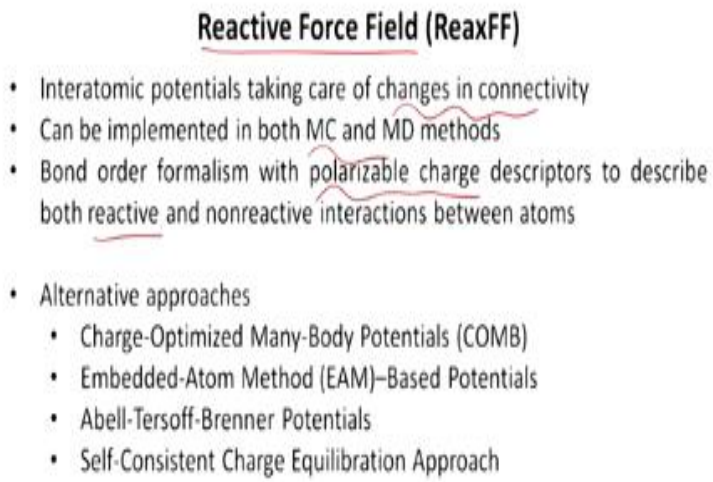
Advanced Thermodynamics and Molecular Simulations
Prof. Prateek Kumar Jha
Department of Chemical Engineering
Indian Institute of Technology, Roorkee

Lecture - 60
Reactive Force Fields; Ab Initio Molecular Dynamics and Other Advanced Methods; Molecular Simulations in Chemical Engineering; Concluding Remarks

Hello all of you, from the last lecture we have discussed the kinetic Monte Carlo method of simulating chemical reactions. In this lecture, I will start with the reactive force fields simulations which are essentially again some classical way of modelling chemical reactions and we will also summarize this course towards the end of the lecture because this is going to be the last lecture of this course.

(Refer Slide Time: 00:55)



So, basically the reactive force field is pretty much following the idea of force field as we had discussed in the atomistic simulation part for the MD simulations, but in here we somehow have the potentials taking care of the changes in connectivity. Now, of course that is somewhat empirical to think about because we are still doing like atoms may not be electrons still but we have some way to account for the changes in connectivity that means the bonds can form and break that is built inside the definition of potential.

So, again, this is somewhat crude way of doing the chemical reactions, but it is quite successful because this has been parameterized for many, many chemical reactions and so on nowadays

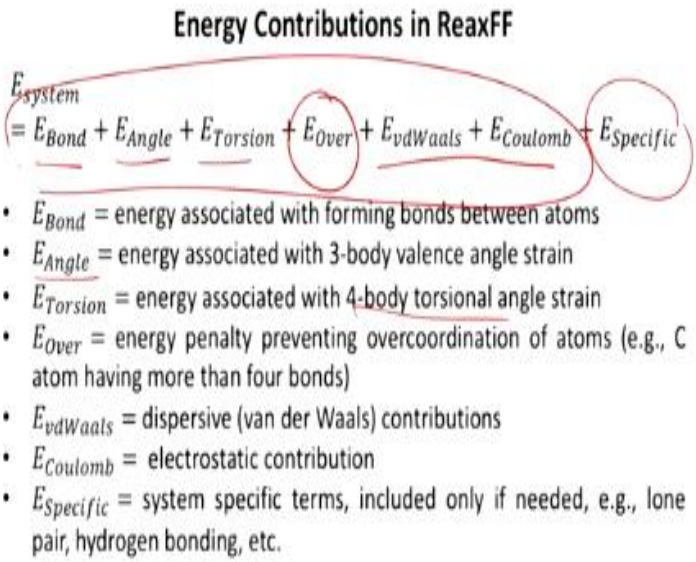
particularly for hydrocarbons and therefore it turns out that the efficiency of this force field at least for the compounds for which it has not correct parameterized is pretty good in many cases comparable to quantum chemistry simulations.

So, this can be implemented within both Monte Carlo and molecular dynamics because it is indeed a force field. The formalism is called what is known as bond order formalism and I will come to that in a minute, but we also include the effect of the charge polarization and we look at both the reactive and non-reactive interactions the reactive are the ones where the bonds are bonding and breaking apart. There are other methods along the same lines such as the COMB potential, the EAM based method, the Abell Tersoff, Brenner potentials and so on that also achieves the same thing but ReaxFF that is I would say the most modern and one of the most common course fields nowadays the model chemical reactions using atomistic simulation.

