TABLE OF CONTENTS

1 day

1 day

1/2 day

DISE CECM

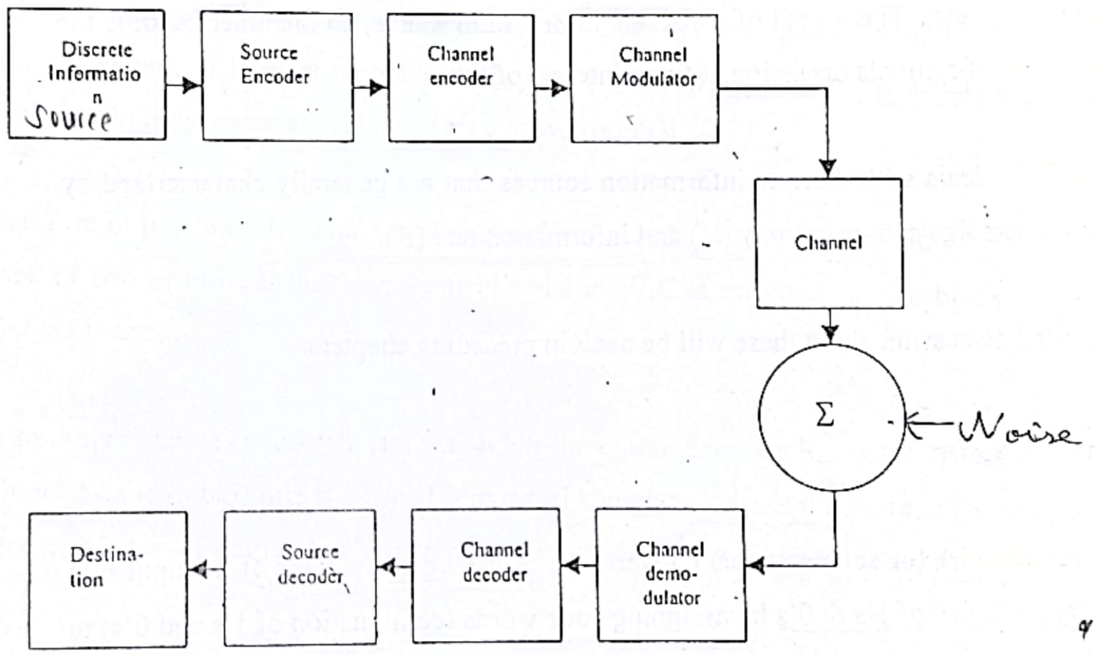
Disom

Chapter 1 Digital Communication Systems

1.1 Digital Communication Sources, Transmitters, Transmission Channels, and Receivers

Digital Communication Systems (DCS) is meant to transmit signals (message or sequence of symbols) in discrete form coming out from a source to a pre- assigned destination with maximum possible rate and accuracy.

The overall functional Block diagram of a DSC is shown below.



The signals at the output of the information source and at the input of the destination are sequence of symbols in discrete form (or sampled version of continuous signal).

The communication channel is a physical media that accept signal in the form of electrical parameters (voltage , current) for the case of wire communication & electro- magnetic wave (EMW) for the case of wireless communication.

Scanned with CamScanner

The signal at the output of communication channel is usually distorted and full of additive noise. The receiver block is so designed that it detects the presence of useful signal in a mixture of signal and noise with highest possible accuracy.

Information source:

There is a standard classification of information source based on the nature of electrical signal produced at its output.

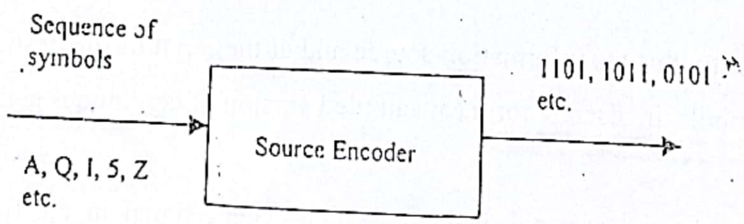
The output of an analog information source is electrical signal, which is a continuous function of time. Examples are the output of a microphone, TV, camera, radio receivers, etc. The output of a discrete information source, on the other hand, is the sequence of symbols occurring at fixed interval of time.

The DCS deals with discrete information sources that are generally characterized by parameters like source entropy (H) and information rate (R).

Detailed discussion about these will be dealt in preceding chapters.

Source encoder:

Source encoder (or source coder) converts the sequence of symbols at its input into binary sequence of 1's & 0's by assigning code words (combination of 1's and 0's) to each symbol at its input.



Scanned with CamScanner

For example the output of an English Teletype is sequence of symbols containing total 32 different symbols. To assign code words to each symbol we need at least 5-bit code i. e. form 00000 to 11111.

Now if the symbol rate is 10 symbols \ sec then the data rate at the output of source encoder will de 10 x 5=50 bits /sec.

This is an example of fixed length coding (FLC). FLC is efficient only if the input symbols occur with equal probability and are statistically independent. For example in English text, probability of occurrence of symbols a, e, i, etc. is much more higher than symbols Q or V or Z. Secondly the symbol occurrence sequence is statistically not independent. It is most probable that symbol 'U' will occur after symbol Q and consonant in most cases are followed by a vowel symbol.

Therefore in practice FLC is not used, instead variable length code (VLC) in which a block of two or more symbols are considered and VLC is assigned to each block of symbols is used.

An optimum source encoder is that which has the output data rate R_{out} nearly or equal to input data (symbol) rate R_i. The efficiency of source encoder is therefore measured in terms of ratio of output data rate to minimum theoretically achievable data rate.

Source decoder

Source decoder converts the binary output of the channel decoder into sequence of symbols with minimum error (due to bit errors) and maximum efficiency.

Channel encoder

Channel encoder (CE) is used to enhance reliability & efficiency of high-speed digital signal transmission. In general the output of source encoder in the binary form can be

Scanned with CamScanner

directly fed to modulator. In this case loss or improper detection of any of the information bearing bit in the receiver side may distort the complete word or even the complete message.

The CE therefore adds to the bit stream of source encoder output some error control bits (redundancy) that does not carry any information These error control bits make possible for the receiver to detect and-, in most cases, correct some of the errors in message bearing bits.

Usually in each block of 'k' information bearing bits, 'r' error control bits are added. This type of coding method is called block coding. In other method, called convolution method, information bit & error control bits are continuously inter-leaved. All channel encoding methods require storage & processing of the information.

Basic parameters of channel encoder are : method of coding, rate or efficiency of the coder, error control capabilities, complexity etc.

Channel decoder:

Channel decoder (CD) recovers information bearing bit streams from coded bit streams with minimum error and maximum efficiency. Complexity of the decoder and the time delay are the two basic design criteria of the channel decoder

Channel modulator:

Channel modulator (CM) is intended to convert bit streams from CE to electrical waveform suitable for transmission over communication channel. The proper design of CM can effectively minimize effects of noise, increase matching of signal characteristic (frequency, power and bandwidth) with channel characteristics and provide multiple data communication over the same physical channel. (multiplexing).

Scanned with CamScanner

~~Channel decoder~~ Channel Demodulator

Channel decoder (CD) converts received electrical signal into sequence of bits with minimum error & maximum efficiency.

~~The channel~~ :

The communication channel (CH) is a physical media (cable for wire communication or free space for radio communication) with parameters that limit the speed or rate of data communication.

CH has finite frequency bandwidth. The signal power is attenuated as it travels along the CH and noise is introduced (multiplicative). Because of power attenuation and noise, the received signal is distorted.

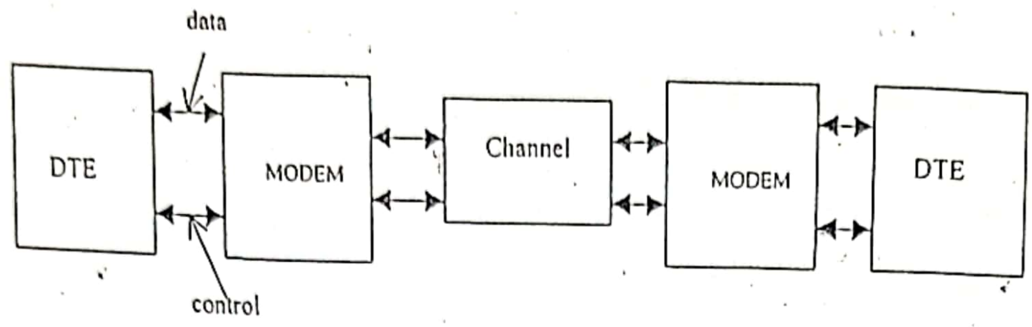
Shannon, Hartley, & Nyquist played significant role in characterizing CH in terms of its capacity to transmit signal with minimum acceptable signal to noise ratio (SNR) or equivalent Bit Error Rate (BER).

The Shannon channel capacity theorem state that for given *B* (channel bandwidth) and required level of *SNR*, the maximum speed of data transmission (or the channel capacity) *C* is limited i.e. for given *B* & *SNR* only *C* can maintain errorless communication.

$$\boxed{C = B \log_2 (1+SNR) \text{ bits/S}}$$

From users point of view the complete DCS comprises of Data Terminal Equipment (Teletype, computer , data-logger etc) and modulator-demodulator (MODEM) units.

Scanned with CamScanner



The DTE generates & receives information signal as well as the control signal and the communication between DTE & MODMS are made through standard interface, such as RS-232C port as specified by Electronics Industries Association.

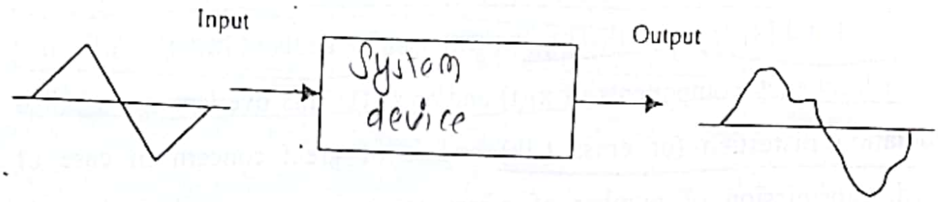
The control signals include request to send data, confirmation of readiness to receive data, busy DTE etc. The overall control of information flow, error detection, error correction and coding are generally performed by DTE.

Scanned with CamScanner

1.1 Distortion, Noise and Interference

Distortion

Distortion is the unwanted change in the waveform at the output of any device in comparison to the input waveform.



Distortion is categorized into Linear and Non-linear type. Linear distortion is produced by linear devices and result in scaling of amplitude and phase shift of the input signal. Linear devices because of non-uniform frequency and phase responses produce linear distortion. Because of non-uniformity of the amplitude and phase responses of the system, different frequency components of the input signal experience different gain (amplitude distortion) and phase shift (phase or delay distortion). A linear device is said to be distortionless if for some limited frequency band $f < f_{max}$, the magnitude of the transfer function is constant and the phase is linearly proportional to the frequency. For digital transmission, the phase distortion or the delay distortion is more critical than the amplitude distortion.

Non-linear (NL) distortion is produced due to non-linear characteristics of the input-output characteristics of the system or device. Because of this additional frequency components, not present in the input signal spectrum, are produced at the output of the NL device.

In general, the input-output characteristics of a device/system can be approximated using the following equation:

Scanned with CamScanner

$$y(t) = a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \ldots\ldots\ldots\ldots$$

The first component of the right hand side of the above equation produces the exact replica of the input signal scaled by some constant factor. The remaining other terms produce non-linear distortion.

Special case of NL distortion is Inter-modulation (IM) distortion. When the sum of two different signals $x_1(t)$ and $x_2(t)$ having different frequency bandwidths pass through a NL device, the output may contain, apart from $x_1(t)$ and $x_2(t)$, components like $x_1^2(t)$, $x_2^2(t)$ and $[x_1(t) \times x_2(t)]$. The frequency components of $[x_1(t) \times x_2(t)]$ may overlap with frequency components of $x_1(t)$ and/or $x_2(t)$. This overlapping is called Inter-modulation distortion (or cross talk) and is of great concern in case of multiplexed transmission of number of messages over common channel. It is especially matter of concern if IM product like $2f_1 \pm f_2$ or $2f_2 \pm f_1$ are very significant.

Harmonic distortion (HD), i.e. generation of harmonics of fundamental frequency of the input signal due to non-linearity of the device can be evaluated as the ratio of the amplitude of n-th harmonic component and the amplitude of the fundamental frequency expressed in %.

$$\%\text{ n-th harmonic distortion} = \%\, D_n = \frac{|A_n|}{|A_1|} \times 100\%$$

Total harmonic distortion (THD) is expressed as

$$\%THD = \sqrt{D_2^{\,2} + D_3^{\,2} + D_4^{\,2} + \ldots + \ldots} \times 100\%$$

Interference

Interference is referred to as the contamination of received signal by other extraneous, usually manmade, signals similar to the desired signal. Manmade interference is usually signals from various broadcasting & communication Systems.

Inter symbol Interference (ISI) is the special case of interference in digital signal transmission when the interference is occurred within the system itself. ISI is the effect of all other transmitted bits on decoding of m-th received bit.

Noise

Noise is the unwanted manmade or natural random signal that adds to the received signal and degrades the performance of the Communication Systems.

a) Thermal noise

Thermal noise is produced due to random movement of free electrons within the conducting portion of electrical circuitry of the system. The average thermal noise power can be estimated using the following formula:

$$\boxed{P_n = KTB \text{ watts,}}$$

Where

P_n =noise power (average)

K = Boltzmann's constant (1.38×10^{-23} J/ ^{0}K)

T = Temperature of the conductor in ^{0}k

B = Bandwidth of noise spectrum , Hz

The psdf of thermal noise spectrum is equal to:

$$\boxed{S_n = KT \text{ (Watt/Hz)}}$$

Scanned with CamScanner

b) Shot noise

This is a term originally used to describe anode current noise in vacuum tubes resulting from random fluctuation in electron emission from the cathode. In semiconductor devices the shot noise is used to describe variation in number of electrons crossing potential barrier.

c) Partition noise

Partition noise is the result of random fluctuation in division when current is divided into two or more paths. In semiconductor transistors the shot noise is produced when the emitter current is divided into base and collector currents. That is why in Microwave receivers the received signals are directly fed to the diode mixers (as diodes do not produce partition noise).

d) Flicker Or Low frequency noise

Below frequencies of few kilo Hertz, a noise appear in the devices whose spectral density increases as the frequency is decreased is called Flicker or Low frequency noise (sometimes called 1/f noise). In semiconductors flicker noise is the result of fluctuation in carrier density.

e) Transit-Time or HF noise

In semiconductors, if the signal period is very low (i.e. frequency is very high) some of the carriers may diffuse back to the source before crossing the junction barrier and produce noise. The *psdf* of this kind of noise increases with frequency.

f) Generation – Recombination noise

Random process of generation & recombination of free electrons in semiconductor devices due to random ionization of impurities produce this kind of noise.

Scanned with CamScanner

1.3 Sampling Theory

Sampling of continuous analog signal is first step of transmission of analog signal over digital communication System.

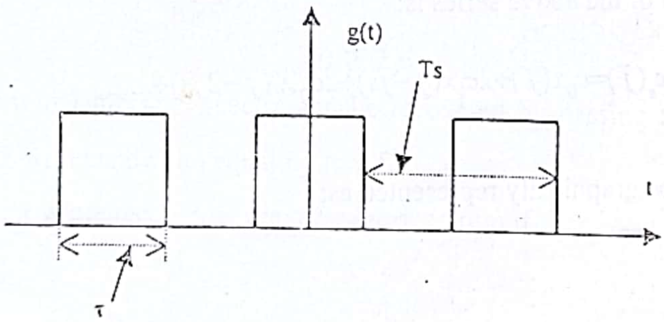
In general, the sampling theorem states "Analog signal can be reproduced from an appropriate set of its samples taken at some fixed intervals of time". This theorem has made possible to transmit only samples of analog signal by changing or encoding these samples into block of code words suitable for digital communication.
In a broad sense the sampling theorem can be stated in two parts.

1. A strictly band limited signal (i.e. for $f > B$, there is no energy) is completely described by the samples (values) of the signal at instants of time separated by $1/2B$ seconds (for transmitting end).
2. The original signal may be recovered if we know its values (samples) taken at the rate of $2B$ per second (for receiving end).

If the signal x(t) to be sampled is band limited then the sampled signal can be represented as:

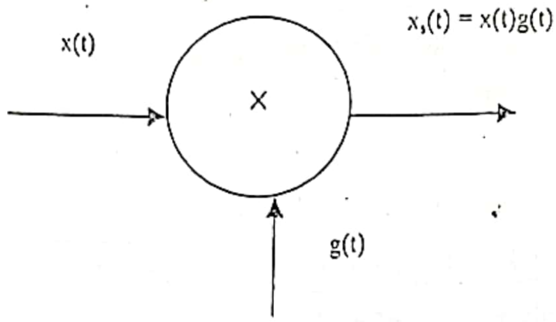
$$X_s(t) = X(t) \times g(t)$$

Where g(t) is the sampling function shown below



Where, Ts is called sampling period and τ-duration of sampling pulse.
The sampler can be implemented using the following arrangement:

Scanned with CamScanner



To prove the two sampling theorem stated above, let's find the spectra of $x_s(t)$. The gate function g(t) can be expressed in terms of Fourier series as:

$$g(t) = c_0 + \sum_{n=1}^{\infty} 2c_n \cos n\omega_s t$$

Where,

$$c_0 = \tau / T_s,\ c_n = f_s \tau\, Sinc[nfs\,\tau],\ \omega_s = 2\pi fs$$

Then the $x_s(t)$ can be expressed as:

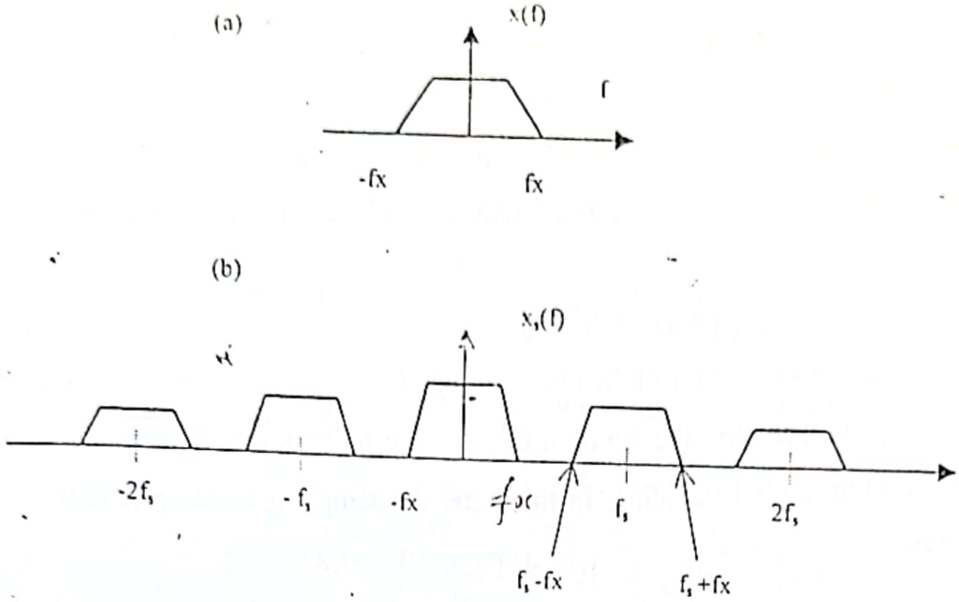
$$x_s(t) = c_0 x(t) + 2c_1 x(t) \cos\omega_s t + 2c_2 x(t) \cos 2\omega_s t + \ldots\ldots + 2c_n x(t) \cos n\omega_n t + \ldots$$

The Fourier transform of the above series is:

$$x_s(f) = c_0 x(f) + 2c_1 x(f - fs) + 2c_2 X(f - 2fs) + \ldots\ldots$$

The above series can be graphically represented as:

Scanned with CamScanner



The figure (a) above represent the FT of the original signal x(t) and the figure (b) is the spectrum of the signal at the output of the sampler. It is clear from the figure (b) that the spectrum of the sampled signal contains the spectrum of the original message signal.

It is evident that for distortionless recovery of original message signal from the spectrum of the sampled signal, the following condition should be met:

$$fs - fx \geq fx$$

Or

$$fs \geq 2fx$$

In this case the original message spectra can be recovered by passing the sampled signal through LPF with bandwidth equaling to $\pm f_x$.

Distortion will occur while recovering the message spectrum if

$$fs - fx \leq fx$$

Or

$$fs \leq 2fx$$

Scanned with CamScanner

The distortion in the above case is caused by overlapping of side bands and message spectra. This distortion is referred to as aliasing effect.

The minimum sampling rate $f_{smin}=2f_x$ is called Nyquist's sampling rate for distortion less recovery of message spectrum. The minimum interval of sampling for a real signal is:

$$T_{smin}=1/2f_{max}$$

where f_{max} is the maximum frequency of the message signal.

Instantaneous sampling (i.e. when the duration of sampling pulse $\tau_s \to 0$ or the delta function) is referred to as ideal sampling. In this case, the sampling function is the train of impulses.

$$s_{\delta}(t) = \sum_{k=-\infty}^{\infty} \delta(t - kTs)$$

Therefore the sampled signal for ideal sampling can be written as:

$$x_{\delta}(t) = x(t) S_{\delta}(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT_s)$$

The Fourier transform of the above expression is the convolution of FT's of $S_\delta(t)$ and x(t)

$$X_\delta(f) = X(f) * S_\delta(f)$$

Where

$$S\delta(f) = fs \sum_{k=-\infty}^{\infty} \delta(f - nfs)$$

$$X_\delta(f) = fs \sum_{n=-\infty}^{\infty} X(f - nfs)$$

The result of the convolution is:

Assuming that the condition

$$f_s \geq 2 f_x$$

Scanned with CamScanner

is satisfied, let us pass the sampled signal through a ideal LPF with the following parameters:

$$H_{LPF}(f) = \{k, \quad for \quad f_x \le B \le f_s - f_x \text{ and 0 otherwise}$$

At the output of the LP filter, we get

$$X_\delta(t)H_{LPF}(f) = [f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)]H_{LPF}(f)$$

As $H_{LPF}(f) = 0$ for frequency components higher than fx, the spectrum at the output of LPF for $k = 1/f_s$ will be x(f) – the spectrum of the input signal.

In time domain,

$$x(t) = FT^{-1}\{X_\delta(f)\, H_{LP}(f)\}$$

$$= x_\delta(t) * h_{LPF}(t), \text{ where}$$

$h_{LPF}(t)$ is the impulse response of ideal LP filter which is expressed as

$$h_{LPF}(t) = 2B\, Ts\, \mathrm{sinc}[2Bt]$$

and the ideally sampled signal $x_\delta(t)$ can be expressed as:

$$x\delta(t) = \sum_{k=-\infty}^{\infty} x(kTs)\delta(t - kTs)$$

Finally, the signal at the output of the ideal LPF will be:

$$x(t) = 2BTsSinc[2Bt] * [\sum_{k=-\infty}^{\infty} x(kTs)\delta(t - kTs)]$$

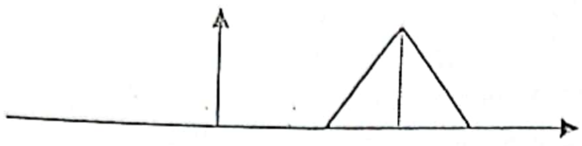
$$= 2BT_s \sum_{k=-\infty}^{\infty} x(kT_s) Sinc\, 2B(t - kT_s)$$

The above equation shows that the original message x(t) can be reconstructed without distortion from its sample values x(kTs}, if the sampling is done at $T_s <= 1/2f_x$.

Scanned with CamScanner

1.8 Sampling of Band Pass Signals

The maximum frequency (f_{xu}) occupied by the band pass signals is usually so high that the Nyquist rate $f_s \geq 2f_{xu}$ becomes quite high for D/A or A/D conversion.



The available solution is either to down convert the signal, sample at lower Nyquist rate, up convert to original position; or apply any other special method. The former method is not very suitable as it requires many steps of signal processing. The special method is referred to as band pass sampling theorem which states that – If a band pass signal is sampled at

$$f_{ss} = \frac{2f_{xu}}{m}, \text{ where}$$

$$m = \frac{f_{xu}}{B}, \text{ is the largest integer below } \frac{f_{xu}}{B}.$$

then the original message signal can be recovered with out distortion. This theorem is also called Sub-sampling theorem . The sub-sampling frequency f_{ss} is much less than the Nyquist rate $f_s > 2\ f_{xu}$.

Practical considerations

All the sampling methods discussed earlier were supposed to be ideal, and therefore recovering the original message signal was distortionless. But in practice due to various implementation constraints, the expected result may differ from ideal. Followings are the differences between ideal and practical sampling:

- The sampled waveform consists of finite amplitude and duration pulses rather than ideal impulses;
- Reconstruction filters are not ideal;and
- The input waveforms are rather time limited than bandlimited

Scanned with CamScanner

The effect of the above deviations can be analyzed in the following manner.

1. The real sampling pulse is flat topped (in flat topped sampling) with finite duration τ. The net result is that

$$x_s(t) = \sum_{k=-\infty}^{\infty} x(kT_s)p(t-kT_s)$$

where p(t) is the sampling pulse of duration τ. In other words, finding convolution of flat-topped pulse and the ideally sampled signal can derive the real sampled signal:

$$x_s(t) = [p(t)] * [\sum_{k=-\infty}^{\infty} x(kT_s)\delta(t-kT_s)]$$

And the spectrum of the sampled signal will be:

$$x_s(f) = P(f)X_\delta(f) = P(f)[f_s \sum_{n=-\infty}^{\infty} X(f-nf_s)]$$

Since P(f) is a sinc function, the primary effect of flat-topped sampling is the attenuation of high frequency components of the message signal. This effect is also called Aperture effect. This effect can be neutralized by using equalizing filter with a transfer function

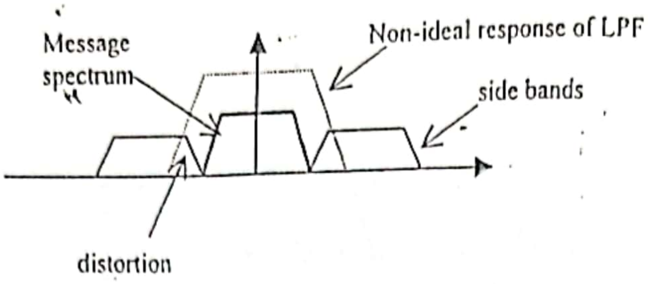
$$H_{eq}(f) = 1/P(f)$$

But if τ << Ts , then P(f) is more or less constant over message frequency band and the aperture effect can be neglected.

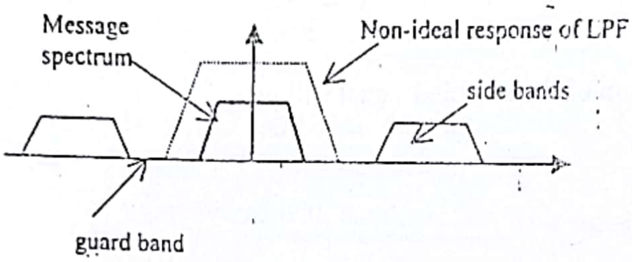
τ << Ts, P(f) is more or less con[illegible] over message-freq. band & aperture effect can be neglected

Scanned with CamScanner

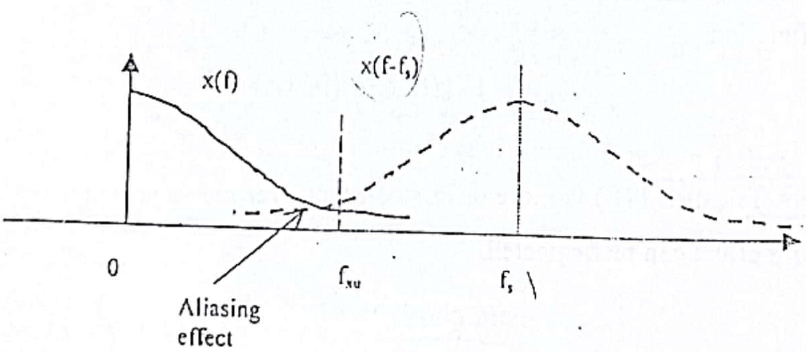
2. Effect of non – ideal reconstruction filter is that the portion of side bands will also be filtered out along with message signal.



Good filter design may eliminate this problem. One of the other way to minimize this effect is to introduce guard band by selecting f_s slightly higher than $2f_x$.



3. Real signals encountered in real life are usually time limited but not band limited. Example may be a pulse of finite duration, whose spectrum is theoretically unlimited. Thus there will be always aliasing effect.



Scanned with CamScanner

This effect is more serious than non- ideal filter effect as in that case the spurious components lies outside the message band spectrum. Where as in this case the spurious components fall within the message band spectrum.

To overcome this effect, a filter called pre-aiias filter is used to attenuate frequency components of the message spectrum higher than the band of interest and by selecting f_s moderately higher than the Nyquist rate.

Scanned with CamScanner

The pulse modulation system comprises of pulse amplitude modulation (PAM), pulse duration modulation (PDM), pulse width modulation (PWM), pulse position modulation (PPM) and pulse code modulation (PCM). Among these, most frequently used is PAM for base band transmission of digital data and PCM for conversion of analog signal into digitally encoded signal. The essential features of PAM signal will be dealt in preceding chapter. In this chapter, detailed analysis of PCM system will be carried out. Readers are advised to consult other literature for PDM,PWM and PPM systems.

