Modern Digital Communication Techniques
Prof. Suvra Sekhar Das
G. S. Sanyal School of Telecommunication
Indian Institute of Technology, Kharagpur

Lecture - 11
Source Coding (Contd.)

Welcome to the lectures on Modern Digital Communication Techniques. In the previous lecture we have discussed about the Kraft inequality, and in that particular discussion we said that given number of code word lengths we have to find out whether using those lengths one can make a prefix free code.

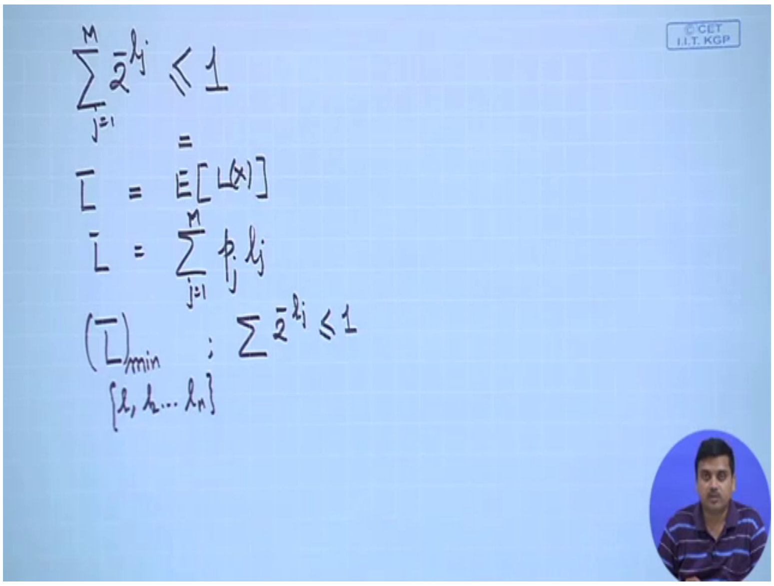
And before those that particular lecture we have discussed why do we need prefix free code and that that you should came up because we are talking about unique decodability and we took up unique decodability, because we are taken up variable length code and we are taken up variable length code because we saw that if one would take the advantage of the probabilistic nature of the source. Then instead of going for a fixed length coding if one goes for variable length coding one could save a number of bits per source symbol, and that would effectively reduce the number of bits that gets transmitted.

So; that means, we started with fixed length coding we moved on to variable length coding; after studying variable length coding we found we require the condition of unique decodability, when we studied unique decodability we defined it we found that we would require prefix free code so that we could instantaneously decode the received bits. Once we studied prefix free codes then we said forth to study that given a set of lengths whether these lengths would satisfy a certain condition known as the Kraft inequality so that we can be assured that whether a prefix free code can be constructed out of these lengths.

Once you assure that one can construct a prefix free code, then one could go ahead and build the binary tree as we have discussed and take each of the leaf nodes as the code words and starting from the root going through the branches through the nodes on to the leaf, one would select the string the binary string as the code word. We also said that any intermediary node would serve as prefix and since intermediary nodes are not code words all the code words produced out of this method are prefix free and then we said that well this conditions are fine, and we further defined something known as a full tree and not a full tree by full tree what we meant is that the tree cannot be short end without destroying the prefix free property and no code word can be added without destroying the prefix free property.

So, with this we moved on further and try to find out that is there number known as the minimum number of bits per source symbol and we denoted it as the L bar min, where L bar was the average number of bits required per source symbol. This average number of bits required per source symbol is used, because we are taking variable length coding. So, once you have variable length coding each code word would be of different length or could be same length, but overall there would be various code word lengths. Of course, one can argue that one particular case can be constructed out of this code tree where all the code word lengths are same and what will find is that there is a special case and the results would be obvious very soon.

