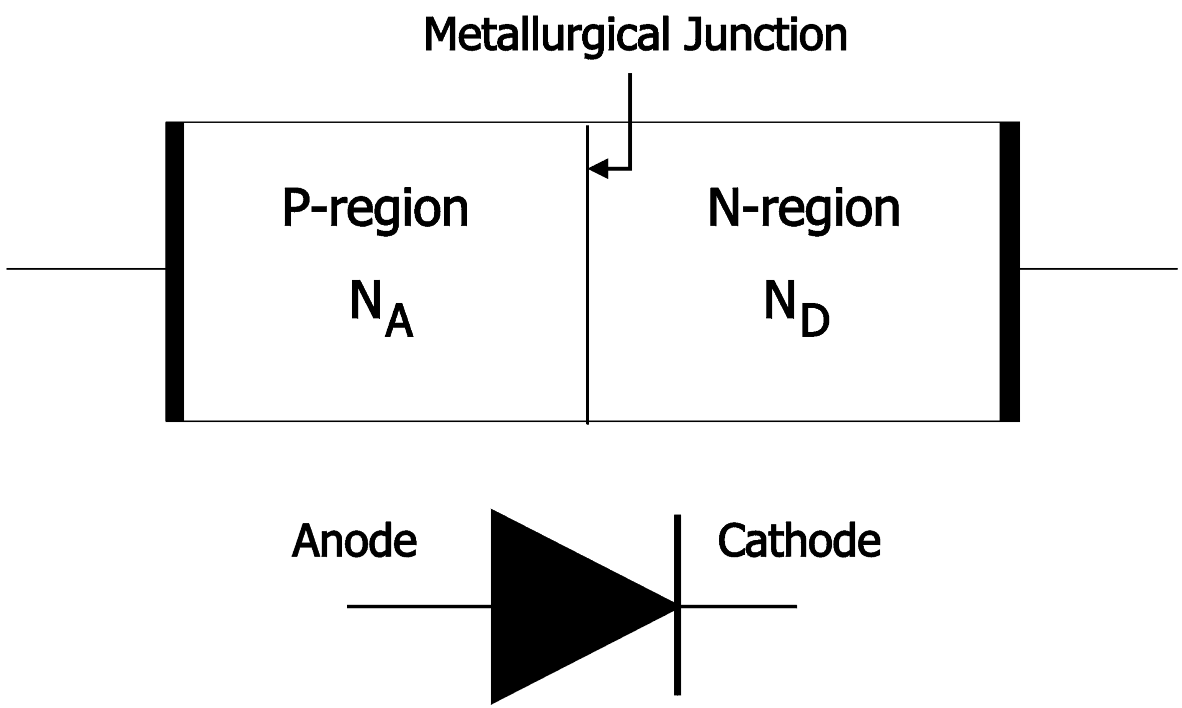
EE2003 Semiconductor Fundamentals

Electrostatics of the P-N Junction

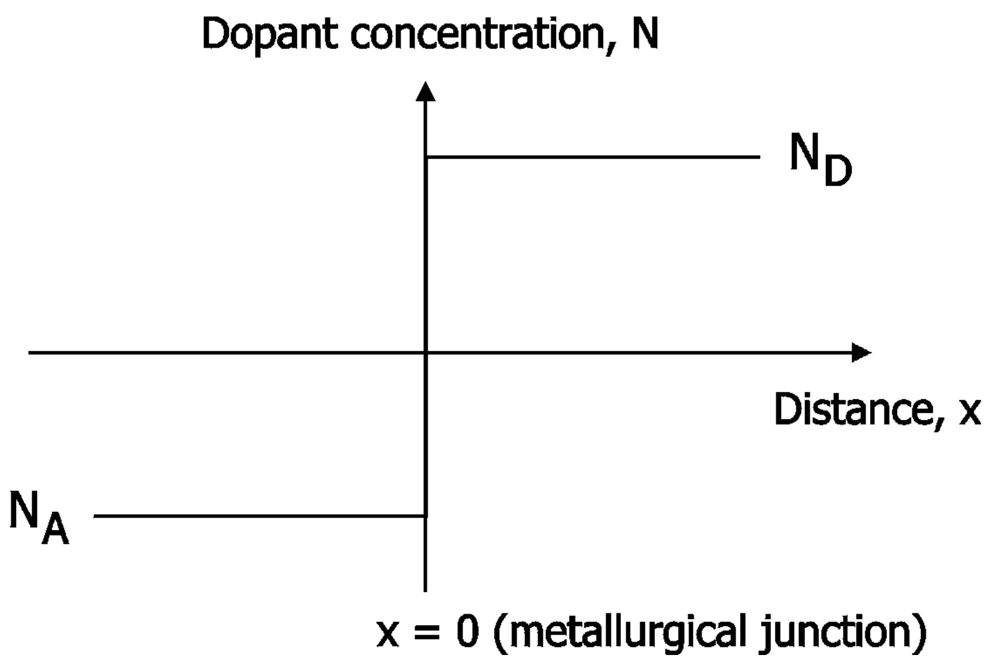
Basic Structure

- A P-N junction is formed when an **n-type** semiconductor region is brought into close contact with a **p-type** semiconductor region.



Basic Structure

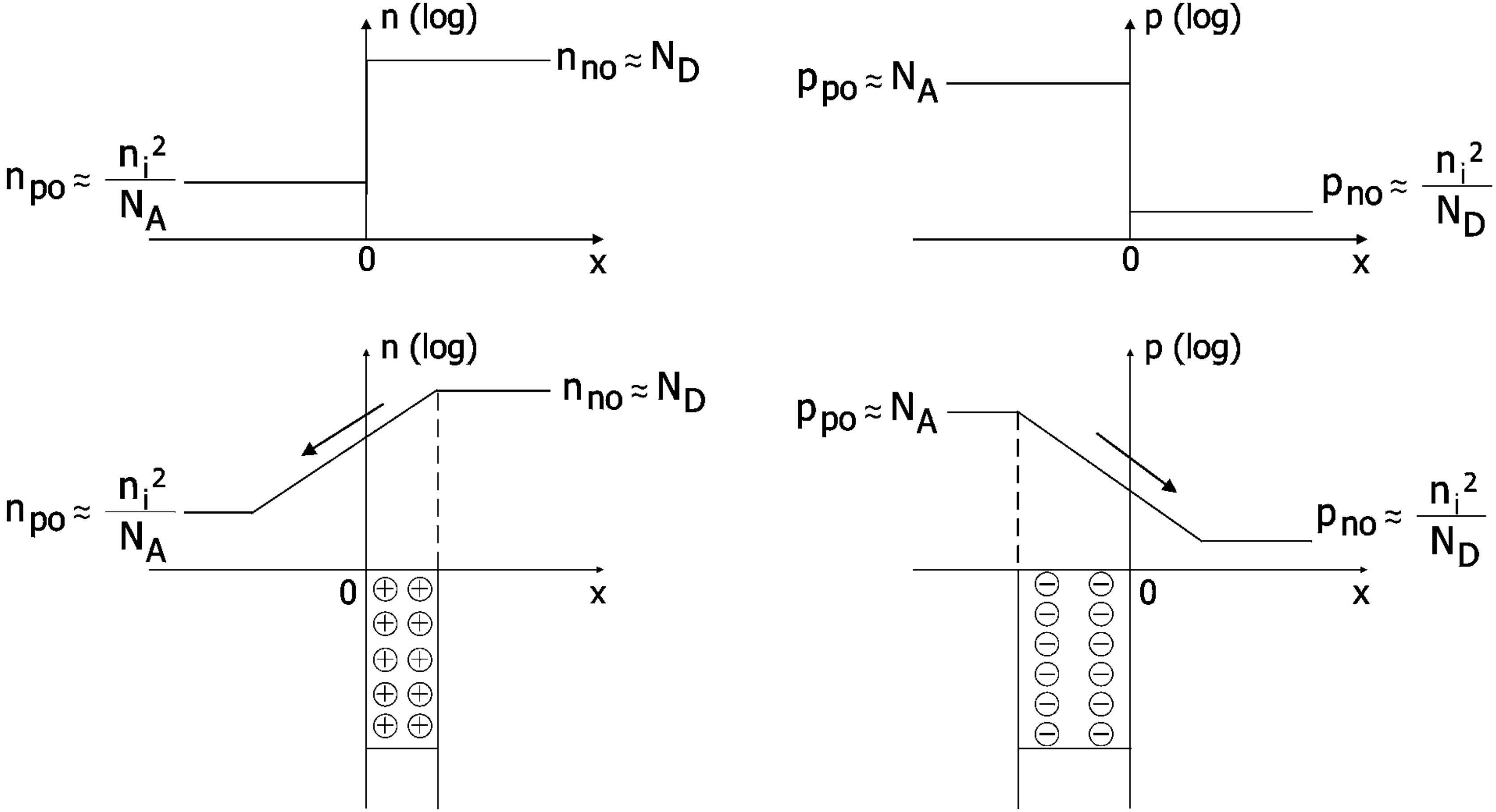
- **Assumptions:**
  - Uniformly doped n and p regions
  - Abrupt or step junction



Important Concepts

- Space charge (depletion) region
- Built-in electric field
- Built-in potential (barrier height)

Formation of the Space Charge Region



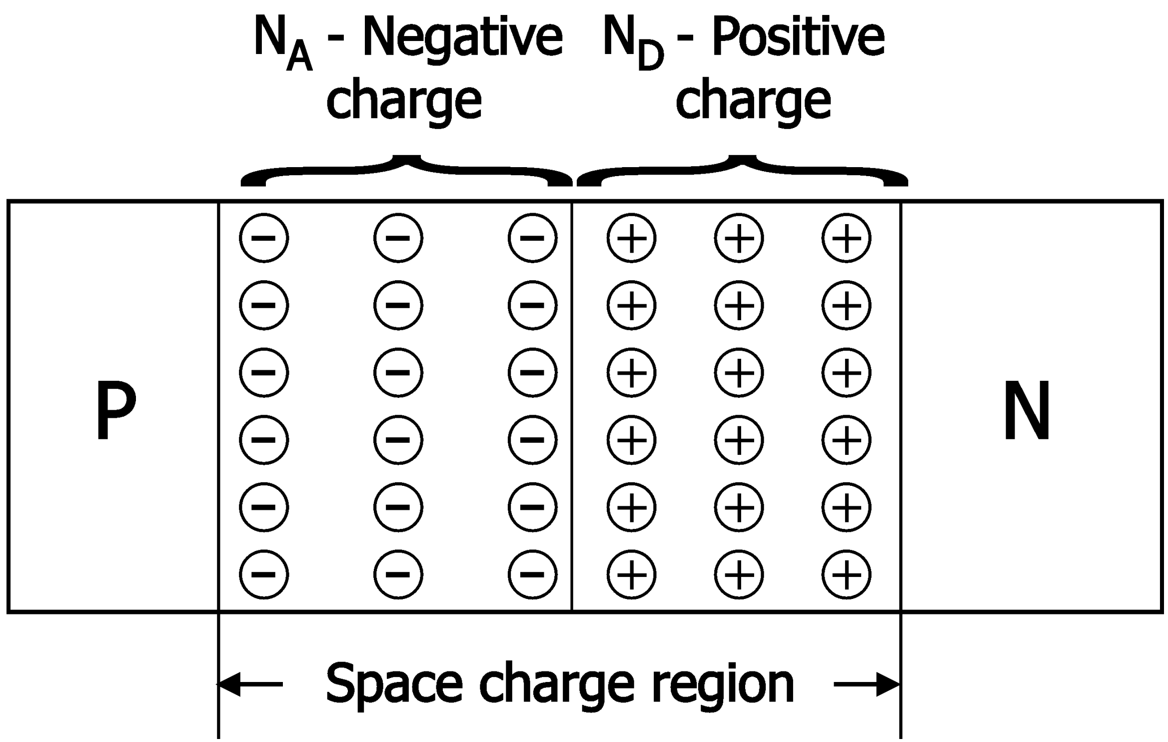
Formation of the Space Charge Region

- N region – High concentration of electrons (majority carriers)
- P region – Low concentration of electrons (minority carriers)
- An electron concentration gradient exists across a PN junction:
  - Electrons from the n region diffuse to the p region.
  - The departure of electrons leaves behind the immobile donor ions (positively charged) in the n region.

Formation of the Space Charge Region

- P region – High concentration of holes (majority carriers)
- N region – Low concentration of holes (minority carriers)
- A hole concentration gradient exists across a PN junction:
  - Holes from the p region diffuse to the n region.
  - The departure of holes leaves behind the immobile acceptor ions (negatively charged) in the p region.

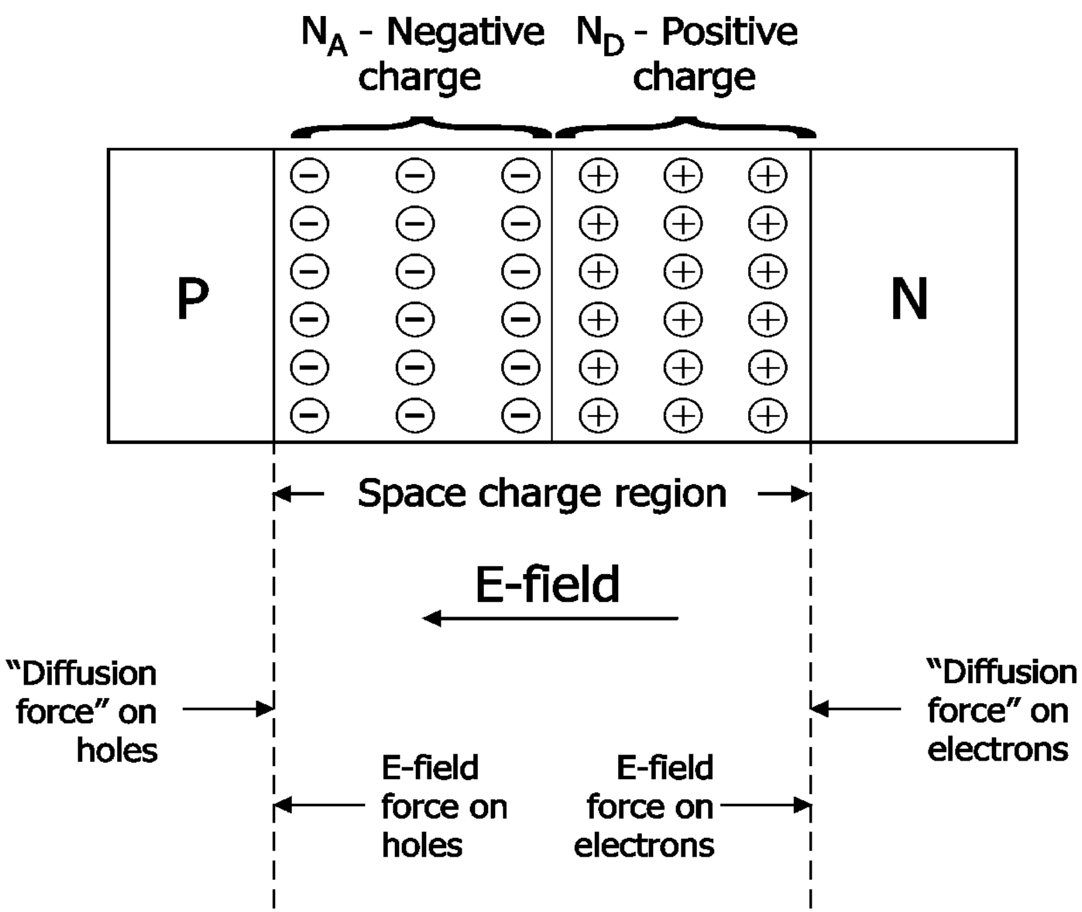
Formation of the Space Charge Region



Formation of the Space Charge Region

- The net positively and negatively charged regions are known as the **space charge regions**.
- Because the concentrations of electrons and holes in this region are lower than that in the respective neutral n and p regions, the space charge region is also known as the **depletion region**.

Built-In Electric Field



Built-In Electric Field

- The net positive and negative space charges in the n and p regions induce an electric field.
- Since this electric field is automatically created when a PN junction is formed, it is known as the **built-in electric field**.
- This electric field points from the positive to the negative charge, i.e. from the n to the p region.

Built-In Electric Field

- The built-in electric field counteracts the electron and hole diffusion processes.
- At **thermal equilibrium**, this counteracting force exactly balances the “diffusion force” exerted by the concentration gradient.
- There is **no net movement** of mobile charges across the PN junction under thermal equilibrium.

Built-In Electric Field

- Under thermal equilibrium,

$$\underbrace{qn\mu_n\xi}_{\text{Electron drift current}} + \underbrace{qD_n\frac{\partial n}{\partial x}}_{\text{Electron diffusion current}} = 0$$

$$\underbrace{qp\mu_p\xi}_{\text{Hole drift current}} - \underbrace{qD_p\frac{\partial p}{\partial x}}_{\text{Hole diffusion current}} = 0$$

Built-In Electric Field

- A mathematical expression for the built-in electric field $\xi(x)$ can be derived by solving the **Poisson's equation**.

$$\frac{\partial^2 V}{\partial x^2} = -\frac{\rho(x)}{\varepsilon_r \varepsilon_0} = -\frac{\partial \xi}{\partial x}$$

  - ρ is the volume charge density/concentration
  - ε_r is the relative permittivity of the semiconductor
  - ε_0 is the permittivity of free space

Solving the Poisson's Equation

- Volume charge density, ρ is given by the sum of **all** positive and negative charges in the space charge region.

$$\rho(x) = \underbrace{qN_D + qp(x)}_{\textbf{positive charges}} - \underbrace{qN_A - qn(x)}_{\textbf{negative charges}}$$

  - Since we have assumed uniform doping concentration, N_A and N_D are independent of distance. However, it is imporant to realise that N_A and N_D are in general functions of distance in practical diodes.
  - *p* and *n* are functions of distance, *x*. The Poisson's equation is thus difficult to solve since we do not exactly know $p(x)$ and $n(x)$.

Solving the Poisson's Equation

- **Assumptions:**
  - **Depletion approximation** – The concentrations of mobile charges in the space charge region are negligible compared to the concentration of the immobile ionic (space) charge.
  - The space charge region ends abruptly at $x = +x_{no}$ and $x = -x_{po}$.

Solving the Poisson's Equation

- Applying the depletion approximation, we arrive at a simplified expression for the volume charge density.

$$\rho \approx qN_D - qN_A$$

- Considering the p-region,

$$\rho \approx -qN_A$$

- Simplified Poisson's equation:

$$\frac{\partial \xi}{\partial x} = -\frac{qN_A}{\varepsilon_r \varepsilon_o}$$

Solving the Poisson's Equation

- Integrating w.r.t. x,

$$\xi(x) = -\frac{qN_A}{\varepsilon_r \varepsilon_0} x + C$$

- C is an integration constant that can be determined by applying an appropriate boundary condition.

- There are two possible boundary conditions: $\xi\,(x = 0)$ and $\xi\,(x = -x_{p0})$.

- Which one should we use?

Solving the Poisson's Equation

- Since we do not yet know what ξ $(x = 0)$ is, the relevant boundary condition to use is:

$$\xi\left(x = -x_{p0}\right) = 0$$

- Integration constant:

$$\xi\left(-x_{p0}\right) = -\frac{qN_A}{\varepsilon_r \varepsilon_0}\left(-x_{p0}\right) + C = 0$$

$$\therefore \quad C = -\frac{qN_A x_{p0}}{\varepsilon_r \varepsilon_0}$$

Solving the Poisson's Equation

- Built-in electric field (p region):

$$\xi(x) = -\frac{qN_A}{\varepsilon_r \varepsilon_0}\left(x + x_{p0}\right), \quad -x_{p0} \le x \le 0$$

  - For an abrupt or step junction, the electric field is a linear function of distance.
  - Note that x is negative. But since $x \ge -x_{p0}$, the electric field ξ is negative.
  - The negative sign denotes the direction of the electric field, from **the n to the p region** or in the negative x-axis direction.

Exercise

- Show that the built-in electric field in the n region can be expressed as:

$$\xi(x) = \frac{qN_D}{\varepsilon_r \varepsilon_0}(x - x_{n0}), \quad 0 \le x \le x_{n0}$$

- Note that since $x > 0$ and $x \le x_{n0}$, ξ is also negative in this case.

