

# Models in condensed matter physics

## 1 “CM Theory of everything”

Let us consider how to write a Hamiltonian for a system that has  $N_A$  atoms. For simplicity, let us assume they are all identical (generalization to multiple species is trivial), each with an atomic number  $Z$ , meaning that we have  $N_A$  nuclei with positive charge  $+Ze$  and  $N_e = ZN_A$  electrons with charge  $-e$ . If we use  $\hat{\vec{R}}_a, \hat{\vec{P}}_a$ ,  $a = 1, \dots, N_A$  to denote the position and momentum operators for the nuclei, and  $\hat{\vec{r}}_i, \hat{\vec{p}}_i$ ,  $i = 1, \dots, N_e$  for the electrons, then we might expect that the answer is:

$$\hat{H} = \sum_{a=1}^{N_A} \frac{\hat{\vec{P}}_a^2}{2M} + \sum_{i=1}^{N_e} \frac{\hat{\vec{p}}_i^2}{2m} + \sum_{a < b} \frac{Z^2 e^2}{|\hat{\vec{R}}_a - \hat{\vec{R}}_b|} + \sum_{i < j} \frac{e^2}{|\hat{\vec{r}}_i - \hat{\vec{r}}_j|} - \sum_{a,i} \frac{Ze^2}{|\hat{\vec{R}}_a - \hat{\vec{r}}_i|} \quad (1)$$

where  $M$  is the nuclear mass and  $m$  is the electron mass. The first two terms are the kinetic energies of the nuclei and the electrons, and the last three describe Coulomb interactions between nuclei, electrons, and nuclei and electrons, respectively.

This abstract notation can be further simplified to the  $r$ -representation that we are used to from quantum mechanics (QM), by choosing to work with wavefunctions  $\Psi(\vec{r}_1, \dots, \vec{r}_{N_e}, \vec{R}_1, \dots, \vec{R}_{N_A}, t)$  which define the amplitude of probability to find the electrons at locations  $\vec{r}_i, i = 1, N_e$  and the nuclei at locations  $\vec{R}_a, a = 1, N_A$  at time  $t$ . (As we'll see soon, this is a really bad idea, but it is what we are being taught to do in introductory courses of quantum physics, so let's go along for a bit to see where it takes us). In this case, we can replace the position operators by their eigenvalues, because  $\hat{\vec{r}}_i \Psi(\vec{r}_1, \dots, \vec{r}_{N_e}, \vec{R}_1, \dots, \vec{R}_{N_A}, t) = \vec{r}_i \Psi(\vec{r}_1, \dots, \vec{r}_{N_e}, \vec{R}_1, \dots, \vec{R}_{N_A}, t)$ , etc., and the momentum operators by the usual derivatives, for example  $\hat{\vec{P}}_a \rightarrow -i\hbar \nabla_a = -i\hbar \left( \frac{\partial}{\partial X_a}, \frac{\partial}{\partial Y_a}, \frac{\partial}{\partial Z_a} \right)$  where  $\vec{R}_a = (X_a, Y_a, Z_a)$ , and similarly for the electrons.

At this point we have to start simplifying notations somewhat, because it is boring to list all electrons' and nuclei' positions (especially considering that in a solid, we have  $\sim 10^{23}$  of each) – so as a first simplification, let me use  $\mathbf{r} \equiv (\vec{r}_1, \dots, \vec{r}_{N_e})$  to be the collection of all electronic locations, and similarly for nuclei  $\mathbf{R} = (\vec{R}_1, \dots, \vec{R}_{N_A})$ . Then, the wavefunction notation is simplified to  $\Psi(\mathbf{r}, \mathbf{R}, t)$ .

