Lecture 5: Charging Effects

Ref. Chapter 1.2 & 1.4
• This terminal is called gate and it is connected to a voltage source. Changing this voltage will change the resistance/conduction of the channel; allowing one to control the amount of current flow.  
• Through years, transistors have gotten smaller and as the dimension of the channel becomes comparable to a couple of 100 nm, concepts like resistance don’t hold any more …
• We need to understand new concepts to describe current flow and this is what this course is about.  
• So how do we actually describe these devices?...

• The main idea in this course is to understand current flow through very small devices (in nanoscale) like this generic transistor (up).  
• This device is composed of a channel to which the two terminals source and drain are connected. By applying a voltage difference between these two terminal current flows. There is a third terminal:
To understand the active region (channel) of the device, we need to understand its electron energy levels which are either filled or empty.

- At equilibrium and 0 K, the filled and empty levels are separated by a hypothetical line which is called the electrochemical potential and has the same value throughout the device.
- Above 0K, this line is not very sharp and spreads out.

(left) The occupancy of levels is described by the Fermi function: It is 0 for levels way above electrochemical potential and 1 for the levels way below electrochemical potential. The transition from 0 to 1 is not sharp (T>0K). And changes smoothly over a range of kT.
Applying a positive voltage to drain relative to source, lowers the energy levels and the Fermi function in the drain. When this happens, current will flow if there are any energy levels in the channel in the particular range... This range is the region between the electrochemical potentials of the two contacts. Current flows because source wants to keep the level filled while drain wants it empty.

Using these concepts, what equations did we derive?
The two equations in the box need to be solved self-consistently. Why?

\[ I = -\frac{q}{h} \int dED(E - U) \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2} (f_1 - f_2) \]

\[ N = \int dED(E - U) \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2} \]

\[ U = U_L + U_0 \Delta N \]

\[ U_0 = \frac{q^2}{C_E} \]

\[ U_L = \frac{C_G (-qV_G) + (-qV_D)C_D}{C_E} \]

\[ C_E = C_S + C_G + C_D \]
• If U was known, it could be substituted in the equation for N to find the number of electrons easily. However, U itself depends on the number of electrons N in the channel. So this equation needs to be solved self-consistently.

• We have to start with an initial guess for U; calculate N and from that calculated the value of U. Then we have to compare the calculate and the guessed values of U. If they are not close, then we have to add or subtracted some amount to our guess. Addition or subtraction come from the fact that the calculated value might be smaller or larger than initial guess.

\[
N = \int dED(E-U) \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2}
\]

\[U = U_L + U_0 \Delta N\]

• In practice, it is very hard to get exactly calculated and initial guess values for U. What we can do is to make them very close to each other. This is what we mean by convergence and we agree on convergence if the two values are with a fraction of kT of each other.

• Once U is known, current can easily be calculated using the value of U because everything else is known.
• Theoretically, the limits of integration for current run form negative infinity to positive infinity. But since we have the factor of \( f_1 - f_2 \) inside the integral we can neglect relatively large and relatively small energies for which \( f_1 - f_2 = 0 \).
• This is how we can understand current flow. There is a difference between the two electrochemical potentials which shows up as a difference in the Fermi functions which in turn give a non-zero \( f_1 - f_2 \) and this results in a current flow.
• What happens if we don’t apply a voltage but make one contact warmer than the other one?
Under these conditions the two electrochemical potentials are at the same energy. But the because \( f_2 \) changes smoothly and \( f_1 \) doesn’t, \( f_1-f_2 \) is not 0 for all energy ranges and this results in a current flow. This current actually can be used as a diagnostic tool for realizing if a device is n or p type.

What does \( f_1-f_2 \) look like now?

Supposing that the density of states is constant in the region of interest, then it could be pulled out of the integral which leaves integral \( f_1-f_2 \) and that is 0.

But what happen if the density of states is not a constant trough region of interest?
Let's say we have a density of states that looks like the following. (n type)

The figure tells us that when \( f_1 - f_2 \) is negative, \( D(E) \) has a constant value but when \( f_1 - f_2 \) is positive, \( D(E) \) is 0. This results in an electron current from drain to source.

The formula for the current is given by:

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

The Fermi distribution function is:

\[
f(E) = \frac{1}{(e^{(E - \mu)/k_B T} + 1)}
\]
Consider a situation where we have broadened level just above the Fermi energy. We want to see what happens as this level is pulled down in the energy scale. Specifically, we are after the relationship between the number of electrons $N$ in the channel and the gate voltage.

Above $\mu$, the level is empty. As we apply a positive voltage the levels lowers and starts being filled as it goes under $\mu$.

To make the problem conceptually easier, we’ll assume that the level is held fixed but $\mu$ moves. Notice that it is the relative movement that matters.

Now the problem how does the number of electrons change as $\mu$ moves in the energy scale.

$$U = U_L + U_0 \Delta N$$
If the self-consistent solution was not being considered, then we would have got something like the red curve:

But because of the term $U_0$ in the self-consistent solution, the level starts floating up as it gets filled with electrons. This makes filling up the level a slower process and we would expect something like the blue curve.

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

$$U = U_L + U_0 \Delta N$$