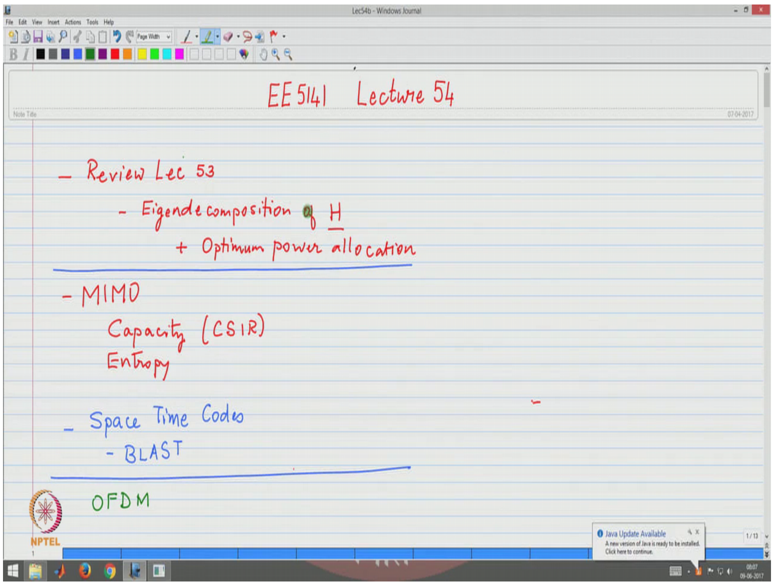
Introduction to Wireless and Cellular Communication
Prof. David Koilpillai
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 54

Good morning and welcome to lecture 54 of our course, in the last lecture we had looked at the subject of MIMO systems and within that a special case where we had done the Eigen decomposition of the MIMO channel and then we did the optimum power allocation for this channel.

(Refer Slide Time: 00:25)

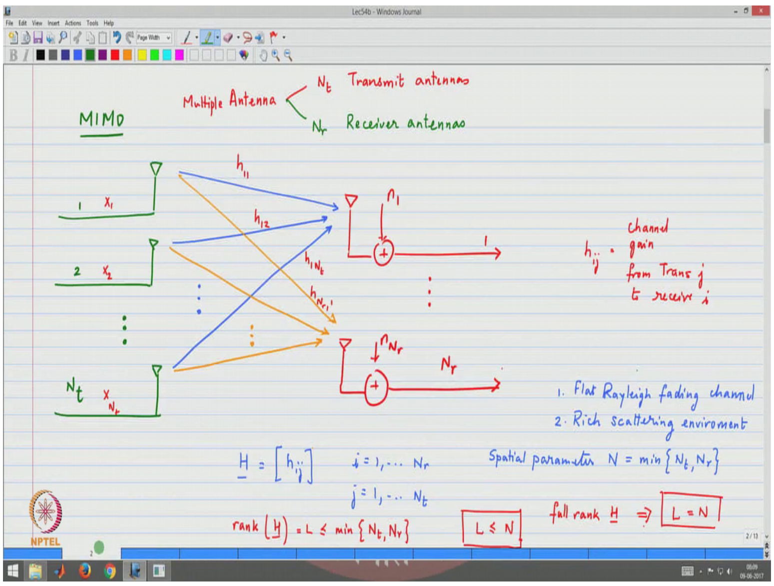


So, what we would like to do in todays lecture is quickly review that material and then build on that foundation. The Eigen decomposition is possible when we have information about the channel at the transmitter. So, therefore, we are able to do the pre coding at the transmitter and the post processing at the receiver.

However if that knowledge is not available at the transmitter, that is the case where you have CSIR and that is what we would like to do in today's lecture. We will also be in unfolding the concept of entropy in our discussion as we have already introduced it in the last lecture; this is the same concept that it would be that is borrowed from the chapter on information theory so that we can understand the capacity of the channel better. We will touch upon the topic of space time codes, when you have multiple antennas at the transmitter and receiver then you have the ability to apply a very special type of coding that method is called the space time codes, and what we would like to do in this course is to introduce the space time codes in a very brief manner, and in particular one of the techniques called the blast technique which was the first demonstration of very large capacities in a MIMO channel.

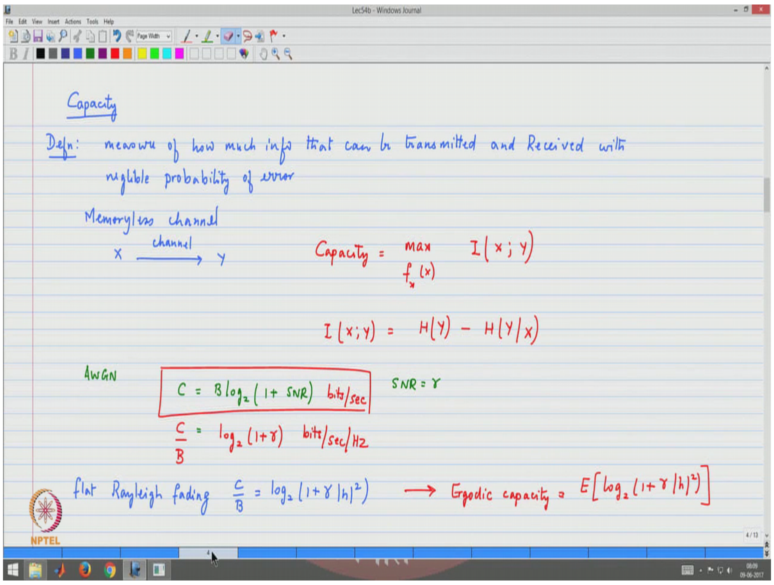
The next chapter that we will be touching upon not in todays lecture, but something for you to read up ahead if you if as you are interested would be on OFDM orthogonal frequency division multiplexing. So, again that is becoming bringing up the last, but one unit of the course and that would be a very useful and interesting study because that is covered in most of the 4G systems.

