Presentation on
Digital communication
(ECE)
III- B.TECH V- Semester
(AUTONOMOUS-R16)

Prepared by,

Dr. S.Vinoth
(Associate Professor)

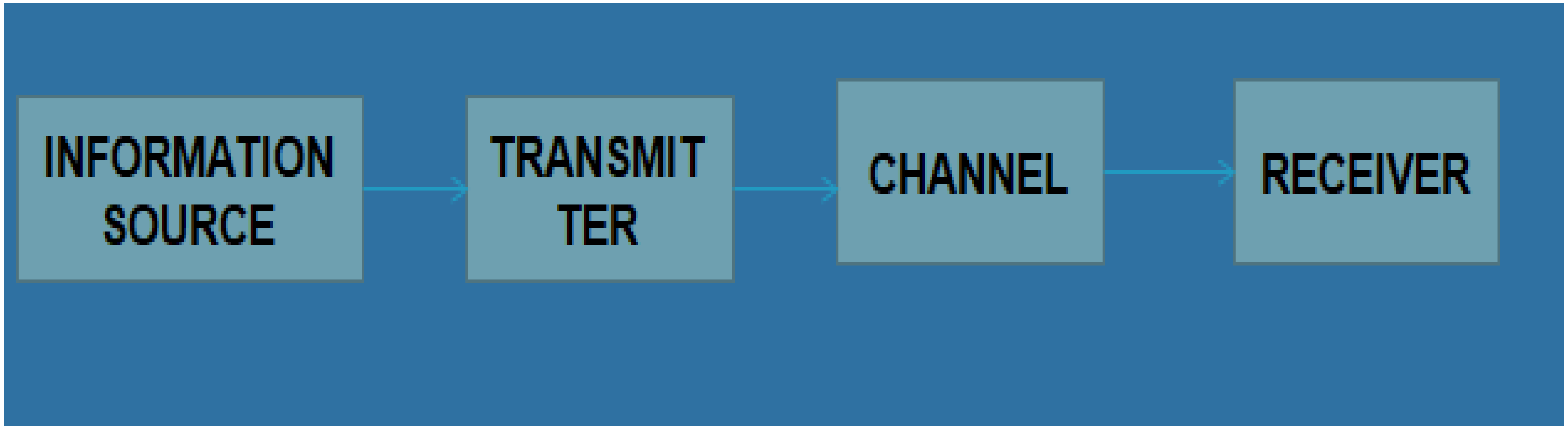
Mr. G.Kiran Kumar
(Assistant Professor)

Pulse Digital Modulation

COMMUNICATION

- The transmission of information from source to the destination through a channel or medium is called communication.

- Basic block diagram of a communication system:



COMMUNICATION cont..

- Source: analog or digital
- Transmitter: transducer, amplifier, modulator, oscillator, power amp., antenna
- Channel: e.g. cable, optical fiber, free space
- Receiver: antenna, amplifier, demodulator, oscillator, power amplifier, transducer
- Destination : e.g. person, (loud) speaker, computer

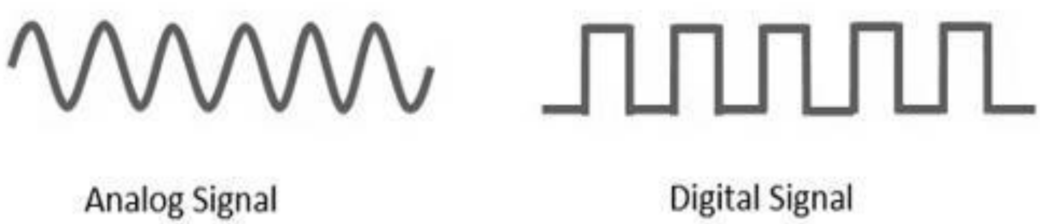
Necessity of Digital communication

- Good processing techniques are available for digital signals, such as medium.
- Data compression (or source coding)
- Error Correction (or channel coding)(A/D conversion)
- Equalization
- Security
- Easy to mix signals and data using digital techniques

Necessity of Digitalization

- In Analog communication, long distance communication suffers from many losses such as distortion, interference & security.
- To overcome these problems, signals are digitalized.

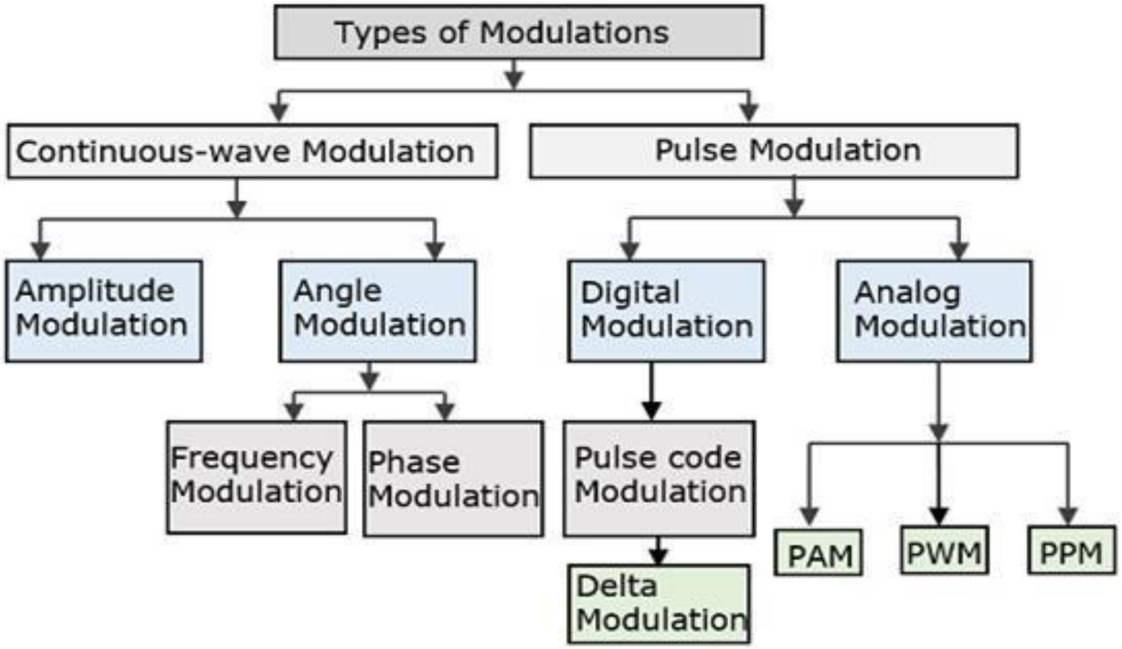
➢ Information is transferred in the form of **signal**.



Representation of Signals

PULSE MODULATION

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



Analog Pulse Modulation

In **Analog modulation** ,

- If the **amplitude**, **duration** or **position** of a pulse is varied in accordance with the instantaneous values of the baseband modulating signal, then such a technique is called as

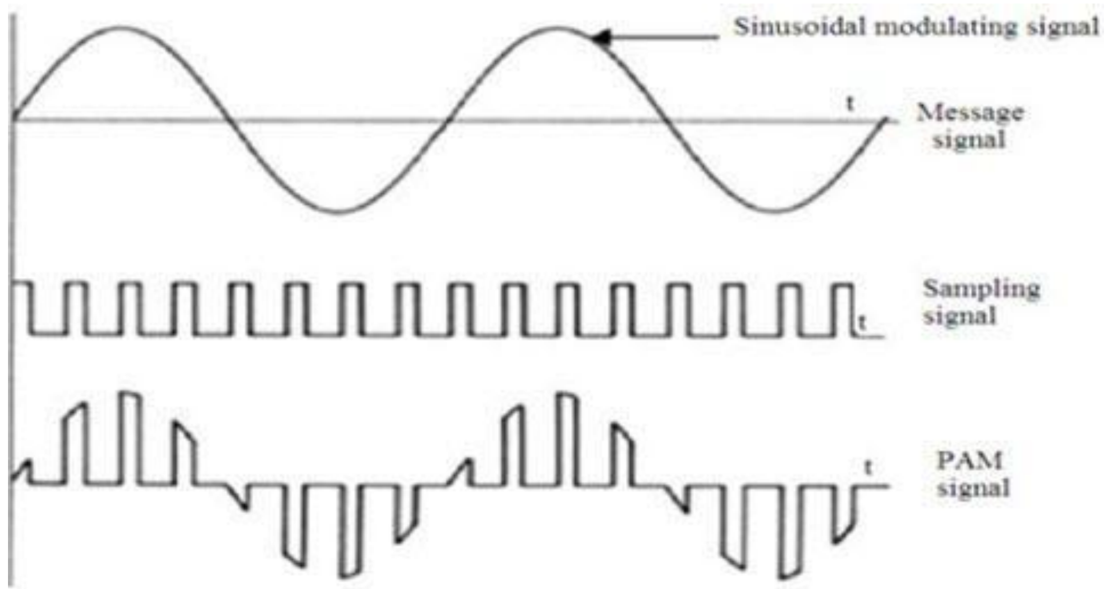
- **Pulse Amplitude Modulation (PAM)** or
- **Pulse Duration/Width Modulation (PDM/PWM)**, or
- **Pulse Position Modulation (PPM)**.

Digital modulation

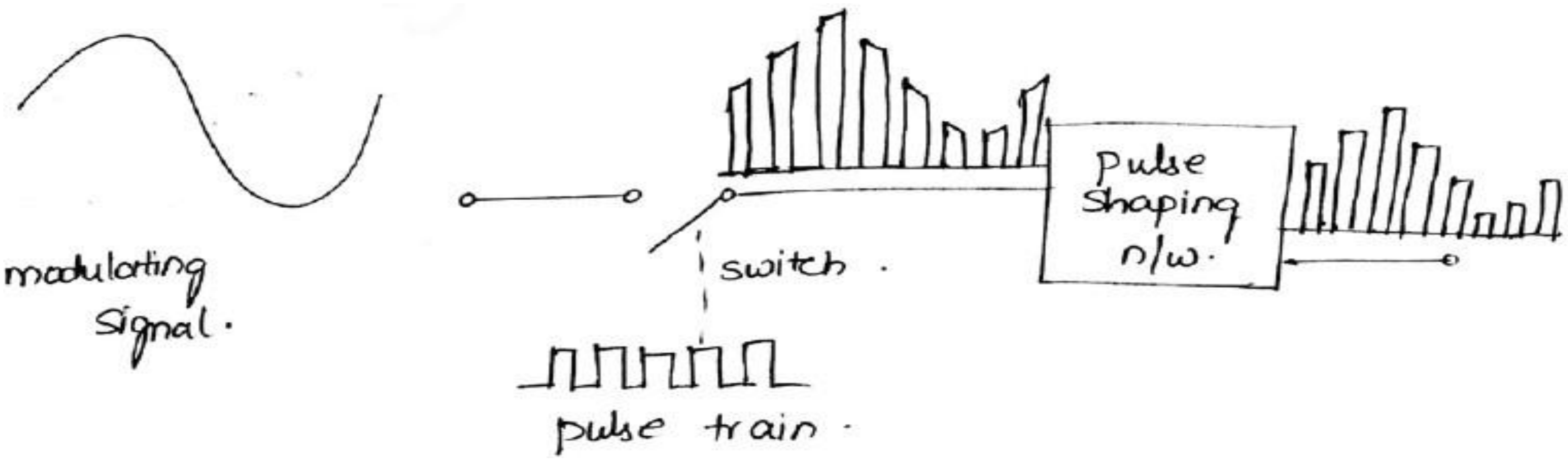
- In **Digital Modulation**, the modulation technique used is **Pulse Code Modulation (PCM)** where the analog signal is converted into digital form of 1s and 0s.
- This is further developed as **Delta Modulation (DM)**,

PULSE AMPLITUDE MODULATION (PAM)

- In PAM, Amplitude of the pulse carrier varies in accordance to the instantaneous amplitude of the message signal.



PAM GENERATION cont...



•The input signal is converted to PAM by the action of switch which is controlled by pulse train.

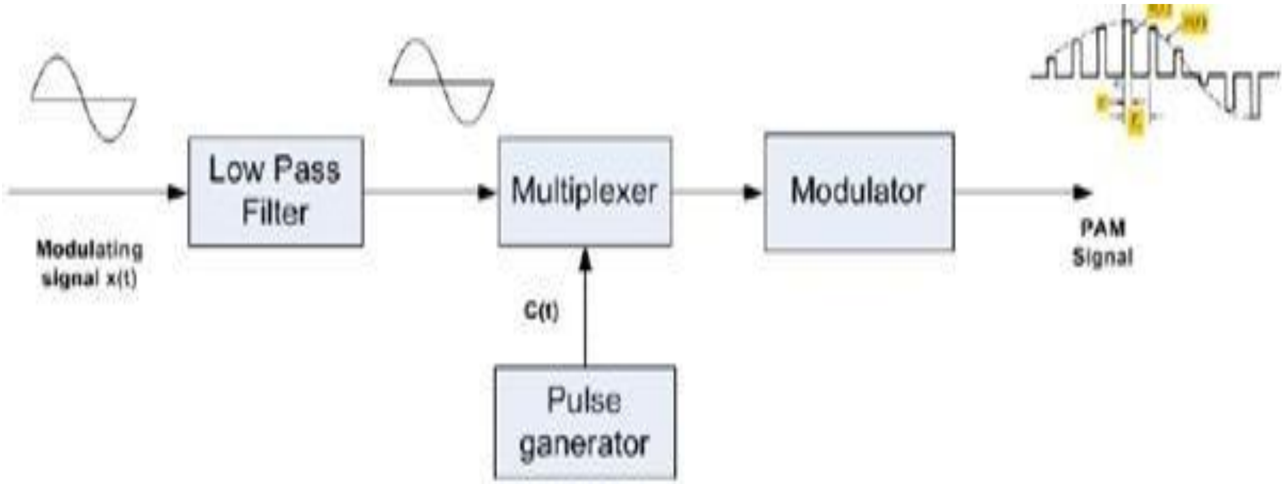
•When pulse is present i.e signal is at high level ,switch is closed and the signal passes through it.

•When pulse is absent i.e signal is at low level ,switch is open ,no signal passes through it.

•With this control action of switch we get PAM waveform at the output of the switch.

•This PAM is passed through pulse shaping n/w for flat tops.

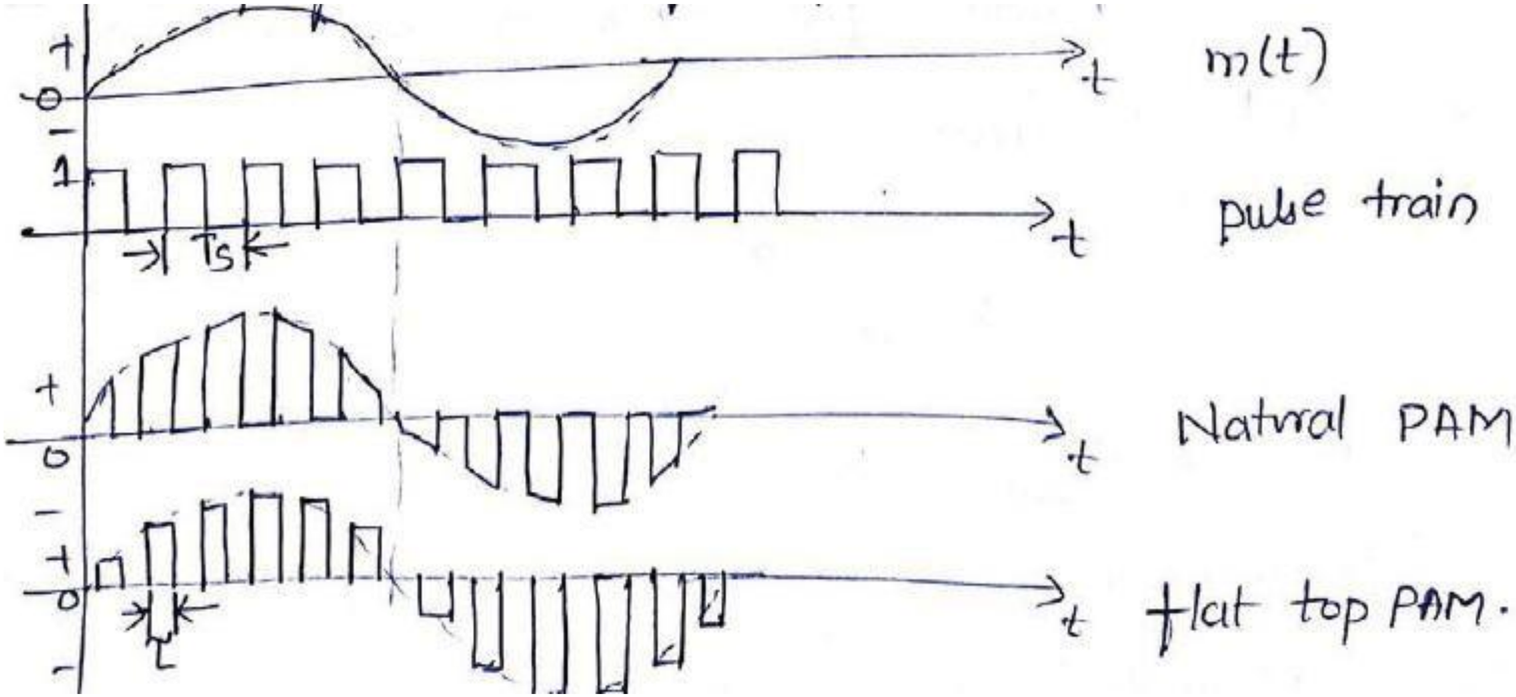
Block schematic of PAM generator



•LPF is used to avoid aliasing effect, during sampling process i.e band limiting is necessary

•LPF removes all frequency components which are higher than f_m , this is known as band limiting .

Different types of PAM signals

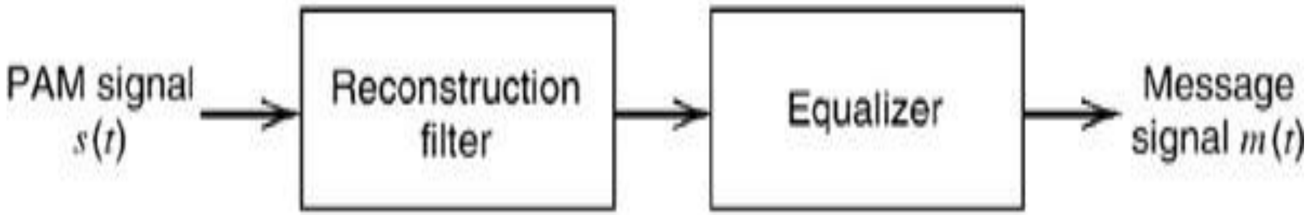


T_S → Time Period Between Two Samples

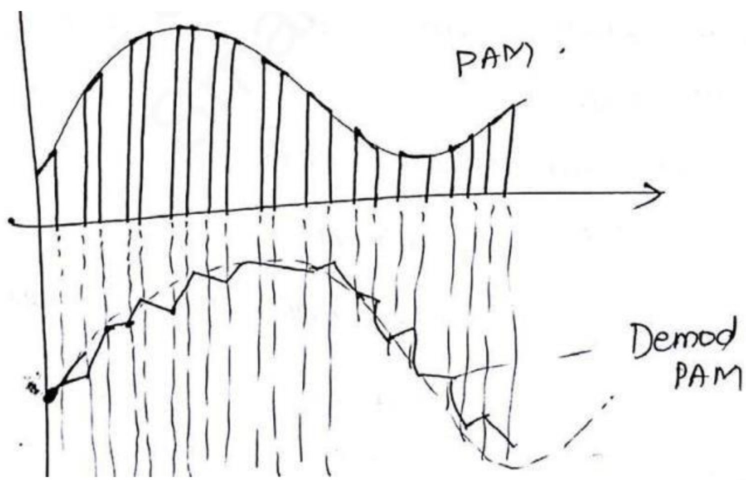
ĩ → Pulse Duration

NOTE: ĩ is very very small compared T_S

PAM Demodulation



•The original Modulating signal can be detected from natural PAM by passing through LPF.
•The LPF with cut-off frequency equal to f_m and it removes high frequency ripples and recovers original modulating signal.
• The Equalizer compensates for aperture effect(Difference b/w input signal and sampled values).



SINGLE POLARITY AND DOUBLE POLARITY PAM

- Single polarity PAM is a situation where a suitable fixed DC bias is added to the signal to ensure that all the pulses are positive.
- Double polarity PAM is a situation where the pulses are both positive and negative.

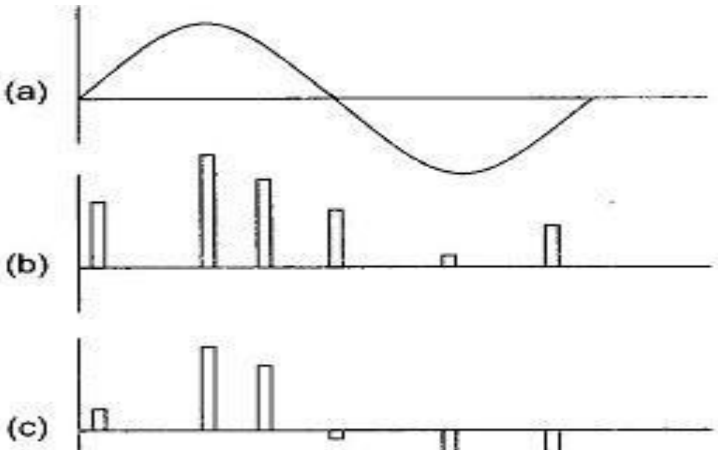
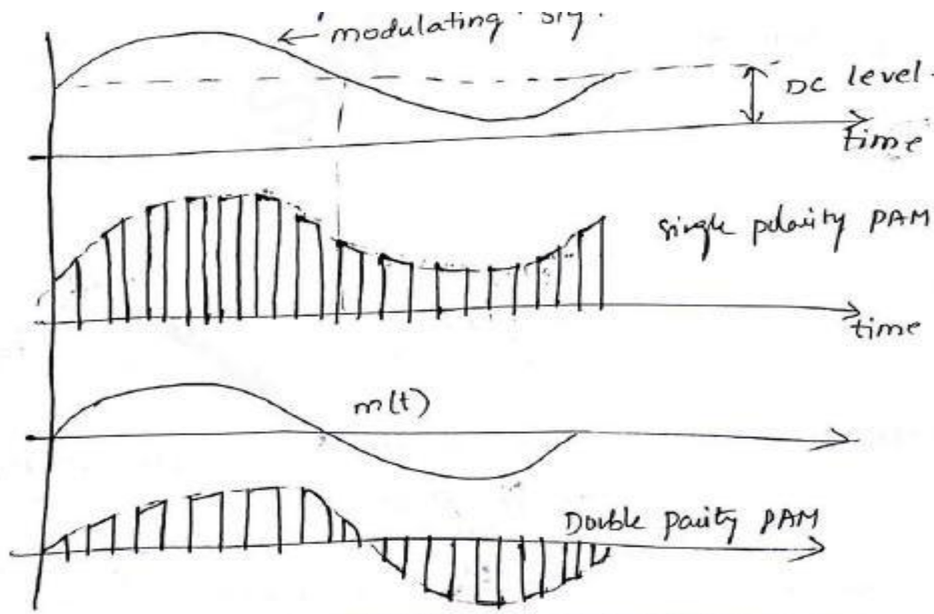


Fig. 18.7

Pulse Amplitude Modualtion (a) Signal
(b) Single Polarity PAM (c) Double Polarity PAM

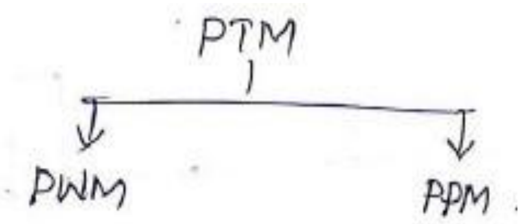


GENERATION OF PWM

- In Pulse Width Modulation (**PWM**) or Pulse Duration Modulation (**PDM**) or Pulse Time Modulation (**PTM**) technique, the **width** or the **duration** or the **time** of the pulse carrier varies in accordance to the instantaneous amplitude of the message signal.

- The width of the pulse varies and the amplitude of the signal remains constant.

- Amplitude limiters are used to make the amplitude of the signal constant. These circuits clip off the amplitude to a desired level, and hence the noise is limited.



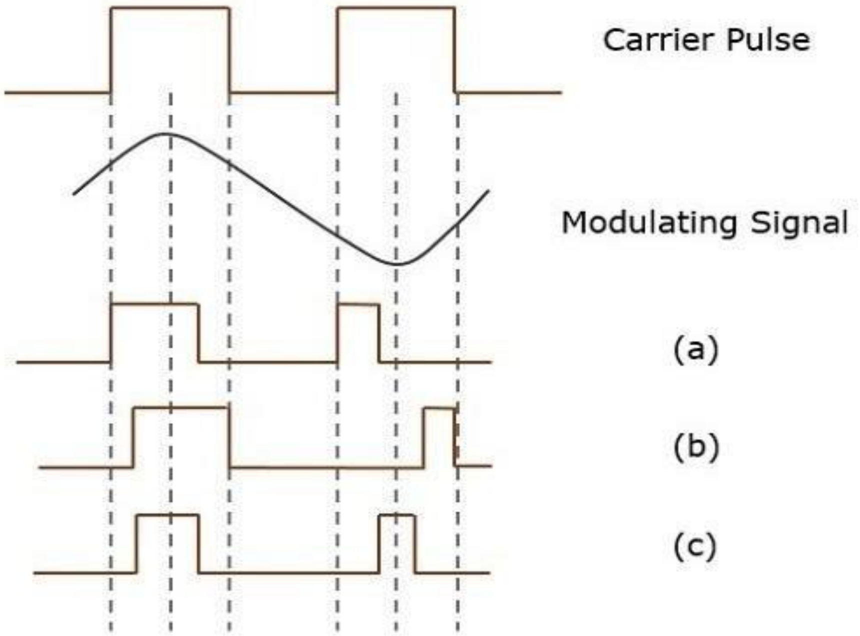
PWM GENERATION cont...

There are three variations of PWM.

a) Leading edge of pulse is held constant and change in pulse width is measured w.r.t leading edge.

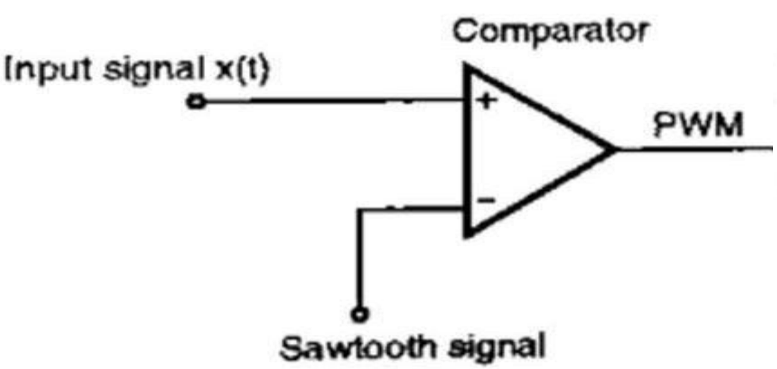
b) Trailing edge is held constant .

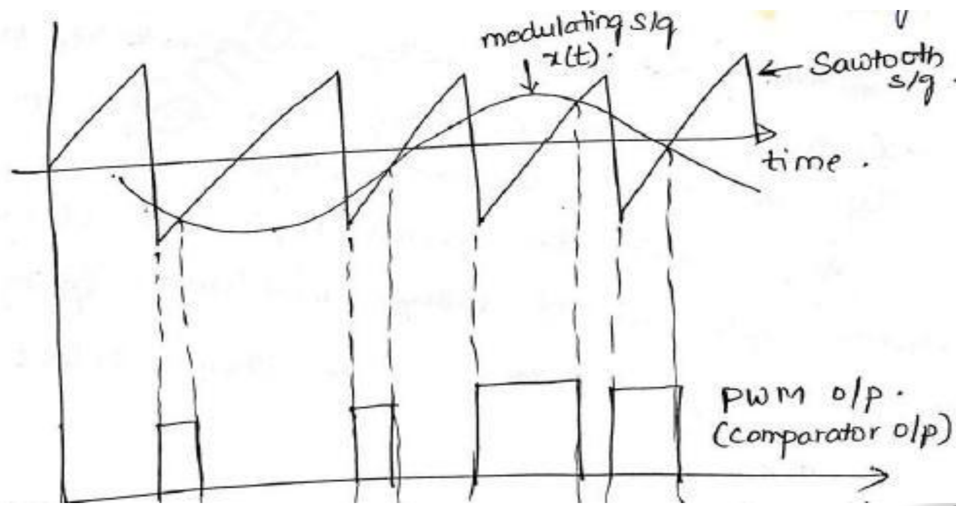
c) Centre of the pulse is held constant and pulse width changes on either side of centre of pulse.



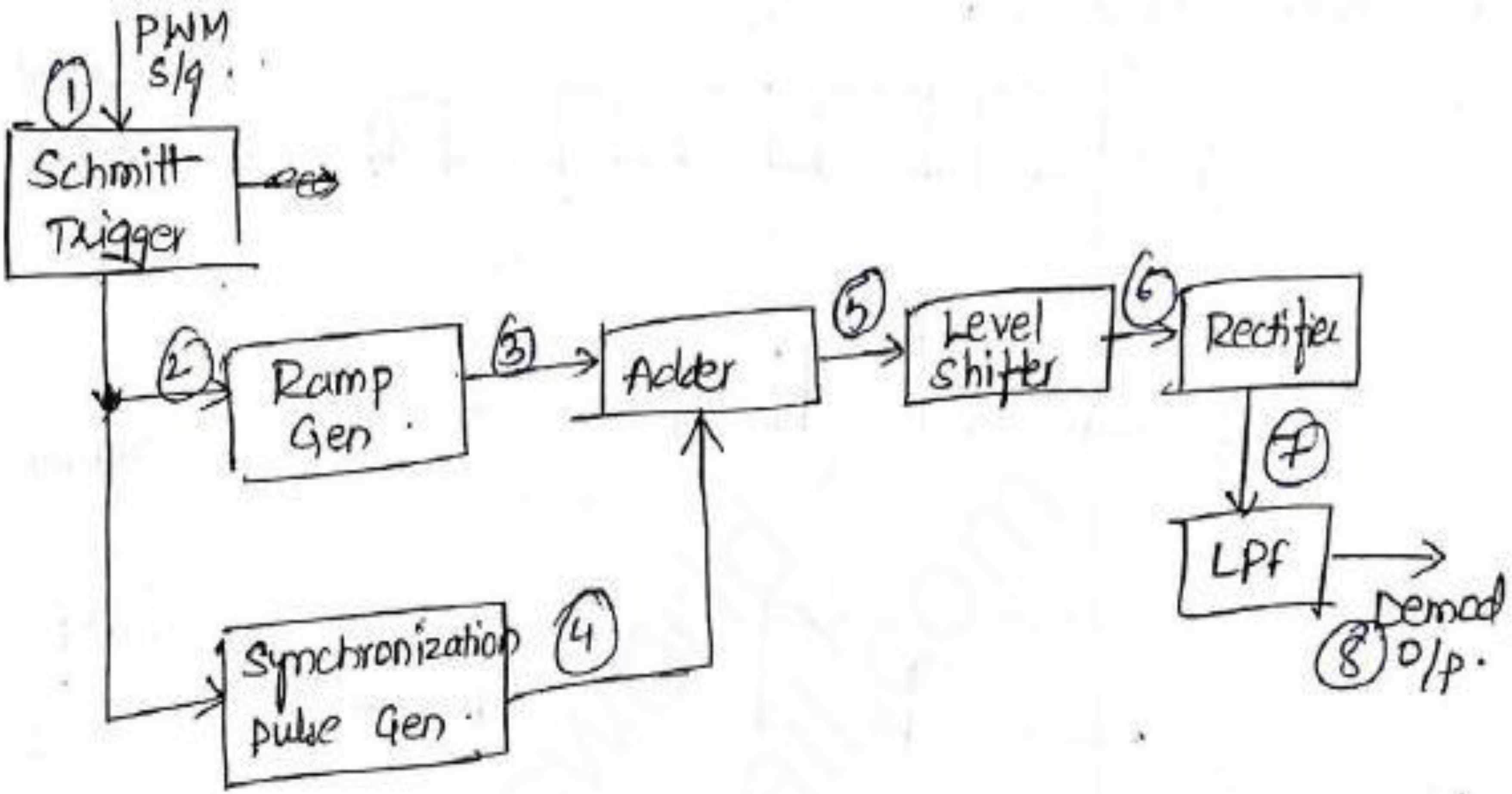
PWM GENERATION cont...

- It consists of sawtooth wave generator, comparator.
- sawtooth generator generates sawtooth signal which is used as sampling signal.
- Comparator compares the amplitude of modulating signal m(t) and amplitude of sawtooth signal.
- Output of the comparator is high as long as the amplitude of m(t) is greater than that of sawtooth signal.
- The duration for which comparator o/p remains high is directly proportional to the m(t).





DEMODULATION OF PWM



DEMODULATION OF PWM cont...

•Schmitt trigger —removes noise in PWM signal and produce square wave pulses.

•Regenerated signal applied to the ramp generator and synchronization pulse generator.

•Ramp generator —produces ramp signal ,the height of the ramps are proportional to the width of the PW pulse.

•The ramp voltage is retained till the next pulse.

•Synchronization pulse generator—produces reference pulses with constant amplitude and width.

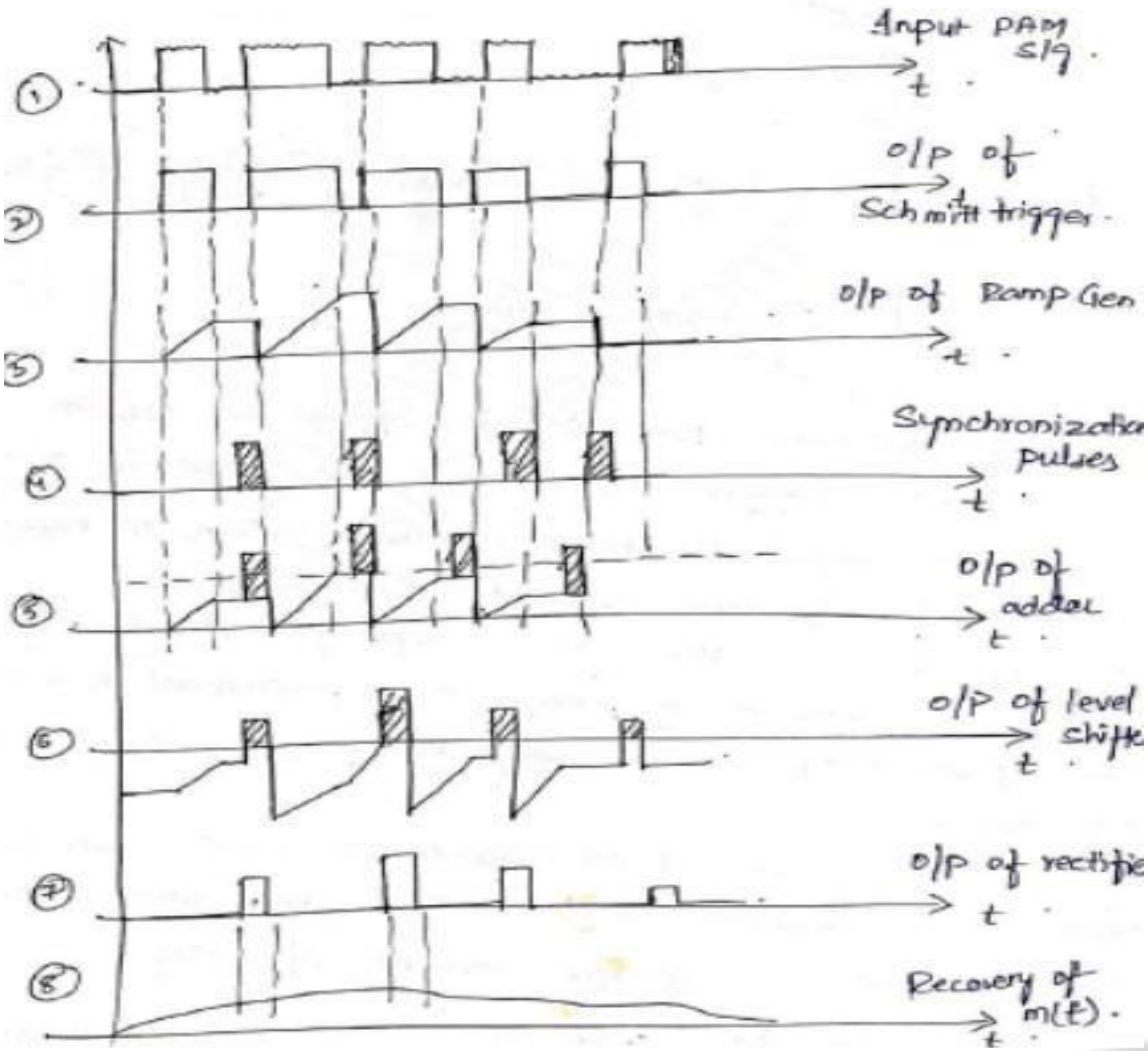
•The input of adder are synchronous pulse and ramp signal.

•Level shifter—here –(ve) offset shift the waveform.

•Rectifier—(-)ve part of waveform is clipped.

•LPF—recovers modulating signal.

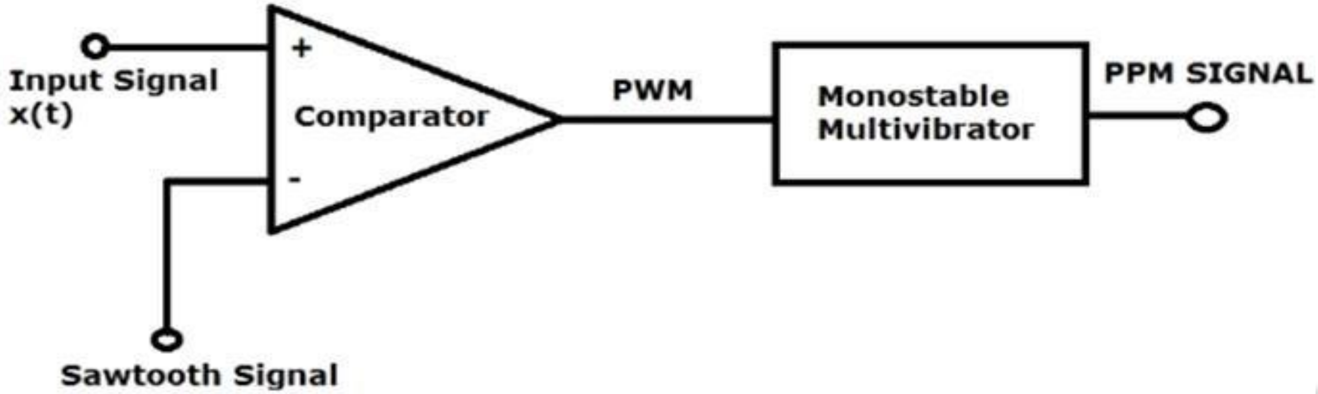
DEMODULATION OF PWM cont...



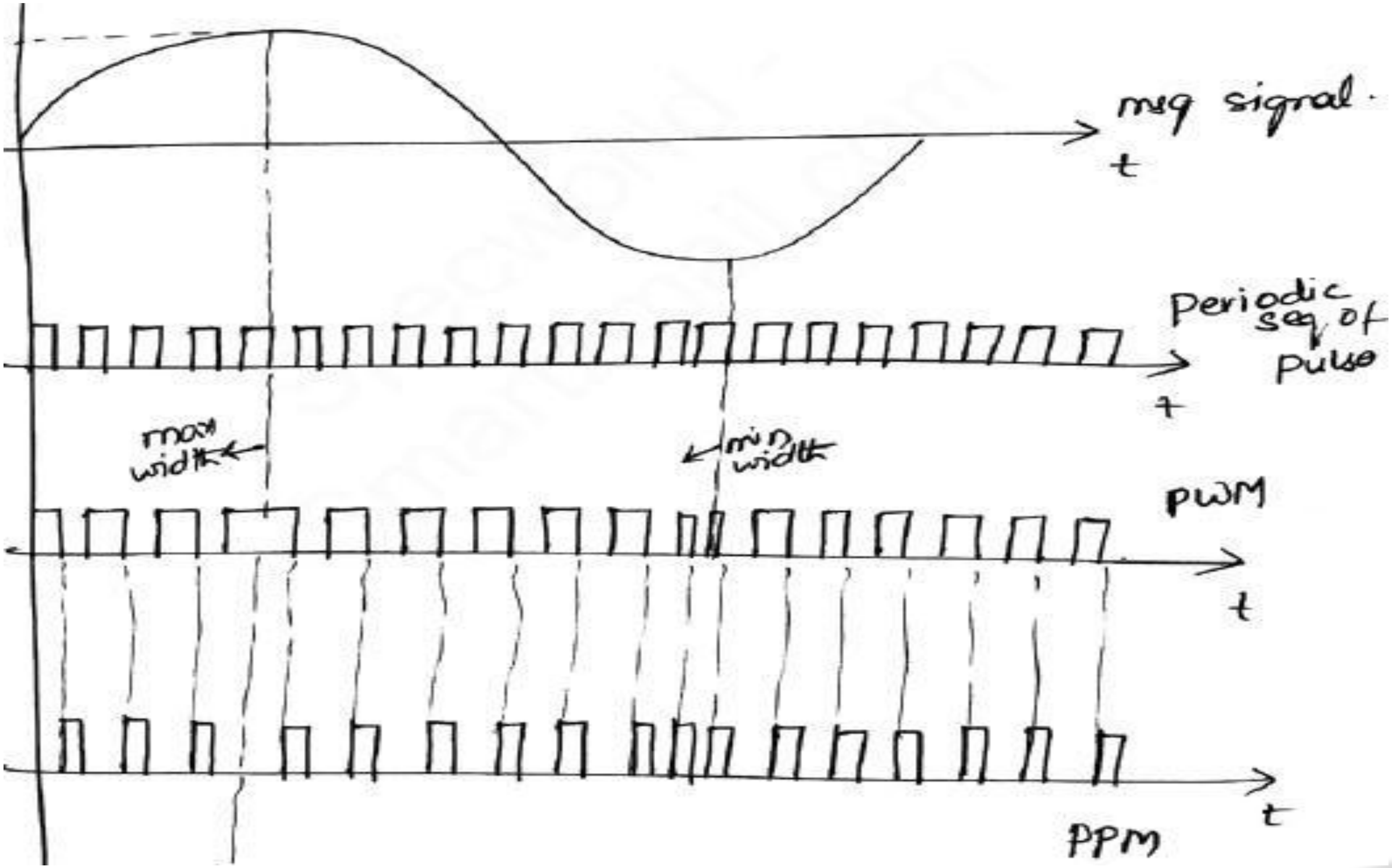
PULSE POSISTION MODULATION(PPM)

- In PPM, Amplitude and width of the pulse are kept constant, while position of each pulse with reference to position of reference pulse is changed according to the m(t).
- PPM is obtained from PWM.
- Each Trailing edge of PWM pulse is a starting point of PPM.
- Position of the pulse is 1:1 proportional to the width of PWM and m(t).

Generation of PPM:

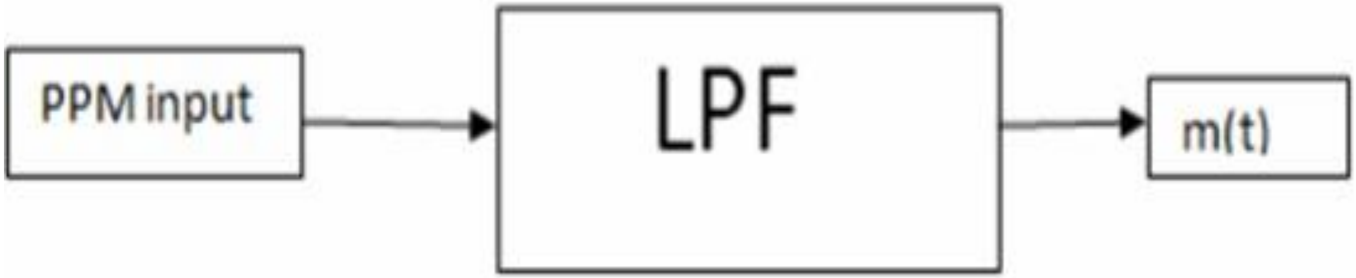


PPM Cont...



msg signal.
t
Periodic seq of pulse
t
max width
min width
PWM
t
t
PPM

PPM Demodulation Cont...



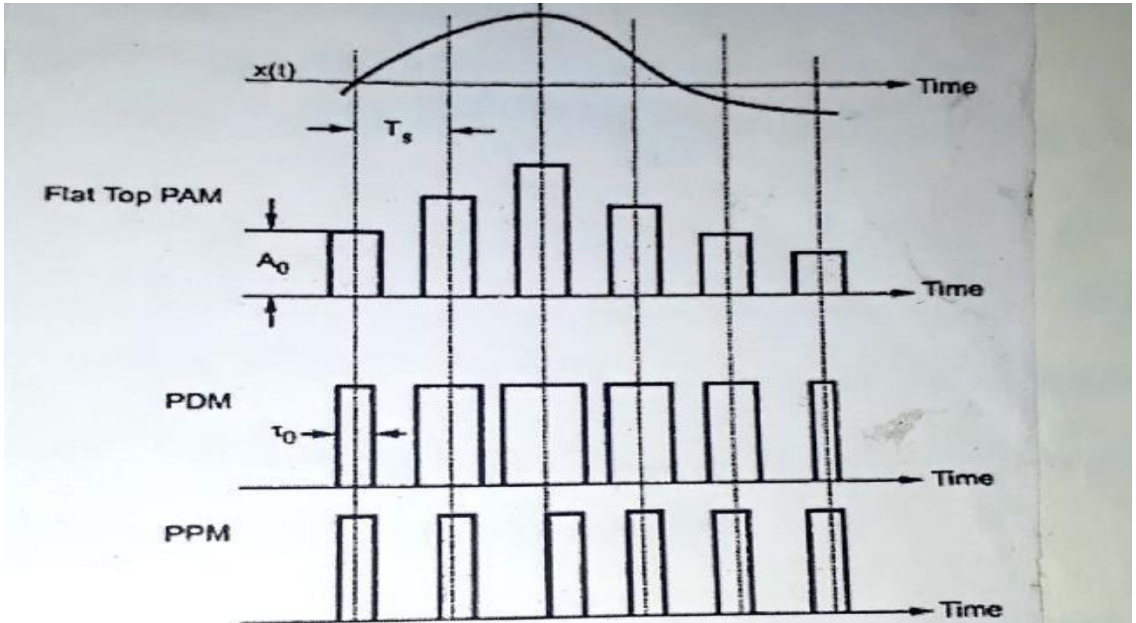
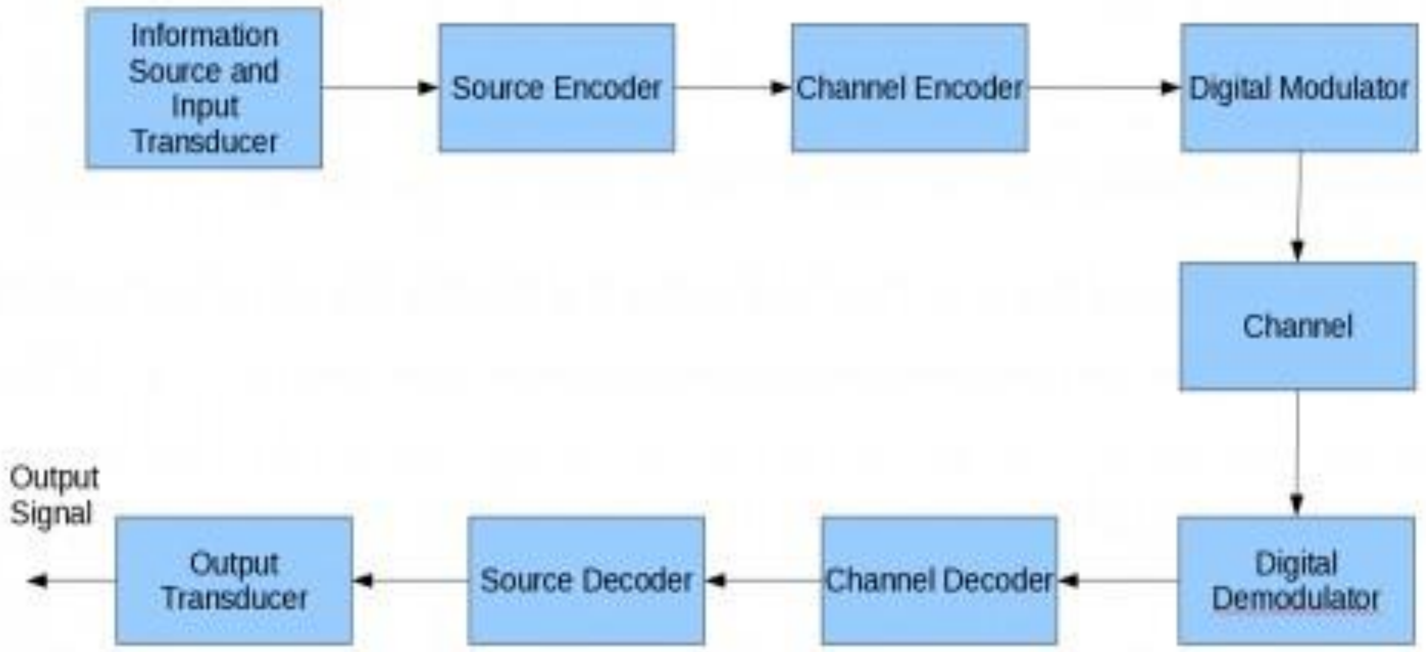


Fig. 1.5.1 Various pulse modulation methods

COMPARE PAM PWM PPM

PAM	PWM	PPM
Amplitude is varied	Width is varied	Position is varied
Bandwidth depends on the width of the pulse	Bandwidth depends on the rise time of the pulse	Bandwidth depends on the rise time of the pulse
Instantaneous transmitter power varies with the amplitude of the pulses	Instantaneous transmitter power varies with the amplitude and width of the pulses	Instantaneous transmitter power remains constant with the width of the pulses
System complexity is high	System complexity is low	System complexity is low
Noise interference is high	Noise interference is low	Noise interference is low
It is similar to amplitude modulation	It is similar to frequency modulation	It is similar to phase modulation

ELEMENTS OF DIGITAL COMMUNICATION SYSTEMS



Basic elements of Digital Communication System

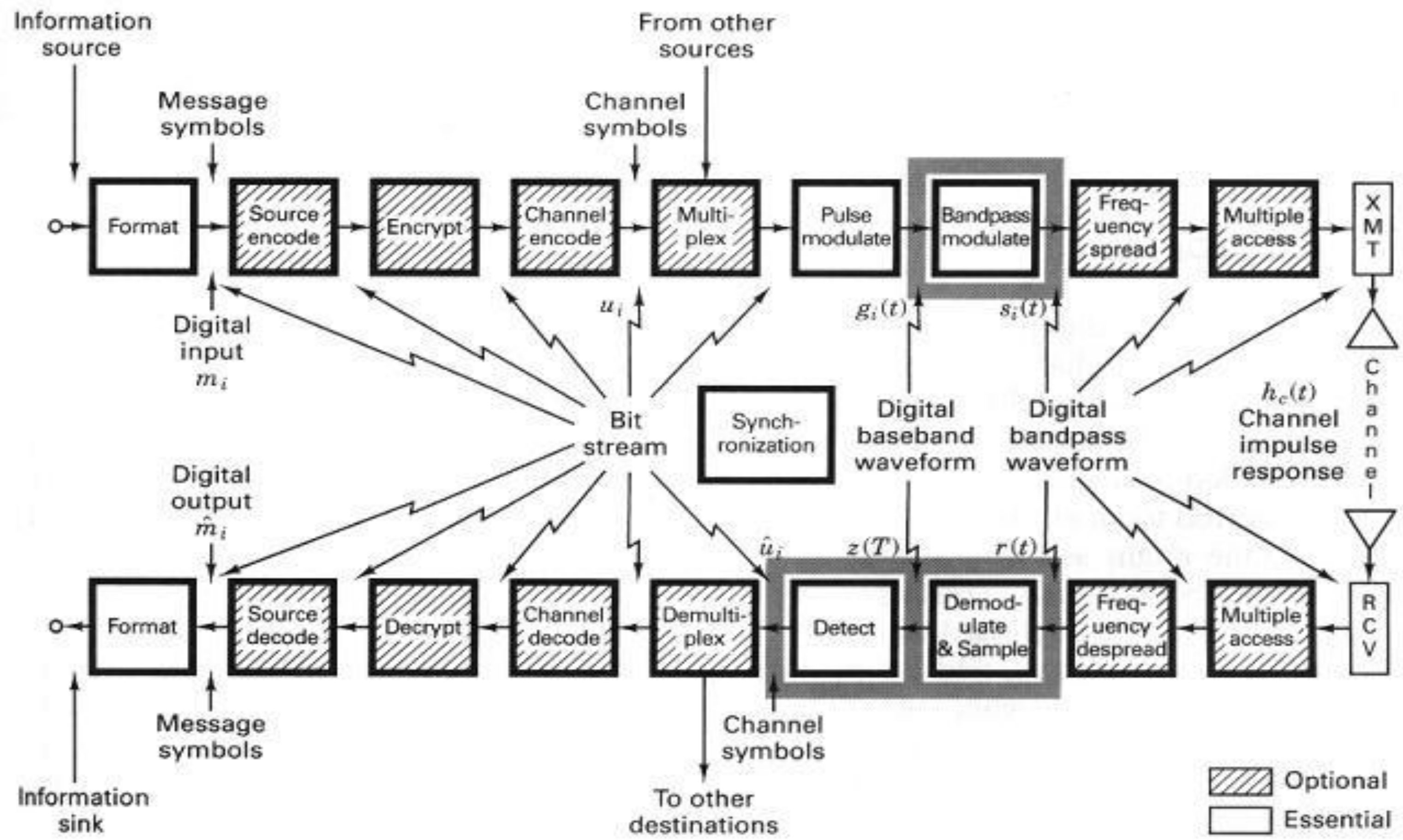
ELEMENTS OF DIGITAL COMMUNICATION SYSTEMS CONT..

