Expanded Comments on The Tight-Binding Model

Students have commented that the book races too quickly through some of the basic parts of condensed matter physics, giving general derivations but not establishing good mental pictures. One example is the tight–binding model, and I am going to try to fix the situation with these notes.

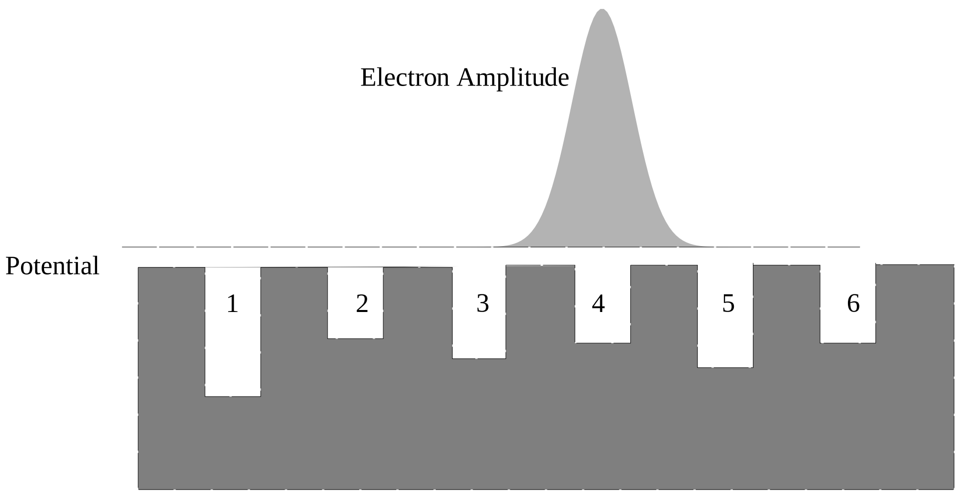

Figure 1. One–dimensional tight–binding model. A single electron sits in a well that nearly holds it captive, but allows it to leak out into other wells at some slow rate.

The tight–binding model is illustrated in Figure 1. Start in one dimension, and think of a metal as a collection of potential wells. The wells may or may not all have the same depth; the analysis will be easiest when they are all the same depth, and the most interesting when they are not. In a perfect crystalline metal, one thinks of all the wells having the same depth, or else alternating depth in a crystalline pattern. Wells of random depth correspond to a solid with a lot of defects and randomness in it.

Figure 1 also shows an electron, which is trapped in one of the wells. OK, that's not quite true. It is almost trapped in one of the wells. When one writes a tight–binding model, the idea is that if there were just a single well sitting in isolation, one could solve the problem of the electron sitting in that well, find its energy, and find the electron wave function. The blob sitting over well #4 is the wave function of an electron trapped in an isolated well sitting at position #4.

It may be helpful at this point to solve the following

Problem concerning wells. Consider the potential

$$U(x) = \begin{cases} 0 & \text{if } x < -a/2 \\ -\mathcal{U} & \text{if } -a/2 \le x \le a/2 \\ 0 & \text{if } x > a/2 \end{cases} . \tag{1}$$

Take $\mathcal{U} > 0$. Solve Schrödinger's equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U(x)\psi = \mathcal{E}\psi, \tag{2}$$

and find the ground state energy $\mathcal{E}_0$.

Questions concerning wells. Here are some additional questions you should think about before you go on, either looking back at a quantum mechanics text if needed, or else working things out for yourself:

- For a potential well of the form in Eq. (1) is there always a bound state; that is, a state of negative energy where the wave function dies of exponentially away from the well?
- Can there be more than one bound state?
- Is there always more than one bound state?
- What happens if $\mathcal{U}$ is positive rather than negative?
- What if the bottom of the well were not perfectly flat, but had some arbitrary bumpy shape? Would the results of the calculation change?
- What if the problem were posed in two dimensions rather than one dimension; the potential looks like a piece of paper with a disc cut out of it, and an electron is attracted to the disc with potential strength $-\mathcal{U}$. Now what happens? Is there still a bound state? What is its energy?
- What if the problem were posed in three dimensions rather than one or two dimensions; the potential looks like a balloon floating in the middle of space, and an electron is attracted to the balloon with potential strength $-\mathcal{U}$. Now what happens? Is there still a bound state? What is its energy?
- OK. So suppose I'm some sort of nano–technology pioneer, and I take a bunch of atoms and I put them in a line a few Å apart, and I put one electron on the line of atoms. Should I think of that as a series of one–dimensional wells like the one in Eq. (1), or should I think of that as a series of three–dimensional wells?
- If I can manipulate atoms, what would I do to create one–dimensional wells, two–dimensional wells, or three–dimensional wells?

Back to the tight–binding model. So, now it is possible to come back to that innocent–looking Figure 1 and explain what it means. The electron blob is the wave function of an electron bound state trapped in a single well. The potential wells are not really indicating the heights of potentials, but are indicating the energies of electron bound states in the wells. That is, the picture is showing $\mathcal{E}_0$ and not $\mathcal{U}$. For this reason, the picture could be describing the result of the nano–technologist's experiment where a bunch of three–dimensional potential wells have been stuck together.

You should be convinced at this point that there is plenty to think about when an electron faces just one well, and it's not trivial to put the electron in contact with an infinite number of wells. This brings up the main idea of the tight–binding model.

Suppose I know the binding energy of an electron in each well given to me. Suppose the electron wave functions fall off really fast away from the well, and by the time the wave functions get to the next well, they are tiny.

Now I'm going to make a bit of a leap, and define a collection of quantum states, $|l\rangle$, where l is an integer ranging from $-\infty$ to ∞. What are these states? Think of them as corresponding to the wave function of the electron bound state in well number l. That is fine just so long as you do not think too hard about it. Unfortunately, if you do, you may have some questions:

Q: So, if $|l\rangle$ corresponds to the wave function of an electron in well l, then the different states $|l\rangle$ will not be orthogonal, and I am going to have trouble with the formalism of quantum mechanics. Is it legal to work with states $|l\rangle$ that are not orthogonal?

A: It is easiest to work with Dirac notation if the states $|l\rangle$ are orthogonal, so they are orthogonal.

Q: Then states $|l\rangle$ are not really the bound states in the wells after all. Then what are they?

A: They are very close to the bound states in the well, but they have been made orthogonal to one another.

Q: Thanks for the help.

A: No, really, this is not such a big deal. There are many different ways to think about it. One way is to think back to your class on linear algebra. You may have learned about something called Gramm–Schmidt orthogonalization, which seemed particularly useless at the time. The way it works is something like this. You have wave functions $\psi_0(x)$,

$\psi_1(x)$, and $\psi_{-1}(x)$ that are the bound states in isolated wells -1, 0, and 1. Now you can modify ψ_0 by writing

$$\phi_0(x) = \psi_0(x) - \psi_1(x) \int dx'\, \psi_1^*(x')\psi_0(x') - \psi_{-1}(x) \int dx'\, \psi_{-1}^*(x')\psi_0(x'). \tag{3}$$

Or, in Dirac notation

$$|\phi_0\rangle = |\psi_0\rangle - |\psi_1\rangle\langle\psi_1|\psi_0\rangle - |\psi_{-1}\rangle\langle\psi_{-1}|\psi_0\rangle \tag{4}$$

Q: Well, $\phi_0(x)$ isn't even normalized any more.

A:

Q: Aren't you going to answer my question?

A: That was not a question. That was a statement[1].

Q: What should I do to normalize ϕ_0?

A:

$$|\phi_0\rangle = \frac{|\psi_0\rangle - |\psi_1\rangle\langle\psi_1|\psi_0\rangle - |\psi_{-1}\rangle\langle\psi_{-1}|\psi_0\rangle}{\sqrt{1 + |\langle\psi_1|\psi_0\rangle|^2 + |\langle\psi_{-1}|\psi_0\rangle|^2}} \tag{5}$$

Q: So how does all this help me?

