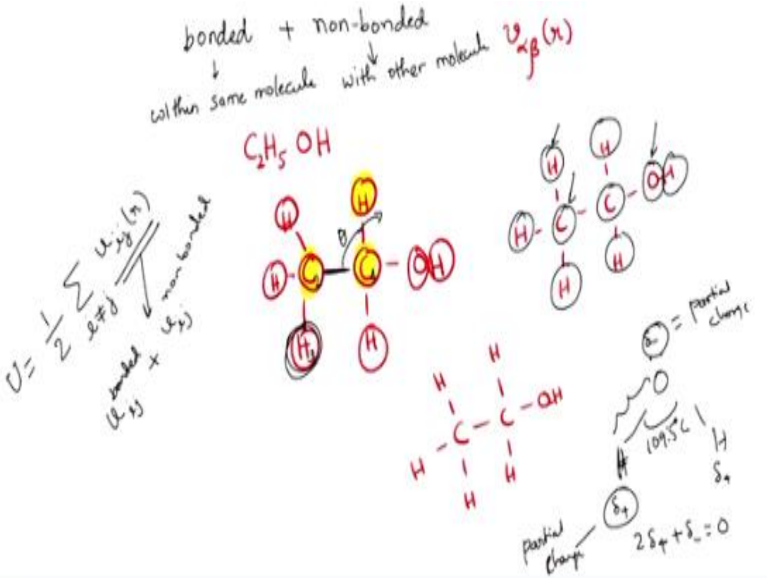
Advanced Thermodynamics and Molecular Simulations
Prof. Prateek Kumar Jha
Department of Chemical Engineering
Indian Institute of Technology, Roorkee

Lecture – 44
MD Simulations – Analysis of Simulation Trajectories (Continued), Case Studies I

Hello all of you. So, in the last lecture we discussed the idea of the pair potentials and how can we efficiently compute the interactions in a system. So, now I want to come to the idea of force fields that is essentially an extension of the idea that we discussed.

(Refer Slide Time: 00:46)



So, earlier I was telling you that we will have pair potentials $v_{\alpha\beta}$ (r) where α and β are different especially in the system, mostly the examples we have done so far we were having a single component system so we only had a v of r. But now let us look at a molecular system and see what kind of interactions do result in that and how can I represent in terms of the pair potentials that we had already discussed and maybe we require some more pair potentials in this case.

Just for the sake of an example I want to think of a system of ethanol. So, let us say you have ethanol in the liquid form present or it can be a gas form as well does not really matter. So, ethanol has the formula C_2H_5OH which can be written as CH_3CH_2OH and now if you imagine liquid ethanol then it contains many of the ethanol molecules. So, you will have other ethanol molecules and I am showing only three of them.

So, now if I want to represent the interaction energy of this system we already have said that we will assume that the electrons are following the nucleus. So, we are dealing with atomic motions. So, the basic entity in our model at the atomistic simulation model is my atoms. So, we think of each of the atoms as the entities. So, I am representing my ethanol as something like 6 hydrogen atoms we have 2 carbons and 1 oxygen atoms.

Now clearly when we look at the interaction between 2 ethanol molecules it is essentially the interaction between the atoms of the ethanol molecule with another ethanol molecule. In fact there can be interactions present between atoms of the same molecule. So, for example let us say if I look at this hydrogen atom here. So, the hydrogen atom if you look at it is forming a bond with this carbon so there is some sort of a bonded interaction with this carbon, even if I look at this particular carbon even if it does not form a direct bond it is kind of very close to that carbon along the molecule. So, if you think about it, this hydrogen is connected to a carbon and that is connected to that carbon. So, let me number this as C1 and C2. So, hydrogen and C1 clearly are connected by a bond. But C2 if you think if you want to look at C2, C2 is not like any other carbon in the system. C2 is connected to the carbon that is connected to the hydrogen that I am interested in.

So, therefore we need to somehow incorporate the fact that when I compute the interactions how the two atoms in a pair are bonded together. So, this is what it call bonded interaction which can be a direct bond like H, C1 in the example or it is like the bonds with a bonded atom this is what I am showing for the case of C 2 here. So, C2 is two atoms apart from this along the or it is two bonds apart from this. And you may imagine that let us say for example if I want to pull the C1 H bond then in that case C2 will is going to be affected. So, whatever the bond at C1 has, has an effect of on C2. So, in a similar manner C2 has an effect on the hydrogen of C1 H.