Elements Of Digital Communication System are ,

- **Information Source**(Analog or Digital**)**

- **Source Encoder(A/D):** 1.Sampler(C.T→D.T)
  2.Quantizer(C.A→D.A)
  3.Eancoder(0,1)
- **Channel Encoder:** Adds Extra bit(security)
- **Modulator**: Modulation (ASK,FSK,PSK)
- **Channel:** Medium
- **Demodulator:** Demodulation
- **Channel decoder:** Removes extra bit, Error correction and Detection
- **Source Decoder:** (D/A)
- **Destination: Receiving end**

Functional block diagram of digital communication systems



ADVANTAGES OF DIGITAL COMMUNICATION SYSTEMS

- Signal processing functions such as encryption and compression are employed in digital circuits to maintain the secrecy of the information.
- The probability of error occurrence is reduced by employing error detecting and error correcting codes.
- Spread spectrum technique is used to avoid signal jamming
- Combining digital signals using Time Division Multiplexing (TDM) is easier than combining analog signals using Frequency Division Multiplexing (FDM).
- The configuring process of digital signals is easier than analog signals.
- The capacity of the channel is effectively utilized by digital signals.

DISADVANTAGES OF DIGITAL COMMUNICATION SYSTEM

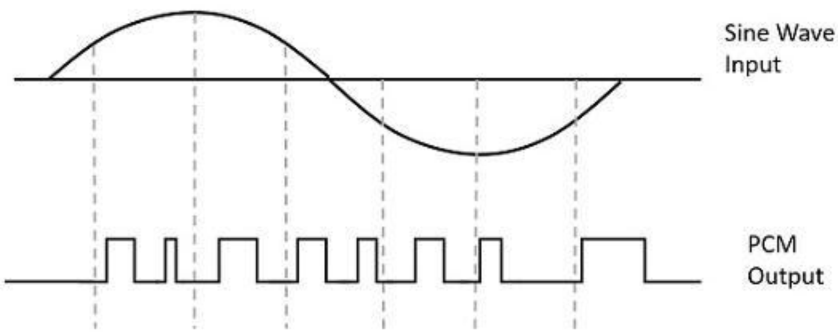
- Generally, more bandwidth is required than that for analog
- High power consumption.
- Complex circuit, more sofisticated device making is also drawbacks of digital system.
- Introduce sampling error.
- As square wave is more affected by noise, That’s why while communicating through channel we send sine waves but while operating on device we use square pulses.

APPLICATIONS OF DIGITAL COMMUNICATION SYSTEMS

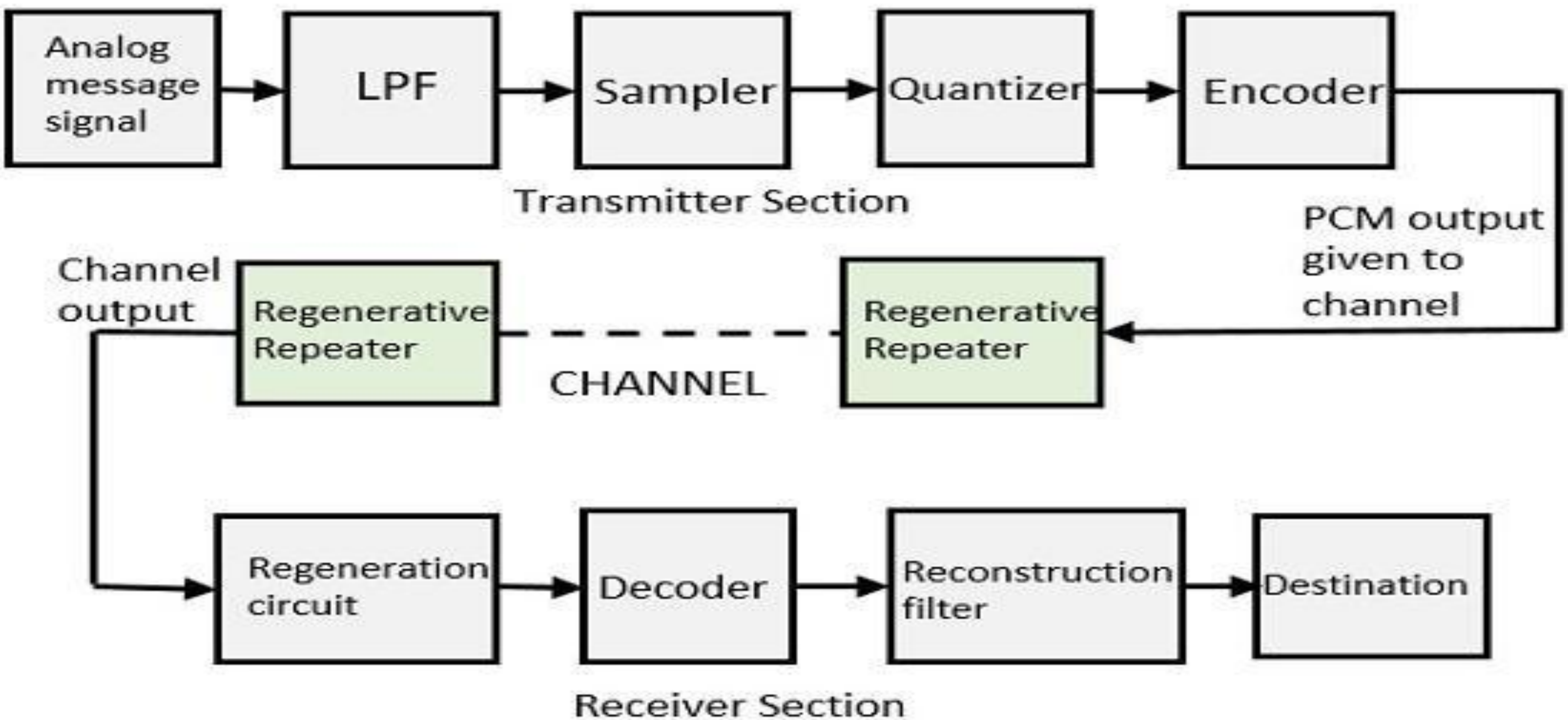
- It is used in military application for secure communication and missile guidance.
- It is used in image processing for pattern recognition, robotic vision and image enhancement.
- The digital communication systems used in telephony for text messaging, etc
- It is used in video compression.
- It is used in digital audio transmission.
- It is used in data compression.

PULSE CODE MODULATION(PCM)

- The system in which sampled and quantized values of analog signals are transmitted via a sequence of **codeword's** is called pulse code modulation(PCM).
- The following figure shows an example of PCM output with respect to instantaneous values of a given sine wave.



Elements of PCM



Elements of PCM cont…

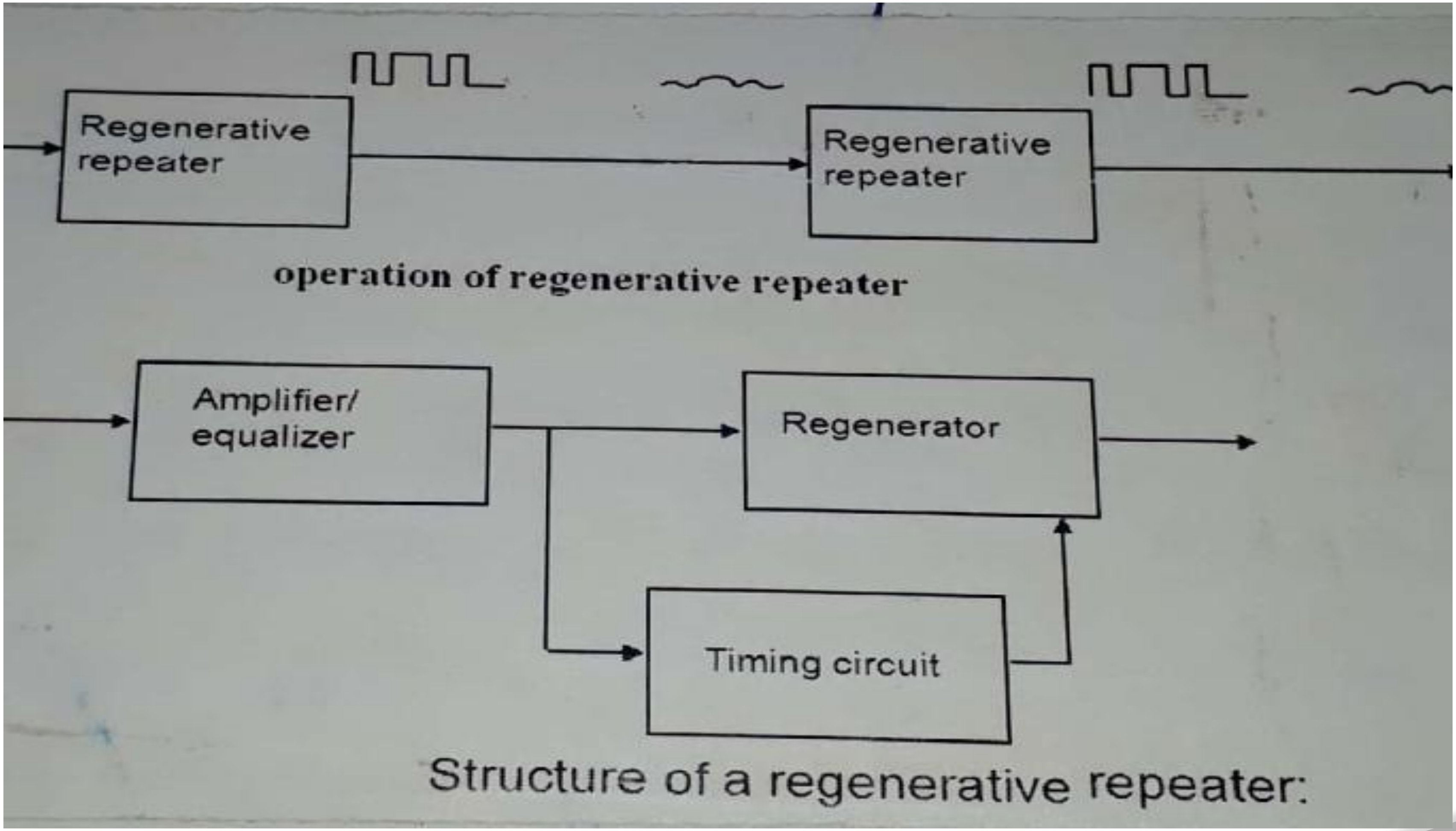
A PCM system can be divided into 3 sections

1. Transmitter: **Sampler,quntizer, Encoder**

2. Transmition path: **Regenerative Repeaters**(controlling the effect of distortion and noise produced by channel)

- ❖ Regenerative Repeaters produces,
  - ❖ Timing (periodic sequences between pulses)
  - ❖ Equalization (shapes the received pulses)
  - ❖ Decision making(takes the decision if pulses are having distortion or not)

Regenerative Repeaters



3.Receiver: decoder, reconstruction filter, destination.

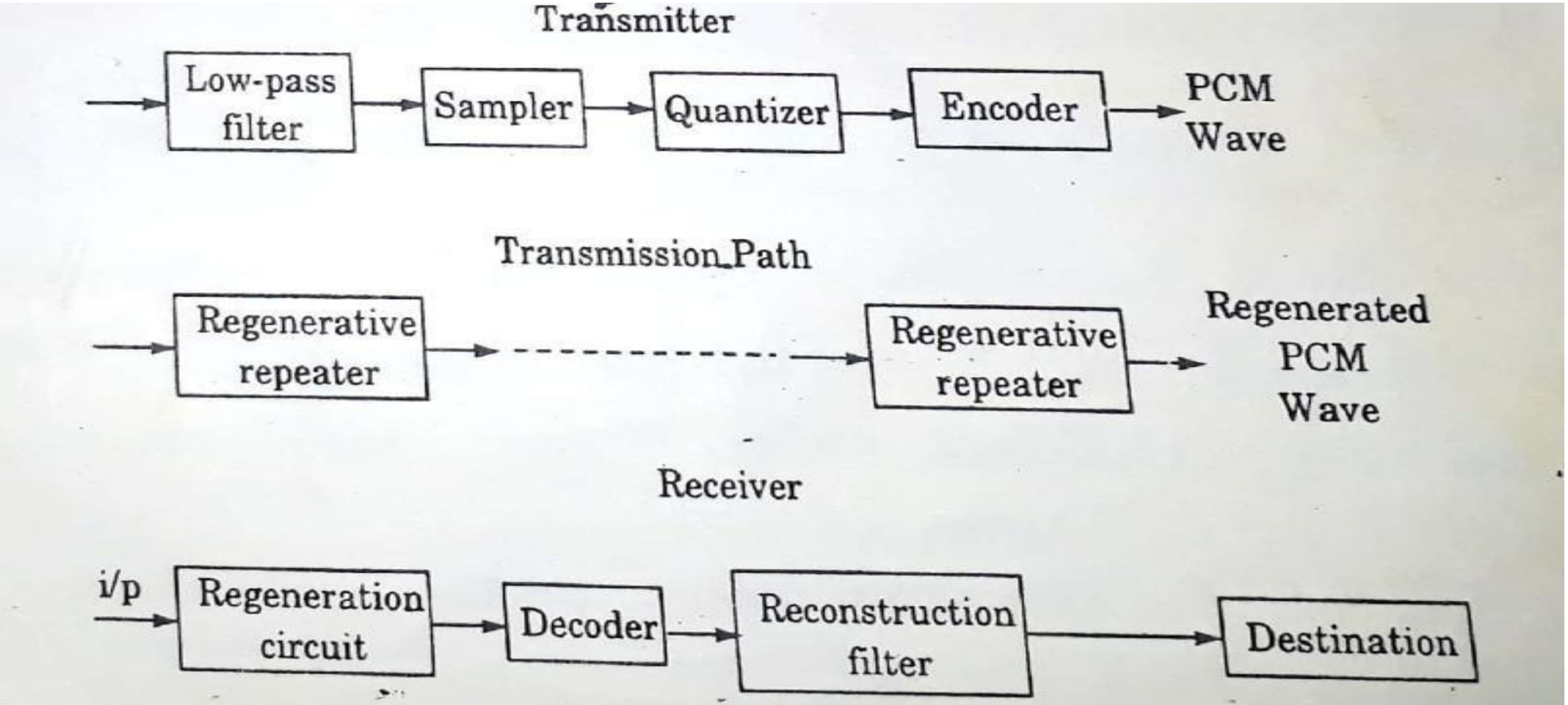
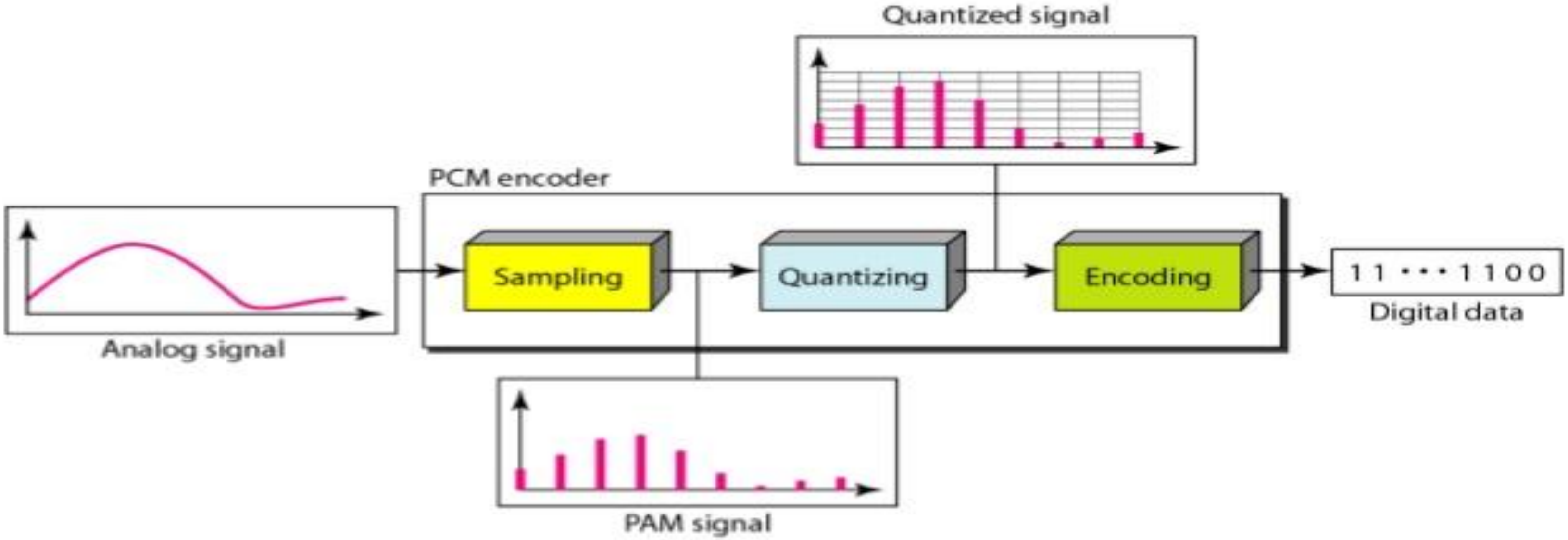


Fig. 4.1 Basic Elements of a PCM system

Elements of PCM cont…

BLOCK DIAGRAM OF PCM



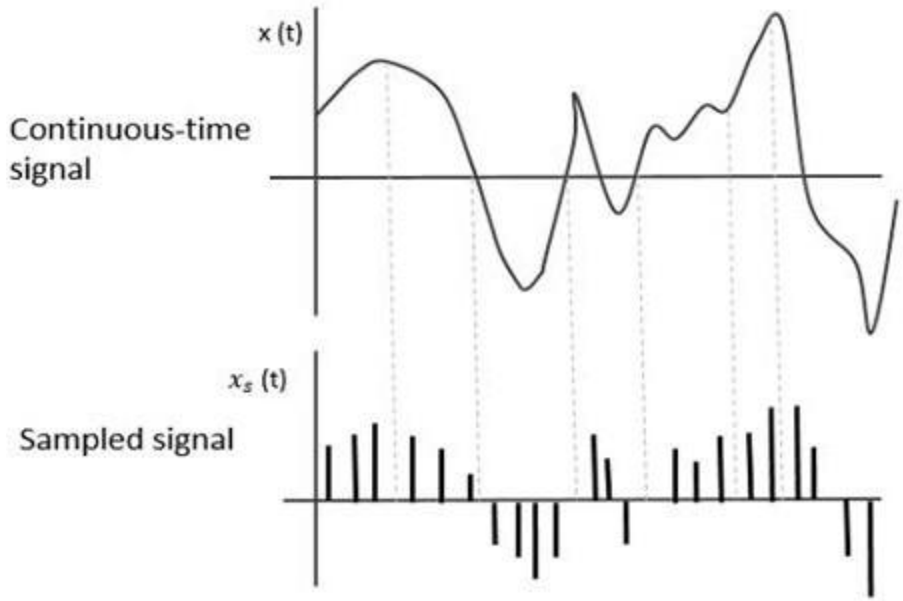
SAMPLING

❖**Sampling** is “the process of converting continuous time signal into discrete time signal”.

❖**Sampling** is defined as, “The process of measuring the instantaneous values of continuous-time signal in a discrete form.”

❖**Sampl**e is a piece of data taken from the whole data which is continuous in the time domain.

❖The following figure indicates a continuous-time signal x (t) and a sampled signal x_s (t). When x (t) is multiplied by a periodic impulse train, the sampled signal x_s (t) is obtained.



❖The **sampling period** is the time difference between two consecutive samples, It is the inverse of the sampling frequency

$$Sampling\ Frequency = \frac{1}{T_s} = f_s$$

Where,

- T_s is the sampling time
- f_s is the sampling frequency or the sampling rate

SAMPLING Cont…

Nyquist Rate: $f_S = 2W$

Where,

- f_S is the sampling rate
- **W** is the highest frequency

This rate of sampling is called as **Nyquist rate**.

Sampling Theorem:

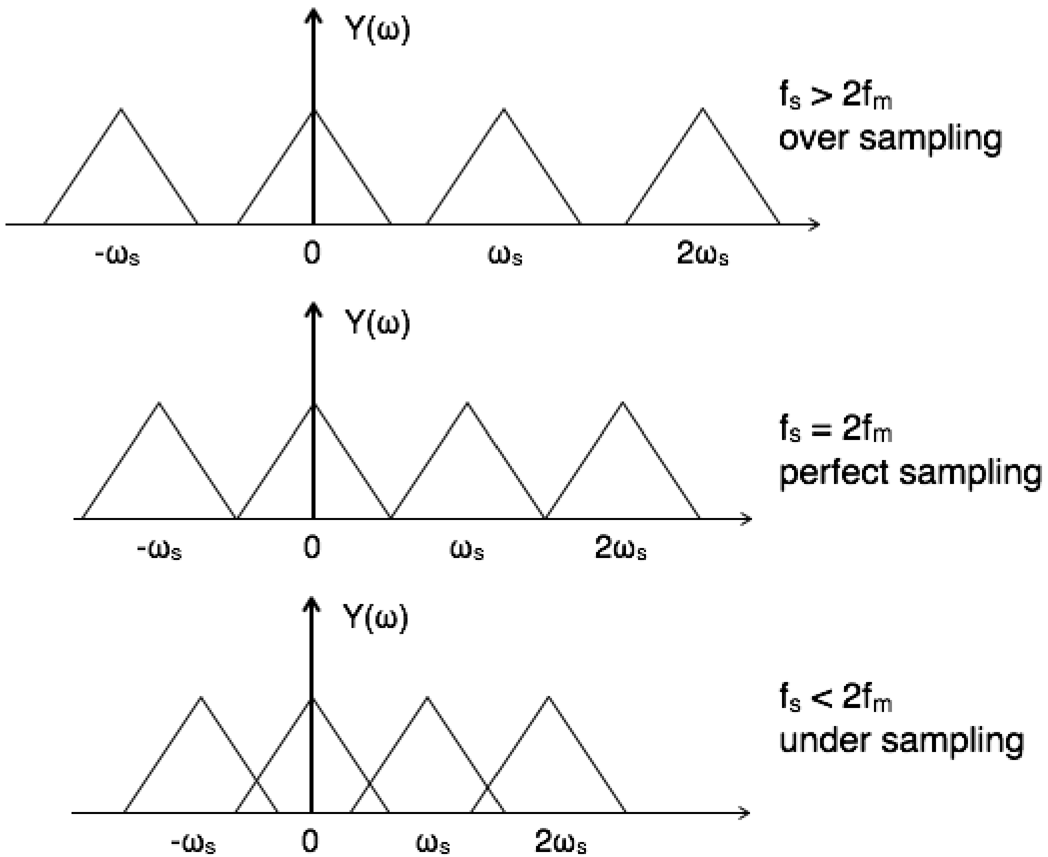
statement: A continuous time signal can be represented in its samples and can be recovered back when sampling frequency f_s is greater than or equal to the twice the highest frequency component of message signal. i. e.

$$f_s \geq 2f_m.$$

Aliasing effect: Aliasing is the phenomenon in which higher frequency components are combined with lower frequency components in spectrum of its sampled version. $f_s < 2f_m$

- By using anti-aliasing filters.

SAMPLING Cont...



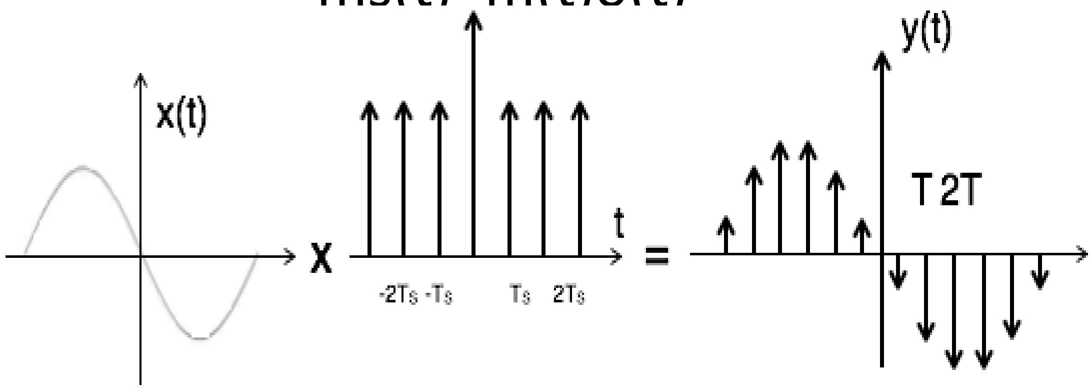
TYPES OF SAMPLING:

There are 3 sampling techniques,They are:

1.Ideal sampling or impulse sampling or instantaneous sampling:

- The instantaneous sampling has a train of impulses.
- The pulse width of the samples has almost zero value.
- It is the product of message signal m(t) with a unit impulse train δ(t) gives sampled signal.

i.e., $ms(t)=m(t)\delta(t)$

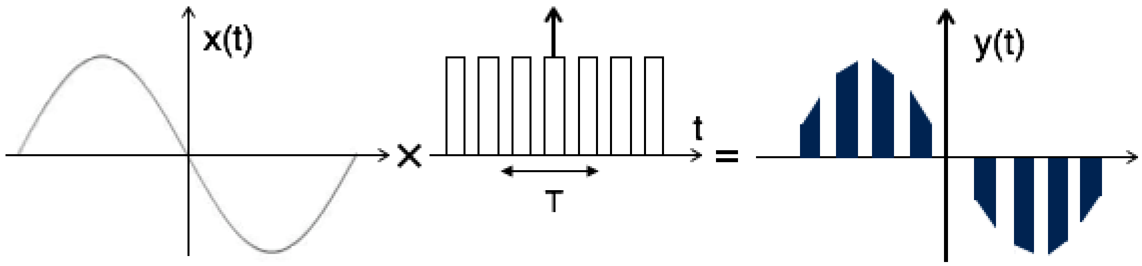


2.Natural sampling:

❖The natural sampling is one which can be represented with respect to amplitude of the analog signal.

❖It is the product of message signal m(t) with a pulse train δ(t) gives sampled signal. i.e.,

$$Xns(t)=m(t).Xp(t)$$

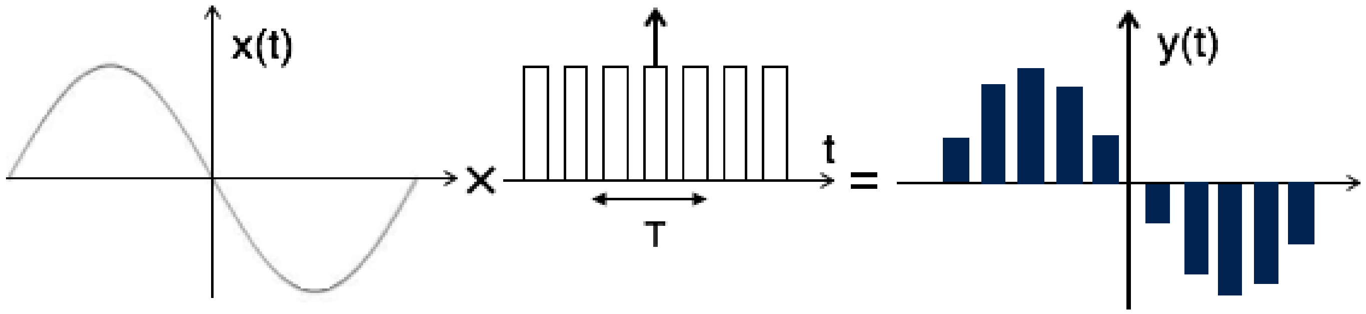


3.Flat -top sampling or practical sampling:

❖It is a product of message signal m(t) with aflat-top pulse train δ(t) gives sampled signal.

❖During transmission, noise is introduced at top of the transmission pulse which can be easily removed if the pulse is in the form of flat top.

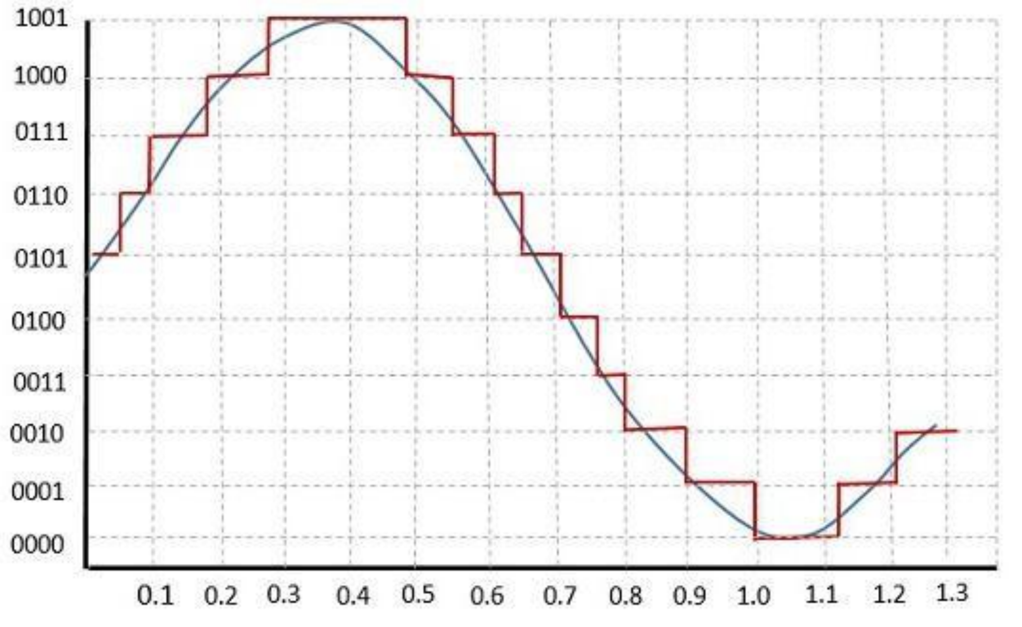
❖Here, the top of the samples are flat i.e. they have constant amplitude. Hence, it is called as flat top sampling or practical sampling.



Quantization

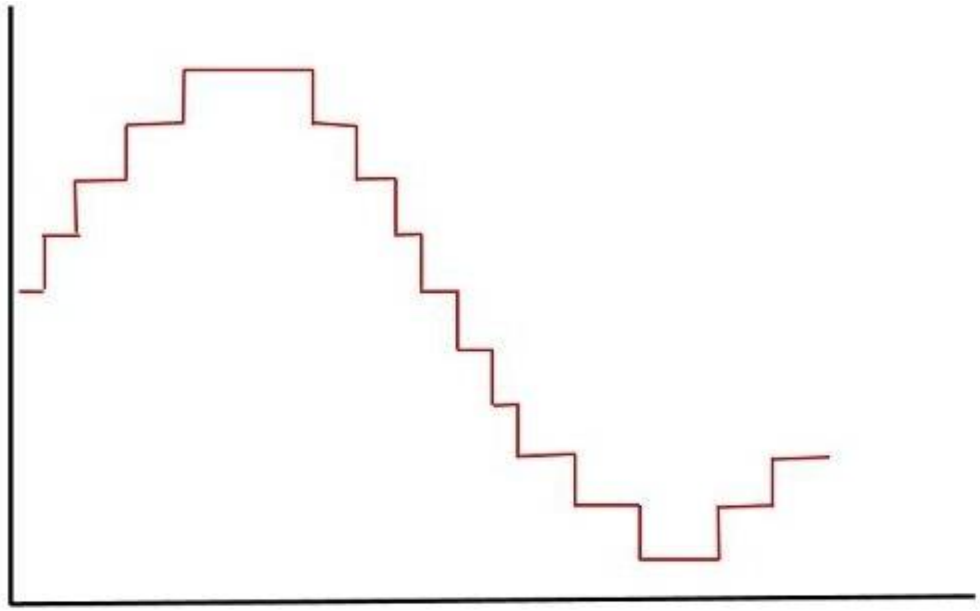
❖**Quantization** is the process of representing the sampled values of the amplitude by a finite set of levels, which means **converting a continuous-amplitude sample into a discrete-Amplitue signal.**

❖The following figure shows how an analog signal gets quantized.



Quantization cont...

The following figure shows the resultant quantized signal which is the digital form for the given analog signal.



This is also called as **Stair-case** waveform, in accordance with its shape

Types of Quantization

There are two types of Quantization ,

1.Uniform Quantization : The type of quantization in which the quantization levels are uniformly spaced is termed as a **Uniform Quantization**.

2.Non-uniform Quantization: The type of quantization in which the quantization levels are unequal is termed as a **Non-uniform Quantization**.

•There are two types of **uniform** quantization, There are two types of uniform quantization. They are,

- Mid-Rise type
- Mid-Tread type.

Uniform Quantization Types

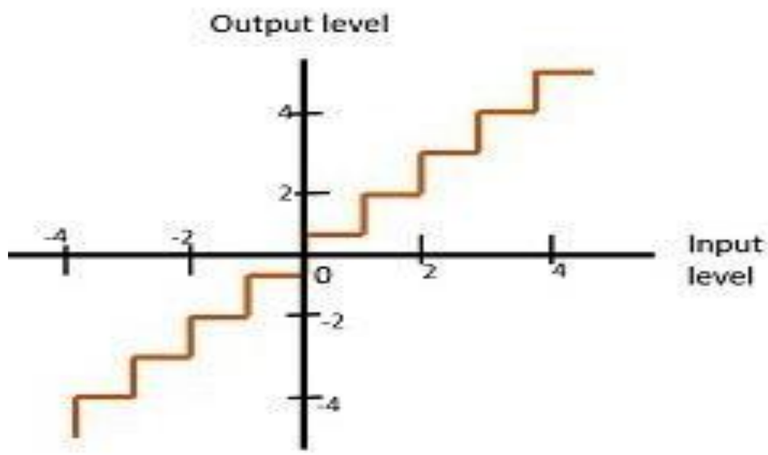


Fig 1 : Mid-Rise type Uniform Quantization

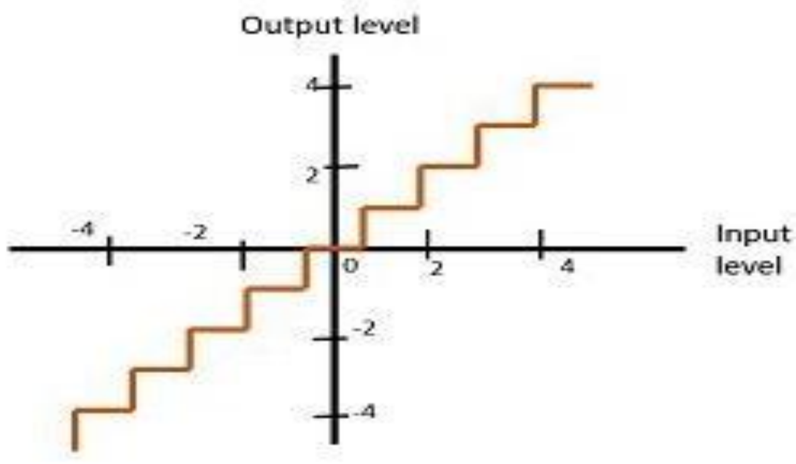


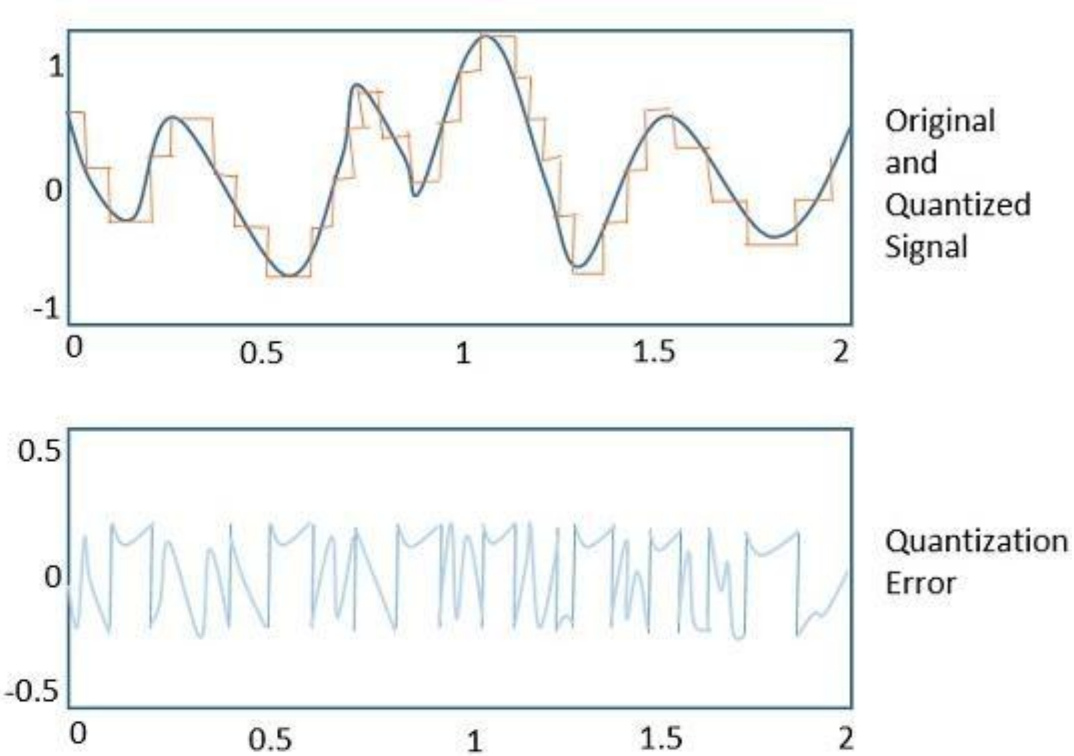
Fig 2 : Mid-Tread type Uniform Quantization

- **Mid-Rise** : the origin lies in the middle of a raising part of the stair-case. The quantization levels in this type are even in number.
- **Mid-tread** :the origin lies in the middle of a tread of the stair-case. The quantization levels in this type are odd in number

Quantization Error

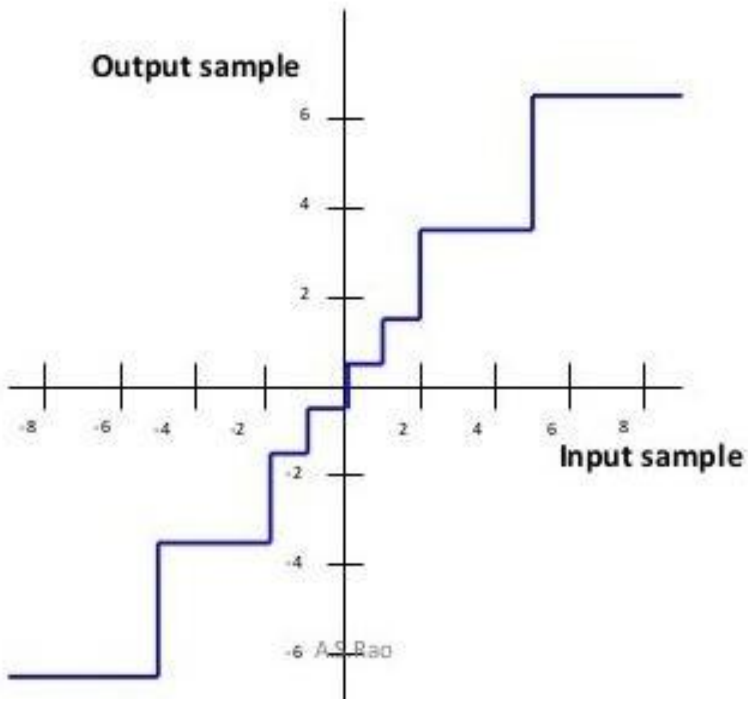
The difference between an input value and its quantized value is called a **Quantization Error**.

$$\mathcal{E} = X_q(nT_s) - X(nT_s)$$



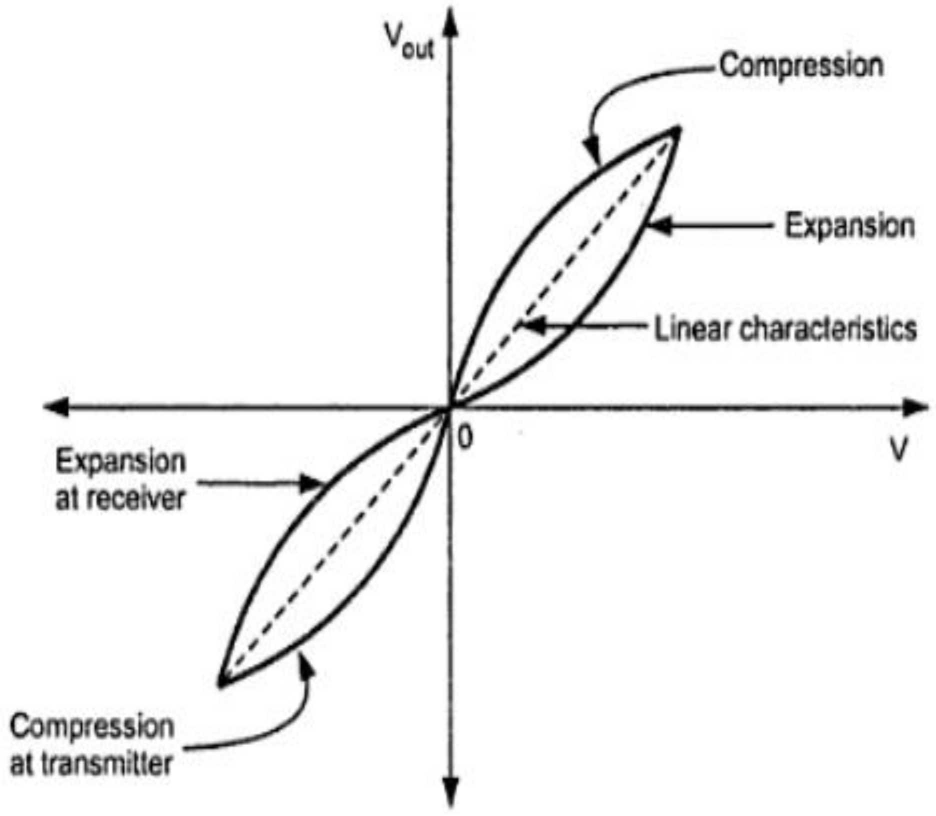
Non Uniform Quantization…

- Variable step size



COMPANDING

- The word **Companding** is a combination of Compressing and Expanding



COMPANDING Cont..

•The signal is amplified at low level, attenuated at high level to get original signal as output.
•More step size at low level and small step size at high level signals .
•At receiver reverse process is done to get original signal

❖There are two types of Companding techniques. They are ,

A-law Companding Technique:(piecewise compressor)

μ-law Companding Technique:(speech signals ,music signals)

COMPANDING Cont..

A-law Companding Technique: liner segments for low level signals ,logarithmic segments for high level

$$Z(x) = \begin{cases} \dfrac{A|x|}{1+ln A} & for\ 0 \le |x| \le \dfrac{1}{A} \\ \dfrac{1+ln(A|x|)}{1+ln A} & for\ \dfrac{1}{A} \le |x| \le 1 \end{cases}$$

When A=1 , uniform quantization, practically A=87.56

μ-law Companding Technique:

$$Z(x) = (\text{Sgn } x) \frac{ln(1+\mu|x|)}{ln(1+\mu)} \quad |x| \le 1$$

Where ,

Z(x)= Output of compressor

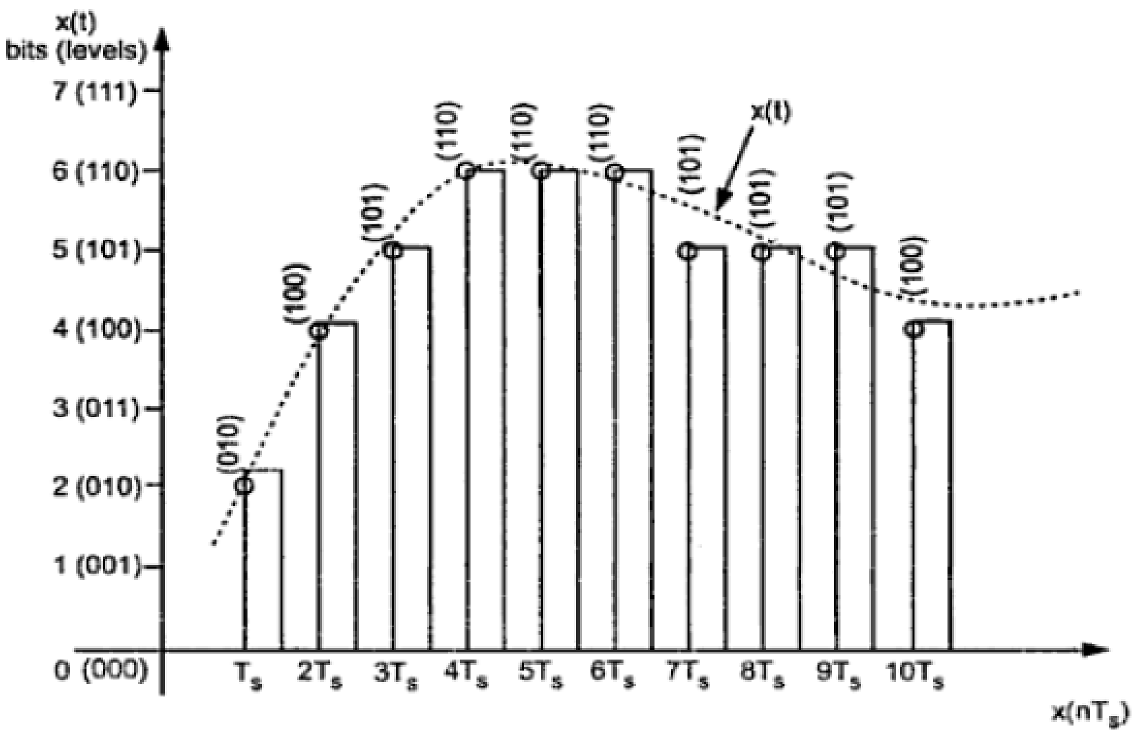
x= input of the compressor

Sgn(x)= positive and negative values of i/p & o/p i.e ± 1

|X|/ Xmax= normalised value of i/p

REDUNDANCY

- The signal's value from the present sample to next sample does not differ by a large amount.
- The adjacent/different samples of the signal carry the same information with a small difference



DIFFERENTIAL PCM

•If the redundancy is reduced, then the overall bitrate will decrease and the number of bits required to transmit one sample will also reduce.

•This type of digital pulse modulation technique is called differential pulse code modulation.

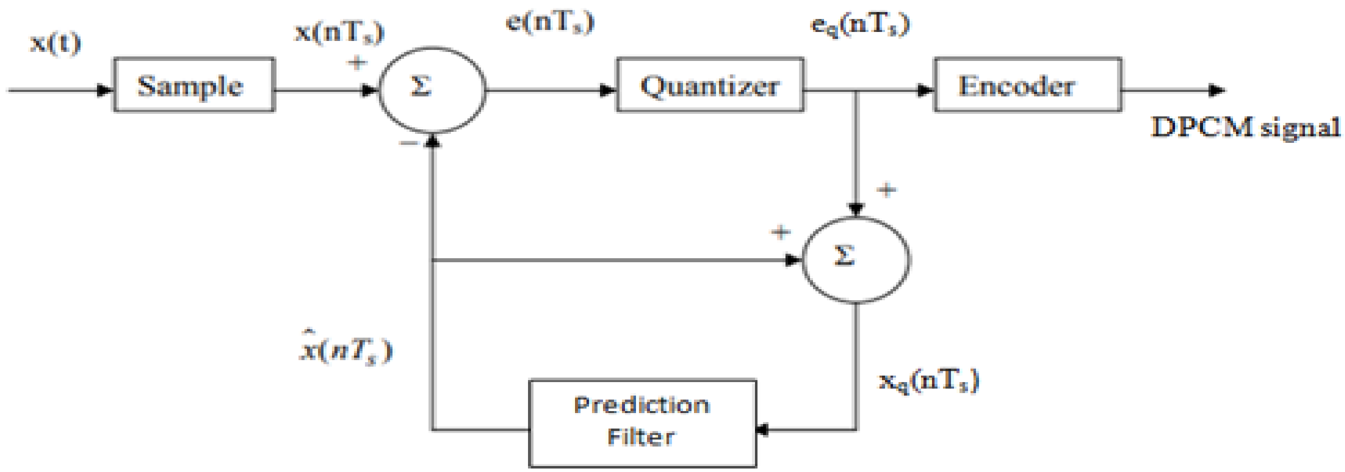


Fig: DPCM transmitter

DIFFERENTIAL PCM cont..

- The sampled signal is denoted by x(nTs) , the predicted signal is indicated by x^(nTs) and signal error is denoted by e(nTs).
- The comparator finds out the difference between the actual sample value x(nTs) and the predicted value x^(nTs).

i.e., e(nTs)= x(nTs)- x^(nTs) …….(1)

- The quantized error signal eq(nTs) is very small and can be encoded by using a small number of bits.
- The quantizer output would be written as,

eq(nTs)= e(nTs)+ q(nTs) ……(2)

[Here q(nTs) is quantization error]

- From the above block diagram the prediction filter input xq(nTs) is obtained by sum of x^(nTs) and the quantizer output eq(nTs).

i.e, xq(nTs) = x^(nTs)+ eq(nTs)………. (3)

DIFFERENTIAL PCM cont..

- by substituting the value of eq(nTs) from the equation (2) in equation (3) we get,

$$xq(nTs) = x\hat{}(nTs) + e(nTs) + q(nTs) \ldots\ldots (4)$$

Equation (1) can written as,

$$e(nTs) + x\hat{}(nTs) = x(nTs) \ldots\ldots (5)$$

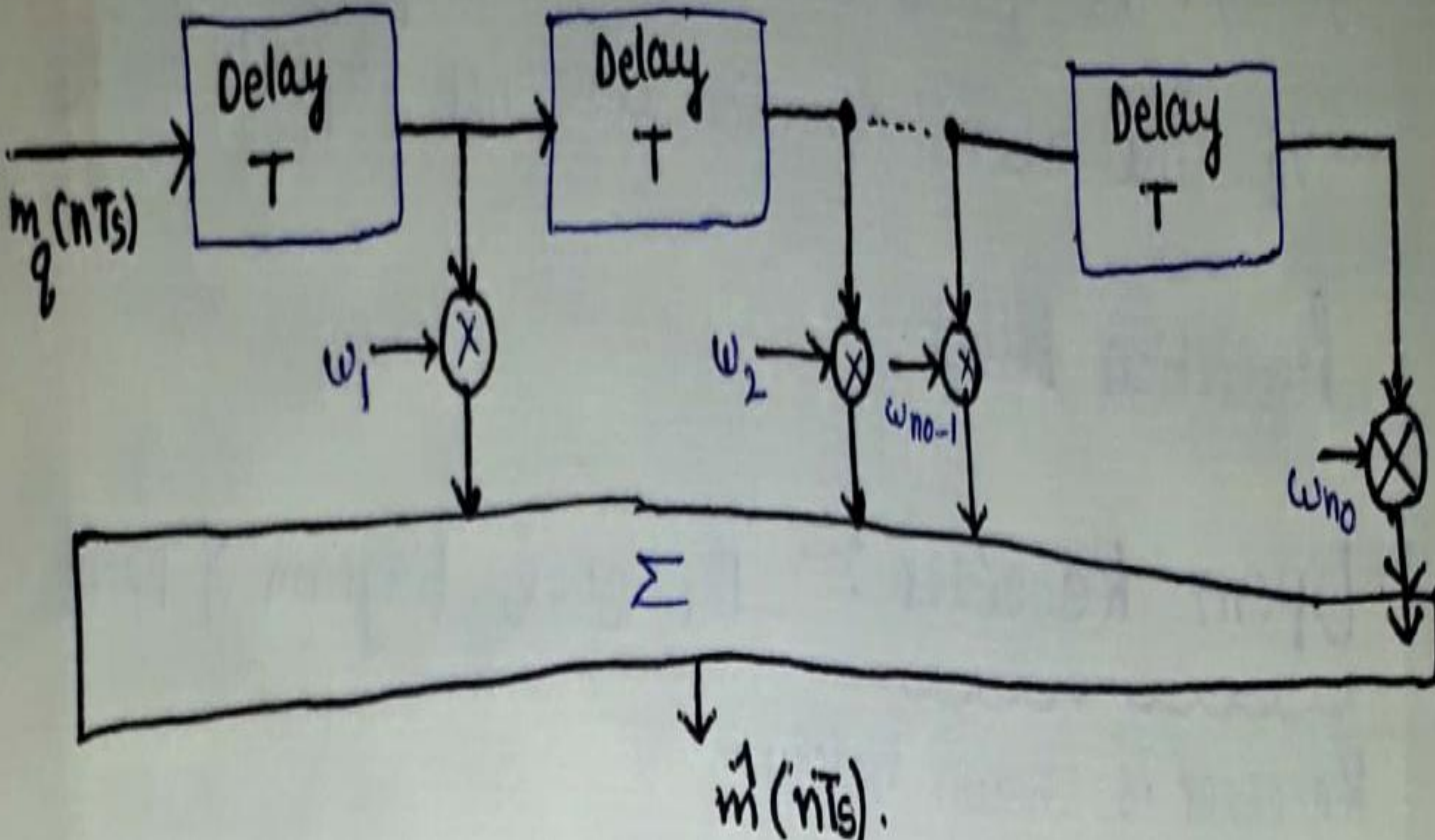
from the above equations 4 and 5 we get,

$$xq(nTs) = x(nTs) + q(nTs)$$

Therefore, the quantized version of signal xq(nTs) is the sum of original sample value and quantized error q(nTs).

The quantized error can be positive or negative.

TAPPED DELAY LINE FILTER



DELTA MODULATION

- In delta modulation it transmits **one bit per sample**, Present sample is **compared** with previous sample value.
- The type of modulation, where the sampling rate is much higher and in which the step size after quantization is of a smaller value **Δ**, such a modulation is termed as **delta modulation**.
- The difference between input signal X(t) and staircase approximation signal U(t) is confined to two levels i.e +δ and –δ.
- If the **difference** is **positive**, the U(t) signal **increased** by one step **'δ'**
- If the **difference** is **negative** , the U(t) signal is **decreased** by one step **'δ'**
- When step size is increased **'1'** is transmitted
- When step size is decreased **'0'** is transmitted
- Thus for each sample, **only one** binary bit is transmitted

DELTA MODULATION cont...