So, in the energy contributions in the reacts that have force field looks something like this-

$$E_{system} = E_{bond} + E_{angle} + E_{torsion} + E_{over} + E_{vdWaals} + E_{coulomb} + E_{specific}$$

(Refer Slide Time: 03:02)

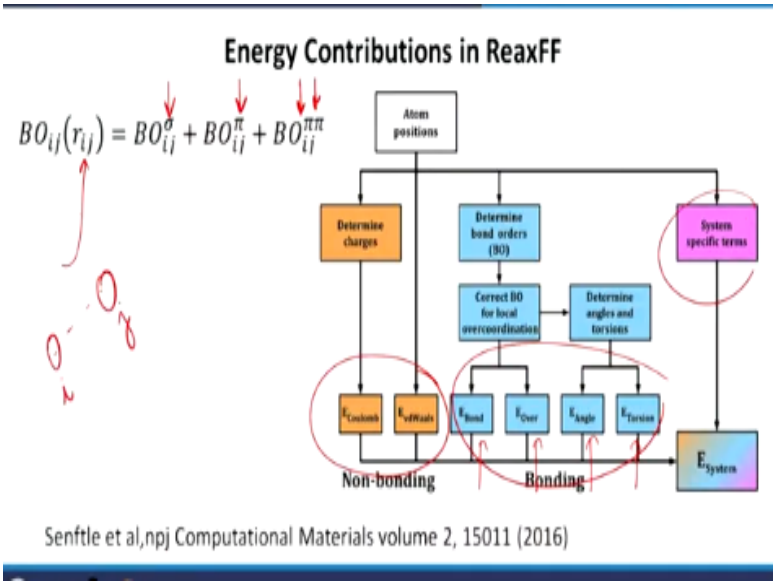


So, you have a bond, angle, torsion contribution as we discussed in the earlier force field descriptions we had but, now in this case we are looking at when I look at the bond energy we are looking at the energy of forming bonds.

Similarly, when we are looking at the angle energy we are looking at energy with forming an angle, energy with forming the torsion or torsion energy and so on, we also include an over

energy that basically represent an energy penalty that prevents the over coordination atoms, for example a carbon atom cannot form more than 4 bonds, so therefore we paralyzed in terms of energy whenever the carbon is having more than 4 bonds for that particular process that particular thing will be not favoured in the simulation. Then we also have some non-bonded interactions such has Van der Waals or dispersive interactions and the electrostatic interactions in addition to that, in certain cases we may require some other interactions such as lone pairs, hydrogen bonds and so on. If they are required for a particular application or if the existing force field is not being able to capture the physics that we are interested but that term is not very common anyway, the other terms are commonly present in most reactions force fields.

(Refer Slide Time: 04:32)



So, the bond order from formalism is something like this-

$$BO_{ij}(r_{ij}) = BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi}$$

-

so we defined the bond order as a function of the distance between the atoms there are forming a bond the r_{ij}. So, let us say this is my i and this is my j and it consist of three parts one is the bond order corresponding to a σ bond the other is for π and the third is for π, π. So, basically we are looking at basically the bond order corresponding to single bond, double bond, triple bond so on.

So, now there are two non bonded interactions that does not depend on the bond order because they are not related to bonding of atoms they are present between the atoms which are not bonded. So, they depend on the charges of the atoms or the partial charges as we had discussed earlier, but the other bonded interactions now depend on this bond order that signifies the

presence of bonds in there and in there we have the bond angle torsion and the over coordination and we of course we correct the bond order for the local over coordination whenever we have that.