Built-In Electric Field

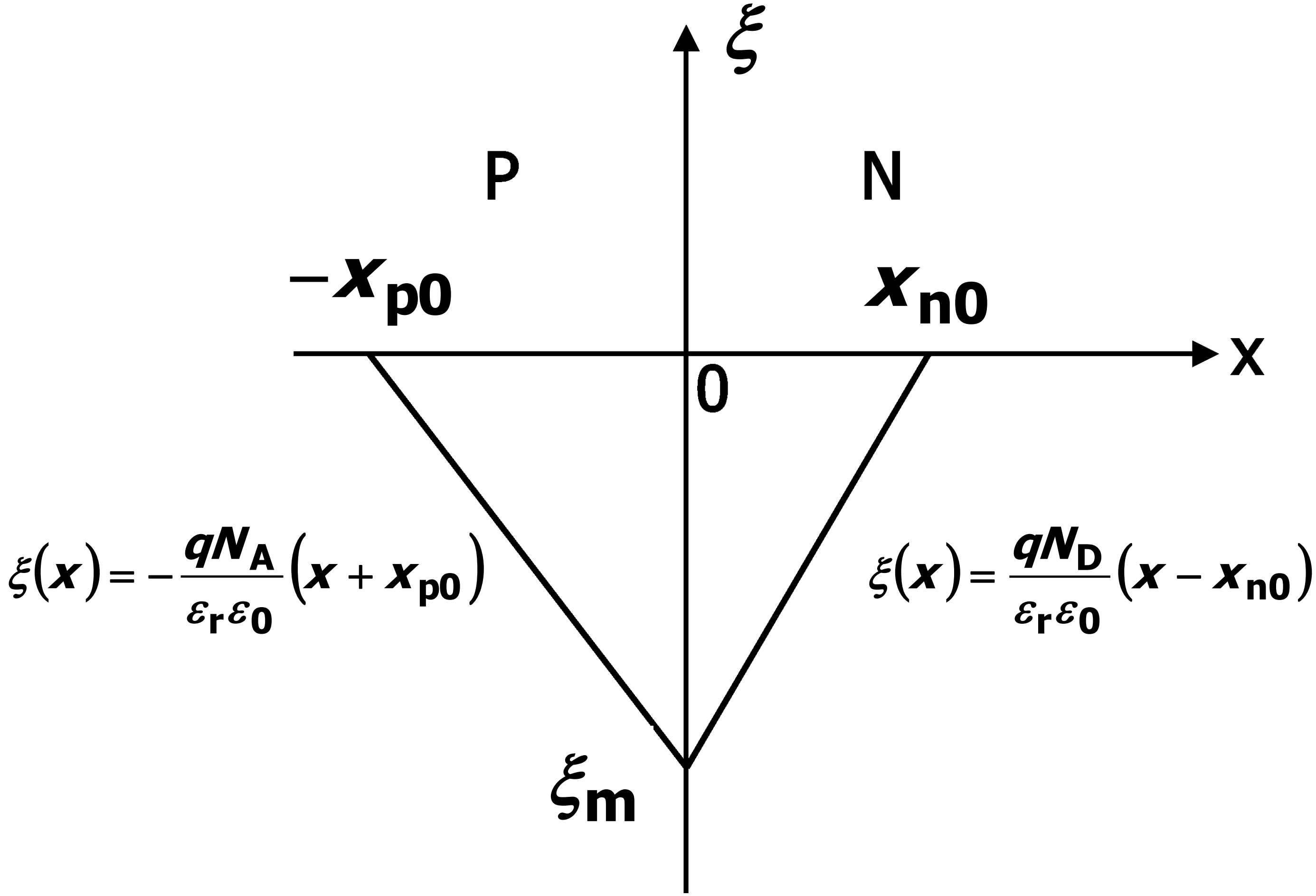
- **Charge neutrality**

$$\underbrace{qN_A x_{p0}}_{\substack{\text{net negative charge}\\ \text{on p-type side}}} = \underbrace{qN_D x_{n0}}_{\substack{\text{net positive charge}\\ \text{on n-type side}}}$$

$$\underbrace{\frac{qN_A x_{p0}}{\varepsilon_r \varepsilon_0}}_{\substack{|\xi(x=0)|\text{ using electric field}\\ \text{expression on p-type side}}} = \underbrace{\frac{qN_D x_{n0}}{\varepsilon_r \varepsilon_0}}_{\substack{|\xi(x=0)|\text{ using electric field}\\ \text{expression on n-type side}}}$$

- $\therefore$ the electric field is **continuous** at $x = 0$ (metallurgical junction).

Built-In Electric Field



Maximum Built-In Field

- The **maximum** built-in field occurs at the **metallurgical junction** ($x = 0$).
- Remarks:
  - Although we arrive at this conclusion by assuming a step junction, the same conclusion applies to PN junction with arbitrary junction doping profile.
  - The maximum built-in field is important as it determines the breakdown voltage of the PN junction.

Built-In Voltage/Potential

- The separation of positive and negative charges in the space charge region of a PN junction induces a built-in electric field.
- The built-in electric field in turn causes a potential difference between the n and p regions. This potential difference is called the built-in voltage of a PN junction.
- **Question**:
  - Which side (n or p) is at a higher potential?

Built-In Voltage/Potential

- The built-in voltage can be evaluated using the following fundamental relationship:

$$V_{bi} = -\int \xi \;\; \partial x$$

- Geometrically, the above integral gives the **area** under the electric field versus distance plot.
- In the case of the step junction,

$$V_{bi} = \frac{1}{2} \cdot |\xi_m| \cdot \left(x_{n0} + x_{p0}\right) = \frac{|\xi_m| \cdot W_0}{2}$$

Built-In Voltage/Potential

- Recall that the maximum electric field,

$$\left|\xi_{\mathbf{m}}\right| = \left|\xi(x = 0)\right| = \frac{qN_{\mathrm{A}}x_{\mathrm{p0}}}{\varepsilon_{\mathrm{r}}\varepsilon_{0}} = \frac{qN_{\mathrm{D}}x_{\mathrm{n0}}}{\varepsilon_{\mathrm{r}}\varepsilon_{0}}$$

- Therefore, the built-in voltage,

$$V_{\mathrm{bi}} = \frac{qN_{\mathrm{A}}x_{\mathrm{p0}}}{2\varepsilon_{\mathrm{r}}\varepsilon_{0}}\left(x_{\mathrm{n0}} + x_{\mathrm{p0}}\right) = \frac{qN_{\mathrm{D}}x_{\mathrm{n0}}}{2\varepsilon_{\mathrm{r}}\varepsilon_{0}}\left(x_{\mathrm{n0}} + x_{\mathrm{p0}}\right)$$

Space Charge Width

- The distance that the space charge region extends into the p and n regions can be determined by solving the following two equations:

$$V_{bi} = \frac{qN_A x_{p0}}{2\varepsilon_r \varepsilon_0}\left(x_{n0} + x_{p0}\right) \qquad (1)$$

$$x_{n0} N_D = x_{p0} N_A \qquad (2)$$

Space Charge Width

- Width of the depletion region extending into the n region:

$$x_{n0} = \left\{ \frac{2\varepsilon_r \varepsilon_0 V_{bi}}{q} \left[\frac{N_A}{N_D \left(N_A + N_D \right)} \right] \right\}^{1/2}$$

- Width of the depletion region extending into the p region:

$$x_{p0} = \left\{ \frac{2\varepsilon_r \varepsilon_0 V_{bi}}{q} \left[\frac{N_D}{N_A \left(N_A + N_D \right)} \right] \right\}^{1/2}$$

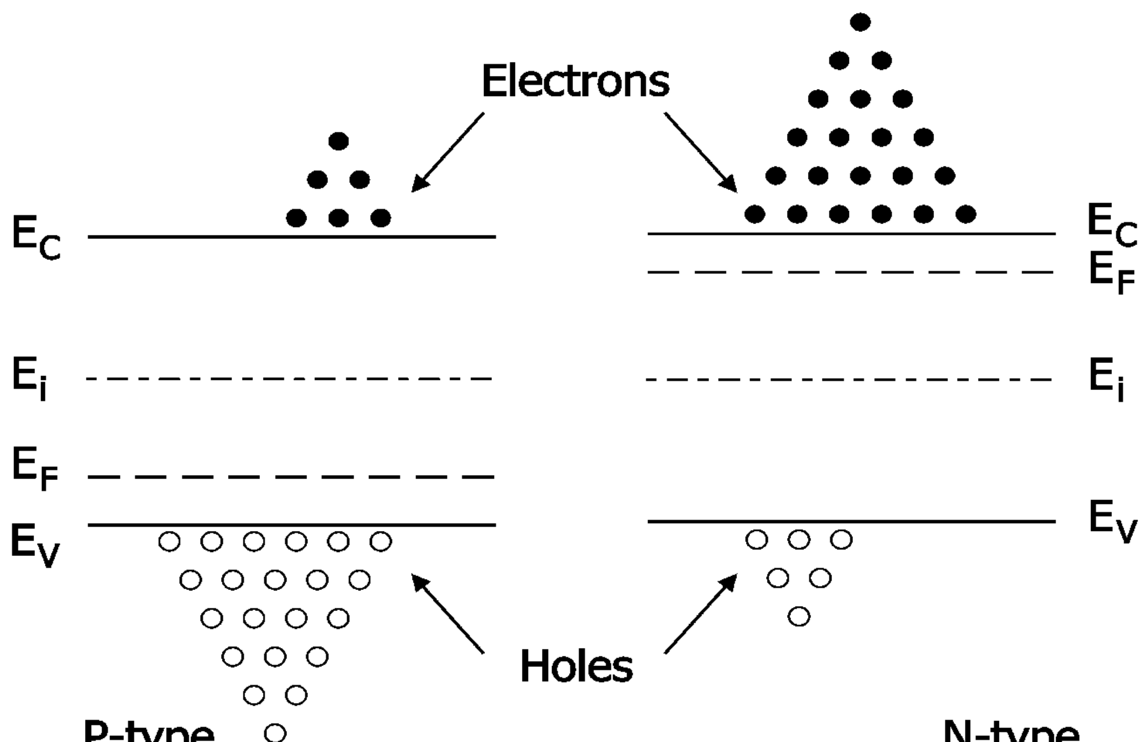
Space Charge Width

- Total space charge width:

$$\boldsymbol{W_0 = x_{n0} + x_{p0}}$$

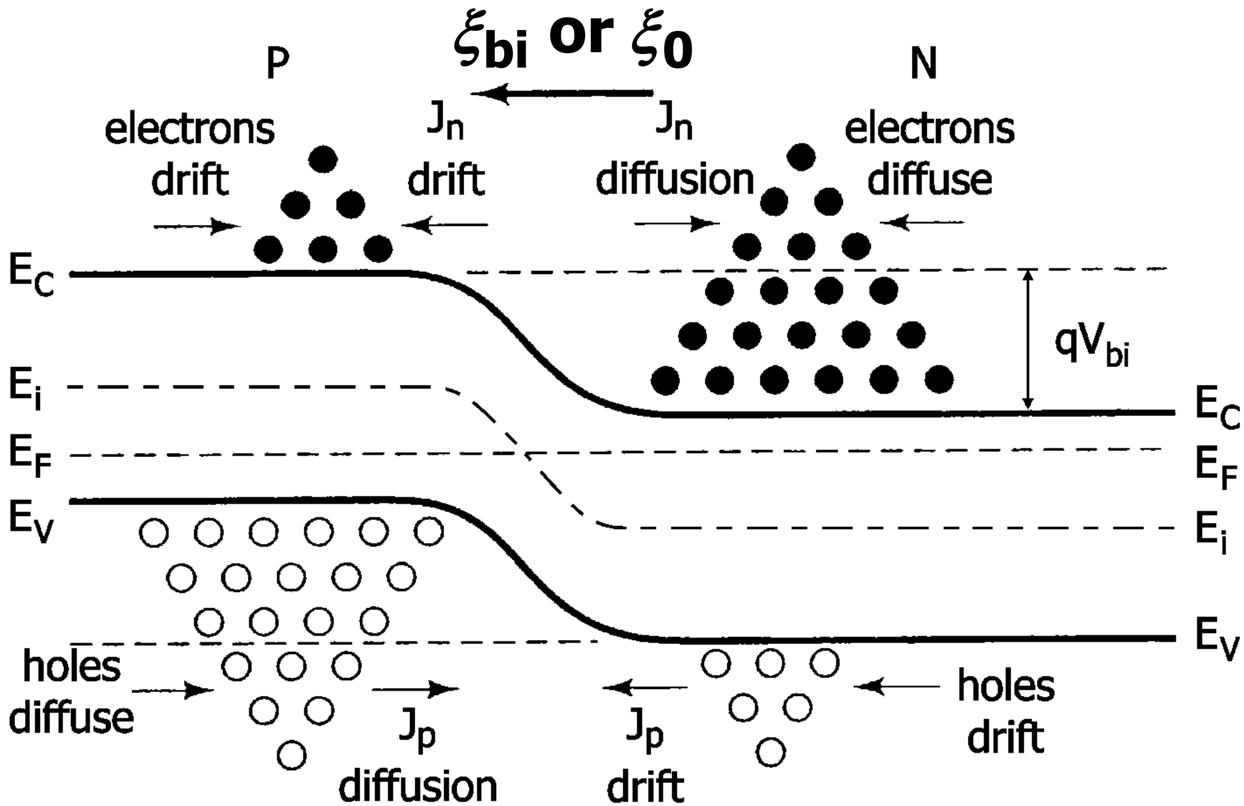
$$= \left\{ \frac{2\varepsilon_r \varepsilon_0 V_{bi}}{q} \left[\frac{N_A + N_D}{N_A N_D} \right] \right\}^{1/2}$$

Energy Band Diagram



Energy band diagrams of the n and p regions before contact

Energy Band Diagram



Energy band diagram of a pn junction under thermal equilibrium

Energy Band Diagram

- Built-in voltage:

$$V_{\mathrm{bi}} = \frac{(E_{\mathrm{i}} - E_{\mathrm{F}})_{\mathrm{p}} + (E_{\mathrm{F}} - E_{\mathrm{i}})_{\mathrm{n}}}{q}$$

- Recall Maxwell-Boltzmann equations:

$$p = N_{\mathrm{V}} \exp\left[\frac{-(E_{\mathrm{F}} - E_{\mathrm{V}})}{kT}\right], \quad n = N_{\mathrm{C}} \exp\left[\frac{-(E_{\mathrm{C}} - E_{\mathrm{F}})}{kT}\right]$$

Energy Band Diagram

- Rewriting:

$$p = N_V \exp\left[\frac{-(E_F - E_i + E_i - E_V)}{kT}\right]$$

$$= N_V \exp\left[\frac{-(E_i - E_V)}{kT}\right] \exp\left[\frac{E_i - E_F}{kT}\right]$$

$$= n_i \exp\left[\frac{E_i - E_F}{kT}\right]$$

$$n = n_i \exp\left[\frac{E_F - E_i}{kT}\right]$$

Energy Band Diagram

- For complete ionization of dopants,

$$N_A = n_i \exp\left[\frac{E_i - E_F}{kT}\right], \quad N_D = n_i \exp\left[\frac{E_F - E_i}{kT}\right]$$

- Taking the natural log on both sides, and re-arranging:

$$(E_i - E_F)_p = kT \ln\left(\frac{N_A}{n_i}\right), \quad (E_F - E_i)_n = kT \ln\left(\frac{N_D}{n_i}\right)$$

Energy Band Diagram

- Built-in voltage:

$$V_{bi} = \frac{kT\ln(N_A / n_i) + kT\ln(N_D / n_i)}{q}$$

$$= \frac{kT}{q}\ln\left(\frac{N_A N_D}{n_i^2}\right)$$

Example 1

- Consider a silicon p-n junction at T=300 K with doping densities $N_A=10^{18}$ cm^{-3} and $N_D=10^{15}$ cm^{-3}. Given that $n_i=1.5\text{x}10^{10}$ cm^{-3} at T=300 K.
  - Calculate the built-in potential barrier.
  - If we change the acceptor doping from 10^{18} cm^{-3} to 10^{16} cm^{-3}, what is the new built-in potential barrier?

Example 1

- The built-in potential barrier is determined using the following relation:

$$V_{bi,1} = \frac{kT}{q}\ln\left(\frac{N_A N_D}{n_i^2}\right)$$

$$= (0.0259)\ln\left[\frac{(10^{18})(10^{15})}{(1.5\times 10^{10})^2}\right]$$

$$= 0.754\ \text{V}$$

Example 1

- If the acceptor doping is 10^{16} cm^{-3} instead of 10^{18} cm^{-3}, the built-in potential barrier is

$$V_{bi,2} = (0.0259)\ln\left[\frac{(10^{16})(10^{15})}{(1.5\times 10^{10})^2}\right]$$

$$= 0.635\text{ V}$$

  - Note that $V_{bi,1}$-$V_{bi,2}$=119mV for a 100 times decrease in acceptor doping concentration.
  - This weak dependence is because of the logarithmic function.

Example 2

- Consider a silicon p-n junction at T=300 K with doping concentrations of $N_A=10^{16}$cm^{-3} and $N_D=10^{15}$ cm^{-3}.
  - Calculate the space charge width
  - Calculate the maximum electric field

- In example 1, we have already determined the built-in voltage as V_{bi}=0.635 V.

Example 2

- Since we know the built-in voltage and the doping concentrations of the n and p regions, the space charge width can be easily calculated:

$$W_0 = \left\{ \frac{2\varepsilon_r \varepsilon_0 V_{bi}}{q} \left[\frac{N_A + N_D}{N_A N_D} \right] \right\}^{1/2}$$

$$= \left\{ \frac{2(11.7)(8.85 \times 10^{-14})(0.635)}{1.6 \times 10^{-19}} \left[\frac{10^{16} + 10^{15}}{(10^{16})(10^{15})} \right] \right\}^{1/2}$$

$$= 0.951 \times 10^{-4} \text{ cm} = 0.951\ \mu\text{m}$$

Example 2

- We can also calculate the respective space charge widths in the n and p regions:

$$x_{n0} = 0.864\ \mu m, \qquad x_{p0} = 0.086\ \mu m$$

- Note that $x_{n0} >> x_{p0}$.

- **Why?** Because of space charge neutrality:

$$\underbrace{qN_A x_{p0}}_{\text{Net negative charge on p-type side}} = \underbrace{qN_D x_{n0}}_{\text{Net positive charge on n-type side}}$$

Example 2

- The maximum electric field occurs at the metallurgical junction, i.e. $x = 0$.

$$
\begin{aligned}
\xi_{\mathrm{m}} \text{ or } \xi_{\mathrm{max}} &= \xi(x = 0) \\
&= -\frac{qN_{\mathrm{D}}x_{\mathrm{n0}}}{\varepsilon_{\mathrm{r}}\varepsilon_{0}} \\
&= \frac{(1.6\times10^{-19})(10^{15})(0.864\times10^{-4})}{(11.7)(8.85\times10^{-14})} \\
&= -1.34\times10^{4}\ \mathrm{V/cm}
\end{aligned}
$$

Summary

- Electrostatics of the pn junction:
  - Space charge/depletion region
  - Built-in electric field
    - Maximum electric field @ metallurgical junction
  - Built-in potential barrier
    - Area under the electric field distribution plot
    - Energy band diagram
  - Space charge width

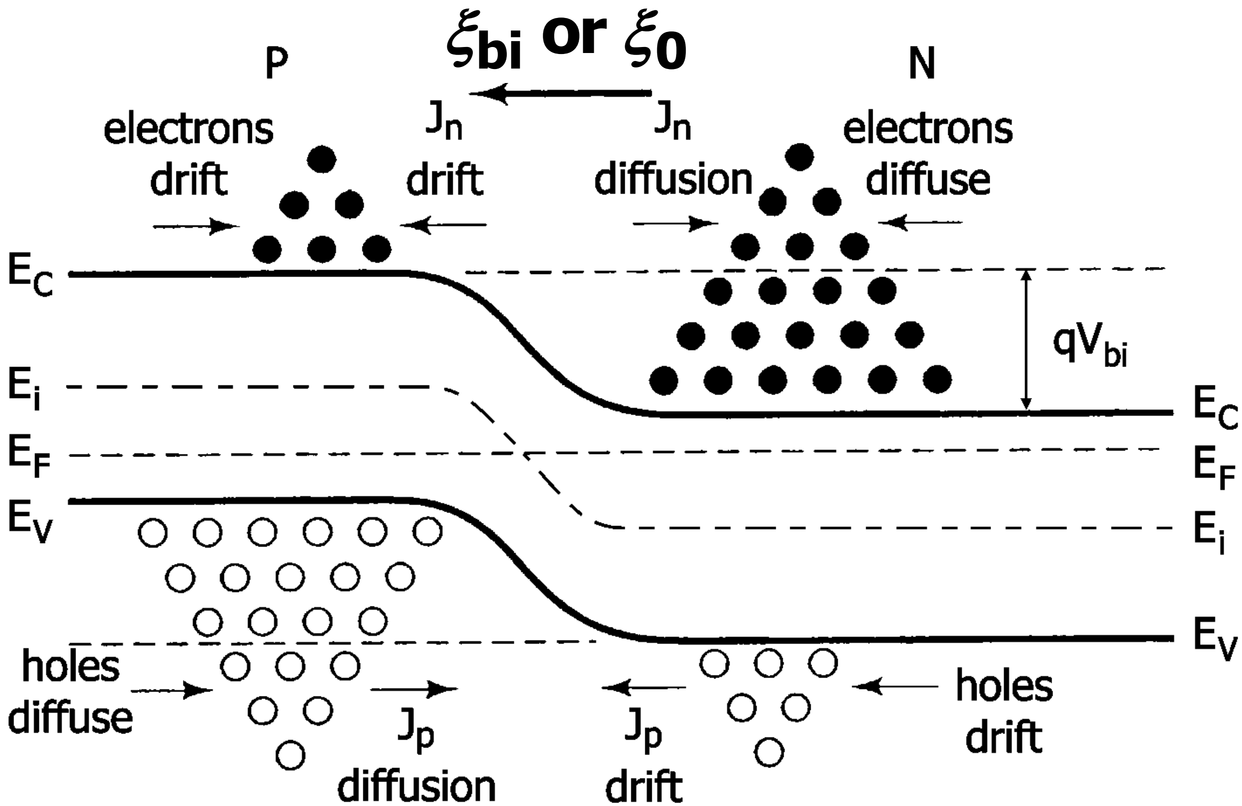
EE2003 Semiconductor Fundamentals

Effect of an Externally Applied Voltage on the P-N Junction

Objectives

- What happens when an external voltage is applied to the terminals of the p-n junction?
- Why does the diode conduct only in one direction?
- Two cases:
  - **Forward bias** – when the p region is made more positive with respect to the n region (large current flow).
  - **Reverse bias** – when the p region is made more negative with respect to the n region (very small current flow)

