Program : **B.Tech**

Subject Name: **Digital Communication**

Subject Code: **EC-502**

Semester: **5th**

Unit 4

Syllabus:

Other Digital Techniques: Pulse shaping to reduce inter channel and inter symbol interference- Duo binary encoding, Nyquist criterion and partial response signaling, Quadrature Partial Response (QPR) encoder decoder, Regenerative Repeater- eye pattern, equalizers

Optimum Reception of Digital Signals: Baseband signal receiver, probability of error, maximum likelihood detector, Bayes theorem, optimum receiver for both baseband and pass band receiver- matched filter and correlator, probability of error calculation for BPSK and BFSK.

4.1 Inter Symbol Interference

This is a form of distortion of a signal, in which one or more symbols interfere with subsequent signals, causing noise or delivering a poor output.

Causes of ISI

The main causes of ISI are −

- Multi-path Propagation
- Non-linear frequency in channels

The ISI is unwanted and should be completely eliminated to get a clean output. The causes of ISI should also be resolved in order to minimize its effect.

To view ISI in a mathematical form present in the receiver output, we can consider the receiver output.

The receiving filter output $y(t)$

is sampled at time $t_i=iT_b$

(with **i** taking on integer values), yielding −

$$y(t_i) = \mu \sum_{k=-\infty}^{\infty} a_k p(iT_b - kT_b)]$$

$$= \mu a_i + \mu \sum_{\substack{k=-\infty \\ k\neq i}}^{\infty} a_k p(iT_b - kT_b)]$$

In the above equation, the first term μa_i is produced by the i^{th} transmitted bit.

The second term represents the residual effect of all other transmitted bits on the decoding of the i^{th} bit. This residual effect is called as Inter Symbol Interference.

In the absence of ISI, the output will be −

$$y(t_i) = \mu a_i$$

This equation shows that the i^{th} bit transmitted is correctly reproduced. However, the presence of ISI introduces bit errors and distortions in the output.

While designing the transmitter or a receiver, it is important that you minimize the effects of ISI, so as to receive the output with the least possible error rate.

Correlative Coding

So far, we’ve discussed that ISI is an unwanted phenomenon and degrades the signal. But the same ISI if used in a controlled manner is possible to achieve a bit rate of 2W bits per second in a channel of bandwidth W Hertz. Such a scheme is called as Correlative Coding or Partial response signaling schemes.

Correlative Level Coding:

Correlative-level coding (partial response signaling) – adding ISI to the transmitted signal in a controlled manner Since ISI introduced into the transmitted signal is known, its effect can be interpreted at the receiver. A practical method of achieving the theoretical maximum signaling rate of 2W symbol per second in a bandwidth of W Hertz.

Using realizable and perturbation-tolerant filters

Since the amount of ISI is known, it is easy to design the receiver according to the requirement so as to avoid the effect of ISI on the signal. The basic idea of correlative coding is achieved by considering an example of Duo-binary Signaling.

4.2 Duo-binary Signaling

If f_M is the frequency of the maximum frequency spectral component of the baseband waveform, then, in AM, the bandwidth is B = $2f_M$. In frequency modulation, if the modulating waveform were a sinusoid of frequency f_M, and if the frequency deviation was Δf, then bandwidth would be

$$B = 2\Delta f + 2f_M \qquad \text{...4.2.1}$$

Altogether, it is apparent that bandwidth decreases with decreasing f_M regardless of the modulation technique employed. We consider now a mode of encoding a binary bit stream, called duobinary encoding which effects a reduction of the maximum frequency in comparison to the maximum frequency of the un-encoded data. Thus, if a carrier is amplitude or frequency modulated by a duobinary encoded waveform, the bandwidth of the modulated waveform will be smaller than if the un-encoded data were used to AM or FM modulate the carrier.

There are a number of methods available for duobinary encoding and decoding. One popular scheme is shown in Fig. 4.2.1. The signal d(k) is the data bit stream with bit duration T_b. It makes excursions between logic 1 and logic 0, and, as had been our custom, we take the corresponding voltage levels to be + 1V and -1V. The signal b(k), at the output of the differential encoder also makes excursions between + 1V and -1V. The waveform $v_d(k)$ is therefore

$$v_d(k) = b(k) + b(k-1) \qquad \text{...4.2.2}$$

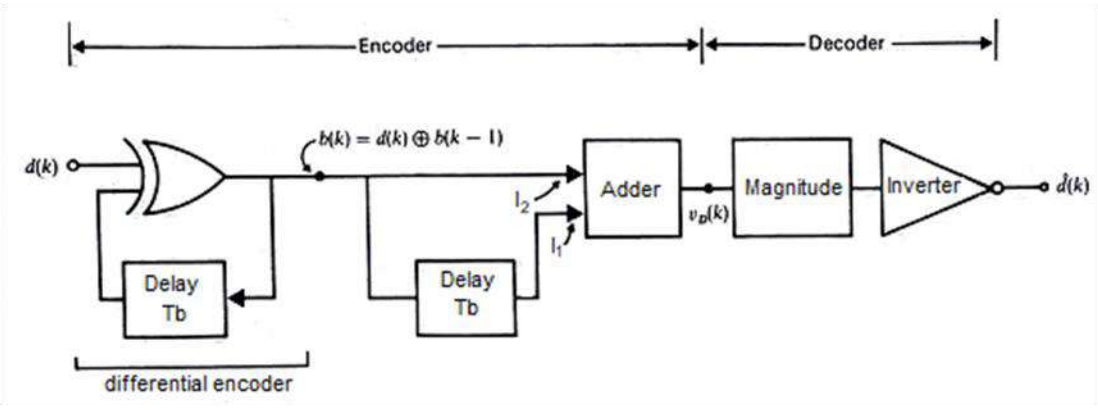

Figure 4.2.1 The Duobinary Encoder Decoder System

which can take on the values $v_D(k)$ = +2V, 0V and - 2V. The value of $v_D(k)$ in any interval k depends on both b(k) and b(k - 1). Hence there is a correlation between the values of $v_D(k)$ in any two successive intervals. For this reason the coding of Fig. 4.2.1 is referred to as correlative coding.

The correlation can be made apparent in another way. When the transition is made from one interval to the next, it is not possible for $v_D(k)$ to change from +2V to - 2V or vice versa. In short, in any interval, $v_D(k)$ cannot always assume any of the possible levels independently of its level in the previous intervals. Finally, we note that the term duobinary is appropriate since in each bit interval, the generated voltage $v_D(k)$ results from the combination of two bits.

The decoder, shown in Fig. 4.2.1, consists of a device that provides at its output the magnitude (absolute value) of its input cascaded with a logical inverter. For the inverter we take it that logic 1 is + 1V or greater and logic 0 is 0V. We can now verify that the decoded data $\hat{d}(\mathrm{k})$ is indeed the input data d(k). For this purpose we prepare the following truth table:

Truth Table For Duobinary Signaling

Adder Input 1 I_1		**Adder Input 2** I_2		**Adder output** $v_D(k)$	**Magnitude Output (Inverter output)**		**Inverter output** $d(k)$
Voltage	Logic	Voltage	Logic	Input Voltage	Voltage	Logic	Logic
-1V	0	-1V	0	-2V	2V	1	0
-1V	0	1V	1	0	0V	0	1
1V	1	-1V	0	0	0V	0	1
1V	1	1V	1	2V	2V	1	0

From the table we see that the inverter output is $I_1 \oplus I_2$.The differential encoder (called a precoder in the present application) output is:

$$I_1 = b(k) = d(k) \oplus b(k-1) \qquad \text{...4.2.3}$$

The input I_2 = b(k - 1) so that the inverter output $\hat{d}(\mathrm{k})$ is:

$$\hat{d}(\mathrm{k}) = I_1 \oplus I_2 = d(k) \oplus b(k-1) \oplus b(k-1) = d(k) \qquad \text{...4.2.4}$$

Waveforms of Duobinary Signaling

The more rapidly d(k) switches back and forth between logic levels the higher will be the frequencies of the spectral components generated. When d(k) switches at each time T_b, the switching speed is at a maximum. The waveform d(k), under such circumstances, has the appearance of a square wave of period $2T_b$ and frequency 1/2 T_b as shown in Fig. 4.2.2a.

