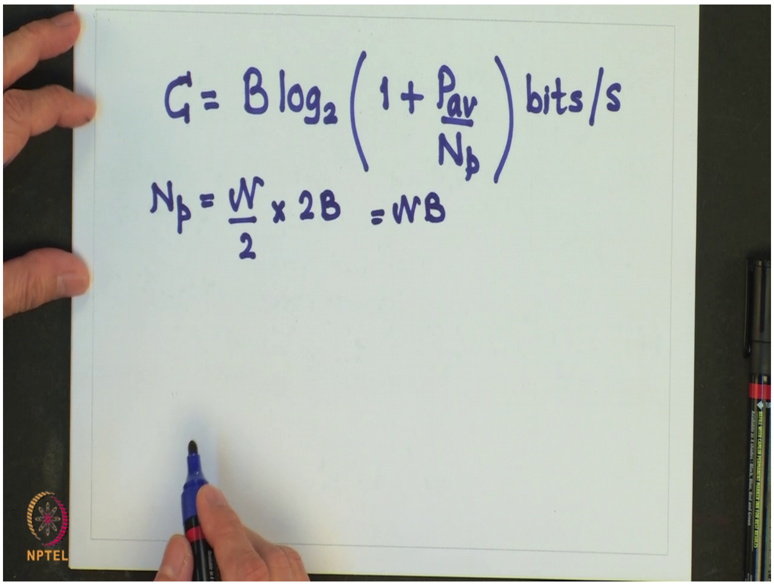
Principles of Digital Communications
Prof. Shabbir N. Merchant
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture - 14
Summary of Information Theory

We have seen that when the channel noise is Additive white and Gaussian with mean square value given say as N p, the channel capacity C of a band limited channel under the constraint of given signal power P average is given by this relationship.

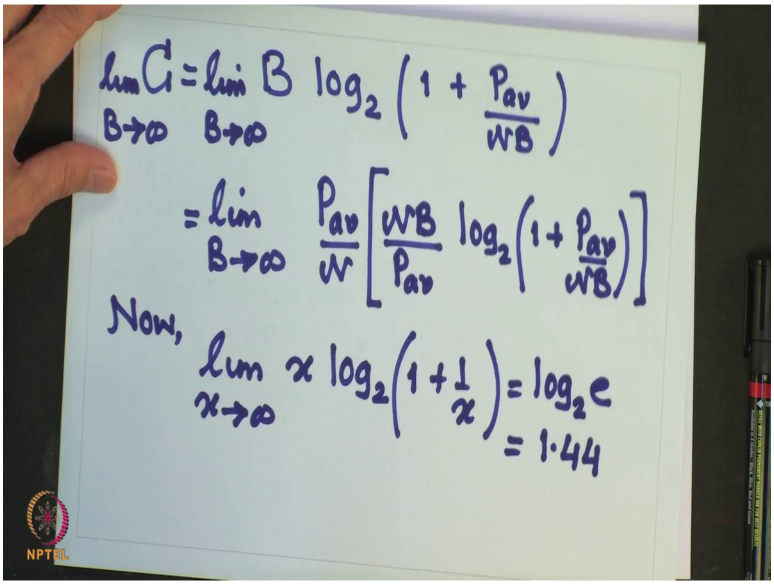
(Refer Slide Time: 00:38)



This is the signal power, this is the noise power bits per second and this is your channel bandwidth in hertz.

Now, the maximum rate of transmission can be realized only if the input signal is a white Gaussian signal ok. The question is now what happens to the channel capacity, if I increase a signal power; from this expression it is very clear that when I increase the signal power, my capacity will also increase. But what will happen if bandwidth increases? Say it goes to infinity; will channel capacity also go to infinity? Superficially the equation seems to indicate that, but this is not however, true correct. The reason for that is basically to realize that the noise power N p is equal to power spectral density multiplied by twice the bandwidth correct. So, when the bandwidth increases the noise power also increases correct.