Thermal Equilibrium

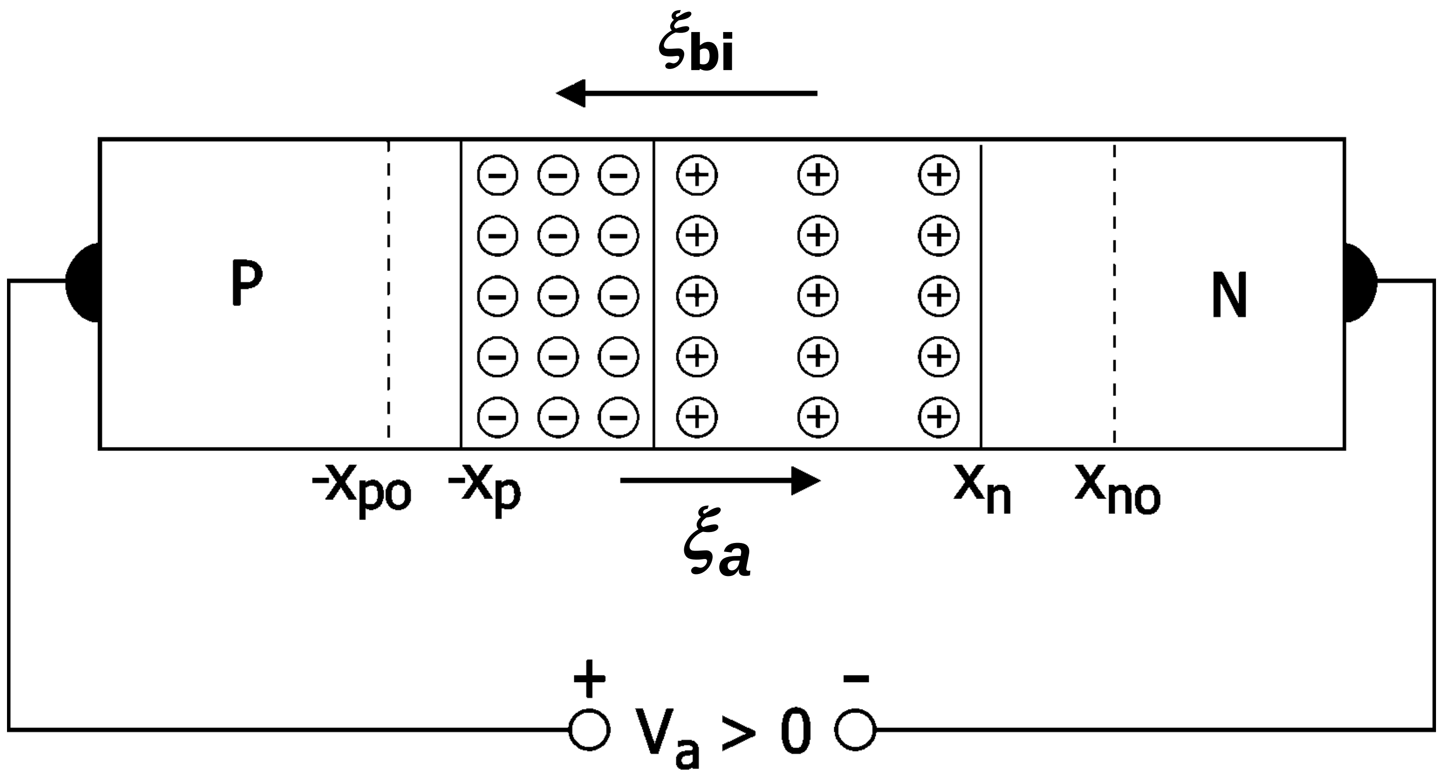


Energy band diagram of a pn junction under thermal equilibrium

Thermal Equilibrium

- Conditions associated with thermal equilibrium:
  - Only those electrons with energy greater than the energy barrier (qV_{bi}) can diffuse from the n to the p region.
  - This diffusion is balanced by the drift of electrons from the p to the n region.
  - Net electron current across the junction = 0
  - A similar situation exists for the holes.

Forward Bias



Schematic of a p-n junction being applied with a forward voltage bias. The directions of the applied and built-in electric field are shown.

Forward Bias

- Conditions associated with forward bias:
  - The applied field opposes the built-in field. The net electric field at the junction becomes smaller, i.e. $\xi = \xi_0 - \xi_a$.
  - Because the net electric field has reduced, the space charge width decreases, i.e. $W < W_0$.
    - The magnitude of the electric field is determined by the amount of positive and negative charges in the n and p regions respectively.
    - Since the overall electric field has decreased, the amount of positive and negative charges must decrease correspondingly.
    - Therefore, the space charge width decreases under forward bias.

$$\left|\xi_{\mathbf{m}}\right| = \left|\xi(\boldsymbol{x} = \mathbf{0})\right| = \frac{qN_A x_p}{\varepsilon_r \varepsilon_0} = \frac{qN_D x_n}{\varepsilon_r \varepsilon_0}$$

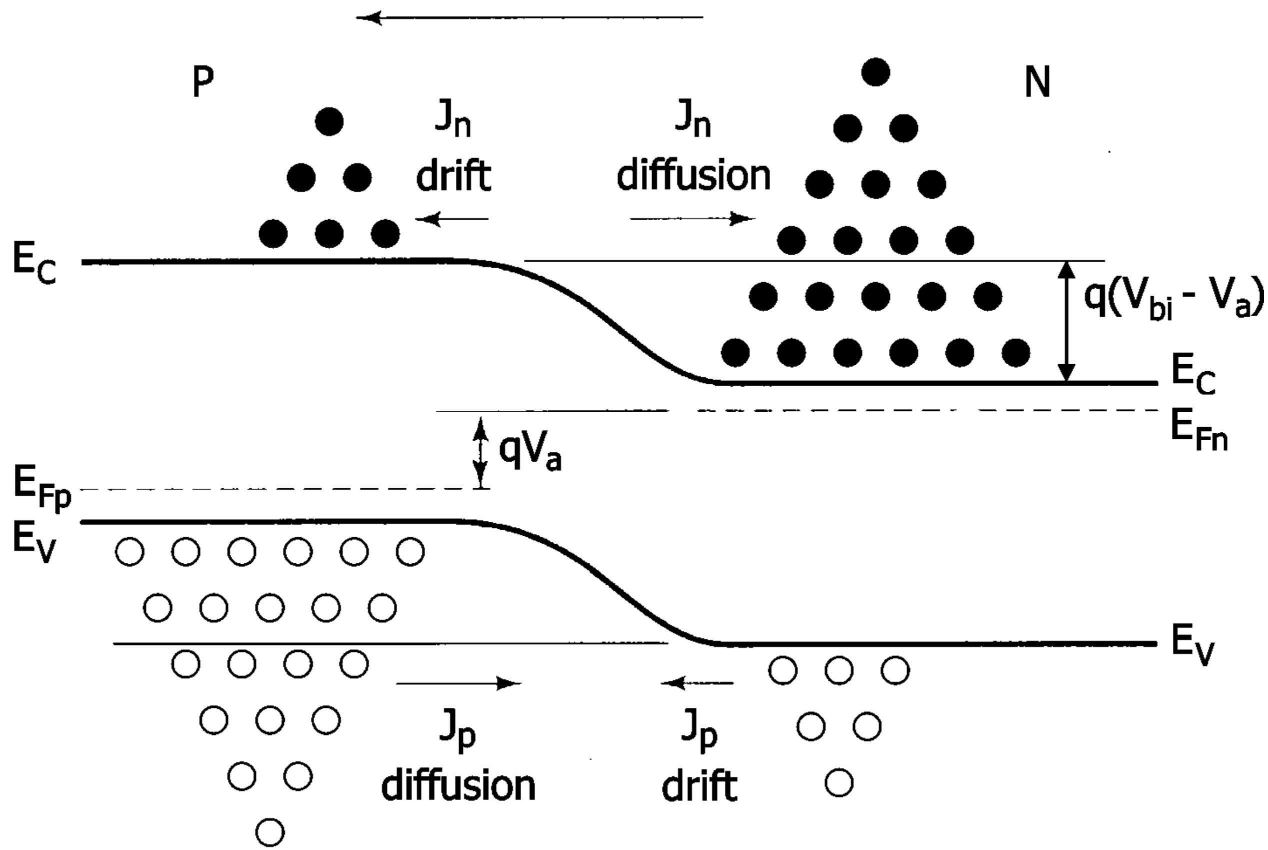
Forward Bias

- Conditions associated with forward bias (cont'd):
  - Since the area under the electric field distribution plot is now lesser than that under thermal equilibrium, the junction potential also decreases, i.e. $V_j = V_{bi} - V_a$.
  - The energy band diagram has to be modified accordingly:
    - Because the electric field is now smaller, the slope of the energy band in the space charge region must decrease.
    - Using the n region as reference, the energy levels in the p region must therefore be **lowered** relative to those in the n region.
    - The Fermi level is no longer constant across all three regions.
    - The difference in the Fermi levels in the neutral n and p regions is proportional to the applied voltage.

$$\xi = \frac{1}{q}\frac{dE}{dx}$$

Forward Bias

$$\xi = \xi_0 - \xi_a$$

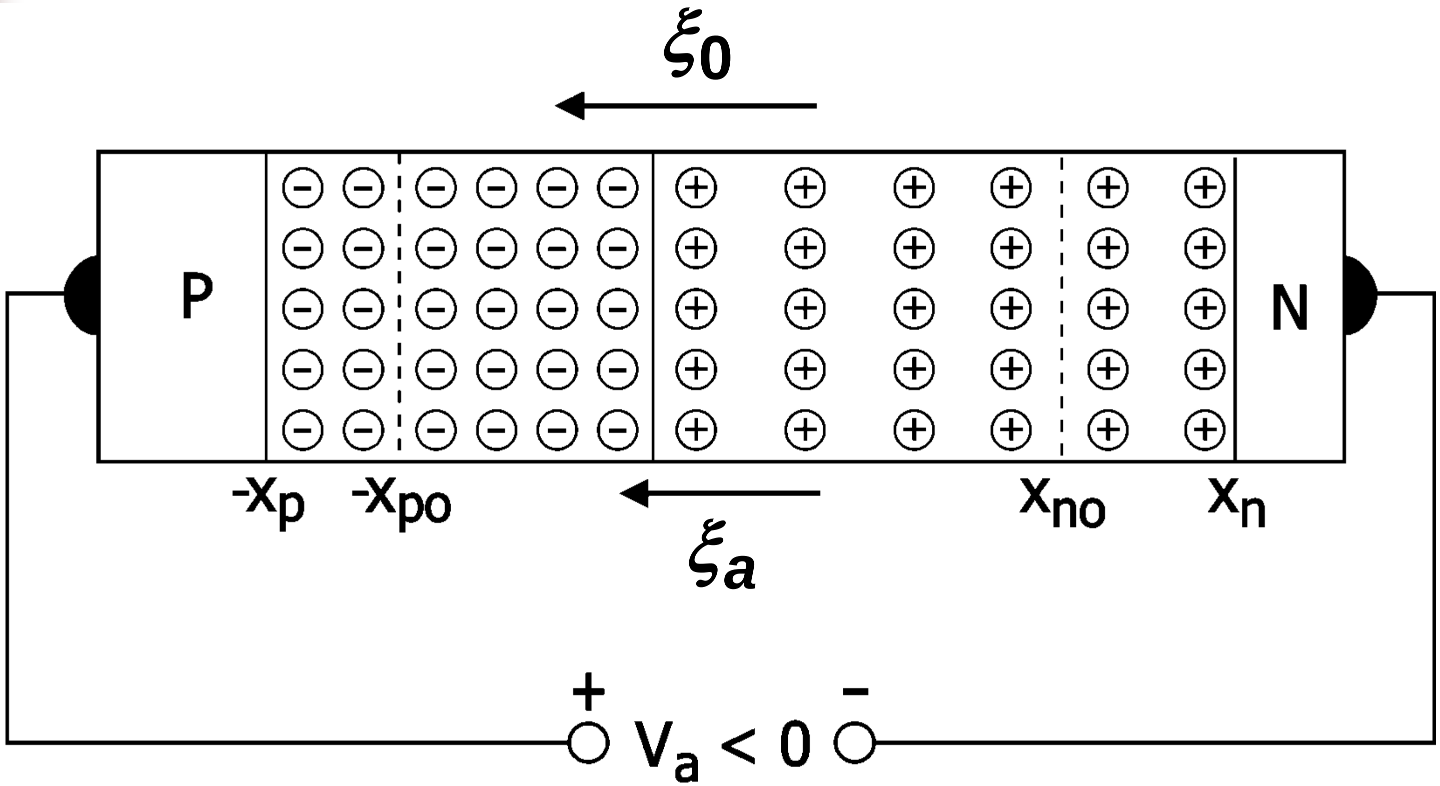


Energy band diagram of a pn junction under forward bias

Forward Bias

- Conditions associated with forward bias (cont'd):
  - Because the **potential barrier** is reduced under forward bias,
    - More electrons are now able to diffuse from the n to the p region.
    - Likewise, more holes are now able to diffuse from the p to the n region.
    - A **net diffusion** of charges across the p-n junction takes place.
    - We commonly say that when a p-n junction is forward biased, electrons are **injected** from the n into the p region and holes are injected from the p into the n region, i.e. under a forward bias, we have **carrier injection** across the p-n junction.
    - Carrier injection is the reason behind the ability of the diode to conduct a large current under a forward bias.

Reverse Bias



A pn junction with an applied reverse-bias voltage showing the directions of the applied field and the built-in field.

Reverse Bias

- Conditions associated with a reverse bias:
  - The applied field appears in the same direction as the built-in field. The net electric field at the junction becomes larger, i.e. $\xi = \xi_0 + \xi_a$.
  - Because the net electric field has increased, the space charge region widens, i.e. $W > W_0$.
  - Since the area under the electric field distribution plot would now be greater than that under thermal equilibrium, the junction potential also increases, i.e. $V_j = V_{bi} - V_a = V_{bi} + |V_a|$.

Reverse Bias

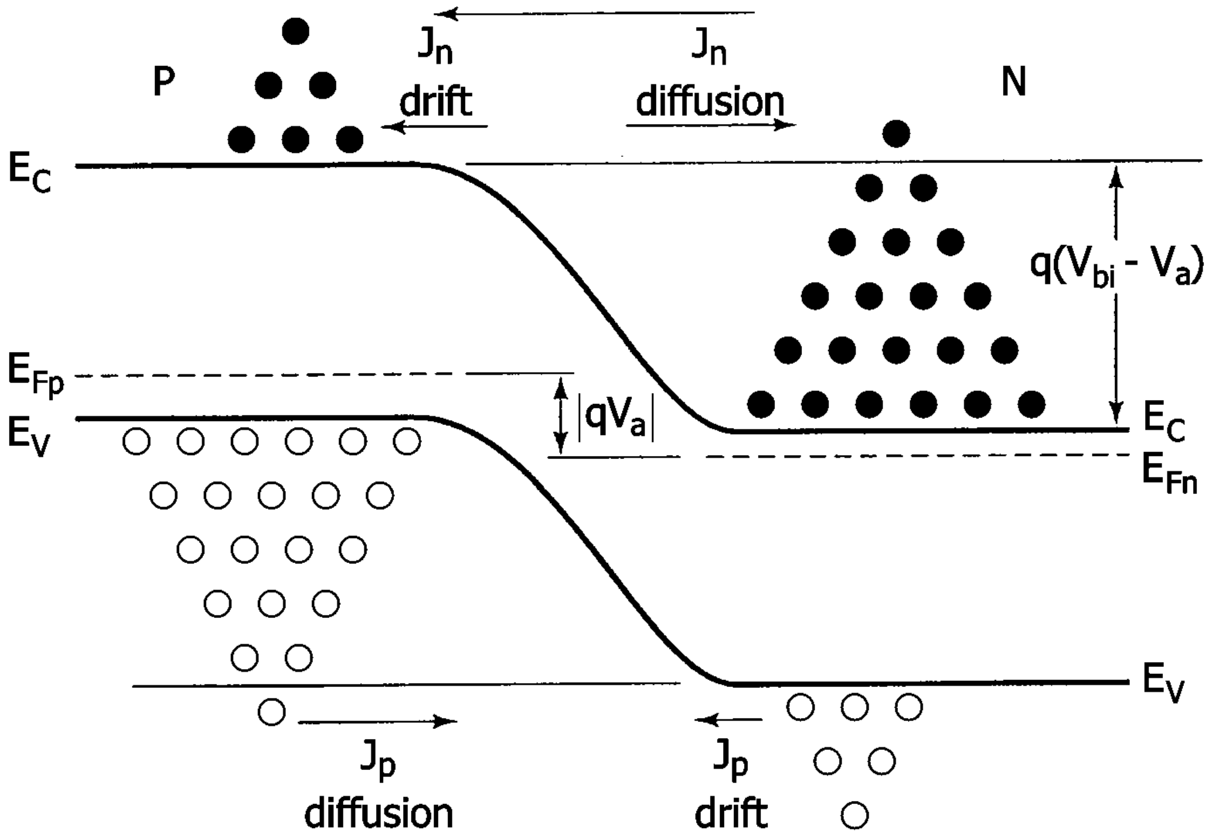
- Conditions associated with a reverse bias (cont'd):
  - The energy band diagram has to be modified accordingly:
    - Because the electric field is now larger, the slope of the energy band in the space charge region must increase.
    - Using the n region as reference, the energy levels in the p region must therefore be **raised**, relative to those in the n region.

Reverse Bias

- Conditions associated with a reverse bias (cont'd):
  - Because the **potential barrier** is enhanced under reverse bias,
    - Fewer electrons can diffuse from n to the p region. Similarly, fewer holes are able to diffuse from the p to the n region.
    - The net current is due to the diffusion of (i) electrons from the p to the edge of the space charge region and then drift across to the n region and (ii) diffusion of holes from the n to the edge of the space charge region and then drift across to the p region.
    - These are minority carriers that are present in very small quantities.
    - Hence, a reverse biased p-n junction does not conduct any current except for a small leakage current.

Reverse Bias

$$\xi = \xi_0 + \xi_a$$



Energy band diagram of a pn junction under reverse bias

Modified Formulae

- All the equations derived previously still apply, but some modifications are needed.
- Since the application of an external bias changes the junction field and hence the junction potential, the term V_{bi} has to be replaced with V_{bi}-V_a.
- **Please remember:**
  - Under forward bias, V_a is positive $\Rightarrow V_{bi} - V_a < V_{bi}$
  - Under reverse bias, V_a is negative $\Rightarrow V_{bi} - V_a > V_{bi}$

Modified Formulae

- Width of space charge regions:

$$x_n = \left\{ \frac{2\varepsilon_r \varepsilon_0 \left(V_{bi} - V_a\right)}{q} \left[\frac{N_A}{N_D \left(N_A + N_D\right)} \right] \right\}^{1/2}$$

$$x_p = \left\{ \frac{2\varepsilon_r \varepsilon_0 \left(V_{bi} - V_a\right)}{q} \left[\frac{N_D}{N_A \left(N_A + N_D\right)} \right] \right\}^{1/2}$$

$$W = \left\{ \frac{2\varepsilon_r \varepsilon_0 \left(V_{bi} - V_a\right)}{q} \left[\frac{N_A + N_D}{N_A N_D} \right] \right\}^{1/2}$$

Modified Formulae

- Maximum electric field:

$$\left|\xi_{\mathbf{m}}\right| = \left|\xi(\boldsymbol{x} = \mathbf{0})\right| = \frac{q N_{\mathrm{A}} x_{\mathrm{p}}}{\varepsilon_{\mathrm{r}} \varepsilon_{\mathbf{0}}} = \frac{q N_{\mathrm{D}} x_{\mathrm{n}}}{\varepsilon_{\mathrm{r}} \varepsilon_{\mathbf{0}}}$$

  - The same expression for the maximum field applies, except that x_{n} replaces x_{n0} and x_{p} replaces x_{p0}.

Example 1

- Consider a silicon p-n junction at T=300 K with doping concentrations of $N_A=10^{16}$ cm^{-3} and $N_D=10^{15}$ cm^{-3}. You may assume $n_i=1.5\text{x}10^{10}$cm^{-3} at T=300 K.
  - Calculate the width of the space charge region when a reverse-bias voltage of 5 V is applied.
  - What is the space charge width if the reverse bias is increased to 15 V?

Example 1

- The space charge width can be determined using the following equation:

$$\boldsymbol{W} = \left\{ \frac{2\varepsilon_r \varepsilon_0 \left(V_{bi} - V_a \right)}{q} \left[\frac{N_A + N_D}{N_A N_D} \right] \right\}^{1/2}$$

Not very sure what this is

  - For reverse bias, V_a = –5 V.
  - To proceed, the built-in voltage, V_{bi} should first be calculated.
  - From example 1 (Electrostatics chapter), we have V_{bi} = 0.635 V.

Example 1

- Therefore,

$$W = \left\{ \frac{2(11.7)(8.85\times10^{-14})(0.635+5)}{1.6\times10^{-19}} \left[\frac{10^{16}+10^{15}}{(10^{16})(10^{15})} \right] \right\}^{1/2}$$

$$= 2.83\times10^{-4}\ \text{cm} = 2.83\ \mu\text{m}$$

- Comments:
  - From example 2 (Electrostatics chapter), we determined W_0=0.951 μm.
  - Thus, the space charge region has increased from 0.951 μm at zero bias to 2.83 μm at a reverse bias of 5 V.

Example 1

- We need not re-calculate the new space charge width all over again using the previous equation. Instead, one should note that:

$$W \propto \left(V_{bi} - V_a\right)^{1/2}$$

- Thus, the new space charge width at $V_a = -15$ V is

$$W^* = W \cdot \left[\frac{V_{bi} + 15}{V_{bi} + 5}\right]^{1/2}$$
$$= 4.71\ \mu\text{m}$$

Example 2

- Consider a silicon p-n junction at T=300 K with a p-type doping concentration of $N_A=10^{18}$ cm^{-3}.
  - Determine the n-type doping concentration such that the maximum electric field does not exceed $3\text{x}10^5$ V/cm at a reverse-bias voltage of 25 V.

