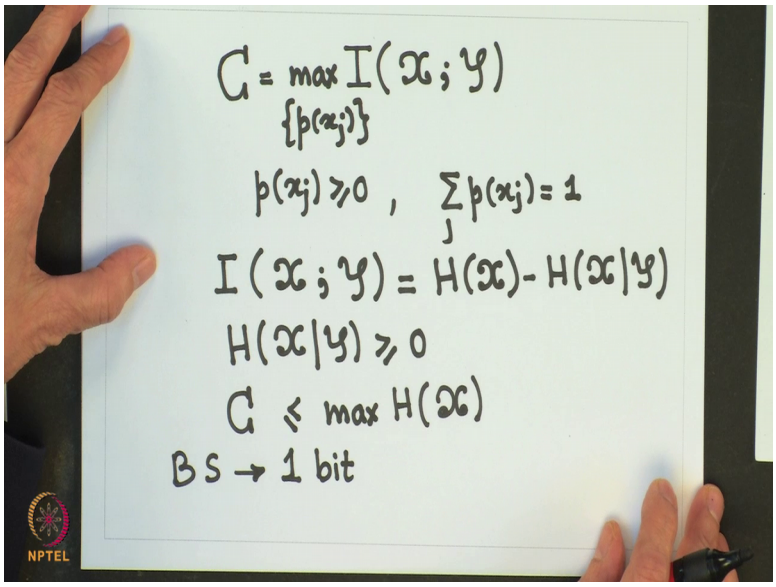
Principles of Digital Communications
Prof. Shabbir N. Merchant
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture - 08
Channel Capacity - II

We defined mutual information and then we defined what is known as a capacity of a communication channel. The definition is as follows.

(Refer Slide Time: 00:39)

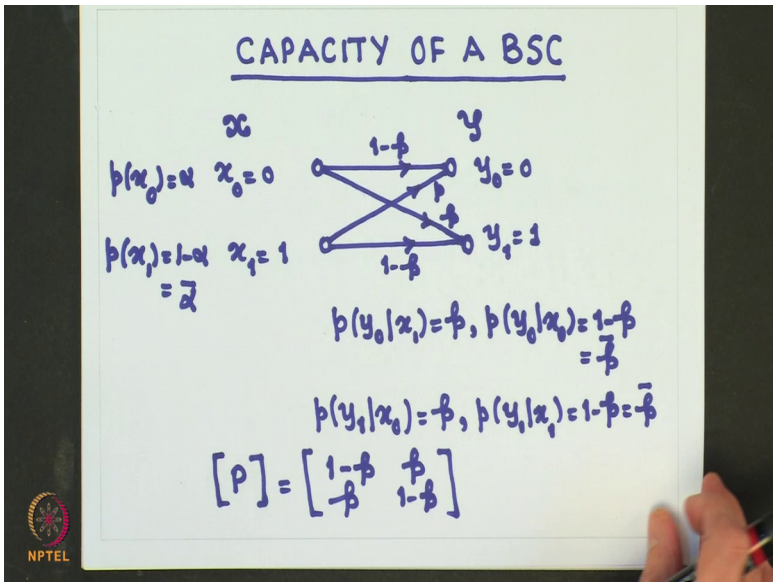


This maximum over the input symbols distribution and with the inequality constraint p x j greater than or equal to 0 and equality constraint summation of p x j is equal to 1.

Now, we also know that mutual information is equal to entropy of x minus conditional entropy of x given y. Now, note that conditional entropy x given y is always greater than equal to 0. So, what this implies that channel capacity will be always less than or equal to maximum of H x. So, if you consider a binary source then we know that maximum value of entropy, H x is equal to 1 bit which implies that the channel capacity of a binary symmetric channel cannot be more than 1 bit.

Let us formally calculate the capacity of our binary symmetric channel.

