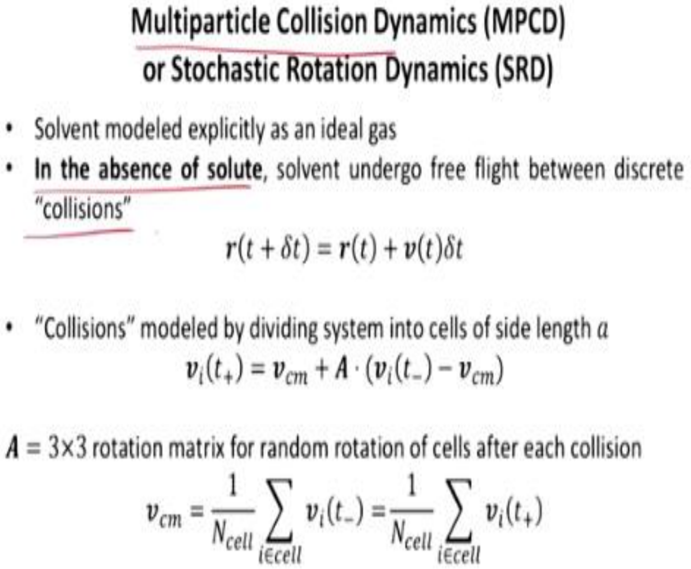
Advanced Thermodynamics and Molecular Simulations
Prof. Prateek Kumar Jha
Department of Chemical Engineering
Indian Institute of Technology, Roorkee

Lecture – 57
Multiparticle Collision Dynamics; Lattice Boltzmann Method; Coarse-Graining

Hello all of you. In the last lecture we discussed some methods of mesoscale simulations, the Langevin method or Langevin equation the Brownian dynamics and the dissipative particle dynamics method.

Now, I will discuss some more methods along those lines and then we come to the idea of coarse graining. So, one of the methods for simulating hydrodynamic behavior or the effect of solvent is the multiparticle collision dynamics that is called MPCD. It is also known as the Stochastic Rotation Dynamics called SRD.

(Refer Slide Time: 00:49)



Multiparticle Collision Dynamics (MPCD) or Stochastic Rotation Dynamics (SRD)

- Solvent modeled explicitly as an ideal gas
- **In the absence of solute**, solvent undergo free flight between discrete "collisions"

$$\boldsymbol{r}(t+\delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t)\delta t$$

- "Collisions" modeled by dividing system into cells of side length a

$$\boldsymbol{v}_i(t_+) = \boldsymbol{v}_{cm} + \boldsymbol{A}\cdot(\boldsymbol{v}_i(t_-) - \boldsymbol{v}_{cm})$$

$\boldsymbol{A}$ = 3×3 rotation matrix for random rotation of cells after each collision

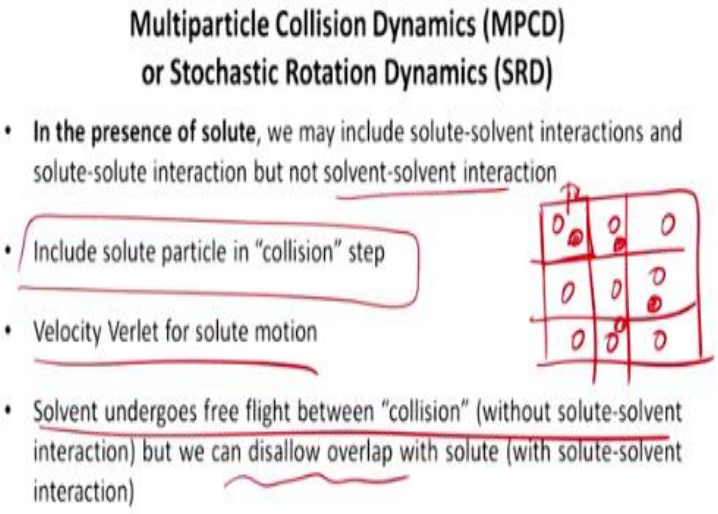
$$\boldsymbol{v}_{cm} = \frac{1}{N_{cell}}\sum_{i\in cell}\boldsymbol{v}_i(t_-) = \frac{1}{N_{cell}}\sum_{i\in cell}\boldsymbol{v}_i(t_+)$$

