

PROFESSOR: One more lecture from me. And so remember that on Thursday we'll be here. And Chris Wolverton from Ford will be lecturing, straight live direct from Ford. And so you better have good questions for him because he only lectures about an hour. And so we usually-- in the past, we've had quite some good discussion afterwards, but that's going to depend on you.

So what I'm going to do today is a variety of topics, but they all have to do with the same idea of bigger and faster and better-- not necessarily better. So when you deal with atoms, both because of the size and the time, if you go away, especially if you go away from periodic systems, you always run into this problem that it's too small [INAUDIBLE] system, that the simulation is too short.

I always like to quote Professor Marzari on this. When people come to him with nano problems, he always says, nano is just a little too big for us. And so I think that's telling you something about the state of what we do.

So you've actually seen certain solutions already to the sort of size and time problem. Maybe they were fairly obvious. But if you think of the time problem, what I've already showed you is how to essentially coarse grain that away, which is it's one thing to coarse grain over time and space. It's another thing, and it's actually easier to almost remove it by, integrating in a way.

And when you essentially integrate time away completely, you end up with thermodynamics. I mean, think about it. That's what integration over the partition function is.

You're essentially assuming the system has had enough time to sample all its excitations, so we essentially integrate over that ensemble. You integrate over all the excitations, and you end up with the thermodynamic functions. And the examples that I showed you with the cluster expansion were essentially examples like that.

So that part, I would say, for crystalline systems is a solved problem. It's not a trivial solution. It's still a lot of work, but it's essentially a solved problem.

I would say most of the spatial problem is solved if you're going to deal with static properties, static properties and things such as elastic behavior. You essentially coarse grain fairly easily and end up with continuum theory. The real problem lies in what sort of happens in between here.

When you want to keep some level of detail that's not continuum or in the time max is not thermodynamic, but you can afford to stay all the way down at the atomistic scale and the scale of the electrons, that's where I think the real challenge exists. And I'll tell you a little more about that. That's when you break your computer.

So first, I'm going to tell you there's obviously the brute force approaches, which is just throw more computer power at it. And these days the way you do that is not buying faster computers. You buy more of them. So you have to parallelize.

Parallelization over space is obviously the most obvious one. If you have a really big system, you can divide it in chunks. OK, you know? You'd probably do it a little more systematic, but you can divide it in chunks and give each chunk of atoms to a CPU, CPU1, CPU2, CPU3, et cetera.

Of course, that implies that the coupling between regions is not too severe. Because when you calculate-- let's say you're doing electrodynamics. When you calculate forces on the atoms in the region of CPU1, the atoms that live in the boundary layer obviously see forces coming from atoms in the other regions.

So how efficient this method is will depend very much on how well you can divide the sort of spatial origin of the forces and the energy. So you can already see that this is going to be easier when you do it with pair potentials, say. With pair potentials, you can directly assign, when you have the force on an atom, the component coming from other atoms. You can say the force on this atom comes-- I have a piece coming from the potential with that atom, the potential with that atom, et cetera.

You'll see that in the standard approximation, for example, the quantum mechanics, this is virtually impossible. When we work with block states, those are, by definition, extended over the whole system, OK? So you can't partition them over pieces of space.

So to do this sort of divide and conquer approach with quantum mechanics, you have to go to alternative approaches. People invest a lot of time now in just real space grids. In theory, the Schrodinger equation is a partial differential equation. Like any other partial differential equation, you should be able to solve it on a real space grid. There's all kinds of complications that arise, but people are doing a lot-- there's a lot of work actually being done on that.

The other is to go to localized basis sets. And you know, maximally localized Wannier functions fall in that category. And Professor Marzari is an expert on that, almost essentially invented-- tight binding approaches.

So you'll see a lot of work being done on real space methods, which is, in essence, directed largely towards this sort of divide and conquer approach. If you can go a real space in quantum mechanics, you can divide space up. So space is maybe, in the end, not the biggest problem we have to deal with. Time is a much harder one. Because how do you parallelize over time?

Time is, by nature, sequential. So it's really hard to do something that's in the future when you don't know the present yet. But you'll see in a second-- well, not in a second, a little more than a second, but later in the lecture. There are some approaches to paralyzing time, but they're all, I would say, much more limited in scope than the methods for parallelizing over space.

So you see time will be the big problem. Time is easy to deal with when it's either short or very long. Because when it's very long, you essentially integrate it away. But in between is where it tends to get harder.

So let me give you sort of a couple of examples of ideas that have been kicked around for coarse-graining space. And then I'll say a little more about the time. And with time comes temperature problem.

So you could imagine, if you do an impurity calculation-- so you take a sort of crystal of these bluish atoms, and you stick a big red atom in. You want to calculate the energy [INAUDIBLE] with quantum mechanics. Well, those shells around the red atom kind of relax a lot. And then as you sort of go further, this is largely an elastic problem.

You've put a ball in a hole that's too small, and sort of the system relaxes around it. We tend to sort of brute force that typically. We do quantum mechanics on bigger and bigger units cells.

But it's actually remarkable if you ever plot the displacements. You will find that these systems behave remarkably elastically in very close shells to the defect already. Often, from the second, third neighbor, if you were to apply elasticity theory, you'd still get it right. It's amazing how applicable elasticity theory is at the atomistic level.

So that gave people the idea that really when they do calculations with broken symmetry, so around defects, maybe you really don't need all the atoms that we tend to put in calculations. So the idea that's been kicked around is that let's say that this is a piece of your simulation that's far away from whatever perturbation you're looking at. Maybe it's a grain boundary. Maybe it's a dislocation moving. Maybe you just did nuclear fusion somewhere in the material.