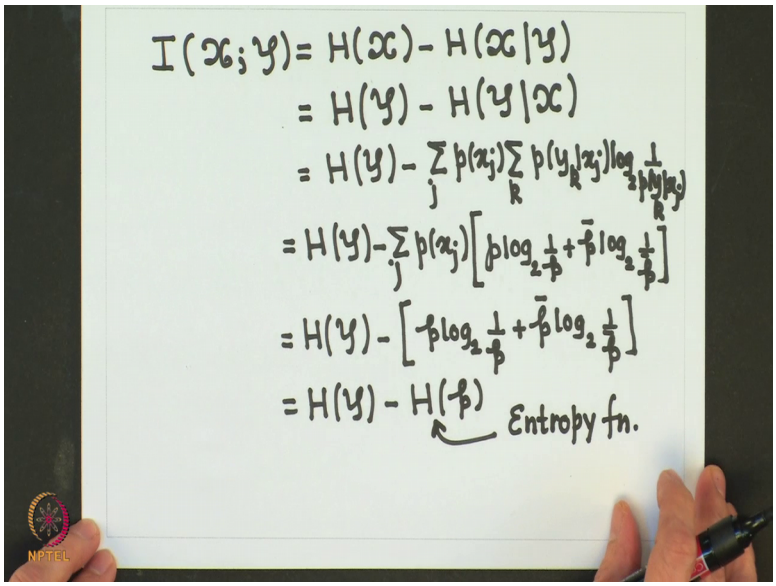
(Refer Slide Time: 02:36)



Here I show the transition probability diagram for a binary symmetric channel we have two inputs x 0, x 1. We have two outputs y 0, y 1 and we have the probability of x 0 given alpha probability of x 1 will be equal to 1 minus alpha, I call it as alpha bar. The transition probabilities are also specified here probability of y 0 given x 1 is equal to p and probability of y one given x 0 is also p. Therefore, it is a symmetric channel this is basically conditional probability of errors also known as crossover probabilities and probability of y 0 given x 0 is equal to 1 minus p which I call it as p bar.

So, the channel matrix or statistic matrix is given as shown here. So, for given for this channel let us calculate the capacity.

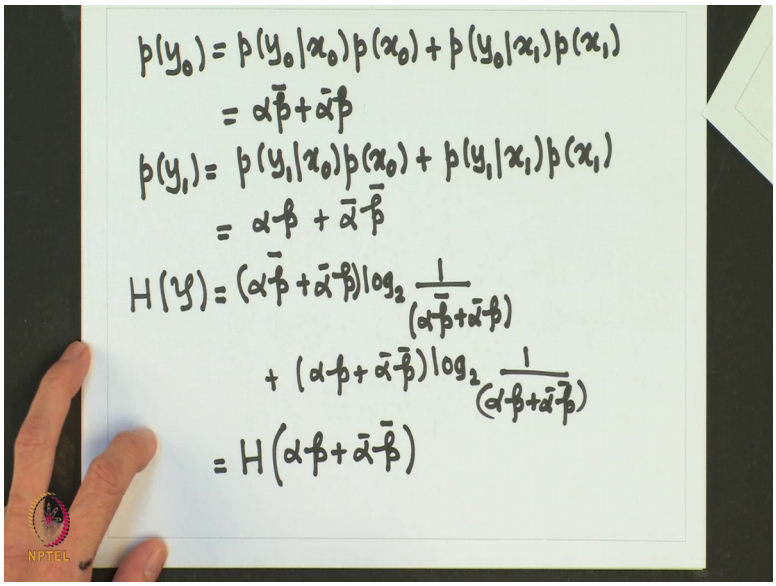
(Refer Slide Time: 03:54)



So, to do that basically let us first calculate the mutual information which is equal to H x minus H of x given y this is same as entropy of the output alphabet minus conditional entropy of y given x. It is to be remember that many a times this relationship helps us to calculate the mutual information in a much faster way, ok.

So, this is equal to this I can write it as this summation over k. Now, since their channel is symmetric we can easily write this as this can be written as. So, this I can rewrite it as and we know that this is nothing, but the entropy function. So, I will write it as H p, this is a entropy function. Now, I need to calculate the entropy y.