So, in this case, unlike the previous methods, we do explicitly put the solvent in the simulations. However, they are modeled as an ideal gas; that means they do not interact with each other; they pretty much move like an ideal gas; they do not occupy any size the way it helps is like automatically takes care of the thermal fluctuations and all that, but it does not clearly have the solvent-solvent interactions in the picture.

$$r(t+\delta t) = r(t) + v(t)\delta t$$

So if there is no solute then in that case the solvent pretty much undergo a free flight that is it moves at a constant velocity between what is defined as collisions. The collisions in here is not really the collision of the solvent molecules with themselves but the collision is defined in somewhat I would say interesting way and that is the following so, we divide the system into cells and that is just the hypothetical construction there is in reality there is no cell, the particles can be much grow go from anywhere to anywhere but that at any given time we will have particles present within these cells.

(Refer Slide Time: 01:59)



Multiparticle Collision Dynamics (MPCD)
or Stochastic Rotation Dynamics (SRD)

- **In the presence of solute**, we may include solute-solvent interactions and solute-solute interaction but not solvent-solvent interaction
- Include solute particle in "collision" step
- Velocity Verlet for solute motion
- Solvent undergoes free flight between "collision" (without solute-solvent interaction) but we can disallow overlap with solute (with solute-solvent interaction)

These particles in cells will keep on changing with time because as particles can keep moving between the cells but in a collision step one of the cells is chosen and it is given a random rotation to the entire cell is rotated and therefore the configuration of all the solvent molecules in that cell we are not doing any solute in this case right now is going to change, because the whole cell is rotated.

So, it is like a change applied over all the solvent molecules within a particular cell what this mathematically means is that if t minus is my old time and the velocity of the particular molecule is like v_i at the particular time then the velocity is first subtracted with the center of mass velocity to make sure at the center of mass does not really move. So, we have to ensure that the center of mass is same as earlier.

$$v_i(t_+) = v_{cm} + A\,.\,(v_i(t_-) - v_{cm})$$

So, we simply look at the difference from the center of mass velocity and multiply with what is known as rotation matrix that is going to be a three times three matrix for three dimensional rotation and then we add back the center of mass velocity that ensures that the center of mass velocity is unchanged after the collision same as before the collision but the molecules have undergone a rotation so their velocities have equally changed. So, for example if it was going in this direction it will now go in for example this direction. The velocity will undergo a change of the molecules but center of mass velocity that is computed over all the particles in a cell, that is not changed.

Now center of mass velocity is the arithmetic average of the velocity of all the particles in a cell before to the collision and that is same as the arithmetic average of the velocity of all the molecules or solvent particles in the cell after the collision that is being that is being held in there.

$$v_{cm} = \frac{1}{N_{cell}} \sum_{i \in cell} v_i(t_-) = \frac{1}{N_{cell}} \sum_{i \in cell} v_i(t_+)$$

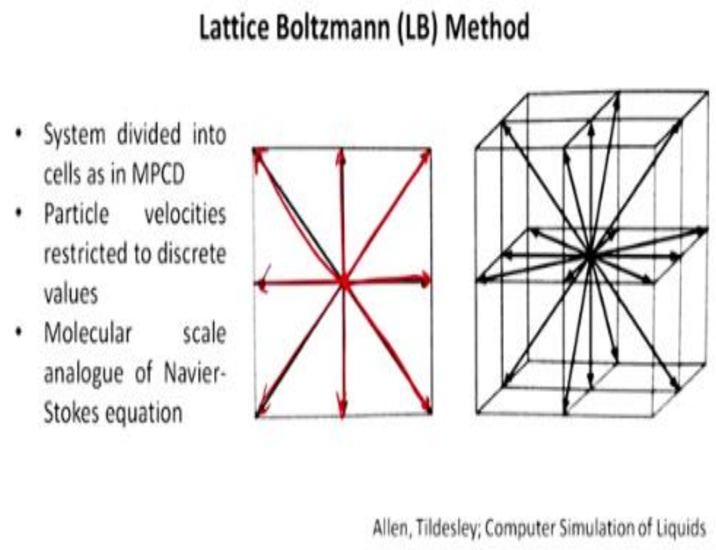
So, in the presence of a solute we may include if we want the solute solvent interactions, but the solvent is still modeled as an ideal gas. So, we will not have any solvent-solvent interactions, we can have solute solvent interactions and we can of course have solute-solute interactions but we will still not have the solvent-solvent interactions. So, in this case the computation will be somewhat slower in comparison to the no solute case. But nonetheless since the solvents already interactions are not being accounted we still have quite an efficient scheme.

