# ECE 5205 Spring 2014

# Secondary Effects in Si p-n Junctions

Marius Orlowski
Virginia Polytechnic Institute and State University

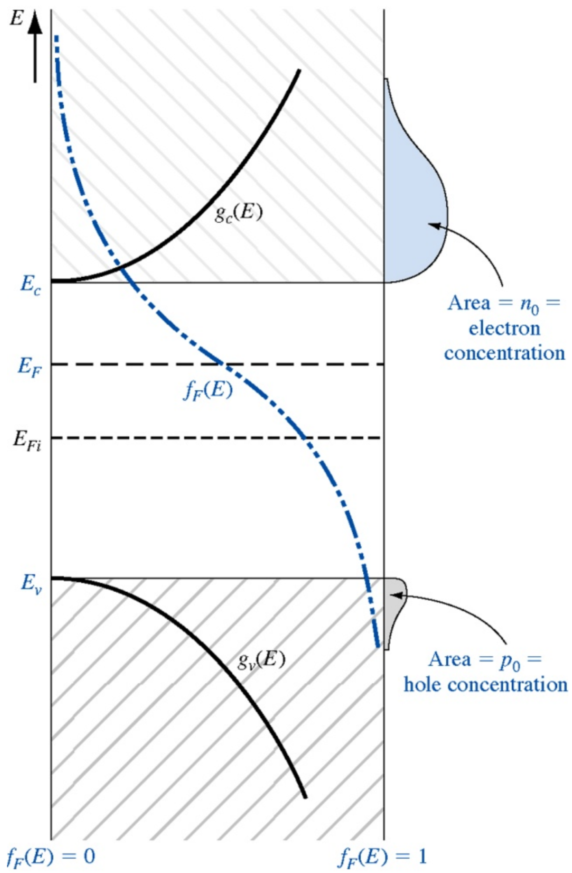
# pn Junctions

## Contents

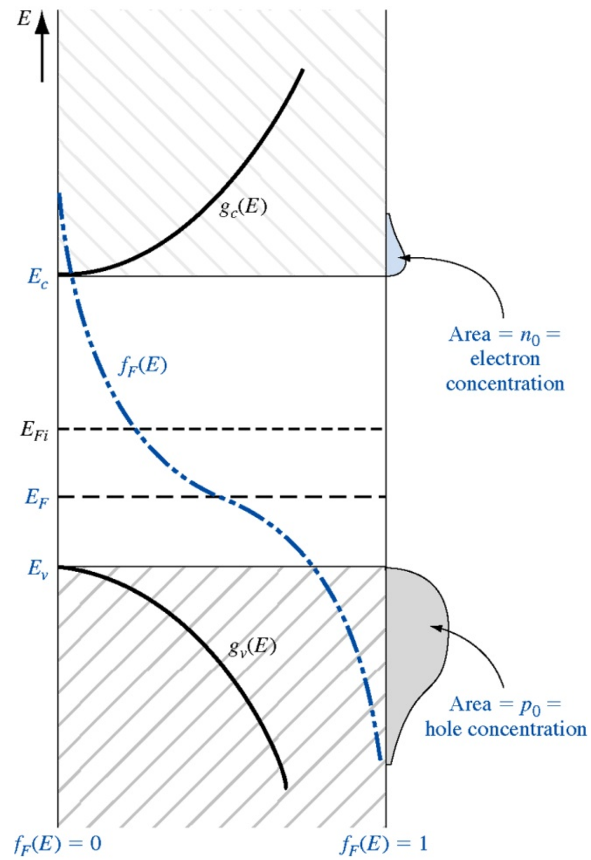
- pn junction and diode revisited
- pn junction diode - non-idealities
- junction breakdown
- impact ionziation and avalanche and
- band-to-band tunneling
- back-to-back pn junctions

# Extrinsic Semiconductor

Adding donor or acceptor impurity atoms will change the distribution of electrons and holes. Since the Fermi level is related to the distribution function of the carriers, the Fermi energy level will change as dopant atoms are added.




**Distribution of electron and hole concentrations when $E_F$ is above the midgap energy**.




**Distribution of electron and hole concentrations when $E_F$ is below the midgap energy**.

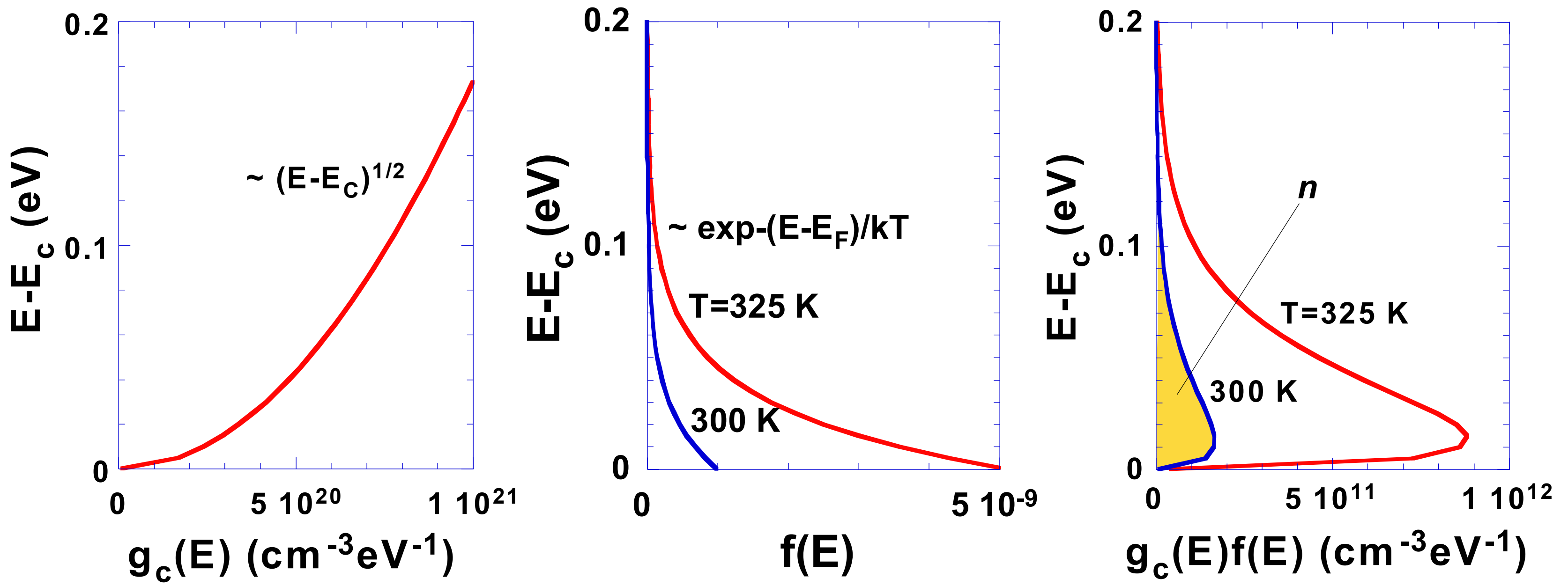
# Carrier Densities

## Electron Density

- The electron density depends on the density of states and on the probability of these states being occupied
- n is the area under the $(E - E_c)$ versus $g_c(E)f(E)$ curve

$$n = \int_{E_c}^{E_{top}} g_c(E) f(E) dE$$

Fermi Integral

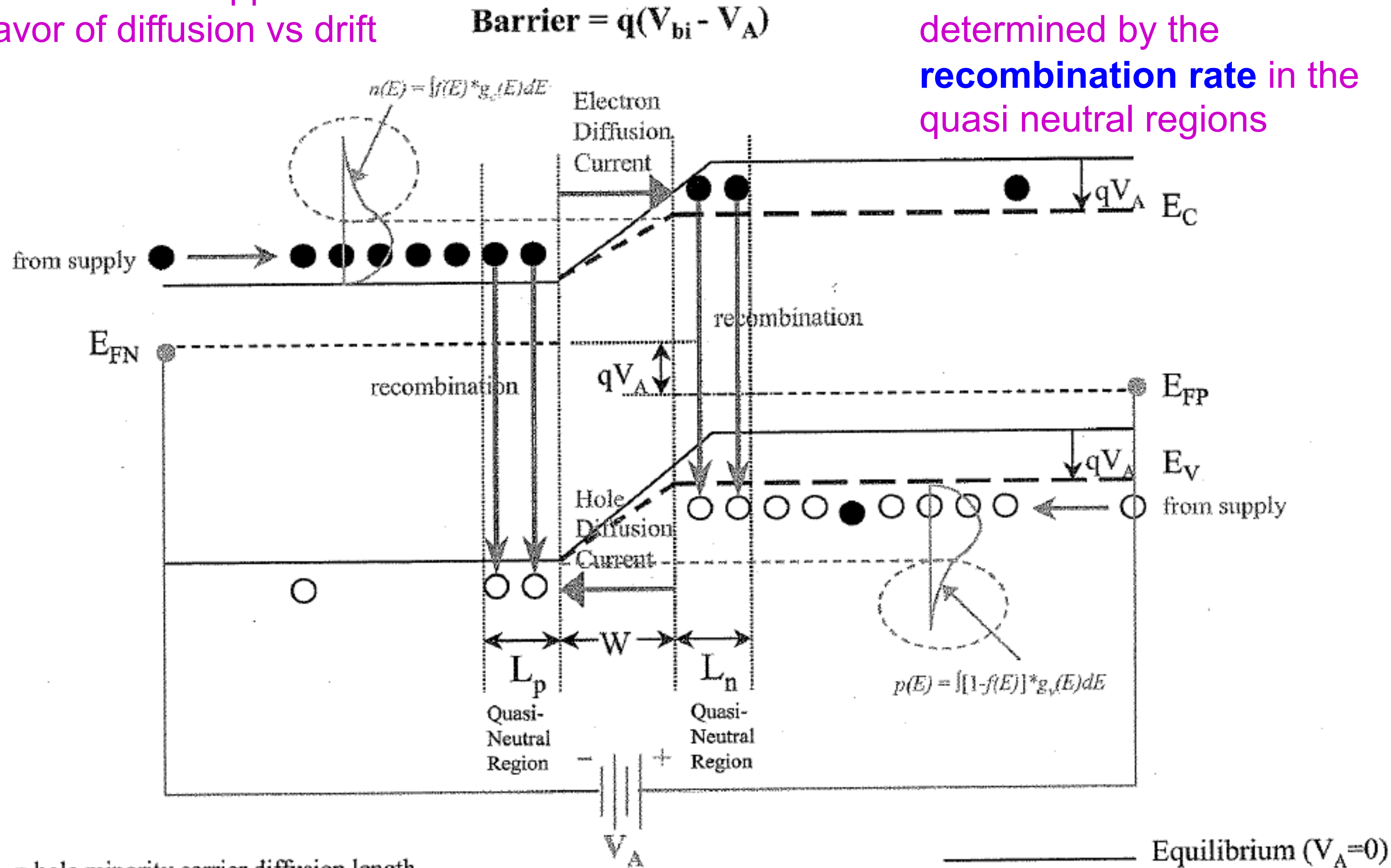



Density of states | probability of occupation | electron density

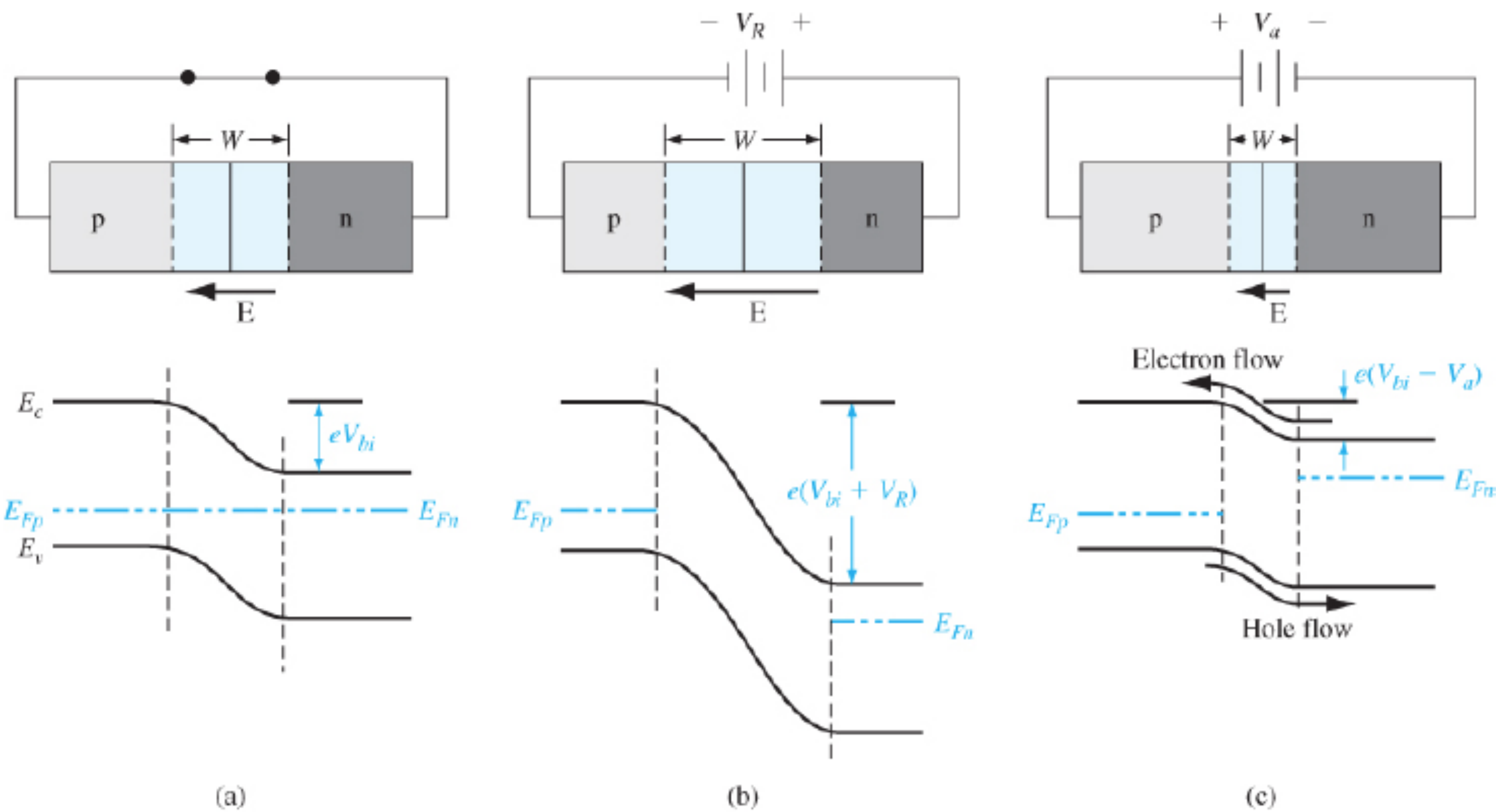
# Forward Bias, A Closer Look . . .

Forward-bias tipps the scales in favor of diffusion vs drift

Forward-bias diffusion current is determined by the **recombination rate** in the quasi neutral regions



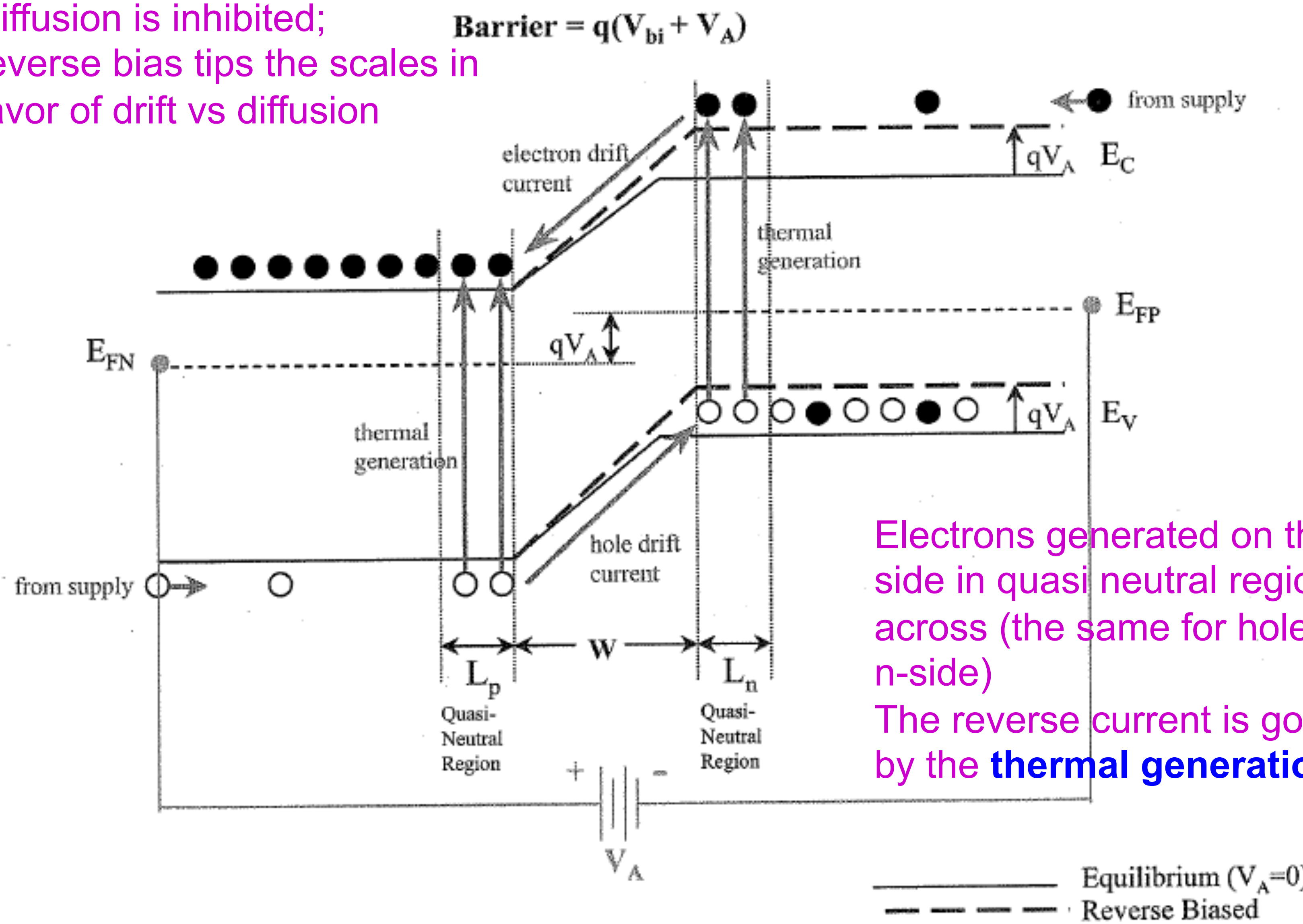

$L_p$= hole minority carrier diffusion length
$L_n$ = electron minority carrier diffusion length




zero bias | reverse bias | forward bias

# Reverse Bias, A Closer Look . . .

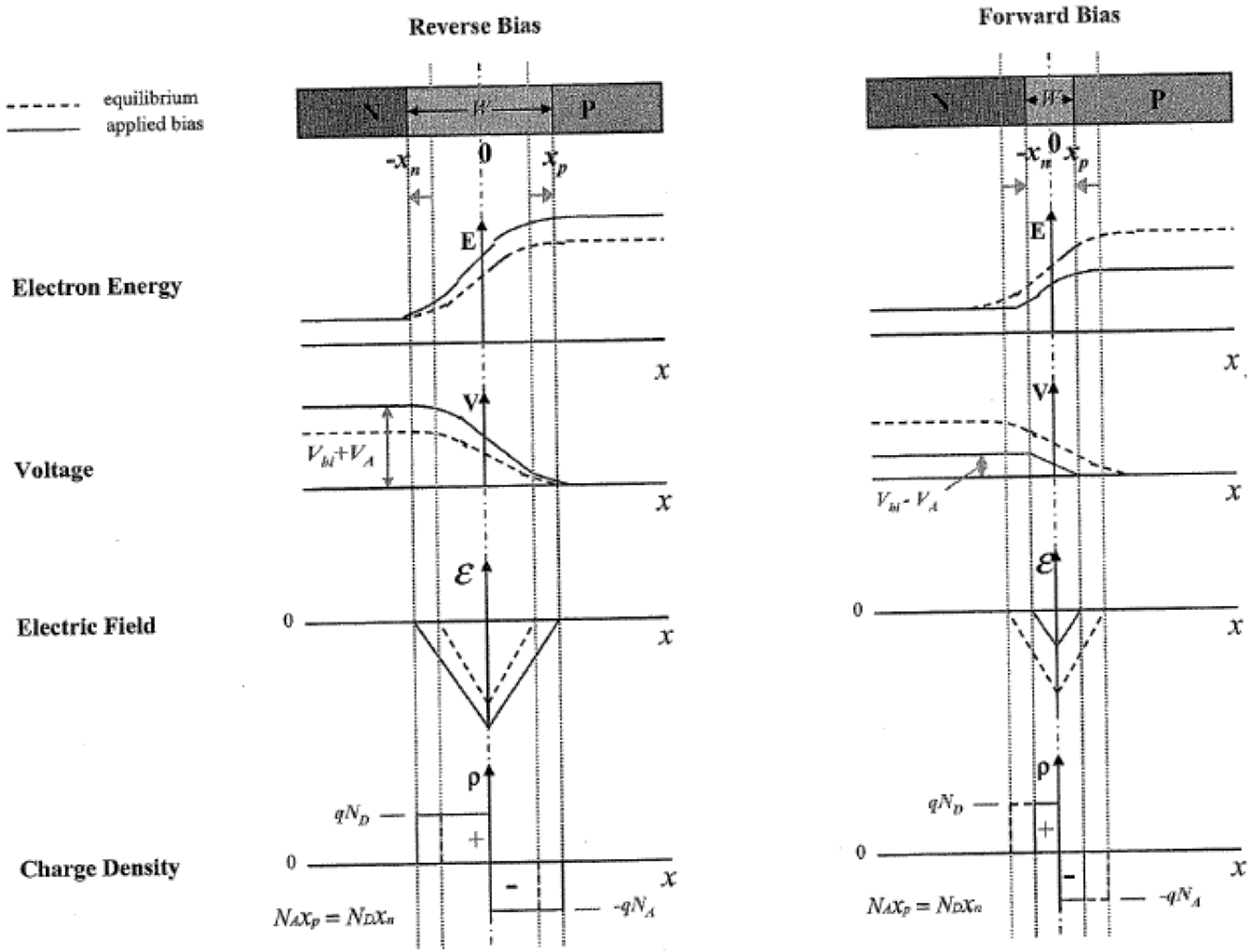
Diffusion is inhibited; reverse bias tips the scales in favor of drift vs diffusion




Electrons generated on the p-side in quasi neutral region drift across (the same for holes on the n-side)

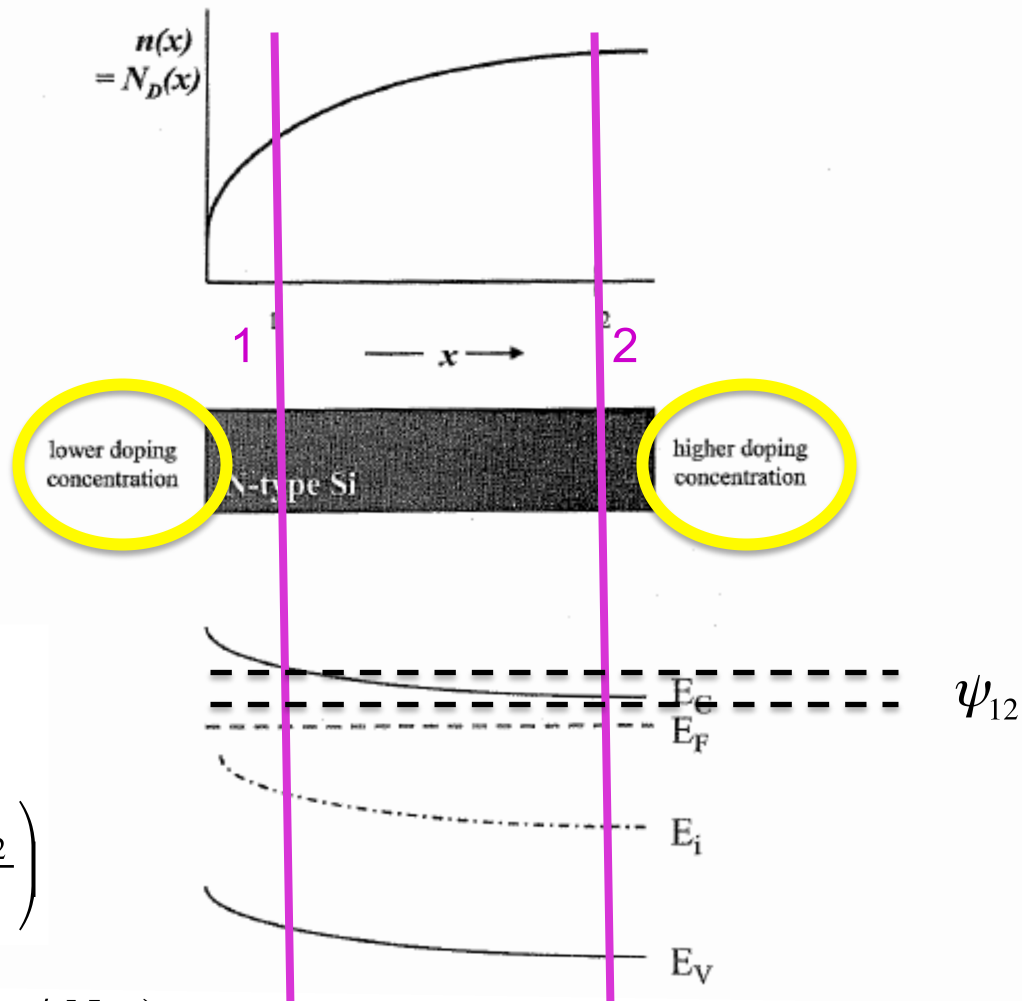
The reverse current is governed by the **thermal generation** term.

