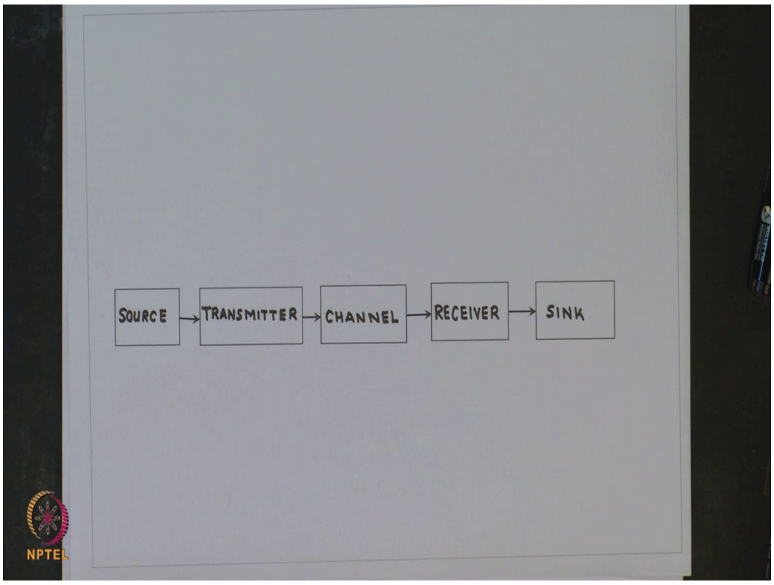
Principles of Digital Communications
Prof. Shabbir N. Merchant
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture – 01
Course Overview

Welcome to this course titled Principles of Digital Communications. The basic goal of this course is to provide some understanding of the basic principles of digital communication and study techniques to design and analyze digital communication systems. A communication system either an analog or digital is represented by the general block diagram as shown in the figure here.

(Refer Slide Time: 00:53)



We have a source, transmitter, channel, receiver, and the destination or the recipient or the sink

The pioneer telephone network and radio broadcasting systems employed analog communication methods such as amplitude modulation and frequency modulation, that is AM FM for transmission of analog voice. And analog TV broadcasting system used vestigial sideband AM for picture transmission. The quality of the message such as speech and images at the analog receiver depends on how well the waveform that carries the message over the physical channel can be reproduced. Some of the physical channels are twisted pair telephone wires, coaxial cables, fiber optic cables, space, water and so on.

In addition the fidelity of the received message depends on the signal to noise ratio at the receiver input. For good analog communications the signal to noise ratio at the receiver must be large and this requires high power transmitters; such as used in AM Radio and TV broadcasting. For FM Radio broadcasting a large frequency spectrum is used such as 200 kilohertz for radio broadcasting, this shows that analog communications do not use power and bandwidth efficiently.

Furthermore, the advent of the internet requires audio, video, imagery, and text messages to be integrated for transmission over a common channel and this in effect rules out analog communications such as AM and FM. In analog communications the message signal requires an infinite set of continuous time waveforms, for transmission over a physical channel. This is because the message it itself such as audio or video must first be converted into a voltage baseband waveform with a continuous range in amplitude, that has infinite or countless possible values.

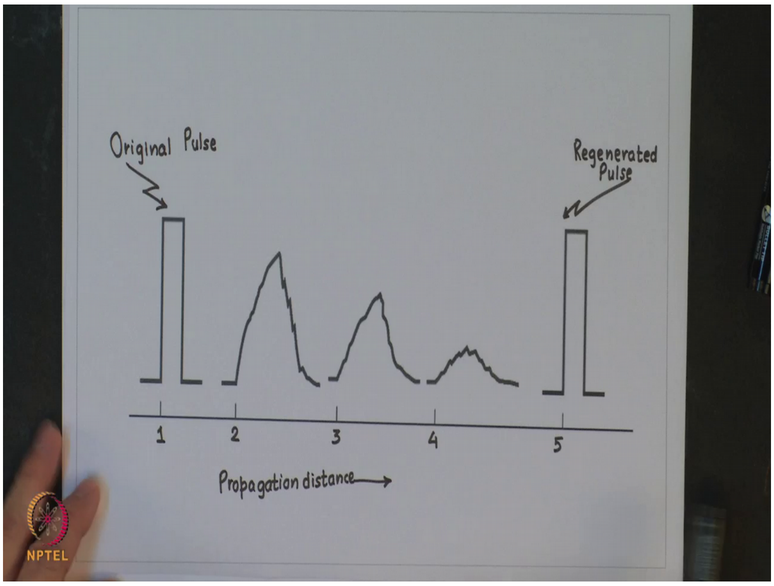
Now, when the baseband voltage waveform is used to modulate an RF carrier for transmission such as in AM and FM; the modulated RF signal transmitted over the physical channel also has infinite or countless possible values in both it is amplitude and frequency range. Now, the only way to recover the message signal is to faithfully reproduce, the baseband waveform from the modulated signal. Now, this can be done easily in the case of no noise or imperfections on the physical channel, but otherwise the fidelity of the message signal will be reduced.

