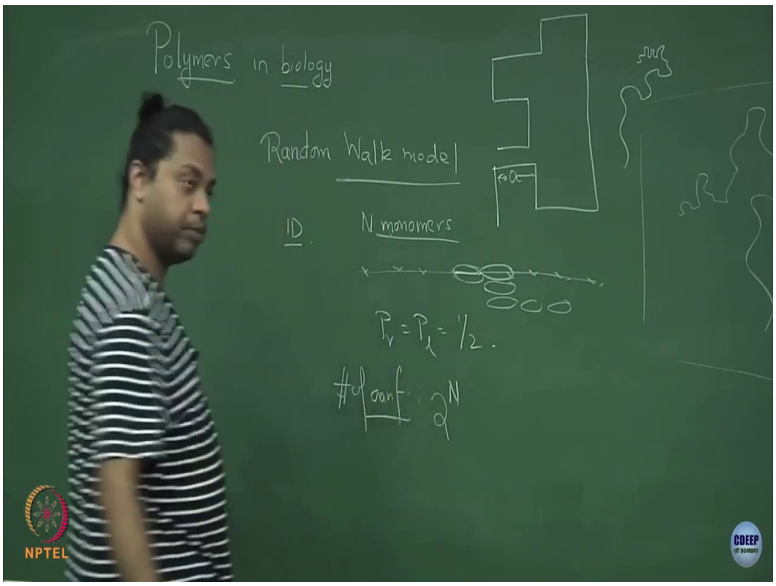
Physics of Biological Systems
Prof. Mithun Mitra
Department of Physics
Indian Institute of Technology, Bombay

Lecture - 29
Freely Jointed Polymer Model

So, what I thought we will start, with a slightly different topic which is polymers and biology.

(Refer Slide Time: 00:27)



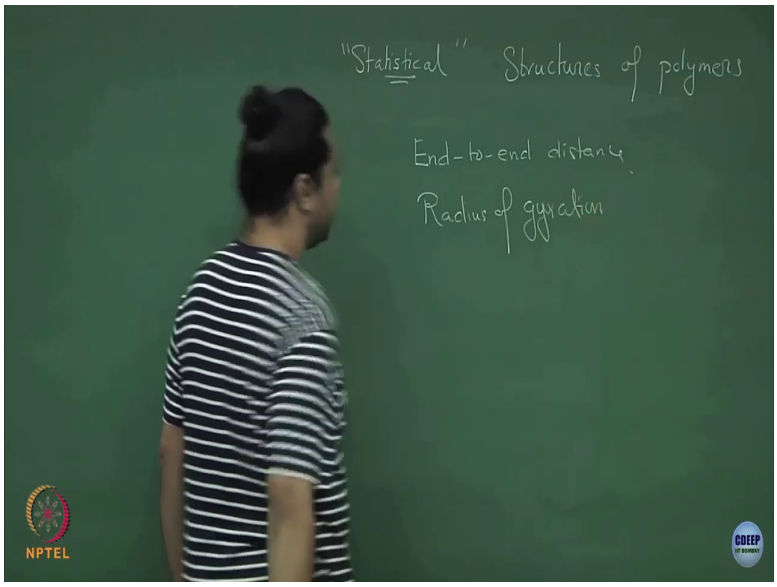
So, today, what I will try to do is try to start with very simple models of polymers and how to write statistical descriptions of polymers and then the next couple of classes will go more into examples of biological situations, where such calculations will be relevant. So, what are biopolymers for example, so, the most common one is of course, DNA right.

(Refer Slide Time: 01:05)

A polymer is of course, as you know it is something which is made up of repeating subunits, a long chain made up of some, repeating subunit. Monomers and when you connect many of these monomers together, what you get is some sort of a continuous chain which is my polymer.

So, DNA is of course, one such polymer then proteins are of also polymers for example, slightly different kind and as you will see the sort of description, that we will use for proteins is slightly different from what we use for DNA, but nevertheless these are also polymers. So, what I want to look at is how to describe sort of structures of these polymeric objects, but structures in a very specific sense. So, structures of polymers, but structures in the sense of statistical structures.