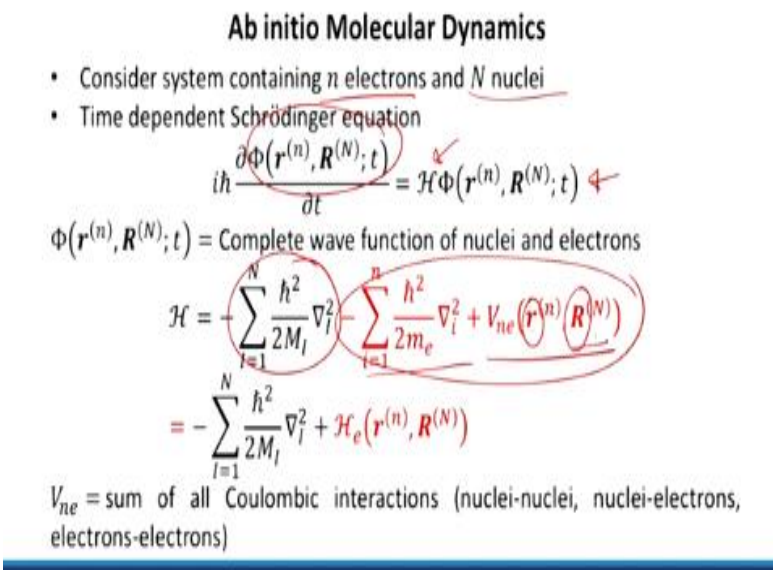
Then, we can also have some system specific terms whenever it is needed and we get the total energy system and once we have that we can pretty much follow the same kind of method that we can we are we have followed let us say a Monte Carlo or molecular dynamics we simply replace the force field by this and now since the method allows the bonds to form and break because we are looking at atoms as an entity but we are defining some kind of a bond order signifying the presence of bonds and so on.

So, we are pretty much following the same classical picture but we have some mechanism to account for the formation of bonds not only bonds but we actually looking at specifics of the bond it can be a σ bond or a π bond or a double π bond or a single bond, double bond or a triple bond.

So, this method has been applied particularly for hydrocarbons, but also has been there for many other molecules. In fact, it can pretty much cover now, good number of atoms in the periodic table and more parameterizations are coming and many software's are now including the reactive force fields. So, particularly for the atoms where these force field parameters are available they can be used with confidence, especially for the reactions where they have already been used or for which the method has been parameterized.

So, with that I conclude the discussion on the chemical reactions and now I want to briefly cover what comes under the quantum chemistry simulations, because there are a whole methods of the class what is known as Ab Initio molecular dynamics that reads like molecular dynamics but the method itself is very different because we are no longer solving the equation of motion from the Newton's laws we are solving the Schrodinger equation.

(Refer Slide Time: 08:13)



So, in this case we are interested in getting the density of electron clouds as a post to the coordinates of atoms and so on but still it borrows the name of molecular dynamics because some of the ideas are common with the molecular dynamics that as we know.

So, in this case we consider a system containing small n electrons and large N nuclei and then we can of course write the time dependent Schrodinger equation as essentially a function of what is known as the complete wave function-

$$ih\frac{\partial \phi(r^{(n)}, R^{(n)}, t)}{\partial t} = H\phi(r^{(n)}, R^{(n)}; t)$$

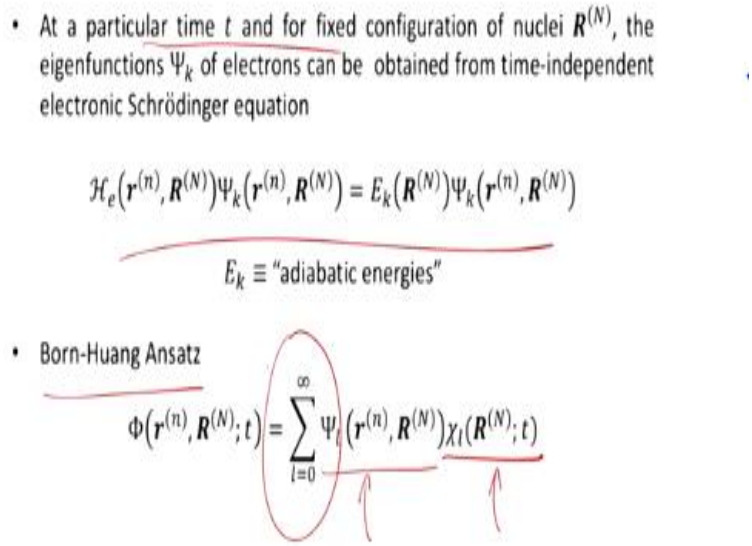
that includes the wave function of nuclei and electrons or the degrees of freedom of nuclear and electrons and then there is an Hamiltonian that appears there that is an operator whenever we are doing quantum mechanics in classical mechanics the Hamiltonian was in a scalar quantity. But in this case it is going to be an operator which I can pretty much write as-

$$H = -\sum_{I=1}^{N} \frac{h^2}{2M_I}\nabla_I^2 - \sum_{i=1}^{n} \frac{h^2}{2m_e}\nabla_i^2 + V_{ne}(r^{(n)}, R^{(n)})$$

