8.06.18

UNIT - I

DIGITAL PULSE MODULATION

Introduction :

Communication is the transformation of information from source to destination or from transmitter to reciever.

Modulation :

In modulation a low frequency information is transfered to high frequency carrier signal, during this one of the parameter (amplitude, frequency or phase) is changed in accordance with the instantaneous amplitude of the message signal.

Modulation
- Continuous Wave Modulation
  - Linear
    - AM
    - DSB-SC
    - SSB-SC
    - VSB-SC
  - Non-linear
    - FM
    - PM
- Pulse Modulation
  - Analog
    - PAM
    - PTM
      - PWM
      - PPM
  - Digital
    - Base Band
    - Band pass

Base Band
- PCM (pulse core modulation)
- DPCM
- DM (Delta Modulation)
- ADM

Band pass
- Binary
  - ASK
  - FSK
  - PSK
- M-ary

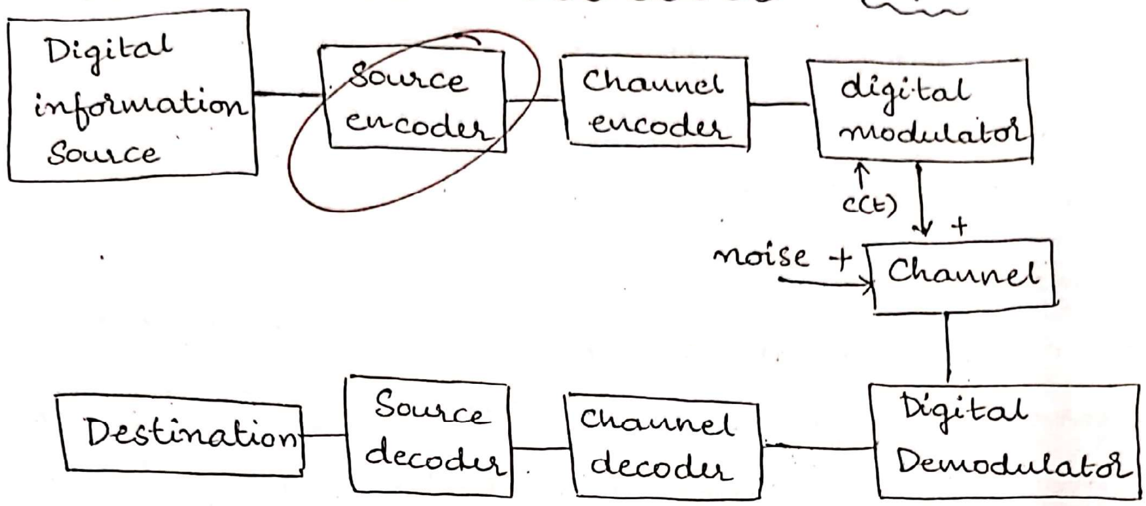
ADM - Adaptive delta Modulation

SK - Shift key

DPCM - Differential Pulse Core Modulation

PCM and DM convert pure analog signal to digital signal.

Elements in digital Communication System:



1. Information Source:

There are two types of sources. They are

(i) Digital Source

(ii) Analog Source

In digital communication the data is to be transmitted in digital form.

The analog sources are Speech Signal from microscope (or) a vedio signal captured in a vedio camera can be converted into digital or binary by A to D convertor involves Sampling and quantization.

Digital Source:

It is a binary source i.e, sequence of bit.

Eg: English text can be converted into binary file by using ASCII code.

2. Source Encoder or Decoder:

⇒ The Source encoder compresses the data into minimum no. of bits. This process helps in effective utilization of bandwidth. It removes redendend bits (unnecessary bits)

⇒ The redendency in the data is reduced by using some variable length coding techniques like arithematic coding, Shanan--fano coding, huffman coding.

⇒ At the reciever Side the Source decoder converts the binary output into a Symbol Sequence.

3. Channel Encoder or Decoder:

The channel encoder does the coding ~~(after)~~ for error control. During the transmission of the Signal due to noise in the channel the Signal may get altered and hence to avoid this the channel encoder adds some redendend bits to the transmitted data.

At the channel decoder error detection and correction is possible with the help of parity bits.

4. Digital Modulator or Demodulator:

The modulator accepts a bitstream as its input and converts into a electrical wave form which is Suitable for transmission over communicatio Channels.

Some of digital modulation techniques are

→ ASK (Amplitude Shift key)

→ FSK (Frequency Shift key)

→ PSK (Phase Shift key)

At the demodulator side we can convert a bandpass signal into baseband signal.

5. Communication Channel:

The communication channel provides the electrical connection between the transmitter and reciever.
The channel is of two types.
(i) Wired channel Eg: Co-axial cable, optical fibre etc,
(ii) Wireless Channel Eg: free space (medium)

Advantages of Digital Communication System:

1. Digital communication systems are simpler and cheaper compare to analog systems to advancements in digital IC technologies.

2. Using multiplexing the speech, vedio and the other data can be merge and transmitted over common channel.

3. High level of privacy due to data encryption (very useful in military applications).

4. Easy to detect and correct errors due to channel coding or decoding

5. Great quality of service provided by digital communication system over analog communication system because it is less effected by noise and for long distance communication can effectively use regenerated repeaters.

6. In digital communication by using different compression techniques, we will able to reduce the required bit rate which inturn reduces bandwidth

7. Digital communication is adaptive to other advanced branches of data processing such as digital signal processing, image signal processing, data compression etc.,

Disadvantages of Digital Communication System: Make B.W

1. Because of analog to digital ~~commu~~ conversion the data rate becomes high. Hence more transmission **bandwidth** is required for digital communication.
2. Synchronization is required.
3. High power consumption (due to various conversion stages is used)
4. Sampling error / Quantization error is introduced.

Sampling theorem:

Sampling theorem —
- low pass signal
- Bandpass signal.

Sampling theorem for low pass signal:

According to sampling theorem, any band limited continuous message signal m(t) having highest frequency component ω_m Hz then it can be sampled and if these samples are transmitted at reciever original signal can be reconstructed provided that sampling frequency $\omega_s \geq 2\omega_m$

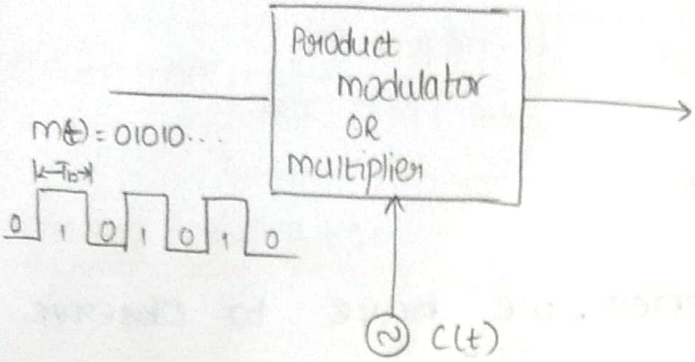
Proof:

message signal
band limited

Amplitude shift keying (ASK):-

* In this modulation, binary '1' is represented by presence of carrier and binary '0' by absence of carrier.

ASK transmitter:-



$1 \rightarrow s_1(t) = A_c \cos 2\pi f_c t$

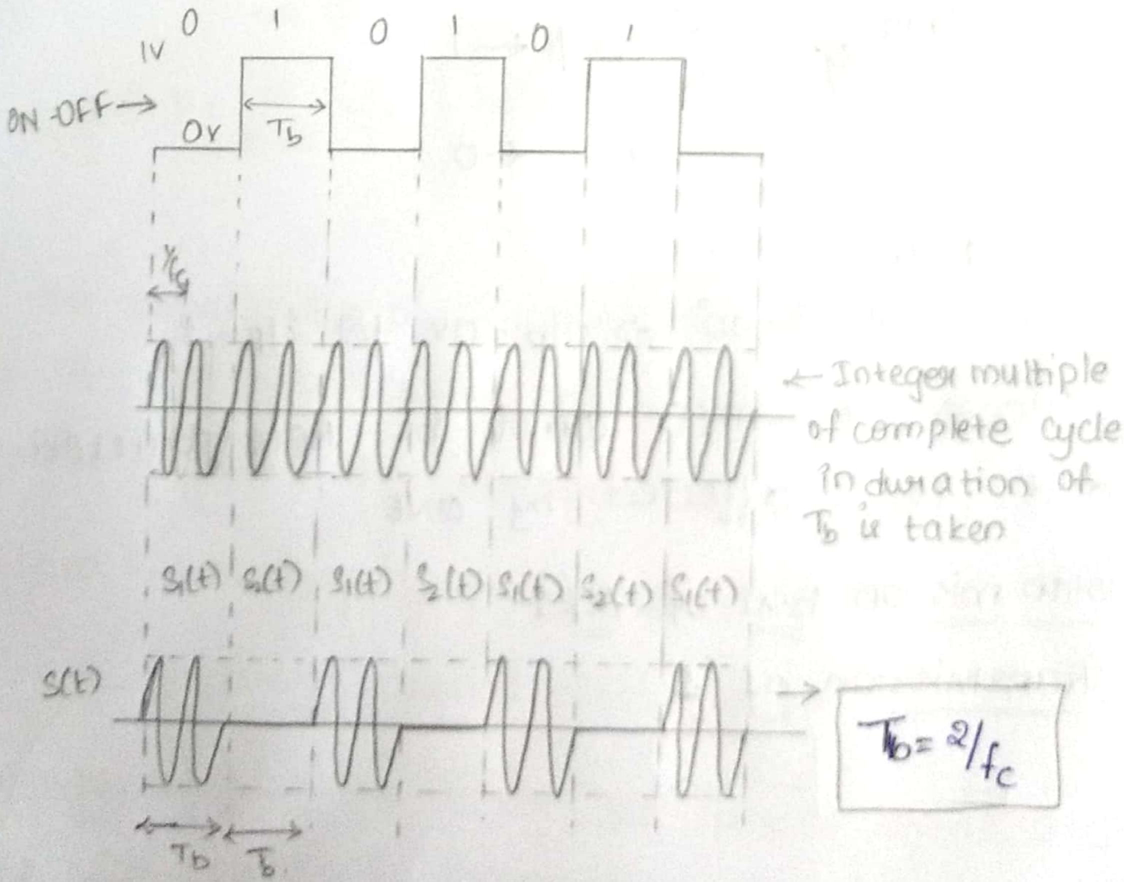
$0 \rightarrow s_2(t) = 0$

* Electrical signal representation technique is ON-OFF coding in which

$1 \rightarrow +ve$

$0 \rightarrow -ve$.

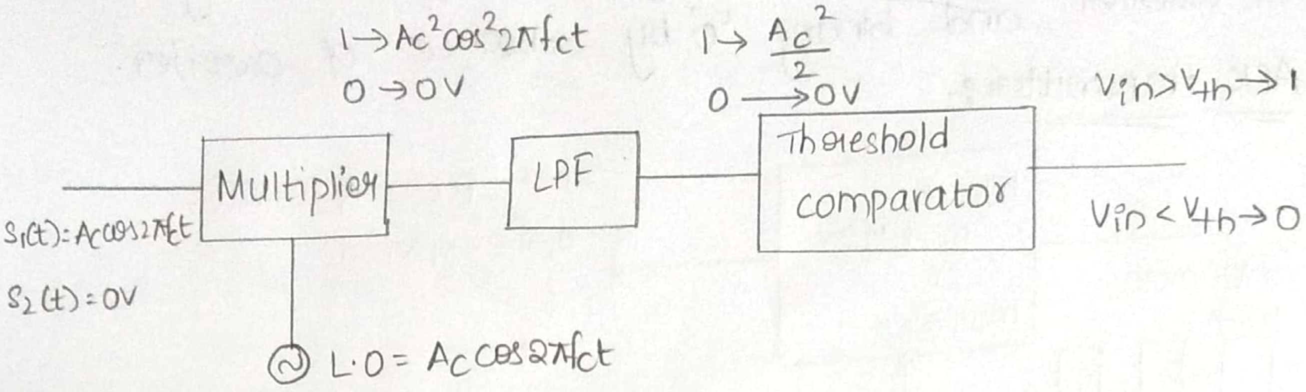
Graphical representation:-



$\boxed{T_b = 2/f_c}$

ASK Receiver:-

For demodulation of ASK, SD will be used.



Since, SD is used hence, we have to observe that either ONE occurs or not.

For this let

$$(L.O)_{o/p} = A_c \cos(2\pi f_c t + \phi)$$

So, $(mul)_{o/p} = A_c^2 \cos^2 2\pi f_c t \cos(2\pi f_c t + \phi) \leftarrow 1$

$0 \leftarrow 0$

$$(LPF)_{o/p} = \frac{A_c^2}{2} \cos\phi \leftarrow 1$$

$0 \leftarrow 0$

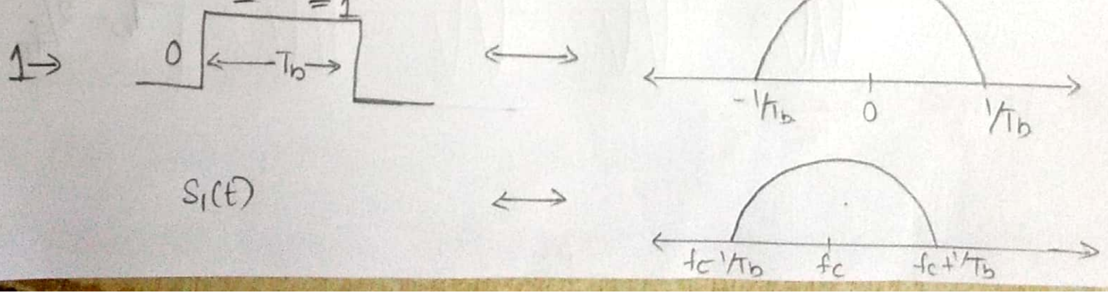
Now,

if $\phi = 90^\circ \Rightarrow o/p = 0V$ for $i/p = 1$

Hence, the $m(t)$ cannot be Reconstructed back and it is affected by ONE.

Transmission B.W of ASK:-

Transmission of '1':



Transmission of '0':

$$S_0(t) = 0$$

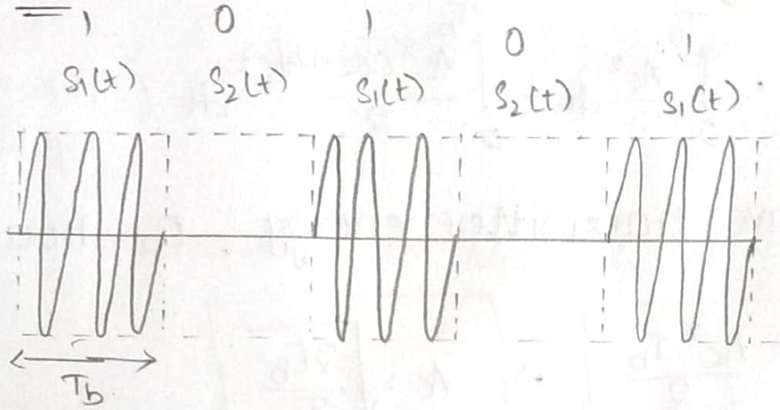
So,

$$\text{ASK B.W} = (f_c + 1/T_b) - (f_c - 1/T_b)$$

$$= 2/T_b$$

$$\boxed{\text{ASK B.W} = 2R_b}$$

Energy per Bit:-



Now,

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

$$\& \; P = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |x^2(t)|\, dt = \frac{E}{T}\Bigg|_{T\to\infty} \quad \left\{ \begin{array}{l} \text{when} \\ T=\infty \end{array} \right\}$$

Note:-

* When the signal is of infinite duration, then the energy of that particular signal is above '∞'.
* For finite signals the energy is also finite.
* $0 < E < \infty \rightarrow$ energy signal (power $= 0$)

  $0 < P < \infty \rightarrow$ power signal (Energy $= \infty$).
* An energy signal can never be power signal and also the vice-versa is correct
* All periodic signals are power signals and its energy is

Transmission of '1':

$$E_b = \int_0^{T_b} s_1^2(t)\,dt$$

$$= \int_0^{T_b} (A_c \cos 2\pi f_c t)^2\,dt$$

$$= \int_0^{T_b} A_c^2 \cos^2 2\pi f_c t\,dt$$

$$= \int_0^{T_b} \frac{A_c^2}{2}\,dt + \cancel{\int_0^{T_b} \frac{A_c^2 \cos 4\pi f_c t}{2}}\,dt$$ (since complete cycles)

Area=0

To save transmitter energy, E_b should be small.