Once we've done this, the Schrödinger equation  $i\hbar \frac{\partial \Psi(t)}{\partial t} = \hat{H} \Psi(t)$  reduces to the partial differential equation (PDE) for the wavefunction  $\Psi(\mathbf{r}, \mathbf{R}, t) = \langle \mathbf{r}, \mathbf{R} | \Psi(t) \rangle$ :

$$i\hbar \frac{\partial \Psi(\mathbf{r}, \mathbf{R}, t)}{\partial t} = \left[ -\frac{\hbar^2}{2M} \sum_{a=1}^{N_A} \nabla_a^2 - \frac{\hbar^2}{2m} \sum_{i=1}^{N_e} \nabla_i^2 + \sum_{a < b} \frac{Z^2 e^2}{|\vec{R}_a - \vec{R}_b|} + \sum_{i < j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} - \sum_{a,i} \frac{Ze^2}{|\vec{R}_a - \vec{r}_i|} \right] \Psi(\mathbf{r}, \mathbf{R}, t) \quad (2)$$

Sometimes this is called the **CM Theory of everything**, because any condensed matter system can be formulated in these terms. Solving this equation, it may seem, is now just a computational matter and should be easy to deal with provided that we have good enough computers.

The problems with this last statement are two-fold: (i) we cannot possibly solve this PDE numerically; and (b) this Hamiltonian already has lots of approximations that may or may not be justified – so more care may be required in many specific cases.

Let's examine these claims. First, let me simplify the problem tremendously and assume I am interested not in a solid with  $10^{23}$  atoms, but in a single atom of iron, with  $Z = 26$ . If we choose the reference system where the nucleus is at the origin, we then have a wavefunction depending

only on the 26 positions of the 26 electrons. Of course, if we solve this numerically we need to use a grid to specify locations - let's use a very rough  $10 \times 10 \times 10$  grid for each electron (so long as we're considering states where the electrons are bound to the nucleus, it is reasonable to assume that the wavefunction goes to zero if any electron is far from the nucleus, so we only need to calculate it inside a finite volume centered on the nucleus, which we can divide in  $10 \times 10 \times 10$  pieces). Then, to know the wavefunction means to calculate and store its  $\sim 1000^{26}$  values corresponding to each possible position of each electron inside this grid (there are symmetries in the wavefunction that would decrease this number somewhat, but not that much). But  $1000^{26} = 10^{78}$ , and assuming we want double precision then we would need  $16 \times 10^{78} \sim 10^{79}$  bits to store the wavefunction at just one instant of time (or, alternatively, to store one eigenfunction). The problem is that there are only  $\sim 10^{79}$  atoms in the whole universe, so far as we know, so even if we could store 1 bit per atom we would need to use all the universe for storing this information (never mind how we could do that, or how long it would take to calculate all these numbers etc). Clearly, this is not the way to go forward even for a single atom, never mind for a solid with a macroscopic number of atoms - we'll have to be a lot smarter than to assume that brute-force numerics will solve all our problems. We'll spend a good part of this course looking at some simple approximations one can use to make this problem tractable, and discuss when they are good approximations. Roughly speaking, that is the case when the interactions are "weak" - we'll come back to what is this supposed to mean. In further courses and then on your own - especially if you become a theorist - you'll learn many more advanced techniques. However, you should know that at this point in time, we do not yet have general methods to deal with "strong" interactions. This is a major challenge in our (and other) areas of physics and a lot of current research is devoted to it.

But before discussing methods to deal with such complicated many-body Hamiltonians, let's see what is the matter with the Hamiltonian we wrote. As I said, the problem is that it already contains lots of approximations, and it is very important to keep these in mind. Some will turn out to be very good in most cases; some may be good in some circumstances but not others, so we have to be careful to use them appropriately; and some are definitely wrong.

Let's look at some examples.