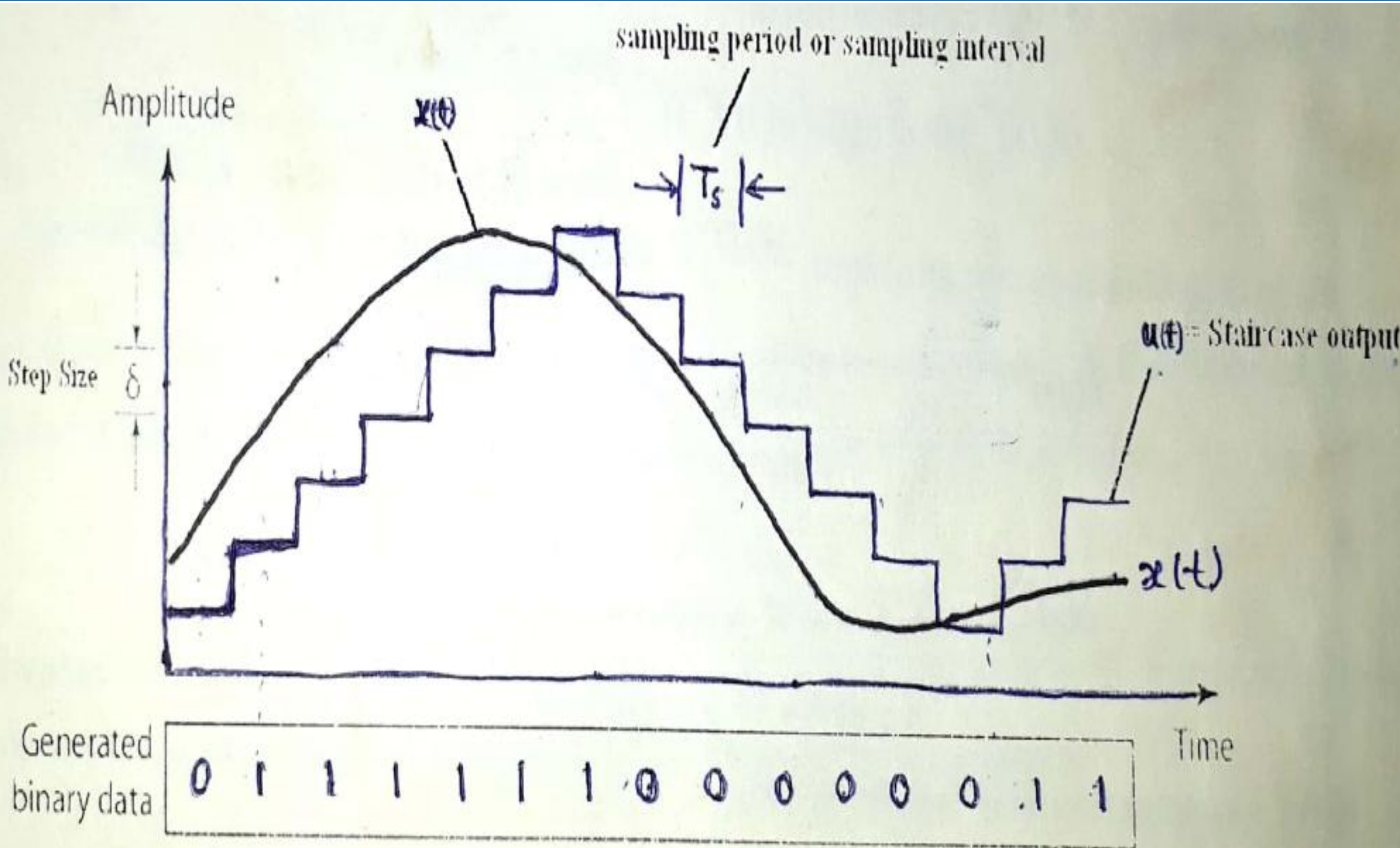
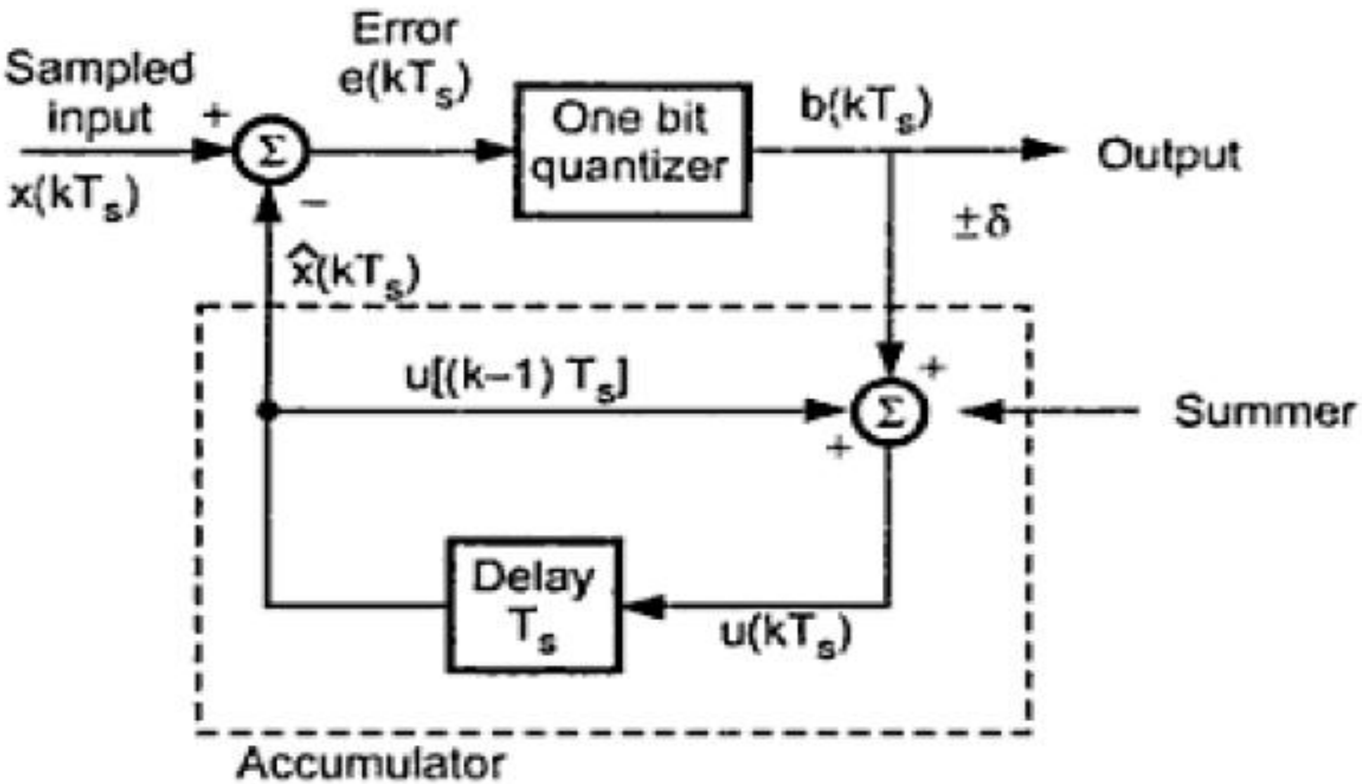


Fig: Delta modulation waveform

DELTA MODULATION cont...



DELTA MODULATION cont

$$
\begin{aligned}
u(nT_s) &= u[(n-1)T_s] + [\pm\delta] \\
&= u[nT_s - T_s] + [\pm\delta] \\
&= u(nT_s - T_s) + b(nT_s)
\end{aligned}
$$

$$
b(nT_s) = \begin{cases} +\delta & \text{if } x(nT_s) \geq \hat{x}(nT_s) \\ -\delta & \text{if } x(nT_s) < \hat{x}(nT_s) \end{cases}
$$

Depending on the sign of e(nTs) One Bit Quantizer Produces Output Step Of +δ or –δ,

+δ→ 1
-δ→ 0

DM Advantages And Disadvantages

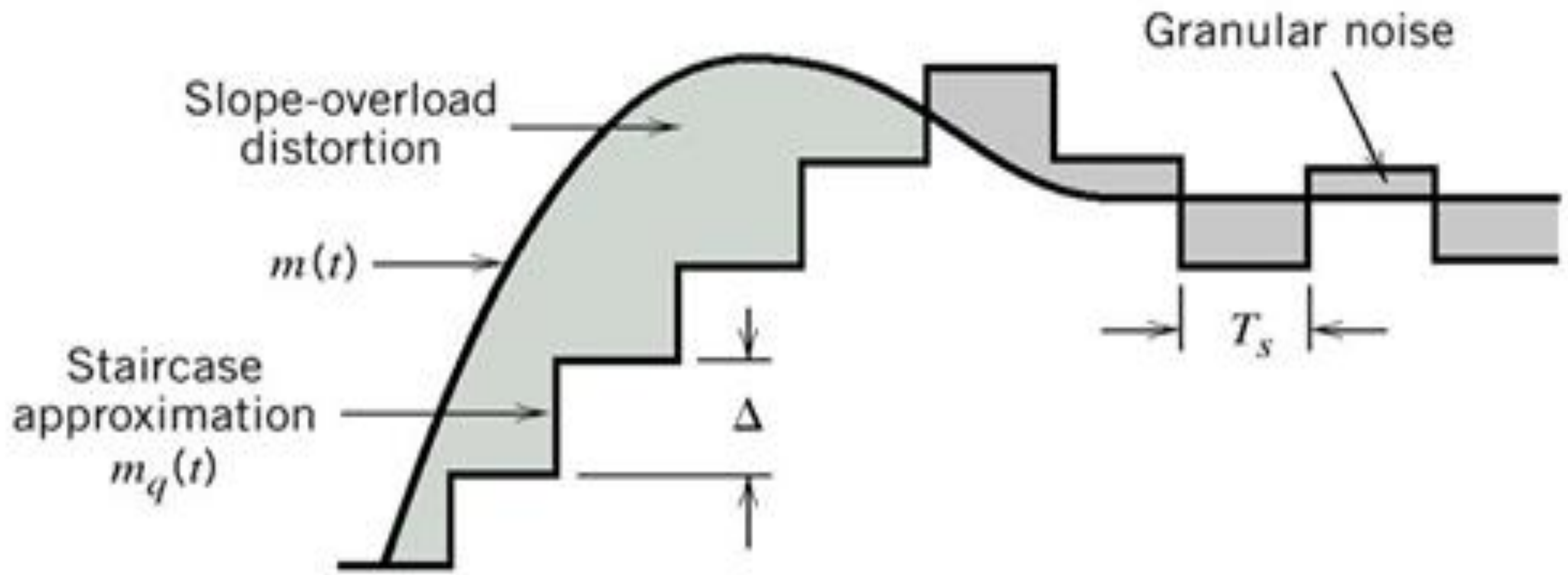
The delta modulation has following advantages over PCM,

1. Delta modulation transmits only one bit for one sample. Thus the signaling rate and transmission channel bandwidth is quite small for delta modulation.
2. The transmitter and receiver implementation is very much simple for delta modulation. There is no analog to digital converter involved in delta modulation.

Delta modulation has two major drawbacks that are:

- **Slope Over load distortion** (when **Δ** is small): (**start-up error**)Slope overload noise occurs when the slope of the input signal is greater than the delta modulator is capable of reproducing.
- **Granular noise** (when **Δ** is large): It is referred to as idle noise or **hunting. Granular noise** exists because the decoded output signal can assume only a specified number of levels in the range of interest.

QUANTIZATION ERRORS IN DM



QUANTIZATION ERRORS IN DM

Positive **slope overload**→

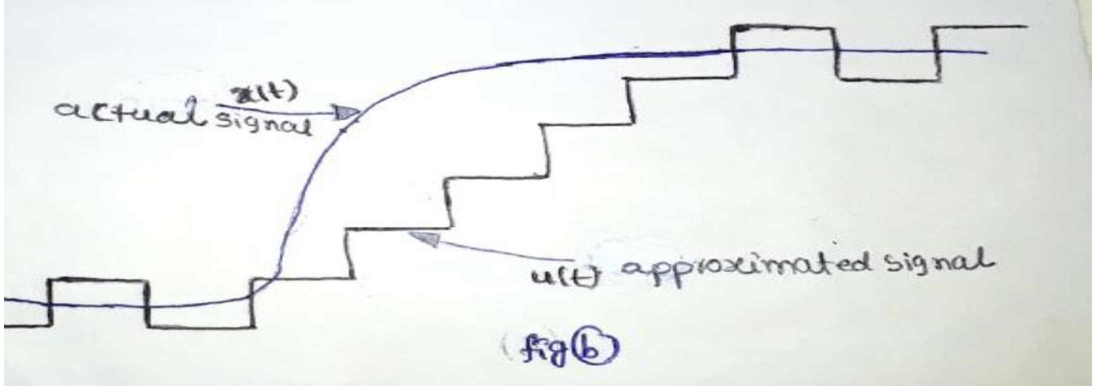


Fig: Positive Slope-Overload Distortion

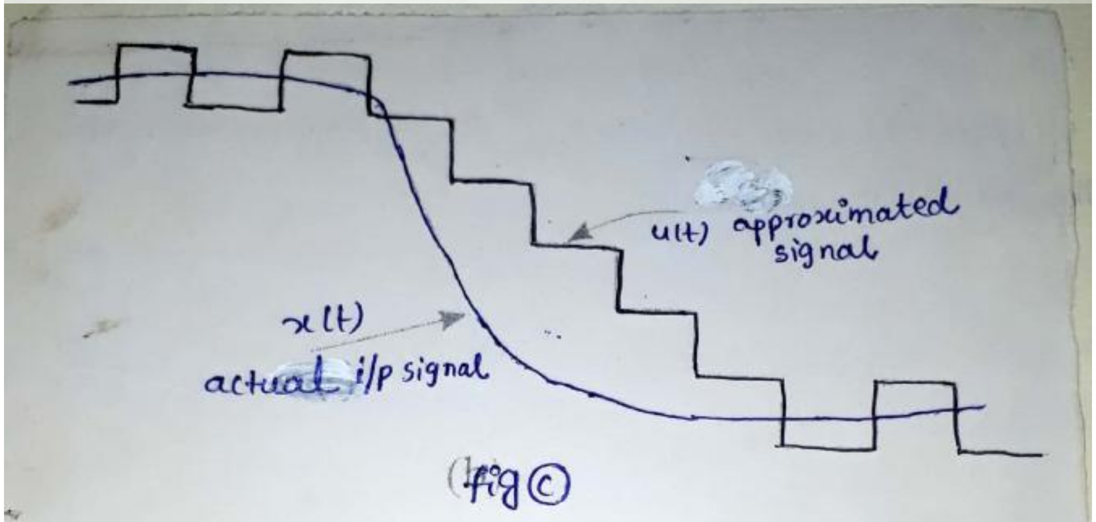
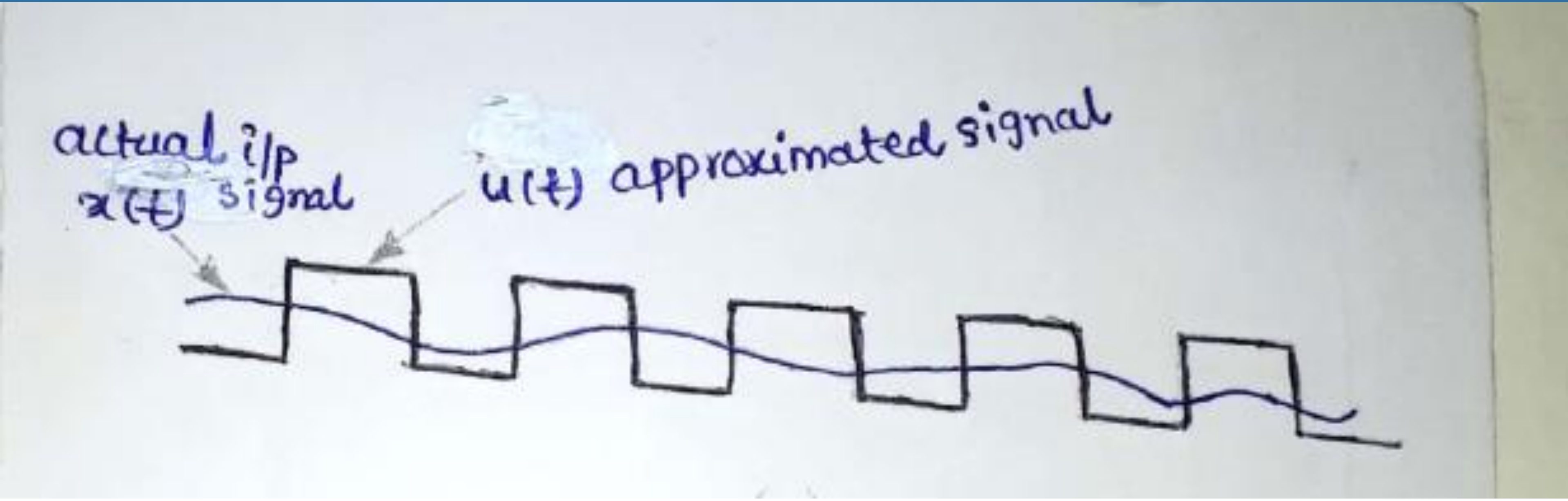


Fig: Negative Slope overload Distortion

negative slope -←overload

Granular noise in DM

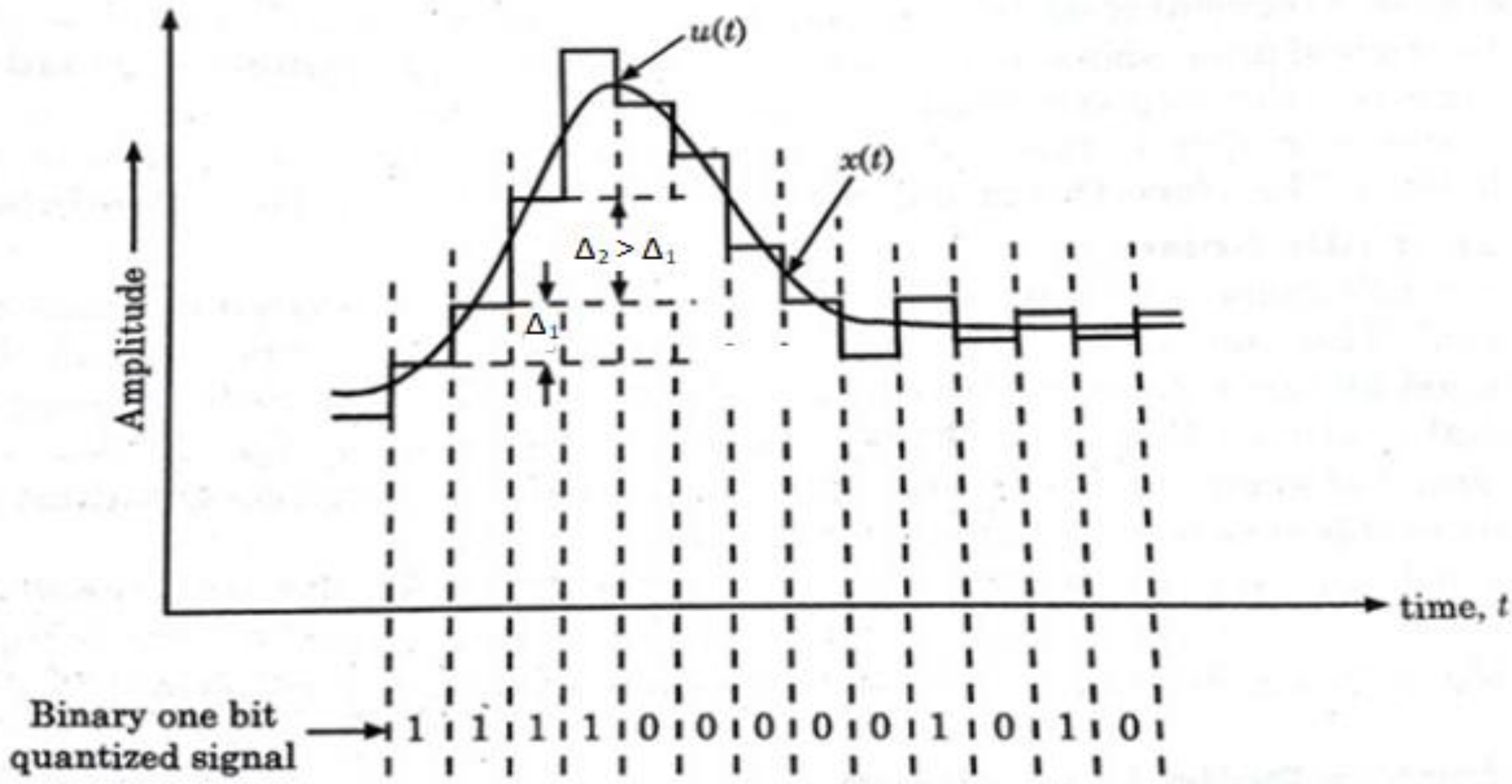


These two errors can be reduced by using ADM

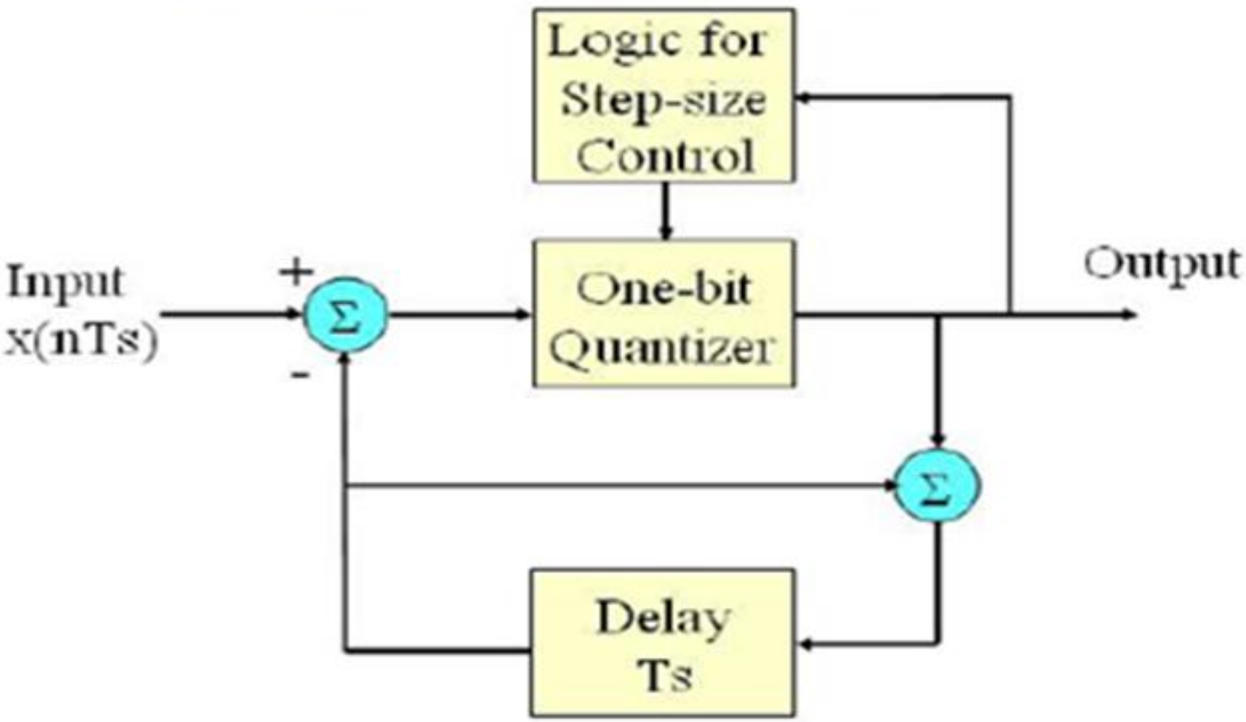
Adaptive delta modulation ADM

- The logic for step size control is added in the diagram.
- The step size increases or decreases according to a specified rule depending on one bit quantizer output.
- For an example, if one bit quantizer output is high (i.e., 1), then step size may be doubled for next sample.
- If one bit quantizer output is low, then step size may be reduced by one step.

Adaptive delta modulation ADM



Adaptive delta modulation ADM



COMPARISION OF DM AND PCM

S.NO	PARAMETER	PCM	DM
1	No. of bits per sample	4/8/16 bits	1 bit
2	No. of levels	Depends on no. of bits	Two levels
3	Step size	Fixed/variable	Fixed
4	Transmission b.w	High	Less
5	Feedback	Does not exist	Exist
6	Quantization noise	Depends on no. of bits	Slop overload and granual noise
7	Complexity of implementation	Complex	Simple

NOISE IN PCM AND DM

The signal to quantization noise ratio is given as:

$$\boxed{\frac{S}{N_q} = \frac{\text{Normalized signal power}}{\text{Normalized noise power}}}$$

$$= \frac{\text{Normalized signal power}}{\frac{\Delta^2}{12}}$$

The number of quantization value is equal to:

$$q = 2^v$$

Putting this value in eq(6), we get:

$$\Delta = \frac{2X_{max}}{2^V}$$

Substitute this value in eq, we get

$$\frac{S}{N_q} = \frac{\text{Normalized signal power}}{\left[\frac{2X_{max}}{2^v}\right]^2 * \frac{1}{12}}$$

Let the normalized signal power is equal to P then the signal to quantization noise will be given by:

$$\boxed{\frac{S}{N_q} = \frac{3P * 2^{2v}}{X_{max}^2}}$$

Digital Modulation Techniques

What is modulation ?

- **Modulation** is the process of changing the characteristics of the carrier signal, in accordance with the instantaneous values of the modulating signal.
- **Signals in the Modulation Process:**

➢ **Message or Modulating Signal**

- The signal which contains a message to be transmitted, is called as a **message signal**.it is also called as the **modulating signal**.

➢ **Carrier Signal**

- The high frequency signal which has a certain phase, frequency, and amplitude but contains no information, is called a **carrier signal**. It is an empty signal. It is just used to carry the signal to the receiver after modulation.

➢ **Modulated Signal**

- The resultant signal after the process of modulation, is called as the **modulated signal**. This signal is a combination of the modulating signal and the carrier signal.

Need for Digital Modulation

- Better performance and more cost effective than analog modulation methods (AM, FM, etc..).
- Used in modern cellular systems.
- Advancements in VLSI, DSP, etc. have made digital solutions practical and affordable.

Performance advantages:

- 1) Resistant to noise, fading, & interference.
- 2) Can combine multiple information types (voice, data, & video) in a single transmission channel.

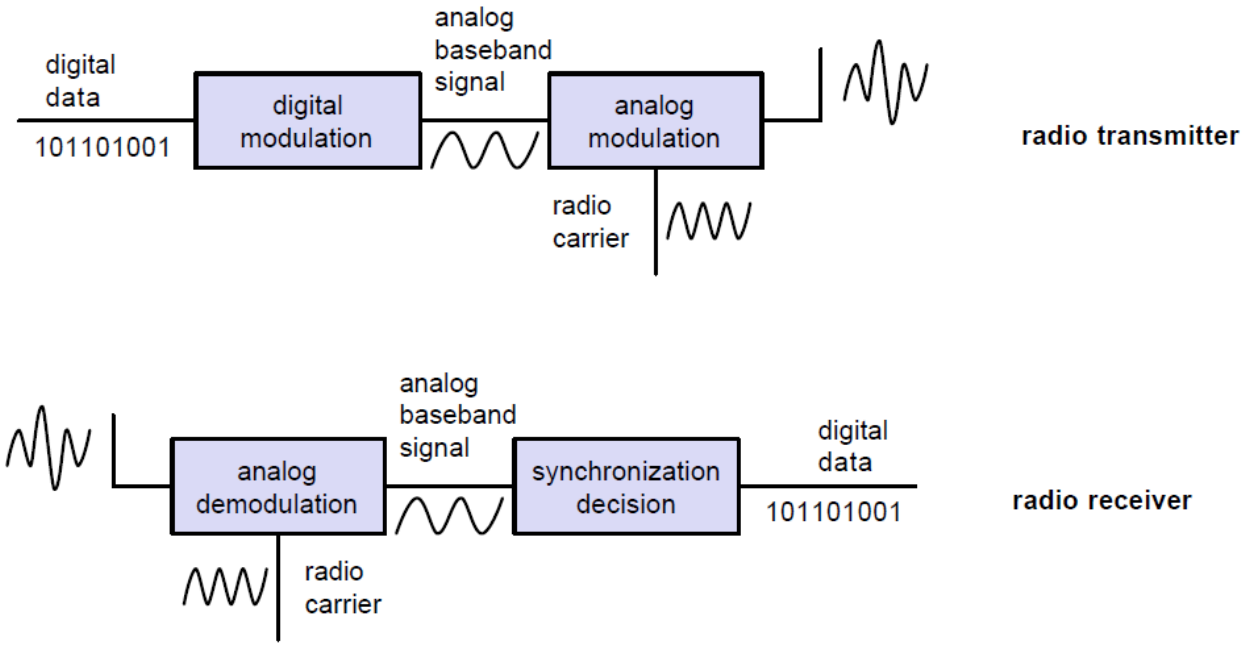
Performance advantages Continued…

- 3) Improved security (e.g. Encryption)
- 4) Error coding is used to detect/correct transmission errors.
- 5) Can implement modulation/demodulation functions using DSP software (instead of hardware circuits).

Performance factors to consider:

- 1) low Bit Error Rate (BER) at low S/N.
- 2) occupying a minimum amount of BW.
- 3) easy and cheap to implement in mobile unit.
- 4) efficient use of battery power in mobile unit.

Digital modulation and demodulation model



There are many types of digital modulation techniques and also their combinations, depending upon the need. Some major techniques are,

- ASK – Amplitude Shift Keying.
- FSK – Frequency Shift Keying.
- PSK – Phase Shift Keying.

Digital Modulation Techniques

Amplitude Shift Keying (ASK) is a type of Amplitude Modulation which represents the binary data in the form of variations in the amplitude of a signal.

- ASK is implemented by changing the amplitude of a carrier signal to reflect amplitude levels in the digital signal.
- For example: a digital “1” could not affect the signal, whereas a digital “0” would, by making it zero.
- ASK is also know as ON-OFF Keying(OOK)
- In ASK ,frequency is kept constant, amplitude has 2 levels (for bit 1 and for bit 0) i.e

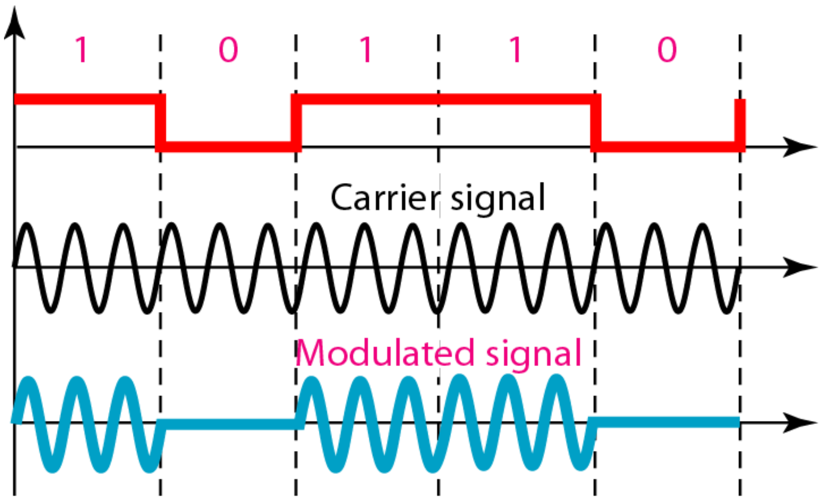
$$s(t)_{ASK} = \overbrace{m(t)}^{\text{Information}} A_c \cos(2\pi \overbrace{f_c}^{\text{Carrier frequency}} t) = \begin{cases} A_c \cos(2\pi f_c t) & m(nT_b) = 1 \\ 0 & m(nT_b) = 0 \end{cases}$$

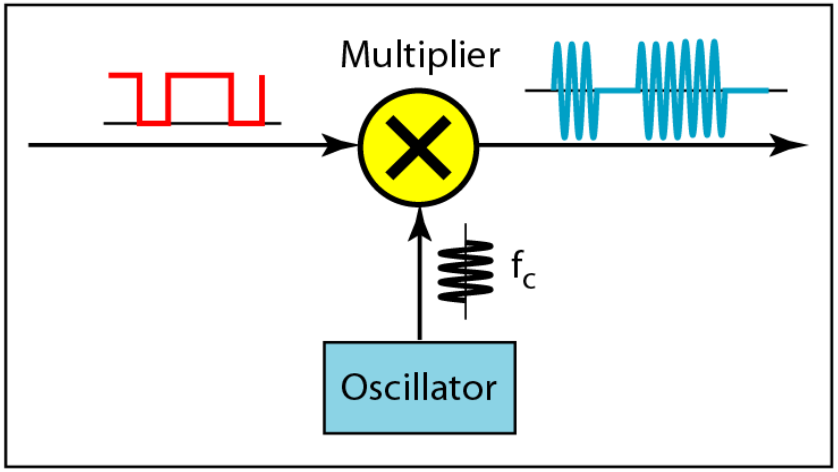
Note: *Signal power* = $A_c^2/2 = E/T_b$, E = Energy, Assume R = 1 Ohm.

Therefore, $A_c = \sqrt{\dfrac{2E}{T_b}}$

ASK MODULATOR

- The ASK modulator block diagram comprises of the carrier signal generator, the binary sequence from the message signal and the band-limited filter. Following is the block diagram of the ASK Modulator.





- The carrier generator, sends a continuous high-frequency carrier. The binary sequence from the message signal makes the unipolar input to be either High or Low.
- The high signal closes the switch, allowing a carrier wave to the output. Hence, the output will be the carrier signal at high input.
- When there is low input, the switch opens, allowing no voltage to appear. Hence, the output will be low.

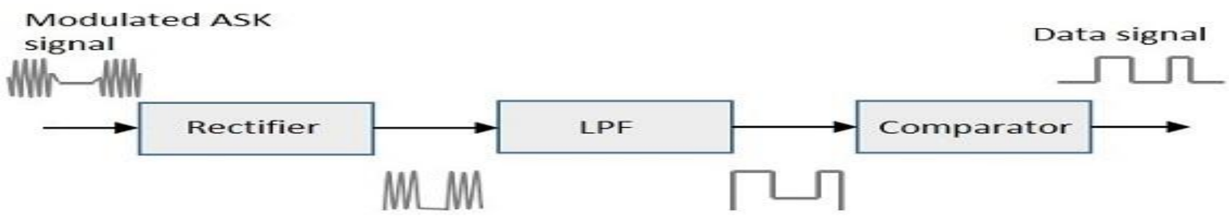
ASK Demodulator

- There are two types of ASK Demodulation techniques. They are,
  - •Asynchronous ASK Demodulation/detection.
  - •Synchronous ASK Demodulation/detection.

- The clock frequency at the transmitter when matches with the clock frequency at the receiver, it is known as a **Synchronous method**, as the frequency gets synchronized. Otherwise, it is known as **Asynchronous**.

Asynchronous ASK Demodulator:

The Asynchronous ASK detector consists of a half-wave rectifier, a low pass filter, and a comparator. Following is the block diagram for the same.

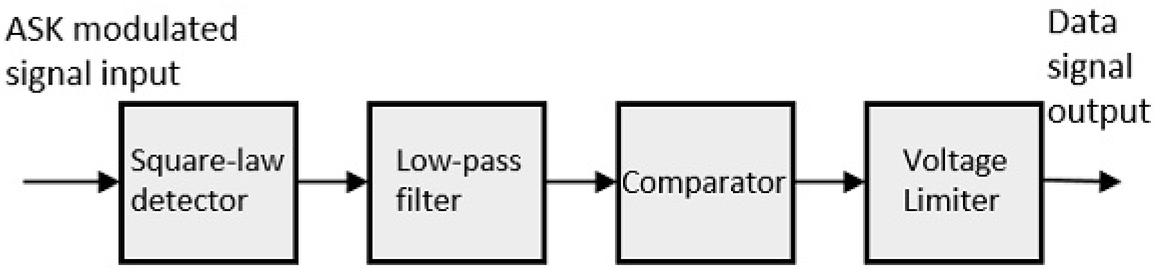


Asynchronous ASK detector

- The modulated ASK signal is given to the half-wave rectifier, which delivers a positive half output.
- The low pass filter suppresses the higher frequencies and gives an envelope detected output from which the comparator delivers a digital output.

Synchronous ASK Demodulator

- Synchronous ASK detector consists of a Square law detector, low pass filter, a comparator, and a voltage limiter. Following is the block diagram for the same.



Synchronous ASK detector

Synchronous ASK Demodulator continued…

- The ASK modulated input signal is given to the Square law detector.
- A square law detector is one whose output voltage is proportional to the square of the amplitude modulated input voltage.
- The low pass filter minimizes the higher frequencies.
- The comparator and the voltage limiter help to get a clean digital output.
- Applications
- Advantages
- disadvantages

Bandwidth of ASK:

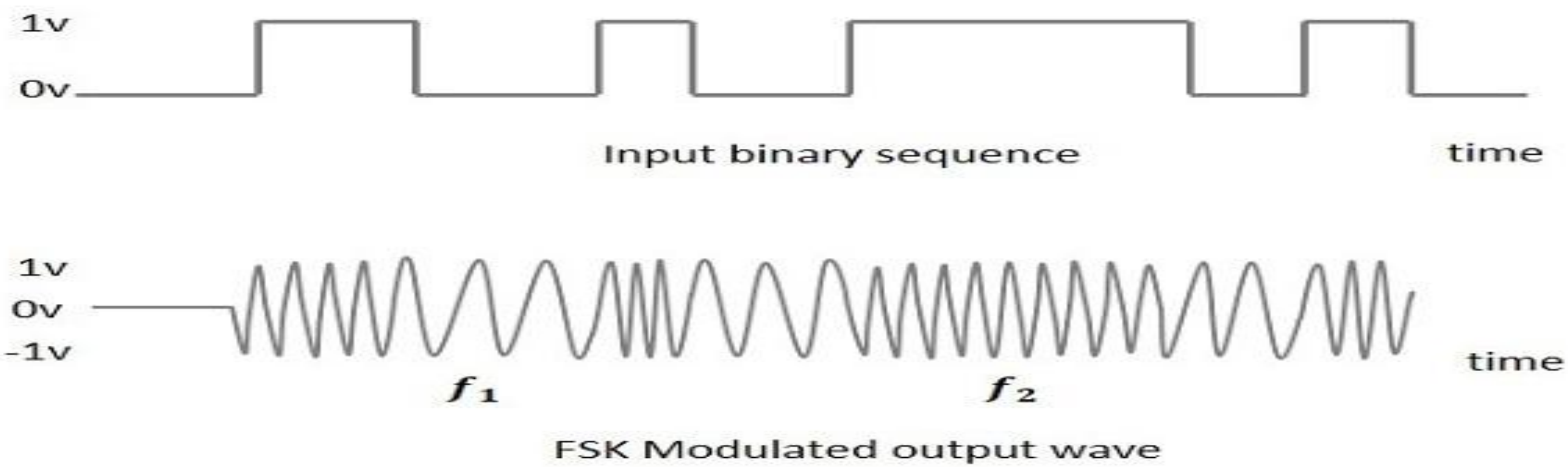
In Amplitude Shift Keying, the bandwidth required is given by

B=(1+d)SB=(1+d)S,

where B is bandwidth, S is the signal rate and d is a value of either 00 or 11.

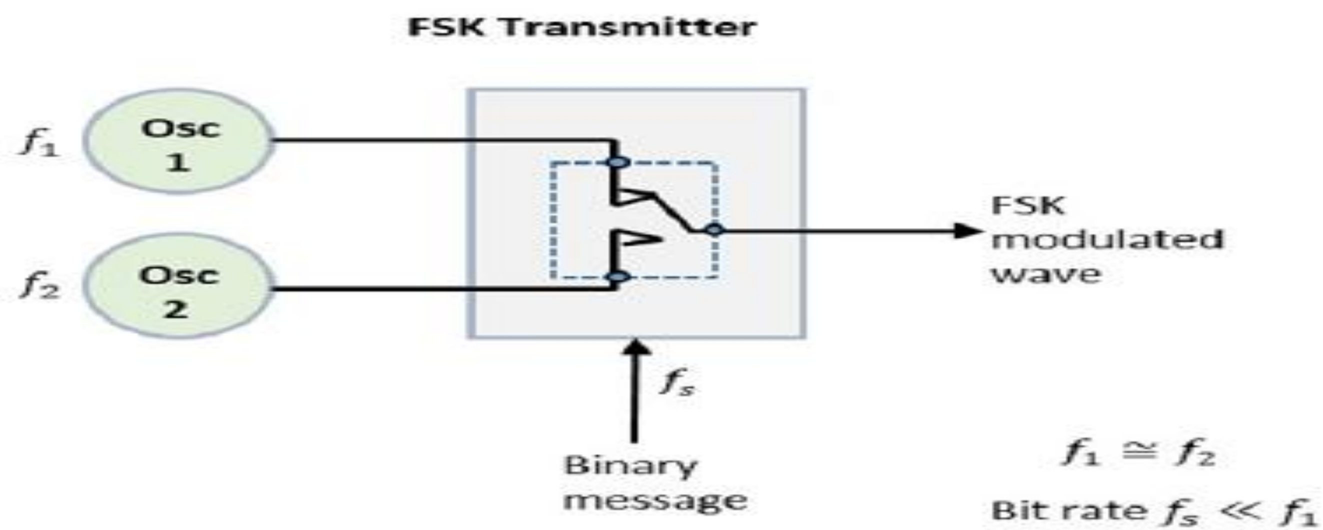
Frequency Shift Keying (FSK)

- **Frequency Shift Keying (FSK)** is the digital modulation technique in which the frequency of the carrier signal varies according to the digital signal changes. FSK is a scheme of frequency modulation.
- The output of a FSK modulated wave is high in frequency for a binary High input and is low in frequency for a binary Low input. The binary **1s** and **0s** are called Mark and Space frequencies.
- The following image is the diagrammatic representation of FSK modulated waveform along with its input.



FSK Modulator

- The FSK modulator block diagram comprises of two oscillators with a clock and the input binary sequence. Following is its block diagram.



- In frequency-shift keying, the signals transmitted for marks (binary ones) and spaces (binary zeros) .

$$S_{FSK}(t) = Ac \operatorname{Cos} m w_c t \qquad 0<t<T_b \text{ for '1'}$$
$$= Ac \operatorname{Cos} n w_c t \qquad 0<t<T_b \text{ for '0'}$$

FSK Modulator Continued...

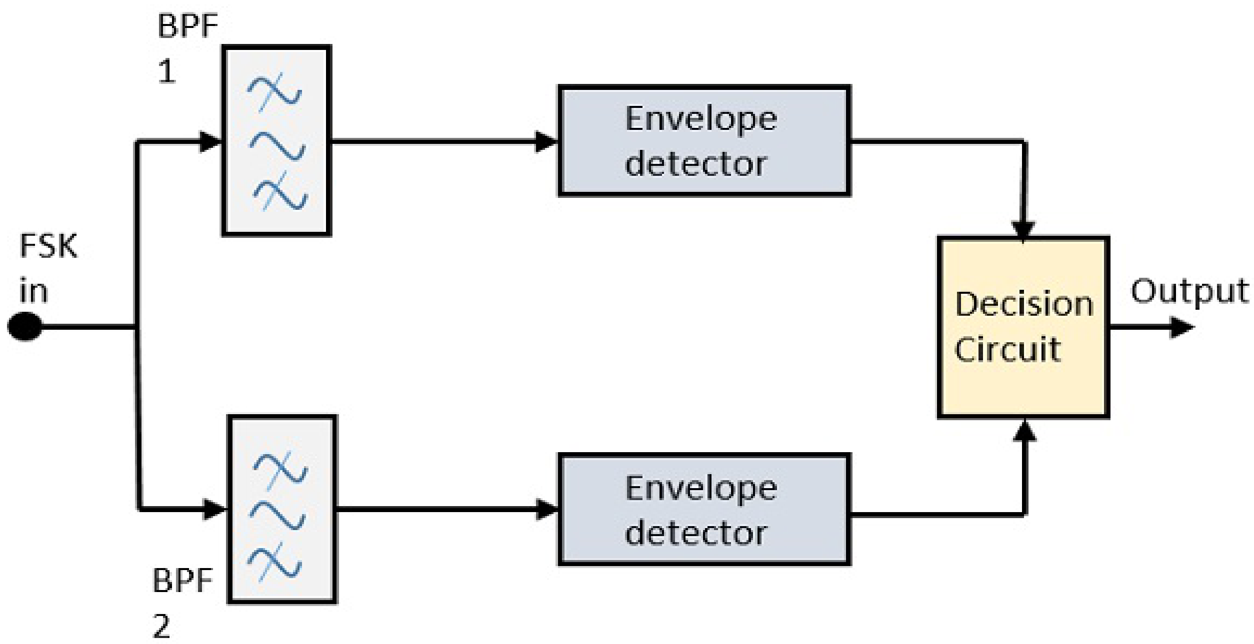
- The two oscillators, producing a higher and a lower frequency signals, are connected to a switch along with an internal clock.
- To avoid the abrupt phase discontinuities of the output waveform during the transmission of the message, a clock is applied to both the oscillators, internally.
- The binary input sequence is applied to the transmitter so as to choose the frequencies according to the binary input.

FSK Demodulator

- There are different methods for demodulating a FSK wave.
- The main methods of FSK detection are asynchronous detector and synchronous detector.
- The synchronous detector is a coherent one, while asynchronous detector is a non-coherent one.

Asynchronous FSK Detector

The block diagram of Asynchronous FSK detector consists of two band pass filters, two envelope detectors, and a decision circuit. Following is the diagrammatic representation.

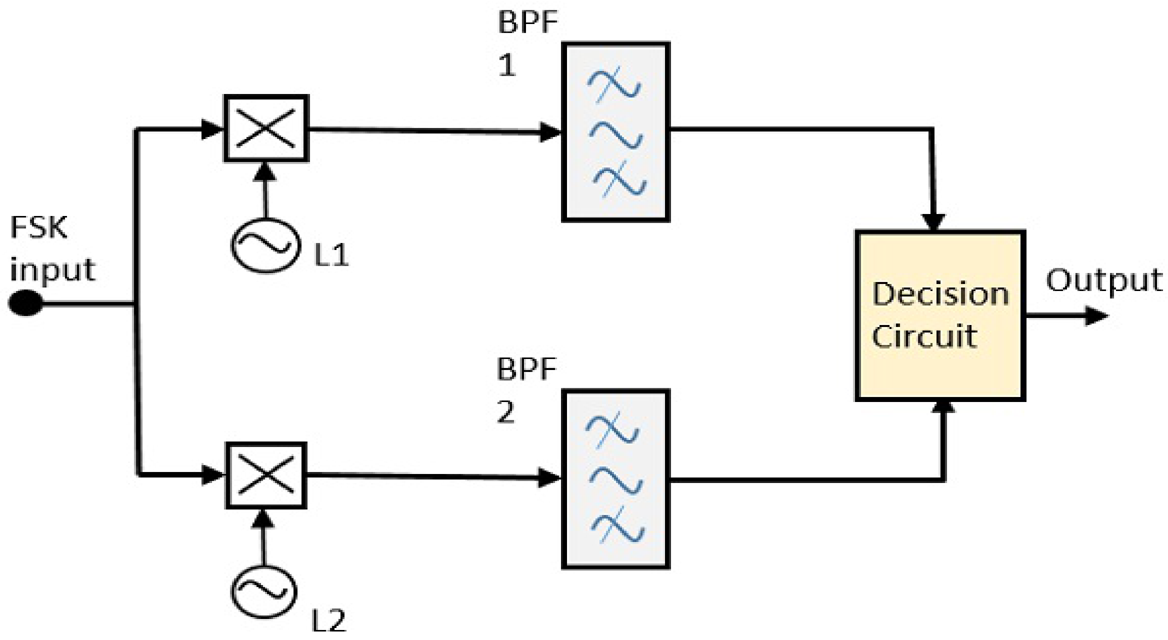


FSK Demodulator continued...

- The FSK signal is passed through the two Band Pass Filters (BPFs), tuned to **Space** and **Mark** frequencies.
- The output from these two BPFs look like ASK signal, which is given to the envelope detector.
- The signal in each envelope detector is modulated asynchronously.
- The decision circuit chooses which output is more likely and selects it from any one of the envelope detectors.
- It also re-shapes the waveform to a rectangular one.

Synchronous FSK Detector

•The block diagram of Synchronous FSK detector consists of two mixers with local oscillator circuits, two band pass filters and a decision circuit. Following is the diagrammatic representation.



FSK Demodulator continued...

- The FSK signal input is given to the two mixers with local oscillator circuits. These two are connected to two band pass filters.
- These combinations act as demodulators and the decision circuit chooses which output is more likely and selects it from any one of the detectors. The two signals have a minimum frequency separation.

❖ For both of the demodulators, the bandwidth of each of them depends on their bit rate.

❖ This synchronous demodulator is a bit complex than asynchronous type demodulators.

Phase Shift Keying

- **Phase Shift Keying (PSK)** is the digital modulation technique in which the phase of the carrier signal is changed by varying the sine and cosine inputs at a particular time.
- PSK technique is widely used for wireless LANs, bio-metric, contactless operations, along with RFID and Bluetooth communications.
- PSK is of two types, depending upon the phases the signal gets shifted. They are

❖ **Binary Phase Shift Keying (BPSK)**

❖ **Quadrature Phase Shift Keying (QPSK)**

1. Binary Phase Shift Keying (BPSK)

- This is also called as 2-phase PSK or Phase Reversal Keying. In this technique, the sine wave carrier takes two phase reversals such as 0° and 180°.
- BPSK is basically a Double Side Band Suppressed Carrier (DSBSC) modulation scheme, for message being the digital information.

2.Quadrature Phase Shift Keying (QPSK)

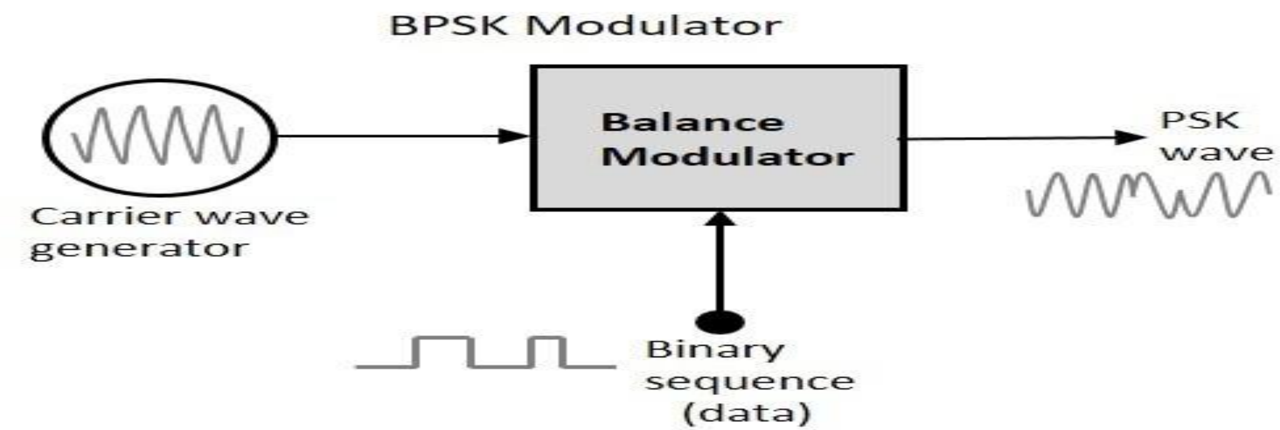
- This is the phase shift keying technique, in which the sine wave carrier takes four phase reversals such as 0°, 90°, 180°, and 270°.
- If this kind of techniques are further extended, PSK can be done by eight or sixteen values also, depending upon the requirement.

BPSK Modulator

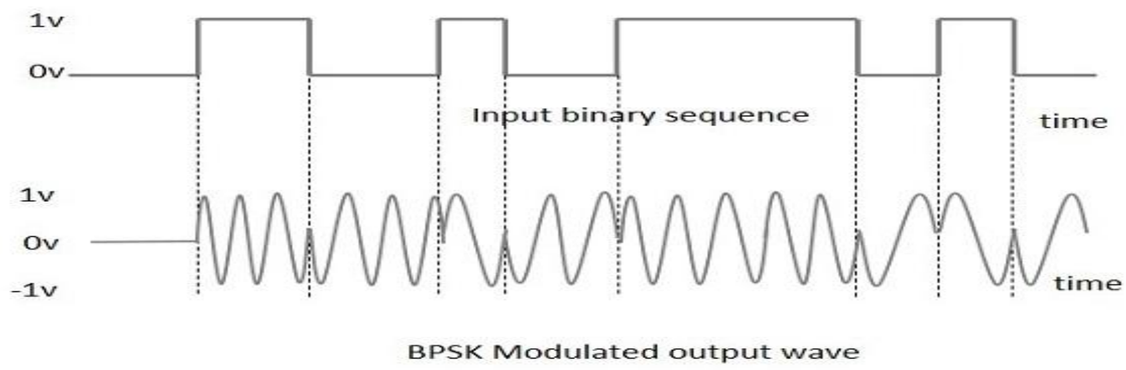
- The block diagram of Binary Phase Shift Keying consists of the balance modulator which has the carrier sine wave as one input and the binary sequence as the other input. Following is the diagrammatic representation.
- The modulation of BPSK is done using a balance modulator, which multiplies the two signals applied at the input. For a zero binary input, the phase will be 0° and for a high input, the phase reversal is of 180°.

$$s(t)_{PSK} = A_c m(t)\cos(2\pi f_c t) = \begin{cases} A_c \cos(2\pi f_c t) & m(nT_b) = 1 \\ A_c \cos(2\pi f_c t + \pi) & m(nT_b) = -1 \end{cases}$$

BPSK Modulator continued…



following is the diagrammatic representation of BPSK Modulated output wave along with its given input.