- The maximum electric field occurs at the metallurgical junction and is given by

$$\xi_m = -\left\{ \frac{2q(V_{bi} - V_a)}{\varepsilon_r \varepsilon_0} \left[\frac{N_A N_D}{N_A + N_D} \right] \right\}^{1/2}$$

Recall (Electrostatics Chapter)

Thermal Equilibrium

$$\left|\xi_{\mathrm{m}}\right| = \left|\xi(x=0)\right| = \frac{qN_{\mathrm{A}}x_{\mathrm{p0}}}{\varepsilon_{\mathrm{r}}\varepsilon_0} = \frac{qN_{\mathrm{D}}x_{\mathrm{n0}}}{\varepsilon_{\mathrm{r}}\varepsilon_0}$$

$$x_{\mathrm{n0}} = \left\{\frac{2\varepsilon_{\mathrm{r}}\varepsilon_0 V_{\mathrm{bi}}}{q}\left[\frac{N_{\mathrm{A}}}{N_{\mathrm{D}}(N_{\mathrm{A}}+N_{\mathrm{D}})}\right]\right\}^{1/2}$$

$$x_{\mathrm{p0}} = \left\{\frac{2\varepsilon_{\mathrm{r}}\varepsilon_0 V_{\mathrm{bi}}}{q}\left[\frac{N_{\mathrm{D}}}{N_{\mathrm{A}}(N_{\mathrm{A}}+N_{\mathrm{D}})}\right]\right\}^{1/2}$$

With Applied Voltage

$$\left|\xi_{\mathrm{m}}\right| = \left|\xi(x=0)\right| = \frac{qN_{\mathrm{A}}x_{\mathrm{p}}}{\varepsilon_{\mathrm{r}}\varepsilon_0} = \frac{qN_{\mathrm{D}}x_{\mathrm{n}}}{\varepsilon_{\mathrm{r}}\varepsilon_0}$$

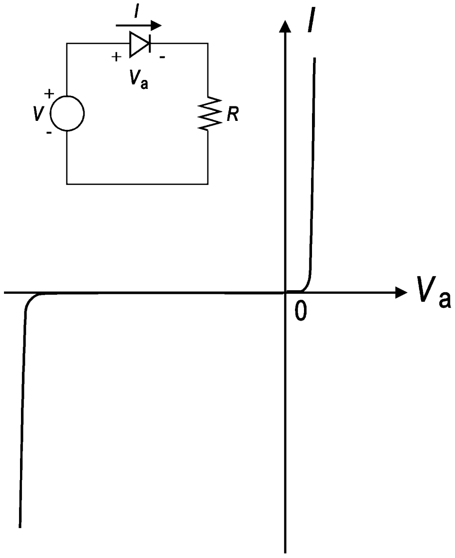
$$x_{\mathrm{n}} = \left\{\frac{2\varepsilon_{\mathrm{r}}\varepsilon_0 (V_{\mathrm{bi}}-V_{\mathrm{a}})}{q}\left[\frac{N_{\mathrm{A}}}{N_{\mathrm{D}}(N_{\mathrm{A}}+N_{\mathrm{D}})}\right]\right\}^{1/2}$$

$$x_{\mathrm{p}} = \left\{\frac{2\varepsilon_{\mathrm{r}}\varepsilon_0 (V_{\mathrm{bi}}-V_{\mathrm{a}})}{q}\left[\frac{N_{\mathrm{D}}}{N_{\mathrm{A}}(N_{\mathrm{A}}+N_{\mathrm{D}})}\right]\right\}^{1/2}$$

Example 2

- It should be noted that for a given N_A, ξ_m depends on both V_{bi} and N_D.
- V_{bi} is in turn dependent on N_D via the relation: $V_{bi} = (kT/q)\ln(N_A N_D/n_i^2)$.
- Hence, the problem is quite complicated and cannot be easily solved unless an iterative method is used.
- However, such a precise solution is often not needed in practice.

Example 2



- Since $-V_a = 25 >> V_{bi}$ (~0.7 V),

$$V_{bi} - V_a >> V_{bi} \Rightarrow V_{bi} - V_a \approx -V_a$$

- Therefore,

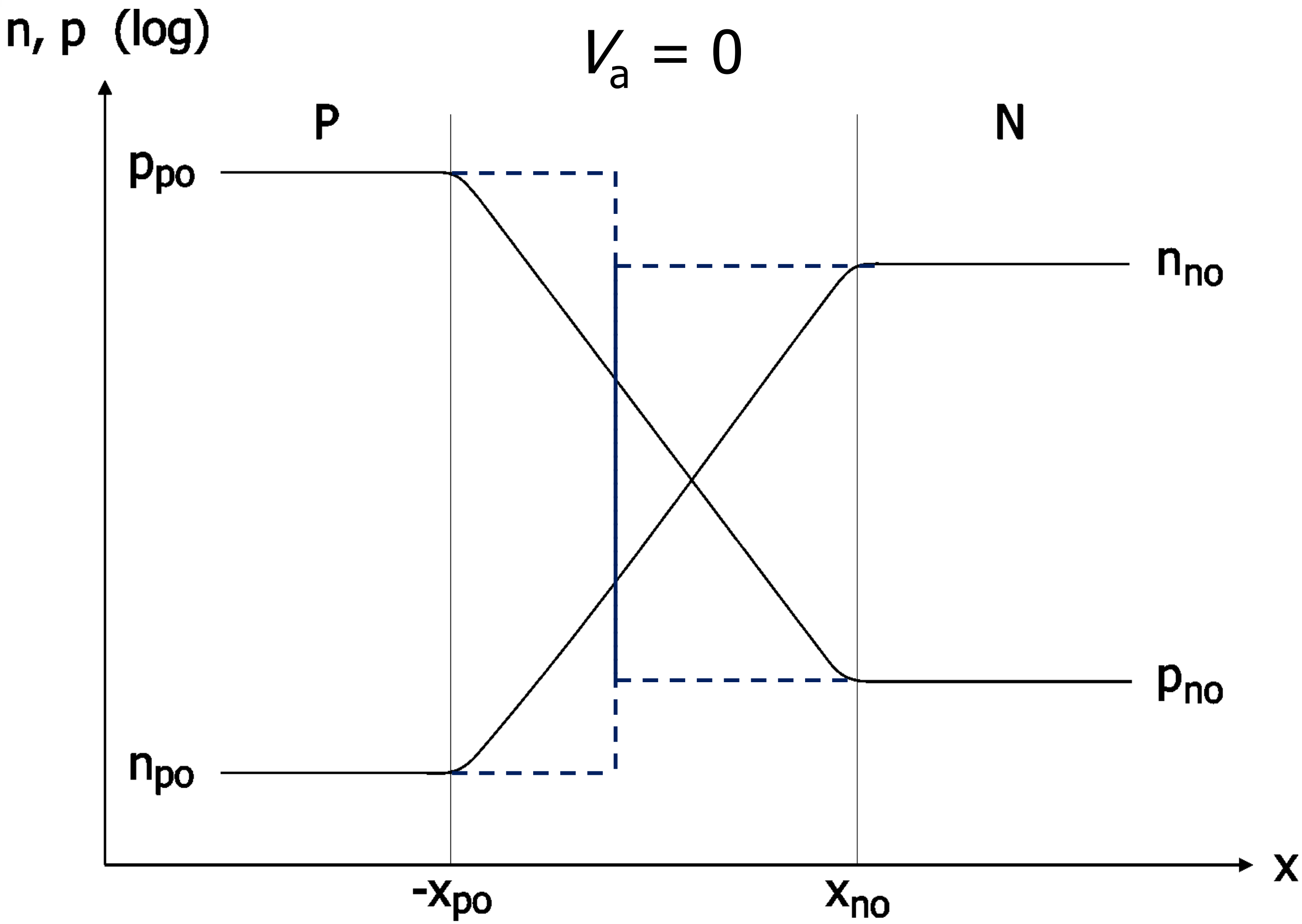
$$\xi_m = -\left\{ \frac{2q(-V_a)}{\varepsilon_r \varepsilon_0} \left[\frac{N_A N_D}{N_A + N_D} \right] \right\}^{1/2}$$

- We require

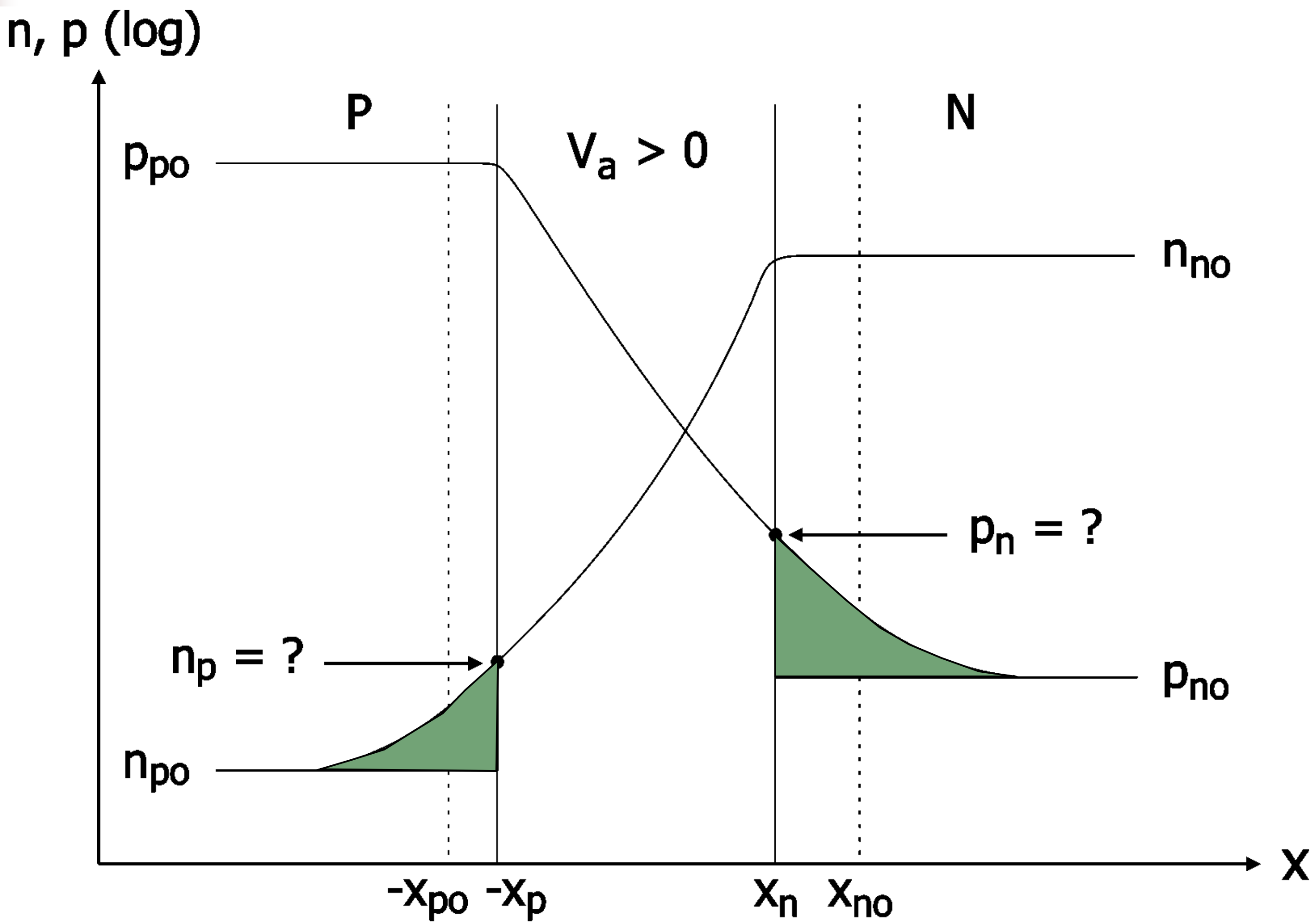
$$\xi_m = -\left\{ \frac{2q(-V_a)}{\varepsilon_r \varepsilon_0} \left[\frac{N_A N_D}{N_A + N_D} \right] \right\}^{1/2} \leq 3 \times 10^5 \text{ V/cm}$$

$$N_D \leq 1.18 \times 10^{18} \text{ cm}^{-3}$$

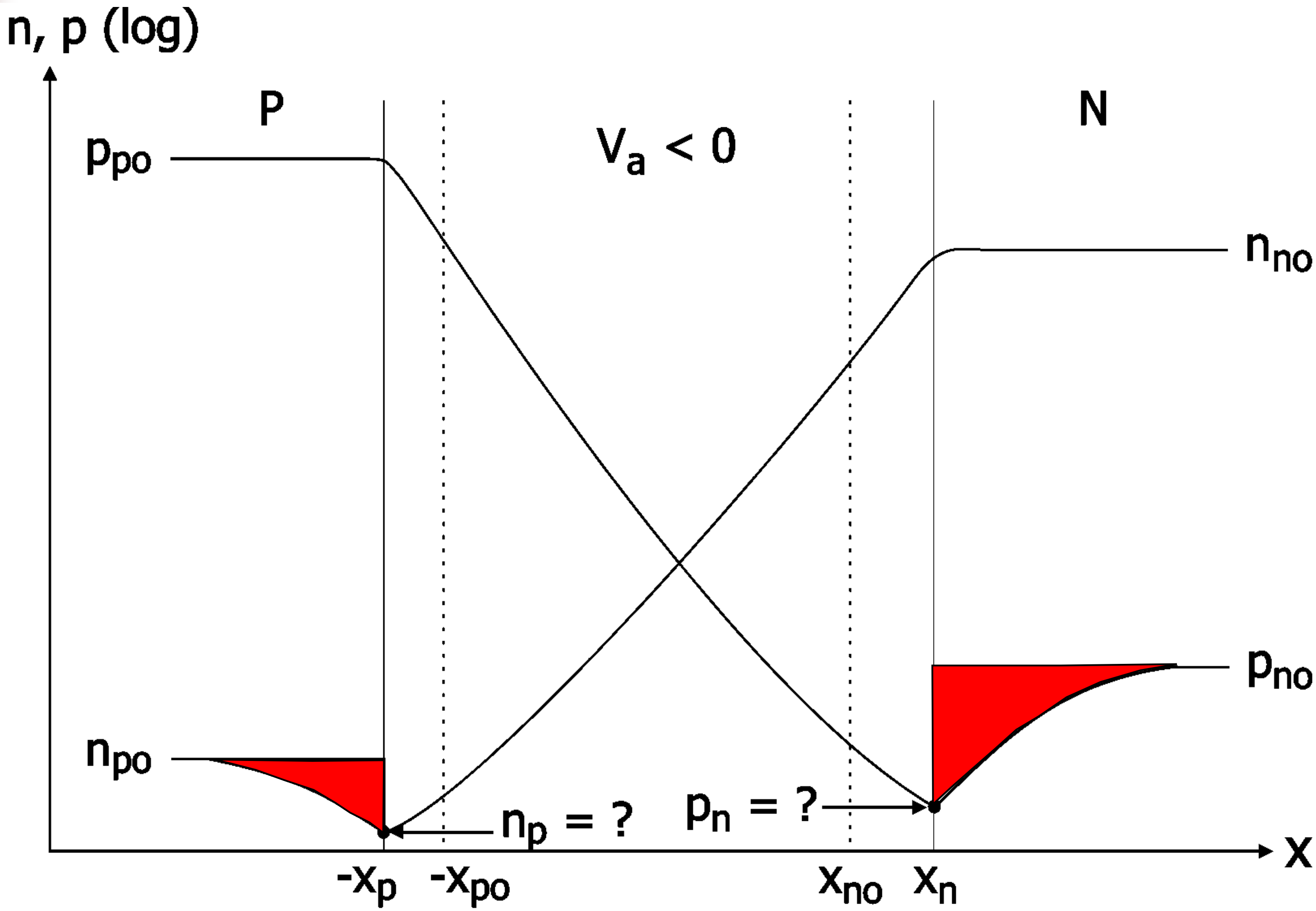
Carrier Distributions – Thermal Equilibrium



Carrier Distributions – Forward Bias



Carrier Distributions – Reverse Bias



Minority Carrier Densities

- Under an applied voltage bias, the minority carrier densities at the depletion edges are modified, either due to charge injection under forward bias, or charge extraction under reverse bias.
- As we will see in the next section, the subsequent transport **(diffusion)** of these minority carriers in the respective quasi-neutral n and p regions determines the amount of current flow in the diode at a given bias.
- Hence, it is of interests to be able to relate the change in minority carrier densities at the depletion edges to the applied voltage bias as this would help establish the boundary condition needed to solve the continuity equations.

Minority Carrier Densities

- Recall:

$$V_{bi} = \frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right)$$

- At normal operating temperature (for example 300 K and above), the dopant atoms are fully ionized. Moreover, the dopant concentration is usually in the order of 10^{16} cm^{-3} >> n_i (=1.5x10^{10} cm^{-3}). Hence,

$$n_{no} \approx N_D\,, \quad p_{no} \approx \frac{n_i^2}{N_D}$$

$$p_{po} \approx N_A\,, \quad n_{po} \approx \frac{n_i^2}{N_A}$$

Minority Carrier Densities

- Rewriting the expression for the built-in potential:

$$V_{bi} = \frac{kT}{q}\ln\left(\frac{p_{p0}}{p_{n0}}\right) = \frac{kT}{q}\ln\left(\frac{n_{n0}}{n_{p0}}\right)$$

$$\Rightarrow \qquad p_{n0} = p_{p0}e^{-qV_{bi}/(kT)}$$

$$n_{p0} = n_{n0}e^{-qV_{bi}/(kT)}$$

  - The minority carrier concentration at the depletion edge is exponentially related to the majority carrier concentration at the opposite depletion edge through the built-in potential V_{bi}.

Minority Carrier Densities

- The same relationship applies under an external bias, with V_{bi} replaced by V_{bi} - V_a, i.e.

$$p_n = p_p e^{-q(V_{bi}-V_a)/(kT)}, \; n_p = n_n e^{-q(V_{bi}-V_a)/(kT)}$$

- If the applied bias is moderately low such that **low-level injection** condition prevails, $p_p \approx p_{p0}$ and $n_n \approx n_{n0}$. Therefore,

$$p_n = p_{p0} e^{\frac{-q(V_{bi}-V_a)}{kT}} = p_{p0} e^{\frac{-qV_{bi}}{kT}} e^{\frac{qV_a}{kT}} = p_{n0} e^{\frac{qV_a}{kT}}$$

$$n_p = n_{n0} e^{\frac{-q(V_{bi}-V_a)}{kT}} = n_{n0} e^{\frac{-qV_{bi}}{kT}} e^{\frac{qV_a}{kT}} = n_{p0} e^{\frac{qV_a}{kT}}$$

Note: $e^{-\frac{q(V_{bi}-V_a)}{kT}} = e^{-\frac{qV_{bi}}{kT}} \cdot e^{\frac{qV_a}{kT}} \quad \left(e^{a+b} = e^a \cdot e^b\right)$

Minority Carrier Densities

- '**Excess**' minority carrier densities:

$$\Delta p_{\text{n}} = p_{\text{n0}} e^{qV_{\text{a}}/kT} - p_{\text{n0}} = p_{\text{n0}} \left(e^{qV_{\text{a}}/kT} - 1 \right)$$

$$\Delta n_{\text{p}} = n_{\text{p0}} e^{qV_{\text{a}}/kT} - n_{\text{p0}} = n_{\text{p0}} \left(e^{qV_{\text{a}}/kT} - 1 \right)$$

- Forward bias: $V_a > 0$, V_a usually >> kT/q

$$\Delta p_{\text{n}} = p_{\text{n0}} e^{qV_{\text{a}}/kT}, \ \Delta n_{\text{p}} = n_{\text{p0}} e^{qV_{\text{a}}/kT}$$

- Reverse bias: $V_a < 0$, $|V_a|$ usually >> kT/q

$$\Delta p_{\text{n}} = -p_{\text{n0}}, \ \Delta n_{\text{p}} = -n_{\text{p0}}$$

Summary

- P-N junction under non-equilibrium:
  - Why a p-n junction conducts only in the forward-bias mode, and not in the reverse-bias mode.
  - Understand how the electrostatics are changed under a non-zero voltage bias.
  - Understand the terms carrier “injection” and “extraction”, as applied to forward and reverse bias, respectively.
  - How the minority carrier concentrations at the edges of the space charge region under a given voltage bias are calculated.

EE2003 Semiconductor Fundamentals

Ideal Diode Equation

Review of Basic Concepts

- By now, you should be able to do the following:
  - Explain why a p-n junction conducts only under forward bias and not under reverse bias.
  - Draw the energy band diagrams of a pn junction under forward and reverse biases.
  - How should the various formulae derived in section 2 be modified under an applied bias.
  - Relate the ‘excess’ minority carrier concentrations at the respective depletion edges to the applied voltage.
  - Explain the meaning of low-level injection.

Further Objective

- To analyze the p-n junction under an externally applied voltage bias **quantitatively**
  - To obtain a mathematical expression that relates the current flow to the applied bias.
- The problem at hand is one that concerns the transport (diffusion) of ‘excess’ carriers and could be analyzed by solving the continuity equations for electrons and holes.

Assumptions

- The n and p regions are uniformly doped all the way to the metallurgical junction, i.e. the p-n junction is abrupt.
- The lengths of the n an p regions are infinitely long compared to the minority carrier diffusion lengths.
- The applied voltage is sufficiently low such that low-level injection is valid.
- There is negligible or no voltage drop across the quasi-neutral n and p regions (i.e. no electric field in these regions).
- Thermal equilibrium prevails at the end contacts.

Assumptions

- There are no other external sources of excitation except the applied voltage.
- The excess carriers do not recombine in the space charge region (forward bias).
- No generation of carriers takes place in the space charge region (reverse bias).
- The p-n junction has reached steady state.

