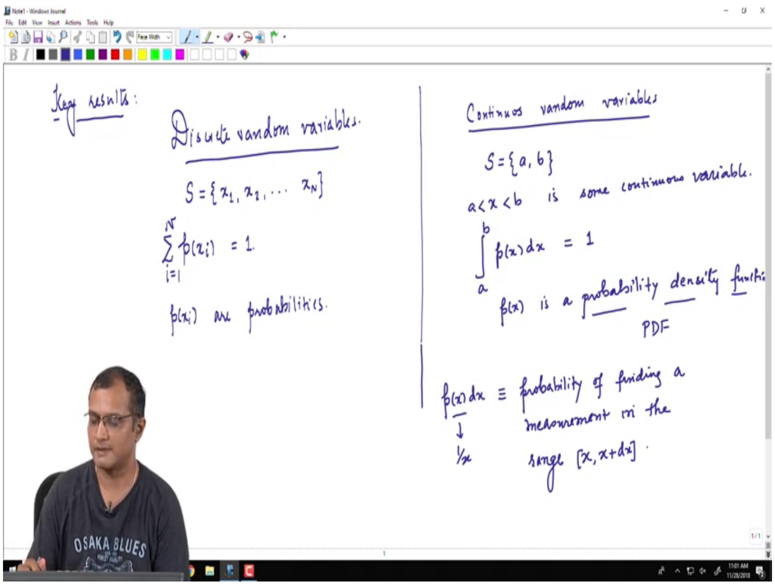
Statistical Mechanics
Prof. Ashwin Joy
Department of Physics
Indian Institute of Technology, Madras

Lecture - 03
Characteristic Functions

Good morning. So, today we will extend yesterday's discussion on continuous random variables. So, just to recall what we discussed in the yesterday's class I will bring some key results before proceeding further ok.

(Refer Slide Time: 00:37)

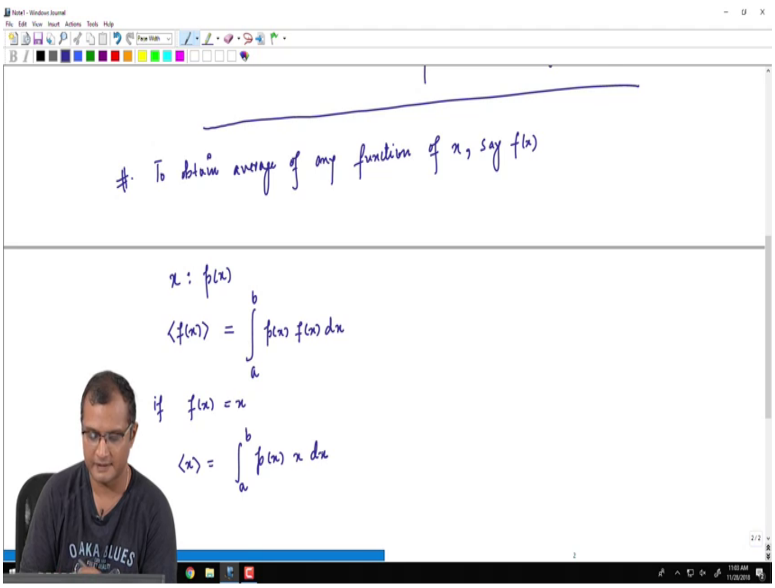


So, some of the key results of yesterday's discussion was if you have the case of discrete random variables then you basically have a set of outcomes which are discrete. So, the outcomes could be x 1 x 2 all the way to x N and the probability of each outcome if it is a sumed over all possible outcomes is normalized unity ok.

So, this is the normalization condition and this p of x i are basically probabilities ok. In the context of continuous random variables, we saw that the set of outcomes are not discrete numbers they could be let us say between some, you know it could be a bounded interval. So, our random variable would be between a and b ok, a is some continuous variable. And, in that case to say that the probability is a properly normalized the equation that I have written on the left becomes an integral, where the lower limit is the lower limit of random variable a, the upper limit is the upper limit of the interval it is b p of x d x and this is 1 ok. Here p of x is not probability, but is a probability density function, this is the difference this is now a dimensional quantity we are also calling this as PDF’s ok.

So, this is a dimensional quantity, you have a dimensions of 1 upon x and the notion of p of x is basically in the following, from probability density you can derive probability as you know you can call p of x dx as the probability. Now this is a probability, this is dimensionless because p x already has dimension of you know 1 by x. So, p x dx is probability which is dimensionless, probability of finding a measurement in the range x and x plus dx that is the interpretation of p x dx ok, p x dx is the probability not p x itself. So, this is the important distinction that we made in the context of continuous random variable and discrete random variables ok.

