



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

School of Electrical and Electronics
Department of Electronics and Communication

SEC1313- Digital Communication

UNIT I – SAMPLING AND QUANTIZATION - SEC1313

SAMPLING AND QUANTIZATION

Introduction:

In Communication systems we studied the term Modulation, which is helpful in transmitting the signals at higher frequencies rather than at baseband frequencies. The following are the outputs of Amplitude modulation (AM), Frequency modulation (FM) and Phase modulation (PM),

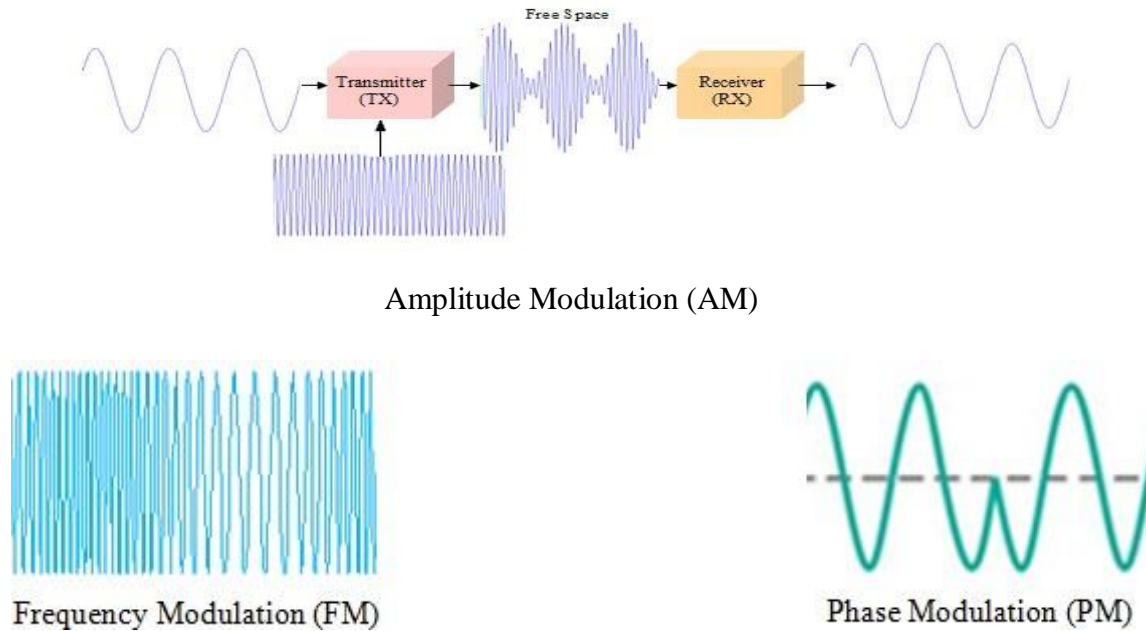


Fig.1 Types of Modulation

Analog systems use electrical signals to represent these natural patterns. The electrical signals are referred to as analog signals.



Fig. 1.1 Analog Signal

Digital systems use electrical signals that represent discrete, binary values. The electrical signals are referred to as **digital signals**. Specifically, binary digital signals use two discrete voltage levels to represent binary 1 or 0. Combining multiple bits into words permits us to represent larger values. Digital circuits operate on digital signals performing logic and arithmetic functions like we saw with our digital logic gates.



Fig. 1.2 Digital Signal

Digital advantages

1. Noise immunity: Noise immunity is the **most important advantage** of digital communications. Receiver circuitry can distinguish between a binary 0 and 1 with a significant amount of noise.

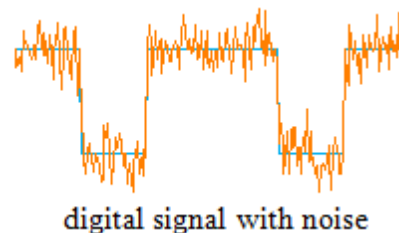
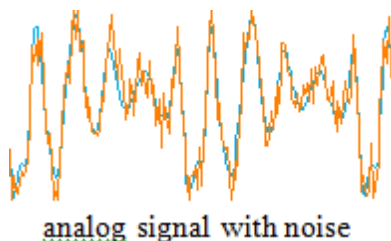


Fig. 1.3 Noise Signals

2. Error detection/corrections: Even if a digital signal does contain bit errors, many of these errors can be fixed at the receiver through the use of *error correcting codes*. Error correcting codes allows CDs with minor scratches to be played without errors.

3. Easier multiplexing: Multiplexing is the process of allowing multiple signals to share the same transmission channel, which we will talk about in our next lesson.

4. Easier to process and store: Since digital is the how computers work, digital signals can be easily processed (aka digital signal processing (DSP)) by computers. Similarly the digital format for easier storage of communication signals.

There are some disadvantages to digital communication; the most important is that the bandwidth size required is larger by almost twice that of analog communication. The digital circuitry is also more complex, but generally smaller and less expensive.

Conversion from analog to digital

Before we can use digital transmission, we must convert the signal of interest into a digital format. The natural pattern (i.e. speech) that we want to transmit will be acquired using an analog device. The analog signal will be translated into a digital signal using a method called **analog-to-digital (A/D) conversion**, digitizing a signal, or encoding. The device used to perform this translation is known as an analog-to-digital converter or ADC. Through A/D conversion continuously variable analog signals are changed into a series of binary numbers.

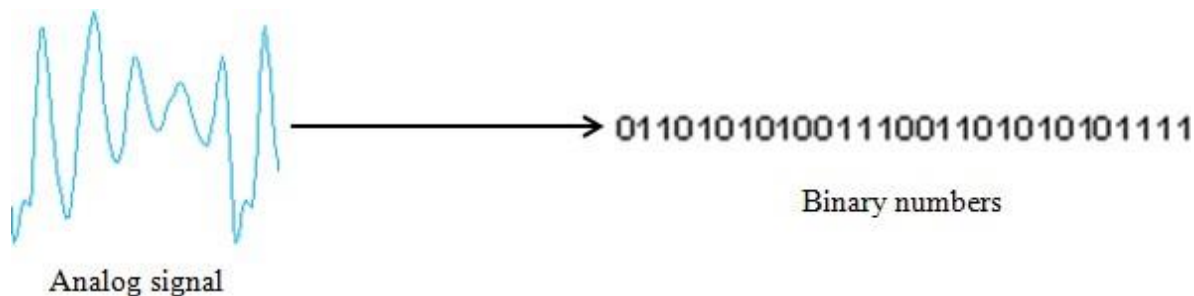


Fig. 1.4 A/D conversion

There are two steps we will use to convert an analog signal to a digital format as represented by binary numbers.

1. Sampling
2. Quantization

Sampling Process:

Sampling Theorem

Consider a band-limited signal with no frequency components above a certain frequency f_m . The sampling theorem states that this signal can be recovered completely from a set of samples of its amplitude, if the samples are taken at the rate of $f_s > 2f_m$ samples per second. This is often called the uniform sampling theorem for baseband or low-pass signals.

The minimum sampling rate, $2f_m$ samples per second, is called the Nyquist sampling rate (or Nyquist frequency); its reciprocal $1/(2f_m)$ (measured in seconds) is called the Nyquist interval.

$f_s = 2 * f_m$ is called the Nyquist sampling rate.

For telephone speech the standard sampling rate is 8 kHz (or one sample every 125 μ s).

Sampling Methods

Suppose we have an arbitrary signal (**the baseband signal $m(t)$**) which has a spectrum **$M(f)$** . Take infinitesimally short samples of the signal $m(t)$ at a uniform rate once every t_s seconds i.e. at a frequency f_s . This is the ideal form of sampling, it is called instantaneous (or impulse) sampling.

In effect the signal $m(t)$ is multiplied by a train of impulses giving rise to a train of pulses as in the lower line of the diagram. The train of sampling impulses has a frequency spectrum consisting of all harmonics or multiples of f_s and all are at the same amplitude.

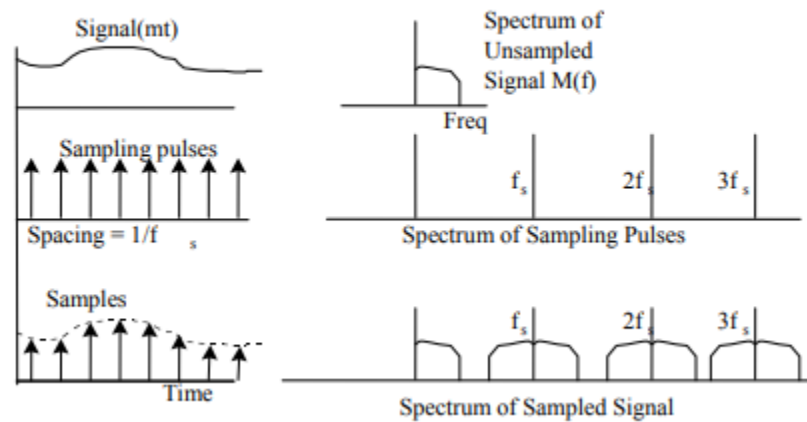


Fig. 1.5 Sampling

This **sampled signal** has a spectrum as shown where **$M(f)$** is **repeated unattenuated periodically** and appears around all multiples of the sampling frequency ($f_s = 1/t_s$).

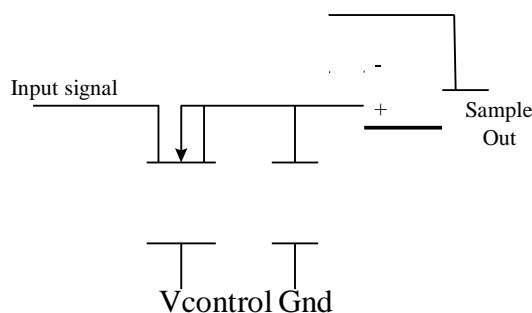
To recover $m(t)$ from the sampled signal we need only pass the sampled signal through a low pass filter with a stop frequency of $f_s/2$. All of the higher frequency components will be dropped. In the diagram, if f_s is greater than twice the highest frequency in $m(t)$ the repetitions of the sampled spectra around the harmonics of the f_s do not overlap.

Flat - top Sampling

An Analogue to Digital Converter requires that the sample value be held constant for a fixed time until the conversion is completed. This requires a flat-top sampled signal. This has approximately the same repeated frequency spectrum as with the instantaneous sampling above, but with each repetition slightly spread out.

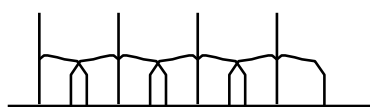
The simplest and most common sampling method is performed by a functional block termed a Sample and Hold (S/H) circuit.

Fig 1.6 Flat top sampling



The output from the circuit must be held at a constant level for the sampling duration. Vcontrol switches the MOSFET ON until the charge on C is equal to the amplitude of the sampled voltage. Vcontrol then goes LOW, the MOSFET is OFF and the charge is held by the capacitor. The charge held on the capacitor puts a voltage across the capacitor, and it is held at that value until the next time that Vcontrol switches the MOSFET ON. This is called a sample and hold circuit and is usually used as the input to an ADC.

Aliasing Error



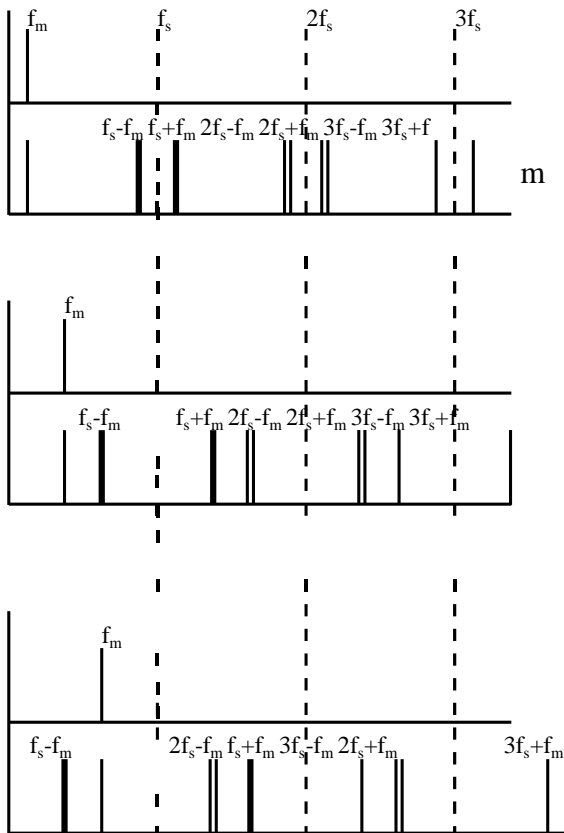
Spectrum of Sampled Signal

If a signal is under sampled (sampled at a rate below the Nyquist rate), the spectrum consists of overlapping repetitions of the sampled spectrum. Because of the overlapping tails a single repetition of the spectrum no longer has the complete information about the unsampled signal, and it is no longer possible to recover it from the sampled signal. To recover the original signal at the receiving end the sampled signal is passed through a lowpass filter with a cut off of $f_s/2$, we get a spectrum that is not the sampled signal but is a different version due to:

- Loss of the tail of the sampled signal spectrum beyond $f_s/2$
- This same tail appears inverted, or folded, onto the spectrum at the cut-off frequency. This tail inversion is known as aliasing, (or spectral folding or fold over distortion).

The aliasing distortion can be eliminated by cutting the tail (i.e. filtering) of the sampled signal beyond $f > f_s/2$ before the signal is sampled. By so doing, the overlap of

successive cycles in the sampled signal is avoided. The only error in the recovery of the unsampled signal is that caused by the missing tail above $f_s/2$.



It is simpler to consider aliasing by considering a single frequency component of $m(t)$. We will look at the frequency f_m and it is sampled at a rate f_s . The diagrams show the frequencies

which will be present in the sampled signal. There will be frequency components at f_m , $f_s - f_m$, $f_s + f_m$, $2f_s - f_m$, $2f_s + f_m$, $3f_s - f_m$, $3f_s + f_m$, etc. etc.

In the first case f_m is very much less than f_s , so that $f_s - f_m$ is much higher than the cut off of the filter ($f_s/2$).

In the second case f_m is below, but close to $f_s/2$, so that a sharp cut off filter is required to ensure that f_m is passed but $f_s - f_m$ is stopped.

In the third case f_m is higher than $f_s/2$, so that $f_s - f_m$ is less than $f_s/2$. The low pass filter with a cutoff of $f_s/2$ will therefore block f_m (the actual signal frequency) but will pass a signal with frequency $f_s - f_m$.

This is **aliasing**

Frequency

Fig. 1.7 Aliasing

Strictly speaking, a band limited signal does not exist in reality. It can be shown that if a signal is time limited it cannot be band limited. All physical signals are necessarily time limited because they begin at some finite instant and must terminate at some other finite instant. Hence, all practical signals are theoretically non band limited.

A real signal contains a finite amount of energy, therefore its frequency spectrum must decay at higher frequencies. Most of the signal energy resides in a finite band, and the spectrum at higher frequencies contributes little. The error introduced by cutting off the tail beyond a certain frequency B can be made negligible by making B sufficiently large.

Thus, for all practical purposes a signal can be considered to be essentially band limited at some value B, the choice of which depends upon the accuracy desired. A practical example of this is a speech signal. Theoretically, a speech signal, being a finite time signal, has an infinite bandwidth. But frequency components beyond 3400 Hz contribute a small fraction of the total energy. When speech signals are transmitted by PCM they are first passed through a lowpass filter of bandwidth of 3500 Hz. (This filter is called an **anti aliasing filter**). Higher sampling rates (i.e. 8000 samples/sec) permits recovery of the signal from its samples using relatively simple filters i.e. it allows for guard bands between the repetitions of the spectrum (otherwise recovering signals sampled at the Nyquist rate would require very sharp cut-off (ideal) filters).

In summary, aliasing distortion produces frequency components in the desired frequency band that did not exist in the original waveform. Aliasing problems are not confined to speech digitization processes. The potential for aliasing is present in any sample data system.

Motion picture taking, for example, is another sampling system that can produce aliasing. A common example occurs when filming a rotating wheel. Often the sampling process (the picture refresh rate) is too slow to keep up with the wheel movements and spurious rotational rates are produced. If the wheel rotates 355° between frames, it looks to the eye as if it has moved backwards 5° .

Pulse Modulation of Signals:

In many cases, bandwidth of communication link is much greater than signal bandwidth. The signal can be transmitted using short pulses with low duty cycle:

- ▲ Pulse amplitude modulation: width fixed, amplitude varies
- ▲ Pulse width modulation: position fixed, width varies
- ▲ Pulse position modulation: width fixed, position varies

All three methods can be used with time-division multiplexing to carry multiple signals over a

single channel

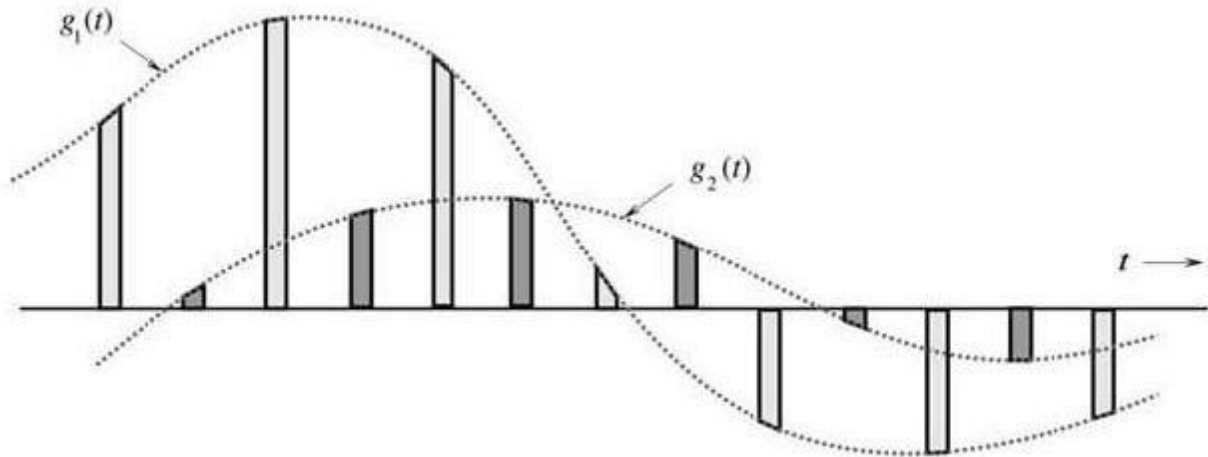
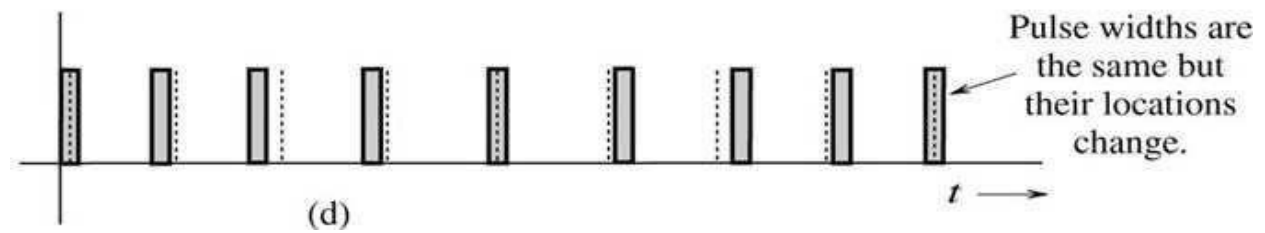
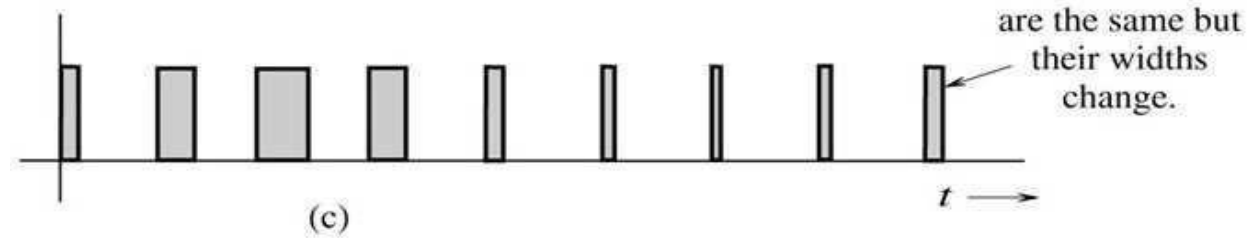
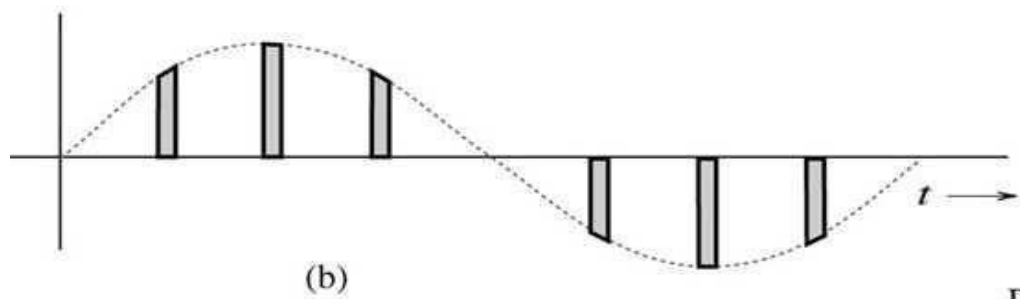
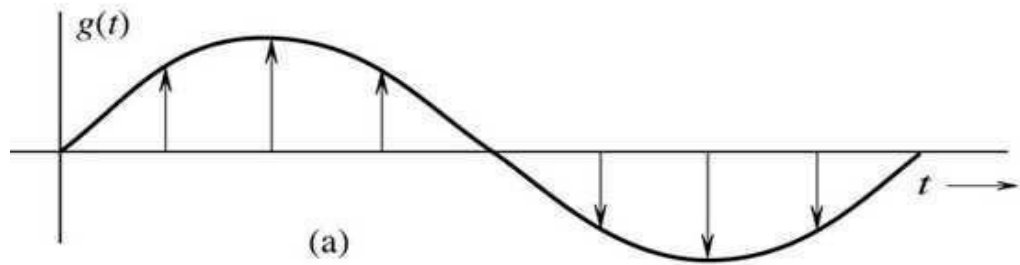


Fig. 1.9 PAM,PWM,PPM: Amplitude, Width and Position



Pulse Amplitude Modulation

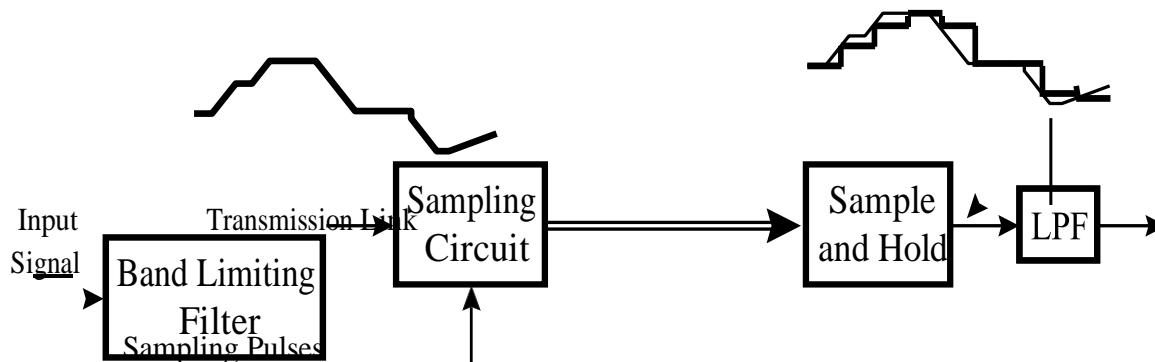


Fig. 1.10 PAM

A sampled signal consists of a train of pulses, where each pulse corresponds to the amplitude of the signal at the corresponding sampling time. The signal sent to line is modulated in amplitude and hence the name **Pulse Amplitude Modulation (PAM)**.

A complete PAM system must include a band limiting (or **anti aliasing**) filter before sampling to ensure that no spurious or source-related signals get folded back into the desired signal bandwidth - no aliasing. The input filter may also be designed to cut off very low frequencies to removed 50 Hz hum from power lines.

Several PAM signals can be multiplexed together as long as they are kept distinct and are recoverable at the receiving end. This system is one example of Time Division Multiplex (TDM) transmission (although it has never been widely used fo speech, it has applications in remote monitoring and telemetry).

The sample-and-hold circuit takes in each pulse and holds the amplitude of that pulse until the arrival of the next pulse. It produces a staircase approximation to the sampled wave form. With use of the staircase approximation, the power level of the signal coming out of the reconstructive filter (LPF) is nearly the same as the level of the sampled input signal.

The filters are assumed to have ideal characteristics - not like real filters. Filters with real attenuation slopes at the band edge can be used if the input signal is slightly over sampled. If sampling frequency is greater than twice the bandwidth, the spectral bands are sufficiently separated from each other that filters with gradual roll-off characteristics can be used.

As an example, sampled voice systems typically use band limiting filters with a 3 dB cut- off around 3.4 kHz and a sampling rate of 8 kHz. Thus the sampled signal is sufficiently attenuated at of 4 kHz to adequately reduce the energy level of the foldover spectrum.

Pulse Width Modulation (PWM)

Pulse width modulation is also called pulse duration modulation (PDM). PWM is more often used for control than for communication. A signal can be recovered exactly from its PWM samples at rate $2B$, provided the bandwidth is $\leq 0.637B$. PWM output can be generated by a sawtooth signal gating the input. Below the pulse width varies from nearly 0 to $1/2$ the pulse period.

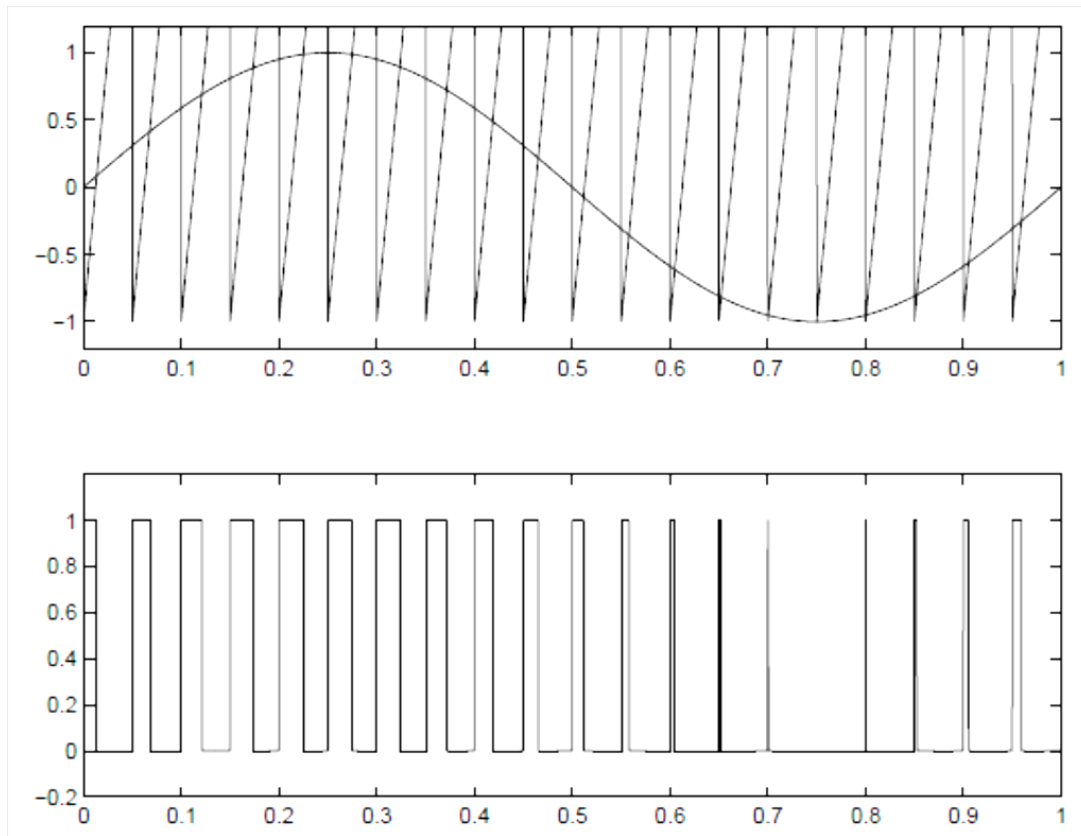


Fig.1.11 PWM

Pulse Position Modulation (PPM)

The value of the signal determines the delay of the pulse from the clock. Electrical circuits use timers such as 555. Microcontrollers can generate PPM (and PWM) in software.

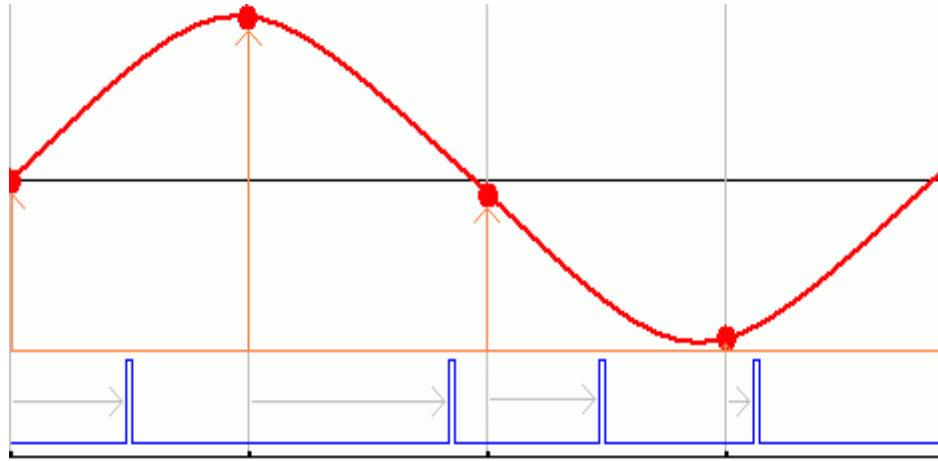


Fig. 1.12 PPM

Bandwidth Noise Tradeoff:

The noise analysis of PPM are as follows:

1. The figure of merit is proportional to square of the ratio $[B_T/W]$.
2. As the signal to noise ratio is reduced, the systems exhibit threshold effect.
3. With digital pulse modulation, the better noise performance than square law can be obtained.
4. The digital pulse modulation such as PCM gives negligible noise effect by increasing the average power in binary PCM signal.
5. With PCM, the bandwidth noise tradeoff can be related by exponential law.

Pulse Code Modulation

Pulse Code Modulation (PCM) is an extension of PAM wherein each analogue sample value is quantized into a discrete value for representation as a digital code word.

Thus, as shown below, a PAM system can be converted into a PCM system by adding a suitable analogue-to-digital (A/D) converter at the source and a digital-to-analogue (D/A) converter at the destination.

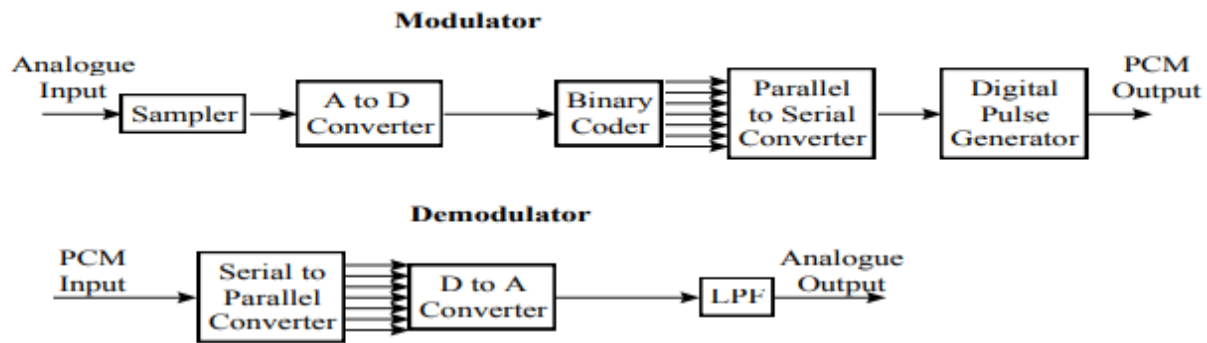


Fig 1.13 PCM

PCM is a true digital process as compared to PAM. In PCM the speech signal is converted from analogue to digital form.

PCM is standardized for telephony by the ITU-T (International Telecommunications Union - Telecoms, a branch of the UN), in a series of recommendations called the G series. For example the ITU-T recommendations for out-of-band signal rejection in PCM voice coders require that 14 dB of attenuation is provided at 4 kHz. Also, the ITU-T transmission quality specification for telephony terminals require that the frequency response of the handset microphone has a sharp roll-off from 3.4 kHz.

In quantization the levels are assigned a binary codeword. All sample values falling between two quantization levels are considered to be located at the centre of the quantization interval. In this manner the quantization process introduces a certain amount of error or distortion into the signal samples. This error known as quantization noise, is minimised by establishing a large number of small quantization intervals. Of course, as the number of quantization intervals increase, so must the number of bits increase to uniquely identify the quantization intervals. For example, if an analogue voltage level is to be converted to a digital system with 8 discrete levels or quantization steps three bits are required. In the ITU-T version there are 256 quantization steps, 128 positive and 128 negative, requiring 8 bits. A positive level is represented by having bit 8 (MSB) at 0, and for a negative level the MSB is 1.

Quantization

This is the process of setting the sample amplitude, which can be continuously variable to a discrete value. Look at Uniform Quantization first, where the discrete values are evenly spaced.

Uniform Quantization

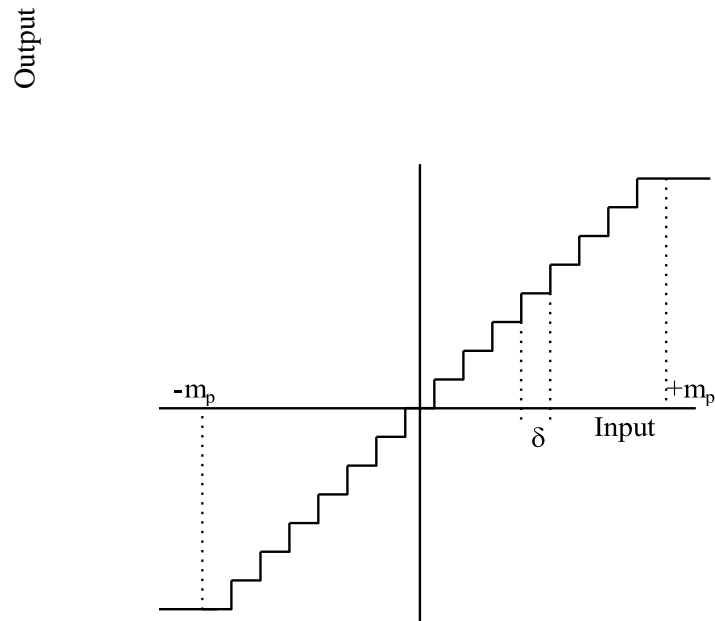


Fig.1.14 Quantization

We assume that the amplitude of the signal $m(t)$ is confined to the range $(-m_p, +m_p)$. This range $(2m_p)$ is divided into L levels, each of step size δ , given by

$$\delta = 2 m_p / L$$

A sample amplitude value is approximated by the midpoint of the interval in which it lies. The input/output characteristics of a uniform quantizer are shown.