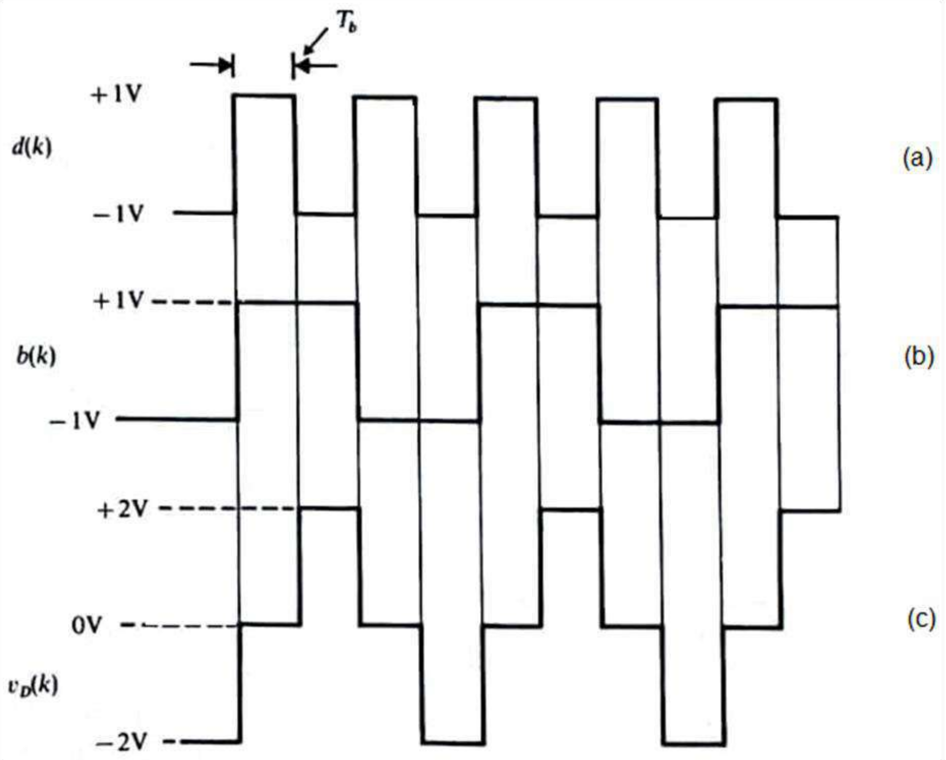

Figure 4.2.2 Waveforms of *d(k), b(k)* and $v_D(k)$

If d(k) is the input to the duobinary encoder of Fig. 4.2.1 then, as can be verified, b(k) appears as in Fig. 4.2.2b and the waveform, $v_D(k)$ which is to be transmitted appears as in Fig. 4.2.2c. Observe that the period of $v_D(k)$ is 4 T_b with corresponding frequency 1/4 T_b. Thus the frequency of $v_D(k)$ is half the frequency of the original unencoded waveform d(k). The waveform d(k) may be a sinusoid of frequency 1/2 T_b and waveform $v_D(k)$ as a sinusoid of frequency $1/4T_b$. If we were free to select either d(k) or $v_D(k)$ as a modulating waveform for a carrier, and if we were interested in conserving bandwidth, we would choose $v_D(k)$. If amplitude modulation were involved, the bandwidth of the modulated waveform would be $2(1/4T_b) = f_b/2$ using $v_D(k)$ since the modulating frequency is $f_M = 1/4T_b$ and would be $2(1/2T_b) = f_b$ using d(k). With frequency modulation, if the peak-to peak carrier frequency deviation were $2\Delta f$, then, the modulated carrier would have a bandwidth $2(\Delta f) + 2(1/2T_b)$ with d(k) as the modulating signal, as in BFSK; and $2(\Delta f) + 2(1/4T_b)$with $v_D(k)$ as the modulating signal.

4.3 Partial Response Signaling

Suppose, that corresponding to each bit of duration Tb, of a data stream we generate a positive impulse of strength +1 whenever the bit is at logic 1 and a negative impulse -1 whenever the bit is at logic 0. Suppose, further, that these impulses are applied to the input of the cosine filter. In Fig. 6.4.3 we have drawn the filter responses individually to five successive positive impulses. For simplicity, we have in each case drawn only the central lobe-and we have indicated by dots all the places where the individual response waveforms pass through zero. Where there is no dot, the waveforms has a finite value. The peaks of the responses are separated by times Tb and the widths of the central lobes are 3Tb.

The total response is, of course, simply the sum of the individual responses.

We can make the following observations from Fig. 4.3.1:

1. If we sample the total response at a time when an individual response is at its peak, the sample will have contributions from all the individual responses.
2. There is no possible time at which a sample of the total response is due only to a single individual response.
3. Importantly, if we sample the total response midway between times when the individual responses are at peak value, i.e., at $t = (2k-1)T_b/2$, then the sample value will have contributions in equal amount from only the two individual responses that straddle the sampling time. These sampling times are indicated in Fig. 4.3.1, by the light vertical lines. One such sampling time, yielding contributions from individual responses 2 and 3 is explicitly marked. It can be calculated that at the sampling time the contribution from each of the straddling individual responses will be a voltage If_b. Note that in sampling at the indicated times, we sample when the individual responses are not at peak value. For this reason, the present signal processing is referred to as Partial Response Signaling.

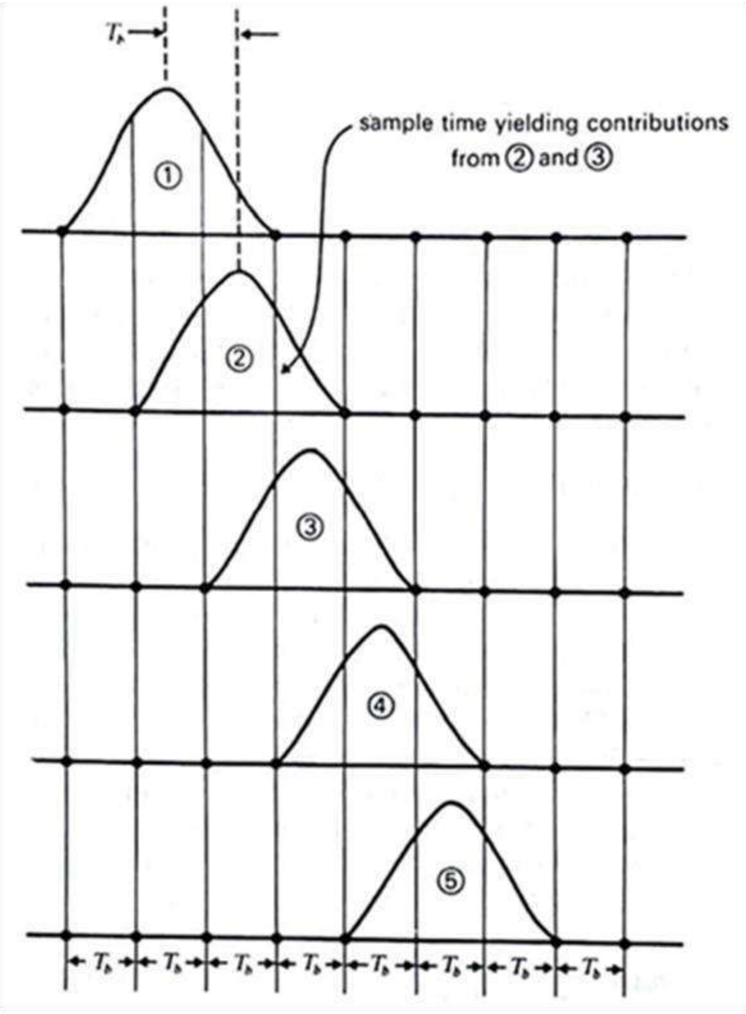

Figure 4.3.1 Filter responses to Five Different Impulses

In partial-response signaling, we shall transmit a signal during each bit interval that has contributions from two successive bits of an original baseband waveform. But this superposition need not prevent us from disentangling the individual original baseband waveform bits. A complete (baseband) partial-response signaling communications system is shown in Fig. 4.3.2.

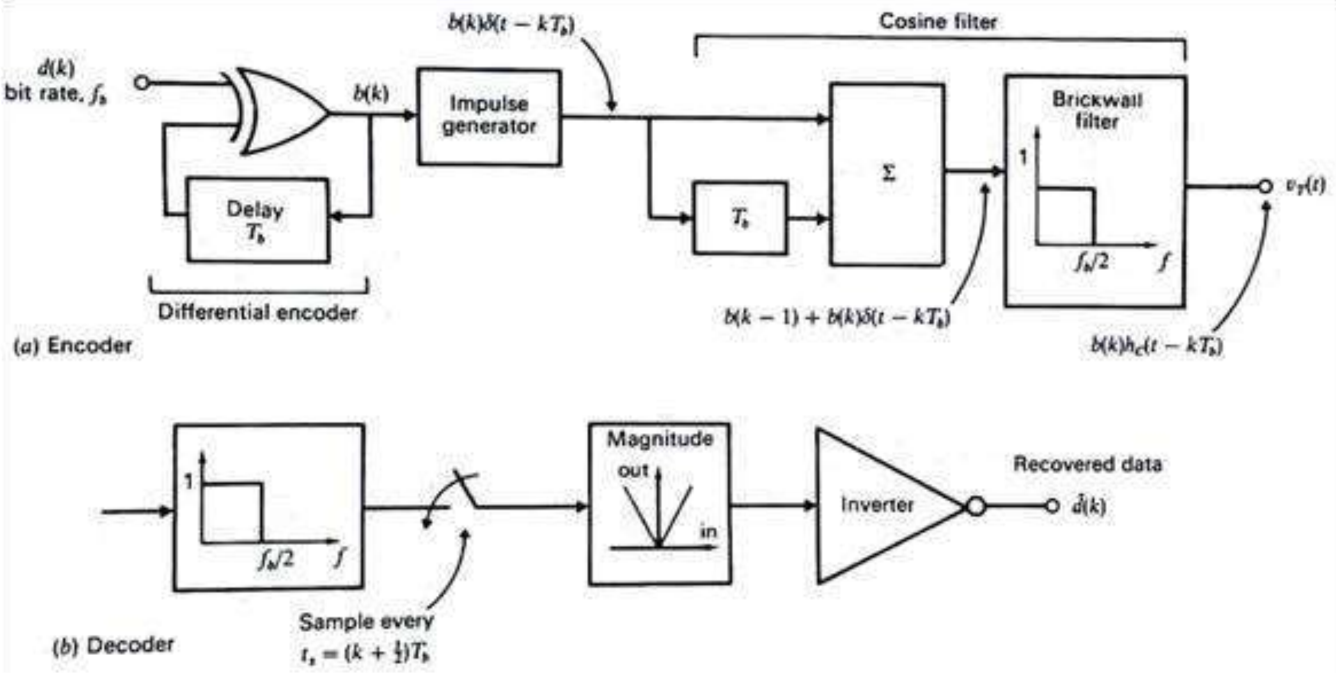

Figure 4.3.2 Duo Binary Encoder and Decoder Using Cosine filter

It is seen to be just an adaptation of duobinary encoding and decoding. The cosine filter employed a delay and an advance of the impulse by amount Tb/2, the total time between delayed and advanced impulses being Tb. Since, in the real world, a time advance is not possible, we have employed only a delay by amount Tb. The brickwall filter at the receiver input serves to remove any out of band noise added to the signal during transmission. It can be shown, that the output data $\hat{d}(\mathrm{k})$ = d(k).

4.4 Quadrature Partial Response (QPR) Encoder and Decoder

Amplitude Modulation of Partial Response Signal

The baseband partial response (duobinary) signal may be used to amplitude or frequency modulate a carrier. If amplitude modulation is employed, either double sideband suppressed carrier DSB/SC or quadrature amplitude modulation QAM can be employed.

For the case of DSB/SC the duobinary signal, $V_T(t)$, shown in Fig. 4.3.2a, is multiplied by the carrier $\sqrt{2}\cos\omega_0\,(t)$. The resulting signal is

$$v_{DSB}(t) = \sqrt{2}v_T(t)\cos\omega_0\,(t) \qquad ...4.4.1$$

The bandwidth required to transmit the signal is twice the bandwidth of the baseband duobinary signal which is $f_b/2$. Hence the bandwidth B_{DSB} of an amplitude modulated duobinary signal is

$$B_{DSB} = 2(f_b/2) = f_b \qquad ...4.4.2$$

If the duobinary signal is to amplitude modulate two carriers in quadrature, the circuit shown in Fig. 4.4.1 is used and the resulting encoder is called a "quadrature partial response" (QPR) encoder.