Now, digital communication does not involve the faithful rep reproduction of the baseband waveform in the presence of noise. The reason is that digital communication operates instead with a finite set of continuous time modulation waveforms for transmission over a physical channel. What this implies is that the message signal must be represented by a finite set of voltage baseband waveforms, which represent a finite set of what is commonly referred to as symbols or letters, which consists of a fixed number of binary digits or bits.

So, in digital communication no attempt is made to reproduce the finite set of voltage baseband waveforms. Instead the receiver detects the energy content of each baseband waveform in the presence of noise and then makes a best estimate of which transmitted symbol was received. So, if the signal to noise ratio per symbol is reasonably large a symbol will most likely be detected correctly with high probability if not a symbol error may occur.

So, this is the essence of digital communication

(Refer Slide Time: 06:12)



If, you look at this figure this is the original pulse, which we are transmitting say representing a message symbol 1. Then on the transmission channel as the signal progresses towards the destination due to the imperfections, and noise on the physical channel the pulse gets distorted as shown here as it progresses towards the destination. At a destination we are interested whether the pulse was transmitted or it was not transmitted. So, it is not important for us to preserve the exact shape of the pulse.

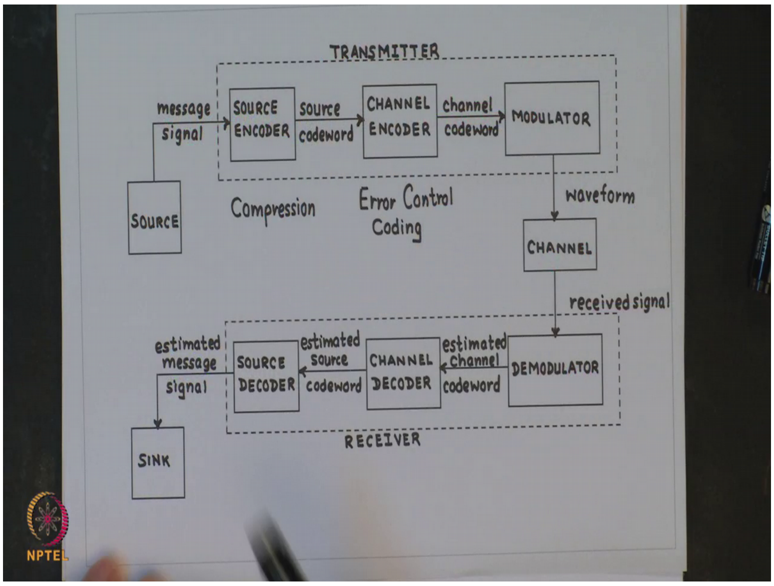
So, in this case looking at this pulse it is decided by the receiver that one the pulse of this shape was transmitted and at the receiver the decision is in the favor of this pulse or if it is a repeater which are used at regular intervals then, that repeater this pulse will be regenerated and retransmitted forward. So, this is the difference between the digital communication and analog communication.

So, for a given signal to noise ratio and analog communication receiver attempts to reproduce the voltage baseband waveform with certain subjective fidelity whereas, on the other hand for a given signal to noise ratio per symbol a digital communication receiver, produces symbols with quantitative error rate.