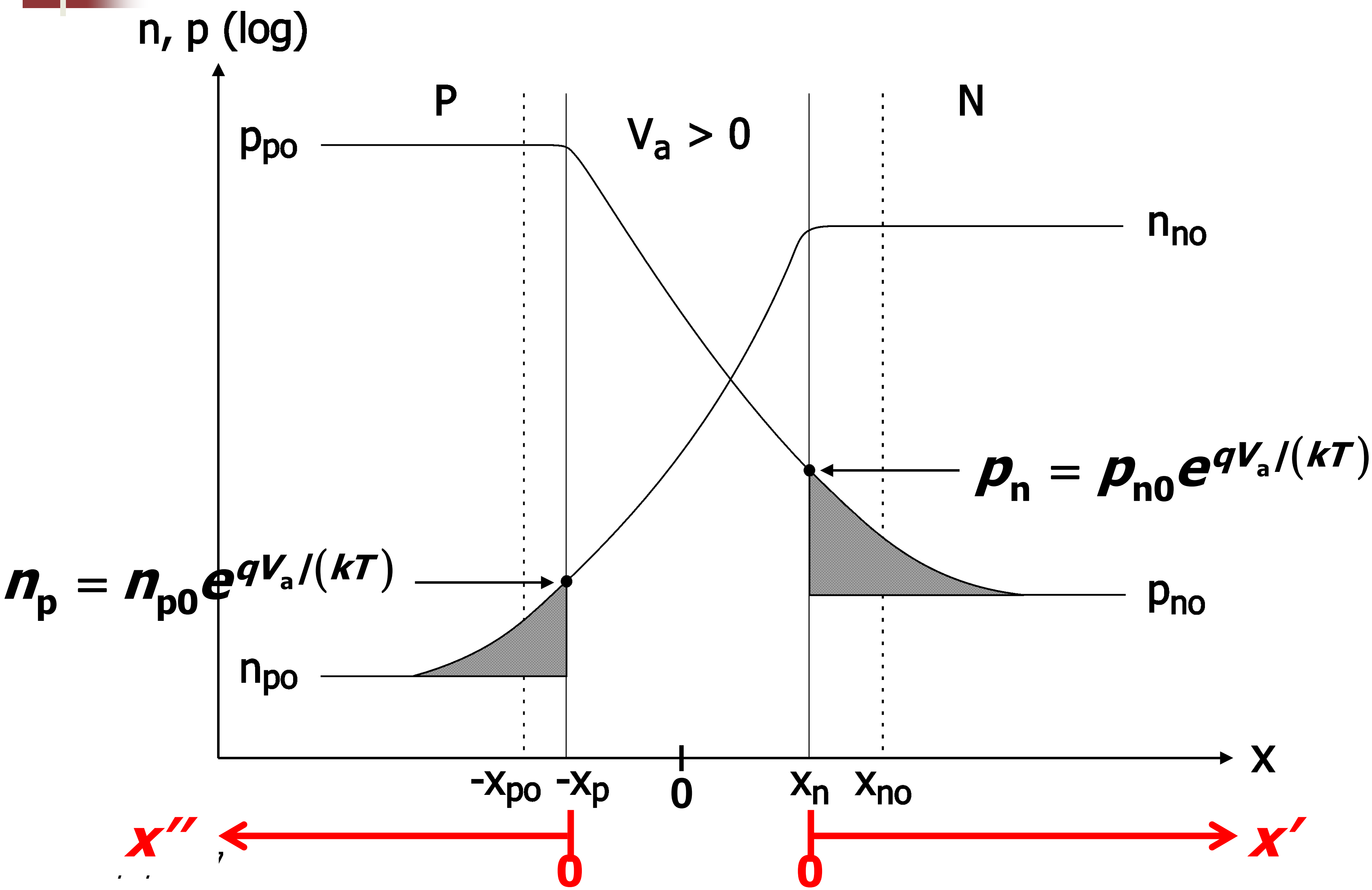
'Excess' Hole Distribution

- The steady-state 'excess' hole distribution in the quasi-neutral n region can be obtained by solving the continuity equation:

$$\frac{\partial p_n}{\partial t} = -\mu_p \left(\xi \frac{\partial p_n}{\partial x'} + p_n \frac{\partial \xi}{\partial x'} \right) + D_p \frac{\partial^2 p_n}{\partial x'^2} + G_p - \frac{\Delta p_n}{\tau_p}$$

- In order to solve this equation analytically, we need to reduce it to a simpler form using some of the assumptions made earlier.

Carrier Distributions – Forward Bias



'Excess' Hole Distribution

- Steady-state operation $\Rightarrow \partial p_n/\partial t = 0$
- No other external sources of excitation $\Rightarrow G_p = 0$
- No electric field in quasi-neutral region $\Rightarrow \xi = 0$
- Uniform doping $\Rightarrow \partial^2 p_n/\partial x' = \partial^2 \Delta p_n/\partial x'$
- Simplified continuity equation:

$$D_p \frac{\partial^2 \Delta p_n}{\partial x'^2} - \frac{\Delta p_n}{\tau_p} = 0$$

$$\frac{\partial^2 \Delta p_n}{\partial x'^2} = \frac{\Delta p_n}{D_p \tau_p} = \frac{\Delta p_n}{L_p^2}$$

‘Excess’ Hole Distribution

- The general solution is of the form:

$$\Delta p_{n}(x') = C_{1}e^{-x'/L_{p}} + C_{2}e^{x'/L_{p}}$$

- C_1 and C_2 are integration constants that can be easily determined by applying the appropriate boundary conditions at $x' = 0$ (depletion edge) and $x' \rightarrow \infty$ (end contact).

'Excess' Hole Distribution

- **Boundary condition 1:** Thermal equilibrium prevails at the end contact.

$$\Delta p_n(x' \to \infty) = 0$$

- Note that $\exp(-x'/L_p) \to 0$ as $x' \to \infty$.
- However, $\exp(x'/L_p) \to \infty$ as $x' \to \infty$.
- Since thermal equilibrium prevails at the end contact, **constant C_2 must be zero** in order for Δp_n to be zero at the end contact.

'Excess' Hole Distribution

- **Boundary condition 2:**

$$\Delta p_{\mathrm{n}}\left(x'=0\right)=p_{\mathrm{n}}-p_{\mathrm{no}}$$
$$=p_{\mathrm{no}}\left(e^{qV_{\mathrm{a}}/(kT)}-1\right)$$

- Therefore,

$$C_1=p_{\mathrm{no}}\left(e^{qV_{\mathrm{a}}/(kT)}-1\right)$$

‘Excess’ Hole Distribution

$$\Delta p_{\mathrm{n}}(x') = p_{\mathrm{no}}\left(e^{qV_{\mathrm{a}}/(kT)} - 1\right) \cdot e^{-x'/L_{\mathrm{p}}}$$

- In a long-base diode, the ‘excess’ holes in the quasi-neutral n region decays exponentially with distance.
- The constant $L_{\mathrm{p}} = (D_{\mathrm{p}}\tau_{\mathrm{p}})^{1/2}$ has a unit of distance and is commonly referred to as the **minority carrier (hole) diffusion length**.

‘Excess’ Electron Distribution

- Through a similar approach, the ‘excess’ electron distribution in the quasi-neutral p region can be obtained:

$$\Delta n_{\mathrm{p}}\left(x''\right) = n_{\mathrm{po}}\left(e^{qV_{\mathrm{a}}/(kT)} - 1\right) \cdot e^{-x''/L_{\mathrm{n}}}$$

Current Components – Forward Bias

- When a p-n junction is forward biased, electrons (holes) are injected from the n (p) to the p (n) region.
- Owing to a difference in the minority carrier concentration at the depletion edge and the end contact, the excess holes and electrons diffuse towards the end contacts after they have been injected into the respective n and p regions.
- The diffusion process gives rise to the minority carrier current components, as shown by the solid lines.

Current Components – Forward Bias

- Minority-carrier current components:

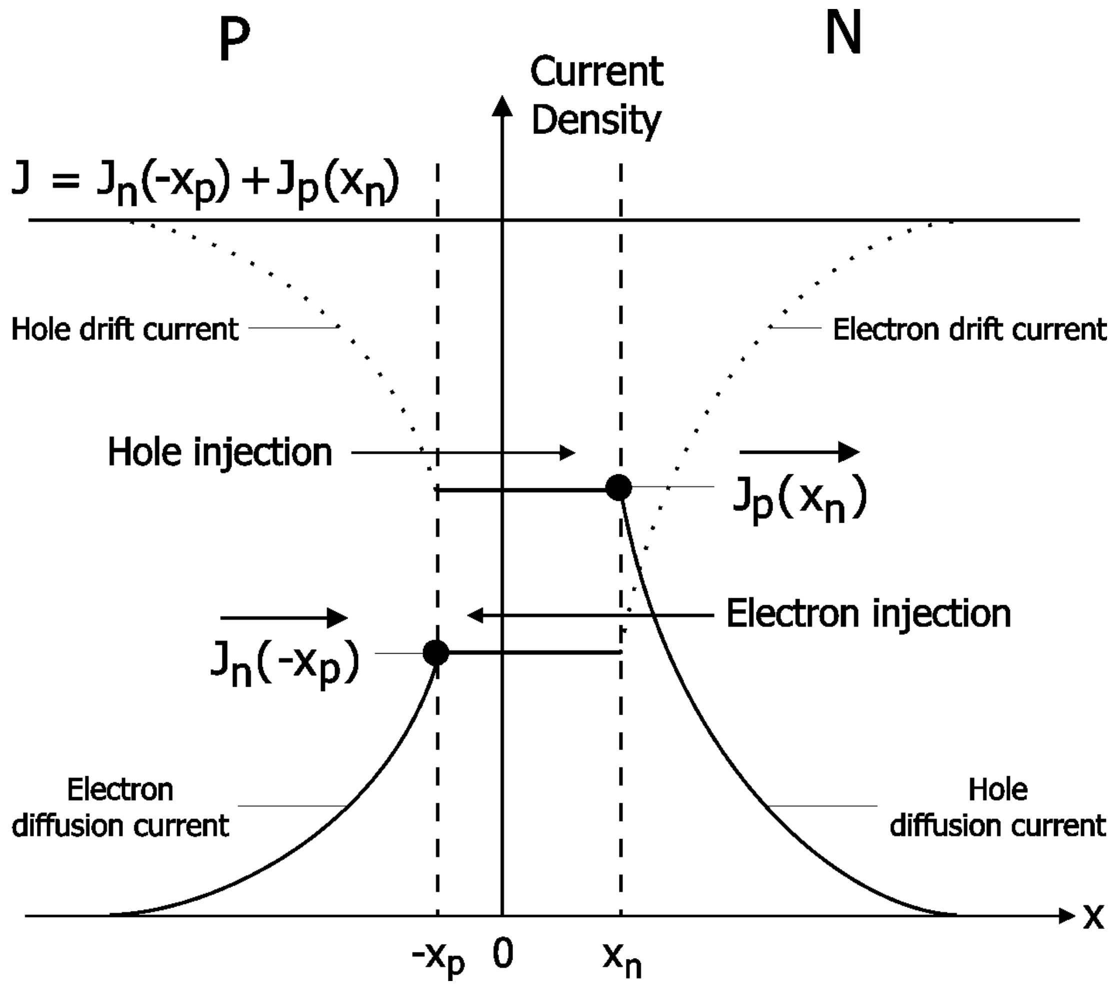
$$J_p(x') = qD_p \left| \frac{\partial p_n}{\partial x'} \right| = q \cdot \frac{D_p}{L_p} \cdot p_{no}\left(e^{qV_a/kT} - 1\right) \cdot e^{-x'/L_p}$$

$$J_n(x'') = qD_n \left| \frac{\partial n_p}{\partial x''} \right| = q \cdot \frac{D_n}{L_n} \cdot n_{po}\left(e^{qV_a/kT} - 1\right) \cdot e^{-x''/L_n}$$

  - The electron or hole diffusion current is not a constant, but varies as a function of distance from the edge of the depletion region.
  - This is because as the minority carriers diffuse along the length of the quasi-neutral region, they recombine with the majority carriers.

$$\partial^2 p_n/\partial x'^2 = \partial^2 \Delta p_n/\partial x'^2; \partial^2 n_p/\partial x''^2 = \partial^2 \Delta n_p/\partial x''^2$$

Current Components – Forward Bias



Current Components – Forward Bias

- Since there is no recombination in the space charge region, the minority-carrier currents flowing through it are constant.
- The individual minority-carrier current flow is completed by the majority carrier component, which supplies the charge injection into the respective n and p regions (dotted lines).
- The sum of the majority- and minority-carrier current components at any point gives the total current, which is independent of position and time under steady state condition.

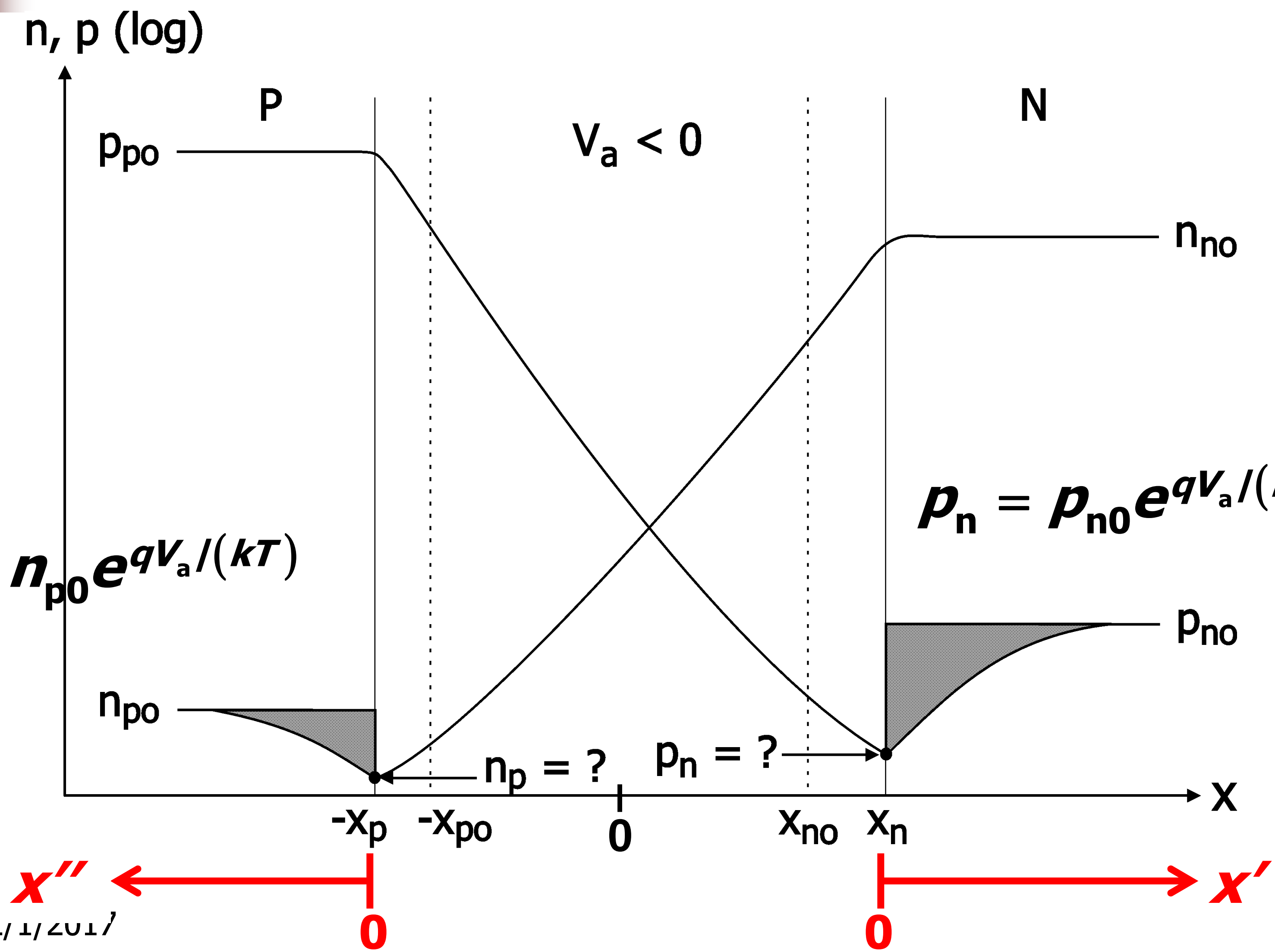
Current Components – Forward Bias

- Total diode current:

$$J = J_p(x' = 0) + J_n(x'' = 0)$$
$$= q\left(\frac{D_p p_{no}}{L_p} + \frac{D_n n_{po}}{L_n}\right)\left(e^{qV_a/(kT)} - 1\right)$$
$$= J_0\left(e^{qV_a/(kT)} - 1\right)$$

- This is the **ideal diode equation**.

Carrier Distributions – Reverse Bias



$$p_n = p_{n0} e^{qV_a/(kT)}$$

$$n_p = n_{p0} e^{qV_a/(kT)}$$

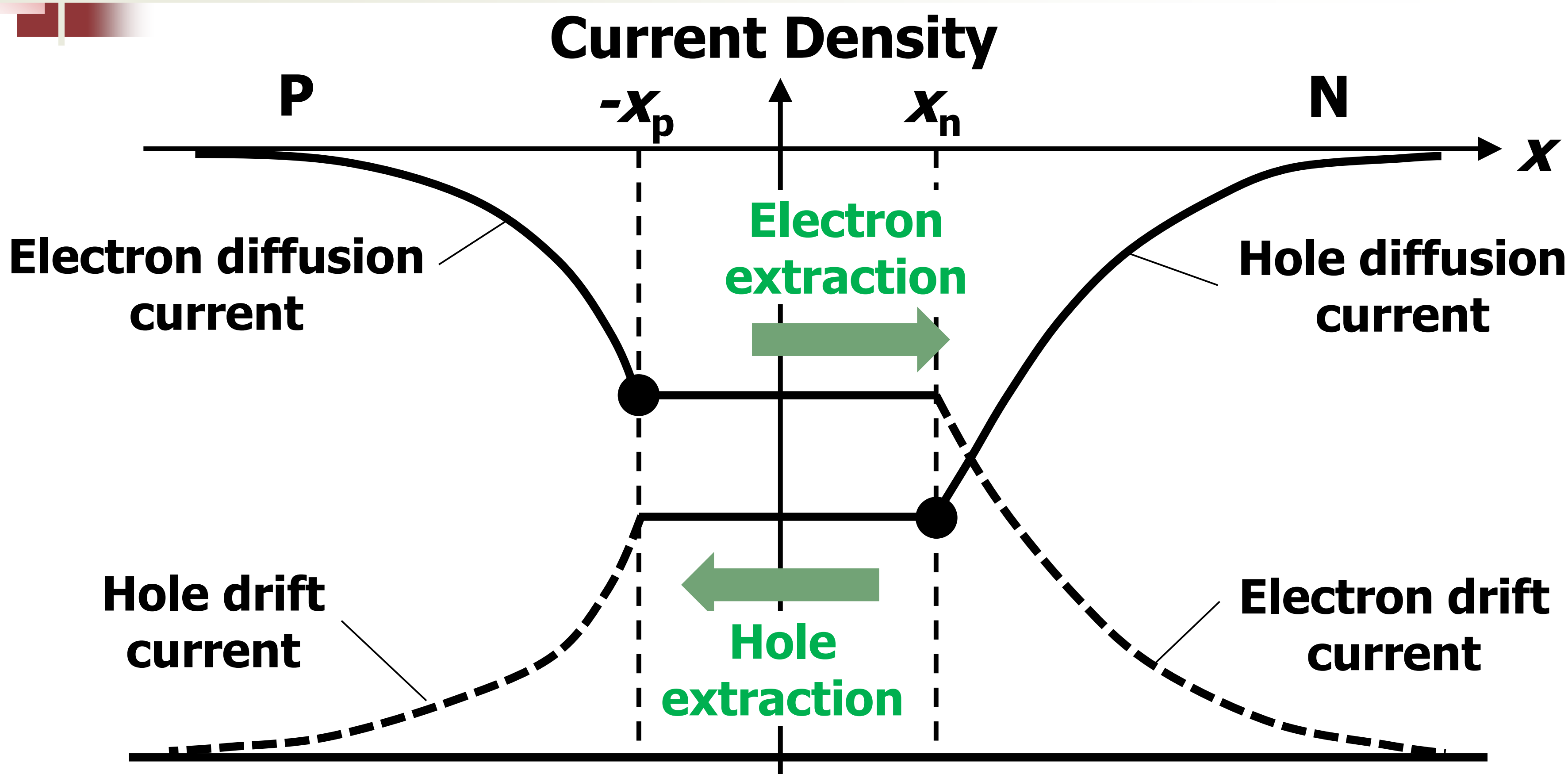
Current Components – Reverse Bias

- Under reverse bias, minority carriers at the edge of the depletion region are extracted across the depletion region.
- This causes a difference in the minority carrier concentration at the depletion edge and the rest of the quasi-neutral region.
- Hence, minority carriers, which are within one diffusion length from the depletion edge, diffuse to the depletion region.
- This diffusion process gives rise to a **small** minority-carrier current component.

Current Components – Reverse Bias

- Upon reaching the depletion region, the minority carriers are extracted across to the opposite region, where they become the majority carriers.
- Since there is no generation within the space charge region, the minority carrier current would be a constant throughout the region.
- The current flow is completed by the drift of the majority carriers towards the end contact of the opposite region.
- Except for a reversal in current direction, the same current distribution plot on page 16 describes the current flow of a reverse-biased pn junction.