Now this hydrogen also interacts with the hydrogen the carbon and the oxygen of the other ethanol molecules. Now if you look at those hydrogen atoms in this molecule the black guys here this hydrogen this carbon or this oxygen in fact it interacts with all the atoms in all the molecules. Then there is no bonds in between them. That is first of all there is no bond, there is no immediate bond but there is no bond even with the bonded atoms. There is no connection between this hydrogen to the other atoms in the other molecules. Those are what I call non-bonded interactions.

So, this is within same molecule now although I have looked only until C2 here. But let us say even if I look at for example this hydrogen atom let me give it a name H3. Now even H3 is connected to the H that I was looking at because H is connected to C1, C1 is connected to C2, C2 is connected to H3. So, in some way H is linked to H 3 as well. So, all the atoms within the same molecule will form some type of bonded interaction because any disturbance in any of the atoms in the molecule will affect clearly all the atoms in the molecule. Clearly, the nearest neighbor in the molecule will have stronger interaction the farther we go along the molecule let us say HC1 will be strongest HC2 will be weaker than that HH 3 will be even weaker.

So, as the distance along the molecule increases clearly that the interactions should become weaker. But that does not mean that there is no interaction. So, all the interactions between atoms of the same molecule is what I classify as bonded interactions. And all the interactions with atoms of other molecules is what I call the non-bonded interaction. Therefore the energy of the system that is I can write in terms of the pair interactions u_{ij} r just like I was looking at the hydrogen atom here in a same manner I can pick any of the atoms in the system. And I can have to compute the interaction between every other atom in the system.

$$U = \frac{1}{2}\sum_{i \neq j} U_{ij}(r)$$

So, in fact all these individual pair potentials can be broken into two parts, a bonded part and a non-bonded part. Now what does this bonded part consist of and what the bonded part should do?

So, first of all each of these bonds if you look at quantum mechanically they will have a very strong overlap of electron clouds. So, there should be a very strong interaction between the bonded pair of atoms like HC1 should be a strong interaction, a covalent bond it is a strong interaction. So, that means that I cannot really change the bond length to like any value. Whenever I try to stretch any of the atom in the bond it is going to give me a large energetic penalty.

So, therefore we need to somehow incorporate the fact that any deviation from the bond length will come at the expense of the energy penalty. The same is true also for the second nearest neighbor that forms like an angle. So, let us say for example if I look at this particular angle between at HC1 and C2. So, that angle is formed because of some interaction between at HC1 and C2 and in fact with other atoms also in the molecule and I cannot simply change the angle to any particular value whenever I try to change the angle I am trying to distort the electron clouds of the true bonds that make up this particular angle. And this clearly is not easily possible this has to come at an energetic penalty. So, we need to have some way to incorporate for that particular angle or some penalty for that particular angle.

Similarly, now I can think of four atoms along there so let us say I look at H3 and H. And now you can define some kind of a twisting motion. So, you focus on the C2 C1 bond and you try to move the C1 H out of the plane kind of a twisting movement. So, you keep the C1 C2 bond fixed and you try to rotate the C1 H across the plane of C2 H3. And this will give you some kind of a torsion and that also should come at an energetic penalty. The way to think about it is that there is a whole mesh of electron clouds within a molecule that I found the various bonds. So, any disruption to the electron clouds are going to give me a large energetic penalty. So, even though we are not explicitly modeling the electron clouds we need to capture that particular effect that the bond lengths cannot change significantly, the angles cannot change significantly, the torsions cannot change significantly.

So, whenever I am thinking in this particular way this is what is known as a molecular mechanics kind of an approach. Now clearly we need to have data for the bond length, bond angle, we need to have data for the torsion angles and all these things, they define the geometry of the molecule. But once the geometry is known for us let us say for example if I am doing water and I know that the angle H O H of water is like 109.5 degree and I want to simulate water molecule in our simulation. We should ensure that my oxygen hydrogen bond length and hydrogen oxygen hydrogen angle should remain fixed or should vary in a very limited range because any large variation should come at a strong energetic penalty because realistically speaking molecule cannot depart from its equilibrium geometry significantly. Such departures should be penalized because that is not what physically happens.

So, even though we are not doing full quantum chemistry we should keep the geometry of molecule kind of intact or closer to its equilibrium configuration or minimum energy configuration that we can obtain for example by quantum chemistry calculations. So, in molecular mechanics approach this is how we enforce it we add some potentials which are arbitrary to begin with if you think about it but they ensure that the bonds the angles the torsions are such that they are around their minimum energy configuration that we can obtain by quantum chemistry or their actual geometry that we can find from experiments or their bond length or vibration data that we can get from experiments.

