PROFESSOR 1:   --labs based on electronic structure, and this will cover the next two computational labs. And on Thursday, we'll start looking at some of the finite temperature ideas that are based on the energy models that we have seen.

This, by the way, is [? Roberto ?] [? Bajo ?] missing a penalty shot. So a reminder of what we have seen in the last class. We have introduced the idea of pseudopotentials. In order to remove the cost of carrying out calculation that include the core electrons that are very many, especially in sort of larger atoms, and that have exceedingly high oscillations around the nucleus due to the orthogonality constraint. what we have done is we have substituted what would be the z over r Coulomb potential with-- inside of the core region of the nucleus-- a pseudopotential. That is a potential that reproduces the effect of the nucleus, and of the frozen core electrons.

And as you can see, this pseudopotential tends to be repulsive close to the origin, basically again, reproducing this sort of angular momentum push outwards of the core electrons to the valence electrons. And in order to make this pseudopotential very accurate, there has been this idea that has been developed of norm conserving pseudopotential that will act differently on the different components of a valence electron wave function.

So you have an incoming electron, so you have a sort of ground state valence wave function. You can decompose it. You can decide how much of it is SP or the component, and you can act differently on the different slices of this wave function.

And so with this, basically, we can solve a new problem in an effective external potential in which the lowest energy ground state will actually be identical in energy to the valence eigenstate of the original so-called all electron on the atom with all its electrons. And the wave function for this pseudo-atom will be identical to the wave function of the real atom outside of the core radius-- outside, in this case, three atomic units in this slide.

So this was sort of one of the first important technical proofs that were introduced in the '70s and '80s to make this calculation really feasible. The other sort of idea that I want to sort of remind you is that whenever we deal with extended systems-- solids, liquids, and in particular, when we have periodic boundary condition-- that is we have our unit cell periodically repeated in all dimensions-- the eigenstates of our Hamiltonian take the form of Bloch Theorem, and get classified according to two quantum numbers. A discrete number-- the bounding that's number n-- and the continuous number k.

And the overall eigenstates can be written as a product of two functions. One is just a plane wave that has the periodicity of the so-called crystal momentum K, one of these quantum numbers. And the other is the periodic part of this Bloch orbital written here as u. And that this is a function with the same periodicity of your unit cell.

So overall, the orbital itself-- the [INAUDIBLE] orbital psi can have any periodicity, but it can always be decomposed into a part of a well-defined wavelength times a periodic part. Periodic part is going to be smooth. Again, we don't have any more core electrons, so close to the nuclei, it will look just as 1s, 2s, 2p orbitals. And that periodic part can actually be expanded in a set of plane waves-- plane waves that need to satisfy the periodic boundary condition of our system.

And so we have seen that we can write the functions e to the iGr, where the G is a linear combination of primitive reciprocal lattice vectors with integer coefficients, and all those G vectors are such that e to the iGr has the same periodicity of your direct lattice. And these are the coefficient of the series expansion.

And we sort of can systematically improve the quality of this basic set expansion by increasing the number of G vectors that we use, and we can do that systematically by taking G vectors with longer and longer lengths, or as we say, with higher and higher energy cutoff, because higher and higher energy [INAUDIBLE] longer and longer moduli for the G vectors corresponds to plane waves with finer and finer periodicity, so with finer and finer resolution.

So this is a distinct advantage of plane waves. And in addition, this basis sets do not depend on atomic position. So when we calculate forces, say, to do molecular dynamics or a structural relaxation calculation, we don't need to take into account the fact that the basis set changes with the position of the atoms.

Plane waves are not the only choices. There are a number of other choices that are very successful. In particular, a lot of the quantum chemistry could use a Gaussian basis sets. So atom-centered orbitals that decay as Gaussians. This is very convenient to do calculations like the exchange term in Hartree-Fock.

But of course, you could have just a finite difference representation. So you could represent your orbitals on a grid of points in real space. That tends to be very efficient to do parallel calculation, but it's very difficult to do accurate calculation of the second derivative.

If you remember, when you expand a function in plane waves, you have an analytic expression for the second derivative just because the first and second derivative of plane waves are just ig or minus g squared times the plane wave itself. And that's very important. So that's when sort of real space representation becomes a little bit trickier.

And then there are a number of sort of more approximate approaches based on a number of sort of atomic-like localized orbitals. Or there are a number of sort of accurate approaches that are based usually on a combination of a basic set that has a plane wave-like characteristic in the regions that are far away from the nuclei, and then it has atomic like characteristics in the region inside the nuclei. And this tends to be the most accurate, but also slightly more expensive approaches.

This concludes one of the technical points that you need to be careful in a practical calculation. The other point that sort of comes over and over again has to do with Brillouin Zone integration. Let me give you a first example of a molecule.

So if you are calculating the electronic structure of a molecule, in the course of your calculation, you'll need to calculate integrated quantity like the charge density. The charge density of a molecule is going to be the sum of the square moduli of all the single particle orbitals.

So if you are sort of studying a molecule like just hydrogen 2, well, you just take the sum of the square moduli of the first orbital-- you have more orbital, more complex molecule. You need to sum over all the orbitals up to the [INAUDIBLE]-- the highest occupied molecular orbital.

You do a calculation in a solid, you have to do exactly the same thing. That is, you need to calculate the charge density by summing over the occupied states.

Now as I said, what are the quantum numbers that describe a solid are a [INAUDIBLE] and a continuous index scale that we call the quasi-momentum. And we usually represent, say, the energy of the states in a solid with a band diagram. And I've plotted here on the left the band diagram for silicon.

In particular, what we are looking here is in the Brillouin Zone of silicon along certain high symmetry direction. I presume these are the high symmetry points, like gamma would be 0, 0, 0. [? Alpha ?] would be 1/2, 1/2, 1/2.

Hope this is correct. This should be x, and it is 1, 0, 0. So we plot the energies of all the occupied bands of silicon in different directions along the Brillouin Zone.

Silicon has two atoms per unit cell. Each atom has four valence electrons, so we have eight valence electrons per unit cell. It's a system in which there is basically spin degeneracy.

So there is really sort of the same spatial part for spin up and spin down. So what we usually say is that we can accommodate two electrons for each space orbital, and they will just have different spin quantum number.

So with these eight valence electrons, we end up with four valence. And sometimes, because of symmetry, you have degeneracy, but you see that somewhere like here at a sort of arbitrary low symmetry point in the Brillouin Zone, you can clearly see four bands.

So if we want to calculate the charge density of the system, we need to sum over all the possible occupied states. That is, we need to sum over all the four bands, and that's trivial. But we also need to integrate over all the possible k vectors. That is, we need to make an integral in the Brillouin Zone.