The solute particles may be included in the collision step, so let us say for example we have now both solute and solvent. Let us say these guys are my solvent and these guys are my solute then when I perform a collision I look at one of these cells and I rotate it and basically I will rotate all the solvent molecules and all the solute molecules within that particular cell and for the motion of solute is not going to be a free flight because they can collide with each other and they have some size. So, they are not really point particles as solvent particles so for this case we can use the standard velocity verlet integrator that we use in the molecular dynamic simulation.

For the solvent we can still assume the free flight between every two collision that we perform. Again I reemphasize that the collision in this method basically refers to rotation of the cell. It does not refer to the collision between the molecules you may have solute-solute collisions happening the real collisions happening and that is being considered but solvent-solvent collisions are meaningless, because they are considered as an ideal gas you can also have a solute solvent collision that is meaningful but will not have a solvent-solvent collision.

So, we can account for basically the solid solvent interaction by basically disallowing the overlap with this solute in that case, we are accounting for the interaction or we can allow the overlap of solvent with solute and in that case we are disregarding the solute solvent interactions.

(Refer Slide Time: 07:30)



Lattice Boltzmann (LB) Method

- System divided into cells as in MPCD
- Particle velocities restricted to discrete values
- Molecular scale analogue of Navier-Stokes equation

Allen, Tildesley; Computer Simulation of Liquids

So, this is one of the ways of doing it the other way along similar lines, but slightly more elaborate and I would say picking pace in recent years is the Lattice Boltzmann scheme. And in this case the system is again divided into cells, but now the particle velocities are restricted to discrete values that means a particle can go into only some of the possible directions which are represented by these lines here that means the particles are essentially having velocity along some lattice vector and again the math is somewhat complicated and elaborate and I will not go into details in here.

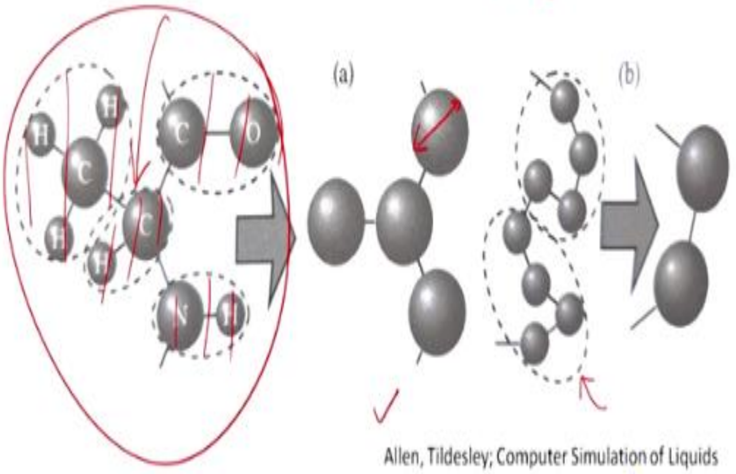
But it turns out that the Lattice Boltzmann scheme is analogous to Navier Stokes scheme at the molecular scale. So, if I write the molecular scale analog of Navier Stokes scheme, it will be the Lattice Boltzmann method pretty much same method we can get if I simply write the Navier Stokes equation for the molecule scale, just analogous to do that.

(Refer Slide Time: 08:47)

Coarse-Graining

Generic Coarse-Graining

Systematic Coarse-Grain



Allen, Tildesley; Computer Simulation of Liquids

So, next we come to the class of methods called the course graining methods and in this case we can preserve both solute and solvent molecules and we can pretty much also have all types of interactions between there between solute-solute, solute solvent and all that. In most typical cases we are not accounting for hydro dynamics and so on and in this case, we are mostly again interested in the equilibrium behavior but we are interested in simulating larger systems than we can simulate using atomistic simulations, this is where the coarse graining become particularly useful but in certain cases where we are interested in dynamics of a system we can also use the coarse graining models.

So, coarse graining essentially is done in two different ways, one of them is a generic coarse graining which is that I simply model my system as comprised of certain entities with some forces between them one of the examples of that is the bead spring model of polymers.