Figure 4.4.1 shows that the data d(t) at the bit rate fb is first separated into an even and an odd bit stream de(t) and do(t) each operating with the bit rate fb /2, Both de(t) and do(t) are then separately duobinary encoded into signals $V_{Te}(t)$ and $V_{To}(t)$.

Each duobinary encoder is similar to the encoder shown in Fig. 4.3.2a except that each delay is now 2Tb, rather than Tb, the data rate of the input is $f_b/2$ rather than f_b and the bandwidth of the brick wall filter is now $(1/2)(f_b/2)= f_b/4$ rather than $f_b/2$. Thus the bandwidth required to pass $V_{Te}(t)$ and $V_{To}(t)$ is $f_b/4$. Each duobinary signal is then modulated using the quadrature carrier signals cos ω_ot and sin ω_ot.

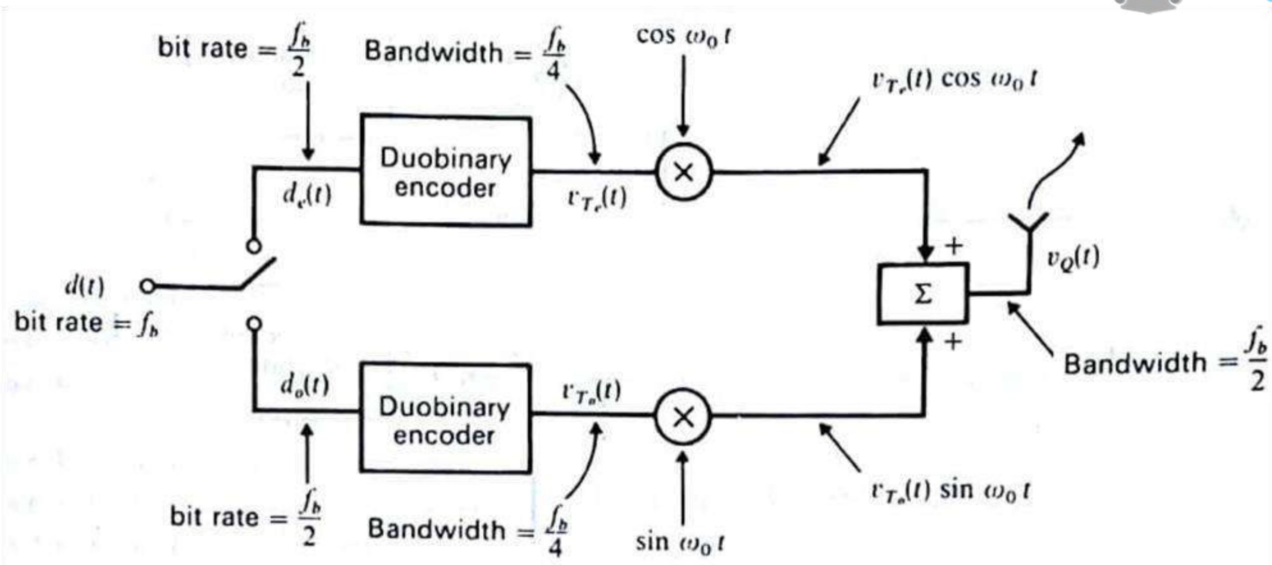

Figure 4.4.1 QPR Encoder

The bandwidth of each of the quadrature amplitude modulated signals is

$$B_{QPR} = 2(f_b/4) = f_b/2 \qquad \text{...4.4.3}$$

Hence the total bandwidth required to pass a QPR signal is also B_{QPR}, since the two quadrature components occupy the same frequency band.

It should be noted that if QPSK, rather than QPR, were used to encode the data d(t), the bandwidth required would be $B_{QPSK} = f_b$. However, if 16 QAM or 16 PSK were used to encode the data the required bandwidth would be $B_{16QAM} = B_{16PSK} = fb/2$. Thus the spectrum required to pass a QPR signal is similar to that required to pass 16 QAM or 16 PSK. However, the QPR signal displays no (or in practice very small) side lobes which makes QPR the encoding system of choice when spectrum width is the major problem. The drawback in using QPR is that the transmitted signal envelope is not a constant but varies with time.

QPR Decoder

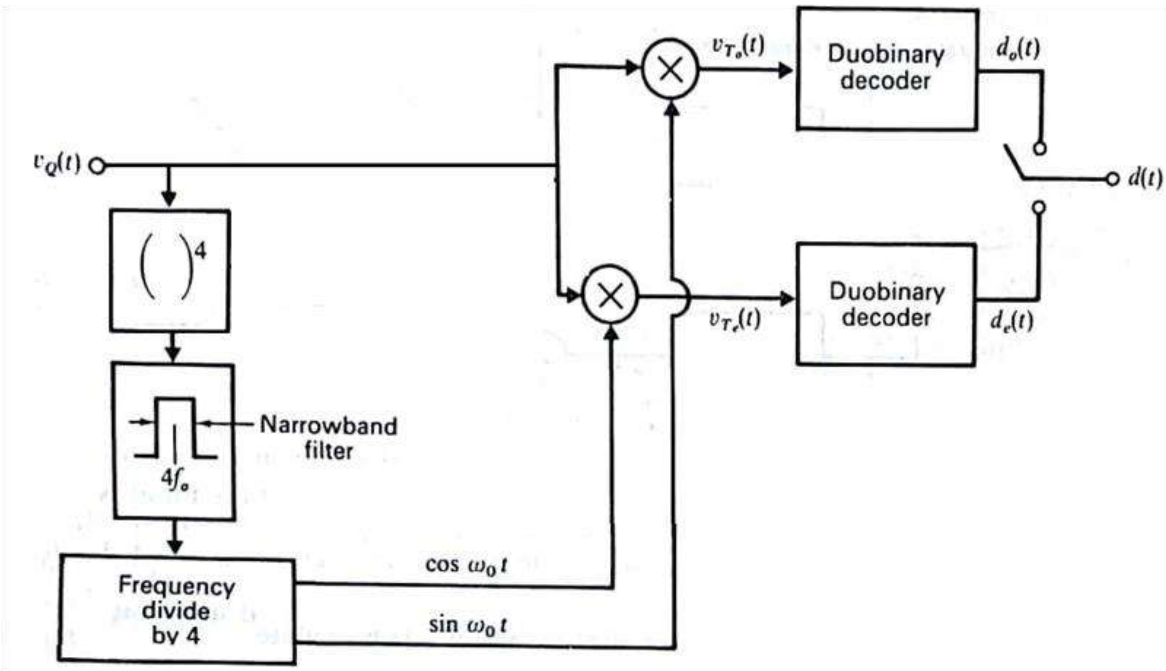

Figure 4.4.2 QPR Decoder

A QPR decoder is shown in Fig. 4.4.2. As in 16-QAM and 16-PSK to decode the input signal, $V_Q(t)$ is first raised to the fourth power, filtered and then frequency divided by 4. The result yields the two quadrature carriers: cos ω_ot and sin ω_ot. Using the two quadrature carriers we demodulate $V_Q(t)$ and obtain the two baseband duobinary signals $V_{Te}(t)$ and $V_{To}(t)$. Duobinary decoding then takes place; each duobinary decoder being similar to the decoder shown in Fig. 4.3.2b except that they operate at fb/2 rather than at fb. The reconstructed data do(t) and de(t) is then combined to yield the data d(t).

4.5 Eye Pattern

An effective way to study the effects of ISI is the Eye Pattern. The name Eye Pattern was given from its resemblance to the human eye for binary waves. The interior region of the eye pattern is called the eye opening. The following figure shows the image of an eye-pattern.

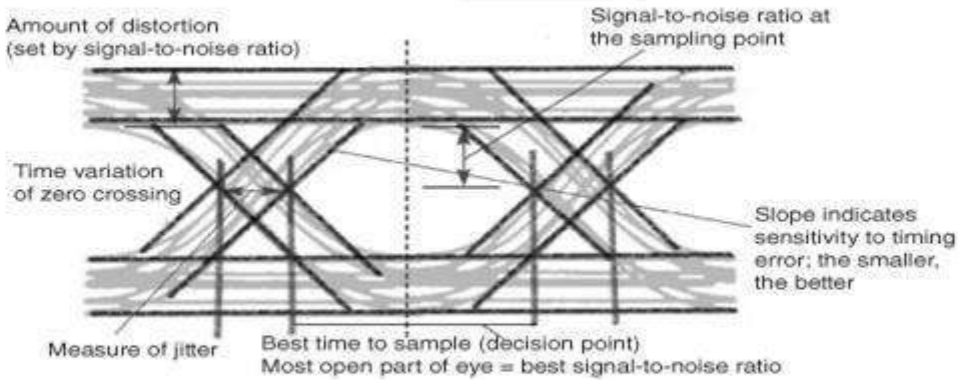

Figure 4.6 Image of eye pattern

Jitter is the short-term variation of the instant of digital signal, from its ideal position, which may lead to data errors.
When the effect of ISI increases, traces from the upper portion to the lower portion of the eye opening increases and the eye gets completely closed, if ISI is very high.
An eye pattern provides the following information about a particular system.

- Actual eye patterns are used to estimate the bit error rate and the signal-to-noise ratio.
- The width of the eye opening defines the time interval over which the received wave can be sampled without error from ISI.
- The instant of time when the eye opening is wide, will be the preferred time for sampling.
- The rate of the closure of the eye, according to the sampling time, determines how sensitive the system is to the timing error.
- The height of the eye opening, at a specified sampling time, defines the margin over noise.

Hence, the interpretation of eye pattern is an important consideration.

4.6 Equalization

For reliable communication to be established, we need to have a quality output. The transmission losses of the channel and other factors affecting the quality of the signal have to be treated. The most occurring loss, as we have discussed, is the ISI.
To make the signal free from ISI, and to ensure a maximum signal to noise ratio, we need to implement a method called Equalization. The following figure shows an equalizer in the receiver portion of the communication system.

