Principles of Digital Communications

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

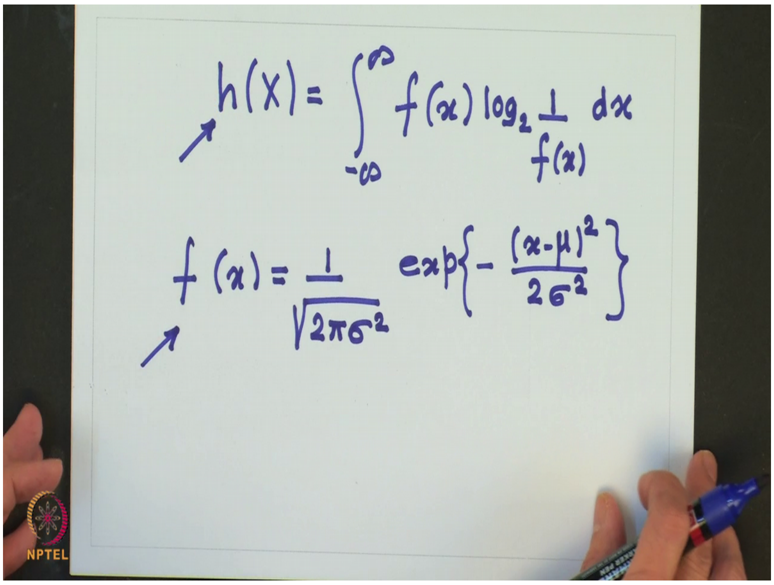
Lecture - 11
Differential Entropy - II

Hello, welcome back in the previous class we studied that absolute entropy for a continuous information source has the value which is infinite and therefore, practically it is of little use and therefore, we defined another measure of this average information in terms of what is known as differential or relative entropy of a continuous source. It is important to note that this differential or relative entropy was defined with a reference that was minus limit delta x tending to 0 of log to the base 2 delta x.

Now, we are concerned with communication and in communication basically we are interested in transmission of information. Now we have seen in the discrete case that this transmission of information which we termed as mutual information turned out to be differences between the entropies that mutual information was defined as entropy of the source minus the conditional entropy of the source given the observed output that is the channel output.

So, we will see that in the continuous case also this transmission of information will be called as mutual information, will be related to this differential or relative entropies and now for this differential relative entropies if the reference is kept fixed is the same, then the difference of the entropies will have significance to define the mutual information correct. So, as far as the continuous source is concerned; our measure of average information will be in terms of differential or relative entropy correct; which is as follows.

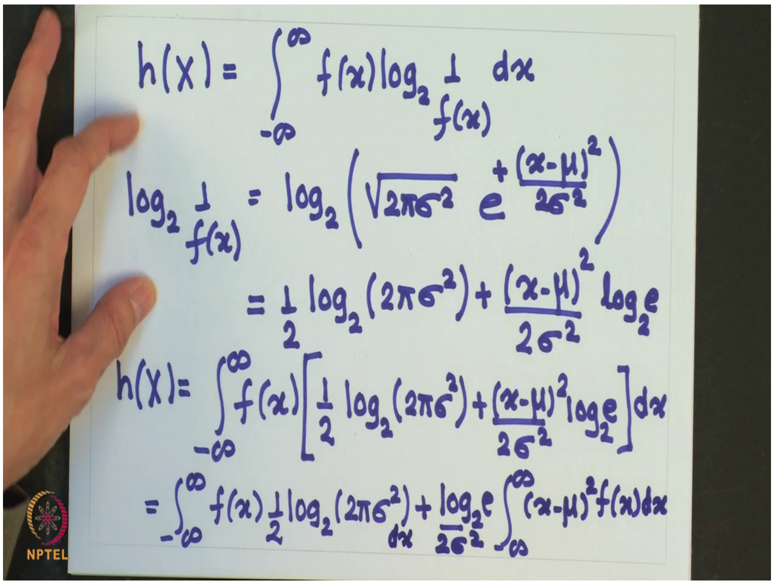
(Refer Slide Time: 02:52)



Where f x is the PDF of the random variable x and we also studied that to maximize this differential entropy for a given constraint on the random variable X in the form of variance; that means, we specify that the variance of the random variable is a constant equal to say sigma square, then we found out that PDF which satisfies this constraint and which maximizes differential entropy h X.