In the first category, there are several candidates. For instance, we ignored gravitational forces between the various particles - they act, of course, but this is a very good approximation because those forces are extremely weak compared to Coulomb forces. Another approximation is that we're treating the nuclei as if they have no structure, whereas of course they are made of protons and neutrons, which themselves are made of quarks, etc. If we want to be more precise, we could allow each of those particles to have its own term in the Hamiltonian, and add the strong, short-range nuclear forces that bind nuclei together. Ignoring all these complications is a good approximation because we will only be interested in what is known as **low-energy physics**. This means that we will be interested in what happens to such a solid-state system when it's either in the ground-state, or in low-lying excited states that can be reached in normal experiments, eg. by heating the material to room temperature or by coupling it to fairly weak magnetic or electric fields, to measure its conductivity or magnetic susceptibility. The kinds of energies exchanged in such processes are many many orders of magnitude smaller than those needed to excite a nucleus out of its ground-state, never mind to split it into its components. So treating the nucleus as a structureless, point-like object with a given mass and charge is also a very good approximation, unless we decide to bombard our system with something carrying lots of energy - then we have to revisit it. We also ignored all other objects in the universe with which our solid interacts, but again this is usually a very reasonable approximation.

We have also assumed that all particles are non-relativistic: we used Schrödinger's instead of

Dirac's equation. As you may know from atomic physics, this is a good approximation for lighter atoms, but heavier ones accelerate their electrons enough to require relativistic corrections, such as including spin-orbit coupling, and this remains true in solids (do some reading if you do not remember what spin-orbit coupling is. Interestingly, it used to be considered rather un-important in condensed matter, until the recent discovery of the topological insulators, where it plays a key role. This is a good illustration of the danger of approximations: a universe without spin-orbit coupling would not have materials with those sorts of topological properties. And it took so long for our community to realize that such materials do exist precisely because of the assumption that this effect is not important, therefore it was not included in the model, therefore its consequences were not investigated). And speaking of relativistic effects, in everything we discussed so far we completely forgot that the electrons and the nuclei have spin! It is relatively easy to fix this in what we said so far, we would simply need to add the quantum numbers for all spins to those for the positions, in the arguments of the wavefunction, in order to fully specify the actual wavefunction. But if the nuclei have non-zero spin (when does that happen?) there will also be hyper-fine couplings between those spins and the spins of the electrons. Depending on circumstances, those may or may not be important enough to need to be included in the Hamiltonian.

Another important aspect not included in this Hamiltonian are interactions with E&M fields. Of course, if we apply external fields we need to add the corresponding terms (which you hopefully remember from QM; if you know how to do it for a single charged particle then it is straightforward to generalize to any number of such particles. We will definitely review this soon enough). But even if there are no external E&M fields, the system is coupled to the E&M vacuum. If it wasn't, it would mean that if somehow or other we placed the system in an excited state, it would stay in that state forever (or until we acted again on it). Of course, in reality systems always de-excite by emitting photons to lower their energy (well, some care is needed with selections rules), and this is only possible because even if there are no applied E&M fields, the system is not isolated but interacts with the E&M vacuum. Can we ignore such processes? Again, that depends on circumstances. There will be a characteristic decay time from the excited states (the stronger the coupling to the E&M fields, the shorter this time). If that time is long compared to the typical timescale of the processes we're interested in, then it is a good approximation to pretend that it is infinite, i.e. to ignore this coupling. Etc etc etc. The key message here is that you always have to think critically and evaluate if things make sense for the particular situation you are considering: an approximation might be excellent in one situation but a disaster in a different situation.

In the last category (definitely wrong) we have to include forgetting the spin of the electrons and nuclei, as already mentioned. Because of these degrees of freedom, the Hilbert space is enlarged and this affects the Hamiltonian as well. Even if there is no spin-orbit coupling or hyperfine couplings and no external magnetic fields, so that we do not have to add extra terms in the Hamiltonian, the operators have parts describing their action on the spin components (this is the identity, for positions and momenta operators). Their spins determine whether particles are fermions or bosons, and this is very important for the symmetry of the wavefunctions. This is why we cannot "forget" them anymore, like we usually do in introductory QM when we deal with just one particle and the spin degree of freedom is treated more like an afterthought.

