

**NICOLA  
MARZARI:**

To calculate ensemble averages-- that is to calculate thermodynamical properties. And we have seen also a few movies on how we can calculate reaction barriers and how we can follow in real time chemical kinetics. The last application of molecular dynamics is likely more different.

In particular, one can use a final temperature, classic in molecular dynamics, as an optimization scheme. In general, suppose that you have a complex potential energy surface. This is obviously in one dimension, but you could think this as being in a multi-dimensional space.

And actually, your potential energy function could represent a complex cost function. Again, it could be, say, the problem of optimizing the usage of your planes on several routes for one airline company. And obviously, depending on where you put your planes, you have a certain cost.

And so this cost function tends to be very complex depending on your coordinates, how you arrange your planes. And finding either the global minimum or sort of one of the lowest minima can be a very complex affair. And one of the powerful techniques that has been introduced in the early '80s is something called a simulated annealing.

That is basically a technique that basis itself on a thermodynamical analogy. That is suppose that you want to find this minimum. And if you were to use a deterministic recipe, it would be very difficult. You would start from a certain point and then go down according to the gradient until you end up in what would be a local minima.

And so what, in particular, Kilpatrick, Gelatt, and Vecchi introduced was a thermodynamical technique in which you really populate your phase space with what we call walkers. So these are sort of dynamical systems. So those are the points that represent those dynamical systems. You give them some coordinates.

That is they are going to have some potential energy. And then you give them some temperature. So they are going to have some kinetic energy. You can think this as a swarm of skiers sort of going around your phase space.

And if you have a lot of them and if these have a lot of temperature, they will move all over phase space. But by the time you start sort of very slowly cooling this dynamical systems down, you start removing temperature for them, they are going to condensate really in whichever local minima they find. So some of them may end up, say, here. And you cool them down, and slowly they find themselves here.

But if you cool them sufficiently slow-- and really in order to make this into an exact schema, you have to pull them almost infinitely slowly. So it's never going to be an exact approach, but just a stochastic approach. But, you know, with some luck, you'll find that some of them start condensing in interesting minima. And then you choose basically your lowest minima possible.

And this is actually a very sort of practical and useful approach. And so it's sort of fairly widely used in optimization problems. OK, so this sort of concludes the sort of set of applications of molecular dynamics.

Now, I wanted to sort of hint at some of the advanced statistical mechanics techniques that you'll actually see later in the class that, generally speaking, go under the umbrella name of Green-Kubo techniques. And sort of I'm taking one of the simplest examples in which, actually, Albert Einstein was involved. So it's an interesting approach.

Well, this has to do on starting a macroscopic property that is really diffusion in a solid. So you could think of a silicon crystal. And we are doping that silicon with gallium or phosphorus, and so you are interested in studying how the impurities in this system move around.

And so if you want, what you have is that you have a perturbation. You put a sort of gallium on the surface of silicon, and then this diffuses into the solid. And sort of the general Green-Kubo approach is a formalism that relates some of these macroscopic properties, like a diffusion coefficient, a transpose property, to microscopic properties. And you'll see that as sort of fluctuations of the equilibrium distribution.

OK. So suppose that we are looking at the diffusion of, again, something like gallium and silicon. What you would define as your sort of fundamental variable is the concentration of impurities. So that's a sort of four-dimensional function is a concentration function of the [INAUDIBLE] space you are interested and function of time. And then, obviously, there is sort of the diffusion law for this that says that basically the current is really proportional to the gradient of the concentration.

And then you put this together with a sort of continuity equation that tells you that the derivative of the concentration with respect to time plus the divergence of your current-- that is your flux-- needs to be constant. That equals-- let's say it's going to be equal to 0. I mean, this very simply says that, if an infinitesimal volume the concentration changes, it's that because there is a current going out of that infinitesimal volume. And so the divergence of the current measures this.

And with fixed diffusion law, we can put the 2 and 2 together and basically obtain this relation, this differential equation, for the concentration profile in time. So if you want, these are starting macroscopic relationship. And then what becomes interesting is the derivation that Einstein did in order to recover the connection with the microscopic dynamics.

And this is just a little bit of algebra, which I will go over fairly quickly. But, basically, what we are doing here, we are actually multiplying the left term and the right-hand terms by  $r$  square and integrating it over all space. And once we do that, well, what we have is that on the right-hand side, this requires some calculus.

So I won't go really into the derivation. But by sort of integration by parts, you can actually show that this integral provided, say, the concentration is normalized to 1 when integrated in all space, what we really obtain is just 2 times the dimensionality of your problem. That is, if you are in one dimension, sort of small  $d$  is going to be equal to 1. If you are in two dimensions, it's going to be equal to 2, to 3, and so on and so forth.

So this integral really just gives us the dimension of the space. And again, we won't go into how we prove this. What we have on the left-hand side, on the contrary, is really, you see, the average value of  $r$  square.

What we mean by this bracket that you see is the integral over all space of  $r$  square with the probability. If you want the concentration, it's identical to the probability of having a certain particle in a certain position in a certain time. So what this integral means is really the average value over all space of  $r$  square, how much your particles has moved.

And we have the derivative and the time. And so you see we start seeing somehow a connection between a diffusion coefficient that's the macroscopic property and the microscopic properties, the dynamic piece, how much the particles are moving once you put them in somewhere.

Let me actually give you an example of what would this quantity look like and how we would calculate it in a molecular dynamic simulation. Again, you can think of your gallium atoms being at the surface. At the beginning, you put everything in motion. And they are going to move with a sort of random walk towards the inside.

And if you want to calculate this sort of average value of what is called the mean square displacement, the average value of  $r^2$ , well, you need really to do the average over all your particle at a certain time of  $\Delta r^2$ .  $\Delta r^2$  is really how far all these particles have moved.

And you know, this would look at something like this in a typical, say, solid. Actually, the example that I have here is slightly more exotic, but it was a sort of good example to show you the difference between a liquid and a solid. And it's the case of silver iodide that is, in particular, an ionic crystal in which the blue iodine atoms sit on a BCC lattice.

And for each iodine, there is a silver atom. But above, say, 420 Kelvin, this system is in a very peculiar state of matter, what is called a super ionic state in which this iodine atom oscillates around an equilibrium position exactly as atoms do in a solid, while the silver atoms sort of move around exactly like a liquid. So it's a system that is really solid. It's crystalline. But one sublattice, the silver sublattice, is sort of moving around as a liquid.

