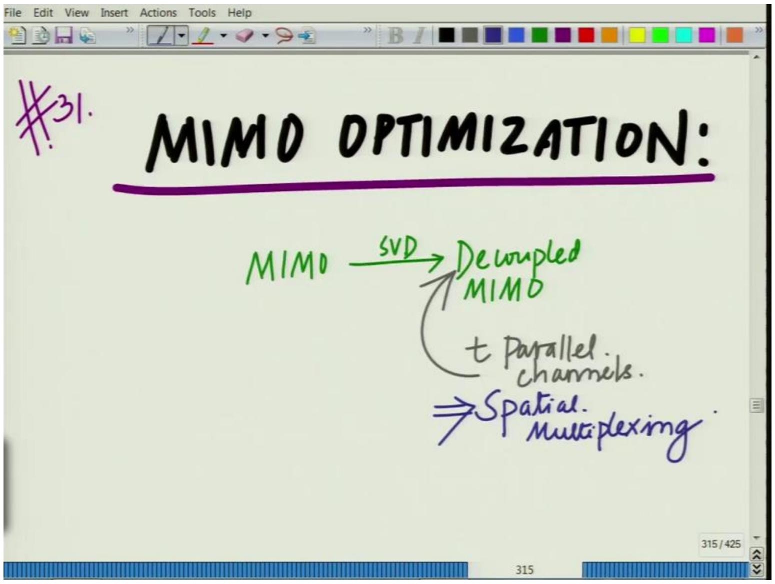
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur
Lecture 31
SVD for MIMO Wireless Optimization, Water Filling Algorithm, Optimal Power Allocation

Hello, welcome to another module in this massive open online course, so let us continue our discussion in the application of SVD singular value decomposition for MIMO wireless communication and let us specifically look at MIMO optimization, that is how to optimize the performance of the multiple input multiple output wireless communication system.

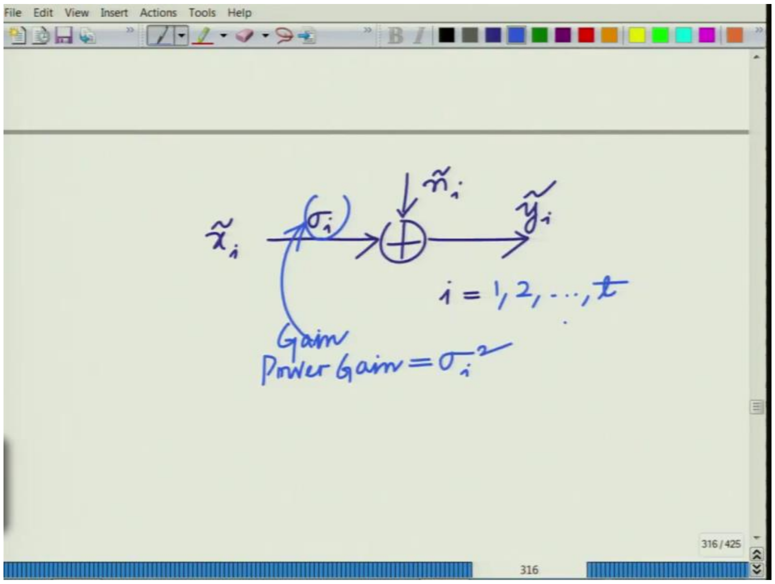
(Refer Slide Time: 00:37)

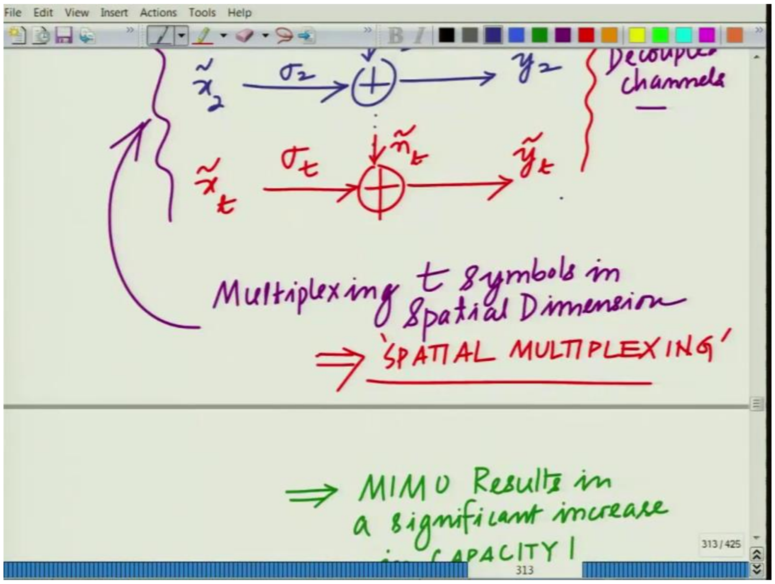


So, we want to look at an application of SVD and more specifically we are looking at how to optimize the performance of the MIMO wireless communication system. And so what we have seen so far in the previous module is that applying the SVD you can take the MIMO system once you apply the SVD this becomes your decoupled MIMO comprising of the t parallel channels.

So, that is playing the MIMO using the precoding at the transmitter and the post-processing combiner at the combining at the receiver you can convert this coupled MIMO system into a decoupled MIMO system comprising of t parallel channels and then therefore we have we had seen that essentially implies that you can multiplex or simultaneously transmit t symbols over the same time and frequency and this therefore leads to spatial multiplexing, you remember. So, you can recall that this is what we had termed as spatial multiplexing.

