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

Lecture - 60
Channel Coding - I

We begin our study of channel coding. Channel coding improves digital communication performance by transforming their transmitted signals, so, that they are more robust to withstand the effect of various channel impairments such as noise, interference and feeding. Channel coding helps in achieving desirable system trades off example error performance versus bandwidth power versus bandwidth.

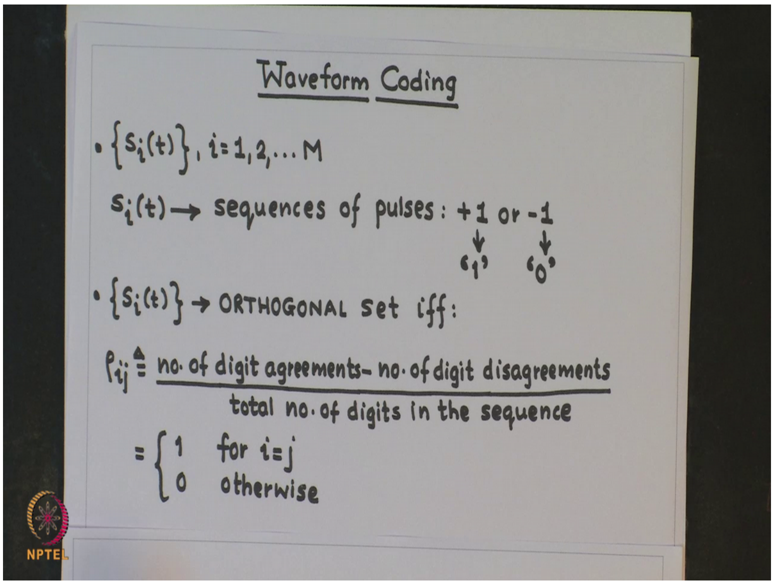
Now, channel coding can be partitioned into 2 study areas, waveform coding and error control coding. In waveform coding a set of waveforms are transformed into another so called improved set of waveforms, which have better error performance properties. In error control coding message data sequences are transformed into better structured sequences, which have structured redundancy in the form of redundant bits added to the original data sequences. The redundant bits are used for the detection and correction of errors.

Now, the main goal of these encoding procedures, both waveform coding and error control coding is to provide the coded signal with better distant properties, then those of their un-coded counter parts; this will improve the error performance property of the coded signal. Now the most popular technique of waveform coding is referred to as orthogonal coding. The motivation is to make each of the waveforms in the signal set dissimilar; that is to make the cross correlation coefficient among all pairs of the signals as small as possible.

Remember that the cross correlation between two message signals is a measure of the distance between the message signal vectors. The smaller the cross correlation the more distant are this message vectors from each other and consequently possess better error performance characteristics. Now obviously the smallest possible value of the cross correlation coefficient occurs when the signals are antipodal.

However, this can be achieved only when the number of symbols or signals in the set is 2, but in general it is possible to make all the cross correlation coefficients between the pair of the signals in the signal set equal to 0, and such a set is then said to be orthogonal. Now antipodal signals are optimum in the sense that each signal is most distance from the other signal in the set. So, compared with the antipodal signals, the distance properties of orthogonal signal sets are quite acceptable for a given level of signal energy. We will take one example of waveform coding.

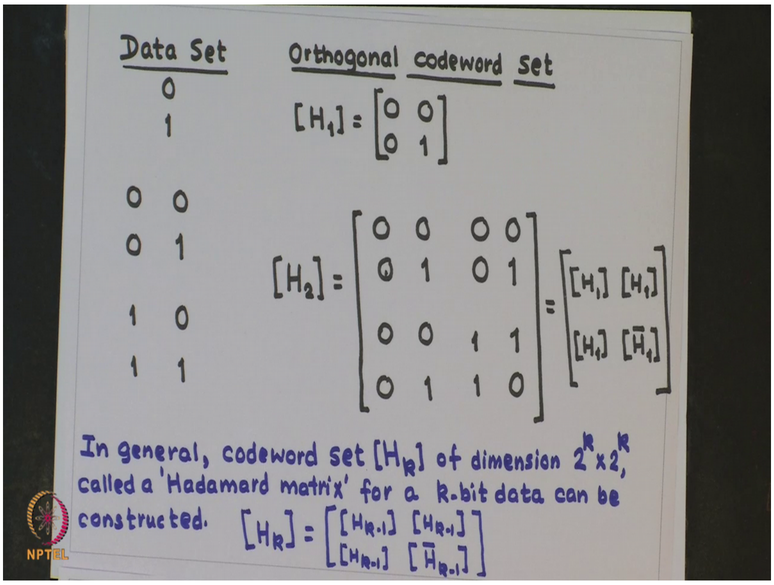
(Refer Slide Time: 04:13)