That is, we need to sort of really sum over all these possible states. An integral is obviously an analytical operation. In practice, on a computer, what we do is just we discretize that integral, and we take a sum.

So it means that-- and again, I'll use two dimension as an example. It means that if this were my Brillouin Zone-- two-dimensional Brillouin Zone, and my k vector can be anywhere in this Brillouin Zone, what I would need to do is an integral overall the possibilities inside there.

But in practice, what I'll do, I'll just take a discretization. And I'll calculate my band structure at each one of these points, and say a regular aqueous space dimension of k points.

And that's an expensive operation. So each calculation at each k point will require a self-consistent diagonalization of your problem. And so really, the cost of your calculation is linearly scaling in the number of k points. So in reality, you want to use as few k points as possible.

For a system like a semiconductor or an insulator, in which if you want-- the band gap structure is very smooth-- it becomes fairly easy to integrate that smooth bands with very coarse measures. So for something like silicon, you might be already very happy when you start to sample the three-dimensional Brillouin Zone with a uniform mesh of k points that could have four k points in each direction-- 4 by 4, by 4, 6 by 6, by 6, 8 by 8, by 8. And this is sort of the order of magnitude of the k points that you need to use.

If you were to study for a moment unit cell of silicon, you might want to study, say, vacancy information energy like you were doing in your first computational lab. Then you are not going to use two atom units. You are going to use a larger unit cell.

And without sort of dwelling that much on it, suppose that, again, in two dimensions, you double the size of your unit cell. What's really happening is that the reciprocal lattice vectors will become one half in length. So remember this general concept. When you sort of increase the size of the unit cell, the reciprocal cell becomes smaller.

Suppose that we have doubled-- we have gone from two atoms in each-- two atoms in our unit cell to-- well, if we are in two dimensions, eight atoms in the unit cell. We have doubled in real space-- in reciprocal space. Our Brillouin Zone becomes 4 times smaller.

And so if we want to keep the same quality of integration, actually, it means that our k point sampling needs only to be one fourth the number of points that we had before. That is we still use the blue points that are included in this sort of smaller blackish unit cell.

So that also means that, say, if 64 k points is a good number for a regular set of unit cell of two atoms of silicon, if now you are going to use a, say, unit cell that has maybe 128 atoms-- much larger to calculate the vacancy formation energy-- you really need only one k point to calculate your total energy with the same accuracy. So there is this general idea of scaling and folding. You make your real space calculation larger. You actually need to use fewer k points.

And if you are familiar with sort of some of the solid state ideas, that also just means that when you double the size of your unit cell, you are really refolding some of this band structure in a smaller space. So you actually sum over a number of bands that increases.

The situation is slightly more complex for something like a metal. The sort of fundamental difference when you deal with a metal as opposed to a semiconductor or an insulator is that there is something called a Fermi energy. That is now there isn't any more a gap, so the total charge density is again given by sum of the states, sum over all the bands, and integral over all the k space.

But really, the total charge density-- so the ultimate integral that would be the number of electrons depends really on where we stop. And there isn't any more sort of natural separation between empty and occupied states. So there is going to be in a metal an energy level that determines what is occupied and what is empty, and that's the Fermi energy.

So for copper I think has 11 electrons per unit cell sort of FCC metal with one atom per-- one atom in each unit cell. What we really need to do in our electrons calculation is find what is this energy level. That is find what is this Fermi energy, such that the integral over all the state below that level gives us the right charge density-- in particular, the right number of electrons.

And this is sort of one of the difficulties that come out in metals, because now what we are trying to do is we are trying to integrate the bands below a level-- below the black line. And so you really introduce a discontinuity in your integral.

And so to calculate integrals of discontinuous function usually requires a much finer accuracy in your k point sampling. And so the calculation tends to become much more expensive. Nothing else. So the general solution to this problem besides using a larger number of k points, and particular accuracy to this sampling issue is that of introducing what is called a finite electron temperature.

So what is actually done in sort of every practical calculation is introduce a small amount of temperature so that in reality, there's a sharp discontinuity become smoother. Because when you have an electron temperature, states above the Fermi energy can be slightly occupied with the Fermi [INAUDIBLE] occupation, and states just below the Fermi energy can be slightly empty.

So to summarize, I mean, we need to be careful in sort of sampling. And that's the other sort of fundamental parameter of your calculation. If you are dealing with a metal, you need to be particularly careful, and you need to use electron temperature techniques.

If you are using sort of-- if you are studying an insulator a semiconductor, what you usually do is just use a regular equi-spaced mesh of k points. And that sort of known in the community with a technical name. They are called [INAUDIBLE] for some very good reason. Actually, there are some actually fairly beautiful symmetry thoughts on why equi-spaced coarse mesh can work very well.

But in practice, it's nothing else than choosing the blue points inside of the green Brillouin zone. If you are studying a really large system, you can actually reduce yourself to sampling only one point in the Brillouin Zone. Brillouin Zone has become so small that just taking sort of-- how do you call it?

A mean-- the theorem of the mean. That is, you can substitute the integral with the value of the function at one point. And usually, you have two choices. You can just use the gamma point. This means that 0, 0, 0-- that has a computational advantage.

When you sample things at the gamma point, you can choose your functions to be real instead of complex. And so you have right away your computational cost. Or you can choose sort of what could be the best single point for your given symmetry that's sometimes called the Baldereschi point. And that can be again sort of useful if you need to do an accurate calculation in a large scale system.

Once you have set these two fundamental parameters, you are really ready to do actually a practical calculation. And so as I said, we have a self-consistent Hamiltonian, and so what you need to do is you need to iterate your problem until the eigenstates that you find give you a charge density that is identical to the charge density that you have done before.

So in practice, how would your electronic structure code work? Well, first you would tell your code where the atoms are. So you need to specify the position of the atom. Suppose that you are studying silicon, you could sort of-- since you know actually what is the structure of silicon, you could already put them in the origin, and in the position one fourth, one fourth, one fourth.

And then you need to specify in particular which flavor of non-local pseudopotential you are going to use. That is, there will be a library of pseudopotential that basically represents a silicon atom with all the core electrons frozen. And there are sort of a number of technicalities that you'll see in the lab on which one you should choose, but they are sort of more or less all the same, at least from this point of view.

Once the code knows where the atoms are-- that is, knows the position of the atoms inside the unit cell, knows what is the shape of the unit cell, and what is the length of the dotted lattice vectors, this infinite array of atoms, the infinite crystalline, or amorphous, or disordered extended system is set. And what we really need to do is throw the electron [? seeds ?] and let the electron find their own ground state.

And so we need to make sure that we have the right basis set cutoff. That is, we are going to describe the orbitals [? anchored. ?] We have the right sampling. And at this point, we can sort of start the self-consistent procedure.

