Applied Optimization for Wireless, Machine Learning, Big Data

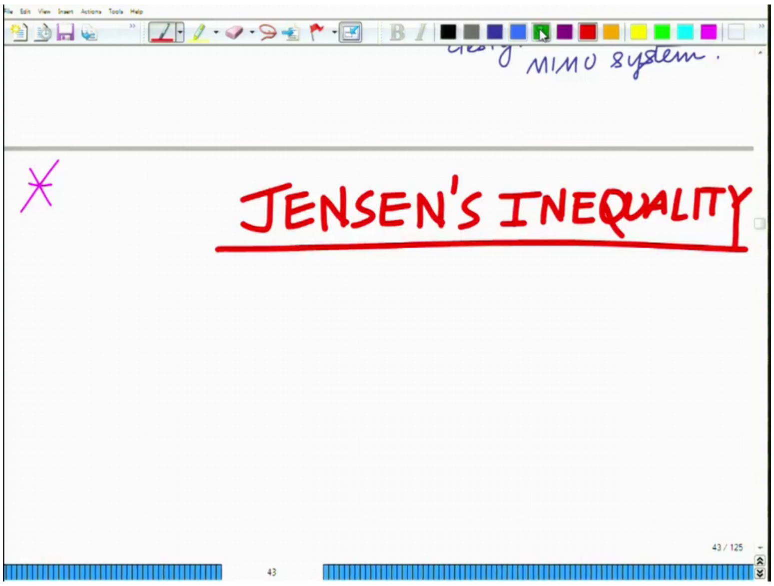
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 27

Jensen's Inequality and Practical Application: BER calculation in Wired and Wireless Scenario

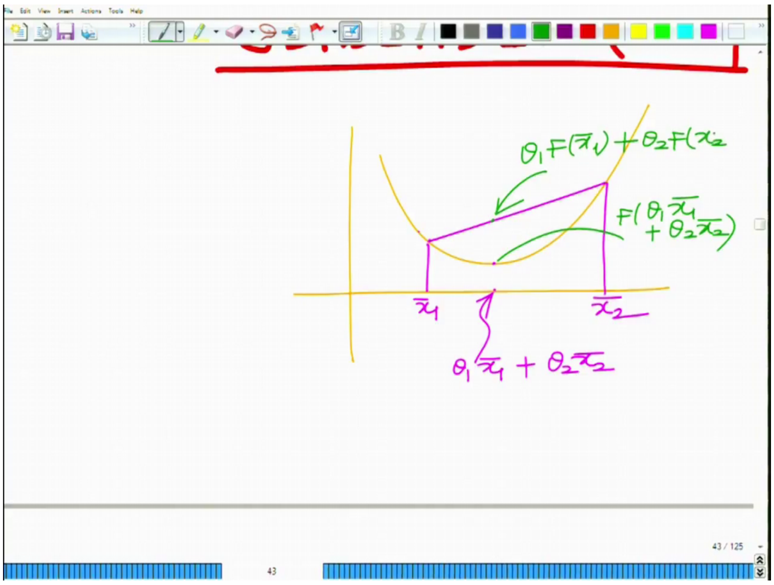
Hello. Welcome to another module in this massive open online course. So, we are looking at convex functions and particular convex functions of a vector, variable vector, correct. And out at the test for convex, if there is a function which convex function, the test of convexity for a function of a vector, alright. So, let us change tracks a little bit and look at something known as Jensen's inequality which has significant applications in various areas, ok.

(Refer Slide Time: 00:49)