# Energy, Voltage, Field, & Charge Revisited



# Aside: Non-Uniformly Doped Silicon

Bands are bent at equilibrium, indicating that there is non-zero built-in potential (analogous to pn-junction)




$$n_1 = n_i \exp(E_F - E_{i1})/kT)$$

$$n_2 = n_i \exp(E_F - E_{i2})/kT)$$

$$\frac{n_2}{n_1} = \exp\left(\frac{E_{i1} - E_{i2}}{kT}\right) = \exp\left(\frac{q\psi_{12}}{kT}\right)$$

$$\psi_{12} = (E_{i1} - E_{i2})/q = kT/q(\ln N_{d1}/N_{d2})$$

$$E = (kT/q)\frac{1}{N_d}\frac{dN_d}{dx}$$

EXAMPLE

Doping varies from $10^{18}$ to $10^{17}$ within 0.1μm

E=0.026x1/$10^{18}$x($10^{18}$-$10^{17}$/$10^{-5}$)=2.6x$10^{3}$ V/cm

# Debye Length - reprised

1-30-2014

$$\frac{d^2\phi}{d^2x} = -\frac{q}{\varepsilon_s}(p - n + N_D - N_A)$$

$$N_D = n = n_i \exp(\frac{q\varphi_n}{kT})$$

$$n = n_i \exp(\frac{q\phi}{kT})$$

$$\frac{d^2\phi}{d^2x} = \frac{q}{\varepsilon_s}(2n_i \sinh(\frac{q\phi}{kT}) + N_A - N_D)$$

$$N_A = n = n_i \exp(\frac{q\varphi_p}{kT})$$

Two special cases:

**A)** dopant concentration varies gradually with position (small gradient case) This is the case of **quasi-neutrality.** The electric fields are of the order of $10^3$ - $10^4$ V/cm. In this case the electron concentration n can be approximated by $N_D$.

Electric field $$E_x = -\frac{D_n}{\mu_n}\frac{1}{n}\frac{dn}{dx} = -\frac{kT}{q}\frac{1}{n}\frac{dn}{dx} \approx \frac{kT}{q}\frac{1}{N_D}\frac{dN_D}{dx}$$

# Debye Length

Two special cases:

**B) Step junction**; we consider an analytical approximation for the potential **near the edges** of the space-charge region (i.e. near $x_n$ or $-x_p$). We consider small variations of potential from $\phi_n$ near $x=x_n$. We rewrite Poisson equation by neglecting the minority-carrier concentration p and letting

$$\varphi = \varphi_n - \varphi' \quad \textit{be the deviation from } \varphi_n$$

$$\frac{d^2\varphi'}{d^2x} = \frac{q}{\varepsilon_s}(N_D - n) = \frac{q}{\varepsilon_s}[N_D - n_i \exp(\frac{q(\varphi_n - \varphi')}{kT})] =$$

$$= \frac{q}{\varepsilon_s} N_D[1 - \exp(-\frac{\varphi'}{kT})]$$

Taylor expansion for small $\phi'$

$$\frac{d^2\varphi'}{d^2x} = \frac{q}{\varepsilon_s} N_D \frac{\varphi'}{kT} = \frac{\varphi'}{L_D^2}$$

$$L_D = \sqrt{\frac{\varepsilon_s kT}{q^2 N_D}}$$

$$\varphi' \sim \exp(-x / L_D)$$

Debye length $L_D$ is a characteristic length associated with spatial variations of potential or carrier distribution. We see that the potential does not drop abruptly but varies exponentially with distance near the edges of the space-charge region. The carrier concentration changes rapidly from that of the dopant concentration to essentially zero within a few Debye lengths, $\exp(-x/L_D)$.

Therefore, the depletion approximation is questionable only within a few extrinsic Debye lengths of the edges of space charge region near $x_n$ or $-x_p$ .

# PN Junction - Useful Equations Revisited

The following equations are valid for an abrupt (step) junction with uniform doping on both sides.

## Depletion Region Width

$$x_n = \sqrt{\frac{2\varepsilon_s(V_{bi}-V_A)}{q}\cdot\frac{N_A}{N_D(N_A+N_D)}}$$

$$x_p = \sqrt{\frac{2\varepsilon_s(V_{bi}-V_A)}{q}\cdot\frac{N_D}{N_A(N_A+N_D)}}$$

$$W = x_n + x_p = \sqrt{\frac{2\varepsilon_s(V_{bi}-V_A)(N_A+N_D)}{qN_AN_D}}$$

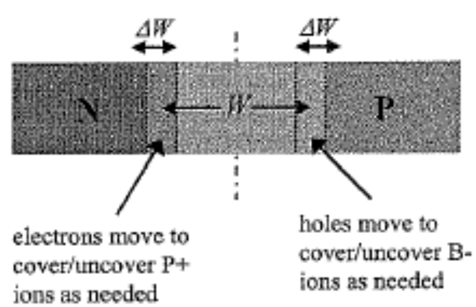
$V_A$ = applied voltage (equilibrium $V_A$=0; fwd bias $V_A$>0; rev bias $V_A$<0)

## Junction Capacitance

**Capacitance formula reformulated**

$$C_j = \frac{\varepsilon_s A}{W} = \frac{C_{j0}}{\left[1+\frac{V_A}{V_{bi}}\right]^{\frac{1}{2}}} \quad \text{where} \quad C_{j0} = \frac{\varepsilon_s A}{\sqrt{\frac{2\varepsilon_s V_{bi}(N_A+N_D)}{qN_AN_D}}}$$

$\Delta V_A \rightarrow \Delta W \rightarrow \Delta Q$



# PN Junction - Drift & Diffusion Currents

$$J = J_p + J_n$$

The total current through the junction is the sum of the electron current and hole current.

where:

$$J_n = J_n(drift) + J_n(diff)$$

$$J_p = J_p(drift) + J_p(diff)$$

The electron current, in turn, is made up of drift and diffusion current components. So is the hole current.

where:

J=qnv

v=μE

$$J_n(drift) = \sigma\mathcal{E} = q\mu_n n\mathcal{E}$$

$$J_p(drift) = \sigma\mathcal{E} = q\mu_p p\mathcal{E}$$

The drift current is calculated from Ohm's Law.

q = elementary charge (magnitude) = $1.602\times10^{-19}$ coul.

$\boxed{\begin{array}{l}\mu_n = \text{electron mobility } [= 1350\ cm^2/(V{\cdot}s) \text{ for Si, room temp.}] \\ \mu_p = \text{hole mobility } [= 480\ cm^2/(V{\cdot}s) \text{ for Si, room temp.}]\end{array}}$

and:

$$J_n(diff) = qD_n \frac{dn}{dx}$$

$$J_p(diff) = -qD_p \frac{dp}{dx}$$

The diffusion current is calculated from Fick's law.

$\boxed{\begin{array}{l}D_N = \text{electron diffusion coefficient } [= 33.75\ cm^2/s \text{ for Si at 300 K}] \\ D_p = \text{hole diffusion coefficient } [=12.4\ cm^2/s \text{ for Si at 300 K}]\end{array}}$

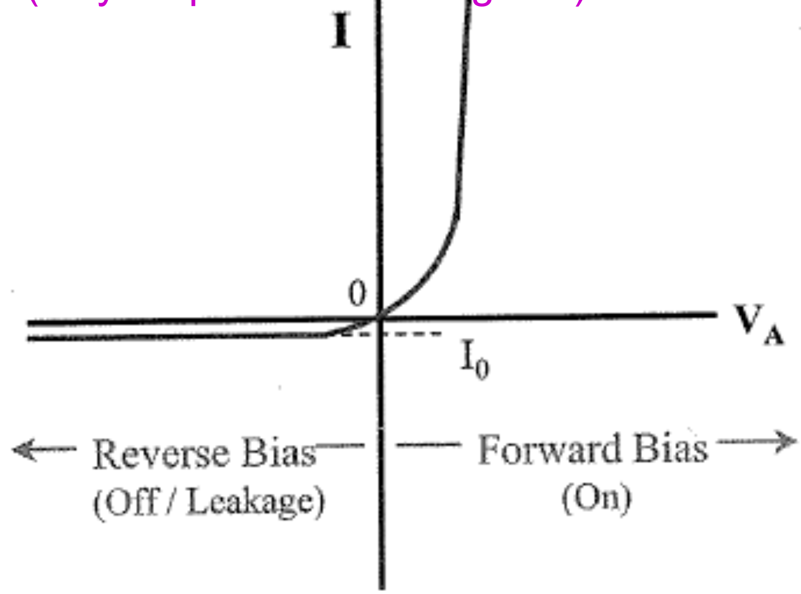
μ mobility

# Ideal Diode - Current-Voltage Characteristic

**Ideal diode assumptions:** a) Low-level injection $n_p \ll p_p$ and $p_n \ll n_n$
(b) No series resistance, voltage drop only across the depletion region
(c) No generation/recombination in depletion region (only in quasi-neutral regions)

**Ideal Diode: I-V Characteristic**

$$I = I_0\left[\exp\left(\frac{qV_A}{k_BT}\right)-1\right]$$




**Ideal Diode: Reverse saturation/leakage current**

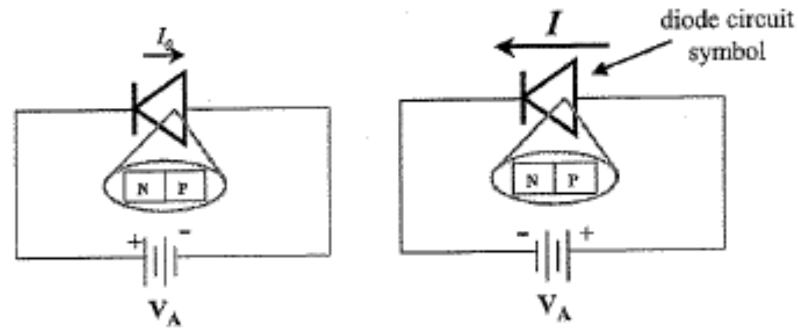
$$I_0 = qA\,n_i^2\left[\frac{D_n}{L_nN_A}+\frac{D_p}{L_pN_D}\right]$$

$D_n$ = electron diffusion coefficient [= 33.75 cm²/s at 300 K]
$D_p$ = hole diffusion coefficient [=12.4 cm²/s at 300 K]
$L_n$ = electron minority carrier diffusion length [~ 10 µm]
$L_p$ = hole minority carrier diffusion length [~10 µm]

$I_o$ larger for semiconductor with smaller band gap (because larger $n_i$)

**Diffusion coefficient can be calculated from mobility**

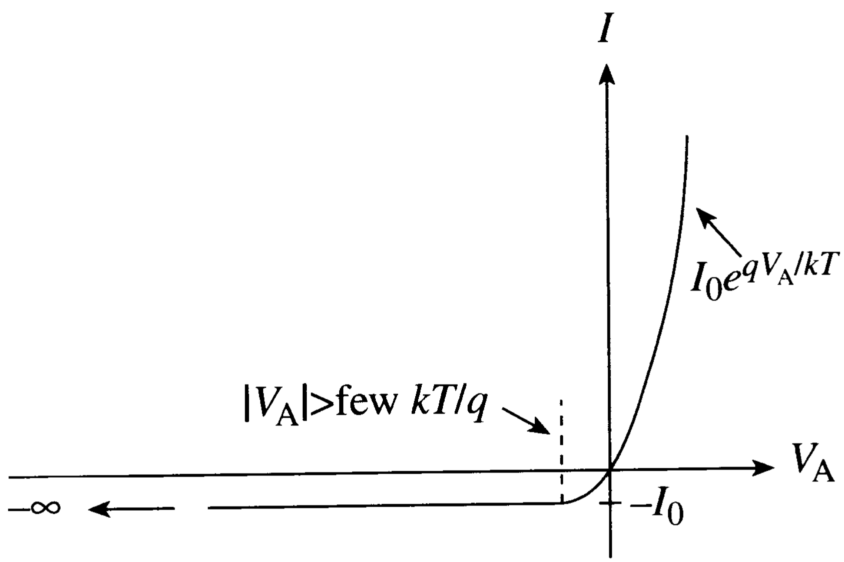
$$\frac{D}{\mu}=\frac{k_BT}{q} \qquad \text{(Einstein relationship)}$$

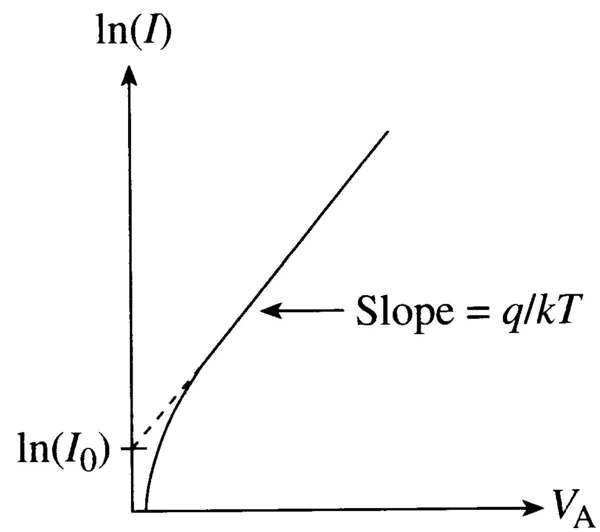



Reverse Forward

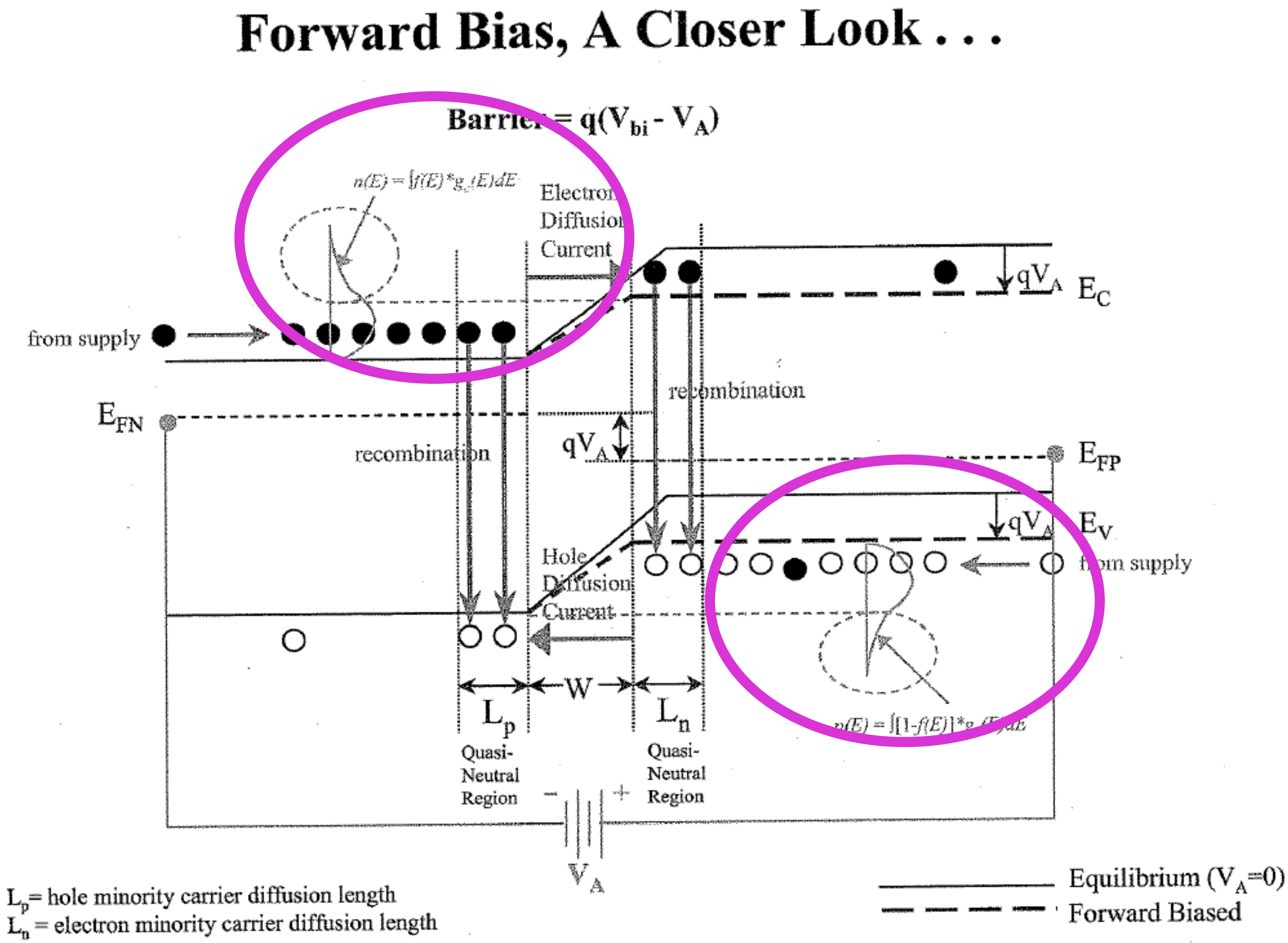
Exponential behavior on $V_A$: the barrier is lowered down the Fermi distribution tail

# Ideal Diode I-V Characteristics



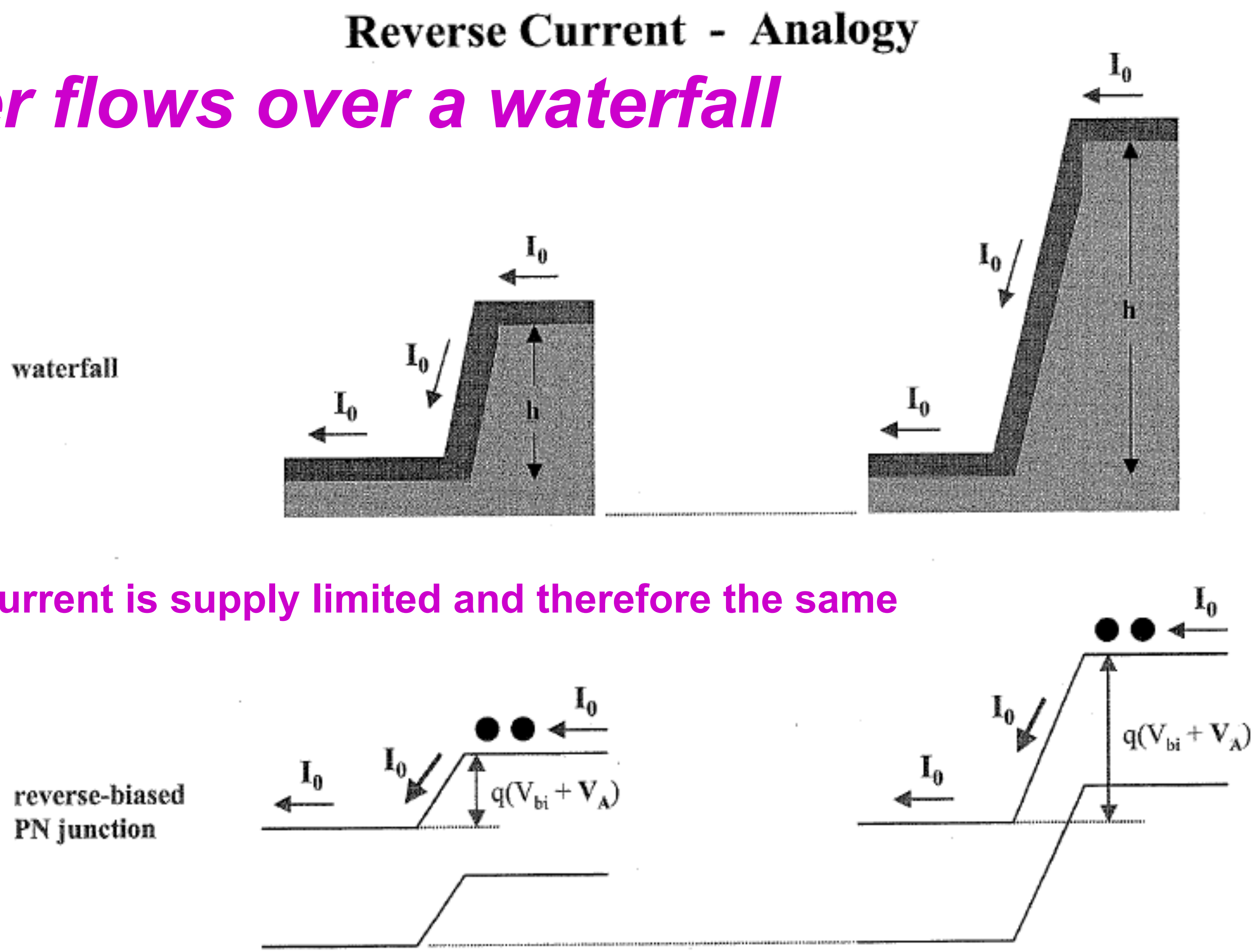



# **Intuitive picture** *of exponential behavior on forward $V_A$: the barrier is lowered down the Fermi distribution tail*



# Intuitive Picture for reverse bias:

*river flows over a waterfall*



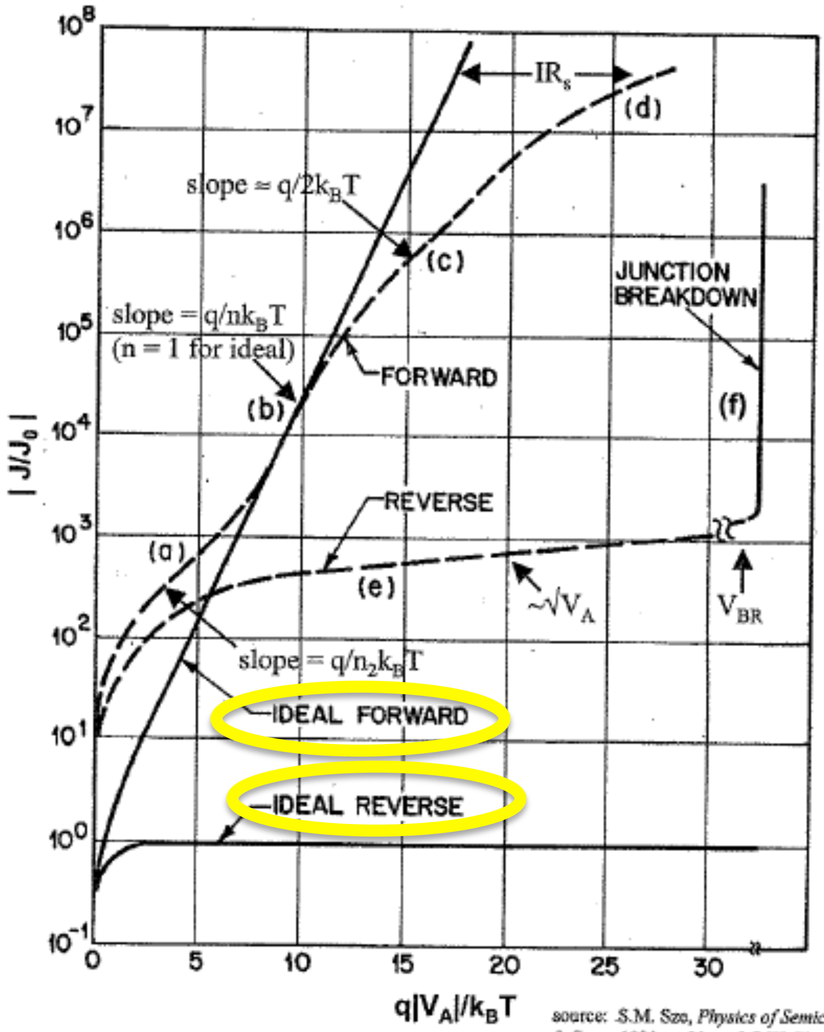

Current is supply limited and therefore the same

**But kinetic energy of the falling water could matter !!!**

# Assumptions of the ideal diode equation

- **Abrupt pn junction, uniformly and non-degenerately doped on both sides**
- **Low level injection ($n_p << p_p$ and $p_n << n_n$)**
- **No series resistance (no resistance in neutral and quasi-neutral regions**
- **No generation/recombination in depletion region (only in quasi-neutral)**
- **No external sources of carriers**

# PN Junction Diode - Non-Idealities




source: S.M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, 1981, p. 91; and G.W. Pierret, *Modular Series on Solid State Devices*, vol. 2, Addison-Wesley Publishing, 1988, p. 83-85.

