**Mathematical modelling and Simulation of Chemical Engineering Process**

**Professor Doctor Sourav Mondal**
**Department of Chemical Engineering**
**Indian Institute of Technology, Kharagpur**

**Lecture 56**
**Kinetic monte carlo simulation**

Hello everyone, in this week we are going to learn about some modeling methods or simulation methods I would say, which is involved in the cases or something which we use when there is no concrete background on the physics of the system. So, for example, we do not know much about the physical system but we want to model the dynamic event. So, these class of problems generally belong to the stochastic process. If you recall in one of the first week we talked about when we had these introductory lectures, we talked about the difference between stochastic and deterministic process.

So, to begin with, today, we are going to learn about this kinetic monte carlo simulation. So, this Monte Carlo is a method actually, which is based on the random number. So, you choose different processes in your system randomly. And so it is like a random sampling of your state space or like your system, and then you try to carry out certain events, again randomly, and whether these occur events are acceptable or not, I mean, I am not going to talk about details right away, but just to give you a glimpse that whether these events are possible or not is again decided by some random numbers, and then finally, you do this sampling over a lot of state space, both in space as well as in time and this is… the final result that you get in this case is average of all these sampling.

So, essentially it is a class of problem which are solved with the help of random numbers. So, random numbers play actually a very big role in this kinetic Monte Carlo simulation. So, what essentially is this you know, this Monte Carlo is that let us say you have a chemical reaction and in the chemical reaction, there are possibilities of several intermediate states or intermediate products or compounds which can form.

Now, from the classical analysis you can only find out whether certain intermediate products are feasible or not based on the calculation of the free energy and all those thermodynamic things, but in this Monte Carlo method, you can apply certain random energies to the different states and you can obtain a series of intermediate products and then you can evaluate their feasibility.

So, in this way it is possible to get any unrealistic states also. So, it is also possible, because you give these energies which may not be thermodynamically possible and then you allow the system to jump to different states and finally, you do averaging of all the… based on the again the probabilities of certain states which are forming or which is possible to form, you assign some weights based on those possibilities and then finally, you try to do a averaging and find out that what are the different fraction of intermediates that are possible or to what extent they are possible and what is the most desirable or the most feasible product in this reaction.

Another example could be like this, dispersed phase systems like when you have let us say a breakage event, a breakage event taking place, sort of, it is a system which is where certain particles are breaking and slowly the population is growing. So, let us say we do not have a very well-defined model about how the system population is increasing with time. Of course, we can have the population balance equation, but these required very detailed understanding of let us say the growth process or the breakage functions and a lot of considerations or mathematical implications are needed.

The other possibility is to use certain breakage event let us say for this case or these particle level events to follow randomly. And then at the end, you try to see that you allow these process to have first generation, second generation breakage and subsequent breakage steps, you allow this to grow and reach a particular time. So, these you do it for different types of random number or different kind of random numbers, and then you do a finally you do an averaging of your result. So, that result or that averaging will be representative of what ideally you would expect, if you had done actual calculation of the process with respect to time.

So, in this regard, I would like to also highlight something known as the ergodic hypothesis. So, Ergodic hypothesis tells you that something which is average over time, a small quantity or a small process or a small set of particles or their behavior, their dynamics, which is actually happening, if they are averaged over time is equivalent to spatial average of a lot of particles at a certain time. So, this is what Ergodic has proposed and it is generally valid for micro or molecular level events.

And the second part of this ergodic hypothesis, which I said that something which is sampled or average over a large spatial domain is actually the presence what is something we are trying to do in this kinetic monte carlo process. So, this process, I mean this simulation

method or this technique can be used for problems such as adsorption, diffusion, material growth, crack preparation, where there is no very well-defined physics about the happening or the progress of the process. For example, crack formation.

I mean, we know when crack forms, but how does the crack propagate? How does it initiate or propagate is something, its bit of randomness is there in that. We all have to accept that from all these solid dynamics and all, we can study that when the crack will form, but even small amounts of instability in the process or small amount of perturbation the process can change the crack growth or propagation drastically. Same is true for any instability-based process. When you have instability in the system, it is very hard to define that, how the instability will grow.