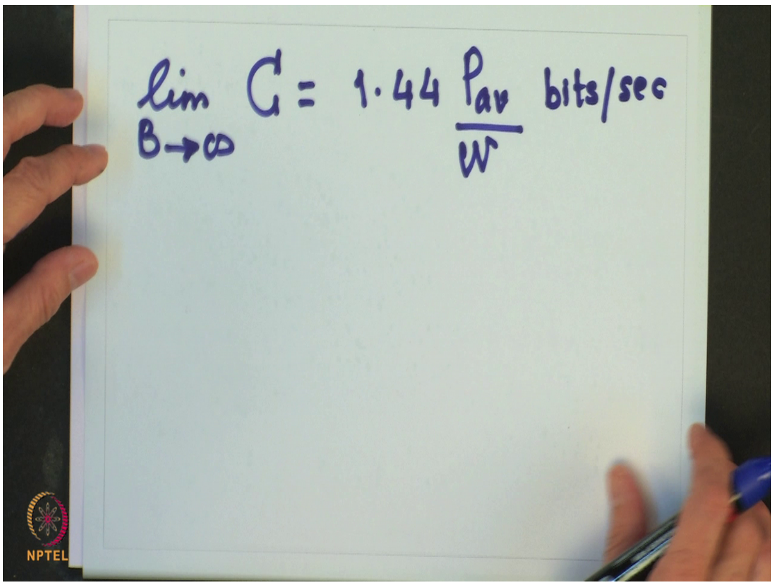
(Refer Slide Time: 02:13)



So, let us try to evaluate the channel capacity when B tends to infinity fine. So, our C is equal to B log to the base 2, 1 plus signal power divided by noise power. Let us take the limit of this for B tending to infinity. This I can rewrite it as and the reason for doing this is basically I want to exploit this identity, which says that limit x tending to infinity of x log to the base 2 of 1 plus 1 by x is equal to this is equal to 1.44 correct.

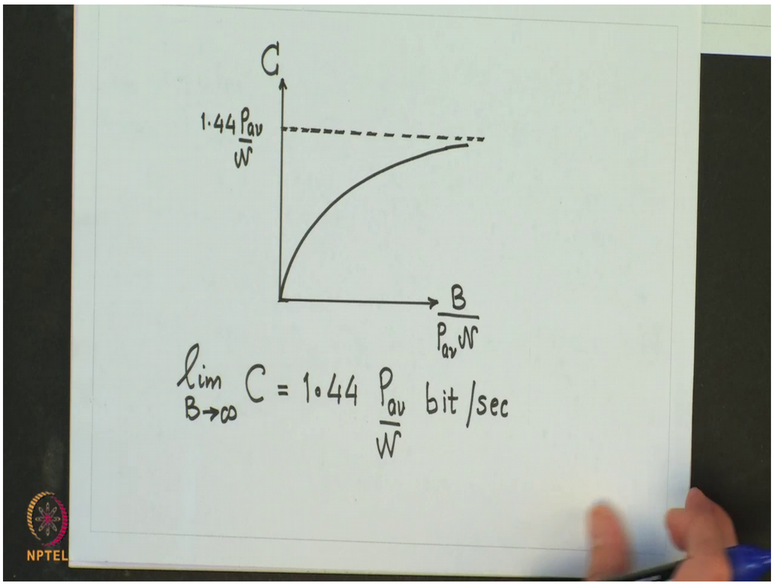
So, now if I use this identity it is very clear from here right that my capacity when B tends to infinity is equal to 1.44 P average by N bits per second.

(Refer Slide Time: 03:47)



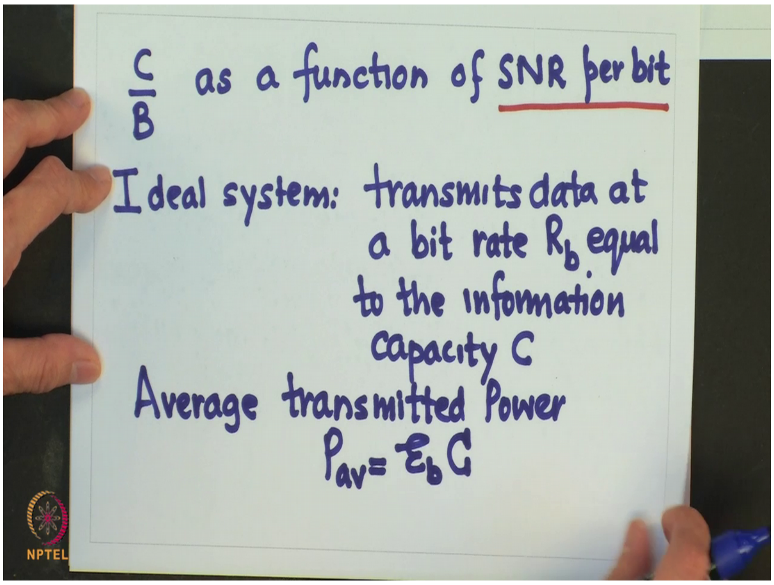
So, for a white Gaussian channel noise the channel capacity C approaches this limit as B tends to infinity. So, the variation of channel capacity with bandwidth is shown in this figure.

(Refer Slide Time: 04:23)



So, from this figure it is evident, that a capacity can be made infinite only by increasing the signal power to infinity. For finite signal and noise power, the capacity will always remain finite and this relationship we just derived.

