These first several weeks we will give an overview of some of the key ideas in solid state physics. Before we can understand why systems act differently at the nm scale than in bulk, it is necessary to lay the groundwork by discussing bulk properties and how they arise. These ideas form the intellectual foundation upon which most nanoscale physics has been based.

A major caveat: It’s extraordinarily difficult to do a complete and modern treatment of these ideas even in a two-semester condensed matter physics course. There also is no single textbook that covers all these topics in a unified way, with an eye toward the nanoscale. These notes will contain numerous references to other books and articles, and we will supplement liberally with material from other sources.

1 The fundamentals

This course will assume that you’re familiar with undergraduate-level quantum mechanics and statistical mechanics. For those of you who have already taken a solid state physics course, hopefully this will refresh your memory and present material in a slightly different light than you’ve seen it previously.

Quantum mechanical topics you should be familiar with include: the Schrödinger equation; particle-in-a-box; simple harmonic oscillator; time-independent perturbation theory; time-dependent perturbation theory. For these last two, I’ve got separate review documents. If you need to brush up on quantum mechanics, I recommend Refs. [1–4].

Statistical physics ideas that are particularly relevant are: ensembles, temperature, and the chemical potential; Maxwell-Boltzmann distribution; Fermi-Dirac distribution; Bose-Einstein distribution; kinetic theory; phonons. Good references for statistical physics include Refs. [5–9].

2 Ideal Fermi Gas

It’s pretty amazing that we can understand anything at all about the properties of condensed matter. Consider a cubic centimeter of copper, for example. It contains roughly $10^{23}$ ion cores and over $10^{24}$ electrons, all of which interact (\textit{a priori} not necessarily weakly) through the long-range Coulomb interaction. However, through the power of statistical physics and quantum mechanics, we can actually understand a tremendous amount about that copper’s electrical, thermal, and optical properties.

In fact, because of some lucky breaks we can even get remarkably far by making some idealizations that at first glance might seem almost unphysical. In the rest of this section, we’ll
start with a simple model system: noninteracting electrons at zero temperature in an infinitely high potential well. That is, the “particle-in-a-box” problem. Gradually we will relax our ideal conditions and approach a more realistic description of solids. Along the way, we’ll hit on some important concepts and get a better idea of why condensed matter physics is tractable at all.

By starting with noninteracting electrons, we’re able to pick a model Hamiltonian for the single particle problem, solve it, and then pretend that the many-particle solution is simply related to that solution. Remember, too, that electrons are fermions: The Pauli Principle tells us that the many-particle wave function must change sign if the positions of any two identical fermions are interchanged. For noninteracting fermions, the many-particle state can then be written as a totally antisymmetrized product of single-particle states. This is called a “Slater determinant”.

Recall, too, that the electrons have spin \( s = 1/2 \). Picking a convenient reference axis and calling it \( z \), we can write the spin state of a particular electron as \( | \uparrow \rangle \) (spin-up), or \( | \downarrow \rangle \) (spin-down). A spin-up electron has spin angular momentum with magnitude \( |s^2| = s(s + 1)\hbar^2 = 3/4\hbar^2 \), and a projection of \( +1/2\hbar \) along the \( z \) axis. Any given single particle spatial state may be occupied by an electron that is either spin-up or spin-down (but not, obviously, by two electrons that are both spin-up; that would violate the Pauli principle). In many standard situations (indeed, in nearly all existing microelectronics technology) spin is comparatively irrelevant, typically introducing “a factor of 2 for spin”. Later, we will see both why this is so, and why, in magnetic systems, spin becomes crucial for understanding many effects, including the operating principle behind the read heads in computer hard drives.

For now, however, we can get much of the essential physics out by not worrying about Slater determinants. Instead, when we want to count electrons in the many-particle ground state, we just think about filling up each single-particle spatial state with one spin-up (\( \uparrow \)) and one spin-down (\( \downarrow \)) electron. This is akin to what we do all the time when we write e.g. the electronic structure of carbon as \( 1s^22s^22p^2 \). In chemistry this is called the aufbau principle, from the German word for “structure”.

Consider an infinitely tall 1d potential well with a flat bottom of length \( L \) (the generalization to 2d and 3d is simple, and we’ll get the results below). What (single-particle) states do we use here? Intro quantum mechanics tells us that an eigenfunction with momentum \( p \) for a free particle is a plane wave with a wavevector \( k = p/\hbar \). (A wavevector, if you recall, points in the direction of wave propagation, and has a magnitude given by \( 2\pi/\lambda \) for a plane wave. Generally, if a wave propagates along a displacement vector \( \mathbf{r} \), the amount of phase accumulated by the wave is given by \( \mathbf{k} \cdot \mathbf{r} \).) The eigenfunctions of the well have to be built out of planewaves, since inside the box the particles feel no potential. The boundary conditions are that the wavefunction \( \psi(x) \) has to vanish at the edge of the well and outside (the infinite potential step means we’re allowed to relax the usual condition that \( \psi'(x) \) has to be continuous). The wavefunctions that satisfy these conditions are:

\[
\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{\pi}{L} nx
\]

where \( n > 0 \) is an integer. So, the allowed values of \( k \) are quantized due to the boundary conditions, and the states are spaced in \( k \) by \( \pi/L \).

One can more easily see what this means by looking at the two-dimensional case. For a particle in a two-dimensional square box of side \( L \), the spatial wavefunctions that satisfy the hard wall boundary conditions are of the form:

\[
\psi_{n_x, n_y}(x, y) = \frac{2}{L} \sin \frac{\pi}{L} n_x x \sin \frac{\pi}{L} n_y y.
\]
Figure 1: This is a k-space plot. Each dot represents an allowed combination of \((k_x, k_y)\) such that the wavefunction of Eq. 3 would satisfy hard wall boundary conditions in a square box of side \(L\).

We can rewrite that as

\[
\psi(x, y) = \frac{2}{L} \sin k_x x \sin k_y y, \tag{3}
\]

as long as we restrict \(k_x\) and \(k_y\) to the discrete values given by integer multiples of \(\pi/L\). Now look at Figure 1. Here we’re looking in “k-space” or “reciprocal space” (so-called because \(k\) has units of \(1/L\)). Each allowed point, \((k_x, k_y)\), corresponds to a spatial state that meets the boundary conditions at the walls of a 2d box with side length \(L\). If we want to know how many allowed spatial states \(N\) are within an interval \(\Delta k_x, \Delta k_y\), it’s clear from the figure that we take the area in k-space, and divide by \((\pi/L)^2\). That is,

\[
N = \frac{\Delta k_x \Delta k_y}{(\pi/L)^2}. \tag{4}
\]

Similarly, in 3 dimensions, we have states described by \(k = (k_x, k_y, k_z)\), where \(k_x = (\pi/L)n_x\), etc. Now each spatial state in \(k\) each takes up a “volume” of \((\pi/L)^3\). As usual, the kinetic energy of an electron in such a state is

\[
E = \frac{p^2}{2m} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2). \tag{5}
\]

An important feature here is that the larger the box, the closer the spacing in energy of single particle states, and vice-versa. This is our first example of a “quantum confinement effect” in the electronic structure of solids, and it has two clear consequences:

- For macroscopic 3d boxes, the spacing between single-particle levels is extremely small, as we’ll see shortly. Therefore it’s not a bad approximation to assume a continuum of allowed energies. What that really means is that, when we want to count states, we can replace sums (discretely counting each dot in Fig. 1) by integrals (computing an area in the figure, and using the dot spacing to calculate the number of states).
For truly nanoscale boxes, the single-particle level spacing can completely dominate the energy scales of the system. For example, the spacing of levels in an atomic-sized 3d box is on the order of several eV, an effective temperature scale of tens of thousands of degrees.

Now consider $N$ of our ideal electrons in a 3d box, in the many-particle ground state. Remember, so far we’re neglecting interactions between the point-like electrons, so we don’t have to worry about electron-electron collisions. This system is called an ideal Fermi gas. Filling the single particle states from the bottom up, what is the energy of the highest occupied single-particle state? We can count single-particle states in $k$-space, and for values of $k$ that are large compared to the spacing, we do this counting using an integral rather than a sum. Calling the highest occupied $k$ value the Fermi wavevector, $k_F$, and knowing that each spatial state can hold two electrons, we can write $N$ in terms of $k_F$ as:

$$N = 2 \times \left( \frac{L}{\pi} \right)^3 \int_0^{k_F} \frac{1}{8} 4\pi k^2 dk.$$  \hspace{1cm} (6)

In the first line, the 2 comes from spin, and the integrand comes from computing the volume of an octant $(k_x, k_y, k_z > 0)$ in k-space. So, given the number of electrons and the size of our box, we can find the wavelength of the highest energy electrons in the box, $2\pi/k_F$.

If we define $n_{3d} \equiv N/V$ and the Fermi energy as

$$E_F \equiv \frac{\hbar^2 k_F^2}{2m},$$  \hspace{1cm} (7)

we can manipulate Eq. 6 to find the density of states at the Fermi level:

$$n_{3d}(E) = \frac{1}{3\pi^2} \left( \frac{2mE}{\hbar^2} \right)^{3/2}, \quad \rightarrow$$  \hspace{1cm} (8)

$$\nu_{3d}(E = E_F) \equiv \frac{dn_{3d}}{dE}|_{E=E_F} = \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} E_F^{-1/2}.$$  \hspace{1cm} (9)

The density of states $\nu_{3d}(E)$ as defined above is the number of single-particle states available per unit volume per unit energy. This is a very important quantity because, as we will see later, the rates of many processes are proportional to $\nu(E_F)$. Intuitively, $\nu(E)$ represents the number of available states into which an electron can scatter from some initial state. Looking at Eqs. (9,9) we see that, in 3d, increasing the density of electrons increases both $E_F$ and $\nu_{3d}$.

Let’s review. We’ve dumped $N$ particles into a box with infinitely high walls and let them fill up the lowest possible states, with the Pauli restriction that each spatial state can only hold two electrons of opposite spin. We then figured out the energy of the highest occupied single-particle state, $E_F$, and through $\nu(E_F)$ and the sample size we can say how far away in energy the nearest unoccupied spatial state is from $E_F$.

Thinking semiclassically for a moment, we can ask, what is the speed of the electron in that highest occupied state? The momentum of that state is called the Fermi momentum, $p_F = \hbar k_F$,
and so we can find a speed by:

\[
v_F = \frac{\hbar k_F}{m}.
\]

in 3d = \frac{\hbar (3\pi^2 n_{3d})^{1/3}}{m}.

(10)

So, the higher the electron density, the faster the electrons are moving. As we’ll see later, this semiclassical picture of electron motion can often be a useful way of thinking about conduction.

We can redo this analysis for 2d and 1d, where \( n_{2d} \equiv N/A \) and \( n_{1d} \equiv N/L \), respectively. The results are summarized in Table 2. Two remarks: notice that \( \nu_{2d}(E_F) \) is independent of \( n_{2d} \); further, notice that \( \nu_{1d} \) actually decreases with increasing \( n_{1d} \). This latter property is just a restatement of something you already knew: the states of an infinite 1d well get farther and farther apart in energy the higher you go.