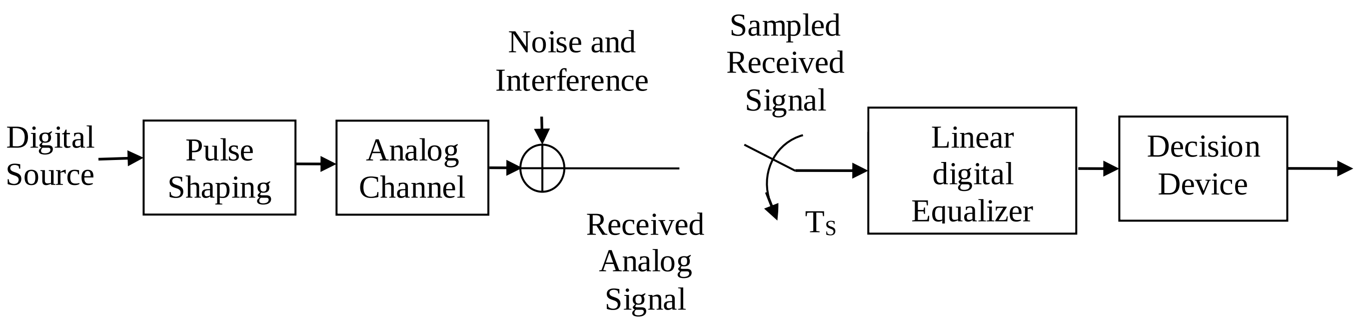

Figure 4.7 Equalization

The noise and interferences which are denoted in the figure are likely to occur, during transmission. The regenerative repeater has an equalizer circuit, which compensates the transmission losses by shaping the circuit. The Equalizer is feasible to get implemented.

Error Probability and Figure-of-merit

The rate at which data can be communicated is called the data rate. The rate at which error occurs in the bits, while transmitting data is called the Bit Error Rate (BER).

The probability of the occurrence of BER is the Error Probability. The increase in Signal to Noise Ratio (SNR) decreases the BER, hence the Error Probability also gets decreased.

In an Analog receiver, the figure of merit at the detection process can be termed as the ratio of output SNR to the input SNR. A greater value of figure-of-merit will be an advantage.

Regenerative Repeater

For any communication system to be reliable, it should transmit and receive the signals effectively, without any loss. A PCM wave, after transmitting through a channel, gets distorted due to the noise introduced by the channel.

The regenerative pulse compared with the original and received pulse, will be as shown in the following figure.

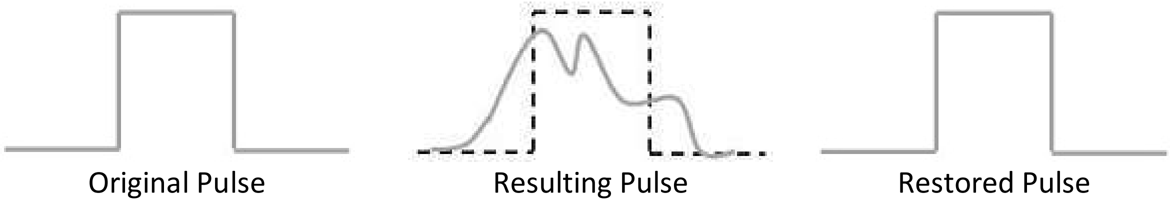

Figure 4.8 Regenerative Repeater

For a better reproduction of the signal, a circuit called as regenerative repeater is employed in the path before the receiver. This helps in restoring the signals from the losses occurred. Following is the diagrammatical representation.

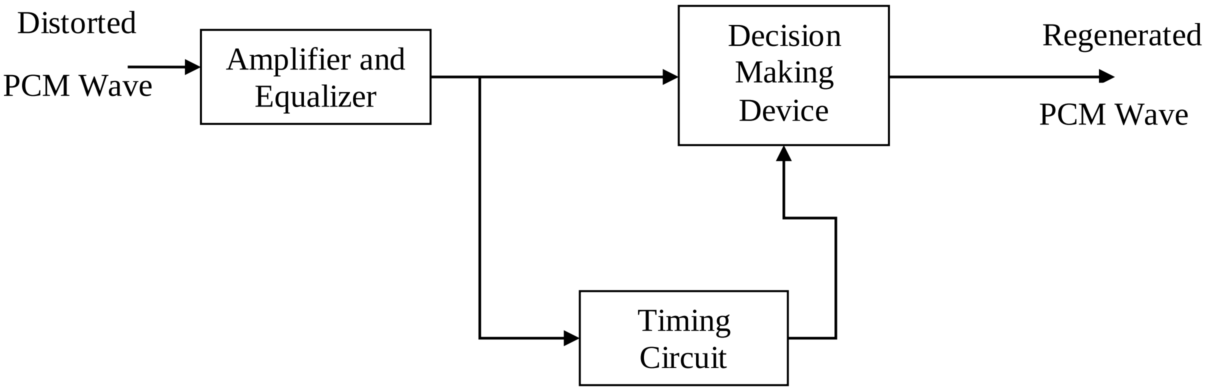

Figure 4.9 Block Diagram of Regenerative Repeater

This consists of an equalizer along with an amplifier, a timing circuit, and a decision making device. Their working of each of the components is detailed as follows.

Equalizer

The channel produces amplitude and phase distortions to the signals. This is due to the transmission characteristics of the channel. The Equalizer circuit compensates these losses by shaping the received pulses.

Timing Circuit

To obtain a quality output, the sampling of the pulses should be done where the signal to noise ratio (SNR) is maximum. To achieve this perfect sampling, a periodic pulse train has to be derived from the received pulses, which is done by the timing circuit.
Hence, the timing circuit allots the timing interval for sampling at high SNR, through the received pulses.

Decision Device

The timing circuit determines the sampling times. The decision device is enabled at these sampling times. The decision device decides its output based on whether the amplitude of the quantized pulse and the noise, exceeds a pre-determined value or not.

4.7 Baseband Signal Receiver:

Consider that a binary-encoded signal consists of a time sequence of voltage levels +V or-V. If there is a guard interval between the bits, the signal forms a sequence of positive and negative pulses. In either case there is no particular interest in preserving the waveform of the signal after reception. We are interested only in knowing within each bit interval whether the transmitted voltage was + V or - V. With noise present, the received signal and noise together will yield sample values generally different from ± V. In this case, what deduction shall we make from the sample value concerning the transmitted bit?
Suppose that the noise is Gaussian and therefore the noise voltage has a probability density which is entirely symmetrical with respect to zero volts. Then the probability that the noise has increased the sample value is the same as the probability that the noise has decreased the sample value. It then seems entirely reasonable that we can do no better than to assume that if the sample value is positive the transmitted level was + V, and if the sample value is negative the transmitted level was - V. It is, of course, possible that at the sampling time the noise voltage may be of magnitude larger than V and of a polarity opposite to the polarity assigned to the transmitted bit. In this case an error will be made as indicated in Fig. 4.7.1. Here the transmitted bit is represented by the voltage + V which is sustained over an interval T from t_1 to t_2. Noise has been superimposed on the level + V so that the voltage *v* represents the received signal and noise. If now the sampling should happen to take place at a time $t=t_1+\Delta t$; an error will have been made.

We can reduce the probability of error by processing the received signal plus noise in such a manner that we are then able to find a sample time where the sample voltage due to the signal is emphasized relative to the sample voltage due to the noise. Such a processer (receiver) is shown in Fig. 4.7.2. The signal input during a bit interval is indicated. As a matter of convenience we have set t = q at the beginning of the interval. The waveform of the signal s(t) before t =0 and after t = T has not been indicated since, as will appear, the operation of the receiver during each bit interval is independent of the waveform during past and future bit intervals.

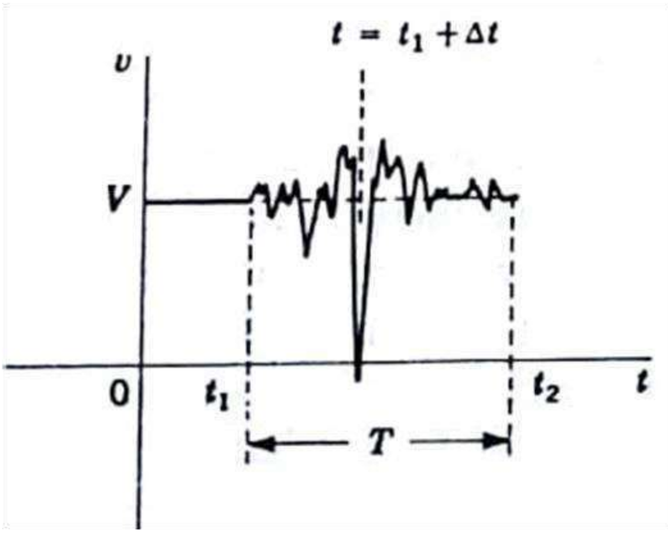

Figure 4.7.1 Illustration that noise may cause an error in determination of transmitted voltage level

The signal s(t) with added white gaussian noise n(t) of power spectral density η/2 is presented to an integrator. At time t = 0 + we require that capacitor C be uncharged. Such a discharged condition may be ensured by a brief closing of switch SW_1 at time t = 0-, thus relieving C of any charge it may have acquired

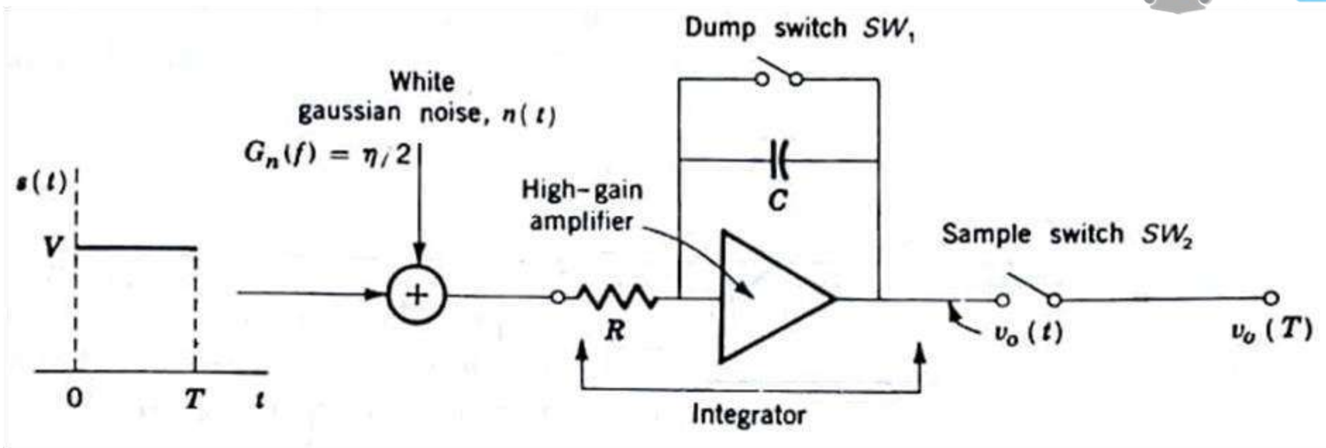

Figure 4.7.2 A Receiver for Binary Coded Signal

during the 'previous interval. The sample is taken at the output of the integrator by closing this sampling switch SW_2. This sample is taken at the end of the bit interval, at t = T. The signal processing indicated in Fig. 4.7.2 is described by the phrase integrate and dump, the term dump referring to the abrupt discharge of the capacitor after each sampling.

