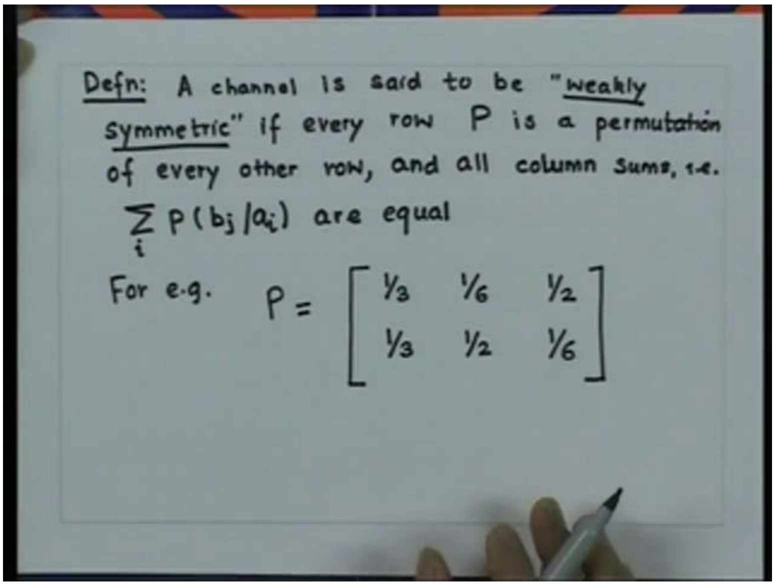
Information Theory and Coding
Prof. S. N. Merchant
Electrical Engineering
Indian institute of Technology, Bombay

Lecture - 25
Calculation of Channel Capacity for Different Information Channels

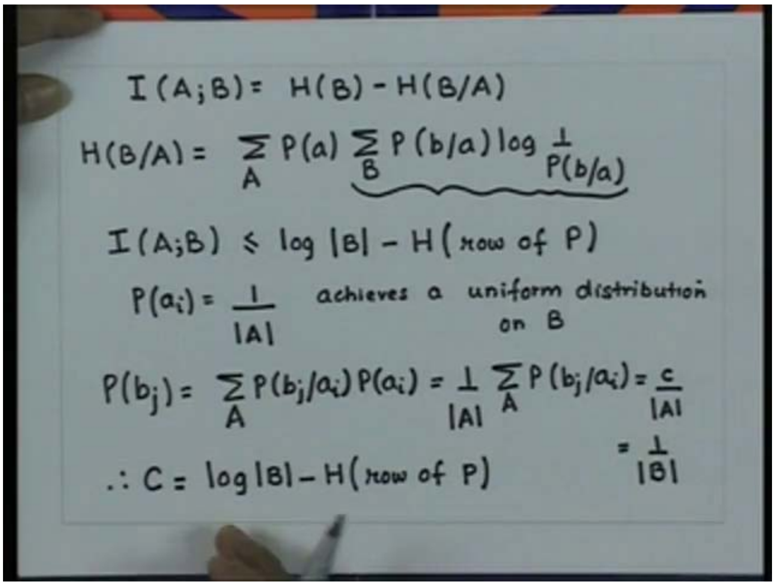
In the previous class, we studied about channel capacity and looked at the procedure to calculate the same for a binary symmetric channel and an array symmetric channel. We also looked at a definition of uniform channel as a channel, whose channel matrix consists of rows and columns, which are permutation of the first row. In literature, this uniform channel is also known as a symmetric channel. Today, we look at the procedure to calculate channel capacity for a few more channels of importance. One such channel is a weekly symmetric channel. So, let us look at the definition for the same.

(Refer Slide Time: 01:38)



A channel is said to weakly symmetric if every row of the channel matrix P is a permutation of every other row and all columns sums th at is probability of b j given a i for all i are equal. For example, a channel is with the channel matrix given by one third, one sixth, half, one third, half, one sixth. In this channel, both the rows are permutations of each other. The sum of column is equal to two third for j equal to 1, 2 and 3. This channel is weakly symmetric, but not symmetric. Let us calculate the capacity for weakly symmetric channel mutual information.

(Refer Slide Time: 04:10)



By definition, I is equal to H B minus H of B given A. The capacity of a channel is maximum of I A semicolon B over input probability. So, let us look at the calculation of first H of B given A. H of B given A is equal to probability of a multiplied by summation probability b given a log of 1 by probability of b given a summation of B summation over A.

