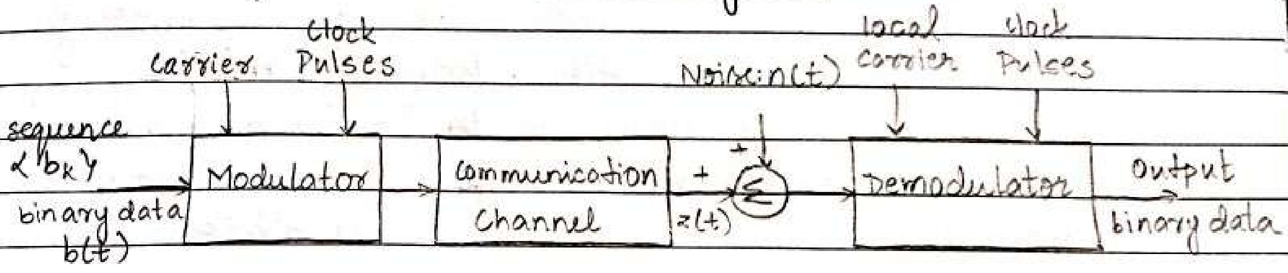
Keerthana Ashok

Page No.:
Date:

UNIT - 01

Digital Carrier Modulation Schemes

* ## General Expression and Block Diagram:



The input to the system is a binary sequence $\{b_k\}$ with an amplitude A and bit duration T_b. The output of the modulator during k^{th} interval depends on k^{th} input bit b_k. The modulator output $z(t)$ during the k^{th} interval is the shifted version of one of the two basic waveforms $s_1(t)$ or $s_2(t)$ and is given by:

$$z(t) = \begin{cases} s_1(t-(k-1)T_b) & \text{if } b_k = 0 \\ s_2(t-(k-1)T_b) & \text{if } b_k = 1 \end{cases}$$

The shape of the waveform depends on the type of modulation used.

– Choice of signaling waveforms for various types of Digital Modulation Schemes:

The frequency of the carrier f_c is assumed to be a multiple of reciprocal of T_b where $r_b = 1/T_b$

	$s_1(t)$; $0 \leq t \leq T_b$	$s_2(t)$; $0 \leq t \leq T_b$	Type of modulation
1.	0	$A_c \cos\omega_c t$ / $A_c \sin\omega_c t$	ASK
2.	$-A_c \cos\omega_c t$ / $-A_c \sin\omega_c t$	$A_c \cos\omega_c t$ / $A_c \sin\omega_c t$	PSK

	$s_1(t)$; $0 \le t \le T_b$	$s_2(t)$; $0 \le t \le T_b$	Type of Modulation
3.	$A_c \cos(\omega_c - \omega_d)t$ / $A_c \sin(\omega_c - \omega_d)t$	$A_c \cos(\omega_c + \omega_d)t$ / $A_c \sin(\omega_c + \omega_d)t$	FSK

The output of the modulator passes through a bandpass communication channel [$H_c(f)$]

The channel is assumed to be ideal with an adequate bandwidth so that the signal passes through the channel without any distortion other than propagation delay.

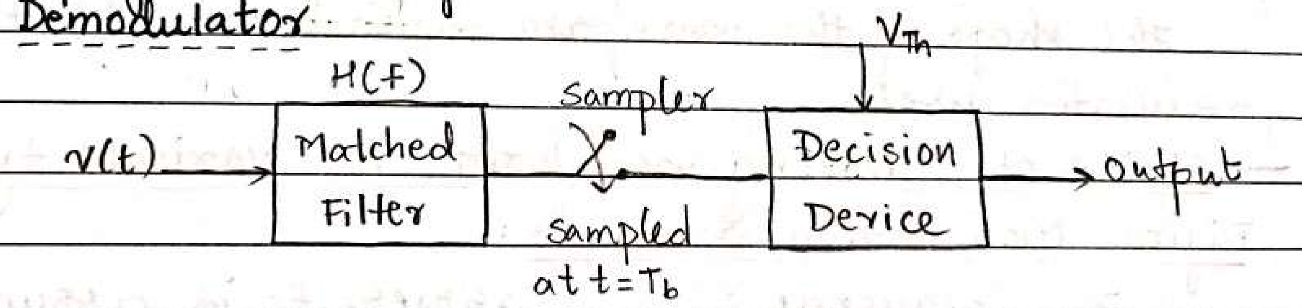
The input to the demodulator is nothing but the received signal $x(t)$ and additive noise $n(t)$.

The input to the demodulator is given by:

$$v(t) = \begin{cases} s_1(t-(k-1)T_b - T_d) + n(t) \\ s_2(t-(k-1)T_b - T_d) + n(t) \end{cases}$$

where $n(t)$ is the channel noise which is nothing but zero mean Gaussian Noise and T_d is the propagation delay.

Demodulator



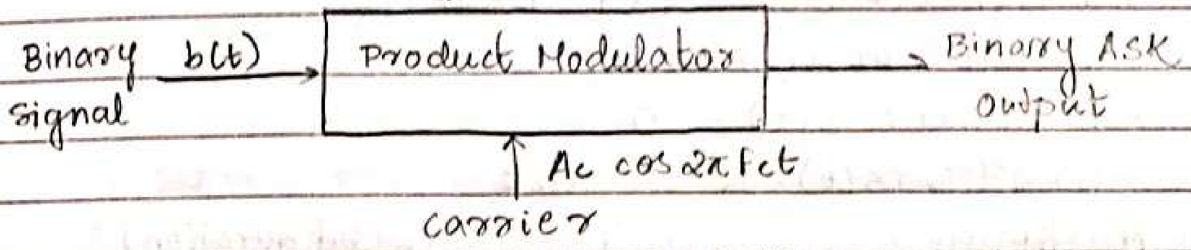
Any digital modulation scheme must satisfy the following important requirements:

- Maximum data rate
- Minimum probability of error
- Minimum transmitted power
- Minimum channel bandwidth
- Maximum resistance to interferance signals
- Minimum circuit complexity

* ASK Generation and Detection:

The amplitude of the carrier changes in accordance with the binary or discrete data.

The presence of carrier indicates transmission of 1 and the absence of carrier indicates transmission of 0.



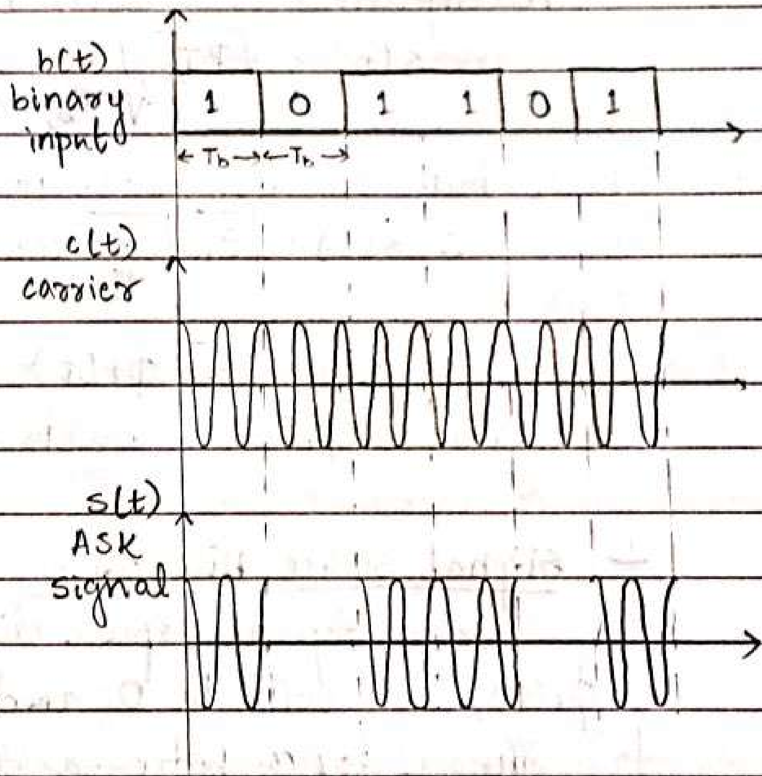
Let b(t) be the binary signal applied to the product modulator as the input.

Let c(t) be the carrier given by:

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

where A_c is the peak amplitude of the carrier and f_c is the frequency of the carrier.

Assume $f_c > 1/T_b$



To represent the amplitude of the carrier in terms of power P, consider the power dissipated across a standard 1Ω resister.

wkt $P = \frac{V^2}{R} = V^2 = \left(\frac{A_c}{\sqrt{2}}\right)^2 = \frac{A_c^2}{2}$

Therefore $A_c^2 = 2P$

$A_c = \sqrt{2P}$

Hence

$c(t) = \sqrt{2P} \cos 2\pi f_c t$

The product modulator modulates the amplitude of sinusoidal carrier using v(t) as modulating signal. The modulator passes the carrier when the input bit is 1 and blocks the carrier when input bit is 0.

For bit 1 or symbol 1

ASK signal : $s(t) = \sqrt{2P} \cos 2\pi f_c t$

For bit 0 or symbol 0

ASK signal : $s(t) = 0$

consider $s(t) = \sqrt{2P} \cos 2\pi f_c t$

Multiplying and dividing by T_b (bit duration)

$$s(t) = \sqrt{PT_b} \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t$$

but $E_b = PT_b$ (bit energy)

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

$$s(t) = \sqrt{E_b}\, \phi_1(t)$$

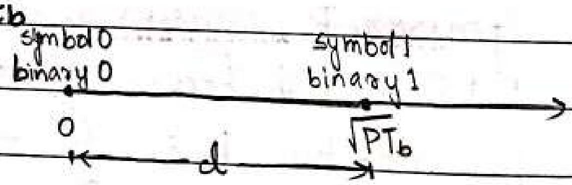
where $\phi_1(t)$ is Basis function $= \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t$

– Signal Space Diagram

The signal space diagram will have two points on $\phi_1(t)$, one will be 0 and the other will be at $\sqrt{PT_b}$.

The distance between these two points is d and is given by :

$$d = \sqrt{PT_b} - 0 = \sqrt{PT_b} = \sqrt{E_b}$$

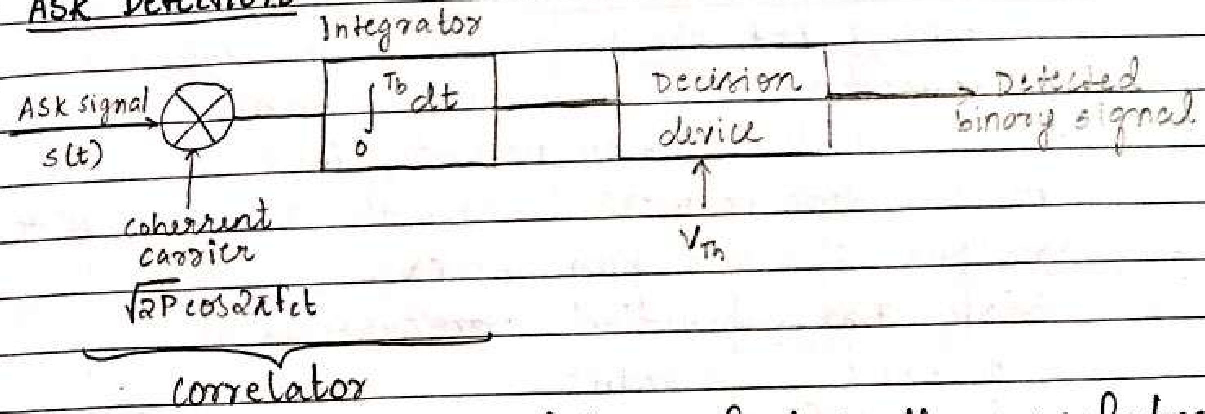


– Transmission Bandwidth:

Minimum transmission bandwidth required to pass ASK signal is $2f_b$ where $f_b = 1/T_b$.

Keerthana Ashok

- ASK Detection



The binary ASK signal is applied to the correlator which is nothing but a multiplier followed by an integrator. The output of the multiplier is integrated over the bit period and the decision device compares the output of the integrator with built in threshold V_{Th} and uses the following decision logic.

Output of an integrator or input to the decision device greater than V_{Th} : The decision is in favour of bit 1. Similarly output of the integrator or input to the decision device lesser than V_{Th} : The decision is in favour of bit 0.

- Advantages of ASK
  - High bandwidth efficiency
  - ASK modulation and demodulation is relatively inexpensive.
  - ~~Used to transmit digital data over optic fiber.~~

- Disadvantages of ASK
  - Low power efficiency.
  - susceptable to noise interference.

(BFSK is less susceptible to noise interference)

- Applications:
  - used to transmit digital data over optic fiber.
  - used in multichannel telegraph systems.

* **BFSK : Binary Frequency shift Keying:**

In Binary FSK the frequency of the carrier is shifted according to the binary symbol. The phase and amplitude of the carrier is unaffected.

It uses two carriers having the same amplitude and phase but different frequencies.

Using Basis function, the carriers are represented by:

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos 2\pi f_H t$$

$$\phi_2(t) = \sqrt{\frac{2}{T_b}} \cos 2\pi f_L t$$

where f_H and f_L should be integer multiple of the base band frequency $f_b = 1/T_b$.

Symbol 1 or bit 1 can be represented by

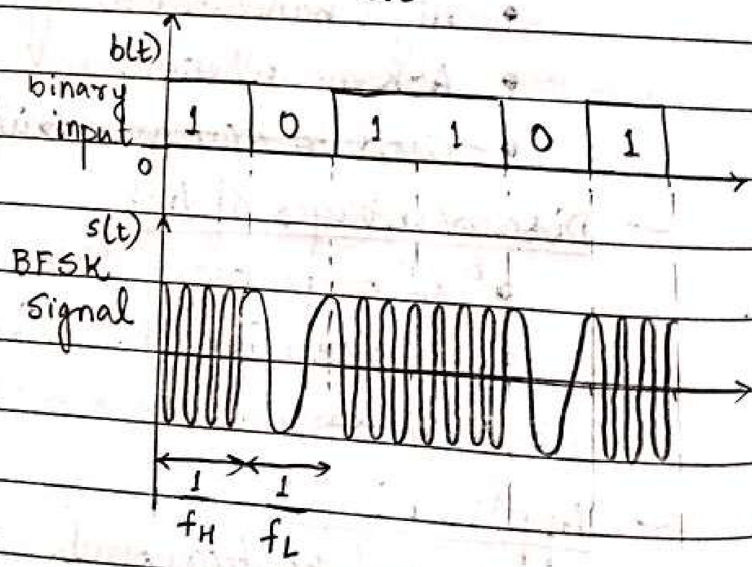
$$s_H(t) = \sqrt{PT_b}\sqrt{\frac{2}{T_b}} \cos 2\pi f_H t$$

Similarly symbol 0 or bit 0 can be represented by

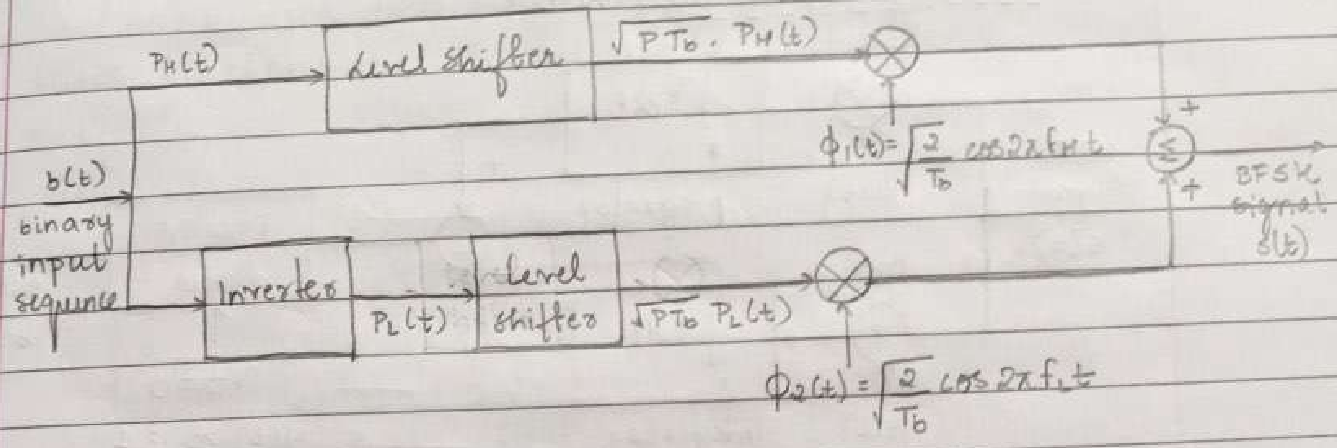
$$s_L(t) = \sqrt{PT_b}\sqrt{\frac{2}{T_b}} \cos 2\pi f_L t$$

where $f_H = f_0 + \frac{\Omega}{2\pi}$ and $f_L = f_0 - \frac{\Omega}{2\pi}$

In binary frequency shift keying (BFSK), the bit 1 or symbol 1 and bit 0 or symbol 0 are represented by carrier frequency being shifted above or below the centre frequency (f_0) respectively.



- BFSK Transmitter:



From the figure we see that $P_H(t) = P(t)$ and $P_L(t)$ is the inverted version of $P(t)$.

An inverter is added after $b(t)$ to get $P_L(t)$.

$P_H(t)$ and $P_L(t)$ are unipolar signals.

The level shifter converts $+1$ level to $\sqrt{PT_b}$ and zero level is unaffected. Thus the output of the level shifter will be either $\sqrt{PT_b}$ for binary 1 or zero for binary 0.

The two product modulators use carriers $\phi_1(t)$ and $\phi_2(t)$ respectively. These carriers are orthogonal to each other.

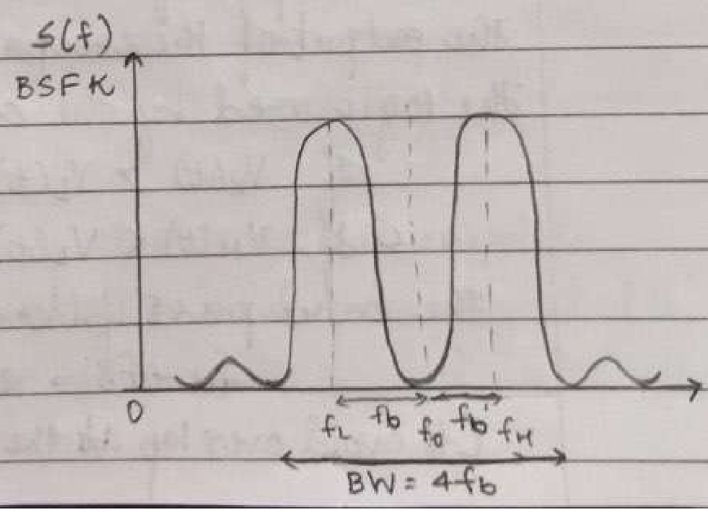
The output of these modulators are given to the summer to get BSFK output.

Output of the transmitter is

$$s_{BFSK}(t) = \sqrt{PT_b}\ P_H(t)\,\phi_1(t) + \sqrt{PT_b}\ P_L(t)\,\phi_2(t)$$

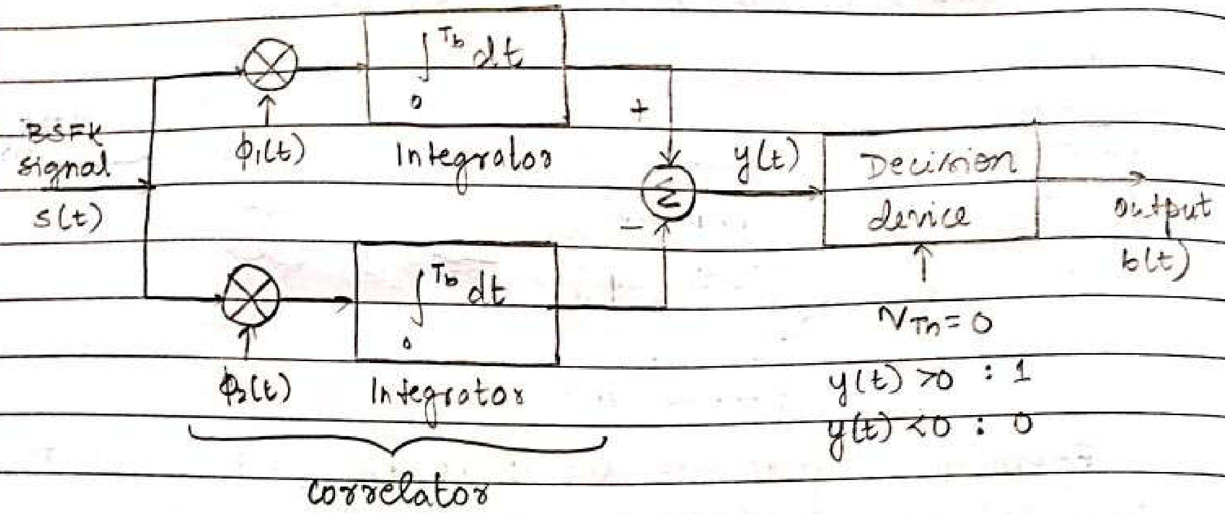
- Bandwidth of BSFK:

Bandwidth of BSFK is $4f_b$ which is two times the bandwidth of ASK.

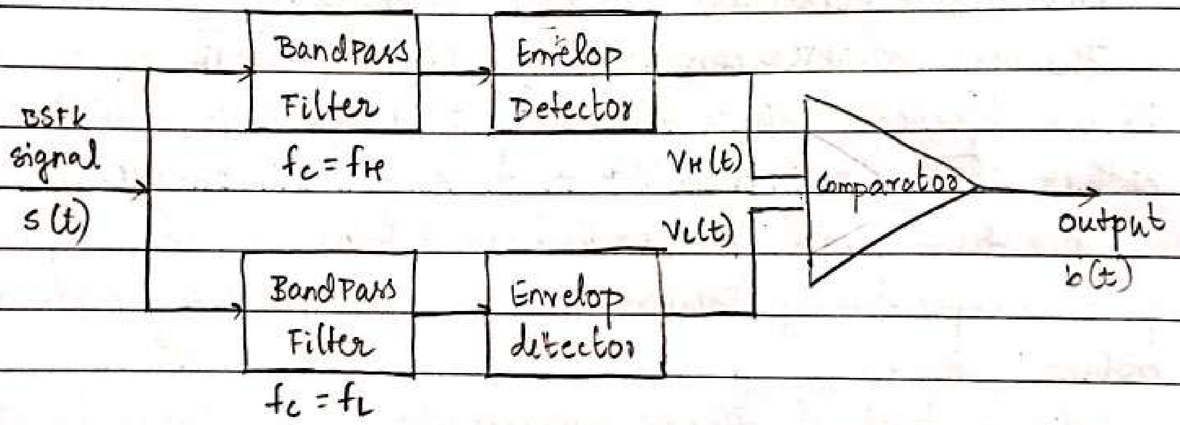


– BSFK detection:

1. coherent BSFK receiver:



2. Non coherent FSK receiver:



Band pass filter frequency variations into amplitude variations and the enelop detectors detects these amplitude variations.

The output of the detectors are compared by the comparator. The output of the comparator is the bit sequence b(t). The logic used by the comparator is as follows.

