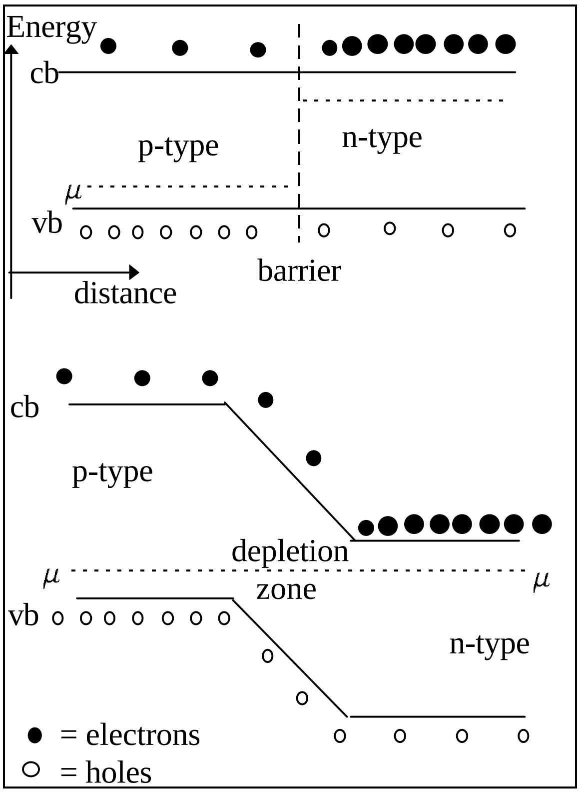
Semiconductor Devices

It is clear that the conductivity of semiconductors can be precisely controlled. This is the basis of a whole range of devices.

The pn junction



There are a number of ways of producing a boundary between p-type and n-type semiconductor. This boundary is important because of the behaviour of the electron energy levels near the junction. The relative positions are controlled by the fact that the chemical potential of the system must remain constant. If the n-type and p-type material are put together with a barrier between, the chemical potential of the p-type material is close to the top of the valence band (vb in the diagram) and the chemical potential of the n-type material is close to the bottom of the conduction band (cb in the diagram). When the barrier is removed, the levels must adjust themselves to ensure that the chemical potential is the same throughout the joined system (see lower half of the diagram). In a pn junction, this is done by a small transfer of electrons from the n-region to the p-region (where they annihilate some holes). This gives a *depletion layer*. The removal of electrons leaves the n-region positively charged and the removal of the holes leaves the p-region negatively charged. The total potential difference needed to obtain a uniform chemical potential can be deduced from the expressions for the chemical potential of electrons and holes. If we let the donor concentration on the n-side be N_D (i.e. $n_e = N_D$), then the chemical potential *relative to the valence band edge* on the n-side is (from equation (11) in the section on semiconductors, but subtracting E_V.

$$\mu_n = E_g - k_B T \ln\left[(1/4N_D)(2m_e k_B T/\pi\hbar^2)^{3/2}\right] \qquad (1a)$$

Similarly on the p side we obtain (for a concentration N_A of acceptors)

$$\mu_p = k_B T \ln\left[(1/4N_A)(2m_h k_B T/\pi\hbar^2)^{3/2}\right] \qquad (1b)$$

Thus the potential difference required is

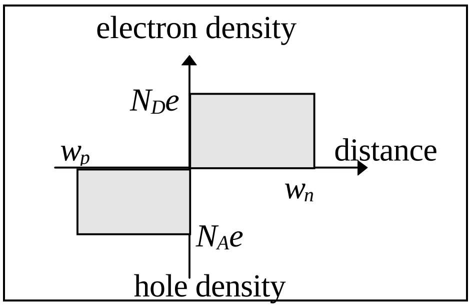
$$e\Delta\phi_0 = \mu_n - \mu_p = E_g + k_B T \ln\left[2N_D N_A(\hbar^2/2\pi k_B T)^3(m_e m_h)^{-3/2}\right]$$

which can be rearranged to give (using equation (10) from the semiconductor notes and recalling that for an intrinsic semiconductor $n_i^2 = n_e n_h$)

$$\Delta\phi_0 = \frac{k_B T}{e}\ln\frac{N_D N_A}{n_i^2} \qquad (2)$$

For Si, the value of this is about 0.7V for typical donor and acceptor concentrations. This is a junction (contact) potential; in a complete circuit other junctions will provide compensating potentials and a current will not flow.