So, each S i t consists of sequence of pulses, where each pulse is designated with a level plus 1 or minus 1, which in turn represents the binary digit 1 or 0 respectively. Now this S i t will be an orthogonal set if and only if this correlation coefficient defined by this relationship, where in the numerator we have the difference of number of digit agreements and number of digit disagreements, and the denominator we have total number of digits in the sequence.

So, if this will be equal to 1 for i equal to j and it will be 0 otherwise. If this condition is satisfied for all i and j in this signal set then this is known as orthogonal set. Now as mentioned that orthogonal coding is a popular form of waveform coding. So, using the orthogonal coding 1 bit data set can be transformed using orthogonal code words of 2 digits each described by the rows of this matrix H1.

(Refer Slide Time: 05:48)



And similarly we can have a 2 bits data set transform using orthogonal code words of 4 digits each described by the rows of matrix H 2. Now H 2 is obtained from the matrix H 1 by this relationship, the bar on H 1 denotes that all the elements in the matrix H 1 are complemented. So, H 1 corresponds to this matrix, so H 1 bar corresponds to this matrix.

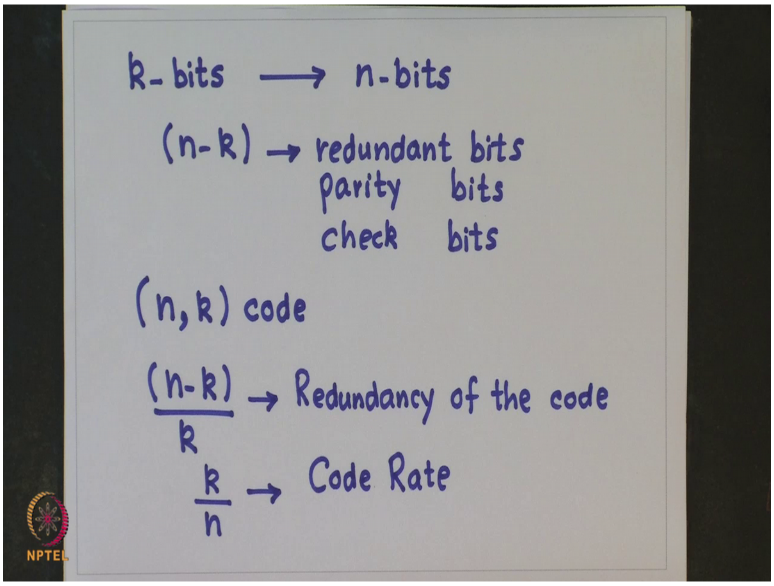
So, in general code word set given by the matrix H k which will be of dimension 2 raised to k by 2 raise to k and also known as Hadamard matrix can be constructed for coding k bit data. And this Hadamard matrix can be recursively obtained from the Hadamard matrix of order H k minus 1 as given by this relationship.

In our study we will restrict to error control coding. Now there are 2 basic ways in which structured redundancy can be used for controlling errors; the first is error detection and retransmission. This uses parity bits as redundant bits added to the data to detect the occurrence of an error. Now in this case the receiver does not attempt to correct the error, but it simply requests the transmitter to retransmit the data. And for it is implementation 2 way link is required for such handshake between the transmitter and the receiver.

The second type of error control coding is what is known as forward error correction FEC and this requires only one way link. Since, in this case the parity bits are designed for both the detection and correction of errors. Of course, not all error patterns can be corrected. Now error correcting codes are classified according to the error correcting capabilities. There are three sub categories these are block codes convolution codes and turbo codes, now for our study we will restrict to block codes.