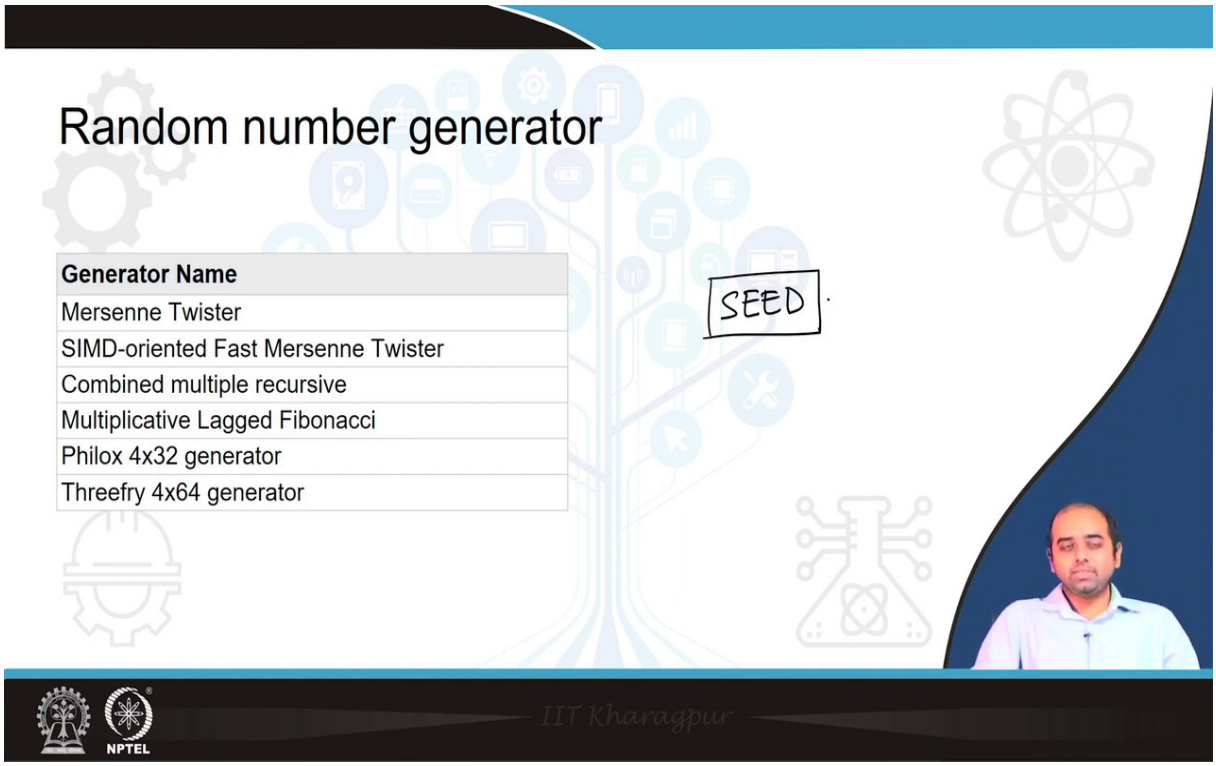
We all know when the instability will happen, when it will exceed certain criteria and all. So, where and how it will grow is something not very well defined or how the instability will progress is not reproducible. So, the sudden instability is progressing in this way, but in the second, another case, you are seeing that it is progressing in a different way. So, then their features or their progressive characteristics are not seen. So, there is a certain level of randomness still in the processor, sometimes we call the process to be a little bit chaotic also.

So, all these process even dendritic growth processes, bacterial growth of process, something which is actually influenced by a lot of factors or phenomena and even small things matter too much, then the final artifact, I mean of the process or the resultant process that we see appears to be to sort of, a little bit random.

So, all of these events or all such processes can be essentially handled by this sort of random technique known as the Kinetic Monte Carlo. So, essentially Monte Carlo is, this random sampling and doing the averaging and the kinetic part actually helps us to have sort of a progress with respect to time.

But please note that this dynamic nature obtained from the Kinetic Monte Carlo simulation may not represent the true dynamic behavior of the process because this is something which you again try to do or try to calculate whatever this time step or the growth rate with respect to time is based on the random numbers again, but this actually helps you to at least track the evolution of the process in some way.

(Refer Slide Time: 09:50)


So, now, let us try to look into what do we actually mean by these random numbers. So, this is something which I always say, this random number by nature is a random quantity, it is random, it cannot, I mean you cannot generate a random number using any formula, then how it is possible that a random number is actually produced by a computer? Because we all know

that the computer only works on certain algorithms or some formula or some relation. So, if you apply a formula to create a number, let us say create this sort of random number, then in the true sense, it is not random, is not it?