A: It only helps if the overlap integrals such as $\langle\psi_1|\psi_0\rangle$ are small. If these are of order 10^{-2}, then the corrections needed to normalize ϕ_0 in Eq. (5) are of order 10^{-4}, which is why I didn't even write them down at first. Also, if $\langle\psi_1|\psi_0\rangle$ is of order 10^{-2}, one would expect $\langle\psi_1|\psi_{-1}\rangle$ to be of order 10^{-4}, and overlaps between more distant sites to be smaller yet. So the difference between ϕ_0 and ψ_0 is that for ψ_0 the overlaps with neighboring wave functions were of order 10^{-2}, but after I have done the subtractions in Eq. (5), the overlaps have dropped down to order 10^{-4}. Check it out by acting from the left with $\langle\phi_0|$, $\langle\psi_1|$, and $\langle\psi_{-1}|$ and see what you get.

Q: The Gramm-Schmidt algorithm is a systematic procedure for making any collection of vectors orthogonal to one another. Why are you doing an approximate version of the algorithm, and making comments about overlap integrals being small?

[1]Dirac is supposed to have done this once during a colloquium.

A: Because if the Gramm-Schmidt procedure is carried out with wave functions that look like the wave functions one expects to arise for bound electrons sitting in wells, then at the end of the day, the resulting orthogonal wave functions should look very much like the wave functions with which one began, but with tiny corrections. For the corrections to be tiny, the overlap integrals have to be small.

Q: This looked as if it was going to be simple at first, and now it's gotten messy. Isn't there a simpler way to go through it?

A: Well, that was the idea of Section 8.4 (Tightly Bound Electrons, p. 194), which uses formal properties of Wannier functions to try to avoid all these questions. Something like a tunnel through a mountain of complexity. But if you're a student seeing it for the first time you sense you're going through a tunnel, but haven't seen the mountain, and you wonder why on earth the textbook author built a tunnel, and why the instructor is driving you through it. There is also Section 10.2.2 (Linear Combination of Atomic Orbitals, p. 235) which uses a starting point more similar to the one in these notes and also comes up with the tight–binding model at the end.

Back to the Tight–Binding Model Again. What I have at this point is a collection of wave functions $|l\rangle$. They look almost exactly like the bound states over isolated wells at location l, but I have manipulated them so that they are all orthogonal to one another. They are no longer perfect eigenfunctions for the isolated wells, and they are not perfect eigenfunctions for the infinite collection of wells lying side–by–side either. Why? I just made a bunch of wave functions orthogonal to one another. I haven't solved Schrödinger's equation yet.

Now I am going to write down an approximate Hamiltonian for an electron in the presence of all the wells, which is

$$\hat{\mathcal{H}}_{\text{tb}} = \sum_l U_l |l\rangle\langle l|. \tag{6}$$

This is not the "tight–binding Hamiltonian", but the "too–boring Hamiltonian". I will use it to introduce some notation. I will be denoting by U_l the energy of the bound state sitting at site l, and I will understand that U_l is usually negative. Now, why is this Hamiltonian boring? It doesn't do anything. The eigenstates are $|l\rangle$, the energies are U_l, and if an electron starts at site l it stays there forever.

Looking back at Figure 1, it should not seem possible to arrange things so that an electron that starts at site l simply sits at site l forever. Its wave function leaks over into the next well, and interacts with the potential in that well. There must be some probability for the electron to tunnel into the two neighboring wells, even if the probability is small. You can look back at pages 78 and 79 for a brief summary of some formulas having to do with tunneling, and references to textbooks that talk about it at greater length. Just as I previously imagine solving the problem of an electron in a well and finding a binding energy U_l, now I imagine solving a tunneling problem, and finding the the rate at which electron amplitude leaks from well 0 to well 1 is $\mathfrak{t}$. That is, suppose an electron is in state $|l\rangle$ at time $t = 0$. Denote the wave function of the electron by $|\psi(t)\rangle$. Then at later times I expect something like

$$\hbar i \frac{\partial}{\partial t} \langle 1|\psi(t)\rangle = \mathfrak{t}\langle 0|\psi(t)\rangle. \tag{7}$$