Now if I look at the non-bonded part of it, now in there it is a whole different game. Now two atoms in two different molecules will not interact because of any of the bonded interaction but they will have some other interaction coming from somewhere else. Let us say for example they can have some Van Der Waals interactions or the electrons in one of the atoms may have some electrostatic interaction with electrons in the other atom. Now whenever I say electrons here keep in mind that even when we have a neutral atom so as to speak that also has an electron cloud. So, when the ethanol itself is electro neutral there is no charge on ethanol. If I look at a hydrogen atom the electron distribution around it will have certain shape.

So, if I look at any particular location in the electron cloud there they may have some kind of a charge. And that will give rise to some electrostatic interaction with other atoms in there and this particular idea is referred as the idea of partial charges in the molecular mechanics framework.

So, the way to think about it is let us say if I look at my water molecule, we know that there is a lone pair of electrons for oxygen there. So, although the water molecule itself is quite neutral we can imagine that there is an accumulation of a negative charge here. Let me call it some delta minus and some positive charge there clearly the net sum of that is equal to 0 that is what is necessary for just the water to be electro neutral.

But if I look at the interaction of this oxygen with some other oxygen in some other water molecule that oxygen will see this delta minus charge and this is what we refer as the partial charge. So, whenever we have any molecule there is electron sharing between atoms and because of the electron sharing every atom appears to occupy some positive or negative charge and that charge gives rise to some electrostatic interaction when I look at the non-bonded part of interaction clearly, this also contributes the bonded interactions. But more often than not the bonded interactions are more important than the corresponding electrostatic or Van Der Waals interaction. In fact this is the mechanism by which molecules interact. So, the reason why water molecules form hydrogen bonds is because oxygen is slightly negative and hydrogen is slightly positive. So, this oxygen is happy to form a hydrogen bond with hydrogen of some other water molecule and this is purely electrostatic in nature so, as to speak.

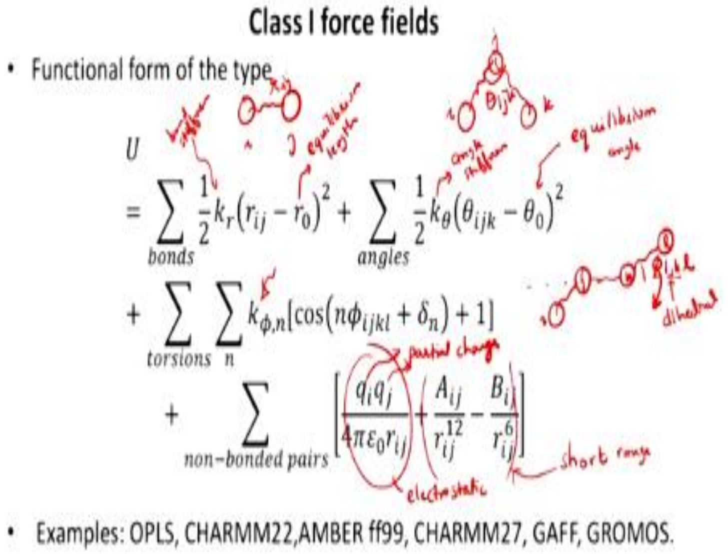
So, we need to have some mechanism to account for electrostatic interactions Van Der Waals interactions or other kinds of the short range interactions that can be present in the system. So, now this is using this particular ideas I can first construct some kind of a mathematical form of a force field and after that the idea remains same as what I was showing for the argon example. So, we know the ab-initio data of argon and we fit a Lennard Jones function and I get the parameters of Lennard Jones potential and using that I can simulate argon. So, in a similar manner I can imagine a mathematical form of a force field which kind of captures the physics I have mentioned that it should have some bonded interaction, some non-bonded interaction but it is still arbitrary in terms of the mathematical form we have chosen when we fit it to some ab-initio data or some experimental data then only we can put that force field to work.

So, it is capturing the basic physics in terms of the behavior we want to capture but the mathematical form is still arbitrary. It really depends on how are we fitting that to the experimental data or some ab-initio data or quantum chemistry data that we have.