And so if you calculate what are the average mean square displacements-- that is if you average, say, all the iodine atoms, what is their average  $\Delta r^2$ , how much they move around-- well, you'll discover that really, in a sense, that they are a crystalline solid. They are just going to oscillate around their equilibrium position. And so you calculate that average.

During a molecular dynamic simulation, you see there is a sort of first very short period of thermalization. The atoms start moving, but then they are just oscillating around. They are not going anywhere. So their instantaneous mean squared displacement as a value that is just a constant.

So this is sort of typical of a crystalline solid. If you do the same thing for the silvers, you see that, as a function of increase in temperature, you see, at 250 Kelvin below the superionic transition, also the silvers sort of look more or less flatter. But when you increase the temperature, their mean square displacement starts becoming a linear function of time.

So the silvers are really moving around. They are diffusing away as a liquid. And so in the mean square displacement, you can see the signature of a liquid when they increase linearly with time. Or you can see the signature of a solid where they just are not going anywhere. They just vibrate around the equilibrium position.

And so this is the quantity that we could calculate in a molecular dynamic simulation. The average of the mean square displacement, I have written it here. Remember, this is going to be equal to 2 times the diffusion coefficient times the dimensionality of your system.

And now, we sort of work a little bit with this. And in particular, we introduce this sort of alternative definition of the displacement of an atom. This is actually very useful.

If you are sort of working in periodic boundary conditions, you need to be careful. Often, your molecular dynamics codes, for simplicity, always show you the coordinates of the ions as they were sitting in your first unit cells. If an atom is diffusing as a liquid and it sort of moves around and from one cell moves in the second unit cell, your code usually will actually sort of bring it back by a lattice translation in the first unit cell.

So a universal way to sort of take into account what the actual position of an atom is is actually write the position that is usually a vector that satisfies this periodic boundary condition as an integral of the velocity. So in that sense, the position of your particle is always taking correctly into account all the distance that has been traversed.

So if you have a unit cell really and your particle is moving, you'll sort of integrate correctly your trajectory by doing the integral of the velocity, while sort of your algorithm would actually bring back the particle once it crosses a boundary. OK. Nothing deep here, we are just sort of writing the position as the integral of the velocity.

But with this definition, we can actually go back and look at what is the expression for the average of the mean square displacement. That is the  $\Delta x^2$ . And now, we write it actually as the square of the integral of the velocity at the instant of  $t$ .

And so I'm just sort of writing this integral out explicitly when I sort of expand the square. I have an integral in  $t$  and  $t'$ . And again, the bracket means just the ensemble average on all your particles, so it commutes with the integral.

And so what we are calculating here is really the sum over all the particles and averaging that. That is dividing by the total number of particles of the velocity at certain instant  $t'$  times the velocity of a certain instant  $t$ . And we can actually, for computational convenience, this is really an integral on a square, integral from 0 to  $t$  in one dimension and the integral from 0 to  $t$  in the other dimension.

But this expression is symmetrical. So we can actually integrate this only on half the square. We are integrating it only on the triangle.

If this is were  $t$  and  $t'$  and our integration interval in this, we can just sort, for simplicity, integrate into half of this. OK. So what we have achieved here is we have written this average mean squared displacement as an integral of the average product between the velocity at a certain instant and the velocity at a different instant. And this is sort of where our connection with the equilibrium properties is starting to emerge.

And you'll see this in a moment. This is called velocity autocorrelation function. And I'm writing it more explicitly.

All of this has been done, again to keep the algebra simple, in two dimensions. So remember, the Einstein relation. Sorry, we are in one dimension. So remember, the Einstein relation. The small  $d$ , the dimensionality of your system, is 1.

So we have the 2 times the diffusion coefficient is equal to the derivative of the mean square displacement. And mean square displacement in itself has been written as the double integral. But when you take the derivative with respect to  $t$ , one of the integral cancels out.

So if you want, this is our final relation. Let me actually sort of write it over here and remove the factor of 2. We have the diffusion coefficient as written as this integral. We can sort of exploit the translational invariance in time.

Really, if we are looking at what is the average value of the product of the velocity at a certain instant,  $d_1$  prime, times the velocity of another instant  $d_2$  prime, well, that average product is not going to be different if we translate it in time. So we can refer it to an arbitrary origin. And so we just do a translation in which we shift  $d_2$  prime into 0 and so on.

So this is sort of our final expression. And this is what is called a velocity-velocity autocorrelation function. So you see, what we have is that the macroscopic property, the diffusion coefficient, that is really telling you how a system responded to a perturbation.

You put some gallium on your surface of silicon, and then there is this macroscopic diffusion transfer process. That is there is a perturbation, and the system evolves. It can actually be related to an equilibrium property of the system. That is really, ultimately, what are the mean square displacements, how atom microscopically move around.

And in particular, the quantity that is playing here is this velocity-velocity autocorrelation. You see, what this is looking at is suppose that you have a velocity at a certain instant. And then you look at that instant that is sort of  $\tau$  away in time from sort of your instant 0.

And you look at the product of these two velocities. Now, if  $\tau$  is very small, the velocity is not going to have changed very much. So in the limit of small  $\tau$ , your velocity at 0 and your velocity at  $\tau$  are going to be very similar.

So if you want, there is a lot of correlation. When you take this product, it's going to look a lot like  $v_0$  squared. And then all these things are normalized properly. So it could look like 1. But as you sort of move away in time, you go towards longer and longer times, there is going to be no correlation.

Your velocity at 0 and your velocity at an instant that is very far away in time is not going to be correlated. And so the product of this can be, if you want, any number. And when you take the average, you see the sort of ensemble average, the bracket.

When you average this quantity on all your particles in the system, this is just going to be 0 because this thing can have any positive or negative value. Because there is no correlation between the velocity at the large  $\tau$ . And so this thing is going to be 0.

OK. So the limit of this velocity-velocity autocorrelation for very large  $\tau$  is going to be equal to 0. The limit for sort of very small  $\tau$  is going to be equal to 1 or whatever your normalization factor is. And then there should be some interesting structure at a certain time.

And we'll see it in a moment that suppose your system is actually sort of oscillating. You are looking not at the liquid part of your silver iodide, but you are looking just at the sort of crystalline iodine sort of oscillating around. Well, suppose that your iodine atom has a certain period of oscillation. So it sort of keeps going back and forth.