And in the sort of simplest form, well, we first need to figure out what is our Hamiltonian operator. The Hamiltonian, remember, depends on the charge density itself, because some of the terms in the Hamiltonian, like the Hartree energy, the Hartree potential of the exchange correlation potential depends on the density.

So we need to pick an initial guess for a trial charge density. It could just be a superposition of atomic charge density. Once we have the charge density, we can construct the Hamiltonian, the [INAUDIBLE] Hamiltonian. The kinetic energy operator is-- we always know it.

But we can construct this Hartree and exchange correlation terms that depend on the charge density, and then we'll have the external potential that is given by this array of non-local pseudopotential. At this point, we have the Hamiltonian, and we try to find the lowest energy eigenstates for Hamiltonian.

In particular, we just need to calculate a number of states that is equal to the number of occupied orbitals if we are dealing with a semiconductor, or it's equal to the number of sort of electrons plus 20%, 30% in a metal to make sure that at different points in the Brillouin Zone, we calculate all the bends that could be below our Fermi energy.

So we solve this. And this is really the expensive step-- and I mean very expensive step in any electronic structure calculation. And there are sort of a number of ways of diagonalizing a matrix, of solving this eigenstate equation.

In a basic set, it tends to be very large. When you do sort of a realistic calculation, even for silicon, you could have hundreds of plane waves, So. Hundreds of basic set elements. And large scale calculation like you would doing in research would contain tens of thousands of plane waves.

And actually, you don't-- you can't really diagonalize on a regular computer even a matrix that has [INAUDIBLE]. If you think a matrix that has a dimension of 1,000 requires one million elements. And one number-- a complex number requires 16 bytes. So just a matrix that has 1,000 sides that will require 16 megabytes to be described, and this number explodes quadratically very quickly.

So you can construct the full Hamiltonian, and you don't want to calculate-- if you have a matrix of dimension 1,000, it will have 1,000 eigenstates. But you only care, say, if you are studying silicon, on the lowest four eigenstates. So you want to have numerical techniques that calculate for you only the lowest energy eigenstates. And there are a number of them that are well-established.

So now you have obtained with one of these techniques your lowest energy eigenstates, and you sum their square moduli to obtain the new charge density. And with this, you go back to the step. You construct the Hamiltonian operator again, you diagonalize new density, and iterate.

Of course, naturally, a recipe like this would most likely never converge. So one needs to develop a mixing approach-- mixing approaches that sort make the change in the charge density at every iterative step smoother than what I've described.

So if you calculate a new charge density and you diagonalize it again, your second new charge density will be even more different from anything that you have obtained before. So what you really do is you need to find some schemes to evolve in a very smooth way your charge density.

So maybe once you have calculated the new sum of eigenstates, you don't take that as huge as density, but you just update your old charge density with 10% of what you have calculated now to try to make the iteration to self-consistent very smooth. And I have to say a lot of the know-how in electronic structure calculation in the '90s has really gone into trying to find sort of mixing approaches that evolve our charge density to solve consistency, and that converge under a large variety of circumstances for large or complex systems, especially for metals.

So stripped to the bare elements, an electronic structure code really needs to do two things. It needs to diagonalize inexpensively a Hamiltonian that expressed on a plane wave basis is a very large order.

And I just mentioned the names of some of the [? algorithm. ?] Things like the [INAUDIBLE], and the [INAUDIBLE], or some of the conjugate gradient algorithms are all algorithms that give us reasonable cost the lowest energy eigenstates of that Hamiltonian.

And then once you have that eigenstates, you need to calculate charge density, and you need to have a mixing strategy. You need to have a strategy to evolve your charge density towards self-consistency. And that is also a very tricky approach.

There is a sort of completely different approach to the problem that sees the ground state solution not as a self-consistent iteration, but as a nonlinear direct minimization of the functional. If you remember, we have the energy function-- or I think it's written here in the next slide. No, it's not.

We have sort of written the density functional of theory energy functional, and it is a well-defined expression of the orbitals only. It will be in one of the-- so the following slides will pick it up again. And so we can also see the problem as the problem of minimization of that functional in a space that is very large, because the sort of variables that we really deal with are the coefficients of our plane wave expansion.

But in principle-- and I'll show you that in a moment-- we can actually write out a minimization algorithm. The advantage of this approach, if it's done properly, it has always a solution. If you keep minimizing your energy, at the end, you will get to a global or to a local minimum.

So this approach tends to sort of converge under every circumstances if done properly, and done properly is not trivial. But then sort of the efficiency of the different things is really system dependent.

And I guess without wanting to bore you with sort of math-- so just I wanted to remind you again what happens in our computer. That is what happens when we say we want to solve this eigenstate equation. Supposing let's say we are in a self-consistent diagonalization approach.

And as always, you have to remember we expand our wave function in a well-defined set of orbitals-- that is our basis set. I represented here them as phi. And it could be plane waves, it could be atomic orbitals. We use plane waves all the time.

So really, in our computer, our unknowns are these coefficients of this basis set expansion. And so our eigenstate equation, once we multiply on the left by phi and star, and integrate, is really a matrix problem. So I've written it here as just the same eigenfunction equation written over there.

And if we call hmn at the matrix element of the Hamiltonian between, say, two plane waves of different wavelengths, this is what our problem means. It's just a linear algebra problem. We need to find the eigenvalues for which there is a possible solution, and the possible solution will be eigenstates. And an eigenstate is nothing else than an appropriate set of coefficients that satisfy this equation, and those coefficients put back in here will give us actually what are the full eigenstates of our problem.

So this was the sort of self-consistent diagonalization. As I just said a moment ago, we can also look at the problem as a nonlinear minimization problem. Once we have decided on an approximation for our exchange correlation functional, could be a local density approximation, could be a generalized gradient approximation.

This is a well-defined quantity in which, again, the external potential is given by this array of non-local pseudopotential. And the Hartree energy is just the functional of the charge density. And the charge density itself is thus the sum of the square modulus of the orbitals.

So in reality, this energy is a function of that psi, or in other terms, is nothing else than a very complex function of those c1 to cn coefficients of each eigenvector. So this is nothing else than a minimization problem, again, on a number of variables that can be 1,000 if you are studying two atoms of silicon. And it can be in the tens of hundreds of thousands if you start to do really serious calculations.

So again, it's a fairly complex problem. A huge number of variables that you need to deal with, and a nonlinear expression for the energy. But again, in principle, if we have this explicit expression for the energy, E of psi, where the psi that we consider around only the occupied orbitals, what we can do is nothing else than take the functional derivative with respect to that psi. I'll consider them real here just to avoid sort of complex conjugate numbers.