So, $\boxed{E_b = \frac{A_c^2 T_b}{2}} \Rightarrow \boxed{A_c = \sqrt{\frac{2E_b}{T_b}}}$

Transmission of '0':

$$E_b = \int_0^{T_b} s_2^2(t)\,dt$$

$$\boxed{E_b = 0}$$

Constellation Diagram:-

$$1 \rightarrow s_1(t) = A_c \cos 2\pi f_c t = \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_c t$$

$$0 \rightarrow s_2(t) = 0$$

As $E_b = \int_0^{T_b} s_1^2(t)\,dt = \int_0^{T_b} \left\{ \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_c t \right\}^2 dt$

$$E_b = E_b \int_0^{T_b} \left\{ \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t \right\}^2 dt$$

So, $\int_0 \left\{ \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t \right\} dt = 1$

So,

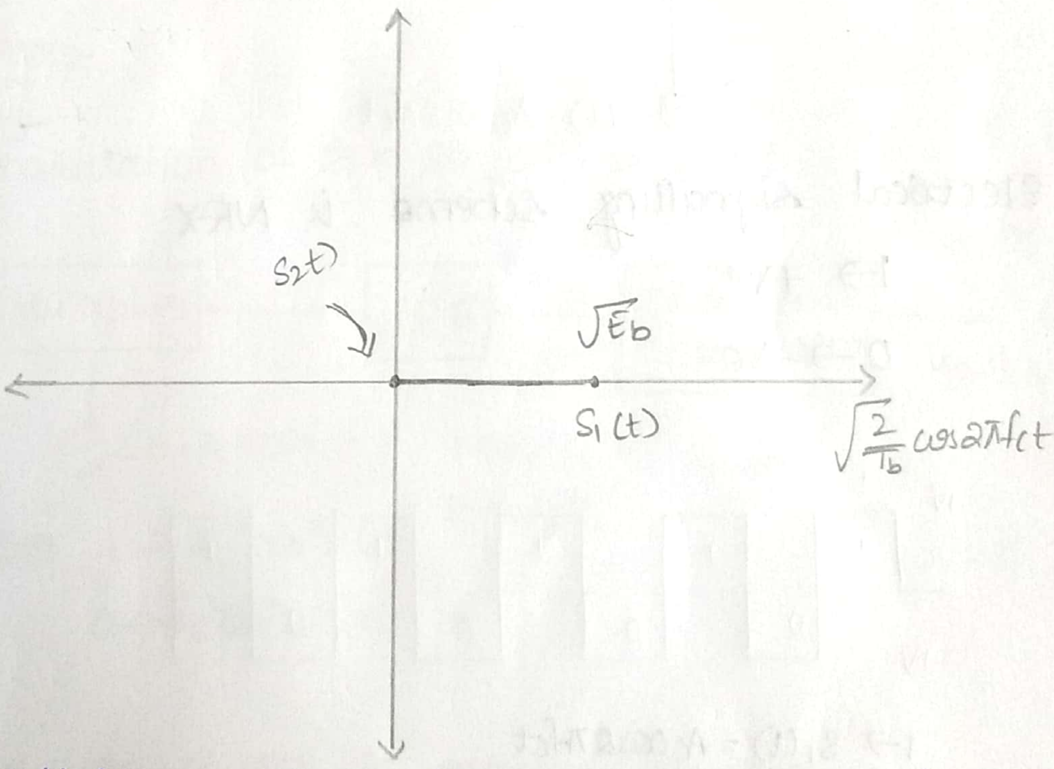
$$\boxed{\text{Energy}\left\{\sqrt{\frac{2}{T_b}}\cos 2\pi f_c t\right\} = 1}$$

* In constellation diagram the function whose energy is equal to 1 is said to be as Normalized function.

Now,

$$s_1(t) = \sqrt{E_b}\underbrace{\sqrt{\frac{2}{T_b}}\cos 2\pi f_c t}_{f(t)} \; ; \quad s_2(t) = 0$$

* each axis corresponds to Normalized function.



In constellation diagram the reference axis corresponds to Normalized functions.

Conclusion:-

* distance of $s_1(t)$ from the origin $= \sqrt{E_b}$

$\boxed{\text{Energy}\{s_1(t)\} = (\sqrt{E_b})^2 = E_b}$ { square of distance }

* Distance of $s_2(t)$ from origin $= 0$

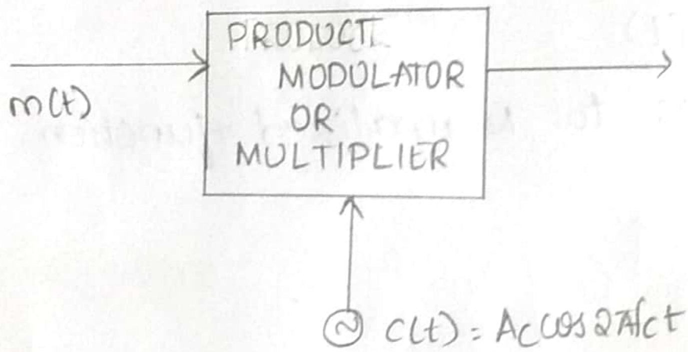
$$\boxed{\text{Energy}\{s_2(t)\} = 0}$$

* Distance btw signalling points $\boxed{d_{12} = \sqrt{E_b}}$

Phase shift keying (PSK):-

* In PSK, Binary 1 is represented by Actual carrier and Binary '0' by 180° phase shift of carrier.

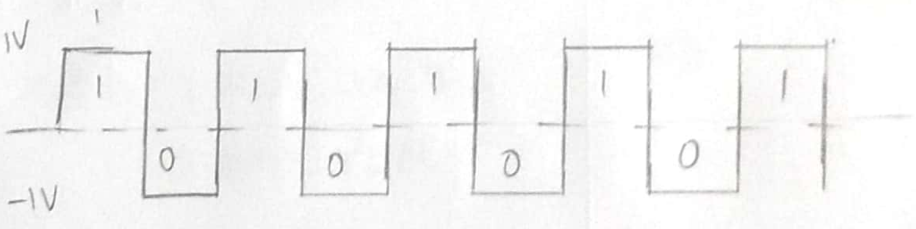
PSK Transmitter:-



* Electrical signalling scheme is NRZ.

$1 \rightarrow +ve$

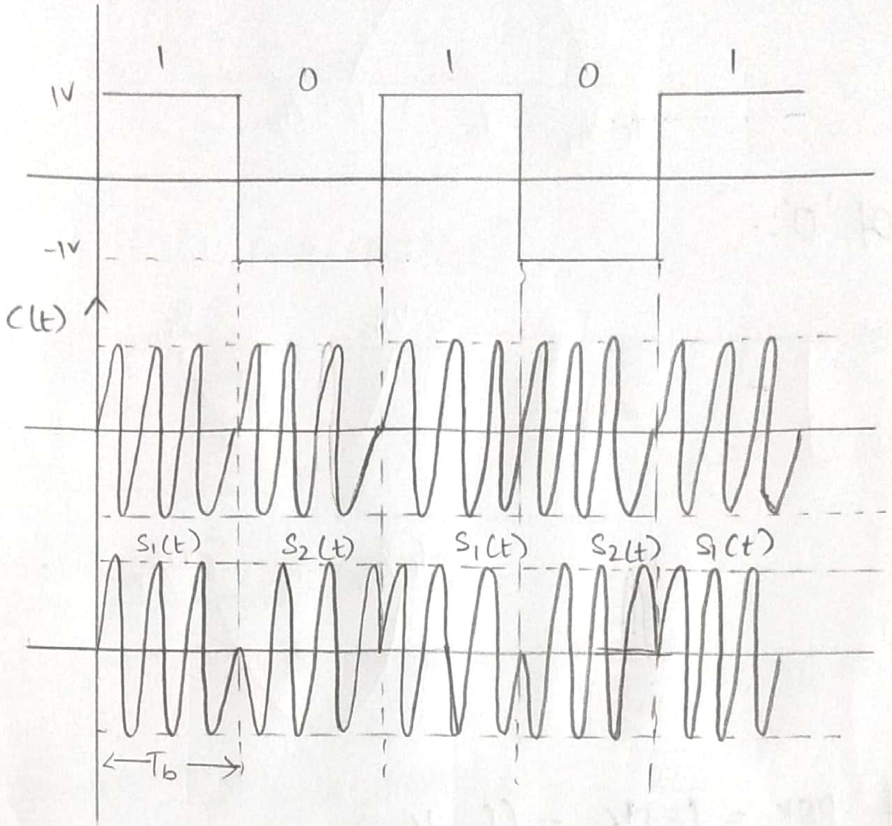
$0 \rightarrow -ve$.



$1 \rightarrow s_1(t) = A_c \cos 2\pi f_c t$

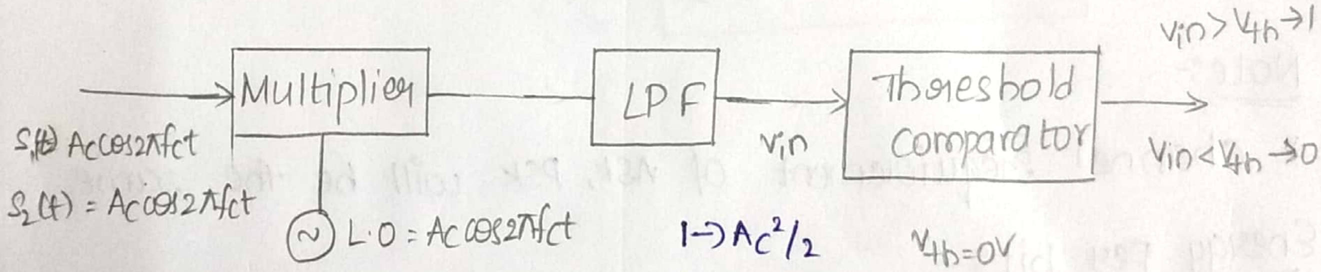
$0 \rightarrow s_2(t) = -A_c \cos 2\pi f_c t = A_c \cos\{2\pi f_c t + 180°\}$

Graphical Representation:-



PSK Receiver:-

* For demodulation of PSK, SD will be used.



$1 \to A_c^2/2$

$0 \to -A_c^2/2$

$V_{th} = 0V$

$1 \to A_c^2\cos^2 2\pi f_c t$

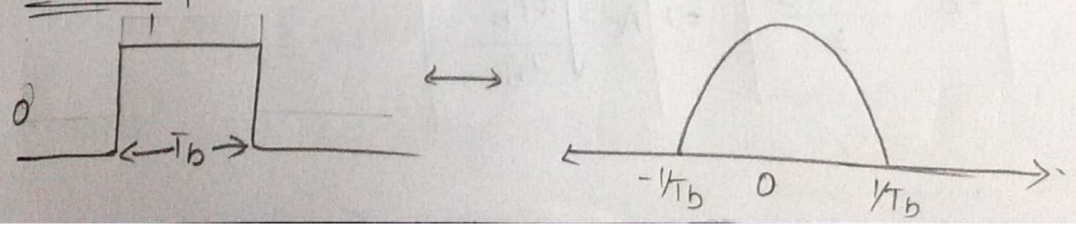
$0 \to A_c^2\cos^2 2\pi f_c t$

Note:-

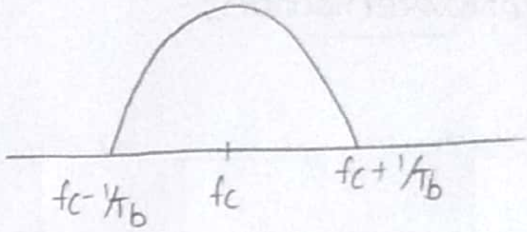
* Demodulation of PSK is affected by QNE.

Bandwidth of PSK:-

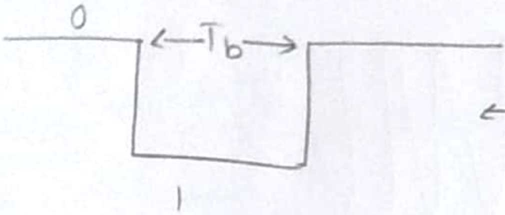
a) Transmission of '1':



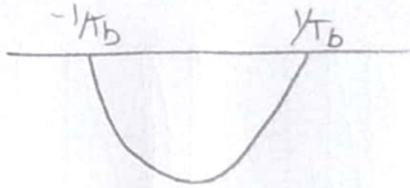
$S_1(t) \longleftrightarrow$



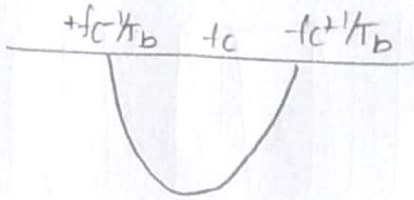
Transmission of '0':-



$\longleftrightarrow$



$S_2(t) \longleftrightarrow$



So, BW of PSK $= f_c + 1/T_b - (f_c - 1/T_b)$

$$= 2/T_b = 2R_b$$

$$\boxed{\text{BW of PSK} = 2R_b}$$

Note:-

* channel requirement of ASK, PSK will be the same.

Energy per bit:-

a) Transmission of '1':-

$$E_b = \int_0^{T_b} S_1^2(t)\, dt$$

$$= \int_0^{T_b} \{A_c \cos 2\pi f_c t\}^2\, dt$$

$$\boxed{E_b = \frac{A_c^2 T_b}{2}} \Rightarrow \boxed{A_c = \sqrt{\frac{2E_b}{T_b}}}$$

(b) Transmission of '0':.

$$E_b = \int_0^{T_b} (s_2(t))^2 \, dt$$

$$= \int_0^{T_b} (-A_c \cos 2\pi f_c t)^2 \, dt$$

$$\boxed{E_b = \frac{A_c^2 T_b}{2}} \Rightarrow \boxed{A_c = \sqrt{\frac{2E_b}{T_b}}}$$

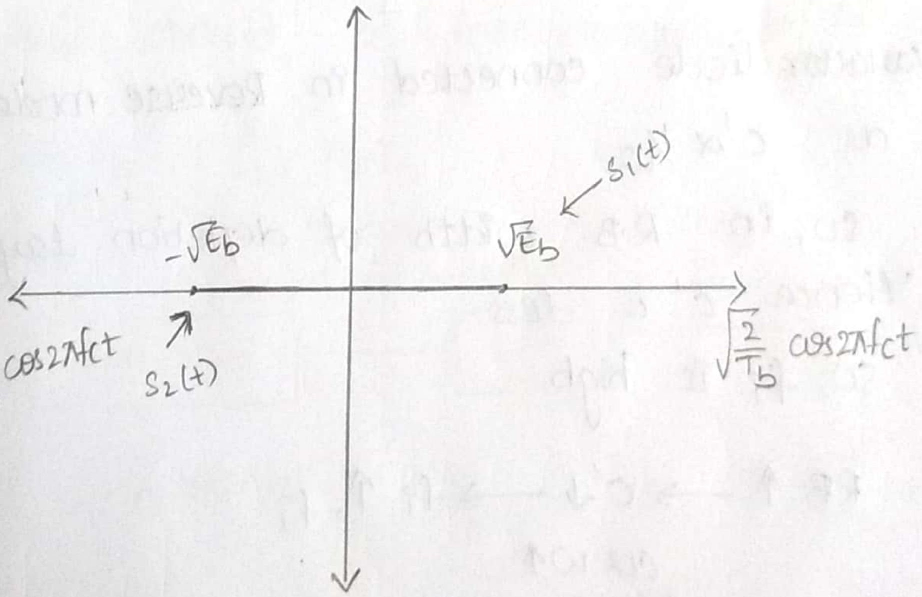
Constellation Diagram:-

$$1 \leftarrow s_1(t) = A_c \cos 2\pi f_c t = \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_c t$$

$$0 \leftarrow s_2(t) = -A_c \cos 2\pi f_c t = -\sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_c t$$

So, $$s_1(t) = \sqrt{E_b} \cdot \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t$$

$$s_2(t) = -\sqrt{E_b} \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t$$



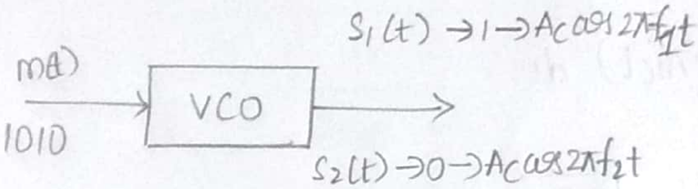
So, energy $\{s_1(t)\} = (\sqrt{E_b})^2 = E_b$

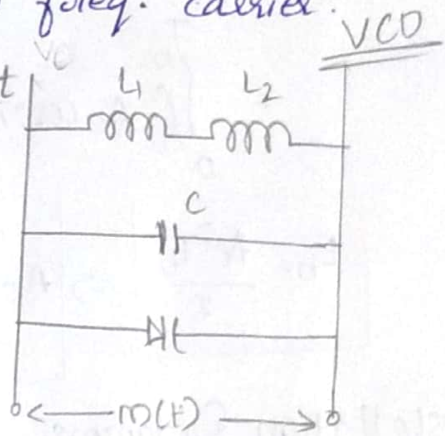
energy $\{s_2(t)\} = (-\sqrt{E_b})^2 = E_b$

* Distance btw signalling points, $\boxed{d_{12} = 2\sqrt{E_b}}$

Frequency Shift keying (FSK):-

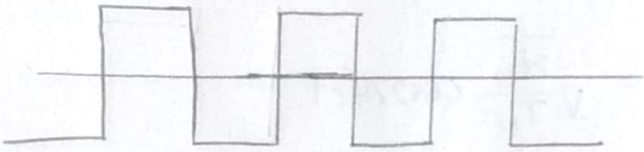
* In the binary 1 is represented by high frequency carrier & binary 0 by low freq. carrier.