But far away, you believe that the system acts almost in a continuum mode. And so the idea is that, what if I kept track only of these red atoms? Let's say I'm looking at energy of an elastic displacement or even not elastic.

If I knew how these four items moved, I could probably know something about how the other atoms move, OK? So I could interpolate, say, their displacement from the displacement of the corner atoms. So one thing you can do is say, well, if I know how these atoms move, I assume that the energy of the atoms in between is essentially the same as an energy of a sort of macroscopic solid that I deform with that displacement.

You can do things in between. You can actually sort of calculate the energy of that unit cell as it's displaced. But the whole idea that essentially you start keeping only track of a limited number of atoms. And you can think of these as nodes.

So these are called nodal atoms. And what happens to all the atoms you removed-- you essentially get by interpolation. And once you're happy with using elasticity theory, then essentially you know the energy of that cube, how that changes as you change the four corners.

Typically, people use triangles because people will tend to use triangular measures. But I wanted to give the example with a square since it's a little simpler.

So this is essentially the idea of the quasi continuum method. Let's say out here you have some perturbation, something you want to study at the atomistic level. As you go farther and farther away, you want to coarse grain more and more.

So essentially, you want to keep less and less nodal atoms and get more and more information simply by interpolation. So you could say this is coarse graining at the level of one out of four when you're here. And then you have a few regions of that.

And then you coarse grain at a higher level. So you keep, say, I should have removed-- sorry, these should not be red. These should be blue, sorry. So you only keep these four atoms, and you sort of assume that that solid undergoes a homogeneous deformation within that box.

What you see is that somewhere in this limit you're going to end up with a finite element approaches, where essentially you do continuum theory. And you essentially say that, if I know some displacements of elements, of the corner points, the vertices of elements, I know how to that element deforms.

So what you have is essentially a sort of continuous transition here in coarse graining from our atomistic behavior to a finite element approach. And people have put that together. And this is called a quasi-continuum approach.

This is sort of a pictorial version of it. Let's say this is some boundary where you do something. You have very fine resolution. You could say the nodes of your finite element mesh are the atoms.

And then as you go farther and farther away, you coarse grain more and more. And you keep less and less information. I'll give you a reference at the end, but the people who developed this were essentially a fairly small group of people, Sheno, Tadmor, and Rob Phillips, and Michael Ortiz at Caltech. And there's some excellent review papers about this if you'd like to read more about this.

It's now been fairly well-implemented, these kind of approaches, mainly with empirical energy methods. People have tried to implement it with quantum mechanics. And some people have actually done it, but it's not all that effective really. It's extremely hard to couple the atomistic region where you do quantum mechanics to a sort of more coarse grained region.

But with potentials and with things like embedded atom method, people have coupled these things fairly well together. This is, for example, a grain boundary-- sorry, a crack impinging on a grain boundary. See, this is a symmetric grain boundary.

And you have the crack in the middle there. It comes in, hits the grain boundary, and essentially deflects in the grain boundary. So that's a simulation done with quasi-continuum. You know, if you want to think about it, what it really does for you quasi-continuum, is that it's a way of getting the boundary conditions right on your atomistic simulation.

By slowly integrating it, essentially with a continuum theory, you, first of all, have adaptive-- your boundary conditions or your atomistic simulation change during the simulation. And because you have a better embedding theory, you have a better continuum theory, they're sort of more appropriate. And especially in mechanical problems, that is very important to have the right elastic boundary conditions.

So let me sort of position this for you to show you where we have the solution and where we still have significant problems. You know, I would essentially say, when we work purely at the microscopic scale, we have most things under control. We know well how to deal with the energetics.

We fairly well how to deal with the dynamics. Because for pretty much everything, except things like hydrogen, you can use Newtonian dynamics even on the atomistic scale. You could argue we don't really know how to deal well with electron dynamics, but that's another problem.

If you're willing to go all the way to the continuum scale, we also know the equations that describe matter. It's essentially thermodynamics and elasticity. And you could think of thermodynamics as the integrator of those.

It's when we live at this inhomogeneous coarse grain scale that we're really in trouble. And I'll give you an example. Think of temperature. In continuum theory, temperature is a field.

In a microscopic simulation, it's kinetic energy of the atoms. You can already see a problem appear. Essentially, as you inhomogeneously coarse grain around that atomistic region, that temperature has to evolve from being the kinetic energy of motion to a field.

And it's [INAUDIBLE] tricky to know what you do in between. I mean, I know how to deal with temperature as a field. I know how to deal with this kinetic energy.

But what is it in between? What is it when you say a factor of 10 coarse grain, so you've removed every 10th atom? And I'm going to go a little deeper in some of these things, but those are some of the problems.

We also don't know what the dynamics here is at this scale. Because think of how do you dissipate energy. How do you dissipate energy in a microscopic system?

You dissipate it, essentially, by non-harmonic vibrations. That's how you dissipate energy. How do you dissipate it here?

Actually, you don't unless you put it in explicitly. And in perfectly elasticity theory, you don't have dissipation. So again, energy dissipation is a big deal in inhomogeneously coarse-grained systems. Because the mechanism by which it happens at the atomistic and at the continuum scale is different.

So quasi-continuum approaches have been extremely successful for elastic static properties and even nonlinear elastic problems. It's when people try to put temperature and dynamics in that life gets a lot more complicated. So there are essentially two approaches that have been suggested and tried. And both sort of seem reasonable, but have significant failures.