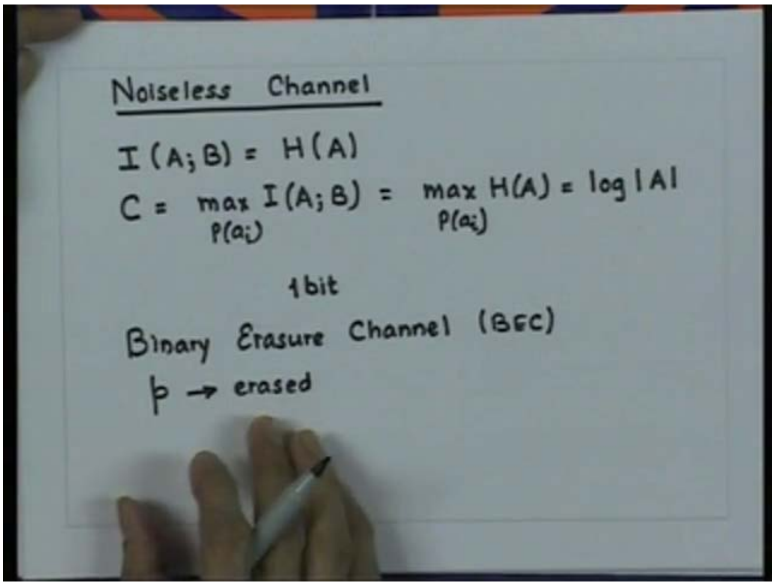
Now, the summation over B is a summation for each a i of the terms in the ith row of the channel matrix. Now, for a weakly symmetric channel however, this summation is independent of i. So, in that case, I of A semicolon B will be always less than equal to log of mod B minus H of row of matrix P. Mod b is the size of the output alphabet B. We know that the maximum value of H B is equal to log times the size of the alphabet. So, this quantity is H of row of channel matrix P. Now, equality will hold if and only if the output distribution is uniform, in which case we will get the capacity for the channel.

Now, in this case, equality it can be shown in this case that if probability of a i is equal to 1 by mod A whereby definition again mod A is equal to size of the input alphabet A. Then if P a i is equal to 1 by A, then this distribution achieves a uniform distribution on B. This is very clear. If we note that probability of b j is equal to probability of b j given a i multiplied by probability of a i sum over A, this is equal to 1 by mod A probability of b j given a i sum over A.

Now, for a weakly symmetric channel, this quantity is constant for each b j. So, this can be written as a constant, which is C by mod A. Since, this is true for all j equal to 1 to s where s denotes the size of the output alphabet B, then this can written as 1 by the size of output alphabet B. Therefore, the capacity of weakly symmetric channel is equal to log of mod B.

Mod B denotes the size of the output alphabet minus H times row of channel matrix P. This capacity is achieved for the input distribution which is uniform. Now, let us look at another channel. Now, let us look at the capacity for a noiseless channel. We had seen earlier that mutual information of a noiseless channel is given by H of A.

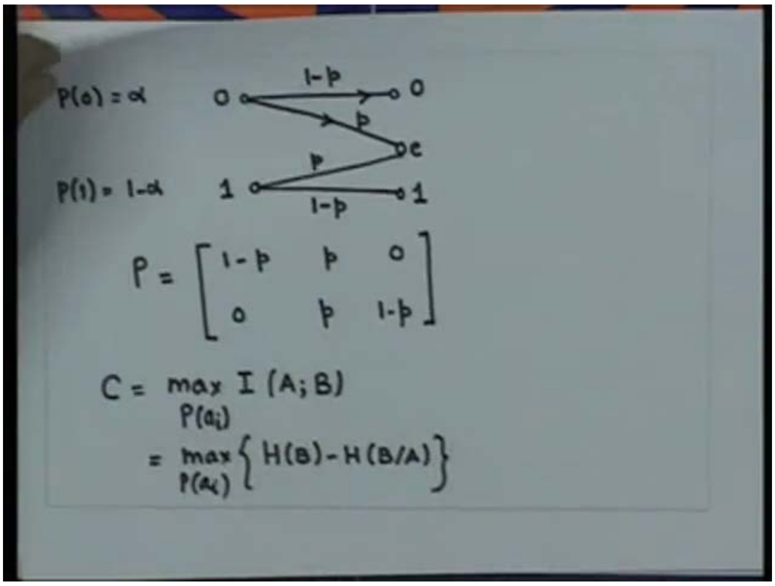
(Refer Slide Time: 09:33)



Therefore, capacity for a noiseless channel is maximum. I A semicolon B over P a i is equal to maximum H A over P a i. This is equal to log of mod A, where mod A denotes the size of the input alphabet A. So, if you have a noiseless binary channel, the capacity of a noiseless binary channel will be equal to 1 bit.