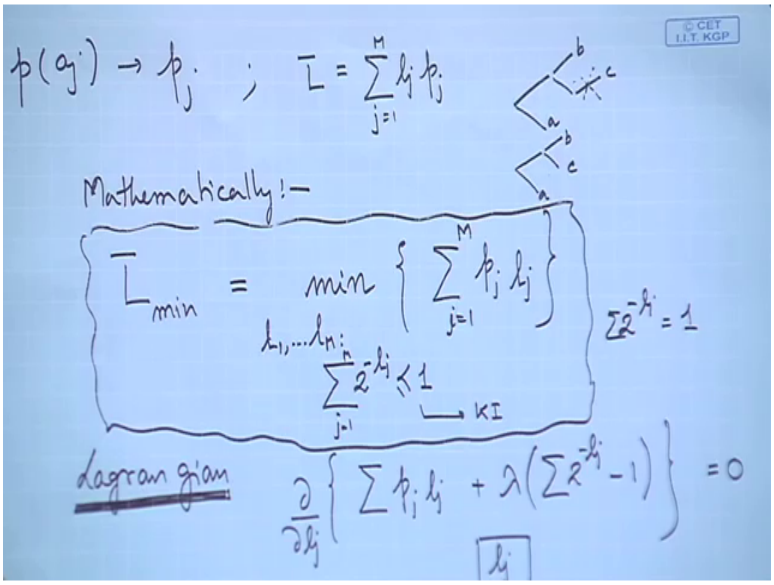
(Refer Slide Time: 03:43)



So, when we were discussing these things we looked at the Kraft inequality and the condition that we had was sum of 2 to the power of minus l j j equals 1 to m, m is the number of symbols in the alphabet should be less than or equal to 1. So, with this condition we found that these lengths which satisfy this condition could per could create a prefix free code and when this holds with equality; that means, if this is equal to 1 in that case we found that a full tree can be constructed. Then we moved ahead and we wanted to find L bar, L bar is basically the expected or expectation of L of x and what we were where we reached is that L bar could be written as sum over p j l j where p j is the probability of the j-th symbol and l j is the length of the j-th code word.

So, you wanted to find the L bar min as a minimum value of L bar a subject to Kraft inequality condition. So, we wanted to find all the set of code word lengths that would minimize the expected length subject to this Kraft inequality. Now why we have Kraft inequality the reason is we wanted to find the minimum length over all prefix free codes.

(Refer Slide Time: 05:19)

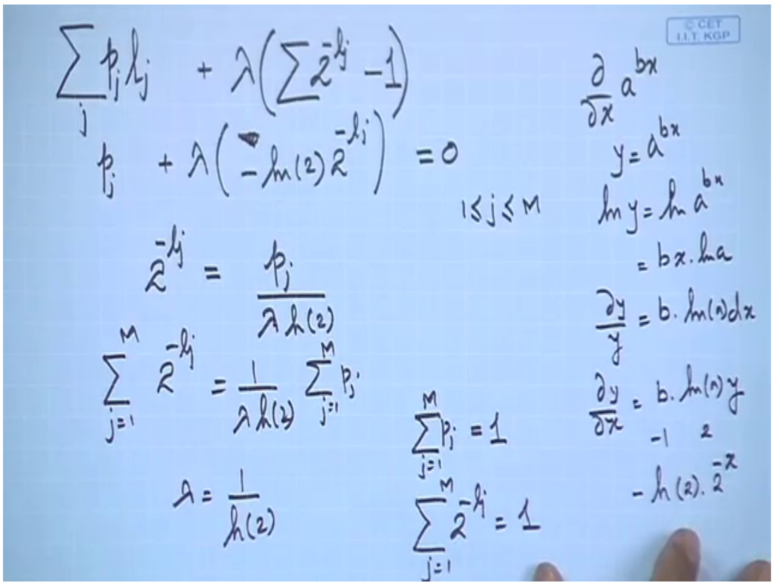


Now prefix free codes the reasons we have already explained and we set up the mathematical problem in the lecture as this that L bar min is equal to the minimum of this summation which we have just defined overall possible lengths subject to Kraft inequality as we have written over here, and while proceeding with this we said that this minimization is achieved when we rather have this condition being equal to 1, because if this is not equal to 1. That means one of the code word lengths could be short end example was taken over here that this particular code word length could be short end. So, if this could be short end then it becomes a full tree in case of full tree this condition is valid and therefore, our condition could be reconstructed instead of this using this.

