Computational Physics

Dr. Apratim Chatterji
Dr. Prasenjit Ghosh
Department of Physics
Indian Institute of Science Education and Research, Pune

Lecture – 55

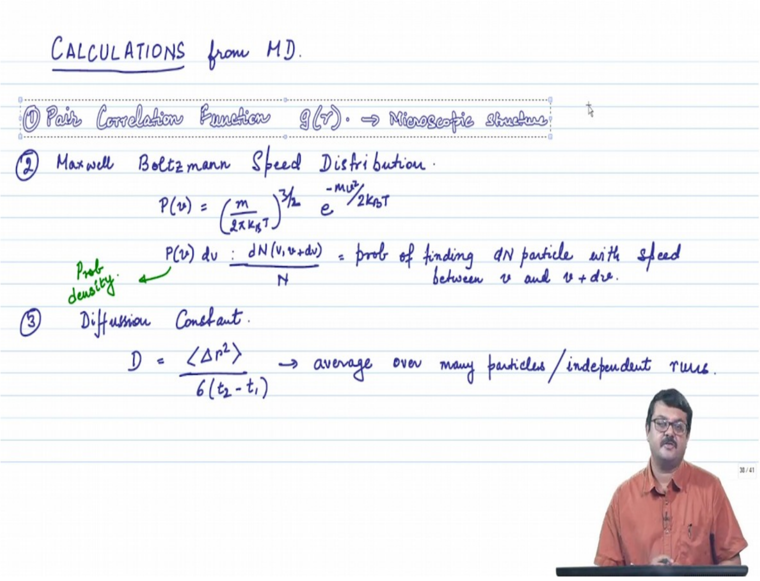
Molecular Dynamics: Calculation of Thermodynamics Quantities

Welcome back to the class. So, in the last class you learned about neighbor lists, which would help you increase the speed of molecular dynamics simulations significantly a factor of 10 or 15. And that would enable you to run your molecular dynamics code for large number of iterations, which means for long times.

So, which means that you can follow the trajectory of n particles where n could be 1000 or 10000 over an extremely long period of time. And why do you want to follow it over a large period of time, so that you can calculate average statistical quantities from it, right. So, you could also run the code without your neighbor list and then well you can do up till 1000 particles at the most and even that takes quite some amount of time if you run it long.

But, if you have above 1000 particles running a molecular dynamics code without neighbor list is practically impossibility; I mean the code becomes extremely slow and you will waste a lot of time waiting for the code to give you results, right . Now, suppose we have the trajectory, then what? Then we could calculate statistical quantities as I have been telling you.

(Refer Slide Time: 01:42)



CALCULATIONS from MD.

① Pair Correlation Function $g(r)$. → Microscopic structure

② Maxwell Boltzmann Speed Distribution.

$$P(v) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-mv^2/2k_BT}$$

$$\underset{\text{Prob density.}}{P(v)\,dv} : \frac{dN(v, v+dv)}{N} = \text{prob of finding } dN \text{ particle with speed between } v \text{ and } v+dv.$$

③ Diffussion Constant.

$$D = \frac{\langle \Delta r^2 \rangle}{6(t_2 - t_1)} \rightarrow \text{average over many particles / independent runs.}$$

In today's class we shall discuss the discussion of how to calculate three statistical quantities just as an example. In that you will calculate the pair correlation function, which essentially gives you information about the microscopic structure of the particles, right. And basically the Fourier transform of the pair correlation function is the structure factor.

So, when you do a scattering experiment to find out the structure of a solid or maybe of a liquid; structure of a solid most probably has been taught in class and you get the so called (Refer Time: 02:23); but here we have been studying liquids, dense liquids or gases. And if you basically measure, as I am using a similar scattering experiment, the structure of the fluid, so structure factor, by measuring the structure factor and then you can take the Fourier transform of it and calculate the g of r.

On the other hand, you can also essentially put in a computer code, have a guess for the potential and calculate the g of r, given the match the densities as with that of your experiments and check whether they are matching or not. If they match, you would say that you have at least in terms of structure a reasonable or a good model of for your fluid at the right temperature and so on and so forth.