And at the end, this is nothing else than calculates the derivative of the energy with respect to all the coefficients, say, for i that goes 1 up to the cutoff, all the coefficients of all the occupied orbitals. So you see, the larger your system becomes, the more basis elements you'll need to use.

I mean, if you double the size of your system, if you look at the math, you'll actually need the twice the number of plane waves, as it makes sense, to describe a charge density or a wave function with the same resolution. So you double the size of your system. The number of plane waves to describe a single particle orbital becomes twice as large, but now you will also have twice as many occupied orbitals. So just the number of this coefficient has become four times as large.

So again, this calculation becomes very expensive very quickly. But again, this is well-defined. I mean, we can just actually write explicitly the nonlinear function of the previous slide in terms of the coefficients of the plane waves. This is actually done in one of the article posted. There is a review of modern physics by Mike Payne and coworkers-- among others, John [? Gianopulos ?] at MIT. And they actually work out the algebra of all these derivatives.

And then once you have the derivatives, you have the gradients, and you know how to move along and go to the minimum following the gradients. The only difference with a sort of regular minimization problem is that this is a constraint problem. That is what we have because these are really electrons, and not just sort of arbitrary functions.

The electrons need to be meaningful quantum states, so they need to be orthonormal. So this derivatives with respect to the c need actually to take place on the hyper surface where these conditions are satisfied. That is, if you were to evolve the coefficients of the plane waves, what you would find is that as soon as you have sort of changed them by a little amount, your orbitals per se are not going to be orthonormal anymore.

So this constraint is sort of fundamental, and this is what ultimately limits sort of-- or determines the computational costs of our calculation. Because again, if we double the size of the system, we'll have twice as many plane waves, and we'll have twice as many occupied orbitals. So we have already a cost of fours. But those occupied orbitals will need to be orthogonal to each other.

And so the number of these matrix elements that you need to calculate has become also twice-- or the number of orbitals in this matrix have become twice, so the number of matrix elements has become four times as large. You see, we double the size of the system, we'll have twice as many orbitals here, twice as many orbitals here. And this integral is going to take place on a region in space that's twice as large.

So 2 by 2, by 2 gives us a factor of 8, and so gives us the, ultimately, cubic scaling of cost of density functional calculation. You go from two atoms of silicon to four atoms of silicon, your calculation has become eight times more expensive.

Hartree-Fock in its original formulation scales as the fourth power. Other quantum chemistry approach scales as the fifth, sixth, or seventh power. So it becomes very easy to sort of reach, really, the limit of calculation that you can do on a regular computer or even on a regular supercomputer.

And there is a lot of effort to develop what are called linear scaling approaches. That these electronic structure algorithm that scale linearly as the size of the system. And somehow, they are all based on the idea that sort of physics or quantum mechanics is local.

So if your orbital at the end is ultimately localized in a certain region of space, it will be automatically orthogonal to orbitals that are very far away. Because if this is localized somewhere here, and this psi i is localized somewhere there, their overlap will be 0 by definition, and so we don't need to worry about orthogonality. So somehow, locality of physics, locality of quantum mechanics in principle tells us that there are linear scaling approaches that could work, although, none of them have really made into sort of production electronic structure at this stage, although there is a lot of ongoing effort in many groups.

So with this [INAUDIBLE] we conclude also all the technicalities. And what we'll do in the rest of the class will give you a sort of panorama of what typical applications of density function of theory calculation are going to do.

And I'll go very quickly over this. I have here sort of cases in which we could be interested in structural excitations. So when you start warming up a system, a molecule, or a solid, you start exciting the different normal modes of your molecule or of a solid. I've shown you some of the possibilities for something like a carbon nanotube, because that band-- we have a banding mode up there.

We have a pinching mode, and we have a breathing mode. And so these are all the possible structural excitations, and you can actually calculate this structural excitations using density functional theory. And I've given here a comparison between the case of diamond-- what is calculated with the sort of state of the art [INAUDIBLE] code like you are going to see in your laboratory, and what is measured actually with the neutron scattering, the red dots. So you see without really any input parameters.

And once you have really phonon dispersion, you can calculate all the thermodynamics of solids. That is really basically based on the statistics of excitation of this vibrational degrees of freedom. So you could calculate, say, how your elastic constant for your bulk models changes with temperature, and this is the calculated black line, and you could compare it with experiments.

Or you could take one of your slabs like you have seen in the first laboratory in which you were calculating the surface energy, and you could actually put it in motion. You could follow at a given time temperature the dynamics of the atoms.

And you want to have a slab thick enough so the atoms in the middle really act as bulk atoms. They don't see the presence of the surface, and then it becomes very easy to investigate what's happening on the surface.

And so you can have sort of a snapshot of sort of how the atoms are moving on the outer layers, what are the sort of typical displacements of typical mean square displacements. Or you can say study how the distance between the layers evolved with temperature. So you can look at, say, how, as you increase the temperature of your slab, the distance between the surface layer and the second layer changes with temperature. And this would be the computation, and here we have the experimental value.

And there are sort of lots of interesting physics that takes place, again, if you look at the distance between the second and third layer, if the red line above and the system expands there. And sort of with computation, you can actually probe your system deeper and deeper, where experiment starts to become very difficult. It's almost impossible to look at what the fourth layer in a surface does, and what the fifth layer of a surface does.

I think in order to keep the balance of the lecture, I'll actually switch on to Professor [? Sidor's ?] part so he can show you some of the other application. And if we have time either in one of the next classes, I'll show you some of the other applications that we have mentioned. So I think I'll pass over the lecture to Professor [? Sidor. ?]

PROFESSOR 2: So Professor [? Mazari ?] already gave you some generic applications. What I want to do for the rest of the class is actually give you some numbers, and really look seriously at what the typical accuracies are that you can expect from what we'll call density functional theory, but what we mean with that is sort of standard density function theory as we kind of explain it to you in the class in the local density or generalized gradient approximation.

So these are sort of staples of electronic structure methods now. But that doesn't mean that there are, in some cases, not already better forms out there, but they're often very much in the research stage, and there wouldn't be things that you would either easily do on real problems. You wouldn't get your hands on them, and you wouldn't necessarily easily learn from them.

So again, what we're going to talk about is the kind of standard staple of electronic structure. Before I did that, I want to say a little bit about a topic we haven't touched about, which is what's called the spin polarized version of density functional theory. If you remember, the Hohenberg-Kohn theorem, everything is, in essence, expressed in terms of the charge density. Everything is a function of the charge density. And the electron spin never explicitly appears in there.

