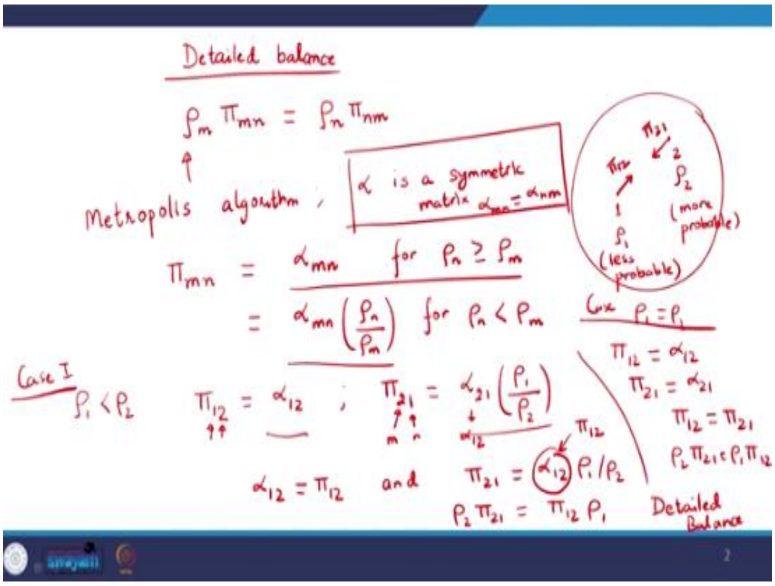
Advanced Thermodynamics and Molecular Simulations
Prof. Prateek Kumar Jha
Department of Chemical Engineering
Indian Institute of Technology, Roorkee

Lecture - 35
Metropolis Algorithm, Periodic Boundary Condition

Hello all of you, so, in the last class we have been discussing the Monte Carlo scheme in somewhat more detail. We discussed the detailed balance condition and we defined the criteria that the transition probability must follow for the detailed balance to be valid.

(Refer Slide Time: 00:57)



Detailed balance

$\rho_m \pi_{mn} = \rho_n \pi_{nm}$

Metropolis algorithm ; $\boxed{\alpha \text{ is a symmetric matrix } \alpha_{mn} = \alpha_{nm}}$

$\pi_{mn} = \alpha_{mn}$ for $\rho_n \geq \rho_m$

$= \alpha_{mn}\left(\frac{\rho_n}{\rho_m}\right)$ for $\rho_n < \rho_m$

Case I $\rho_1 < \rho_2$

$\pi_{12} = \alpha_{12}$; $\pi_{21} = \alpha_{21}\left(\frac{\rho_1}{\rho_2}\right)$

$\alpha_{12} = \pi_{12}$ and $\pi_{21} = \boxed{\alpha_{12}}\, \rho_1/\rho_2$

$\rho_2 \pi_{21} = \pi_{12} \rho_1$

Case $\rho_1 = \rho_1$

$\pi_{12} = \alpha_{12}$

$\pi_{21} = \alpha_{21}$

$\pi_{12} = \pi_{21}$

$\rho_2 \pi_{21} = \rho_1 \pi_{12}$

Detailed Balance

So, now I will take it further and the goal is to get a workable Monte Carlo is scheme and in the next week, we will see how can we numerically implement that so again, I want to start with the detailed balance condition that is, the number of states going from m to n, that is given by the probability density of states m multiplied with transition probability-

$$\rho_m \pi_{mn} = \rho_n \pi_{nm}$$

So, let us choose a π_{mn} value and see whether it satisfies this. So, clearly as I was telling you there are multiple ways in which we can pick this as long as this condition is valid we can rest assured that the Monte Carlo will kind of work whether it is going to be efficient or not that will depend on the choice of π but this condition should be met. This is not met at least a weaker condition that we discussed should be met and so on.

So, let us start with one of the original prescriptions by the scientist named Metropolis, it is called Metropolis algorithm and it goes like this- so π_{mn} is assumed to be-

$$\pi_{mn} = \alpha_{mn} \quad for\ \rho_n \geq \rho_m$$

And,

$$\pi_{mn} = \alpha_{mn}\left(\frac{\rho_n}{\rho_m}\right) for\ \rho_n < \rho_m$$

Let us first check whether this satisfies the detailed balance or not. So, let us say we have two states 1 and 2 and they have some probability density ρ_1 and ρ_2 and π_{12} is transient probability of going from 1 to 2, π_{21} is for the opposite and we are interested in checking whether this particular choice of the algorithm satisfies the detailed balance or not. By the way, Metropolis made one assumption here and you will see how this is useful, he assumed that α is a symmetric matrix that says $\alpha_{mn} = \alpha_{nm}$.