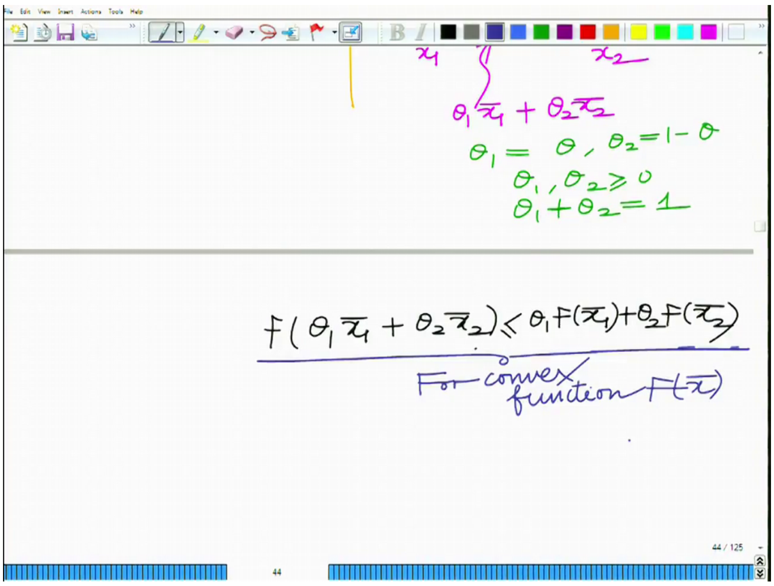
So, what we want to look at in this module is Jensen's inequality for you can say, convex as well as which is a handy tool that arises in several scenarios. In fact, we will try to justify this by looking at a practical application. So, this is a Jensen's inequality and the Jensen's inequality is very interesting and that is as follows.

(Refer Slide Time: 01:17)



That is, remember if you go back and look at the definition of the convex function, it is this bowl shaped function we said and if you look at 2 points x 1 and x 2 and you take any linear combination, that is, let us make this as vectors theta 1 x 1 bar plus theta 2 x 2 bar and this is the value of the function at the linear combination and this is the value of the function that is this is remember, F of theta 1 x 1 bar plus theta 2 x 2 bar. And this is theta 1 times F of x 1 bar plus theta 2 times F of x 2 bar.

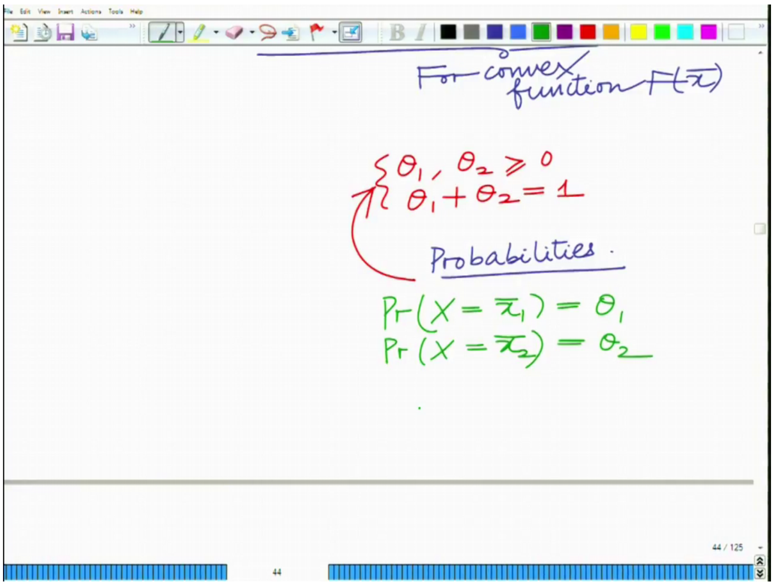
(Refer Slide Time: 02:18)



Where we can say theta 1, you can set it as theta. Remember, theta 2 equals 1 minus theta or you can simply say theta 1 comma theta 2 greater than equal to 0 and theta 1 and theta 2 satisfy this condition that theta 1 plus theta 2 equals 1. So, this is a convex combination of 2 points, theta 1 times x 1 bar plus theta 2 times x 2 bar where, theta 1 theta 1 theta 2 are 2 scalar called 2 numbers which are greater than or equal to 0 which are non-negative.

And sum theta 1 plus theta 2 equals 1 ok, which automatically implies is a non negative and sum to 1 automatically implies that both theta 1 and theta 2 lie in the interval 0 to 1 ok. Further, now you can see that therefore, the inequality for the basic inequality is that F of theta 1 x 1 theta 1 x 1 bar plus theta 2 x 2 bar is less than or equal to theta 1 F of less than or equal to theta 1 F of x 1 bar theta 2 F of x 2 bar this is the inequality for a convex function, for convex function F ok, F of x bar.