The results in Table 2 are surprisingly more general than you might expect at first. One can redo the entire analysis starting with Eq. (1) and use periodic boundary conditions (\( \psi(x = L) = \psi(x = 0) = 0; \psi'(x = L) = \psi'(x = 0) \)). When this is done carefully, the results in Table 2 are reproduced exactly. This is quite interesting: it’s not the specific boundary condition that’s all that significant; just the existence of the boundary is enough to impose “quantum confinement” effects like the spacing of single-particle levels.

From now on, we will generally talk about the periodic boundary condition case, which allows both positive and negative values of \( k_x, k_y, \) and \( k_z \), and has the spacing between states in \( k \) be \( 2\pi/L \). In the 3d case, this means the ground state of the ideal Fermi gas can be represented as a sphere of radius \( k_F \) in \( k \)-space, with each state satisfying \( k < k_F \) being occupied by a spin-up and a spin-down electron. States corresponding to \( k > k_F \) are unoccupied. The boundary between occupied and unoccupied states is called the Fermi surface.

Notice that the total momentum of the ideal Fermi gas is essentially zero; the Fermi sphere (disk/line in 2d/1d) is centered on zero momentum. If an electric field is applied in, say, the \( x \)-direction, because there are available states, the Fermi sphere will shift, as in Fig. 2. The fact that the sphere remains a sphere and the picture represents the case of an equilibrium current in the \( x \)-direction is discussed in various solid state books.
Figure 2: Looking down the z-axis of the 3d Fermi sphere, before and after the application of an electric field in the \( x \)-direction. Because there were allowed \( k \) states available, the Fermi sphere was able to shift its center to a nonzero value of \( k_x \).

### 2.1 Distributions and chemical potential

One other point. Extending the above discussion lets us introduce the familiar idea of a distribution function, \( f(T, E) \), the probability that a particular state with energy \( E \) is occupied by an electron in equilibrium at a particular temperature \( T \). For our electrons-in-a-box system, at absolute zero the ground state is as we’ve been discussing, with filled single-particle states up to the Fermi energy, and empty states above that. Labeling spin-up and spin-down occupied states as distinct, mathematically,

\[
f(0, E) = \Theta(E_F - E),
\]

where \( \Theta \) is the Heaviside step function:

\[
\Theta(x) = \begin{cases} 
0, & x < 0 \\
1, & x > 0.
\end{cases}
\]

Note that \( f \) is normalized so that

\[
\int_0^\infty f(T, E)\nu(E)dE = N.
\]

At finite temperature the situation is more complicated. Some states with \( E < E_F \) are empty and some states above \( E_F \) are occupied because thermal energy is available. In general, one really has to do the statistical mechanics problem of maximizing the entropy by distributing \( N \) indistinguishable electrons among the available states at fixed \( T \). This is discussed in detail in many stat mech books, and corresponds to minimizing the Helmholtz free energy. The answer for fermions is the Fermi distribution function:

\[
f(T, E) = \frac{1}{\exp[(E - \mu(T))/k_B T] + 1},
\]

where \( \mu \) is the chemical potential. The chemical potential takes on the value which satisfies the constraint of Eq. (13). At \( T = 0 \), you can see that \( \mu(T = 0) = E_F \). See Fig. 2.1.
A physical interpretation of $\mu$ is the average change in the free energy of a system caused by adding one more particle. For a thermodynamic way of thinking about $\mu$, start by considering why temperature is a useful idea. Consider two systems, 1 and 2. These systems are in thermodynamic equilibrium if, when they’re allowed to exchange energy, the entropy of the combined system is already maximized. That is,

$$\delta S_{\text{tot}} = \left( \frac{\partial S_1}{\partial E_1} \right) \delta E_1 + \left( \frac{\partial S_2}{\partial E_2} \right) \delta E_2$$

$$= \left[ \left( \frac{\partial S_1}{\partial E_1} \right) - \left( \frac{\partial S_2}{\partial E_2} \right) \right] \delta E_1$$

$$= 0, \quad (15)$$

implying

$$\left( \frac{\partial S_1}{\partial E_1} \right) = \left( \frac{\partial S_2}{\partial E_2} \right). \quad (16)$$

We define the two sides of this equation to be $1/k_B T_1$ and $1/k_B T_2$, and see that in thermodynamic equilibrium $T_1 = T_2$. With further analysis one can show that when two systems of the same temperature are brought into contact, on average there is no net flow of energy between the systems.

We can run through the same sort of analysis, only instead of allowing the two systems to exchange energy such that the total energy is conserved, we allow them to exchange energy and particles so that total energy and particle number are conserved. Solving the analogy of Eq. (15) we find that equilibrium between the two systems implies both $T_1 = T_2$ and $\mu_1 = \mu_2$. Again, with further analysis one can see that when two systems at the same $T$ and $\mu$ are brought into contact, on average there is no net flow of energy or particles between the systems.

### 2.2 Fermi surfaces: consequences

Once the Fermi energy scale has been set, one can define a Fermi temperature,

$$T_F \equiv \frac{E_F}{k_B}. \quad (17)$$

For $T \ll T_F$, the system is nearly in its ground state, and $f(E)$ is well-approximated by the step-function form. Such a Fermi gas is said to be nearly degenerate. It is interesting to
compute the Fermi temperature for a bulk metal like copper. Estimating the electronic density as \(8 \times 10^{22}\ \text{cm}^{-3}\) and using the free electron mass, we find \(T_F \approx 80000\ \text{K}\). Clearly it’s fair to consider such an electron gas as degenerate. As we shall see later, \(T_F\) in semiconductors may be much lower, requiring cryogenic conditions to achieve degeneracy.

The existence of a Fermi surface, itself a necessary consequence of the Pauli principle, has major consequences for the thermodynamic and electronic properties of solids. As we will see soon through some examples, the Pauli principle means that electrons may only scatter from occupied (obviously) single-particle states into empty single-particle states. That means that the electrons “buried” deep within the Fermi sea cannot be scattered into unoccupied single-particle states without an energy cost comparable to \(E_F\) and a momentum exchange comparable to \(|k_F|\). The result is that the electrons that dominate most processes have energies within \(k_B T\) of \(E_F\).

A standard solid state physics example of this effect is the calculation of the electronic heat capacity of a metal. While a classical gas of \(N\) (noninteracting) electrons would have a heat capacity given by \(3/2 N k_B\), independent of temperature, the nearly degenerate Fermi gas has a heat capacity that varies \(\sim \nu(E_F) T\). The density of states at \(E_F\) is of critical importance here. Similarly, if one uses Fermi’s Golden Rule to calculate the transition rates of some process with an energy scale \(\ll E_F\), the results usually end up proportional to \(\nu(E_F)\).

2.3 An illustrative example

Now we see that the variation of the \(E_F\) dependence of the density of states with dimensionality (Table 2) is really quite profound! Suppose we had a nanoscale potential well that was strongly confined in \(z\) and \(y\), but was extended in \(x\). Structures like this are achievable, and this is not too poor a way of modeling certain properties of, for example, a carbon nanotube. We can approximate the single-particle eigenstates of this potential well as products of 1d particle-in-a-box type states:

\[ \psi(n_x, n_y, n_z) = A \phi_{n_x}(x) \cdot \phi_{n_y}(y) \cdot \phi_{n_z}(z), \]

where \(n_x, n_y, n_z = 1, 2, 3, \ldots\). We start adding electrons to this potential well, filling up states from the bottom. If \(L_x >> L_y >> L_z\), we know that the first states that get filled are \(\psi(1, 1, 1), \psi(2, 1, 1), \psi(3, 1, 1), \ldots\). Because, roughly speaking, the \(x\) states are closer in energy than the \(y\) and \(z\) states. In this range of filling, the density of states looks like the 1d example from above.

Eventually we reach a level of filling such that the next electron will spill over into the second \(y\)-subband. From then on we’ll start filling states like \(\psi(1, 2, 1), \psi(2, 2, 1), \psi(3, 2, 1)\ldots\). Recall that the 1d density of states goes like \(\nu_{1d}(E) \sim E^{-1/2}\). At this crossover to the second subband, the total density of states has a divergence called a Van Hove singularity. This dramatic behavior of the density of states can be seen, for example, in optical studies of carbon nanotubes. Optical absorption has very strong peaks when incident photon energies are large enough to promote carriers up to a van Hove singularity in the density of states.

2.4 The story so far

So, thus far we’ve learned a number of things:

- One useful model for electrons in solids is an ideal Fermi gas. Starting from simple particle-in-a-box considerations we can calculate properties of the ground state of this
system. We find a Fermi sea, with full single particle states up to some highest occupied level whose energy is $E_F$.

- We see that quantum confinement (the effect of having a bounded system) determines the spacing between single-particle levels in the non-interacting case.
- We also calculate the spacing of states near the Fermi energy and the (semiclassical) speed of electrons in this highest state.
- The dimensionality of the system dramatically alters the qualitative dependences of system properties ($e.g.$ the dependence of $\nu(E)$).
- We introduce the idea of a distribution function for calculating finite-temperature properties of the electron gas.
- Finally, we see the chemical potential, which determines whether particles flow between two systems when they’re brought into contact.
- The density of states at the Fermi level determines many measured properties, and therefore the dimensionality of the system can have profound experimental consequences.

3 How ideal are real Fermi gases?

Obviously we wouldn’t spend time examining the ideal Fermi gas if it wasn’t a useful tool. It turns out that the big concepts from the previous section generally persist even when complications like actual atomic structure and electron-electron interactions are introduced. Why are classical and/or highly idealized quantum mechanical treatments like these of any use at all, given the background potential of the ion cores and the mutual electron-electron Coulomb interaction in any real material?

Band theory provides some of the answers. First we will worry about the effects of the potential of the ion cores for a single electron moving in a solid. We will see how, in perfect solids, a picture of electron motion quite similar to the free Fermi gas case can be relevant.

To get to band theory, we’ll start from the opposite extreme: the energy levels in small molecules, relevant to Huckel theory. We will then extend the ideas of Huckel theory to crystalline solids, arriving at the “tight-binding” model for bandstructures. Finally, we’ll consider the case most like the ideal Fermi gases described above, “nearly-free electron” theory that is surprisingly good at describing electrical properties of metals and free carriers in semiconductors.

3.1 Huckel theory - molecules

3.1.1 Born-Oppenheimer

Before we tackle electronic states of molecules, it’s important to mention another approximation underlying nearly everything we’re going to do: the Born-Oppenheimer approximation (BOA). The BOA treats the atoms as if the nuclear and electronic wavefunctions are completely separable, and ignores direct nuclear-nuclear interactions.

Instead of worrying about the nuclear positions being influenced by the electronic states, we treat the nuclear positions $R_i$ as parameters. We can solve the electronic problem as a
function of $R_i$, and then say that the nuclei move in a potential created by the electrons. This is reasonable for many cases because nuclei are very heavy compared to electrons, and the timescale for electronic rearrangements is often much faster than that for significant nuclear motion. In what follows, consider the BOA to be in effect. (A side note: people doing research on femtosecond chemistry and reaction mechanisms involving excited complexes often need to take a more sophisticated approach than the BOA.)

