaling rate for a given* ***S*** *is 1.443 bits/sec/Hz in the bandwidth over which the signal power can be spread without its falling below the noise level*".

4.7 CYCLIC CODES:

In coding theory, cyclic codes are linear block error-correcting codes that have convenient algebraic structures for efficient error detection and correction.

Let C be a linear code over a finite field $GF(q)^n$ of block length *n*. C is called a cyclic code, if for every codeword $c=(c_1,...,c_n)$ from *C*, the word $(c_n,c_1,...,c_{n-1})$ in $GF(q)^n$ obtained by a cyclic right shift of components is again a codeword. Same goes for left shifts. One right shift is equal to $n - 1$ left shifts and vice versa. Therefore the linear code C is cyclic precisely when it is invariant under all cyclic shifts.

Cyclic Codes have some additional structural constraint on the codes. They are based on Galois fields and because of their structural properties they are very useful for error controls. Their structure is strongly related to Galois fields because of which the encoding and decoding algorithms for cyclic codes are computationally efficient.

Cyclic code for correcting error:

a) For correcting single error

The cyclic codes explicitly with error detection and correction. Cyclic codes can be used to correct errors, like Hamming codes as a cyclic codes can be used for correcting single error. Likewise, they are also used to correct double errors and burst errors. All types of error corrections are covered briefly in the further subsections.

The Hamming code has a generator polynomial $g(x)=x^3+x+1$. This polynomial has a zero in Galois extension field GF(8) at the primitive element α, and all codewords satisfy .

$C(\alpha)=0$ Cyclic codes can also be used to correct double errors over the field GF(2). Blocklength will be n equal to $2^m - 1$and primitive elements $\alpha$and α^3as zeros in the GF(2^m) because we are considering the case of two errors here, so each will represent one error.The received word is a polynomial of degree n-1 given as

$$v(x) = a(x)g(x) + e(x)$$

where e(x) can have at most two nonzero coefficients corresponding to 2 errors.

Syndrome Polynomial, S(x) as the remainder of polynomial v(x) when divided by the generator polynomial g(x) i.e.

S(x)=v(x) mod g(x)= (a(x)g(x)+e(x)) mod g(x)=e(x) mod g(x) as (a(x)g(x)) mod g(x) is zero

b) For correcting two errors

Let the field elements $X_1$and X_2 be the two error location numbers. If only one error occurs then $X_2$is equal to zero and if none occurs both are zero.

Let $S_1 = v(\alpha)$ and $S_3 = v(\alpha^3)$.

These field elements are called "syndromes". Now because $g(x)$is zero at primitive elements $\alpha$and α^3, so we can write $S_1 = e(\alpha)$ and $S_3 = e(\alpha^3)$. If say two errors occur, then

$$S_1 = \alpha^i + \alpha^{i'} \text{ and } S_3 = \alpha^{3i} + \alpha^{3i'}.$$

And these two can be considered as two pair of equations in $GF(2^m)$with two unknowns and hence we can write

$S_1 = X_1 + X_2$ and $S_3 = (X_1)^3 + (X_2)^3$.

Hence if the two pair of nonlinear equations can be solved cyclic codes can used to correct two errors.

Syndrome Calculator

We know *r* (*X*) = *q*(*X*)*g*(*X*) + *S*(*X*).

We can also write *r* (*X*) = *c*(*X*) + *e*(*X*).

Rewrite the expression to:

(*X*) = *c*(*X*) + *r* (*X*)

= *c*(*X*) + *q*(*X*)*g*(*X*) + *S*(*X*)

= (*f* (*X*) + *q*(*X*))*g*(*X*) + *S*(*X*)

If error polynomial *e*(*X*) is divided by generator polynomial *g*(*X*), the remainder is the syndrome polynomial *S*(*X*).

Error Detection

r (X) = c(X) + e(X) = q(X)g(X) + S(X)

e(X) = c(X) + q(X)g(X) + S(X) = (f (X) + q(X))g(X) + S(X)

Investigate error detecting capability of cyclic code:

Assuming *e*(*X*) is a burst of length *n* - *k* or less, i.e., errors are confined to *n* - *k* or fewer consecutive positions;

e(*X*) can be expressed by *e*(*X*) = *XjB*(*X*), here *B*(*X*) is a polynomial of degree *n* – *k* – 1 or less;

Xj cannot divided by *g*(*X*), *B*(*X*) cannot divided by *g*(*X*) neither,

e(*X*) = X^{jj}*B*(*X*) is NOT divisible by *g*(*X*);

Thus the syndrome polynomial is not equal to zero.

It means that a cyclic code *Ccyc* (*n*,*k*) can detect any error burst of length *n* - *k* or less.

A cyclic code *Ccyc* (*n; k*) can also detect all the *end-around* error bursts of length *n* - *k* or less.

e = (1 1 0 0 0 0 0 0 1 0 1)

Cyclic Redundancy Check (CRC) codes:

Cyclic redundancy .check codes are extremely well suited for "error detection". The two important reasons for this statement are,

(1) they can be designed to detect many combinations of likely errors.

(2) The implementation of both encoding and error detecting circuits is practical.

Accordingly, all error detecting codes used in practice, virtually, are of theCRC -type. In ann-bit received word if a contiguous sequence of ‘b-bits’ in which the first and the last bits and any number of intermediate bits are received in error, then we say aCRC "error burst' of length ‘b’

has occurred. Such an error burst may also include an end-shifted version of the contiguous sequence.

In any event, Binary (n, k)CRC codes are capable of detecting the following error patterns:

1. All CRC error bursts of length (n-k) or less.
2. A fraction of (1 - 2 (n –k - 1)) of CRC error bursts of length (n – k + 1).
3. A fraction (1-2(n – k)) of CRC error bursts of length greater than (n – k + 1).
4. All combinations of (d min – 1) or fewer errors.
5. All error patterns with an odd number of errors if the generator polynomial g (X) has an even number of non zero coefficients.

Generator polynomials of three CRC codes, internationally accepted as standards are listed below.

