DIGITAL COMMUNICATION TECHNIQUES (3-1-0)

Module-I (8 Hours) Sampling Theorem, Low Pass Signal, Band Pass Signal, Signal Reconstruction, Practical Difficulties, The Treachery of Aliasing, The Anti-aliasing Filter, Application of Sampling Theorem, PAM, PWM and PPM Signal Generation and Detection.

Module-II (12 Hours) Pulse Code Modulation: Quantization of Signals, Uniform and Non-Uniform Quantization, The Compander, The encoder, Transmission Bandwidth and output SNR, Digital multiplexer, Synchronizing and Signaling, Differential PCM, Delta Modulation, Adaptive Delta Modulation, Output SNR, Comparison with PCM. Noise in PCM and DM: Calculation of Quantization Noise Power, Output Signal Power, and the Thermal Noise Power, Output SNR of PCM using different modulation techniques. Output SNR of DM.

Module-III (12 Hours) Principles of Digital Data Transmission: A Digital Communication System, Line Coding-Various line codes, Polar Signaling, ON-OFF Signaling, Bipolar Signaling, Pulse Shaping: Nyquist Criterion for zero ISI, Scrambling, Regenerative Repeater- Preamplifier, Equalizer, Eye diagram, Timing Extraction, Timing Jitter, A Base-band Signal Receiver, Peak Signal to RMS Noise output voltage ratio, The Optimum Filter, White Noise, The Matched Filter- Probability of Error of the Matched Filter, Coherent Reception.

Module-IV (10 Hours) Digital Modulation Technique: Generation, Transmission, Reception, Spectrum and Geometrical Representation in the signal space of BPSK, BFSK, Differentially- Encoded PSK, QPSK, Minimum Shifting Keying (MSK), M-ary PSK, M-ary FSK, Use of Signal Space to calculate probability of Error for BPSK and BFSK.

Text Books:

1. Principles of Communication Systems by Taub & Schiling, 2nd Edition, Tata Mc Graw Hill. Selected portion from Chapter 5, 6, 11, and 12.

2. Communication System by Simon Haykin, 4th Edition, John Wiley & Sons, Inc.

3. Modern Digital and Analogue Communication Systems by B.P.Lathi, 3rd Edition, Oxford University Press. Selected Portion from Ch.2, 3, 6, 7, 13, and 14.

Reference Books:

1. Communication System, Analogue and Digital, R.P.Singh & S.D. Sapre, TMH.

2. Digital and Analogue Communication System, Leon W.Couch-II, 6th Edition, Pearson.

<u>Module I</u> **<u>8 hours</u>**

Sampling Theorem:

All pulse modulation scheme undergoes sampling process. Sampling of low frequency(LF) signal is achieved using a pulse train. Sampling process provides samples of the message signal. Sampling rate of sampling process must be proper to get original signal back. Sampling theorem defines the sampling rate of sampling process in order to recover the message signal. The solution to sampling rate was provided by Shannon.

Basically there are two types of message signal, such as-

(i) Low-pass (baseband) signal,
(ii) Band-pass (passband) signal.

- **<u>Sampling rate for Low-Pass Signal:--</u>**
  Sampling theorem states that if g(t) being a lowpass signal of finite energy and is band limited to W Hz, then the signal can be completely described by and recovered from its sampled values taken at a rate of 2W samples or more per second.

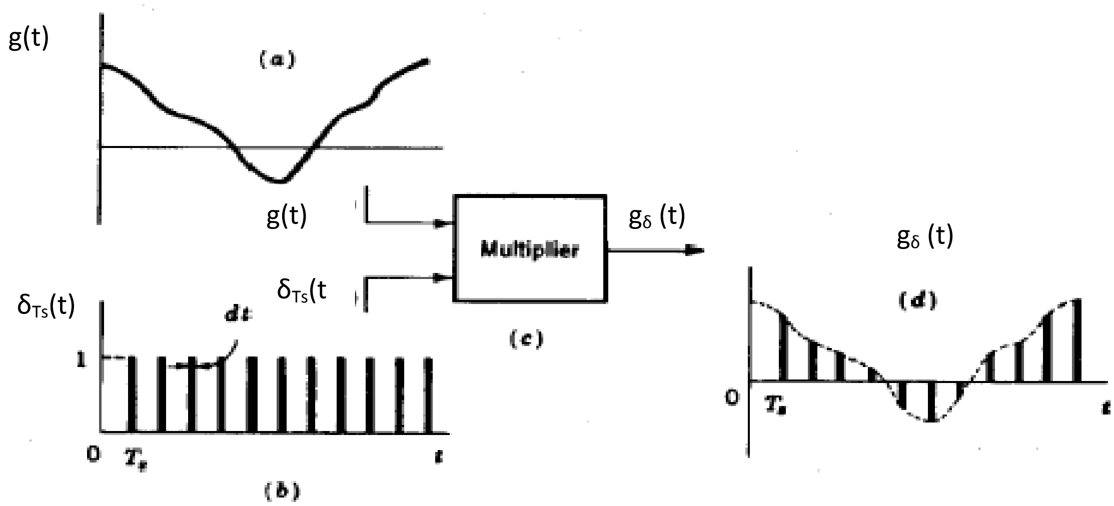


Fig. 1.1 Representation of sampling process.

Thus the time period of sampled signal must be, $T_s \leq 1/(2W)$.

Considering a signal g(t) as shown be a low pass signal where fourier transform of g(t),

$$G(f) = 0, \quad \text{for } f > W$$
$$= \text{finite}, \quad \text{for } f \leq W.$$

Ideally, we can get sampled values of g(t) at a regular time interval of time T_s if we multiply a train of pulses δ_{Ts} to g(t) as shown.

The product signal [$g_\delta(t)$], ie, the sampled values can be written as,

$$g_\delta(t) = g(t)\ \ \delta_{Ts}(t) \quad (1.1a)$$

or, $$g_\delta(t) = g(t) \quad (1.1b)$$

If we denote $g(nT_s)$ as the weights of low pass signal at sampled interval, then we can write,

$$g_\delta(t) = \quad (1.2)$$

Taking the fourier transform of equation 1.2, we get

$$G_r(f) = G(f)$$

Or, $$Gr(f) = 1/T_s$$

or, $$Gr(f) = 1/T_s \quad (1.3)$$

Now, we can draw graphically the frequency components of both the original signal and the sampled signal as follows,

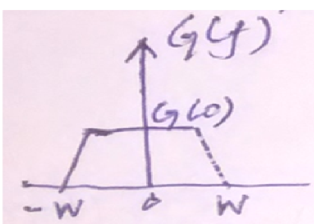


Fig. 1.1a Spectrum of original signal.

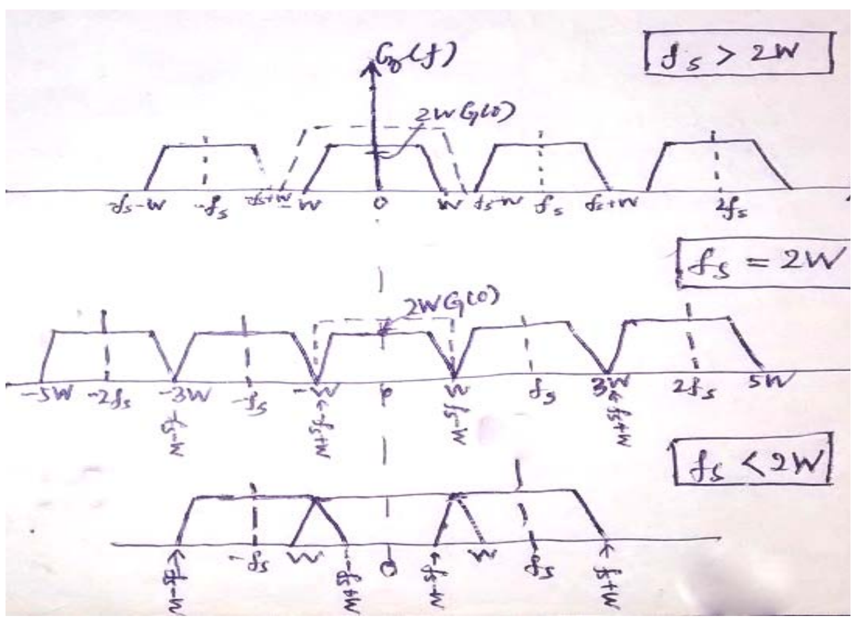


Fig. 1.1b Spectrum of Sampled signal.

<u>Note:-</u> The process of uniformly sampling a baseband signal in time domain results in a periodic spectrum in the frequency domain with a period, $f_s=1/T_s$, where T_s is the sampling period in time domain and $\leq 1/2W$.

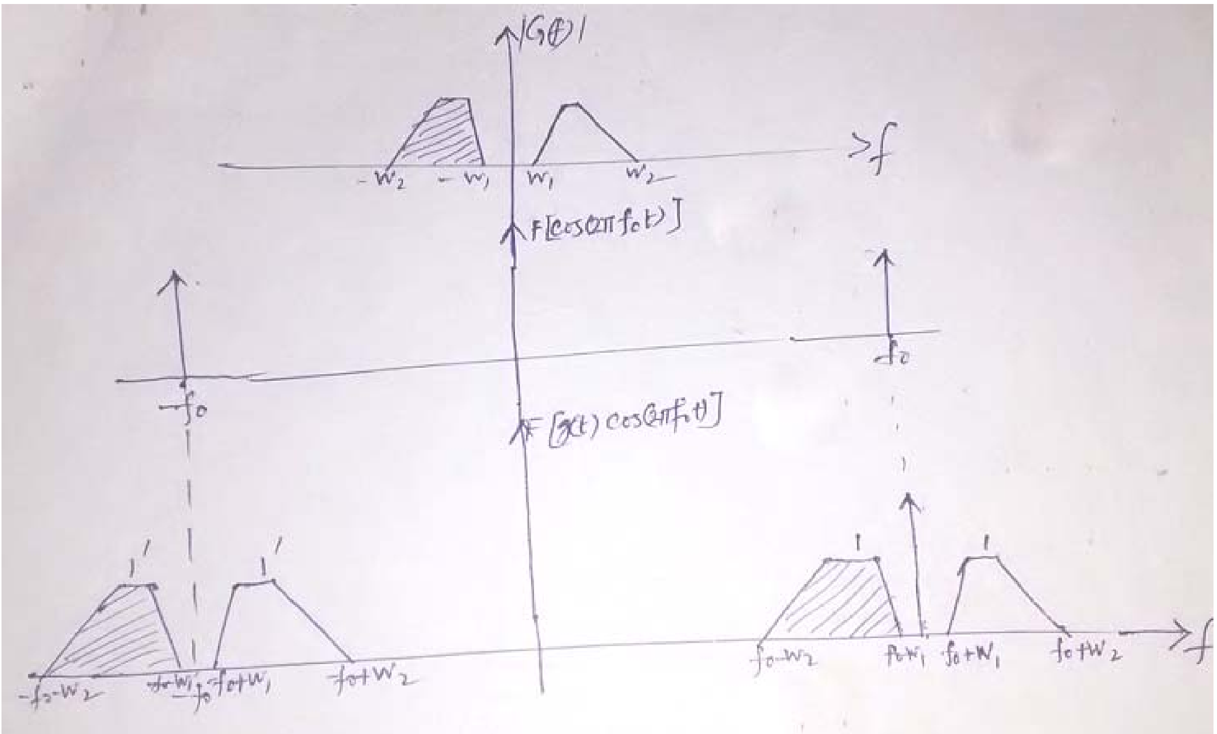


Fig. 1.1c Spectrum of baseband, carrier and modulated carrier signal.

- **<u>Sampling of Bandpass Signal:</u>**

  If the spectral range of a signal extends from 10 MHz to 10.1 MHz, the signal may ne recovered from samples taken at a frequency f_s=2{10.1 - 10} = 0.2 MHz.
  The sampling signal $\delta_{Ts}(t)$ is periodic. So,

$$\delta_{Ts} = dt/ds + 2.dt/ds(\cos 2\pi t/Ts + \cos(2.2\ \pi t/T_s) + \cos(3.2\ \pi t/T_s) + \ldots\ldots)$$
$$= f_s dt + 2f_s dt(\cos 2\pi f_{st} + \cos(2\pi.2f_s t) + \cos 2\pi.3f_s t + \ldots\ldots)$$

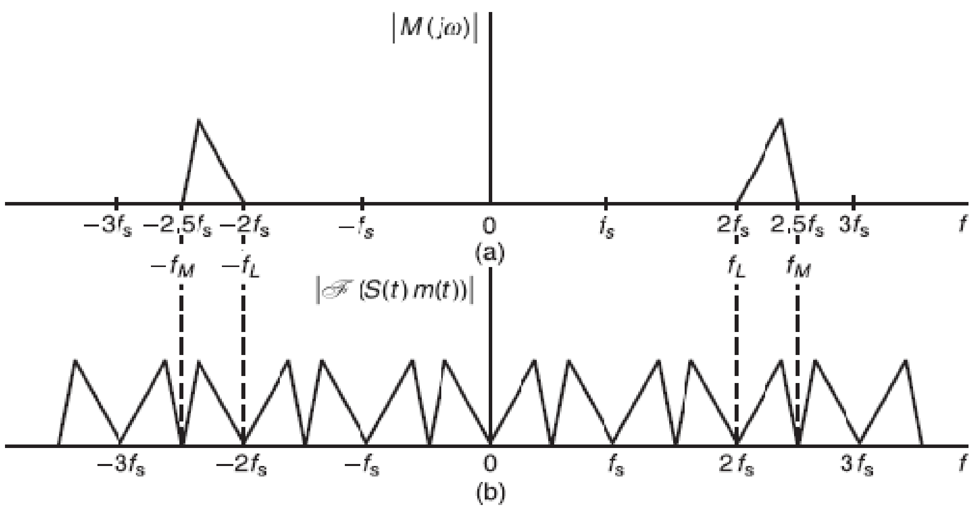


Fig. 1.2 Spectrum of bandpass and its sampled version signal

In fig. 1.2 the spectrum of g(t) extends over the first half of the frequency interval between harmonics of the sampling frequency, that is, from $2f_s$ to $2.5f_s$. As a result, there is no spectrum overlap, and signal recovery is possible. It may also be seen from the figure that if the spectral range of g(t) extends over the second half of the interval from 2.5 f_s to $3f_s$, there would similarly be no overlap. Suppose, however that the spectrum of g(t) were confined neither to the first half nor to the second half of the interval between sampling frequency harmonics. In such a case, there would be overlap between the spectrum patterns, and signal recovery would not be possible. Hence the minimum sampling frequency allowable is $f_s=2(f_M - f_L)$ provided that either f_M or f_L is a harmonic of f_s.

If neither f_M nor f_L is a harmonic of f_S, a more general analysis is required. In fig 1.3a, we have reproduced the spectral pattern of fig 1.2. The positive frequency part and negative frequency part of the spectrum are called PS and NS respectively. Let us , for simplicity, consider separately PS and NS and the manner in which they are shifted due to the sampling and let us consider initially what constraints must be imposed so that we cause no overlay over, say, PS. The product of g(t) and the dc component of the sampling waveform leaves PS unmoved, which will be considered to reproduce the original signal. If we select the minimum value of $f_s=2(f_m - f_L) = 2B$, then the shifted Ps patterns will not overlap

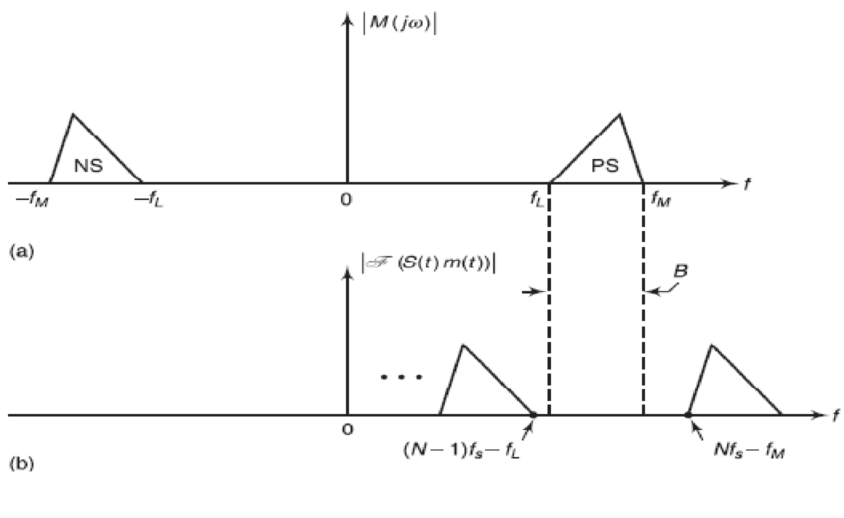


Fig. 1.3 (a) Spectrum of the bandpass signal (b) Spectrum of NS shifted by the (N-1)st and the Nth harmonic of the sampling waveform.

PS. The NS will also generate a series of shifted patterns to the left and to the right. The left shiftings can not cause an overlap with unmoved PS. However, the right shifting of NS might cause an overlap and these right shifting of NS are the only possible source of such overlap over PS. Shown in fig. 1.3b, are the right shifted patterns of NS due to the (N-1)th and Nth harmonics of the sampling waveform. It is clear that to avoid overlap it is necessary that,

$$(N-1)f_s - f_L \le f_L \qquad (1.4a)$$

and,

$$Nf_s - f_M \ge f_M \qquad (1.4b)$$

So that, with $B = f_M - f_L$, we have

$$(N - 1)f_s \leq 2(f_M - B) \quad (1.4c)$$

and, $$Nf_s \geq 2f_M \quad (1.4d)$$

If we let $k = f_M/B$, eqn. (1.4c) & (1.4d) become

$$f_S \leq 2B(K-1)/(N-1) \quad (1.4e)$$

and, $$fs \leq 2B(K/N) \quad (1.4f)$$

In which $k \geq N$, since $fs \geq 2B$. Eqn. (1.4e) and (1.4f) establish the constraint which must be observed to avoid an overlap on PS. It is clear from the symmetry of the initial spectrum and the symmetry of the shiftings required that this same constraint assumes the there will be no overlap on NS. Eqn.(1.4e) and (1.4f) has been plotted in fig. 1.4 for several values of N.
Let us take a case where f_L=2.5 KHz and f_M=3.5 KHz. So, B=1 KHz and $K=f_M/ =$ 3.5. On the plot of fig. 1.4 line for k=3.5 has been erected vertically. For this value of k if f_s = 2B, then overlapping occurs. If f_s is increased in the range of 3.5 to 5 KHz, then no overlap occurs corresponding to N=2. If f_s is 7B or more then no overlap occurs.

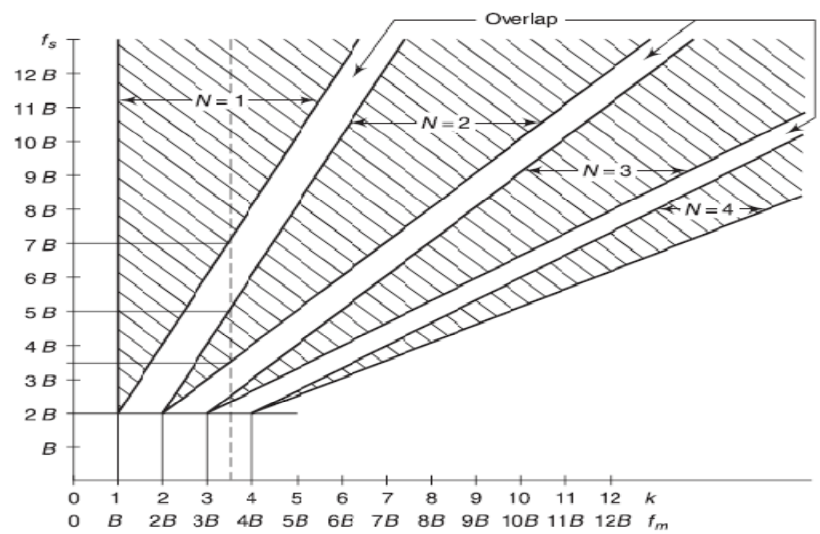


Fig. 1.4 The shaded region are the regions where the constraints eqn. (1.4e) and (1.4f) are satisfied.

From this discussion, we can write bandpass sampling theorem as follows---A bandpass signal with highest frequency f_H and bandwidth B, can be recovered from its samples through bandpasss filtering by sampling it with frequency f_s=2 f_H/k, where k is the largest integer not exceeding f_H/B. All frequencies higher than f_s but below $2f_H$(lower limit from low pass sampling theorem) may or may not be useful for bandpass sampling depending on overlap of shifted spectrums.

m(t) – low pass signal band limits to f_M.
s(t) – impulse train

$s(t) = \Delta t/Ts + 2.\ \Delta t/Ts(\cos 2\pi t/Ts + \cos(2.2\ \pi t/T_s) + \cos(3.2\ \pi t/T_s) + \ldots\ldots)$
$= \Delta t.fs + 2.\ \Delta t.fs(\cos 2\pi.fs.t + \cos(2.2\ \pi.fs.t) + \cos(3.2\ \pi.fs.t) + \ldots\ldots)$ (1.4g)

Product of m(t) and s(t) si the sampled m(t) ie, $m_s(t)$
$m_s(t) = m(t).s(t)$
$= \Delta t/Ts.m(t) + \Delta t/Ts[2.m(t)\cos 2\pi.fs.t + 2.m(t).\cos(2\pi.2.fs.t) + 2.m(t).\cos(2\pi.3.fs.t) + \ldots\ldots]$ (1.4h)

By using a low pass filter(ideal) with cut-off frequency at f_m then $\Delta t/Ts.m(t)$ will be passed so the m(t) can be recovered from the sample.

Band pass m(t) with lower frequency 'f_L' & upper frequency 'f_H', $f_H - f_L = B$. The minimum sampling frequency allowable is $f_s = 2(f_H - f_L)$ provided that either f_H or f_L is a harmonic of f_s.

A bandpass signal with highest frequency f_H and bandwidth B, can be recovered from its samples through bandpass filtering by sampling it with frequency $f_s = 2.f_H/k$, where k is the largest integer not exceeding f_H/B. All frequencies higher than f_s but below $2.f_H$(lower limit from low pass sampling theorem) may or may not be useful for bandpass sampling depending on overlap of shifted spectrum.

Eg. Let us say, f_L=2.5 KHz and f_H=3.5 KHz.
So, B=1 KHz, k=f_M / B =3.5.
Selecting fs = 2B = 2 KHz cause overlap.
If k is taken as 3 then f_s = 2*3.5 kHz/3 = 7/3 kHz cause no overlap.
If k is taken as 2 then f_s = 2*3.5 KHz/2 = 3.5 KHz cause no overlap.

- **Aliasing Effect:-**

From the spectrum of $G_s(f)$ we can filter out one of the spectrum, say $-W<f<W$, using a low pass filter and can reconstruct the time domain representation of it after doing inverse fourier transform of the spectrum. This is possible only when $f_s >= 2W$.
But when $f_s < 2W$, ie, $T_s > 1/2W$, then there will be overlap of adjacent spectrums. Here high frequency part of 1st spectrum interfere with low frequency part of 2nd spectrum. This phenomenon is the aliasing effect. In such a case the original signal g(t) cannot be recovered exactly from its sampled values $g_s(t)$.

➢ Signal Reconstruction :

The process of reconstructing a continuous time signal g(t)[bandlimited to W Hz] from its samples is also known as interpolation. This is done by passing the sampled signal through an ideal low pass filter of bandwidth W Hz. As seen from eqn. 1.4, the sampled signal contains a component $1/T_s\, G(f)$, and to recover G(f)[or g(t)], the sampled signal must be passed through on ideal low-pass filter of bandwidth W hz and gain T_s.

Thus the reconstruction(or interpolating) filter transfer function is,

$$H(f) = T_s \, rect(f/2W) \quad (1.5)$$

The interpolation process here is expressed in the frequency domain as a filtering operation.

Let the signal interpolating (reconstruction) filter impulse response be h(t). Thus, if we were to pass the sampled signal $g_r(t)$ through this filter, its response would be g(t).

Let us now consider a very simple interpolating filter whose impulse response is $rect(t/T_s)$, as shown in fig. 1.5. This is a gate pulse of unit height, cantered at the origin, and of width T_s(the sampling interval). Each sample in $g_\delta(t)$, being an impulse generates a gate pulse of the height equal to the strength of the sample. For instance the kth sample is an impulse of strength $g(kT_s)$ located at $t=kT_s$, and can be expressed as $g(kT_s)\ \delta(t-kT_s)$. When this impulse passes thorugh the filter, it generates and ouput of $g(kT_s)\ rect(t/T_s)$. This is a gate pulse of height $g(kT_s)$, centred at $t=kT_s$(shown shaded in fig. 1.5).

Each sample in $g_\delta(t)$ will generate a corresponding gate pulse resulting in an output,

$$y(t) = \sum_k g(k..Ts) \, rect\left(\frac{t}{Ts}\right) \quad (1.6)$$

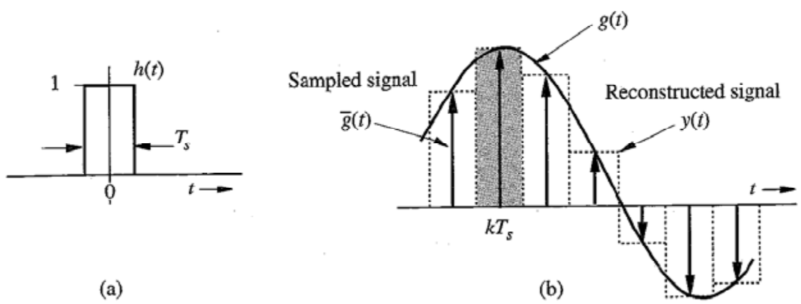


Fig. 1.5 Simple interpolation using zero-order hold circuit

The filter output is a staircase approximation of g(t), shown dotted in fig. 1.5b. This filter thus provides a crude form of interpolation.

The transfer function of this filter H(f) is the fourier transform of the impulse response $rect(t/T_s)$. Assuming the Nyquist sampling rate, ie, $T_s = 1/2W$,

$$W(t) = rec(t/Ts) = rect(2Wt)$$

and, $$H(f) = T_s.sinc(\pi.f.Ts) = 1/(2W).sinc(\pi f/2W) \quad (1.7)$$

The amplitude response |H(f)| for this filter shown in fig. 1.6, explains the reason for the crudeness of this interpolation. This filter is also known as the zero order hold filter, is a poor approximation of the ideal low pass filter(as shown double shaded in fig. 1.6).

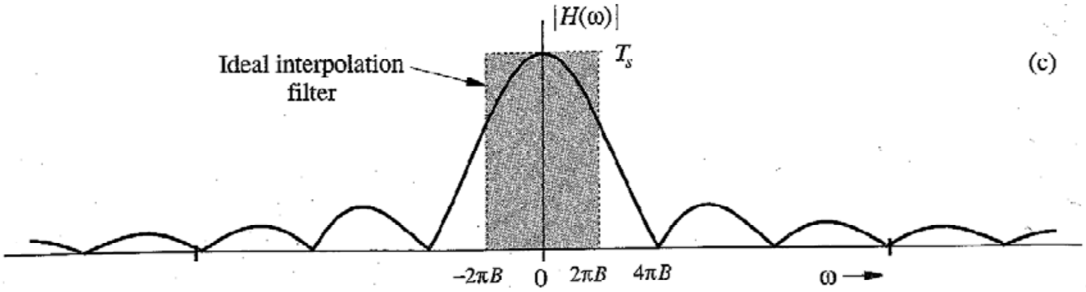


Fig. 1.6 Amplitude response of interpolation filter.

We can improve on the zero order hold filter by using the first order hold filter, which results in a linear interpolation instead of the staircase interpolation. The linear interpolator, whose impulse response is a triangular pulse $\Delta(t/2T_s)$, results in an interpolation in which successive sample tops are connected by straight line segments. The ideal interpolation filter transfer function found in eqn. 1.5 is shown in fig. 1.7a. The impulse response of this filter, the inverse fourier transform of H(f) is,

$$h(t) = 2.W.Ts.sinc(Wt),$$

Assuming the Nyquist sampling rate, ie, $2WT_s = 1$, then

$$h(t) = sinc(Wt) \tag{1.8}$$

This h(t) is shown in fig. 1.7b.

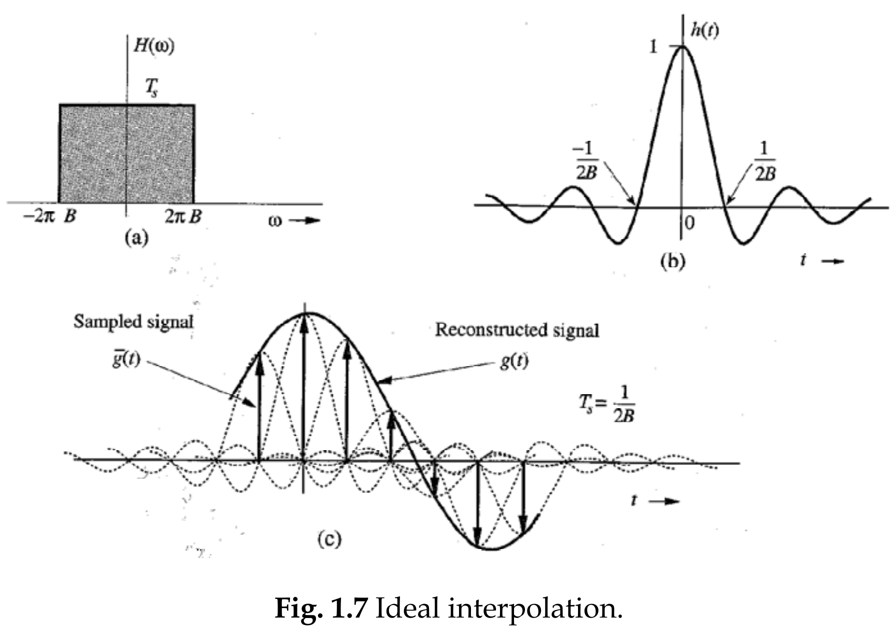


Fig. 1.7 Ideal interpolation.

The very interesting fact we observe is that, h(t) = 0 at all Nyquist sampling instants(t = ±n/2W) except at t=0. When the sampled signal $g_\delta(t)$ is applied at the input of this filter, the output is g(t). Each sample in $g_\delta(t)$, being an impulse, generates a sine pulse of height equal to the strength of the sample, as shown fig. 1.7c.

The process is identical to that shown in fig. 1.7b, except that h(t) is a sine pulse instead of gate pulse. Addition of the sine pulses generated by all the samples results in g(t). The k^{th} sample of the input $g_\delta(t)$ is the impulse $g(kT_s)\delta(t-kT_s)$; the filter output of this impulse is $g(kT_s)h(t-kT_s)$. Hence, the filter output to $g_\delta(t)$, which is g(t), can now be expressed as a sum.

$$g(t) = \sum_k g(k.Ts)\,\mathrm{h(t - KTs)}$$

$$= \sum_k g(k.Ts)\,\mathrm{sinc[W(t - KTs)]} \quad (1.9a)$$

$$= \sum_k g(k.Ts)\,\mathrm{sinc[Wt - K/2]} \quad (1.9b)$$

Eqn. 1.9 is the interpolation formula, which yields values of g(t) between samples as a weighted sum of all the sample values.

- **<u>Practical Difficulties:</u>**

  If a signal is sampled at the Nyquist rate f_s = 2W hz, the spectrum $G_\delta(f)$ without any gap between successive cycles.. To recover g(t) from $g_\delta(t)$, we need to pass the sampled signal $g_\delta(t)$ through an ideal low pass filter. Such filter is unrealizable; it can be closely approximated only with infinite time delay in the response. This means that we can recover the signal g(t) from its samples with infinite time delay.
  A practical solution to this problem is to sample the signal at a rate higher h=than the Nyquist rate(f_s > 2W). This yields $G_\delta(f)$, consisting of repetition of G(t) with a finite band gap between successive cycles. We can now recover G(g) from $G_\delta(f)$ from $G_\delta(f)$ using a low pass filter with a gradual cut-off characteristics. But even in this case, the filter gain is required to be zero beyond the first cycle of G(f). By Paley-Wiener criterion, it is also impossible to realize even this filter. The only advantage in this case is that the required filter can be closely approximated with a smaller time delay.
  This indicated that it is impossible in practice to recover a band limited signal $g_\delta(t)$ exactly from its samples even if sampling rate is higher than the Nyquist rate. However as the sampling rate increases, the recovered signal approaches the desired signal more closely.

- <u>The Treachery of Aliasing:</u>

  There is another fundamental practical difficulty in reconstructing a signal from its samples. The sampling theorem was proved on the assumption that the signal g(t) is bandlimited. All practical signals are time limited, ie, they are of finite duration width. A signal cannot be time-limited and band-limited simultaneously. If a signal is time limited, it cannot be band limited and vice-versa(but it can be simultaneously non time limited and non band limited). This means that all practical signals which are time limited are non band limited; they have infinite bandwidth and the spectrum $G_\delta(f)$ consists of overlapping cycles of G(f) repeating every f_s hz(the sampling frequency) as shown in fig. 1.8.

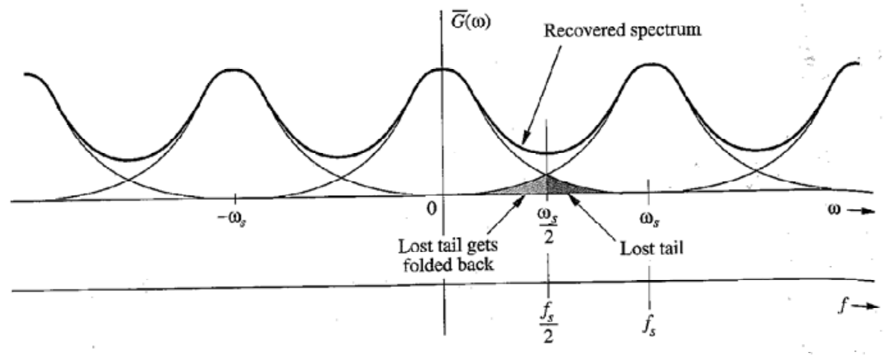


Fig. 1.8 Aliasing effect

Because of the overlapping tails, $G_\delta(f)$ no longer has complete information about G(f) and it is no longer possible even theoretically to recover g(t) from the sampled signal $g_\delta(t)$. If the sampled signal is passed through and ideal low pass filter the output is not G(f) but a version of G(f) distorted as a result of two separate causes:

1. The loss of the tail of G(f) beyond $|f| > f_s/2$ Hz.
2. The reappearance of this tail inverted or folded onto the spectrum.

The spectra cross at frequency $f_s/2 = 1/2T_s$ Hz, is called the folding frequency. The spectrum, therefore, folds onto itself at the folding frequency. In fig. 1.8, the components of frequencies above $f_s/2$ reappear as components of frequencies below $f_s/2$. This tail inversion, known as spectral folding or aliasing is shown shaded in fig. 1.8. In this process of aliasing, we are not only losing all the components of frequencies above $f_s/2$ Hz, but these very components reappear(aliased) as lower frequency components also as in fig. 1.8.

- <u>A Solution: The Antialiasing Filter</u>

The potential defectors are all the frequency components beyond $f_s/2 = 1/2T_s$ Hz. We should eliminate (suppress) these components from g(t) before sampling g(t). This way, we lose only the components beyond the folding frequency $f_s/2$ Hz. These components now cannot reappear to corrupt the components with frequencies below the folding frequency. This suppression of higher frequencies can be accomplished by an ideal low pass filtr of bandwidth $f_s/2$ hz. This filter is called the antialiasing filter. This antialiasing operation must be performed before the signal is sampled.

The antialiasing filter, being an ideal filter, is unrealizable. In practice we use a steep cut off filter which leaves a sharply attenuated residual spectrum beyond the folding frequency $f_\delta/2$.

Even using antialiasing filter, the original signal may not be recovered if $T_s > 1/2W$, ie, $f_s < 2W$. For this case also aliasing will occur. To avoid this sampling frequency f_s should be always greater than or atleast equal to 2W, where W is the highest frequency component available in information signal.

➢ Some Applications of the Sampling Theorem:

In the field of digital communication the transmission of a continuous time message is replaced by the transmission of a sequence of numbers. These open doors to many new techniques of communicating continuous time signals by pulse trains. The continuous time signal g(t) is sampled, and samples values are used to modify certain parameters of a periodic pulse train. As per these parameters, we have pulse amplitude modulation (PAM), pulse width modulation (PWM) and pulse position modulation (PPM). In all these cases instead of transmitting g(t), we transmit the corresponding pulse modulated signal. One advantage of using pulse modulation is that it permits the simultaneous transmission of several signals on a time sharing basis-time division multiplexing (TDM) which is the dual of FDM.

➢ Pulse Amplitude Modulation(PAM) :

In PAM, the amplitude of regularly spaced rectangular pulses vary with the instantaneous sample value of a continuous message signal in one to one fashion.

$$V_{PAM}(t) = \sum_{n=-\infty k}^{\infty}[1 + Ka.g(n.Ts)]\delta(t - n.Ts)$$

Where $g(nT_S)$ represents the nth sample of the message signal g(t), T_s is the sampling time, k_a is a constant called the amplitude sensitivity(or modulation index of PAM) and $\delta_{TS}(t)$ demotes the pulse train. 'k_a' is chosen so as to maintain a single polarity, ie, $\{1+k_a g(nT_s)\} > 0$ for all values of $g(nT_s)$.

Different forms of pulse analog modulation (PAM, PWM & PPM) are illustrated below:-

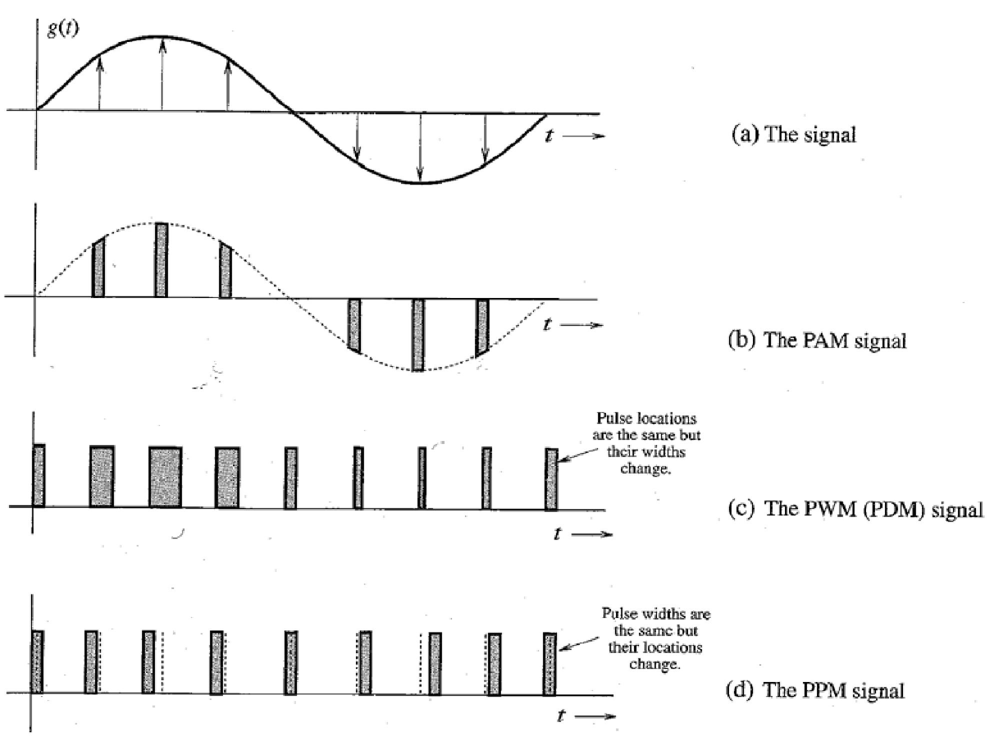


Fig. 1.9 Pulse modulated signals.