So, let us do one case where for example ρ_1 is less than ρ_2. Now in this particular case what I find is π_{12}. So, n is the n is 2 here and m is 1 here and clearly ρ_2 is higher than ρ_1. So, what it is meaning physically is this state is more probable and this state is less probable. So, therefore I will use the first relation because ρ_n here is greater than ρ_m, so this will be equal to-

$$\pi_{12} = \alpha_{12}\ and\ \pi_{21} = \alpha_{21}\left(\frac{\rho_1}{\rho_2}\right)$$

Now since α_{12} is equal to α_{21} we can pretty much replace this quantity by α_{12}, and if I use now these two expressions what I have is-

$$\alpha_{12} = \pi_{12}\ and\ \pi_{21} = \alpha_{12}\left(\frac{\rho_1}{\rho_2}\right)$$

And π_{21} is α_{12} ρ_1 by ρ_2 so I can replace this α_{12} with π_{12} using the first equation and with therefore obtain-

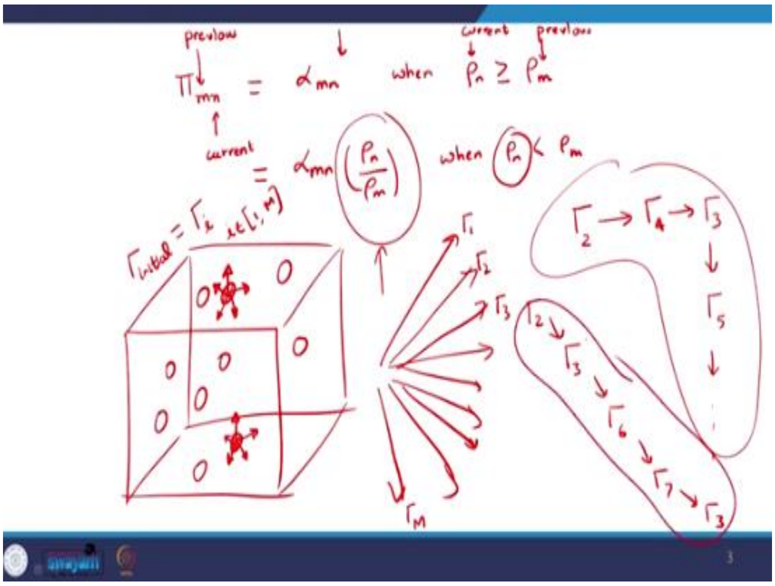
$$\rho_2\pi_{21} = \pi_{12}\rho_1$$

that is exactly the detailed balance condition for this particular case.

Now you can see even if the opposite where true where ρ_1 was greater than ρ_2, then also you will have the same equation. So, then I will use second equation for the first thing here and the first equation for second thing here and nonetheless will have the same equation. What will happen when they are equal? What will happen when they are equal is, when ρ_1 is equal to ρ_2. Then in that case π_{12} is equal to α_{12} and π_{21} is also equal to α_{21} So, basically π_{12} is equal to π_{21} and the densities are already equal, so therefore again we can get the detailed balance met.

So, by choosing α to be a symmetric matrix we are making sure that the detailed balance condition is met but we can choose still α to be anything. There is still some room in which are, we have some choice by which α can be chosen. Now let us see why we are doing that for a minute, why this particular algorithm so what exactly it is doing?