So, in the case of block codes the source data are segmented into block of k data digits also called information digits or message digits, in our study we will restrict to the binary case, so our digits are going to be chosen from the alphabet 0 and 1. So, in this case the block will consist of k data bits, so each block can represent any of the 2 raise to k distinct messages. Now the job of the encoder is to transform each k bit data block into a larger block of n bits called code bits or channel symbols.

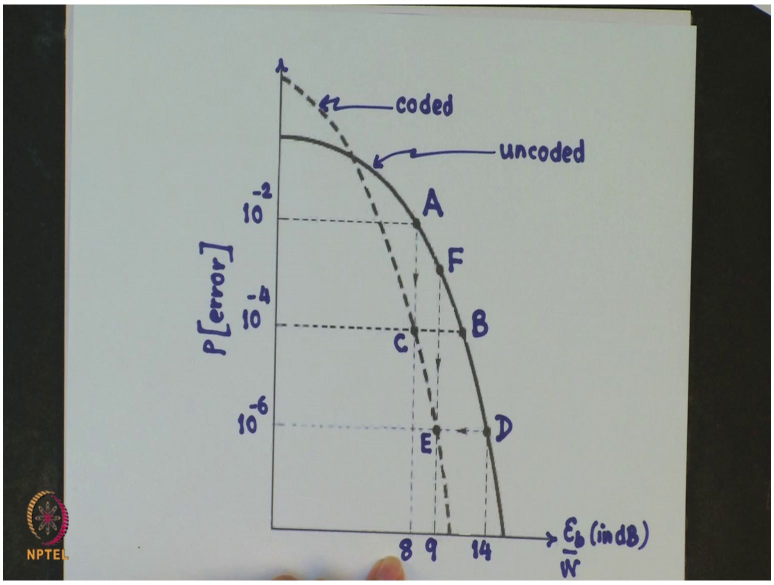
(Refer Slide Time: 10:07)



So, we have message sequences of k bits and this gets transformed to n bits sequence, where n is larger than k. Now the n minus k bits which the encoder adds to each data block are called redundant bits are also known as parity bits or check bits and the code is referred to as n comma k code.

The ratio of the redundant bits to data bits which is denoted by n minus k by k, within a block is called the redundancy of the code. The ratio of data bits which is k by n this is called the code rate. So, the code rate can be thought of as the portion of a code bit that constitutes information. For example, in a rate half code each code bit carries half bit of information. Now block coding increases the transmission bandwidth requirement for example, rate half code would require twice the bandwidth assuming 0 delay constraint. Now, before we go ahead let us look at the advantages of the error control coding.

(Refer Slide Time: 12:45)



Now the figure out here depicts 2 curves, one curve is for the un-coded system and the other dotted curve is for the coded system. And here we have plotted the probability of error versus signal to noise ratio per bit that is Eb by italic N in dB.

Now let us see the advantages of the coded system. Let us say we have some simple inexpensive communication system which is operating at point A. So, the probability of error is 10 raised to minus 2 and SNR per bit is 8 dB. But after few trials we find that this equipment is not performing well as far as the quality of the service is concerned, and it is desired that we reduced the probability of error to say 10 raise to minus 4. So, if we still use the un-coded system we would be required to go from the point A to point B. Now this would demand higher SNR per bit.

So, if we cannot afford to provide that kind of high SNR per bit, what we could do instead of coming on this curve to point B we could move from this point to this point on the coded curve, where we can get 10 raise to minus 4 probability of error, but with the same SNR per bit. But in doing so, when we land up at point C the requirement of bandwidth will increase assuming that we are working with a real time system and no delay is acceptable. So, we get now error performance improvement, but at the cost of bandwidth when you move from point A to point C.