Transmission BW in PAM

We know $\tau << T_s \leq 1/2W$

Considering 'ON' and 'OFF' time of PAM it is velar the maximum frequency of PAM is $f_{max} = 1/2\tau$.

So transmission BW $\geq f_{max} = 1/2\tau >> W$.

Noise performance of PAM is never better than the baseband signal transmission.

However we need PAM for message processing for a TDM system, from which PCM can be easily generated or other form of pulse modulation can be generated.

Be it single or multi user system the detection should be done in synchronism. So synchronization between transmitter and receiver is an important requirement.

Pusle Width Modulation(PWM):

In pulse width modulation, the instantaneous sample values of the message signal are used to vary the duration of the individual pulses. This form of modulation is also referred to as pulse duration modulation (PDM) or pulse length modulation (PLM).

Here the modulating wave may vary the time of occurrence of leading edge, the trailing edge or both edges of the pulse.

Disadvantage – In PWM, long pulses (more width) expand considerable power during the pulse transmission while bearing no additional information.

$$V_{PWM} = P(t - n.T_s) = \delta(t - n.Ts) \quad \text{for } nTs < t < (nTs + kn.g(nTs))$$

$$= 0 \quad \text{for } [nT_s + k_w.g(nT_S)] \leq t \leq (n+1)T_s$$

Generation of PWM and PPM waves:

The figure below depicts the generation of PWM and PPM waves. Hence for the PWM wave the trailing edge is varied according to the sample value of the message.

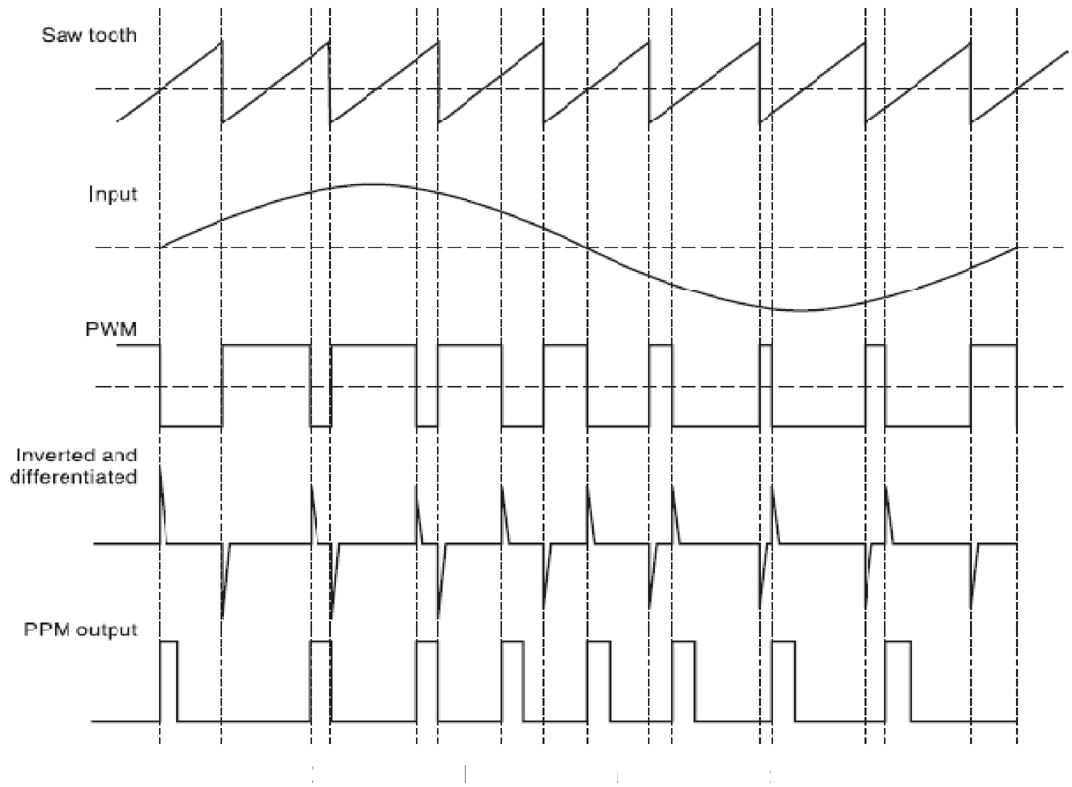


Fig. 1.10 Principle of PWM and PPM generation.

The saw tooth generator generates the sawtooth signal of frequency f_s ($f_s = 1/T_s$). If sawtooth waveform is reversed, then leading edge of the pulse will be varied with samples of the signal and if the sawtooth waveform is replaced by a triangular waveform then both the edged will vary according to samples.
PPM waveform is generated when PWM wave is used as the trigger input to a monostable multivibrator. The monostable multivibrator is triggered on the falling (trailing edge) of PWM. The output of monostable is then switches to positive saturation value and remain there for a fixed period and then goes low. Thus a pulse is generated which occurs at a time which occurs at a time which depends upon the amplitude of the sampled value.

Demodulation of PWM waves

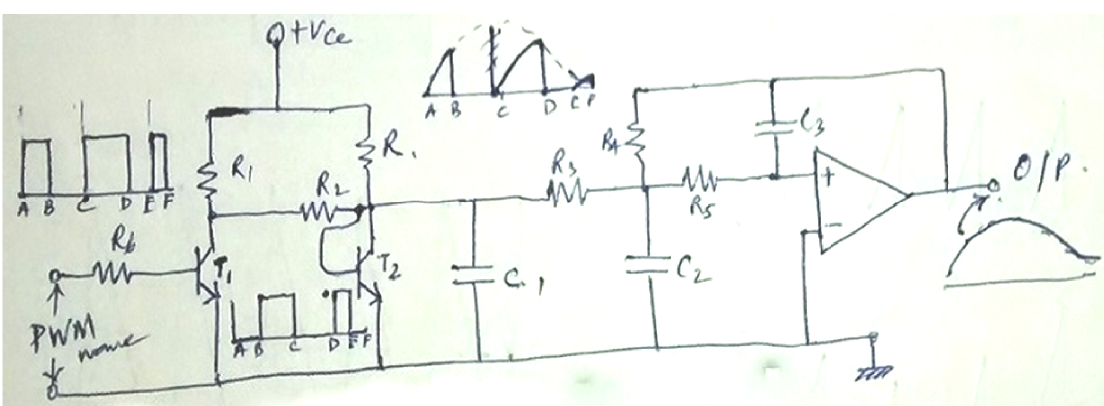


Fig. 1.10a A PWM demodulator circuit.

Here transistor T1 acts as an inverter. Hence when transfer is off capacitor C1 will chase through R as when it is 'on' C1 discharges quickly through T1 as the resistance in the path is very small. This produces a sawtooth wave at the output of T2. This sawtooth wave when passed through an op-amp with 2nd order LPF produces the desired wave at the output.

Demodulation of PPM waves:

Since in PPM the gaps in between pulses contains information, so during the gaps, say OA, BC and DE the transfer T remains off and capacitor the capacitor C gets charged. The voltage across the capacitor depends on time of charging as the value of R and C. Rest of the operation is same as above.

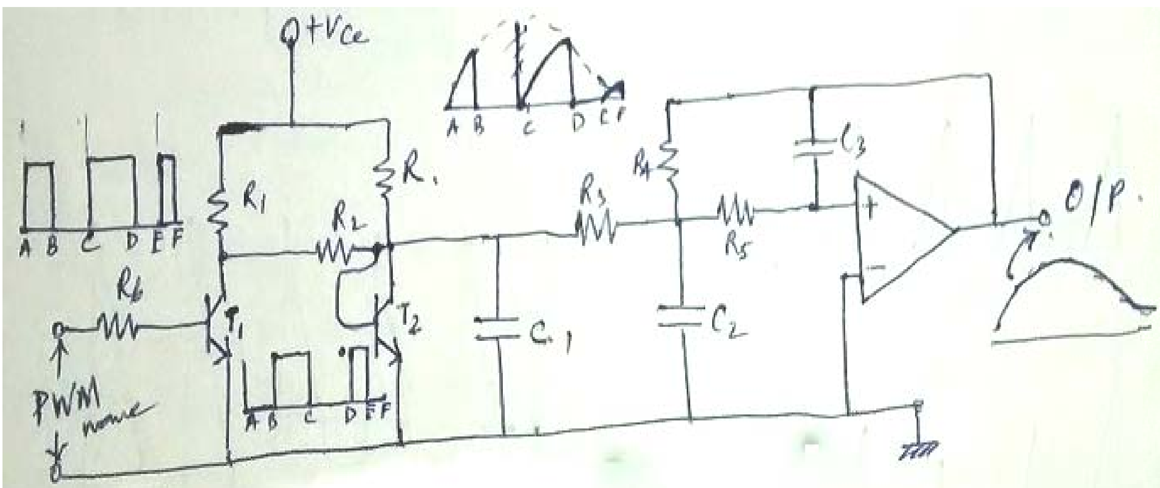


Fig. 1.10b A PPM demodulator circuit.

MODULE – II **12 hours**

PCM is the most useful and widely used of all the pulse modulations mentioned. Basically, PCM is a method of converting an analog signal into a digital signal (A/D conversion).An analog signal is characterized by the fact that its amplitude can take on any value over a continuous range. This means that it can take on an infinite number of values. On the other hand, digital signal amplitude can take on only a finite number of values. An analog signal can be converted into a digital signal by means of sampling and quantizing, that is, rounding off its value to one of the closest permissible numbers (or quantized levels) as shown in fig 2.1.

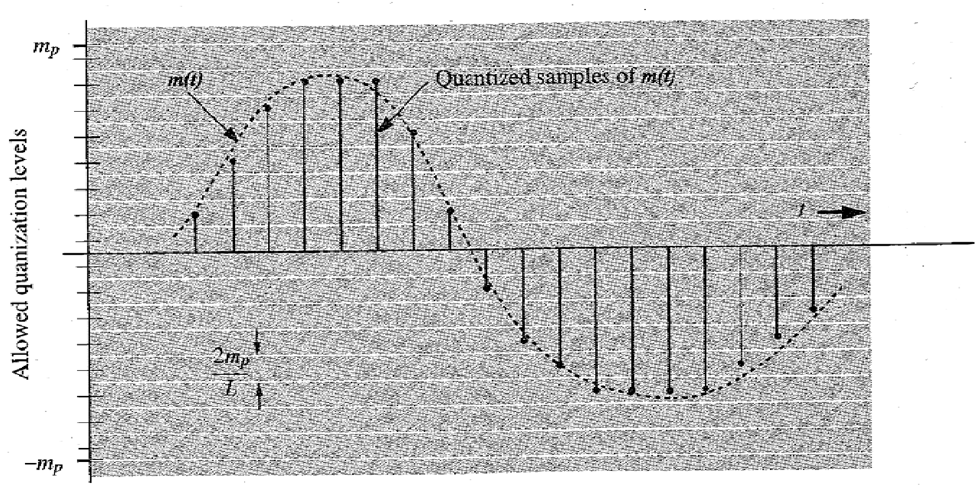


Fig. 2.1 Quantization of a sampled analog signal.

Quantization is of two types:--uniform and non-uniform quantization.

- **Uniform Quantization :--**

  Amplitude quantizing is the task of mapping samples of a continuous amplitude waveform to a finite set of amplitudes. The hardware that performs the mapping is the analog-to-digital converter (ADC or A-to-D). The amplitude quantizing occurs after the sample-and-hold operation. The simplest quantizer to visualize performs an instantaneous mapping from each continuous input sample level to one of the preassigned equally spaced output levels. Quantizers that exhibit equally spaced increments between possible quantized output levels are called uniform or linear quantizers.

  Possible instantaneous input-output characteristics are easily visualized by a simple staircase graph consisting of risers and treads of the types shown in Fig 2.2. Fig 2.2 a, b, and d show quantizers with uniform quantizing steps, while fig 2.2c is a quantizer with nonuniform quantizing steps.

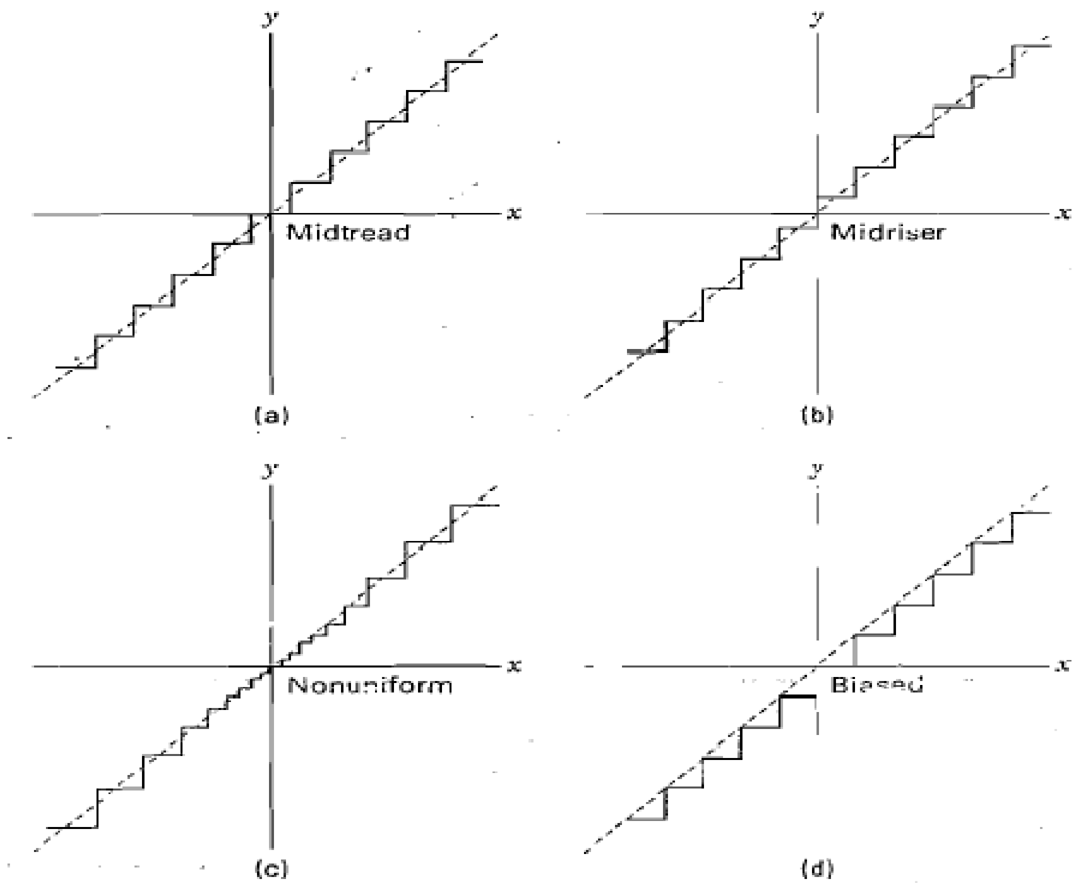


Fig. 2.2 Various quantizers transfer functions.

- **<u>Non Uniform Quantization:</u>**

For many classes of signals the uniform quantization is not efficient, for example, in speech communication it is found(statistically) that smaller amplitudes predominate in speech and that larger amplitudes are relatively rare. The uniform quantizing scheme is thus wasteful for speech signals; many of the quantizing levels are rarely used. An efficient scheme is to employ a non uniform quantizing method in which smaller steps for small amplitudes are used.

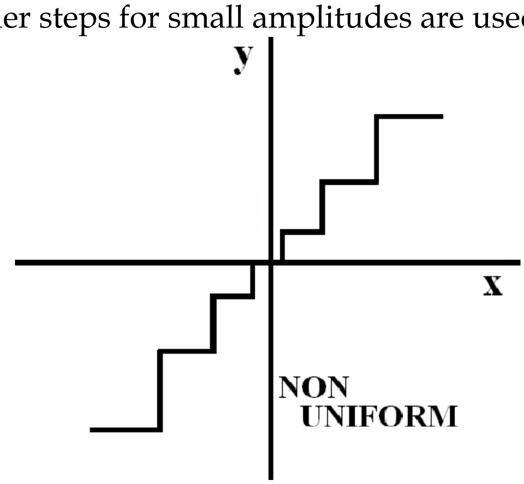


Fig. 2.3. Non-uniform quantization

The same result can be achieved by first compressing the signal samples and then using a uniform quantizing. The input-output characteristics of a compressor are shown in below fig. 2.4

The same result can be achieved by first compressing the signal samples and then using a uniform quantizing. The input output characteristics of a compressor are shown in fig. The horizontal axis is the normalized input signal (ie, g/g_P), and the vertical axis is the output signal y. The compressor maps input signal increment Δg, into larger increment Δy for small signal input signals and small increments for larger input signals. Hence, by applying the compressed signals to a uniform quantizer a given interval Δg contains a larger no. of steps (or smaller step-size) when g is small.

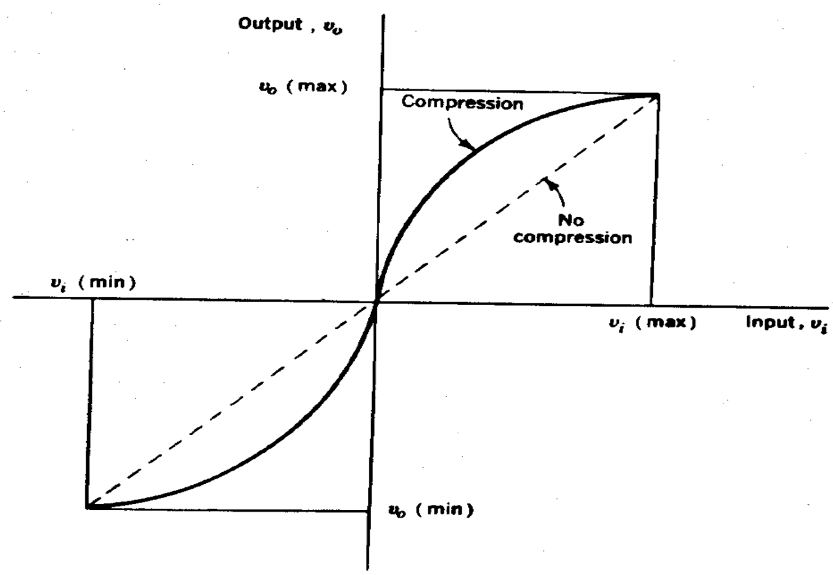


Fig. 2.4 Characteristics of Compressor.

A particular form of compression law that is used in practice (in North America and Japan) in the so called μ law (μ law compressor), defined by

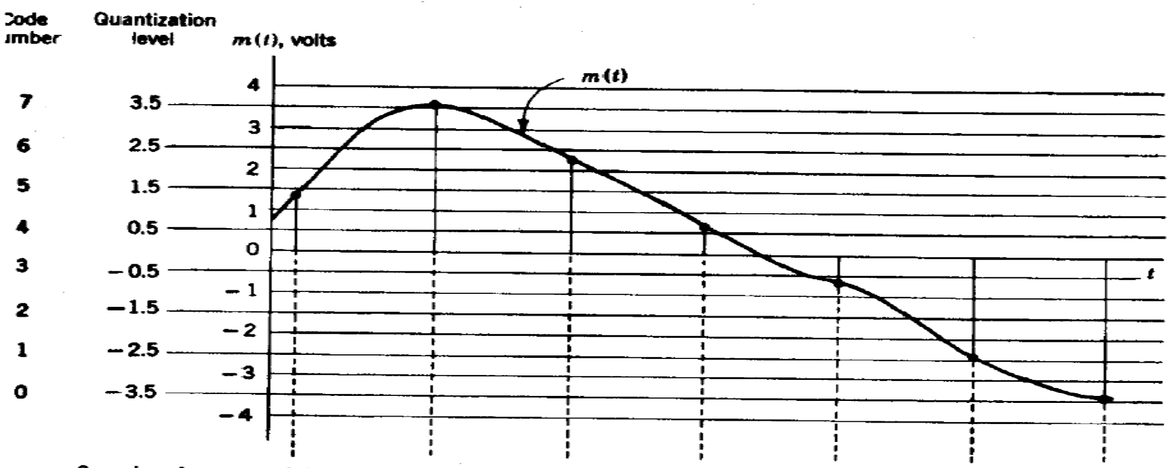
$y = \ln(1 + \mu|g/g_P|)/\ln(1 + \mu).\text{sgn}(g)$ for $|g/g_P| \leq 1$ where, μ is a +ve constant and sgn(g) is a signum function.

Another compression law popular in Europe is the so A-law, defined by,

$$y = A/(1+\ln A).(g/g_P) \quad \text{for } 0 \leq g/g_P \leq 1/A$$
$$= (1+\ln A|g/g_P|/(1+\ln A)).\text{sgn}(g) \quad \text{for } 1/A \leq |g/g_P| \leq 1 \qquad (2.1)$$

The values of μ & A are selected to obtain a nearby constant output signal to quantizing noise ratio over an input signal power dynamic range of 40 dB.
To restore the signal samples to their correct relative level, an expander with a characteristic complementary to that of compressor is used in the receiver. The combination of compression and expansion is called companding.

➢ Encoding:-



Sample value	1.3	3.6	2.3	0.7	−0.7	−2.4	−3.4
Nearest quantization level	1.5	3.5	2.5	0.5	−0.5	−2.5	−3.5
Code number	5	7	6	4	3	1	0
Binary representation	101	111	110	100	011	001	000

Fig. 2.5 Representation of each sample by its quantized value and binary representation.

A signal g(t) bandlimited to B hz is sampled by a periodic pulse train $P_{Ts}(t)$ made up of a rectangular pulse of width 1/8B seconds (cantered at origin), amplitude 1 unit repeating at the Nyquist rate(2B pulses per second. Show that the sampled signal is given by,

$$\bar{g}(t) = \tfrac{1}{4}.g(t) + \sum_{n=1}^{\infty}\left(\frac{2}{n\pi}.\sin(n\pi/4)g(t)\cos n.w_s.t\right) \quad (2.2)$$

➢ Quantizing Noise or Quantizing Error :

We assume that the amplitude of g(t) is confined to the range($-g_p$, g_p). This range is divided into L no. of equal segments. Each segment is having step size Δ, given by,

$$\Delta = 2.g_p/L \quad (2.3)$$

A sample amplitude value is approximated by the mid-points of the interval in which it lies. The input-output characteristic of a midrise uniform quantizer is shown in fig.

The difference between the input and output signals of the quantizer becomes the quantizing error or quantizing noise.

It is apparent that with a random input signal, the quantizing error 'q_e' varies randomly within the interval,

$$-\Delta/2 \le q_e \le \Delta/2 \quad (2.4)$$

Assuming that the error is equally likely to lie anywhere in the range $(-\Delta/2, \Delta/2)$, the mean square quantizing error $\langle q_e^2 \rangle$ in given by,

$$\langle q^2_e \rangle = 1/\Delta \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} q_e^2 \, dq_e = \Delta^2/12 \qquad (2.5)$$

Substituting eqn.(2.3) in eqn.(2.5) we get,

$$\langle q^2_e \rangle = g^2_p/(3L^2)$$

$$S_i = \langle g^2(t) \rangle = \int_{-g_p}^{g_p} g^2(t) \,.\, \frac{1}{2} \,.\, g.dg = g^2_p/3$$

➢ Transmission Bandwidth and the output SNR :

For binary PCM, we assign a distinct group of 'n' binary digits(bits) to each of the L quantization levels. Because a sequence of n binary digits can be arranged in 2^n distinct patterns,

$$L = 2^n \text{ or } n = \log_2 L \qquad (2.6)$$

Each quantized sample is thus, encoded into 'n' bits. Because a signal g(t) bandlimited to W Hz requires a minimum of 2W samples second, we require a total of 2nW bits per second(bps), ie, 2nW pieces of information per second. Because a unit bandwidth (1 Hz) can transmit a maximum of two pieces of information per second, we require a minimum channel of bandwidth B_T Hz, given by,

$$B_T = n.W \quad \text{Hz} \qquad (2.7)$$

This is the theoretical minimum transmission bandwidth required to transmit the PCM signal. We shall see that for practical reasons we may use transmission bandwidth higher than as in eqn.(2.7).

$$\text{Quantizing Noise} = N_o = \langle q^2_e \rangle = g^2_p/(3.L^2) \qquad (2.8)$$

Assuming the pulse detection error at the receiver is negligible, the reconstructed signal $\hat{g}$ (t) at the receiver output is,

$$\hat{g}(t) = g(t) + q_e(t) \qquad (2.9)$$

The desired signal at the output is g(t), and the (quantizing) noise is $q_e(t)$. Since the power of the message signal g(t) is $\langle g^2(t) \rangle$, then

$$S_0 = \langle g^2(t) \rangle \qquad (2.10)$$

So, $SNR = S_o/N_o = \langle g^2(t)\rangle / (g^2_p/(3.L^2)) = 3L^2 \langle g^2(t)\rangle / g^2_p$ (2.11)

$S_o/N_o(dB) = 10.\log(3L^2) \langle g^2(t)\rangle / g^2_p$

Signal to noise ration can be written as,

$$S_o/N_o = 3.2^{(2n)} \langle g^2(t)\rangle / g^2_p \quad (2.12)$$

$$= C(2)^{2n} \quad (2.13)$$

Where,

$C = 3.\langle g^2(t) \rangle / g^2_p$ (uncompressed case, as in eqn.(2.12))

$= 3/[\ln(1+\mu)]^2$ (compressed case)

For a μ-law compander, the output SNR is,

$S_o/N_o = 3.l^2/[\ln(1+\mu)]^2 \qquad \mu^2 >> g^2_p/\langle g^2(t)\rangle$

Substituting eqn.(2.7) in eqn.(2.12), we find

$$S_o/N_o = C(2)^{2.B_T/W} \quad (2.14)$$

From eqn.(2.14), it is observed that SNR increases almost exponentially with the transmission bandwidth B_T. This trade-off SNR with bandwidth is attractive and come close to the upper theoretical limit. A small increase in bandwidth yields a large benefit in terms of SNR. This trade relationship is clearly seen by rewriting eqn.(2.14) using decibel scale as,

S_0/N_0 (dB) $= 10.\log(S_o/N_o)$

$= 10\log(C2^{2n})$

$= 10\log C + 20\log 2$

$= (\alpha + 6n)$ dB (2.15)

Where, $\alpha = 10\log C$. This shows that increasing n by 1, quadruples the output SNR(6 dB increase). Thus if we increase 'n' from 8 to 9, the SNR quadruples, but the transmission bandwidth increases only from 32 to 36 Khz(an increase of only 12.5%). This shows that in PCM, SNR can be controlled by transmission bandwidth. We shall see later that frequency and phase modulation also do this. But it requires a doubling of the bandwidth to quadruple the SNR. In this respect, PCM is strikingly superior to FM or PM.

➢ Digital Multiplexer :--

This is a device which multiplexers or combines several low bit rate signals to form one high bit rate signal to be transmitted over a high frequency medium. Because of the medium is time shared by various incoming signals, this is a case of time-division multiplexing (TDM. The signals from various incoming channels may be such diverse nature as digitized voice signal (PCM), a computer output, telemetry data, a digital facsimile and so on. The bit rates of the various tributaries (channels) need not be the same.

Multiplexing can be done on a bit-by-bit basis(known as bit or digit interleaving) or on a word-by-word basis(known as byte or word interleaving). The third category is interleaving channel having different bit rate.

T1 carrier system:-- The input to the (fast) 13-bit ADC comes from an analog multiplexer. The digital processor compresses the digital value according to μ-law.

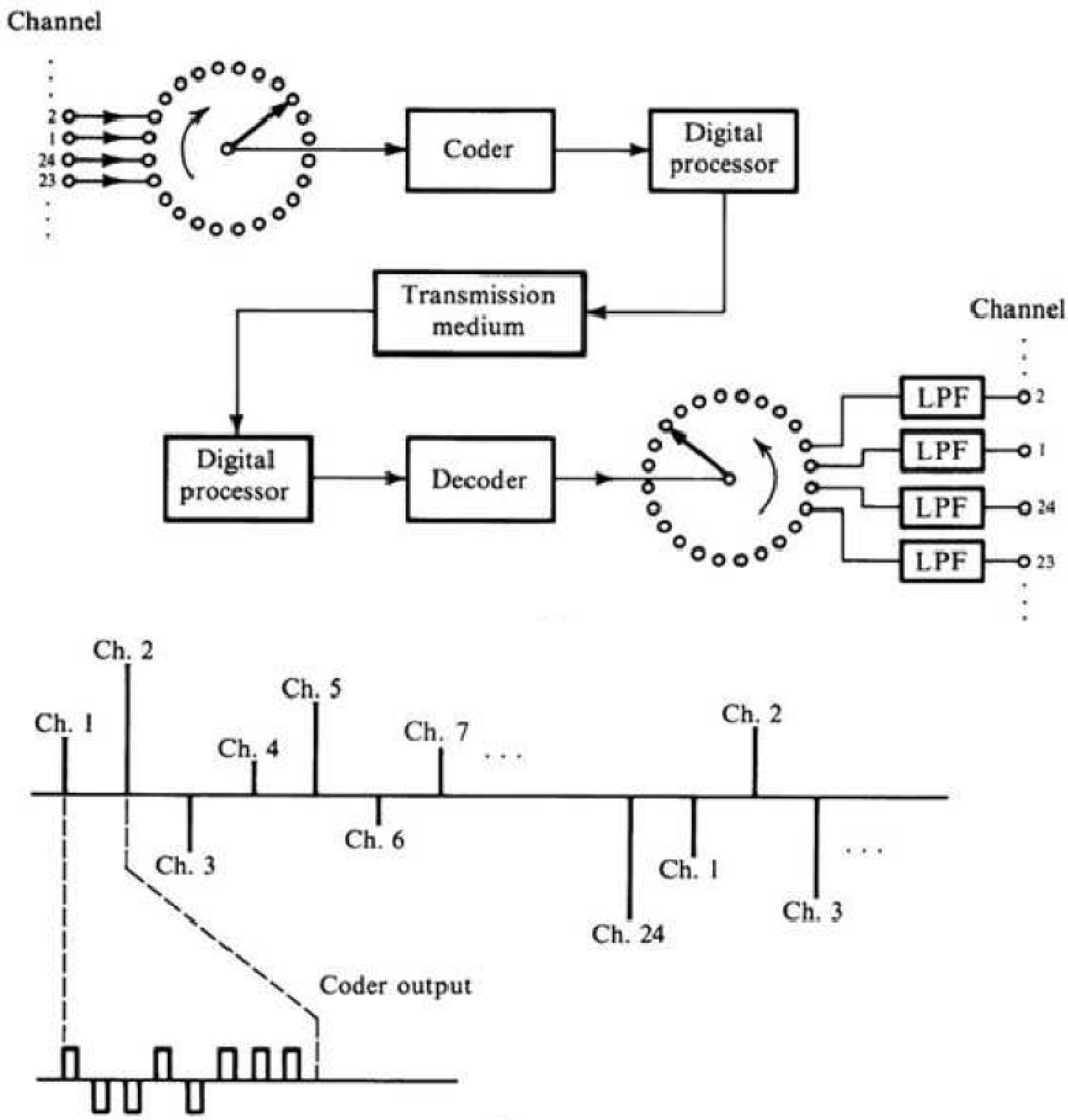


Fig. 2.6 T-1 carrier system.

The 8-bit compressed voice values are sent consecutively, MSB first. The samples of all 24 inputs comprise a frame. Most serial communications transmits data LSB first ("little endian").

➢ Synchronizing & Signalling :

Binary code words corresponding to samples of each of the 24 channels are multiplexed in a sequence as shown in fig 2.7. A segment containing one codeword (corresponding to one sample) from each of the 24 channels is called a frame. Each frame has 24x8 = 192 information bits. Because the sampling rate is 8000 samples per second, each frame takes 125 μs. At the receiver it is necessary to be sure where each frame begins in order to separate information bits separately. For this purpose, s framing bit is added at the beginning of each frame. This makes a total of 193 bits per frame. Framing bits are chosen so that a sequence of framing bits, one at the beginning of each frame, forms a special pattern that is unlikely to be formed in a speech channel.

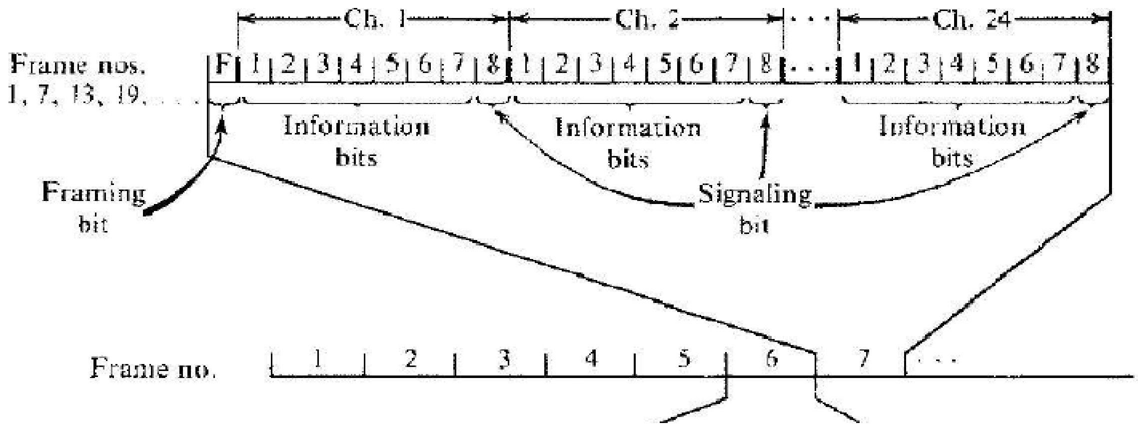


Fig. 2.7 T-1 frame.

The sequence formed by the first bit from each frame is examined by the logic of the receiving terminal. If this sequence does not follow the given coded pattern (framing bit pattern), then a synchronization lost is detected and the next position is examined to determine whether it is actually the framing bit. It takes about 0.4 to 6 ms to detect and about 50 ms (in the worst possible case) to reframe.

In addition to information and framing bits we need to transmit signalling bits corresponding to dialling pulses, as well as telephone on-hook/off-hook signals. When channels developed by this system are used to transmit signals between telephone switching systems, the switches must be able to communicate with each other to use the channels effectively. Since all eight bits are now used for transmission instead of the seven bits used in the earlier version, the signalling channel provided by the eighth bit is no longer available. Since only a rather low speed signalling channel is required, rather than create extra time slots for this information we use one information bit(the least significant bit) of every sixth sample of a signal to transmit this information. This means every sixth sample of each voice signal will have a possible error corresponding to the least significant digit. Every sixth frame, therefore, has 7x24 = 168 information bits, 24 signalling bits and 1 framing bit. In all the remaining frames, there are 192 information bits and 1 framing bit. This technique is called 75/6 bit encoding and the signalling channel so derived is called robbed-bit signalling. The slight SNR degradation suffered by impairing one out of six frame is considered to be an acceptable penalty. The signalling bits for each signal occur at a rate of 8000/6 = 1333 bits/sec.

In such above case detection of boundary of frames is important. A new framing structure called the super frame was developed to take care of this. The framing bits are transmitted at the 8 kbps rate as before (earlier case) and occupy the first bit of each frame. The framing bits form a special pattern which repeats in twelve frames: 100011011100. The pattern thus allows the identification of frame boundaries as before, but also allows the determination of the locations of the sixth and twelfth frames within the superframe. Since two signalling frames are used so two specific job can be initiated. The odd numbered frames are used for frame and sample synchronization and the even numbered frames are used to identify the A & B channel signalling frames(frames 6 & 12).

A new superframe structure called the extended superframe (ESF) format was introduced during 1970s to take advantage of the reduced framing bandwidth requirement. An ESG is 24 frames in length and carries signalling bits in the eighth bit of each channel in frames 6, 12, 18 and 24. Sixteen state signalling is thus possible. Out of 24 framing bits 4^{th}, 8^{th}, 12^{th}, 16^{th}, 20^{th} and 24^{th}(2 kbps) are used for frame synchronization and have a bit sequence 001011. Framing bits 1, 5, 9, 13, 17 and 21(2 kbps) are for error detection code. 12 remaining bits are for management purpose and called as facility data link(FDL). The function of signalling is also the common channel interoffice signalling (CCIS).

Differential Pulse Code Modulation :

In analog messages we can make a good guess about a sample value from a knowledge of the past sample values. In other words, the sample values are not independent and there is a great deal of redundancy in the Nyquist samples. Proper exploitation of this redundancy leads to encoding a signal with a lesser number of bits. Consider a simple scheme where instead of transmitting the sample values, we transmit the difference between the successive sample values.

If g[k] is the kth sample instead of transmitting g[k], we transmit the difference d[k] = g[k] – g[k-1]. At the receiver, knowing the d[k] and the sample value g[k-1], we can construct g[k]. Thus form the knowledge of the difference d[k], we can reconstruct g[k] iteratively at he receiver. Now the difference between successive samples is generally much smaller than the sample values. Thus peak amplitude, g_p of the transmitted values is reduced considerably. Because the quantization interval $\Delta = g_p/L$, for a given L(or n) this reduces the quantization interval Δ. Thus, reducing the quantization noise which is given by $\Delta^2/12$.

This means that for a given n(or transmission bandwidth), we can increase the SNR or for a given SNR we can reduce n(or transmission bandwidth).

We can improve upon scheme by estimating the value of the kth sample g[k] from knowledge of the previous sample values. If this estimate is $\hat{g}$[k], then we transmit the difference (prediction error) d[k] = g[k] – $\hat{g}$[k]. At the receiver also we determine the estimate $\hat{g}$[k] from the previous sample values and then generate g[k] by adding the received d[k] to the estimate $\hat{g}$[k]. Thus we reconstruct the samples at the receiver iteratively. If our prediction is worthful the predicted value $\hat{g}$[k]will be close to g[k] and their difference (prediction error) d[k] will be even smaller than the difference between the successive samples. Consequently this scheme known as the differential PCM(DPCM) is superior to that described in the previous paragraph which is a special case of DPCM, where the estimate of a sample value is taken as the previous sample value, ie, $\hat{g}$[k]=g[k-1].

Consider for example a signal g(t) which has derivative of all orders at 't'. Using Taylor series for this signal, we can express $g(t+T_s)$ s,

$$g(t+T_s) = g(t) + T_s.\dot{g}(t) + T^2{}_s/2!\,\ddot{g}(t) + \ldots\ldots \quad (2.16)$$

$$= g(t) + T_s.\dot{g}(t) \qquad \text{for small } T_s. \quad (2.17)$$

So from eqn.(2.16) it is clear a future signal can be predicted from the present signal and its all derivatives. Even if we know the first derivative we can predict the approximated signal.

Let us denote the kth sample of g(t0 by g[k], ie, $g[kT_s]$ =g[k] and $g(kT_s \pm T_s)$ = g[k ± 1] and so on. Setting $t=kT_s$ in eqn.(2.17) and recognizing $g(kT_s) \approx [g(kT_s) - g(kT_s - T_s)]/T_s$.

