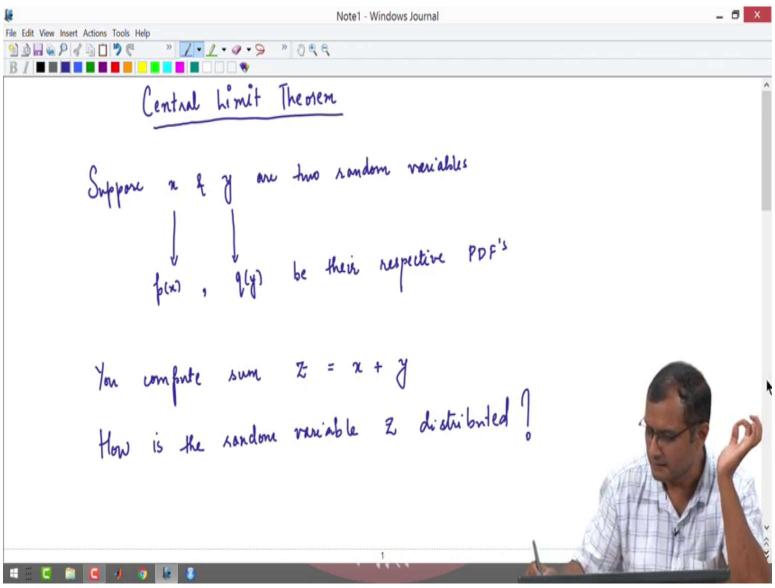
Statistical Mechanics
Prof. Ashwin Joy
Department of Physics
Indian Institute of Technology, Madras

Lecture - 07
Central Limit Theorem

So, good morning to all of you; today we will discuss a very important concept in Statistical Mechanics which is actually used in large areas of you know physics, it has a great range of applications from soil state physics to high energy physics. And the theorem that I am going to discuss at the end of this lecture is called as a Central Limit Theorem.

So, before I formulize central limit theorem, I would like to initiate the discussion by discussing what happens if you add two random variables and want to know how the distribution of the sum behaves like ok. So, we will introduce ourselves to distribution of some of random variables and when the discussion matures to a stage which is appropriate to discuss central limit theorem, I shall alert all of you ok. So, let us state with the discussion on central limit theorem.

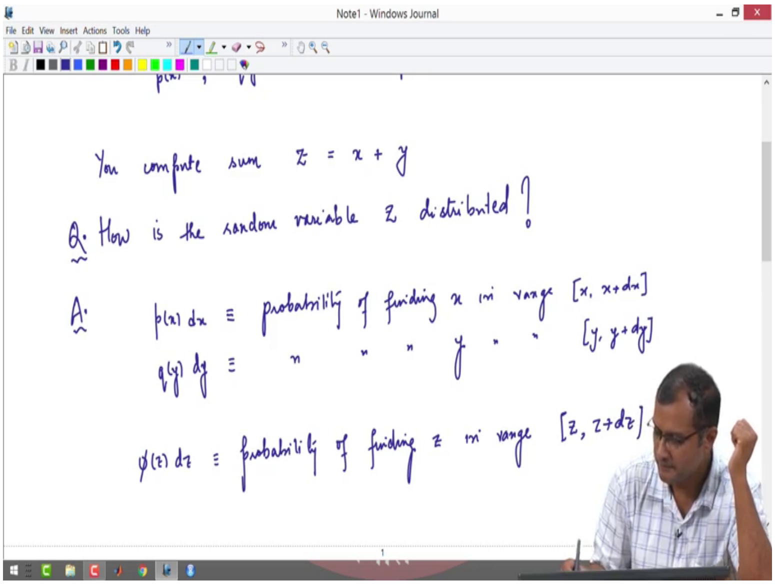
(Refer Slide Time: 01:25)



This is the agenda of today's lecture, a very important theorem in statistical mechanics it has applications at our wide spread. Now suppose you take two random variables x and y are two random variables and the distribution of x is p of x. So, x is distributed as p of x and y is distributed as some q of y ok.

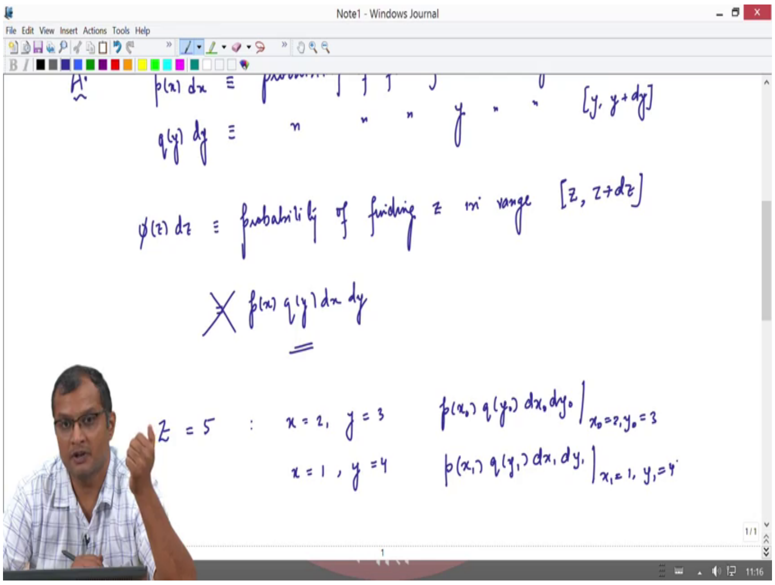
So, p of x and q of y be their respective PDFs, Probability Density Functions and now you compute a sum of this two random variables call the sum as z which is equal to x plus y ok. Simple you take 2 PDFs p x and q of y, you draw random variable from p x call it as x and you draw random variable from distribution q call it y. You add them up call the sum as z and ask a question how or what is the PDF of z, I think the better a question would be how is z distributed. How is the random variable z distributed? Ok, that the question I am going to answer that is my question and the answer is very simple.