All three contain (1 +X) as a prime factor. TheCRC-12 code is used when the character lengths is 6- bits. The others are used for 8-bit characters.

* CRC-12 code: $g(X) = 1 + X + X^2+X^3 + X^{11} + X^{12}$*

*CRC-16 code: $g(X) = 1 +X^2 + X^{15} + X^{16}$

*CRC-CCITT code: $g(X) = 1 + X^5 + x^{12}+ X^{26}$.

Definition 1 An (n, k) linear block code C is said to be **cyclic** if for every code word $\mathbf{c} = (c_0, c_1, \ldots, c_{n-1})$ in C, there is also a code word $\mathbf{c}' = (c_{n-1}, c_0, \ldots, c_{n-2})$ that is also in C. ($\mathbf{c}'$ is a cyclic shift of $\mathbf{c}$.) □

It will be convenient to represent our codewords as *polynomials*. The codeword

$$\mathbf{c} = (c_0, c_1, \ldots, c_{n-1})$$

is represented by the polynomial

$$c(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$$

using the obvious one-to-one correspondence. A cyclic shift can therefore be represented as follows. Observe that

$$xc(x) = c_0x + c_1x + \cdots + c_{n-1}x^n.$$

If we not take this product modulo $x^n - 1$ we get

$$xc(x) \pmod{x^n - 1} = c_{n-1} + c_0x + \cdots + x_{n-2}x^{n-1}.$$

So multiplication by x in the ring $GF(q)[x]/(x^n - 1)$ corresponds to a cyclic shift. Furthermore, any power of x times a codeword yields a codeword (apply the definition recursively), so that, for example,

$$\begin{aligned}
(c_{n-1}, c_0, c_1, \ldots, c_{n-2}) &\leftrightarrow xc(x) \\
(c_{n-2}, c_{n-1}, c_0, \ldots, c_{n-3}) &\leftrightarrow x^2c(x) \\
&\vdots \\
(c_1, c_2, \ldots, c_{n-1}, c_0) &\leftrightarrow x^{n-1}c(x)
\end{aligned}$$

where the arithmetic is done in the ring $GF(q)[x]/(x^n - 1)$. Now observe that if we take an polynomial $a(x) \in GF(q)[x]$ of the form

$$a(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$$

then

$$c(x)a(a)$$

is simply a linear combination of cyclic shifts of $c(x)$ and hence, must also be a codeword. Hence: **a cyclic code is an ideal** in $GF(q)[x]/(x^n - 1)$. Because of what we know about ideals in $GF(q)[x]/(x^n - 1)$ we can immediately make some observations about cyclic codes:

- A cyclic code has a generator polynomial $g(x)$, which is the generator of the ideal. Let the degree of g be r, where $r < n$.

- Every code polynomial in the code can be expressed as a multiple of the generator.

$$c(x) = m(x)g(x),$$

where $m(x)$ is the *message polynomial*. The degree of m is less than $n - r$.

- The generator is a factor of $x^n - 1$ in $GF(q)[x]$.

Example 1 We will consider cyclic codes of length 15 with binary coefficients. We need to find the factors of $x^n - 1$ in some field. Observe that

$$15|2^4 - 1,$$

so we are dealing in the field $GF(16)$. The conjugacy classes in $GF(16)$ are

$$\begin{aligned}
\{1\} &\leftrightarrow x + 1 \\
\{\alpha, \alpha^2, \alpha^4, \alpha^8\} &\leftrightarrow 1 + x + x^4 \\
\{\alpha^3, \alpha^6, \alpha^9, \alpha^{12}\} &\leftrightarrow 1 + x + x^2 + x^3 + x^4 \\
\{\alpha^5, \alpha^{10}\} &\leftrightarrow 1 + x + x^2 \\
\{\alpha^7, \alpha^{14}, \alpha^{13}, \alpha^{11}\} &\leftrightarrow 1 + x^3 + x^4
\end{aligned}$$

Thus

$$x^{15} - 1 = (x + 1)(1 + x + x^4)(1 + x + x^2 + x^3 + x^4)(1 + x + x^2)(1 + x^3 + x^4)$$

So: degrees 1,2,4,4,4. If we want a generator of, say, degree 9, we could take

$$g(x) = (x + 1)(1 + x + x^4)(1 + x + x^2 + x^3 + x^4)$$

If we want a generator of degree 5 we could take

$$g(x) = (x + 1)(1 + x + x^4)$$

or

$$g(x) = (x + 1)(1 + x + x^2 + x^3 + x^4)$$

In fact, in this case, we can get generator polynomials of any degree from 1 to 15. So we have codes

$$(15, 0)(15, 1), \ldots, (15, 15)$$

□

A message sequence $(m_0, m_1, \ldots, m_{k-1})$ (where $k = n - r$) corresponds to a message polynomial

$$m(x) = m_0 + \cdots + m_{k-1}x^{k-1}.$$

Then the message polynomial corresponding to m is

$$c_m(x) = m(x)g(x).$$

We can write this as

$$c_m(x) = [m_0, m_1, \ldots, m_{k-1}] \begin{bmatrix} g(x) \\ xg(x) \\ \vdots \\ x^{k-1}g(x) \end{bmatrix}$$

Taking the next step, we can go back to a matrix representation,

$$\mathbf{c}_m = [m_0, m_1, \ldots, m_{k-1}] \begin{bmatrix} g_0 & g_1 & \cdots & g_r & & & & & \\ & g_0 & g_1 & \cdots & g_r & & & & \\ & & g_0 & g_1 & \cdots & g_r & & & \\ & & & \ddots & \ddots & & \ddots & & \\ & & & & g_0 & g_1 & \cdots & g_r & \\ & & & & & g_0 & g_1 & \cdots & g_r \end{bmatrix} = \mathbf{m}G$$

So we have a linear code, and can write the generator matrix corresponding to it. Note: G is $k \times n$.