Well, then there will be a lot of correlation if you look at your velocity at a time 0 and your velocity at an instant that is roughly equal to a period of oscillation. Because if this thing were to oscillate exactly around this equilibrium, every sort of period you would have that the velocity has become the same. OK. So you will see a very definite structure in this correlation function.

If your system is not a crystal, if your system is sort of diffusing away as a liquid, as your time increases, the correlation becomes 0. And even if it's a crystal, but it's not really oscillating perfectly, like a harmonic oscillator around the period, when you really got to very, very long time away, you start losing this correlation that sort of takes place at every beta. And that average quantity starts in itself to be 0.

So how does our velocity-velocity autocorrelation function looks like? Well, it looks like something like this. I need to sort of graph it here.

So again, when you are really sort of far away in time-- so, you know, what we are plotting is the ensemble average of  $v(t) \cdot v(0)$ . So very, very far away in time, it's going to be 0. There is no correlation between the velocity of the atoms.

At very, very small times, there is really maximum correlation. And then what is in between really depends on the dynamics of the system. And it can look very different in different systems.

But Einstein relation and the Green-Kubo formula that we have seen before actually relates, first of all, the integral from 0 to infinity of the expansion to the diffusion coefficient. So again, these functions really represent a [INAUDIBLE] microscopic fluctuation at equilibrium, sort of how the velocities are correlated, so how your vibrate around.

But all the sort of algebra we have seem before shows also that the integral of this quantity gives you the diffusion coefficient [INAUDIBLE]. And that's one way of calculating the diffusion coefficient. The other way, if you go back to your slides, would be just calculating the derivative with respect to time of the mean squared displacement.

And so in a liquid, you have seen that your mean square displacements sort of become linear if you sort of weighed enough time with respect to time. And so the diffusion coefficient is also equivalently given by the slope of this system. One of the interesting things that you obtain from the sort of velocity-velocity autocorrelation function that you don't obtain from the slope of your mean square displacement is that you can obtain what we call the power spectrum.

That is you can look at a system like a liquid and figure out what are the typical vibrations in your system. Again, this is because, you know, what we have said. If a system, if an atom, is oscillating around, say, an equilibrium position, there is going to be a lot of correlation.

That is, if we look at time 0 and if you look after a period  $t$ , the velocities are, again, going to be very similar. If you look again at time  $2t$ , they are, again, going to be similar. If you look at the time  $3t$ , they are, again, going to be similar.

So in the velocity-velocity autocorrelation function, actually you will see not only a peak at 0, but a peak at  $t$ , at  $2t$ , and  $3t$  and so on, and then sort of slowly decay. But so that means that the velocity-velocity autocorrelation function will sort of show somehow some periodic features that are related to what is the typical dynamics of your particles.

And so if you do the Fourier transform of that, the Fourier transform of a function, it picks up what are the relevant frequencies in that function. And so if you do that, you actually find out that what is the vibrational density of states, what are the typical frequencies of your system. And this is sort of an example that comes from a molecular dynamics simulation of water.

So what you have is you have your liquid water. Remember, that in a system like liquid water, you actually need only 30, 40, 50 molecules in a unit cell to basically simulate the infinite system as long as you are sort of enough far away from the melting or freezing transition. So you take the system, you let it evolve.

And then you calculate at every instant in time  $t$  the product of the velocity. You average that on all the molecules, and you have the velocity-velocity autocorrelation function. And that, as a function of time, will show typical beatings that have to do with the frequencies in your system.

You do the Fourier transform of this, and you find a power spectrum. You find a vibrational spectrum of your liquid in which you see you can identify a very large peak, a very sort of typical set of vibrations that have really to do-- this is actually heavy water, so it's got deuterium-- with the stretching mode of the deuterium oxygen distance in the water molecule, so the modes in which the two atoms vibrate one against the other. So this would be the optical intermolecular modes.

And then you see another peak that has to do with the vibration modes, the fact that the water molecules act as a scissors. And then you find a lot of lower energy modes. But again, just the Fourier transform of this velocity-velocity autocorrelation function gives you right away what is the vibrational density of states and, again, sort of a snapshot of what are the important vibrations in the system.

And a lot of spectroscopic properties would be correlated with this vibrational density of states. Because suppose that some of these states are, say, what we call infrared active. That is, when the atoms move around, they create a little polarization. They create a little local electric field.

Well, then those modes would interact and couple very strongly with sort of electromagnetic radiation. And so depending on your frequency of your electromagnetic radiation, you would couple very strongly with the appropriate frequencies of your liquid system. And so this is, again, one of the sort of very important quantities that you might want to extract from a molecular dynamics simulation.

And if, instead of having a liquid, you had a solid, again, that would be the sort of vibrational density of states of your phonon modes. OK. This sort of concludes this parenthesis on Green-Kubo approach. You'll see more of this in one of the later classes.

But again, if there is one thing that you need to remember, it's this, that this general sort of set of relation make a connection between a macroscopic property and, in particular, a response property of the system. The diffusion coefficient is a response property of the system to a concentration inhomogeneity. And these macroscopic property, these response properties, can be connected to equilibrium fluctuations or something like the velocity-velocity autocorrelation function.

And there are a lot of interesting quantities that you can obtain from Green-Kubo relations besides the diffusion coefficient. So in particular, you could find out the viscosity of your system. Suppose that you want to study liquid iron because you want to understand how seismic waves propagate in this sort of liquid inner core of Earth.

Well, you are very interested in the viscosity of the system because you want to know how sort of shear propagation takes place. And just from the fluctuations in the stress tensor, if you have your unit cell with all the iron liquid atoms moving around, the fluctuations instantaneous due to the thermal motion in the stress tensor for that system give you the sheer viscosity.

Or say you want to study how, say, one of the systems like water would couple and what would be its infrared absorption. Well, you can actually look at the instantaneous fluctuations in the total polarization in the system. Because what is the microscopic local electric field there and how it couples?

So again, you can sort of find out macroscopic properties from microscopic fluctuation or sort of electrical thermal conductivities. Again, sort of macroscopic transfer properties can be found out from fluctuations in the autocorrelation functions for the sort of electrical charge or a thermal carriers in a system.

OK. This basically concludes the classical molecular dynamic part. And what I wanted to show you next is how we actually do first principle molecular dynamics. That is how we sort of evolve atoms in time not using a classic field, but using our favorite electronic structure methods.