(Refer Time Slide: 08:12)


previous
current
current
previous
$\pi_{mn} = \alpha_{mn}$ when $\rho_n \geq \rho_m$
$= \alpha_{mn}\left(\frac{\rho_n}{\rho_m}\right)$ when $\rho_n < \rho_m$
$\Gamma_{initial} = \Gamma_k$ $k \in [1, M]$
Γ_1
Γ_2
Γ_3
Γ_M
$\Gamma_2 \rightarrow \Gamma_4 \rightarrow \Gamma_3 \rightarrow \Gamma_5$
$\Gamma_2 \rightarrow \Gamma_3 \rightarrow \Gamma_6 \rightarrow \Gamma_7 \rightarrow \Gamma_3$

So, if I look at my π_{mn}, m is my previous state and n is my current state. And I am keeping as α_{mn}, when the probability of the current state is greater than equal to the probability of the previous state. On the other hand, I am multiplying this quantity, ρ_n by ρ_m. When ρ_n is less than ρ_m and now I am referring to the stationary ρ distribution, of course. I will begin with something else, but ultimately we reach is stationary state. So, at this stationary is state, what do I want to do?

So, I want to sample more in the region where you have higher probability densities and we want to simple less in regions which will have lower probabilities. So, if the states have less probability to happen then why do I want to sample there more, so therefore when the probability in the current state that is like I am starting to previous state and I am making some kind of a movement from or transition from that state to the current state. So, if the probability in the current state happens to be less than the previous state so that means that I am trying to move to a state of a lower probability. I am multiplying with ρ_n and by ρ_m. And as the ratio of probability of the current versus the previous decreases, I am basically using a smaller transition probability in the limit when ρ_n is very smaller than ρ_m in that case, there will be very less probability of going from m to n.

So, over a period of time what is going to happen is by time here, I mean the computation time. What is going to happen is I will be sampling more and more in the regions of higher probability and at when we are at the stationary solution then also I am sampling in regions of higher probability more than in regions of lower probability. So, essentially this ρ_m by ρ_n basically is able to wait my transition probability in a manner that gives more preference to the transitions from a state of low probability to high probability because I did not add any factor here. And gives less preference to movements from high probability to less probability. Keep in mind that the movements or transitions may still happen.

So, it is not completely stopping that to happen, so I can go from a state of high probability to low probability because that happens to be important for the conditional detailed balance but the transition probability decreases for those kind of movements as opposed to the movements from states of low probability to states of higher probability, that is I would say a very important point.

Now, let us see how exactly we will do that in practice. So, all of this so far is kind of very theoretical. Let us say I want to apply this metropolis algorithm in practice, how will I start?

So, just like molecular dynamic simulation, I can still imagine that there would be a box and there would be particles in the box. The setting up of the system when I begin may look very similar to the molecular dynamic setup. So, as soon as I do that clearly I am at a particular state gamma, let me call it some Γ_{initial}.

Now however, unlike molecular dynamics I will not solve the Newton's laws of motion. What I say this system starting from here can go to many many possible states. So, let us say I have numbered the states Γ_1, Γ_2, Γ_3 to some Γ_m. This Γ_{initial} could be any Γ_i where i is in the range 1 to m, so any one of them could have been picked whenever I am doing the simulation setup. How exactly we will do that, will come back to that point. One of the possibilities is we randomly generate particles in a cube, we will revisit that point, but keep in mind that whatever my initial gamma is that is chosen among the state space.

Now I will make some movement and that is a artificial movement nothing to do with the actual dynamics system that leads to a transition from this initial state to some new state. Let us say for example, one of the possibility is I started from some Γ_2 from here I go to Γ_4. And since it is a Markov process now, when I am at Γ_4 I do not care whether I come from Γ_2 or $\Gamma_{3.}$ The memory, there is no memory of the previous step. So, now from Γ_4 again I can go to any of the possibilities let us say if I go to Γ_3 and from there let us say if I go to Γ_5 and so on and I will continue this kind of a process.

One of the possibilities by which this can be done is that, I pick one of the particles and I move it in a random direction and we will take particular cases in more detail but this is one of the schemes. So, at any given step I pick one of the particles and move it a bit, as soon as I am doing that I am disturbing the phase point because phase point was p q and now I move to a new coordinate by the way momentum does not play any role in Monte Carlo because there is no dynamics here, so as to speak so coordinates are the only thing of concern.