(Refer Slide Time: 03:51)



So, recall that p x d x is the probability of finding x in the neighborhood ok, in the neighborhood of a or in the interval x to x plus dx. And, similarly q of y dy is the probability of finding y in range y to y plus dy, we have to be care full here. If I am asking what is the PDF of the random variable z which is basically answer to the question about how is random variable z distributed, where it is distributed as its PDF, the answer is let us say phi of z is the PDF of z. Then phi of dz is the probability of finding z in range z to z plus dz ok, which should tell you that you have to be care full here that in the since that you cannot say phi of z dz is p of x q of y into dx dy.

(Refer Slide Time: 05:41)

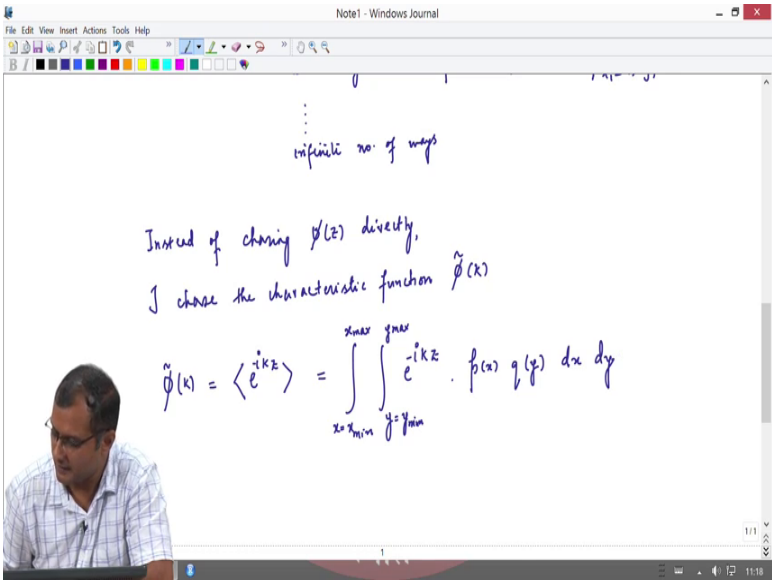


You are attempted to say that since p x d x is the probability of getting x in the interval x plus d x q y d y the certainly the probability of getting y in the range y and y plus dy. So, phi is the dz where d z is a some of x plus y is given as dx dx into dy dz, no that is not true this is not the probability of getting z in the neighborhood of z and z plus dz. For the simple reason is that while this might be the probability of getting joint probability of getting x and y in the neighborhood of x and y, I can compute the same sum by changing the values of x and y. For example, if you take z as 5 ok, this can be done by taking x as 2 and y as 3 ok.

Now, also be done, it can also be obtain by taking x as 1 and y as 4 ok. So, if I take p of p which is a distribution of x p of 2 and q of 3 into you know, I will take it right this as a. Now this would give a probability, joint probability of getting x in the neighborhood of 2 and y in the neighborhood of 3, which will give me the sum as 5 in the neighborhood of 5. But if I consider this p at sum x 1 q at y 1 d at x 1 and d at y 1 where x 1 is equal to 1 and y 1 equals to 4. Well, this will also give me a probability of getting z in the neighborhood of 5 because 1 plus 4 is also 5. So, certainly what I have written here is not right ok.

So, that is why I cannot just write down the probability of getting the sum in the neighborhood of z as p x q y into dx dy ok.

(Refer Slide Time: 08:39)

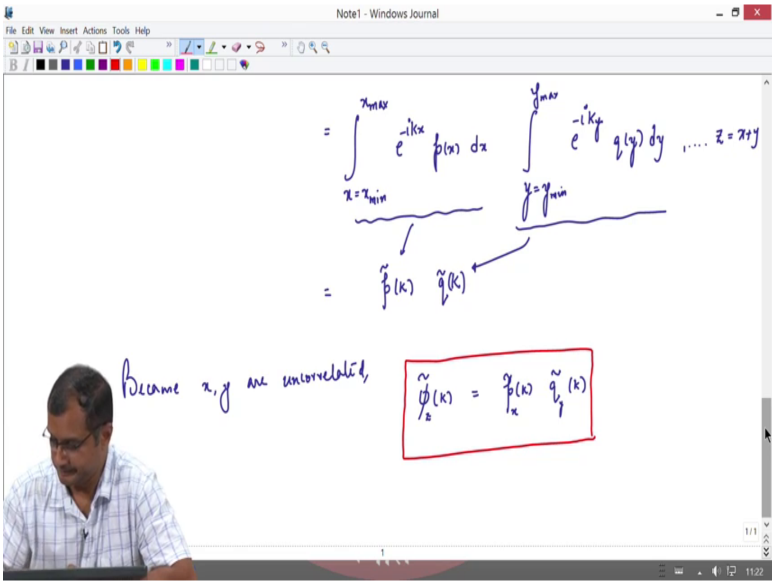


So, this problem has to as a infinite number of ways in which I can do this. So, there infinite number of ways as x and y are continuity, are continues variables ok. So, what I do know is that whatever probability is take the sum of x plus y constraint to be z ok. So, how about dealing this problem slightly differently? Suppose instead of a chasing the PDF directly say one the PDF that is the goal of my exercise ok. I chase the characteristic function so I am changing my strategy, why? It will be clear in a short while ok.

