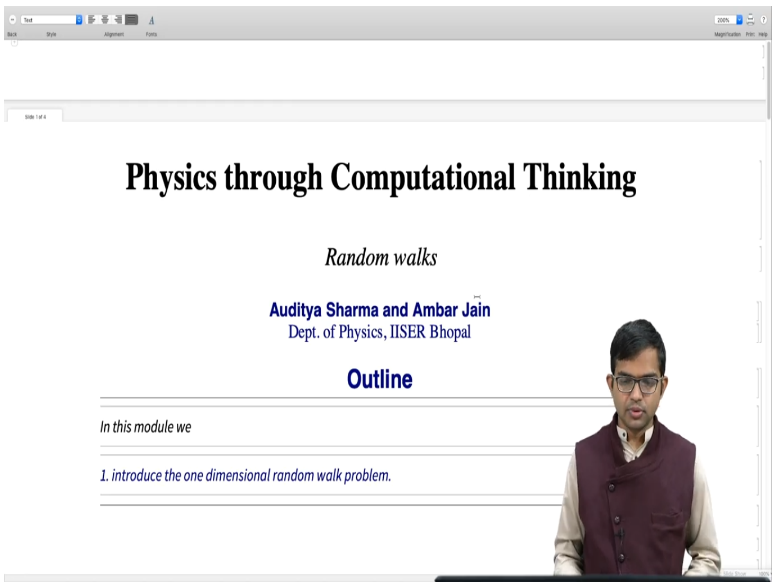
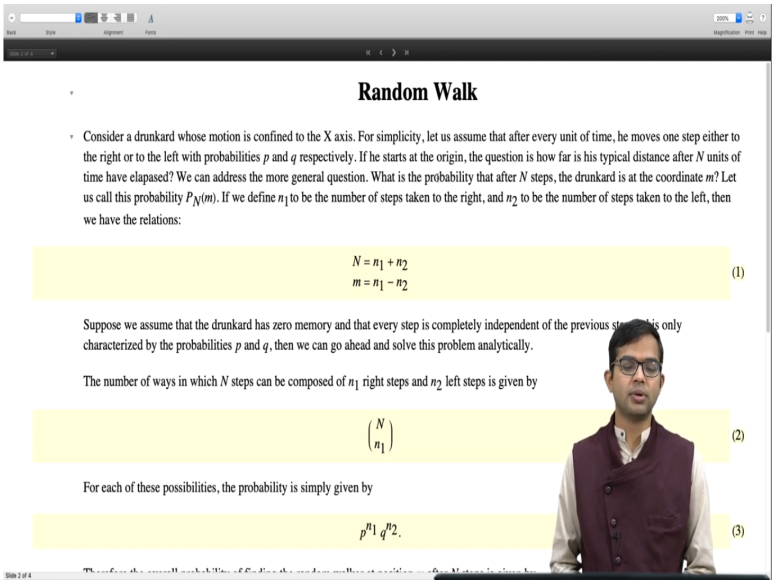
**Physics through Computational Thinking**
**Random Walks**
**By Dr. Auditya Sharma and Dr. Ambar Jain**
**Department of Physics**
**Indian Institute of Science Education and Research, Bhopal**

(Refer Slide Time: 0:27)




Hey, guys. So, in these next few modules, we are going to continue from where we left off last time. Where we introduce the technique, some basic techniques on Monte-Carlo simulations and then you know take consider somewhat more advanced approach to Monte-Carlo you can say. But I mean first of all we deal with the notion over random walks and do some theoretical analysis with them and then extract the main key results of random walks. Pretty much everybody should know I guess and try to make contact with this using some simulation, using Mathematica. So, in this module the agenda is to just introduce this random walk model in 1D. And then build up the model so that it can be put on a computer next. All right.

(Refer Slide Time: 01:40)




Okay so, think of a person who is on the origin. Let us say he is a drunkard who does not have a sense of purpose in life. So, with probability p he chooses to go to the right. So, let us say he is confined to live on the x axis. I mean of course you can generalize this to higher dimensions. Suppose he is confined to walk along the x axis but at every unit of time. So, he does not keep on moving, it is not continuous motion in time or space, we will consider this discrete problem first. So, he can be either at the origin or he can be at +1, +2, +3 so on, right in some units does not matter what. And time also most forward in units of one unit, which could be whatever, one second or a minute or whatever you want to choose.

And after every unit time. So, our friend makes a choice whether to go right or go left. And so, let's say after N number of steps. So, we want to ask how what is a typical distance which our friend would manage to cover, if he is given you know N steps? So, this is the type of question. So, probably its a stochastic random variable. It is not going to be, there is no definite answer for exactly where he is located is not something that you can calculate because there is a, even he does not know where he will be exactly, precisely after and such.