And this particular problem as we see we are talking about lengths in terms of number of bits and obviously, this termer is an integer number. So, if it is an integer problem it is not possible to solve it using standard methods you have to search over and find the numbers. So, one easier way could be relax to relax this particular constraint and say that let us have real lengths, and once we find real lengths then from that we can go to the nearest integer as our solution or those special cases where this lengths would be integer that would be that will be the solution for the minimum. And one could also say that for real lengths one should have this condition; that means, the lengths that we find along with the probabilities should be less than or equal to the case than when these lengths are integer so that means, this particular solution would give a lower bound on the integer problem.

So, proceeding with this with that particular relaxation, we said that one could proceed with the Lagrangian method and well established method. So, there we take the cost function which is the cost function over here j equals 1 to m and we add this Lagrangian multiplier to this constraints and we have this Lagrangian expression we take the derivative with respect to l j, because l j is what we are trying to find set of l j s which would minimize this, this minimized over the choice of l j s set this equal to 0 this is what we are discussed in the last lecture. So, we move on from this particular point and we see how do we proceed.

(Refer Slide Time: 07:59)



So, what we have with us is sum over p j l j plus lambda times sum over 2 to the power of minus l j, since we said that minimization happens when there is equality constraint otherwise it is the case where it is not be minimum because it is not a full tree, see if it is a full tree it has to be the equality constraint. So, we use the equality constraint and we have this expression.

So, when you take the derivative with respect to l j equal to get p j we taking for every j, j equals 1 to m plus lambda and here we have to take the derivative of this. So, just a reminder that what we need is basically the derivative of a to the power of b x is what we have. So, although we can use the formula straight away which is well known it is still like to revise a little bit. So, what we can do is we can say that y equals to a 2 the power of b x and we want to take the derivative of y with respect to x that is what we want to find. So, one simple way would be to take the l n of this, and that would lead to l n of a to the power of b x which would be b x times l n a, and if you take the derivative on both the sides you can get d y upon y because this is 1 by y and this will be b times l n a and d of x.

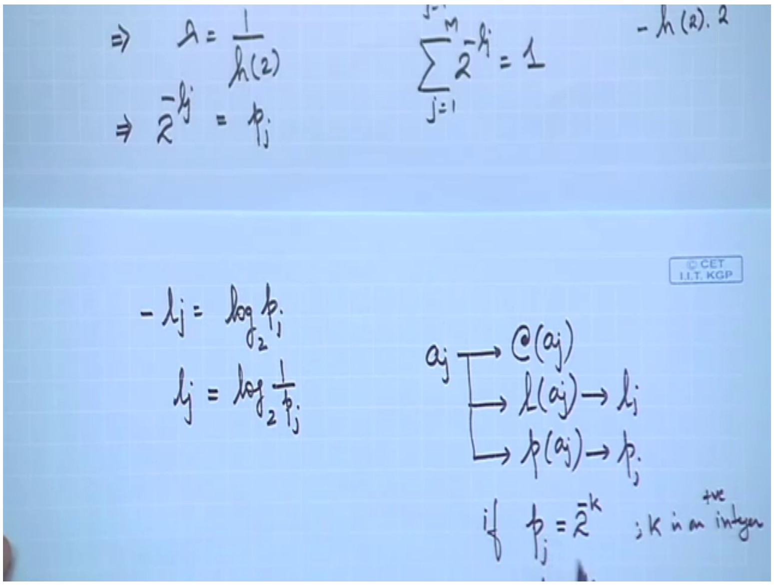
So, what we want is d of y by d of x, and this would be b times l n a times y. So, b in our case is minus 1 and a in our case is 2. So, what we have is minus l n of 2 times 2 to the power of minus x right. So, when we translate that over here what we get is a minus l n of 2 times 2 to the power of minus l j. Look at this this is the same and of course, this term is 0 and we said this equal to 0 as per our previous step take this derivative and set it equals to 0. So, if taken the derivative and now we have set this equal to 0 and of course, this is true for all j in the range of 1 to m.