Let $h(x)$ be **parity check polynomial**, that is a polynomial such that

$$x^n - 1 = g(x)h(x).$$

Since codewords are multiples of $g(x)$, then for a codeword,

$$c(x)h(x) = m(x)g(x)h(x) = m(x)(x^n - 1) \equiv 0 \pmod{x^n - 1}.$$

We let

$$s(x) = c(x)h(x) \pmod{x^n - 1}.$$

be the **syndrome polynomial**. If $s(x)$ is identically zero, then $c(x)$ is a codeword. Now let's put this in matrix form.

$$s(x) = c(x)h(x) = \sum_{i=0}^{n-1} c_i x^i \sum_{j=0}^{n-1} h_j x^j \pmod{x^n - 1}.$$

Performing the multiplication,

$$s_k = \sum_{i=0}^{n-1} c_i h_{((k-i))_n} \qquad k = 0, 1, \ldots, n-1.$$

Writing the last $n - k$ of these out, we have

$$[s_k, s_{k+1}, \ldots, s_{n-1}] = [c_0, c_1, \ldots, c_{n-1}] \begin{bmatrix} h_k & h_{k-1} & \cdots & h_1 & h_0 & & & & & \\ & h_k & h_{k-1} & \cdots & h_1 & h_0 & & & & \\ & & \ddots & \ddots & & & & & & \\ & & & & h_k & h_{k-1} & \cdots & h_1 & h_0 & \\ & & & & & h_k & h_{k-1} & \cdots & h_1 & h_0 \end{bmatrix}^T = \mathbf{c}H^T.$$

4.8 LINEAR BLOCK CODES.

Error-Control Coding

Error-control coding techniques are used to detect and/or correct errors that occur in the message transmission in a digital communication system. The transmitting side of the error-control coding adds redundant bits or symbols to the original information signal sequence. The receiving side of the error-control coding uses these redundant bits or symbols to detect and/or correct the errors that occurred during transmission. The transmission coding process is known as *encoding*, and the receiving coding process is known as *decoding*.

There are two major classes in error-control code: block and convolutional. In block coding, successive blocks of K information (message) symbols are formed.

The coding algorithm then transforms each block into a codeword consisting of n symbols where $n>k$. This structure is called an (n,k) code. The ratio k/n is called the code rate. A key point is that each codeword is formed independently from other codewords.

An error-control code is a *linear code* if the transmitted signals are a linear function of the information symbols. The code is called a *systematic code* if the information symbols are transmitted without being altered. Most block codes are systematic, whereas most convolutional codes are nonsystematic.

Almost all codes used for error control are linear. The symbols in a code can be either binary or non-binary. Binary symbols are the familiar '0' and '1'.

Linear Block Codes

Linear block coding is a generic coding method. Other coding methods, such as Hamming and BCH codes, are special cases of linear block coding. The codeword vector of a linear block code is a linear mapping of the message vector. The codeword ***x*** and the message ***m*** have the relationship

$$\mathbf{x} = \mathbf{mG}$$

where $\mathbf{G}$ is a K-by-N matrix and is known as the *generator matrix*.

Linear block code is called a systematic linear code if the generator matrix has the form

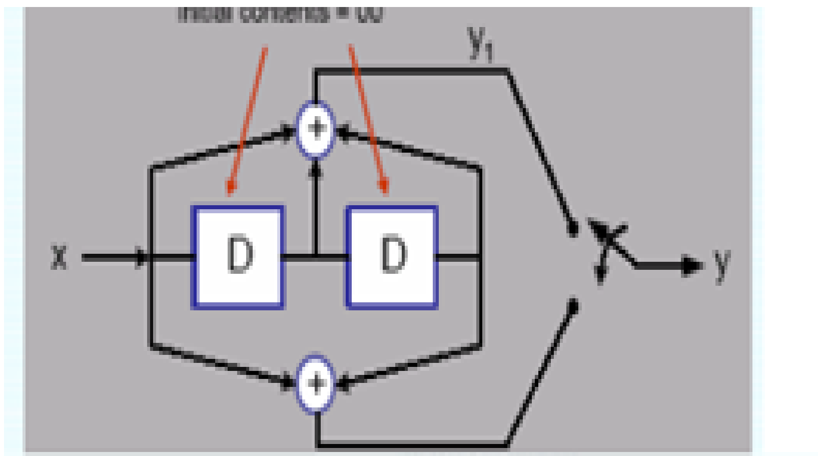
$$\mathbf{G} = [\mathbf{P} \vdots \mathbf{I}_k]$$

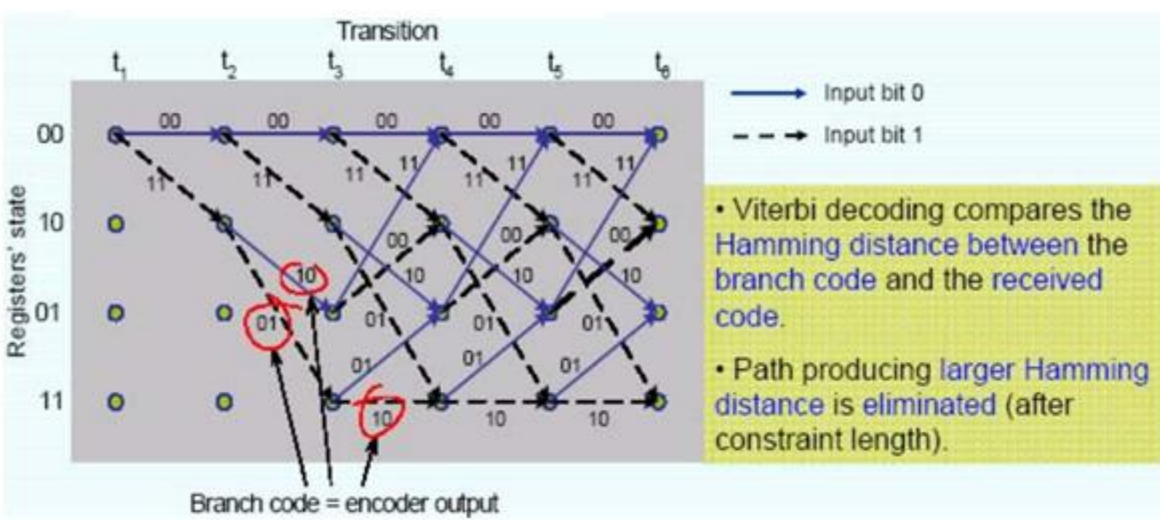
where $\mathbf{P}$ is an *(n-k)-by-k* matrix and $\mathbf{I}_k$ is a *k-by-k* identity matrix. A systematic linear code renders a length k message into a length n codeword where the last k bits are exactly the original message and the first $(n-k)$ bits are redundant. These redundant bits serve as parity-check digits.