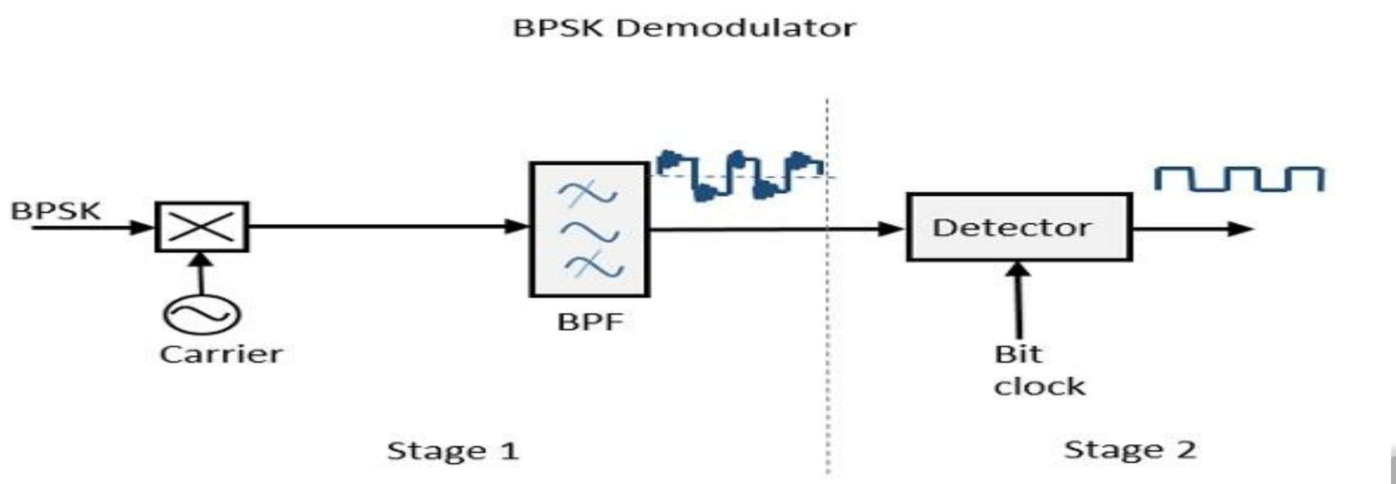


BPSK Modulator continued...

- The output sine wave of the modulator will be the direct input carrier or the inverted (180° phase shifted) input carrier, which is a function of the data signal.

BPSK Demodulator

- The block diagram of BPSK demodulator consists of a mixer with local oscillator circuit, a band pass filter, a two-input detector circuit. The diagram is as follows.



BPSK Demodulator continued…

- By recovering the band-limited message signal, with the help of the mixer circuit and the band pass filter, the first stage of demodulation gets completed.
- The base band signal which is band limited is obtained and this signal is used to regenerate the binary message bit stream.
- In the next stage of demodulation, the bit clock rate is needed at the detector circuit to produce the original binary message signal.
- If the bit rate is a sub-multiple of the carrier frequency, then the bit clock regeneration is simplified.
- To make the circuit easily understandable, a decision-making circuit may also be inserted at the 2^{nd} stage of detection.

Quadrature Phase Shift Keying (QPSK)

- The **Quadrature Phase Shift Keying (QPSK)** is a variation of BPSK, and it is also a Double Side Band Suppressed Carrier (DSBSC) modulation scheme, which sends two bits of digital information at a time, called as **dibits**.

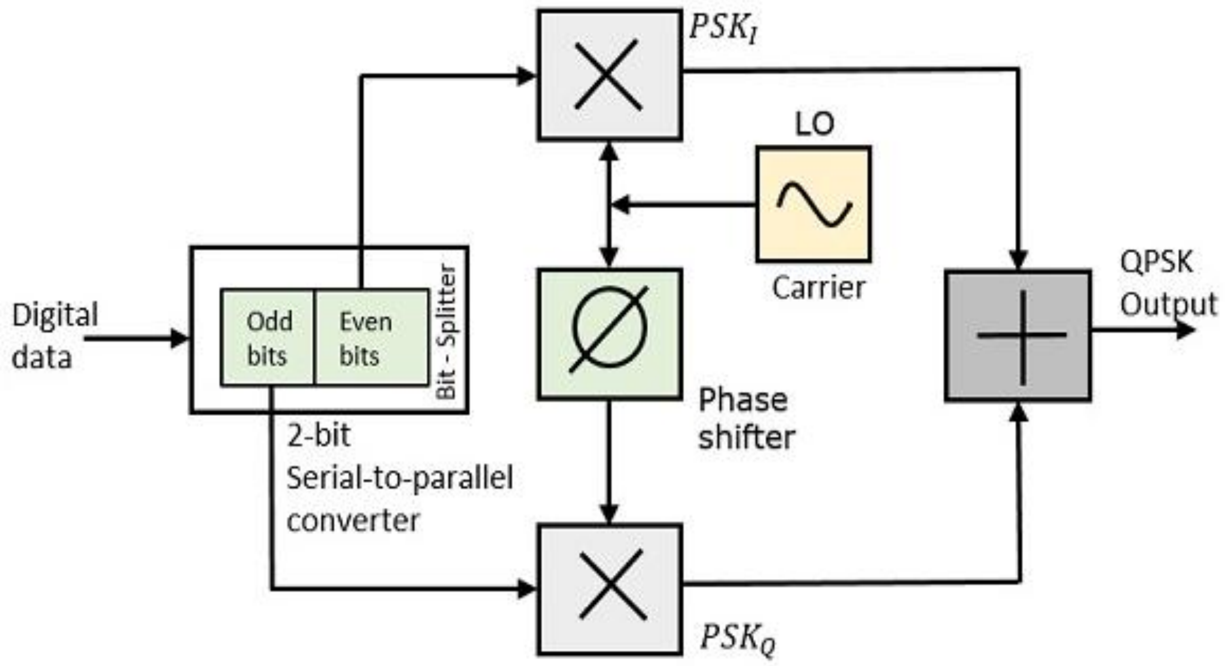
- Instead of the conversion of digital bits into a series of digital stream, it converts them into bit pairs. This decreases the data bit rate to half, which allows space for the other users.

- This modulation scheme is very important for developing concepts of two-dimensional I-Q modulations as well as for its practical relevance.

- A group of two bits is often called a ‘dibit’. So, four dibits are possible. Each symbol carries same energy.

QPSK Modulator

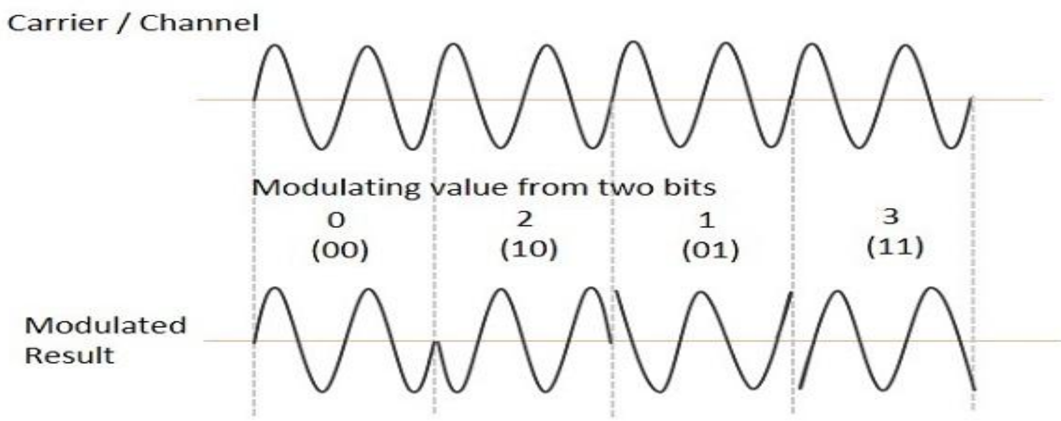
- The QPSK Modulator uses a bit-splitter, two multipliers with local oscillator, a 2-bit serial to parallel converter, and a summer circuit. Following is the block diagram for the same.



QPSK Modulator continued…

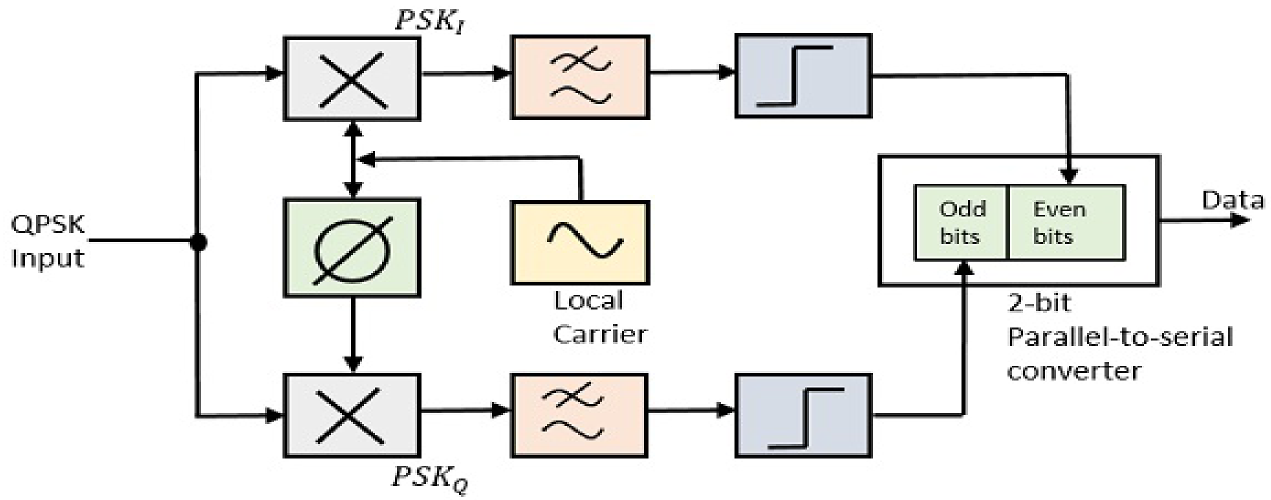
- At the modulator's input, the message signal's even bits (i.e., 2nd bit, 4th bit, 6th bit, etc.) and odd bits (i.e., 1st bit, 3rd bit, 5th bit, etc.) are separated by the bits splitter and are multiplied with the same carrier to generate odd BPSK (called as **PSK_I**) and even BPSK (called as **PSK_Q**).
- The **PSK_Q** signal is anyhow phase shifted by 90° before being modulated.

❖ The QPSK waveform for two-bits input is as follows, which shows the modulated result for different instances of binary inputs.



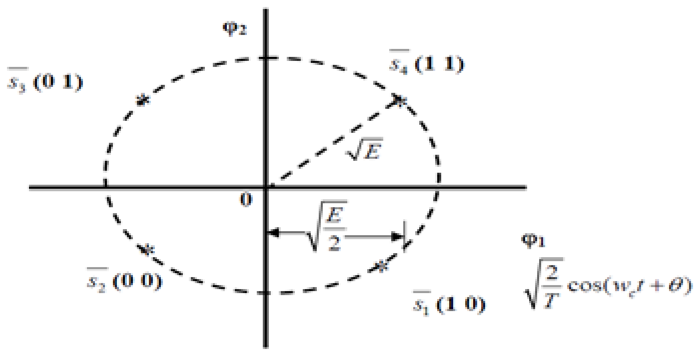
QPSK Demodulator

- The QPSK Demodulator uses two product demodulator circuits with local oscillator, two band pass filters, two integrator circuits, and a 2-bit parallel to serial converter.
- Following is the diagram for the qpsk demodulator.



- The two product detectors at the input of demodulator simultaneously demodulate the two BPSK signals.
- The pair of bits are recovered here from the original data. These signals after processing, are passed to the parallel to serial converter.

Signal constellation for QPSK

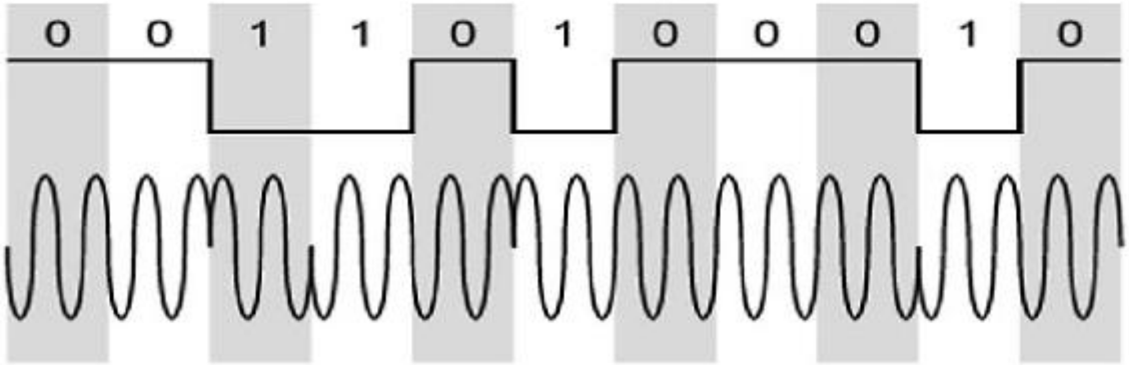


Input	Dibit		Phase of QPSK	Coordinates of signal points		
	(b_0)	(b_e)		s_{i1}	s_{i2}	i
$\bar{s}_1$	1	0	$\pi/4$	$+\sqrt{E/2}$	$-\sqrt{E/2}$	1
$\bar{s}_2$	0	0	$3\pi/4$	$-\sqrt{E/2}$	$-\sqrt{E/2}$	2
$\bar{s}_3$	0	1	$5\pi/4$	$-\sqrt{E/2}$	$+\sqrt{E/2}$	3
$\bar{s}_4$	1	1	$7\pi/4$	$+\sqrt{E/2}$	$+\sqrt{E/2}$	4

Differential Phase Shift Keying (DPSK)

- **Differential Phase Shift Keying (DPSK)** the phase of the modulated signal is shifted relative to the previous signal element.
- No reference signal is considered here. The signal phase follows the high or low state of the previous element.
- This DPSK technique doesn't need a reference oscillator.

- The following figure represents the model waveform of DPSK



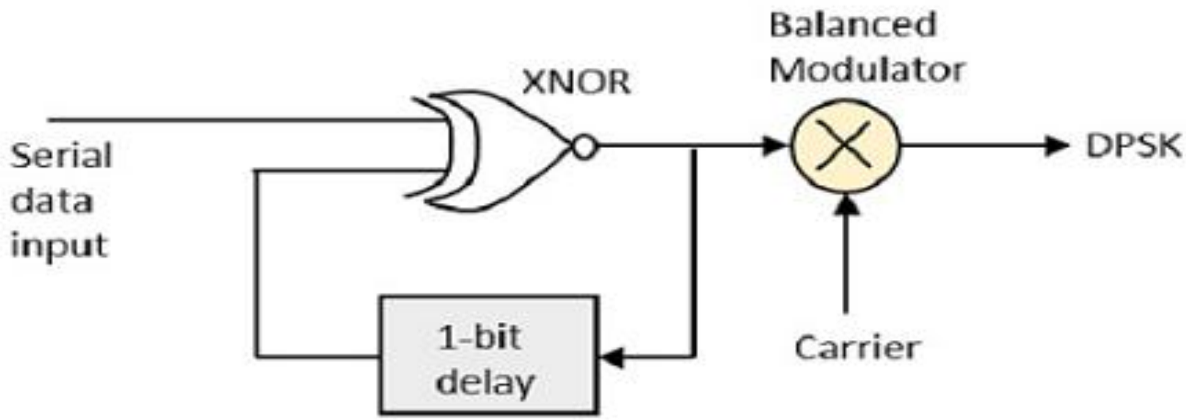
DPSK continued…

- It is seen from the above figure that, if the data bit is Low i.e., 0, then the phase of the signal is not reversed, but continued as it was.
- If the data is a High i.e., 1, then the phase of the signal is reversed, as with NRZI, invert on 1 (a form of differential encoding).
- If we observe the above waveform, we can say that the High state represents an **M** in the modulating signal and the Low state represents a **W** in the modulating signal.

DPSK Modulator

- DPSK is a technique of BPSK, in which there is no reference phase signal. Here, the transmitted signal itself can be used as a reference signal.

- Following is the diagram of DPSK Modulator.



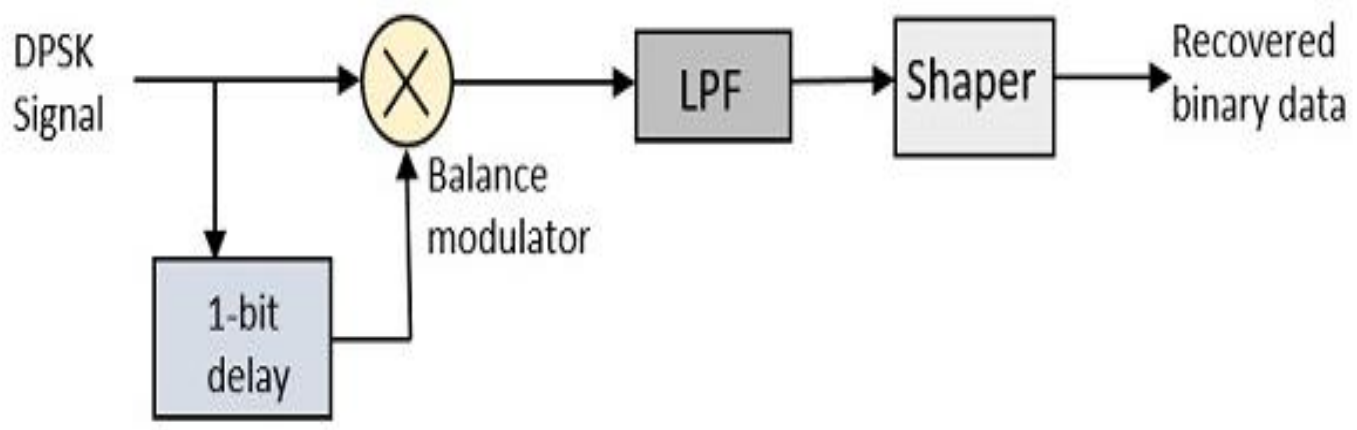
DPSK Modulator

DPSK Modulator continued…

- DPSK encodes two distinct signals, i.e., the carrier and the modulating signal with 180° phase shift each.
- The serial data input is given to the XNOR gate and the output is again fed back to the other input through 1-bit delay.
- The output of the XNOR gate along with the carrier signal is given to the balance modulator, to produce the DPSK modulated signal.

DPSK Demodulator

- In DPSK demodulator, the phase of the reversed bit is compared with the phase of the previous bit.
- Following is the block diagram of DPSK demodulator.



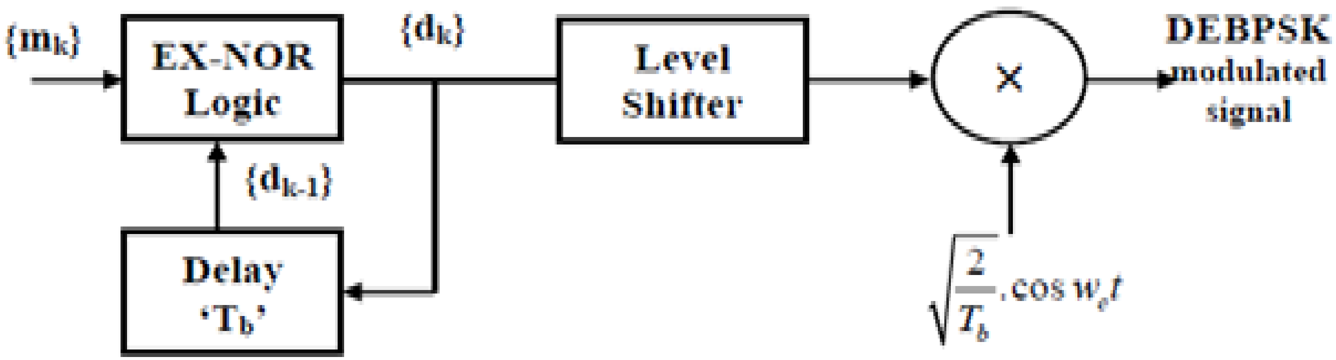
DPSK Demodulator continued…

- From the above figure, it is evident that the balance modulator is given the DPSK signal along with 1-bit delay input.

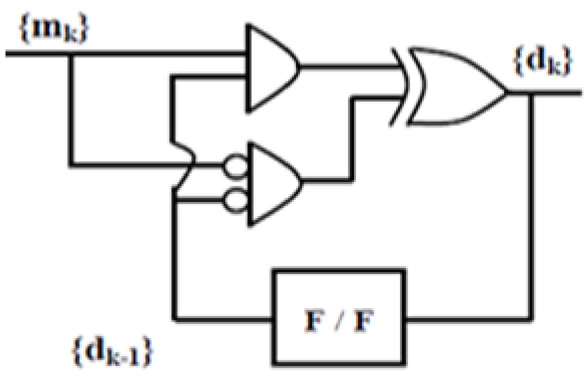
- That signal is made to confine to lower frequencies with the help of LPF. Then it is passed to a shaper circuit, which is a comparator or a Schmitt trigger circuit, to recover the original binary data as the output.

DEPSK Modulator

- When we apply differential encoding, the encoded binary '1' will be transmitted by adding 0c to the current phase of the carrier and an encoded binary '0' will be transmitted by adding πc to the current phase of the carrier.



- The transmitter (modulator) of a DEPSK system is identical to the DPSK transmitter.
- The differential phase shift keying (DPSK) is a modulation of BPSK.
- Fig shows the block diagram of DPSK generator and the relevant waveforms.



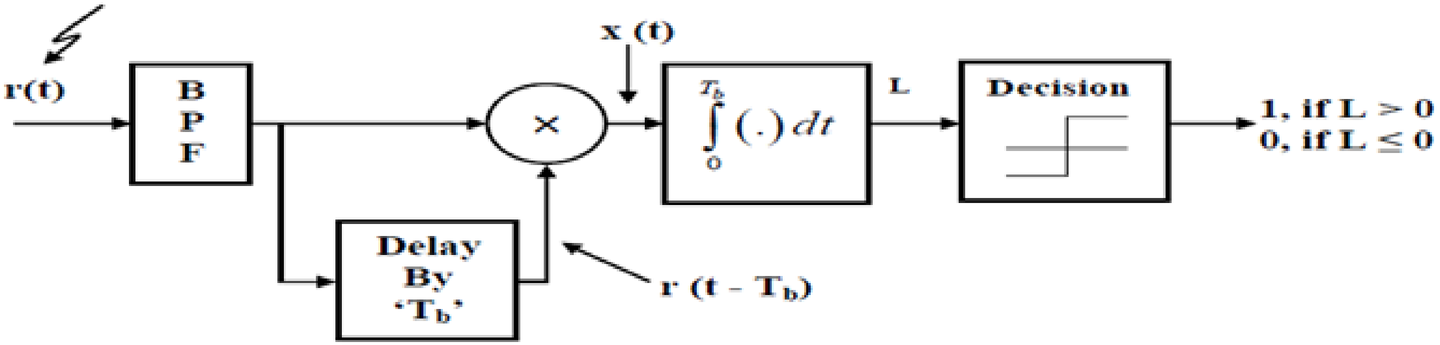
$\{m_k\}$: 1 0 1 1 0 1 0 0

$\{d_k\}$: 1 1 0 0 0 1 1 0 1

Tx.Phase : 0 0 π π π 0 0 π 0

Demod. Data : 1 0 1 1 0 1 0 0

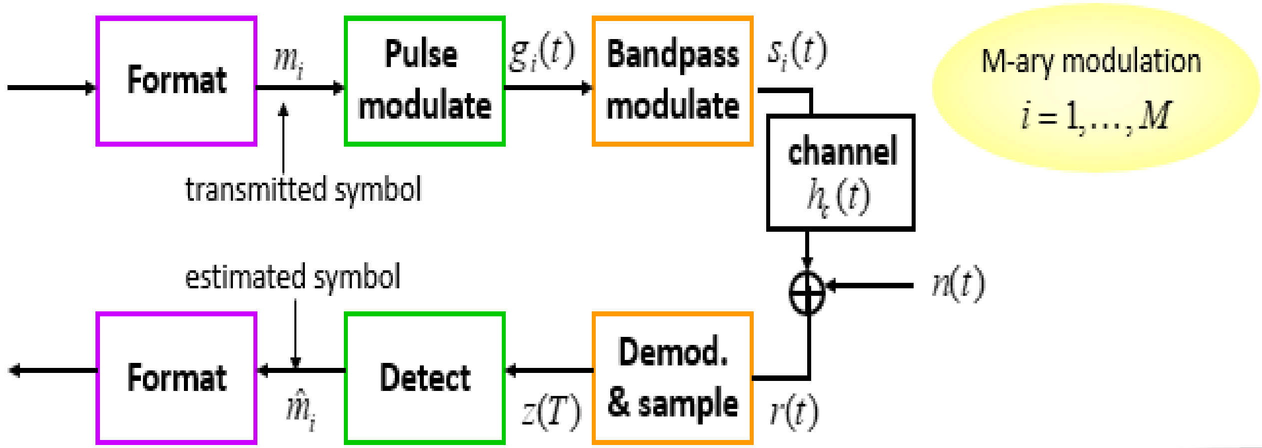
- The signal b(t) is also applied to a time delay circuit and the delayed signal b(t−Tb)b(t−Tb) is applied to the other input of the EX-OR gate as shown below ,



- If b(t)= b(t−Tb)b(t−Tb) then output of the EX-OR gate will be 0
  d (t) =0 ………if b(t)=b(t−Tb)(t)=b(t−Tb)
- if b(t) = b(t−Tb)b(t−Tb) then output of the EX-OR gate will be 1
  d (t) =1 ……… b(t) = b(t−Tb)

OPTIMAL RECEPTION OF DIGITAL SIGNALS

- Receiver structure
  - Demodulation (and sampling)
  - Detection
- First step for designing the receiver
  - Matched filter receiver
  - Correlator receiver



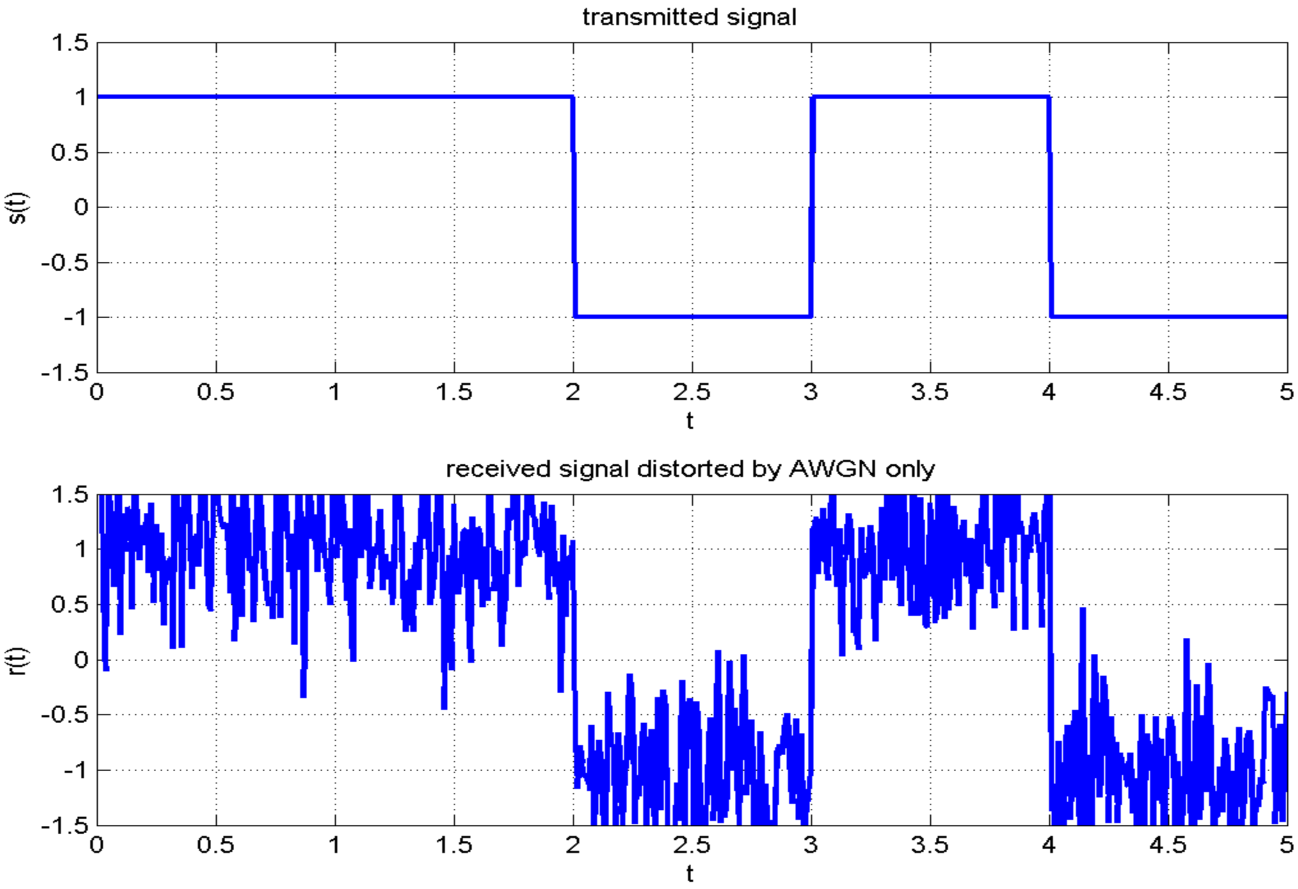
Continued…

- Major sources of errors:

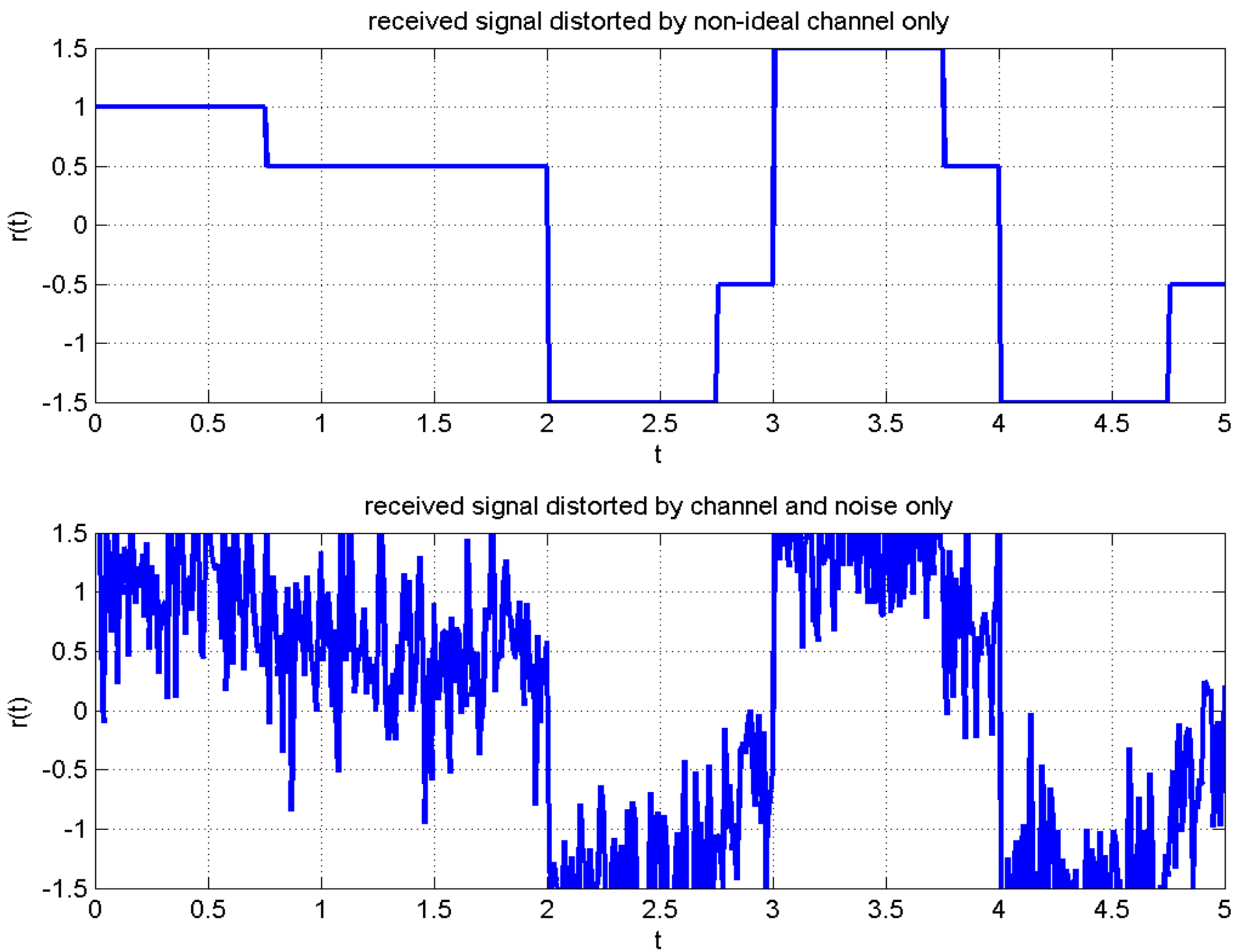
  - Thermal noise (AWGN)
    - disturbs the signal in an additive fashion (Additive)
    - has flat spectral density for all frequencies of interest (White)
    - is modelled by Gaussian random process (Gaussian Noise)
  - Inter-Symbol Interference (ISI)
    - Due to the filtering effect of transmitter, channel and receiver, symbols are "smeared".

- The receiver block consists of a processor whose structure that maximizes the output signal to noise ratio at the end of each symbol time (T) can be determined as follows.
- The input to the receiver block is the signals1(contaminated by an additive white Gaussian noise (AWGN),n(t), having a two sided power spectral density N0/2.

Noise Impact of the channel



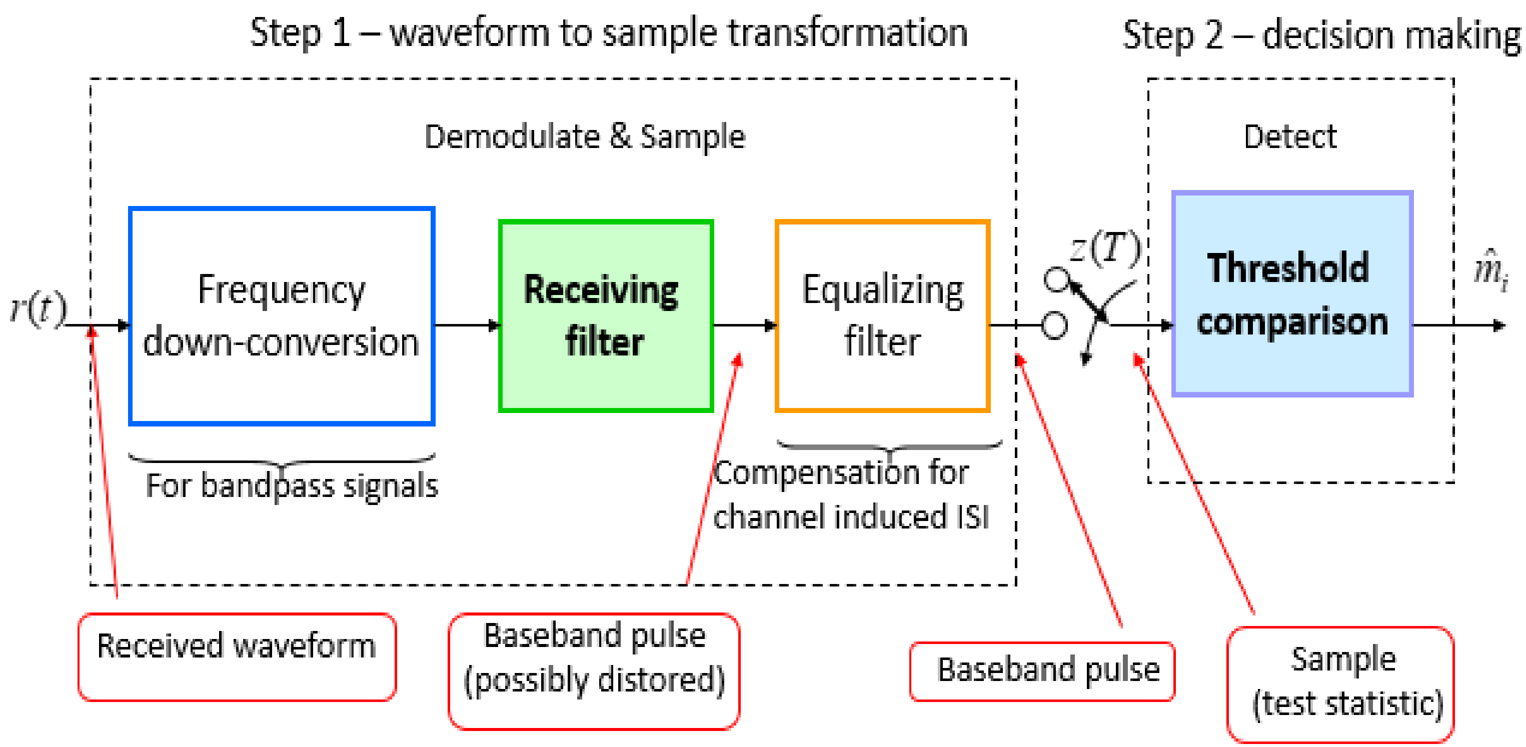
Noise Impact of the channel



Receiver tasks

- Demodulation and sampling:
  - Waveform recovery and preparing the received signal for detection:
    - Improving the signal power to the noise power (SNR) using matched filter
    - Reducing ISI using equalizer
    - Sampling the recovered waveform
- Detection
  - Estimate the transmitted symbol based on the received sample

Receiver structure

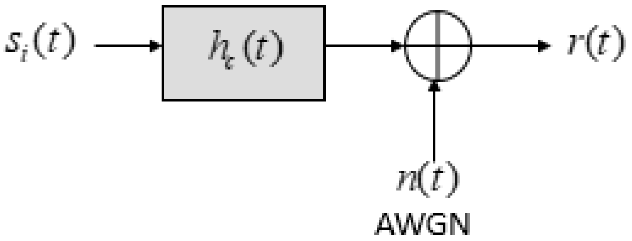


Steps in designing the receiver

- Find optimum solution for receiver design with the following goals:
  1. Maximize SNR
  2. Minimize ISI
- Steps in design:
  – Model the received signal
  – Find separate solutions for each of the goals.
- First, we focus on designing a receiver which maximizes the SNR.

Design the receiver filter to maximize the SNR

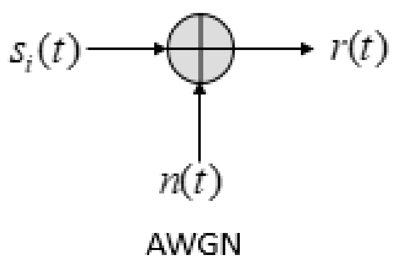
- Model the received signal



$$\boxed{r(t) = s_i(t) * h_c(t) + n(t)}$$

- Simplify the model:
  - Received signal in AWGN

Ideal channels
$h_c(t) = \delta(t)$



$$\boxed{r(t) = s_i(t) + n(t)}$$

Error Analysis

- Two ways in which errors occur:
  - A is transmitted, AT+N<0 (0 received,1 sent)
  - A is transmitted, -AT+N>0 (1 received,0 sent)

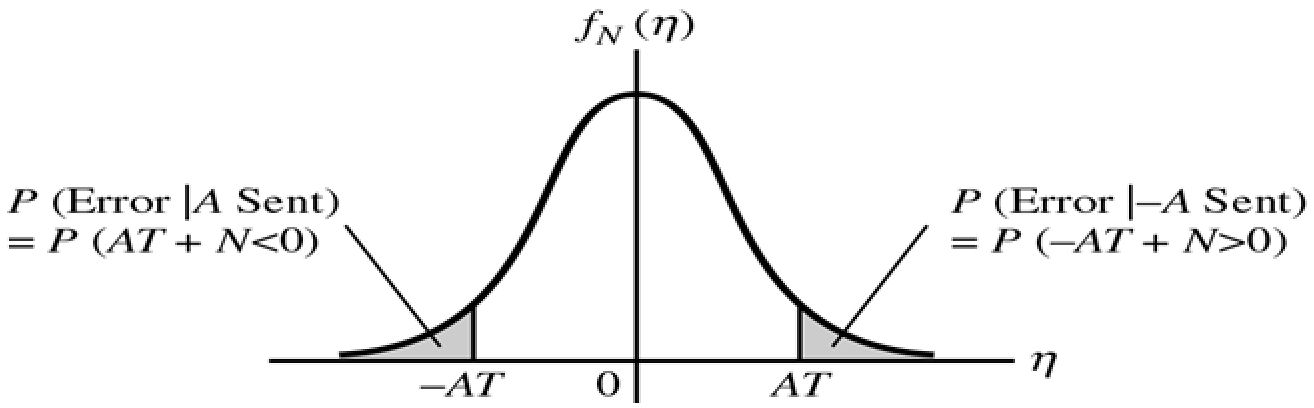


Figure 7-4 Illustration of error probabilities for binary signaling.

Error Analysis continued...

$$P(Error\,|\,A) = \int_{-\infty}^{-AT} \frac{e^{-n^2/N_0T}}{\sqrt{\pi N_0 T}} dn = Q\left(\sqrt{\frac{2A^2T}{N_0}}\right)$$

Similarly,

$$P(Error\,|-A) = \int_{AT}^{\infty} \frac{e^{-n^2/N_0T}}{\sqrt{\pi N_0 T}} dn = Q\left(\sqrt{\frac{2A^2T}{N_0}}\right)$$

The average probability of error:

$$\begin{aligned} P_E &= P(E\,|\,A)P(A) + P(E\,|-A)P(-A) \\ &= Q\left(\sqrt{\frac{2A^2T}{N_0}}\right) \end{aligned}$$

Probability of Error

- Two types of errors:

$$P(E \mid s_1(t)) = \int_k^{\infty} \frac{e^{-[v-s_{01}(T)]^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} dv = Q\left(\frac{k - s_{01}(T)}{\sigma}\right)$$

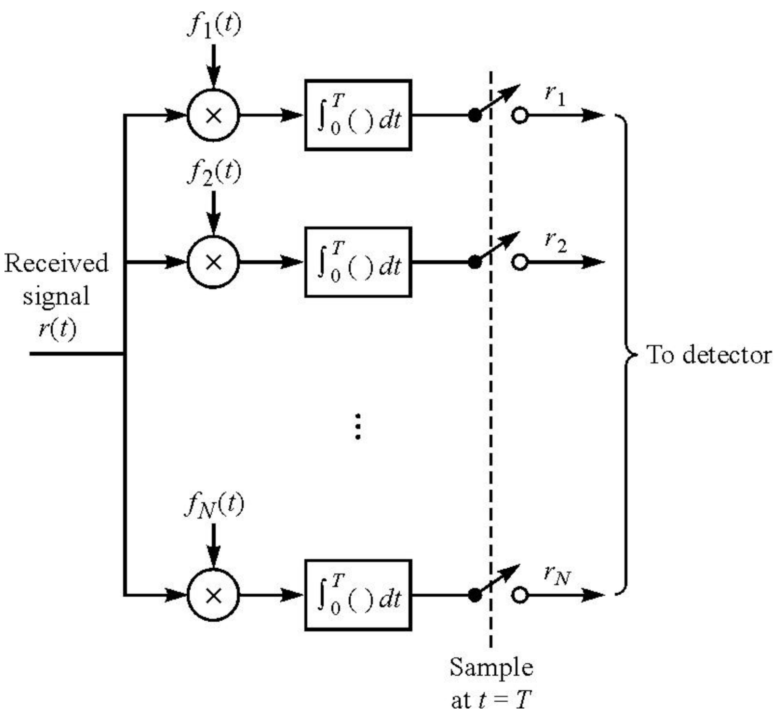
$$P(E \mid s_2(t)) = \int_{-\infty}^{k} \frac{e^{-[v-s_{02}(T)]^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} dv = 1 - Q\left(\frac{k - s_{02}(T)}{\sigma}\right)$$

- The average probability of error:

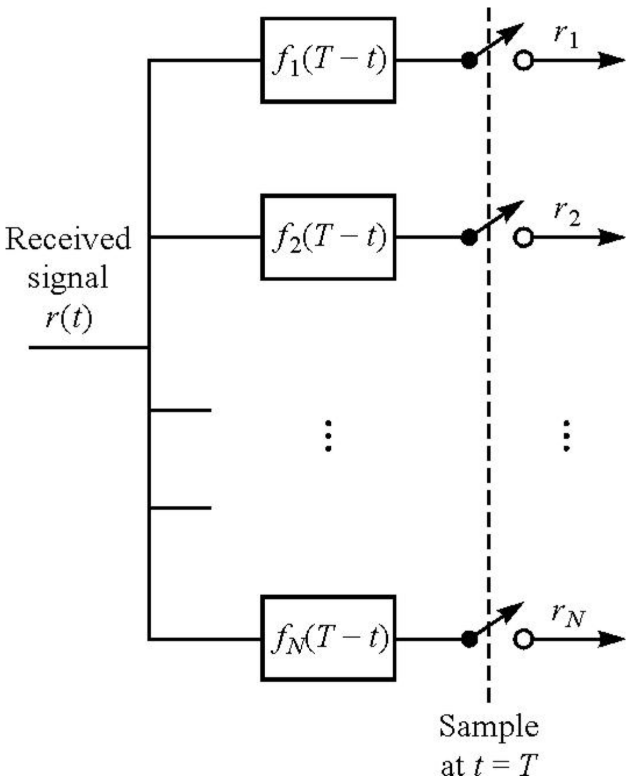
$$P_E = \frac{1}{2} P[E \mid s_1(t)] + \frac{1}{2} P[E \mid s_2(t)]$$

Demodulators

- Correlation Demodulator

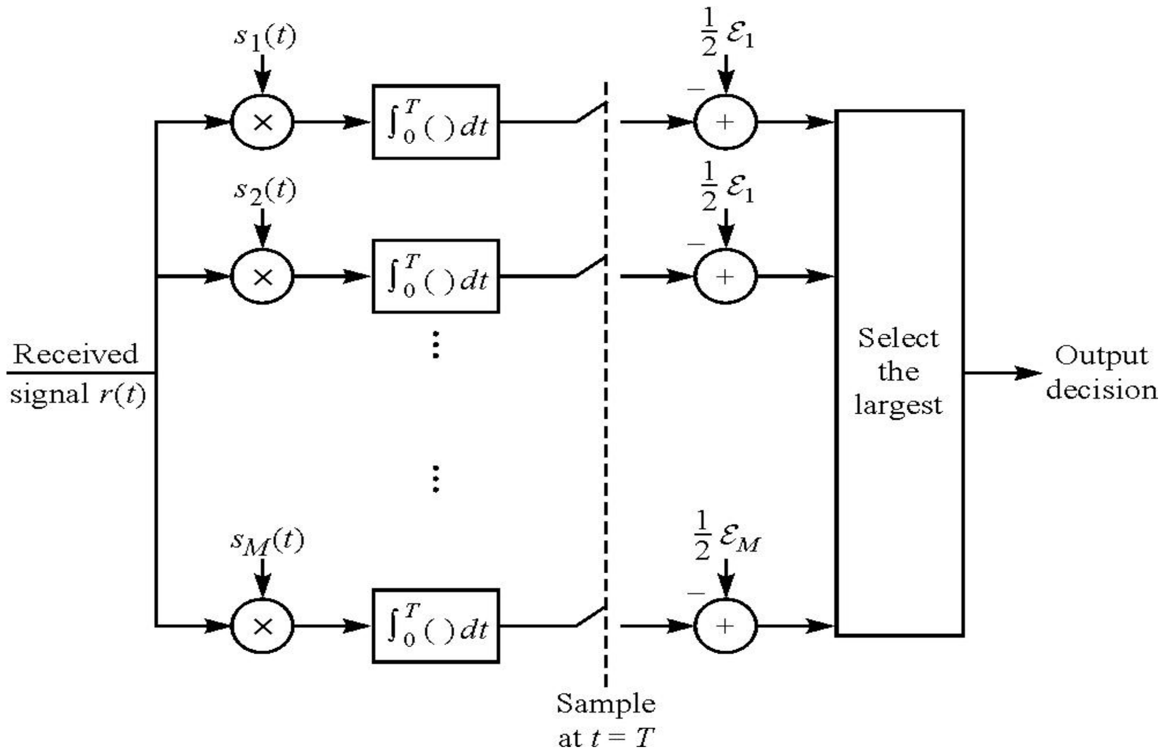


- Matched filter



Optimal Detector

- Thus get new type of correlation demodulator using symbols *not* the basis functions:



Matched Filter

- The Matched Filter is the linear filter that maximizes:

$$\left(\frac{S}{N}\right)_{out} = \frac{s_o^2(t)}{n_o^2(t)}$$

- Recall

$$y(t) = h(t) * x(t) \Leftrightarrow Y(f) = H(f)X(f)$$

$$S_y(f) = |H(f)|^2 S_x(f)$$

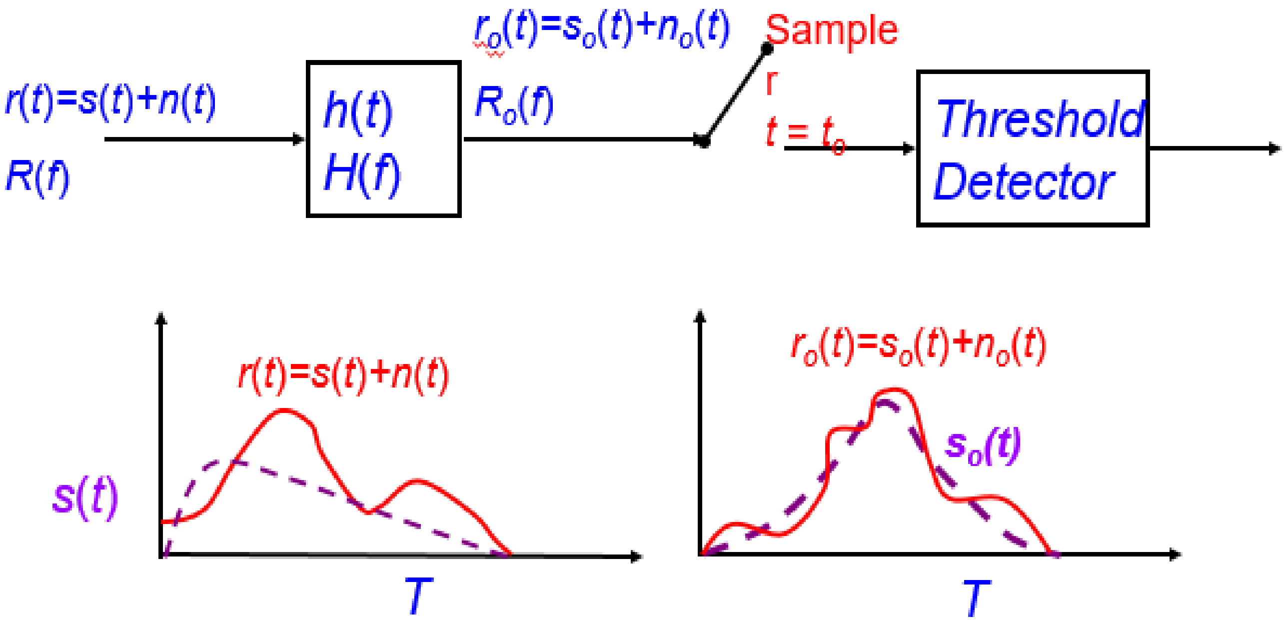
Matched Filter continued..

- Design a linear filter to minimize the effect of noise while maximizing the signal.
- $s(t)$ is the input signal and $s_0(t)$ is the output signal.
- The signal is assumed to be known and absolutely time limited and zero otherwise.
- The PSD, $P_n(f)$ of the additive input noise is also assumed to be known.
- Design the filter such that instantaneous output signal power is maximized at a sampling instant t_0, compared with the average output noise power:

$$\left(\frac{S}{N}\right)_{out} = \frac{s_o^{\,2}(t)}{n_o^{\,2}(t)}$$

Matched Filter continued..

- The goal is maximize $(S/N)_{out}$



Matched Filter continued..

- The matched filter does not preserve the input signal shape.
- The objective is to maximize the output signal-to-noise ratio.
- The matched filter is the linear filter that maximizes $(S/N)_{out}$ and has a transfer function given by:

$$H(f) = K \frac{S^*(f) e^{-j\omega t_o}}{P_n(f)}$$

- where $S(f) = F[s(t)]$ of duration T sec.
- t_0 is the sampling time
- K is an arbitrary, real, nonzero constant.
- The filter may not be realizable.