(Refer Slide Time: 01:56)



So, for example, for a protein you know, you have an amino acid sequence that comprises the protein and then if you do some X ray crystallography you can get some sort of a 3 dimensional structure for the protein. And some sort of a structure and that is what you think of is this, there is a structure of this particular protein let us say. But these are often somewhat static descriptions in reality in inside cells in vivo conditions the structure is not often not one fixed thing; it can go from one conformation to another depending on what environmental conditions pH salt and so on that you are changing.

Similarly, if you were to think of DNA for example, the DNA which is coiled inside a nucleus so, let us say this is my nucleus and here is my DNA chain, that is lying inside this. Often it is not that let us say whatever the 100 th monomer, the 100 th base pair of this DNA is at some location are 100 and 200 th one is somewhere else are 200 and so on. It is not a fixed structure in that sense. Thing sort of fluctuate because of thermal energy and you know often

you will find this monomer here or maybe at some later instants of time will you find that same monomer somewhere else and so on.

So, what we are interested what we are going to look at is somewhat statistical properties, that will characterize for us in a very coarse grained sense. Things about the structure of this polymer such as basically; things like what is the end to end distance of this polymer, things like what is the end to end distance or something like what is the radius of gyration of this polymer and I will define all of these terms, right. So, for example, if I say that well I take a bacterial chromosome.

(Refer Slide Time: 04:10)

Let us say which has some N 1 number of base pairs and I take a human chromosome, which as this some different N 2 number of base pairs. How large will this object be when I put it in solution, what is going to be a characteristic size scale, that describes this object versus

something that would describe that object. How does that vary with the number of monomers or the number of base pairs that you have, ok?.

So, questions like that is roughly, what we will be asking in this section, or how likely is it for example, for two monomers, which are some distance apart on the backbone, how likely is it for them to be close to one another something that is called the contact probability, ok. So, we will look at statistical sort of descriptions ensemble average descriptions of these structural quantities.

So, at a very basic level what we will start off with is actually, something that we have done which is the random walk model, which is the random walk model all right. We have done the random walk model in the case of transport. So, diffusion we started from a random walk and we went to the diffusion equation and that gave us time trajectories of particles, right.

So, if I consider a bacteria is doing random walk, it looks something like this in time the trajectory, but if I were to think in terms of sort of space then this could equivalently represent the confirmation of a polymer, right. So, the underlying mathematics is very similar; what was time in my earlier case R square grew with time will be replaced by the number of monomers in your polymer in this case. So, that is what we will start off with.

This is the most basic model in some sense of a polymer freely jointed chain, a random walk polymer or a Gaussian polymer and then we will try to see; how to build in more and more complexity as you go along, ok. So, let us start with this random walk polymer. So, let us say I have and again let me start off with 1D and I will generalize later, ok.

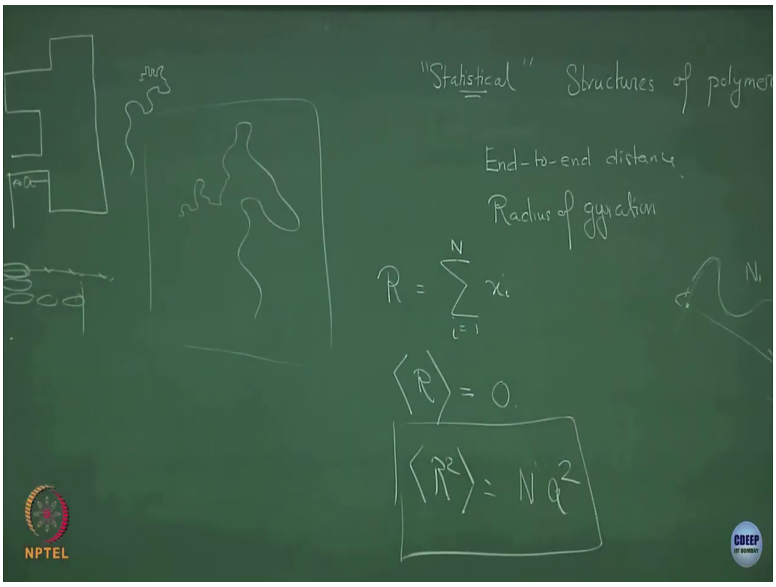
So, here is let us say my lattice on which I will do my random walk these are my lattice sites, all right. I start off from somewhere let us say here and I place one link, which is like one monomer, either to the right or to the left. So, let us say I have placed one link to the right then I placed another one to the right and then maybe I place one back to the left and then maybe 3 to the right. So, this is what a polymer in 1D would look like.