4.9 VITERBI ALGORITHM.

ML algorithm is too complex to search all available pathes.

- End to end calculation.
- Viterbi algorithm performs ML decoding by reducing its complexity.
- Eliminate least likely trellis path at each transmission stage.
- Reduce decoding complexity with early rejection of unlike pathes.
- Viterbi algorithm gets its efficiency via concentrating on suvival paths of the trellis.



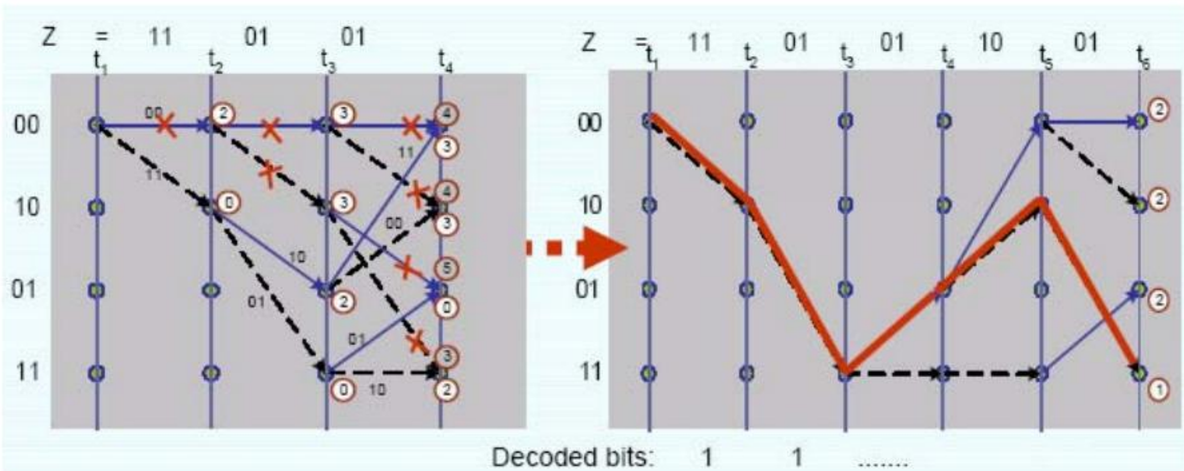


Example of viterbi Decoding:

Input data: m =1 1 0 1 1

Codeword: X = 11 01 01 00 01

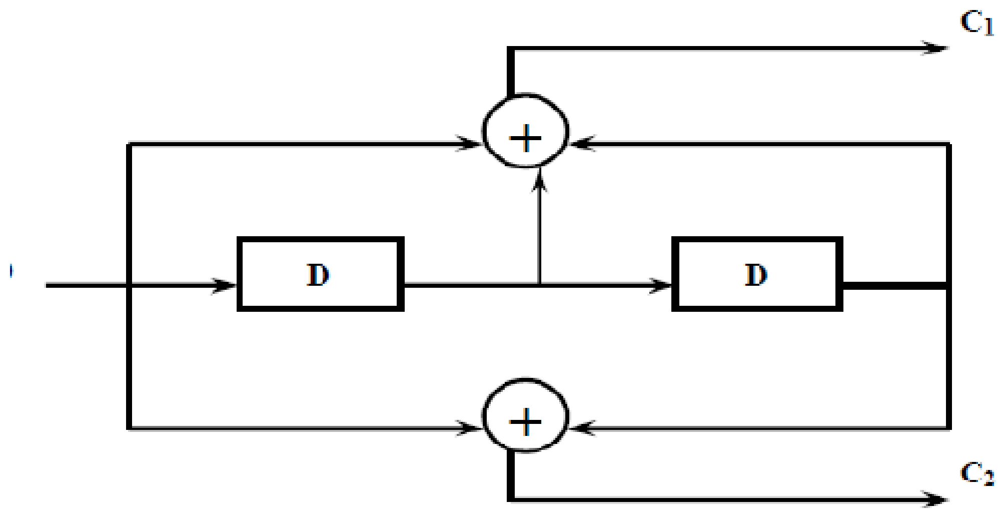
Received code: Z = 11 01 01 10 01



4.10 CONVOLUTIONAL CODES.

* Convolutional codes are widely used as channel codes in practical communication systems for error correction.
* The encoded bits depend on the current k input bits and a few past input bits.
* The main decoding strategy for convolutional codes is based on the widely used Viterbi algorithm.
* Convolutional codes are commonly described using two parameters: the code rate and the constraint length. The code rate, k/n, is expressed as a ratio of the number of bits into the convolutional encoder (k) to the number of channel symbols output by the convolutional encoder (n) in a given encoder cycle.
* The constraint length parameter, K, denotes the "length" of the convolutional encoder, i.e. how many k-bit stages are available to feed the combinatorial logic that produces the output symbols. Closely related to K is the parameter m, which can be thought of as the memory length of the encoder. A simple convolutional encoder is shown below(fig 3.1).

The information bits are fed in small groups of k-bits at a time to a shift register. The output encoded bits are obtained by modulo-2 addition (EXCLUSIVE-OR operation) of the input information bits and the contents of the shift registers which are a few previous information bits.