Current Components – Reverse Bias



$J = J_n(-x_p) + J_p(x_n)$

Ideal Diode Equation

$$J = J_0\left(e^{qV_a/(kT)} - 1\right)$$

$$J_o = q\left(\frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n}\right) = qn_i^2\left(\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A}\right)$$

- The constant J_o is commonly known as the **reverse saturation current density**.
- J_o depends on the dopant concentrations, carrier diffusivities and lifetimes.
- It also depends on the temperature and bandgap energy of the semiconductor material through the intrinsic carrier concentration, n_i. Recall: $n_i = \sqrt{N_c N_v} \cdot e^{-E_g/2kT}$

Ideal Diode Equation

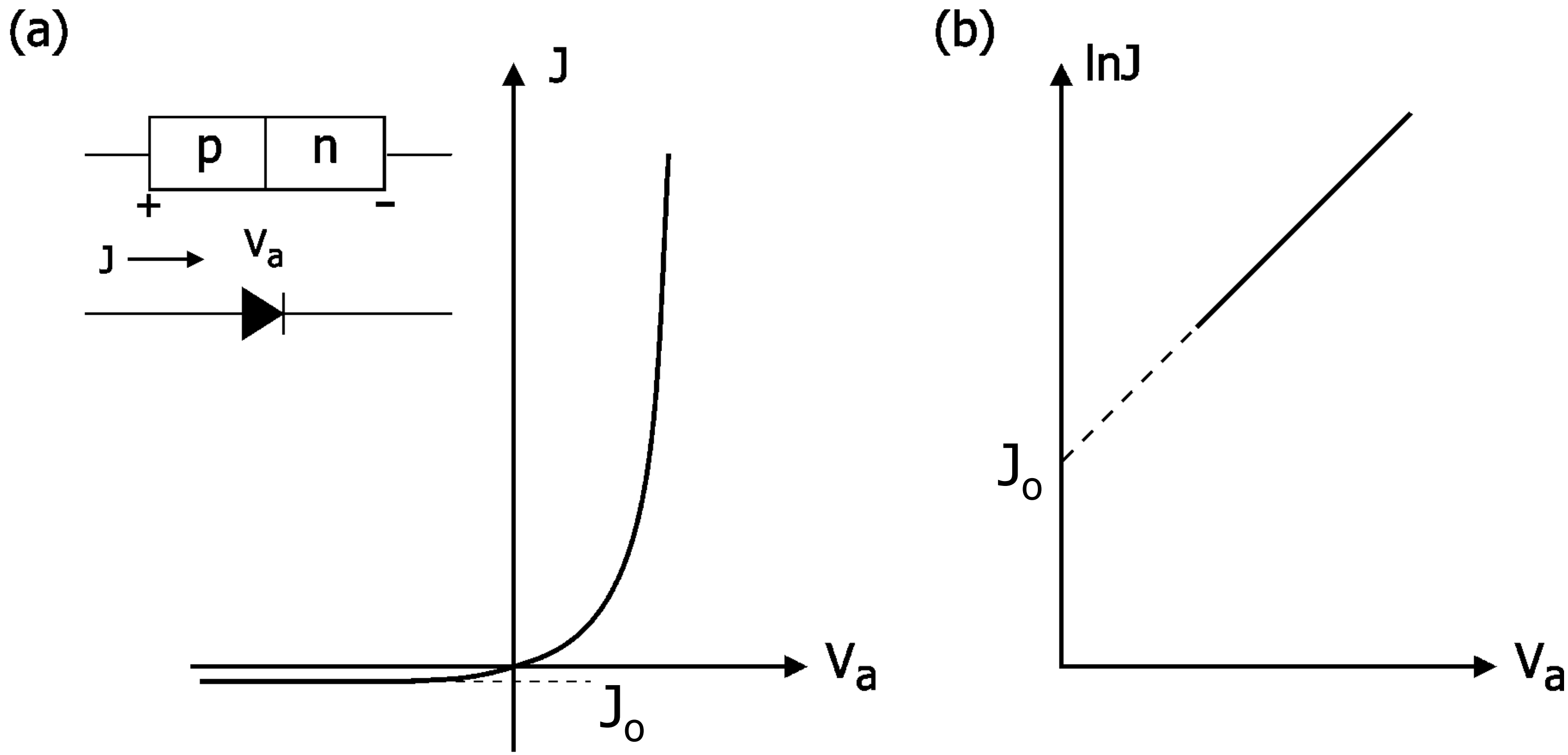
- **Forward bias:**
  - For an applied forward voltage of a few kT/q or larger, the exponential term in the bracket is much greater than 1. The ideal diode equation becomes:

$$J \approx J_0 e^{qV_a/(kT)} \Rightarrow \ln J \approx \ln J_0 + \frac{qV_a}{kT}$$

  - Plotting $\ln J$ versus V_a should therefore yield a straight line with intercept $\ln J_0$ and slope $q/(kT)$.
- **Reverse bias:**
  - For an applied reverse voltage of a few kT/q or larger (note that V_a is negative in this case), the exponential term is much smaller than 1. Hence,

$$J \approx -J_0$$

Ideal Diode Characteristics



(a) Current-voltage (I-V) characteristics of an ideal pn junction diode. (b) Forward I-V characteristics plotted on a log-linear scale.

Example 1

- To determine the reverse saturation current density at $T = 300$ K, of an ideal silicon p-n junction with the following parameters:
  - $N_A = N_D = 10^{16}$ cm^{-3}
  - $D_p = 10$ cm^2/s
  - $D_n = 25$ cm^2/s
  - $\tau_p = \tau_n = 0.5$ μs
- Assume $n_i = 1.5 \times 10^{10}$ cm^{-3} at $T = 300$ K.

Example 1

- The reverse saturation current density is given by

$$J_o = qn_i^2\left(\frac{D_p}{L_pN_D} + \frac{D_n}{L_nN_A}\right)$$

$$D\tau = L^2$$

$$\frac{D}{L} = \frac{L}{\tau} = \frac{\sqrt{D\tau}}{\tau} = \sqrt{\frac{D}{\tau}}$$

which could be re-written as

$$J_o = qn_i^2\left(\frac{1}{N_D}\cdot\sqrt{\frac{D_p}{\tau_p}} + \frac{1}{N_A}\cdot\sqrt{\frac{D_n}{\tau_n}}\right)$$

- Substituting the parameters gives $J_0 = 4.15\text{x}10^{-11}$ A/cm^2.
- Comment:
  - The reverse saturation current density is very small. If the cross-sectional area is 10^{-4} cm^2, for example, the current would be $4.15\text{x}10^{-15}$ A or 4.15 fA.

Example 2

- Consider a silicon p-n junction diode at $T = 300$ K. Design the diode such that $J_n = 20$ A/cm^2 and $J_p = 0.2$ A/cm^2 at $V_a = 0.65$. The following parameters apply:
  - $D_p = 10$ cm^2/s
  - $D_n = 25$ cm^2/s
  - $\tau_p = \tau_n = 0.5\ \mu$s
- Assume $n_i = 1.5\text{x}10^{10}$ cm^{-3} at $T = 300$ K.

Example 2

- The electron diffusion current density:

$$J_n = q \cdot \frac{D_n}{L_n} \cdot \frac{n_i^2}{N_A}\left(e^{qV_a/(kT)} - 1\right) = q \cdot \sqrt{\frac{D_n}{\tau_n}} \cdot \frac{n_i^2}{N_A}\left(e^{qV_a/(kT)} - 1\right)$$

- Substituting in the numbers

$$20 = \left(1.6 \times 10^{-19}\right) \cdot \sqrt{\frac{25}{5 \times 10^{-7}}} \cdot \frac{\left(1.5 \times 10^{10}\right)^2}{N_A}\left(e^{0.65/0.0259} - 1\right)$$

yields $N_A = 1.01\text{x}10^{15}\ \text{cm}^{-3}$.

- To obtain N_D, the same method can be used. However, it may be useful to note that

$$J_n \propto \frac{1}{N_A} \cdot \sqrt{\frac{D_n}{\tau_n}}, \quad J_p \propto \frac{1}{N_D} \cdot \sqrt{\frac{D_p}{\tau_p}}$$

Example 2

- Taking the ratio of the two current densities

$$\frac{J_n}{J_p} = \frac{N_D}{N_A} \cdot \sqrt{\frac{D_n}{\tau_n}} \cdot \sqrt{\frac{\tau_p}{D_p}}$$

yields

$$N_D = N_A \cdot \frac{J_n}{J_p} \cdot \sqrt{\frac{D_p}{D_n}}$$

$$= \left(1.01 \times 10^{15}\right)\left(\frac{20}{0.2}\right)\sqrt{\frac{10}{25}}$$

$$= 6.39 \times 10^{16}\ \mathrm{cm}^{-3}$$

Example 2

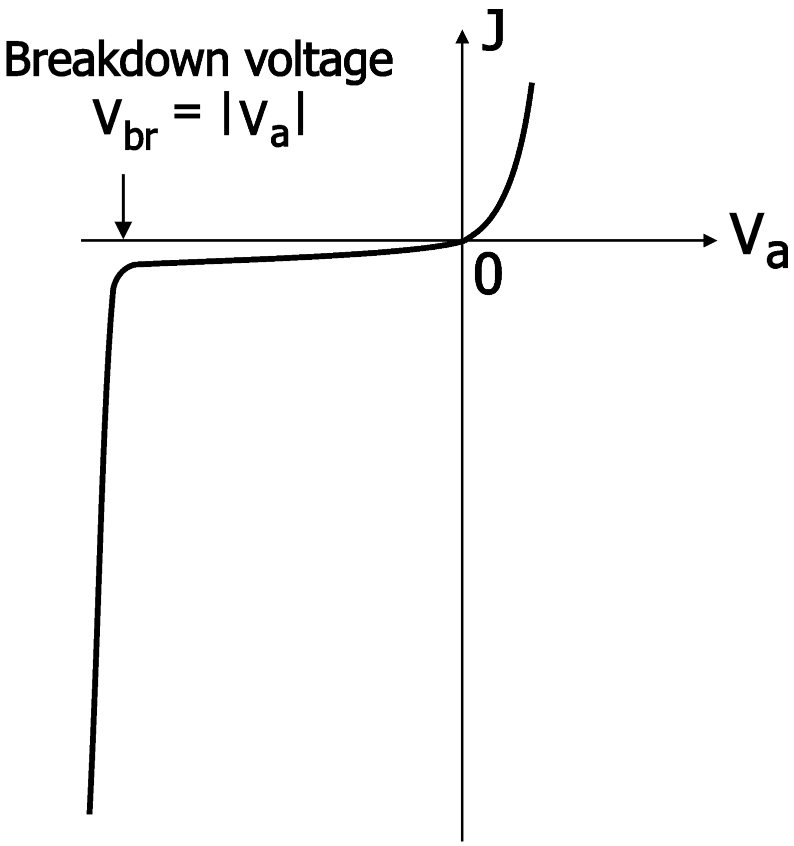
- Comments:
  - The magnitude of the electron and hole current densities through a diode can be varied by changing the doping concentrations in the device.
  - In general, for a diode with one region more heavily doped than the other, the current would be dominated by the injection of carriers from the heavily doped region to the lowly doped region. In the above example, $N_D > N_A$, and hence the current is mainly made up of electron diffusion current in the lowly doped p region.

EE2003 Semiconductor Fundamentals

Reverse Breakdown of the P-N Junction

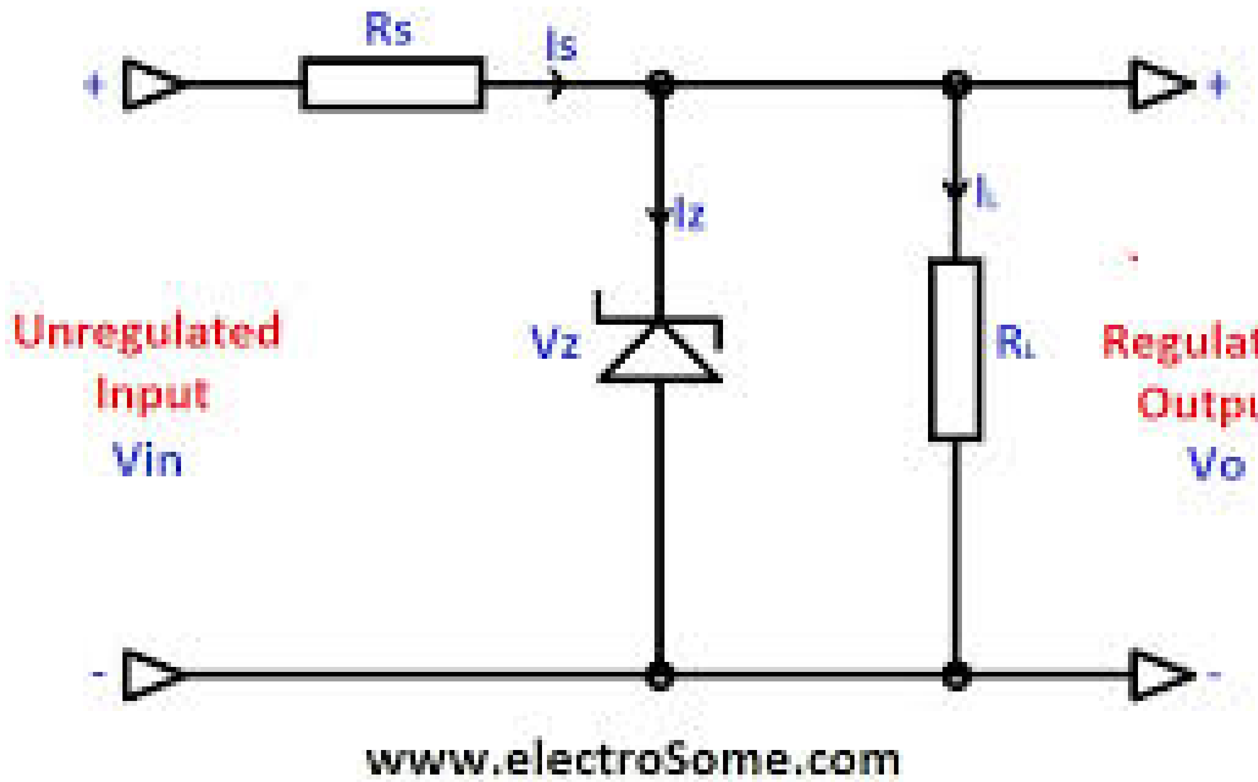
Reverse Breakdown

- The reverse leakage current of a practical diode remains small only for a limited range of applied voltage.
- Beyond a particular critical voltage, the reverse current increases rapidly as shown. The voltage at this point is known as the **breakdown voltage**.
- Mechanisms::
  - **Zener effect**
  - **Avalanche effect**



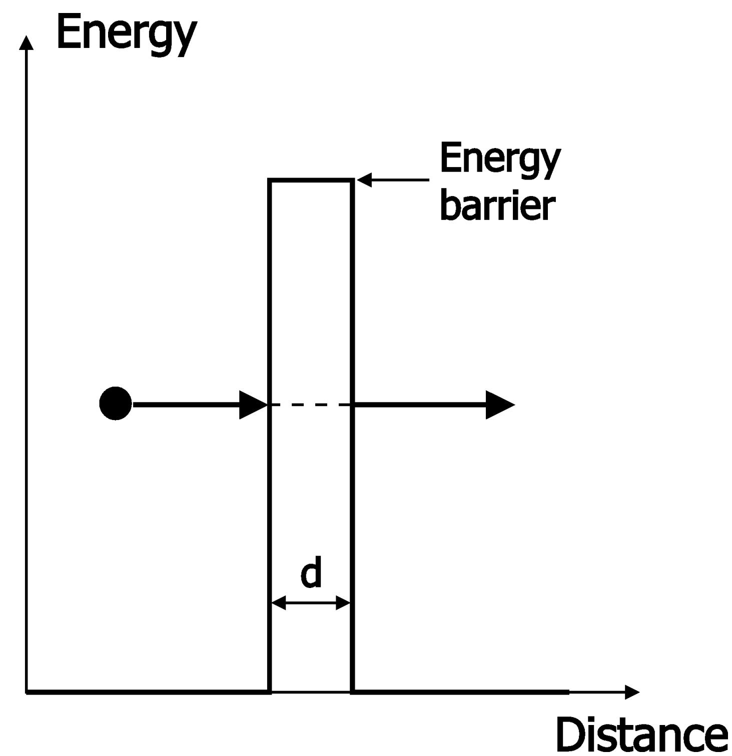
Voltage Regulator

An example of a practical application where a p-n junction diode is operated in the reverse breakdown mode.



Reverse Breakdown

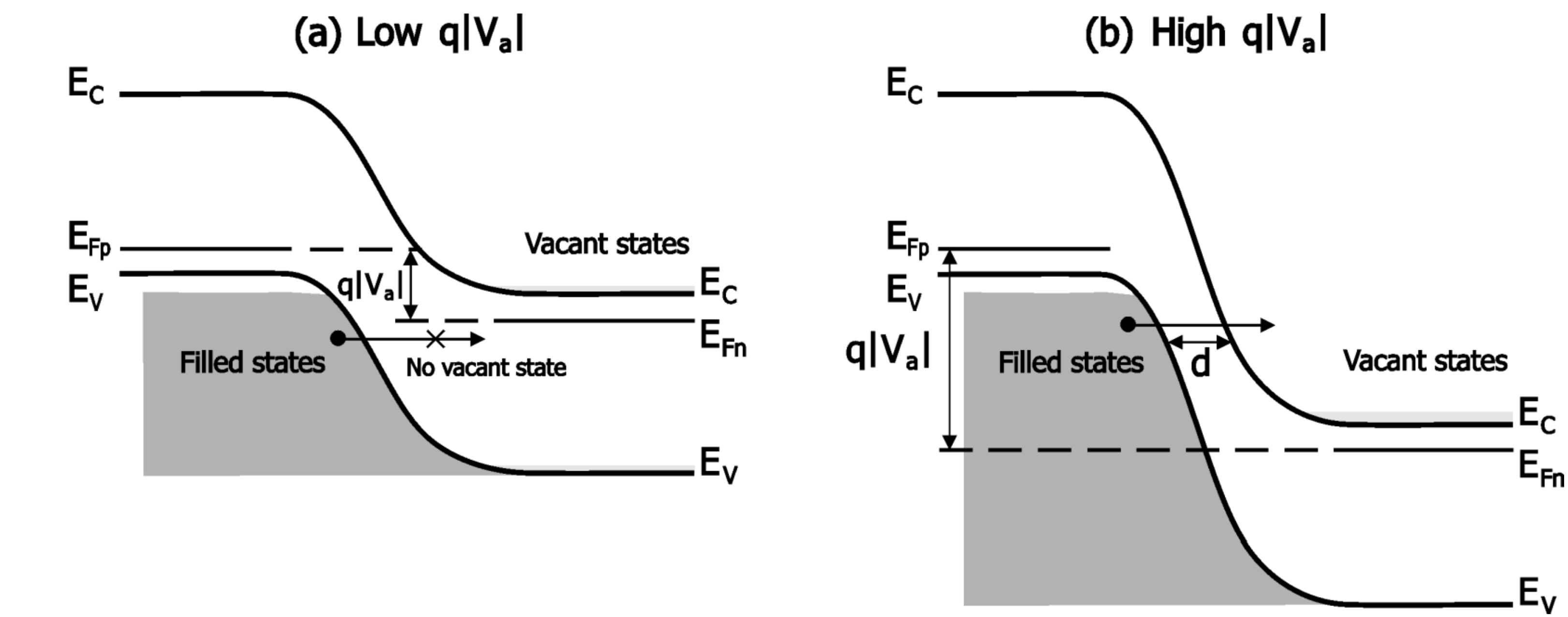
- **Zener breakdown** occurs in a highly doped p-n junction, via **quantum mechanical tunneling**.
  - Even when the potential energy of an electron is less than that of the barrier, there exists a non-zero probability that the electron will penetrate (**tunnel through**) the barrier, and continue its motion on the other side as shown.
  - An analogy is the passage of a bullet through a thin wall.



Reverse Breakdown

- The **tunneling probability** depends on:
  - Thickness of the barrier
    - The thicker the barrier, the smaller the probability.
  - Magnitude of the applied electric field
    - The electric field provides a force that tends to “pull” the electron across the barrier.
  - The presence of an **empty** energy state at the other side of the barrier.
    - After passing through the barrier, the electron must have a place to “reside” in.

Reverse Breakdown



Energy band diagrams of a p-n junction. In (a), because the applied reverse voltage is relatively low, no tunneling can occur since there is no vacant state to which the electron can tunnel to. In (b), the applied reverse voltage causes E_C in the n region to be lower than E_V in the p region. Electrons in the valence band of the p region can easily tunnel to the empty states in the conduction band of the n region.

Reverse Breakdown

- Zener breakdown occurs when the applied reverse bias is such that it lowers the conduction band edge of the n region **below** the valence band edge of the p region.
- When this happens, electrons in the valence band of the p region could tunnel to the empty states in the conduction band of the n region.
- At relatively low reverse bias, this tunneling cannot occur since there are no empty states to which the electrons could tunnel to.

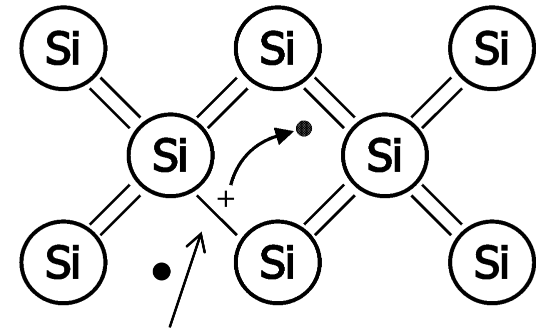
$$\left|E_{\mathrm{m}}\right| = \left|E\left(x = 0\right)\right| = \frac{qN_{\mathrm{A}}x_{\mathrm{p0}}}{\varepsilon_{\mathrm{r}}\varepsilon_0} = \frac{qN_D x_{\mathrm{n0}}}{\varepsilon_{\mathrm{r}}\varepsilon_0}$$

Reverse Breakdown

$$W_0 = \left[\frac{2\varepsilon_{\mathrm{r}}\varepsilon_0 V_{\mathrm{bi}}}{q}\left(\frac{1}{N_{\mathrm{A}}} + \frac{1}{N_{\mathrm{D}}}\right)\right]^{1/2}$$

- Heavily doped n and p regions:
  - Narrower depletion width
  - Higher electric field for a given applied reverse bias voltage
- The above factors make it easier for an electron to tunnel through the energy barrier (see diagram) at a given applied reverse bias voltage.
  - More tunneling $\Rightarrow$ higher reverse leakage current
- In practice, the desired breakdown voltage can be achieved through varying the doping concentrations of the n and p regions.