If you think in 3D or let us say 2D then of course, you can go up, but top bottom left right. So, if you keep placing links your polymer could look something like this, right. So, you take each step let us say a step of length a, which represents the size of the monomers of this polymer that comprise this polymer. And, then you place them along a lattice and then if you step far back and you look at it will so, if you if these step sizes are much smaller compared to the size of the polymer so, it will look something like a continuous chain, right.

So, I will work it out in 1D first. So, let us say that this polymer in 1D can take state steps to the right or to the left and let us say I take steps with equal probability. So, P r equal to P l is equal to half, ok. And, let us say this is a polymer with N subunits so, N monomers. So, this is a N number of monomers. So, the number of configurations that are possible of this polymer is what its 2 to the power of N, right. So, it is 2 to the power of N.

So, this is the total number of confirmation, this is the number of confirmations this is 2 to the power of N. And, like you start off with the my so, these each such each of these 2 to the power of N confirmations, would be one microstate of this system, right. And, if I start from the micro canonical ensemble, then each of these microstates a priori are equally likely to occur, ok. So, then I could ask that well I let us say I am interested in something like the end to end distance, of this polymer.

(Refer Slide Time: 08:40)



So, I construct I define my end to end distance R sorry, which is the sum over all these steps x i, i going from 1 to N, right I have N monomers. So, I take N such steps I do left to the right and at the end I have sum so, here for example, I have taken 1, 2, 3, 4, 5, 6 steps and this is my end to end distance 1, 2, 3, 4, right.

So, the end of that I will get some end to end distance and I can ask what is the expectation value of that and so on. So, for example, in the walk like this in a random walk polymer what would be the mean value of this end to end distance? It would be?

Student: (Refer Time: 09:21).

0, right because you have equal probability of going to the right or to the left. On the other hand what would be the variance of this random walk polymer would be?

Student: N square.

N square.

Student: (Refer Time: 09:40).

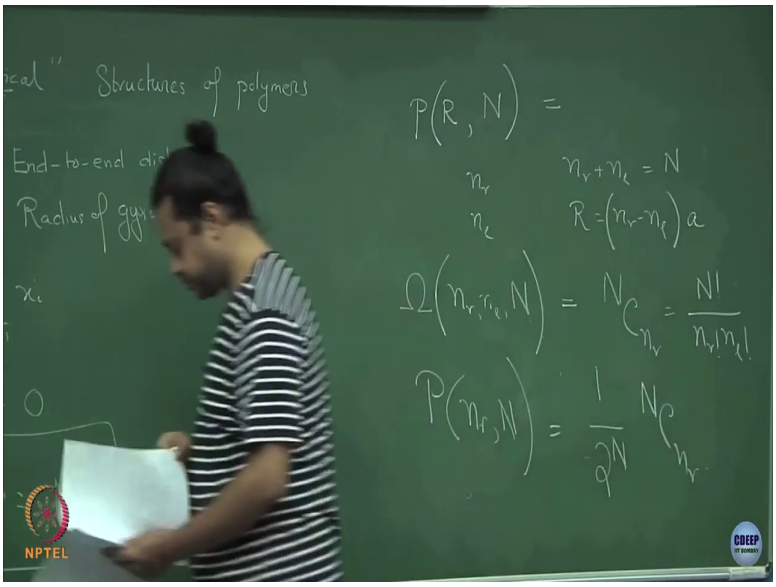
Into my step let us say the steps are a fast.