(Refer Slide Time: 02:17)



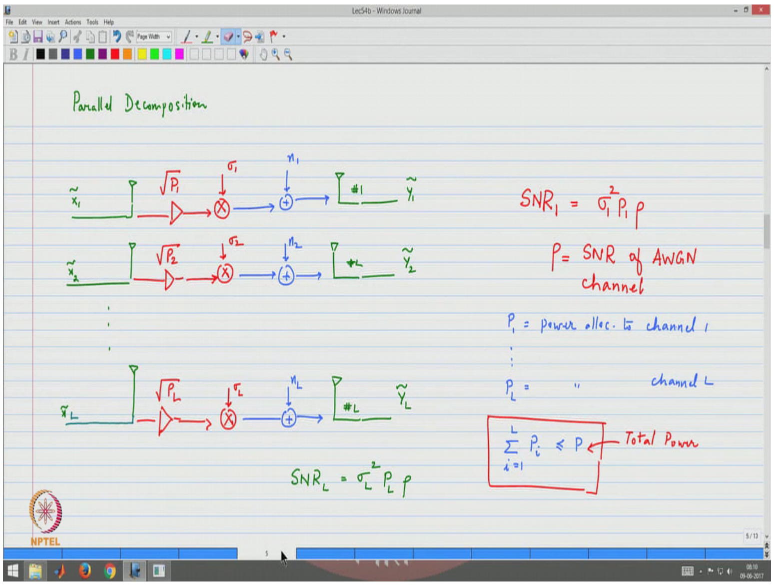
So, a quick review of what we have discussed in the last lecture. The MIMO framework again just for reminder we have N t transmitted antennas N r transmit receive antennas and what we would like to do is characterize this channel by a matrix and then we look at the capacity of this channel when the information is known at the transmitter.

(Refer Slide Time: 02:39)



So, what we are doing with the MIMO channel? When we have knowledge at the transmitter is to do the singular value decomposition, and the singular value decomposition helps us to singular value decomposition helps us to apply the channel in a in a parallel form and obtain L parallel channels.

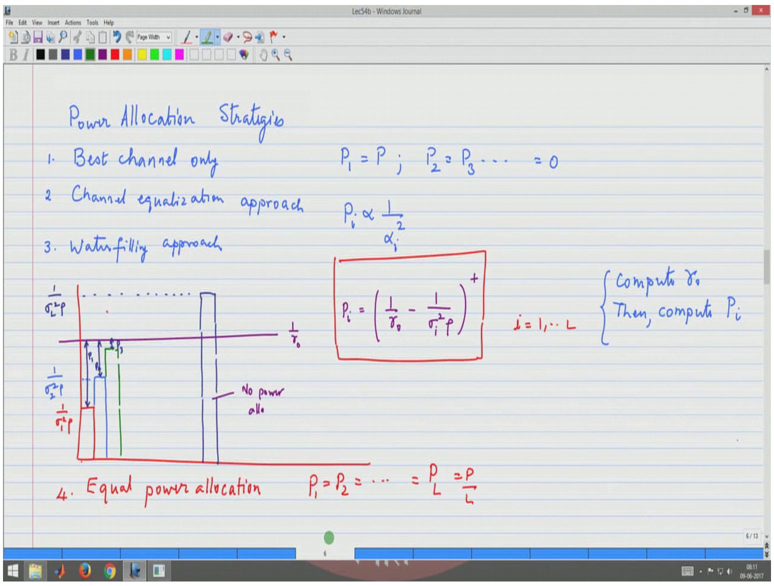
(Refer Slide Time: 03:04)



Now, each of these an L corresponds to the rank of h, and corresponds to the non singular non zero singular values that are present in the singular value decomposition. Now each of these can be treated as parallel channels independent of each other, and in the last class we also saw the benefits of introducing the additional power allocation to each of these channels, now given that there is a total power constraint the total power constraint is that the power transmitted in all the L channels together cannot exceed a total value uppercase P, and within each channel depending upon how much power has been allocated we then derive an expression for the SNR.

Now the reason for the and the importance of the SNR is that this will help us apply the capacity the estimate the capacity because the capacity is related to logarithm base 2 of 1 plus SNR. So, again that is the reason why we are focused on the SNR.