## Deviations from the ideal:

Generation / Recombination in Depletion Region

$$I = I_0\left[\exp\left(\frac{qV_A}{nk_BT}\right)-1\right]+\frac{qAn_iW}{2\tau_0}\left[\exp\left(\frac{qV_A}{n_2k_BT}\right)-1\right]$$

Ideality Factor

$$where \quad \tau_0 = \frac{\tau_n+\tau_p}{2}$$

**Forward bias ($V_A > 0$):** Exponential increase in recombination current.
**Reverse bias ($V_A < 0$):** Increase in generation current goes as $W \sim \sqrt{V_A}$.

n = diode ideality factor [n = 1 for ideal diode, n > 1 for non-ideal]
$n_2$ = a slope factor for the second term [$n_2 \to 2$ is typical]
$\tau_n$ = electron minority carrier lifetime [~1μs typical]
$\tau_p$ = hole minority carrier lifetime [~1μs typical]
$\tau_0$ = effective (average) minority carrier lifetime

*aside:* $L_n \equiv \sqrt{D_n\tau_n}$

$L_p \equiv \sqrt{D_p\tau_p}$

# Ideal vs Non-ideal Diode

$$I = I_o\left[\exp\left(\frac{qVA}{kT}\right)-1\right]$$

carrier generation term
in depleted region

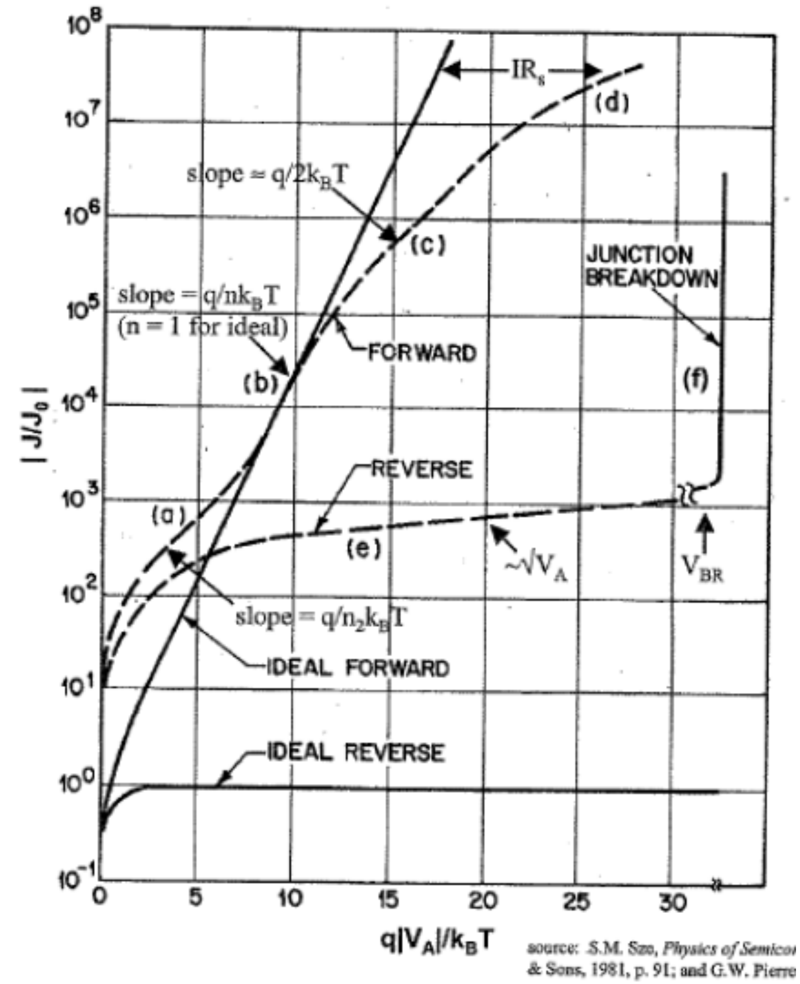
$$I = I_o\left[\exp\left(\frac{qV_A}{nkT}\right)-1\right]+\frac{qAn_iW}{2\tau}\left[\exp\left(\frac{qV_A}{nkT}\right)-1\right]$$

n – non-ideality factor 1<n<2
τ - average electron/hole life time

**•Ideality factor n**
n=1 ideal diode, real diodes 1<n<2
(grading of the junction {not really abrupt}
+ Recombination)

**•Second term added**
Generation and recombination in the
depletion region (e,f regions)

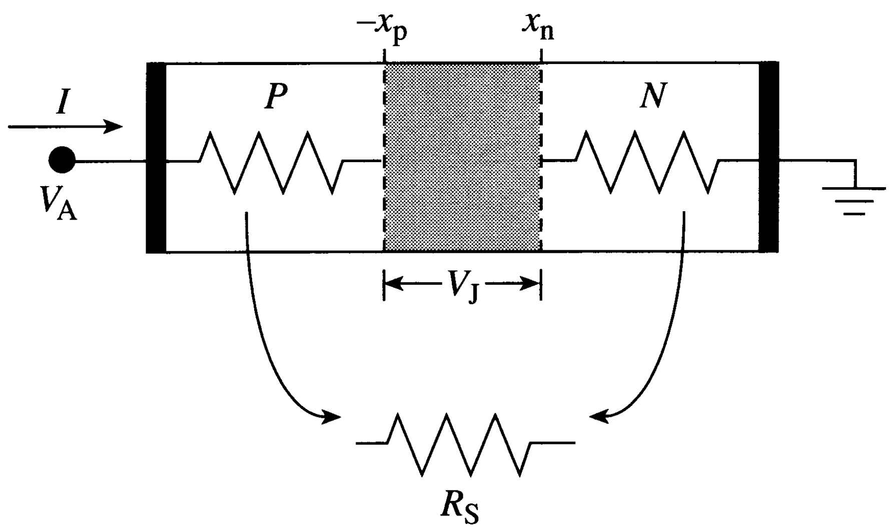
**• Non-idealities - reverse bias**
Generation and recombination in the
Depletion region (region e)
Junction breakdown (region f)




**• Non-idealities – forward bias**
Recombination in the depletion region
(comprehended in n factor) at low $V_F$ (region a)
Ideal behavior in region b
High injection (region c). The low level injection assumption does not hold at high $V_F$.
High injection reduces the recombination efficiency, hence reducing the current vs ideal

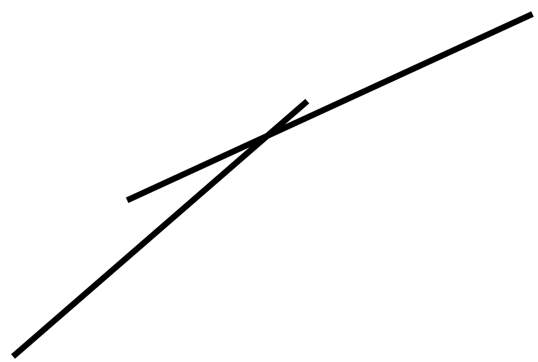
Series resistance (region d): voltage drop across the neutral regions and contacts
cannot be ignored → resistance $R_s$ (additional voltage drop $I \cdot R_s$).)

# Forward Bias with High Currents: **Series Resistance**




$$I = I_0[e^{qV/kT}-1]$$

$$I = I_0[e^{q(V-IR_S)/kT}-1]$$



# High Level Injection

High injection of carriers[1] causes to violate one of the approximations made in the derivation of the ideal diode characteristics, namely that the majority carrier density equals the thermal equilibrium value. Excess carriers will dominate the electron and hole concentration and can be expressed in the following way:

$$n_p p_p = n_i^2 \exp(\frac{V_a}{V_t}) = n_p (p_{p0} + n_p)$$

$$n_n p_n = n_i^2 \exp(\frac{V_a}{V_t}) = p_n (n_{n0} + p_n)$$

Solving the resulting quadratic equation yields:

$$n_p = \frac{N_a}{2}(\sqrt{1 + \frac{4n_i^2 \exp(\frac{V_a}{V_t})}{N_a^2}} - 1) \cong n_i \exp(\frac{V_a}{2V_t})$$

$$p_n = \frac{N_d}{2}(\sqrt{1 + \frac{4n_i^2 \exp(\frac{V_a}{V_t})}{N_d^2}} - 1) \cong n_i \exp(\frac{V_a}{2V_t})$$

$$J_n + J_p = q(\frac{D_n}{L_n} + \frac{D_p}{L_p}) n_i \exp(\frac{V_a}{2V_t})$$

This means that high injection in a p-n diode will reduce the slope on the current-voltage characteristic on a semi-logarithmic scale to 119mV/decade.

# High Level Injection

$$np = n_i^2 \exp\left(\frac{V_a}{V_t}\right)$$

We have that $n = n_o + \delta n$ and $p = p_o + \delta p$, so that

$$(n_o + \delta n)(p_o + \delta p) = n_i^2 \exp\left(\frac{V_a}{V_t}\right)$$

Under high-level injection, we may have $\delta n > n_o$ and $\delta p > p_o$ so that Equation (8.60) becomes approximately

$$(\delta n)(\delta p) \cong n_i^2 \exp\left(\frac{V_a}{V_t}\right)$$

Since $\delta n = \delta p$, then

$$\delta n = \delta p \cong n_i \exp\left(\frac{V_a}{2V_t}\right)$$

The diode current is proportional to the excess carrier concentration so that, under high-level injection, we have

$$I \propto \exp\left(\frac{V_a}{2V_t}\right)$$

**In the high-level injection region, it takes a larger increase in diode voltage to produce a given increase in diode current.**

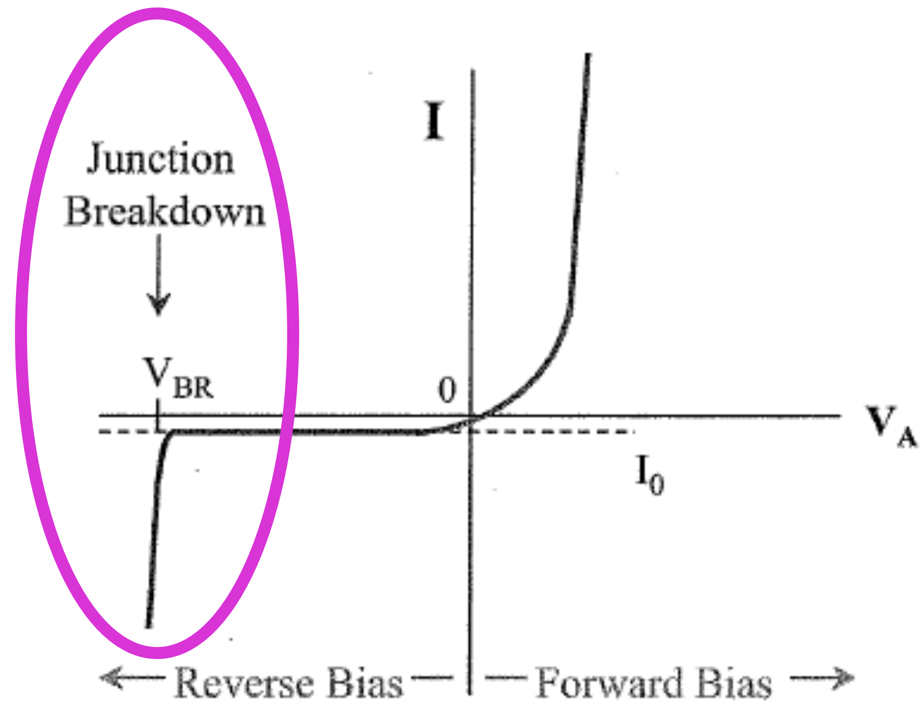
# PN Junction Diode - Reverse Breakdown

**Avalanche Breakdown Voltage**

$$V_{BR} = \frac{\varepsilon_s \mathcal{E}_{CR}^{\ 2}}{2q} \frac{N_A + N_D}{N_A N_D}$$

$\mathcal{E}_{CR}$ = critical field for impact ionization [= 4x10$^5$ V/cm for Si]

Increasing the doping on either or both sides of the junction decreases the breakdown voltage.

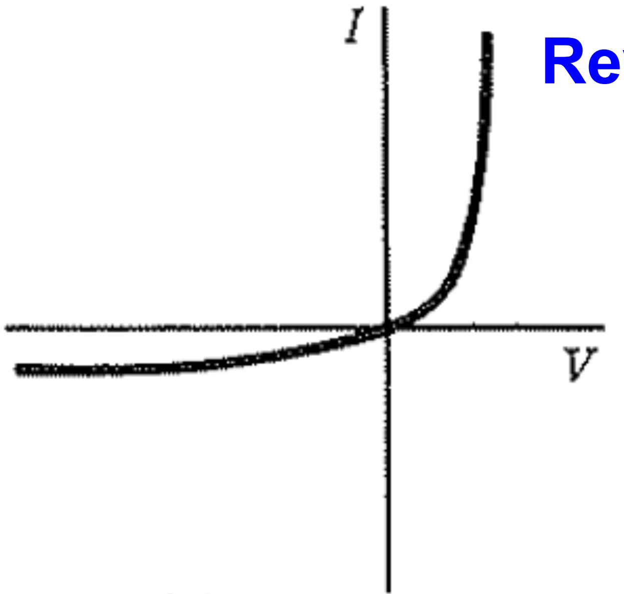



**Two types of breakdown:**

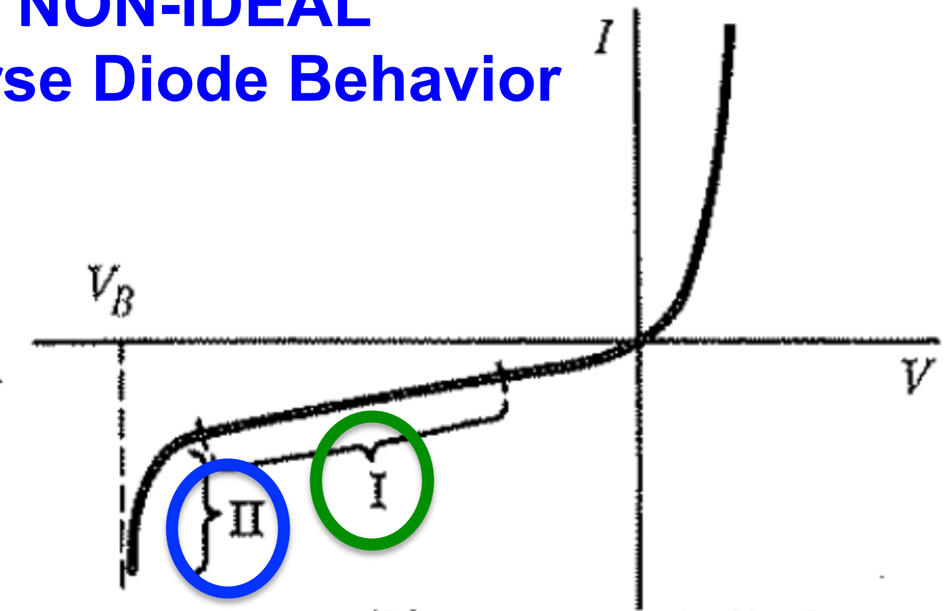
**Avalanche Breakdown (or impact ionization) most common**

**Zener Breakdown (tunneling) if both sides are heavily doped**

# NON-IDEAL
## Reverse Diode Behavior




(a) Ideal diodes



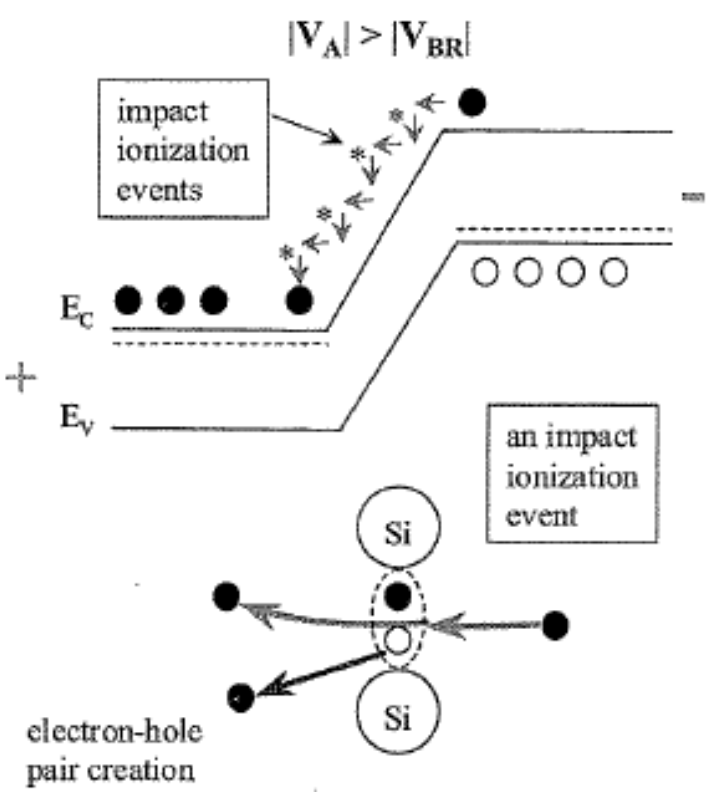

(b) Practical diodes

**Region I : slow rise of current in region I is due to space charge region increase with increased reverse bias and hence raising the reverse saturation current.**

**Region II: rapid rise of current in region II is due to a phenomenon called breakdown; there are two mechanisms for the breakdown:**

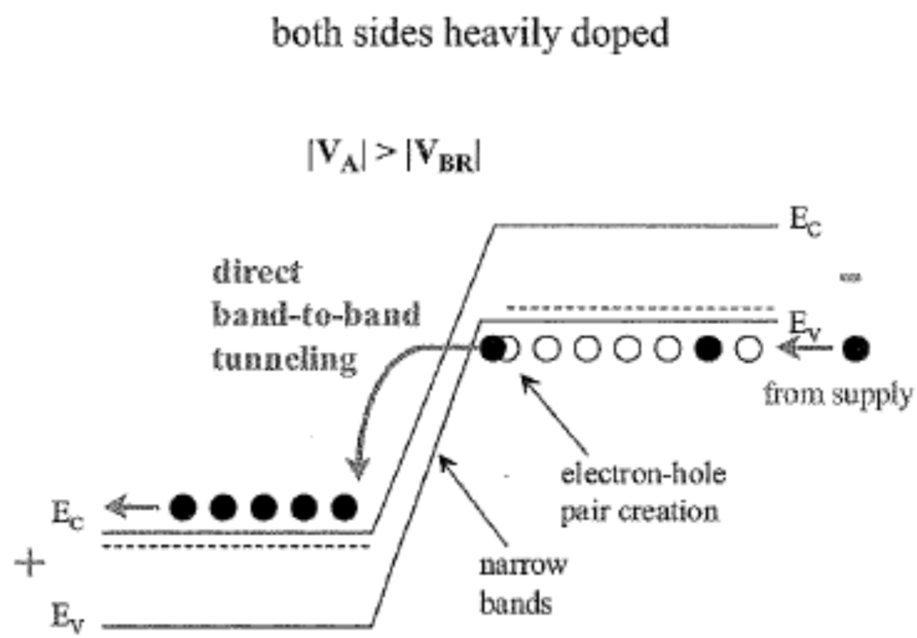
- **Avalanche**
- **Tunneling**

# PN Junction Diode - Reverse Breakdown (cont.)




Hot electron collides with Si atom

**Avalanche Breakdown**




Barrier sufficiently thin for electron tunneling

**Zener Breakdown**

# Break-down Voltage $V_{BR}$

Recall that for a step function for $N^+/P$ or $P^+/N$ junction we have

$$\tfrac{1}{2} W \cdot E_{\max} = (\phi_{bi} - V_a)$$

$$E_{\max} = \frac{2(\phi_{bi} - V_a)}{W}$$

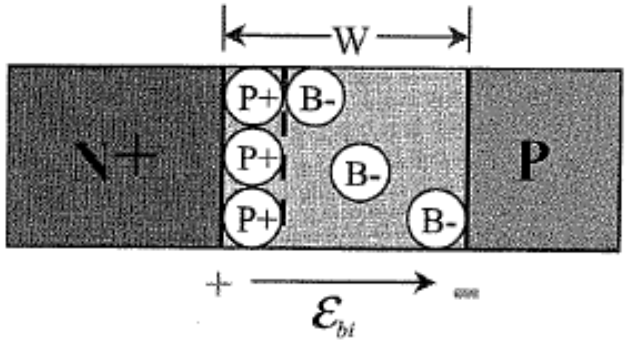
$$E_{\max} \approx \left( \frac{2qN_A |V_R|}{\varepsilon_s} \right)^{1/2} \rightarrow E_{crit}$$

$$V_{BR} = \frac{\varepsilon_s E_{crit}^2}{2qN_A}$$

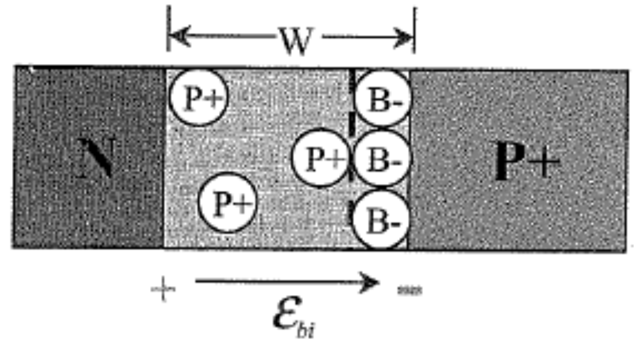
# PN Junction Diode - One-Sided Junction

## Break down Voltage

$N_D >> N_A$




$N_A >> N_D$




$$W = \sqrt{\frac{2\varepsilon_s(V_{bi} - V_A)}{qN_A}}$$

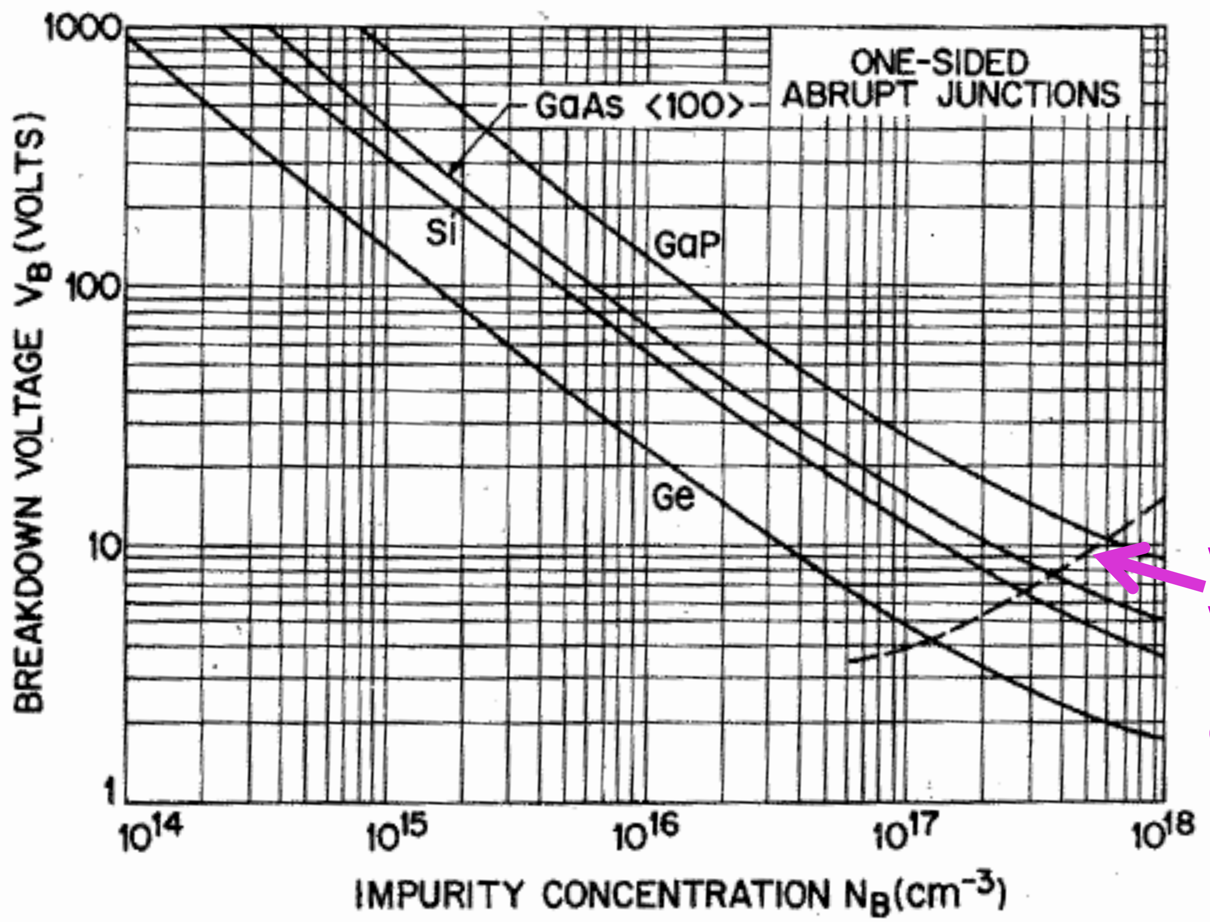
$$W = \sqrt{\frac{2\varepsilon_s(V_{bi} - V_A)}{qN_D}}$$

$$V_{BR} = \frac{\varepsilon_s \mathcal{E}_{CR}^{\;2}}{2qN_A}$$

$$V_{BR} = \frac{\varepsilon_s \mathcal{E}_{CR}^{\;2}}{2qN_D}$$

**$E_{CR} > 2 \times 10^5$ V/cm**

# Breakdown Voltage vs. Dopant Concentration

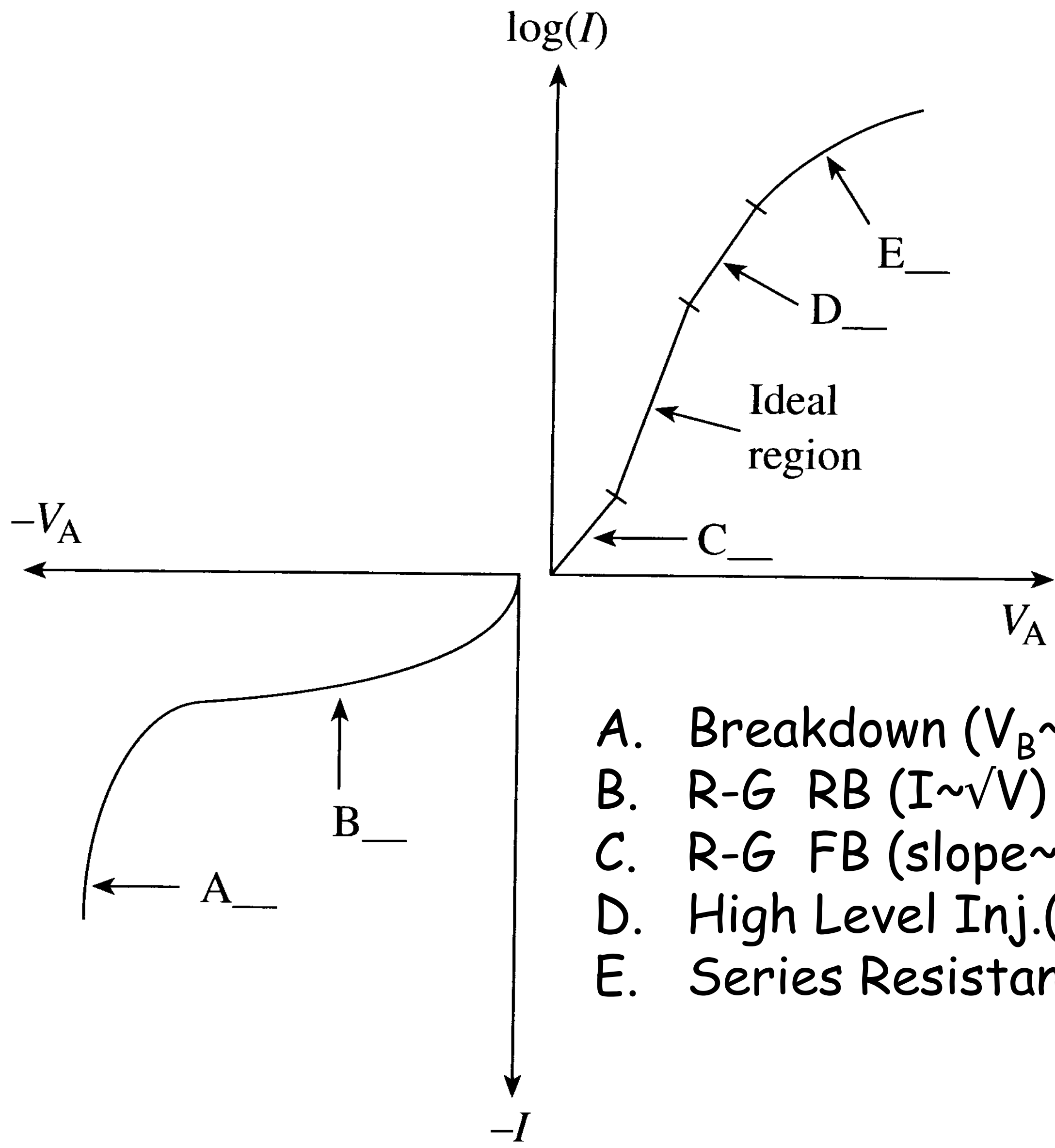



Doping levels beyond which Zener tunneling will dominate the breakdown characteristics

The doping on the lighter-doped side

source: S.M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, 1981.