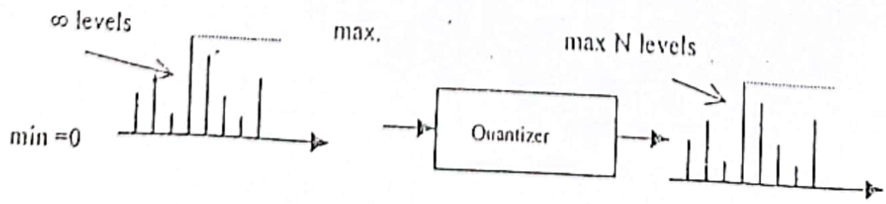
2.1 Pulse Code Modulation

Pulse Code Modulation is a technique by which analog signals are converted into digitally encoded signals.

The three basic and essential operations in a PCM system are Sampling, Quantization and encoding. These three operations are sometimes also called analog to digital (A/D) conversion. The reverse process of recovering message signal from PCM is called D/A conversion.

Quantization:

Quantization is the process of representing the analog sampled values by a finite set of levels. The sampling process converts a continuous time signal to a discrete time signal (with amplitude that can take any values from 0 to maximum level) and the quantization process converts continuous amplitude samples to a finite set (discrete) amplitude values.



Scanned with CamScanner

Uniform Quantization

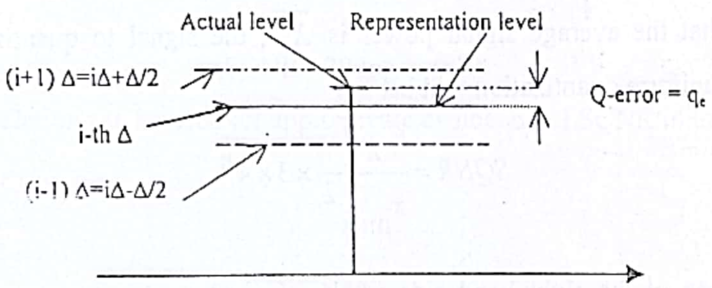
In uniform quantization it is assumed that the range of input sample is $[-x_{max}, x_{max}]$ and the number of quantization level (Q-level) $N=2^n$, where 'n' is the number of bits per source sample. Then the step size (Δ)or the length of the Q- level

$$\Delta=2x_{max}/N = x_{max}/2^{n-1}$$

The step size Δ is also called quantum.

The quantization error (Q-error) is the difference between the input signal level and the level of the quantized version. It is evident that the maximum Q-error could be only $\Delta/2$. In uniform quantization, the step size Δ is constant for entire dynamic range of input discrete signal level.

Q -error and hence the Q -noise is produced during the process of quantization because of rounding -off of the sampled values of a continuous message signal to the nearest representation level.



It is evident that q_e lies between $-\Delta/2$, $\Delta/2$ in random manner with zero mean and the mean square value determined by the step size Δ.

The average power of Q noise is therefore equals to:

Scanned with CamScanner

$$P_q = \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} q_e^2 \, dq_e = \frac{\Delta^2}{12}$$

It is seen from the above expressions that q-noise P_q in uniform quantization is dependent only upon the step size. Reducing step size (or increasing number of representation level) can reduce Pq for given range of input signal levels.

Signal to quantization noise ratio (SQNR)

The average quantization noise power can be expressed interms of signal level and number of bits per representation level as:

$$P_q = \frac{\Delta^2}{12} = \frac{x_{max}^2}{3 \times 4^n}$$

Assuming that the average signal power is $\overline{X^2}$, the signal to quantization noise (SQNR) for uniform quantization will be:

$$SQNR = \frac{\overline{X^2}}{x_{max}^2} \times 3 \times 4^n$$

The first ratio of the right hand side of the above equation can be replaced by normalized signal power:

$$\overline{\hat{x}^2} = \frac{\overline{X^2}}{x_{max}^2}$$

Then the SQNR will be:

$$SQNR = 3 \times 4^n \times \overline{\hat{x}^2}$$

Since x_{max} is the maximum level of input signal, the normalized signal power is always less or equal to unity, therefore the upper limit of SQNR will be equal to:

$$SQNR = 3 \times 4^n = 3 \times N^2$$

Scanned with CamScanner

In terms of dB , the SQNR will be:

$$SQNR_{(dB)} = P_{x\,(dB)} + 10\log 3 + 10 \times n \times \log 4$$
$$\approx P_{x(dB)} + 4.8_{(dB)} + 6 \times n_{(dB)}$$

It means that for each extra bit (n) used for representing each quantization level (coding), the SQNR increases by 6dB.

In terms of level of quantization N, this ratio can be expressed as

$$SQNR \approx P_{x\,(dB)} + 10\log(3N^2) = P_{x\,(dB)} + 10\log 3 + 20\log N$$

$$= P_{x\,(dB)} + 4.8 + 20\log N$$

And for N>>1,

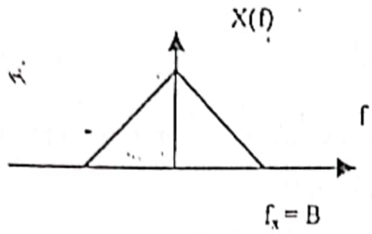
$$SQNR = 20\log N\ (dB)$$

The above relation can be used for approximate evaluation of SQNR in uniform quantization.

Scanned with CamScanner

2.2 Bandwidth Consideration and Signalling Rate for PCM.

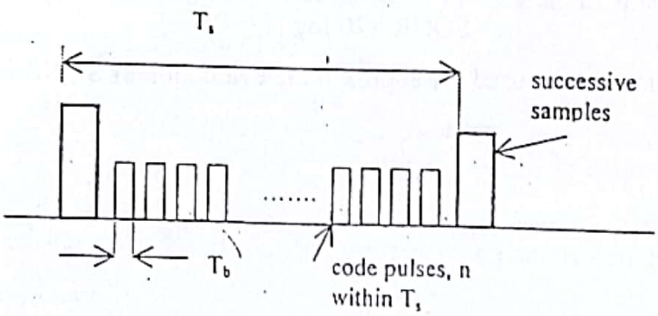
Lets consider a band limited signal with upper maximum frequency f_x and bandwidth B:



The minimum sampling frequency is then:

$$f_s \geq 2f_x = 2B$$

Each sample is represented by n- bit code word, each code consisting of a single pulse of the duration T_b. Within a time duration of successive samples $T_s = 1/f_s = 1/2B$, n pulses of T_b duration has to transmitted. The maximum possible pulse duration is then:



$$T_{b\,max} = \frac{T_s}{n} = \frac{1}{2Bn}$$

Signaling rate (R) is the number of symbols transmitted per second. For binary waveform R is expressed in bits per second.

For example if the time taken to transmit one bit is τ ($\tau = T_b$-bit duration) then the signalling rate

Scanned with CamScanner

$$R = 1/\tau = 1/T_b = n\, f_s \text{ (bits/sec)}$$

Then the minimum bandwidth required as per Nyquist criteria for zero ISI is:

$$BW_{PCM} \geq R/2 = n\, f_s / 2$$

As fs = 2B

$$BW_{PCM} \geq nB$$

In practice

$$BW_{PCM} = (1+\rho)\, nB$$

Where ρ is called Roll- off factor.

Now for a telephone voice channel with highest frequency 3400 hz, the bandwidth is standarized at 4khz. For 8 bit coding of the speech and with $\rho=1$

$$BW_{PCM} \text{ (voice)} = (1+1) \times 8 \times 4 = 64 \text{ khz.}$$

$$SQNR \approx 50 \text{ dB.}$$

Scanned with CamScanner

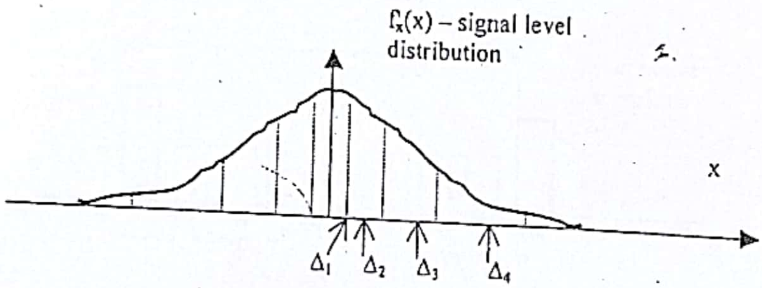
2.3 Non Uniform Quantization – Companding

The uniform quantizer produces highest or optimum SQNR provided the input signal has a uniform power density function.

The average quantization noise power q_e is fixed at $\Delta^2/12$ regardless of the value (level) of the sample being quantized. In such case, if at most of the time the signal level remains small, the apparent SQNR will be much lower than the design value. This situation is prominent in voice signal where medium voice level prevails at most of the time wheres as loud voice levels occur rarely. The ratio of peak to RMS value, called crest factor, in audio signals. is very large. For such signals uniform quantization may not yield optimum result.

For these type of signals it is advantageous to taper quantization steps (make it non-uniform), making them small near zero and larger at extremes.

A non – uniform quantizer employ variable step size quantization. In this case the average SQNR can be maintained high and constant over entire dynamic range of the input signal.



For the signal levels with highest probabilities, the step size is made smaller. As the probability decreases the step size is increased correspondingly, so that:

$$\Delta_1 < \Delta_2 < \Delta_3 < \Delta_4$$

Direct realization of variable step size quantization is very complicated. The indirect method of non-uniform quantization involves compressing of base band signal and

Scanned with CamScanner

then applying compressed signal to uniform quantizer. At the receiving end the decoded message is expanded in the reverse manner to recover original base band signal. The process of compressing & expanding is calling companding.

There are two compression laws, called μ and A laws, that are commonly used in practice.

In μ law the output signal 'y' is related to the normalized input signal level in the following manner:

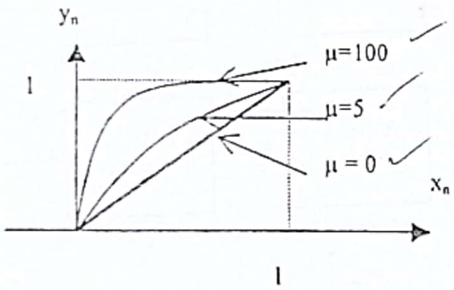
$$|y| = \frac{\log\left(1+\mu\left|\frac{x}{x_{max}}\right|\right)}{\log(1+\mu)}$$

If x and y are normalized to (+1,-1), then

$$|y_n| = \frac{\log(1+\mu|x_n|)}{\log(1+\mu)}$$

or in general,

$$y_n = \frac{\log(1+\mu|x_n|)}{\log(1+\mu)}\operatorname{sgn}(x_n)$$

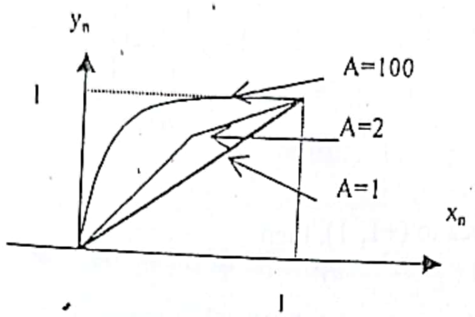


μ = 0 is the case of uniform quantization with no compressing effect. Practical values of μ is within 100 to 300. Standard PCM employ μ = 255 compander with 7 bits (128 levels) quantization yielding system improvement (SQNR) of about 24 dB.

The A law compander utilizes the following relationship between input and output signal levels:

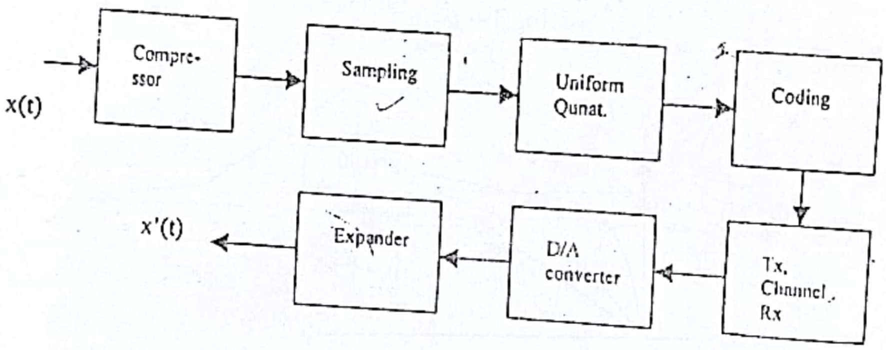
Scanned with CamScanner

$$|y_n| = \begin{cases} \dfrac{A|x_n|}{1+\log A}, & for \quad 0 \le |x_n| \le \dfrac{1}{A} \\ \dfrac{1+\log(A|x_n|)}{1+\log A}, & for \quad \dfrac{1}{A} \le |x_n| \le 1 \end{cases}$$



For practical purpose the value of A is chosen around 100 and the improvement is about 25 dB.

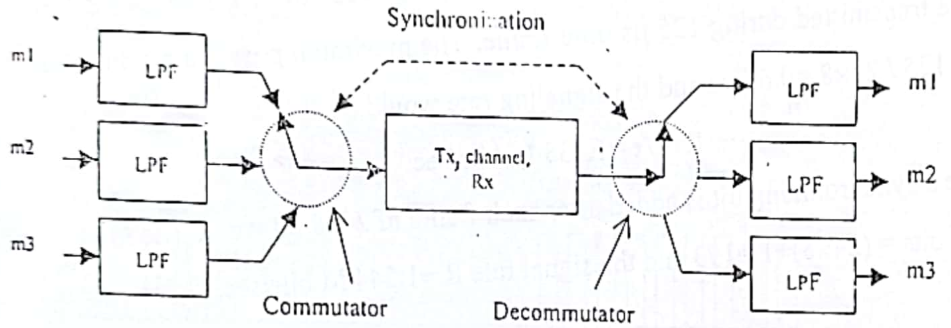
The overall functional block diagram of a complete P.CM system is shown below.



Scanned with CamScanner

2.4 Time Division Multiplexing

Sampling theorem states the possibility of the transmission of samples of message signal at some fixed interval of time. Thus over the same communication channel, samples of number of message signals can be transmitted serially and then recovered and separated at the receiving end. In other words time interval between two adjacent samples of one message can be used to transmit samples of other message and such technique is known as Time Division Multiplexing (TDM).



The signals to be multiplexed are first individually bandlimited by low pass filters. The commutator takes samples of each signal sequentially at fixed interval of time. These samples are then transmitted through the common channel using digital transmission technique. The decommutator at the receiving end separates the sequentially transmitted signal into individual signals. The commutator and decommutator are synchronized using timing signal. The low pass filters at the receiving end convert the samples of the message signal into original continuous signal.

Signaling rate and bandwidth consideration in TDM

Let m1 be the voice message signal band limited to 3400 Hz. Standard sampling frequency is taken to be 8 kHz. The time duration between successive samples is therefore 1/8 kHz = 125 μs. Let each sample be converted into 8 bit code word. It is

Scanned with CamScanner

evident that within a span of 125 μs eight pulses are required to be transmitted. The maximum duration of each bit would be $\tau = 125/8 = 15.6\ \mu s$.

The minimum signaling rate now becomes -

$$R = 1/\tau = 1/15.6\ \mu s = 0.064\ \text{Mbits/sec} = 64\ \text{kbits/sec (kbps)}.$$

The bandwidth of PCM in this case will be –

$$BW_{PCM} = (1+\rho)\ n\ B \approx 64\ \text{kHz}$$

One telephone channel thus requires signaling rate of 64 kbits/sec and channel BW of 64 khz. Now suppose 24 voice channels are to be TDM'ed. In this case 24x 8 bits need to be transmitted during 125 μs time frame. The maximum possible bit duration would be $125 / 24\times8 = 0.65\ \mu s$ and the signaling rate would be

$$R = 1/\tau = 1.538\ \text{M bits/sec}$$

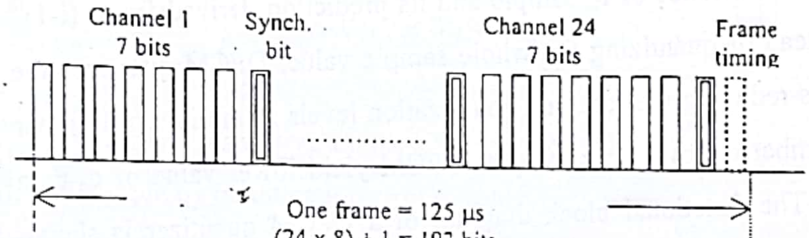
In practice a synchronizing bit is added after each frame of 24x8 bit word. The total number of bits = $(24\times8)+1 = 193$ and the signal rate $R \approx 1.544$ M bits/sec.

Scanned with CamScanner

The T1 TDM-PCM Telephone System

The breakthrough in TDM-PCM Telephone system is the 24 voice channel TDM multiplexer known as T1 system, designed by AT & T (American Telephone & Telegraph Co. Bell Telephone Systems pioneered the original concept of the TDM-PCM multiplexing.

In T1 system 24 voice channels are sampled at f_s = 8khz (T_s= 125 μs). Each sample is quantized and converted into 7 bit PCM code word. The eighth additional bit is added for synchronization purpose. Thus total number of bits per voice channel is 8.



The T1 Frame

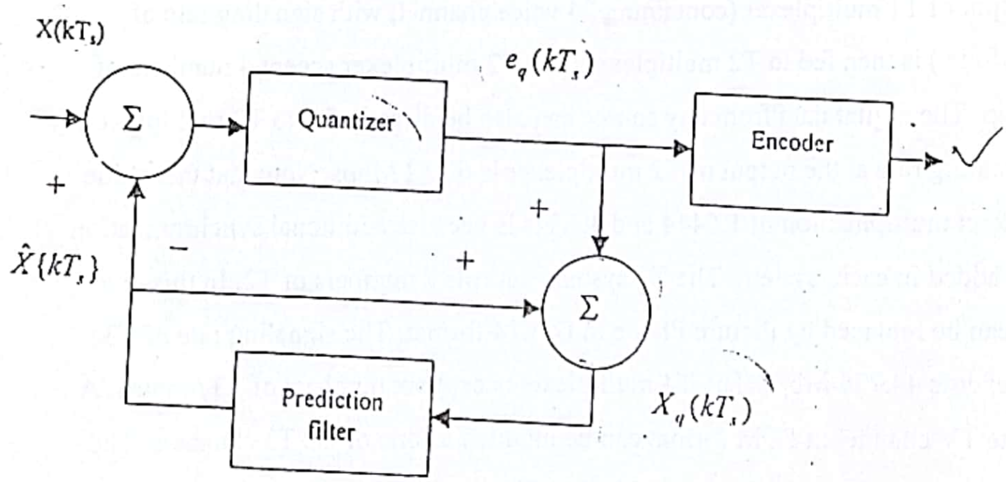
The output of T1 multiplexer (containing 24 voice channel, with signaling rate of 1.544 Mbits) is then fed to T2 multiplexer. The T2 multiplexer accept 4 numbers of T1 group. The digital data from any source can also be directly fed to T2 multiplexer. The signaling rate at the output of T2 multiplexer is 6.312 Mbps. Note that this value is not exact multiplication of 1.5444 and 4. This is because additional synchronization bits are added in each system. The T3 system accepts 7 numbers of T2. In this level the T2 can be replaced by Picture Phone in DPCM format. The signaling rate of T3 multiplexer is 44.736 Mbps. The T4 multiplexer accept six numbers of T3 outputs. A complete TV channel in PCM format can be inputted as one of the T3 channels. The data rate at the output of T4 is 274.176 Mbps. Finally the T5 level multiplexer accept two T4 level signals and the data rate at the output of T5 system is 560.160 Mbps.

Scanned with CamScanner

2.6 Differential Pulse Code Modulation

In ordinary PCM, after sampling the message signal, each sample is quantized in independent manner. It means, previous sample values have no effect on the quantization of new samples. But in practice the band limited random signals, when sampled at Nyquist rate or higher rate, produce highly correlated sample values. This is true in most cases except when the spectrum of message signal is flat within the bandwidth of interest (Flat spectrum represents highly un-correlated samples). For example the samples of speech signal does not change abruptly from sample to sample. It means there is redundancy in samples that could be removed for efficient quantization & coding.

The Differential PCM (DPCM) takes into account of these facts and the quantization is performed into the difference of i^{th} sample and its prediction derived from $(i-1)^{th}$ sample. That is instead of quantizing the whole sample value, DPCM quantizes the difference only, thus reducing the required quantization levels to minimum. In other words for given number of levels per sample, DPCM yield lower value of q_e than direct quantization. The functional block diagram of a DPCM quantizer is shown below.



Let $X\{kT_s\}$ be the sampled sequence of a message signal $x(t)$ sampled at T_s intervals. Assume that $\hat{X}(kT_s)$ be the prediction of $X\{kT_s\}$ derived from the previous $X\{(k-1)T_s\}$ sample. Then the input to the quantizer would be :

$$e\{kT_s\} = X\{kT_s\} - \hat{X}\{kT_s\}.$$

The input to quantizer $e\{kT_s\}$ is called prediction error. The quantizer output in general can be represented as :

$$e_q\{kT_s\} = e\{kT_s\} + q_e\{kT_s\}$$

Where $q_e\{kT_s\}$ is the quantization error.

The input to the Prediction Filter (PF) is the sum of $e_q\{kT_s\}$ and $\hat{X}\{kT_s\}$

$$\boxed{X_q\{kT_s\} = e_q\{kT_s\} + \hat{X}\{kT_s\}}$$

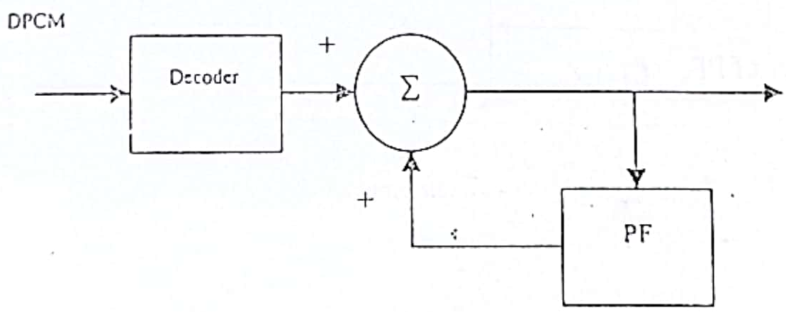
$$X_q\{kT_s\} = \hat{X}\{kT_s\} + e\{kT_s\} + q_e\{kT_s\} =$$
$$= X\{kT_s\} + q_e\{kT_s\}$$

In other words irrespective of the properties of PF the input to it differs from the original sample by quantization error q_e only.

Now if the prediction is good then $e_q\{kT_s\}$ becomes smaller. It means, we can now quantize smaller change in e_q than quantizing greater changes in X.

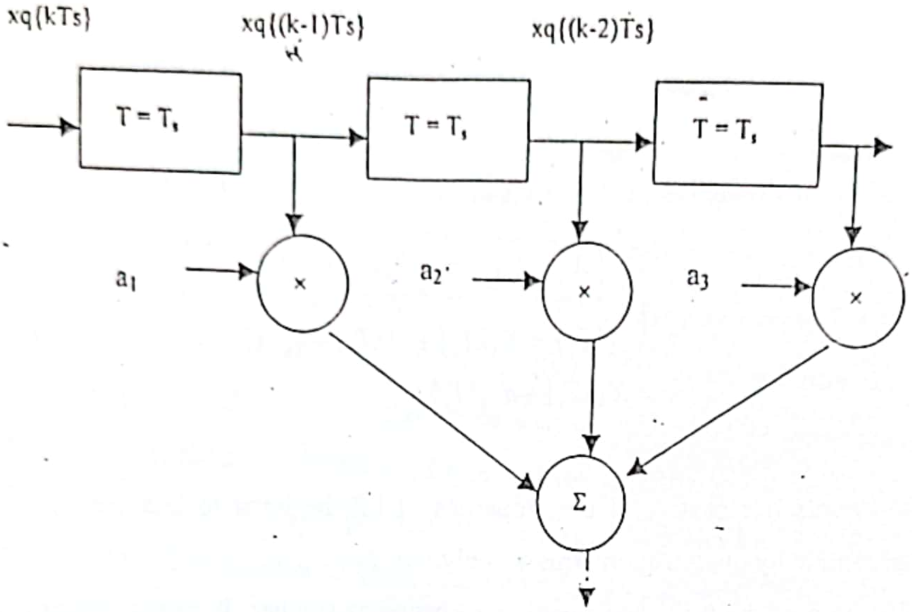
The straightforward advantage of the DPCM is that an eight bit encoded signal (256 levels) may be transmitted using 4 bits only and in both cases quantization noise power remains same.

The receiver of DPCM utilizes the same prediction filter as used in transmitter.



The output of the decoder is the quantized version of the prediction error. This value is added to the predicted version of i-th quantized level derived from the prediction filter.

The prediction filter in general is of tapped delay line filters type, where the predicted value $X\{kT_s\}$ is modeled as linear combination of past values of quantized inputs.



The output of the above PF is the linear sum of n previous values of samples scaled by some coefficients.

$$\hat{X}\{nT_s\} = \sum_{k=1}^{P} a_k Xq\{nT_s - kT_s\}$$

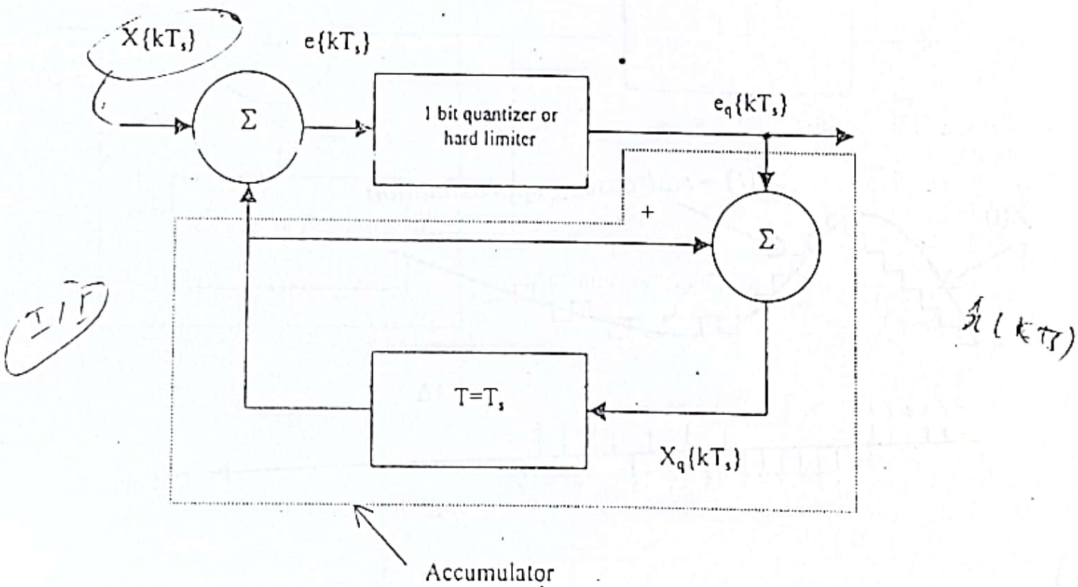
Where P is the order of PF.

Scanned with CamScanner

2.7 Delta Modulation

Delta Modulation (DM) is the simplified version of DPCM. In DM the difference between original sample and its approximation .(derived from immediate past sample) is quantized in one of two possible levels $+\Delta$ or $-\Delta$ and each level is converted into 1bit code word (e.g. $+\Delta \rightarrow 1$, $-\Delta \rightarrow 0$). Thus the DM uses only one bit to represent each sampled level.

Staircase approximation with step size Δ is used in DM. Depending upon the sign of the difference between $X\{kT_s\}$ and its prediction $\hat{X}\{kT_s\}$, the approximation is increased by $+\Delta$ or decreased by $-\Delta$ to follow the signal. This technique is feasible (i.e. signal to quantization noise ratio becomes acceptable) only when sample to sample correlation is very high (i.e. $e_q\{kT_s\}$ is very low). To make the samples highly redundant (high correlation), the sampling rate is chosen much higher than the Nyquist rate. But since the coding is 1 bit coding, the bandwidth requirement as well as the signaling rate for given signal with DM becomes considerably low.



