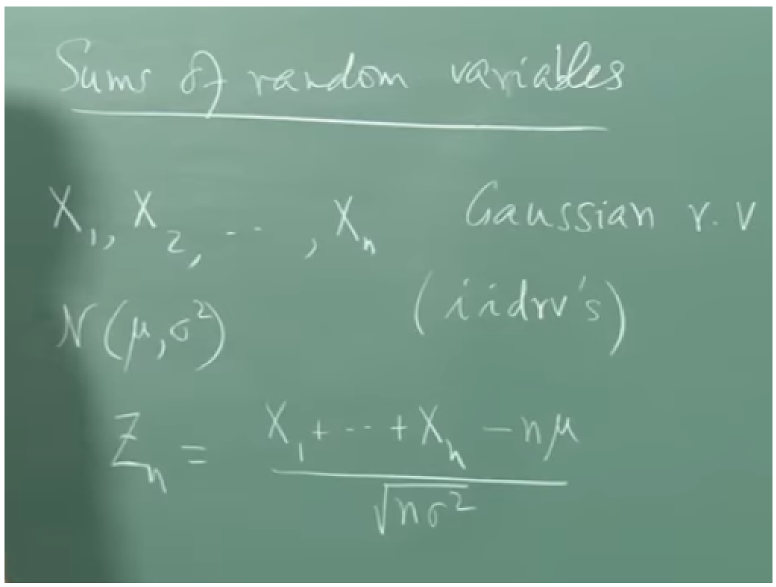
Physical Applications of Stochastic Processes
Prof. V. Balakrishnan
Department of Physics
Indian Institute of Technology-Madras

Lecture - 04
Central Limit Theorem

Okay, we start today by looking at sums of random variables. So and I said we are heading towards deducing our at least understanding some limit laws such as how the Gaussian emerges in many cases etc. So I start by saying, asking a very simple question and that is the following.

(Refer Slide Time: 00:52)



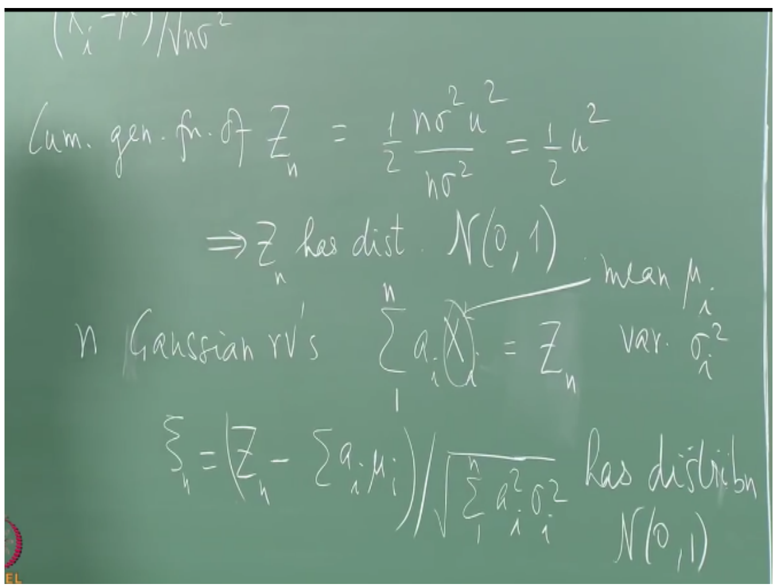
Suppose I have a set of random variables, Gaussian random variables and let us denote them by X 1, X 2 up to X n and each of them is a Gaussian random variable all of which are identically distributed okay. So they are independent, statistically independent, identically distributed random variables and there is an abbreviation for this. So independent identically distributed random variables, iidrv's each of which is a has a Gaussian distribution.

And it is a normal distribution with a mean mu and a variance sigma square and is generally denoted in this fashion. A normal distribution with mean mu and sigma square. So that is notation to say that the random variable is a Gaussian with the specified mean and variance okay and each of them has exactly the same distribution.

Then the question is what is the distribution of a variable which I will define as Z n this is equal to X 1 plus etc up to X n - n times the mean divided by and I must scale it out to make it dimensionsless here, dvided by n sigma square square root. Remember sigma is the standard deviation for each one of them and this is just the sum of the variances out here n sigma square because the variances of independent random variables add and I take a square root.

So what is the distribution of this Z n, that is the question okay and now of course you can do this the hard way. You can write its probability distribution by writing a product of the distributions of all these variables out here integrating over all the sample spaces of each of these variables and then imposing a delta function constraint that this combination should be equal to that. That is a painful way of doing it in this case.

(Refer Slide Time: 03:03)



On the other hand I argue that each of these variables if I consider a variable like X i - mu I subtract the mean out in this case. These are all independently distributed and each of these fellows, these variables has a variance sigma square and therefore the sum of all these fellows has variance n times this and I immediately write down what the characteristic function is or what the moment generating function is or what better still the cumulant generating function for this whole thing.

So the cumulant generating function of Z n this thing here is equal to, there is no mean because the mean is 0, I have subtracted it out and then it is equal to one half; sigma square u square was the formula for single Gaussian variable, for n of them it is n sigma square u square but remember I have scaled out each variable by this thing here. So X i – mu/n sigma square. I multiplied this random variable by a constant 1 over root n sigma square and therefore in the variance it is just this constant square.

So this is equal to, this fellow, which is equal to one half u square. So we have a statement that the cumulant generating function of this sum we scaled by shifting the constant out by removing the mean and dividing by the square root of n sigma square is just one half u square which implies immediately that Z n has distribution, normal distribution of 0 mean and unit variance which is called the standard Gaussian distribution N of 0, 1 okay.

So our first interesting result that if you have n of these Gaussian random variables and you subtract out the mean value and divide by, suitably rescale it by this factor here which is square root of n times the variance of each of these then this combination is normally distributed. For every n this is true okay. You can generalize this result immediately. Consider for example this quantity.

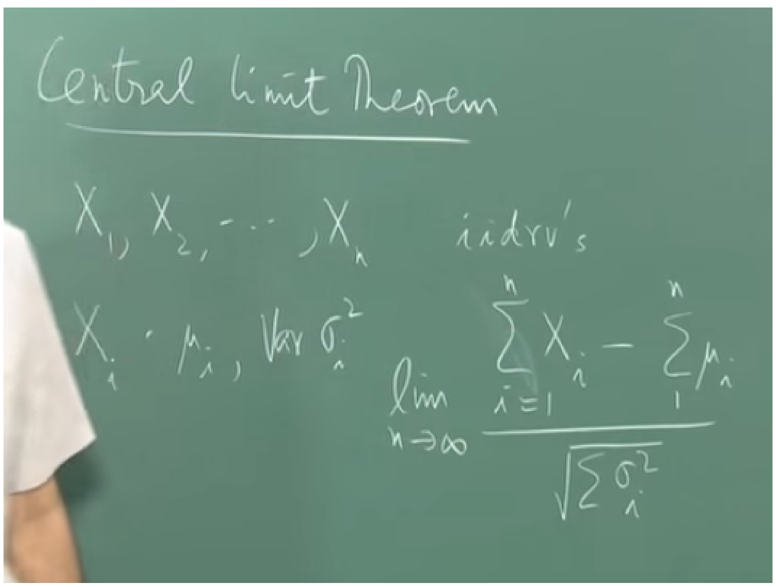
So let us suppose that you have n Gaussian random variables and we have a a i X of i summed from 1 to n, so the X's are all Gaussian random variables and let us suppose that the mean value of any of the X's is some new i and the variance is some sigma i, but they are all Gaussian random variables, then what is the distribution of this combination here okay. That combination is not a very good one. We need to do something better.