But it turns out that do you still know a surprising amount of information you know, in terms of, you know the average location on average distance from the origin and so on. You can ask lot of questions and there are very precise answers to some of these questions one can formulate in stochastic way. So, that is the beauty of you know stochastic methods already comes out in something very simple like the random walk problem.

So, let us say that n1 steps are taken to the right and n2 steps are taken to the left. So, clearly N = n1 + n2. Because if n1 is if there is no option of not taking a step, that is the simplest model to consider. You can also make this model little more complicated and say that you know with some probability you know some other probability which is just such that p + q plus this new probability are introducing is equal to 1.

The person stays at a location. You can say that with probability p, he moves to the right and with probability q he moves to the left and probability $\acute{q}$ , let us say he stays where he is so that p + q + $\acute{q}$ = 1. So, that option we are not considering here. We saying that p + q = 1 and so the person either move to the right or to the left. At in every time instance when a decision needs to be made.

So, therefore N = n1 + n2. So, a quantity of interest for us is n, which is n1 - n2. So, you are interested in you know, what is a typical distance that our friend manages to cover, as a function of time. It could be a negative number, but you will see in some time that actually we care about the modulus of our or even more convenient is actually $m^2$

So, you are interested in finding out what is a typical distance squared of you know distance covered by this person as a function of N, if you move, if you take M number of steps, what is a typical M. It is a distance, in fact that you can, you can compute the full distribution itself. The probability distribution. So, let us say that there is no, no memory in the motion of the random walker; is that every instant of time, the person basically does not care about what they did in the previous step.

So, they make, you can think of it as tossing a coin every time he goes to a step then the before he makes up his mind next time he just tosses the coin and with probability p the toss is going to give him heads and him go to the right. And with probability q is going to get tails and he is going to move to the left. So, these kinds of, you know, the step of a model can well be applied to the motion of stock prices, for example.

But there it turns out that may this kind of a memory less model. Perhaps it is not the most accurate, although some people have argued there are you know mathematical models of the of stock prices where people argue that it is basically a completely random walk. And you know, the anecdotes where, you know, there was group of academicians who just took up, cooked up data using random walk and then do random walks. And then they just, you know, showed this data to some financial analysts and they predicted all this is we see a pattern here and then you should put your money in here and it is going to go up or whatever.

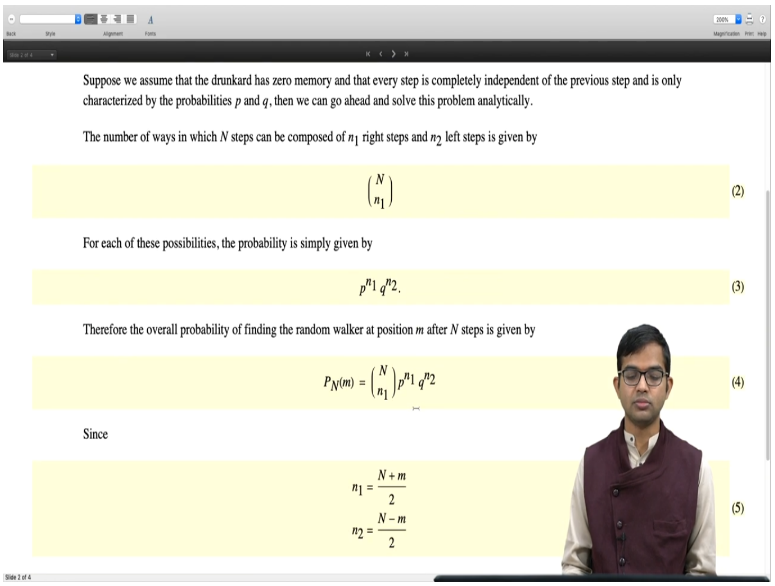
So, and basically their argument was that, their data generated from a random walks were completely indistinguishable from realistic data where people claim that, you know, there are there is some hidden order there, which is predictable. Predictability is what is of great interest in the context of stock prices. But it turns out that, in fact, there are other academic discussions where people have argued that in fact there is. It is not really random walk. There is some predictability involved.

And in fact, that has to do with the fact that there is often some memory in your in the motion of stock price or whatever. At every instant it is not going to be completely cut off from what happened in the prior instance, so it is typically there is going to be some speed associated with it and keeps going in a certain direction. And then, of course, there is a randomness associated with it as well. And there it, you know if sufficient people start seeing a trend, then they try to exploit it and therefore the trend gets reversed there.