Peak Signal to RMS Noise Output Voltage Ratio

The integrator yields an output which is the integral of its input multiplied by 1/RC. Using τ = RC, we have

$$v_o(T) = \frac{1}{\tau}\int_0^T [s(t) + n(t)]dt = \frac{1}{\tau}\int_0^T s(t)dt + \frac{1}{\tau}\int_0^T n(t)dt \qquad \text{...4.7.1}$$

The sample voltage due to the signal is

$$s_o(T) = \frac{1}{\tau}\int_0^T V dt = \frac{VT}{\tau} \qquad \text{...4.7.2}$$

The sample voltage due to the noise is

$$s_n(T) = \frac{1}{\tau}\int_0^T n(t)dt \qquad \text{...4.7.3}$$

This noise-sampling voltage $n_o(T)$ is a Gaussian random variable in contrast with n(t) which is a Gaussian random process.

The variance of $n_o(T)$ is given by

$$\sigma_0^2 = \overline{n_0^2(t)} = \frac{nT}{2\tau^2} \qquad \text{...4.7.4}$$

It has a Gaussian probability density.

The output, of the integrator, before the sampling switch, is $v_0(t) = S_0(t) + n_0(t)$. As shown in Fig. 4.7.3a, the signal output So(t) is a ramp, in each bit interval, of duration T. At the end of the interval the ramp attains the voltage $S_0(t)$ which is + VT/τ or - VT/τ, depending on whether the bit is a 1 or a 0. At the end of each interval the switch SW_1 in Fig. 4.7.2 closes momentarily to discharge the capacitor so that so(t) drops to zero. The noise $n_0(t)$ shown in Fig. 4.7.3b, also starts each interval with no(0) = 0 and has the random value $n_0(t)$ at the end of each interval. The sampling switch SW_2 closes briefly just before the closing of SW_1 and hence reads the voltage

$$v_o(T) = s_o(T) + n_o(T) \qquad \text{...4.7.5}$$

We would naturally like the output signal voltage to be as large as possible in comparison with the noise voltage. Hence a figure of merit of interest is the signal-to-noise ratio

$$\frac{[s_o(T)]^2}{\overline{[n_o(T)]^2}} = \frac{2}{\eta}V^2T \qquad \text{...4.7.6}$$

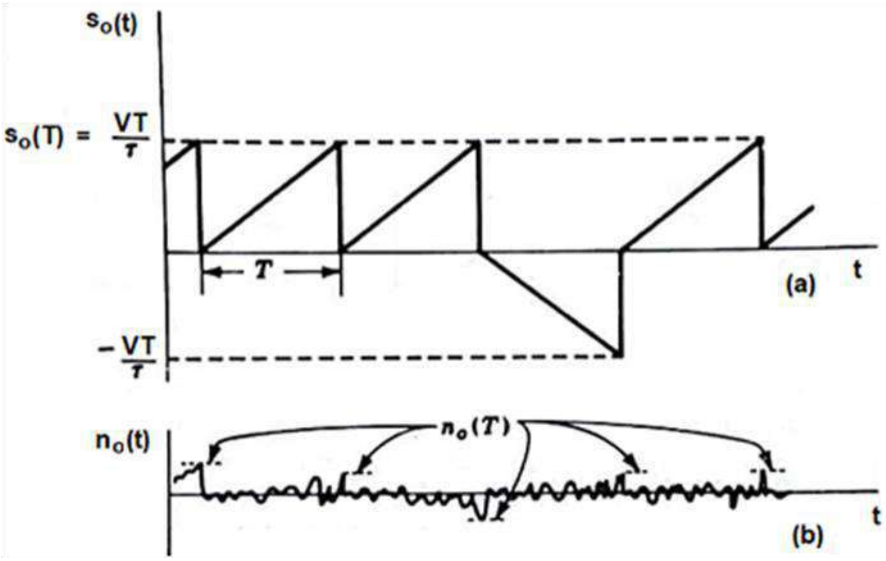

Figure 4.7.3 (a) The Signal Output and (b) the Noise Output of the integrator

This result is calculated from Eqs. (4.7.2) and (4.7.4). Note that the signal-to noise ratio increases with increasing bit duration T and that it depends on V^2T which is the normalized energy of the bit signal. Therefore, a bit represented by a narrow, high amplitude signal and one by a wide, low amplitude signal are equally effective, provided V^2T is kept constant. It is instructive to note that the integrator filters the signal and the noise such that the signal voltage increases linearly with time, while the standard deviation (rms value) of the noise increases more slowly, as $\sqrt{T}$. Thus, the integrator enhances the signal relative to the noise, and this enhancement increases with time as shown in Eq. (4.7.6).

4.8 Probability of Error:

Since the function of a receiver of a data transmission is to distinguish the bit 1 from the bit 0 in the presence of noise, a most important characteristic is the probability that an error will be made in such a determination. We now calculate this error probability P, for the integrate-and-dump receiver of Fig. 4.7.2. We have seen .that the probability density of the noise sample *no(T)* is Gaussian and hence appears as in Fig. 4.7.1.The density is therefore given by

$$f[n_o(T)] = \frac{e^{-n_0^2(T)/2\sigma_0^2}}{\sqrt{2\pi\sigma_0^2}} \qquad \text{...4.8.1}$$

Where σ_0^2 the variance is given by $\sigma_0^2 = \overline{\sigma_0^2(t)}$. Suppose, then, that during some bit interval the input-signal voltage is held at, say, - V. Then, at the sample time, the signal sample voltage is So(T)= -VT/τ, while the noise sample is no(T). If no(T) is positive and larger in magnitude than VT/ τ, the total sample voltage vo(T) = so(T) + no(T) will be positive. Such a positive sample voltage will result in an error, since as noted earlier, we have instructed the receiver to interpret such a positive sample voltage to mean that the signal voltage was + V during the bit interval. The probability of such a misinterpretation, that is, the probability that no(T) > VT/ τ, is given by the area of the shaded region in Fig. 4.8.1. The probability of error is, using Eq. (4.8.1).

$$P_e = \int_{VT/\tau}^{\infty} f[n_o(T)]\, dn_o(T) = \int_{VT/\tau}^{\infty} \frac{e^{-n_0^2(T)/2\sigma_0^2}}{\sqrt{2\pi\sigma_0^2}} dn_o(T) \qquad \text{...4.8.2}$$

Defining $x \equiv \frac{n_o(T)}{\sqrt{2}\sigma_0}$ and using equations 4.7.4 equation 4.8.2 may be written as

$$P_e = \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{x=VT/\tau}^{\infty} e^{-x^2}\, dx = \frac{1}{2} erfc\left(V\sqrt{\frac{T}{\eta}}\right) = \frac{1}{2} erfc\left(\frac{VT}{\eta}\right)^{1/2} \frac{1}{2} erfc\left(\frac{E_S}{\eta}\right)^{1/2} \qquad \text{...4.8.3}$$

In which $E_s=V^2T$ is the signal energy of a bit.

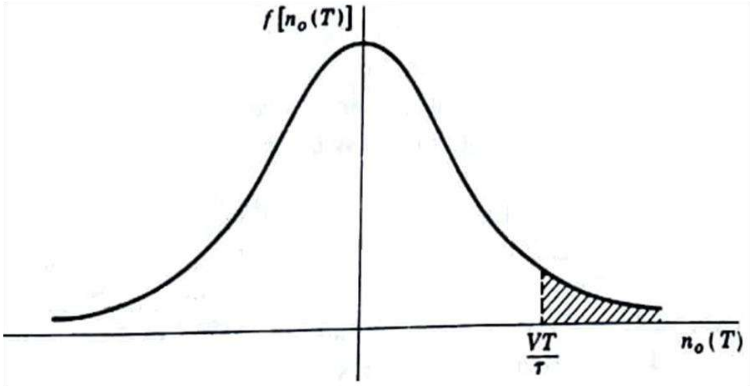

Figure 4.8.1 The Gaussian Probability Density of the noise sample $n_0(T)$

If the signal voltage were held instead at + V during some bit interval, then it is clear from the symmetry of the situation that the probability of error would again be given by P, in Eq. (4.8.3). Hence Eq. (4.8.3) gives P, quite generally.

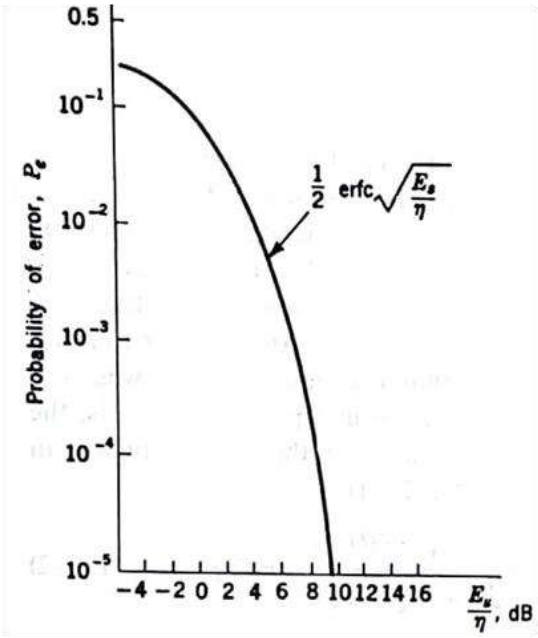

Figure 4.8.2 Variation of P_e versus Es/η

The probability of error Pe as given in Eq. (4.8.3), is plotted in Fig. 4.8.2. Note that Pe decreases rapidly as Es/η increases. The maximum value of Pe is 1/2. Thus, even if the signal is entirely lost in the noise so that any determination of the receiver is a sheer guess, the receiver cannot be wrong more than half the time on the average.