We need to scale out the mean value etc. So let us call this equal to let us call this Z n just for convenience and let me define a random variable psi n = Z n minus summation a i mu i, I subtract that out and I divide by the square root of, what should I divide by the square root of? Assuming that this fellow x i has mean mu i and variance sigma i square what should I divide by here?

Ya I should sum from 1 to n a i square sigma i square okay. What would be the distribution of psi sub n, the standard normal distribution right. So has distribution N of 0 comma 1 and that is true for every n okay. So we see that if you do a rescaling and a translation suitably subtracting the constant a sum of Gaussian random variables or a linear combination more generally a linear combination of Gaussian random variables has a Gaussian distribution.

So in this sense the Gaussian is extremely robust okay. You take 2 independent Gaussian random variables, do a linear combination then some rescaled version of this sum has a Gaussian distribution once again okay. It reproduces itself as you can see. The theorem is actually more general than that um and the statement is the following and this is one statement of a famous theorem called the central limit theorem.

(Refer Slide Time: 08:15)



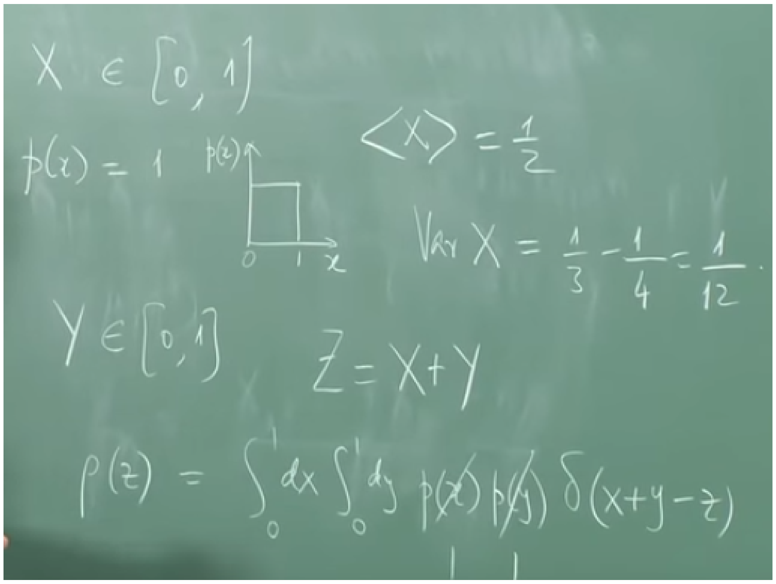
I am saying this very loosely there is a regress way of stating this theorem in its all its generality but what I am going to do here is simply state it in very very simple terms and that is that if u give me a n random variables X 1 to X n, they need not be identically distributed, X 1, X 2, …X n and let us to start with look at iidrv's does not matter, iidrv's. Each of them X i has a mean mu i and a variance sigma i sigma i square which are finite.

So we want the first 2 moments of every the variance and the mean of every one of these variables to be some finite value. Then one can ask what happens to the sum sigma X i, i = 1 to n minus the mean value summation mu i 1 to n rescaled by this summation sigma i square okay and now if you say these excess are distributed by any arbitrary distribution which has a finite mean and a finite variance independent of what their distribution is, the statement is that the limit as n tends to infinity of this combination here has a normal distribution okay.

So that is the statement of the central limit theorem. It says no matter what distribution you start with you will end up with the Gaussian distribution and what we would like to do is to look at a couple of toy models and see if this is really true and how it comes about rather than how to prove this statement is a regress proof of this theorem and it is called generalizations.

What we will do is to look at the simplest instance, simplest possible example toy model and see exactly where it comes from, where does it become, how does it become Gaussian and so on and that is an instructive lesson. Now let us do the following. Let us look at a set of n variables. Let us first do the following.

(Refer Slide Time: 11:02)



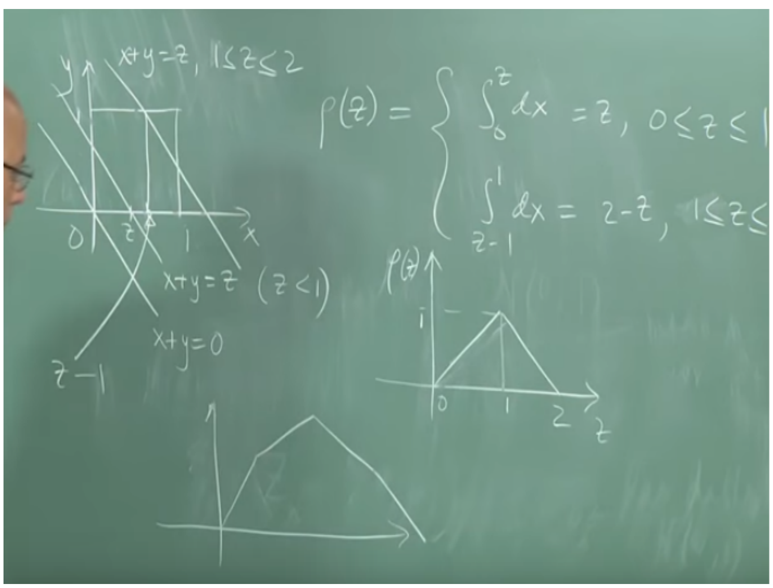
Let us look at a variable X which is in the sample space 0 to 1 and p(x) = 1. So this is uniformly distributed in the unit interval between 0 and 1. So if I plot this p(x) is just this guy. So the average value is half and so on and so forth. What is the variance? So it is clear that X =1/2 because it is just uniformly distributed and the mean square is the integral of X square from 0 to 1 which is one third.

And you subtract half whole square so it says the variance of X equal to one third minus one fourth equal to 1 over 12 in this case. Now I ask what happens if I have 2 of these variables. If I add 2 of these variables what does the distribution look like? So let us call, let us have another such variable Y which is also let us say we got a Y which is also element of 0, 1 distributed in exactly the same way out here.

Then the question is what is the distribution of a variable X which is defined as X + Y. Now what is the sample space of Z? 0 to 2 in this case. So we immediately write down p(z) the probability density function of this variable Z must be equal to integral 0 to 1 d x, integral 0 to 1 d you p(x), well it is 1 in this case, p(u) that is also 1 in this case. I do not want to confuse with this symbol here.

Let me call this rho and you have p(x), p(y) delta function of x + y - z and this is 1, this is 1. So we got to do this integral. It is as simple as that. But we got to be little cautious doing this delta function integral and that is done as follows.

(Refer Slide Time: 13:51)



I draw a picture and here is x, here is y, here is the origin and the integral is over the unit square 0 to 1. Those are the boundaries of, that is the interval in which each of these variables takes values and now I try to impose the condition that x + y is some number z between 0 and 2 okay.

If it is 0 the graph looks like this. So this is the curve x + y = 0 or y is - x 45 degree line 135 degree line and as long as z is less than 1 it is clear that the line goes something like this.

So this is x + y = z for z less than 1 and what is the value of this integral. Well, look at it this way. I do the y integral first and then I do the x integral. When I do the y integral I must see if the delta function fires or not. So as I move up in y for every value of x there exists a legitimate value of y such that x + y is equal to z provided x is less than this value which is z and then I can just finish off the integral completely.

But as soon as x exceeds z the curve looks like this, as soon as it exceeds 1 it starts looking in cutting these 2 points. So till z hits 1 you are fine. It is just an integral in x running from 0 to 1. So it says that you can get rid of the delta function constraint and write rho of z equal to integral 0 to z dx because the y integration is gone, the delta function has taken care of it but the region of integration over x has been curtailed not from 0 to 1 but 0 to z okay.

So this is equal to z itself right up to 1 okay. What happens beyond that is that the curve looks like this. So this is x + y = z and z is 1 less than equal to z less than equal to 2. What is the range of integration now in x if I finish this delta function constraint? As I move up in y I have a contribution as long as x is bigger than this value whatever it is only if x is less bigger than that value right up to 1, right up to this point 1 okay.