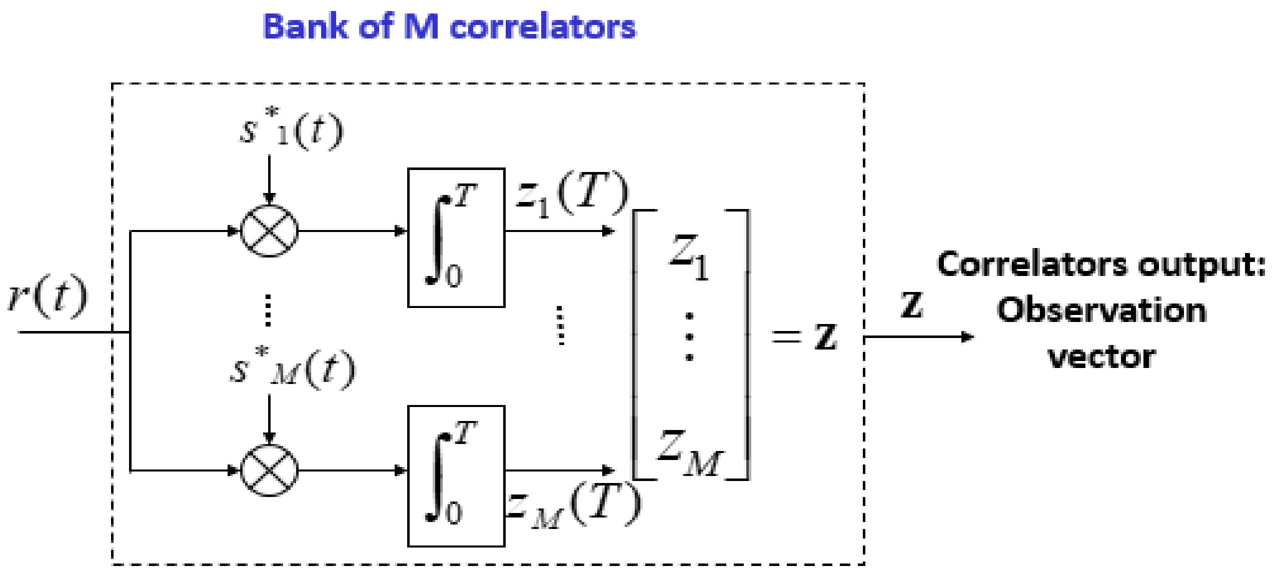
Correlator receiver

- The matched filter output at the sampling time, can be realized as the correlator output.

$$z(T) = h_{opt}(T) * r(T)$$

$$= \int_0^T r(\tau) s_i^*(\tau) d\tau = < r(t), s(t) >$$

Implementation of correlator receiver

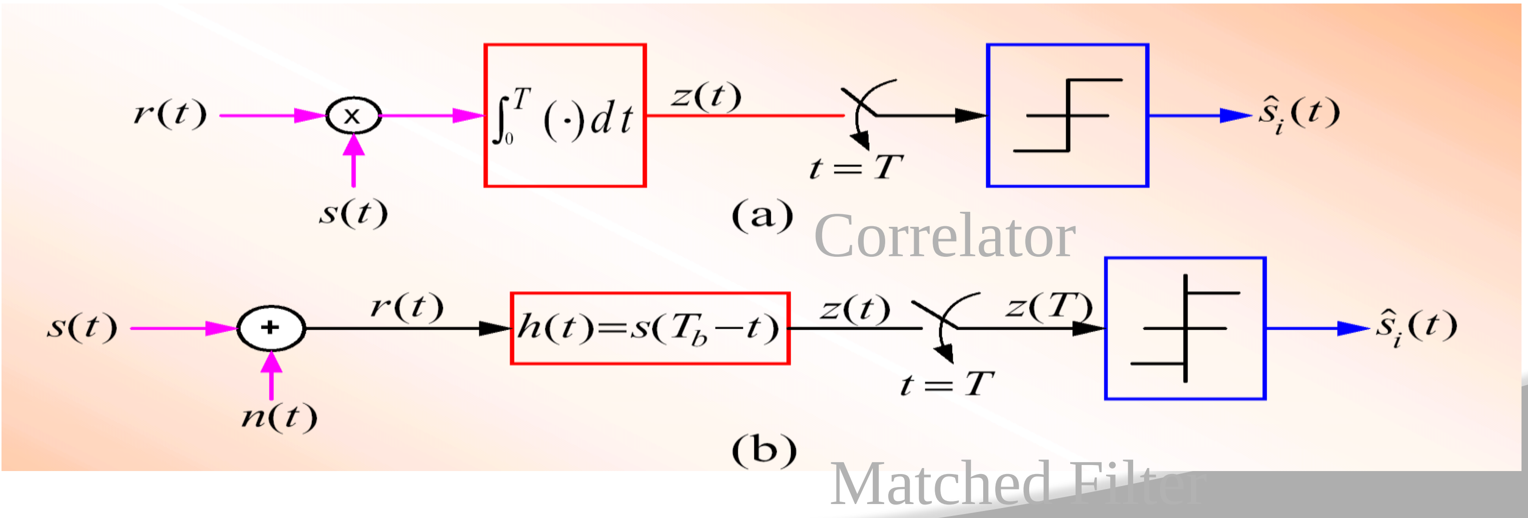


$$\mathbf{z} = (z_1(T), z_2(T), \ldots, z_M(T)) = (z_1, z_2, \ldots, z_M)$$

$$z_i = \int_0^T r(t) s_i(t) dt \qquad i = 1, \ldots, M$$

Correlator versus Matched Filter

- The functions of the ***correlator*** and ***matched filter***
- The mathematical operation of Correlator is correlation, where a signal is correlated with its replica.
- Whereas the operation of Matched filter is Convolution, where signal is convolved with filter impulse response.
- But the o/p of both is same at t=T so the functions of correlator and matched filter is same.



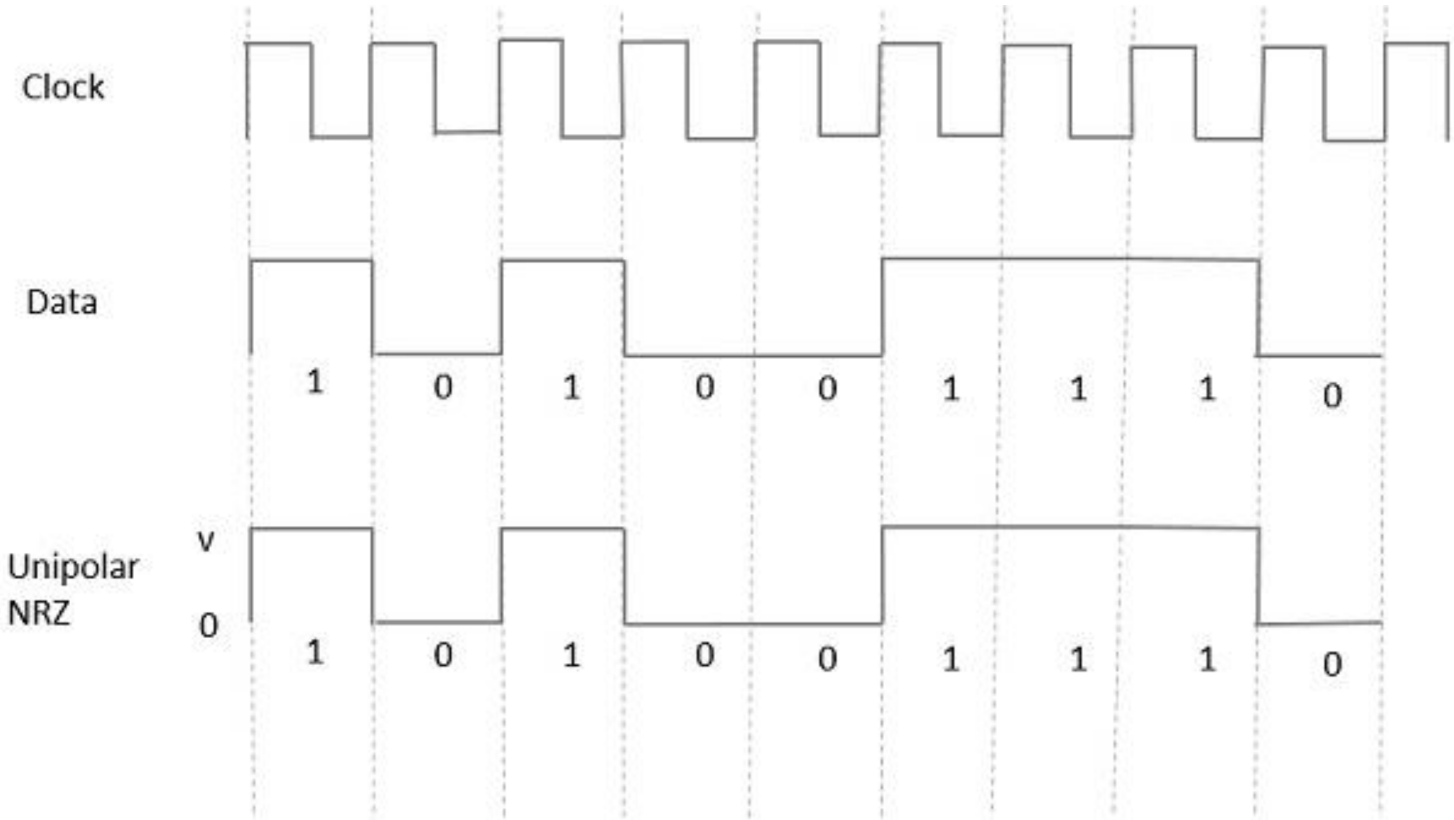
Properties of Line Coding

- As the coding is done to make more bits transmit on a single signal, the bandwidth used is much reduced.
- For a given bandwidth, the power is efficiently used.
- The probability of error is much reduced.
- Error detection is done and the bipolar too has a correction capability.
- Power density is much favorable.
- The timing content is adequate.
- Long strings of **1s** and **0s** is avoided to maintain transparency.

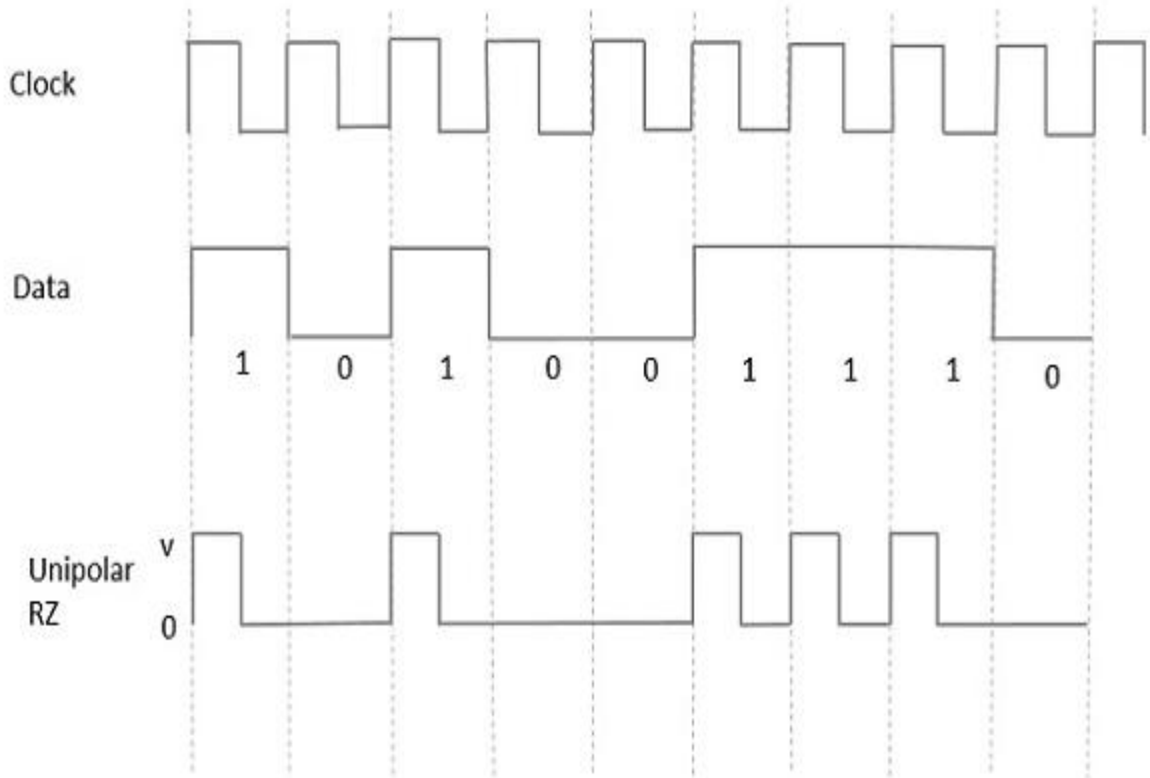
Types of Line Coding

- Unipolar-
  - ➢ Non Return to Zero (NRZ)
  - ➢ Return to Zero (RZ)
- Polar
  - ➢ Polar NRZ
  - ➢ Polar RZ
- Bi-polar
  - ➢ Bipolar NRZ
  - ➢ Bipolar RZ

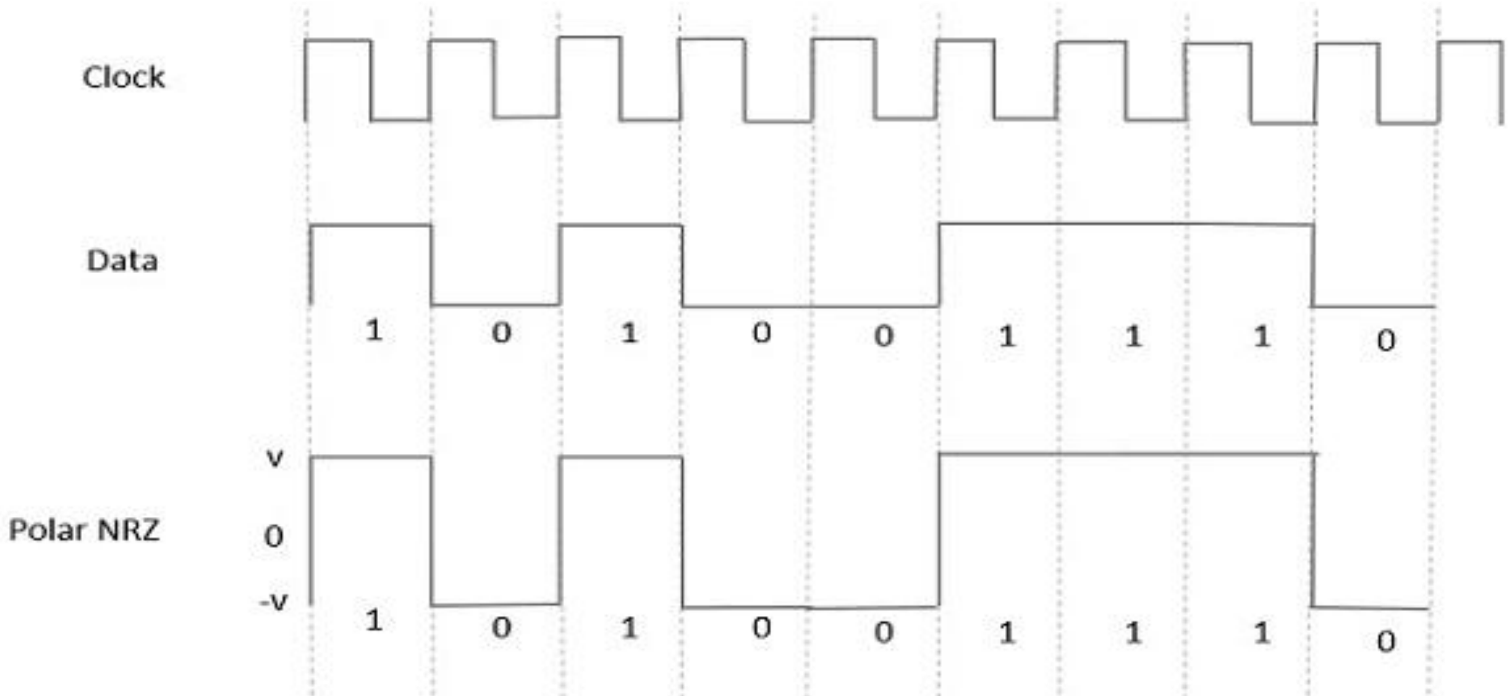
Unipolar Non-Return to Zero (NRZ)



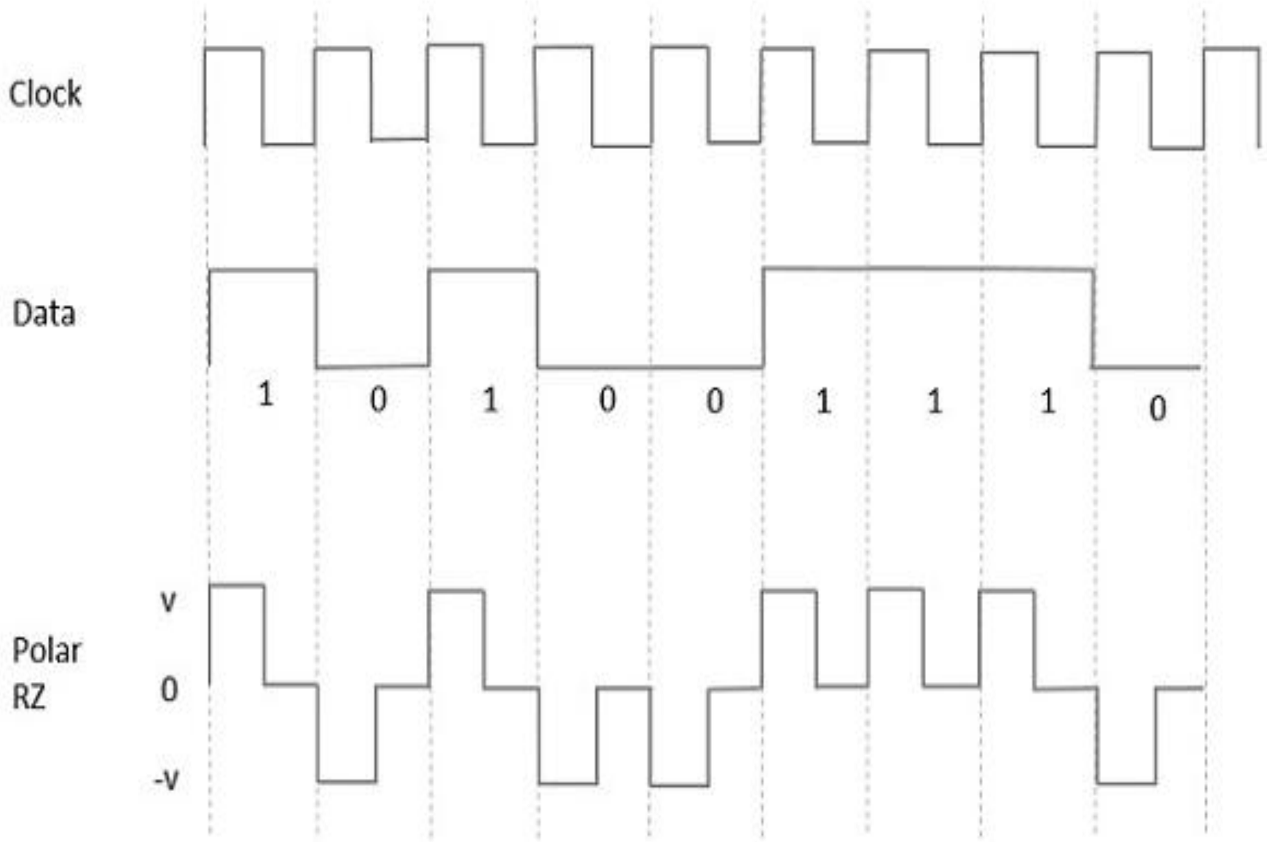
Unipolar Return to Zero (RZ)



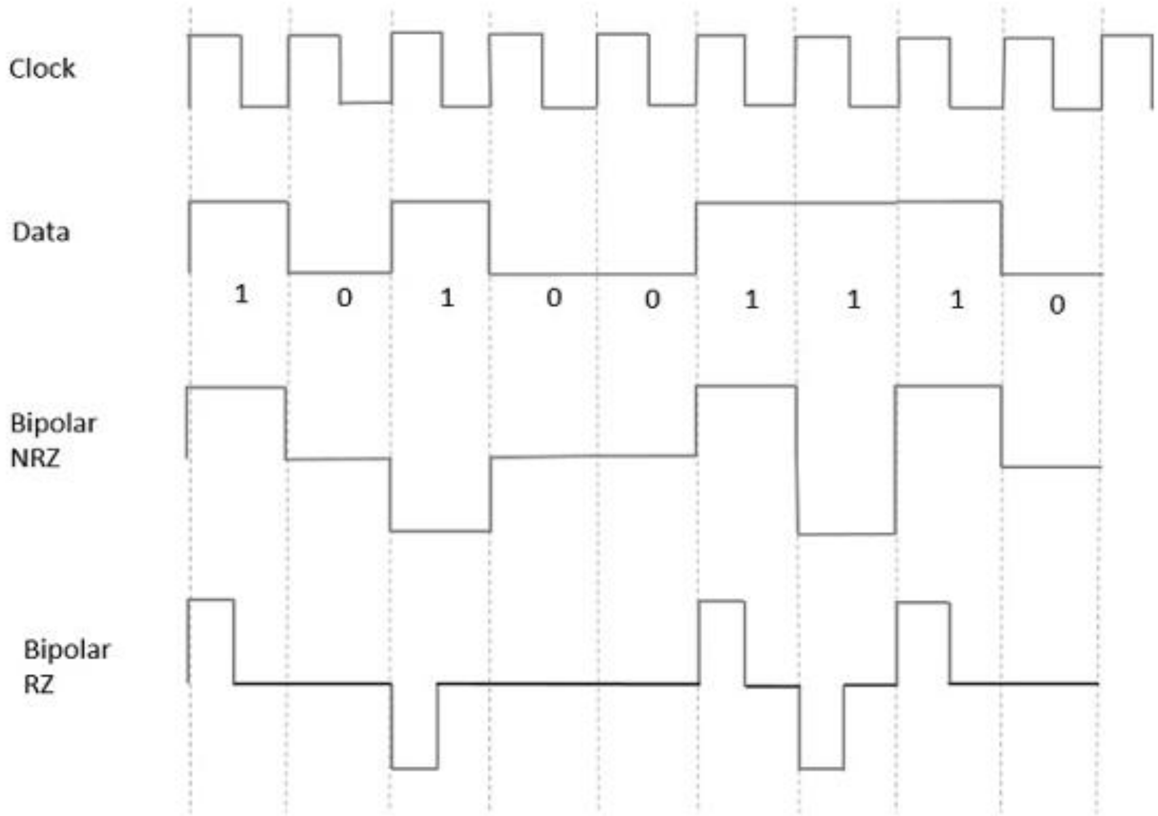
Polar NRZ



Polar RZ



Bipolar Signaling



Probability of Error in ASK

- The expression for BER (or probability of error) normally contains the energy-to-noise ratio (E/N_o)
- The unit energy is:

$$E = ST_b \quad \text{Energy/bit}$$

$$S = \text{Signal power} = A_c^2/2 = E/T_b, \quad \text{Assume } R = 1 \text{ Ohm}$$

Or in terms of signal to noise ratio (SNR) $\dfrac{E}{N_o} = \dfrac{ST_b}{N_o}$

Probability of error for ASK with coherent detection is:

$$P_{e-ASK} = Q\left(\sqrt{\frac{d_{01}^2}{2N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right) = 0.5erfc\left(\sqrt{\frac{E_b}{2N_0}}\right)$$

Probability of Error in ASK Cont..

For multi-level ASK with coherent detection

$$P_{e-MASK} = \frac{2(M-1)}{Mlog_2M} Q\left(\sqrt{\frac{(6log_2M)E_b}{(M^2-1)N_0}}\right) = \frac{(M-1)}{Mlog_2M} erfc\left(\sqrt{\frac{(2log_2M)E_b}{(M^2-1)N_0}}\right)$$

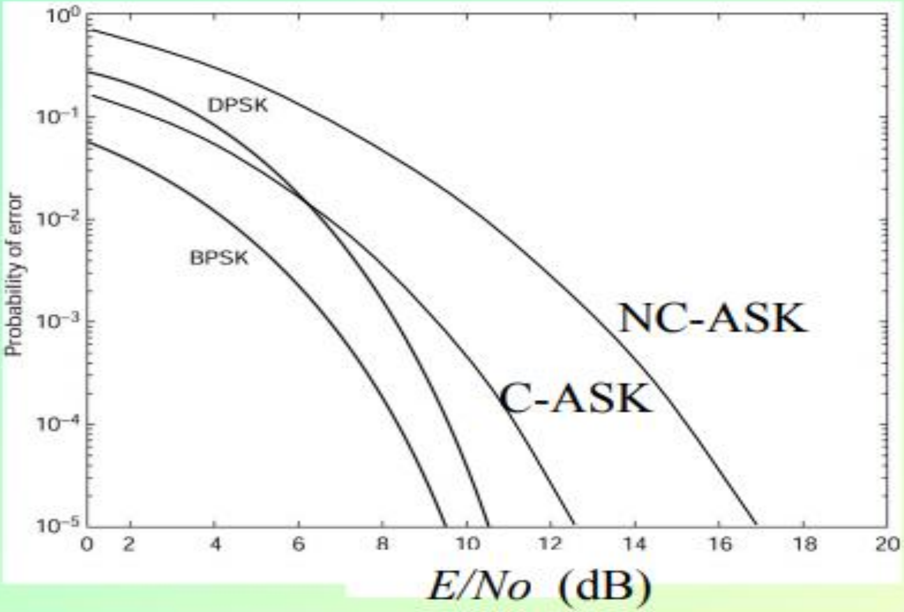
Same transmission bandwidth, yet more information at the cost of higher SNR

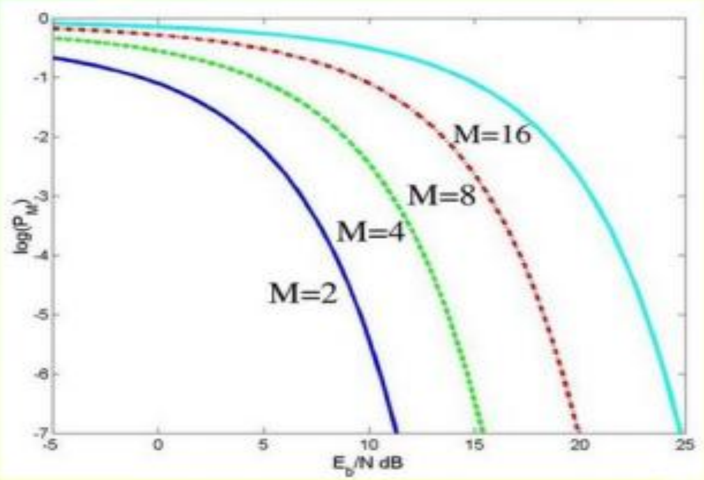
- **Non-Coherent**

$$P_{s-NCASK} = 0.5e^{-(E/4N)} 0.5 erfc\left(\sqrt{\frac{E}{2N_o}}\right)$$

Probability of Error in ASK Cont..

BER Vs. Signal -to-Noise Ratio

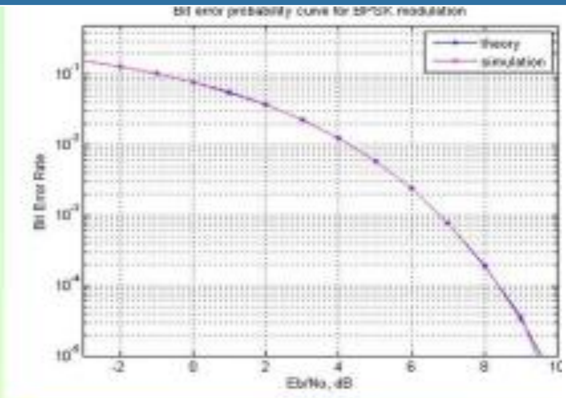




Probability of Error in BPSK

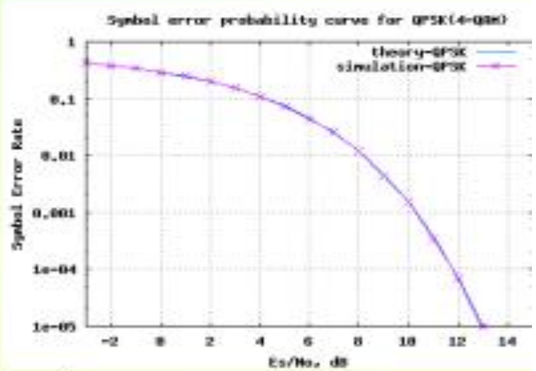
- Coherent

$$P_{e-PSK} = Q\left(\sqrt{\frac{d_{01}^2}{2N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = 0.5erfc\left(\sqrt{\frac{E_b}{N_0}}\right)$$



Symbol error rate

$$P_{s-QPSK} = 2Q\left(\sqrt{\frac{E_s}{N_0}}\right) = erfc\left(\sqrt{\frac{E_s}{2N_0}}\right)$$



$$P_{e-QPSK} = Q\left(\sqrt{\frac{E_s}{N_0}}\right) = 0.5erfc\left(\sqrt{\frac{E_s}{2N_0}}\right) = 0.5erfc\left(\sqrt{\frac{E_s}{N_0}}\right)$$

BPSK and QPSK have the same bit error probability (BER) because QPSK is configured as two BPSK signals modulating orthogonal components of the carrier.

Probability of Error in BPSK Cont..

For multi-level PSK with coherent detection

$$P_{e-MASK} = \frac{2}{log_2 M} Q\left(\sqrt{\frac{E_s}{N_0}} \sin(\frac{\pi}{M})\right) = \frac{1}{log_2 M} erfc\left(\sqrt{\frac{(log_2 M)E_b}{N_0}} \sin(\frac{\pi}{M})\right)$$

Same transmission bandwidth, yet more information at the cost of higher SNR

- The average energy / bit is given as:

$$E = \int_0^{T_b} S_{FSK}(t)^2\, dt = \int_0^{T_b} A_c^{\,2} \sin^2(\omega_c t)\, dt = \frac{A_c^{\,2}}{2} T_b$$

Coherent $$P_{e-FSK} = Q\left(\sqrt{\frac{d_{01}^2}{2N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right) = 0.5erfc\left(\sqrt{\frac{E_b}{2N_0}}\right)$$

Non-coherent $$P_{e-NCFSK} = \frac{1}{2} e^{\frac{-E}{2N_o}}$$

Probability of Error in FSK Cont..

Coherent

Symbol error rate

$$P_{s-MFSK} = (M-1)Q\left(\sqrt{\frac{E_s}{N_0}}\right) \leq 0.5(M-1)erfc\left(\sqrt{\frac{E_s}{2N_0}}\right) \leq 0.5erfc\left(\sqrt{\frac{(log_2 M)E_b}{2N_0}}\right)$$

BER

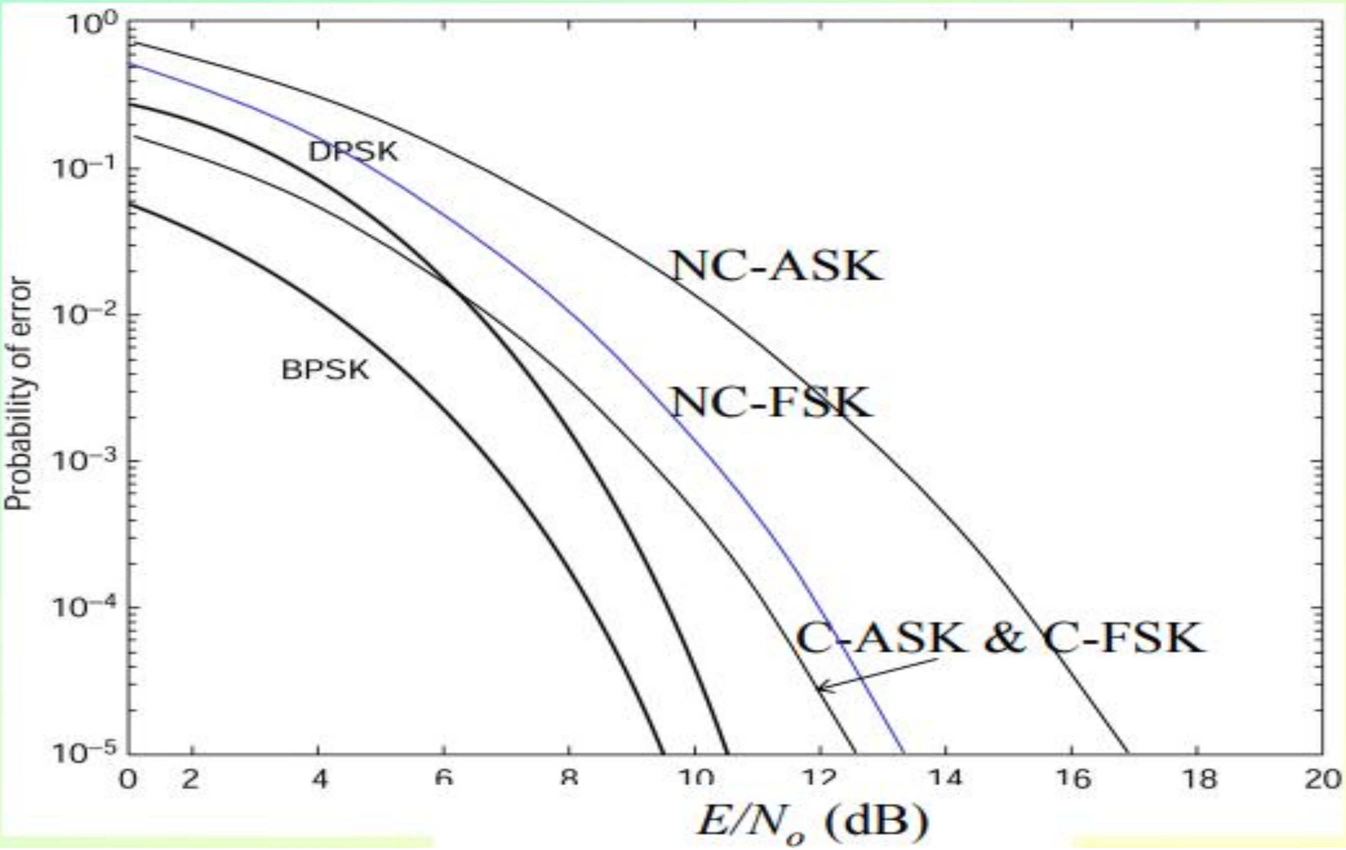
$$P_{e-MFSK} = \frac{M/2}{M-1} P_{s-MFSK}$$

Average BER

$$P_{b-MFSK} \leq 0.5MQ\left(\sqrt{\frac{E_s}{N_0}}\right) \leq 0.25Merfc\left(\sqrt{\frac{E_s}{2N_0}}\right) \leq 0.25Merfc\left(\sqrt{\frac{(log_2 M)E_b}{2N_0}}\right)$$

BER in ASK,BPSK,FSK

- **Equal *E***



BASE BAND TRANSMISSION AND PULSE SHAPING

Baseband Transmission

- Baseband Transmission is a signaling technology that sends digital signals over a single frequency as discrete electrical pulses.

Advantages of Baseband

- Simplicity
- Low cost
- Ease of installation and maintenance

line code

- A line code is the code used for data transmission of a digital signal over a transmission line. This process of coding is chosen so as to avoid overlap and distortion of signal such as inter-symbol interference.
- Properties of Line Coding

Following are the properties of line coding −

- As the coding is done to make more bits transmit on a single signal, the bandwidth used is much reduced.
- For a given bandwidth, the power is efficiently used.
- The probability of error is much reduced.
- Error detection is done and the bipolar too has a correction capability.
- Power density is much favorable.
- The timing content is adequate.

line code cont…

Long strings of 1s and 0s is avoided to maintain transparency.

Types of Line Coding

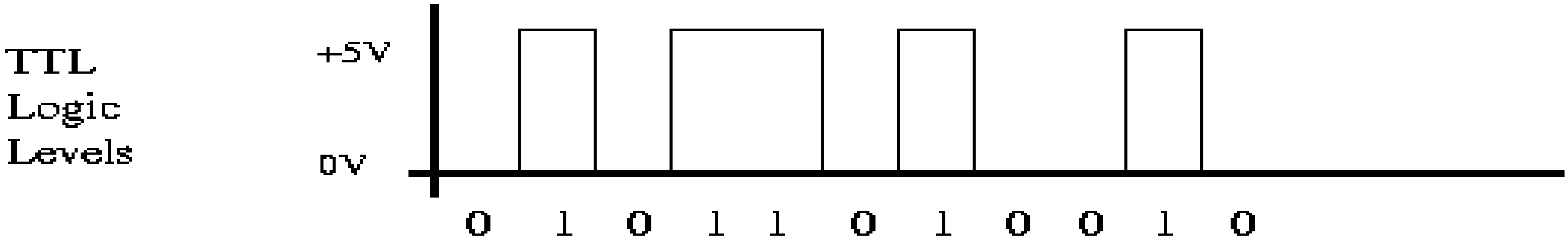
There are 3 types of Line Coding

- Unipolar
- Polar
- Bi-polar

Unipolar Encoding: Unipolar encoding has 2 voltage states, with one of the states being 0 volts.

- Since Unipolar line encoding has one of its states at 0 Volts, it is also called Return to Zero (RTZ).
- A common example of Unipolar line encoding is the TTL logic levels used in computers and digital logic.

Unipolar Encoding cont..



•Unipolar line encoding works well for inside machines--where the signal path is short-- but is unsuitable for long distances, due to the presence of stray capacitance in the transmission medium.

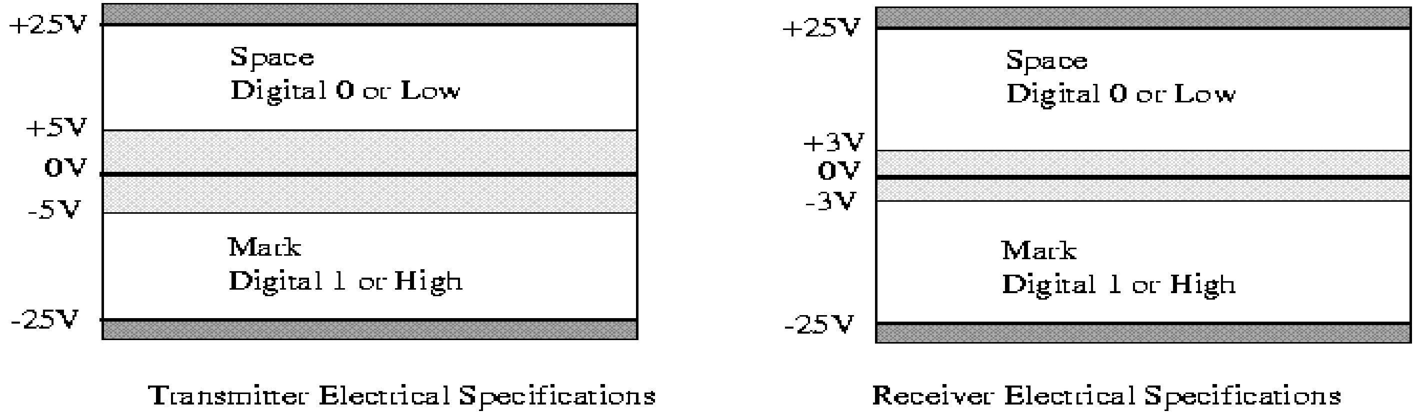
• On long transmission paths, the constant level shift from 0 to 5 volts, which causes the stray capacitance to charge up (remember, the capacitor charging formula is: 1-e-t/RC !).

Polar Encoding

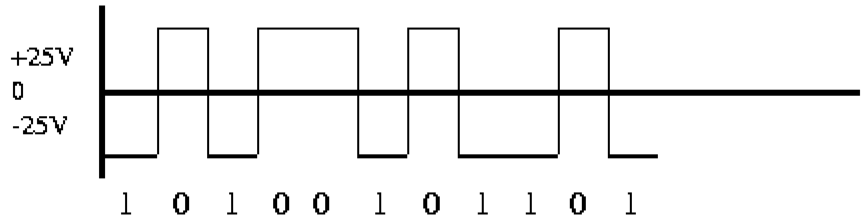
•When The digital encoding is symmetrical--around 0 Volts--it is called a Polar Code. For example, the RS-232D interface uses Polar line encoding.

•The signal does not return to zero; it is either a +ve voltage or a -ve voltage. Polar line encoding is also called None Return To Zero (NRZ).

• Polar line encoding is the simplest pattern that eliminates most of the residual DC problem.



Polar Encoding cont..

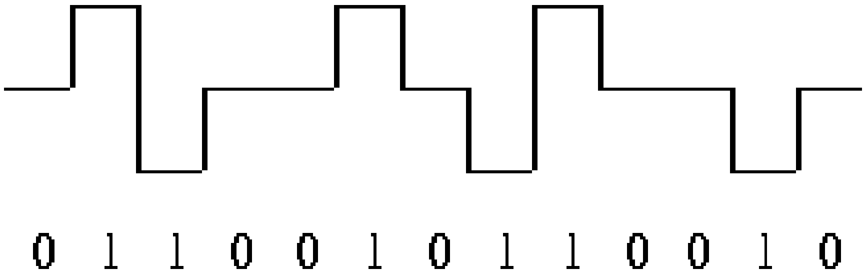


Bipolar Line Encoding:

•Bipolar line encoding has 3 voltage levels. A low or 0 is represented by a 0 Volt level.

•a 1 is represented by alternating polarity pulses.

•By alternating the polarity of the pulses for 1s, the residual DC component cancels.



Bipolar Line Encoding

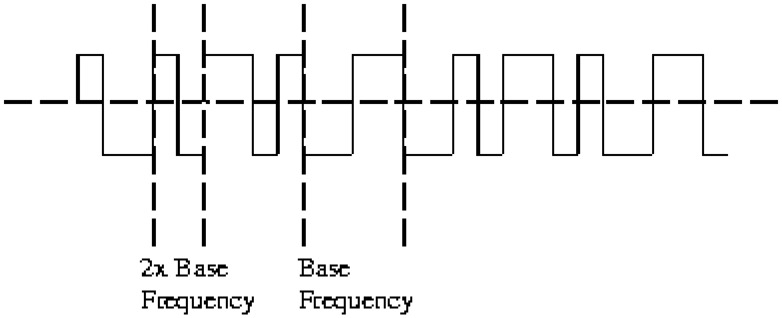
•Synchronization of receive and transmit clocks is greatly improved--except if there is a long string of 0s transmitted.

•Bipolar line encoding is also called Alternate Mark Inversion (AMI).

Manchester Line Encoding

•In Manchester Line Encoding, there is a transition at the middle of each bit period.

•The mid-bit transition serves as a clocking mechanism (and also as data) a low to high transition represents a 1 and a high to low transition represents a 0.



What is Scrambling?

- Computer network is designed to send information from one point to another. Data that we send can either be digital or analog.
- Also signals that represent data can also be digital or analog.
- we must able to convert data into signals this conversion can be Analog to Analog, Analog to Digital, Digital to Analog or Digital to Digital.
- Digital to Digital conversion involves three techniques – Line Coding, Block Coding, and Scrambling.
- Line coding is always needed, whereas Block Coding and Scrambling may or may not be needed depending upon need.

Scrambling

•Scrambling is a technique that does not increase the number of bits and does provide synchronization.

•Problem with technique like Bipolar AMI(Alternate Mark Inversion) is that continuous sequence of zero’s create synchronization problems one solution to this is Scrambling.

There are two common scrambling techniques:

•B8ZS(Bipolar with 8-zero substitution)

•HDB3(High-density bipolar3-zero)

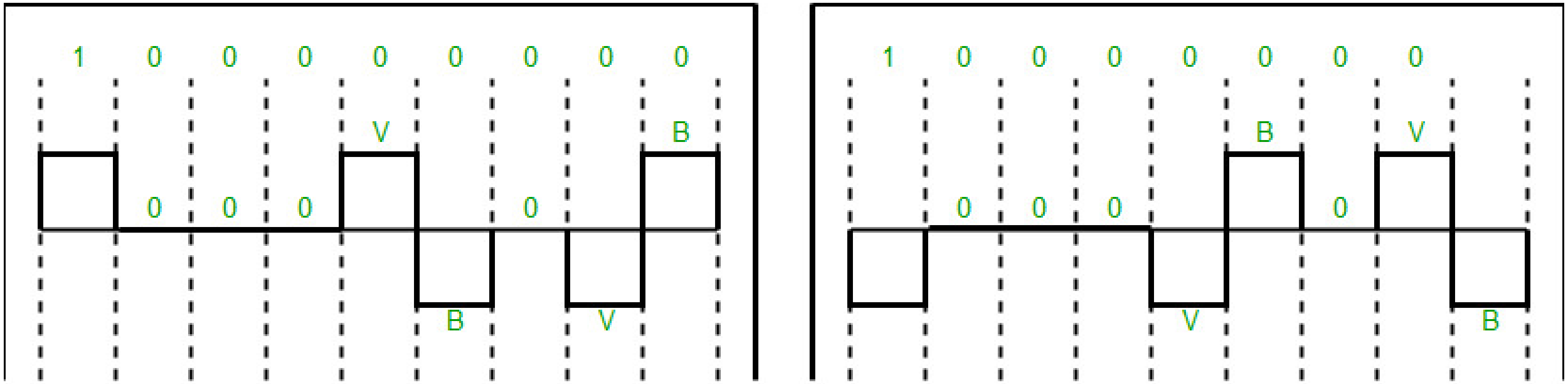
DESCRIPTION

- B8ZS(Bipolar with 8-zero substitution) –This technique is similar to Bipolar AMI except when eight consecutive zero-level voltages are encountered they are replaced by the sequence,"000VB0VB".

- Note –V(Violation), is a non-zero voltage which means signal have same polarity as the previous non-zero voltage. Thus it is violation of general AMI technique.

- B(Bipolar), also non-zero voltage level which is in accordance with the AMI rule (i.e.,opposite polarity from the previous non-zero voltage).

Scrambling



Example: Data = 100000000

Note – Both figures (left and right one) are correct, depending upon last non-zero voltage signal of previous data sequence (i.e., sequence before current data sequence "100000000").

HDB3(High-density bipolar3-zero)

- In this technique four consecutive zero-level voltages are replaced with a sequence “000V” or “B00V”.

Rules for using these sequences:

- If the number of non-zero pulses after the last substitution is odd, the substitution pattern will be “000V”, this helps maintaining total number of nonzero pulses even.

- If the number of non-zero pulses after the last substitution is even, the substitution pattern will be “B00V”. Hence even number of nonzero pulses is maintained again.

HDB3(High-density bipolar3-zero)

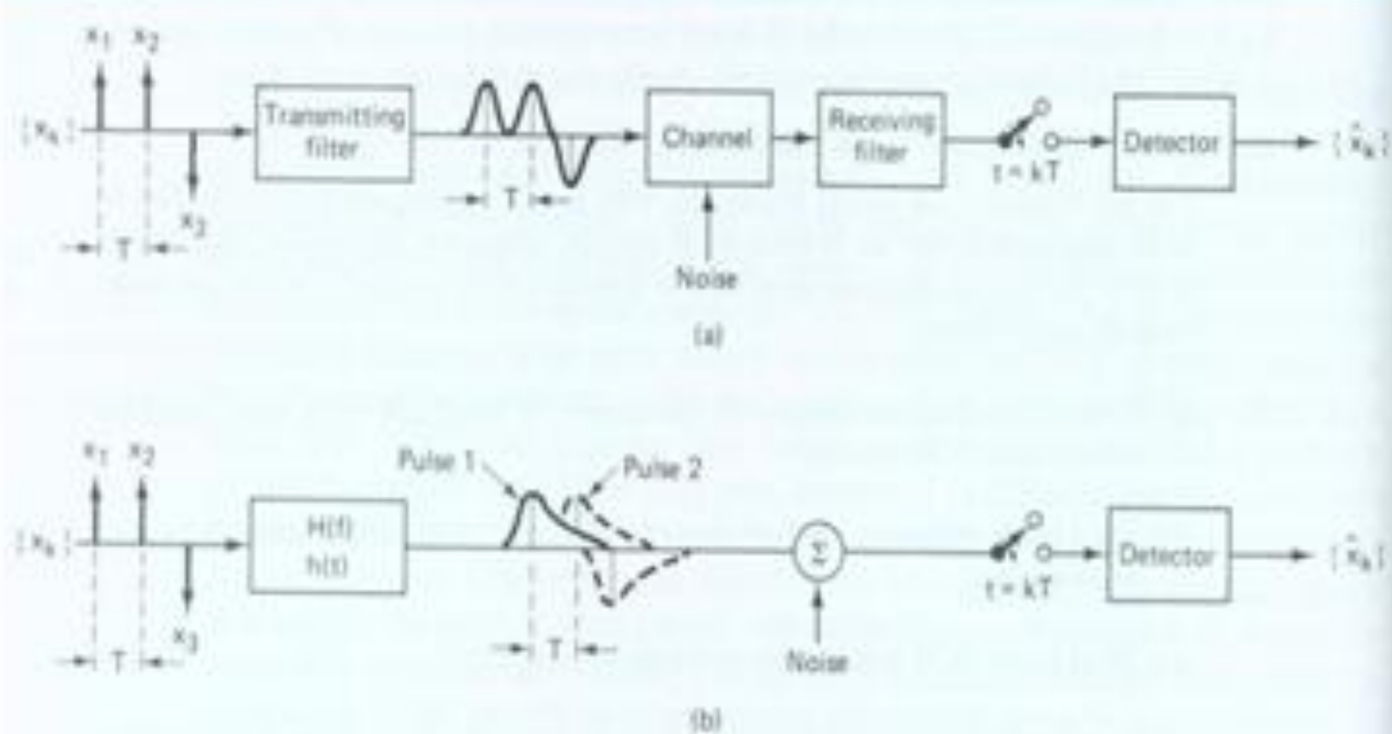


Figure 2.32 Intersymbol interference in the detection process. (a) Typical baseband digital system. (b) Equivalent model.

Power spectral density

- The power spectral density (PSD) of the signal describes the power present in the signal as a function of frequency, per unit frequency.

- Power spectral density is commonly expressed in watts per hertz (W/Hz). .

Power spectral density of Unipolar NRZ line code

•Unipolar is the simplest line coding scheme which has the advantage of being compatible with TTL logic.

•Unipolar coding uses a positive rectangular pulse $p(t)$ to represent binary **1**, and the absence of a pulse (i.e., zero voltage) to represent a binary **0**.

Two possibilities for the pulse $p(t)$ exist:

• Non-Return-to-Zero (NRZ) rectangular pulse and Return-to-Zero (RZ) rectangular pulse.

•The difference between Unipolar NRZ and Unipolar RZ codes is that the rectangular pulse in NRZ stays at a positive value (e.g., +5V) for the full duration of the logic **1** bit, while the pulse in RZ drops from +5V to 0V in the middle of the bit time.

Unipolar and Polar RZ & NRZ

•A drawback of unipolar (RZ and NRZ) is that its average value is not zero, which means it creates a significant DC-component at the receiver .

• Polar signals have more power than unipolar signals, and hence have better SNR at the receiver.

•Polar NRZ signals have more power compared to polar RZ signals.

• The drawback of polar NRZ, is that it lacks clock information especially when a long sequence of **0**‟s or **1**‟s is transmitted.

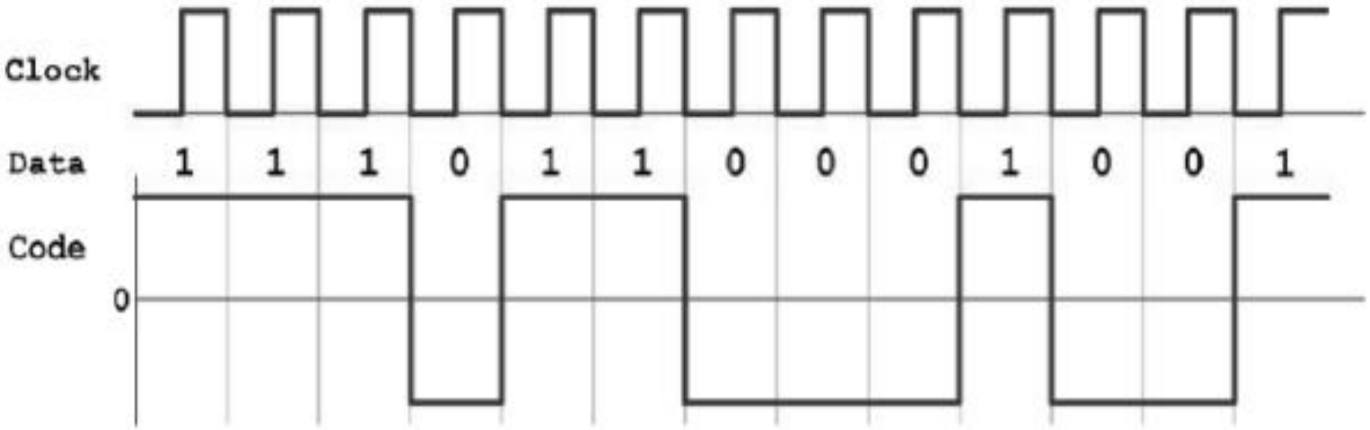
Non-Return-to-Zero, Inverted (NRZI):

- **Non-Return-to-Zero, Inverted (NRZI):** NRZI is a variant of Polar NRZ. In NRZI there are two possible pulses, $p(t)$ and $-p(t)$.

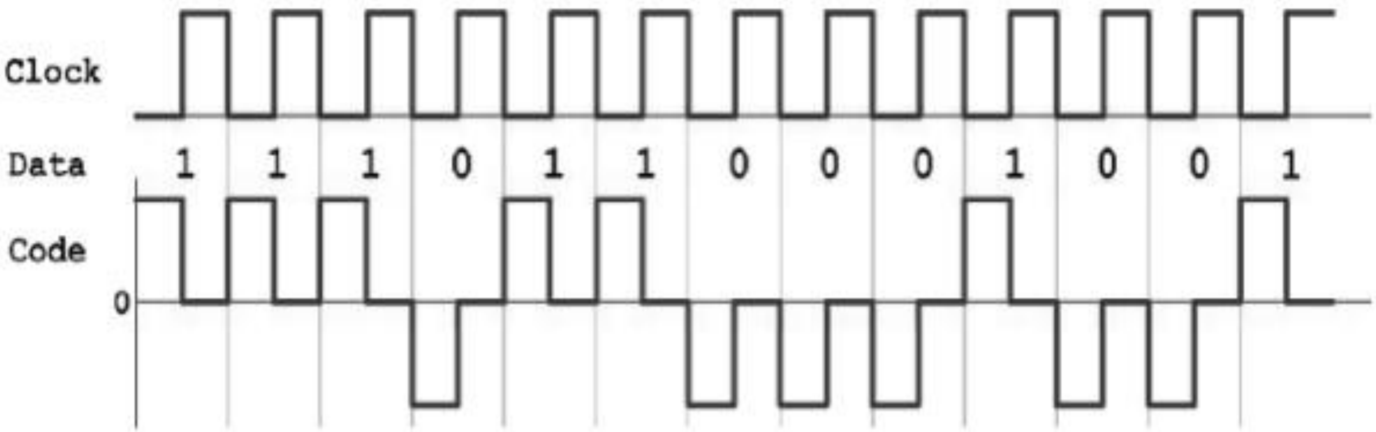
- A transition from one pulse to the other happens if the bit being transmitted is logic **1**,

- No transition happens if the bit being transmitted is a logic **0**.

Polar NRZ & RZ



Polar NRZ Code



Polar RZ Code

Pulse shaping

• Pulse shaping is the process of changing the waveform of transmitted pulses.

•Its purpose is to make the transmitted signal better suits to the communication channel, typically by limiting the effective bandwidth of the transmission.

•By filtering the transmitted pulses, the intersymbol interference caused by the channel is kept under control.

•In RF communication, pulse shaping is essential for making the signal fit in its frequency band.

•Typically pulse shaping occurs after Line coding and Modulation.