The other thing which we shall calculate is the Max well Boltzmann speed distribution, right. So, in a statistical mechanics course you must have already learnt about the

Maxwell Boltzmann speed distribution and what it gives is essentially the probability that a certain number of particles are moving with speed v and v plus d v, right.

So, here this is the standard expression for P, where P is the probability density actually alright. And, so basically probability density has this expression, if you have any confusion just look up your stat mech book, this is nothing but a pre factor into e to the power minus e; e being energy and here you have kinetic energy by to K B T, right half m v square by 2 is kinetic energy and as the energy of the particles by K B T, it is the Boltzmann factor essentially.

So, we shall be discussing how to calculate that quantity and P v d v right, P v is the probability density and P v d v, d v being the speed interval; is basically the number of particles moving with velocity v and v plus d v by N, the total number of particles and that is the quantity which gives the probability of finding d N particles with speed between v and d v.

So, P v is probability density and P v d v is the probability. And the last thing which we will calculate or I will tell you how to calculate is the so called diffusion constant and Maxwell Boltzmann speed distribution. These two both are dynamical quantities, means it is it basically is a measure of how the particles are moving with time, right.

So, you; so diffusion constant basically gives a measure of how much a particle has diffused or moved or got displaced. So, delta r square is a measure of the displacement of a particle, of course you have an expectation value here, you need that expectation value; because you are averaging over many particles right. You it is an n symbol average quantity, because if a particle is diffusing is doing a Brownian motion, you cannot predict how much that particle will move in the time interval t 2 minus t 1 in a certain time interval say, which I have denoted by t 2 minus t 1.

But if you average over many particles, then you can have an expectation value; what is the average expectation value? So, this is a measure of how much you can expect a particle to get displaced due to diffusion Brownian motion in time t 2 minus t 1. Here we are calculating P v d v Maxwell Boltzmann distribution, it is basically the number of particles moving with velocity v plus d v, right.

So, it is impossible for you to calculate P v d v, Maxwell Boltzmann speed distribution, if you did Monte Carlo simulations; a Monte Carlo off lattice simulations of this particles, which is also possible, right. From there you can calculate the pair correlation function from if you simulate n particles. If you simulate the equilibrium distribution or the positional distribution of particles in a Monte Carlo off lattice; Monte Carlo simulations with the same number of particles, you can get this quantity, pair correlation function you can get.

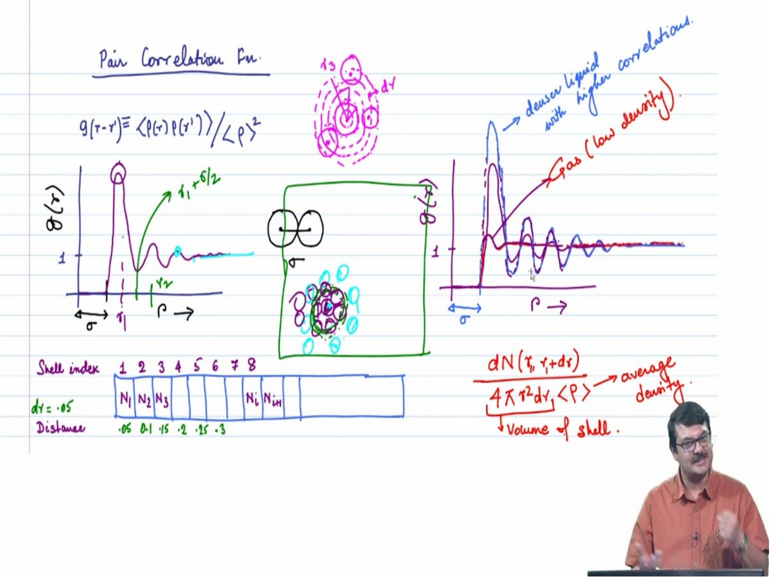
But you cannot possibly get Maxwell Boltzmann speed distribution, because you do not have speed in Monte Carlo simulations. Please refer to your first lecture where I gave a 15 minute 10 minute brief about how to do off lattice Monte Carlo simulation with particles I mean basically the same kind of things.