So, in some sense, the whole system enforces a certain amount of randomness is an essential aspect of that system. But okay, so that is just a small digression into how a model of this kind can be applied in the context of stock markets. But for our model, we were very simple. So, we do not have any memory effects here at every instant, it is a fresh coin which is tossed with probability p you go to the right and with probability q go to the left.

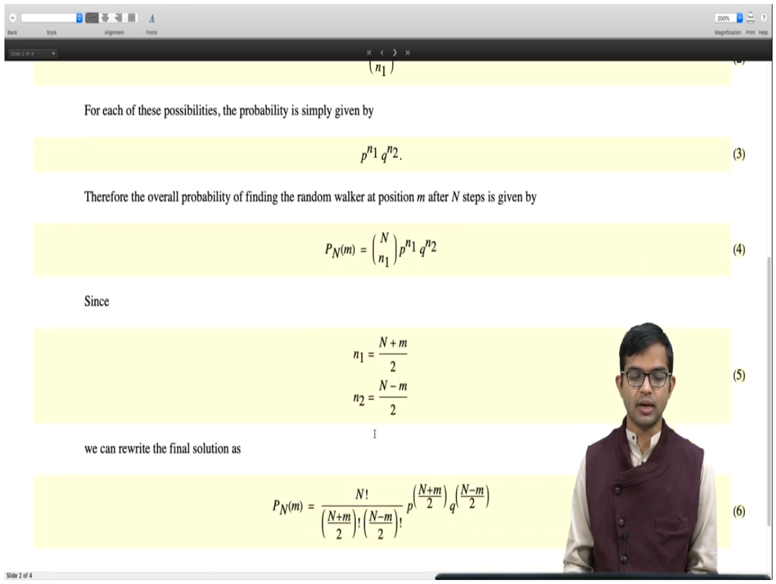
So, the number of ways in which N steps can be composed of small n1 right steps and small n2 left steps is simply N choose n1. Now for each of these possibilities, the probability is given by, you know p times, p times, p times p so on up n1 times and q times, q times q so on n2 times. So, you have to tag along all these various possibilities. You have to count them, which we already did and $\binom{N}{n1} p^{n1} q^{n2}$.

(Refer Slide Time: 09:39)




So, therefore in fact that is all there is to it. Therefore, the overall probability of finding the random walker at position m after N steps is simply given by $\binom{N}{n1} p^{n1} q^{n2}$. So, this is already a bit of a result. So, now we can invert this n1 and n2 write n1 as (N + m)/2. And n2 as (N - m)/2 all I have done is just to solve for n1 and n2 an equation one straightforward.

(Refer Slide Time: 09:54)




And then we can write the final solution in a more convenient form. So, ultimately you are interested in probability of you know be at position m given that you have taken N steps is given by this expression $\frac{N!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!} p^{\frac{N+m}{2}} q^{\frac{N-m}{2}}$. So, of course here, notice that if N is even, then m is also going to be even.

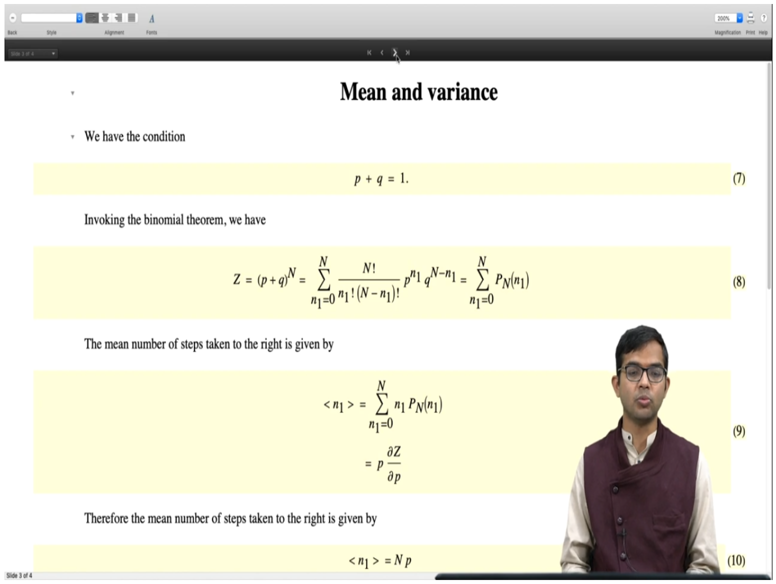
So, you cannot have m odd. So, for example, so, let us just think of the case where you just walk once you take one step, if you take one step, then for sure you can, you could have gone either to the right or to the left. So, there is no question of the walker being at zero. So, the so you can ask whether they at 1 or whether they are at 3. So, after two steps, then for sure, you are guaranteed that your walker is going to be at an even position, there is no odd there is no possibility of them being at the odd position.