Let $X(kT_s)$ be the original sample and $\hat{X}(kT_s)$ be its prediction. Then the quantizer input

$$e(kT_s) = X(kT_s) - X_q(kTs\text{-}T_s),$$

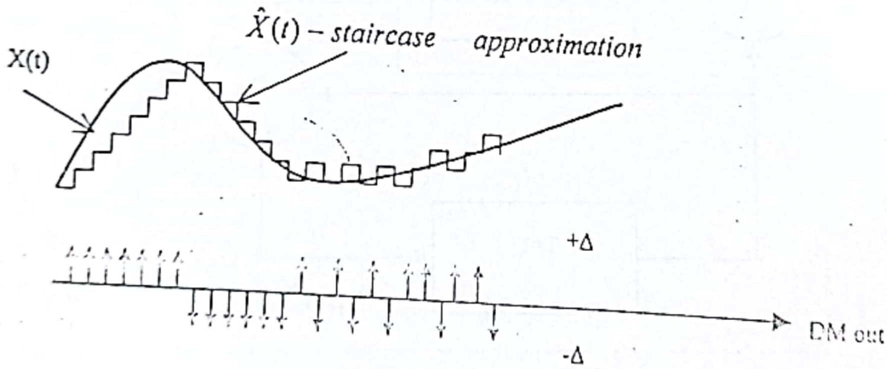
Where

$$X_q(kTs\text{-}T_s) = X_q\ \{(k\text{-}1)Ts\}$$

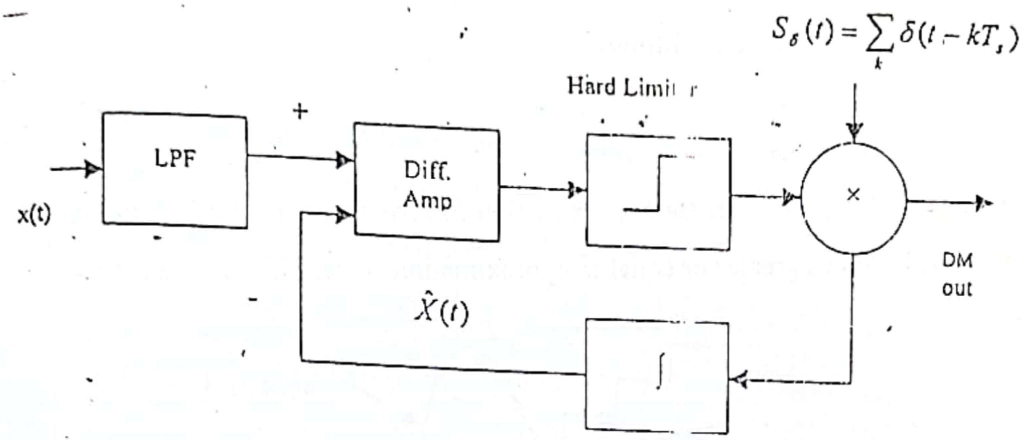
is quantized version of previous sample. The output of the one bit quantizer (also called hard limiter) will be:

$$e_q(kT_s) = \Delta\ \mathrm{Sgn}\ [e(kT_s)],$$

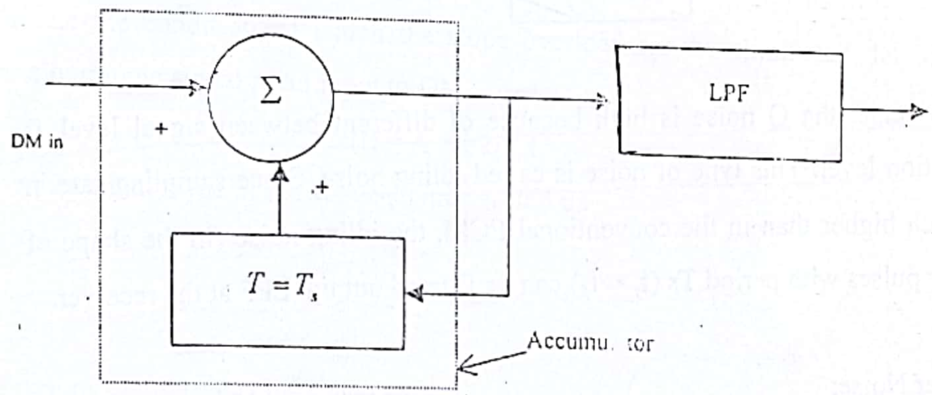
where Δ is the step of quantization and 'Sgn' function represents the sign of $e(kT_s)$. Technically speaking the DM modulator is nothing but the staircase approximator of the input wave form, in which each stair is represented by 1 or 0 depending upon the rise or the fall of the steps of the stair.



Practically the DM modulator is implemented as follows. The accumulator is replaced by a RC integrator.



The receiver of DM consists of an accumulator and LPF.



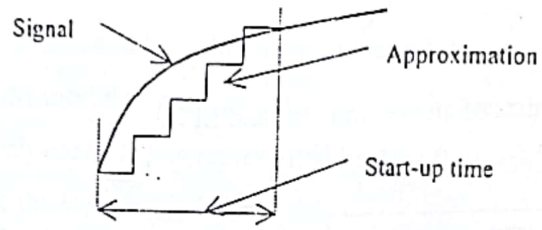
Scanned with CamScanner

2.8 Quantization Noise in DM

The different conditions in following the signal by staircase approximation, that produce noise, can be described as follows

a) Start-up

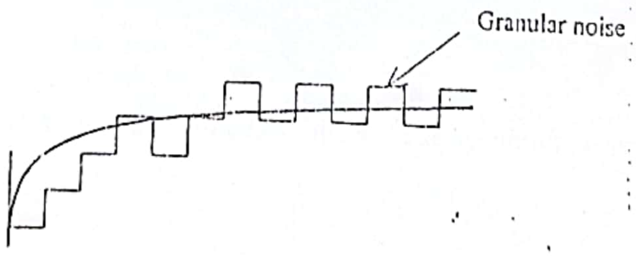
At initial stage of DM, the staircase approximation follows the signal with step-rise by Δ unless x (t) becomes greater or equal to approximation stage. This is called start-up condition.



In start-up stage, the Q noise is high because of different between signal level & approximation level. This type of noise is called idling noise. Since sampling rate in DM is much higher than in the conventional PCM, the idling noise (in the shape of rectangular pulses with period Ts (f_s>>f_x) can be filtered out by LPF at the receiver.

b) Granular Noise:

When the local slope of the signal is low (i.e. signal is almost constant with respect to t) and the Δ is relatively high, the approximation starts swinging from -Δ to Δ causing high noise levels.

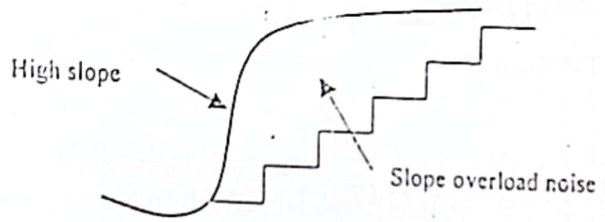


Scanned with CamScanner

It is evident that granular noise can be minimized by reducing the step size Δ.

c) Slop over load noise

If the local signal slope is high then the given step size may not be sufficient to follow rate of change of the signal. In this case a condition called slope overload is present and the noise produced by such overload is called slope overload noise.



To evaluate the condition, at which the slope overload can be eliminated, let us consider a harmonic signal as an input to DM

$$x(t) = A\cos 2\pi f_x t$$

Then the maximum slope of the above signal would be

$$\left[\frac{dx(t)}{dt}\right]_{max} = A\,2\pi f_x t$$

The maximum sample to sample change in value of $x(t)$ would be $A\,2\pi f_x T_s'$, where T'_s is the sampling period determined by the sampling frequency higher than Nyquist frequency

If the step size is ……… then the max. slope

$$A\,2\pi f_x T_s' < \Delta$$

then the slope overload noise can be minimized. For this to happen the amplitude of the signal should satisfy the following condition:

Scanned with CamScanner

reduced

$$A_{max} = \Delta f_s' / 2\pi f_x$$

Slope overload can be alleviated by filtering the signal to limit its maximum rate of change or by increasing the step size and/or the sampling rate. These measures will result in poor signal resolution and higher bandwidth requirements.

A technique called adaptive DM, where by the step size is made variable (i.e. made larger) upon detection of slope overload, is more effective.

Imp

Signal to Quantization Noise Ratio in DM

Since in DM, the output of the quantizer is

$$e_q(kT_s) = \Delta \operatorname{sgn}\{ e(kT_s)\}$$

and that the value of e_q (kT_s) may be $+\Delta$ or $-\Delta$, the total swing is 2Δ. The total noise power is therefore

$$\boxed{Pq = \frac{1}{2\Delta}\int_{-\Delta}^{\Delta} e^2 q\, de_q = \Delta^2/3}$$

Experiments have indicated that in case of DM normalized power of $e_q(t)$ is uniformly distributed over f interval of $(0, f_s')$. Then the psdf of e_q (t) is -

$$G_{eq}(f) = \{\frac{P_q}{2 f_s'} \ for \ |f| \leq fs'$$

or,

$$G_{eq}(f) = \frac{\Delta^2}{3 \times 2 fs'}$$

Scanned with CamScanner

If the output receiver filter is ideal over the frequency range 0, f_x, then the normalized $[h_{LPF}(f)=1]$ average power at the output of filter will be-

$$P_{q(LPF)} = \int_{-fx}^{fx} G_{eq}(f)df$$

$$= \frac{\Delta^2}{3} \times \frac{2 f_x}{2 f_s'} = \frac{\Delta^2 f_x}{3 f_s'}$$

Regarding the signal power let us consider the case when all the signal power is concentrated at the upper end of the bandwidth of message signal, i.e.,

$$x(t) = A \cos 2\pi f_x t$$

At the output of ideal LPF

$$x_{LPF}(t) = A \cos 2\pi f_x t$$

The power of this signal is

$$P_x = A^2/2$$

To avoid slope overload the maximum value of A should be less or equal to

$$\Delta f_s'/2\pi f_x$$

In this case the signal power (for minimum slope overload) will be equal to:

$$P_{xo} = \frac{\Delta^2}{8\pi^2}\left(\frac{fs'}{fx}\right)$$

Then the SQNR offered by DM will be:

Scanned with CamScanner

$$SQNR(DM) = \frac{\Delta^2}{8\pi^2}\left(\frac{fs'}{fx}\right)^2 \times \frac{3 fs'}{\Delta^2 fx}$$

$$= \frac{3}{8\pi^2}\left(\frac{fs'}{fx}\right)^3$$

Comparison of PCM & DM :

Comparison of any two systems are basically made in terms of SQNR in similar conditions (i.e. for given f_x and R) and the equipment complexity.

SQNR :

For n bit PCM:

$$f_s = 2f_x$$

and the bit rate

$$R_{PCM} = nf_s$$

For DM the bit rate

$$R_{DM} = fs' \text{ (as n=1)}$$

The sampling frequency in DM chosen n times higher than in PCM for identical bandwidth requirements for DM & PCM-

$$f_s' = 2nf_x = nf_s$$

The SQNR expression for PCM is:

$$SQNR_{PCM} \approx 6n \text{ dB for } n >> 1$$

And the SQNR expression for DM is:

Scanned with CamScanner

$$SQNR_{DM} = \frac{3}{8\pi^2}\left(\frac{fs'}{fx}\right)^3 = \frac{3}{8\pi^2}\left(\frac{nfs}{fx}\right)^3 = \left(\frac{2nfx}{fx}\right)^3 \times \frac{3}{8\pi^2} =$$

$$= \frac{3}{\pi^2} \times n^3 \cong 0.3n^3 = 10\log 0.3 + 30\log n$$

Now assuming n = 8, we get

$$(SQNR)_{PCM} \approx 48dB$$

$$(SQNR)_{DM} \approx 22dB$$

It shows that under similar condition the performance of PCM is superior to DM. With ordinary DM, the comparable quality comparable to PCM can be achieved only when the signal rate is around 100 kbits/s.

Superiority of PCM over DM is highlighted only when the number of bits per sample is greater than 5. At n<5, PCM & DM yield almost same performance level.

Equipment complexity:

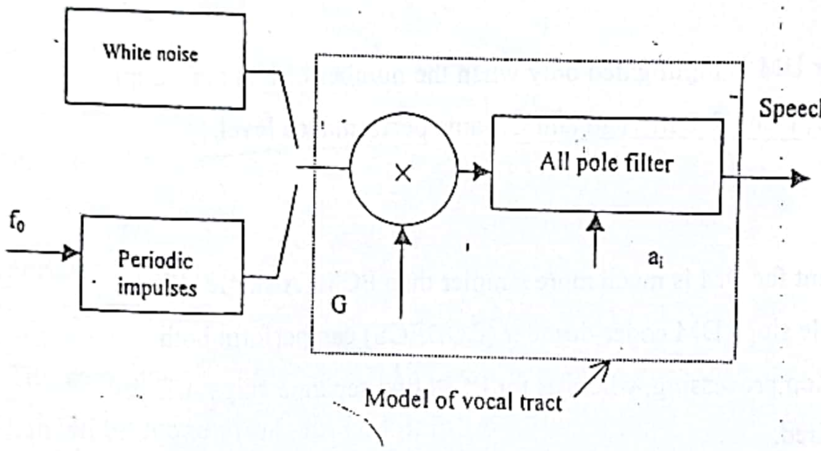
The hardware requirement for DM is much more simpler than PCM. A single chip called continuous variable slope DM coder-decoder (CODECS) can perform both transmission and reception processing, where as for PCM two separate chips A/D & D/A converters are required.

Scanned with CamScanner

The transmission of speech signals (as well as the video signals) at relatively lower bit rate with acceptable quality is possible by the Analysis-Synthesis method.

The speech signal first is modeled (or analyzed) and its parameter (parameters of the model) is extracted. These parameters are then transmitted using PCM and at receiver end the original voice signal is predicted or synthesized using the parameters of model. Such technique of coding of speech signal is called Linear Prediction Coding and theory associated with it called Linear PredictionTheory.

A speech signal can be modeled as a system consisting of short pulses of repetition frequency f_0 ($1/f_0$ represents the pitch of the speech), an all-pole filter with coefficients a_i and the gain parameters G.



Since the speech signals are known to be stationary (i.e. the mean value does not change with shift in time) for a short period of 20-30 mS. It means that within this time frame the filter coefficients can be assumed to be constant. We can therefore analyze (estimate) the values of f_0, a_i and G for these period and transmit these parameter using PCM. Then at the receiving end, using same Prediction filter, we reconstruct the original speech signal.

Example:

Suppose speech signal is filtered at 3 khz and sampled at 8000 samples/sec. A block of 160 samples corresponding to 20ms interval is considered to be a frame for

Scanned with CamScanner

prediction. The average filter coefficient ai, G and f_0 for this interval is estimated and converted into PCM for transmission. The bit wise representation of each parameters are as follows

Voice-unvoiced information	- 1 bit
Pitch f_0	-6 bits
Gain parameter G	-5 bits (log companding)
Filter coefficients	-8-10 bits/coefficient.

Now if 1st order filter is used, the total number of bits to represent a frame of 20 mS would be 22 bits. Since Frame rate is 160 frames/sec., the bit rate would be 3520 bits/sec. This bit rate can further be compressed to as low as 2400 bits/sec with tolerable voice quality. Where as for similar quality the PCM requires 64000 bits/sec. Application : Cellular (mobile) Telephone systems uses LPC at 4800 bits/sec signal rate for good quality speech.

Scanned with CamScanner

The two fundamental questions related to the communication theory are:

a) What is the rate at which the given source is emitting "Information"?
b) What is the maximum rate of information transmission over a noisy channel ?

The answer to the first question is the Information theory, which deals with the definition of the information, measure of information and the modeling of information. The channel capacity theorem is the answer to the second question.

Imp

3.1 Information Theory

Information content of a Message:

A message is a sequence of symbols intended to reduce uncertainty of the receiver. If the sequence of symbols does not change the uncertainty level of the receiver then the message does not contain any information.

Example:

You are planning to go to Biratnagar during summer vacation. You called your friend in Biratnagar to know the weather conditions of Biratnagar. Assume that you could have received the following messages regarding the weather condition in Biratnagar in summer:

a) It is sunny and hot;
b) It is cold; or
c) Snow is falling down.

Among these three messages, the content of information is very different. From the past experience about the weather conditions of Biratnagar you know that in summer in Biratnagar is sunny & hot and therefore the first message would contain almost no information for you.

Scanned with CamScanner

The information content in message (b) is may relatively high because you do not expect, in general, cold weather during summer in Biratnagar, but may happen rarely.

Message (c) contains highest level of information (of course among the three possible messages you could have received) as snow during summer and in Biratnagar is almost next impossible (but still could happen).

Therefore the information content of any message is closely related to the past knowledge of the occurrence of event and the level of uncertainty it contains with respect to the recipient of the message. The same message, if given to a person living in Switzerland and who never have heard of Biratnagar and have no plan to go there, would have contained almost zero information.

Thus in general "On an intuitive basis, the amount of information received from the knowledge of occurrence of an event is related to the probability or the likelihood of occurrence of the event".

In other words the message related to an event least likely to occur (with respect to the knowledge of occurring such events in the past) contains more information.

Let $m_1, \ldots, m_q$ be the one of the q possible messages emitted by a source with probabilities of occurrence $p_1, p_2, \ldots, p_q$ such that

$$p_1+p_2+\ldots+p_q=1.$$

Let $I(m_k)$ be the amount of information contained in k-th message.

Then for $I(m_k)$ to represent information content of m_k message, based on intuitions, the following conditions should be met

a) $I(m_k)>I(m_j)$, if $p_k<p_j$

It means the information content of an event with less probability is higher than that of with high probability of occurrence.

Scanned with CamScanner

b) $I(m_k) \rightarrow 0$ if $p_k \rightarrow 1$

The information content in an event with the probability of occurrence near to unity, is near to zero. The example is the message to a person living in the earth that " Sun rises from the east".

c) $I(m_k) \rightarrow 1$ if $p_k \rightarrow 0$

The example is the information content in the message about an event which almost impossible – "Sun rises from the west"

d) $I(m_k) \geq 0$ when $0 \leq p_k \leq 1$

It means any message should contain some information i.e. I (m_k) should be non-negative.

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

It means, if m_k and m_j are two independent messages coming from the same source, then the total information received from two message is equal to the sum of information contents in each message. For example, in weather forecast from Radio Nepal you hear "It will be rainy today and sunny tomorrow". There are two messages in it and if weather of today does not affect the weather of tomorrow, then the total information in this message is equal to sum of the two individual informations.

To satisfy all the conditions mentioned above, we could relate $I(m_k)$ and p_k in the following manner.

$$I(m_k) = \log(1/p_k) = -\log (p_k)$$

The unit of $I(m_k)$ depends upon the base assigned to log:

a) base is 'e' , i.e. logarithm is natural (ln) -- The unit of information is 'nat'

b) base is 10 – The unit is Hartley or decit

c) base is 2- The unit is bit

If two binary digits 1 and 0 occur with equal probability and are correctly detected at the receiving end, then the information content in each digit is 1 bit

$$I\,(0 \text{ or } 1) = -\log_2(1/2) = 1 \text{ bit.}$$

Scanned with CamScanner

The Entropy or Average Information Content :

The average information content of a sequence of symbols is called entropy.

Let a source emit one of M possible symbols (long independent sequence of symbols) $S_1, S_2, \ldots\ldots\ldots S_M$ in statistically independent sequence (i.e. probability of occurrence of S_i does not depend upon previous occurrence of S_k or future occur of S_j) with probabilities $p_1, p_2, \ldots\ldots\ldots p_M$.

Now in a long message sequence containing N symbols, the rate of occurrence of S_i symbol is p_iN and the information content contributed by S_i symbol in that sequence is

$$I(s_i)=p_iN\log_2(1/p_i)$$

Then the total information Content of N sequence of M symbols is log24

$$I_{Total} = \sum_{i=1}^{m} Np_i \log_2\left(\frac{1}{p_i}\right)$$

The average information content is

$$I_{average} = H = \frac{I_{total}}{N} = \sum_{i=1}^{M} p_i \log\left(\frac{1}{p_i}\right)$$

Where H is called source entropy.

If the symbol rate (i.e. message rate) is R_s, then the information rate is

$$R_{inf} = R_s \times H \quad bits/sec$$

Scanned with CamScanner

Shannon's Channel Capacity Theorem

For a channel with additive Gaussian white noise, the relationship between channel capacity C, channel bandwidth B and the received signal to noise ratio (SNR) is :

$$C = B \log_2(1 + SNR) \text{ bits/sec}$$

This relationship is called Shannon-Hartley Channel Capacity Theorem.

Implication of Theorem:

1) Indicates the upper limit of data transmission for reliable communication. A designer thus can estimate C for required SNR and B for reliable communication.

2) Trade off between B and SNR for given C

The channel capacity is limited by various extraneous factors that the designer can not play with. For example the maximum frequency that can be transmitted over a pair of cable is limited by its construction. But the other two parameters B and SNR are in hand of the designer. Bandwidth of the signal can be compressed and the SNR can be increased by increasing signal power or by introducing low noise devices. Therefore the designer can trade-off between B and SNR in optimum way for given C.

Example:

A signal with the data rate R=10,000 bits/sec is required to transmit over a channel with limited bandwidth of B=3,000 Hz. Theoretically the absolute minimum bandwidth required for transmission of the above data is 10,000/2 = 5000 Hz. Now for transmission of R bits/sec, the minimum channel capacity should also be equal to this value , i.e., C_{min} = R. In this case the required SNR will be:

$$SNR = 2^{(C/B)} - 1 = 2^{(10{,}000/3000)} - 1 \approx 9$$

means the signal power must be 9 times higher than the noise power.

But now if we consider a channel with B=10,000 Hz, then the required SNR will be 1. It means if B is higher, same quality of signal transmission can be achieved with less signal power (9 times). In other words bandwidth compression from 10,000 to 5,000 Hz is possible but at the cost of increasing signal power by 9 times.

3) Bandwidth compression:

Shannon channel capacity theorem indicate that it is possible to transmit signal with upper frequency f_{max} through a channel having bandwidth less than f_{max} :

Example:

Let a signal x(t) have upper frequency limit of f_{max}

Let us sample x(t) at $3f_{max}$ (i.e. 1.5 times greater than the Nyquist rate)

Then the data rate will be

$$R = 3n\, f_{max}$$

where n is the number of bits/sample.

Now if

$$B = f_{max} / 2$$

then for $C \geq 3n\, f_{max}$ and for n = 6, the required SNR would be:

$$SNR = 2^{(C/B)} - 1 = 2^{3 X 6\, f_{max} X 2 / f_{max}} - 1 = 2^{36} - 1 \approx 6.8 X 10^{10}$$

If we increase signal power by $7 x 10^{10}$ times in comparison to the noise power we can transmit a signal through a channel having bandwidth equal to half of the signal bandwidth. Which is although possible, but very impracticable.

N.B. Theoretical limits of Shannon's channel capacity theorem:

1. As the noise in the channel tends to zero, the value of SNR will tend to infinity. Subsequently the channel capacity C will tend to infinity. It means that the noise less channel has an infinite capacity. This type of channel is referred to as ideal channel.
2. As the bandwidth of the channel B tend to infinity, the channel capacity reaches an upper limit C_{max}. This is because Noise power is proportional to the bandwidth

Scanned with CamScanner

and as the bandwidth is increased, the noise power also increases correspondingly.

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

where S is the signal power and η is the psdf of the white noise. The above expression can be written as:

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

Replacing $\frac{S}{\eta B}$ by Y and considering the limit:

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

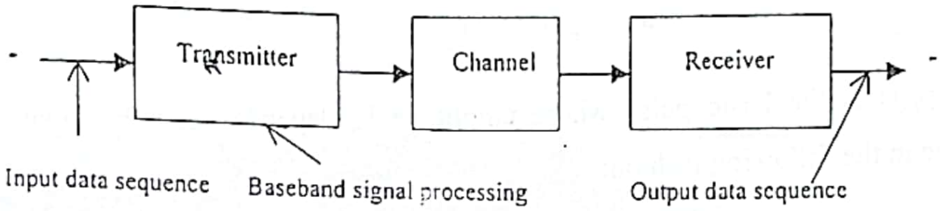
we get

$$\lim_{B\to\infty} C = C_{\max} = \frac{S}{\eta}\log_2 e = 1.44\left(\frac{S}{\eta}\right)$$

Scanned with CamScanner

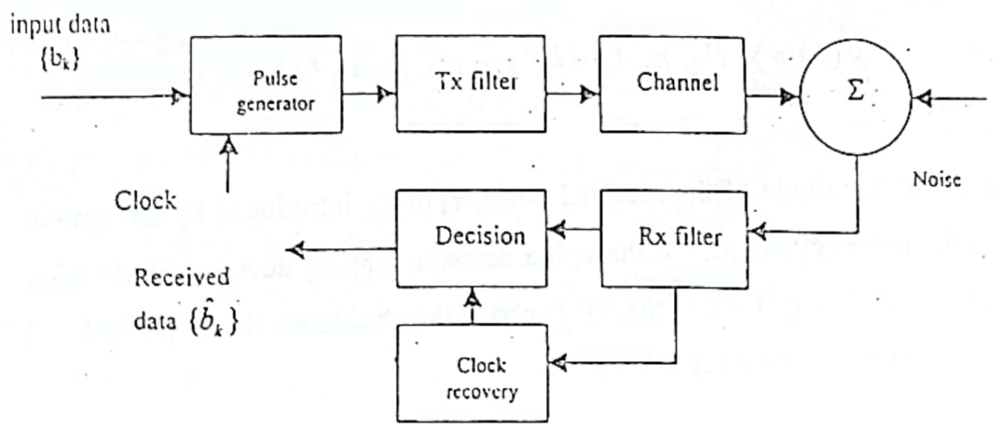
3.3 Baseband Digital Communication Systems

Baseband digital communication system refers to a system in which transmission and reception of digital signals over band limited channel is accomplished without employing carrier modulation (band pass) technique.



Before transmitting the data sequence over band limited channel, some transformation in the data wave- form has to be performed to make them suitable for transmission over the channel. This is required primarily because the input data sequence are usually in the form of voltage or current with limited amplitude values where as for transmission over the channel the amplitude level has to be increased to compensate transmission losses. These transform for baseband (BB) communication can be accomplished using pulse modulation (PM) technique. Among PM, the PAM is considered to be simplest and efficient. Therefore further discussions will be based on PAM only.

Baseband data communication system using PAM have the following functional blocks.



Scanned with CamScanner

The input to the pulse generator clocked at the bit rate is sequence of digital pulses (binary data sequence). The output of the pulse generator is PAM signal represented as:

$$x(t) = \sum_{k=-\infty}^{\infty} b_k \, p_g(t - kT_b)$$

Where $p_g(t)$ is the basic pulse whose amplitude b_k depends upon the input data sequence in the following fashion:

$$b_k = +b \text{ if k-th bit is 1}$$
$$= -b \text{ if k-th bit is 0}$$

For convenience lets consider that $p_g(t)$ to be normalized in such way that

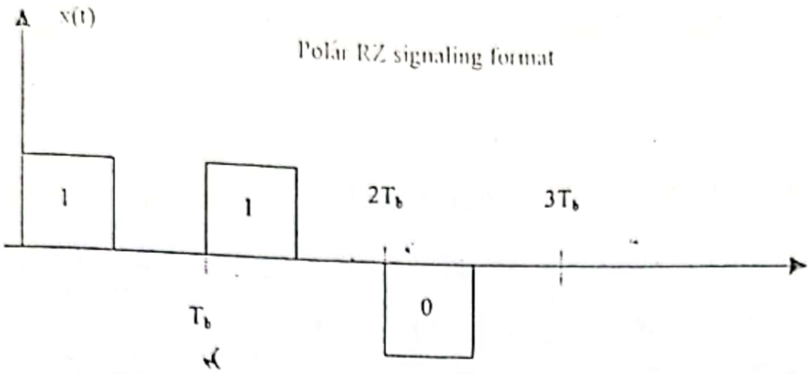
$$p_g(0) = 1$$

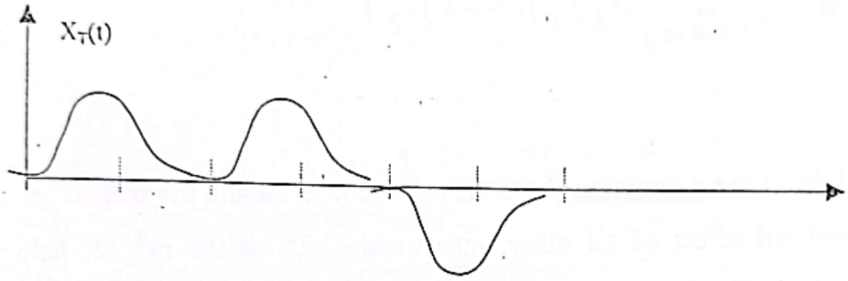
The signal x(t) from the output of the pulse generator is passed through transmitting filter, the channel and the receiving filter having impulse responses $h_T(t)$, $h_C(t)$ and $h_r(t)$ and the transfer functions $H_T(f)$, $H_C(f)$ and $H_r(f)$. Furthermore the additive noise is added to the signal in the channel. The output of R_x filter is thus

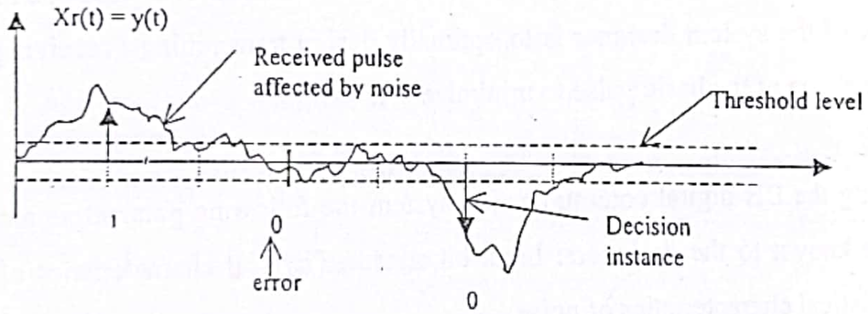
$$y(t) = \sum_k B_k \, p_r(t - kT_b - \tau_d) + n_0(t)$$

Where B_K is the amplitude of the received pulse, τ_d delay introduced by the system and $n_0(t)$ is the noise. y(t) is passed through a decision-making device with decision making instance of timing clock. If the y(t) is above threshold then the output is 1 and if below threshold then the output is 0.

