

Computational Chemistry & Classical Molecular Dynamics
Prof. B. L. Tembe
Department of Chemistry
Indian Institute of Technology – Bombay

Lecture – 31

Classical Molecular Dynamics – 4: Algorithms for Constraints, Thermodynamic Properties and the Pair Distribution Functions

Hello and welcome to this continuation lecture. In the last lecture we considered the details of the MD simulation in particular the Verlet algorithm and other details, details of the forces and how to calculate the new step in terms of the old step and one of the points that we considered was if there is a molecule like water or CH₃ CL or CCL₄. As these molecules move the bond lengths and bond angle should not change.

So for this purpose a method called SHAKE is used. SHAKE is a constraint method that is what I have given on this slide okay.

(Refer Slide Time: 00:54)

Constraints and Shake

If certain motions are considered unimportant or "unphysical", constrained MD can be more efficient
e.g. SHAKE algorithm - bond length constraints
Rigid bodies can be used e.g. Eulers methods and quaternion algorithms

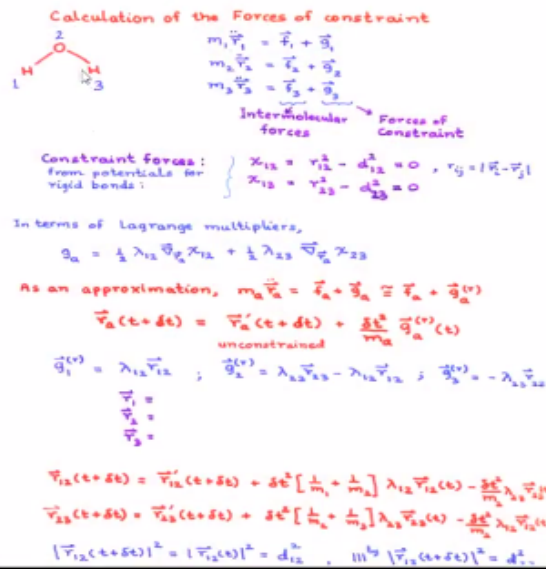
Statistical Mechanics

The prime purpose of MD is to sample the *phase space* of the statistical mechanical ensemble.
Most physical properties are obtained as specific averages
Structural properties are obtained from spatial correlation functions e.g. radial distribution functions.
Time dependent properties (transport coefficients) are obtained via time correlation functions e.g. velocity autocorrelation function.

So for this particular MD simulation the OH bond angle changes or the OH length changes are not important, so we are only interested in the molecule moving as a whole not in much shorter range motions. So in that case we want to keep the bond lengths and bond angles constant as if the molecule is moving as a rigid body okay. So that is what we will discuss now and we also commented on the methods of statistical mechanics.

This you will see in another course on statistical mechanics. So once you have lot of constraints that is indicated here, have this molecule water HOH.

(Refer Slide Time: 01:35)



As the molecule moves around I do not want to change the OH bond length nor do I want to change the HOH bond angle. So instead of calling a bond angle as a separate constraint I want to say that one, the distance between oxygen and hydrogen is fixed, between second oxygen and hydrogen is fixed and also the HH length is fixed. So if the 3 lengths are fixed automatically the bond length is fixed.

So for this purpose we use a constraint method so the rate of change of the first coordinate is given by the forces on the first coordinate plus the constraint on the first atom. Similarly, there is a constraint on the second atom, g_2 , constraint on the third atom g_3 , now what are these constraints, these constraints are given here. So the distance between 1 and 2 is $r_{12}^2 - d_{12}^2$ square, d_{12}^2 square is the bond length okay.

So this bond length has to be fixed, distance between 1 and 3 is the second bond length that should be fixed and similarly the bond length between 3 and 1 has to be fixed. So these are the constraint equations. So how do I get the constraint. So we use a Lagrange equation of motion and these are the Lagrange multipliers. So the constraint is given in terms of some constant and the gradient of the potential.

So this gradient of this particular constraint gives me a potential, so there is a constraint to give me an overall force. So there is a force on each atom due to all the constraints on that atom. So in particular, so this is g_a , constraint on g_a is gradient with respect to r_a of X_{12} and

also gradient with respect to r_{12} and r_{23} , both constraints have to be used. So this is a rather complicated problem.

So what is done the whole thing is linearized okay, so this is a method of linearization. So you need to basically find λ_{12} and λ_{23} . So you rewrite the equations for constraint and the new positions now, the new r_{12} and the new r_{23} okay that should be the unconstrained move plus the constraints. So the unconstrained move would be due to the forces themselves directly.

