

FoP3B Part II Lecture 5: pn junction (part I)

The **pn junction** is the basic building block of semiconductor devices. It is used in solar cells, light emitting diodes as well as part of more complex device architectures, such as bipolar transistors. In this and the next lecture we will examine the structure of a pn junction and how this influences electric current transport.

A pn junction is formed by bringing together *n*- and *p*-type material to form an **interface** (Figure 1a)¹. The electron and hole concentration either side of the interface is different, which results in **majority carrier** electrons in the *n*-layer flowing into the *p*-layer, a process called **diffusion**. Similarly, majority carrier holes will diffuse from the *p*-layer into the *n*-layer. During this process ionised donor and **acceptor** atoms are left behind. For example, removal of the donor electron from the *n*-layer means that the layer acquires a net *positive* charge due to the Group V ionised donors. Similarly, removal of a hole from the *p*-layer, or equivalently acceptance of an electron from the *n*-layer, results in *negative* charge due to ionised Group III acceptors. The diffusion process takes place over a region close to the interface and is known as the **space charge region** or **depletion region** (Figure 1b).

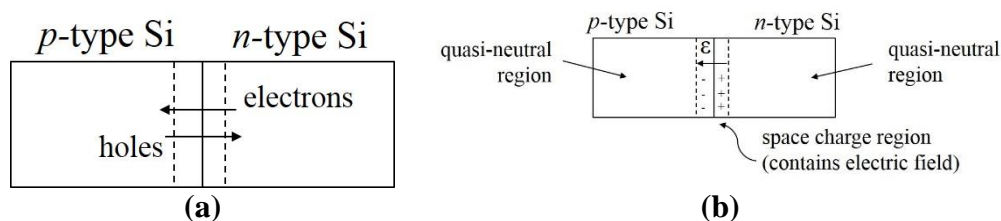


Fig 1: (a) and (b) is a schematic of a silicon pn junction.

*The presence of (unscreened) ionised donors and acceptors means that there is an internal electric field or equivalently potential difference within the space charge region. The electric field prevents further diffusion of electrons and holes. The region outside the space charge region is called the **quasi-neutral region**. The space charge region is an important feature of a pn junction. Its properties, i.e. electric field and potential, can be calculated using standard electrostatic methods provided we know the charge distribution. However, this is quite a complex problem, since the charge carrying particles consist of conduction band electrons, valence band holes as well as ionised donors and acceptors. The **depletion approximation** is therefore used to simplify the problem. This states that the space charge region is free of electrons and holes. The only charged particles are therefore ionised acceptors on the p-side and ionised donors on the n-side. Furthermore, we assume that all donors and acceptors in the space charge region are ionised. This is reasonable since at room temperature the material is in the **saturation regime**.*

Calculation of Electric Field

Let us now calculate the electric field distribution within the space charge region. This can be done using **Gauss' Law**:

¹ In reality pn junctions are formed by starting with one material, say p-type silicon, and doping the surface n-type by either ion implantation or diffusion of donor 'impurities'.

$$\vec{\nabla} \cdot \vec{\varepsilon} = \frac{d\varepsilon}{dx} = \frac{\rho(x)}{\epsilon_r \epsilon_0} \quad \dots (1)$$

where ε is the electric field, ϵ_r is the **dielectric constant** or relative permittivity, ϵ_0 is the permittivity of free space and ρ is the charge density. Due to the depletion approximation ρ takes the form:

$$\rho(x) = \begin{cases} -eN_A & -w_p < x < 0 \text{ (} p\text{-side)} \\ eN_D & 0 < x < w_n \text{ (} n\text{-side)} \end{cases} \quad \dots (2)$$

where $-w_p$ and w_n are the space charge boundaries for p - and n -sides (Figure 2a). N_A and N_D are acceptor and donor concentrations respectively (note that the ionised acceptors on the p -side are negatively charged and hence $\rho(x) = -eN_A$ in this region). The electric field in the p -side is determined by integrating Equation (1):

$$\varepsilon(x) = \int \frac{-eN_A}{\epsilon_r \epsilon_0} dx = \frac{-eN_A}{\epsilon_r \epsilon_0} x + A \quad \dots (3)$$

where A is the constant of integration. To determine its value we make use of the **boundary condition** that *the electric field at the space charge edge is zero*, i.e. $\varepsilon(-w_p) = 0$. Using this condition we find:

$$\varepsilon(x) = \frac{-eN_A}{\epsilon_r \epsilon_0} (x + w_p) \quad \dots (4)$$

Using the boundary condition $\varepsilon(w_n) = 0$, the electric field on the n -side is:

$$\varepsilon(x) = \frac{eN_D}{\epsilon_r \epsilon_0} (x - w_n) \quad \dots (5)$$

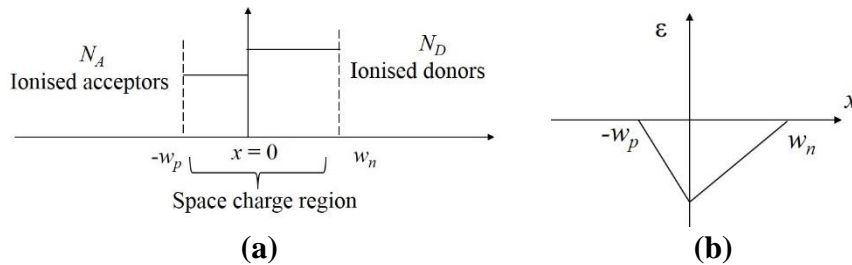


Fig 2: (a) Notation used for calculating space charge region properties and (b) the electric field distribution.