If $V_H(t) > V_L(t)$: bit 1

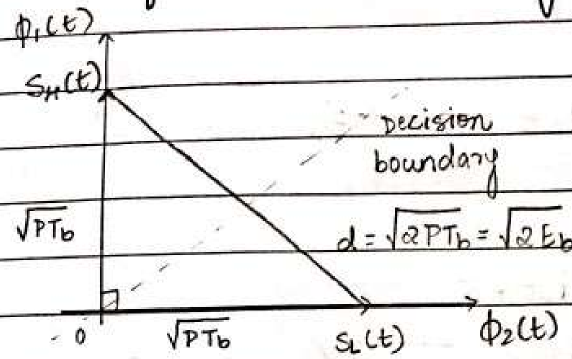
If $V_H(t) < V_L(t)$: bit 0

The band pass filters are selected in such a way that $f_H - f_L = 2f_b$ is the center frequency (to avoid over lap at the output of the filters.

Keerthana Ashok

- **Signal Space Diagram**

BSFK uses two basis function $\phi_1(t)$ and $\phi_2(t)$. These basis functions are orthogonal to each other



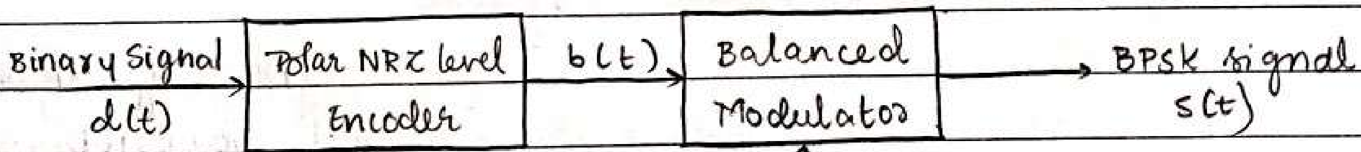
★ **Binary PSK or PSK (Phase Shift keying):**

The phase of the carrier changes in accordance with the binary input data.

- **Phase transition table:**

Binary Inputs	Carrier Phase change
$0 \rightarrow 0$	No phase change
$0 \rightarrow 1$	} phase change of 180°
$1 \rightarrow 0$	
$1 \rightarrow 1$	No phase change

- **BPSK or PSK Transmitter**



Binary 1 → +ve pulse
Binary 0 → −ve pulse

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

carrier $= \sqrt{2P} \cos 2\pi f_c t$

Binary 1

$b(t) = +1 \; ; 0 \le t \le T_b$

Binary 0

$b(t) = -1 \; ; 0 \le t \le T_b$

Polar NRZ level encoder converts unipolar signal $d(t)$ into polar signal $b(t)$

Keerthana Ashok

In BPSK when the symbol or bit is changed from 0 to 1 or 1 to 0, the phase of the carrier is changed by 180°

Therefore

for symbol 1 or bit 1

$$s_1(t) = \sqrt{2P}\cos 2\pi f_c t \quad \text{——} (1)$$

Similarly

for symbol 0 or bit 0

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

$$= -\sqrt{2P}\cos 2\pi f_c t \quad \text{——} (2)$$

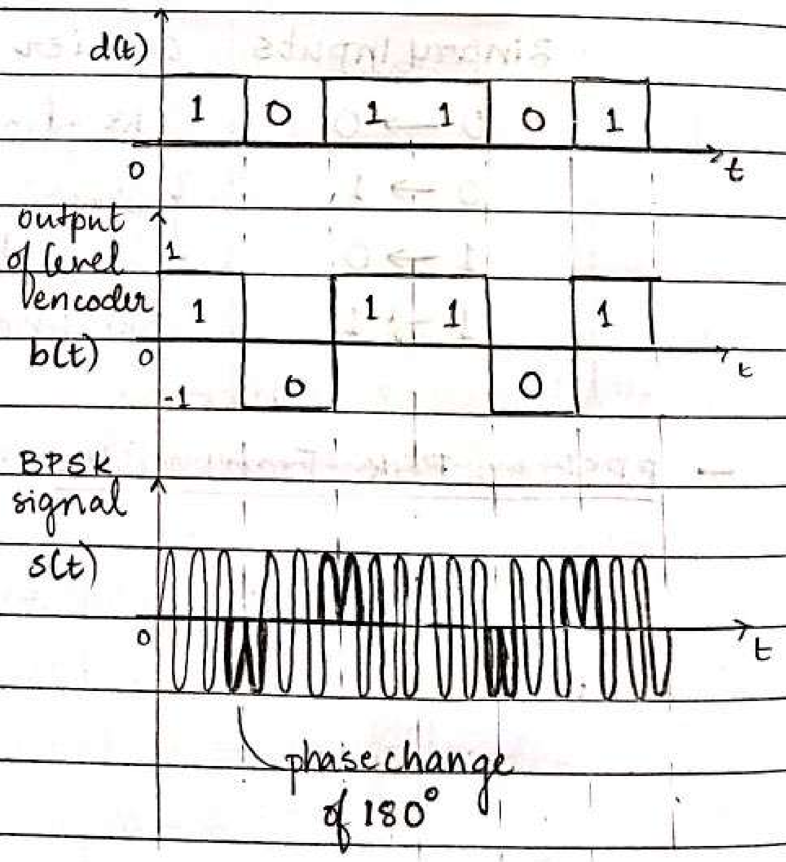
From eq (1) and eq (2), defining BPSK signal as

$$s(t) = b(t)\sqrt{2P}\cos 2\pi f_c t \quad \text{——} (3)$$

where $b(t) = \pm 1$

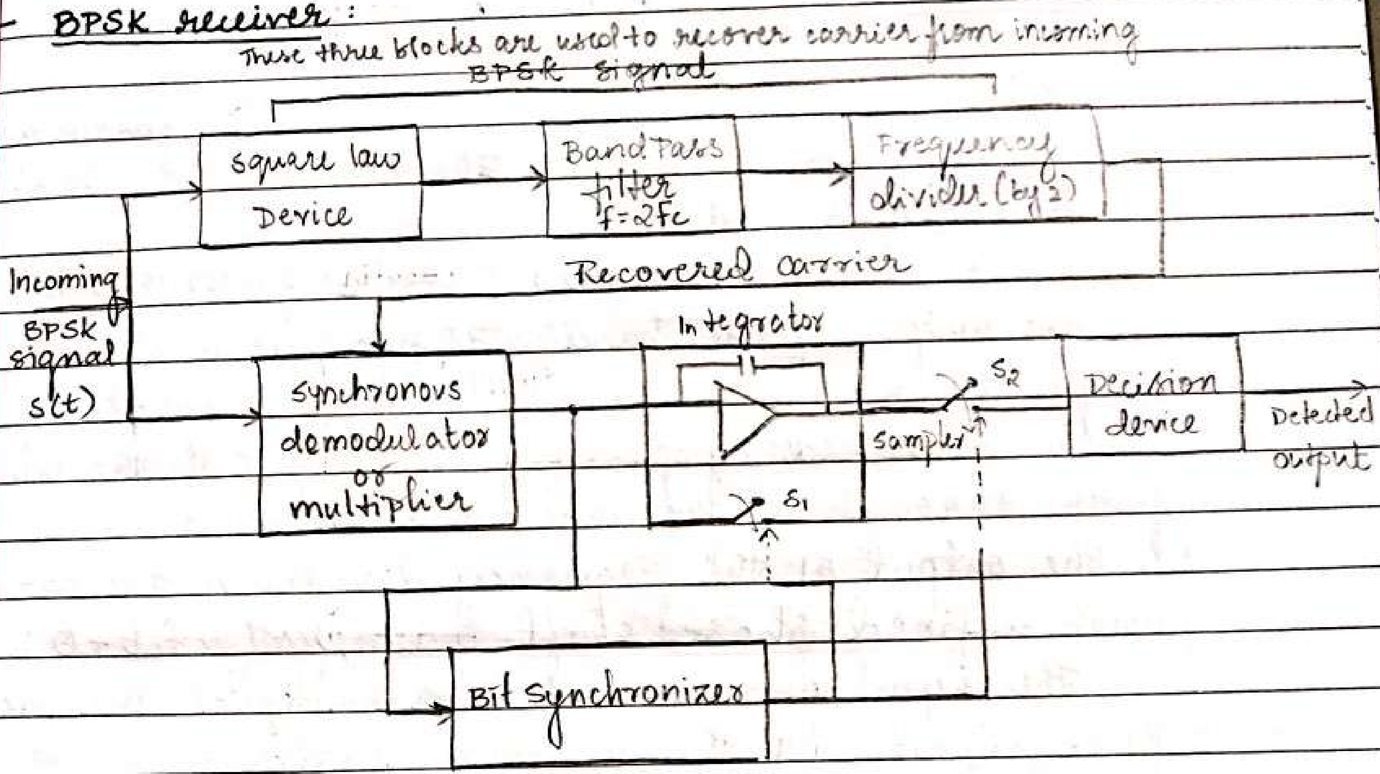
wkt $b(t) = +1$: when bit 1 is transmitted

$b(t) = -1$: when bit 0 is transmitted



Keerthana Ashok

- **BPSK receiver:**



The circuit uses synchronous detection, the carrier used for the detection is derived from the incoming BPSK signal.

The transmitted BPSK signal is given by:

~~BPSK Receiver~~):

$$s(t) = b(t)\sqrt{2P}\cos 2\pi f_c t \quad \text{——— (1)}$$

where $f_c \gg f_b$ and $f_b = 1/T_b$

This signal undergoes the phase change depending upon the time delay from transmitter to receiver. This phase change is fixed and is indicated by θ.

Therefore the BPSK signal at the input of the receiver is:

$$s(t) = b(t)\sqrt{2P}\cos(2\pi f_c t + \theta) \quad \text{——— (2)}$$

This signal is passed through the square law device and the output of the square law device will be

$$\cos^2(2\pi f_c t + \theta) \quad \text{(magnitude is neglected)}$$

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

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

the dc component (1/2) is eliminated by passing through a bandpass filter. The output of the band pass filter is

$\cos 2(2\pi f_c t+\theta)$ (scaling factor is eliminated)

The output of the bandpass filter is a signal of frequency $2f_c$ and hence the original carrier frequency f_c is recovered by passing the output of the bandpass filter through a frequency divider by 2.

The output of the frequency divider is the carrier with a fixed phase shift θ i.e., $\cos(2\pi f_c t+\theta)$.

The synchronous demodulator multiplies the incoming BPSK signal with the recovered carrier.

Therefore the output of the synchronous demodulator is:

$$s(t)\cos(2\pi f_c t + \theta)$$
$$= b(t)\sqrt{2P}\cos(2\pi f_c t+\theta)\cos(2\pi f_c t+\theta)$$
$$= b(t)\sqrt{2P}\cos^2(2\pi f_c t+\theta)$$
$$= b(t)\sqrt{2P}\left[\frac{1+\cos 2(2\pi f_c t+\theta)}{2}\right]$$
$$= b(t)\sqrt{\frac{P}{2}}\left[1+\cos 2(2\pi f_c t+\theta)\right]$$

The output of the synchronous demodulator or multiplier is given to the integrator as input. The integrates the signal over the bit period.

The bit synchronizer takes care of the starting and ending time of a bit. At the end of bit duration the bit sychronizer closes the switch S_2 (acts as a sampler) temporarily. This connects the output of the integrator to the decision device (which is equivalent to samping the output of the integrator)

The bit synchronizer opens switch S2 and closes switch S1 temporarily. This resets the integrator voltage to zero. In the kth bit interval, the output of the integrator is:

$$s_0(kT_b) = b(kT_b)\sqrt{\frac{P}{2}}\int_{(k-1)T_b}^{kT_b}\left(1+\cos 2(2\pi f_c t+\theta)\right)dt$$

$$= b(kT_b)\sqrt{\frac{P}{2}}\left[t + \frac{\sin 2(2\pi f_c t+\theta)}{2(2\pi f_c)}\right]_{(k-1)T_b}^{kT_b}$$

$$= b(kT_b)\sqrt{\frac{P}{2}}\left[T_b\right.$$

$$=$$

$$= b(kT_b)\sqrt{\frac{P}{2}}\,T_b$$

For a given bit duration and power, the output of the integrator: $s_0(kT_b) \propto b(kT_b)$

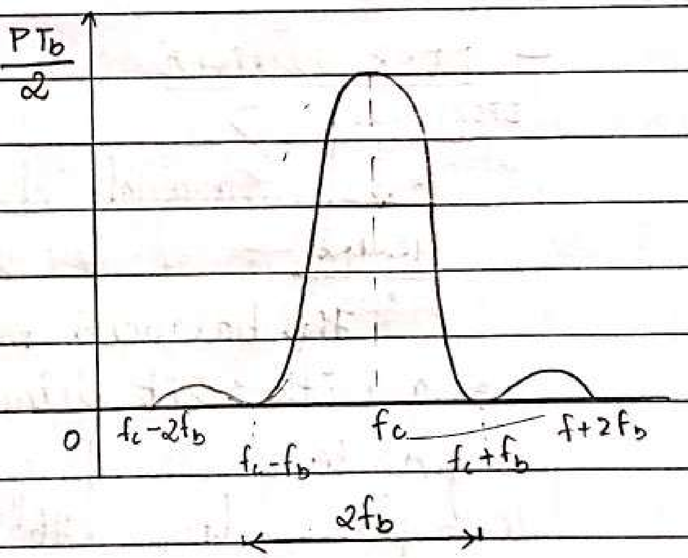
Depending upon the value of $b(kT_b)$, the output $s_0(kT_b)$ is generated at the receiver. This signal is then given to the decision device which decides whether the transmitted bit is 0 or 1.

- Bandwidth of BPSK:

Bandwidth of BPSK

$= (f_c+f_b)-(f_c-f_b)$

$= 2f_b$



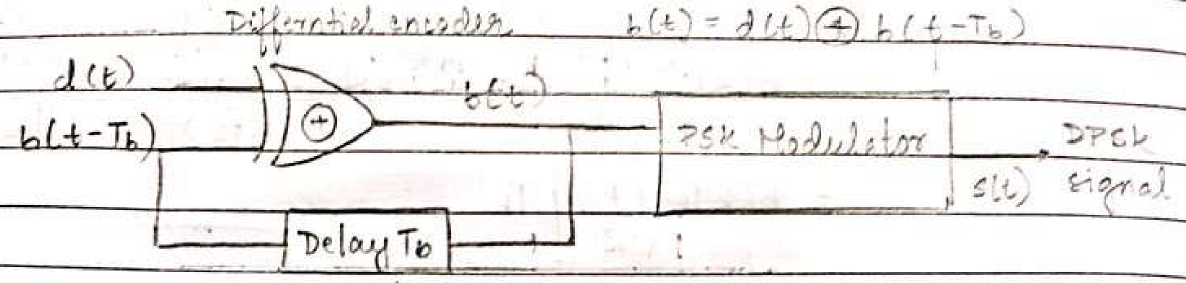
- Drawbacks:
  - Bandwidth is more.
  - For a given error rate, the signal power required is 4 times as that of ASK.

Keerthana Ashok

★ Differential PSK: (DPSK)

Prior to PSK modulator, the circuit uses a differential encoder and hence the name differential PSK (DPSK).

– DPSK modulator (transmitter)



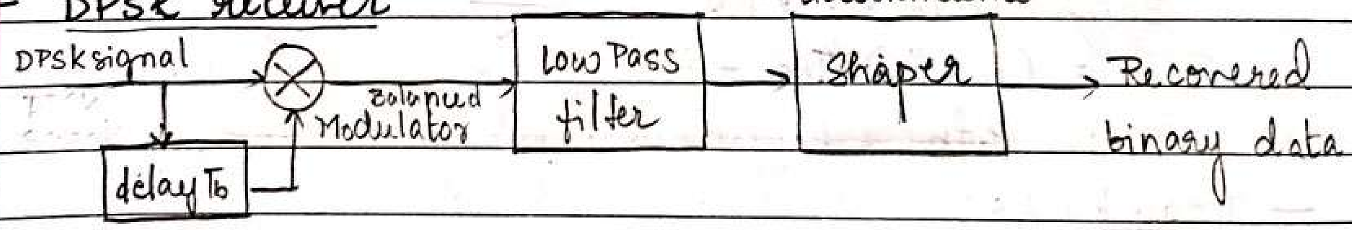
Assume initial bit / implied bit as 0.

$d(t)$	:	0	1	0	0	1	1	0	0	1	
$b(t-T_b)$	:	0 →	0 →	1 →	1 →	1 →	0 →	1 →	1 →	1 →	0
$b(t)$	:	0	1	1	1	0	1	1	1	0	

– Phase transition table for PSK Modulator

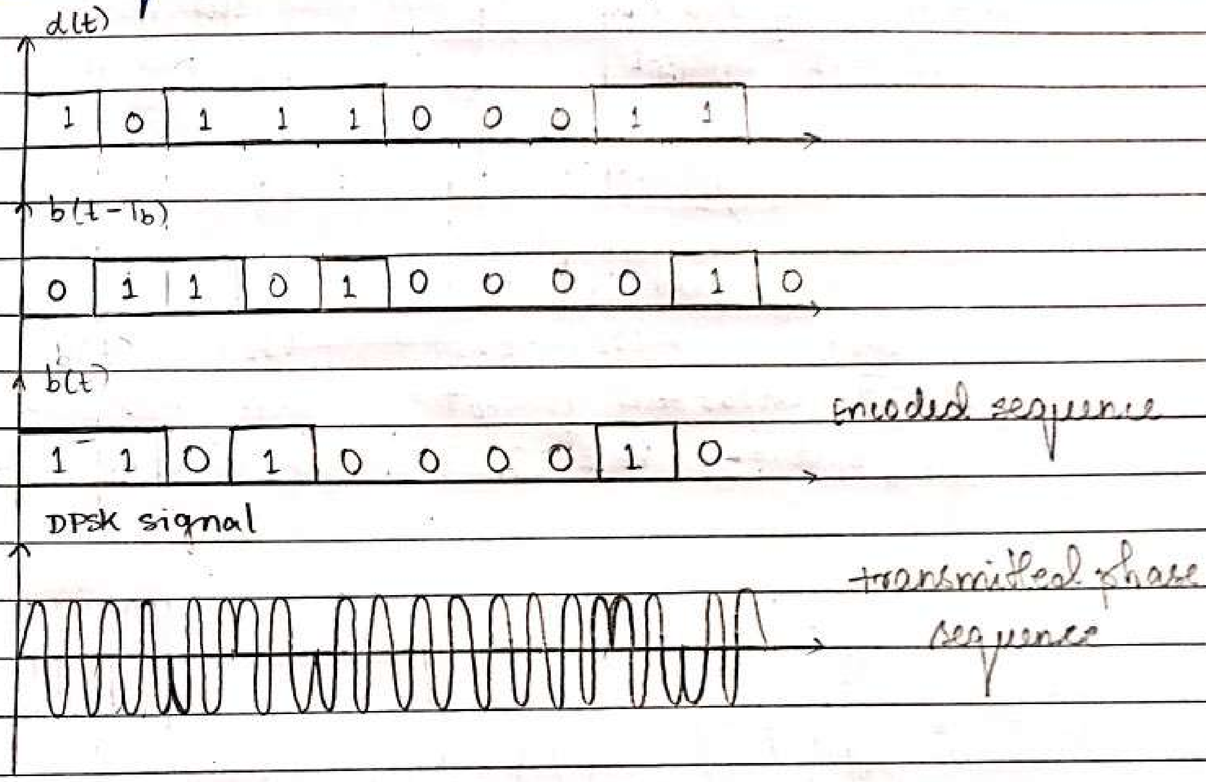
Inputs	Carrier Phase change
0 → 0	No change
0 → 1	Phase change of 180°
1 → 0	Phase change of 180°
1 → 1	No change

– DPSK receiver



The balanced modulator is given the DPSK signal along with 1 bit delay input. That signal is made to confine to lower frequencies with the help of the low pass filter. It is then passed to a shaper circuit which is a comparator or a schmitt trigger circuit, to recover the original binary data as the output.

Q: The bit stream 1011100011 is to be transmitted using DPSK. Determine the encoded sequence and the transmitted phase sequence.



★ <u>Quadrature PSK</u> : (QPSK)

In ASK, FSK, PSK and DPSK modulation techniques one bit is transmitted at a time. To increase the capacity, instead of transmitting one bit, transmit two bits (dibits) at a time.

Now, there are four distinct symbols (00, 01, 10, 11). These symbols are represented by using four quadrature carriers having the same frequency and amplitude but different in phase.