So, we all see that any this programming software has options to generate random numbers, but please note that these random numbers are actually used by, are actually developed or produced by a generator or algorithm. So, in the in the literal sense, these are not true random numbers.

So, we generally call random number which is produced by a computer is actually pseudo random numbers. But anyway, these acts I mean, these are quite good enough to be classified or to be categorized as random number for any reasonable calculations, but please keep in mind that a random number is something which cannot be generated by a formula or a correlation or we cannot have a relation between two random numbers in any broadest terms.

So, it is naturally, if something is generated by a formula, there is a relation between the two numbers that is generated and then in the proper spirit or in the actual spirit, this cannot be termed as actually a random number. So, we call them to be pseudo random numbers. So, here in the screen that you see, I have listed down several of these random numbers I mean generators are name of the algorithms which are used by different programming languages to get the different types or categories of random number.

One important thing I mean, we will not go into details of these algorithms of this, but you can generally google it up to see that what is the background of these algorithms which is generated, each one of them has certain advantage, some disadvantages, So, one of the important thing in the random number is essentially the seed. So, a seed is actually this sort of the initialization of the random number and based on that initial value the random number can produce I mean, this random number generator can essentially produce a random number as an output.

So, if you change the seed, it is quite likely that the same random number will not be generated. And generally, the seed to a random number is given by default as the current time. So, generally when you try to run the same code or same random number generator, I mean, second time twice, there will be some difference in the time, because we cannot operate at the same time. So, the final random number, the output that it comes appears to be

different, but ideally that is something one would always see, if something is random, then we always get this feel that if I execute a certain code, I am getting different numbers.

So, I feel like this is a random number because, if it is based on some formula, then whatever I do every time I should be getting the same answer or output. But the trick is that these seeds or this initialization parameter that is set for a random number generator is by default, I mean, that is how this is devised in most random number generators or in the software or the programming tools that we are using to be used as the current time. Now, when you run the same code second time and based on the time, the second time will be different, the clock time will be different.

So, then the seed will be different and finally, the output becomes different. So, it appears that even if running the same code, the numbers I am getting different. So, please look into it. It is a very easy thing that you can do in any coding software that you know that if you try to fix the seed of the random number, or the initialization parameter to be used as same, then rerunning the same code will always give you the same number. So, this is a check and this also helps you to realize that essentially these are not true random numbers because if it is a true random number it is not possible that you have you always generate the same random number.

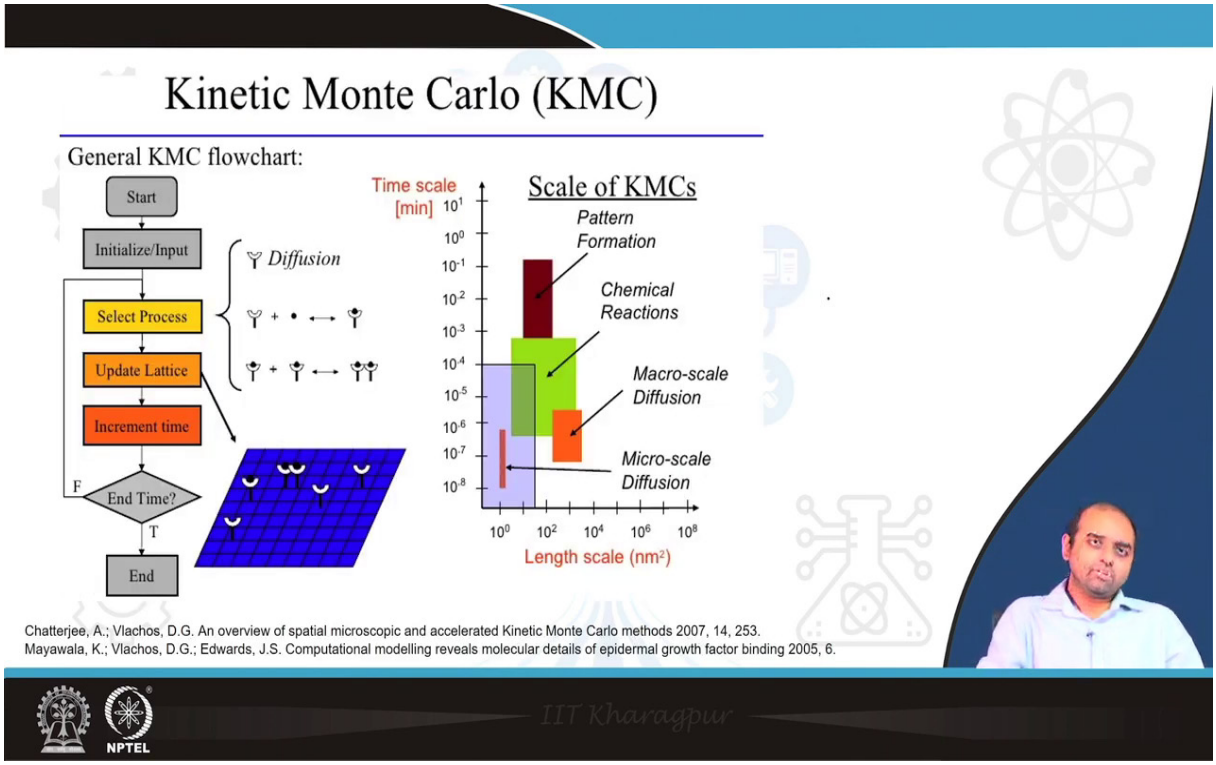
So, if you fix your seed, you will be getting the same random number even if you are running the code n number of times. So, that is something that I wanted to highlight and seed actually plays a big role. So, if you are fixing the seed, if you are getting a set of, generally we try to take a set of random numbers.