Nonuniform quantizers

They have unequally spaced levels. The spacing can be chosen to optimize the Signal- to-Noise Ratio for a particular type of signal. It is characterized by:

Variable step size and Quantizer size depend on signal size

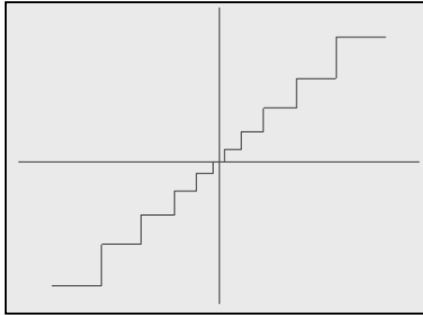


Fig. 1.15 Non uniform quantizer

Companding

In a uniform or linear PCM system the size of every quantization interval is determined by the SQR requirement of the lowest signal to be encoded. This interval is also for the largest signal - which therefore has a much better SQR.

Example: A 26 dB SQR for small signals and a 30 dB dynamic range produces a 56 dB SQR for the maximum amplitude signal.

In this way a uniform PCM system provides unneeded quality for large signals. In speech the max amplitude signals are the least likely to occur. The code space in a uniform PCM system is very inefficiently utilised.

A more efficient coding is achieved if the quantization intervals increase with the sample value. When the quantization interval is directly proportional to the sample value (assign small quantization intervals to small signals and large intervals to large signals) the SQR is constant for all signal levels. With this technique fewer bits per sample are required to provide a specified SQR for small signals and an adequate dynamic range for large signals (but still with the SQR as for the small signals). The quantization intervals are not constant and there will be a non linear relationship between the code words and the values they represent.

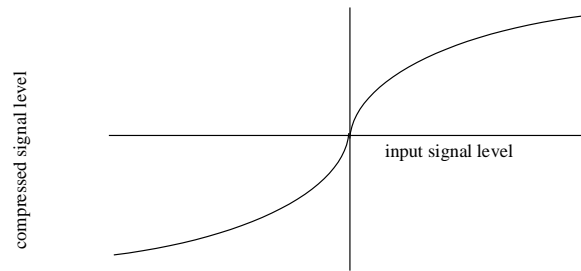


Fig. 1.16 Companding

Originally to produce the non linear quantization the baseband signal was passed through a non-linear amplifier with input/output characteristics as shown before the samples were taken. Low level signals were amplified and high level signals were attenuated. The larger the sample value the more it is **compressed** before encoding. The PCM decoder **expands** the compressed value using an inverse compression characteristic to recover the original sample value. The two processes are called **companding**.

There are 2 companding schemes to describe the curve above:

1. μ -Law Companding (also called log-PCM)

This is used in North America and Japan. It uses a logarithmic compression curve which is ideal in the sense that quantization intervals and hence quantization noise is directly proportional to signal level (and so a constant SQR).

2. A- Law Companding

This is the ITU-T standard. It is used in Europe and most of the rest of the world. It is very similar to the μ -Law coding. It is represented by straight line segments to facilitate digital companding. Originally the non linear function was obtained using non linear devices such as special diodes. These days in a PCM system the A to D and D to A converters (ADC and DAC) include a companding function.

Differential pulse coding schemes

PCM transmits the absolute value of the signal for each frame. Instead we can transmit information about the difference between each sample. The two main differential coding schemes are:

- Delta Modulation
- Differential PCM and Adaptive Differential PCM (ADPCM)

Delta Modulation

Delta modulation converts an analogue signal, normally voice, into a digital signal.

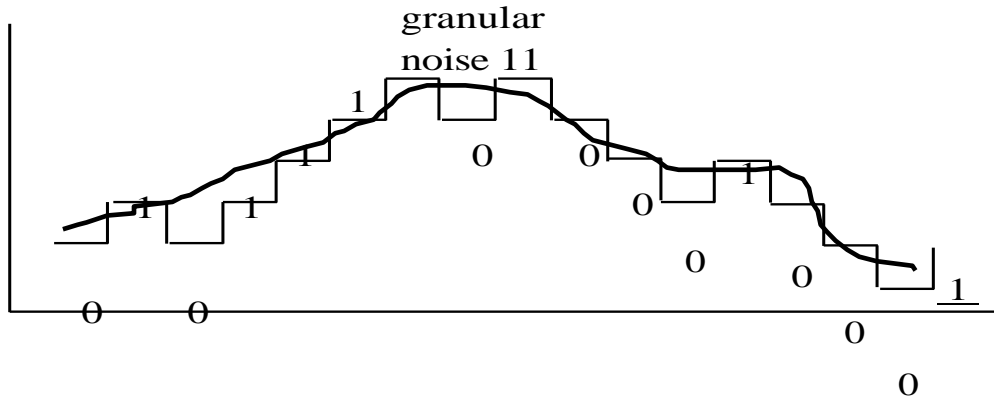


Fig. 1.17 Delta Modulation

The analogue signal is sampled as in the PCM process. Then the sample is compared with the previous sample. The result of the comparison is quantified using a one bit coder. If the sample is greater than the previous sample a 1 is generated. Otherwise a 0 is generated. The advantage of delta modulation over PCM is its simplicity and lower cost. But the noise performance is not as good as PCM.

To reconstruct the original from the quantization, if a 1 is received the signal is increased by a step of size q , if a 0 is received the output is reduced by the same size step. Slope overload occurs when the encoded waveform is more than a step size away from the input signal. This condition happens when the rate of change of the input exceeds the maximum change that can be generated by the output.

Overload will occur if:

$$\max \left| \frac{d}{dt} x(t) \right| > \frac{\delta}{T_s}$$

$$\max \left| \frac{d}{dt} A_m \sin(2\pi f_m t) \right| > \frac{\delta}{T_s}$$

$$\max |A_m 2\pi f_m \cos(2\pi f_m t)| > \frac{\delta}{T_s}$$

$$A_m 2\pi f_m > \frac{\delta}{T_s}$$

or

$$A_m > \frac{\delta}{2\pi f_m T_s}$$

where: $x(t)$ = input signal, δ = step size, T = period between samples, f_s = sampling frequency

Example $A = 2$ V, $F = 3.4$ kHz, and the signal is sampled 1,000,000 times per second, requires $q > 2 * 3.14 * 3,400 * 2 / 1,000,000$ V > 42.7 mV

Granular noise occurs if the slope changes more slowly than the step size. The reconstructed signal oscillates by 1 step size in every sample. It can be reduced by decreasing the step size. This requires that the sample rate be increased. Delta Modulation requires a sampling rate much higher than twice the bandwidth. It requires oversampling in order to obtain an accurate prediction of the next input, since each encoded sample contains a relatively small amount of information. Delta Modulation requires higher sampling rates than PCM.

Differential PCM (DPCM) and ADPCM

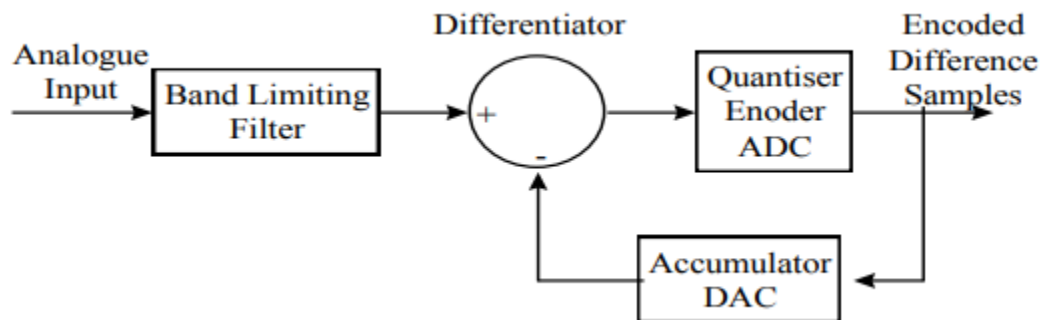


Fig. 1.18 DPCM

DPCM is also designed to take advantage of the redundancies in a typical speech waveform. In DPCM the differences between samples are quantized with fewer bits that would be used for quantizing an individual amplitude sample. The sampling rate is often the same as for a comparable PCM system, unlike Delta Modulation.

Adaptive Differential Pulse Code Modulation ADPCM is standardised by ITU-T recommendations G.721 and G.726. The method uses 32,000 bits/s per voice channel, as compared to standard PCM's 64,000 bits/s. Four bits are used to describe each sample, which represents the difference between two adjacent samples. Sampling is 8,000 times a second. It makes it possible to reduce the bit flow by half while maintaining an acceptable quality. While the use of ADPCM (rather than PCM) is imperceptible to humans, it can significantly reduce the throughput of high speed modems and fax transmissions.

The principle of ADPCM is to use our knowledge of the signal in the past time to predict the signal one sample period later, in the future. The predicted signal is then compared

with the actual signal. The difference between these is the signal which is sent to line - it is the error in the prediction. However this is not done by making comparisons on the incoming audio signal - the comparisons are done after PCM coding.

To implement ADPCM the original (audio) signal is sampled as for PCM to produce a code word. This code word is manipulated to produce the predicted code word for the next sample. The new predicted code word is compared with the code word of the second sample. The result of this comparison is sent to line. Therefore we need to perform PCM before ADPCM.

The ADPCM word represents the prediction error of the signal, and has no significance itself. Instead the decoder must be able to predict the voltage of the recovered signal from the previous samples received, and then determine the actual value of the recovered signal from this prediction and the error signal, and then to reconstruct the original waveform.

ADPCM is sometimes used by telecom operators to fit two speech channels onto a single 64 kbit/s link. This was very common for transatlantic phone calls via satellite up until a few years ago. Now, nearly all calls use fibre optic channels at 64 kbit/s.

QUESTION BANK UNIT I

SL.NO	Part A
1.	Summarize the advantages of digital communication.
2.	Outline the demerit of digital communication.
3.	Illustrate sampling theorem for band limited signals and the filters to avoid aliasing.
4.	Demonstrate companding? sketch the input-output characteristic of a compressor and expander.
5.	Make use of sampling theorem the condition for aliasing?
6.	Distinguish natural and flat top sampling?
7.	Illustrate the difference between uniform and non-uniform quantization?
8.	How would you show your understanding of the components required for signal reconstruction
9.	Describe non uniform quantization.
10.	Interpret the principle of ADM.
11.	Illustrate the difference between DM and ADM.

12.	A band pass signal has the spectral range that extends from 20 kHz to 82 kHz. Find the acceptance range of sampling frequency f_s
13.	A certain low pass bandlimited signal $x(t)$ is sampled and the spectrum of the sampled version has the first guard band from 1500Hz to 1900Hz. How will you determine the sampling frequency and the maximum frequency of the signal?
14.	Find the Nyquist sampling rate for a band limited signal given by $x(t) = \cos 1000t$.
15.	Give the interpolation formula for the reconstruction of the original signal
16.	Find the Nyquist sampling rate for a band limited signal given by $x(t) = \cos 1000t$.
17.	Suggest the minimum sampling frequency for a band pass signal with frequency range from 200kHz to 300kHz.
18.	Specify the Nyquist sampling rate for the following signals: $x(t) = \sin(200t)$ and b) $x(t) \sin^2(200t)$

SL.NO	Part B
1.	(i) Develop your understanding on PCM transmitter and receiver with neat sketch(10) (ii) Narrate the merits and demerits of digital communication system(6)
2.	Elucidate the working of each functional blocks of a digital communication system with suitable illustration.
3.	A band limited analog signal $\sin 2\pi f_m t$ is naturally sampled using an impulse train with frequency f_s . Derive the expression for the sampled signal $s(t)$ and plot the spectrum of $s(t)$ for the three conditions: i) $f_s > 2f_m$ ii) $f_s = 2f_m$ and iii) $f_s < 2f_m$
4.	Make use of suitable derivation explain in detail about sampling theorem and discuss the reconstruction of the signal from its samples.
5.	Outline the process of natural sampling and how the message can be reconstructed from its samples. Also illustrate the effect of aliasing with neat sketch.
6.	(i) Describe delta modulation system in detail with a neat block diagram. (10) (ii) Derive the derivation for slope overload error(6).
7.	Explain with neat diagram the adaptive delta modulation and demodulation system in detail.
8.	(i) Illustrate the functioning of DPCM system with block diagram(10) (ii) Explain the structure of linear predictor?(6)
9.	Explain quantization error and derive an expression for maximum signal to noise ratio in PCM system that uses linear quantization.

UNIT II – BASEBAND PULSE TRANSMISSION - SEC1313

UNIT 2 BASEBAND PULSE TRANSMISSION

Naturally, noise and ISI arise in the system simultaneously. However, to understand how they affect the performance of the system, we first consider them separately channel noise. Naturally, noise and ISI arise in the system simultaneously. However, to understand how they affect the performance of the system, we first consider them separately. The device for the optimum detection of such a pulse immersed in additive white noise involves the use of a linear-time-invariant filter known as a matched filter, which is so called because its impulse response is matched to the pulse signal.

2.1 Matched Filter

A basic problem that often arises in the study of communication systems is that of detecting a pulse transmitted over a channel that is corrupted by channel noise (i.e., additive noise at the front end of the receiver). For the purpose of the discussion presented in this section, we assume that the major source of system limitation is the channel noise.

Consider then the receiver model shown in Figure 2.1, involving a linear time-invariant filter of impulse response $h(t)$. The filter input $x(t)$ consists of a pulse signal $g(t)$ corrupted by additive channel noise $w(t)$, as shown by

$$x(t) = g(t) + w(t), 0 < t < T \quad (2.1)$$

where T is an arbitrary observation interval. The pulse signal $g(t)$ may represent a binary symbol 1 or in a digital communication system. The $w(t)$ is the sample function of a white noise process of zero mean and power spectral density $N/2$. It is assumed that the receiver has knowledge of the waveform of the pulse signal $g(t)$. The source of uncertainty lies in the noise $w(t)$. The function of the receiver is to detect the pulse signal $g(t)$ in an optimum manner, given the received signal $x(t)$. To satisfy this requirement, we have to optimize the design of the filter so as to minimize the effects of noise at the filter output in some statistical sense, and thereby enhance the detection of the pulse signal $g(t)$. Since the filter is linear, the resulting output $y(t)$ may be expressed as

$$y(t) = g(t) + n(t) \quad (2.2)$$

Where $g(t)$ and $n(t)$ are produced by the signal and noise components of the input $x(t)$, respectively. A simple way of describing the requirement that the output signal component $g(t)$ be considerably greater than the output noise component $n(t)$ is to have the filter make the instantaneous power in the output signal $g(t)$, measured at time $t = T$, as large as possible compared with the average

power of the output noise $n(t)$. This is equivalent to maximizing the peak pulse signal-to-noise ratio, defined as

$$\eta = \frac{|g_o(T)|^2}{E[n^2(t)]} \quad (2.3)$$

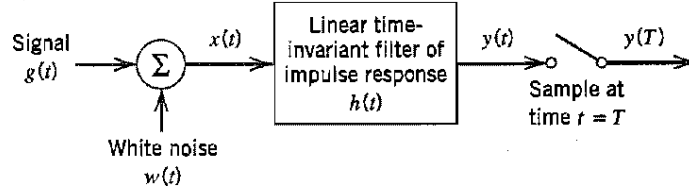


Figure 2.1 Linear receiver.

where $|g_o(T)|^2$ is the instantaneous power in the output signal, E is the statistical expectation operator, and $E[n^2(t)]$ is a measure of the average output noise power. The requirement is to specify the impulse response $h(t)$ of the filter such that the output signal-to-noise ratio in Equation (2.3) is maximized.

Let $G(f)$ denote the Fourier transform of the known signal $g(t)$, and $H(f)$ denote the frequency response of the filter. Then the Fourier transform of the output signal $g_o(t)$ is equal to $H(f)G(f)$, and $g_o(t)$ is itself given by the inverse Fourier transform

$$g_o(t) = \int_{-\infty}^{\infty} H(f)G(f) \exp(j2\pi ft) df \quad (2.4)$$

Hence, when the filter output is sampled at time $t = T$, we have (in the absence of channel noise) Consider next the effect on the filter output due to the noise $w(t)$ acting alone. The power spectral density $S_N(f)$ of the output noise $n(t)$ is equal to the power spectral density of the input noise $w(t)$ times the squared magnitude response $|H(f)|^2$

Since $w(t)$ is white with constant power spectral density $N_0/2$, it follows that

$$S_N(f) = \frac{N_0}{2} |H(f)|^2 \quad (2.6)$$

The average power of the output noise $n(t)$ is therefore

$$E[n^2(t)] = \int_{-\infty}^{\infty} S_N(f) df$$

Thus substituting Equations (2.5) and (2.7) into (2.3), we may rewrite the expression for the peak pulse signal-to-noise ratio as

$$\begin{aligned} E[n^2(t)] &= \int_{-\infty}^{\infty} S_N(f) df \\ &= \frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df \end{aligned} \quad (2.7)$$

Our problem is to find, for a given $G(f)$, the particular form of the frequency response $H(f)$ of the filter that makes η a maximum. To find the solution to this optimization problem, we apply a mathematical result known as Schwarz's inequality to the numerator of Equation (2.8).

if we have two complex functions $u(x)$ and $v(x)$ in the real variable x , satisfying the conditions

where k is an arbitrary constant, and the asterisk denotes complex conjugation.

$$\eta = \frac{\left| \int_{-\infty}^{\infty} H(f)G(f) \exp(j2\pi fT) df \right|^2}{\frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df} \quad (4.8)$$

2.2 Error Rate due to noise

To derive a formula for the error rate in a matched filter due to noise.

To proceed with the analysis, consider a binary PCM system based on polar non- return-to-zero (NRZ) signaling. In this form of signaling, symbols 1 and 0 are represented by positive and negative rectangular pulses of equal amplitude and equal duration. The channel noise is modeled as additive white Gaussian noise $w(t)$ of zero mean and power spectral density $N/2$; $0 \leq t \leq T_b$, the Gaussian assumption is needed for later calculations. In the signaling interval, the received signal is thus written as follows:

$$x(t) = \begin{cases} +A + w(t), & \text{symbol 1 was sent} \\ -A + w(t), & \text{symbol 0 was sent} \end{cases}$$

where T_b is the bit duration, and A is the transmitted pulse amplitude. It is assumed that the receiver has acquired knowledge of the starting and ending times of each transmitted pulse; in other words, the receiver has prior knowledge of the pulse shape, but not its polarity. Given the noisy signal $x(t)$, the receiver is required to make a decision in each signaling interval as to whether the transmitted symbol is a 1 or a 0.

The structure of the receiver used to perform this decision-making process is shown in Figure 2.4. It consists of a matched filter followed by a sampler, and then finally a

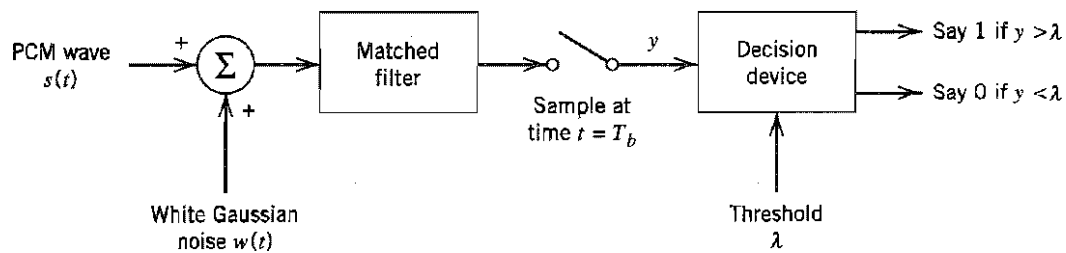


Figure 2.4 Receiver for baseband transmission of binary-encoded PCM wave using polar NRZ signaling.

decision device. The filter is matched to a rectangular pulse of amplitude A and duration T exploiting the bit-timing information available to the receiver. The resulting matched filter output is sampled at the end of each signaling interval. The presence of channel noise $w(t)$ adds randomness to the matched filter output.

Let y denote the sample value obtained at the end of a signaling interval. The sample value y is compared to a preset threshold A in the decision device. If the threshold is exceeded, the receiver makes a decision in favor of symbol 1; if not, a decision is made in favor of symbol 0. We adopt the convention that when the sample value y is exactly equal to the threshold A , the receiver just makes a guess as to which symbol was transmitted such a decision is the same as that obtained by flipping a fair coin, the outcome of which will not alter the average probability of error.

There are two possible kinds of error to be considered:

1. Symbol 1 is chosen when a 0 was actually transmitted; we refer to this error as an error of the first kind.
2. Symbol 0 is chosen when a 1 was actually transmitted; we refer to this error as an error of the second kind.

To determine the average probability of error, we consider these two situations separately.

Suppose that symbol was sent. Then, according to Equation (2.21), the received signal is

$$x(t) = -A + w(t), \quad 0 \leq t \leq T_b \quad (2.22)$$

Correspondingly, the matched filter output, sampled at time $t = T_b$, is given by (in light of Example 2.1 with kAT_b set equal to unity for convenience of presentation)

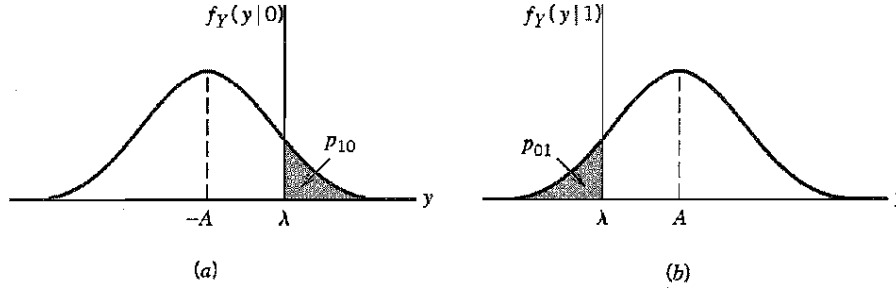


Figure 4.5 Noise analysis of PCM system (a) Probability density function of random variable y at matched filter output when 0 is transmitted (b) pdf of y when 1 is transmitted

2.3 Intersymbol Interference

The next source of bit errors in a baseband-pulse transmission system that we wish to study is intersymbol interference (ISI), which arises when the communication channel is dispersive. First of all, however, we need to address a key question: Given a pulse shape of interest, how do we use it to transmit data in M-ary form?

The answer lies in the use of discrete pulse modulation, in which the amplitude, duration, or position of the transmitted pulses is varied in a discrete manner in accordance with the given data stream. However, for the baseband transmission of digital data, the use of discrete pulse-amplitude modulation (PAM) is one of the most efficient schemes in terms of power and bandwidth utilization. Accordingly, we confine our attention to discrete PAM systems. We begin

the study by first considering the case of binary data; later in the chapter, we consider the more general case of M-ary data.

Consider then a baseband binary PAM system, a generic form of which is shown in Figure 2.7. The incoming binary sequence $\{b_k\}$ consists of symbols 1 and 0, each of duration T_b . The pulse-amplitude modulator modifies this binary sequence into a new sequence of short pulses (approximating a unit impulse), whose amplitude a_k is represented in the polar form

$$a_k = \begin{cases} +1 & \text{if symbol } b_k \text{ is 1} \\ -1 & \text{if symbol } b_k \text{ is 0} \end{cases} \quad (2.42)$$

The sequence of short pulses so produced is applied to a transmit filter of impulse response $g(t)$, producing the transmitted signal

$$s(t) = \sum_k a_k g(t - kT_b) \quad (2.43)$$

The signal $s(t)$ is modified as a result of transmission through the channel of impulse response $h(t)$. In addition, the channel adds random noise to the signal at the receiver input. The noisy signal $x(t)$ is then passed through a receive filter of impulse response $c(t)$. The resulting filter output $y(t)$ is sampled synchronously with the transmitter, with the sampling instants being determined by a clock or timing signal that is usually extracted from the receive filter output. Finally, the sequence of samples thus obtained is used to reconstruct the original data sequence by means of a decision device. Specifically, the amplitude of each sample is compared to a threshold A . If the threshold A is exceeded, a decision is made in favor of symbol

1. If the threshold A is not exceeded, a decision is made in favor of symbol 0. If the sample amplitude equals the threshold exactly, the flip of a

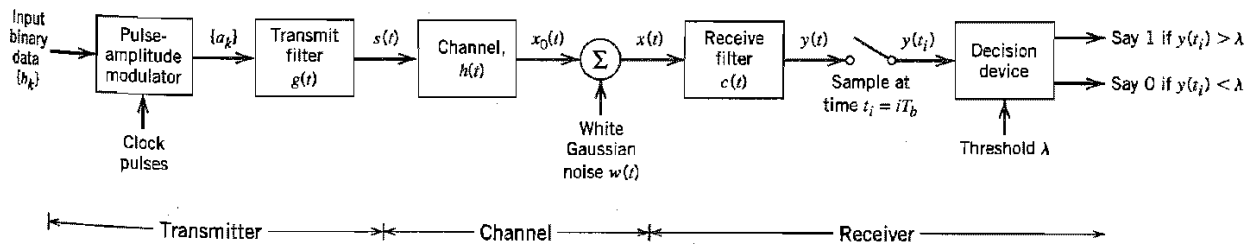


FIGURE 2.7 Baseband binary data transmission system.

fair coin will determine which symbol was transmitted (i.e., the receiver simply makes a random guess).

The receive filter output is written as

$$y(t) = \mu \sum_k a_k p(t - kT_b) + n(t)$$

where μ is a scaling factor, and the pulse $p(t)$ is to be defined.

The scaled pulse $\mu p(t)$ is obtained by a double convolution involving the impulse response $g(t)$ of the transmit filter, the impulse response $h(t)$ of the channel, and the impulse response $c(t)$ of the receive filter, as shown by

$$\mu p(t) = g(t) \cdot h(t) \cdot c(t) \quad (2.45)$$

where the star denotes convolution. We assume that the pulse $p(t)$ is normalized by setting

Since convolution in the time domain is transformed into multiplication in the frequency domain, we may use the Fourier transform to change Equation (2.45) into the equivalent form

$$P(f) = G(f)H(f)C(f) \quad (2.47)$$

where $P(f)$, $G(f)$, $H(f)$, and $C(f)$ are the Fourier transforms of $p(t)$, $g(t)$, $h(t)$, and $c(t)$, respectively.

Finally, the term $n(t)$ in Equation (2.44) is the noise produced at the output of the receive filter due to the channel noise $w(t)$. It is customary to model $w(t)$ as a white Gaussian noise of zero mean.

The receive filter output $y(t)$ is sampled at time $t_i = iT_b$ (with i taking on integer values), yielding [in light of Equation (2.46)]

$$\begin{aligned} y(t_i) &= \mu \sum_{k=-\infty}^{\infty} a_k p[(i-k)T_b] + n(t_i) \\ &= \mu a_i + \mu \sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} a_k p[(i-k)T_b] + n(t_i) \end{aligned} \quad (2.48)$$

In Equation (2.48), the first term μa_i represents the contribution of 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 due to the occurrence of pulses before and after the sampling instant t_i is called intersymbol interference (ISI). The last term $n(t_i)$ represents the noise sample at time t_i .

2.4 Nyquist's criterion for Distortionless Base band Binary Transmission

Typically, the frequency response of the channel and the transmitted pulse shape are specified, and the problem is to determine the frequency responses of the transmit and receive filters so as to reconstruct the original binary data sequence $\{b_k\}$. The receiver does this by extracting and then decoding the corresponding sequence of coefficients, $\{a_k\}$, from the output $y(t)$. The extraction involves sampling the output $y(t)$ at time $t = iT_b$. The decoding requires that the weighted pulse contribution $a_k p(t - kT_b)$ be free from ISI due to the overlapping tails of all other weighted pulse contributions represented by $k \neq i$.

$$a_k p(iT_b - kT_b) = 0 \quad k \neq i.$$

This, in turn, requires that we control the overall pulse $p(t)$, as shown by

$$p(iT_b - kT_b) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \quad (2.49)$$

where $p(0) = 1$, by normalization. If $p(t)$ satisfies the conditions of Equation (2.49), the receiver output $y(t)$ given in Equation (2.48) simplifies to (ignoring the noise term)

$$y(t_i) = \mu a_i \quad \text{for all } i$$

which implies zero intersymbol interference. Hence, the two conditions of Equation (2.49) ensure perfect reception in the absence of noise.

Ideal Nyquist Channel

The simplest way of satisfying Equation (2.53) is to specify the frequency function $P(f)$ to be in the form of a rectangular function, as shown by

$$\begin{aligned} P(f) &= \begin{cases} \frac{1}{2W}, & -W < f < W \\ 0, & |f| > W \end{cases} \\ &= \frac{1}{2W} \text{rect}\left(\frac{f}{2W}\right) \end{aligned} \quad (2.54)$$

where $\text{rect}(\cdot)$ stands for a rectangular function of unit amplitude and unit support centered on $f = 0$, and the overall system bandwidth W is defined by

$$W = \frac{R_b}{2} = \frac{1}{2T_b} \quad (2.55)$$

According to the solution described by Equations (2.54) and (2.55), no frequencies of absolute value exceeding half the bit rate are needed. Hence, from Fourier-transform pair 2 of Table A6.3 we find that a signal waveform that produces zero intersymbol interference is defined by the sinc function:

$$\begin{aligned} p(t) &= \frac{\sin(2\pi Wt)}{2\pi Wt} \\ &= \text{sinc}(2Wt) \end{aligned} \quad (2.56)$$

The special value of the bit rate $R_b = 2W$ is called the Nyquist rate, and W is itself called the Nyquist bandwidth. Correspondingly, the ideal baseband pulse transmission system described by Equation (2.54) in the frequency domain or, equivalently, Equation (2.56) in the time domain, is called the ideal Nyquist channel.

Figures 2.8a and 2.8b show plots of $P(f)$ and $p(t)$, respectively. In Figure 2.8a, the normalized form of the frequency function $P\{f\}$ is plotted for positive and negative frequencies. In Figure 2.8b, we have also included the signaling intervals and the corresponding centered sampling instants. The function $p(t)$ can be regarded as the impulse response of an ideal low-pass filter with passband magnitude response $1/2W$ and bandwidth W .

The function $p\{t\}$ has its peak value at the origin and goes through zero at integer multiply of the bit duration T_b . It is apparent that if the received waveform $y\{t\}$ is sampled at the instants of time $t = 0, \pm T_b, \pm 2T_b, \dots$, then the pulses defined by $fip(t - iT_b)$ with arbitrary amplitude p and index $i = 0, \pm 1, \pm 2, \dots$, will not interfere with each other. This condition is illustrated in Figure 2.9 for the binary sequence 1011010.

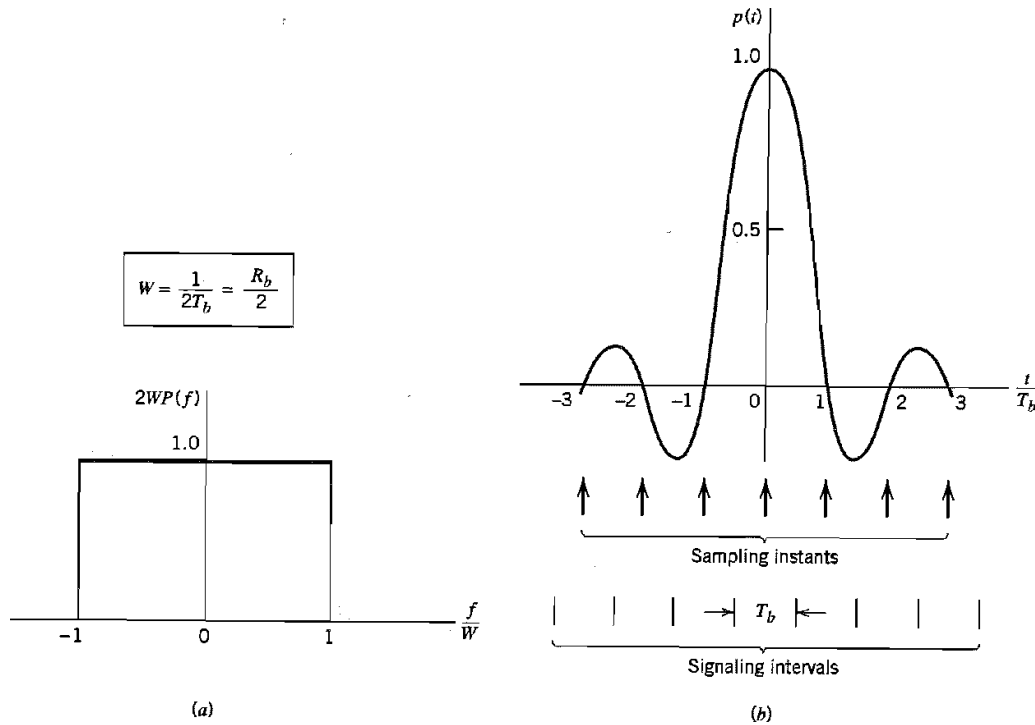


Figure 2.8 (a) Ideal magnitude response (b) Ideal basic pulse shape

There are two practical difficulties that make it an undesirable objective for system design:

1. It requires that the magnitude characteristic of $P(f)$ be flat from $-W$ to W , and zero elsewhere. This is physically unrealizable because of the abrupt transitions at frequency band edges $\pm W$.
2. The function $p(t)$ decreases as $1/|f|$ for large $|f|$, resulting in a slow rate of decay. This is also caused by the discontinuity of $P(f)$ at $\pm W$. Accordingly, there is practically no margin of error in sampling times in the receiver.

Raised Cosine Spectrum

We may overcome the practical difficulties encountered with the ideal Nyquist channel by extending the bandwidth from the minimum value $W = R_b/2$ to an adjustable value between W and $2W$. We now specify the overall frequency response $P(f)$ to satisfy a condition more elaborate than that for the ideal Nyquist channel

This frequency response consists of a flat portion and a rolloff portion that has a sinusoidal form, as follows:

$$P(f) = \begin{cases} \frac{1}{2W}, & 0 \leq |f| < f_1 \\ \frac{1}{4W} \left\{ 1 - \sin \left[\frac{\pi(|f| - W)}{2W - 2f_1} \right] \right\}, & f_1 \leq |f| < 2W - f_1 \\ 0, & |f| \geq 2W - f_1 \end{cases}$$

The frequency parameter f_1 and bandwidth W are related by

$$\alpha = 1 - \frac{f_1}{W}$$

The parameter α is called the rolloff factor; it indicates the excess bandwidth over the ideal solution, W . Specifically, the transmission bandwidth B_T is defined by

$$\begin{aligned} B_T &= 2W - f_1 \\ &= W(1 + \alpha) \end{aligned}$$

The frequency response $P(f)$, normalized by multiplying it by $2W$, is plotted in Figure 2.10a for three values of α , namely, 0, 0.5, and 1. We see that for $\alpha = 0.5$ or 1,

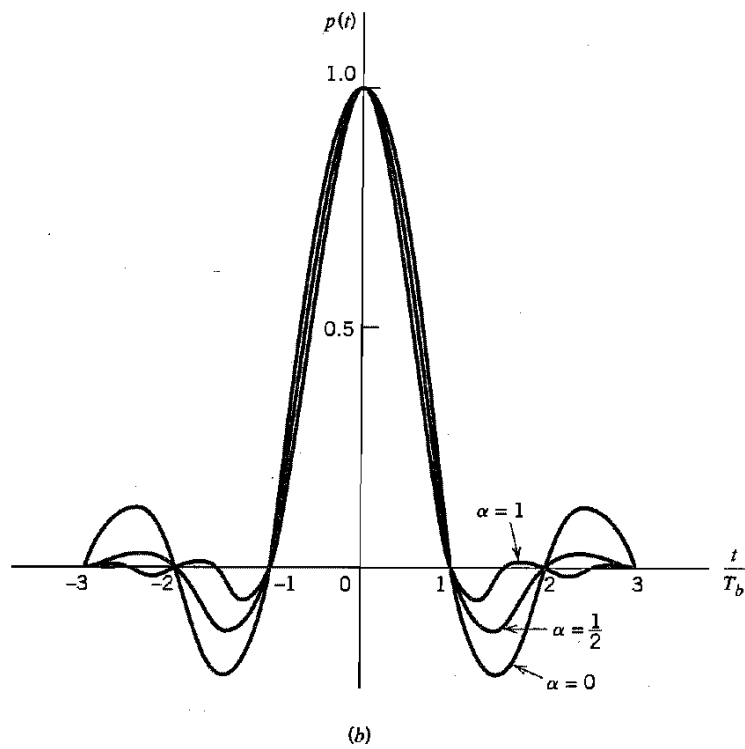
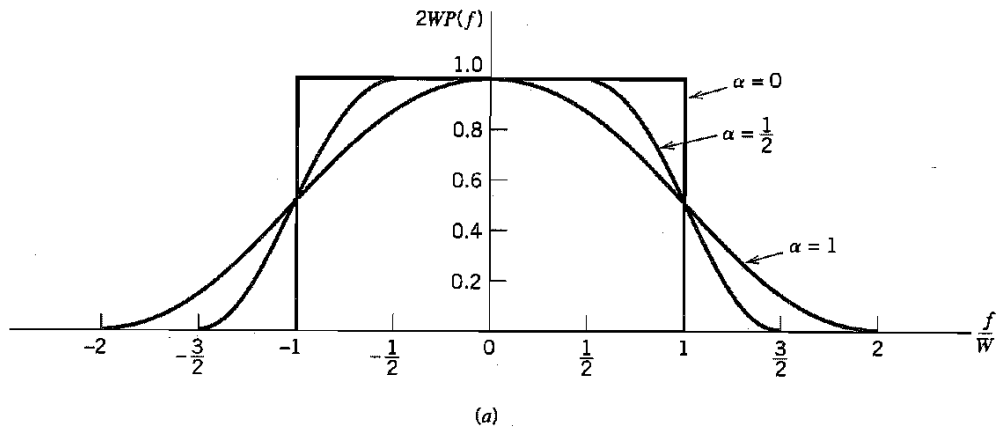


Figure 2.10 Response for different rolloff factors. (a) Frequency response (b) Time response

2.5 Correlative level coding

Thus far we have treated intersymbol interference as an undesirable phenomenon that

produces a degradation in system performance. Indeed, its very name connotes a nuisance effect. Nevertheless, by adding intersymbol interference to the transmitted signal in a controlled manner, it is possible to achieve a signaling rate equal to the Nyquist rate of $2W$ symbols per second in a channel of bandwidth W Hertz. Such schemes are called correlative-level coding or partial-response signaling schemes. The design of these schemes is based on the following premise: Since intersymbol interference introduced into the transmitted signal is known, its effect can be interpreted at the receiver in a deterministic way. Thus correlative-level coding may be regarded as a practical method of achieving the theoretical maximum signaling rate of $2W$ symbols per second in a bandwidth of W Hertz as postulated by Nyquist, using realizable and perturbation-tolerant filters.