Now, it is important to know in advance the lower bound of the signal to noise ratio per symbol, for a specified error rate irrespective of the type and size of the set of modulation waveforms. In 1948 Claude Shannon established this lower bound and also provided the channel capacity for reliable transmission.

Now, for a digital communication their transmitter and receiver blocks of a general communication system is further subdivided as shown in the figure here.

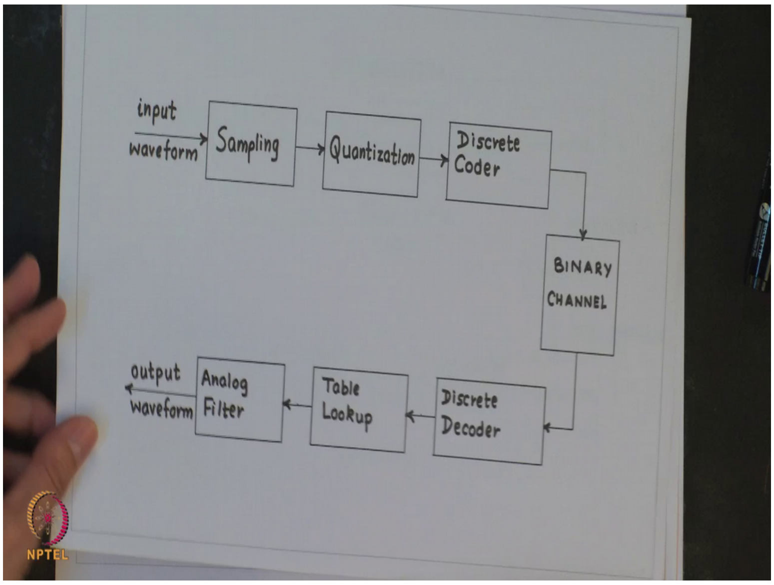
(Refer Slide Time: 09:11)



So, we have a source and the transmitter is subdivided into 3 blocks; one is the source encoder, the next is channel encoder followed by a modulator and all 3 together forms what is known as the transmitter? And the output of the modulator goes over the channel and at the receiver the receiver module is again subdivided into 3 blocks, we have demodulator, the equivalent of the modulator, we have a channel decoder, the inverse of the channel encoder, and the source decoder, the inverse for the source encoder, and finally, the recipient or the sink or the destination.

So, the source is typically first pass through a source encoder, which formats the source messages. The encoder details depend on the source and may be further subdivided as shown in the figure here.

(Refer Slide Time: 10:40)



So, the source encoder itself could be subdivided as sampler followed by a quantizer and followed by a discreet coder. So, as far as the source encoder is concerned all this could be clubbed in the model for the channel and that is what is shown here. And, then we have the inverse operation at the receiver we have again discreet decoder, in place of quantizer, we will have table lookup and in place of the sampler. We will have what is known as the reconstruction analog filter the output of which will go to the sink.

So, ideally the task of the source encoder should be to remove all redundancy from the source message and represented by symbol drawn from a finite alphabet and transmit this symbol every TS seconds. Typically, this alphabet is binary and the source encoder output is a bit stream or sequence.

Now, removal of the redundancy implies that the source rate typically measured in bit per second will be reduced, which in turn as we will learn later on means, that the required frequency spectrum bandwidth will be reduced. Now having removed some of the redundancy by means of the source encoder channel encoder adds redundancy back. So, this might appear to be countered interviewed, but this redundancy, which is added by channel encoder is done in a controlled fashion to take care of error detection and or error correction process at the receiver. So, the channel encoder maps the input symbol sequence into an output symbol sequence.

Now, to transmit this symbol sequence across a physical channel requires some energy and this is the function of the modulator block. So, the modulator block takes the symbol occurring in each TS second or symbol interval and maps it into a continuous time waveform, which is then sent across the channel. At the receiver one simply passes the received signal through the inverse of the operations at the transmitter.

Now, this is not very simple due to the influence of the channel. If the channel did not filter or distort the signal or did not add noise to it and if there was no interference from other users than, this process at the receiver would be very simple. However, some of these degradations are always present on physical channels. Therefore, one must attempt to overcome the degradations in the design process itself. So, for the design or analysis one needs engineering models of this channel.