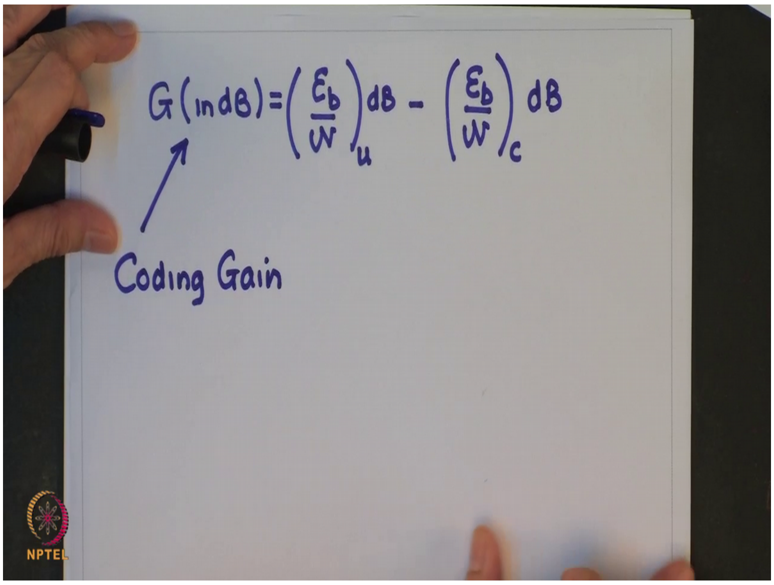
Now, let us say that let us take another advantage of this coding. Let us say that we are at point D, where the probability of error is 10 raised to minus 6 and SNR per bit is 14 dB.

But we find that due to this high SNR per bits the equipment gets heated up and there are frequent breakdowns.

So, in that case we will have to bring down SNR per bit, but we also warned that error performance should not suffer. So, in that case what we could do? We could move from point D to point E.

Now when we move from point D to point E which is now the coded system, the error performance remains the same 10 raise to minus 6, but now we get SNR per bit requirement to be just 9 dB. So, in this case we get the coding gain, so coding gain G is defined in dB as follows.

(Refer Slide Time: 16:28)



This is equal to SNR per bit required for the un-coded system and difference of this width, the SNR per bit required for the coded system, so this is known as coding gain. So, we get a coding gain of 5 dB for the same error performance, but lower SNR per bit. So, by moving from point D to point E we have lowered the SNR per bit requirement, but again this is done at the cost of bandwidth because, the bandwidth requirement will again be higher assuming real time system.

Now, for both these cases, it should be remember that it is possible to obtain improved bit error performance or reduce power by paying the price of delay instead of bandwidth.

So, for non real systems we can avoid the increase of bandwidth, but then we will have to pay in terms of delay.

Now let us take another advantage of coding. Say let us say we are operating at point D again and the operation is all fine the SNR per bit can be withstand by the equipment and the performance is 10 raise to minus 6 as for the error is concerned. But now at this point we required to increase the data rate. So, if we increase the data rate what will happen? More number of bits have to be squeezed in the same duration, so this will imply that your SNR per bit will fall down. So, it is possible that point D would move to the point F. Now once you move to the point F your error performance will deteriorate and since we want to keep the same error performance then what will be required that will be required to again come back to point E.

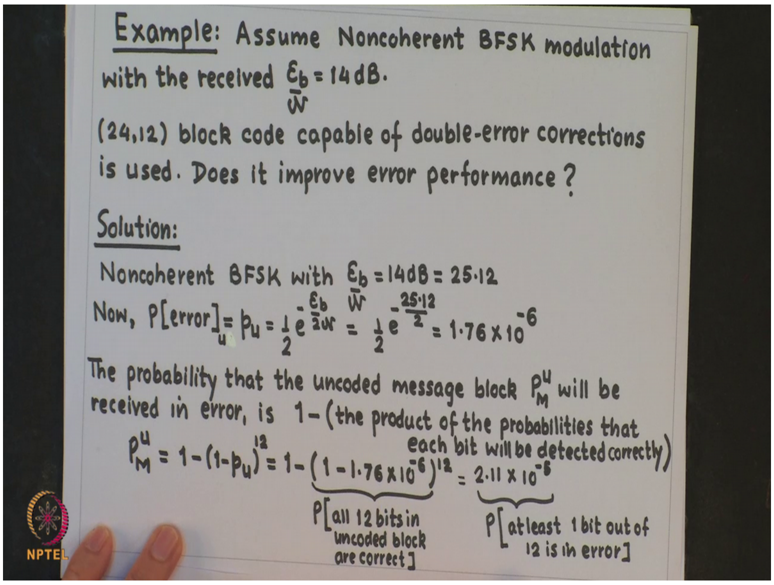
So, again we see here that we can get the higher data rate, but we will have to pay in terms of the bandwidth and we will get one more advantage is that SNR per bit will also reduce and now we look at a fourth advantage of the error control coding, and that is the increase in capacity at the expense of bandwidth.