3.1.2 Basic Huckel

Start by thinking about two hydrogen atoms very far from one another. Each atom has an occupied 1s orbital, and a number of unoccupied higher orbitals ($p, d, \text{etc.}$). If the atoms are moved sufficiently close that wavefunction overlap can be significant (a separation comparable to a couple of Bohr radii), an accurate set of energy eigenstates requires by hybrid orbitals.

That is, do a perturbation theory calculation accounting for the Coulomb interaction between the electrons in this geometry. Instead of the ground state energy being minimized by two single-electron states (1s) of identical energy, we find two states ($\sigma, \sigma^*$) that differ in energy. The energy of two electrons (with antisymmetrized spins) in the $\sigma$ state is lower than that of each electron in a 1s state. The $\sigma$ and $\sigma^*$ states are bonding and antibonding molecular orbitals.

Now think about combinations of many-electron atoms. It’s reasonable to think about an ion core containing the nucleus and electrons that are firmly stuck in localized states around that nucleus, and valence electrons, which are more loosely bound to the ion core and can in principal overlap with their neighbors. We can then consider using the valence orbitals of those electrons as a basis for constructing molecular orbitals.

In the 1930s Huckel came up with an approximation procedure based on this idea of linear combinations of atomic orbitals (LCAO) in an attempt to calculate the electronic levels of conjugated hydrocarbons. Rather than doing a full perturbation theory calculation, the idea is to use a variational method: a trial wavefunction for the molecule is assumed to be a LCAO. The coefficients of the individual orbitals are found by minimizing the energy expectation value with respect to changes in the coefficients, using a parametrized model for the Hamiltonian.

**Note:** this method should be viewed as largely qualitative, giving some intuitive sense of level spacings, ionization energies, and molecular orbital shapes. This is not how electronic structure calculations are done anymore, since we now have the computing power to do far more sophisticated things.

Let’s do an example to make this concrete. Consider the allyl radical consisting of three carbons as shown here. Suppose that the Hamiltonian for the whole radical is $H$. The wavefunction $|i\rangle$ is the atomic $p$-orbital for the carbon at position $i$. Our trial wavefunction is assumed to be

$$|\Psi\rangle = c_1|1\rangle + c_2|2\rangle + c_3|3\rangle,$$

where the $c_i$ are the (real) coefficients in question. We can then write the energy expectation value,

$$\langle E \rangle = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{(c_1|1\rangle + c_2|2\rangle + c_3|3\rangle)(\sum c_i|j\rangle H(c_1|1\rangle + c_2|2\rangle + c_3|3\rangle))}{(c_1|1\rangle + c_2|2\rangle + c_3|3\rangle)(\sum c_i|j\rangle)}.$$

Let’s deal with the denominator first. Since the $|i\rangle$ correspond to orbitals around different
atoms, it is a common approximation to assume

$$\langle i | j \rangle = \delta_{ij}. \quad (21)$$

Clearly this is not necessarily accurate if the two atoms are in close proximity, but since we’re doing a perturbative calculation, we won’t really care to first order. With this assumption, the denominator is just \( c_1^2 + c_2^2 + c_3^2 \).

The numerator is more complicated, with lots of cross-terms. A simplifying approximation is to worry only about interactions between nearest neighbor atoms. Define

$$\alpha = \langle i | H | i \rangle, \quad \beta = \langle i | H | j \rangle, \quad i \neq j \quad (22)$$

for \( i, j \) nearest neighbors. Other matrix elements are approximated as zero. The idea is that \( \alpha \) and \( \beta \) can be compiled phenomenologically for different pairs of atoms. Rewriting,

$$\langle E \rangle \approx \frac{c_1^2 \alpha + 2c_1c_2\beta + c_2^2 \alpha + 2c_2c_3\beta + c_3^2 \alpha}{c_1^2 + c_2^2 + c_3^2}. \quad (23)$$

We’ve parametrized the interactions here! Note, too, that nowhere have we included the antisymmetrization of the total wavefunction - the Pauli principle is thrown in ad-hoc at the end. See below for more caveats.

We now consider varying each of the \( c_i \), and for each variation we set

$$\frac{\partial \langle E \rangle}{\partial c_i} = 0. \quad (24)$$

(Extremizing \( \langle E \rangle \) subject to the normalization constraint is equivalent to solving the time-independent Schroedinger equation.) This gives us a system of equations that, in our example, looks like

$$\begin{pmatrix} X & 1 & 0 \\ 1 & X & 1 \\ 0 & 1 & X \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0, \quad (25)$$

where

$$X \equiv \frac{\alpha - \langle E \rangle}{\beta}. \quad (26)$$

The matrix is called the Huckel matrix. For there to be a solution to this equation, the determinant of the Huckel matrix must equal zero. Notice that for a molecule with \( N \) atoms this results in an \( N \)th order polynomial equation, and \( N \) resulting values for \( X \) (and thus \( \langle E \rangle \)). The remaining constraint that we must remember is the normalization condition for \( |\Psi\rangle \).

For any given \( X \) that is a solution, we can solve for the \( c_i \) and construct a trial wavefunction called a Huckel molecular orbital. One can then look at the molecular energy levels, count valence electrons, and figure out what the occupancy of the of the orbitals is.

### 3.1.3 Caveats and the take-home message

This approach is not what’s currently used by people calculating electronic structures today! It omits many things, most notably the antisymmetry condition on the total wavefunction. It also
parameterizes the electronic interactions into $\alpha$ and $\beta$ terms. This approach is equivalent to a parameterized Hartree method. It’s reliance on independent electrons is its biggest weakness.

Far more sophisticated methods exist today that incorporate the Pauli principle (Hartree-Fock, for example). Similarly, parametrizing the matrix elements above has now largely been replaced by empirical interaction terms (on the one hand) or direct first-principles calculations (on the other). Using approaches like HF and density functional theory, today far more realistic and quantitatively accurate calculations may be done than is possible with simple Huckel theory.

The Huckel approach does get certain qualitative trends correct, though:

- If Huckel is performed on large molecules, one finds that the predicted energy levels cluster into groups. The number of such groups goes like the number of orbitals included in the calculation.
- As overlap between adjacent atoms is increased, the width in energy of those groups of levels increases.
- As the number of atoms in the molecule increases toward infinity (the bulk limit), the level clusters evolve into bands of allowed energies separated by regions of energy containing no allowed states.

### 3.2 A tight-binding picture

For small numbers of atoms, one can consider the different types of bonding that can occur. Which actually takes place depends on the details of the atoms in question:

- Van der Waals bonding. This doesn’t involve any significant change to the electronic wavefunctions; atom A and atom B remain intact and interact via fluctuating electric dipole forces. Only included on this list for completeness.

- Ionic bonding. For Coulomb energy reasons atom A donates a valence electron that is accepted by atom B. Atom A is positively charged and its remaining electrons are tightly bound to the ion core; atom B is negatively charged, and all the electrons are tightly bound. Atoms A and B “stick” to each other by electrostatics, but the wavefunction overlap between their electrons is minimal.

- Covalent bonding. As in the hydrogen case described above, it becomes more useful to describe the valence electrons in terms of molecular orbitals, where to some degree the valence electrons are delocalized over more than one ion core. In large molecules there tends to be clustering of energy levels with intervening gaps in energy containing no allowed states. There is a highest occupied molecular orbital (HOMO) and a lowest unoccupied molecular orbital (LUMO).

- Metallic bonding. Like covalent bonding only more extreme; the delocalized molecular orbitals extend over many atomic spacings.

Now let’s really get serious and consider very large numbers of atoms arranged in a periodic array, as in a crystal. This arrangement has lots of interesting consequences. Typically one thinks of the valence electrons as seeing a periodic potential due to the ion cores.
There are two broad approaches to calculating electronic structures. One of these is the “tight-binding” picture, reminiscent of the molecular orbital case described above. We’re interested in approximating the ground state of the many-body system of interacting electrons plus the potential due to the ion cores. In tight binding we consider our solution to be built out of superpositions of localized single-electron wave functions centered on the positions of the atoms in the solid. That is, if the atoms in the solid are at positions $r_n$, we try to write the ground state wave function as:

$$\psi(r) = \sum_{r_n} a(r_n) \phi(r - r_n),$$  \hspace{1cm} (27)

where $\phi(r)$ is a localized atomic orbital wave function centered on the origin. The periodic nature of the lattice potential in a crystal enters through restrictions on the $a_n$.

For a more complete description of the tight binding method, refer to many solid state books, such as Ashcroft and Mermin. For now, we only state the results and make a few qualitative observations. Suppose that the full potential energy at location $r$ is given by the sum of two pieces:

$$U(r) = U_{\text{atomic}}(r) + \delta U(r),$$ \hspace{1cm} (28)

where the first term is the potential due to the atom located at $r$, and $\delta U$ is the potential energy at that site due to all the other atoms in the solid. Suppose further that the overlap between orbitals on adjacent atoms is small. Consider a very simple case involving only a single atomic orbital for each atom $\phi(r)$ (for example, the 1s orbital). Then we can define some parameters:

$$\alpha(R) \equiv \langle \phi(r) | \phi(r - R) \rangle,$$  \hspace{1cm} (29)

$$\beta \equiv -\langle \phi(r) | \delta U | \phi(r) \rangle,$$ \hspace{1cm} (30)

$$\gamma(R) \equiv \langle \phi(r) | \delta U | \phi(r - R) \rangle.$$  \hspace{1cm} (31)

Here we’ve kept the $r$ label to remind you that the $\phi(r)$ are atomic orbitals centered on a particular location.

This is essentially a variation on the Huckel approach. The idea is that these parameters may be calculated for pairs of atoms within a particular model. The main result in the case of a crystalline solid ends up giving us the relationship between energy and wavevector. For the simple case above,

$$E(k) = E_\phi - \frac{\beta + \sum_{nn} \gamma(R) e^{ikR}}{1 + \sum_{nn} \alpha(R) e^{ikR}}.$$  \hspace{1cm} (32)

The first term is the energy eigenvalue of the particular atomic orbital we’re using. The sums are over nearest-neighbors, if we assume that the overlap is small enough that other terms are negligible. If we assume only using $s$ orbitals, and remember that the $\alpha$ terms in the denominator are small, we can find:

$$E(k) \approx E_\phi - \beta - \sum_{nn} \gamma(R) \cos k \cdot R.$$  \hspace{1cm} (33)

Note that what had been a single energy level $E_\phi$ is now broadened out into a band of allowed energies, with the band width being given by $\gamma$. Notice, too, how this expression labels states with a variable $k$, like the wavevector in the free electron gas case we talked about before. However, it’s in the form of corrections to the atomic energy levels used in the calculation. We shall see below that the energy spectrum of the nearly-free electron model looks even closer to that of the free electron case.

General remarks:
• When allowing more than one kind of atomic orbital per atom in a tight-binding calculation, Eq. 32 becomes a matrix equation. Using three $p$-orbitals means a three-by-three matrix; five $d$-orbitals means a five-by-five matrix, etc.

• The bandwidth is proportional to the overlap integrals $\gamma$. This means that higher orbitals, which tend to be larger, lead to broader bands. Once the overlap of orbitals is substantial over an atomic spacing, our treatment of the tight binding parameters is of questionable validity.