One is to simply-- let's say, how would you do molecular dynamics on a coarse grain system like this? Well, remember, you've removed these atoms. You could lump their mass into the nodes.

So essentially, the nodes of your mesh now become heavier. And then you do MD on those. The problem with that is that, as you coarse grain more and more, you get the heavier and heavier nodes.

And after a while, they don't move anymore. Because if they have the same temperature, mv^2 average over 2 is the temperature. So at the same temperature, the v becomes very small. The other one is a sort of often used approach to use static optimizations. But rather than say get all the positions of these atoms-- rather than get them from the energy, get them from minimizing the free energy.

Say, the force would be now the gradient of something like the free energy rather than the gradient of the energy. So you have some amount of temperature effects in there then, but you don't truly have dynamics. So one thing this would do for you, for example, is you'd get the right bond length and lattice parameters.

OK. I want to show you some other ideas that have been out there. This is one that I've worked on for a short while. But the generic idea has really been out there in the community a long time. And it's very similar to what we've done in the coarse graining of time.

The idea is essentially think of three atoms on a chain, 1, 2, 3 there. What if you want to remove atom 2, and so coarse grain and only keep 1 and 3 as the nodes? The question really is, what should your potential-- let's say you do this with potentials. Don't even worry about quantum mechanics.

What should your potential between 1 and 3 be, so that they behave the same way as if 2 were there? So you're going to take 2 out, but you want to set up a potential between 1 and 3 that acts like if 2 were there. Well, you can get the solution again from thermodynamics, essentially requiring that the partition function for the motion of 1 and 3 is the same with or without 2 there.

And that's essentially what's written out here. If you think of q 's and p 's for the coordinates, remember q 's are spatial coordinates. p 's are velocities, essentially the moment.

Essentially, what you're saying is that the free energy of the Hamiltonian with only q_1 and q_3 -- so now, you keep only the positions of 1 and 3-- is obtained by integrating away the coordinates of 2. So you integrate over the position and over the momentum of 2. And if the motions are classical, then the integration over the momentum is trivial.

You always get the same factors from momentum. So it's really the integration over the coordinates. And so if you integrate this away, then formally this integral has no q_2 dependence by definition.

You've essentially integrated over all possible displacements of 2 to get the interaction between 1 and 3. And so that then defines a potential just between 1 and 3. Now, there's all kinds of practical issues.

You can really only do this effectively if you have interactions that are spatially limited. Because, otherwise, you really don't know what you all have to integrate. So this works great with pair potentials then.

So you know, what does this physically mean? It really means that, if you, say, displace atom three and you look at the force that it produces on atom 1, that you're looking at the kind of screening by atom 2 is included. In the real system, atom 2 would sort of maybe screen that displacement away somewhat. And that's not included.

And people have done this before, looking at interactions of-- when you look at interactions of ions and fluids, the way to look at the interaction between two ions is by integrating over the possible displacement of all the other ones in between. And that's how you end up with screening. So this is sort of the screening equivalent in displacement fields.

And what you get at is very much what you'd expected. This is the potentials for different levels of coarse graining in normalized distances. So if the length of the bond length is in units of 1, so this would be the normal potential.

If you coarse grain one level, so you remove every other atom, you actually end up with a potential that's twice the bond length. Because, remember, now you have atom 1, 2, 3. You've removed 2, so the equilibrium distance between 1 and 3 is about 2 times the bond length.

And you keep on coarse graining, and you end up with a potential that's 4 times the bond length, 8 times the bond length. OK. It turns out that special coarse graining always tends to cause some kind of time quartering. That's a much more difficult issue. I don't want to say much.

In 1D, it's easy to remove atoms and add up. So when you remove atom 2 that's between 1 and 3, essentially that defines a potential between 1 and 3. In 2D, it's a little harder how you partition bonds, but you can do all kinds of bond moving approximations. This is a simple one in a triangular lattice where, if you have a triangle, you essentially want to remove these guys here.

And so you can define some scheme by which you sort of collapse the bonds onto these sides. You can also define other schemes. So here's an example of an inhomogeneously coarse-grained system in 2D where, in this case, it's fine at the edges and it's coarse in the middle.

OK. There are essentially two major problems with almost any coarse-graining scheme. And the first one is the worst, is that whenever you go from a-- let's say you're atomistic out here. Here, you've coarse-grained level 4, coarse-grained another factor of 4.

You cannot sustain short wavelength phonons in the coarse regions. Let's say you have a phonon come in here that's, say, has a wavelength like this. Since here you only have these nodes, this is essentially the minimum wavelength you can have. It's the minimum bond distance.

So at every interface between a fine region and of coarser region, what happens is that there's a series of phonons. The ones that have a wavelength that's too short get reflected back because they cannot go into the coarser region.

And ultimately, you hit the continuum region, and you can't have any phonons. Because there they should all be dissipated as temperature. So what does that cause?

Well, if you look at dynamical processes, let's say you have a fine region in which you have some reaction that dissipates energy or that creates energy, the way that that energy has to get out is by phonon transmission. You have to send lattice vibrations out which, through their n harmonicity, have to get dissipated somewhere. And so any time you block phonons, you're essentially reducing the thermal conductivity of the system.

So this coarse graining gives you lower thermal conductivity. And I'll show you some examples in a second. But you know, it does worse things. It's essentially elastically confining, also, your system somewhat, but only in a dynamical sense.

So it's like you're having, for certain, dynamical nodes, a small unit cell. And so that means that they reflect back of the coarse graining interface. And they hit your region of action fairly quickly as soon as they're reflected back.