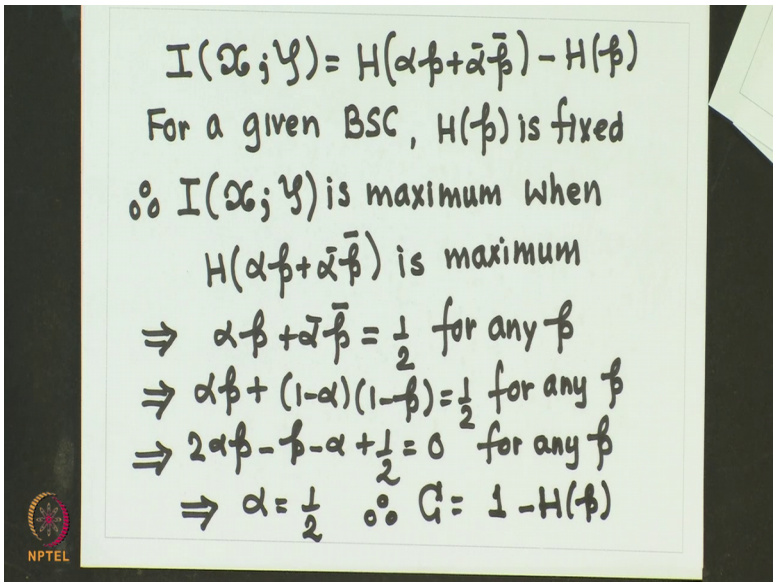
(Refer Slide Time: 07:00)



To do that I need to calculate the probabilities of the output symbols this is equal to and we get the probability of y 1 is equal to probability of y 1 given x 0 multiplied by probability x 0 plus probability of y 1 given x 1 probability x 1 and this is equal to alpha p, plus alpha dash p dash. Note p y 0 and p y 1; that means probability of y 0 and probability of y 1, summation of that is equal to 1.

So, I can write the entropy of the output as that is one term corresponding this other term corresponding this would be equal to ok. So, this basically I can identify as entropy function, ok.

(Refer Slide Time: 09:34)

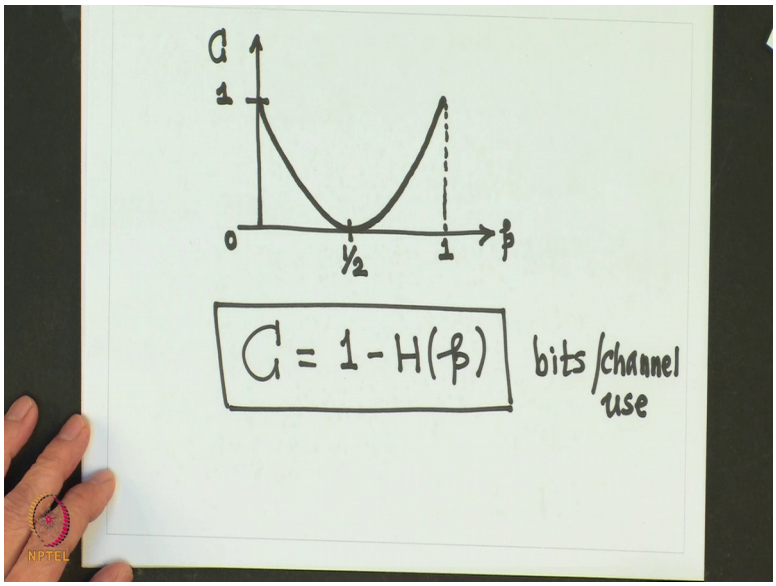


So, I have my mutual information equal to entropy function given by this minus now, remember for a given binary symmetric channel H p is fixed. Therefore, the mutual information is maximum when H alpha p plus this quantity is maximum.

Now, we know from the property of entropy function that this implies that alpha p plus alpha bar p bar should be equal to half for any p. So, this implies that alpha p plus 1 minus alpha 1 minus p is equal to half for any p. This implies that 2 alpha p minus p minus alpha plus half is equal to 0 for any p and this implies that alpha should be equal to half and therefore, my c turns out to be equal to 1 minus H p.

So, this is the capacity which I get for a binary symmetric channel.

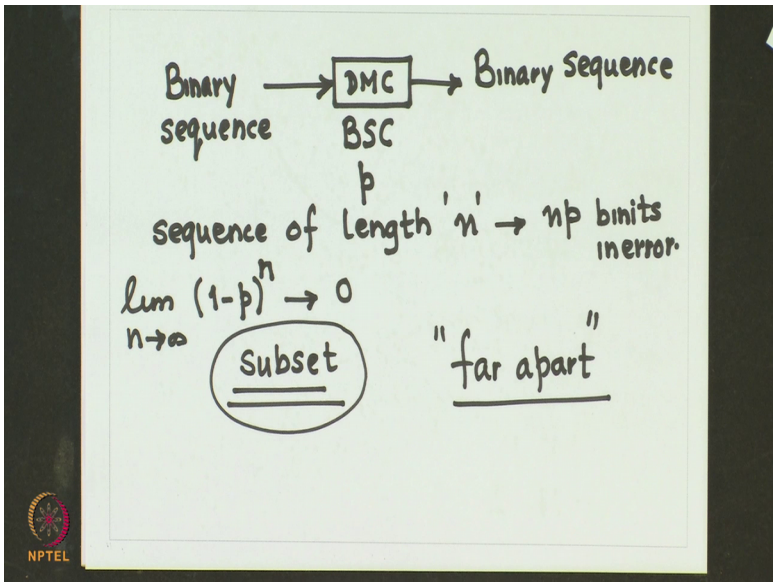
(Refer Slide Time: 12:05)