• Complicated crystal structures (and hence nearest neighbor arrangements) in solids can split a single atomic orbital into more than one band.

We may return to the tight binding method again later. For now we will concentrate on what is usually described as the *nearly free electrons* picture. Here we treat the lattice potential as a perturbation on the free electron gas that we’ve been discussing. To worry about bulk properties for now, we’ll ignore the edges of our crystal by imposing periodic boundary conditions. Clearly this won’t work well if the material in question is truly of nanoscale size, such that most of the participating atoms are on the boundary. We’ll deal with the validity of this approximation later, noting that a rigorous proof is nontrivial.

### 3.3 Nearly free electrons: Band theory

Clearly our infinite square well model of the potential seen by the electrons is an oversimplification. When the underlying lattice structure of (crystalline) solids is actually included, the electronic structure is a bit more complicated, and is typically well-described by *band theory*.

When solving the Schrödinger equation for this situation, the eigenfunctions are plane waves (like our old free Fermi gas case) multiplied by a function that’s periodic with the same period as the lattice:

$$\psi_k(r) = u_k(r) \exp(ik \cdot r),$$

$$u_k(r) = u_k(r + r_n).$$  \hspace{1cm} (34)

These wavefunctions are called *Bloch waves*, and like the free Fermi gas wavefunctions are labeled with a wavevector $k$. See Fig. 3.3

The really grungy work is two-fold: finding out what the the function $u_k(r)$ looks like for a particular arrangement of particular ion cores, and figuring out what the corresponding allowed energy eigenvalues are. In practice this is done by a combination of approximations and numerical techniques. It turns out that while getting the details of the energy spectrum right is extremely challenging, there are certain general features that persist.

First, not all values of the energy are allowed. There are *bands* of energy for which Bloch wave solutions exist, and between them are *band gaps*, for which no Bloch wave solutions with real $k$ are found. Plotting energy vs. $k$ in the 1d general case typically looks like Fig. 3.3.

The details of the allowed energy bands and forbidden band gaps are set by the interaction of the electrons with the lattice potential. In fact, looking closely at Fig. 3.3 we see that the gaps really “open up” for Bloch waves whose wavevectors are close to harmonics of the lattice potential. The Coulomb interactions between the electrons only matter here in the indirect sense that the electrons screen the ion cores and self-consistently contribute to the periodic potential.
Figure 4: The components of a Bloch wave, and the resulting wavefunction. From Ibach and Luth.

Figure 5: Allowed energy vs. wavevector in general 1d periodic potential, from Ibach and Luth.
Figure 6: Filling of allowed states in two different systems. On the left, the electrons just fill all the states in the valence band, so that the next unoccupied state is separated by a band gap; this is an insulator. On the right, the electrons spill over into the conduction band, leaving the system metallic.

Figure 7: More possibilities. On the left is a band insulator, as before. Next is an intrinsic semiconductor, followed by two doped semiconductors.

Now we consider dropping in electrons and ask what the highest occupied single-particle states are, as we did in the free Fermi gas case. The situation here isn't too different, though the properties of the ground state will end up depending dramatically on the band structure. Notice, too, that the Fermi sea is no longer necessarily spherical, since the lattice potential felt by the electrons is not necessarily isotropic.

Figure 3.3 shows two possibilities. In the first, the number of electrons is just enough to exactly fill the valence band. Because there is an energy gap to the nearest allowed empty states, this system is a band insulator: if an electric field is applied, the Fermi surface can't shift around as in Fig. 2 because there aren't available allowed states. Therefore the system can't develop a net current in response to the applied field. A good example of a band insulator is diamond, which has a gap of around 10 eV. The second case represents a metal, and because of the available states near the Fermi level, it can support a current in the same way as the ideal Fermi gas in Fig. 2.

Other possibilities exist, as shown in Fig. 3.3. One can imagine a band insulator where the gap is quite small, so small that at room temperature a detectable number of carriers can be promoted from the valence band into the conduction band. Such a system is called an intrinsic semiconductor. A good example is Si, which has a gap of around 1.1 eV, and a carrier density at room temperature of around $1 \times 10^{10} \text{ cm}^{-3}$.

Further, it is also possible to dope semiconductors by introducing impurities into the lattice. A donor such as phosphorus in silicon can add an electron to the conduction band. At zero temperature, this electron is bound to the P donor, but the binding is weak enough to be broken...
at higher temperatures, leading to usual electronic conduction. This is the third case shown in Fig. 3.3. Similarly, an acceptor such as boron in silicon can grab an electron out of the valence band, leaving behind a positively charged hole. This hole acts like a carrier with charge +e; at zero temperature it is weakly bound to the B acceptor, but at higher temperatures it can be freed, leading to hole conduction. One more exotic possibility, not shown, is semimetallic behavior, as in bismuth. Because of the funny shape of its Fermi surface, parts of Bi’s valence band can have holes at the same time that parts of Bi’s conduction band can have electrons (!).

3.3.1 Dispersion relations and k-space

A full discussion of these issues is covered in numerous solid state physics texts, including Kittel, Ashcroft and Mermin, Goodstein, etc. Here we present only a brief introduction to the key concepts. We start with a discussion of the 1d case, but the main ideas all generalize to higher dimensions.

We’ve already seen from our discussion of the ideal Fermi gas that it is useful to plot filled and empty states in k-space. Since our Bloch waves are labeled by a parameter k, this continues to be a useful point of view. A related type of plot is that of dispersion relations, plotting allowed energy as a function of k. For an ideal free 1-d Fermi gas, this is simply a parabola centered on k = 0, as shown by the bold line in Fig. 3.3. The functional form of this curve is $E = \frac{\hbar^2 k^2}{2m}$. All energies are allowed, and for every value of k there is a single allowed energy.

In the nearly free Fermi gas, the periodicity of the lattice potential introduces a complication. We’ve already said that for real k not all energies are allowed in this case. How does this manifest itself? Suppose the lattice parameter is a. For k-values near integer multiples of $2\pi/a$ (so-called reciprocal lattice vectors), the allowed single-electron states approach standing waves rather than travelling waves. Put another way, electrons with wave vectors equal to integer multiples of $2\pi/a$ Bragg scatter (diffract) off the lattice potential.

The result is shown in Fig. 3.3, and the key observations to make is: The lattice potential opens up gaps at particular values of the energy, so that not all energy values correspond to allowed single-particle states. Now, the ensemble of actual allowed wave functions will not change if we shift our origin in k-space by a reciprocal lattice vector. This means we can effectively plot all the allowed states in what’s called a reduced zone scheme. The part of k-space from $-\pi/a$ to $\pi/a$ is the first Brillouin zone for this 1d example. The range $-2\pi/a$ to $\pi/a$ and $\pi/a$ to $2\pi/a$ is the second zone, and so forth. Instead of allowing k to range over all k-space, we account for all the allowed single-particle states by restricting k to the first zone and having multiple bands. As we stated above we can label the allowed single particle states by a value of k (limited to the first zone) and a band index.

This illustration was for a simple, infinite, periodic lattice potential in 1d. For higher dimensionalities and more complicated crystal structures the qualitative features (bands, gaps at zone boundaries) remain the same, but important details can change, as we’ll discuss below. Defining zone boundaries when bond lengths in different directions are different is a subject covered in generality in most solid state textbooks. There will be a handout distributed showing the way points within the zone are defined.

Notice now that while the existence of Bloch waves means it is possible to label each eigenstate with a wavevector k, the energy of that state may not depend quadratically on k as in Eq. (5). The approximation that $E(k) \sim k^2$ is called the effective mass or parabolic bands.
approximation. As you might expect, the effective mass is defined by:

\[
E(k) = \hbar^2 \left( \frac{k_x^2}{2m_x} + \frac{k_y^2}{2m_y} + \frac{k_z^2}{2m_z} \right),
\]

where we’re explicitly showing that the effective mass \( m^* \) isn’t necessarily isotropic. Note that the effective mass is not restricted to be positive!

We can define a parameter called crystal momentum for our single-particle states, given as \( \hbar k \), in analogy to the momentum for plane waves. One important difference: crystal momentum is only conserved to within a reciprocal lattice vector! Physically, what this means is that the lattice can exchange momentum freely with the single-particle state in chunks of size \( 2\pi/a \) (for our 1d example). This is just another way of saying that the electrons experience Bragg diffraction off the lattice potential. Total momentum is, of course, conserved.

Detailed crystal structure can have a more profound impact than simply altering or introducing anisotropy into the effective mass. Consider our 1d case from above. If we consider bands 2 and 3 as “valence” and “conduction” bands, respectively, we see that the valence band maximum and the conduction band minimum both occur at \( k = 0 \). This situation is called a direct gap. In this case the band gap is defined as the energy difference between the valence band maximum and the conduction band minimum both at \( k = 0 \). An example of a direct gap material is GaAs. A direct gap means that the minimum energy transition between the bands can occur without some additional excitation to account for momentum conservation.

Conversely, silicon is an indirect gap system. The valence band maximum does occur at \( |k| = 0 \), but there are six minima in the conduction band, and those minima are displaced from \( |k| = 0 \). The presence of these energetically equivalent minima is called the “valley degeneracy”. The upshot of this is that bulk silicon has an indirect gap - the minimum energy transition from the valence to the conduction band requires a substantial \( \Delta k \). This is profound technologically, and is the reason that optoelectronic devices like light emitting diodes are not made from pure Si! Even more interesting for our purposes: in silicon nanoparticles and in nanoporous silicon, the change in crystal structure from the bulk can turn Si from an indirect to a direct gap.

### 3.3.2 Work functions

There is one more energy scale in the electronic structure of real materials that we explicitly ignored in our ideal Fermi gas model. The work function, \( \Phi \), is defined as the energy difference between the vacuum (a free electron outside the sample) and the Fermi energy, \( E_{\text{vac}} - E_F \). Our toy infinite square well model artificially sets \( \Phi = \infty \). Unsurprisingly, \( \Phi \) also depends strongly on the details of the material’s structure, and can vary from as low as 2.4 eV in Li to over 10 eV in band insulators. From a tight-binding perspective, one can see where \( \Phi \) comes from, qualitatively: it evolves from the single-atom ionization energy.

### 3.3.3 Capsule summary

Big picture ideas to take away from this subsection are:

- Bonding in small systems is crucially affected by electronic binding energies and Coulomb interactions.
- Large systems typically have energy level distributions well-described by bands and gaps.
Figure 8: Surface states on copper, imaged by Crommie and Eigler at IBM with a scanning tunneling microscope.

- Single-particle eigenstates in systems with periodicity are Bloch waves that can be labeled by a wavevector $k$.
- Whether a system is conducting, insulating, or semiconducting depends critically on the details of its band structure, including the number of available carriers. Systems can exhibit either electronic or hole conduction depending on structure and the presence of impurities.
- The energy needed to actually remove an electron from a material to the vacuum also reflects the structure of that material.

3.4 Structural issues

The previous section deliberately neglected a number of what I’ll call “structural issues”. These include nonidealities of structure such as boundaries, impurities, and other kinds of structural disorder. Further, we’ve treated the ion cores as providing a static background potential, when in fact they can have important dynamics associated with them.