That, actually for most of this class, has been density functional theory, but doesn't really have to be density functional theory. Any of the electronic structure methods would work. Density functional theory tends to be the simplest and most efficient to implement.

Sadly, in order to do this, I need to give you some other reminders of formal classical mechanics. Because especially in first principle molecular dynamics, we use a lot of the concept of extended Lagrangian and extended Hamiltonians. That is we'll derive the equation of motion from an appropriate functional that includes sometimes very exotic degrees of freedom.

So let me remind you how you have seen this in some of your physics or mechanics class. But let me show you how, in general, one can think at evolution in phase space and sort of find out the equations that integrated the trajectory. And up to now, we have really just seen Newton equation of motion, forces equal to the mass times acceleration.

But there is a sort of more complex and, if you want, more elegant formalism to derive the equation of motion for a system. And in particular, what I'm showing here is sort of what is called Lagrangian dynamics. In fact, this is actually very, very simple. And the way Lagrangian dynamics works is this.

First of all, you have to construct your Lagrangian. That is the functional that drives the evolution of your systems. And there are various ways in which one constructs these functionals. And there are even sort of equivalent ways. One can construct different Lagrangians that give you the same equation of motion of the same trajectories.

But, see, what is important here is the standard way. That is the way we construct this functional, it's not very different from thermodynamics. We just take the kinetic energy of the system,  $T$ . We subtract the potential energy of the system. And this is sort of our Lagrangian,  $T$  minus  $V$ .

In general, the potential energy, as you have seen it, tends to be a function of position only. So usually, we written as a function of, if we have  $n$  particles,  $r_1$  to  $r_n$ . So you know, this is what is called a conservative field.

If you are sort of a particle living in a gravitational field, well, you are going to be in a certain position. You are going to feel a certain potential, or you are going to feel a certain force. That's the gradient of that potential.

You go somewhere else, you'll feel a different potential. You see a different force. And the work that you sort of make in going from one place to the other, it's just the integral of that force. And it's independent of the trajectory.

So this is sort of, you know, a very general sort of potential function that you have seen. And you know, again, the kinetic energy, you have seen it and tends to be a function of the square of the velocities. OK.

So again, if you have only one particle, its kinetic energy is going to be  $1/2$  times the mass times the square velocity. We usually, in the Lagrangian formulation, don't use the sort of notation for the positions  $r_1, r_n$ . But instead, say in particular, we use the other notation in which the coordinates are given by  $q_1, q_2, q_n$ . And then the velocities we just indicate them as  $\dot{q}$ .

And the reason we call them  $q$  is that what you sometimes want to do is not use your regular coordinates as the description of your position for your dynamical system, but you might want to use generalized coordinates. Say you study water molecules. And all of a sudden you want to describe this liquid of water molecules as rigid molecules.

So you want to say that the angle between the hydrogen, oxygen, and the hydrogen doesn't change. And if you want to say that the distance between the hydrogen and oxygen doesn't change, well, then you want to sort of develop a dynamic in which what you really move around are not the position of the atoms, but you move around the center of mass of your water molecules. And you move around the orientation.

This is actually very important. If you remember in sort of last class, I've told you that, when we study water, actually because water at regular temperature is still a quantum system, is still a system that has most of its vibrational states frozen in the sort of 0 point motion quantum state, you actually tend to describe better liquid water if you describe it as a set of rigid molecules moving around.

This is, again, an approximation. But it's actually a better approximation of the true dynamics of the system than an approximation which you let also the internal degrees of freedom change. So suppose you want to sort of simulate the rigid water around.

You need to find out what are the equation of motion for this generalized set of coordinates in which what you really move around when you move water is the center of mass and their orientation. And you know, it would be very difficult to do sort of using Newton's equation of motion with a constraint. And so what you do, you use your Lagrangian formulation in a generalized formalism of generalized coordinates  $q$  and generalized velocity  $\dot{q}$ .

So what Lagrangian dynamics tell us is that we construct our Lagrangian function  $T$  minus  $V$ . And then the equation of motion are given by these. These are the Lagrangian equation. We want to derive them, but this is how they are written.

The derivative with respect to time of the partial derivatives of the Lagrangian with respect to the generalized velocity  $\dot{q}$  minus the partial derivatives of the Lagrangian with respect to the generalized coordinates needs to be equal to 0. And you know, this also focuses us on just constructing the two scalar function, kinetic energy and potential energy, for a system. And that is just, you know, straightforward algebra to derive this equation.

And I've actually done the derivation for the sort of simple case of Newtonian dynamics. You actually see how trivially the Lagrange equation that is written here turns into Newton's equation of motion when you plug in for your Lagrangian  $\frac{1}{2} m \dot{x}^2$  minus your potential energy. And you see, when you take the partial derivatives of the Lagrangian with respect to the generalized velocity, since it is a partial derivative, you only need to take the derivative of the kinetic energy with respect to the velocity, with respect to  $\dot{x}$ .

And then you need to take partial derivatives with respect to the position. And there is no position in the kinetic energy functional. So there is only the position in that potential energy.

So what you see is that, from the left-hand term, you have the derivative with respect to time of mass times velocity. And so that's nothing else than mass times acceleration. And then on the right, the right-hand term, you have minus the gradient of your conservative potential field. And so this is nothing else than the force.

So we have force equal to mass times acceleration. And we have recovered. Just by applying Lagrange equations to the Lagrangian of classical dynamics, kinetics minus potential, we have derived Newton's equation of motion.

So this is one way of deriving equation of motion, but, again, are going to be second-order differential equation with respect to time. Because, if you think, you are taking the derivative of a kinetic energy with respect to  $\dot{q}$ . There is a second formulation of classical mechanics-- and I also need to sort of present this here-- that is called the Hamiltonian formulation.

And it's sort of, again, very easy to see this if you think of a thermodynamic analogy. Say, when you're sort of looking at thermodynamics, suppose that you are in the microcanonical ensemble. Your thermodynamical functional is the energy.

And so when you sort of look at that thermodynamical ensemble, it means that, say, you are looking at the system that has a constant energy, constant number of particle, and constant volume. But many times it becomes appropriate to look at, actually, sort of a system that lives at constant pressure or constant temperature. And so you transform your thermodynamical functional from the energy to one, say, of the Helmholtz or Gibbs free energies.