But of course, electrons have spin. If you don't consider any coupling with the angular momentum, it's really just up or down spin, so plus or minus a half [INAUDIBLE]. And we'll treat here, in the lecture, spin as a scalar quantity. So really, spin is just up or down-- plus 1 or minus 1, or plus a 1/2, or minus a 1/2 in the appropriate units.

And in reality, spin, as soon as it applies to the angular momentum, becomes a vector quantity. Also as soon as it couples to a magnetic field from the environment, it becomes a vector quantity. And people do that now already, treating spin as a vector quantity, but most codes that you will use, spin will simply be treated as a scalar, which is fine for most purposes.

We tend to refer to them as just up and down spin. I'll often write that with the either up and down arrow. Now why do you actually need to treat the electron spin? Let me just sort of give you a refresher of why you need to deal with the spin.

Well, the reason is the Pauli Exclusion Principle really. Is that the Pauli Exclusion Principle tells you that two electrons cannot be in exactly the same quantum state. Remember that? So that means that if you have two up electrons approaching each other, versus say an up and a down electron, these will approach each other differently.

Because these two are in the same spin state, so if you bring them very close together, they essentially now get the same coordinate as well. So they almost have the same quantum numbers now, and the Pauli exclusion principle prevents that.

So the Pauli Exclusion Principle keeps electrons with parallel spin essentially away from each other. Whereas if you have electrons with empty parallel spin, even if these are at the same position, they don't have the same spin, so they don't have the same set of quantum numbers, so the Pauli Exclusion Principle doesn't act on them.

And the Pauli Exclusion Principle is essentially something that keeps the electrons away without an explicit term for it in the Hamiltonian. It's not like-- the coulombic interaction, of course, keeps electrons away from each other. But the Pauli Exclusion Principle is essentially something on top of that that comes from anti-symmeterizing the wave.

You don't see it directly in the form-- as a term in the form of the Hamiltonian. And if you remember, consequences of-- it's going to take forever-- the Pauli Exclusion Principle is Hund's rule. That if you remember atomic d level, say, for example, if you add electrons to, say, five d levels, you're going to add them with parallel spin first, because again, then the Pauli Exclusion Principle is satisfied. And then you start filling them with the anti-parallel levels.

So this is going to carry over in atoms-- in solids-- I'm sorry. Basically, if those five, say-- let's focus on d levels. If they remain degenerate-- so they remain roughly at the same level, you're going to fill them according to Hund's rule. And that's what I've shown here.

So in solids, these d levels will split a little, which is what I've shown here, but they don't split a lot. Then you're actually going to fill them with parallel spin. And so if that's the case, you have a lot of magnetic moment on your ion.

You have five electron spins here. You have no down spin, so you have a strong magnetic moment. And I'll show in a second where that becomes important.

On the other hand, let's say you're in an environment that splits off two of these, but so much higher. At some point, you won't satisfy Hund's rule anymore, because the energy cost of-- let's say after you put in these three green electrons, now you have to add two more. You have five electrons.

To fill them with parallel spin, you would have to put them here. But since those levers are so much higher, you basically want to pay the exchange penalty-- they call it the Hund's rule penalty-- to put them in a lower orbit when you put them in with anti-parallel.

So in some sense, whether you get a lot of magnetic spin-- a lot of magnetic moment left over depends on how much your orbital will split in the end. But I'm going to show you in the end that it can actually have significant consequences on the physical properties of your material.

So in your density functional calculation, you will carry a magnetic moment locally when the up density and the down density are not the same. And so I've sort of already given away here how this problem is dealt with in density functional theory. The interesting thing is that if you think about it very hard, you shouldn't have to deal with spin in density functional theory.

Somehow, it put me back at my old picture. In principle, remember where we told you that-- Professor [? Mazari ?] told you that the energy and the potential is a function of the charge density?

So the charge itself should actually also contain the information about electron spin and magnetic moment, even though it doesn't explicitly contain that. But for a given charge density, there would probably be a certain amount of spin polarization. So it should all be in that functional, but remember, that's the functional that we don't know.

And so in practice, that doesn't work very well. So what we do is we really helped density functional theory along by treating the up and the down density separately. But again, you should keep in mind that in principle, in the formalism of density functional theory, you wouldn't have to do that.

So if you do what's called, say, the local spin density approximation, which is the spin polarized version of the Local Density Approximation-- and so goes under the name LSD-- Lucy in the Sky with Diamonds, or LSDA. And there's a version of that for the GGA, which nobody-- often people will just call it LDA or GGA, but they will call it sometimes spin polarized LDA or GGA.

What you do there is that you have a separate density for the up electrons and a separate density for the down electrons, and the two will interact differently. So the up/up will-- up will interact differently with up, with up, then up with down. And that comes from the way the exchange correlation potentials are defined. So I think Professor [? Mazari ?] mentioned restricted and unrestricted to Hartree-Fock before very briefly, and this is essentially the same idea.

Sort of one quick tip I want to give you is that if you have spin polarized materials, it's often much more useful to look at spin densities than at charge densities. One of the really cool things about doing quantum mechanics is that you can actually look at the charge densities, and look at the electrons, which most people get very excited about the first time they do quantum mechanics.

It's kind of cool-- you can look at where the electrons go. Well, the first thing you learn is that you don't see much when you look at charge densities. You typically see big blobs of charge, and it's very hard to see any fine structure of bonding in blobs of charge density.

And I wanted to show you an example. Here's lithium cobalt oxide. It's a transition metal oxide. It's a layered material-- layers of oxygen here-- the red things. And then layers of cobalt, and layers of lithium.

And if you plot the charges in a plane-- this is actually a plane in the plane of the figure. If you look at the charges-- this is a picture of the charge density. You see big blobs, and you see the oxygen layers here.

And where the oxygens are 2 minus, so remember that you're showing only the valence electrons. So these have a lot of valence electrons, and so you see a lot of intensity. Cobalt has somewhat less valence electrons on it, so you see less intensity here. And lithium is ionized to lithium plus, so it has no valence electrons on it, so you see almost nothing here.

But in essence, this doesn't give you a lot of detail. If you actually take a material like this, and rather than plot the charges, you plot the spin polarization density. So let's say up minus down, or down minus up. So it's how much magnetic moment there is locally. You get much cleaner pictures.

This is the same picture. It's a slightly different material with different ions in it. It's with nickel and manganese in it. But here's the oxygen. By the way, I should have told you red is the neutral color in this picture, so it's 0.

So now you don't see the oxygen at all, and the reason is oxygen doesn't have spin. It's a filled shell. Every time you have filled shells, you don't have spin. So looking at spin density often allows you to filter out certain ions, and you really-- the transition metals tend to have spin on them, and you see that very clearly. So it's often a trick that I just wanted to share with you.