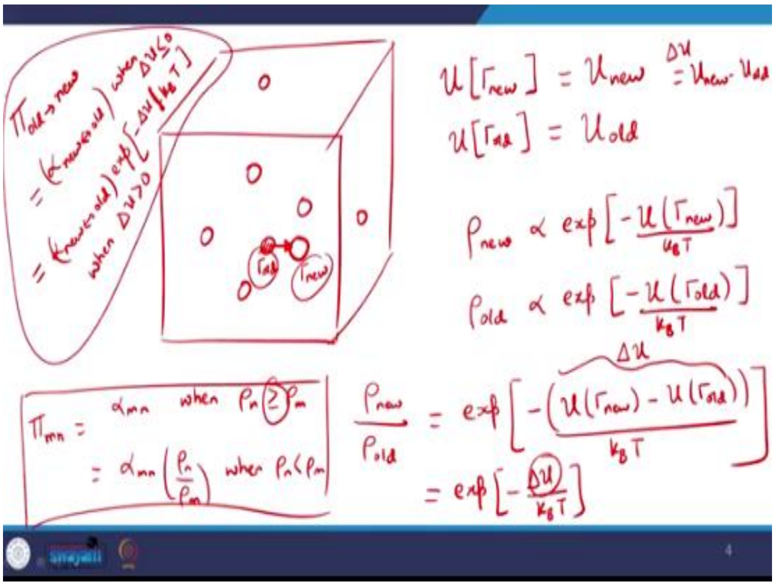
So, once I have moved from 1 point move 1 point from here to there I have moved in the, my state from something to something else. Now clearly that is state has to be among one of the state space. So, whatever movement I am doing should consider how I have constructed my state space I should be making that happen there. But let say if it is going to valid state in my state space then when I go to the next step, I can again pick some other particle and I can move it randomly in any one of the direction I can pick how much to move, and I can pick the direction as long as I am going to a new state and as well I can remain wherever I am.

So, then, how it is a Monte Carlo scheme? How we have lost memory? The way we have lost memory is in any given step, I am picking a random particle from the box and I am moving it in a random direction. In the next step, I do not care what happened previously I again pick a new particle randomly and again moving random direction. So, it does not really matter what happened previously whenever I am making the new step there is that is what I say, it is a Markov process.

Now clearly if this is what I am doing then let us say if I repeat the simulation let us say if I re-run my code. Again, what will happen is I will pick a new particle randomly and I will move it. Now this time since it is a random generation, random picking of particle and a random movement, the new movement is not going to be the same as what happened in the previous Monte Carlo simulation. So, if I repeat this I may follow some other pathway let us say this.

So, every time I am running my Monte Carlo simulation, I am essentially sampling different reasons of the phase space and that is where Monte Carlo is quite different from molecular dynamics in molecular dynamics, it is fully deterministic because we are solving Newton's laws of motion. As soon as we pick the initial state, we are doing the true dynamics so we know where we will go in a Monte Carlo simulation since we are doing a random generation or random picking and random movement every time I am doing it I am essentially capturing different regions in the phase space and since there is no memory of the previous step. I can pretty much go in a very zig-zag manner in the phase space as opposed to a more deterministic movement in the case of a molecular dynamic simulation.

(Refer Time Slide: 19:04)



$$\Pi_{old\to new} = (\alpha_{new\to old}) \text{ when } \Delta U \le 0$$
$$= (\alpha_{new\to old}) \exp[-\Delta U / k_B T] \text{ when } \Delta U > 0$$

$$U[\Gamma_{new}] = U_{new} \quad \overset{\Delta U}{=} U_{new} - U_{old}$$
$$U[\Gamma_{old}] = U_{old}$$

$$\rho_{new} \propto \exp\left[-\frac{U(\Gamma_{new})}{k_B T}\right]$$
$$\rho_{old} \propto \exp\left[-\frac{U(\Gamma_{old})}{k_B T}\right]$$

$$\Pi_{mn} = \alpha_{mn} \text{ when } \rho_n \boxed{\ge} \rho_m$$
$$= \alpha_{mn}\left(\frac{\rho_n}{\rho_m}\right) \text{ when } \rho_n < \rho_m$$

$$\frac{\rho_{new}}{\rho_{old}} = \exp\left[-\frac{\overbrace{(U(\Gamma_{new}) - U(\Gamma_{old}))}^{\Delta U}}{k_B T}\right]$$
$$= \exp\left[-\frac{\boxed{\Delta U}}{k_B T}\right]$$

So, let us say if this is the Monte Carlo scheme that I am doing in that case, let us see what the densities should look like. So, let us say I start with this system and I have moved this particle from here to there. So, there are two states a new state where this particular particle has moved and an old state, where this particular particle is here, by the way, the state is representation of coordinates of all the particles in the system. So, even though I am labeling this i as gamma old and gamma new this is referring to all the coordinates of all the particles in the system in the new state. Although I am moving only one, everything else has changed in the system because all the forces that involve this particular particle will change the separations have changed. So, it is not that the motion of every particle is somewhat individual. The motion of every particle even if it is artificial movement as in Monte Carlo it is affecting the entire system behaviour.