We obtain,

$$g[k+1] \approx g[k] + T_s[\{g[k] - g[k-1\}/T_s]]$$

$$= 2g[k] - g[k-1] \quad (2.18)$$

This shows that we can find a crude prediction of the (k+1)th sample from two previous samples. The approximation in eqn.(2.17) improves as we add more terms in the series on the right hand side. To determine the higher order derivatives in the series, we require more samples in the past. The larger the member of past samples we use, the better will be the prediction. Thus, in general we can express the prediction formula as,

$$g[k] \approx a_1g[k-1] + a_2g[k-2]+\ldots\ldots\ldots+ a_Ng[k-N] \quad (2.19)$$

The right hand side of eqn.(2.19), is , $\hat{g}$[k, the predicted value of g[k]. Thus,

$$\hat{g}[k] = a_1g[k-1] + a_2g[k-2]+\ldots\ldots\ldots+ a_Ng[k-N] \quad (2.20)$$

This is the eqn. of an Nth order predictor. Larger n would result in better prediction in general. The output of this filter (predictor) is $\hat{g}$[k], the predicted value of g[k]. the input is the previous samples g[k-1], g[k-2],……,g[k-n], although it is customary to say that the input is g[k] and the output is $\hat{g}$[k].

Eqn.(2.20) reduces to $\hat{g}$[k] = g[k-1] for the 1st order predictor. This is similar to eqn.(2.17). This means a_1 = 1 and the 1at order predictor is a simple time delay.

The predictor described in eqn.(2.20) is called a linear predictor. It is basically a transversal filter(a tapped delay line), where the tap gains are set equal to the prediction coefficients as shown in fig. 2.8.

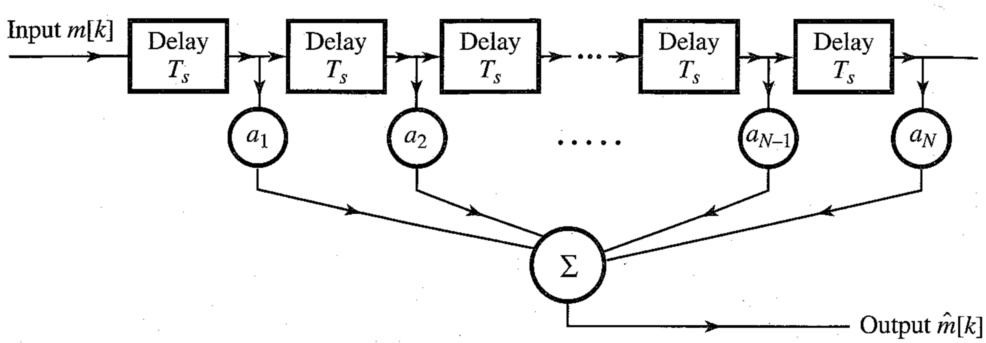


Fig. 2.8 Transversal filter(tapped delay line) used as a liner predictor

➢ Analysis of DPCM :

As mentioned earlier, in DPCM we transmit not the present sample g[k] but d[k] (the difference between g[k] and its predicted value $\hat{g}$[k]). At the receiver, we generate $\hat{g}$[k] from the past sample values to which the received d[k] is added to generate g[k]. There is, however, one difficulty in this scheme. At the receiver, instead of the past samples g[k-1], g[k-2],……. as well as d[k], we have their quantized versions g_p[k-1], g_p[k-2],….. Hence, we cannot determine $\hat{g}$[k]. We can only determine g_p[k], the estimate of the quantized sample g_q[k] in terms of the quantized samples g_q[k-1], g_q[k-2],………. This will increase the error in reconstruction. In such a case, a better strategy is to determine $\hat{g}_q$[k], the estimate of g_q[k](instead of g[k]), at the transmitter also from the quantized samples g_q[k-1], g_q[k-2],………. The difference d[k] = g[k] - g_q[k-2],……. is now transmitted using PCM. At the receiver we can generate $\hat{g}_q$[k], and from the received d[k], we can reconstruct g_q[k].

Fig 2.9 shows a DPCM transmitter. We shall soon see that the predictor input is g_q[k]. Naturally its output is $\hat{g}_q$[k], the predicted value of g_q[k]. The difference,

$$d[k] = g[k] - \hat{g}_q[k] \qquad (2.21)$$

is quantized to yield

$$d_q[k] = d[k] + q[k] \tag{2.22}$$

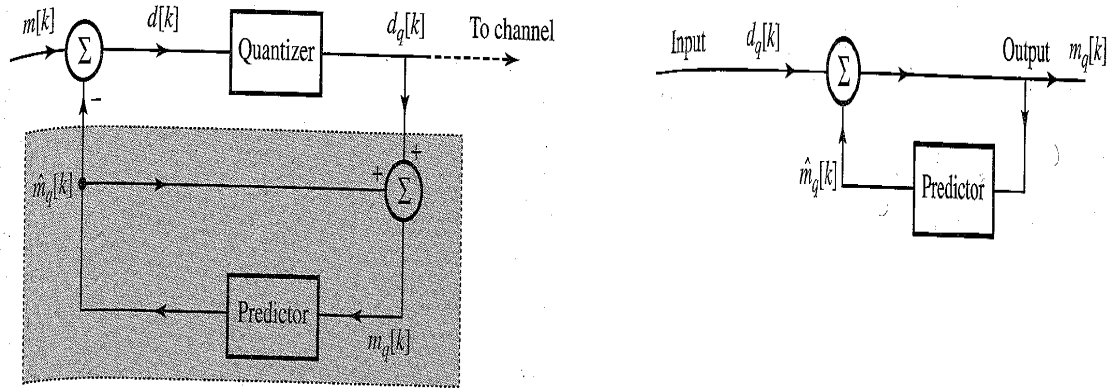


Fig. 2.9 DPCM system—Tansmitter and Receriver

In eqn.(2.22) q[k] is the quantization error. The predictor output $\hat{g}_q[k]$ is fed back to its input so that the predictor input $g_q[k]$ is,

$$\begin{aligned} g_q[k] &= \hat{g}_q[k] + d_q[k] \\ &= g[k] - d[k] + d_q[k] \\ &= g[k] + q[k] \end{aligned} \tag{2.23}$$

This shows that $g_q[k]$ is a quantized version of g[k]. The predictor input is indeed $g_q[k]$ as assumed. The quantized signal $d_q[k]$ is now transmitted over the channel. The receiver shown in fig 2.9 is identical to the shaded portion of the transmitter. The input in both cases is also the same, viz., $d_q[k]$. Therefore, the predictor output must be $\hat{g}_q[k]$ (the same as the predictor output at the transmitter). Hence, the receiver output (which is the predictor input) is also the same, viz., $g_q[k] = g[k] + q[k]$, as found in eqn.(2.23). This shows that we are able to receive the desired signal g[k] plus the quantization noise q[k]. This is the quantization noise associated with the difference signal d[k], which is much smaller than g[k]. The received samples are decoded and passed through a low pass filter of D/A conversion.

- **<u>SNR Improvement :</u>**

  To determine the improvement in DPCM over PCM, let g_p and d_p be the peak amplitudes of g(t) and d(t). If we use the same value of 'L' in both cases, the quantization step Δ in DPCM is reduced by the factor g_p/d_p. Because the quantization noise power is $\Delta^2/12$, the quantization noise in DPCM reduced by the factor $(g_p/d_p)^2$ and the SNR increases by the same factor. Moreover, the signal power is proportional to its peak value squared (assuming other statistical properties invariant). Therefore, G_p(SNR improvement due to prediction) is

$$G_p = P_g/P_d \tag{2.24}$$

Where P_g and P_d are the powers of g(t) and d(t) respectively. In terms of dB units, this means that the SNR increases by $10\log(P_m/P_d)$ dB.
For PCM,

$$(S_0/N_0) = \alpha + Gn \qquad \text{where, } \alpha = 10\log C \qquad (2.25)$$

In case of PCM the value of α is higher by $10\log(P_g/P_d)$ dB. A second order predictor processor for speech signals can provide the SNR improvement of around 5.6 dB. In practice, the SNR improvement may be as high as 25 dB. Alternately, for the same SNR, the bit rate for DPCM could be lower than that for PCM by 3 to 4 bits per sample. Thus telephone systems using DPCM can often operate at 32 kbits/s or even 24 kbits/s.

➢ **Delta Modulation:**

Sample correlation used in DPCM is further exploited in delta modulation(DM) by oversampling(typically 4 times the Nyquist rate) the baseband signal. This increases the correlation between adjacent samples, which results in a small prediction error that can be encoded using only one bit (L=2) for quantization of the $g[k] - \hat{g}_q[k]$. In comparison to PCM even DPCM, it us very simple and inexpensive method of A/D conversion. A 1-bit code word in DM makes word framing unnecessary at the transmitter and the receiver. This strategy allows us to use fewer bits per sample for encoding a baseband signal.

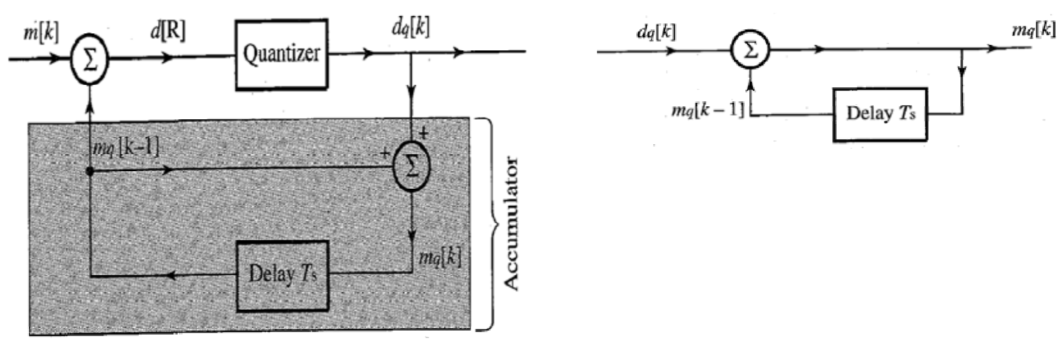


Fig. 2.10 Delta Modulation is a special case of DPCM

In DM, we use a first order predictor which as seen earlier is just a time delay of T_s(the sampling interval). Thus, the DM transmitter (modulator) and the receiver (demodulator) are identical to those of the DPCM in fig2.9 with a time delay for the predictor as shown in fig 2.10. From this figure, we obtain,

$$\hat{g}_q[k] = g_q[k-1] + d_q[k] \qquad (2.26)$$

Hence, $$g_q[k-1] = g_q[k-2] + d_q[k-1] \qquad (2.27)$$

Substituting eqn.(2.27) into eqn.(2.26) yields

$g_q[k] = g_q[k-2] + d_q[k] + d_q[k-1]$ (2.28)

Proceeding iteratively in this manner and assuming zero initial condition, ie, $g_q[0] = 0$, yields,

$g_q[k] = \sum_{g=0}^{k} dq[g]$ (2.29)

This shows that the receiver(demodulator) is just an accumulator(adder). If the output $d_{q=}[k]$ is represented by an integrator because its output is the sum of the strengths of the input impulses(sum of the areas under the impulses). We may also replace the feedback portion of the modulator (which is identical to the demodulator) by an integrator. The demodulator output is $g_p[k]$, which when passed through a low pass filter yields the desired signal reconstructed from the quantized samples.

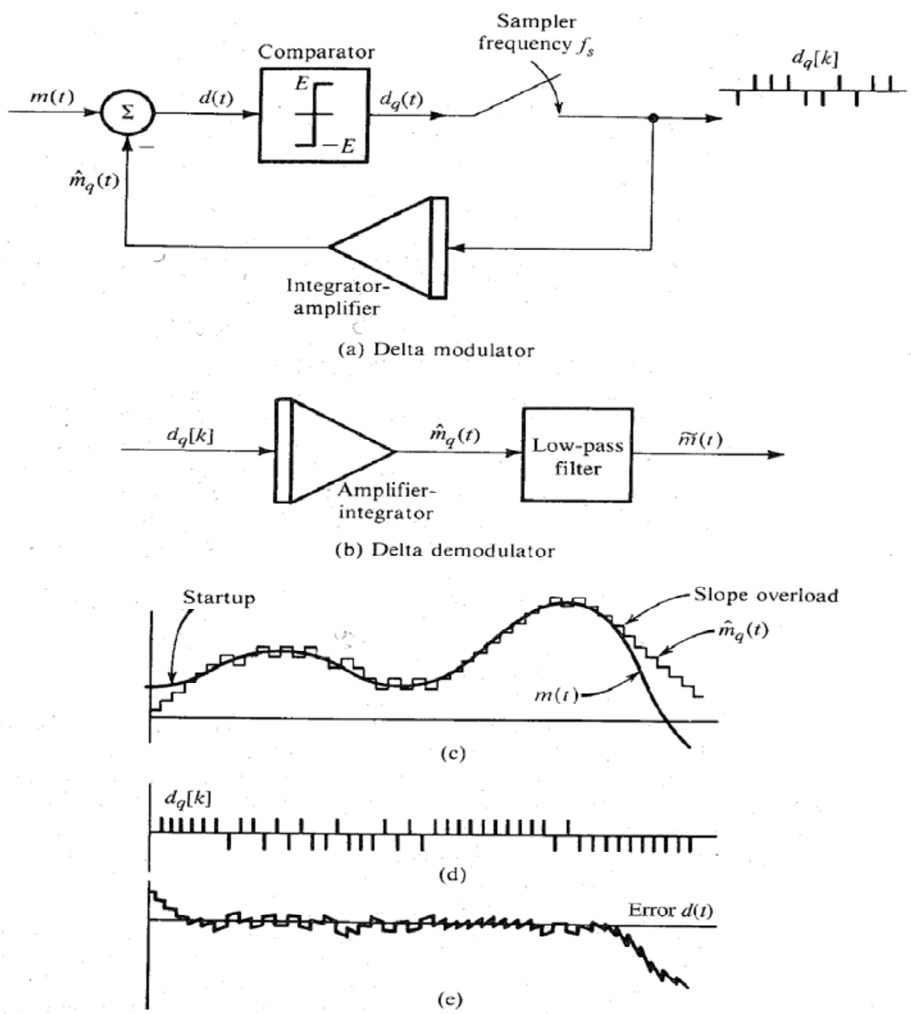


Fig. 2.11 Delta Modulation

Fig 2.11 shows a practical implementation of the delta modulator and demodulator. As discussed earlier, the first order predictor is replaced by a low cost integrator circuit (such as and RC integrator). The modulator consists of a comparator and a sampler in the direct path and an integrator amplifier n the feedback path. Let us see how this delta modulator works.

The analog signal g(t) is compared with the feedback signal(which served as a predicted signal) $\hat{g}_q[k]$. The error signal $d(t) = g(t) - \hat{g}_q[k]$ is applied to a comparator. If d(t) is +ve, the comparator output is a constant signal of amplitude E, and if d(t) is –ve, the comparator output is –E. Thus, the difference is a binary signal [L = 2] that is needed to generate a 1-bit DPCM. The comparator output is sampled by a sampler at a rate of f_s samples per second. The sampler thus produces a train of narrow pulses $d_q[k]$ with a positive pulse when $g(t) > \hat{g}_q[k]$ and a negative pulse when $g(t) < \hat{g}_q[k]$. The pulse train $d_q(t)$ is the delta modulated pulse train. The modulated signal $d_q(t)$ is amplified and integrated in the feedback path to generate $\hat{g}_q[k]$ which tries to follow g(t).

To understand how this works we note that each pulse in $d_q[k]$ at the input of the integrator gives rise to a step function (positive or negative depending on pulse polarity) in $\hat{g}_q[k]$. If, eg, $g(t) > \hat{g}_q[k]$, a positive pulse is generated in $d_q[k]$, which gives rise to a positive step in $\hat{g}_q[k]$, trying to equalize $\hat{g}_q[k]$ to g(t) in small steps at every sampling instant as shown in fig 2.11. It can be seen that $\hat{g}_q[k]$ is a kind of staircase approximation of g(t). The demodulator at the receiver consists of an amplifier integrator (identical to that in the feedback path of the modulator) followed by a low pass filter.

- **DM transmits the derivative of g(t)**

  In DM, the modulated signal carries information not about the signal samples but about the difference between successive samples. If the difference is positive or negative a positive or negative pulse (respectively) is generated in the modulated signal $d_q[k]$. Basically, therefore, DM carries the information about the derivative of g(t) and , hence, the name delta modulation. This can also be seen from the face that integration of the delta modulated signal yields $g_q(t)$, which is an approximation of g(t).

- **Threshold of coding and overloading**

  Threshold and overloading effects can be clearly seen in fig 2.11c. Variation in g(t) smaller than the step value(threshold coding) are lost in DM. Moreover, if g(t) changes too fast ie, $\hat{g}_q[k]$ is too high, $\hat{g}_q[k]$ cannot follow g(t), and overloading occurs. This is the so called slope overload which gives rise to slope overload noise. This noise is one of the basic limiting factors in the performance of DM. We should expect slope overload rather than amplitude overload in DM, because DM basically carries the information about $\hat{g}_q[k]$. The granular nature of the output signal gives rise to the granular noise similar to the quantization noise. The slope overload noise can be reduced by increasing the step size Δ. This unfortunately increases granular noise. There is an

optimum value of Δ, which yields the best compromise giving the minimum overall noise. This optimum value of Δ depends on the sampling frequency f_s and the nature of the signal.
The slope overload occurs when $\hat{g}_q[k]$ cannot follow g(t). During the sampling interval Ts, $\hat{g}_q[k]$ is capable of changing by Δ, where Δ is the height of the step Hence, the maximum slope that $\hat{g}_q[k]$ can follow is $\Delta/T_{s=,}$ or Δf_s, where f_s is the sampling frequency. Hence, no overload occurs if

$$|\dot{g}(t)| < \Delta f_s \qquad (2.30)$$

Consider the case of a single tone modulation,

ie, $g(t) = A.\cos(wt)$

The condition for no overload is

$$|\dot{g}(t)|_{max} = wA < \Delta f_s \qquad (2.31)$$

Hence, the maximum amplitude 'A_{max}' of this signal that can be tolerated without overload is given by

$$A_{max} = \Delta f_s/W \qquad (2.32)$$

The overload amplitude of the modulating signal is inversely proportional to the frequency W. For higher modulating frequencies, the overload occurs for smaller amplitudes. For voice signals, which contain all frequency components up to(say) 4 KHz, calculating A_{max} by using W = 2.pi.4000 in eqn.(2.32) will give an overly conservative value. It has been shown by De Jager that 'A_{max}' for voice signals can be calculated by using W_r = 2.pi.800 in eqn.(2.32),

$$[A_{max}]_{voice} \approx \Delta f_s/w_r \qquad (2.33)$$

Thus, the maximum voice signal amplitude 'A_{max}' that can be used without causing slope overload in DM is the same as the maximum amplitude of a sinusoidal signal of reference frequency f_r(f_r = 800 Hz) that can be used without causing slope overload in the same system.

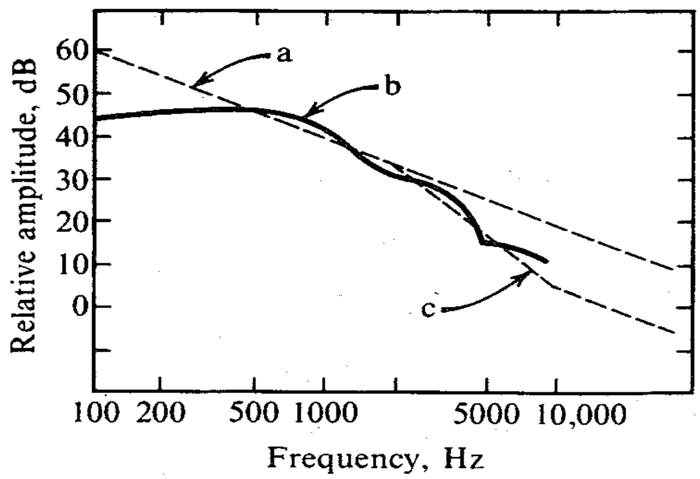


Fig. 2.12 Voice Signal Spectrum

Fortunately, the voice spectrum (as well as the TV video signal) also decays with frequency and closely follows the overload characteristics (curve c, fig 2.11). For this reason, DM is well suited for voice (and TV) signals. Actually, the voice signal spectrum (curve b) decrease as 1/W upto 2000 Hz, land beyond this frequency, it decreases as $1/W^2$. Hence, a better match between the voice spectrum and the overload characteristics is achieved by using a single integration up to 2000 Hz and a double interaction beyond 2000 Hz. Such a circuit (the double integration) is fast responding, but has a tendency to instability, which can be reduced by using some lower order prediction along with double integration. The double integrator can be built by placing in cascade tow low pass RC integrators with the time constant R_1C_1 = 1/2000.pi and R_2C_2 = 1/4000.pi, respectively. This result in single integration from 100 Hz to 2000 Hz and double integration beyond 2000 Hz.

- **Adaptive Delta Modulation**

The DM discussed so far suffers from one serious disadvantage. The dynamic range of amplitudes is too small because of the threshold and overload effects discussed earlier. To correct this problem, some type of signal compression is necessary. In DM a suitable method appears to be the adaptation of the step value 'Δ' according to the level of the input signal derivative. For example in fig.2.11c when the signal g(t) is falling rapidly, slope overload occurs. If we can increase the step size during this period, this could be avoided. On the other hand, if the slope of g(t) is small, a reduction of step size will reduce the threshold level as well as the granular noise. The slope overload causes dq[k] to have several pulses of same polarity in succession. This call for increased step size. Similarly, pulses in dq[k] alternating continuously in polarity indicates small amplitude variations, requiring a reduction in step size. This results in a much larger dynamic range for DM.

- **Output SNR**

The error d(t) caused by the granular noise in DM, (excluding slope overload), lies in the range (-Δ,Δ) , where Δ is the step height in $g_q(t)$. The situation is similar to that encountered in PCM, where the quantization error amplitude was in the range from –Δ/2 to Δ/2. The quantization noise is,

$$\langle q^2_e\rangle = 1/\Delta\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} q_e^2 dq_e = \Delta^2/12 \qquad (2.34)$$

Similarly the granular noise power $\langle g^2_n\rangle$ is

$$\langle g^2_n\rangle = 1/(2A) \int_{-\Delta}^{\Delta} g_n^2 . dg_n = \Delta^3/3 \qquad (2.35)$$

The granular noise PSD has continuous spectrum, with most of the power in the frequency range extending well beyond the sampling frequency 'fs'. At the output, most of this will be suppressed by the baseband filter of bandwidth W. Hence the granular noise power N_0 will be well below that indicated in equation (18). To compute N_0 we shall assume that PSD of the quantization noise is uniform and concentrated in the band of 0 to fs Hz. This assumption has been verified experimentally. Because the total power $\Delta^3/3$ is uniformly spread over the bandwidth f_s, the power within the baseband W is

$$N_0 = (\Delta^3/3)W/f_s = \Delta^2 . W/(3f_s) \qquad (2.36)$$

The output signal power is $S_0 = \langle g^2(t)\rangle$. Assuming no slope overload distortion

$$S_0/N_0 = 3.f_s\langle g^2(t)\rangle/(\Delta^2.W) \qquad (2.37)$$

If g_p is the peak signal amplitude, then eqn. (2.33) an be written as,

$$g_p = \Delta f_s/W_r$$

& $$S_0/N_0 = 3.f^3_s\langle g^2(t)\rangle/(W^2_r.W.g^2_p) \qquad (2.38)$$

Because we need to transmit f_s pulses per second, the minimum transmission bandwidth $B_T = f_s/2$. Also for voice signals, W=4000 and W_r =2.pi.800 =1600.pi. Hence,

$$S_0/N_0 = [3.(2B_T)^3\langle g^2(t)\rangle]/[1600x1600.\pi^2 W g^2_p]$$
$$=150/\ \pi^2.(B_T/W)^3.\langle g^2(t)\rangle/g^2_p \qquad (2.39)$$

Thus the output SNR varies as the cube of the bandwidth expansion ratio B_T/W. This result is derived for the single integration case. For double integration DM, Greefkes and De Jager have shown that,

$$S_0/N_0 = 5.34(B_T/W)^5\langle g^2(t)\rangle/g^2_p \qquad (2.40)$$

It should be remembered that these results are valid only for voice signals. In all the preceding developments, we have ignored the pulse detection error at the receiver.

- **Comparison With PCM**

The SNR in DM varies as a power of B_T/W, being proportional to $(B_T/W)^3$ for single integration and $(B_T/W)^5$ for double integration. In PCM on the other hand the SNR varies exponentially with B_T/W. Whatever the initial value, the exponential will always outrun the power variation. Clearly for higher values of B_T/W, PCM is expected to be superior to DM. The output SNR for voice signals as a function of the bandwidth expansion ratio B_T/W is plotted in fig. for tone modulation, for which $\langle g^2 \rangle / g_p^2 = 0.5$. The transmission band is assumed to be the theoretical minimum bandwidth for DM as well as PCM. It is clear that DM with double integration has a performance superior to companded PCM(which is the practical case) for lower valued of $B_T/W = 10$. In practice, the crossover value is lower than 10, usually between 6 & 7(f_s = 50 kbits/s). This is true only for voice and TV signals, for which DM is ideally suited. For other types of signals, DM does not comparable as well with PCM.

Because the DM signal is digital signal, it has all the advantages of digital system, such as the use of regenerative repeaters and other advantages as mentioned earlier. As far as detection of errors are concerned, DM is more immune to this kind of error than PCM, where weight of the detection error depends on the digit location; thus for n=8, the error in the first digit is 128 times as large as the error in the last digit.

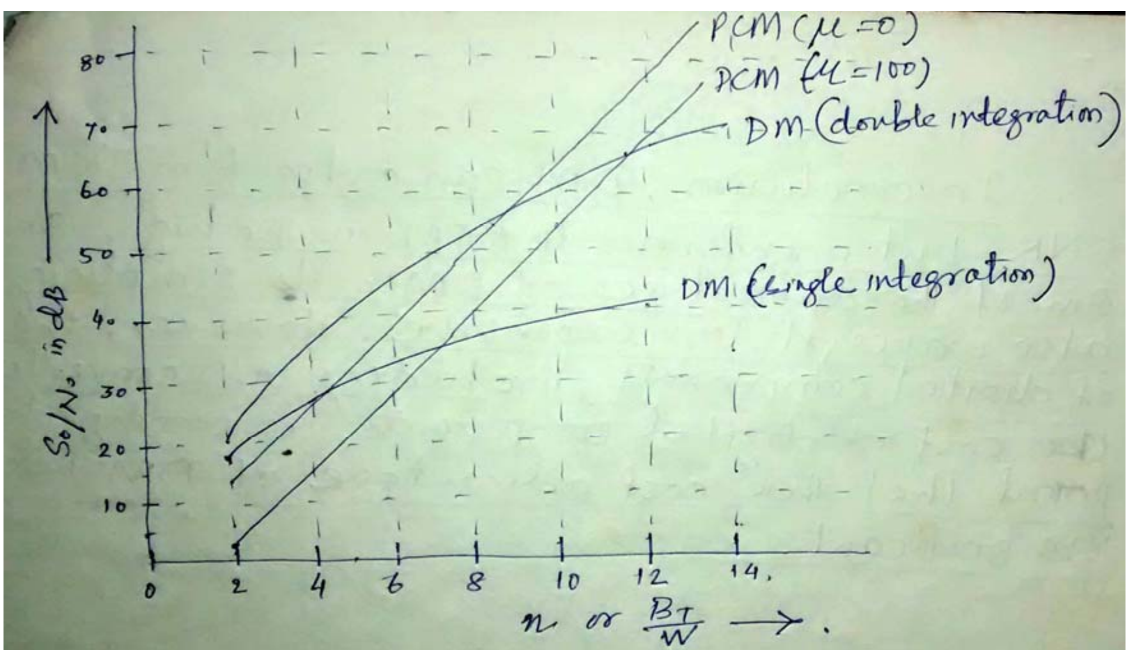


Fig. 2.21a Comparison of DM and PCM.

For DM, on the other hand, each digit has equal importance. Experiments have shown that an error probability 'Pe' on the order of 10^{-1} does not affect the

intelligibility of voice signals in DM, where as 'Pe' as low as 10^{-4} can cause serious error, leading to threshold in PCM. For multiplexing several channels, however, DM suffers from the fact that each channel requires its own coder and decoder, whereas for PCM, one coder and one decoder are shared by each channel. But his very fact of an individual coder and decoder for each channel also permits more flexibility in DM. On the route between terminals, it is easy to drop one or more channels and insert other incoming channels. For PCM, such operations can be performed at the terminals. This is particularly attractive for rural areas with low population density and where population grows progressively. The individual coder-decoder also avoids cross-talk, thus alleviating the stringent design requirements in the multiplexing circuits in PCM.
In conclusion, DM can outperform PCM at low SNR, but is inferior to PCM in the high SNR case. One of the advantages of DM is its simplicity, which also makes it less expensive. However, the cost of digital components, including A/D converters, ie, coming down to the point that the cost advantage of DM becomes insignificant.

- **Noise in PCM and DM**

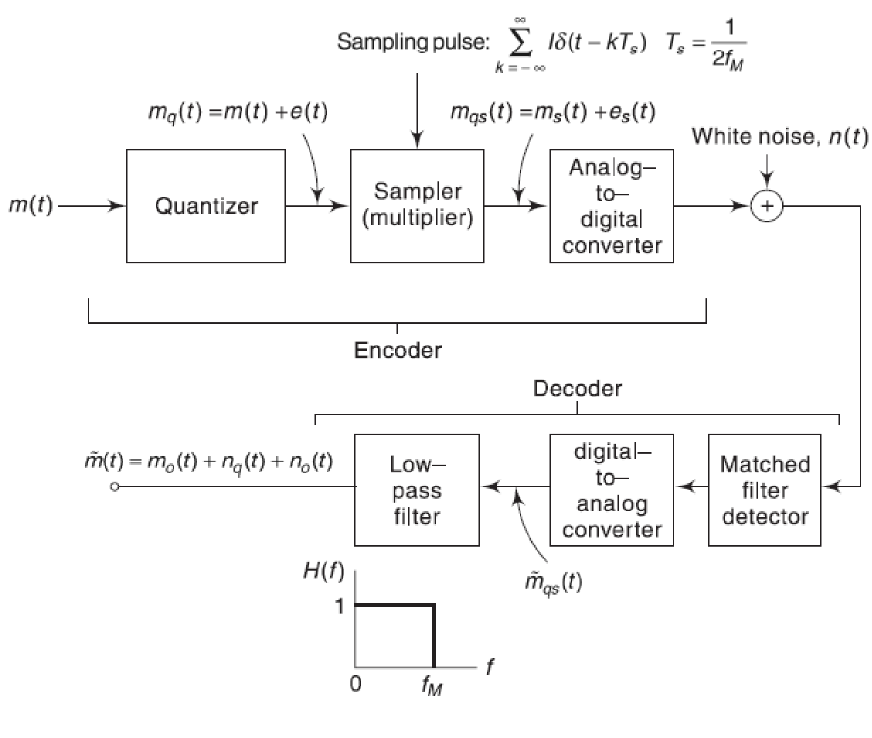


Fig. 2.13 A binary PCM encoder-decoder.

In the above figure m(t) is same as g(t). The baseband signal g(t) is quantized, giving rise to quantized signal $g_q(t)$, where

$$g_q(t) = g(t) + e(t) \qquad (2.41)$$

(e(t) is same as $q_e(t)$ as discussed earlier).
The sampling interval is $T_s=1/2f_m$, where f_m is the frequency to which the signal g(t) is bandlimited.

The sampling pulses considered here are narrow enough so that the sampling may be considered as instantaneous. With such instantaneous sampling, the sampled signal may be reconstructed exactly by passing the sequence of samples through a low pass filter with cut off frequency of f_m. Now as a matter of mathematical convenience, we shall represent each sampling pulse as an impulse. The area of such an impulse is called its strength, and an impulse of strength I is written as Iδ(t).
The sampling impulse train is therefore s(t), given by,

$$s(t) = I\sum_{-\infty}^{\infty} \delta(t-k)\, T_s \quad (2.42)$$

Where, $T_s = 1/(2.f_m)$

From equation 1 and 2 , the quantized signal $g_q(t)$ after sampling becomes $g_{qs}(t)$, written as,

$$g_{qs}(t) = g(t)I\sum_{k=-\infty}^{\infty} \delta(t-kT_s) + e(t)I\sum_{k=-\infty}^{\infty} \delta(t-kT_s) \quad (2.43a)$$

$$= g_s(t) + e_s(t) \quad (2.43b)$$

The binary output of the A/D converter is transmitted over a communication channel and arrives at the receiver contaminated as a result of the addition of white thermal noise W(t). Transmission may be direct as indicated in fig.2.13, or the binary output signal may be used to modulate a carrier as in PSK or FSK.
In any event the received signal is detected by a matched filter to minimize errors in determining each binary bit and thereafter passed on to a D/A converter. The output of a D/A converter is called $g_{qs}(t)$. In the absence of thermal noise and assuming unity gain from the input to the A/D converter to the output of the D/A converter, we should have $g\sim_{qs}(t) = g_{qs}(t)$. Finally the signal $g\sim_{qs}(t)$ is passed through the low pass baseband filter. At the output of the filter we find a signal $g_0(t)$ which aside from a possible difference in amplitude has exactly the waveform of the original baseband signal g(t). This output signal however in accompanied by a noise waveform $W_q(t)$ due to thermal noise.

- **<u>Calculation of Quantization Noise</u>**

Let us calculate the output power due to the quantization noise in the PCM system as in fig.2.14 ignoring the effect of thermal noise.
The sampled quantization error waveform, as given by eq^n (2.43b),

$$e_s(t) = e(t)I\sum_{k=-\infty}^{\infty} g(t-kT_s) \quad (2.44)$$

It is to be noted that if the sampling rate is selected to be the nyquist rate for the baseband signal g(t) the sampling rate will be inadequate to allow reconstruction of the error signal e(t) from its sample $e_s(t)$. In fi.2 the quantization levels are separated by amount Δ. We observe that e(t) executes a complete cycle and exhibits an abrupt discontinuity every time g(t) makes an excursion of amount Δ. Hence spectral range of e(t) extends for beyond the band limit f_m of g(t).

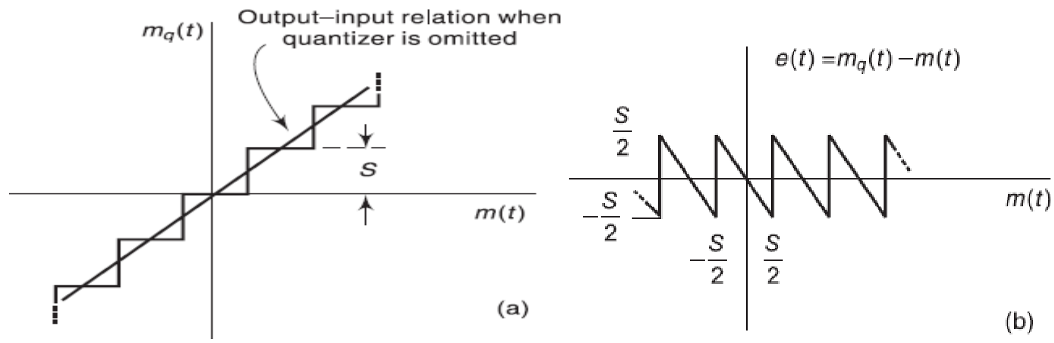


Fig. 2.14 Plot of $m_q(t)$ and e(t) as a function of m(t).

To find the quantization noise output power N_q, we require the PSD of the sampled quantization error $e_s(t)$ given in eq^n (2.44).

Since $\delta(t-kT_s) = 0$ except when $t=kT_s$ $e_s(t)$ may be written as,

$$e_s(t) = I.\sum_{k=-\infty}^{\infty} e(kT_s)\delta(t - kT_s) \quad (2.45)$$

The waveform of eq^n (2.45) consists of a sequence of impulses of area=A=$e(kT_s)$ I occurring at intervals T_s. The quantity $e(kT_s)$ is the quantization error at sampling time and is a random variable.
The PSD $G_{es}(f)$ of the sampled quantization error is,

$$G_{e_s}(f) = \frac{I^2}{T_s}\overline{e^2(kT_s)} \quad (2.46)$$

and,

$$e^2(t) = e^2(kT_s) = \frac{S^2}{12}$$

For a step size of Δ the quantization error is

$$e^2(t) = \Delta^2/12 \quad (2.47)$$

Equation 6 involves $<e^2(kT_s)>$ rather than $<e^2(t)>$. However since the probability density of e(t) does not depend on time the variance of e(t) is equal to the variance of $e(t= kT_s)$.

Thus, $<e^2(t)> = <e^2(kT_s)> = \Delta^2/12$ (2.48)
From eqn. (2.46) and eqn. (2.49) we have,

$$G_{es}(f) = I^2\Delta^2/(T_s.12) \quad (2.49)$$

Finally the quantization noise N_q is, from eqn. (2.50),

$$N_q = \int_{-f_M}^{f_M} G_{e_s}(f)\,df = \frac{I^2}{T_s}\frac{S^2}{12}2f_M \quad (2.50)$$

$$= \frac{I^2}{T_s^2}\frac{S^2}{12}$$

[take 'S' as 'Δ']

The Output Signal Power

The sampled signal which appears at the input to the baseband filter shown in fig.2.14 is given by $g_s(t)$ in $eq^n(2.43)$ as.

$$g_s(t) = g(t).I.\sum_{k=-\infty}^{\infty} \delta(t - kT_s) \qquad (2.51)$$

Since the impulse train is periodic it can be represented by a fourier series. Because the impulses have strength I and are separated by a time T_s, the first term in Fourier series is the dc component which is $1/T_s$. Hence the signal $g_0(t)$ at the output of the baseband filter is

$$g_0(t) = I/T_s.g(t) \qquad (2.52)$$

Since $T_s=1/2f_m$, other terms in the series of equation 11 lie outside the passband of the filter. The normalised signal output power is from eq^n (2.52),

$$\overline{g_0^2(t)} = I^2/T^2 .\overline{g^2(t)} \qquad (2.53)$$

We can now express $\overline{g^2(t)}$ in terms of the number M of quantization levels and the step size Δ. To do this we can say that the signal can vary from -mΔ/2 to mΔ/2, i.e we assume that the instantaneous value of g(t) may fall anywhere in its allowable range of 'mΔ' volts with equal likelihood. Then the probability density of the instantaneous value of g in f(g) given by,

$$f(g) = 1/(M\Delta)$$

The variance σ^2 of g(t), ie, $\overline{g^2(t)}$ is,

$$\overline{g^2(t)} = \int_{-\frac{M\Delta}{2}}^{\frac{M\Delta}{2}} g^2 f(g)dg = M^2.\Delta^2/12 \qquad (2.54)$$

Hence from eqn. (2.53), the output signal power is

$$S_0 = \overline{g_0^2(t)} = I^2/T^2 \,.\, M^2.\Delta^2/12 \qquad (2.55)$$

From eqn.(2.50) and (2.55) we find the signal to quantization noise ratio is

$$S_o/N_q = M^2 = (2^N)^2 \qquad (2.56)$$

where, N is the number of binary digits needed to assign individual binary code designations to the M quantization levels.

The Effects of Thermal Noise

The effect of additive thermal noise is to calculate the matched filter detector of fig.2.14 to make an occasional error in determining whether a binary 1 or binary 0 was transmitted. If the thermal noise is white and Gaussian the probability of such an error depends on the ratio E_b/η. Where E_b is signal energy transmitted during a bit and $\eta/2$ is the two sided power spectral density of the noise. The probability depends also on the type of modulation employed.