So, those set of random numbers are actually unique for that seed. So, if you change the seed, you will get a different set of random numbers. So, let us say I want to take 10 random numbers or I want to generate 10 random numbers between 1 to 100 and I have defined my seed, so these random numbers which I have got is from the same seed.

So, if I am changing the seed, the set of random numbers will different, but these random numbers will stay the same even if I rerun the code. So, it is always advisable in Kinetic Monte Carlo simulations, that whatever set of random numbers that you try to get for a particular process should be from the same seed, you should not be changing the seed to generate different sets of random number. So, these random numbers are required at different

stages in the calculation when we discuss in the algorithm, I will talk to you, but these should be actually from the same seed.

(Refer Slide Time: 16:33)




So, what essentially this, I mean, in the broader perspective, this Monte Carlo is something that, I already talked about this in the beginning of this class that in any process, whether it is diffusion, or whether it is crystal growth, whether it is absorption, in any sort of process that can be carried out for a particular process, and then you try to assign certain rates to each of these process. So, that rate, it could have a maximum and minimum to certain these transition, and then you assign some probabilities of the execution of these states, or these transitions or events and then you define a random number.

So, if this probability is more than this random number, you allow that event to happen; if not do not allow, that is it. The event can be anything, it could be diffusion, it could be segregation, it could be breakage, it could be any sort of event, it could be even molecular motion, and then there is a formula, a certain formula has been devised, then you try to update your time, and then you carry out the next event. So, like that you proceed and reach a certain end time.

So, this is how it works, it is based on random numbers, we will talk about the algorithm detail. So, this right-hand chart that you see here, tells you that, what sort of different

