Lecture 4: Charging Effects
Ref. Chapter 1.4
We’ve been talking about current flow through a small device. For understanding current flow and conduction or any electronic property the main idea is to understand the allowed energy levels. Some of these levels are full and some are empty. At equilibrium, the empty and full levels are separated by electrochemical potential. This Fermi level has the property that all levels above it are empty and levels below it are full (at 0 K).

The contacts at the two sides of the channel are large objects. For reasons that will be understood later, large objects have a continuous set of allowed energy levels unlike the channel where the energy levels are discrete. What happens to these contacts when voltage is applied?

We usually take one of the contacts as a reference and investigate the effect of applied bias on the other one.

If the applied voltage is positive, then all the energy levels in that contact move down and if the applied voltage is negative, the levels move up. A common sense reason to see this is that a positive voltage makes it easier for the electrons to get in, hence the levels should be lowered.
• Current through one of the energy levels that is in between the contact electrochemical potentials can be written as:

\[ I = -\frac{q}{\hbar} \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2) \]

• While the average number of electrons in the channel is

\[ N = \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2} \]

• In the above equations, \( f_1 \) and \( f_2 \) are the Fermi functions of contacts 1 and 2. Gamma’s represent how easily an electron can get into the contacts. On other words they are measure of the strength of coupling of the channel to the contacts.

• \( N \) is the weighted average of \( f_1 \) and \( f_2 \). Contact 1 wants to see \( f_1 \) electrons in the channel and contact 2 wants to see \( f_2 \) electrons. It is a weighted average because the Gamma’s give the strength (weight) of the coupling of each contact to the channel.
What we just saw is fine so long as the coupling to the contacts is small. If the coupling is strong, we have to worry about another effect: broadening.

At this point the easiest way to rationalize this effect is to look at the wavefunction in the energy level. Ordinary we’d have:

$$\psi = e^{-iEt/\hbar}$$

which has a constant magnitude. Fourier transforming (FT) this to energy domain will give a single delta function.

When the channel is coupled to a contact, the electron doesn’t stay there forever but it gradually leaks out:

FT of above is a Lorentzian function in the energy domain:

Stronger coupling means more broadening and this limits the amount of current flow because as the level broadens some of it lies outside of the range between the Fermi levels of the contacts and doesn’t contribute to the current.
• We can incorporate broadening in the expressions for current and the average number of electrons.

\[ I = -\frac{q}{\hbar} \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2) \]

\[ N = \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2} \]

• This is what we had for one level. What happens now is that we have a density of states \( D(E) \) in the energy range \( dE \).

\[ dI = -\frac{q}{\hbar} D(E) dE \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2) \]

• To get the total current we need to sum over energy range of interest which means integrating over \( dE \).

\[ I = -\frac{q}{\hbar} \int D(E) dE \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2) \]

• Same story for \( N \):

\[ N = \int D(E) dE \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2} \]
Today we want to talk about another point which is needed to get the right expressions and complete the picture. That is the self consistent potential. To discuss this, consider the one level problem:

What does the current-voltage characteristic look like in a structure like this?

\[
I = \frac{-q}{\hbar} \int D(E) dE \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2)
\]

For large enough voltage \( f_1=1 \) and \( f_2=0 \) in the energy range of interest. We are left with:

\[
I = \frac{-q}{\hbar} \int D(E) dE \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2}
\]

If Gamma’s are constant, they can be pulled out of the integral. We have:

\[
I = \frac{-q}{\hbar} \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} \int D(E) dE
\]

The integral over \( D(E) \) is one because it is over a distribution sum of which amounts to one level because we only had one to start with.
Finally what we are left with is:

\[ I = \frac{-q}{\hbar} \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} \]

which is the maximum value of the current.

The energy range over which the change from minimum to maximum value occurs is about a few kT. In the event of strong coupling the energy level is broadened and the spread becomes wider by the amount of broadening.

What would the current be for a negative voltage? Based on the picture we have right now, there won’t be any current because the energy levels of the drain will be shifted…

…up and the energy level in the channel will not be in between the two Fermi levels. No energy levels => no conduction => no current. So we’d expect:
In the experiments, under certain conditions we can get what we get is not:

But looks like:

This could be expected in the sense that as long as the two contacts are the same, then there should be no reason for the current to flow under a positive voltage and not flow under negative bias.