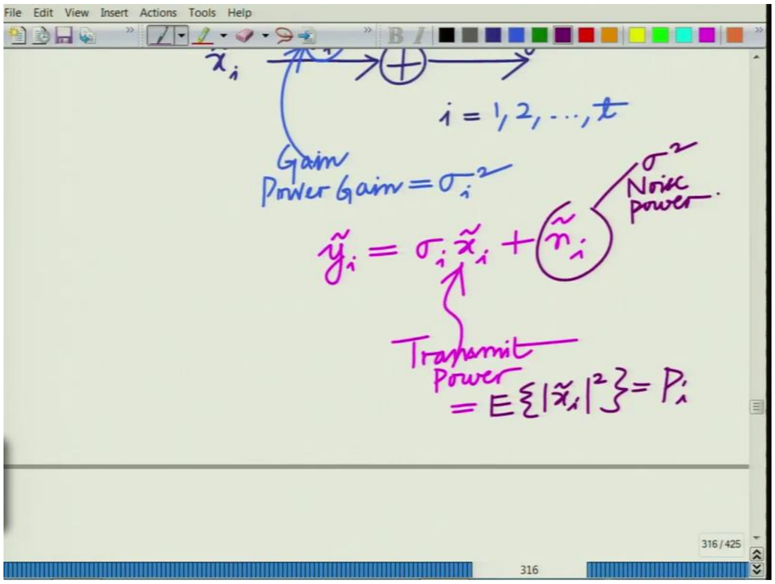
(Refer Slide Time: 02:40)





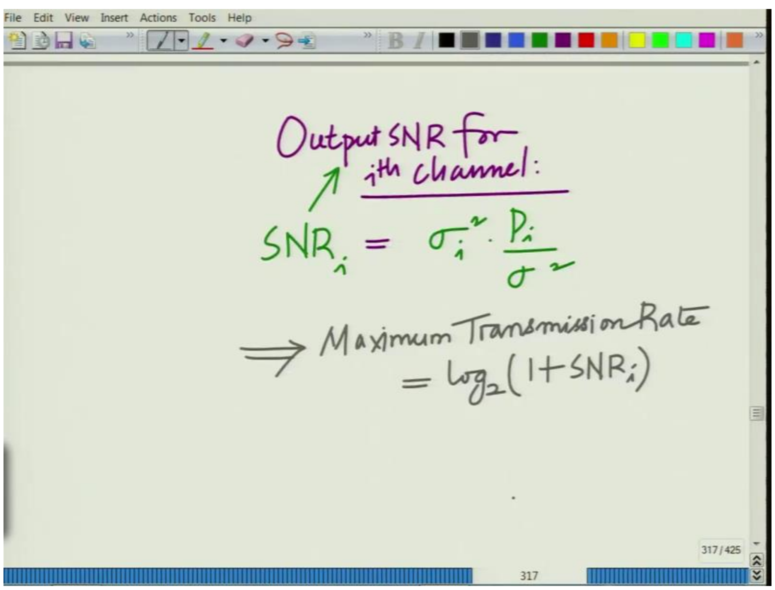
And each of this channel is given as if you can recall that is we have effectively we have xi tilde that is being transmitted through a channel with gain sigma I and then you have the noise which is n tilde i and you have the output which is basically your yi, that is essentially what we had seen here, yi tilde is the output. And this is therefore i equal to 1, 2 up to t and the sigma i you can see what is this sigma i, this sigma i is basically the amplitude gain, the power gain will be gain in the power or the SNR will be sigma i square.

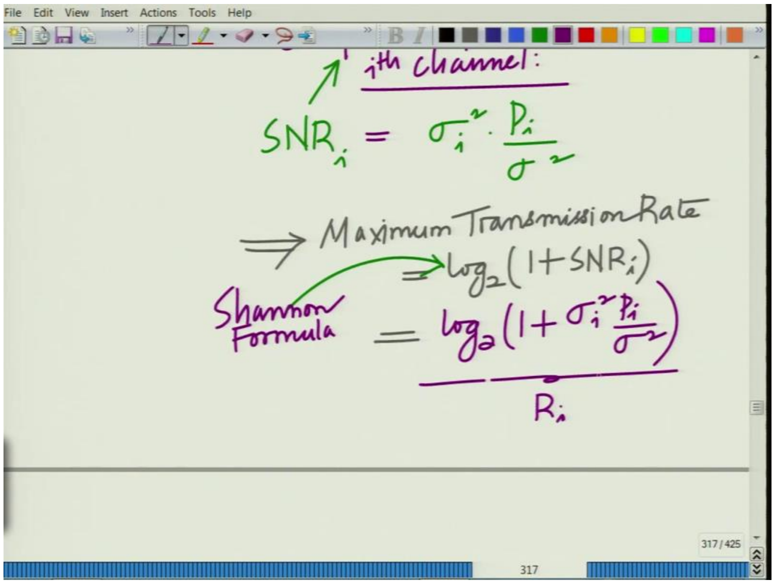
(Refer Slide Time: 03:47)



The model is yi tilde equals sigma i xi tilde plus n tilde i, let us now assume that the transmit power of the ith symbol this is equal to expected value of magnitude xi tilde square let this equal to Pi, let us call this pi and let us call the noise power as sigma square, this is the noise power.

(Refer Slide Time: 04:35)