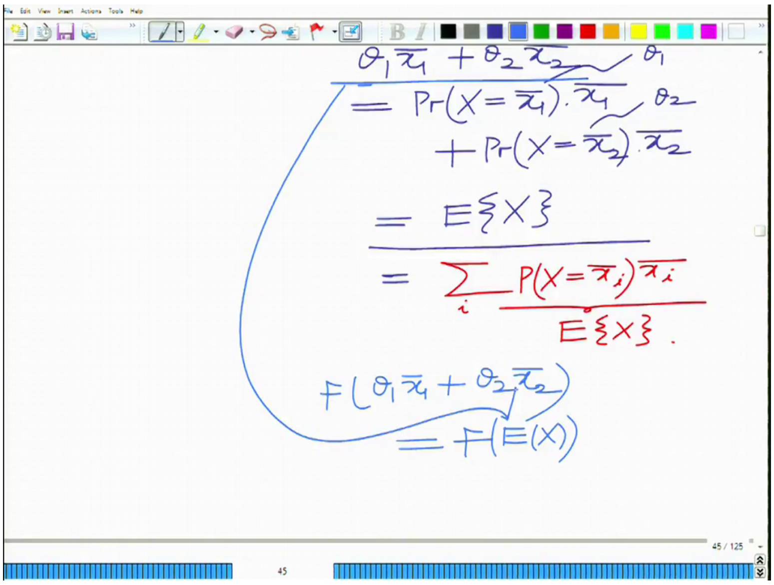
(Refer Slide Time: 03:50)



Now, you can see that, now look at look go back and look at this quantity theta 1 plus theta 2 greater than or equal to 0 and theta 1 plus theta 2 equals 1. Remember, this should remind you of something, these remind you of probabilities. Reason being, you have 2 quantities theta 1 plus theta 2 which are non-negative greater than equal to 0 and the sum to 1.

So, this remind you of the it should remind you of a probability distribution or a probability mass function in this case which you can say that, you can consider a distribution probability x equals x 1 bar equals theta 1 and the probability x equals x 2 bar equals theta 1, that is probability random variable x takes x 1 bar is x 1 bar with probability theta 1 and it is x 2 bar with probability theta 2. Some of the probabilities are greater than equal to 0 and some of the probabilities as well.

(Refer Slide Time: 05:01)

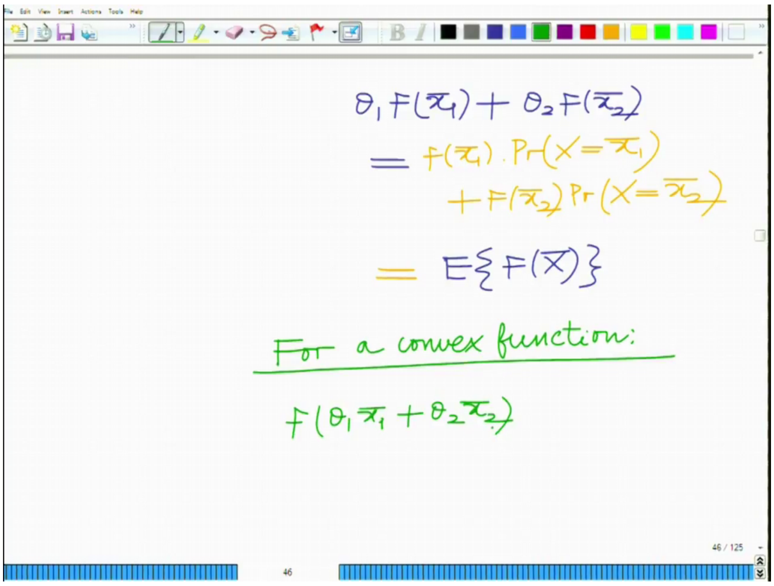


And now, therefore, if you look at this quantity, now, it is interesting if you look at this quantity, theta 1 times x 1 bar plus theta 2 times x 2 bar, this is equal to, what is this equal to? Well, this is if you look at this is x 1 bar or this is probability, X equals x 1 bar times theta. I am sorry times theta 1 is probability X equal to x 1 bar times x 1 bar plus probability X equals x 2 bar times x 2 bar, ok. This probability X equal to x 1 bar is theta 1 probability X equal to x 2, once theta that is what we said about.

And therefore, this is nothing but if you can look at this, this is the expected value of the random variable X. Because, remember what is the expected value? Expected value is nothing but you take each poise possible value, multiply it by the corresponding probability and sum, right. That is expected value. This is nothing but summation over i probability X equal to x bar i times x bar i. This is your definition for the expected value of X ok.