And, if I plot it I will get it as shown here. So, it will be a convex function this will be 1, this will be 0. So, this will be 1 and this will be equal to half. So, when the crossover probability is equal to 0, correct 1 we will get capacity to be equal to 1 bit per channel use and when the probability is equal to half I get the channel capacity to be equal to 0, ok, fine.

So, for a binary symmetric channel our capacity turns out to be this expression bits per channel use. Let us have a physical insight into this channel capacity formula and let us do it for a binary symmetric channel, ok.

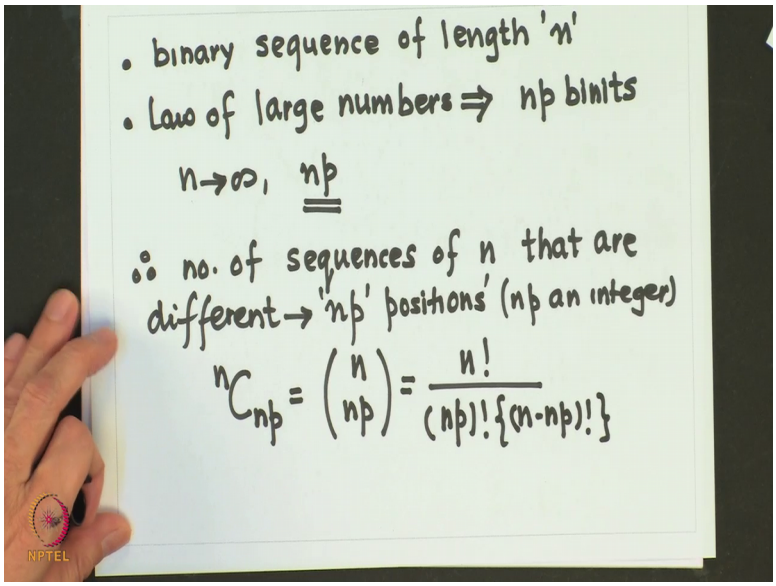
(Refer Slide Time: 13:41)



So, given a discrete memoryless source with the input as binary source; that means, I have a binary sequence here and let us consider this discrete memory channel to be a binary symmetric channel. So, the output will be also a binary stream and let us assume that the cross over probability is p. So, what this means basically that if I have a sequence of length n and n is a large then the probability of errors for each binary digit is p correct. So, what it implies that np binary bits are likely to go in error. So, the probability of correct reception of a binary digit is going to be 1 minus p. So, if you want the complete sequence of length n to arrive at the output of our discrete memoryless channels correctly the probability of such a thing to happen would be given by 1 minus p raised to n, correct.

So, it means that the probability of correct reception of the sequence of length n for n large is going to tend towards 0. So, one approach to improve the performance is not to use all the binary sequences of length n as possible inputs, but to use only a subset of the sequence of length n. Now, the subset has to be selected in a manner such that the sequences in this subset are in some sense far apart. So, if it satisfies this condition then it should be possible for us to recognize and correctly detect at the receiver the transmitted sequence even in the presence of channel errors.