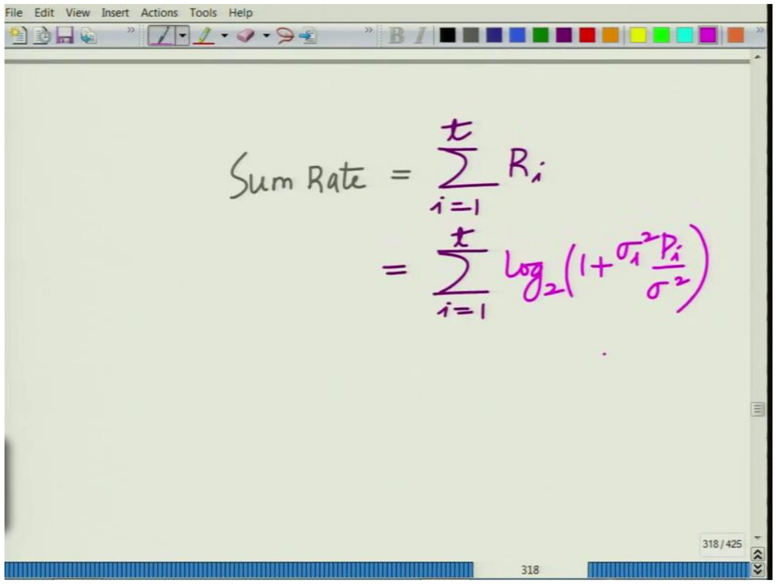


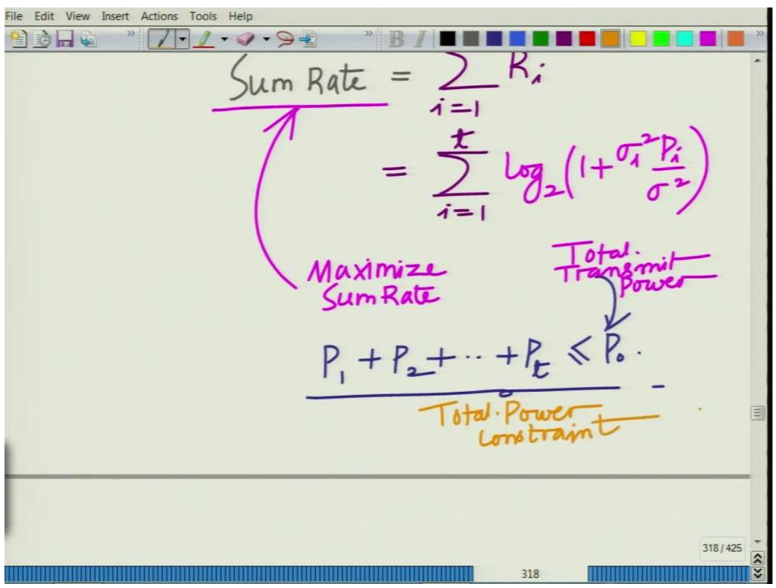


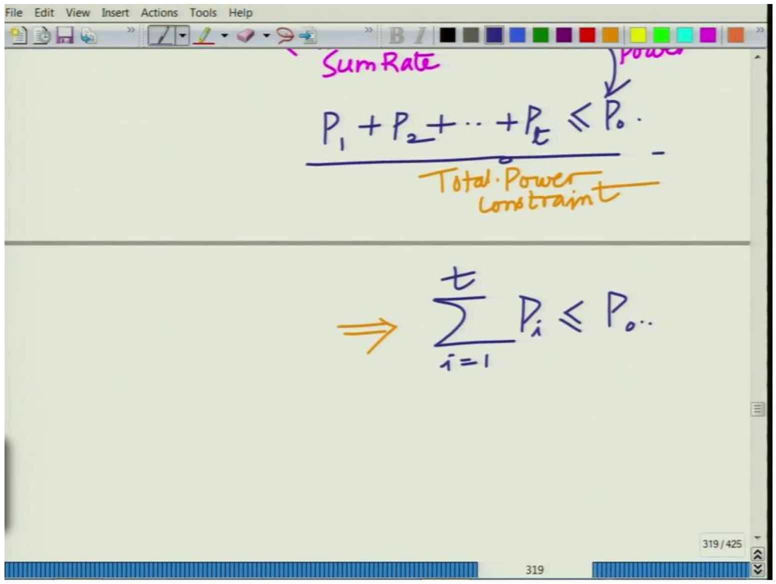
So, the output SNR for the ith channel, this will be equal to because we have the amplitude gain sigma i so we will have a factor the gain of the power will be sigma i square sigma i square times the input power which is Pi divided by sigma square which is the noise power, so that is the output SNR for the ith channel, sigma i square pi divided by sigma square this is the output SNR for the ith channel which implies the maximum rate from the Shannon formula which implies the maximum transmission rate for error-free decoding which is given by the Shannon capacity.

It is given as log to the base 2 1 plus SNR i which is essentially now if you substitute the expression for so this is essentially your Shannon capacity formula, this is the well known channel formula which is given as therefore substituting the value for SNR i, so you have sigma i square pi divided by sigma square. So, this is the maximum rate i for the ith channel. And now the some rate, so this is the rate i for each channel i and therefore the some rate is given as is basically obtained by the sum of the rates of these individual t channels others.

(Refer Slide Time: 07:16)





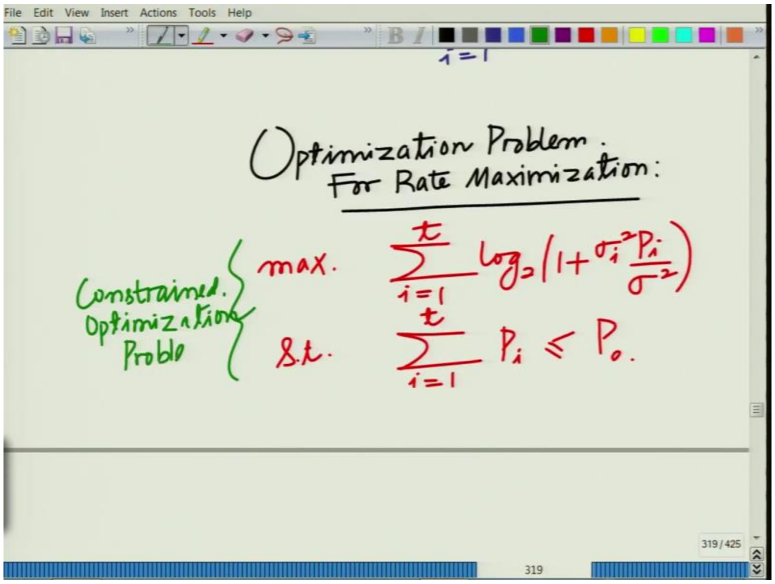


So therefore, the sum rate, this is equal to summation i equal to 1 to t Ri which is summation i equal to 1 to t log to the base 2 1 plus sigma i square pi over sigma square, this is the maximum this is the sum rate and now we want to maximize the sum rate, now typically whenever we look at a wireless communication there is a constraint on the transmit power, the transmit power is not unlimited.

So, which essentially implies that this if you look at all these t parallel channels, which are powers P1 P2 Pt the total transmit power has to be constrained by a maximum power that is we look at summation that is P1 plus P2 plus all up to Pt that has to be less than equal to the total transmit power, let us call this as P naught.

So, typically we have the total transmit power that is if you look at the powers of these t channels P1 plus P2 plus so on up to Pt this has to be constrained to be less than or equal to P naught you can call this as the total transmit power. So, this is the total power constraint, so this implies that you will have summation i equal to 1 to t Pi is less than or equal to P naught.