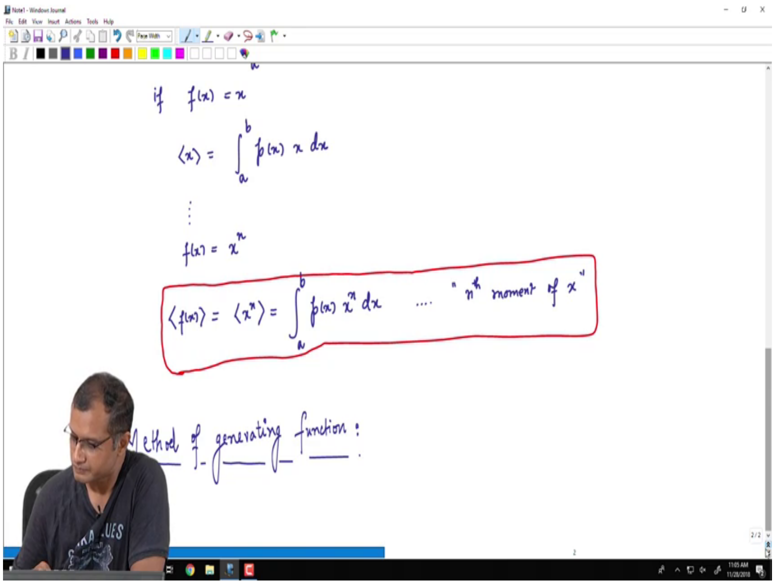
(Refer Slide Time: 04:59)



Now we can use this to create, to obtain average of any function of x for example, if you interested in average of x square or average of x cube or average of any function of x ok, let us say f of x. Now x is a random variable which is distributed with the PDF p of x and I want to know: what is the average of sum function of x ok. So, this average of sum function of x is given as the integral of p x f x dx ok, is the definition of average function of x.

So, if you think for example, f of x as x itself, this gives average of x as integral a to b p of x x dx and if you take in the following way f of x as some power of x.

(Refer Slide Time: 06:21)

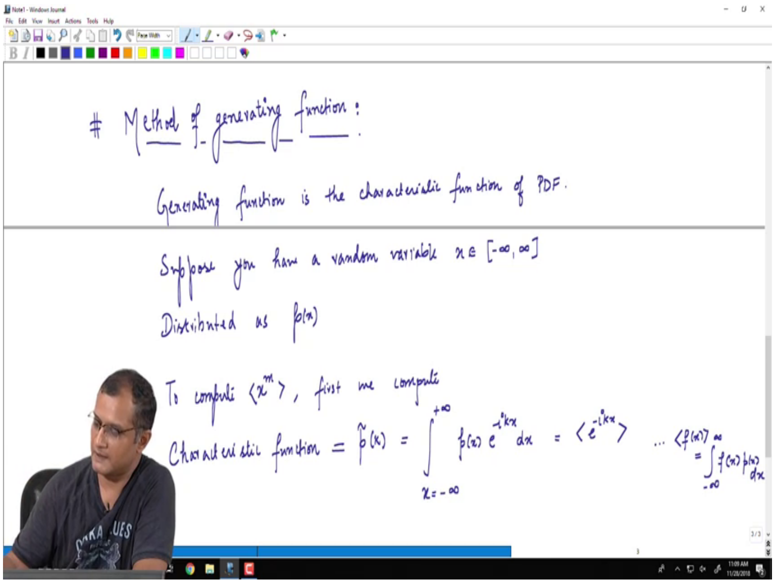


Let us say x to the power n then the average of f of x would simply be the average of x to the power of n which by the definition of averages is integral a to b p of x x to the power n dx. And, the language of statistics this is called as the n th moment of x ok. So, is like the average of the nth power of x called as nth moment of x fine.

So, today we will discuss some important mathematical laws that allow you to derive these moments with different route, sometimes the method of obtaining these averages by Brute force integration may not be feasible ok. We will give you some examples where the technique that will going to develop would be would make life much easier for you if you want compute the nth moment ok. So, this is one way to compute the average or the average of the n th power which is also call as the nth moment of a random variable.

So, let me record it and we will compare this with the ok. So, this is the formula for the nth moment and today we will discuss an alternative formula, one that is more powerful through the method of generating functions. So, I will discuss the method of generating functions and I am going to discuss method of generating functions to compute both the moments and cumulates ok. So, I will describe what are cumulates in a short while ok. So, the generating function of moments is nothing, but the characteristic polynomial ok. So, this is so the generating function.

(Refer Slide Time: 09:07)



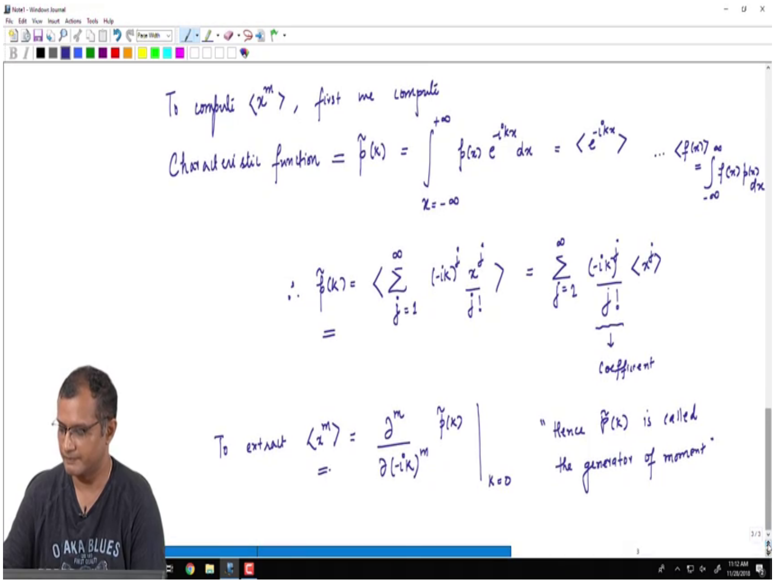
Now, this is called the generating function because it will allow you to generate moments of a random variable is the characteristic function also called as a characteristic function of the PDF whatever PDF ok. So, suppose you start with I will take an example, suppose you are you have a random variable x and this is distributed as PDF, this is a random variable is p of x ok. So, you have a random variable x and it is distributed as p of x that is the distribution of this random variable.