The electric field distribution is sketched in Figure 2b and shows a linear variation either side of the pn junction. It has a negative value since the electric field points in the $-x$ direction (Figure 1b). Note that the electric field must be continuous at the pn junction interface in order

to be consistent with **Maxwell's boundary conditions**² for the electric field. Equating (4) and (5) for $x = 0$ gives:

$$N_A w_p = N_D w_n \quad \dots (6)$$

This is the condition for **charge neutrality**, i.e. *it states that the negative charge due to ionised acceptors is equal to the positive charge due to ionised donors* (recall that we are using the depletion approximation, where it is assumed that there are no electrons or holes within the space charge region).

Calculation of Potential Field

The electric potential ϕ is easily calculated from the electric field using:

$$\vec{E} = -\vec{\nabla}\phi \quad \dots (7)$$

For example, the p -side potential is obtained by integrating the electric field in Equation (4):

$$\phi(x) = \frac{eN_A}{\epsilon_r \epsilon_0} \int (x + w_p) dx = \frac{eN_A}{2\epsilon_r \epsilon_0} (x^2 + 2w_p x) + B \quad \dots (8)$$

where B is a constant of integration. We can *arbitrarily* set the potential at $x = -w_p$ to zero and therefore:

$$\phi(x) = \frac{eN_A}{2\epsilon_r \epsilon_0} (x + w_p)^2 \quad \dots (9)$$

Similarly, the n -side potential can be shown to be:

$$\phi(x) = \phi_{bi} - \frac{eN_D}{2\epsilon_r \epsilon_0} (x - w_n)^2 \quad \dots (10)$$

where ϕ_{bi} is the **built-in potential** at the $x = w_n$ space charge edge on the n -side. Figure 3a is a schematic of the potential field. The potential energy E of an electron in a potential field ϕ is given by $E = -e\phi$, i.e. the energy profile is inverted w.r.t the potential (Figure 3b).

² Strictly speaking this is a special case of Maxwell's boundary condition and is only valid when the p - and n -sides are of the same type of material (e.g. silicon).

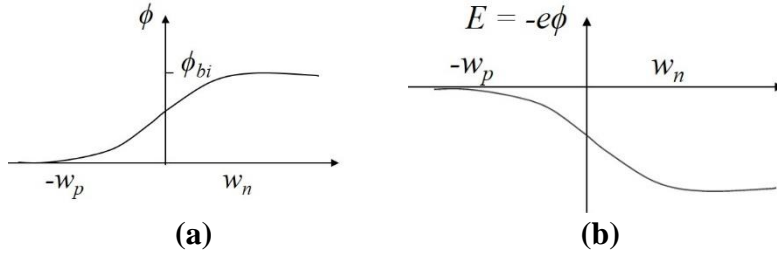


Fig 3: (a) Potential field ϕ and (b) energy E of an electron in the space charge region.

Using Figure 3b we can plot the shape of the conduction band minimum E_c and valence band maximum E_v before and after the p - and n -side make contact to form the pn junction (see Figure 4). *Before making contact the chemical potential μ of the p - and n -sides are different, but once the pn junction has formed the band bending due to the potential field creates a constant chemical potential throughout the system.* This can be understood by noting that the chemical potential is the free energy (G) change due to the addition of one mole of particles to the system, i.e. $\mu = (dG/dn)$. For the system to be in equilibrium the chemical potential must therefore be constant everywhere, otherwise electrons and holes will move from regions of high to low chemical potential in order to minimise their energy. In fact, the initial process of electron and hole diffusion to create the space charge region can also be thought of as being driven by the difference in chemical potential, forcing the system into a lower energy configuration.

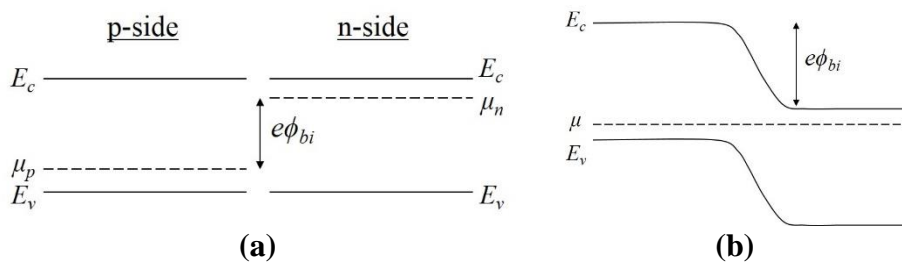


Fig 4: Chemical potential μ and band edges (a) before and (b) after the p - and n -sides come together to form a pn junction.