So it's exactly the same as putting your dynamics in a very small box except that the size of the box that the system fields is different for different wavelengths. So a certain amount of elastic behavior gets reflected back fairly quickly. So the phonon transmission problem is a serious one.

And the phonon transmission one is usually recognized very well. There's one that's not as well recognized, but is probably in the long-term just as severe, is any time you remove degrees of freedom you remove entropy. Because the entropy is essentially a count of your degrees of freedom.

You may not worry about that. But the problem is that, if you do dynamics, your system will equilibrate along certain derivatives of the entropy. And you'll get them wrong. If you get the entropy wrong, you'll get its derivatives wrong.

And some of the ones that you may worry about are heat capacity, which is the temperature derivative of the entropy, but also this one, which you may worry a lot more about if you do mechanical behavior, thermal expansion, which is the volume derivative of the entropy. Remember your Maxwell relations? The sdv is essentially dv/dt up to some factors.

So how the entropy changed with volume determines the thermal expansion. And that's why, if you actually look back, the slide I showed you, when we define the effective potential between atoms 1 and 3, there's actually a potential. And then there's a part that doesn't contain the coordinates of 1 and 3.

And it's essentially the entropy you've lost by removing node 2. You have to actually keep that entropy in the system if you want to get the total thermodynamic qualities right. It's sort of like you see that this is obvious if you take the complete limit of removing all nodes or going to 1 node. Then you end up with continuum behavior.

And the degrees of freedom of continuing behavior are not the entropy. You've actually removed all the dynamic motion when you go through continuum theory. So unless you've kept track of the degrees of freedom and the entropy that they include, you've lost all your information about that.

So when you do something like this here, these will actually, if you define the potentials at a given temperature, because the renormalization potential depends on temperature-- if you define this at a given temperature and you run this at any other temperature, these regions will have different thermal expansions. And so you'll build up internal strain in your system without knowing it.

I get a new color scheme. There we go. So if you take a simple pair potential model, these coarse graining schemes work pretty well. Here's essentially a static property even though it has temperature in it, but this is the strain in that 2D system versus the stress.

And the black curve underlying this is the full system, so no coarse graining. And the red curve is that coarse grain system that I showed you. And you'll see even in the non-linear part does this very well.

And it makes sense, you know. This kind of mechanical behavior, strain versus stress, is largely dominated by the potentials in this kind of solid system. There's very little dynamical factors that play a role in it.

So as long as you normalize the potentials right, you will get the stress-strain relation right even in the non-linear regime. So that's not a big surprise. If you keep track of the entropy right, you'll get things like the heat capacities right.

This is the ratio between the heat capacity in the non-homogeneous system, which is the coarse grain system, divided by the exact result, you could say, the heat capacity in the homogeneous. It essentially oscillates around one as a function of temperature. So by construction you pretty much have that right.

So here gets more interesting. If you look at heat conduction through a coarse grained interface-- so let's say you're fine at this side. Let's say these are atoms. And here we've removed a bunch of nodes with coarse grains.

And so you look at the heat conduction through this. So you set up a temperature gradient and essentially look at the heat flux when you do dynamics. So what this is showing here is the heat conduction in the non-homogeneous system divided by the heat conduction in the homogeneous system.

And this is pretty much always lower than 1. So the non-homogeneous system has less heat conduction. And the reason is you scatter phonons on this interface.

So here there is a certain amount of phonons that get reflected back. But it's interesting. If you show this here as a function of the length of the homogeneous system, you do better and better as this system gets bigger.

And that makes sense because then the interface becomes less and less important. But, you know, look at the size here. To get up to sort of about 1, you need about 1,500 atoms in the homogeneous region.

So this is a major problem with full phonons in general. I think of any atomistic level phenomena phonons are probably the ones that have the largest length scale. The phonon mean free [INAUDIBLE] at a low temperature is enormous. It's thousands of atoms far.

That means that phonons are usually the first thing to see something far away. They're the first thing to see finite size effects. They're the first to see scattering of interfaces. Because the reason is that a lot of systems behave at low temperature fairly harmonic.

That means that these phonons just travel and travel and travel. And their probability of being scattered is pretty low because it's only as they start becoming enharmonic that they scatter. Of course, they could scatter off disorder. So if you have the disordered systems, the mean free path is a lot lower. But phonons you often see hundreds of angstrom far, essentially.

OK, here's the same result as a function of temperature. So the heat capacity versus the normalized temperature, this is the real system. So this is a fully atomistic level system. And this, in red here, is the partially coarse grained system, in this case with two interfaces.

Now, what you see is the coarse grained system always has lower heat conductance. They tend to come a little closer together at higher temperature. And the reason is at higher temperature you have more enharmonicity. So the phonons have a shorter mean free path. So they tend to homogenize easier and don't see the interfaces as much.

OK, let me skip this. OK. So that's some of the efforts going on in coarse graining over space. I think if there's an area in which you want to work and have high impact, this is probably one.

I haven't seen a single good idea in this field in the last 10 years. I probably shouldn't say this on tape, but, you know, this is a field where progress is desperately needed. And it's a really, really hard problem.

Because you're essentially asking the question, what is the dynamics of partially coarse-grained systems? And the reason it's so hard is that it's a mixture of sort of Newtonian mechanics, which is energy conserving, and dissipated dynamics. And you have to sort of-- and that how you mix those essentially depends on the frequency of your motion.