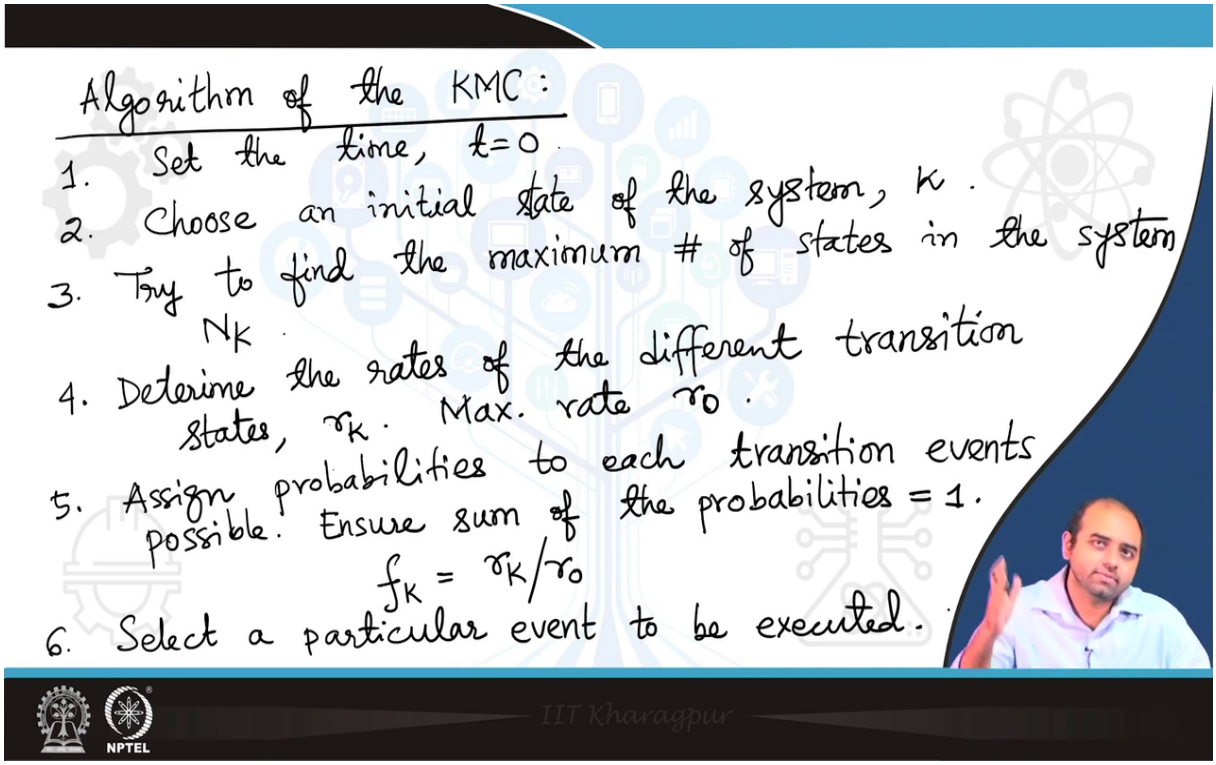
processes that we can actually look into. So, this gives you some idea on the length scale and the timescale. So, diffusion whatever, when we talk about diffusion, again happens at different length scales and timescales. So, generally, micro scale diffusion happens at a very small length scales, so it is mostly the molecular diffusion, Brownian motion and all.

So, that happens in a very small timescale and a very small landscape also. Macroscale diffusion is something that we see around, like diffusion of gases in this room. So, that happens at a larger length scale, but timescale is more or less same. Chemical reaction is something which happens at a larger time scale, but again these at a slightly higher length scale, but at least significantly higher timescale compared to this sort of micro scale diffusion.

But again, the chemical kinetics of the chemical reaction and macro scale diffusion can be compared in certain length scale and that is something very interesting. Pattern formation is generally very slow, it does not happen at a very low timescale, it generally happens at a larger timescale.

So, these actually, this picture is a sort of schematic that helps you to understand that what is the timescale of the kinetic Monte Carlo process that we are actually looking into. And this helps us to determine what are the possible rates or the kinetics of the individual events that we intend to carry out as a part of this simulation.

(Refer Slide Time: 19:52)




So, now, let us look into the algorithm of the Kinetic Monte Carlo. So, let me write down or list down the different steps. The first step is you set the time to 0, when you start. Then you choose an initial state of the system. Let us call this state of the system as k. Next is try to find out the maximum number of states in the system, let us mark this as N k.

Next is to assign or find determine the rates of these different transmission states, determine the rates of the different transition states. So, let us call, since the different states could be, so if you are having a chemical reaction it could be, a state could be a reactant, could be a product, could be an intermediate or whatever.