Now, let us say the energy as-

$$U[\Gamma_{new}] = U_{new}$$

And,

$$U[\Gamma_{old}] = U_{old}$$

So, now what do we expect when we are in the stationary state? What we expect is the probability in the new state should be proportional to the Boltzmann factor that is-

$$\rho_{new} \propto \exp[-U[(\Gamma_{new})]$$

$$\rho_{old} \propto \exp[-U(\Gamma_{old})]$$

In fact if I want to find the actual value, I have to know the partition function. But if I am only concerned about the ratio of this density, then that denominator will cancel. So-

$$\frac{\rho_{new}}{\rho_{old}} = \exp\left[\frac{-\left(U(\Gamma_{new}) - U(\Gamma_{old})\right)}{k_B T}\right]$$

you can represent this quantity as some Δu, that is the energy change from the old state to the new state. This going to be—

$$\frac{\rho_{new}}{\rho_{old}} = \exp\left[-\frac{\Delta U}{k_B T}\right]$$

Now let us go back to the metropolis algorithm, I have said my transition probability is-

$$\pi_{mn} = \alpha_{mn}\ when\ \rho_n \geq \rho_m$$

And,

$$\pi_{mn} = \alpha_{mn}\left(\frac{\rho_n}{\rho_m}\right)\ when\ \rho_n < \rho_m$$

So, what is going to be the transition probability therefore of going from an old state to a new state. So, if the new state happens to have ρn greater than ρm then in that case we will have the probability as α_{mn} and when will that happen that will happen when Δu is less than 0 because when Δu is less than 0 then this exponential quantity is going to be larger than 1. That means the ρ_{new} is more than ρ_{old}.

So, this quantity will be α and now it is a symmetric matrix, so it is α new to old or old to new whatever want to write-

$$\pi_{old\rightarrow new} = (\alpha_{new\leftrightarrow old})\ when\ \Delta U \leq 0$$

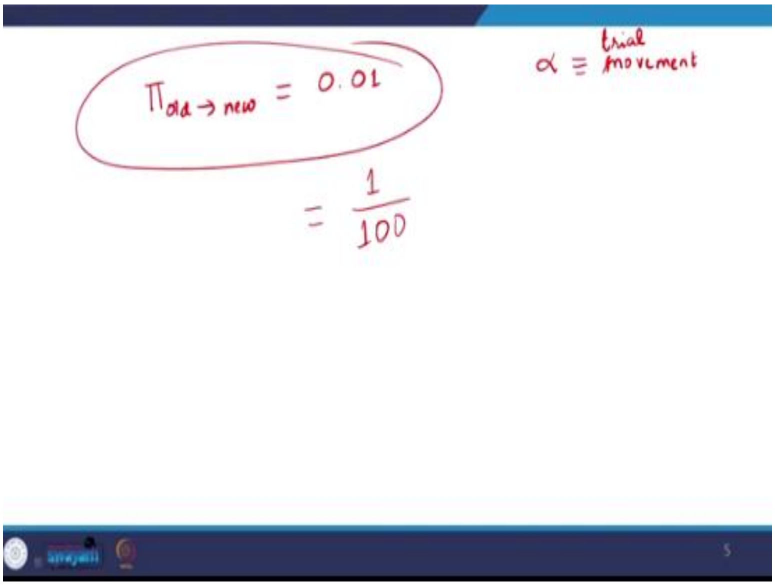
And,

$$\pi_{old\rightarrow new} = (\alpha_{new\leftrightarrow old})\exp - \frac{\Delta U}{[k_B T]}\ when\ \Delta U > 0$$

So, now how will I therefore implement that, what is that α that is there, now α in a sense that we have already chosen. So, when I was making the movement I was making with some mechanism. I was picking a random direction on a computer there can be only certain number of direction. I mean, it is going to be continuous but on computer it has to be discrete. So, let say there are one million possible ways in which I can move a particle in any of those directions so every time I am making a random move the probability of that to happen is one by that million, so α is determined by how I am doing the Monte Carlo movement. Everything else is well determined because I can find the energy changes once I know the coordinates and coordinates are known because I start with an initial coordinates and I know how I am making a movement, so the coordinates are known α is ensured the only thing we have to make sure in choosing α is the reverse should be possible that is when it is symmetric, so let us say for example if there is some probability of a particle going to the right the probability of it going to the left should be so be equal. As long as there is no directionality or some bias in the way I am choosing to move a particle, there is no problem.