Now, the channel model depends on the physical medium use for transmission of modulated signals. For example, in guided media such as twisted pair coaxial cable and optical fiber the background noise is the only kind of imperfection and this could be modeled as Gaussian PDF. However, they may exhibit a signal distortion where the transmitted signal is smeared out and causes what is known as inter symbol interference? That is signals in adjacent or nearby time slots interfere with each other a space channel example satellite communication. Typically only adds Gaussian noise to the received signal fading where the received signal strength varies with time is the predominant degradation in mobile communications, where the signal path from the transmitter to the receiver changes rapidly.

Let me quickly provide an overview of this course. Main purpose of communication is to transfer information from a source to destination via physical channel or medium. Therefore, it is appropriate that we begin our study of this course with the fundamentals of information theory. We will see that information theory will pay the way for many important concepts and topics to be discussed in this course. All of us have an intuitive notion of the meaning of information. However, performance analysis of communication systems can hardly be conceived without a quantitative measure of information and mathematical modeling of information sources.

So, in our course we will quantitatively define information and address an important issue in communication theory that is the representation of data generated by a discrete source of information. We will study Shannon’s first theorem that is source coding theorem which imposes the bound on the average number of bits per symbol necessary to represent or discrete memory less source. We will study that we should assign short code words to those messages, which have high probability and long code words to those messages, which have low probability.

We will also study Shannon’s second theorem that is the Channel coding theorem, which is both the most surprising and the single most important result of information theory. It gives the designers of digital communication systems, the freedom to choose the set of modulation waveforms that achieve either the best power of bandwidth efficiency or trade off combination of both.

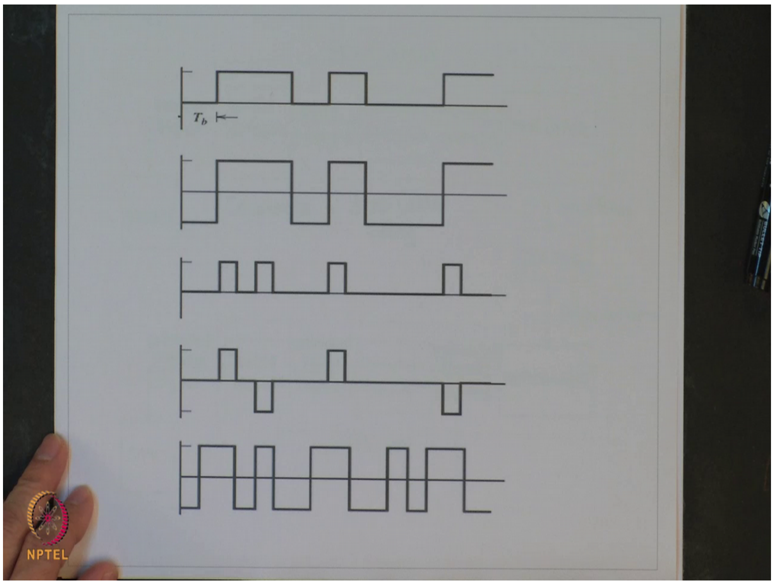
Now, what this theorem says is that as long as the transmission rate is below the channel capacity and the signal to noise ratio per symbol is above the Shannon limit, then reliable communication is possible with an arbitrarily small error rate this is very important. In the analysis of digital communication systems it is often convenient to represent a set of finite energy digital signals by a corresponding set of vectors. Each vector represents a digital signal, thus the time dependency can be removed and these signals define, what is known as signal space or signal constellation.

So, to obtain the signal vectors each digital signal must be represented by an orthogonal series in terms of what is known has basis functions? In this course we will learn how to achieve this orthogonalization process by means of the graham smith procedure, which is similar to the one used in vector algebra.