And the third step, the guarantee to be at either 1 or -1 or +3 or -3. All these possibilities arise. But for sure, they cannot be at the origin and they cannot be 2 or 4 or whatever. So, this is ingrained into this equation. So, OK so, here is a homework question for you. How do you how can you be sure that this is a reasonable probability here? So what? Can you check this? Is it already normalized or is there some more work one has to do to normalize it?

Some claiming that it is already normalized because I have used a nice argument involved in probabilities to get to another probability and I believe that I have done my counting correctly. But you should check this and the way to check this is just some over all possibilities here. So, probability is a quantity which must add up to 1 if you know some over the entire sample space in this case. So, m can only go from the maximum value of small m is of course N. So, if all steps are taken to the, to the right, then n1 - n2 = m because n2 is zero.

Now on the other hand, if so, the smallest value of N in magnitude is zero but it can also be -N. If all steps were taken to the left, it could be -N. So, what you should do is just some over all possibilities of m. First start small, take N to 2 and 3 and 4 and so on. And then play with it for yourself and then you see a pattern. Why, this should just sum up to 1. And then you will see that there is a simple, direct way to solve for this. That is homework for you.

(Refer Slide Time: 13:35)




Okay, let us move on. So, what I want to do is show you some tricks here. Some of you might have seen this but it is a, it's a very clever. You know, the way to compute mean and variance of this distribution. So, we have the condition p + q = 1. I have said it, but it also sort of implicitly clear that there is no possibility of the random walker at any instant of time to not make a choice. He either goes to the right or the left. Therefore, p + q = 1.

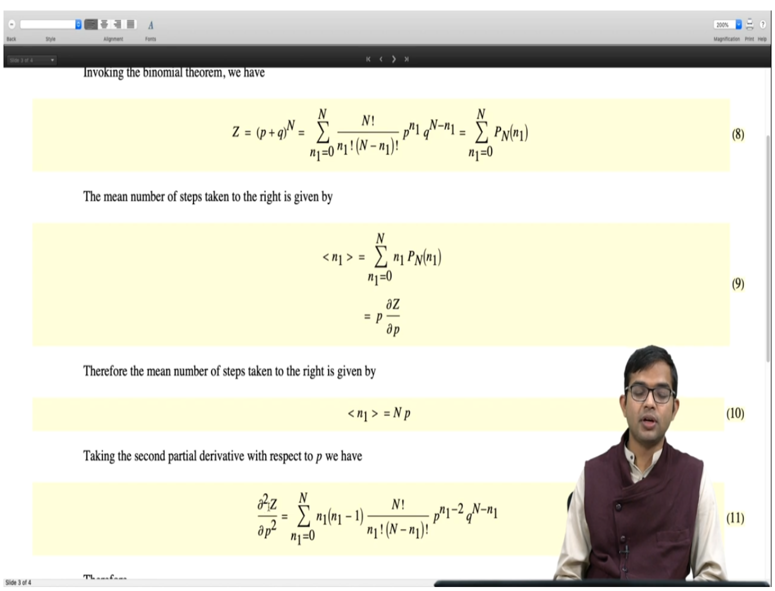
Now consider this quantity z. I mean, you will see in a moment why I am doing this at all. So, I have $z=(p+q)^N$ for now you think of this as you know, just some two variables p and q. You are not already putting p + q = 1. We will do that later on. So, it is just some trickery involved here. So, I expand $(p+q)^N$ and write as this series, $\sum_{n1=0}^{N} \cdot \frac{N!}{n1!(N-n1)!} p^{n1} q^{N-n1}$ . So, this let me call this as equal to $\sum_{n1=0}^{N} P_N(n1)$ because you see this is exactly the expression I have for, I have for my $P_N(m)$ or n1 is the same.

So, so if you are paying attention, this already gives you proof of why this probability is normalized? You see that? So, if you see that, do not see that spend a minute, you can pause the video here and go back to this homework problem and then connect to this expression for z I have. And then quickly check that indeed. It is a normalized probability distribution, probably, okay.

So, the mean number of steps taken to the right is just given by summation over n1 and it is you have to do $n1\times P_N(n1)$. Which you can see, that now you see the power of this expanding this type of a quantity called z. See, you think of this z as simply an algebraic expression for a moment, you are not supposed to put in p + q = 1. Do not think of it as I am taking the derivative of one with respect just think of it as a derivative of some function of two variables p and q and then you just do $\partial z/\partial p$.