NRZ ⇒ 1 = +ve

0 = -ve

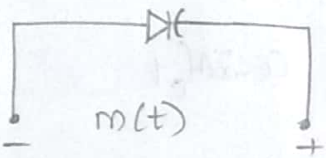


$$\boxed{f_i = \frac{1}{2\pi\sqrt{(L_1+L_2)(C+C')}}}$$

(i) Transmission of '1':-

m(t) is +ve

then



Varactor diode connected in Reverse mode.

And, as $C' \propto 1/\omega$

so, in R.B width of depletion layer is high. Hence c' is less.

so, f_i is high

$RB\uparrow \longrightarrow C'\downarrow \longrightarrow f_i\uparrow = f_1$

as $\omega\uparrow$

(ii) Transmission of '0':-

m(t) is -ve

$F.B\uparrow \to C'\uparrow \to f_i\downarrow = f_2$

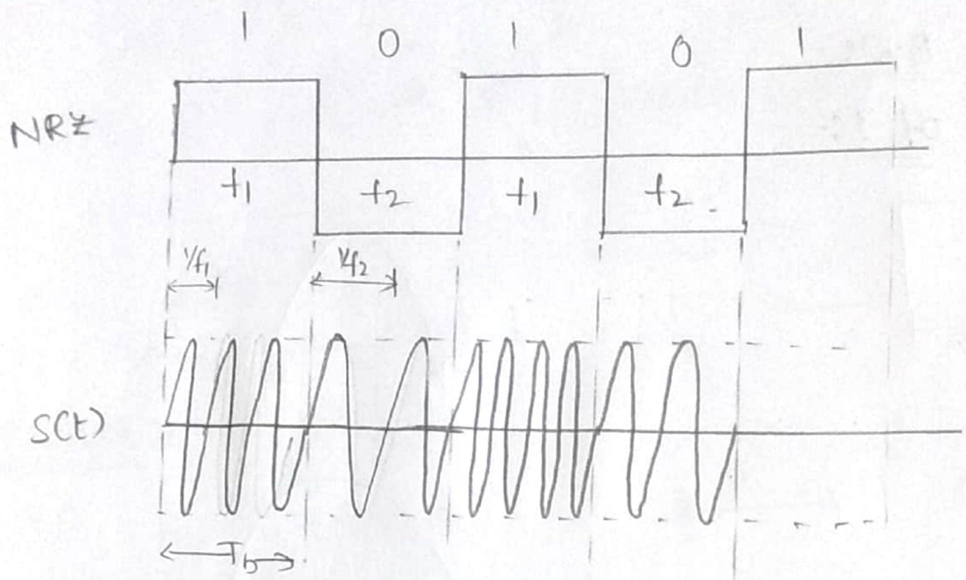
as $\omega\downarrow$

$\Rightarrow f_1 >> f_2$

as $f_1 >> f_2$, then also both the frequencies should be in the Range of MHz

Scanned by CamScanner

Graphical Representation:-



$T_b = 4/f_1$; $T_b = 2/f_2$

So, In general

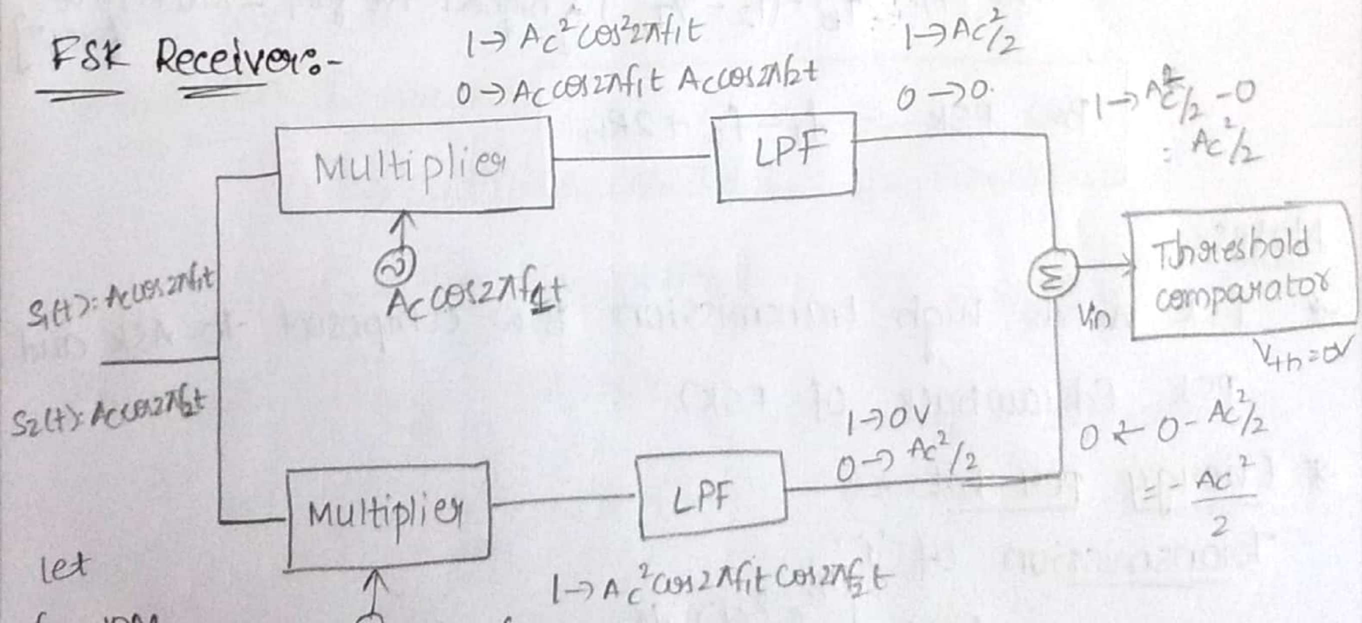
$$T_b = \frac{n_1}{f_1} \; ; \; T_b = \frac{n_2}{f_2}$$

$$\boxed{f_1 = \frac{n_1}{T_b}} \; ; \; \boxed{f_2 = \frac{n_2}{T_b}}$$

f_1 & f_2 should be integer multiple of the Bit rate

$f_1 = n_1 T_b$; $f_2 = n_2 T_b$

FSK Receiver:-



let

$f_1 = 10M$

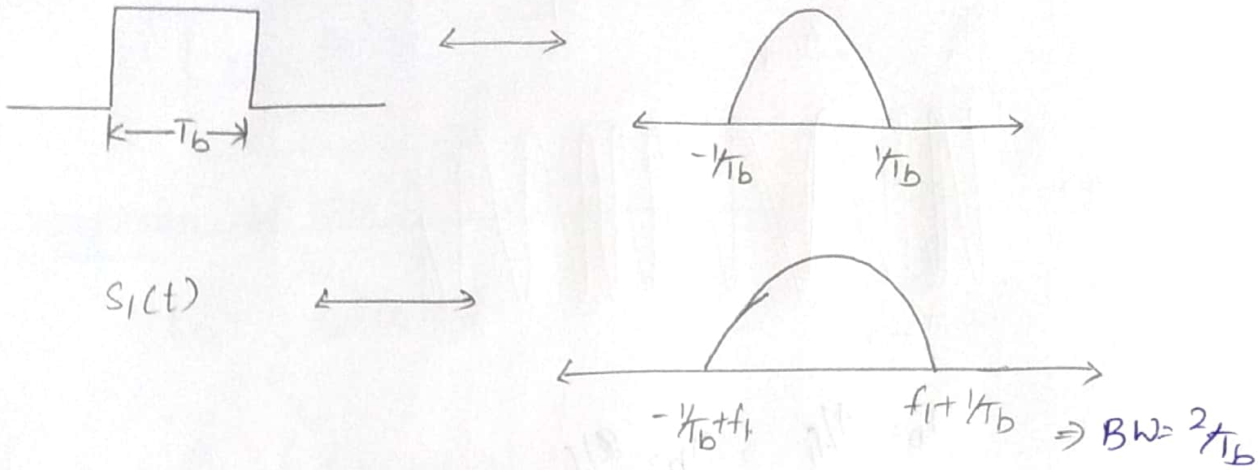
$f_2 = 2M$

Note:-

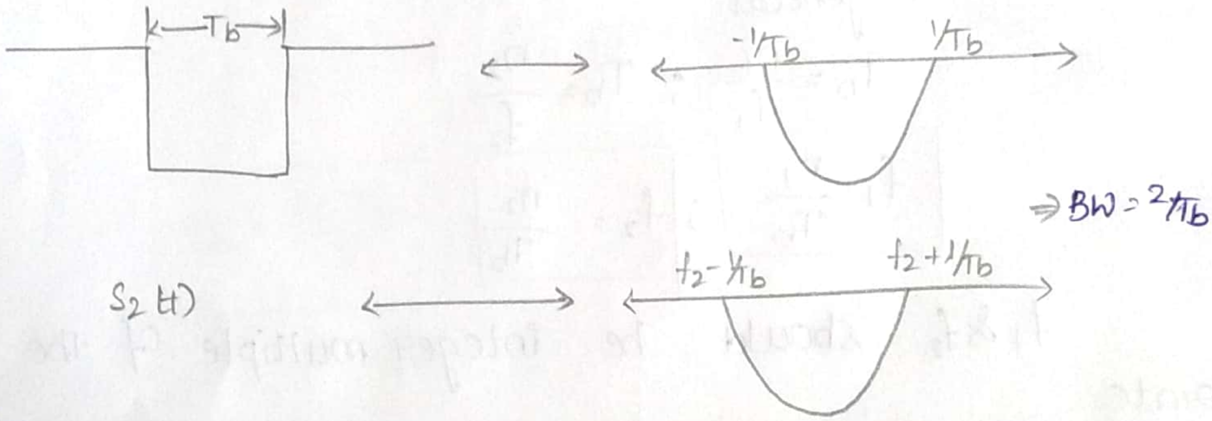
The demodulation of FSK is affected by QNE.

Transmission B.W:-

Transmission of 1:-



Transmission of '0':.



So,
the FSK $BW = (f_1 + 1/T_b) - (f_2 - 1/T_b)$ {highest +ve freq^n – lowest +ve freq^n}

$$\boxed{BW\ FSK = f_1 - f_2 + 2R_b}$$

Note:-

* FSK needs high transmission B.W composed to ASK and PSK (drawback of FSK).

* Energy per bit:.

Transmission of '1':

$$E_b = \int_0^{T_b} S_1^2(t)\,dt$$

$$= \int_0^{T_b} (A_c \cos 2\pi f_1 t)^2 \, dt$$

$$= \int_0^{T_b} \frac{A_c^2}{2} \, dt + \int_0^{T_b} \frac{A_c^2}{2} \cancel{\cos 4\pi f_1 t} \, dt \quad \{ \because \text{complete cycle} = \text{Area} \}$$

$$\boxed{E_b = \frac{A_c^2 T_b}{2}}$$

Transmission of '0':-

$$E_b = \int_0^{T_b} s_2^2(t) \, dt$$

$$= \int_0^{T_b} (A_c \cos 2\pi f_2 t)^2 \, dt$$

$$= \int_0^{T_b} \frac{A_c^2}{2} + \int_0^{T_b} \frac{A_c^2}{2} \cancel{\cos 4\pi f_2 t} \, dt \quad \{ \because \text{complete cycle} = \text{Area} \}$$

$$\boxed{E_b = \frac{A_c^2 T_b}{2}}$$

Notes:-

* Transmitter energy requirements of PSK and FSK will be the same.

Constellation Diagram:-

$$1 \rightarrow s_1(t) = A_c \cos 2\pi f_1 t = \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_1 t$$

$$0 \rightarrow s_2(t) = A_c \cos 2\pi f_2 t = \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_2 t$$

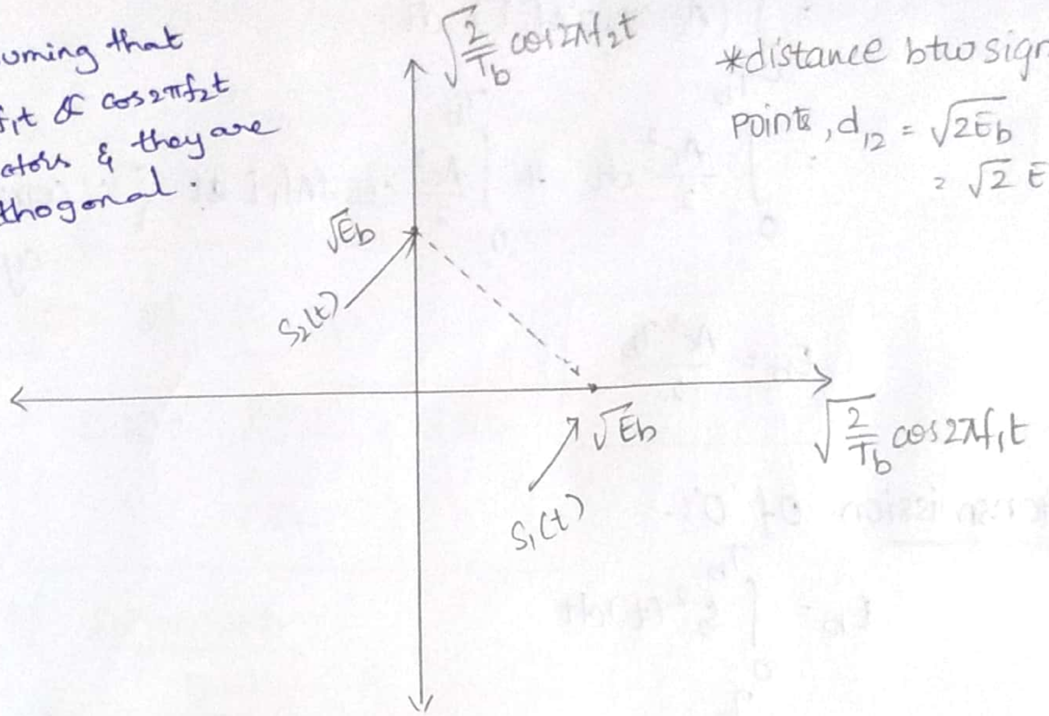
So, $s_1(t) = \sqrt{E_b} \sqrt{\frac{2}{T_b}} \cos 2\pi f_1 t$

$s_2(t) = \sqrt{E_b} \sqrt{\frac{2}{T_b}} \cos 2\pi f_2 t$

} In terms of Normalized functions.

By assuming that $\cos 2\pi f_1 t$ & $\cos 2\pi f_2 t$ are vectors & they are orthogonal.

*distance btw signalling points, $d_{12} = \sqrt{2E_b} = \sqrt{2}\, E_b$.