And what we showed was that it turned out to be a Gaussian or normal PDF sigma squared is the variance which has been specified, mu is the average of the random variable ok. So, this we had seen in the previous class. Now, what we will do is basically assuming that this is my PDF, let us try to evaluate the entropy or differential entropy for this PDF. My h X is equal to your log 2 1 by f x will be equal to we are assuming the Gaussian PDF.

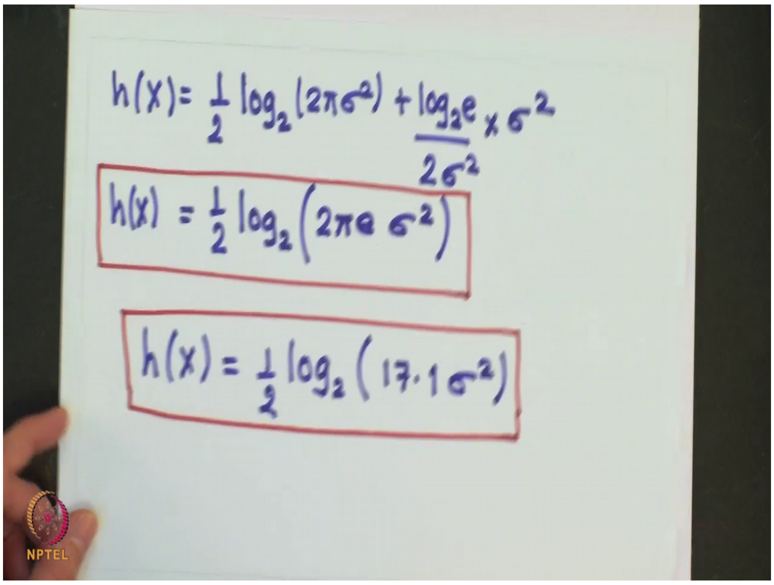
(Refer Slide Time: 04:31)



So, this will be equal to this expression which I am writing here and this can be written as this term comes because of the change of logarithmic to the base 2 to the natural log.

So, we can write my h X is equal to so, we integrate this term ok. So, we can add this; this is equal to that will be the one term and the second term would be fine ok. So now, sorry this is dx here ok. So, this basically because we know that PDF has to integrate over minus infinity to plus infinity equal to 1.

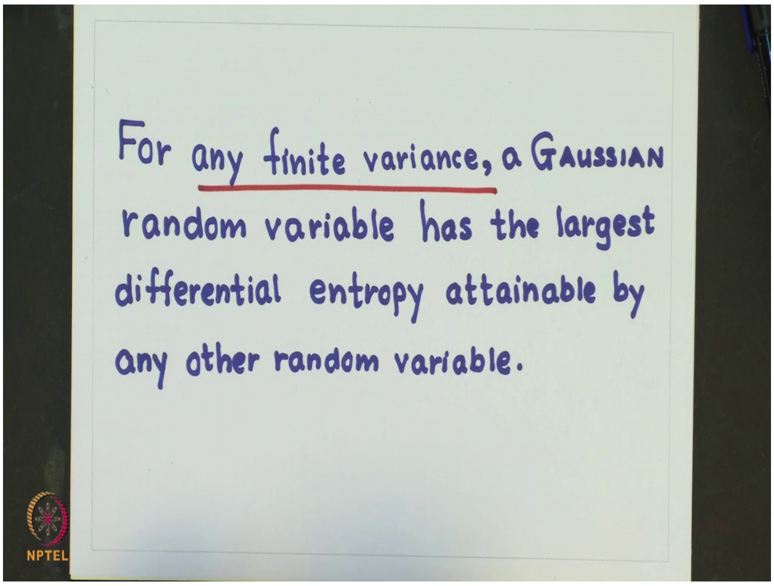
(Refer Slide Time: 07:25)