Then you see that, you will get $n1\times p^{n1-1}$ so you this n1 here is you know, it appears in the sum when you do this do $\partial z/\partial p$ and then but you are losing a power of p. So, you have to multiply by p so that you adjust appropriately for all the powers. And then you see you can evaluate this expression. So, what I would suggest is just do do $\partial z/\partial p$ and then multiplied by p and then check for yourself that indeed that is equal to this $\sum_{n1=0}^{N} n1\times P_N(n1)$.

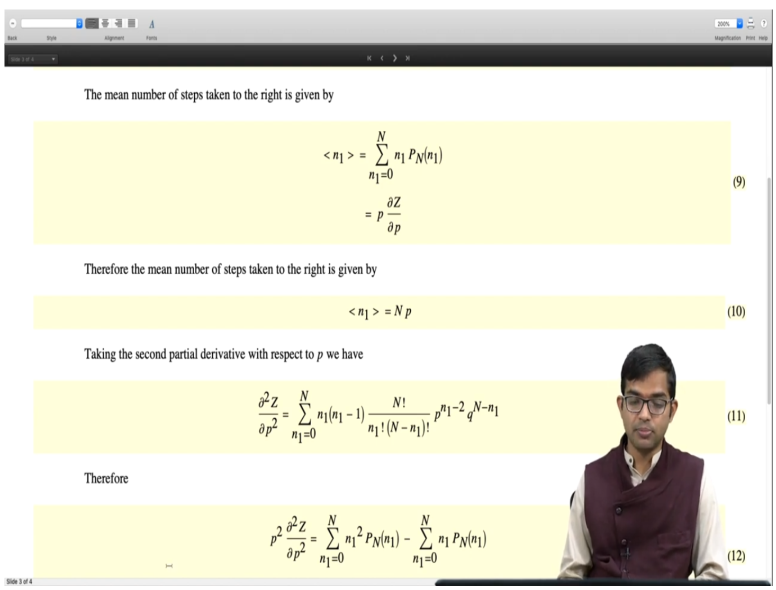
(Refer Slide Time: 17:22)




And then you can of course very quickly evaluate it. Now that you have this. You can evaluate it and then you see that the mean number of steps taken to the right is simply given by average of n1 = Np. Because $\partial z/\partial p$ is going to be just N. You see that? So, it is just you can if you take a derivative it is $N(p+q)^{N-1}$. And now you put in p + q =1. So, you will just remain N for $\partial z/\partial p$ it, just Np. It is some clever trickery. There must surely there are other ways of doing this as well. But this is one nice way of doing it.

Okay, you can go ahead and actually find the second moment as well. It is a great importance here is to find the average of $n1^2$ So, here you just simply take the derivative of with respect to p once again of $\partial z/\partial p$ So, in other words, you are looking at $\partial^2 z/\partial p^2$ the second derivative partial derivative. And so now you see that you have n1(n1-1). So, actually you are interested in $n1^2$. Now, but we will see that using n1 information and you know, some trickery off this of the left-hand side. We can actually evaluate this.

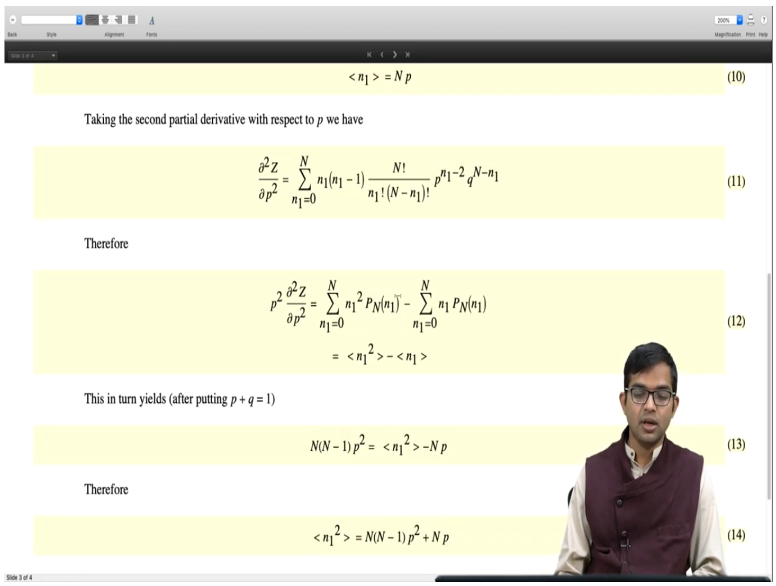
(Refer Slide Time: 17:59)




So, you should check this. What does it involve? Simply taking a derivative and over this sum and writing it explicitly like this. And then you multiply by $p^2$. When you multiply by $p^2$, you are going to get $p^{n1}$. You have compensated appropriately for the loss in the power of p, and then you have to you know, collect these terms while on the one hand, you have $n1^2$ but you have -n1.