And that's why this is such a hard problem. People have tried things with boundaries that have complex impedances, where essentially your transmission through the boundary is frequency dependent. And you can see, with that one, you can definitely sort of tailor better what the phonons do. Because, now, you have a much more complex interface literally to control.

The problem is that some of these things you can make work, but they get so complicated to implement that they're not necessarily very practical. Other people do things with large overlapping regions. You could say one way to solve this problem is to not have these sharp boundaries between different levels of coarse graining, but overlap the two regions and slowly force them in the overlap region. The solution, say, to the displacements fields are the same.

OK. The next thing I want to talk about is accelerating time. I said without Einstein, but I think Einstein only could slow down time. Is isn't that right? I'm not sure we can accelerate, no? I don't know.

I don't know my theory of relativity too well anymore. I think I don't have to convince you-- you've done MD-- why you need to speed up time sometimes. Most of what I'm going to say comes out of this review article by Art Voter, who's probably been the guy to work on accelerated MD methods. And I'll give you the references again at the end.

This is a great review if you want to read about this because there's some really clever stuff in here that I think, even if you don't want to do accelerated MD, that you could use for other things. So first of all, what's the problem? The problem is obviously that, in most systems that are fairly dense, you have well-defined minimum phase space. And often the barriers between them are pretty high.

So you sample the transitions between them at a very low rate. The rate is essentially proportional to the exponential of the activation barrier between the minimum. And like I said, in a lot of condensed systems, those activation barriers are quite high and lead to timescales that you can't quite sample.

So I'm going to talk briefly about three different methods to speed up time parallel replica dynamics, hyperdynamics, and temperature accelerated dynamics. OK. Let me sort of quickly review what you would do with transition state theory. So if you knew the transitions that a system had to make, like in, say, a simple molecular reaction, then you really don't need to do a simulation. Because if the activation barrier, this quantity and you have some idea of the frequency with which the system tries to cross that barrier, you can do transition state theory and essentially say that the crossing rate is going to be something like the attempt rate, ν , times the success rate, which is this Boltzmann factor.

And the chemists among you and chemical engineers know there's all kinds of approximations in there. But none of them are particularly severe for most things to study. Like, one approximation just assumes you don't cross back.

Often, when you sit at the transition state, you're essentially assuming that you fall over and don't cross back. The other approximation is that this assumes that the transitions are infrequent compared to the vibrations in the minimum.

So you assume that when the system goes from here to here it's sort of equilibrates here, vibrates around until it makes the next transition. And the reason you need to do that is that you need to equilibrate first in every potential well before you go to the next one. Because, otherwise, you have dynamical memory left from the previous transition, which would essentially change your statistic, would make your statistical mechanics approach invalid.

But usually those are easy to deal with. It's sort of obvious. Because if you think about it, if the system crosses very rapidly compared to the time it vibrates, then you just do molecular dynamics because then you have very fast transition. So who cares about transition state theory?

OK. So the parallel replica method is really kind of elegant in its triviality almost. If you think of what is time, time is waiting for events to happen. And if you think of a really trivial example, let's say the system wants to move from this state to that state. You're just sitting there.

The atoms are vibrating around. And you're waiting for a system to transition. Why not wait on many processors at a time?

So the idea is you run this simulation on a bunch of processors. And if you run on, say, 100, you could say that's kind of like waiting 100 times the time you wait on one processor. And then let's say it happens on one processor. The system transitions.

Then you essentially add up all the time all the systems have waited, and that's your total waiting time. And in a statistical sense, that's actually right. If you do queuing theory and things like that, you'll find that this is the solution.

It's actually easy to see if you turn this all into probabilities for transfer. Essentially, you could say, when you do it 100 times, you've 100 times attempted or given the system chances in a given time to make a transition. So you add up all the time, and then you just restart.

So you put the system in the next state. And you again run all those replicas. You have to introduce a little bit of randomization, of course. You don't want these to follow exactly the same trajectory through phase space. So there's a certain amount of randomization that goes on.

So essentially, this is linear time scaling. Overall, you pretty much-- if you have n processors, time is accelerated by a factor of n you could say. It's a little less because there's overhead. You have to, first of all, do the non-trivial thing, which is detect the transition.

When you think about it, this is kind of not that trivial. You're doing your MD simulation. And you've got to sort of ask your system, oh, has something happened?

And maybe it's, like, an atom diffusing, whatever. So you've got to do quite a bit of work to sort of check whether a transition that happens. So you've got overhead there. And then you have overhead when you sort of copy the transition state into all the other ones. You have to do a little bit of randomization here.

So typically you may restart with an actual Boltzmann distribution of velocities and so. But you have to initialize that. You have to run that for a very short time. And so you have a bit of overhead there. So you'll get a little less than end scaling.

But it's not bad. It's not great either. But like I said, with 1,000 processors, you can go from nanoseconds to microseconds. So yeah, that's not bad. I mean, you can see that you're never going to make it to seconds because of the linear scaling. But it is a sort of intermediate timescale that can be useful.

Here's an example. This is planarization of silver on silver. And this comes straight out of Art Voter's article.

So essentially, this is a 1, 1, 1 layer of silver. That's the white atoms. And then on top of that is a first layer, which is the blue atoms. And then the yellow atoms is another layer on top. OK, so you're seeing three layers.

And so this study tracks how the system planarizes to lower its surface energy. So the top atoms should sort of come down and form all one island. And so this is perfect for accelerated MD.

I mean, researchers know why they pick problems because they can solve them. Because this is a set of discrete events. You know, atoms diffuse and maybe even collectively.