Let’s deal with structural “defects” first. We’ll treat some specific effects of these defects later on. First, let’s ask what are the general consequences of not having an infinite perfect crystal lattice.

*Bloch waves infinite in extent are no longer exact single-particle eigenstates of the system.* That’s not necessarily a big deal. Intuitively, an isolated defect in the middle of a large crystal isn’t going to profoundly alter the nature of the entire electronic structure. In fact, the idea that there are localized electronic states around the ion cores and there can be delocalized (extended) states which span many atomic spacings is still true even without any lattice at all; this happens in amorphous and liquid metals.

*Special states can exist at free surfaces.* Unsurprisingly, these are called surface states. The most famous experimental demonstration of this is shown in Fig. 3.4. Suppose the
surface is the $x-y$ plane. Because of the binding energy of electrons in the material, states exist which are pinned to the surface, having small $z$ extent and wavelike (or localized, depending in surface disorder) character in $x$ and $y$. Surface states can have dramatic implications when (a) samples are very small, so that the number of surface states is comparable to the number of “bulk” states; and (b) the total number of carriers is very small, as in some semiconductors, so that an appreciable fraction of the carriers can end up in surface states rather than bulk states. This latter case is the reason for “depletion layers” near the surfaces of doped semiconductors; the doped carriers occupy (usually localized) surface states, leading to a reduced carrier density near free surfaces.

Interfaces between different materials can also produce dramatic effects. The boundary between a material and vacuum is just the limiting case of the interface between two materials with different work functions. When joining two dissimilar materials together there are two conditions one has to keep in mind:

(a) The “vacuum” for materials A and B is the same, though $\Phi_A \neq \Phi_B$ in general. That means that prior to contact the Fermi levels of A and B are usually different.

(b) Two systems that can exchange particles are only in equilibrium once their chemical potentials are equal. That implies that when contact is made between A and B, charge will flow between the two to equalize their Fermi levels.

The effect of (a) and (b) is that near interfaces “space charge” layers can develop which bend the bands to equalize the Fermi levels across the junction. The details of these space charge layers (e.g. how thick are they, and what is the charge density profile) depend on the availability of carriers (is there doping?) and the dielectric functions of the two materials. In general one has to solve Poisson’s equation self-consistently while considering the details of the band structure of the two materials.

A term related to all this is a Schottky barrier. It is possible to have band parameters of two materials be such that the space charge layer which forms upon their contact can act like a substantial potential barrier to electronic transport from one material to the other. For example, pure Au on GaAs forms such a barrier. These barriers have very nonlinear $I - V$ characteristics, in contrast to “Ohmic” contacts between materials (an example would be In on GaAs). A great deal of semiconductor lore exists about what combinations of materials form Schottky barriers and what combinations form Ohmic contacts.

We also need to worry about the dynamics of the ion cores, rather than necessarily treating them as a static background of positive charge. The quantized vibrations of the lattice are known as phonons. We won’t go into a detailed treatment of phonons, but rather will highlight some important terminology and properties.

A unit cell is the smallest grouping of atoms in a crystal that exhibits all the symmetries of the crystal and, when replicated periodically, reproduces the positions of all the atoms.

The vibrational modes of the lattice fall into two main categories, distinguished by their dispersion curves, $\omega(q)$, where $q$ is the wavenumber of the wave. When $q = 0$, we’re talking about a motion such that the displacements of the atoms in each unit cell is identical to those in any other unit cell. Acoustic branches have $\omega(0) = 0$. There are three acoustic branches, two transverse and one longitudinal. Optical branches have $\omega(0) \neq 0$, and there typically three optical branches, too. The term “optical” is historic in origin, though optical modes of a particular $q$ are typically of higher energy than acoustic modes with the same wavenumber.
For our purposes, it’s usually going to be sufficient to think of phonons as representing a bath of excitations that can interact with the electrons. Intuitively, the coupling between electrons and phonons comes about because the distortions of the lattice due to the phonons show up as slight variations in the lattice potential through which the electrons move. Electron-phonon scattering can be an important process in nanoscale systems, as we shall see.

The Debye model of phonons does a nice job at describing the low energy behavior of acoustic modes, which tend to dominate below room temperature. The idea is to assume a simple dispersion relation, \( \omega = v_{L,T} q \), where \( v \) is either the longitudinal or transverse sound speed. This relation is assumed to hold up to some high frequency (short wavelength) cutoff, \( \omega_D \), the Debye frequency, set by the requirement that the total number of acoustic modes = \( 3rN_{uc} \), where \( r \) is the number of atoms per unit cell, and \( N_{uc} \) is the number of unit cells in the solid. Without going into details here, the main result of Debye theory for us is that at low temperatures \( T < T_D \), the heat capacity of 3d phonons varies as \( T^3 \).

One more dynamical issue: for extremely small structures, like clusters of tens of atoms, figuring out the equilibrium positions of the atoms requires a self-consistent calculation that also includes electronic structure. That is, the scales of electronic Coulomb contributions and ionic displacement energies become comparable. Electronic transitions can alter the equilibrium conformations of the atoms in such systems.

The important points of this subsection are:

- Special states can exist at surfaces, and in small or low carrier density systems these states can be very important.
- Interfaces between different materials can be very complicated, involving issues of charge transfer, band bending, and the possible formation of potential barriers to transport.
- The ion cores can have dynamical degrees of freedom which couple to the electrons, and the energy content of those modes can be strongly temperature dependent.
- In very small systems, it may be necessary to self-consistently account for both the electronic and structural degrees of freedom because of strong couplings between the two.

### 3.5 Interactions?

We have only been treating Coulomb interactions between the electrons indirectly so far. Why have we been able to get away with this? As we shall see, one key is the fact that our electronic Bloch waves act so much like a cold \( (T < T_F \equiv E_F/k_B) \) Fermi gas. Another relevant piece of the physics is the screening of point charges that can take place when the electrons are free to move (e.g. particularly in metals).

Let’s look at that second piece first, working in 3d for now. Suppose there’s some isolated perturbation \( \delta U(r) \) to the background electrical potential seen by the electrons. For small perturbations, we can think of this as causing a change in the local electron density that we can find using the density of states:

\[
\delta n(r) = \nu_{3d}(E_F)|e|\delta U(r).
\]

Now we can use Poisson’s equation factor in the change in the potential caused by the response of the electron gas:

\[
\nabla^2(\delta U(r)) = \frac{e}{\epsilon_0}\delta n(r) = \nu_{3d}(E_F)|e|\delta U(r).
\]
In 3d, the solution to this is of the form $U(r) \sim (1/r)(\exp(-\lambda r))$, where $\lambda \equiv 1/r_{TF}$, defining the Thomas-Fermi screening length:

$$r_{TF} = \left( \frac{e^2}{\epsilon_0 \nu_{3d}(E_F)} \right)^{-1/2}. \quad (38)$$

If we plug in our results from Eqs. (9,9) for the free Fermi gas in 3d, we find

$$r_{TF} \simeq 0.5 \left( \frac{n}{a_0} \right)^{-1/6}, \quad (39)$$

where $a_0$ is the Bohr radius. Plugging in $n \sim 10^{23} \text{ cm}^{-3}$, typical for a metal like Cu or Ag, we find $r_{TF} \sim 1 \text{ Å}$. So, the typical lengthscale for screening in a bulk metal can be of atomic dimensions, which explains why the ion cores are so effectively screened.

Looking at Eq. (38) it should be clear that screening is strongly affected by the density of states at the Fermi level. This means screening in band insulators is extremely poor, since $\nu \sim 0$ in the gap. Further, because of the changes in $\nu$ with dimensionality, screening in 2d and 1d systems is nontrivial.

Now let’s think about electron-electron scattering again. We have to conserve energy, of course, and we have to conserve momentum, though because of the presence of the lattice in certain circumstances we can “dump” momentum there. The real “crystal momentum” conservation that must be obeyed is:

$$k_1 + k_2 = k_3 + k_4 + G, \quad (40)$$

where $G$ is a reciprocal lattice vector. For zero temperature, there just aren’t any open states the electrons can scatter into that satisfy the conservation conditions! At finite temperature, where the Fermi distribution smears out the Fermi surface slightly, some scattering can now occur, but (a) screening reduces the scattering cross-section from the “bare” value; and (b) the Pauli principle reduces it further by a factor of $(k_B T/E_F)^2$. This is why ignoring electron-electron scattering in equilibrium behavior of solids isn’t a bad approximation. We will come back to this subject soon, though, because sometimes this “weak” scattering process is the only game in town, and can have profound implications for quantum effects.

The full treatment of electron-electron interactions and their consequences is called \textit{Landau Fermi Liquid Theory}. We won’t get into this in any significant detail, except to state some of the main ideas. In LFLT, we consider starting from a noninteracting Fermi gas, and adiabatically turn on electron-electron interactions. Landau and Luttinger argued that the ground state of the noninteracting system smoothly evolves into the ground state of the interacting system, and that the excitations above the ground state also evolve smoothly. The excitations of the noninteracting Fermi gas were (electron-above-$E_F$, hole-below-$E_F$), so-called particle-hole excitations. In the interacting system, the excitations are \textit{quasiparticles} and \textit{quasiholes}. Rather than a lone electron above the Fermi surface, a LFLT quasiparticle is such an electron plus the correlated rearrangement of all the other electrons that then takes place due to interactions.

This correlated rearrangement of the other electrons is called \textit{dressing}. An electron injected from the vacuum into a Fermi liquid is then “dressed” to form a quasiparticle. The effect of these correlations is to renormalize certain quantities like the effective mass and the compressibility. The correlations can be quantified by a small number of “Fermi liquid parameters” that can be evaluated by experiment. Other than these relatively mild corrections, quasiparticles usually act very much like electrons.
Note that this theory can break down under various circumstances. In particular, LFLT fails if a new candidate ground state can exist that has lower energy than the Fermi liquid ground state, and is separated from the LFLT ground state by a broken symmetry. The symmetry breaking means that the smooth evolution mentioned above isn’t possible. The classic example of this is the transition to superconductivity.

In 1d systems with strong interactions, Fermi liquid theory can also break down. The proposed ground state in such circumstances is called a Luttinger liquid. We will hopefully get a chance to touch on this later.

The main points here are:

- Screening depends strongly on $\nu(E_F)$, and so on the dimensionality and structure of materials.
- Electron-electron scattering is pretty small in many circumstances because of screening and the Pauli principle.
- LFLT accounts for electron-electron interactions by dressing the bare excitations.

### 3.6 Transitions and rates

Often we’re interested in calculating the rate of some scattering process that takes a system from an initial state $|i\rangle$ takes it to a final state $|f\rangle$. Also, often we’re not really interested in the details of the final state; rather, we want to consider all available final states that satisfy energy conservation. The result from first order time-dependent perturbation theory is often called Fermi’s Golden Rule. Perturbation theory is reviewed and this result is derived in the appendix provided. If the potential associated with the scattering is $V_s$, the rate is approximately:

$$S = \frac{2\pi}{\hbar}|\langle i|V_s|i\rangle|^2\delta(\Delta E),$$

where the $\delta$-function (only strictly a $\delta$-function as the time after the collision $\to \infty$) takes care of energy conservation. If we want to account for all final states that satisfy the energy condition, the total rate is given by the integral over energy of the right hand side of Eq. (41) times the density of states for final states.