And the constraints will correct for them. So constraint will correct for them so this λ_{12} and λ_{23} are obtained iteratively so that the bond length is fixed. So the new bond length r_{12} after the move will be identical to the old bond length, okay. So the theory you can study later. So we just want to mention that you use SHAKE constraints or some other types of constraints so that the bond length and bond angles are fixed okay.

(Refer Slide Time: 04:27)

• Thermodynamic Properties

- Kinetic Energy:

$$\langle K.E. \rangle = \left\langle \frac{1}{2} \sum_i^N m_i v_i^2 \right\rangle = \left(\frac{3}{2} \right) k_B T$$

- Temperature calculated from K. E:

$$T = \frac{2}{3 N k_B} \langle K.E. \rangle$$

Configurational or internal Energy:

$$U_c = \left\langle \sum_i \sum_{j>i}^N V(r_{ij}) \right\rangle$$



So next we want to study thermodynamic properties, so what are the thermodynamic properties you want to study, one would be the kinetic energy, so kinetic energy is nothing but the sum of i going from 1 to N , $m_i v_i^2$ so that should be $= 3/2 kT$ by the equipartition theorem. So it allows me to calculate the temperature from calculation of the kinetic energy. So temperature is really obtained by kinetic energy $* 2/3 NKB$.

So this is how I calculate the temperature for every step in my simulation. Similarly, how do I calculate the internal energy, so remember all my particles are interacting with one another.

So if I calculate the total interaction energy between all the particles this is i and j summed over all the particles, calculate the average interaction energy over the entire trajectory then it gives me what is called the configuration energy or it is nothing but the internal energy, okay.

So internal energy is an average of all the interactive forces. For an ideal gas this internal energy is 0, okay, so this is really your thermodynamic energy.

(Refer Slide Time: 05:35)

Pressure:

$$PV = Nk_B T - \frac{1}{3} \left\langle \sum_{i=1}^{N-1} \sum_{j>i}^N r_{ij} \cdot f_{ij} \right\rangle$$

Heat Capacity

$$\langle \delta(U_c)^2 \rangle = \frac{3}{2} Nk_B^2 T^2 \left(1 - \frac{3Nk_B}{2C_v} \right)$$



Then I can also calculate pressure through my simulation. So this is an equation, this is a virial equation for pressure. If there was no interaction f_{ij} is the force between molecule i and molecule j , r_{ij} is the distance between i and j . So this is called a virial. So this is a standard statistical mechanical formula. So if there were no forces again $PV = NKT$, if there would be no forces.

But when the forces act it corrects for the ideal gas equation of state through this virial corrections. So this is an equation for pressure and since in each step we are calculating r_{ij} between particles, f_{ij} between particles, so this dot product can be calculated at every step. You get the average of this quantity so at the end of the simulation I get pressure in terms of this quantity.

So evaluating pressure, thermodynamic pressure is very easy in an MD simulation. So similarly now suppose I want to get the heat capacity. So heat capacity is given in terms of this formula. So δU_c square is the average of c square – average of U_c whole square. So this is a fluctuation in energy. Fluctuation in energy is obtained by average of the square of

energy – square of the average of energy. So that fluctuation is given by $3/2 Nk$ square T square * $1-3 NKB/C_v$, so C_v is my heat capacity at constant volume.

So this is how you calculate heat capacity in my calculations. So in addition to my trajectories which are positions and momenta as a function of time, there is now a formula to calculate pressure as well as heat capacity.

(Refer Slide Time: 07:23)



Pair correlation (Radial Distribution Function):

$$g(r) = \frac{\langle n(r) \rangle}{4\pi r^2 \Delta r \rho} = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i}^N \delta(r - r_{ij}) \right\rangle$$

Static Structure factor:

$$S(k) = 1 + 4\pi\rho \int_0^\infty \frac{\sin(kr)}{kr} (g(r) - 1) r^2 dr$$

S(k) available from x-ray diffraction (experimental data)

So then now we come to the probably the most important term in this entire simulation which is called the radial distribution function. So we will spend some time in understanding what this radial distribution function is. So this radial distribution function at a given distance are is the local density at the distance; local density, how do you get the local density, find the density at that distance and divided by 4π square dr , which is the volume at that distance r .