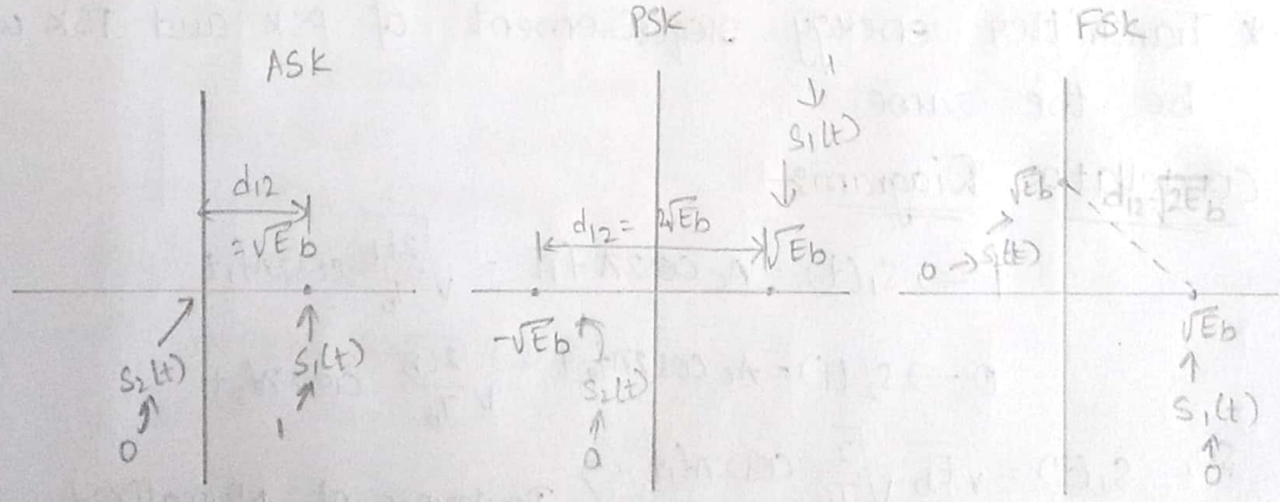


* We can't locate $\cos 2\pi f_2 t$ as all the axis correspon-ds to f_1 freq^n.

* $\sqrt{\frac{2}{T_b}} \cos 2\pi f_1 t$ and $\sqrt{\frac{2}{T_b}} \cos 2\pi f_2 t$ are orthogonal functns ins the interval $(0, T_b)$

* By interpreting these functions as vectors, the phase angle b/w resulting vectors will be 90°.

conclusion:-



In a constellation diagram if the distance b/w signalling points is less; then P_e will be more and vice versa P_e d

P_e depends upon the distance b/w signalling points

so,

$$\boxed{Pe: ASK > FSK > PSK}$$

<u>Comparision of B·W:-</u>

$$\boxed{B \cdot W:- \; ASK \;\&\; PSK < FSK}$$

<u>Usage of schemes:-</u>

	B·W	Pe
ASK	✓	✗
FSK	✗	✓ (moderate)
PSK	✓	✓

<u>Note:-</u>

PSK is much preferred signalling scheme compared to ASK and FSK.

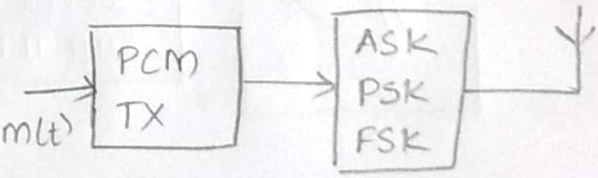
Q1. A message signal of $8\cos 8\pi\times10^3 t$ is given to 10-bit-PCM system the resulting PCM signal is transmitted through free space, by using Band Pass modulation scheme. Find the TX signal B·W. if modulation schemes is

(a) ASK (b) PSK (C) FSK with $f_H = 2MHz$, $F_L = 1MHz$.

<u>sol:</u> Given $m(t) = 8\cos 8\pi\times10^3 t$

$A_m = 8$, $f_m = 4k$

$n = 10$



∵ Sampling rate is not given

so, $f_s = NR = 2f_m = 8k$

so, $R_b = nf_s = 10\times 8k$

$= 80\,kbps$

So, for ASK,

$$\boxed{BW = 2R_b = 160k};$$

for PSK

$$\boxed{BW = 160K}$$

for FSK

$$BW = f_H - f_L + 2R_B$$

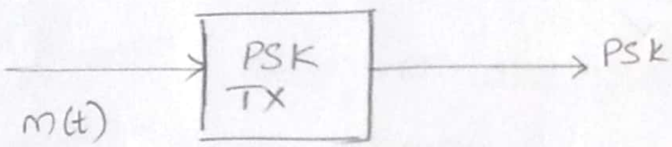
$$= (2-1)M + 160k$$

$$\boxed{B.W = 1.16M}$$

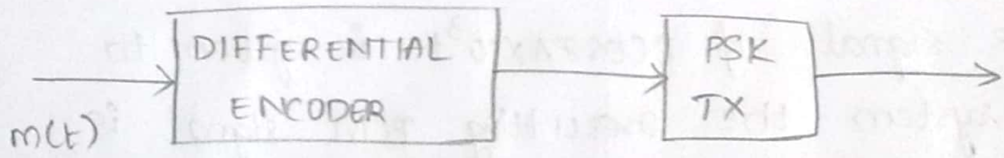
Differential phase shift keying (DPSK):-

* The advantage of DPSK over PSK is NO ONE.

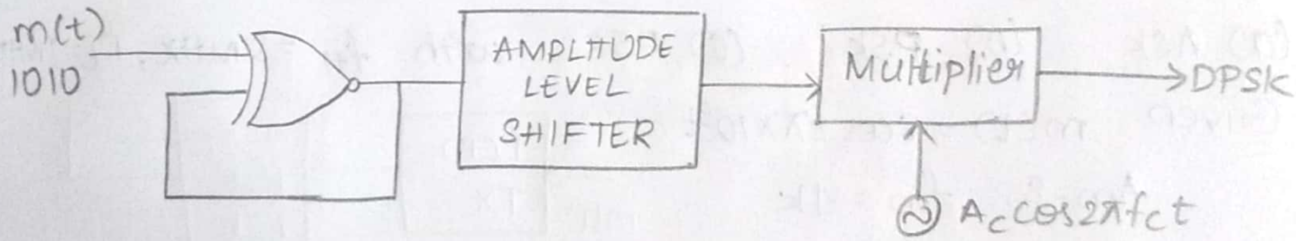
PSK:-

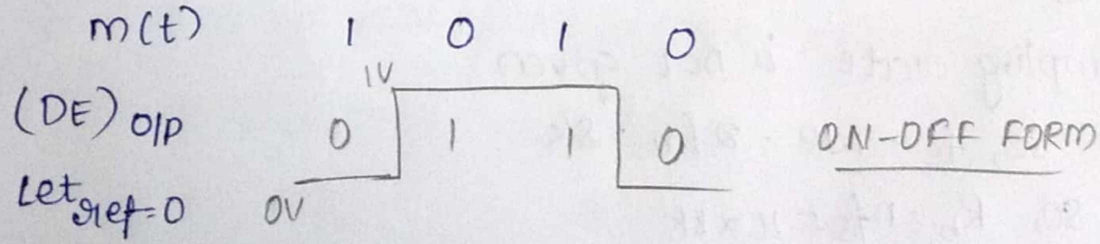


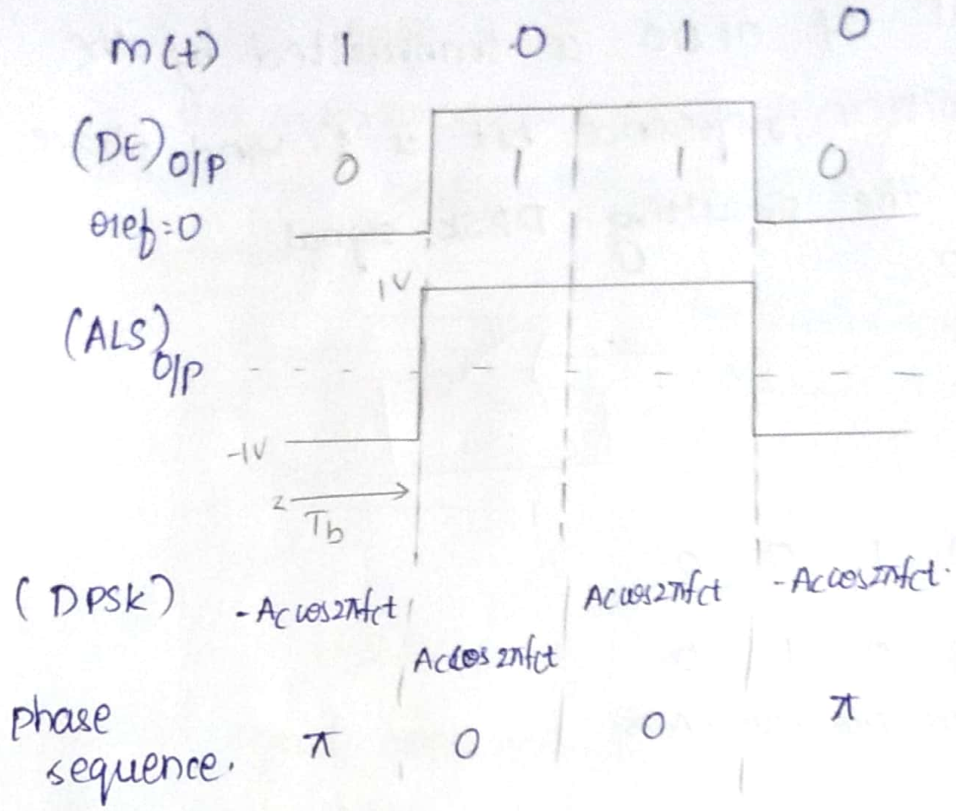
DPSK:-



Internal Cktary:-



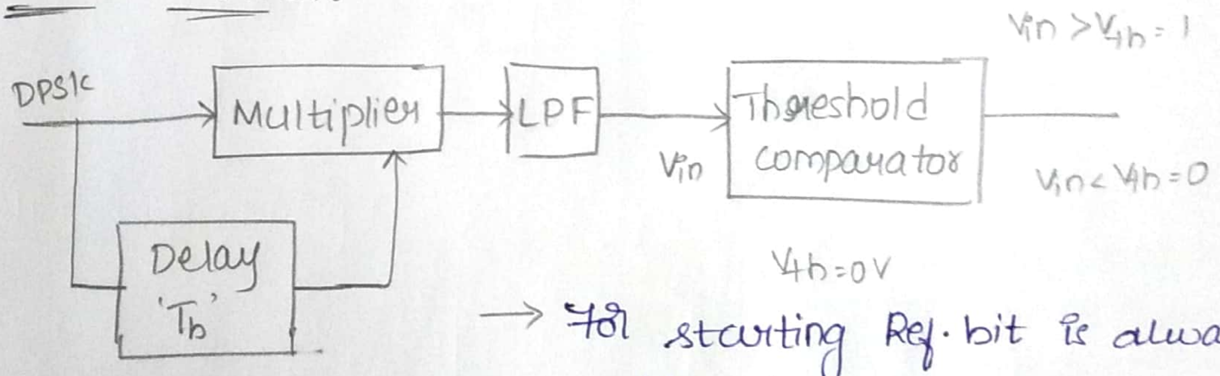




Note:-

m(t) will be given and ref bit will above be given and the phase sequence of resulting DPSK will be asked.

DPSK Receiver:-



→ For starting Ref. bit is always considered ref. Bit = 0 = $-A_c \cos 2\pi f_c t$

Analysis:-

DPSK →	$-A_c \cos 2\pi f_c t$	$A_c \cos 2\pi f_c t$	$A_c \cos 2\pi f_c t$	$-A_c \cos 2\pi f_c t$
$(\text{mul})_{o/p}$ →	$A_c^2 \cos^2 2\pi f_c t$	$-A_c^2 \cos^2 2\pi f_c t$	$A_c^2 \cos^2 2\pi f_c t$	$-A_c^2 \cos^2 2\pi f_c t$
$(\text{LPF})_{o/p}$ →	$A_c^2/2$	$-A_c^2/2$	$A_c^2/2$	$-A_c^2/2$
final o/p →	1	0	1	0

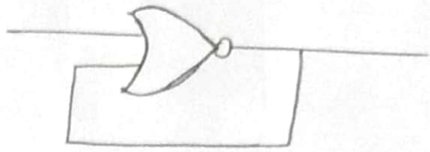
Q. A Binary signal of 0100 is transmitted by the DPSK transmitter; reference bit is '1'. find phase sequence of the resulting DPSK signal.

(a) $0\ \pi\ 0\ 0$

(b) $\pi\ 0\ \pi\ \pi$

(c) $0\ 0\ \pi\ 0$

(d) $\pi\ \pi\ 0\ \pi$



sol:-

$m(t) = 0 \quad 1 \quad 0 \quad 0$

$(D.E)_{o/p} = 0 \quad 0 \quad 1 \quad 0$

ref: 1 $\quad -Ac\cos.. \quad -Ac\cos.. \quad Ac\cos.. \quad -Ac\cos..$

ϕ sequence	π	π	0	π

4 Data Transmission

The analog signal is converted to digital or binary waveform by means of waveform coding techniques. In first and second chapter we have seen such waveform coding techniques. They are PCM, DM, ADM, DPCM etc. This digital data is then converted to RZ, NRZ, AMI etc. type of signal waveforms. The digital (binary) signal then can be transmitted either using baseband transmission or using bandpass transmission.

In *bandpass transmission,* the digital signal modulates high frequency sinusoidal carrier. The analysis of such techniques we have seen in previous chapter. They are called digital modulation techniques. With the help of such techniques, it is possible to transmit data over long distances. In *baseband transmission,* the data is transmitted without modulation.

During the transmission of data over the channel, it is corrupted by noise. Hence at the receiver, the noisy signal is received. Therefore correct detection of the transmitted signal is difficult. For example consider the transmitted signal and received noisy signal as shown in Fig. 4.1 (a) and (b).

The received signal $\hat{x}(t)$ is a noisy signal at the receiver. Let us consider that, the detector checks $\hat{x}(t)$ at 'T' during every bit interval. In above figure observe that the decision in first interval will be correct i.e. symbol '1'. But in second interval, the decision will be '1' but it is wrong. At the time when detector checks $\hat{x}(t)$ [i.e. at $t = T$], noise pulse is detected and decision is taken in favour of '1'. But actually symbol '0' is transmitted in second interval as shown in Fig. 4.1(a). Thus errors are introduced because of noise. The detecting method of the baseband signal perform following jobs:

(i) The detection method should attenuate noise and amplify signal, i.e. it should improve signal to noise ratio of the received signal.

(ii) The detection method should check the received signal at the time instant in the bit interval when signal to noise ratio is maximum.

(iii) The detection should be performed with minimum error probability.

In this chapter we will study some methods for detection of digital signals. We will also compare these methods on the basis of their performances.

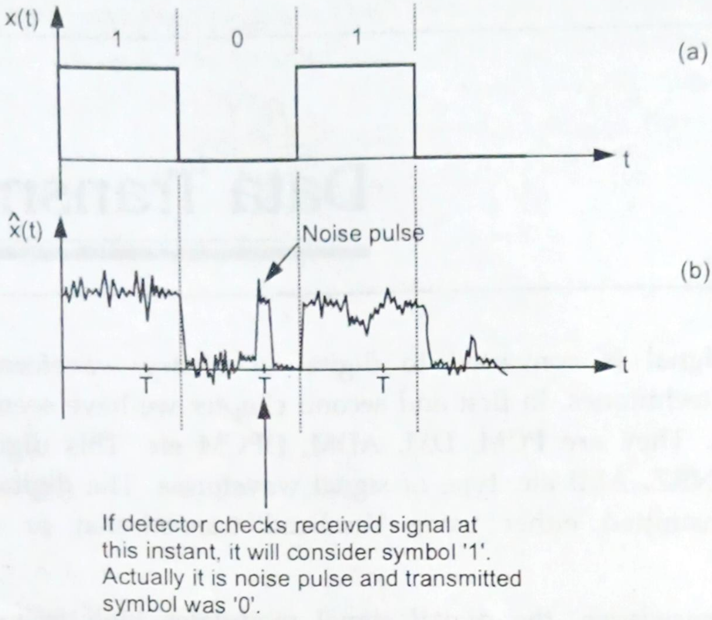


Fig. 4.1 Effect of noise at the receiver can make wrong decision
(a) The signal transmitted at the transmitter
(b) The received noisy signal at the receiver

4.1 Baseband Signal Receiver

Consider a very simple and basic detector for the detection of digital signals. Fig. 4.1.1 shows the circuit diagram of such integrate and dump filter.