(Refer Slide Time: 09:56)

So, in this case when we build the model we assume that we have some beads connected by some springs and these beads are not really referring to some particular chemistry of the polymer. They are simply a physics based model they can be mapped into a particular chemistry but they are derived with much general description and the idea is to capture the qualitative behavior of course I can make my spring that in this case represents the elasticity of polymer chain I can make my spring stiffer, I can expect make my spring softer, I can make my spring more flexible and so on. But that would be simply capturing the qualitative behavior and not particularly applicable to any particular chemistry.

When we are doing what is known as the systematic coarse graining in that case, we are basically aiming to map the actual molecular structure to a coarse grained system of units where every unit in the coarse grained system will be comprised of many atoms but then there is a clear mapping between the atoms in the atomic picture and the coarse gradient that we have.

So, in generic coarse graining, there is no clear connection between the two although I can build one if necessary but in this systematic coarse graining there is a clear connection, there is one to one mapping between the atoms and the coarse graining unit.

So, let us say I am going to examples here, so in this case we had an amino acid and the dotted lines represent the different coarse green units that have been defined. In this case we have forced coarse grained unit, so these are the ones which are showing up here and simulation is performed over these 4 units.

Now clearly since we have embedded multiple atoms into one coarse grained unit, we cannot say anything regarding what is happening in these three or four atoms or we do not have the resolution at the level of the three, four atoms that are forming the CG unit but we still have the resolution at length scales beyond the CG unit. So, the global behavior of the molecules may be captured at length scales beyond the size of the units that are forming but for the molecular scale feature or atomic scale features less than the CG unit size we have to again resolve to atomistic resolution.

The same thing is done here for another molecule here in this case, we have two units. The question you may ask is how many units, so that I put in a CG unit and that is a difficult thing to answer. In principle you could have said that this entire amino acid is one CG unit is that a worse model than in comparison to this model and there is no clear answer to this because it depends on what question that we are trying to answer. If the computational efficiency is of concern then clearly we should go for just one ball as opposed to four balls connected together but clearly it is not going to have the resolution of anything less than the lengthy scale of that ball that means I will not know anything regarding the molecular configuration of that molecule. Any confirmation change is happening inside the molecule, it is not going to be captured using one ball.

If on the other hand I am using four balls, of course I am capturing change in conformations beyond the unit length that is beyond four atoms because I still have four units to play with I can say how for example a chemical group is changing configuration with respect to another chemical group. For example it can be assist to change where we can if we have a CG unit in the sis configuration it can go to the trans configuration.

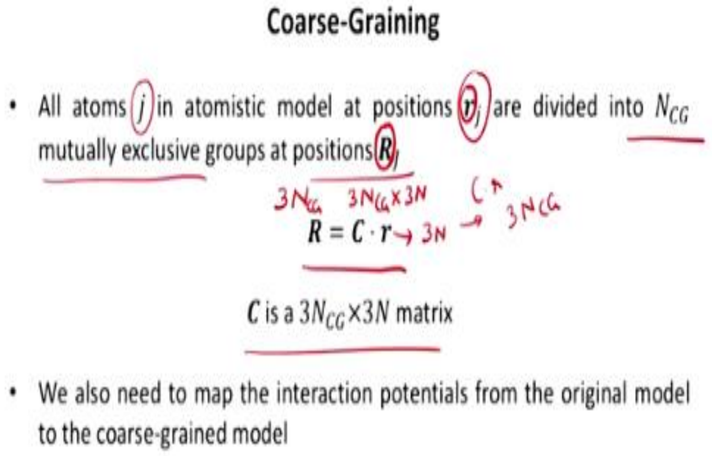
So, using an appropriate choice of the CG unit if your number of atoms are small enough we can capture that kind of confirmation change. So, it really depends on what exactly we seek to address. If we are really interested in long length scale phenomena at which the confirmation changes within the molecule is not so important in that case, we should go for as low number of CG units as possible.

On the other hand when we are interested in the molecular level or atomic level changes. In that case, we should go for larger number of GG units. Now, of course the largest number of

CG units we can take is basically the case of atomistic simulation where every atom is one CG unit. So, that is of course possible, but that is going to be computationally inefficient so that is the reason why we are going for the CG simulation and that is where the compromise comes in and one has to think of what phenomena that we are trying to address and will the CG model be able to capture that phenomena or not that is the first question that we should try to address.