So, we can write this as the h X is equal to half to the base 2 2 pi sigma squared plus log 2 base to e by 2 sigma squared and please note that this term by definition is equal to sigma squared. So, what I get is 2 pi e sigma squared ok. So, this is an important relationship; you can also write this as substituting the values of pi e n fine.

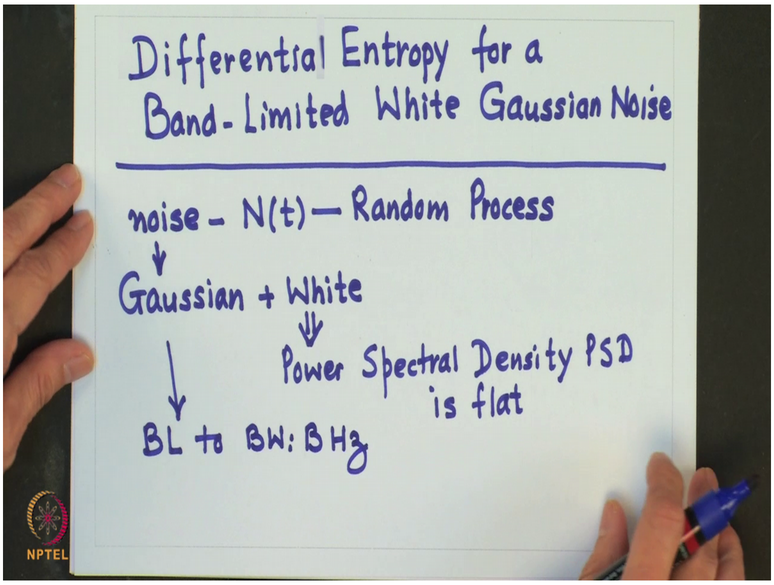
So, now for what it says basically that for Gaussian PDF this will be your differential entropy and that is a maximum value which you can get fine. So, we can say that for any finite variance or Gaussian random variable has the largest differential entropy attainable by any other random variable correct that is important to note this for any finite variance.

(Refer Slide Time: 08:58)



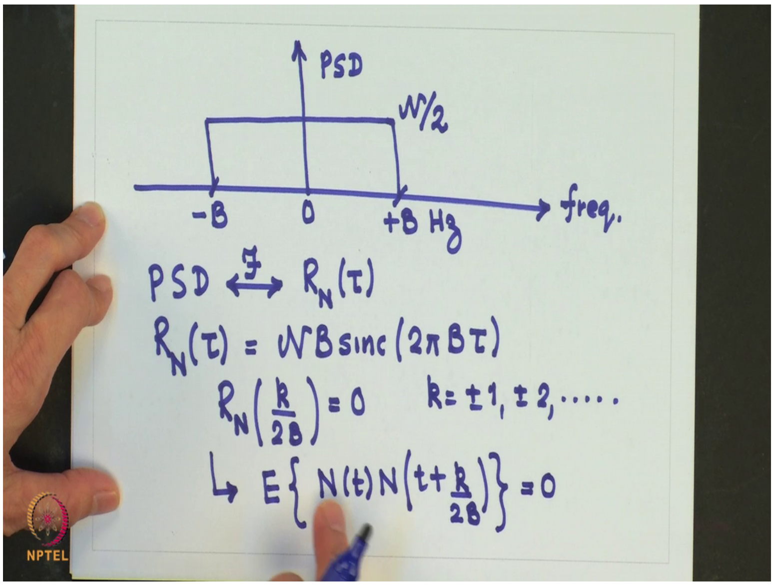
Let us try to evaluate for a practical signals this entropy and we will take a case of a Gaussian noise correct which is a band limited correct. So, what I am interested is basically in calculation of differential entropy for a band limited white Gaussian noise.