Now the transition has to be happening with the probability that we have listed. Now how will we make sure let us say if π is 0.01. How will we make sure that, I will make that transition with a probability of point 0.01?

(Refer Time Slide: 27:00)



$\Pi_{old \to new} = 0.01$

$= \frac{1}{100}$

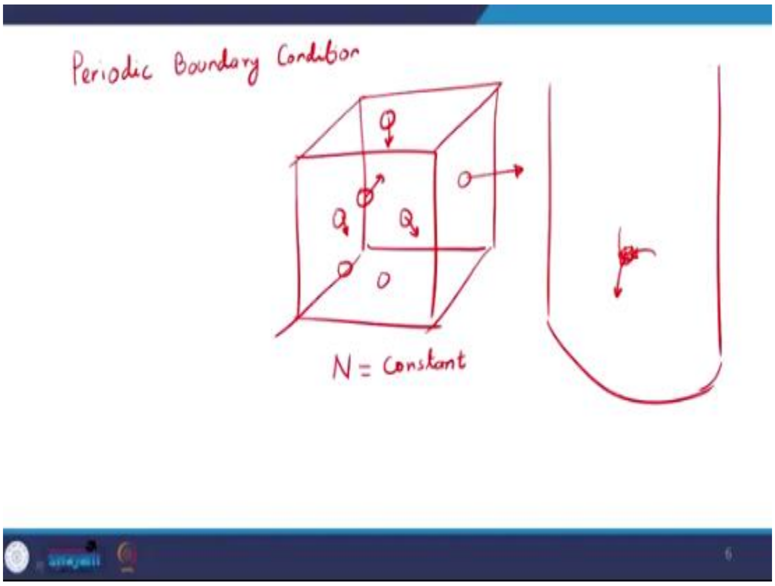
$\alpha =$ trial movment

So, the answer to this lies in here so let us say for example. If $\pi_{\text{old to new}}$ come out to be 0.01, so α refers to the movement and I am saying that that has to be symmetric. So, I should be able to move for every particle to the right or left top or bottom there should not be known in in-directionality to this. So, then how will I ensure the transition probability is 0.01, that means I should be accepting only 1 out of 100 moves that I make and the answer lies in the word I already use I only accept those number of moves which are like 1 by 100 in number.

So, let us say for example, if I am making 100 movements I call them a trial movement and a trial is accepted with a probability of π. What I do then is I do not move the particle, I try to move the particle. I just imagine that the particle has moved from here to there. Now what I compute, I compute the energy change that would result if that movement is made. Now on the basis of the energy change, we compute π that tells me the probability by which this movement should be made. Now with that probability I accept the move. If the probability is high I accept more, if the probability is less I accept less.

So, that acceptance and rejection is something that is novel to Monte Carlo that does not happen in molecular dynamics. In molecular dynamics particles move they move i mean, there is nothing like I want to move there and then I can judge whether I should go there or not. In a Monte Carlo simulation, we always make a trial movement. I see whether I want to get there or not and I make the jump from here to there only with a certain probability and this is what makes the Monte Carlo simulation special, because now as I have been telling you we have a control on this π_{mn} in terms of whatever saying I can have a lower probability of going from a high probability state to low probability state so I have got in full control, so I can accept the movements from highly probable regions in phase space to less probable reasoning in phase space with a lower transition probability on the other hand the movements from a lower probable reason in phase space to higher probable reason will be accepted with a higher probability or transition probability and it is therefore this we will be able to sample the phase space more efficiently.

(Refer Time Slide: 30:36)



So, before I conclude there is one detail that is extremely important in both Monte Carlo and molecular dynamics and that is whenever I say my system is like this that it contains some molecules, I tend to draw some kind of a box, imagine that you have some carton containing balls just make it scale down to very very small length scales and the balls you make it like molecules and they will assemble the system that I am talking about.