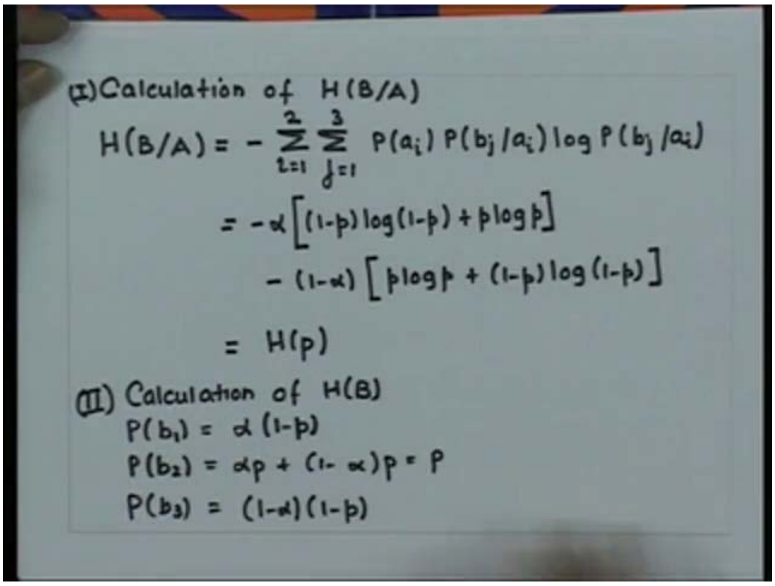
Another important channel of practical use is binary erasure channel. This is analogue of the binary symmetric channel in which some bits are loss rather than corrupted in the binary erasure channel that is BEC. A fraction of p of the bits is erased. The receiver knows which bits have been erased. The BEC has 2 inputs and 3 outputs. The channel diagram for binary erasure channel is as follows.

(Refer Slide Time: 11:53)



We have 2 inputs, 1 and 1. We have 3 outputs. 1 e denotes the erased bit and 1. This probability is given by p. This probability if it is symmetric binary channel is again p. Therefore, this is 1 minus p and 1 minus p. So, let us calculate the capacity of a binary erasure channel. The channel matrix for this channel will be as follows 1 minus p, p, 1, 1, p, 1 p. Let us assume probability of 1 at the input is equal to alpha. Probability of 1 is equal to 1 minus alpha. So, calculation of the capacity requires calculation of mutual information and maximum over a i. So, this is equal to maximum of P a i of H of B minus H of B given A. Let us look at the calculation for H of B given A.

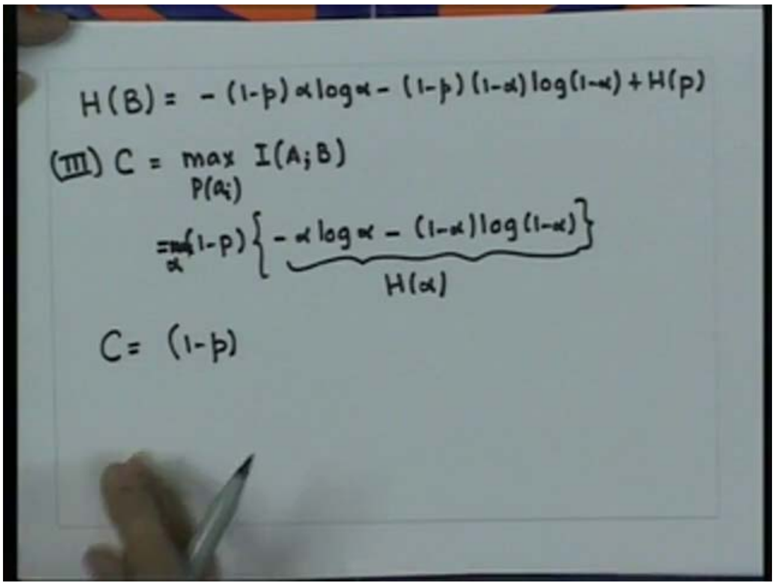
(Refer Slide Time: 14:02)



This is equal to minus double summation probability of a i probability of b j given a i log of probability b j given a i from this channel matrix and with this probability model. For the input alphabet A, we get H of B given A is equal to minus alpha 1 minus p log 1 minus p plus p log p.

This quantity corresponds to the first row of the channel matrix minus 1 minus alpha p log p plus 1 minus p log 1 minus p. This can be simplified as H of P. Now, to calculate H of B, we need to calculate output probability distribution. So, probability of b 1 is equal to alpha 1 minus p. Probability of b 2 is equal to alpha p plus 1 minus alpha p is equal to p and probability of b 3 is equal to 1 minus alpha into 1 minus p.

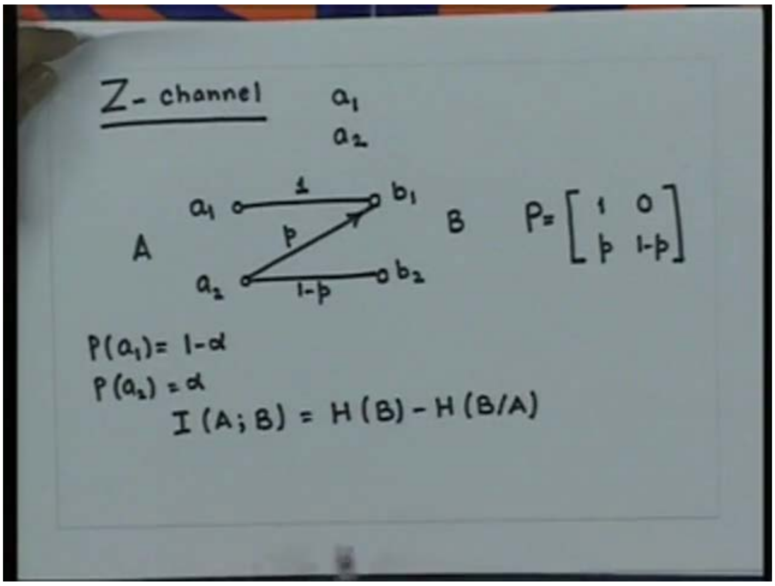
(Refer Slide Time: 16:54)