Duobinary Signaling

The basic idea of correlative-level coding will now be illustrated by considering the specific example of duobinary signaling, where "duo" implies doubling of the transmission capacity of a straight binary system. This particular form of correlative-level coding is also called class I partial response.

Consider a binary input sequence $\{b_k\}$ consisting of uncorrelated binary symbols 1 and 0, each having duration T_b . As before, this sequence is applied to a pulse-amplitude modulator producing a two-level sequence of short pulses (approximating a unit impulse), whose amplitude a_k is defined by

$$a_k = \begin{cases} +1 & \text{if symbol } b_k \text{ is 1} \\ -1 & \text{if symbol } b_k \text{ is 0} \end{cases} \quad (2.65)$$

When this sequence is applied to a duobinary encoder, it is converted into a three-level output, namely, -2 , 0 , and $+2$. To produce this transformation, we may use the scheme shown in Figure 2.11. The two-level sequence $\{a_k\}$ is first passed through a simple filter involving a single delay element and summer. For every unit impulse applied to the

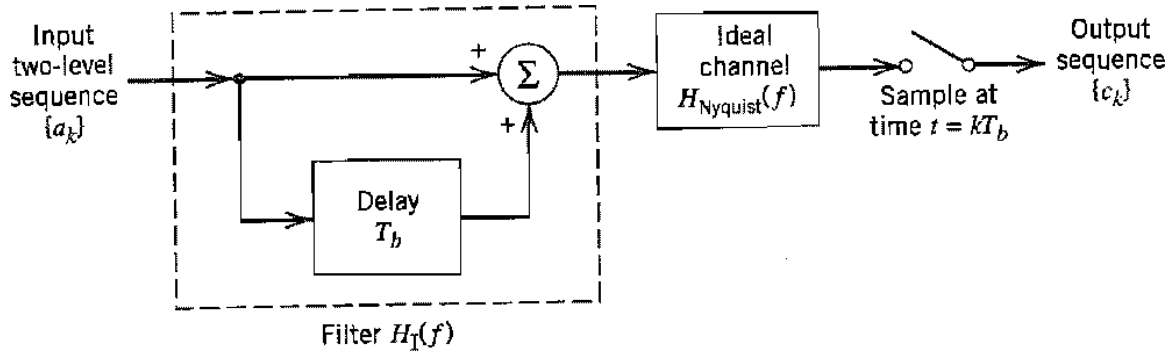


Figure 2.11 Duo binary signaling scheme.

input of this filter, we get two unit impulses spaced T_b seconds apart at the filter output. We may therefore express the duobinary coder output c_k as the sum of the present input pulse a_k and its previous value a_{k-1} as shown by

$$c_k = a_k + a_{k-1} \quad (2.66)$$

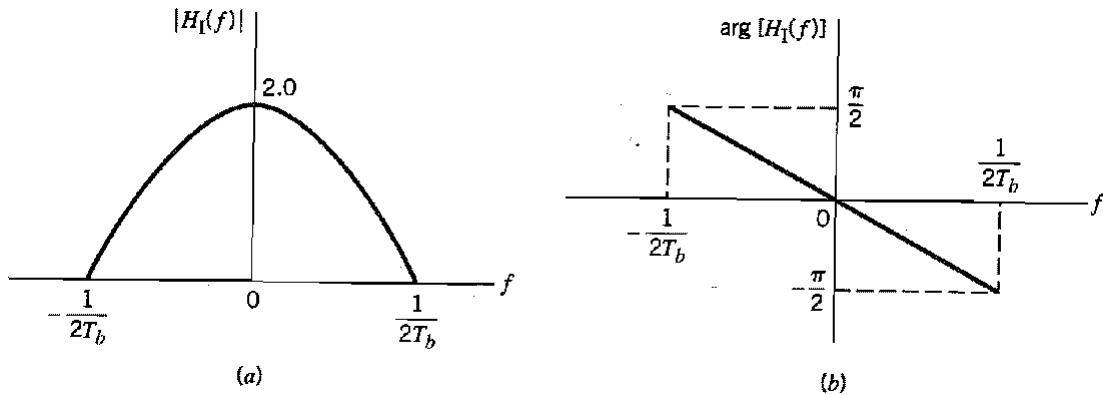


Figure 2.12 Frequency response of the duobinary conversion filter. (a) Magnitude response. (b) Phase response.

Note that unlike the linear operation of duobinary coding, the precoding described by eqn 2.72 is a nonlinear operation.

$$c_k = \begin{cases} 0 & \text{if data symbol } b_k \text{ is 1} \\ \pm 2 & \text{if data symbol } b_k \text{ is 0} \end{cases} \quad (2.72)$$

Which is illustrated in example 2.3 . From eqn 2.75 we deduce the following decision rule for detecting the original binary sequence $\{b_k\}$ from $\{c_k\}$

If $|c_k| < 1$, say symbol b_k is 1
 If $|c_k| > 1$, say symbol b_k is 0

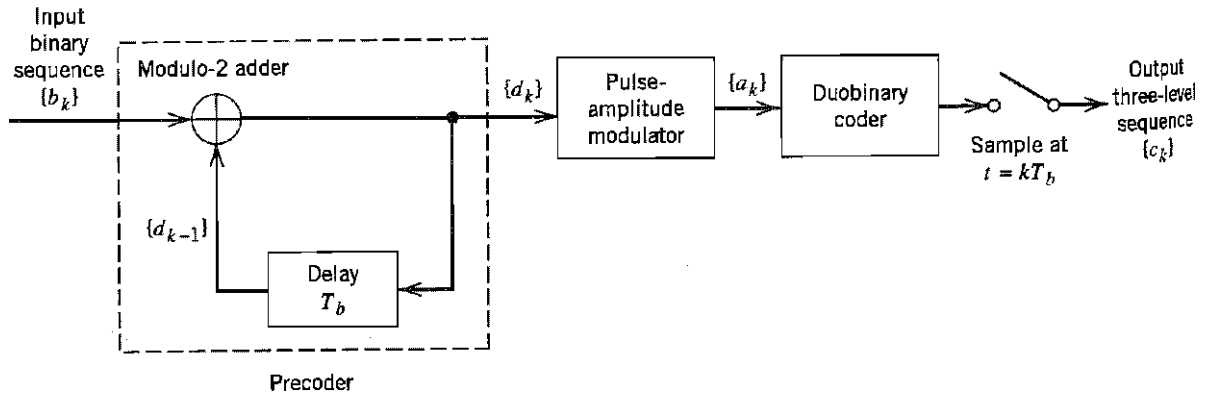


Figure 2.14 A precoded duobinary scheme

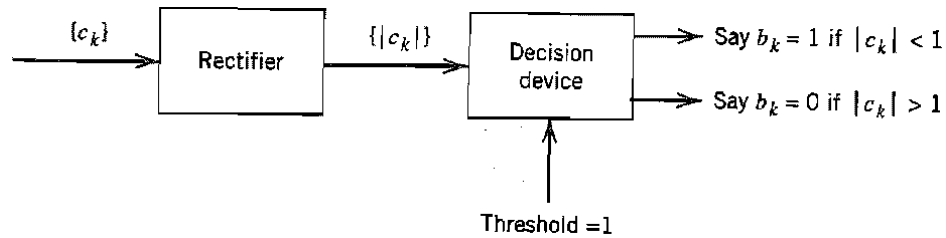


Figure 2.15 Detector for recovering original binary sequence from the precoded duobinary coder output.

When $|c_k| = 1$, the receiver simply makes a random guess in favor of symbol 1 or 0. According to this decision rule, the detector consists of a rectifier, the output of which is compared in a decision device to a threshold of 1. A block diagram of the detector is shown in Figure 2.15. A useful feature of this detector is that no knowledge of any input sample other than the present one is required. Hence, error propagation cannot occur in the detector of Figure 2.15.

Modified Duobinary Signaling

In the duobinary signaling technique the frequency response $H(f)$, and consequently the power spectral density of the transmitted pulse, is nonzero at the origin. This is considered to be an undesirable feature in some applications, since many communications channels cannot transmit a DC component. We may correct for this deficiency by using the class IV partial response or modified duobinary technique, which involves a correlation span of two binary digits. This special form of correlation is achieved by subtracting amplitude modulated pulses spaced T_b seconds apart, as indicated in the block diagram of Figure 2.16

TABLE 2. 1 Illustrating Example 4.3 on duobinary coding

Binary sequence $\{b_k\}$		0	0	1	0	1	1	0
Precoded sequence $\{d_k\}$	1	1	1	0	0	1	0	0
Two-level sequence $\{a_k\}$	+1	+1	+1	-1	-1	+1	-1	-1
Duobinary coder output $\{c_k\}$		+2	+2	0	-2	0	0	-2
Binary sequence obtained by applying decision rule of Eq. (4.76)		0	0	1	0	1	1	0

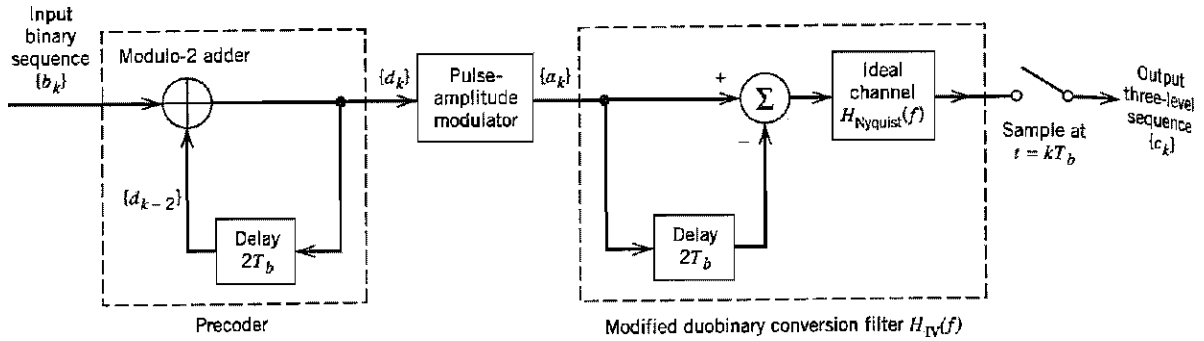


Figure 2.16 Modified duobinary signaling scheme.

2.6 Baseband M-ary PAM transmission

In the baseband binary PAM system of Figure 2.7, the pulse-amplitude modulator produces binary pulses, that is, pulses with one of two possible amplitude levels. On the other hand, in a baseband M-ary PAM system, the pulse-amplitude modulator produces one of M possible amplitude levels with $M > 2$. This form of pulse modulation is illustrated in Figure 2.20(a) for the case of a quaternary ($M = 4$) system and the binary data sequence 0010110111. The waveform shown in Figure 2.20a is based on the electrical representation for each of the four possible dibits (pairs of bits) given in Figure 4.20(b). Note that this representation is Gray encoded, which means that any dibit in the quaternary alphabet differs from an adjacent dibit in a single bit position.

In an M-ary system, the information source emits a sequence of symbols from an alphabet that consists of M symbols. Each amplitude level at the pulse-amplitude modulator output corresponds to a distinct symbol, so that there are M distinct amplitude levels to be transmitted. Consider then an M-ary PAM system with a signal alphabet that contains M equally likely and statistically independent symbols, with the symbol duration denoted by T seconds. We refer to 1/T as the signaling rate of the system, which is expressed in symbols per second, or bauds. It is informative to relate the signaling rate of this system to that of an equivalent binary PAM system for which the value of M is 2 and the successive binary symbols 1 and 0 are equally likely and statistically independent, with the duration of either symbol denoted by T_b seconds. Under the conditions described here, the binary PAM system produces information at the rate of $1/T_b$ bits per second. We also observe that in the case of a quaternary PAM system, for example, the four possible symbols may be identified with the dibits 00, 01, 10, and 11. We thus see that each symbol represents 2 bits of information, and 1 baud is equal to 2 bits per second. We may generalize this result by stating that in an M-ary PAM system, 1 baud is equal to $\log_2 M$ bits per second, and the symbol duration T of the M-ary PAM system is related to the bit duration T_b of the equivalent binary PAM system as

$$T = T_b \log_2 M \quad (2.84)$$

Therefore, in a given channel bandwidth, we find that by using an M-ary PAM system, we are able to transmit information at a rate that is $\log_2 M$ faster than the corresponding binary PAM system. However, to realize the same average probability of symbol error, an M-ary PAM system requires more transmitted power. Specifically, we find that for M much larger than 2 and an average probability of symbol error small compared to 1, the transmitted power must be increased by the factor $M / \log_2 M$, compared to a binary PAM system.

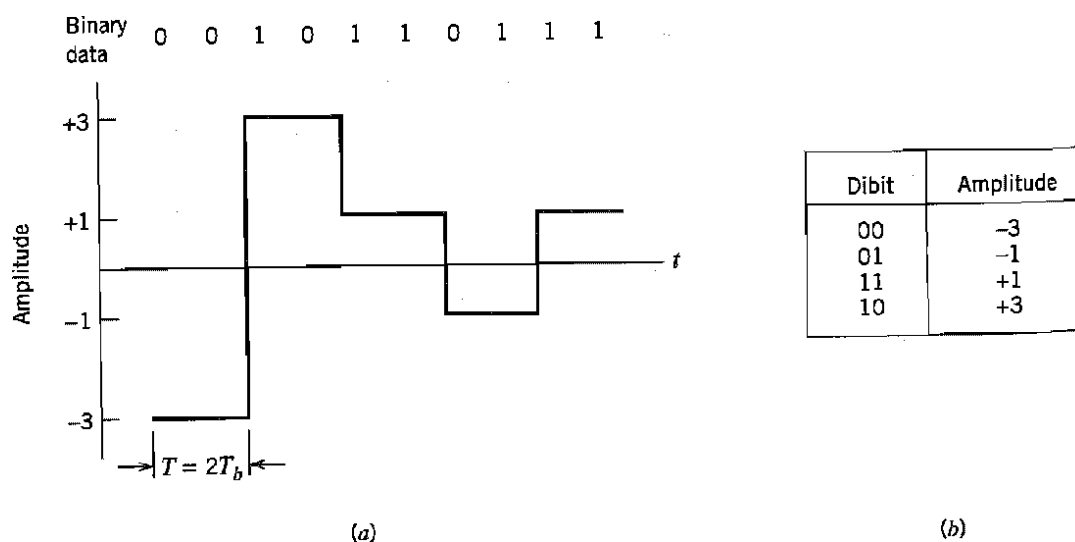


Figure 2.20 Output of a quaternary system. (a) Waveform.
(b) Representation of the 4 possible dibits, based on gray encoding.

In a baseband M-ary system, first of all, the sequence of symbols emitted by the information source is converted into an M-level PAM pulse train by a pulse-amplitude modulator at the transmitter input. Next, as with the binary PAM system, this pulse train is shaped by a transmit filter and then transmitted over the communication channel, which corrupts the signal waveform with both noise and distortion. The received signal is passed through a receive filter and then sampled at an appropriate rate in synchronism with the transmitter. Each sample is compared with preset threshold values (also called slicing levels), and a decision is made as to which symbol was transmitted. We therefore find that the designs of the pulse-amplitude modulator and the decision-making device in an M-ary PAM are more complex than those in a binary PAM system. Intersymbol interference, noise, and imperfect synchronization cause errors to appear at the receiver output. The transmit and receive filters are designed to minimize these errors. Procedures used for the design of these filters are similar to those discussed in Sections 4.5 and 4.6 for baseband binary PAM systems.

2.7 Adaptive Equalization

In this section we develop a simple and yet effective algorithm for the adaptive equalization of a linear channel of unknown characteristics. Figure 2.28 shows the structure of an adaptive synchronous equalizer, which incorporates the matched filtering action. The algorithm used to adjust the equalizer coefficients assumes the availability of a desired response. One's first reaction to the availability of a replica of the transmitted signal is: If such a signal is available at the receiver, why do we need adaptive equalization? To answer this question, we first note that a typical telephone channel changes little during an average data call. Accordingly, prior to data transmission, the equalizer is adjusted under the guidance

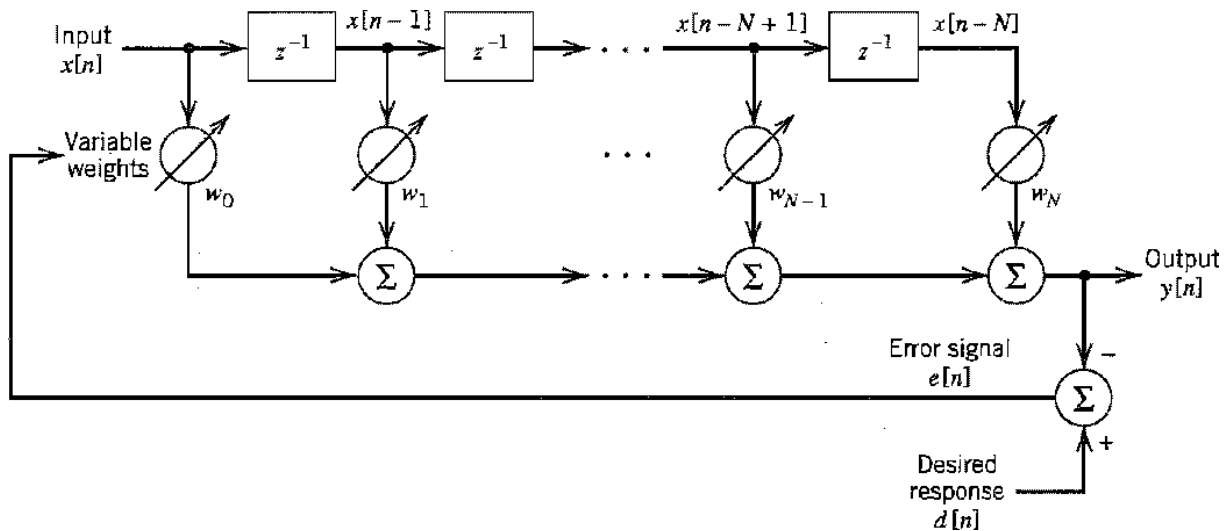


Figure 2.28 Block diagram of adaptive equalizer.

of a training sequence transmitted through the channel. A synchronized version of this training sequence is generated at the receiver, where (after a time shift equal to the transmission delay through the channel) it is applied to the equalizer as the desired response. A training sequence commonly used in practice is the pseudonoise (PN) sequence, which consists of a deterministic periodic sequence with noise-like characteristics. Two identical PN sequence generators are used, one at the transmitter and the other at the receiver. When the training process is completed, the PN sequence generator is switched off, and the adaptive equalizer is ready for normal data transmission.

Least-Mean-Square Algorithm

To simplify notational matters, we let

$$\begin{aligned}x[n] &= x(nT) \\ y[n] &= y(nT)\end{aligned}$$

Then, the output $y[n]$ of the tapped-delay-line equalizer in response to the input sequence $\{x[n]\}$ is defined by the discrete convolution sum (see Figure 4.28)

$$y[n] = \sum_{k=0}^N w_k x[n - k] \quad (4.112)$$

where w_k is the weight at the k th tap, and $N + 1$ is the total number of taps. The tap-weights constitute the adaptive filter coefficients. We assume that the input sequence $\{x[n]\}$ has finite energy. We have used a notation for the equalizer weights in Figure 4.28 that is different from the corresponding notation in Figure 4.27 to emphasize the fact that the equalizer in Figure 4.28 also incorporates matched filtering.

Let $a[n]$ denote the desired response defined as the polar representation of the transmitted binary symbol. Let $e[n]$ denote the error signal defined as the difference between the desired response $a[n]$ and the actual response $y[n]$ of the equalizer, as shown in

$$e[n] = a[n] - y[n]$$

In the least-mean-square (LMS) algorithm for adaptive equalization, the error signal* actuates the adjustments applied to the individual tap weights of the equalizer as algorithm proceeds from one iteration to the next. A derivation of the LMS algorithm adaptive prediction was presented in Section 3.13. Recasting Equation (3.72) into its most general form, we may state the formula for the LMS algorithm in words as follows:

$$\begin{pmatrix} \text{Updated value} \\ \text{of } k\text{th tap-} \\ \text{weight} \end{pmatrix} = \begin{pmatrix} \text{Old value} \\ \text{of } k\text{th tap-} \\ \text{weight} \end{pmatrix} + \begin{pmatrix} \text{Step-size} \\ \text{parameter} \end{pmatrix} \cdot \begin{pmatrix} \text{Input signal} \\ \text{applied to} \\ k\text{th tap-} \\ \text{weight} \end{pmatrix} \begin{pmatrix} \text{Error} \\ \text{signal} \end{pmatrix}$$

2.114

now summarize the LMS algorithm for adaptive equalization as follows:

1. Initialize the algorithm by setting $\hat{\mathbf{w}}[1] = 0$ (i.e., set all the tap-weights of the equalizer to zero at $n = 1$, which corresponds to time $t = T$).
2. For $n = 1, 2, \dots$, compute

$$\begin{aligned} y[n] &= \mathbf{x}^T[n] \hat{\mathbf{w}}[n] \\ e[n] &= a[n] - y[n] \\ \hat{\mathbf{w}}[n + 1] &= \hat{\mathbf{w}}[n] + \mu e[n] \mathbf{x}[n] \end{aligned}$$

where μ is the step-size parameter.

3. Continue the iterative computation until the equalizer reaches a “steady state,” by which we mean that the actual mean-square error of the equalizer essentially reaches a constant value.

The LMS algorithm is an example of a feedback system, as illustrated in the block diagram of Figure 4.29, which pertains to the k th filter coefficient. It is therefore possible

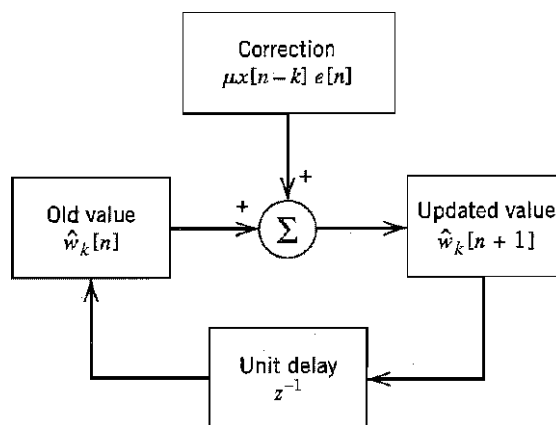


Figure 2.29 Signal flow graph representation of the LMS algorithm involving the k th tap weight.

■ OPERATION OF THE EQUALIZER

There are two modes of operation for an adaptive equalizer, namely, the training mode and decision-directed mode, as shown in Figure 4.30. During the *training mode*, as explained previously, a known PN sequence is transmitted and a synchronized version of it is generated in the receiver, where (after a time shift equal to the transmission delay) it is applied to the adaptive equalizer as the desired response; the tap-weights of the equalizer are thereby adjusted in accordance with the LMS algorithm.

When the training process is completed, the adaptive equalizer is switched to its second mode of operation: the *decision-directed mode*. In this mode of operation, the error signal is defined by

$$e[n] = \hat{a}[n] - y[n] \quad (4.120)$$

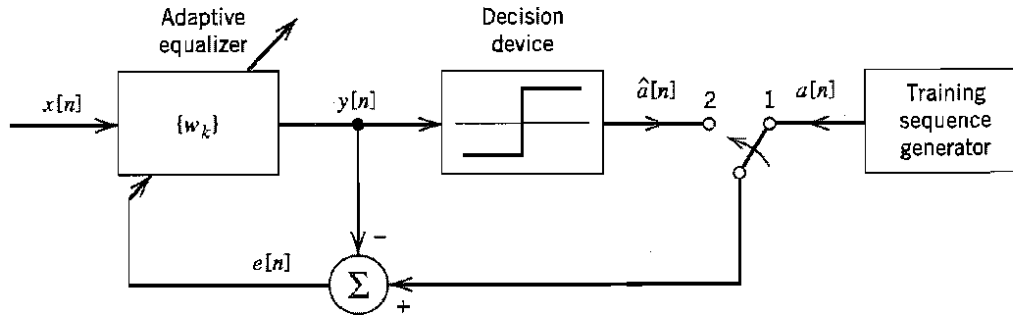


Figure 2.30 Illustrating the two operating modes of an adaptive equalizer, for the training mode, the switch is in position 1; and for the tracking mode, it is moved to position 2.

2.8 Eye patterns

In previous sections of this chapter we have discussed various techniques for dealing with the effects of channel noise and intersymbol interference on the performance of a baseband pulse-transmission system. In the final analysis, what really matters is how to evaluate the combined effect of these impairments on overall system performance in an operational environment. An experimental tool for such an evaluation in an insightful manner is the so-called eye pattern, which is defined as the synchronized superposition of all possible realizations of the signal of interest (e.g., received signal, receiver output) viewed within a particular signaling interval. The eye pattern derives its name from the fact that it resembles the human eye for binary waves. The interior region of the eye pattern is called the eye opening.

An eye pattern provides a great deal of useful information about the performance of a data transmission system, as described in Figure 2.33. Specifically, we make the following statements:

► The width of the eye opening defines the time interval over which the received signal can be sampled without error from intersymbol interference; it is apparent that the preferred time for sampling is the instant of time at which the eye is open the widest.

*■ The sensitivity of the system to timing errors is determined by the rate of closure of the eye as the sampling time is varied.

> The height of the eye opening, at a specified sampling time, defines the noise margin of the system.

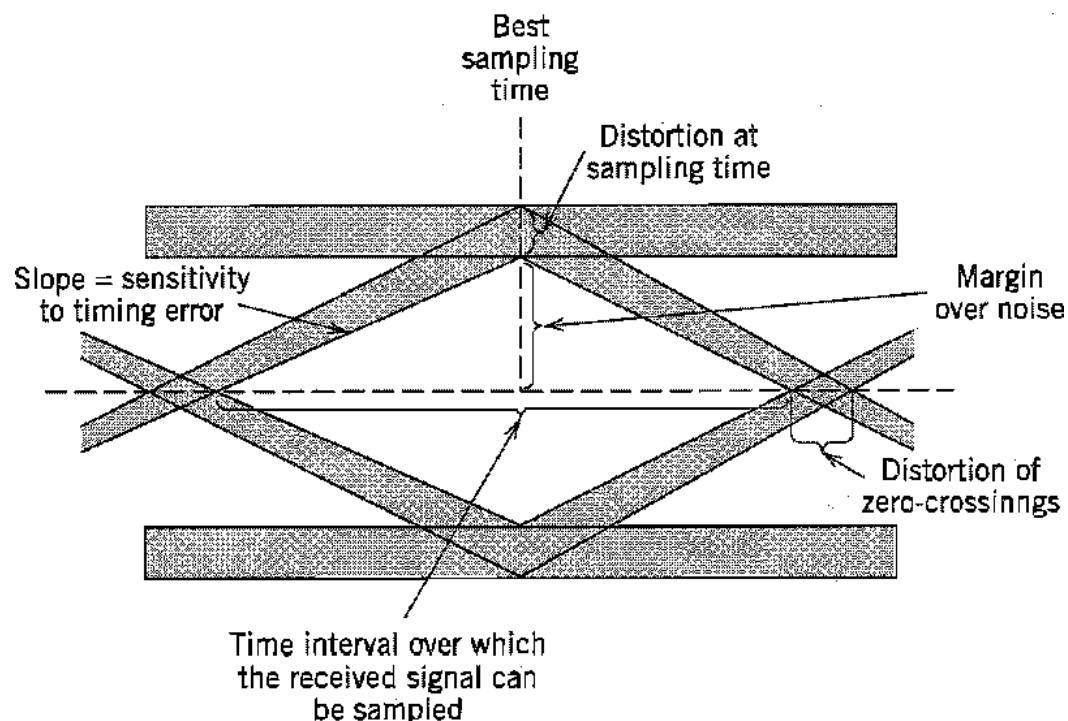
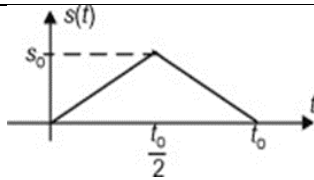


Figure 2.33 Interpretation of the eye pattern.

QUESTION BANK UNIT II

PART-A

1. State Nyquist's criterion for distortion less transmission of baseband data.
2. Outline the causes for inter symbol interference.
3. Give any two properties of Matched filter.
4. Plot the impulse response of the matched filter for the input pulse given.



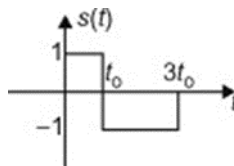
5. Distinguish matched filter receiver and correlation receiver.
6. Summarize the need of for channel Line Coding.
7. Draw the Manchester code format for the data 1101011.
8. For the binary data 10110001, show the unipolar-RZ format representation.
9. Outline the role of equalizers in digital communication receivers.
10. Narrate the conditions that should be satisfied for reducing inter symbol interference.

PART-B

1. Draw the time domain representation of the binary code '11100110' for following line-encoding schemes and outline the rules for each scheme. (16 marks)

- i. Unipolar NRZ-L, NRZ-M, NRZ-S
- ii. Bipolar NRZ-L, NRZ-M, NRZ-S
- iii. Unipolar Return to zero, Bipolar Return to zero
- iv. Return to Zero AMI
- v. Biphas-L, Biphas-M and Biphas-S

2. a) Determine the impulse response of the filter matched to the input given below. Assume that the filter is designed to maximize the output SNR at time $t = 3t_0$. Also, plot the output of the matched filter. (8 marks)



b) Derive the expression for probability of error of a matched filter for a bipolar NRZ data. (8 marks)

3. a) Discuss in detail about inter symbol interference (ISI) and the Nyquist criterion for minimizing ISI. (8 marks)

b) Describe how correlative coding could eliminate ISI. (8 marks)

4. a) A binary PAM wave is to be transmitted over a low-pass channel with bandwidth of 75 kHz. The bit duration is $10\mu\text{s}$. Find a raised-cosine pulse spectrum that satisfies these requirements. (8 marks)

b) Illustrate the principle of duo-binary encoding scheme with an example. (8 marks)

5. a) Describe the realisation of a non-uniform quantizer using a compressor and a uniform quantizer and plot its input and output characteristics. (10 marks)

b) Distinguish A-law and μ -law companding.

(6 marks)

6. a) Elucidate the structure of eye pattern and explain how would you analyse the level of ISI degradation, dispersion and jitter from the eye diagram. (12 marks)

b. Explain how equalizers can be used to undo the distortions introduced in the channel.

(4 marks)

UNIT III – DIGITAL MODULATION TECHNIQUES - SEC1313

UNIT 3 DIGITAL MODULATION TECHNIQUES

Introduction - ASK- FSK - PSK- coherent modulation techniques-BFSK-BPSK-signal space diagram-probability of error-Coherent Quadrature modulation techniques- QPSK-signal space diagram-probability of error- Non coherent modulation techniques-M-ary modulation techniques - Vectorial view of MPSK and MFSK - error performance.

DIGITAL MODULATION TECHNIQUES

Given a binary source that emits symbols 0 and 1, the modulation process involves switching or keying the amplitude, phase, or frequency of a sinusoidal carrier wave between a pair of possible values in accordance with symbols 0 and 1. To be more specific, consider the sinusoidal carrier

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

------(1)

Where A_c is the carrier amplitude, f_c is the carrier frequency, and ϕ_c is the carrier phase. Given these three parameters of the carrier(t), we may now identify three distinct forms of binary modulation:

1. Binary amplitude shift-keying (BASK), in which the carrier frequency and carrier phase are both maintained constant, while the carrier amplitude is keyed between the two possible values used to represent symbols 0 and 1.
2. Binary phase-shift keying (BPSK), in which the carrier amplitude and carrier frequency are both maintained constant, while the carrier phase is keyed between the two possible values (e.g., 0° and 180°) used to represent symbols 0 and 1.
3. Binary frequency-shift keying (BFSK), in which the carrier amplitude and carrier phase are both maintained constant, while the carrier frequency is keyed between the two possible values used to represent symbols 0 and 1.