Now, to compute mth moment we first compute the characteristic function ok. So, first we will compute the characteristic function which is nothing, but the Fourier transform of the PDF ok. So, this is the definition of the characteristic function I will call it as Fourier transform of the PDF. So, our p of x is transformed to p tilde of k, where tilde here this denotes as this is a Fourier transform an amplitude of original function p of x.

So, you can take the Fourier transforms as following. So, let us say our random variable x is distributed between minus infinity to plus infinity ok, that is the range of my hand very well and so I will going I am going to take a Fourier transform of this random variable. So, this would be e to the power minus i k x dx ok. Now if you look at the definition of average of a function this is nothing, but the average value of e raise to minus I kx ok. Since we have the we are using definition that a functions average is given as integral minus infinity to plus infinity function into px dx ok.

So, this characteristic function is nothing, but the average value of the complex exponential e to the power minus i kx.

(Refer Slide Time: 12:19)

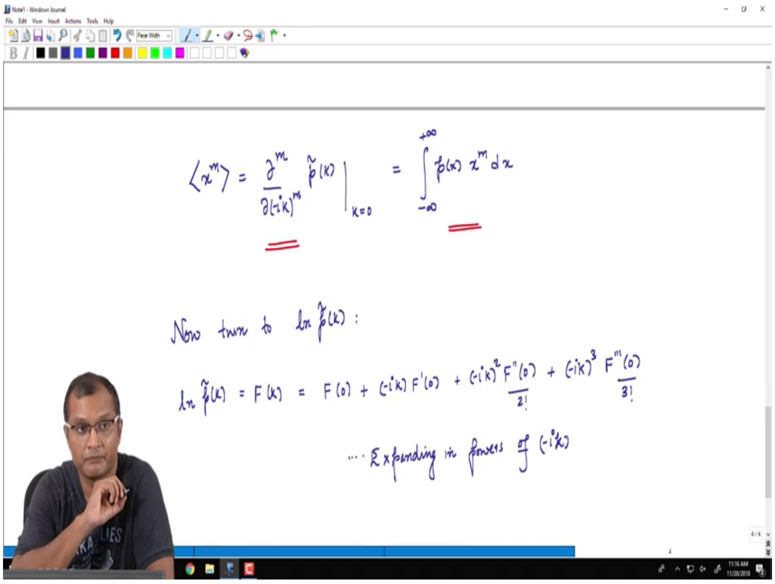


Therefore, I can write down p of x as summation i going from, let me change the variable here summation i going from, j going from 1 to infinity e to the power minus i k to the power j x to the power j over j factorial ok. And you can clearly see that this is nothing, but the averages are simply going to walk over the constants and what you get here is nothing, but summation j going from 1 to infinity minus ik to the power j upon j factorial into the average of x to the power j.

So, now you can clearly see that this is the co efficient in the expansion of the characteristic function p of k. And so, if I want to extract the mth moment from this power series, the extraction would require just taking the mth derivative of my characteristic function with respect to my k and then evaluating this derivative at k is equal to 0, it should precisely give me the this would eliminate the coefficient and what give me is basically the mth moment of the random variable.

So, now we can see that this method why, now we can see that this is called as generator of you know moments. So, this p of k is called generator of moments because the left hand side is a moment and by taking the mth derivative of this moment generating function I can obtain the mth moment, which is this and that is the reason why this is called as the moment generating function or characteristic function right

(Refer Slide Time: 15:29)



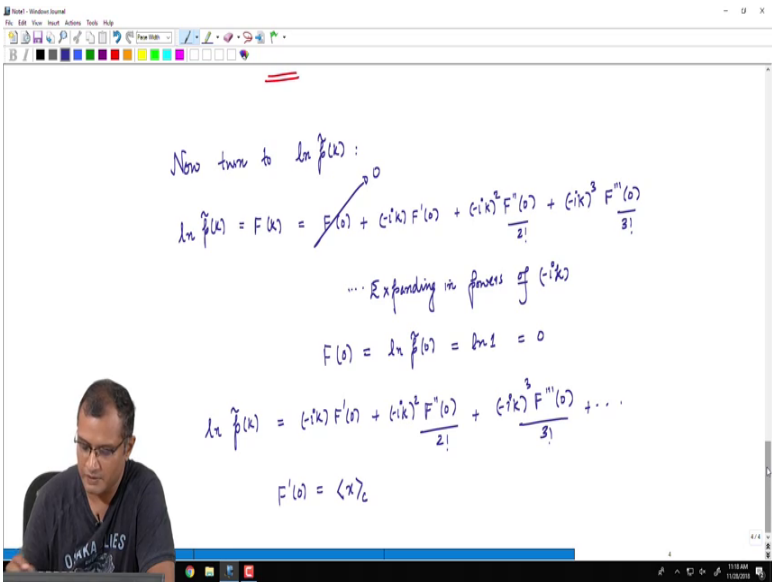
So, this is a if you compile everything that we have now is basically what you have just derive is basically the formula that the mth moment of a random variable, is the mth derivative of the characteristic function at k equals to 0. And, this is equal to the Brute force integration of weighed integration of our mth power of the random variable. So, whether you can use, whether you want to use the definition which is on the extreme right inside or you want to use the definition that we have description that you just derived, it is up to you it depends on the problem at hand may be in some problems the brute force integration is easier.