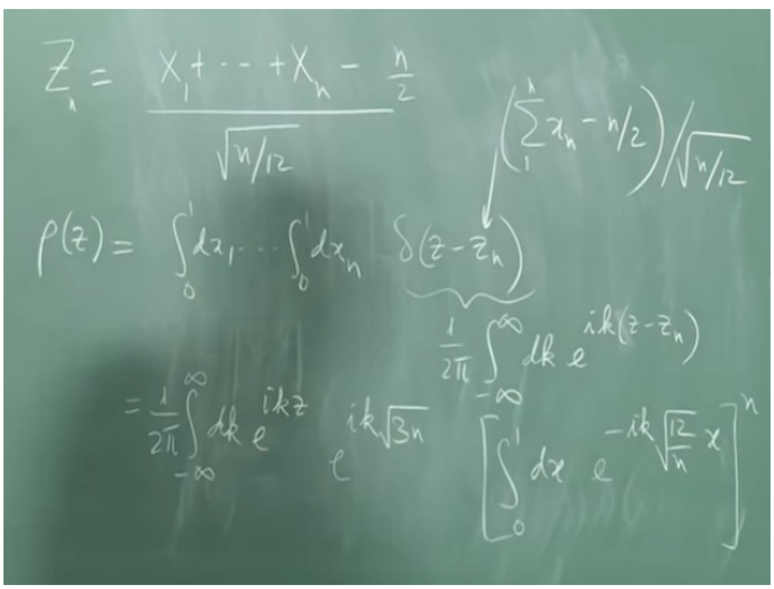
And what is this value? This straight line here this point here y is 1 um and you are on the curve x + y is this guy here. So what is this value here? Z - 1. So this is equal to z - 1 up to 1 dx = 2 - z okay. So you see how the shape changes as soon as you have this z crossing 1. So therefore there is a point where this thing is sharply peaked this value and if you write what the distribution looks like.

So here is z, here is rho z and is actually something like this. At z = 1 it goes up to 1so it is this, looks like that. So what was flat, as soon as you add two of them has become a little conical thing like that, triangular distribution okay and it is normalized because the area under this curve is half the base and times the height. The height is unity and the base is 2. So it is 1 okay. So this is what is going to happen successively.

I add more and more of these variables but you see if I had one more variable I had x, y, z and then you call w the sum for example then you are going to have an integral which is in the first instance going to constrain this to 1 minus something and then 1 minus x minus whatever so it is going to be more and more complicated as I go down here. So it is a very foolish way of doing this when you have more than 2 variables.

In fact what will happen at the next stage and you should do this numerically to see what the shape changes like. At the next stage it is going to go something like and so on, will go right up to 3 etc. I can get rid of that spread all the way by subtracting the means each time but you see the shape changing all the time. It is getting the range is getting wider and wider and the shape is changing. So we would like to see what happens when you put a very large number of these and then take the limit when the number goes to infinity and the way to do this is as follows.

(Refer Slide Time: 19:22)



So let us define Z n to be equal to X 1 plus up to X n minus the mean and the mean for each of these is a half. So let us define this to be n over 2 divided by the variance. What is the variance of each of these. It is one twelfth as we saw and for n of them it is just n over 12 and I want to divide by variance so can be defined Z n in this fashion and I ask now what is the distribution of this Z n, probability density function okay. Let us call that variable is value Z.

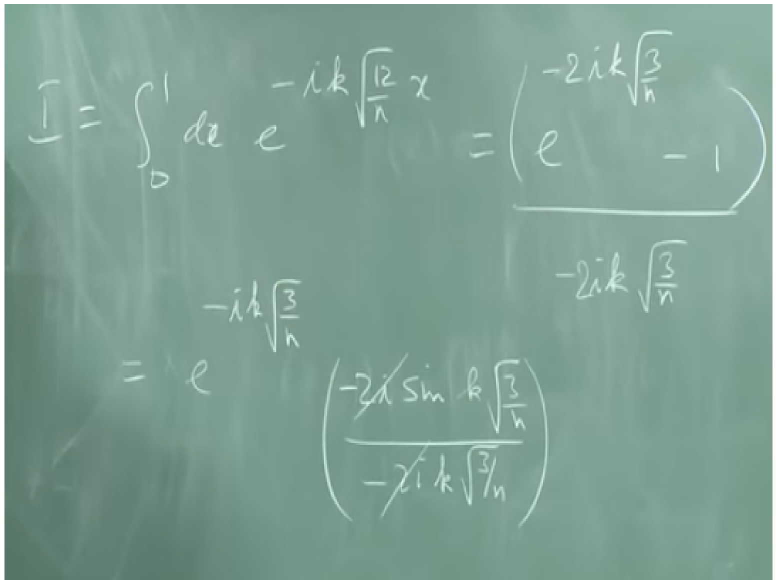
So this is equal to integral 0 to 1 d x 1, 0 to 1, d x n in this fashion and then the p's and all of them are unity. This is all uniformly distributed over the unit interval and then a delta function of z - z n where little z n is the summation x of n 1 to n - n over 2/ root n over 2. That is the integral I have to do okay. Well it is sort of foolish to try to do it as it is because as you can see the constraints are going to keep on multiplying here.

So the way to do it is to try to factor this into a product of integrals and we do that by writing a representation for this as the fourier representation a minus infinity to infinity dk e to the i kx e to the i k whatever minus z n. So I replaced the delta function by that and what happens to rho of z? This becomes equal to 1 over 2 pi integral minus infinity to infinity dk e to the ikz that comes out and then I have an integral 0 to 1 dx 1 e to the z n is going to be e to the ik, I am just concerned about that n over 2, what happened to that.

Oh ya it is sitting there, it is very much there. So we got to take care of that. So let us do that. This is e to the ik. There is a minus z and so does a plus n over 2 there e to the ik n over 2. This minus and that minus goes away and then root 12 over n and this is 2 root 3 and then there is an n. So let us write this as square root of 3n. That factor is there, we cannot avoid it and then an integral from 0 to 1 say dx 1 for instance.

Let us write that first e to the - ik x 1 and then there is this guy root 12 over n x 1 okay yes in this fashion and that is it but then the same integral gets repeated right. So let us just write this as dx and take this to the power n. It is perfectly alright. So this becomes and that is a trivial integral to do. So rho of z is let us first simplify this integral and see what this is.

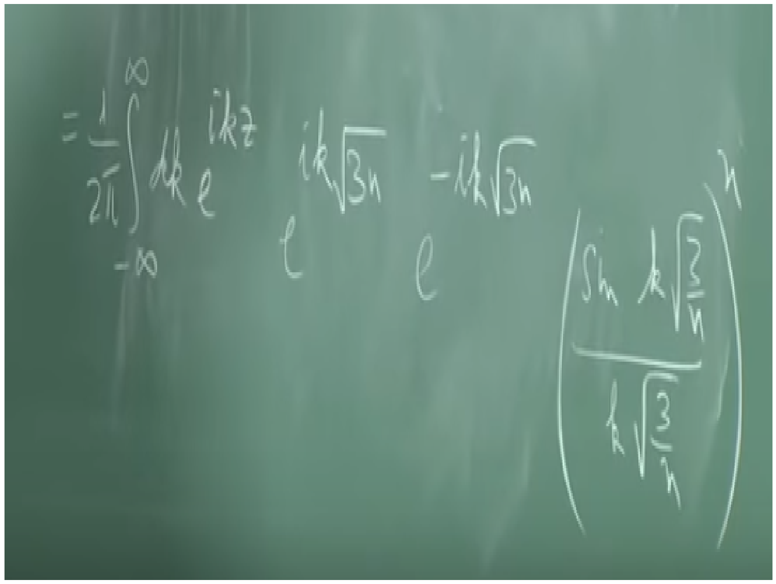
(Refer Slide Time: 23:52)



Let us put I equal to integral 0 to 1 dx e to the minus ik and what do you get here. This is 2 root 3 right root 12 over n x which is equal to e to the - ik - 2ik root 3 over n – 1/- ik 2ik root 3 over n doing this definite integral right and let us pull out this factor. So this is equal to e to - ik root 3 over n and then this is going to appear with a plus exponent there. There is a minus here. That is minus 2i sin of this guy right.