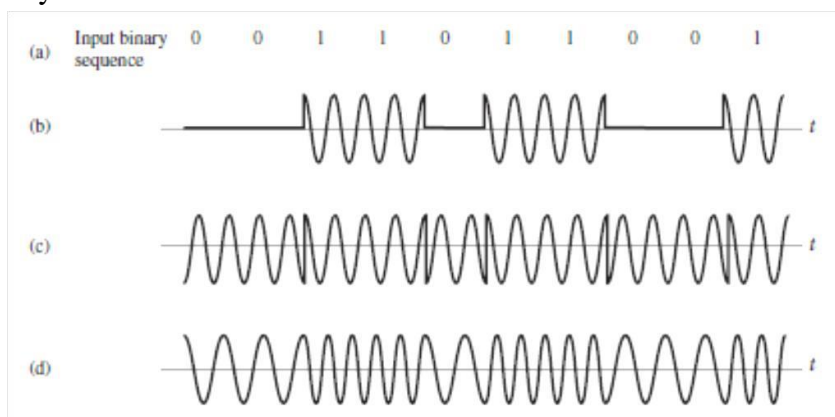


Fig. 3 Types of Binary signaling Information

The three basic forms of signaling binary information. (a) Binary data stream. (b) Amplitude-shift keying. (c) Phase-shift keying. (d) Frequency-shift keying with continuous phase. In the digital communications literature, the usual practice is to assume that the carrier $c(t)$ has unit energy measured over one symbol (bit) duration.

$$A_c = \sqrt{\frac{2}{T_b}}$$

------(2)

Where T_b is the bit duration. Using the terminology of Eq. (2), we may thus express the Carrier $c(t)$ in the equivalent form

$$c(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t + \phi_c) \quad \text{-----}(3)$$

The spectrum of a digitally modulated wave, exemplified by BASK, BPSK and BFSK, is centered on the carrier frequency f_c implicitly or explicitly. Moreover, as with analog modulation, it is normal practice to assume that the carrier frequency f_c is large compared with the “bandwidth” of the incoming binary data stream that acts as the modulating signal. This band-pass assumption has certain implications, as discussed next. To be specific, consider a linear modulation scheme for which the modulated wave is defined by

$$s(t) = b(t)c(t) \quad \text{-----}(4)$$

where $b(t)$ denotes an incoming binary wave. Then, setting the carrier phase ϕ_c for convenience of presentation, we may use Eq. (3) to express the modulated wave as

$$s(t) = \sqrt{\frac{2}{T_b}} b(t) \cos(2\pi f_c t) \quad \text{-----}(5)$$

Under the assumption where $f_c \gg w$ is the bandwidth of the binary wave there will be no spectral overlap in the generation of $s(t)$ (i.e., the spectral content of the modulated wave for positive frequencies is essentially separated from its spectral content for negative frequencies). Another implication of the band-pass assumption is that we may express the transmitted signal energy per bit as

$$\begin{aligned} E_b &= \int_0^{T_b} |s(t)|^2 dt \\ &= \frac{2}{T_b} \int_0^{T_b} |b(t)|^2 \cos^2(2\pi f_c t) dt \end{aligned} \quad \text{-----}(6)$$

Using the trigonometric identity

$$\cos^2(2\pi f_c t) = \frac{1}{2} [1 + \cos(4\pi f_c t)]$$

we may rewrite Eq. (6) as

$$E_b = \frac{1}{T_b} \int_0^{T_b} |b(t)|^2 dt + \frac{1}{T_b} \int_0^{T_b} |b(t)|^2 \cos(4\pi f_c t) dt \quad \text{-----}(7)$$

The band-pass assumption implies that $|b(t)|^2$ is essentially constant over one complete cycle of the sinusoidal wave $\cos(4\pi f_c t)$ which, in turn, means that

$$\int_0^{T_b} |b(t)|^2 \cos(4\pi f_c t) dt \approx 0$$

Accordingly, we may approximate Eq. (7) as

$$E_b \approx \frac{1}{T_b} \int_0^{T_b} |b(t)|^2 dt \quad \text{-----}(8)$$

In words, for linear digital modulation schemes governed by Eq. (5), the transmitted signal energy (on a per bit basis) is a scaled version of the energy in the incoming binary wave responsible for modulating the sinusoidal carrier.

BINARY AMPLITUDE-SHIFT KEYING

Binary amplitude-shift keying (BASK) is one of the earliest forms of digital modulation used in radio telegraphy at the beginning of the twentieth century. To formally describe BASK, consider a binary data stream $b(t)$ which is of the ON–OFF signaling variety.

$$b(t) = \begin{cases} \sqrt{E_b}, & \text{for binary symbol 1} \\ 0, & \text{for binary symbol 0} \end{cases}$$

That is, $b(t)$ is defined by

Then, multiplying $b(t)$ by the sinusoidal carrier wave of Eq. (3) with the phase set ϕ_c equal to zero for convenience of presentation, we get the BASK wave

$$s(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), & \text{for symbol 1} \\ 0, & \text{for symbol 0} \end{cases} \quad \text{-----}(9)$$

The carrier frequency f_c may have an arbitrary value, consistent with transmitting the modulated signal anywhere in the electromagnetic radio spectrum, so long as it satisfies the band-pass assumption.

When a bit duration is occupied by symbol 1, the transmitted signal energy E_b is When the bit duration is occupied by symbol 0, the transmitted signal energy is zero. On this basis, we may express the average transmitted signal energy as

$$E_{av} = \frac{E_b}{2}$$

------(10)

GENERATION AND DETECTION OF ASK SIGNALS

From Eqs. (8) and (9), we readily see that a BASK signal is readily generated by using a product modulator with two inputs. One input, the ON–OFF signal of Eq. (8), is the modulating signal. The sinusoidal carrier wave supplies the other input.

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

------(11)

SPECTRAL ANALYSIS OF BASK

Consider a binary data stream that consists of a square wave, the amplitude of which alternates between the constant levels ± 1 and zero every T_b seconds. The square wave is centered on the origin for convenience of the presentation. The objective of the experiment is twofold:

- (i) To investigate the effect of varying the carrier frequency on the power spectrum of the BASK signal assuming that the square wave is fixed. Recall that the power spectrum of a signal (expressed in decibels) is defined as 10 times the logarithm (to base 10) of the squared magnitude (amplitude) spectrum of the signal.
- (ii) To investigate the effect of varying the frequency of the square wave on the spectrum of the BASK signal, assuming that the sinusoidal carrier wave is fixed. For the purpose of computer evaluation, we set the carrier frequency $f_c = n/T_b$ where n is an integer. This choice of the carrier frequency f_c permits the simulation of a band-pass system on a digital computer without requiring $f_c \gg 1/T_b$ the only restriction on the choice is to make sure that spectral overlap is avoided. (We follow this practice when performing computer experiments as we go forward in the study of other digital modulation schemes.)

BINARY PHASE-SHIFT KEYING (BPSK)

In the simplest form of phase-shift keying known as binary phase-shift keying (BPSK), the pair of signals $s_1(t)$ and $s_2(t)$ used to represent symbols 1 and 0, respectively, are defined by

$$s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), & \text{for symbol 1 corresponding to } i = 1 \\ \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t + \pi) = -\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), & \text{for symbol 0 corresponding to } i = 2 \end{cases}$$

------(11)

Where T_b denoting the bit duration E_b and denoting the transmitted signal energy per bit; see the waveform of Fig. 7.1 (c) for a representation example of BPSK. A pair of sinusoidal waves, $s_1(t)$ and $s_2(t)$ which differ only in a relative phase-shift of π radians as defined in Eq. (7.12), are referred to as antipodal signals. BPSK differs from BASK in an important respect: the envelope of the modulated signal $s(t)$ is maintained constant at the value

for all time t . This property, which follows directly from Eq. (7.12), has two important consequences:

1. The transmitted energy per bit, E_b is constant; equivalently, the average transmitted power is constant.
2. Demodulation of BPSK cannot be performed using envelope detection; rather, we have to look to coherent detection as described next.

GENERATION AND COHERENT DETECTION OF BPSK SIGNALS

(i) Generation

To generate the BPSK signal, we build on the fact that the BPSK signal is a special case of DSB-SC modulation. Specifically, we use a product modulator consisting of two components

- (i) Non-return-to-zero level encoder, whereby the input binary data sequence is encoded in polar form with symbols 1 and 0 represented by the constant-amplitude levels: and , respectively.

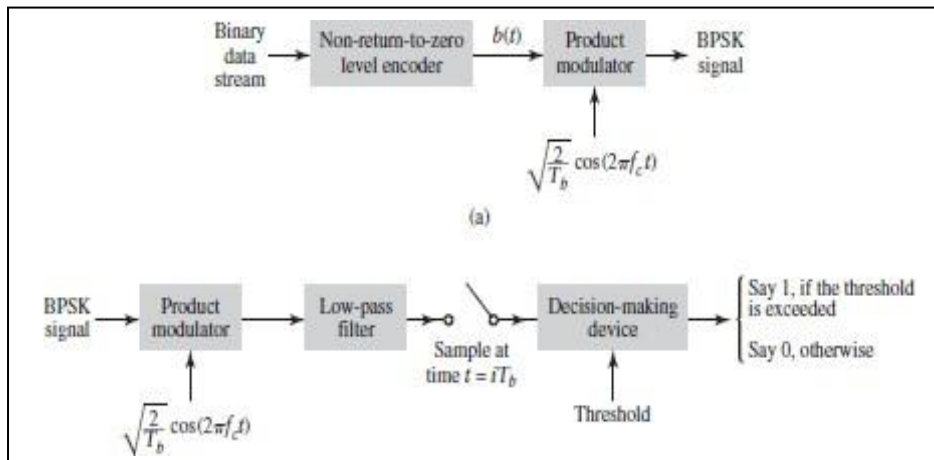


Fig 3.1 (a) BPSK modulator. (b) Coherent detector for BPSK; for the sampler, integer $i = 0, 1, 2, \dots$

- (ii) Product modulator, which multiplies the level-encoded binary wave by the sinusoidal carrier $c(t)$ of amplitude to produce the BPSK signal. The timing pulses used to generate the level-encoded binary wave and the sinusoidal carrier wave are usually, but not necessarily, extracted from a common master clock.

(ii) Detection

To detect the original binary sequence of 1s and 0s, the BPSK signal $x(t)$ at the channel output is applied to a receiver that consists of four sections, as depicted in Fig. 3.1(b)

- (i) Product modulator, which is also supplied with a locally generated reference signal that is a replica of the carrier wave $c(t)$.

(ii) Low-pass filter, designed to remove the double-frequency components of the product modulator output and pass the zero-frequency components.

(iii) Sampler, which uniformly samples the output of the low-pass filter at $t=iT_b$ where $i=0,1,2,\dots$; the local clock governing the operation of the sampler is synchronized with the clock responsible for bit-timing in the transmitter.

(iv) Decision-making device, which compares the sampled value of the low-pass filter's output to an externally supplied threshold, every T_b seconds. If the threshold is exceeded, the device decides in favor of symbol 1; otherwise, it decides in favour of symbol 0.

The BPSK receiver described in Fig. 3.1 is said to be coherent in the sense that the sinusoidal reference signal applied to the product modulator in the demodulator is synchronous in phase (and, of course, frequency) with the carrier wave used in the modulator. In addition to synchrony with respect to carrier phase, the receiver also has an accurate knowledge of the interval occupied by each binary symbol. The operation of the coherent BPSK receiver in Fig. 3.4(b) follows a procedure similar to that described for the demodulation of a double-sideband suppressed-carrier (DSBSC) modulated wave) with a couple of important additions: sampler and decision-making device. The rationale for this similarity builds on what we have already stated: BPSK is simply another form of DSB-SC modulation. However, an issue that needs particular attention is how to design the low-pass filter in Fig. 3.4(b)

SPECTRAL ANALYSIS OF BPSK

As with the experiment on BASK, consider a binary data stream that consists of a square wave, the amplitude of which alternates between $\pm A_m$ every T_b seconds. The square wave is centered on the origin. The objectives of this second experiment are similar to those of Computer Experiment I on BASK:

(i) To evaluate the effect of varying the carrier frequency on the power spectrum of the BPSK signal, for a fixed square modulating wave.

(ii) To evaluate the effect of varying modulation frequency on the power spectrum of the BPSK signal for a fixed carrier frequency.

BASK and BPSK signals occupy the same transmission bandwidth—namely, $2/T_b$ —which defines the width of the main lobe of the sinc-shaped power spectra. The BASK spectrum includes a carrier component, whereas this component is absent from the BPSK spectrum. With this observation we are merely restating the fact that BASK is an example of amplitude modulation, whereas BPSK is an example of double sideband-suppressed carrier modulation.

QUADRIPHASE-SHIFT KEYING

An important goal of digital communication is the efficient utilization of channel bandwidth. This goal is attained by a bandwidth-conserving modulation scheme known as

quadrature-shift keying.

In quadrature-shift keying (QPSK), as with BPSK, information carried by the transmitted signal is contained in the phase of a sinusoidal carrier. In particular, the phase of the sinusoidal carrier takes on one of four equally spaced values, such as $\pi/4, 3\pi/4, 5\pi/4$, and $7\pi/4$. For this set of values, we define the transmitted signal as

$$s_i(t) = \begin{cases} \sqrt{\frac{2E}{T}} \cos\left[2\pi f_c t + (2i - 1)\frac{\pi}{4}\right], & 0 \leq t \leq T \\ 0, & \text{elsewhere} \end{cases} \quad \text{-----}(12)$$

Where $i=1,2,3,4$; E is the transmitted signal energy per symbol and T is the symbol duration. Each one of the four equally spaced phase values corresponds to a unique pair of bits called a dibit. For example, we may choose the foregoing set of phase values to represent the Gray encoded set of dibits: 10, 00, 01, and 11. In this form of encoding, we see that only a single bit is changed from one dibit to the next. Note that the symbol duration (i.e., the duration of each dibit) is twice the bit duration, as shown by

$$T = 2T_b \quad \text{-----}(13)$$

Using a well-known trigonometric identity, we may recast the transmitted signal in the interval in the expanded form

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left[(2i - 1)\frac{\pi}{4}\right] \cos(2\pi f_c t) - \sqrt{\frac{2E}{T}} \sin\left[(2i - 1)\frac{\pi}{4}\right] \sin(2\pi f_c t) \quad \text{--}(14)$$

where $i=1,2,3,4$. Based on the expanded form of Eq. (14), we can make some important observations:

1. In reality, the QPSK signal consists of the sum of two BPSK signals.
2. One BPSK signal, represented by the first term

$$\sqrt{2E/T} \cos[(2i - 1)\pi/4] \cos[(2\pi f_c t)],$$

defines the product of modulating a binary wave by the sinusoidal carrier π which has unit energy over the symbol duration T . We also recognize that

$$\sqrt{E} \cos\left[(2i - 1)\frac{\pi}{4}\right] = \begin{cases} \sqrt{E/2} & \text{for } i = 1, 4 \\ -\sqrt{E/2} & \text{for } i = 2, 3 \end{cases} \quad \text{-----}(15)$$

We therefore see that this binary wave has an amplitude equal to

3. The other BPSK signal, represented by the second term

$$-\sqrt{2E/T} \sin\left[(2i - 1)\frac{\pi}{4}\right] \sin(2\pi f_c t),$$

defines the product of modulating a different binary wave by the sinusoidal carrier π , which also has unit energy per symbol. This time, we recognize that

$$-\sqrt{E} \sin\left[(2i-1)\frac{\pi}{4}\right] = \begin{cases} -\sqrt{E/2} & \text{for } i = 1, 2 \\ \sqrt{E/2} & \text{for } i = 3, 4 \end{cases} \text{-----}(16)$$

We therefore see that this second binary wave also has an amplitude equal to, albeit in a different way with respect to the index i .

5. The two binary waves defined in Eqs. (15) and (16) share a common value for the symbol duration—namely, T .

6. The two sinusoidal carrier waves identified under points 2 and 3 are in phase quadrature with respect to each other. Moreover, they both have unit energy per symbol duration. We may therefore state that these two carrier waves constitute an orthonormal pair of basis functions.

6. For each possible value of the index i , Eqs. (15) and (16) identify the corresponding dibit, as outlined in Table 1. This table also includes other related entries pertaining to the phase of the QPSK signal, and the amplitudes of the two binary waves identified under points 2 and 3.

Table:3.1 Relation between index i and identity of corresponding dibit and other related matters

Index i	Phase of QPSK signal (radians)	Amplitudes of constituent binary waves		Input dibit $0 \leq t < T$
		Binary wave 1 $a_1(t)$	Binary wave 2 $a_2(t)$	
1	$\pi/4$	$+\sqrt{E/2}$	$-\sqrt{E/2}$	10
2	$3\pi/4$	$-\sqrt{E/2}$	$-\sqrt{E/2}$	00
3	$5\pi/4$	$-\sqrt{E/2}$	$+\sqrt{E/2}$	01
4	$7\pi/4$	$+\sqrt{E/2}$	$+\sqrt{E/2}$	11

GENERATION AND COHERENT DETECTION OF QPSK SIGNALS

Generation and coherent detection of QPSK signals, as described here:

(i) Generation

To generate the QPSK signal, the incoming binary data stream is first converted into polar form by a non-return-to-zero level encoder; the encoder output is denoted by $b(t)$.

Symbols 1 and 0 are thereby represented by $\pm\sqrt{E_b/2}$ where $\overline{E_b}=E/2$. The resulting binary wave is next divided by means of a demultiplexer (consisting of a serial- to-parallel converter) into two separate binary waves consisting of the odd- and even numbered input bits of $b(t)$. These two binary waves, referred to as the demultiplexed components of the input binary wave, are denoted by $a_1(t)$ and $a_2(t)$. In any signalling interval, the amplitudes of $a_1(t)$ and $a_2(t)$ are

determined in accordance with columns 3 and 4 of Table 3.1, depending on the particular dibit that is being transmitted. The demultiplexed binary waves $a_1(t)$ and $a_2(t)$ are used to modulate the pair of quadrature carriers—namely, $\pi/4, 3\pi/4, 5\pi/4$ and $7\pi/4$. Finally, the two BPSK signals are subtracted to produce the desired QPSK signals, as depicted in Fig 3.2

(a).

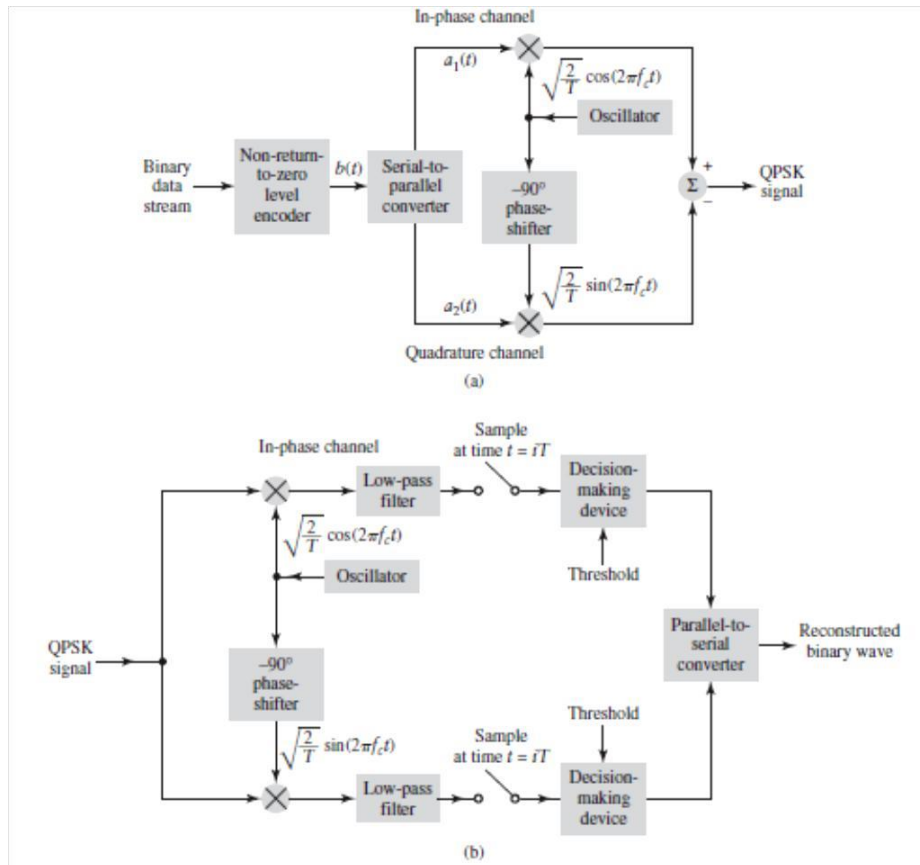


Fig 3.2 Block diagrams of (a) QPSK transmitter and (b) coherent QPSK receiver;

(ii) Detection

The QPSK receiver consists of an in-phase (I)-channel and quadrature (Q)-channel with a common input, as depicted in Fig.(b). Each channel is itself made up of a product modulator, low-pass filter, sampler, and decision-making device. Under ideal conditions,

For the two synchronous samplers, integer i = the I- and Q-channels of the receiver, respectively, recover the demultiplexed components $a_1(t)$ and $a_2(t)$ responsible for modulating the orthogonal pair of carriers in the transmitter.

Accordingly, by applying the outputs of these two channels to a multiplexer (consisting of a parallel-to-serial converter), the receiver recovers the original binary sequence. The design of the QPSK receiver builds on the strategy described for the coherent BPSK receiver. Specifically, each of the two low-pass filters in the coherent QPSK receiver of Fig.3.2 (b) must be assigned a bandwidth equal to or greater than the reciprocal of the symbol duration T for satisfactory operation of the receiver.

OFFSET QUADRIPHASE-SHIFT KEYING

In QPSK, the carrier amplitude is maintained constant. However, the carrier phase may jump by 45° for every two-bit (dibit) duration. This latter property can be of particular concern when the QPSK signal is filtered during the course of transmission over a communication channel. Unfortunately, such a filtering action can cause the carrier amplitude, and therefore the envelope

of the QPSK signal, to fluctuate. When the data transmission system contains nonlinear components, fluctuations of this kind are undesirable as they tend to distort the received signal; the net result is a reduced opening of the eye diagram.

The extent of amplitude fluctuations exhibited by QPSK signals may be reduced by using a variant of quadriphase-shift keying known as the offset quadriphase-shift keying (OQPSK).¹ In OQPSK, the demultiplexed binary wave labeled $a_2(t)$ in Fig.3.2(a) is delayed (i.e., offset) by one bit duration with respect to the other demultiplexed binary wave labelled $a_1(t)$ in that figure. This modification has the effect of confining the likely occurrence of phase transitions to 0° and 180° . However, the phase transitions in OQPSK occur twice as frequently but with a reduced range of amplitude fluctuations, compared with QPSK. In addition to the phase transitions, there are also phase transitions in QPSK. We therefore find that amplitude fluctuations in OQPSK due to filtering have a smaller amplitude than in QPSK.

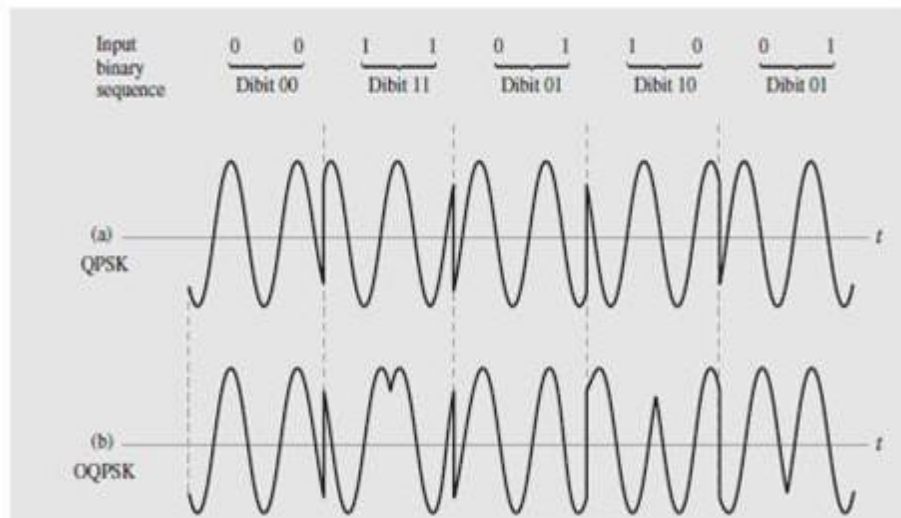


Figure 3.3 Graphical representation of phase transitions in QPSK and OQPSK

FREQUENCY-SHIFT KEYING

BINARY FREQUENCY-SHIFT KEYING (BFSK)

In the simplest form of frequency-shift keying known as binary frequency-shift keying (BFSK), symbols 0 and 1 are distinguished from each other by transmitting one of two sinusoidal waves that differ in frequency by a fixed amount. A typical pair of sinusoidal waves is described by

$$s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_1 t), & \text{for symbol 1 corresponding to } i = 1 \\ \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_2 t), & \text{for symbol 0 corresponding to } i = 2 \end{cases} \quad \text{-----(17)}$$

Where E_b is the transmitted signal energy per bit. When the frequencies f_1 and f_2 are chosen in such a way that they differ from each other by an amount equal to the reciprocal of the bit duration T_b , the BFSK signal is referred to as Sunde's BFSK after its originator. This modulated signal is a continuous-phase signal in the sense that phase continuity is always maintained, including the inter-bit switching times.

Waveform Figure plots the waveform of Sunde's BFSK produced by the input binary sequence 0011011001 for a bit duration Part (a) of the figure displays the waveform of the input sequence,

and part (b) displays the corresponding waveform of the BFSK signal. The latter part of the figure clearly displays the phase-continuous property of Sunde's BFSK.

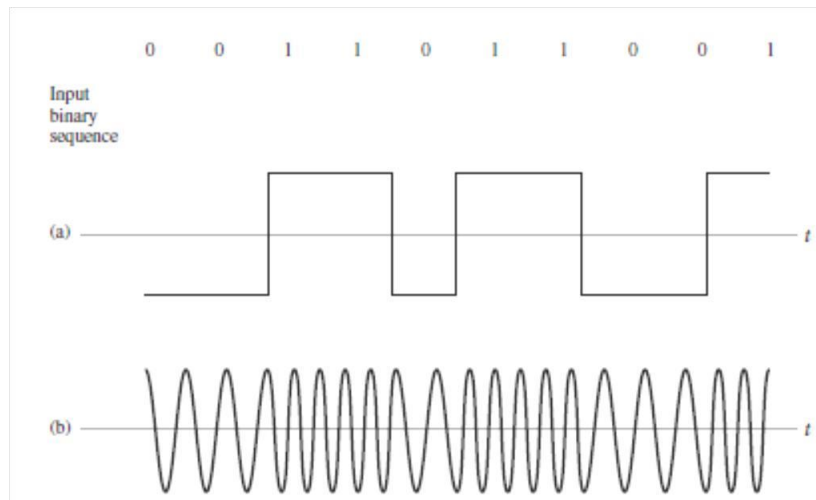


Fig. 3.4 (a) Binary sequence and its non-return-to-zero level-encoded waveform.(b) Sunde's BFSK signal.

CONTINUOUS-PHASE FREQUENCY-SHIFT KEYING

Sunde's BFSK is the simplest form of a family of digitally modulated signals known collectively as continuous-phase frequency-shift keying (CPFSK) signals, which exhibit the following distinctive property:

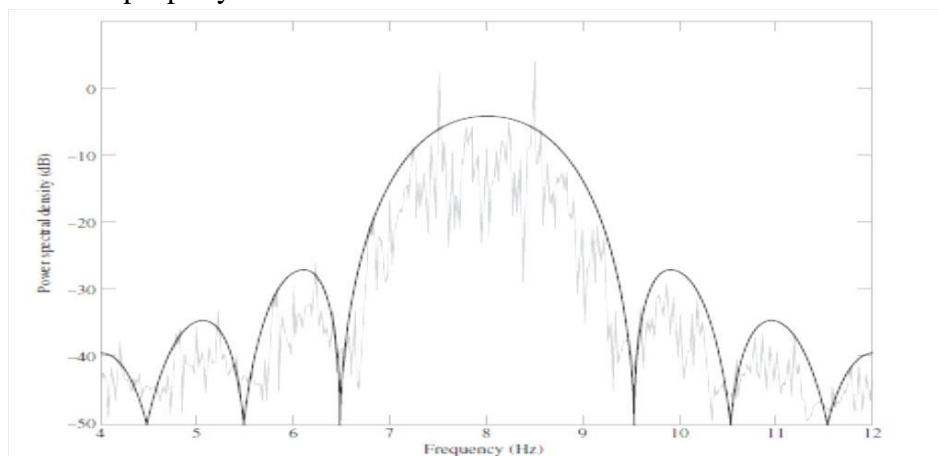


Fig.3.5 Power spectrum

Power spectrum of Sunde's BFSK produced by square wave as the modulating signal for the following parameters: $f_c = 8$ Hz and $T_b = 1$ s. In other words, the CPFSK signal is a continuous-wave modulated wave like any other angle-modulated wave experienced in the analog world, despite the fact that the modulating wave is itself discontinuous. In Sunde's BFSK, the overall excursion in the transmitted frequency from symbol 0 to symbol 1, or vice versa, is equal to the bit rate of the incoming data stream. In another special form of CPFSK known as minimum shift keying (MSK), the binary modulation process uses a different value for the frequency excursion f with the result that this new modulated wave offers superior spectral properties to Sunde's BFSK.

MINIMUM-SHIFT KEYING

In MSK, the overall frequency excursion δf from binary symbol 1 to symbol 0, or vice versa, is one half the bit rate, as shown by

$$\begin{aligned}\delta f &= f_1 - f_2 \\ &= \frac{1}{2T_b}\end{aligned}\quad \text{-----(18)}$$

The unmodulated carrier frequency is the arithmetic mean of the two transmitted frequencies f_1 and f_2 that is,

$$f_c = \frac{1}{2}(f_1 + f_2)\quad \text{-----(19)}$$

Expressing f_1 and f_2 in terms of the carrier frequency f_c and overall frequency excursion we have

$$\begin{aligned}f_1 &= f_c + \frac{\delta f}{2}, \quad \text{for symbol 1} \\ f_2 &= f_c - \frac{\delta f}{2}, \quad \text{for symbol 0}\end{aligned}\quad \text{-----(20) and (21)}$$

Accordingly, we formally define the MSK signal as the angle-modulated wave

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \theta(t)]\quad \text{-----(22)}$$

Where $\theta(t)$ is the phase of the MSK signal. In particular, when frequency f_1 is transmitted, corresponding to symbol 1, we find from Eqs. (20) and (22) that the phase assumes the value

$$\begin{aligned}\theta(t) &= 2\pi \left(\frac{\delta f}{2} \right) t \\ &= \frac{\pi t}{2T_b}, \quad \text{for symbol 1}\end{aligned}$$

In words, this means that at time $t=T_b$, the transmission of symbol 1 increases the phase of the MSK signal $s(t)$ by π radians. By the same token, when frequency f_2 is transmitted, corresponding to symbol 0, we find from Eqs. (21) and (22) that the phase assumes the value

$$\begin{aligned}\theta(t) &= 2\pi\left(-\frac{\delta f}{2}\right)t \\ &= -\frac{\pi t}{2T_b}, \quad \text{for symbol 0}\end{aligned}\quad \text{-----(24)}$$

This means that at time $t=T_b$ the transmission of symbol 0 decreases the phase of $s(t)$ by $\pi/2$ radians.

The phase changes described in Eqs. 23 and 24) for MSK are radically different from their corresponding counterparts for Sunde's BFSK. Specifically, the phase of Sunde's BFSK signal changes by $-\pi$ radians at the termination of the interval representing symbol 0, and $+\pi$ radians for symbol 1. However, the changes $-\pi$ and $+\pi$ are exactly the same, modulo 2π . This observation, in effect, means that Sunde's BFSK has no memory; in other words, knowing which particular change occurred in the previous bit interval provides no help in the current bit interval. In contrast, we have a completely different situation in the case of MSK by virtue of the different ways in which the transmissions of symbols 1 and 0 affect the phase as shown in Eqs. (23) and (24). Note also that the overall frequency excursion in MSK is the minimum frequency spacing between symbols 0 and 1 that allows their FSK representations to be coherently orthogonal, hence the terminology "minimum-shift keying."

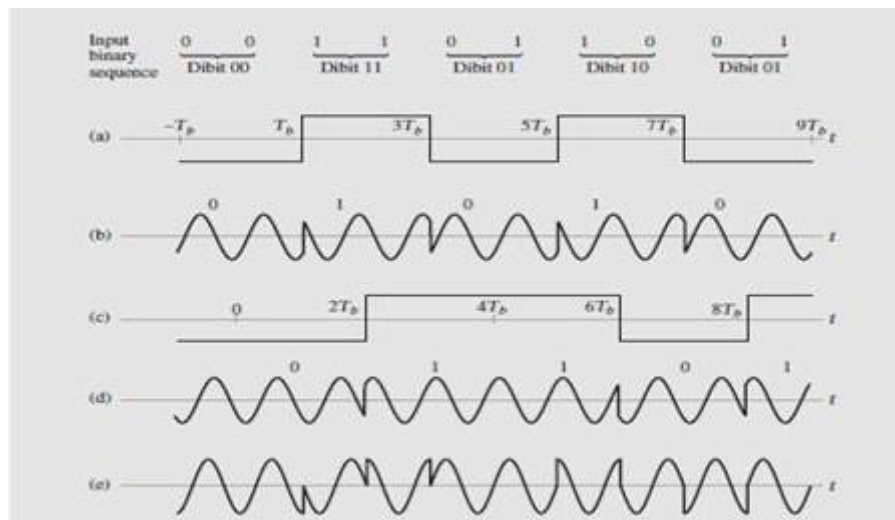


Figure 3.6 OQPSK signal components: (a) modulating signal for in phase component (b) modulated waveform of in phase component (c) modulating signal for quadrature component (d)

modulated waveform of quadrature component (e) waveform of OQPSK signal obtained by subtracting d from b.

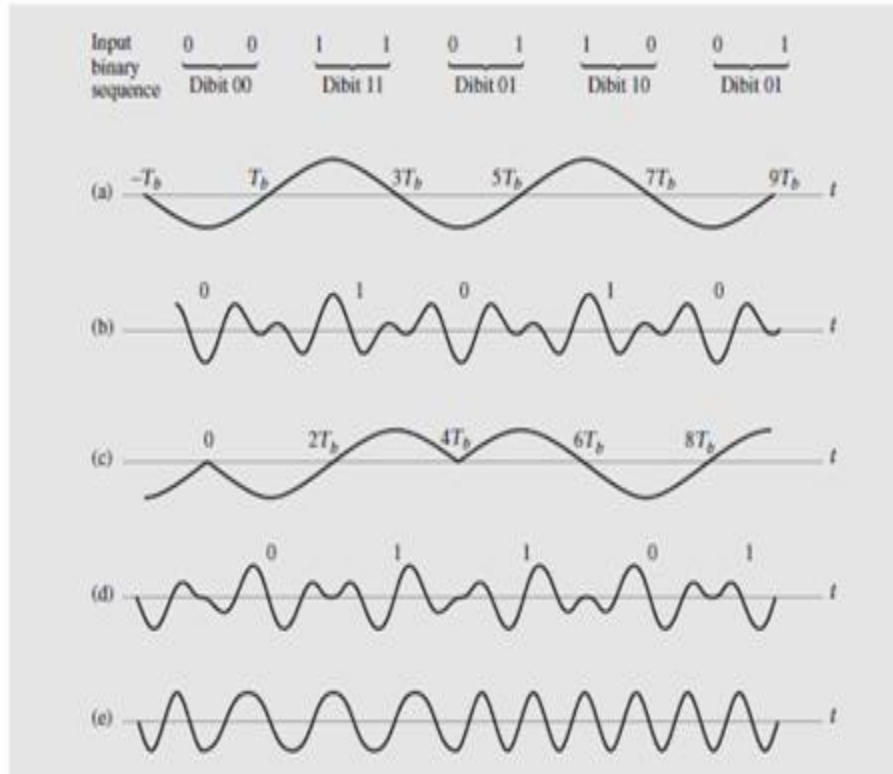


Figure 3.7 MSK signal components (a) modulating signal for in phase component (b) modulated waveform of in phase component (c) modulating signal for quadrature component (d) modulated waveform of quadrature component (e) waveform of MSK signal obtained by subtracting d from b.