So, this is the idea on which the force fields are developed and fortunately for any molecule you want to simulate it is most likely that someone has developed a force field for that and you can pretty much borrow the force field from there most software's that we work with have a capability to work with different kinds of force fields and they have now mechanisms to import data of different force fields and all that so, pretty much all the parameters of the force field of most standard molecules you can find and for the molecules that you cannot find the force field parameters you may have to develop the force field yourself that means you have to compute ab-initio simulations or conduct simulistic calculations or use some experimental data to fit some mathematical arbitrary form most of the application scientists in the area of molecular dynamics and Monte Carlo, do not do that the force development work is typically done by different research groups. So, this is one example of a force field called the class 1 force field.

$$
\begin{aligned}
U = \sum_{bonds} \frac{1}{2} k_r \left(r_{ij} - r_0\right)^2 \\
&+ \sum_{angles} \frac{1}{2} k_\theta \left(\theta_{ijk} - \theta_0\right)^2 \\
&+ \sum_{torsions} \sum_{n} k_{\phi,n} \left[\cos\left(n\phi_{ijkl} + \delta_n\right) + 1\right] \\
&+ \sum_{non-bonded\ pairs} \left[\frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right]
\end{aligned}
$$

(Refer Slide Time: 19:32)



Class I force fields

- Functional form of the type

$$U = \sum_{bonds} \frac{1}{2} k_r \left(r_{ij} - r_0\right)^2 + \sum_{angles} \frac{1}{2} k_\theta \left(\theta_{ijk} - \theta_0\right)^2 + \sum_{torsions} \sum_{n} k_{\phi,n} \left[\cos\left(n\phi_{ijkl} + \delta_n\right) + 1\right] + \sum_{non-bonded\ pairs} \left[\frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right]$$

- Examples: OPLS, CHARMM22,AMBER ff99, CHARMM27, GAFF, GROMOS.

So, in here if you notice I have included the bonds so this is if for example i and j forms an bond and r_{ij} is the is the bond length r_0 is some equilibrium length. So, I want my bond length r_{ij} to fluctuate around this particular value and k_r is a measure of the bond stiffness. So, if I increase k_r, I will have a more stiff bond that means we will allow lesser fluctuation around the average value, as soon as the fluctuation is large we are going to have a strong energetic penalty on the other hand if I use smaller value of k_r my bond is more flexible but in both these cases the bond length will fluctuate around r 0.

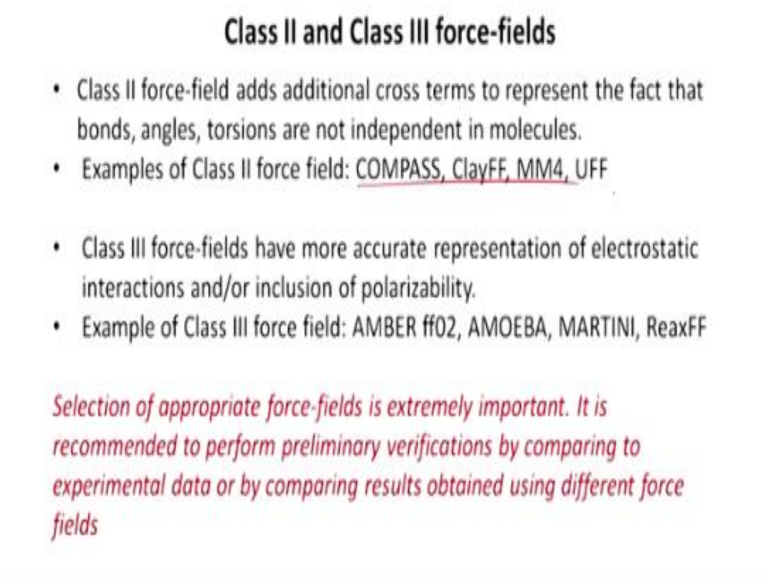
Similarly for a triplet of atoms i j k, I can define the actual angle as θ_{ijk} and θ_0 can be an equilibrium angle and k_θ represents the angle stiffness. The particular forms are called the harmonic forms they are simply used because of convenience. There is no strong reason to use this particular form. It is just mathematically convenient to implement.

This term is for the torsion that is for four atoms so let us say if you have i, j, k and l. So, basically we are looking at the rotation of k l along the j k axis and this will be having some torsion angle Φ_{ijkl} which will give rise to this particular form again k is some kind of the stiffness of the dihedral angle, so Φ is called the dihedral angle or torsion angle and then once we are done with the bonded part we go to the non-bonded part.

So, in there you can have electrostatic interactions where q_i and q_j are the partial charges on the atoms, I want to emphasize again that this is not the actual charge. Even if a molecule is neutral the atoms may have a partial charge and then we may have some other form of short range potential A_{ij} and B_{ij} are some arbitrary constants and we hope that using this functional form we are able to fit different kind of molecules.