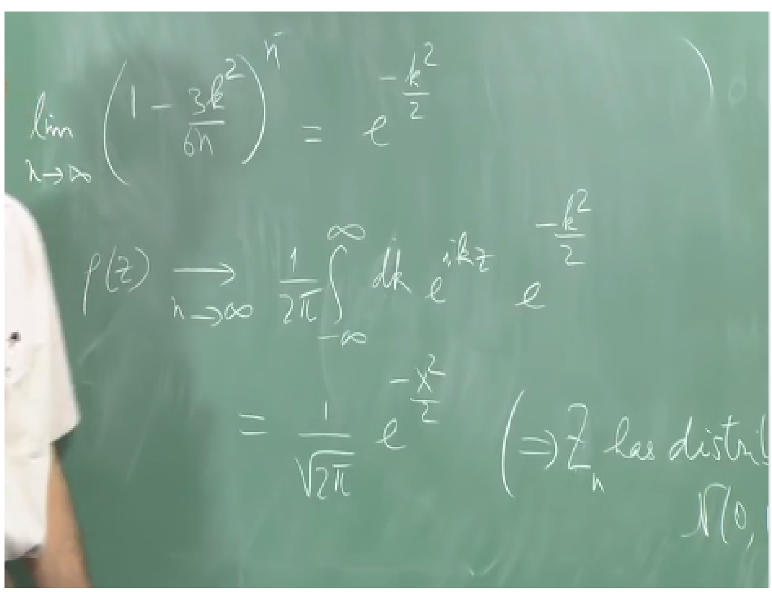
So - 2i sin of k root 3 over n/- 2ik root 3 over n. Watch all the minus signs and so on because I do not swear to this thing. Think it is okay. If it is wrong I blame you. So this cancels out here and that is i but what is appearing is i to the power n. So this whole thing raised to the power n, let us put that in. If I raise this to the power n, this becomes square root of 3n.

(Refer Slide Time: 25:52)



So that neatly become e to the - ik square root of 3n. That is going to cancel against this and then there is a sin k root 3 over n over k root 3 over n to the power n. That the answer which is equal to. So that is rho of z. It is obviously the fourier transform of fourier inverse fourier transform of this guy. That is what rho of z is. Now we are in good shape because this factor cancels against that and all you got to do is to look at the behaviour of this as n becomes larger and larger. What happens to this guy as n becomes larger. The argument goes to 0 as you can see. So it is like sin x over x okay and as n and as x goes to 0 the limit is 1 right. The next term will be a correction of order 1 over n because sin x is x - x cubed over 6 the next term.

(Refer Slide Time: 27:12)



So what we need to do is to find this ratio here becomes 1 minus the square of this guy so k square 3k square over n and then there is a 6 out here, that is the next term raised to the power n and we need limit n tends to infinity this guy which is equal to e to the minus k square over 2 right. So we are home because it immediately says that rho of z in the limit as n tends to infinity tends to 1 over 2 pi integral minus infinity to infinity dk e to the ikz e to the minus k square over 2 okay.

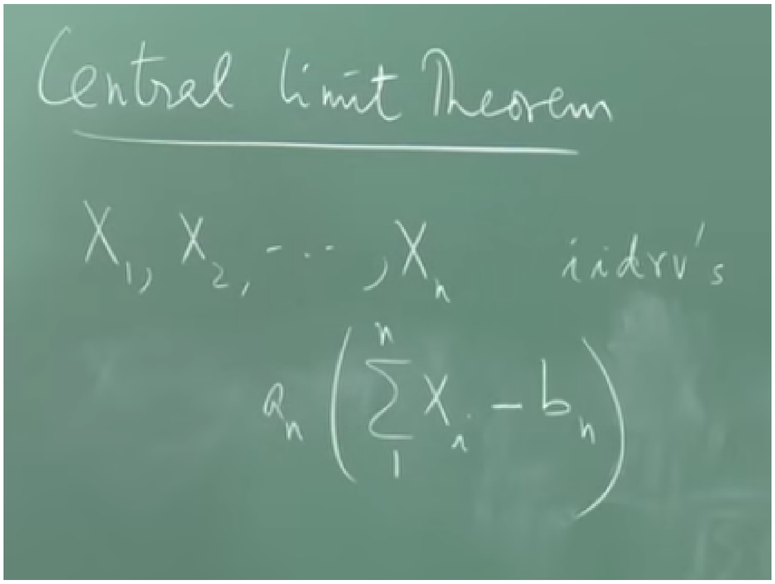
Well that is a Gaussian with variance equal to 1 out here because the inverse transform is also a Gaussian which implies this is equal to 1 over square root of 2 pi e to the minus which implies that Z n has distribution N 0, 1. 0 mean unit variance okay. That is precisely what the inverse transform of this guy is. By the way you can do this integral by completing squares etc. and you will end up with this result here. This integral will give you a square root of pi that cancels and gives you a 2 square root of pi in the denominator.

So you see we started with a distribution, which look nothing like a Gaussian, flat distribution and the variable was 0 to 1. But yet as you kept adding stuff to it we are ending up with a Gaussian distribution and the random variable can take on values all the way from minus infinity to infinity because we made sure that negative values could be reached because we subtracted out that mean part. So it got centered at the origin and now the variable can take on all values minus infinity to infinity.

Very important to note that this is the only limit in which there is a distribution possible. You see what happens is that we subtracted out something which is linear in n because means were all subtracted but divided by something which was a square root of n and that gave the right scaling so that you end up with a Gaussian distribution. Actually the theorem is a little more general than that. We could have variables which had different distributions and still you would have similar properties but that gets a little more intricate.

Generally you are used you encounter cases where they are all iidrv’s some kind. So although I have done this for the case of this uniform distribution the same thing is actually true if you had any distribution which has a finite variance and mean okay with suitable rescaling.

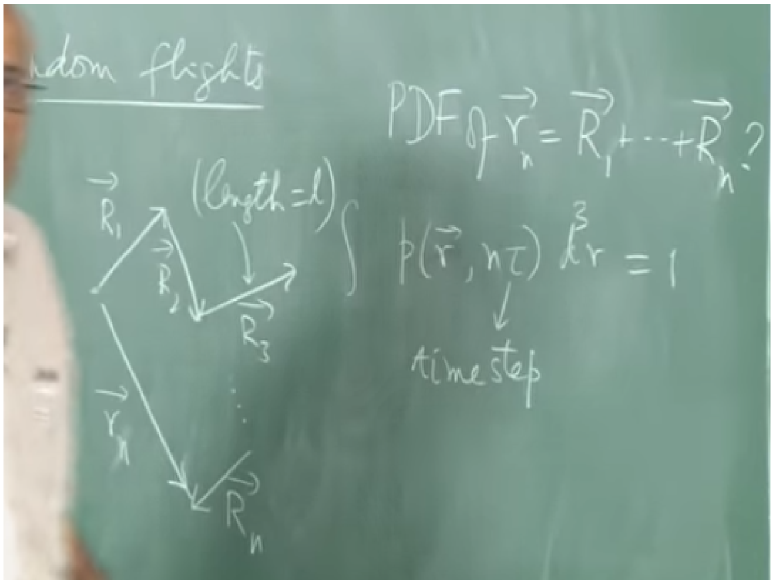
(Refer Slide Time: 30:49)



So the statement is that if you have iid's of this kind then if you have iidrv's of this kind then there exist constants a and b such that the variable summation X i 1 to n minus some constant b rescaled by some factor a sub n independent positive quantity and that is some real number b sub n there exist a n and b n such that the limit of this as n tends to infinity goes to a Gaussian distribution. So that is the statement of the central limit theorem.