We now consider the width of the depletion layer and the variation of the electrostatic potential $\phi(x)$. We assume that the boundary itself is abrupt and also that the edges of the depletion layer are sharp. In other words, we assume that there are no majority carriers within the depletion layer. and that the charge density near the junction can be approximated by



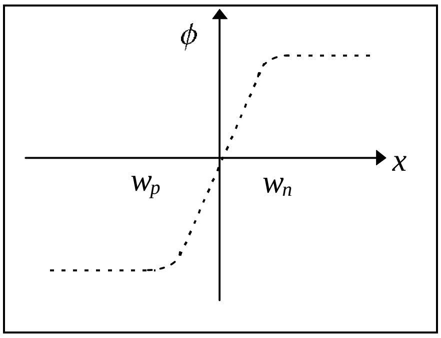
$$\rho(x) = -N_A e \quad \text{for } -w_p < x < 0 \text{ and}$$
$$\rho(x) = +N_D e \quad \text{for } 0 < x < w_n \text{ and } \rho(\mathrm{x}) = 0 \text{ otherwise} \tag{3}$$

where w_n, w_p are the widths of the depletion layer on the n-type and p-type sides respectively. We use Poisson's equation on the system. For the p-type region we have

$$\frac{d^2\phi}{dx^2} = \frac{N_A e}{\varepsilon\varepsilon_0}. \tag{4}$$

We choose the zero of potential so that the potential in the p-type region beyond the depletion layer is zero, $\phi(-w_p) = 0$; we also have $d\phi / dx|_{x=w_n} = 0$ (since the electric field must vanish at the boundary of the depletion layer) this can be solved to give

$$\phi = (N_A e / 2\varepsilon\varepsilon_0)(x + w_p)^2 \tag{5a}$$



in the p-type region. For the n-type region, we know that the total potential difference across the junction must be $\Delta\phi_0$; i.e. $\phi(w_n) = \Delta\phi_0$ and again, the electric field must vanish at the boundary. This gives

$$\phi = \Delta\phi_0 - (N_D e / 2\varepsilon\varepsilon_0)(x - w_n)^2 \tag{5b}$$

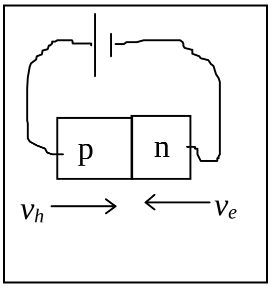
in the n-type region. We also know that the potential must be continuous across the junction. Thus if we set $x = 0$ for both these expressions we have derived we have $(N_A e / 2\varepsilon\varepsilon_0)w_p^2 = \Delta\phi_0 - (N_D e / 2\varepsilon\varepsilon_0)w_n^2$ which connects the widths of the two depleted regions and the potential drop. Since the system must be electrically neutral overall, we also know that $N_A e w_p = N_D e w_n$. These can be used as two simultaneous equations to determine the widths of the depletion layer, w_n, w_p. This gives

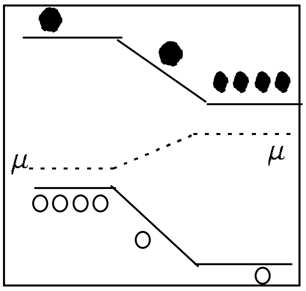
$$w_n = \left(\frac{2\varepsilon\varepsilon_0 N_A \Delta\phi_0}{eN_D(N_A + N_D)} \right)^{1/2} \quad \text{and} \quad w_p = \left(\frac{2\varepsilon\varepsilon_0 N_D \Delta\phi_0}{eN_A(N_A + N_D)} \right)^{1/2}$$

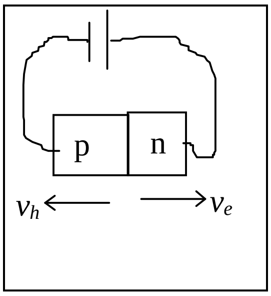
This gives a value of the width of the depletion zones of about a micron or so for typical doping levels. Note that the lighter the doping, the *wider* the depletion zone.