(Refer Slide Time: 04:11)

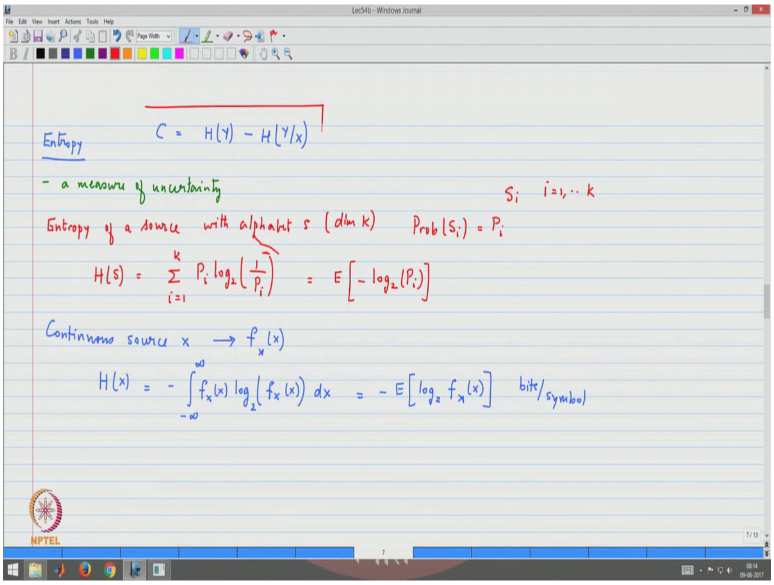


Now, the power allocation strategies we have seen four of them, the first one was give power only to the best channel. So, you give all your if your singular values are written in decreasing order, then the channel number one is the one that has got the best SNR and you give the total power to channel number one. Now the channel equalization approach, where we if we saw that if you apply the power allocation such that it cancels out the effect of the singular value; basically what we would say is the power allocated is in proportional to 1 by sigma i square. So, if you remember the SNR is sigma i square times P i, now if Pi was 1 by sigma i square then those 2 terms will cancel giving you rho as the SNR of each channel. So, more or less what the channel equalization approaches does is to allocate enough power. So, that all the channels appear to have equal SNR.

Now, again whether this is a good strategy we will have to see the best option that we have is for us to do water filling as we have done in the case of the single antenna system we find that the different parallel channels, have got different SNRs and the water filling approach depends on us computing gamma naught, which is the water filling level and then computing the power to be allocated to each of these channels. Now if any of the channels has an SNR that falls below a threshold what we are plotting here is 1 by sigma, sigma 2 square rho. So, it is like the reciprocal of the SNR. So, which means that it is the reciprocal of the SNR exceeds the water level no power allocation is done.

So, for example, this blue channel does not have any power allocation done to it; however, the other three channels do have power allocation as per this figure, and the best channel of course, gets the most power allocation. And the fourth one which we mentioned in the case of an asymptotically, where all the L channels got good SNR was to have equal power allocation each of these channels P 1 to p L have got a 1 P divided by L equal power allocation. Again 4 approaches each of them giving us different insights, but most importantly the best approach be the water filling approach, which says that this is what will help us get the maximum performance out of the system.

(Refer Slide Time: 07:02)

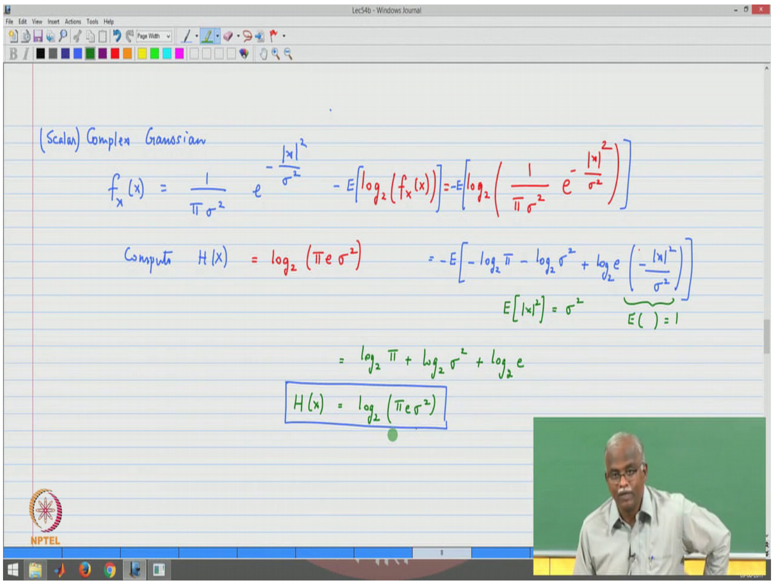


Now, we would like to focus on the case where we do not have the information at the transmitter, the channel information of the transmitter. So, which means that I cannot do parallel decomposition.

So, this is the case where we would have to derive the channel capacity based on the entropy of the channel. The definition of entropy that we have introduced in the last class is channel capacity is entropy of course, being a measure of uncertainty, the channel capacity given by the entropy of y minus the entropy of y given x. So, that is the definition that we have and what we would like to do is build on that. Now the if it was a source with a finite alphabet, then each of those alphabets were having a probability Pi, then the entropy of the source would be summation Pi logarithm base2 of 1 over Pi, which can be expressed as expected value of minus logarithm base 2 of Pi of. So, that is the for a discrete source. For continuous source the probabilities are expressed to the p d f, f subscript x of x. So, the entropy would be minus integral of minus infinity to infinity f x of x logarithm base 2 f x of x, d x and this can also be written very compactly as expected value of the logarithm base 2 of the p d f and that would be the number of bits per symbol.

