P-N junctions: Looking under the rug

We have derived the current-voltage relationship of ideal p-n junctions in the class and found it to be given by the famous diode equation

$$J = J_S[\exp(\frac{qV}{kT}) - 1],$$

where V is the applied voltage, and the saturation current is given by

$$J_S = q(n_{p0}\frac{D_n}{L_n} + p_{n0}\frac{D_p}{L_p})$$

for a junction with long (width >> diffusion lengths) emitter and base regions. For short base diodes, the base side minority diffusion length of J_S is replaced by the base width, W_B. To derive this relation, we made sweeping approximations, and swept many apparently contradictory facts under the rug. It is time to look under the rug.

The central two tricks in deriving the diode equation were the Shockley Boundary Conditions (SBC – now a phone company!), and the fact that the electron and hole currents (J_n, J_p) did not change across the depletion region. We return to these two approximations now.

Shockley Boundary Conditions

Shockley boundary conditions related the minority carrier density at the depletion edges to the applied voltage –

$$n(x_p) = n_{p0}\exp(\frac{qV}{kT})$$
$$p(x_n) = p_{n0}\exp(\frac{qV}{kT})$$

which required the assumption that $n(x_n) \approx n_{n0}$, $p(x_p) \approx p_{p0}$, i.e., the majority carrier densities at (and near) the depletion edges *do not change appreciably*. When can we make this assumption? To answer that, we note that the minority carrier density is typically a factor of 10^{10} or 10^{11} less than the majority carrier density. So, unless the exponential factor of SBC is as large, it is a safe approximation. At room temperature, the forward bias voltage required to achieve this is $V = (kT/q)\ln(10^{11}) \approx 0.7 Volts$, which is certainly not very high. In this lies the first approximation – the ideal diode equation is valid as long as we have LOW LEVEL INJECTION, i.e., when the minority carrier density injected from the emitter to the base is not as large as the majority carrier density in the base.

What happens at HIGH INJECTION conditions? If the minority carriers injected into the base (lightly doped side) of the diode becomes very large, then in order to maintain charge-neutrality, the electron and hole concentration will become equal, i.e.,

$$n \approx p \approx n_i \exp(\frac{qV}{2kT}),$$

and this effectively *relocates* the depletion edge of the p-n junction! This interesting high-injection effect will return to haunt us when we look at bipolar transistors operating at large currents.

Constant current components across depletion region

Finally, it is time to justify the shape of the quasi-Fermi levels E_{Fn}, E_{Fp} that we assumed to derive the ideal diode current. First, we note that the electron and hole current components are given by $J_n = n\mu_n dE_{Fn}(x)/dx$, and $J_p = n\mu_p dE_{Fp}(x)/dx$. Now assume that we are driving a current of density 10^3A/cm^2 through the diode in forward bias. Considering the neutral region of the n-side of the junction (far from depletion region), say the donor doping results in n=10^{17}/cm^3 electrons, with mobility 1000cm^2/Vs. Then, the gradient in the quasi-Fermi level in that region will be $dE_{Fn}(x)/dx = J/n\mu_n \approx 63\text{eV/cm}$. On the other hand, we know that the typical built-in voltage in a p-n junction is of the order of 1Volt, and it drops over a typical depletion width of 100nm, leading to a gradient of the bands of 10^5eV/cm. We can see right away why we assume that the quasi-Fermi levels are essentially flat in the neutral regions, even though a current is flowing. In reality, they have a very small slope, but can be neglected when compared to other slopes in the band diagram.

In the depletion region, we assumed the quasi-Fermi level E_{Fn} to continue to be flat till the p-side depletion edge is reached. We know that J_n and J_p are constant across the depletion region as long as there are no R-G processes occurring there. Thus, we must have

$$J_n(x_n) = J_n(x_p) \Rightarrow n(x_n)dE_{Fn}(x_n)/dx = n(x_p)dE_{Fn}(x_p)/dx.$$

Since $n(x_n) \gg n(x_p)$, the electron quasi-Fermi level MUST have a VERY HIGH slope in the p-side near the depletion edge, and a VERY SMALL slope in the n-side. This implies that the quasi-Fermi level is ALMOST flat in through the depletion region. Note that the quasi-Fermi level has a non-zero slope in the depletion region, otherwise there would not be a current, similar to the neutral region. However, the boundary conditions tell us that the slope in the depletion region is negligible. Similar argument holds for the hole quasi-Fermi level. Finally, we note that whenever we have R-G processes in the depletion region, the two current components J_n and J_p are no more constant over the depletion region.