So, what molecular dynamics gives, as I said even in the first lecture you have these dynamical quantities. So, instead of calculating just structural quantities which you can also do in a Monte Carlo simulation, I have intentionally chosen two quantities which you can calculate when using only molecular dynamics. Having said that I will just add a small k v 8, people do calculate diffusion constant and dynamics because it is about matter of displacement even in Monte Carlo simulations; where t 2 minus t 1 is replaced by the number of Monte Carlo iterations.

So, that is also done, it is not unknown; but then you cannot really ascribe any time to the number of Monte Carlo iterations, but that is a topic for some other day and some other class.

So, if you want to do dynamics typically you choose molecular dynamics.

(Refer Slide Time: 08:56)



Having said that me give you details of what is a pair correlation function. The pair correlation function is typically denoted by g of r, as I have written here, r minus r dash, where r is the position of one particle and r dash is the position of another particle and it is also referred to the density-density correlation function.

So, the density at point r, right and the density at point r dash and you are taking an average expectation value over it and normalizing by the average density square. Now, so in statistical physics books you might more often find discussion about the density-density correlation function that is what you typically calculate when you can calculate density as a function of space.

So, I should not write equal to I should be writing equivalent to, where as in particle based simulations you typically calculate g of r which again tells you that if suppose there is a particle at point suppose r, then what is the probability, what is the, how are the other particles in it is neighborhood distributed. So, basically with respect to this one particle, you are trying to find out the probability of finding other particles at a certain distance r.

So, what you calculate in the pair correlation function is essentially how the particles are distributed about one particular particle. Of course you have to take an average; to calculate the pair correlation function you have to take an average over all particles. Because then it will say, I mean that, pair correlation function will give you the information that around any one particle and all particles you assume to be statistically similar; around any one particle what is the distribution of its neighbors around the particle.

Now just for an example just, suppose you are looking at this particular red particle here and basically what you do to calculate the pair correlation function is, draw concentric circles which are marked in green here around this red particle, right. And you figure out within each of these concentric circles, how many of it is neighbors are there.

Now, this concentrate circles are of width d r, right and they will be at distances r 1, r 2 and r 3 and r 4 and r 5 further and further away from the center of your particle, of the particle in choice. And then you basically calculate within each shell, that is at distance of r 1 and r 1 plus d r, thus the basically the width of a shell and r 2 and r 2 plus d r and r 3 and r 3 plus d r, how many particles are there, ok.

And this is what you have to calculate and suppose you calculate this around particle number 1, suppose this red particle here is particle number 1 and you find out that particle number 2, 3, 4 and maybe 500 and 600, as the particles move around in space is around particle number 1. And you are calculating what is the average density of particles, how many particles are there at distance r 1, r 2, r 3 on an average.

When you are doing on an average, then as you do, you draw concentric circles around particle 1; similarly you can draw concentric circles around say particle number 2, you draw concentric circles and figure out how many particles are there at a distance of r 1, at a distance of r 2; when I say r 2, I actually mean r 2 plus d r, when I say r 1, I mean r 1 plus d r.

So, how many particles are there within those concentric circles and then you repeat that thing for particle number 3 and 4 and 5 and 6 and for n particles; where n could be around 1000 or 5000 it depends upon the number of particles in your box. And then to get an average, you can average over how many particles were there around 1 2 3 4 5 and you can average over these, right; because for each particle you can calculate how many particles where there at a distance of r 1 r 2 r 3.

So, then you basically calculate an average over all these particles right, then you get a pair correlation function for a particular microstate. Now typically you will see that it will be extremely noisy and so it would be noisy and you have to average over many microstates; because what you are calculating is the statistical quantity right, how many particles on an average are at a distance r, that is what you are trying to calculate.

So, you not only average over a number of particles, but you run the simulations, assume that it has reached equilibrium, it would have reached equilibrium when we had a discussion about how to check for equilibrium; basically kinetic energy, potential energy etcetera would fluctuate about in average, they would not evolve especially if you have a thermostat, right.

And assuming it has reached equilibrium, you can calculate the pair correlation function and it will give you an average number of particles that you find in a at a distance r on an average from any particle, right. It is basically telling you the distribution of neighbors around your any particles. So, how would it look, how should it look?