Student: That is N into step.

N into step length square right. So, if you look about you think about the correspondence this N is like my time. In my space diffusive works this was it R squared grow as time here that time is replaced by the number of monomer links the time space, keeping or placing. So, R square would grow is the number of monomers that you have in your polymer. So, this would be my R square, would be N times a square, ok. So, this is exactly similar to what we do in diffusion.

But of course, again we can do better than this we can do better than calculating just the first few moments; we can calculate the full distribution, all right. So, let us say I want to derive, what is the probability distribution?

(Refer Slide Time: 10:36)



That the end to end distance for a polymer, which is N monomers N subunits is R, ok. What would that be for? So, I have a polymer made up of N sub units or N monomers and at the end I want to calculate the probability that my end to end distance is R. So, if you think in terms of this polymer is a 1D polymer remember so, it can take steps to the right or to the left. So, let us say it has taken n r steps to the right and n l steps to the left, right.

Then can I express this n r and n l in terms of R and N I can write. So, n r plus n l would be the total number of steps, which would be the number of monomers. So, this would be N and what would be R? R would be let us say n r minus n l, right times of course, my step length it has to have the dimensions of length.

So, n r minus n l into a. So, now, it would be sort of becomes easier, that I have this 1D random walk of N steps out of which n r are to the right and n l are to the left and I know,

what is nr and n l in terms of R and N from these equations. So, then I could. So, I can write that what is the for example, the number of ways omega in which I can have n r steps to the right and n l steps to the left, such that my total number of steps is N and what is that? This is?

Student: N C.

N C n r.

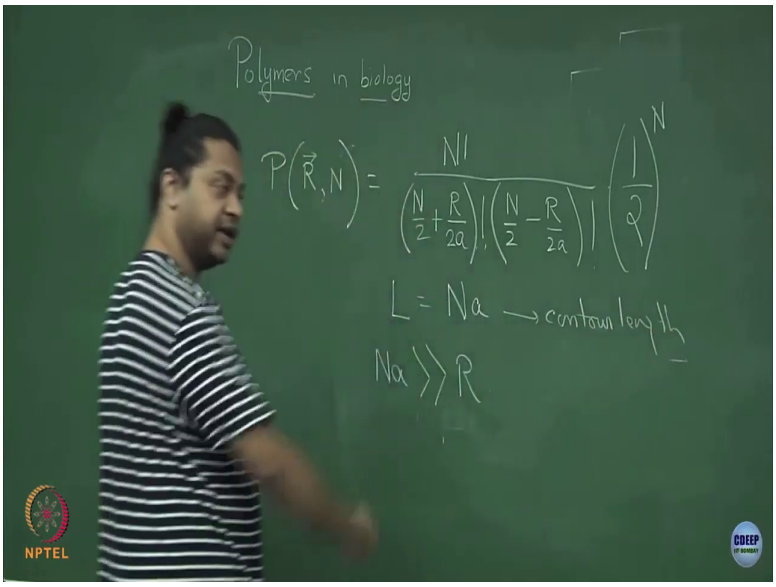
Student: (Refer Time: 12:30).

Or N C n l; whichever let me keep N these are not independent coordinates n l is just n minus n r. So, the N C nr right. Which is N factorial by n r factorial into n minus n r factorial so, n l factorial, right. There is a total number of ways in which I can take nr steps to the right given that the total number of steps is capital N. So, then what would be the probability of a configuration like this? What would be the probability of a configuration where I have taken n r steps to the right out of a total of capital N steps? Yes.

Student: (Refer Time: 13:24).

Divided by 2 to the power of capital N, that is the total number of configurations that are possible N C n r right. So, you can now, just write it in terms of R and N since that is those are the sort of physical quantities that we are interested in.