(Refer Slide Time: 04:52)

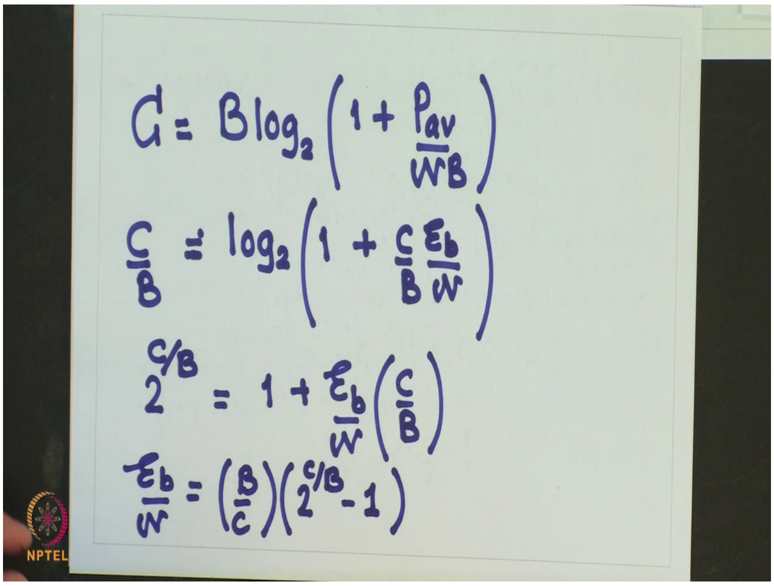


Now, it is also instructive to express the normalized channel capacity which is C by B as a function of signal to noise ratio per bit right. In digital communication; unlike analog communication, this plays an important role. We will see this later in our course also; signal to noise ratio per bit is an important parameter and we will see what is that ok.

So, let us assume that we have an ideal system, correct what is that ideal system it transmits data at a bit rate, this information bit rate R b, equal to the information or channel capacity C correct. So now if we assume this then we also know this relationship Average transmitted Power which is P average is nothing but energy per bit multiplied by the rate and in our case we are assuming ideal system; an ideal system the rate is same as the channel capacity C ok.

So, if that is the case then I can write my channel capacity expression as follows.

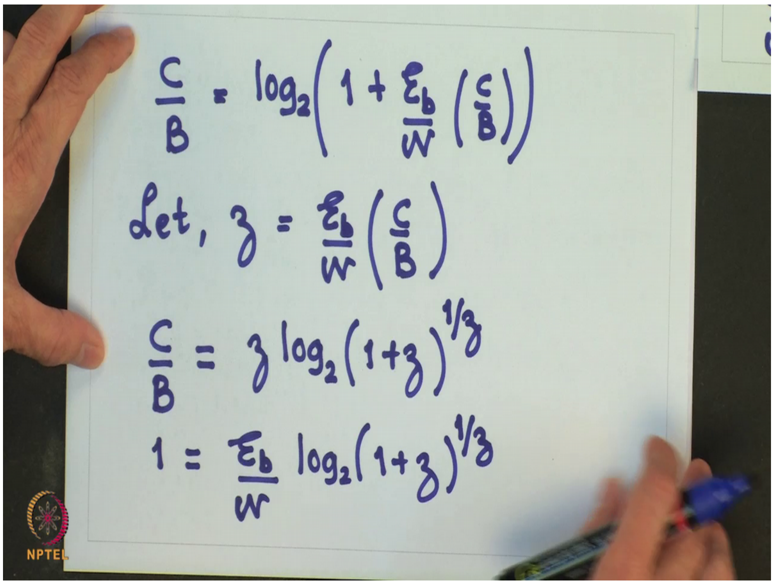
(Refer Slide Time: 07:00)



So, C by B for P average I substitute this quantity and E b by this italic N is known as signal to noise ratio for bit ok fine. So, from this relationship it is easy to see that I will get; so from this I get it as signal to noise ratio per bit is equal to I get this quantity fine.