So I want to go in sort of the last half hour-- is go through the kind of numerical accuracy, and then slowly try to connect that to the physical accuracy that you get in properties. So if you want to use-- have initial methods in density functional theory to get to engineering properties, a lot of steps you have to make, because in the end, we calculate simple things.

We've got to get charge densities, and band structures, and energies. And you'll talk later to somebody, and they want to know what's the corrosion resistance of this. And corrosion resistance is not a quantum operator, so you need to take a lot of steps to go from the simple, what I would call primitive output, to engineering properties.

But before you even take that step, you need to understand the kind of accuracy-- how reliable your output is. And so I collected a lot of results. And I was going to start with the simple things-- the energies of the atoms.

So here's a collection, and it looks like a bunch of numbers, but there's a very systematic trend in it. So what I show for a bunch of atoms-- and these all should have a minus sign in front of them, because if you sum up all the electronic states in the atom, they're obviously binding, so they should be negative.

So there's always the experimental line, at least for most of them. The LDA number, and then the GGA with a fairly recent implementation of the exchange correlation function.

And so if you look carefully at the numbers, let's take one here-- let's take carbon for example. So experiment is 275688. Oops, that wasn't good. LDA, the binding energy is somewhat weaker-- almost is an electron volt weaker.

And GGA is slightly closer to the experiment. That's typically what you'll see. If you look at all the other atoms, you'll see a very systematic trend. In the LDA, in the atoms, the electrons are not bound enough. In the GGA, They are somewhat closer to the experiment.

It gets more interesting when you look at molecules, because now you can talk about a physical binding energy. And so the one we look at here is for very simple diatomic molecules. What's their binding energy? So what's their energy to pull them apart?

So if you think about it, you have a molecule-- has an ab. This is the vector between them. You'll have sort of something that looks like this. And so we're looking at what's that well depth here-- the binding energy.

If you look at hydrogen, H2, you have the experimental number here, the LDA number. You see now that the LDA number-- that the binding is too strong. The H2 molecule is bound too strongly. It's not so bad. It's only about 5% in hydrogen. And the GGA, the binding is too weak.

For reference, I've also put uncorrected Hartree-Fock here. Professor [? Mazari ?] showed you essentially what Hartree-Fock is, which is essentially having the Hartree term for the self-consistent Coulomb interaction from the other electrons, and the exact exchange with no correlation effects.

This is an interesting one that you will actually often use-- O2-- any time you look at oxidation reactions, for example. Experiment-- oxygen is only about minus 5.2 ev binding. LDA binds by a whopping 7 and 1/2. So you're more than two electron volts off. GGA gets you a little closer in this case.

Uncorrected Hartree-Fock is off the charts. And this is something you'll generically see. Uncorrected Hartree-Fock, very few people would actually use that any more. It's sort of for binding energies way off the charts.

These things are important. If you now-- let's say you want to look at an oxidation reaction. So that means that at some point, you're going to calculate the state of an oxide, and compare the chemical potential of the oxygen there to that of oxygen gas.

So of course, you have a big error in the oxygen gas. The question is, how much of that error carries over to the solid? If you make exactly the same error in the solid, then the reaction energy is perfect, Because. You're going to subtract the two. And a lot of practical things you do with density functional theory depend on error cancellation.

The thing is that you will have more error cancellation as the states that you subtract are more physically similar. But the problem is, let's say, you look at oxidation of a metal-- aluminum plus oxygen going to aluminum oxide. The oxygen in aluminum oxide is very different from the oxygen in the O2 molecule, so not all the error will cancel.

Let's say it was so bad that you kept 2 ev error. So 2 ev error in the molecule is one ev per oxygen if you want to think of it that way. If you're not careful about it, that's an enormous effect.

Think about it-- the chemical potential which relates directly to the energy goes like the logarithm of the partial pressure of oxygen. Remember that mu is mu 0 plus rt log p? So if now you invert that, that means your-- if you wanted to calculate, say, a partial pressure of oxygen at which something oxidized, your oxygen pressure goes exponentially with the energetics.

So if you have a 1 ev error at room temperature, your error in the oxygen pressure is the exponential of one ev over kt, which is off the charts. So you have to be a little careful with these kinds of error calibrations.

Fortunately, we'll see later when we look at reaction between solids, most of the error tends to cancel, and we get much, much better accuracy. If all our reaction energies were wrong by one ev, we wouldn't be here. We'd be out of business.

But you have to keep in mind that you get less error cancellation as the states you're comparing are different. The more they're different, the less error cancellation. That's sort of a rule of thumb. And so going from a gas to a solid is a significant difference.

Often, what people do is that if you want to get practical results, they'll add a correction to this which they fit at one point. So these are the small molecules. Let's go to the solids.

Here I don't have the binding energies, but I have the lattice parameters. But you'll see something that's very consistent with the molecules. If you look at the lattice parameters, you compare, say, the experimental ones versus the LDA ones. What you'll see is that the LDA ones almost always-- and I think always in this case are smaller.

Actually, yeah, because the difference is actually negative. The GGA results are always bigger. This is rather consistent whatever material you do.

You will find almost always that the LDA gives you lattice parameters that are too small by a factor of a percent, sometimes, 2%. And so people refer to that as the over-binding of LDA. LDA binds somewhat too strongly.

Remember you saw that in the molecules as well. Oxygen had a 7 ev binding energy, and it should only have a 5 ev binding energy. And so in solids, the way that comes out is that your equilibrium lattice parameters are slightly too small.

Actually, I'm not sure that I know of a single result where LDA gives a lattice parameter that's too big. I've seen that on some occasions in papers, but it's almost always an indication that the people did the calculation wrong.

Actually, an LDA lattice parameter that agrees with experiment is usually wrong-- a wrong calculation. GGA is much more unpredictable. The ones that I've shown here because they're simple metals and semiconductors give you a lattice parameter that's too large.

In GGA, it's actually also possible to get a lattice parameter that's too small, although it's rare. Most of the time, you're on the higher side, but it's less predictable, and that's sort of slightly problematic with the GGA.

In LDA, a good guess of the lattice parameters-- you calculate your lattice parameters, and you know you're on the low side. You always know that the real lattice parameter is going to be bigger. In GGA, it's slightly more difficult to predict on which side you are. But in metals, you do tend to be on the high side.

That actually has consequences for other properties, like the bulk modulus. If you compare, say, the experimental bulk modulus to the LDA one, s you'll find is that in almost all cases-- and I think in all cases that I've shown-- well, all cases except silicon-- the LDA bulk modulus is too large, so the material is too stiff that means.

And that kind of goes together with the over-binding. Remember, the bonding energy is too high, the lattice parameter is too small. All that is kind of in agreement with the material also being too stiff. As you compress the material, it gets stiffer.