4.9 The Optimum Receiver

In the receiver system of Fig. 4.7.2, the signal was passed through a filter (i.e. the integrator), so that at the sampling time the signal voltage might be emphasized in comparison with the noise voltage. We are naturally led to ask whether the integrator is the optimum filter for the purpose of minimizing the probability of error. We shall find that for the received signal contemplated in the system of Fig. 4.7.2 the integrator is indeed the optimum filter.

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. 4.7.2, $S_1(t) = +V$, while $S_2(t) = -V$; for other modulation systems, different waveforms are transmitted. For example, for PSK signalling, $S_1(t) = A\cos\omega_0 t$ and $S_2(t) = -A\cos\omega_0 t$; while for FSK, $S_1(t) = A\cos(\omega_0+\Omega)t$ and $S_2(t) = A\cos(\omega_0-\Omega)t$.

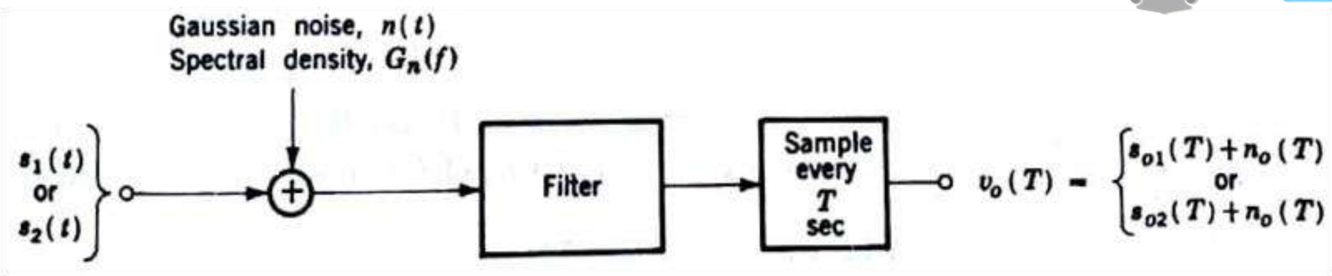

Figure 4.9.1 A Receiver for binary coded signaling

As shown in Fig. 4.9.1 the input, which is $S_1(t)$ or $S_2(t)$, is corrupted by the addition of noise n(t). The noise is gaussian and has a spectral density G(f). [In most cases of interest the noise is white, so that G(J) = η/2. However, we shall assume the more general possibility, since it introduces no complication to do so.] The signal and noise are filtered and then sampled at the end of each bit interval. The output sample is either vo(T) = S_{01}(T) + n_0(T) or v_o(T) = $S_0$2(T) + n_o(T). We assume that immediately after each sample, every energy-storing element in the filter has been discharged.

We note that in the absence of noise the output sample would be v_o(T) = S_{o1}(T) or S_{02}(T). When noise is present we have shown that to minimize the probability of error one should assume that S_1(t) has been transmitted if v_0(T) is closer to S_{01}(T) than to S_{02}(T). Similarly, we assume S_2(t) has been transmitted if v_o(T) is closer to S_{02}(T). The decision boundary is therefore midway between S_{o1}(T) and S_{02}(T). For example, in the baseband system of Fig. 4.7.2, where S_{01}(T) = VT/τ and S_{02}(T) = - VT/τ, the decision boundary is v_o(T) = 0. In general. we shall take the decision boundary to be

$$v_o(T) = \frac{s_{o1}(T) + s_{o2}(T)}{2} \qquad \text{...4.9.1}$$

The probability of error for this general case may be deduced as an extension of the considerations used in the baseband case. Suppose that S_{01}(T) > S_{02}(T) and that S_2(t) was transmitted. If, at the sampling time the noise n_o(T) is positive and larger in magnitude than the voltage difference (1/2)[S_{01}(T)+ S_{02}(T)] - S_{02}(T), an error will have' been made. That is, an error will result if

$$n_o(T) \geq \frac{s_{o1}(T) - s_{o2}(T)}{2} \qquad \text{...4.9.2}$$

Hence the probability of error is

$$P_e = \int_{[s_{o1}(T)-s_{o2}(T)]/2}^{\infty} \frac{e^{-n_0^2(T)/2\sigma_0^2}}{\sqrt{2\pi\sigma_0^2}} dn_o(T) \qquad \text{...4.9.3}$$

If we make the substitution $x \equiv \frac{n_o(T)}{\sqrt{2}\sigma_0}$, then above equation becomes,

$$P_e = \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{[s_{o1}(T)-s_{o2}(T)]/2\sqrt{2}\sigma_0}^{\infty} e^{-x^2}\, dx \qquad \text{...4.9.4a}$$

$$P_e = \frac{1}{2} erfc\left[\frac{s_{o1}(T) - s_{o2}(T)}{2\sqrt{2}\sigma_0}\right] \qquad \text{...4.9.4b}$$

Note that for the case S_{01}(T) = VT/τ and S_{02}(T) = - VT/τ, and, using Eq. (4.7.4), Eq. (4.9.4b) reduces to Eq. (4.8.3) as expected.

The complementary error function is a monotonically decreasing function of its argument. (See Fig. 4.8.2.) Hence, as is to be anticipated, Pe decreases as the difference S_{01}(T) - S_{02}(T) becomes larger and as the rms noise voltage σ_0 becomes smaller. The optimum filter, then, is the filter which maximizes the ratio

$$\gamma = \frac{s_{o1}(T) - s_{o2}(T)}{\sigma_0} \qquad \text{...4.9.5}$$

We now calculate the transfer function H(f) of this optimum filter.

Optimum Filter Transfer Function *H(f)*

The fundamental requirement we make of a binary encoded data receiver is that it distinguishes the voltages $S_1(t) + n(t)$ and $S_2(t) + n(t)$. We have seen that the ability of the receiver to do so depends on how large a particular receiver can make γ. It is important to note that it is proportional not to $S_1(t)$ nor to $S_2(t)$, but rather to the difference between them. For example, in the baseband system we represented the signals by voltage levels + V and - V. But clearly, if our only interest was in distinguishing levels, we would do just as well to use +2 volts and 0 volt, or +8 volts and +6 volts, etc, (The + V and - V levels, however, have the advantage of requiring the least average power to be transmitted.) Hence, while $S_1(t)$ or $S_2(t)$ is the received signal, the signal which is to be compared with the noise, i.e., the signal which is relevant in all our error-probability calculations, is the difference signal

$$p(T) \equiv s_1(T) - s_2(T) \qquad \text{...4.9.6}$$

Thus, for the purpose of calculating the minimum error probability, we shall assume that the Input Signal to the optimum filter is p(t). The corresponding output signal of the filter is then

$$p_0(T) \equiv s_{01}(T) - s_{02}(T) \qquad \text{...4.9.7}$$

Let *P(f)* and *P₀(f)* ne the Fourier transform of *p(f)* and *P₀(f)* respectively. If H(f) is the transfer function of the filter

$$P_0(f) = H(f)P(f) \qquad \text{...4.9.8}$$

And

$$p_0(T) = \int_{-\infty}^{\infty} P_0(f)\, e^{j2\pi fT}\, df = \int_{-\infty}^{\infty} H(f)P(f)\, e^{j2\pi fT}\, df \qquad \text{...4.9.9}$$

The input noise to the optimum filler is n(t). The output noise is $n_0(t)$ which has a power spectral density $Gn_0(f)$ and is related to the power spectral density of the input noise Gn(f) by

$$G_{n0}(f) = |H(f)|^2 G_n(f) \qquad \text{...4.9.10}$$

Using Parseval's theorem, we find that the normalized output noise power, i.e., the noise varianceσ_0^2, is

$$\sigma_0^2 = \int_{-\infty}^{\infty} G_{n0}(f)\, df = \int_{-\infty}^{\infty} |H(f)|^2 G_n(f)\, df \qquad \text{...4.9.11}$$

From equation 9 and 11,

$$\gamma^2 = \frac{p_0^2(T)}{\sigma_0^2} = \frac{\left[\int_{-\infty}^{\infty} H(f)P(f)\, e^{j2\pi fT}\, df\right]^2}{\int_{-\infty}^{\infty} |H(f)|^2 G_n(f)\, df} \qquad \text{...4.9.12}$$

Equation 4.9.12 is unaltered by the inclusion or deletion of the absolute value sign in the numerator since the quantity within the magnitude sign $p_0(T)$ is a positive real number. The sign has been included, however, in order to allow further development of the equation through the use of the Schwarz inequality.

The Schwarz inequality state that given arbitrary complex functions X(f) and Y(f) of a common variable f, then

$$\left|\int_{-\infty}^{\infty} X(f)Y(f)\, df\right|^2 \le \int_{-\infty}^{\infty} |X(f)|^2\, df \int_{-\infty}^{\infty} |Y(f)|^2\, df \qquad \text{...4.9.13}$$

The equal sign applies when

$$X(f) = KY^*(f) \qquad \text{...4.9.14}$$

where K is an arbitrary constant and Y*(f) is the complex conjugate of Y(f).