Using these probabilities, we can show that H of B is equal to minus 1 minus p alpha log alpha minus 1 minus p into 1 minus alpha log of 1 minus alpha plus H of p. Therefore, to calculate C, which is equal to maximum of P a i I A semicolon B, this is equal to 1 minus p into minus alpha log alpha minus 1 minus alpha log 1 minus alpha.

So, maximum over alpha, although we can determine the optimal value of alpha by taking the derivative of I A semicolon B with respect to alpha and equating to 0. We can find it directly from the expression from mutual information I A semicolon B by realising the part between the curly brackets is really H alpha. Therefore, its maximum value is 1. Hence, C is equal to 1 minus p. This value is achieved when alpha is equal to half; that is the input probability distribution is uniform.

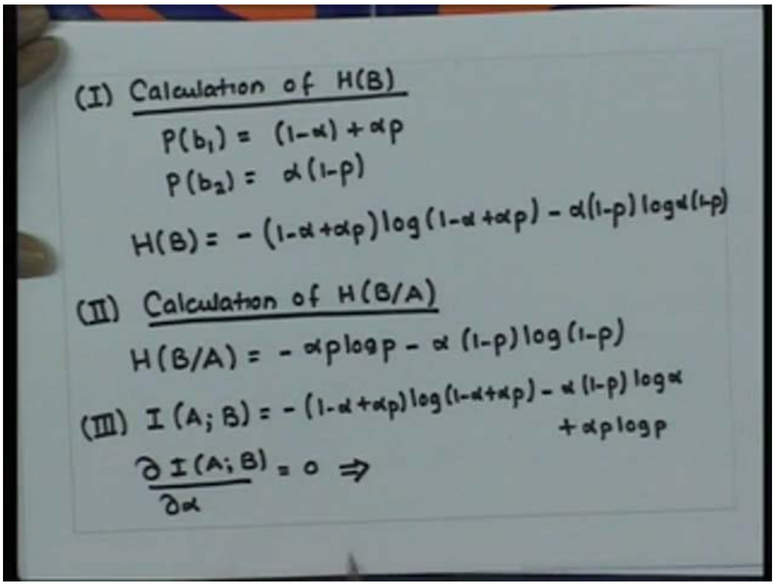
(Refer Slide Time: 19:34)



Now another channel of importance is known as Z channel. For this model of a binary channel, it is assumed that exactly 1 of the 2 input symbols say a 1 is transmitted without errors, while the other symbol say a 2 may be incorrectly received with a probability P. So, channel model for such a channel is depicted in a figure as shown here. a 1 and a 2 are input alphabets. b 1 and b 2 are output alphabets. This probability is 1. This is p and this is 1 minus p. So, the channel matrix for this channel is denoted by 1, 0, p, 1 minus p.

Let us determine the channel capacity for Z channel with the assumption that probability of a 1 is equal to 1 minus alpha and probability of a 2 is equal to alpha. So, as usual we require calculating mutual information, which is H B minus H of B given A.

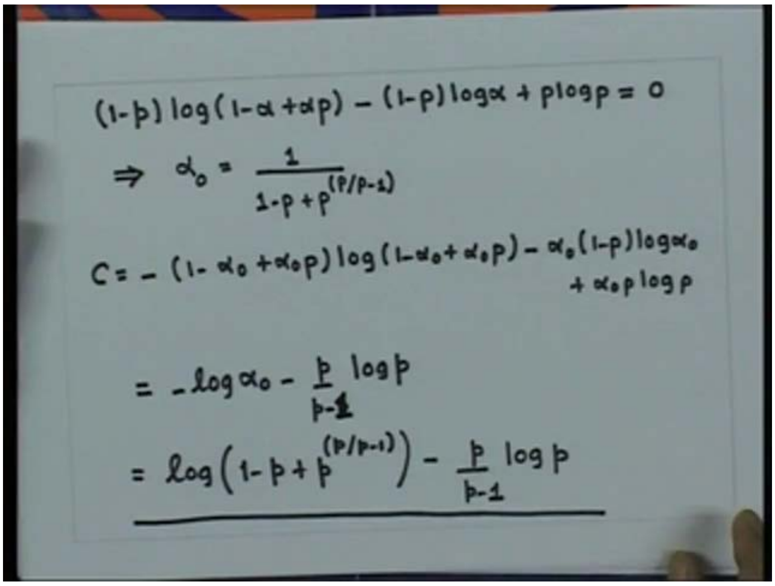
(Refer Slide Time: 21:31)



So, calculation of H B requires calculation of probability of output symbol b 1 and b 2. It is given by P of b 1 is equal to 1 minus alpha plus alpha p. Probability of b 2 is equal to alpha into 1 minus p. Using this probability, we can write down H of B as minus 1 minus alpha plus alpha p log of 1 minus alpha plus alpha p minus alpha times 1 minus p log of alpha times 1 minus p. The second quantity, which we require to calculate, is H of B given A.