So, what we would like to do is pick up from here, and then develop the concepts and in terms of understanding the MIMO channel and all of the elegant results that we can obtain with that. So, as a first step I had requested you to look at the scalar complex Gaussian with a p d f, f x of x equal to one over pi sigma square e power minus mod x squared by sigma square.

(Refer Slide Time: 09:08)

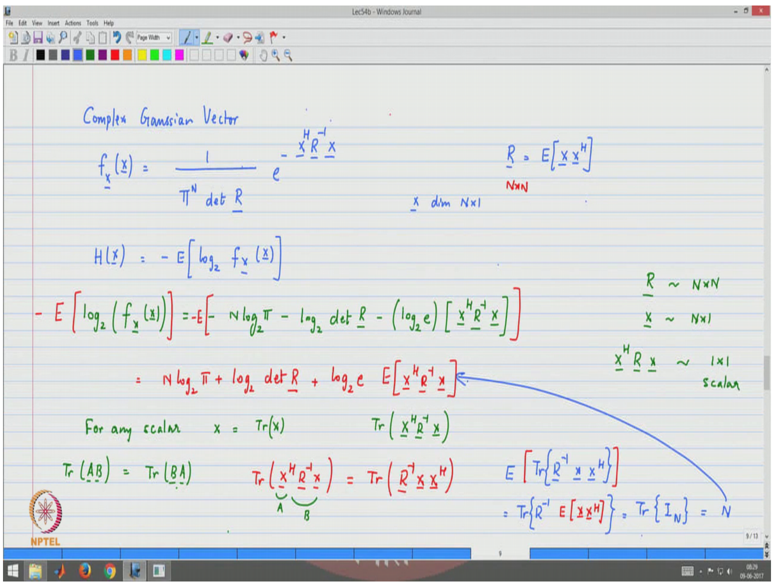


So, basically we are interested in computing logarithm base 2 of f x of x. So, logarithm base 2 let us do the quick calculation, logarithm base 2 of f x, of x that would be logarithm base 2 of 1 by pi sigma squared e power minus mod x the whole squared by sigma squared logarithm, and then expected value. So, after we take the logarithm we would then be interested in taking because. So, because basically what we are looking for is minus expected value of the logarithm base 2. So, this would be minus expected value of this expression ok.

So, if you were to calculate this will do the logarithm for as a first step and then the expected value at the next step, the first step gives me minus logarithm base 2 of pi is the denominator, minus logarithm base 2 of sigma squared and then the final term is logarithm base 2 of ah e power minus of an exponent. So, what we would like to do is write it in terms of the logarithm base 2 I split it as logarithm base 2 and logarithm base e. So, using the result that we had given in the last lecture, this would be plus logarithm e base 2 times the natural logarithm base e of e power minus x square by sigma squared, and so that would give us the following result minus mod x squared divided by sigma square. Now take the expected value inside and make note of the fact that expected value of mod x squared is equal to sigma squared.

So, therefore, the when you take the expected value this term within this bracket the expected value of this term becomes equal to 1, and when we do the simplification what we find is what we have is logarithm base 2 of pi plus logarithm base 2 of sigma squared plus logarithm base 2 of e because the next term is actually equal to 1. So, this of course, being using the property of logarithm, we can write it down as logarithm base 2 pi e sigma square. So, this is the expression for the entropy of scalar complex Gaussian source. So, this is H of x, and this is a very useful result because this will also tell us how we can expand this to the case where we are dealing with not scalar signals, but with vector signals because in the case of MIMO the transmitted signal received signal are all vectors. So, that is an important extension that we would like to do. So, basically if I have a scalar Gaussian complex Gaussian source I know the probability distribution, I know that the entropy is given as the minus expected value of logarithm base 2 of f x of x, now the we took the expression substituted and simplified and showed that the entropy of the source is equal to log of them base 2 pi e sigma squared.

(Refer Slide Time: 13:45)



Now, follow along with me for the next step which would be the case where we have a complex Gaussian vector source. So, first step would be for us to write down the p d f of a complex Gaussian vector, and then take the logarithm to the base 2 and then compute the expression. Again these are a set of simple steps, but I would like you to follow along and we will be using some of the results from linear algebra in our derivation. So, the first step would be to write down the p d f I am writing x with the under bar to show that this is a vector. So, the p d f, f x bar of f x of x is given by this is a n dimensional vector. So, you will see the value n appearing in the expression. So, this is equal to pi raised to the power N determinant of matrix R, where R is defined as the mac is defined as expected value of x x Hermitian. So, that is the correlation matrix. So, R is 1 by pi over N determinant of e power minus x Hermitian R inverse times x.

So, this is the expression that we have and of course, you can definitely plug in the value when it is a scalar and you can verify that when you substitute the value n equal to 1,we will get back the expression that we had in the previous case. So, again that is a verification, but this is a n dimensional vector where x has dimension N cross 1 it is a N dimensional vector. So, this is the expression that we have and H of x, H of x can be now obtained in the following way, notice that this is now a scalar this is a vector and this would be equal to minus expected value of logarithm base 2 of the vector p d f. Identically to the case when for the scalar case, which is extending it to the n dimensional case.