Rather typically, PCM system operate with error probabilities which are small enough so that we may ignore the likelihood that more than a single bit error will occur with in a single word. For example, if the error probability $P_e=10^{-5}$ and a word of 8 bits we would expect on the average that 1 word would be in error for every 12500 word transmitted. Indeed the probability of two words being transmitted in error in the same 8 bit word is $28*10^{-10}$.

Let us assume that a code word used to identify a quantization level has N bits. We assume further that the assignment of code words to levels is in the order of numerical significance of the word. Thus we assign 00.....00 to the most negative level to the next higher level until the most positive level is assigned the codeword 1 1.....1 1.

An error which occurs in the least significant bit of the code word corresponds to an incorrect determination by amount 'Δ' in the quantized value $g_s(t)$ of the sampled signal. An error in the next higher significant bit corresponds to an error 2Δ; in the next higher, 4Δ, etc.

Let us call the error δg_s. Then assuming that an error may occur with equal likelihood inany bit of the word, the variance of the error is,

$$<\delta g^2_s> = 1/N.[\, \Delta^2 + (2\Delta)^2 + (4\Delta)^2 + \text{..........}+(2^{N-1}\Delta)^2]$$
$$= \Delta^2/N.[\, 1^2 + (2)^2 + (4)^2 + \text{..........}+(2^{N-1})^2] \qquad (2.57)$$

The sum of the geometric progression in eqn.(2.57),

$$<\delta g^2_s> = \Delta^2/N.2^{(2N-1)}/(2^2-1) = 2^{2N}.\Delta^2/(3N), \quad \text{for } N \geq 2 \qquad (2.58a)$$

The preceding discussion indicates that the effect of thermal noise errors may be taken into account by adding at the input to the A/D converter in fig. 2.14, an error voltage δg_s , and by detecting the white noise source and the matched filter. We have assumed unity gain from the input to the A/D converter to the output of the D.A converter. Thus the same error voltage appears at the input to the lowpass baseband filter. The results of a succession of errors is a train of impulses, each of strength $I(\delta g_s)$. These impulses are of random amplitude and of random time of occurrence.

A thermal noise error impulse occurs on each occasion when a word is in error. With P_e the probability of a bit error, the mean separation between bits which are in errors is $1/P_e$.

With N bits per word , the mean separation between words which are in error is $1/N P_e$ words. Words are separated in time by the sampling interval T_s. Hence the mean time between words which are in error is T, given by

$$T = \frac{T_s}{NP_e} \qquad (2.58b)$$

The power spectral density of the thermal noise error impulses train is, using eqn.(2.58a) and(2.58b),

$$G_{th}(f) = I^2/T < \delta.g_s^2 > = NP_e I^2/T_s <\delta.g_s^2> \qquad (2.59)$$

using eqn.(2.58a), we have

$G_{th}(f) = 2^{2N}\Delta^2 P_e I^2/(3T^2_e)$ (2.60)

Finally, the output power due to the thermal error noise is,

$N_{th} = \int_{-fm}^{fm} G_{th}(f)df \ = 2^{2N}.\Delta^2 P_e I^2/(3.T_s^{\ 2})$ (2.61)

- **<u>Output Signal To Noise Ratio in PCM</u>**

The output SNR including both quantization and thermal noise , is found by combining equation 10,16 and 23. The result is

$$\frac{S_o}{N_o} = \frac{S_o}{N_q + N_{th}} = \frac{(I^2/T_s^2)(M^2 S^2/12)}{(I^2/T_s^2)(S^2/12) + (I^2/T_s^2)(P_e 2^{2N} S^2/3)}$$

[replace 'S' by 'Δ'; S is same as Δ]

$$= \frac{2^{2N}}{1 + 4P_e\ 2^{2N}} \quad (2.62)$$

In PSK(or for direct transmission) we have,

$$(P_e)_{\text{PSK}} = \frac{1}{2}\operatorname{erfc}\sqrt{\frac{E_b}{\eta}} \quad (2.63)$$

Where, E_b is the signal energy of a bit and η/2 is the two sided thermal noise power spectral density. Also, for coherent reception of FSK we have,

$$(P_e)_{\text{FSK}} = \frac{1}{2}\operatorname{erfc}\sqrt{0.6\frac{E_b}{\eta}} \quad (2.64)$$

To calculate $E_{b,}$ we note that if a sample is taken at intervals of T_s and the code word of N bit occupies the total interval between samples, then a bit has a duration T_s/N. If the received signal power is S_i , energy associated with a single bit is

$$E_b = S_i\frac{T_s}{N} = S_i\frac{1}{2f_M N} \quad (2.65)$$

Combining eqns. (2.62), (2.63) & (2.65), we find,

$$\left(\frac{S_o}{N_o}\right)_{\text{PSK}} = \frac{2^{2N}}{1 + 2^{2N+1}\operatorname{erfc}\sqrt{(1/2N)(S_i/\eta f_M)}} \quad (2.66)$$

using eqn.(2.64) in place of eqn.(2.63), we have

$$\left(\frac{S_o}{N_o}\right)_{\text{FSK}} = \frac{2^{2N}}{1 + 2^{2N+1}\operatorname{erfc}\sqrt{(0.3/N)(S_i/\eta f_M)}} \quad (2.67)$$

Note that for $S_i/\eta f_M \gg 1$ and $N = 8$

$$\left(\frac{S_o}{N_o}\right)_{\text{PSK, FSK}} = 10 \log (2^{16}) = 48 \text{ dB} \qquad (2.68)$$

From fig. we find both the PCM system exhibit threshold, FSK threshold occurring at a $S_i/\eta f_m$ which is 2.2 dB greater than that for PSK. Experimentally, the onset of threshold in PCM is marked by an abrupt increase in a crackling noise analogous to the clicking noise heard below threshold in analogue FM systems.

- **Delta Modulation:**

A delta modulation system including a thermal noise source is shown in fig.2.15. The impulse generator applies the modulator a continuous sequence of impulses $p_i(t)$ of time separation τ. The modulator output is a sequence of pulses $P_0(t)$ whose polarity depends on the polarity of the difference signal $\delta(t)=g(t) - \tilde{g}(t)$, where $\tilde{g}(t)$ is the integrator output. We assume that the integrator has been adjusted so that its response to an input impulse of strength I is a step size Δ; i.e. $\tilde{g}(t) = (\Delta/I)\int P_0(t)dt$.

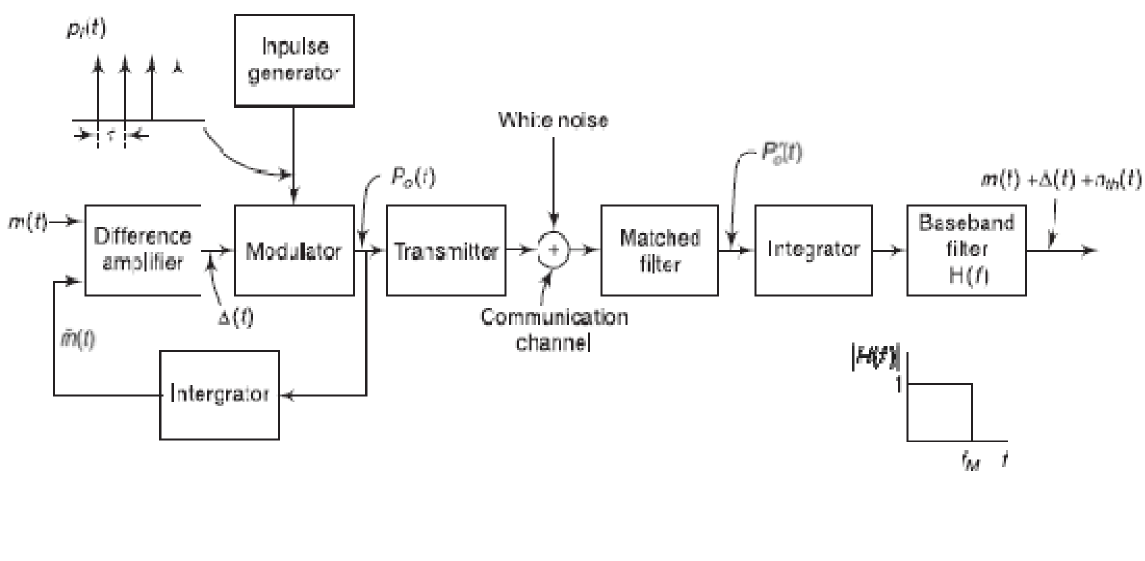


Fig. 2.15 A delta modulation system.

A typical impulse train $P_0(t)$ is shown in fig.2.16(a). Before transmission, the impulse waveform will be converted to the two level waveform of fig.2.16(b). Since this latter waveform has much greater power than a train of narrow pulses. This conversion is accomplished by the block in fig.2.15 marked “transmitter”. The transmitter in principle need be nothing more complicated than a bistable multivibrator. We may readily

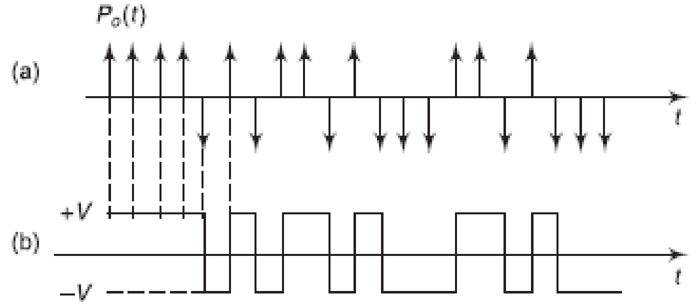


Fig.2.16 (a) A typical impulse train $p_0(t)$appearing at the modulator output in previous fig.
(b) The two-level signal transmitted over the communication channel.

arrange that two positive impulses set the flip-flop into one of its stable states, while the negative impulses reset the flip-flop to its other stable state. The binary waveform of fig.2.16(b) will be transmitted directly or used to modulate as a carrier in FSK or PSK. After detection by the matched filter shown in fig.2.15, the binary waveform will be reconverted to a sequence of impulses $P_0'(t)$. In the absence of thermal noise $P_0'(t)=P_0(t)$, and the signal $\tilde{g}(t)$ is recovered at the receiver by passing $P_0'(t)$ through an integrator. We assume that transmitter and receiver integrators are identical and that the input to each consists of a train of impulses of strength +I or -I. Hence in the absence of thermal noise , the output of both the integrators are identical.

➢ Quantization Noise in Delta Modulation

Here in fig. 2.17 $\tilde{g}(t)$ in the delta modulator approximation to g(t). Fig 2.17 shows the error waveform δ(t) given by,

$$\delta(t) = g(t) - \hat{g}(t) \qquad (2.69)$$

This error waveform is the source for quantization noise.

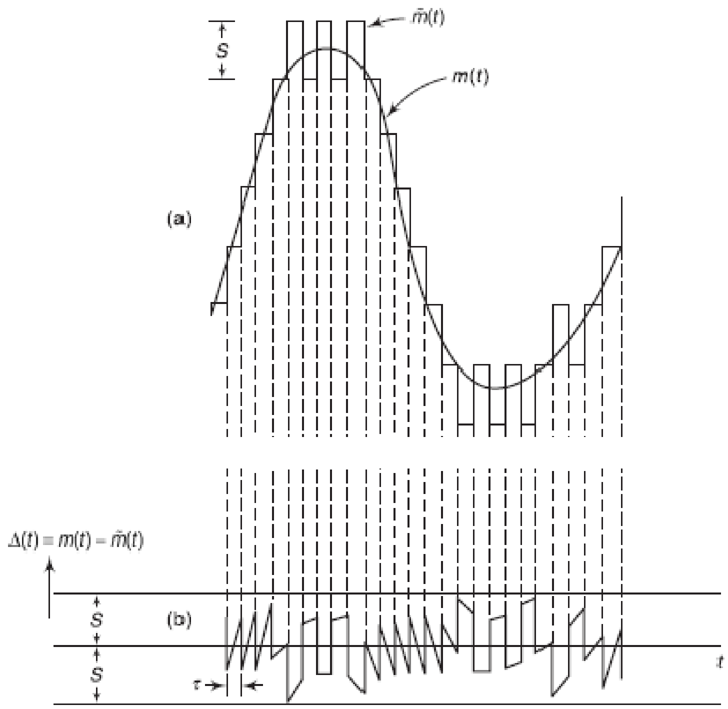


Fig. 2.17 The estimate $\hat{g}$(t) and error Δ(t) when g(t) is sinusoidal.

We observe that, as long as slope overloading is avoided, the error δ(t) is always less than the step size Δ. We shall assume that δ(t) takes on all values between -Δ and +Δ with equal likelihood. So we can assume the probability δ(t) is,

$$f(\delta) = 1/(2\Delta), \qquad -\Delta \le \delta(t) \le \Delta \tag{2.70}$$

The normalization power of the waveform δ(t) is then,

$$<[\delta(t)]^2> = \int_{-\Delta}^{\Delta} f(\delta)\, \delta^2 d\delta \;=\; \Delta^2/3 \tag{2.71}$$

Our interest is in estimating how much of this power will pass through a baseband filter. For this purpose we need to know something about the PSD of δ(t) .

In fig. 2.17 the period of the sinusoidal waveform g(t) i.e. T has been selected so that T is an integral multiple of step duration τ. We then observe that the δ(t) is periodic with fundamental period T, and is of course, rich in harmonics. Suppose, however, that the period T is charged very slightly by amount δT. Then the fundamental period of δ(t) will not be T but will be instead T * τ/δT corresponding to a fundamental frequency near zero as δT tends to 0. And again, of course δ(t) will be rich in harmonics. Hence, in the general case, especially with g(t) a random signal, it is reasonable to assume that δ(t) has a spectrum which extends continuously over a frequency which begins near zero.

To get some idea of the upper frequency range of the spectrum of the waveform δ(t). Let us contemplate passing δ(t) through a LPF of adjustable cutoff frequency. Suppose that initially the cutoff frequency is high enough so that δ(t) may pass with nominally no distortion. As we lower the cutoff frequency, the first type of distortion we would note is that the abrupt discontinuities in the waveform would exhibit finite rise and fall times. Such is the case since it is the abrupt changes which contribute the high frequency power content of the signal. To keep the distortion within reasonable limits, let us arrange that the rise time be rather smaller than the interval τ. To satisfy this condition we require the filter cutoff frequency f_c be of the order of $f_c=1/\tau$, since the transmitted bit rate $f_b=1/\tau$, $f_c=f_b$ as expected.

We now have made it appear reasonable, by a rather heuristic arguments that the spectrum of δ(t) extends rather continuously from nominally zero to $f_c = f_b$. We shall assume further that over this range the spectrum is white. It has indeed been established experimentally that the spectrum of δ(t) is approximately white over the frequency range indicated.

We may now finally calculate the quantization noise that will appear at the output of a baseband filter of cutoff frequency f_m. Since the quantization noise power in a frequency range f_b is $\Delta^3/3$ as given by equation 32, the output noise power in the baseband frequency range f_m is

$$N_q = \frac{S^2}{3}\frac{f_M}{f_b} = \frac{S^2}{3}\frac{f_M}{f_b} \quad \text{[replace 'S' with '}\Delta\text{']} \qquad (2.72)$$

We may note also, in passing, that the two-sided power spectral density of δ(t) is,

$$G_\delta(f) = \Delta^2/(3.2.f_b) = \Delta^2/(6.f_b), \qquad -f_b \le f \le f_b \qquad (2.73)$$

➢ The Output Signal Power

In PCM, the signal power is determined by the step size and the number of quantization levels. Thus, with step size Δ and M levels, the signal could make excursion only between -MΔ/2 and MΔ/2. In delta modulation there is no similar restriction on the amplitude of the signal waveform, because the number of levels is not fixed. On the other hand, in delta modulation there is a limitation on the slope of the signal wave form which must be observed if slope overload is to be avoided. If however, the signal waveform changes slowly , there is normally no limit to the signal power which may be transmitted.

Let us consider a worst case for delta modulation. We assume that the signal power is concentrated at the upper end of the baseband. Specifically let the signal be,

$$g(t) = A.\sin(w_m t)$$

With 'A' the amplitude and $\omega_m = 2\pi f_m$, where f_m is the upper limit of the baseband frequency range. Then the output signal power

$$S_0(t) = \overline{g^2(t)} = A^2/2 \qquad (2.74a)$$

The maximum slope of g(t) is $\omega_m A$. The maximum average slope of the delta modulator approximation $\tilde{g}(t)$ is $\Delta/\tau = \Delta f_b$, where Δ is step size and f_b the bit rate. The limiting value of 'A' just before the onset of slope overload is, therefore given by the condition,

$$w_M \cdot A = \Delta f_b \tag{2.74b}$$

From eqns.(2.74a) and (2.74b), we have that the maximum power which may be transmitted in,

$$S_0 = \Delta^2 f^2_b/(2w^2_M) \tag{2.75}$$

The condition specified in equation 37 is unduly severe. A design procedure, more often employed, is to select the Δf_b product to be equal to the rms value of the slope g(t). In this case the output signal power can be increased above the value given in equation 38.

➢ Output Signal to Quantization Noise Ratio for Delta Modulation

The output signal to quantization noise ratio for delta modulation is found by dividing eqn.(2.75) by eqn.(2.72). The result is

$$\frac{S_o}{N_q} = \frac{5}{8\pi^2}\left(\frac{f_b}{f_M}\right)^3 \cong \frac{3}{80}\left(\frac{f_b}{f_M}\right)^3 \tag{2.76}$$

It is of interest to note that when our heuristic analysis is replaced by a rigorous analysis, it is found that eqn. 39 continues to apply, except with a factor 3/80 replaced by 3/64, corresponding to a difference of less than 1dB.

The dependence of S_0/N_q on the product f_b/f_m should be anticipated. For suppose that the signal amplitude were adjusted to the point of slope overload, if now, say, f_m were increased by some order to continue to avoid overload.

Let us now make a comparison of the performance of PCM and DM in the matter of the ratio S_0/N_q. We observe that the transmitted signals in DM and in PCM are of the same waveform, a binary pulse train. In PCM a voltage level, corresponding to a single bit persists for the time duration allocated to one bit of codeword. With sampling at the Nyquist rate $1/2f_m$ s , and with N bits per code word , the PCM bit rate is $f_b=2f_mN$. In DM, a voltage corresponding to a single bit is held for a duration τ which is the interval between samples. Thus the DM system operates at a bit rate $f_b=1/\tau$.

If the communication channel is of limited bandwidth, then there is a possibility of interference in either DM or PCM. Whether such inter-symbol interference occurs in DM depends on the ratio of f_b to the bandwidth of the channel and similarly in PCM on the ratio of f_b to the channel bandwidth. For a fixed channel bandwidth, if inter-symbol interference is to be equal in the two cases, DM or PCM , we require that both systems operate at the same bit rate or

$$f_b=f'_b=2f_mN \tag{2.77}$$

Combining eq 17 and 40 for PCM yields

$$S_0/N_q=2^{2N}=2^{fb/fm} \tag{2.78}$$

Combining eq 39 and 40 for delta modulation yields

$$S_0/N_q = N^3\,(3/\pi^2) \tag{2.79}$$

Comparing equation 41 with 42 , we observe that for a fixed channel bandwidth the performance of DM is always poorer than PCM. For example if a channel is adequate to accommodate code words in PCM with N=8, equation 41 gives S_0/N_q = 48dB. The same channel used for DM would, from equation 42 yield S_0/N_q =22dB.

- **Comparison of DM and PCM for Voice**

when signal to be transmitted is the waveform generated by voice, the comparison between DM and PCM is overly pessimistic against DM. For as appears in the discussion leading to equation 37, in our concern to avoid slope overload under any possible circumstances, we have allowed for the very worst possible case. We have provided for the possibility that all the signal power might be concentrated at the angular frequency ω_m which is the upper edge of the signal bandwidth. Such is certainly not the case for voice. Actually for speech a bandwidth f_m = 3200Hz is adequate and the voice spectrum has a pronounced peak at 800Hz = $f_m/4$. If we replace ω_m by $\omega_m/4$ in eqn. (2.74b) we have,

$$w_M\,.A/4 = \Delta.f_b$$

The amplitude 'A' will now be four times larger than before and the allowed signal power before slope overload will be increased by a factor of it(12dB). Correspondingly, equation 39 now becomes,

$$S_0/N_q = 6/\pi^2.(f_b/f_M)^3 = 0.6(f_b/f_M)^3 = 5N^3 \tag{2.80}$$

It may be readily verified that for $(f_b/f_m)\leq 8$ the signal to noise ratio for DM , SNR(δ), given by eqn.(2.80) is larger than SNR(PCM) given by eq^n (2.78). At about (f_b/f_m) = 4 the ratio SNR(DM)/ SNR(PCM) has maximum value 2.4 corresponding to 3.8db advantage. Thus if we allow f_m = 4KHz for voice, then to avail ourselves of this maximum advantage offered by DM we would take f_b = 16KHz.

In our derivation of the SNR in PCM we assumed that at all times the signal is strong enough to range widely through its allowable excursion. As a matter of fact, we specifically assumed that the distribution function f(g) for the instantaneous signal value g(t) was uniform throughout the allowable signal range. As a matter of practice, such would hardly be the case. The commercial PCM systems using companding, are designed so that the SNR remains at about 30dB over a 40dB range of signal power. In short while eq^n (2.78) predicts a continuous increase in SNR(PCM) with increase in f_b/f_m, this result is for uncompanded PCM and in practice SNR(PCM) is approximately constant at 30dB. The linear DM discussed above has a dynamic range of 15dB. In order to widen this dynamic range to 40dB one employs adaptive DM(ADM), which yields advantages similar to the companding of PCM. When adaptive DM is employed, the SNR is comparable to the SNR of companded PCM. Today the satellite business system employs ADM operating at 32kb/s rather than companded PCM which operates at 64kb/s thereby providing twice as many voice channels in a given frequency band.

- **The Effect of Thermal Noise in DM**

When thermal noise is present, the matched filter in the receiver will occasionally make an error in determining the polarity of the transmitted waveform. Whenever such an error occurs , the received impulse stream $P_0'(t)$ will exhibit an impulse of incorrect polarity. The received impulse stream is then

$$P_0'(t)=P_0(t) + P_{th}(t) \qquad (2.81)$$

In which $P_{th}(t)$ is the error impulse stream due to thermal noise. If the strength of the individual impulses is I, then each impulse in P_{th} is of strength 2I and occurs only at each error. The factor of two results from the fact that an error reverses the polarity of the impulse.

The thermal error noise appears as a stream of impulses of of random time of occurrence and of strength ±2I. The average time of separation between these impulses is τ/P_e, where P_e is the bit error probability and τ is the time duration of a bit. The PSD of thermal noise impulses is

$$G_{pth}(f) = \frac{Pe}{\tau}(2I)^2 \qquad (2.82)$$

Now the integrators (assumed identical in both the DM transmitter and receiver) as having the property that when the input is an impulse of strength the output is a step of amplitude Δ is

$$F\{\Delta u(t)\} = \Delta/j\omega \quad ; \ \omega\neq 0$$
$$= \Delta\pi\delta(\omega) \quad ; \ \omega=0 \qquad (2.83)$$

We may ignore the dc component in the transform since such dc components will not be transmitted through the baseband filter. Hence we may take the transfer function of the integrator to be $H_i(f)$ given by

$$H_i(f) = \frac{\Delta}{I}\frac{1}{j\omega} \quad ; \quad \omega \neq 0 \ - \qquad (2.84)$$

And $$|H_i(f)|^2 = \left(\frac{\Delta}{I}\right)^2\frac{1}{\omega^2} \quad ; \omega=0 \qquad (2.85)$$

From equation 46 and 49 we find that the PSD of the thermal noise at the input to the baseband filter is $G_{th}(f)$ given by

$$G_{th}(f) = |H_i(f)|^2 G_{pth}(f) = \frac{4\Delta^2 Pe}{\tau\omega^2} \qquad (2.86)$$

It would now appear that to find the thermal noise output, we need not to integrate $G_{th}(f)$ over the passband of the baseband filter. During integration we have extended the range of integration from $-f_m$ through f=0 to $+f_m$, even though we recognised that baseband filter does not pass dc and eventually has a low frequency cutoff f_l. However in other cases the PSD of the noise near f=0 is not inordinately large in comparison with the density throughout the baseband range generally. Hence, it as is normally the case, $f_l \ll f_m$, the procedure is certainly justified as a good approximation. We observe however that in the

present case [eq^n (2.86)], G_{th} (f) $\rightarrow \infty$ at $\omega \rightarrow 0$, and more importantly that the integral of $G_{th}(f)$, over a range which include $\omega \rightarrow 0$,is infinite. Let us then explicitly take account of the low frequency cutoff f_l of the baseband filter. The thermal noise output is using eq^n (2.86) with $\omega = 2\pi f$ and since $f_b = \frac{1}{\tau}$,

$$N_{th} = \frac{\Delta^2 P_e}{\pi^2 \tau}\left(\int_{-f_m}^{-f_l} \frac{df}{f^2} + \int_{f_l}^{f_m} \frac{df}{f^2}\right) \tag{2.87}$$

$$= \frac{2\Delta^2 P_e}{\pi^2 \tau}\left(\frac{1}{f_l} - \frac{1}{f_m}\right) \tag{2.88}$$

$$= \frac{2\Delta^2 P_e}{\pi^2 \tau f_l} = \frac{2\Delta^2 P_e f_b}{\pi^2 f_l} \tag{2.89}$$

If $f_l \ll f_m$, unlike the situation encountered in all other earlier cases, the thermal noise output in delta modulation depends upon the low frequency cutoff rather than the higher frequency limit of the baseband range. In many application such as voice encoder where the voice signal is typically band limited from 300 to 3200 H_z, the use of band pass output filter(f_l=300H_z) is common place.

➢ <u>Output Signal-to-Noise ratio in DM</u>

The o/p SNR is obtained by combining eq^n (2.72), (2.80) and (2.89), the result is

$$\frac{S_0}{N_0} = \frac{S_0}{N_q + N_{th}} = \frac{(2\Delta^2/\pi)(f_b/f_m)^2}{(\Delta^2 f_m/3f_b) + (2\Delta^2 P_e f_b/\pi^2 f_l)} \tag{2.90}$$

Which may be written as

$$\frac{S_0}{N_0} = \frac{0.6(f_b/f_m)^3}{1 + 0.6P_e(f_b^2/f_m f_l)} \tag{2.91}$$

If transmission is direct or by means of PSK,

$$P_e = \frac{1}{2} erfc\sqrt{E_s/\eta} \tag{2.92}$$

Where E_s is the signal energy is a bit, is related to the received signal power S_i

By $E_s = S_i T_b = S_i/f_b$ (2.93)

Combining eq^n (2.91), (2.92) and (2.93), we have

$$\frac{S_0}{N_0} = \frac{0.6(f_b / f_m)^3}{1 + [0.3 f_b^2 / f_m f_1] \text{erfc} \sqrt{S_i / \eta f_b}} \tag{2.94}$$

➢ Comparison of PCM and DM

We can now compare the output signal SNR I PCM and DM by comparing eq^n(2.66)and (2.94). To ensure that the communications channels bandwidth required is same in the two cases, we use the condition, given in eq^n(2.77), that 2N = f_b/f_m . Then eq^n(2.66) can be written as

$$\frac{S_0}{N_0} = \frac{2^{\frac{f_b}{f_m}}}{1 + 2(2^{\frac{f_b}{f_m}}) \text{erfc} \sqrt{S_i / \eta f_b}} \tag{2.95}$$

Eq^n (2.95) and (2.94) are compared in fig.2.18 for N=8(f_b(DM)=48 Kb/s) : to obtain the thermal performance of the delta modulator system, we assume voice transmission where f_m=300 H_z and f_l = 300 H_z .

Thus f_b/f_m =16 (2.96)
And f_m/f_l = 10 (2.97)

Let us compare the ratios S_0/N_0 for PCM and DM for case of voice transmission. We assume that f_m=3000 H_z, f_l = 2Nf_m= 48 x 10^3 H_z. Using these numbers and resulting that the probability of an error in a bit as $P_{eb} = \frac{1}{2} erfc\sqrt{S_i/\eta f_b}$ we have from eq^n (2.94) & (2.95) the result for DM is,

$$(\frac{S_0}{N_0})_{DM} = \frac{2457.6}{1 + 768 erfc\sqrt{S_i / \eta f_b}} = \frac{2457.6}{1 + 1536 P_e} \tag{2.98}$$

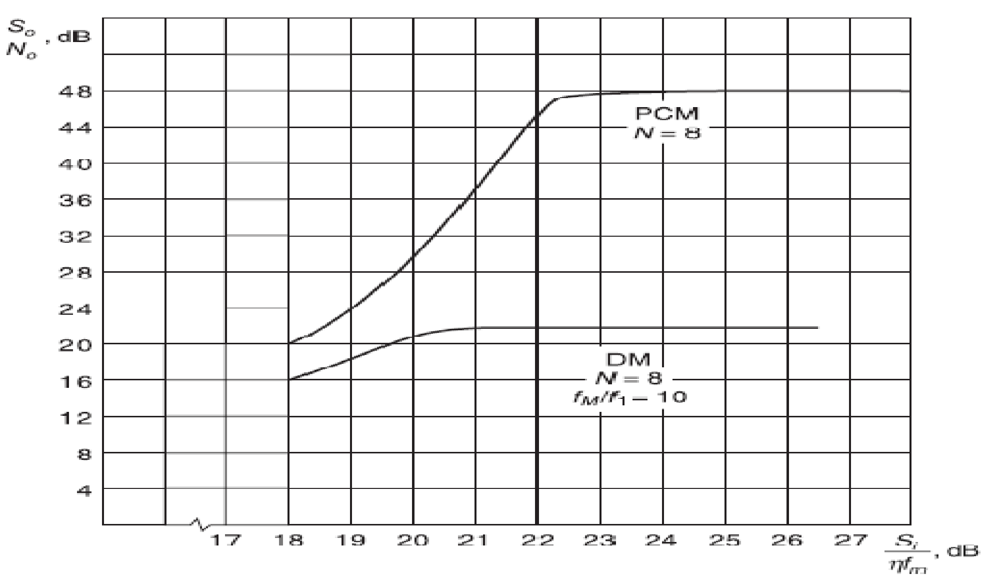


Fig.2.18 A comparison of PCM & DM

And for PCM

$$(\frac{S_0}{N_0})_{PCM} = \frac{65,536}{1+131,072 erfc\sqrt{S_i / \eta f_b}} = \frac{65536}{1+262144 P_e} \qquad (2.99)$$

When the probability of bit error is very small, the PCM system is seen to have higher output SNR than the DM system. Indeed the o/p SNR for PCM system is 48 dB and only about 33 dB for DM system. However, an o/p SNR of 30 dB is all that is required in a communication system. Indeed if commanded PCM is employed the o/p SNR will decrease by about 12 dB to 36 dB for PCM system. Thus eq^n (2.99) indicates that the output SNR is higher for PCM system, the output SNR. In practice, can we consider as being comparable.

With regard to the threshold, we see that when $P_e \sim 10^{-6}$ the PCM system has reached threshold with the DM system reaches threshold when $P_e \sim 10^{-4}$. In practice, we find that our ear does not detect threshold P_e is about 10^{-4} for PCM and 10^{-2} for DM and ADM. Some ADM systems can actually produce understandable speech at error rates as high as 10^{-1-}. Fig.2.18 shows a comparison of PCM and DM for N=8 and $f_m/f_1 = 10$.

Module-III (12 Hours)

Principles of Digital Data Transmission:

A Digital Communication System

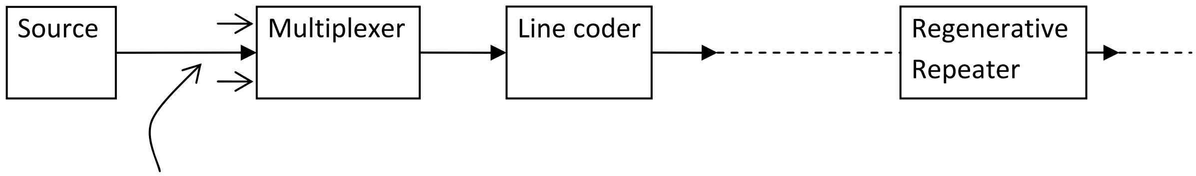


Figure 3.0 A Simple Digital Communication System

1. Digital Data Set
2. Computer output
3. Digital Voice Signal (PCM or DM)
4. Digital facsimile signal
5. Digital TV signal
6. Telemetry equipment signal
7. Etc.

Line Coding

Digital data can be transmitted by various line codes

Desirable properties from a line code

1. Transmission bandwidth – It should be as small as possible
2. Transmitted power – It should be as small as possible
3. Error detection and correction capability – It must be good
4. Favorable PSD – It is desirable to have zero power spectral density (PSD) at $\omega = 0$, because AC coupling and transformers are used at the repeaters. Significant powers in low frequency components cause DC wander in the pulse stream when AC coupling is used.
5. Adequate timing content – It should be possible to extract timing and clock information from the signal
6. Transparency – It should be possible to transmit a digital signal correctly regardless of the pattern of 1's and 0's. If the data are so coded that for every possible sequence of data the coded signal is received faithfully, the code is then transparent.

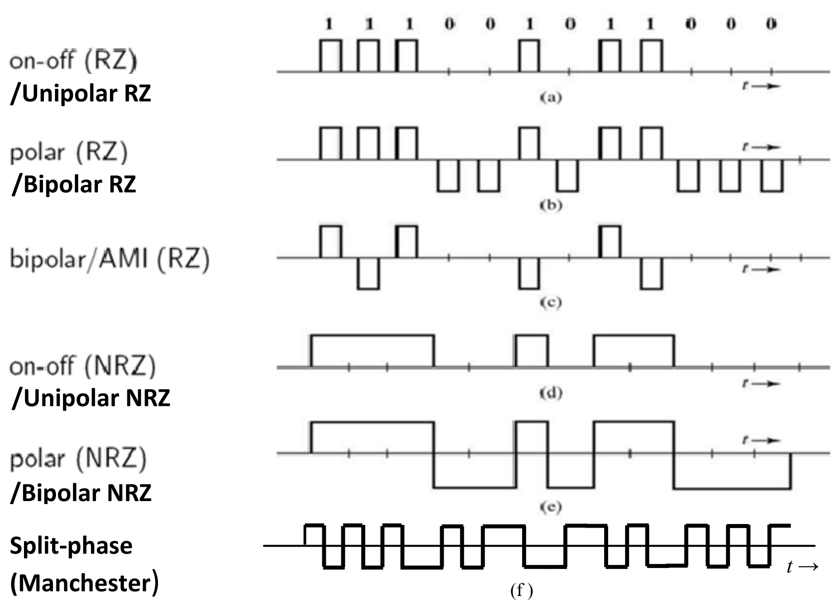


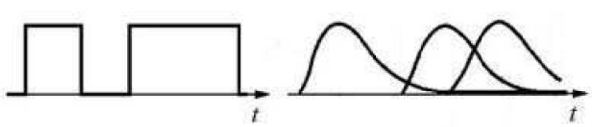
Figure 3.1 Binary signaling formats

Various line codes

Various line codes are as shown in Figure 3.1

Power Spectral Density (PSD) of Line Codes

1. The output distortion of a communication channel depends on the power spectral density of the input signal
2. Input PSD depends on
   i) pulse rate (spectrum widens with pulse rate)
   ii) pulse shape (smoother pulses have narrower PSD)
   iii) pulse distribution
3. Distortion can result in smeared channel output; output pulses are (much) longer than input pulses
4. Inter symbol interference (ISI): received pulse is affected by previous input symbols

Figure 3.2 

Power Spectral Density (review)

For an energy signal g(t) the energy spectral density is the Fourier transform of the autocorrelation:

$$\psi_g(t) = R_g(t) = \int_{-\infty}^{\infty} g(u)g(u+t)\,du \;\Rightarrow\; |G(f)|^2 = \mathcal{F}\{R_g(t)\} \tag{3.1}$$

The autocorrelation of a periodic signal is periodic.

$$R_g(t) = \frac{1}{T}\int_0^T g(u)g(u+t)\,du = \sum_{n=-\infty}^{\infty} |G_n|^2 e^{j2\pi nt} \tag{3.2}$$

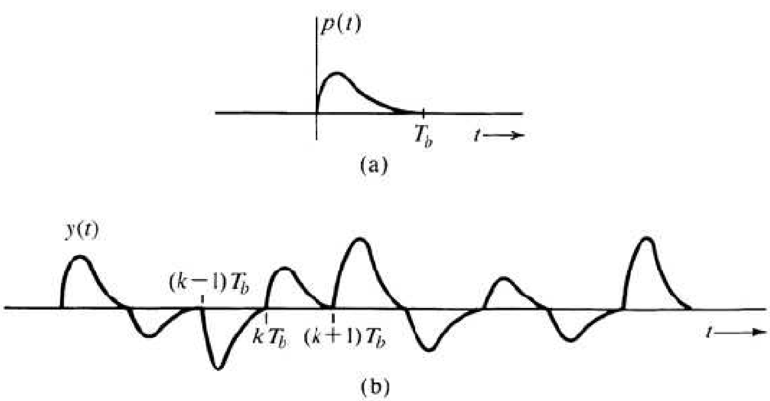
For a power signal, autocorrelation and PSD are average over time. Defines

$$g_T(t) = \Pi(t/T)g(t) = \begin{cases} g(t) & |t| < T/2 \\ 0 & |t| > T/2 \end{cases} \tag{3.3}$$

Then, $$R_g(t) = \lim_{T\to\infty} \frac{R_{g_T}(t)}{T} \;\Rightarrow\; S_g(f) = \lim_{T\to\infty} \frac{|G_T(f)|^2}{T} \tag{3.4}$$

PSD of Line Codes

The PSD of a line code depends on the shapes of the pulses that correspond to digital values. Assume PAM.



$$y(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT_b) \tag{3.5}$$

The transmitted signal is the sum of weighted, shifted pulses. Where, T_b is spacing between pulses. (Pulse may be wider than T_b.) PSD depends on pulse shape, rate, and digital values $\{a_k\}$. We can simplify analysis by representing $\{a_k\}$ as impulse train as shown in figure 3.3(c).

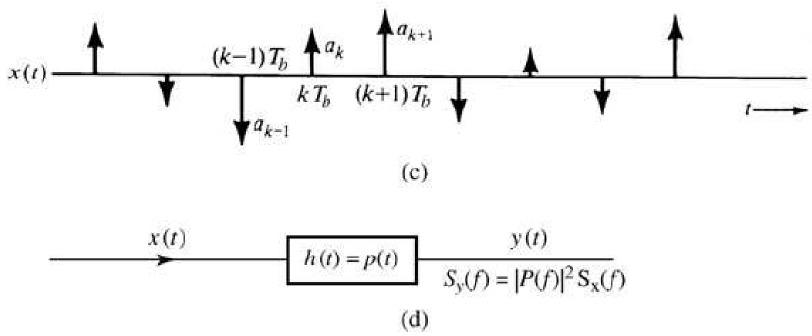


Figure 3.3

PSD of $y(t)$ is $S_y(f) = |P(f)|^2 S_x(f)$.

- $P(f)$ depends only on the pulse, independent of digital values or rate.
- $S_x(f)$ increases linearly with rate $1/T_b$ and depends on distribution of values of $\{ a_k \}$. E.g., $a_k = 1$ for all k has narrower PSD.