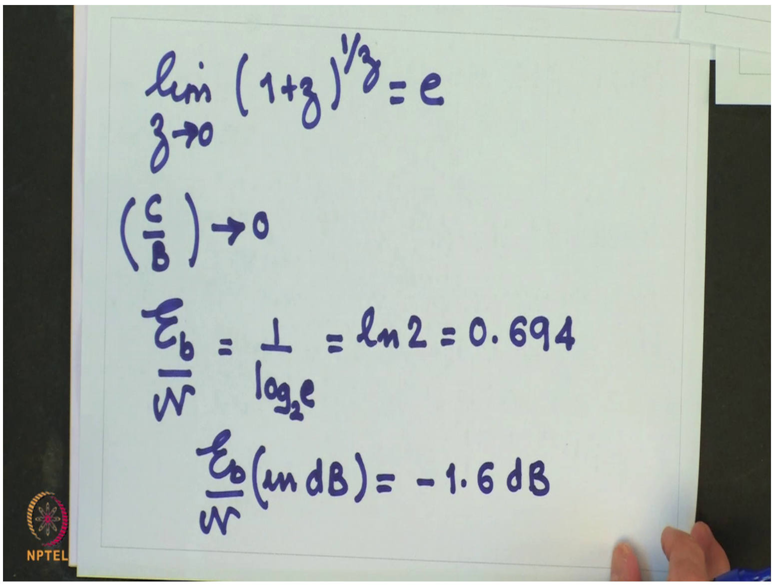
Now, I am interested basically I can we will see the plot of the signal to noise ratio per bit and normalize rate correct. But before we do that let me try to evaluate what would happen to this signal to noise ratio per bit; when the bandwidth becomes infinite correct.

(Refer Slide Time: 08:55)



Now I will this C by B is equal to log to the base 2 of 1 plus ok. Now, let Z equal to this quantity ok. So, I can write my C B from here as equal to fine very simple. So, this is equal to this will get cancelled and I will get this quantity, which is equal to this ok fine. So, what I get from here that 1 is equal to I Z is equal to this quantity; so I can write this. Why I am doing all this because, I want to see what is happens through this E b by N that signal to noise ratio per bit, when this quantity tends to 0 that is be tending to infinity.

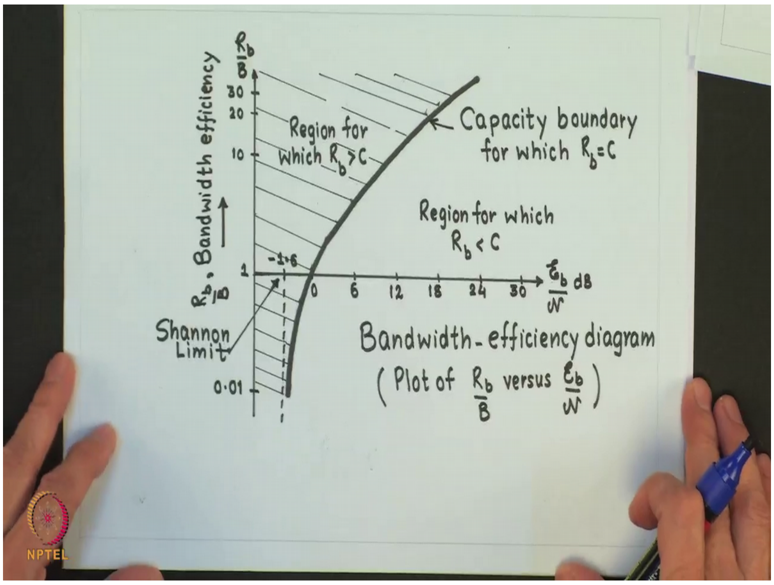
(Refer Slide Time: 10:36)



Now, we know that again this is relationship limit Z tending to 0 of 1 plus Z raised to 1 by Z is equal to e correct. So, in our case when C by B tends to 0 that is Z tending to 0, we will get it as E b by N this quantity correct. So, is equal to 1 by log to the base 2 e which is same as log 2, this is natural algorithmic and this is equal to 0.694. So, E b by N in d B that is N log of this quantity turns out to be minus 1.6 d B fine.

So, now coming back to this equation let us plot this equation correct. So, here is the plot for that.

(Refer Slide Time: 11:52)