So, first step would be to calculate the logarithm base 2 of this expression. So, let us do that. So, logarithm base 2 of f x of x can be obtained as the first term minus n log of pi, minus n log base we are doing log base 2o. So, n log base 2 of pi plus or minus logarithm base 2 of the determinant of R plus logarithm base 2 of e times the natural logarithm of this expression. So, that would give us x Hermitian R inverse x. So, this is the expression that we have this is logarithm base 2 is one term and then what we have. Now what we are required of there is a minus sign there is a minus sign.

So, now we want to do the expected value with a minus sign. So, that will be minus expected value of this quantity. So, what we will get is N log base 2 of pi plus log base 2 of the determinant of R plus log base 2 to of e. Now we have to take expected value of x

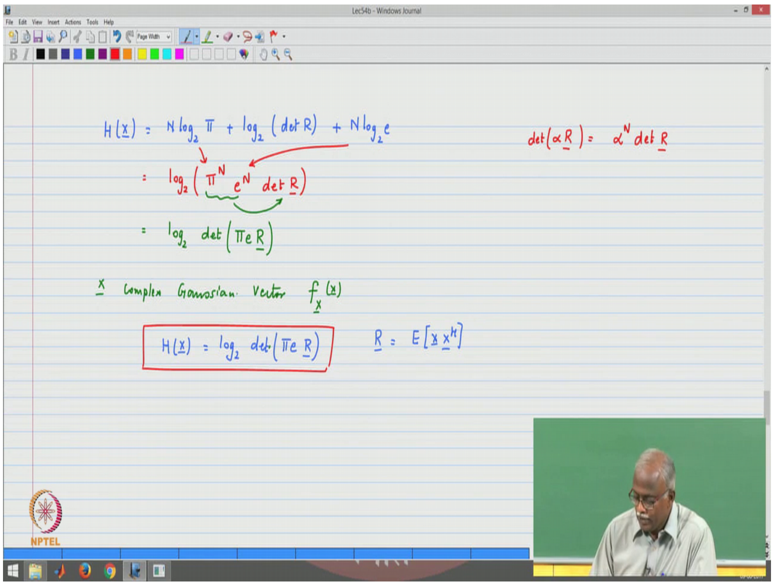
Hermitian R inverse x. Now in order for us to simplify this expression, I would now like to make a following observation. now R is a N by N matrix the dimensionality of R is N cross N. Now if you now then look at the term that we have if R is an N cross N matrix the matrix R has got the dimensions of N cross N, and x has got the dimension of N cross 1, now if you then look at x Hermitian R times x this will have dimension 1 cross 1 or in other words it is a scalar, x Hermitian R x is actually a scalar. So, we are take taking the expected value of a scalar quantity, though the derivation is for a vector quantity what is in the within the brackets is a scalar quantity.

Now, for any scalar we can write down that x is equal to the trace of x sort of trivially because trace is the addition of the diagonal elements since x is a scalar there is only one element which is also the diagonal element. So, we can write this. So, once we can write this as a the trace if you were to look at the trace of X Hermitian R inverse x, gives us a very very powerful result because trace has got in the following property. So, whenever you have trace of a product of 2 matrices AB, assuming that the interchange the order permits us to interchange the order. So, this also is equal to the trace of BA this is a result that we have from linear algebra. So, once we have this result then we use it in the following manner, trace of x hermitian R inverse x.

If I were to treat the this as my matrix A and this as matrix B then what we have is the following, that I can interchange A and B and write this as trace of R inverse times x times x Hermitian. I am using the property that trace of AB is equal to the trace of BA. So, once we write this write it in this form then we are able to look at the expected value. Expected value of the trace of R inverse x x Hermitian, you can take the expected value inside. If you take the expected value inside what we are left with is the trace of our inverse R is already a constant matrix. So, expected the expectation does not affect it then what we have. So, basically the bracket should be here expected value of the trace. So, this is the there is a trace and then the expected value, and taking the expectation inside. So, if what we are left with is a trace of R inverse times expected value of x x Hermitian.

Now, notice that R is defined as expected value of x x Hermitian. So, when you do R inverse times R, we will get the I will get the identity matrix. So, this will be the trace of the identity matrix of dimension n. So, the trace is the addition of the diagonal elements. So, therefore, this is actually equal to N. So, if you were to now use this result in this expression we can now write the following.

(Refer Slide Time: 23:00)