(Refer Slide Time: 09:38)



Let us try to evaluate the differential entropy for this. So, my assumption is that I have a noise correct, I denote it by N t; this is a random process and this noise is Gaussian correct and it is also white; so because it is white it implies that if you take the Power Spectral Density of it right; this is going to be is flat correct and we are also assuming that this noise is band limited to say a bandwidth B hertz fine.

(Refer Slide Time: 11:32)

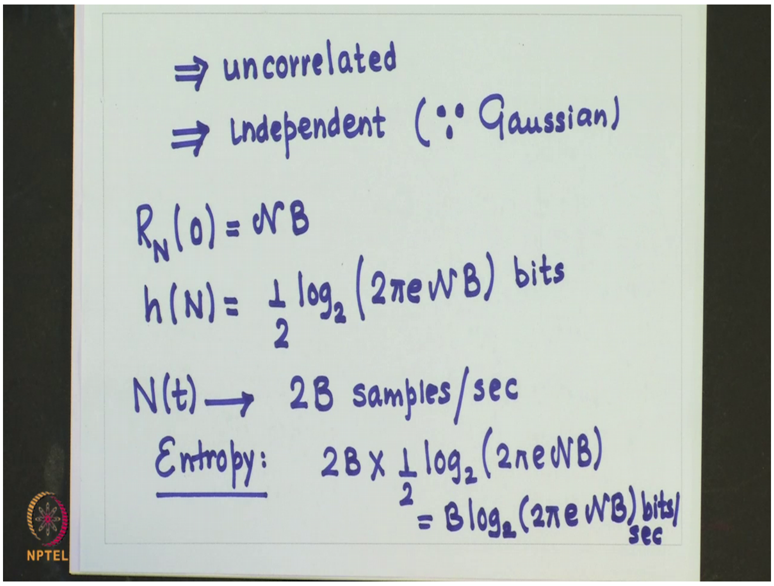


So, the power spectral density for this would look like this frequency and this is my power spectral density correct and I assume that the spectral density is given by italic N by 2 and the bandwidth is minus B to plus B hertz; so band limited to B hertz.

Now we know that power spectral density is the Fourier transform of the autocorrelation of the random process and because this is a flat out here correct. So, the autocorrelation will be the inverse Fourier transform of this power spectral density and that would be equal to a sinc function correct and this sinc function will have the following property that this autocorrelation will be 0 for all these values where k is equal to plus minus 1, plus minus 2 and all that ok.

Now, so this is the property we will get. Now we know what is an autocorrelation? Autocorrelation by definition is the expectation of N t and N t displaced k by 2 B correct this correct; so by this means that this is equal to 0 correct. So, now we are assuming let us assume that your noise process is also 0 mean without loss of generality.

(Refer Slide Time: 14:14)