<u>PSD of Impulse Train</u>

The autocorrelation of $$x(t) = \sum_{k=-\infty}^{\infty} a_k \delta(t - kT_b) \tag{3.6}$$

can be found as the limit of the autocorrelation of pulse trains:

$$\hat{x}(t) = \sum_{k=-\infty}^{\infty} a_k \frac{\Pi((t - kT_b)/\epsilon)}{\epsilon} \tag{3.7}$$

The autocorrelation of this pulse train (a power signal) is

$$R_{\hat{x}}(t) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \hat{x}(u)\hat{x}(u+t)\, du \tag{3.8}$$

Therefore, $$R_x(t) = \lim_{\epsilon\to 0} R_{\hat{x}}(t). \tag{3.9}$$

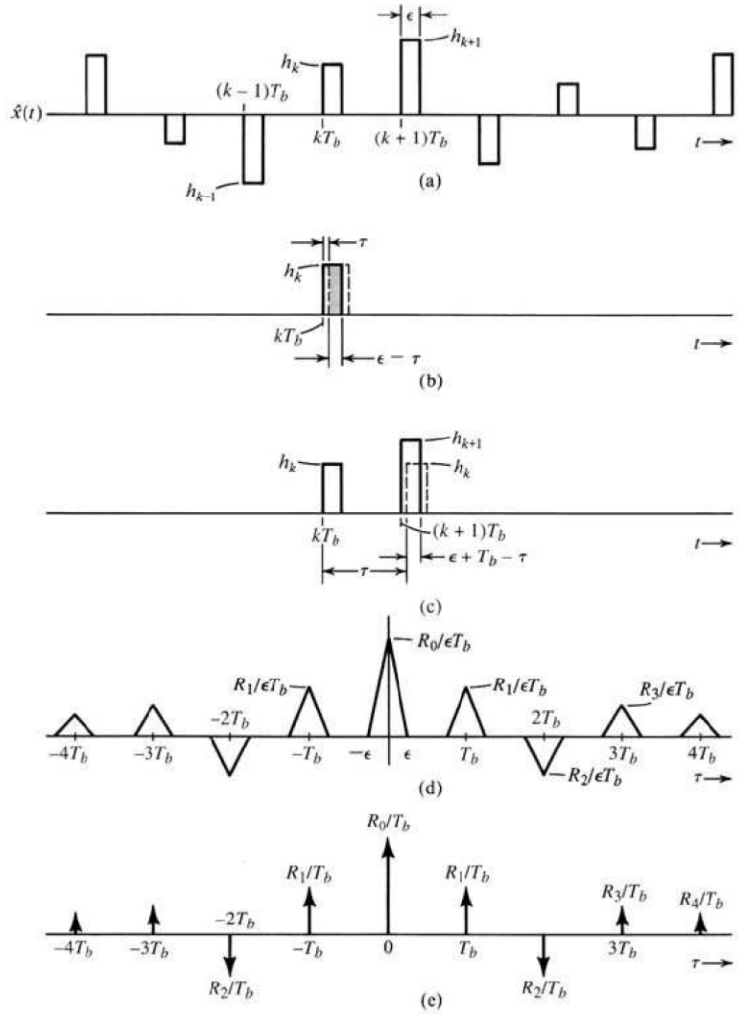


Figure 3.4

$$R_0 = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k^2$$

$$R_1 = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k a_{k+1}$$

and in general

$$R_n = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k a_{k+n} \tag{3.10}$$

The autocorrelation is discrete.
Therefore PSD is periodic in frequency.

The PSD of pulse signal is product

$$S_y(f) = \frac{|P(f)|^2}{T_b} \sum_n R_n e^{j2\pi n f T_b} \tag{3.11}$$

PSD of Polar Signaling

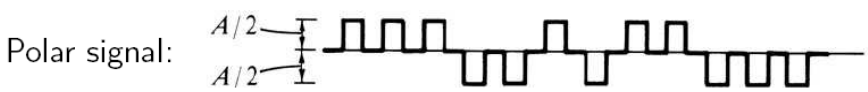


Figure 3.5

- $1 \to +p(t)$, $0 \to -p(t)$
- Since a_k and a_{k+n} ($n \neq 0$) are independent and equally likely,

$$R_0 = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k^2 = \lim_{N\to\infty} \sum_k 1 = 1$$

$$R_n = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k a_{k+n} = 0 \tag{3.12}$$

$$S_y(f) = \frac{|P(f)|^2}{T_b} R_0 = \frac{|P(f)|^2}{T_b} \tag{3.13}$$

- Example: NRZ (100% pulse) $p(t) = \Pi(t/T_b)$

$$P(f) = T_b \operatorname{sinc}(\pi T_b f) \Rightarrow |P(f)|^2 = T_b^2 \operatorname{sinc}^2(\pi T_b f) \tag{3.14}$$

- Half-width: $p(t) = \Pi(t/(T_b/2))$

$$P(f) = \tfrac{1}{2} T_b \operatorname{sinc}(\tfrac{1}{2}\pi T_b f) \Rightarrow |P(f)| = \tfrac{1}{4} T_b^2 \operatorname{sinc}^2(\tfrac{1}{2}\pi T_b f) \tag{3.15}$$

Power spectral density of Polar Signaling (Half-Width Pulse)

For NRZ, $$S_y(f) = \frac{|P(f)|^2}{T_b} = \frac{\frac{1}{4} T_b^2 \operatorname{sinc}^2(\frac{1}{2}\pi T_b f)}{T_b} = \frac{T_b}{4} \operatorname{sinc}^2\left(\frac{\pi T_b f}{2}\right) \tag{3.16}$$

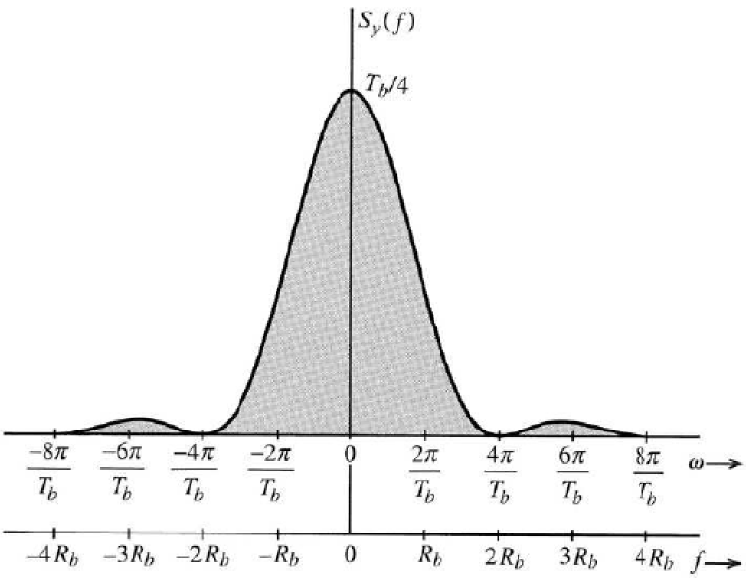


Figure 3.6 PSD of Polar Signaling (Half-Width Pulse)

The bandwidth $2R_b$ is 4× theoretical minimum of 2 bits/Hz/sec.

PSD of On-Off Signaling

- On-off signaling is shifted polar signaling:

$$y_{\text{on-off}}(t) = \tfrac{1}{2}\left(1 + y_{\text{polar}}(t)\right) \tag{3.17}$$

- The DC term results in impulses in the PSD:

$$S_y(f) = \frac{|P(f)|^2}{4T_b}\left(1 + \sum_n \delta(f - n/T_b)\right) \tag{3.18}$$

- We can eliminate impulses by using a pulse $p(t)$ with

$$P\left(\frac{n}{T_b}\right) = 0, \quad n = 0, \pm 1, \pm 2, \ldots \tag{3.19}$$

- Overall, on-off is inferior to polar. For a given average power, noise immunity is less than for bipolar signaling.

Alternate Mark Inversion (Bipolar) Signaling

AMI encodes 0 as 0 V and 1 as +V or −V, with alternating signs.

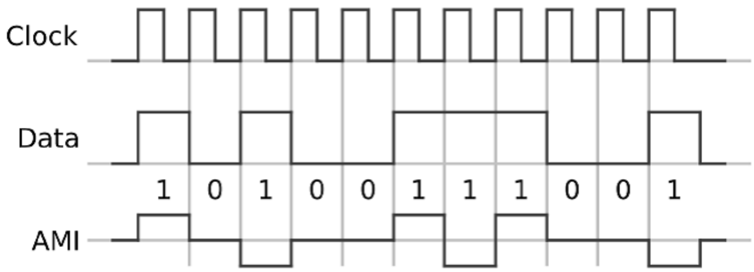


Figure 3.7 AMI signaling

AMI was used in early PCM systems.

- Eliminates DC build up on cable.
- Reduces bandwidth compared to polar.
- Provides error detecting; every bit error results in bipolar violation.
- Guarantees transitions for timing recovery with long runs of ones.

AMI is also called bipolar and pseudo-ternary.

PSD of AMI Signaling

If the data sequence { a_k } is equally likely and independent 0s and 1s, then the autocorrelation function of the sequence is

$$R_0 = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k^2 = \frac{1}{2}$$

$$R_1 = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k a_{k+1} = -\frac{1}{4}$$

$$R_n = \lim_{T\to\infty} \frac{T_b}{T} \sum_k a_k a_{k+n} = 0\,,\ n \geq 2 \tag{3.20}$$

Therefore, $$S_y(f)\frac{|P(f)|^2}{2T_b}(1-\cos 2\pi T_b f) = \frac{|P(f)|^2}{T_b}\sin^2(\pi T_b f) \tag{3.21}$$

This PSD falls off faster than sinc ($\pi T_b f$). Further, the PSD has a null at DC, which aids in transformer coupling.

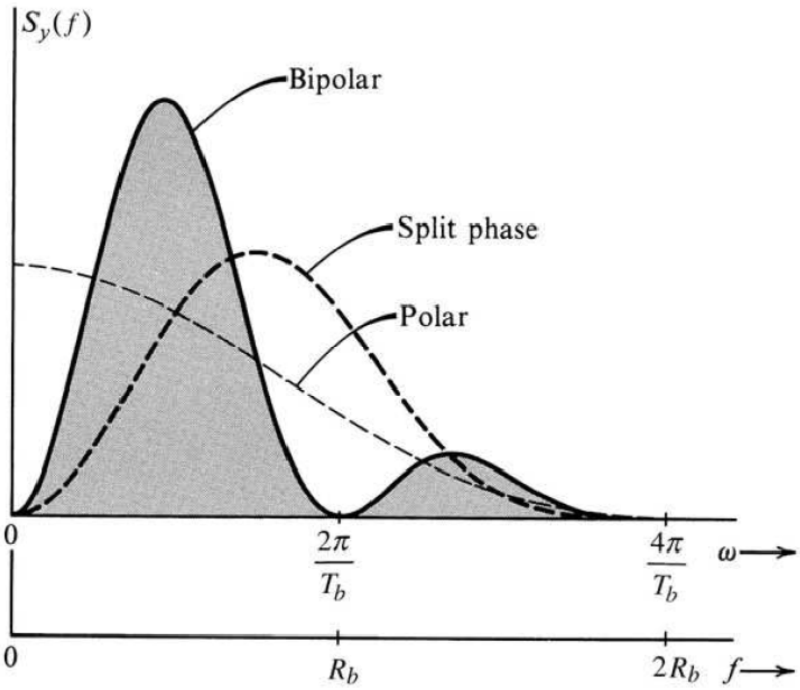


Figure 3.8 PSD of bipolar, polar, and split phase signals normalized for equal power. (Half width rectangular pulses are used)

<u>Nyquist First Criterion</u>

<u>Reducing ISI: Pulse Shaping</u>

- A time-limited pulse cannot be band-limited
- Linear channel distortion results in spread out, overlapping pulses
- Nyquist introduced three criteria for dealing with ISI.

The first criterion was that each pulse is zero at the sampling time of other pulses.

$$p(t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm kT_b\,,\ k = \pm 1, \pm 2, \ldots \end{cases} \tag{3.22}$$

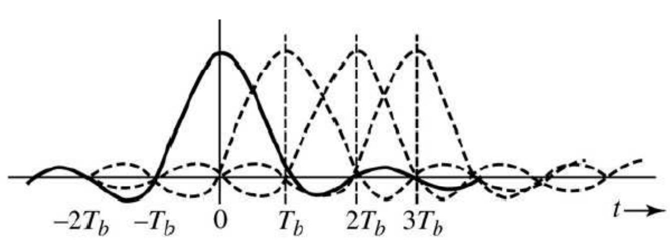


Figure 3.9

<u>Pulse Shaping: Sinc Pulse</u>

- Let $R_b = 1/T_b$. The sinc pulse, $\text{sinc}(\pi R_b t)$ satisfies Nyquist's first crierion for zero ISI:

$$\text{sinc}(\pi R_b t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm kT_b\,,\ k = \pm 1, \pm 2, \ldots \end{cases} \tag{3.23}$$

- This pulse is band-limited. Its Fourier transform is $P(f) = \frac{1}{R_b}\Pi\left(\frac{f}{R_b}\right)$ (3.24)

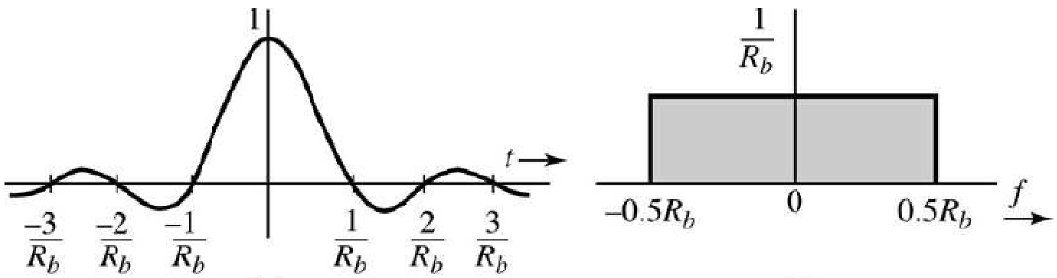


Figure 3.10 Sic pulse (minimum bandwidth pulse) and its Fourier transform.

- Unfortunately, this pulse has infinite width in time and decays slowly.

<u>Nyquist Pulse</u>
Nyquist increased the width of the spectrum in order to make the pulse fall off more rapidly. The Nyquist pulse has spectrum width (1/2) $(1 + r)R_b$, where $0 < r < 1$.

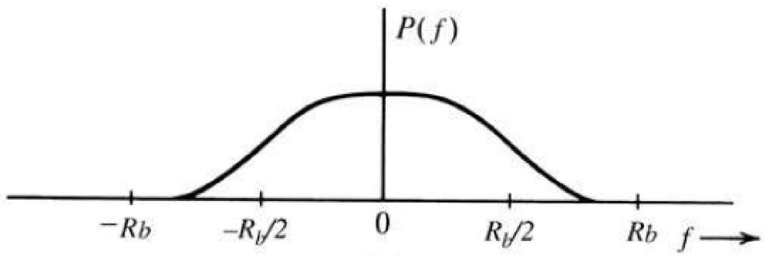


Figure 3.11 Proposed Nyquist pulse

If we sample the pulse $p(t)$ at rate $R_b = 1/T_b$, then $\overline{p}(t) = p(t)\,\text{III}_{T_b}(t) = p(t)\delta(t) = \delta(t)$ (3.25)

The Fourier transform of the sampled signal is $\overline{P}(f) = 1 = \sum_{k=-\infty}^{\infty} P(f - kR_b)$ (3.26)

Since we are sampling below the Nyquist rate $2R_b$, the shifted transforms overlap. Nyquist's criterion requires pulses whose overlaps add to 1 for all f.

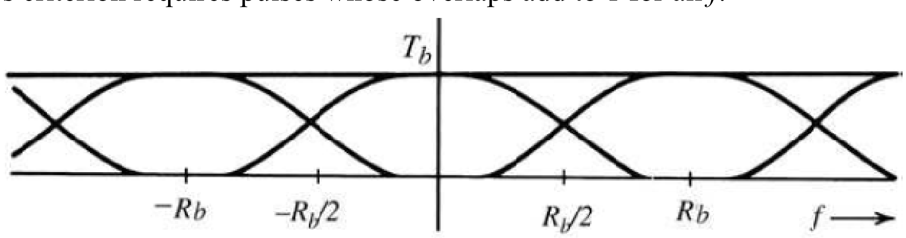


Figure 3.12 Sampled Nyquist pulse

For parameter r with $0 < r < 1$, the resulting pulse has bandwidth $B_r = \frac{1}{2}(R_b + rR_b)$ (3.27)

The parameter r is called *roll-off factor* and controls how sharply the pulse spectrum declines above $(1/2)R_b$.

There are many pulse spectra satisfying this condition. e.g., trapezoid:

$$P(f) = \begin{cases} 1 & |f| < \frac{1}{2}(1-r)R_b \\ 1 - \frac{|f|-(1-r)R_b}{2R_b} & \frac{1}{2}(1-r)R_b < |f| < \frac{1}{2}(1+r)R_b \\ 0 & |f| > \frac{1}{2}(1-r)R_b \end{cases} \tag{3.28}$$

A trapezoid is the difference of two triangles. Thus the pulse with trapezoidal Fourier transform is the difference of two sinc^2 pulses.
Example: for $r = 1/2$,

$$P(f) = \tfrac{3}{2}\Lambda\left(\frac{f}{\frac{3}{2}R_b}\right) - \tfrac{1}{2}\Lambda\left(\frac{f}{\frac{1}{2}R_b}\right) \tag{3.29}$$

So the pulse is, $p(t) = \frac{9}{4}\,\text{sinc}^2(\frac{3}{2}R_b t) - \frac{1}{4}\,\text{sinc}^2(\frac{1}{2}R_b t)$ (3.30)

This pulse falls off as $1/t^2$

Nyquist chose a pulse with a "vestigial" raised cosine transform. This transform is smoother than a trapezoid, so the pulse decays more rapidly.
The Nyquist pulse is parameterized by r. Let $f_x = rR_b/2$.

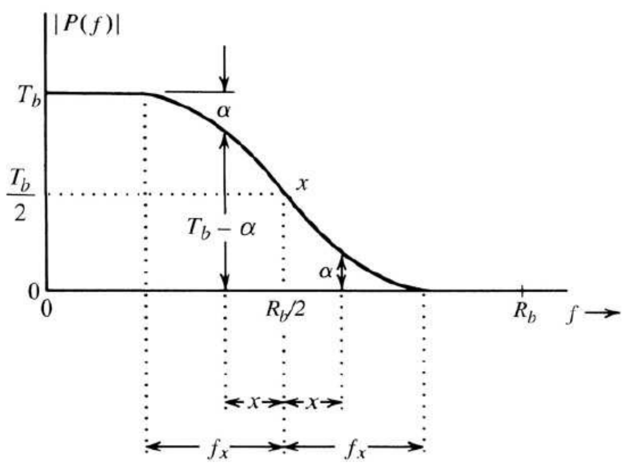


Figure 3.13 Vestigial spectrum

Nyquist pulse spectrum is raised cosine pulse with flat porch.

$$P(f) = \begin{cases} 1 & |f| < \frac{1}{2}R_b - f_x \\ \frac{1}{2}\left(1 - \sin\pi\left(\frac{f - \frac{1}{2}R_b}{2f_x}\right)\right) & |f| - \frac{1}{2}R_b| < f_x \\ 0 & |f| > \frac{1}{2}R_b + f_x \end{cases} \tag{3.31}$$

The transform $P(f)$ is differentiable, so the pulse decays as $1/t^2$.

Special case of Nyquist pulse is $r = 1$: *full-cosine roll-off.*

$$\begin{aligned} P(f) &= \tfrac{1}{2}(1 + \cos\pi T_b f)\Pi(f/R_b) \\ &= \cos^2(\tfrac{1}{2}\pi T_b f)\,\Pi(\tfrac{1}{2}T_b f) \end{aligned} \tag{3.32}$$

This transform *P*(*f*) has a second derivative so the pulse decays as $1/t^3$.

$$p(t) = R_b \frac{\cos \pi R_b t}{1 - 4R_b^2 t^2} \operatorname{sinc}(\pi R_b t) = \frac{\sin(2\pi R_b t)}{2\pi t(1 - 4R_b^2 t^2)} \tag{3.33}$$

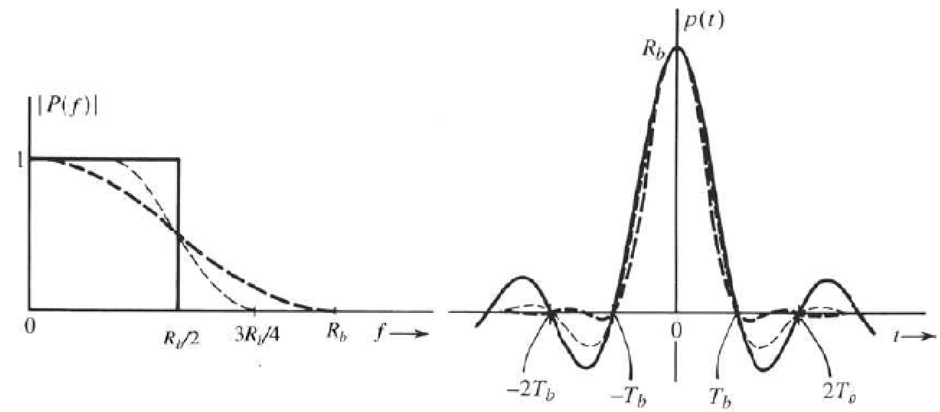


Figure 3.14 Pulses satisfying the Nyquist criterion

Controlled ISI (Partial Response Signaling)

We can reduce bandwidth by using an even wider pulse. This introduces ISI, which can be canceled using knowledge of the pulse shape.

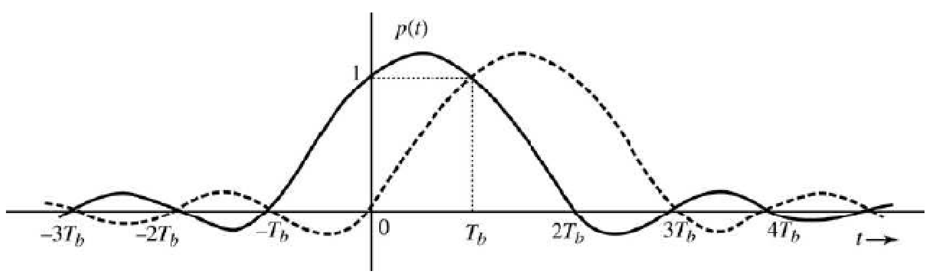


Figure 3.15 Duo-binary pulse

The value of $y(t)$ at time nT_b is $a_{n-2} + a_{n-1}$. Decision rule:

$$\hat{a}_{n-1} = \begin{cases} 1 & y(nT_b) > 0 \\ 0 & y(nT_b) < 0 \\ (\hat{a}_{n-2})' & y(nT_b) = 0 \end{cases} \tag{3.34}$$

A related approach is decision feedback equalization: once a bit has been detected, its contribution to the received signal is subtracted. The ideal duo-binary pulse is

$$p(t) = \frac{\sin \pi R_b t}{\pi R_b t(1 - R_b t)} \tag{3.35}$$

The Fourier transform of $p(t)$ is

$$P(f) = \frac{2}{R_b} \cos\left(\frac{\pi f}{R_b}\right) \Pi\left(\frac{f}{R_b}\right) e^{-j\pi f/R_b} \tag{3.36}$$

The spectrum is confined to the theoretical minimum of $R_b/2$.

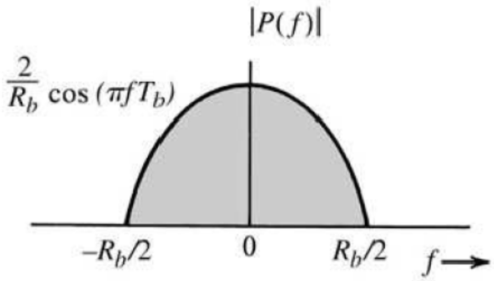


Figure 3.16 Minimum bandwidth pulse that satisfies the duo-binary pulse spectrum

Zero-ISI, Duobinary, Modified Duobinary Pulses

Suppose $p_a(t)$ satisfies Nyquist's first criterion (zero ISI). Then

$$p_b(\mathrm{t}) = p_a(\mathrm{t}) + p_a(t - T_b) \tag{3.37}$$

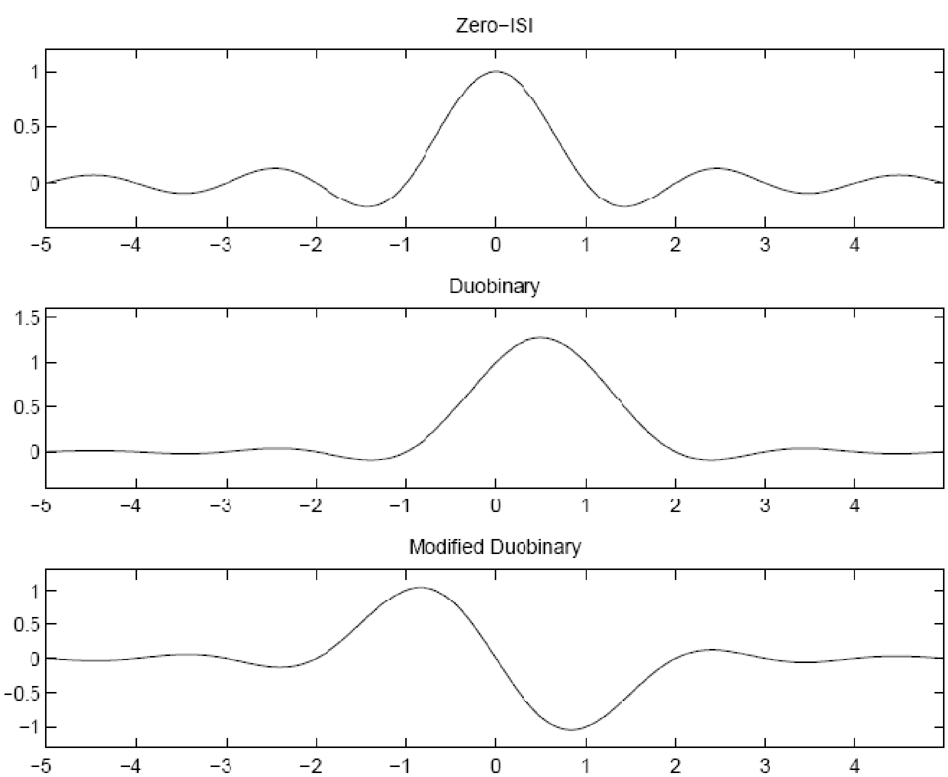
is a duo-binary pulse with controlled ISI. By shift theorem,

$$P_b(f) = P_a(1 + \mathrm{e}^{-j2\pi T_b f}) \tag{3.38}$$

Since $P_b(R_b/2) = 0$, most (or all) of the pulse energy is below $R_b/2$. We can eliminate unwanted DC component using modified duo-binary, where $p_c(-T_b) = 1$, $p_c(T_b) = -1$, and $p_c(\mathrm{n}T_B) = 0$ for other integers n.

$$p_c(t) = p_a(t + T_b) - p_a(t - T_b) \Rightarrow P_c(f) = 2jP_a(f)\sin 2\pi T_b f \tag{3.39}$$

The transform of $p_c(t)$ has nulls at 0 and $\pm R_b/2$.



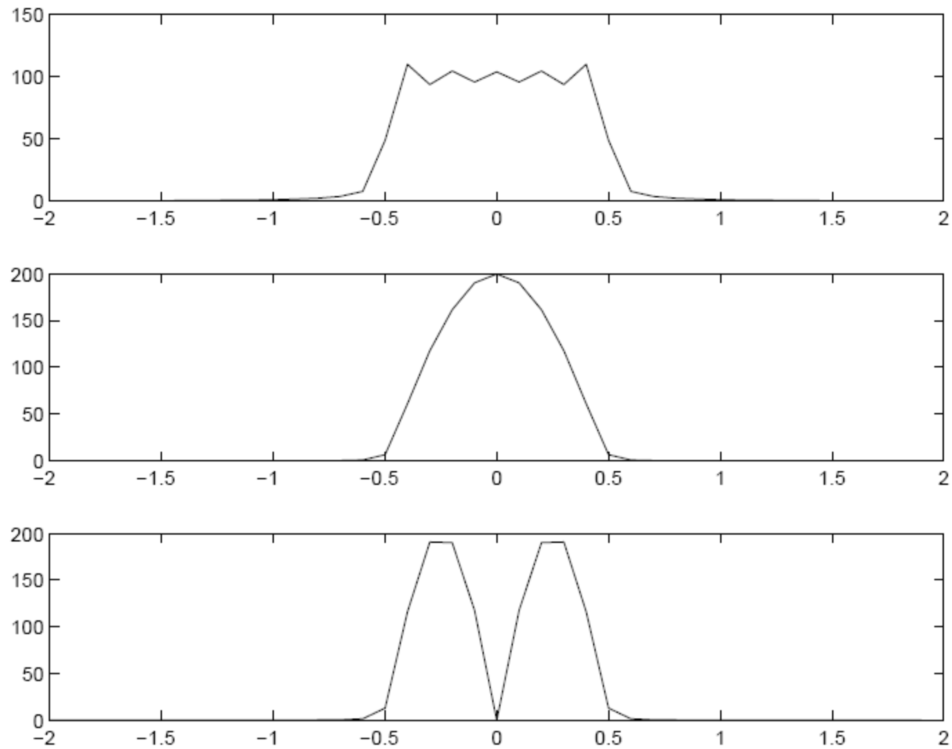


Figure 3.17 Zero-ISI, Duobinary, Modified Duobinary and other Pulses

<u>Partial Response Signaling Detection</u>

Suppose that sequence 0010110 is transmitted (first bit is startup digit).

Digit x_k	0	0	1	0	1	1	0
Bipolar amplitude -	-1	-1	1	-1	1	1	-1
Combined amplitude		-2	0	0	0	2	0
Decoded values		-2	0	2	0	0	2
Decode sequence		0	1	0	1	1	0

Partial response signaling is susceptible to error propagation. If a nonzero value is mis-detected, zeros will be mis-detected until the next nonzero value.
Error propagation is eliminated by pre-coding the data: $p_k = x_k \oplus p_{k-1}$.

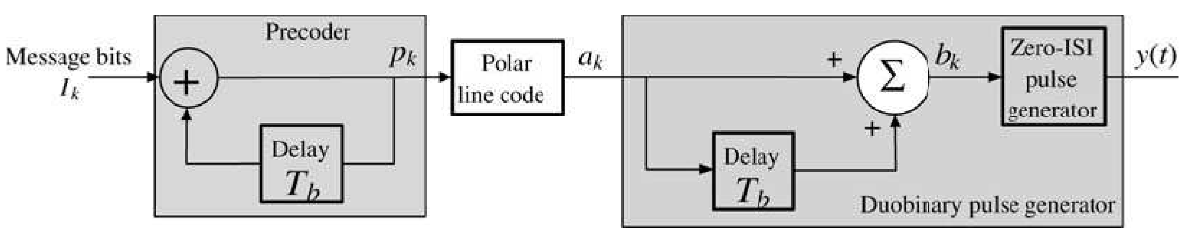


Figure 3.18 Duo-binary pulse generator

<u>Scrambling</u>

In general, a scrambler tends to make the data more random by removing long strings of 1's or 0's. Scrambling can be helpful in timing extraction by removing long strings of 0's in data. Scramblers, however, are primarily used for preventing unauthorized access to the data, and are optimized for that purpose. Such optimization may actually result in the generation of a long string of zeros in the data. The digital network must be able to cope with these long zero strings using zero suppression techniques as discussed in case of high density bipolar (HDB) signaling and binary with 8 zeros substitution (B8ZS) signaling.

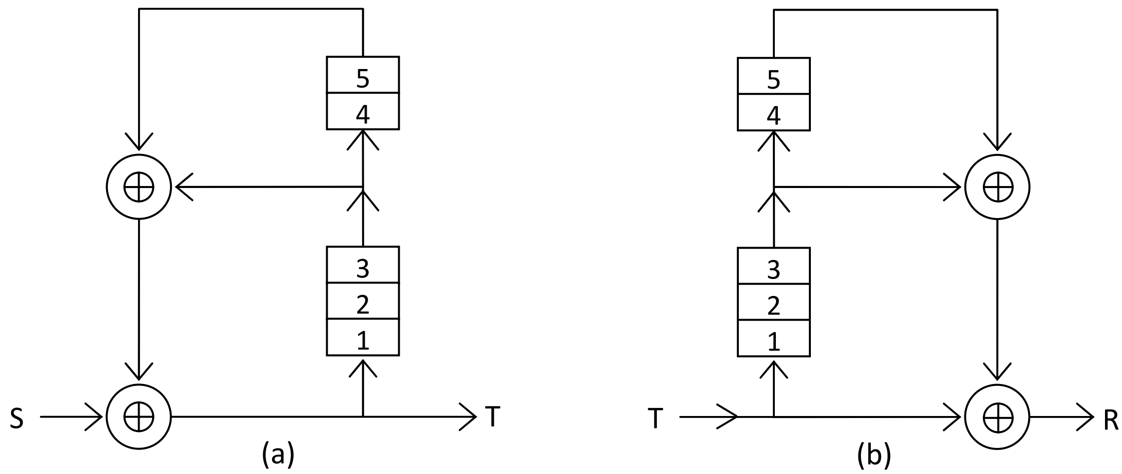


Figure 3.19 Scrambler and Descrambler

Above figure 3.19 shows a typical scrambler and descrambler. The scrambler consists of a feedback shift register, and the matching descrambler has a feed-forward shift register as indicated. Each stage in the shift register delays a bit by one unit. To analyze the scrambler and the matched descrambler, consider the output sequence T of the scrambler [figure 3.19 (a)]. If S is the input sequence to the scrambler, then

$$S \oplus D^3T \oplus D^5T = T \tag{3.40}$$

Where, D represents the delay operator; i.e., D^nT is the sequence T delayed by 'n' units. The symbol ⊕ indicates modulo 2 sum. Now recall that the modulo 2 sum of any sequence with itself gives a sequence of all 0's. Modulo 2 addition of $(D^3 \oplus D^5)T$ to both sides of the above equation, we get

$$\begin{aligned} S &= T \oplus (D^3 \oplus D^5)T \\ &= [1 \oplus (D^3 \oplus D^5)]T \\ &= (1 \oplus F)T \quad ; \text{where, } F = D^3 \oplus D^5 \end{aligned} \tag{3.41}$$

To design the descrambler at the receiver side, we start with T, the sequence received at the descrambler. Now we can see that received signal after descrambling i.e. R is same as S.

$$R = T \oplus (D^3 \oplus D^5)T = T \oplus FT = (1 \oplus F)T = S \tag{3.42}$$

Regenerative Repeater

Basically, a regenerative repeater performs three functions.

1. Reshaping incoming pulse by means of equalizer
2. The extraction of timing information required to sample incoming pulses at optimum instants.
3. Decision making based on the pulse samples.

The schematic of a repeater is shown in the following figure. A complete repeater also includes provision for the separation of DC power from AC signals. This is normally accomplished using transformer by coupling the signals and bypassing the DC around transformers to the power supply circuitry.

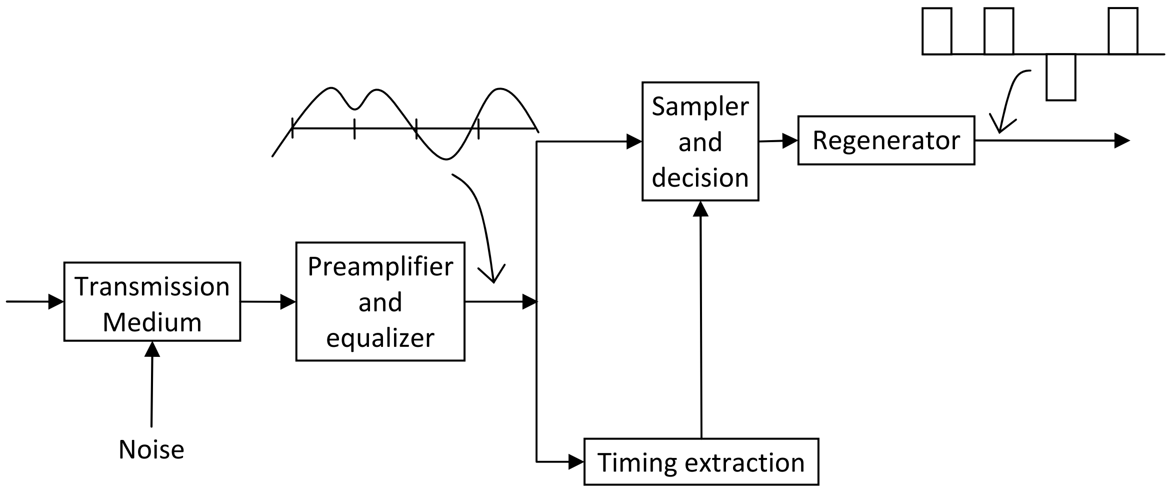


Figure 3.20 Regenerative Repeater

Preamplifier

Preamplifier, as the name suggests, is an electronic device to amplify very weak signal. The output from it becomes the input for another amplifier.

A signal is modulated by superimposing a known frequency on it and the amplifier is set to detect only those signals on which the selected frequency is superimposed. Such an amplifier is known as lock-in-amplifier. Noise not modulated by the selected frequency will not be amplified. Therefore it will be filtered off.

Equalization

As discussed in the *Pulse Shaping*, a properly shaped transmit pulse resembles a sinc function, and direct superposition of these pulses results in no ISI at properly selected sample points.

In practice, however, the received pulse response is distorted in the transmission process and may be combined with additive noise. Because the raised cosine pulses are distorted in the time domain, you may find that the received signal exhibits ISI. If you can define the channel impulse response, you can implement an inverse filter to counter its ill effect. This is the job of the equalizer. See figure 9 below, which depicts the response to a single transmit pulse at various points in the system.

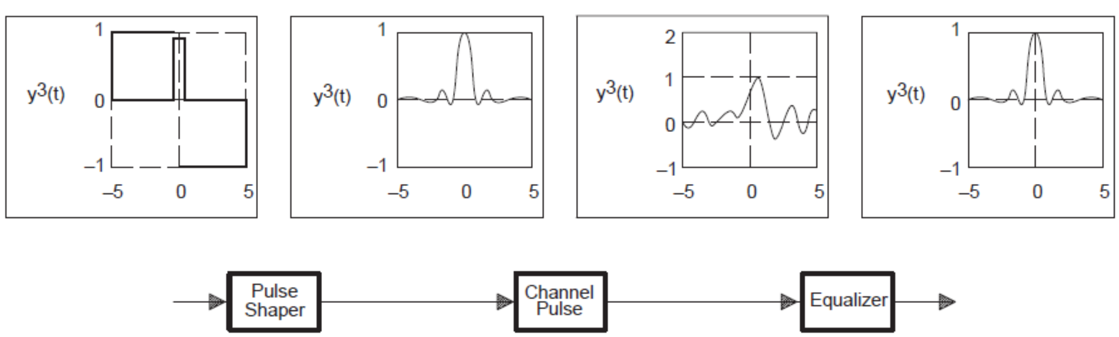


Figure 3.21 Transmission process with pulse responses example

The original rectangular pulse is shaped by the raised cosine filter before transmission. This ensures that the sampled spectra do not alias and therefore there is no ISI. This third waveform portrays the distorted impulse response received at the input of the equalizer. This distortion can be caused by spectral shaping due to a non-flat frequency response or multipath reception of the channel. This distortion can be removed by applying a filter that is the exact inverse (multiplicative inverse in spectral domain) of the channel frequency response.