GGA, most of the time, if you see from deviations, has to be on the other side. It tends to be too soft. And bulk modulus effects that will transfer, for example, also into vibrational frequencies.

In material, when you're too hard, too stiff, you'll have higher vibrational frequency. When you're too soft, you'll have lower vibrational frequencies. Here's the same for oxides. You're not exactly learning a lot new by looking at them, except that in oxides, the errors just tend to be slightly larger.

So here's the summary for geometry prediction. You almost always-- and I would probably say always-- under-predict with LDA. Less systematic errors with GGA. For normal materials like semiconductors and metals, often, your errors are confined to order 1% to 2%.

In transition metal oxides-- and if I have a chance, I'll say a little bit more about that later because they have electronic structures where the LDA and GGA approximations are particularly harsh on. You tend to have somewhat bigger errors. But I may say a little more about that if we get to it.

So I want to say something about predicting structure, and about the energy scale that's required. So this is often something you want to do. You want to know if I put my energy in this arrangement, is that lower in energy than some other arrangement, so I can kind of predict what the most stable arrangement is.

So that could be [INAUDIBLE] crystal structure, but it's the same for if you look, say, at a surface. So I wanted to give you a feeling for the scale of energetic differences. So for vanadium, I've listed the atomic energy here in [INAUDIBLE]. This is the energy of all its electrons, so not just the valence electrons actually.

The energy for FCC vanadium-- so remember, the first line is the atom, not in a solid. The second line is the FCC iron, and the third one is BCC iron.

So look at the differences. First of all, if you go from the atom to the solid, your first four digits don't even change. And again, that's a reflection of-- well, a lot of your deep core states don't change.

But you would see something similar even with a pseudopotential approximation where you only deal with the valence electrons. So the cohesive energy is only 0.03% of the total energy. So if you're calculating-- the reason I'm saying this-- if you're calculating the cohesive energy by first calculating the total energy of a solid, and then calculating atomic energy, you'd better do these things damn accurate, because you're going to subtract them, and most of what you subtract is the same.

So to get any significance in your result, you need to have high numerical accuracy. And that's not a big problem with a lot of code, but I want you to keep that in mind.

But few people care about the cohesive energy. Let's say you want to know whether vanadium is FCC or BCC. So you could calculate it as BCC. Now the FCC/BCC difference is only 0.001% of the total energy.

And these are not complicated structures. So in many cases, we're going to work with energy differences that are fractions of-- so that are 10 to the minus 6, 10 to the minus 7 times the total energy. So it's sort of telling you something about how much numerical accuracy you need.

If you want to look at mixing energies-- let's say I mix vanadium with something else-- platinum-- and you want to know what's the mixing enthalpy, because that sets the whole temperature scale for mixing, the whole phase diagram topology. That tends to be a fraction-- 10 to minus 6, 10 to the minus 7 of the total energy.

So my former advisor used to compare this to-- let's say you want to know the weight of a captain that sails a big supertanker. It's like weighing the tanker with the captain and without the captain, and looking at the difference, and that's the weight of the captain. You're almost at a kind of relative scale like that here.

So the [? cute ?] thing really is that all these approximations we make to density functional theory are obviously not this accurate. The total energy is not accurate up to a fraction of 10 to the minus 6. The only reason we're here and we can get physical behavior right is because a lot of the error the density functional theory makes in LDA and GGA is systematic.

And so a lot of it cancels away when you take energy differences. When I do FCC and BCC vanadium, yes, I may have an error of 10 to minus 4 in the energy, but most of it cancels away when I take the energy difference. And that's why we're lucky. But you have to keep that in mind, because again, the less cancellation you have, the bigger your error on the result.

So again, let me show you how well or how badly it does. So I did a very simple thing. I looked at how well does it predict, say, the structure of the elements. This is done in GGA, a standard [? Trudel ?] potential method. So this comes out-- you may have to look at your hand out because this is extremely fuzzy on the screen.

In red, I did metals that are experimentally FCC, and green I did metals that are experimentally BCC. Now what I show you is the calculated energy difference between FCC and BCC, and it's actually the first line below every element kind of like this.

And so when that's positive, the BCC energy is higher than FCC, so it's going to-- FCC is preferred over BCC. If it's negative like here, then BCC is preferred over FCC. And so if you look-- so the color is the experimental result, the number is calculated.

So if you look at them carefully, they're all correct. It's negative when we have green, it's positive when we have red. You can do a more subtle comparison. Look at the difference between HCP-- Hexagonal Close Pack-- and FCC.

And the reason that that's more subtle is HCP FCC are much more alike. They're all close pack. It's just the difference in stacking. ab, ab, versus abc, abc.

And so again, you'll see that they're all correct. The red ones are the FCC ones. They're the ones where that first line is positive. By the way, that number is in kilojoules per mole here. The yellow ones, that number is negative.

So we get the structure of the elements essentially correct. There are notable exceptions. In LDA, iron is wrong. Iron is FCC in LDA, not BCC, but in GGA, that's corrected.

And then there are the weirdos. If you go deep down in the periodic table, especially f electron metals have-- the f states are extremely localized even in metals, and so electron correlation becomes very important here. And I may say a little more about that.

And so there you'll start to see failures LDA and GGA. An important one is plutonium. Plutonium is kind of important for obvious reasons, especially if you work at national labs these days.

And so people are building more sophisticated methods to deal with materials such as plutonium. Typically, when you work with f electron metals, sometimes, you'll get the answer right, sometimes, you won't. But you should be a little more careful.

Let me skip this. You can get the-- most of the time, you'll get the structure of compounds right. If you go to transition metal oxides-- so I sort of went from metallic elements now to transition metal oxides-- most of the time, you also get the structure right, but things get more subtle.

In transition metal oxides, the transition metal has local d states. And I showed you before they often have significant spin polarization, so the first thing you need to do is turn spin polarization on, or you really get the wrong answer.

But it often gets worse. Remember that your spin is a scalar, so it's up or down. So now you have a spatial degree of freedom of how to organize that spin on the ions. If you have a bunch of ions, you could put them all with the same direction. That's a ferromagnet, or you could put them with sort of alternating direction as a kind of a anti-ferromagnet.

And then there's many ways to make them anti-ferromagnetic. And unfortunately, in transition metal oxide, it often matters because there's not only a strong spin polarization effect on the energy, but there's a fairly strong effect of the interaction between spins on different ions.

And I'm showing you a result here. This is a simple crystal structure. It's a structure of lithium manganese oxide. It's an ordered rock salt. These are still very simple.

But the correct answer-- I'm showing the comparison here between two structures only labeled by their symmetry, unfortunately. One is C2/m, and one is pmmn. There are similar structures, but one is orthorhombic, and one is monoclinic. The correct answer is pmmn-- is the real crystal structure.