FORMULATION OF MINIMUM-SHIFT KEYING

To proceed with the formulation, we refer back to Eq. (22), and use a well-known trigonometric identity to expand the angle-modulated wave (i.e., MSK signal) $s(t)$ as

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos(\theta(t)) \cos(2\pi f_c t) - \sqrt{\frac{2E_b}{T_b}} \sin(\theta(t)) \sin(2\pi f_c t) \quad \text{-----(25)}$$

In light of this equation, we make two identifications:

(i) $S_I(t) = \sqrt{\frac{2E_b}{T_b}} \cos(\theta(t))$ is the in-phase (I) component associated with the carrier π ----- (26)

(ii) $S_Q(t) = \sqrt{\frac{2E_b}{T_b}} \sin(\theta(t))$ is the quadrature (Q) component associated with the 90° - Phase shifted carrier----- (27)

To highlight the bearing of the incoming binary data stream on the composition of $s_1(t)$ and $s_2(t)$, we reformulate them respectively as follows:

$$s_1(t) = a_1(t) \cos(2\pi f_0 t) \quad \text{-----(28) and}$$

$$s_Q(t) = a_2(t) \sin(2\pi f_0 t) \quad \text{-----}(29)$$

The $a_1(t)$ and $a_2(t)$ are two binary waves that are extracted from the incoming binary data stream through demultiplexing and offsetting, in a manner similar to OQPSK. As such, they take on the value +1 or -1 in symbol (i.e., dibit) intervals of duration $T=2T_b$, where T_b is the bit duration of the incoming binary data stream. The two data signals $a_1(t)$ and $a_2(t)$ are respectively weighted by the sinusoidal functions $\cos(2\pi$ and $\sin(2\pi f_0 t)$, where the frequency f_0 is to be determined. To define f_0 , we use Eqs. (28) and (29) to reconstruct the original anglemodulated wave $s(t)$ in terms of the data signals $a_1(t)$ and $a_2(t)$. In so doing, we obtain

$$\begin{aligned} \theta(t) &= -\tan^{-1} \left[\frac{s_Q(t)}{s_I(t)} \right] \\ &= -\tan^{-1} \left[\frac{a_2(t)}{a_1(t)} \tan(2\pi f_0 t) \right] \quad \text{-----}(30) \end{aligned}$$

on the basis of which we recognize two possible scenarios that can arise:

1. $a_1(t)=a_2(t)$, This scenario arises when two successive binary symbols (constituting a dibit) in the incoming data stream are the same (i.e., both are 0s or 1s); hence, Eq. (30) reduces to

$$\begin{aligned} \theta(t) &= -\tan^{-1} [\tan(2\pi f_0 t)] \\ &= -2\pi f_0 t \quad \text{-----}(31) \end{aligned}$$

2. $a_2(t)=-a_1(t)$, This second scenario arises when two successive binary symbols (i.e., dibits) in the incoming data stream are different; hence, Eq. (7.31) reduces to

$$\begin{aligned} \theta(t) &= -\tan^{-1} [-\tan(2\pi f_0 t)] \\ &= 2\pi f_0 t \quad \text{-----}(32) \end{aligned}$$

Equations (31) and (32) are respectively of similar mathematical forms as Eqs. (24) and (23). Accordingly, we may now formally define

$$f_0 = \frac{1}{4T_b} \quad \text{-----}(33)$$

To sum up, given a non-return-to-zero level encoded binary wave $b(t)$ of prescribed bit T_b duration and a sinusoidal carrier wave of frequency f_c , we may formulate the MSK signal by proceeding as follows:

1. Use the given binary wave $b(t)$ to construct the binary demultiplexed-offset waves $a_1(t)$ and $a_2(t)$

2. Use Eq. (33) to determine the frequency f_0 .
3. Use Eqs. (28) and (29) to determine the in-phase component $s_I(t)$ and quadrature component $s_Q(t)$ respectively from which the MSK signal $s(t)$ follows.

NONCOHERENT DIGITAL MODULATION SCHEMES

Coherent receivers, exemplified by the schemes shown in Figs. 3.4(b) and 3.5(b), require knowledge of the carrier wave's phase reference to establish synchronism with their respective transmitters. However, in some communication environments, it is either impractical or too expensive to phase-synchronize a receiver to its transmitter. In situations of this kind, we resort to the use of noncoherent detection by abandoning the use of phase synchronization between the receiver and its transmitter, knowing that when we do so the receiver performance is degraded in the presence of channel noise..

NONCOHERENT DETECTION OF BASK SIGNAL

The generation of BASK signals involves the use of a single sinusoidal carrier of frequency f_c for symbol 1 and switching off the transmission for symbol 0. Now, the system designer would have knowledge of two system parameters:

- The carrier frequency f_c

The transmission bandwidth, which is determined by the bit duration T_b

It is therefore natural to make use of these known parameters in designing the noncoherent receiver for BASK. Specifically, the receiver consists of a band-pass filter, followed by an envelope detector, then a sampler, and finally a decision-making device, as depicted in Fig. 3.6. The band-pass filter is designed to have a mid-band frequency equal to the carrier frequency f_c and a bandwidth equal to the transmission bandwidth of the BASK signal. Moreover, it is assumed that the intersymbol interference (ISI) produced by the filter is negligible, which, in turn, requires that the rise time and decay time of the response of the filter to a rectangular pulse be short compared to the bit duration T_b . Under these conditions, we find that in response to the incoming BASK signal (assumed to be noise-free), the band-pass filter produces a pulsed sinusoid for symbol 1 and, ideally, no output for symbol 0. Next, the envelope detector traces the envelope of the filtered version of the BASK signal.

Finally, the decision-making device working in conjunction with the sampler, regenerates the original binary data stream by comparing the sampled envelope-detector output against a preset threshold every T_b seconds; this operation assumes the availability of bit-timing in the receiver. If the threshold is exceeded at time $t = iT_{b,i=0}$, the receiver decides in favor of symbol 1; otherwise, it decides in favor of symbol 0. In the absence of channel noise and channel distortion, the receiver output (on a bit-by-bit basis) would be an exact replica of the original binary data stream applied to the transmitter, subject to the abovementioned assumptions on the band-pass filter.

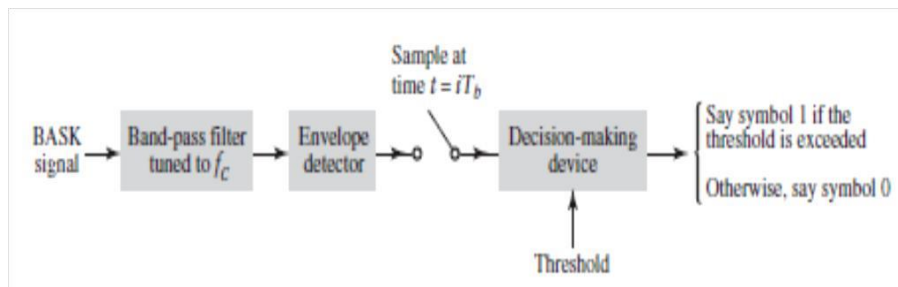


Figure 3.8 Noncoherent BASK receiver; the integer i for the sampler equals 0,

NONCOHERENT DETECTION OF BFSK SIGNALS

In the case of BFSK, the transmissions of symbols 1 and 0 are represented by two carrier waves of frequencies f_1 and f_2 respectively, with adequate spacing between them. In light of this characterization, we may build on the noncoherent detection of BASK by formulating the noncoherent BFSK receiver of Fig. 3.5. The receiver consists of two paths, one dealing with frequency f_1 (i.e., symbol 1) and the other dealing with frequency f_2 (i.e., symbol 0)

- Path 1 uses a band-pass filter of mid-band frequency f_1 . The filtered version of the incoming BFSK signal is envelope-detected and then sampled at time $t = iT_b, i=0, 1, \dots$, to produce the output v_1 .
- Path 2 uses a band-pass filter of mid-band frequency f_2 . As with path 2, the filtered version of the BFSK signal is envelope-detected and then sampled at time $t = iT_b, i=0, 1, \dots$, to produce the output v_2 .

The two band-pass filters have the same bandwidth, equal to the transmission bandwidth of the BFSK signal. Moreover, as pointed out in dealing with BASK, the intersymbol interference produced by the filters is assumed to be negligible. The outputs of the two paths, v_1 and v_2 are applied to a comparator, where decisions on the composition of the BFSK signal are repeated every T_b seconds. Here again, the availability of bit timing is assumed in the receiver. Recognizing that the upper path corresponds to symbol 1 and the lower path corresponds to symbol 0, the comparator decides in favour of symbol 1 if v_1 is greater than v_2 at the specified bit-timing instant; otherwise, the decision is made in favor of symbol 0. In a noise-free environment and no channel distortion, the receiver output (on a bit-by-bit basis) would again be a replica of the original binary data stream applied to the transmitter input.

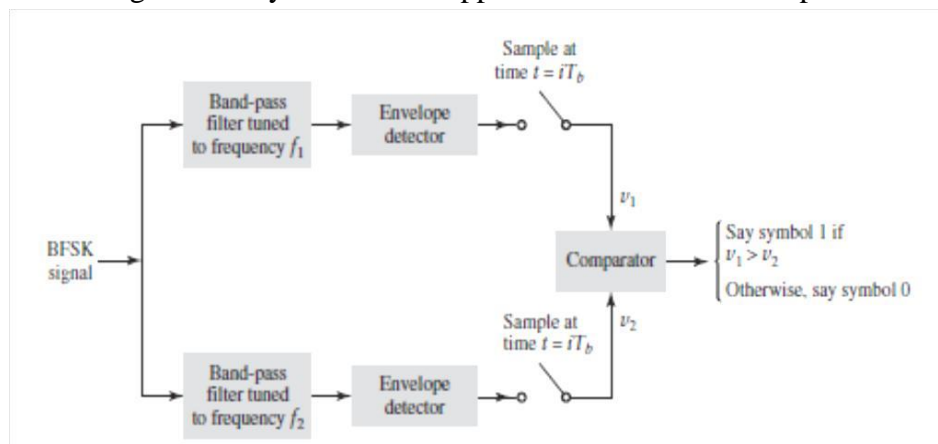


Figure 3.9 Noncoherent BFSK receiver; the two samplers operate synchronously, with $i=0, 1, \dots$,

DIFFERENTIAL PHASE-SHIFT KEYING

From the above discussion, we see that both amplitude-shift keying and frequency-shift keying lend themselves naturally to noncoherent detection whenever it is impractical to maintain carrier-phase synchronization of the receiver to the transmitter. But in the case of phase-shift keying, we cannot have noncoherent detection in the traditional sense because the term “noncoherent” means having to do without carrier-phase information. To get around this difficulty, we employ a “pseudo PSK” technique known as differential phase-shift keying (DPSK), which, in a loose sense, does permit the use of noncoherent detection.

DPSK eliminates the need for a coherent reference signal at the receiver by combining two basic operations at the transmitter:

- Differential encoding of the input binary wave
- Phase-shift keying

It is because of this combination that we speak of “differential phase-shift keying.” In effect, to send symbol 0, we phase advance the current signal waveform by 180 degrees, and to send symbol 1 we leave the phase of the current signal waveform unchanged. Correspondingly, the receiver is equipped with a storage capability (i.e., memory) designed to measure the relative phase difference between the waveforms received during two successive bit intervals. Provided the unknown phase varies slowly (i.e., slow enough for it to be considered essentially constant over two bit intervals), the phase difference between waveforms received in two successive bit intervals will be essentially independent of .

GENERATION AND DETECTION OF DPSK SIGNALS

(i) Generation

The differential encoding process at the transmitter input starts with an arbitrary first bit, serving merely as reference. Let $\{d_k\}$ denote the differentially encoded sequence with this added reference bit. To generate this sequence, the transmitter performs the following two operations:

- If the incoming binary symbol $\{b_k\}$ is 1, then the symbol $\{d_k\}$ is unchanged with respect to the previous symbol d_{k-1} .
- If the incoming binary symbol $\{b_k\}$ is 0, then the symbol $\{d_k\}$ is changed with respect to the previous symbol d_{k-1} .

The differentially encoded sequence $\{d_k\}$, thus generated is used to phase-shift a sinusoidal carrier wave with phase angles 0 and radians, representing symbols 1 and 0, respectively. The block diagram of the DPSK transmitter is shown in Fig. 3.5(a). It consists, in part, of a logic network and a one-bit delay element (acting as the memory unit) interconnected so as to convert the raw binary sequence $\{b_k\}$ into a differentially encoded sequence $\{d_k\}$. This sequence is amplitude-level encoded and then used to modulate a carrier wave of frequency f_c thereby producing the desired DPSK signal.

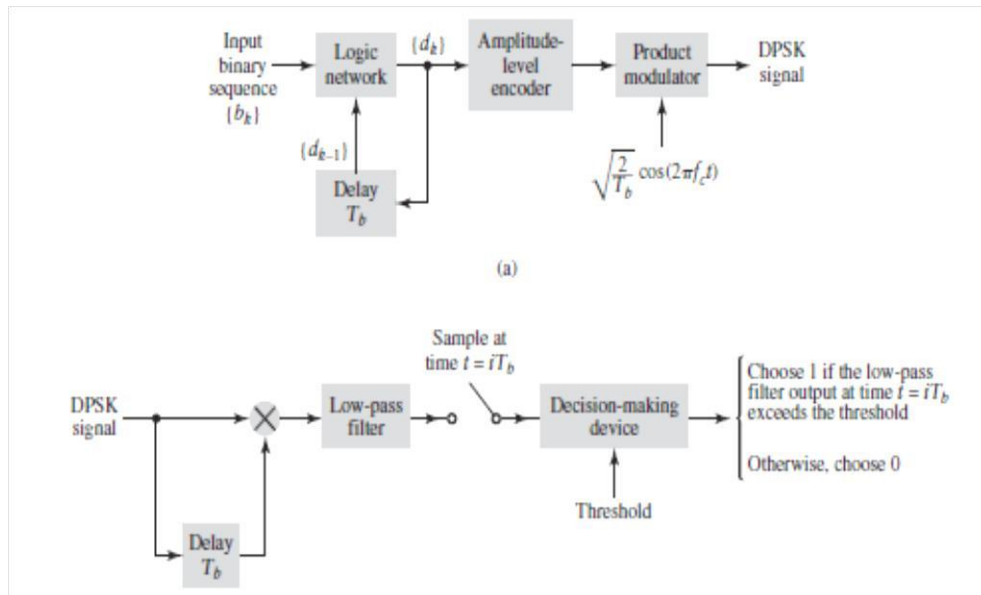


figure 3.11 Block diagrams for (a) DPSK transmitter and (b) DPSK receiver; for the sampler, integer $i=0$,

(ii) Detection

For the detection of DPSK signals, we take advantage of the fact that the phase-modulated pulses pertaining to two successive bits are identical except for a possible sign reversal. Hence, the incoming pulse is multiplied by the preceding pulse, which, in effect, means that the preceding pulse serves the purpose of a locally generated reference signal. On this basis, we may formulate the receiver of Fig. 3.8(b) for the detection of DPSK signals. Comparing the DPSK detector of Fig. 3.8(b) and the coherent BPSK detector of Fig. 3.3(b), we see that the two receiver structures are similar except for the source of the locally generated reference signal. According to Fig. 3.8(b), the DPSK signal is detectable, given knowledge of the reference bit, which, as mentioned previously, is inserted at the very beginning of the incoming binary data stream. In particular, applying the sampled output of the low-pass filter to a decision-making device supplied with a prescribed threshold, detection of the DPSK signal is accomplished. If the threshold is exceeded, the receiver decides in favor of symbol 1; otherwise, the decision is made in favor of symbol 0. Here again, it is assumed that the receiver is supplied with bit-timing information for the sampler to work properly.

M-ARY DIGITAL MODULATION SCHEMES

By definition, in an M-ary digital modulation scheme, we send any one of M possible signals $s_1(t), s_2(t), \dots, s_M(t)$, during each signaling (symbol) interval of duration T. In almost all applications, $M=2^m$, where m is an integer. Under this condition, the symbol duration, $T=MT_b$, where T_b is the bit duration. M-ary modulation schemes are preferred over binary modulation schemes for transmitting digital data over band-pass channels when the requirement is to conserve bandwidth at the expense of both increased power and increased system complexity.

In practice, we rarely find a communication channel that has the exact bandwidth required for transmitting the output of an information-bearing source by means of binary modulation schemes. Thus, when the bandwidth of the channel is less than the required value, we resort to an M-ary modulation scheme for maximum bandwidth conservation.

M-ARY PHASE-SHIFT KEYING

To illustrate the capability of M-ary modulation schemes for bandwidth conservation, consider first the transmission of information consisting of a binary sequence with bit duration T_b . If we were to transmit this information by means of binary PSK, for example, we would require a channel bandwidth that is inversely proportional to the bit duration T_b . However, if we take blocks of m bits to produce a symbol and use an M-ary PSK scheme $M=2^m$ with symbol duration $T=mT_b$, then the bandwidth required is proportional to $1/(mT_b)$. This simple argument shows that the use of M-ary PSK provides a reduction in transmission bandwidth by a factor $m=\log_2 M$ over binary PSK. In M-ary PSK, the available phase of 2π radians is apportioned equally and in a discrete way among the M transmitted signals, as shown by the phase-modulated signal

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{2\pi}{M} i\right), \quad \begin{matrix} i = 0, 1, \dots, M-1 \\ 0 \leq t \leq T \end{matrix} \quad \text{-----}(34)$$

where E is the signal energy per symbol, and f_c is the carrier frequency. Using a well-known trigonometric identity, we may expand Eq. (34) as

$$s_i(t) = \left[\sqrt{E} \cos\left(\frac{2\pi}{M} i\right) \right] \left[\sqrt{\frac{2}{T}} \cos(2\pi f_c t) \right] - \left[\sqrt{E} \sin\left(\frac{2\pi}{M} i\right) \right] \left[\sqrt{\frac{2}{T}} \sin(2\pi f_c t) \right], \quad \begin{matrix} i = 0, 1, \dots, M-1 \\ 0 \leq t \leq T \end{matrix} \quad \text{-----}(35)$$

The discrete coefficients $\sqrt{E} \cos\left(\frac{2\pi}{M} i\right)$ and $-\sqrt{E} \sin\left(\frac{2\pi}{M} i\right)$ are respectively referred to as the in-phase and quadrature components of the M-ary PSK signal $s_i(t)$. We now recognize that

$$\left\{ \left[\sqrt{E} \cos\left(\frac{2\pi}{M} i\right) \right]^2 + \left[\sqrt{E} \sin\left(\frac{2\pi}{M} i\right) \right]^2 \right\}^{1/2} = \sqrt{E}, \quad \text{for all } i \quad \text{-----}(36)$$

Accordingly, M-ary PSK modulation has the unique property that the in-phase and quadrature components of the modulated signal $s_i(t)$ are interrelated in such a way that the discrete envelope of the signal is constrained to remain constant at the value \sqrt{E} for all M .

SIGNAL-SPACE DIAGRAM

The result described in Eq. (36) combined with the fact that the in-phase and quadrature components of M-ary PSK are discrete, leads to an insightful geometric portrayal of M-ary PSK. To explain, suppose we construct a two-dimensional diagram with the horizontal and vertical axes respectively defined by the following pair of orthonormal functions:

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T$$

And

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad 0 \leq t \leq T$$

------(38)

where the band-pass assumption implies orthogonality; the scaling factor assures unit energy

over the interval T for both $\phi_1(t)$ and $\phi_2(t)$. On this basis, we may represent the in-phase

component π and quadrature component π for $i=0,1,\dots,M-1$, as a set of

points in this two-dimensional diagram, as illustrated in Fig. 7.19 for $M = 8$. Such a diagram is referred to as a signal-space diagram.

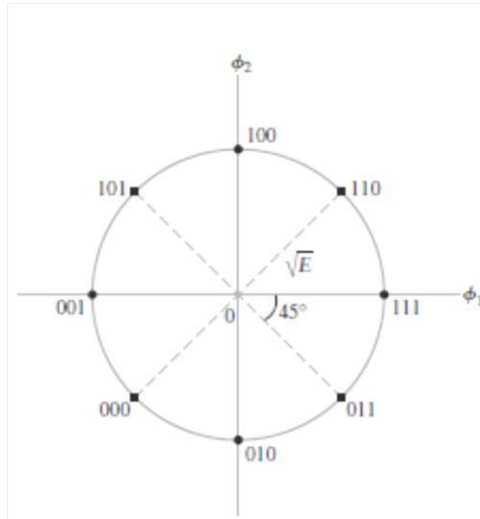


Fig 3.12 Signal-space diagram of 8-PSK.

Figure 3.9 leads us to make three important observations:

1. M-ary PSK is described in geometric terms by a constellation of M signal points distributed uniformly on a circle of radius
2. Each signal point in the figure corresponds to the signal $S_i(t)$ of Eq. (34) for a particular value of the index i .
3. The squared length from the origin to each signal point is equal to the signal energy E . In light of these observations, we may now formally state that the signal-space diagram of Fig. 3.6 completely sums up the geometric description of M-ary PSK in an insightful manner. Note that the 3-bit sequences corresponding to the 8 signal points are Gray-encoded, with only a single bit changing as we move along the constellation in the figure from one signal point to an adjacent one.

M-ARY QUADRATURE AMPLITUDE MODULATION

Suppose next that the constraint of Eq. (36) that characterizes M-ary PSK modulation is removed. Then, the in-phase and quadrature components of the resulting M-ary modulated signal are permitted to be independent of each other. Specifically, the mathematical description of the new

$$s_i(t) = \sqrt{\frac{2E_0}{T}} a_i \cos(2\pi f_c t) - \sqrt{\frac{2E_0}{T}} b_i \sin(2\pi f_c t), \quad i = 0, 1, \dots, M-1$$

------(39)

where the level parameter in the in-phase component and the level parameter a_i in the quadrature component b_i are independent of each other for all i . This new modulation scheme is called M-

ary quadrature amplitude modulation (QAM). Note also that the constant E_0 is the energy of the signal pertaining to a particular value of the index i for which the amplitude of the modulated signal is the lowest. M-ary QAM is a hybrid form of M-ary modulation, in the sense that it combines amplitude-shift keying and phase-shift keying. It includes two special cases:

- (i) If $b_i=0$, for all i , the modulated signal $S_i(t)$ of Eq. (39) reduces to

$$s_i(t) = \sqrt{\frac{2E_0}{T}} a_i \cos(2\pi f_c t) \quad i = 0, 1, \dots, M-1$$

which defines M-ary amplitude-shift keying (M-ary ASK).

- (ii) If $E_0=E$ and the constraint

$$\text{for all } i \quad (Ea_i^2 + Eb_i^2)^{1/2} = \sqrt{E},$$

is satisfied, then the modulated signal $S_i(t)$ of Eq. (39) reduces to M-ary PSK.

SIGNAL-SPACE DIAGRAM

Figure 3.10 portrays the signal-space representation of M-ary QAM for $M=16$, with each signal point being defined by a pair of level parameters a_i and b_i where $i=1,2,3,4$. This time, we see that the signal points are distributed uniformly on a rectangular grid. The rectangular property of the signal-space diagram is testimony to the fact that the in-phase and quadrature components of M-ary QAM are independent of each other. Moreover, we see from Fig. 7.21 that, unlike M-ary PSK, the different signal points of M-ary QAM are characterized by different energy levels, and so they should be. Note also that each signal point in the constellation corresponds to a specific quadbit, which is made up of 4 bits. Assuming the use of Gray encoding, only one bit is changed as we go from each signal point in the constellation horizontally or vertically to an adjacent point, as illustrated in Fig. 3.7

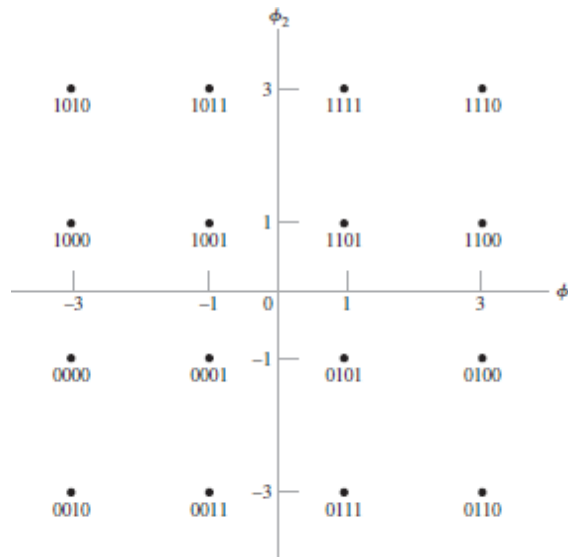


Figure 3.10 Signal-space diagram of Gray encoded M-ary QAM for $M=16$.

M-ARY FREQUENCY-SHIFT KEYING

However, when we consider the M-ary version of frequency-shift keying, the picture is quite different from that described for M-ary PSK or M-ary QAM. Specifically, in one form of M-ary FSK, the transmitted signals are defined for some fixed integer n as follows:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos \left[\frac{\pi}{T} (n + i)t \right], \quad i = 0, 1, \dots, M-1$$

$$0 \leq t \leq T$$

-----(40)

The M transmitted signals are all of equal duration T and equal energy E. With the individual signal frequencies separated from each other by 1/2T hertz, the signals in Eq.(40) are orthogonal; that is, they satisfy the condition

$$\int_0^T s_i(t)s_j(t) dt = \begin{cases} E & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$

-----(41)

Like M-ary PSK, the envelope of M-ary FSK is constant for all M, which follows directly from Eq. (40). Hence, both of these M-ary modulation strategies can be used over nonlinear channels. On the other hand, M-ary QAM can only be used over linear channels because its discrete envelope varies with the index i (i.e., the particular signal point chosen for transmission).

SIGNAL-SPACE DIAGRAM

To develop a geometric representation of M-ary FSK, we start with Eq. (40). In terms of the signals $S_i(t)$ defined therein, we introduce a complete set of orthonormal functions:

$$\phi_i(t) = \frac{1}{\sqrt{E}} s_i(t) \quad i = 0, 1, \dots, M-1$$

$$0 \leq t \leq T$$

-----(42)

Unlike M-ary PSK and M-ary QAM, we now find that M-ary FSK is described by an M-dimensional signal-space diagram, where the number of signal points is equal to the number of coordinates. The visualization of such a diagram is difficult beyond M=3.

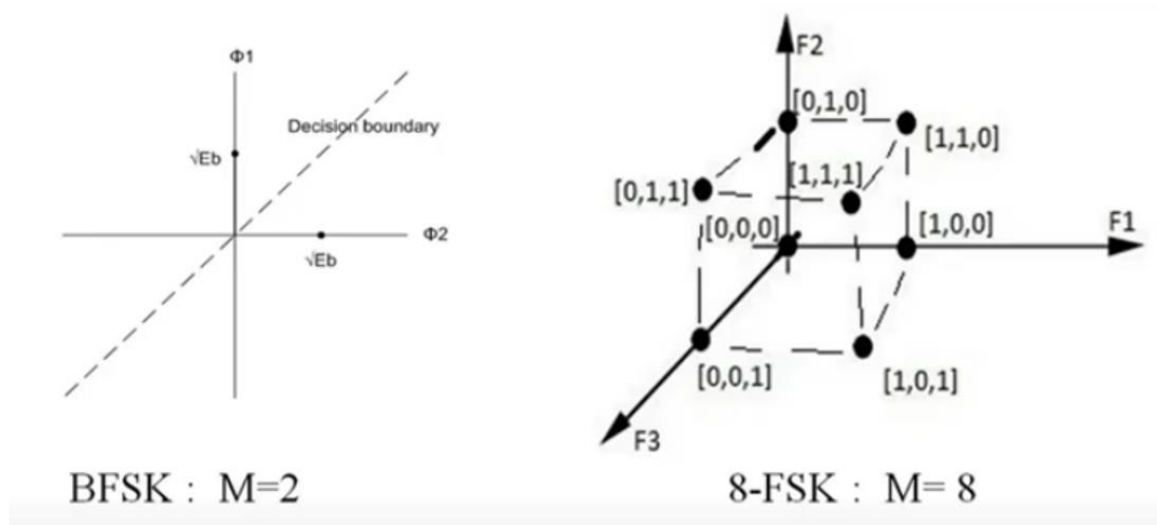


Figure 3.11 Signal-space diagram of BFSK and M-ary FSK

QUESTION BANK

UNIT IV

PART-A

1. Compare coherent and non-coherent detection.
2. A BFSK system employs two signalling frequencies f_1 and f_2 . The lower frequency f_1 is 1200 Hz and signaling rate is 500 Baud. Compute f_2 .
3. Outline the concept of M-ary modulation systems.
4. Formulate the probability of error for ASK, PSK and FSK modulation schemes.
5. Plot the binary ASK waveform for the bit sequence "1011011".
6. Plot the binary FSK waveform for the bit sequence "1010110".
7. What would be the advantage of the BFSK system over the BASK system, given that they operate on the same binary data stream and the same communication channel? Justify your answer.
8. Plot the binary PSK waveform for the bit sequence "1011001".
9. Consider a binary bit stream, with bit duration $T_b=1\text{ms}$, is modulated by a sinusoidal carrier $\sin 2\pi 4000t$. Compute the bandwidth of the modulated signal for BPSK and QPSK.
10. Mention the merits and demerits of DPSK.
11. Compare QPSK and offset QPSK in-terms of phase changes that occurs in the modulated signal and power spectrum.
12. Draw the signal space diagram of 8-PSK.
13. QAM is a combination of which two types of modulation?
14. In 64-PSK, how many bits are transmitted per symbol?

PART-B

1. a) Explain the generation and non-coherent detection of binary amplitude shift keying (BASK) signal and derive the equation for probability of error in the presence of AWGN channel noise.
(12marks)

a) A binary sequence 11100101 is applied to a BASK modulator. The bit duration is $1\mu\text{s}$ and the sinusoidal carrier wave used to represent symbol 1 has a frequency equal to 7MHz. Plot the time domain and frequency domain representation of the transmitted ASK signal.
(4 marks)
- 2 a) Illustrate the generation and non-coherent detection of binary frequency shift keying (BFSK) signal and derive the equation for probability of error in the presence of AWGN channel noise.
(12marks)

b) The binary sequence 10110101 is applied to a BFSK modulator. The bit duration is $2\mu\text{s}$. The carrier frequencies used to represent symbols 0 and 1 are 2.5 MHz and 3.5 MHz, respectively. (a) Calculate the transmission bandwidth of the BFSK signal. (b) Plot the waveform of the BFSK signal.
(4 marks)
3. a) Explain the generation and coherent detection of binary phase shift keying (BPSK) signal and derive the equation for probability of error in the presence of AWGN channel noise.

(12marks)

3.b The binary sequence 10100110 is applied to a BPSK system. The bit duration is $10\mu\text{s}$. The carrier frequency is 600KHz. (a) Calculate the transmission bandwidth of the system. (b) Plot the waveform of the transmitted signal. (4 marks)

4. a) Elucidate the generation and detection of quadrature phase shift keying (QPSK) signal and derive the expression for probability of error in the presence of AWGN channel noise. (12marks)

4. b) The binary sequence 11000110 with bit duration $1\mu\text{s}$ is applied to a QPSK modulator. The carrier frequency is 4MHz. Plot the waveform of the QPSK signal and calculate its transmission bandwidth. (4 marks)

5. Discuss the generalized structure of an M-ary FSK transmitter and receiver. Also develop the expression for transmission bandwidth and probability of error. (16 marks)

6 a) Elucidate generation, detection and bandwidth requirements of M-ary PSK modulation. (12 marks)

6.b) Mention the factors that should be considered in choosing the value of M in an M-ary system? (4 marks)

UNIT IV – SYNCHRONIZATION - SEC1313

SYNCHRONIZATION:

Synchronization: Receiver synchronization - Coherent systems - Symbol and frame synchronization - Network synchronization - Open and closed loop transmitter synchronization - Tracking and acquisition in spread spectrum system.

Synchronization

The signals from various sources are transmitted on the single channel by multiplexing. This requires synchronization between transmitter and receiver. So the transmitter and Receiver is said to be synchronous when an event occur simultaneously in both transmitter and receiver at a particular instant of time. The process of making a situation synchronous and maintaining it in this condition is called synchronization.

If special synchronization bits are added in the transmitted signal for this purpose means it is known as transmitter synchronization.

Synchronization is also required for detectors to recover the digital data properly from the modulated signal. In almost every receiver or demodulator performance, some level of signal synchronization is used. If the parameters are adjusted at the receiver environment for synchronous then it is called receiver synchronization. For example in the case of coherent phase demodulation (PSK), the receiver is assumed to be able to generate reference signals whose phases are identical (except perhaps for a constant offset) to those of the signaling alphabet at the transmitter.

Digital communication systems using coherent modulation require three levels of synchronization 1-Phase, 2-Symbol, 3-Frame synchronization. (2&3 combinely known as Time synchronization)

Digital communication systems using noncoherent modulation require two levels of synchronization 1-Frequency, 2-Symbol, 3-Frame synchronization.

Since the modulation is not coherent, accurate phase lock is not required. Instead, noncoherent require frequency synchronization. Frequency synchronization differs from phase synchronization in that the replica of the carrier that is generated by the receiver is allowed to have arbitrary constant phase offset from the received carrier. Receiver designs can be simplified by removing the requirement to determine the exact value of the incoming carrier phase.

In digital communication systems, the receiver, has accurate knowledge about incoming symbol Time period. This knowledge is required in order to know the proper symbol integration interval (the interval over which energy is integrated prior to making symbol decision).

It can be seen that symbol synchronization and phase synchronization are similar in that both involve producing in the receiver a replica of a portion of the transmitted signal.

For phase synchronization, it is an accurate replica of the carrier.

For symbol synchronization, it is a square wave at the symbol transition rate (the receiver must be able to produce a square wave for each incoming signals transitions between symbols). A receiver that is able to do this can be said to have symbol synchronization, or to be in symbol lock.

In many communication systems an even higher level of synchronization is required. This is usually called frame synchronization. Frame synchronization is required when the information is organized in blocks, or messages of some uniform number of symbols (this occur, for example, if a block code is used for forward error control, or if the communication channel is used for several user as in TDMA). In the case of block coding, the decoder needs to know the boundaries between code words in order to decode the message correctly. In the case TDMA, it is necessary to know where the location of boundaries between channel users are, in order to route the information correctly.

Similar to symbol synchronization, frame synchronization is equivalent to being able to generate a square wave at the frame rate, with the zero crossing coincident with the transitions from one frame to the next.

All of the discussion thus far had been oriented toward the receiving end of a communication link. There are many cases, when the transmitter assumes the more active role in synchronization by varying the time and frequency of its transmissions to receiver, as in the case of satellite communication network. In most of these cases, the transmitter knows on a return path from the receiver to determine the accuracy of its synchronization (two way communications are used in the transmitter synchronization). The transmitter synchronization is often called network synchronization.

Synchronization costs and benefits

- The most obvious costs is in the need for additional hardware or software in the receiver for acquisition and tracking.
- Extra time is required to achieve synchronization before commencing communications.
- Error control coding is used.

Frequency and phase synchronization (Carrier Synchronization)

The carrier synchronization is required in coherent detection methods to generate a coherent reference at the receiver. In this method the data bearing signal is modulated on the carrier in such a way that the power spectrum of the modulated carrier signal contains a discrete component at the carrier frequency. That is the Fourier transform of the modulated signal contains one component at f_c also. Then the phase lock loop can be used to track this component f_c . the output frequency of phase locked loop is thus locked to the carrier frequency f_c in the transmitted signal. This output frequency of phase locked loop is used as a coherent reference signal for detection in the receiver.

The heart of all synchronization circuits is the phase locked loop (PLL). A schematic diagram of the basic PLL is shown in fig. PLL is a servo control loop, whose controlled parameter is the phase of a locally generated replica of the incoming carrier signal. PLL has three basic components:-

Phase detector: is a device that produce a measure of the difference in phase between an incoming signal and the local replica. As the incoming signal and the local replica change with respect to each other, the phase difference (or phase error) becomes a time varying signal into the loop filter.

Loop filter: is the device governs the PLL response to these variations in the error signal.