So, this would give you that 2 to the power of minus l j is equal to p j we take this on the left or we take p j on the left take away the minus sign p j upon lambda l n of 2, because of this minus it is a plus on both the sides. Now since we know that or simply we can take this summation of j equals to 1 to m on both the sides on left hand side and right hand side. So, 2 to the power of minus l j is equal to 1 upon lambda l n 2 this is not a function of j sum over p j. So, this is what we have.

Now, sum over p j since we have taken the full tree p j equals to 1 to m p j should be equal to 1. So, what it means is simply that we are not taking any probabilities which is non-zero sorry which is not 0. So, we are only taking probabilities which are not zero; that means, lying between 0 and 1 and if there is any symbol which does not occur it start in the probability space, it is not part of the source symbols because it is not occurring. So, for the for it to the probability space this summation must be equal to 1 and also what we have is by the Kraft inequality or the Kraft equality in this case 2 to the power of minus l j is also equal to 1. So, these 2 conditions are there with us.

Now, since these 2 conditions there with us, we use it to get a lambda is equal to 1 upon l n of 2; this summation becomes 1 this summation becomes 1. So, this is the result that we get out of this exercise. So, once we have this result we can use it further and we will simply bring this back into this equation. So, when we bring it back into this equation we get lambda equals to 1 upon l n 2 or this cancels out. So, if you look at this lambda equals to 1 upon l n 2 when you replace it here this cancels out.

(Refer Slide Time: 13:17)



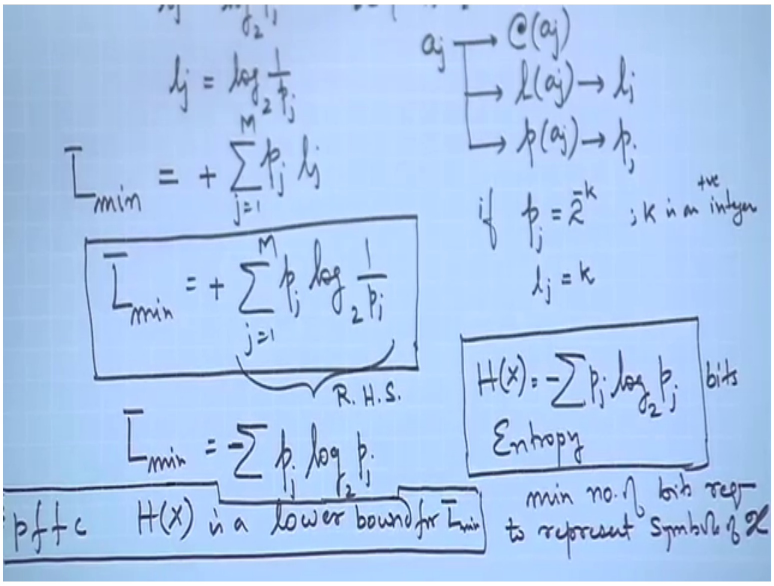
So, we can write that 2 to the power of minus l j is equal to p j right. So, this is what we get or we can also state that l j or you can take It as minus l j is equal to log base 2 of p j or we simply write log because whenever we write log in this particular situation we will be using log base 2 or you could also write l j is equal to log base 2 of 1 upon p j.

Now, just to clarify few things since p j is less than or equal to 1. So, this whole number is less than 0 or equal to 0. So, this is less than or equal to 0 so that means, l j in minus l j is equal to number which is negative number therefor l j is a positive number. So, at least it clarifies that this is the positive number and here also we can see that since it is p j is lying between 0 and 1. So, this overall number is a positive number now this number is greater than 1. So, it is positive number.

So, l j is a positive number. So, this this number that we have indicates the length. So, it gives us the relationship between the length of the j-th symbol and the probability of the j-th symbol. So, all it tells is that for the symbol a j we have the code which is c of a j, and it also has a length which is l of a j and it has a probability p of a j for simplicity we have written it as l j and this is p j. So, basically it connects the probability of occurrence of the symbol to the length of code word that one should use in order to minimize the l bar.

So, if we have p j if p j is equal to 2 to the power of some minus k, in that case what we have is where k is an integer is a positive integer. So, is a positive integer. So, l j is equal to k in that case right. So, what it means is that when p j is some 2 to the power of something in that case this holds with equality and we have the minimization in other case it is lower bound.