The phase transition table is as shown below:

Input dibits	symbol	Phase shift in carrier
10	S_1	$\pi/4$ or 90°
00	S_2	$3\pi/4$ or 0°
01	S_3	$5\pi/4$ or 270°
11	S_4	$7\pi/4$ or 180°

Keerthana Ashok

- constellation diagram and phasor diagram:

(00)

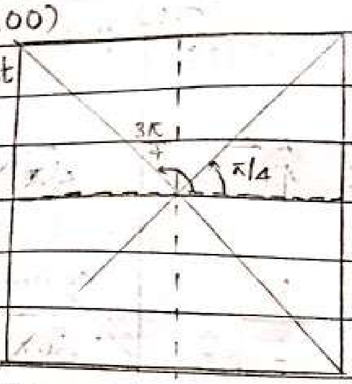
$s(t) = -\sqrt{P}\cos 2\pi f_0 t - \sqrt{P}\sin 2\pi f_0 t$

$b_o(t) = -1 \quad b_e(t) = -1$

(10)

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

$b_o(t) = 1$ (odd) $\quad b_e(t) = -1$ (even)



(01)

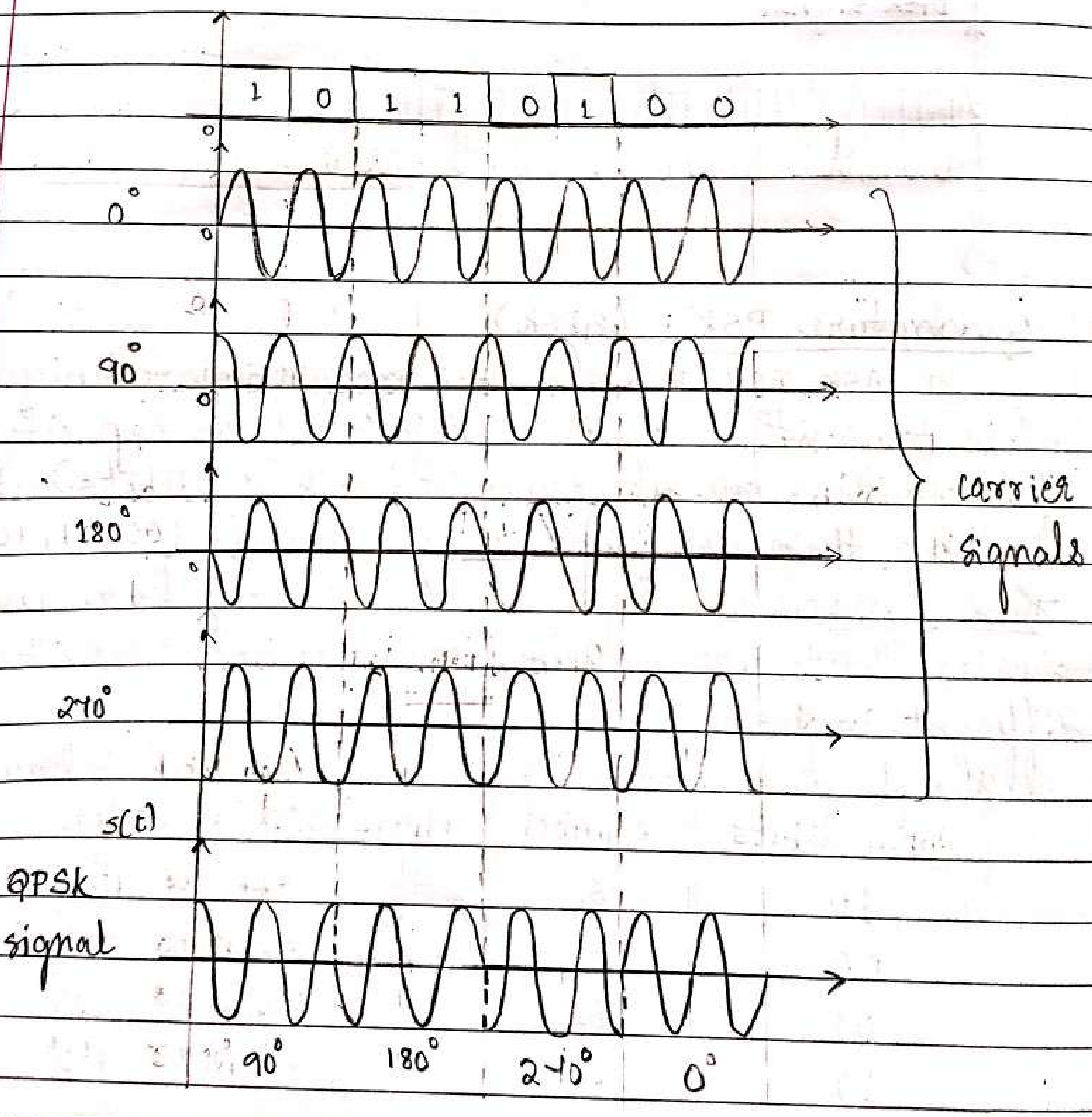
$s(t) = -\sqrt{P}\cos 2\pi f_0 t + \sqrt{P}\sin 2\pi f_0 t$

$b_o(t) = -1 \quad b_e(t) = 1$

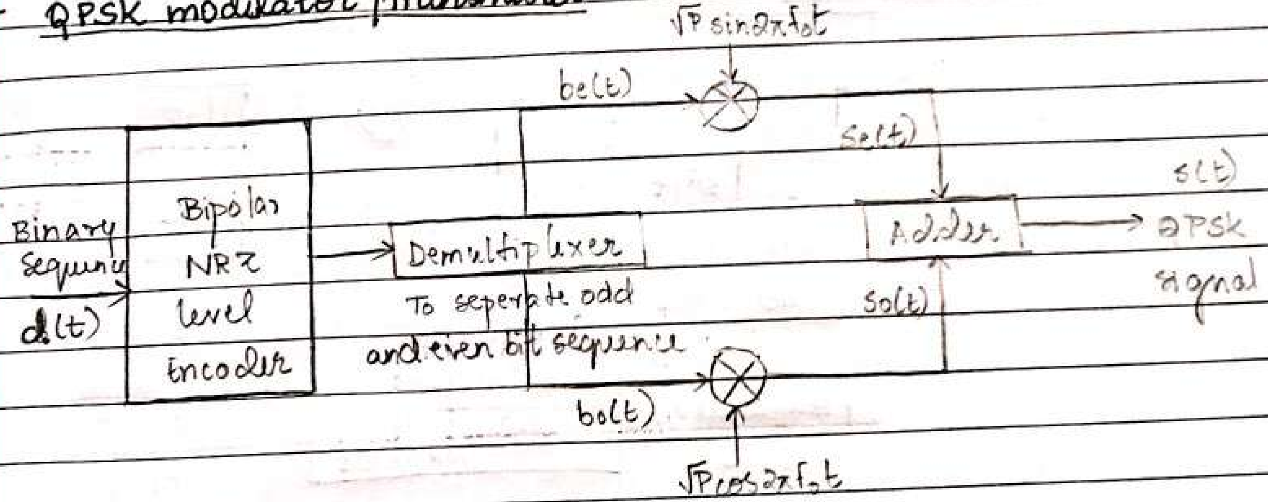
(11)

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

$b_o(t) = 1 \quad b_e(t) = 1$



- QPSK modulator/transmitter:

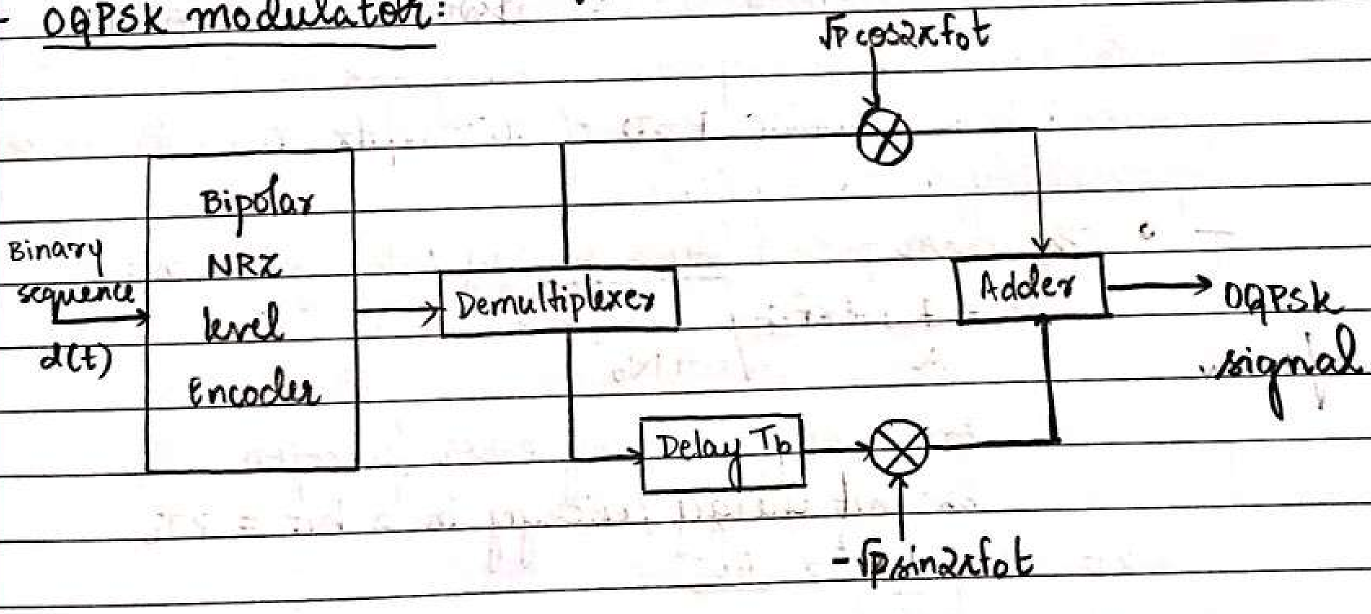


* Offset QPSK: (OQPSK)

It is impossible to get odd bit and even bit at a time. One bit delay exists between the odd and the even bits and hence the name.

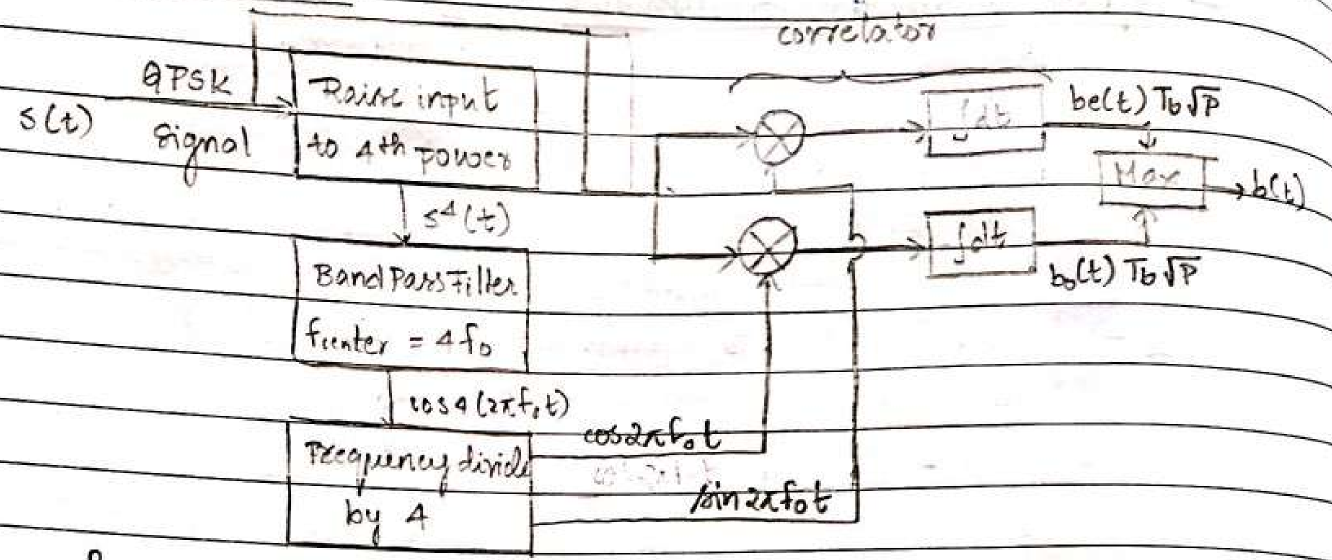
change in even sequence and change in odd sequence occur because of the offset.

- OQPSK modulator:



Keerthana Ashok

- QPSK receiver:



Inphase carrier $\cos 2\pi f_0 t$ and quadrature carrier $\sin 2\pi f_0 t$ are recovered from the incoming QPSK signal.

* Error Performance of Binary systems:

- Probability of error: (Pe)

Q: Binary data is transmitted over an RF bandpass channel having a bandwidth of 10MHz at a rate of 4.8Mbps, with 1mV carrier amplitude using ASK technique. The channel has a noise PSD of 10^{-15} W/Hz. Find the error probability of the system.

- • The error probability of binary ASK system is:

$$P_e = \frac{1}{2} \operatorname{Erfc} \sqrt{\frac{E_b}{4N_0}}$$

Erfc: complimentary error function

E_b: bit energy / energy in a bit $= PT_b$

wkt $P = \left(\frac{A_c}{\sqrt{2}}\right)^2 = \frac{A_c^2}{2}$

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

Scanned with CamScanner

Given: bit rate : $r_b = 4.8\ Mbps = 4.8 \times 10^6\ bps$

bit period : $T_b = \frac{1}{r_b} = \frac{1}{4.8 \times 10^6} = 2.083 \times 10^{-7}\ sec$

bit energy : $E_b = PT_b$

$E_b = \frac{A_c^2 T_b}{2} = \frac{[1 \times 10^{-3}]^2 2.083 \times 10^{-7}}{2} = 1.0415 \times 10^{-13}\ J$

noise PSD : $\frac{N_0}{2} = 10^{-15}\ W/Hz$

$N_0 = 2 \times 10^{-15}\ W/Hz$

Therefore, the error probability of binary ASK system is :

$P_e = \frac{1}{2}\, erfc \sqrt{\frac{E_b}{4N_0}}$

$P_e = \frac{1}{2}\, erfc \sqrt{\frac{1.0415 \times 10^{-13}}{4 \times 2 \times 10^{-15}}}$

$P_e = \frac{1}{2}\, erfc\,(3.6)$

wkt $erfc\,(u) = 1 - erf(u)$ erf : error function

Assume $erf\,(3.6) = 0.9999998$

$erfc\,(3.6) = 1 - erf\,(3.6) = 1 - 0.9999998 = 0.0000002$

$erfc\,(3.6) = 2 \times 10^{-7}$

Therefore

$P_e = \frac{1}{2}(2 \times 10^{-7}) = 10^{-7}$

NOTE: Greater the carrier amplitude / power lesser the probability error. Lesser the noise PSD, lesser the probabity error.

Q: Binary data is transmitted over a microwave link at a rate of 10^6 bps and the noise PSD at the receiver input is $10^{-10}\ W/Hz$. Find the average carrier power required to maintain an average probability of error $P_e \leq 10^{-4}$ for coherent binary FSK. What is the required channel bandwidth

Keerthana Ashok

- The probability of error for a coherent binary FSK system is:

$$P_e = \frac{1}{2} \operatorname{Erfc} \sqrt{\frac{0.6 E_b}{N_0}}$$

Given:

data rate: $r_b = \frac{1}{T_b} = 10^6$ bps.

noise PSD: $\frac{N_0}{2} = 10^{-10}$ watt/Hz

$\Rightarrow N_0 = 2 \times 10^{-10}$ watt/Hz

Probability of error : $P_e \leq 10^{-4}$

$$P_e = \frac{1}{2} \operatorname{Erfc} \sqrt{\frac{0.6 E_b}{N_0}}$$

$$10^{-4} = \frac{1}{2} \operatorname{Erfc} \sqrt{\frac{0.6 E_b}{2 \times 10^{-10}}}$$

$$\operatorname{Erfc} \sqrt{0.3 \times 10^{10} E_b} = 2 \times 10^{-4}$$

$$1 - \operatorname{Erf} \sqrt{0.3 \times 10^{10} E_b} = 2 \times 10^{-4}$$

$$\operatorname{Erf} \sqrt{0.3 \times 10^{10} E_b} = 0.9998$$

From the error function table we have

$$\operatorname{Erf}(2.8) \simeq 0.9998$$

Therefore

$$\sqrt{0.3 \times 10^{10} E_b} = 2.8$$

$$E_b = \frac{(2.8)^2}{0.3 \times 10^{10}}$$

$$\underline{E_b = 2.613 \times 10^{-9} \text{ J}}$$

Average carrier power

$$P = \frac{E_b}{T_b} = 2.613 \times 10^{-9} (10^6)$$

$$\underline{P = 2.613 \times 10^{-3} \text{ Watt}}$$

To achieve $P_e \leq 10^{-4}$, average power of the carrier must be $P \geq 2.613 \times 10^{-3}$ watt.

The channel bandwidth is approximately equal to the bit rate i.e., 10^6 Hz.

For FSK, $BW = 4f_b = \frac{4}{T_b} = 4 \times 10^6$ Hz

Q: Binary data is transmitted 10^6 bits per sec over a channel of bandwidth 3MHz. Assume that the noise PSD at the receiver is 10^{-10} W/Hz. Find the average carrier power required at the receiver input for coherent PSK and non coherent DPSK signaling sequence to maintain a probability error $P_e = 10^{-4}$.

- The probability of error for coherent PSK is:

$$P_e = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{N_0}}$$

Given: For coherent PSK signaling.

data rate: $r_b = \frac{1}{T_b} = 10^6$ bps

noise PSD: $\frac{N_0}{2} = 10^{-10}$ watt/Hz

$$\Rightarrow N_0 = 2 \times 10^{-10} \text{ watt/Hz}$$

Probability of error: $P_e = 10^{-4}$

$$P_e = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{N_0}}$$

$$10^{-4} = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{2 \times 10^{-10}}}$$

$$\operatorname{erfc} \sqrt{0.5 \times 10^{10} E_b} = 2 \times 10^{-4}$$

$$1 - \operatorname{erf} \sqrt{0.5 \times 10^{10} E_b} = 2 \times 10^{-4}$$

$$\operatorname{erf} \sqrt{0.5 \times 10^{10} E_b} = 0.9998$$

From the error function table we have

$$\operatorname{erf}(2.6) \simeq 0.9998$$

Therefore

$$\sqrt{0.5 \times 10^{10} E_b} = 2.6$$

$$E_b = 1.35 \times 10^{-9} \text{ J}$$

Average carrier power

$$P = \frac{E_b}{T_b} = 1.35 \times 10^{-9} \times 10^{6}$$

$$P = 1.35\,m\,Watt$$

To achieve $P_e = 10^{-4}$, average power of the carrier must be $P = 1.35\,m\,Watt$.

- The probability of error for noncoherent DPSK is

$$P_e = \frac{1}{2} e^{-E_b/N_0}$$

Therefore: $10^{-4} = \frac{1}{2} e^{-E_b/2\times10^{-10}}$

$$e^{-E_b/2\times10^{-10}} = 2\times10^{-4}$$

$$\frac{-E_b}{2\times10^{-10}} = -8.51$$

$$E_b = 1.7\times10^{-9}\,J$$

Average carrier power

$$P = \frac{E_b}{T_b} = 1.7\times10^{-9}\times10^{6}$$

$$P = 1.7\,m\,Watt$$

To maintain the probability of error of 10^{-4} noncoherent DPSK signalling scheme requires more carrier power (1.7 mWatt) that of coherent PSK signaling scheme.

Q: Compare the average power requirements of binary non coherent ASK, coherent PSK, non coherent DPSK and non coherent FSK signalling schemes operating at a rate of 10^3 bps over a channel of bandwidth 3MHz. Assume noise PSD $= 10^{-10}$ W/Hz and probability of error $P_e = 10^{-5}$. Calculate T_b.

Keerthana Ashok

Given:

data rate: $r_b = 10^3$ bps Probability of error $P_e = 10^{-5}$

$T_b = \frac{1}{r_b} = 10^{-3}$ sec

noise PSD: $\frac{N_0}{2} = 10^{-10}$ W/Hz

$\Rightarrow N_0 = 2\times10^{-10}$ W/Hz

1. Non coherent ASK

- The probability of error for noncoherent ASK is

$$P_e = \frac{1}{2} e^{(-E_b/2N_0)}$$

$$10^{-5} = \frac{1}{2} e^{-E_b/2\times2\times10^{-10}}$$

$$e^{-E_b/4\times10^{-10}} = 2\times10^{-5}$$

$$\frac{-E_b}{4\times10^{-10}} = -10.82$$

$$E_b = 4.32\times10^{-9}\ J$$

Average carrier power: $P = \frac{E_b}{T_b} = \frac{4.32\times10^{-9}}{10^{-3}} = 4.32\times10^{-6}$ watt

2. coherent PSK

Propability of error is:

$$P_e = \frac{1}{2}\,\text{Erfc}\sqrt{\frac{E_b}{N_0}}$$

$$10^{-5} = \frac{1}{2}\,\text{Erfc}\sqrt{\frac{E_b}{2\times10^{-10}}}$$

$$\text{Erfc}\sqrt{0.5\times10^{10} E_b} = 2\times10^{-5}$$

$$1 - \text{Erf}\sqrt{0.5\times10^{10} E_b} = 2\times10^{-5}$$

$$\text{Erf}\sqrt{0.5\times10^{10} E_b} = 0.99998$$

From the error function table

$$\text{Erf}(3) = 0.99998$$

$$\therefore \sqrt{0.5\times10^{10} E_b} = 3.$$

$$E_b = 1.8\times10^{-9}\ J$$

Average carrier power

$$P = \frac{E_b}{T_b} = \frac{1.8 \times 10^{-9}}{10^{-3}} = 1.8 \times 10^{-6} \text{ watt}$$

3. Non coherent DPSK

Probability of error is

$$P_e = \frac{1}{2} e^{-E_b/N_0}$$

$$10^{-5} = \frac{1}{2} e^{-E_b/2\times10^{-10}}$$

$$e^{-E_b/2\times10^{-10}} = 2\times10^{-5}$$

$$\frac{-E_b}{2\times10^{-10}} = -10.82$$

$$E_b = 2.16 \times 10^{-9} \text{ J}$$

Average carrier power

$$P = \frac{E_b}{T_b} = \frac{2.16\times10^{-9}}{10^{-3}} = 2.16 \times 10^{-6} \text{ watt}$$

4. Non coherent FSK

- Probability of error for non coherent FSK is

$$P_e = \frac{1}{2} e^{-E_b/2N_0}$$

$$10^{-5} = \frac{1}{2} e^{-E_b/2\times2\times10^{-10}}$$

$$e^{-E_b/4\times10^{-10}} = 2\times10^{-5}$$

$$\frac{-E_b}{4\times10^{-10}} = -10.82$$

$$E_b = 4.32\times10^{-9} \text{ J}$$

Average carrier power

$$P = \frac{E_b}{T_b} = \frac{4.32\times10^{-9}}{10^{-3}} = 4.32\times10^{-6} \text{ watt}$$

Keerthana Ashok

Page No.:
Date:

* Matched Filter

A matched filter is optimum receiver. It maximizes signal to noise power ratio and provide minimum probability of error. It is called noise filter since its impulse response is matched to the shape of input signal.

- Impulse Response to matched filter

The transfer function of matched filter is such that it maximizes to signal to noise power ratio which is given by:

$$SNR = \rho \leq \int_{-\infty}^{\infty} \frac{|X(f)|^2}{N_0/2}\, df \quad \text{—— (1)}$$

The SNR is maximum when

$$\theta_1(f) = k\,\theta_2^*(f) \quad \text{—— (2)}$$

* indicates complex conjugate

where $\theta_1(f) = \sqrt{\frac{N_0}{2}}\, H(f)$ —— (3)

$$\theta_2(f) = \frac{1}{\sqrt{N_0/2}}\, X(f)\, e^{2j\pi f t} \quad \text{—— (4)}$$

substituting eq (3) and eq (4) in eq (2)

$$\sqrt{\frac{N_0}{2}}\, H(f) = \frac{k}{\sqrt{N_0/2}}\, X^*(f)\, e^{2\pi j f t} \quad \text{—— (5)}$$

From eq (5)

$$H(f) = \frac{2k}{N_0}\, X^*(f)\, e^{j2\pi f t} \quad \text{—— (6)}$$

$$H(f) = \frac{2k}{N_0}\, X^*(-f)\, e^{-j2\pi f t} \quad \text{—— (7)}$$

Taking inverse Fourier Transform

$$h(t) = IFT\{H(f)\}$$

wkt $IFT\{X(-f)\} = x(-t)$

IFT of $e^{-j2\pi f t}$ represents the time shift of t seconds.