I can arrange for this to be true by adding a term $\mathfrak{t}|1\rangle\langle 0|$ to the Hamiltonian. To make the operator Hermitian, I will have to add $\mathfrak{t}|0\rangle\langle 1|$ as well. This means that the rate at which a particle leaks from well 0 to well 1 is the same as the rate at which a particle leaks from well 1 to well 0. At this point I have

$$\hat{\mathcal{H}}_{\text{tb}'} = \mathfrak{t}|0\rangle\langle 1| + \mathfrak{t}\langle 1||0\rangle + \sum_l U_l |l\rangle\langle l|. \tag{8}$$

I'm finally ready to write down the tight–binding model in one dimension, as

$$\hat{\mathcal{H}}_{\text{TB}} = \sum_l \left[\mathfrak{t}|l\rangle\langle l+1| + \mathfrak{t}|l+1\rangle\langle l|\right] + \sum_l U_l |l\rangle\langle l|. \tag{9}$$

A Few More Questions.

Q: Shouldn't you compute a different tunneling matrix element $\mathfrak{t}$ for every single pair of sites?

A: Yes. But the model has enough rich behavior as it stands that I simply don't do that in this book. Letting the tunneling matrix elements vary from site to site does not introduce any additional qualitative behavior, just changes numerical results.

Q: What equation in the book corresponds to Eq. (9)?

A: Equation (8.35) or stuff in the vicinity of Equation (10.25).

Q: Are there some problems in the book that will give me some practice with the tight–binding model?

A: Look at Problems 8.6 and 18.2, 18.3, and 18.4. I really ought to have a problem in Chapter 10 also, but never made one up.

Solution in a Simple Case. In the simplest possible case, the tight–binding model becomes

$$\hat{\mathcal{H}}_{\mathrm{TB}} = \sum_l \mathfrak{t}|l\rangle\langle l+1| + \mathfrak{t}|l+1\rangle\langle l| + \sum_l U_0|l\rangle\langle l|. \tag{10}$$

All the wells are the same depth. Therefore, the problem has the sort of translational symmetry that lets Bloch's theorem work, and I can solve it. Actually, I remember the first time I ran into this in graduate school. Embarrassing story, really. I got to a problem like this, and I had no idea what to do. I just sat staring it it for hour after hour. The problem was due the next day. At something like 2 in the morning another student came into the office, and I told him I was stuck and just sitting there with no idea what to do. He said, "Oh, it's just a like the undergraduate wave problems in Crawford." So I looked back at the text for my sophomore waves text, and sure enough there was exactly the same math describing a bunch of discrete masses connected by springs. I copied the solution, and had it worked out in a few minutes.

The main idea is that whenever you have some form of translational symmetry, you can use Fourier transforms. Guess a solution in the form

$$|\psi_k\rangle = \sum_{l'} e^{ikl'}|l'\rangle. \tag{11}$$

Let Eq. (10) act on this. You get

$$\hat{\mathcal{H}}_{\mathrm{TB}}|\psi_k\rangle = \sum_{ll'} e^{ikl'} \left\{\mathfrak{t}|l\rangle\langle l+1|l'\rangle + \mathfrak{t}|l+1\rangle\langle l|l'\rangle + U_0|l\rangle\langle l|l'\rangle\right\}. \tag{12}$$

The great thing about having orthonormal states is that

$$\langle l|l'\rangle = \delta_{ll'} \tag{13}$$

So now I have

$$\hat{\mathcal{H}}_{\mathrm{TB}}|\psi_k\rangle = \sum_{ll'} e^{ikl'} \left\{\mathfrak{t}|l\rangle\delta_{l+1,l'} + \mathfrak{t}|l+1\rangle\delta_{l,l'} + U_0|l\rangle\delta_{l,l'}\right\}. \tag{14}$$