So, we can work in two ways, we can try to start with the minimum possible number of CG units and see how much of the physics we can capture and if we are able to capture the physics, we can leave with that and then start increasing the number of CG units to capture more physics as needed or we can start from the atomistic simulations and systematically think of first lumping two atoms then three atoms than four atoms and look at how much loss in resolution is happening as I increase the number of CG units and both these strategies have been used it is simply a matter of art as opposed to I would say a clear design rule of what should be my CG unit and for every molecular class of systems, we will have a whole host of CG configuration that are possible.

(Refer Slide Time: 17:07)



Coarse-Graining

- All atoms j in atomistic model at positions r_j are divided into N_{CG} mutually exclusive groups at positions R_J

$$R = C \cdot r$$

C is a $3N_{CG} \times 3N$ matrix

- We also need to map the interaction potentials from the original model to the coarse-grained model

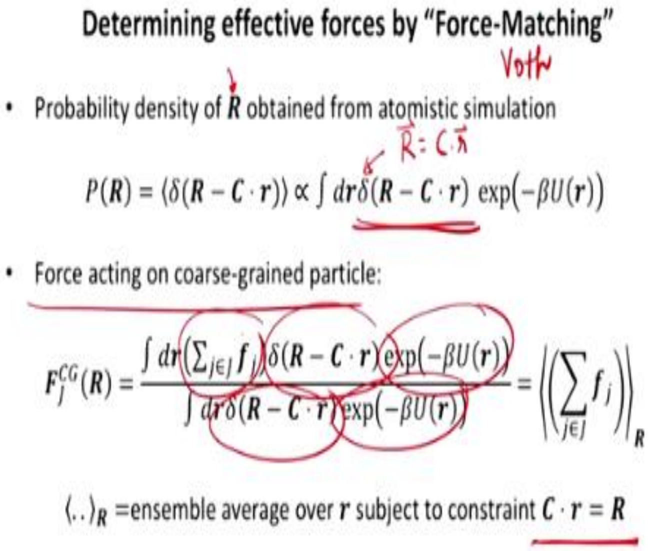
So, in general if I define the small j as the index of the atoms in an atomistic model and they are located at positions r_j. And if I am saying that I want to represent using some N_{CG} coarse grained units which are of course mutually exclusive groups. That means that no atoms are common between the two CG units located at positions R_J. Then I can pretty much map the positions of the CG units using the positions of the atoms in the system

$$R = C\,.r$$

And that can be written in general in this particular form where C is going to be a $3N_{CG}$ multiplied by 3N matrix because R is basically a 3N vector and $3N_{CG}$ vector and C if it is $3N_{CG}$ times 3N matrix and this r is 3 N vector. So, basically taking the product will give me C dot r as a $3N_{CG}$ vector.

So, not only we have to do the structure mapping that is we have already done using this diffusion of R but we also have to map the interaction potentials between the atomistic model and the coarse grain model and there are multiple ways we can do that.

(Refer Slide Time: 18:41)



Determining effective forces by "Force-Matching"

Voth

- Probability density of $\boldsymbol{R}$ obtained from atomistic simulation

$R : C\cdot\vec{r}$

$$P(\boldsymbol{R}) = \langle \delta(\boldsymbol{R} - \boldsymbol{C}\cdot\boldsymbol{r})\rangle \propto \int d\boldsymbol{r}\,\delta(\boldsymbol{R} - \boldsymbol{C}\cdot\boldsymbol{r}) \exp(-\beta U(\boldsymbol{r}))$$

- Force acting on coarse-grained particle:

$$\boldsymbol{F}_J^{CG}(\boldsymbol{R}) = \frac{\int d\boldsymbol{r}\left(\sum_{j\in J}\boldsymbol{f}_j\right)\delta(\boldsymbol{R} - \boldsymbol{C}\cdot\boldsymbol{r})\exp(-\beta U(\boldsymbol{r}))}{\int d\boldsymbol{r}\,\delta(\boldsymbol{R} - \boldsymbol{C}\cdot\boldsymbol{r})\exp(-\beta U(\boldsymbol{r}))} = \left\langle\left(\sum_{j\in J}\boldsymbol{f}_j\right)\right\rangle_R$$

$\langle..\rangle_R$ =ensemble average over $\boldsymbol{r}$ subject to constraint $\boldsymbol{C}\cdot\boldsymbol{r} = \boldsymbol{R}$