(Refer Slide Time: 13:53)



So, the probability of having an end to end distance of R in a polymer which has N steps is then N C n r. So, N factorial n r is what. So, N plus N by 2 plus R by 2a factorial and then N by 2 minus R by 2a, all right. I just solve for n r and n l from those two equations and I write it in terms of there.

So, this is the probability distribution to have an end to end distance R given that the total number of monomers is N. This is just a simple binomial distribution and of course, I forgotten this 1 by 2 to the power of N, ok. Now, let us make an approximation. So, I can treat this is sort of polymer is this continuous sort of an object, when the length of the polymer is much larger than these individual step lengths that you are taken, right.

So, for example, in the DNA each individual is step is like a base pair length and remember a base pair is around one third of a nanometer 0.33 nanometers and the total length of the DNA

is billions of base pairs, right. So, that is my L. So, L is like the total number in to the step length, and this so, let me now, work in the limit that this length is going to be much larger than the end to end distance that I have.

So, if I work in this limit that N a much much larger than R then I can massage this equation and get it to a slightly better form, ok. So, I this is in the limit basically, when the number of steps of this polymer is very large. Yes.

Student: (Refer Time: 16:00).

Student: (Refer Time: 16:04).

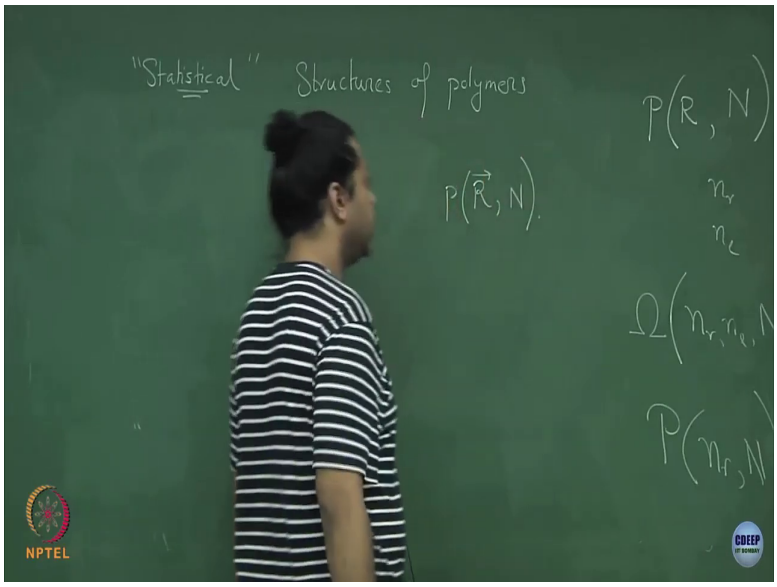
Is R assigned quantity?

Student: (Refer Time: 16:09).

Student: (Refer Time: 16:11).

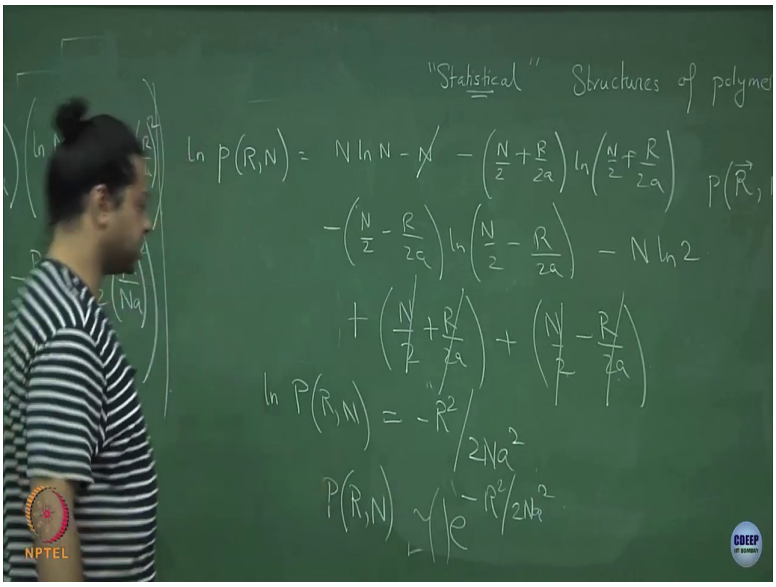
So, it is in principle a vector quantity.