A convolutional encoder with k=1, n=2 and r=1/2

The operation of a convolutional encoder can be explained in several but equivalent ways such as, by

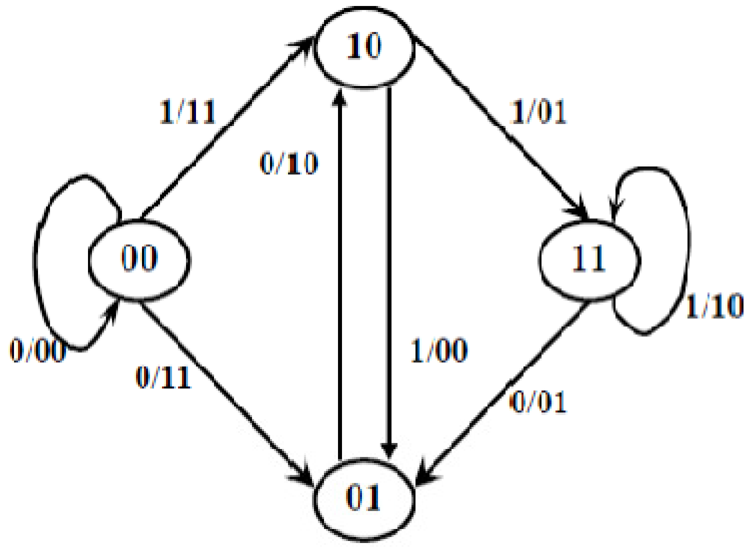
a) state diagram representation. b) tree diagramrepresentation.

c) trellis diagram representation.

a) **State Diagram Representation:** A convolutional encoder may be defined as a finite state machine. Contents of the rightmost (K-1) shift register stages define the states of the encoder. So, the encoder in **Fig. 3.1** has four states.

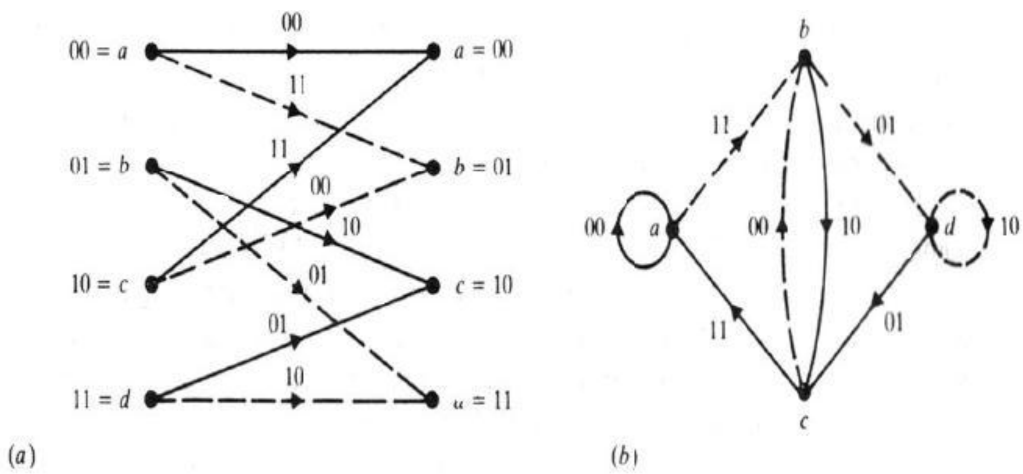
The transition of an encoder from one state to another, as caused by input bits, is depicted in the state diagram.**Fig. 3.2** shows the state diagram of the encoder in **Fig. 3.1**.

A new input bit causes a transition from one state to another.



State diagram representation for the encoder.

b) **Tree Diagram Representation:** The tree diagram representation shows all possible information and encoded sequences for the convolutional encoder. **Fig. 3.3** shows the tree diagram for the encoder in **Fig. 3.1**. The encoded bits are labeled on the branches of the tree. Given an nput sequence, the encoded sequence can be directly read from the tree. **Representing convolutional codes compactly: code trellis and state diagram:** State diagram



(a) (b)

Input	1	1	0	1	1	1	0	0	1	0	0	0	
State	a	b	d	c	b	d	d	c	a	b	c	a	a
Output	11	01	01	00	01	10	01	11	11	10	11	00	

(c)

Inspecting state diagram: Structural properties of convolutional codes:

• Each new block of k input bits causes a transition into new state

• Hence there are $2k$ branches leaving each state

• Assuming encoder zero initial state, encoded word for any input of k bits can thus be obtained. For instance, below for **u**=(1 1 1 0 1), encoded word **v**=(1 1, 1 0, 0 1, 0 1, 1 1, 1 0, 1 1, 1 1) is produced:

- encoder state diagram for (n,k,L)=(2,1,2) code - note that the number of states is $2L+1$ = 8

Distance for some convolutional codes:

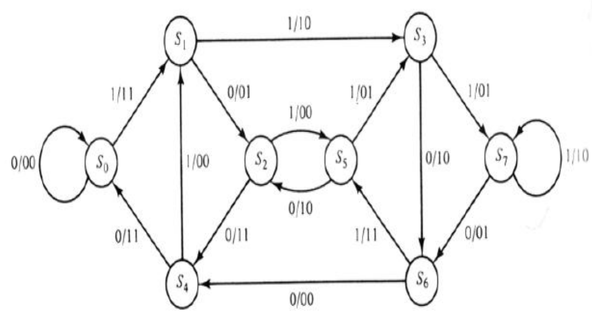


Fig. A tree diagram for the encoder

c) **Trellis Diagram Representation:**

The trellis diagram of a convolutional code is obtained from its state diagram. All state transitions at each time step are explicitly shown in the diagram to retain the time dimension, as is present in the corresponding tree diagram.

Usually, supporting descriptions on state transitions, corresponding input and output bits etc. are labeled in the trellis diagram.

It is interesting to note that the trellis diagram, which describes the operation of the encoder, is very convenient for describing the behavior of the corresponding decoder, especially when the famous „Viterbi Algorithm (VA)‟ is followed.