VCO: is the device that produces the carrier replica. The VCO is a sinusoidal oscillator whose frequency is controlled by a voltage level at the device input. In figure the phase detector is shown as a multiplier, the loop filter is described by its impulse response function $f(t)$, with Fourier transform $F(\omega)$, and the voltage controlled oscillator (VCO) is also indicated. A VCO is an oscillator whose output frequency is a linear function of its input voltage over a certain range. The output of VCO may be a digital signal (i.e not a sinusoidal), but it may be implemented as a read only memory whose output frequency controlled by a combination of a clock and the error signal (as a variable address).

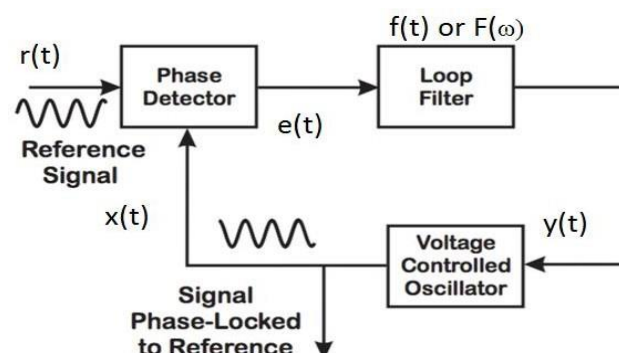


Figure 1. Basic PLL

Let us consider the input signal to the PLL is given by

$$r(t) = 2 \cos[\omega_0 t + \theta(t)]$$

Where $\omega_0 t$ is the nominal carrier frequency and $\theta(t)$ is a slowly varying phase. Similarly, consider a normalized VCO output of the form .

$$x(t) = 2 \sin[\omega_0 t + \hat{\theta}(t)]$$

Where $\hat{\theta}(t)$ is the estimated phase.

These signals will produce an output error signal at the phase detector output of the form

$$\begin{aligned} e(t) &= x(t)r(t) = 2 \sin[\omega_0 t + \hat{\theta}(t)] * 2 \cos[\omega_0 t + \theta(t)] \\ e(t) &= x(t)r(t) = \sin[\theta(t) + \hat{\theta}(t)] + \sin[2\omega_0 t + \theta(t) + \hat{\theta}(t)] \end{aligned}$$

Assuming that the loop filter is low pass filter, the second term will be filtered out. Thus the low pass filter provides error signal that is a function of the difference phases between input and the output, only. This error signal is applied to the input VCO. The VCO output frequency is the time derivative of the of the argument (angle) of the sine function. If we assume that ω_0 is the output frequency of the VCO when the voltage is zero, we can express the difference in the VCO output frequency from ω_0 as the differential of the phase term $\hat{\theta}(t)$.

Therefore, since an input voltage of zero produces an output frequency of ω_0 , the difference in the output frequency ω_0 form will be proportional to the value of the input voltage $y(t)$.

$$\Delta\omega(t) = \frac{d}{dt} [\hat{\theta}(t)] = K y(t) = K e(t) * f(t)$$

$$\begin{aligned} &\frac{d}{dt} [\hat{\theta}(t)] \\ &\approx K_0 [\theta(t) - \hat{\theta}(t)] * f(t) \end{aligned}$$

Where $\Delta\omega(t)$ =the frequency difference, $*$ = convolution, K_0 = the gain the VCO

$f(t)$ =the loop filter impulse response

This linear differential equation in $\hat{\theta}(t)$ (if small angle approximation) is known as the linearized loop equation. It is used to determine the loop behaviour during normal operation (where the phase error is small). Consider the Fourier transform of above equations

$$F \left\{ \frac{d[\hat{\theta}(t)]}{dt} \right\} = F \{ K [\theta(t) - \hat{\theta}(t)] * f(t) \}$$

$$\begin{aligned} &\frac{d}{dt} \hat{\theta}(t) \\ &jW \hat{\theta}(\omega) = K_0 [\theta(\omega) - \hat{\theta}(\omega)] F(\omega) \end{aligned}$$

$$\hat{\theta}(\omega) [j\omega + K_0 F(\omega)] = K_0 \theta(\omega) F(\omega)$$

$$\hat{\theta}(t) = \frac{K_0 F(\omega)}{j\omega + K_0 F(\omega)} \theta(t)$$

$$\overline{\theta(t)} = j\omega + K_0 \quad F(\omega) = H(j\omega)$$

Where $H(j\omega)$ is the closed loop transfer function of the PLL

Example :

Show that for appropriately chosen K_0 , and $f(t)$ the linearized loop equation (for PLL) demonstrates a tendency toward phase lock-that is, the phase difference between the incoming signal and the VCO output tends to decrease.

Solution

$$\Delta(t) \approx K_0 [\theta(t) - \hat{\theta}(t)] * f(t)$$

1. Consider the case where the phase of the input signal is $\theta(t)$ slowly varying with time.
2. If the phase difference on the right hand side of above equation is positive [i.e $\theta(t) > \hat{\theta}(t)$], so that $\hat{\theta}(t)$ will increase with time, which will tend to reduce the magnitude of the difference $\theta(t) - \hat{\theta}(t)$
3. If $\theta(t) = \hat{\theta}(t)$ then the equation indicates that $\hat{\theta}(t)$ will not change with time ,and the equality will be maintained.

Steady state tracking characteristics.

By using above equation we can obtain an expression for the Fourier transform of the phase error $E(\omega)$

$$\frac{\hat{\theta}(\omega)}{\theta(\omega)} = \frac{K_0 F(\omega)}{j\omega + K_0 F(\omega)} = H(j\omega)$$

$$\begin{aligned} E(\omega) &= Fe(t) = [\theta(\omega) - \hat{\theta}(\omega)] \\ &= [1 - H(\omega)]\theta(\omega) \\ &= \frac{j\omega\theta(\omega)}{j\omega + K_0 F(\omega)} \end{aligned}$$

The above equation can be used to determine the steady state error response of a loop to a variety of possible input characteristics. The steady state error is the residual error after all transients have died away.

$$\log_{t \rightarrow \infty} e(t) = \log_{j\omega \rightarrow} j\omega E(\omega)$$

Carrier synchronization using Mth power loop

Fig 2 shows the block diagram of carrier recovery or carrier synchronization circuit.

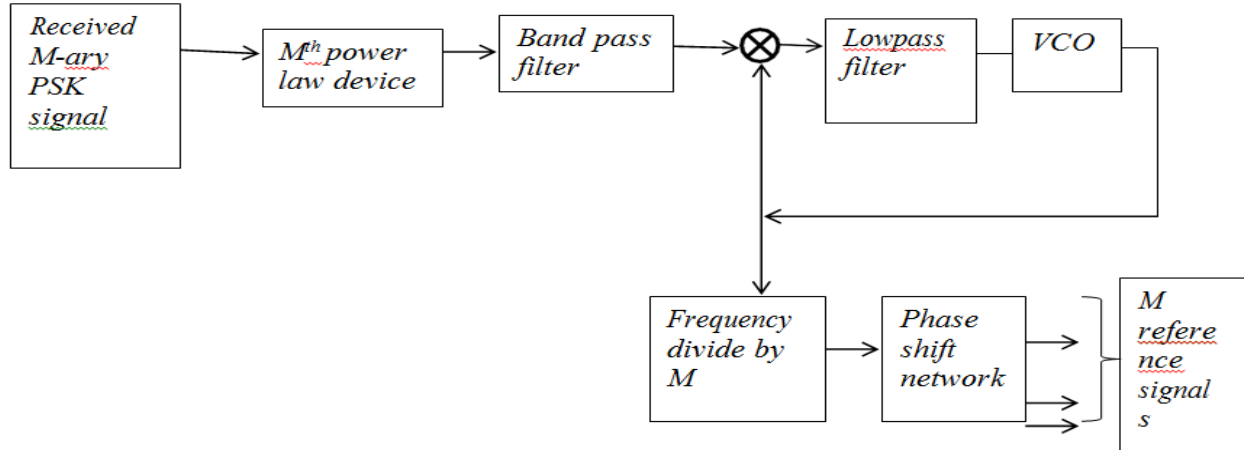


Fig 2: Block diagram of Mth power loop

Fig 2 shows the block diagram of carrier recovery circuit for M-ary PSK*. This circuit is called the Mth power loop. When $M = 2$, then it is called squaring loop. When $M = 2$, the M-ary PSK is then called as binary PSK. As shown in diagram, the input signal is first raised to the Mth power by the mth power law device with the frequency f_c . The phase locked loop consists of a phase detector, low-pass filter and VCO. The phase locked loop tracks the carrier frequency. Then the output of a voltage controlled oscillator (VCO) is the carrier frequency. The output frequency of VCO is divided by M. this is done because the Mth power of the input signal multiplies carrier frequency by M. the phase shift network then separates 'M' reference signals for the 'M' correlation receivers. In this technique the power of the input signal is raised to some power say 'M'. Let us say $M = 2$, then the input signal is squared. Because of this, the sign of the recovered carrier is always independent of sign of the input signal carrier since it is squared. Therefore there can be 180 degree error in the output.

Suppressed of carrier loops

In the case of a phase modulation, if the carrier phase variation due to the modulation less than $\pi/2$ radians, there will be positive energy at the carrier frequency. This is called a system design that has a residual carrier component. A diagram of the signal space for a binary phase modulated with a residual carrier components is shown in figure, for modulating angle of $\gamma \leq \pi/2$. However, the residual component is wasted energy, it is not to transmit the information, only to transmit the carrier. Thus most modern phase modulated systems are suppressed carrier systems. This means that there is no average energy transmitted at the carrier frequency also there is no longer any signal for the basic PLL of figure to track.

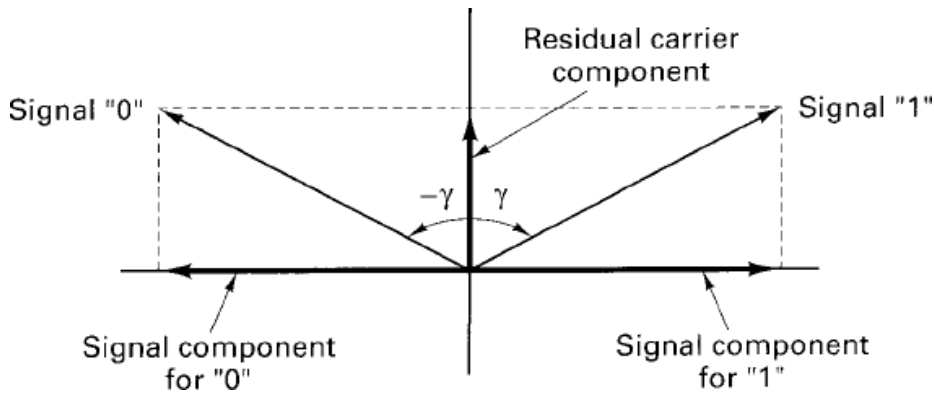


Figure 3 Residual Carrier

Consider, as an example, a BPSK

$$r(t) = m(t) \sin(\omega_0 t + \theta) + n(t)$$

Where $m(t) = 1$ with equal probability. This is suppressed carrier with the average energy at radian frequency ω_0 is zero. This is represented graphically in the above figure when $\gamma \leq \pi/2$, and in this case the residual carrier will vanish.

To acquire and track the phase of the carrier, the effects of the modulation must be eliminated. One way to eliminate the modulation is to square the signal.

$$r^2(t) = m^2(t) \sin^2(\omega_0 t + \theta) + n^2(t) + 2n(t)m(t)\sin(\omega_0 t + \theta)$$

$$r^2(t) = \frac{1}{2} - \frac{1}{2} \cos(2\omega_0 t + 2\theta) + n^2(t) + 2n(t)m(t)\sin(\omega_0 t + \theta)$$

If $m^2(t) = 1$, the second term on the right hand side of the above equation is a carrier related term (at twice the original carrier frequency) that can be acquired and tracked with a basic PLL of the type illustrated in figure 1. Such an arrangement is illustrated in figure 4.

The problems with this schematic are:-

- All phase angles have been doubled. Thus, the phase noise and phase jitter have been doubled.
- Larger S/N is required for this PLL to overcome this larger internal variation of the noise.
- False lock. This additional loss due to the terms in the above equation

$n^2(t) + 2n(t)m(t)\sin(\omega_0 t + \theta)$ which is called the loop squaring loss (SL)

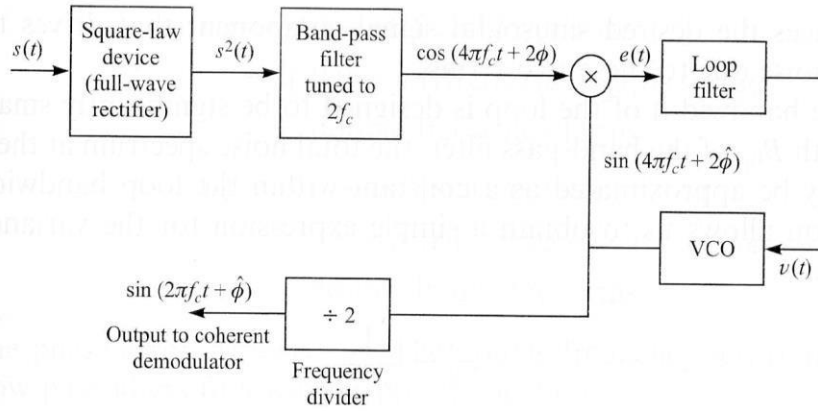


Figure 4 Suppressed of carrier loops

Costas loop

This loop design is important because it eliminates the square law device, which can be difficult to implement at carrier frequencies, and replaces it with a multiplier and relatively simple low pass filters. The problem in this loop is to implement the two LPF identical. This problem can be solved by using the digital filter. The decision as to whether to implement a Costas loop or the classical design of figure 5 amount to a design decision between the difficulty of implementing the squaring device and the difficulty of implementing closely matched arm filter

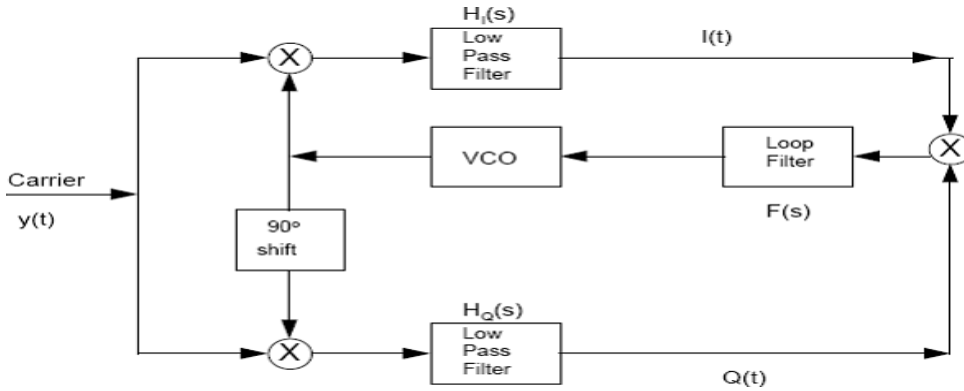


Figure 5 Costas loop

$\frac{1}{2} \frac{\sqrt{2}E}{T} \cos(\varphi - \theta)$ and $\frac{1}{2} \frac{\sqrt{2}E}{T} \sin(\varphi - \theta)$. The multiplier output is given as,

$$\begin{aligned} V_m &= \frac{1}{4} \times \frac{2E}{T} \sin(\varphi - \theta) \cos(\varphi - \theta) \\ &= \frac{E}{2T} \cdot \frac{1}{2} \sin 2(\varphi - \theta) \\ &= \frac{E}{4T} \sin 2(\varphi - \theta) \end{aligned}$$

The power 'P' of the signal over the period T is given by,

$$P = E/T$$

Therefore equation V_m can be written as,

$$V_m = P/4 \sin 2(\varphi - \theta)$$

If there is some difference between the VCO frequency and the input carrier frequency then the phase difference $(\phi - \theta)$ is changed proportionally. The change in $(\phi - \theta)$ causes V_m to increase or decrease VCO frequency such that synchronization is maintained.

Higher order suppressed carrier loops

Squaring the signal a second time (equivalent to taking the original signal to the fourth power) can be seen to produce a term with a carrier component at four times the transmitted carrier's frequency. The loss for fourth power loop is

$$S_L \leq 1 + \frac{9}{\gamma_i^2} + \frac{6}{\gamma_i^2} + \frac{3}{2\gamma_i^3}$$

Example

$$10 \log$$

$$\frac{S}{N}$$

For the squaring loop :

$$S_L = 1 + \frac{1}{2P_i} \\ = 10 \log$$

For the fourth loop :

$$S_L \leq 1 +$$

$$S_L \leq 1 + \frac{9}{10} + \frac{6}{10^2} + \frac{3}{2 \times 10^3} \\ 10 \log$$

For the squaring loop :

Thus, an input signal to noise ratio of 10dB is suitable to keep losses small for the squaring loop, the same signal to noise ratio may allow significant losses for the fourth power loop.

Symbol synchronization (clock recovery)

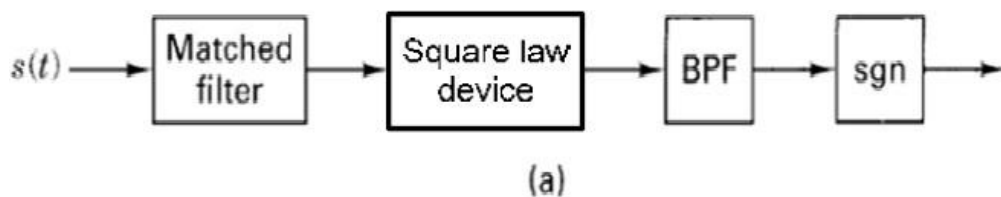
In the last section we have seen how to recover the carrier for coherent detection. In a matched filter or correlation receiver, the incoming signal is sampled at the end of one bit or symbol duration. This is because the output of matched filter is maximum at $t_m = T$. here T is the symbol or bit duration. Therefore the receiver has to know the instants of time at which a symbol or bit is transmitted. That is the instants at which a particular bit or symbol starts and when it is ended. The estimation of these times of bit or symbol is called symbol or bit synchronization. The symbol or bit timing is controlled by a clock at the transmitter. The receiver has to know the phase of this clock. Therefore symbol or bit synchronization is also called clock recovery.

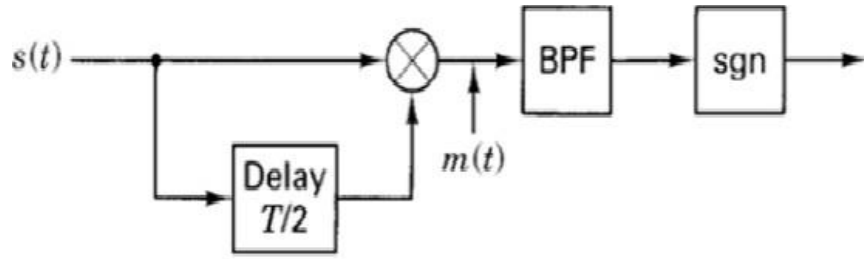
The symbol synchronizers can be classified into two basic groups according to the type of the loop:-

1. Open loop synchronizers: These circuits recover a replica of the transmitter data clock output directly from operation on the incoming data stream.
2. Closed loop synchronizers: These synchronizers attempt to lock a local data clock to the incoming signal by use of comparative measurements on the local and incoming signals.

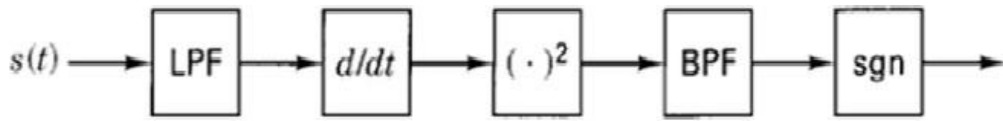
Open loop symbol synchronizers -Open loop symbol synchronizers are also called filter synchronizers.

This class of synchronizers generates a frequency component at the symbol rate by operating on the incoming baseband sequence with a combination of filtering and a nonlinear device. In this synchronizer, the desired frequency component, at the data symbol rate, is filtered by the bandpass filter, and shaped with a high gain saturating amplifier. The shaping recovers the square wave appearance of the data clock signal. Three examples of open loop bit synchronizers are shown in Following figures





(b)



(c)

In the first example the incoming signal is filtered with matched filter and rectified by the square device. The output waveform from square law device will contain a Fourier component at the fundamental frequency of the data clock. This frequency is filtered by BPF and shaped with an ideal saturating amplifier, with transfer function (signum function)

$$Sgn(t) = \begin{cases} 1 & \text{for } t > 0 \\ 0 & t = 0 \\ -1 & t < 0 \end{cases}$$

The second example shown in figure produces a Fourier component at the data clock frequency by means of a delay and multiply. The delay shown in figure is half a bit period, which is the best value because it provides the strongest Fourier component. The third example with the main operations are those differentiation and rectification by using a square law device. For a square wave input, the differentiator will produce positive or negative spikes at all symbol (bit) transitions. When rectified, the resulting sequence of positive spikes will have a Fourier component at the data symbol (bit) rate.

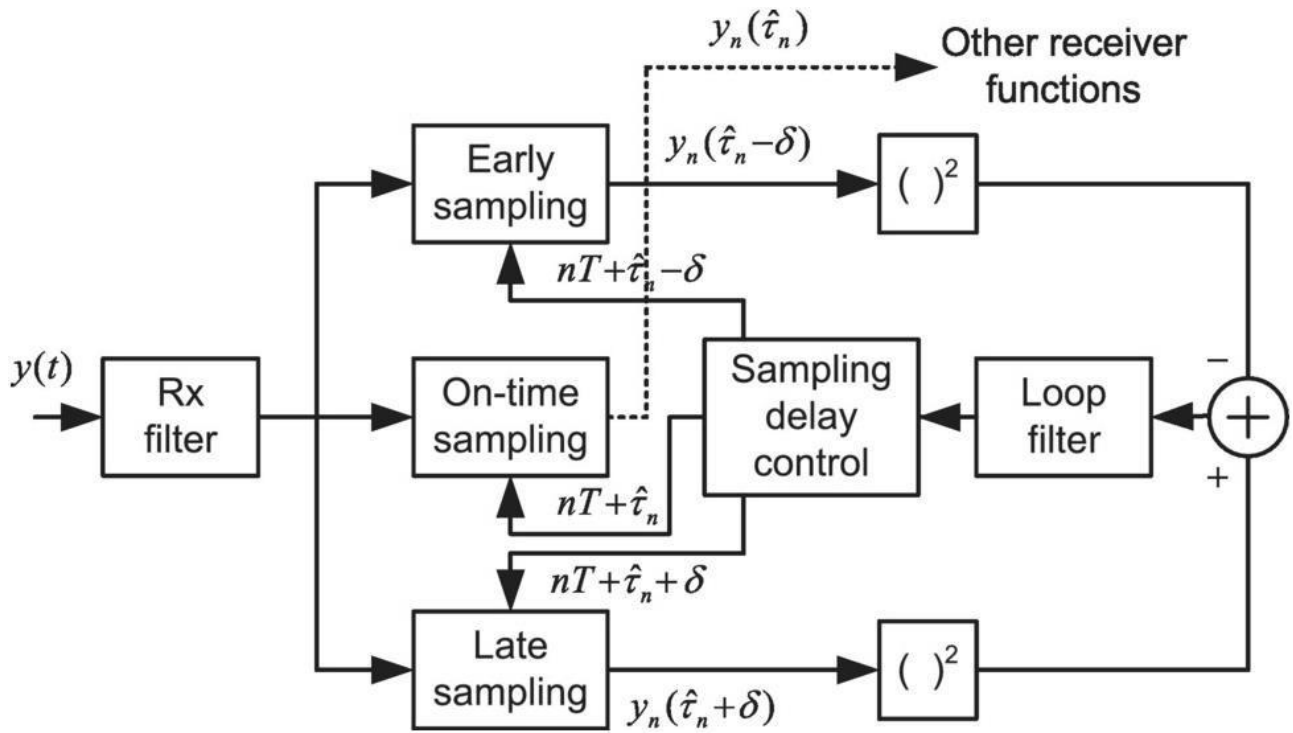
Closed loop symbols synchronizers

The primary disadvantage of open loop symbol synchronization methods is that there is an unavoidable non-zero mean tracking. Closed loop, symbol data synchronizers use comparative measurements on the incoming signal and a locally generated data lock signal to bring the locally generated signal into synchronism with incoming data transitions.

Early Late Synchronizer

The most popular of the closed loop symbol synchronizers is the early/late gate synchronizer as shown in following figure. The synchronizer operates by performing two separate integrations of the incoming signal energy over two different ($T-d$) second portions of a symbol

interval. The first integration (the early gate) begins from zero to $T-d$. The second integration (the late gate) begins from d to T . The difference in the absolute values of the outputs of these two integrations Y_x and Y_z is a measure of the receiver's symbol timing error, and it can be fed back to the loop's timing reference to correct loop timing.



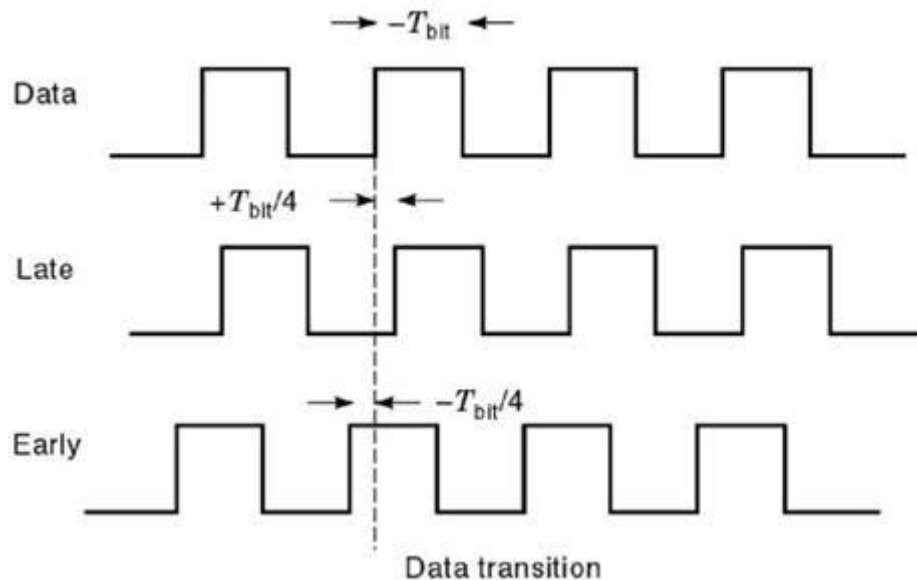


Fig shows the block diagram of early late synchronizer. This type of synchronizer does not need any zero crossings. We know that in the matched filter the output is maximum at the sampling time. Actually the sampling time is obtained by bit synchronization. The prefix allows the time for bit sync acquisition and indicates the beginning of a frame. After the prefix, there is one more codeword. It indicates the start of message. The message bits follow the special codeword. To maintain the synchronization of message bits, the sync words are inserted periodically in the bit stream.

Disadvantages:

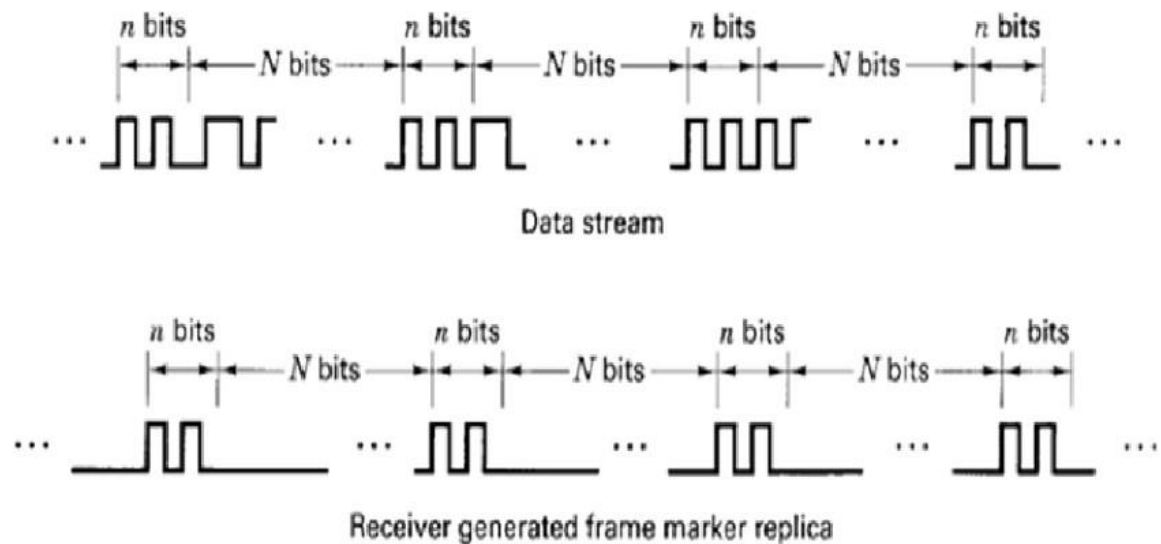
If there is a long string of 1's or 0's, then $y(t)$ has no zero crossings and synchronization may be lost. If zero crossing of $y(t)$ are not placed at integer multiples of T_b , the synchronization suffers from timing jitter.

Frame synchronization

All digital data streams have some sort of frame structure. The data streams is organized into uniformly sized groups of bit as shown previously. Computer data are typically organized into words of some number of 8 bit bytes. To receive the incoming data stream, the receiver needs to be synchronized with the data stream's frame structure.

Frame synchronization is usually accomplished with the aid of some special signaling procedure from the transmitter.

The simplest frame synchronization aid is the frame marker shown in following figure. The frame marker is a single bit or a short pattern of bits that the transmitter injects periodically into the data stream. The receiver, having achieved data synchronization, correlates the known pattern with the incoming data stream at the known injection interval.



The advantage of the frame marker is its simplicity. Even a single bit can suffice as a frame marker if a sufficient number of correlations are accumulated before deciding whether or not the system has achieved synchronization.

The major drawback is that the sufficient number may be very large and thus the expected time required to acquire synchronization would be long. Therefore, frame markers are most useful in systems that transmit data continuously, like many telephony and computer links.

Network synchronization

For systems using coherent modulation techniques, one direction communication such as most microwave links, land line links or fiber optics links, the synchronization architecture is performed at the receiver. For systems using noncoherent modulation techniques or that involves many user accessing a central communication node, such as many satellite communication systems, the synchronization is performed at a terminal transmitter. This means that the terminal transmitter parameters are modified to achieve synchronization rather than modifying the central nodes receiver parameters. This approach is used for time division multiple access (TDMA). In TDMA each user is allotted a segment of time in which to transmit its information. The terminal transmitter must be synchronized with the system in order for its transmitted burst data to arrive at the central node at the time when the node is prepared to receive the data.

Synchronization of the terminal transmitter also makes sense with n systems that combine signal processing at the central node with frequency division multiple access (FDMA). If the terminals precorrect their transmission to be synchronized with the central node, the node can use a fixed set of channel filters and a single timing reference for the processing of all channels. Otherwise, the node would require a separate time and frequency acquisition and tracking capability for each incoming channel. It seems clear that terminal transmitter synchronization is often more reasonable system approach to synchronizing network Transmitter synchronization procedures may be classified being either open loop or closed loop;-

Open loop network synchronization techniques do not depend on any measurement of the arriving signal parameters at the central node. The terminal precorrect its transmission based on stored knowledge of link parameters that have been provided by some external authority but many possible be modified by observations of a return signal from the central node. Open loop techniques depend on link parameters being accurately known. The main advantages of the open loop methods are that acquisition is fast the procedure can work without a return link and the amount of real time computation that is required is small. The disadvantages of the open loop methods are that they require knowledge of the required link parameters variation.

Closed loop network synchronization techniques involve measurements of the synchronization accuracy of the incoming transmission from the terminal upon that their arrival at the central node, and the return of the results of these measurements to the terminal via a return path. Thus, closed loop methods require a return path that provides a response to the terminal's transmissions, the ability in the terminal to recognize the response for what it is, and the ability in the terminal to modify the transmitter characteristics based on the response. The advantages of the closed loop are that no external source of knowledge is required for the system to work and the responses on the return link allow the system adapt easily and quickly to changing the characteristics. The disadvantages of closed loop methods are that they require a large amount of the real time processing and require two way links to every terminal.

Tracking and Acquisition in spread spectrum system :

- Spread spectrum techniques are essentially synchronous. The Pseudo Noise sequences generated at the receiver must be phase locked with PN sequence used in transmitter and the synchronization of spread spectrum will be considered in two parts, Acquisition and Tracking.
- Usually the problem of timing acquisition is solved via a two-step approach:
 - Initial code acquisition (coarse acquisition or coarse synchronization) which synchronizes the transmitter and receiver to within an uncertainty of T_c .
 - Code tracking which performs and maintains fine synchronization between the transmitter and receiver.
- Acquisition means initial synchronization of spread spectrum. It is also called coarse synchronization.
- Tracking starts after acquisition. It maintains the PN generator at the receiver in synchronous with the transmitter. It is also called fine synchronization.

Acquisition in spread spectrum system :

- Given the initial acquisition, code tracking is a relatively easy task and is usually accomplished by a delay lock loop (DLL).
- If the channel changes abruptly, the delay lock loop will lose track of the correct timing and

initial acquisition will be re-performed.

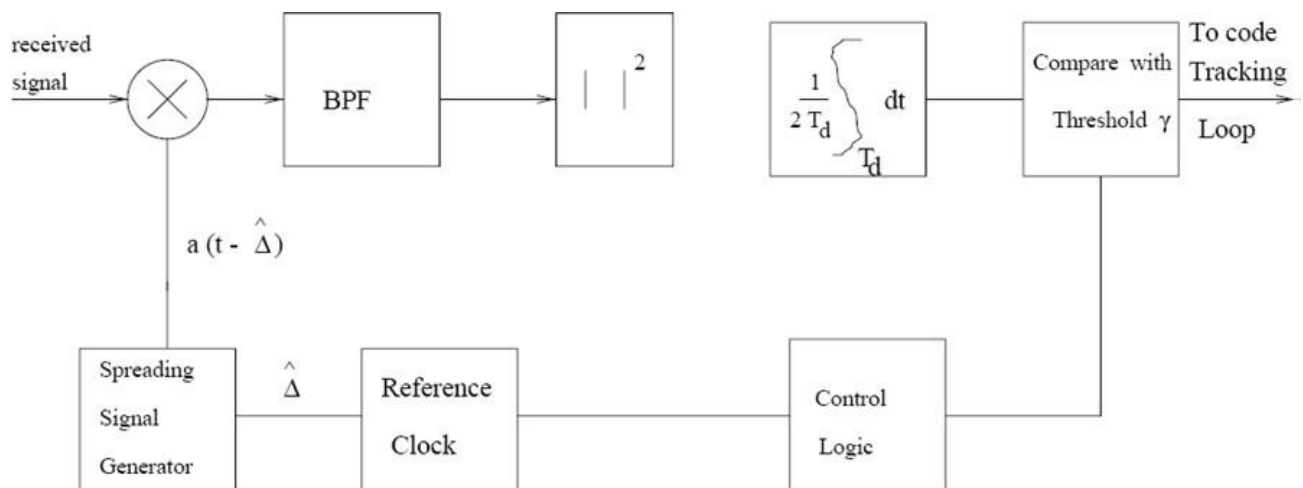
- Compared to code tracking, initial code acquisition in a spread spectrum system is usually very difficult.