So, let me find out the characteristic function. Now I know the characteristic function of the z distribution is nothing, but the average value of e to the power minus ik z it is the definition of the characteristic function and this is nothing, but the average value of the complex expansion and then naturally I have to take the complex exponential in this distribution of x and y.

So, I will take this as x going from x minimum to x maximum and y going from y minimum to y maximum; p of x q of y dx dy fine. Because this is the distribution join distribution x and y, I am simply weighing it is to minus ik z in that it is like the two dimensional Fourier transform of p x and q y is nothing, but two dimension Fourier transform.

(Refer Slide Time: 11:05)



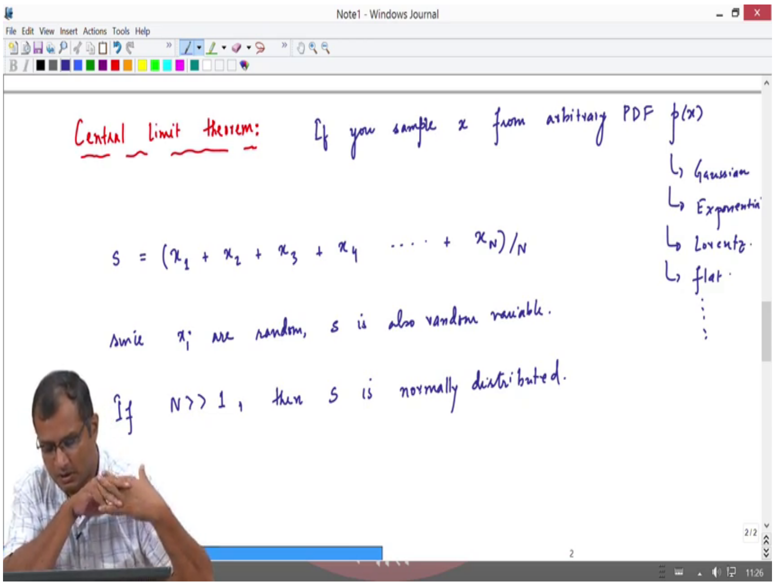
Now, I can write this as a e to the power. So, I can separate out this integrals and write this as y going from y minimum to y maximum e to the power minus ik y q of y dy. I can do that because, the random variable z was x plus I x plus y as I have just separated the two degrees of freedom in the 2 separate integrals.

Now, x and y are uncorrelated. So, the separation was possible and if you want visualize the two integrals their individual is nothing, but characteristics function of the x distribution and y distribution themselves ok. So, this was nothing, but this entire integral is nothing, but characteristics function of the p distribution and this guy is are nothing, but characteristics function of the q distribution. Look at beautiful result that we obtained because x and y were completely uncorrelated, the PDF of the sum has a characteristics distribution which is nothing, but the product of individual characteristics distribution.

So, the PDF of the characteristics distribution of, I will put a subscript here as a z it tells you that this is the characteristics function of the z distribution. This came out to be simple product of the individual characteristics functions, the subscript x y and z are only there to alert you that they where the, they where the characteristics functions of the respective random variable distributions ok.

So, this is the very important result and I am going to use this to derive the central limit theorem ok. So, this is very important result. So now, we are in the position to formally derive, formally state and derive the central limit theorem ok. So, I am going to basically make the definition here or state the central limit theorem here.

(Refer Slide Time: 14:09)



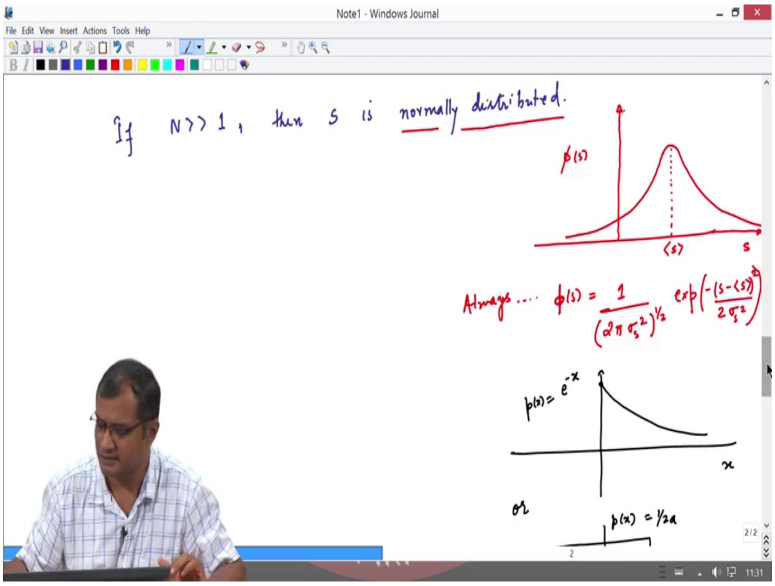
So, central limit theorem is a beautiful theorem and we will come back to it many many times in this course. So, I think the reason why I want to discuss this theorem and give it sufficient amount time that it truly deserves, because we will be coming back to it in the next few chapters repeatedly ok.