(Refer Slide Time: 16:16)



So, ideally I should write P of R comma N, and in fact, when we write 3D polymers that is what we will do. So, if because, I was writing in 1D, but even there is a plus or a minus. So, that if you take more steps to the left; you will get a minus sign for the length, ok. So, if I do this sort of an approximation that is a very long polymer this contour this L is called the contour length of the polymer, the contour length is much larger than this end to end distance that you have. Then I can go ahead and sort a few Sterling's formula, right. So, let me just do that.

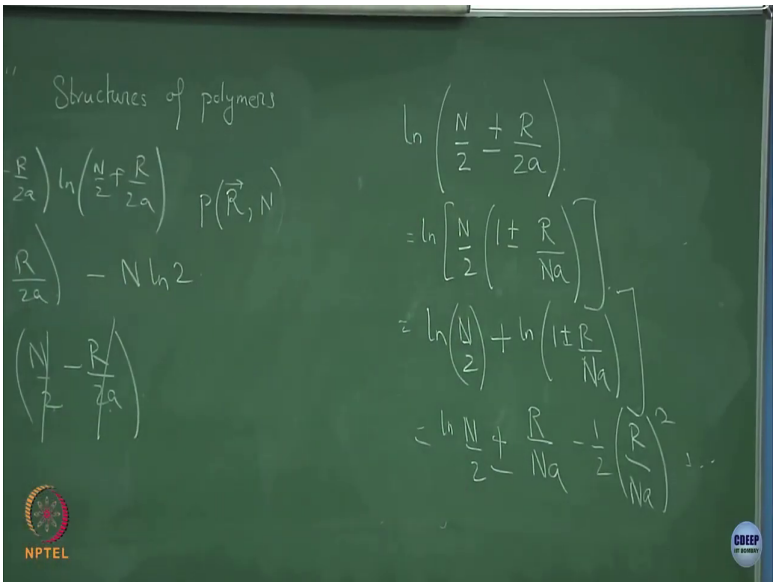
(Refer Slide Time: 16:59)



So, log of P log of P R comma N is going to be N log N N log N minus N minus N by 2 plus R by 2a log N by 2 plus R by 2a minus N by 2 minus R by 2a log N by 2 minus R by 2a and minus N log 2, ok. So, that is the log of the probability. These N terms anyway cancel hopefully; I have missed out something, right. I have missed out these things N log N minus N. So, I have missed out minus minus plus N by 2 plus R by 2a and one from here plus N by 2 minus R by 2a, right and N log N minus N, ok.

So, then this N will cancel out with this N by 2 and this N by 2. And, we can think of what these logs will look like actually, this R by 2 anyway cancels out that R by 2a. So, I do not bother about that, ok. So, now let us look at what this logs look like. So, I have terms which are like log of N by 2 plus or minus R by 2. So, if terms which are like log of N by 2 plus or minus R by 2a, right.

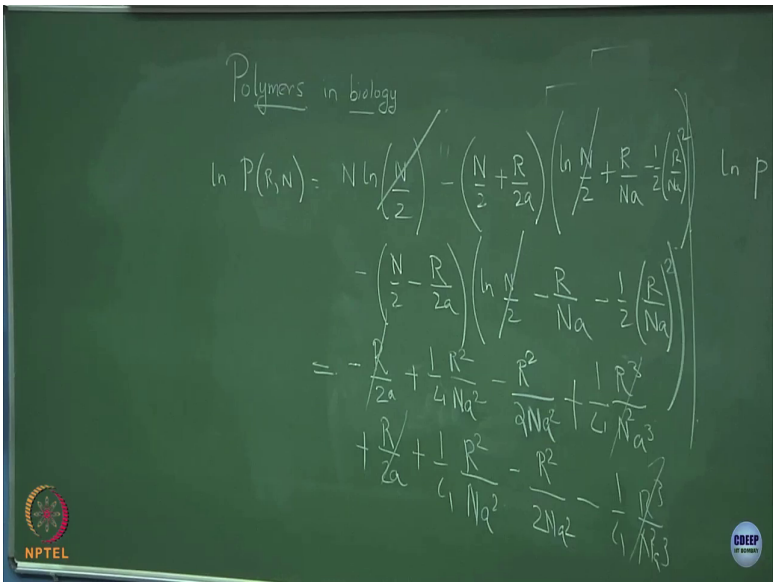
(Refer Slide Time: 18:45)