So, here I have drawn schematically; of course, you will calculate it computationally, but here I have drawn it schematically, how it would look for a simple liquid. Simple liquid, Lennard Jones would you would call a simple liquid; where basically the potential is radially symmetric, because you could have also particles of a liquid which are rod like, but we are talking about simple liquid. So, we are not going there.

So, for a simple liquid like Lennard Jones fluid at a suitable density is suppose around 0.2 or 0.3, what you would find; so, is if I plot g of r pair correlation function on the y axis versus r, the distance between two particles on the x axis. Then basically for g of r would be 0 at very small distances till about the distance till r becomes order of sigma. And why is that, because basically two particles as in here, they cannot approach each other beyond the below a distance where the distance between their centers is approximately is equal to their diameter, right.

So, why cannot they approach, but because as you see as the center between the particles becomes less than sigma; we have chosen sigma equal to 1, just to remind you. As the distance between the centers becomes less than sigma, this is a sharp r to the power 12; 1 upon r to the power 12, strong repulsive force which will make the particles move away from each other that is a repulsive force.

So, hence you see that nearly till up till sigma, the g of r is nearly equal to 0, which means, two particles cannot approach each other below this distance 1, when sigma equal to 1. Well, you can, so basically you do find, so the particles can approach distances of say 0.98 or 0.99 sigma. It depends upon the value, how strongly the potential goes up and also depends the value of 4 epsilon; the prefactor of the Lennard Jones potential.

But basically from around 0.98, the g of r starts to increase, which means that with rare occasions you do find where the distance between two particles, where the distance between the centers of two particles becomes around 0.98 or 0.97 or 0.99, but it is a rarity. And then you find a peak just beyond r greater than sigma, here you find a peak, right.

And the reason you find a peak is that, if it is a liquid, then each particle will be surrounded by many other particles. So, if this is the; this is your central sphere in your particle of interest, just around it would be around 12 neighbors, if it is a dense liquid; less if it is a less dense liquid. But you would find a large number of neighbors just sitting around it, right.

So, what would be the schematic of a liquid? It would be something like this, surrounded by identical particles, right. So, somewhere here, some are slightly further apart, some are closer and this is the approximate schematic diagram of a liquid. So corresponding, correspondingly just like around this particle right, you find these other particles, you find a first shell of particles which is surrounding the particle of your interest and corresponding to that you see a peak here, right.

And if you have a large number of particles just surrounding your any particular particle; I mean this central particle would be actually any particle, because an average you are calculating how many particles would you find at a distance r, where you have to take the both the average over the number of particles and the time average, right. And if you have a particle; if you have a large number of particles at say distance r 1 right, which is basically a measure of your first neighbor shell, then just at a distance beyond r 1.

So, it is slightly larger; suppose r 1 plus sigma by 2, sigma by 2 is the radius, right. So, if you are supposed finding a large number of particles at such a distance around the central particle; then you are not going to find the center of particles at such a distance, right.

Because most of the particles are just simply very close and sticking, there is Lennard Jones attraction, they are sticking to it.

So, at a slightly higher distance around, if this was r 1 then around so, if this first shell the peak was at r 1 at a distance and approximately r 1 plus sigma by 2 distance, you are going to find a minima; because not you would not find the center of other particles at that distance.

But at the position of r 2 suppose, which is the second neighbor shell. So, which is basically the second range of neighbors, so you have a range of parts, this is your first range of neighbors and you would have suppose a more neighbors somewhere sitting here right; so, which is the second range of neighbors. For this particle this would be the first range of neighbors, but for this particle this is the second range of neighbor, you will find a large number of particles around distance r 2. Of course, they are going to move around they want to fluctuate their position in space and for that you will see a peak.

Similarly this speak if you have one, then it corresponds to a third range of neighbors. Of course, if the density of the liquid is lower, then this peak will also be lower, because you are going to find fewer number of particles at distance r 1 and a fewer number of particles at a distance r 2. And at large distances here, right at large distances what you will find is basically that it goes to one, there are no correlation; so on an average at very far distances or at all distances there is an equal probability.