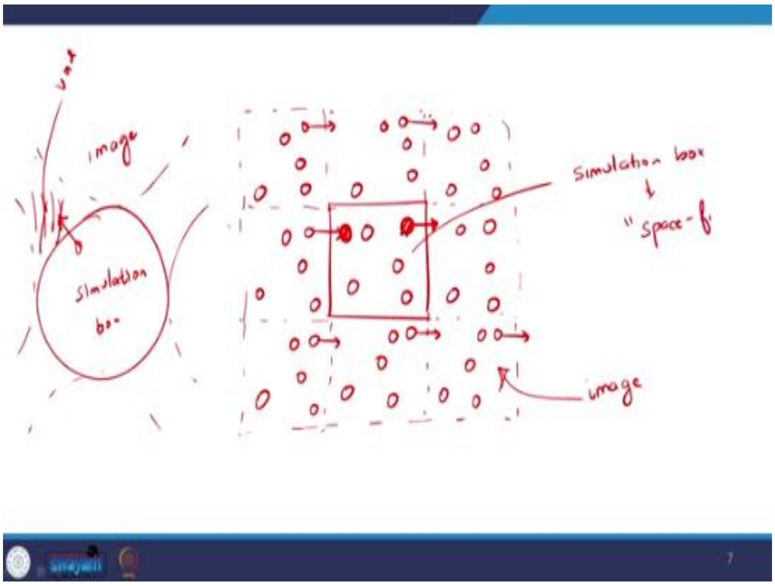
Now, clearly in molecular dynamics in the molecules are indeed moving. In Monte Carlo, molecules are still given or particles are still given an artificial movement and therefore, it is quite likely that they will hit the walls of whatever container I am imagining. Now what exactly that means, does it mean that we really have a cubic container containing the molecules? Probably not, because yes in principle you can have a cubic container containing the molecules which you are simulating and in that case that picture is very valid that the molecules hit the wall and get rebound. But in more general situation since I have been telling that the number of particles or molecules you can simulate is pretty small at least not comparable to any macroscopic system you are imagining in most cases, you are doing a tiny bit of a large system. Let us say for example, if I think of a beaker containing a liquid I am simulating a small very very tiny volume within that beaker what that means is that there is no walls really because you can have molecules coming in here molecules leaving out from there and so on.

In fact, even though the pictures are just that there is a wall there is in principle no wall first of all because there is no physical wall out there, because I am considering a small cross section of the system and secondly because that system not an isolation. So, that system exchanges molecules with the other nearby systems even though I am choosing to simulate one small part of my large volume of any liquid or solid or whatever, that is not isolated in terms of the exchange of the number of molecules or the energy even when we discuss the ensemble and all that, all we say that we have certain control variables that control can be applied for a macroscopic volume, but when we look at that particular cubic box or whatever we are not trying to suggest that this works exists in isolation clearly that box is, I would say a small representation of a large microscopic system.

So, it turns out that better boundary conditions is what is known as Periodic boundary condition, is a much better bit in these kind of situations and what this says is that the number of molecules over all in the box remains constant, yes molecules can come in, molecules may go out but overall we can assume that N is constant. So, as long as we have started with system containing large enough number of molecules, the fluctuation in N is any way going to be small. So, therefore keeping N as constant is not really taking away any much realism from the picture that you have of course in the grant canonical ensemble we will not have and constant and we will make the same assumption there but for the canonical ensemble we are keeping N constant and so we will do in a simulation but what we say is that, let us say if some molecule want to go out then some other molecule has to come in to fill the void that is created by that molecule.

So, let us say this is my box and some molecule is crossing one of the one of the walls because walls technically there is no one physically there is no wall. So, if one molecule leave something else comes from the opposite end, the other wall opposite wall and enters the system so every molecule that is lost or every particle that is lost is compensated by another particle that comes from the opposite end of the system and that is the essence of the periodic boundary condition.

(Refer Time Slide: 35:55)



So, what this picture essentially mean if you want to think about it is something like this that is seeing that in fact if I look at the entire system, let us say a beaker or whatever I am trying to simulate. I am simulating only a small part of that system that is my simulation box right, that can be a cubic volume and I am now doing in just 2d for the sake of illustration and this contains some configuration of molecule.