We now apply the Schwarz inequality to Eq. (4.9.12) by making the identification

$$X(f) \equiv \sqrt{G_n(f)}\, H(f) \qquad \text{...4.9.15}$$

and

$$Y(f) \equiv \frac{1}{\sqrt{G_n(f)}}\, P(f) e^{j2\pi Tf} \qquad \text{...4.9.16}$$

Using equation 15 and 16 and using Schwarz inequality equation 13, we may write equation 12 as,

$$\frac{p_0^2(T)}{\sigma_0^2} = \frac{\left[\int_{-\infty}^{\infty} X(f)Y(f)\,df\right]^2}{\int_{-\infty}^{\infty} |H(f)|^2\,df} \le \int_{-\infty}^{\infty} |Y(f)|^2\,df \qquad ...4.9.17$$

Using equation 16,

$$\frac{p_0^2(T)}{\sigma_0^2} \le \int_{-\infty}^{\infty} |Y(f)|^2\,df = \int_{-\infty}^{\infty} \frac{[P(f)]^2}{G_n(f)}\,df \qquad ...4.9.18$$

The ratio $p_0^2(T)/\sigma_0^2$; will attain its maximum value when the equal sign in Eq. (4.9.18) may be employed as is the case when X(f) = K Y*(f). We then find from Eqs. (4.9.15) and (4.9.16) that the optimum filter which yields such a maximum ratio $p_0^2(T)/\sigma_0^2$; has a transfer function

$$H(f) = K\frac{P^*(f)}{G_n(f)}e^{-j2\pi fT} \qquad ...4.9.19$$

Correspondingly, the maximum ratio is, from Eq. (4.9.18),

$$\left[\frac{p_0^2(T)}{\sigma_0^2}\right]_{max} = \int_{-\infty}^{\infty} \frac{[P(f)]^2}{G_n(f)}\,df \qquad ...4.9.20$$

4.10 White Noise: The Matched Filter

An optimum filter which yields a maximum ratio $p_0^2(T)/\sigma_0^2$is called a matched filler when the input noise is white. In this case $G_n(f) = \eta/2$, and Eq. (4.9.19) becomes

$$H(f) = K\frac{P^*(f)}{\eta/2}e^{-j2\pi fT} \qquad ...4.10.1$$

The impulsive response of this filter, i.e., the response of the filter to a unit strength impulse applied at t = 0, is

$$h(t) = F^{-1}[H(f)] = \frac{2K}{\eta}\int_{-\infty}^{\infty} P^*(f)e^{-j2\pi fT}\,e^{-j2\pi ft}\,df \qquad ...4.10.2(a)$$

$$= \frac{2K}{\eta}\int_{-\infty}^{\infty} P^*(f)e^{j2\pi f(t-T)}\,df \qquad ...4.10.2(b)$$

A physically realizable filter will have an impulse response which is real, i.e., not complex. Therefore h(t) = h*(t). Replacing the right-hand member of Eq. (4.10.2b) by its complex conjugate, an operation which leaves the equation unaltered, we have

$$h(t) = \frac{2K}{\eta}\int_{-\infty}^{\infty} P(f)e^{j2\pi f(t-T)}\,df \qquad ...4.10.3(a)$$

$$= \frac{2K}{\eta}p(T-t) \qquad ...4.10.3(b)$$

Finally since $p(t)=s_1(t) - s_2(t)$, we have

$$h(t) = \frac{2K}{\eta}[s_1(T-t) - s_2(T-t)] \qquad ...4.10.4$$

As shown in Fig. 4.10.1a, the $s_1(t)$ is a triangular waveform of duration T, while $s_2(t)$, (Fig. 4.10.1b), is of identical form except of reversed polarity. Then p(t) is as shown in Fig. 4.10.1c, and p(-t) appears in Fig. 4.10.1d. The waveform p(-t) is the waveform p(t) rotated around the axis t =0. Finally, the waveform p(T - t) called for as the impulse response of the filter in Eq. (4.10.3b) is this rotated waveform p(-t) translated in the positive t direction by amount T. This last translation ensures that h(t) = 0 for t < 0 as is required for a causal filter.

In general, the impulsive response of the matched filter consists of p(t) rotated about t = 0 and then delayed long enough (i.e., a time T) to make the filter realizable. We may note in passing, that any additional delay that a filter might introduce would in no way interfere with the performance of the filter, for both signal and noise would be delayed by the same amount, and at the sampling the ratio of signal to noise would remain unaltered.

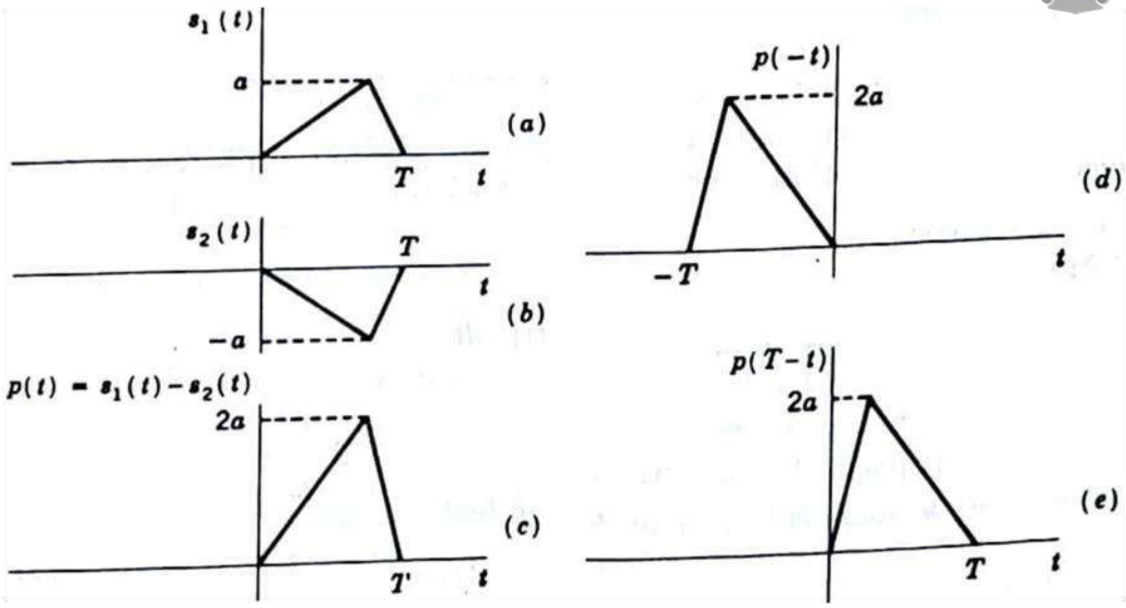

Figure 4.10.1 The signals (a) $s_1(t)$, (b) $s_2(t)$, (c) $p(t)=s_1(t)- s_2(t)$, (d) p(t) rotated about the axis t=0, (e) The waveform of (d) translated to right by amount T.

4.11 Correlator

Coherent Detection: Correlation

Coherent detection is an alternative type of receiving system, which is identical in performance with the matched filter receiver. Again, as shown in Fig. 4.11.1, the input is a binary data waveform $S_1(t)$ or $S_2(t)$ corrupted by noise n(t). The bit length is T. The received signal plus noise $v_i(t)$ is multiplied by a locally generated waveform $S_1(t)$ - $S_2(t)$. The output of the multiplier is passed through an integrator whose output is sampled at t = T. As before, immediately after each sampling, at the beginning of each new bit interval, all energy-storing elements in the integrator are discharged. This type of receiver is called a correlator, since we are correlating the received signal and noise with the waveform $S_1(t)$- $S_2(t)$.

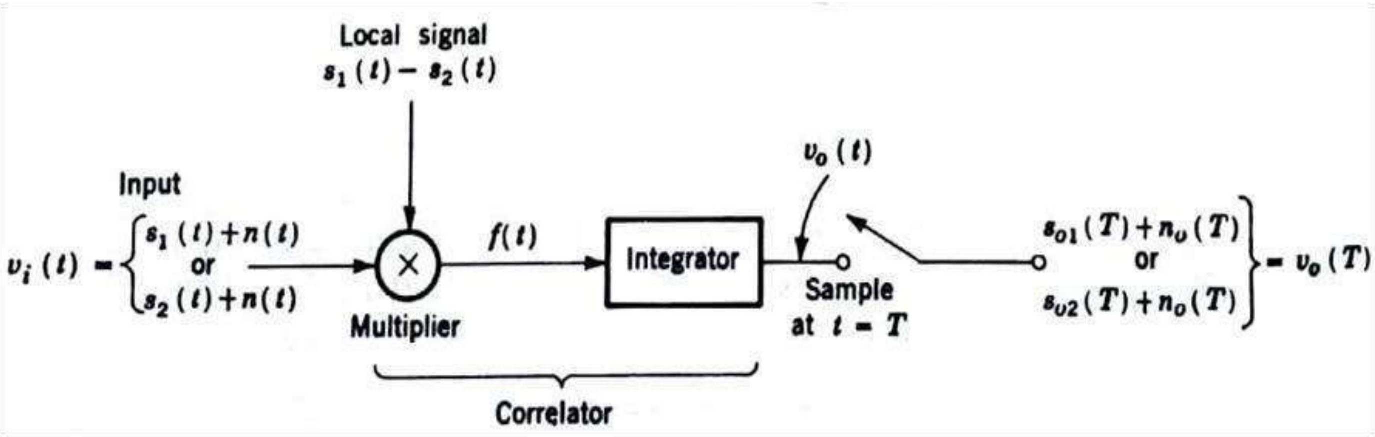

Figure 4.11.1 A Coherent System of Signal Reception

The output signal and noise of the correlator shown in Fig. 4.11.1 are

$$s_0(t) = \frac{1}{\tau}\int_0^T S_i(t)[S_1(t) - S_2(t)]\,dt \qquad \text{...4.11.1}$$

$$n_0(t) = \frac{1}{\tau}\int_0^T n(t)[S_1(t) - S_2(t)]\,dt \qquad \text{...4.11.2}$$

where $S_i(t)$ is either $S_1(t)$ or $S_2(t)$, and where τ is the constant of the integrator (i.e., the integrator output is l/τ times the integral of its input). We now compare these outputs with the matched filter outputs.

If h(t) is the impulsive response of the matched filter, then the output of the matched filter $v_o(t)$ can be found using the convolution integral. We have

$$v_0(t) = \int_{-\infty}^{\infty} v_i(\lambda)h(t-\lambda)\,d\lambda = \int_0^T v_i(\lambda)h(t-\lambda)\,d\lambda \qquad ...4.11.3$$

The limits on the integral have been changed to 0 and T since we are interested in the filter response to a bit which extends only over that interval. Using Eq. (4.10.4) which gives h(t) for the matched filter, we have

$$h(t) = \frac{2K}{\eta}[s_1(T-t) - s_2(T-t)] \qquad ...4.11.4$$

So that

$$h(t-\lambda) = \frac{2K}{\eta}[s_1(T-t+\lambda) - s_2(T-t+\lambda)] \qquad ...4.11.5$$

Submitting equation 4.11.5, in equation 4.11.3

$$v_0(t) = \frac{2K}{\eta}\int_0^T v_i(\lambda)[s_1(T-t+\lambda) - s_2(T-t+\lambda)]\,d\lambda \qquad ...4.11.6$$

Since $v_i(\lambda) = s_i(\lambda)+n(\lambda)$, and $v_0(t) = s_0(t)+n_0(t)$, setting t=T yields,

$$s_0(t) = \frac{2K}{\eta}\int_0^T s_i(\lambda)[s_1(\lambda) - s_2(\lambda)]d\lambda \qquad ...4.11.7$$

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

$$n_0(t) = \frac{2K}{\eta}\int_0^T n(\lambda)[s_1(\lambda) - s_2(\lambda)]d\lambda \qquad ...4.11.8$$

Thus as we can see from above equations $s_o(T)$ and $n_o(T)$, are identical. Hence the performances of the two systems are identical.