So the next slide will explain this in greater detail. So I get the local density given by this term, divided by the bulk density that is my pair distribution function. So the same thing can be obtained in terms of counting all the ij pairs, r_{ij} is the distance and this $\delta(r - r_{ij}) = 1$, if $r = r_{ij}$ and it is $= 0$ if r is not equal to r_{ij} . So you count all the pair distances through this delta function, Kronecker delta function.

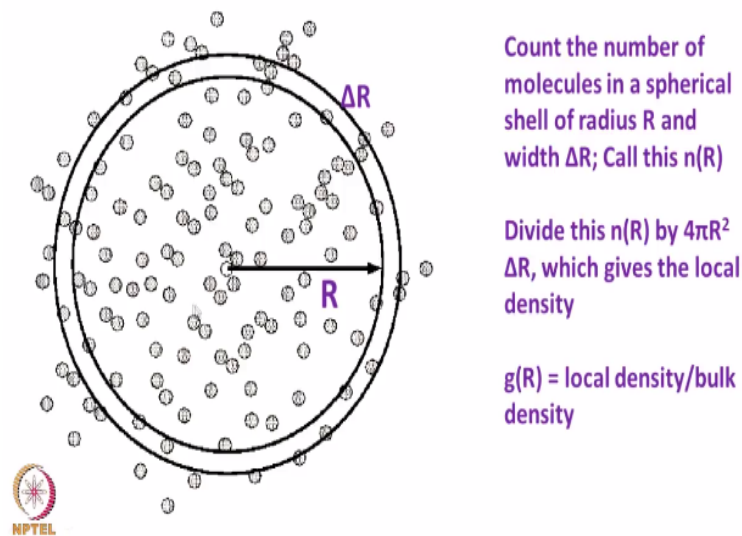
Then multiply by v divide by n square. So when you execute this formula you will also get the pair distribution function. So now this pair distribution function really tells you probability of finding the second particle at a distance r if one particle is located the origin.

So this is pair because there is one particle at the origin and one particle at r . So what is the probability of finding these 2 molecules at this distance r .

So that is what the pair correlation function or the radius distribution function tells us. So now what is the use of this distribution function. So when you take a sin transform of that pair correlation function what you get is a structure factor. This S of k is the structure factor and structure factor can be directly obtained from experimental x-ray diffraction data. So this pair correlation function is very important because you can directly relate the transform of this correlation function.

This is like a Fourier transform to the structure factor, which can be measured experimentally. So this is the connection between your microscopic pair correlation function with a structure factor. So let us now analyze how to calculate this pair correlation function again.

(Refer Slide Time: 09:40)



So this slide really explains what is the meaning of this pair correlation function. Imagine this white circle which is located at the center and all the other molecules are distributed around the central molecule. Now what I want to do is to calculate a local density at this distance R then divide by the global density. So what is the local density at this distance R . So at this distance R , I construct a shell, this is a spherical shell.

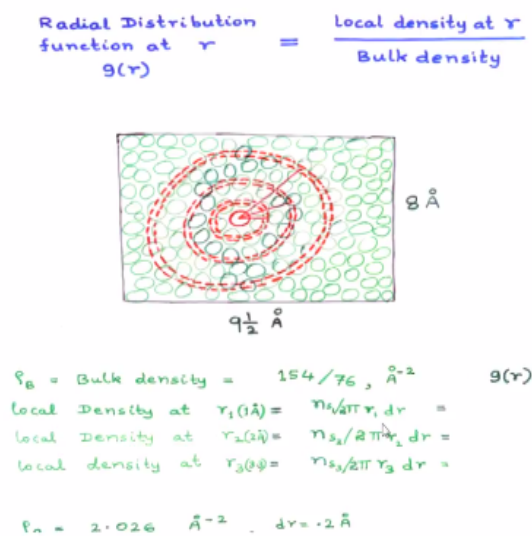
So in 2 dimensional it looks like 2 circles, but this is a spherical shell and what is the volume of that spherical shell, volume would be area multiplied by this delta R . So what is the area of this sphere $4\pi R^2$. So $4\pi R^2$ is the area of this sphere, to that I multiply by this

delta R, so that gives me the volume of this shell. So that is the volume of the shell, $4 \pi R^2 \Delta R$ is the volume of that shell.

Now what I want to do, I want to count how many particles are there in this particular shell. How will I count? Let us say I have the central particle, I have another particle. If this distance between these 2 particles is between R and R+ delta R then this will be in that shell. So all the molecules whose distances are between R and R+ delta R will be in this particular shell. So once I count that number let me call that number as nR.