We did this in a very simple scalar way just you know one dimensional variables and so on but it does not matter, what the dimensionality is does not matter at all and this is the famous random walk problem about which we are going to say a great deal. So let me do that step by step and then we will see how the central limit theorem emerges okay. We are going to talk about the random walk on discrete lattices but right now in connection with the central limit theorem let us just look at a problem of a random flight.

(Refer Slide Time: 31:53)



So let us call this problem of random flights and the rule is the following. I start in 3 dimensional space. Actually it is independent of the dimensionality of the space as well. I start in 3 dimensional space at some origin and I take a step of fixed length l in some arbitrary direction in 3 dimensional space. So I fly from this point to that point and this vector let us call it R 1 okay.

Having reached this, at the next time step when I do this in discrete time at the end of every time step I take a discrete step whose length is l magnitude is l but the direction is anywhere in space. So the next time around I go here and then I go there. I can cross my path earlier path it does not matter. No constraints at all whatever. So this is vector R 2. This is vector R 3 and so on and I do this for some n steps okay.

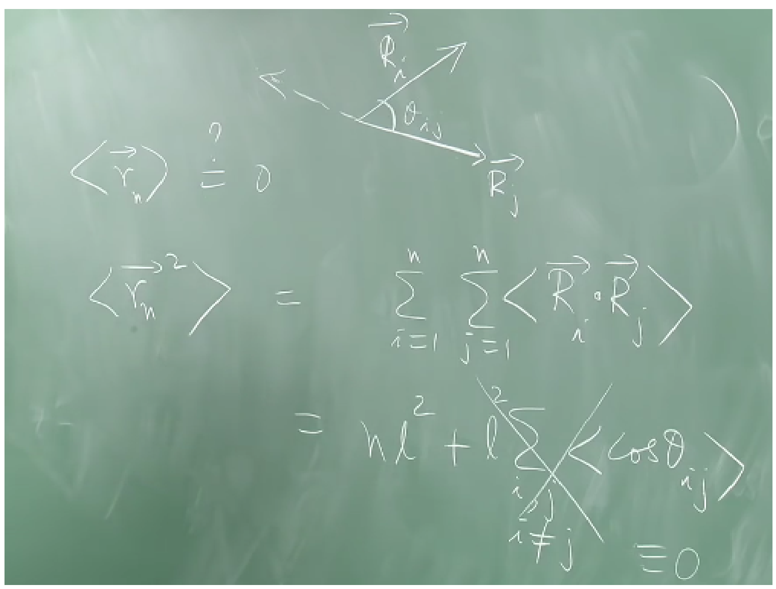
And let us suppose that at the end of the nth step I am here. This is R n etc. and the end to end distance from here to here let us call it r. I should put a subscript n here actually. So let us do that and I ask what is the probability density function of this vector R n, R sub n okay. It is a 3 dimensional vector so I need to know its distribution in space as well as in in direction as well as in magnitude.

So I ask what is the PDF of r sub n which is equal to R 1 plus dot dot plus R n. It will be some quantity. Let me denote it by the following. Let me denote it as p(r), n because it is dependent on the number of time steps that I have and I would like to look at it and the limit little n tends to infinity okay. I should really put a time step also so that I have time included in a dimensional way. So some time step tau for example.

We can put that later but just to remind myself of this let us put a tau here and call this a time step. I am doing this in anticipation of the fact that later we are going to look at things in continuous time. So I need to have a quantity of dimensions time. Right now for this problem it is only the n that is relevant. It is only the number of steps that is relevant. I want to know what this quantity is. It should clearly be normalized.

So it is clear that this integrated over all space d 3 r should be equal to 1 at every tau at every n. This should be the case okay. So the whole thing should be normalized and the random variable is the sum of n random variables, all of which are identically distributed except each variable is a vector in 3 dimensions. Only the length is fixed but the direction is arbitrary completely. So let us go about it in the following way. What would be the average value of this r sub n?

(Refer Slide Time: 35:28)



Intuitively what would be the average value? It should be 0 because it is a vector so it is as likely to point in one direction as in the opposite direction with equal weight so I expect this should be 0. When we of course find the distribution exactly we should verify if this is so but it is trivially true that this is equal to 0. What would be the mean square value that means the dot product of this with itself. What would that be?

So what we now asking for is the expectation value of R n dotted with itself. So summation i = 1 to n summation j = 1 to n R i dot R j. This is what we are asking for but this summation is commutes with the operation of taking averages, so it is this okay. What is this angular bracket over, average over what? Over the collection and sample of what? All possible realizations. All possible random walks, random flights of n steps. All possible this thing.

That is what I have to take the arithmetic average over but what we are trying to do is to find the probability distribution itself so that these ensemble averages can be replaced by averages over the weighted averages over the probability densities okay. But now this guy here is equal to let us suppose the length of each of these is l, length equal to l okay then this dot product has n diagonal terms where i is equal to j and then of course R i dot R i is just the square of the magnitude of this vector which is l square and then n of these guys.

So this is nl square plus summation i, j but i not equal to j in this case R i dot R j. Now if R i is a vector in some direction like this and R j is a vector in some direction like this bringing them down to the same common tail here and theta ij is the angle between these 2 vectors then since the magnitudes of these vectors is fixed all we have to do is to multiply this by l square and take the average value of cos theta ij and that is what the dot product means.

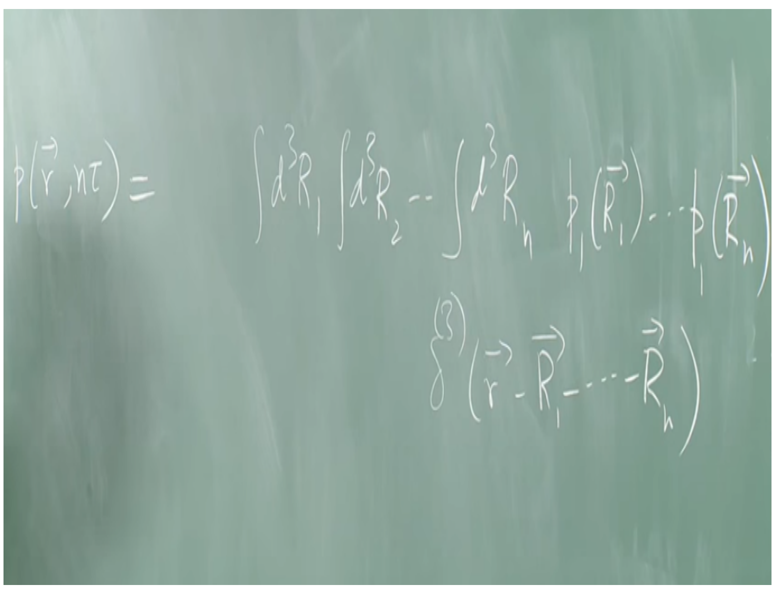
But what is this average equal to? It has got to be 0 because for every such configuration there is an equal equally probable configuration like that in which the other fellow points in the other direction okay and cos theta is - cos pi - theta. So these 2 contributions cancel each other right. So this goes away immediately. This is identically 0 and you end up with r n square is n times l square and since the average is 0, average value of the vector R n is 0 this is the variance.

So if you take a random flight and ask for the end to end vector, displacement vector the variance of that vector is proportional to n okay, not n square but n. So the standard deviation is turning out to be square root of n. Root mean square displacement is turning out to be proportional to the square root of n. Now if n is proportional to time you can see that the distance covered, root mean square distance covered is going like the square root of time which is the famous random walk because if I walk purposefully in this direction at a constant speed the distance I cover is proportional to the time.