Need for pulse shaping

•Transmitting a signal at high modulation rate through a band-limited channel can create intersymbol interference.

•As the modulation rate increases, the signal's bandwidth increases.

•When the signal's bandwidth becomes larger than the channel bandwidth, the channel starts to introduce distortion to the signal.

•This distortion usually manifests itself as intersymbol interference.

•Transmitted symbols are usually represented as a time sequence of dirac delta pulses.

• This theoretical signal is then filtered with the pulse shaping filter, producing the transmitted signal.

Pulse shaping

•In base band communication systems, the pulse shaping filter is implicitly a boxcar filter.

• Its Fourier transform is of the form *sin(x)/x*, and has significant signal power at frequencies higher than symbol rate.

•However, in RF communications this would waste bandwidth, and only specified frequency bands are used for single transmissions.

•Therefore, the channel for the signal is band-limited. Thus, better filters have been developed to minimise the bandwidth needed for a certain symbol rate

Inter Symbol Interference (ISI)

•Inter-Symbol Interference (ISI) arises because of imperfections in the overall frequency response of the system.

•When a short pulse (Tb secs), is transmitted through a band-limited system, the frequency components constituting the input pulse are differentially attenuated and differentially delayed by the system.

•Consequently, the pulse appearing at the output of the system is dispersed over an interval longer than Tb seconds, thereby resulting in intersymbol interference.

•Even in the absence of noise, imperfect filtering and system bandwidth constraints lead to ISI.

Nyquist Channels for Zero ISI

- Nyquist channel is not physically realizable since it dictates a rectangular bandwidth characteristic and an infinite time delay.
- Detection process would be very sensitive to small timing errors.

Solution: Raised Cosine Filter.

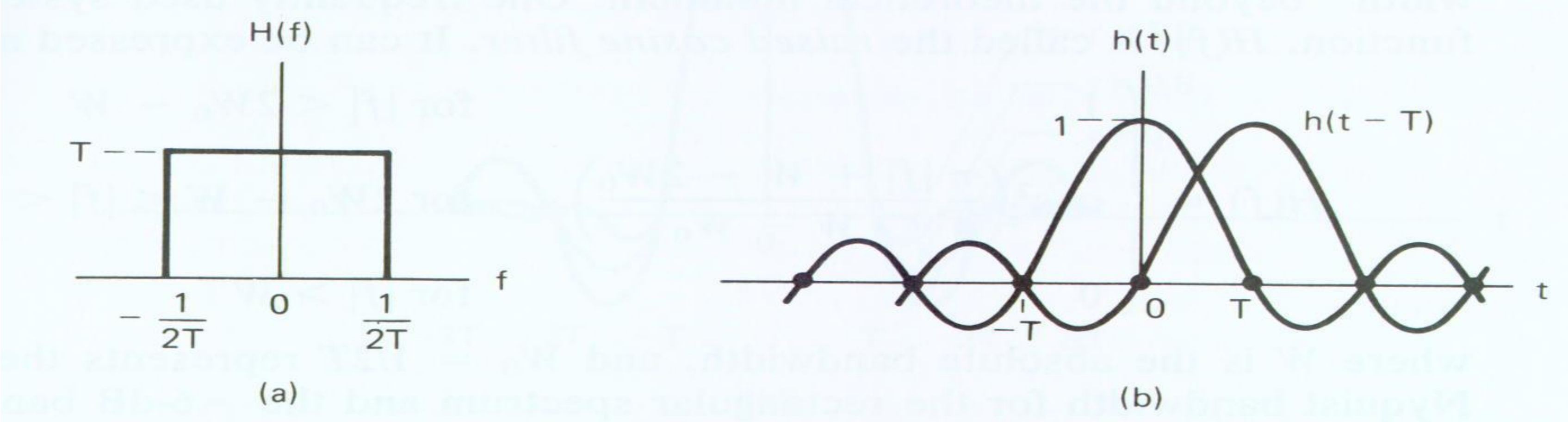


Figure 2.33 Nyquist channels for zero ISI. (a) Rectangular system transfer function $H(f)$. (b) Received pulse shape $h(t) = \text{sinc}(t/T)$.

Raised Cosine Filter

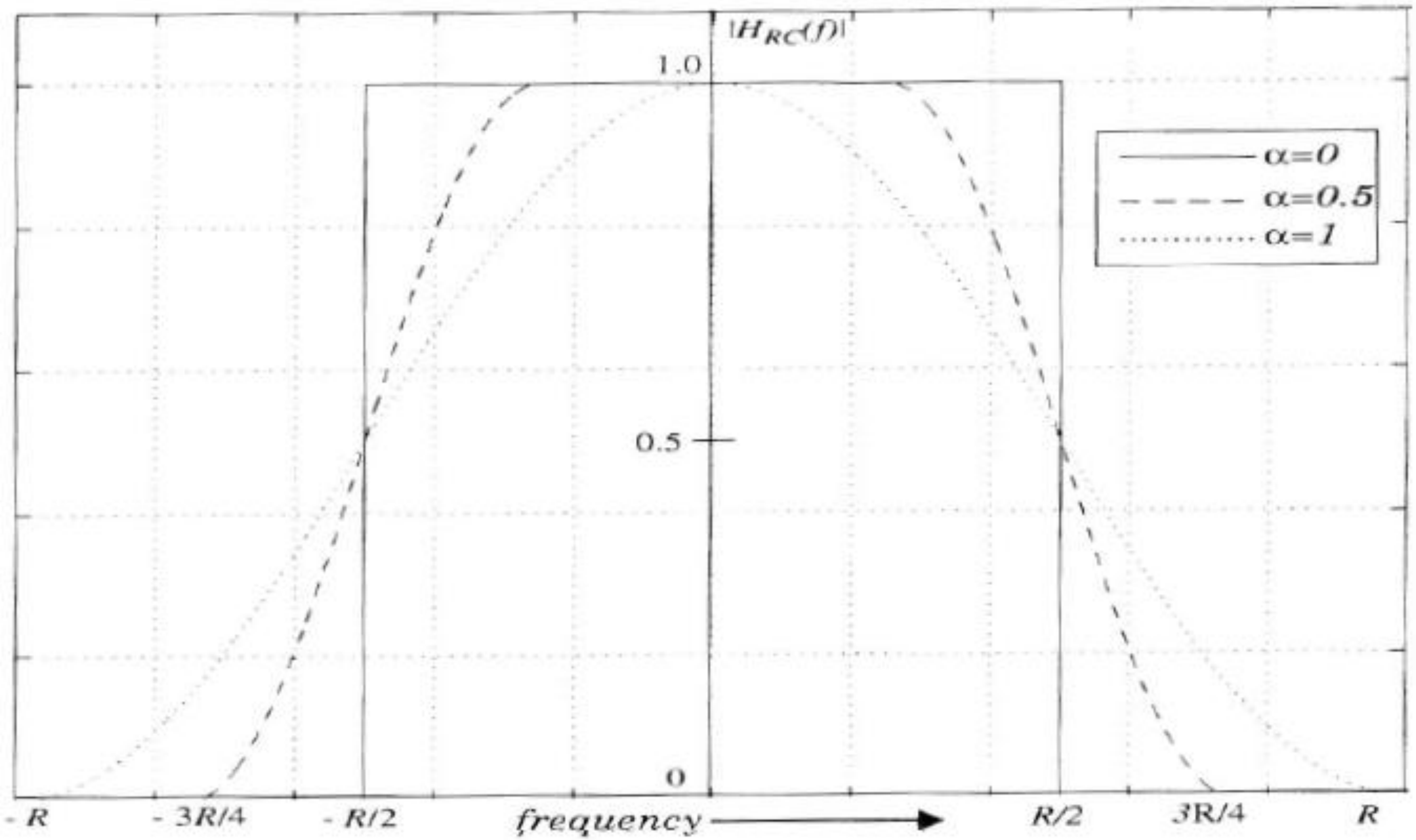
$$H(f)=\begin{cases} 1 & \text{for } |f|<2W_0-W \\ \cos^2\left(\frac{\pi}{4}\frac{|f|+W-2W_0}{W-W_0}\right) & \text{for } 2W_0-W<|f|<W \\ 0 & \text{for } |f|>W \end{cases}$$

$$W_0=\frac{1}{2T}$$

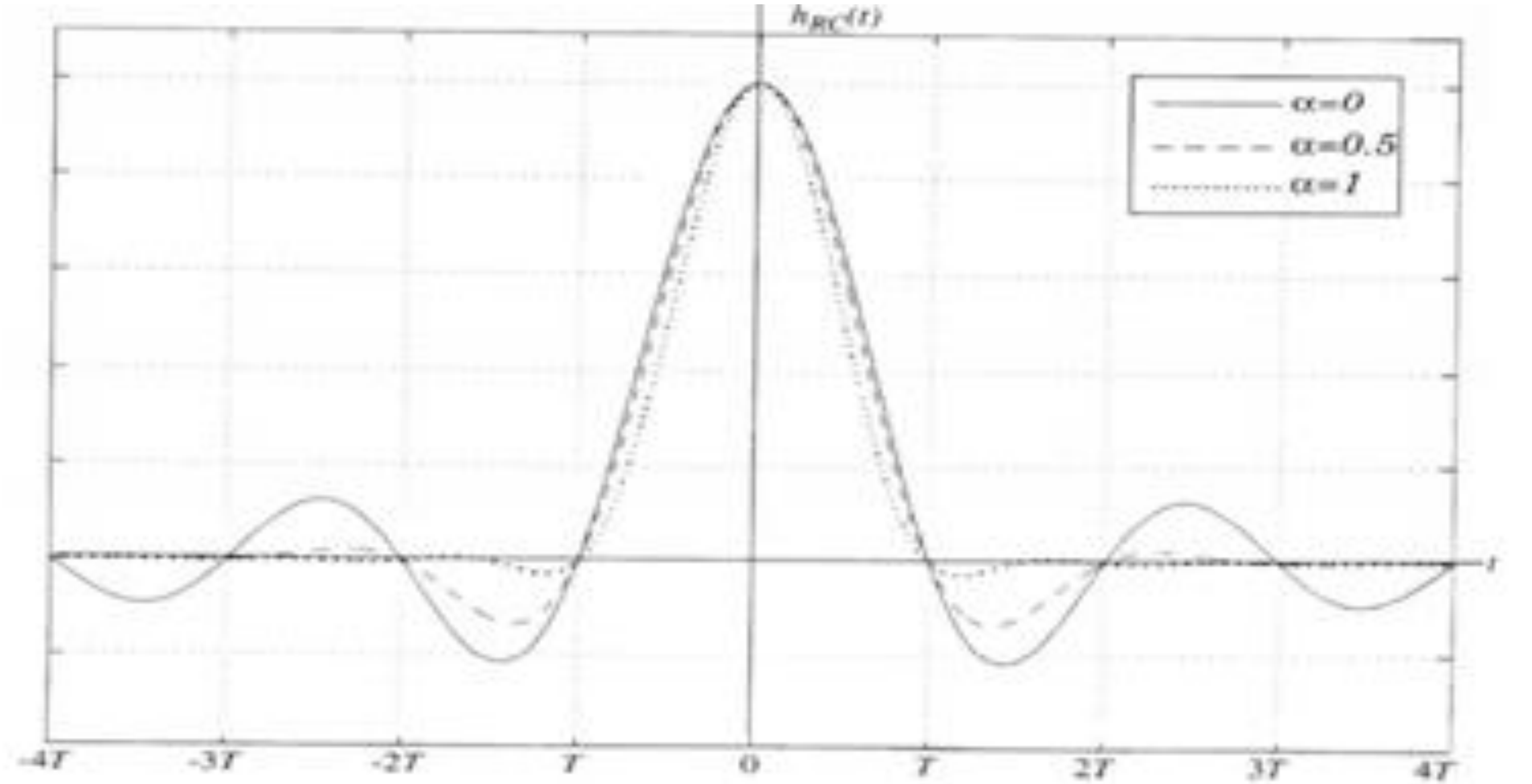
$$\text{Excess Bandwidth}: W-W_0$$

$$\text{Roll-Off Factor}:\ r=\frac{W-W_0}{W_0}$$

Raised Cosine Filter Characteristics



Raised Cosine Filter Characteristics



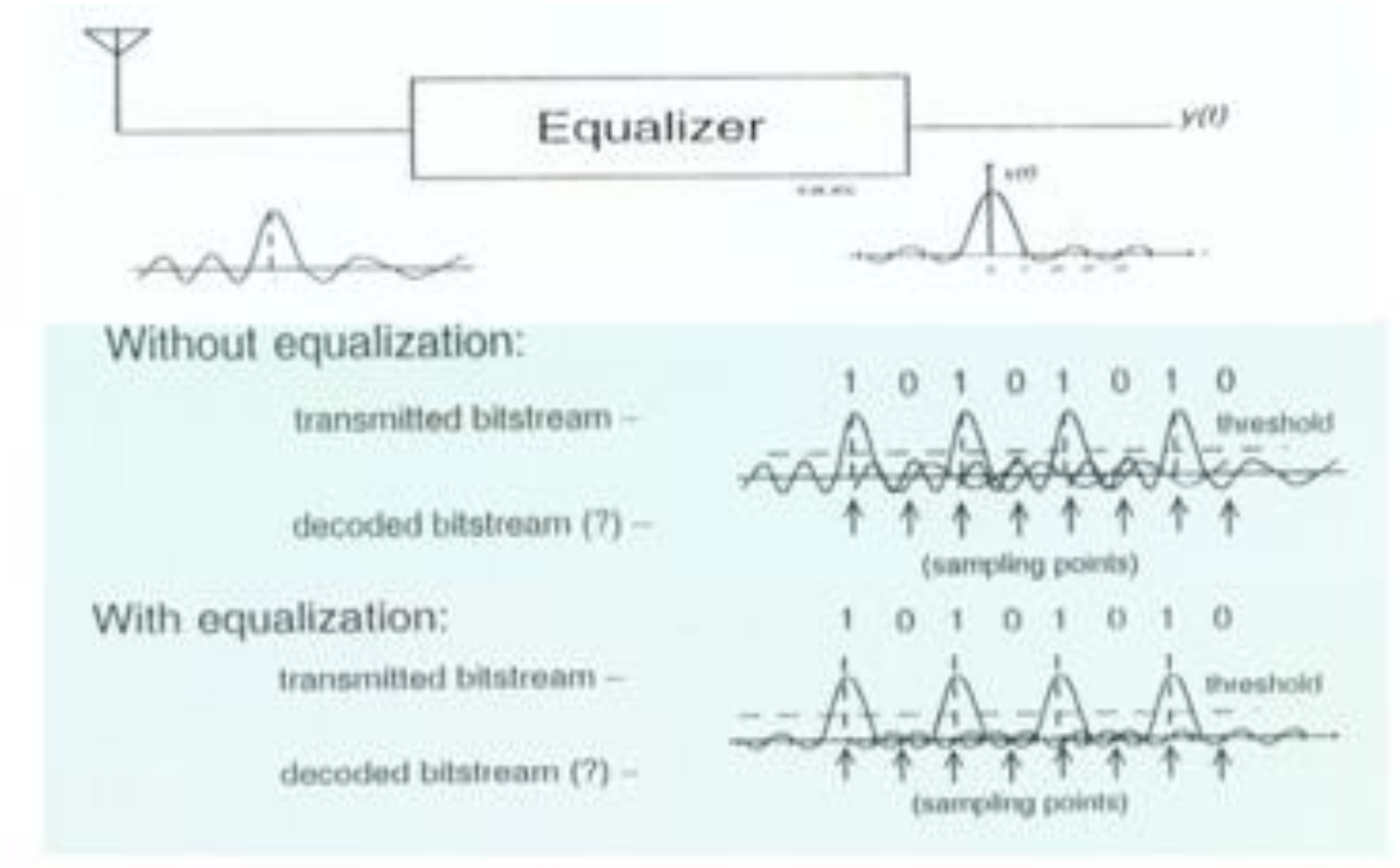
Equalization

- In practical systems, the frequency response of the channel is not known to allow for a receiver design that will compensate for the ISI.

- The filter for handling ISI at the receiver contains various parameters that are adjusted with the channel characteristics.

- The process of correcting the channel-induced distortion is called equalization.

Equalization



Correlative-Level Coding

- Correlative-level coding (partial response signaling)

- Adding ISI to the transmitted signal in a controlled manner

- Since ISI introduced into the transmitted signal is known, its effect can be interpreted at the receiver

- A practical method of achieving the theoretical maximum signaling rate of 2W symbol per second in a bandwidth of W Hertz

- Using realizable and perturbation-tolerant filters

Duobinary Signaling

- Duobinary Signaling

  Doubling of the transmission capacity of a straight binary system

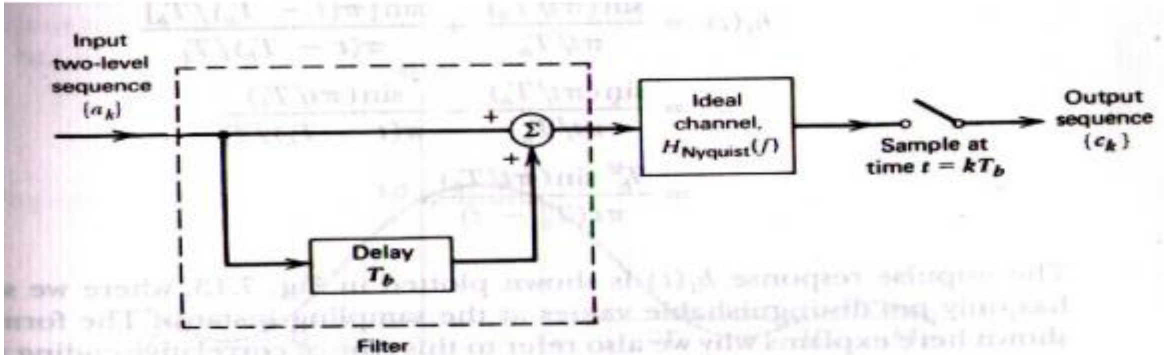


Figure 7.11 **Duobinary signaling scheme.**

- Binary input sequence {bk } : uncorrelated binary symbol 1, 0

$$a_k = \begin{cases} +1 & \text{if symbol } b_k \text{ is } 1 \\ -1 & \text{if symbol } b_k \text{ is } 0 \end{cases} \qquad c_k = a_k + a_{k-1}$$

Correlative-Level Coding

- Duobinary Signaling: The tails of $h_I(t)$ decay as $1/|t|^2$, which is a faster rate of decay than $1/|t|$ encountered in the ideal Nyquist channel.

- Decision feedback : technique of using a stored estimate of the previous symbol

- Propagate : drawback, once error are made, they tend to propagate through the output

- Precoding : practical means of avoiding the error propagation phenomenon before the duobinary coding

Correlative-Level Coding

Duobinary Signaling

$$d_k = b_k \oplus d_{k-1}$$

D_k is applied to a pulse-amplitude modulator, producing a corresponding two-level sequence of short pulse $\{a_k\}$,
where +1 or –1 as before.

c_k|=1 : random guess in favor of symbol 1 or 0

If $|\,c_k\,| = 1$, *say symbol* b_k *is* 1

If $|\,c_k\,| = 1$, *say symbol* b_k *is* 0

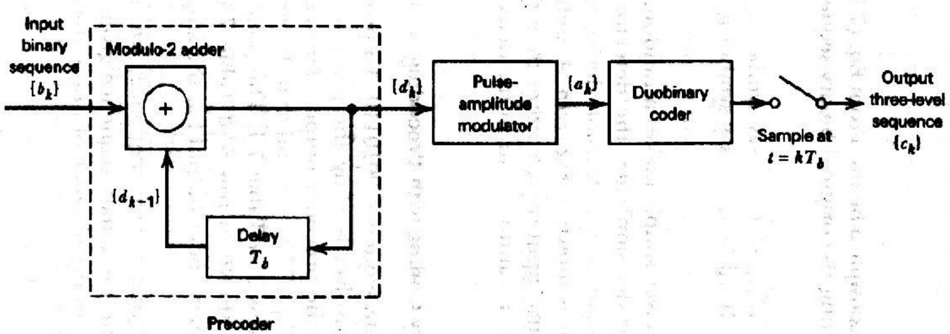


Figure 7.14 A precoded duobinary scheme; details of the duobinary coder are given in Figure 7.11.

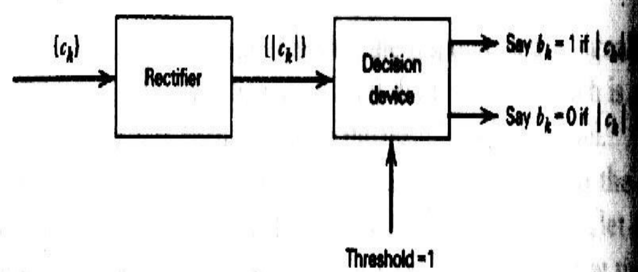


Figure 7.15 Detector for recovering original binary seque from the precoded duobinary coder output.

Modified Duobinary Signaling

- Nonzero at the origin : undesirable
- Subtracting amplitude-modulated pulses spaced $2T_b$ second

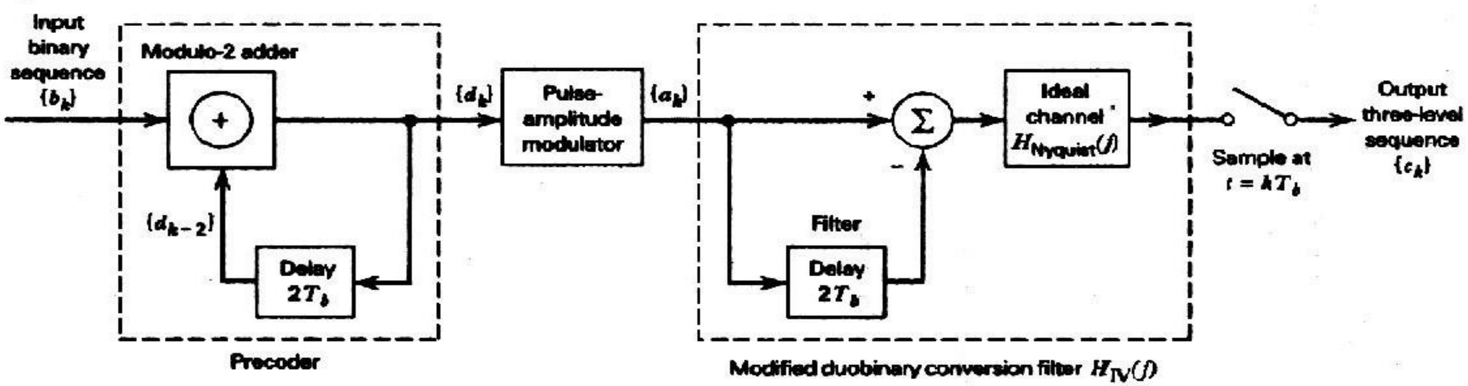


Figure 7.16 Modified duobinary signaling scheme.

Eye diagrams

- In telecommunication, an **eye pattern**, also known as an **eye diagram**, is an oscilloscope display in which a digital signal from a receiver is repetitively sampled and applied to the vertical input.

- It is so called because, for several types of coding, the pattern looks like a series of eyes between a pair of rails.

- It is a tool for the evaluation of the combined effects of channel noise and intersymbol interference on the performance of a baseband pulse-transmission system.

- It is the synchronised superposition of all possible realisations of the signal of interest viewed within a particular signaling interval.

Eye diagrams

- Several system performance measures can be derived by analyzing the display.

- If the signals are too long, too short, poorly synchronized with the system clock, too high, too low, too noisy, or too slow to change, or have too much undershoot or overshoot, this can be observed from the eye diagram.

- An open eye pattern corresponds to minimal signal distortion.

- Distortion of the signal waveform due to intersymbol interference and noise appears as closure of the eye pattern

Features of Eye diagrams

Eye-diagram feature	What it measures
Eye opening (height, peak to peak)	Additive noise in the signal
Eye overshoot/undershoot	Peak distortion due to interruptions in the signal path
Eye width	Timing synchronization & jitter effects
Eye closure	Intersymbol interference, additive noise

Digital Carrier Modulation

Eye diagrams

Pulse amplitude modulation (PAM)

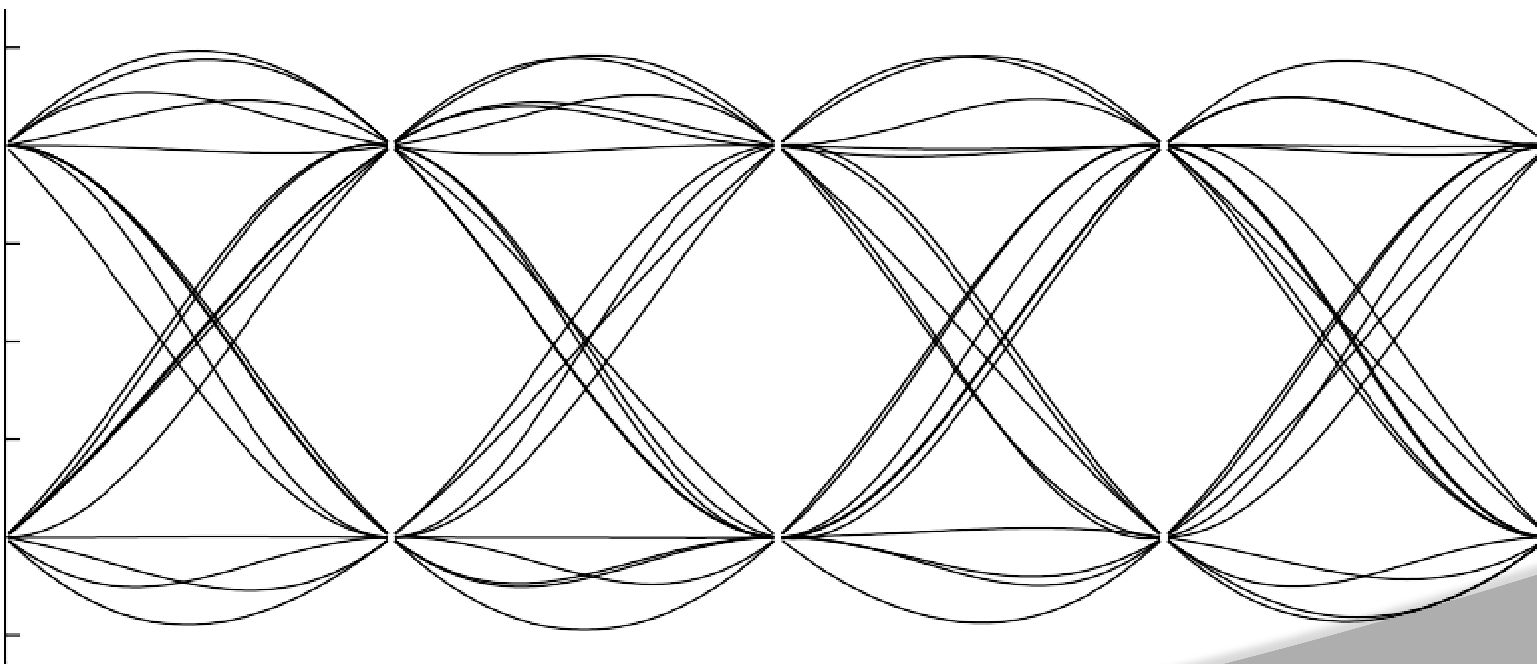
Binary digital modulation

- Amplitude shift keying (ASK)
- Frequency shift keying (FSK)
- Phase shift keying (PSK)
- Quadrature PSK and QAM

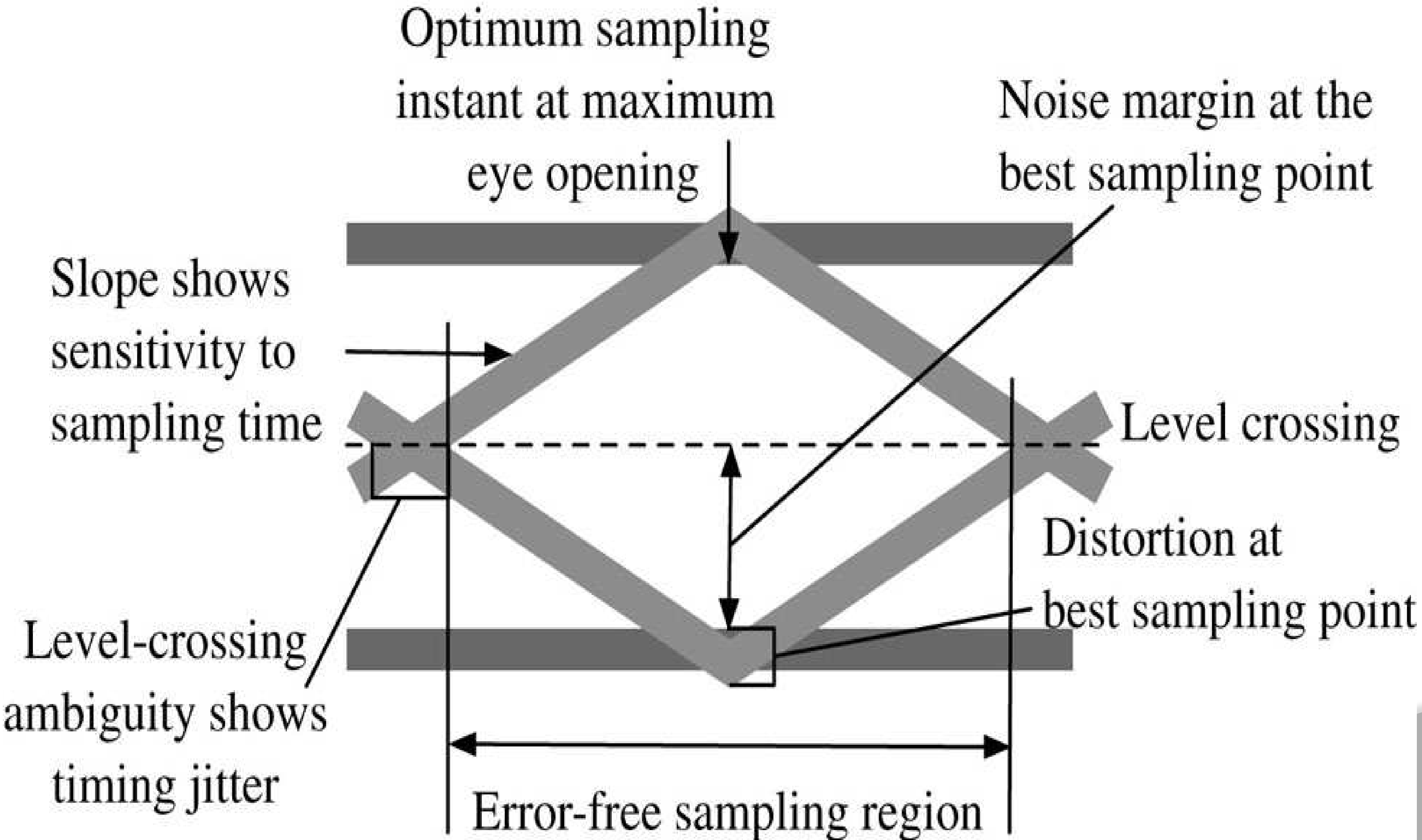
Polar Signaling with Raised Cosine Transform ($r = 0.5$)

The pulse corresponding to $P(t)$ is

$$p(t) = \mathrm{sinc}(\pi R_b t)\,\frac{b}{1 - 4r^2 R_b^2 t^2}$$



Eye Diagram Measurements



Eye Diagram Measurements

- Maximum opening affects noise margin
- Slope of signal determines sensitivity to timing jitter
- Level crossing timing jitter affects clock extraction
- Area of opening is also related to noise margin

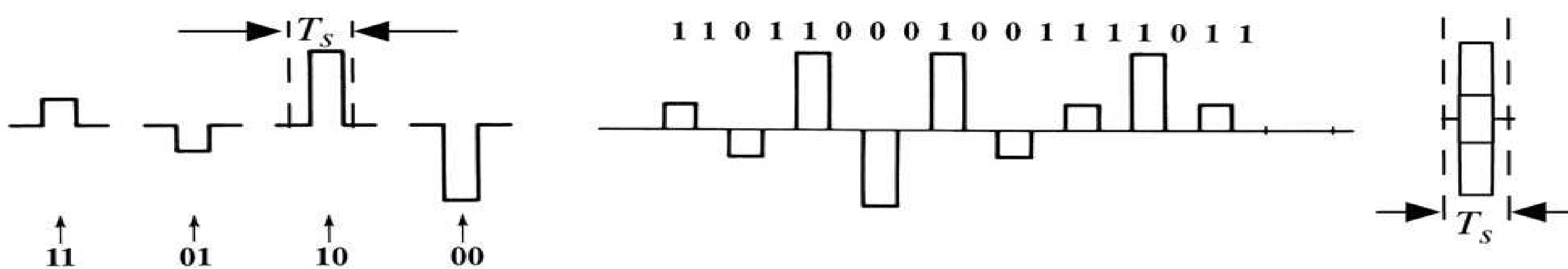
We can generalize polar signaling to

$$y(t) = \quad a\, p(t - kT\ )$$

where a_k is chosen from a set of more than two values (i.e., not just ±1).

Example: one widely used encoding of two bits into four levels is

$$\begin{array}{ll} -3 & \text{message bits } 00 \\ -1 & \text{message bits } 01 \\ +1 & \text{message bits } 11 \\ +3 & \text{message bits } 10 \end{array}$$



UNIT-IV

INFORMATION THEORY AND SOURCE CODING

WHAT IS INFORMATION

- Information is the source of a communication system, whether it is analog or digital.

- Information theory is a mathematical approach to the study of coding of information along with the quantification, storage, and communication of information.

INFORMATION

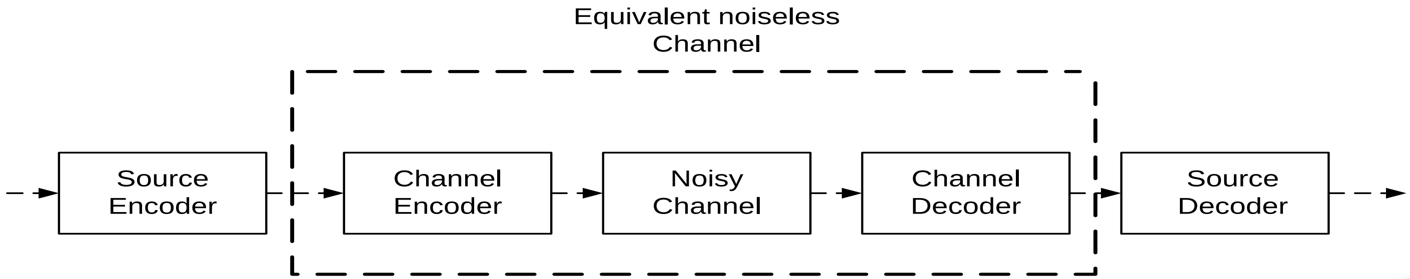
- Information Theory deals with

1.) The Measure of Source Information

2.) The Information Capacity of the channel

3.) Coding

- If The rate of Information from a source does not exceed the capacity of the Channel, then there exist a Coding Scheme such that Information can be transmitted over the Communication Channel with arbitrary small amount of errors despite the presence of Noise



INFORMATION MEASURE

- This is utilized to determine the information rate of discrete Sources

Consider two Messages

- A Dog Bites a Man → High probability → Less information
- A Man Bites a Dog → Less probability → High Information

So we can say that.

- Information α (1/Probability of Occurrence)

INFORMATION MEASURE

Also we can state the three law from Intution,

- Rule 1: Information $I(m_k)$ approaches to 0 as P_k approaches infinity.
  Mathematically $I(m_k) = 0$ as $P_k \rightarrow 1$
  e.g. Sun Rises in East

- Rule 2: The Information Content $I(m_k)$ must be Non Negative contity.
- It may be zero
  Mathematically $I(m_k) >= 0$ as $0 <= P_k <=1$
  e.g. Sun Rises in West.

- Higher probability is less than the Information Content of Message having Lower probability
- Mathematically $I(m_k) > I(m_j)$

- Also we can state for the Sum of two messages that the information content in the two combined messages is same as the sum of information content of each message Provided the occurrence is mutually independent.

e.g. There will be Sunny weather Today.

There will be Cloudy weather Tomorrow

Mathematically,

$$I\,(m_k \text{ and } m_j) = I(m_k\, m_j) = I(m_k)+I(m_j)$$

INFORMATION MEASURE

- So question is which function that we can use that measure the Information?

  Information = F(1/Probability)

- Requirement that function must satisfy

1. Its output must be non negative Quantity.
2. Minimum Value is 0.
3. It Should make Product into summation.

- Information $I(m_k) = \text{Log}_b (1/ P_k)$

  Here b may be 2, e or 10

  If b = 2 then unit is bits

  b = e then unit is nats

  b = 10 then unit is decit

INFORMATION

Conditions of Occurrence of Events

- If we consider an event, there are three conditions of occurrence.

- If the event has not occurred, there is a condition of uncertainty.

- If the event has just occurred, there is a condition of surprise.

- If the event has occurred, a time back, there is a condition of having some information.

- These three events occur at different times. The difference in these conditions help us gain knowledge on the probabilities of the occurrence of events.

Entropy

- When we observe the possibilities of the occurrence of an event, how surprising or uncertain it would be, it means that we are trying to have an idea on the average content of the information from the source of the event.

- Entropy can be defined as a measure of the average information content per source symbol.

Average Information Content

- It is necessary to define the information content of the particular symbol as communication channel deals with symbol.

Here we make following assumption,

- The Source is stationery, so Probability remains constant with time.

- The Successive symbols are statistically independent and come out at avg rate of r symbols per second

AVERAGE INFORMATION CONTENT

- Suppose a source emits M Possible symbols s1, s2, …..SM having
- Probability of occurrence, p1,p2,…….pm

$$\sum_{i=1}^{M} Pi = 1$$

- For a long message having symbols N (>>M)
- s1 will occur P1N times, like also
- s2 will occur P2N times and so on

Information Rate

- Information Rate = Total Information/ time taken
- Here Time Taken n bits are transmitted with r symbols per second.
- Total Information is nH.
- Information rate,

$$R = \frac{nH}{\left(\frac{n}{r}\right)}$$

$$R = rH$$ Bits/sec

ENTROPY

- Claude Shannon, the “father of the Information Theory”, provided a formula for it as − ipilogbpiH=−∑ipilogbpi.

- Where p_i is the probability of the occurrence of character number i from a given stream of characters and b is the base of the algorithm used. Hence, this is also called as Shannon’s Entropy.

- The amount of uncertainty remaining about the channel input after observing the channel output, is called as Conditional Entropy.

- It is denoted by H(x|y).

Mutual Information

- Let us consider a channel whose output is Y and input is X.

- Let the entropy for prior uncertainty be X = H(x).

(This is assumed before the input is applied).

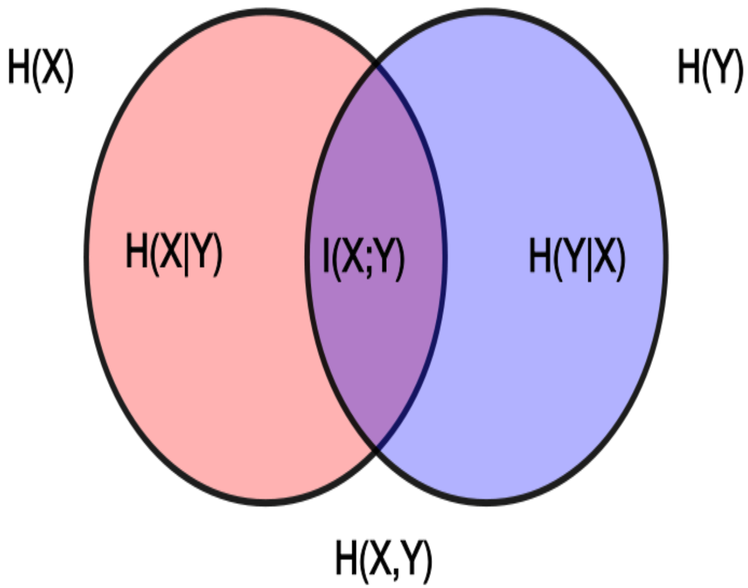
- To know about the uncertainty of the output, after the input is applied, let us consider Conditional Entropy, given that $Y = y_k$

Mutual Information

- Two discrete random variables: *X* and *Y*
- Measures the information knowing either variables provides about the other
- What if *X* and *Y* are fully independent or dependent?

$$
\begin{aligned}
I(X;Y) &= \sum \sum \mathrm{P}[X = x, Y = y] I(x, y) \\
&= \sum \sum \mathrm{P}[X = x, Y = y] \log \frac{\mathrm{P}[x \mid y]}{\mathrm{P}[x]} \\
&= \sum \sum \mathrm{P}[X = x, Y = y] \log \frac{\mathrm{P}[x, y]}{\mathrm{P}[x]\mathrm{P}[y]}
\end{aligned}
$$

Mutual Information



$$
\begin{aligned}
I(X;Y) &= H(X) - H(X \mid Y) \\
&= H(Y) - H(Y \mid X) \\
&= H(X) + H(Y) - H(X,Y)
\end{aligned}
$$

Some properties

$$I(X;Y) = I(Y;X)$$

$$I(X;Y) \geq 0$$

$$I(X;X) = H(X)$$

$$I(X;Y) \leq \min\{H(X), H(Y)\}$$

$$0 \leq H(X) \leq \log|\mathrm{X}|$$

$$\text{If } Y = g(X), \text{ then } H(Y) \leq H(X)$$

Entropy is maximum, when probabilities are equal

Joint and conditional entropy

- Joint entropy

$$H(X,Y) = \sum \mathrm{P}[X = x, Y = y] \log \mathrm{P}[X = x, Y = y]$$

- Conditional entropy of *Y* given *X*

$$H(Y \mid X) = \sum \mathrm{P}[X = x] H(Y \mid X = x)$$
$$= -\sum \mathrm{P}[X = x, Y = y] \log P[Y = y \mid X = x]$$

Joint and conditional entropy

- Chain rule for entropies

$$H(X_1, X_2, \ldots, X_n) = H(X_1) + H(X_2 \mid X_1) + H(X_3 \mid X_1, X_2) + \cdots + H(X_n \mid X_1, X_2, \ldots, X_{n-1})$$

- Therefore,

$$H(X_1, X_2, \ldots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

- If X_i are iid

$$H(X_1, X_2, \ldots, X_n) = nH(X)$$

CHANNEL CAPACITY

- We have so far discussed mutual information. The maximum average mutual information,

- An instant of a signaling interval, when transmitted by a discrete memory less channel, the probabilities of the rate of maximum reliable transmission of data, can be understood as the channel capacity.

- It is denoted by C and is measured in bits per channel use.

INFORMATION CHANNEL CAPACITY

- We define the *"information" channel capacity* of a discrete memoryless channel as, where the maximum is taken over all possible input distributions *p(x)*.

- This means: the capacity is the maximum entropy of Y, reduced by the contribution of information given by Y.

- An operational definition of channel capacity as the highest rate in bits per channel use at which information can be sent with arbitrarily low probability of error.

- Shannon's second theorem establishes that the information channel capacity is equal to the operational channel capacity.

DATA COMPRESSION AND TRANSMISSION

- There is a duality between the problems of data compression and data transmission.

- During compression, we remove all the redundancy in the data to form the most compressed version possible.

- During data transmission, we add redundancy in a controlled fashion to combat errors in the channel.

Examples of Channel Capacity

- Noiseless binary channel. Suppose that we have a channel whose the binary input is reproduced exactly at the output.

- In this case, any transmitted bit is received without error. Hence, oneerror-free bit can be transmitted per use of the channel, and the capacity is 1 bit.

- We can also calculate the information capacity $C = \max I(X; Y) = 1$ bit, which is achieved by using $p(x) = (1/2, 1/2)$.

EXAMPLES OF CHANNEL CAPACITY

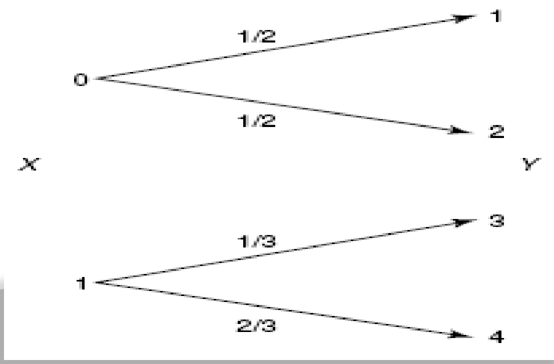
- Noisy Channel with Nonoverlapping Outputs. This channel has two possible outputs corresponding to each of the two inputs.

- Even though the output of the channel is a random consequence of the input, the input can be determined from the output, and hence every transmitted bit can be recovered without error.

- The capacity of this channel is also 1 bit per transmission.
- We can also calculate the information capacity

$C = \max I(X; Y) = 1$ bit, which is achieved by using

$p(x) = (1/2, 1/2)$.



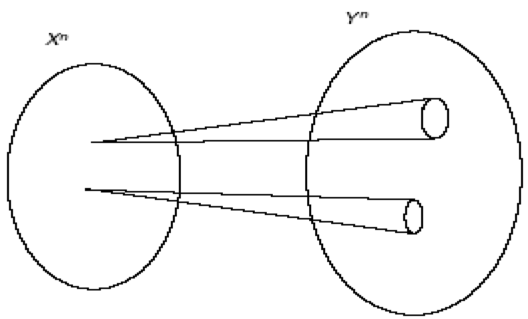
PROPERTIES OF CHANNEL CAPACITY

- 1. $C \geq 0$ since $I(X; Y) \geq 0$.
- 2. $C \leq \log |X|$ since $C = \max I(X; Y) \leq \max H(X) = \log |X|$.
- 3. $C \leq \log |Y|$ for the same reason.
- 4. $I(X; Y)$ is a continuous function of $p(x)$.
- 5. $I(X; Y)$ is a concave function of $p(x)$

WHY C IS THE CAPACITY?

- If we consider an input sequence of n symbols that we want to transmit over the channel, there are approximately $2^{nH(Y|X)}$ possible Y sequences for each typical n-sequence, all of them equally likely.

- We wish to ensure that no two *X* sequences produce the same *Y* output sequence. Otherwise, we will not be able to decide which *X* sequence was sent.



WHY C IS THE CAPACITY?

- The total number of possible (typical) Y sequences is $\approx 2^{nH(Y)}$.
- This set has to be divided into sets of size $2^{nH(Y|X)}$ corresponding to the different input X sequences that are the producers.
- The total number of disjoint sets is less than or equal to

$$2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$$

- Hence, we can send at most $\approx 2^{nI(X;Y)}$ distinguishable sequences of length n.
- In fact, the maximum number of disjoint sets is the maximum number of "independent" output sets.

SHANNON-HARTLEY THEOREM

- Channel capacity, the tightest upper bound on information rate (excluding error correcting codes) of arbitrarily low bit error rate data that can be sent with a given average signal power S through an additive white Gaussian noise channel of power N, is:

$$C = B \log_2 \left(1 + \frac{S}{N}\right)$$

- C is the channel capacity in bits per second
- B is the bandwidth of the channel in hertz
- S is the total received signal power over bandwidth, in watts

SHANNON'S THEOREM: EXAMPLE

- For SNR of 0, 10, 20, 30 dB, one can achieve C/B of 1, 3.46, 6.66, 9.97 bps/Hz, respectively.

Example:

- Consider the operation of a modem on an ordinary telephone line. The SNR is usually about 1000. The bandwidth is 3.4 KHz.

Therefore:

$$C = 3400 \text{ X } \log_2(1 + 1000)$$

$$= (3400)(9.97)$$

$$\approx 34 \text{ kbps}$$

BANDWIDTH AND SNR TRADEOFF

The channel capacity is

$$C = \omega \, log\left(1 + \frac{S}{N}\right) = \omega \, log\left(1 + \frac{S}{\eta\omega}\right)$$

Now if $S/N = \infty$ the capacity C should be ∞. But it is not. Again, on the other hand C is not ∞ even if $\omega = \infty$. Thus for a fixed signal power in presence of Gaussian noise the channel capacity approaches upper limit called as Shannons limit. This can be analysed as fallows:

$$C = \omega \, log\left(1 + \frac{S}{\eta\omega}\right)$$

$$C = \frac{S}{\eta}\frac{\eta\omega}{S} \, log\left(1 + \frac{S}{\eta\omega}\right)$$

$$C = \frac{S}{\eta} \, log\left(1 + \frac{S}{\eta\omega}\right)^{\eta\omega/S}$$

BANDWIDTH AND SNR TRADEOFF

$$\text{Let } \frac{S}{\eta\omega} = x \text{ so}$$

$$\lim_{x\to 0}(1+x)^{x} = e$$

If $x = \dfrac{S}{\eta\omega}$, then if ω approaches infinity, x approaches zero. Hence,

$$\lim_{\omega\to 0} C = \frac{S}{\eta} log\, e = 1.44\frac{S}{\eta} = R_{max}$$

HUFFMAN CODING

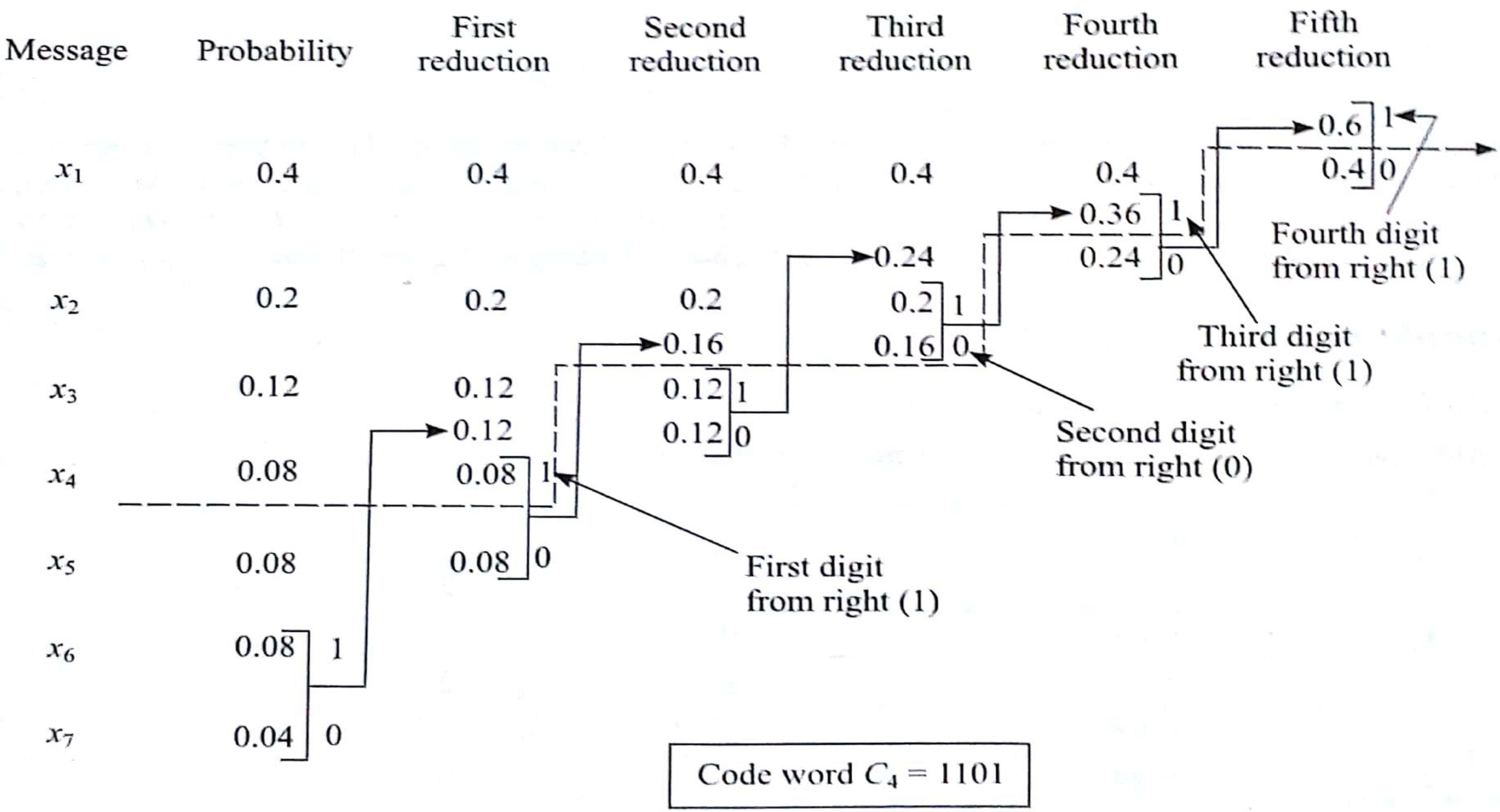
- In the field of data compression, Shannon–Fano coding, named after Claude Shannon and Robert Fano.

- It is a technique for constructing a prefix code based on a set of symbols and their probabilities (estimated or measured).

- It is suboptimal in the sense that it does not achieve the lowest possible expected code word length like Huffman coding.

HUFFMAN CODING

Example 11.3.1 Solve Ex. 11.2.2 by the Huffman method.



Shannon-Fano Algorithm

- Shannon-Fano Algorithm: A Shannon–Fano tree is built according to a
- specification designed to define an effective code table.

The actual algorithm is simple:

1. For a given list of symbols, develop a corresponding list of probabilities or frequency counts so that each symbol's relative frequency of occurrence is known.

2. Sort the lists of symbols according to frequency, with the most frequently occurring symbols at the left and the least common at the right.

Shannon-Fano Algorithm

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

4. The left part of the list is assigned the binary digit 0, and the right part is assigned the digit 1. This means that the codes for the symbols in the first part will all start with 0, and the codes in the second part will all start with 1.

5. Recursively apply the steps 3 and 4 to each of the two halves, subdividing groups and adding bits to the codes until each symbol has become a corresponding code leaf on the tree.