Now there are several force fields that use pretty much the same functional form but they differ in the magnitude of the parameters. So, different force fields are developed for different class of molecules it may be that even for the same class there can be two different force fields for that. Those parameters need not be the same because we have so many variables here that there is no unique value of the kr, kθ and all these parameters in there that will fit the behavior that we are interested in. So, the choice of parameterization the choice of parameters of the force field is not unique. Different force fields will have different choice of parameters. And therefore we should take that force fields as empirical at best. So, they are empirical functions that are able to fit the behavior of interest. They are as good as how we are parameterizing it and we should not simply blindly use any force field we should pick the force field that has been shown to work well for the class of molecules that we are interested in simulating. So, for example OPLS, CHARM, AMBER, GAFF and GROMOS force fields. Then there are class 2 and class 3 force fields that are even more elaborate and therefore more accurate but computationally more expensive in general.

(Refer Slide Time: 24:34)



Class II and Class III force-fields

- Class II force-field adds additional cross terms to represent the fact that bonds, angles, torsions are not independent in molecules.
- Examples of Class II force field: COMPASS, ClayFF, MM4, UFF

- Class III force-fields have more accurate representation of electrostatic interactions and/or inclusion of polarizability.
- Example of Class III force field: AMBER ff02, AMOEBA, MARTINI, ReaxFF

Selection of appropriate force-fields is extremely important. It is recommended to perform preliminary verifications by comparing to experimental data or by comparing results obtained using different force fields

So, class two force fields account for the fact that it is not that the bonds angles and torsions are independent of each other. Two bonds may affect each other, a bond may affect the angle, an angle may affect the torsion. So, to account for those dependence between two bonds, a bond in an angle or two angles or angle in a torsion we also can incorporate some cross terms which will essentially mathematically speaking add more terms in our force fields and there are more parameters to work with. So, clearly by introducing more parameters we can get even better fit of the experimental or ab-initio data but it is going to add a burden on the forced computation part of the molecular dynamics or Monte Carlo simulation. And here are some examples of that COMPASS, Clay, MM4 and UFF.

And then class 3 is slightly more elaborate. It has a more accurate representation of electrostatics and it also includes the effect of polarizability that is missing in class 1 and class 2 force fields. And again there are some examples in here.

It is really recommended that any force field you pick for any particular problem it must be verified before we put to actual use and you can verify in several ways. If you have some data from experiments you can run a simulation and see whether you are getting the behavior as experiments or not. And secondly if that is not possible then at least we can compare the performance of different force fields for the particular system.

So, let us say for example if there are three or four possible choices of force fields, we should try running for all these three or four force fields to see how different the results are. Clearly, different force fields will not give you the same result but as long as the results are more or less same or there is no new feature coming in a particular force field, there is nothing to be alarmed about. But if you start to see that the results are drastically different if the features in different force fields are different. In that case we have to try to verify the force field by some other means instead of simply doing comparisons. So, this exercise is compulsory before we look at the results of the molecular synthesis.

(Refer Slide Time: 27:11)

List of commonly used Force Fields

Force field family	Variants	Original Developer
CHARMM	CGenFF, CHARMM22, CHARMM27, CHARMM36	Martin Karplus, Harvard University
OPLS	OPLS-AA, OPLS-UA	William Jorgensen, Purdue University & Yale University
AMBER	GAFF, GLYCAM, ff94, ff96, ff98, ff99, ff02, ff02EP, etc.	Peter Kollman, University of California San Francisco
Consistent force fields	COMPASS, COMPASS II, CFF, CFF91, PCFF	Molecular Simulations Inc. (now Biovia)
UFF	UFF	William Andrew Goddard III, California Institute of Technology
MMFF	MMFF, MMFF94	Merck Research Laboratories
GROMOS	GROMOS 43a1, 43a2, 43b1, 45a3, 53a5 and 53a6	Wilfred van Gunsteren, University of Groningen & ETH Zürich
MARTINI	MARTINI	Siewert J. Marrink, University of Groningen, and Peter Tieleman, University of Calgary

So, here I list some of the commonly used force fields some of them are of class 1 some of class 2 and class 3. Class 1 force fields are the most likely used force fields in fact the GROMOS and lamps which are the most common software’s for MD they mostly incorporate the class 1 force fields. But in certain cases you may have to go to class 2 or class 3 force fields as well. So, with this particular idea I want to conclude today's lecture.

In the next lecture I will start discussing the molecular dynamics simulations and what is the benefit of MD over Monte Carlo simulation. And then in the next week we will start seeing how can we apply to real problems and again the focus would be more on how can we apply the simulations as opposed to the math behind it. But we will touch upon some math as the times comes in.

So, with that I want to conclude, thank you.