So, this is the definition for the expected value of X. Further, if you look at this quantity and therefore, if you look at this quantity that is F of theta 1 x 1 bar plus theta 2 x 2 bar, you can now write that as F of well that is F of simply expected value of X ok. This random variable x because, remember we have shown that theta 1 x 1 bar plus theta 2 x 2 bar, that is basically equal to this F of that is the expected value of X.

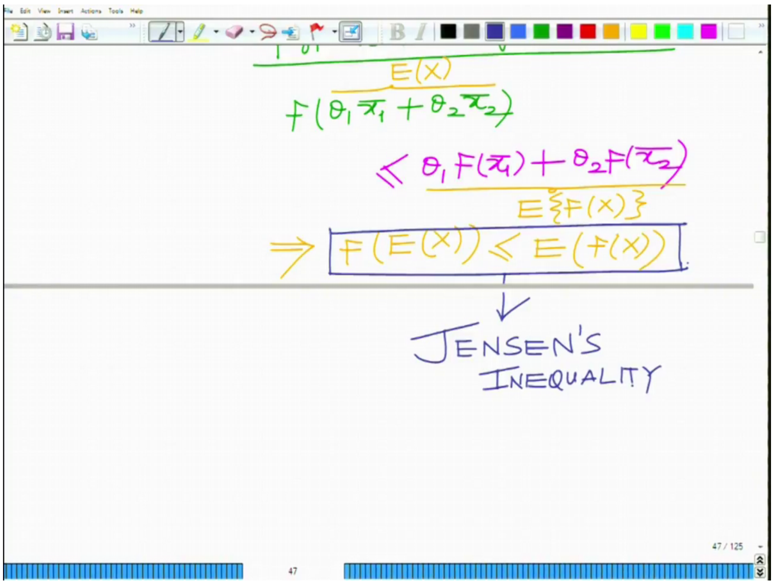
(Refer Slide Time: 07:10)



Now, on the other hand, if you look at this quantity, that is theta 1 F of x 1 bar plus theta 2 F of x 2 bar, well this is equal to again, I can write this similarly, F of x 1 bar times the probability that X equals x 1 bar which is theta 1 plus F of x 2 bar times the probability that X equals x 2 bar. And this is equal to the expected value of F of x bar [noise,] this is equal to the expected value of F of x bar.

And, therefore, what we have now? If you look at for a convex function right; so, what we have? Just shown is that theta 1 F of x 1 bar plus theta 2 of F of x 2 bar, that is the convex combinations of the 2 points, that is a point on the cord is nothing but the expected value of F of x bar. And for a convex function, we know for a convex function we have F of theta 1 x 1 bar plus theta 2 x 2 bar this is less than or equal to theta 1 F of x 1 bar plus theta 2 F of x 2 bar well.

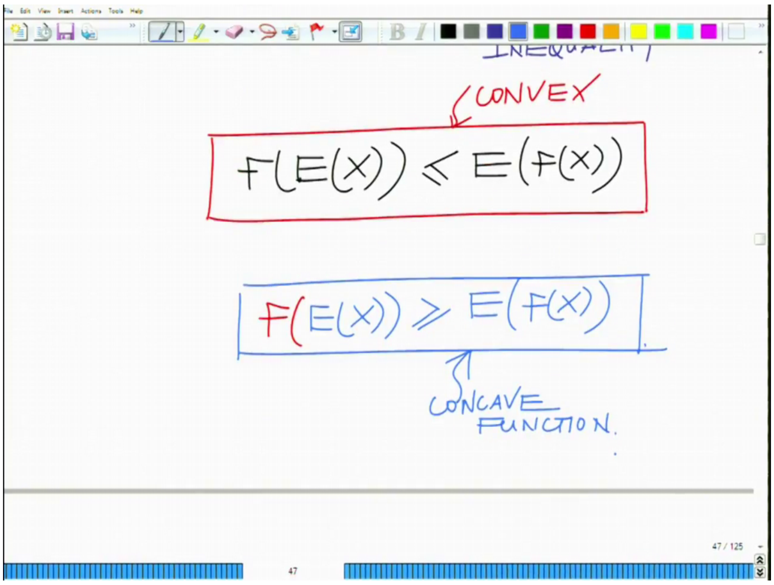
(Refer Slide Time: 08:58)



What is this we have seen is expected value of X and this is expected value of F of X. Therefore, we can represent the same thing as F of expected value of X is less than or equal to the expected value of F of X and this is basically Jensen's inequality for a concave function. Convex function, I am sorry this is basically your Jensen's inequality so, very important and a very handy tool as we just will see in a practical application. This is Jensen's inequality and it is frequently used in signal processing and communications.