Using this channel matrix and using this input probability model, we can easily write down H of B given A is equal to minus alpha p log p minus alpha 1 minus p log of 1 minus p. Therefore, we get mutual information I A semicolon B equal to minus 1 minus alpha plus alpha p log of 1 minus alpha plus alpha p minus alpha 1 minus p log of alpha plus alpha p log p. This is the expression, which we get after subtracting this quantity from this quantity and simplifying. Now, in order to calculate the capacity, we have to take the derivative of this quantity with respect to the variable alpha.

(Refer Slide Time: 24:58)

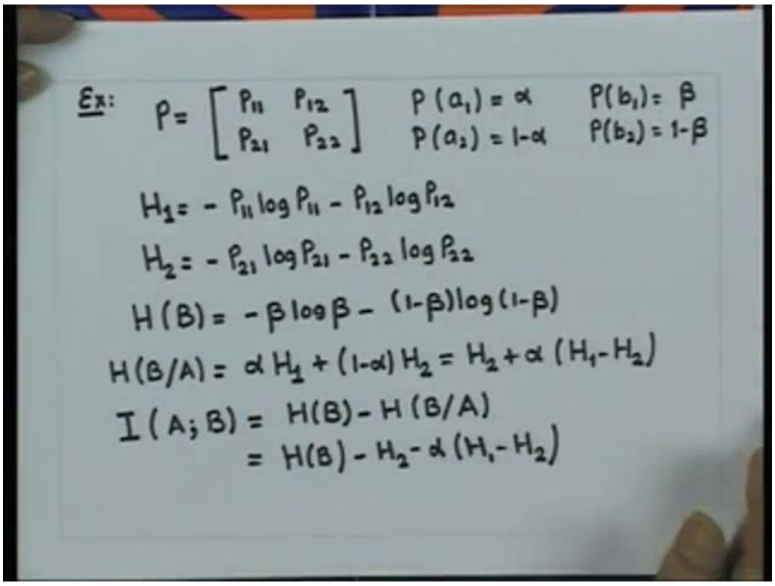


So, if we take the derivative of I A semicolon B with respect to alpha and equate to 0; this will imply that 1 minus p into log of 1 minus alpha plus alpha p minus 1 minus p log alpha plus p log p is equal to 0. This on solving for the optimum value of alpha, which we call it as alpha naught we get the relationship as follows.

So, if we take this relationship of alpha naught and plug it into the channel capacity expression, which is given by 1 minus alpha naught plus alpha naught p log of 1 minus alpha naught plus alpha naught p minus alpha naught 1 minus p log alpha naught plus alpha naught p log p. We get capacity equal to minus log alpha naught minus p times p minus p by p minus 1 log of p is equal to log of 1 minus p plus p raise to p upon p minus 1 minus p by p minus 1 log p.

So, this is the final expression, which we get for the channel capacity for a Z channel. Earlier we have looked at the procedure to calculate the channel capacity for a binary symmetric channel. Let us look at the methodology to calculate the channel capacity for a general binary channel, which is non symmetric.

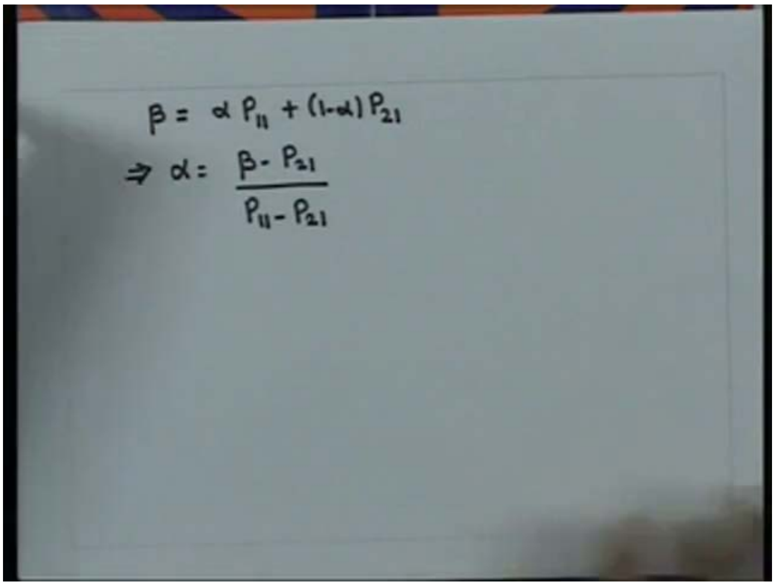
(Refer Slide Time: 27:46)



So, if we have a binary channel with a channel matrix given by P equal to P 1 1, P 1 2, P 2 1, P 2 2 where P 1 2 is not necessarily equal to P 2 1. We assume that input alphabet P a 1 is equal to alpha. So, P of a 2 is equal to 1 minus alpha. For the output alphabet b, probability of b 1 is equal to beta. So, probability of b 2 is equal to 1 minus beta. For calculation of capacity for such a general binary channel, we need to calculate H of B.