But in some problems where integration becomes difficult whatever n becomes large you would want to take derivatives of the generating function because in derivative successive derivative easier definitely compared to taking you know computing integration by parts if m is large. So, these are the benefits of characteristic function over a brute force integration.

Now, one can also generate cumulants ok. So, cumulants are important objects and they defined as basically they are connected to with there connected with the correlations in random variables and I will I will just give a definition of defining logarithm of the characteristic function as the generator. I think it is better to first write down logarithm of p of k and then we will introduce the definition, it is perhaps better that way ok. So, you can now you know turn your attention to logarithm of p of k and I am going to tell you that this is also an important very important object and it will become clear in the moment why it is important ok.

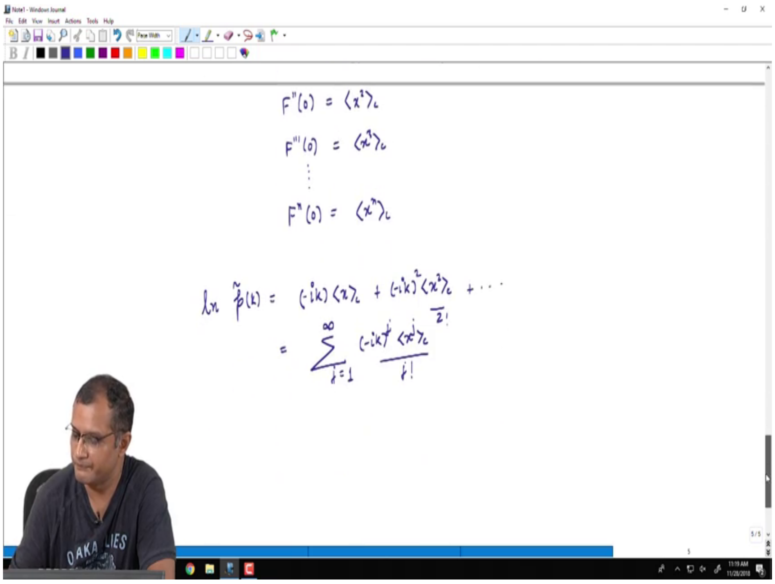
So, so far we have found a formula prescription to generate moments from characteristic function, let us see what is logarithm of a characteristic function ln p of k. So, let us denote this logarithm p of k as sum function of k, we know it is a function of k because p of k is the function of k. Now if I want to expand this power series ok, I can write it as F of 0 that is a leading term now expanding around origin plus minus i k. So, I am expanding in powers of minus i k i k times F prime is 0 plus minus i k to the power 2 F double prime 0 over factorial 2 plus minus i k to the power 3 f triple prime 0 upon. So, the right hand side is basically expanding in powers of minus i k ok.

(Refer Slide Time: 19:41)



And then if I ask you to associate these coefficients of the expansion as F of 0 is basically if you see this is nothing, but ln of p of 0 which is nothing, but ln of 1 because p of 0 is 1 and that is nothing, but 0. So, this leading term is 0. So, what we have here is basically ln of p of k the left hand side has minus ik into F prime 0 plus minus i k to the power 2 F double prime 0 1 factorial 2 plus minus i k to the power 3 F triple prime 0 upon factorial 3 and so on ok. So, if you denote these coefficients as you know the first cumulant lets call them as the denote them as some cumulants.

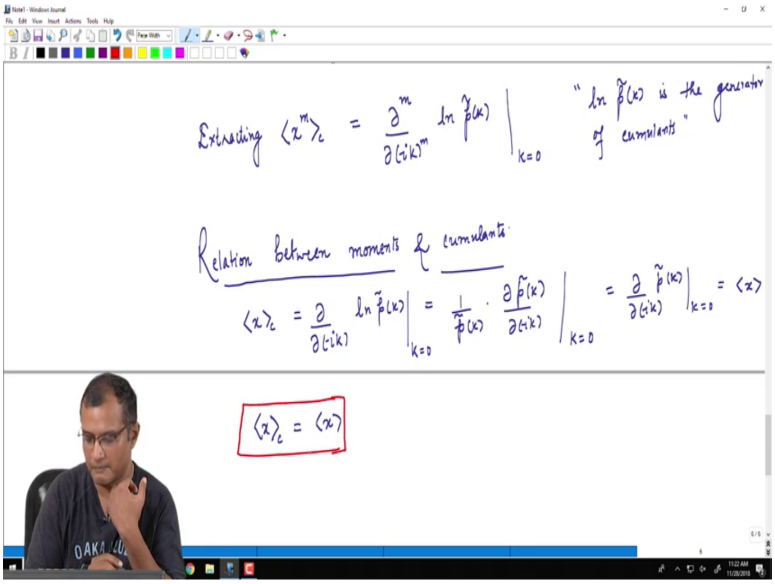
(Refer Slide Time: 20:59)