But if I walk meaninglessly meandering in random directions then my mean displacement is 0 but the end to end distance that is not 0, the mean, root mean square distance is going to be proportional to the square root of the time elapsed when I do a random walk. This is typical behaviour of typically diffusive behaviour okay as we will see there are exceptions to it when there are certain physical conditions met.

But in general this is going to happen that the root mean square displacement is going to become proportional to the square root of the time right or the mean square distance is going to go like the time itself, first power of time. So this is crucial. That is just a dimensional constant but this is crucial. So we already have some valuable information. What we need however is this distribution, the full probability distribution function. So let us write it out and see what happens.

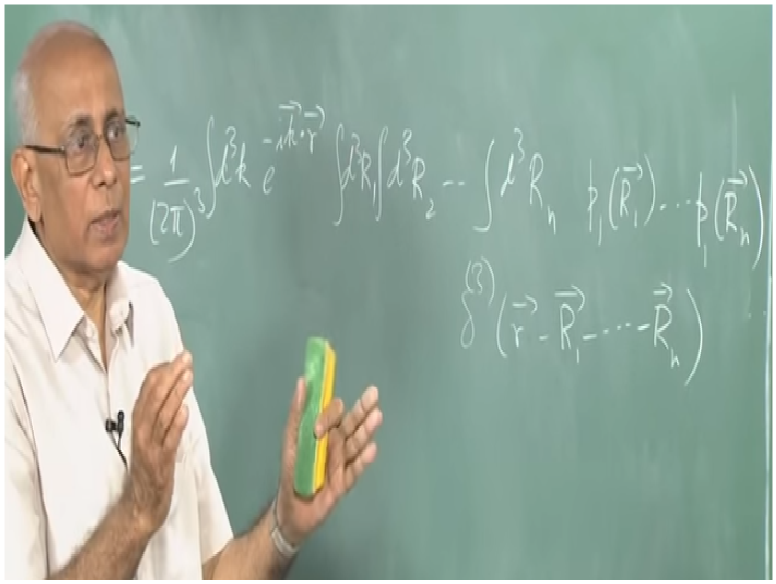
(Refer Slide Time: 40:46)



We need now to integrate over all these guys so we put in a delta function constraint and we will say a p(r) n tau is equal to the delta function of r minus this vector here and the standard trick of course is to write that delta function in a fourier representation immediately. So I am going to have d 3 R 1, d 3 R 2 up to d 3 R n times the distribution of each one of these guys that is needed for each of these steps and since it is a one-step quantity let me call it p 1 of R 1 p 1 of R n.

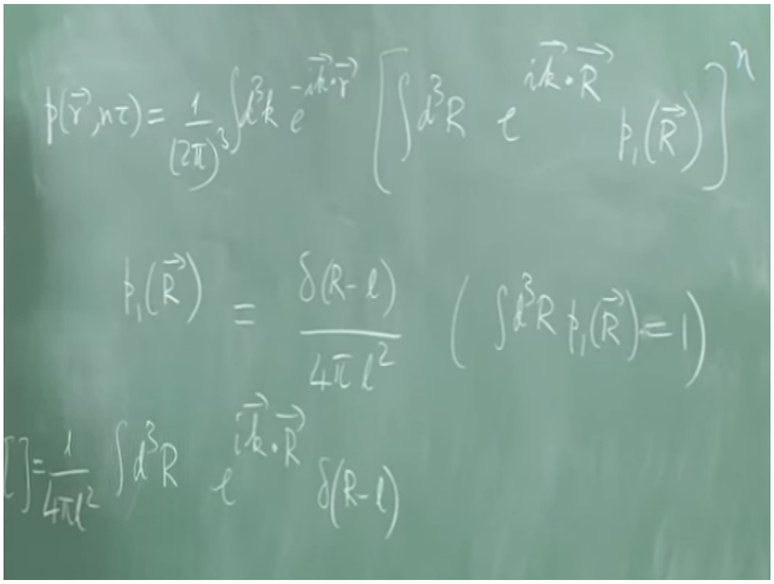
You still got to write down the probability density function of a single step, that is going to be my input and then a delta function, the 3 dimensional delta function of r - R 1 - - R n in this fashion. That is what this guy is and I write a fourier representation for this and bring that integral to the left hand side.

(Refer Slide Time: 42:08)



So we have a p(r) and tau = 1 over 2 pi whole cube, it is in 3 dimensions, integral d 3 k times e to the power - ik dot whatever it is r so e to the - ik dot r and then I have e to the - ik dot r and then I have e to the - ik dot R 1, e to ik dot R 1, e to the ik dot R 2 etc. times this fellow here and it is the same integral repeated for all of them. So it is some integral to the power n and what is that integral?

(Refer Slide Time: 42:55)



It is equal to integral d 3 R, forget the index now, not need; e to the power ik dot R p 1 of R raised to the power n. this is what we need okay. Now what is p 1 going to be? First of all what is going to be the physical dimensions of p 1? Because I want p 1 of R integrated over all components of R to be equal to 1 so it has got to be 1 over length cubed. I need a third power. But what we do know about this is that you are asking for the distribution of a vector whose length is l and which is moving about its got all directions possible, it can be, it is tip can be anywhere on a sphere of on the surface of a sphere of radius r.

So this has got to be proportional to a delta function of R - l. It has got to be proportional to that right and what is the constant of proportionality. How do you discover what that is? You normalize, you normalize this but over all directions and what is the total solid angle, 4 pi. So you divide by 4 pi and the volume element also has an r square dr but r’s got to be equal to l otherwise it is 0 right and that is got to be divided out. So it is 4 pi l square and that is all it is okay.

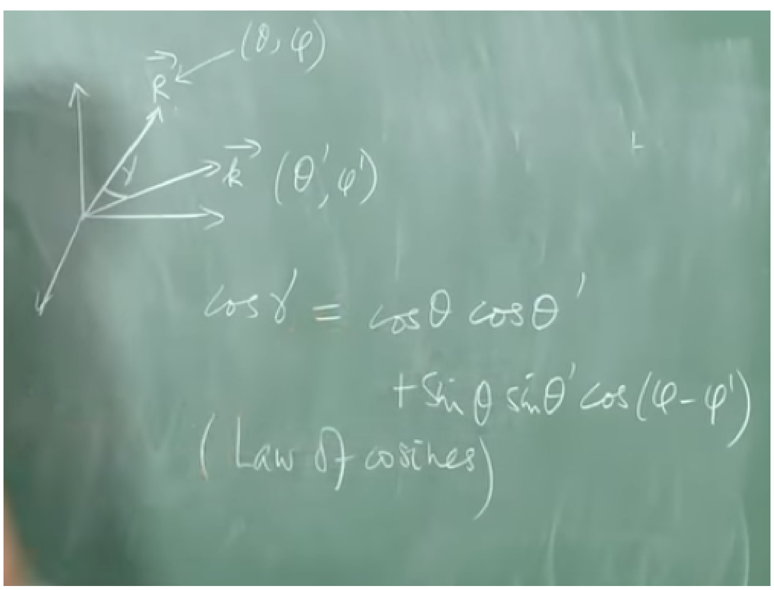
So that is p1 of each of these vectors. It has got dimensions of 1 over length cubed because this is 1 over length square and this delta function of course has physical dimensions of 1 over the argument which is 1 over length okay. So it is a sharply defined sharp you know at one point there is a spike in r and that is it okay. So it is trivial to verify that integral d 3 R p 1 of R = 1. We use that fact. So let us put that in in this integral and see what happens.