Example

Example : The source of information A generates the symbols {A0, A1, A2, A3 and A4} with the corresponding probabilities {0.4, 0.3, 0.15, 0.1 and 0.05}. Encoding the source symbols using binary encoder and Shannon-Fano

Example

Source Symbol	P_i	Binary Code	Shannon-Fano
A0	0.4	000	0
A1	0.3	001	10
A2	0.15	010	110
A3	0.1	011	1110
A4	0.05	100	1111
L_{avg}	H = 2.0087	3	2.05

The Entropy of the source is

$$H = -\sum_{i=0}^{4} Pi \log_2 Pi = 2.0087 \text{ bit/symbol}$$

Since we have 5 symbols ($5<8=2^3$), we need 3 bits at least to represent each symbol in binary (fixed-length code). Hence the average length of the binary code is

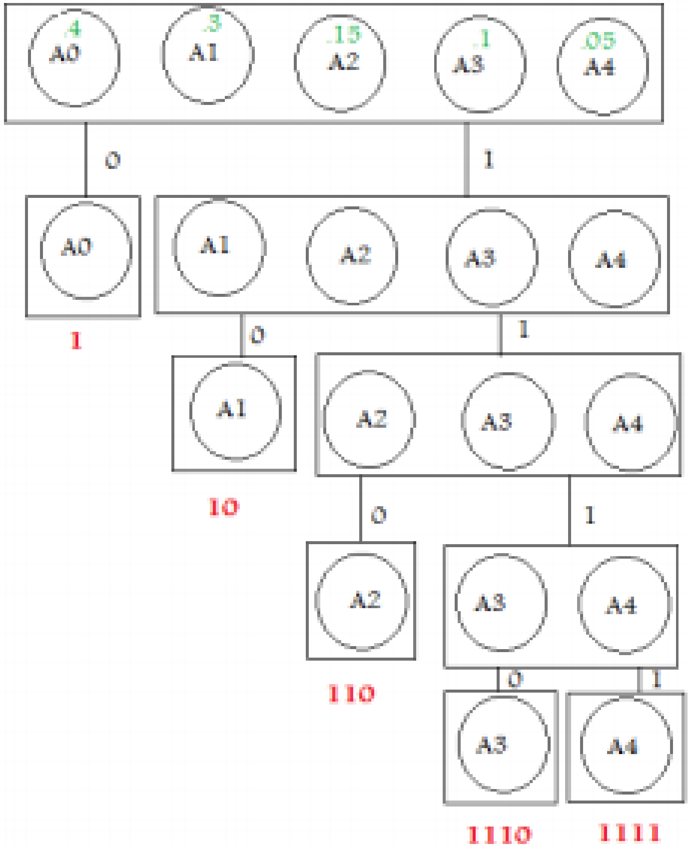
$$Lavg = \sum_{i=0}^{4} Pi\ li = 3\ (0.4 + 0.3 + 0.15 + 0.1 + 0.05) = 3 \text{ bit/symbol}$$

Thus the efficiency of the binary code is

$$\eta = \frac{H}{Lavg} = \frac{2.0087}{3} = 67\%$$

Cont.

Shannon-Fano code is a top-down approach. Constructing the code tree, we get



Cont.

The average length of the Shannon-Fano code is

$$Lavg = \sum_{i=0}^{4} Pi\,li = 0.4 * 1 + 0.3 * 2 + 0.15 * 3 + 0.1 * 4 + 0.05 * 4 = 2.05 \text{ bit/symbol}$$

Thus the efficiency of the Shannon-Fano code is

$$\eta = \frac{H}{Lavg} = \frac{2.0087}{2.05} = 98\%$$

This example demonstrates that the efficiency of the Shannon-Fano encoder is much higher than that of the binary encoder.

INTRODUCTION - SPREAD SPECTRUM

- Important form of encoding for wireless communication.
- In conventional wireless communication we use a single frequency for transmission.

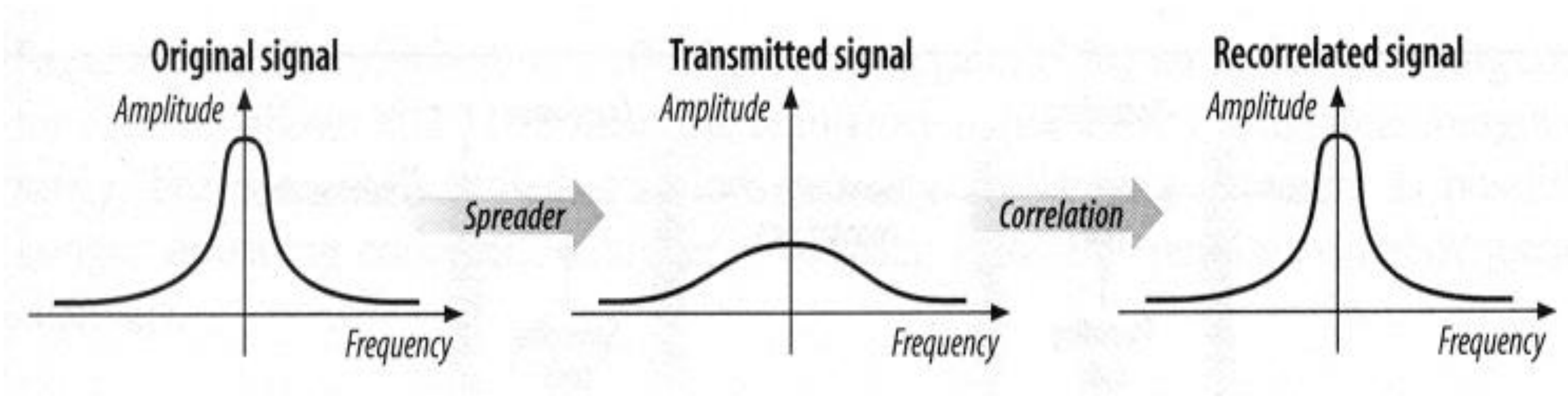
  For example 93.5FM, Every one can tune this frequency and they will get all signals from FM station.

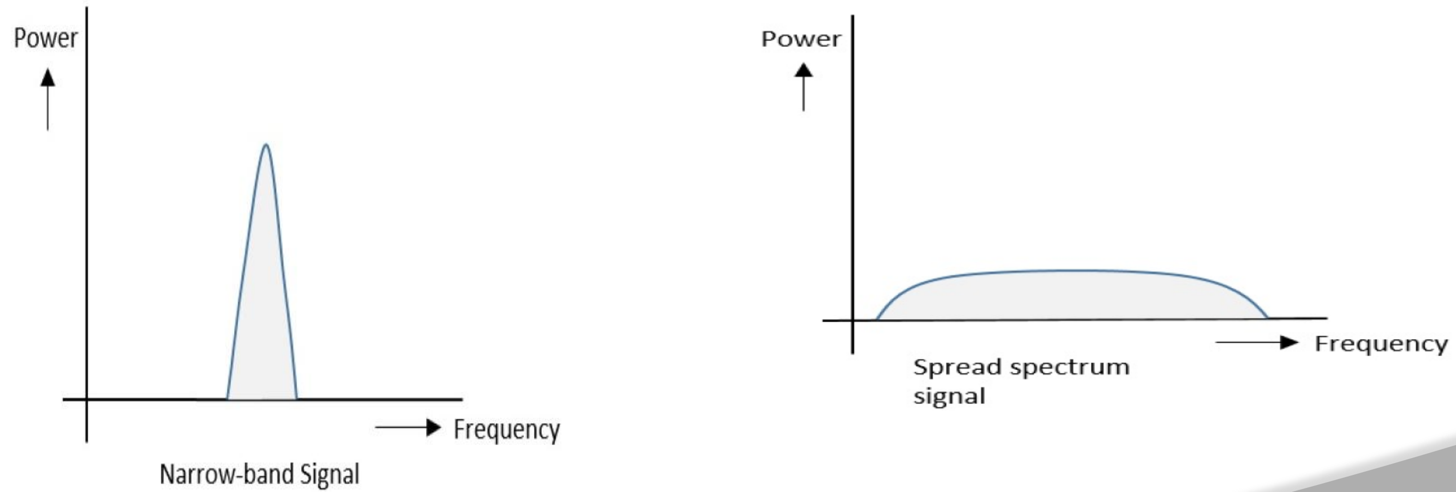
INTRODUCTION - SPREAD SPECTRUM

- A collective class of signaling techniques are employed before transmitting a signal to provide a secure communication, known as the Spread Spectrum Modulation.

- Spread spectrum is a communication technique that spreads a narrowband communication signal over a wide range of frequencies for transmission then de-spreads it into the original data bandwidth at the receive.

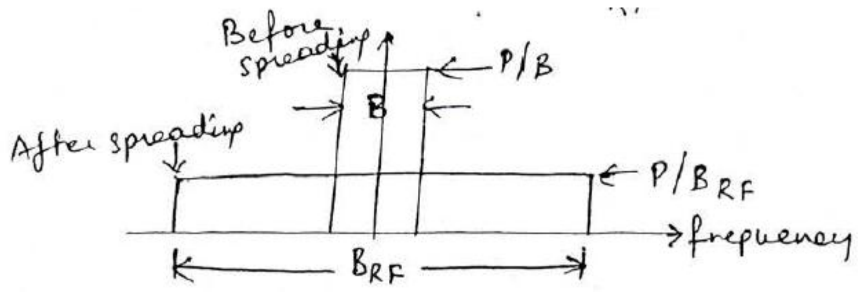
SPREAD SPECTRUM





Principle of Spread Spectrum

- The basic idea behind this spread spectrum is to take the energy in Bandwidth ‘B’ and Spread it over the wide bandwidth BRF.



- There are three techniques to accomplishing the Spread spectrum. They are

1. Direct Sequence Spread Spectrum(DSSS)
2. Frequency Hopping Spread Spectrum(FHSS)
3. Time Hopping Spread Spectrum(THSS)

SPREAD SPECTRUM

- Spread Spectrum is a technique in which a transmitted signal occupies a bandwidth which is kept much larger than that which is required by the baseband information signal.

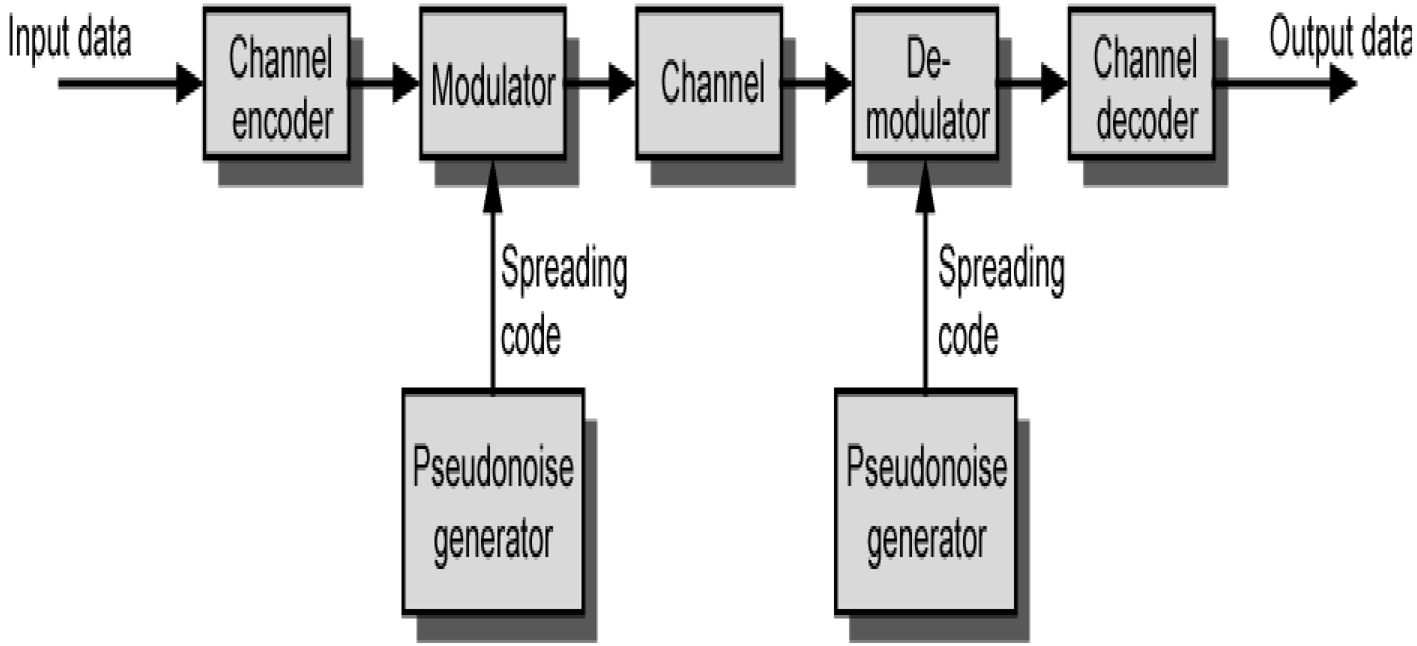
- A communication system is considered a spread spectrum system if it satisfies the following two criteria:

SPREAD SPECTRUM

1.Bandwidth of the spread spectrum signal has to be greater than the information bandwidth

2.The spectrum spreading is accomplished before transmission through the use of a code that is independent of data sequence. The same code is used in receiver to dispread the received signal.

GENERAL MODEL OF SPREAD SPECTRUM SYSTEM



DESCRIPTION OF SS

- Input fed into channel encoder
  - Produces narrow bandwidth analog signal around central frequency
- Signal modulated using sequence of digits
  - Spreading code/sequence
  - Typically generated by pseudonoise/pseudorandom number generator
- Increases bandwidth significantly
  - Spreads the spectrum
- Receiver uses same sequence to demodulate signal
- Demodulated signal fed into channel decoder

ADVANTAGE OF SPREAD SPECTRUM

- The main advantage of spread spectrum communication technique is to prevent “interference” whether it is intentional or unintentional.

- “Spread” radio signal over a wide frequency range
- Several magnitudes higher than minimum requirement
- Gained popularity by the needs of military communication
- Proved resistant against hostile jammers
- Ratio of information bandwidth and spreading bandwidth is identified as spreading gain or processing gain

ADVANTAGES OF SPREAD SPECTRUM

- Crosstalk elimination
- Better output with data integrity
- Reduced effect of multipath fading
- Better security
- Reduction in noise
- Coexistence with other systems
- Longer operative distances
- Hard to detect
- Not easy to demodulate/decode
- Difficult to jam the signals

APPLICATIONS OF SPREAD SPECTRUM

- Interference

  Prevents interference at specific frequencies

  E.g. other radio users, electrical systems

- Military

  Prevents signal jamming

Scrambling of ‘secret’ messages

- Wireless LAN security

Prevents ‘eavesdropping’ of wireless links

– Prevents ‘hacking’ into wireless LANs

- CDMA (Code Division Multiple Access)

– Multiple separate channels in same medium using different spreading codes

DIRECT SEQUENCE SPREAD SPECTRUM(DSSS)

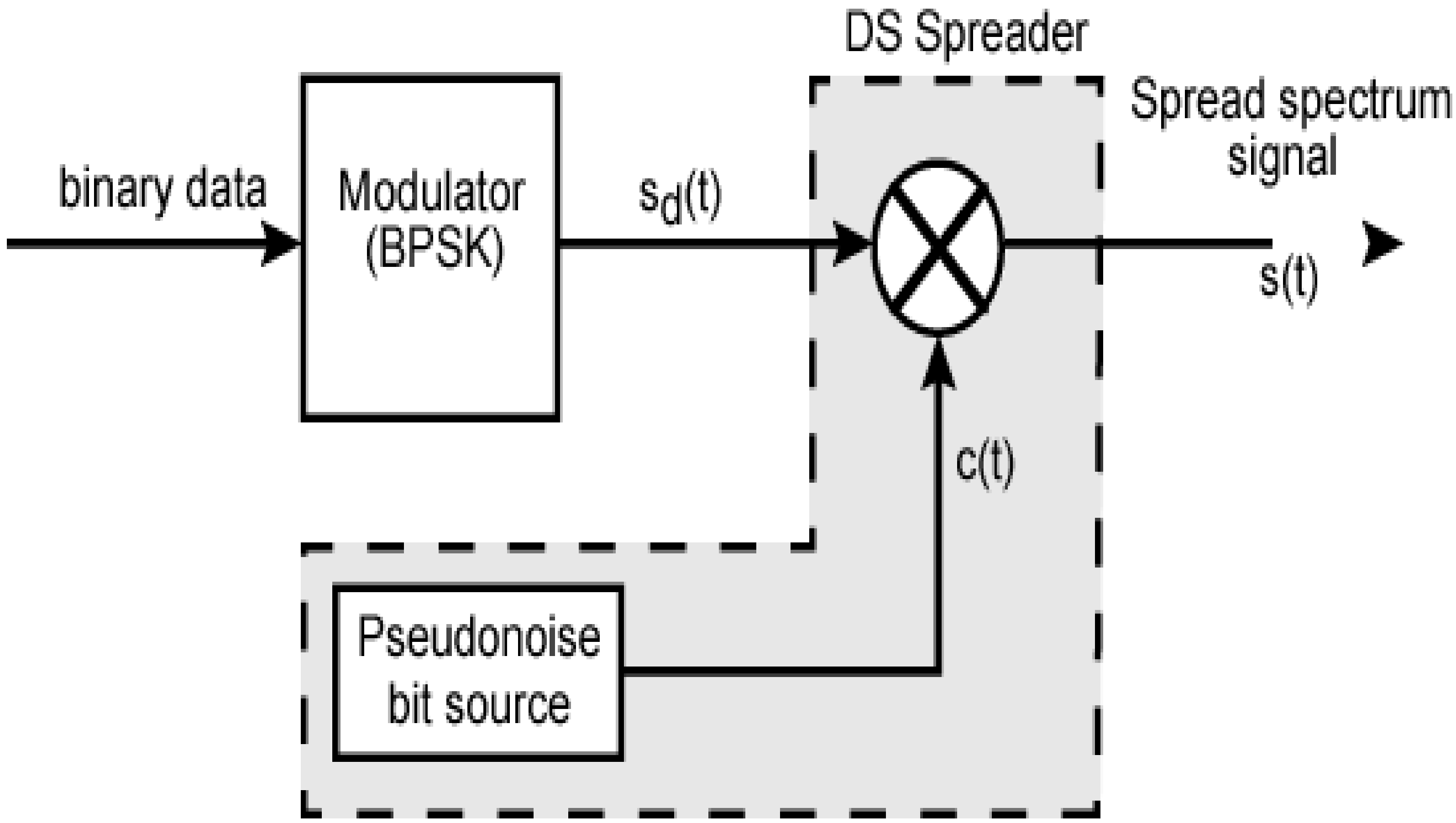
- Whenever a user wants to send data using this DSSS technique, each and every bit of the user data is multiplied by a secret code, called as chipping code.

- This chipping code is nothing but the spreading code which is multiplied with the original message and transmitted.

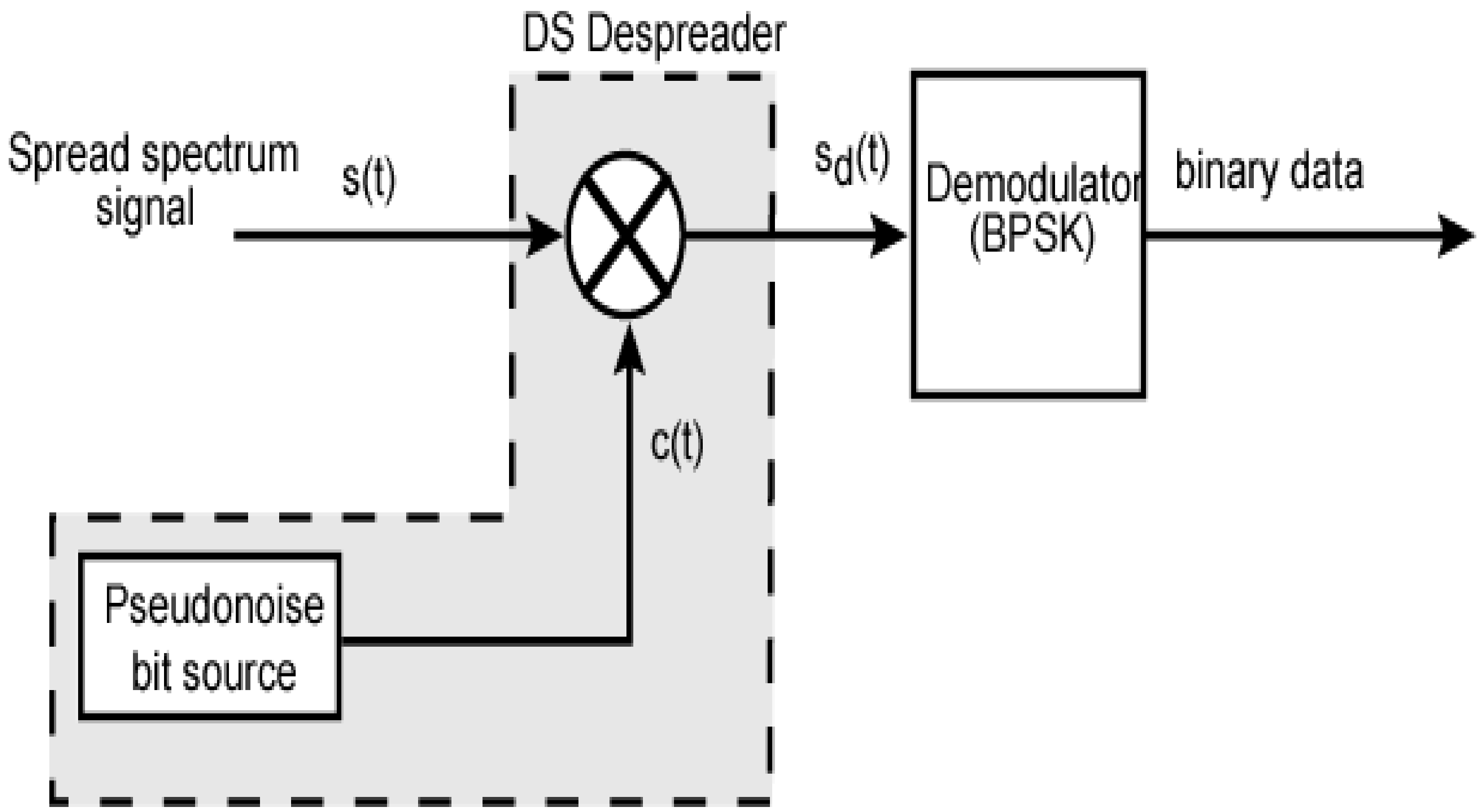
- The receiver uses the same code to retrieve the original message.

- Each input bit represented by multiple bits using spreading code

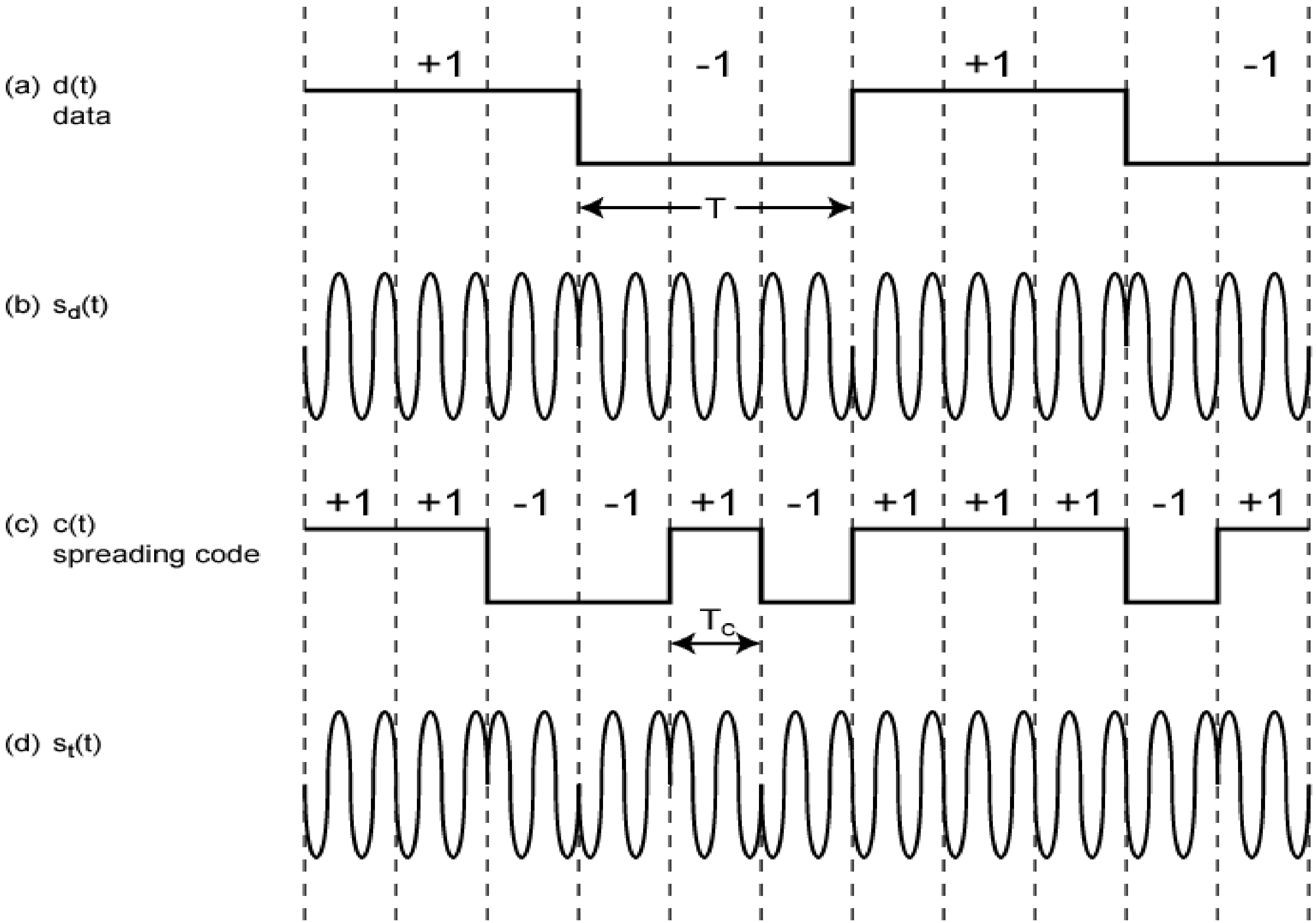
DIRECT SEQUENCE SPREAD SPECTRUM: TRANSMITTER



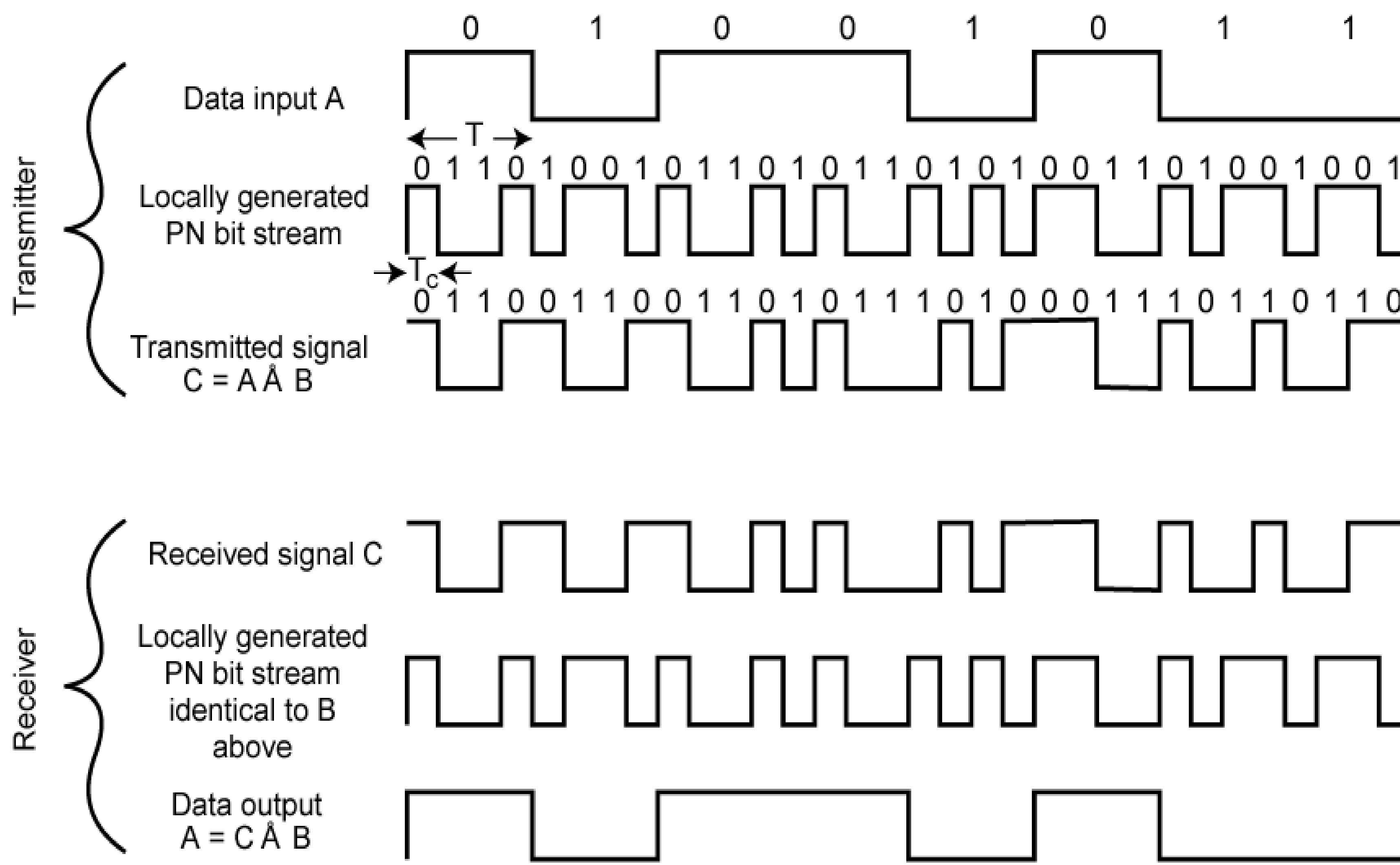
DIRECT SEQUENCE SPREAD SPECTRUM RECEIVER



DIRECT SEQUENCE SPREAD SPECTRUM USING BPSK EXAMPLE:



DIRECT SEQUENCE SPREAD SPECTRUM EXAMPLE

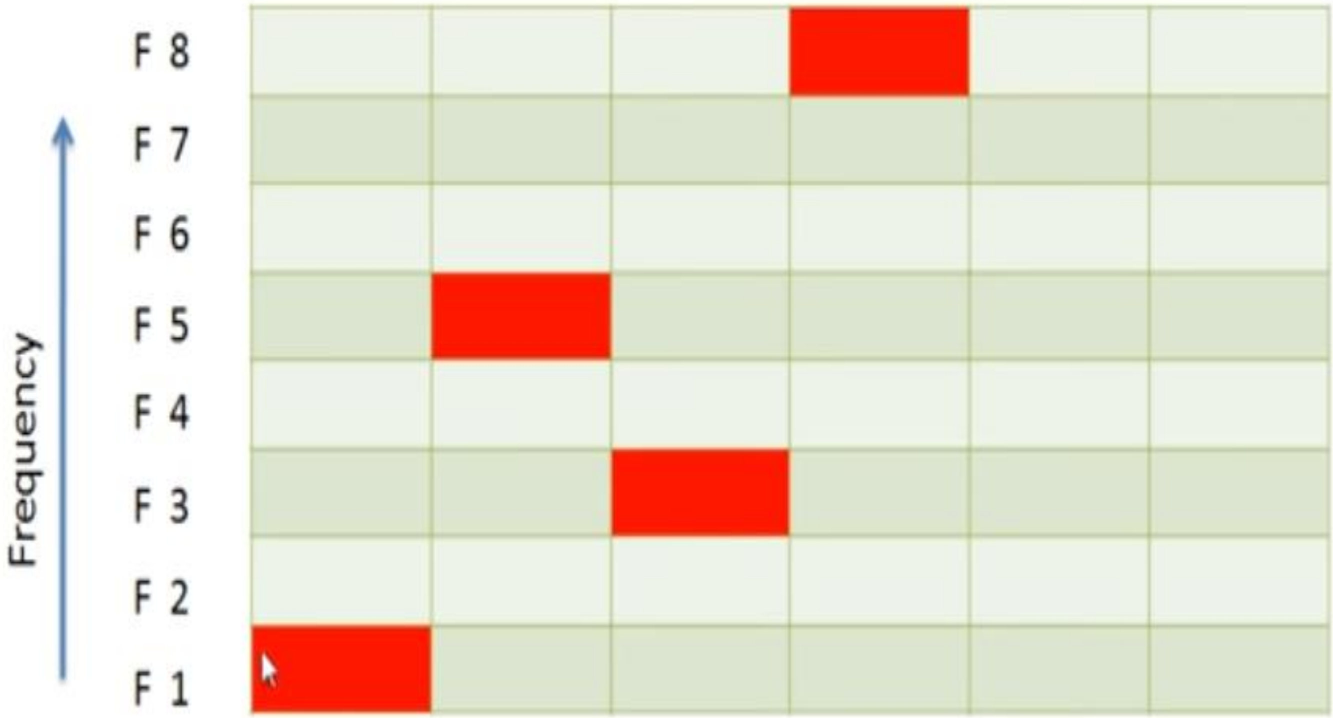


FREQUENCY HOPPING SPREAD SPECTRUM (FHSS)

- The information signal is transmitted on different frequencies.

- The type of spread spectrum in which the carrier hops randomly from one frequency to another is called frequency-hopping (FH) spread spectrum.

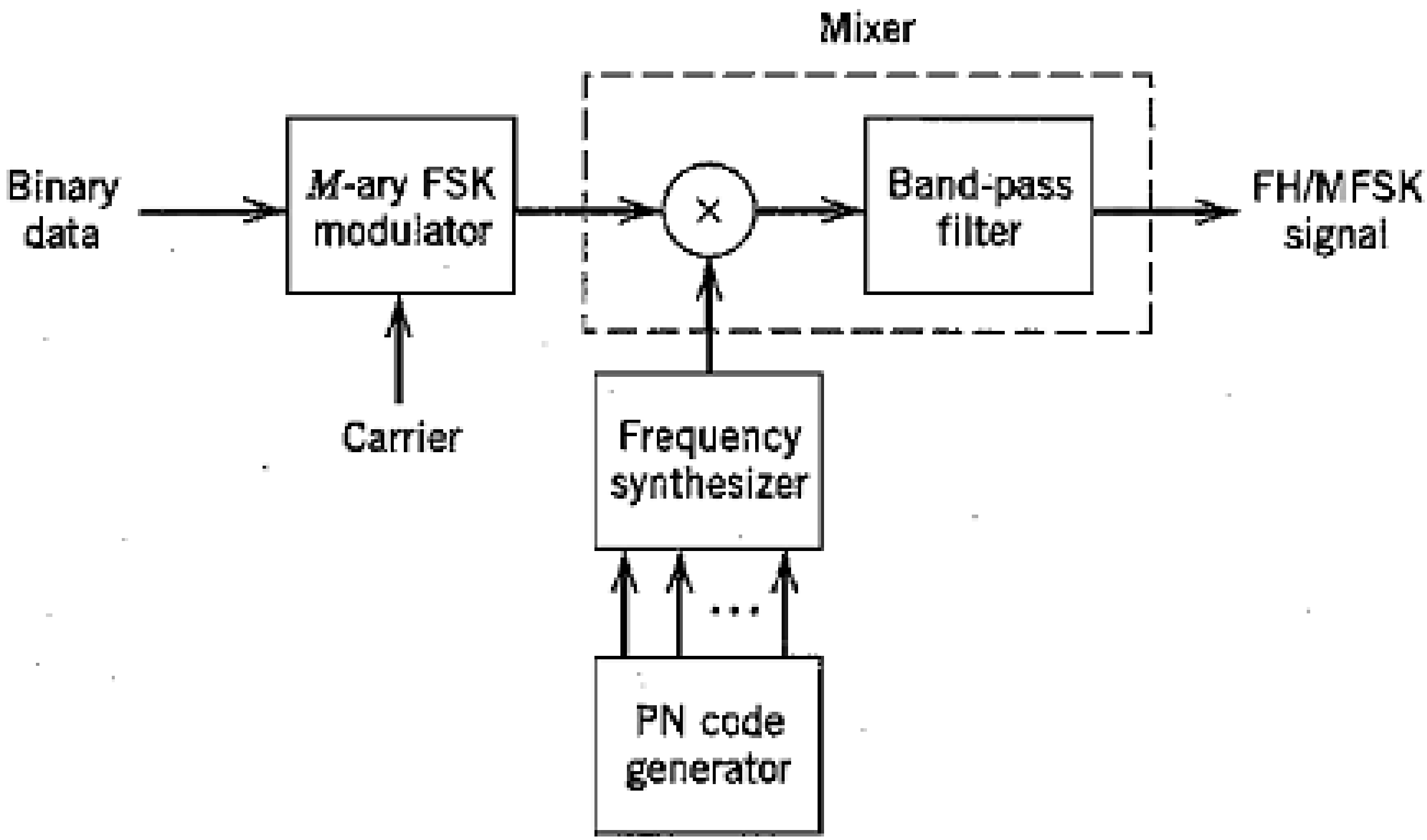
Example: Sender A want to send some data. Hopping sequence for A is F1F5F3F8



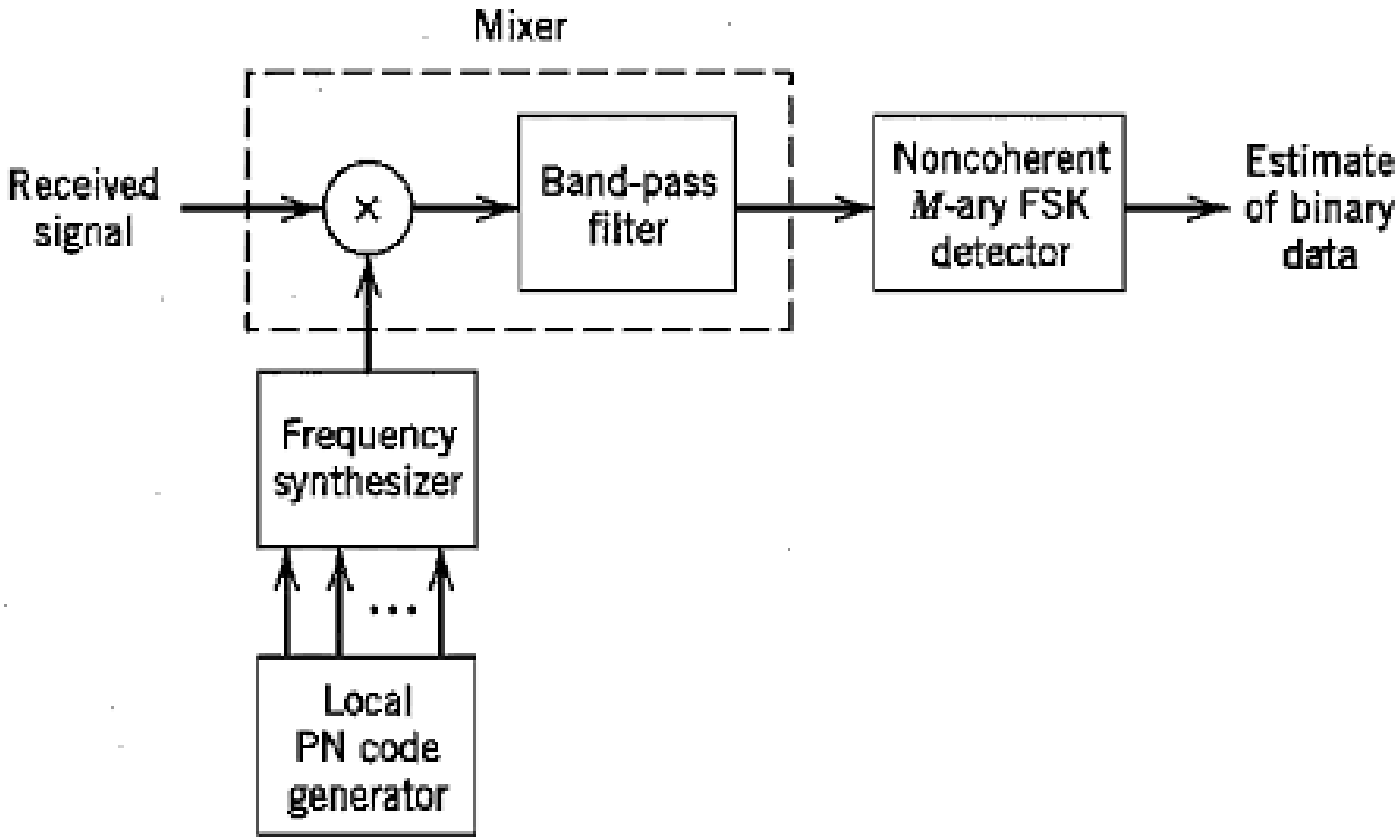
TYPES OF FHSS

- Slow-frequency hopping, in which the symbol rate R_s, of the MFSK signal is an integer multiple of the hop rate R_h.
- That is, several symbols are transmitted on each frequency hop.
- Fast-frequency hopping, in which the hop rate R_h is an integer multiple of the MFSK symbol rate R_s.
- That is, the carrier frequency will change or hop several times during the transmission of one symbol

FREQUENCY HOPPING SPREAD SPECTRUM (FHSS)-TRANSMITTER



FREQUENCY HOPPING SPREAD SPECTRUM (FHSS)-RECEIVER



PSEUDO-NOISE SEQUENCE
PN SEQUENCE

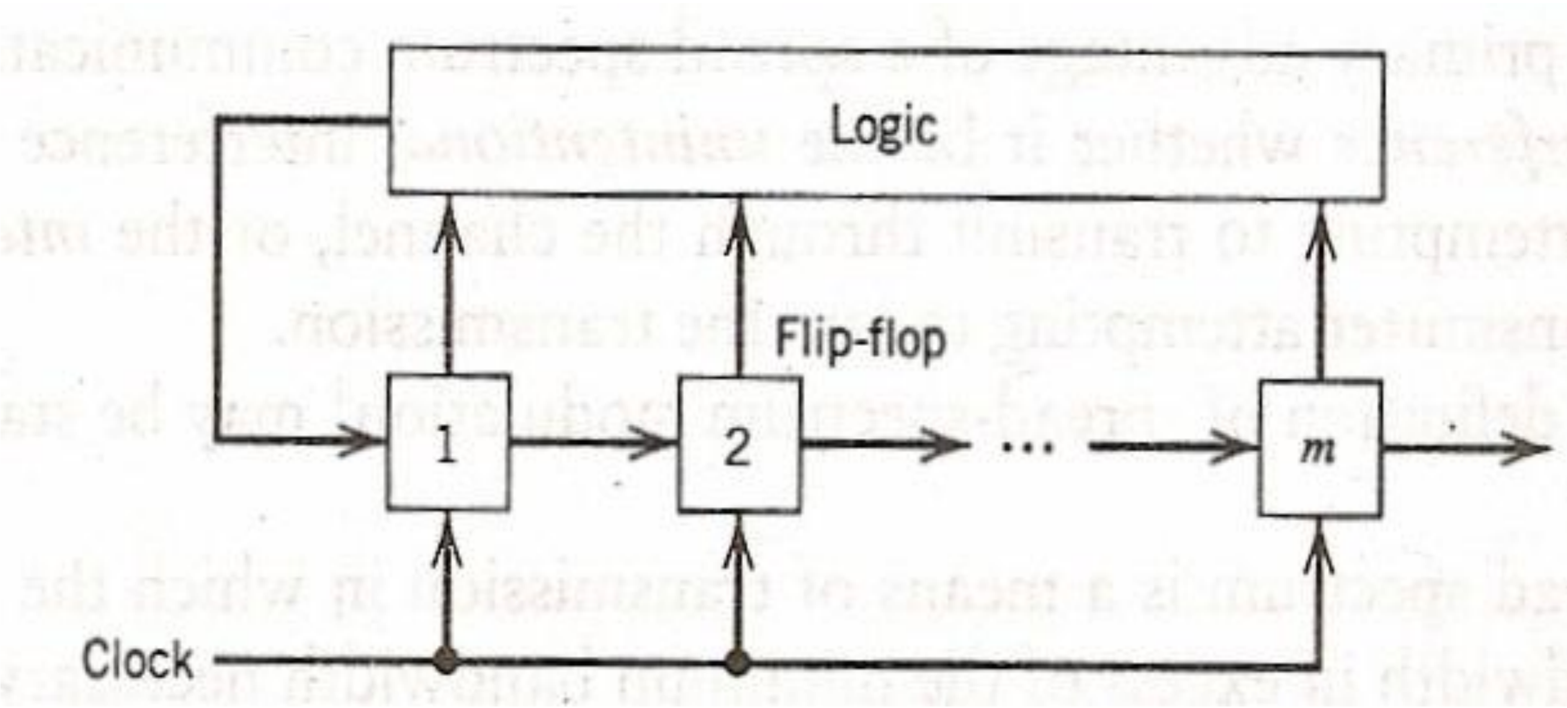
- The pseudo-noise (PN) sequence is a periodic binary sequence with a noise like waveform that is generated by means of a feedback shift register.

- How to generate? Using feedback shift register.

- The feedback shift register consists of m-stage shift registers and a logic circuit

- Flip flops are regulated by clock, At each clock pulse state of the flip-flop change

PN SEQUENCE GENERATOR



EXAMPLE: PN SEQUENCE GENERATION

- Three flip-flops form 3 bit shift register

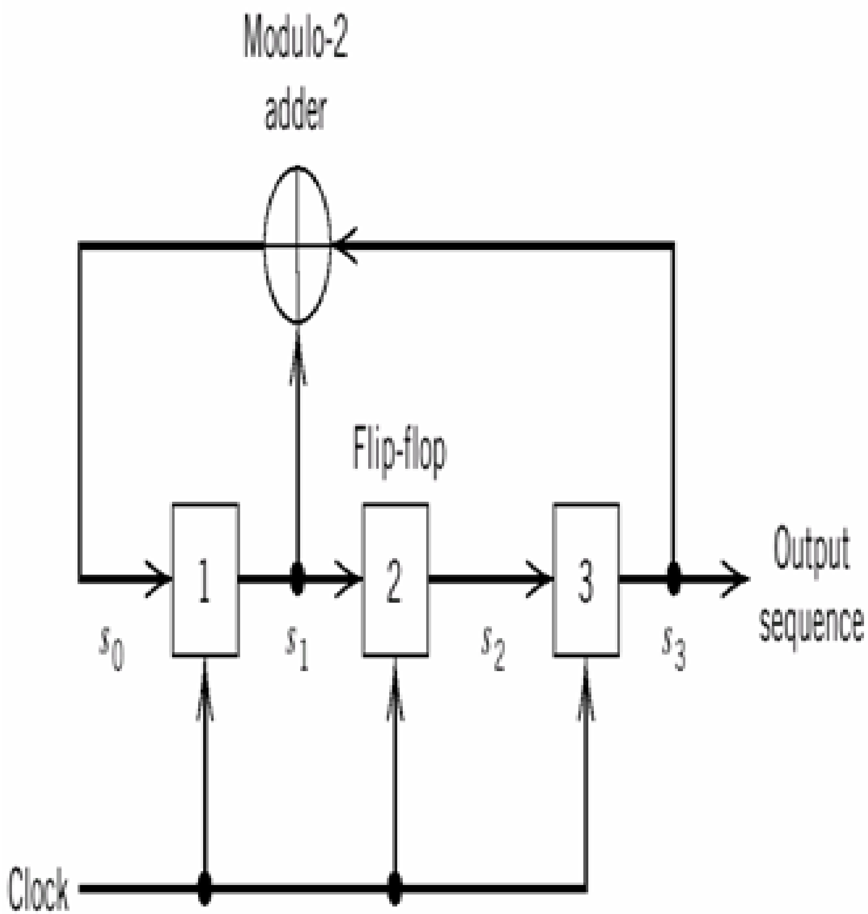
- Logic circuit is modulo-2 adder

- Modulo-2 adder is nothing but 2 input X-OR gate

(inputs to this gate are s1 and s3,output of this gate is s0)

- Initial state is s1=1, s2=0 and s3=0

EXAMPLE: PN SEQUENCE GENERATION



S1	S2	S3
1	0	0
1	1	0
1	1	1
0	1	1
1	0	1
0	1	0
0	0	1
1	0	0

LINEAR BLOCK CODES AND CONVOLUTIONAL CODES

Introduction

Error Control Coding (ECC)

- Extra bits are added to the data at the transmitter (redundancy) to permit error detection or correction at the receiver.
- Done to prevent the output of erroneous bits despite noise and other imperfections in the channel.
- The positions of the error control coding and decoding are shown in the transmission model.
- Noise or Error is the main problem in the signal, which disturbs the reliability of the communication system.
- **Error control coding** is the coding procedure done to control the occurrences of errors. These techniques help in Error Detection and Error Correction.

Error Control Coding (ECC) Cont...

•There are many different error correcting codes depending upon the mathematical principles applied to them. But, historically, these codes have been classified into **Linear block codes** and **Convolution codes**.

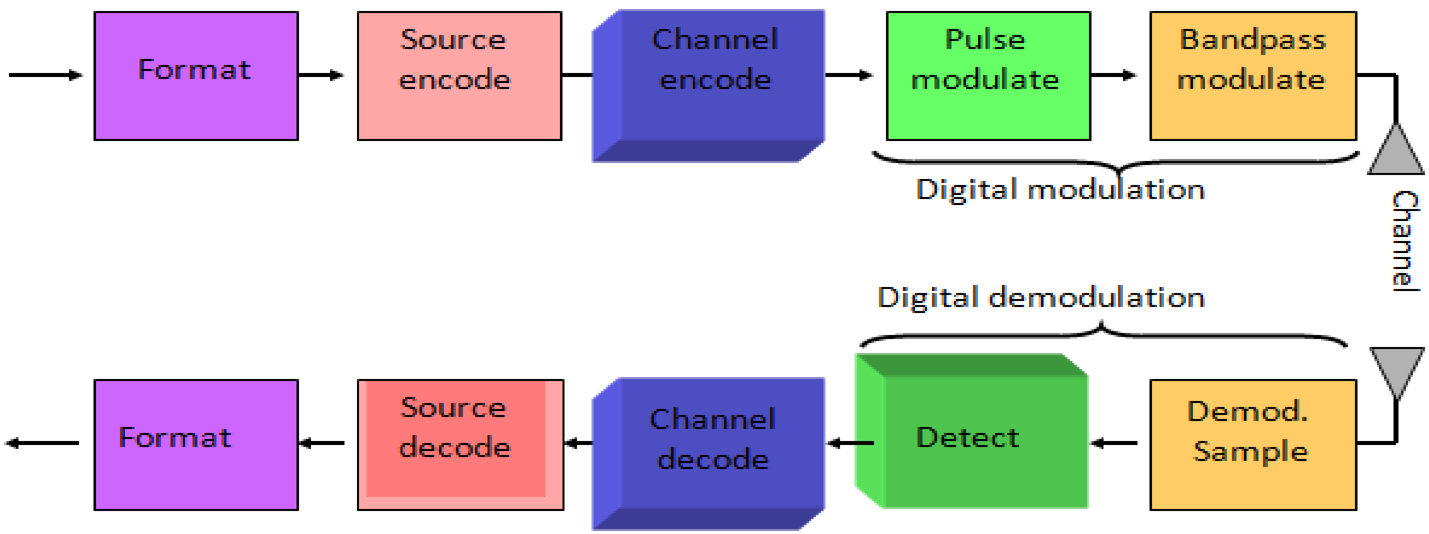


Fig: Digital Communication System

What is channel coding?

Channel coding:

- Transforming signals to improve communications performance by increasing the robustness against channel impairments (noise, interference, fading, ...).
- Waveform coding: Transforming waveforms to better waveforms .
- Structured sequences: Transforming data sequences into better sequences, having structured redundancy.

➢ **channel coding schemes** are listed below:

--Automatic repeat request (ARQ)
--Block coding
--Convolution coding
--Concatenated coding
--Orthogonal coding

Error control techniques

- **Automatic Repeat reQuest (ARQ)**

  Full-duplex connection, error detection codes

  The receiver sends feedback to the transmitter, saying that if any error is detected in the received packet or not (Not-Acknowledgement (NACK) and Acknowledgement (ACK), respectively).

  The transmitter retransmits the previously sent packet if it receives NACK.

- **Forward Error Correction (FEC)**

  Simplex connection, error correction codes

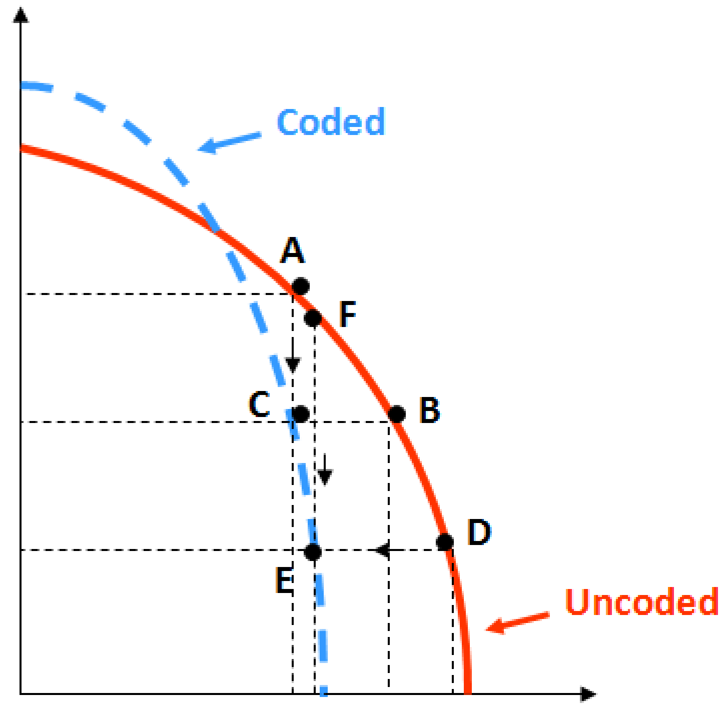
  The receiver tries to correct some errors

- **Hybrid ARQ (ARQ+FEC)**

  Full-duplex, error detection and correction codes

Why using error correction coding?

- Error performance vs. bandwidth
- Power vs. bandwidth
- Data rate vs. bandwidth
- Capacity vs. bandwidth



Coding gain:

For a given bit-error probability, the reduction in the Eb/N0 that can be realized through the use of code:

Channel models

- Discrete memory-less channels
  - Discrete input, discrete output
- Binary Symmetric channels
  - Binary input, binary output
- Gaussian channels
  - Discrete input, continuous output

Linear block codes

•In the linear block codes, the parity bits and message bits have a linear combination, which means that the resultant code word is the linear combination of any two code words.