Especially, if you look at information theory, there are several instances where this Jensen's inequality is fairly handy to prove various results, all right. It states that for a convex function F of the function of an expected value of a random variable is less than expected value of the function of that random variable.

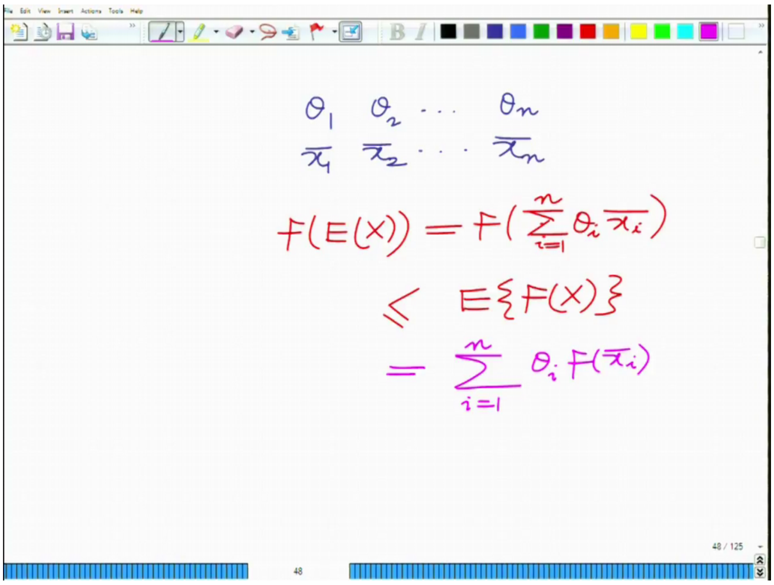
(Refer Slide Time: 10:28)



So, this is again F of expected value of X less than or equal to expected value of F of X ok. For concave or convex and for concave, it is the other way. So, this is for a convex function. For concave, will naturally have the reverse of this is F of expected value of because the cord lies below the curve. So, F of expected value of X is greater than the expected value of F of X. And this is basically for you concave function and this is for your concave function.

So, these are basically you can think of this as basically Jensen's inequality for a convex function and Jensen's inequality for a concave function ok, all right and. So, let us know and you can generalize this too. Well, we have simply considered random variable that takes 2 values, you can generalize it to you can generalize this as follows.

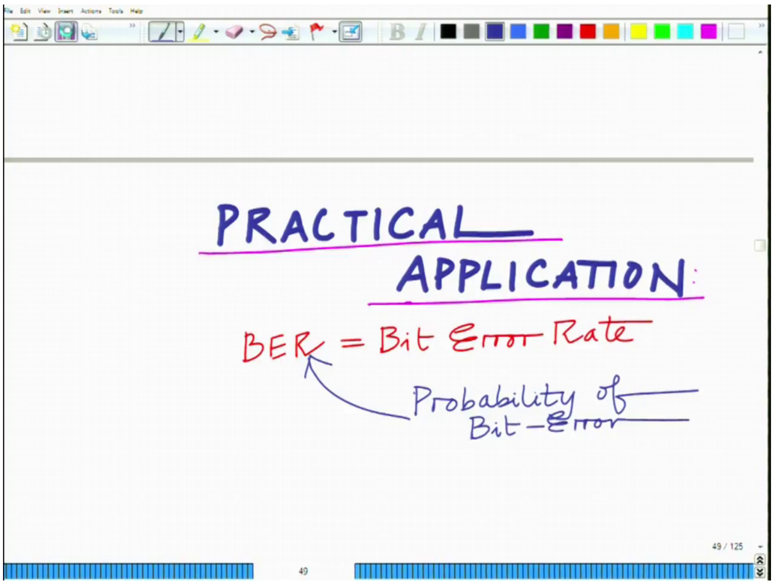
(Refer Slide Time: 11:50)



So, you can consider theta 1 theta 2 theta n as the probabilities and x 1 bar x 2 bar, so on. X 1 bar are the various valuables of the random variable. Then, you have again F of expected value of X equals F of summation i theta i x i bar i equals 1 to n is less than or equal to F of or is less than equal to expected value of again for a convex function ok. Expected value of F of X which is basically summation i equals 1 to n theta i x bar i ok, all right.