You do  $e - Tds$  to obtain a thermodynamic functional that depends on temperature instead of depending on entropy. Or you do  $e + pv$  to obtain the enthalpy that sort of depends on pressure and not on volume. And this is this general concept of Legendre transformation.

If you have a function, let's say, for the Lagrangian was a functional of  $q$  and  $\dot{q}$ , you can construct a new one that doesn't depend, say, on  $\dot{q}$ , but depends only on a new variable that we call it the conjugate variable  $2q$ . So pressure and volume, temperature and entropy, chemical potential and number of particles are all conjugate variables.

And so, say, if you take the Lagrangian and you derive it with respect to  $\dot{q}$ -- remember, the Lagrangian is a function of  $\dot{q}$ -- what you obtain is a conjugate variable that we call a conjugate momentum. And then if you do this operation, this Legendre transform in which you take-- the sign doesn't really matter. In this case, you take a minus sign.

And you sum the product of the conjugate variable times your original variable. You get a new function that doesn't depend on your original  $\dot{q}$  variable, but depends only on its conjugate variable  $p$ . That's how you remove the dependence, say, on the volume and you put in the dependence on the pressure, making the Legendre transformation which you have  $pV$ , or that's how you sort of move on from entropy to temperature in the Helmholtz free energy.

And this is just very simple to do when you take the differential of this quantity. Because you are going to see that really, because of this relation, you remove all the dependency in  $\dot{q}$ . And you put into your system an independence on  $p$ .

Well, why do we do this? Well, because, again with some algebra, we can find that, for this new function that is now called the Hamiltonian, another alternative sets of equation of motions. Remember, from the Lagrangian, we had obtained equation of motion basically for  $\dot{q}$  and  $q$ .

And from the Hamiltonian formulation, if we work this out, we obtain equation of motion for  $q$  and for  $p$ . And sort of the slight difference is that instead of having basically a second order differential equation from the Lagrangian formulation, now we have a double set of differential that are only first-order. So depending on the your problem, they can actually be easier to solve. They can actually be different, but they could lead to the same trajectories.

So again, all these formalisms of Lagrangian and Hamiltonian is that's the very general way to construct functions, either the Lagrangian  $T$  minus  $V$  or the Hamiltonian via the Legendre transform that give us equation of motion in  $q$  and  $\dot{q}$  for the case of the Lagrangian and in  $q$  and  $p$  for the case of the Hamiltonian. And if you actually work this out for the standard case of Newtonian dynamics, you find out that your Hamiltonian is, again, something very trivial.

It's just the kinetic energy plus the potential energy. So at the end of all of this, for all the cases of interest to you, you see that either you construct a Lagrangian  $T$  minus  $V$  and you have the Lagrangian equation of motion, or you construct your Hamiltonian,  $p$  plus  $v$ . And you have the sort of Hamiltonian equation of motion that I have written here.

And one of the reasons that we do this-- so this is-- is that, often, we want actually to simulate the microscopic dynamics not in the microcanonical ensemble, but in different thermodynamical ensembles, say, in which we control the temperature. Or we control the pressure, or maybe we control the number of particles. Or maybe we control the chemical potential.

And so we need to have some of this sort of formalism to do this effectively. In particular, remember when we have discussed the sort of control of temperature. I've told you that there are sort of three different approaches in which you can do a canonical simulation, which you can control the temperature in your system.

And you could have Langevin dynamics. You could have a stochastic approach in which you randomly pick atoms in order to accelerate them or to slow them down. So that in analogy with a thermal bath, they sort of, on average, have the right kinetic target, kinetic energy for your problem.

Or you could do something that is probably even more brutal, although it tends to be very efficient in thermalizing, effectively, your system. You could actually do a dynamics in which, every time you have got new position, you renormalize those new position by renormalizing the velocity of the particle, so that the sum of the kinetic energy is a actually constant. So a constraint method would actually keep, strictly speaking, the temperature of your system sort of constant to your target value.

That can be actually very effective to thermalize your system, to bring it really very close to your equilibrium distribution. But it does have some counter-effects. That is, if you actually look at what is going to be your equilibrium distribution in positions, it's really going to sort of be a canonical distribution according to the Boltzmann canonical ensemble.

But if you look at your distribution of velocity, it's only pseudo-canonical. So often, people, especially those that do very complex and long molecular dynamics, use the sort of most elegant and most accurate approach. That is really coupling your system to an additional dynamical variable using an extended Lagrangian or an extended Hamiltonian.

So we have written our Lagrangian or Hamiltonian in terms of generalized coordinates. All of a sudden, you can add one more generalized coordinate. So you can add, if you want, a pseudo-particle in your system with its own sort of kinetic energy and with its own potential energy.

And you can construct the kinetic energy. And in particular, you can construct the potential energy, so that this additional dynamical system, this additional dynamical variable, interacts with the other dynamical variables, basically exchanging temperature with it to bring the sort of average temperature of the real classical particles to the equilibrium distribution.

And so this is actually how you would write the extended Lagrangian for the case of a canonical simulation, something in which we want to keep the temperature constant. And so you see that's where the power of all this generalized system comes into play. And again, you sort of write out your kinetic energy.

And note that there is this sort of  $s$  squared term that couples the kinetic energy of your real particles with the kinetic energy of this new thermostat, as we call it. So we have kinetic energy minus potential energy. And then this is the new particle. This is the new extended coordinate that is described by a generalized position  $s$  and the generalized velocity  $\dot{s}$ .

And so, if you want, its kinetic energy is written in a trivial way,  $1/2$  our generalized mass times the square velocity. And this is actually how Nosé figured out the potential energy should look like for a system of  $n$  classical particles that you want to keep at an inverse temperature  $\beta$ . Remember,  $\beta$  is just  $1$  over the Boltzmann constant times the temperature.

And so it's a very exotic potential energy. And it's sort of the logarithm of this generalized coordinate. And so if you use these Lagrangian, you can construct the Lagrangian equation of motion.

So you will have equation of motion for your  $3N$  particle that will be very similar to your standard equation of motion, but we'll have in there also terms that contain  $s$ . And then you will have a sort of an equation of motion for your Nosé or Nosé-Hoover variable that is equation of motion for  $s$ . And you let all this system evolve.

And what will happen is that the kinetic energies of your classical particles will start to sort of distribute themselves according to a Maxwell-Boltzmann distribution. That is according to the canonical ensemble. And people have worked out all the statistical mechanics of your problem and so worked out the appropriate coordination.