You see, the nice thing about MD is that you're not imposing the kinetic mechanism. But it's still a set of discrete events. So atoms may help a little-- you can see it in the first few slides there-- and ultimately sort of planarize.

People got all the way up to 1 microsecond. Although it was done with empirical potential. And according to Art Voter, this took only about five days on 32 Pentiums. And these are old. I mean, these are 1 gigahertz Pentiums, so probably do it in maybe one to two days now.

I think you can already see, if I go back, when this will not work. This will not work if the overhead becomes excessive. And how is it going to become excessive-- if you start getting a lot of transitions.

If only after a little bit of simulation you get a transition, then your overhead starts to weigh heavily. And when is that going to happen? Well, it's going to happen, of course, when you have low activation barriers. But there's another problem which is a little more severe, if you make your system bigger and bigger.

See the probability-- let's say you have a system that's sort of locally identical. If you make it bigger and bigger, the probability that a transition occurs-- let's say an atom diffuses-- is proportional to the system size. The probability that an event happens is proportional to the system size. So as you make the system bigger and bigger, this essentially becomes less and less efficient because you get more and more transitions. And that's something you'll see in other versions of accelerated MD as well.

OK. Hyperdynamics, I like the word very much, but I'm not going to say much about it. Hyperdynamics is really a class of methods that is getting very popular not just in accelerating MD now. But the whole idea of modifying the potential surface is getting sort of very popular also for pure optimization methods.

If you think that this is the surface of the original system, this is the energy surface, your problem is that these wells are too deep. Well, solve the problem by lifting them up. So if you can define a potential that tracks the original potential in most of space, especially the activated pieces, but then lifts up the minima of the well, if your system moves on the red potential, it's going to transition a lot faster because the activation barrier goes down a lot.

And you can actually run the system on the red curve and correct for the transition rate. Because, see, the cool thing is that once the MD finds the transitions-- so MD is going back and forth here. But once it's found the transition, you could calculate back what the real barrier should have been. So you know at what rate the system should have transitioned.

So boosting the potential surface with this boost potential-- that's this difference here, boosting it up-- is essentially accelerating the escape from the minimum. And you can correct back by this Boltzmann factor, which is essentially comes from the ratio of the proper rate to the rate that you actually had in your MD simulation.

There's all kinds of tricks involved. This is a simplified version of it. There's a lot of questions about how you define that boost surface. For example, do you want to keep the same frequencies?

So then you'd like to keep the curvature the same. Because then you keep the same attempt frequencies, but you're only changing the success rate of crossing the barrier. That's a very elegant way of doing it. And people have ways that they can do this.

So doing this in practice often means calculating derivatives of potential surfaces, in many cases even Hessians, so matrices of second derivatives. And that's easy to do when you have potentials. It's a lot harder to do when you're doing quantum mechanics.

And so it's, again, one of these things that is usually only done with empirical potentials. Actually, I don't know of any implementation with quantum mechanics, but it may be something I've just missed in the literature. This is used for sort of other schemes as well.

If you for a second don't care about dynamics, there are now schemes out there to use this for optimization. If you think of global optimization of a system is a big problem because there's all kinds of local wells, well, what you can do is, if you fall in a well, you essentially start filling up the well. It's sort of like water filling up the well until you fall out into the next well.

OK. And then you start filling up that well until you fall in next well. And it's sort of like water cascading down an energy surface. And it's a way to do global optimization in a landscape of many minima. I think this method was developed by people at Princeton.

So that's sort of in the category of hyperdynamics. The one that's actually very practical is the next one, which is temperature accelerated dynamics, which is, again, a sort of obvious idea. If stuff runs too slow at low temperature, run it at higher temperatures. It'll go faster. That's actually something people have often done in MD without much justification.

What Art Voter's group did in Los Alamos is essentially give this a justification and show how you can correct for that temperature difference between where you would like the result and where you're actually simulating. And so the idea of TAD, as it's called, is to use the high temperature to find the transitions, but then execute them with the proper rate of the low temperature.

So essentially, you run at high temperature. So you fairly quickly scan the potential surface. And that tells you what the transitions are.

And then you go back to the low temperature, and you do a sort of statistical mechanics, a kind of kinetic Monte Carlo scheme or a transition state theory approach and execute the transitions. So I'll show you an example. The idea is that you run at high temperature until a transition occurs.

Once you find the transition, it's a trivial matter of finding the activation barrier. Then you reverse the transition, and you run again at high temperature. And the idea is that you get this catalog of transitions.

So you get this catalog of possible transitions. Now, why do you need more than one, you say? Well, the reason is that the one you find at high temperature may not be the one that actually gets executed at low temperature. If they have different activation barriers, then their rates may cross with temperature.

Here's an example. If I show the crossing rate versus $1/t$, let's say you run at high temperature. If these behave Arrhenius, then their crossing rate, which is $1/t$ over the time you have to wait, goes like $1/t$.

You may find one that has a high crossing rate at high temperature and this one that has a lower crossing rate at high temperature. But if they have a different activation barrier, they'll extrapolate differently to lower temperature. But what you see at low temperature, the one with the high rate at high temperature has the lower rate at low temperature.

So because transitions have different activation barriers, they may not be in the same order, so the same frequency, the same rate at different temperatures. And that's why you need a catalog of them. You essentially need a window of rates, of transmission rates, that give you a certain certainty that at the lower temperature you found the lowest one.

And you can kind of do back of the envelope calculations. If you make some assumptions about the range in which your EAs, your activation barrier, can vary, then you have the range of slopes of these Arrhenius walls. So that can tell you a little bit about how long you should be running at the high temperature to make sure you found all the transitions or are likely to have found all the transitions that get executed at the low temperature.