It consist of two components one that refers to the kinetic energy of the nuclei, so it only depends on when nuclei, but the second and third that are basically affected by the electrons the second term is the kinetic energy of the electrons and the third term is all sorts of electrostatic interactions that are present with such as the nuclei-nuclei, nuclei-electrons and electron-electron interactions all of them depends on the electrons in the system. Small r represent the positions of the electrons and capital R represents the position of the nuclei in the

system and I use the notation r and to indicate that there are n such positions for an electrons and similarly for capital R.

(Refer Slide Time: 09:50)



So, then we use an approximation and that is at particular time and for a fixed configuration of nuclei we can write the Eigen function of electrons as the solution of this equation-

$$H_e(r^{(n)}, R^{(n)})\psi_k(r^{(n)}, R^{(n)}) = E_k(R^{(n)})\psi_k(r^{(n)}, R^{(n)})$$

where E_k is what is known as adiabatic energies or what is known as the ground state that is considering a time independent Schrodinger equation.

And then we have what is known as a Born Huang Ansatz that basically factors the complete wave function into the nuclei degree of freedom that is the chi term and the electron degree of freedom that is the psi term or the electronic wave function ψ and the nuclei wave function chi and since we are working with Eigen functions, this has to be summed over all possible values of n.

$$\phi(r^{(n)}, R^{(n)}; t) = \sum_{l=0}^{\infty} \psi_l(r^{(n)}, R^{(n)})\chi_l(R^{(n)}; t)$$

(Refer Slide Time: 10:56)

$$i\hbar\frac{\partial\Phi(r^{(n)},R^{(N)};t)}{\partial t}=\mathcal{H}\Phi(r^{(n)},R^{(N)};t)$$

$$\Phi(r^{(n)},R^{(N)};t)=\sum_{l=0}^{\infty}\Psi_l(r^{(n)},R^{(N)})\chi_l(R^{(N)};t)$$

$$\Rightarrow i\hbar\frac{\partial}{\partial t}\sum_{l=0}^{\infty}\Psi_l(r^{(n)},R^{(N)})\chi_l(R^{(N)};t)=\mathcal{H}\sum_{l=0}^{\infty}\Psi_l(r^{(n)},R^{(N)})\chi_l(R^{(N)};t)$$

$$\Rightarrow i\hbar\sum_{l=0}^{\infty}\Psi_l(r^{(n)},R^{(N)})\frac{\partial}{\partial t}\chi_l(R^{(N)};t)=\mathcal{H}\sum_{l=0}^{\infty}\Psi_l(r^{(n)},R^{(N)})\chi_l(R^{(N)};t)$$

So, then we basically put this factorized expression in my time independent equation and I get basically after certain manipulations and equation like this-

$$ih\sum_{l=0}^{\infty}\psi_l(r^{(n)},R^{(n)})\frac{\partial}{\partial t}\chi_l(R^{(n)};t)=\sum_{l=0}^{\infty}\left(-\sum_{I=1}^{N}\frac{h^2}{2M_I}\nabla_I^2+H_e\right)\psi_l(r^{(n)},R^{(n)})\chi_l(R^{(n)};t)$$

(Refer Slide Time: 11:11)

$$i\hbar\sum_{l=0}^{\infty}\Psi_l(r^{(n)},R^{(N)})\frac{\partial}{\partial t}\chi_l(R^{(N)};t)$$

$$=\sum_{l=0}^{\infty}\left(-\sum_{I=1}^{N}\frac{\hbar^2}{2M_I}\nabla_I^2+\mathcal{H}_e\right)\Psi_l(r^{(n)},R^{(N)})\chi_l(R^{(N)};t)$$

Integrating over $r^{(n)}$ and some further assumptions results in **Born-Oppenheimer approximation**