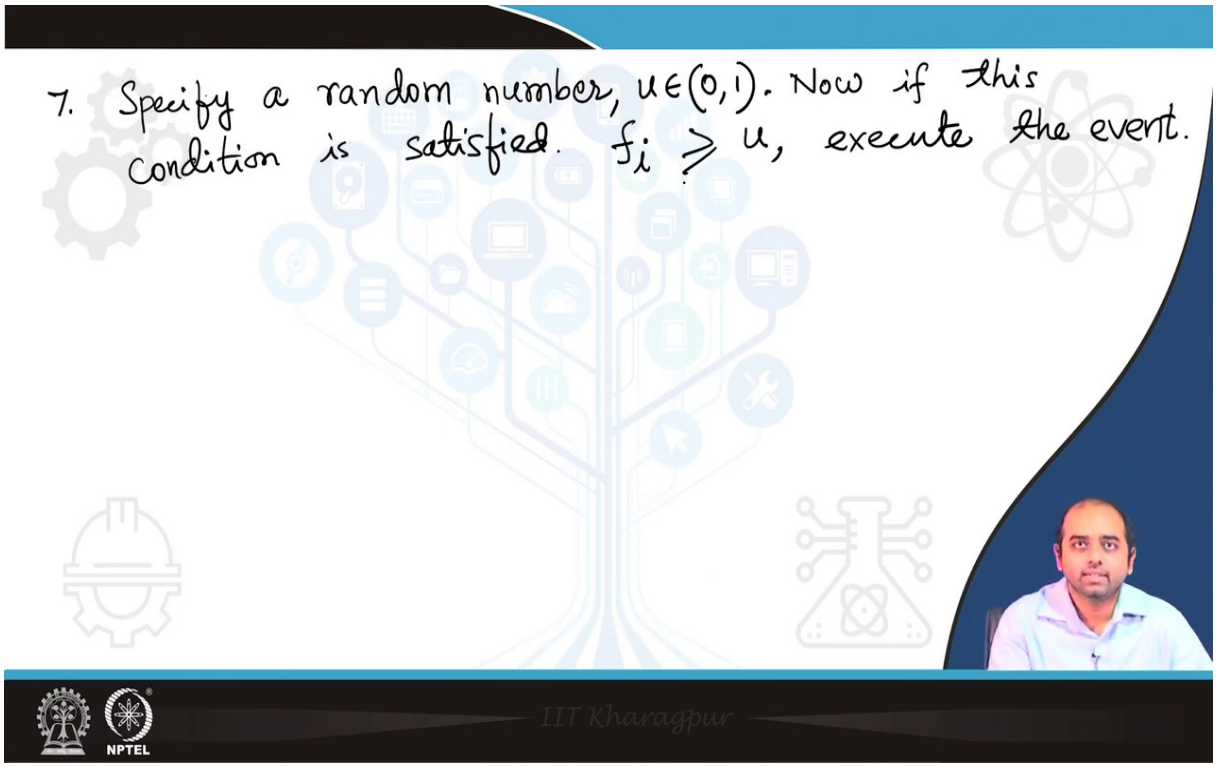
So, what is the rates of the different transition states? Let us call that as Rk and maximum rate let us call it as R0. Next important step is assign probabilities. Assign probabilities to each transition state. Let us say this transition events possible. You must ensure that some of the probabilities should be equal to 1. So, you can define the probabilities at like ratio of the rates. So, that is like the possibilities of a particular event, the degree of the level of the possibility of a particular event to happen.

So, you can write like this, f k is equal to something like R k by R naught, so summation will also give you 1. The next is to select a particular event to execute. So, how do you select a

particular event? It is again with the help of a random number. So, initially you need a set of random numbers for your different steps.