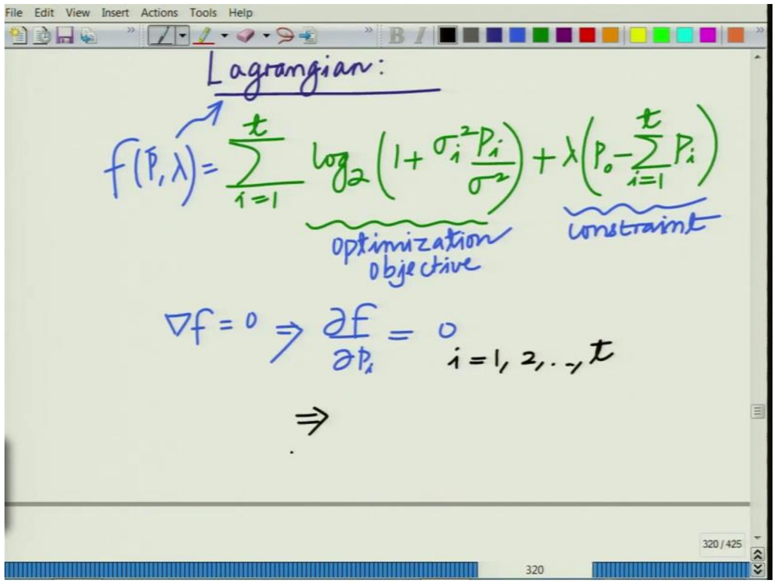
(Refer Slide Time: 09:54)

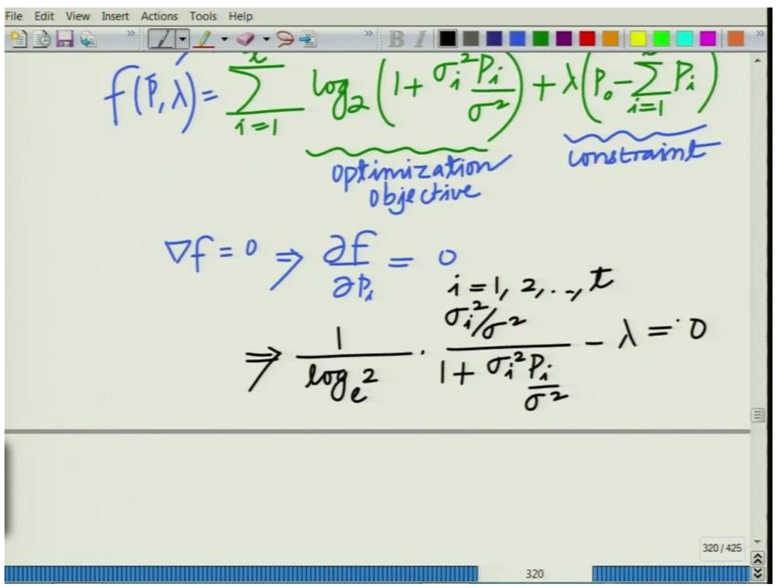


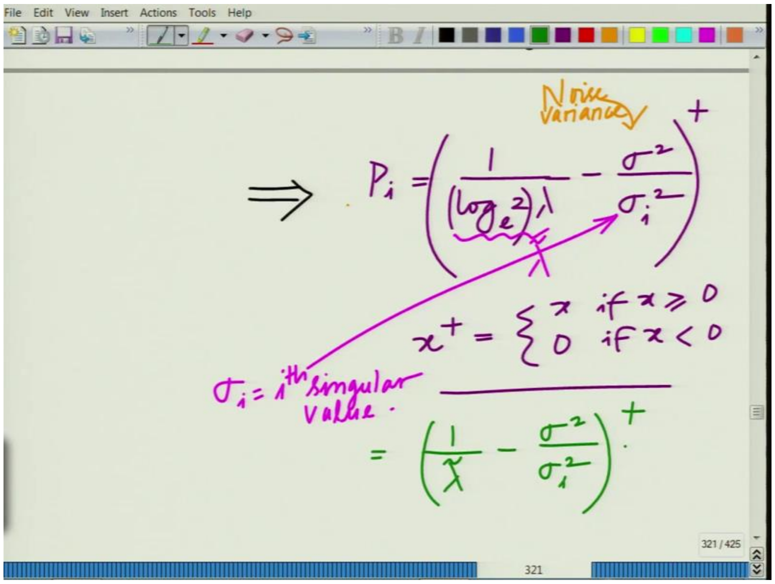
And therefore the optimization problem for my MIMO rate maximization this can be formulated as the optimization problem for rate maximization that will be maximized, some i equal to 1 to t, the some rate log to the base 2 1 plus sigma i square Pi divided by sigma square subject to the constraint i equal to 1 to t Pi less than or equal to the total power P naught and this is again we have a constrained, if you remember this is a constrained optimization problem, summation of Pi less than or equal to P naught.

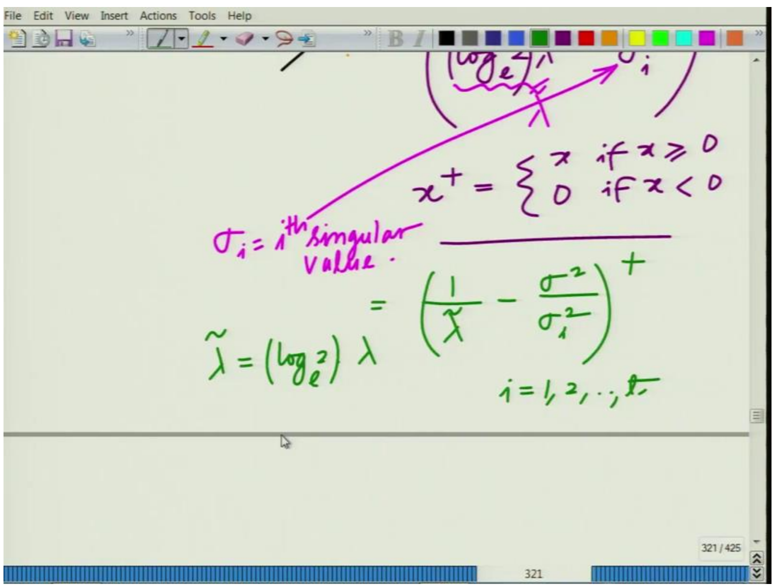
So, in to maximize the sum rate that is sum of the rates some of the transmission rates Ri across the t the t parallel channels, that constitute this MIMO channel subject to the power constraint that is the total transmit power that is summation Pi has to be less than or equal to P naught which is the total transmit power. And as we have seen several times before the way to solve this constraint optimization problem is to in is to introduce the Lagrange which I used to introduce its construct the Lagrangian using the Lagrange multiplier Lambda.

(Refer Slide Time: 12:08)









So, we construct the Lagrangian, and what is the Lagrangian? That is if you look at summation i equal to 1 to t log this is the optimization objective log to the base 2 1 plus sigma i square Pi by sigma square times 1 lambda 1 minus summation i equal to 1 to t Pi or in fact summation lambda P naught minus summation i equal to 1 to t Pi, so this is the if you recall this is the optimization objective and this is the constraint.