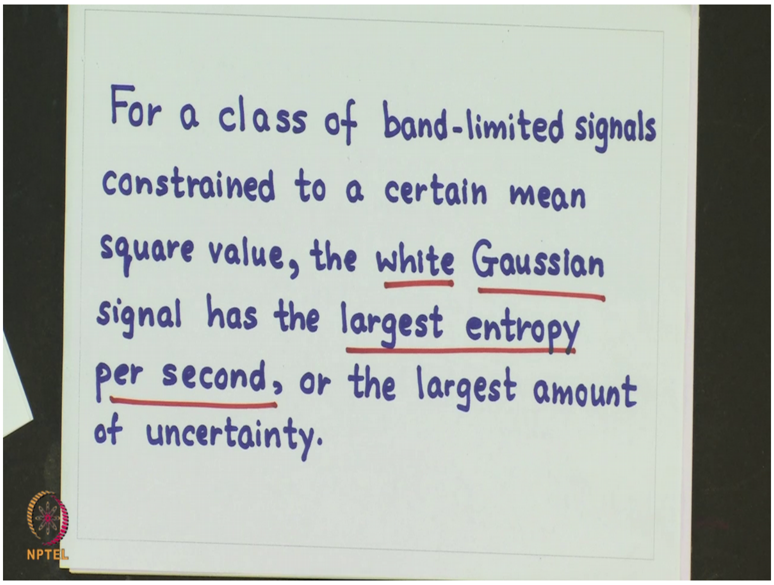
So, what it means that if this condition is satisfied it implies that your samples at the distance given by k by 2 B are uncorrelated because the process is 0 mean process and because it is a Gaussian correct; this also implies they are just independent samples, because it is Gaussian correct. Now from this expression for the autocorrelation, we can also find out the variance; of that random process which you have to evaluate at k equal to 0 and this will be equal to italic N multiplied by B. Now so, I have my samples Gaussian samples and they have this variance.

Now from the earlier result which we derived we know the entropy for that sample is going to be equal to half log to the base 2 2 pi e the variance is nothing, but this is equal to this value ok and now remember that your noise process N t is completely specified by the sampled values correct which are 2 B samples per second you sample and that is because its band limited to B hertz. So, by a Nyquist criterion I can sample it at 2 B samples per second. So, your noise process is completely specified by this. So, the entropy of this noise process is completely specified by this 2 B samples.

Now, each of these 2 B samples are independent correct from here. So, what it implies that the entropy of those 2 B samples is going to be the sum of the entropies of those samples correct. So, you will get the entropy of those 2 B samples equal to 2 B multiplied by this quantity which is half log to the base 2 2 pi e N correct. So, so many this basically remember it is a bits correct. So, I get basically so many samples per second. So, this becomes equal to B log to the base 2 2 pi e [noise] B bits per second because I am taking 2 B samples per second fine.

So, from this now what is the significant conclusion out of this? The significant conclusion is that; that among all signal bind limited 2 B hertz and constrained to have a certain mean square value say sigma squared, the white Gaussian band limited signal has the largest entropy per second correct.

(Refer Slide Time: 18:16)



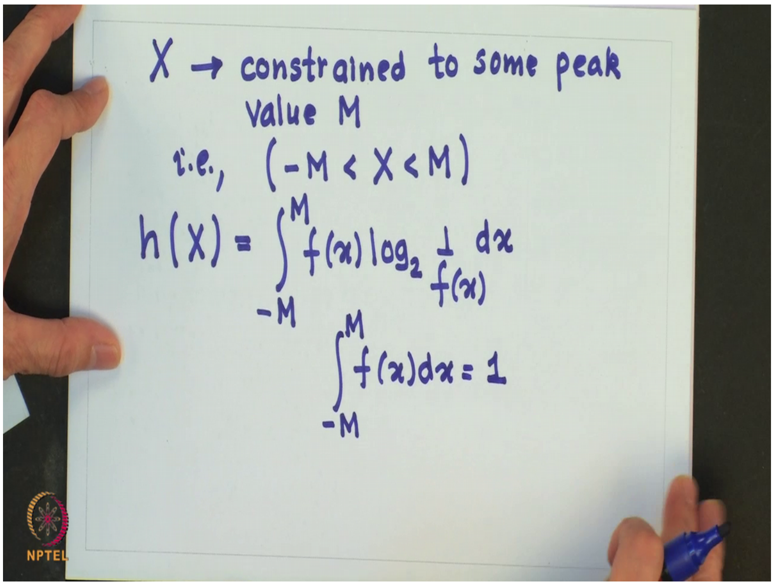
So, I can make this statement for a class of band limited signals constrained to a certain mean square value the white Gaussian signal has the largest entropy per second or the largest amount of uncertainty.