(Refer Slide Time: 15:52)

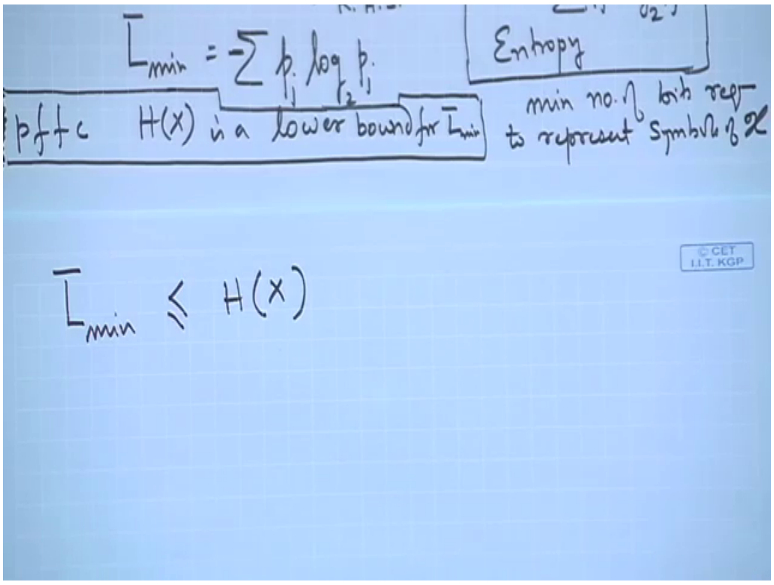


So, what we have now we could write that L bar min is equal to minus summation j equals 1 to m p j times l j this we already had. So, what you have is the minimization solution gives us p j is an inherent property of the source, and l j is what we have found. So, we have write it as log base 2 1 by p j right.

At this point it is important to note that this quantity on the right hand side this right hand side quantity which is the summation of p j times log of sorry this is if i write log base 2 one by p j it is a plus this this is plus because I have p j times l j. So, this is a plus otherwise if i would bring it on the numerator for l min bar is equal to summation of p j times log base 2 p j in this case I have minus which is the same, p j with the minus this is denoted as the entropy of the source. This is by definition that this is the entropy of the source alphabet x denoted by x, where this symbols which contains this symbols x is equal to a 1 a 2 to a m these are the symbols with probabilities p j. So, in that is case this denotes the entropy and where the definition of entropy tells that it is bits because we have used log base 2 or the minimum number of bits required to represent symbols of the alphabet x. So, this is what we have.

So that means, L bar min with the real with real results; that means, not integer lengths is equal to the entropy so that means, for prefix free codes. So, if we are having a prefix free codes then H of x which is the entropy of the source going by this definition is a lower bound for L bar min, this is a very important result is a very very important result that we have with us and we can note that when this particular probabilities are integer lengths of 2 integer powers of 2, in that case we can find these lengths which are exact solutions in that case this L bar min will attain the entropy. In other cases when it is not so, we can say that L bar min in other cases we can say that L bar min is less than or equal to H of x.

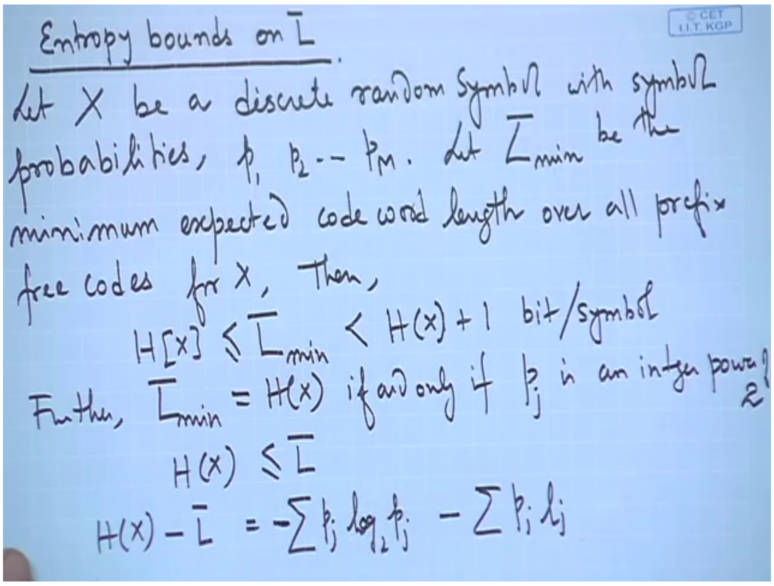
(Refer Slide Time: 19:41)