So, the way pair correlation function is calculated, is basically you look at count on an average, the number of particles that you have between r 1 and r 1 plus d r that is within a shell. Then you divide it by the volume of the shell, because as you go farther and farther away though your d r remains the same, the volume of the shell is 4 pi r squared d r

So, if you are going to larger distances, when r is large 4 pi r square. So, basically the volume of the shell increases, right. And if you have a large volume of a shell, then you are likely to have a larger number of particles as well within that volume. So, to normalize it, to compare across distances what you calculate is; you normalize this d N, which is the number of particles between N between r 1 and r 1 plus d r and you also normalize by the volume of the shell, right. Because you are calculating how many particles are there, what is the density of particles at a distance r that is what pair correlation function is giving you, what is the average number of particles if you integrate.

So, if you integrate over the volume of the shell, you will get the total number of particles which is within the shell at a distance r 1 yeah and you normalize by the each volume of the shell. So, that on an average you can figure out oh, at certain distance r 1 I have a peak which means there is a high number of particles around it, at a slightly larger distance r 1 plus sigma by 2 we have a low number of particles around it. Low density of particles you are normalizing, because if a larger shell you might end up having larger number of particles; but density is lower, and finally, you normalize by the average density.

So, that at large distances r large, you can basically this g of r goes to approximately 1. So, then when if you have over large distances, you will lose any correlations unless you are very near the critical point. And for liquids this will basically go to 1, because it is normalized by the average density; it will any particle at far distances will see an average density, it would not have these correlations.

Now if this liquid was, if the density of the liquid is decreased; then what you will find is the peak will go down or rather if you increase the density of the liquid, then the this peak will increase and you will have a higher peak, the second peak and the third peak and so on so forth. It also depends upon the temperature, how will far the correlations can be seen, but always at large distances for a liquid, for a fluid that is a liquid or a gas, you will go the see that the correlations have gone to 0.

By the way whenever I say, distances, large distances please you have to keep in mind that the largest distance that you can calculate is L by 2, which is half the length of the box. Technically root 3 by 2 if you measure along the diagonal, but typically you do not plot r larger than L by 2; where l is the length of the box, the size of the boxers as so assuming a cubic box.

And, if the box is not cubic, if L 1, L 2, L 3 that three different sizes are different, then you would typically plot up to the smallest dimension of the box by 2, right. And why is that? Because that is because of periodic boundary conditions, so whenever you are calculating r; you have to also keep in mind that two particles could be at two ends of the box and yet be very close to each other, while calculating due to the minimum image convention.

So, whenever you calculate r, when you draw concentric shell and calculate how many particles are there around my shell; remember as the shell grows bigger and bigger, you have to account for particles which are on the other side of the box. right. Specially, if this particle is somewhere sitting, suppose this particle is sitting somewhere. So, this is your simulation box see right and if it is here, then definitely this particle could be some close to some particle here. So, you have to be careful of that.

So, going; so while you have to develop the algorithm to calculate pair correlation function; we shall just tell you that for the pair correlation function you shall need an array of length which is essentially decided by d r; d r is what is the width of the concentric shell.

And suppose the size of the box assume a cubic box for simplicity is supposed 30, then basically L by 2 is 15. And the length of this array by in which you shall store the values of g r while it is being calculated will be 15 or L by 2 divided by d r. right. Because that distance you are going to; half the length of the box, the maximum distance you are going to divide it into small little segments of width d r.

So, the length of this array in which you shall store the pair correlation function will be L by 2 divided by d r. And so, in this different boxes of the array, different indices of the array 1 2 3 4 so on so forth. So, what are you going to store; basically when the index for the array is 1, basically you store and how many particles are there between 0 and 0.5. For in the end; when index for the array the index is 2, you basically store how many particles you find between 0.05 and 0.1; when the array is 3, you basically store how many particles you store between you can find on between 0.1 and 0.15 and so on.

Now, if you have a liquid, you know that this the values of n 1 and n 2 and n 3 will be 0; till approximately the value of sigma. So, write only when you find. So, it is much further down where basically you are checking how many particles are there in distance between 1 and 1.05, so which is somewhere here; that is where you will find a finite number of particles.

Now you have to calculate this array for each particle, then you have to average basically you, there should be another array where you shall call average g of r or whatever right, and there you shall similarly add all these numbers for another particle, and so on so forth for every particle.