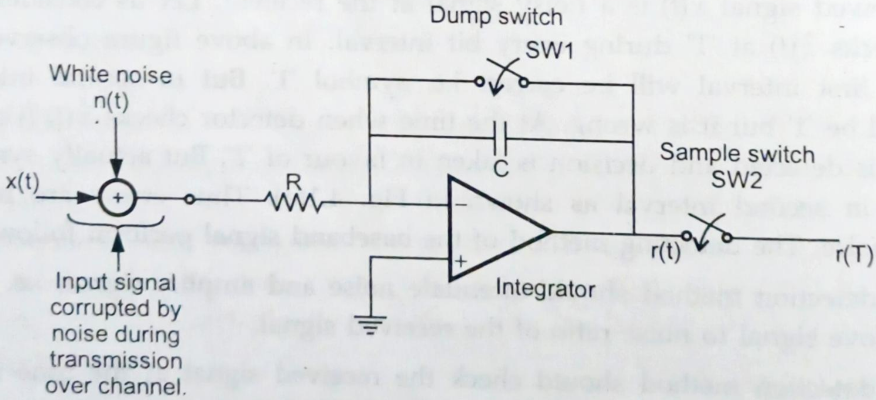


Fig. 4.1.1 Integrate and dump filter

The digital signal $x(t)$ is corrupted by white noise $n(t)$ during transmission over channel. Such noisy signal $[x(t)+n(t)]$ is given to the input of Integrate and Dump Filter. The capacitor is discharged fully at the beginning of the bit interval. This is

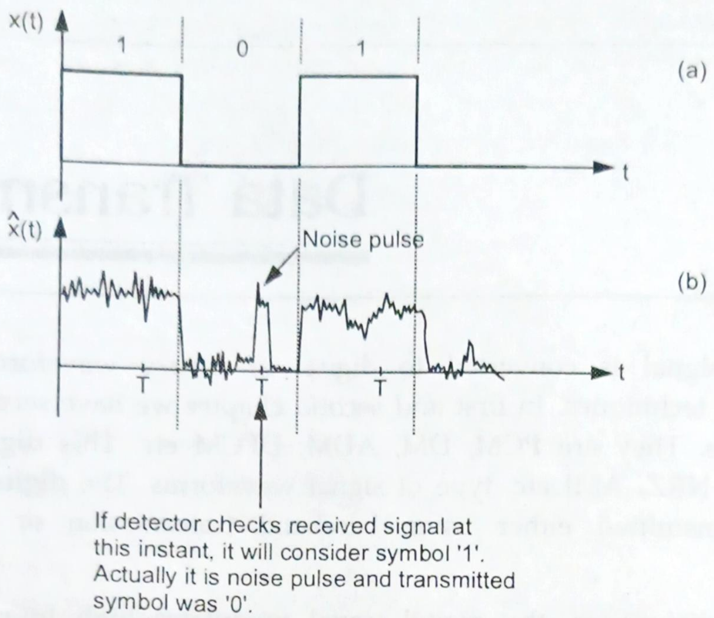


Fig. 4.1 Effect of noise at the receiver can make wrong decision
(a) The signal transmitted at the transmitter
(b) The received noisy signal at the receiver

4.1 Baseband Signal Receiver

Consider a very simple and basic detector for the detection of digital signals. Fig. 4.1.1 shows the circuit diagram of such integrate and dump filter.

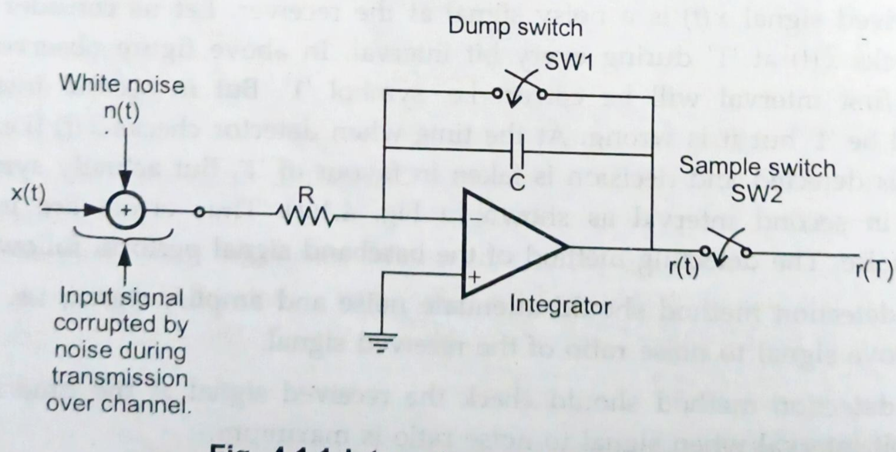


Fig. 4.1.1 Integrate and dump filter

The digital signal $x(t)$ is corrupted by white noise $n(t)$ during transmission over channel. Such noisy signal $[x(t) + n(t)]$ is given to the input of Integrate and Dump Filter. The capacitor is discharged fully at the beginning of the bit interval. This is achieved by temporarily closing switch SW1 at the beginning of the bit interval. The integrator then integrates noisy input signal over one bit period. This integrated signal is shown as $r(t)$ in above figure. For the square pulse input, the output of the integrator will be a triangular pulse as shown in Fig. 4.1.2.

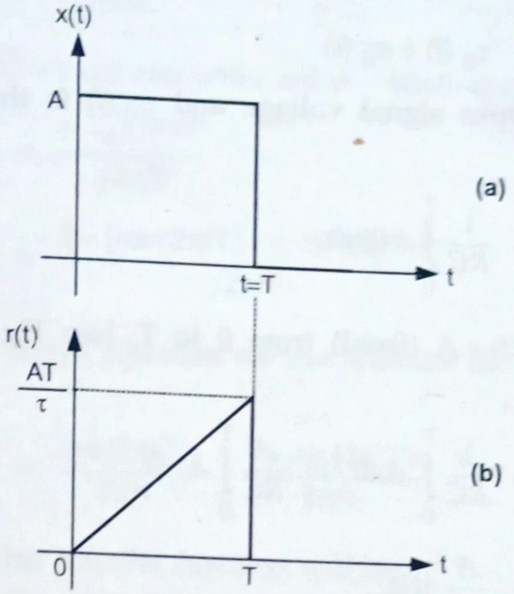


Fig. 4.1.2 (a) Input pulse to the integrator (assuming that the noise is absent). This pulse represents binary '1'. The width of pulse is 'T'

(b) Output of integrator. The initial output is zero. At $t=T$, the output of integrator is $r(t)=AT$

Fig. 4.1.2 (b) shows the waveform of $r(t)$. At the end of bit period i.e. $t=T$, the value of $r(t)$ reaches to its maximum amplitude. Therefore the value of $r(t)$ is sampled at the end of bit period. We will further prove that the signal to noise ratio is maximum at the end of bit period (i.e. $t=T$). Depending upon the value of $r(T)$, the decision is taken. The dump switch SW1 is then closed momentarily to discharge the capacitor to receive next bit. Thus integrator integrates (or generates output) independent of the value of previous bit. This shows that the detection in integrate and dump filter is unaffected by values of previous bits. In Fig. 4.1.2 (b), the output of integrator will decrease after $t>T$.

4.1.1 Signal to Noise Ratio of the Integrator and Dump Filter

The output of the integrator can be written as,

$$r(t) = \frac{1}{RC}\int_{0}^{T} [x(t)+n(t)]\,dt$$

Here the integration is performed over one bit period i.e. from 0 to T. The noisy signal $x(t)+n(t)$ is input to an integrator. We can write the above equation as,

$$r(t) = \frac{1}{RC}\int_0^T x(t)\,dt + \frac{1}{RC}\int_0^T n(t)\,dt \qquad \text{... (4.1.1)}$$

$$= x_0(t) + n_0(t)$$

Here $x_0(t)$ is the output signal voltage and $n_0(t)$ is the output noise. Consider output signal voltage,

$$x_0(t) = \frac{1}{RC}\int_0^T x(t)\,dt$$

Since the value of $x(t) = A$ (fixed) from 0 to T, [see Fig. 4.1.2 (a)], we can write above equation as,

$$x_0(t) = \frac{1}{RC}\int_0^T A\,dt = \frac{A}{RC}\int_0^T 1\cdot dt$$

$$= \frac{A}{RC}[t]_0^T$$

$$= \frac{AT}{RC}$$

Let the time constant $RC = \tau$. Then above equation becomes,

$$x_0(t) = \frac{AT}{\tau} \qquad \text{... (4.1.2)}$$

The normalized signal power in standard 1 Ω resistance will be,

$$\text{Output signal power} = \frac{x_0^2(t)}{1\,\Omega}$$

$$= \frac{A^2T^2}{\tau^2} \qquad \text{... (4.1.3)}$$

Now let us calculate the noise power. But before that we have to evaluate the transfer function of the integrator. A network which performs integration operation has the transfer function of $\frac{1}{j\omega RC}$. [*Note* : This can be very easily obtained by taking the case of RC circuit. An RC low pass filter also basically performs integration operation].

A delay of $t = T$ in time domain is equivalent to $e^{-j\omega T}$ in frequency domain. Thus the network performing integration over the period of T can be represented by the following transfer function.

$$H(f) = \frac{1-e^{-j\omega T}}{j\omega RC}$$

Since $\omega = 2\pi f$ and $RC = \tau$ we can write above equation as,

$$H(f) = \frac{1-e^{-j2\pi fT}}{j2\pi f\tau}$$

$$= \frac{1-[\cos(2\pi fT) - j\sin(2\pi fT)]}{j2\pi f\tau}$$

By rearranging the above equation we can separate the real and imaginary parts as follows:

$$H(f) = \frac{\sin(2\pi fT)}{2\pi f\tau} - j\frac{1-\cos(2\pi fT)}{2\pi f\tau}$$

The magnitude of this transfer function will be,

$$|H(f)|^2 = \frac{\sin^2(2\pi fT) + 1 - 2\cos(2\pi fT) + \cos^2(2\pi fT)}{(2\pi f\tau)^2}$$

On simplifying the above equation we get,

$$|H(f)|^2 = \frac{\sin^2(\pi fT)}{(\pi f\tau)^2} \quad \text{... (4.1.4)}$$

The average power of the output noise signal $n_0(t)$ is obtained by integrating its power density spectrum. i.e.,

$$\text{Power, } P = \int_{-\infty}^{\infty} S(f)\,df \quad \text{By definition.}$$

In standard 1Ω resistance, the noise power will be $\frac{\overline{n_0^2(t)}}{1\Omega} = \overline{n_0^2(t)}$. Here mean square value of noise is taken since it is random signal. i.e.,

$$\text{Noise power, } \overline{n_0^2(t)} = \int_{-\infty}^{\infty} S_{no}(f)\,df \quad \text{... (4.1.5)}$$

The input and output power spectral densities are related as,

$$S_{no}(f) = |H(f)|^2 S_{ni}(f) \quad \text{By standard formula} \quad \ldots (4.1.6)$$

Here $H(f)$ is transfer function of filter,

$S_{no}(f)$ is psd of output noise and

$S_{ni}(f)$ is psd of input noise.

We are assuming that white noise is present. The power spectral density (psd) of this noise is,

$$S_{ni}(f) = \frac{N_0}{2} \quad \ldots (4.1.7)$$

Putting this value in equation 4.1.6 we get,

$$\therefore \quad S_{no}(f) = |H(f)|^2 \cdot \frac{N_0}{2}$$

Putting this value in equation 4.1.5 we get,

$$\overline{n_0^2(t)} = \int_{-\infty}^{\infty} |H(f)|^2 \cdot \frac{N_0}{2} \, df$$

Putting the value of $|H(f)|^2$ from equation 4.1.4 in above equation,

$$\overline{n_0^2(t)} = \int_{-\infty}^{\infty} \frac{\sin^2(\pi f T)}{(\pi f \tau)^2} \cdot \frac{N_0}{2} \, df$$

$$= \frac{N_0}{2} \int_{-\infty}^{\infty} \frac{\sin^2(\pi f T)}{(\pi f \tau)^2} \, df \quad \ldots (4.1.8)$$

Put $\quad \pi f \tau = x$

$\therefore \quad dx = \pi \tau \, df \quad$ or $\quad df = \frac{1}{\pi\tau} dx$

Here $\quad f = \frac{x}{\pi\tau}$

$\therefore \quad \pi f T = \frac{xT}{\tau}$

With these substitutions equation (4.1.8) becomes,

$$\overline{n_0^2(t)} = \frac{N_0}{2} \int_{-\infty}^{\infty} \frac{\sin^2\left(\frac{xT}{\tau}\right)}{x^2} \cdot \frac{1}{\pi\tau} \, dx$$

Let us rearrange the above equation as,

Let us rearrange the above equation as,

$$\overline{n_0^2(t)} = \frac{N_0}{2} \int_{-\infty}^{\infty} \frac{\sin^2\left(\frac{xT}{\tau}\right)}{\left(\frac{xT}{\tau}\right)^2} \cdot \left(\frac{T}{\tau}\right)^2 \cdot \frac{1}{\pi\tau}\, dx$$

$$= \frac{N_0}{2} \cdot \frac{T^2}{\pi\tau^3} \int_{-\infty}^{\infty} \frac{\sin^2\left(\frac{xT}{\tau}\right)}{\left(\frac{xT}{\tau}\right)^2} \cdot dx$$

Let $\frac{xT}{\tau} = u$

$\therefore$ $dx = \frac{\tau}{T} du$ and limits will be unchanged, therefore above equation becomes,

$$\overline{n_0^2(t)} = \frac{N_0}{2} \cdot \frac{T^2}{\pi\tau^3} \int_{-\infty}^{\infty} \frac{\sin^2 u}{u^2} \cdot \frac{\tau}{T}\, du$$

$$= \frac{N_0 T}{2\pi \tau^2} \int_{-\infty}^{\infty} \left(\frac{\sin u}{u}\right)^2 du$$

Since the function $\frac{\sin u}{u}$ is squared, we can write above equation as,

$$\overline{n_0^2(t)} = \frac{N_0 T}{2\pi\tau^2} \cdot 2 \int_{0}^{\infty} \left(\frac{\sin u}{u}\right)^2 du$$

$$= \frac{N_0 T}{2\pi\tau^2} \cdot 2 \cdot \frac{\pi}{2} \quad \text{By equation C-46 in appendix 'C'.}$$

$\therefore$
$$\overline{n_0^2(t)} = \frac{N_0 T}{2\tau^2} \qquad \text{... (4.1.9)}$$

The above relation gives noise power at the output. We obtain the signal to noise power ratio at output of integrator as,

Signal to noise ratio, $\rho = \frac{\text{Signal power}}{\text{Noise power}}$

Putting the values of signal power from equation 4.1.3 and noise power from equation 4.1.9 we get,

$$\rho = \frac{\frac{A^2 T^2}{\tau^2}}{\frac{N_0 T}{2\tau^2}}$$

$$\therefore \qquad \rho = \frac{2A^2T}{N_0} \quad \text{or} \quad \frac{A^2T}{N_0/2}$$

Thus,

$$\boxed{\textbf{Signal to Noise ratio of Integrate and Dump receiver : } \rho = \frac{A^2T}{N_0/2}} \qquad \dots (4.1.10)$$

This signal to noise ratio is also called Figure of Merit.

Comments about signal to noise ratio :

1. The above result shows that signal to noise ratio improves in proportion to sampling period 'T'. It also increases as signal amplitude 'A' is more.
2. Since noise has gaussian distribution and zero mean value at any time, the output of integrator also increases by very small amount at the end of bit interval.

The Fig. 4.1.3 shows the waveforms of $x_0(t)$ and $n_0(t)$ for the input signal $x(t)$.

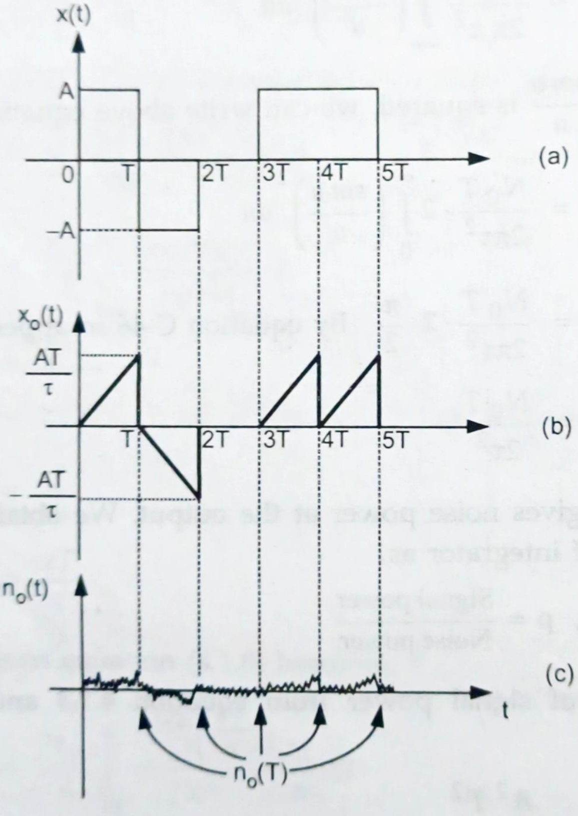


Fig. 4.1.3 (a) Input signal to the integrate and dump filter (receiver)
(b) Output signal of the integrate and dump filter
(c) Output noise of the integrate and dump receiver. Note that the output noise voltage increases to very small value at the sampling instants. i.e. T, 2T,3T,....