The important point here is, as we mentioned above, the density of states plays a crucial role in establishing transition and relaxation rates. One good illustration of this in nanoscale physics is the effect of dimensionality on electron-phonon scattering in, for example, metallic single-walled carbon nanotubes. Because of the peculiarities of their band structure, these objects are predicted to have one-dimensional densities of states. There are only two allowed momenta, forward and back, and this reduction of the Fermi surface to these two points greatly suppresses the electron-phonon scattering rate compared to that in a bulk metal. As a result, despite large phonon populations at room temperature, ballistic transport of carriers is possible in these systems over micron distances. In contrast, the scattering time due to phonons in bulk metals is often three orders of magnitude shorter.

Often this simple perturbation theory picture gives important physical insights. More sophisticated treatments include diagrammatic expansions, equivalent to higher-order perturbation calculations.

One other thing to bear in mind about transition rates. If multiple independent processes exist, and their individual scattering events are uncorrelated and well-separated in time, it is
usually safe to add rates:

\[
\frac{1}{\tau_{\text{tot}}} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \ldots
\]  

(42)

If a single effective rate can be written down, this implies an exponential relaxation in time, as we know from elementary differential equations. This is not always the case, and we’ll see one major example later of a nonexponential relaxation.

References

PHYS 533: Nanostructures and Nanotechnology I - Fall 2005

- Outline of the course
- Logistical details
- Introduction: what is nanotechnology, and why should you care?

Logistics

Prof. Douglas Natelson
Office: Space Sciences Bldg., Rm. 329
Phone: x3214
E-mail: natelson@rice.edu

Lectures: MWF 13:00 - 13:50

Problem sets: mostly weekly, handed out Wednesday due following Wednesday. (50% of grade)

Exams: none

Midterm paper (20% of grade)
Final paper (25% of grade)

Course webpage:
http://www.owlnet.rice.edu/~phys533
More logistics

**Text:** no specific text. There will be some written lecture notes, as well as handouts based on a number of books. Also a large number of papers.

**Goal:** to make you literate about the physics of nanostructures and their current and future roles in technology, to the point where you at least know what to consider and where to look for more information.

General outline of the course

- Review of solid state physics
  - Why is nano different than bulk?
  
- Some fabrication, some materials

- Electronic transport

- Industrial practice and needs: electronics

- Nanoelectronics + molecular electronics

- Basics of magnetism

- Industrial practice and needs: magnetism

- Nanomagnetism and spintronics
Nanostructures and Nanotechnology

“There’s plenty of room at the bottom!”

- R.P. Feynman, 1959

- Materials behave differently when structured at the nm scale than they do in bulk.
- Technologies now exist that allow engineering of materials on the nanometer scale.
- Nanostructured materials are useful now, and will be phenomenally more so in the near future.
- Fundamental overlap between physics, chemistry, biology.
- Current situation may be similar to that of electronics in ~ 1955.

Nanostructure: a piece of material with at least one dimension less than 100 nm in extent.

Sizes of things

- Diameter of human hair
- Diameter of red blood cell
- Visible light wavelengths
- Intel’s newest transistor
- Diameter of DNA, nanotubes
- Bohr diameter

Hair image from Pattern in Nature site at LSU; blood cell from osmosis webpage at Maricopa.edu
Where are we now?  Where are we going and why?

- Microelectronics
- Magnetoelectronics
- Photonics
- Micromechanics
- Bio-nano interaction
- Nano- and molecular electronics
- Spintronics
- Nanophotonics
- Micromechanics
- Bio-nano interaction

**Leaving out:** much chemistry, polymer science, molecular biology

**Microelectronics: Moore’s (1st) law**

Typical lateral feature size on an integrated circuit is reduced in size by a factor of 2 roughly every 18 months.

Data from Intel’s website
Microelectronics

Typical metal-oxide-semiconductor field-effect transistor (MOSFET):

- Features defined by photolithography, etching, and controlled diffusion of impurities.
- Channel length $L \sim 90$ nm.
- Oxide thickness $t \sim 2$ nm.

State-of-the-art, 2004-5:
- $L \sim 65$ nm, $t \sim 1.2$ nm.
- 2007: $L \sim 45$ nm, “high-K”

Problems coming:
- Tunneling, leakage, $t \rightarrow 0$, finite size effects

TEM images from Intel’s website

Photonics: LEDs

Ability to grow materials like GaAs and AlGaAs with atomic precision means engineering of properties to tailor optical emission.

Example: double heterostructure LED.

Figure from SPIE, OE Reports, December 2000.
Photonics: Solid State Lasers and Bragg reflectors

Diode lasers
Dielectric mirrors

Quantum cascade lasers
Band-gap engineering

Magnetoelectronics: giant magnetoresistance (GMR)

- Sub-nm normal metal layer thickness drastically controls magnetic properties of stack.
- Resistance of (I) is ~50% lower than (II).
- Discovered in 1988; in consumer products by 1993!
Bio-nano: DNA manipulation

Molecular biologists have developed many tools for manipulating DNA, deoxyribonucleic acid, and RNA, ribonucleic acid.

With specialized macromolecules like DNA, fine line between biochemistry and nanoengineering.

- Can now purchase sequences made to order!
- Can functionalize inorganic surfaces to allow DNA binding.
- Can pick out specific sequences and bulk produce them at the expense of others (PCR).

Bio-nano: “gene chip”

Images from Affymetrix.
Tools: lithography

- Basic patterning technique: “drawing” in polymer followed by chemical development.
- Photolithography can make ~ 80 nm features using phase contrast masks and chemically amplified resists.
- Electron beam lithography can make ~ 20 nm features.

Tools: atomic force microscopy

- AFM (1986) allows imaging of surface topography with sub-nm $z$ resolution, and lateral resolution (usually) limited by radius of curvature of tip.
- With single-walled nanotube tips, resolution ~ 1nm is possible!
- Under certain conditions, atomic resolution may be achieved.
- Surface manipulation also possible through mechanical force or electrochemistry.
Tools: scanning tunneling microscopy

- STM (1986) allows imaging of surface topography of conductors with atomic resolution.
- Surface manipulation also possible.
- May also be used as atomic-resolution probe of local electronic states.

Tools: material deposition

- Several tools for atomic-scale layer deposition of material.
- Molecular Beam Epitaxy is a great example.
Where are we going?

• Ideas / lab demonstrations now
• Potential industrial importance over the next ten years.
• What are the drivers?

➢ New physics makes old methods no longer apply.
➢ New physics allows advantages in new methods.
➢ Economics!

Nanoelectronics: single-electron devices

• Possible to make transistors switched by single electrons.
• Coulomb blockade physics is relevant.
• Requires very small sizes (few nm) to get behavior at 300 K.
Nanoelectronics: molecular electronics

- Proposes (Aviram & Ratner, 1974) to use small numbers of molecules as active circuit elements.
- A number of challenges: the wiring problem, speed, low currents, sensitivity to impurities.
- Some recent progress:

![Image from IBM.]

![Image from McEuen group @ Cornell.]

![Image from Ralph group @ Cornell.]

Photonics: photonic band gaps

- Nanometer control of materials makes possible 3d manipulation of light.
- Applications include optical waveguides and switching.

![Image from Sandia National Lab.]

![Image from Joannopoulos at MIT.]

Photonics: Si laser (!)

- Potential for Si-based optical interconnects!
- Still requires large optical pump.

Spintronics: spin transistor

- Much interest in using spin of electrons as well as their charge.
- Relevant question: how far does spin polarization survive?
Spintronics: spin Hall effect

Possibility of dissipationless spin currents!

Bio-nano: interface with nanomaterials

- New nanomaterials may be used in vitro for medical purposes.

Nanoshells as cancer therapy.

CdS nanocrystals as fluorophores for medical research.

- Interfacing biomolecules with inorganic electronic materials.
New tools: nanomaterials

- Titania particles for catalysis in environmental cleanup.
- Nanotube fibers for light-strong construction.

![Image from Colvin group.](image1)

![Image from Colvin group.](image2)

New tools: self-assembly

Self-organization may be one key to mass fabrication at the nanometer scale.

Dysprosium Silicide nanowires on Si

Lamellar structure of diblock copolymer

![Dysprosium Silicide nanowires on Si](image3)

![Lamellar structure of diblock copolymer](image4)

*Image from Stan Williams at HP Labs.*

*Image from Ulrich Wiesner at Cornell.*
Conclusions

- Nanostructured systems are already having a dramatic impact on our lives.
- As control over materials at the nm scale improves and more fundamental science is done, nanoscience has the potential to drastically change our lives, on a par with the electronics revolution of the last 60 years.
Ways of treating electrons in nanoscale materials:

“Bottom-up” – start by considering electrons in single atoms and small molecules

Chemistry

Huckel theory

“Tight-binding”

Localized orbitals

“Top-down” – free electrons

Physics / electrical engineering

Free Fermi gas

Nearly-free electrons

Bloch waves

We’ll start with top-down approach first.

• Not bad for many metals and doped semiconductors

• Shows qualitative features that hold true in detailed treatments.

• Successive approximations:

  • “Free particles” – no external potential

  • Independent electron approximation – ignores interactions. Assumes many-particle system can be modeled by starting from single-particle case.

Will start here and see when we have to make adjustments.

First, find allowed single-particle states and energies....
Time-independent Schroedinger equation (TISE):

\[-\frac{\hbar^2}{2m}\nabla^2 \psi + V(r)\psi = E\psi\]

Must solve, consistent with boundary conditions. Free particle means \( V = 0 \). Solutions are of the form

\[\psi \sim Ae^{ikr} + Be^{-ikr}\]

These are traveling waves with wavevector \( k \).

\[|k| = \frac{2\pi}{\lambda}, \text{ where } \lambda \text{ is wavelength.}\]

Energies are then

\[E(k) = \frac{\hbar^2 k^2}{2m}\]

Dispersion relation

No restrictions here on allowed values of \( k \) or \( E \).

Now try a 1d particle in a box potential:

\[V = \begin{cases} 0, 0 < x < L \\ \infty, x \leq 0; x \geq L \end{cases}\]
Interior of box, $V = 0$, so solution must look like superposition of plane waves, but $\Psi$ must vanish at walls. Answer:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{\pi}{L} nx$$

$n = 1, 2, 3, \ldots$

So, allowed $k$ values are

$$k = \frac{n\pi}{L}, n = 1, 2, 3, \ldots$$

meaning allowed energy values are

$$E(k) = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2}$$

Finite sample size drastically alters allowed energy levels!
Main observations:

- Energy spectrum is now discrete rather than continuous.
- Allowed wavevectors are uniformly spaced in \( k \)-space with a separation of \( \pi/L \).
- Sample size \( L \) determines spacing of allowed wavevectors and single-particle energies, with a smaller box giving larger spacings.