# Real Diode I-V curve summary

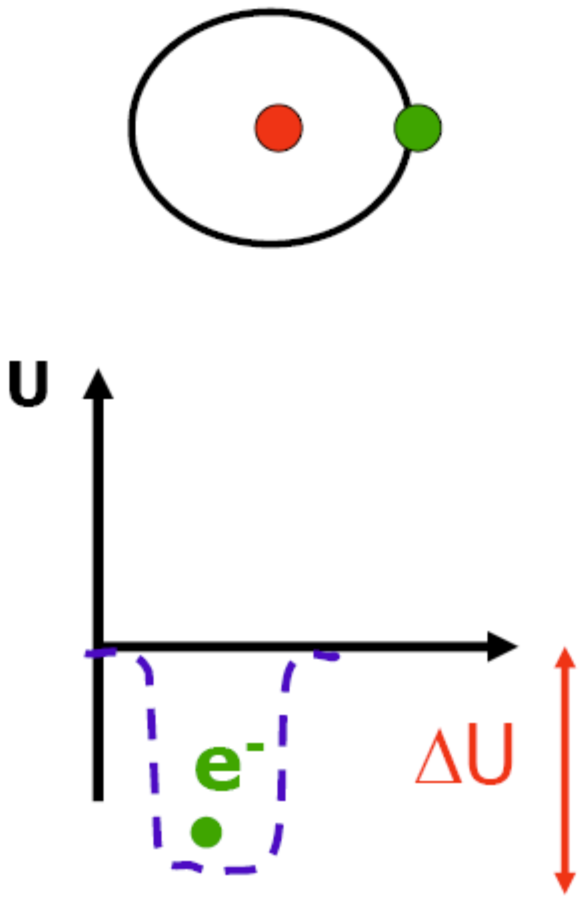



A. Breakdown ($V_B$~$1/N_B$)
B. R-G RB (I~√V)
C. R-G FB (slope~q/2kT)
D. High Level Inj.(slope ~ q/2kT)
E. Series Resistance – slope over

# Impact Ionization
# and
# Avalanche

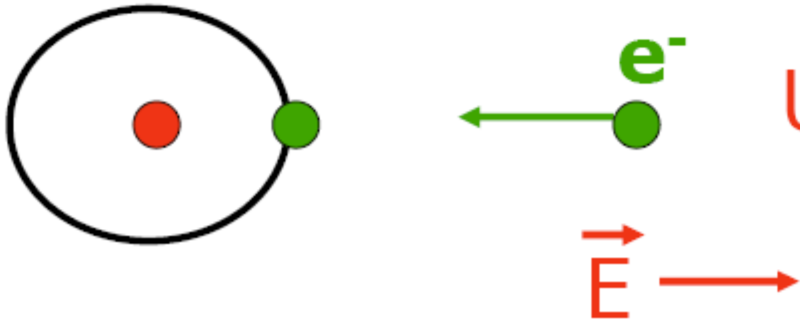
Precursor to avalanche:

# Ionization In Gases

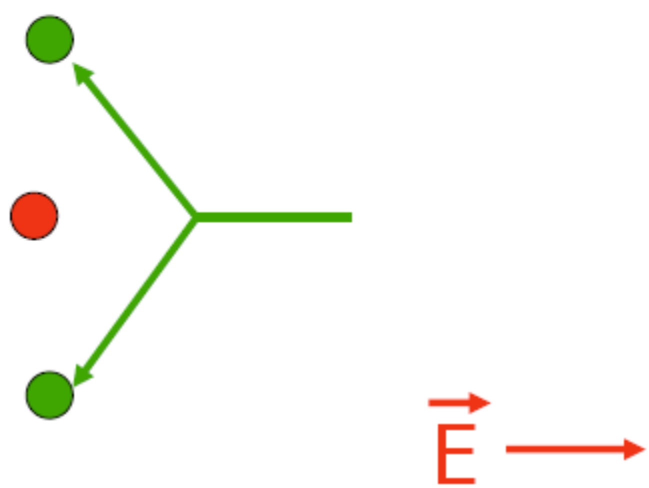



- Electrons and nucleus bound together
- Electrons stuck in potential well of nucleus
- Need energy $\Delta U$ to jump out of well
- How to provide this energy?

# Impact Ionization In Gases




$U_{kin} > \Delta U$




- Define $V_{ion} = \Delta U/q$
- Ionization potential

- One $e^-$ in, two $e^-$ out
- Avalanche?

$$E_{kin} = q\int_0^{\lambda} \vec{E} \cdot d\vec{x}$$

e-h creation in a semiconductor

$$V_{ion} \cdot q \approx \frac{3}{2} E_g$$

# Center of mass kinetic energy

$$K = K_1 + K_2 = \tfrac{1}{2} m_1\mathbf{v_1}^2 + \tfrac{1}{2} m_2\mathbf{v_2}^2$$

$$\mathbf{u_1} = \mathbf{v_1} - \mathbf{v_{CM}}$$
$$\mathbf{u_2} = \mathbf{v_2} - \mathbf{v_{CM}}$$

$$M\mathbf{v_{CM}} = m_1\mathbf{v_1} + m_2\mathbf{v_2} \quad \Rightarrow \quad \mathbf{v_{CM}} = \frac{m_1\mathbf{v_1} + m_2\mathbf{v_2}}{M}$$

$$K_{CM} = \tfrac{1}{2} M\mathbf{v_{CM}}^2$$

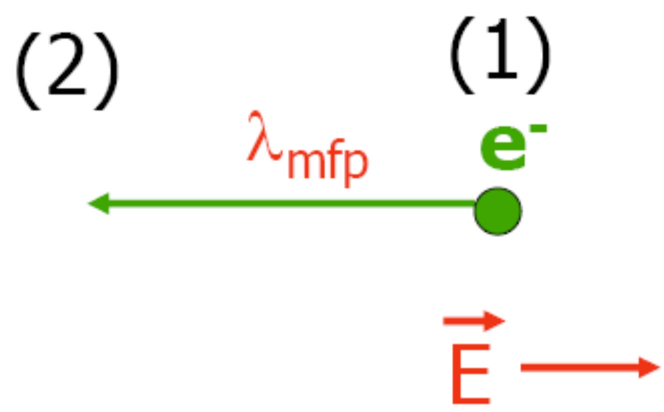
Note that because the total momentum of the system has to be conserved, the velocity, momentum, and kinetic energy of the centre of mass are conserved throughout the collision (this is true for any type of collision, including inelastic and reactive ones). Energy 'tied up' in the motion of the centre of mass is therefore not available for the collision. For a reactive collision, this energy does not help overcome any activation barrier that might be present.

$$K_{rel} = \tfrac{1}{2} \mu\mathbf{v_{rel}}^2$$

reduced mass of the two particles, $\mu = m_1m_2/M$

$$\mathbf{v_{rel}} = \mathbf{v_2} - \mathbf{v_1}$$

# Impact Ionization In Gases




$\lambda_{mfp}$ : Mean Free Path

- To get avalanche we need:

  $\Delta U_{kin}$ between collisions (1) and (2) $> V_{ion} * e$

- Acceleration in uniform Field

  $\Delta U_{kin} = V_2 - V_1 = e\, E\, d_{12}$

- Avalanche condition then

  $E > V_{ion} / \lambda_{mfp}$

## Impact Ionization and Avalanche in Semiconductors

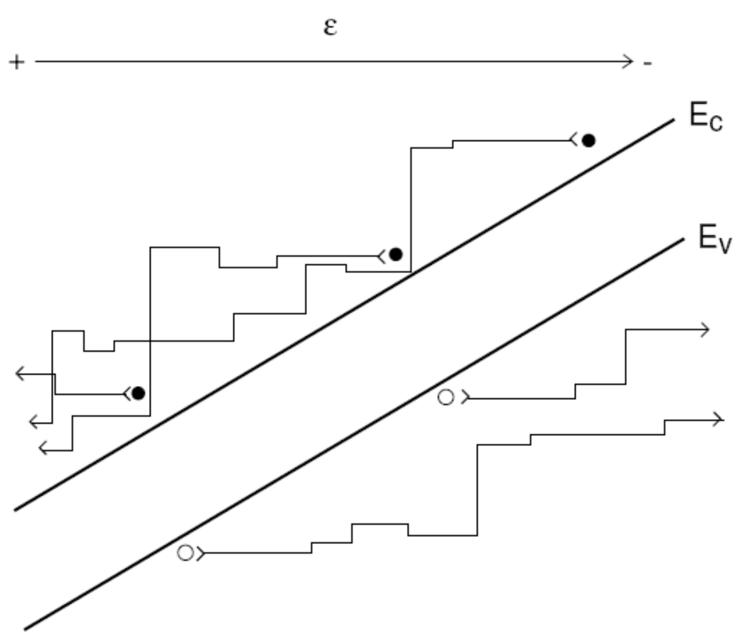
Lattice-scattering effects can result in the creation of e-h pairs if the carrier being scattered has sufficient energy. For example, if the electric field E is large, an electron entering from the p-side may be accelerated to high enough kinetic energy to cause an ionizing collision with the lattice. The carrier multiplication can become high **if carriers generated within the transition region also have ionizing collisions with the lattice. This is an avalanche process, since each incoming carrier can initiate the creation of a large number of new carriers.**

# Impact Ionization and Avalanche in Semiconductors

Assume that a carrier has a probability P of having an ionizing collision while being accelerated over a distance $\lambda$. After first collision we have $n_{in}(1+P)$ carriers. Now the secondary carriers $n_{in}P$ might undergo a ionizing collision resulting in $n_{in}PxP=n_{in}P^2$ of tertiary e-h pairs. Summing up the total number of electrons, we have

$n_{out}=n_{in}(1+P+P^2+P^3+\ldots)= n_{in}/(1-P)$.

We can define a multiplication factor according to: $$M_n = \frac{n_{out}}{n_{in}} = \frac{1}{1-P}$$

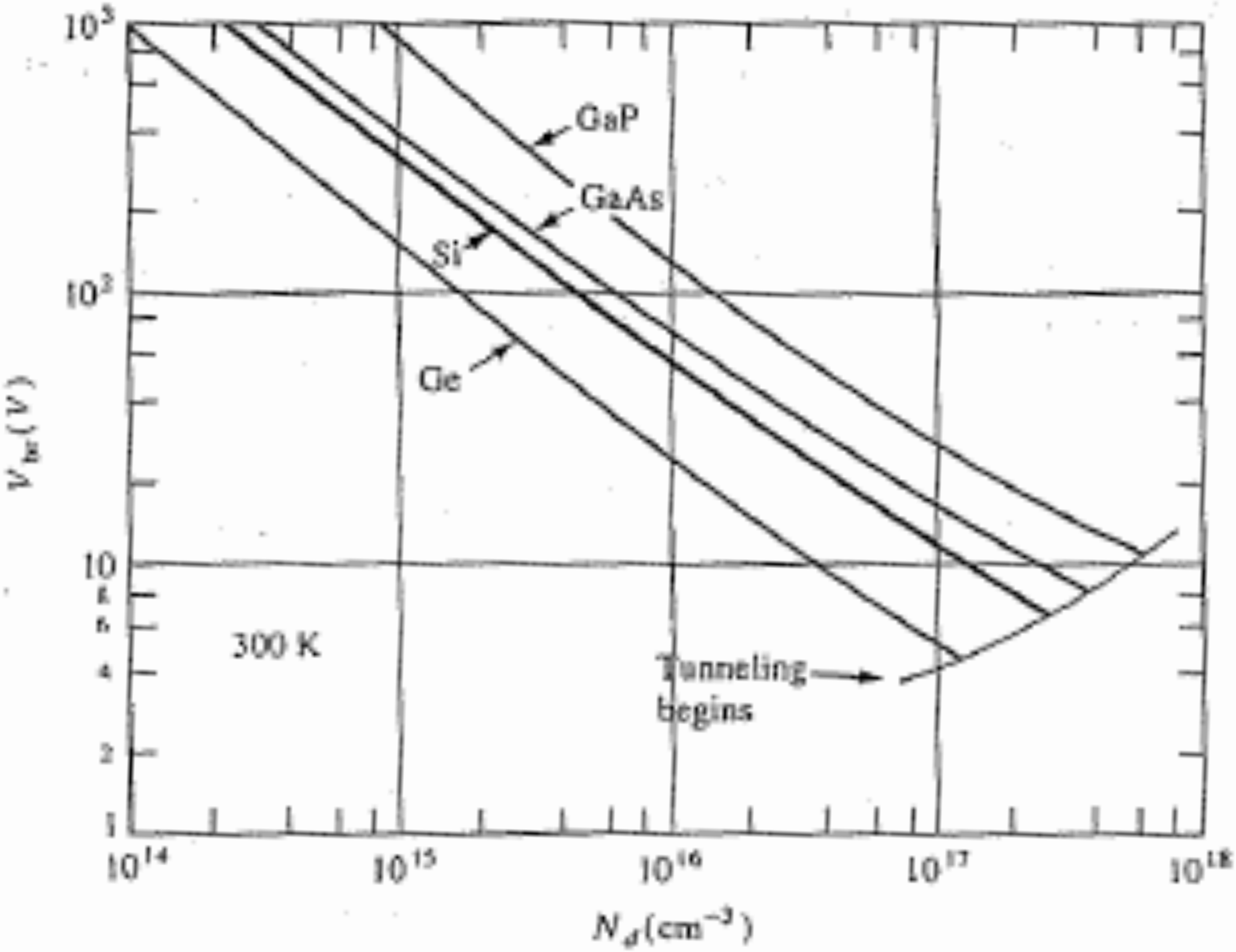



**As the probability of ionization approaches unity the carrier multiplication increases without limit.**

Empirically, one has found

$$M = \frac{1}{1-(V/V_{BR})^n} \quad where \quad n = 3-6$$

**Figure 5-19** Variation of avalanche breakdown voltage in abrupt $p^+$-n junctions, as a function of donor concentration on the n side, for several semiconductors. [After S. M. Sze and G. Gibbons, *Applied Physics Letters*, vol. 8, p. 111 (1966).]



# Impact Ionization Coefficient

$$G_{II} = \alpha_n n v_n + \alpha_p p v_p \qquad \alpha = AE\exp(-\frac{B}{E})$$

**$\alpha_n$ is the electron ionization rate defined as the number of e-h pairs generated by an electron per unit distance traveled. $\alpha_p$ is the hole ionization rate, n and p are electron and hole concentrations and $v_n$ and $v_p$ are the thermal velocities for electrons and holes, respectively.**

# Derivation:

$$n^* = n \cdot \exp(-d/\lambda)$$

$$dn = Cn^* \frac{dx}{d}$$

$$d = \Delta E_{\min}/(qE_{el})$$

$$dn = Cn^* \frac{dx}{d} = \frac{CqE_{el}}{\Delta E_{\min}} \cdot \exp(-\frac{\Delta E_{\min}}{\lambda q E_{el}}) \cdot n \cdot dx$$

$$\boxed{\alpha = AE_{el}\exp\left(-\frac{B}{E_{el}}\right)}$$

**$n^*$ – density of ionizing collisions**

**exp(-d/λ) probabilty for no collision in a distance d**

**d – distance necessary to gain adequate energy**

**$\Delta E_{min}$ minimum energy necessary for an ionizing collision**

**λ is the mean-free path**

**dx/d number of ionizing collisions in the distance d.**

# Derivation of the Impact Ionization Coefficient

n* density of ionizing electrons is n*=n x p(d) where n is the total density of electrons and p(d) is the probability that electron traveling a distance d **has not suffered** a collision

$$p(d) = \exp\left(-\frac{d}{\lambda}\right)$$

where $\lambda$ is the mean free path. In Si $\lambda$ is about 100A.

If d=10A then p(10A)=0.9, if d=100A then p(100A)=0.37; if d=1000A then p(1000) =4.5E-4.

The length d necessary to gain adequate energy to break the bond (cause e-h pair creation, can be found from the minimum ionization energy $E_{min}$ and the electric field $E_{el}$ that accelerates the electron

Example: For $E_{el}=5\text{x}10^5$V/cm and $E_{min}=1.5\text{x}E_g$
We obtain d=33 nm and p(33nm)=0.037

$$E_{\min} = q \cdot d \cdot E_{el} \quad or \quad d = \frac{E_{\min}}{qE_{el}}$$

# Derivation of the Impact Ionization Coefficient

The number of ionizing collisions in the distance dx is proportional to dx/d if we assume that the electron collides immediately (or soon enough) after gaining sufficient energy to ionize an atom.

Then the number of electrons dn created in such collisions will be

$$\delta n \propto n^* \times \frac{dx}{d} \quad or \quad \delta n = A \times n^* \times \frac{dx}{d}$$

where A is a proportionality constant. Thus

$$n^* = n \times p(d) = n \times \exp\left(-\frac{d}{\lambda}\right) \qquad d = \frac{E_{\min}}{qE_{el}}$$

And $\delta n = A \times n^* \times \dfrac{dx}{d} = A \times n \times \exp\left(-\dfrac{E_{\min}}{\lambda q E_{el}}\right) \times \dfrac{qE_{el}}{E_{\min}} dx$

We can write

$$\delta n = \alpha \times n \times dx \quad with \quad \alpha = A \times \exp\left(-\frac{E_{\min}}{\lambda q E_{el}}\right) \times \frac{qE_{el}}{E_{\min}} \quad or \quad \alpha = K \times E_{el} \times \exp\left(-\frac{B}{E_{el}}\right)$$

# Estimate of d

**d-distance necessary to gain ionizing energy**

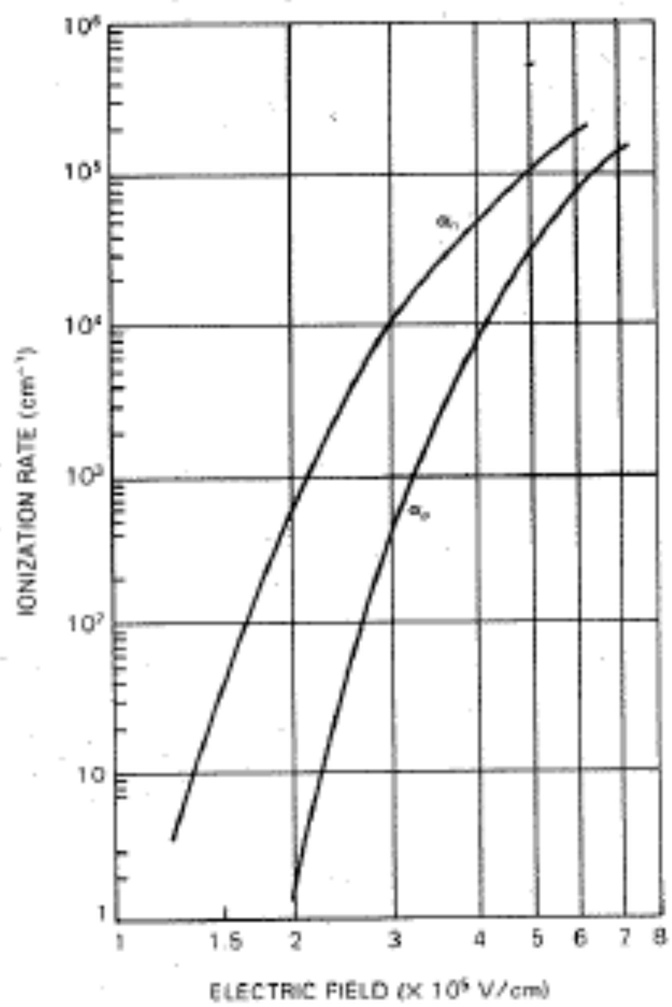
Let $\Delta E_{min}$ be the minimum energy necessary for an ionizing collision and $E_{el}$ be the average electric field that accelerates electron:

$$d = \Delta E_{\min} / (qE_{el})$$

# Impact Ionization Coefficient

**Small increase in field as the voltage approaches breakdown voltage causes sharp increase in current.**

for silicon




$$\alpha_n = 3.8\times10^{6}\exp(-1.75\times10^{6}/E)$$

$$\alpha_p = 2.25\times10^{7}\exp(-3.26\times10^{6}/E)$$

The ionization rates are known to decrease as the temperature increases. (why?)

The ionization rate for electrons is higher than that for holes by one-to-two orders of magnitude.

# Impact Ionization or Multiplication Coefficient

Since the maximum field in a $pn^+$ junction is approximately given by

$$E_{\max} = E_{crit} = \left( \frac{2qN_a |V_R|}{\varepsilon_s} \right)^{1/2}$$

$V_{BR}$ in equation

$$M = \frac{1}{1-(V/V_{BR})^n} = \frac{1}{1-\int_{x_a}^{x_c} \alpha \cdot dx}$$

$$x_c - x_a = W$$

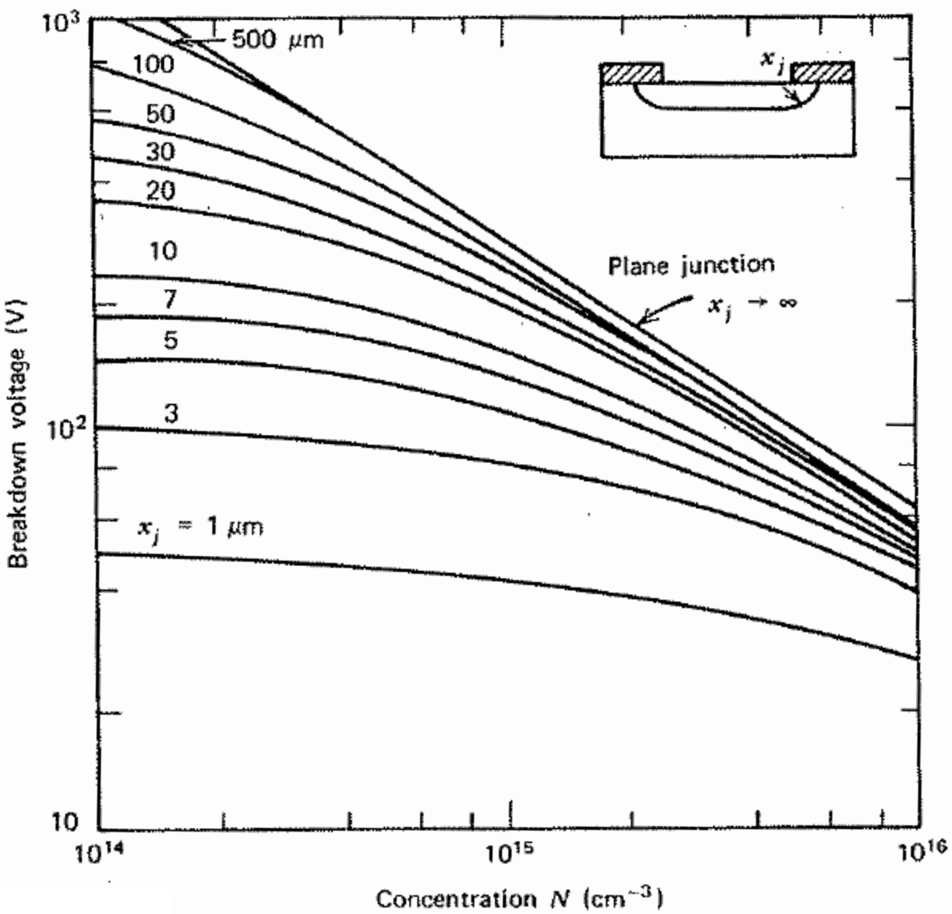
W is width of the depletion region

One obtains approximately

$$V_{BR} = \frac{\varepsilon_s E_{crit}^2}{2qN_a}$$

The last equation illustrates that the breakdown decreases for more heavily doped material.
Although the decrease is not quite as rapid as indicated in the equation above; one finds that $V_{BR}$ varies with doping as $N_a^{-2/3}$. At higher $N_a$ slightly higher critical field $E_{crit}$ is needed.

# Impact of Geometry on Avalanche




Breakdown voltage of one-sided, planar, silicon step junction showing the effect of junction curvature.[4,5]

**The field in the sharp corner regions can be markedly higher than in the rest of the junction, causing breakdown to occur there at unexpectedly low voltages.**

**The reduction in breakdown voltage is especially severe for a shallow junction with a small curvature radius. It becomes less serious as the junction depth increases.**

# Temperature Dependence of Avalanche

Impact ionization coefficient decreases with temperature because the free mean path $\lambda$ decreases with increasing temperature (more scattering of carriers with the lattice).

# Zener Breakdown or Band-to-Band-Tunneling

Two pictures:

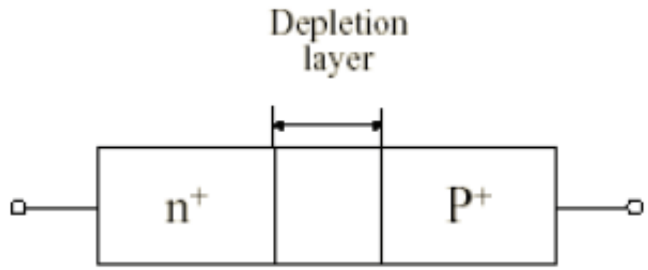
1) Electric field high enough to free (rip off) a (covalently) bound electron from the atom without interaction of any other particle.

2) Electron makes a tunneling transition from a valence band to conduction band if the tunneling barrier is thin enough (requires high electric fields

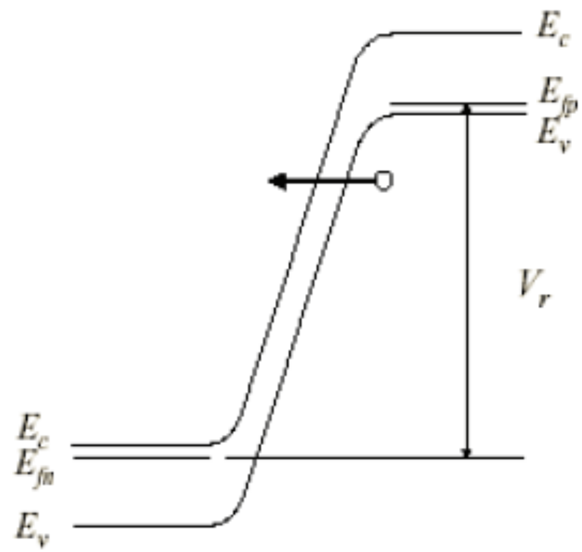
**Both pictures are equivalent !!**

# Zener breakdown

- Band to band (quantum-mechanical tunneling) occurs when valence electrons in p side pass through the depletion layer and reach the n-side conduction band.