Acquisition techniques

- Serial search
- Parallel Search

Serial search

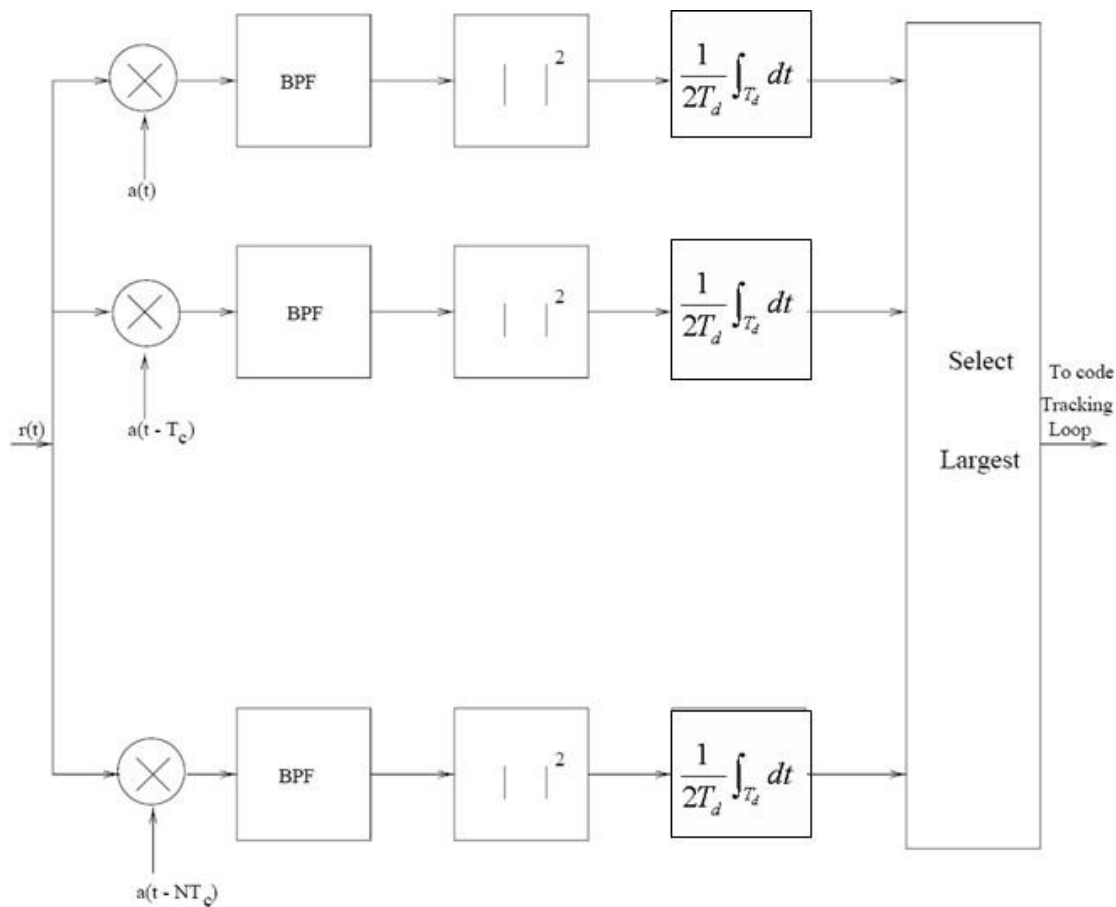
- In this method, the acquisition circuit attempts to cycle through and test all possible phases one by one (serially) as shown in Figure
- The circuit complexity for serial search is low.
- However, penalty time associated with a miss is large.
- Therefore we need to select a larger integration (dwell) time to reduce the miss probability.
- This, together with the serial searching nature, gives a large overall acquisition time (i.e., slow acquisition).
- The received signal is correlated with the generated PN sequence. This cross-correlation is performed over the time interval of NT_c .
- The output of the correlator is compared with a threshold. If it exceeds the threshold, then the required signal is obtained.
- If the threshold is not exceeded, then the PN generator output is advanced by half chip duration ($1/2 T_c$) and the correlation is performed.
- The output of the correlator is again compared with the threshold and the procedure is repeated.



Parallel search

- Unlike serial search, all the possible phases are tested simultaneously in the parallel search strategy as shown in Figure

- Obviously, the circuit complexity of the parallel search is high. The overall acquisition time is much smaller than that of the serial search.



The presence of noise causes two different kinds of errors in the acquisition process:

- A *false alarm* occurs when the integrator output exceeds the threshold for an incorrect hypothesized phase.
- A *miss* occurs when the integrator output falls below the threshold for a correct hypothesized phase.
- A false alarm will cause an incorrect phase to be passed to the code tracking loop which, as a result, will not be able to lock on to the DS-SS signal and will return the control back to the acquisition circuitry eventually.
- However, this process will impose severe time penalty to the overall acquisition time.
- On the other hand, a miss will cause the acquisition circuitry to neglect the current correct hypothesized phase.
- Therefore a correct acquisition will not be achieved until the next correct hypothesized phase comes around.
- The time penalty of a miss depends on acquisition strategy.

Tracking Techniques

- Matched filter energy detector
- Radiometer

Matched filter energy detector

The name is because the combination of the despreader and the integrator is basically an implementation of the matched filter for the spreading signal. It is based on the following basic working principle depicted in the Figure

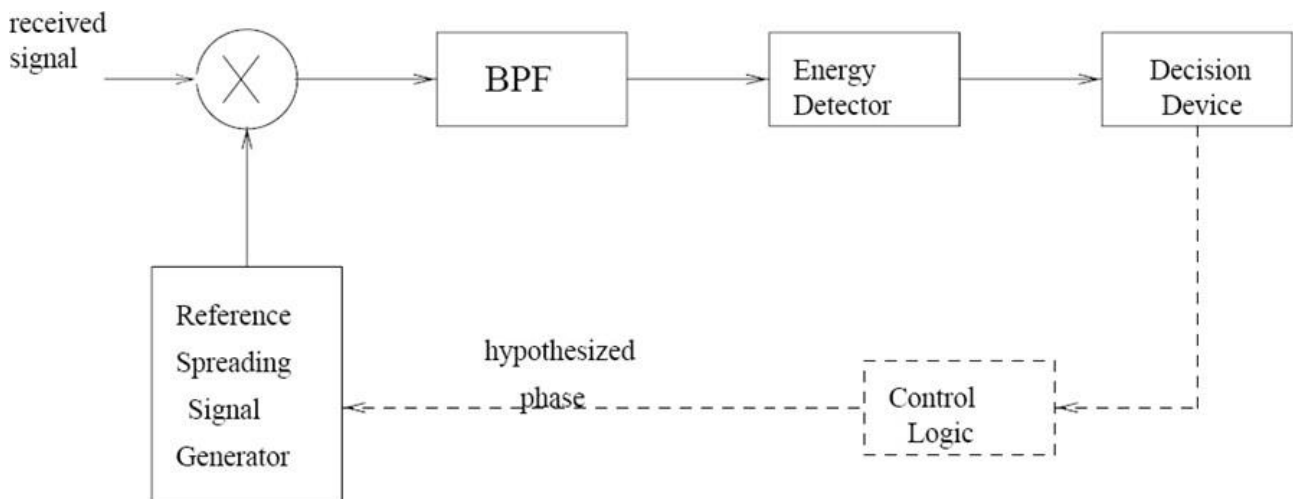


Figure. Generic Acquisition circuit

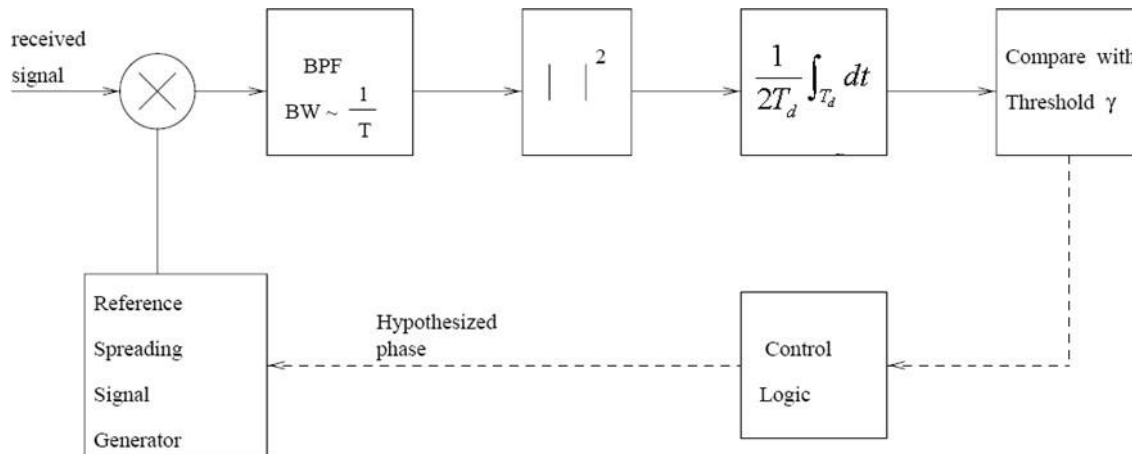
- The receiver hypothesizes a phase of the spreading sequence and attempts to despread the received signal using the hypothesized phase.
- If the hypothesized phase matches the sequence in the received signal, the wide-band spread spectrum signal will be despread correctly to give a narrowband data signal.
- Then a bandpass filter, with a bandwidth similar to that of the narrowband data signal, can be employed to collect the power of the despread signal.
- Since the hypothesized phase matches the received signal, the BPF will collect all the power of the despread signal.
- In this case, the receiver decides a coarse synchronization has been achieved and activates the tracking loop to perform fine synchronization.
- On the other hand, if the hypothesized phase does not match the received signal, the despreader will give a wideband output and the BPF will only be able to collect a small portion of the power of the despread signal.
- Based on this, the receiver decides this hypothesized phase is incorrect and other phases should be tried.

Radiometer

- As before, the receiver hypothesizes a phase of the spreading process.

- The despread signal is bandpass filtered with a bandwidth roughly equal to that of the narrowband data signal.
- The output of the bandpass filter is squared and integrated for a duration of T_d to detect the energy of the despread signal.

Figure. Acquisition circuit using radio meter



- The comparison between the matched filter energy detector and the radiometer brings out several important design issues for initial code acquisition circuits.
- **Dwell time**
 - The time needed to evaluate a single hypothesized phase of the spreading sequence.
- Neglecting the processing time for the other components in the acquisition circuit, the dwell time for the matched filter energy detector and the radiometer are determined by the integration times of the respective integrators in the matched filter energy detector and the radiometer.
- From the discussion above, the *dwell times* for the matched filter energy detector and the radiometer are T and T_d , respectively.
- In practice, we would like the dwell time to be as small as possible.
- For the matched filter energy detector, we cannot significantly reduce the integration time since the spreading sequence is usually designed to have a small out-of-phase autocorrelation magnitude, but the out-of-phase partial autocorrelation magnitude is not guaranteed to be small in standard sequence design methods.
- If we reduce the integration time significantly, the decision statistic would not be a function of the autocorrelation function, but rather a function of the partial autocorrelation function of the spreading sequence.
- Hence it would be difficult to distinguish whether or not we have a match between the hypothesized phase and the received signal, when noise is present.
- On the other hand, since the radiometer uses the BPF directly to detect whether there is a match

or not and the integrator is simply employed to collect energy, we do not need to integrate for the whole period of the sequence as long as we have enough energy.

- The integration time and hence the dwell time can be smaller than T , provided that the BPF can settle to its steady state output in a much shorter duration than T
- Another design issue, which is neglected in the simplified example before, is the effect of noise.

QUESTION BANK

UNIT III

PART-A
<ol style="list-style-type: none"> 1. Mention the levels of synchronisation required in coherent and non-coherent modulation systems. 2. List out the approaches used for carrier synchronization. 3. Illustrate the equation for time error in transmitter synchronization. 4. What is the need for carrier synchronisation in receivers? 5. Define timing phase and timing recovery w.r.t symbol synchronisation. 6. List the importance of symbol and bit synchronisation. 7. Give the characteristic of data aided synchronization. 8. Distinguish tracking and acquisition. 9. State the advantages of Early/late gate symbol synchroniser. 10. Compare and contrast open loop and closed symbol synchronisation. 11. Show how frame synchronisation can be achieved using marker bits.
PART-B
<ol style="list-style-type: none"> 1. Explain the principle of Costas loop and squaring loop method of carrier recovery with suitable sketch. (16 marks) 2. Justify with appropriate proof the need for using second order phase locked loop for synchronization applications. (16 marks) 3. Elucidate the operating principles of various open loop symbol synchronizing techniques with suitable examples. (16 marks) 4. Explain the working principle of early/late-gate based closed loop bit synchronizer. (16 marks) 5. Analyze how open loop and closed methods are employed for providing network synchronization in applications such as broadcast channels and microwave links. (16 marks) 6. Analyze the necessity of tracking and acquisition systems in spread spectrum communication transceivers. (16 marks)

UNIT V – SPREAD SPECTRUM MODULATION - SEC1313

Introduction:

Initially developed for military applications during II world war, that was less sensitive to intentional interference or jamming by third parties. Spread spectrum technology has blossomed into one of the fundamental building blocks in current and next-generation wireless systems.

Problem of radio transmission

Narrow band can be wiped out due to interference. To disrupt the communication, the adversary needs to do two things,

- (a) to detect that a transmission is taking place and
- (b) to transmit a jamming signal which is designed to confuse the receiver.

Solution

A spread spectrum system is therefore designed to make these tasks as difficult as possible.

Firstly, the transmitted signal should be difficult to detect by an adversary/jammer, i.e., the signal should have a low probability of intercept (LPI).

Secondly, the signal should be difficult to disturb with a jamming signal, i.e., the transmitted signal should possess an anti-jamming (AJ) property

Remedy

spread the narrow band signal into a broad band to protect against interference

In a digital communication system the primary resources are **Bandwidth** and **Power**. The study of digital communication system deals with efficient utilization of these two resources, but there are situations where it is necessary to sacrifice their efficient utilization in order to meet certain other design objectives.

For example to provide a form of secure communication (i.e. the transmitted signal is not easily detected or recognized by unwanted listeners) the bandwidth of the transmitted signal is increased in excess of the minimum bandwidth necessary to transmit it. This requirement is catered by a technique known as “**Spread Spectrum Modulation**”.

The primary advantage of a Spread – Spectrum communication system is its ability to reject ‘**Interference**’ whether it be the unintentional or the intentional interference.

The definition of Spread – Spectrum modulation may be stated in two parts.

1. Spread Spectrum is a mean of transmission in which the data sequence occupies a BW (Bandwidth) in excess of the minimum BW necessary to transmit it.
2. The Spectrum Spreading is accomplished before transmission through the use of a code that is independent of the data sequence. The Same code is used in the receiver to despread the received signal so that the original data sequence may be recovered.

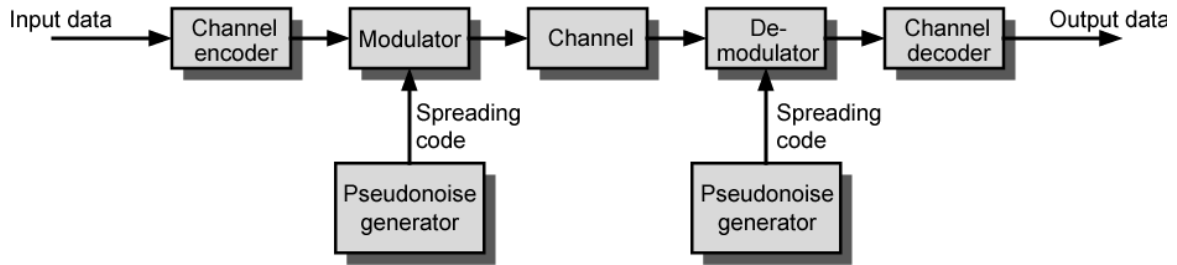


Fig. 5.1Block diagram for spread spectrum communication

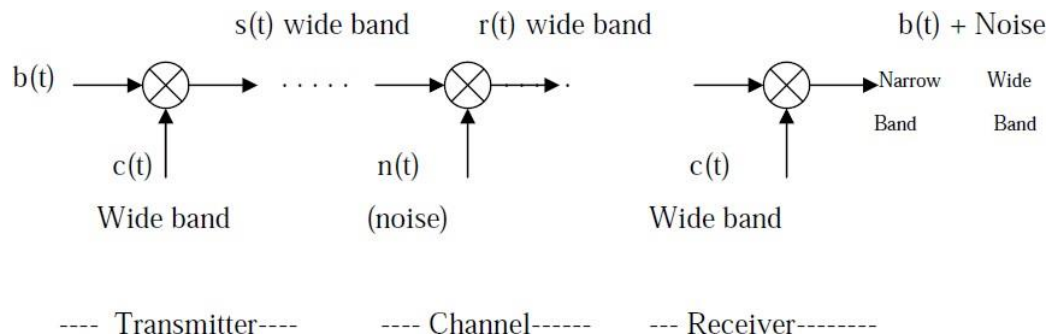


Fig:5.2 Spread spectrum technique.

$b(t)$ = Data Sequence to be transmitted (Narrow Band);

$c(t)$ = Wide Band code ;

$$s(t) = c(t) * b(t) - (\text{wide Band})$$

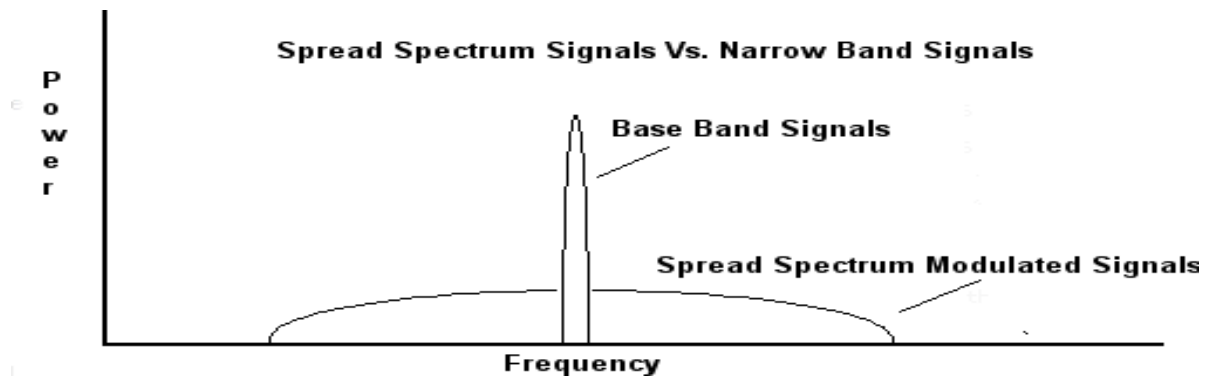


Fig:5.3 Spectrum of signal before & after spreading

PSUEDO-NOISE SEQUENCE:

Generation of PN sequence:

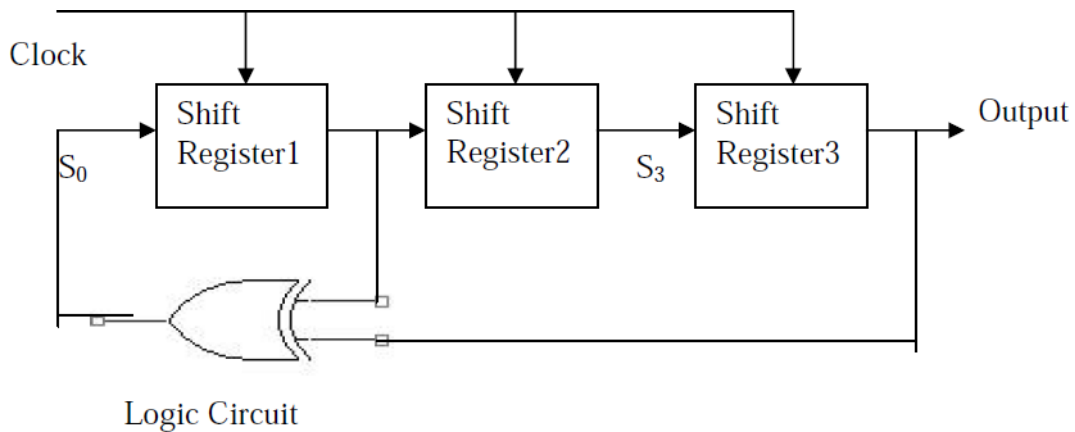


Fig: 5.4 Maximum-length sequence generator for n=3

A feedback shift register is said to be Linear when the feedback logic consists of entirely mod-2-adds (Ex-or gates). In such a case, the zero state is not permitted. The period of a PN sequence produced by a linear feedback shift register with 'n' flip flops cannot exceed $2^n - 1$.

When the period is exactly $2^n - 1$, the PN sequence is called a '**maximum length sequence**' or '**m-sequence**'.

Example1: Consider the linear feedback shift register shown in above figure

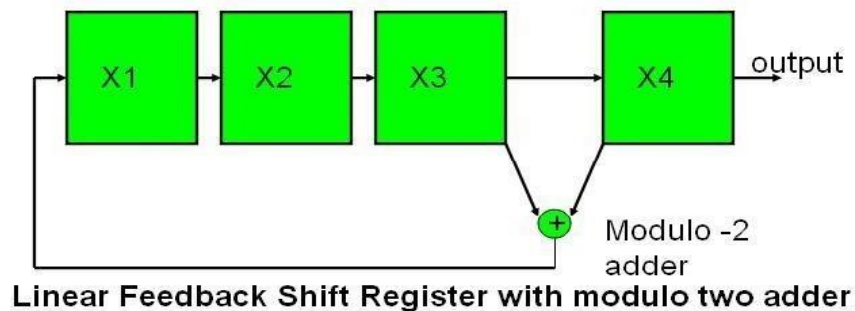
Involve three flip-flops. The input s_0 is equal to the mod-2 sum of S_1 and S_3 . If the initial state of the shift register is 100. Then the succession of states will be as follows.

100,110,011,011,101,010,001,100

The output sequence (output S_3) is therefore. 00111010 Which repeats itself with period $2^3 - 1 = 7$ ($n=3$). Maximal length codes are commonly used PN codes In binary shift register, the maximum length sequence is

$$\underline{N = 2^m - 1}$$

chips, where m is the number of stages of flip-flops in the shift register.



At each clock pulse

- Contents of register shifts one bit right.
- Contents of required stages are modulo 2 added and fed back to input.

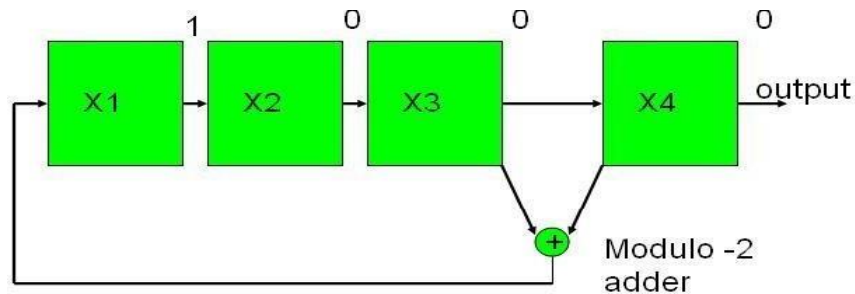


Fig: 5.5 Initial stages of Shift registers 1000

Let initial status of shift register be 1000

1	0	0	0
0	1	0	0
0	0	1	0
1	0	0	1
1	1	0	0
0	1	1	0
1	0	1	1
0	1	0	1
1	0	1	0
1	1	0	1
1	1	1	0
1	1	1	1
0	1	1	1
0	0	1	1
0	0	0	1
1	0	0	0

- We can see for shift Register of length $m=4$.
At each clock the change in state of flip-flop is shown.

- Feed back function is modulo two of X_3 and X_4 .

- After 15 clock pulses the sequence repeats.

Output sequence is

0 0 0 1 0 0 1 1 0 1 0 1 1 1 1

Properties of PN Sequence

Randomness of PN sequence is tested by following properties

1. Balance property
2. Run length property
3. Autocorrelation property

1. Balance property

In each Period of the sequence, number of binary ones differ from binary zeros by at most one digit.

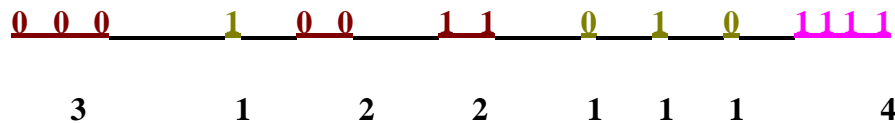
Consider output of shift register 0 0 0 1 0 0 1 1 0 1 0 1 1 1 1 Seven zeros and eight ones -meets balance condition.

2. Run length property

Among the runs of ones and zeros in each period, it is desirable that about one half the runs of each type are of length 1, one-fourth are of length 2 and one-eighth are of length 3 and so-on.

Consider output of shift register

Number of runs = 8



3. Auto correlation property

Auto correlation function of a maximal length sequence is periodic and binary valued. Autocorrelation sequence of binary sequence in polar format is given by

$$R_c(k) = \frac{1}{N} \sum_{n=1}^N c_n c_{n-k}$$

Where N is length or the period of the sequence, k is the lag of auto correlation function.

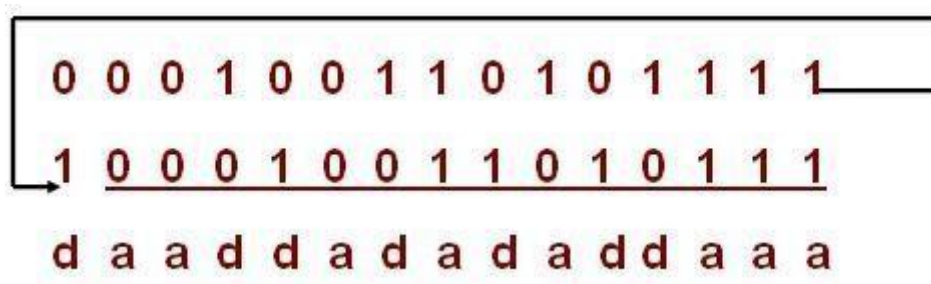
$$R_c(k) = \begin{cases} 1 & \text{if } k = 1N \\ -\frac{1}{N} & \text{if } k \neq 1N \end{cases}$$

Where 1 is any Integer. We can also state the auto correlation function is

$$R_c(k) = \frac{1}{N}$$

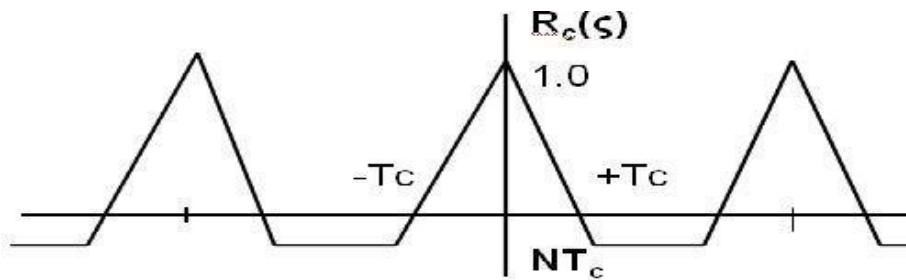
{ No. of agreements – No. of disagreements in comparison of one full period }

Consider output of shift register for l=1



$$R_c(k) = \frac{1}{15} \quad 7 - 8 = -\frac{1}{15}$$

Yields PN autocorrelation as



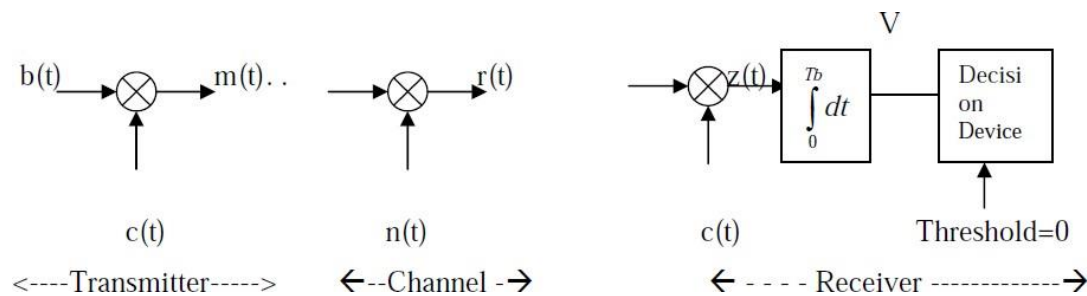
PN autocorrelation function.

Range of PN Sequence Lengths

Length Of Shift Register, m	PN Sequence Length,
7	127
8	255
9	511
10	1023
11	2047
12	4095
13	8191
17	131071
19	524287

Notion of Spread Spectrum:

An important attribute of Spread Spectrum modulation is that it can provide protection against externally generated interfering signals with finite power. Protection against jamming (interfering) waveforms is provided by purposely making the information – bearing signal occupy a BW far in excess of the minimum BW necessary to transmit it. This has the effect of making the transmitted signal a noise like appearance so as to blend into the background. Therefore Spread Spectrum is a method of ‘camouflaging’ the information – bearing signal.



Let $\{ b_K \}$ denotes a binary data sequence.

$\{ c_K \}$ denotes a PN sequence.

$b(t)$ and $c(t)$ denotes their NRZ polar representation respectively.

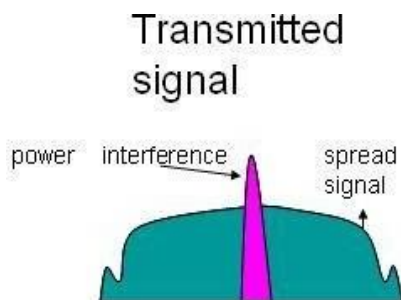
The desired modulation is achieved by applying the data signal $b(t)$ and PN signal $c(t)$ to a product modulator or multiplier. If the message signal $b(t)$ is narrowband and the PN sequence signal $c(t)$ is wide band, the product signal $m(t)$ is also wide band. The PN sequence performs the role of a **‘Spreading Code’**.

For base band transmission, the product signal $m(t)$ represents the transmitted signal. Therefore $m(t) = c(t).b(t)$

The received signal $r(t)$ consists of the transmitted signal $m(t)$ plus an additive interference noise $n(t)$, Hence

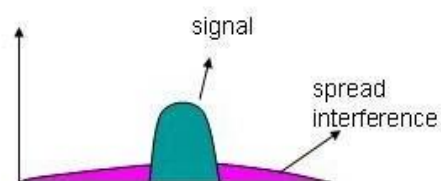
$$r(t) = m(t) + n(t)$$

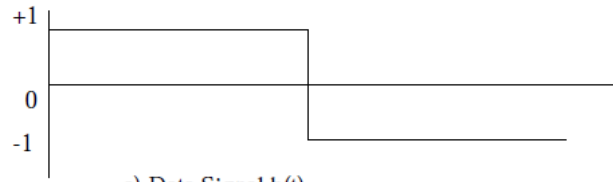
$$= c(t).b(t) + n(t)$$



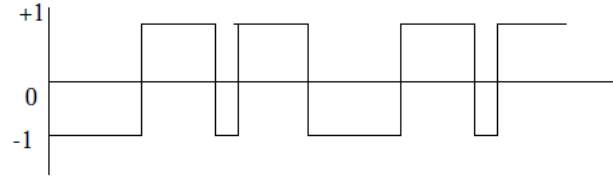
Narrow band interference

Received signal after demodulation

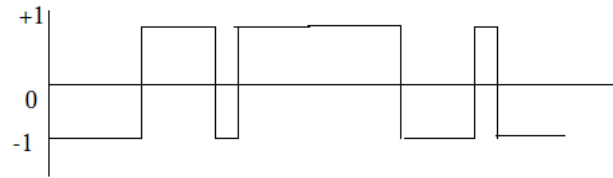




a) Data Signal b(t)



b) Spreading Code c(t)



c) Product signal or base band transmitted signal m(t)

To recover the original message signal $b(t)$, the received signal $r(t)$ is applied to a demodulator that consists of a multiplier followed by an integrator and a decision device. The multiplier is supplied with a locally generated PN sequence that is exact replica of that used in the transmitter. The multiplier output is given by

$$\begin{aligned} Z(t) &= r(t) \cdot c(t) \\ &= [b(t) \cdot c(t) + n(t)] \cdot c(t) = c^2(t) \cdot b(t) + c(t) \cdot n(t) \end{aligned}$$

The data signal $b(t)$ is multiplied twice by the PN signal $c(t)$, where as unwanted signal $n(t)$ is multiplied only once. But $c^2(t) = 1$, hence the above equation reduces to

$$Z(t) = b(t) + c(t) \cdot n(t)$$

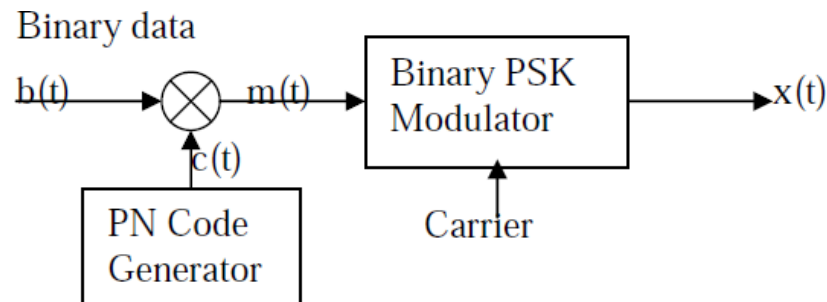
Now the data component $b(t)$ is narrowband, where as the spurious component $c(t)n(t)$ is wide band. Hence by applying the multiplier output to a base band (low pass) filter most of the power in the spurious component $c(t)n(t)$ is filtered out. Thus the effect of the interference $n(t)$ is thus significantly reduced at the receiver output.

The integration is carried out for the bit interval $0 \leq t \leq T_b$ to provide the sample value V . Finally, a decision is made by the receiver.

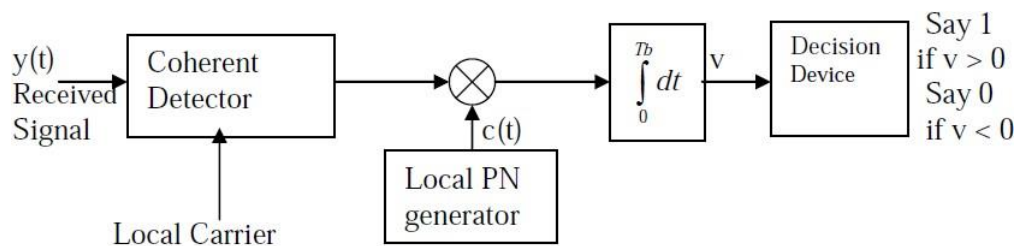
If $V > \text{Threshold Value '0'}$, say binary symbol '1' If $V < \text{Threshold Value '0'}$, say binary symbol '0'

Direct – Sequence Spread Spectrum with coherent binary Phase shift

Keying:-



a) Transmitter



b) Receiver

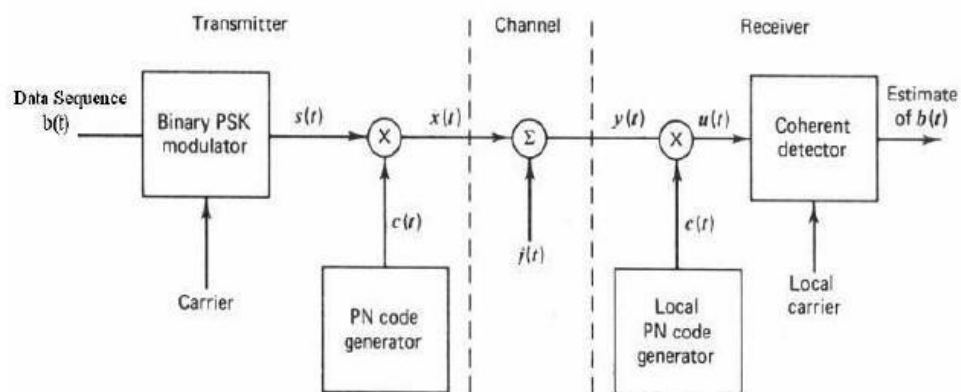


Fig: model of direct – sequence spread binary PSK system(alternative form)

To provide band pass transmission, the base band data sequence is multiplied by a Carrier by means of shift keying. Normally binary phase shift keying (PSK) is used because of its advantages. The transmitter first converts the incoming binary data sequence $\{b_k\}$ into an NRZ waveform $b(t)$, which is followed by two stages of modulation.