If you do a non-spin polarized calculation-- so that's not even allowing spin on the ions-- you get a whopping error. I mean, C2/m is lower in energy by 250 mill-electron volt per formal unit. That's very large. In kilojoules, that'd be 25 kilojoules per mole. It's a very large error.

If you turn on spin polarization, will make them ferromagnetic. They're degenerate. And if you make them anti-ferromagnetic, this one is the lowest in energy.

Now here's a very common mistake that people make. If you take this material at room temperature, it's paramagnetic. Some people say, well, it's paramagnetic, so I shouldn't have any spin polarization-- there's no net moment.

That is so wrong, because a paramagnet still has a local moment. The ions still have a moment on them, it's just randomly oriented. So you still need to represent that moment.

Because it turns out that's the biggest effect on the energy is the fact that you have that local moment. It's not necessarily how they're arranged. You can actually see that here. How they're arranged makes you go from this difference to this difference.

But turning on the local moment makes you go from this difference to that difference. So never fall in that trap. It's really only non-magnetic materials or diamagnetic materials for which you don't really need spin polarization.

Now why is this effect so important? It's really because if you spin polarize an ion, you fill different orbitals. I mean, I've shown that here with a bunch of d orbitals.

And this is typically how they split in most oxides. Every time an ion is octahedral-- five d orbitals tend to split in pairs of 3, 2-- 3 down, 2 up. In some cases, 2 down, 3 up. But let's say you have to put four-- five electrons in there.

How many do I have? I'm missing-- no, four electrons. If you put them-- that's called high spin. So all parallel spin, you put them like this. If you put them low spin, you put them like this. So here we have no moment.

So these two ions have different chemical properties, because the electrons occupy different orbitals. These are different d orbitals, and so this orbital points in a different direction, for example, than this one.

So by spin polarizing, you create essentially a different ion. It's not really an issue of magnetism, because magnetic effects tend to be small in materials. But it's the fact that you create chemically a different ion because you fill different levels. That's really why these energy differences are so big.

I sort of want to end with showing you some reaction energies very quickly. And I'm going to sort make it systematically harder. So here's a simple one. A metal, lithium, BCC, with another metal, aluminum forming a lithium-aluminum compound.

Here's the experimental reaction energy, here's the LDA one. 10% off. That's classic. Metallic reaction energies, you're somewhere in the range 5% to 15%. The one I show here, copper-gold, is a notable exception where you're over 50% off. But most of them, it's much simpler.

So in metals, you tend to get very good reaction energies. I want to show you the case where things go wrong. Where your errors become bigger is in redox reactions. And I've shown sort of here three different ones. They're all related.

They're essentially a reaction between an oxide, or in this case, a phosphate with a metal to react the two together. And if you look at the reaction energies, you're not considerably off. GGA gives you 2.8 electronvolts for this reaction. Experiment is 3.5.

This one, which is very similar, the error is 30%. You get 3.3 electronvolts. Experiment is 4.6. Why is that? Well, it has to do with the lack of error cancellation. If you look in detail what happens to the electronic structure in these materials, these are redox reactions.

So if you do the math on the valences-- and believe me, this ion here is 3 plus. Phosphorus, 5 plus. Oxygen is 2 minus, so you can do the math.

Iron here is 2 plus, and lithium is 1 plus. So what has happened in this reaction? Well, you've taken essentially an electron from lithium in it's metallic state, and put it on the iron 3 plus to make iron 2 plus.

So essentially, you've transferred an electron from metallic lithium and ionized lithium to the iron to reduce it from 3 plus to 2 plus. But think about what's that doing? That electron in lithium-- lithium is there in the alkaline metals. That's an s electron. So that's a wide delocalized orbital, and I think it's metallic-- and you're transferring it to a localized d state on the iron.

So that electron is essentially being transferred between extremely different states, and this is what's killing you. Because you transfer between such different states, you start losing a lot of the sort of cancellation of errors that you need. That's the functional theory.

And in particular, the error here comes from something quite particular. It comes from what we call the self-interaction error. So I'm trying to sort make you understand where these errors come from so that when you work on your application, you get a bit of a feeling for.

If you remember how we solve all these quantum mechanical equations, we reduce them to one electronic equations, where you have the kinetic energy, the nuclear potential. And then this effective potential which, remember what all goes in there. The effective the potential-- that's the one that has the exchange correlation in it, but it also has the Hartree field, so the coulombic field from all the electrons.

Well, that field includes the electron itself. That's the sort of oddity in essence. When you calculate the charge density, that's the charge density of all the electrons. We then calculate the potential coming from the charge density, that's the potential coming from all the electrons.

But you operate that now on a single electron. So the electron is feeling its own potential. Part of the exchange correlation term corrects for that, but not all of it. And the problem is that the correction doesn't operate as well on different forms of the charge density.

In a metal, you have a small self-interaction error. And the reason is if you look at a state in a metal, sort of very delocalized. So very spread out charge density. So if you want to think of it, the part of the electron here doesn't feel much of the charge density coming-- of the potential coming from that piece of the charge density, because they're very far away.

Whereas if you do a very localized state, in some sense, then the potential from the electron is very high where the electron itself is sitting, because it's all very close. If you put an electron in a delta function if you-- if you didn't have an uncertainty principle, and you calculate its potential, it's basically sitting on top of itself then. You'd have an infinite self-interaction.

So the more local the state is, the more self-interaction you have. And the exchange correlation functional can't quite correct these two in the same way. And remember that the exchange correlation correction comes from homogeneous charge densities, so it tends to correct the metallic state better than the localized state.

And so this is why that redox reaction I showed you had a big error, because we were transferring from a state that was metallic-- the electron went from the lithium state-- to the transition metal state. And somehow, the self-interaction error doesn't cancel. And you will see things like that whenever you transfer electrons between quite different states, so that's something to keep in mind.

I think I'm running out of time here, so yeah, let me stop here. So I'm summarizing here because this is the stuff I went over before. In general, you do pretty well. I think if I'd given this summary 10 years ago, I would have been even more optimistic, because most people worked on metals and semiconductors which tend to be fairly delocalized state, so LDA and GGA do quite well.

I think as we dig into more complicated materials, we have learned more about the errors of LDA and GGA. But on sort of classic metals, you do pretty well with lattice constants, reaction energies, and cohesive energies. But now there is a series of methods under development-- and if we have some time, we might sort of just broach them-- to deal better with the correlation energy, with the self-interaction energy to solve these problems of both energetics and also electronic structure-- things such as the band gap.

So I'll end here. And remember, on Tuesday, we have lab. So you meet in the lab, and then Thursday, we'll be back here.