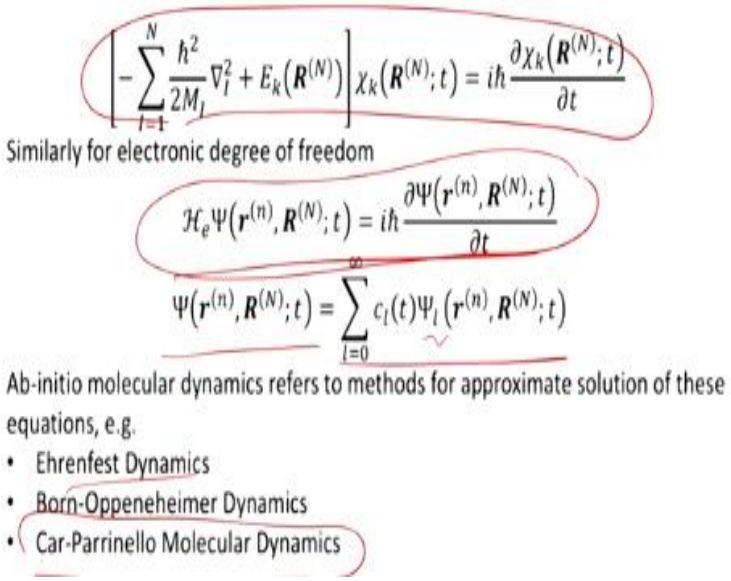
$$\left[-\sum_{I=1}^{N}\frac{\hbar^2}{2M_I}\nabla_I^2+E_k(R^{(N)})\right]\chi_k(R^{(N)};t)=i\hbar\frac{\partial\chi_k(R^{(N)};t)}{\partial t}$$

And, after that basically integrate over all the coordinates of the electrons and we make some more approximations along those lines and what we then get is the Born Oppeneheimer approximation that is an equation that gives me nuclei degree of freedom, because it is written in terms of the chi variable that chi is the nuclei wave function.

I do not want you to understand the complete math behind it but the basic idea that we; are essentially separating the complete wave function into a nuclei part and an electron part so, that

we can write the basic equations that we can solve in the simulations. So, all of this is done to make this Schrodinger equation simpler enough to be able to numerically solve it.

(Refer Slide Time: 12:16)



So, then we can do a similar exercise for the electron wave functions as well and we can get an equation for that-

$$\left[-\sum_{I=1}^{N}\frac{h^2}{2M_I}\nabla_I^2 + E_k(R^{(N)}\right]\chi_k(R^{(N)};t) = ih\frac{\partial\chi_k(R^{(N)};t)}{\partial t}$$

So, now we have written the Schrodinger equation as basically two equation one for the nuclei degree of freedom and one for the electronic degree of freedom and then for the electronic wave function ψ-

$$\psi(r^{(n)},R^{(n)};t) = \sum_{l=0}^{\infty} c_l(t)\psi_l(r^{(n)},R^{(n)};t)$$

I again represent that into what is known as basis functions ψ_l and it is these basis functions that we solve for in any quantum mechanics software.

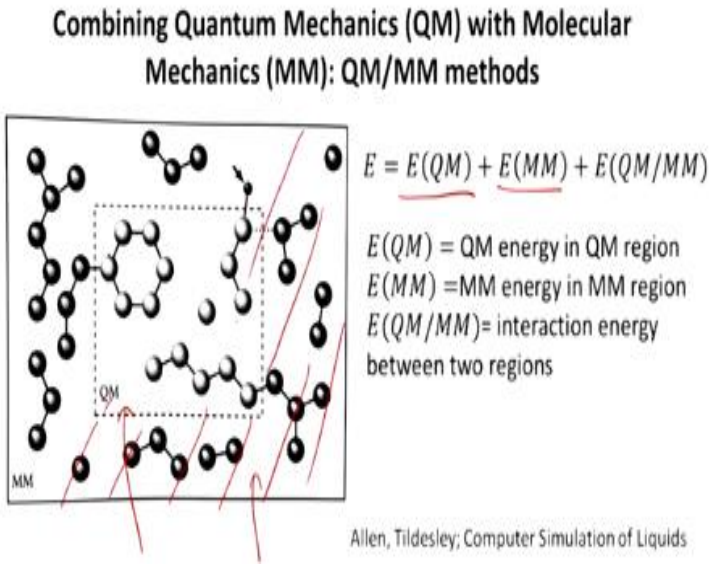
So, this solution is what is known as the ab initio molecular dynamics and the idea why I wanted to present this is to say that we are solving very different equations than in comparison to what we are doing in the classical mechanics world, ab initio molecular dynamics is clearly more accurate for the quantum description of the system in the classical mechanics picture, we of

course miss anything to do with electrons and nuclei because atoms are the entities.