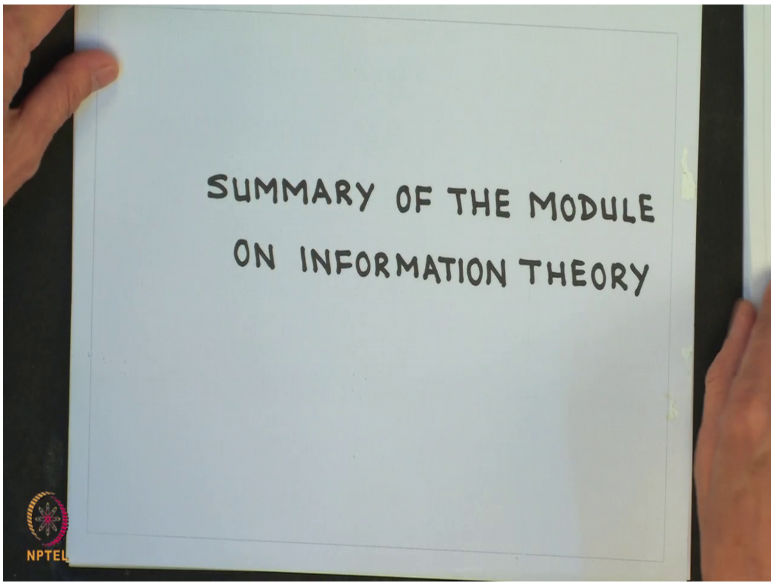
So, if I plot that equation correct. So, this is the curve which I will get this is the curve which I get correct. So, this is the curve which I get for this equation correct. So, E b by N is here and this C by B basically is this correct. So, this is your ideal system where my transmission rate is the same as the channel capacity C and now we also seen basically that in the limit in the limit, what happens when this quantity bandwidth goes to infinity; that means, this quantity goes to 0, this E by N signal to noise ratio per bit goes to this, and that is why you see that basically as this quantity goes to 0 correct on this axis y axis correct; this will tend to this value will tend to minus 1.6 d B that is what happened so this is the curve fine.

Now, so on the x axis we have signal to noise ratio per bit in d B and y axis basically what I am showing you is the, the normalized channel capacity correct. So, for general if you the transmission rate is not equal to channel capacity C, then it implies that you will be lying in this region you will be in this region above this correct and when your rate is less than the channel capacity C you would be operating in this region.

Now, by Shannon's coding theorem the rate of transmission has to be always less than equal to the channel capacity, this we have seen earlier correct. So, your region of operation is going to be this side whereas, this portion is the region where you cannot get error free transmission correct. So, this is a Shannon's limit in terms of for the signal to noise ratio per bit and this is known as bandwidth efficiency diagram; which is a plot of R b by B versus E b by italic N correct and this remember is a capacity boundary for the ideal system that is your R b is equal to C fine.

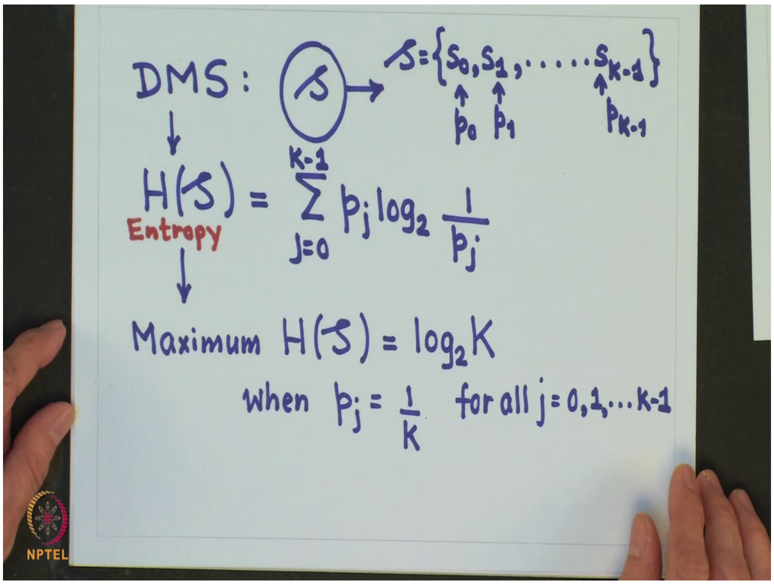
So, prohibited region of operation and the region of operation will be below this curve and this is known as your capacity boundary, this is a critical rate R b equal to C. So, with this we come to the end of the module on information theory; let us summarize our study of this module on information theory.

(Refer Slide Time: 15:00)



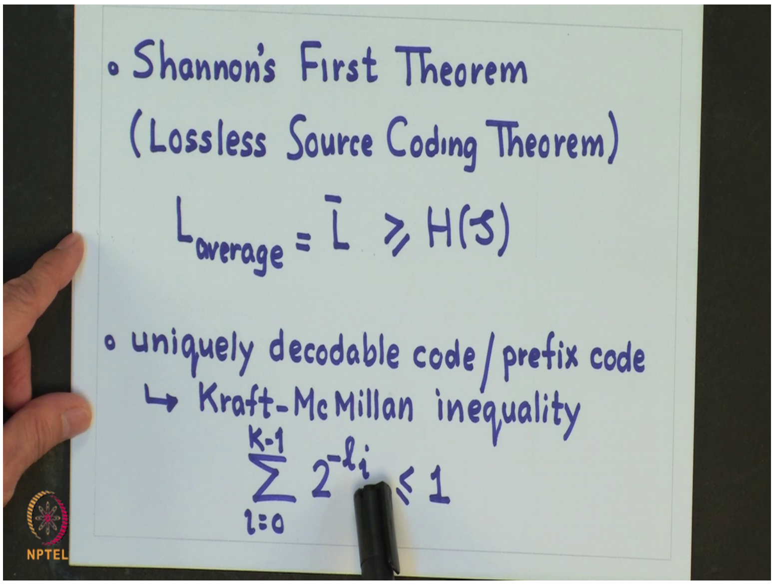
So, the motivation for studying information theory was to address 2 key issues in the evaluation of the performance of a digital communication system and this key issues were the efficiency with which information from a given source can be represented and the rate at which information can be transmitted, reliably over a noisy channel.