You would need to have a sort of Maxwell-Boltzmann distribution. This is in the moduli of your system. And really, if you do all of this and you sort of take an average from your molecular dynamic simulation, you really see that both the position and the velocity of your classical particles are distributed in the appropriate way.

So it's very sort of useful to calculate, say, response suppose property of the system. Say, again, if you want to calculate the diffusion coefficient from the velocity-velocity autocorrelation function, you really need to have your velocity distributed according to the appropriate thermodynamical ensemble. So it's sort of the, if you want, most accurate and careful way of doing the dynamics.

Nosé-Hoover, so it's probably the most common way of thermostating your problem. It's not the most robust. It's the approach that gives you the long time thermodynamical properties correctly. But it tends to be very poorly, say, for a system that is very harmonic.

OK. A system is very harmonic when its potential energy is really a quadratic function of its coordinate. So if you have, say, a single particle that sits in a parabolic well, what we call a harmonic oscillator also from a standard pendulum, that is a perfect harmonic system. A solid at very, very low temperature is also very harmonic.

If you want the potential energy of each particle, it's just a quadratic function of the [INAUDIBLE] displacement from its equilibrium position. And so what are the trajectories in a harmonic oscillator? Well, they look something like this.

If you look now, if you go back to a one-dimensional position and momentum representation, well, you know, a pendulum does really this over and over again. So this would be the trajectory moving in time. So the position sort of oscillates back and forth.

And the momentum or the velocity oscillate back and forth. And they are in a positional phase. When the elongation is maximum, your momentum is linear and so on. I actually sort of made, I guess, incorrect axis. But thanks to the power of the [INAUDIBLE] tablet, now this looks much better.

OK. Now, suppose that you sort of studying not a single oscillator, but you are studying a solid at very low temperature. What your Nosé-Hoover thermostat would give you is something that is very different from the equilibrium distribution. So the sort of equilibrium canonical distribution is going to look something in which there's of momenta and velocity of the particle are distributed.

Because, now, there is obviously, if we have a real solid, a small amount of interaction. So each atom, each harmonic oscillator, is going to be talking with its neighbors. And so they are not going to be all in phase and perpetually in phase, but they are going to sort of exchange energy between one and the other.

And so a large harmonic solid, but slightly sort of different from 0 temperature, will have a distribution of velocity and momentum that doesn't sit in a perfect circle, but is sort of distributed around. If you do this with a Nosé-Hoover thermostat, sadly there will be no exchange of energy. There will be no thermalization between all different atoms.

They are going to be doing perpetually their own thing in synchrony. And so the Nosé-Hoover thermostat works very poorly for systems. The more harmonic your system is, the poorer the equilibration that comes from the Nosé thermostat is.

And so you have to pay a lot of attention to system at low temperature. And sort of one of the usual solutions that people mention is using Nosé-Hoover chains in which you actually have your dynamical system. You have your thermostat.

And then you have another thermostat, thermostating your thermostat to thermostate for your particle. And then you have a third thermostat that thermostates the second thermostat. And suddenly, you have to do this.

Harmonic solids are actually very important. So there is a reason why a lot of effort has been put into this. Because in particular-- and this is, again, something that we'll see in one of the last classes. In order to calculate the free energy of a system at finite temperature, often you need to do an integration from 0 temperature to the present temperature.

And so you really need to start from a harmonic solid that you need to describe correctly. And if you don't do that so if you don't use your proper thermostating techniques, this is not going to work out. And there is really a lot of dark [INAUDIBLE] in all of this.

You really need to be careful because the temperature is a global quantity. But if you have, say, a system with 100 particles, you can have a temperature of 300 Kelvin by having sort of all particles distributed along the same temperature. But you can have also the same temperature if part of your system is very cold and part of your system is very hot.

And if you wanted, the thermostat has this role of making sure that the energy flow goes around, so that all part of your system are really at the equilibrium. And you know, this can be actually a trickier thing than sort of it actually seems. Obviously, all these dots and [INAUDIBLE].

OK. These are sort of a summary of all the books in which you can find a description of these concepts. If you have to read one for this class-- the sort of free primer that Furio Ercolessi has put on his website. I'm actually very proud of this because Furio moved recently to this very small town in northern Italy that is sort of my hometown. But anyhow, this is available and free.

It's 30 or 40 pages. I posted it also on the Stella website. And it gives you a very clean introduction to molecular dynamics.

And then I think this is a field in which there are, in particular, two exemplary books. So if you really want to do this for a living, there are a couple of books that are very, very good-- the Allen and Tildesley book that, by now, it's very old. It's sort of really written at the beginning of the '80s when you can imagine computer capabilities were very different.

And it's actually remarkable somehow how the clarity that you develop by dealing with systems that are very primitive from the computational point of view actually develops your intellectual skills. So this is really very, very good. And equally good is a much more recent book that is in its third edition from last year, the Frenkel and Smit that also has many more advanced topics. But either of these are really sort of very, very good references.

And now, in the last 20 minutes, I'll give you really the flavor on how we do first principle molecular dynamics. That is really nothing else than a classical molecular dynamics with a lot of additional variables. So we use this concept of extended Hamiltonians.

And you know, somehow it's, by now, a little bit old. I had sort of once downloaded this. This is actually the citations of first principle molecular dynamics, of initial molecular dynamics or the citation of the first paper that developed this concept that goes under the name of the [INAUDIBLE] techniques that is from 1985.

And you see we are sort of very pleased with this exponential explosion. It means that basically there is a lot of useless literature out there that you can look at. So in order to do first principle molecular dynamics, I need to give you a reminder of your favorite topic. That is the expansion in plane waves of the total energy of a system.

And hopefully, you still remind something about this. But if you have a crystal that's going to be defined by the three primitive direct lattice vectors-- so you could have an FCC crystal and they point to  $1/2, 1/2, 0, 0, 0$ , and the third one. And then we can define reciprocal space and the reciprocal lattice whose three primitive vectors, this  $g$  vectors, are just defined as the duals of your direct lattice vector. So the scalar total between the reciprocals and the direct lattice vectors needs to be equal to 0 or  $2\pi$ .

OK. So this is the definition. Once we have this reciprocal lattice vectors that, for simplicity, I'll draw here as being two-dimensional and square, what we have is, again, sort of these points here represents all possible integer combination of reciprocal lattice vectors. That, for my problem at hand here, say, this would be the Brillouin zone. This would be the unit cell of the reciprocal lattice.