So where are we now? On one hand, we concluded that the Hamiltonian could be quite a bit more complicated than the one usually presented as the "CM Theory of everything", see Eqs. (1) and (2). As we will discuss in this course, many times we actually deal with much simpler Hamiltonians, so called **model Hamiltonians** (for instance, you may have heard of tight-binding Hamiltonians, or the Hubbard Hamiltonian; we will discuss these and other examples and their link to the "Theory of Everything" soon). Such model Hamiltonians are obtained from so-called

**first principles** Hamiltonians like that of Eq. (1) upon making further suitable approximations, by ignoring aspects that are not expected to matter for the physics we are interested in. Unlike Eq. (1), where the only information we need in order to specify the Hamiltonian is what types of atoms and how many there are (this is why they are called first-principles), in model Hamiltonians there are parameters that cannot be calculated, instead they need to be extracted from comparison with experimental results (examples coming soon). This is one disadvantage of such models. The advantage is that solving such model Hamiltonians is a simpler task (as I just argued, we cannot do the job for Eq. (2)). A very few have exact solutions, but even for the ones that don't, finding reasonable approximations is easier than trying to deal with the "CM Theory of everything". The problem is that often times, especially for complex materials, it is very difficult to figure out what is the correct model Hamiltonian. This is usually an iterative process where one starts with the simplest possible form and then solves it to see if its predictions agree with experiments or not (for acceptable values of the parameters). If they don't, then some important aspects have not been included in the model, so one has to go back and fix it accordingly. Figuring out the "right" model and the appropriate methods to deal with it can take many decades of work, for hard problems.

Hopefully this brief discussion explains why condensed matter is basically a conglomerate of approximations. These are needed both to figure out what model Hamiltonian is appropriate for the problem of interest, and how to deal with it (how to solve the resulting Schrödinger equations); a lot of care needs to be exercised to use the correct approximations in the proper circumstances. From this point of view, CM is quite different from high-energy physics, E&M and several other such areas, where the Hamiltonians (or the equations that need to be solved, like Maxwell's equations in E&M) are known exactly, and approximations are only (sometimes) needed to get their solutions. By contrast, the "CM Theory of everything" and its generalizations, besides not being "exact" in that sense, are also unsolvable so we need to replace them with simpler model Hamiltonians, for which we then need to find accurate approximate solutions.

There is a very good saying that you should always keep in mind: "All models are wrong, but some of them are useful" (G. E. P. Box and N. R. Draper, *Empirical Model-Building and Response Surfaces*) – because we are dealing with such complicated systems, we are always forced to ignore some aspects of the physics, so the models we study in CM are definitely not "right". But we have been very successful in figuring out how some classes of materials work, otherwise all the gadgets in our lives, such as computers or solar cells, would not exist. As I said, we do not know yet how to deal with all such problems, but with the typical optimism of our profession, we believe that it is just a matter of time and effort before that will be accomplished.

From now on we will proceed to discuss such approximations. We will start with two which are virtually always used, and then we'll be in a position to sketch a fairly general approach to the further steps that are necessary. After we get a handle on the types of resulting model Hamiltonians, we will start to discuss methods to deal with them.

## 2 Core vs. valence electrons

Consider what happens when we "make" a solid by bringing together atoms that were initially very far apart. When at large distances, the atoms are essentially isolated so each will be neutral, binding its  $Z$  electrons close to its nucleus. As we know from atomic physics, in the ground state these electrons occupy the lowest available shells: the first two go into the  $1s$  level (one with spin-up, one with spin-down), and the next occupy  $2s, 2p, 3s, \dots$  orbitals as necessary. Of course, the lower shells have lower energy because their wavefunctions are spatially closer to the nucleus, so those electrons feel more of its Coulomb attraction, although one needs to be a bit careful with

symmetries: *eg.*,  $3d$  levels have higher energies than  $4s$  ones but are spatially closer to the nucleus because they are the first set of levels with angular momentum  $l = 2$  (so, their energy is higher because of the rotational part of the kinetic energy, not because they are far from the nucleus).