(Refer Slide Time: 15:32)



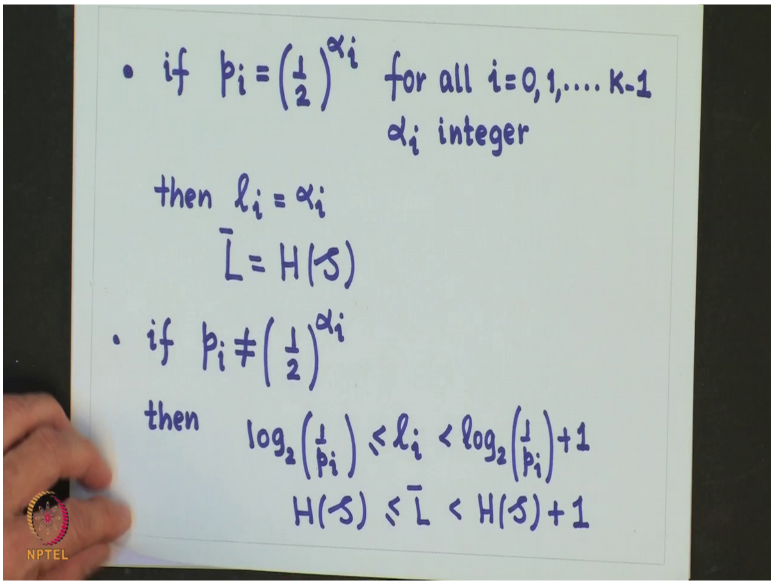
We began our study with a discrete memoryless source and we provided a mathematical model for the same, in the form of source alphabet consisting of source symbols and the associated probability distribution with this symbols and we defined a measure of information for the output symbol s j in the form of log to the base 2 1 by p j, where p j is the probability of occurrence of the symbol s j and then we went on to define the entropy of the source which is nothing but the average information per symbol and we showed that this entropy is maximum when the probability distribution is uniform meaning that p js is equal to 1 by K for all j equal to 0 1 up to k minus 1, where K is the number of symbols or letters in the source alphabet S.

(Refer Slide Time: 16:54)



Then we studied Shannon's first theorem, which is also known as Lossless Source Coding Theorem and what it said was that, the average length of the source code can never be less than the entropy of the source correct. Then we also mentioned Kraft-Mc Millan inequality, which is the necessary and sufficient condition to be satisfied by all uniquely decodable code or prefix code and this condition is basically on the length l i of the code words of the code.

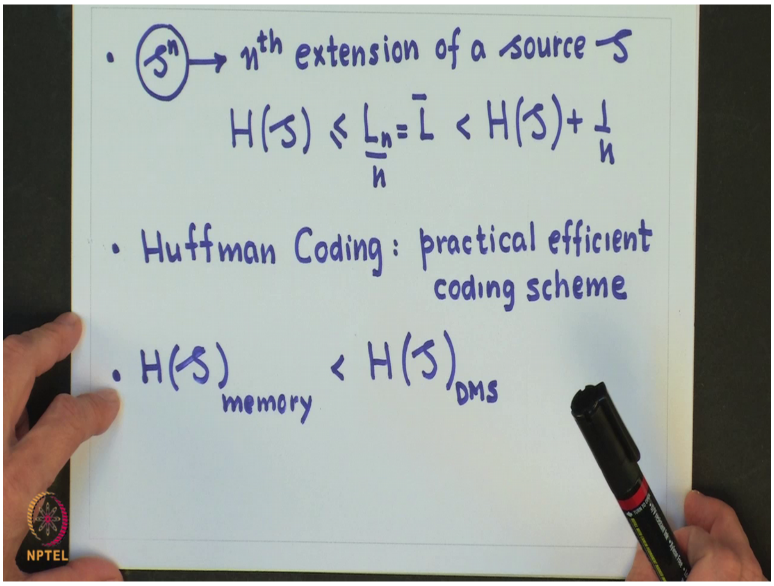
(Refer Slide Time: 17:37)



And what we saw there was that, if we had P i equal to half raised to the power of integer alpha i; for all i equal to 0 to k minus 1, where alpha is integer then the length of the code words could be chosen to be equal to that integer alpha i and in that case, the average length of the code turns out to be equal to the entropy of the source, which is the minimum value which he can have.