And so all these  $g$  vectors in blue, remember, are such that the complex function  $e$  to the  $iG$  times  $r$ , where  $G$  is any of these vectors represented by one of these point, this function in real space is compatible with your periodic boundary condition. It's going to have either one oscillation inside your unit cell, 2 oscillation, 3 oscillation, 10 oscillations. And it's going to be, generally speaking, in three dimensions.

And so if you remember, we expand our wave function into linear combination of these plane waves because these plane waves are all the possible-- it's a complete set of functions that describe orbitals that have the periodicity of the direct lattice.

OK. So what was our quantum mechanical system? We keep it simple. We won't go into density functional theory. We'll just think again at the operator.

You see the Hamiltonian raises its head again. It's just going to be the quantum kinetic energy minus  $1/2$  the Laplacian plus the potential energy and with the caveat that we develop, we expand, the periodic part of your wave function, what we call sort of, via the Bloch theorem, the periodic. Sometimes I call it  $u$ .

But what is the periodic component? We expand it. We write it as a linear combination of plane waves with appropriate coefficients.

So if you want, your quantum mechanical problem is, yet again, nothing else than trying to find that these numbers. Once you have defined your basis sector, then you have written an expansion. And then all your algebra and all your computational problem is finding out what this coefficient should be.

And this tends to be a very good basis sector because it can be made more and more accurate that by including  $G$  vectors of larger and larger models that correspond to plane waves of higher and higher resolution. Not only that, but it's actually a very manageable basis sector to use. Because it's very easy to take first derivatives, second derivatives of a wave function of square moduli.

Because, say, if we take a first derivative of  $\psi$  with respect to  $r$ , the only thing that we have is that each term in  $G$  gets down from here a factor  $i$  times  $G$ . And the second derivative gives us a factor  $iG$  scalar product  $iG$ . That is minus  $G$  squared.

OK. So our quantum mechanical problem basis for our first principle molecular dynamics has really nothing else to do than finding what is the ground state energy. That, not really in a density functional formalism, but in a sort of simplified formalism, I write out just as a sum of the eigenvalues of our single particle orbitals. That is just the sum of all the electrons of this expectation value.

If you remember, in density functional theory, the energy is slightly different because the Hamiltonian in that approach has actually become self-consistent. The Hamiltonian depends on the charge density itself. And so there are some additional collective terms.

But you know, from the point of view of our sort of [INAUDIBLE] here, it doesn't really matter that we include these additional terms. And so, you know, what is that we have to deal with here? Well, we have, again, a quantum kinetic energy term and a potential energy term.

And so the quantum kinetic energy is going to be the sum of the kinetic energy of each orbital, as written here. And you know what I was saying before. It's actually trivial to calculate the kinetic energy if you have a wave function that is expanded in plane waves because this is it.

This is the expectation value. Remember, it's the integral over all your Hilbert space. That is the integral in space of this.

This is the [INAUDIBLE]. So this is the complex conjugate of the plane waves  $e$  to the  $iGr$ . So it's  $e$  to the minus  $iGr$  times the action of this operator to the plane wave  $e$  to the  $iG$  prime  $r$ .

And so, you know, this second derivative, so with respect to  $r$ , gives us an  $iG$  prime times  $iG$  prime that is a minus  $G$  square. The minus sign cancels out. And so we have  $1/2$   $g$  square times the integral with respect to space of  $e$  to the minus  $iGr$  times  $e$  to the  $iG$  prime  $r$ .

And actually, if  $G$  is equal to  $G$  prime, that is just going to be 1. And so the integral, it's actually-- I sort of skipped all the normalization. So the integral is going to be 1. You can work out the algebra. If  $G$  is different from  $G$  prime, it's going to be 0.

So this expectation value of the kinetic energy operator between two plane waves is just going to be  $1/2 G$  square between two identical plane waves. And it's going to be 0 if the plane waves are different. This is what we say when we say that the kinetic energy operator in a matrix representation in which the rows and the columns are the different plane waves is actually diagonal.

That is, if you calculate this, so you would have, say,  $G$  here and  $G$  prime here, all the terms in this matrix function of  $G$  and  $G$  prime has 0 outside the diagonal. And the diagonal is  $1/2 G$  square, so very trivial to do in plane waves. If we want to calculate the second term in the energy, we need to calculate the potential energy.

And the potential energy is going to be, again, sort of the sum over all the single particle terms in the potential energy. And I'm doing the same thing here. I'm writing this expectation value between two plane waves.

And so I have same as before, the exponential of the complex conjugate minus  $iGr$  times  $e$  of  $r$  times  $e$  to the  $iG$  prime  $r$ . And if you look at this, this is nothing else than the definition of the Fourier transform coefficient of your potential energy with respect to the wave vector  $G$  minus  $G$  prime. So now, this part, the potential energy in a plane wave basis sector, is not anymore diagonal.

It's going to have off-diagonal terms. That is, if you look in the previous sort of matrix representation in  $G$  and  $G$  prime, what we have along the diagonal is just the kinetic energy terms  $T$  that are  $1/2 G$  squared. But we are going to have a number of diagonal terms and also on-diagonal diagonal terms from the potential.

Obviously, the more you go outside the diagonal, the more  $G$  minus  $G$  prime is going to be large. So the more you're going to look at high frequency components of your potential. And in general, unless your potential is really ill-defined, the high frequency components are going to be smaller and smaller.

So again, sort of the overall Hamiltonian matrix kinetic plus potential in a plane wave basis tends to be diagonally dominant. As we move farther and farther away, it tends to be smaller and smaller. So we have actually written all the terms. And then we can take the sum over all these expectation value.

And so we have here sort of a [INAUDIBLE], if you want, the total energy as a function of this coefficient of the plane waves. So we have the kinetic energy term and the potential energy term. And the fundamental approach that we take in first principle calculation is not trying to diagonalize this Hamiltonian to find eigenvalues and eigenvectors because that operation would scale as the cube of the number of plane waves.

And it gives us not only the 10 or 20 lowest eigenvalues that have the occupied states. But, say, if we have a simple system like the silicon atom-- remember, you have seen semiconductors in your lab two, where, at the end, what you need is really the four lowest eigenvalues and the four lowest eigenvectors. Well, if you diagonalize this Hamiltonian and you have plane wave basis set with 1,000 elements, you get all 1,000 eigenvalues and eigenvectors.