The Fig. 4.1.3 shows the waveforms of integrate and dump filter receiver for the rectangular pulses input. Observe that the output signal voltage reaches to the value of $\pm \frac{AT}{\tau}$ at the sampling instant. This is the maximum signal voltage. But the noise voltage $n_0(t)$ does not increase in same proportion. This is because the noise has zero average value and gaussian distribution.

Example 4.1.1 : *Prove that the maximum signal to noise ratio of the integrate and dump filter receiver is given as,*

$$\rho_{max} = \frac{2E}{N_0}$$

When the input signal $x(t)$ is rectangular pulses of amplitudes $\pm A$ and duration 'T'.

Solution : We know that energy of the signal $x(t)$ is given by standard relation as,

$$E = \int_{-\infty}^{\infty} x^2(t)\, dt \qquad \text{... (4.1.11)}$$

Since $x(t)$ is a rectangular pulse of amplitude $\pm A$ and duration T, we can write above equation as,

$$E = \int_0^T (\pm A)^2\, dt$$

$$= A^2 \int_0^T dt = A^2 \, [t]_0^T$$

$$= A^2 T \qquad \text{... (4.1.12)}$$

The signal to noise ratio of integrate and dump filter receiver is given by equation 4.1.10 as,

$$\rho = \frac{A^2 T}{N_0 / 2}$$

Putting $A^2 T = E$ from equation 4.1.12 in above equation,

$$\rho = \frac{E}{N_0 / 2} = \frac{2E}{N_0}$$

Since output is maximum at sampling instant T, the above value is maximum i.e.,

$$\rho_{max} = \frac{2E}{N_0}$$

7

Linear Block Codes

7.1 Introduction

- Errors are introduced in the data when it passes through the channel. The channel noise interferes the signal. The signal power is also reduced.
- Hence errors are introduced. In this chapter we will study various types of error detection and correction techniques.

7.1.1 Rationale for Coding and Types and Codes

- The transmission of the data over the channel depends upon two parameters. They are transmitted power and channel bandwidth. The power spectral density of channel noise and these two parameters determine signal to noise power ratio.
- The signal to noise power ratio determine the probability of error of the modulation scheme. For the given signal to noise ratio, the error probability can be reduced further by using coding techniques. The coding techniques also reduce signal to noise power ratio for fixed probability of error.

Fig. 7.1.1 shows the block diagram of the digital communication system which uses channel coding.

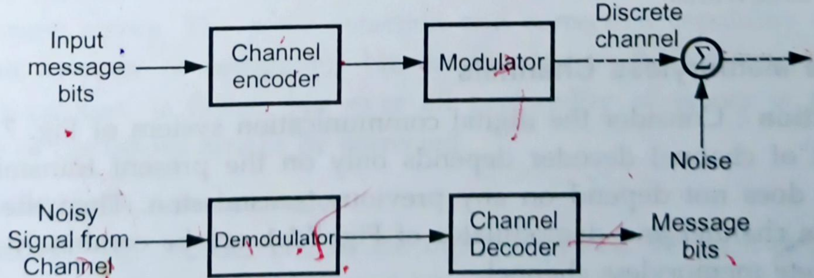


Fig. 7.1.1 Digital communication system with channel encoding

1. Channel encoder

The channel encoder adds extra bits (redundancy) to the message bits. The encoded signal is then transmitted over the noisy channel.

2. Channel decoder

- The channel decoder identifies the redundant bits and uses them to detect and correct the errors in the message bits if any.
- Thus the number of errors introduced due to channel noise are minimized by encoder and decoder. Due to the redundant bits, the overall data rate increases. Hence channel has to accommodate this increased data rate. The systems become slightly complex because of coding techniques.

7.1.2 Types of Codes

The codes are mainly classified as block codes and convolutional codes.

i) Block codes : These codes consists of 'n' number of bits in one block or codeword. This codeword consists of 'k' message bits and (n – k) redundant bits. Such block codes are called (n, k) block codes.

ii) Convolutional codes : The coding operation is discrete time convolution of input sequence with the impulse response of the encoder. The convolutional encoder accepts the message bits continuously and generates the encoded sequence continuously.

The codes can also be classified as linear or nonlinear codes.

i) Linear code : If the two codewords of the linear code are added by modulo-2 arithmetic, then it produces third codeword in the code.

This is very important property of the codes, since other codewords can be obtained by addition of existing codewords.

ii) Nonlinear code : Addition of the nonlinear codewords does not necessarily produce third codeword.

7.1.3 Discrete Memoryless Channels

- **Definition** : Consider the digital communication system of Fig. 7.1.1. Let the output of channel decoder depends only on the present transmitted signal, and it does not depend on any previous transmission. Then the modulator, discrete channel and demodulator of Fig. 7.1.1 can be combinely modeled as a discrete memoryless channel.
- We know that such channel is completely described by transition probabilities $P(y_j / x_i)$. Here x_i is the input symbol to modulator from channel encoder. And y_j is the output of demodulator and input to the channel decoder. $P(y_j / x_i)$ represents the probability of receiving symbol y_j, given that symbol x_i was transmitted.

7.1.4 Examples of Error Control Coding

- **Principle of redundancy :**

Let us consider the error control coding scheme which transmits 000 to transmit symbol '0' and 111 to transmit symbol '1'. Here note that there are two redundant bits in every message (symbol) being transmitted. The decoder checks the received triplets and takes the decision in favour of majority of the bits. For example, if the triplet is 110, then there are two 1's. Hence decision is taken in favour of 1. Here note that there is certainly error introduced in the last bit. Similarly if the received triplet is 001 or 100 or 010, then the decision is taken in favour of symbol '0'. The message symbol is received correctly if no more than one bit in each triplet is in error. If the message would have been transmitted without coding, then it is difficult to recover the original transmitted symbols. Thus the redundancy in the transmitted message reduces probability of error at the receiver.

- **Important aspects of error control coding :**

i. The redundancy bits in the message are called check bits. Errors can be detected and corrected with the help of these bits.

ii. It is not possible to detect and correct all the error in the message. Errors upto certain limit can only be detected and corrected.

iii. The check bits reduce the data rate through the channel.

7.1.5 Methods of Controlling Errors

There are two main methods used for error control coding : Forward acting error correction and Error detection with transmission.

i) Forward acting error correction

In this method, the errors are detected and corrected by proper coding techniques at the receiver (decoder). The check bits or redundant bits are used by the receiver to detect and correct errors. The error detection and correction capability of the receiver depends upon number of redundant bits in the transmitted message. The forward acting error correction is faster, but over all probability of errors is higher. This is because some of the errors cannot be corrected.

ii) Error detection with retransmission

In this method, the decoder checks the input sequence. When it detects any error, it discards that part of the sequence and requests the transmitter for retransmission. The transmitter then again transmits the part of the sequence in which error was detected. Here note that, the decoder does not correct the errors. It just detects the errors and sends requests to transmitter. This method has lower probability of error, but it is slow.

7.1.6 Types of Errors

There are mainly two types of errors introduced during transmission on the data : random errors and burst errors.

i) Random errors : These errors are created due to white gaussian noise in the channel. The errors generated due to white gaussian noise in the particular interval does not affect the performance of the system in subsequent intervals. In other words, these errors are totally uncorrelated. Hence they are also called as random errors.

ii) Burst errors : These errors are generated due to impulsive noise in the channel. These impulse noise (bursts) are generated due to lightning and switching transients. These noise bursts affect several successive symbols. Such errors are called burst errors. The burst errors are dependent on each other in successive message intervals.

7.1.7 Some of the Important Terms used in Error Control Coding

The terms which are regularly used in error control coding are defined next.

Codeword : The encoded block of 'n' bits is called a codeword. It contains message bits and redundant bits.

Block length : The number of bits 'n' after coding is called the block length of the code.

Code rate : The ratio of message bits (k) and the encoder output bits (n) is called code rate. Code rate is defined by 'r' i.e.,

$$\boxed{r = \frac{k}{n}} \quad \text{... (7.1.1)}$$

we find that $0 < r < 1$.

Channel data rate : It is the bit rate at the output of encoder. If the bit rate at the input of encoder is R_s, then channel data rate will be,

$$\text{Channel data rate } (R_o) = \frac{n}{k} R_s \quad \text{... (7.1.2)}$$

Code vectors : An 'n' bit codeword can be visualized in an n-dimensional space as a vector whose elements or co-ordinates are the bits in the codeword. It is simpler to visualize the 3-bit codewords. Fig. 7.1.2 shows the 3-bit codevectors. There will be distinct '8' codewords (since number of codewords = 2^k). If we let bits b_0 on x-axis, b_1 on y-axis and b_2 on z-axis, then the following table gives various points as code vectors in the 3-dimensional space.

Sr.No.	Bits of code vector		
	$b_2 = z$	$b_1 = y$	$b_0 = x$
1	0	0	0
2	0	0	1
3	0	1	0
4	0	1	1
5	1	0	0
6	1	0	1
7	1	1	0
8	1	1	1

Table 7.1.1 Code vectors in 3-dimensional space

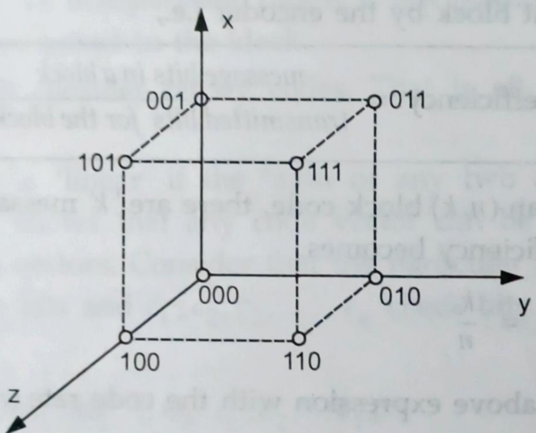


Fig. 7.1.2 Code vectors representing 3-bit codewords

Hamming distance : The hamming distance between the two code vectors is equal to the number of elements in which they differ. For example let $X = (101)$ and $Y = (110)$. The two code vectors differ in second and third bits. Therefore hamming distance between X and Y is 'two'. Hamming distance is denoted as $d(X, Y)$ or simply 'd'. i.e.

$$d(X, Y) = d = 2$$

Thus we observe from Fig. 7.1.2 that the hamming distance between (100) and (011) is maximum i.e. 3. This is indicated by the vector diagram also.

Minimum distance (d_{min}) : It is the smallest hamming distance between the valid code vectors.

Error detection is possible if the received vector is not equal to some other code vector. This shows that the transmission errors in the received code vector should be less than minimum distance d_{min}. The following table lists some of the requirements of error control capability of the code.

Sr.No.	Name of errors detected / corrected	Distance requirement
1	Detect upto 's' errors per word	$d_{min} \geq s+1$
2	Correct upto 't' errors per word	$d_{min} \geq 2t+1$
3	Correct upto 't' errors and detect $s > t$ errors per word	$d_{min} \geq t+s+1$

Table 7.1.2 Error control capabilities

For the (n, k) block code, the minimum distance is given as,

$$d_{min} \leq n-k+1 \quad \text{... (7.1.3)}$$

Code efficiency : The code efficiency is the ratio of message bits in a block to the transmitted bits for that block by the encoder i.e.,

$$\boxed{\text{Code efficiency} = \frac{\textit{message bits in a block}}{\textit{transmitted bits for the block}}}$$

We know that for an (n, k) block code, there are 'k' message bits and 'n' transmitted bits. Therefore code efficiency becomes,

$$\text{Code efficiency} = \frac{k}{n} \quad \text{... (7.1.4)}$$

If we compare the above expression with the code rate (r) of equation 7.1.1 we find that,

$$\text{Code efficiency} = \text{code rate} = \frac{k}{n} \quad \text{... (7.1.5)}$$

Weight of the code : The number of non-zero elements in the transmitted code vector is called vector weight. It is denoted by $w(X)$ where X is the code vector. For example if $X = 0111\,0101$, then weight of this code vector will be $w(X) = 5$.

7.2 Linear Block Codes

Principle of block coding :

For the block of k message bits, $(n-k)$ parity bits or check bits are added. Hence the total bits at the output of channel encoder are 'n'. Such codes are called (n, k) block codes. Fig.7.2.1 illustrates this concept.

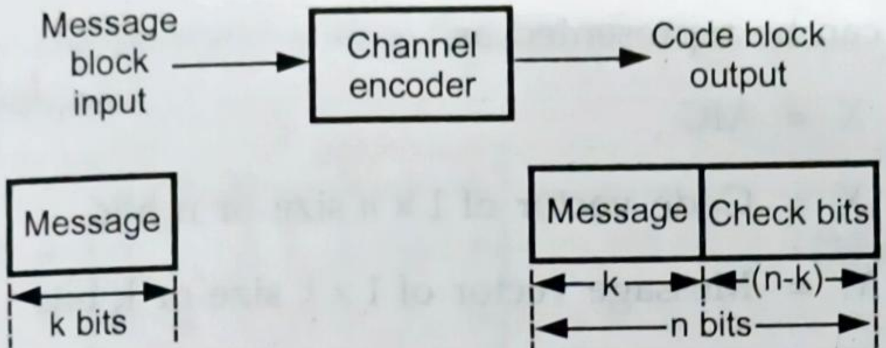


Fig. 7.2.1 Functional block diagram of block coder

Systematic codes : In the systematic block code, the message bits appear at the beginning of the code word. Thus as shown in Fig. 7.2.1 , the message bits appear first and then check bits are transmitted in a block. This type of code is called systematic code.

Nonsystematic codes : In *nonsystematic code* it is not possible to identify message bits and check bits. They are mixed in the block.

In this section we will consider binary codes. That is all transmitted digits are binary.

5.6 Discrete Memoryless Channels

The discrete memoryless channel has input X and output Y. Both X and Y are the random variables. The channel is discrete when both X and Y are discrete. The channel is called memoryless (zero memory) when current output depends only on current input.

The channel is described in terms of input alphabet, output alphabet and the set of transition probabilities. The transition probability $P(y_j / x_i)$ is the conditional probability of y_j is received, given that x_i was transmitted. If $i = j$ then $P(y_j / x_i)$ represents conditional probability of correct reception. And if $i \neq j$, then $P(y_j / x_i)$ represents a conditional probability of error. The transition probabilities of the channel can be represented by a matrix as follows :

$$P = \begin{bmatrix} P(y_1 / x_1) & P(y_2 / x_1) & --- & P(y_m / x_1) \\ P(y_1 / x_2) & P(y_2 / x_2) & --- & P(y_m / x_2) \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ P(y_1 / x_n) & P(y_2 / x_n) & --- & P(y_m / x_n) \end{bmatrix} \quad \text{.... (5.6.1)}$$

The above matrix has the size of $n \times m$. It is called the *channel matrix* or *probability transition matrix.* Each row of above matrix represents fixed input. And each column of above matrix represents fixed output. The summation of all transition probabilities along the row is equal to 1 i.e.,

$$P(y_1 / x_1) + P(y_2 / x_1) + ...+ P(y_m / x_1) = 1.$$

This is applicable to other rows also. Hence we can write,

$$\sum_{j=1}^{m} P(y_j / x_i) = 1 \quad \text{.... (5.6.2)}$$

For the fixed input x_j, the output can be any of $y_1, y_2, y_3, \ldots\ldots y_m$. The summation of all these possibilities is equal to 1. From the probability theory we know that,

$$P(AB) = P(B / A)\, P(A) \qquad \ldots (5.6.3)$$

Here $P(AB)$ is the joint probability of A and B. Here if we let $A = x_i$ and $B = y_i$, then,

$$P(x_i\ y_j) = P(y_j / x_i)\, P(x_i) \qquad \ldots (5.6.4)$$

Here $P(x_i\ y_j)$ is the joint probability of x_i and y_j. If we add all the joint probabilities for fixed y_j then we get $P(y_j)$ i.e.,

$$\sum_{i=1}^{n} P(x_i, y_j) = P(y_j) \qquad \ldots (5.6.5)$$

This is written as per the standard probability relations. The above equation gives probability of getting symbol y_j. From equation 5.6.4 and above equation we can write,

$$P(y_j) = \sum_{i=1}^{n} P(y_j / x_i)\, P(x_i) \qquad \ldots (5.6.6)$$

Here $j = 1, 2, \ldots\ldots m.$

Thus if we are given the probabilities of input symbols and transition probabilities, then it is possible to calculate the probabilities of output symbols. Error will result when i^{th} symbol is transmitted but j^{th} symbol is received.