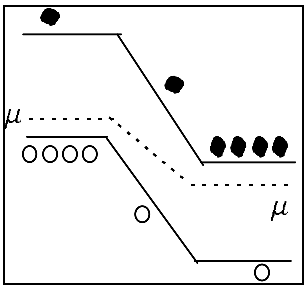
The pn junction with an applied voltage; Diodes

If we apply a voltage to a pn junction a current will flow. We can apply this voltage V in either of two ways. If we apply the positive terminal of a battery to the p side of the junction, V is taken to be positive and the junction is *forward biased*. The holes and electrons move towards one another and tend to annihilate.

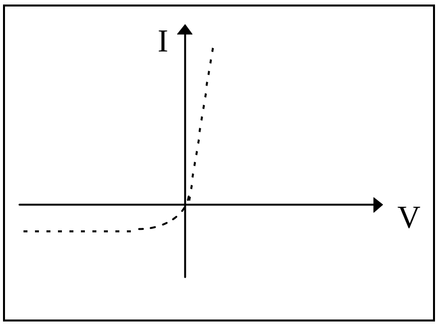








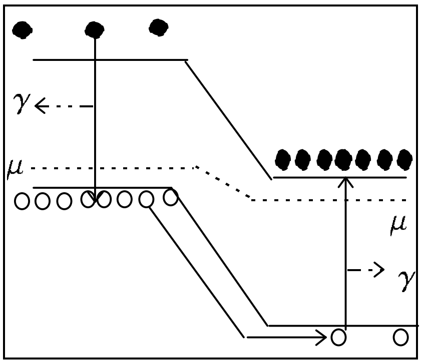
If the positive terminal is applied to the n side, then V is negative and the junction is *reversed biased.* The total potential difference across the depletion layer (given the conventions above) is $\phi = \Delta\phi_0 - V$ Forward bias reduces the potential difference whereas reverse bias increases it. Also, replacing $\Delta\phi_0$ by ϕ in the expressions for the depletion widths above, it is obvious that forward bias reduces the widths and reverse bias increases them. This obviously changes the charges on either side of the junction (since the depletion layers are where the ionised donors/acceptors are not compensated by free electrons). That is, the change in charge is related to a change in voltage - we have a variable capacitance.



Much more important is the effect on the current passed through the junction. In the absence of bias, the potential difference $\Delta\phi_0$ acts as a barrier to the flow of conduction electrons from n to p. A few electrons can jump the barrier giving an electron current I_e^0. If there is no bias, there is no current overall and so this must be balanced by an electron current from p to n. Although there are very few electrons on the p side, those that do drift into the depletion zone are immediately swept across to the n side by the potential difference. If the junction is forward biased, the barrier to the electron current from n to p falls. The current therefore rises exponentially, given by $I = I_e^0(\exp(eV / k_B T) - 1)$ but there is no corresponding change in the reverse current since there is already no barrier it. If, on the other hand, a reverse bias is applied, the barrier is made higher and there is an exponential fall in the current (V is negative). A similar argument applies for the hole current. This *rectifying behaviour* is of fundamental importance in devices.

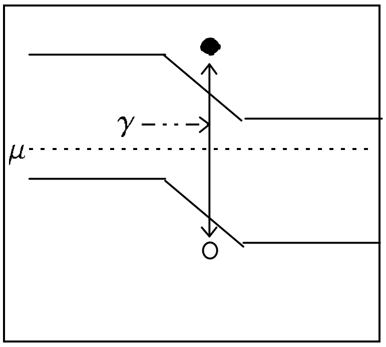
Light-emitting diodes and laser diodes

When electrons in a forward-biased diode flow into the p region, they combine with the holes. This recombination will release the energy of the band-gap E_g and, in certain cases, this is emitted as a photon. It is possible to design materials to produce light of different colours. If the forward bias is large enough, it is possible to drive enough electrons across the junction into states near the conduction band edge on the p side exceeds that in states at the top of the valence band. This is a *population inversion* and stimulated recombination in this case can produce lasing action.



Solar cells

Here, the photon is used to create an electron-hole pair in the depletion zone. After creation, the electrons move into the n-region and the holes into the p-region because of the potential difference $\Delta\phi_0$. This gives a current from n to p which is proportional to the flux of photons.



Field-effect transistor

In this device, the current in a thin strip of n-type material is controlled by an attached (reverse-biased) pn junction. If we change the bias voltage, we change the width of the depletion layer (see above). This in turn alters the width of the channel (*pinch-off*) along which the current can flow down the n-type strip and hence alters the current flow down the strip.