The matched filter and the correlator are not simply two distinct, independent techniques which happen to yield the same result. In fact they are two techniques of synthesizing the optimum filter h(t).

4.12 Probability of error calculation for BPSK and BFSK

(i) BPSK

The synchronous detector for BPSK is shown in figure 4.12.1(b). Since the BPSK signal is one dimensional, The only relevant noise in the present case is

$$n(t) = n_0 u(t) = n_0\sqrt{2/T_b}\cos\omega_0 t \qquad ...4.12.1$$

where n_o is a Gaussian random variable of variance $\sigma_0^2 = \eta/2$. Now let us suppose that S_2 was transmitted.

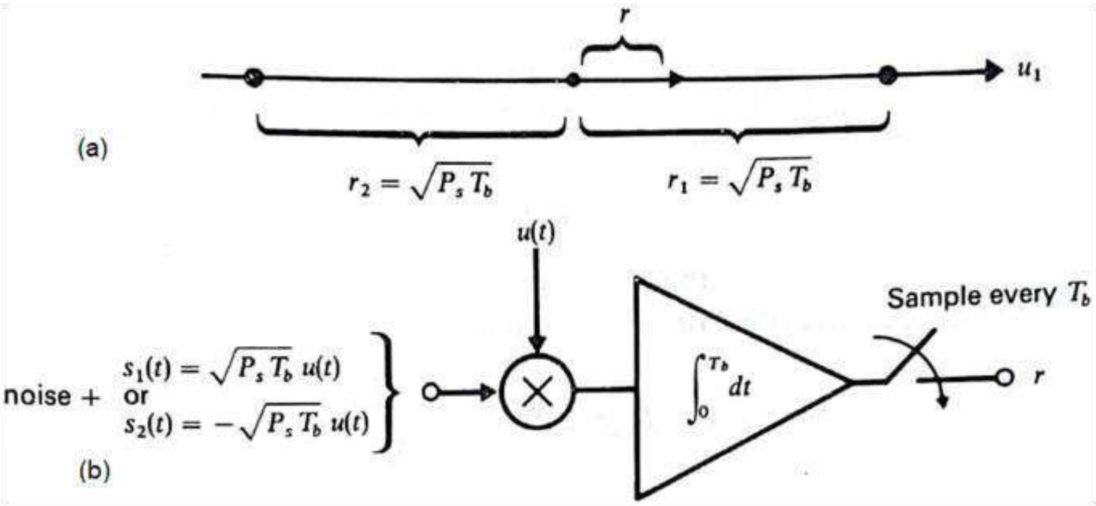

Figure 4.12.1 (a) BPSK representation in signal space showing r_1 and r_2 (b)Correlator receiver for BPSK showing that $r=r_1+n_0$ or r_2+n_0

The error probability, i.e., the probability that the signal is mistakenly judged to be S_1 is the probability that $n_0 > \sqrt{P_s T_b}$. Thus the error probability P_e, is

$$P_e = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\sqrt{P_s T_b}}^{\infty} e^{-n_0^2/2\sigma_0^2}\,dn_0 = \frac{1}{\sqrt{\pi\eta}}\int_{\sqrt{P_s T_b}}^{\infty} e^{-n_0^2/\eta}\,dn_0 \qquad ...4.12.2$$

Let $y^2 = n_0^2/2\sigma_0^2$, then

$$P_e = \frac{1}{\sqrt{\pi}}\int_{\sqrt{P_sT_b/\eta}}^{\infty} e^{-y^2}\,dy = \frac{1}{2}\frac{2}{\sqrt{\pi}}\int_{\sqrt{P_sT_b/\eta}}^{\infty} e^{-y^2}\,dy = \frac{1}{2}erfc\,\sqrt{P_sT_b/\eta} \qquad ...4.12.3$$

The signal energy is $E_b = P_sT_b$ and the distance between end points of the signal vectors in Fig. 4.12.1 is = $2\sqrt{P_sT_b}$. Accordingly we find that

$$P_e = \frac{1}{2}erfc\,\sqrt{E_b/\eta} = \frac{1}{2}erfc\,\sqrt{d^2/4\eta} \qquad ...4.12.4$$

The error probability is thus seen to fall off monotonically with an increase in distance between signals.

(ii) BFSK

The case of synchronous detection of orthogonal binary FSK is represented in Fig. 4.12.2. The signal space is shown in (a). The unit vectors are

$$u_1(t) = \sqrt{2/T_b}\cos\omega_1 t \qquad ...4.12.5a$$

and

$$u_2(t) = \sqrt{2/T_b}\cos\omega_2 t \qquad ...4.12.5b$$

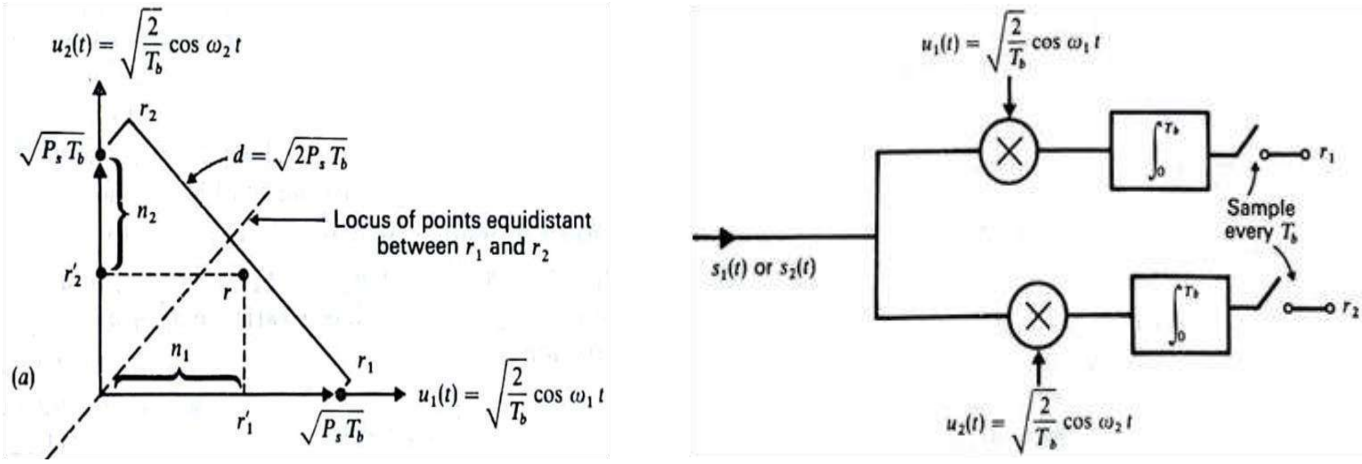

Figure 4.12.2 (a) Signal Space representation of BFSK (b) Correlator Receiver for BFSK

Orthogonality over the interval T_b having been insured by the selection of ω_1 and ω_2. The transmitted signals s_1 and s_2 are of power P_s, and are

$$s_1(t) = \sqrt{2P_s}\cos\omega_1 t = \sqrt{P_sT_b}\sqrt{2/T_b}\cos\omega_1 t = \sqrt{P_sT_b}\,u_1(t) \qquad ...4.12.6a$$

and

$$s_2(t) = \sqrt{2P_s}\cos\omega_2 t = \sqrt{P_sT_b}\sqrt{2/T_b}\cos\omega_2 t = \sqrt{P_sT_b}\,u_2(t) \qquad ...4.12.6b$$

Detection is accomplished in the manner shown in Fig. 4.12.2 (b). The outputs are r_1 and r_2. In the absence of noise when $s_1(t)$ is received, $r_2 = 0$ and $r_1 = \sqrt{P_sT_b}$. For $S_2(t)$, $r_1 = 0$ and $r_2 = \sqrt{P_sT_b}$. Hence the vectors representing r_1 and r_2 are of length $\sqrt{P_sT_b}$ as shown in Fig. 4.12.2(a).

Since the signal is two dimensional the relevant noise in the present case is

$$n(t) = n_1u_1(t) + n_2u_2(t) \qquad ...4.12.7$$

In which n_1 and n_2 are Gaussian random variables each of variance $\sigma_1^2 = \sigma_2^2 = \eta/2$. Now let us suppose that $S_2(t)$ is transmitted and that the observed voltages at the output of the processor are r'_1 and r'_2 as shown in Fig. 4.12.2a. We find that $r'_2 \neq r_2$ because of the noise n_2 and $r'_1 \neq 0$ because of the noise n_1. We have drawn the locus of points equidistant from r_1 and r_2 and suppose, that the received voltage r, is closer to r_1 than 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 whenever $n_1 > r_2 - n_2$ or $n_1 + n_2 > \sqrt{P_sT_b}$. Since n_1 and n_2 are uncorrelated, the random variable $n_0 = n_1 + n_2$ has a variance $\sigma_0^2 = \sigma_1^2 + \sigma_2^2 = \eta$ and its probability density function is

$$f(n_0) = \frac{1}{\sqrt{2\pi n}}e^{-n_0^2/2\eta} \qquad ...4.12.8$$

The probability of error is

$$P_e = \frac{1}{\sqrt{2\pi\eta}} \int_{\sqrt{P_s T_b}}^{\infty} e^{-n_0^2/2\eta}\, dn_0 \qquad \text{...4.12.9}$$

Again we have $E_b = P_s T_b$ and in the present case the distance between r_1 and r_2 is $d = \sqrt{2}\sqrt{P_s T_b}$. Accordingly, proceeding as in Eq. (4.12.2) we find that

$$P_e = \frac{1}{2} erfc\sqrt{E_b/2\eta} \qquad \text{...4.12.10a}$$

$$= \frac{1}{2} erfc\sqrt{d^2/2\eta} \qquad \text{...4.12.10b}$$

Comparing Eqs. (4.12.10b) and (4.12.4) we see that when expressed in terms of the distance d, the error probabilities are the same for BPSK and BFSK.

----------X----------