Scanned with CamScanner







The amplitude of the received pulse is

$$B_k = k_c\ b_k$$

and the shape of the received pulse is

$$p_r(t) = k_c\, p_g(t)$$

Where k_c is the cumulative response of transmitting filter, channel and receiving filter to the x(t). For simplicity of further analysis we assume that $\tau_d = 0$ and the channel is noiseless i.e. $n_0(t) = 0$. Then in frequency domain the received pulse can be expressed as:

$$p_r(f) = k_c\, p_g(f)\, H_T(f)\, H_r(f)\, H_C(f)$$

When y(t) is compared with the threshold, the output of the decision making device at $t = m\,T_b$ is

$$y(t = mT_b) = B_m + \sum_{k=-\infty\ (k \neq m)}^{\infty} B_k P_r[mT_b - kT_b]$$

$$= B_m + \sum_{k\,(k \neq m)} B_k\, p_r[(m-k)T_b]$$

The first term of the above expression B_m is the m[th] decoded bit and the second term represents the residual effect of all other transmitted bits on the m[th] bit being decoded. The residual effect is called Inter symbol Interference (ISI)

ISI arises due to dispersion of pulse shape by the filters and channel. Therefore one of the major task of the system designer is to optimally design transmitting / receiving filters and the shape of the basic pulse to minimize ISI.

Design Criteria:

While designing the BB digital communication system the following parameters are assumed to be known to the designers: Input bit stream, Physical characteristics of channel & statistical characteristics of noise.

The objective of the design is to: select optimum $H_R(f)$,$H_T(f)$, $P_g(t)$, $P_r(t)$ for minimum ISI and minimum bit error rate [or bit error probability $p(\hat{b}_k \neq b_k)$]

The design criteria are:

Data rate-maximize, bandwidth requirement –optimize, error rate-minimize, transmitter power- minimize, signal to noise ratio –maximize, circuit complexity- make simple or trade-off between these parameters to come to an optimum solution in given constraints.

Scanned with CamScanner

3.4 Nyquist pulse shaping criteria or Nyquist condition for zero ISI

The basic equation for decision making in BB data communication is:

$$y(t = mT_b) = B_m + \sum_{k=-\infty\,(k \neq m)}^{\infty} B_k\, p_r[(m-k)T_b]$$

From the above equation it is evident that for zero ISI while detecting m^{th} bit, the following condition need to be satisfied:

$$p_r[(m-k)T_b] = \begin{cases} 1, & k = m \\ 0, & k \neq m \end{cases}$$

or in general ,

$$p_r(nT_b) = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases}$$

The pulse $p_r(nT_b)$ with above condition is nothing but a sinc function-

$$p(t) = Sinc(2\pi B_0 t) = \frac{\sin(2\pi B_o t)}{2\pi B_o t}$$

where $B_0 = 1/2T_b = R_b/2$ $\quad$ ($1/T_b = R_b$-bit rate)

The FT of p(t) is :

$$p(f) = \begin{cases} \dfrac{1}{2B_0}, & 0 \leq |f| \leq B_0 \\ 0, & B_0 < |f| \end{cases}$$

The ideal Sinc function is non-casual and therefore physically non-realizable. But we can approximate p(t) so that the shape of the pulse is near to Sinc function.

Scanned with CamScanner

The idea is to make p(f) to roll- off at the ends gradually than abruptly and have maximum span of flat top.

One of such approximation is called raised cosine frequency characteristics defined as:

$$P(f) = \frac{1}{2B_0} \; for \quad 0 \le |f| < f_1,$$

$$P(f) = \frac{1}{4B_0}\left(1+\cos\left[\frac{\pi(|f|-f_1)}{2B_0-2f_1}\right]\right) for \quad f_1 < |f| \le 2B_0 - f_1$$

$$P(f) = 0 \quad for \quad 2B_0 - f_1 < |f|$$

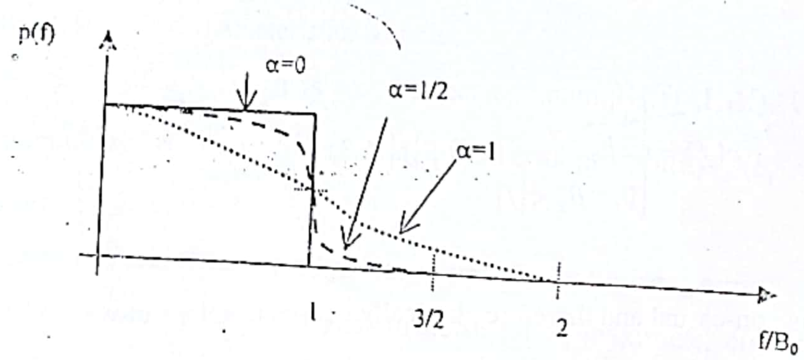
Here f_l & B_0 are related as

$$\alpha = 1 - f_l/B_0,$$

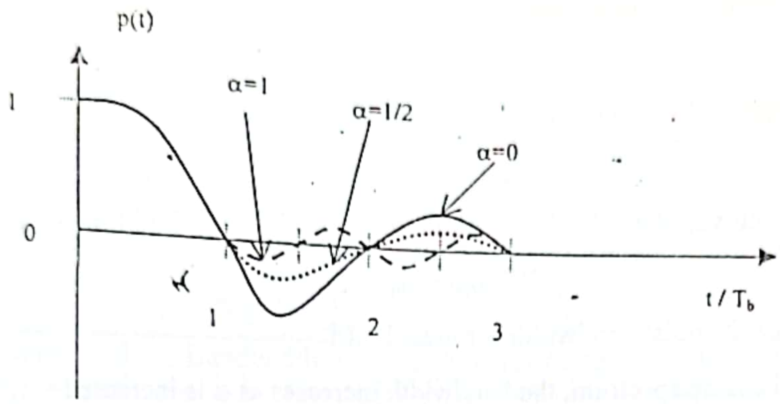
where α is the Roll off factor. For $\alpha=0$, $f_l=B_0$ is the absolute minimum bandwidth required for zero ISI.

With above p(f), the pulse shape p (t) will be:

$$p_r(t) = \frac{Sinc(4B_0 t)}{1-16B_0^2 t^2}$$



Scanned with CamScanner



Characteristics of p(f)

1. for α=0.5, p(f) cuts-off gradually and therefore physically realizable,
2. p(f) has odd symmetry with respect to cut-off frequency of ideal filter (B_0)
3. the value of p(f) at $f / B_0 = 1$ equals to ½ of its maximum value for any α.
4. p(f) is real i.e. non-negative and

$$\int_{-\infty}^{\infty} p(f)\, df = 1$$

Characteristics of p(t) :

1. at

$$t = \pm T_b / 2, p(t) = 0.5$$

i.e. pulse width measured at half-amplitude is exactly $T_b/2$.

2. for α =1, there is additional zero crossing of p(t) at

$$\pm 3 Tb / 2, \pm 5 Tb / 2$$

Scanned with CamScanner

(usual zero crossing at $\pm T_b, \pm 2T_b$ ....)

These additional zero crossing is useful for generating timing signal for synchronization from received signal.

Transmission bandwidth consideration:

As seen from p(f) curve, the absolute minimum bandwidth required for PAM is

$$BW_{PAM} = B_0,$$

where $B_0 = 1/2\,T_b$ is Nyquist bandwidth for ideal PAM

Because of raised cosine spectrum, the bandwidth increases as α is increased.

For α=1,

$$BW_{PAM} = 2B_0$$

And in general,

$$BW_{PAM} = B_0\,(1+\alpha)$$

$$= (1+\alpha)/2T_b = R_b(1+\alpha)/2$$

where R_b-bit rate.

$$R_b = 1/T_b = 2B_0 \text{ for zero ISI}$$

Conclusion:

I. For practically realizable pulse shaping filter either the bandwidth of the channel for given R_b should be double the Nyquist bandwidth for zero ISI (B_0

II. If the channel bandwidth is B_0 then for practically realizable filters the data rate should be $R_b/2$ (both cases for α=1).

Scanned with CamScanner

3.5 Correlative Coding Techniques - Duo- Binary & Modified Duo Binary Encoding.

As per Nyquist condition for zero ISI, for the ideal case (α=0), which is of course practically unrealizable, for R_b data rate the absolute minimum bandwidth is $R_b/2$ and with raised cosine pulse shaping the minimum bandwidth would be R_b (α=1) .

Data rate	Bandwidth	ISI	Condition
R_b	$R_b/2$	Zero	Ideal (unrealizable)
R_b	R_b	Zero	Raised cosine

It is practically possible to transmit R_b data rate over the channel having $R_b/2$ bandwidth by introducing controlled amount of ISI. By intentionally introducing controlled amount of ISI, to maintain given bit error probability, we need to increase signal power corresponding. Introduction of controlled amount of ISI makes possible to take care of it at the receiving end. One of the techniques to introduce controlled amount of and then by arriving at practically realizable pulse shaping filter is called duo -binary encoding or correlative coding or partial response signaling.

Duo binary encoding (DBE):

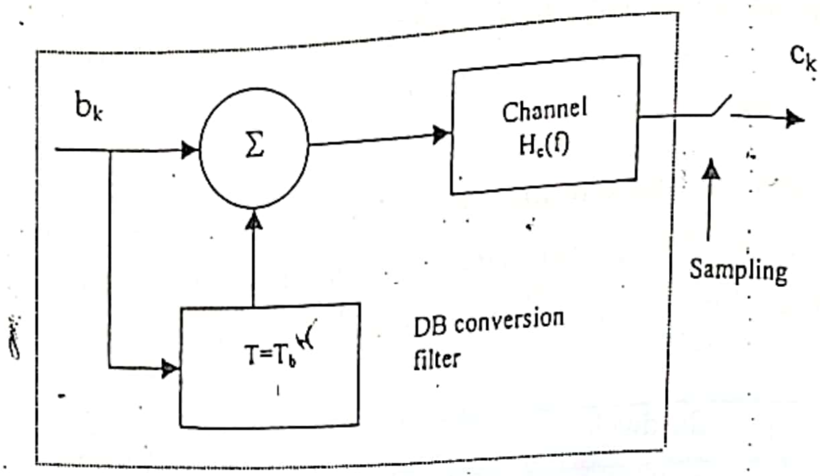
In DBE, input bit streams b_k represented as + A for digit 1 and –A for digit 0 is converted into three level bit sequence C_k having three levels +2A, 0 and –2A by employing the following technique:

$$C_k = b_k \oplus b_{k-1},$$

Such that

$$C_k = \begin{cases} +2A & \text{if } b_k \;\&\; b_{k-1} \text{ are both } 1 \\ 0 & \text{if } b_k \;\&\; b_{k-1} \text{ are different} \\ -2A & \text{if } b_k \;\&\; b_{k-1} \text{ are both } 0. \end{cases}$$

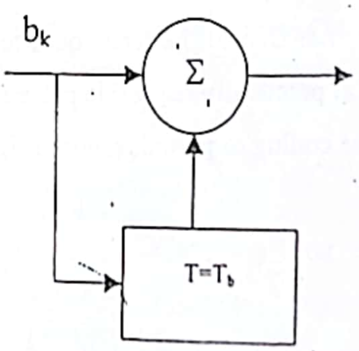
Scanned with CamScanner



Now the frequency response of a filter consisting of DB conversion & channel:

$$H(f) = H_{DB}(f).H_c(f)$$

Where $H_{DB}(f)$ is the frequency response consisting of a adder and delay line network-



and is equal to

$$H_{DB}(f)=1+\exp(-j2\pi fT_b)$$

Then

$$H(f)=H_c(f)[1+\exp(-j2\pi fT_b)]$$

$$=H_c(f)[\exp(j\pi fT_b)+\exp(-j\pi fT_b)]\exp(-j\pi fT_b)$$

$$=H_c(f).2\cos(\pi fT_b)\exp(-j\pi fT_b)$$

$$=2H_c(f)\cos(\pi fT_b)\exp(-j\pi fT_b)$$

Scanned with CamScanner

Assuming that $H_c(f)$ is ideal

$$H_c(f) = \begin{cases} 1 & \text{for } |f| \le \frac{1}{2T_b} = B_0 \\ 0 & \text{for } |f| > \frac{1}{2T_b} \end{cases}$$

We get

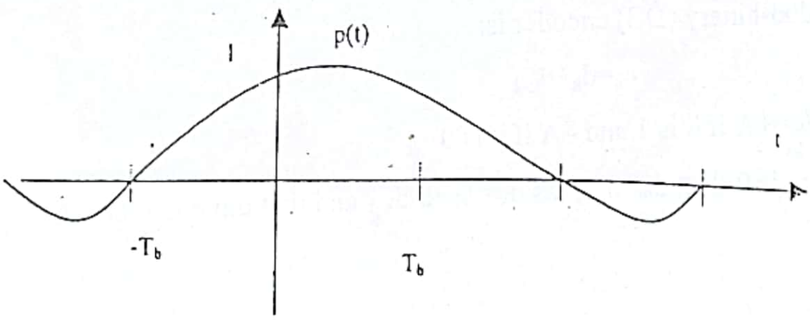
$$H(f) = \begin{cases} 2\cos(\pi f T_b)\,\exp(-j\pi f T_b) & \text{for } |f| \le \frac{1}{2T_b} \\ 0 & \text{elsewhere} \end{cases}$$

The advantage of this transfer function H(f) is that for transmission of R_b signals the required bandwidth is only B_0 (Nyquist bandwidth) and that H(f) is easily approximated and implemented.

The impulse response of DB conversion filter with the transfer function H(f) is-

$$h(t) = \frac{\sin(\pi t/T_b)}{\pi t/T_b} + \frac{\sin[\pi(t-T_b)/T_b]}{\pi(t-T_b)/T_b}$$

$$= \frac{T_b^2 \sin(\pi t/T_b)}{\pi t(T_b - t)}$$



Scanned with CamScanner

The above figure show that p(t) have two distinguished values at sampling instances $+T_b$ & $-T_b$.

Now let us discuss the method of recovering transmitted bit in duo- binary encoding. It is apparent that the original bit b_k can be recovered from C_k by subtracting previously decoded bit from currently received digit C_k -

$$\hat{b}_k = c_k - \hat{b}_{k-1}$$

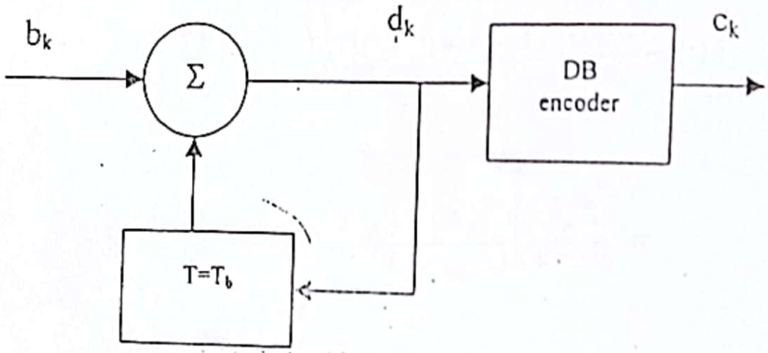
The above equation will yield correct b_k only if the previous bit b_{k-1} was correctly decoded at sampling instance $t=(k-1)T_b$

As decoding of present data require previously decoded data, any error in decoding previous data will affect the present decoding, resulting an error. The draw back of this system is that the errors tend to propagate.

To overcome this problem we use pre- coding of input sequence in such way that the input to duo- binary coder is a sequence represented as

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

The operation $\oplus$ is modulo –2 addition and equivalent to Exclusive OR operation .



Then the output of the duo-binary (DB) encoder is:

$$c_k = d_k + d_{k-1}$$

Now let's assume that d_k =+A if it is 1 and –A if it is 0.

Then if m-th input bit b_m is 0 then $d_m = d_{m-1}$ (as $d_k = b_k \oplus d_{k-1}$ and that dm could be +A or –A)

Scanned with CamScanner

Then as per the rule

$$c_k = \begin{cases} 2A & \text{if } d_k \text{ and } d_{k-1} \text{ are both } 1 \\ 0 & \text{if } d_k \text{ and } d_{k-1} \text{ are different} \\ -2A & \text{if } d_k \text{ and } d_{k-1} \text{ are both } 0 \end{cases}$$

it is evident that for $b_m = 0$, c_m will be either 2A or –2A, i.e ±2A.

Now if $b_m=1$, then d_m is the complement of d_{m-1} i.e. d_m and d_m-1 will be always different, resulting $c_m=0$

Therefore rectifying c_k and setting the threshold level at A and -A we can correctly decode b_m by applying rule :

$$b_m=0 \text{ if } c_m > \pm A$$

$$b_m=1 \text{ if } c_m < \pm A$$

It can be shown now that the decoding of present symbol (b_k) is not affected by correctives of decoding of immediate past symbol (b_k-1)

Assumption: $d_{k-1}=1$

Count	(k-1)	k					
Input seq.b_k		0	0	1	0	1	0
d_k	1	1	1	0	0	1	1
Rep.of d_k	+A	+A	+A	-A	-A	+A	+A
Output of DB encoder ($c_k=d_k+d_{k-1}$)		+2A	+2A	0	-2A	0	+2A
decoded bit		0	0	1	0	1	0

Scanned with CamScanner

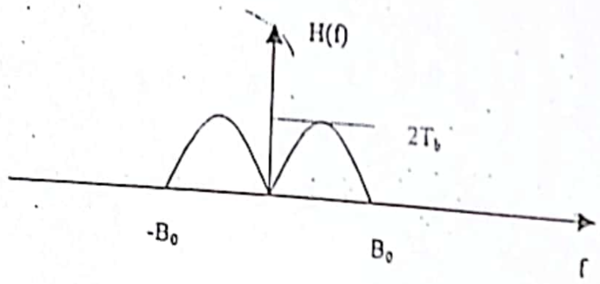
Assumption: $d_{k-1}=0$

Count	(k-1)	k					
Input seq.b_k		0	0	1	0	1	0
d_k	0	0	0	1	1	0	0
Rep.of d_k	-A	-A	-A	+A	+A	-A	-A
Output of DB encoder ($c_k=d_k+d_{k-1}$)		-2A	-2A	0	+2A	0	-2A
decoded bit		0	0	1	0	1	0

Modified Duo binary signaling:

The DB signaling scheme has non-zero frequency response at f=0, it means it is not suitable for circuitry with no DC path. As most of the circuitry uses non-DC path, another signaling scheme called modified DB (MDB)is employed

In MDB instead of one T_b delay, $2T_b$ is employed. With this modification the frequency response becomes symmetrical with respect to the origin (f=0) with zero magnitude at the origin (f=0).



Scanned with CamScanner

3.6 M-ary Signaling

In binary signaling, the output of the pulse generator can have one of two possible levels, where as in M-ary signaling the output can have one of M possible levels. In M-ary system the input source emits one of M distinct symbols and each symbol is assigned a distinct level out of M possible levels. For example, for M=4, the signal representation can be defined as:

Input symbol	Representation Level
B	-3A
C	-A
D	+A
E	+3A

Combination of two bits can be represented by one level in M-ary signaling system. For M=4, the combination of input bit stream in binary form can be represented by slicing the stream into four groups of two bits (00, 01, 10, 11). Each group is then assigned a fixed level.

In M-ary PAM, as in case of binary PAM, the pulse generator output level takes one of four possible levels. These levels are encoded (shaped) and transmitted through the channel.

At receiving end, the decoded output is compared with preset threshold values (slicing levels) and decision is made.

The signaling rate and the bandwidth requirement:

Let R_s (symbols/sec or Baud) be the rate of symbols emitted by the source. Now if the M symbols emitted are equi-probable and statistically independent, then the source entropy will be:

Scanned with CamScanner

$$H = \frac{N \sum_{i=1}^{N} p_i \log_2(1/p_i)}{N} = \log_2 M$$

Where $p_i = 1/M$ (for equi -probability) and

$$\sum_{i=1}^{N} p_i = 1$$

Therefore the source entropy is:

$$H = \log_2 M$$

R_s – Symbol generation Rate

And the information rate for M-ary system will be:

$$R_M = R_s \log_2 M \quad \text{(bits/sec.)}$$

The signaling interval duration $T_s = 1/R_s$ is same for both binary and M-ary systems. Therefore the absolute minimum bandwidth required to transmit $R_s \log_2 M$ bits/sec of information is $R_s/2$ Hz.

Under similar conditions ($T_s = T_b$), the signaling rate for binary system is-

$$R_b = R_s \log_2 2 = R_s$$

and the bandwidth is also $R_s/2$.

It means M-ary signaling can transmit data $\log_2 M$ times faster than binary system under similar conditions.

The price paid for higher speed (or subsequently less bandwidth compared to binary) in M- ary system is the power required to transmit M-ary signal.

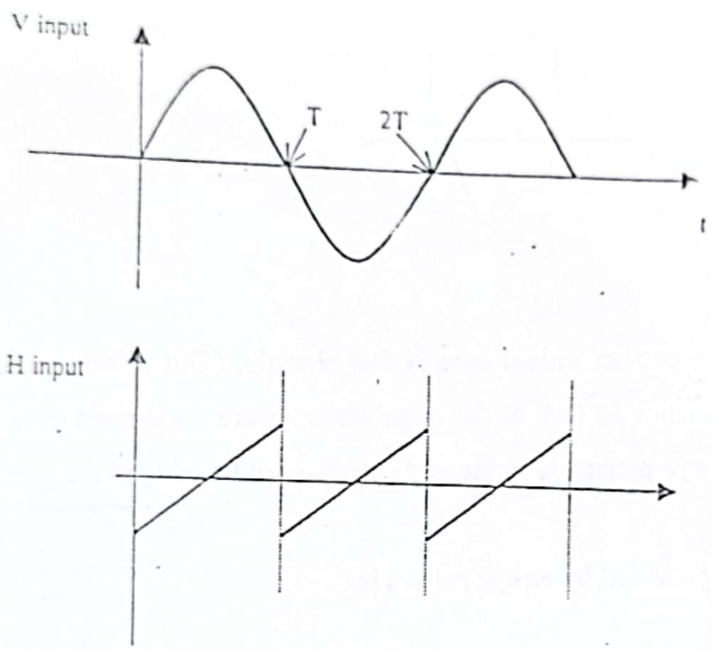
In general for $M \gg 1$ and the error probability $P_e \ll 1$, the power required in M- ary is nearly $M^3 / 3\log_2 M$ times greater than in binary.

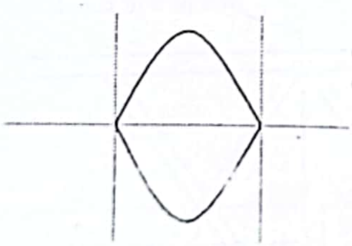
M-ary systems are more complex since it requires (M-1) comparators (level slicers) at the receiving end.

Scanned with CamScanner

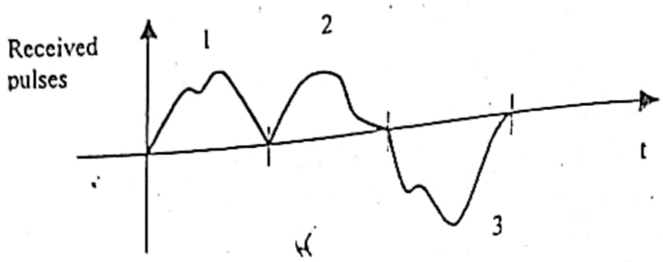
3.7 The Eye Diagram

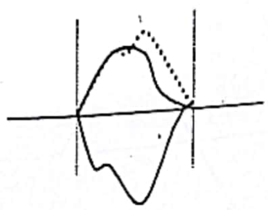
The performance of baseband PAM can be analyzed and observed on the screen of an oscilloscope by applying received pulses to the vertical input and a sawtooth signal with duration $T=T_b$ or the fraction of it (like $2T_b$)at the horizontal input. In such case all the pulses to be displayed in one trace in a normal scope will be converted into one single interval of T.





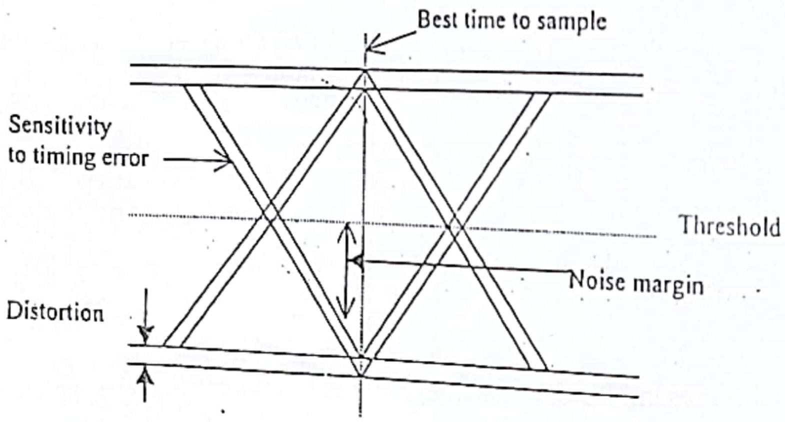
The resulting display resembles the human eye and is therefore called eye patterns. The above example is for the case when the input signal is sinewave. Now let us assume that the received signal is sequence of pulses in distorted form as shown below. The resulting eye pattern is also illustrated.





The resulting eye pattern in real case is fast changing. But if we consider that the channel characteristics as well as the noise statistics are unchanged over the time of observation, the eye pattern is stable and can be used to analyze the performance of the channel.

The information provided by an eye pattern is:



Scanned with CamScanner

1. The best time to sample is the instance when the eye opening is largest.
2. The maximum distortion and ISI are indicated by vertical width of the two branches at sampling time.
3. Immunity to noise (noise margin) is indicated by the width of eye opening.
4. Sensitivity to timing error is indicated by rate of closing of the eye.

Non-linearity in the channel will produce asymmetries in successive eye patterns (strictly linear channel and random data will yield identical eye openings)

Scanned with CamScanner

Chapter 4. Representation of Random Signals and Noise in Communication System

4.1 Random Variables and Processes

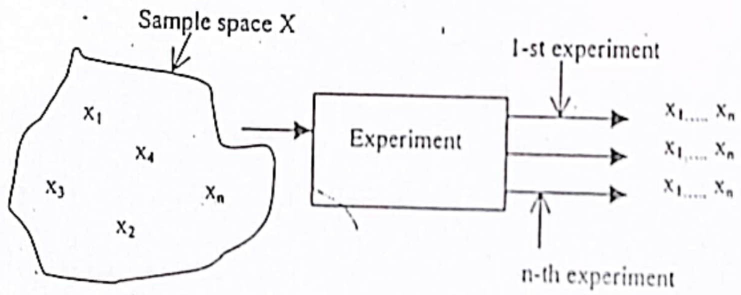
Variables are classified in to two categories: deterministic and random. If the outcome of an experiment can be exactly predicted or counted then the result is called deterministic. Example is the result of the examinations, in which the students will pass or fail and score the numbers in between 0 to 100.

If the outcome of an experiment can only be predicted with some probabilities, then the outcome of that experiment is referred to as random variable.

Example: A die with six different faces (1-6) is cast - An experiment

The fact that the outcome of this experiment will be one of the six possible numbers is deterministic variable. But the result of the experiment - which of the possible six numbers appears- can be predicted with probability only.

In general a function whose domain is a sample space and whose range is some real number is called a random variable of the experiment.



In the above example $x_1 \ldots x_n$ are the real numbers.

If $x_1 \ldots x_n$ is countable then the random variable is called discrete random variable. If $x_1 \ldots x_n$ is uncountable, it is called continuous random variable. The probability $P(X \leq x)$ that in an experiment $X \leq x$, where x is pre assigned or fixed, is called distribution function.

$$F_x(x) = p(X \leq x)$$

Scanned with CamScanner

Properties of $F_x(x)$

1. $0 \le F_x(x) \le 1$
2. $F_x(x_1) \le F_x(x2)$, if $x_1 < x_2$, i.e. $F_x(x)$ is non-decreasing.

The derivative of $F_x(x)$ is called probability density function (pdf):

$$\frac{dF_x(x)}{dx} = f_x(x)$$

$f_x(x)$ in general indicate the probability that the outcome of an experiment X lies between x_2 & x_1.

$f_x(x).dx = p(x < X \le x+dx)$

or $p(X \le x_2) - p(X \le x_1) = F_x(x_2) - F_x(x_1)$

$$= \int_{x1}^{x2} f_x(x)dx$$

And that

$$\boxed{\int_{\infty}^{\infty} f_x(x)dx = 1}$$

~~Since, $F_x(\infty) = 1$ sure event~~

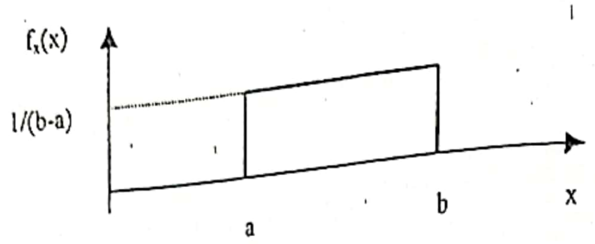
~~$F_x(-\infty) = 0$ impossible~~

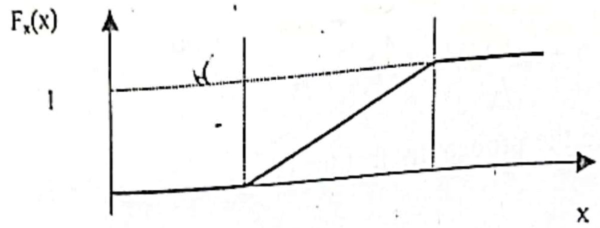
Example: Uniformly distributed random variable

$$f_x(x) = \begin{cases} \dfrac{1}{b-a} & , \quad a \le x \le b \quad (b > a) \\ 0 \quad elsewhere & \end{cases}$$

Then,

$$F_X(x) = \begin{cases} 0 & if\ x < a \\ \dfrac{x-a}{b-a} & if\ a \le x \le b \\ 0 & if\ x > b \end{cases}$$





Moments of probability density function:

The nth moment of a pdf $f_x(x)$ is defined as:

$$E[X^n] = \int_{-\infty}^{\infty} X^n f_X(x)dx$$

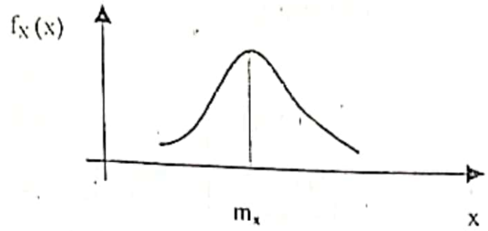
where E[.] is the Expectation operator.

