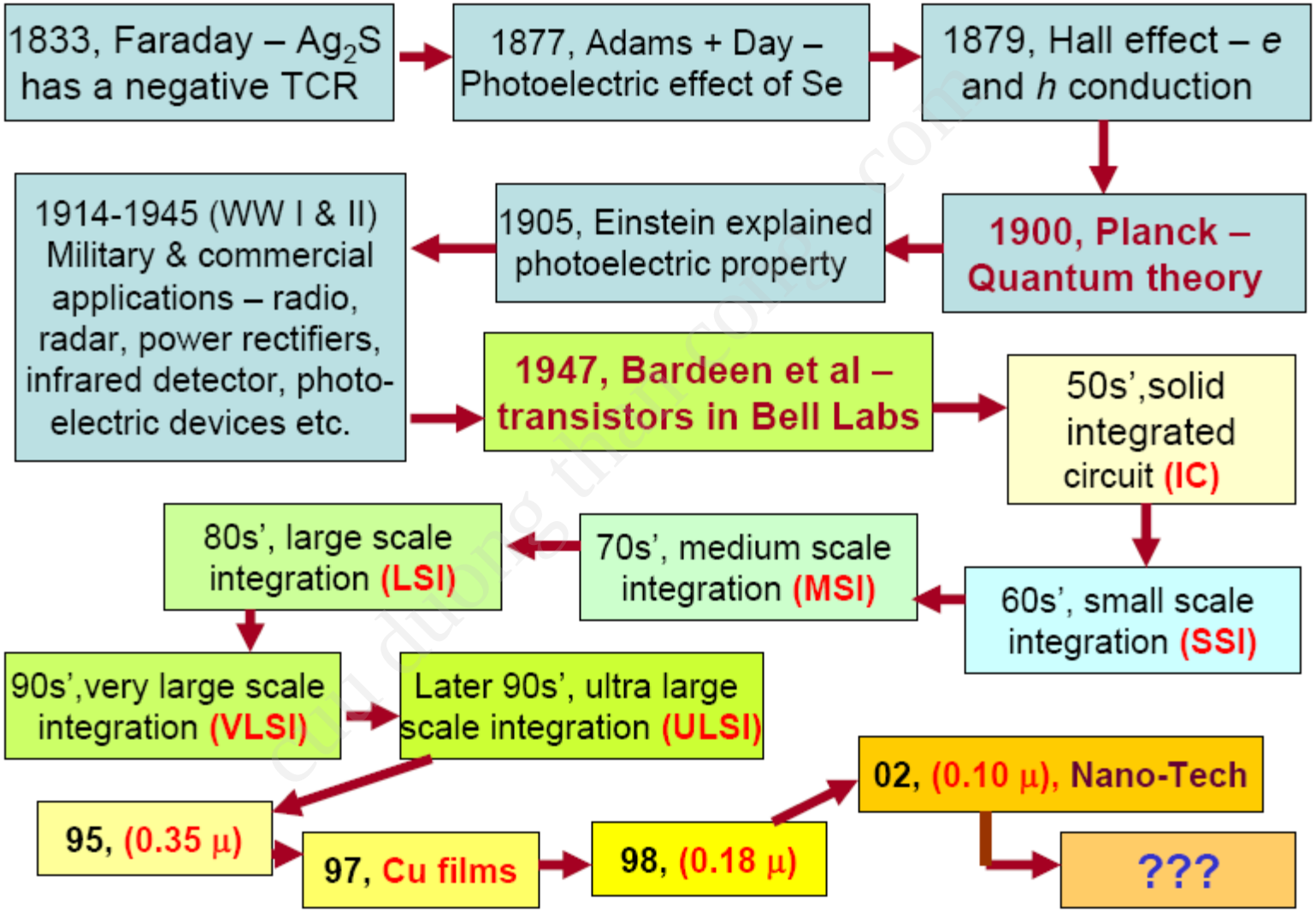
Vật liệu bán dẫn

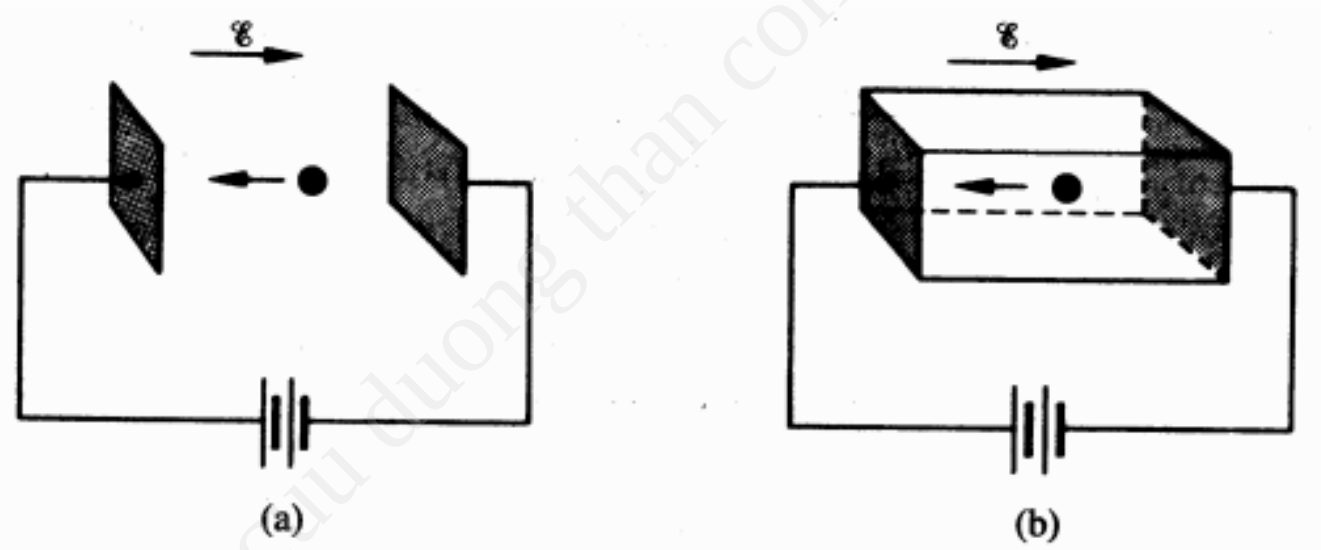
PGS.TS. Phan Bách Thắng

History of Semiconductor Materials



Mass like charge is a very basic property of electrons and holes. The mass of electrons in a semiconductor may be different than its mass in vacuum.

Effective mass concept



$$F = -q\mathcal{E} = m_0 \frac{dv}{dt}$$

$$F = -q\mathcal{E} = m_n^* \frac{dv}{dt}$$

Carrier Movement in Free Space

Newtons second law

$$F = -qE = m_o \frac{dv}{dt}$$

$$F \equiv force, v \equiv velocity, t \equiv time,$$

$$q \equiv electronic\ charge, m_o \equiv electron\ mass$$

Carrier Movement Within the Crystal

$$F = -qE = m_n^* \frac{dv}{dt}$$

$F \equiv force, v \equiv velocity, t \equiv time,$

$q \equiv electronic\ charge,$

$m_n^* \equiv electron\ effective\ mass$

$$F = qE = m_p^* \frac{dv}{dt}$$

$F \equiv force, v \equiv velocity, t \equiv time,$

$q \equiv electronic\ charge,$

$m_p^* \equiv hole\ effective\ mass$

Table 2.1 Density of States Effective Masses at 300 K.

Material	m_n^*/m_0	m_p^*/m_0
Si	1.18	0.81
Ge	0.55	0.36
GaAs	0.066	0.52

Ge and GaAs have "lighter electrons" than Si which results in faster devices

Consider a specific solution for the free space (no electrostatic potential, V=0) wave solution (electron traveling in the +x direction in 1D only):

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\Psi = E\Psi$$

$$\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} - E\Psi = 0$$

$$\Psi(x) = Ae^{ikx} + Be^{-ikx}$$

$$\text{where k} = \frac{2\pi}{\lambda} = \sqrt{\frac{2mE}{\hbar^2}} \text{ or E} = \frac{\hbar^2 k^2}{2m}$$

Since we have to add our time dependent portion (see (*) previous) our total solution is:

$$\Psi = \Psi(x)w(t) = Ae^{-i(\omega t - kx)} + Be^{-i(\omega t + kx)}$$

This is a standard wave equation with one wave traveling in the +x direction and one wave traveling in the –x direction. Since our problem stated that the electron was only traveling in the +x direction, B=0.