Now, the theorem is very simple and very powerful, it states that if you take let us see you sample some random variable x from an arbitrary PDF ok. So, you sample likes from a arbitrary PDF, this PDF could be anything, it could be could be Gaussian, it could be exponential, it could be Lorentz distribution, it could be a flat distribution and so on. You could take it to be any distribution observed in nature and you sample the random variable from this distribution.

Let us say you take a N variables x 1 x 2 x 3 x 4 all the way to x N you take N random variables from this distribution p of x and you compute the sum and take the mean of this sum and call it a new variable S ok. So, since these all of them where random in nature x 1 to x i ok, S is also random, the sum of random variable is also a random variable central limit theorem states that if N is very large the sum of this random variables is normally distributed, ok.

Normally means it is distributed as a Gaussian distribution, normal distribution and Gaussian distribution are the same. The Gaussian distribution is a general distribution with an arbitrary mean and an arbitrary variance, if you take the mean to be 0 and variance to be 1, you can call it as a normal distribution or the standard normal distribution. So, I would like to rephrase this so restate this important theorem.

(Refer Slide Time: 17:55)



That if you take a large number of random variables from any distribution does not matter whether it is Gaussian or exponential or a Lorentz or flat, you add them up the sum if you have taken large number of such variables will be always normally distributed. Which means what? Which mean the following that if you compute the PDF of S, this will be always a Gaussian distribution.

So, this PDF would be ok. So, I am going to write down the Gaussian distribution explicitly. So, p of S is nothing, but we shall see at the end of this discussion, now it is 1 up on 2 pi sigma S square e to the power minus S minus ok. So, that is the beautiful result. So, this is the p of S as I have plotted, your p of x could be let us call this distribution S is phi of S. So, that we do not use this same some going to call this distribution of S as phi of S because p is the distribution taken for x ok.

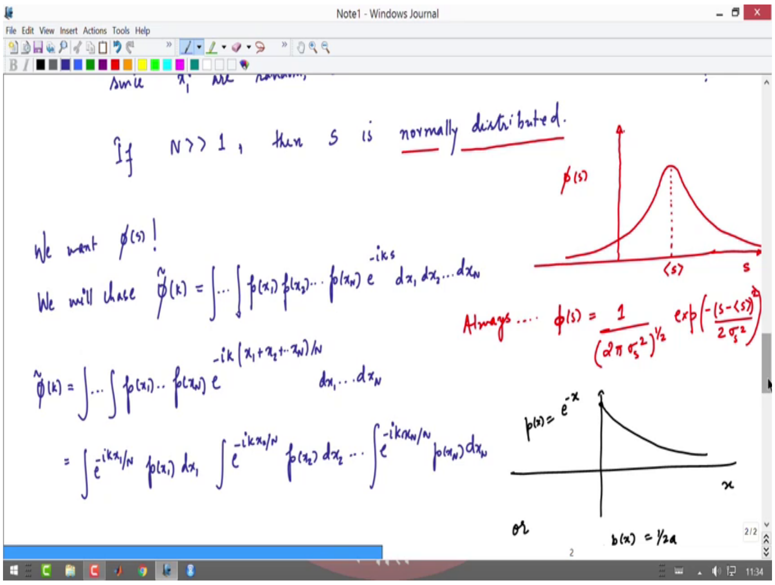
So, this to so phi of S which is the distribution of S will be always normally distributed, no matter what is your p x ok. So, p x could be anything. So, let me also plot p of x, you can take any p of x ok. So, you could take p of x as an exponential distribution if you like or you could take p of x as flat distribution does not matter, does not matter you can any p of x your phi of S will be always normally distributed ok. Beautiful result, it does not care which PDF you take as long you take a large number of them the sum will be normally distributed.

Now, this theorem has large number of applications in solid states physics and as I said in many areas physics whether it is fluid mechanics, turbulence, high energy physics you have central limit theorem only present. And this is often the reason why you see co lactic behavior in systems at low temperatures, where you see long wave length prohibitions which are excited. These are the whole mark of all low temperature prohibitions, low temperate prohibitions and low frequency in nature which means you have long wave lengths. Which means in each long wave length you may have a large number of particles that are assaulting or having an randomly having random motions. These particles could be on a lathers and because this is vibrations are themselves p dimensional random variables in a long wave length that are many of them.

So, the fact that this are large number of random variables will always be normally in some of this random variables will be always normal distributed, can be attributed to why you see low temperature universal scaling lows in such as heat capacity or any transport coefficient that I seen to universally behave at low temperatures. By universal I mean it does not depend on the material property, whether it is iron or aluminum or nickel as long as a temperatures are sufficiently low their transport coefficients will be universally behaving with some scaling law.

So, in some since the implications of central limit theorem are very deep. So, let us formula is all that with some very simple mathematics. So, what I am going to do now is that define this random variable, it is already defined here. So, let us call this as equation 1 ok. So, if I want to write down a characteristic function of this random variable distribution.