At this stage they just variables to you because I have really defined physically: what is a cumulant, F double prime as a second cumulant, F triple prime as the third cumulant and so on. So, nth derivative at 0 would be nothing, but the n th cumulant, this way our series becomes logarithm of p of k as minus i k into the first cumulant plus minus I k to the power 2 into the second cumulant by factorial 2 and so on.

So, I can write this as a series as a summation sum j going from 1 to infinity minus i k to the power j x j c upon j factorial ok. So, now, you can see there extract any jth cumulant by simply taking derivatives of the appropriate order.

(Refer Slide Time: 22:09)



So, extracting any mth cumulant here become extremely straight forward all you have to do is take derivative mth order derivative of your of the left hand side in that eliminates the coefficient. And you can write it down automatically as d to the power d m by d I kth power m ln p of k and this is evaluated at k equals to 0 ok.

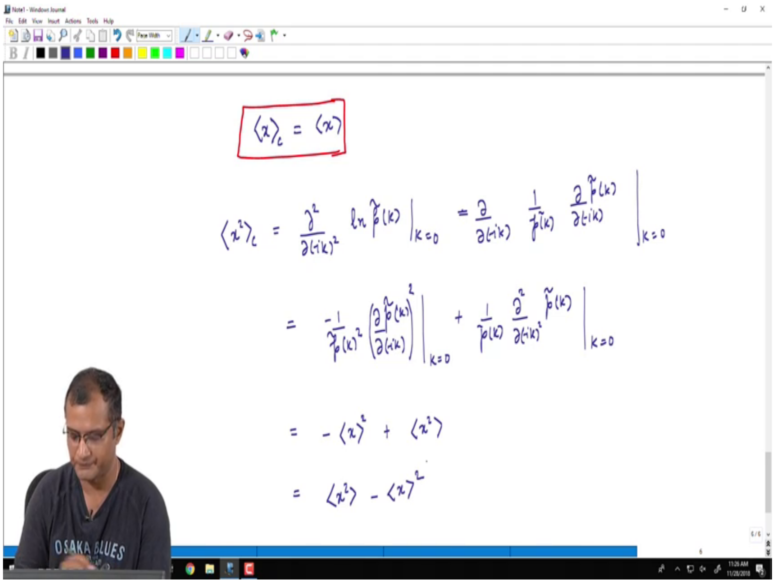
So, this is then classified as so this that is why we can say that ln p of k is the generator of cumulants now. So, for cumulants are again is fictitious objects which have not been related to any physical observation. So, we will now connect these moments with cumulants and natural meaning of this cumulants will automatically arise.

So, let us find out the relation between moments and cumulants and you will be surprised to see that some this cumulants you already know ok. So, so it is going to be my first agend ok. So, let is find out the first cumulant. So, the first cumulant is nothing, but m equal to 1 ok. So, that could be just its first derivative if I want to use the previous derivation it is just the first derivation of ln p of k ok; which I know is nothing, but one upon 1 p of k derivative at k equal to 0.

So, is derivative is nothing, but p of k into the p k by minus I k and this is entirely taken at k equals to 0; now characteristic function k equal to 0 is just 1. So, this is nothing, but d over minus d k of p of k at k equal to 0 which is nothing, but my by definition which is nothing, but my mean ok.

So, the first cumulant is nothing, but the first mean, this is the relationship that we got. So, this stage it is not very exciting we have the first cumulant is nothing, but the first mean, let us compute second cumulant.

(Refer Slide Time: 25:33)

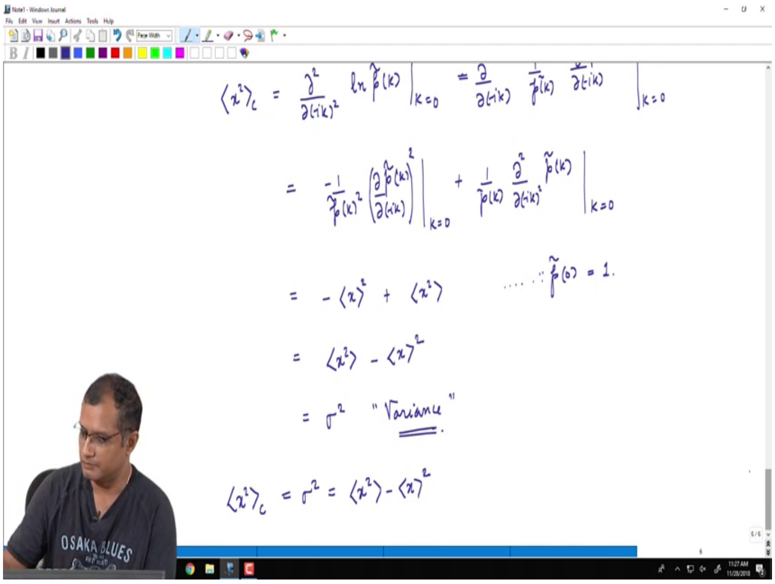


So, the second cumulant by this definition the definition which is here this is the definition of the mth cumulant. So, you put m equal to 2 here you get the definition of the second cumulant ok. So, let us use this definition to or advantage. So, we can write this as d square by d of minus I k in the whole square ln of p k at k equals to 0, let us find this out. Now we already has a first derivative extracted.