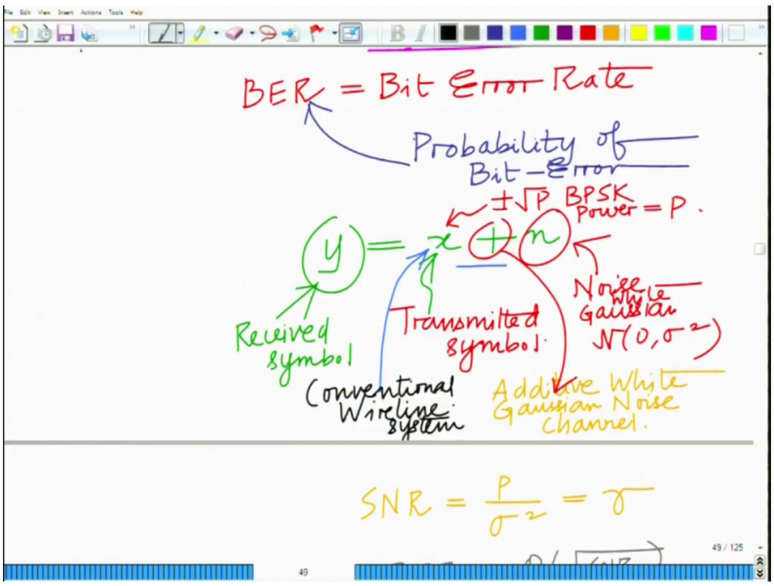
So, this is the Jensen's inequality. In fact, it holds even for a continuous random variable X. In fact, that is what makes this very interesting and very powerful inequality all right. And, in fact, that is what we are going to see shortly in a practical application in the context of a communication system, alright. So, what we want to see now is, we want to see a practical application of this to demonstrate the applicability.

(Refer Slide Time: 13:33)



So, you want to see a practical application ok. And therefore, consider the channel y equals well consider a communication and this practical application is in the context of the BER what is termed as the bit error rate. I think all of you are familiar with this or most of you are familiar, those who work or those who are familiar with the properties of a communication system or the performance analysis of communication systems, this is basically a bit error rate which also is the probability. That is, it denotes the probability with which a bit is received in error over a communication channel, all right.

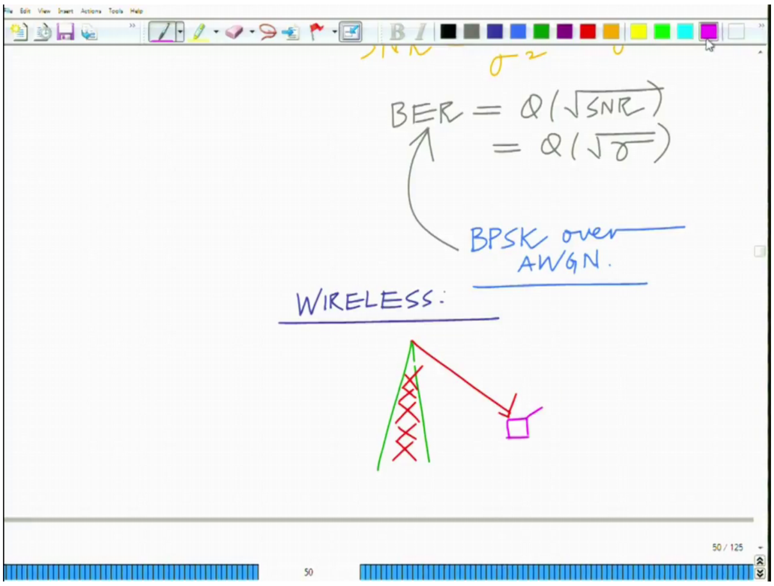
(Refer Slide Time: 14:50)



So, you have consider a very simple channel which we y is the received symbol x is the transmitted symbol and n is the noise correct. So, y this is your received symbol, that is the channel output. This is your transmitted symbol and this is your noise ok, which is typically assumed to be white Gaussian, correct. This is white Gaussian with mean 0 and variance sigma square x can be a symbol typically take this as plus or minus square root of P that is BPSK symbol.