(Refer Slide Time: 23:13)

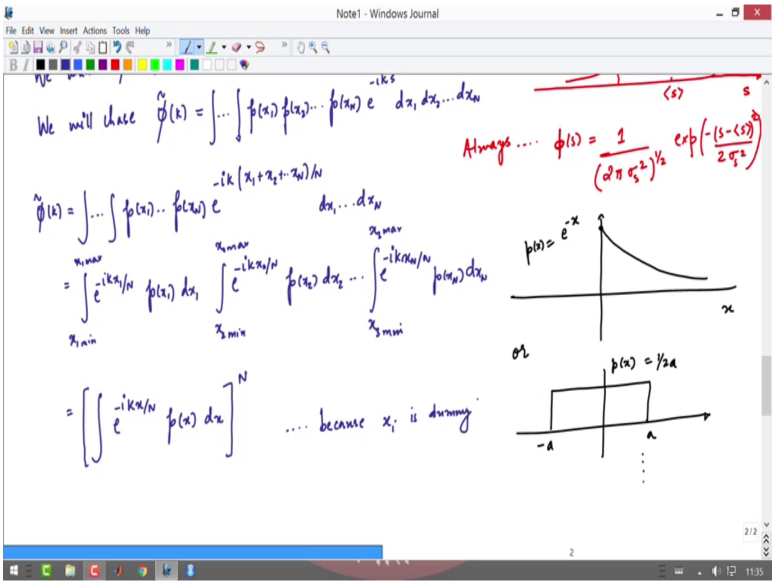


So, remember what we want is the following, we want phi of z, phi of S our sum is know S we want phi of S ok, but we will start with the characteristics function right. So, now, the characteristics function can be written as these N integrals, I am not writing the limits here because the limits are understood, they are from the low value to the x lower to the x maximum.

So, I am not writing it, but you should understand that this integrals are all different integrals and you will say p of x 1 into p of x 2 up to p of x N d of e to the power minus ik s d of x 1, d of x 2 all the way d of x N ok. And this can be written as phi of k as all this an integrals p of x 1 up to p of x N and I can write down this variable S as x 1 plus x 2 all the way to the x N divided by N into d x 1 all the way to d x N ok.

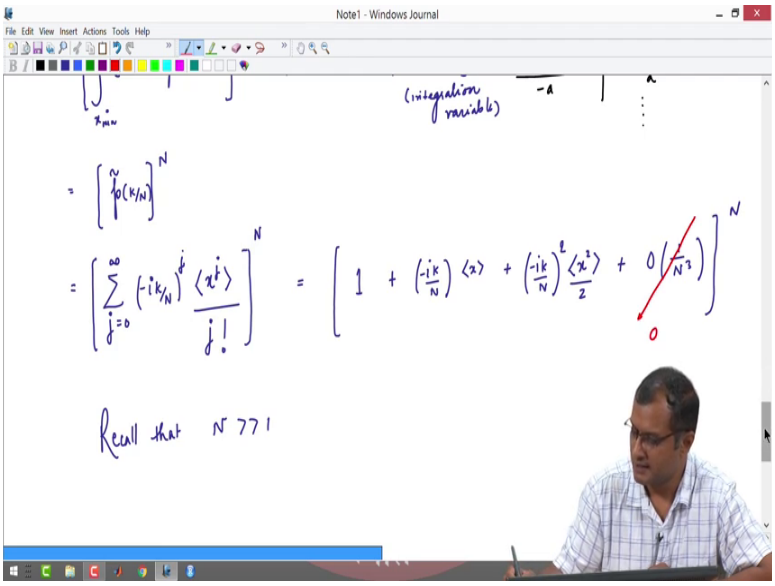
And since these variables are uncorrelated I can split up this exponentials and write them as e to the power minus ik x 1 p of x 1 into e raised minus ikx 2 p of x 2 dx 2 all the way to e raised to minus ik xn p of x N into d of x N ok. So, if you look at these definite integrals, each one of them is from x 1 minimum to x 1 maximum ok.

(Refer Slide Time: 26:17)



Similar this is from x 2 minimum to x 1 maximum and this is also from x 3 minimum to x 3 maximum, if you look at this definite integrals there all the same because the integration variable is dummy ok, the integration variable in a definite integral is dummy. So, this is like integral e raised to minus ikx by N into p of x dx whole raise to power N because their N such integrals ok, because the integration variable is dummy, fine.

(Refer Slide Time: 27:19)



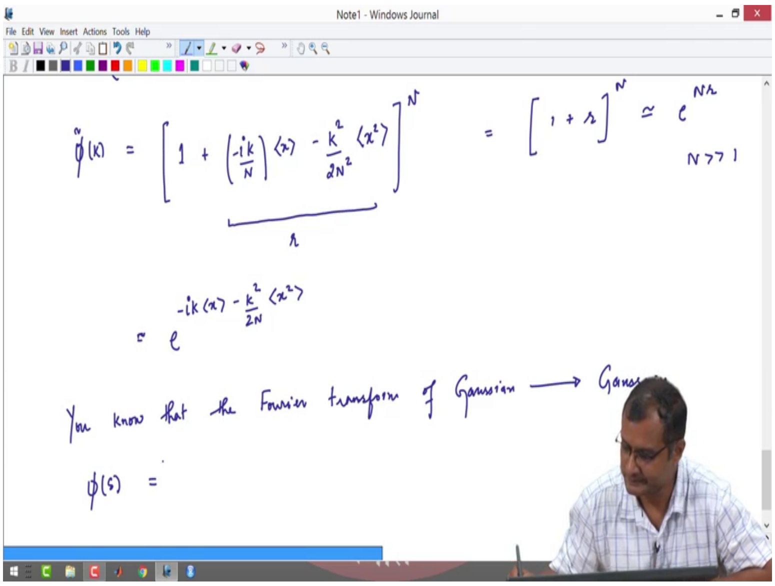
Now, look at this, integral is nothing but to remind immediately that this integral is nothing, but the definition of the characteristics function of my PDF except for that fact that this is the characteristics function evaluated at d wave number k by N. So, this is the characteristics function of my PDF they can at k by N and the entire thing is to the raised the power N ok.