What about a 2d box of side \( L \)?

\[
\psi_{n_x,n_y}(x) = \frac{2}{L} \left( \sin \frac{\pi}{L} n_x x \right) \left( \sin \frac{\pi}{L} n_y y \right)
\]

\[
k_x \equiv \frac{n_x \pi}{L}, n_x = 1, 2, 3, \ldots
\]

\[
k_y \equiv \frac{n_y \pi}{L}, n_y = 1, 2, 3, \ldots
\]

Allowed \( k \) values now spaced uniformly in 2d \( k \)-space, one allowed value per “area” \( (\pi/L)^2 \).

\[
E(k) = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2) = \frac{\hbar^2 \pi^2}{2mL^2} (n_x^2 + n_y^2)
\]
Finite sample size still alters allowed energy levels!

Dispersion relation plot
($E$ vs $k$)

Energy level diagram
(allowed $E$ values)

$k$-space plot
(allowed $k$ values)

You can see where this is going for the 3d case:

$$\psi_{n_x,n_y,n_z}(x) = \left(\frac{2}{L}\right)^{3/2} \left(\sin \frac{\pi}{L} n_x x\right) \left(\sin \frac{\pi}{L} n_y y\right) \left(\sin \frac{\pi}{L} n_z z\right)$$

$$k_x = \frac{n_x \pi}{L}, n_x = 1,2,3...$$

$$k_y = \frac{n_y \pi}{L}, n_y = 1,2,3...$$

$$k_z = \frac{n_z \pi}{L}, n_z = 1,2,3...$$

Allowed $k$ values now spaced uniformly in 3d $k$-space, one allowed value per $k$-space “volume” $(\pi/L)^3$.

$$E(k) = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} \left(k_x^2 + k_y^2 + k_z^2\right)$$
Non-interacting many-electron systems

- Interested in ground state of many-electron system.
- No interactions $\rightarrow$ many-body eigenstates should be linear combinations of products of single-particle eigenstates.
- Still need to obey Pauli principle $\rightarrow$ correct total wavefunction should be antisymmetric under exchange of any two particles.

Rigorous answer: many-body eigenstates should be linear combinations of *Slater determinants* built out of single-particle eigenstates.

Approximation (or shorthand) for now: start filling each single-particle state (“orbital”) from the lowest energy, each with one spin-up and one spin-down electron.

Consider the *ground state* of the 2d case. What is the energy of the highest occupied single-particle state, if we drop $N_{2d}$ electrons into the box?

Start filling allowed $k$-states from bottom up, two electrons per state (one spin-up, one spin-down).

Easiest to count in $k$-space.

$$N_{2d} = 2\times\left(\frac{L}{\pi}\right)^2 \int_{0}^{k_F} \frac{1}{2\pi} (2\pi k) dk = \frac{1}{2\pi} L^2 k_F^2$$

Using an integral to approximate the sum is valid for large $N$. 

For spin

Number of $k$ states per $k$-space area

$k$-space area occupied

Using an integral to approximate the sum is valid for large $N$. 

For spin

Number of $k$ states per $k$-space area

$k$-space area occupied
Fermi energy
\[ E_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2 (2m_{2d})}{2m} \]
Fermi momentum/velocity
\[ p_F = \hbar k_F = \hbar (2m_{2d})^{1/2} \]
\[ v_F = \frac{p_F}{m} \]

- Fermi energy and momentum increase with density of electrons!
- All single-particle states with \( E < E_F \) are occupied at \( T=0 \). These states are called the Fermi Sea. The set of points in \( k \)-space that divides empty and full states is the Fermi surface.
- Exactly how \( E_F \) depends on density depends drastically on dimensionality.

Let’s look at the 3d case.
\[ N_{3d} = 2 \times \left( \frac{L}{\pi} \right)^3 \int_0^{k_F} \frac{1}{8} (4\pi k^2) dk = \frac{1}{3\pi^2} L^3 k_F^3 \]

Fermi energy
\[ E_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2 (3\pi^2 n_{3d})^{2/3}}{2m} \]
Fermi momentum
\[ p_F = \hbar k_F = \hbar (3\pi^2 n_{3d})^{1/3} \]
Pauli exclusion says that there are no doubly occupied sets of quantum numbers.

Result: processes we care about (e.g. electrical conduction) can only involve transferring electrons from occupied to unoccupied single-particle states.

A useful quantity: the density of states, $v$.

\[ v(E) \, dE \, dS = \text{the number of allowed single-particle states with energies between } E \text{ and } E+dE, \text{in a length/area/volume of size } dS. \]

Put another way, from our expressions for $N(E)$, we can find

\[ \nu_d (E) = \left( \frac{\partial n_d (E)}{\partial E} \right) \]  

where $n_d$ is the spatial electron density in $d$ dimensions.

From this definition, we can find the spacing of single-particle levels in a piece of material! All other things being equal, a higher $\nu$ means levels are more closely-spaced.

\[ \nu_d (E) = \left( \frac{\partial n_d (E)}{\partial E} \right) \]

\begin{align*}
    d &= 1 & \nu_{1d} (E_F) &= \frac{1}{\pi} \left( \frac{2m}{\hbar^2} \right)^{1/2} E_F^{1/2} \\
    2 & & \nu_{2d} (E_F) &= \frac{m}{\pi \hbar^2} \\
    3 & & \nu_{3d} (E_F) &= \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} E_F^{1/2}
\end{align*}

Qualitative change in energy dependence as dimension is varied!
\[ \nu_{1d}(E_F) = \frac{1}{\pi} \left( \frac{2m}{\hbar^2} \right)^{1/2} \frac{1}{E_F^{1/2}} \]

Decreases with increasing \( E_F \), meaning single-particle levels get **farther apart** with increasing \( N \).

That’s just what we know from the 1d particle-in-a-box!

\[ \nu_{2d}(E_F) = \frac{m}{\pi \hbar^2} \]

Typical level spacing is independent of \( N \)!

The approximate degeneracy of 2d levels balances out the tendency for the 1d states to spread out.

\[ \nu_{3d}(E_F) = \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} E_F^{1/2} \]

Typical level spacing is drops with increasing \( N \)!

Let’s plug in some numbers.

Suppose we have 1 cm\(^3\) of sodium, and we only count the valence electrons (one per atom).

\[ n_{3d} \approx 2.65 \times 10^{22} \text{ electrons per cm}^3 \]

Use free electron mass, \( m = 9.1 \times 10^{-31} \text{ kg} \)

\[ E_F = 5.2 \times 10^{-19} \text{ J} = 3.2 \text{ eV} = 38000 \text{ K} \]

\[ \nu_{3d}(E_F) = 7.6 \times 10^{46} \text{ J}^{-1} \text{m}^{-3} \]

\( \Delta = \) single-particle level spacing

\[ = 1/ (\nu_{3d}(E_F) \text{ V}) \]

\[ = 1.3 \times 10^{-41} \text{ J} \]

The single-particle levels are essentially continuous!
Hmm….

Suppose we have $(3 \text{ nm})^3$ of sodium instead.

$$n_{3d} \approx 2.65 \times 10^{22} \text{ electrons per cm}^3$$

Use free electron mass, $m = 9.1 \times 10^{-31} \text{ kg}$

$$E_F = 5.2 \times 10^{-19} \text{ J} = 3.2 \text{ eV} = 38000 \text{ K} (!)$$

$$\nu_{3d}(E_F) = 7.6 \times 10^{16} \text{ J m}^{-3} \text{ m}^{-1}$$

$\Delta$ = single-particle level spacing

$$= 1/ (\nu_{3d}(E_F) \text{ V})$$

$$= 4.9 \times 10^{-22} \text{ J} = 35 \text{ K} = 3 \text{ meV}$$

This is much more significant, and is actually measurable!

At low temperatures, the individual electronic levels in a metal grain can dominate many properties, something that simply never happens at macroscopic sizes.

How does one actually “see” electronic levels like this?

First, come up with a clever way of making a single metal grain.

Then find a way to do spectroscopy on that grain.

**Coulomb Blockade:**

Electrons can tunnel on and off the nanoparticle from one side of the membrane to the other.

By varying voltages, one can see enhanced conductance corresponding to discrete levels in the grain.
Peaks in conductance for each single-particle state.

Can convert $V$ to energy, find level spacing of 0.5 meV, consistent with particle size.

The metal grain case is 3d, in the sense that the grain is still larger in all 3 directions than the Fermi wavelength of the electrons.

What about a 1d case?

30 nm segment of metallic carbon nanotube.

Not bad to treat as 1d “box”.

Measurement done by scanning tunneling microscope (STM).

Tunneling current proportional to local density of states. Presence of an available molecular orbital = higher conductance.

Venema et al., Science 283, 52 (1999).
Tunneling into nanotube segment shows 1d-levels, too!

Each bump = electron density from molec. orbital.

Each bump = 1 C atom.

Venema et al., Science 283, 52 (1999).

Similar effect in 1d chain of metal atoms!

Again, using STM to map single-particle state wavefunctions.


The story so far:

Starting model for electrons: *Ideal Fermi Gas*

- no interactions
- many-particle states built from single-particle states
- finite size of box quantizes allowed values of momenta, energy.
- level spacings set by size and dimensionality of system.
- \( k \)-space is useful way to count single-particle states
- At \( T=0 \), all states below \( E_F \) are full.

Next time:

- Dimensionality - when is it legit to say system is 1d?
- Finite temperatures, distribution functions
- Better boundary conditions
- Rates of processes
Starting model for electrons: *Ideal Fermi Gas*
- no interactions
- many-particle states built from single-particle states
- finite size of box quantizes allowed values of momenta, energy.
- level spacings set by size and dimensionality of system.
- $k$-space is useful way to count single-particle states
- At $T=0$, all states below $E_F$ are full.

On the plate today:
- Dimensionality?
- Better boundary conditions
- Finite temperatures, distribution functions
- Rates of processes

### Dimensional crossovers

Suppose we have some confining potential with $x$-length $L$, $y$-width $w$, and $z$-depth $d$, such that $L \gg w \gg d$.

Start adding electrons at $T = 0$. When should we treat this system as 1d, 2d, or 3d, in terms of electronic properties?

$$E(n_x, n_y, n_z) = \frac{\hbar^2}{2m} \left( \frac{\pi}{L} n_x \right)^2 + \left( \frac{\pi}{w} n_y \right)^2 + \left( \frac{\pi}{d} n_z \right)^2$$

First single-particle states filled are those at lowest energies; these states are those with the lowest allowed $y$ and $z$ momenta: $n_x = n_y = 1$. *System acts 1d* in this limit.
First single-particle states filled are those at lowest energies; these states are those with the lowest allowed $y$ and $z$ momenta: $n_y = n_z = 1$. *System acts 1d* in this limit.

Cross-over to 2d type behavior

Once $E_F$ is high enough, the next “$y$-subband” begins to be occupied. That is, once

$$E(n_x, n_y, n_z) = \frac{\hbar^2}{2m} \left( \left( \frac{\pi}{L} n_x \right)^2 + \left( \frac{\pi}{w} n_y \right)^2 + \left( \frac{\pi}{d} n_z \right)^2 \right)$$

it becomes energetically favorable for the next electron to occupy the $n_y = 2$ state with $n_x = 1$.

As higher $n_x$ and $n_y$ states are occupied, the system acts roughly like a 2d system of area $L \times w$.