Reverse Breakdown

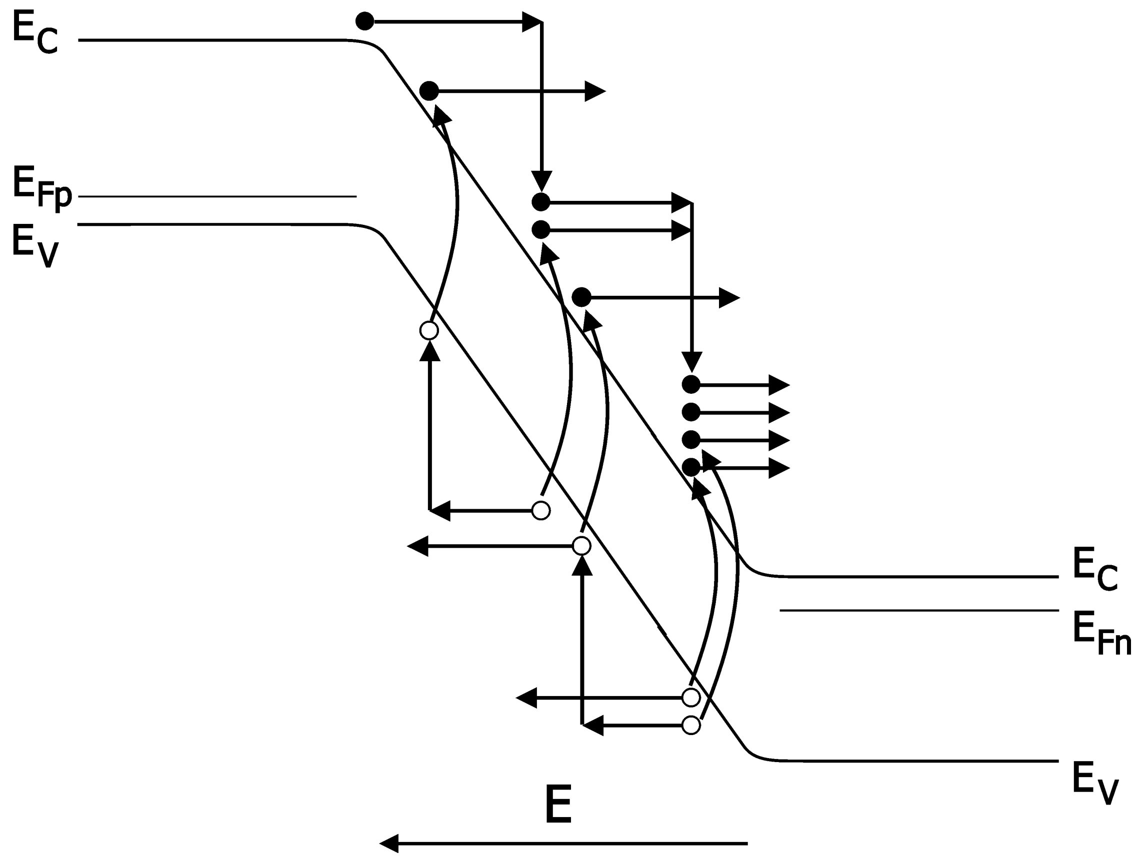


- **Avalanche breakdown**:
  - Mobile charges in the space charge region are accelerated to high speed by the high electric field.
  - When these energetic carriers collide with the lattice atoms, electrons which are involved in the covalent bonding could receive appropriate amount of energy from the colliding electrons/holes and be ‘freed’.
    - Bonds are broken $\Rightarrow$ electrons are promoted from the valence band to the conduction band (electron-hole pair generation).
    - This process through which electron-hole pairs are created via the collision of energetic carriers is known as **impact ionization**.

Reverse Breakdown

- **Avalanche breakdown:**
  - The newly generated carriers are in turn accelerated by the same field, leading to the generation of more 'free' carriers in the semiconductor.
  - A chain reaction (**positive feedback** path) is thus set-up, leading to a rapid increase in the diode reverse current $\Rightarrow$ diode breaks down.
  - **Snow ball analogy**

Reverse Breakdown



An illustration of the avalanche multiplication of carriers.

Reverse Breakdown

- **Breakdown voltage:**
  - The diode breaks down when the applied reverse voltage is such that the electric field established reaches or exceeds the **critical electric field**.
  - The critical electric field is at which significant impact ionization starts to take place.
  - For silicon, the critical electric field, $E_c \approx 3 \text{x} 10^5$ V/cm, and is dependent on the doping concentration.

Reverse Breakdown

- **Breakdown voltage:**
  - Maximum electric field:

$$\left|E_{max}\right| = \frac{qN_D}{\varepsilon_r \varepsilon_o} x_n = \frac{qN_A}{\varepsilon_r \varepsilon_o} x_p$$

  - x_n and x_p are the widths of the depletion regions in the n and p regions respectively:

$$x_n = \left[\frac{2\varepsilon_r \varepsilon_o \left(V_{bi} - V_a\right)}{q} \cdot \frac{N_A}{N_D\left(N_A + N_D\right)} \right]^{1/2}$$

$$x_p = \left[\frac{2\varepsilon_r \varepsilon_o \left(V_{bi} - V_a\right)}{q} \cdot \frac{N_D}{N_A\left(N_A + N_D\right)} \right]^{1/2}$$

Reverse Breakdown

- **Breakdown voltage:**
  - Substituting the expression for x_n or x_p into that of E_{max} gives:

$$E_{max}^{\ 2} = \frac{2q\left(V_{bi} - V_a\right)}{\varepsilon_r \varepsilon_o} \cdot \frac{N_A N_D}{\left(N_A + N_D\right)}$$

$$V_{bi} - V_a = \frac{\varepsilon_r \varepsilon_o \left(N_A + N_D\right)}{2qN_A N_D} E_{max}^{\ 2}$$

  - For reverse bias, $V_r = -V_a >> V_{bi}$, $V_{bi} - V_a \approx V_r$.

$$V_r \approx \frac{\varepsilon_r \varepsilon_o \left(N_A + N_D\right)}{2qN_A N_D} E_{max}^{\ 2}$$

Reverse Breakdown

- **Breakdown voltage:**
  - Breakdown occurs when the maximum field in the space charge region reaches the critical field. Equating E_{max} to E_c thus allows us to determine the reverse breakdown voltage V_{br} according to

$$V_{br} \approx \frac{\varepsilon_r \varepsilon_o (N_A + N_D)}{2qN_A N_D} E_c^{\,2}$$

Example 1

- An abrupt p-n junction has a uniform doping concentration of 10^{16} cm^{-3} for the n and p regions. Calculate the avalanche breakdown voltage. The critical field is 3.5×10^5 V/cm.

- The breakdown voltage is

$$V_{br} \approx \frac{\varepsilon_r \varepsilon_o \left(N_A + N_D\right)}{2qN_A N_D} E_c^{\ 2}$$

$$= \frac{(11.7)(8.85 \times 10^{-14})(10^{16} + 10^{16})}{2(1.6 \times 10^{-19})(10^{16})(10^{16})} (3.5 \times 10^5)^2$$

$$= 79.3 \text{ V}$$

Example 1

- Comments:
  - The avalanche breakdown voltage is typically in the range of several tens to hundreds of volts. In comparison, the Zener breakdown voltage is much lower (in the range of 5-20V).

Example 2

- Design an n^+p abrupt junction to achieve a breakdown voltage of 100 V. The critical field is 3.5×10^5 V/cm.
- For an n^+p junction, $N_D >> N_A$. The breakdown voltage is given by

$$V_{br} \approx \frac{\varepsilon_r \varepsilon_o (N_A + N_D)}{2qN_A N_D} E_c^{\ 2} \approx \frac{\varepsilon_r \varepsilon_o}{2qN_A} E_c^{\ 2}$$

- The required doping for the p region is therefore

$$N_A \approx \frac{\varepsilon_r \varepsilon_o}{2qV_r} E_c^{\ 2}$$

$$= \frac{(11.7)(8.85 \times 10^{-14})}{2(1.6 \times 10^{-19})(100)} (3.5 \times 10^5)^2$$

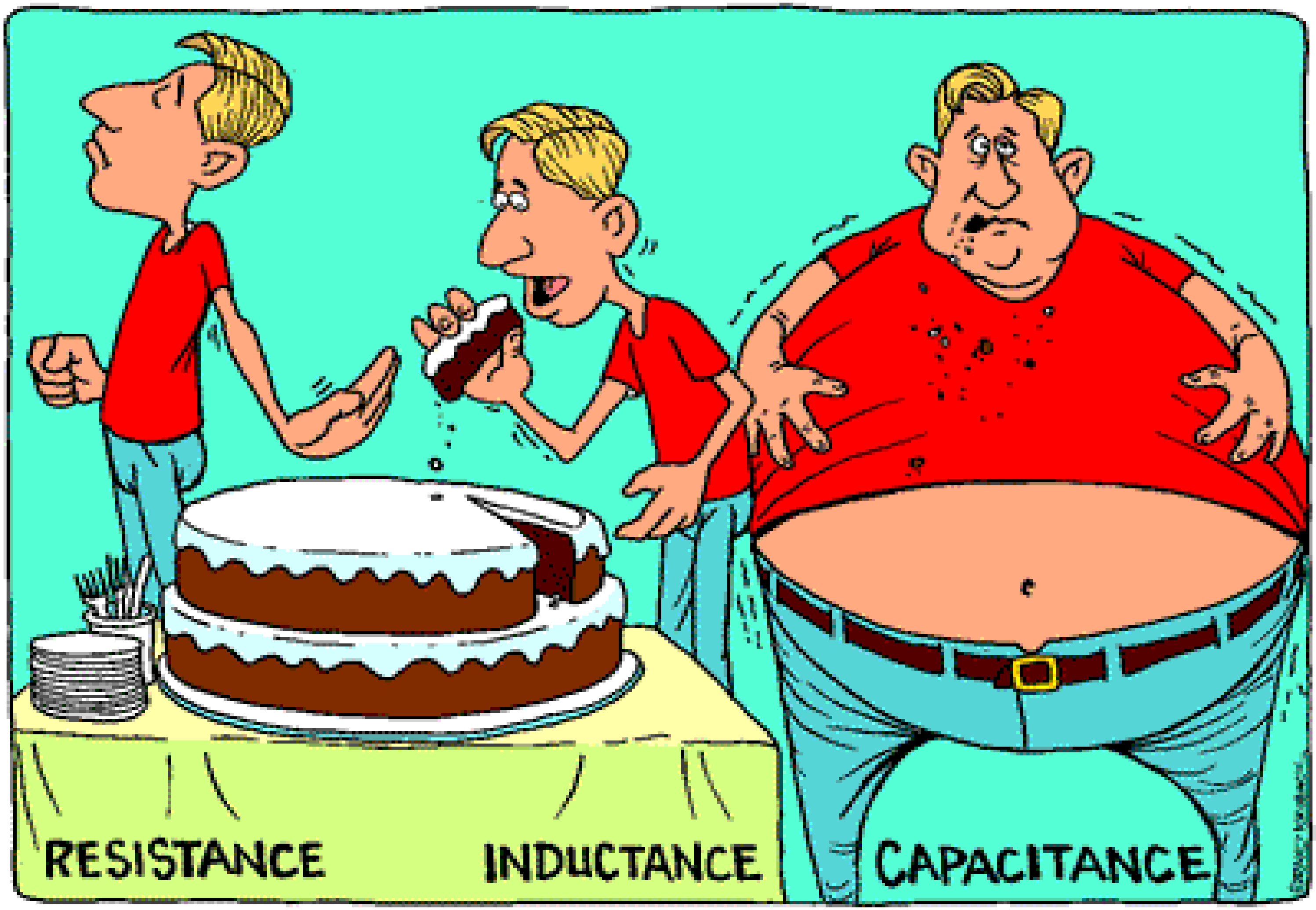
$$= 3.96 \times 10^{15}\ \text{cm}^{-3}$$

Example 2

- Comments:
  - The avalanche breakdown voltage of a one-sided pn junction is controlled by the doping concentration of the lowly doped side.
  - In this example, we see that the breakdown voltage is inversely proportional to the doping concentration of the lowly doped p region, i.e.

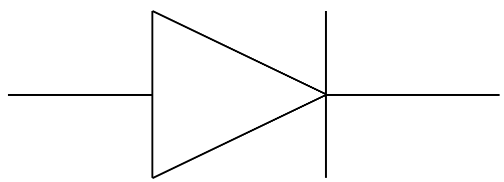
$$\boldsymbol{V_{\mathrm{br}} \propto \frac{1}{N_{\mathrm{A}}}}$$

Capacitance of the P-N Junction

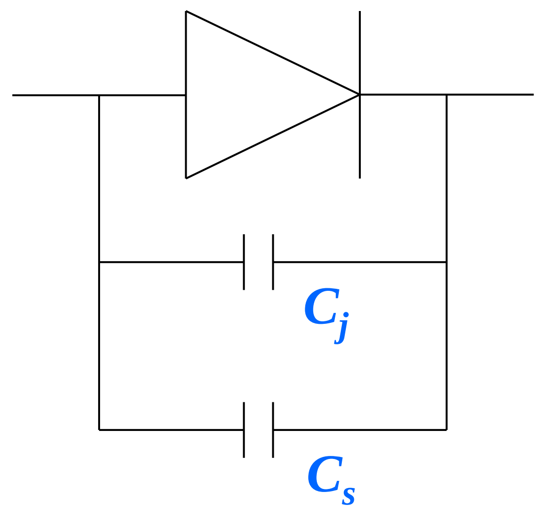


Capacitance of the P-N Junction

Two kinds of capacitances are associated with pn junctions. These are <u>junction capacitance</u> (or transition capacitance) and <u>charge storage capacitance</u> (or diffusion capacitance). These are modeled by two capacitors, C_j and C_s respectively, connected parallel to the p-n junction.



Ideal p-n junction



Practical p-n junction

Junction capacitance has physical characteristics similar to that of a parallel-plate capacitor and the expression that defines this capacitance is identical to that of a parallel-plate capacitor.

The storage capacitance also possesses the charge storing property of a capacitor but has no characteristics similar to those of a parallel-plate capacitor.

Capacitance is present in any device if a small change in the applied voltage ***dV*** across it results in a change in some charge ***dQ*** stored in it. By definition, the capacitance ***C*** is given by

$$C = \frac{dQ}{dV}$$

In the special case of a parallel plate capacitor, since ***Q*** is proportional to ***V***,

$$C = \frac{Q}{V}$$

$$\boldsymbol{Q \propto V \Rightarrow Q = kV}$$

$$\boldsymbol{C = \frac{dQ}{dV} = k = \frac{Q}{V}}$$

As the voltage $\boldsymbol{V}$ across a p-n junction varies, there are two stored charges in it, $\boldsymbol{Q_j}$ and $\boldsymbol{Q_s}$, that will be affected. Hence, there are two associated capacitances.

$$C_j = \frac{dQ_j}{dV} \quad , \quad C_s = \frac{dQ_s}{dV}$$

From circuit point of view, the voltage across the p-n junction cannot change instantaneously due to the presence of $\boldsymbol{C_j}$ and $\boldsymbol{C_s}$. This implies that there will be delay in circuits involving p-n junctions as the capacitors $\boldsymbol{C_j}$ and $\boldsymbol{C_s}$ will need time to be charged up or discharged.

Physically, this delay is related to the fact that <u>a finite time is required for current to flow and charges to be transported to or away from the p-n junction, so that the stored charge $\boldsymbol{Q_j}$ and $\boldsymbol{Q_s}$ can be built-up or reduced to their new value upon changing $\boldsymbol{V}$.</u>

Let's look at the source of $\boldsymbol{Q_j}$ and $\boldsymbol{Q_s}$ that give rise to $\boldsymbol{C_j}$ and $\boldsymbol{C_s}$.

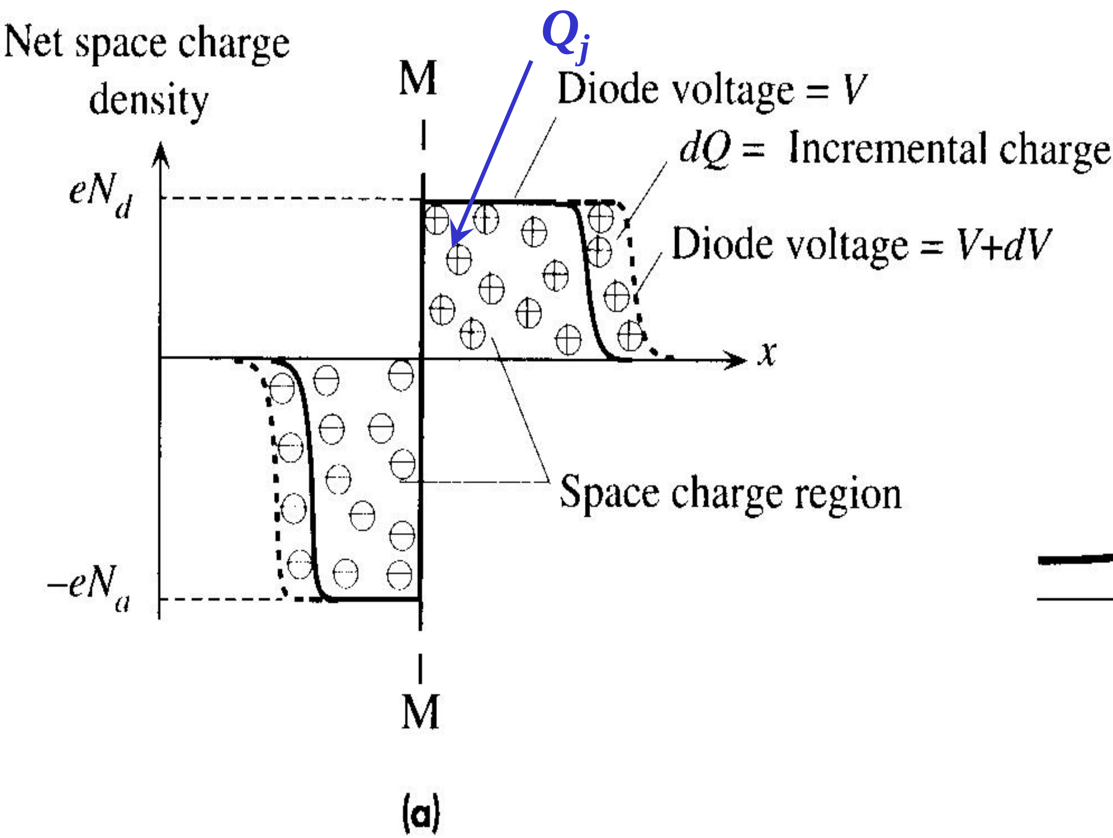
a) Junction Capacitance C_j

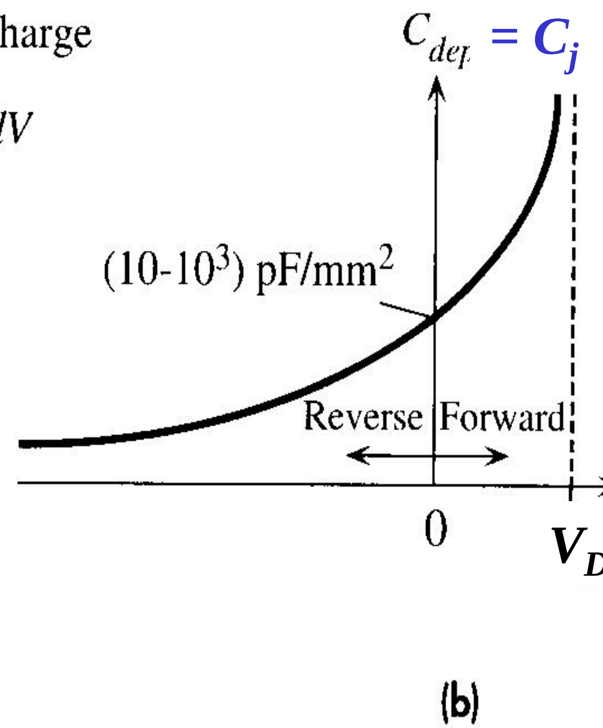
Upon changing the bias $\boldsymbol{V}$ across a p-n junction, the depletion width $\boldsymbol{W}$ and hence the amount of space charge in the depletion region will be affected.

The space charge $\boldsymbol{Q_j}$ in the depletion region $\boldsymbol{W}$ is given by

$$Q_j = |Q| = qAx_p N_a = qAx_n N_d$$

Under FB, $\boldsymbol{W}$ decreases and hence $\boldsymbol{Q_j}$ decreases. Under RB, $\boldsymbol{W}$ increases and hence $\boldsymbol{Q_j}$ increases. Thus, there is a capacitive effect associated with this change of charge $\boldsymbol{Q_j}$. Unlike the parallel plate capacitor, $\boldsymbol{Q_j}$ is not proportional to $\boldsymbol{V}$, viz., non-linear, and hence the above eqn. must be used to determine $\boldsymbol{C_j}$.





Note:
V_D here is the applied voltage called the diode voltage

The depletion region behaves like a capacitor.