Now, if you recall the definition of the characteristics function from the, from the past few lectures it is nothing, but a summation sum j going from 0 to infinity minus i k by N to the power of j jth moment divided by jth factorial ok. That is the characteristics function for the wave number k by N and the entire thing is raised to the power N ok.

Now, you can write down this summation for a for a few values ok. So, you can write it for a few values as for let us say for you write it j is equal to 0, 1 and 2 for j equal to 0 you simply have 1, for j equal to 1 you have minus of ik by N mod x. And, for j equal to 2 you will have the second moment by 2 plus the third term that you have is basically of the order 1 by N cube, the entire term entire thing is raised to the power N ok.

So, now remember, recall that our assumption was we have taken N to be much larger that was the assumption ok. We are going to derive central limit theorem for very large in if that is the case I can drop the terms of the order N cube 1 up on N cube in comparison to 1 by N square in 1 by N.

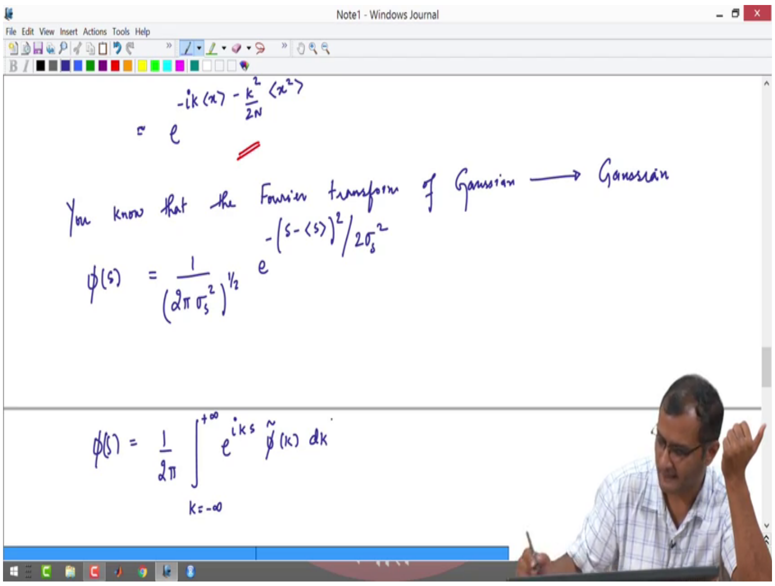
(Refer Slide Time: 30:27)



So, when you drop this you get the characteristics function as 1 plus minus i k by N into mod x and the second term I will write it as minus k square by 2 N square into mode of x square hole raise to N fine. Now you visualize this some of this two terms as some variable r. So, what you have is basically so this basically 1 plus r to the power N ok, what is the visualize the some of this two terms as a r, now on the expansion of 1 plus r to the power N is just e to the power N r if N is much larger than 1 ok.

So, this is an approximation if 1 plus r to the power N is raised to N r now simply put r as what I have written here. So, this becomes proximately e raised to N times r would be just minus i k into mod x minus k square by 2 N into mod x square, this multiply N to r and based to the power e ok. So, this is our characteristics function, now you what you looking here is just a Gaussian characteristics function, you know that that Fourier transform of a Gaussian is also Gaussian, is also Gaussian. So, in some sense our theorem is already proved, if I want to compute the PDF phi of S this will also be a Gaussian distribution, it will also be a Gaussian distribution.

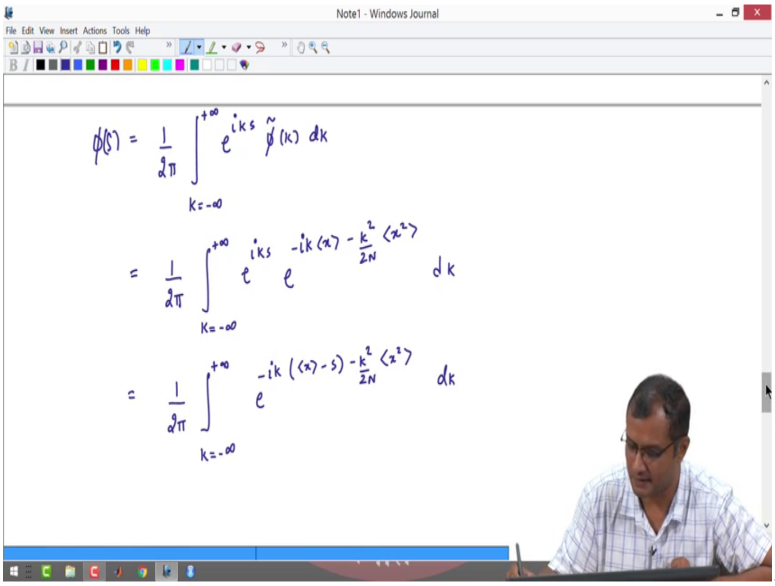
(Refer Slide Time: 33:09)



So, my theorem stands proved as S, but he want to prove if you are the guy who is not convinced, if that you can match your expressions in the discussion on Gaussian distribution we got prissily this expression for the Fourier transform the Gaussian distribution. So, you know that the inverse Fourier transform this would be a Gaussian distribution, but if you want to see explicitly how this is possible while you can do that also ok.