The first moment (n=1) indicate the mean or expected value of the random experiment:

$$m_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$$

m_x is the indication of location of center of gravity of the area under pdf curve, i.e. the average of the value that a random variable can take, in a large number of experiments.

Scanned with CamScanner



The expectation operator for n=2 gives mean square value:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx \;.$$

Central moment is the moment of difference between X and its mean value:

$$\boxed{E[(x-m_X)^n] = \int_{-\infty}^{\infty} (x-m_x)^n f_X(x)\, dx}$$

The first central moment, as seen from above equation, is always zero. The second central moment is called variance:

$$\boxed{\begin{aligned} \sigma_x^2 &= \int_{-\infty}^{\infty} (x-m_x)^2 f_X(x)\, dx \\ \sigma_x^2 &= E[X^2] - m_x^2 \end{aligned}}$$

The variance σ_x^2 is the measure of dispersion (spread) of pdf of X.

Square root of σ_x^2 is called the standard deviation.

The correlation function:

Joint moment of first order of two random variables X & Y is called correlation.

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\; f_{XY}(x,y)\, dx\, dy$$

Joint central moment of X & Y is called covariance of X & Y.

$$\mathrm{Cov}[XY] = E\{[(x-E(x))]\,[Y-E(Y)]\}$$

$$= E[(x-m_x)(y-m_y)]$$

$=E[XY]-m_x m_y$

$$\boxed{\rho_{xy} = \frac{Cov[XY]}{\sigma_x \sigma_y}}$$

The correlation coefficient ρ_{xy} is defined as:

where σ_x, σ_y are standard deviations of X & Y. Followings are the basic statements related to the correlation between X and Y:

i. X & Y are uncorrelated, if and only if cov[XY] is zero.

ii. X & Y are orthogonal if and only if E[X,Y]=0

iii. If both X & Y have zero mean (i.e. $m_x=m_y=0$) and if they are orthogonal random variable (E[XY]=0), then they are uncorrelated and vice versa.

Example of probability density function of some standard random variables.

i. Uniform distribution

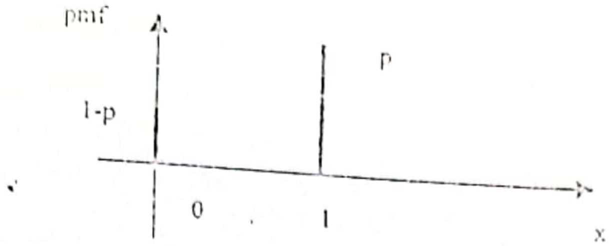
$$f_X(x) = \begin{cases} \frac{1}{(b-a)} & \text{if } a \le x \le b \\ 0 & \text{elsewhere} \end{cases}$$

$$m_x = \frac{b+a}{2}, \quad \sigma_x^2 = \frac{(b-a)^2}{12}$$

$$\sigma_x = \frac{(b-a)}{\sqrt{12}}$$

ii. Bernoulli random variable : Discrete random variable that takes only two values 1and 0 with probabilities p & 1-p. The probability mass function (pmf) is

Scanned with CamScanner

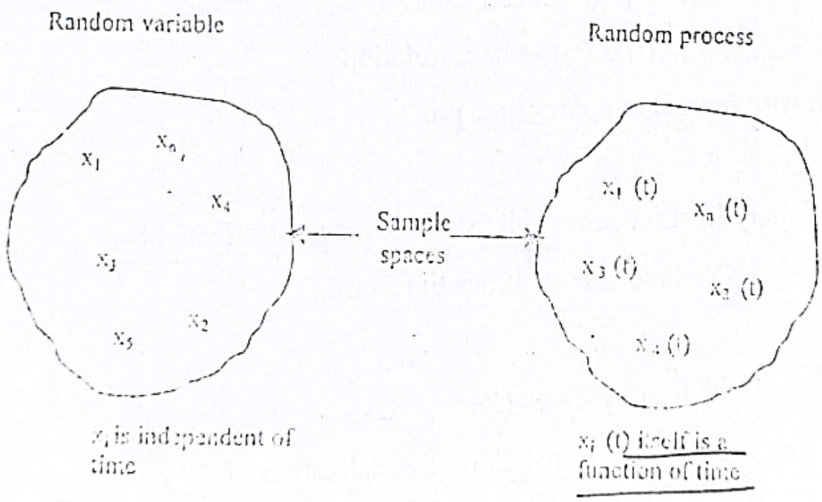


iii. Gaussian (Normal) random variable

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_x}} e^{\frac{-(x-m_x)^2}{2\sigma_x^2}}$$

Random process

Random variable deal with only one variable that can take any value within the given range. In practice, especially in communication, the source of signal as well as noise are functions of time also. It means that at any instance t_i, the value of the signal is random variable. Such process or signal is called random process.



The observation of instantaneous values of a Random Process (RP) within a certain interval of time is called its realization. Theoretically such set of realization

B-4207

Scanned with CamScanner

is infinity. The realizations $x_1(t)$.........$x_n(t)$, $n \to \infty$ form statistical ensemble of realization, where in $x_1(t)$.........$x_n(t)$ are itself time dependent random variables. In ther words a RP's realization is the outcome of an experiment from sample space λ

$$x(\lambda, t) \xrightarrow[\lambda]{i^{th}} x(\lambda_i, t) = x_{\lambda_i}(t)$$

The realization is a function of time and for any given $t = t_0$, the realization is a random variable (number)

$$x_{\lambda_i}(t = t_0) = x(\lambda_i, t_0) \to number$$

Joint Distribution function of a RP can be described in terms of joint probability distribution function.

$P\{X(t_1) \leq a_1, X(t_2) \leq a_2, \ldots\ldots\ldots\ldots X(t_n) \leq a_n\}$ for any n and t_n

A RP can be described using its statistical averages (ensemble averages). The most common are means & autocorrelation functions.

$m_X(t)=E[X(t)]$-mean value

$R_{XX}(t)=E[X(t_1)X(t_2)]$-autocorrelation

These values are taken with respect to appropriate pdf.

Stationarity:

A RP is said to be stationary in strict sense if its statistics are not affected by shift in time origin i.e. $X(t)=X(t+\epsilon)$ where ϵ is arbitrary time shift.

In general,

$$f_{X(t)}(X)=f_{X(t+\epsilon)}(X)$$

Physically stationarity implies that time translation of a realization of RP results in another realization having same statistical properties. Here the statistical properties refer to all moments. In practice we deal with simple moments (mean, covariance and autocorrelation), therefore it is easier to limit stationarity to these statistical

Scanned with CamScanner

averages only. In this sense a RP is called stationary in wide sense if its mean & AC function do not vary with a shift in time origin. i.e.

$$E\{X(t)\}=E\{X(t+\tau)\}=\text{constant}$$

$$R_{XX}(t_1 t_2)=R_{XY}\{t+t_1, t+t_2\}=R_{XY}\{|t_2-t_1|\}=R_{XX}(\tau)$$

$$\text{or} \quad R_{XX}(\tau)\triangleq E\{X(t)X(t+\tau)\}$$

Naturally all strict sense stationary process (SSSP) is wide sense stationary process (WSSP) but vice versa is not true always.

AC => Autocorrelation

Properties of AC function of a WSSP:

i. $R_{XX}(\tau)=R_{XX}(-\tau)$- symmetry condition
ii. $R_{xx}(0)=E\{X(t-0)x(t)\}=E\{X^2(t)\}$- mean square value
iii. $|R_{XX}(\tau)| \leq R_X(0)$-maximum value at zero time lag

Time Averaging:

Consider a SSRP whose mean value and AC function are described as :

$$m_X = \int_{-\infty}^{\infty} f_{X(t)}(x)\, dx \quad \text{for at } -\infty < t < \infty$$

$$\text{and} \quad R_X(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{x(t)x(t-\tau)}(x,y)\, dx\, dy$$

$$\text{for all } -\infty < t < +\infty$$

Evaluation of these parameters by ensemble averaging requires prior knowledge of all possible realizations or knowledge of pdf's. In most cases these pdf's are simply not available (since they are random in nature). In practice the only available data is one sample (realization) recording of RP and it is therefore more convenient to describe RP by time averaging its parameters:

$$\langle m_x\rangle = \langle x(t)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)\, dt$$

$$\langle R_{XX}(\tau)\rangle = \langle x(t)\, x(t-\tau)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)\, x(t-\tau)\, dt$$

Scanned with CamScanner

where 2T is the observation time.

$\langle m_X \rangle$ and $\langle R_{XX}(\tau) \rangle$ are random variables in the sense that these values depend upon the particular realization of RP for which the average is taken.

Ergodicity:

In general ensemble average and time average are not equal for most of RP, except for the random process called the Ergodic Random Process.

A RP is said to be ergodic in general form, if its statistical average mean value m_X and AC $R_{XX}(\tau)$ are equal to the time average values $\langle m_X \rangle$ and $\langle R_{XX}(\tau) \rangle$. In other words:

$$E[\langle m_X \rangle]=m_X$$

$$E[\langle x(t)x(t-\tau) \rangle]=R_{XX}(\tau)$$

$$E[\langle m_X \rangle] = E\left[\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}x(t)\,dt\right] \xrightarrow{Stationarity}$$

$$= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}E[x(t)]\,dt = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}m_x\,dt$$

$$= m_x \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}dt = m_x$$

The AC and psdf of an ergodic process:

For RP, the FT of its AC function is called psdf

$$S_x(f) = \int_{-\infty}^{\infty} R_{xx}(\tau)\exp(-2\pi j\tau)\,d\tau$$

provided

$$\int_{-\infty}^{\infty} R_{xx}(\tau)\,d\tau < \infty$$

In fact $R_{XX}(\tau)$ is the measure of average power dissipated in 1Ω resistance

Scanned with CamScanner

$$E\left[x^2(t)\Big|_{t=0}\right] = R_{XY}(0) = \int_{-\infty}^{\infty} S_X(f)\, df$$

If the process is ergodic then

$$\langle S_X(f)\rangle = S_X(f)$$

$$\langle S_X(f)\rangle = \lim_{T\to\infty}\frac{1}{T}\left[\int_{-T/2}^{T/2} X_T(t)\exp(-j2\pi ft)\, dt\right]^2$$

Interpretation of various time average means of an ergodic process:

i: The mean value $m_X=\langle X(t)\rangle$ represents the DC component of the signal

ii. The mean square value $\langle X_2(t)\rangle = R_{XX}(0)$ represents the total average power

iii. The square of mean value $[\langle X_2(t)\rangle]^2$ represents the DC power

iv: The variance $\langle X^2(t)\rangle - \langle X(t)\rangle^2$ represents the AC power

v. The standard deviation

$$\sqrt{\langle X^2(t)\rangle - \langle X(t)\rangle^2}$$

represents the RMS value of the random signal.

The random processes, used to simulate signals in communication, are considered to be stationary and ergodic in most general form. Therefore further dealing of RP will be time averaged.

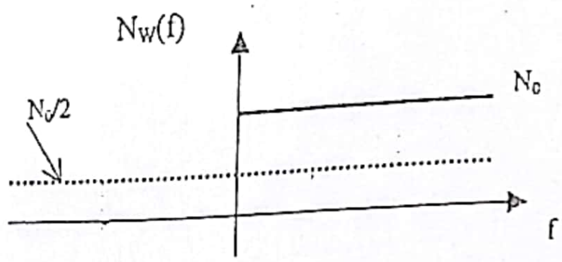
Scanned with CamScanner

4.2 Noise in Communication: Representation, Characteristic and PSDF

White noise (WN) is an ideal case of description of noise in communication. WN has flat spectral density $N_W(f)$ over $-\infty < f < \infty$ and has zero mean value.

$$N_W(f) = N_0/2 \text{ for } -\infty < f < \infty$$

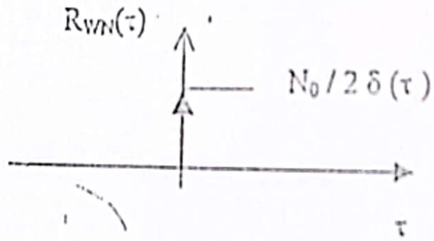
$$= N_0 \text{ for } 0 \leq f < \infty \text{ or } -\infty < f < 0$$



The AC function of the white noise is therefore a delta function.

$$\boxed{R_{WN}(\tau) = N_0/2\,\delta(\tau)}$$

$R_{WN}(\tau) = \frac{N_0}{2}\delta(t)$



As $R_{WN}(\tau)=0$ for $\tau\neq0$, two different samples of WN, no matter how close they are in time shift ($\tau\to0$), are uncorrelated.

The origin of the terminology "white noise" is white light where all color frequencies exist.

White noise has no physical significance, as it does not exist in nature.

Thermal noise:

The noise produced by the random movement of electrons in any conducting material due to thermal energy has psdf expressed as :

Scanned with CamScanner

$$N_T(f) = \frac{hf}{2\left(e^{hf/KT} - 1\right)} \; watt/hz$$

where,

h - Planck's constant

K- Boltzman's constant

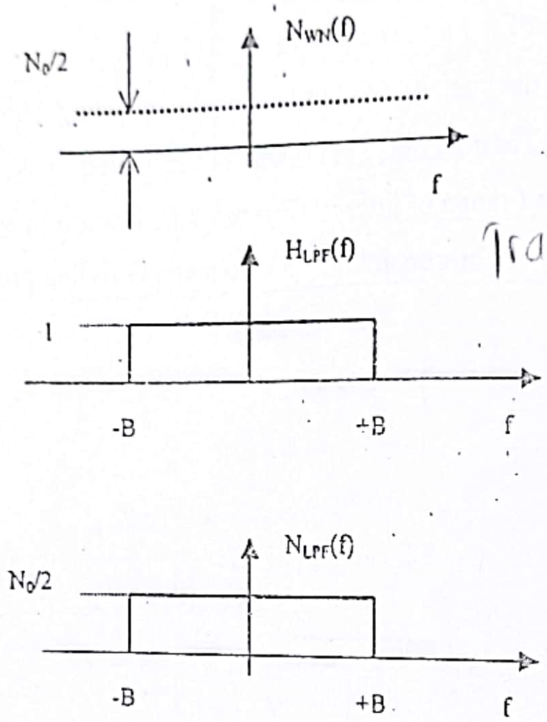
T- Temperature in ^0K

$N_T(f)$ has its maximum at f = 0 and is equal to $KT/2=N_0/2$. As the frequency is increased, $N_T(f)$ slowly decline and tend to zero as the frequency tend to infinity. For all practical purpose the Thermal noise (TN) can be considered as WN with flat spectrum KT/2 for entire f range of interest. For most application in communication engineering, white noise can be considered as zero mean Gaussian process with psdf equaling to $N_0/2$.

4.3 RC Filtering of White Noise

Ideal Low Pass filtering of White Noise:

Let N(f) be the spectrum of a White Noise (WN) with pdf equaling to $N_0/2$. The WN is applied to the input of an ideal LPF with bandwidth *B*.



Transfer function of ·LPF

It is obvious from the above illustration that the pdf of the noise at the output of the filter is:

$$N_{LPF}(f) = \begin{cases} \dfrac{N_0}{2} & for\ -B < f < B \\ 0 & elsewhere \end{cases}$$

Scanned with CamScanner

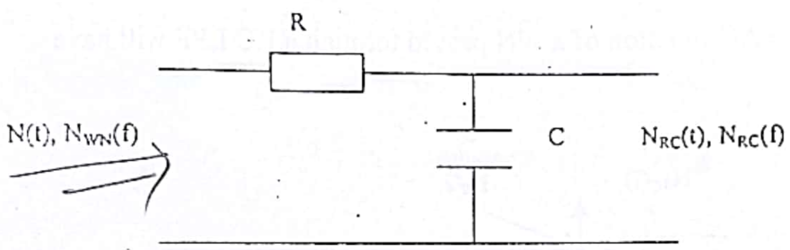
And the autocorrelation function is:

$$R_{NLPF}(\tau) = \int_{-B}^{B} \frac{N_0}{2} \exp(j2\pi f\tau)\, df$$
$$= N_0\, B\, Sinc\,(2B\tau)$$

$N_0\, B\, sinc\,(2\pi B\tau)$

RC Filtering

Let the white noise N(t) with spectrum $N_{WN}(f)$ pass through a RC LPF as shown below:



The frequency response of the RC filter can be expressed as:

$$H(f) = \frac{1}{1 + j2\pi fRC}$$

$$|H(f)| = \frac{1}{1 + (2\pi fRC)^2}$$

And the psd of the noise at the output of the RC filter would be:

$$N_{RC}(f) = N_{WN}(f) \times \frac{1}{1 + (2\pi fRC)^2} = \frac{N_0}{2} \times \frac{1}{1 + (2\pi fRC)^2}$$

Scanned with CamScanner

The AC function of the output noise is thus expressed as:

$$R_{RC}(\tau) = \int_{-\infty}^{\infty} \frac{N_0/2}{1+(2\pi f RC)^2} \exp(j2\pi f\tau)df = \frac{N_0}{4\pi}\int_{-\infty}^{\infty}\frac{\exp(j\omega\tau)}{1+(\omega RC)^2}d\omega$$

as $w = 2\pi f$

$\therefore \frac{dw}{2\pi} = df$

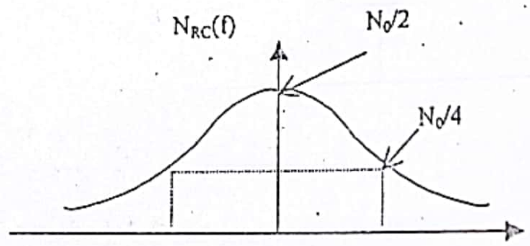
or,

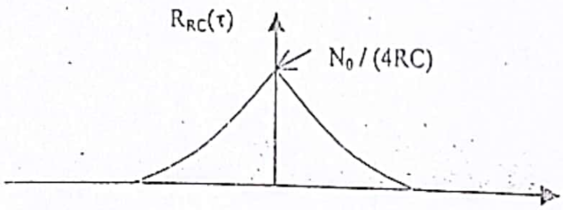
$$R_{RC}(\tau) = \frac{N_0}{2\pi}\int_{-\infty}^{\infty}\frac{\cos\omega\tau}{1+(\omega RC)^2}d\omega$$

The above integral is approximated as the following and is tabulated:

$$R_{RC}(\tau) = \frac{N_0}{2\pi}\times\frac{\pi}{2RC}\exp\left(\frac{-|\tau|}{RC}\right) = \frac{N_0}{4RC}\exp\left[\frac{-|\tau|}{RC}\right]$$

Thus the spectrum and the AC function of a WN passed through a RC LPF will have the following shape:





Scanned with CamScanner

Noise Equivalent Bandwidth

From the previous chapter we can conclude that the noise power at the output of an ideal LPF is finite and proportional to the bandwidth of the filter:

$$P_{Nideal}=\frac{N_0}{2}\times 2B=N_0B$$

whereas the noise power at the output of a RC filter is theoretically infinite and is defined only by the transfer function of the RC filter:

$$P_{NRC}=\frac{N_0}{2}\int_{-\infty}^{\infty}\left|H_{RC}(f)\right|^2 df=N_0\int_{0}^{\infty}\left|H_{RC}(f)\right|^2 df$$

To generalize the noise power, we need to define some standard parameter called noise equivalent bandwidth B_N that can be used to calculate average noise power.

In case of a RC filter example we can replace the transfer function of the RC filter with an ideal LP filter having the transfer H(0) and the bandwidth equal to B_N so that the noise powers at the output of this idealized filter and the RC filter are equal.

$$P_{Nideal}=N_0B_NH^2(0)$$

$$P_{Nreal}=N_0\int_{0}^{\infty}\left|H(f)\right|^2 df$$

Since the assumption is that both the above noise powers are equal, we get

$$B_N=\frac{\int_{0}^{\infty}\left|H(f)\right|^2 df}{\left|H^2(0)\right|}$$

Similarly for a Band pass filter, the equivalent noise bandwidth can be calculated as:

$$B_N = \frac{\int_0^{\infty} |H(f)|^2 df}{|H^2(f_c)|}$$

And in general, the equivalent noise bandwidth is expressed as:

$$B_N = \frac{1}{g_a} \int_0^{\infty} |H(f)|^2 df$$

where g_a is called maximum available power gain of the filter and represents the maximum value of $|H(0)|^2$ or $|H(f_c)|^2$. Therefore in terms of equivalent noise bandwidth, the noise power at the output of any filter is expressed as:

$$P_N = N_0 B_N g_a$$

The effect of passing the white noise through a filter can be summarized as follows:

- the output noise power is proportional to the power gain of the filter;
- the output noise power is proportional to the noise equivalent bandwidth; and
- the effect of the noise can be reduced by narrowing system bandwidth, i.e. by making the system more selective.

Let a wide sense stationary random process (WSSP) x (t) is applied to the input of a LTI having the impulse response and transfer function h (t) and H (f) respectively. We now find the mean, autocorrelation function and the spectrum of the output of the LTI. By definition of the LTI, the output is the convolution of the input signal and the impulse response of the system:

Noise through LTI

$$y(t) = \int_{-\infty}^{\infty} x(t-\alpha)h(\alpha)d\alpha \overset{stationarity}{=} \int_{-\infty}^{\infty} x(\alpha)h(t-\alpha)d\alpha$$

The mean value of the y(t) is thus equal to:

$$E[y(t)] = E\left[\int_{-\infty}^{\infty} x(\alpha)h(t-\alpha)d\alpha\right]$$

Since the expectation operation E[.] is a linear operation, we can re-write the above equation as:

$$E[y(t)] = \int_{-\infty}^{\infty} E[x(\alpha)]h(t-\alpha)d\alpha = m_x \int_{-\infty}^{\infty} h(t-\alpha)d\alpha = m_x H(0),$$

because-

$$H(f) = \int_{-\infty}^{\infty} h(\alpha)\exp(-j2\pi f\alpha)d\alpha,$$ and for f = 0, we get,

$$H(0) = \int_{-\infty}^{\infty} h(\alpha)d\alpha$$

It means the mean value of the output signal is also independent of the shift in time.

For finding the AC function of y(t), first of all let us determine the cross correlation between the input x(t) and the output y(t).

Cross Correlation

$$R_{xy}(\tau) = E[x(t_1)y(t_2)] = E\left[x(t_1)\int_{-\infty}^{\infty} x(s)h(t_2 - s)ds\right] = \int_{-\infty}^{\infty} E[x(t_1)x(s)]h(t_2 - s)ds =$$

$$= \int_{-\infty}^{\infty} R_{xx}(\tau)h(t_2 - s)ds \quad = \int_{-\infty}^{\infty} R_{xx}(t_1 - s)h(t_2 - s)\,ds$$

Now substituting the variable of integration by $u = s - t_2$, we get

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} R_{xx}(t_1 - t_2 - u)h(-u)du =$$

$$= \int_{-\infty}^{\infty} R_{xx}(\tau - u)h(-u)du = R_{xx}(\tau) * h(-\tau)$$

The AC function of the output signal y(t) can be determined as:

$$R_{yy}(t_1, t_2) = E[y(t_1)y(t_2)] = E\left[y(t_2)\int_{-\infty}^{\infty} x(s)h(t_1 - s)ds\right] = \int_{-\infty}^{\infty} E[x(s)y(t_2)]h(t_1 - s)ds =$$

$$= \int_{-\infty}^{\infty} R_{xy}(s - t_2)h(t_1 - s)ds$$

Substituting the variable of integration by $u = s - t_2$, we get

$$R_{yy}(t_1, t_2) = \int_{-\infty}^{\infty} R_{xy}(u)h(\overbrace{t_1 - t_2}^{(\tau)} - u)du = R_{xy}(\tau) * h(\tau)$$

Substituting the expression for $R_{xy}(\tau)$ in the above equation, we get,

$$R_{yy}(\tau) = R_{xx}(\tau) * h(-\tau) * h(\tau)$$

It is evident from the above derivations for mean value and AC function that both these parameters for the output signal are not dependent upon the shift in time. Therefore it can be concluded that the output of a LTI to a WSSP excitation is also a WSSP.

Scanned with CamScanner

In frequency domain analysis, the spectrum of the output signal can be derived from the AC function in the following way. From the analysis of random signal, we know that the pdf and AC function are related by Fourier transform, i.e.,

$$S(f) = FT\,[R(\tau)]$$

Similarly it is established that if

$$FT[h(\tau)] = H(f), \text{ then}$$

$$FT[h(-\tau)] = H^{*}(f)$$

where $H^{*}(f)$ is the complex conjugate of $H(f)$.

Finally the pdf of the output signal will be:

$$S_y(f) = FT\left[R_{xx}(\tau) * h(\tau) * h(-\tau)\right] = FT\left[R_{xx}(\tau)\right] \times FT\left[h(\tau)\right] \times FT\left[h(-\tau)\right] =$$
$$= S_x(f) \times H(f) \times H^{*}(f) = S_x(f) \times \left|H(f)\right|^2$$

4.5 Analytical Representation of Narrow Band Noise

To obtain a narrow band noise signal, we can pass the white noise through a bandpass filter having extremely narrow bandwidth. The AC function of the output noise from the BPF can be expressed as:

$$R_{yy}(\tau) = F^{-1}[S_y(f)] = 2[S_y(f_c)\Delta f](\cos\omega_c\tau)(Sinc\Delta f\tau)$$

For $\Delta f \rightarrow 0$, Sinc $\Delta f\tau$ is nearly equal to unity and therefore

$$R_{yy}(\tau) \approx 2S_y(f_c)\Delta f(\cos\omega_c\tau)$$

The above expression is the AC function of the randomly phased sinusoid. Therefore the output of the BPF can be expressed as:

$$y(t) \approx C\cos(\omega_c t + \theta) = A\cos\omega_c t + B\sin\omega_c t$$

where C,A,B and θ are random variables.

Further analysis show that the narrow band noise can be represented as :

$$y(t) = \sum_{i=-n}^{n} A_i\cos(\omega_i t) + B_i\sin(\omega_i t)$$

where A_i and B_i are uncorrelated, zero mean random variables.

For the convenience of the mathematical analysis, the narrow band noise is often expressed in quadrature form:

$$y(t) = Y_c(t)\cos\omega_c t - Y_s(t)\sin\omega_c t$$

Scanned with CamScanner

Here Yc(t) and Ys(t) are uncorrelated, zero mean Gaussian processes having the following explicit expressions:

$$Y_c(t) = \sum_{i=-n}^{n} \sqrt{A_i^2 + B_i^2} \cos(2\pi i \Delta f t - \theta_i)$$

$$Y_s(t) = \sum_{i=-n}^{n} \sqrt{A_i^2 + B_i^2} \sin(2\pi i \Delta f t - \theta_i)$$

where $\theta_i = \tan^{-1}\left(\frac{B_i}{A_i}\right)$.

In the envelope and phase representation, the narrow band noise is expressed as:

$$y(t) = R(t)\cos[\omega_c t + \theta(t)]$$

The envelop R(t) and the phase θ (t) are given by:

$$R(t) = \sqrt{[Y_c(t)]^2 + [Y_s(t)]^2}$$

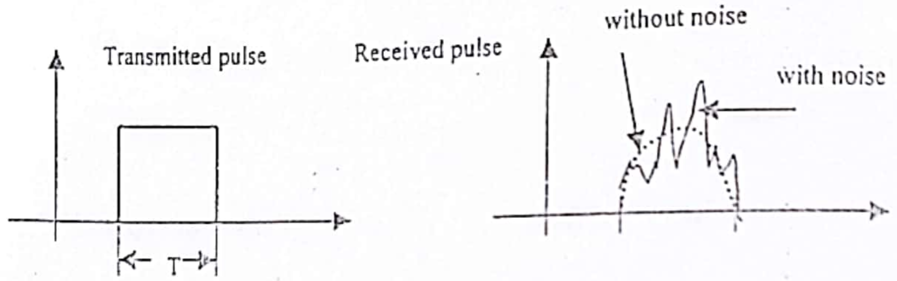
$$\theta(t) = \tan^{-1}\left\{\frac{Y_s(t)}{Y_c(t)}\right\}$$

Scanned with CamScanner

4.6 Optimum Detection of a Pulse in Additive White Noise

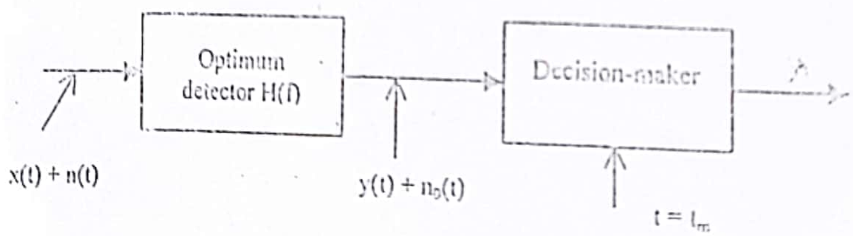
In digital communication message are transmitted by using two symbols 1 & 0. In general, the presence of the signal is represented by a pulse of fixed height and duration and the absence by no pulse. Therefore in digital communication the detection of exact replica of transmitted pulse is not important, as the pulse shape is already known. The only required function of the receiver is to decide the presence or absence of the pulses in a mix of received signal and noise.