So what is nR? nR is the number of molecules in this spherical shell. What is the shell? The space between an inner area and an outer area. So since this NR is the number of particles in that shell, if I divide that by that volume, so I get a local density. So that will be the density in that shell. So that density divided by bulk density is called the pair distribution function. So this pair distribution function will allow you to calculate many quantities.

(Refer Slide Time: 11:46)



So in the next slide I will illustrate the same calculation once again, so this is a 2 dimensional area. So in that area what I have done, I have taken a central particle here. The red particle is the central particle. So at several distances from the central particle I construct these shells. Now in 2 dimensions the shell is nothing but that area. So what will be the area. Area will be $\pi R^2 dr$ will be that area now.

So for the inner shell it will be $\pi R^2 dr$, for the outer shell it will be πR^2 , this R is bigger than the inner R, so in a 2 dimension, so instead of a volume I have an area. So how

do I calculate the local density? Number of particles in that area divided by the volume in 2 dimension which is that area now okay. So density at R divided by bulk density. So it would be very easy to calculate now.

So let us first calculate the bulk density. So what is the bulk density in 2 dimensions? Number of particles in this entire box divided by the area. So area is $8 \text{ angstroms} * 9 \text{ and } 1/2 \text{ angstroms}$ so that is 154 that is 76. So 76 angstrom square is the area of this box and now what I want to do, how do I get the bulk density? Total number of particles, so there are 154 particles in that area divided by that area so this is my bulk density.

Bulk density is the total density, the total number divided by total area. Now how do I get a local density at r, so the first one is at 1 angstrom, the first one is at 1 angstrom, second one is the 2 angstrom, third one is at 3.3 angstrom okay. So I have calculated 3 particular areas now, 3 particular distances. So in the first area that is the inner red circle, you will see that in the inner red circle there is hardly any particle in the inner red circle because some particle has to be inside that area.

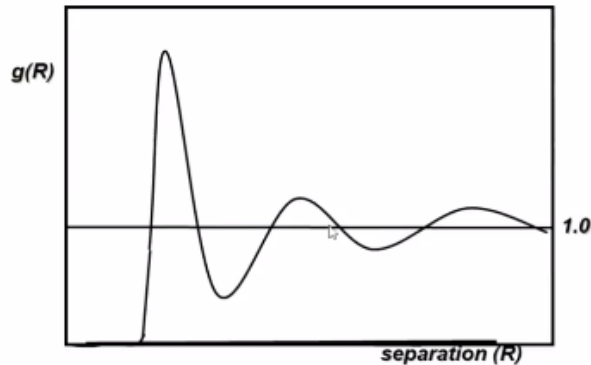
Only if the center is inside the area I will count it inside. So this is my second area, third area, so what you have to do, count the number 1 in the first area, number 2 in the second area, number 3 in the third area and divide. So that number divided by the area will give you the local density. So you can calculate actually for this example here. Calculate the local density and divide by your bulk density.

Bulk density is $154/76$, so that is your pair distribution function. So your bulk density is already calculated here. You will see that it is 2.026 angstrom to the -2. So this is the first one. So calculate all the local densities divided by the bulk density you will get a gfr. So now one point I want to make is that if you are very close to this center, you will see that no other particle will be located at this center.

So therefore you expect the local density to be 0 as r becomes smaller and smaller because when r becomes less than the size of the particle you cannot expect any other particle to be inside that first particle. So therefore this local density will be 0 for a number of points near the origin, then it will have oscillations.

(Refer Slide Time: 15:16)

A Radial Distribution Function



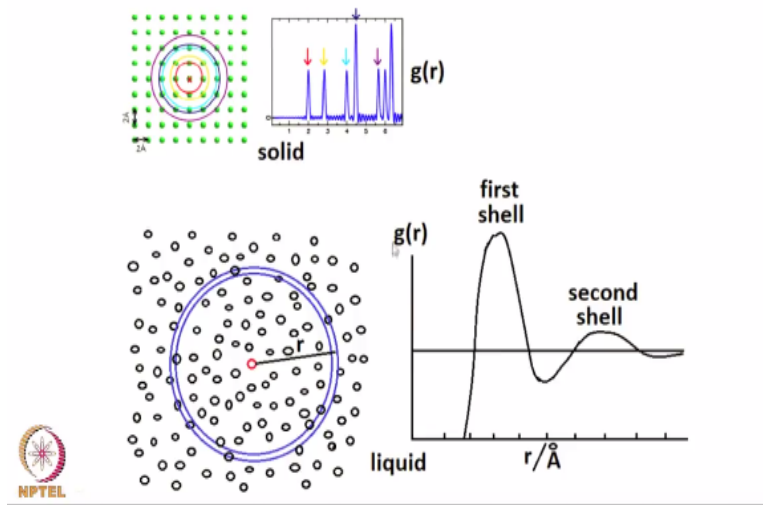
So one example will be shown here, so this is how that radial distribution function will look like. When you start near the origin, the local density is 0 because there is no other particle immediately near the center or $= 0$. So now when you go to some distance that local density will become a maximum. Why does it become a maximum? Go back, you will see that in some particular circle there will be many neighbours in that circle.