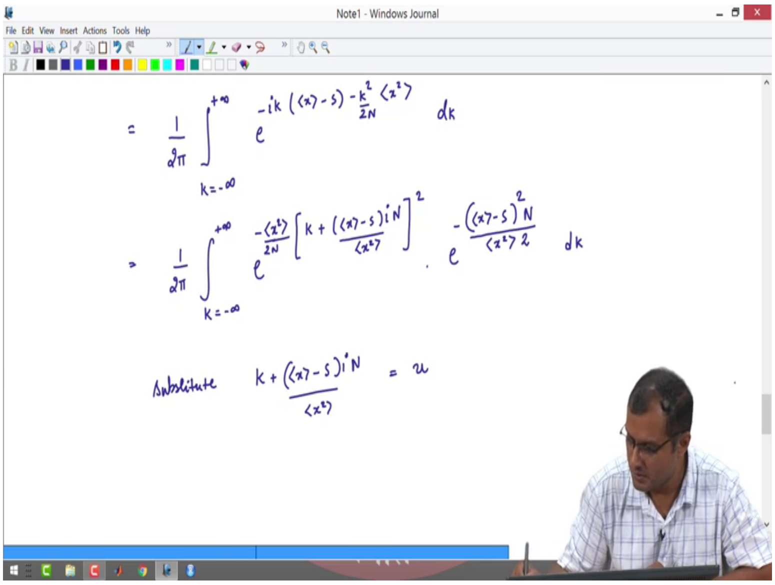
So, explicitly you can show that this is phi of S. So, remember I wrote down this expression without solving, but you can also get it by doing a Fourier transform. So, let us do that. So, you can write down phi of S as the inverse Fourier transform of phi of k re inverse Fourier transform is nothing, but k going for minus infinity to plus infinity e to the power iks phi of k N dk if you get that.

(Refer Slide Time: 34:35)



So, let us do it. So, you have a 1 up on 2 pi k going from minus infinity to plus infinity e to the power i k s and phi of k is just derived here I am going to just write it here I just copy it form here dk ok, just need some manipulations here which is very simple. So, you write this again as minus infinity to plus infinity and you can see that this is easily written as e raise to minus i k into mode x minus S I have taken the first exponential as this is 2 minus k square by 2 N to mod x square d of k.

(Refer Slide Time: 36:01)

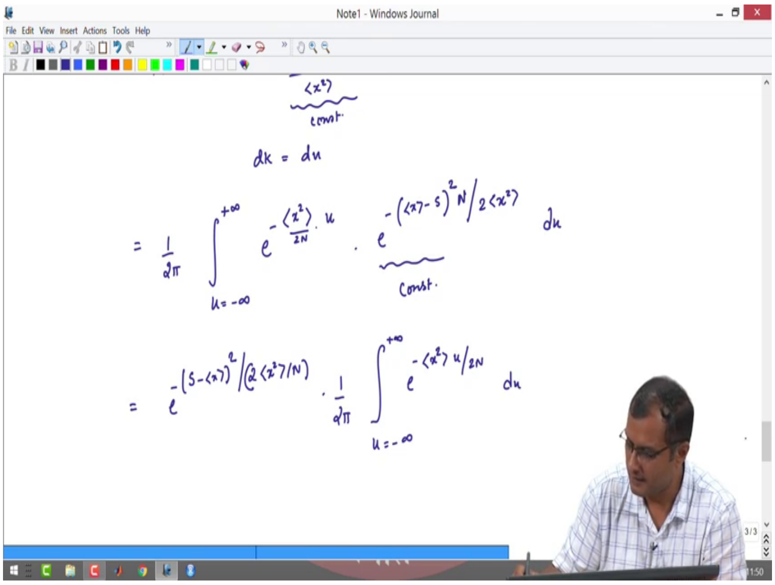


And, these are very simple integrals. So, you can perform a, you can construct the perfect square in the in the numerator that is always a trick to solve this Gaussian integrals. So, the perfect square that I am going to construct is on k square. So, I am going to take the pre factor here as minus mod x square over 2 N and I am going to write it as k minus. So, the k square term will give me second term here, twice of ik with a minus sign becomes minus 2 by k and have a 2 here outside to get canceled off, but I need a N here because I have to cancel by N here.

So, the 2 times of this would give me just ik into mod x minus S with the minus sign becomes and o also need a division by mod x square. So, that cancels with the outside mod x square and to cancel with 2 and cancel with N I did this ok. Now, there is a aquatic term which will come extra and that would be just entire thing square, but as a i square that cancels of with minus 1 i square so I get a plus sign. So, here is a minus sign mod x minus S the whole square N square by N is N and I will have divided by just mod x square by 2 ok.

So, the cordatic term would be N squares cancels with N that gives me N here, i square becomes minus 1 and I will have an x square denominator, so that is it ok. Also you can check this, I have a computed this as the factorization and now you can see you can substitute k plus mod x minus S into i, this is a constant ok. So, this is so this is basically ok so, this is my variable S k. So, the entire thing here is substitute as sum u.