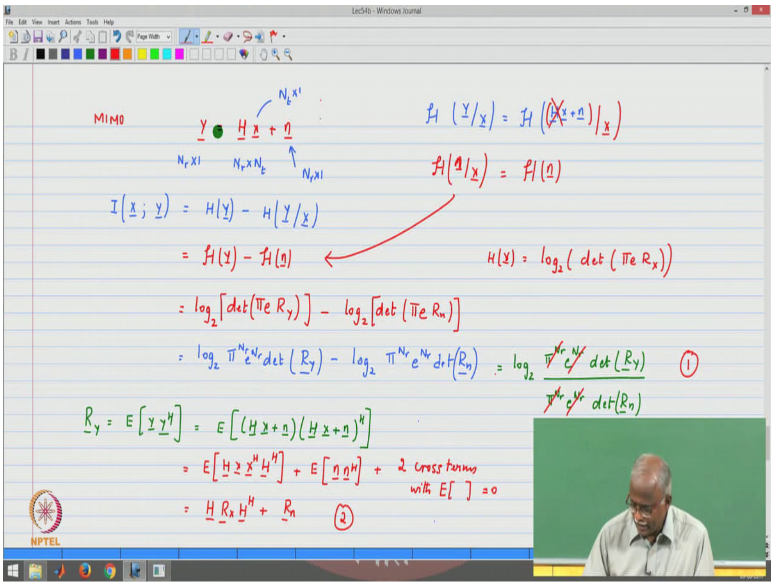
We can write that the entropy of a vector source is given by n times logarithm base 2 of pi, plus logarithm base 2 of the determinant of R, plus N times N coming from the trace of the identity matrix logarithm of e base 2. So, we will rewrite this in the following form write the first term and basically combine the terms in inside the logarithm, the first term will be pi raised to the power N that is the first term. The second term gives us e raised to the power N as you can see, and then the third term gives us a determinant of R ok.

Now, there is another result which says that if I multiply all of the elements of a matrix by a constant alpha. So, basically if I take alpha times R and then take its determinant, then what we get is alpha raised to the power N times the determinant of (Refer Time: 00:00), because every element in your determinant computation has got alpha power N because each of them has got the factor alpha. So, based on this result we can rewrite this into the following form, the form that we would like to write it is logarithm base 2.

Now, I am taking these 2 constants inside the determinant. So, it logarithm base 2 determinant of pi e times R. So, pi e is the multiplicative factor, it multiplies each of the elements of R. So, therefore, when you take the determinant you what you will get is pi power n e power n determinant of R, basically using this razor. So, the important result for us is that when you have a complex Gaussian vector, x is a complex Gaussian vector with the vector p d f, f x of x being given by the expression that we have for Gaussian. Then we can compute the entropy of such a source, H of this vector source in terms of these following expression logarithm base 2 determinant of pi e R, where R is equal to expected value of x times x Hermitian. So, this is a very useful result and we will use this very extensively in our discussion of the capacity of the of the MIMOs channels system when we do not have the information about the transmitter about the channel at the transmitter side ok.

I hope you will have a chance to relook at this and derive the results. So, that you are very comfortable with this result. Now let us move back to the MIMO system that we have.

(Refer Slide Time: 26:40)



So, the MIMO system that we have gives us a vector y equal to H times x plus the noise vector again this is how what we have we have expressed the. So, just the dimensionality just for to remind ourselves, this is N r cross 1, H will be N r cross N t x is N t cross 1 and n is N R cross 1 again this for completeness the dimensionality. So, now, if I wanted to understand the capacity of this MIMO channel H then I would look at the entrope or the mutual information between the vector that was transmitted and the vector that was received. Mutual information between x and y from the scalar case we now have expression in the form of a vector case.

So, this can be again using the same principles as for the scalar case, we can write this as H of Y sorry H of Y minus H of Y given x, all of these are vectors H of Y given x. Now I would like you to look at the second term and we will make a simplification of the second term. So, just the second term alone H of the vector y given the input vector x can be written as the entropy of the matrix H, maybe just to avoid confusion just in this part I will use the script H for the entropy just. So, that we do not get confused because there is a matrix H as well. So, the script H represents the entropy and the plain H and channel matrix. So, this is equal to H x plus in that is the expression for y given x. So, that is the entropy that we are interested in computing. Now entropy is the measure of uncertainty now if you look at this first term H x, once you are given x and you know H there is no uncertainty in H x. So, therefore, what we can say is the first term can be omitted because there is no uncertainty after once you know x once you are given x there is no uncertainty in H x.

So, the uncertainty only will be because of the noise and therefore, it will be the entropy of n given x. Now we have also made the assumption that we also made the assumption that there is no correlation between the noise and the input sources. So, therefore, what we can write is H of n given x basically there is no information that you can get about the noise from the input vector x. So, therefore, this can be written as the information of the entropy of the noise vector itself. So, just by a 2 step argument we have shown that y given x is the same as n itself.

So, this expression for the mutual information can now be written in terms of the entropy of y minus the entropy of n, using the result from this discussion. So, given this and the fact that we know that for any vector source, we said that for any vector source the expression for the entropy is given by logarithm base 2 of pi e times sorry pi determinant of logarithm base 2 of the determinant of pi e times R. So, that was the we will write it as

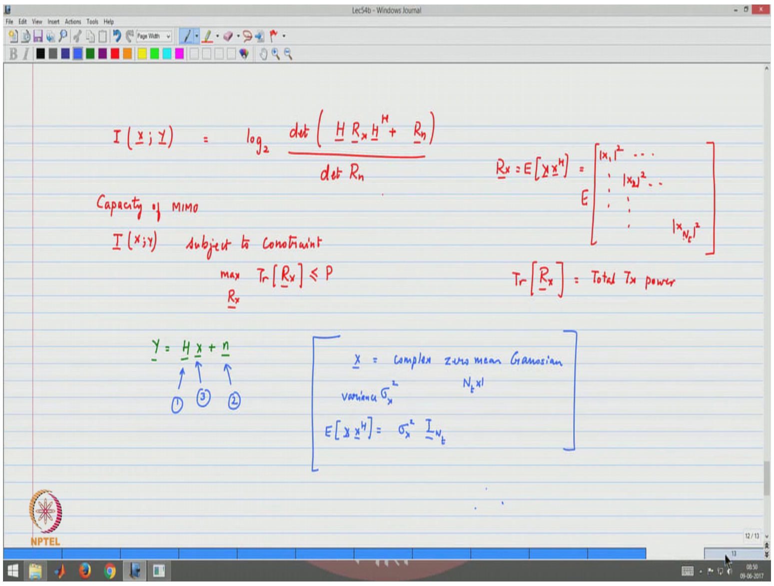
R subscript x. So, when you have the we want to compute the entropy of the vector x, we then have the this expression that we have. So, basically now we should be able to write down the mutual information in the following manner.