$\therefore$ IFT of $X(-f)\, e^{-2j\pi f t} = x(T-t)$

$$h(t) = \frac{2k}{N_0}\, x(T-t)$$

The impulse response of matched filter is nothing but the shifted version of input signal with a scaling factor $\frac{2k}{N_0}$

If $x(t) = x_1(t) - x_2(t)$

The impulse response of the filter

$$h(t) = \frac{2k}{N_0} \{ x_1(T-t) - x_2(T-t) \}$$

wkt matched filter is a LTI system

Therefore the output of matched filter is convolution of input $x(t)$ and the impulse response of a filter $h(t)$

i.e., the output of the filter is

$$x(t) = x(t) * h(t)$$

Q: The input to matched filter is a polar NRZ signal where binary 1 is represented by a positive rectangular and binary 0 by negative rectangular pulse. Find the impulse response of matched filter and sketch the same

— $x_1(t)$ represents positive rectangular pulse with the amplitude A and duration T seconds which is as shown in figure.

similarly $x_2(t)$ represents negative rectangular pulse with the amplitude $-A$ and duration T seconds

$$x(t) = x_1(t) - x_2(t)$$

$$= A - (-A) = 2A \quad ; \quad 0 \le t \le T$$

The time reverse version of $x(t)$ is $x(-t)$ and is given by: $x(-t) = 2A \quad ; \quad -T \le t \le 0$

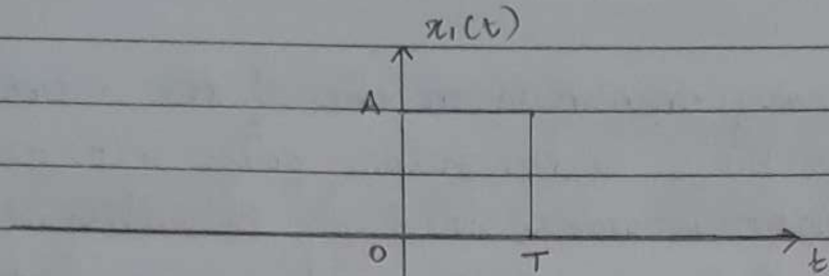
Delay of $x(-t)$ by T seconds.

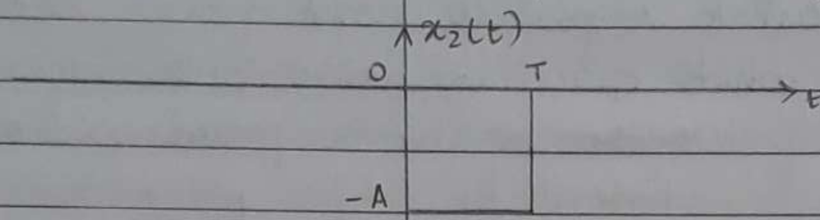
$$x(T-t) = 2A \quad ; \quad 0 \le t \le T$$

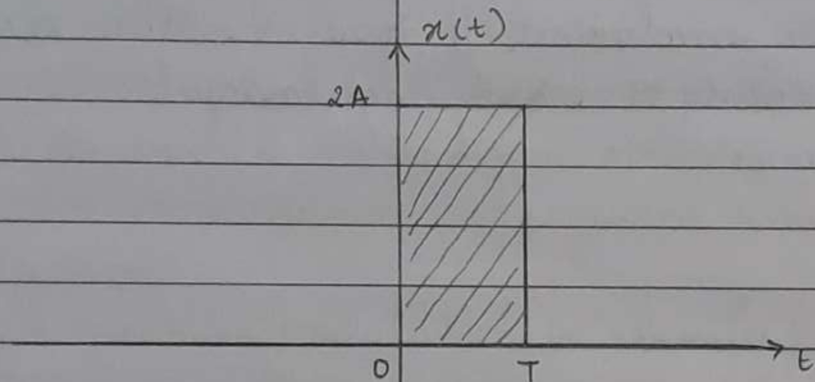
$$h(t) = \frac{2k}{N_0} x(T-t) = \frac{2k}{N_0} \times 2A = \frac{4Ak}{N_0} \quad ; \quad 0 \le t \le T$$

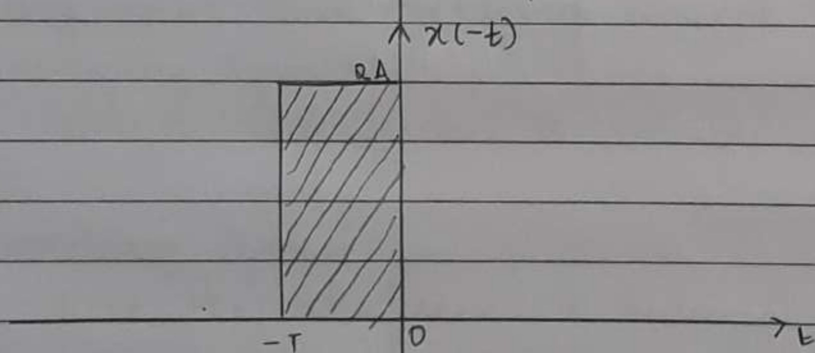
with $\frac{2k}{N_0} = 1$ impulse response of matched signal is nothing but input signal and hence the name.

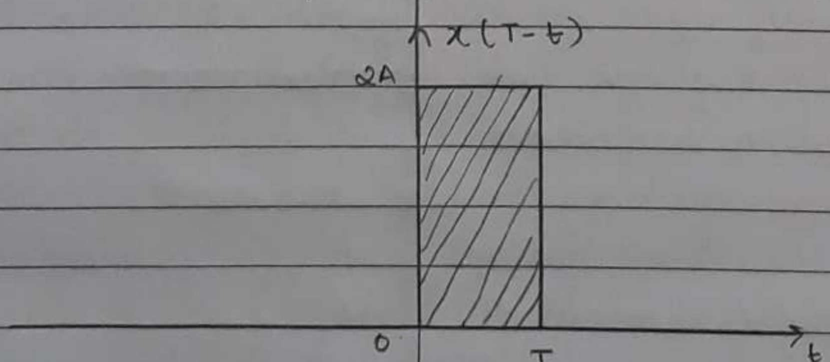
Keerthana Ashok

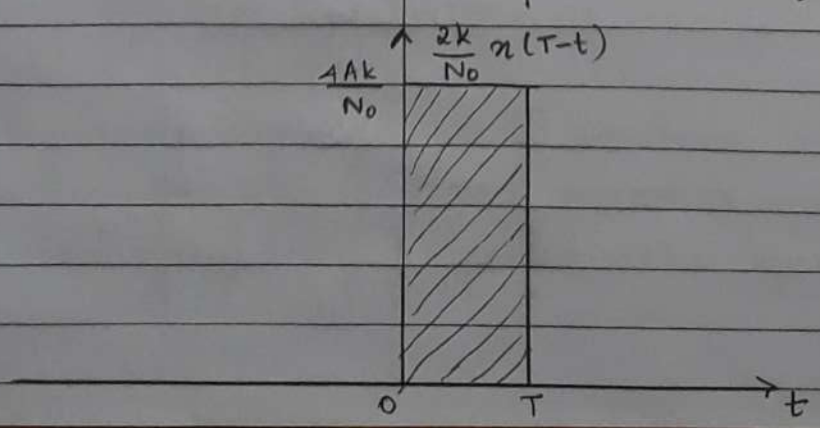












Q: In a binary transmission one of the message is represented by a rectangular pulse $x(t)$ and the other message is transmitted by the absence of the pulse. Calculate signal to noise power ratio at $t = T$. Assuming white Gaussian noise with noise PSD $N_0/2$. Find the sketch the impulse response and output of the filter.

Hint: the convolution of two similar rectangular pulse is nothing but a triangular pulse.

Keerthana Ashok

UNIT - 02

Spread Spectrum Modulation

* **Spread Spectrum Modulation:**

spread spectrum Modulation can be defined as two parts:

- spreading the signal at the transmitter and despreading the signal at the receiver and which is obtained by a special code which is independent of data sequence.
- the transmitted data signal occupies a much more bandwidth than minimum required bandwidth.

- Classification:

→ spread spectrum Modulation is classified into following types depending upon their operating concept.

- Averaging systems

Here the interference is reduced by averaging the signals.

- Avoiding systems

Here the interference is reduced by making the signal avoid the interference by large fractions of time.

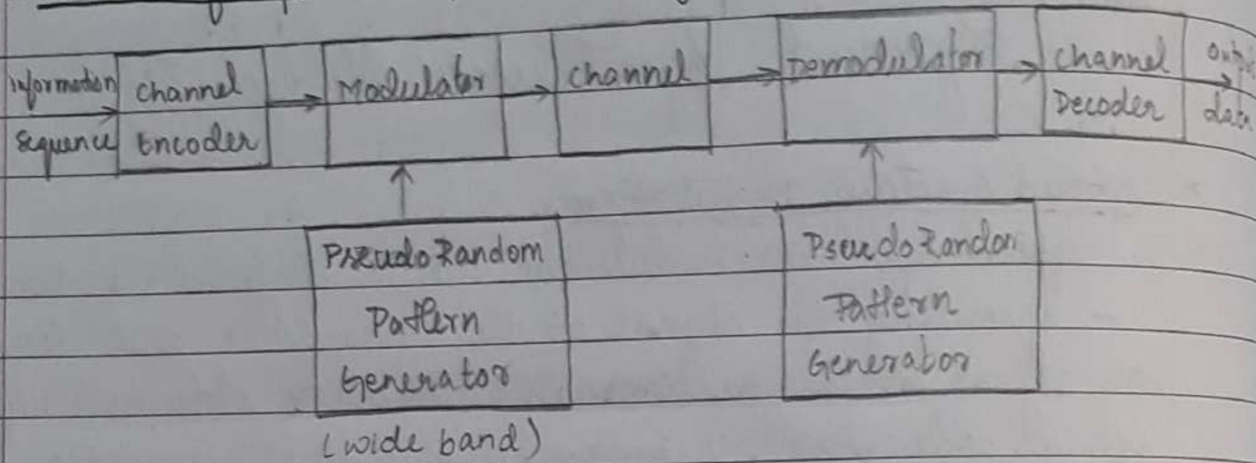
→ Based on Modulation techniques employed, the spread spectrum modulation is classified as:

- Direct sequence spread spectrum
- Frequency Hopping Spread Spectrum
- Time Hopping spread spectrum
- Chirp spread Spectrum
- Hybrid methods

Frequency Hopping Spread spectrum are classified into:

- slow frequency Hopped spread spectrum
- fast frequency hopped spread spectrum.

* Model of Spread Spectrum Digital Communication system:



Channel Encoder: It adds some redundant bits to the input bit sequence in properly defined format.

Pseudorandom pattern generator: It produces a pseudorandom or pseudonoise binary valued sequence.

Modulator and Demodulator: Pseudonoise binary valued sequence is impressed on the transmitted signal at modulator and the impressed PN signal is removed from received signal at demod.

Channel Decoder: It uses the coded bits to reconstruct error free accurate bit sequence and reduce the effects of channel noise and distortion.

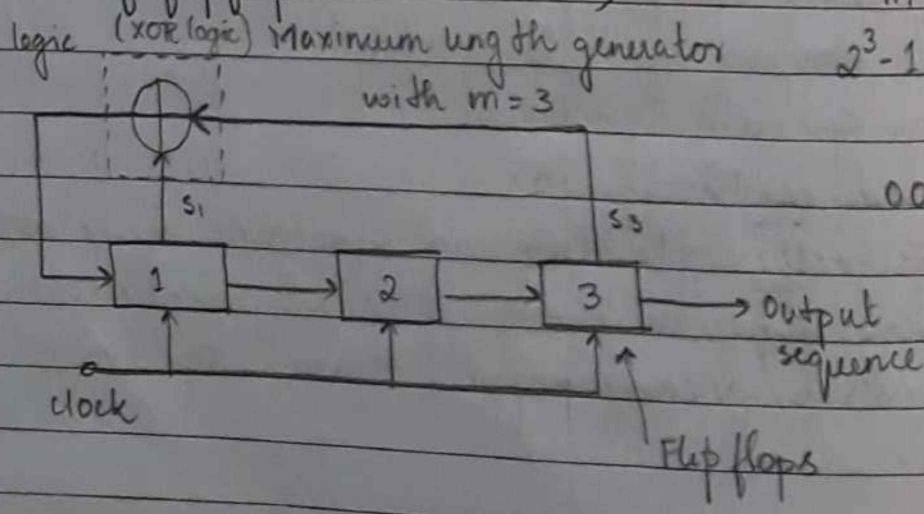
- Generation of pseudo noise sequence:

When pseudo noise sequence generated by linear feedback shift register has the length of $2^m - 1$ it is called maximum length sequence. (where m indicates the number of flipflops in the circuit)

$m = 3$

$2^3 - 1 = 7$ different combinations

$000 \rightarrow$ avoided



Here the pseudo noise sequence repeats after seven different combinations as it is periodic in nature.

States of Flip Flops			output Pseudonoise sequence (S_3)
S1	S2	S3	
1	0	0	0
1	1	0	0
1	1	1	1
0	1	1	1
1	0	1	1
0	1	0	0
0	0	1	1

– Properties of Pseudonoise sequence

- Balance Property:

In each period of a maximum length sequence, the number of 1's is always one more than the number of 0's.

of the $2^m - 1$ terms ($2^3 - 1 = 7$)

$2^{m-1} = 2^{3-1} = 4$ are 1's

$2^{m-1} - 1 = 2^{3-1} - 1 = 3$ are 0's

- Run Property:

Among the runs of 1's and 0's in each period of maximum length sequence, one half the runs of each kind are of length 1, 1/4 the runs of each kind are of length 2, 1/8 are of length 3 ....

Relative frequency of run length n (zero or ones) is

$1/2^n$ for $n \leq m-1$ and

$1/(2^m - 1)$ for $n = m$

One run length each of $r-1$ zeros and r ones occurs. There are no run lengths for $n > r$.

- Correlation Property:

The number of disagreements exceeds the number of agreements by unity.

★ Direct sequence Spread spectrum: (A notion of spread sequence)

In this method the data sequence directly modulates the pseudo noise sequence. Let the data sequence be represented by $\{b_k\}$. Using NRZ bipolar format, the data sequence can be represented as:

Keerthana Ashok

when $b_k = 1$ $\quad b(t) = +1 \; ; \; 0 \le t \le T_b$

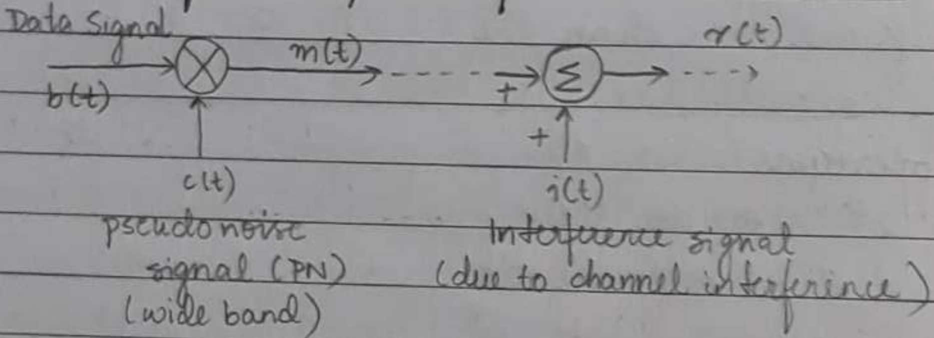
$b_k = 0$ $\quad b(t) = -1 \; ; \; 0 \le t \le T_b$

Similarly the pseudo noise sequence can be represented by $\{c_k\}$. Using NRZ bipolar format, the pseudo noise sequence can be represented as:

when $c_k = 1$ $\quad c(t) = +1 \; ; \; 0 \le t \le T_c$

$c_k = 0$ $\quad c(t) = -1 \; ; \; 0 \le t \le T_c$

- **Direct sequence spread spectrum Transmitter:**



The output of the multiplier can be written as

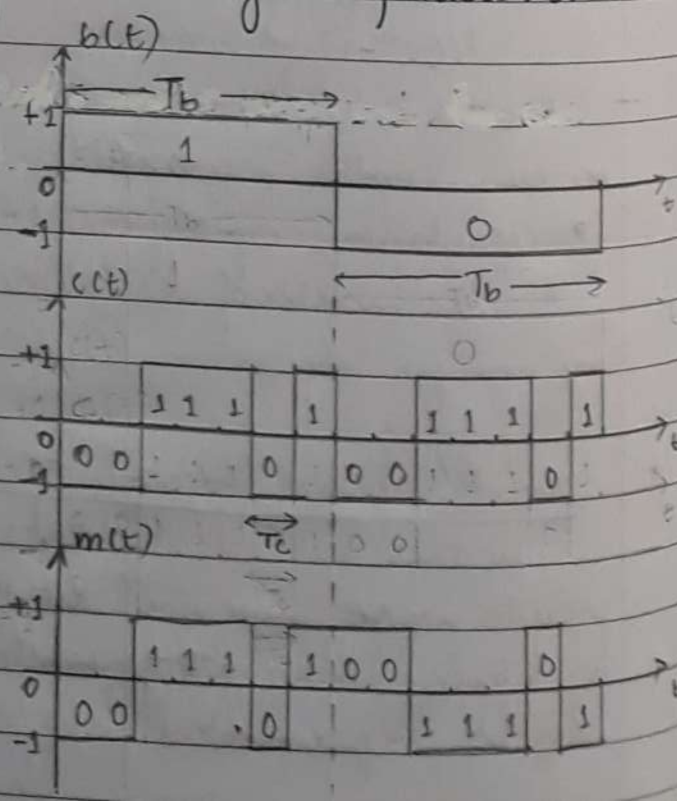
$$m(t) = c(t)\, b(t)$$

Let T_b be the one bit period of data signal and T_c be the one bit period of pseudo noise signal.

Suppose the sequence generated by the pseudo noise generator is 0011101.

There are 4 bits of pseudonoise sequence in one bit period of data signal.

The pseudonoise signal c(t) is a wide band signal. When it is multiplied with the narrow band data signal b(t) resulting in a wide band signal m(t) and has wide spectrum as large as that of c(t).



Keerthana Ashok

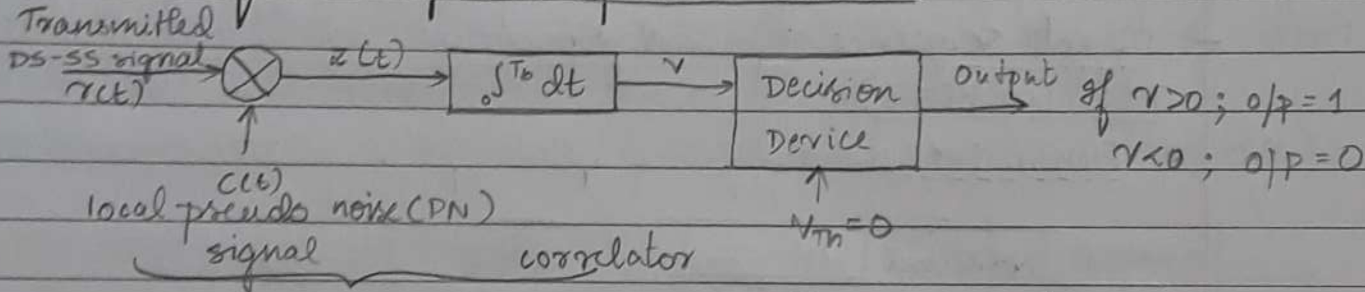
Page No.:
Date:

Thus the narrow band signal b(t) is converted into wide band message signal m(t). This signal can be transmitted directly by using base band transmission.
During the communication process, the interference signal i(t) is added and is as shown in the figure
The output transmitted sequence is given by:

$$r(t) = m(t) + i(t)$$

$$r(t) = b(t)c(t) + i(t)$$

- **Direct sequence spread spectrum Receiver:**



The receiver consists of a multiplier, integrator and decision device. The locally generated pseudo noise signal is applied to the multiplier. This signal is an exact replica of that used in the transmitter. The output of the multiplier is given by:

$$z(t) = r(t)\, c(t)$$

$$z(t) = [m(t) + i(t)]\, c(t)$$

$$z(t) = [\, b(t)c(t) + i(t)]\, c(t)$$

$$z(t) = b(t)c^2(t) + i(t)c(t)$$

wkt irrespective of the value of c(t) (either +1 or −1),

$$c^2(t) = +1 \text{ always}$$

Therefore

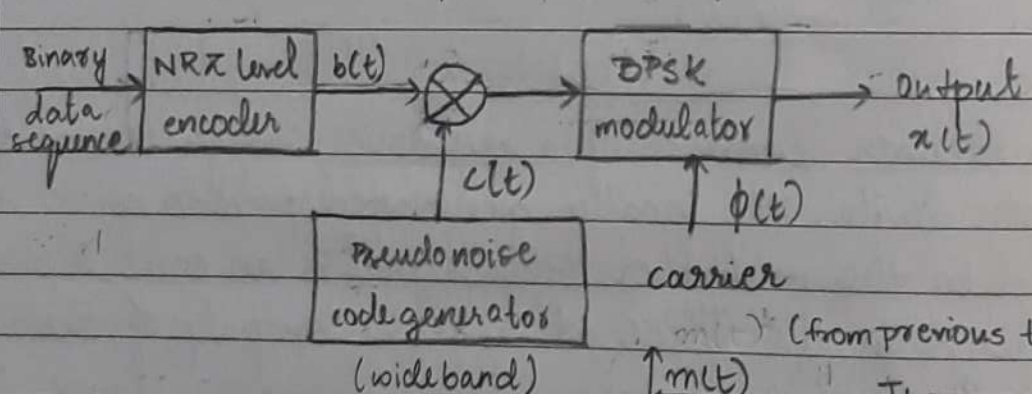
$$z(t) = b(t) + i(t)c(t)$$

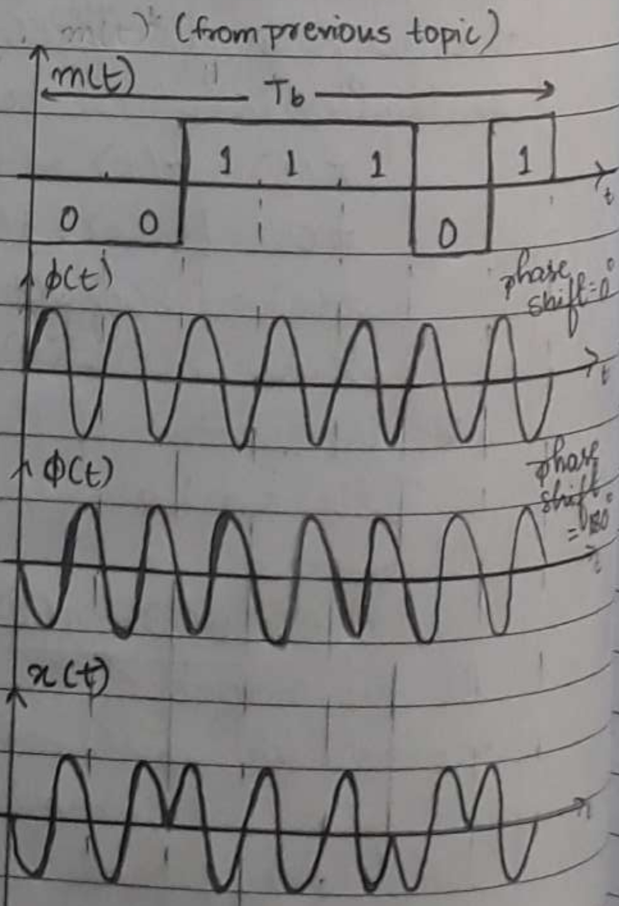
At the output of the multiplier the data signal is reproduced with noise signal i(t)c(t). We know that c(t) is a wide band signal. After multiplying with interference signal i(t), the product c(t) i(t) is also a wideband signal.

The output of the multiplier $x(t)$ is passed through the integrator which integrates over one bit period. Due to low pass action of integrator the wide band noise signal $c(t)\, i(t)$ is removed. The output of the integrator is given to the decision device as input and the decision device compares with the threshold voltage equal to zero. The output of the decision device is: 1 if $v>0$
0 if $v<0$.