So when there is a first coordination shell there will be a maximum in gfr, second coordination shell there will be another maximum. So that is what this graph says. So around the first particle whenever the first shell is completely formed I have a maximum. So second shell I have another maximum, third shell I have another maximum. So between the first shell and the second shell as I am trying to move out there comes a point where this density is less than the bulk density okay.

There will be a region between the first peak and the second peak where the local density will be very small because immediately after the first neighbours there will be some gap till you find the second neighbour around a particle. So that is why this minimum is there. So this is my gfr.

(Refer Slide Time: 16:29)

Radial Distribution Function $g(r) = \exp -W(r)/kT$

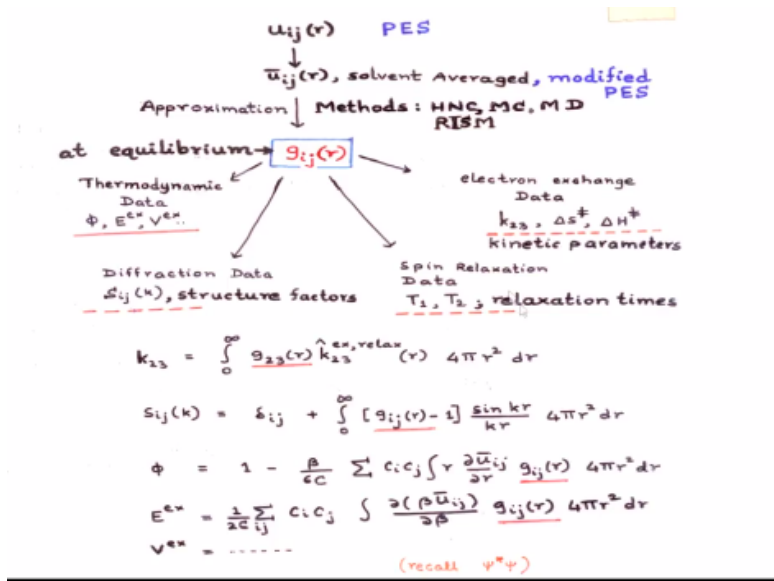


So the same thing now is explained here. This is a gfr for several systems okay. Now if it is a solid you will see that around a solid there are particles at some points and no particle at some points. So gfr for a solid will be oscillatory like this, whereas for a liquid gfr is continuous because these molecules keep on moving all the time. Molecules keep on moving all the time so there will be no place where there is no molecule.

So therefore for a liquid the gfr is a continuous function, whereas for a solid it is a discrete function like this. So now the other physical meaning of the gfr is that gfr can be written as exponential of $-Wr/kT$, so this Wr is really the work done to bring a particle up to distance r okay. So this is called a potential of mean force so you will have more of this in your later GROMACS calculations as well.

So what is the physical meaning of gfr? So if I take exponential of $-Wr/kT$ which is gfr. See this Wr tells me the work done to bring the particle from infinity to some given distance. So that is what this shows, it is a work term, work done to bring a particle from infinity to the distance r , so that is my meaning of Wr , which is related to gfr. So remember this in the future calculations.

(Refer Slide Time: 17:59)



So this slide is a representative slide which tells you how this gfr is related to many properties. Lots of theory is there behind this. You do not have to bother too much. All I am saying is that if you know gfr, you know the structure factor, I have already mentioned gfr can be given as a Fourier transform of the structure factor. So if you know gfr you can calculate the osmotic coefficient.

You can calculate the excess energy. You can calculate the excess volume and in certain cases using gfr you can also calculate rate constants. So this gfr is a very crucial part in MD simulations what you have done, you have started with a potential energy function. You have done the averaging and obtain this gfr then all these properties can be calculated.