The electrons from the last occupied shell are thus generally the furthest out and so they are the ones most affected by interactions with the other atoms, when the atoms are brought closer together. We call such electrons **valence electrons**, whereas the ones from inner, totally filled shells are called **core electrons**. The precise identification of which is which depends on particular circumstances; the same element placed in different materials might have different numbers of valence electrons. Figuring out which is which is one of the first decisions in going to a simpler model, and one uses intuition and knowledge about other similar materials to come up with a good hypothesis, which then needs to be tested. In any event, let's assume that  $Z_v$  of the  $Z$  electrons are valence electrons. It is then a reasonable step, in many cases, to “compactify” the remaining  $Z_c = Z - Z_v$  core electrons from the inner filled core shells together with their nuclei and treat them as ions. In other words, we assume that these core electrons will always be in their ground state because interactions with other atoms are not providing enough energy to excite them to higher levels; so we don't need to worry about their state because it is basically the same as in the isolated atom. (Of course, there are circumstances when this is not true, for instance when high-energy photons are used to excite electrons out of the core levels – then we have to explicitly keep them into the model, obviously. We won't discuss such cases in this course, but you have to always keep in mind that there are no universally valid approximations).

With this approximation, i.e. thinking of our solid as being made out of ions of charge  $+Z_v e$  plus  $Z_v$  valence electrons each, we can simplify the Hamiltonian of Eq. (1) quite a bit: it will still have the same form but we now have  $NZ_v$  instead of  $NZ$  electrons to deal with! In many cases  $Z_v = 1$  or  $2$ , so this has reduced tremendously the number of electrons whose state we're trying to track. However, we gain this at the expense of a complication: we now no longer know what is the correct  $Z$  value to be used in the last term, describing the attraction between the valence electrons and the ions. Using  $Z_v$  implies that the valence electrons are always found to be outside all the core electrons so that they experience the full repulsion from the filled spherical core shells (this is known as **screening**). But this is not always true, often times the wavefunctions of the valence electrons have some probability to find them close to nuclei, meaning they only experience partial screening from the core electrons, i.e. we should use a  $Z^* \neq Z_v$ . Now you see the problem: we cannot know what is the correct  $Z^*$  until we figure out where are the valence electrons, and in particular how close to the nuclei they get. But to find that, we need to know the Hamiltonian from which we can get their wavefunctions that carry this information. So, we need to know  $Z^*$  if we want the wavefunction, but we need the wavefunction to be able to calculate  $Z^*$  – a vicious circle. One way to deal with this is to use  $Z^*$  as a free parameter, solve the problem and then either try to estimate its resulting value and iterate until we reach self-consistency, or (what is done in most cases) use comparison with some experimental data to infer what the correct  $Z^*$  is for that material.

This is a clear illustration of what I said before about the difference between first-principles Hamiltonians, where we know all the values of all the parameters needed, and model Hamiltonians – these are simpler but they depend on parameters whose values we do not know a priori, and have to either calculate or infer from fits to experiments.

So now we have a much simpler Hamiltonian describing ions and valence electrons:

$$\tilde{H} = \sum_{a=1}^{N_a} \frac{\hat{P}_a^2}{2M} + \sum_{i=1}^{N_a Z_v} \frac{\hat{p}_i^2}{2m} + \sum_{a < b} \frac{Z_v^2 e^2}{|\hat{R}_a - \hat{R}_b|} + \sum_{i < j} \frac{e^2}{|\hat{r}_i - \hat{r}_j|} - \sum_{a,i} \frac{Z^* e^2}{|\hat{R}_a - \hat{r}_i|} \quad (3)$$

which at low energies of interest to us, should give the same prediction for what happens (for the correct  $Z^*$ ) like the full Hamiltonian  $\hat{H}$  does, but at significantly reduced computational cost.

Of course, you can easily convince yourself that this approximation is nowhere near enough to allow us a full brute-force solution, because even if  $Z_v = 1$ , we still have  $\sim 10^{23}$  ions and valence electrons to deal with. So we need to continue simplifying. The second approximation that is virtually always used is the Born-Oppenheimer (BOA) approximation. We discuss this next.