- **Direct sequence spread spectrum with coherent PSK (or BPSK):**

- Transmitter (Direct sequence spread spectrum BPSK transmitter)





For binary 1 $b(t) = +1$
For binary 0 $b(t) = -1$

The NRZ level encoder converts binary data sequence into a bipolar NRZ waveform $b(t)$.

The multiplier multiplies the data signal $b(t)$ with a pseudo noise signal $c(t)$. The output of the multiplier is direct sequence spread signal $m(t)$ (wide band)

Page No.: Keerthana Ashok
Date:

This signal is given as modulating signal to BPSK modulator. The direct sequence BPSK signal is generated at the output of the modulator.

Let the carrier be represented by

$$\phi(t) = \sqrt{2P}\,\sin 2\pi f_c t$$

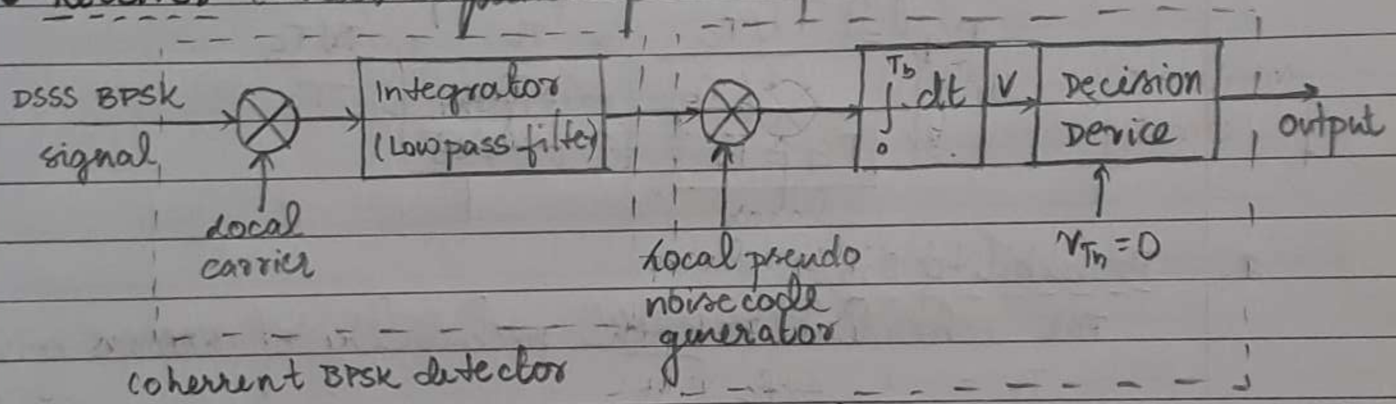
The output of the modulator i.e, transmitted signal is

$$x(t) = \sqrt{2P}\, m(t) \sin 2\pi f_c t$$

Therefore when $m(t)$ is positive, there is phase shift of $0°$ and when $m(t)$ is negative, there is phase shift of $180°$.

From the waveform we see that one bit period of $b(t)$ (i.e, T_b) will accomodate seven bit periods of $m(t)$

- **Receiver (Direct sequence Spread spectrum BPSK receiver):**



Despreading the signal

Decision device logic
output is 1 if $v > 0$
output is 0 if $v < 0$

Bandwidth of low pass filter is the bandwidth of $m(t)$.

- Performance of Direct sequence spread spectrum

1. **Processing Gain (PG):**

It is defined as the ratio of bandwidth of spread message signal to bandwidth of unspread data signal.

$$\text{Processing gain: } PG = \frac{BW_{(\text{spread signal})}}{BW_{(\text{unspread signal})}} = \frac{BW \text{ of } m(t)}{BW \text{ of } b(t)}$$

The bandwidth of unspread signal is same as bandwidth of data signal and is given by:

$$\text{BW of unspread signal (BW of data signal)} = \frac{1}{\text{one bit period}} = \frac{1}{T_b}$$

The Bandwidth of spread message signal is same as bandwidth of message signal and is given by

$$\text{BW of spread signal (BW of message signal)} = \frac{1}{\text{one bit period}} = \frac{1}{T_c}$$

(BW of $m(t)$ is same as BW of pseudonoise signal $c(t)$)

We know that, one bit period of data signal (T_b) is equal to N bit periods of pseudo noise signal or message signal $m(t)$. i.e., $T_b = NT_c$.

Therefore

Processing gain

$$PG = \frac{1/T_c}{1/T_b} = \frac{T_b}{T_c} = \frac{NT_c}{T_c}$$

$$\therefore \boxed{PG = N}$$

2. Probability Error

For coherent PSK the probability of error is given by:

$$P_e = \frac{1}{2} \operatorname{Erfc} \sqrt{\frac{E_b}{N_0}}$$

where E_b is bit energy

and $N_0/2$ is noise spectral density.

For direct sequence spread spectrum modulation, the noise spectral density is given by:

$$\frac{N_0}{2} = \frac{JT_c}{2}$$

where J is average interference power

$$\therefore N_0 = JT_c$$

For direct sequence spread spectrum with coherent PSK, the probability of error is given as.

$$\boxed{P_e = \frac{1}{2} \operatorname{Erfc} \sqrt{\frac{E_b}{JT_c}}}$$

3. Jamming Margin

Jamming margin is defined as the ratio of average powers of interference (J) and data signal (P_s).

$$\boxed{\text{Jamming margin} = \frac{J}{P_s}}$$

we know that the energy in a bit / bit energy $E_b = P_s T_b$ where T_b is the bit duration of the data signal $b(t)$.

consider the ratio of energy / bit to the noise spectral density, i.e.,

$$\frac{E_b}{N_0} = \frac{P_s T_b}{N_0}$$

For direct sequence spread spectrum, $N_0 = J T_c$

Therefore,

$$\frac{E_b}{N_0} = \frac{P_s T_b}{J T_c}$$

$$\frac{J}{P_s} = \frac{T_b / T_c}{E_b / N_0}$$

$$\boxed{\begin{aligned} \text{Jamming margin} &= \frac{\text{Processing Gain}}{E_b/N_0} \\ (\text{Jamming margin})_{dB} &= (\text{Processing Gain})_{dB} - 10\log_{10}\frac{E_b}{N_0} \end{aligned}}$$

Q: A DSSS system has following parameters:

i. Data sequence bit duration $T_b = 4.095\, ms$

ii. Pseudo noise chip duration $T_c = 1\,\mu s$

iii. $E_b/N_0 = 10$ for average probability of error less than 10^{-5}

calculate processing gain and jamming margin

– sometimes 1 bit period of pseudonoise sequence is also called as one chip.

Processing gain: $PG = \frac{BW(\text{spread signal})}{BW(\text{unspread signal})}$

Keerthana Ashok

$$PG = \frac{T_b}{T_c} = \frac{4.095 \times 10^{-3}}{1 \times 10^{-6}} = \underline{\underline{4095}}$$

since $PG = \frac{T_b}{T_c} = N$, hence the length of the bit sequence is 4095.

$$\text{Jamming Margin} = \frac{J}{P_S} = \frac{PG}{E_b/N_0}$$

$$= \frac{4095}{10} = 409.5$$

In dB

$$(\text{Jamming Margin})_{dB} = (\text{Processing gain})_{dB} - 10 \log_{10} \frac{E_b}{N_0}$$

$$= 10 \log_{10}(4095) - 10 \log_{10} 10$$

$$= 36.12 - 10$$

$$= \underline{\underline{26.12 \, dB}}$$

Jamming Margin = 409.5 shows that the information bits at the receiver output can be detected with the probability of error less than 10^{-5} even when noise interference is upto 409.5 times the received signal power

* **Frequency Hopped Spread spectrum system:**

Frequency hopping means to transmit the data bits in different frequency slots or hops. The total bandwidth of the output signal is the sum of all these frequency slots or hops. Since the frequency hopping is random (which is known only to the transmitter and recognized receiver).

Since the bandwidth is in GHz, the unwanted receivers find it difficult to receive frequency hop signal.

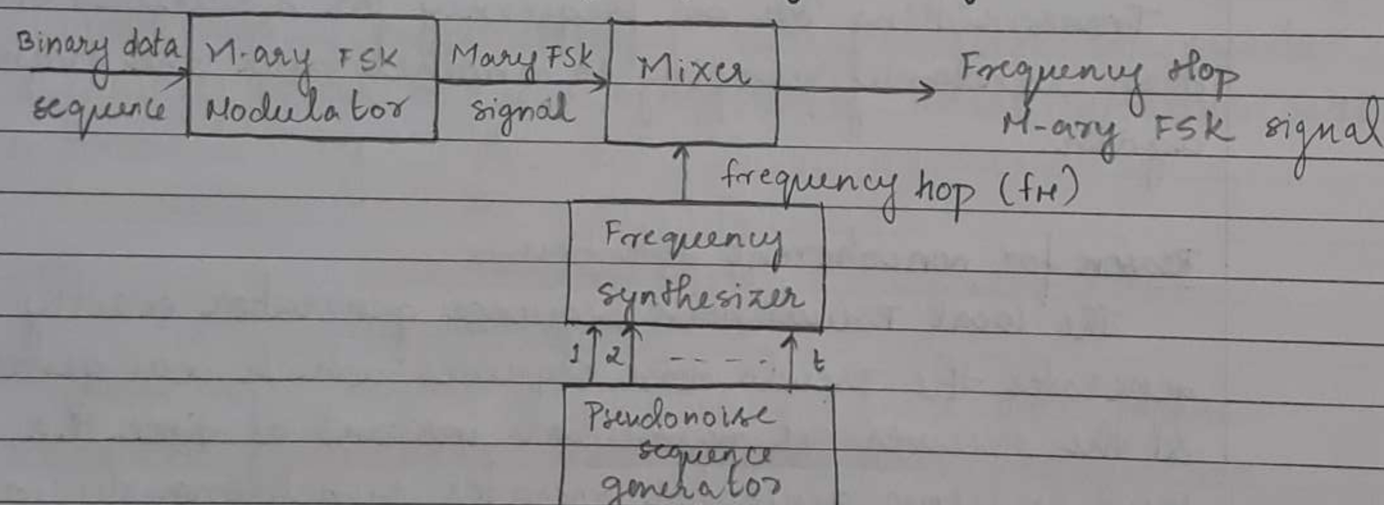
Thus in Frequency Hopped spread spectrum the carrier frequency hops randomly from one frequency to another.

Normally M-ary FSK is used along with frequency hopped spread spectrum.

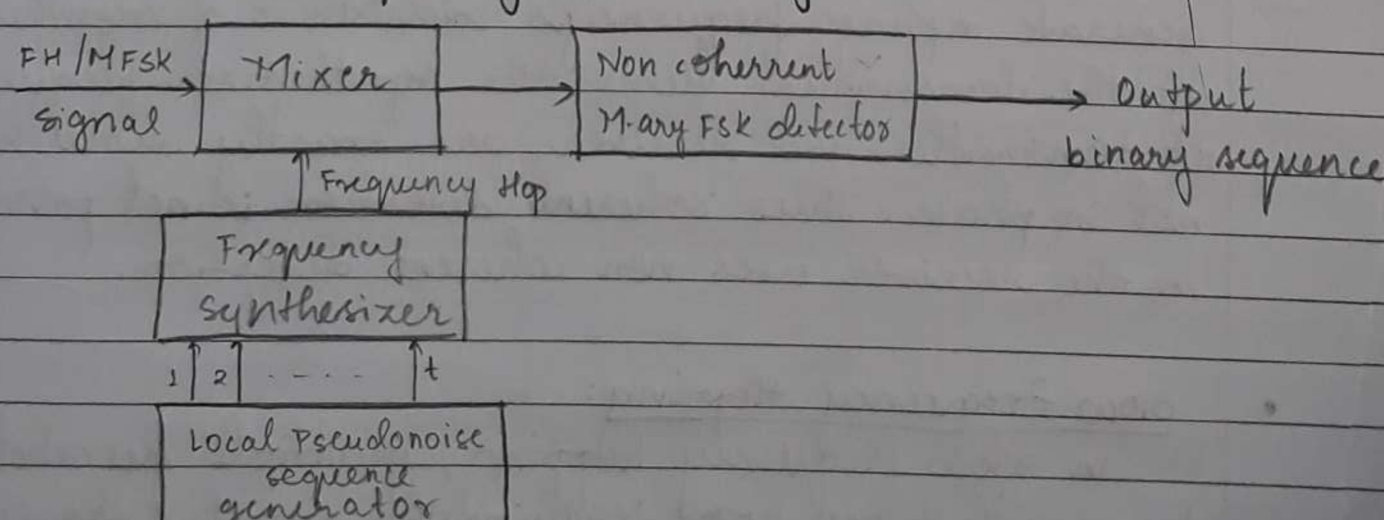
There are two types of Frequency Hopped spread spectrum:

- Slow Frequency Hopping
- Fast Frequency Hopping

- Transmitter of Frequency Hopping M-ary FSK:



- Receiver of Frequency Hop M-ary FSK:



Keerthana Ashok

The carrier frequency is determined by the output sequence from a pseudonoise generator.

Slow hopping system has a hopping rate that is lower than the information rate (symbol rate). Several information symbols are transmitted by the same carrier frequency.

Fast hopping system has a hopping rate that is higher than the information rate. One information symbol is transmitted by different carrier frequencies.

The biggest advantage of frequency hoping lies in the coexistence of several access points in the same area.

Transmitting on one frequency for a certain time, then randomly jumping to another and then transmitting again.

Reason for noncoherent detection

The local Pseudonoise sequence generator exactly generates the pseudo noise sequence which was generated at the receiver at particular instant of time. i.e, the local sequence generator operates synchronously with the transmitter but the frequency synthesizer doesnot generate inphase frequencies with that of synthesizer at the transmitter. Hence the synthesizer frequencies at transmitter and receiver are exactly same but not in phase. Thus coherent detection is not possible so the receiver uses non coherent detection.

- Slow frequency Hopping:

In slow frequency hopping, multiple symbols are transmitted per hop. Within the single hop every symbol will have independent frequency.

Let R_s is the symbol rate and R_c is the chip rate

For slow frequency hopping:

$$R_c = R_s \quad \text{and} \quad R_s = R_b / k$$

where R_b is the input bitrate and k is the number of bits per symbol.

Example:

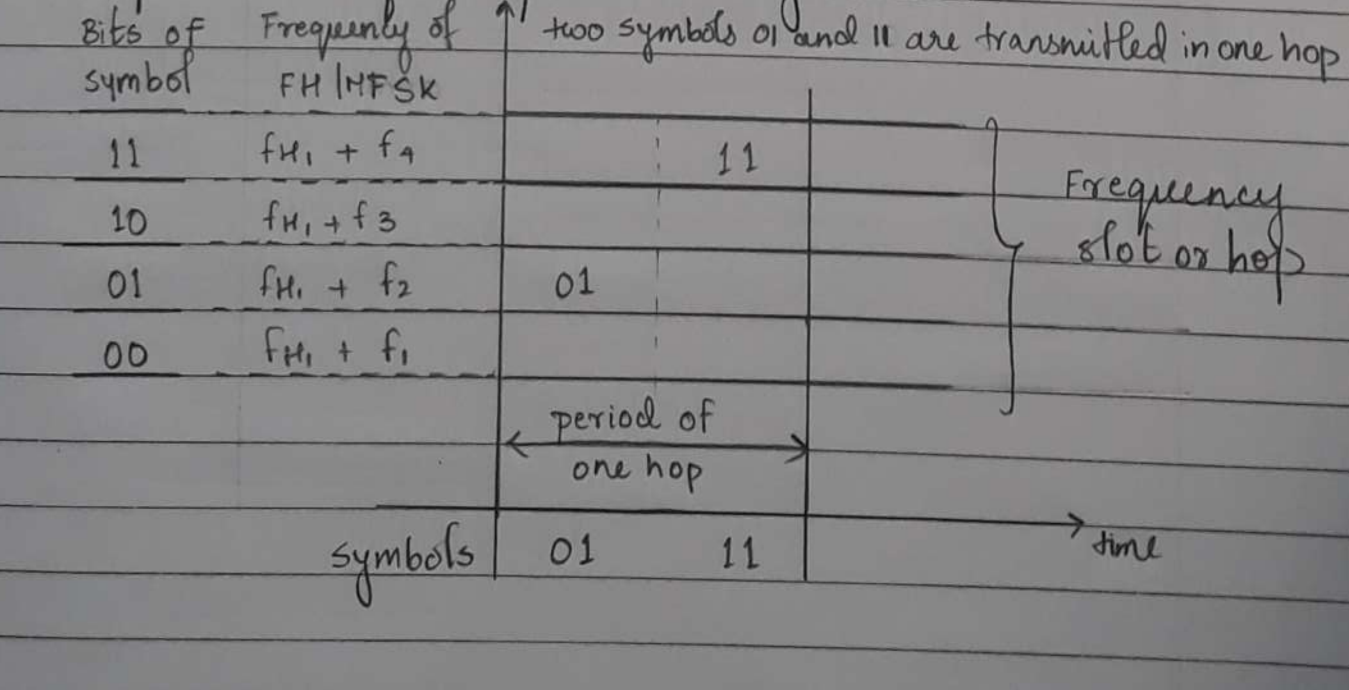
The figure shows the variation of frequency of output with respect to the input binary data symbols.

Three bits of pseudonoise sequence is used to select a hop. Therefore there are $2^3 = 8$ different hops over the complete FH bandwidth.

Two bits of input binary data represent one symbol. Therefore total $M = 2^2 = 4$ symbols. These symbols are represented by:

S_1	00	: f_1
S_2	01	: f_2
S_3	10	: f_3
S_4	11	: f_4

Thus in a single frequency hop there are four different frequencies and these four frequencies correspond to four possible symbols.



UNIT - 04

Measure of Information, Source Coding and Channel Capacity

An information source can be:
- analog information source
- digital information source

Info Source → Encoder → Transmitter → Channel → Decoder → Receiver

symbol sequence → binary sequence   binary sequence → symbol sequence

* <u>Measure of information</u>:

<u>Information content</u>: It is the various messages that are produced by the source.

Let us consider the communication system which transmits messages $m_1, m_2, m_3, \ldots,$ with probabilities of occurance $p_1, p_2, p_3, \ldots$. The amount of information transmitted through the message m_k with the probability p_k is given as:

Amount of Information : $\boxed{I_k = \log_2 \frac{1}{p_k}}$ bits.

<u>note</u>:

$I_k = \log_2 1/p_k$: in bits

$I_k = \log_{10} 1/p_k$: in Hartley / decits

$I_k = \log_e 1/p_k$: in nats

$I_k = \log_r 1/p_k$: in rary

Q1: If the binary symbol 0 and 1 are transmitted with probabilities 1/4 and 3/4 respeckively, then find the corresponding self information / amount of information

— Here binit '0' has $P_1 = 1/4$

binit '1' has $P_2 = 3/4$

The amount of information is

$$I_k = \log_2 \frac{1}{P_k}$$

Therefore

$$I_1 = \log_2 \frac{1}{P_1} = \log_2 \frac{1}{1/4} = \log_2 4 = \underline{2 \text{ bits}}$$

$$I_2 = \log_2 \frac{1}{P_2} = \log_2 \frac{1}{3/4} = \log_2 \frac{4}{3} = \underline{0.415 \text{ bits}}$$

NOTE: If probability of occurance is less information carried is more and vice versa

Q2: The binary source has source alphabets s_1 and s_2 with the probabilities 2/5 and 1/2 respectively calculate self information / amount of information

— For source alphabet S_1

$P_1 = 2/5$

$$\therefore I_1 = \log_2 \frac{1}{P_1} = \log_2 \frac{5}{2} = \underline{1.322 \text{ bits}}$$

For source alphabet S_2

$P_2 = 1/2$

$$\therefore I_2 = \log_2 \frac{1}{P_2} = \log_2 \frac{1}{1/2} = \log_2 2 = \underline{1 \text{ bit}}$$

— If I_1 is the information carried by message m_1 and I_2 is the information carried by message m_2, then the amount of Information carried compositely due to m_1 and m_2 is $I = I_1 + I_2$

Proof: consider two messages m_1 and m_2 and the information carried by them as I_1 and I_2 respectively.

Total amount of information.

$$I_k = \log_2 \frac{1}{P_k}$$

Individual amounts carried

$$I_1 = \log_2 \frac{1}{P_1} \quad \text{and} \quad I_2 = \log_2 \frac{1}{P_2}$$

where P_1 and P_2 are probabilities of message m_1 and m_2 respectively.

Therefore the information carried is

$$I = \log_2 \left[\frac{1}{P_1 P_2}\right]$$

$$I = \log_2 \left[\left(\frac{1}{P_1}\right)\left(\frac{1}{P_2}\right)\right]$$

$$I = \log_2 \frac{1}{P_1} + \log_2 \frac{1}{P_2}$$

$$I = I_1 + I_2$$

* **Entropy**: (Average Information)

It is defined as the ratio of total information to the number of messages.

$$\text{Entropy} = \frac{\text{Total information}}{\text{Number of messages}}$$

consider m number of messages such as $m_1, m_2, \dots, m_M$. with probabilities of occurance as $p_1, p_2, \dots p_m$. Let a sequence of L messages is transmitted.

$p_1 L$ messages of m_1 are transmitted
$p_2 L$ messages of m_2 are transmitted
⋮
$p_m L$ messages of m_m are transmitted.

Hence the information due to message m_1 will be

$$I_1 = \log_2 \frac{1}{P_1}$$

Since there are $P_1 L$ number of messages of m_1, the total information due to all messages of m_1 will be,

$$I_{1(total)} = P_1 L \log_2 \frac{1}{P_1}$$

Similarly: $I_{2(total)} = P_2 L \log_2 \frac{1}{P_2}$ and so on.

Thus the total information carried due to the sequence of L messages will be,

$$I_{(total)} = I_{1(total)} + I_{2(total)} + \dots + I_{m(total)}$$

$$\therefore I_{(total)} = P_1 L \log_2 \frac{1}{P_1} + P_2 L \log_2 \frac{1}{P_2} + \dots + P_m L \log_2 \frac{1}{P_m}$$

$$\text{Entropy} = \frac{\text{Total information}}{\text{number of messages}} = \frac{I_{(total)}}{L}$$

$$\therefore H = P_1 \log_2 \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2} + \dots + P_m \log_2 \frac{1}{P_m}$$

$$\boxed{H = \sum_{k=1}^{m} P_k \log_2 \frac{1}{P_k}}$$ Entropy (bits / message symbol)

Q3: Find the entropy with respect to source alphabet $S = \{s_1, s_2\}$ with the probabilities of $P = \{\frac{1}{256}, \frac{255}{256}\}$.

For S_1 : $P_1 = \frac{1}{256}$

For S_2 : $P_2 = \frac{255}{256}$

Entropy $H = \sum_{k=1}^{m} P_k \log_2 \frac{1}{P_k}$

$$H = \sum_{k=1}^{2} P_k \log_2 \frac{1}{P_k} = P_1 \log_2 \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2}$$

$$H = \frac{1}{256} \log_2 256 + \frac{255}{256} \log_2 \frac{256}{255}$$

$H = 0.03125 + 0.00562$

$H = 0.03687$ bits / message symbol

Q.4: consider the sources $S = \{s_1, s_2, s_3\}$ with probabilities $P = \{1/2, 1/4, 1/4\}$, then find:

a: self information of each message

b: entropy of source S.