So you have integral d 3 R e to the ik dot R and then a delta function of R - l and a 1 over 4 pi l square. That is what this square bracket is and I need to do this integral okay. Obviously I should do this in spherical polar coordinates it is the simplest because this constraint here involves this magnitude R but I have to take care of the angular variables integrations right. Now what should I do. I should choose, I should choose my polar coordinates in such a way that the polar axis is along the vector that is sticking out.

There is a vector k sitting here which is not being integrated over and since this whole thing is spherically symmetric this region of integration is all space, spherically symmetric. This thing is spherically symmetric and that is a scalar. So it is invariant and the rotations of the coordinate axis which implies that I can choose the orientation of my axis as I please and the integral does not change.

And the convenient thing to do is to choose it along the direction of whatever vector is sticking out because then that ik dot R just becomes magnitude of k magnitude of r times the polar angle and that is theta in this case in spherical polar coordinate. Otherwise you are in trouble because if you did not do that the answer does not change but it is needless complication. Just to tell you what the complication can be.

(Refer Slide Time: 47:00)



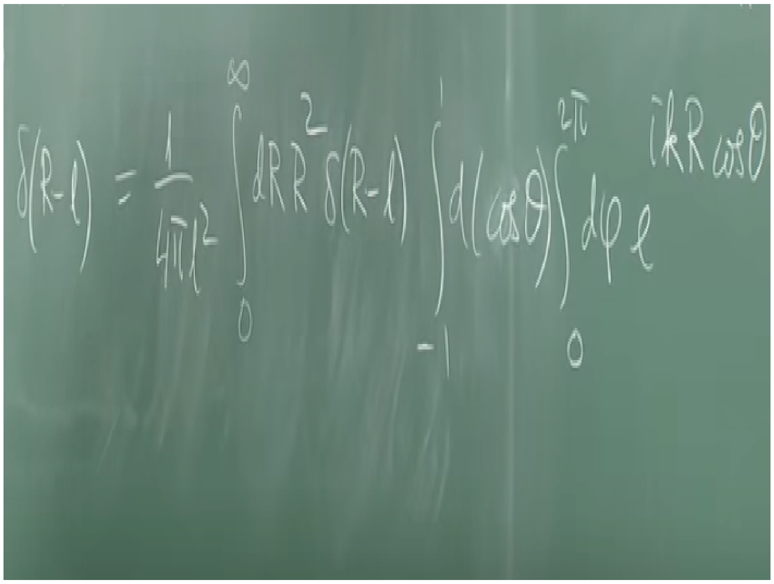
If I choose the axis in some random direction and this vector is k say and that is my vector R and that is the polar axis, the z axis, then if this R has got spherical polar coordinates theta and phi and this fellow has spherical polar coordinates theta prime and phi prime then the angle between these two the dihedral angle if you call it some gamma the addition the law of cosines which you learn in high school or where cos gamma right in spherical trigonometry we learn this formula for cosine.

It is the first it is the simplest instance of the addition theorem for Legendre polynomials which I presume everybody knows. So what is this equal to? What is cos gamma in terms of these guys? It is called the law of cosines. **"Professor - student conversation starts"** You learn this in spherical trigonometry right in high school or wherever in first year college or something yes? No? You do not learn it? You do not learn it in college? No, I am outdated. So now tell me. Yes, I mean UK probably knew it and I am sure he knew it. Various other people knew it but the internet generation does not. **"Professor - student conversation ends"**.

This is the addition theorem for p 1 of k, it is equal to cos theta cos theta prime plus sin. It is called the law of cosines alright. So if you got 2 vectors sticking out randomly in space and you know the polar coordinates of each of these the polar and azimuthal angles you can tell what the cosine of the dihedral angle between them is. So it is given by this guy okay. So what you have to do?

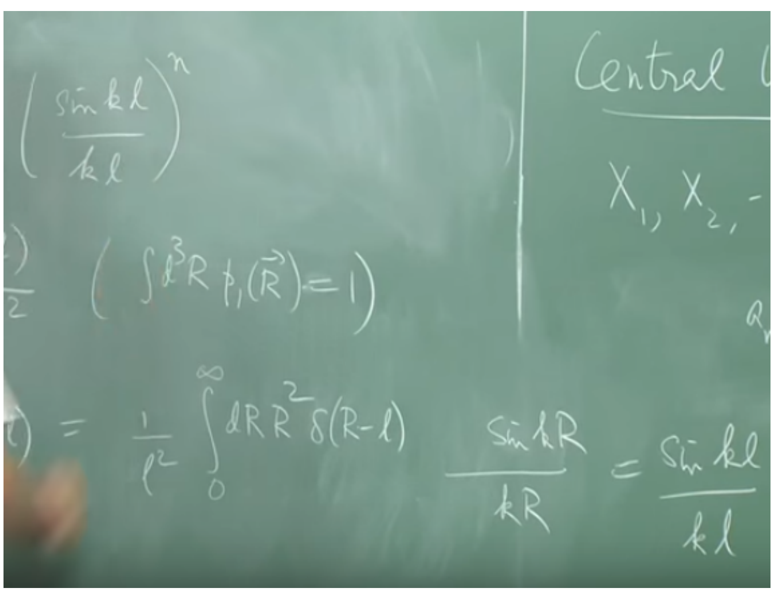
For doing this integral if you choose the polar coordinates some polar axis in some arbitrary direction is to say that for this vector this vector and then k dot R is k R cos gamma and for cos gamma you have to substitute this in the exponent and then try to do the integrals. So obviously horrible mess. Much simpler to exploit the fact that you have spherical symmetry in this case and there is a vector sticking out in the direction of k so you can without loss of generality choose the polar axis along k which means that the theta prime the theta for k this this quantity the prime is 0.

(Refer Slide Time: 50:05)



So this goes away and cos gamma becomes cos theta okay which immediately implies that you can write this as equal to 1 over 4 pi l square integral um 0 to infinity d R R square is sitting out there. There is a delta of R - l of course multiplied by e to the ik R cos theta, this is also an integral over theta. What is the range of theta, the polar angle? 0 to pi. So there is sin theta d theta but I like to write it as minus 1 to 1 d of cos theta okay and then there is an integral over phi, 0 to 2 pi d phi e to the ik R cos theta. Now you can do the phi integral trivially. You get a 2 pi factor.

(Refer Slide Time: 51:10)



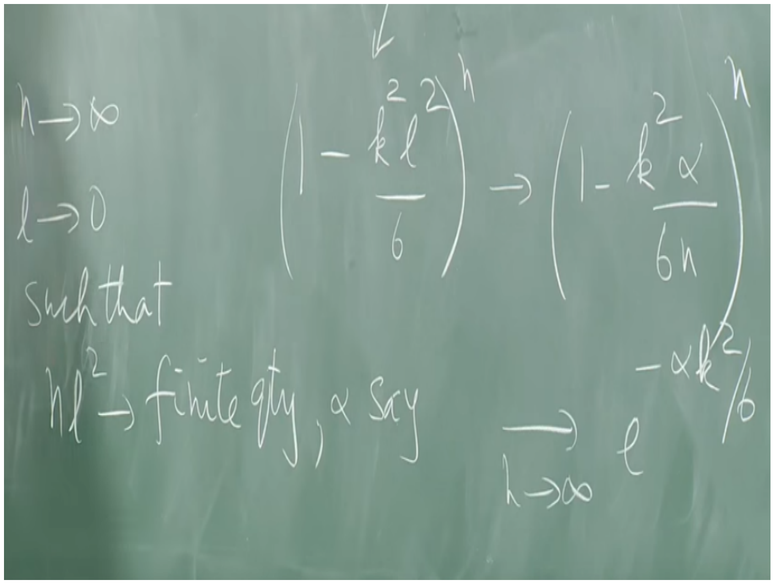
So this gives you a 2 pi and the phi goes away. This gives you an e to the ikR cos theta but cos theta is the integration variable. So it gives you e to the ikR - e to the - ikR over ikR. So this integral goes away and you get 2i sin kR over ikR okay. The 2 pi goes away with this 4 pi so it is 1 over l square. The 2 goes away here the i goes away here okay and you got a delta of R - l.