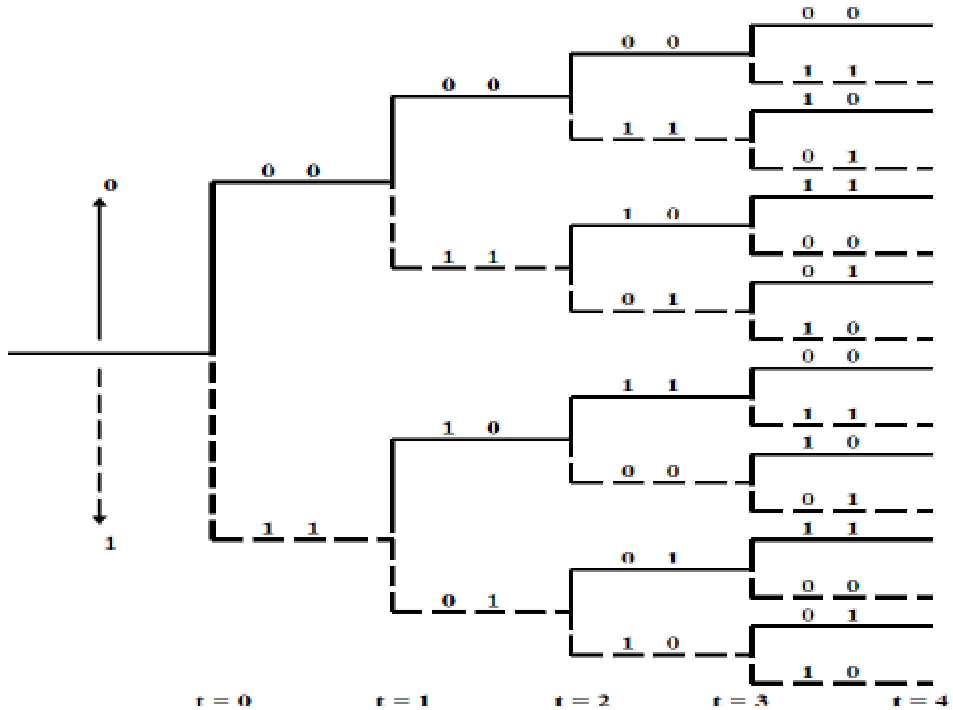


Fig.. Trellis diagram for the encoder in Fig. 3.1

1/3 rate convolutional encoder has the following generator g1=(1 0 0), g2=(1 0 1), g3=(1 1 1)

i)Draw the encoder circuit corresponding to this code (3)

ii) Draw the code tree (3) iii) Draw the state Diagram (3) v) Draw the trellis Diagram (3)

v)This code is used for transmission over a Awgn channel with hard decision decoder. Using viterbi algorithm

----

UNIT V MULTI-USER RADIO COMMUNICATION

5.1 INTRODUCTION:

Global system for mobile communication (GSM) is a globally accepted standard for digital cellular communication. GSM is the name of a standardization group established in 1982 to create a common European mobile telephone standard that would formulate specifications for a pan-European mobile cellular radio system operating at 900 MHz. It is estimated that many countries outside of Europe will join the GSM partnership.

Subscriber Identity Module (SIM). It is a memory device that stores information such as the subscribers identification number, the networks and countries where the subscriber is entitled to service, privacy tax and other user specific information A subscriber uses the SIM with a four digit personal ID number to activate service from GSM phone.

5.2 Advanced Mobile Phone Services (AMPS).

AMPS is a Standard Cellular Telephone Service (CTS). The AMPS system uses a seven cell reuse pattern with provisions for sectoring and cell splitting to increase channel when needed. AMPS uses frequency modulation and frequency division duplex for radio transmission.

The concept of AMPS.

- AMPS channel
  - Explanation
- N AMPS
  - Explanation
- Voice modulation and Demodulation
  - Compander
  - Pre-emphasis
  - Post Deviation Limiter filter.
  - Supervisory signal.
  - Division Limiter.

5.3 Cellular Concept and Frequency Reuse, Channel Assignment and Handoffs

• When a user/call moves to a new cell, then a new base station and new channel should be assigned (handoff)

• Handoffs should be transparent to users, while their number should be kept to minimum

• A threshold in the received power (Pr, handoff) should be determined to trigger the handoff process. This threshold value should be larger than the minimum acceptable received power (Pr, acceptable)

• Define: Δ=Pr,handoff - Pr,acceptable

– If Δ is large then too many handoffs

– If Δ is small then insufficient time to complete a handoff

• In order to correctly determine the beginning of handoff, we need to determine that a drop in the signal strength is not due to the momentary (temporary) bad channel condition, but it is due to the fact that the mobile is moving away from BS.

• Thus the BS needs to monitor the signal level for a certain period of time before initiating a handoff. The length of the time (running average measurements of signal) and handoff process depends on speed and moving pattern.

• First generation systems typical time interval to make a handoff was 10 seconds (large Δ). Second generations and after typical time interval to make a handoff is 1-2 seconds (small Δ).

• **First generation systems:** handoff decision was made by BS by measuring the signal strength in reverse channels.

• **Second generation and after:** Mobile Assisted Hand-Off (MAHO).

Mobiles measure the signal strength from different neighboring BSs. Handoff is initiated if the signal strength from a neighboring BS is higher than the current BS's signal strength.

Cell Dwell Time

• It is the time over which a call maybe maintained within a cell (without handoff).

• It depends on: propagation, interference, distance between BS and MS, speed and moving pattern (direction), etc.

• Highway moving pattern: the cell dwell time is ar.v. with distribution highly concentrated around the mean.

• Other micro-cell moving patterns mix of different user types with large variations of dwell time (around the mean).

Prioritizing Handoffs

• **Guard Channels:** Fraction of total bandwidth in a cell is reserved for exclusive use of handoff calls. Therefore, total carried traffic is reduced if fixed channel assignment is used. However, if dynamic channel assignment is used the guard channel mechanisms may offer efficient spectrum utilization.

– Number of channels to be reserved: If it is low (under-reservation) the QoS on handoff call blocking probability can not be met. If reservation is high (over-reservation) may result in waste of resources and rejection of large number of new calls.

– Static and Dynamic schemes: Advantage of static scheme is its simplicity since no communication and computation overheads are involved.

- However problems of under- reservation and over reservations may occur if traffic does not conform to prior knowledge.
- Dynamic schemes may adjust better to changing traffic conditions.