— self information

For source s_1 : $P_1 = 1/2$

$$I_1 = \log_2 \frac{1}{P_1} = \log_2 \frac{1}{1/2} = \log_2 2 = 1 \text{ bit}$$

For source s_2 : $P_2 = 1/4$

$$I_2 = \log_2 \frac{1}{P_2} = \log_2 \frac{1}{1/4} = \log_2 4 = 2 \text{ bits}$$

For source S_3 : $P_3 = 1/4$

$$I_3 = \log_2 \frac{1}{P_3} = \log_2 \frac{1}{1/4} = \log_2 4 = 2 \text{ bits}$$

Entropy of source S

$$H(s) = \sum_{k=1}^{m} P_k \log_2 \frac{1}{P_k}$$

$$H(s) = \sum_{k=1}^{3} P_k \log_2 \frac{1}{P_k}$$

$$H(s) = P_1 \log_2 \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2} + P_3 \log_2 \frac{1}{P_3}$$

$$H(s) = \frac{1}{2} \log_2 \frac{1}{1/2} + \frac{1}{4} \log_2 \frac{1}{1/4} + \frac{1}{4} \log_2 \frac{1}{1/4}$$

$$H(s) = \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{4} \log_2 4$$

$$H(s) = \frac{1}{2} + \frac{2}{4} + \frac{2}{4} = 1.5 \text{ bits / message symbol}$$

— Memoryless source / Zero Memory Source

If the symbol emitted by the source during successive signalling intervals are statically independent it is a memoryless source.

Q5: Let the source S_1 and S_2 have probabilities $P = \{1/256, 255/256\}$ and source S_3 and S_4 has $P = \{7/16, 9/16\}$ and source S_5 and S_6 has $P = \{1/2, 1/2\}$. then find the entropy of all sources and compare all the values with each other.

— For sources S_1 and S_2

$$P_1 = 1/256 \quad P_2 = 255/256$$

$$H = \sum_{k=1}^{2} P_K \log_2 \frac{1}{P_K}$$

$$H = P_1 \log_2 \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2}$$

$$H = \frac{1}{256} \log_2 256 + \frac{255}{256} \log_2 \frac{256}{255}$$

$$H = 0.03125 + 0.00562$$

$$\underline{H = 0.03687} \text{ bits/message symbol}$$

For sources S_3 and S_4

$$P_3 = 7/16 \text{ and } P_4 = 9/16$$

$$H = \sum_{k=3}^{4} P_K \log_2 \frac{1}{P_K}$$

$$H = P_3 \cancel{\log} \log_2 \frac{1}{P_3} + P_4 \log_2 \frac{1}{P_4}$$

$$H = \frac{7}{16} \log_2 \frac{16}{7} + \frac{9}{16} \log_2 \frac{16}{9}$$

$$H = 0.5217 + 0.4669$$

$$\underline{H = 0.9886} \text{ bits/message symbol}$$

For sources S_5 and S_6

$P_5 = 1/2$ and $P_6 = 1/2$

$$H = \sum_{k=5}^{6} P_k \log_2 \frac{1}{P_k}$$

$$H = P_5 \log_2 \frac{1}{P_5} + P_6 \log_2 \frac{1}{P_6}$$

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

$$H = 1/2 + 1/2$$

$$\underline{\underline{H = 1}} \text{ bit / symbol}$$

* **Uncertainty / surprise**

conditions of occurrence of events:

- If an event has not occurred, there is a condition of uncertainty.
- If an event has just occurred, there is a condition of surprise.
- If an event has occurred, a time back, there is a condition of having some information.

* **Information Rate:**

The information rate is given as:

$$R = r_s H(S)$$

where R: information rate

r_s: rate at which messages are generated

H(S): Entropy / average information.

The information rate is the average number of bits of information per second.

$$R = \left[r \text{ in } \frac{\text{message}}{\text{second}} \right] \left[H \text{ in } \frac{\text{Information bits}}{\text{message}} \right]$$

Keerthana Ashok

Q6: A discrete source emits one of 6 symbols every millisecond with probabilities of 1/2, 1/4, 1/8, 1/16, 1/32 and 1/32 respectively. Find source entropy and information rate.

Given: $r_s = 1$ message/m sec $= 10^3$ msg/sec

P_1	P_2	P_3	P_4	P_5	P_6
1/2	1/4	1/8	1/16	1/32	1/32

Source Entropy

$$H = \sum_{k=1}^{6} P_K \log_2 \frac{1}{P_K}$$

$$H = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{8}\log_2 8 + \frac{1}{16}\log_2 16 + \frac{1}{32}\log_2 32 + \frac{1}{32}\log_2 32$$

$$H = \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{1}{4} + \frac{5}{32} + \frac{5}{32}$$

$H = 1.9375$ bits/msg symbol

Therefore information rate is

$R = r_s H(s)$

$R = 10^3 (1.9375)$

$R = 1937.5$ bits/symbol

* **Properties of information:**

If there is more uncertainty about message information then message information carried is also more. If receiver knows the message being transmitted the amount of information carried is zero. If I_1 is information carried by message 1 and I_2 is the information carried by message 2, then the amount of information carried combining together is $I_1 + I_2$.

If there are $m = 2^n$ equally likely messages then amount of information carried by each message will be n bits.

Keerthana Ashok

* Relationship between Hartleys, nats and bits:

$$1 \text{ Hartley} = \log_{10} \frac{1}{P} = 1$$

$$1 \text{ Nats} = \log_e \frac{1}{P} = 1$$

$$1 \text{ Bit} = \log_2 \frac{1}{P} = 1$$

$$1 \text{ Hartley} = \frac{1}{\log_{10}(1/P)} = \frac{\log_e 1/P}{\log_{10} 1/P} = \frac{-\log_e P}{-\log_{10} P}$$

$$= \frac{\log_P 10}{\log_P e} \quad \left(\text{because } \log_a b = \frac{1}{\log_b a}\right)$$

$$= \log_e 10 \quad \left(\text{because } \log_a b = \frac{\log b}{\log a}\right)$$

Therefore

$$\boxed{1 \text{ Hartley} = \log_e 10 \text{ nats} = 2.3 \text{ nats}}$$

similarly

$$\boxed{1 \text{ nat} = 0.434 \text{ Hartley}}$$

$$1 \text{ Hartley} = \frac{1}{\log_{10}(1/P)} = \frac{\log_2 1/P}{\log_{10} 1/P} = \frac{-\log_2 P}{-\log_{10} P}$$

$$= \frac{\log_P 10}{\log_P 2} = \log_2 10$$

Therefore

$$\boxed{1 \text{ Hartley} = \log_2 10 \text{ bits} = 3.32 \text{ bits}}$$

similarly

$$\boxed{1 \text{ bit} = 0.3 \text{ Hartley}}$$

$$1 \text{ nat} = \frac{1}{\log_e 1/P} = \frac{\log_2 1/P}{\log_e 1/P} = \frac{-\log_2 P}{-\log_e P}$$

$$= \frac{\log_P e}{\log_P 2} = \log_2 e$$

Therefore

$$\boxed{1\,nat = \log_2 e \text{ bits} = 1.44 \text{ bits}}$$

Similarly

$$\boxed{1\,bit = 0.69\,nat}$$

Q7: Calculate entropy when $P_k = 1$ and $P_k = 0$.

— Entropy

$$H(s) = \sum_{k=1}^{n} P_k \log_2 \frac{1}{P_k}$$

For $P_k = 0$: $H(s) = 0$ bits / msg symbol

For $P_k = 1$: $H(s) = 0$ bits / msg symbol.

Therefore no uncertainty and no amount of surprise.

Q8: If x represents the outcome of a single roll of a fair die, find the entropy.

— For a die $P_k = 1/6$

Therefore, Entropy is

$$H(s) = P_k \log_2 \frac{1}{P_k} = \frac{1}{6} \log_2 6 = \underline{\underline{0.43 \text{ bits / msg symbol}}}$$

Q9: Find the entropy of source in nats if the source emits 1 out of 4 symbols (a, b, c, d) in statistically independent sequence with probabilities 1/2, 1/4, 1/8 and 1/8 respectively.

— Given: P_1 P_2 P_3 P_4

$1/2$ $1/4$ $1/8$ $1/8$

Entropy of source

$$H(s) = \sum_{k=1}^{4} P_k \log_2 \frac{1}{P_k}$$

$$H(s) = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{8}\log_2 8 + \frac{1}{8}\log_2 8$$

$$H(s) = \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8}$$

$\therefore H(s) = 1.75$ bits / msg symbol.

wkt 1 nat = 1.44 bits

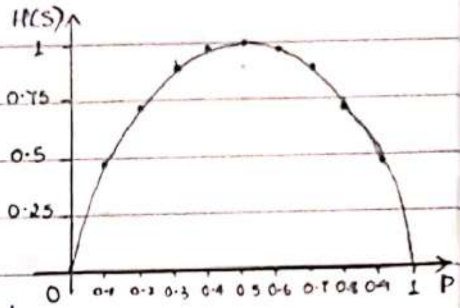
Therefore

$$H(s) = \frac{1.75}{1.44} = \underline{1.215} \text{ nats / msg symbol}$$

Q10 The binary source is emitting an independent sequence of 0's and 1's with probabilities P and 1-P respectively. Plot the entropy of source versus P.

— Source Entropy

$$H(s) = \sum_{k=1}^{2} P_k \log_2 \frac{1}{P_k}$$



For 0 : $P_1 = P$

For 1 : $P_2 = 1-P$

$$\therefore H(S) = P \log_2 \frac{1}{P} + (1-P) \log_2 \frac{1}{1-P}$$

P	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
H(S)	0	0.46	0.72	0.88	0.97	1	0.97	0.88	0.72	0.46

* Properties of Entropy:

The entropy function H(s) for the source alphabets $s = s_1, s_2, \ldots s_q$ with probabilities $P = P_1, P_2, \ldots P_q$, (where q is the number of source symbols) is given as:

$$H(s) = \sum_{i=1}^{q} P_i \log \frac{1}{P_i}$$

The entropy is continuous for every independent variable P_k in the interval (0, 1) i.e., if P_k varies continuously from 0 to 1 so does the entropy function. (entropy function vanishes at both $P_k = 0$ and $P_k = 1$).

The maximum value of entropy can be calculated by using $\boxed{H(s)_{max} = \log_2 q.}$ bits / message symbol. when all symbols become equally probable.

Keerthana Ashok

Date
Page

The source efficiency can be calculated using

$$\boxed{\eta_s = \frac{H(S)}{H(S)_{max}}}$$

Redundancy:

$$\boxed{R(S) = 1 - \eta_s}$$

Q11: For the sources $S = S_1, S_2, S_3$ and S_4 with probabilities $1/2, 1/4, 1/8, 1/8$ respectively. Find the entropy, efficiency and redundancy of the source.

Given: $P_1 \quad P_2 \quad P_3 \quad P_4 \qquad q = 4$

$1/2 \quad 1/4 \quad 1/8 \quad 1/8$

Source Entropy

$$H(S) = \sum_{k=1}^{4} P_k \log_2 \frac{1}{P_k}$$

$$H(S) = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{8}\log_2 8 + \frac{1}{8}\log_2 8$$

$$H(S) = 1/2 + 1/2 + 3/8 + 3/8 = \underline{1.75 \text{ bits/msg symbol}}$$

For $H(S)_{max}$

$$H(S)_{max} = \log_2 q = \log_2 4 = \underline{2} \text{ bits/msg symbol}$$

Source Efficiency

$$\eta_s = \frac{H(S)}{H(S)_{max}} = \frac{1.75}{2} = \underline{87.5\%}$$

Redundancy

$$R(S) = 1 - \eta_s = \underline{12.5\%}$$

Q12: A black and white TV picture consists of 525 lines of picture information. Assume that each line consists of 525 picture elements (pixels) in that each element has 256 brightness levels. The pictures are repeated at the rate of 30 frames/sec. calculate the average rate of information conveyed by the TV to the viewer

- Total number of pixels = 525

Number of frames = 525 x 525 = 275625 frames.

Total brightness levels = 256

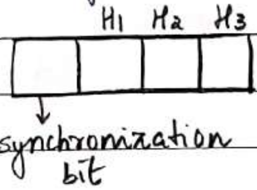
Number of source symbols

$$q = (256)^{275625}$$

Source Entropy

$$H(S) = \log_2 q = \log_2 (256)^{275625} = \underline{\underline{2205000}} \text{ bits/msg symbol}$$

Q13: A certain data source has 8 symbols that are produced in the blocks of codes at a rate of 500 blocks/sec. If the first symbol in each block is always same assume it as synchronization, the remaining three are filled by any of the 8 symbols with equal probability. What is the entropy of the source?



- Given: $q = 8$

Total entropy of the source is

$H = H_1 + H_2 + H_3$

But $H_1 = H_2 = H_3$ as the symbols are equiprobable.

Hence

$$H = 3\,H(S)_{max} = 3 \log_2 q = 3 \log_2 8 = 3 \times 3$$

$$\underline{\underline{H = 9 \text{ bits/msg symbol}}}$$

Q14: A code is composed of dots and dashes. Assuming dash is three times dot and has $1/3$rd probability of occurrence of dot. Calculate

a. information in dot and dash

b. entropy of dot and dash code

c. the average rate of information if a dot lasts for 10m sec and the time allocated between symbols is also 10 m sec.

- a. information in dot and dash

$$P_{dot} + P_{dash} = 1$$

but given $P_{dash} = \frac{1}{3} P_{dot}$

$$\therefore P_{dot} + \frac{P_{dot}}{3} = 1$$

$$P_{dot} = 3/4 \rightarrow P_{dash} = 1/4$$

Information in dot

$$I_{dot} = \log_2 \frac{1}{P_{dot}} = \log_2 \frac{4}{3} = 0.415 \text{ bits}$$

Information in dash

$$I_{dash} = \log_2 \frac{1}{P_{dash}} = \log_2 4 = 2 \text{ bits}$$

b. entropy of dot and dash code

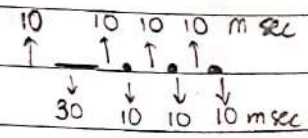
$$H(S) = P_{dot} \log_2 \frac{1}{P_{dot}} + P_{dash} \log_2 \frac{1}{P_{dash}}$$

$$H(S) = \frac{3}{4} \log_2 \frac{4}{3} + \frac{1}{4} \log_2 4$$

$$H(S) = 0.311 + 0.5$$

$$H(S) = 0.811 \text{ bits/message symbol}$$

c. Rate of information



total = 100m sec

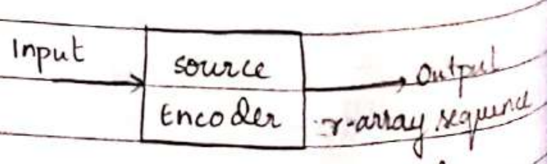
$$R = r_s H(S)$$

$$R = \frac{4}{100m} \; 0.811$$

$$R = 32.44 \text{ bits/sec}$$

* Source Encoding:

The process of efficiently converting the output of analog or digital source into a sequence of binary digits is known as source encoding.



sources: $S = \{S_1, S_2 \cdots S_q\}$

where q is number of sources

* code efficiency and redundancy:

- The average length of the code is given by

$$\boxed{L = \sum_{i=1}^{q} P_i l_i} \text{ binits/ msg symbol}$$

where

$P_i : \{P_1, P_2 \ldots P_q\}$: respective probabilities of the source symbols $s_1, s_2, \ldots s_q$

$l_i : \{l_1, l_2, \ldots l_q\}$: the respective word lengths in binits of the code word $s_1, s_2 \ldots s_q$.

- The entropy is given by

$$\boxed{H(s) = \sum_{i=1}^{q} P_i \log \frac{1}{P_i}} \text{ bits/ msg symbol}$$

- we have a condition that

$L \geq H(s)$: Binary code

$L \geq H_r(s)$: r-array code

where $H_r(s)$ is the entropy in r-array

$$\boxed{H_r(s) = \frac{H(s)}{\log_2 r}} \quad \text{where } H(s) \text{ is in bits/msg symbol}$$

r is the number of different symbols used in the code. alphabets. The code efficiency is calculated using

$$\boxed{\eta = \frac{H(s)}{L}} \text{ : Binary code}$$

$$\boxed{\eta = \frac{H_r(s)}{L}} \text{ : r-array code}$$

- Redundancy is given by

$$\boxed{R(s) = 1 - \eta}$$

Q15: A source having alphabets s_1, s_2, s_3, s_4 and s_5 with probabilities of 1/2, 1/6, 1/6, 1/9 and 1/18 respectively. They have word lengths of $l_1 = 1, l_2 = 2, l_3 = 3, l_4 = 4$ and $l_5 = 5$ binits. calculate the code efficiency and redundancy.

- Given:

	P_1	P_2	P_3	P_4	P_5
	1/2	1/6	1/6	1/9	1/18
l_i:	1	2	3	4	4

Average length of code is

$$L = \sum_{i=1}^{5} P_i l_i$$

$$L = \frac{1}{2}(1) + \frac{1}{6}(2) + \frac{1}{6}(3) + \frac{1}{9}(4) + \frac{1}{18}(4)$$

$$L = 0.5 + 0.33 + 0.5 + 0.44 + 0.22$$

$$L = 2 \text{ binits / msg symbol}$$

Source entropy

$$H(S) = \sum_{i=1}^{5} P_i \log_2 \frac{1}{P_i}$$

$$H(S) = \frac{1}{2}\log_2 2 + \frac{1}{6}\log_2 6 + \frac{1}{6}\log_2 6 + \frac{1}{9}\log_2 9 + \frac{1}{18}\log_2 18$$

$$H(S) = 0.5 + 0.43 + 0.43 + 0.35 + 0.23$$

$$H(S) = 1.94 \text{ bits / msg symbol}$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{1.94}{2} = 97.1\%$$

Redundancy

$$R = 1 - \eta = 2.9\%$$

★ Shannon's encoding algorithm: (binary)

Step 1: List the source symbol in the order of nonincreasing probability.

Step 2: Compute the sequence

$$\alpha_1 = 0$$

$$\alpha_2 = P_1 + \alpha_1$$

$$\alpha_3 = P_2 + \alpha_2$$

$$\alpha_4 = P_3 + \alpha_3$$

$$\vdots$$

$$\alpha_{q+1} = P_q + \alpha_q$$

Step 3: Determine the smallest integer value of l_i to satisfy, $2^{l_i} \geq \frac{1}{P_i}$ for all $i = 1$ to q.

Step 4: Expand the decimal number α_i in binary upto l_i places neglecting places beyond l_i.

Step 5: Remove the binary point to get the desired code.

Q16: Apply shannon's encoding algorithm for the following messages s_1, s_2, s_3 with the probabilities of 0.5, 0.3, 0.2 respectively. Calculate the efficiency and redundancy.

— Step 1: non increasing order of probabilities.

S_1	S_2	S_3
P_1	P_2	P_3
0.5	0.3	0.2

Step 2: $\alpha_1 = 0$

$\alpha_2 = P_1 + \alpha_1 = 0.5 + 0 = 0.5$

$\alpha_3 = P_2 + \alpha_2 = 0.3 + 0.5 = 0.8$

$\alpha_4 = P_3 + \alpha_3 = 0.2 + 0.8 = 1$

Step 3: $2^{l_i} \geq 1/P_i$

For $i = 1$

$l_1 = 1$ $\quad 2^1 \geq 1/0.5 = 2$ ✓ $\quad \therefore l_1 = 1$

For $i = 2$

$l_2 = 1$ $\quad 2^1 \geq 1/0.3 = 3.33$ ✗

$l_2 = 2$ $\quad 2^2 = 4 \geq 1/0.3 = 3.33$ ✓ $\quad \therefore l_2 = 2$

For $i = 3$

$l_3 = 2$ $\quad 2^2 = 4 \geq 1/0.2 = 5$ ✗

$l_3 = 3$ $\quad 2^3 = 8 \geq 1/0.2 = 5$ ✓ $\quad \therefore l_3 = 3$

Step 4:

$\rightarrow \alpha_1 = 0$

$\rightarrow \alpha_2 = (0.5)_{10} = 0.5 \times 2 = 1 \longrightarrow 1$ $\quad \alpha_2 = (0.1)_2$

$\rightarrow \alpha_3 = (0.8)_{10} = 0.8 \times 2 = 1.6 \longrightarrow 1$

$0.6 \times 2 = 1.2 \longrightarrow 1$ $\quad \alpha_3 = (0.11)_2$

$\rightarrow \alpha_4 = (1)_{10} = (1.00)_2$

Step 5:

$\alpha_1 = 0$

$\alpha_2 = 1$

$\alpha_3 = 11$

$\alpha_4 = 100$

Average length

$$L = \sum_{i=1}^{3} P_i l_i$$

$$L = 0.5(1) + (0.3)(2) + (0.2)3$$

$$L = 0.5 + 0.6 + 0.6$$

$L = 1.7$ binits / msg symbol

Entropy

$$H(S) = \sum_{i=1}^{3} P_i \log_2 \frac{1}{P_i}$$

$$H(S) = 0.5 \log_2 \frac{1}{0.5} + 0.3 \log_2 \frac{1}{0.3} + 0.2 \log_2 \frac{1}{0.2}$$

$$H(S) = 0.5 + 0.52 + 0.46$$

$H(S) = 1.48$ bits / message symbol

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{1.48}{1.7} = 87.06\%$$

Redundancy

$$R(S) = 1 - \eta = 12.94\%$$

Q17: Construct binary code for the following source using Shannon's binary encoding procedure. Calculate code efficiency and redundancy for $S = \{S_1, S_2, S_3, S_4, S_5\}$ with probabilities $P = \{0.4, 0.25, 0.15, 0.12, 0.08\}$ respectively

S_1	S_2	S_3	S_4	S_5
0.4	0.25	0.15	0.12	0.08
P_1	P_2	P_3	P_4	P_5

Keerthana Ashok

$\alpha_1 = 0$

$\alpha_2 = P_1 + \alpha_1 = 0.4 + 0 = 0.4$

$\alpha_3 = P_2 + \alpha_2 = 0.25 + 0.4 = 0.65$

$\alpha_4 = P_3 + \alpha_3 = 0.15 + 0.65 = 0.8$

$\alpha_5 = P_4 + \alpha_4 = 0.12 + 0.8 = 0.92$

$\alpha_6 = P_5 + \alpha_5 = 0.08 + 0.92 = 1$

For the condition.