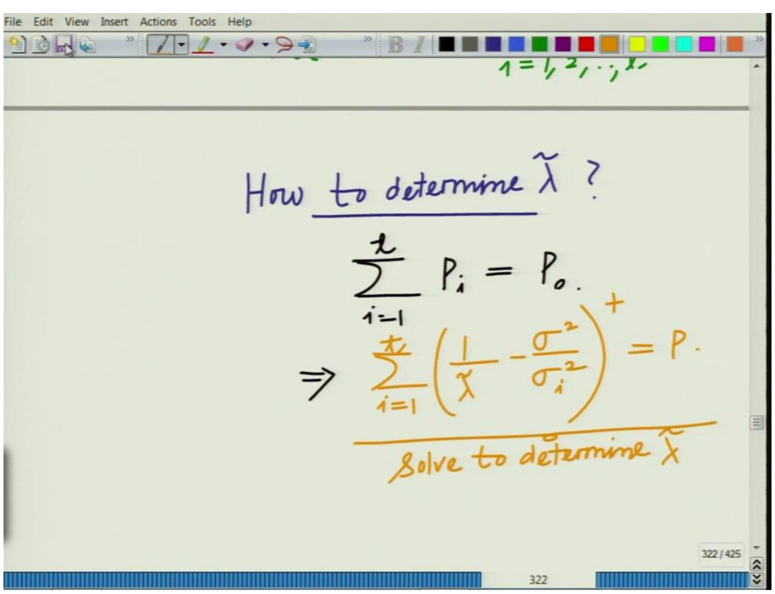
And therefore now this is a function of your this is basically this is the Lagrangian and now we have to take the partial derivative of f with respect to so the condition the KKT condition is that the gradient of f has to vanish which implies that the partial derivative with respect to each Pi this has to be equal to 0, for i equal to 1, 2 up to t.

So, this implies now you differentiate this with respect to Pi this implies that 1 over log 2 to the base e times differentiate this sigma i square over sigma square divided by 1 plus sigma i square Pi over sigma square minus lambda equal to 0, this is the condition you get when you take the partial derivative with respect to Pi and basically set it equal to 0.

And when you solve this, you will get the condition this implies basically that Pi solving for Pi you will get Pi equal to 1 over log 2 to the base e times lambda minus sigma square over sigma i square and here I will introduce this plus because the power has to be positive non-negative, so where this x plus equal to x, if x is greater than or equal to 0, 0 if x is less than 0. So, this is this quantity 1 over log 2 to the base e times lambda minus sigma square that is the noise variance divided by sigma i square which is essentially the power gain the gain of the channel and sigma i you recall is the ith singular value.

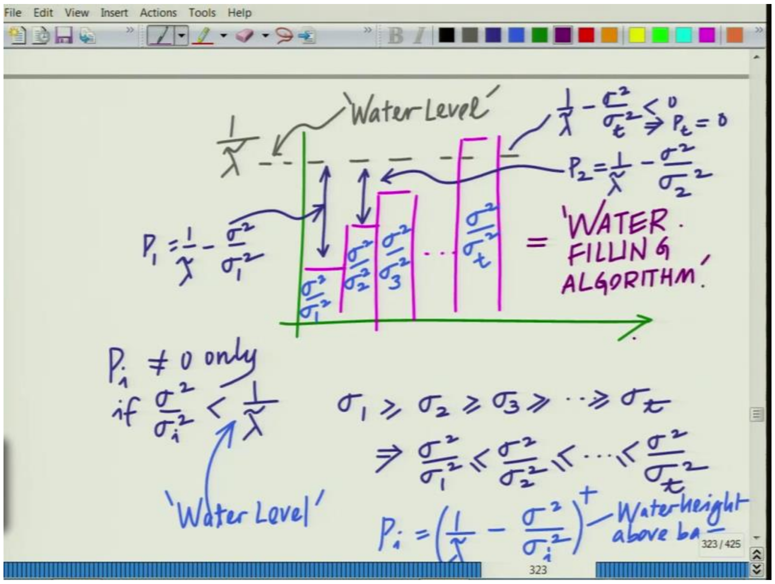
And this plus essentially used to indicate that it can be this quantity only if it is non negative. If it is negative, then the power cannot be negative so you have to set the power equal to 0. So, let us recall these different quantities sigma i square this is the noise variance and if you recall this sigma i, sigma i equals ith in fact this log 2 to the base e lambda you can set it as a parameter you can simply call this as lambda tilde, because this is just a parameter so you can call this as 1 over lambda tilde minus sigma square over sigma i square plus, this is for i equal to 1, 2 to t where lambda tilde equals log 2 to the base e times lambda, and how to find lambda tilde?

(Refer Slide Time: 17:35)



And of course now we ask the question how to determine lambda tilde. Remember we still have the power constraint, so we have summation i equal to 1 to t Pi is equal to P naught which implies now substitute for the expression for Pi summation i equal to 1 to t 1 over lambda tilde minus sigma square over sigma square this is equal to P, so solve this to determine lambda tilde, 1 over lambda tilde that is the summation of the powers all the transmit powers must be equal to P naught which is the total transmit power.

(Refer Slide Time: 18:57)



Now, this if you see this has a very interesting structure this power allocation, if you see this is a very interesting structure this is known that is if you plot this power allocation you will realize that let us first plot the different quantities the different bars each of these bars is of different height this is of height sigma square by sigma 1 square this is of height sigma square by sigma 2 square sigma square by sigma 3 square sigma square by sigma t square this is the last bar.

Now, remember sigma 1 singular values are arranged in the decreasing order sigma 1 greater than equal to sigma 2 greater than equal to sigma 3 so on greater than equal to sigma t which implies sigma square by sigma 1 square is less than or equal to sigma square by sigma 2 square is less than or equal to sigma square by sigma t square.