Hence the error probability P_e can be obtained as,

$$P_e = \sum_{\substack{j=1 \\ j \neq i}}^{m} P(y_j) \qquad \ldots (5.6.7)$$

Thus all the probabilities will contribute to error except $i = j$. This is because in case of $i = j$, correct symbol is received. From equation 5.6.6 and above equation,

$$P_e = \sum_{\substack{j=1 \\ j \neq i}}^{m} \sum_{i=1}^{n} P(y_j / x_i)\, P(x_i) \qquad \ldots (5.6.8)$$

And the probability of correct reception will be,

$$P_c = 1 - P_e \qquad \ldots (5.6.9)$$

5.6.1 Binary Communication Channel

Consider the case of the discrete channel where there are only two symbols transmitted. Fig. 5.6.1 shows the diagram of a binary communication channel.

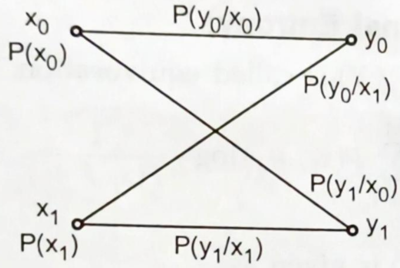


Fig. 5.6.1 Binary communication channel

We can write the equations for probabilities of y_0 and y_1 as,

$$P(y_0) = P(y_0 / x_0)\, P(x_0) + P(y_0 / x_1)\, P(x_1) \quad \text{... (5.6.10)}$$

and

$$P(y_1) = P(y_1 / x_1)\, P(x_1) + P(y_1 / x_0)\, P(x_0) \quad \text{... (5.6.11)}$$

Above equations can be written in the matrix form as,

$$\begin{bmatrix} P(y_0) \\ P(y_1) \end{bmatrix} = \begin{bmatrix} P(x_0) & P(x_1) \end{bmatrix} \begin{bmatrix} P(y_0 / x_0) & P(y_1 / x_0) \\ P(y_0 / x_1) & P(y_1 / x_1) \end{bmatrix} \quad \text{... (5.6.12)}$$

Note that the 2×2 matrix in above equation is a probability transition matrix. It is similar to equation 5.6.1.

Binary symmetric channel :

The binary communication channel of Fig. 5.6.1 is said to be symmetric if $P(y_0 / x_0) = P(y_1 / x_1) = p$. Such channel is shown in Fig. 5.6.2.

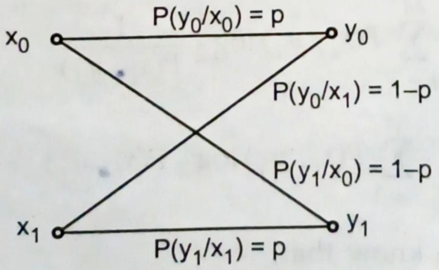


Fig. 5.6.2 Binary symmetric channel

For the above channel, we can write equation 5.6.12 as,

$$\begin{bmatrix} P(y_0) \\ P(y_1) \end{bmatrix} = \begin{bmatrix} P(x_0) & P(x_1) \end{bmatrix} \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix} \quad \text{... (5.6.13)}$$

5.6.2 Equivocation (Conditional Entropy)

The conditional entropy $H(X / Y)$ is called equivocation. It is defined as,

$$H(X / Y) = \sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(x_i / y_j)} \quad \text{... (5.6.14)}$$

And the joint entropy $H(X, Y)$ is given as,

$$H(X, Y) = \sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(x_i\, y_j)} \quad \text{... (5.6.15)}$$

The conditional entropy $H(X / Y)$ represents uncertainty of X, on average, when Y is known. Similarly the conditional entropy $H(Y / X)$ represents uncertainty of Y, on average, when X is transmitted. $H(Y / X)$ can be given as,

$$H(Y / X) = \sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(y_j / x_i)} \quad \text{... (5.6.16)}$$

The conditional entropy $H(X / Y)$ is an average measure of uncertainty in X after Y is received. In other words $H(X / Y)$ represents the information lost in the noisy channel.

Example 5.6.1 : *Prove that*

$$H(X, Y) = H(X/Y) + H(Y)$$
$$= H(Y/X) + H(X)$$

Solution : Consider equation 5.6.15,

$$H(X, Y) = \sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(x_i \cdot y_j)}$$

$$= -\sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i, y_j) \log_2 P(x_i \cdot y_j) \quad \text{... (5.6.17)}$$

From probability theory we know that,

$$P(AB) = P(A / B)\, P(B)$$

$$\therefore \quad P(x_i, y_j) = P(x_i / y_j) P(y_j)$$

Putting this result in the $\log_2$ term of equation 5.6.17 we get,

$$H(X,Y) = -\sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 [P(x_i / y_j) P(y_j)]$$

We know that $\log_2 \left[P\left(x_i / y_j\right) P\left(y_j\right)\right] = \log_2 P\left(x_i / y_j\right) + \log_2 P\left(y_j\right)$.

Hence above equation becomes,

$$H(X,Y) = -\sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 P(x_i / y_j)$$

$$-\sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 P(y_j)$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(x_i / y_j)}$$

$$-\sum_{j=1}^{M}\left\{\sum_{i=1}^{M} P(x_i, y_j)\right\} \log_2 P(y_j)$$

The first term in above equation is $H(X / Y)$ as per equation 5.6.14. From the standard probability theory,

$$\sum_{i=1}^{M} P(x_i, y_j) = P(y_j)$$

Hence $H(X, Y)$ will be written as,

$$H(X,Y) = H(X / Y) - \sum_{j=1}^{M} P(y_j) \log_2 P(y_j)$$

$$= H(X / Y) + \sum_{j=1}^{M} P(y_j) \log_2 \frac{1}{P(y_j)}$$

As per the definition of entropy, the second term in the above equation is $H(Y)$. Hence,

$$H(X,Y) = H(X / Y) + H(Y) \qquad \text{... (5.6.18)}$$

Thus the first given equation is proved. From the probability theory we know that

$$P(AB) = P(B / A) P(A)$$

$$\therefore \quad P(x_i, y_j) = P(y_j / x_i) P(x_i)$$

Putting this result in the $\log_2$ term of equation 5.6.17 we get,

$$H(X,Y) = -\sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 [P(y_j / x_i) P(x_i)]$$

$$= -\sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 P(y_j / x_i)$$

$$-\sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 P(x_i)$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(y_j / x_i)}$$

$$-\sum_{i=1}^{M}\left\{\sum_{j=1}^{M} P(x_i, y_j)\right\} \log_2 P(x_i)$$

As per equation 5.6.16, the first term of above equation is $H(Y / X)$. And from standard probability theory,

$$\sum_{j=1}^{M} P(x_i, y_j) = P(x_i)$$

Hence $H(X,Y)$ will be written as,

$$H(X,Y) = H(Y / X) - \sum_{i=1}^{M} P(x_i) \log_2 P(x_i)$$

$$= H(Y / X) + \sum_{i=1}^{M} P(x_i) \log_2 \frac{1}{P(x_i)}$$

As per the definition of entropy, the second term in the above equation is $H(X)$. Hence

$$H(X,Y) = H(Y / X) + H(X) \quad \text{... (5.6.19)}$$

Thus the second part of the given equation is proved.

Example 5.6.2 : *Two BSC's are connected in cascade as shown in Fig. 5.6.3.*

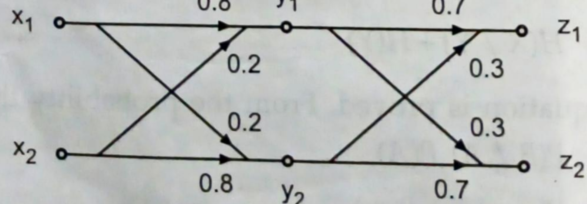


Fig. 5.6.3 BSC of example 5.6.2

i) Find the channel matrix of resultant channel.

ii) Find $P(z_1)$ and $P(z_2)$ if $P(x_1) = 0.6$ and $P(x_2) = 0.4$.

Solution : i) To obtain channel matrix using equation 5.6.1 :

Here we can write two matrices of the channel as follows :

$$P(Y/X) = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

and $$P(Z/Y) = \begin{bmatrix} P(z_1/y_1) & P(z_2/y_1) \\ P(z_1/y_2) & P(z_2/y_2) \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

Hence resultant channel matrix is given as,

$$P(Z/X) = P(Y/X) \cdot P(Z/Y)$$

$$= \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

$$= \begin{bmatrix} 0.62 & 0.38 \\ 0.38 & 0.62 \end{bmatrix}$$

ii) To obtain $P(z_1)$ and $P(z_2)$:

The probabilities of Z_1 and Z_2 are given as,

$$P(Z) = P(X)\, P(Z/X)$$

$$= [P(x_1)\ P(x_2)] \begin{bmatrix} 0.62 & 0.38 \\ 0.38 & 0.62 \end{bmatrix}$$

$$= [0.6\ \ 0.4] \begin{bmatrix} 0.62 & 0.38 \\ 0.38 & 0.62 \end{bmatrix}$$

$$= [0.524\ \ 0.476]$$

Thus $P(z_1) = 0.524$ and

$P(z_2) = 0.476.$

5.6.3 Rate of Information Transmission Over a Discrete Channel

The entropy of the symbol gives average amount of information going into the channel. i.e.,

$$H(X) = \sum_{i=1}^{M} p_i \log_2\left(\frac{1}{p_i}\right) \quad \text{... (5.6.20)}$$

Let the symbols be generated at the rate of 'r' symbols per second. Then the average rate of information going into the channel is given as,

$$D_{in} = rH(X) \text{ bits/sec} \quad \text{... (5.6.21)}$$

Errors are introduced in the data during the transmission. Because of these errors, some information is lost in the channel. The conditional entropy $H(X/Y)$ is the measure of information lost in the channel. Hence the information transmitted over the channel will be,

$$\text{Transmitted information} = H(X) - H(X/Y) \quad \dots (5.6.22)$$

Hence the average rate of information transmission D_t across the channel will be,

$$D_t = [H(X) - H(X/Y)]r \text{ bits/sec} \quad \dots (5.6.23)$$

When the noise becomes very large, then X and Y become statistically independent. Then $H(X/Y) = H(X)$ and hence no information is transmitted over the channel. In case of errorless transmission $H(X/Y) = 0$, hence $D_{in} = D_t$. That is the input information rate is same as information rate across the channel. No information is lost when $H(X/Y) = 0$.

Example 5.6.3 : *Fig. 5.6.4 shows the binary symmetric channel. Find the rate of information transmission across this channel for p = 0.8 and 0.6. The symbols are generated at the rate of 1000 per second.* $P(x_0) = P(x_1) = \frac{1}{2}$ *Also determine channel input information rate.*

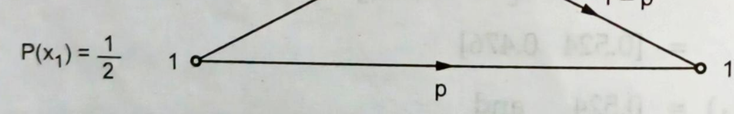


Fig. 5.6.4 Binary symmetric channel of Ex. 5.6.3

Solution : (i) To obtain entropy of the source :

The entropy of the source can be obtained as,

$$H(X) = P(x_0)\log_2 \frac{1}{P(x_0)} + P(x_1)\log_2 \frac{1}{P(x_1)}$$

$$= \frac{1}{2}\log_2 2 + \frac{1}{2}\log_2 2$$

$$= 1 \text{ bit/symbol.}$$

(ii) To obtain input information rate :

The input information rate is,

$$D_{in} = rH(X)$$

Here $r = 1000$ symbols sec. Hence above equation will be,

$$D_{in} = 1000 \times 1 = 1000 \text{ bits/sec.}$$

(iii) To obtain $P(y_0)$ and $P(y_1)$:

The probability transition matrix for binary symmetric channel of Fig. 5.6.4 can be written as,

$$P = \begin{bmatrix} P(y_0 / x_0) & P(y_1 / x_0) \\ P(y_0 / x_1) & P(y_1 / x_1) \end{bmatrix}$$

$$= \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

From equation 5.6.13 we can write the probabilities of output symbols as,

$$\begin{bmatrix} P(y_0) \\ P(y_1) \end{bmatrix} = \begin{bmatrix} P(x_0) & P(x_1) \end{bmatrix} \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} p + \frac{1}{2}(1-p) \\ \frac{1}{2}(1-p) + \frac{1}{2} p \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Thus $P(y_0) = \frac{1}{2}$ and $P(y_1) = \frac{1}{2}$.

(iv) To obtain $P(x_i, y_j)$ and $P(x_i / y_j)$:

Now let us obtain the conditional probabilities $P(X / Y)$. From the probability theory we know that,

$$P(AB) = P(A / B)\, P(B) = P(B / A)\, P(A) \quad \text{... (5.6.24)}$$

$$\therefore \quad P(x_0\, y_0) = P(y_0 / x_0)\, P(x_0)$$

$$= p \times \frac{1}{2} = \frac{1}{2} p$$

Similarly $P(x_1\, y_0) = P(y_0 / x_1)\, P(x_1)$

$$= (1-p) \times \frac{1}{2}$$

$$= \frac{1}{2}(1-p)$$

$$P(x_0\ y_1) = P(y_1 / x_0)\ P(x_0)$$

$$= (1-p) \times \frac{1}{2} = \frac{1}{2}(1-p)$$

and $$P(x_1\ y_1) = P(y_1 / x_1)\ P(x_1)$$

$$= p \times \frac{1}{2} = \frac{1}{2}p$$

Now from equation 5.6.24 we can obtain the condition probabilities $P(X / Y)$.

$$P(x_0 / y_0) P(y_0) = P(y_0 / x_0)\ P(x_0)$$

$$\therefore \quad P(x_0 / y_0) = \frac{P(y_0 / x_0)\ P(x_0)}{P(y_0)}$$

$$= \frac{p \times \frac{1}{2}}{\frac{1}{2}} = p$$

Similarly, $$P(x_1/y_0) = \frac{P(y_0 / x_1)\ P(x_1)}{P(y_0)}$$

$$= \frac{(1-p) \times \frac{1}{2}}{\frac{1}{2}} = 1-p$$

$$P(x_0 / y_1) = \frac{P(y_1 / x_0)\ P(x_0)}{P(y_1)}$$

$$= \frac{(1-p) \times \frac{1}{2}}{\frac{1}{2}} = 1-p$$

and $$P(x_1 / y_1) = \frac{P(y_1 / x_1)\ P(x_1)}{P(y_1)}$$

$$= \frac{p \times \frac{1}{2}}{\frac{1}{2}} = p$$

All the calculated results are given in Table 5.6.1 below :

Quantity	Value
$P(x_0) = P(x_1)$	$\frac{1}{2}$
$P(y_0) = P(y_1)$	$\frac{1}{2}$
$P(y_1 / x_1) = P(y_0 / x_0)$	p
$P(y_1 / x_0) = P(y_0 / x_1)$	$1-p$
$P(x_0 y_0)$	$\frac{1}{2}p$
$P(x_1 y_0)$	$\frac{1}{2}(1-p)$
$P(x_0 y_1)$	$\frac{1}{2}(1-p)$
$P(x_1 y_1)$	$\frac{1}{2}p$
$P(x_0 / y_0)$	p
$P(x_0 / y_1)$	$1-p$
$P(x_1 / y_0)$	$1-p$
$P(x_1 / y_1)$	p

Table 5.6.1 Conditional probabilities of Fig. 5.6.4

(v) To obtain information rate across the channel :

The conditional entropy $H(X / Y)$ is given by equation 5.6.14 as,

$$H(X / Y) = \sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i, y_j) \log_2 \frac{1}{P(x_i / y_j)}$$

Here $M = 2$ and expanding the above equation we get,

$$H(X / Y) = P(x_0, y_0) \log_2 \frac{1}{P(x_0 / y_0)} + P(x_0, y_1) \log_2 \frac{1}{P(x_0 / y_1)}$$

$$+ P(x_1, y_0) \log_2 \frac{1}{P(x_1 / y_0)} + P(x_1, y_1) \log_2 \frac{1}{P(x_1 / y_1)}$$

Putting the values in above equation from Table 5.6.1,

$$H(X / Y) = \frac{1}{2} p \log_2 \frac{1}{p} + \frac{1}{2}(1-p) \log_2 \frac{1}{1-p} + \frac{1}{2}(1-p) \log_2 \frac{1}{1-p} + \frac{1}{2} p \log_2 \frac{1}{p}$$

i.e.
$$\boxed{H(X/Y) = p\log_2\frac{1}{p} + (1-p)\log_2\frac{1}{1-p}} \quad \dots (5.6.25)$$

For p = 0.8 , $H(X/Y)$ becomes,

$$H(X/Y) = 0.8\log_2\frac{1}{0.8} + (1-0.8)\log_2\frac{1}{(1-0.8)}$$
$$= 0.721928 \text{ bits/symbol}$$

Hence from equation 5.6.23, the information transmission rate across the channel will be,

$$D_t = (1-0.721928)\times 1000$$
$$= 278 \text{ bits/sec.}$$

For $p = 0.6$, $H(X/Y)$ of equation 5.6.25 becomes,

$$H(X/Y) = 0.6\log_2\frac{1}{0.6} + (1-0.6)\log_2\frac{1}{(1-0.6)}$$
$$= 0.97 \text{ bits/symbol}$$

Hence from equation 5.6.23, the information transmission rate across the channel will be,

$$D_t = (1-0.97)\times 1000$$
$$= 29 \text{ bits/sec.}$$

The above results indicate that the information transmission rate decreases rapidly as p approaches $\frac{1}{2}$.