$2^{l_i} \geq 1/P_i \quad i = 1 \text{ to } 5$

For $i=1$

$l_1 = 1 \quad 2^1 \geq 1/0.4 = 2.5$ ×

$l_1 = 2 \quad 2^2 = 4 \geq 1/0.4 = 2.5$ ✓ $\quad \therefore l_1 = 2$

For $i=2$

$l_2 = 2 \quad 2^2 = 4 \geq 1/0.25 = 4$ ✓ $\quad \therefore l_2 = 2$

For $i=3$

$l_3 = 3 \quad 2^3 = 8 \geq 1/0.15 = 6.67$ ✓ $\quad \therefore l_3 = 3$

For $i=4$

$l_4 = 3 \quad 2^3 = 8 \geq 1/0.12 = 8.33$ ×

$l_4 = 4 \quad 2^4 = 16 \geq 1/0.12 = 8.33$ ✓ $\quad \therefore l_4 = 4$

For $i=5$

$l_5 = 4 \quad 2^4 = 16 \geq 1/0.08 = 12.5$ ✓ $\quad \therefore l_5 = 4$

Expand α_i in binary number

$\rightarrow \alpha_1 = 0$

$\rightarrow \alpha_2 = (0.4)_{10} = 0.4 \times 2 = 0.8 \rightarrow 0$

$0.8 \times 2 = 1.6 \rightarrow 1 \quad \rightarrow \alpha_2 = (0.01)_2$

$\rightarrow \alpha_3 = (0.65)_{10} = 0.65 \times 2 = 1.3 \rightarrow 1$

$0.3 \times 2 = 0.6 \rightarrow 0$

$0.6 \times 2 = 1.2 \rightarrow 1 \quad \rightarrow \alpha_3 = (0.101)_2$

$\rightarrow \alpha_4 = (0.8)_{10} = 0.8 \times 2 = 1.6 \rightarrow 1$

$0.6 \times 2 = 1.2 \rightarrow 1$

$0.2 \times 2 = 0.4 \rightarrow 0$

$0.4 \times 2 = 0.8 \rightarrow 0 \quad \rightarrow \alpha_4 = (0.1100)_2$

$\rightarrow \alpha_5 = (0.92)_{10} = 0.92 \times 2 = 1.84 \rightarrow 1$

$0.84 \times 2 = 1.68 \rightarrow 1$

Keerthana Ashok

Date

Page

$0.68 \times 2 = 1.36 \rightarrow 1$

$0.36 \times 2 = 0.72 \rightarrow 0 \rightarrow \alpha_5 = (0.1110)_2$

The desired code is

$\alpha_1 = 0$ $\alpha_4 = 1100$

$\alpha_2 = 01$ $\alpha_5 = 1110$

$\alpha_3 = 101$

Average length

$$L = \sum_{i=1}^{5} P_i l_i$$

$L = 0.4(2) + 0.25(2) + 0.15(3) + 0.12(4) + 0.08(4)$

$L = 0.8 + 0.5 + 0.45 + 0.48 + 0.32$

$L = 2.55$ binits/msg symbol

Entropy

$$H(S) = \sum_{i=1}^{5} P_i \log_2 \frac{1}{P_i}$$

$$H(S) = 0.4 \log_2 \frac{1}{0.4} + 0.25 \log_2 \frac{1}{0.25} + 0.15 \log_2 \frac{1}{0.15} + 0.12 \log_2 \frac{1}{0.12} + 0.08 \log_2 \frac{1}{0.08}$$

$H(S) = 0.53 + 0.5 + 0.41 + 0.36 + 0.29$

$H(S) = 2.1$ bits/msg symbol

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{2.1}{2.55} = 82.35\%$$

Redundancy

$1 - \eta = 17.65\%$

Q18: compute the shannon's algorithm. Also calculate efficiency and redundancy for the following data.

	m_1	m_2	m_3	m_4	m_5
P:	1/8	1/16	3/16	1/4	3/8

Keerthana Ashok

Step 1: non increasing order of probabilities

given:

m_1	m_2	m_3	m_4	m_5
2/16	1/16	3/16	4/16	6/16

therefore

P_1	P_2	P_3	P_4	P_5
6/16	4/16	3/16	2/16	1/16

Step 2:

$\alpha_1 = 0$

$\alpha_2 = P_1 + \alpha_1 = 6/16 + 0 = 6/16 = 0.375$

$\alpha_3 = P_2 + \alpha_2 = 4/16 + 6/16 = 10/16 = 0.625$

$\alpha_4 = P_3 + \alpha_3 = 3/16 + 10/16 = 13/16 = 0.8125$

$\alpha_5 = P_4 + \alpha_4 = 2/16 + 13/16 = 15/16 = 0.9375$

$\alpha_6 = P_5 + \alpha_5 = 15/16 + 1/16 = 16/16 = 1$

Step 3:

To satisfy $2^{l_i} \geq 1/P_i$ condition for $i = 1$ to 5

For $i = 1$

$l_1 = 1 \quad 2^1 \geq 1/P_1 = 16/6 = 2.67$ ×

$l_1 = 2 \quad 2^2 = 4 \geq 1/0.375 = 2.67$ ✓ $\quad l_1 = 2$

For $i = 2$

$l_2 = 2 \quad 2^2 = 4 \geq 1/P_2 = 16/4 = 4$ ✓ $\quad l_2 = 2$

For $i = 3$

$l_3 = 3 \quad 2^3 = 8 \geq 1/P_3 = 16/3 = 5.33$ ✓ $\quad l_3 = 3$

For $i = 4$

$l_4 = 3 \quad 2^3 = 8 \geq 1/P_4 = 16/2 = 8$ ✓ $\quad l_4 = 3$

For $i = 5$

$l_5 = 4 \quad 2^4 = 16 \geq 1/P_5 = 16/1 = 16$ ✓ $\quad l_5 = 4$

Step 4:

→ $\alpha_1 = 0$

→ $\alpha_2 = (0.375)_{10} =$ $0.375 \times 2 = 0.75 \rightarrow 0$
$0.75 \times 2 = 1.5 \rightarrow 1$ ↳ $\alpha_2 = (0.01)_2$

→ $\alpha_3 = (0.625)_{10} =$ $0.625 \times 2 = 1.25 \rightarrow 1$
$0.25 \times 2 = 0.5 \rightarrow 0$
$0.5 \times 2 = 1 \rightarrow 1$ ↳ $\alpha_3 = (0.101)_2$

Keerthana Ashok

$\rightarrow \alpha_4 = (0.8125)_{10} = 0.8125 \times 2 = 1.625 \rightarrow 1$
$0.625 \times 2 = 1.25 \rightarrow 1$
$0.25 \times 2 = 0.5 \rightarrow 0$
$\alpha_4 = (0.110)_2$

$\rightarrow \alpha_5 = (0.9375)_{10} = 0.9375 \times 2 = 1.875 \rightarrow 1$
$0.875 \times 2 = 1.75 \rightarrow 1$
$0.75 \times 2 = 1.5 \rightarrow 1$
$0.5 \times 2 = 1 \rightarrow 1$
$\alpha_5 = (0.1111)_2$

Step 5: desired code

$\alpha_1 = 0$ $\alpha_4 = 110$
$\alpha_2 = 01$ $\alpha_5 = 1111$
$\alpha_3 = 101$

Average length

$$L = \sum_{i=1}^{5} P_i l_i$$

$$L = \frac{6}{16}(2) + \frac{4}{16}(2) + \frac{3}{16}(3) + \frac{2}{16}(3) + \frac{1}{16}(4)$$

$$L = 0.75 + 0.5 + 0.562 + 0.375 + 0.25$$

$$L = 2.4375 \text{ binits/msg symbol}$$

Entropy

$$H(S) = \sum_{i=1}^{5} P_i \log \frac{1}{P_i}$$

$$H(S) = \frac{6}{16} \log_2 \frac{16}{6} + \frac{4}{16} \log_2 \frac{16}{4} + \frac{3}{16} \log_2 \frac{16}{3} + \frac{2}{16} \log_2 \frac{16}{2} + \frac{1}{16} \log_2 16$$

$$H(S) = 0.53 + 0.5 + 0.45 + 0.375 + 0.25$$

$$H(S) = 2.105 \text{ bits/msg symbol}$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{2.105}{2.4375} = 86.35\%$$

Redundancy

$$R(S) = 1 - \eta = 13.65\%$$

* Extension of zero memory source:

Let us consider a binary source S emitting symbols s1 and s2 with probabilities of P1 and P2 respectively such that $P_1 + P_2 = 1$ ——(1)

Then the 2^{nd} extension of this binary source will have $2^2 = 4$ number of symbols given by:

$s_1 s_1 \quad : \quad P_1 P_1 = P_1^2$

$s_1 s_2 \quad : \quad P_1 P_2$

$s_2 s_1 \quad : \quad P_2 P_1$

$s_2 s_2 \quad : \quad P_2 P_2 = P_2^2$

The sum of all probabilities of the 2^{nd} extended source is $P_1^2 + 2P_1P_2 + P_2^2 = (P_1 + P_2)^2 = 1$ ——(2)

Entropy of the basic binary source

$$H(S) = \sum_{i=1}^{2} P_i \log \frac{1}{P_i}$$

$$H(S) = P_1 \log \frac{1}{P_1} + P_2 \log \frac{1}{P_2} \quad \text{——(3)}$$

The entropy of the 2^{nd} extended source is

$$H(S^2) = \sum_{i=1}^{4} P_i \log \frac{1}{P_i}$$

$$H(S^2) = P_1^2 \log \frac{1}{P_1^2} + P_1P_2 \log \frac{1}{P_1P_2} + P_2P_1 \log \frac{1}{P_2P_1} + P_2^2 \log \frac{1}{P_2^2}$$

$$H(S^2) = 2P_1^2 \log \frac{1}{P_1} + 2P_1P_2 \log \frac{1}{P_1P_2} + 2P_2^2 \log \frac{1}{P_2}$$

$$H(S^2) = 2P_1^2 \log \frac{1}{P_1} + 2P_1P_2 \log \frac{1}{P_1} + 2P_1P_2 \log \frac{1}{P_2} + 2P_2^2 \log \frac{1}{P_2}$$

$$H(S^2) = 2P_1 \underbrace{(P_1 + P_2)}_{1} \log \frac{1}{P_1} + 2P_2 \underbrace{(P_1 + P_2)}_{1} \log \frac{1}{P_2}$$

$$H(S^2) = 2\left[P_1 \log \frac{1}{P_1} + P_2 \log \frac{1}{P_2}\right]$$

From eq ③

$$H(S^2) = 2H(S)$$

similarly for third extension

$$H(S^3) = 3H(S)$$

Hence generally.

For n^{th} extension

$$\boxed{H(S^n) = nH(S)}$$

Ex:

S_1	S_2	S_3
0.5	0.3	0.2

2nd Extension:

S_1S_1 : $0.5(0.5) = 0.25$

S_1S_2 : $0.5(0.3) = 0.15$

S_1S_3 : $0.5(0.2) = 0.10$

S_2S_1 : $0.3(0.5) = 0.15$

S_2S_2 : $0.3(0.3) = 0.09$

S_2S_3 : $0.3(0.2) = 0.06$

S_3S_1 : $0.2(0.5) = 0.10$

S_3S_2 : $0.2(0.3) = 0.06$

S_3S_3 : $0.2(0.2) = 0.04$

★ Shannon - Fano encoding algorithm:

Step1 : List the source symbol in the order of nonincreasing probabilities.

Step2 : Split the list into two parts, with the total probability of both the parts being as close to each other as possible.

Step 3: Assign the value 1 to the first group and 0 to the second group.

Step 4: Repeat step 2 and step 3 for each group until all the symbols are split into individual subgroups.

NOTE: The shannon codes are considered accurate if the code of each symbol is unique.

Q19: given the message symbols $x_1, x_2, x_3, x_4, x_5, x_6$ with probabilities 0.4, 0.2, 0.2, 0.1, 0.07, 0.03. construct the binary code by applying shannon-Fano encoding algorithm and find the efficiency.

- step1: non increasing order of probabilities

step2, 3&4:

x_1:	$P_1 = 0.4$	1									
x_2:	$P_2 = 0.2$	0	0.2	1							
x_3:	$P_3 = 0.2$	0	0.2	0	0.2	1					
x_4:	$P_4 = 0.1$	0	0.1	0	0.1	0	0.1	1			
x_5:	$P_5 = 0.07$	0	0.07	0	0.07	0	0.07	0	0.07	1	
x_6:	$P_6 = 0.03$	0	0.03	0	0.03	0	0.03	0	0.03	0	

	code	l_i
x_1	1	1
x_2	01	2
x_3	001	3
x_4	0001	4
x_5	00001	5
x_6	00000	5

Average length

$$L = \sum_{i=1}^{6} P_i l_i$$

$$L = 0.4(1) + 0.2(2) + 0.2(3) + 0.1(4) + 0.07(5) + (0.03)5$$

$$L = 0.4 + 0.4 + 0.6 + 0.4 + 0.35 + 0.15$$

$$L = 2.3 \text{ bits/msg symbol}$$

Entropy

$$H(s) = \sum_{i=1}^{6} P_i \log_2 \frac{1}{P_i}$$

$$H(s) = 0.4 \log_2 \frac{1}{0.4} + 0.2 \log_2 \frac{1}{0.2} + 0.2 \log_2 \frac{1}{0.2} + 0.1 \log_2 \frac{1}{0.1} + 0.07 \log_2 \frac{1}{0.07} + 0.03 \log_2 \frac{1}{0.03}$$

$$H(s) = 0.53 + 0.46 + 0.46 + 0.33 + 0.27 + 0.15$$

$$H(s) = 2.19 \text{ bits/msg symbol}$$

Efficiency

$$\eta = \frac{H(s)}{L} = \frac{2.19}{2.3} = 95.2\%$$

Keerthana Ashok

Q.20: construct the binary code by applying Shannon-Fano coding algorithm for the following data

x_1	x_2	x_3	x_4
0.1	0.2	0.3	0.4

calculate efficiency. Also find the probability of 0's and 1's in the code.

x_4 :	$P_1 = 0.4$	1				
x_3 :	$P_2 = 0.3$	0	0.3	1		
x_2 :	$P_3 = 0.2$	0	0.2	0	0.2	1
x_1 :	$P_4 = 0.1$	0	0.1	0	0.1	0

	code	l_i
x_4	1	1
x_3	01	2
x_2	001	3
x_1	000	3

Average length

$$L = \sum_{i=1}^{4} P_i l_i$$

$$L = 0.4(1) + 0.3(2) + 0.2(3) + 0.1(3)$$

$$L = 0.4 + 0.6 + 0.6 + 0.3$$

$$L = 1.9 \text{ binits/msg symbol}$$

Entropy:

$$H(S) = \sum_{i=1}^{4} P_i \log_2 \frac{1}{P_i}$$

$$H(S) = 0.4 \log_2 \frac{1}{0.4} + 0.3 \log_2 \frac{1}{0.3} + 0.2 \log_2 \frac{1}{0.2} + 0.1 \log_2 \frac{1}{0.1}$$

$$H(S) = 0.52 + 0.52 + 0.46 + 0.33$$

$$H(S) = 1.83 \text{ bits/msg symbol}$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{1.83}{1.9} = 96.31\%$$

Probability of 0's

$$P(0) = \frac{1}{L} \sum_{i=1}^{4} (\text{number of 0's in the code}) \cdot P_i$$

$$P(0) = \frac{1}{1.9} \left[0.4(0) + 0.3(1) + 0.2(2) + 0.1(3) \right]$$

$$P(0) = \frac{1}{1.9} \left[0.3 + 0.4 + 0.3 \right] = \frac{1}{1.9} = 0.526$$

Probability of 1's

$P(1) = 1 - P(0) = 1 - 0.526 = \underline{0.474}$

Q21: The discrete source has the alphabets of 7 symbols with probability for its output listed below.

symbols	s_1	s_2	s_3	s_4	s_5	s_6	s_7
probabilities	0.25	0.25	0.125	0.125	0.125	0.0625	0.0625

compute shannon-Fano encoding algorithm.

s_1 :	$P_1 = 0.25$	1	0.25	1				
s_2 :	$P_2 = 0.25$	1	0.25	0				
s_3 :	$P_3 = 0.125$	0	0.125	1	0.125	1		
s_4 :	$P_4 = 0.125$	0	0.125	1	0.125	0		
s_5 :	$P_5 = 0.125$	0	0.125	0	0.125	1		
s_6 :	$P_6 = 0.0625$	0	0.0625	0	0.0625	0	0.0625	1
s_7 :	$P_7 = 0.0625$	0	0.0625	0	0.0625	0	0.0625	0

	code	l_i
s_1	11	2
s_2	10	2
s_3	011	3
s_4	010	3
s_5	001	3
s_6	0001	4
s_7	0000	4

Average length

$$L = \sum_{i=1}^{7} P_i l_i$$

$$L = 0.25(2) + 0.25(2) + 0.125(3) + 0.125(3) + 0.125(3) + 0.0625(4) + 0.0625(4)$$

$$L = 0.5 + 0.5 + 3(0.375) + 2(0.25)$$

$$\underline{L = 2.625 \text{ binits/msg symbol}}$$

Entropy

$$H(S) = \sum_{i=1}^{7} P_i \log_2 \frac{1}{P_i}$$

$$H(S) = 0.25 \log_2 \frac{1}{0.25} + 0.25 \log_2 \frac{1}{0.25} + 0.125 \log_2 \frac{1}{0.125} + 0.125 \log_2 \frac{1}{0.125} + 0.125 \log_2 \frac{1}{0.125} + 0.0625 \log_2 \frac{1}{0.0625} + 0.0625 \log_2 \frac{1}{0.0625}$$

$$H(S) = 2(0.25)(2) + 3(0.125)(3) + 2(0.0625)(4)$$

$$H(S) = 1 + 1.125 + 0.5$$

$$\underline{H(S) = 2.625 \text{ bits/msg symbol}}$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{2.625}{2.625} = \underline{100\%} \quad (\text{because } H(S) = L)$$

Q22: Source produces 2 symbols S_1 and S_2 with probabilities 7/8 and 1/8 respectively. compute Shannon-Fano encoding algorithm to get the coding effeciency of atleast 75%

S_1 : $\boxed{P_1 = 7/8}$ 1

S_2 : $\boxed{P_2 = 1/8}$ 0

	code	l_i
S_1	1	1
S_2	0	1

Average length

$$L = \sum_{i=1}^{2} P_i l_i = \frac{7}{8}(1) + \frac{1}{8}(1)$$

$$\underline{L = 1 \text{ binit / msg symbol}}$$

Entropy $= H(S) = \sum_{i=1}^{2} P_i \log_2 \frac{1}{P_i}$

$$H(S) = \frac{7}{8} \log_2 \frac{8}{7} + \frac{1}{8} \log_2 8$$

$$H(S) = 0.168 + 0.375 = \underline{0.543} \text{ bits / msg symbol}$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{0.543}{1} = 54.3\%$$

As the efficiency is less than 75% let us consider the 2nd extension.

S_1S_1 : $P_1P_1 = (7/8)(7/8) = \boxed{0.7656}$ 1

S_1S_2 : $P_1P_2 = (7/8)(1/8) = \boxed{0.1093}$ 0 $\boxed{0.1093}$ 1

S_2S_1 : $P_2P_1 = (1/8)(7/8) = \boxed{0.1093}$ 0 $\boxed{0.1093}$ 0 $\boxed{0.1093}$ 1

S_2S_2 : $P_2P_2 = (1/8)(1/8) = \boxed{0.0156}$ 0 $\boxed{0.0156}$ 0 $\boxed{0.0156}$ 0

	code	l_i
S_1S_1 :	1	1
S_1S_2	01	2
S_2S_1	001	3
S_2S_2	000	3

Average length

$$L = \sum_{i=1}^{4} P_i l_i$$

$$L = 0.7656(1) + 0.1093(2) + 0.1093(3) + 0.0156(3)$$

Keerthana Ashok

$L = 0.7656 + 0.5465 + 0.0468$

$\underline{L = 1.3589 \text{ binits/msg symbol}}$

Entropy

$$H(S^2) = \sum_{i=1}^{4} P_i \log_2 \frac{1}{P_i}$$

$$H(S^2) = 0.7656 \log_2 \frac{1}{0.7656} + 0.1093 \log_2 \frac{1}{0.1093}$$

$$+ 0.1093 \log_2 \frac{1}{0.1093} + 0.0156 \log_2 \frac{1}{0.0156}$$

$$H(S^2) = 0.295 + 2(0.349) + 0.093$$

$$\underline{H(S^2) = 1.086 \text{ bits/msg symbol}} = 2H(S) = 2(0.543)$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{1.086}{1.3589} = \underline{\underline{79.91\%}}$$

★ Huffman coding: / compact coding:

Step 1: Arrange the symbols in nonincreasing order of probabilities.

Step 2: Add the two lowest values and remove the symbols. At the end we will be left with only two values.

Step 3: Moving backwards assign the bits to the values.

compression technique

Ex:

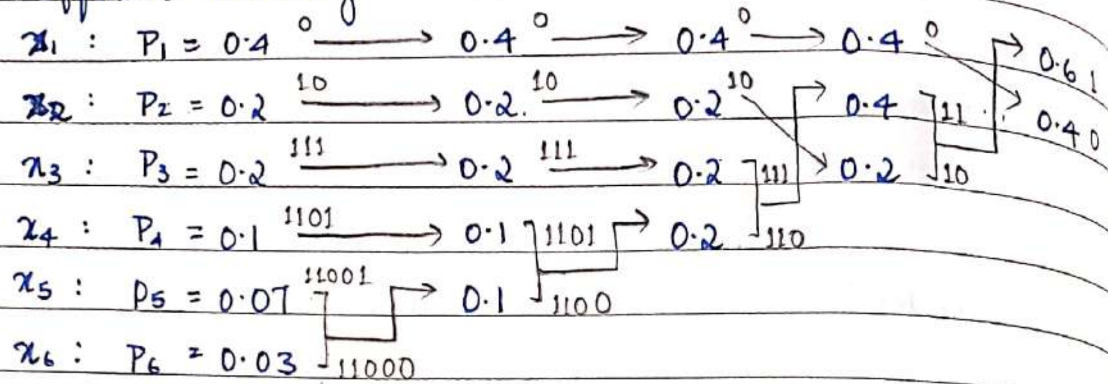
	P_i			code	l_i
S_1	0.4 (0)	0.4 (0)	0.6 (1)	0	1
S_2	0.3 (11)	0.3 (11)	0.4 (0)	11	2
S_3	0.2 (101)	0.3 (10)		101	3
S_4	0.1 (100)			100	3

Q23: For the given symbols, construct binary code by applying Huffman encoding procedure. Determine the efficiency and redundancy of the code formed.

x_1	x_2	x_3	x_4	x_5	x_6
0.4	0.2	0.2	0.1	0.07	0.03

Keerthana Ashok

- Huffman encoding:



	code	l_i
x_1	0	1
x_2	10	2
x_3	111	3
x_4	1101	4
x_5	11001	5
x_6	11000	5

Average length

$$L = \sum_{i=1}^{6} P_i l_i$$

$$L = 0.4(1) + 0.2(2) + 0.2(3) + 0.1(4) + 0.07(5) + 0.03(5)$$

$$L = 0.4 + 0.4 + 0.6 + 0.4 + 0.35 + 0.15$$

$$\underline{L = 2.3 \text{ binits} / \text{msg symbol}}$$

Entropy

$$H(S) = \sum_{i=1}^{6} P_i \log \frac{1}{P_i}$$

$$H(S) = 0.4 \log_2 \frac{1}{0.4} + 0.2 \log_2 \frac{1}{0.2} + 0.2 \log_2 \frac{1}{0.2} + 0.1 \log_2 \frac{1}{0.1} + 0.07 \log_2 \frac{1}{0.07} + 0.03 \log_2 \frac{1}{0.03}$$

$$H(S) = 0.52 + 0.46 + 0.46 + 0.33 + 0.27 + 0.15$$

$$\underline{H(S) = 2.19 \text{ bits} / \text{msg symbol}}$$

Efficiency

$$\eta = \frac{H(S)}{L} = \frac{2.19}{2.3} = \underline{95.2\%}$$

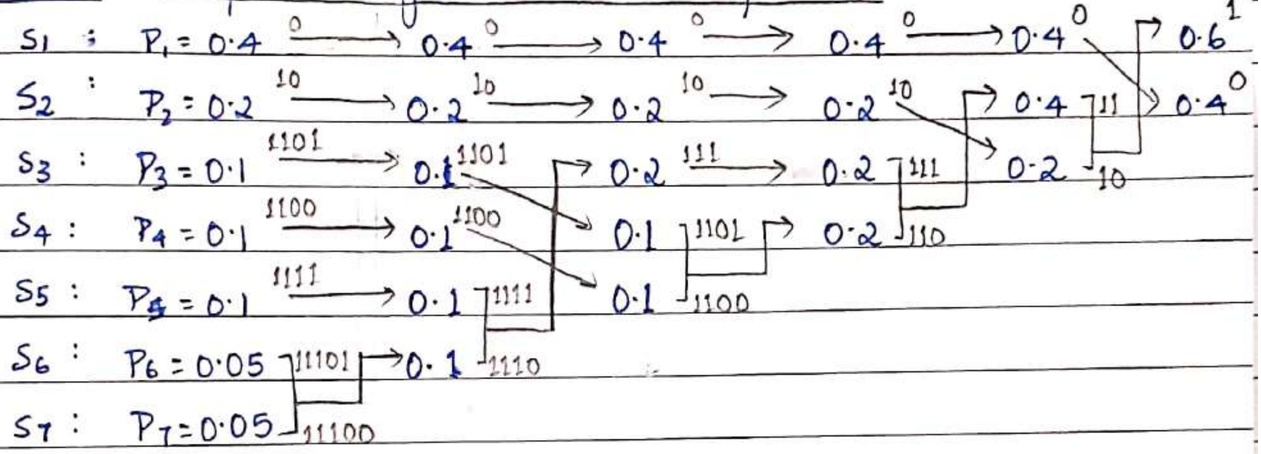
Redundancy

$$R(S) = 1 - \eta = \underline{4.8\%}$$

Keerthana Ashok

Q24: consider a zero memory source with source symbol $S_1, S_2, S_3, S_4, S_5, S_6, S_7$ with probabilities 0.4, 0.2, 0.1, 0.1, 0.1, 0.05, 0.05 respectively. construct the binary Huffman code by placing the composite symbol as low as you can and repeat by moving the composite symbol as high as possible. Compute the variance of word length.