For the middle term, I replace the dummy index l by $l-1$. Now I have

$$\hat{\mathcal{H}}_{\mathrm{TB}}|\psi_k\rangle = \sum_{ll'} e^{ikl'} \left\{\mathfrak{t}|l\rangle\delta_{l+1,l'} + \mathfrak{t}|l\rangle\delta_{l-1,l'} + U_0|l\rangle\delta_{l,l'}\right\} \tag{15}$$

$$= \sum_l \left\{ \mathfrak{t}|l\rangle e^{ik(l+1)} + \mathfrak{t}|l\rangle e^{ik(l-1)} + e^{ikl} U_0 |l\rangle \right\} \tag{16}$$

$$= \sum_l e^{ikl}|l\rangle \left\{ \mathfrak{t}e^{ik1} + \mathfrak{t}e^{-ik} + U_0 \right\} \tag{17}$$

$$= \sum_l e^{ikl}|l\rangle \left\{ 2\mathfrak{t} \cos k + U_0 \right\} \tag{18}$$

$$= \left\{ 2\mathfrak{t} \cos k + U_0 \right\} |\psi_k\rangle \tag{19}$$

So ψ_k is an eigenfunction with energy

$$\mathcal{E}_k = 2\mathfrak{t} \cos k + U_0. \tag{20}$$

Final Questions.

Q: So I begin with electrons localized on sites, and I end up with something like plane waves extending everywhere. Isn't that weird?

A: That's Bloch's theorem. This might be a good time to look back at the beginning of Chapter 7.

Q: I've seen this before. You have to have a in the answer where a is the lattice spacing.

A: No I don't. But I can. Just replace k everywhere by ka, except in the subscript to ψ. In what I just did, l is a dimensionless index, and k is dimensionless.

Q: But if I make wells really deep, shouldn't electrons get stuck on their sites forever?

A: If wells are really deep, then the hopping matrix element $\mathfrak{t}$ becomes exponentially small, the band energies $\mathcal{E}_k$ are all nearly equal to U_0, and the electron takes a time of order $1/\mathfrak{t}$ to leave the place where it begins. One says that the electrons are in very narrow bands, and very tightly bound. For the simple problem I just solved, they are never stuck forever. Chapter 18 provides some cases where they really get stuck forever.

Q: What happens when all the energies U_l are not the same? Can I still solve the problem?

A: This is the subject of Sections 18.3.2 and on to the end of Chapter 18. P. W. Anderson won a Nobel prize for two big contributions between around 1955 and 1965, and one of them was to ask this question about random energies and make a first serious stab at answering it.

Q: So far you have just talked about one electron in this potential. What if I put more than one electron in the potential?

A: Well, electrons are fermions, so I can't put two of them in the same state. They do have spin, so I can put two of them into any give state indexed by k, one with spin up, one with spin down. To build a ground state, I put the first electrons into the state where $\mathcal{E}_k$ is the lowest, and then keep populating states in ascending order until I run out of electrons.

Q: How would I set that problem up formally?

A: Use second quantization. If you turn toward the end of the book you can see variants of the tight–binding model in second–quantized notation. The Anderson Model in Equation (26.83) is just the tight–binding model, with some added terms in front. The Hubbard model, (26.127) is another variant, with a term thrown away and another term added. Anderson was probably the first person to write down (26.127) too, but Hubbard studied it more thoroughly, so Hubbard's name has stuck to it. The amazing thing about the Hubbard model is that just by adding one little extra term to the tight–binding model so that electrons at different sites interact with each other, not only can the model not be solved exactly in more than one dimension, but there is still no consensus on qualitative features of the solution after 40 years of trying.

Q: Who first wrote down the tight–binding model?

A: I believe it was Bloch, in the first paper about what is known as Bloch's theorem, as a solvable example. I don't know why the model is called the "tight–binding model" rather than the "Bloch model." The connection between scientists' names and equations is curious and random.