So, this will be logarithm base 2 of the determinant of pi e R y logarithm base 2, minus again for the vector n it will be logarithm base 2, determinant of pi e of R n. Now keep in mind that y is has dimension N R cross 1, N has dimension N r cross 1. So, if you were to simplify these expressions or then this can be actually written in the following form the pulling out the constants pi and e. So, it will be logarithm base 2 pi power N, e power N determinant of R y minus logarithm base 2 pi power N actually it should be N r, pi power N r e power N r times determinant of R n. Now again a simple simplification step combining the 2 logarithm terms, this can be written as logarithm base 2, you divide one term by the other the constants pi power N r and e power N r get cancelled, what we are left without let me just write it down, pi power N r e power N r determinant of our y divided by pi power N r, e power N r determinant of R n. Of course, the simplifications cancel these 2 terms and what we are left with as the mutual information or the capacity is logarithm base 2, determinant of R y can be written in terms of the x (Refer Time: 34:06) maybe we should obtain the expression for the determinant of our of R y. So, let me just take one more step.

So, R y would be given by expected value of y y Hermitian substitute the expression for y H x plus n, expected value of H x plus n times H x plus n Hermitian. Now when we multiply these 2 terms there will be 4 terms out of which 2 terms have got the cross products, the cross product of x with n, x n n. Now when we take the expected value both x and n if they are n is a zero mean complex Gaussian vector. So, therefore, we are then able to say that the expected value and they are uncorrelated.

So, therefore, when you take the expected value those cross terms will go to 0, leaving us with only 2 terms that we need to be careful about. So, this would be equal to expected value of the following expression H times x times x Hermitian times H Hermitian x mission times H Hermitian. So, that is one term that is will be nonzero, and then the second term that will be nonzero that we expected value of n n Hermitian. The there would be 2 cross terms which plus 2 cross terms with expectation 0, we will just mention that plus 2 cross terms with expected value with expectation equal to 0 because of the zero mean property of then of the noise vector. So, given this then what we can write down is take the expectation inside this is nothing, but H times R x times H Hermitian plus the second term becomes the R of the or correlate autocorrelation matrix of the noise vector.

(Refer Slide Time: 36:53)



So, using this result if this was equation 1 and this was equation 2 then we can write down the combined expression for the channel capacity. I of x comma y mutual information or the channel capacity can be written in the following form. It can be written as the logarithm base 2 of the determinant of the following where you have H times R x plus the H Hermitian plus R n divided by the determinant of R n. This is a very very useful expression for us in terms of the capacity of the MIMO channel. So, and keep in mind that we would like to maximize the mutual this mutual information and thereby maximize the capacity. So, we can now say that the capacity, the capacity of a MIMO channel. Capacity of the MIMO channel can be obtained as the maximization of the or the mutual information x y subject to the constraint that we have a maximum power limit, and the maximum power limit if you notice consists of the power with which we have transmitted each of the transmit signals, x one through x L and that comes in terms of the total power level. So, if you take expected value of x times x Hermitian, what we will have along the diagonal n elements will this will be expected value of x 1 squared other terms, other terms, second term will be mod x 2 squared other terms and then finally, we have mod x N t whole squared mod x N t squared. So, this is basically what you have along the diagonals are the powers of the signals transmitted by each of the antennas.

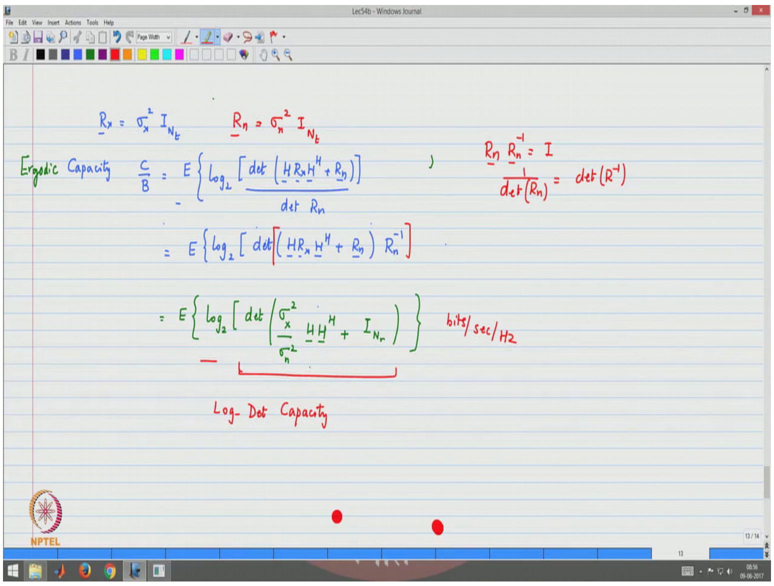
So, if you take look at R x, this is R x and you look at and take the trace of R x. The trace of R x actually gives you the total transmit power this is equal to the total transmit power, because you have N t and transmit antennas and you have taken expected value of mod x squared from each of these antennas the transmission. So, what it says is the power constraint can be written in a very compact form such that you are trying to maximize the mutual information sub using the constraint that the transmit power the trace. The trace of R x is less than or equal to the total power P, and we will maximize this over all possible R x that means, that you are looking at various power distributions and based on this power distribution you want to find out such you want to maximize the mutual information under the constraint that you have a total power constraint. So, this is a very very useful and powerful formulation, in order for us to simplify the expression we will now make a important assumption which will help us get a feel for the capacity of the channels.