One of the methods is called the Force Matching method coming from the group of Greg Voth in University of Chicago and in this method the probability density for the value of R of CG unit that is the coordinates of the CG units are obtained from the atomistic simulations. And in this case it is given by essentially the average of the δ function of r minus C dot r that is giving me the probability density.

$$P(R) = \langle \delta(R - C\,.r)\rangle \propto \int dr\delta(R - C\,.r)\exp(-\beta U(r))$$

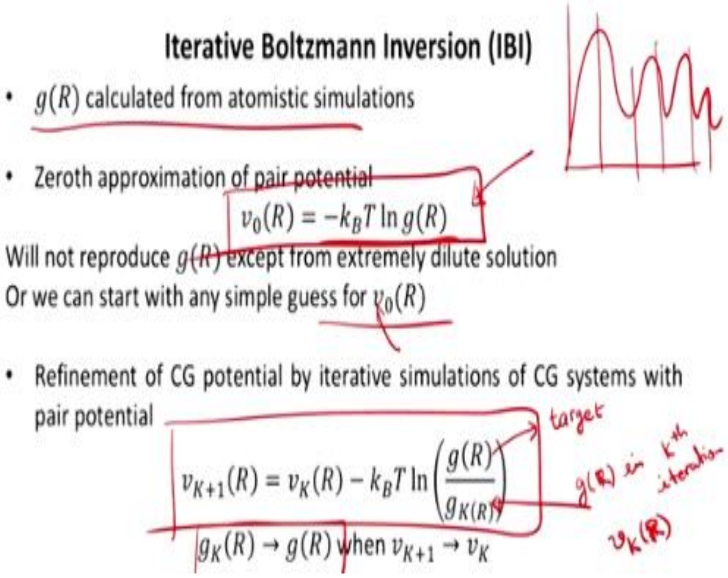
Similarly, I can find the force acting on the coarse grained particle, but now the ensemble averaging will be done subject to the constraint that R is equal to C dot r, that is the mapping that we are trying to achieve. So, therefore again, we introduce a δ function along with the Boltzmann weight that is already there and this becomes the modified partition function when I set R is equal to C dot r.

$$F_J^{CG}(R) = \frac{\int dr\,\left(\sum j \in J\, f_j\right)\delta(R - C\,.r)\ \exp\left(-\beta \mathrm{U}(\mathrm{r})\right)}{\int dr\delta(R - C\,.r)\exp(-\beta U(r))} = \left\langle\left(\sum_{j\in J} f_j\right)\right\rangle$$

And then finally we can compute the total force on a particular CG unit as sum of the forces acting on the individual atoms in that unit that means if I go back to this picture, let us say if I am looking at this particular unit 1 then in that case, I can find the force on 1 as the force of this carbon on 1 this hydrogen on 1, this hydrogen on one and this hydrogen on 1.

So, we can find forces acting in the atomistic simulation on these atoms and if I add them together, I get the force acting the CG unit. Besides what is done there and then if I basically average over that canonically we can get the force acting on the CG unit or the effective force acting on the CG unit.

(Refer Slide Time: 21:22)



Iterative Boltzmann Inversion (IBI)

- $g(R)$ calculated from atomistic simulations
- Zeroth approximation of pair potential

$$v_0(R) = -k_B T \ln g(R)$$

Will not reproduce $g(R)$ except from extremely dilute solution

Or we can start with any simple guess for $v_0(R)$

- Refinement of CG potential by iterative simulations of CG systems with pair potential

$$v_{K+1}(R) = v_K(R) - k_B T \ln\left(\frac{g(R)}{g_{K(R)}}\right)$$

$$g_K(R) \rightarrow g(R) \text{ when } v_{K+1} \rightarrow v_K$$

There is one more approach we can do where the goal is not to match the forces but instead match the structure of the system and in that case what we do is we compute the pair distribution function g of R from atomistic simulations and then we use the idea that we discussed in the context of the potential of mean force we can find the 0th approximation I can say of the pair potential as something like-

$$v_0(R) = -k_B T \ln g(R)$$

This will not precisely reproduce gR except for the extremely dilute solution because of course you will have beyond 2 particle interactions in the system. So, in that case, we have to somehow correct this particular pair potential, but that can be still a very good starting point. So, now there what we do is we refine this potential successively in a manner that pretty much fits the g of R, I am interested in. It turns out that we may not even have to start with this particular potential, I may start with any simple guess for the pair potential. Let us say I want to work with a Lennard Jones potential and even starting from that can give me the correct pair potential in the end because it is an iterative scheme so the way the iterative scheme works is this g of R is a kind of becomes my target value and g k R is the value I get in the kth iteration and for every iteration I am performing a whole CG simulation.