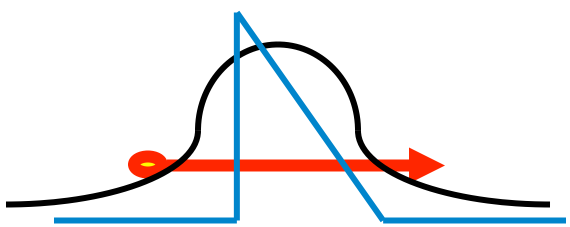

- This mechanism is possible if the depletion layer is very thin and the electric field is high.
- High impurity concentration on both the p and n sides ***(> 5x10$^{17}$ cm$^{-3}$)***.
- Zener Breakdown field is about ***10$^{6}$ V/cm***.
- The depletion layer is thin and the breakdown field is reached approximately with a reverse bias less than $4E_g/q$ ***(≈ 6V)***.



# Band-To-Band Tunneling

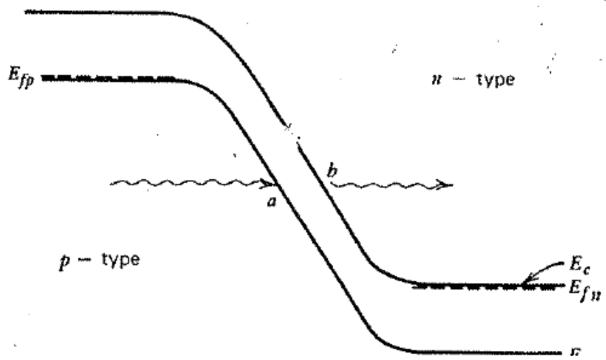
- Tunneling is a quantum mechanical phenomenon with no analog in classical physics
- Occurs when an electron passes through a potential barrier without having enough energy to do so

Golf ball “tunnels” through the mound.



Triangular barrier

# Band-To-Band Tunneling

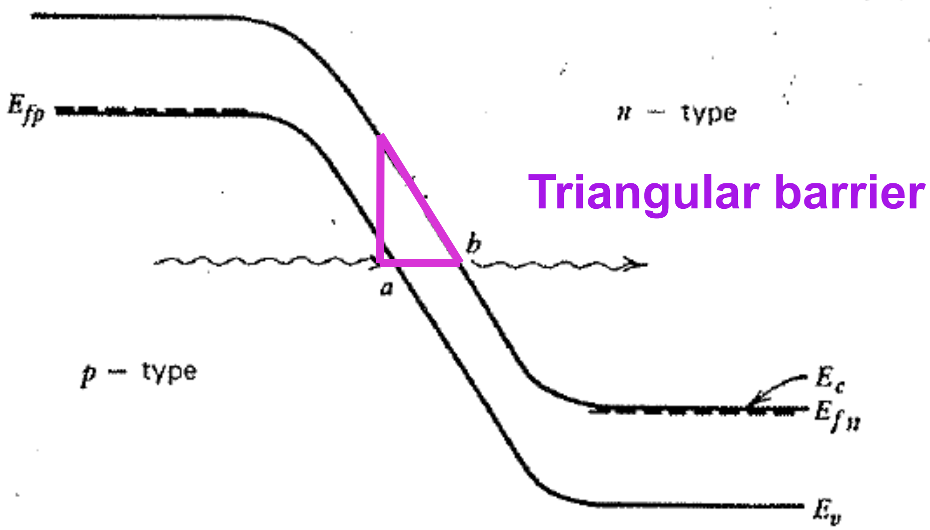



Zener breakdown occurs when the force exerted by the electric field is strong enough to rip an electron from its covalent bond to create electron-hole pair.

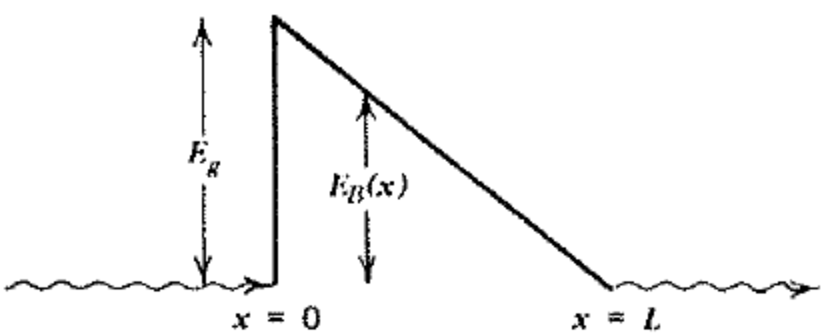
A large number of valence electrons in the p-region are separated by a narrow depletion region from empty allowed states at the same energy in the conduction band of the n-type side.

At a given reverse bias the energy bands in the depletion region are bent steeply downward. **Because of the wave nature of the electron, there is a probability that the valence electron of the p-material can tunnel through the forbidden region and appear at the same energy in the conduction band of the n-material.**

# BTBT - Q.M. Treatment




Energy-band diagram of a reverse-biased junction that has a high dopant concentration on both sides. Tunneling or Zener breakdown is likely in this type of junction.




The height of the energy barrier $E_g$ decreases linearly from $E_g$ at x=0 to zero at x=L, and the average field is $E_{el}=E_g/qL$.

The probability of tunneling across the junction may be approximated by considering tunneling through a triangular barrier.

# Derivation of the Band-to-Band Tunneling

We derive of the tunnel probability from the time independent Schrödinger equation:

$$-\frac{\hbar^2}{2m^*}\frac{d^2\Psi}{dx^2}+V(x)\Psi=E\Psi$$

which can be rewritten as

$$\frac{d^2\Psi}{dx^2}=\frac{2m^*(V-E)}{\hbar^2}\Psi$$

Assuming that *V(x)-E* is independent of position **(i.e. slow varying)** in a section between *x* and *x+dx* this equation can be solved yielding:

$$\Psi(x+dx)=\Psi(x)\exp(-k\,dx)\quad \textbf{with } k=\frac{\sqrt{2m^*[V(x)-E]}}{\hbar}$$

The minus sign is chosen since we assume the particle to move from left to right.
For a slowly varying potential the amplitude of the wave function at *x = L* can be related to the wave function at *x = 0* :

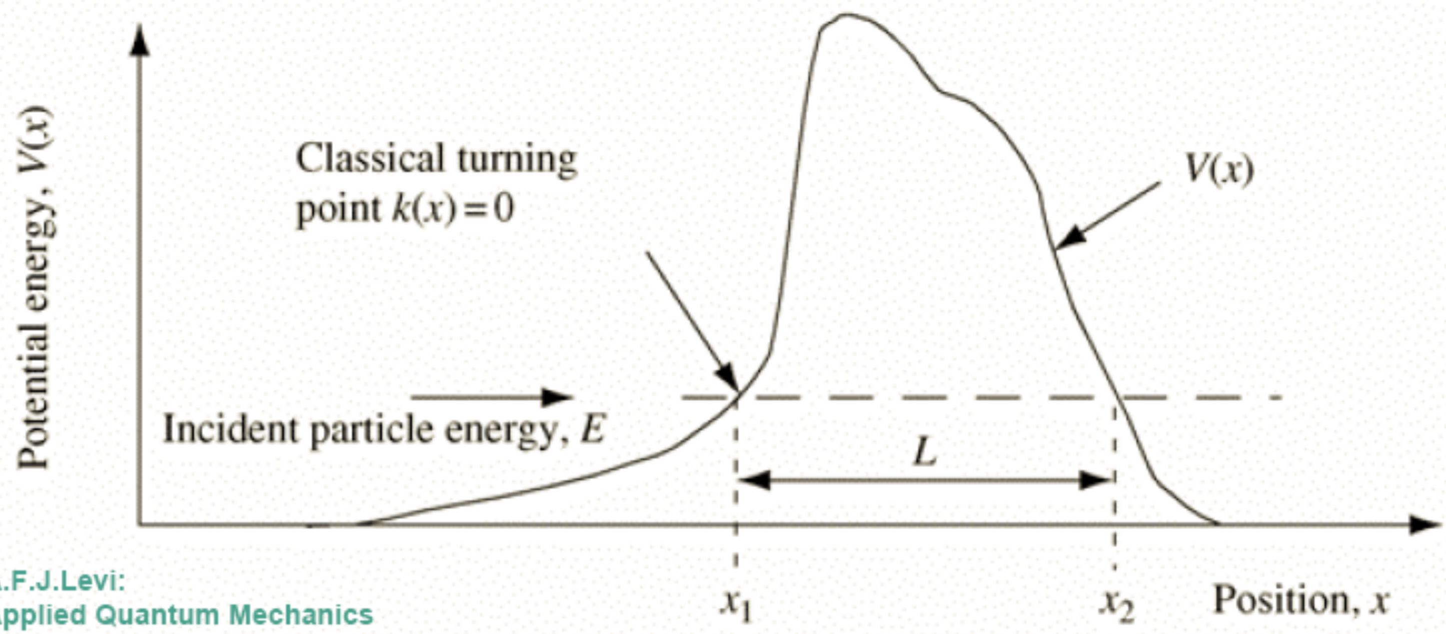
$$\Psi(L)=\Psi(0)\exp\left(-\int_0^L\frac{\sqrt{2m^*[V(x)-E]}}{\hbar}dx\right)$$

This equation is referred to as the WKB (Wigner, Kramers, Brillouin) approximation.

# Tunneling Probability for a barrier with arbitrary shape

$$\Theta = \frac{\Psi(L)\Psi^*(L)}{\Psi(0)\Psi^*(0)} = \quad \cong \exp\left( -2\int_{x_1}^{x_2} \sqrt{\frac{2m(V(x)-E)}{\hbar^2}}\, dx \right)$$

known as Wentzel-Kramers-Brillouin (WKB) approximation




A.F.J.Levi:
Applied Quantum Mechanics

**WKB valid for potential that vary slowly on a scale of the wavelength**

# Derivation of the Band-to-Band Tunneling

From the wavefunction the tunneling probability, $\Theta$, can be calculated for a triangular barrier for which *$V(x)-E = q\phi_B\,(1- x/L)$* the tunneling probability then becomes:

$$\Theta = \frac{\Psi(L)\Psi^*(L)}{\Psi(0)\Psi^*(0)} = \exp\left(-2\frac{\sqrt{2m^*}}{h/2\pi}\int_o^L\sqrt{qE_g(1-\frac{x}{L})}dx\right) \quad \Theta = \exp(-\frac{4}{3}\frac{q\sqrt{2m^*}\cdot E_g^{3/2}}{\hbar\cdot E_{el}})$$

The tunneling current is obtained from the product of the carrier charge, velocity and density. The velocity equals the Richardson velocity, the velocity with which on average the carriers approach the barrier while the carrier density equals the density of available electrons multiplied with the tunneling probability, yielding: (since $E_{el}=E_g/qL$)

$$\Theta = \exp(-\frac{4}{3}\frac{q\sqrt{2m^*}\cdot E_g^{3/2}}{\hbar\cdot E_{el}}) = \exp(-\frac{4}{3}\frac{\sqrt{2m^*}\cdot E_g^{1/2}\cdot L}{\hbar})$$

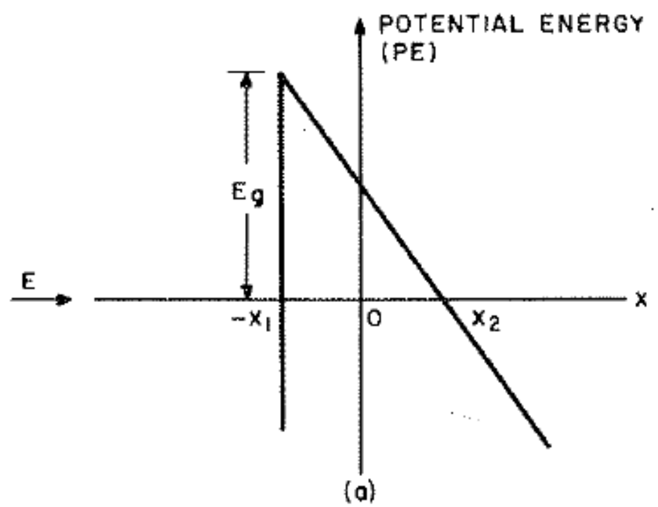
$$E_{el} = E_g/(L\cdot q)$$

**The current is the product of the area, the electron charge, the number of valence-band electrons $N_v$ in the p-region arriving per second at the barrier to "see" empty states across the barrier in the conduction band of the n-material, and the probability that each electron tunnels through the barrier.**

$$I_{BTBT} = q \cdot A \cdot N_v \cdot v_{therm} \cdot \Theta$$

## Exercise for $E_g$(Si, Ge, GaAs)

$I_{BTBT}$=10mA across a junction of $10^{-5}$ $cm^2$, $N_v$(Si)=$10^{22}$ $cm^{-3}$, $v_{therm}$(Si)=$10^7$ cm/sec. What is the tunneling probability $\Theta$, the tunnel thickness L, and electric field $E_{el}$?



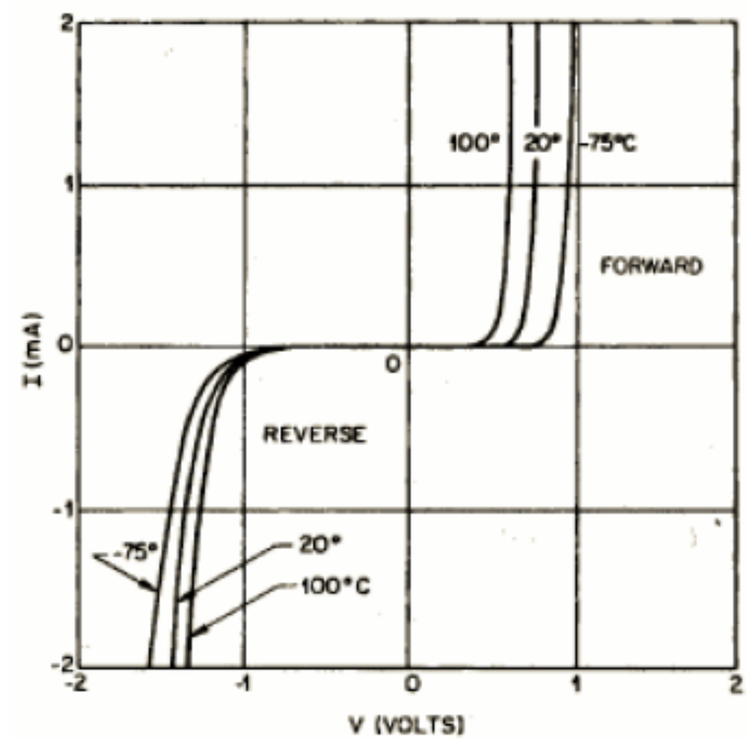
# Zener Breakdown – Temperature Dependence

The tunneling current density can be expressed as :

$$J_T = \frac{\sqrt{2m^*}\, q^3 E V}{4\pi^2 \hbar^2 E_g^{1/2}} \exp\left( -\frac{4\sqrt{2m^* E_g^{3/2}}}{3\hbar E q} \right)$$

Where $E$ is the electric field at the junction and $V$ is the applied voltage

- Since the energy bandgap *Eg* ***decreases*** with ***increasing*** temperature, the Zener breakdown voltage ***decreases*** with ***increasing*** temperature.

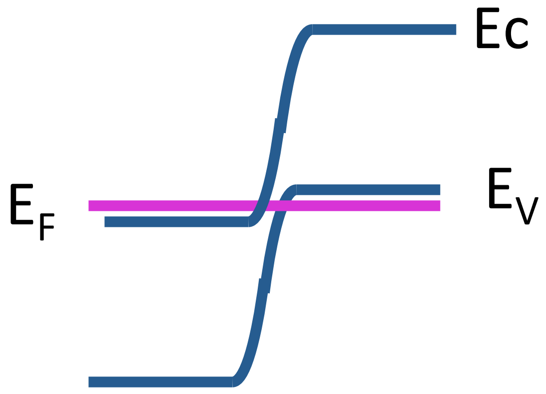



**Opposite temperature effect to the impact ionization effect.**

**Measurements at different temperatures can reveal which mechanism is at work!**

# (Esaki 1958) Tunnel Diodes (TD)

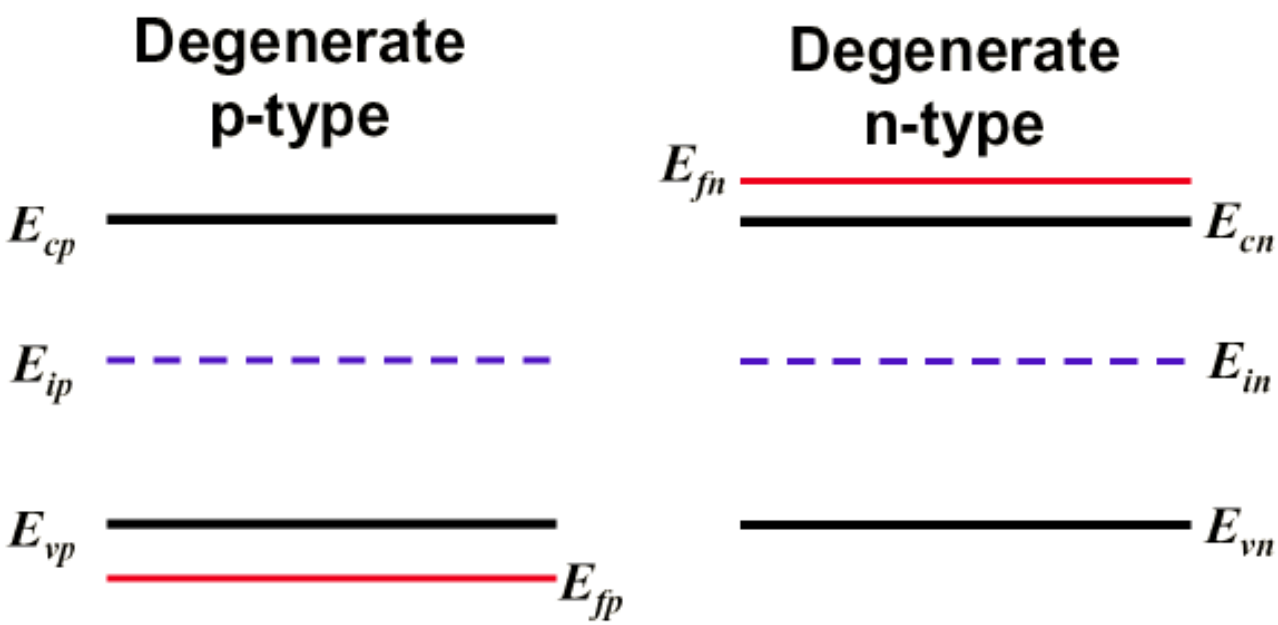
Nobel prize




- Simplest tunneling device
- Heavily-doped pn junction
  - Leads to overlap of conduction and valence bands
- Carriers are able to tunnel band to band
- Tunneling goes exponentially with tunneling distance
  - Requires junction to be abrupt to assure a thin enough tunneling barrier

# tunnel diode

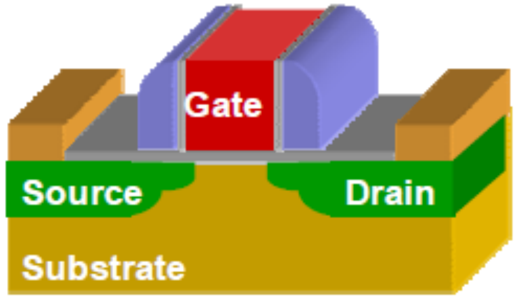
- pn-diode with very high (degeneration) donor and acceptor concentrations
- very sharp transition in metallic junction
- very thin depletion region (5-10 nm), which allows tunneling between conduction band of n-type and valence band of p-type semiconductor