– CASE 1 : composite symbol as low as possible



	code	l_i
S_1	0	1
S_2	10	2
S_3	1101	4
S_4	1100	4
S_5	1111	4
S_6	11101	5
S_7	11100	5

Average length

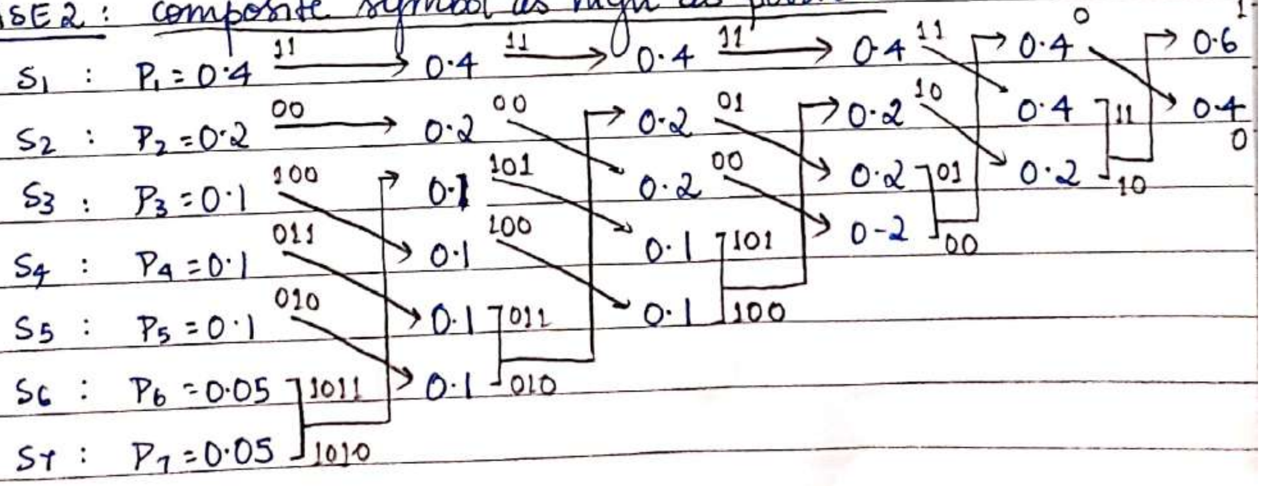
$$L = \sum_{i=1}^{7} P_i l_i$$

$$L = 0.4(1) + 0.2(2) + 0.1(4) + 0.1(4) + 0.1(4) + 0.05(5) + 0.05(5)$$

$$L = 0.4 + 0.4 + 1.2 + 0.5$$

$$\underline{L = 2.5 \text{ binits / msg symbol}}$$

CASE 2 : composite symbol as high as possible



Keerthana Ashok

	code	l_i
S_1	11	2
S_2	00	2
S_3	100	3
S_4	011	3
S_5	010	3
S_6	1011	4
S_7	1010	4

Average length

$$L = \sum_{i=1}^{7} P_i l_i$$

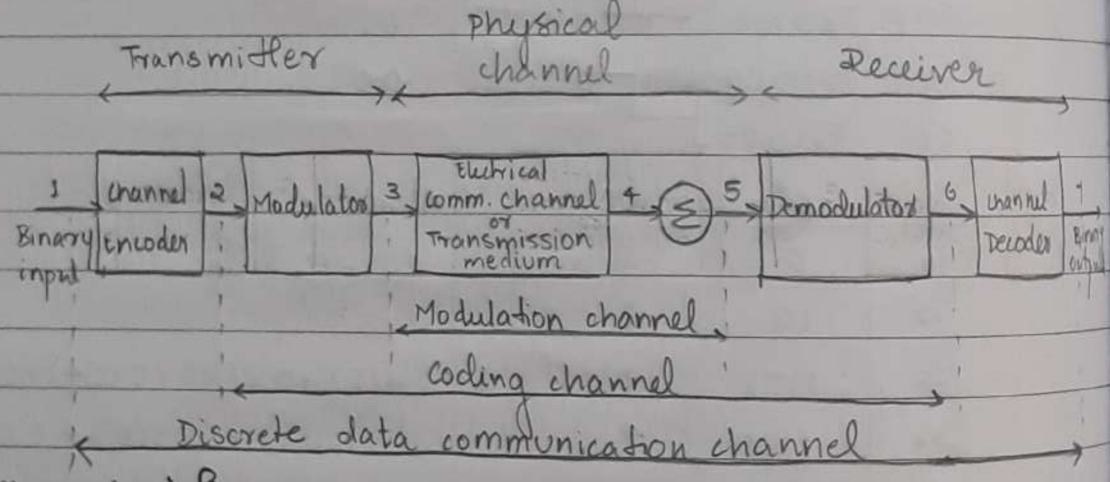
$$L = 0.4(2) + 0.2(2) + 0.1(3) + 0.1(3) + 0.1(3) + 0.05(4) + 0.05(4)$$

$$L = 0.8 + 0.4 + 0.9 + 0.4$$

$$\underline{\underline{L = 2.5 \text{ binits/msg symbol}}}$$

$$\text{Variance}(L) = \sum_{i=1}^{7} P_i (l_i - L) = 0$$

★ **Communication Channel:**

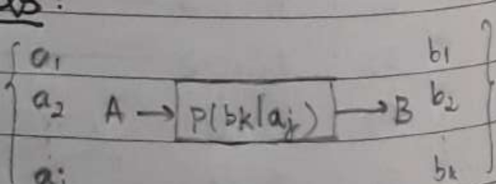


Characteristics:

- Channel Capacity: maximum rate of data transmitted with minimum probability of error.

★ **Discrete Memoryless Channels:**

It is statistical model with an input A and an output B that is a noisy version of A. Every unit of time, the channel accepts an input symbol and in response, it emits an output symbol.



Discrete memoryless channel

The channel is said to be 'discrete' when both of the alphabets A and B have finite sizes. It is said to be 'memoryless' when the current output symbol depends only on the current input symbol and not any of the previous ones.

The channel is described in terms of an input alphabet

$A = \{a_1, a_2, \ldots, a_j\}$ where j is number of symbols.

an output alphabet

$B = \{b_1, b_2, \ldots, b_k\}$ where k is number of symbols

and a set of transition probabilities.

$p(b_k | a_j) = P(B = b_k | A = a_j)$.

A convenient way of describing a discrete memoryless channel is to arrange the various transition/conditional probabilities of the channel in the form of a matrix:

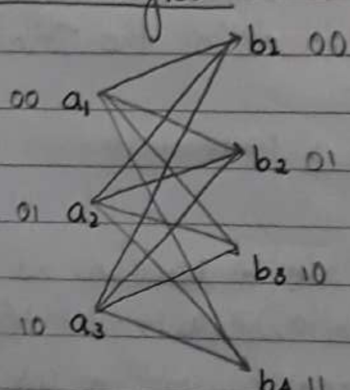
$$P(b_k|a_j) = \begin{array}{c c} \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ \\ a_j \end{array} & \begin{bmatrix} P(b_1|a_1) & P(b_2|a_1) & \cdots & P(b_k|a_1) \\ P(b_1|a_2) & P(b_2|a_2) & \cdots & P(b_k|a_2) \\ \vdots & \vdots & \vdots & \vdots \\ P(b_1|a_j) & P(b_2|a_j) & \cdots & P(b_k|a_j) \end{bmatrix} \\ & \begin{array}{cccc} b_1 & \quad\quad b_2 & \quad\quad \cdots & \quad\quad b_k \end{array} \end{array}$$

Matrix P is called the channel matrix. Each row corresponds to fixed channel input and each column corresponds to fixed channel output.

property: The sum of the elements along any row of the matrix is always equal to one.

i.e., $\sum_{k=1}^{k} P(b_k|a_j) = 1$ for all j

Noise Diagram



$P_{11} = P(b_1|a_1)$ $\quad P_{21} = P(b_1|a_2)$

$P_{12} = P(b_2|a_1)$ $\quad P_{22} = P(b_2|a_2)$

$P_{13} = P(b_3|a_1)$ $\quad P_{23} = P(b_3|a_2)$

$P_{14} = P(b_4|a_1)$ $\quad P_{24} = P(b_4|a_2)$

$P_{31} = P(b_1|a_3)$

$P_{32} = P(b_2|a_3)$

$P_{33} = P(b_3|a_3)$

$P_{34} = P(b_4|a_3)$

Keerthana Ashok

By the property of channel matrix

$P_{11} + P_{12} + P_{13} + P_{14} = 1$

$P(b_1|a_1) + P(b_2|a_1) + P(b_3|a_1) + P(b_4|a_1) = 1$

i.e. $\sum_{r=1}^{k} P(b_r|a_1) = 1$ —— ①

but $P(a_1) + P(a_2) + \cdots + P(a_j) = 1$

i.e. $\sum_{s=1}^{j} P(a_s) = 1$ —— ②

The joint probability of both A and B is

$P(a_j, b_k) = P(a_j|b_k)P(b_k) = P(b_k|a_j)P(a_j)$

$\therefore \boxed{P(a_j|b_k) = \frac{P(b_k|a_j)P(a_j)}{P(b_k)}}$ input conditional probability

The joint probability matrix is given by:

$$P(a_j, b_k) = \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_j \\ \\ \end{array}\begin{bmatrix} P(a_1,b_1) & P(a_1,b_2) & \cdots & P(a_1,b_k) \\ P(a_2,b_1) & P(a_2,b_2) & \cdots & P(a_2,b_k) \\ \vdots & \vdots & & \vdots \\ P(a_j,b_1) & P(a_j,b_2) & \cdots & P(a_j,b_k) \end{bmatrix}$$
$$\quad\quad b_1 \quad\quad b_2 \quad \cdots \quad b_k$$

Properties:

- By adding the elements of Joint Probability Matrix column wise, we can obtain the probability of the output symbol.
$$\sum_{m=1}^{j} P(a_m, b_k) = P(b_k)$$
- By adding the elements of Joint Probability Matrix row wise, we can obtain the probability of the input symbol. $\sum_{n=1}^{k} P(a_j, b_n) = P(a_j)$
- The sum of all the elements of the Joint Probabity matrix is unity.
$$\sum_{m=1}^{j}\sum_{n=1}^{k} P(a_m, b_n) = 1$$

Marginal Probabilities:

- The output symbol $P(b_1)$ can be written as
$P(b_1) = P(b_1|a_1)P(a_1) + P(b_1|a_2)P(a_2) + \cdots + P(b_1|a_j)P(a_j)$
- The input symbol $P(a_1)$ can be written as
$P(a_1) = P(a_1|b_1)P(b_1) + P(a_1|b_2)P(b_2) + \cdots + P(a_1|b_k)P(b_k)$

Q: In the communication system, a transmitter has three symbols $a = \{a_1, a_2, a_3\}$ and receiver has three output symbols $b = \{b_1, b_2, b_3\}$. The matrix given shows the joint probability matrix with some marginal probabilities. Find:

ⓐ Missing probabilities.

ⓑ. $P(b_3/a_1)$ and $P(a_1|b_3)$

ⓒ Are the events a_1 and b_1 statistically independent? If yes why?

	b_1	b_2	b_3
a_1	$1/12$	*	$5/36$
a_2	$5/36$	$1/9$	$5/36$
a_3	*	$1/6$	*

$P(a_1, b_2)$ $P(a_3, b_1)$ $P(a_3, b_3)$

$P(b_j)$	$1/3$	$14/36$	*

$P(b_3)$

— ⓐ

wkt sum of output probabilities is unity.

$$P(b_j) = P(b_1) + P(b_2) + P(b_3) = 1$$

$$1/3 + 14/36 + P(b_3) = 1$$

$$\therefore P(b_3) = 5/18 = 2.44$$

The probability of output symbol b_1 is given by

$$P(b_1) = P(a_1, b_1) + P(a_2, b_1) + P(a_3, b_1)$$

$$1/3 = 1/12 + 5/36 + P(a_3, b_1)$$

$$\therefore P(a_3, b_1) = 1/9 = 0.11$$

The probability of output symbol b_2 is given by

$$P(b_2) = P(a_1, b_2) + P(a_2, b_2) + P(a_3, b_2)$$

$$14/36 = P(a_1, b_2) + 1/9 + 1/6$$

$$\therefore P(a_1, b_2) = 1/9 = 0.11$$

The probability of output symbol b_3 is given by

$$P(b_3) = P(a_1, b_3) + P(a_2, b_3) + P(a_3, b_3)$$

$$5/18 = 5/36 + 5/36 + P(a_3, b_3)$$

$$\therefore P(a_3, b_3) = 0$$

ⓑ

$$P(b_3|a_1) = \frac{P(a_1, b_3)}{P(a_1)}$$

$$P(a_1) = P(a_1, b_1) + P(a_1, b_2) + P(a_1, b_3)$$

$$= 1/12 + 1/9 + 5/36$$

$$= 1/3$$

$$P(b_3|a_1) = \frac{5/36}{1/3} = \frac{5}{12}$$

Keerthana Ashok

$$P(a_1|b_3) = \frac{P(a_1,b_3)}{P(b_3)} = \frac{5/36}{5/18} = \frac{1}{2}$$

(c) If a_1 and b_1 are statistically independent then it has to satisfy the following condition:

$$P(a_1, b_1) = P(a_1)P(b_1)$$

$$\frac{1}{12} = \frac{1}{3}\left(\frac{1}{3}\right)$$

$$\frac{1}{12} \neq \frac{1}{9}$$

note: If a_j and b_k are statistically independent then

$$P(a_j, b_k) = P(a_j)P(b_k)$$

Hence a_1 and b_1 are not statistically independent.

* **Mutual Information:**

The mutual information between two random variables measures the amount of information that one conveys about each other.

It conveys the average reduction in uncertainty about A when B is known. It is defined as:

$$I(A;B) = \sum_{a,b} P(a,b) \log_2 \frac{P(a,b)}{P(a)P(b)}$$

In case of A and B being independent random variables, $P(a,b) = P(a)P(b)$. Therefore the mutual information is zero for statistically independent random variables.

Properties

- Mutual information is always ≥ 0.
- The mutual information of a random variable with itself is its entropy

$$I(A;A) = H(A) \quad \text{(self information)}$$

Hence

$$I(A;B) = H(A) - H(A|B)$$

$$I(A;B) = H(B) - H(B|A) = I(B;A)$$

$$I(A;B) = H(A) + H(B) - H(A|B)$$

conditional entropy

★ Channel Capacity:

Channel capacity is defined as maximum mutual information $I(A;B)$ in any single use of the channel (i.e., signaling interval), where the maximization is over all possible input probability distributions $(P(a_i))$ on X. The channel capacity is denoted as C:

$$C = \max_{(P(a_j))} I(A;B)$$

unit: bits per channel use
bits per transmission.

★ Shannon Hartley Theorem / Channel capacity Theorem:

note:

- white noise: due to thermal motion of electron
- shot noise: due to movement of electrons across the semiconductor
- additive white gaussian noise:
- fade noise:

Channel capacity of bandwidth limited channel with Additive White Gaussian Noise is given by the Shannon Hartley Theorem.

$$C = B \log_2 \left[1 + \frac{S}{N}\right] \text{ bits/second}$$

where S/N : mean square signal to noise ratio
B : bandwidth

Q: Alphanumeric data are entered into a computer from a remote terminal through a voice grade telephone channel. Channel has a bandwidth of 3.4kHz and the output S/N ratio of 20dB. The terminal has total of 128 symbols. Assume that the symbols are equiprobable and the successive transmission are statistically independent. Calculate:

i. Channel capacity
ii. Average information content

iii. Maximum sample rate for which error free transmission over the channel is possible.

— Given:

BW = 3.4 kHz $\quad$ S/N = 20 dB = 100 $\quad$ q = 128

i. channel capacity.

$$C = B\log_2\left[1 + \frac{S}{N}\right]$$

$$C = 3.4 \times 10^3 \log_2\left[1 + 100\right]$$

$$\underline{C = 22.638\text{ k bits/sec}}$$

ii. Average information content

$$H_{max} = \log_2 q = \log_2 128$$

$$\underline{H_{max} = 7 \text{ bits/msg symbol}}$$

iii. Maximum sample rate

$R = r_s H(s)$: Information rate.

For error free transmission

$$R < C \Rightarrow r_s H(s) < C$$

$$r_s < \frac{C}{H(s)}$$

$$r_s < \frac{22.638 \times 10^3}{7}$$

$$\underline{r_s < 3.234\text{k bits/sec}}$$

Q: The CRT terminal is used to enter alphanumeric data into communication channel. The CRT is connected to voice grade telephone with bandwidth of 3k Hz and output S/N of 10dB, 128 symbols and data is sent in independent manner with equal probability.

Find: i. Average information per character.

ii. channel capacity.

iii. Maximum rate at which data can be sent from terminal without error.

— Given: $BW = 3 \times 10^3$ Hz

$S/N = 10$ dB $= 10$

$q = 128$.

i. Average information per character.

$$H_{max} = \log_2 q = \log_2 128$$

$$H_{max} = 7 \text{ bits/msg symbol}$$

ii. Channel capacity.

$$C = B \log_2 \left[1 + \frac{S}{N}\right]$$

$$C = 3 \times 10^3 \log_2 [1 + 10]$$

$$C = 10.378 \text{ k bits/sec}$$

iii. Maximum rate of transmission.

$$r_s < \frac{C}{H(S)}$$

$$r_s < \frac{10.378 \times 10^3}{7}$$

$$r_s < 1.48 \text{k bits/sec}$$

Q: Calculate the channel capacity for gaussian channel of bandwidth 1MHz and $S/N = 30$ dB.

— Given: $BW = 1 \times 10^6$ Hz

$S/N = 30$ dB $= 10^3$

channel capacity

$$C = B \log_2 \left[1 + \frac{S}{N}\right]$$

$$C = 1 \times 10^6 \log_2 [1 + 1000]$$

$$C = 9.967 \text{ M bits/sec}$$

Keerthana Ashok

Q: A friend of yours says he can design a system for transmitting output of minicomputer to a line printer at a rate of 30 lines/min over a voice grade telephone with a bandwidth of 3.5kHz and S/N of 30dB. Assume that the printer prints 8 bits of data per character and prints 80 character per line.

Find: i. Channel capacity

ii. Information rate.

iii. Maximum rate of transmission without error.

— Given:

$BW = 3.5 \times 10^3 \, Hz$

$S/N = 30 dB = 10^3$

i. Channel capacity.

$$C = B \log_2 \left[1 + \frac{S}{N}\right]$$

$$C = 3.5 \times 10^3 \log_2 \left[1 + 1000\right]$$

$$\underline{C = 34.885 \text{ k bits/sec}}$$

ii. Information rate.

$H = 8 \times 80 = \underline{640}$ bits/msg sym

$r_s = 30$ lines/min $= \frac{30}{60} = 0.5$ lines/sec

$$R = r_s H(S)$$

$$R = 0.5 (640)$$

$$\underline{R = 320 \text{ bits/sec}}$$

iii. Maximum transmission rate

$$r_s < \frac{C}{H}$$

$$r_s < \frac{34.885 \times 10^3}{640}$$

$$\underline{r_s < 54.5 \text{ bits/sec}}$$

Keerthana Ashok

Date
Page

Q: A voice grade channel of a telephone network has a bandwidth of 3.4 kHz. Calculate:

i. channel capacity of telephone channel for S/N of 30dB

ii. minimum S/N required to support through the telephone channel at a rate of 4800 bits/sec.

— Given:

$BW = 3.4 \times 10^3 \ Hz$

i. channel capacity

Given: $S/N = 30dB = 10^3$

$$C = B \log_2 \left[1 + \frac{S}{N} \right]$$

$$C = 3.4 \times 10^3 \log_2 [1 + 1000]$$

$$\underline{\underline{C = 33.888 k \text{ bits/sec}}}$$

ii. Channel capacity $C = 4800$ bits/sec.

$$\therefore C = B \log_2 \left[1 + \frac{S}{N} \right]$$

$$4800 = 3.4 \times 10^3 \log_2 \left[1 + \frac{S}{N} \right]$$

$$\log_2 \left[1 + \frac{S}{N} \right] = 1.412$$

$$\frac{\log_e [1 + S/N]}{\log_e (2)} = 1.412$$

$$\log_e \left[1 + \frac{S}{N} \right] = 0.978$$

$$\left[1 + \frac{S}{N} \right] = 2.659$$

$$\frac{S}{N} = 1.659 = \underline{\underline{2.198 dB}}$$