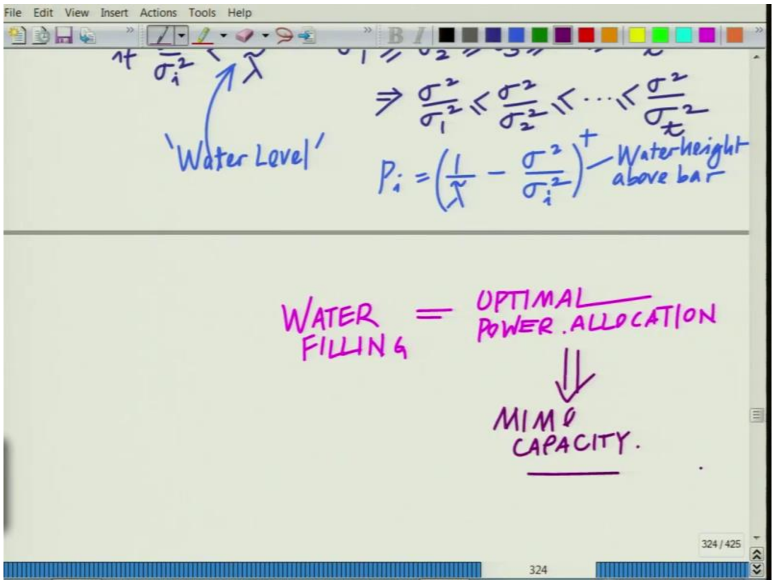
So, this bars are in increasing order of height, sigma square by sigma 1 square is less than or equal to sigma square by sigma 2 square so on and so forth is less than equal to sigma square by sigma t square. Now, if you call 1 over lambda tilde as this level if you think of filling this boil with water and if you call 1 over lambda tilde as the water level, it is convenient to think of this 1 over lambda tilde as the water level, then the power allocated Pi is the height of the water level.

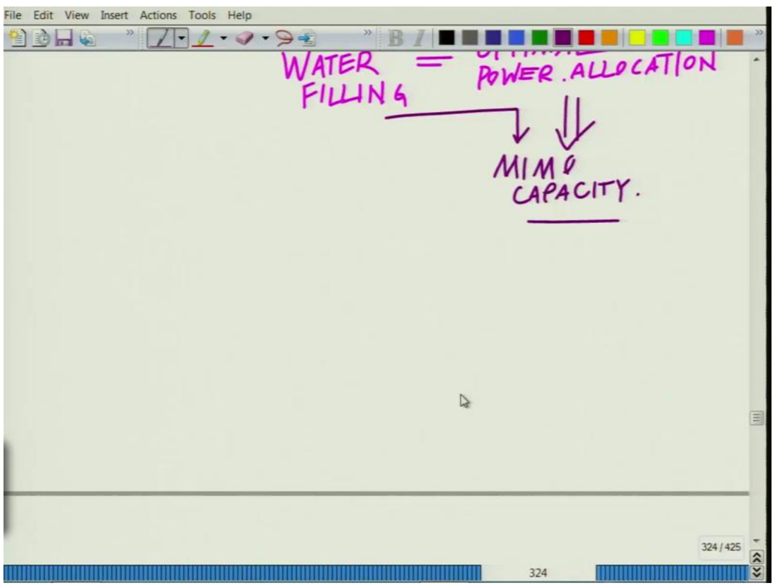
So, Pi P1 equal to 1 over lambda tilde minus sigma square over sigma 1 square this is equal to P2 equal to 1 over lambda tilde minus sigma square by sigma 2 square and now you can see if sigma square over sigma t square is less than is greater than the water level 1 over lambda tilde then the power allocated is 0. Because 1 over lambda tilde minus sigma square over sigma t square is negative.

In this case if you look at this 1 over lambda tilde minus sigma square over sigma t square less than 0 implies Pt equal to 0, implies a transmit power. So, power is allocated only if sigma square or sigma i square is less than the water level 1 over lambda, this is an important, so power is non-zero, Pi is not equal to 0 only if sigma square over sigma i square is less than 1 over lambda tilde, what is this 1 over lambda tilde this is basically your water level and this algorithm is known as the water filling algorithm.

So, you can think of this Pi equal to 1 over lambda tilde minus sigma square over sigma i square sigma i square plus as the water about the bar, water height about bar at this is essentially therefore this is termed as the water filling algorithm, this is basically you this is a very celebrated very popular algorithm for optimal power allocation, this is termed as the water filling algorithm, water filling algorithm for optimal power allocation to allocate the power to maximize the some rate or to maximize the to achieve the capacity of the MIMO system. So, what does the water filling algorithm do?

(Refer Slide Time: 23:54)



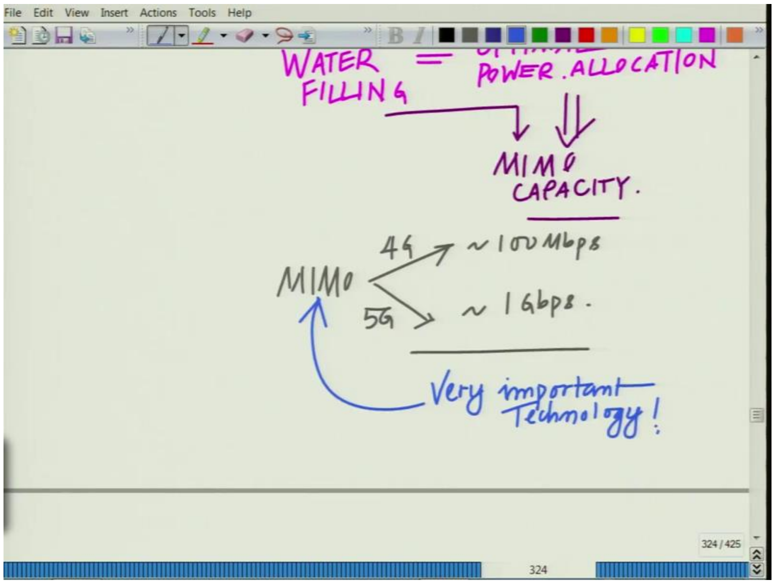


Water filling algorithm it allocates the powers to the sub channels optimally, so this is allocates this is for optimal power allocation which implies this achieves the MIMO this achieves the capacity of the MIMO channel. So, this is basically the water filling algorithm needs to optimal power allocation which achieves the capacity of the MIMO channel.

So, therefore that is essentially how the SVD can be used to maximize the rate of the MIMO system and then we therefore the SVD indeed you can see has a very important tool for optimization of the performance. So, SVD can be used to decouple the MIMO system convert this coupled MIMO system it would be called decoupled MIMO system that is the t parallel channels, transmitting independent symbols x1 tilde, x2 tilde, x3 tilde.