So, let me take that N by 2 common. So, this is let us say log of N by 2 1 plus minus R by Na, ok. So, this is log of N by 2 and then plus log of 1 plus or minus R by Na, And, remember that I am working in the limit that it is a very long polymer. So, R is much much smaller than Na. So, then I know what is the expansion of log 1, 1 plus x or 1 minus x. So, I can substitute with that.

So, let me say what I get is log of N by 2 plus or minus R by Na minus half R by Na whole squared plus higher order terms (Refer Time: 19:52), right. So, then let me substitute that here; let me keep this so, that I written from here. So, I have an N log N over here and I have a minus N log 2.

(Refer Slide Time: 20:18)



So, that is log of P of R comma N is N log N by 2, if I combine those two terms let me get that out of the way. Minus N by 2 plus R by 2a into log of that so, into log of N by 2 plus R by Na minus half R by Na whole square. Let us this term and then I have another minus N by 2 minus R by 2a log of N by 2 minus R by Na minus half R by whole square, right.

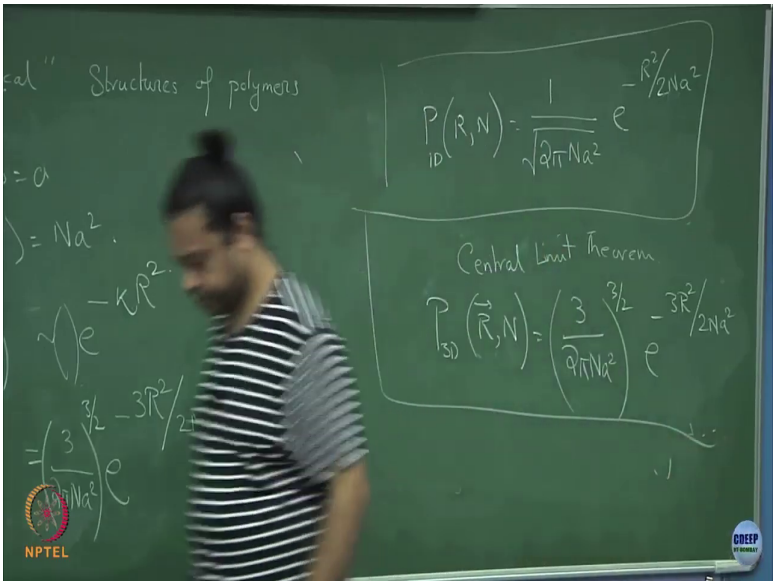
So, then what cancels now. This N by 2 log N by 2 and this one add up to cancel this first one which is an N log N by 2. So, this cancels out this cancels out. What else remain anything else cancels out; obviously, let us see. So, what do I get? Minus R by 2 a and minus R by 2a then minus minus plus plus 1 by 4 R square by Na square minus R square by 2 Na square and then here a plus 1 by 4 R cubed by N square a cubed hopefully; that will cancel out at some point.

And, then from here I get a minus minus plus R by 2a then another plus 1 by 4 R square by Na square then plus and a minus so, minus R square by 2 Na square and a plus and a minus so,

minus 1 by 4 R cube by N square a cube. So, hopefully; if I have done it correctly this cancels this cancels. So, what am I left with minus R square by 2 Na square, right. So, this gives me roughly log of P R comma N is equal to minus R square by 2 N a square, right. If you add up these 4 terms that is what you get.