But if this condition is not true; that means, if p i is not of the form half raised to say some integer alpha i, then we said that if we follow this strategy of choosing the length of the code word l i to be the smallest integer greater than equal to log to the base 2 of 1 by p i then we can show that the average length of the code turns out to be within 1 binit of the entropy of the source.

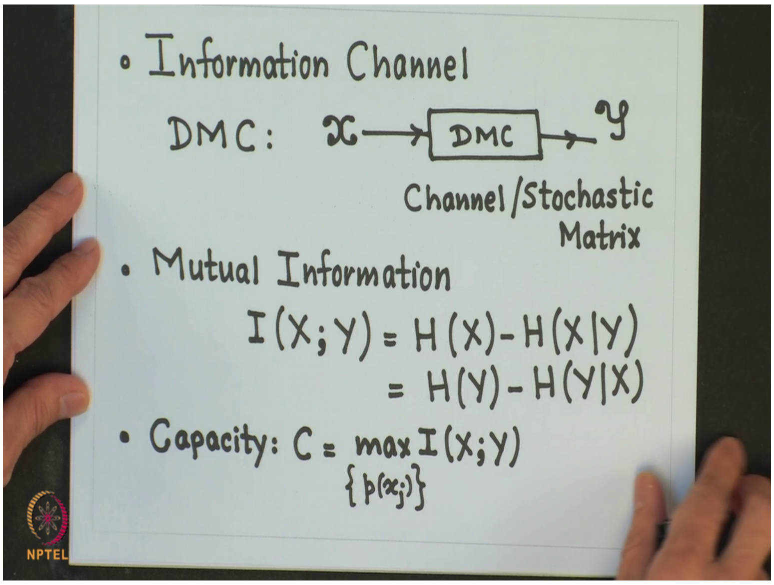
(Refer Slide Time: 18:52)



And using this strategy, we said that we could get the average length of the code as small as possible correct, but not less than the entropy of the source. So, what we mean that, as close as possible to the entropy of the source by using the nth extension of the source for coding, but in the process we will be increasing the complexity of the coding.

And then we studied Huffman Coding, which is a practical efficient coding scheme for obtaining the compact code and we also mentioned that the entropy of a source with memory is always less than the entropy of the source for discrete memoryless source.

(Refer Slide Time: 19:30)

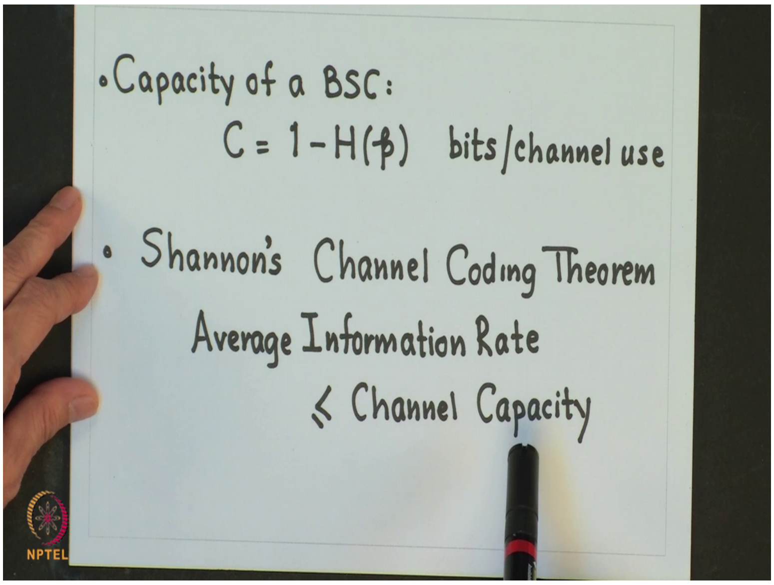


We define information channel in terms of input alphabet and output alphabet and the relation between the 2 is given in the form of channel or Stochastic Matrix.

The element in this channel matrix is nothing but the conditional probability of the output symbol given the input symbol. Then we went to define what is mutual information; this is the amount of information which gets transmitted on a channel and we showed that this mutual information turns out to be the difference between the entropy of the input alphabet and the conditional entropy of the input alphabet given the output alphabet right.

And then we define the capacity of the channel to be the maximum value of this mutual information and we showed that for a given channel, this maximization takes place over the probability distribution of the input alphabet.