Power BPSK stands for binary phase shift keying power is P. The noise is white Gaussian and it is additive in nature, correct. Therefore, this is termed as an additive white Gaussian noise channel. Therefore, this is termed as an additive white Gaussian noise channel and since this is an additive white Gaussian noise channel.

(Refer Slide Time: 16:36)

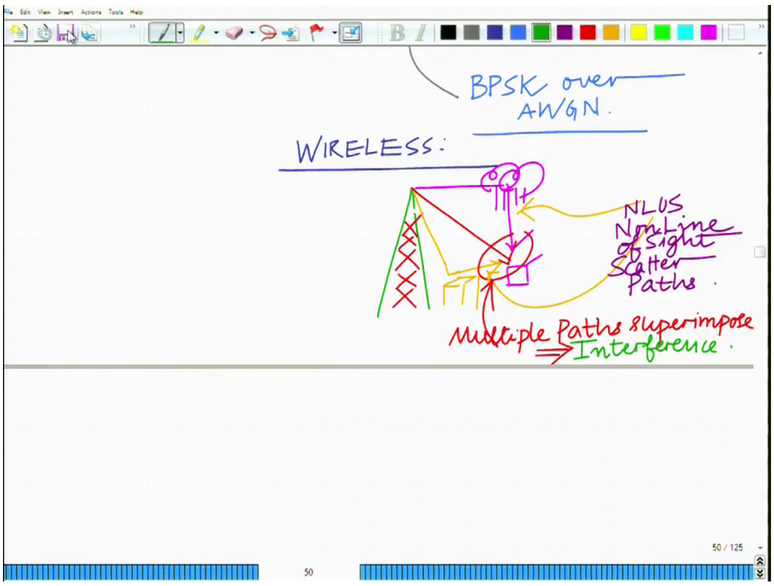


If you are transmitting BPSK symbols power square root of P, then we have the SNR is nothing but the symbol power signal power divided by noise power. That is, sigma square, all right. And we can denote this by gamma. And the interesting thing about this is, if you look at the bit error rate of an additive white BPSK of communication of digital community BPSK or transmit the transmission of BPSK modulated digital symbols, over an additive white Gaussian channel that is given by the well-known expression, that is the bit error rate is given as Q of square root of SNR, that is equal to Q of gamma Q of square root of this is the bit error rate of BPSK over AWGN.

And now, however so now, this model additive white Gaussian noise model is generally associated with a conventional digital communication system in which they, in which there is a wire medium between the transmitter and the receiver. This is also known as a wired channel, all right. A wire based channel such as the twisted copper pair or a coaxial cable and so on it or your conventional telephone where there is a wire that connects the telephone to the local exchange and so on, all right.

So, this is a model for a conventional communication system or a wire line what we call a conventional or basically or this is your conventional or basically your wire line system. Now, on the other hand, what happens in if you look at a wireless system, something interesting happens in a wireless system. In a wireless system, you have a base station which is transmitting. This is your base station which is transmitting to let us say, a mobile in the cell, then in addition to this signal which is you can call this as a direct path.

(Refer Slide Time: 19:14)



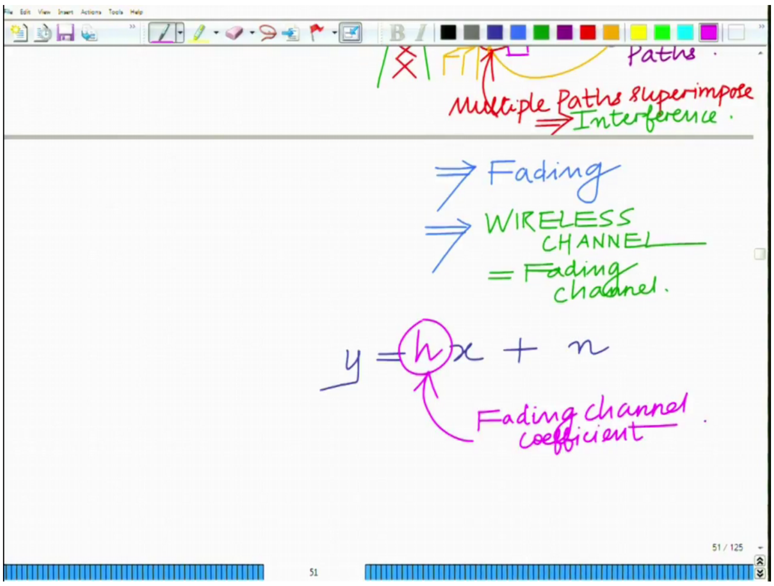
They are also going to be several reflections or what are also known as scattered paths. So, these are also known as the scatters such as large buildings. So, these gives ray give rise to this non-line of sights the NLOS non line of sight scatter. So, there is a line of sight. Now, what happens? When these multiple paths the signals from multiple paths, they superimpose, correct.