So, then you will have an average g of r, g of r is anyway the pair correlation function g of r is anyway an average quantity; nobody is typically interested to know what is going on around one particular particle, it is an average quantity. And so, you have to average over all the particles in this array keep on storing for each particle, the end you have to divide by the total number of particles to get an average g of r at one particular microstate.

And then as you allow the system to evolve, say you can and the system will access different microstates due to it is time evolution. And then what you do is at each iteration, every 50 iteration say; because one iteration to the other the positions will be rather correlated, it would not be an independent microstate because the particles would have moved only a little v bit, right. So, after say around 50 iterations or after 100 iterations where the, so that the configuration of these particles are slightly different; you reach another microstate, independent microstate and then again this value of g of r, you add this array to another calculated value of g of r, so that you can take a time average, right.

So, time average is basically not very different from what we are doing in a Monte Carlo. You are doing, you are basically calculating some quantity at a particular number of iterations and then that quantity after again a certain number of iterations and again in the value of that quantity, again after certain number of iterations and at the end you were adding all these things up and dividing by the number of snapshots, right.

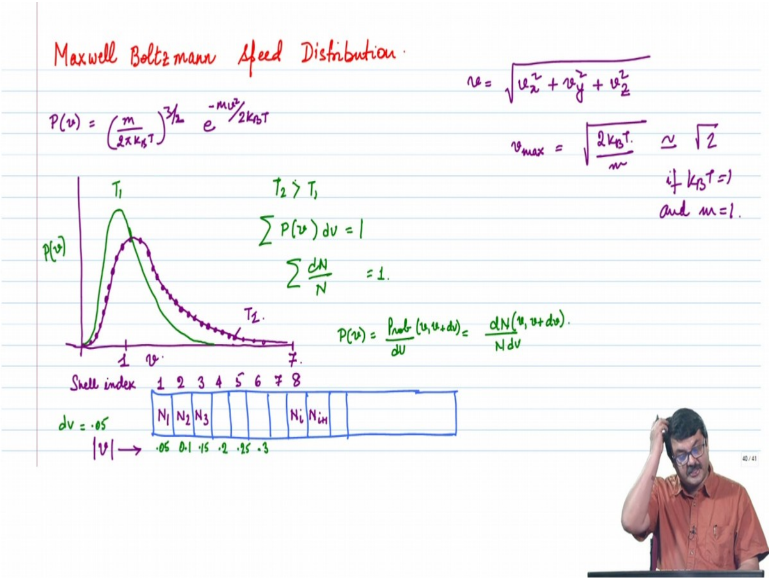
Here the same thing you are doing except that you are doing it for an arrays. So, it is a long array and when you have to plot, you plot basically this average values of N 1, N 2, N 3, N 4, N i, N i plus 1, particle average and time averaged versus distance. Because your first one, this one in to d r will say how many particles are there in; how many particles are there between 0 and 0.5.

Similarly if you multiply the eighth or ninth or i th index of the array by d r right, that will tell you how many particles are there in that box and so on so forth ok. So, that is how you calculate pair correlation function that will be given as an assignment for different densities. And you should see that if you decrease the density, then this peak, the position of this peak also keeps on decreasing.

And so that when it is a gas right, it is really dilute, then you basically the pair correlation one function beyond distance sigma becomes practically equal to 1. So, here I have shown, so very small peak and thereafter it becomes practically one here all over and if it is a real gas it would be 0 up till distance 1 and 1 onwards.

So, for r equal to 1 and after r equal to 1 the value of g of r will remain 1 from r greater than 1 to all large distances this is half the length of the box. What does that mean; that means, that the probability of finding a particle at any distance greater than r equal to sigma is the same; it is an average density, there are no local spatial correlations as seen in a liquid. In fact, microscopically that is how you differentiate between liquid and a gas; because for a gas the pair correlation function for r greater than sigma will be absolutely equal to 1, ok.