So, what the periodic boundary condition actually employs is the assumption that the system will be containing many replica of the box that I am simulating and that fills the entire volume. So, I can imagine that the same configuration of the molecule is repeated throughout the volume of the beaker, because whatever I am simulating is a representation of molecules in the beaker. I may imagine that whatever cubic simulation box I am doing that is like being repeated throughout the beaker or whatever system I am trying to say that means, every system in there has the same configuration of molecules so as to speak, it has pretty much the same configuration just being repeated and all these guys which I am essentially not simulating, so these are basically referred as images of the simulation box.

So, this is purely a hypothetical picture for the purpose of simulation I am only simulating the simulation box in the centre but one can imagine that is the same configuration is being repeated throughout the entire volume and that would be a representation of the bulk system that I am trying to simulate and it is then we can say that, let's say for example if one of the particles leaves out then since the configuration of the images is same as the simulation box, what means is that all these images also undergo the same movement and since the images also undergo the same movement as soon as this guy leaves another guy comes in and fills the void from the other end that is just an implication of the assumption of the images of the system.

Now if you think about it, this is not really that is something that is highly realistic because in some sense what I am assuming is some sort of a crystalline behaviour because I am saying that the configurations are being repeated only difference from real crystals is that, that crystallinity is applied over units of the simulation box that contains huge number of molecules. So, a crystal will have repeating units of the order of one or two molecules, but in this case I am assuming that there are millions or billions of particles in simulation box that characterize its configuration and that entire thing is being repeated. So, yes it is introducing some kind of crystalline behaviour but it is much better than assuming that there is a physical valve, where particles hit and get rebounded because there is no wall in reality.

So, that is one of the points here the second point that also comes from the periodic boundary condition is this sets a limitation on the shape of the box because ultimately I am trying to represent the actual system, that is the beaker I was giving the example of by using the simulation box and replicas of that or images of that, so I am basically trying to fill the beaker with repetition of the simulation box that will be valid if the shape of the simulation box is such that it is space filling, that means I can stock when like hypothetically speaking I can stock many cubes or cubical cartons and I can fill the entire space. If that is possible then we say that shape of simulation boxes space filling, so cube naturally is therefore space filling. On the other hand let us say if I would have taken my simulation box to be spherical in that case, you may imagine that let say this is my simulation box. Now this thing will repeat so you will have images on all sides but you will also have regions of like some kind of a void that has no physical meaning. What if a molecule happens to go here? What will be with the model those molecules it is neither the simulation box nor its image so the periodic boundary condition will not work in that situation.

So, therefore that assumption of the periodic boundary condition requires that my simulation box is a space filling and cubic is clearly a great choice, but it does not have to be cube. It can be many other kinds of shape, but clearly not a spherical shape because the spherical shape will not fail the entire volume. So, periodic boundary condition is employed. I would say in most of the simulations however, if I want to simulate things like interfaces, we are clearly the behaviour is not of a repetition there is a discontinuity in the behaviour as soon as an interface comes in, so in those situation the idea of the periodic boundary condition will not work. It may still work in the other directions apart from the interface direction, but it will not work in the along the interface across the interface that is one of the limitations and secondly in many cases, the use of periodic boundary condition is not even needed because I am more interested in the behaviour around a molecule.

Let say for example, I am simulating a protein and I am interested in how the protein interacts with the solvent around it. So, I am not interested in something like protein in bulk and I am trying to simulate one small representation of that thing instead my interest is in knowing how protein interacts with the solvent molecules.

In those cases I do not need to make this sort of representation because there is no macroscopic system I am trying to simulate. I want to get a behaviour that can be gotten with much simpler boundary conditions actually in that case more realistic boundary condition, and it turns out that in that case we also do not need a space filling box you can just take a protein molecule and some solvent molecules around it to solve it the protein molecule and so on that also works pretty well.

And in special situations, we also know some other types of boundary conditions, for example a reflexive boundary condition that is pretty much a rebound kind of an arrangement. We can also add the effect of actual physical walls if they are present by applying special wall boundary conditions and so on we will discuss them briefly later in the course. So, with that I want to conclude the discussion today.

In the next lecture, we will see how we actually numerically implement a Monte Carlo simulation. And after that we will be doing some case studies around this idea.

So, I am going to conclude with this, thank you.