Now, what is the reason for this? Recall that for a given mean square value Gaussian samples has the largest entropy ok. Now moreover all the 2 2 B samples of a Gaussian band limited process are independent. So, this implies that the entropy per second is the sum of the entropies of all the 2 B samples. Now in processes that are not white what will happen that the Nyquist samples will become correlated and hence the entropy per second is going to be less than the sum of the entropies of 2 B samples.

Now, now we relax one more condition; so if it is not white the entropy will be less and now if the signal is not Gaussian then the samples are not Gaussian and hence the entropy per sample is also less than the maximum possible and entropy for a given mean square value ok; is it ok? So, this basically based on this argument we can make this statement that a band limited signal constrained to a certain mean square value, the white Gaussian signal this will make the independent because uncorrelated Gaussian correct has the largest entropy per second or the largest amount of uncertainty.

And this is also the reason why white Gaussian noise is the worst possible noise in terms of interference with signal transmission and that is why we use this model very often in communication ok. Now, let us take one more example of maximization of this differential entropy, but with a different constraint correct.

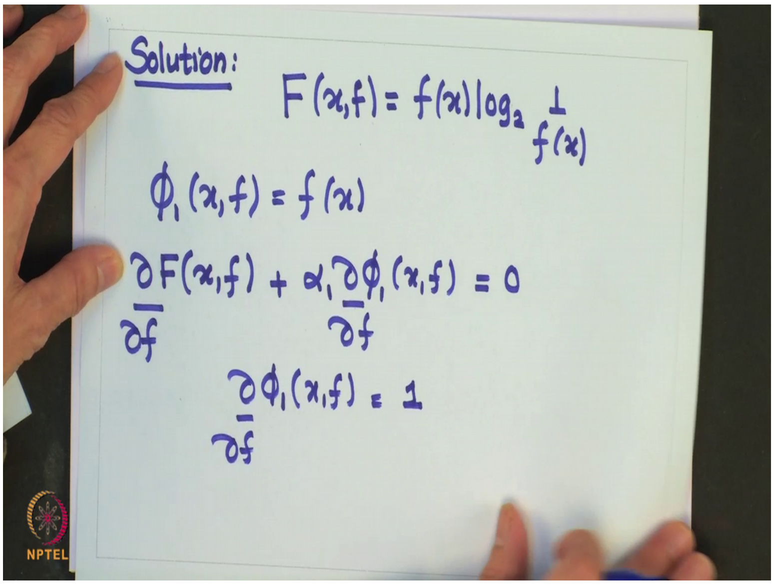
(Refer Slide Time: 20:46)



So, let me assume that I have a random variable X which is constrained to some peak value; let us call that peak value to be M; what this means that the random variable can take the value between minus M to plus M fine. And now our problem is basically I want to maximize this differential entropy and I want to find out the PDF which will achieve this.

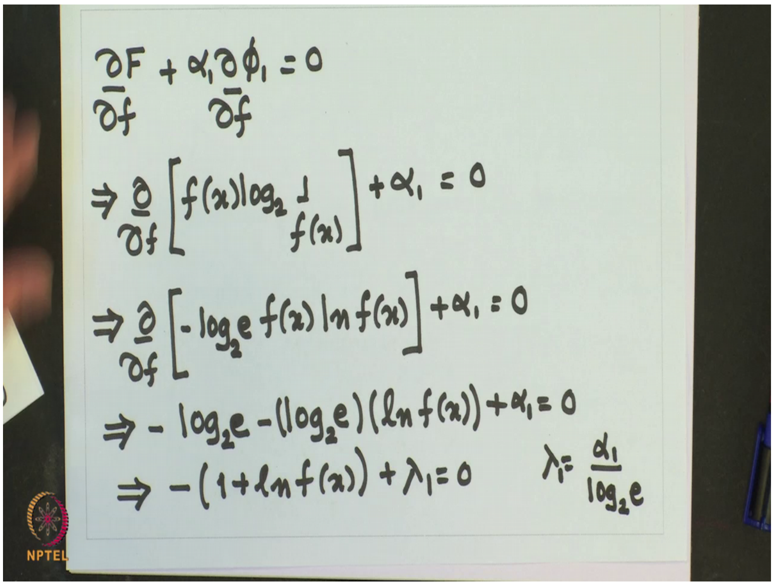
So, my job is to find out the PDF f x such that this quantity will be maximized; now please remember that the constraint of this random variable constrained to some peak value M has been incorporated in this limits of the integral ok and we have one more constraint on the PDF as usual and that would be this constraint this should be equal to 1 fine. So, let us try to evaluate the PDF for this.