So does everyone see the idea? You're essentially using molecular dynamics as a way of finding transitions. And that's the hard problem in any kinetic from. It's finding the transitions.

Once you find them, you can calculate their activation barrier. And you can execute them with standard rate constant theory. But it's finding them. And MD, because it's unbiased, is great at doing that.

OK. So the approximations of the method are fairly obvious. To do this extrapolation, you assume harmonic transition state theory. So you essentially assume that you have a simple Arrhenius extrapolation between temperature, so that the exponential factor is constant. And like I said before, you have to make sure that you found all the mechanisms, that you found enough mechanisms at high temperature, so you definitely have the fastest one at low temperature.

OK. And here's another application out of Art Voter's work. This is copper on copper deposition. So they literally, I think, add atoms to a surface and then equilibrate them. And they run-- so they do direct MD for 2 picoseconds. That's how the deposit the atom, but then run 0.3 seconds of intervening time with TAD.

Now, you may think, 0.3 seconds, that's a lot. Well, it's not because you're really doing transition state theory. If I have barriers of 1 electron volt, then pretty much at room temperature 1 electron volt, that tends to give you rates around 1 per second. That's kind of the time scale.

So if you have barriers of 1 electron volt, then every time you execute a step, you're 1 second further on average. So the 0.3 seconds, that is not that impressive. You know, I can do 1,000 seconds with one step just by having a system at a very high activation barrier.

OK. You can look at the different events in the simulation. And one of the nice things is, because you're doing MD, your system is unbiased. And one thing we've learned more and more by doing MD on surfaces is that the diffusive processes are not at all the way people thought they were.

People think of atoms hopping by themselves. But because surfaces have so much open space, have such a low symmetry, you see a lot of collective behavior. For example, here's one. This blue atom here pushes in the surface, pushing that orange atom or kind of-- what is it-- brown.

And that pushes then these two out. And that's all one collective event. So these are not things you would easily guess at. Here's another collective event of sort of a kind of roll of atoms three at a time at a step edge, essentially kind of moving down.

OK. So you typically need MD methods whenever you don't know or you're not absolutely sure what the transmission mechanisms are. If you know what the transmission mechanisms are, then it tends to be much more efficient to do Monte Carlo. You can set up kinetic Monte Carlo schemes that execute transitions with the right transfer rate. It's just that you have to know what those transitions are.

So you've already done Monte Carlo on lattice models, but we saw it as a sampling method. You could also, now, see it as a kinetic method. Let's say you have a lattice of atoms. Let's make it easier even, atoms and vacancies.

So we've shown you how to do the thermodynamics. Well, let's say you want to study diffusion. If you know microscopically how this atom would migrate to there, through which path it would do that, then you just calculate the energy along that path. And you can now do a Monte Carlo where the way you travel through phase space is by a real kinetic mechanism.

OK. So you could now take your exchange rate, dependent not only on the energy of the initial and the final state, which is the way we did it in a Metropolis algorithm, but also have a prefactor now that's kinetic. So that has a frequency times an activation barrier. And that's essentially kinetic Monte Carlo.

OK. But you see, kinetic Monte Carlo implies that you know what your possible migration mechanisms are between states. And usually you know that reasonably well in things like crystalline solids. So there kinetic Monte Carlo is very applicable. People also use it on surfaces even though it gets a little more dicey there.

But once, of course, you go to sort of very disordered systems, it really gets way out of hand what the kinetic mechanisms are. And this wouldn't work anymore. So I was going to show you one example to sort of end up. And on the continuity, I was going to take the same example that we did for the phase stability. If you remember this when we talked about the cluster expansion, we looked at the phase diagram of lithium and lithium vacancies and lithium cobalt oxide.

So remember the issue is that, if you take lithium out, you create vacancies there. And the issue was how did they organize. And I showed you how to use a cluster expansion to do the phase diagram, the thermodynamics of that.

Let's say you want to worry about the kinetics. So essentially, how fast can you get that lithium in and out? So you essentially want a diffusion constant in that material.

Well, this is fairly well-suited for kinetic Monte Carlo rather than molecular dynamics. First of all, the rates are fairly slow. So doing anything with transition state theory is going to help you a lot.

And you have transitions between well-defined positions. The atoms go from one well-defined site to another site. So if you had dilute diffusion, you could just use random walk theory.

So in that case, diffusion is sort of trivial. All you need is an activation barrier and a prefactor. All the rest you know. You have the lattice constant, which is actually the jump length. And you have the geometric correlation factor.

These are all known. The things in green you would have to calculate. But let's do calculation. Random walks only apply when you have dilute diffusion. What I mean with dilute diffusion-- it's when the carrier of diffusion is very dilute, so it doesn't interact with itself.

So if you have a substitutional diffusion with very dilute vacancy concentration, then this would apply. The problem, of course, that we're going to look at the lithium cobalt [INAUDIBLE] is by definition not dilute. You essentially can go all the way from all lithium on the site to all vacancies on the site. So you go to very non-dilute regimes.

If you want to do that, you need to actually simulate the diffusion-- and you may have even done that in the MD homework-- by doing some kinetic model and then tracking the root mean square displacement. This is sort of a simple form for the self-diffusion. The chemical diffusion constant, which is the one you would use in macroscopic theories, is the self-diffusion times the thermodynamic factor. But the thermodynamic factor is essentially the derivative of the chemical potential of the composition, so that you already have from the thermodynamic calculation.