Equalizers

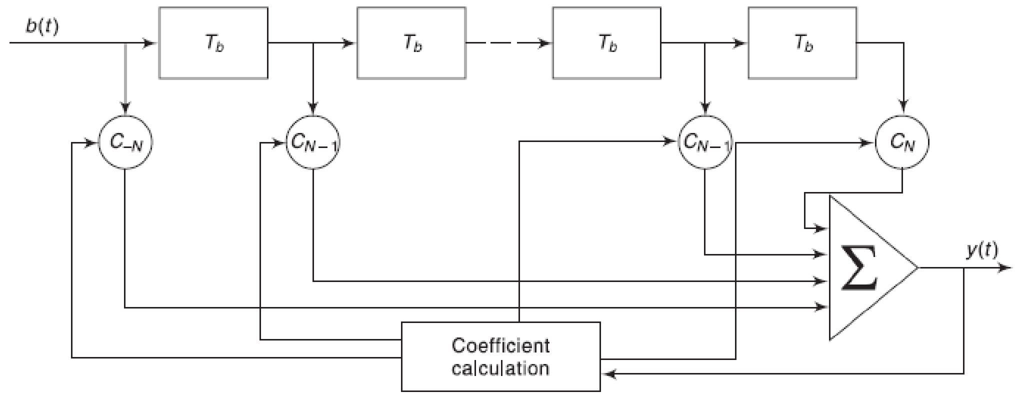


Figure 3.22 Block diagram of a tap delay equalizer

Zero Forcing Equalizer

$$y(t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm T_b, \pm 2T_b, \ldots, \pm NT_b \end{cases} \tag{3.43}$$

$$\begin{bmatrix} b(0) & b(-T_b) & \cdots & b(-2NT_b) \\ b(T_b) & b(0) & \cdots & b[(-2N+1)T_b] \\ \cdots & \cdots & & \cdots \\ b(NT_b) & b[(N-1)T_b] & \cdots & b(-NT_b) \\ \cdots & \cdots & & \cdots \\ b(2NT_b) & b[(2N-1)T_b] & \cdots & b(0) \end{bmatrix} \begin{bmatrix} C_{-N} \\ C_{-N+1} \\ \cdots \\ C_0 \\ \cdots \\ C_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{bmatrix} \tag{3.44}$$

In the above matrix represents $2N + 1$ independent equations as many number of tap weights C_i which are uniquely determined by solving the matrix.

Mean square and Adaptive Equalizer

Mean square equalizer Instead of forcing zero crossing this method tries to minimize mean square error by a set of output samples solving simultaneous equations.

Adaptive equalizer This is useful when channel characteristics is changing. This involves sending pre-assigned pulses at periodic intervals prior to data transmission which adjusts tap weights by an iterative procedure that minimizes ISI.

Eye Diagrams

Polar Signaling with Raised Cosine Transform (r = 0.5)

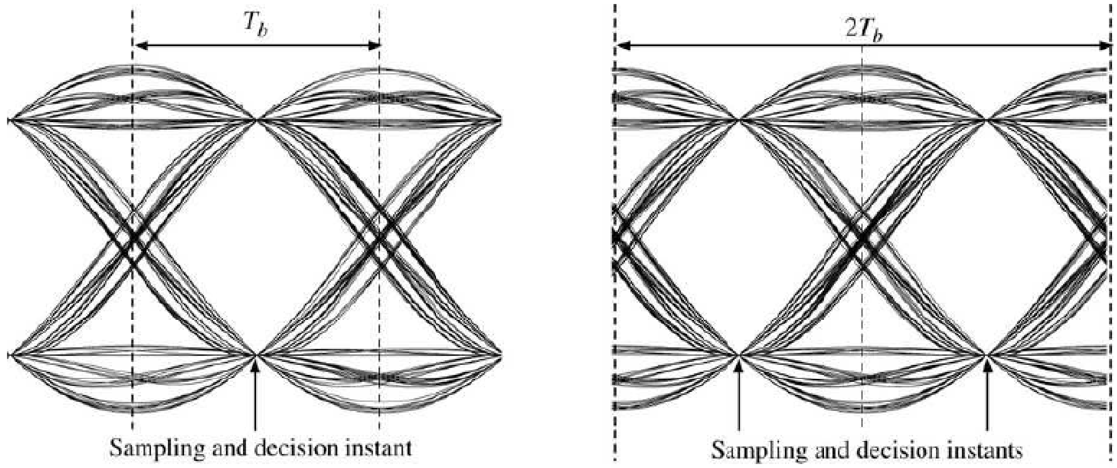


Figure 3.23 Eye diagram of Polar Signaling with Raised Cosine Transform (single window)

$$P(f) = \begin{cases} 1 & |f| < \frac{1}{4}R_b \\ \frac{1}{2}\left(1 - \sin\pi\left(\frac{f - \frac{1}{2}R_b}{R_b}\right)\right) & \left||f| - \frac{1}{2}R_b\right| < \frac{1}{2}R_b \\ 0 & |f| > \frac{3}{4}R_b \end{cases} \quad (3.45)$$

Polar Signaling with Raised Cosine Transform (r = 0.5). The pulse corresponding to $P(f)$ is

$$p(t) = \text{sinc}(\pi R_b t)\,\frac{\cos(\pi r R_b t)}{1 - 4r^2 R_b^2 t^2} \quad (3.46)$$

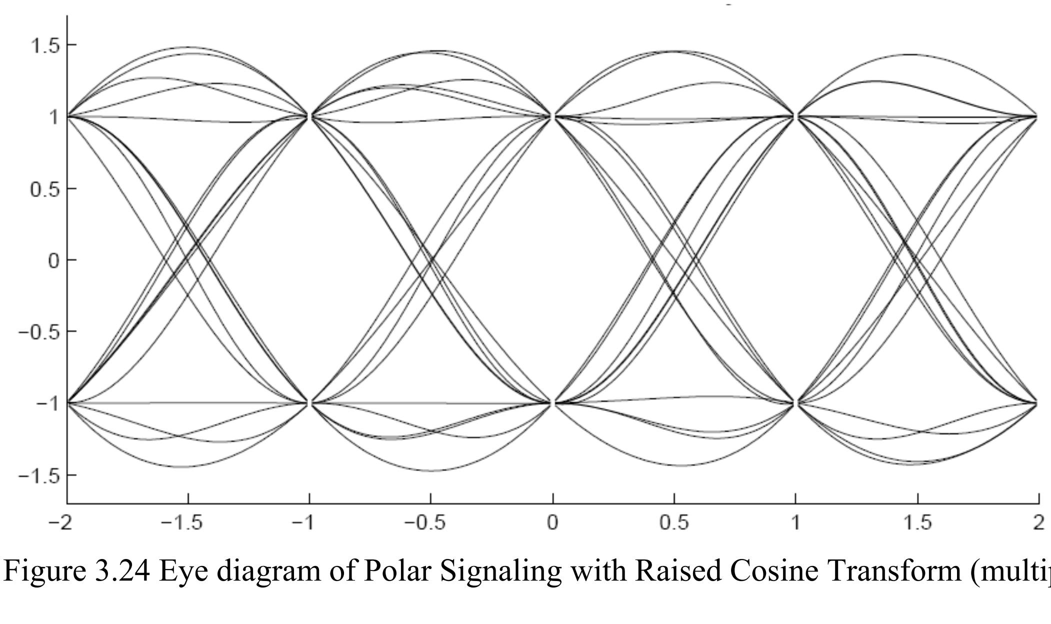


Figure 3.24 Eye diagram of Polar Signaling with Raised Cosine Transform (multiple window)

Eye Diagram Measurements

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

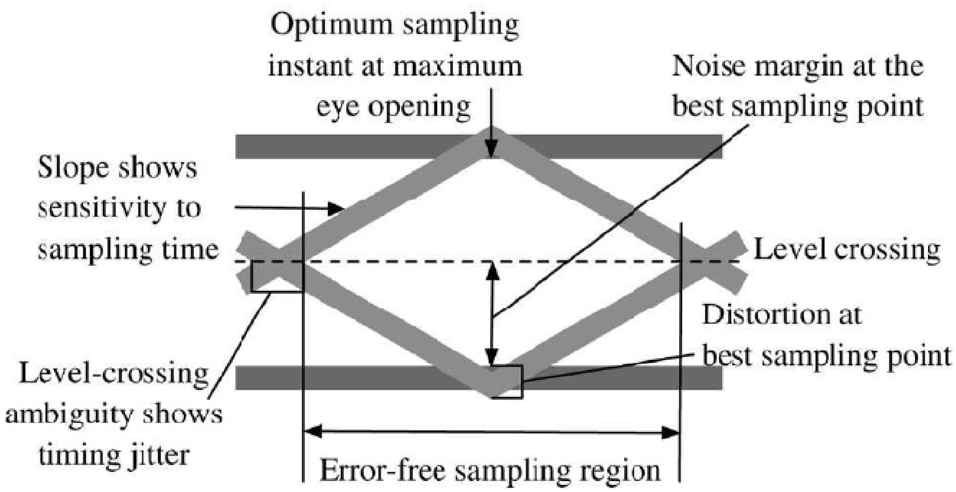


Figure 3.25 Measurement using Eye diagram

Timing Extraction

The received digital signal needs to be sampled at précised instants. This requires a clock signal at the receiver in synchronism with the clock signal at the transmitter (Symbol or bit synchronization). Three general methods of synchronization exist.

1. Derivation from a primary or a secondary standard (e.g. transmitter and receiver slaved to a master timing source)
2. Transmitting a separate synchronizing signal (Pilot clock)
3. Self synchronization, where the timing information is extracted from the received signal itself.

The first method is suitable for large volume of data and high speed communication systems because of its high cost. In the second method, part of the channel capacity is used to transmit timing information and is suitable when the available capacity is large compared to the data rate. The third method is a very efficient method of timing extraction or clock recovery because the timing is derived from the digital signal itself.

Timing Jitter

Variations of pulse positions or sampling instants cause timing jitter. This results from several causes, some of which are dependent on the pulse pattern being transmitted where as others are not. The former are cumulative along the chain of regenerative repeaters because all the repeaters are affected in the same way, where as the forms of jitter are random from regenerator to regenerator and therefore tend to partially cancel out their mutual effects over a long-haul link. Random forms of jitter are caused by noise, interference, and mistuning of clock circuits. The pattern-depend jitter results from clock mistuning, amplitude-to-phase conversion in the clock circuit, and ISI, which alters the position of the peaks of the input signal according to the pattern. The r.m.s. value of the jitter over a long chain of 'N' repeaters can be shown to increase as $\sqrt{N}$.

Jitter accumulation over a digital link may be reduced by buffering the link with an elastic store and clocking out the digital stream under the control of highly stable PLL. Jitter reduction is necessary about every 200 miles in a long digital link to keep the maximum jitter with reasonable limits.

A Baseband Signal Receiver

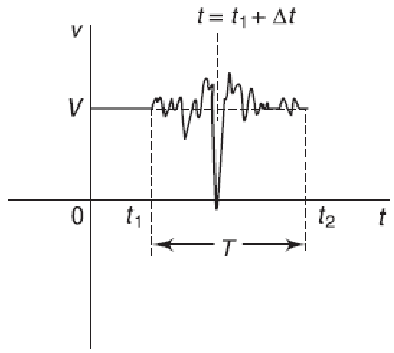


Figure 3.26 Transmitted pulse with noise

The above figure explains that noise may cause an error in the determination of a transmitted voltage level.

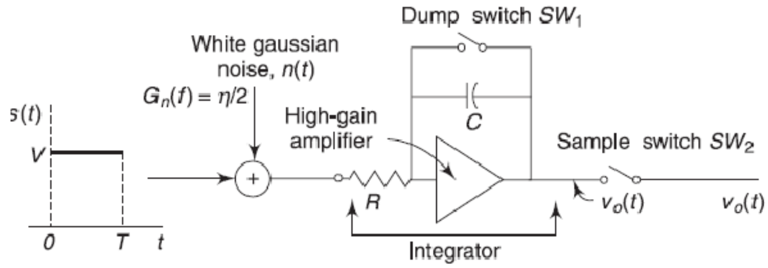


Figure 3.27 A receiver for a binary coded signal.

Peak SNR

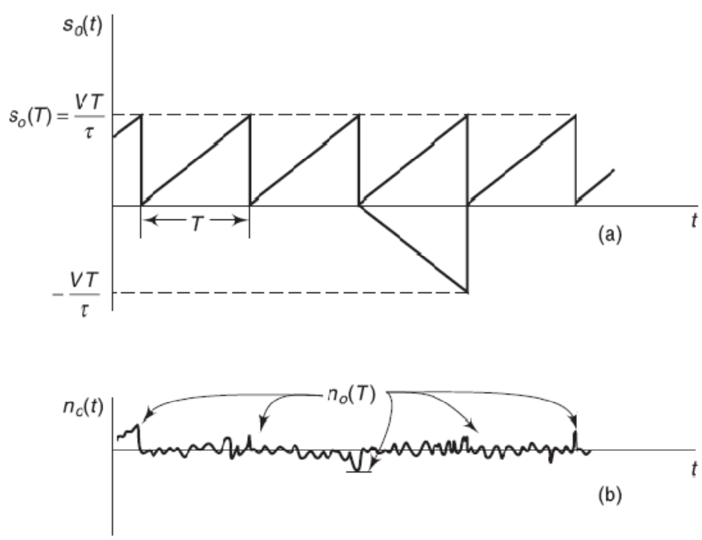


Figure 3.28 (a)The signal output (b) the noise output of the integrator as shown in figure 3.27

Using $\tau = RC$, $v_o(T) = \frac{1}{\tau}\int_0^T [s(t) + n(t)]\, dt = \frac{1}{\tau}\int_0^T s(t)\, dt + \frac{1}{\tau}\int_0^T n(t)\, dt$ (3.47)

The sample voltage due to the signal is $s_o(T) = \frac{1}{\tau}\int_0^T V\, dt = \frac{VT}{\tau}$ (3.48)

The sample voltage due to the noise is $n_o(T) = \frac{1}{\tau}\int_0^T n(t)\, dt$ (3.49)

The variance of noise is $n_o(T)$ is known to us and is $\sigma_o^2 = \overline{n_o^2(T)} = \dfrac{\eta T}{2\tau^2}$ (3.50)

$$v_o(T) = s_o(T) + n_o(T) \tag{3.51}$$

Figure of merit is $\dfrac{[s_o(T)]^2}{[n_o(T)]^2} = \dfrac{2}{\eta} V^2 T$ (3.52)

<u>Probability of Error</u>

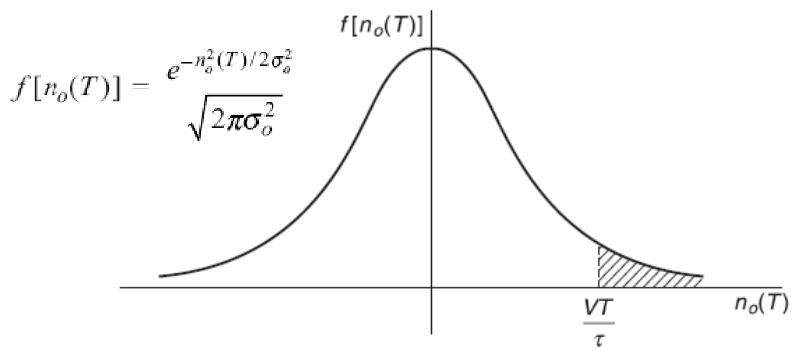


Figure 3.29 The Gaussian probability density of the noise sample $n_o(T)$

$$P_e = \int_{VT/\tau}^{\infty} f[n_o(T)]\, dn_o(T) = \int_{VT/\tau}^{\infty} \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}}\, dn_o(T) \tag{3.53}$$

Defining $x \equiv n_o(T)/\sqrt{2}\sigma_o$,

$$\begin{aligned} P_e &= \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{x=V\sqrt{T/\eta}}^{\infty} e^{-x^2}\, dx \\ &= \frac{1}{2}\operatorname{erfc}\left(V\sqrt{\frac{T}{\eta}}\right) \\ &= \frac{1}{2}\operatorname{erfc}\left(\frac{V^2 T}{\eta}\right)^{1/2} \\ &= \frac{1}{2}\operatorname{erfc}\left(\frac{E_s}{\eta}\right)^{1/2} \end{aligned} \tag{3.54}$$

Note that P_e decreases rapidly as E_s/η increases. The maximum value of P_e is $\frac{1}{2}$.

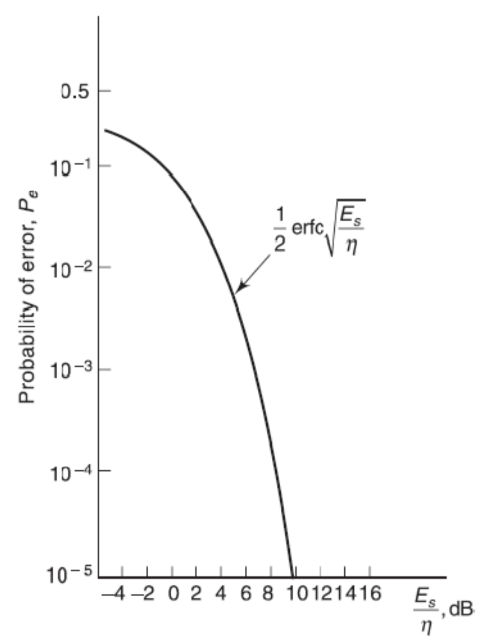


Figure 3.30 Variation of P_e versus E_s/η

Optimum Threshold

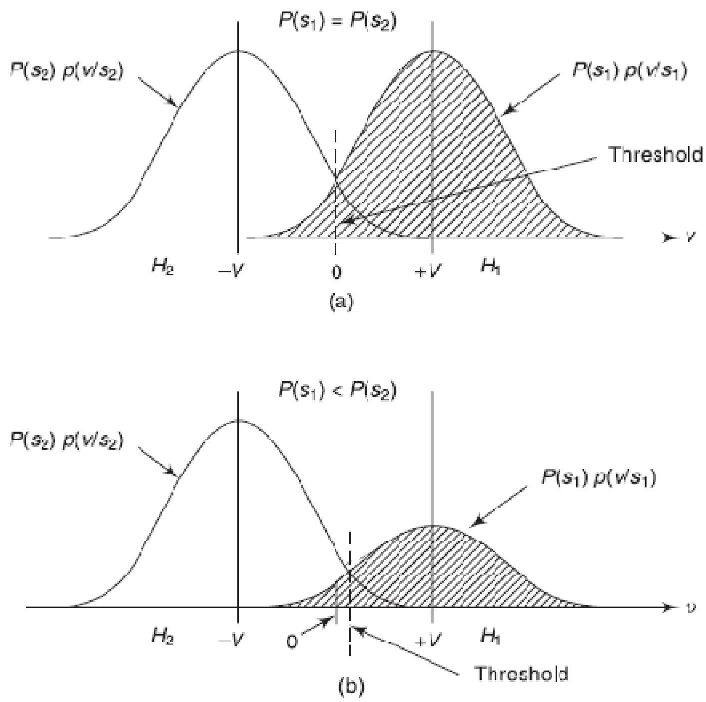


Figure 3.31 Decision threshold when apriori probability are (a) equal (b) unequal

Consider, when symbol sent is s_1, the probability of receiving voltage v is $P(v/s_1)$ and for symbol sent s_2 it is $P(v/s_2)$. We define apriori probability of presence of these symbols as $P(s_1)$ and $P(s_2)$ respectively. The decision threshold λ is such that for $v > \lambda$, symbol s_1 is selected and for $v < \lambda$, symbol s_2 is selected. Then probability of error

$$P_e = P(s_1)\int_{v<\lambda} p(v/s_1)dv + P(s_2)\int_{v>\lambda} p(v/s_2)dv \tag{3.55}$$

$$\int_{v>\lambda} p(v/s_1)dv + \int_{v<\lambda} p(v/s_1)dv = 1 \tag{3.56}$$

$$P_e = P(s_1)\left[1 - \int_{v>\lambda} p(v/s_1)\,dv\right] + P(s_2)\int_{v>\lambda} p(v/s_2)dv$$

$$= P(s_1) + \int_{v>\lambda} [P(s_2)p(v/s_2) - P(s_1)p(v/s_1)dv]dv \tag{3.57}$$

probability of error is minimum if for every $v > \lambda$,

$$P(s_1)p(v/s_1) > P(s_2)p(v/s_2) \quad \text{Or,} \quad \frac{p(v/s_1)}{p(v/s_2)} > \frac{P(s_2)}{P(s_1)} \tag{3.58}$$

maximum likelihood detector $$\frac{p(v/s_1)}{p(v/s_2)} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(s_2)}{P(s_1)} \tag{3.59}$$

generalized *Bayes receiver* $$C_{21}P(s_1)p(v/s_1) > C_{12}P(s_2)p(v/s_2) \tag{3.60}$$

Optimum Receiver

We assume that the received signal is a binary waveform. One binary digit (bit) is represented by a signal waveform $s_1(t)$ which persists for time T, while the other bit is represented by the waveform $s_2(t)$ which also lasts for an interval T. For example, in the case of transmission at baseband, as shown in Fig. 3.27, $s_1(t) = +V$, while $s_2(t) = -V$; for other modulation systems, different waveforms are transmitted. For example, for PSK signaling, $s_1(t) = A \cos \omega_0 t$ and $s_2(t) = -A \cos\omega_0 t$; while for FSK, $s_1(t) = A \cos(\omega_0 + \Omega)t$ and $s_2(t) = A \cos(\omega_0 - \Omega)t$.

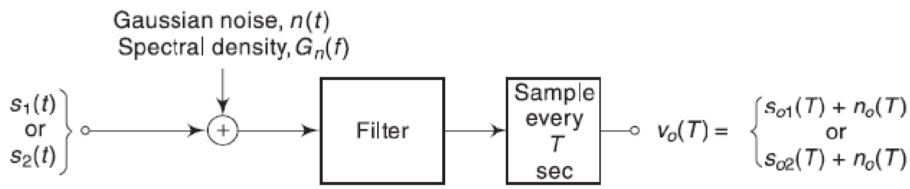


Figure 3.32 A receiver for binary coded signaling

An error [we decide $s_1(t)$ is transmitted rather than $s_2(t)$] will result if

$$n_o(T) \geq \frac{s_{o1}(T) - s_{o2}(T)}{2} \tag{3.61}$$

probability of error is $$P_e = \int_{[s_{o1}(T)-s_{o2}(T)]/2}^{\infty} \frac{e^{-n_o^2(T)/2\sigma_o^2}}{\sqrt{2\pi\sigma_o^2}} dn_o(T) \quad = \frac{1}{2}\operatorname{erfc}\left[\frac{s_{o1}(T) - s_{o2}(T)}{2\sqrt{2}\sigma_o}\right]$$

for the case $s_{o1}(T) = VT/\tau$ and $s_{o2}(T) = -VT/\tau$, $$P_e = \frac{1}{2}\operatorname{erfc}\left(\frac{V^2 T}{\eta}\right)^{1/2}$$

The complementary error function is monotonically decreasing function of its argument (indicated in Fig. 3.30). Hence, as is to be anticipated, P_e decreases as the difference $s_{o1}(T) - s_{o2}(T)$ becomes larger and as the r.m.s. noise voltage σ_o becomes smaller. The optimum filter, then, is the filter which maximizes the ratio

$$\gamma = \frac{s_{o1}(T) - s_{o2}(T)}{\sigma_o} \tag{3.62}$$

We shall now calculate the transfer function $H(f)$ of this optimum filter. As a matter of mathematical convenience we shall actually maximize γ^2 rather than γ

Calculation of the Optimum-Filter Transfer Function $H(f)$

Signal to the optimum filter is $p(t) \equiv s_1(t) - s_2(t)$

Corresponding *output signal* of the filter is $p_o(t) \equiv s_{o1}(t) - s_{o2}(t)$

Let $P(f)$ and $P_o(f)$ be the Fourier transforms, respectively, of $p(t)$ and $p_o(t)$. Then

$$P_o(f) = H(f)P(f) \tag{3.63}$$

$$p_o(T) = \int_{-\infty}^{\infty} P_o(f)e^{j2\pi fT}\,df = \int_{-\infty}^{\infty} H(f)P(f)e^{j2\pi fT}\,df \qquad (3.64)$$

$$G_{n_o}(f) = |H(f)|^2 G_n(f)\,df \qquad (3.65)$$

Normalized output noise power $\sigma_o^2 = \int_{-\infty}^{\infty} G_{n_o}(f)\,df = \int_{-\infty}^{\infty} |H(f)|^2 G_n(f)\,df$ (3.66)

$$\gamma^2 = \frac{p_o^2(T)}{\sigma_o^2} = \frac{\left|\int_{-\infty}^{\infty} H(f)P(f)e^{j2\pi Tf}\,df\right|^2}{\int_{-\infty}^{\infty} |H(f)|^2 G_n(f)\,df} \qquad (3.67)$$

Schwarz inequality defines $\left|\int_{-\infty}^{\infty} X(f)Y(f)\,df\right|^2 \le \int_{-\infty}^{\infty} |X(f)|^2\,df \int_{-\infty}^{\infty} |Y(f)|^2\,df$ (3.68)

The equal sign applies when $X(f) = KY^*(f)$ (3.69)

$$\frac{p_o^2(T)}{\sigma_o^2} = \frac{\left|\int_{-\infty}^{\infty} X(f)Y(f)\,df\right|^2}{\int_{\infty}^{\infty} |X(f)|^2\,df} \le \int_{-\infty}^{\infty} |Y(f)|^2\,df \qquad (3.70)$$

Or, $$\frac{p_o^2(T)}{\sigma_o^2} = \int_{-\infty}^{\infty} |Y(f)|^2\,df = \int_{-\infty}^{\infty} \frac{|P(f)|^2}{G_n(f)}\,df \qquad (3.71)$$

$$X(f) \equiv \sqrt{G_n(f)}\,H(f) \qquad (3.72)$$

$$Y(f) \equiv \frac{1}{\sqrt{G_n(f)}} P(f)e^{j2\pi Tf}$$

The ratio $p_o^2(T)/\sigma_o^2$ will attain its maximum value when $H(f) = K\frac{P^*(f)}{G_n(f)}e^{-j2\pi fT}$ (3.73)

Optimum Filter using Matched Filter

An optimum filter which yields a maximum ratio $p_o^2(T)/\sigma_o^2$ is called a matched filter when the input noise is white. In this case $G_n(f) = \eta/2$, and equation (3.73) becomes

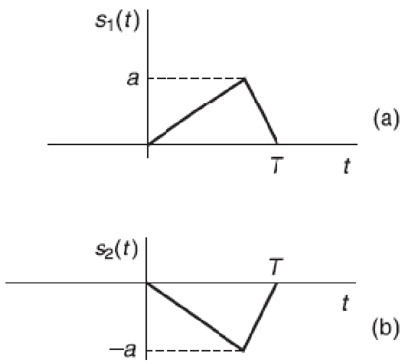
$$H(f) = K\frac{P^*(f)}{\eta/2}e^{-j2\pi fT} \qquad (3.74)$$

$$h(t) = \mathscr{F}^{-1}[H(f)] = \frac{2K}{\eta}\int_{-\infty}^{\infty} P^*(f)e^{-j2\pi fT}e^{j2\pi ft}\,df = \frac{2K}{\eta}\int_{-\infty}^{\infty} P^*(f)e^{j2\pi f(t-T)}\,df \qquad (3.75)$$

A physically realizable filter will have an impulse response which is real, $h(t) - h^*(t)$

$$h(t) = \frac{2K}{\eta}\int_{-\infty}^{\infty} P(f)e^{j2\pi f(T-t)}\,df = \frac{2K}{\eta}p(T-t) \qquad (3.76)$$

since $p(t) \equiv s_1(t) - s_2(t) \quad h(t) = \dfrac{2K}{\eta}[s_1(T-t) - s_2(T-t)]$ (3.77)



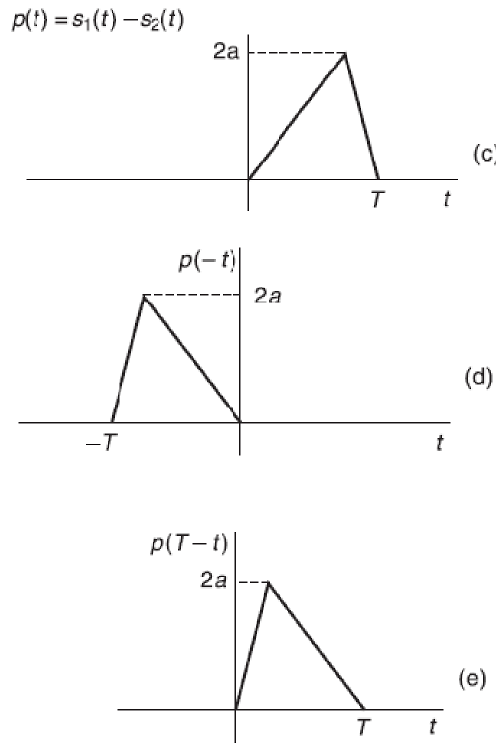


Figure 3.33

The signals (a) $s_1(t)$, (b) $s_2(t)$, and (c) $p(t) = s_1(t) - s_2(t)$ (d) $p(t)$ rotated about the axis $t = 0$. (e) The waveform in (d) translated to the right by amount T.

Probability of Error of Matched Filter

With $G_n(f) = \eta/2$, $\left[\dfrac{p_o^2(T)}{\sigma_o^2}\right]_{max} = \dfrac{2}{\eta}\displaystyle\int_{-\infty}^{\infty} |P(f)|^2\, df$ (3.78)

From Parseval's theorem, $\displaystyle\int_{-\infty}^{\infty} |P(f)|^2\, df = \int_{-\infty}^{\infty} p^2(t)\, dt = \int_0^T p^2(t)\, dt$ (3.79)

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{max} = \frac{2}{\eta}\int_0^T [s_1(t) - s_2(t)]^2\, dt \qquad (11.50a)$$

$$= \frac{2}{\eta}\left[\int_0^T s_1^2(t)\,dt + \int_0^T s_2^2(t)\,dt - 2\int_0^T s_1(t)\,s_2(t)\,dt\right]$$

$$= \frac{2}{\eta}(E_{s1} + E_{s2} - 2E_{s12})$$

(3.80)

The optimum choice of $s_2(t)$ is as given by $s_2(t) = -\,s_1(t)$ (3.81)

Hence, $E_{s1} = E_{s2} = -E_{s12} \equiv E_s$ (3.82)

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = \frac{8E_s}{\eta} \tag{3.83}$$

$$(P_e)_{\min} = \frac{1}{2}\operatorname{erfc}\left\{\frac{1}{8}\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max}\right\}^{1/2} = \frac{1}{2}\operatorname{erfc}\left(\frac{E_s}{\eta}\right)^{1/2} \tag{3.84}$$

Integrator as Matched Filter

When we have, $s_1(t) = V \quad 0 \le t \le T$
$s_2(t) = -V \quad 0 \le t \le T$ (3.85)

Impulse response of the matched filter is, $h(t) = \frac{2K}{\eta}[s_1(T-t) - s_2(T-t)]$ (3.86)

$s_1(T-t) - s_2(T-t)$ is a pulse of amplitude $2V$ extending from $t = 0$ to $t = T$

Hence, $h(t) = \frac{2K}{\eta}(2V)[u(t) - u(t-T)]$ (3.87)

The inverse transform of $h(t)$, that is, the transfer function of the filter, becomes, with s the Laplace transform variable,

$$H(s) = \frac{1}{s} - \frac{e^{-sT}}{s} \tag{3.88}$$

The first term in equation (3.88) represents an integration beginning at $t = 0$, while the second term represents an integration with reverse polarity beginning at $t = T$.

Optimum Filter using Correlator

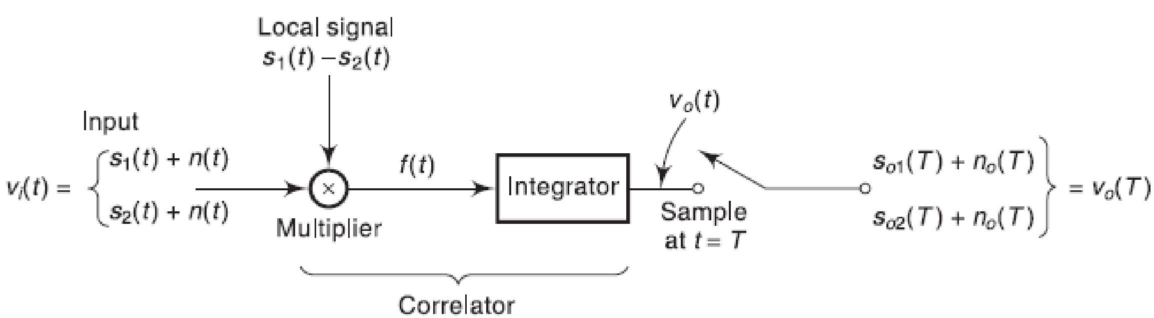


Figure 3.34 A coherent system of signal reception

$$S_o(T) = \frac{1}{\tau}\int_0^T s_i(t)[s_1(t) - s_2(t)]\,dt \tag{3.89}$$

$$n_o(T) = \frac{1}{\tau}\int_0^T n(t)[s_1(t) - s_2(t)]\,dt \tag{3.90}$$

If $h(t)$ is the impulsive response of the matched filter, then

$$v_o(T) = \int_{-\infty}^{\infty} v_i(\lambda)h(t-\lambda)\,d\lambda = \int_0^T v_i(\lambda)\,(t-\lambda)\,d\lambda \tag{3.91}$$

$$h(t) = \frac{2K}{\eta}[s_1(T-t) - s_2(T-t)] \tag{3.92}$$

$$v_o(t) = \frac{2K}{\eta}\int_0^T v_i(\lambda)[s_1(T-t+\lambda) - s_2(T-t+\lambda)]\, d\lambda \tag{3.93}$$

$$v_o(t) = \frac{2K}{\eta}\int_0^T v_i(\lambda)[s_1(T-t+\lambda) - s_2(T-t+\lambda)]\, d\lambda \tag{3.94}$$

Since $v_i(\lambda) = s_i(\lambda) + n(\lambda)$, and $v_o(t) = s_o(t) + n_o(t)$, setting $t = T$ yields

$$s_o(T) = \frac{2K}{\eta}\int_0^T s_i(\lambda)[s_1(\lambda) - s_2(\lambda)]d\lambda \tag{3.95}$$

Where, $s_i(\lambda)$ is equal to $s_1(\lambda)$ or $s_2(\lambda)$

Similarly, $n_o(T) = \frac{2K}{\eta}\int_0^T n(\lambda)[s_1(\lambda) - s_2(\lambda)]d\lambda$ (3.96)

Thus $s_o(t)$ and $n_o(t)$, as calculated from equations (3.89) and (3.90) for the correlation receiver, and as calculated from equations (3.95) and (3.96) for the matched filter receiver, are identical. Hence the performances of the two systems are identical.

Optimal Coherent Reception: PSK

The input signal is $s_1(t) = A\cos\omega_0 t$, $s_2(t) = -A\cos\omega_0 t$ (3.97)

In PSK, $s_1(t) = -s_2(t)$, Equation (3.84) gives the error probability as in base band transmission

$$P_e = \frac{1}{2}\text{erfc}\sqrt{\frac{E_s}{\eta}} = \frac{1}{2}\text{erfc}\sqrt{\frac{A^2T}{2\eta}} \tag{3.98}$$

Imperfect Phase Synchronization $P_e = \frac{1}{2}\text{erfc}\sqrt{\frac{E_s}{\eta}\cos^2\phi}$ (3.99)

Imperfect Bit Synchronization

$$\begin{aligned} s_o(T+\tau) &= \frac{2K}{\eta}\int_\tau^T A\cos\omega_0 t[2A\cos\omega_0 t]\,dt - \frac{2K}{\eta}\int_T^{T+\tau} A\cos\omega_0 t[2A\cos\omega_0 t] \\ &= \frac{2K}{\eta}[A^2(T-\tau) - A^2\tau] = \frac{2K}{\eta}[A^2T]\left[1 - \frac{2\tau}{T}\right] \end{aligned} \tag{3.100}$$

If the overlap is in the other direction, integration extends from $-\tau$ to $T-\tau$

$$s_o(T+\tau) = \frac{2K}{\eta}[A^2T]\left[1 - \frac{2|\tau|}{T}\right] \tag{3.101}$$

Correspondingly, $P_e = \frac{1}{2}\text{erfc}\sqrt{\left(\frac{E_s}{\eta}\right)\left(1 - \frac{2|\tau|}{T}\right)^2}$ (3.102)

If $\tau = 0.05T$, the probability of error is increased by a factor 10

If both phase error and timing error are present, then

Probability of error $P_e = \frac{1}{2}\text{erfc}\left[\left(\frac{E_s}{\eta}\right)(\cos^2\phi)\left(1-\frac{2|\tau|}{T}\right)^2\right]^{1/2}$ (3.103)

Optimal Coherent Reception: FSK

$$s_1(t) = A\cos(\omega_0 + \Omega)t$$
$$s_2(t) = A\cos(\omega_0 - \Omega)t \tag{3.104}$$

Local waveform is $s_1(t) - s_2(t) = A\cos(\omega_0 + \Omega)t - A\cos(\omega_0 - \Omega)t$ (3.105)

$s_1(t) = -s_2(t)$ assumption is obviously not valid for FSK

We start with $\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = \frac{2}{\eta}\int_0^T [s_1(t) - s_2(t)]^2\, dt$ (3.106)

Substituting $s_1(t)$ and $s_2(t)$

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = \frac{2A^2T}{\eta}\left[1 - \frac{\sin 2\Omega T}{2\Omega T} + \frac{1}{2}\frac{\sin[2(\omega_o + \Omega)T}{2(\omega_o + \Omega)T} - \frac{1}{2}\frac{\sin[2(\omega_o - \Omega)T}{2(\omega_o - \Omega)T} - \frac{\sin 2\omega_o T}{2\omega_o T}\right] \tag{3.107}$$

If we assume that the offset angular frequency Ω is very small $\omega_0 T \gg 1$

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = \frac{2A^2T}{\eta}\left(1 - \frac{\sin 2\Omega T}{2\Omega T}\right) \tag{3.108}$$

Largest value when Ω is selected so that $2\Omega T = 3\pi/2$

$$\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max} = 2.42\frac{A^2T}{\eta} = 4.84\frac{(A^2/2)T}{\eta} \tag{3.109}$$

$$P_e = \frac{1}{2}\text{erfc}\left\{\frac{1}{8}\left[\frac{p_o^2(T)}{\sigma_o^2}\right]_{\max}\right\}^{1/2} \simeq \frac{1}{2}\text{erfc}\left(0.6\frac{E_s}{\eta}\right)^{1/2} \tag{3.110}$$

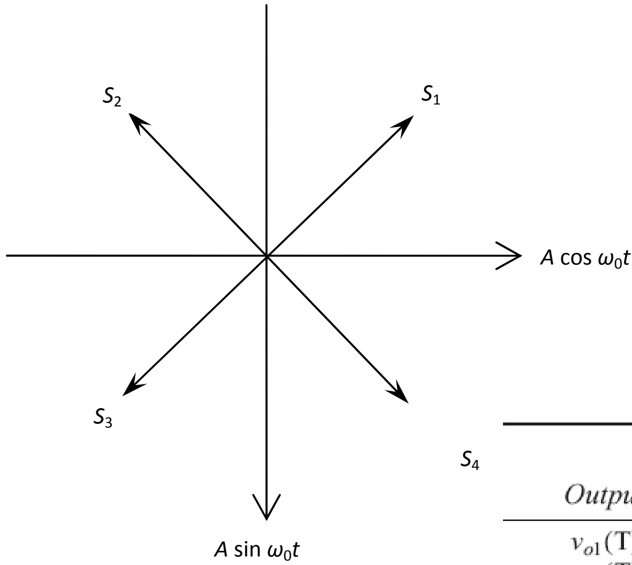
Where, the signal energy is $E_s = A^2T/2$

When one of two *orthogonal* frequencies are transmitted, $2\Omega T = m\pi$

$$P_e = \frac{1}{2}\text{erfc}\left(\frac{E_s}{2\eta}\right)^{1/2} \tag{3.111}$$

Comparing the probability of error obtained for FSK [Eq. (3.110)] with probability of error obtained for PSK [Eq. (3.98)], we see that equal probability of error in each system can be achieved if the signal energy in the PSK signal is 0.6 times as large as the signal energy in FSK. As a result, a 2 dB increase in the transmitted signal power is required for FSK. Why is FSK inferior to PSK? The answer is that in PSK, $s_1(t) = -s_2(t)$, while in FSK this condition is not satisfied. Thus, although an optimum filter is used in each case, PSK results in considerable improvement compared with FSK.