So if you set R = l this factor cancels against that and you are left with sin kl over kl okay. So this whole thing finally becomes something extremely simple. So this is equal to sin kl over kl and this whole thing here becomes sin kl over kl to the power n. That is it. So it says p(R) is an inverse fourier transform of this quantity sin kl over kl to the power n okay and it is well behaved as k goes to 0 this fellow goes to 1. So there is no singularity.

Otherwise you will be in trouble. You got to do an integration over k and if there is something in the denominator that blows up faster than 1 over k you are in trouble but that does not happen even though it is raised to the power n this guy here has got a well-defined limit okay and this sink function as you know what it behaves like and you raise it to the nth power. Now you can see that as you increase n if you increase n here, this number is always less than 1.

So it is essentially going to go to 0 unless expect for the contribution from the region near k = 0 where this thing goes to 1. So the dominant portion is near k = 0 and in what manner should you take a limit? You should take a limit when l goes to 0 and n goes to infinity in such a way that this guy goes to a finite limit. Then there is a respectable limit here. Now what does it do for small values of k?

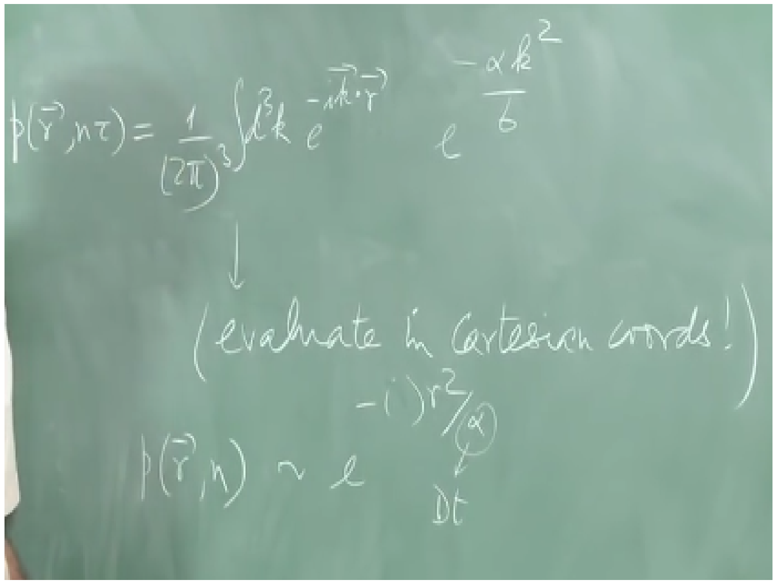
(Refer Slide Time: 53:58)



This goes like 1 - k square l square over 6 out here. So if you take a limit n goes to infinity, l goes to 0 such that nl square tends to a finite quantity. Then I can replace this l square by that finite quantity over n and raising it to and raising it to the power n is going to give me e to the minus k square and that is a Gaussian and it is inverse fourier transform is another Gaussian immediately.

So let us call this let us call this alpha say. So this fellow is this guy which is essentially 1 - k square so I put this equal to this so it is alpha over 6, 6 n to the power n and that tends in the limit n tends to limit e to the minus alpha k square over 6 and the inverse transform of the Gaussian is another Gaussian. That is an interesting exercise to do. How should you do that integral? How do you actually establish that, that it is a Gaussian?

(Refer Slide Time: 55:42)

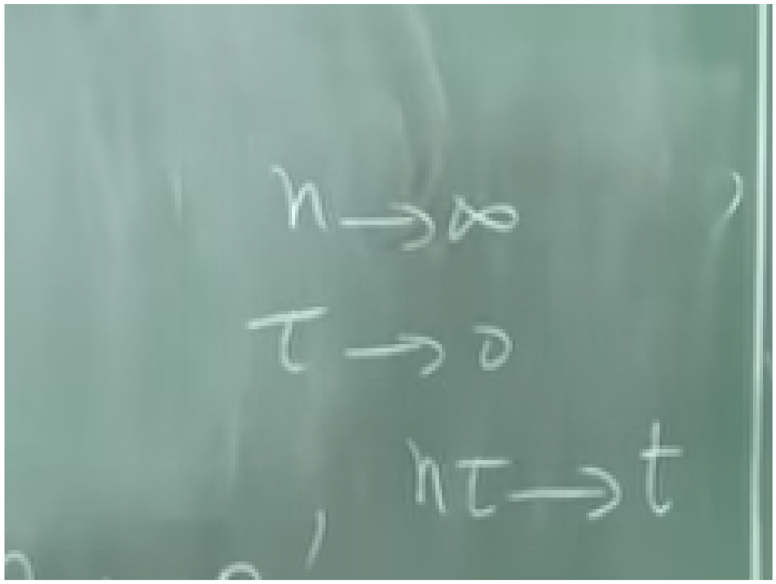


So let us suppose we have done this and we would finally want to find what is the inverse transform of e to the minus alpha k square over 6. What should I do now? I should not of course I know that I know for sure that the inverse transform of a Gaussian is a Gaussian even in high dimensions but what coordinate system should I choose for this. No, no because there is no need to. That makes it very hard, that makes it very hard if you try to choose polar coordinates now.

You should choose Cartesian coordinates because this here will become k 1 x + k 2 y + k 3 z and it will factor and that guy is already in factored form so it will be the product of 3 Gaussians which you can write down easily. So at this stage you should use Cartesian, go back and use Cartesian coordinates, evaluate in Cartesians. You could do it for the coordinates but and you will end up with something which say is p(r), n e to the minus something times r square.

Crucial point, where will alpha sit? Will it sit up here or down here? Downstairs ya. So this is some constant divided by alpha because the width here is 1 over alpha, the width here is alpha out here. We will see what that is when we do the diffusion equation.

(Refer Slide Time: 57:35)



This will turn out to be proportional to time because we are taking the limit n goes to infinity tau goes to 0 such that n tau goes to the time t and then this constant here will become Dt the diffusion constant times t. We will see this explicitly as we go here. So this is how random flights in the limit of 0 times step and 0 length step tends to the solution of a diffusion equation in continuous space and time eventually but it is actually a consequence of the central limit theorem; really arose because you add all these steps and you end up with.

Remember that each step was not distributed in a Gaussian or anything like that but this is true. But now I also mention that the individual random variables could have other distributions. We took them to have fixed lengths tends out it is of incredible generality. The individual steps need not have the same distribution at all. They need not have the same fixed length. They could in fact be distributed themselves.

The steps could themselves have different lengths distributed such that that distribution of step length has a finite variance. That is all you need. The result is also independent of the space dimensionality. Does not matter how many dimensions you are in in space. It is still true. It is still true apart from some numeric factors you would still get a Gaussian variable here. You could have distributions in time. You could have distributions in space.

As long as they are respectable distributions with finite variances and so on this Gaussian this emergence of this Gaussian form is completely robust. Only when very fundamental quantities like if I did for example I took a problem of random flights in which the different steps are of different lengths chosen from some drawn from some distribution themselves and that distribution does not have a finite variance.

This means that there do occur steps of arbitrarily large length such that the mean square length is not finite then you have a very different behaviour altogether. The Gaussian gets destroyed and this is what happens when you have the so called levy flights and so on. We will talk a little bit about that later on. This is like the strategy used by bacteria when they are in a nutrient solution they wander round taking random steps but then after a while they take one big hop, a long distance and then start doing it again.

Now if you have these clusters of clusters of clusters in a self-similar way it is possible to have random walks flights which have non Gaussian behaviour in the limit and they are also of practical importance will come back. But otherwise the lesson here is that the Gaussian is very robust, huge number of generalizations okay. I will stop here and we will resume from here.