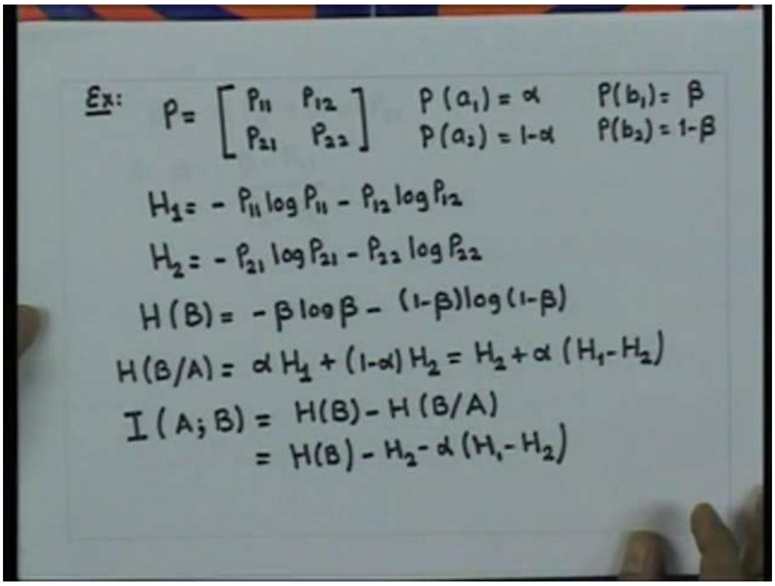
So, H of B can be calculated. If we define H 1 as minus P 1 1 log of P 1 1 minus P 1 2 log P 1 2, H 2 define as entropy of this row is minus P 2 1 log P 2 1 minus P 2 2 log 2 2. Once we have H 1 and H 2, we obtain H B. H of B given A is equal to alpha times H 1 plus 1 minus alpha times H 2, which is equal to H 2 plus alpha times H 1 minus H 2. Now, for this, we get the mutual information I A semicolon B, which is equal to H of B minus H of B given A as H of B minus H 2 minus alpha times H 1 minus H 2.

(Refer Slide Time: 31:16)



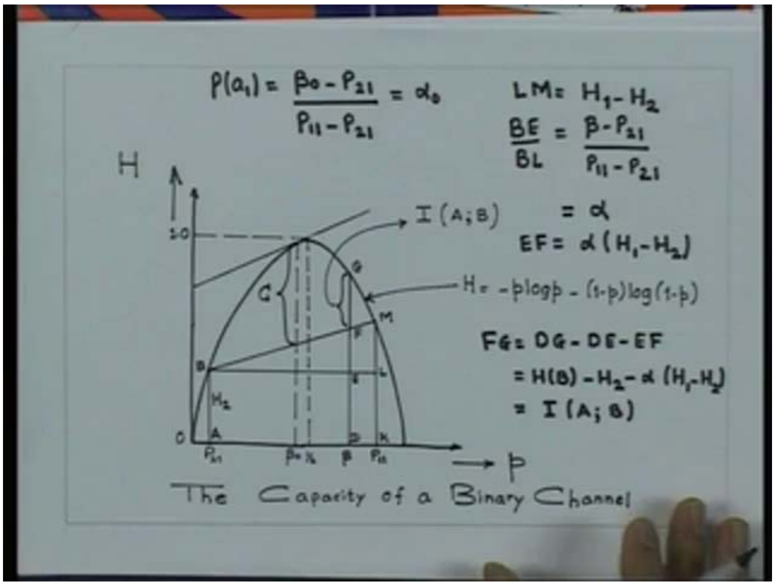
Now, with this channel matrix and with this input probability distribution probability of b 1, which is beta, can be easily calculated. It is alpha times P 1 1 plus 1 minus alpha P 2 1. From this, it implies that alpha is equal to beta times minus P 2 1 upon P 1 1 minus P 2 1.

(Refer Slide Time: 31:41)



Now, all this entropies can be depicted in a figure as shown here. In this figure, the length of B G represents the value of entropy of the output alphabet that is H B. The length of A B represents the value of H 2. Similarly, the length K M represents the value of H 1.