Optimal Coherent Reception: QPSK



	Signal			
Output	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$
$v_{o1}(\mathrm{T})$	$+V_o$	$-V_o$	$-V_o$	$+V_o$
$v_{o2}(\mathrm{T})$	$-V_o$	$-V_o$	$+V_o$	$+V_o$

Figure 3.35 A phasor diagram representation of the signals in QPSK

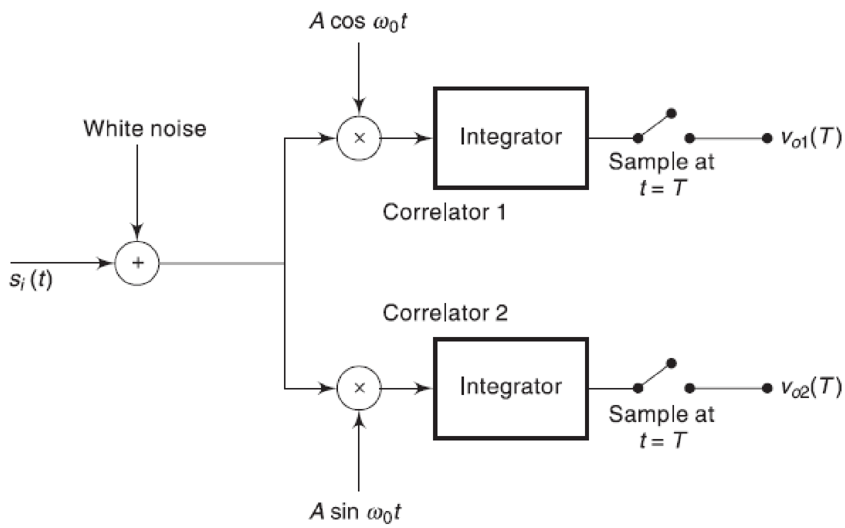


Figure 3.36 A correlation receiver for QPSK

We note from Fig. 3.35, that the reference waveform of correlator 1 is an angle $\phi = 45^0$ to the axes of orientation of all of the four possible signals. Hence, from equation (3.99), since $(\cos 45^0)^2$ = ½, the probability that correlator 1 or correlator 2 will make an error is

$$P'_{e1} = P'_{e2} = \frac{1}{2}\operatorname{erfc}\sqrt{\frac{A^2 T_s}{4\eta}} \tag{3.112}$$

to compare this result to the result obtained for BPSK $\quad T_s = 2T$

$$P'_{e1} = P'_{e2} = \frac{1}{2}\operatorname{erfc}\sqrt{\frac{A^2 T}{2\eta}} = P_e(\text{BPSK}) \tag{3.113}$$

$$P_c = (1 - P_e')(1 - P_e') = 1 - 2P_e' + P_e'^2 \tag{3.114}$$

$$P_e' \ll 1, \quad P_e(\text{QPSK}) = 1 - P_c \simeq 2P_e' = \operatorname{erfc}\sqrt{\frac{A^2 T}{4\eta}} \tag{3.115}$$

Module – IV **12 hours**

BPSK

Generation

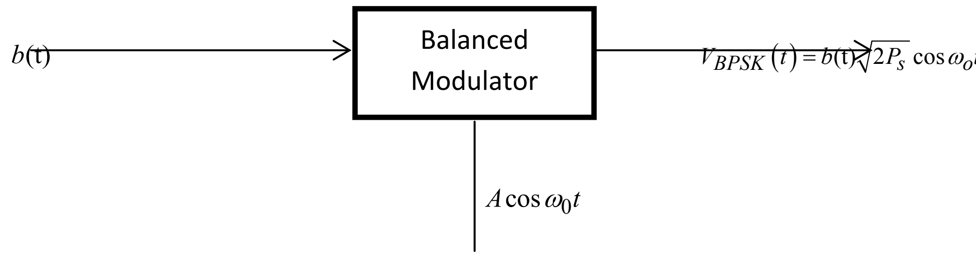


Fig.4.1 Balanced Modulator

Transmitted signals are

$$V_H(t) = V_{BPSK}(t) = \sqrt{2P_s}\cos\omega_o t \quad (4.1)$$

$$V_H(t) = V_{BPSK}(t) = \sqrt{2P_s}(\cos\omega_o t + \pi) = -\sqrt{2P_s}\cos\omega_o t \quad (4.2)$$

In BPSK the data b(t) in a stream of binary digit with voltage levels which as a matter of convenience, we take +1 V and -1 V. So BPSK can be written as

$$V_{BPSK}(t) = b(t)\sqrt{2P_s}\cos\omega_o t \quad (4.3)$$

Transmission

This $V_{BPSK}(t)$ signal is transmitted through the channel. While it moves in the transmission path of the channel, the phase of the carrier may be changed at the output of the receiver. So the BPSK signal received at the input of the receiver can be taken as $V_{BPSK}(t) = b(t)\sqrt{2P_s}\cos(\omega_o t + \phi)$ where $\Delta t = \phi / \omega_o$ is the time delay.

Receiver

$$b(t)\sqrt{2P_s}\cos^2(\omega_o t + \phi) = b(t)\sqrt{2P_s}\left[\frac{1}{2} + \frac{1}{2}\cos 2(\omega_o t + \phi)\right]$$

$$V_o(kT_b) = b(kT_b)\sqrt{2P_s}\int_{(k-1)T_b}^{kT_b}\frac{1}{2}dt + b(kT_b)\sqrt{2P_s}\int_{(k-1)T_b}^{kT_b}\frac{1}{2}\cos 2(\omega_o t + \phi)dt = b(kT_b)\sqrt{\frac{P_s}{2}}T_b \quad (4.4)$$

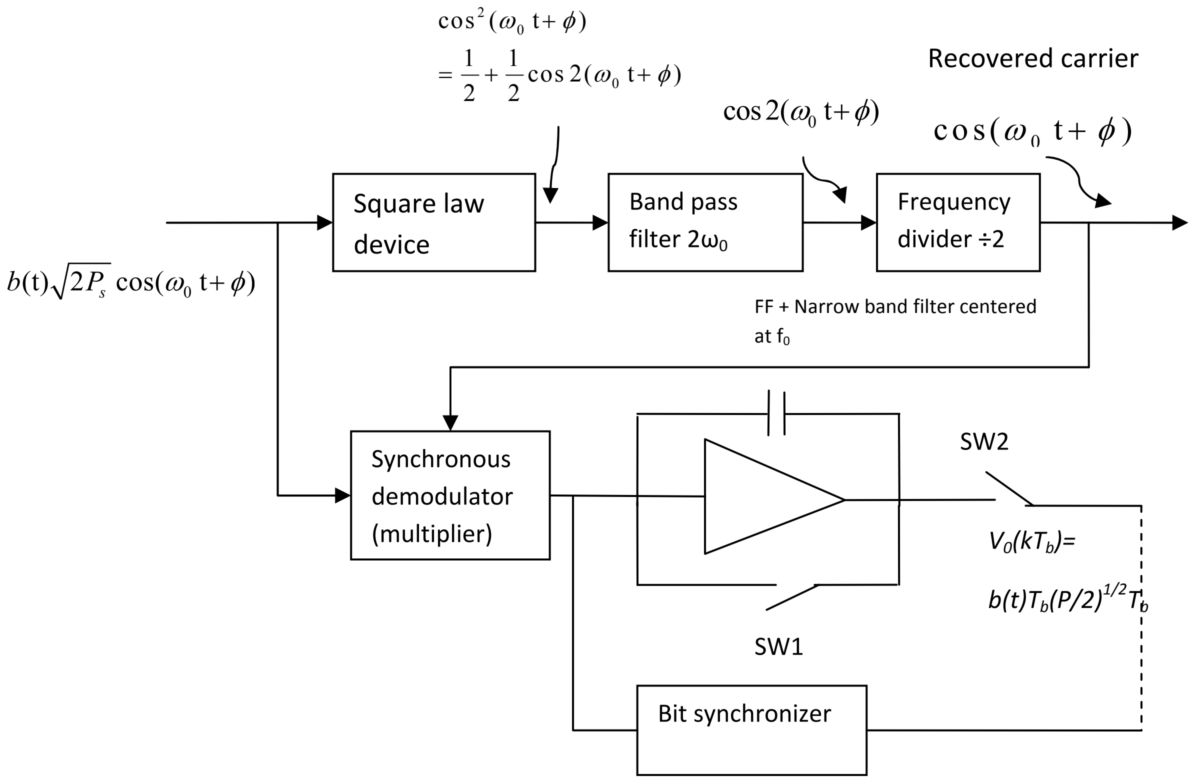


Fig. 4.2 Synchronous demodulator

Spectrum

The waveform b(t) is a NRZ binary waveform which makes an excursion between $+\sqrt{P_S}$ and $-\sqrt{P_S}$. The PSD of this waveform

$$G_b(\mathrm{f}) = \mathrm{P}_s\mathrm{T}_b\left(\frac{\sin\pi f T_b}{\pi f T_b}\right)^2 \tag{4.5}$$

The BPSK waveform is the NRZ waveform multiplied by $\sqrt{2P_S}\cos\omega_o t$. Thus the power spectral density of the BPSK signal is

$$G_{BPSK}(\mathrm{f}) = \mathrm{P}_s\mathrm{T}_b/2\left\{\left[\frac{\sin\pi(f-f_o)T_b}{\pi(f-f_o)T_b}\right]^2+\left[\frac{\sin\pi(f+f_o)T_b}{\pi(f+f_o)T_b}\right]^2\right\} \tag{4.6}$$

Simultaneous bit transmission and thereafter overlapping of spectra is known as inter-channel interference. Restricting the overlapping by considering the principal lobe to transmit 90% of power ultimately cause inter-symbol interference.

Geometrical representation of BPSK signals:

When BPSK signal can be represented, in terms of one orthogonal signal

$u_1(t)=\sqrt{2/T_b}\cos\omega_o t$ as

$$V_{BPSK}(t)=\left[\sqrt{P_sT_b}\,b(\mathrm{t})\right]\sqrt{\frac{2}{T_b}}\cos\omega_o t=\left[\sqrt{P_sT_b}\,b(\mathrm{t})\right]u_1(\mathrm{t}) \tag{4.7}$$

The distance between signals is

$$d = 2\sqrt{P_s T_b} = 2\sqrt{E_b} \tag{4.8}$$

$$d \propto \frac{1}{P_e} \tag{4.9}$$

DPSK(Differential Phase Shift Keying)

In BPSK receiver to regenerate the carrier we start by squaring $b(t)\sqrt{2P_s}\cos\omega_o t$ .but when the received signal were instead $-b(t)\sqrt{2P_s}\cos\omega_o t$, the recovered carrier would remain as before. Therefore we shall not be able to determine whether the received baseband is transmitted signal b(t) or its negative i.e. -b(t). DPSK and DEPSK are modification of BPSK which have the merit that they eliminate the ambiguity about whether the demodulated data is actual or inverted. In addition DPSK avoids the need to provide the synchronous carrier required at the demodulator for detecting a BPSK signal.

Transmitter (Generation)

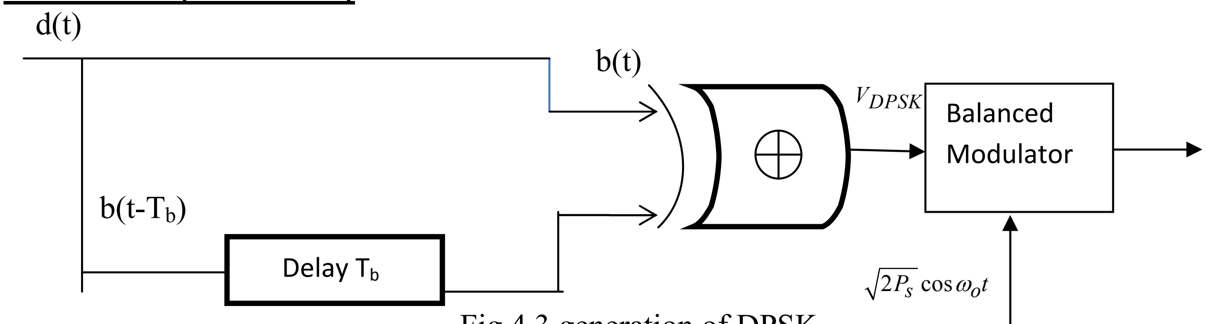


Fig.4.3 generation of DPSK

Here, $b(t) = d(t) \oplus b(t - T_b)$ (4.10)

b(0) cannot be found unless we know d(0) and b(-1). Here we have b(0) = 0, b(+1) = 0 so d(1) should be 0.

In this Fig. 4.3, d(0) & b(-1) is not shown. Here we have chosen b(0) = 0. If we choose b(0) = 1, then there is no problem in detection of b(t).

$$\begin{aligned} V_{DPSK}(t) &= b(t)\sqrt{2P_s}\cos\omega_o t \\ &= \pm\sqrt{2P_s}\cos\omega_o t \end{aligned} \tag{4.11}$$

Transmission

When $V_{DPSK}(t)$ is transmitted from the generator to the channel, at passes through the channel, then b(t) may be changed to –b(t) before reaching receiver.

Receiver

$b(t)b(t-T_b) = 1,$ if d(t) = 0

but $b(t)b(t-T_b) = -1,$ if d(t) = 1

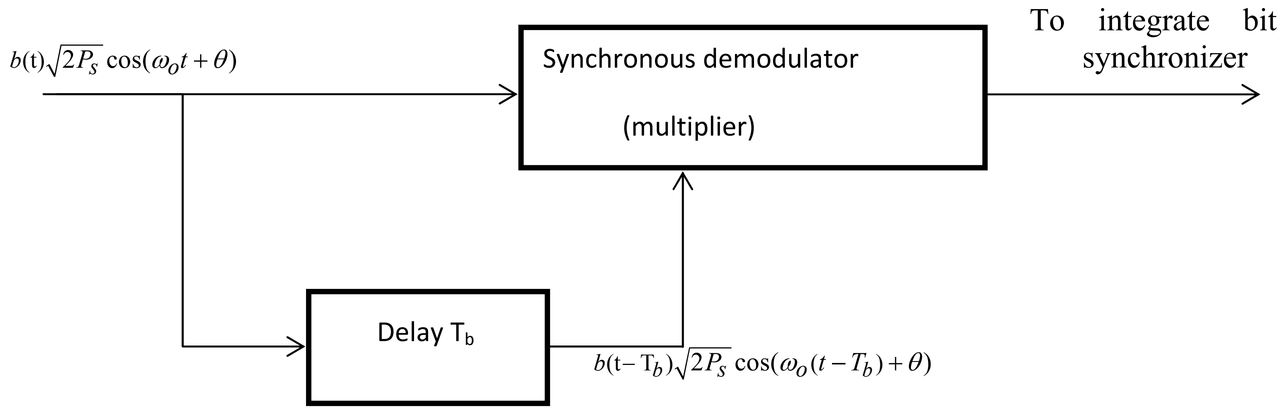


Fig. 4.6 Receiver of DPSK

<u>Advantage of DPSK over BPSK</u>

1. Local carrier generation not required and receiver circuit is simple.
2. If whole of the bits of b(t) is inverted then also correct d(t) can be recovered.

<u>Disadvantage</u>

1. Noise in one bit interval may cause errors to two bit determination that is a tendency for bit errors to occur in pairs. The single errors are also possible.
2. Specrum of DPSK is same as BPSK .the geometrical representation of DPSK is same as BPSK.

<u>DEPSK (Differentially Encoded Phase Shift Keying)</u>

DPSK demodulator requires a device which operates at the carrier frequency and provides a delay of T_b. Differentially encoded PSK eliminates the need for such a piece of hardware Transmitter or generator is same as DPSK

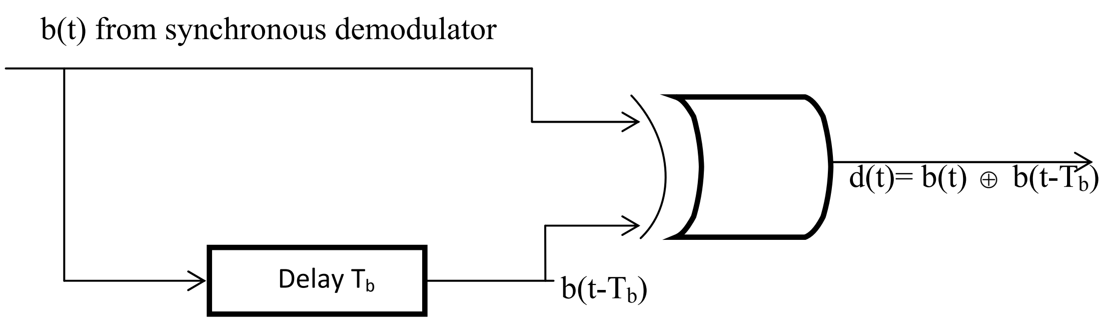


Fig 4.7a Generation of DEPSK

<u>QPSK (Quadrature Phase Shift Keying)</u>

The transmission bandwidth of bit NRZ signal is f_b. So the transmission rate is $2f_b$bps.Hence to transmit BPSK signal the channel must have a bandwidth of $2f_b$. QPSK has been formulated to allow the bits to be transmitted using half the bandwidth. D-flip flop is used in QPSK transmitter to operate as one bit storage device.

Generation

QPSK transmitter

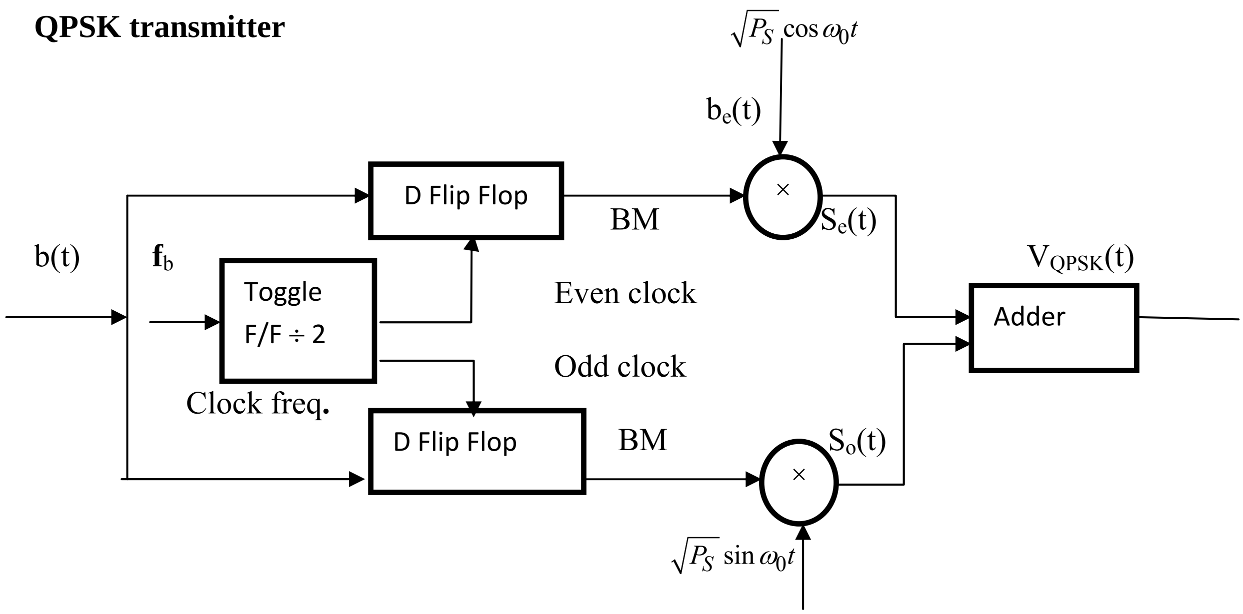


Fig 4.7b Generation of QPSK

Transmission

Due to finite distance between generator and receiver the signal available at receiver may have some phase change so,

$$V_{QPSK}(t) = k_1\sqrt{P_s}\,b_o(t)\sin(w_0 t+\theta) + k_2\sqrt{P_s}\,b_e(t)\cos(w_0 t+\theta) \qquad (4.12)$$

Reception

QPSK receiver

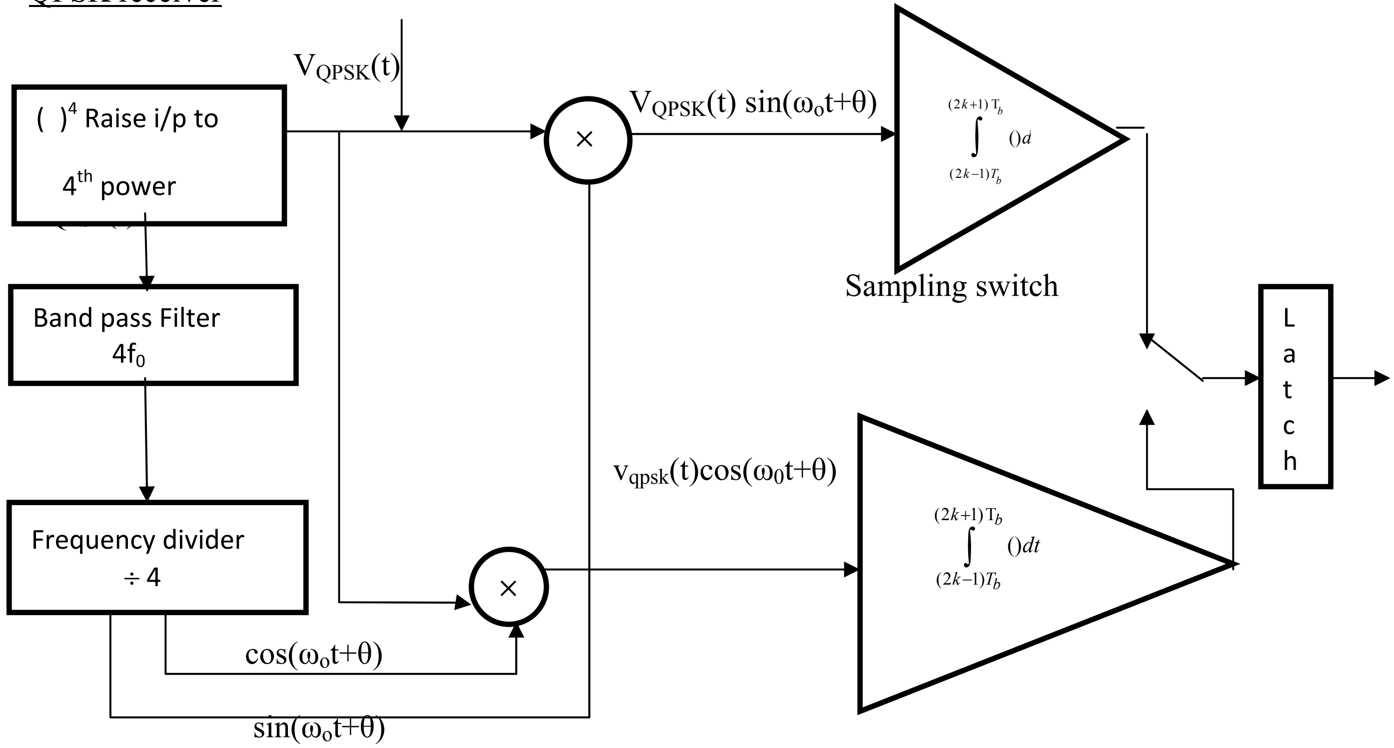


Fig 4.8 Reception of QPSK

Samples are taken alternatively from one and the other integrator output at the end of each bit time T_b and these samples are half in the latch for the bit time T_b and these samples half in the latch for the bit time T_b. Each individual integrator output is sampled at intervals $2T_b$. The latch output is the recovered bit stream b(t).

Spectrum:

The waveform $b_o(t)$ or $b_e(t)$ (if NRZ) is binary waveform makes an excursion $+\sqrt{P_s}$ and $-\sqrt{P_s}$. The PSD of this waveform

$$G_{b_o}(f) = G_{b_e}(f) = P_s(2T_b)\left\{\frac{\sin \pi f(2T_b)}{\pi f(2T_b)}\right\}^2 \tag{4.13}$$

When QPSK signal is multiplied by $\cos\omega_o t$. Then the PSD of the QPSK signal is

$$G_{QPSK}(f) = P_s T_b\left[\left\{\frac{\sin \pi(f-f_o)(2T_b)}{\pi(f-f_o)(2T_b)}\right\}^2 + \left\{\frac{\sin \pi(f+f_o)(2T_b)}{\pi(f+f_o)(2T_b)}\right\}^2\right] \tag{4.14}$$

Symbol versus bit transmission

In BPSK we deal with each bit individually in its duration T_b. In QPSK we lump two bits together to form what is termed a symbol. The symbol can have any one of four possible values corresponding to the two bit sequence 00, 01, 10 and 11. We therefore arrange to make four distinct signals available for transmission. At the receiver each signal represents one symbol and correspondingly two bits. When bits are transmitted, as in BPSK, the signal changes occur at the bit rate. When symbols are transmitted the changes occur at the symbol rate which is one half the bit rate. Thus the symbol time is $T_s = 2T_b$ (OQPSK). $T_s = T_b$(QPSK)

Geometrical representation of QPSK signals in signal space

Four symbols are four quadrature signals. These are to be represented in signal space. One possibility representing the QPSK signal in one equation is

$$V_{QPSK} = \sqrt{2P_s}\cos\left[\omega_o t + (2m+1)\tfrac{\pi}{4}\right] \quad ; m=0, 1, 2, 3 \tag{4.15}$$

$$V_{QPSK} = \sqrt{2P_s}\cos[(2m+1)\frac{\pi}{4}]\cos w_0 t - \sqrt{2P_s}\sin\{(2m+1)\frac{\pi}{4}\}\sin w_0 t \tag{4.16}$$

To represent this signal in signal space, two ortho-normal signals are be selected. They can be

$$U_1(t) = \sqrt{\frac{2}{T}}\cos w_0 t \text{ and } U_2(t) = \sqrt{\frac{2}{T}}\sin w_0 t$$

So V_{QPSK} can be written as

$$V_{QPSK} = \left[\sqrt{P_s T}\cos(2m+1)\tfrac{\pi}{4}\right]\sqrt{\tfrac{2}{T}}\cos\omega_0 t - \left[\sqrt{P_s T}\sin(2m+1)\tfrac{\pi}{4}\right]\sqrt{\tfrac{2}{T}}\sin\omega_0 t \tag{4.17}$$

b_o and b_e take values as +1 or -1. So we can write the same V_{QPSK} signal as

$$V_{QPSK} = \sqrt{E_b} b_e(\mathrm{t}).\mathrm{u}_1(\mathrm{t}) - \sqrt{E_b} b_o(\mathrm{t}).\mathrm{u}_2(\mathrm{t}) \qquad (4.18)$$

Where,

$$b_e(\mathrm{t}) = \sqrt{2}\cos(2m+1)\frac{\pi}{4}$$
$$b_o(\mathrm{t}) = -\sqrt{2}\sin(2m+1)\frac{\pi}{4} \qquad (4.19)$$

In the above equations $T = 2T_b$. Working at above signals four symbols can be shown in signal space as shown below. Four dots in the signal space represents four symbol. The distance of signal point form the origin is $\sqrt{E_s}$, which in the square root of the signal energy associated with the symbol. i.e $E_s = P_s T_s = 2 P_s T_b$. The signal points which differ in a signal bit are separates by the distance $d = \sqrt{P_s T_b} = \sqrt{E_b}$. Noise immunity in QPSK is same as BPSK.

M-ary Phase shift keying

In BPSK we transmit each bit individually. Depending on Whether b(t) is logic 0 or logic 1, We transmit one or another of sinusoid for the bit time T_b, the sinusoids differ in phase by $2\pi / 2 = 180^{\cdot}$. In QPSK We lump together two bits. Depending on which of the four two-bit words develops, we transmit one or another of four sinusoids of duration $2\pi / \mathrm{M}$,the sinusoids differing in phase by amount $2\pi / 4 = 90^{\cdot}$. The scheme can be extended. Let us lump together N bits so that in this N- bit symbol, extending over the $\mathrm{N}T_b$, there are $2^N = M$ possible symbol as shown in Fig. 4.9. Now let us represent the symbols by sinusoids of duration $\mathrm{N}T_b = T_s$ which differ from one another by the phase $2\pi / \mathrm{M}$. Hardware to accomplish this M-ary communication is available. So

$$V_{M-aryPSK} = (\sqrt{2P_s}\cos\phi_m)\cos w_0 t - (\sqrt{2P_s}\sin\phi_m)\sin w_0 t \qquad \mathrm{m} = 0,1,2,3 - - - - (\mathrm{M}-1) \qquad (4.20)$$

Where, $\phi_m == (2m+1)\frac{\pi}{M}$

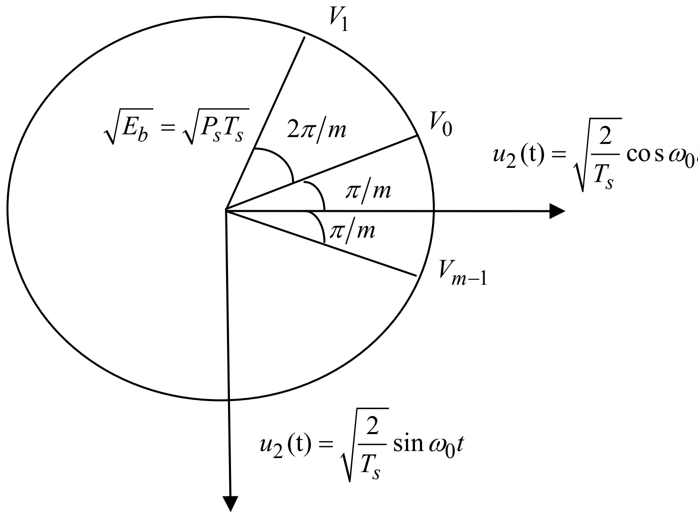


Fig 4.9 Spectrum of M-ary PSK

The co-ordinate are the orthogonal waveforms $u_1(\mathrm{t}) = \sqrt{\frac{2}{T_S}} \cos w_0 t$ and $u_2(\mathrm{t}) = \sqrt{\frac{2}{T_S}} \sin w_0 t$.

$$V_{M-aryPSK} = (\sqrt{2P_s} \cos \phi_m) \cos w_0 t - (\sqrt{2P_s} \sin \phi_m) \sin w_0 t \tag{4.21}$$

$$= P_e \cos w_0 t - P_o \sin w_0 t$$

$$\begin{aligned} &where, P_e = \sqrt{2P_s} \cos \phi_m \\ &P_o = \sqrt{2P_s} \sin \phi_m \end{aligned} \tag{4.22}$$

<u>Spectrum</u>

$$G_e(\mathrm{t}) = \mathrm{G}_o(\mathrm{t}) = \mathrm{P}_s\, T_s \left(\frac{\sin \pi f T_s}{\pi f T_s} \right)^2 \tag{4.23}$$

When carrier multiplied to bit , the resultant spectrum is centered at the carrier frequency and extends normally over a $BW = B = \frac{2}{T_s} = 2 f_s = \frac{2 f_b}{N}$.

The distance between symbol signal points

$$d = \sqrt{4E_s \sin^2\left(\frac{\pi}{M}\right)} = \sqrt{4NE_b \sin^2\left(\frac{\pi}{2^N}\right)} \tag{4.24}$$

M-ary PSK Transmitter and Receiver

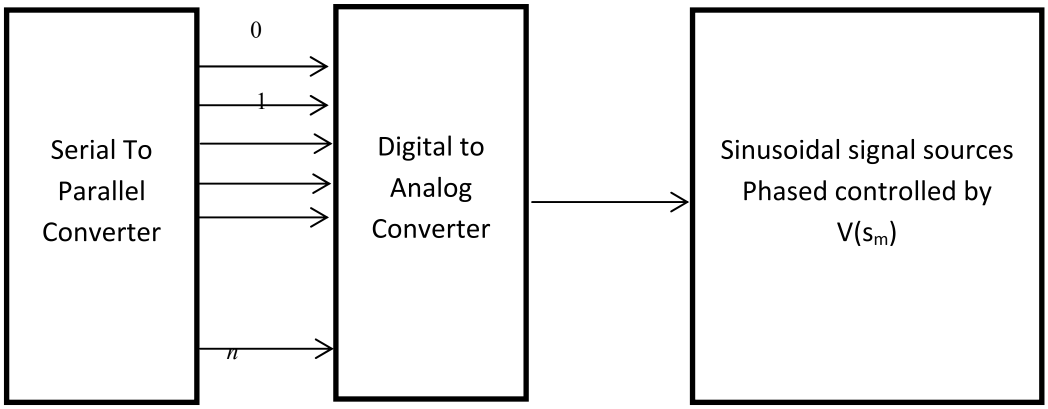


Fig 4.10 Transmission of M-ary PSK

Finally $v(s_m)$ is applied as a control input to a special type of constant amplitude sinusoidal signal source whose phase ϕ_m is determined by $v(s_m)$. Altogether, then the output is fixed amplitude, sinusoidal waveform, whose phase has a one to one correspondence to the assembled N-bit symbol. The phase can change once per symbol time.

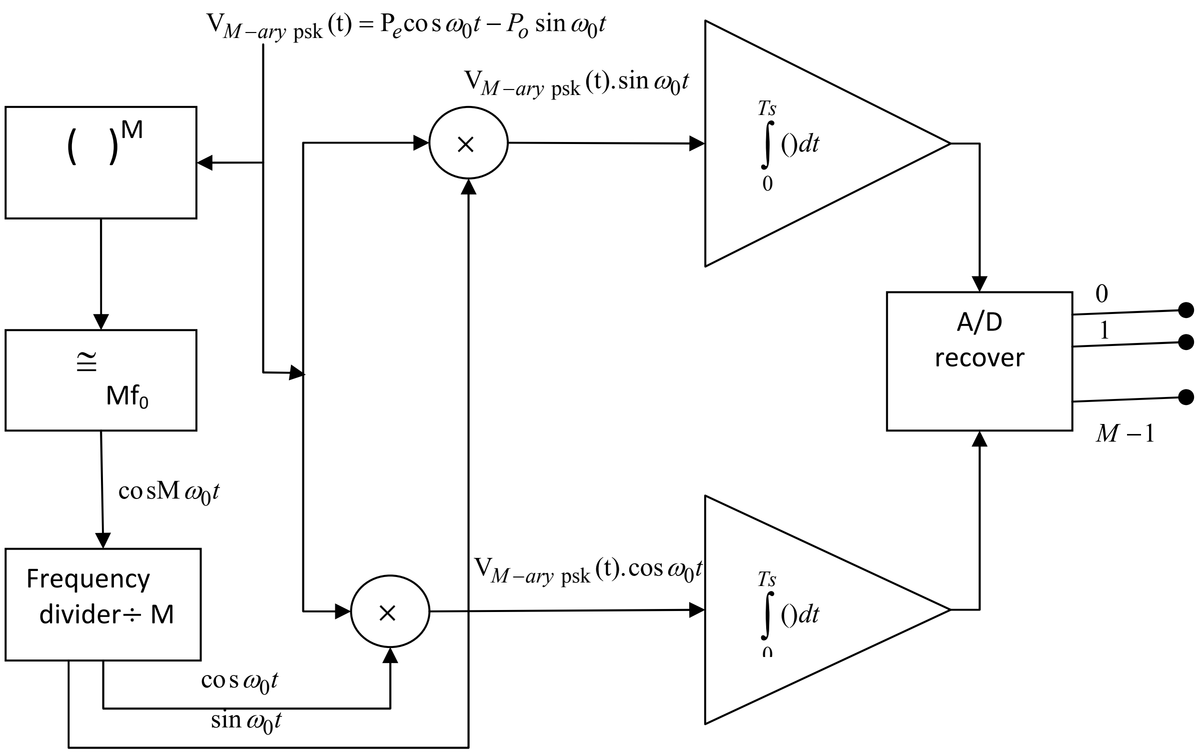


Fig 4.11 Reception of M-ary PSK

The integrator outputs are voltages whose amplitudes are proportional to $T_s P_e$ and $T_s P_o$ respectively and charge at the symbol rate. These voltages measure the components of the received signal in the directions of the quadrature phasors $\sin w_0 t$ & $\cos w_0 t$. Finally the signals $T_s P_e$ and $T_s P_o$ are applies to advice which reconstructs the digital N-bit signal which constitutes the transmitted signal.

<u>BFSK (Frequency shift keying)</u>

The BFSK signal can be represented for binary data waveform b(t) as

$$V_{BFSK}(\mathrm{t}) = \sqrt{2P_s}\cos(w_0 t + b(t)\Omega t) \tag{4.25}$$

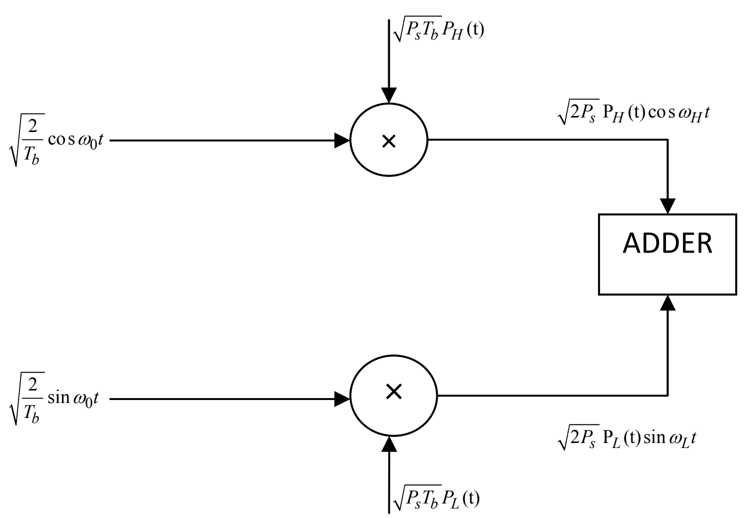
Where b(t)=+1 or -1 corresponding to the logic level 0 and 1. The transmitted signal is of amplitude $\sqrt{2P_s}$ and is either

$$\begin{aligned} V_{BFSK}(\mathrm{t}) &= \mathrm{V}_H(\mathrm{t}) = \sqrt{2P_s}\cos(w_0 + \Omega)t \\ V_{BFSK}(\mathrm{t}) &= \mathrm{V}_L(\mathrm{t}) = \sqrt{2P_s}\cos(w_0 - \Omega)t \end{aligned} \tag{4.26}$$

And thus fhas an angular frequency $w_0 + \Omega$ or $w_0 - \Omega$ with Ω a constant offset from the normal carrier frequency w_0. So, $w_H = w_0 + \Omega$ & $f_L = f_0 + \dfrac{\Omega}{2\pi} = f_0 + f_{b.} = w_0 - \Omega$

<u>Transmitter (Generation of BFSK)</u>

At any time $P_H(\mathrm{t})\,\mathrm{or}\,\mathrm{P}_L(\mathrm{t})$ is 1 but not both so that the generated signal is either at angular frequency w_H or at w_L .



b(t)	P_H(t)	P_L(t)
+1v	+1v	0v
-1v	0v	+1v

Fig 4.12 Transmission of BFSK

Receiver (Reception of BFSK)

$$f_H = f_0 + \frac{\Omega}{2\pi} = f_0 + f_{b.}$$

The BFSK signal is applied to two band pass filters one with frequency at f_H the other at f_b. Here we have assumed, that f_H - $T_s P_o$=2 f_b.The filter frequency ranges selected do not overlap and each filter has a pass band wide enough to encompass a main lobe in the spectrum of BFSK.