So, you should verify this and check that indeed you have $n1^2$ times of the probability n of n1 - n1 X $Pn(n1)$ , Right. And in fact, this quantity $p^2$ , $\partial^2 z/\partial p^2$ is actually nothing but the variance of this random variable n1. n1 is the number of steps taken to the right. Okay so, this in turn yields after putting p + q = 1. All you have to do is take this derivative $N(p+q)^{N-1}$ .

(Refer Slide Time: 19:08)




And then again, take another derivative will get $N(N-1)(p+q)^{N-2}$ . And then you put p + q = 1 right. You should reserve this stuff of putting p + q = 1 to the last, right. This is some trickery to ponder over this. And then you will see why it makes sense, Right. Then you see once you have this, you just plug in here $N(N-1)p^2$ . So, there is a second p also which comes in here, which comes in because you have $p^2$ here, $p^2$ sitting here.

So, you have $N(N-1)p^2 = \langle n1 \rangle^2 - Np$ because you already evaluated this and therefore you see that average of n1 square is actually now known. It is going to be just $N(N-1)p^2 + Np$ . So, once we have this expression, we will go back and we will find out what is, we will find out what is... First of all, let us find the variance $\sigma^2 = \langle n1^2 \rangle - \langle n1 \rangle^2$ .

So, check that this indeed make sense. So finally, the expression is just $N(N-1)p^2+Np-N^2p^2$ which after some manipulation can be shown to be just simply Npq okay. So, this is for you have to check this all you have to do is use the fact that p + q = 1 and then you can show this. So, this is homework to go from equation fifteen to equation sixteen using p plus q equal to 1, okay.

So, we have obtained an expression for the mean of n1. We obtain an expression for the mean of the square of n1. And therefore, using that we have on the variance in n1. Now, actually we are interested in the mean and variance of the random variable m which is, you know, the typical distance, which is a measure of the typical distance you are from the origin. So, it is actually n1-n2. You can go ahead and compute this you know in terms of n2 and then you put in variable m as n1-n2. And that is one way of doing it.

So, another way is to do the following. You just solve for m in terms of n1. So, since n is it is already given here. So, now from which we have m is equal to actually 2 you eliminate n2 because n2 we have not really found out the information about its mean and variance. So, you just compute m = 2n1 - N. Now the $\langle m\rangle = 2\langle n1\rangle - N$ and $\langle n1\rangle = Np$.

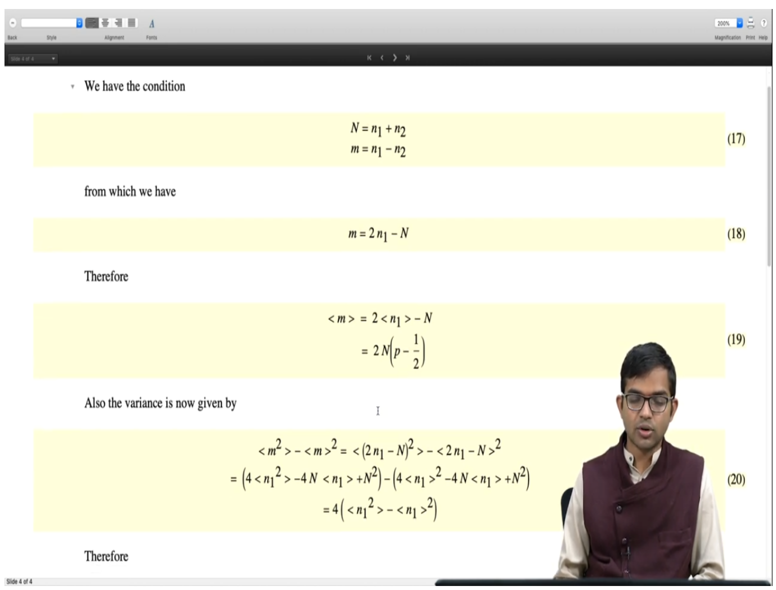
By the way, the fact that $\langle n1\rangle = Np$ is also very, very reasonable physically. So, basically what are you saying? You are saying that should take one step to the right, the probability of going to the right is p. So, the probability that after N steps, the probability that you have taken you know N. What is a number of, what is a average number of steps you have taken to the right is just going to be Np.