So what is missing? What we’ve ignored is that we’ve assumed the channel energy levels don’t move as voltage is applied to a contact. A better way to think about this is that the cannal potential should be half way between 0 and 1 V which is 0.5V. So as the levels in the drain are lowered, the channel levels should also be lowered.
So if we take that into account, then we would expect a symmetric I-V characteristic. As the pictures illustrate, the channel energy levels move up and down with the drain levels (source is kept fixed as a reference). The sense of movement depends on the sign of the applied voltage.
So what we need is a way to write the potential $U$ in the channel. If we knew this how should $I$ and $N$ expressions be modified to incorporate the effect of channel’s potential?

$$I = \frac{-q}{\hbar} \int D(E) dE \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2)$$

$$N = \int D(E) dE \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2}$$

Thinking in a simple way, the effect of the potential $U$ is to move the entire density of states up or down (depending on the sign of $U$)

\[ \begin{array}{c}
E \\
U \\
0 \\
\end{array} \]  
\[ \begin{array}{c}
D(E-U) \\
D(E) \\
\end{array} \]

- Notice that the correct expression is $D(E-U)$ and not $D(E+U)$ because the peak of the density of states should appear when its argument is 0. Including the effect of potential, we’ll have:

$$I = \frac{-q}{\hbar} \int D(E-U) dE \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2)$$

$$N = \int D(E-U) dE \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2}$$

Next we need to figure out how to write $U$…
• We need to find an expression for U. In what follows we’ll make an approximation: we’ll assume all regions of the channel from left to right have the same potential. In other words, we attribute only 1 quantity to U for all of the channel at a given applied voltage. 
• If we’d want to find U in general, we’d have to solve Poisson’s equation:
\[
\frac{d^2 V}{dx^2} = -\frac{(-q)\Delta n}{\varepsilon}
\]
• Here we’ll go for 1 number for U and use a circuit model to find it.

• The amount of charge in the channel is:
\[
Q = -q\Delta n = C_s V + C_G (V - V_G) + C_D (V - V_D)
\]

• And the potential is (Vs is grounded):
\[
V = \frac{C_G V_G + C_D V_D}{C_S + C_G + C_D} + \frac{-q\Delta n}{C_S + C_G + C_D}
\]

• The first term shows the effect of the source, gate and drain voltages. Note that Vs doesn’t appear because it has been set to 0 as a reference. But then if the number of electrons in the channel change, that itself changes the channel’s potential. This is why there is the second term. Decreasing number of electrons in the channel makes it more positive and the levels in the channel go down.
• Notice that V and U are different quantities.
Potential Energy in the Channel

- V represents the electric potential in a certain region whereas U is the potential energy. Their relationship is \( U = -qV \).

So to get the potential energy U in the channel from the potential V, multiply everything by \(-q\):

\[
U = -qV = -q \left( C_G V_G + C_D V_D + \frac{-q \Delta n}{C_s + C_G + C_D} \right)
\]

For later use, we’ll right U as follows:

\[
U = U_L + \frac{q^2}{C_E} \Delta n, \quad C_E \equiv C_s + C_G + C_D
\]

- The quantity \( q^2/CE \) is called single electron charging energy. For small devices, CE is small and this quantity is big. In the even that its value is larger than a certain amount, it can affect the physics of the device.

If the potential in the channel is 1V, then the potential energy is -1eV. This again represent the fact that a positive voltage will bring down all the levels. A negative voltage will push all the levels up.
Let’s talk about Laplace potential.

Setting the source as reference, the potential in the channel depends on how far it is from the gate and how far it is from the drain. The closer it is to a contact, the larger will be the capacitance due the voltage applied to that contact.

This is very important in designing transistors. Consider a semiconductor as the channel material. Its density of states DOS look like:

- The on state means that a positive voltage has been applied and DOS has been lowered.
- How does the I-V curve look like?
This is exactly what we want from a good transistor from the circuit point of view. But in small transistors what people get is that current keeps increasing. The reason is that as the drain levels are lowered, the DOS in the channel also wants to slide down. This can be seen in the expression for UL:

\[ U_L = -q \left( \frac{C_G V_G + C_D V_D}{C_S + C_G + C_D} \right) \]

As DOS is lowered, there will be more available states; hence more current.
• To design a good transistor, one has to stop the density of states in the channel from sliding. The way to achieve this is making the first term of UL as big as possible (by increasing $C_g$) to make the effect of $V_d$ negligible.

$$U_L = -q \left( \frac{C_G V_G + C_D V_D}{C_S + C_G + C_D} \right)$$

• That is why designers like to have the gate as close as possible to the channel (to increase $C_g$). If $L=500$ Å, then gate should be as close as 20 Å to the channel. To make the length of the channel smaller, the gate should be even closer to the channel. At the same time, we don’t want any leakage current to pass through the oxide and this balance of things is hard to achieve.