Now code division multiple access that is CDMA is a popular spread spectrum multiple access technique, which is used in cellular communication. In this CDMA, multiple users share the same spectrum and each user acts as an interferer to the other users in the same or nearby cell. So, in such a scenario the capacity which is maximum number of users per cell is inversely proportional to SNR per bit.

It is desirable that each user operates at as small as possible low power level. Therefore, it is recommended to use a coded system to achieve certain error performance because, this will be achieved at a lower SNR per bit. So, consequently the capacity will increase, but here again the cost is bandwidth, but it is important to note that in this case transmission bandwidth expansion due to error correcting code is small, compared with the most significant spread spectrum bandwidth expansion. And consequently in this case there is in significant impact on the transmission bandwidth expansion due to error correcting code.

Now let us take one example to understand how coding helps and also to see if there is a possibility when coding could deteriorate the performance.

(Refer Slide Time: 21:54)



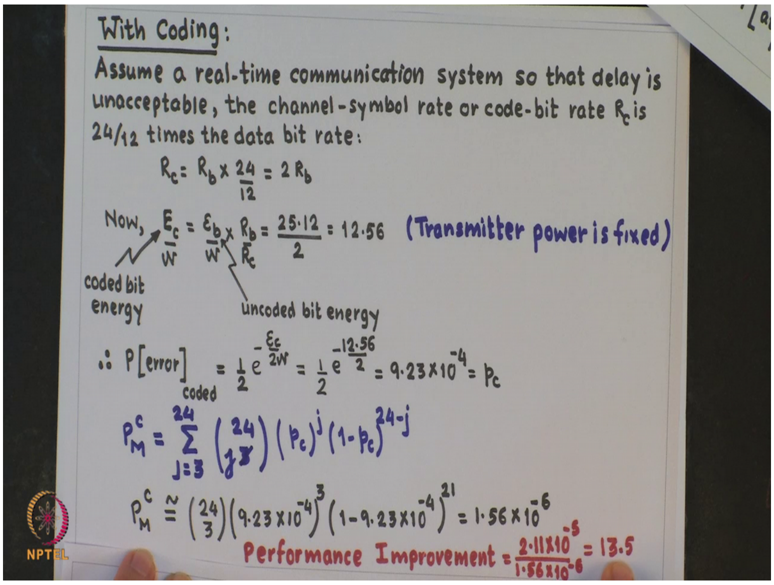
So, let us assume that we have a non coherent BFSK modulation with the received SNR per bit to be 14 dB. And we are also permitted to use 24 comma 12 block code which is capable of double error corrections and we will see whether this will improve the error performance.

So the solution is as follows, we know that for the non coherent BPSK with SNR equal to 14 dB, this corresponds to 25.12. The probability of error that would it is a bit error because, we have only 2 symbols now un-coded and I call it as pu is equal to half e raise to minus 1 by 2 SNR per bit; and if you plug in this value we get this as the probability of symbol error or bit error.

Now the probability that the un-coded message block of 12 bits will be received in error is as follows. So, we will denote that error as by this notation and that is equal to 1 minus the product of the probabilities that each bit will be detected correctly in the block of 12 bits, and that would be solved using this expression. So, this is the error this is the probability of correct detection of all the bits. So, 1 minus that and if we just plug in this values we get 2.11 into 10 raise to minus 5.

So, this is the probability that all 12 bits in the un-coded block are correct and this is the probability that it is 1 bit out of 12 is in error. Now we will see what happens when we use the coding.