The first stage consists of a multiplier with data signal $b(t)$ and the PN signal $c(t)$ as inputs. The output of multiplier is $m(t)$ is a wideband signal. Thus a narrow – band data sequence is transformed into a noise like wide band signal.

The second stage consists of a binary Phase Shift Keying (PSK) modulator. Which converts base band signal $m(t)$ into band pass signal $x(t)$. The transmitted signal $x(t)$ is thus a direct – sequence spread binary PSK signal. The phase modulation $\theta(t)$ of $x(t)$ has one of the two values ‘0’ and ‘ π ’ (180°) depending upon the polarity of the message signal $b(t)$ and PN signal $c(t)$ at time t .

Polarity of PN & Polarity of PN signal both +, + or - - Phase ‘0’

Polarity of PN & Polarity of PN signal both +, - or - + Phase ‘ π ’

Polarity of data sequence $b(t)$			
		+	-
Polarity of PN sequence $C(t)$	+	0	π
	-	π	0

The receiver consists of two stages of demodulation.

In the first stage the received signal $y(t)$ and a locally generated carrier are applied to a coherent detector (a product modulator followed by a low pass filter), Which converts band pass signal into base band signal.

The second stage of demodulation performs Spectrum despread by multiplying the output of low-pass filter by a locally generated replica of the PN signal $c(t)$, followed by integration over a bit interval T_b and finally a decision device is used to get binary sequence.

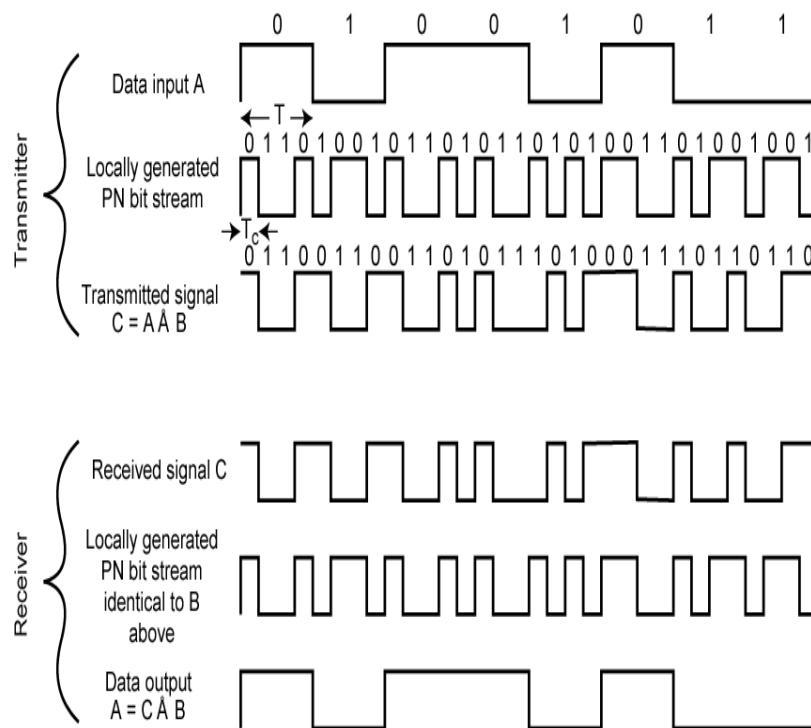


Fig : Direct Sequence Spread Spectrum Example

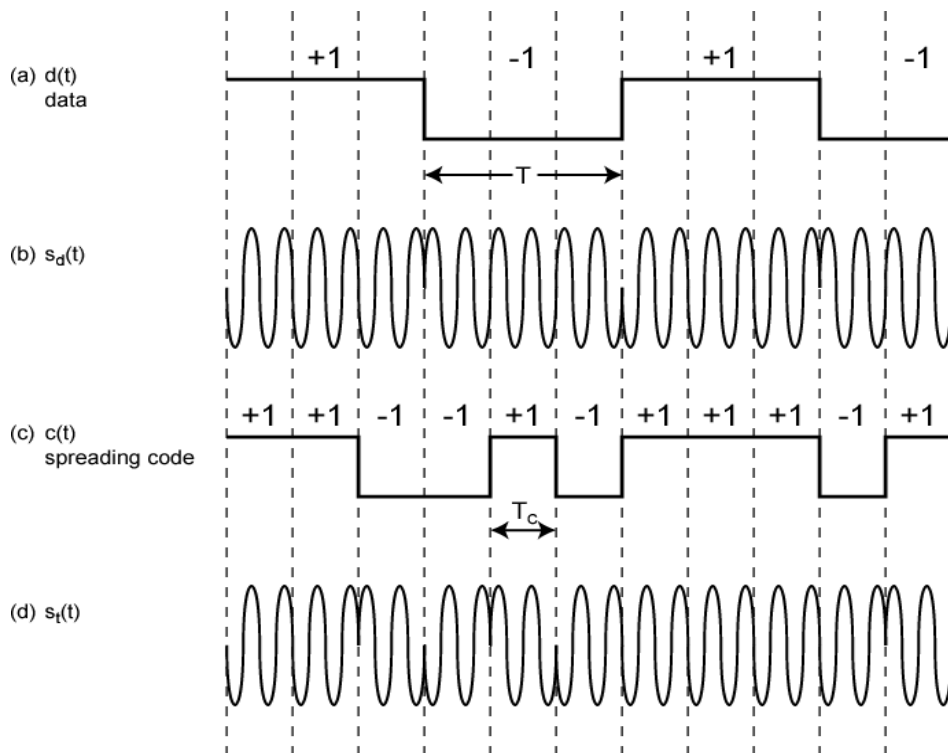


Fig : 5.5 Direct Sequence Spread Spectrum Using BPSK Example

Signal Space Dimensionality and Processing Gain

- Fundamental issue in SS systems is how much protection spreading can provide against interference.
- SS technique distribute low dimensional signal into large dimensional signal space (hide the signal).
- Jammer has only one option; to jam the entire space with fixed total power or to jam portion of signal space with large power.

Consider set of orthonormal basis functions;

$$\varphi_k(t) = \begin{cases} \frac{2}{T_c} \cos 2\pi f_c t & 0 \leq t \leq T_c \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi_k(t) = \begin{cases} \frac{2}{T_c} \sin 2\pi f_c t & 0 \leq t \leq T_c \\ 0 & \text{otherwise} \end{cases} \quad k = 0, 1, \dots, N-1$$

Where T_c is chip duration, N is number of chips per bit. Transmitted

signal $x(t)$ for the interval of an information bit is

$$x(t) = c(t)s(t)$$

$$\varphi_k(t) = \pm \frac{2}{T_c} c(t) \cos 2\pi f_c t$$

$$\overline{E_b} \quad N-1$$

$$s(t) = \sum_{k=0}^{N-1} c_k \phi_k(t) \quad 0 \leq t \leq T_b$$

where, E_b is signal energy per bit.

PN Code sequence $\{c_0, c_1, \dots, c_{N-1}\}$ with $c_k = \pm 1$, Transmitted signal $x(t)$ is therefore N dimensional and requires N orthonormal functions to represent it. $j(t)$ represent interfering signal (jammer). As said jammer tries to place all its available energy in exactly same N dimension signal space. But jammer has no knowledge of signal phase. Hence tries to place equal energy in two phase coordinates that is cosine and sine.

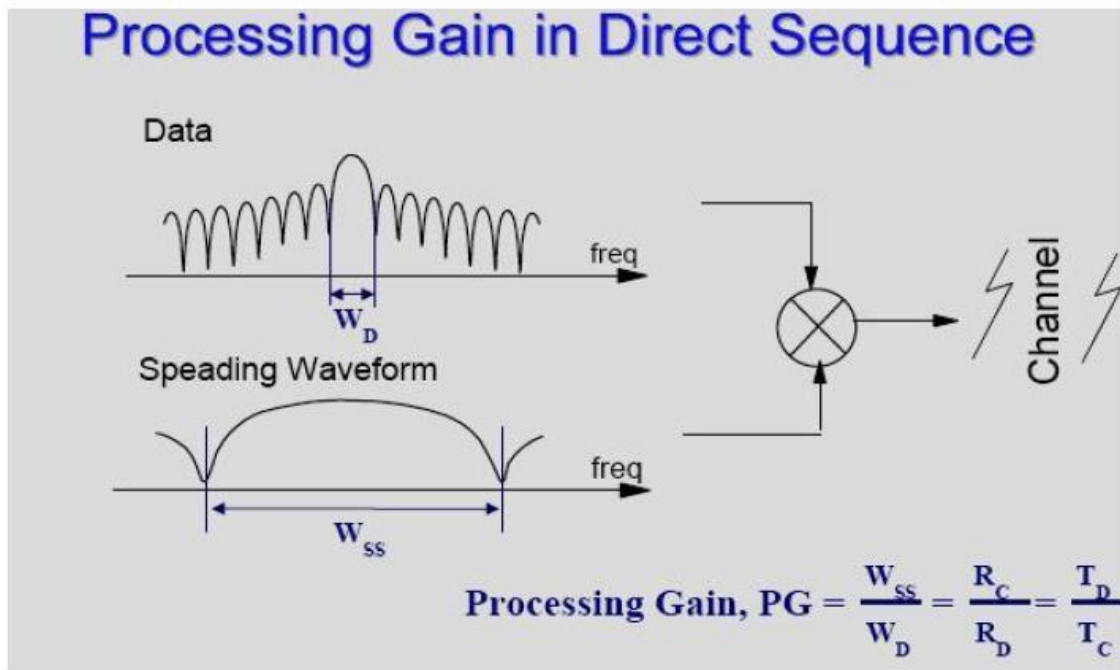
$$(SNR)_0$$

Expressing SNR in decibels

$$10 \log_{10}(SNR)_0 = 10 \log_{10}(SNR)_i + 3 + 10 \log_{10} PG, dB$$

$$\text{Where } P = \frac{E_b}{T_c}$$

3db term on right side accounts for gain in SNR due to coherent detection. Last term accounts for gain in SNR by use of spread spectrum. **PG is called Processing Gain.**



The bandwidth of PN sequence $c(t)$, of main lobe is W_c

To calculate probability of error, we consider output component v of coherent detector as sample value of random variable

$$V = \pm \sqrt{E_b} + V_{cj}$$

E_b is signal energy per bit and V_{cj} is noise component

Decision rule is, if detector output exceeds a threshold of zero volts; received bit is symbol 1 else decision is favored for zero.

- Average probability of error P_e is nothing but conditional probability which depends on random variable V_{cj} .
- As a result receiver makes decision in favor of symbol 1 when symbol 0 transmitted and vice versa
- Random variable V_{cj} is sum of N such random variables. Hence for Large N it can assume Gaussian distribution.
- As mean and variance has already been discussed, zero mean and variance

Probability of error can be calculated from simple formula for DS/BPSK system

Example1

A pseudo random sequence is generated using a feed back shift register of length $m=4$. The chip rate is 107 chips per second. Find the following a) PN sequence length b) Chip duration of PN sequence c) PN sequence period

Solution

- a) Length of PN sequence $N = 2^m - 1 = 2^4 - 1 = 15$
b) Chip duration $T_c = 1/\text{chip rate} = 1/107 = 0.1\mu\text{sec}$
c) PN sequence period $T = NT_c$
 $= 15 \times 0.1\mu\text{sec} = 1.5\mu\text{sec}$

Example2

A direct sequence spread binary phase shift keying system uses a feedback shift register

of length 19 for the generation of PN sequence. Calculate the processing gain of the system.

Solution

Given length of shift register = m = 19 Therefore

$$\text{length of PN sequence } N = 2^m - 1 \\ = 2^{19} - 1$$

$$\text{Processing gain } PG = T_b/T_c = N \text{ in db} = 10 \log_{10} N = 10 \log_{10} (2^{19}) = 57 \text{db}$$

Example3

A Spread spectrum communication system has the following parameters. Information bit duration $T_b = 1.024$ msecs and PN chip duration of 1µsecs. The average probability of error of system is not to exceed 10^{-5} . calculate a) Length of shift register b) Processing gain c) jamming margin

Solution

$$\text{Processing gain } PG = N = T_b/T_c = 1024$$

corresponding length of shift register m = 10

In case of coherent BPSK For Probability of error 10^{-5} . [Referring to error function table] $E_b/N_0 = 10.8$

Therefore jamming margin

$$\text{jamming margin (dB)} = \text{Processing gain (dB)} - 10 \log_{10} \frac{E_b}{N_0} \\ \text{jamming margin} = 10 \log_{10} PG_{dB} - 10 \log_{10} \frac{E_b}{N_0} \\ \text{dB} \quad \quad \quad \min$$

min

$$jamming\ margin_{dB} = 10 \log_{10} 1024 - 10 \log_{10} 10.8$$

$$jamming\ margin_{dB} = 30.10 - 10.33 = 19.8\ dB$$

Frequency – Hop Spread Spectrum:

In a frequency – hop Spread – Spectrum technique, the spectrum of data modulated carrier is widened by changing the carrier frequency in a pseudo – random manner. The type of spread – spectrum in which the carrier hops randomly from one frequency to another is called **Frequency – Hop (FH) Spread Spectrum.**

Since frequency hopping does not covers the entire spread spectrum instantaneously. We are led to consider the rate at which the hop occurs. Depending upon this we have two types of frequency hop.

1. **Slow frequency hopping:-** In which the symbol rate R_s of the MFSK signal is an integer multiple of the hop rate R_h . That is several symbols are transmitted on each frequency hop.
2. **Fast – Frequency hopping:-** In which the hop rate R_h is an integral multiple of the MFSK symbol rate R_s . That is the carrier frequency will hop several times during the transmission of one symbol. A common modulation format for frequency hopping system is that of M- ary frequency – shift – keying (MFSK).

Slow frequency hopping:-

Fig Shows the block diagram of an FH / MFSK transmitter, which involves frequency modulation followed by mixing.

The incoming binary data are applied to an M-ary FSK modulator. The resulting modulated wave and the output from a digital frequency synthesizer are then applied to a mixer that consists of a multiplier followed by a band – pass filter. The filter is designed to select the sum frequency component resulting from the multiplication process as the transmitted signal. An ‘k’ bit segments of a PN sequence drive the frequency synthesizer, which enables the carrier frequency to hop over 2^k distinct values. Since frequency synthesizers are unable to maintain phase coherence over successive hops, most frequency hops spread spectrum communication system use non coherent M-ary modulation system.

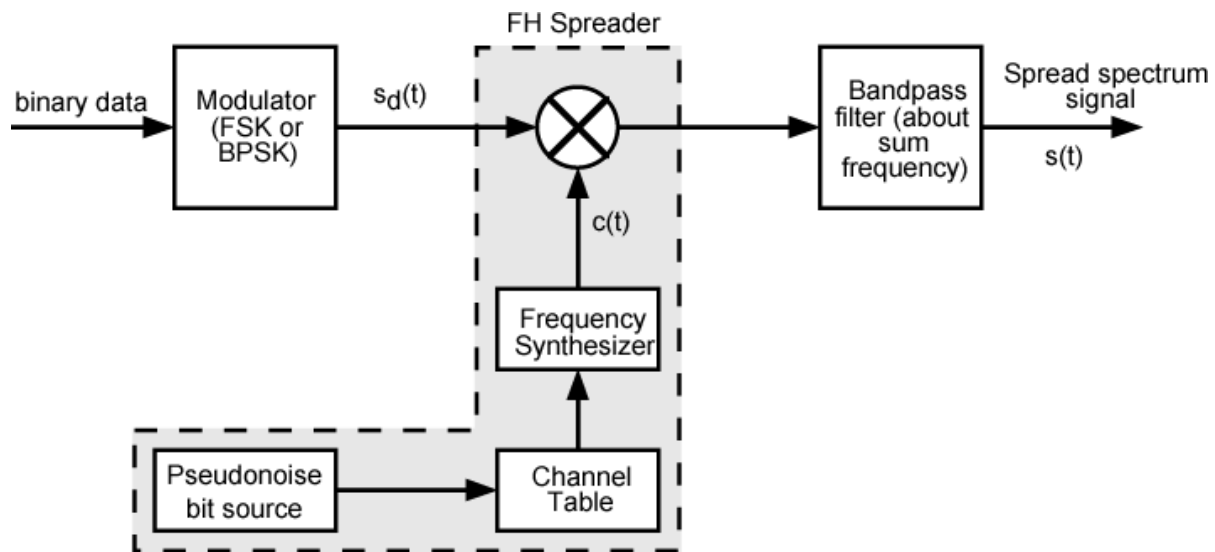


Fig 5.6:- Frequency hop spread transmitter

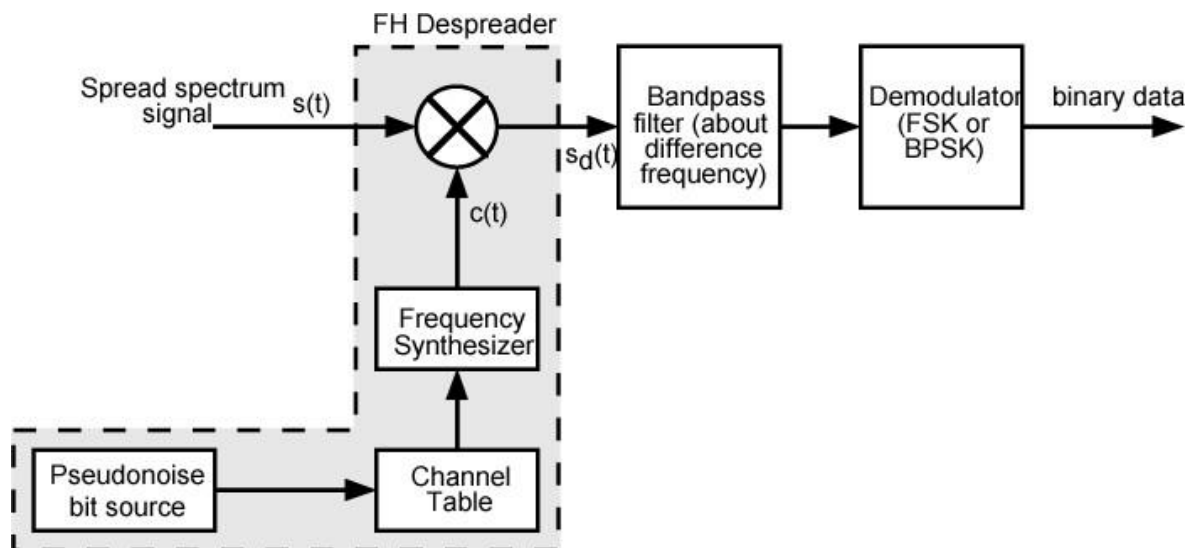


Fig 5.7 :- Frequency hop spread receiver

In the receiver the frequency hopping is first removed by mixing the received signal with the output of a local frequency synthesizer that is synchronized with the transmitter. The resulting output is then band pass filtered and subsequently processed by a non coherent M-ary FSK demodulator. To implement this M-ary detector, a bank of M non coherent matched filters, each of which is matched to one of the MFSK tones is used. By selecting the largest filtered output, the original transmitted signal is estimated.

An individual FH / MFSK tone of shortest duration is referred as a chip. The chip rate R_c for an FH / MFSK system is defined by

$$R_c = \text{Max}(R_h, R_s)$$

Where R_h is the hop rate and R_s is Symbol Rate

In a slow rate frequency hopping multiple symbols are transmitted per hop. Hence each symbol of a slow FH / MFSK signal is a chip. The bit rate R_b of the incoming binary data. The symbol rate R_s of the MFSK signal, the chip rate R_c and the hop rate R_h are related by

$$R_c = R_s = R_b / k \geq R_h \text{ where}$$

$$k = \log_2 M$$

Fast frequency hopping:-

A fast FH / MFSK system differs from a slow FH / MFSK system in that there are multiple hops per m-ary symbol. Hence in a fast FH / MFSK system each hop is a chip.

Fast Frequency Hopping	Slow Frequency Hopping
Several frequency hops Per modulation	Several modulation symbols per hop
Shortest uninterrupted waveform in the system is that of hop	Shortest uninterrupted waveform in the system is that of data symbol
Chip duration =hop duration	Chip duration=bit duration.

The following figure illustrates the variation of the frequency of a slow FH/MFSK signal with time for one complete period of the PN sequence. The period of the PN sequence is $2^4 - 1 = 15$.

The FH/MFSK signal has the following parameters:

Number of bits per MFSK symbol $K = 2$. Number of MFSK tones $M = 2^K = 4$

Length of PN segment per hop $k = 3$; Total number of frequency hops $2^k = 8$

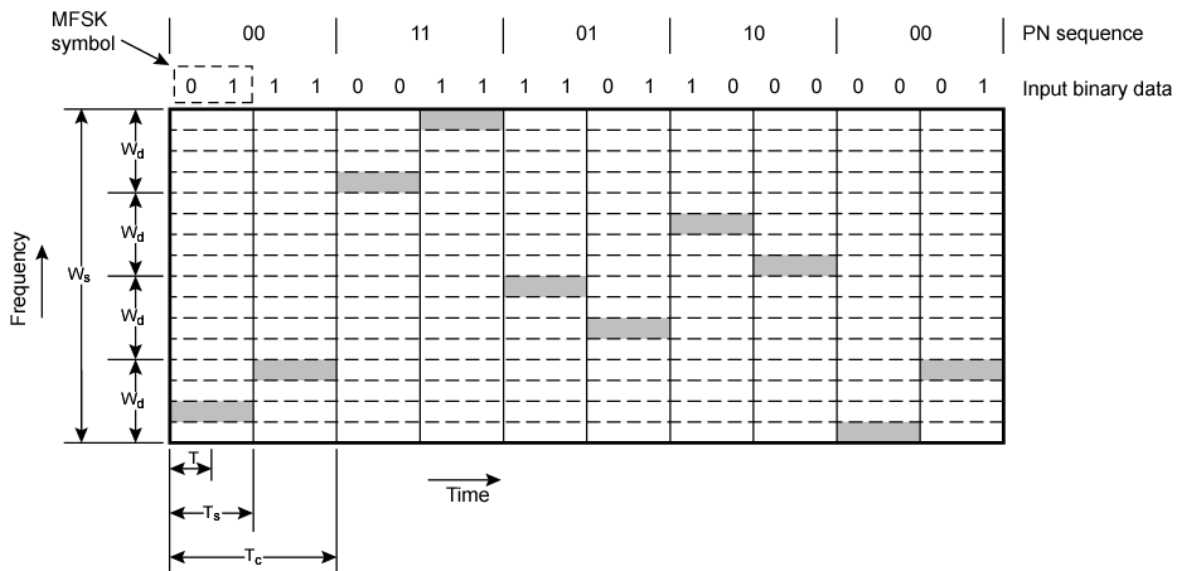
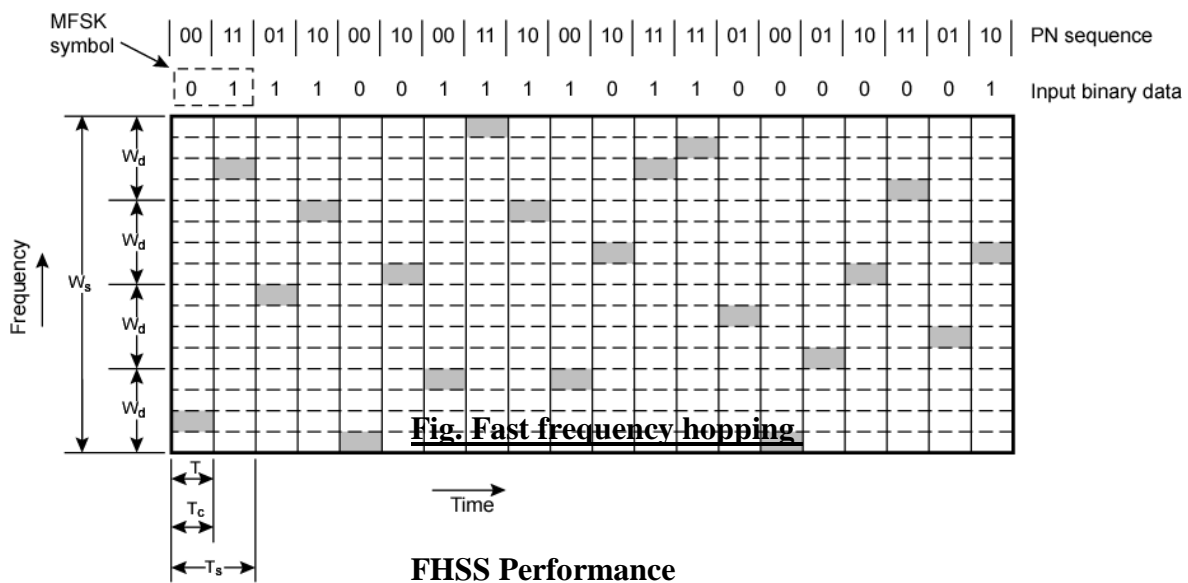


Fig.5.8 Slow frequency hopping

The following figure illustrates the variation of the transmitted frequency of a fast FH/MFSK signal with time.

The signal has the following parameters:

Number of bits per MFSK symbol $K = 2$. Number of MFSK tones $M = 2^K = 4$ Length of PN segment per hop $k = 3$; Total number of frequency hops $2^k = 8$



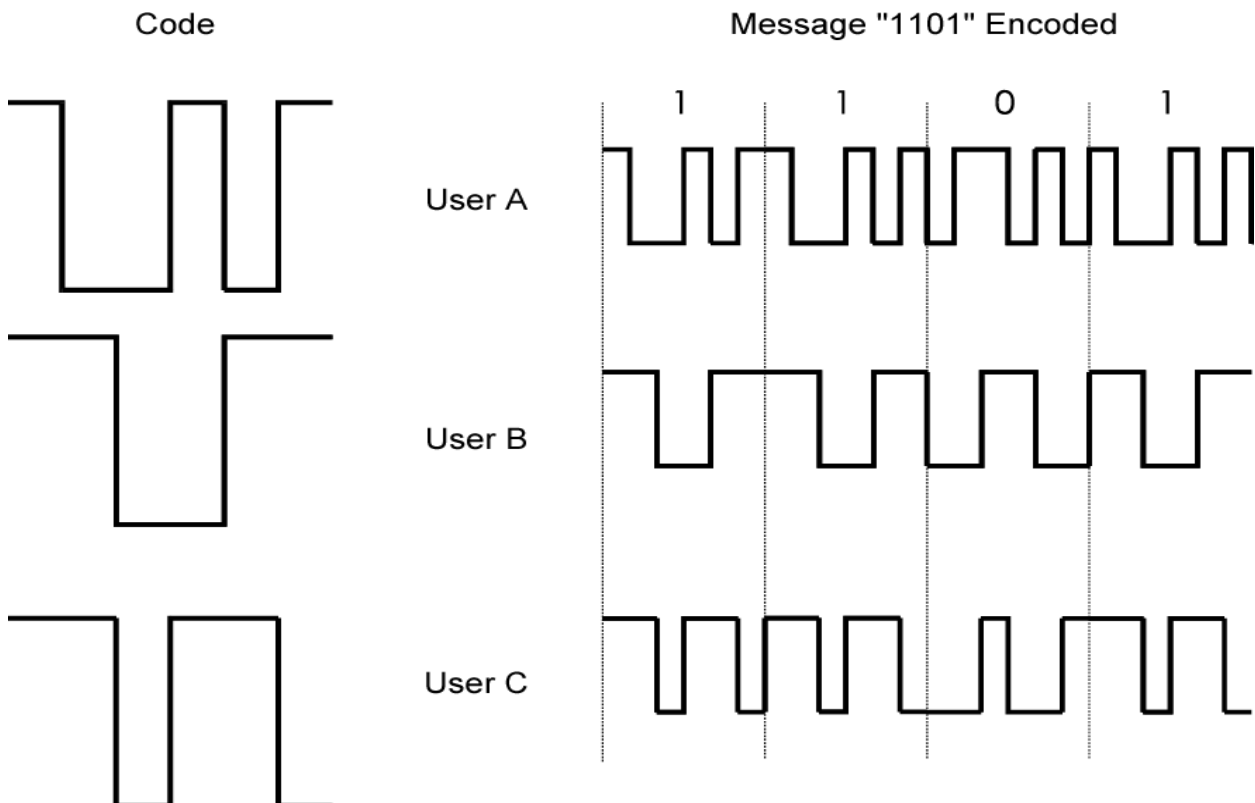
Considerations:

- Typically large number of frequencies used
 - Improved resistance to jamming

Code Division Multiple Access (CDMA):

- Multiplexing Technique used with spread spectrum
- Start with data signal rate D
 - Called bit data rate
- Break each bit into k chips according to fixed pattern specific to each user
 - User's code
- New channel has chip data rate kD chips per second
- E.g. $k=6$, three users (A,B,C) communicating with base receiver R
- Code for A = $\langle 1, -1, -1, 1, -1, 1 \rangle$
- Code for B = $\langle 1, 1, -1, -1, 1, 1 \rangle$
- Code for C = $\langle 1, 1, -1, 1, 1, -1 \rangle$

CDMA Example:



- Consider A communicating with base
- Base knows A's code
- Assume communication already synchronized
- A wants to send a 1
 - Send chip pattern $\langle 1, -1, -1, 1, -1, 1 \rangle$
 - A's code
- A wants to send 0
 - Send chip pattern $\langle -1, 1, 1, -1, 1, -1 \rangle$
 - Complement of A's code
- Decoder ignores other sources when using A's code to decode
 - Orthogonal codes
 -

CDMA for DSSS:

- n users each using different orthogonal PN sequence
- Modulate each users data stream
 - Using BPSK
- Multiply by spreading code of user

CDMA in a DSSS Environment:

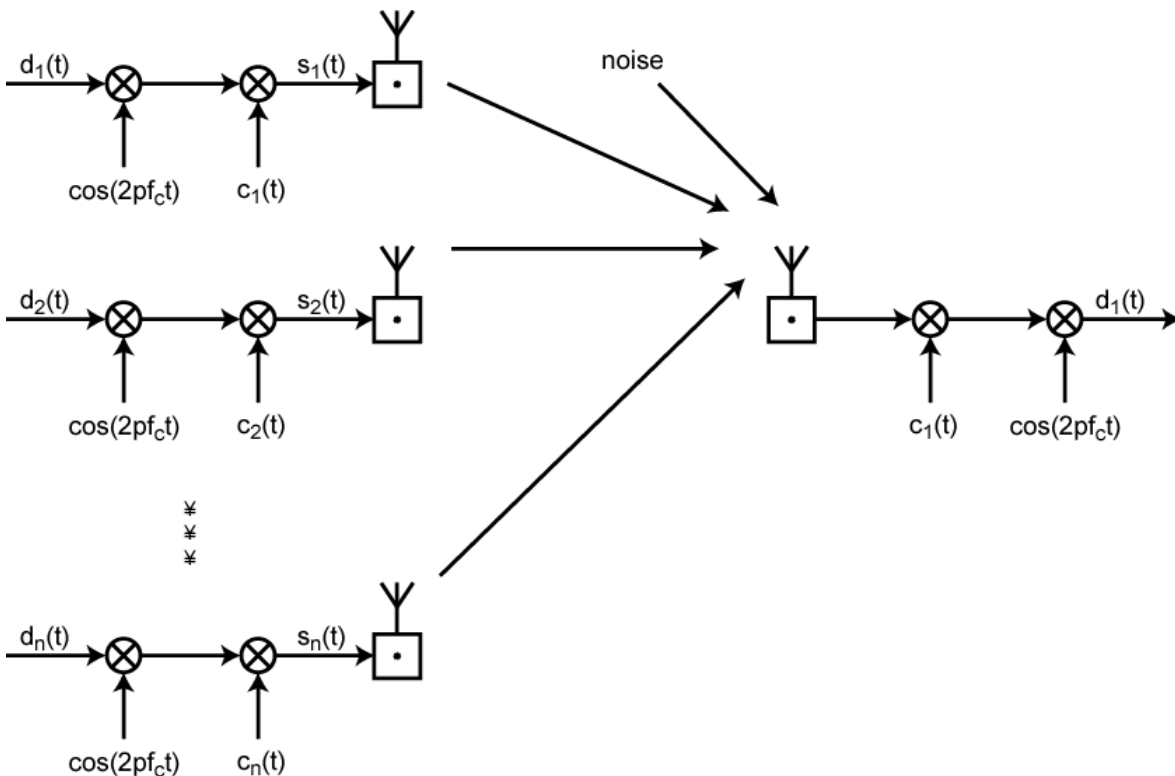


Fig 5.9 CDMA in a DSSS Environment

DS-CDMA has various features and benefits. First, DSSSS is robust to frequency-selective fading. Second, DSSSS compensates for the effect of a multipath propagation at the receiver by exploiting rake filters, which can collect the transmitted energy spread over multiple rays. Third, DS-SS also allows receivers to distinguish among signals simultaneously transmitted by multiple devices. Because of these reasons, SS increases the number of reuse channels and decreases the number of packet retransmissions. Therefore SS results in decreased energy consumption and increased network throughputs. Finally, DS-SS has an excellent security, noise/jamming immunity.

TEXT / REFERENCE BOOKS

1. Simon Haykins, "Communication Systems" John Wiley, 5th Edition, March 2009.
2. Taub. HDL Schilling, G Saha, "Principles of Communication" 3rd edition, 2007.
3. John G. Proakis, Masoud Salehi, "Digital Communication", McGraw Hill 5th edition November 6, 2007.
4. Bernard Sklar, "Digital Communication, Fundamentals and Application", Pearson Education Asia, 2nd Edition, Jan. 21, 2001.

QUESTION BANK

UNIT V

PART-A	
1. Give the beneficial attributes of spread spectrum systems.	
2. Write the properties of pseudo noise (PN) sequences.	
3. What advantage is there in making the message bit rate a sub-multiple of the PN sequence rate ?	
4. Given the data rate is 100Kbps and the chip rate is 1Mbps, calculate the processing gain.	
5. What is pulse jamming?	
6. Compare frequency hopping and direct sequence spread spectrum.	
7. Define processing gain?	
8. Define run length property of PN sequence.	
9. What do you mean by slow frequency hopping?	
10. Give examples of real life wireless systems that employ DSSS or FHSS.	
PART-B	
1. Discuss about the potential applications of spread spectrum techniques in wireless communications.	(16 marks)
2. a) Describe how linear feedback shift registers can be used to generate PN sequence for given degree and sequence length.	(8marks)
2.b) Discuss briefly about the importance of primitive polynomials in PN sequence generation.	(8 marks)
3.a) Design a PN sequence generator circuit using shift registers for the primitive polynomial	

x^4+x+1 of degree 4. (8 marks)

3.b) Plot the output PN sequence of your circuit w.r.t to a clock frequency of 1 KHz and initial seed value [1 0 0 0]. (8 marks)

3. Discuss the hardware system configuration of a Direct sequence spread spectrum transmitter and receiver with Binary Phase shift keying modulation scheme in detail. (16 marks)

4. Derive the expression for processing gain of Direct sequence spread spectrum system.

(16

marks)

5. Explain the transmitter and receiver architecture of a frequency hopping spread spectrum system. (16 marks)

6. Illustrate fast frequency hopping and slow frequency hopping spread spectrum with examples. (16 marks)

7. Explain how spread spectrum systems can reliably receive signals that are “buried in the noise”. (16 marks)

8. Assume that in order to obtain a decent BER, the required E_b/N_o is 12 dB. Furthermore, assume that the duration of the information bit T_b is 0.5 ms. Given that the required jamming margin is approximately 21.11 dB, what is the required processing gain? What is the resulting chip-duration? What is the required length of the shift register that produces the corresponding maximum length PN sequence? (16 marks)