•Let us consider some blocks of data, which contains **k** bits in each block. These bits are mapped with the blocks which has **n** bits in each block. Here **n** is greater than **k**.

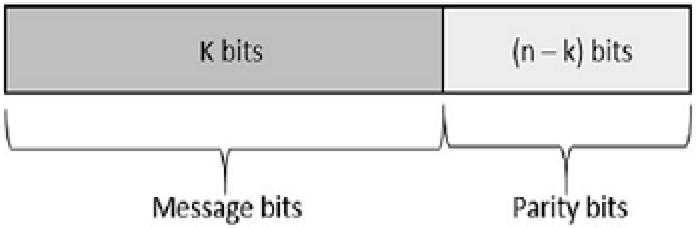
•The transmitter adds redundant bits which are **(n-k)** bits. The ratio **k/n** is the **code rate**. It is denoted by **r** and the value of **r** is **r < 1**.

•The **(n-k)** bits added here, are **parity bits**. Parity bits help in error detection and error correction, and also in locating the data. In the data being transmitted, the left most bits of the code word correspond to the message bits, and the right most bits of the code word correspond to the parity bits.

Systematic Code

- Any linear block code can be a systematic code, until it is altered. Hence, an unaltered block code is called as a **systematic code**.
- Following is the representation of the **structure of code word**, according to their allocation.

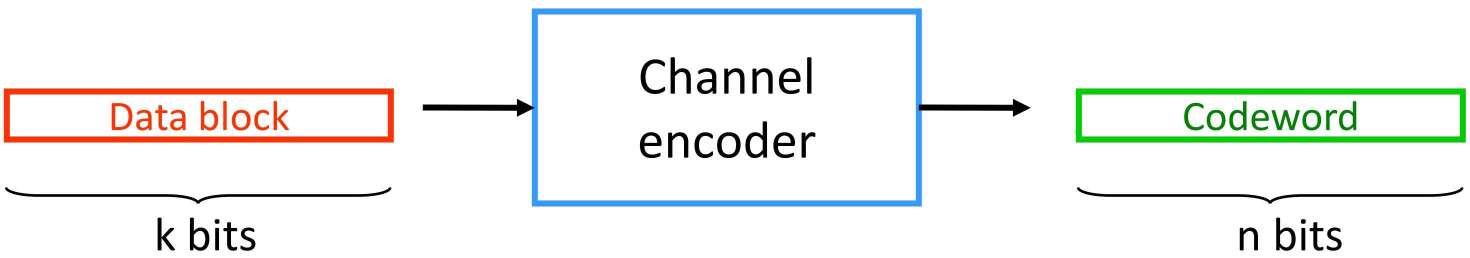
Structure of code word



- If the message is not altered, then it is called as systematic code. It means, the encryption of the data should not change the data.

Linear block codes cont...

- The information bit stream is chopped into blocks of k bits.
- Each block is encoded to a larger block of n bits.
- The coded bits are modulated and sent over the channel.
- The reverse procedure is done at the receiver.



$n\text{-}k$ Redundant bits

$$R_c = \frac{k}{n} \quad \text{Code rate}$$

Linear block codes cont...

- The Hamming weight of the vector **U**, denoted by w(**U**), is the number of non-zero elements in **U**.
- The Hamming distance between two vectors **U** and **V**, is the number of elements in which they differ.

$$d(\mathbf{U},\mathbf{V}) = w(\mathbf{U} \oplus \mathbf{V})$$

- The minimum distance of a block code is

$$d_{\min} = \min_{i \neq j} d(\mathbf{U}_i, \mathbf{U}_j) = \min_i w(\mathbf{U}_i)$$

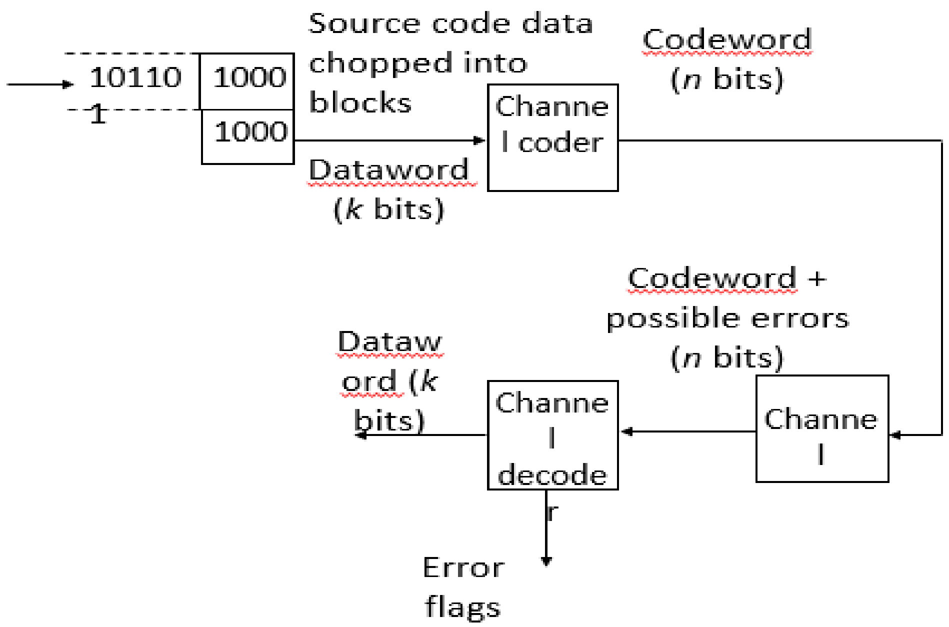
- Error detection capability is given by

$$e = d_{\min} - 1$$

- Error correcting-capability **t** of a code is defined as the maximum number of guaranteed correctable errors per code word, that is

$$t = \left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor$$

Error Control Process



Matrix description of Linear Block Codes

- In Linear Block Code,

➢ The block length C of the Linear Block Code is

$$C = m\, G$$

- where m is the information code word block length, G is the generator matrix.

$$G = [\mathrm{I}_k \,/\, P]_{k \times n}$$

- where P^i = Remainder of $[x^{n-k+i-1}/g(x)]$ for i=1, 2, .., k, and I is unit or identity matrix.

➢ At the receiving end, parity check matrix can be given as:

$H = [P^{\mathrm{T}} \mid \mathrm{I}_{n-k}]$, where P^{T} is the transpose of the matrix P.

Matrix description of Linear Block Codes cont…

- Encoding in (n,k) block code,

$$\mathbf{U} = \mathbf{mG}$$

$$(u_1, u_2, \ldots, u_n) = (m_1, m_2, \ldots, m_k) \cdot \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \vdots \\ \mathbf{V}_k \end{bmatrix}$$

$$(u_1, u_2, \ldots, u_n) = m_1 \cdot \mathbf{V}_1 + m_2 \cdot \mathbf{V}_2 + \ldots + m_2 \cdot \mathbf{V}_k$$

– The rows of G are linearly independent.

Matrix description of Linear Block Codes cont...

- Example: Block code (6,3)

$$\mathbf{G} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \mathbf{V}_3 \end{bmatrix} = \begin{bmatrix} 1\,1\,0\,1\,0\,0 \\ 0\,1\,1\,0\,1\,0 \\ 1\,0\,1\,0\,0\,1 \end{bmatrix}$$

Message vector	Codeword
000	000000
100	110100
010	011010
110	101110
001	101001
101	011101
011	110011
111	000111

LBC Example

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \mathbf{g}_2 \\ \mathbf{g}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

u= [0 1 1 0]

Linear Block Encoder (**v**=**u.G**)

$\mathbf{v} = \mathbf{g}_1 + \mathbf{g}_2$

v= [1 0 0 0 1 1 0]

Message	Codeword	
0000	0000000	
0001	**1010001**	$\mathbf{g}_3$
0010	**1110010**	$\mathbf{g}_2$
0011	0100011	
0100	**0110100**	$\mathbf{g}_1$
0101	1100101	
0110	1000110	
0111	0010111	
1000	**1101000**	$\mathbf{g}_0$
1001	0111001	
1010	0011010	
1011	1001011	
1100	1011100	
1101	0001101	
1110	0101110	
1111	1111111	

Linear Systematic Block Codes

n-k bits	*k* bits
Redundant Checking Part	Message Part

p-matrix *kxk- identity matrix*

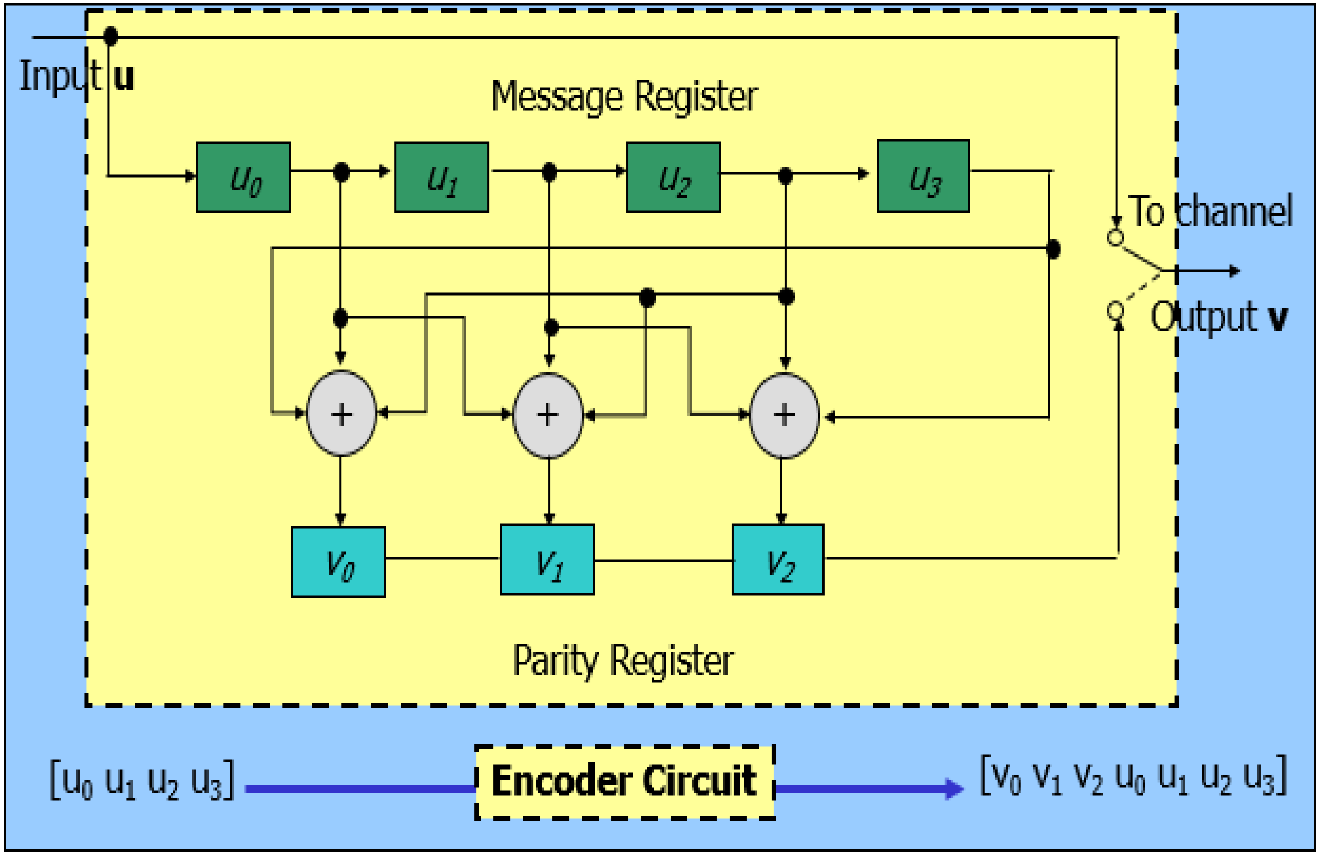
$$\mathbf{G}=[\mathbf{P}\ \mathbf{I}_k] = \left[\begin{array}{cccc|cccc} p_{00} & p_{01} & \cdots & p_{0,n-k-1} & 1 & 0 & \cdots & 0 \\ p_{10} & p_{11} & \cdots & p_{1,n-k-1} & 0 & 1 & \cdots & 0 \\ . & . & & . & . & . & & . \\ . & . & & . & . & . & & . \\ . & . & & . & . & . & & . \\ p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} & 0 & 0 & \cdots & 1 \end{array}\right]$$

The Parity Check Matrix

- For any *k x n* matrix **G** with *k* linearly independent rows, there exists an (*n-k*) *x n* matrix **H** (Parity Check Matrix), such that
  - **G.H^T=0**

$$\mathbf{H}=\left[\mathbf{I}_k \ \mathbf{P}^T\right]=\begin{bmatrix} 1 & 0 & \dots & 0 & p_{00} & p_{01} & \dots & p_{k-1,0} \\ 0 & 1 & \dots & 0 & p_{01} & p_{11} & \dots & p_{k-1,1} \\ . & . & & . & . & . & & . \\ . & . & & . & . & . & & . \\ . & . & & . & . & . & & . \\ 0 & 0 & \dots & 1 & p_{0,n-k-1} & p_{1,n-k-1} & \dots & p_{k-1,n-k-1} \end{bmatrix}$$

Encoding Circuit



$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Error Correcting Power of LBC

- The Hamming distance of a linear block code (LBC) is simply the minimum Hamming weight (number of 1's or equivalently the distance from the all 0 codewords) of the non-zero codewords
- Note $d(c_1,c_2) = w(c_1 + c_2)$ as

Hamming Codes

- Hamming codes are a subclass of linear block codes and belong to the category of perfect codes.
- Hamming codes are expressed as a function of a single integer,

$$m \geq 2$$

$$\begin{array}{ll}
\text{Code length:} & n = 2^m - 1 \\
\text{Number of information bits:} & k = 2^m - m - 1 \\
\text{Number of parity bits:} & n\text{-}k = m \\
\text{Error correction capability:} & t = 1
\end{array}$$

- The columns of the parity-check matrix, **H**, consist of all non-zero binary m-tuples.

Hamming codes

- Example: Systematic Hamming code (7,4)

$$\mathbf{H} = \left[\begin{array}{ccc:cccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{array}\right] = [\mathbf{I}_{3\times3} \;\vdots\; \mathbf{P}^T]$$

$$\mathbf{G} = \left[\begin{array}{ccc:cccc} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] = [\mathbf{P} \;\vdots\; \mathbf{I}_{4\times4}]$$

Cyclic block codes

- Cyclic codes are a subclass of linear block codes.
- Encoding and syndrome calculation are easily performed using feedback shift-registers.
- Hence, relatively long block codes can be implemented with a reasonable complexity.
- BCH and Reed-Solomon codes are cyclic codes.
- A linear (n,k) code is called a Cyclic code if all cyclic shifts of a codeword are also codewords.

$$\boxed{\mathbf{U} = (u_0, u_1, u_2, \ldots, u_{n-1})}$$

"*i*" cyclic shifts of **U**

$$\boxed{\mathbf{U}^{(i)} = (u_{n-i}, u_{n-i+1}, \ldots, u_{n-1}, u_0, u_1, u_2, \ldots, u_{n-i-1})}$$

– Example:

$$\mathbf{U} = (1101)$$

$$\mathbf{U}^{(1)} = (1110) \quad \mathbf{U}^{(2)} = (0111) \quad \mathbf{U}^{(3)} = (1011) \quad \mathbf{U}^{(4)} = (1101) = \mathbf{U}$$

Algebraic structure of Cyclic codes

- Algebraic structure of Cyclic codes, implies expressing codewords in polynomial form

$$\boxed{\mathbf{U}(X) = u_0 + u_1 X + u_2 X^2 + \ldots + u_{n-1} X^{n-1} \qquad \text{degree } (n\text{-}1)}$$

- Relationship between a codeword and its cyclic shifts:

$$\begin{aligned} X\mathbf{U}(X) &= u_0 X + u_1 X^2 + \ldots, u_{n-2} X^{n-1} + u_{n-1} X^n \\ &= \underbrace{u_{n-1} + u_0 X + u_1 X^2 + \ldots + u_{n-2} X^{n-1}}_{\mathbf{U}^{(1)}(X)} + \underbrace{u_{n-1} X^n + u_{n-1}}_{u_{n-1}(X^n+1)} \\ &= \mathbf{U}^{(1)}(X) + u_{n-1}(X^n + 1) \end{aligned}$$

- Hence:

$$\mathbf{U}^{(1)}(X) = X\mathbf{U}(X) \text{ modulo } (X^n + 1)$$

By extension

$$\boxed{\mathbf{U}^{(i)}(X) = X^i \mathbf{U}(X) \text{ modulo } (X^n + 1)}$$

Algebraic structure of Cyclic codes cont..

Definition A code *C* is cyclic if

(i) *C* is a linear code;

(ii) any cyclic shift of a codeword is also a codeword, i.e. whenever $a_0, \ldots a_{n-1}$ Î *C*, then also $a_{n-1}\, a_0 \ldots a_{n-2}$ Î *C*

Example

(i) Code C = {000, 101, 011, 110} is cyclic.

(ii) Hamming code Ham(3, 2): with the generator matrix is equivalent to a cyclic code.

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

(iii) The binary linear code {0000, 1001, 0110, 1111} is not a cyclic, but it is equivalent to a cyclic code.

Error Syndrome

- For error correcting codes we need a method to compute the required correction
- To do this we use the Error Syndrome ‘s’ of a received codeword ‘cr’

$$s = c_r H^T$$

- If cr is corrupted by the addition of an error vector, e, then

$$c_r = c + e$$

and

$$s = (c + e)\, H^T = cH^T + eH^T$$

$$s = 0 + eH^T$$

- Syndrome depends only on the error

Error Syndrome cont…

- That is, we can add the same error pattern to different codewords and get the same syndrome.
- There are 2(n - k) syndromes but 2n error patterns.
- For example for a (3,2) code there are 2 syndromes and 8 error patterns.
- Clearly no error correction possible in this case.
- Another example. A (7,4) code has 8 syndromes and 128 error patterns.
- With 8 syndromes we can provide a different value to indicate single errors in any of the 7 bit positions as well as the zero value to indicate no errors.
- Now need to determine which error pattern caused the syndrome.

Error Syndrome cont…

- For systematic linear block codes, H is constructed as follows,

  $G = [\, I \mid P]$ and so $H = [-P^T \mid I]$

  where I is the $k{*}k$ identity for G and the $R{*}R$ identity for H

- Example, (7,4) code, $d_{min} = 3$

$$G = [I \mid P] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad H = [-P^T \mid I] = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Error Syndrome - Example

- For a correct received codeword c_r = [1101001]

  In this case,

$$s = c_r H^T = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

Error Syndrome - Example

- For the same codeword, this time with an error in the first bit position, i.e.,

  c_r = [1101000]

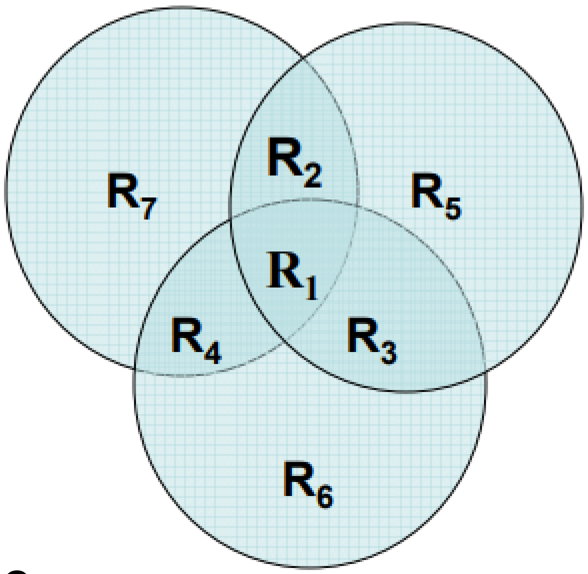
$$s = c_r H^T = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

- In this case a syndrome 001 indicates an error in bit 1 of the codeword

SYNDROME DECODING

- Let R 1 R 2 R 3 R4 R 5 R 6 R7 be the received block of binary digits, possibly with errors.
- Counting 1's in the circles is the same as computing the result of the following equations:

$$R_1 \oplus R_2 \oplus R_3 \oplus R_5 = S_1$$
$$R_1 \oplus R_3 \oplus R_4 \oplus R_6 = S_2$$
$$R_1 \oplus R_2 \oplus R_4 \oplus R_7 = S_3$$



- S 1, S 2 and S 3 is called the syndrome

SYNDROME DECODING cont..

- Thus to correct a single error based upon the received sequence R 1, R 2, R 3, R 4, R5, R6, R7.
- One can first compute the syndrome S1, S 2, S3 , and then compare it with the columns of the parity check matrix.
- The matching column is where the error occurred.
- This technique will work for any single error correcting code.

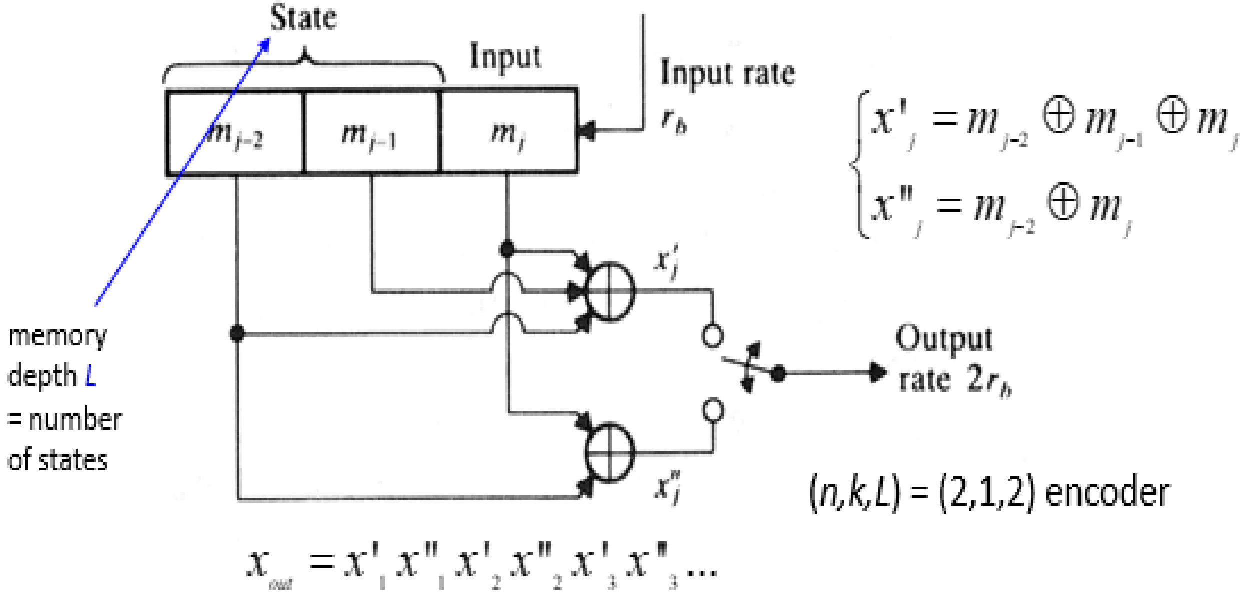
Convolutional codes

- Convolutional codes map information to code bits sequentially by convolving a sequence of information bits with "generator" sequences.
- A convolutional encoder encodes K information bits to N>K code bits at one time step.
- Convolutional codes can be regarded as block codes for which the encoder has a certain structure such that we can express the encoding operation as convolution.
- Convolutional codes are applied in applications that require good performance with low implementation complexity. They operate on code streams (not in blocks).

Convolutional codes

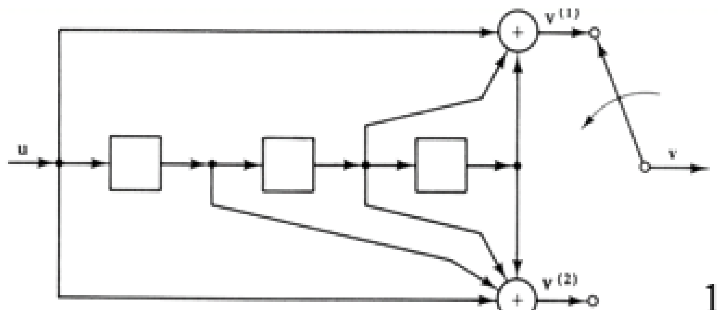
- Convolution codes have memory that utilizes previous bits to encode or decode following bits (block codes are memoryless)
- Convolutional codes are denoted by (n,k,L), where L is code (or encoder) Memory depth (number of register stages).
- Constraint length C=n(L+1) is defined as the number of encoded bits a message bit can influence to.
- Convolutional codes achieve good performance by expanding their memory depth.
- Convolutional encoder is a finite state machine (FSM) processing information bits in a serial manner.

Example: Convolutional encoder, k = 1, n = 2



- Thus the generated code is a function of input and the state of the FSM.
- In this $(n,k,L) = (2,1,2)$ encoder each message bit influences a span of *C= n(L+1)=6* successive output bits = **constraint length *C*.**
- Thus, for generation of *n*-bit output, we require in this example *n* shift registers in *k* = 1 convolutional encoder.

Time domain Approach Using generator matrix



$$\begin{pmatrix} \mathbf{g}^{(1)} = [1\ 0\ 1\ 1] \\ \mathbf{g}^{(2)} = [1\ 1\ 1\ 1] \end{pmatrix}$$

If $\mathbf{u} = (1\ \ 0\ \ 1\ \ 1\ \ 1)$, then

$$\mathbf{v} = \mathbf{u}\mathbf{G}$$

$$\underbrace{\underbrace{11 \oplus 00}_{11} \oplus \underbrace{01 \oplus 11}_{10}}_{01} = 01$$

$\mathbf{g}_0^{(1)}\mathbf{g}_0^{(2)}\ \ \mathbf{g}_1^{(1)}\mathbf{g}_1^{(2)}\ \ \ \ \mathbf{g}_m^{(1)}\mathbf{g}_m^{(2)}$

$$v_l^{(j)} = u_l g_0^{(j)} + u_{l-1} g_1^{(j)} + \dots + u_{l-m} g_m^{(j)}$$

$$= \underset{u_{l-m}}{(1\ \ 0\ \ 1\ \ 1\ \ 1)} \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & & & & & & & & \\ & & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & & & & & & \\ & & & & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & & & & \\ & & & & & & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & & \\ & & & & & & & & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= (1\ \ 1,\ \ 0\ \ 1,\ \ 0\ \ 0,\ \ \boxed{0}\ \ 1,\ \ 0\ \ 1,\ \ 0\ \ 1,\ \ 0\ \ 0,\ \ 1\ \ 1),$$

Transform Domain Approach cont..

The D-transform is a function of the indeterminate D (the delay operator) and is defined as:

$$u_i \multimap U_i(D) = \sum_{t=0}^{\infty} u_{i,t} D^t$$

$$v_j \multimap V_j(D) = \sum_{t=0}^{\infty} v_{j,t} D^t$$

$$g_j^{(i)} \multimap G_{i,j}(D) = \sum_{t=0}^{\infty} g_{j,t}^{(i)} D^t$$

Transform Domain Approach cont..

- The convolutional relation of the D transform
- $D\{u * g\} = U(D)G(D)$ is used to transform the convolution of input sequences and generator sequences to a multiplication in the D domain.
- Ex: (2,1,2) convolutional code encoder :

$$G^{(1)}(D) = 1 + D + D^2$$

$$G^{(2)}(D) = 1 + D^2$$

$$G(D) = [1 + D + D^2, 1 + D^2]$$

$$U(D) = 1 + D + D^2 + D^4$$

Transform Domain Approach cont..

- Ex: (2,1,2) convolutional code encoder :

$$V(D) = U(D) \cdot G(D)$$

$$V1(D) = U(D) \cdot G(1)(D)$$

$$V1(D) = (1 + D + D^2 + D^4) \text{ x } (= 1 + D + D^2)$$

$$= (1 + D + D^2 + D^4 + D + D^2 + D^3 + D^5 + D^2 + D^3 + D^4 + D^6)$$

$$V1(D) = 1 + D^2 + D^5 + D^6$$

Transform Domain Approach cont..

- Ex: (2,1,2) convolutional code encoder :

  V (D) = U(D) • G(D)

  V1(D)= U(D) • G(1)(D)

  V1(D)= $(1 + D + D^2 + D^4)$ x $(= 1 + D + D^2)$

  $= (1 + D + D^2 + D^4 + D + D^2 + D^3 + D^5 + D^2 + D^3 + D^4 + D^6)$

Transform Domain Approach cont..

- Ex: (2,1,2) convolutional code encoder :

  $V1(D) = 1 + D^2 + D^5 + D^6$

  $V2(D) = 1 + D \quad + D^3 + D^6$

  In the time domain :

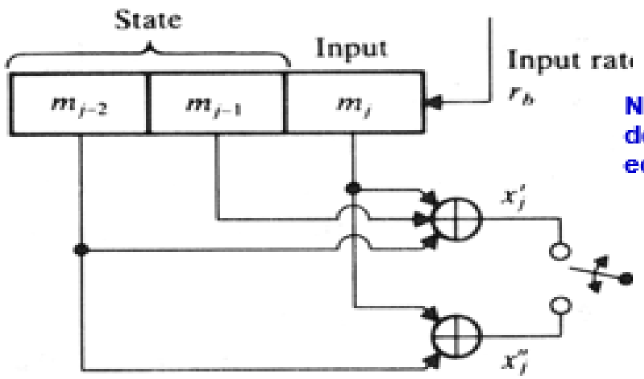
  $v_1 = (1010011)$

  $v_2 = (1101001)$

  Then the code word is :

  $v \quad = (11\ 01\ 10\ 01\ 11\ 10\ 11)$

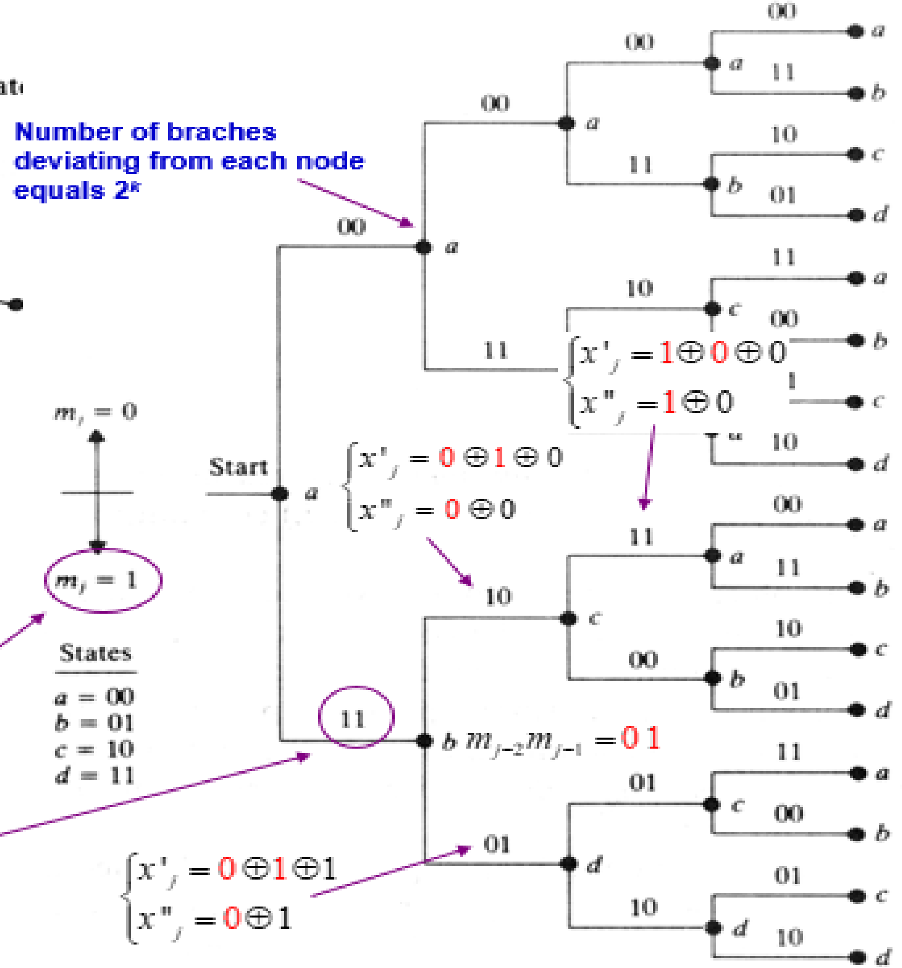
Representing convolutional codes: Code tree



$(n,k,L) = (2,1,2)$ encoder

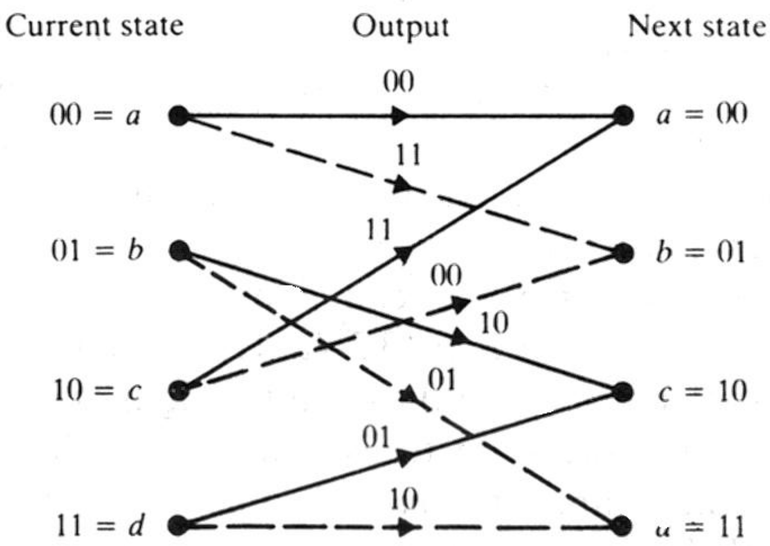
$$\begin{cases} x'_j = m_{j-2} \oplus m_{j-1} \oplus m_j \\ x''_j = m_{j-2} \oplus m_j \end{cases}$$

$$x_{out} = x'_1 x''_1 x'_2 x''_2 x'_3 x''_3 \ldots$$

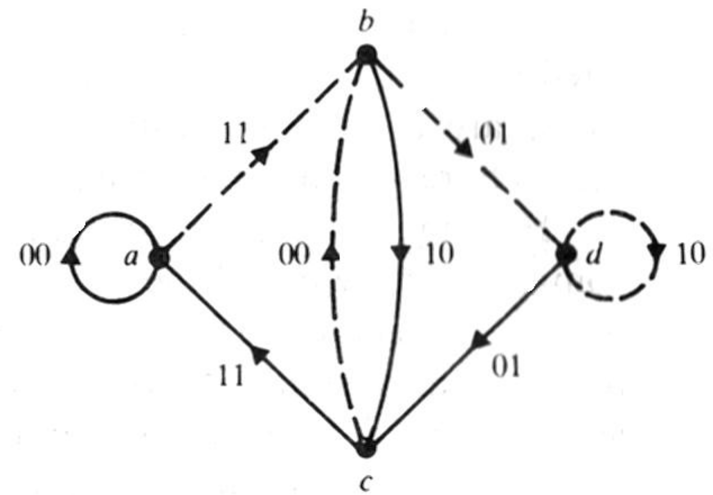


This tells how **one input bit**
is transformed into **two output bits**
(initially register is all zero)

Code trellis and state diagram



(a)



(b)

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

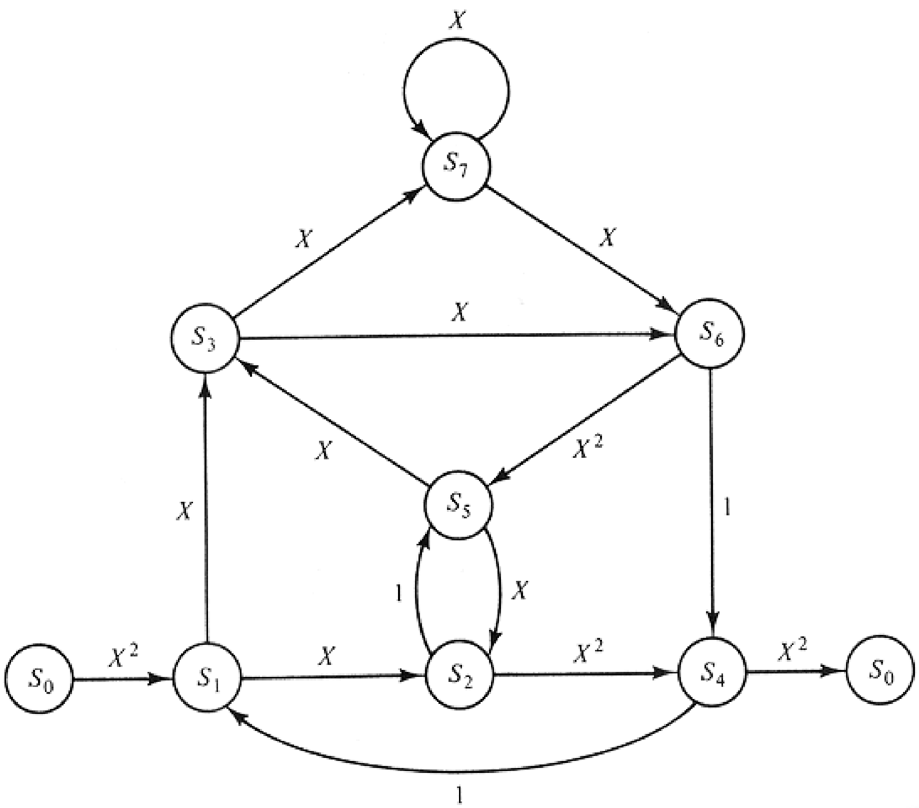
(c)

Viterbi algorithm

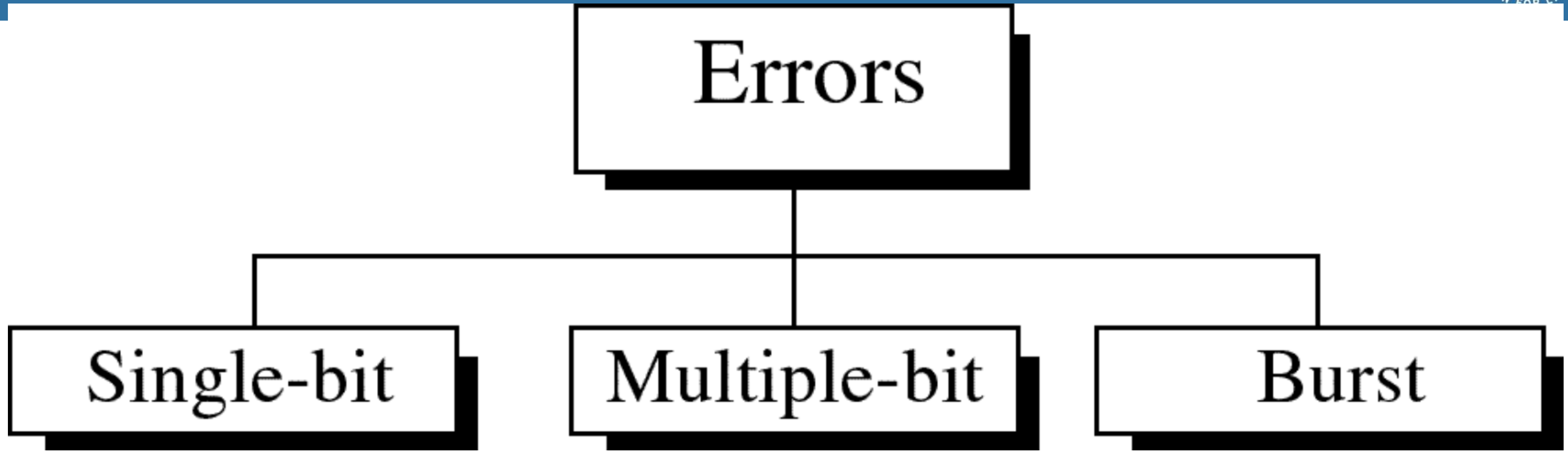
- Problem of optimum decoding is to find the minimum distance path from the initial state back to the initial state (below from S_0 to S_0). The minimum distance is one of the sums of all path metrics from S_0 to S_0.
- Exhaustive maximum likelihood method must search **all the paths** in phase trellis (2^k paths emerging/entering from 2^{L+1} states for an (n,k,L) code).
- The Viterbi algorithm gets improvement in computational efficiency via concentrating into **survivor paths** of the trellis.
- The most commonly employed decoding technique that can be implemented using either software or digital hardware.

Viterbi algorithm Cont..

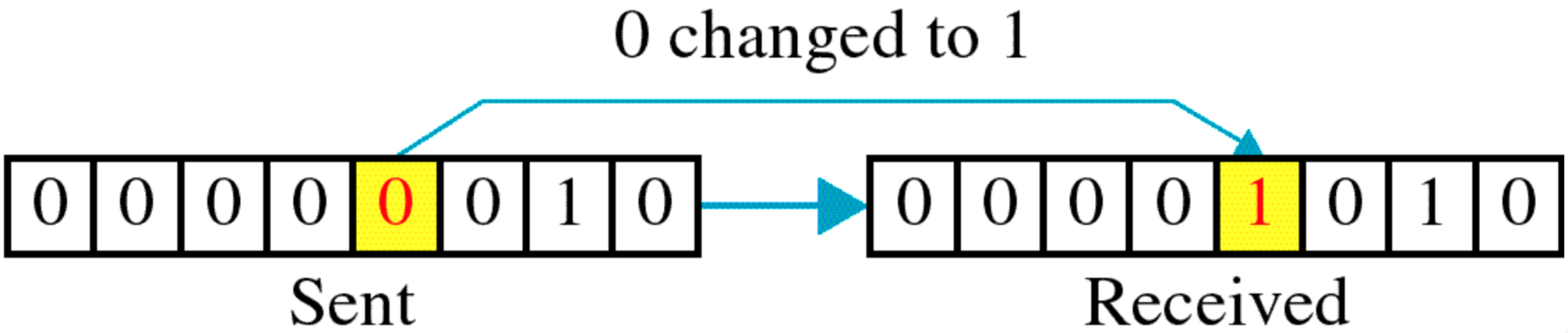
- VA uses the trellis diagram (Fig.2) and can theoretically perform maximum likelihood decoding.
- It finds the most likely path by means of suitable distance metric between the received sequence and all the trellis paths.



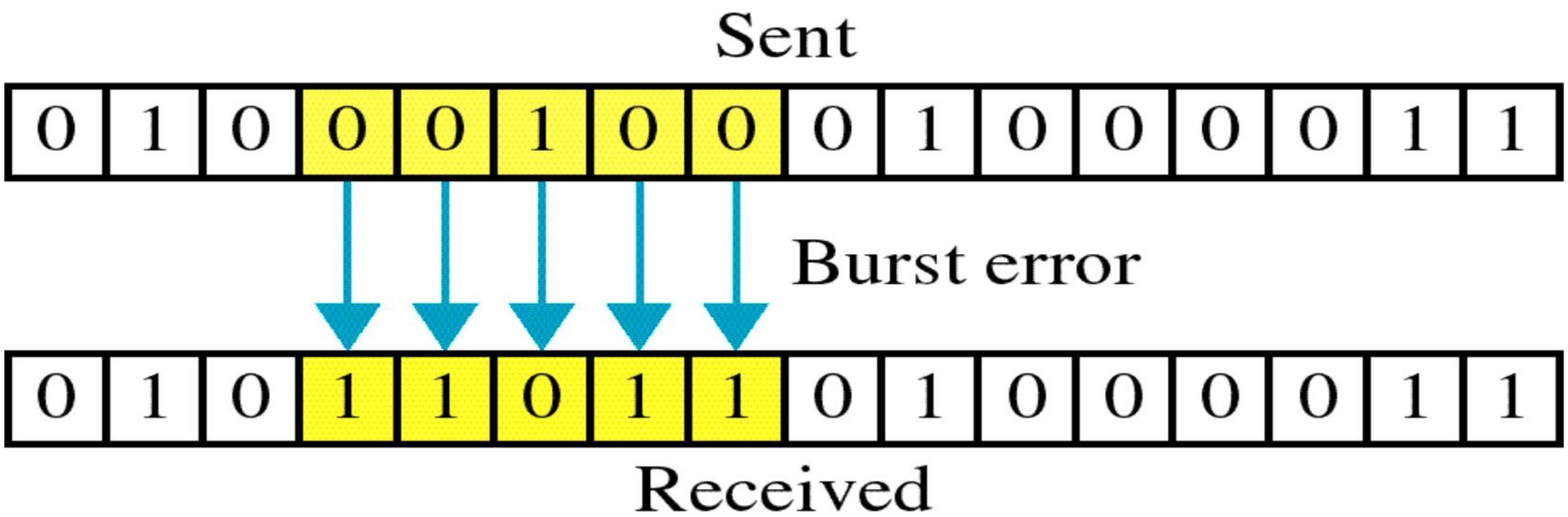
Types of Errors



Single-bit error:



Burst error



- The term burst error means that two or more bits in the data unit have changed from 1 to 0 or from 0 to 1.
- Burst errors does not necessarily mean that the errors occur in consecutive bits, the length of the burst is measured from the first corrupted bit to the last corrupted bit.
- Some bits in between may not have been corrupted.

Error detection & Error Correction

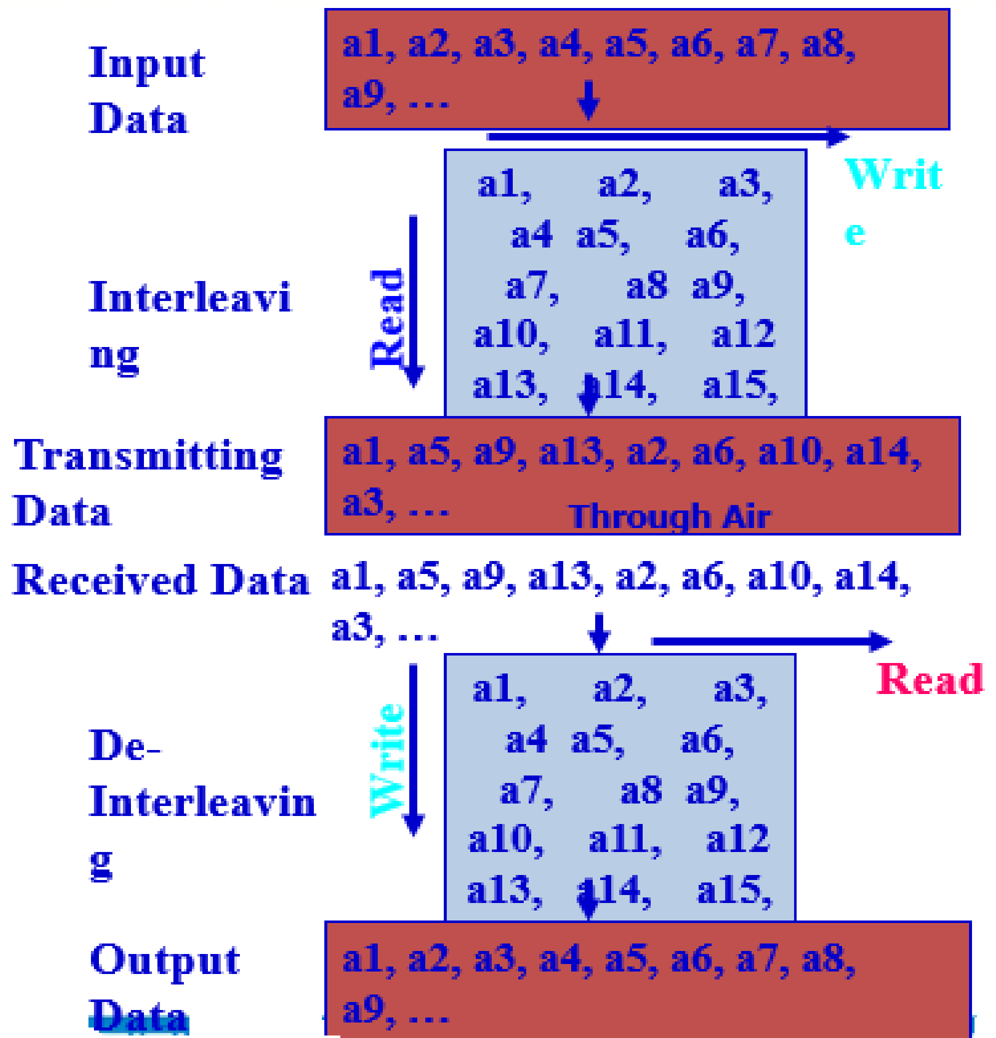
- **Error detection** means to decide whether the received data is correct or not without having a copy of the original message.
- Error detection **uses the concept of redundancy**, **which means** adding extra bits for detecting errors at the destination.

Error Correction

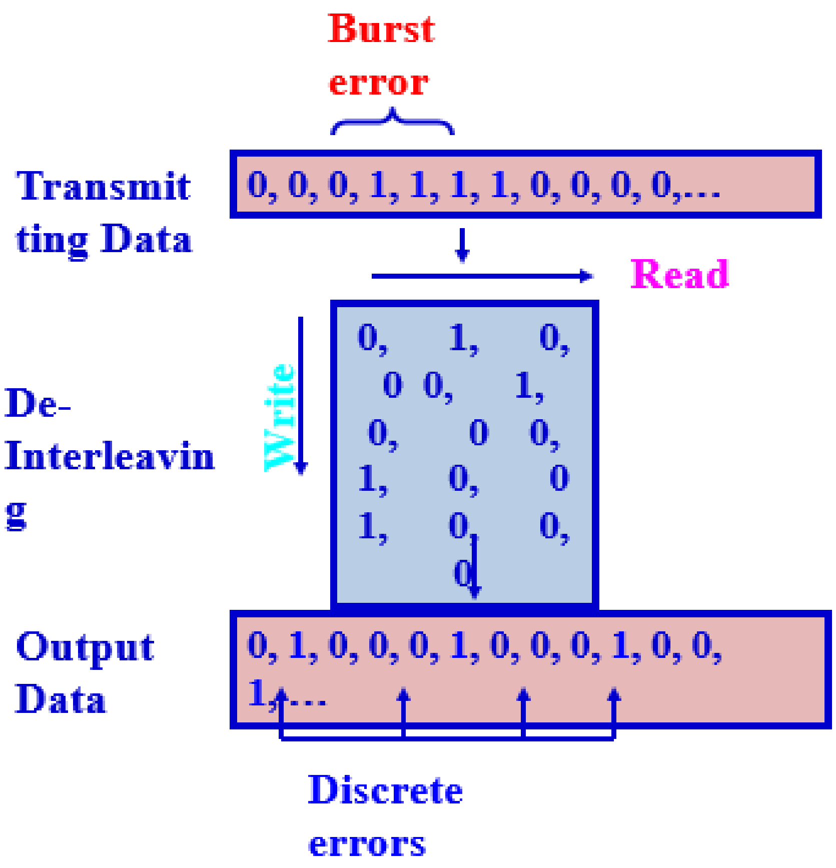
It can be handled in two ways:

1) receiver can have the sender retransmit the entire data unit.
2) The receiver can use an error-correcting code, which automatically corrects certain errors.

Interleaving



Interleaving Example



Block Interleaving

- Data written to and read from memory in different orders
- Data bits and corresponding check bits are interspersed with bits from other blocks
- At receiver, data are de-interleaved to recover original order
- A burst error that may occur is spread out over a number of blocks, making error correction possible

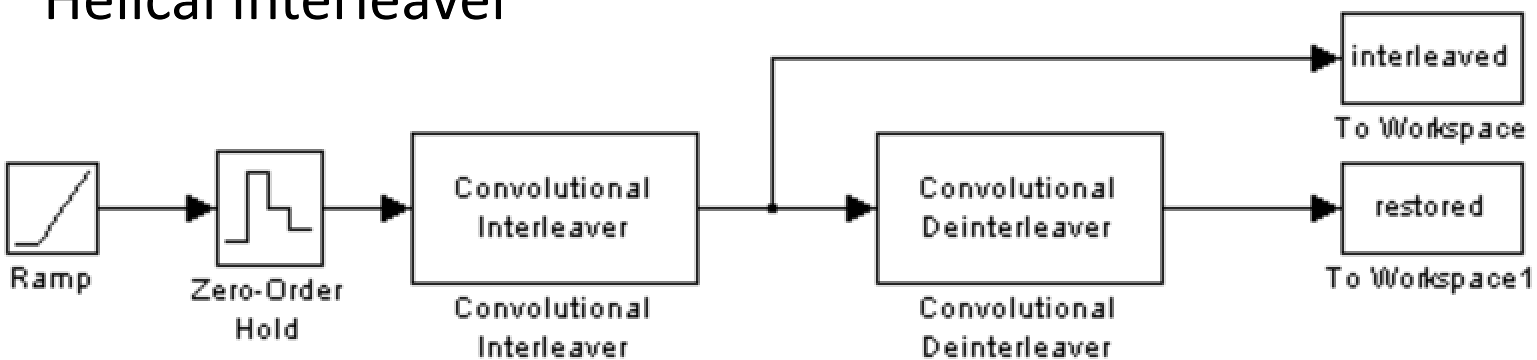
Convolutional Interleavers

- A convolutional interleaver consists of a set of shift registers, each with a fixed delay.
- In a typical convolutional interleaver, the delays are nonnegative integer multiples of a fixed integer (although a general multiplexed interleaver allows arbitrary delay values).
- Each new symbol from the input signal feeds into the next shift register and the oldest symbol in that register becomes part of the output signal.
- The schematic below depicts the structure of a convolutional interleaver by showing the set of shift registers and their delay values D(1), D(2),..., D(N).
- The blocks in this library have mask parameters that indicate the delay for each shift register. The delay is measured in samples.

Types of Convolutional Interleavers

- The set of convolutional interleavers in this library includes a general interleaver/deinterleaver pair as well as several special cases.
- Each special-case block uses the same computational code that its more general counterpart uses, but provides an interface that is more suitable for the special case.

1. General Multiplexed Interleaver
2. Convolutional Interleaver
3. Helical Interleaver




Thank you