(Refer Slide Time: 24:18)



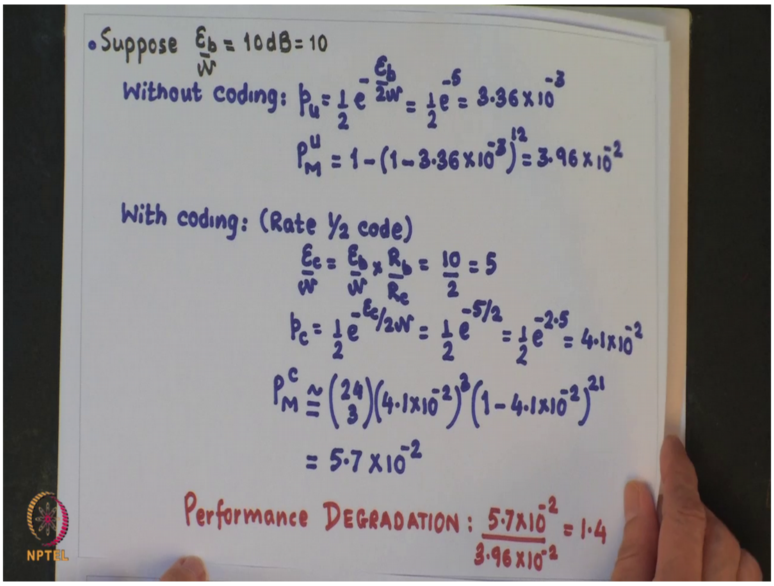
So now assume a real time communication system, so the delay is unacceptable so, the channel symbol rate of code bit rate R c is now 24 by 12 times the data bit rate. So, Rc is equal to Rb multiplied by 24 by 12 it will be twice the rate because it is a real time system no delay is expected. Now we know if the transfer of power is fixed, then the relationship between the coded SNR per bit and un-coded SNR per bit is equal to this quantity we have seen this earlier.

So, this turns out to be 12.56 so SNR per bit has reduced in the coded case. And now we use the same we will compute the probability of symbol error or bit error for the coded case and we use this formula and we get this quantity. And now from here we see that the probability of the bit error or the symbol error in the coded case is; obviously higher than the 1 which we get in the un-coded this is expected, because the energy has got distorted in more number of bits we will see basically what happens the overall probability of error in the message block.

Now the probability of error which we will get after using the coding would be given by this formula;, now this formula is written based on the assumption that 24 comma 12 block can correct double errors. So, only when there are more than 3 or more errors you will get error in the block. So, this tells you how many ways I can find the combination of errors. So, this should be J this is not 3 this should be J. So, J out of 25 combinations this gives me the probability that exactly J bits have gone into error and this is the probability that 24 minus J bits have been received correctly.

So, this is the probability of error which you will get. Now this can be approximated by just taking the first term we get this and if you just plug in this values you will get this value. And now if you look at the performance improvement, if you take the earlier probability of error which had and this probability of error we get the performance improvement of 13.5. Suppose now we decrease the SNR per bit at the receiver end correct and let us say that that turns out to be this value is 10 dB.

(Refer Slide Time: 27:17)



So, this is equal to 10 and now let us recalculate everything. So, without the coding we can find out the values, we will find out the probability of error in the message correct in the un-coded case this turns out to be this and with the coding the procedure remains exactly the same except that this SNR per bit has reduced to 10 and now if we do the computation out here correct follow this computation, and finally again we make an approximation that there are neglect the more than 3 errors so, we get this and the final value we get is this. And now we see that this quantity turns out to be larger than without coding. So, if we see that now there is a performance degradation by the factor of 1 by 4.

So, if the received SNR per bit is very low correct, then we see that the coding does not help correct. And this is fact is also reflected in that figure and we saw here correct. So, we see that there is a threshold of SNR per bit, below which we see that the performance of the coded system is degraded compared to the uncoated system. So, in the figure this crossover is there the reason for existence of such a threshold is that every code has some fixed error correcting capability. So, if there are more errors then the code is capable of correcting the system will perform poorly.

So, for example, if SNR per bit is continuously reduced, then at the output of the demodulators there are more and more errors. So, consequently continual decrease of SNR per bit will eventually cause some threshold to be reached, when there will be many errors at the output of the decoder, when such a threshold is reached, we could interpret that such a degraded performance is being caused by the redundant bits consuming energy, but giving back nothing beneficial in return.

Now, with this introduction of channel coding and it is importance we will study a class of block codes known as linear block codes and this we will do next time.

Thank you.