### 3 The Born-Oppenheimer approximation

BOA was proposed very early on, in 1927, as the formalism of QM was basically being finalized. It was originally intended for small molecules but it works the same in systems with macroscopic numbers of atoms, like the ones we are interested in. The basic idea is that there is a major difference between ions and valence electrons in the Hamiltonian above, namely their mass: ions are 3 or more orders of magnitude heavier than electrons, therefore they should move much much more slowly than electrons do.

In many textbooks, at this point one jumps directly to saying that we will therefore assume the ions to be “frozen” in their equilibrium positions. Note that we will only consider systems where these are on a regular lattice, i.e. we only study crystalline solids. It is reasonable to expect that the ground-state of a system with a large number of identical atoms has such a high degree of order, because if the ions are identical then each will prefer to be in an environment identical to the one seen by all the other ions, and that can only be achieved in a highly ordered state. Of course, ions located near the surfaces cannot be in the same environment so physics there will be affected, but for this course we will mostly ignore such complications. Note also that there are plenty of materials that are not crystalline – that could be either because they are alloys, i.e. a mix of many elements; or that they were grown in conditions which did not allow them to reach the best (most orderly) ground-state structure, etc. We will not consider such cases.

Finally, note that we would now need to know what is this periodic structure! (again, this is a feature of a model Hamiltonian, where we need to input information known from elsewhere. A first-principle Hamiltonian, if we could solve it, would predict how the nuclei order in their ground-state, we would not need to put this information in the model by hand). I hope you remember from your undergrad solid-state physics course how one goes about inferring that structure based on the type of bonding, and how it can be measured with  $X$ -rays or neutron scattering; also you need to know what are the Bravais lattices, how we define unit cells and bases, what is the Brillouin zone, etc. because we will be using these concepts later on.

Going back to the BOA of freezing the ions in their equilibrium positions (which from now on we assume to be known, i.e. the lattice structure is given to us), there is clearly something very wrong with this approach: the ions are quantum objects so they can’t possibly be at rest in well-specified locations because that violates Heisenberg’s uncertainty principle. So when we make this approximation we are definitely missing something, which is the fact that the ions must be oscillating about their equilibrium positions and that their motion should, presumably, influence the behavior of the valence electrons somehow.

The way to make progress is to proceed through a sequence of approximations:

(i) first, we will assume that the ions are indeed frozen in their equilibrium positions, and study the motion of the valence electrons in the resulting periodic potential. This step will itself necessitate lots of further approximations and we’ll spend a good part of the course on these. As we’ll see, this is already enough to explain many properties of some classes of materials.

(ii) second, we will consider these lattice vibrations, i.e. the motion of the ions around their equilibrium positions (while ignoring the valence electrons). This leads to the discussion of phonon modes, which you should already be familiar with at least for simple lattices; we will briefly review the general formulation and solution for this problem.

(iii) third, we will consider the effects of the so-called **electron-phonon coupling**, i.e. the effects of the lattice vibrations on the motion of the valence electrons, and viceversa. As we will see, these give rise to spectacular properties such as conventional superconductivity, which we will briefly discuss.

(iv) in the remaining time we will also look at some magnetic properties, i.e. consequences of the fact that electrons have spins. For this we'll only be able to consider step (i), i.e. how this degree of freedom will influence the behavior of electrons if the ions are assumed to be frozen. Of course, one should then go on to study the effects of lattice vibrations on these magnetic properties and viceversa. We won't have time for that, but at least you should have some idea how that is likely to be approached, and be able to make progress on your own on this topic, if needed for your research.

But before we can do all this, we need to learn better notation: when dealing with such huge numbers of particles, using wavefunctions like we do in systems with 1 or 2 particles is way too cumbersome and prevents progress simply because it makes any calculations time consuming and very painful. Using the so-called **second quantization** formalism will take care of this problem, so we will first acquaint ourselves with this new notation.