(Refer Slide Time: 36:11)



Maxwell Boltzmann Speed Distribution.

$$P(v) = \left(\frac{m}{2\pi k_BT}\right)^{3/2} e^{-mv^2/2k_BT}$$

$$v = \sqrt{v_x^2 + v_y^2 + v_z^2}$$

$$v_{max} = \sqrt{\frac{2k_BT}{m}} \simeq \sqrt{2}$$

if $k_BT = 1$ and $m = 1$.

$T_2 > T_1$

$$\sum P(v)\, dv = 1$$

$$\sum \frac{dN}{N} = 1.$$

$$P(v) = \frac{Prob\,(v, v+dv)}{dv} = \frac{dN(v, v+dv)}{N\,dv}.$$

$dv = .05$

So, what about the next quantity this, if you have understood how to calculate pair correlation function; then calculating the Maxwell Boltzmann speed distribution is should be a cakewalk. Because in the Maxwell Boltzmann speed distribution where the probability density function P v is, this is the form which you can check from your statistical mechanics book. And if you plot it, this is how it looks right; there has to be something missing here, and I had missed v square even here, if I am calculating the speed distribution, ok.

And because this quantity is an exponential with increasing v, it is sharply falling; but v square is an increasing function you get a peak, right. And it is not a symmetric function, speed cannot be less than 0, so it has to start from 0 to larger and larger values. And when I say that you have to calculate P v in the computer simulations, this again means that the particles are moving around in space, they are colliding. Hence, they are exchanging velocities with each other, there is momentum transfer from one particle to the other. So, hence the particle; each particle will keep on changing it is speed as it moves along.

Moreover, on an average different particles will be moving with different speeds and of course, velocities itself. So, on an average to calculate P v, you have to calculate how many particles d N, how many particles d N are moving with speeds between v and v plus d v, right. So, what is v; v is nothing but v x square plus v y square plus v z square root, and just as previously, just as in pair correlation function; you have to average over different times, right.

So, you allow the system to evolve as the particles move around every 50 iterations or every 100 iterations depends upon the value of d t. As soon as you have a different different microstate of the system; you calculate for a particular microstate, how many particles are moving with velocity v and v plus d v. And of course, this d N quantity you averaged over different microstates at different time.

So, finally, you keep on adding an array at the value of d N 1, 2, 3, 4, 5, 6 for different values of velocity and each value you calculate at different times, average them divided by the total number of snapshots; not very different from what we did in Monte Carlo or the previous step.

So, here you will again have an array where the first index of the array will store how many particles are moving with speed 0 and 0.05. So, suppose d v is 0.05 and then how many particles are moving with speed between 0.05 and 0.1, how many particles are moving with speed between 0.1 and 0.15 and so on so forth.

So, again what will be the size of the array? Because you have to initialize it, how can you estimate it. Now we were working with essentially K B T equal to 1 and mass equal to 1. And you know that the position of this peak of the Maxwell rules wind speed distribution is root 2 K B T by M, right, so this point, if K B T is 1 and M equal to 1, so the position of this peak will be at root 2, right.

Now, if the temperature changed, so if the temperature became less then what you would get of course, then the peak will also shift, so though you will get a different Maxwell Boltzmann distribution. So, here this data, this one is for essentially temperature T 1; the one with dots where this one is with temperature T 2, where T 2 is greater than T 1.

And of course, the probability under the curve which is P v d v right, as to be and the some of it has to be 1 because it is probability, right. This is probability density function; that is probability divided by d v, right What is probability? It is number of particles moving with velocity v 1 and v 1 plus v or v plus d v divided by the total number of particles on an, this N will be fixed, this d N is the quantity whose average you have to calculate; and then all that you have to do is basically divided by d v, which will give you the probability density.

Because, if you chose d v to be smaller, say instead of 0.05 you chose it to be 0.001; then you are going to find a lot fewer number of particles between 0 and 0 0 1, right. Because all the particles then suppose N 1 was the number of particles, an average number which was moving with velocity with speed between 0 and 0.05.

Now if you choose your d v to be 0.01 say, then this N 1 particle will be divided into 3, into five partitions right; because now you are calculating how many particles are moving between 0 and 0.01, 0.01 and 0.02, between 0.02 and 0.03. If we add up all these numbers, the average from point from 0 to 0.05, then you will get. So, you have to divide by d v, be aware of that, right.