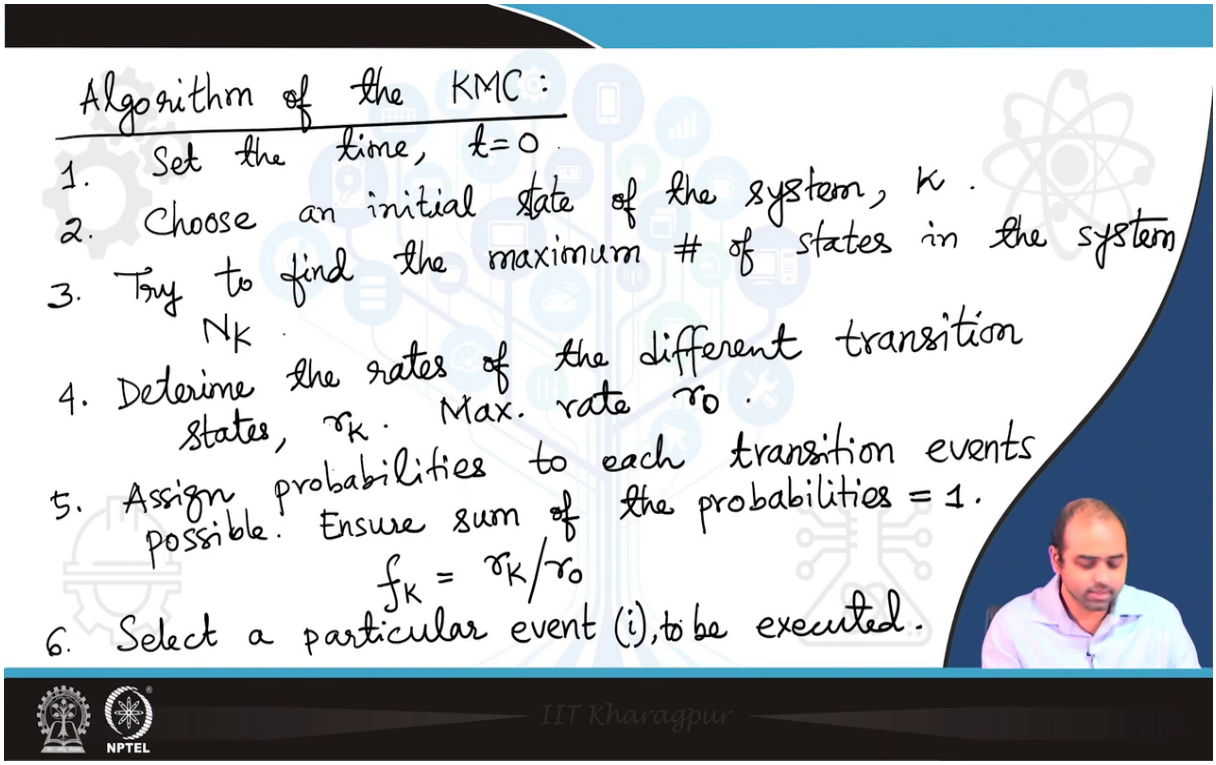
And from there you choose a particular random number and based on the random letter, say there are 10 different possible states, you choose a particular state or a particular event, choose a particular event actually to be executed. So, from that random number you pick a particular event let us say the different events are listed down with some index and you choose a particular event.

(Refer Slide Time: 24:49)