(Refer Slide Time: 22:27)



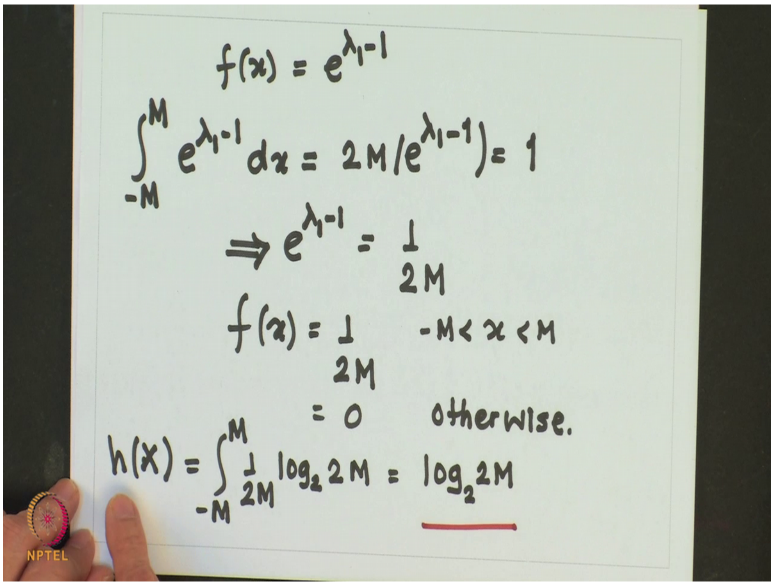
So, solution now again our F x f is same as earlier case which we had looked into; we have phi 1 x f is equal to f x. Now we know that the solution to this constraint optimization problem will be obtained by this theorem from calculus of variation this should be equal to 0.

(Refer Slide Time: 23:30)



Now again in our case this will turn out to be 1 fine and we have to evaluate this equal to 0 implies; first I converted it to a natural logarithm from, divide everything by log to the base 2 e and this you will get it as; where now lambda 1 is equal to alpha 1 by log to the base 2 e this is another constants fine. So, from here basically you can see this will be equal to now I get it my f x turns out to be e raised to lambda 1 minus 1 ok.

(Refer Slide Time: 25:18)



So, and now so this is the thing let us solve for this and we have our constraint on the PDF f x dx is equal to 1. So, if I substitute that dx is equal to 2 M e raised to lambda 1 minus 1 is equal to M this implies that e raise to lambda 1; quantity is equal to 1 by 2 M correct. So, my PDF is going to be 1 by 2 M which is a uniform distribution between for this range and is equal to 0 otherwise correct.

And for this let us evaluate the entropy differential entropy this is my PDF log to the base 2 1 by f x. So, that will be 2 M and this will be equal to log to the base 2 2 M. So, I get this result correct fine. So, if I change my constraint then this is the PDF which I get for that constraint that is the constraint to the peak value M and then this is the entropy which I will get that is a differential entropy fine.

So, now we have looked into the definition of average information for a continuous information source in terms of differential or relative entropy. Now our next concern is basically transmission of this continuous information on a communication channel correct. So, we will take up this study in the next class.

Thank you.