Next, we will consider the transmission of the digital information signals over communication channels that are characterized as additive white Gaussian noise that is AWGN channels. And study the signal space approach for design of optimum demodulation and detection of the signals and evaluation of the performance in terms of the probability of error. Now pulse code modulation is a discrete time discrete amplitude waveform coding process by means of which an analog signal is directly represented by a sequence of coded pulses. And this is made possible to the combined use of sampling and amplitude quantization. Amplitude quantization is defined as the process of transforming the sample amplitude of a message signal at the discrete time into a discrete amplitude taken from a finite set of possible amplitudes. We will study both uniform and optimum non uniform quantizers including compounders.

We will also study differential pulse code modulation that is DPCM and delta modulation, which are useful forms of digital pulse modulation. Both this modulation scheme exploits the idea of prediction to reduce the channel bandwidth requirements of PCM signals. The output of the source and channel encoder is a binary sequence or stream of bits, which represent the information contained, but this are abstract or intangible quantities and need to be converted into electrical waveforms for effective transmission across the physical channel, this in fact, is a form of baseband modulation.

(Refer Slide Time: 22:38)



So, the figure here shows different types of pulse signals of waveforms used to represent the binary signals. So, here this would mean 0 and this means 1 followed by 1 again 0 1 and then maybe 2 0s again followed by 1. So, this is one type of the waveform which we could have used for representing this binary stream, but there is another type of a waveform which is shown here and there is a third type of a waveform shown here, and there is a fourth type and there is a fifth type of the waveform shown here. So, this is known as baseband modulation.

Now, what type to use is governed by many factors of which the most important one is available transmission bandwidth of the communication channel? We will study these different factors to be considered in the choice of a particular line code.

Student: (Refer Time: 23:36).

These are known as line code and also study how to evaluate the power spectral density of this different line codes?

Now, when signals are transmitted over channels of infinite bandwidth or at least large in our bandwidth, then any signal distortion is negligible and can be ignored. During some situations this assumption is reasonable satellite communication is a common example band limitation is also common. The classical example is telephone channel, where the twisted pair wires used as a transmission medium have a bandwidth on the order of kilohertz, but even a medium such as optical fiber exhibits a phenomena called dispersion, which results in effect very analogous to band limitation. It is important to note, that band limitation depends not only on the channel medium, but also on the source especially the source rate, that is the number of symbols per second.

One common measure of the bandwidth needed by a source is the source rate itself. So, as source rate keeps increasing to accommodate more data eventually any channel starts to look band limited. Band limitation can also be imposed on a communication channel by regulatory requirements; a user is usually allotted only some specific amount of bandwidth in which to transmit the information?

So, the general effect of ban limitation on a transmitted signal of finite duration is to disperse it or spread it out. Therefore, the signal transmitted in a particular time slot or symbol interval will interfere with signals in other time slots resulting in what is called inter symbol interference or ISI. We will study nyquist criteria for 0 ISI and provide the design an optimum demodulator that can achieve both maximum output signal to noise ratio and 0 ISI. We will also study how to mitigate the effects of ISI by deliberately allowing some ISI, but in a controlled manner resulting in what is known as partial response signaling?

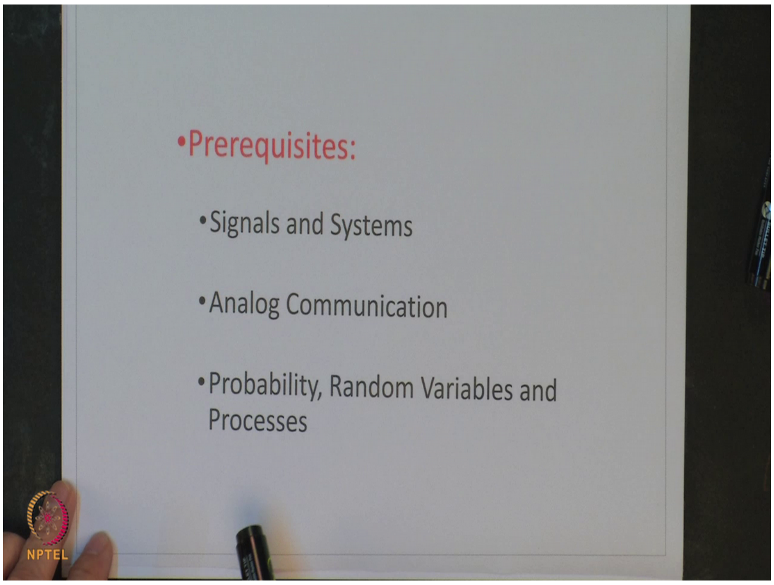
Now, it is desirable in many digital communication systems for the same reasons as in analog communication systems, for the transmitted signal to lie in a frequency band toward the high end of the spectrum. Therefore, we will study the digital modulation techniques for transmitting information over the physical channels. We will study binary modulation techniques like BSK also called on and off key a technique used in fiber optic communication, binary phase shift keying that is BPSK a popular modulation, that is employed widely in practice in wireless 8 0 2.11 a and 8 0 2.11 g standards, binary FSK and minimum shift keying used in cellular communication in the form of Gaussian MSK.