(Refer Slide Time: 18:45)

Time Correlation functions C(t)

$$C(t) = \frac{1}{T_\infty} \int_0^{T_\infty} (f(t+\tau) - f_{av})(f(\tau) - f_{av}) d\tau$$

or

$$C(t) = \langle f(t)f(0) \rangle - \langle f_{av} \rangle^2$$

Mean squared displacement

$$2Dt = \frac{1}{3} \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle$$

Diffusion constants (D)

Velocity Autocorrelation

$$D = \frac{1}{3} \int_0^\infty \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle dt$$

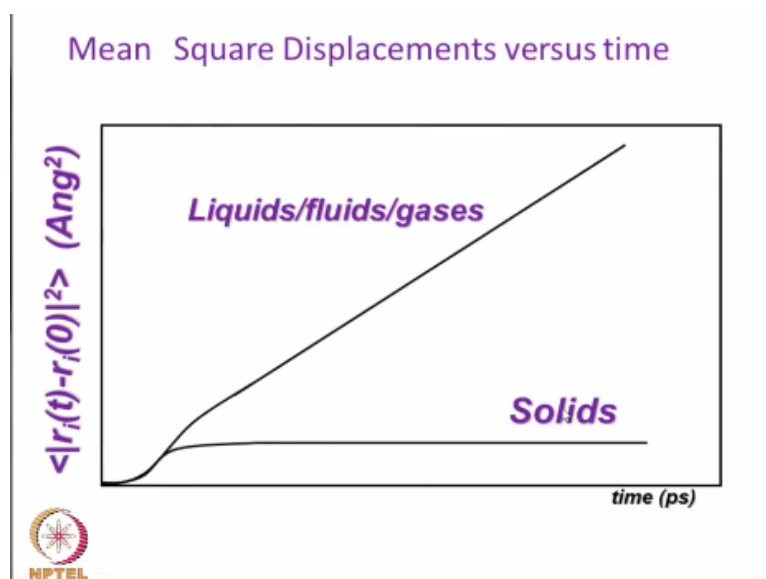
So next we want to go to a second physical quantity which is called a time correlation function. So what is the physical meaning of time correlation function. Suppose I have a function which depends on time. So this f now is the function which depends on time. So f at a given time t will be different from f at a later time $t + \tau$ okay. So this is f of τ , value of the function at time $t = \tau$.

This value is the function at a value of $t + \tau$. So between this value and the second value there is always a gap of τ . So I multiply the product and then divide by integrate from 0 to infinity for all values of τ . So what I get is called a time correlation function. So time correlation function basically tells me what are the fluctuations in time. So $g(r)$ told me what are the fluctuations in space.

Time correlation function tells me what are the fluctuations in time as time becomes very large you expect that the correlation become smaller and smaller, okay. Now what is the advantages of these time correlation function? One value is given here, so this is called a velocity autocorrelation function. If you know the velocity autocorrelation function you can integrate the time correlation function for velocity with respect to time and get the diffusion constant.

And the same diffusion constant can also be obtained from what are called mean square displacements. So I will have a slide now to explain the meaning of the time correlation function okay.

(Refer Slide Time: 20:26)



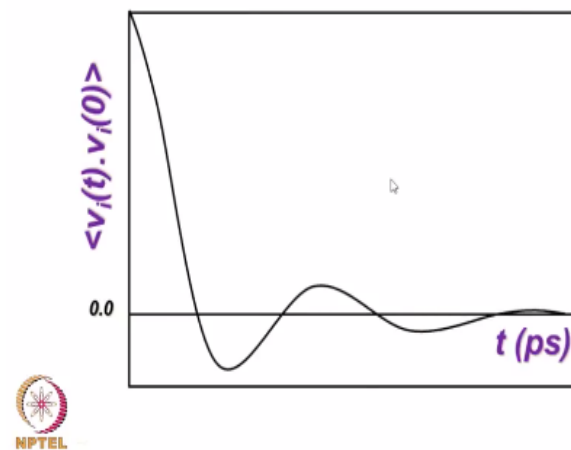
But before I do that so what I will show you, I will show how these mean square displacement change with time, this is given here, so mean square displacement is the value of position at a later time minus the value at an earlier time square that and average. So that is the mean square displacement. The meaning of mean square displacement is position at time 0 okay. This is the initial position.

Take the position at time t , take the difference between the two and square it and average over all particles. So that is the mean square displacement. You will see that in a solid the position of the particle is not going to change at all okay. So since the position does not change at all. The mean square displacement almost is constant, because there is hardly some fluctuations will be there around the mean position.