The detector is therefore a decision making devices that examines the entire duration of pulses and gives decision whether the pulse is present or absent (of course the decision has to be made in presence of additive noise).



Optimum detector of the pulse is a device that has least probability of errors in making decision in favor of 1 or 0.

The basic idea is to pass the received signal through a network (or filter) that suppress the noise and give sharp peak to the signal at the decision making instance (if the pulse is present) thus creating maximum S/N ratio at decision making instance.



The objective of the optimum detector is to make:

$$\left. y^2(t) \middle/ n_0^2(t) \right|_{\max} \quad at\ t = t_m$$

$$or\ for\ t = t_m, \quad \max\{SNR_0\} = \frac{y^2(t_m)}{n_0^2(t_m)}$$

Now the output of the optimum detector can be derived by taking inverse Fourier transform of product of X(ω) and H(ω):

$$y(t) = F^{-1}[X(\omega)H(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)X(\omega)e^{j\omega t}\,d\omega$$

And for $t = t_m$,

$$y(t_m) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)X(\omega)e^{j\omega t_m}\,d\omega$$

The output noise power at the decision making instance can be expressed as:

$$\overline{n_0^{\,2}(t_m)} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{N}{2}|H(\omega)|^2\,d\omega = \frac{N}{4\pi}\int_{-\infty}^{\infty}|H(\omega)|^2\,d\omega$$

Finally the output signal to noise ratio will be:

$$SNR_0 = \frac{\overline{y^2(t_m)}}{\overline{n_0^{\,2}(t_m)}} = \frac{1}{N\pi}\,\frac{\left|\int_{-\infty}^{\infty} H(\omega)X(\omega)e^{j\omega t_m}\,d\omega\right|^2}{\int_{-\infty}^{\infty}|H(\omega)|^2\,d\omega}$$

To maximize the SNR_0 we can apply Schwarz's inequality principle, which states:

$$\frac{\left|\int_{-\infty}^{\infty} F_1(\omega) F_2(\omega) d\omega\right|^2}{\int_{-\infty}^{\infty} |F_1(\omega)|^2 d\omega} \le \int_{-\infty}^{\infty} |F_2(\omega)|^2 d\omega$$

and holds true only if

$$F_1(\omega) = K F_2^*(\omega),$$

where K is an arbitrary constant and $F_2^*(\omega)$ is the complex conjugate of $F_2(\omega)$.

By substituting $F_1(\omega)$ by $H(\omega)$ and $F_2(\omega)$ by $X(\omega)\exp(j\omega t_m)$ we get,

$$\left|\int_{-\infty}^{\infty} H(\omega) X(\omega) e^{j\omega t_m} d\omega\right|^2 \le \int_{-\infty}^{\infty} |H(\omega)|^2 d\omega \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega$$

$$\text{or } \frac{1}{\pi n} \frac{\left|\int_{-\infty}^{\infty} H(\omega) X(\omega) e^{j\omega t_m} d\omega\right|^2}{\int_{-\infty}^{\infty} |H(\omega)|^2 d\omega} \le \frac{1}{\pi n} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega$$

$$\text{or } SNR_0 \le \frac{1}{\pi n} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega$$

The maximum value of the SNR would be:

$$SNR_0\Big|_{\max} = \frac{1}{\pi n} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega$$

The above expression remains valid only when the prerequisite condition of the Schwarz's inequality remains valid, i.e.:

$$H(\omega) = K \times \left\{X(\omega) e^{j\omega t_m}\right\}^* = K \times X^*(\omega) e^{-j\omega t_m} = K \times X(-\omega) e^{-j\omega t_m}$$

The impulse response of the optimum detector corresponding to the transfer function $H(\omega)$ will be:

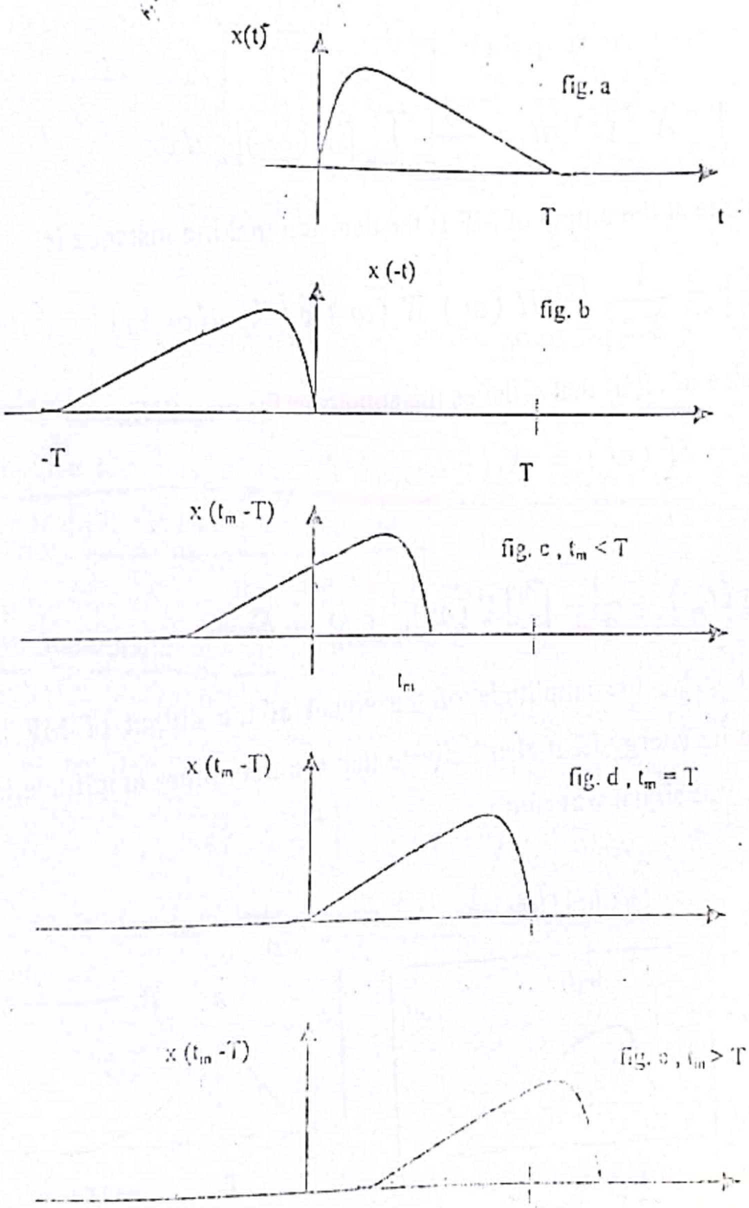
$$h(t) = FT^{-1}\left[H(\omega)\right] = FT^{-1}\left[K \times X(-\omega) e^{-j\omega t_m}\right]$$

Scanned with CamScanner

As the inverse FT of $X(-\omega)$ is $x(-t)$ and $e^{-j\omega t_m}$ gives time shift of t_m, the impulse response will be:

$$h(t)=K\,X(t_m-t)$$

Assuming K=1 for conveniences, the impulse response of the optimum detector network (filter) is the replica of in coming signal shifted by t_m. It is therefore, the optimum detector of a pulse is also called the Matched Filter.



Scanned with CamScanner

The above sequence of figures shows how to achieve the impulse response h(t). Fig. c is the case when $t_m < T$. As $x(t_m - T)$ in this case is non- casual and therefore physically non-realizable. In practice cases with $t_m \geq T$ is used for optimum detection.

For a rectangular pulse, the maximum SNR that could be achieved by the matched filter is:

$$SNR_{max} = \frac{1}{\pi n}\int_{-\infty}^{\infty}|X(\omega)|^2\, d\omega = \frac{2E}{N} = \frac{E}{N/2},$$

- where E is the energy of the signal (pulse):

$$E = \int_{-\infty}^{\infty} X^2(t)\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}|X(\omega)|^2\, d\omega$$

And the signal amplitude at the output of MF at the decision making instance is:

$$y(t_m) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\, X(\omega)\, e^{j\omega t_m}\, d\omega$$

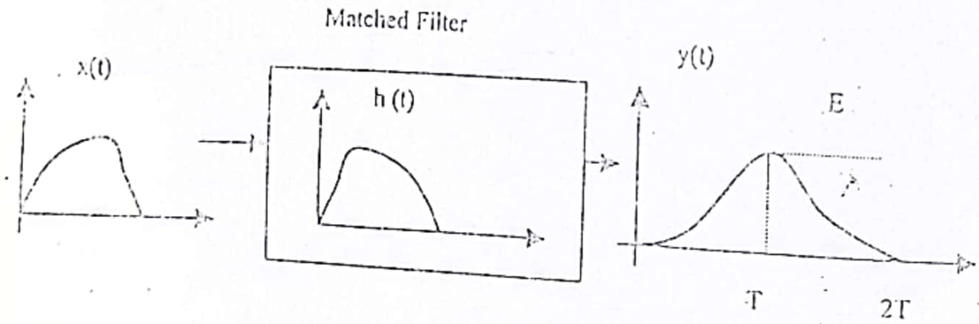
By substituting the value of $H(\omega)$ that satisfies the condition for max SNR -

$$H(\omega) = X(-\omega)e^{-j\omega t_m}$$

We get

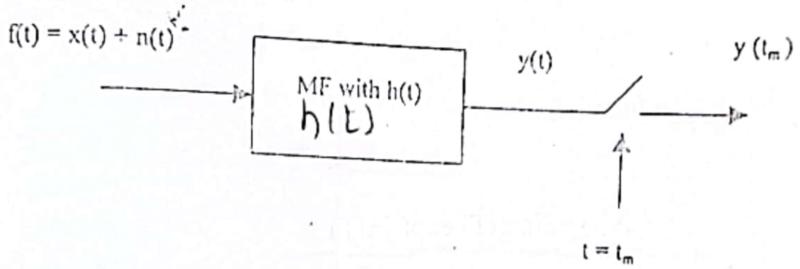
$$y(t_m) = \frac{1}{2\pi}\int_{-\infty}^{\infty}|X(\omega)|^2\, d\omega = E$$

It means at instance $t = t_m$, the amplitude of the signal at the output of MF is maximum and equal to its energy E. It also indicate that the maximum amplitude is independent of the received signal waveform.



Duration of y(t) indicates that the entire input signal must pass through MF in order to achieve maximum SNR. This gives practical constraint when T is large enough.

Realization of Matched filter



The output signal in the above arrangement can be expressed as:

$$y(t) = \int_{-\infty}^{\infty} f(s)\, h(t-s)\, ds$$

As the impulse response of the MF is time-shifted version of the input signal (excluding noise), it can be expressed as:

$$h(t) = x(t_m - t)$$

The time-shifted version of h(t) will be:

$$h(t-s) = x(t_m - t + s) = x(s + t_m - t)$$

And when the decision is made at $t = t_m$, the output signal will be:

$$y(t_m) = \int_{-\infty}^{\infty} f(s)\, x(s)\, ds$$

The above arrangement is called time correlator and obviously is a synchronous (or coherent) detector.

Scanned with CamScanner

Matched Filter for Rectangular Pulse

Suppose the input to the MF is a rectangular pulse with duration T and having amplitude A and that it has unit area i.e. $A \times T=1$, such that

$$x(t) = \begin{cases} A & for\ 0 \le t \le T \\ 0 & elsewhere \end{cases}$$

The FT of x(t) is then a sinc function-

$$X(f) = \text{Sinc}(fT)\exp(-j\pi fT)$$

It would not be difficult to show that the impulse response of MF for this pulse is same as x(t)

$$h(t) = \begin{cases} A, & for\quad 0 \le t \le T \\ 0, & \text{elsewhere} \end{cases}$$

And the transfer function of MF (for AT=1) is:

$$H(f) = \text{sinc}(fT)\exp(-j\pi fT)$$

The spectrum of output signal of MF:

$$Y(f) = X(f) \times H(f) = \text{sinc}^2(fT)\exp(-j2\pi fT),$$

where $\text{sinc}^2(fT)$ is the energy spectrum density of a rectangular pulse x(t) having unit area.

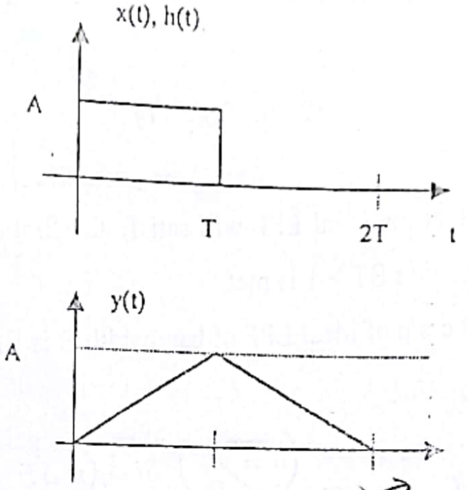
The output waveform can be evaluated by taking convolution of x(t) and h(t).

$$y(t) = \int_{-\infty}^{\infty} x(\alpha)\, h(t-\alpha)\, d\alpha$$

And for AT = 1, the output signal will be:

Scanned with CamScanner

$$y(t) = \begin{cases} \dfrac{At}{T} & 0 < t \le T \\ A\left(2 - \dfrac{t}{T}\right) & \text{for} \quad T \le t \le 2T \\ 0 & \text{elsewhere} \end{cases}$$



Comparison between MF and ideal LPF

A true matched filter gives maximum SNR only when the entire signal component has entered the MF. In other words if the signal duration is T then decision can be made only after time elapse of 2T. This may be impractical when T is too large. Secondly the time correlator MF is difficult to realize as it requires prior knowledge of the shape of received signal (not transmitted). Therefore, for a MF to be practically realizable and feasible the following conditions should be met.

1. The maximum SNR should be achieved at some instance less than T.
2. It should be easily (practically) realizable.

Now let us evaluate the ideal LPF as a matched filter. The transfer function of a MF for a rectangular pulse is a sinc function -

$$H_{MF}(f)=Sinc(fT)\exp(-j\pi fT)$$

where as the transfer function of an ideal LPF is a rectangle. In this sense, the SNR produced by the ideal LPF will be lower than that produced by MF.

$$SNR_{MF}>SNR_{LPF}.$$

When a rectangular pulse of amplitude A and duration T is passed through an ideal LPF, its response is maximum (i.e..output is maximum) at t=T/2 for BT≥1. And the maximum value is-

$$(2A/\pi)\ Si\ (\pi BT).$$

It is therefore evident that the ideal LPF will satisfy the first condition of realization aspect of MF if the condition BT > 1 is met.

As the noise power at the o/p of ideal LPF of bandwidth B is BN_0, the SNR produced by a ideal LPF will be:

$$SNR_{LPF}=\frac{\left(2A/\pi\right)^2 Si^2(\pi BT)}{BN_0}$$

The SNR produced by a MF for a rectangular pulse as a input is:

$$SNR_{MF}=\frac{2E}{N_0}$$

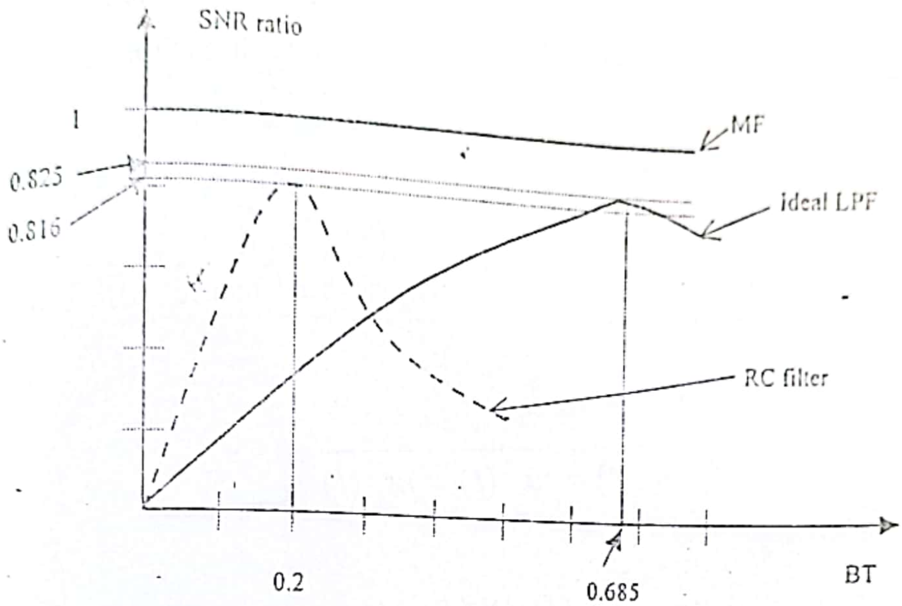
$$where \quad E=\int_0^{\infty} x^2(t)\,dt=\int_0^T A^2\,dt=A^2T$$

$$or \quad SNR_{MF}=2A^2T/N_0$$

Then the ratios of SNR produced by ideal LPF and MF will be :

$$\frac{SNR_{LPF}}{SNR_{MF}}=\frac{2}{\pi^2 BT}Si^2(\pi BT)$$

Scanned with CamScanner



The above plot indicates that for BT=0.625, the ideal LPF gives almost same performance as a MF. The degradation in performance is only 0.84 dB compared to a true matched filter. Moreover a simple RC with BT=0.2 also give almost same performance as a true MF and the degradation in terms of dB is 0.9 dB. The performance of ideal LPF and a real RC as a MF is comparable (the degradation is only 0.06dB).

Therefore a simple RC filter with BT=0.2 can substitute a MF and the degradation in performance is only 0.9 dB.

Scanned with CamScanner

Chapter-7

5. Noise Performance of Analog and Digital Communication Systems

As the most of the communication systems are band pass systems, we represent the noise as narrow band bandpass random signal and express mathematically as:

$$N(t) = n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t$$

where $n_c(t)$, $n_s(t)$ are low frequency random signals band limited to f_m with the respective powers:

$$\overline{n^2(t)} = \overline{n_c^2(t)} = \overline{n_s^2(t)}$$

and that

$$N_c(f) = N_s(f) = N_0$$

Performance evaluation method

The performance evaluation method employed here is to find the SNR at the input and the output of the demodulator and calculate the gain

gain $\gamma = SNR_0/SNR_i$

provided by particular demodulation / modulation scheme.

5.1 Amplitude Modulation

DSB with full carrier:

Input to the demodulator is the sum of signal and noise:

$$x_i(t) = [A_c + x_m(t)]\cos\omega_c t + n_i(t)$$

The signal power at the input is:

$$P_{si} = \frac{A_c^2}{2} + \frac{\overline{x_m^2(t)}}{2}$$

Scanned with CamScanner

And the input noise power

$$P_{Ni} = \overline{n_i^2(t)}$$

Therefore the input SNR will be:

$$SNR_i = \frac{A_c^2 + \overline{x_m^2(t)}}{2 \times \overline{n_i^2(t)}} \quad = \quad \frac{P_{si}}{P_{Ni}}$$

Now let us find P_{S0} and P_{N0} for the cases when the demodulator is:

(a) Envelop type ; and (Non-linear detection)

(b) Synchronous type (Linear detection)

a) Envelop detection:

The input signal can be written as:

$$\begin{aligned} x_i(t) &= [A_c + x_m(t)]\cos\omega_c t + n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t \\ &= [A_c + x_m(t) + n_c(t)]\cos\omega_c t + ns(t)\sin\omega_c t \\ &= V(t)\cos[\omega_c t + \varphi(t)] \end{aligned}$$

where,

$$V(t) = \sqrt{[AC + x_m(t) + n_c(t)]^2 + ns^2(t)}$$

$$\varphi(t) = -\arctan\left[\frac{n_s(t)}{A_c + x_m(t) + n_c(t)}\right]$$

Case A: Let us assume that noise is small, i.e.-

$$A_c + x_m(t) >> n_i(t)$$

In this case

$$[A_c + x_m(t) + n_c(t)]^2 >> n_s^2(t)$$

and therefore

$$V(t) \cong A_c + x_m(t) + n_c(t)$$

$$\varphi(t) \cong 0$$

Then the output of a ideal envelope detector is

$$V(t) = A_c + x_m(t) + n_c(t)$$

Scanned with CamScanner

where A_c is the DC component which will be filtered out by LPF. Finally the noise power and the signal power at the output of the demodulator will be:

$$P_{S0} = \overline{x_m^2(t)}$$

$$P_{N0} = \overline{n_c^2(t)}$$

The output signal to noise ratio will be:

$$SNR_o = \frac{P_{So}}{P_{No}} = \frac{\overline{x_m^2(t)}}{\overline{n_c^2(t)}}$$

The detection gain

$$\gamma = \frac{\overline{x_m^2(t)} \Big/ \overline{n_c^2(t)}}{A_c^2 + \overline{x_m^2(t)} \Big/ \overline{2n_i^2(t)}}$$

$$\frac{\overline{x_m^2(t)} \times \overline{2n_i^2(t)}}{\overline{n_c^2(t)}\left(A_c^2 + \overline{x_m^2(t)}\right)}$$

As $\overline{n_i^2(t)} = \overline{n_c^2(t)}$,

$$\gamma = \frac{2\overline{x_m^2(t)}}{A_c^2 + \overline{x_m^2(t)}}$$

It is clear from above expression that γ increases as A_c is decreased. Since for an envelop detector for distortionless transmission

$$A_c \geq |x_m(t)|_{\max}$$

the component

Scanned with CamScanner

$$\eta = \frac{\overline{x_m^2(t)}}{A_c^2 + \overline{x_m^2(t)}}$$

is the efficiency of DSB-full carrier AM. Therefore the system gain is

$$\gamma = 2\eta$$

For extreme case when modulation index is 100% and the modulating signal is sinusoidal waveform, the value of η is 1/3, therefore the maximum system gain provided by the DSB full carrier is:

$$\gamma = 2 \times \frac{1}{3} < 1$$

Therefore even for 100% modulation the detector gain is less than 1.

b) Synchronous detection:

In synchronous detection, the received AM signal is multiplied by cos $\omega_c t$ and then passed to LPF. Considering that the synchronization is ideal and that the transfer function of LPF is flat within the message bandwidth and zero outside it, we get

$$x_{dem}(t)\big|_{LPF} = x_i(t)\cos\omega_c t$$

$$= \frac{1}{2}x_m(t) + \frac{1}{2}n_c(t)$$

$\cos^2 \omega_c t$ आउँछ र यसलाई $\cos 2\omega_c t$ मा लगेर higher frequency + Ac term लाई filter गरेको हो ।

$$\text{Therefore} \quad P_{S0} = \frac{\overline{x_m^2(t)}}{4} \quad \text{and} \quad P_{N0} = \frac{\overline{n_c^2(t)}}{4}$$

$$\& \ SNR_0 = \overline{x_m^2(t)} \Big/ \overline{n_c^2(t)}$$

$$\text{and} \quad \gamma = \frac{2\overline{x_m^2(t)}}{A_c^2 + \overline{x_m^2(t)}}$$

The above expression shows that the gain is same as in case of envelope detection. It means the full carrier AM is the least efficient system, even the optimum detection

Scanned with CamScanner

ordinary detection method (envelop detection).

ii) DSB-SC

For DSB-SC the detection method is always synchronous. In this case the signal representation, the signal and noise powers at the input of the demodulator are:

$$x_i(t) = x_m(t)\cos\omega_c t + n_i(t)$$

$$\text{Therefore} \quad P_{Si} = \overline{\left[x_m(t)\cos\omega_c t\right]^2} = \overline{x_m^2(t)}/2$$

$$\text{and} \quad P_{Ni} = \overline{n_i^2(t)}$$

At the output of demodulator the useful signal component is $x_m(t)/2$ and therefore

$$Ps_0 = \overline{x_m^2(t)}/4 = Ps_i/2$$

The noise component of the input signal $n_i(t)$ is also multiplied by $\cos\omega_c t$ in the process of detection.

$$\begin{aligned} n_{det.}(t) &= n_i(t)\cos\omega_c t \\ &= n_c(t)\cos^2\omega_c t + n_s(t)\sin\omega_c t\cos\omega_c t \\ &= \frac{1}{2}\left[n_c(t) + n_c(t)\cos 2\omega_c t + n_s(t)\sin 2\omega_c t\right] \end{aligned}$$

In the above expression all other components except $n_c(t)/2$ shall be filtered out by LPF. Therefore the output noise power will be:

$$P_{N0} = \overline{n_c^2(t)}/4$$

The input /output SNR and the detection gain are :

Scanned with CamScanner

$$SNR_i = \frac{\overline{x_m^2(t)}}{2\overline{n_i^2(t)}}$$

$$SNR_0 = \frac{4\overline{x_m^2(t)}}{4\overline{n_c^2(t)}} = \frac{\overline{x_m^2(t)}}{\overline{n_c^2(t)}}$$

$$\gamma = {SNR_0}/{SNR_i} = 2, \text{ because } \left[\overline{n_c^2(t)} = \overline{n_i^2(t)}\right]$$

The gain provided by DSB-SC equaling to 2 comes from the fact that input signal consisted of two side bands and noise was also distributed in both side bands. After demodulation the noise bandwidth is reduced to half, and hence the noise power is reduced 4 times where as the signal power is reduced only 2 times.

SSB

5.2 Signal to Noise Ratio for Single Side-Band Modulation

The input signal to the demodulator (only synchronous) is: $x_i(t) = x_m(t)\cos\omega_c t + \hat{x}_m(t)\sin\omega_c t + n_i(t)$

$$x_i(t) = x_m(t)\cos\omega_c t \pm \hat{x}(t)\sin\omega_c t + n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t$$

The input signal , noise powers and the SNR are:

$$P_{Si} = \frac{\overline{x_m^2(t)}}{2} + \frac{\overline{\hat{x}_m^2(t)}}{2} = \overline{x_m^2(t)}$$

$$as\ \overline{x_m^2(t)} = \overline{x_m^2(\hat{t})},$$

$$P_{Ni} = \overline{n_i^2(t)}$$

$$SNR_i = \frac{\overline{x_m^2(t)}}{\overline{n_i^2(t)}}$$

After synchronous demodulation the output will be:

$$x_o(t) = x_m(t)\ or\ \hat{x}_m(t) + n_c(t)$$

The output signal, noise powers and the SNR are:

$$Ps_0 = \overline{x_m^2(t)}$$

$$P_{N0} = \overline{n_c^2(t)} = \overline{n_i^2(t)}$$

$$SNR_0 = \frac{\overline{x_m^2(t)}}{\overline{n_i^2(t)}}$$

Finally the detection gain is:

$$\gamma = \frac{SNR_0}{SNR_i} = 1$$

Scanned with CamScanner

It means the SSB does not provide noise improvement in noise performance.

It may seem that DSB-SC is superior to SSB-SC in terms of noise performance. But it is not true as in case of DSB-SC the input noise power is twice than that in case of SSB. Therefore performance of DSB-SC & SSB-SC are identical from noise improvement point view.

5.3 Threshold Effects in Detecting AM with Non- linear Detection Methods

In evaluating the noise performance of DSB-AM, we have assumed that the noise level is much more less than the signal level. Now let's consider the case of large noise when the noise level is comparable or even greater than the signal level. The input signal to the demodulator is:

$$x_i(t) = V(t) \quad \cos[\omega_c t + \varphi(t)]$$

$$where \quad V(t) = \sqrt{[A_c + x_m(t) + n_c(t)]^2 + n_s^2(t)}$$

$$\varphi(t) = -\arctan\left[\frac{n_s(t)}{A_c + x_m(t) + n_c(t)}\right]$$

In case of large noise :

$$n_i(t) >> [A_c + x_m(t)] \quad and$$

$$hence \quad n_C(t), n_S(t) >> [A_c + x_m(t)]$$

Then the envelope of the input signal can be expressed as: $[A_c + x_m(t) + n_c(t)]^2$ लाई expand गरेको ।

$$V(t) = \sqrt{[A_c + x_m(t)]^2 + 2[A_c + x_m(t)]n_c(t) + n_c^2(t) + n_s^2(t)}$$

$$V(t) \cong \sqrt{n_c^2(t) + n_s^2(t) + 2[A_c + x_m(t)] \cos\theta(t)}$$

$$\cong R(t)\sqrt{1 + \frac{2[A_c + x_m(t)]}{R(t)}\cos\theta(t)}$$

where,

$$R(t) = \sqrt{n_c^2(t) + n_s^2(t)}$$

$$\theta(t) = -\arctan\left[\frac{n_s(t)}{n_c(t)}\right]$$

Scanned with CamScanner

And since

$$R(t) >> [Ac + x(t)]$$

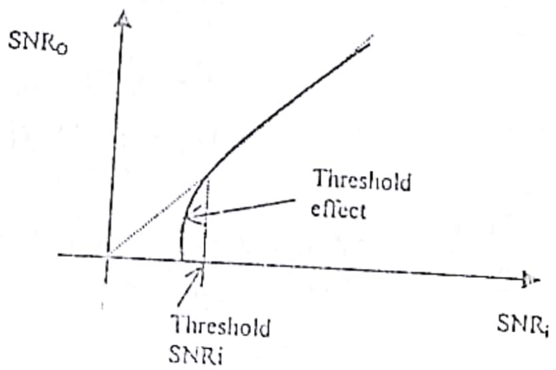
after expanding.

$$V(t) \cong R(t)\left[1 + \frac{Ac + x_m(t)}{R(t)} \cos\theta(t)\right]$$

$$= \left[R(t) + [Ac + x_m(t)]\cos\theta(t)\right]$$

· The above equation indicates that the envelope of the input signal (which will be the actually demodulated signal) contain no independent message signal $x_m(t)$. The message signal now is modulated by random noise component $\cos\theta(t)$. In other words the useful signal component is completely mutilated by the envelope detector. Such behavior of the envelope detector present so called threshold effect. The threshold effect deteriorate the output SNR more rapidly than the input SNR when the input SNR is below certain level, called the threshold level.