So, only if it is a integer power of 2 it is equal to the entropy, in other cases it is lower bounded by the entropy. So, that is a very very important result that if you can calculate the entropy that will tell you that what is the minimum number of bits that will be required to represent this particular source, and if by providence or if by chance you have the source symbols such that the sorry such that the probabilities of this particular source are integer powers of 2 which is not always guaranteed, you can find lengths which would actually meat the entropy of this particular source.

So with this we would like to move forward and we would like to go further and discuss about what are the entropy bounds on this L bar min; that means, we have found one minimum length of it. So, that is the minimum the lower bound of which is the entropy and which will also show that yes indeed it is the minimum, we will also show what is the upper bound of this minimum L bar.

(Refer Slide Time: 21:18)

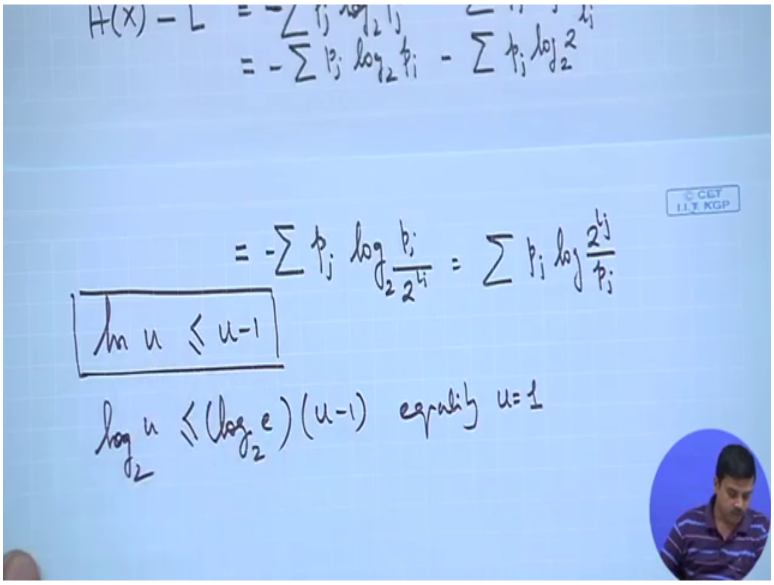


So, what we are going to look at is the. So, for this we are not going to create anything any new set up we will continue with the existing set up that we have. So, for this we say that let a X be a discreet random symbol with symbols probabilities p 1, p 2 up to p m in other words it tells us that we are taking the same alphabet with m m symbols, and each of the symbols are having the probabilities of p 1, p 2 up to p m. So, this all that we have indicated here and we say that let L bar min be the minimum expected code word length over all prefix free codes for x then what we can say is that H of x is less than or equal to L bar min and it is less than h x plus one bit per symbol.

So, what it tells us is that the entropy bound on L bar tells us that it is lower bounded by the entropy and it is upper bounded at most one bit from entropy, that it is it is not taking more than one bit from entropy now this is the very important result which relates this the minimum length of the code word over all possible prefix free code word lengths to entropy. So, it also tells we can also add that further that L bar min we have already discussed this is equal to H of x we have already explained this if and only if p of j is an integer power of 2. This we have already discussed in the previous note that we were writing that and we have also seen that when because l j is log base 2 of 1 by p j. So, if p j is integer power of 2 because we are doing log base 2 we are going to get this integer number in the length ok.