Now, if you look back to 2 lectures ago we made in the equation Y is equal to H x plus n. We made three Gaussian assumptions, the first Gaussian assumption was the entries of H that these are complex Gaussians which will give us a Rayleigh distributed channel coefficient between each of the pair of antennas. So, that was the first Gaussian assumption, the second Gaussian assumption was for regarding the noise that these are zero mean uncorrelated Gaussian noise samples, and then the third one that we said was for the ease of analysis we will also make the assumption that the vector x is also Gaussian has got a Gaussian distribution. So, that was the third Gaussian assumption that we had made.

Now, if you make the assumption that x is Gaussian, then we are also able to simplify this expression because in that case the capacity of a Gaussian distributed random variable at the input then tells us that we can write down this maximization in a very very compact form. So, if x is a complex Gaussian, has got a complex zero mean Gaussian. If this was the distribution or the statistical characterization of x, then we can write down the following the x has got a variance sigma squared it is a zero mean complex Gaussian with variance sigma square and therefore, maybe we write it as sigma x square avoid confusion, we can write down expected value of x x Hermitian. Now because they are zero mean the cross terms will go to 0 the diagonal terms will be equal to sigma x squared times the dimensionality of x.

So, if x is a complex zero mean complex Gaussian. So, basically if x is N t cross one this will be I of N t. So, this is a very very useful result and based on this result we can now write down the channel capacity because we are now dealing with a complex Gaussian vector very easy for us to compute the maximum entropy and therefore, the channel capacity. So, the R x is given by sigma x squared times I N t.

(Refer Slide Time: 44:02)



So, the expression for the channel capacity, the channel capacity given by C divided by B, this is equal to the expected value of logarithm base 2 of the determinant of H R x times H Hermitian plus R n, H Hermitian plus R n divided by the determinant of R n and if you now simplify this expression as we have indicated, then what we can write this down in the following way. That this is equal to the expected value of the logarithm base 2 of determinant H R x, H Hermitian plus R n times R n inverse and this is within the bracket. So, again using the property that R the R n inverse R n inverse is equal to the identity matrix. So, therefore, the determinant of R n 1 by determinant of R n is equal to determinant of R inverse.

So, using this result we can simplify this expression and basically take R n inside also write down the simple substitute for the R x expression. So, this we can write down as the in the following manner, this is equal to expected value the of the logarithm base 2 this is the Ergodic Capacitythat is why there is a second expectation. Ergodic capacity is expected value of logarithm base 2 of determinant sigma x squared just for reference R n also has got the similar structure it is sigma n squared times I N t. So, we can write this down as sigma x squared by sigma n squared what is within the inside is H times H hermitian plus sorry this is R xs H times H Hermitian, plus I times N r determinant of this expression. So, this is the and the units of this are in bits per second per hertz and this is also called as the log debt capacity the reason why it is called log debt capacity is because we are taking logarithm of a determinant. So, this is called the log debt capacity of a MIMO channel ok.

So, what are the key elements of the log debt capacity we have said that there is information it is a N t transmit antennas N r receive antennas, there is no information about the channel is not, available at the transmitter. So, therefore, we would have to go for Ergodic capacity, we have obtained the expression for the mutual information. Once we have the expression for the mutual information we are then able to in a through a series of steps we are able to get the expression for the mutual information, then if you make the assumption that the input vector is also a complex Gaussian then we get the following simplification where we can write down the argotic capacity in terms of the log that formula log debt capacity formula. Logarithm base 2 of the determinant sigma x squared variance of the vector x components of x sigma n squared various of the component of the variance of the components of n H which is the matrix times H Hermitian times I N r ok.

So, this is our baseline expression this is what we have as the log debt capacity, we would now like to build on this in the next lecture for a brief period, to understand how this log that capacity can be validated in terms of the capacities that we are familiar with for a MIMO channel. For example, the typical MIMO channel could be received diversity, we also have talked about transmit diversity, now Alamouti scheme how does that fit in into this. So, basically give given this a framework of a MIMO channel, that we are now able to derive the log debt capacity under the conditions of only CSIR that means, the channel information is only available at the receiver not at the transmitter. So, based on that we now will be able to do the rest of the analysis, and that we will pick it up in the next lecture.

Thank you.