We don't need that. It's too expensive. We need only 4 out of 1,000. OK. So the approach that most or all electronic structure approaches take is actually an approach in which we find what are these coefficients of the relevant occupied orbitals.

You see, here, the total energy, that's the function of the occupied orbital. So for silicon, we would sum only on the four lowest orbitals. And by minimizing this quantity, we find what are the coefficients.

So the electronic structure problem, again, it's a minimum problem in which you have an expression for the energy that, in density functional theory, is slightly more complex, but at the end depends only on the coefficients of your plane wave expansion and on the matrix elements of the potential between plane waves and of the kinetic energy between plane waves. You calculate these ones for all or, in density functional theory, at every step. Because your potential, being self-consistent, changes up every time step.

But if you have this and you have this, it's nothing else than a problem of finding the minimum with respect to these variables. Sadly, you have sort of thousands or tens of thousands or hundreds of thousands of them. So it tends to be a very expansive problem of minimization of a non-linear function that depends on basically zillions of variables. So how do we do this?

Well, let's see. We use a molecular dynamics idea. That is we have an energy, a function of thousands of variables. So we have multi-dimensional space with really thousands of coordinates.

And as a function of those thousands of coordinates, these  $c$  and  $G$  variables, what we have is a total energy surface. So our problem looks nothing else like this. We have a total energy that is this sort of hyper surface in black that is non-linear functional of these things.

In density functional theory, for a standard LDA, GGA calculation, this system tends to have only one minimum. If you are doing a spin polarized calculation, you will have different minima that corresponds to a non-magnetic solution or all the magnetic solution with different set of values for the magnetization. Or maybe have anti-ferromagnetic solution with a total magnetization of 0, but different spins and so on and so forth.

So this potential energy surface can start to develop some of the complex minima that I was showing you when we were talking about simulated annealing. And that's why actually finding the ground state of magnetic system tend to become, very quickly, much more complex. In LDA or GGA, it has only one minimum. And we need to find that minimum as a function of this very many coordinates.

And so we use molecular dynamics techniques and molecular dynamic analogies. And you know, I'll do this in one dimension, but really what we want to do is find an equation of motion, something that evolves our coefficients,  $c$ ,  $G$ , so that they move towards the minimum of that potential energy surface. And in order to do that, we have nothing else to do than to calculate the derivative of that Bloch potential energy surface with respect to each and every variable, with respect to each and every plane wave coefficient.

Again-- sort of lots of algebra that I'm hitting in there. But actually, and very pleasantly, it's actually very simple. And we sort of write this generalized force. That is the derivative with the minus signs, the gradient of the energy with respect to each and every coefficient. And it's, by chance and nothing else than, the application of the Hamiltonian itself to the wave function coefficient by coefficient. And so it's actually sort of very easy to calculate.

And then once we have that, well, so if we look at this problem in one dimension, what we have is our Bloch potential energy surface. And we need to find the minimum. So this is the energy as a function of all this coefficient of plane waves.

And we can choose a random value for this coefficient, calculate the energy. And what we will have is a value for this energy. And now, what we want is really to evolve, to move the value of each and every coefficient. So that the total energy of our system reaches the minimum.

And in order to do that, we need the gradient that we have written in the previous expression that is nothing than minus the Hamiltonian applied to the wave function. That is the force. And it's a perfect dynamical analogy.

If you are somewhere and you know what is the force that drives you downhill, well, then you have molecular dynamics techniques to get to the bottom of your potential energy surface. And this is particularly easy if, as in LDA or GGA, your potential energy surface has only one minimum. And there are actually two different approaches that you could use.

You could use standard molecular dynamics. So if you are here, you put yourself in motion. And what you start doing is going down.

But in a pure conservative molecular dynamics, you are going to overshoot your minimum. Go back and oscillate perennially. So what you do, you actually put some friction, so that you sort of slowly lose energy. And you basically, again, condense down to your sort of lowest energy solution.

And so a standard molecular dynamics technique applied on the coefficient of the plane waves with a friction term added is one approach that brings all these additional degrees of freedom of plane waves down to the ground state. This would be one approach. The other approach would be actually evolving your system in a way in which the velocity is proportional to the force.

And so, again, you can think that, when you are at the minimum, your force is 0. And so your velocity is 0. And so if instead of accelerating your system proportionally to the force and putting a friction that is something that will oscillate and dump down to the minimum, you move with a velocity that is proportional to your force.

It's something that, again, it will sort of move you towards the ground state. Once you are here, your driving force is going to be 0. So your velocity is going to be 0.

So somehow it's a different approach that also brings you to the 0. And sort of the two approaches have a slightly different performance depending on the kind of system that you are doing. In this case, in the sort of dynamical way, you need to make sure that your friction is good enough.

That is you don't use too much friction, so you sort of flow down and never reach the minimum. And you have to be sure that there is sort of enough friction, so you don't oscillate forever. And that's one way in this sort of approach that you need to make sure that your system, again, doesn't really sort of stop short of getting to the minimum.

Because the closer you get to the minimum, the closer your velocity is to 0. And so you tend to just get very close, but not right to there. And there are advanced techniques to deal with both of these systems.

And with this, I think I'll actually conclude here because I wanted to sort of tell you more about how we actually evolve in time this problem when we start moving the atoms. Up to now, we were sort of thinking at just how to get out at the ground state solution. But then we have the additional problem of having the atoms move, so our potential starts changing in real time.

And we need to sort of follow in real time what happens to the solution. And again, there are sort of two general approaches that go under the name of sort of Born-Oppenheimer molecular dynamics or sort of Car-Parrinello molecular dynamics. And I think we'll leave that at this stage for one of the next lectures.

So let me just sort of remind you that another of the terms that you'll see for this kind of dynamics in which the velocity drives the system towards the minimum goes actually under the name of steepest descent or conjugate gradient minimization. And that's, again, a sort of technique that you see very often in minimization problem where sort of this is really a molecular dynamic kind of simulated [INAUDIBLE]. With this, I conclude.

We'll keep some of the sort of ab initio molecular dynamic techniques for one of the last lectures when we actually go back and look at case studies of this. And in the last transparencies, in the last few graph of your notes, you have, again, some freely available bibliography of quantum molecular dynamics primers, in particular, the sort of primer of Dominik Marx on his website. The last note is something very completely and very well written. With this, I conclude. Have a very happy weekend, and we'll see you on Tuesday in the lab in 115.