The threshold effect becomes pronounced when the carrier power to noise power approaches unity.



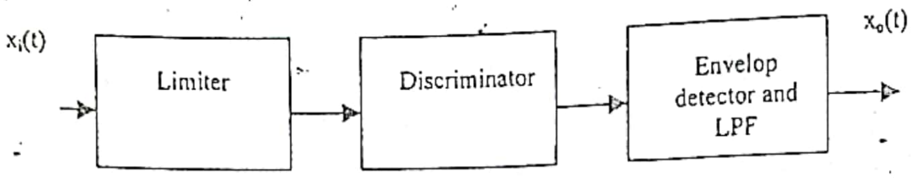
Threshold effect isn't occured in synchronous (linear) detection only in envelope (non-linear) detection

Such effect is not pronounced in synchronous detection (linear detection). Therefore for small noise application envelope detector offer same performance as synchronous detector. And for large noise application synchronous detector is superior to envelop detector. The reader is advised to do the analysis of the threshold effect for square law detector.

Scanned with CamScanner

Ans 5 ## Signal to Noise Ratio for Frequency Modulation
[Non-Coherent]

The functional block diagram of a standard FM demodulator is shown below.



The input signal to the demodulator is the sum of FM signal and noise:

$$x_i(t) = A_c \cos[\omega_c t + \phi(t)] + n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t$$

where,

$$\phi(t) = 2\pi k_f \int_{-\infty}^{t} x_m(\alpha)\,d\alpha$$

Therefore the input signal and noise powers and the SNR can be calculated as follows:

B- Bandwidth
N_0 - psd of noise

$$P_{si} = A_c^2/2 \,,$$

$$P_{ni} = \frac{\overline{n_c^2(t)}}{2} + \frac{\overline{n_s^2(t)}}{2} = \overline{n_c^2(t)} = 2BN_0$$

$$SNR_i = \frac{A_c^2}{4BN_0}$$

To calculate the output powers, let us first consider the case when the input noise is absent i.e.,

$$n_c(t)=n_s(t)=0$$

In this case the input to the demodulator is only the FM signal:

$$x_i(t)=A_c \cos[\omega_c t+\phi(t)]$$

The output of the discriminator is:

$$x_{dis}(t)=\frac{dx_i(t)}{dt}=A_c\left[\omega_c+\frac{d\phi(t)}{dt}\right]\sin[\omega_c t+\phi(t)]$$

Assuming that the envelope detector is ideal and that LPF remove components centered at ω_c and DC, at the output of the LPF we get:.

$$x_0(t)\big|_{LPF}=A_c\frac{d\phi(t)}{dt}=A_c 2\pi\, k_f x_m(t)$$

As $\rightarrow P=\frac{V^2}{R}$ and $R\rightarrow 1$.

The output signal power is therefore equal to:

$$P_{s0}=4\pi^2 A_c^2 k_f^2\,\overline{x_m^2(t)}$$

Now let us assume that input is only the un-modulated FM and noise i.e.:

$$\begin{aligned}x_i(t)&=A_c\cos\omega_c t+n_c(t)\cos\omega_c t+n_s(t)\sin\omega_c t\\&=[A_c+n_c(t)]\cos\omega_c t+n_s(t)\sin\omega_c t\\&=R_n(t)\cos[\omega_c t+\theta(t)]\end{aligned}$$

$\phi(t)$ isn't present so;

where,

$$R_n(t)=\sqrt{[Ac+n_c(t)]^2+n_s^2(t)}$$

$$\theta(t)=-\arctan\left[\frac{n_s(t)}{A_c+n_c(t)}\right]$$

$R_n(t)$ will be removed and high i/p SNR

Since there is the limiter before discriminator, we can assume that $R_n(t)$ will be removed. Assuming high input SNR, i.e.,

Scanned with CamScanner

$$A_c >> n_c(t)$$

we get:

$$\theta(t) \cong \tan^{-1}\left(\frac{n_s(t)}{A_c}\right) \cong \frac{n_s(t)}{A_c}$$

With these assumptions, the input to the discriminator will be:

$$x_{dis}(t) = \cos\left(\omega_c t + \frac{n_s(t)}{A_c}\right)$$

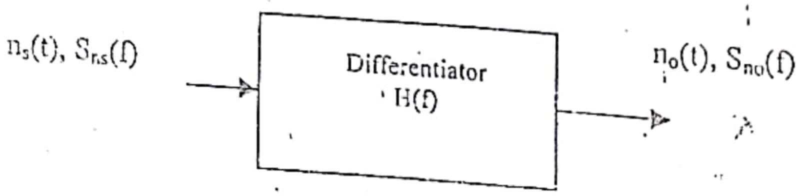
The output of the discriminator-

$$y_{dis}(t) = \frac{dx_i(t)}{dt} = \left[\omega_c + \frac{1}{A_c}\frac{dn_s(t)}{dt}\right]\sin(\cdots)$$

Finally, assuming ideal envelope detector and LPF, the output of the demodulator will be:

$$x_0(t)\big|_{LPF} = \frac{1}{A_c}\frac{dn_s(t)}{dt} = n_0(t)$$

But since there is no mathematical representation of $n_s(t)$ for differentiation, we now use spectral method, that is use the psdf approach.



The psdf of the signal at the output of the differentiator is:

Scanned with CamScanner

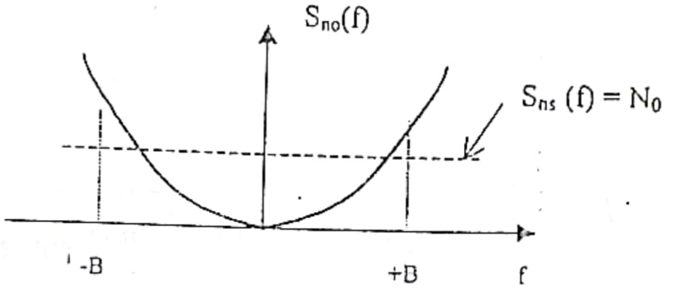
$$S_{no}(f) = S_{ns}(f)\left|H(f)\right|^2$$

The transfer function of a differentiator is:

Therefore- $H(f)=j2\pi f$

$$S_{no}(f) = N_0\left|j2\pi f\right|^2 = 4\pi^2 N_0 f^2$$

The above equation show that noise psd at the output of a FM discriminator is proportional to f^2.



It means that for low frequency components of the message signal, noise will be attenuated where as for high frequency components, the signal will be masked by noise. The effect of high noise level in high frequency range is compensated by using *pre-emphasis and de-emphasis* networks. Detailed discussion about these networks is presented in the course manual – Communication Systems I. The noise power within the message bandwidth can be calculated by taking integral of the output psdf within the message bandwidth:

$$P_{NO} = \frac{1}{A_c^2}\int_{-B}^{B} S_{no}(f)\,df = \frac{1}{A_c^2}4\pi^2 N_0 \int_{-B}^{B} f^2\,df = \frac{8\pi^2 N_0 B^3}{3A_c^2}$$

Then the output SNR will be:

$$P_{NO} = \frac{1}{A_c^2}\int_{-B}^{B} S_{no}(f)\,df$$

$$= 1/A_c^2 \int_{-B}^{B} 4\pi^2 \cdot N_0 f^2\,df$$

$$= 1/A_c^2 \cdot 4\pi^2 \cdot N_0 \cdot 2B^3/3$$

$$= \frac{8\pi^2 N_0 B^3}{3A_c^2}$$

Scanned with CamScanner

$$SNR_0 = \frac{4\pi^2 A_c^2 k_f^2 \overline{x_m^2(t)}}{8\pi^2 N_0 B^3} \times 3A_c^2 = \frac{3A_c^4 k_f^2 \overline{x_m^2(t)}}{2N_0 B^3}$$

And the detection gain

$$\gamma = SNR_0 / SNR_i = \frac{3A_c^4 k_f^2 \overline{x_m^2(t)}}{2N_0 B^3} \times \frac{4BN_0}{A_c^2} = \frac{6A_c^2 k_f^2 \overline{x_m^2(t)}}{B^2}$$

The parameter k_f can be expressed in terms of modulation index β by using the following relationships:

$$\beta = \frac{\Delta f}{B} = \frac{k_f |x_m(t)|_{\max}}{B}, \quad k_f = \frac{\beta B}{|x_m(t)|_{\max}}$$

The output SNR in terms of the modulation index can be expressed as:

$$\gamma = \frac{6A_c^2 \beta^2 B^2 \overline{x_m^2(t)}}{|x_m(t)|_{\max}^2 \times B^2} = \gamma = 6A_c^2 \beta^2 \overline{x_n^2(t)},$$

where

$$\overline{x_n^2(t)} = \frac{\overline{x_m^2(t)}}{|x_m(t)|^2_{\max}}$$

Since $6A_c^2 \overline{x_n^2(t)}$ is constant for given type of FM, it can be concluded that the gain provided by the FM is directly proportional to the square of the modulation index.

$\gamma \propto \beta^2$

Scanned with CamScanner

$$\gamma \propto \beta^2$$

This proportionality indicate that for narrow band FM, where β<1, the system does not provide any gain, but in fact the system performance is degraded.

But for wide band FM where β>> 1, the system provides high gain on noise performance . This is therefore, the FM system is preferred over AM for high quality (high fidelity) signal transmission (broadcasting).

$\gamma < 1$ system performance degraded.

$\gamma > 1$ system performance improved.

Scanned with CamScanner

5.5 Threshold Effect in FM

As the detection gain γ in the FM is proportional to β^2, for given SNR_i rise in β will increase γ or the output signal to noise ratio SNR_O. But with increase in β, the system bandwidth also increases at Carlson's rule rate:

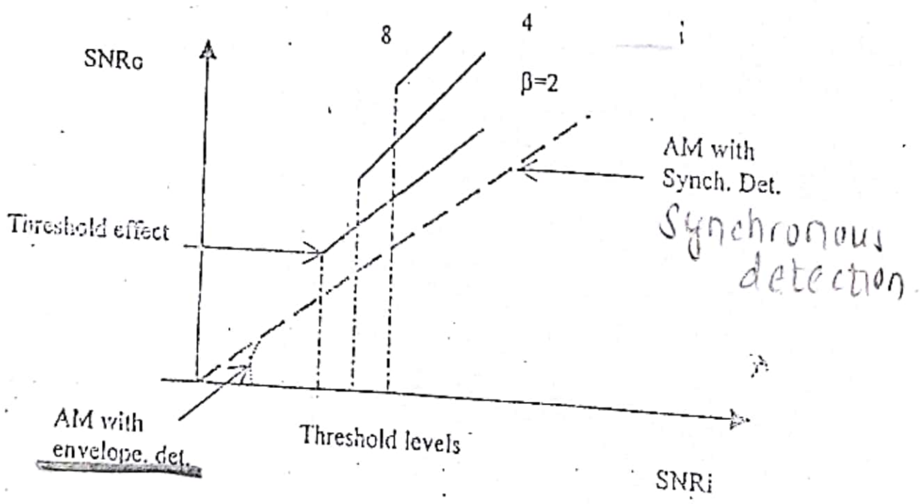
$$B_{FM} = 2(\beta+1)\, fm$$

Increase in bandwidth will subsequently increase the input noise power and decrease SNR_i.

Initially we have assumed that the carrier amplitude Ac is considerably greater than the noise $n_c(t)$. This assumption will no more remain valid if we increase β indiscriminately. That is at some point of increase in β, the input SNR will be so low that the signal, as in case of large noise AM, will be mutilated by noise and the output SNR will be reduced drastically causing the reception impossible.

In other words there exists the threshold effect that for SNR_i less than the threshold level FM system will no more detect the message signal.

General threshold level for FM is 10 dB, i.e. if $SNR_i \leq 10$dB, the system the system will not receive the signal.



Scanned with CamScanner

Before reaching the threshold level., the FM system produces noise clicks. As SNR_i is decreased further (experimentally by increasing noise level), individual clicks are produced, which is converted into crackling sound as the SNR_i is further decreased. And finally when the threshold level is reached, the receiver breaks (i.e., the received signal is completely mutilated by the noise).

As discussed earlier, the FM system, owing to the threshold effect, exhibit also the locking property i.e., if the input SNR_i is below threshold level the FM receiver locks to the noise.

If there is interference in the form of other stations with comparable carrier frequencies, the FM system will enhance one of the stations depending upon the intensities of the signals at the receiver.

If the level of interference is low compared to the desired signal then FM captures the desired signal and remains locked . But if the level of interference becomes greater than desired signal then the FM captures the interference. This effect of capturing the strong station is called capture effect.

The threshold effect in FM can be considerably reduced by employing the following techniques:

(i) use of PLL as demodulator- PLL can further reduce the threshold level by order of 5-7 dB,

(ii) use of pre-emphases, de-emphasis networks - Improvement in system performance is about 6 dB.

Scanned with CamScanner

5.6 Comparison of Linear Modulation (AM, DSB-SC, SSB) and Non–linear (FM,PM) Modulation Systems

AM (DSB-FC, DSB-SC, SSB) , FM (Narrow and Wide Band)

Any modulation system can be compared with another in terms of certain fixed criteria. The basic criteria for comparison could be :

(i) band width efficiency

(ii) power efficiency

(iii) system complexity.

(i) In terms of bandwidth efficiency ,the SSB-SC shows the best performance as in this case the channel bandwidth is equal to the message bandwidth. Therefore the SSB-SC is widely used in application where bandwidth is the major constraint (microwave links , satellite communications , point to point communication). For the applications requiring near to DC transmission (such as TV broadcasting), VSB-SC technique is used.

(ii) In terms of power efficiency DSB –AM is the least efficient system where as the FM has highest efficiency due to its high level of noise immunity (at the cost of larger channel bandwidth). Therefore in power critical applications such as space vehicle communication, FM broadcasting etc. FM is preferred over other modulation techniques.

(iii) System complexity plays vital role when the number of receivers is tremendous. AS DSB-AM reception is least complicated, it is used for commercial ratio broadcasting. In terms of system complexity SSB-SC is slightly more complex than the DSB-SC. But since DSB-SC is less effective in terms of bandwidth, it is almost not used in practice.

Scanned with CamScanner

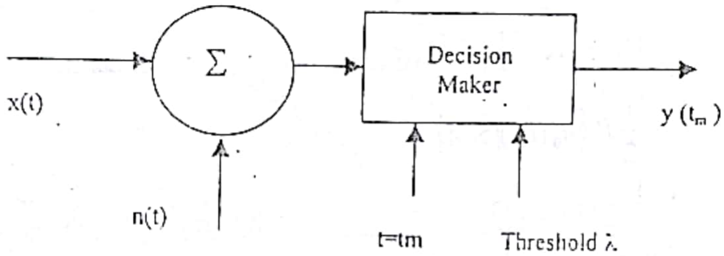
Error probability.

5.7 Noise in Modulated Digital System

Performance of digital communication system is evaluated in terms of error probability i.e. the probability that the decoder, in presence of noise, wrongly decodes received bit.

The error probability will solely depend upon the input SNR (input to decoder) and the complexity of decoder.

Let us first start with base-band data communication where no carrier modulation is employed (PAM, PDM, PPM, PCM). The most frequently used is PAM, therefore we will start with the analysis of this system.



The signal at the output of the decision maker at $t = t_m$ is :

$$y(t_m) = A_m + n_0(t_m),$$

$$\text{where} \quad A_m = +A \quad \textit{if } B_m = 1 \quad (\textit{bit } 1 \text{ transmitted})$$

$$A_m = -A \quad \textit{if } B_m = 0 \quad (\textit{bit } 0 \text{ transmitted})$$

In absence of noise, the decision making device compares the $y(t\ )$ against the threshold level λ (which is "0v"). If $y(t_m)>0$ then it is decided that bit 1 was transmitted and if $y(t_m)<0$ then decision is made in favor of 0 bit. The presence of noise and its strength may introduce error in decision making. The error will occur when decision is made in favor of bit 1 when actually bit 0 was transmitted and in favor of bit 0 when actually bit 1 was transmitted. In other words the condition for occurrence of the error can be stated as:

Condition for occurrence of error can be stated as:

Scanned with CamScanner

$Y(t_m) < 0$ when $A_m = +A$ (i.e. bit 1 transmitted)

$Y(t_m) > 0$ when $A_m = -A$ (i.e. bit 0 transmitted)

As the bit sequence is independent, the total probability of error is sum of each error probabilities:

$$P_e = P[y(t_m) > 0/b_m = 0] \times P(b_m = 0) + P[y(t_m) < 0/b_m = 1] \times P(b_m = 1)$$

Since 1 and 0 bits are equi-probable

$$P_e = \frac{1}{2} P\left[y(t_m) > 0/b_m = 0\right] + \frac{1}{2} P\left[y(t_m) < 0/b_m = 1\right]$$

$$= \frac{1}{2}\{P\left[n_0(t_m) > A\right] + P\left[n_0(t_m) < -A\right]\}$$

$$or \quad p_e = \frac{1}{2} P\left\{|n_0(t_m)| > A\right\}$$

Since n(t) is assumed to be zero mean Gaussian process with variance N_0, the noise $n_0(t)$ will also be a Gaussian process with variance N_0.

Therefore,

$$p_e = \frac{1}{2}\int_{|x|>A}^{\infty} \frac{1}{\sqrt{2\pi N_0}} \exp(-x^2/2N_0)\,dx$$

$$= \int_{A}^{\infty} \frac{1}{\sqrt{2\pi N_0}} \exp(-x^2/2N_0)\,dx$$

By changing the variable

$Z = x/\sqrt{N_0} \Rightarrow x = Z\sqrt{N_0}$, we get

$$p_e = \int_{\frac{A}{\sqrt{N_0}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-Z^2/2)\,dZ = erfc\left(\frac{A}{\sqrt{N_0}}\right)$$

Scanned with CamScanner

where,

$$erfc(u) = \int_u^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-Z^2/2)\, dZ$$

is called the complimentary error function.

The error probability can be calculated by using error function erf(u) described as:

$$erf(u) = \int_0^u \frac{1}{\sqrt{2\pi}} \exp(-Z^2/2)\, dZ$$

This error function is available in tabular form for different values of 'u'.

It can be seen from the expression for error probability that P_e can be minimized by maximizing the power ratio A^2/N_0.

With known P_e, we can estimate the number of errorly decoded bit in a stream of N bits:

$$N_e = p_e \times N$$

e.g. for u=3.3 erf(u)=0.999998 and therefore erfc(u)= 1 - 0.999998 $= 2\times10^{-6} = P_e$

Error probability for M-ary system

For M-ary data transmission the error probability, P_e can be evaluated first by deriving P_e for M=4 and then generalizing it for M.

For M=4, the decoding algorithm is

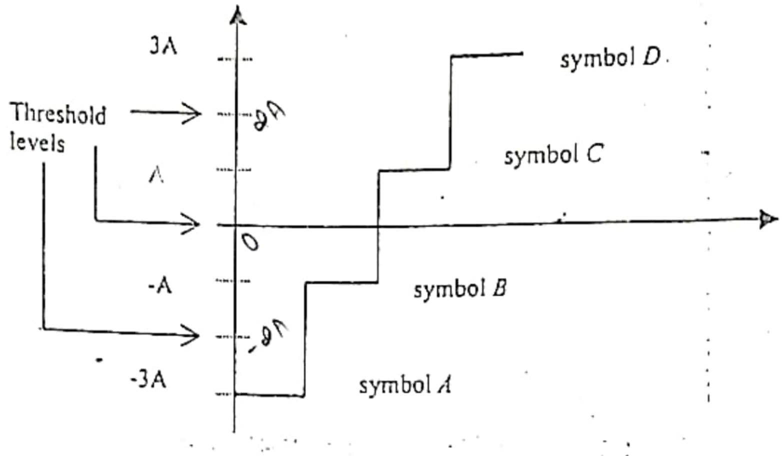
$y(t_m) > 2A$ - symbol D

$0 < y(t_m) \leq 2A$ - symbol C

$-2A < y(t_m) \leq 0$ - symbol B

$y(t_m) \leq -2A$ - symbol A

Scanned with CamScanner



Then the probability of error is:

$$p_e = p\,(\text{error}/D\ \text{sent}) \times p\,(D\ \text{sent}) + \ldots\ldots + \ldots\ldots + p\,(\text{error}/A\ \text{sent}) \times p\,(A\ \text{sent}).$$

As

$$p\,(D\ \text{sent}) = p\,(C\ \text{sent}) = p\,(B\ \text{sent}) = p\,(A\ \text{sent}) = 1/4, \text{ we get:}$$

D C

$$P_e = \frac{1}{4}\begin{bmatrix} p\left(y(t_m) \le 2A / A_m = 3A\right) + p\left(y(t_m) > 2A \text{ or } \le 0 / A_m = A\right) + \\ + p\left(y(t_m) > 0 \text{ or } \le -2A / A_m = -A\right) + p\left(y(t_m) > -2A / A_m = -3A\right) \end{bmatrix}$$

B A

Repeating the calculations similar to that made for binary PAM for each of the above cases, we get:

$$P_e = \frac{1}{4}(6)\,\operatorname{erfc}\left(A \big/ \sqrt{N_0} \right) \quad (\text{for } M = 4)$$

3

And generalizing it for M, we get:

Scanned with CamScanner

$$p_{e\,M-ary} = \frac{2\,(M-1)}{M}\, erfc\left(A \Big/ \sqrt{N_0} \right) \simeq$$

$$p_{e\,M-ary\,(min)} = \frac{2\,(M-1)}{M}\, erfc\left(\sqrt{A \Big/ N_0} \right)_{max} \simeq$$

Comparison between binary & M-ary scheme:

The following table summarize the merits /demerits of M-ary and binary systems:

	Binary	M-ary
Data rate	r_b, bits/sec	r_s, symbols/sec (= r_b bits/sec in terms of bits)
Bandwidth	r_b, Hz	r_b / k, Hz →bandwidth reduction $(M = 2^k)$
p_e	erfc $(A/\sqrt{N_0})$	2 (M-1) / M erfc $(A/\sqrt{N_0})$→high error probability.
Transmitter power required for given p_e.	Less	High
Equipment complexity	Less	High

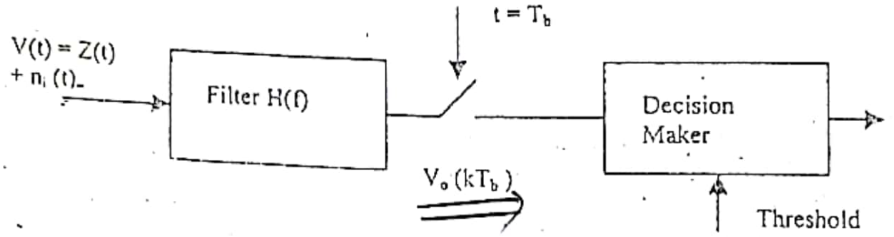
1. For given p_e, M-ary requires $M^3/3 \log_2 M$ times more signal power than binary.
2. The p_e for M-ary is higher than for binary by 2(M-1) / M times.
3. M-ary system is more complex as the number of threshold levels required for demodulation is (M-1)

For equal r_b, the bandwidth requirement in M-ary system is 'k' times smaller than binary.

Scanned with CamScanner

5.8 Effect of Noise in Modulated Digital Communication Systems

Basic types of digital carrier modulation are ASK, FSK & PSK. The receiver structure for the digital carrier modulation is shown below:



Output of modulator Z(t) depends upon the k-th bit b_k. At instantaneous moment z(t) is the one bit shifted version of input bit. In other words, the present status of z(t) is the response to $(k-1)^{th}$ input bit.

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

The signal representation of $s_1(t)$ & $s_2(t)$ depend upon type of carrier modulation used.

	ASK	FSK	PSK
$s_1(t), 0 \le t \le T_b$ $\{b_k=0\}$	0	$A\cos(\omega_c-\Delta\omega)t$ or $A\sin(\omega_c-\Delta\omega)t$	$-A\cos\omega_c t$ or $-A\sin\omega_c t$
$s_2(t), 0 \le t \le T_b$ $\{b_k=1\}$	$A\cos\omega_c t$ or $A\sin\omega_c t$	$A\cos(\omega_c+\Delta\omega)t$ or $A\sin(\omega_c+\Delta\omega)t$	$A\cos\omega_c t$ or $A\sin\omega_c t$

Then, in general,

$$V(t)=\begin{cases} s_1[t-(k-1)T_b]+n(t) \\ s_2[t-(k-1)T_b]+n(t) \end{cases} ,for\ (k-1)T_b \le t \le T_b$$

Scanned with CamScanner

Since s_1, s_2 have finite duration of T_b, i.e. for $t \notin [0, T_b]$; $s_1 = s_2 = 0$, therefore they have finite energy:

$$E_1 = \int_0^{T_b} [s_1(t)]^2 \, dt < \infty$$

$$E_2 = \int_0^{T_b} [s_2(t)]^2 \, dt < \infty$$

With the following assumptions:

1. Input bit streams $\{b_k\}$ are equiprobable i.e. $s_1(t)$ & $s_2(t)$ are equiprobable and statistically independent.
2. The channel noise is zero mean Gaussian with psdf $N(f)$.
3. ISI introduced by $H_c(f)$, $H_T(f)$ and $H_R(f)$ are negligible.

the output of the receiver filter at $t = kT_b$ will be:

$$V_0(kT_b) = s_0(kT_b) + n_0(kT_b)$$

$$s_0(kT_b) = \int_{-\infty}^{kT_b} Z(\varepsilon)\, h(kT_b - \varepsilon)\, d\varepsilon$$

$$= \int_{(k-1)T_b}^{kT_b} Z(\varepsilon) h(kT_b - \varepsilon) d\varepsilon + ISI \ terms \rightarrow \approx 0$$

$$\text{or} \quad s_o(kT_b) \cong \int_{(k-1)T_b}^{kT_b} Z(\varepsilon)\, h(kT_b - \varepsilon)\, d\varepsilon$$

Substituting the value of $Z(t)$ in above

$$s_0(kT_b) = \begin{cases} \int_0^{T_b} s_1(\varepsilon)\, h(T_b - \varepsilon)\, d\varepsilon = s_{01}(kT_b) & for\ b_k = 0 \\ \int_0^{T_b} s_2(\varepsilon)\, h(T_b - \varepsilon)\, d\varepsilon = s_{02}(kT_b) & for\ b_k = 1 \end{cases}$$

The output noise can be expressed as:

Scanned with CamScanner

$$n_0(kT_b) = \int_{-\infty}^{kT_b} n(\varepsilon)\, h(kT_b - \varepsilon)\, d\varepsilon$$

and the noise pdf:

$$N_0(x) = \frac{1}{\sqrt{2\pi N_0}} \exp\left(-x^2 / 2N_0\right) \qquad -\infty < x < \infty$$

The output of the receiver filter $V_0(kT_b)$ is compared with a threshold λ, and assuming $s_{01}(kT_b) < s_{02}(KT_b)$, the following decision making algorithm is employed:

$$\text{if } V_0(kT_b) > \lambda \quad \text{then } \hat{b}_k = 1 \text{ and}$$
$$\text{if } V_0(kT_b) < \lambda \quad \text{then } \hat{b}_k = 0$$

$$\text{where,} \quad \lambda = \frac{s_{01} + s_{02}}{2}$$

is the mid point between maximum and minimum values of $V_0(kT_b)$. Then the error probability can be expressed as:

$$p_e = p\{V_0(kT_b) \geq \lambda / b_k = 0 \text{ or } V_0(kT_b) \leq \lambda / b_k = 1\}$$
$$= \frac{1}{2} p\{V_0(kT_b) \geq \lambda / b_k = 0\} + \frac{1}{2} p\{V_0(kT_b) \leq \lambda / b_k = 1\}$$

Now if k^{th} transmitted bit were 0, i.e. $b_k=0$ then $V_0(kT_b) = s_{01}+n_0$ or $n_0 = V_0-s_{01}$, and similarly if $b_k=1, V_0(KT_b)=s_{02}+n_0$ or $n_0=V_0-s_{02}$. Therefore the above equation for p_e can be expressed as:

Scanned with CamScanner

$$p_e = \frac{1}{2}\int_{\lambda}^{\infty} \frac{1}{\sqrt{2\pi N_0}} \exp\left(\frac{-(V_0 - s_{01})^2}{2N_0}\right) dV_0$$

$$+ \frac{1}{2}\int_{-\infty}^{\lambda} \frac{1}{\sqrt{2\pi N_0}} \exp\left(\frac{-(V_0 - s_{02})^2}{2N_0}\right) dV_0$$

$$= \int_{\lambda}^{\infty} \frac{1}{\sqrt{2\pi N_0}} \exp\left(\frac{-(V_0 - s_{01})^2}{2N_0}\right) dV_0$$

Changing the variable

$$Z = \left(V_0 - s_{01}\right) \Big/ \sqrt{N_0}$$

weget

$$p_e = \int_{\frac{(s_{02} - s_{01})}{2\sqrt{N_0}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-Z^2/2\right) dZ$$

The error probability p_e becomes smaller as

$$\frac{s_{02}(T_b) - s_{01}(T_b)}{\sqrt{N_0}}$$

becomes larger. Therefore an optimum receiver of binary modulated signal should maximize the ratio

$$r = \frac{s_{02} - s_{01}}{\sqrt{N_0}}$$

or it's square (r^2). By maximizing r^2, the need for $s_{01} < s_{02}$ can be eliminated and finally:

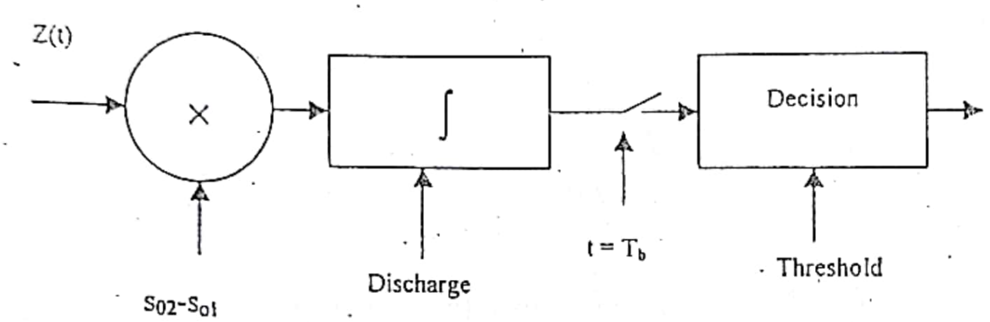
error probability for digital carrier modulation.

$$\boxed{p_e = erfc\left(\frac{r}{2}\right) \quad and \quad p_{e\min} = erfc\left(\frac{r_{max}}{2}\right)}$$

Now if we use matched filter for optimum detection then the impulse response of MF would be :

$$h(t) = s_{02}(T_b - t) - s_{01}(T_b - t)$$

And the implementation (correlation receiver) would be



Individual Cases

Binary ASK

AA. Coherent detection

$$\gamma_{max} = \sqrt{\frac{2(s_{02} - s_{01})}{N_0}}$$

In binary ASK, the signals can be represented as:

$$s_1(t)=0,\ s_2(t)=A\cos\omega_c t\ ,$$

Therefore,

$$s_2(t)-s_1(t)=A\cos\omega_c t$$

The output signals for each case is:

$$s_{01}(kT_b) = \int_0^{T_b} s_1(t)\left[s_2(t)-s_1(t)\right]dt = 0$$

$$s_{02}(kT_b) = \int_0^{T_b} s_2(t)\left[s_2(t)-s_1(t)\right]dt$$

$$= \int_0^{T_b} A^2\cdot\cos^2\omega_c t\ dt = \frac{A^2}{2}T_b$$

The maximum value of signal to noise ratio γ will be:

Scanned with CamScanner

$$r_{max} = A^2T_b/N_0$$

Therefore the error probability will be:

$$p_e = erfc\left(\sqrt{\frac{A^2T_b}{4N_0}}\right)$$

As the signal s_1 and s_2 are equi-probable, the average signal power will be:

$$S_{av}=A^2/4$$

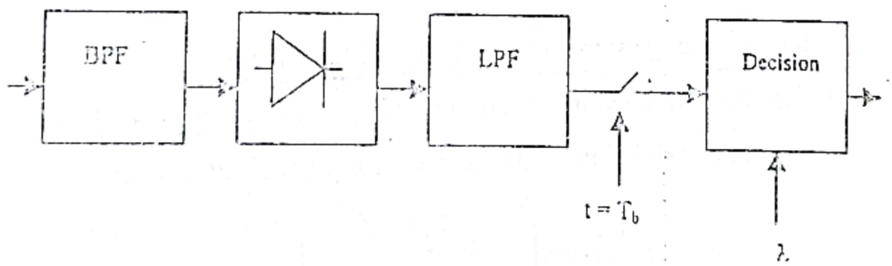
In terms of average signal power, the expression for error probability can be re-written as:

$$p_e = erfc\left(\sqrt{\frac{S_{av}T_b}{N_0}}\right) = erfc\left(\sqrt{\frac{E_{av}}{N_0}}\right)$$

Here $E_{av}=S_{av}.T_b$ is the average signal energy per information bit.