Prioritizing Handoffs

• **Queuing Handoffs:** The objective is to decrease the probability of forced determination of a call due to lack of available channels. When a handoff call (and in some schemes a new call) can not be granted the required resources at the time of its arrival, the request is put in a queue waiting for its admitting conditions to be met.

– This is achieved because there is a finite time interval between the time that the signal of a call drops below the handoff threshold, and the time that the call is terminated due to low (unacceptable) signal level. Queuing and size of buffer depends on traffic and QoS. Queueing in wireless systems is possible because signaling is done on separate control channels (without affecting the data transmission channels).

• According to the types of calls that are queued, queuing priority schemes are classified as: handoff call queuing, new call queuing and handoff/new call queuing (handoff calls are given non-preemptive priority over new calls).

Practical Issues (Capacity/Handoff)

• To increase capacity, use more cells (add extra sites).

• Using different antenna heights and powers, we can provide "large" and "small" cells co-located at a signal location (it is used especially to handle high speed users and low speed users simultaneously.

• Reuse partitioning (use of different reuse patterns)

• Cell splitting: Change cell radius R and keep co-channel reuse ratio (D/R) unchanged. If R'=R/2 than the transmit power needs to be changed by (1/2)4 = 1/16.

• Another way is to keep cell radius R unchanged and decrease D/R ratio required (that is decrease the number of cells in a cluster). To do this it is required to decrease interference without decreasing transmit power.

• Sectoring: Use directional antennas (instead of omni-directional) and therefore you receive interference from only a fraction of the neighboring cells.

• Hard handoffs vs. soft handoffs: more than one BSs handle the call during handoff phase (used in CDMA systems)

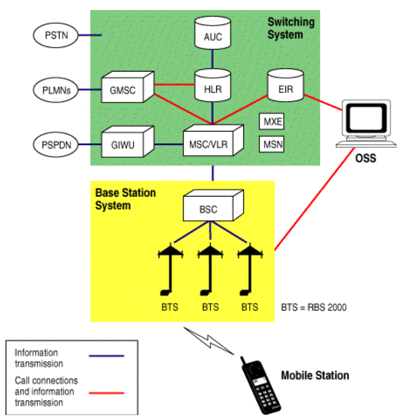
Super audio tone (SAT): SAT is superimposed on the voice signal on both the forward and reverse link and is barely audible to the user

• The particular frequency of the SAT denotes the particular base station location for a given channel and is assigned by the MSC for each call.

5.4 Global System for Mobile Communications (GSM) Code division multiple access (CDMA)

Global system for mobile communication (GSM) is a globally accepted standardfor digital cellular communication. GSM is the name of a standardization groupestablished in 1982 to create a common European mobile telephone standardthat would formulate specifications for a pan-European mobile cellular radiosystem operating at 900 MHz. It is estimated that many countries outside ofEurope will join the GSM partnership.

Throughout the evolution of cellular telecommunications, various systems havebeen developed without the benefit of standardized specifications. This presented many problems directly related to compatibility, especially with the development of digital radio technology. The GSM standard is intended to address these problems.



GSM network elements

GSM SERVICES

- Telephone Services
    - Explanation
- Data Services
    - Explanation
- Supplementary Services
    - Explanation

GSM Features:

- Explanation

GSM System Architecture:

- Base Station Subsystem (BSS)
- Network and Switching Subsystems (NSS)
- Operation Support Subsystem (OSS)

GSM interface:

- Abis interface
- A – interface

GSM channel Types:

- Traffic channels
- Control channels

Frame structure for GSM.

Different types of GSM services

- Telephone services
- Data services
- Supplementary services

Signaling tone :

The signaling tone is a 10 kbps data base which signals call termination by the subscriber. It is a special end of call message consisting of alternating 1s and 0s which is sent on the RVC by the subscriber unit for 200ms. The signaling tone alerts the base station that a subscriber has ended the call.

Telephone services in GSM: Teleservices provides communication between two end user applications according to a standard protocol. GSM mainly focuses on voice oriented tele services. This service includes emergency calling and facsimile. GSM also supports video text and tele text.

Handoff.

When a user/call moves to a new cell, then a new base station and new channel should be assigned (handoff)

• Handoffs should be transparent to users, while their number should be kept to minimum

• A threshold in the received power (Pr, handoff) should be determined to trigger the handoff process. This threshold value should be larger than the minimum acceptable received power (Pr, acceptable)

• Define: Δ=Pr,handoff - Pr,acceptable

– If Δ is large then too many handoffs

– If Δ is small then insufficient time to complete a handoff.

Features of CDMA

- Frequency reuse
- Soft capacity
- Multipath fading
- Data Rate
- Soft Handoff
- Self Jamming
- Flexibility.

5.5 SATELLITE COMMUNICATION

TYPES OF SATELLITES (BASED ON ORBITS) Geostationary or geosynchronous earth orbit (GEO)

GEO satellites are synchronous with respect to earth. Looking from a fixed point from Earth, these satellites appear to be stationary. These satellites are placed in the space in such a way that only three satellites are sufficient to provide connection throughout the surface of the Earth (that is; their footprint is covering almost 1/3rd of the Earth). The orbit of these satellites is circular.

There are three conditions which lead to geostationary satellites. Lifetime expectancy of these satellites is 15 years.

1) The satellite should be placed 37,786 kms (approximated to 36,000 kms) above the surface of the earth.

2) These satellites must travel in the rotational speed of earth, and in the direction of motion of earth, that is eastward.

3) The inclination of satellite with respect to earth must be 00.

Geostationary satellite in practical is termed as geosynchronous as there are multiple factors which make these satellites shift from the ideal geostationary condition.

1) Gravitational pull of sun and moon makes these satellites deviate from their orbit. Over the period of time, they go through a drag. (Earth‟s gravitational force has no effect on these satellites due to their distance from the surface of the Earth.)

2) These satellites experience the centrifugal force due to the rotation of Earth, making them deviate from their orbit.

3) The non-circular shape of the earth leads to continuous adjustment of speed of satellite from the earth station.