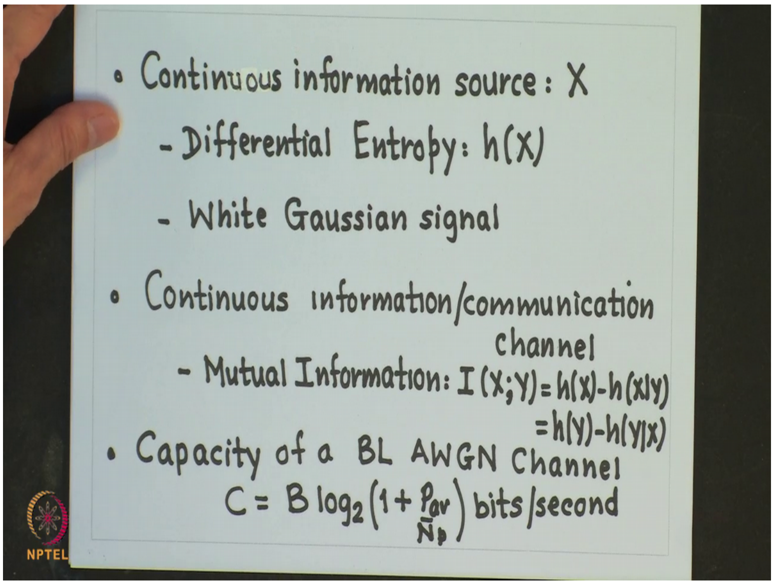
(Refer Slide Time: 20:30)



Having done this basically we computed the capacity of a interesting practical channel that is binary symmetric channel and we showed that the capacity of this channel, basically can be maximum equal to 1 bits per channel use and then we studied another important theorem, which is known as Shannon's Channel Coding Theorem and what that theorem states is that for reliable transmission with the error as small as possible as desired, the average information rate has to be less than or equal to channel capacity.

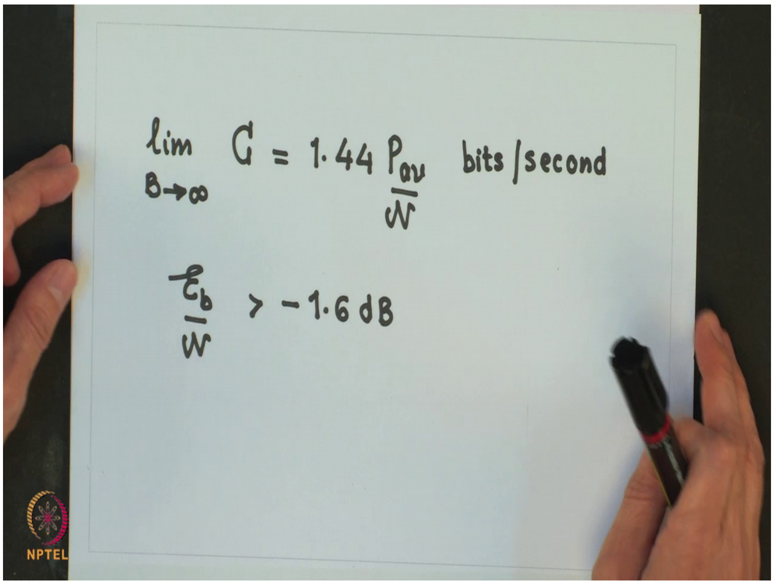
Then we moved over to the continuous information source and we defined the output of the continuous information source in terms of random processes and random signals and random variables, and what we showed was there we could define what is known as a differential entropy correct, which is different from the absolute entropy. It uses some kind of a reference or datum and then we showed that for any band limited signal; which is constrained to have certain mean squared value, the white Gaussian signal which is band limited has the maximum entropy per second.

(Refer Slide Time: 21:20)



And then we moved over to the study of continuous information or communication channel, and then again we showed there that the mutual information turns out to be difference of the differential entropy of the input and conditional differential entropy of the input given the output. And then we evaluated the capacity of a band limited Additive white Gaussian noise channel, which is of practical importance in design of communication systems.

(Refer Slide Time: 22:52)



And we showed that the capacity of a band limited Additive white Gaussian noise channel, when bandwidth tends to infinity is limited by this value given on this slide.

And we also showed that the minimum signal to noise ratio per bit, which we could have for reliable transmission, which is also known as Shannon's limit has to be larger than minus 1.6 d B. So, in short; the Shannon's first theorem, which is lossless source coding theorem and Shannon's channel coding theorem basically address our both the issues of efficiency of information representation from a given source, and the maximum rate at which the information can be transmitted reliably over a noisy channel. So, with this we come to the end of the module on information theory.

Thank you.