So at the most there will be some fluctuations which will remain constant whereas for liquids the particles keep on moving further and further away from the initial position. So therefore in liquid the mean square displacement varies linearly with time okay. So this is an important property of all liquids where there is a lot of uniform and random motion. So whenever you have a random motion the mean square displacement always is linear with respect to time.

(Refer Slide Time: 21:47)

Velocity Autocorrelation function versus time



So now since we talked of velocity autocorrelation function let me also show you a picture how it looks like. The velocity autocorrelation function looks like this. I will take the velocity at time 0 multiply by velocity at time t , average over all particles, then that product which is average is plotted as a function of time. So you will see that when this time $t = 0$, so it will be v_0 square.

So v_0 square average will have some value and as time goes on increasing the new velocity will have less and less connection with the initial velocity. Imaging for example a particle starts moving in the x axis, after some time it will not move only along the x axis, but it will change direction and move to y, move to z, finally the motion will be completely random. So at very large time the velocity at large times will have no connection or correlation with velocity at 0 time.

So that is why these correlation function go to 0. So 0 means no correlation, large value means very large correlation, negative value means negative correlation that is initially the particle starts moving say in one direction then the molecules in that direction will oppose that movement. Once the molecules in that direction oppose the movement the particle will come back.

So therefore velocity at time t will be negative compared to velocity at time 0 that is why the velocity autocorrelation function takes a negative value.

(Refer Slide Time: 23:17)

t	1			2			3		
	v_x	v_y	v_z	v_x	v_y	v_z	v_x	v_y	v_z
1 Δt	-2.7	-0.5	-2.0	-4.7	-8.2	-1.8	1.4	-7.0	-9.6
2 Δt	-2.8	-6.8	-2.1	4.1	-6.8	-1.9	1.43	-6.3	-1.0
3 Δt	-2.8	-8.7	-1.99	3.4	-5.1	-1.89	1.45	-5.2	-1.0
4 Δt	-2.79	-10.6	-1.85	2.6	-3.3	-1.86	1.44	-3.3	-9.0
5 Δt	-2.73	-1.2	-1.6	1.8	-1.5	-1.81	1.43	-0.9	-6.8
6 Δt	-2.6	-1.4	-1.31	0.8	-0.1	-1.74	1.41	-1.8	-3.1
7 Δt	-2.3	-1.6	-0.92	-0	-1.2	-1.65	1.37	4.8	-1.6
8 Δt	-1.8	-1.75	-0.49	-0.8	-1.9	-1.53	1.34	-7.5	-6.8
9 Δt	-1.0	-1.83	-0.11	-1.5	-1.9	-1.37	1.29	9.7	1.16
10 Δt	0.05	-1.95	0.16	-2.1	-1.2	-1.16	1.25	1.13	1.52
11 Δt	1.27	-2.0	0.28	-2.4	0.1	-0.89	1.22	1.22	1.71
12 Δt	2.43	-2.1	0.29	-2.5	+2.0	-0.55	1.18	1.22	1.71

Velocity autocorrelation function	τ	$c(\tau)$
$c(\tau) =$	0	
$\frac{1}{6} \cdot \frac{1}{3} \sum_{\mu=1}^6 \sum_{i=1}^{N_p} \frac{\vec{V}_i(\mu) \cdot \vec{V}_i(\mu+\tau)}{ \vec{V}_i(\mu) ^2}$	Δt	
	$2\Delta t$	
	$3\Delta t$	
	$4\Delta t$	
	$5\Delta t$	

$N_p = \text{No. of particles} = 3$
 $\tau = n \Delta t, n = 0, 1, 2, \dots$

So this slide really tells you what is the meaning of that autocorrelation function. So all these formula have a particular meaning, but let us understand what is the meaning of this autocorrelation function. What I have plotted I took a molecular dynamics trajectory so 3 particle trajectory, and I have written here V_x, V_y, V_z of particle 1, V_x, V_y, V_z of particle 2, V_x, V_y, V_z of particle 3 and I am listing them as a function of time.

So what is the meaning of correlation function at delta t, if delta t is, so this is 1 delta t, this is 2 delta t. So if I want to know how a function at one time is correlated to function at the second time. So how do I find that correlation. Take the first value, multiply by the second value. So this is correlation at delta t = 1 delta t, take at 2 delta t, multiply by 3 delta t, take the product, so take first this product then second product, third product.