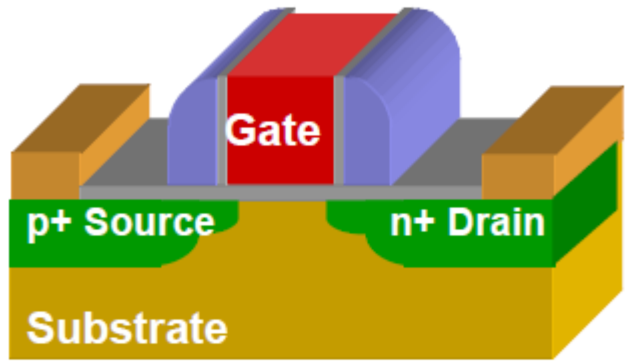

**Amount of degeneracy is typically a few kT; depletion layer less than 100A**

# Replacement of the Conventional MOSFET by a Tunneling MOSFET or TFET

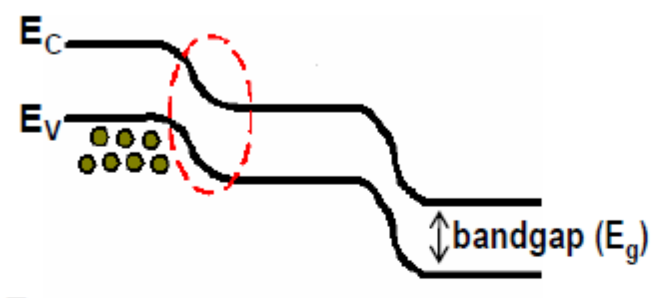
# MOSFET

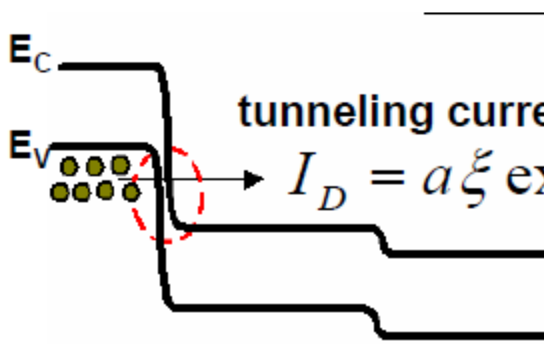



# Tunnel FET



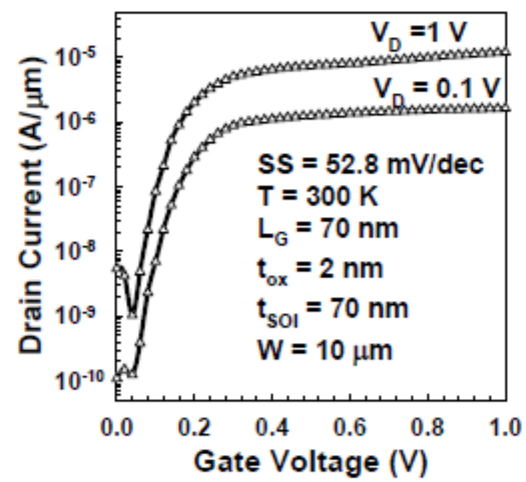

**OFF STATE:**



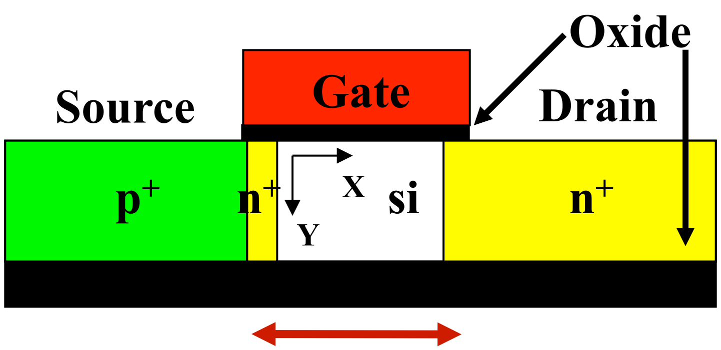




tunneling current

$$I_D = a\xi \exp(-b/\xi)$$

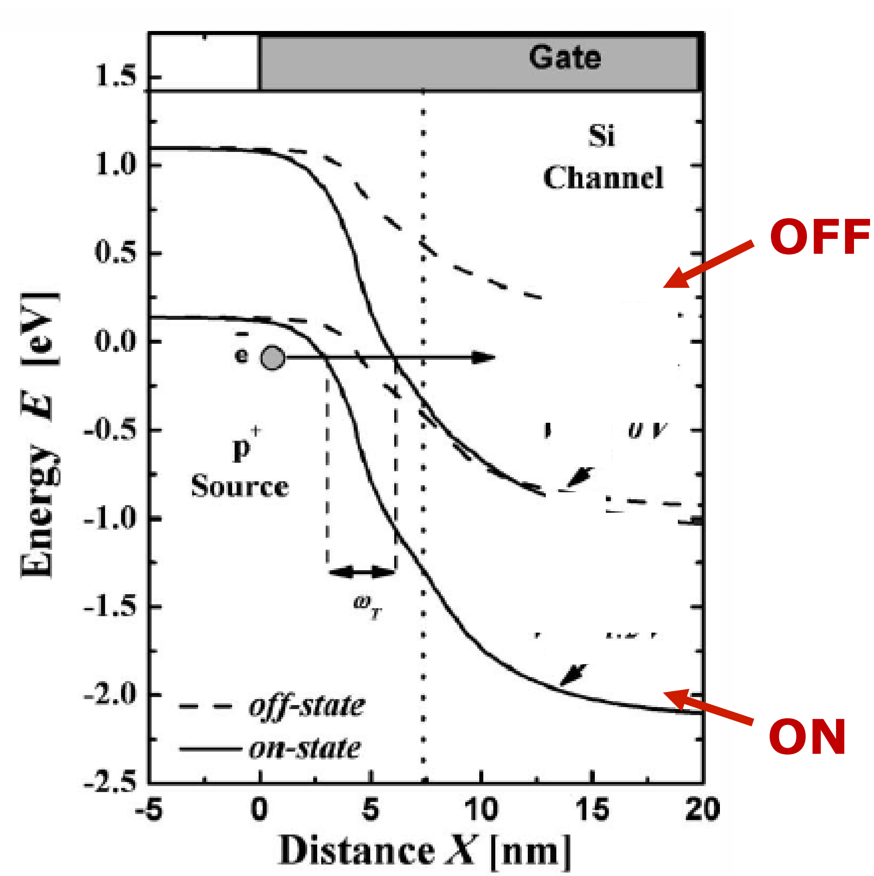


# Lateral SOI TFET




$L_G$ = 30nm

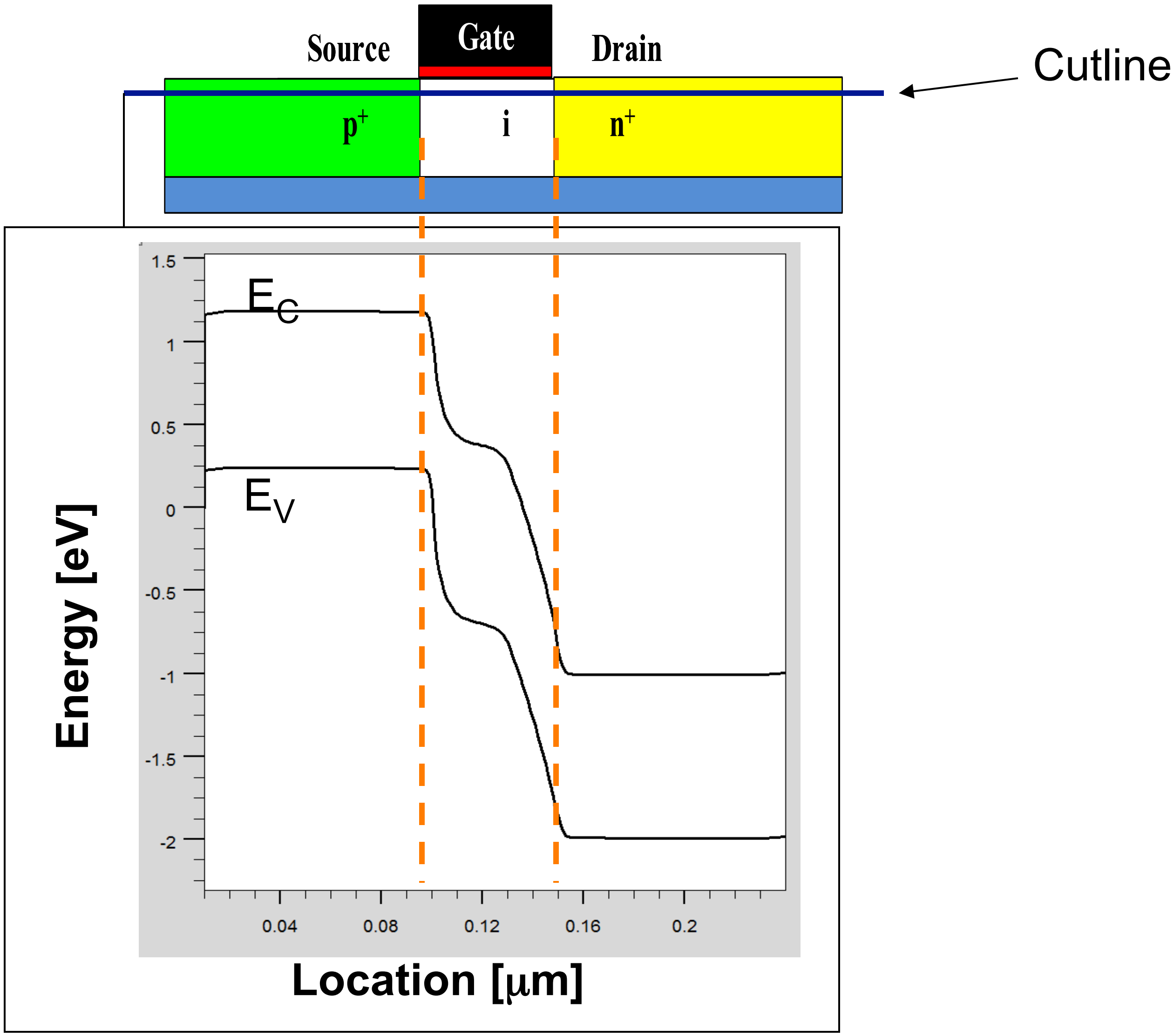
< N-type TFET >




< Band-to-band tunneling >

# TFET – the operation principle

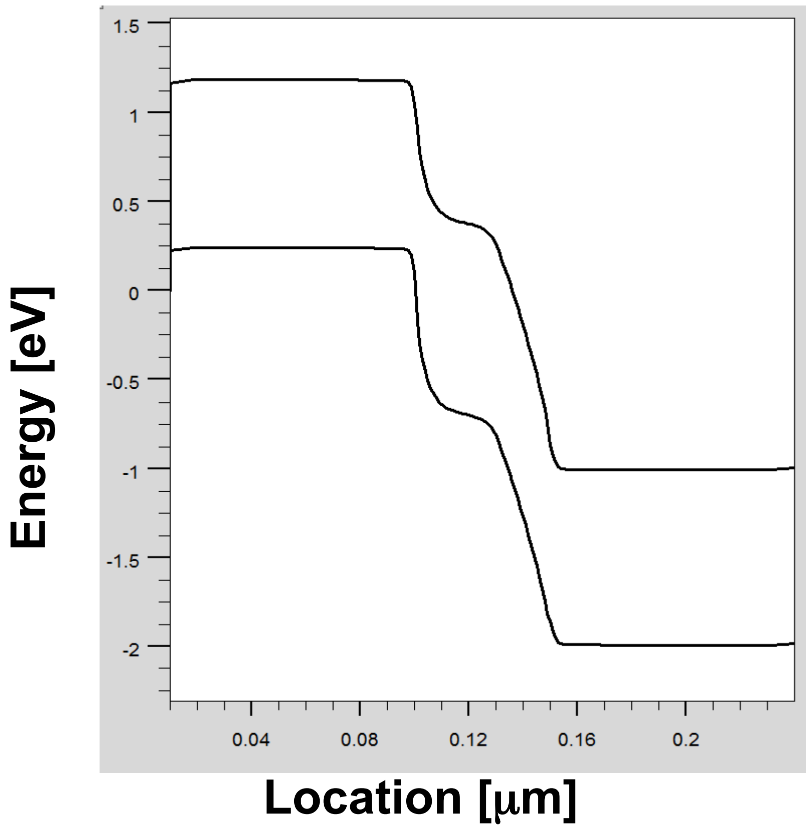
$V_G$=0V
$V_D$=1V



# TFET – the operation principle

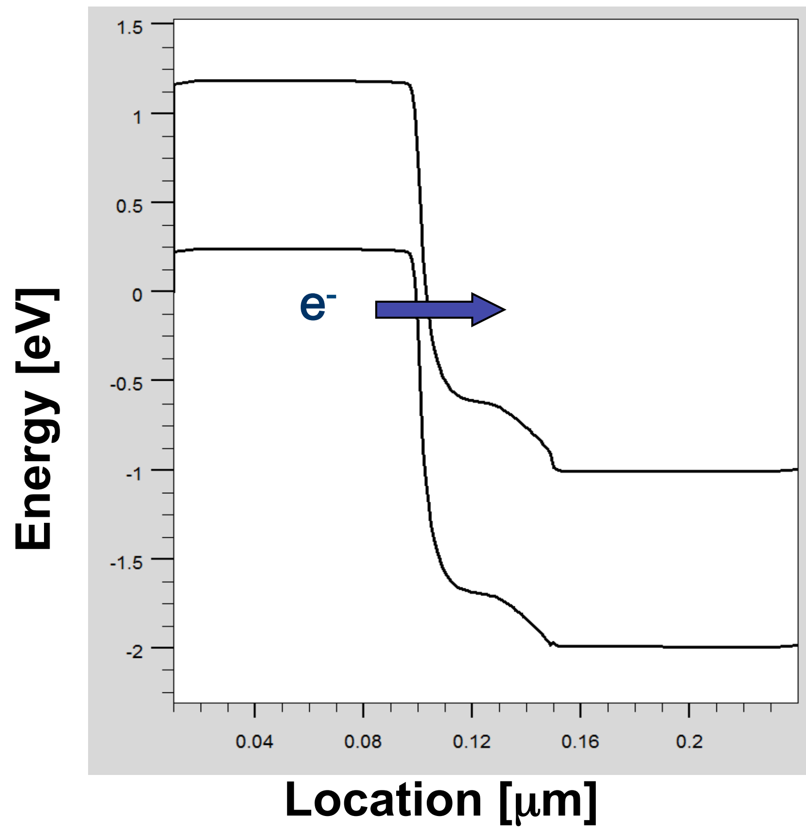
TFET
**OFF** state

$V_G$=0V
$V_D$=1V




TFET
**ON state**

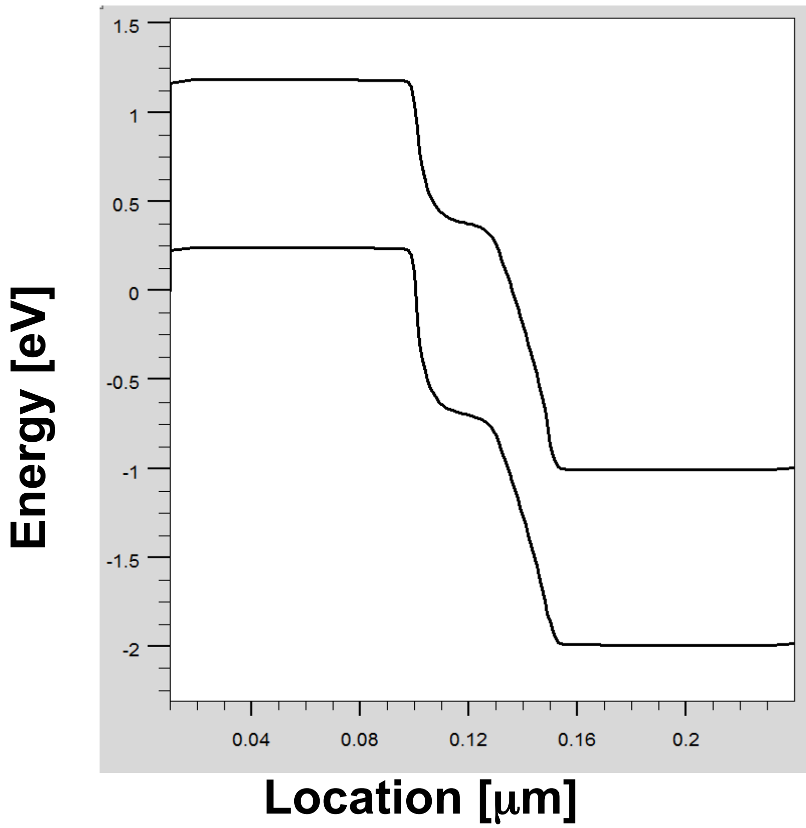
$V_G$=1V
$V_D$=1V



# TFET – the operation principle

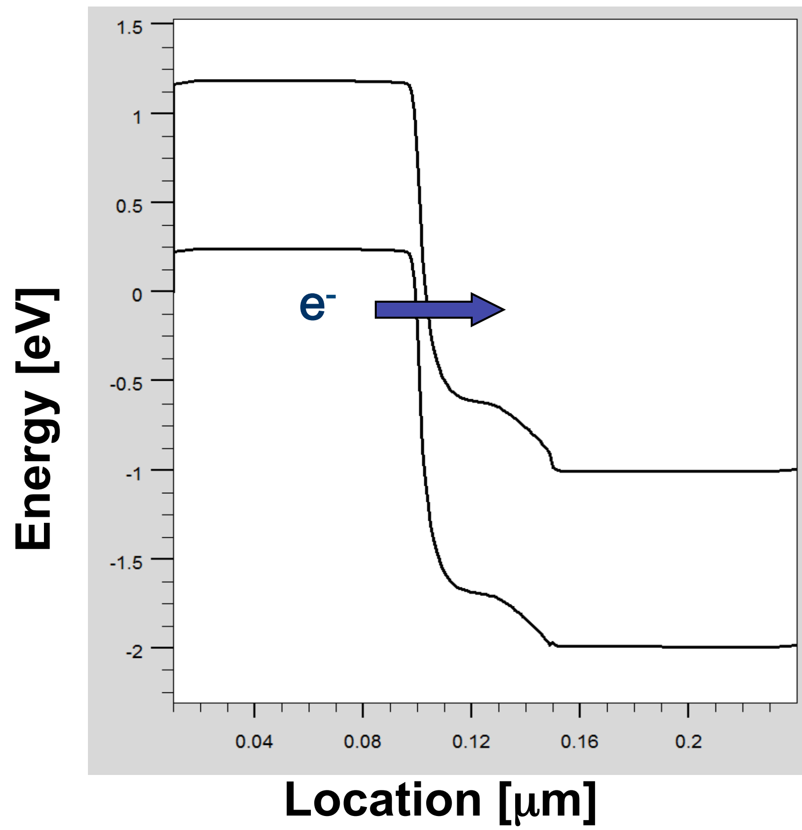
TFET
**OFF** state

$V_G$=0V
$V_D$=1V




TFET
**ON state**

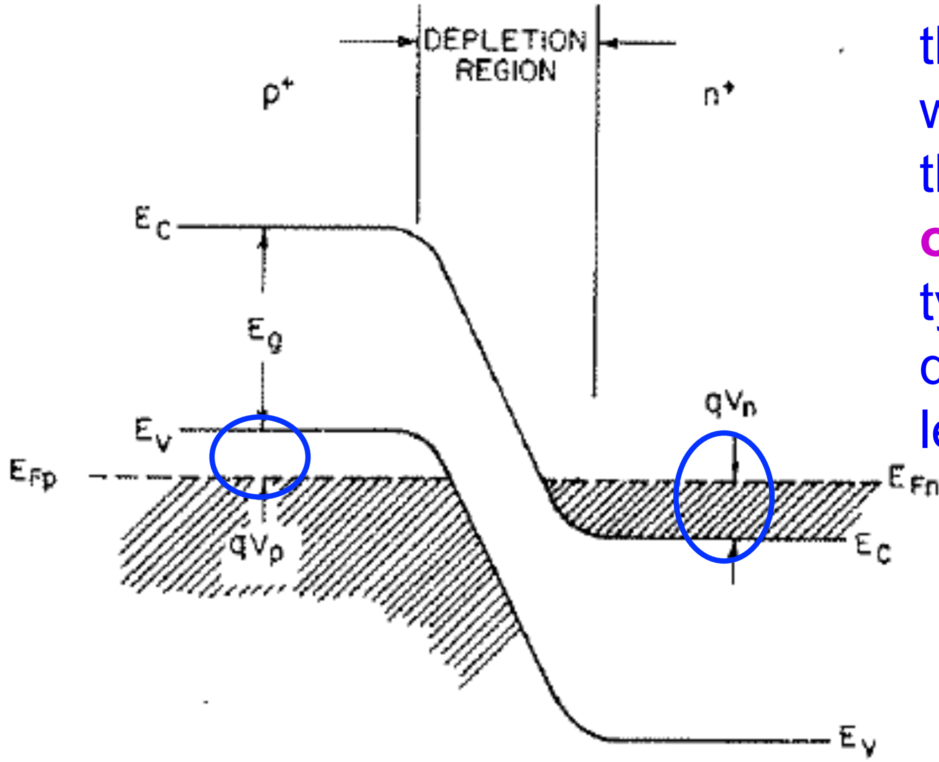
$V_G$=1V
$V_D$=1V



# Back to Esaki Tunnel Diode

p and n-side are degenerate (very highly doped);

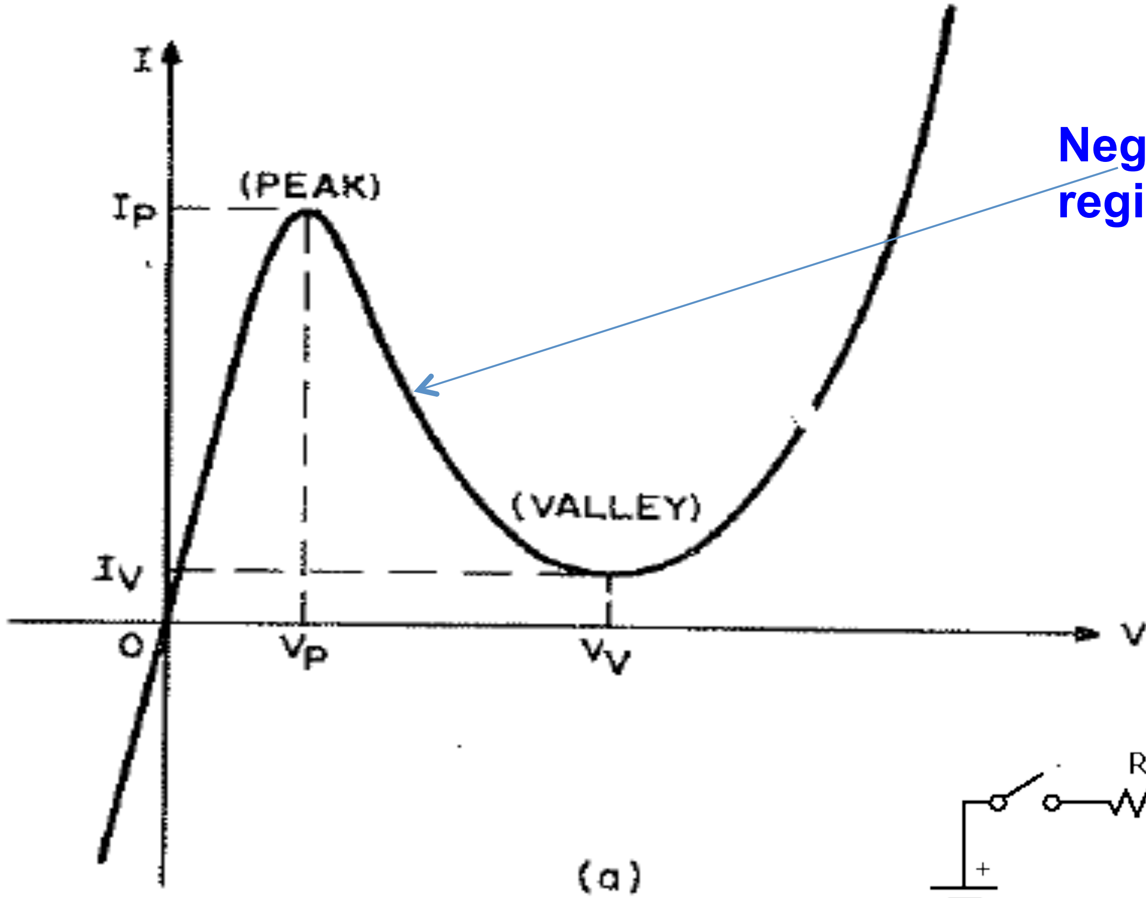
therefore $E_F$ is located within the allowed bands themselves; the **measure of degeneracy** $V_p$ and $V_n$ is typically a few kT. The depletion width is 100 A or less.



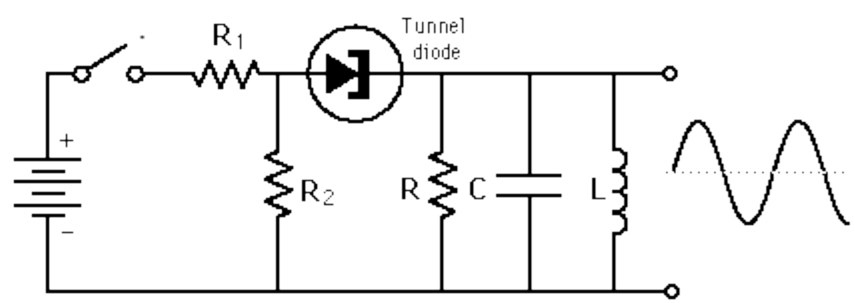

Energy-band diagram of a tunnel diode in thermal equilibrium. $V_p$ and $V_n$ are the degeneracies on the $p$-side and $n$-side, respectively.

## Measure of degeneracy $V_p$ and $V_n$

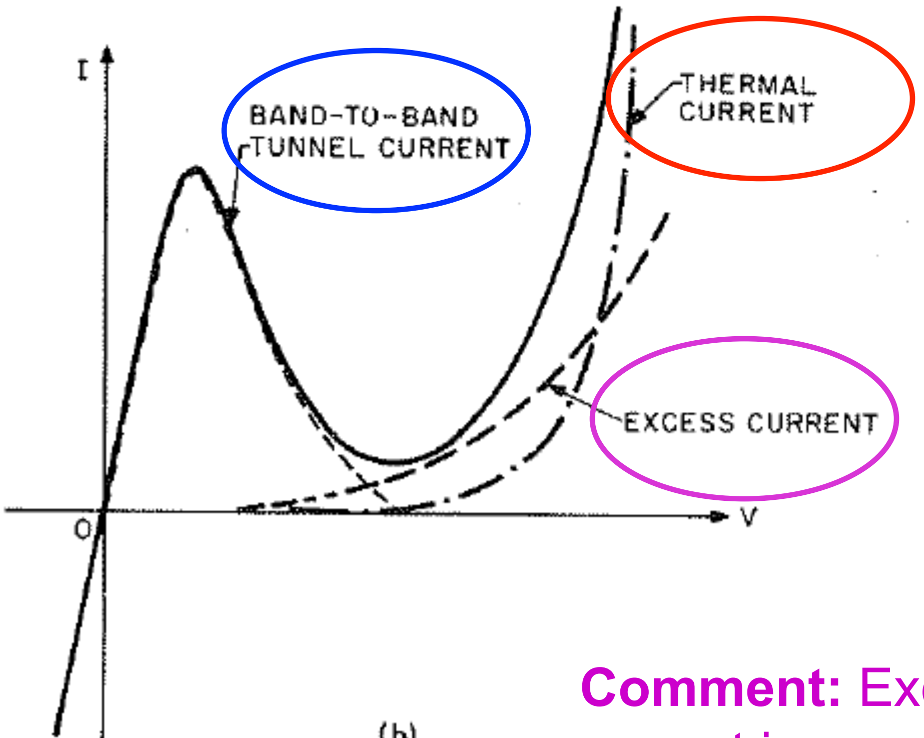
# I-V Characteristics of a Tunnel Diode




**Negative resistance region**




Oscillator design



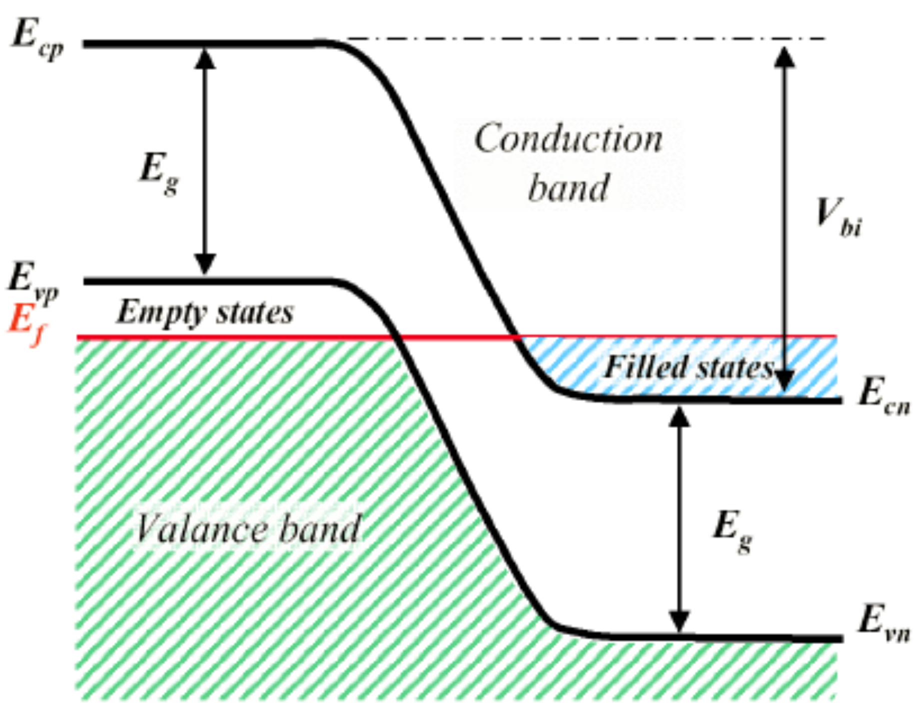
(b)

# Three current components:

- BTBT current
- excess current
- thermal current

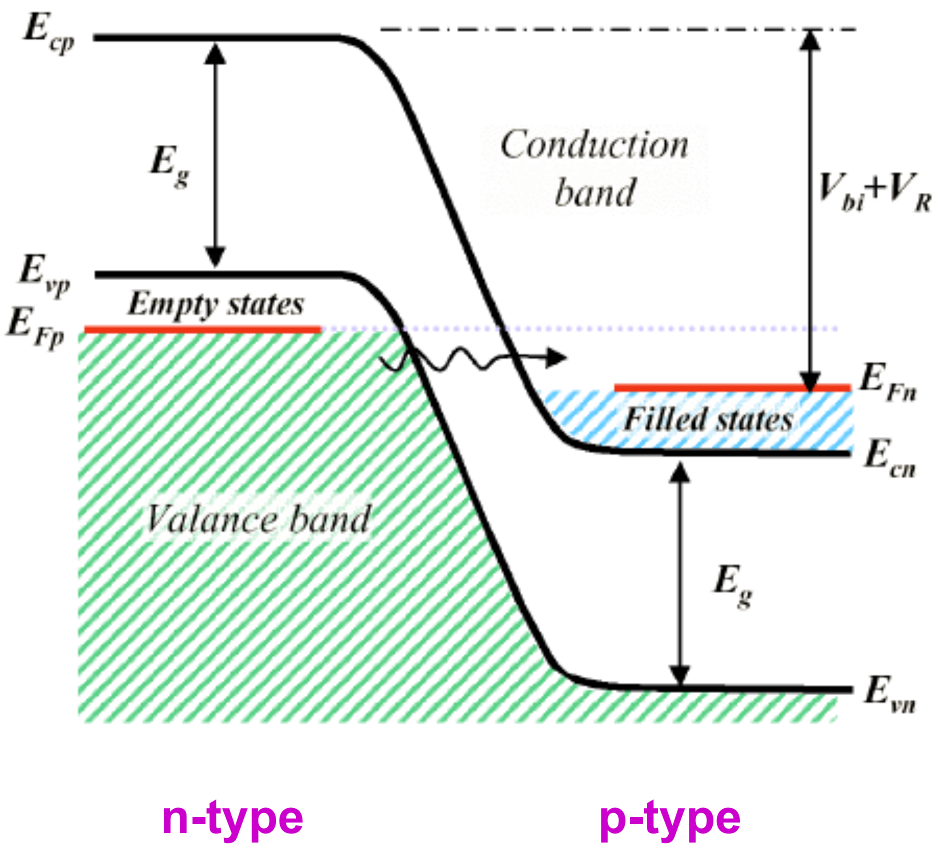
**Comment:** Excess current is if you will a parasitic current in excess of the normal diode current. The excess current is mainly due to carrier tunneling by way of defects (energy states within the forbidden energy gap Eg.