OK. So I have a few more slides here that I already used, so I'm going to [INAUDIBLE] through them. You remember how we did the thermodynamics? We built a lattice model. We calculate a lot of lithium vacancy arrangement, build a cluster expansion, and do Monte Carlo on that. And that gives you the phase diagram.

Let me skip that since we did that already. Those are the interactions. OK. How would you do kinetic Monte Carlo? I sort of also mentioned that. You would now execute all the exchanges in the Monte Carlo that look like diffusive processes. And you would execute them with the proper rate. So you would have an activation barrier in there.

OK. I think I said all this. OK. So here's an example of how that works out. If you calculate barriers in lithium cobalt oxide and you then do a kinetic Monte Carlo simulation, you keep track of the root mean square displacement. You get the macroscopic diffusivity.

And as you see, it varies by orders of magnitude. It's 10 to the minus 7 here centimeters squared per second. And it's 10 to the minus 13 here. You can already see that you would never get to all these time scales with a simple dynamics model.

OK. How you get activation barriers-- typically, we get this with what's called a nudge elastic band model. You know, what's the issue in finding activation barrier? Well, you know the initial state, and you know the final state.

This is not necessarily the state of an atom. You really should think of these as two states in phase space. Now, for simplicity, you can think, well, only one atom moves, is different, between those two states.

But in reality, more stuff is different. Because if the atom has moved, the other atoms may have relaxed around it. So these are really states in phase space.

So what you have to find is the path between the two with the lowest energy maximum along that path. That's essentially the activated state. And a very practical way to find is what's called a nudged elastic band model, which is essentially that you do a simulation on many systems at a time. And all the systems live at intermediate states between the final state and the initial state.

And the way you keep them there is interesting. Rather than optimizing the energy of each system, what you optimize is the sum of their energies, the sum of their Hamiltonian values, plus some spring constant, basically plus some energy penalty, which is associated with the difference in coordinates between the two systems.

And what that tends to do is actually, since this is a quadratic in the difference in coordinates, it tends to spread out the systems along the path between the two states. I mean, think of it. If I got rid of this, so if I set k to 0, then I'm going to minimize this.

So they're all going to go down here. But what do you actually do is you hold one here. You hold one there. And then because of the harmonic potential between them, you tend to end up with systems that live in between. So it's a very elegant way of finding minimum energy paths between states.

It's called the elastic band method or the nudged elastic band method. And there's a whole bunch of variants of it as well. It's the perfect parallelization. You may have noticed.

If you want to do, say, 10 replicas-- so you want to do 10 points along the path-- you run on 10 processors. Because it's the same system, they pretty much run about the same amount of time. They never have to talk to each other. They have to talk to each other on extremely rare occasion when they exchange coordinates. That's it. So it's a perfect parallelization tool.

If you do that in lithium cobalt oxide, you have a complication. Just to show you how the data for that kinetic Monte Carlo simulation was derived, the yellow is the lithium here. These are oxygens, the big red balls.

The way you can diffuse from this side to that side is straight through the oxygen bond. And if you do that, this is obtained by elastic band method. So these are the replicas. So this is initial stage. This is the final state.

And so you get an activation barrier of about 0.85. So it's very nice how you get these activation curves. But the problem is that there's two paths in this material.

So already you're starting to run into a particular problem of kinetic Monte Carlo. Let's say you thought that was the path. Well, you'd be very wrong because there's another path. And that's one when the outcome actually, rather than going straight through this bond, it takes a curve into this tetrahedron and comes back out.

And that's actually a path, if you do elastic band, that is so much lower. It's only about 200 millielectron volts. And it sort of shows the real problem you can run into that, when you do kinetic Monte Carlo, you have to assume what your paths are for transitions.

And if you don't get them right, you can be seriously wrong. And that's one of the nice things of MD, that it's unbiased. So it would find these things for you.

But these are actually the only two paths. And of course, because of the activation barrier, this is so much lower. This is pretty much the one that gets executed almost all the time.

The difference between 0.2 electron volts and 0.8 electron volts at room temperature is amazing. It's, I think, 5, 6 orders of magnitude in rate constant. You just have to calculate the exponential of each of these.

One of the problems you have whenever you have two rate constants is that, when you do your Monte Carlo, you can only scale the fastest one away. So you have essentially one scale factor in the time of your Monte Carlo simulation. So when you do multiple time scales, when you have multiple time scales, it gets very inefficient to do kinetic Monte Carlo.

But this works. You can get diffusion constants. This is with elastic band method. This is the activation barrier for that low energy mechanism as a function of lithium concentration, so very dependent on concentration.

So you can do your kinetic Monte Carlo. And then you have these, and you can go all the way to macroscopic stimulation. Then you can solve a fixed equation and actually look at diffusion on the scale of microns. So it's a fairly simplistic way of course graining.

OK. So I'm going to end just giving you some reference. If you want to read more about the quasi-continuum method, there's these great papers by Miller and Tadmor. These are actually somewhat similar, but the 202 version is an update.

These are not the original papers, but they're actually sometimes more pedagogical than the original papers. But you'll find the references in there to the original papers. The problem of sort of dynamics at that scale is my own work.

The accelerated MD, there's a series of papers by Art Voter. But this annual review of materials research is, I think, a great review if you want to read about it. And the lithium cobalt application is in a few papers, but it's quite elaborate. The whole idea of kinetic Monte Carlo and parameterizing activation barriers is quite elaborately discussed in this paper.

So I think we're done a little early, but that's fine. So I'm going to end here. And I'll see you on Thursday for the lecture from Ford.