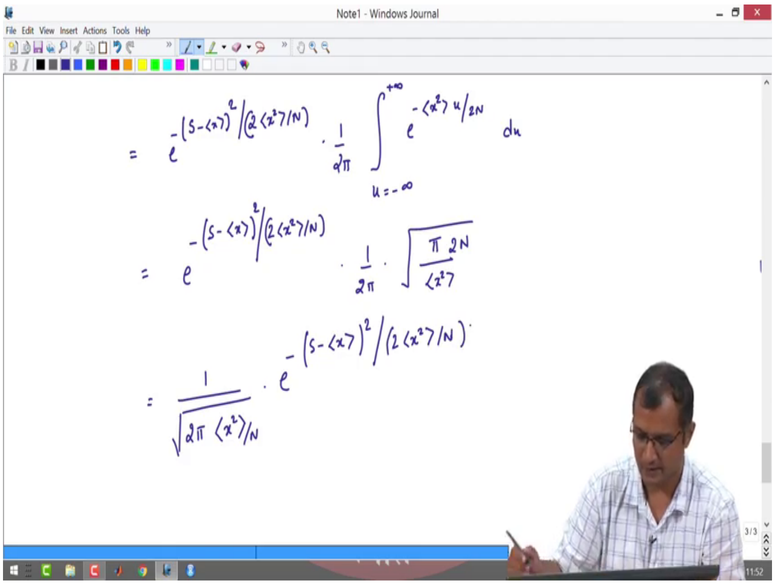
(Refer Slide Time: 40:09)



So, d of k will be just d of u because the entire thing is just a constant is just a constant ok. So, my integration above will become 1 upon 2 pi into u going from minus infinity to plus infinity e raised to minus mod x square by 2 N into u into e raised to minus mod x minus S whole square by N divided by 2 into mod x square d k. Now this exactly is a constant ok. So, for this integral this is also a constant it does not depend on k. So, it will come out of this.

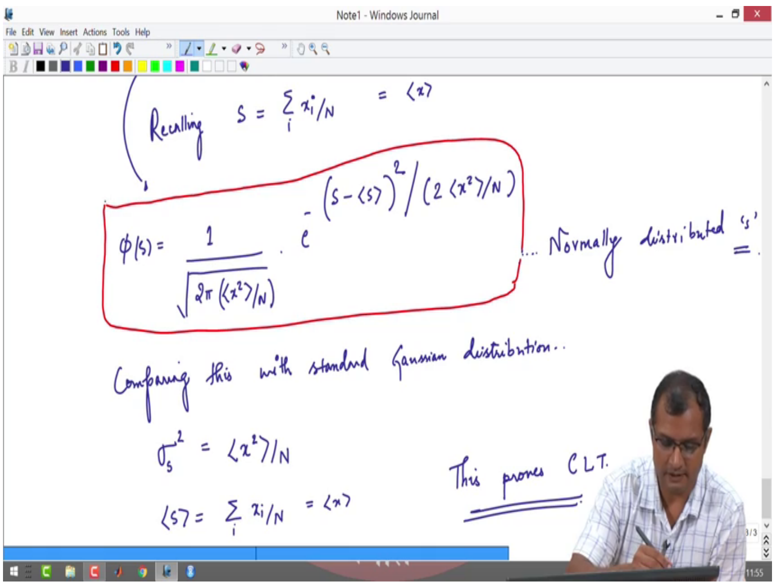
So, what I will have is basically e raised to minus and I can write down mod x minis S the whole square as S minus mod x the whole square of a negative number is same as a square of the positive number ok. So, I can write this as S minus mod x the whole square divided by 2 into x square by N fine and this is multiplied by 1 upon 2 pi and the integral is going from minis infinity to plus infinity e raise to minus x square into u by 2 N d of u. This is a d u this is not d k this is d of u. We have converted the k integral into u and you know the Gaussian integral which is square of pi by a.

(Refer Slide Time: 42:25)



So, you write this as a e raised to minus S minus mod of x, the whole square divided by 2 times into 1 upon 2 by it is square root of pi by a, now a here is x square by 2 N. So, this would be just x square by 2 N ok. So, this would just be if I write it as a 1 upon square root of 2 pi into e raised to minus S minus mod x the whole square divided by 2 times x square by N ok. And if you see the way we have defined our variable S is nothing, but all the summations of x divided by N.

(Refer Slide Time: 44:07)



So, if you recall that S itself was summation of x i by N. So, if I take a modern both sides ok. So this is nothing, but this already nothing, but mod of x ok, but the average value of x because summation x i minus mod of x. So, I can rewrite this as phi of S as 1 upon square root of 2 pi x square mod divided by N onto e raised to minus S minus mod S the whole square divided by 2 times x square mod by N ok.

Now, if you compare this with the definition of standard Gaussian distribution, comparing this with you know standard Gaussian distribution or standard normal distribution you can see that the variance of a of a the S distribution is nothing, but the second moment of the random variables that withdrawn from p of x divided by N. And the mean of this was just nothing, but the mean of the variables themselves ok. So, this proves your result that the distribution itself is normally distributed. So, the take a message here is when I look at this expression that I am going to sort of box it is that is that my distribution of S is also normally distributed ok.

So, this is the take a message of our discussion. So, S is normally distributed ok. So, this is the end of, this proves central limit theorem and that is the, that is the proof of a very important distribution that we started off in the beginning of this lecture ok. So, we end the discussion here. In the next class we will talk about many random variables and the concept of conditional.