# Tunnel Diode V=0



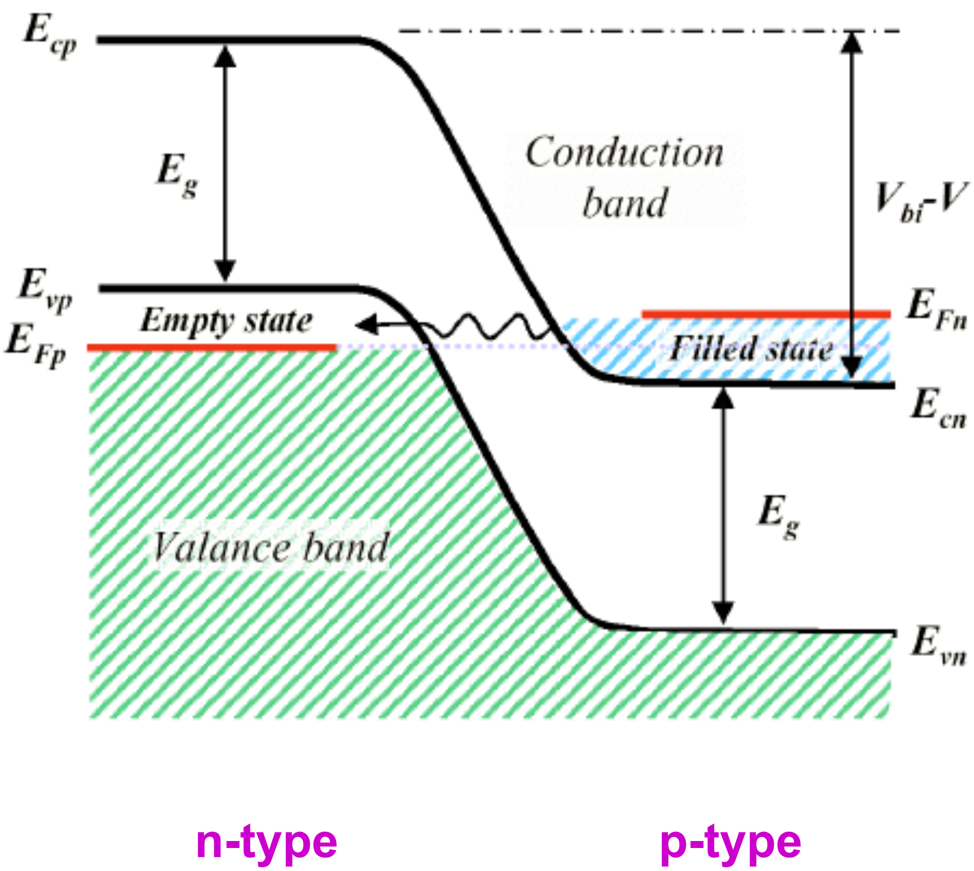

- no external voltage applied to the junction
- Fermi levels in p and n type semiconductor are on the same level
- the tunnel currents from p to n and from n to p are equal

# tunnel diode: reverse polarized



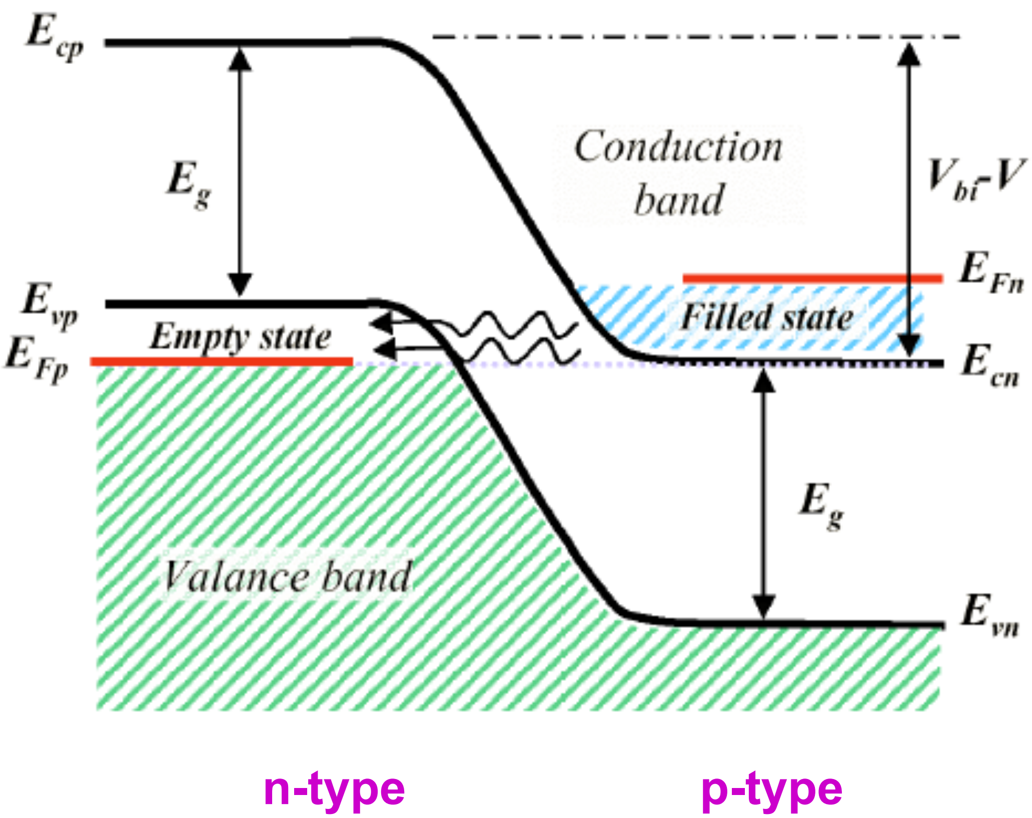

- reverse voltage $V_R$ applied
- Fermi level in p-type semiconductor higher than in n-type one
- tunneling of electrons from the p-type semiconductor valence band to the n-type semiconductor conduction band
- EVB-tunneling
- current increases with increasing voltage

# tunnel diode: forward polarized



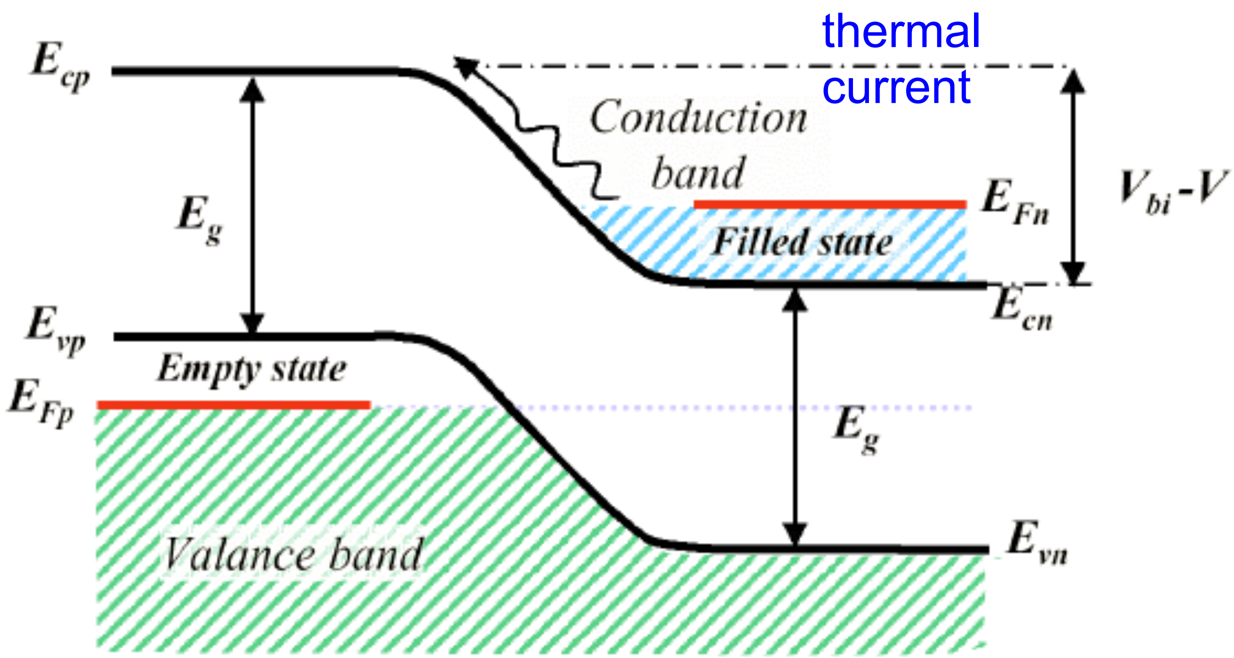

- forward voltage $V_F$ is applied
- Fermi level in p-type semiconductor is below the Fermi level in n-type one
- tunneling of electrons from the n-type semiconductor conduction band to the p-type semiconductor valence band
- ECB-tunneling
- current increases with increasing voltage

# tunnel diode: maximum




- increasing the forward voltage $V_F$ the maximum current $I_P$ is reached for $V_P$
- Fermi level in p-type semiconductor is on the same level with the conduction band edge of the n-type material
- continuing the voltage increases over this value causes reduction of the current because the overlap of the p-type valence band and n-type conduction band decreases

# tunnel diode: forward without tunneling




- forward voltage $V_F$ is further increased
- valence band of p-type semiconductor doesn't overlap with conduction band of the n-type semiconductor: no direct tunneling current
- small current flows due to trap mechanism tunneling
- at sufficient voltage value the classical forward current starts to flow and is equal to the tunnel current for $V_V$

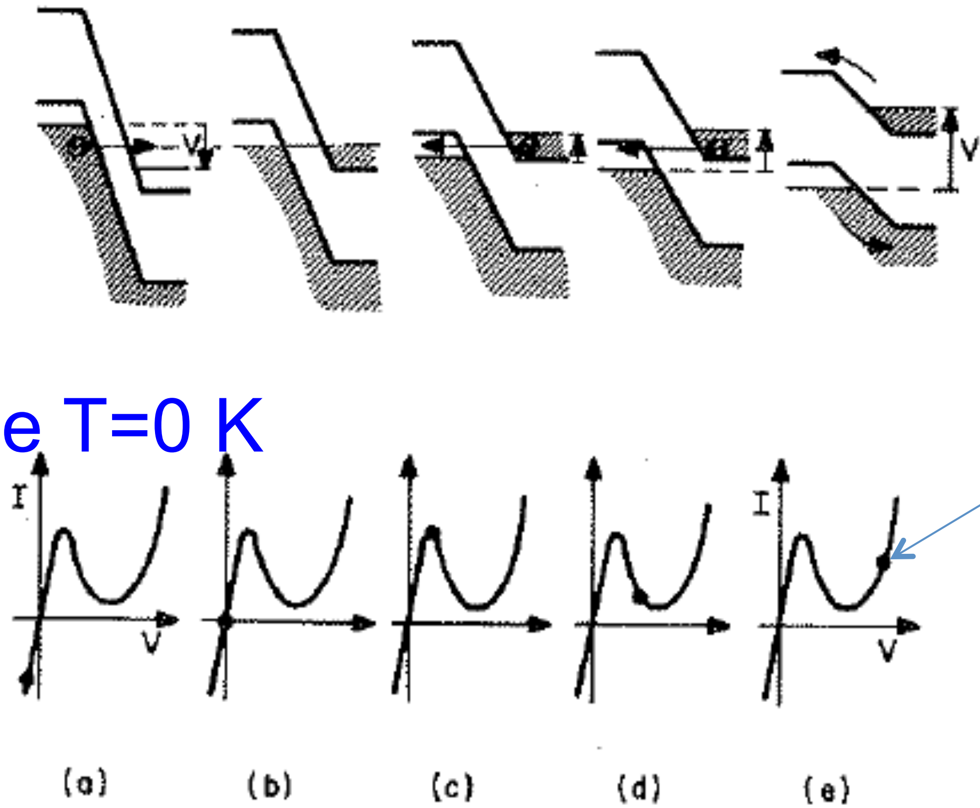
## Tunneling conditions:

**(1) occupied energy states exist on the side from which the electron tunnels**

**(2) unoccupied energy states exist at the same energy level as in (1) on the side to which the electron tunnels**

**(3) The tunneling potential barrier height is low and the barrier width is small enough**

**(4) the momentum is conserved in the tunneling process**




$$V_n = (E_{Fn} - E_c)/q$$

$$V_p = (E_v - E_{Fp})/q$$

**Degeneracy**

Assume T=0 K

$$V = V_n + V_p$$

**Fig. 3** Simplified energy-band diagrams of tunnel diode at (*a*) reverse bias; (*b*) thermal equilibrium, zero bias; (*c*) forward bias such that peak current is obtained; (*d*) forward bias such that valley current is approached; and (*e*) forward bias with thermal current flowing. (After Hall, Ref. 3.)

# **Comment:** Excess current

For larger forward biases we would have the normal diode forward currents caused by forward injection of minority carriers. In practice, the actual current is considerable and in excess of the normal diode current (hence excess current).

The excess current is mainly due to carrier tunneling by way of energy states within the forbidden gap.

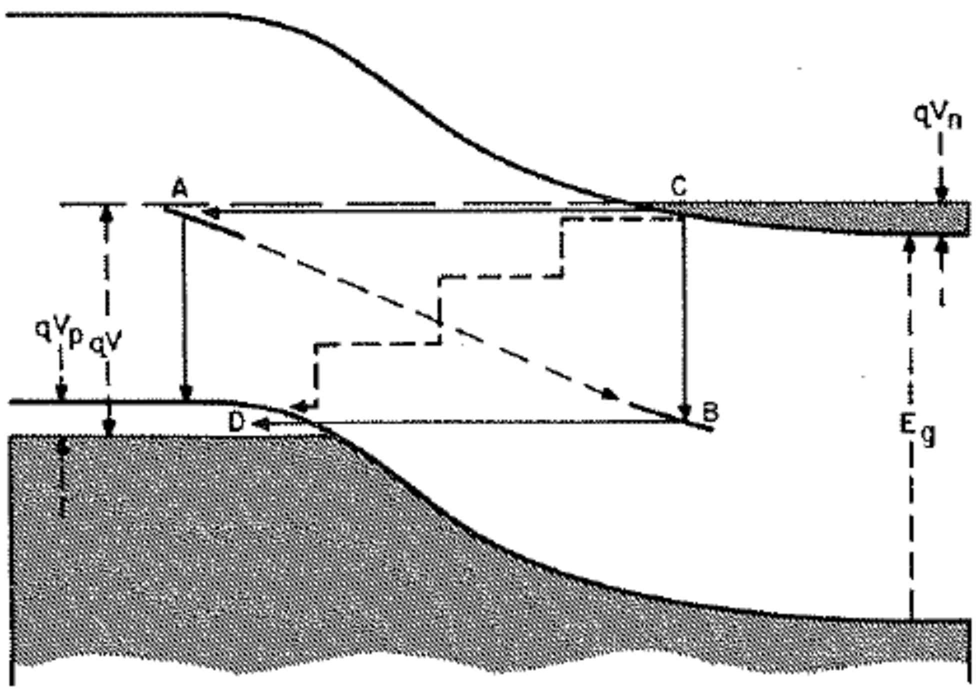


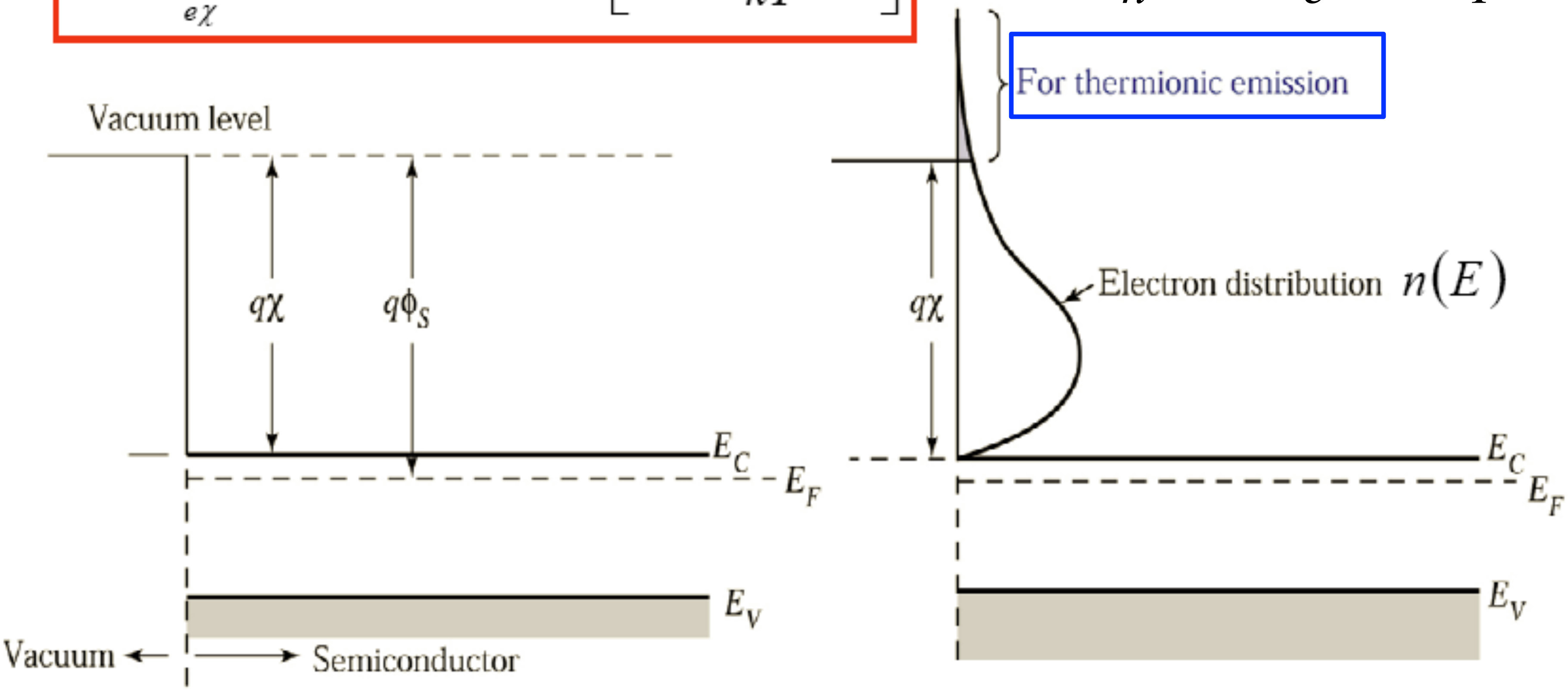

Fig. 9 Band diagram illustrating mechanisms of tunneling via states in the forbidden gap for the excess current flow. (After Chynoweth, Feldmann, and Logan, Ref. 13.)

# **Comment:** Thermionic current

## thermionic emission from semiconductor

$$\boxed{n_{\mathrm{th}} = \int_{e\chi}^{\infty} n(E)\,dE = N_{\mathrm{C}} \exp\left[-\frac{e(\chi - V_{\mathrm{n}})}{kT}\right]}$$

$$qV_n = E_c - E_F$$

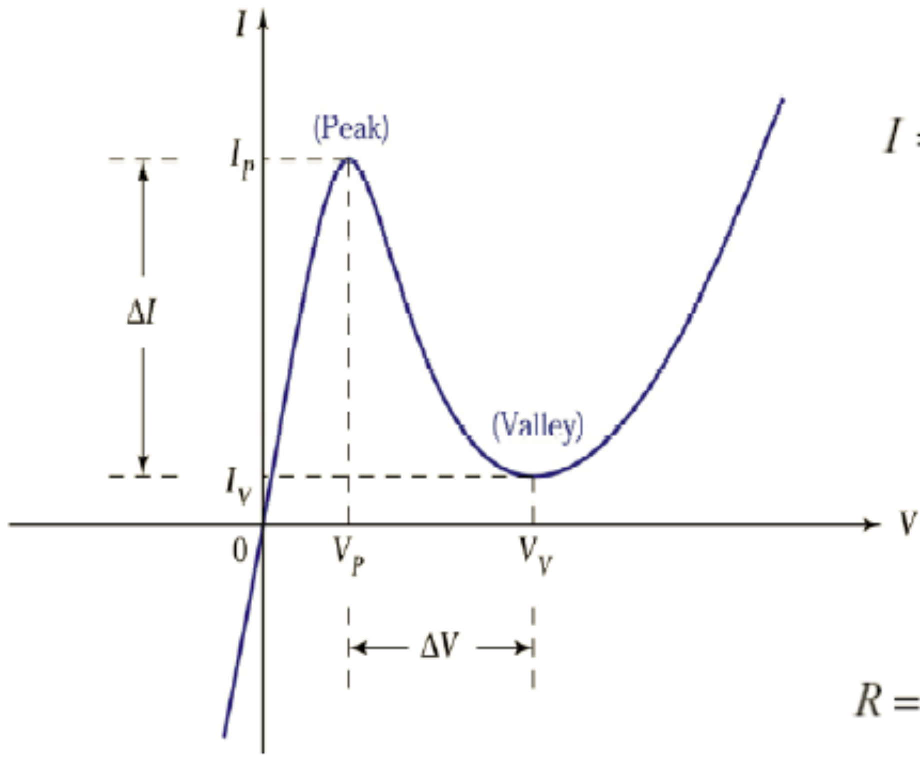


# Comment: Thermionic current

The thermal current is the familiar minority-carrier injection current in pn junctions:

$$J_{th} = J_0(\exp(qV / kT) - 1)$$

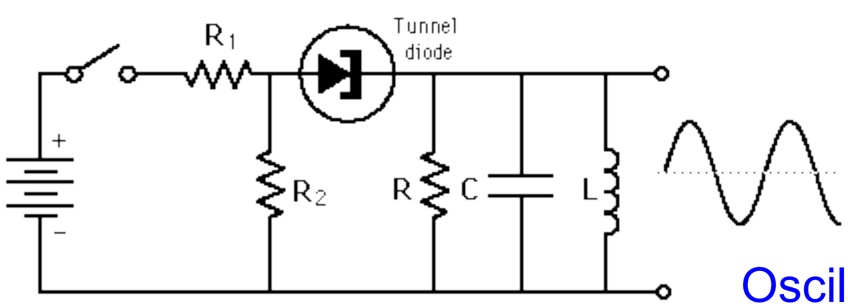
# Useful Analytic Expression for **tunnel diode: I-V characteristics**




$$I = \underbrace{I_p\left(\frac{V}{V_p}\right)\exp\left(1-\frac{V}{V_p}\right)}_{\textit{Tunneling current}} + \underbrace{I_o\exp\left(\frac{qV}{kT}\right)}_{\textit{injection current}}$$

Resistance

$$R = \left(\frac{dI}{dV}\right)^{-1} = -\left[\left(\frac{V}{V_p}-1\right)\frac{I_p}{V_p}\exp\left(1-\frac{V}{V_p}\right)\right]^{-1}$$




Oscillator design

# Comment: Tunneling time (a research topic)

For a classical particle with a velocity v it takes t to traverse a distance W:

$$\tau = W / v$$

Tunneling phenomenon is not governed by this classical rule. Q.M. tunneling time depends rather on the quantum transition probability $\Theta$ per unit time and is proportional to:

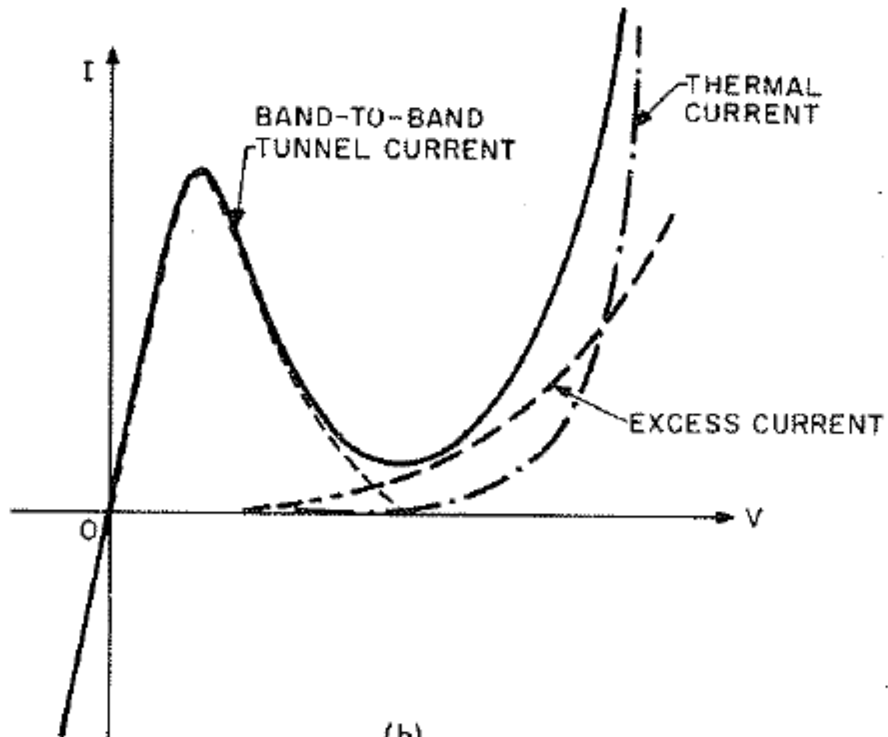
$$\Theta \propto \exp(-2k_{av}(0)W)$$

Where $k_{av}(0)$ is the average value of the momentum along the tunneling path.
The transit time $\tau$ is then by reciprocation:

$$\tau = \frac{1}{\Theta} \propto \exp(2k_{av}(0)W)$$

**This tunneling time is very short, allowing high frequency applications of the tunneling diodes.**

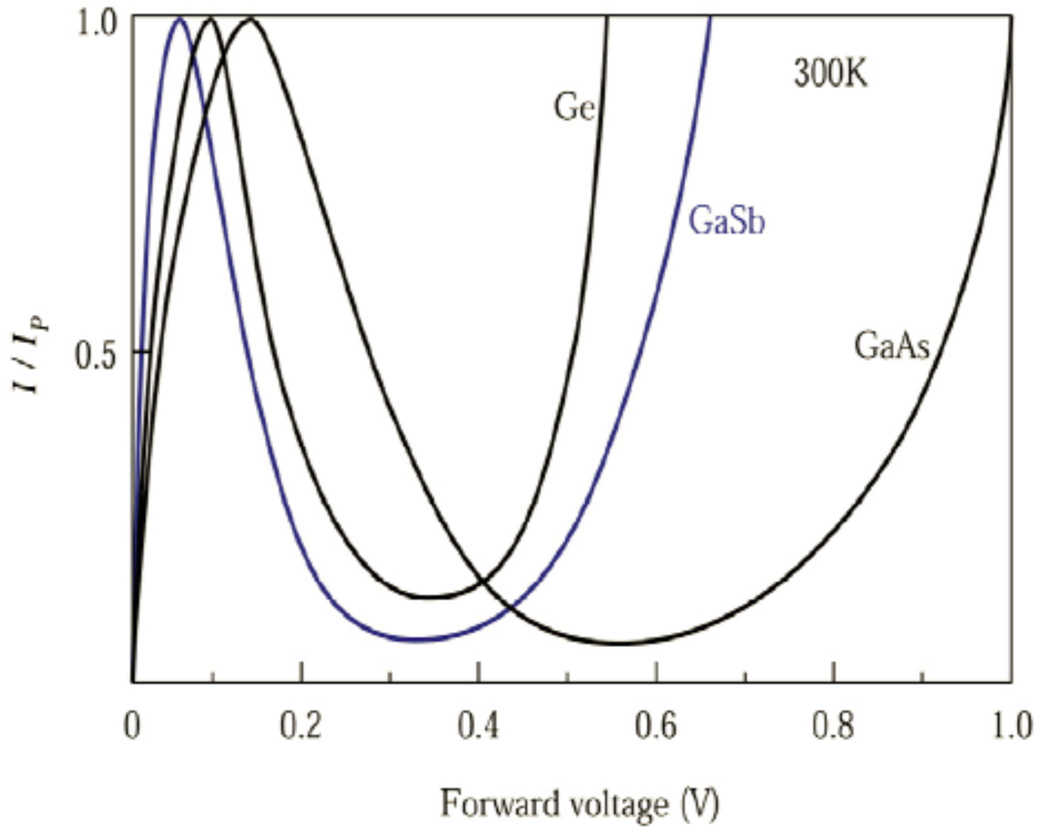
# Can we engineer the tunnel diode characteristics?