BB. Non- coherent detection

The structure of the non-coherent detector of ASK can be simulated by the following functional diagram:



The detailed analysis will result that the error probability in this case can be approximately expressed as:

$$p_e \approx \frac{1}{2}\exp\left(\frac{-A^2}{8N_0}\right) \quad for\ A^2 >> N_0$$

Scanned with CamScanner

The error probability p_e for non-coherent ASK is always higher than in case of coherent detection for similar conditions (signal power, r_b and noise pdf). But the equipment complexity is lower in non-coherent detection method of ASK.

PSK. Binary PSK

The demodulation method used in PSK is always coherent. The input signals can be expressed as:

$$s_1(t) = -A\cos\omega_c t,\ s_2(t) = A\cos\omega_c t$$

The respective output signals therefore will be:

$$s_{01}(kT_b) = -A^2T_b$$
$$s_{02}(kT_b) = A^2T_b$$

The avaerage signal power and the maximum signal to noise ratio will be:

$$S_{av} = A^2/2,\quad r_{max} = 2A\sqrt{\frac{T_b}{N_0}}$$

Finally the error probability in terms of average energy per bit can be approximately expressed as:

$$p_e = erfc\left(\sqrt{2E_{av}/N_0}\right)$$

ASK require double the average signal power than PSK for same p_e.
DPSK is the special case of PSK in which non-coherent detection method can be used. For non-coherent demodulation of DPSK, the error probability will be:

$$p_e = \frac{1}{2}\exp\left(-A^2T_b/2N_0\right)$$

DPSK requires about 1dB more power than PSK for same P_e, but because of the simplicity in demodulation, DPSK is preferred over PSK.

Scanned with CamScanner

Binary FSK

In binary FSK, the input signal representation is:

$$s_1(t)=A\cos(\omega_c-\Delta\omega)t,\quad s_2(t)=A\cos(\omega_c+\Delta\omega)t$$

For the case of coherent demodulation the error probability can be expressed as:

$$p_e = erfc\left(\sqrt{\frac{1.2E_{av}}{N_0}}\right)$$

For non - coherent demodulation the error probability can be expressed as:

$$p_e \approx \frac{1}{2}\exp\left(-A^2/4N_0\right)$$

Scanned with CamScanner

5.9 Comparison of Digital Modulation Systems

The following table summarize the comparison of different digital carrier modulation systems in terms of data rate, bandwidth requirements, required signal to noise ratio for given error probability and the equipment complexity.

System	Band-width	Required SNR (dB) for p_e=10^{-4}	Equipment Complexity / uses
Coherent ASK	$\approx 2r_b$	14.45	Moderate Uses: generally not used
Non-coherent ASK	$\approx 2r_b$	18.33	Minor Uses: generally not used
Coherent FSK	$>2r_b\approx$	10.6	Major Uses: generally not used
Non-coherent FSK	$>2r_b$	15.33	Minor Uses: low speed data
Coherent PSK	$\approx 2r_b$	8.45	Major Uses: high-speed data, best power performance.
DPSK	$\approx 2r_b$	9.30	Moderate Uses: most commonly used for medium speed data

Scanned with CamScanner

5.10 Application of Modems for Data Transmission over Telephone Lines

Modem (*Mo*dulator-*dem*odulator) is an electronic device used to transmit and receive digital data for transmission over public telephone lines.

Different modulation system is used in modem depending upon data speed. Generally for a speed up to 1800 bps NC-FSK is used. C-PSK is used for data rate of about 4800 bps. For 9600 bps,QAM (8-phase) is used and for higher data rates PSK or VSB is generally used.

Modes of operation of Modems:

1. Simplex – Data transmission is only in one direction and no signaling path is available from receiver to transmitter. Therefore there is no possibility of retransmission and error correction. This mode has limited use.
2. Half –duplex – Reverse channel is available but in turn, not simultaneously. Only one channel (bi- directional) is required for this mode of operation. The transmission speed is reduced because of the need to wait for the transmission or reception.
3. Full duplex – In this mode data can be transmitted and received in both directions at same time. It requires two independent channels. This is the most commonly used mode of operation.

Interconnections to communication circuits:

1. Dedicated line- For short distance, Public Telephone Network (PTN) cable pairs dedicated to given modem system could be the effective way of data communication over two fixed terminals. Data do not go through switching. Network of PTN therefore modem complexity could be simple.
2. Dial-up lines - For more flexible operation dial-up modems are used. It uses standard PTN telephone (voice grade) link and the data signal pass through PTN switching network. Since data transfer takes place in voice grade channel, the

modem for this application requires complex circuitry to compensate the degradation in performance due to limit in the bandwidth.

Method of coupling with the communication cable

(a) Hardwired –Modems are permanently connected to PTN line. This arrangement reduces the mobility, as the modems can not be easily shifted to other working space if required.

Acoustically coupled- Special acoustic foam pads are used to couple modem with Telephone set. Therefore this arrangement provides grater mobility for the data source.

Scanned with CamScanner

6.0 Introduction to Coding Theory

The purpose of coding in digital communication is twofold. The first use of the coding is for binary representation of alpha-numeric characters. Standard codes for representing alphabets, numbers and special sign should be used for compatibility. These days modern computers can accept various code standards, convert them to the code standard used in the computer and process. The basic types of codes are:

(a) Baudot code- Baudot code the first coding sequence introduced by J.M.E. Baudot, the pioneer of Telegraph. It is a 5 bit code and maximum combination is 32. Additional shift code is incorporated to indicate the letter case. Therefore total number of combinations available is 64.

(b) BCD code- Binary Coded Decimal (BCD) is used for numerical processing only. Extension of BCD is a 7 bit alphanumeric code.

(c) ASCII code- American standard code for information Interchange (ASCII) is more universal code that uses standard BCD progression. The last four bits of the code word represent the characters (numbers, signs, alphabets) and the 3 MSB bits identify whether the last 4 LSB bits are representing number, alphabet, character or special signs. Although ASCII is a 7 bit code, the eight bit is added as parity bit.

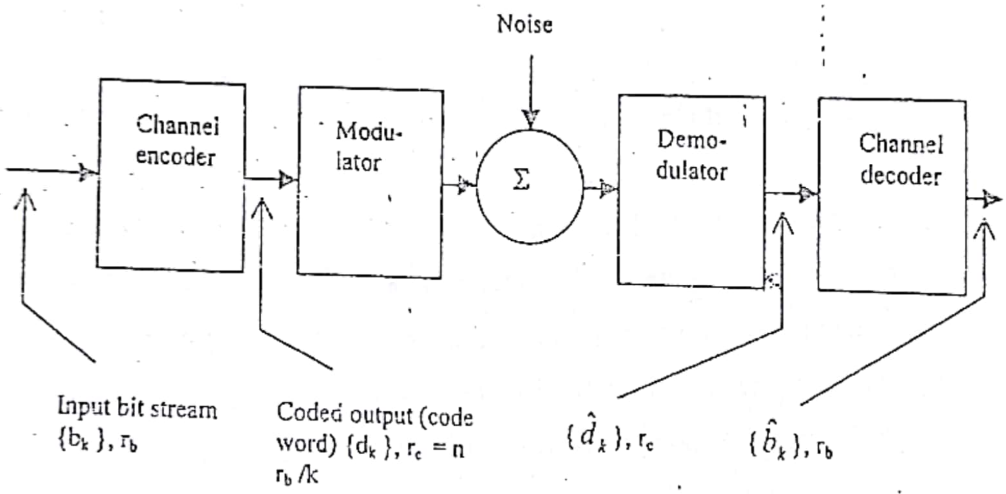
(d) EBCDIC- Extended binary coded decimal interchange code. is a 8 bit code very similar to ASCII, no parity bit is available.

6.1 Coding for error detection and correction

Error probability in digital communication is a function of input SNR and the data rate r_b. In practice maximum signal power as well as the channel bandwidth are restricted to some upper limit by government regulations and the channel noise is fixed by its distribution. Therefore for particular situation the only way to arrive at acceptable error probability p_e is to introduce error control coding technique before data transmission. This type of error control coding is performed by channel encoder at transmitting end and decoding is performed by channel decoder at the receiving end.

The key factor in error control coding is the calculated use of redundancy. The source encoder removes redundancy in the source whereas the channel encoder put back just enough redundancy to arrive at acceptable BEP. Therefore the functions of source encoder and channel encoder are quite different and therefore can be treated in isolation.

The model:



The input bit stream is converted into blocks of k message bits. The channel encoder replaces the k bit message block 'D' with n bit code word 'C', where first k bits are the usual message bits and remaining (n-k) bits are check bits derived from k message bits as per predetermined rule.

The receiver compares the (n-k) check bits derived from received k message bits as per predetermined rule and if it is matched then accepts the message bits and if not sends ARQ (Automatic request for repeat) signal or NAK (do not

Scanned with CamScanner

acknowledge) signal to transmitter for repeat. If the code has error correction capabilities, then the receiver corrects the error without sending ARQ signal.

The followings are the general rules for designing the error control coding:

i) It is possible to detect and correct errors by adding extra bits, which are not bonafide message bits.

ii) It is not possible to correct and detect all errors unless the redundancy is very high (Theoretically impossible).

iii) Addition of extra bits reduces the effective bit rate (Efficiency = r_b/r_c).

Types of errors:

i) If the channel noise is a random Gaussian process then the error caused by this noise will also be random and the coding designed for dealing this type of error is called random error correcting code.

ii) In some cases the nature of the noise is high amplitude bursts, followed by long quite intervals. Such noise burst are induced by man made activities like ignition, switching transients or natural like lightning discharges. Coding system designed for such noise is called burst error correcting codes.

Methods of error control:

a) Error detection and correction: In this coding, the receiver detects the error and then corrects it to the extent possible. This type of error correction technique is called forward acting error correction.

Example: For each 1 or 0 the triplets 111 or 000 are transmitted (i.e. 200% redundancy). The received triplets may be:

Rx	000	001	010	100	011	101	110	111
Decision	0	0	0	0	1	1	1	1

Scanned with CamScanner

i.e. even if the received triplet does not exactly 000 or 111; decision can be made that 1 was transmitted if the received triplets contains more than one 1's and 0 if the received triplets contains were than one 0's.

This system correctly detects and corrects the error if not more than one error is induced in each triplet.

b) Error detection only.

This coding permits error detection but has no provision for correction. In case if error is detected ARQ is sent for retransmission.

The second method yield overall low p_e but reduces the transmission time and requires reverse channel for ARQ.

The first method requires sufficient redundancy to be able to correct the errors.

Scanned with CamScanner

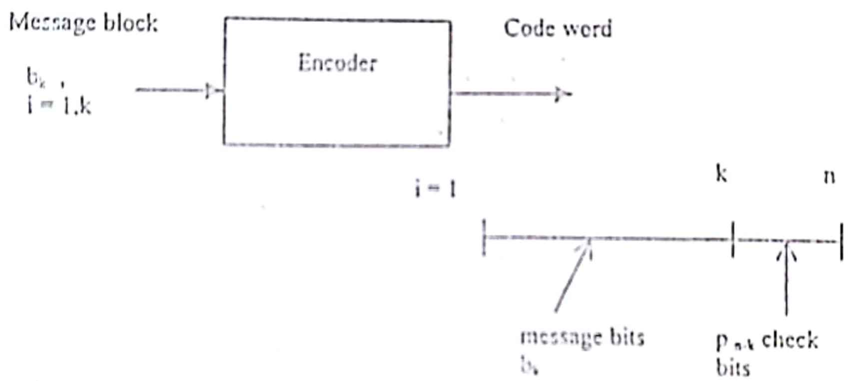
6.2 Block Codes

The basic types of error detection and correction codes are: Block codes & Convolutional codes.

In block codes, block of k message bits are followed by (n-k) check bits of a n bit code word.

In convolutional codes message bits and check bits are continuously interleaved and have memory effect i.e. the check bits are used to verify the correctness of the received information bits not only in the block immediately preceding them but in other blocks as well.

If the linear combination of k code words of 2^k code word vector is also a code word then such code is called linear block code. Moreover if the message bits appear at the beginning /end of a code word then it is called systematic code.



Simple parity check linear block codes:

In this type of code a matrix of (n x m) is derived from the data stream in the form of (m-1) x (n-1) matrix. The last row and column of the matrix are check bits for row, column and check for check bits of row and column.

Scanned with CamScanner

$$\begin{vmatrix} data\ bits_{(m-1)\times(n-1)} & check\ bit\ for\ rows \\ check\ bits\ for\ columns & check\ for\ check\ bits \end{vmatrix}$$

By comparing the parity of received bits in each row and column with that received check bits, single errors can be corrected provided there is only one error in each row or column. But double errors in rows or columns can not be corrected due to ambiguity.

Scanned with CamScanner

6.3 Basic Terminology used in Coding Theory

1. Hamming weight of a code (code vector) - The Hamming Weight (HW) is defined as the number of non-zero components in the code. Example:

   c=101100 , HW=3

   c=111011 , HW=5.

2. Hamming distance (HD)-HD is the number by which a code word differs from other code word of the same code vector. Example:

   C={001, 000} HD=1

   C={000, 111} HD=3

   C={010, 111} HD=2

3. Minimum distance (MD)- MD is the smallest distance between any pair of code words in the code vector. Example:

   C={001, 110, 111, 000}, MD=1

   1 2 3 0

Basic Theorems:

1. The MD of a linear block code is equal to the minimum weight of any non-all-zero word in the code.
2. A linear block code with a MD = d_{min} can correct upto $(d_{min}-1)/2$ errors in each code word.
3. For detecting t error bits or fewer it is necessary & sufficient that HD≥ (t+1)
4. For detecting and correcting t or fewer error : HD≥ (2t+1)
5. For t_2 error detection and t1 error correction ($t_2>t_1$): HD≥ (t_1+t_2+1)

Example: consider a 3-bit code word.

The combinations of codes with HD 2 are 000, 101, 110 & 011

If these code words are used then any single error in any of the code words can easily be detected. i.e. if for 101 the first bit has error i.e. it is received as 001, then as this

Scanned with CamScanner

combination now is not a valid code, it is then decided that there is an error (error detection), but you can not correct it as error in 2nd bit of code $01^e1=00^e1$ will also yield the same received sequence as error in 1st bit of 101.

Now if the code word with HD=3 (000 & 111) are used then we can detect and correct single error in each code word.

Scanned with CamScanner

6.4 Linear Block Codes

In general, group linear block codes can be described in a matrix form. The message bit stream is divided into blocks of k message bits D(i=1,k). Now a code word C of block n is derived from D in the fashion :

$$c_1=d_1,\ c_2=d_2,\ldots\ldots,c_k=d_k$$

And

$$c_{k+1}=p_{11}d_1+p_{21}d_2+p_{31}d_3+\ldots..p_{k,1}d_k$$

$$c_{k+2}=p_{12}d_1+p_{22}d_2+p_{32}d_3+\ldots..p_{k,2}d_k$$

....

$$c_n=p_{1,n-k}d_1+p_{2,n-k}d_2+p_{3,n-k}d_3+\ldots..p_{k,n-k}d_k$$

Where p_{ij} are coefficients of coding and can have 1 or 0 value. All addition and multiplication are modulo-2 operation so that C is also in the form of 1 & 0.

Then in matrix form we can rewrite above equation as

$$[c_1, c_2 \ldots\ldots\ldots c_k] = [d_1, d_2 \ldots\ldots\ldots d_k] \begin{bmatrix} 100\ldots..0 \mid p_{11}, p_{12}\cdots p_{1,n-k} \\ 010\ldots..0 \mid p_{21}, p2\ldots. p_{2,n-k} \\ \ldots\ldots\ldots \\ 0\ldots\ldots\ldots 1 \mid p_{k,1}, p_{k,2}\cdots p_{k,n-k} \end{bmatrix}$$

Or C = DG

where G is called the generator matrix and has the form

$$G=[I_k \mid P_{k,n-k}]_{k\times n}$$

The I_k matrix is called identity or unity matrix and gives 1st k bits of the code word as the message bits and $P_{k,n-k}$ is the arbitrary k×(n-k) matrix that completely defines the code word.

The task of the designer is to specify such a P matrix that is easy to implement, has ability of detect & correct errors, high rate efficiency etc.

Scanned with CamScanner

Syndrome Calculation:

Once the code words are received by the receiver, there should be some mechanism to verify whether the received code word is generated by the predetermined generator matrix G. The parity check matrix H can be used for this purpose. The parity check matrix H is defined as:

$$H = \begin{bmatrix} p_{11}, p_{21} \ldots\ldots p_{k,1} & | & \\ p_{12}, p_{22} \ldots\ldots p_{k,2} & | & I_{n-k} \\ \ldots\ldots\ldots\ldots & & \\ p_{1,n-k}, \ldots\ldots\ldots p_{k,n-k} & | & \end{bmatrix} = \left[P^T \mid I_{n-k} \right]_{n-k \times n}$$

Where P^T is the Transpose of matrix P. Now if C is generated by G=[I_k | P], then it is necessary and sufficient that $CH^T=0$

Now let us consider that the code vector C generated by G=[I_k | P] with check matrix H was received corrupted by noise. The received code vector is then

$$R = C + E,$$

where E is the error vector.

The receiver does not know among R which is C and which is E. The decoding is done by multiplying vector (modulo 2) R with H^T and the result is called Syndrome S.

$$S = RH^T$$

$$S=[C + E]\, H^T=CH^T+EH^T=EH^T \qquad (\text{since } CH^T=0)$$

Therefore the error syndrome S is all zero if there is no error at all. The combination of S can exactly define the error position in received code vector.

Example: Let

$$G= \begin{array}{c} \quad I_4 \qquad\qquad P \\ \begin{bmatrix} 1 & 0 & 0 & 0 & | & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & | & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & | & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 1 & 1 \end{bmatrix}_{7\times4} \end{array}, k=4$$

Scanned with CamScanner

Then H for it is

$$H = \left[\begin{array}{cccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\qquad P^T \qquad\qquad\qquad I_3$$

For a message block D=(1011) , the code vector C will be:

$$C = DG = 1011|001$$

Now if we assume that the code vector is decoded without any error, the syndrome will be:

$$S = CH^T = 000$$

Now suppose the third bit of C has error i.e. $R = 1001001 = 1011001 \oplus 0010000$

C + E

In this case the transpose of parity check matrix will be:

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

The error syndrome for erroneous reception will be

$$S = EH^T = [0010000]\, H^T = 101$$

It means single error in the third bit position of the message is the syndrome that matches the third row of H^T.

In general a single error in i^{th} bit of n bit code word lead to a syndrome identical to i^{th} row of H^T.

Scanned with CamScanner

A single error correcting code require $d_{min} = 3$ and such code is called single error correction Hamming code.

Linear block codes technique requires large storage and processing time. Therefore are complex and in most cases impractical.

Scanned with CamScanner

6.5 Binary Cyclic Codes

Binary cyclic code is the special class of linear block codes where encoding and syndrome calculations can be easily implemented using simple shift registers. A (n,k) linear code C is called cyclic code if cyclic shifts of the code are also code vectors of C. Thus in polynomial form if $C_1(X)$ is a code vector given by

$$C_1(X)=C_0+C_1X+C_2X^2+C_3X^3+\ldots\ldots C_{n-1}X^{n-1} \quad \text{where } C_i=0/1$$

Then the shifted vector

$$C_2(x)=C_{n-1}+C_0X+C_1X^2+C_2X^3+\ldots\ldots C_{n-2}X^{n-1}$$

is also a code vector. Here all addition and multiplication are modulo-2 operation.

Properties of cyclic codes:

1. In an (n,k) cyclic code there exists one and only one generator polynomial $g(x)=1+g_1x+g_2x^2+g_3x^3+\ldots\ldots+g_{n-k}x^{n-k}$
2. The g(x) of an (n,k) cyclic code is a factor of (x^n+1), i.e.$(x^n+1)=g(x)h(x)$
3. The code vector can be expressed as

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

   Where m(x) is the message polynomial.

$$m(x)=m_0+m_1x+m_2x^2+\ldots\ldots m_{k-1}x^{k-1}$$

Example:

Let $g(x)=1+x+x^3$ for a polynomial of (7,4) cyclic code. Find code vector for message vector {1010}, n=7,k=4

Solution:

$$m(x)=m_0+m_1x+m_2x^2+m_3x^3$$
$$=1+0.x+1.x^2+0.x^3$$
$$=1+x^2$$

therefore code vector will be:

$$C(x) = m(x).g(x)$$
$$= (1+x^2)\,(1+x+x^3)$$
$$=1+x^2+x+x^3+x^3+x^5$$

Scanned with CamScanner

As

$$x^3+x^3=(1+1)\,x^3=0.x^3=0,$$

we get

$$C(x)=1+x+x^2+x^5$$

Therefore the code vector for the given generator matrix and data input will be:

$$C(x)=1+x+x^2+0.x^3+0.x^4+x^5+0.x^6$$

$$=1\ 1\ \ 1\ 0\ 0\ 1\ 0$$

The codes generated by simple cyclic generation method are not systematic.

In systematic form the first n-k bits are check bits and remaining k bits are message bits.

$$C(x) \rightarrow \{\underbrace{r_0, r_1, r_2, \ldots\ldots r_{n-k-1}}_{\text{n-k parity check bits}},\ \underbrace{m_0, m_1, \ldots\ldots m_{k-1}}_{\text{k-message bits}}\}$$

where

$$r(x)=r_0+r_1x+r_2x^2+\ldots\ldots+r_{n-k-1}\,x^{n-k-1}$$

is the parity check polynomial for the message polynomial m(x).

The parity check polynomial r(x) is the remainder from dividing x^{n-k} m(x) by g(x)

$$x^{n-k}m(x)=q(x).g(x)+r(x)$$

where q(x) is quotient and r(x) is the remainder of the division.

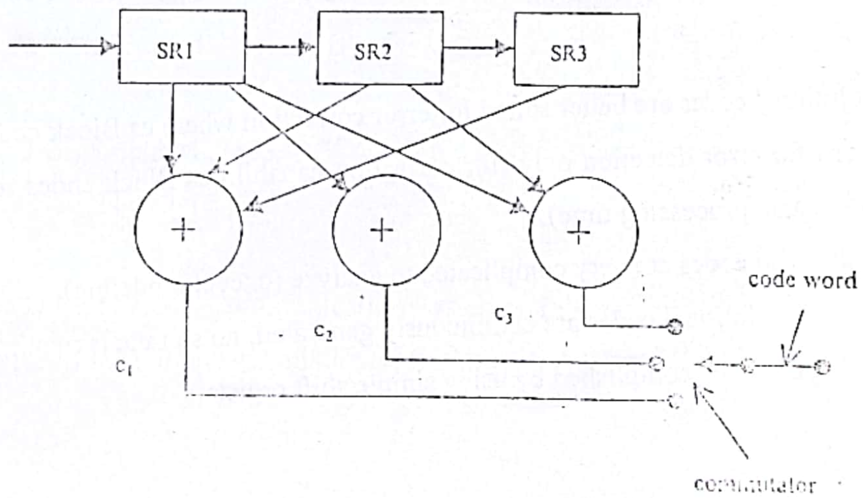
Scanned with CamScanner

6.6 Convolutional Codes

Unlike block codes where block of k message bits are followed by (n-k) check bits and the code pattern is dependent only upon the present k message blocks to be coded, in convolutional codes the message bits and check bits are continuously interleaved and the code words for i -th message bit depend not only upon the i- th input bit but also upon previous (N-1)th bit. The product n x N is called constraint length. N, in general, is the number of shift registers used in the coder and n is the number of code bits per k input message bit. The data rate efficiency of the convolutional code is defined as R = k/n.

Example of a simple convolutional code:

A convolutional code block of size (n,k)=(3,1) having constraint length = 9 (nN) can be constructed as follows:



The output of each of the modulo-2 adders are defined by the following equation (SR being the output of the corresponding shift registers):

$$c_1 = SR1 \oplus SR2 \oplus SR_3$$

$$c_2 = SR1$$

$$c_3 = SR1 \oplus SR2$$

Operation

1. Initially suppose that all shift registers are clear.
2. First message bit enters RS1 and during the bit period of this message bit the commutator samples all three outputs.
3. Therefore a single input bit is converted into 3-bit code word.
4. When the next message bit enter the RS1, its contents are shifted to RS2 and during the bit duration (T_b) of second message bit, the commututor samples C_1, C_2, C_3 and generate another code block.
5. In this way the status of the first bit influence the three blocks of code words containing of a total of 9 bits (including the first set of code word generated immediately after the first bit). Therefore the influence level (constraint length) is equal to n×N, where n-number of commutator & N is the number of shift registers.

Remarks:

(i) Convolutional codes are better suited for error correction where as Block codes are mainly used for error detection only (for correction capabilities Block codes require large storage and processing time).

(ii) Convolutional codes are very complicated to analyze (except modeling).

(iii) As the convolutional codes are continuously generated, no storage is required.

(iv) Encoding can be accomplished by using simple shift registers.

Scanned with CamScanner

REFERENCES

1. Haykin S.,(1994), "An Introduction to Analog and Digital Communications", John Wiley & Sons.

2. Couch II L.W., (1998), "Modern Communication Systems- Principles and Applications", Prentice Hall of India.

3. Lathi B.P., (1985), "Communication System", Wiley Eastern Limited.

4. Roddy D., Coolen J., (1989), "Electronic Communications", Prentice Hall of India.

5. Baskakov S.I., (1986), "Signals and Circuits", Mir Publishers.

6. Das J., Mullick S.K.,Chatarjee P.K., (1992), " Principles of Digital Communication – Signal Representation, Detection and Information Coding", Wiley Eastern Limited.

7. Proakis J.G., Salehi M., (1994), "Communication Systems Engineering", Prentice Hall, New Jeresy.

8. Sharma D.K., (1999), "Communication Systems –I : Course Manual", Institute of Engineering, Tribhuvan University, Kathmandu.

Scanned with CamScanner