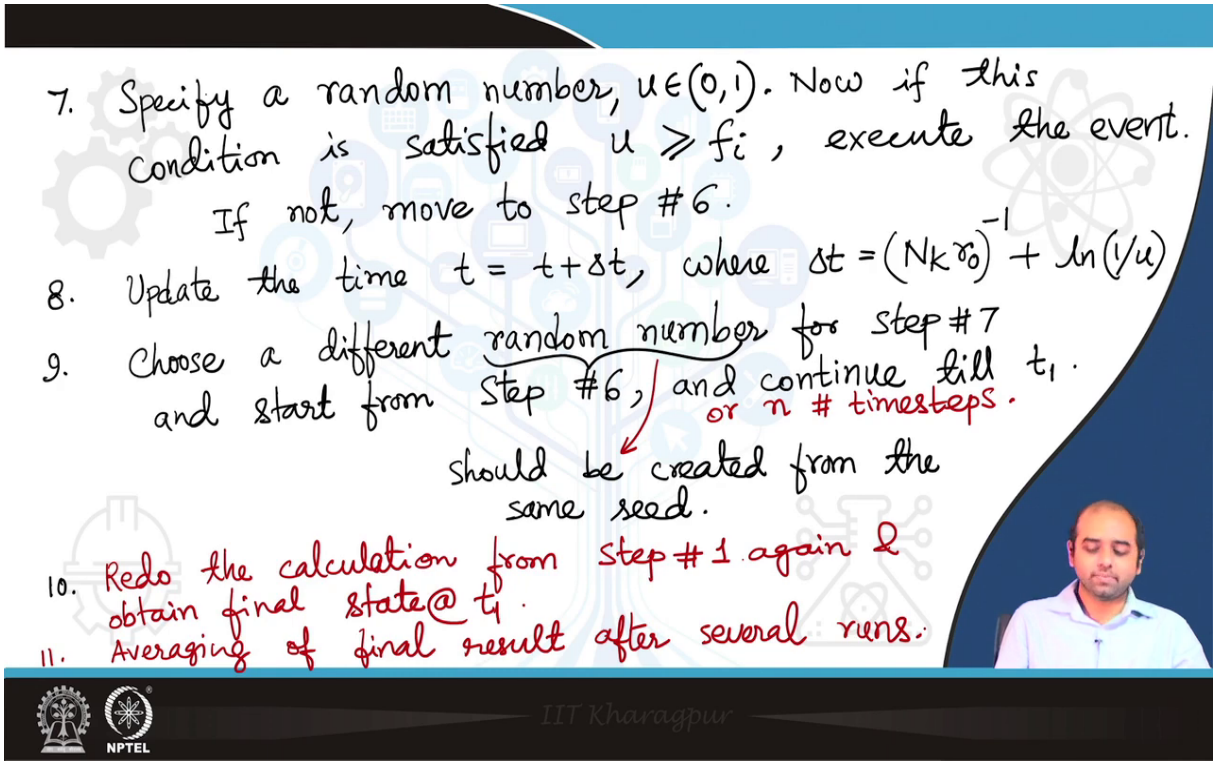

Next is specify a random number in between 0 and 1. Now, let us call that random number as u. If this u, if this condition is satisfied that for this particular event this f k is greater than u, then execute the event. So, if you are executing the event, so this means that the probability of this event or the probability of that particular transmission is more than this random number, then you execute the event. It could be other way around also. You can select, if the random number is more than the probability of that particular event then also you can select in that way, it is not not an issue. So, let us say this particular event instead of writing k, let us not confuse, this event is i.

(Refer Slide Time: 26:49)




So, let me say that this particular event i to be executed.

(Refer Slide Time: 26:52)



And let us say this f of i is greater than this random number, then you execute the event and move to the next step. If not, if this is not the case, then you move to step 5, the previous one, choose another event, move to step 6, a previous event and choose another event to be executed. Now, in the next state, you update the time, t is equal to t plus delta t where delta t is actually N k r naught inverse plus l n 1 by u. This is the formula for the updating of the time and you can see that if the rates are higher, R0 is the maximum rate or the number of possible transition data states then that would be low.

Similarly, it also depends on the value of the random number. So, if the random number is small then this update is reduced. Next is once your time is updated choose a different random number for step 7 and start from step 6. But please remember that this new random number whatever we are specifying should be created from the same seed.

Let us write it in a classical way that u is greater than fi, so guys please note that I have changed this condition, instead of saying that the random number, I mean the probability of that particular event is more than the random number, there is also other possibility that this random number is more than the probability that is assigned.

So, I mean both ways it works but you have to do this to find out that what is the difference in the final result. So, after carrying out these steps over a lot of time, you will get a final state right, all the where you see that most of the possibilities are almost covered and that gives you the idea that what is the likely scenario in this case.

So, after you reach a certain time and the finishing time, again start from and continue let us say continue till a certain time, continue up to let us say t1 time, so which you define that how much time you want to calculate or how many possible intervals, time intervals you want to take, that is something that you need to decide right and something will keep it same.

So, once this is done, then you redo the calculation from step 1 again and obtain the final state. So, after you do this let us say certain number of time, so either you can fix the final time or you can set the number of time steps that you want to do, time t1 or whatever, n number of steps, whatever you decide. So, whatever you decide that should be same for all the different steps that you want to do. So, like this you carry on for, I mean this time progression you carry on that in the final state.

So, similarly, start this again but as you have seen step number 6 and 7, there is a random number involved, you continue, I mean whatever you try to continue or wherever you try to

get every time the results will be different or the final state will also be different. So, that has to be averaged out. So, step 11 is averaging of the final result after several run. So, that will lead to a possible distribution.

So, if you do a calculation several times and you get several results, the final state will be sort of a distribution. So, from that distribution you can find out after doing n number or a number of reruns, you can find out what is the most probable state. So, since this is an event or this is a calculation which is based on the random number, the final answer that we get cannot be interpreted with absolute certainty, but we can say what is the most probable state or the most likelihood of a certain result or a possibility as an outcome can be determined or can be understood.

So, I hope all of you got a fair idea on this Kinetic Monte Carlo simulation. It will be much clearer when we try to, when you actually try to do a hands on some experience or try to work out a small problem, we will see if something like that can be provided in the assignment for you to try out and things will be much more clear.

Thank you, I hope all of you liked this class on the Kinetic Monte Carlo simulation. And it is something that could be useful for determining your transitions of sort of random process, which cannot be moderate actually, with well-defined physical system or well-defined physics of the process or the dynamics cannot be calculated by based on the physics of the system. Thank you. I hope all of you like this class.