Example 5.6.4 : *A channel has the following channel matrix,*

$$[P(Y/X)] = \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$

(i) Draw the channel diagram.

(ii) If the source has equally likely outputs, compute the probabilities associated with the channel outputs for $p = 0.2$.

Solution : Given data

$$P(Y/X) = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) & P(y_3/x_1) \\ P(y_1/x_2) & P(y_2/x_2) & P(y_3/x_2) \end{bmatrix}$$

$$= \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$

(i) To obtain channel diagram :

From above matrices we can prepare the channel diagram as follows :

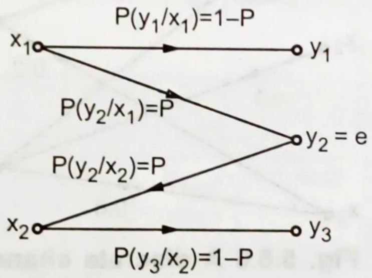


Fig. 5.6.5 Binary erasure channel

In the above channel diagram, observe that there are two input symbols and three output symbols. At the output symbol y_2 represents error (e). Such channel is called *binary erasure channel.*

(ii) To calculate $P(y_1)$, $P(y_2)$ and $P(y_3)$:

It is given that source emits x_1 and x_2 with equal probabilities. Hence,

$$P(x_1) = \frac{1}{2} \quad \text{and} \quad P(x_2) = \frac{1}{2}$$

The output probabilities are given as,

$$\begin{bmatrix} P(y_1) \\ P(y_2) \\ P(y_3) \end{bmatrix} = [P(x_1) \quad P(x_2)] \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$

It is given that p = 0.2. Putting values in above equation,

$$\begin{bmatrix} P(y_1) \\ P(y_2) \\ P(y_3) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} \times 0.8 + 0 \\ \frac{1}{2} \times 0.2 + \frac{1}{2} \times 0.2 \\ \frac{1}{2} \times 0 + \frac{1}{2} \times 0.8 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.2 \\ 0.4 \end{bmatrix}$$

Thus $P(y_1) = 0.4$, $P(y_2) = 0.2$ and $P(y_3) = 0.4$.

Example 5.6.5 : *A discrete memoryless source with three symbols, with probabilities $P(x_1) = P(x_3)$ and $P(x_2) = \alpha$, feeds into discrete memoryless channel shown in Fig. 5.6.6.*

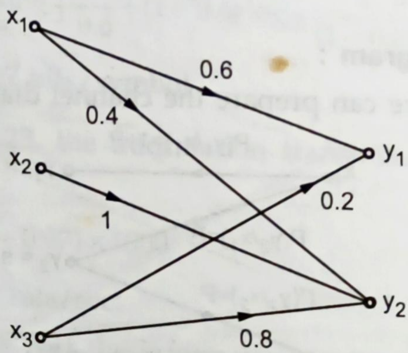


Fig. 5.6.6 A discrete channel

(i) Sketch the variation of $H(X)$ with α. Determine maximum value of $H(X)$ and show it on the sketch.

(ii) Determine the transition matrix for discrete memoryless channel.

(iii) Determine the maximum value of entropy $H(Y)$ at the channel output.

Solution : (i) To obtain variation of $H(X)$ with α :

It is given that,

$$P(x_2) = \alpha$$

We know that, $P(x_1) + P(x_2) + P(x_3) = 1$

Since $P(x_2) = P(x_3)$, $\quad P(x_1) + 2P(x_2) = 1$

$$\therefore \quad \alpha + 2P(x_2) = 1$$

$$\therefore \quad P(x_2) = \frac{1-\alpha}{2}$$

And $\quad P(x_2) = P(x_3) = \frac{1-\alpha}{2}$

Entropy of the source is given as,

$$H(X) = \sum_{k=1}^{3} P(x_k) \log_2 \frac{1}{P(x_k)}$$

$$= \alpha \log_2 \frac{1}{\alpha} + \left(\frac{1-\alpha}{2}\right) \log_2 \left(\frac{2}{1-\alpha}\right) + \left(\frac{1-\alpha}{2}\right) \log_2 \left(\frac{2}{1-\alpha}\right)$$

$$= \alpha \log_2 \frac{1}{\alpha} + (1-\alpha) \log_2 \frac{2}{1-\alpha}$$

Following table shows the calculations of $H(X)$ with respect to α.

α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$H(X)$	1	1.368	1.522	1.581	1.57	1.5	1.37	1.18	0.921	0.569	0

Table 5.6.2 $H(X)$ with respect to α

Fig. 5.6.7 shows the sketch of $H(X)$ versus α as per above calculations.

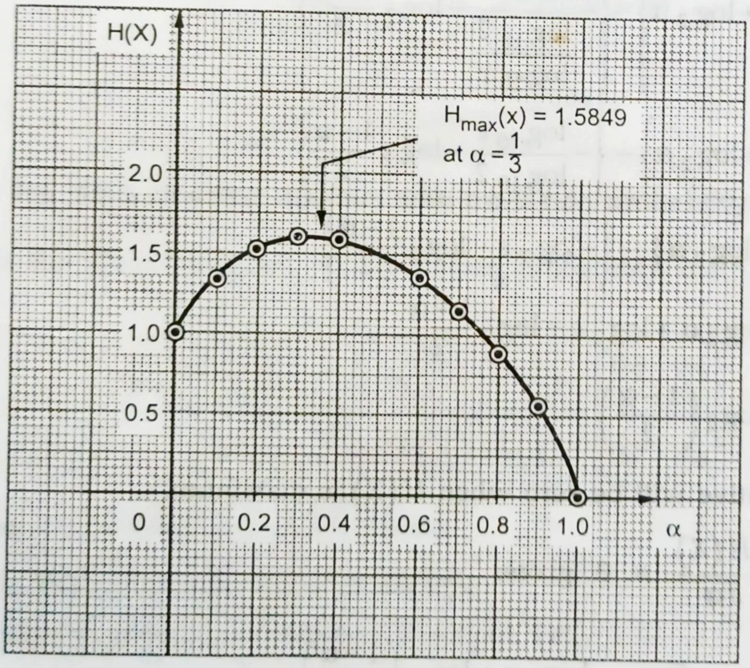


Fig. 5.6.7 Plot of $H(X)$ versus α

To obtain maximum value of $H(X)$:

Maximum value of $H(X)$ can be obtained by differentiating $H(X)$ with respect to α and equating it to zero. i.e.,

$$\frac{dH(X)}{d\alpha} = \frac{d}{d\alpha}\left[\alpha \log_2 \frac{1}{\alpha}\right] + \frac{d}{d\alpha}\left[(1-\alpha)\log_2\left(\frac{2}{1-\alpha}\right)\right]$$

We know that $\log_2 \frac{1}{\alpha} = -\log_2 \alpha$. Hence above equation becomes,

$$\frac{dH(X)}{d\alpha} = \frac{d}{d\alpha}\left[-\alpha \log_2 \alpha\right] + \frac{d}{d\alpha}\left[-(1-\alpha)\log_2\left(\frac{1-\alpha}{2}\right)\right]$$

$$= -\left\{\alpha \cdot \frac{d}{d\alpha}\frac{\log_e \alpha}{\log_e 2} + \log_2 \alpha\right\} - \left\{(1-\alpha)\frac{d}{d\alpha}\frac{\log_e \frac{1-\alpha}{2}}{\log_e 2} + \log_2\left(\frac{1-\alpha}{2}\right)\cdot(-1)\right\}$$

$$= -\left\{\alpha \cdot \frac{1}{\alpha}\cdot\frac{1}{\log_e 2} + \log_2 \alpha\right\} - \left\{(1-\alpha)\cdot\frac{1}{\frac{1-\alpha}{2}}\cdot\frac{1}{\log_e 2}\left(-\frac{1}{2}\right) - \log_2 \frac{1-\alpha}{2}\right\}$$

$$= -\left\{\frac{1}{\frac{\log_{10} 2}{\log_{10} e}} + \log_2 \alpha\right\} - \left\{-\frac{1}{\frac{\log_{10} 2}{\log_{10} e}} - \log_2 \frac{1-\alpha}{2}\right\}$$

$$= -\left\{\frac{\log_{10} e}{\log_{10} 2} + \log_2 \alpha\right\} - \left\{-\frac{\log_{10} e}{\log_{10} 2} - \log_2 \frac{1-\alpha}{2}\right\}$$

$$= -\left\{\log_2 e + \log_2 \alpha\right\} - \left\{-\log_2 e - \log_2 \frac{1-\alpha}{2}\right\}$$

$$= -\log_2 e - \log_2 \alpha + \log_2 e + \log_2 \frac{1-\alpha}{2}$$

$$= -\log_2 \alpha + \log_2\left(\frac{1-\alpha}{2}\right)$$

Above derivative will be zero when $H(X)$ is maximum.

Hence $\frac{dH(X)}{d\alpha} = 0$ gives,

$$0 = -\log_2 \alpha + \log_2\left(\frac{1-\alpha}{2}\right)$$

$$\therefore \quad \log_2 \alpha = \log_2\left(\frac{1-\alpha}{2}\right)$$

$$\therefore \qquad \alpha = \frac{1-\alpha}{2}$$

or
$$2\alpha = 1-\alpha \quad \Rightarrow \quad \alpha = \frac{1}{3}$$

Thus at $\alpha = \frac{1}{3}$, $H(X)$ will be maximum. This maximum value will be,

$$H_{\max}(X) = \frac{1}{3}\log_2 3 + \frac{2}{3}\log_2 \frac{2}{(2/3)}$$

$$= \frac{1}{3}\log_2 3 + \frac{2}{3}\log_2 3$$

$$= \log_2 3 = 1.5849$$

(ii) To determine the transition matrix :

The given channel diagram is redrawn here with transition matrix values.

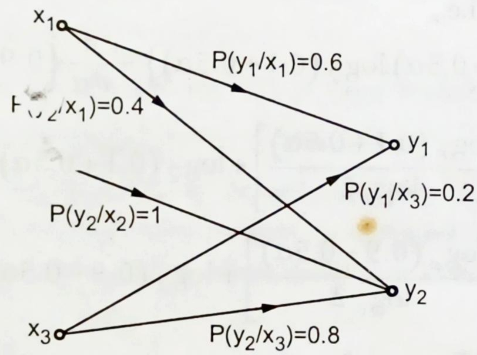


Fig. 5.6.8 Channel diagram with transition probability values

Based on the above diagram, the transition probability matrix can be written as follows :

$$P = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \\ P(y_1/x_3) & P(y_2/x_3) \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ 0 & 1 \\ 0.2 & 0.8 \end{bmatrix}$$

This is the 3×2 channel transition matrix.

(iii) To obtain the maximum value of $H(Y)$:

First we have to obtain probabilities of output. i.e.,

$$\begin{bmatrix} P(y_1) \\ P(y_2) \end{bmatrix} = \begin{bmatrix} P(x_1) & P(x_2) & P(x_3) \end{bmatrix} \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \\ P(y_1/x_3) & P(y_2/x_3) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha & \frac{1-\alpha}{2} & \frac{1-\alpha}{2} \end{bmatrix} \begin{bmatrix} 0.6 & 0.4 \\ 0 & 1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$= \begin{bmatrix} 0.6\alpha + 0 + (0.2)\frac{1-\alpha}{2} \\ 0.4\alpha + \frac{1-\alpha}{2} + (0.8)\frac{1-\alpha}{2} \end{bmatrix} = \begin{bmatrix} 0.1+0.5\alpha \\ 0.9-0.5\alpha \end{bmatrix}$$

i.e. $P(y_1) = 0.1 + 0.5\alpha$ and $P(y_2) = 0.9 - 0.5\alpha$

Hence entropy of output becomes,

$$H(Y) = P(y_1)\log_2 \frac{1}{P(y_1)} + P(y_2)\log_2 \frac{1}{P(y_2)}$$

$$= -P(y_1)\log_2 P(y_1) - P(y_2)\log_2 P(y_2)$$

$$= -(0.1+0.5\alpha)\log_2(0.1+0.5\alpha) - (0.9-0.5\alpha)\log_2(0.9-0.5\alpha)$$

Now the maximum value can be obtained by differentiating $H(Y)$ with respect to α and equating it to zero. i.e.,

$$\frac{dH(Y)}{d\alpha} = -\frac{d}{d\alpha}\{(0.1+0.5\alpha)\log_2(0.1+0.5\alpha)\} - \frac{d}{d\alpha}\{0.9-0.5\alpha\log_2(0.9-0.5\alpha)\}$$

$$= -\left\{(0.1+0.5\alpha)\frac{d}{d\alpha}\left[\frac{\log_e(0.1+0.5\alpha)}{\log_e 2}\right] + \log_2(0.1+0.5\alpha)(0.5)\right\}$$

$$-\left\{(0.9-0.5\alpha)\frac{d}{d\alpha}\left[\frac{\log_e(0.9-0.5\alpha)}{\log_e 2}\right] + \log_2(0.9-0.5\alpha)(-0.5)\right\}$$

$$= -\left\{(0.1+0.5\alpha)\frac{0.5}{(0.1+0.5\alpha)}\frac{1}{\log_e 2} + 0.5\log_2(0.1+0.5\alpha)\right\}$$

$$-\left\{(0.9-0.5\alpha)\frac{-(0.5)}{(0.9-0.5\alpha)}\frac{1}{\log_e 2} - 0.5\log_2(0.9-0.5\alpha)\right\}$$

$$= -\left\{\frac{0.5}{\log_e 2} + 0.5\log_2(0.1+0.5\alpha) - \frac{0.5}{\log_e 2} - 0.5\log_2(0.9-0.5\alpha)\right\}$$

$$= -\{0.5\log_2(0.1+0.5\alpha) - 0.5\log_2(0.9-0.5\alpha)\}$$

For maximum, $\frac{dH(Y)}{d\alpha} = 0$ gives,

$$0 = -\{0.5\log_2(0.1+0.5\alpha) - 0.5\log_2(0.9-0.5\alpha)\}$$

$$\log_2(0.1+0.5\alpha) = \log_2(0.9-0.5\alpha)$$

$$\therefore \quad 0.1+0.5\alpha = 0.9-0.5\alpha$$

$$\therefore \quad \alpha = 0.8$$

And $H(Y)$ at $\alpha = 0.8$ will be,

$$H(Y) = -(0.1+0.5\times 0.8)\log_2(0.1+0.5\times 0.8)-(0.9-0.5\times 0.8)\log_2(0.9-0.5\times 0.8)$$

$$= -0.5\log_2 0.5 - 0.5\log_2 0.5$$

$$= -\log_2 0.5 = 1$$

Thus maximum entropy of Y is $H(Y) = 1$ and it occurs at $\alpha = 0.8$.

5.7 Mutual Information