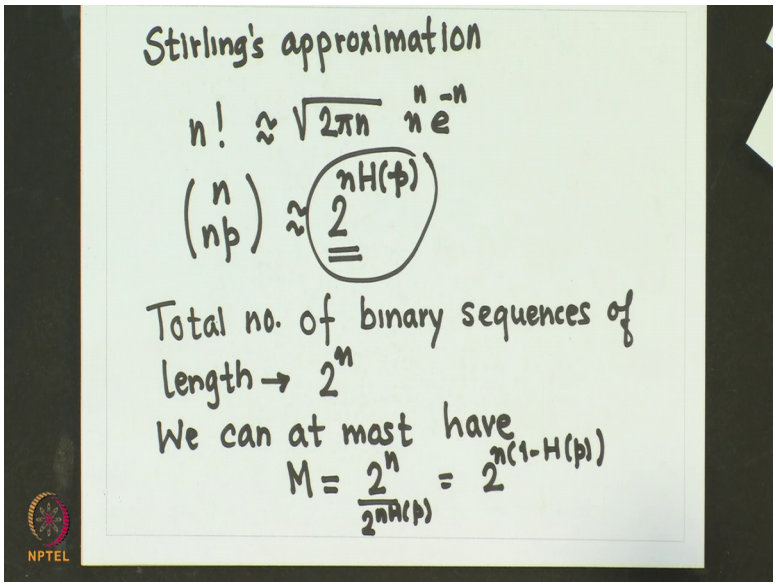
(Refer Slide Time: 17:00)



So, let us consider a binary sequence of length n which is being transmitted on the channel and this channel has a crossover probability of p. So, what it means that by law of large numbers for n large approximately np binits will be received if error in this sequence of length n. So, it means that if n tends to infinity with high probability approaching one the number of binits which differ from the transmitted binary sequence of length n will be equal to n p, ok.

So, what this implies that number of sequences of length n that are different from the transmitted sequence at np positions I am assuming np as an integer will be given by n C np is equal to n factorial by np factorial n minus np factorial, correct, ok.

(Refer Slide Time: 19:30)

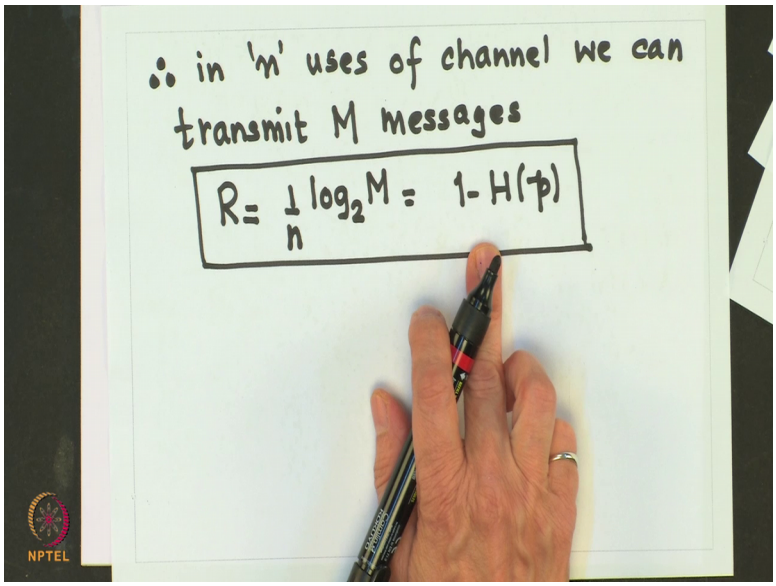


There is a approximation called Stirling's approximation which says that for large n, n factorial can be approximated by root of 2 pi n, n raised to n e raised to minus n and based on this approximation it can be shown that this is approximately equal to. So, this combinations number of combinations can be reduced approximately to this formula, correct. We will not go into the proof of this, ok.

So, this means that when any sequence of length n is transmitted it is highly probable that one of this number of sequences will be received which will be different from the transmitted sequence in np positions, correct. Now, if we insist on using all possible input sequences for this channel then errors are inevitable because there will be consider overlap of the received sequence, but if we decide to use only a subset of the possible sequences and select this subset in a such a manner that the set of highly probable received sequences for each element in this subset is non overlapping then reliable communication is possible, ok.

So, if we can do this then let us see what happens, total number of binary sequences of length n at the channel output is given by 2 raised to n. Now, if we if we follow what we have said earlier then it means that we can at most have M is equal to 2 raised to n 2 raised to 2 n H p which is nothing, but 2 raised to n 1 minus H p sequences of length n transmitted without the corresponding highly probable receive sequences overlapping.

(Refer Slide Time: 23:00)



Therefore, in n uses of channel we can transmit M messages, that means, at the rate that is the information transmitted per use of the channel is given by R is equal to 1 by n log of M, correct and this is equal to 1 minus H p and this is the capacity which we derive for a binary symmetric channel.

So, this discussion basically provides some kind of a physical insight into the concept of capacity and we did it for binary symmetric channel. We will continue this discussion on channel capacity in the next class.

Thank you.