So, first we can first we are going to take this part so that means, we first will show that H of x is less than or equal to L bar now since L bar min is the minimum value of l bar. So, if L bar or if h x is less than or equal to L bar then of course, L bar min is the one which is or H x is less than or equal to L bar min as well. So, what we have is that we will take H of x minus of l bar. So, if we write this what we have is H of x is denoted as p j log base 2 of p j with the minus sign and if you take L bar, L bar is minus p j times l j now look at this h x is defined with respect to this and when I said L bar what we mean is that we have this lengths l j and we have this probabilities p j. So, if we have this lengths l j and we have this probabilities p j, then this quantity is the L bar and this quantity is the is the H x. So, moving on further what we have is this should be equal to minus summation of p j log base 2 of p j.

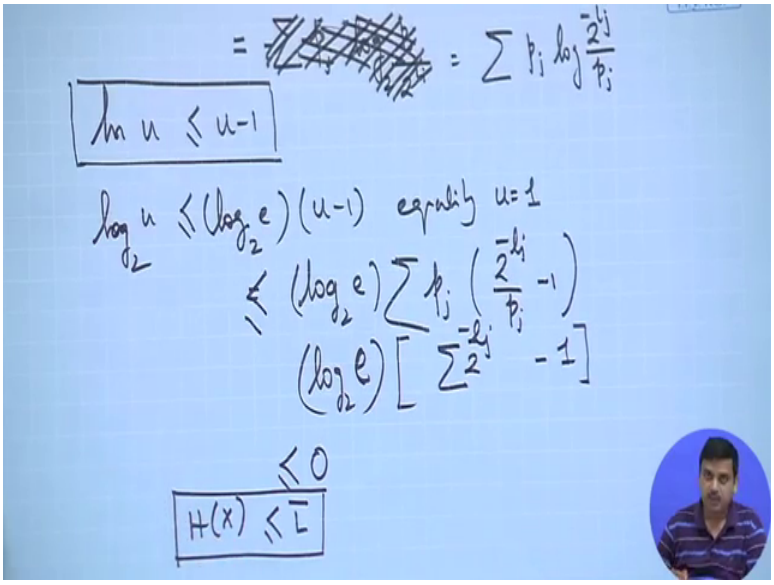
(Refer Slide Time: 26:03)



So, we have not done anything this quantity, only this quantity we change it slightly that p j log base 2, 2 to the power of l j. So, effectively this is l j. So, we have not made any changes there. So, we continue further and what we have is sum of p j that is what we have common both of them is a negative sign log of course, p j upon 2 to the power of l j right it is log base 2 or you can reverse it bring the negative sign inside. So, what you have is p j log of 2 to the power of l j upon p j.

At this point one has to use important result which says that l n of u is less than or equal to u minus 1 now this can be graphically understood also at this point we will not go into that we will rather remain over here l n is equal to this. So, you could write l n of u is less than or equal to log base 2 of e times log sorry log base 2 of u is less than or equal to log e base 2 q minus 1 and this equality holds only at the value of u equals to one because equality for u equals to one. So, when u is equal to 1, you can clearly see the right hand side is 0 and definite the left and side is also equal to 0.

(Refer Slide Time: 28:06)



So, if we use this what we get is this this particular side what we have is this whole thing is less than or equal to because we have a log base 2. So, we are using that. So, log base 2 of e and summation we have p j as p j and this one is what we are going to replace 2 to the power of l j upon p j minus 1.

So, now when we take p j inside the bracket what we have is log base 2 of e multiplied by sum of 2 to the power of l j, let me see if we have done it right what we have is we have made a slight mistake at this point, we could have actually taken this as plus and we could have taken this as minus and then our expression because this is a minus. So, we have a plus we have minus over here. So, this is this equates out and what we have over here is this is not our interest. So, what we have is this one p j log base 2 to the power of minus l j by p j.

So, we have this p j p j cancels out, minus summation of p j is basically 1 and this quantity is less than or equal to 1. So, what we have in other words that this quantity is on the right hand side is less than or equal to 0 or in other words what we started of with h x at this point we can write H x is less than or equal to l bar. So, this is the first part of the proof that we have established. And in the next lecture we will establish the next part of the inequality which is in the upper bound on L bar.

Thank you.