Now, many practical applications require either the higher spectral efficiency, which is defined as amount of bits transmitted per second per hertz of the bandwidth or higher power efficiency that binary modulation techniques can provide. So, we will study ameri modulation which can accommodate both. We will study ameri modulation techniques like quadrature PSK, that is QPSK most popular modulation for satellite communication and also used in the wireless 8 0 2.11 a g and 8 0 2.16 standards, ameri amplitude shift keying, which is employed in HDTV standard in the united states, Am Qam Ameri quadrature amplitude modulation, which are again used in the wireless 8 0 2.11 a g and 8 0 2.16 standards, ameri FSK, QPSK and QAM are used delete.

Now, the structure of each modulation technique will be studied by the signal space approach to design the modulator and optimum coherent demodulator and evaluation of the power spectral density. As, with baseband the focus will be on the power needed to achieve a certain performance measured as the bit error probability and on the bandwidth requirements of the specific modulation.

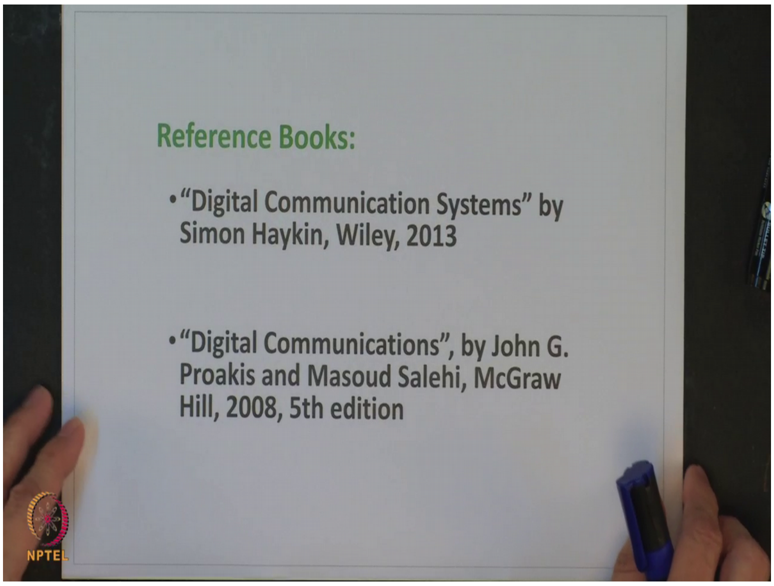
In situations where synchronization between the transmitter and the receiver is difficult, non quadrant form of demodulation for phase shift keying in the form of differential phase shift keying that is DPSK, and FSK will be studied again DPSK is used in the wireless 8 0 2.11 standard. Finally, we will learn channel coding, which introduces redundancy. However, in a controlled manner for error detection and error correction purposes, we will restrict our study to linear block coding.

(Refer Slide Time: 29:22)



Now, the prerequisite for this course is signals and systems, analog communication and probability random variables and random processes. It is presumed that the student would have had an exposure to some of the basic concepts from these topics, like power spectral density to be very specific.

(Refer Slide Time: 29:55)



And there are quite good books, but for this course, the 2 reference books would be digital communication system by Simon Haykin and digital communication by Proakis and Masoud Salehi.

So, with this introduction we will begin the formal study of our course next time.

Thank you.