So, basically then in this continuum limit; what I get is an expression for the probability, to the probability to have an end to end distance R for a polymer of length N goes as e to the power of minus R square by 2 Na square and of course, you can work out the what work out what is this factor in front by just demanding normalization, that the probability must be normalized and then you will get the full probability distribution P R of N for this 1 dimensional polymer.

(Refer Slide Time: 24:44)



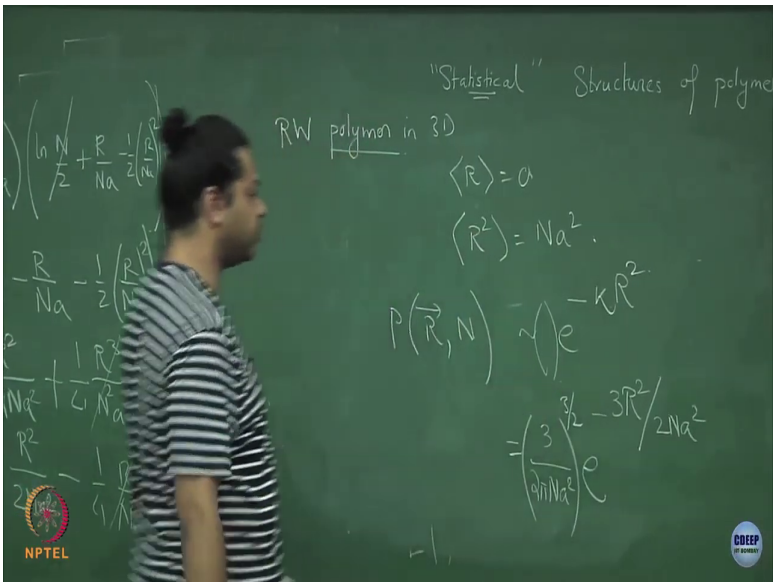
So, you will get that P of R comma N is 1 by square root 2 pi Na square e to the power of minus R square by 2 Na square. Could you have written down this distribution without doing this algebra? What theorem tells us what this distribution will look like?

Student: (Refer Time: 25:12).

Central limit theorem, right. If you take independent and identically distributed random variables then in the limit that the number of steps go to infinity, you should end up with a Gaussian distribution and that is what we have done; we ended up with a Gaussian distribution my remember my variance is Na square. So, this R square by 2 sigma squared, right so, this is like the standard Gaussian distribution. So, you could of got so, this is like consequence of the central limit theorem as well.

You can now, generalize this to this was artificial polymer in the sense of this was doing random walk in 1D you can, but you can generalize this to polymer let us say in 3 dimensions and again so, for a polymer so, for a random walk polymer in 3D random walk polymer let us say in 3 dimensions, because it is a random walk polymer again my R is going to be 0 and my R squared is going to be sum Na square, right.

(Refer Slide Time: 26:08)



Which means; that my probability distribution to have this end to end vector R, given some number of steps N, will again go as sum e to the power of minus something times Rs and should not write K something else times R square with an appropriate normally, with an appropriate normalization factor.

So, if you substitute the normalization and you substitute the variance, you can work out what this k is, what this Kappa is what this normalization factor is and then this; what should you get? So, what is the probability distribution in 3 dimensions? You get e to the power of minus 3 R square by 2 Na square over here and here you will get 3 by 2 pi N a square 2 pi Na square to the power of 3 halfs, ok.

So, it is like each component R x, R y, R z is doing their own random walk. So, for each of them you will get a minus R x square by Na square and so on. And, if you so, rather if you just

put in this variance and this normalization, if you will get this full probability distribution for the random walk in 3 dimensions, ok. So, this is one thing. So, let me just write this is P 1D and let me keep this P 3D as well of the end to end vector is 3 by 2 pi Na square to the power of 3 halfs, e to the power of minus 3 R square by 2 Na square. This is in 3D.