These satellites are used for TV and radio broadcast, weather forecast and also, these satellites are operating as backbones for the telephone networks.

Disadvantages of GEO:

Northern or southern regions of the Earth (poles) have more problems receiving these satellites due to the low elevation above a latitude of 60°, i.e., larger antennas are needed in this case. Shading of the signals is seen in cities due to high buildings and the low elevation further away from the equator limit transmission quality.

The transmit power needed is relatively high which causes problems for battery powered devices. These satellites cannot be used for small mobile phones. The biggest problem for voice and also data communication is the high latency as without having any handovers, the signal has to at least travel 72,000 kms.

Due to the large footprint, either frequencies cannot be reused or the GEO satellite needs special antennas focusing on a smaller footprint. Transferring a GEO into orbit is very expensive.

Low Earth Orbit (LEO) satellites:

These satellites are placed 500-1500 kms above the surface of the earth. As LEOs circulate on a lower orbit, hence they exhibit a much shorter period that is 95 to 120 minutes. LEO systems try to ensure a high elevation for every spot on earth to provide a high quality communication link. Each LEO satellite will only be visible from the earth for around ten minutes.

Using advanced compression schemes, transmission rates of about 2,400 bit/s can be enough for voice communication. LEOs even provide this bandwidth for mobile terminals with Omni-directional antennas using low transmit power in the range of 1W. The delay for packets delivered via a LEO is relatively low (approx 10 ms).

The delay is comparable to long-distance wired connections (about 5–10 ms). Smaller footprints of LEOs allow for better frequency reuse, similar to the concepts used for cellular networks. LEOs can provide a much higher elevation in Polar Regions and so better global coverage.

These satellites are mainly used in remote sensing an providing mobile communication services (due to lower latency).

Disadvantages:

The biggest problem of the LEO concept is the need for many satellites if global coverage is to be reached. Several concepts involve 50–200 or even more satellites in orbit.

The short time of visibility with a high elevation requires additional mechanisms for connection handover between different satellites.

The high number of satellites combined with the fast movements resulting in a high complexity of the whole satellite system.

One general problem of LEOs is the short lifetime of about five to eight years due to atmospheric drag and radiation from the inner Van Allen belt1. Assuming 48 satellites and a lifetime of eight years, a new satellite would be needed every two months.

The low latency via a single LEO is only half of the story. Other factors are the need for routing of data packets from satellite to if a user wants to communicate around the world.

Due to the large footprint, a GEO typically does not need this type of routing, as senders and receivers are most likely in the same footprint.

Medium Earth Orbit (MEO) satellites:

MEOs can be positioned somewhere between LEOs and GEOs, both in terms of their orbit and due to their advantages and disadvantages.

Using orbits around 10,000 km, the system only requires a dozen satellites which is more than a GEO system, but much less than a LEO system. These satellites move more slowly relative to the earth"s rotation allowing a simpler system design (satellite periods are about six hours).

Depending on the inclination, a MEO can cover larger populations, so requiring fewer handovers.

Links in satellite communication

1. Uplink , 2. Downlink & 3. Crosslink.

Disadvantages:

Again, due to the larger distance to the earth, delay increases to about 70 – 80 ms. the satellites need higher transmit power and special antennas for smaller footprints

The three orbits of satellite. Low Earth orbit : Medium Earth orbit & Geosynchronous Earth orbit

Visitor location register (VLR)—The VLR is a database that contains temporary information about subscribers that is needed bythe MSC in order to service visiting subscribers. The VLR is always integrated with the MSC. When a mobile station roams into a new MSC area, the VLR connected to that MSC will request data about the mobile station from the HLR. Later, if the mobile station makes a call, the VLR will have the information needed for call setup without having to interrogate the HLR each time.

Mobile services switching center (MSC)—The MSC performs the telephony switching functions of the system. It controls calls to and from other telephone and data systems. It also performs such functions as toll ticketing, network interfacing, common channel signaling, and others.

Home location register (HLR)—The HLR is a database used for storage and management of subscriptions. The HLR is considered the most important database, as it stores permanent data about subscribers, including a subscriber's service profile, location information, and activity status. When an individual buys a subscription from one of the PCS operators, he or she is registered in the HLR of that operator.

Kepler's laws of planetary motion

1. A satellite will orbit a primary body following an elliptical path
2. For equal intervals of time a satellite will sweep out equal areas in orbital plane
3. The square of the periodic time of orbit is proportional to the cube of the mean distance between the primary and the satellite.

The links in satellite communication?

i) Uplink

ii) Downlink iii) Crosslink

15. Define apogee

The point in an orbit that is located farthest from earth.

The point in an orbit that is located closest to earth .

Satellites are specifically made for telecommunication purpose. They are used for mobile applications such as communication to ships, vehicles, planes, hand-held terminals and for TV and radio broadcasting They are responsible for providing these services to an assigned region (area) on the earth. The power and bandwidth of these satellites depend upon the preferred size of the footprint, complexity of the traffic control protocol schemes and the cost of ground stations.

A satellite works most efficiently when the transmissions are focused with a desired area. When the area is focused, then the emissions don‟t go outside that designated area and thus minimizing the interference to the other systems. This leads more efficient spectrum usage.

Satellite's antenna patterns play an important role and must be designed to best cover the designated geographical area (which is generally irregular in shape). Satellites should be designed by keeping in mind its usability for short and long term effects throughout its life time.

The earth station should be in a position to control the satellite if it drifts from its orbit it is subjected to any kind of drag from the external forces.

Kepler's laws:

Kepler's first law

A satellite will orbit a primary body following an elliptical path

Kepler's second law

For equal intervals of time a satellite will sweep out equal areas in orbital plane

Kepler's third law

The square of the periodic time of orbit is proportional to the cube of the mean distance between the primary and the satellite

5.6 Bluetooth

Bluetooth is a standard developed by a group of electronics manufacturers that allows any sort of electronic equipment from computers and cell phones to keyboards and headphones to make its own connections, without wires, cables or any direct action from a user. Bluetooth is intended to be a standard that works at two levels.