(Refer Slide Time: 31:51)

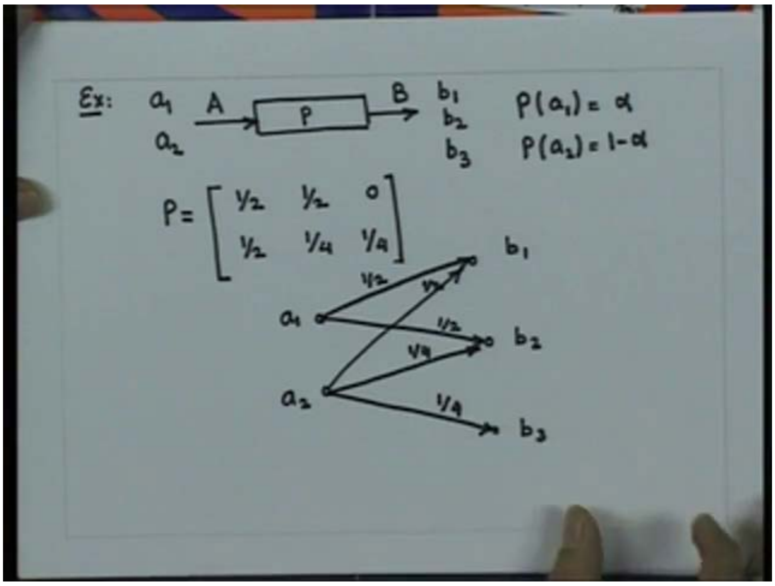


Therefore, L M which is equal to K M minus K L is equal to H 1 minus H 2. Now, triangle B E F and triangle B L M are symmetric. Therefore, B E upon B L is equal to E F upon L M. Now, B E upon B L from this figure is clearly equal to beta minus P 2 1 upon P 1 1 minus P 2 1, which is equal to alpha.

Hence, E F is equal to alpha times H 1 minus H 2. Now, F G is equal to D G minus D E minus E F. this is equal to h b minus h 2 minus alpha times h 1 minus h 2. It is equal to mutual information I A semicolon B. So, in order to obtain the capacity for a general binary channel, which is equal to the maximum of I A semicolon B over P a i, the fine channel capacity C by drawing a line parallel to B M and tangent to this curve which is the entropy function.

The point of contact determines the value of b naught of b of beta that belongs to the capacity. So, the optimal probability distribution of the source is therefore given by probability of a 1 is equal to beta naught minus P 2 1 upon P 1 1 minus P 2 1 is equal to alpha naught. So, this will be general procedure to calculate the capacity of a binary channel based on a graphical approach.

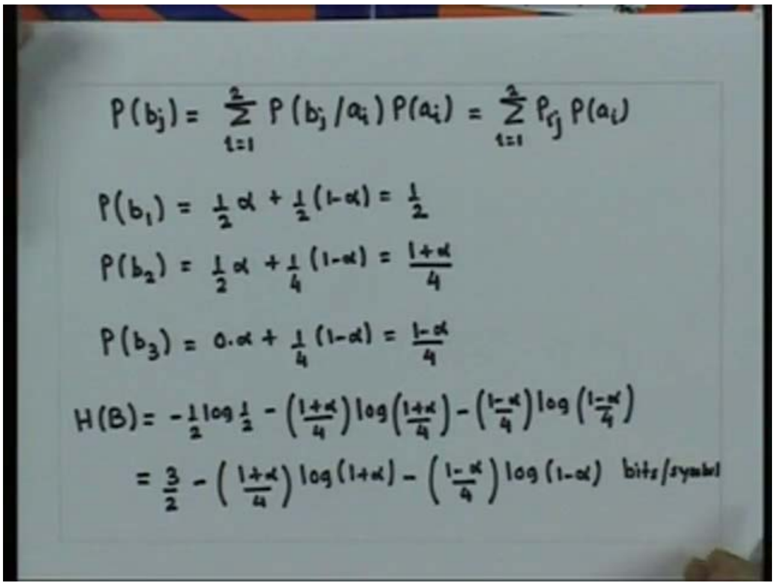
(Refer Slide Time: 36:25)



Now, as a final example of our exercise in calculation of channel capacity, let us take a channel with input alphabet consisting of 2 symbols a 1 and a 2 and output alphabet B consisting of 3 symbols b 1, b 2, b 3 with the channel matrix given as half, half, 0, half, one fourth and one fourth. So, for this probability of a 1 is equal to alpha and probability of a 2 is equal to 1 minus alpha.

Now, for this channel, the channel diagram can be drawn as follows. So, in order to calculate the capacity, we need to calculate the mutual information for which we need to calculate H of B, which in turn demands that we calculate the output probabilities of the symbols b 1, b 2 and b 3.