Because in another way of thinking of that is, it is a bit like you know, Monte Carlo in some sense. So, the average you know the probability. If suppose you did not know what p is? How would you find out if you had a machine which would generate a random walk for you? You just simply count after the first step. You know you will find out whether they went to right or not. And then you, you get data for N steps and you ask yourself how many of these went to the right? So, you will find a certain number which actually is n1.

And then so you will just divide n1/N and say that should be the probability of going to the right. So, that is, that is in the spirit of a Monte-Carlo. This how we found out what is what is you know like how many points fall inside a region, Right. If you want to compute the area of a certain region, for example, with respect to the total number of darts thrown and so on. And so, in fact, if you want to find out whether how biased the coin is? This would be the method right. You would just toss it and then you find that it has you get head to tails. Then you find that again. Then you keep on counting.

So, if you take a very large number of trials of this and you are finding that actually is sixty-five out of hundred are heads. So, then you say that it is a biased coin. It is not a it is not point five and point five but actually is point six one. So, that is going on here. So, now we have. Once you have n1 average of n1, you can also quickly find the average of m. So, we just simply given by 2N(p-1/2).

(Refer Slide Time: 24:12)




So, the variance is a little more involved the computation of variance, but not necessarily out of reach. All you have to do is expand this. So, when you have $\langle(2n1 - N)^2\rangle - \langle 2n1 - N\rangle^2$ . Then you carefully pull out these factors. So, then you have four times, you know, four can come out of the average. Does not matter. So, $4\langle n1^2\rangle - 4N\langle n1\rangle + N2 - 4\langle n1\rangle^2 + 4N\langle N1\rangle - N^2$ .

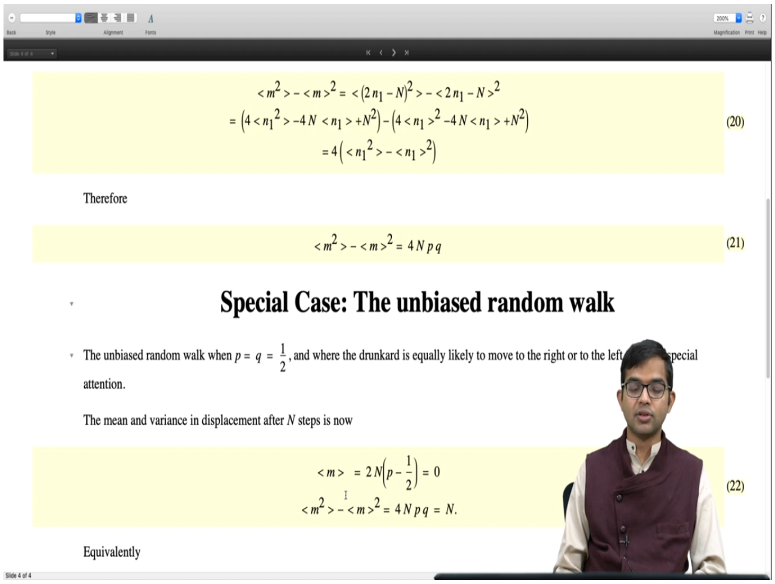
If you collect all these terms carefully so you will find that in fact the variance of m is related to the variance of n1. And it is just a matter of a factor of four. So, just four times the variance n1 is the variance of m. So, you can immediately write down the answer. Which is the variance in m is actually simply 4Npq. And it is mean is 2N(p-1/2). So, one immediate thing you would observe is of course what happens when p=1/2 for the unbiased coin toss or a random walker.

So basically, if somebody who treats both directions you know equally carefully or equally carelessly say they do not care where they going. So, with probability half if he is going to the right and with probability half, he goes to the left. Then on average, you expect that after N number of steps he has basically mode not at all. It would be very close to the origin. So, the mean is indeed at the origin. There is no drift.

And but the spread is, it turns out, is actually the more interesting quantity when you consider this. So, although the mean of this random variable is zero, their variance will turn out to be N so in fact, so that is typically how far away he is going to be from the origin. He is not going to be sitting on the origin like the mean seems to suggest. So, in fact it is the variance that count. So, you have to take the square root of this. And so, he is actually going to be very likely close to a distance of $\sqrt{N}$ away.

It could be on the right side or to the left that cannot be ascertained. But he is not going to be on the origin. So, this is actually an example of how, you know the mean of a distribution is not necessarily always a most meaningful quantity. Sometimes it is actually the variance or the root the root of the variance, which is the standard deviation, which carries more information, more useful information than the mean. And so, this is one such context where actually the typical distance of your random walker after n steps is actually $\sqrt{N}$ .