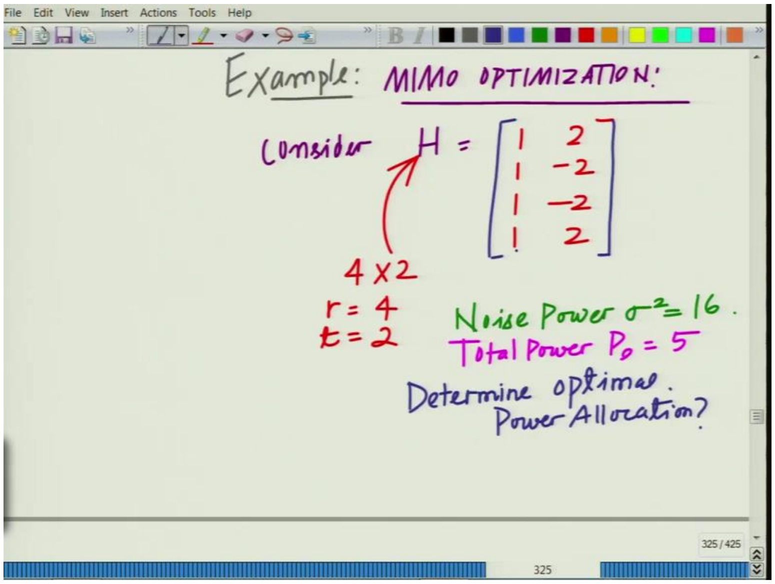
And now you can use the rate maximization framework to allocate the powers in an optimal fashion, so that subject to a total available power P naught using that is subject to this total power constraint P naught the some rate the rate at which you can transmit information reliable you over this MIMO channel multiple input multiple output wireless channel is maximized. And in fact as we have seen MIMO is a very important technology both in 5G 4G and 5G which helps achieve incredibly high rates to the tune of several 100 megabits per second and in fact 5G going up to gigabit per second.

(Refer Slide Time: 25:59)



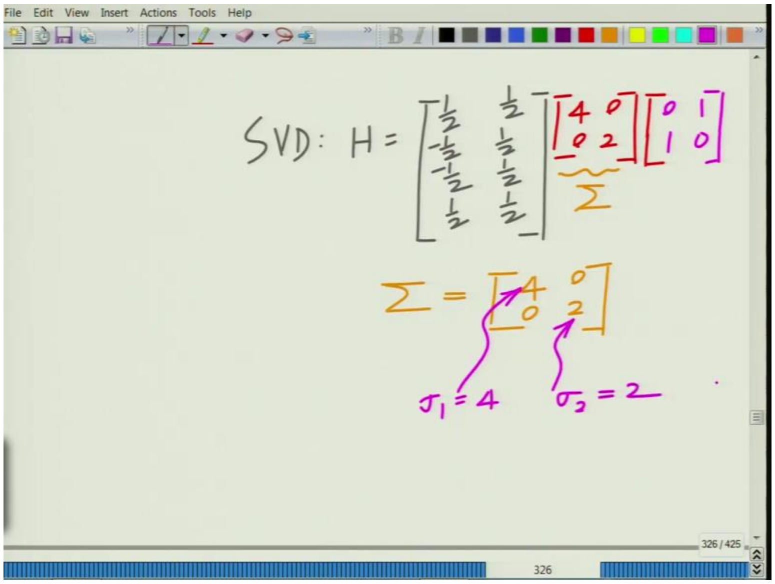
So, MIMO technology in MIMO can be used both in 4G and 5G and under megabit per second and gigabit per second in 5G, so essentially it is a very important this is a very important technology.

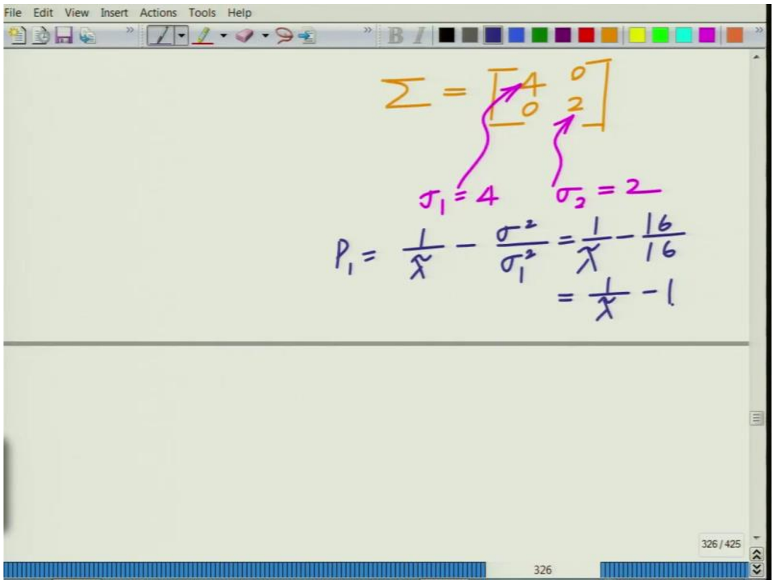
(Refer Slide Time: 26:41)



Let us, look at a very simple example for MIMO optimization, consider the channel 1, 1, 1, 1, 2, minus 2, minus 2, 2, this is our MIMO channel, recalls we already seen this before this is recalled 4 cross 2 MIMO channel which means the number of receive antennas equals 4 number of transmit antennas equals 2, let us now say in this system the noise power sigma square is equal to 16 total power P naught is equal to 5 and now we want to ask the question determine the optimal power allocation. So, we want to determine the optimal power allocation for this system.

(Refer Slide Time: 28:40)