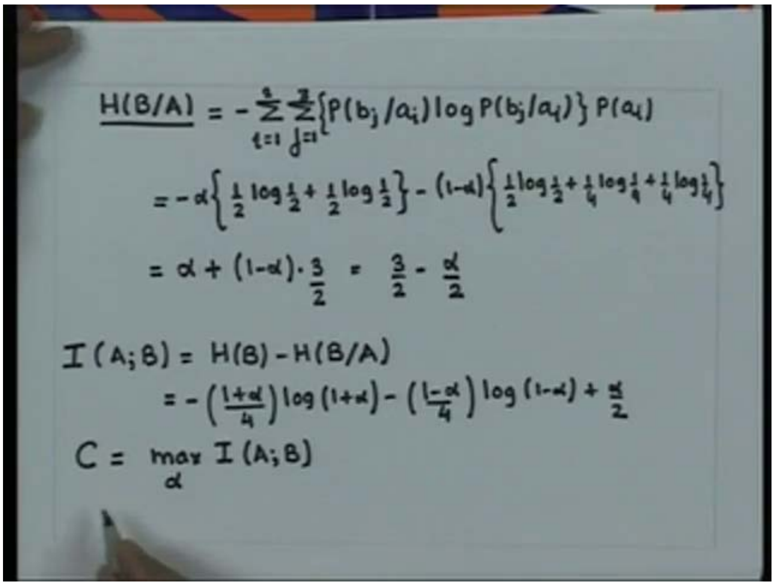
(Refer Slide Time: 38:33)



Probability of b j is equal to probability of b j given a i multiplied by probability a i. i is equal to 1 to 2. It is equal to P i j multiplied by P a i, i equal to 1 to 2 for the given channel matrix. We can show that probability of b 1 is equal to half alpha plus half 1 minus alpha is equal to half. Probability of b 2 is equal to half alpha plus one fourth 1 minus alpha is equal to 1 plus alpha by 4. Finally, probability of b 3 is equal to 0 multiplied by alpha plus one fourth multiplied by 1 minus alpha.

Given these probabilities, we can calculate H of B, which is equal to minus half log half minus 1 plus alpha by 4 log of 1 plus alpha by 4 minus 1 minus alpha by 4 log of 1 minus alpha by 4. This can be simplified as equal to 3 by 2 minus 1 plus alpha by 4 log of 1 plus alpha minus 1 minus alpha by 4 log of 1 minus alpha bits per symbol.

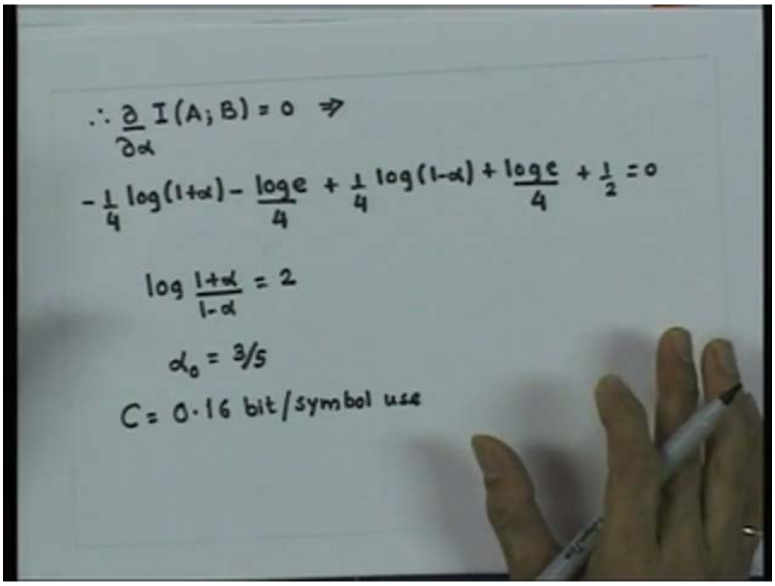
(Refer Slide Time: 41:13)



Next step is to calculate H of B given A, which is equal to double summation of probability b j given a i log of probability b j given a i multiplied by probability of a i. i is equal to 1 to 2, j is equal to 1 to 3 for the given channel matrix. We can show that this quantity is equal to minus alpha times half log half plus half log half minus 1 minus alpha half log of 1 by 2 plus 1 by 4 log of 1 by 4 plus 1 by 4 log of 1 by 4. This can be simplified as equal to alpha plus 1 minus alpha times 3 by 2 is equal to 3 by 2 minus alpha by 2.

So, we get mutual information as H B minus H of B given A equal to minus 1 plus alpha by 4 log of 1 plus alpha minus 1 minus alpha by 4 log of 1 minus alpha plus alpha by 2. Now, channel capacity is maximum of I A semicolon B over P a i, which is equivalent to maximum over alpha.

(Refer Slide Time: 44:13)



Therefore, this implies I A semicolon B equal to 0, which implies that minus one fourth log of 1 plus alpha minus log e by 4 plus 1 by 4 log 1 minus alpha plus log e times log e by 4 plus half is equal to 0. This implies log of 1 plus alpha upon 1 minus alpha is equal to 2. Therefore, the optimum value of alpha is equal to 3 by 5. If we plug in this value into the mutual information, which is given by this expression where alpha is substituted by alpha naught that is the optimum value. We get the channel capacity as 0.16 bit per symbol use.

So, we have looked at a very general procedure of calculating channel capacity for a channel. Now, in the next class, we will begin our study which will lead to channel second theorem; the most surprising as well as the most important single result of information theory. The channel second theorem deals with the amount of error free information, which we can get through a channel. We will also look at the relationship between the error probabilities and the mutual information in a channel.