(b)

# By using different materials:
# Ge, GaSb, GaAS

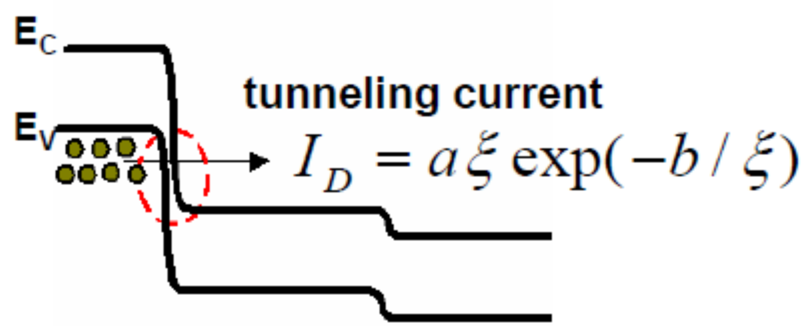
## tunnel diode: applications



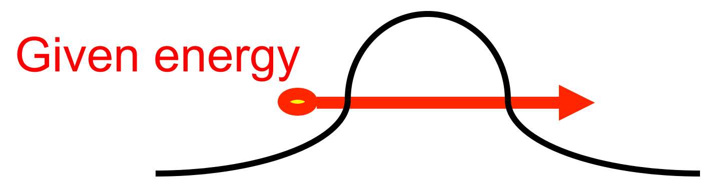

- can be applied in THz range, because the tunneling processes are very fast
- negative dynamic resistance makes tunnel diode useful for oscillators

# Tunneling:
# direct and indirect tunneling

During tunneling energy is conserved



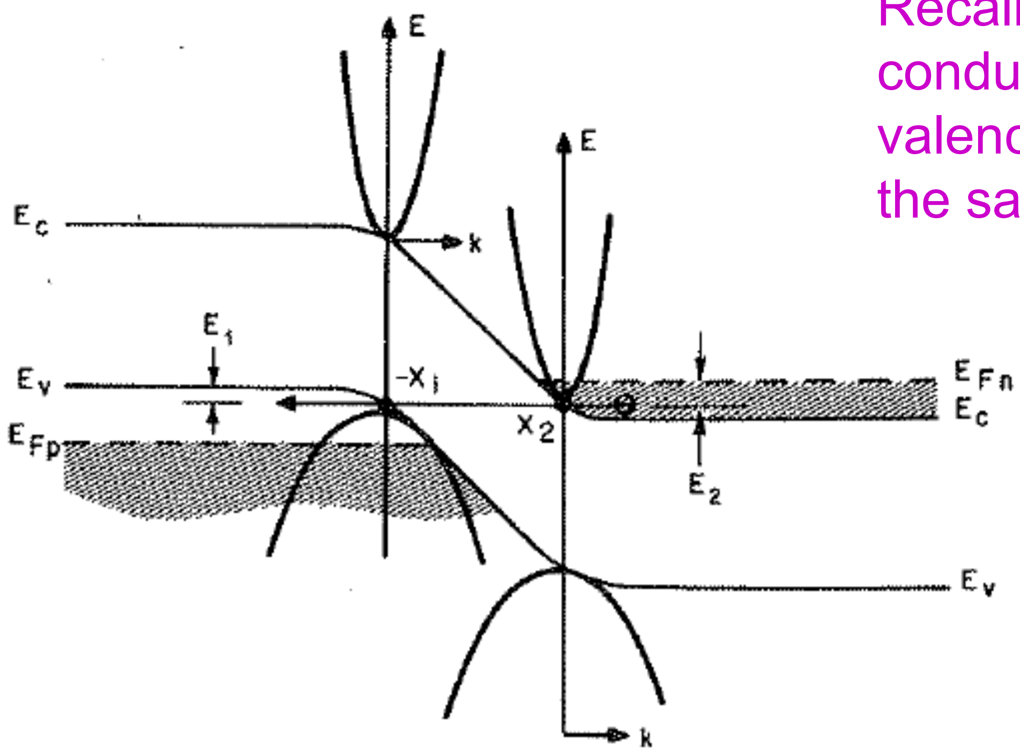

But momentum of the system has also to be conserved.




In solids (and semiconductors) the band diagram
Shows that electrons at the same energy may have different
wave vectors k (p=k(h/2p).

This means that if an electron tunnels from one state E(k) to another state E' (k')
Both energy and momentum have to be conserved.

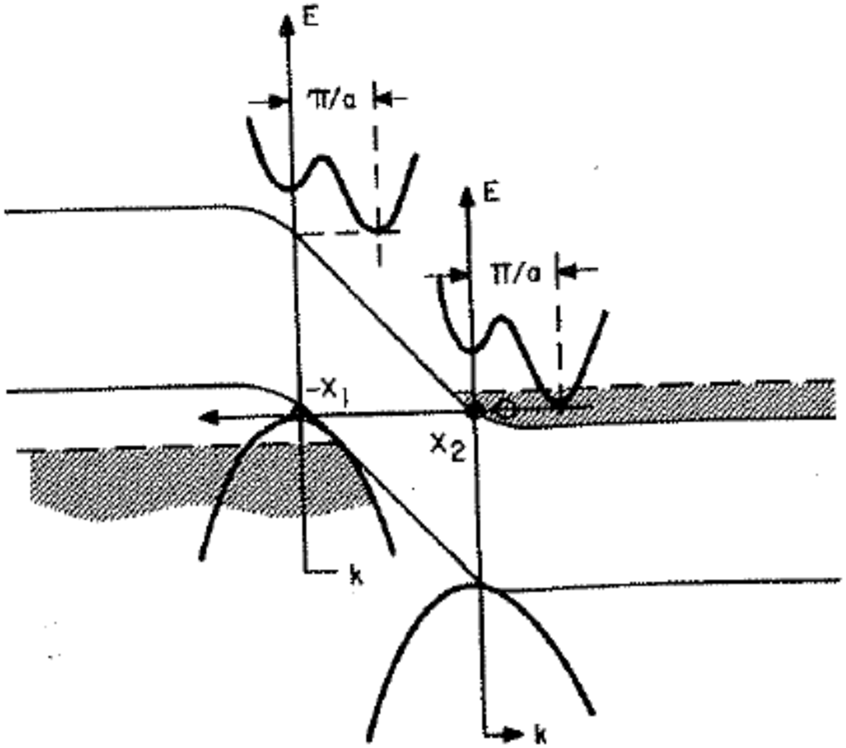
# Direct Tunneling Process: GaAs and GaSb




Recall for direct tunneling the conduction-band minimum and the valence-band maximum must have the same momentum.

E-k relation and E-x superimposed

# Indirect Tunneling Process: Si and Ge




**E-k relation and E-x superimposed**

Here $$k_{\min} \neq k_{\max}$$

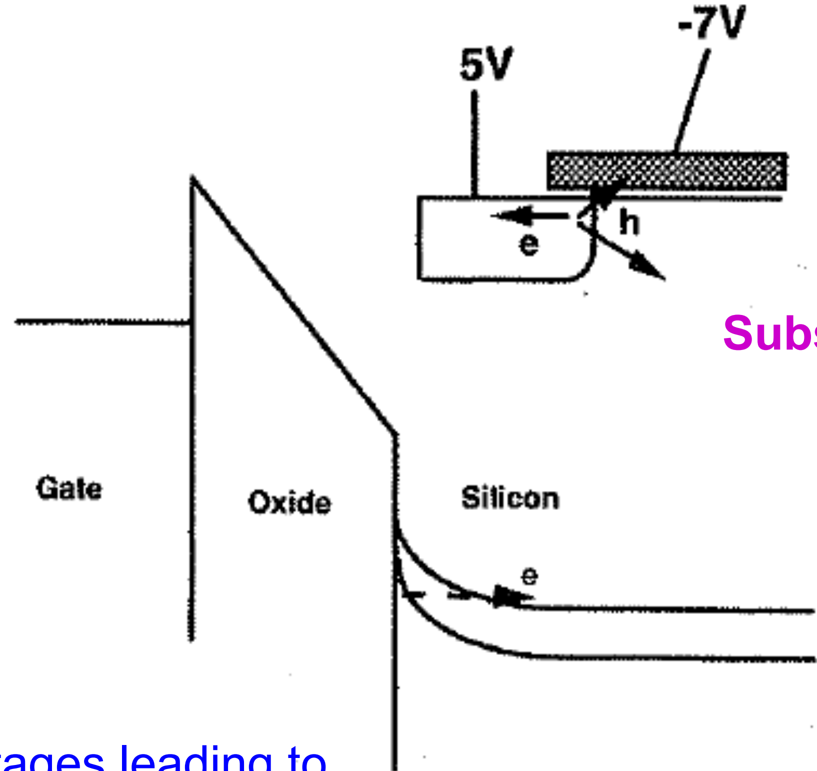
To conserve momentum, the difference between the conduction-band minimum and the valence-band maximum must be supplied by a scattering agents such as phonons or impurities

(phonon-assisted tunneling).

In general, probability for indirect tunneling is much lower than the probability for direct tunneling.

# Gated Diode

(a pn junction with a gate electrode above the metallurgical juntion)




**Substrate grounded**

**Quasi Fermi level of electrons and holes separated by 5 V**

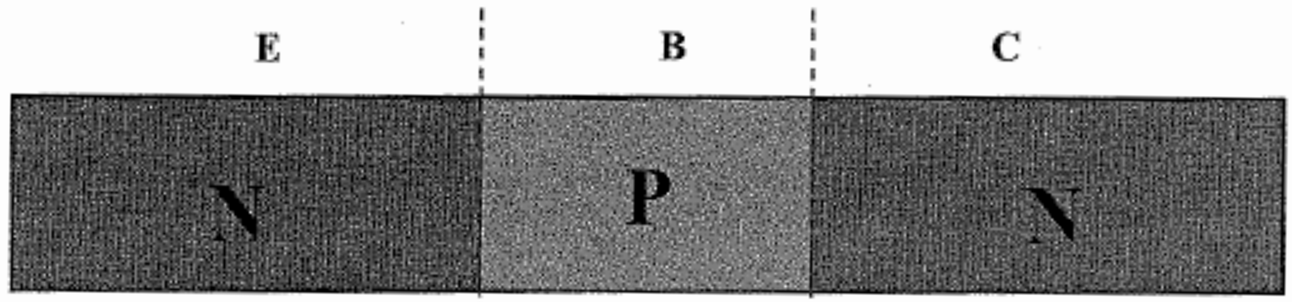
Gate and drain voltages leading to generation of e-h pairs by

**band-to-band tunneling**

**No avalanche**

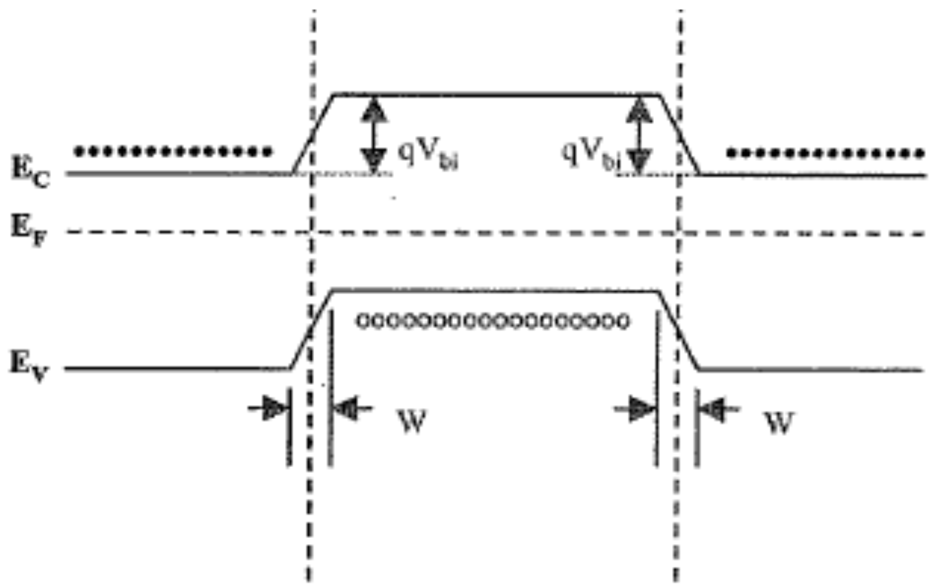
# Construction of a MOSFET

## Exercise: Back-to-Back PN Junctions



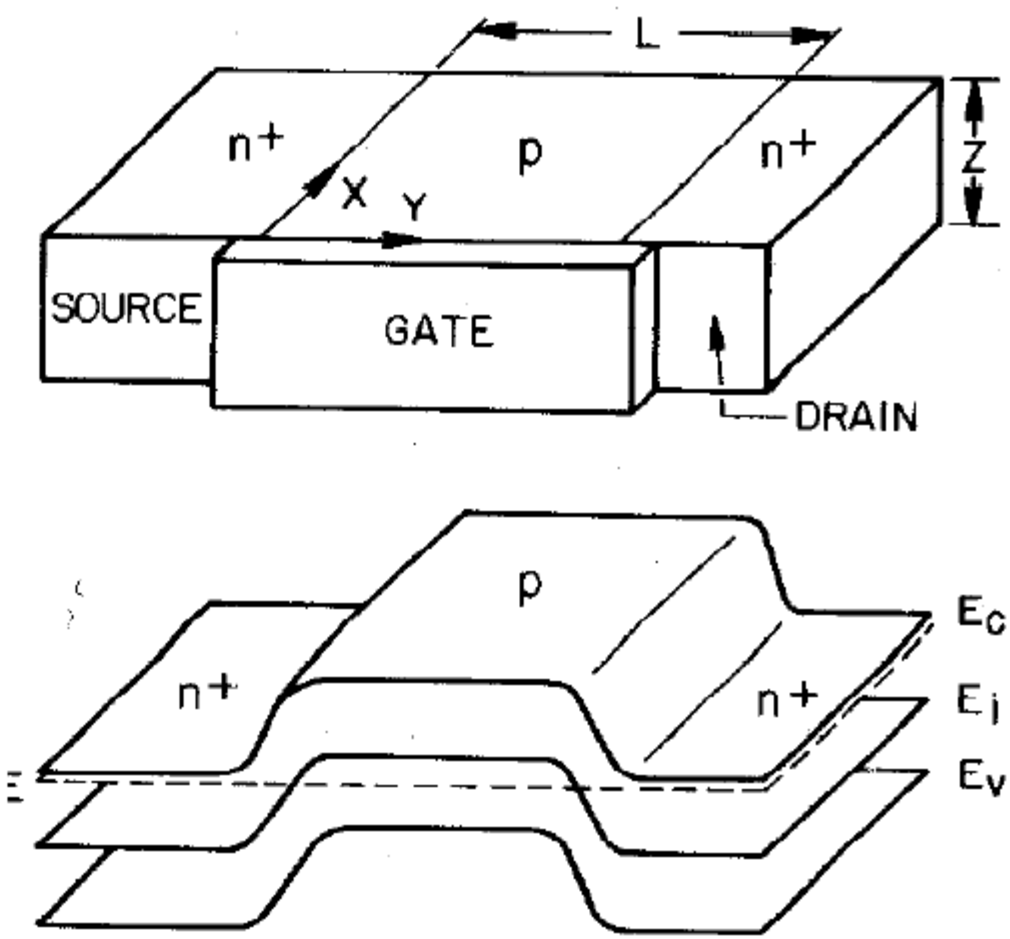

In the space below, draw the band diagram for this structure at equilibrium.

**Equilibrium**



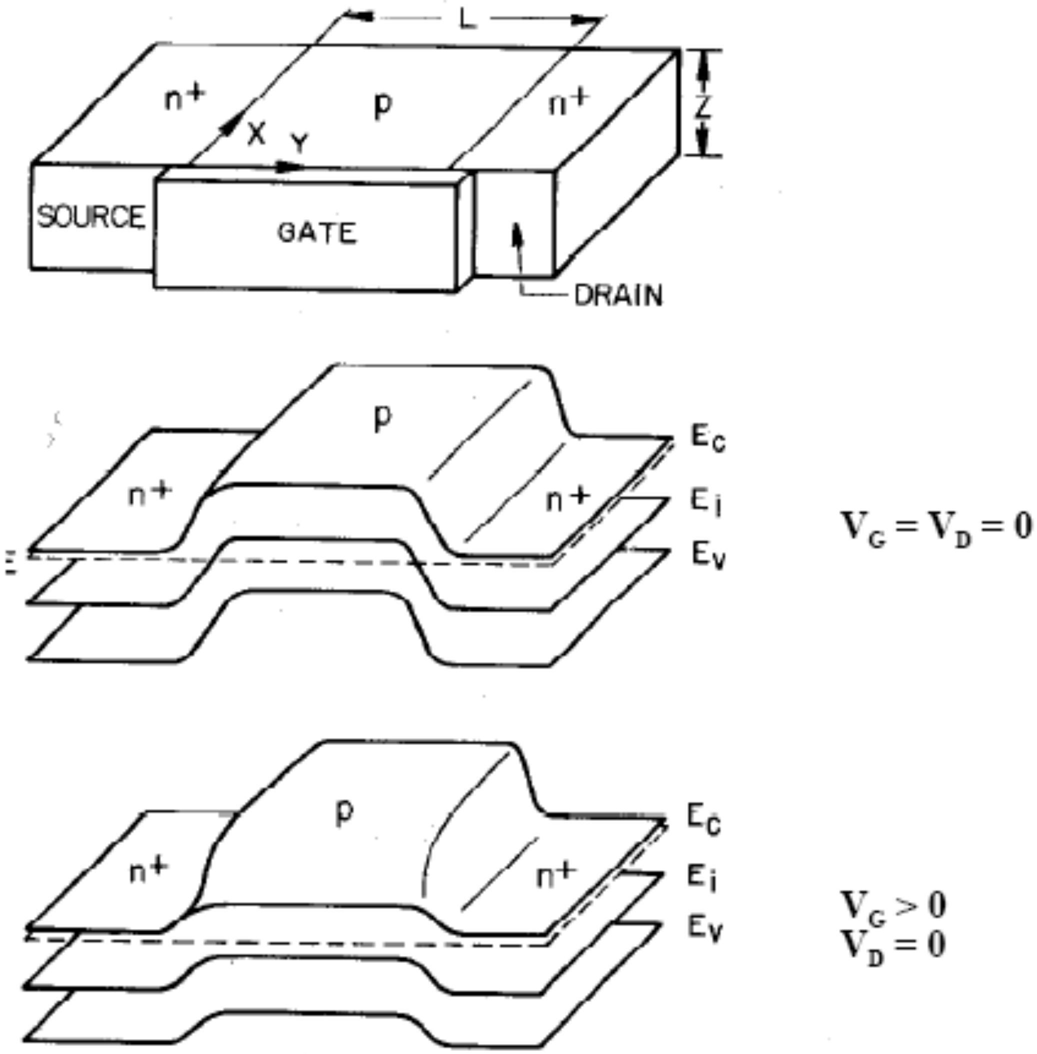
# Back-to-back pn-junction + MOS Capacitor

# = MOSFET




$V_G = V_D = 0$

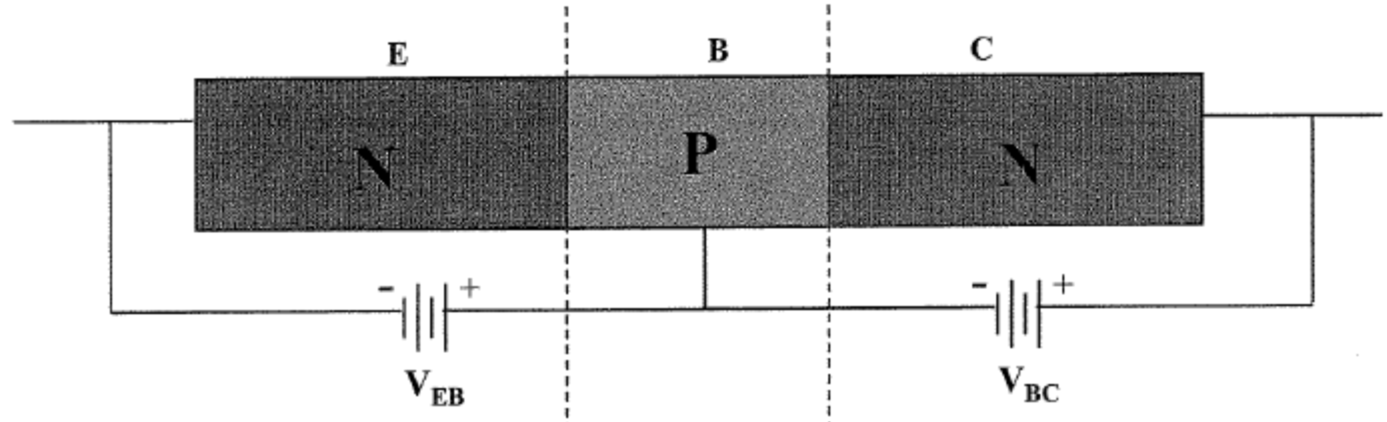
# MOSFET Band Diagram with Gate in inversion condition




MOS Capacitor will be addressed in the next Chapter.

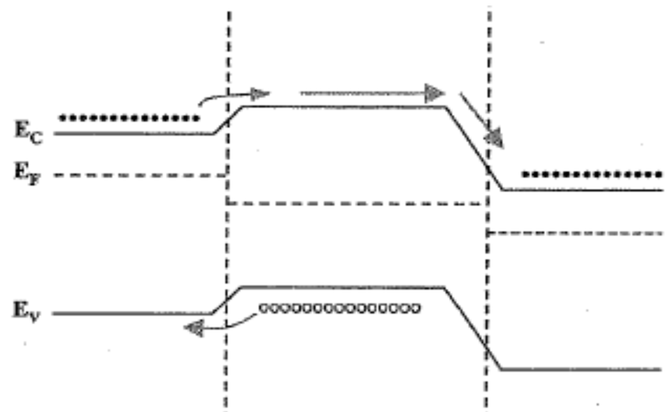
# Apply Collector or Drain Bias

## Back-to-Back PN Junctions

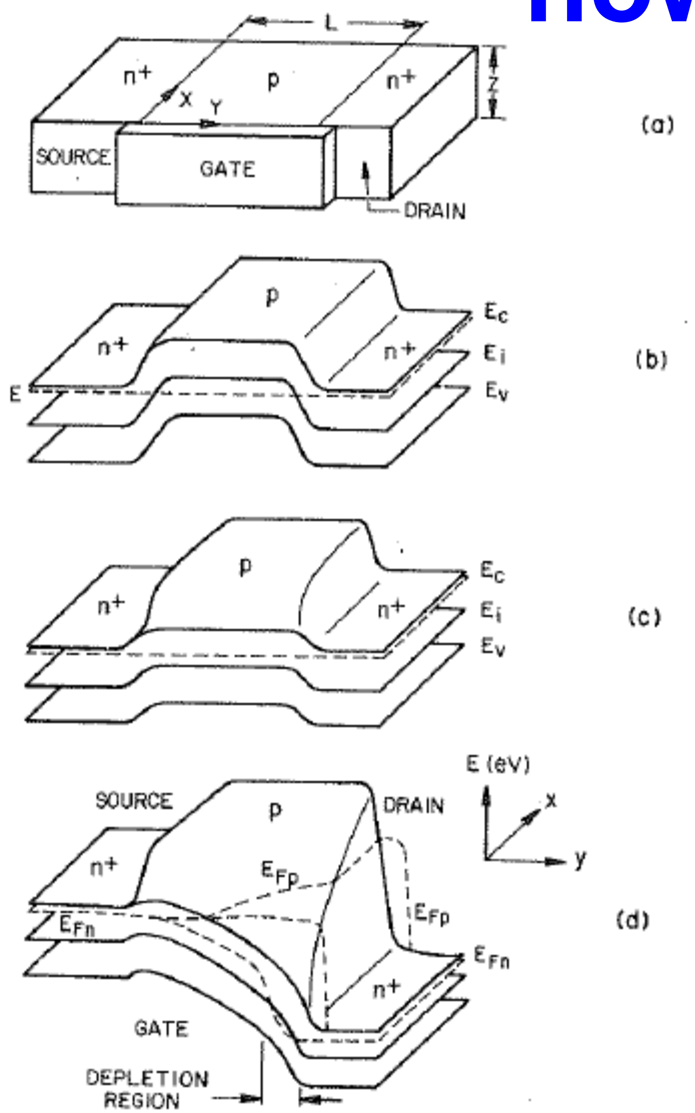



**Draw the band diagram for this structure (*E-B junction forward biased* and *B-C junction reverse biased*).**

**Applied Voltages**



# Gate in inversion condition: now Apply Drain Bias




Energy bands on the drain side are pulled down

# Concepts learned:

- ❑ pn-junction revisted
- ❑ Debye length
- ❑ Ideal pn junction diode
- ❑ Non-idealities of pn junction dode
- ❑ Reverse break-down
- ❑ Impact ionization and avalanche
- ❑ Zener breakdown
- ❑ Band-to-band tunneling
- ❑ Tunnel diode
- ❑ Back-to-back pn junction