$$v_{k+1}(R) = v_k(R) - k_B T \ln\left(\frac{g(R)}{g_k(R)}\right)$$

$$g_k(R) \to g(R) when\, v_{k+1} \to v_k$$

This pair potential I am using the CG simulation or the coarse grained simulation and in one iteration I am using basically something like a v_K the pair potential and that pair potential is refined in the successive iteration. The first iteration we start with v_0 R and after one iteration we will work with v_1 R and then v_2 R and so on and then finally what this solution can tell you is that when this thing converges that means when v_{K+1} becomes equal to v_{KR} or when it approaches v_{KR} in that limit the g of R will also approach g_{KR} or vice versa and this is the thing that we want, we want to capture the pair distribution function.

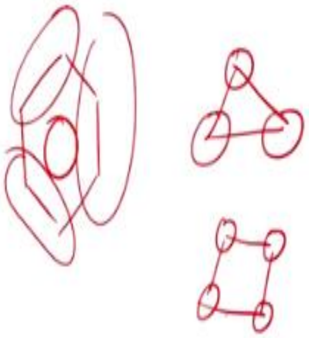
So, we have started with the target pair distribution function and we have refined our pair potential estimates by successively performing or interactively performing the coarse grained simulations. It turns out that it is easier said than done because there are couple of problems in this approach.

The first problem is let us say if I begin with this particular relation then g of R, in fact is actually a distribution and this can have any arbitrary say. So, the pair potential form that I can get from this relation may not be very straight forward, it can be actually quite complicated mathematically to work with so if the pair potential form is some complicated then of course, there is not too much advantage of going for the coarse grained simulation either because every time I have to basically compute using a very complicated pair potential that is the first disadvantage of this that can be gotten away by using educated guess of the pair potential or maybe using the approximation that we had already obtained we can basically fit a function to that and work with that.

The second problem however that is difficult to get away is many of the features in that g of R is going to depend on the atomic or molecular scale features for length scales less than the CG units that we have defined. So you may have, for example have a hydrogen bond happening between two atoms in the system and that is really happening between the atoms. As soon as I, coarse grained we clearly are not capturing that hydrogen bonding unless we have separated them into two different units and even when we have done that we have to be concerned how big the units that we have defined and so on.

The second problem that we are going to have is that although we are trying to keep the same molecular structure, the confirmation changes in a model of the reduced degree of freedom may not perfectly match the conformation changes in the affiliated domestic model and that happens because you are working with a system containing fewer number of particles.

(Refer Slide Time: 26:48)



So, let us say for example we are looking at a benzene molecule. So, benzene molecule as we know is a planar structure. Now, let us say for example, I club the atoms in a way that we have three CG units in presenting that. So, now we are working with a model containing three CG units that clearly ensures that we are going to have some kind of a planar structure, because three units forms a plane, that is not really a problem but let us say we want to work with four units so in that case what we will have is something like this.

Now one of them can very well go out of the plane. Now since we are not including the atomic level interactions and we are working with some kind of a fitted kind of a potential, how do we ensure that the molecular configuration will remain the same as in the original model. It is still to be seen whether we are able to capture the same changes by removing some atoms because if it has been so easy why we are doing atomistic simulation the first place.

So, the reason why we are doing atomistic simulation is because these precise details are not being captured in other simulations. If we could coarse grain in a very systematic way and we keep on removing degree of freedom while keeping the same physics, it would have been the most ideal situation and in general that is never were true so we always make a compromise with the kind of physics that we can capture or that we desire to capture and make sure that the atomistic model and CG model they basically agree to the extent that we can work with of course that will not be 100% agreement, it can be 70% agreement, 50% agreement depending on the requirements on the case to case basis and then further use CG model other scenarios where we want to apply it or where atomistic simulations cannot be performed and that is the typical motivation behind the coarse grained simulation. It is not possible to capture all the aspects of a molecule using a coarse grained description that is completely I would say impossible. So, therefore we always make a compromise when we look at the coarse grained simulation model.

So, with that I want to conclude discussion today. In the next class, we will discuss some case studies of the molecular simulations where I apply some of the techniques that we have discussed in the last couple of lectures, thank you.