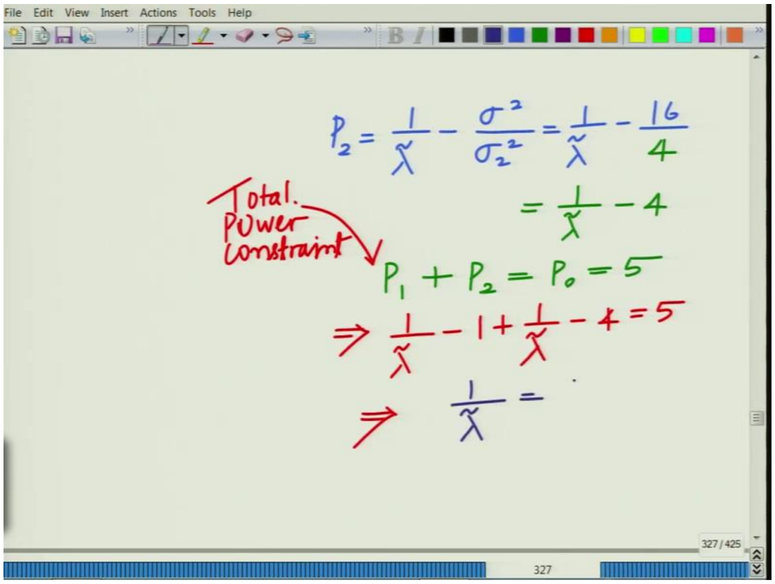


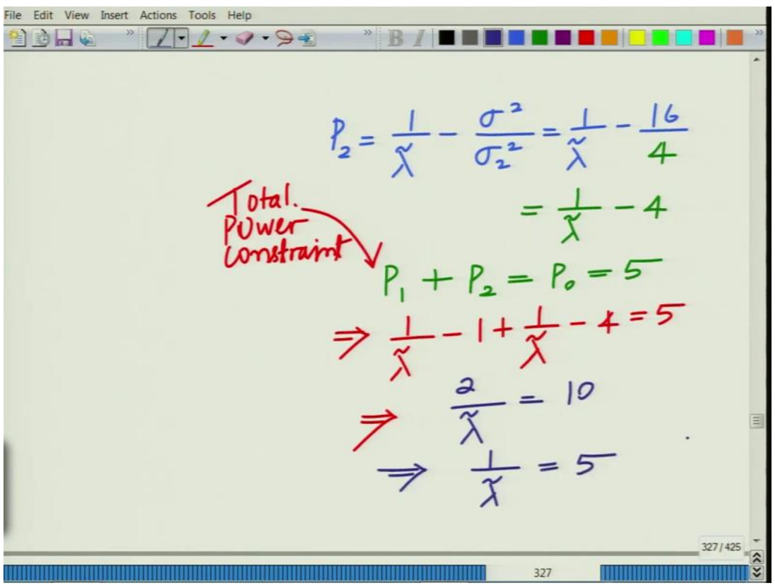


And for that we begin with the SVD and recall we have already performed this video of this in a previous example and the SVD of the channel matrix based on the properties at the columns of H orthogonal, we have already evaluated this video of this you can look at the previous module and the SVD of this is given as half, minus half, minus half, half, half, half, half, half times the singular values matrix of singular values that is 4, 0, 0, 2 times 0, 1, 1, 0 which is the matrix v (())(29:42) and remember these are the singular values.

So, if you look at this is basically your matrix sigma which is the diagonal matrix of singular values, so this is given as 4, 0, 0, 2 and therefore the singular values are sigma 1 equal to 4 this is sigma 2 equal to 2 and therefore now the optimal powers are given as P1 equals 1 over lambda tilde minus sigma square over sigma 1 square which is 1 over lambda tilde minus sigma square that is 16 over sigma 1 square that is 16, so this is 1 over lambda tilde minus 1, of course it is equal to this quantity only if it is greater than equal to 0, if it is less than 0 then it will be 0. So, we have to determine that.

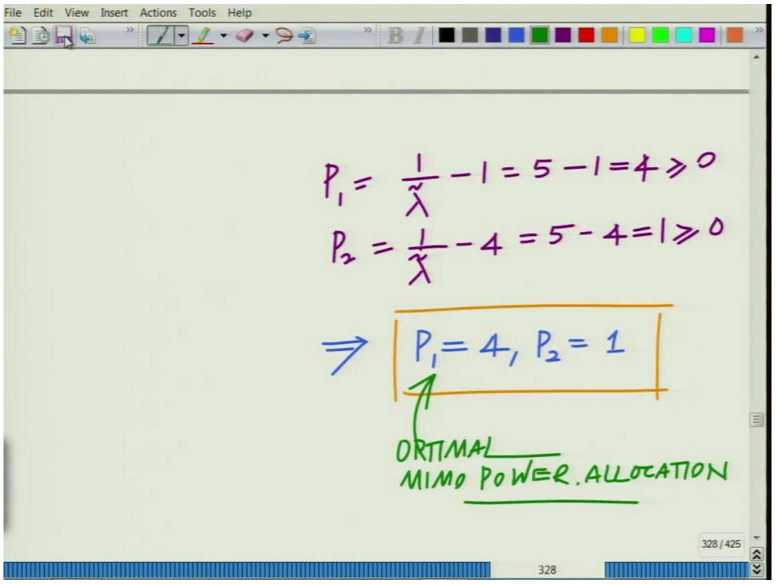
(Refer Slide Time: 30:49)





And P2 equals 1 over lambda tilde minus sigma square over sigma 2 square which is 1 over lambda tilde minus 16 divided by lambda tilde, what is lambda 16 minus sigma 2 square which is 2 square which is 4, so 1 over lambda tilde minus 4. Now, we must have P1 plus P2 equal to P naught which is equal to 5, remember this is our total power constraint, this is remember our which basically implies 1 over lambda tilde minus 1 plus 1 over lambda tilde minus 4 is equal to 5 is basically implies that 1 over lambda tilde equals or 2 over lambda tilde I am sorry, 2 over lambda tilde equals 10, which basically implies 1 over lambda tilde is equal to 5. So, that is essentially what we get.

(Refer Slide Time: 32:17)



And therefore we get P1 substituting back we get P1 equal to 1 over lambda tilde minus 1 which is equal to 5 minus 1 equal to 4 this is greater than equal to 0, so P2 is equals 1 over lambda tilde minus 4 equals 5 minus 4 equal to 1 which is also great than equal to 0, both these values are non-negative, so this is a valid power allocation remember because the power values have to be non-negative.

Otherwise, you have to set those the corresponding power to be 0 if it is negative you have to set the corresponding power to be 0 and then you have to do the problem again. So, essentially this implies that P1 is equal to 4, P2 is equal to 1 and this therefore is basically our optimal MIMO power allocation for a rate maximization, this is essentially the optimal power allocation which is giving by the water filling algorithm.

So, this is a very very interesting and I would say very high impact application of the singular value decomposition you start with the MIMO channel converted into a decoupled channel, look at what is the output SNR for each channel ask what is the output rate maximum transmission rate and look at what is the sum rate and then maximize the sum rate subject to the transmit power constraint that is P naught.

And the optimal power are given by the water filling algorithm and we have also seen an example to understand this water filling power allocation in action and evaluated the powers, the transmit powers of the individual channels, the component channels of this MIMO channel to maximize the sum rate that is to achieve the MIMO capacity. So, let us stop here and let us continue with this discussion in the subsequent modules. Thank you very much.