So, this would be just the derivative of 1 upon p k into d by d i k of p of k ok, all this take at k equals to 0. So, let us take the second derivative, this would be minus 1 by p of k the whole square, that is the derivative by parts into d of p k over minus ik plus 1 upon pk of the second derivative of the characteristic function again this has to be evaluated k equals to 0.

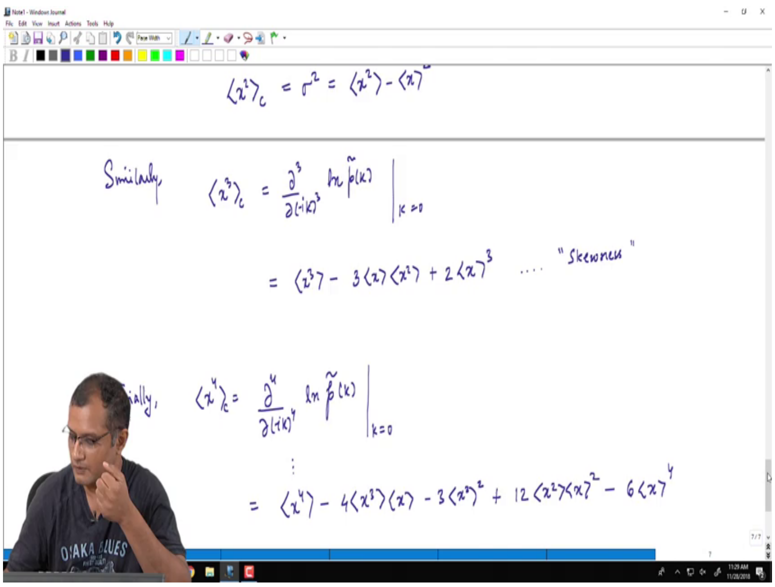
Now, if you look at the formula for moments which is given here over here let me circle this formula ok, uses formula for the moments what you have to just written becomes sorry would be a square here this become just minus of square plus, also write it as average of the second moment or second moment minus this square of the first moment.

(Refer Slide Time: 28:27)



So, this is a familiar and we will be could derive at because the value of p of 0 is nothing, but 1. So, this is also called as sigma square or the variance ok. So, what you have just derived is basically the variance of the random variable x. So, the second cumulant it is now a physical meaning of the variance of random variable, variance should be basically the measure of width of the random variable about the mean ok.

(Refer Slide Time: 29:07)

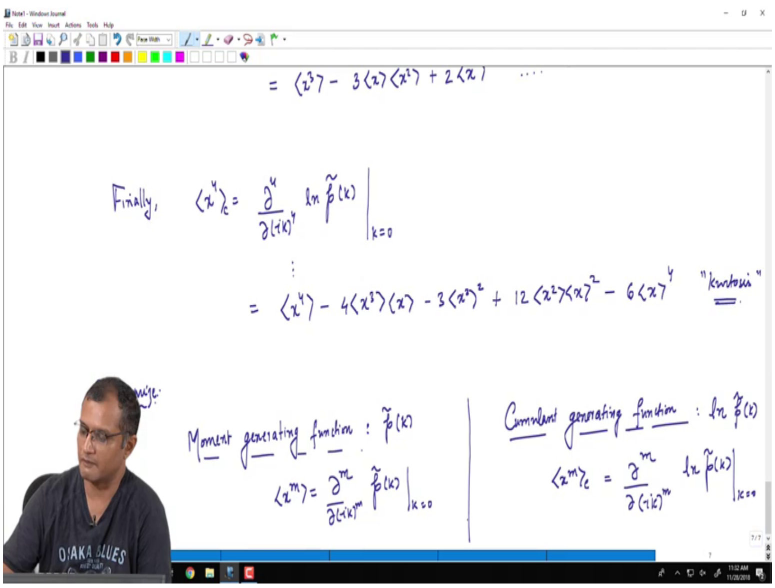


Similarly you can compute higher order cumulants for example, I can compute the third cumulant and that is basically its formula, third cumulant and forth cumulant can be written as. I will leave this as a derivation you to you to do at home ok, you can write this as the third derivative.

You can start the derivation with the derivative of the moment generating function, cumulant generating function taken it k equals to 0 and this can be written as x cube minus 3 mod x mod of x square minus plus 2 of 2 times mod x cube ok. And this quantity has a significant meaning or significance in statistical mechanics this is called as a skewness of a random variable. We will tell you whether distributional skew towards positivity side or negativity in side.