(a) The charge in the depletion region depends on the applied voltage just as in a capacitor.
(b) The incremental capacitance of the depletion region increases with forward bias and decreases with reverse bias. Its value is typically in the range of picofarads per mm^2 of device area.

With applied bias $\boldsymbol{V}$ (+ve when FB and –ve when RB)

$$W = \left[\frac{2\varepsilon(V_o - V)}{q} \left(\frac{N_a + N_d}{N_a N_d} \right) \right]^{1/2}$$

Now, recall the expressions for x_n and x_p

$$x_n = \frac{WN_a}{N_a + N_d} \qquad \text{and} \qquad x_p = \frac{WN_d}{N_a + N_d} \qquad \left(\begin{array}{l} x_n N_d = x_p N_a \\ x_n + x_p = W \end{array} \right)$$

Thus, the space charge $\boldsymbol{Q_j}$ can now be written as

$$Q_j = qA \frac{N_a N_d}{N_a + N_d} W = qA \frac{N_a N_d}{N_a + N_d} \left[\frac{2\varepsilon(V_o - V)}{q} \left(\frac{N_a + N_d}{N_a N_d} \right) \right]^{1/2}$$

$$= A \left[2q\varepsilon(V_o - V) \frac{N_a N_d}{N_a + N_d} \right]^{1/2}$$

The junction capacitance C_j is thus given by

$$C_j = \left|\frac{dQ_j}{dV}\right| = \left|\frac{dQ_j}{d(V_o - V)}\right| = \frac{A}{2}\left[\frac{2q\varepsilon}{(V_o - V)} \quad \frac{N_a N_d}{N_a + N_d}\right]^{1/2}$$

Note that C_j is voltage dependent (**$C_j \propto (V_o - V)^{-1/2}$**).

Rearranging,

$$W = \left[\frac{2\varepsilon(V_o - V)}{q}\left(\frac{N_a + N_d}{N_a N_d}\right)\right]^{1/2}$$

$$C_j = \varepsilon A \left[\frac{q}{2\varepsilon(V_o - V)} \frac{N_a N_d}{N_a + N_d}\right]^{1/2} = \frac{\varepsilon A}{W}$$

The expression for $\boldsymbol{C_j}$ is an exact analogy with that of a parallel plate capacitor, with the depletion width $\boldsymbol{W}$ corresponding to the plate separation of the capacitor.

Physically as we change the bias ***V*** across the p-n junction, it takes time for the majority carriers to respond in order to expose a larger space charge region under RB, or a smaller space charge region under FB.

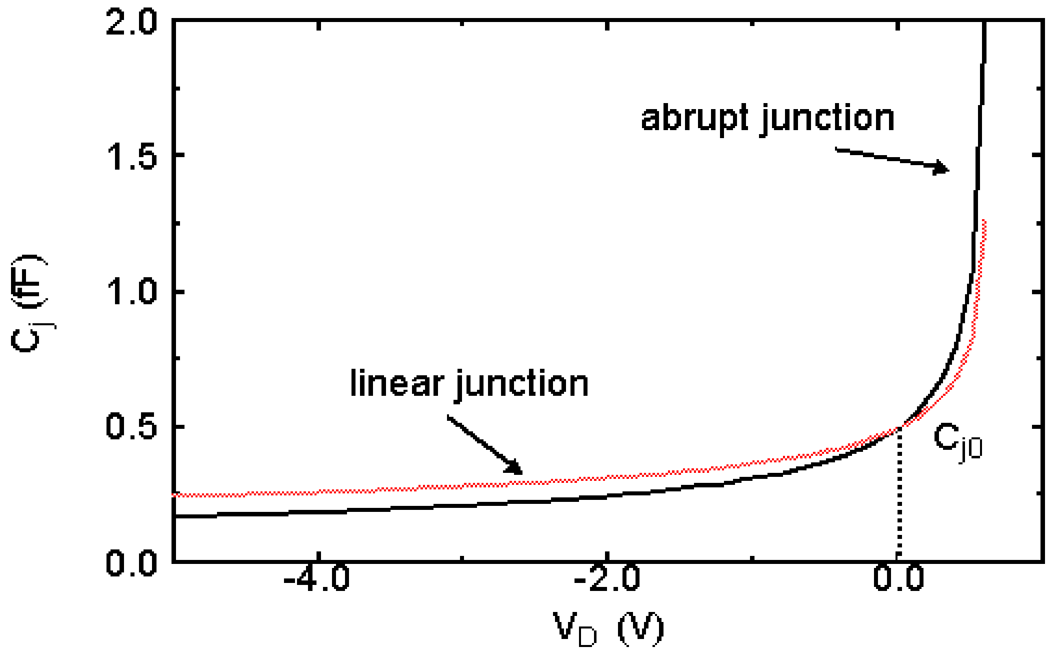
$\boldsymbol{C_j}$ is associated with this delay time for ***W*** to reach its steady state width. In other words, ***W*** cannot change instantaneously in response to a sudden change in the bias.

It is quite common to write the expression for C_j as a function of applied voltage. C_j is often written as

$$C_j(V) = C_{j0}\left[\frac{V_o}{V_o - V}\right]^{1/2} \qquad = \varepsilon A\left[\frac{q\boldsymbol{V_o}}{2\varepsilon \boldsymbol{V_o}(V_o - V)}\left(\frac{N_a N_d}{N_a + N_d}\right)\right]^{1/2}$$

where C_{j0} is the zero-bias junction capacitance, given by

$$C_{j0} = \varepsilon A\left[\frac{q}{2\varepsilon}\frac{1}{V_o}\frac{N_a N_d}{N_a + N_d}\right]^{1/2}$$



$$C_j(V) = C_{jo}\left[\frac{V_o}{V_o - V}\right]^m$$

m = 0.5: abrupt junction
m = 0.33: linear junction

Work Example

The doping densities of an abrupt-junction silicon P-N diode are $N_a = 10^{17}/\text{cm}^3$ and $N_d = 8\times10^{15}/\text{cm}^3$ and the area of the junction is 2×10^{-5} cm^2. Calculate the junction capacitance at (a) zero bias, (b) RB of 6 V and (c) FB of 0.7 V. Assume $n_i = 10^{10}$ cm^{-3}.

Solution: First, we have

$$C_j = \varepsilon A \left[\frac{q}{2\varepsilon(V_o - V)} \frac{N_a N_d}{N_a + N_d} \right]^{1/2} = \frac{\varepsilon A}{W}$$

Also, the equilibrium contact potential V_o is

$$V_o = \frac{kT}{q} \ln\left(\frac{N_a N_d}{n_i^2} \right)$$

Calculate V_o as

$$V_o = 0.0259 \ln[(10^{17} \times 8 \times 10^{15})/(10^{10})^2] = 0.77 \text{ V}$$

(a) At zero bias

$$C_j = 11.8 \times 8.85 \times 10^{-14} \times 2 \times 10^{-5} \times \left[\frac{1.6 \times 10^{-19}}{2 \times 11.8 \times 8.85 \times 10^{-14} \times (0.77 - 0)} \cdot \frac{10^{17} \cdot 8 \times 10^{15}}{10^{17} + 8 \times 10^{15}} \right]^{1/2}$$

$$C_j = 0.56 \text{ pF}$$

(b) With RB of 6 V,

$$C_j = 11.8\times 8.85\times 10^{-14}\times\ 2\times 10^{-5}\times \left[\frac{1.6\times 10^{-19}}{2\times 11.8\times 8.85\times 10^{-14}\,(0.77 + 6)}\cdot\frac{10^{17}\cdot 8\times 10^{15}}{10^{17} + 8\times 10^{15}}\right]^{1/2}$$

$$C_j = 0.188\ \text{pF}$$

(c) With FB of 0.7 V,

$$C_j = 11.8 \times 8.85 \times 10^{-14} \times \; 2 \times 10^{-5} \times$$

$$\left[\frac{1.6 \times 10^{-19}}{2 \times 11.8 \times 8.85 \times 10^{-14} \cdot (0.77 - 0.7)} \cdot \frac{10^{17} \cdot 8 \times 10^{15}}{10^{17} + 8 \times 10^{15}} \right]^{1/2}$$

$$C_j = 1.85 \text{ pF}$$

Note that junction capacitance in FB is much larger than that in RB as the width of the depletion region is much smaller in FB.

b) Charge Storage Capacitance C_s

The application of FB to a diode results in a reduction in the barrier height, a reduction in the depletion region width, and an injection of majority carriers across the depletion region into the opposite region, where they are stored as excess minority carriers. The **density of the excess minority carriers increases with increase in the FB.**

Note that with changing bias ***V***, there will be a change in the concentration of the excess minority carriers stored at both sides of the junction, near the depletion region edges. The concentration is increased upon increasing the FB and decreased upon decreasing the FB.

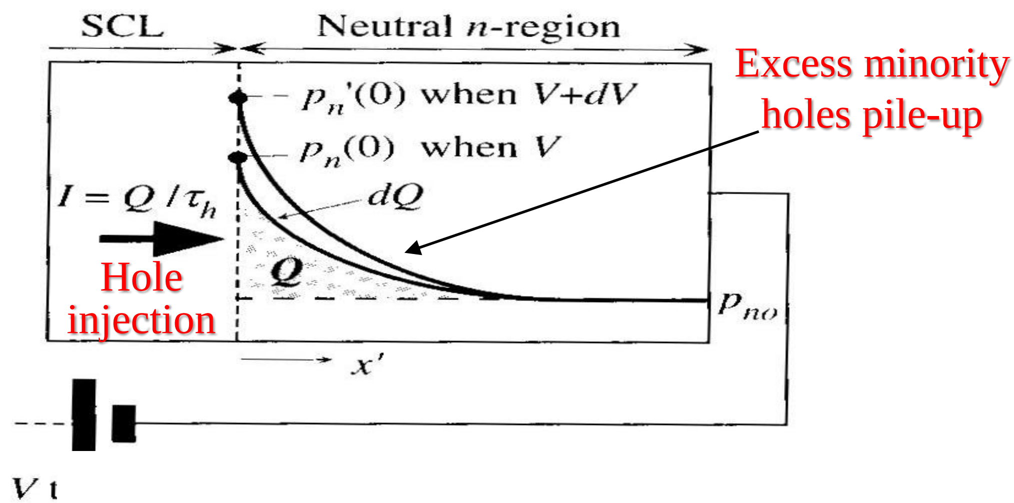


Figure 6.12 Consider the injection of holes into the *n*-side during forward bias. Storage or diffusion capacitance arises because when the diode voltage increases from *V* to *V* + *dV*, more minority carriers are injected and more minority carrier charge is stored in the *n*-region.

The increase in the minority carrier concentration does not take place instantaneously with a sudden change in FB.

A delay time is involved in the minority carrier concentration attaining a steady-state. This delay is due to capacitive effect as a result of stored minority carrier charges in the neutral n and p regions. This is the origin of storage capacitance.

The total excess stored charge for minority holes at n-side $\boldsymbol{Q_p}$ is

$$Q_p = qA\int_0^\infty \Delta p_n\,(x')\,dx' = qA\Delta p_n\left(x'=0\right)\int_0^\infty e^{-x'/L_p}\,dx'$$

$$= qAL_p\Delta p_n\left(x'=0\right) \qquad \left(\Delta p_n(x')=\Delta p_n\left(x'=0\right)e^{-x'/L_p}=p_{n0}(e^{qV/kT}-1)\,e^{-x'/L_p}\right)$$

The total excess stored charge for minority electron at p-side $\boldsymbol{Q_n}$ is

$$Q_n = -qA\int_0^\infty \Delta n_p\,(x'')\,dx_p = -qA\Delta n_p\left(x''=0\right)\int_0^\infty e^{-x''/L_p}\,dx''$$

$$= -qAL_n\Delta n_p\left(x''=0\right) \qquad \left(\Delta n_p(x'')=\Delta n_p\left(x''=0\right)e^{-x''/L_n}=n_{p0}(e^{qV/kT}-1)\,e^{-x''/L_n}\right)$$

Here, $\boldsymbol{\Delta p_n(x')}$ and $\boldsymbol{\Delta n_p(x'')}$ are the steady state profiles of the minority holes and electrons respectively, given by

$$\Delta p\,(x') \;=\; \Delta p_n\left(x'=0\right)e^{-x'/L_p} \;=\; p_{n0}(\,e^{qV/kT}-1\,)\,e^{-x'/L_p}$$

$$\Delta n\,(x'') \;=\; \Delta n_p\left(x''=0\right)e^{-x''/L_n} \;=\; n_{p0}(\,e^{qV/kT}-1\,)\,e^{-x''/L_n}$$

Let $\boldsymbol{Q_s}$ denote the total excess charge stored associated with the <u>minority carriers</u>, viz., $\boldsymbol{Q_s = Q_p + Q_n}$.

Consider the case of a p^+n junction under FB. Recall that <u>the hole injection current will dominate.</u> This transforms to $\boldsymbol{\Delta p_n \gg \Delta n_p}$, and hence $\boldsymbol{Q_p \gg Q_n}$. (and $\boldsymbol{I_p(x' = 0) \gg I_n(x'' = 0)}$ generally).

Hence,

$$C_s = \frac{dQ_s}{dV} \approx \frac{dQ_p}{dV}$$

Recall $\Delta p_n\left(x'=0\right) = p_n(x'=0) - p_{n0} = p_{n0}\,(\,e^{qV/kT} - 1\,)$

$$\approx p_{n0}\, e^{qV/kT}$$

Thus, the storage charge capacitance C_s associated with this change of storage charge is

$$C_S = \frac{d\ Q_p}{d\ V} = \frac{q^2}{kT} A L_p p_{n0} e^{\ qV\ /\ KT} \approx \frac{q}{kT} Q_p$$

Note: $\boldsymbol{Q_p = qAL_p\Delta p_n\left(x'=0\right)}$

$$I_p(x'=0)_{diff} = qA\frac{D_p}{L_p}\Delta p_n(x'=0)e^{-x'/L_p} = qA\frac{D_p}{L_p}\Delta p_n(x'=0)$$

For p$^+$n junction hole injection current will dominate and hence the total current $\boldsymbol{I}$ is given by

$$I \approx \frac{q\ A\ D_p}{L_p}\Delta p_n(x'=0) = \frac{qAL_p\Delta p_n(x'=0)}{\tau_p} \approx \frac{Q_p}{\tau_p}$$

since $L_p = (D_p\tau_p)^{1/2}$

Thus, $\boldsymbol{Q_p \approx I\tau_p}$.

Hence, C_s can also be expressed in terms of $\boldsymbol{I}$ and $\boldsymbol{\tau_p}$ as

$$C_s = \frac{q}{kT}I\tau_p$$

It can be seen that C_s can be very large under FB due to the large I. Physically, this can be understood because the storage charge depends exponentially on the bias voltage V under FB, meaning that there will be a large change in the amount of storage charge (dQ_s) of the minority carriers with changing bias (dV), resulting in a large C_s. The C_s will impose a serious limitation for a FB biased p-n junction in high frequency applications.

The total capacitance across the pn junction is $C_T = C_j + C_s$.

During FB, C_s is dominant ($C_s >> C_j$) due to the large change in the amount of storage charge of the minority carriers, hence $C_T \approx C_s$.

During RB, C_s is not critical ($C_s << C_j$) since there is only very small change in the amount of storage charge of the minority carriers. In comparison, the junction capacitance C_j dominates under RB, viz., $C_T \approx C_j$.

The presence of these capacitances will limit the speed of operation of any circuits using p-n junction diodes.

Work Example

The doping densities of an abrupt-junction silicon PN diode are N_a = 10^{17}/cm^3 and $N_d = 8 \times 10^{15}$/cm^3 and the area of the junction is 2×10^{-5} cm^2. The lifetime of holes in the n region is 0.1 µs. Given the diffusion constant for holes in N is 16 cm^2/s, calculate, at room temperature, the storage capacitance at (a) FB of 0.6 V and (b) 0.65 V. Assume $n_i = 10^{10}$ cm^{-3}.

Solution: Recall

$$C_S = \frac{q^2}{kT} A L_p p_{n0}\, e^{qV/kT}$$

Equilibrium density of holes in the N region is

$p_{n0} = n_i^2/N_d = 10^{20}/8 \times 10^{15} = 1.25 \times 10^4$ /cm^3

Also, $L_p = (D_p \tau_p)^{1/2} = (16 \times 0.1 \times 10^{-6})^{1/2} = 1.26 \times 10^{-3}$ cm

(a) When the FB is 0.6 V, storage capacitance is $C_S = \frac{q^2}{kT} A L_p p_{n0} e^{qV/kT}$

$$C_S = \frac{(1.6\times10^{-19})^2}{0.026\times1.6\times10^{-19}} \cdot 2\times10^{-5} \times 1.26\times10^{-3} \times 1.25\times10^{4} \cdot e^{0.6/0.026}$$

$$C_s = 20.4 \text{ pF}$$

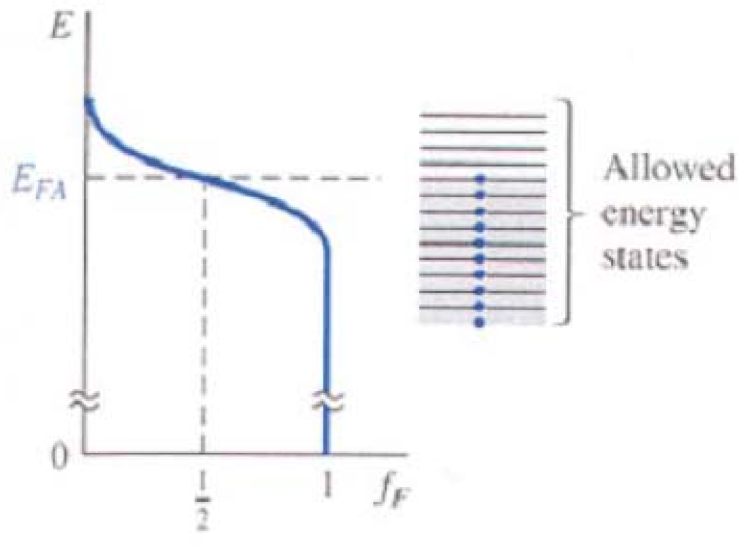
(b) Similarly, at FB of 0.65 V

$$C_S = \frac{(1.6\times10^{-19})^2}{0.026\times1.6\times10^{-19}} \cdot 2\times10^{-5} \times 1.26\times10^{-3} \times 1.25\times10^{4} \cdot e^{0.65/0.026}$$

$$C_s = 139.6 \text{ pF}$$

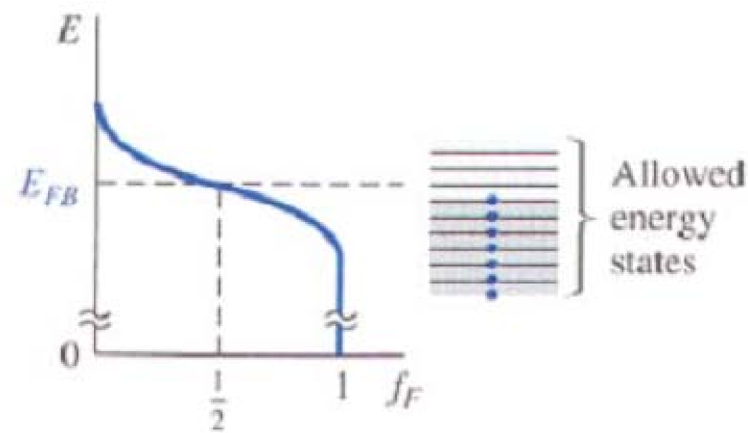
Note the large value of C_s. It is in fact much larger than C_j in FB (compare with the previous example for C_j value. It was 1.85 pF at a FB of 0.7 V).

EF must be equal when different systems are in contact and in thermodynamic equilibrium



(a)

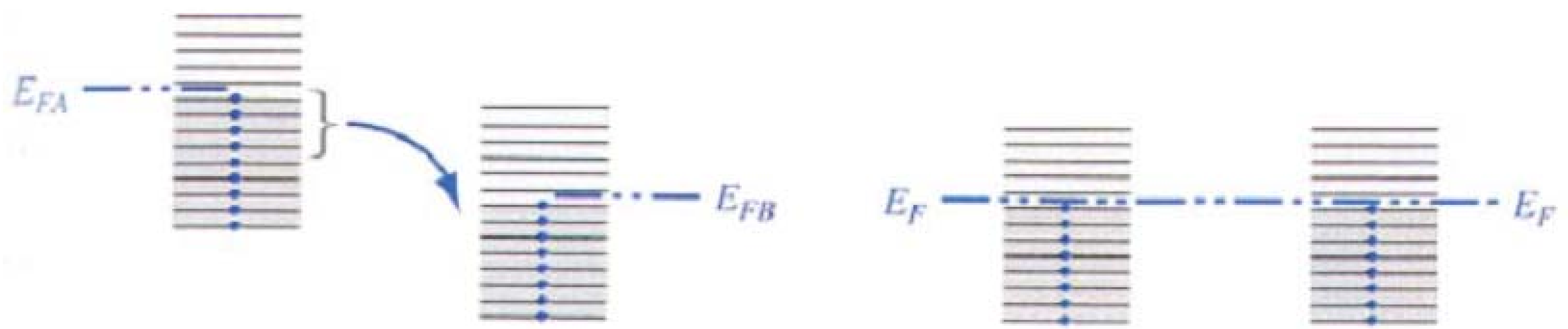
Consider a material A, with Fermi level E_{FA}. Bands below E_{FA} are full and above are empty.



(b)

material B with Fermi level E_{FB}.

EF must be equal when different systems are in contact and in thermodynamic equilibrium



- When A and B are brought in contact, electrons will flow from A into lower energy states of B, until thermal equilibrium is reached.

- Thermal equilibrium → E_F same in A & B