(Refer Slide Time: 27:13)

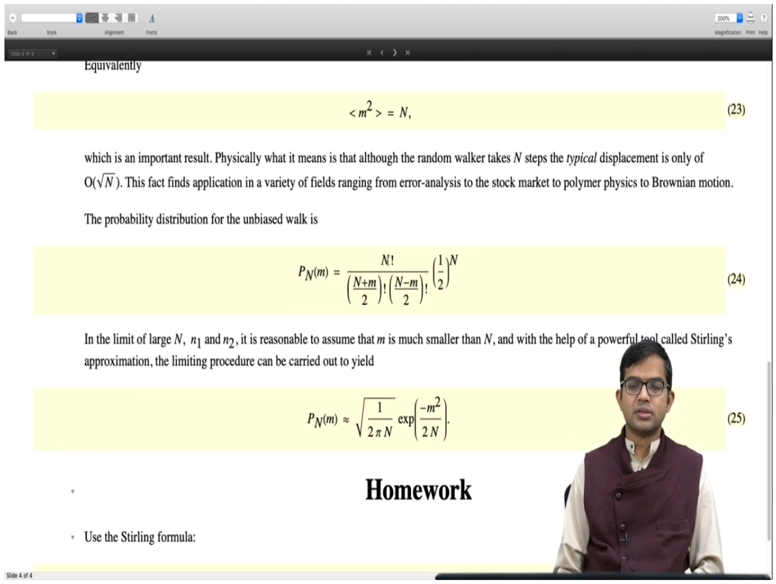



So, let us look at this, this is a special case. I already talked about this. But let us do it more formally now. So, if I put p = q = 1/2. So, indeed $\langle m \rangle = 0$, $\langle m^2 \rangle - \langle m \rangle^2 = N$. And so equivalently this is a very, very important result.

Ultimately with all this simple set of arguments which seem very logical we have manage to show very, very important result which appears in all kinds of contexts whenever there is a stochastic motion involved. And that result is just $\langle m^2 \rangle = N$. So, if you take N steps, in fact, basically, you are moving only $\sqrt{m}$. If you are a random walker. So, if you have no purpose in life and if you are just simply exploring every direction so you can imagine that you would although you take N steps, you would only go $\sqrt{N}$.

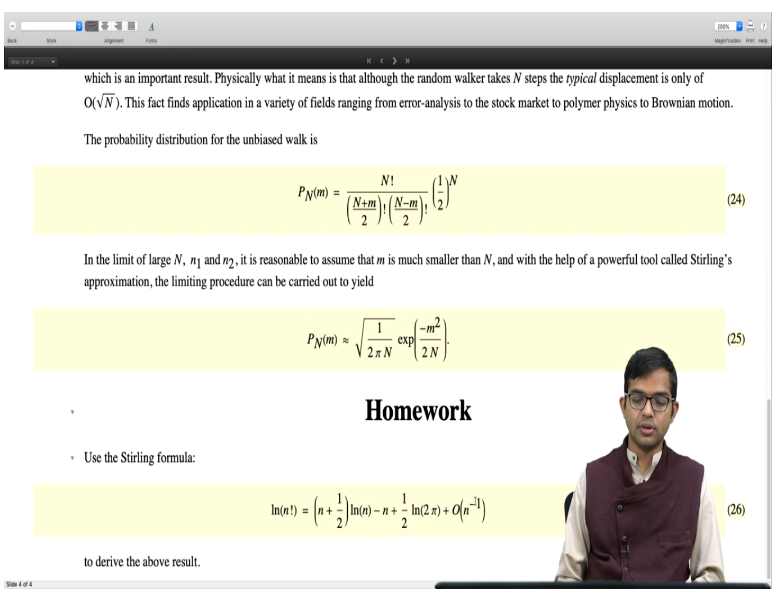
If and the best you can do is, of course to go N, right. So, if it is order N it means that you are more or less going you know all, all your steps, every step that you take does count for your overall motion. So, and that is what is called a ballistic motion. And this is called diffusive motion for a reason, which we will show in a movement. So, it is connected to the diffusion equation and so on, Okay.

(Refer Slide Time: 28:47)




So, so the main result from this discussion is $\langle m^2 \rangle = N$. And in fact, you can do better and get a distribution $P_N(m)$ is equal to the binomial distribution, which using Stirling's approximation, which is going to be homework for you to show is actually the Gaussian distribution. It is a sharply picked distribution. The more steps you take, the more sharply pick your around your, you know origin and the spread is going to become smaller and smaller. You will see this.

(Refer Slide Time: 29:20)



So, this is a stunning approximation is also listed here. So, log of n factorial is equal is given by this complicated looking equation. But it is actually a very nice tool. We all have to do is just plug this in and take care of all the simplifications. And then you can show that it is a Gaussian distribution, All right. Thank you.