So, when these signals superimpose, so, multiple signals from the multiple paths is also known as a multipath environment. Multiple paths superimpose implies this leads to interference, this leads to interference, correct and that is the problem, right. And once you have interference, the signal can be now interference may not be only destructive; interference can also be constructive, alright.

But, the moment you have seen interference, there is uncertainty in the received signal level, all right. The signal level can dip if the interference is destructive or the signal level can rise if the interference is constructive. So, in general, the signal the level of the signal or the power that is the power of the received signal is varying with time, all right. Unlike a conventional wire line communication system where there is no phenomenon of multipath reflection.

In a wireless system, because of this multipath reflection phenomenon, the interference the resulting interference leads to a time varying power for the received signal like this process is termed as a fade, this process is termed as fading, all right and this channel the wireless channel is known as a fading channel, ok.

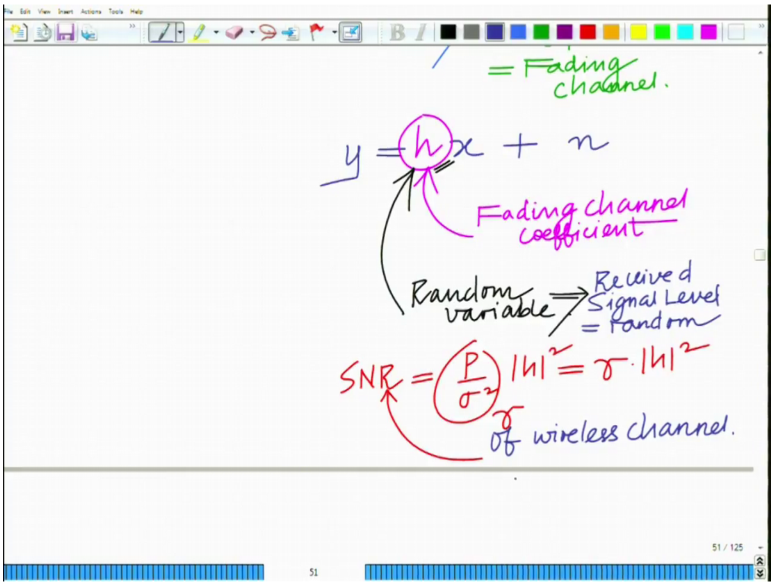
(Refer Slide Time: 21:21)



So, this leads to variation in the received signal which is term as fading implies that the wireless channel is a fading channel as we have seen in some examples before.

Now, therefore, I cannot simply use the AWGN channel model. For the wireless channel, I have to use a different model. So, for instance, they are wire line or a conventional channel model is additive white Gaussian noise that is, your y questions may receive signal y equals transmitter signal x plus noise. In addition, in a wireless channel, I will have the presence of a multiplicative factor h, which is a coefficient this term as the fading channel coefficient. This is term as the fading channel coefficient, ok.

(Refer Slide Time: 22:38)



So, this is termed as a fading channel coefficient and this is a random variable. The important thing to note is, since the received power is random. This fading channel coefficient is a this is a random variable. This implies that, received power is random in nature or received signal level the received signal level is random in nature. And therefore, now you can see the SNR is influenced by this fading channel coefficient.

So, one can write the SNR which was previously P over sigma square will now be multiplied by magnitude h square. So, it is P over sigma square. Remember, this is gamma. So, this is gamma times magnitude h square. So, you can think of this as the SNR of the wireless channel SNR of wireless channel. This is the SNR of the wireless channel.

And, therefore, now what we can do is, we can look at the resulting bit error rate of the bit error rate depending on this SNR of the fading wireless channel and apply Jensen's inequality and derive a suitable conclusion, alright. So, this is the practical scenarios in which amount apply the Jensen’s inequality, alright. And, we will continue this discussion in the subsequent module.

Thank you very much.