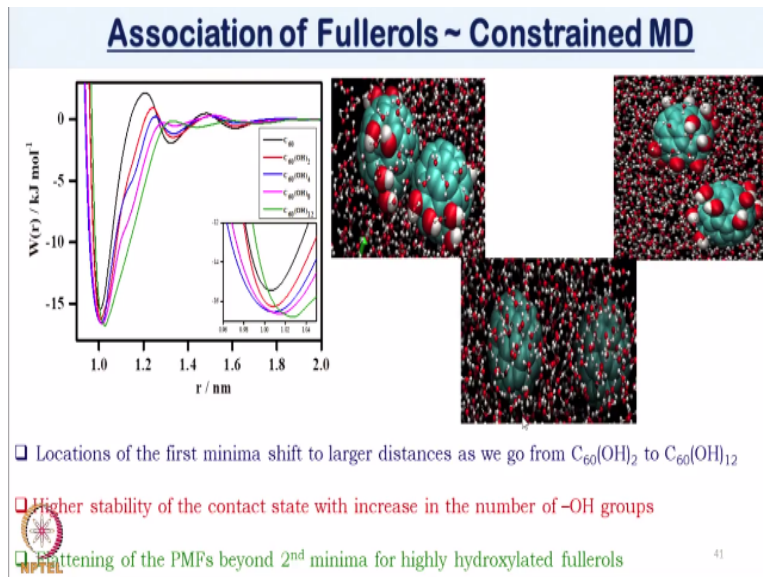
Each one of these pair products and divide by the number of products you have taken. So that is my pair correlation function at delta t. Now if I want pair correlation function at 2 delta t, take the value at here and take the third value, $1*3$, $2*4$, $3*5$, $4*6$. So whenever you take pair separated by 2 steps that is my autocorrelation function at 2 delta t. Now if I want correlation function at 3 delta t take the first value and multiply by the fourth value.

So this is my first point when I take the second value take second and fifth, then third and sixth, fourth and seventh. So take all products which are separated by 3 delta t. So that will be correlation function at 3 delta t. So if you want at 5 delta t, then first value and sixth value, second value and seventh value. So take all the pairs of products which are separated by the same difference in time.

So that is what this formula means, take some tau, now tau is the separation. So you take the products all separated by tau and sum over all values of mu and as well as all values of the particle. So I sum over all these particles, so that is how I will get the autocorrelation function, okay. So then you list all those points. So find tau is 0 at 0 it is just square of the original value.

What is the value at 0? All the value squared okay, that is my initial value then at each delta t I take products as I describe to get the autocorrelation function and it is picture you already know how it looks like okay.

(Refer Slide Time: 26:03)



So this is some example of a molecular dynamics we have taken fullerols here and calculate the pair distribution of fullerols and I told you earlier that pair distribution function can be written as $e^{-W(r)}$, so these really give you the pair correlation function between several fullerol molecules. This is just to given an illustration.

(Refer Slide Time: 26:27)

Computer Simulation of Liquids, M.P. Allen and D.J. Tildesley, Oxford (1989).

Theory of Simple Liquids, J.-P. Hansen and I.R. McDonald, Academic Press (1986).

Classical Mechanics, H. Goldstein, Addison Wesley (1980)

Gromacs manuals

NPTEL Courses on *Statistical Mechanics and Computational chemistry* (~2015)



Now some references for this, references are *Computer Simulation of Liquids* by Allen and Tildesley, it is a very nice book, if you want more theory you can do Hansen and McDonald, there is a *Theory of Simple Liquids* since we are discussing lot of equations of motion *Classical Mechanics* by Goldstein is a fairly good book and we have NPTEL Courses on *Statistical Mechanics* as well as *Computational chemistry* okay.

So you may also refer to that NPTEL Courses on Computational chemistry which will have lots of details of all the things we have said here okay.

(Refer Slide Time: 27:01)

Monte Carlo

- Monte Carlo (MC) simulations usually mean any simulation which utilizes random numbers in the simulation algorithm.
- The term "Monte Carlo" derives its name from the famous casinos in Monte Carlo; They could have called it by some other town as Las Vegas, etc. Metropolis Monte Carlo is most often used.
- Another closely related term for the same purpose is stochastic simulations



So what I want to do next time is to consider Monte Carlo simulations and an application of this MD program. So I will conclude here and next class I will exactly start with this Monte Carlo slide tell you how Monte Carlo simulations differ from molecular dynamics and illustrate it with an execution of a program. I will conclude here. Thank you.