And finally, one more moment I am going to write down which is the fourth cumulant and if you derived this its expression would be x to the power 4 mod x 4 minus 4 times mod x cube into mod x minus 3 x square in the whole square plus 12 mod x square into mod x the whole square minus 6 mod x to the power 4. This is called as kurtosis in statistical mechanics ok.

(Refer Slide Time: 31:27)

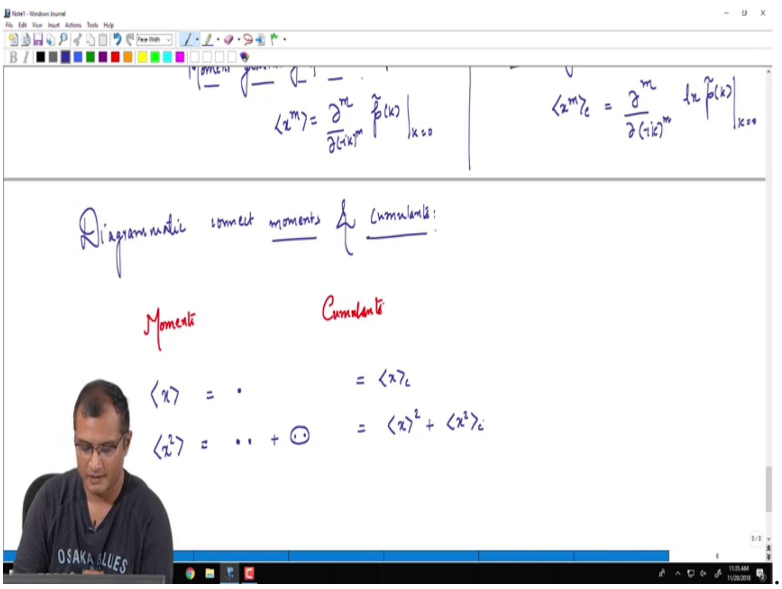


I will give some examples in the in the tutorial to explain what is what are these what is the physical meaning of these cumulants, but it is usually enough to stop the fourth cumulant and one usually it does not go to cumulants beyond order 4. Because all the information that one usually need is available with the 4 cumulants. So, it is enough to in most problems to go up to the second cumulant and in problems such as fluid mechanics, one can actually in compute cumulants up to order 4 which will give important information about physical property of the system.

So, this is the basic idea behind deriving these moments and cumulants. So, what we have if you summaries everything what we are seen so far is basically a moment generating function. So, if you want to generate moments then you basically use just the characteristic function which is the fluid transform of your PDF or if you want cumulant generating function then you can use logarithm of the characteristic function.

To generate mth moment you have to take the m th derivative of this characteristic function at k equals to 0 and to generate the mth cumulant you have to take the mth derivative of the cumulant generating function. So, it is in some sense symmetric in its application fine. So, this is the short summary of what we have learn so far. And, it is important to see that there is very beautiful diagrammatic way of remembering the connection between moments and cumulants.

(Refer Slide Time: 34:25)



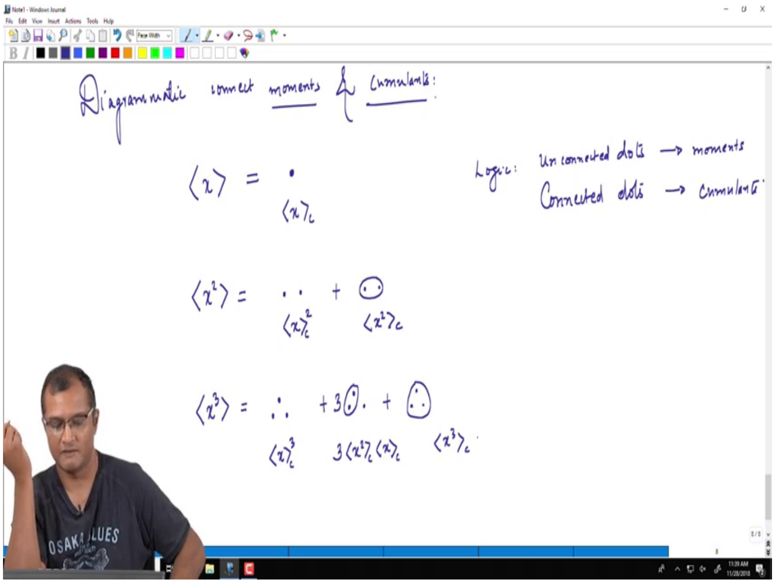
So, just because this method is very high pleasing and easy to remember I will give you this exposition of how to diagrammatically connect moments and cumulant without going into too much derivation. So, the idea is connect these moments and cumulants using some sort of diagram right.

So, now diagram is the followings, suppose I want to write down the definition for connection between first moment and first cumulant ok. So, if you have first moment ok. So, I am going to write down moments here on the left hand side ok. So, let me just rub on this and write down in some sense. So, I am going to write down moments here, I am going to write down cumulants here ok. So, so write down the first moment ok. So, the first moment is actually you can use for first domain you can use a single dot, number of dots is equal to the order of the moment ok.