Hence one filter will pass nearby all the energy in the transmission at f_L. The filter outputs are applied to envelope detectors and finally the envelope detector outputs are compared by a comparator.

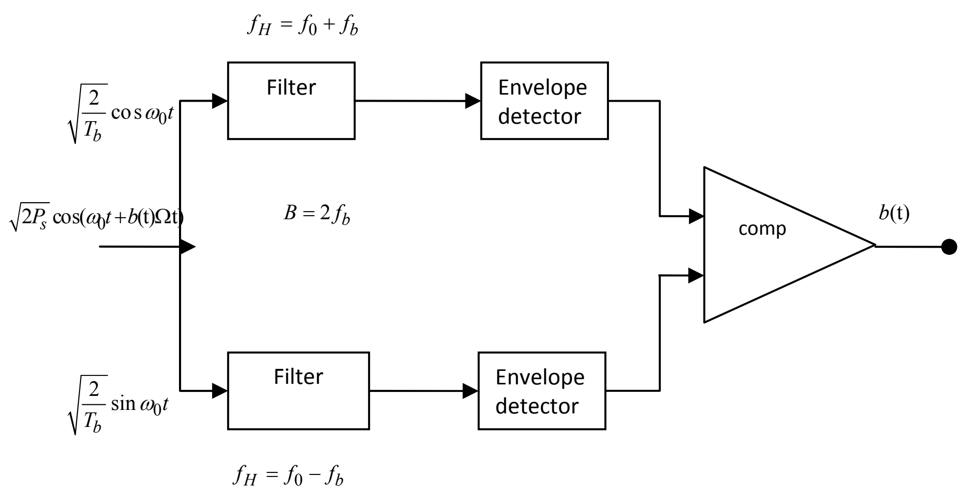


Fig 4.13 Reception of BFSK

When noise is present, the output of the comparator may vary due to the system response to the signal and noise. Thus, practical system use a bit synchronizer and an integrator and sample the comparator output only once at the end of each time interval T_b.

Spectrum(BFSK)

In terms of the variable $P_H(\mathrm{t})\,\&\,\mathrm{P}_L(\mathrm{t})$ the BFSK signal can be written as

$$V_{BFSK}(\mathrm{t}) = \sqrt{2P_s}.P_H.\cos(w_H t + \theta_H) + \sqrt{2P_s}.P_L.\cos(w_L t + \theta_L) \qquad (4.27)$$

Here each of two signals are of independent and random, uniformly distributed phase. E ach of the terms in above equation looks like the signal $\sqrt{2P_s}b(\mathrm{t})\cos\mathrm{w}_0 t$ which we encountered in BPSK, but there is an important difference. In the BPSK case, b(t) is bipolar(it alternates between +1 and-1), while in the present case P_H & P_L are unipolar (it alternates between+1 and 0). We may however, rewrite P_H & P_L as the sum of a constant and a bipolar variable, i.e.

$$P_H(\mathrm{t}) = \frac{1}{2} + \frac{1}{2} P_H{'}(\mathrm{t})$$
$$P_L(\mathrm{t}) = \frac{1}{2} + \frac{1}{2} P_L{'}(\mathrm{t}) \quad (4.28)$$

In the above equation $P_H(\mathrm{t})$ & $P_L{'}(\mathrm{t})$ are bipolar, alternating between +1 and -1 and are complementary. We have then

$$V_{BFSK}(\mathrm{t}) = \sqrt{P_s/2}\cos(w_H t + \theta_H) + \sqrt{P_s/2}\cos(w_c t + \theta_L) + \sqrt{P_s/2}\,P_H{'}\cos(w_H t + \theta_H) + \sqrt{P_s/2}\,P_L{'}\cos(w_c t + \theta_L) \quad (4.29)$$

The first terms in above equation produce a power spectral density which consists of two impulses, one at f_H and one at f_L. The last two terms produce the spectrum of two binary PSK signals, one centered at---- and one about $f_H - f_L = 2 f_b$ is assumed. For this separation $2f_b$ between f_H and f_L we observe that the overlapping between the two parts of the spectra is not large and we may expect to be able, without excessive difficulty, to distinguish the levels of the binary waveforms b(t). in any event, with this separation the bandwidth of BFSK is, $BW_{BFSK} = 4 f_b$

Geometrical representation of orthogonal BFSK in signal space

We know that any signal could be represented as $c_1\, u_1(\mathrm{t}) + c_2\, u_2(\mathrm{t})$ Where $u_1(\mathrm{t}) = \sqrt{2/T_s}\cos w_0 t$ and $u_2(\mathrm{t}) = \sqrt{2/T_s}\sin w_0 t$ are the orthogonal vectors in the signal space. $u_1(\mathrm{t})$ and $u_2(\mathrm{t})$ are orthogonal over the symbol interval T_s and if the symbol is single bit $T_s = T_b$.The coefficients c_1 & c_2 are constants. In M-ary PSK the orthogonality of the vectors u_1 and u_2 results from their phase quadrature. In the present case of BFSK it is appropriate that the orthogonality should result from a special selection of the frequencies of the unit vectors. Accordingly, with m and n integers, let us establish unit vectors.

$$u_1(\mathrm{t}) = \sqrt{\frac{2}{T_s}}\cos w_0 t$$
$$u_2(\mathrm{t}) = \sqrt{\frac{2}{T_s}}\sin w_0 t \quad (4.30)$$

In which, as usual, $f_b = \frac{1}{T_b}$. The vectors u_1 and u_2 at the mth & nth and harmonics of the fundamental frequency f_b. As we are aware, from the principles of Fourier analysis, different harmonics($m \pm n$) are orthogonal over the interval of the fundamental period $T_b = \frac{1}{f_b}$. It now the frequencies f_H and f_L in a BFSK system are selected to be

$f_H = mf_b$
$f_L = nf_b$

Then corresponding signal vectors are

$V_H(\text{t}) = \sqrt{E_b}u_1(\text{t})$ and $V_L(\text{t}) = \sqrt{E_b}u_2(\text{t})$

The signal $V_H(\text{t})$ & $V_L(\text{t})$, like vectors are orthogonal. The distance between signal end points is therefore $d = \sqrt{2E_b}$ which is considerably smaller than the distance separating end points

(i.e $d = \sqrt{2E_b}$) of BPSK signal, which are antipodal.

If we consider Non-orthogonal BFSK and $(\text{w}_H - \text{w}_L)\text{T}_b = \frac{3\pi}{2}$ then distance $d \approx \sqrt{2.4E_b}$

1. Not be as effective as BPSK in the presence of noise. Because in BFSK, since carrier is present in the spectrum and takes some energy, information bearing term is there by diminished.
2. d is less so P_e is more & SNR is less.
3. BW requirement is higher.

M-Ary FSK

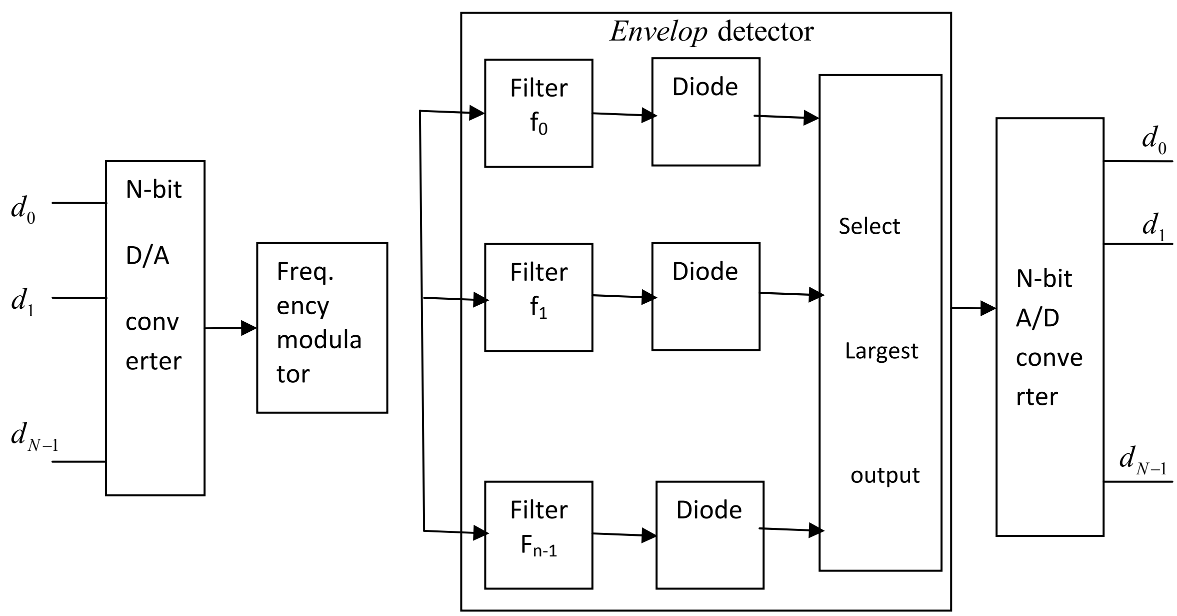


Fig 4.14 M-ary FSK

At the transmitter an N-bit symbol is presented for each T_S to an N-bit D/A converter. The converter output is applied to a frequency modulator, which generates a carrier waveform whose frequency is determined by the modulating waveform. The transmitted signal for the duration of the symbol interval, is of frequency f_0, or $f_1, or\ f_{m-1}$, where $M = 2^N$ $M = 2^N$. At the receiver, the incoming signal is applied to M parallel band pass filter with carrier frequency $f_0, f_1 ..... f_{M-1}$ and each followed by an envelope detector. The envelope detectors apply their outputs to a device which determines which of the detector indication is the largest and transmit that envelope output to an N-bit A/D converter. In this scheme the probability of error is minimized by selecting frequencies $f_0, f_1 ...... f_{M-1}$ so that the M signals are mutually orthogonal. One common employed arrangement simply provides that the carrier frequency be successive even harmonics of the symbol frequency $f_s = 1/T_s$. Thus the lowest frequency, say $f_0 = Kf_s$, while $f_1 = (\mathrm{K}+2)\mathrm{f}_s$ etc. in this case the spectral density patterns of the individual possible transmitted signals overlap, which is an extension of BFSK. It is clear that to pass M-Ary FSK the required spectral range is

$$B = 2Mf_s \qquad (4.31a)$$

Since, $f_s = \dfrac{f_b}{N}$ and $M = 2^N$

So, $B = 2^{N+1} f_b / N$ (4.31b)

M-Ary FSK required a considerably increased BW in comparison with M-Ary PSK. However as we shall see the probability of error for M-Ary FSK decreases as M increases, while for M-Ary PSK, the probability of error increases with M.

Geometrical Representation of M-Ary FSK in Signal Space

The case of M-Ary orthogonal FSK signal is extension of signal space representation for the case of orthogonal binary FSK. We can simply conceive of co-ordinate system with M mutually orthogonal co-ordinate axes. The signal vectors are parallel to these axes. The best we can do pictorially is the 3-dimensional case. The square of the length of the signal vector is the normalized energy and the distance between the signal points is

$$d = \sqrt{2E_s} = \sqrt{2NE_b} \qquad (4.32)$$

This value of d is greater than the value of d calculated for M-Ary PSK.

Minimum Shift Keying (MSK)

The wide spectrum of QPSK is due to the character of baseband signal. This signal consists of abrupt changes, and abrupt changes give rise to spectral components at high frequencies. The problem of interchannel interference in QPSK is so serious that regulatory and standardization energies such as FCC and CCIR will not permit these system will be used except with band pass filtering at carrier frequencies to suppress the side lobe. If we try to pass the baseband signal through a low pass filter to suppress the insignificant side lobes (the main lobe contains 90% of signal energy). Such filtering will cause ISI.

The QPSK is a system which the signal is of constant amplitude, the information content being borne by phase changes. In both QPSK and OQPSK are abrupt phase changes in the signal. In QPSK these changes can occur at the symbol rate $1/T_S = 1/2T_b$ and can be as large as 180^o. In OQPSK phase changes of 90^o can occur at the bit rate. Such abrupt phase changes cause many problems.

There are two difference between QPSK and MSK

1. In MSK the baseband waveform, that multiplies the quadrature carrier, is much smoother than the abrupt rectangular wave form of QPSK. While the spectrum of MSK has a main centre lobe while as 1-5 times as wide the main lobe of QPSK.
2. The wave form of MSK exhibits phase continuity that is there are no abrupt changes in QPSK. As a result we avoid the ISI caused by non-linear amplifier.

The staggering which is optimal in QPSK is essential in MSK. MSK transmitter needs two waveforms $\sin 2\pi(t/4T_b)$ and $\cos 2\pi(t/4T_b)$ to generate smooth baseband. The MSK transmitted signal is

$$V_{MSK}(\mathrm{t}) = \sqrt{2P_s}[b_e(\mathrm{t}).\sin 2\pi(\mathrm{t}/4\,\mathrm{T}_b)]\cos\mathrm{w}_0 t + \sqrt{2P_s}[b_o(\mathrm{t}).\cos 2\pi(\mathrm{t}/4\,\mathrm{T}_b)]\sin\mathrm{w}_0 t \quad (4.33)$$

suppose $2\pi/4T_b = \Omega$. then we can rewrite the above equation as

$$V_{MSK}(\mathrm{t}) = \sqrt{2P_s}[b_e(\mathrm{t}).\sin\Omega t]\cos w_0 t + \sqrt{2P_s}[b_o(\mathrm{t}).\cos\Omega t]\sin w_0 t \quad (4.34)$$

The above equation to be modified form of OQPSK, which we can call “shaped QPSK”. We can call apparent that MSK is an FSK system.

$$\begin{aligned} V_{MSK}(\mathrm{t}) = \sqrt{2P_s}[&{}^{b_e(\mathrm{t})}\!/_2\{\sin w_0 t.\cos\Omega t + \cos w_0 t.\sin\Omega t\} - {}^{b_e(\mathrm{t})}\!/_2\{\sin w_0 t.\cos\Omega t - \cos w_0 t.\sin\Omega t\} \\ &+ {}^{b_o(\mathrm{t})}\!/_2\{\sin w_0 t.\cos\Omega t + \cos w_0 t.\sin\Omega t\} + {}^{b_o(\mathrm{t})}\!/_2\{\sin w_0 t.\cos\Omega t - \cos w_0 t.\sin\Omega t\}] \\ = \sqrt{2P_s}&\left[\frac{b_o(\mathrm{t}) + b_e(\mathrm{t})}{2}\right].\sin(w_0 + \Omega)\,\mathrm{t} + \sqrt{2P_s}\left[\frac{b_o(\mathrm{t}) - b_e(\mathrm{t})}{2}\right].\sin(w_0 - \Omega)\,\mathrm{t} \end{aligned} \quad (4.35)$$

If we define $C_H = \frac{b_o + b_e}{2}, C_L = \frac{b_o - b_e}{2}, w_H = w_0 + \Omega$ & $w_L = w_0 - \Omega$ then the above equation can be written as,

$$V_{MSK}(\mathrm{t}) = \sqrt{2P_s}C_H(\mathrm{t}).\sin w_H t + \sqrt{2P_s}C_L(\mathrm{t}).\sin w_L t \quad (4.36)$$

Here $b_o = \pm 1$ and $b_e = \pm 1$, so it can be easily verified that, if $b_o = b_e =$ then $C_L = 0$ write $C_H = b_o = b_e = \pm 1$, Further if $b_o = b_e$, then $C_H = 0$ and $C_L = b_o = b_e = \pm 1$, Thus depending on the value of the bits w_H and w_L in each bit interval, the transmitted signal is at angular frequency ω_H or at ω_L precisely as in FSK and amplitude is always equal to $\sqrt{2P_s}$.

In MSK, the two frequencies f_H and f_L are chosen to ensure that the two possible signals are orthogonal over the bit interval T_b. That is, we impose the constraint that

$$\int_0^{T_b} \sin w_H t . \sin w_c t = 0 \tag{4.36a}$$

This is possible only when, $2\pi(f_H + f_L)T_b = m\pi$ and $2\pi(f_H - f_L)T_b = n\pi$, (4.37)

where m and n are integers. In equation (4.35)

$$f_H = f_o + \frac{\Omega}{2\pi} = f_o + \frac{f_b}{4}$$
$$f_L = f_o - \frac{\Omega}{2\pi} = f_o - \frac{f_b}{4}$$

$$f_b . T_b = 1$$

AS, $2\pi\left(f_H - f_L\right)T_b = n\pi$

$$\Rightarrow 2\pi_b . {}^{f_b}\!/_2 . T_b = n\pi$$

$$\Rightarrow n = 1 \tag{4.38}$$

Again,

$$2\pi\left(f_H + f_L\right)T_b = m\pi$$

$$\Rightarrow 2\pi_b . 2\,\mathrm{f}_0 . T_b = m\pi$$

$$\Rightarrow f_0 = \frac{m}{4} . f_b \tag{4.39}$$

Eq(38) shows that sincen=1, f_H and f_L are as close together as possible for orthogonality to prevail. It is for this reason that the present system is called "minium shift keying". Equation(4.39) shows that the carrier frequency f_0 is an integral multiple of $f_b/4$. Thus

$$f_H = (m+1) . {}^{f_b}\!/_4$$
$$f_L = (m-1) . {}^{f_b}\!/_4 \tag{4.40}$$

MSK Transmitter & Receiver

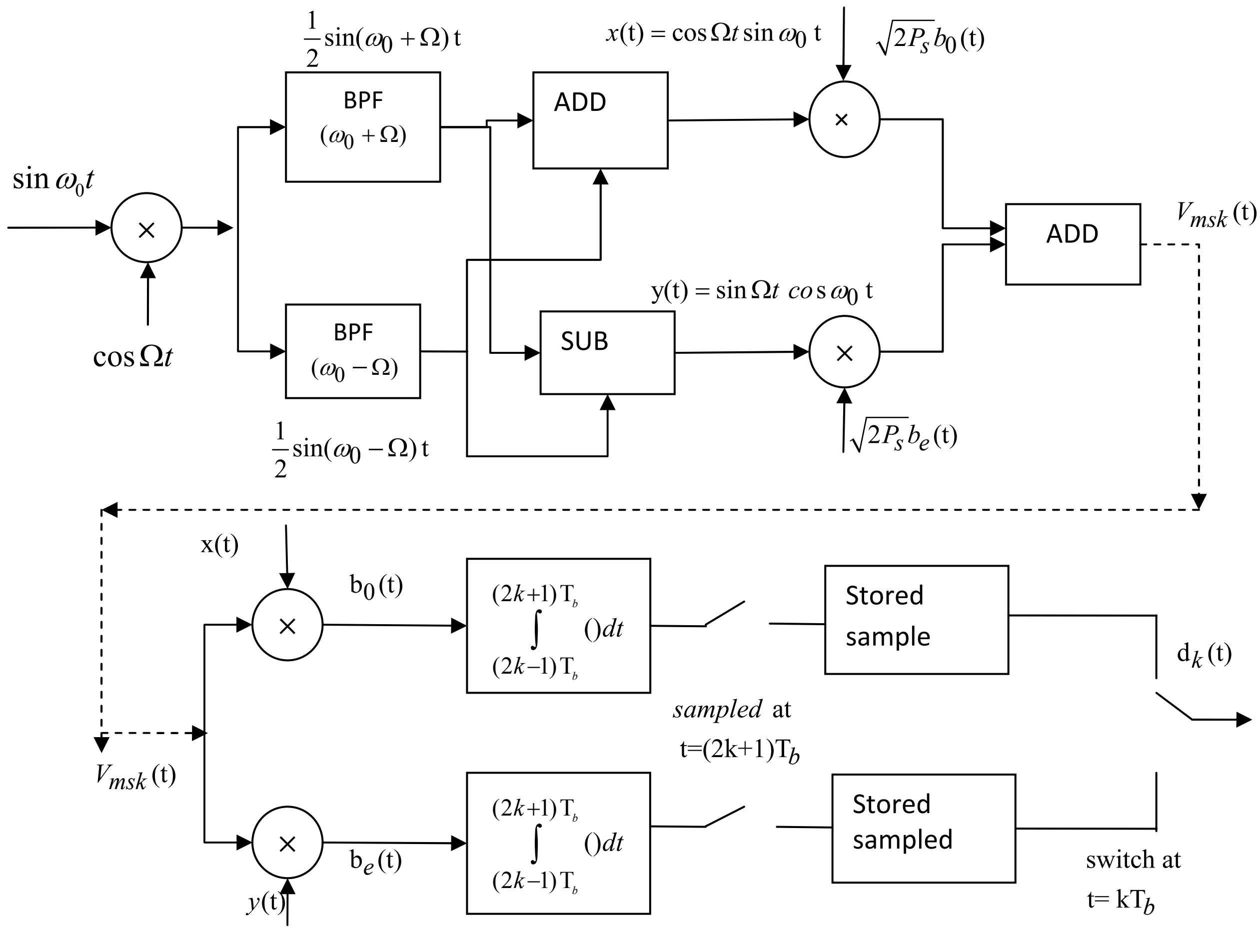


Fig 4.15 Transmission of MSK

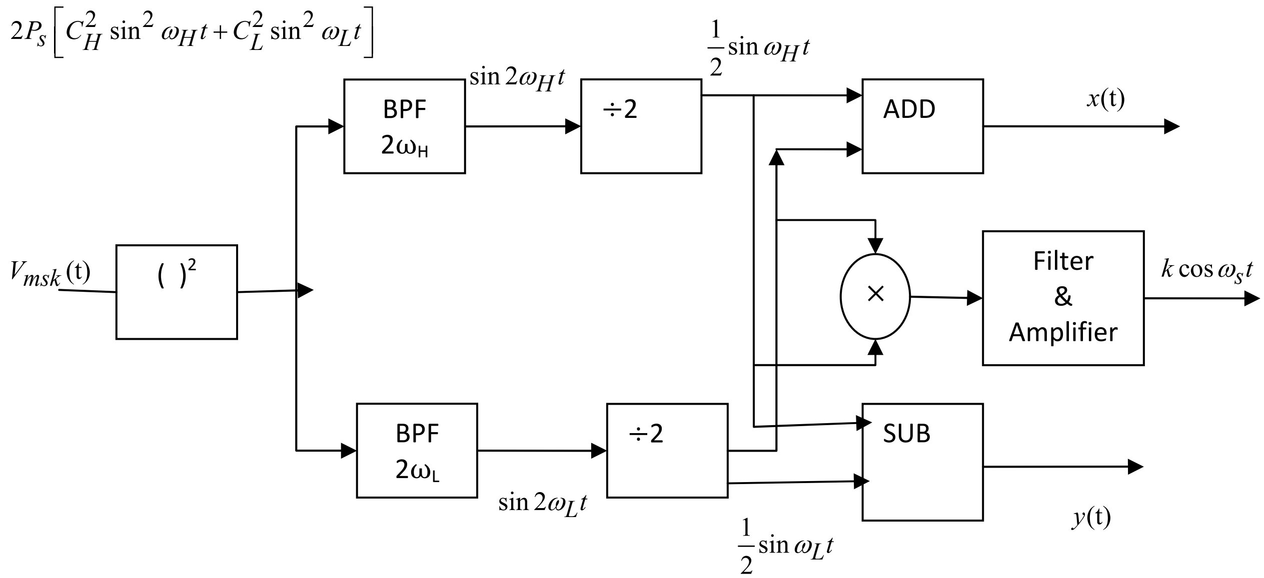


Fig 4.16 Reception of MSK

Spectrum of MSK

We see that the base band waveform which multiplies the sinω_0t in MSK is

$$\rho(t) = \sqrt{2p_s}\, b_0 \cos\frac{\pi}{2} f_b t \qquad -T_b \le t \le T_b \tag{4.41}$$

The waveform ρ(t) has a PSD $G_p(f) = \frac{32E_b}{\pi^2}\left[\frac{\cos 2\pi f / f_b}{1-(4f/f_b)^2}\right]^2$

$G_p(f)$ gives by

$$G_p(f) = \frac{32E_b}{\pi^2}\left[\frac{\cos 2\pi f / f_b}{1-(4f/f_b)^2}\right]^2 \tag{4.42}$$

Then the PSD for the total MSK signal of equation (4.33) is

$$G_{msk}(f) = \frac{8E_b}{\pi^2}\left[\left\{\frac{\cos 2\pi(f-f_0)/f_b}{1-[4(f-f_0)/f_b]^2}\right\}^2 + \left\{\frac{\cos 2\pi(f+f_0)/f_b}{1-[4(f+f_0)/f_b]^2}\right\}\right] \tag{4.43}$$

It is clear from the fig-4.9 that the main loab in MSK is wider than the main lobe in QPSK. In MSK the band width required to accommodate this lobe is $2*3/4f_b=1.5f_b$ while it is only $1f_b$ in QPSK. However in MSK the side lobe are very greatly suppressed in comparision to QPSK. in QPSK ,G(f) falls off as $1/f^2$ while in MSK G(f) falls off as $1/f^4$,It turns out that in MSK 99% of the signal power is contained in a band width of about $1.2f_b$. while in QPSK the corresponding bandwidth is about $8f_b$.

Geometrical representation of MSK in signal space

The signal space representation of MSK is shown in Fig 4.17a. The orthogonal unit vectors of the co-ordinate system are given by u_{ff}(t) and u_l(t). The end point of the four possible signal vectors are indicated by dots. The smallest distance between signal point is $d = \sqrt{2E_s} = 2\sqrt{E_b}$

QPSK generates two BPSK signal which are orthogonal to one another by virtue of the fact that the respective carriers are in phase quadrature. Such phase quadrature can also be charactarised as time quadrature since , at a carrier frequency to a phase shift of $\pi/2$ is accomplished by a time shift in amount $1/4f_0$i.e $\sin 2\pi f_0(t+1/4f_0) = \sin(2\pi f_0 t + \pi/2) = \cos 2\pi f_0 t$ It is of interest to note , in contrast, that in MSK we have again two BPSK signal [i.e the two individual terms in equation 4.36]

Here, however ,the respective carriers are orthogonal to one another by virtue of the fact that they are in frequency quadrature.

Phase continuity in MSK

A most important and useful feature of MSK in its phase continuity. This matter is illustrated in 4.17 b in waveform g, h ,and i. Here we have assumed $f_0=5f_b/4$ so that

$f_H= f_0+f_b/4= 5f_b/4 +f_b/4 =1.5f_b$ (4.44)

$f_L= f_0-f_b/4= 5f_b/4 -f_b/4 =1f_b$ (4.45)

Carriers of f_H and f_L are shown in g & h. We also find form eqn(4.35),that for the various combination of b_0 and b_e , $V_{msk}(t)/\sqrt{2P_s}$. It is clear that because of staging ,b_0 and b_e don't change simultaneously. The waveform $V_{msk}(t)$ is generated in the following way: in each bit interval we determine from eqn (4.36a), whether to use the carrier frequency f_H or f_l and also whether to use carrier waveform is to be inverted. Having made such a determination the waveform of $V_{msk}(t)$ is smooth and exhibits no abrupt changes in phase. Hence, in MSK we avoid the difficulty described above (pulse case),which results from the abrupt phase changes in the waveform of QPSK. We shall now see that the phase continuity and is a general characteristics of MSK. For this purpose we note from table 3 that the $V_{msk}(t)$

Waveform of eqn(4.35) or eqn (4.36) can be written as

$$V_{msk}(t) = b_0(t)\sqrt{2P_s}\sin\left[\omega_0 t + b_0(t)b_e(t)\Omega t\right] \quad (4.46)$$

The instantaneous phase $\phi(t)$ of the sinusoidal in eqn (4.46) is given by

$$\phi(t) = \omega_0 t + b_0(t)b_e(t)\Omega t \quad (4.47)$$

For convergence we represent the two phases as $\phi_+(t)$ or $\phi_-(t)$, where

$$\phi_+(t) = (\omega_0 + \Omega)t \quad ;b_o(t)b_e(t)=+1 \quad (4.48)$$

$$\phi_-(t) = (\omega_0 - \Omega)t \quad ;b_o(t)b_e(t)=-1 \quad (4.49)$$

$b_0(t)$ can take +-1and $b_e(t)$ can take +-1.The term $b_0(t)$,$b_e(t)$ in eqn(4.46) can change at times KT_b(k inis an integer).but they don't change at the same time .consider then ,first a change in $b_e(t)$.such a change will cause a phase change which is a multiple of 2π ,which is equivalent to n_o change at all ($b_e(t)$ can only change when k is even).when $b_0(t)$ changes the phase change in $\phi(t)$ will be an odd multiple of π i.e a phase change of π .but as per eqn (4.46) and its coefficient $b_0(t)$ which multiplies $\sqrt{2P_s}\sin\phi(t)$ .whenever there is a change in $b_0(t)$ to change the phase $\phi(t)$ by π ,the coefficient $b_0(t)$ will also change the sign of ,yielding an additional π phase change. Hence a change in $b_0(t)$ produces no net phase discontinuity.

Use of signal space to calculate probability of error for BPSK & BFSK

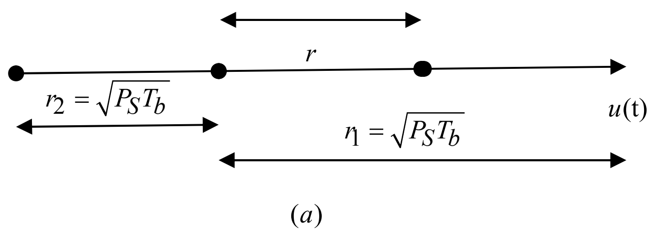
BPSK: in BPSK case,the signal space is one dimensional . The signal s1 & s2 are given by

$$\left.\begin{matrix} s_1(t) \\ s_2(t) \end{matrix}\right\} = \sqrt{2P_s}\, b(t)\cos\omega_0 t \;; \qquad 0<t \le T_b \tag{4.50}$$

Where b(t)=+1 for s_1 and b(t)=-1 for s_2. P_s is the signal power. If we introduce the unit (normalized)

Vector $u(t) = \sqrt{2/T_b}\cos\omega_0 t$,then

$$\left.\begin{matrix} s_1(t) \\ s_2(t) \end{matrix}\right\} = b(t)\sqrt{P_s T_b}\sqrt{\frac{2}{T_b}}\cos\omega_0 t \tag{4.51}$$



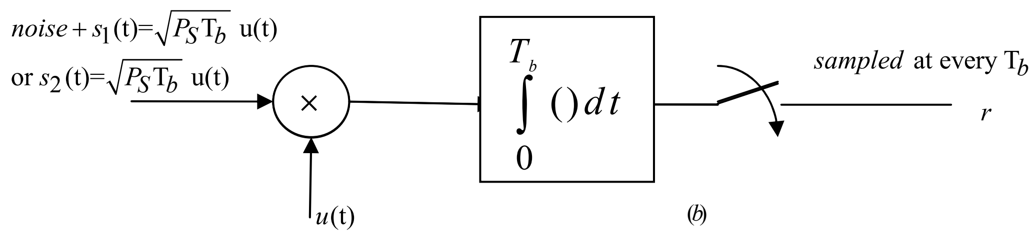


Fig 4.17 (a) Signal Vector (b) Co-relator Receiver

So signal vectors each of length $\sqrt{P_s T_b}$,measured in terms of unit vector u(t).processing at the correlator receiver, we will generate a response r_1 or r_2 for s_1 and s_2 respectively when no. noise is present. Now suppose that in some interval, because of noise a response r is generated.if we find $|r-r_1|<|r-r_2|$,then we determine that $s_1(t)$ was transmitted.

The relevant noise in BPSK case is

$$n(t) = n_0(t)u(t) = n_0\sqrt{\frac{2}{T_b}}\cos\omega_0 t \tag{4.52}$$

Where n_0 is a Gaussian random variable.

Variance of noise power $= \sigma_{0p} = \frac{n}{2}\frac{T_b}{\tau^2} = \frac{n}{2T_b} \qquad (\tau = Rc = T_b)$

Variance of noise energy $=\sigma_{0e} = \frac{n}{2T_b} T_b = \frac{n}{2}$ (4.53)

Let us take $S_2(t)$ was transmitted. The error probability ie the probability that the signal is mistaken or judged as $S_1(t)$.This is possible only when $n_0 > \sqrt{P_s T_b}$ .thus error probability P_e is given by

$$P_e = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\sqrt{P_s T_b}}^{\infty} e^{-n_0^2 / 2\sigma_0^2} dn_0 \quad (4.54)$$

$$P_e = \frac{1}{\sqrt{\pi\eta}} \int_{\sqrt{P_s T_b}}^{\infty} e^{-n_0^2 / \eta} dn_0$$

Let us assume $x^2 = \eta_0^2 / \eta$ then $dx = d\eta_0 / \sqrt{\eta}$ when $\eta_0 = \sqrt{P_s T_b}$ then $x = \sqrt{P_s T_s / \eta}$

$$\therefore P_e = \frac{1}{\sqrt{\pi}} \int_{\sqrt{P_s T_b / \eta}}^{\infty} e^{-x^2} dx \quad (4.55)$$

$$= \frac{1}{2} erfc(\sqrt{\frac{PsT_b}{\eta}}) = \frac{1}{2} erfc(\sqrt{\frac{E_b}{\eta}}) \quad (4.56)$$

As argument of erfc increases ,its value decreases .ie p_e decreases .

Thus error probability is seen to fall off monotonically with an increase in distance between signals.

BFSK

The unit vectors in BFSK considered are

$$\left.\begin{aligned} u_1(\mathrm{t}) &= \sqrt{2/T_b} \cos \omega_1 t \\ u_2(\mathrm{t}) &= \sqrt{2/T_b} \cos \omega_2 t \end{aligned}\right\} \quad (4.57)$$

ω_1 and ω_2 are selected in sucha manner that they are orthogonal over the interval T_b .The transmitted signal $s_1(\mathrm{t})$ and $s_2(t)$ are of power P_s are given by

$$S_1(\mathrm{t}) = \sqrt{2P_s} \cos \omega_1 t = \sqrt{P_s T_b} \cos \omega_1 t = \sqrt{P_s T_b} u_1(\mathrm{t}) \quad (4.58)$$

$$S_2(\mathrm{t}) = \sqrt{2P_s} \cos \omega_2 t = \sqrt{P_s T_b} \cos \omega_2 t = \sqrt{P_s T_b} u_2(\mathrm{t}) \quad (4.59)$$

In the absence of noise , when $s_1(t)$ is received, then $r_2=0$ and $r_2 = \sqrt{P_s T_b}$ .for$s_2(t)$ is received, then $r_1=0$ and $r_1 = \sqrt{P_s T_b}$ .The vectors representing r_1 and r_2 are of length $\sqrt{P_s T_b}$ .since the signal is two dimensional ,the relevant noise in the present case is

$n(t) = n_1(t)u_1(t) + n_2(t)u_2(t)$ (4.60)

Where n_1 and n_2 are Gaussian random variable each of variance $= \sigma_1^2 = \sigma_2^2 = \eta/2$.

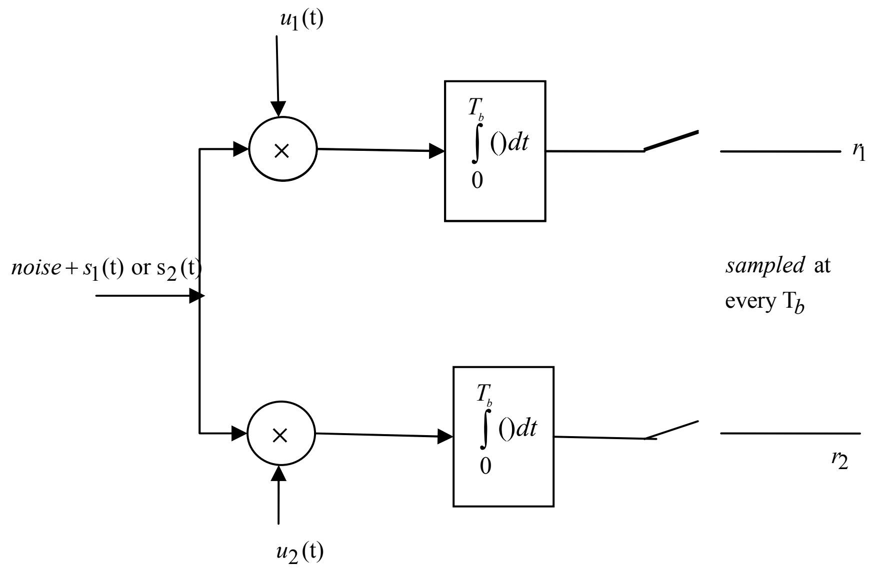


Fig 4.18 Reception in BFSK signal

Now let us suppose that $s_2(t)$ is transmitted and the observed voltage at the output of the receiver are r_1' and r_2' .we find r_2' not equal to r_2 because of the noise n_2 and $\eta_1' \neq 0$ because of noise then n_1 . we have locus of points equidistant from r_1 and r_2 suppose as shown that received voltage r is closer to r_1 to r_2.Then we shall have made an error in estimating which signal was transmitted.It is readily apparent that such an error will occur when ever noise $\eta_1 > r_2 - \eta_2$ or $(\eta_1 + \eta_2) > \sqrt{P_s T_b}$ .since n_1 and n_2 are uncorrelated ,random variable $n_0 = (n_1 + n_2)$ has a variance $\sigma_0^2 = \sigma_1^2 + \sigma_2^2 = n$ and its probability density function

$$f(n_0) = \frac{1}{\sqrt{2\pi\eta}} e^{-n_0^2/2n} \quad (4.61)$$

The probability error is

$$P_e = \frac{1}{\sqrt{2\pi\eta}} \int_{\sqrt{P_s T_b}}^{\infty} e^{-n_0^2/2\eta} dn_0 = \frac{1}{2} \times \frac{1}{\sqrt{\pi}} \int_{\sqrt{P_s T_b / 2\eta}}^{\infty} e^{-x} dx$$

$$= \frac{1}{2} erfc(\sqrt{\frac{PsT_b}{2\eta}}) = \frac{1}{2} erfc(\sqrt{\frac{E_b}{2\eta}}) \quad (4.62)$$

$$= \frac{1}{2} erfc(\sqrt{\frac{d}{4\eta}}) \quad (4.63)$$

For comparison of equation 4.55 & 4.62 should be used. Equation 4.56 & 4.63 are generalized equation.