Again, there are several methods in this area ab initio dynamics such as Ehrenfest dynamics the Born Oppeneheimer dynamics, Car-Parrinello dynamics, the last one is the one that is quite popular and all of these methods basically solve for electron wave functions and therefore in principle we can simulate also chemical reaction using Ab initio methods but these are computationally highly prohibitive and what any system beyond certain number of atoms, this is completely not feasible using the software's that we have but, of course our capabilities are improving this time and we are able to do Ab initio for more and more systems as the computing power improves.

But none the less as I have been telling you in the context of kinetic Monte Carlo simulations, these simulations can be used to train atomistic solutions or for example get the force field parameters for atomistic simulations and to the extent these methods are quite useful because, we can do look at a smaller subset of the system perform a ab initio on that and use that to get the force field and then do an atomistic simulations on top of that and this is an approach that is also becoming quite the popular in the community.

(Refer Slide Time: 14:52)



The next thing that one can do and that is closer to what we have been doing again is that we can treat one part of the system or a small subsection of the system using quantum mechanics and the larger part of the system using what is known as the molecular mechanics, that is basically the atomistic force fields that we have been discussing and by doing that as long as

the subsection of the quantum mechanics is a small enough. Let us say for example, I want to look at solvation around the protein residue then that small bit of the residue I can treat using quantum mechanics and all the solvent I can handle using molecular mechanics or atomistic simulation as long as the quantum mechanics part is small enough that is still quite feasible. But in this case, we are still doing both quantum chemistry and the molecular dynamics for Monte Carlo simulation.

$$E = E(QM) + E(MM) + E(\frac{QM}{MM})$$

And, we define a net energy as the energy of the quantum mechanics reason the QM region here obtained using quantum mechanics, then we have the energy in the MM region here that is the region right here using the molecular mechanics or MD or MC whatever method we are using and then finally since there is some interaction between the QM and MM region, we have to also account for that interaction energy between the two regions and, this can then go into your atomistic simulation so we still need to perform quantum chemistry calculation, but we are doing it only for a small subset of the large system.

So, with that we are pretty much done with things with the quantum mechanics there is yet another scale that one can probe that it is towards the larger end of the spectrum if we want to simulate particles much larger than what we have discussed in this course and, there are set of methods called a discrete element methods which are again quite similar to molecular dynamics, but they are applied to much larger particles, particularly granular particles or powders and so on and in there what is different from the particles that we have been concerned in this course is that the thermal energy is not so important because they are quite heavy particles there is too much contact or friction between the particles which are playing the major role, the shape of particles plays a major role and therefore the methods try to include those factors into the algorithm although it is still solving the equation of motion and in fact that extent it is still like molecular dynamics, but it is a more general class of methods applicable to particles much larger than the particles that we have done in the particular course.

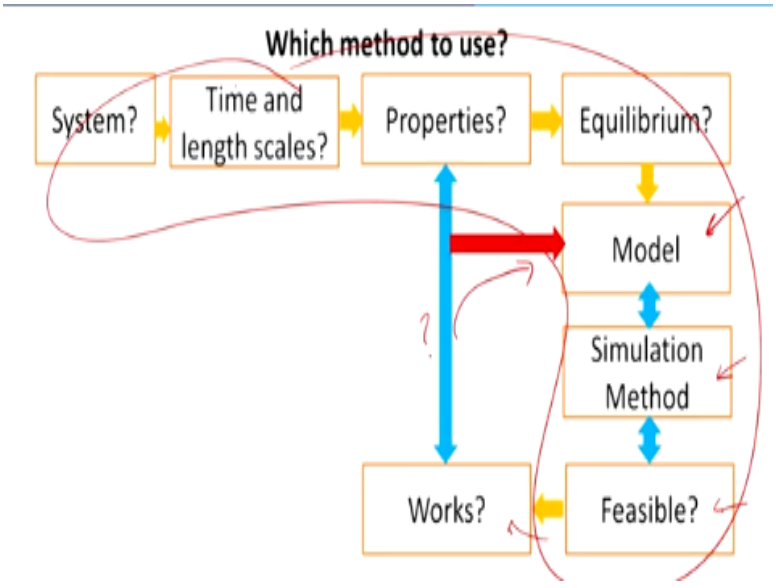
So, with that now, we have a much larger question that we should stick to address as we also come towards the end of this course which method should we use and the answer is not so trivial. So, we first have to think of what system do we want to simulate what exactly should be the boundary of the system and what should be constituent of the system. So, clearly all the

methods have their own limits in terms of systems sizes, they can simulate and, in terms of phenomena that they can simulate the choice of system itself pretty much reduces our choice of the method that we can pick.