Now, this first dot you can only show by you can show the dot as is you cannot anyway, if you have more than one dots you can connect them or you can leave them disconnected. So, in case of just single dot there is no logic there is no point of connection or no connection which is leave it as is ok, to connect dots you need at least 2 dots. So, the first moment is nothing, but the first cumulant itself that is the idea ok. So, I am going to write it as just the first cumulate. Now to second moment I have to take two dots, but I can write down two dots either disconnected or I can connect these dots where only two ways leave them disconnected or connect them.

So, this is the interpretation. So, you leave them disconnected this is nothing, but the moment, but it is squared how many doubts you have this is basically 2 plus the connected dots are second cumulant ok. May be this is better if I write below ok. So, I am going to write it below at this so that ok. So, maybe it is better that I right them below the dots.

(Refer Slide Time: 38:03)



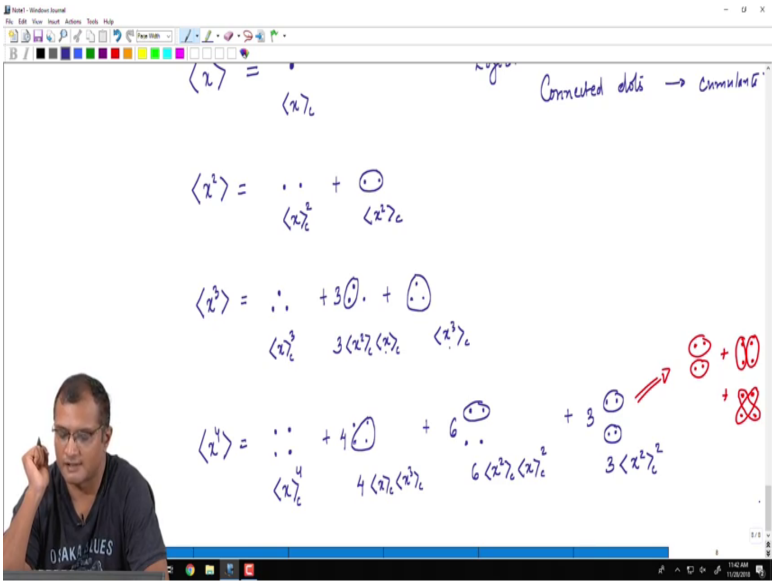
So, the first moment is just a single dot and that is nothing, but the first cumulant and what we a just shown is the second moment. So, for the second moment I need two dots ok, I was say the number of dots taken should be equal to order of the moment, but I can take two dots and also connect them ok. So, if you leave them it is called as the second power of the first moment and if you connect them it is the cumulant ok. So, the logic here is unconnected dots are moments and connected dots are cumulants ok, but I know that the first moment is also first cumulants. So, this is also cumulants square.

Let us go to third moment now for third moment I need 3 dots because I said that number of dots I equal to the order of the moment. Now a three dots I can I have three different possibilities I can leave all of them unconnected or I can connect two of them and leave the three unconnected, but there are precisely three ways of doing it ok. So, I can leave this dot and connect these two or I can leave this dot and connect these two. So, there are three ways. So, this basically is written precisely to account for the other two ways ok.

So, I am going to rub all this I am write down three times this plus I can take all of them together there are no more ways to. So, if you take three disconnected dots this is nothing, but the cube of the first moment ok, but the moment is also cumulant plus 3 times this is nothing, but second cumulant. Because, I have connect two dots into the first moment, this third dot is a empty is disconnected, but I know that the moment is also cumulant. So, you write it as a first cumulant and the final configuration is basically the third cumulant.

So, that is how you write down for the third moment x, third cumulant plus three times second cumulant into first cumulant plus the thirds cumulant.

(Refer Slide Time: 41:39)



Similarly, you can write down for the forth moment as, you know for the forth moment I need four dots now how many ways can you arrange 4 dots when you leave all of them free or disconnected. So, this should be the first moment to the power 4, but I know the moment is also cumulant.

So, you can just put of fourth power of the first cumulant, I can connect 3 of them and leave the third guy ok, but precisely there are four ways of doing it. So, this is nothing, but 4 times the first moment into the third cumulant, I can pick two of them and leave the other and they are basically 4 c to ways which is 6 6 ways of doing it ok. So, this is 6 times second cumulant into the square of the first cumulant plus I can pick 2 and also pick the other 2 ok. There are two ways of doing it, say I am going to write 3 times second cumulant square because each one of them is a second cumulant. So, this is a square of second cumulant.

Now, how can I do it 3 times which is basically one if you if you want to see how can I do it 3 times. So, this can be basically mean that I can take it like this, I can also take it like this, I can also take it like this that is it, only 3 times anything else would be a reputation. So, that is why this is just 3 times of second cumulant square such you can see this is a very nice diagrammatic way of constructing relationship between moments. And cumulants and a very powerful method that I have stated without any proofs, but the proof is also straight forward this is just comparing the moment generating function. And, the cumulant generating function and comparing terms of equal order, which will lead this derivation, which lead to this formulation actually.

So, we stop this lecture at this point and we will continue from the next lecture and derive some characteristic functions of important distribution in statistical mechanics such as the Gaussian distribution, the exponential distribution and the Poisson distribution. So, when we come comeback here next lecture we will start of from some very important distribution that have found in statistical mechanics and we will proceed from there.