Once the Fermi level is sufficiently high that the $z$ subbands can play a role, the system is effectively 3d.
A wavelength picture

Put another way:

• When the Fermi wavelength is small compared to one transverse size of the system, the system acts like a 1d conductor.

• When the Fermi wavelength is small compared to two transverse sizes of the system, the system acts like the 2d case.

• When the Fermi wavelength is small compared to all three transverse sizes of the system, the system acts like the 3d case.

Three points:

• This holds true in real systems, though the overall confining potential also matters.

• This is dimensionality w.r.t. “quantum confinement”. There are other colloquial uses of “dimensionality”!

• A small molecule or dot with only a few nodes in the wavefunction along any direction = “0d”.

Example: quantum wire

Image from IBM-Zurich
Example: quantum point contact

Image from Ferry group, ASU

Example: quantum point contact

Image from Ferry group, ASU
Example: quantum point contact

For fixed $\lambda_p$, varying the width of the QPC controls how many transverse modes can fit through. i.e. How does width compare with $\lambda_p$?

This has been theory - can this actually be seen rather than inferred from calculations or conductance measurements?

Example: quantum point contact

Scan-probe microscopy of structure - look at conductance modulation as function of position; tells you where the electron density is.

Westervelt group at Harvard
Boundary conditions

We’ve been assuming square-well potentials with infinitely high walls.

- Is this physically reasonable?
- Is it necessary? If not, why not?

We know these boundary conditions aren’t right because it is possible to eject electrons from atoms and molecules (ionization) and from larger solids (photoemission, Auger, etc.).

An alternative for bulk solids

In physicist lingo, translational invariance implies momentum conservation.

Boundary conditions break that global translational invariance, though locally inside the potential well it still must hold.

The result: only certain discrete values of momentum (≈ $k$) are allowed.

Many of the properties of the ideal Fermi gas continue to work well even if the boundary conditions are changed.

More useful alternative: **Periodic boundary conditions**

In 1d, \[ \psi(x + L) = \psi(x), \]
\[ \psi'(x + L) = \psi'(x) \]
What’s the difference? Use 2d case as example.

<table>
<thead>
<tr>
<th></th>
<th>Hard-wall</th>
<th>Periodic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowed values of $k$</td>
<td>$k_x = \frac{n\pi}{L}, k_y = \frac{n\pi}{L}$, $n = 1,2,3…$</td>
<td>$k_x = \pm \frac{n\cdot2\pi}{L}, k_y = \pm \frac{n\cdot2\pi}{L}$, $n = 1,2,3…$</td>
</tr>
<tr>
<td>Energy</td>
<td>$E = \frac{\hbar^2(k_x^2 + k_y^2)}{2m}$</td>
<td>$E = \frac{\hbar^2(k_x^2 + k_y^2)}{2m}$</td>
</tr>
<tr>
<td>Fermi surface</td>
<td><img src="image1.png" alt="Fermi surface" /></td>
<td><img src="image2.png" alt="Fermi surface" /></td>
</tr>
</tbody>
</table>

Now have traveling waves instead of standing waves.

Can really think of $k$ as a wavevector, as in the free particle case.

Amazingly, all our expressions for densities of states from the hard-boundary case continue to hold, exactly.

$$V_{1d}(E_F) = \frac{1}{\pi} \left( \frac{2m}{\hbar^2} \right)^{1/2} \frac{1}{E_F^{1/2}}$$

$$V_{2d}(E_F) = \frac{m}{\pi\hbar^2}$$

$$V_{3d}(E_F) = \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} E_F^{1/2}$$
Finite temperatures

What are the excitations of the ideal Fermi gas?

Ground state \((T=0)\)

Particles and Holes

\[ k_x \quad k_y \]

\[ E \quad N(E) \]

\[ k_x \quad k_y \]

\[ E \quad N(E) \]

Distribution functions for finite \(T\)

- Particle-hole excitations are created as \(T\) increases.
- Typically, particles are promoted from within \(k_B T\) of \(E_F\) to an otherwise unoccupied single particle state with \(E > E_F\).
- Particles aren’t promoted from deep within the Fermi sea because there aren’t any available unoccupied states accessible with a typical thermal excitation energy.

Probability of finding a single-particle state of particular spin with energy \(E\) occupied is given by the \textit{Fermi-Dirac distribution}:

\[
f(E, \mu, T) = \frac{1}{e^{\left(\frac{E-\mu}{k_B T}\right)} + 1}
\]

Here \(\mu\) is the \textbf{chemical potential}. 
Fermi distribution

- As $T$ approaches zero, the Fermi-Dirac distribution approaches a step function.
- A Fermi gas well-described by a FD distribution that's approximately step-like is termed degenerate.
- At higher temperatures, the FD distribution smears out and looks more like the classical Maxwell-Boltzmann case.

![Fermi distribution graph](image)

Chemical potential refresher 1

In canonical ensemble, $T$ is defined by looking at $\delta S$ when energy is exchanged with a reservoir:

$$
\delta S_{\text{tot}} = \left( \frac{\partial S}{\partial E} \right)_{\text{system}} \delta E_{\text{sys}} + \left( \frac{\partial S}{\partial E} \right)_{\text{res}} \delta E_{\text{res}}
$$

$$
\left( \frac{\partial S}{\partial E} \right)_{\text{sys}} = \left( \frac{\partial S}{\partial E} \right)_{\text{res}}
$$

$$
\left( \frac{\partial S}{\partial N} \right) = \left( \frac{1}{k_B T} \right)
$$

In grand canonical ensemble, $\mu$ is defined by a similar analysis allowing energy + particle exchange. The result:

$$
\left( \frac{\partial S}{\partial N} \right) = \left( \frac{\mu}{T} \right)
$$
Chemical potential refresher 2

- Diffusive equilibrium at fixed $T$ requires equal chemical potentials.
- Equal $\mu, T$ implies no net flow of particles.
- Particles flow from higher chemical potential regions to lower chemical potential regions.
- “External” potentials (e.g. applied voltage differences) add onto the chemical potential.
- One may interpret the chemical potential as the (average) energy required to add one more particle to a system.
- In a Fermi-Dirac (FD) system, at $T = 0$ it is clear that $\mu = E_F$

Degeneracy - some numbers

- Compare $T$ with Fermi Temperature to answer whether to use FD or MB.

**3d copper at 300 K:** $n_{3d} \sim 8.47 \times 10^{22}$ electrons per cm$^3$

Use free electron mass, $m_0 = 9.1 \times 10^{-31}$ kg

$$E_F = 1.13 \times 10^{-18} \text{ J} = 7 \text{ eV} = 82000 \text{ K} (!)$$

Since 300 K $\ll$ 82000 K, copper is degenerate.

**2d electrons in a Si transistor at 300 K:**

$n_{2d} \sim 10^{13}$ electrons per cm$^2$

Use effective mass, $m_* = 0.33 m_0$

$$E_F = 1.16 \times 10^{-20} \text{ J} = 0.07 \text{ eV} = 840 \text{ K}$$

Electrons at RT not very degenerate in transistors!

We’ll see later that devices at lower $T$ may become important….
Computing averages of various quantities

Often we’re interested in integrals of the form:

$$\int_0^\infty dE \left( \nu_d(E) f(E, T, \mu) \right) g(E)$$

where \(g(E)\) is some function of energy. This adds up the contributions \(g\) from each electron.

Where \(\nu_d(E)\) is the number of single-particle states at energy \(E\) and \(f(E, T, \mu)\) is the occupancy of such a state.

Only electrons affected by \(T\) are within ~ \(k_B T\) of the Fermi surface.

Difference of two FD functions

Also often care about difference between two FD functions, (e.g. one with energy \(E\) and one with \(E + eV\)) closely related to the derivative of a FD function. Why?

$$\int_0^\infty H(E) f(E, T, \mu) dE \approx$$

$$\approx H(E) f(E, T, \mu) \bigg|_{E=\infty}^{E=0} - \int_0^\infty dE \left[ \int_0^E H(E') dE' \right] \frac{\partial f}{\partial E} dE$$

expand around \(E=\mu\):

$$\approx \int_0^\mu H(E) dE + \frac{1}{2} \frac{d}{d\mu} H(\mu) \int_0^\infty dE (E - \mu)^2 \left[ \frac{\partial f}{\partial \mu} \right] + h.o.t.$$
“Thermal broadening” function

The function

\[
\frac{\partial f}{\partial \mu} = -\frac{\partial f}{\partial E} = \frac{1}{4k_BT} \cosh^{-2}\left(\frac{E - \mu}{2k_BT}\right)
\]

is often called the thermal broadening function.

\[
\lim_{\mu \to 0} \left\{ \frac{1}{4k_BT} \cosh^{-2}\left(\frac{E - \mu}{2k_BT}\right) \right\} = \delta(E - \mu)
\]

Time-dependent perturbation theory

- See handout for review.
- Used when calculating rates of processes - consider transitions from one single-particle state to another.
- Assume system starts in eigenstate of unperturbed potential.
- Turn on perturbing potential, \( V(t) \). Expand perturbed state in unperturbed time-varying states:
  \[
  |\Psi(t)\rangle = \sum_j c_j(t) e^{-i\hat{E}_j t/\hbar} |\psi_j^0\rangle
  \]
- \( |c_j(t)|^2 \) = instantaneous probability of finding system in state \( j \).
- Solve time-varying problem to find \( P_{j,a}(t) \), probability of being in state \( j \) at time \( t \) if initially in state \( a \) at \( t=0 \).
Main result: Fermi’s Golden Rule

\[ V_{0,ja} \equiv \langle \psi_j^0 | V | \psi_a^0 \rangle \]

\( V \) is coefficient of a harmonic perturbation

\[
\frac{1}{\tau_{ja}} \equiv \frac{d}{dt} P_{ja}(t) = \frac{2\pi}{\hbar} |V_{0,ja}|^2 V_d(E_j)L^d
\]

density of single-particle states at the final energy in a \( d \)-dimensional box of side \( L \).

Fermi’s Golden Rule caveats

\[
\frac{1}{\tau_{ja}} \equiv \frac{d}{dt} P_{ja}(t) = \frac{2\pi}{\hbar} |V_{0,ja}|^2 V_d(E_j)L^d
\]

• Initial and final states must be mixed by the perturbing potential! Origin of selection rules.

• If the number of available states at the final energy is zero, the rate of the process is zero!

• Energy is conserved.

• Derivation assumes initial state never depleted.
Why this matters

\[ \frac{1}{\tau_{ja}} = \frac{2\pi}{\hbar} |V_{0,ja}|^2 v_d(E_j)L^d \]

\[ \tau \sim \left( \frac{1}{v_d(E_j)L^d} \right) \text{single-particle level spacing} \]

- Some processes can slow down at the nanoscale because single-particle levels become far apart - restricted phase space for scattering.
- Example: apparent ballistic conduction in nanotubes at room temperature.

This time and next time

Today:
- Periodic boundary conditions also work for IFG.
- Sample size vs. wavelength \(\rightarrow\) dimensionality for DOS.
- Chemical potential and Fermi distribution at finite \(T\).
- DOS matters for rates of processes!

Next time:
- Not-so-ideal Fermi gas - electrons in a periodic potential