The next thing we should ask ourselves, what is the time and length scale that is important here, whatever phenomena that is happening in the system or there is our interest what is the time and length scale of that process that will further reduce the number of methods that we can use for the particular problem then we should finally look at which properties are we interested in and these should be the target properties that whatever simulation we are doing should be able to give me.

The next question we should ask are we looking at systems at equilibrium or can we make an equilibrium approximation or not? If we are doing equilibrium, then we pretty much have the most of the methods that we have discussed in this course but if we are at non equilibrium, or equilibrium effects are very important then we have to choose I would say from a very different class of methods.

(Refer Slide Time: 19:42)



Once we have answered all these questions then we have to build a model of the system and the model will be specific to the simulation method. Let us say for example, for the same polymer if I am using atomistic simulation the model is going to be different because now atoms are the entities. On the other hand, if I am doing the course grain model, then the course grain entities become the unit there so, the model itself changes depending on the simulation

method.

Using the model and the simulation method we now have the construction of a simulation that we can possibly perform, but now we have to ask whether that simulation is going to be computationally feasible or not because if it is not feasible then you should rethink about our choice of model and simulation method you should maybe even go for a smaller system size so, that the system would be computationally feasible many a times we may actually do a simulation and realize that it is not feasible but many, times we can also know our previous experience that it is not going to be feasible.

Once we have done that then we finally run the simulation, because although it is cheaper than experiment it is still takes computational power and it takes the t of time so only after making all these checks we should go to run the simulation and, therein we can make the predictions of the properties that we were initially considering to look at if we get a good match or if we get a reasonable agreement so that we are happy with that it is good otherwise we can always go back and define the model.

So, we should not start thinking from the simulation method perspective itself that I want to use Monte Carlo on this problem but instead, this would look at the system and the phenomena you are interested in and then identify the method most appropriate for the particular problem.

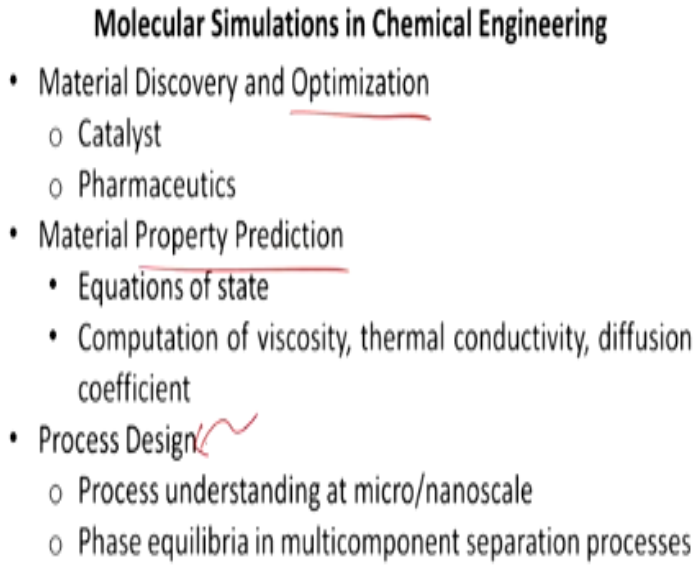
So, the major part of our effort actually lies in the initial part of this where we choose the model or build a model and the simulation method and do a feasibility check because there are software's to do the simulation for us but we have to provide the model and the physics that we want to simulate and that is the, main art of molecular simulations.

So, just to look at some of the possible areas of applications of molecular simulations within the chemical engineering many of them are already in use and many of them are coming up. So, one of the major area is the area of material discovery and optimization. Let us say I want to develop a new catalyst where should I begin?

So, I can perform a molecular simulation and screen among the possible candidates that is a much promising strategy than compared to performing some trial error experiments or some kind of intuitive way of designing catalytic systems. If we want to develop a new drug I think

molecular simulations is the way to go because we can again check many, many candidates without having to do experiments on that and at least get some preliminary understanding of the behaviour the simulations nowadays are pretty fast but still not fast enough to simulate all possible behaviour of the systems under consideration and many a times there are limitations of the methods and the results may not be very accurate and, therefore we still need to perform experiments but even if we can reduce the number of experiments we can pretty much speed up the discovery process of the materials.

(Refer Slide Time: 23:35)



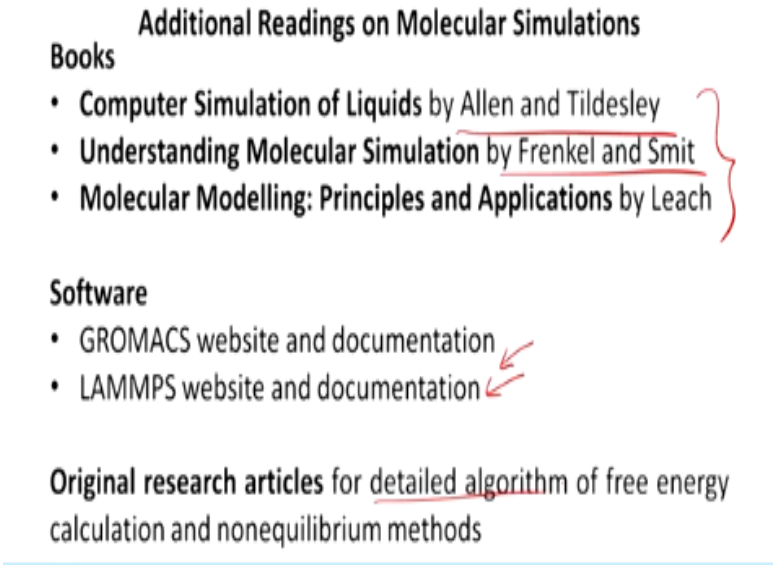
We can also use it for optimization if I know certain combination of materials are providing a good catalyst which combination is the most optimum which one will give you the highest yield. So, in that case, we can play with multiple compositions molecular simulations and figure out which one is the best possible catalyst or best possible drug formed by multiple compounds in question.

It can also be used and it is already being used in the prediction of properties of material such as the equation of a states, the computation of the viscosity, thermal conductivity and diffusion coefficient again that requires separate experiments otherwise much of that we can do I would say with pretty good accuracy using molecular simulations.

Finally, it can be important in the process design if you want to develop understanding the processes at micro and nano scale that is particularly important in the area of nanotechnology and biotechnology or the design of miniature devices in that case, the simulation can provide us the understanding of the process much beyond what can be known from continuum level

models more importantly, they can be used in the analysis of phase equilibrium multi component separation processes because we have a whole host of methods that can represent with good confidence the phase equilibrium behaviour and also compute the chemical potentials and so on.

(Refer Slide Time: 25:17)



So, with that, we are pretty much done with this particular course these are some additional readings on molecular simulations. The references are also there in the course template. These three books I can pretty much swear by they are can start reading this book from cover to cover from any additional reading. Although, it will not be required for the exams of this course, but if you want to learn more about the methods, these are the books to go with. I will particularly recommend the first two books from which I have borrowed a substantial part of the molecular simulations. If you want to start learning about an MD software, you can pretty much go to GROMACS and LAMMP Stock documentations and website and they are both open source software so you can download and at least for preliminary simulations you can only laptop but again, you will require more hardware for detailed simulations of larger systems.

And for some of the algorithms particularly in the last two weeks of the course such as free energy calculations and non-equilibrium methods since these methods are relatively recent they are not present in the books, we may have to go to the original articles or the book may not provide you detailed algorithm. So, you may have to go to the original articles which have been referred whenever I discussed that in this course, but they also present in on web and you can try to find it.

So, with that, we conclude this course on advanced thermodynamics and molecular simulations. I hope to have provided you a molecular level understanding of the principles of thermodynamics and I hope to have given you some understanding of how can we use statistical mechanics to represent a thermodynamic system and then how can we use those concepts to build the foundations of molecular simulation methods and how can we then numerically implement those methods and get the behaviour that we want to get.

And, finally I hope to have at least covered briefly some of the advanced topics like free energy calculations and non-equilibrium methods and mesoscale methods and so on. Although you definitely need to learn much more on that because I have been very, very brief on many of these things towards the end.

So, with that, I conclude this course on advanced thermodynamics and molecular simulations, thank you.