Classically, momentum, p=mv and kinetic energy is $(mv^2)/2 = (p^2)/2m$

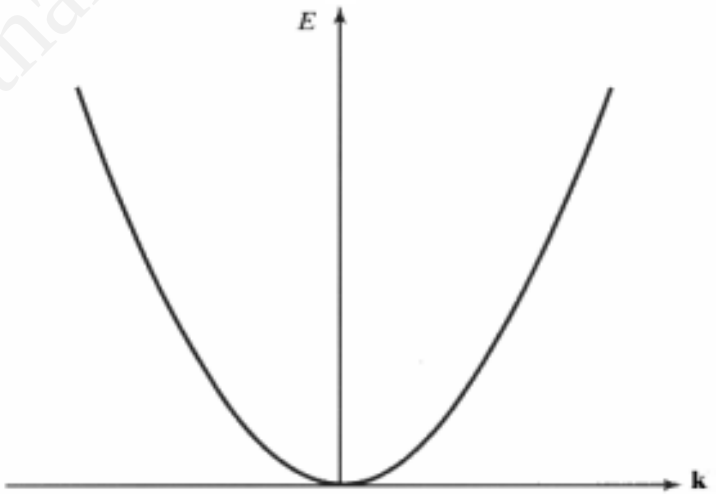
The solution to this free particle example brings out several important observations about the dual wave-particle nature of our universe:

$$\Psi = \Psi(x)w(t) = Ae^{-i(\omega t - kx)}$$

While particles act as waves, their charge is carried as a particle. I.e. you can only say that there is a "probability" of finding an electron in a particular region of space, but if you find it there, it will have all of it's charge there, not just a fraction.

Energy of moving particles follows a square law relationship:

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\langle p \rangle^2}{2m}$$



Energy–momentum relationship for a free particle.

What effect does this “E-k” square law relationship have on electron velocity and mass?

The group velocity (rate of energy delivery) of a wave is:

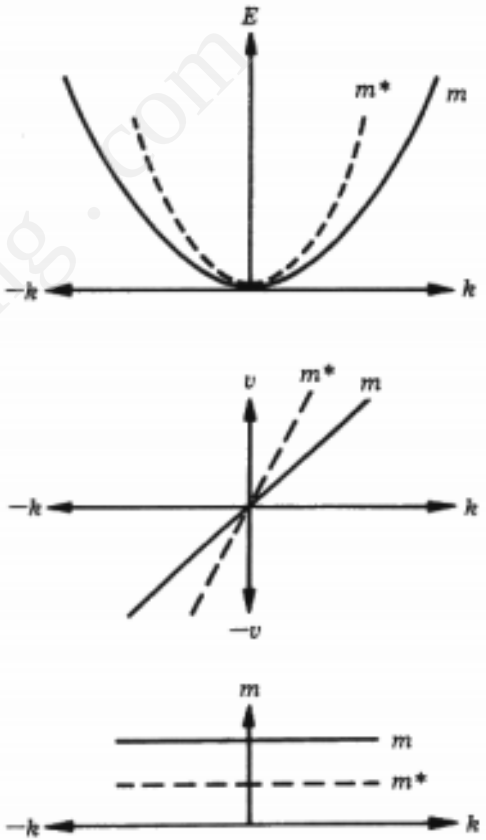
$$v_g = \frac{dE}{dp} = \frac{1}{\hbar}\frac{dE}{dk}$$

So the “speed” of an electron in the direction defined by **p** is found from the slope of the E-k diagram.

Similarly, since $E = \frac{\hbar^2 k^2}{2m}$

$$m^* = \hbar^2 \left(\frac{d^2 E}{dk^2} \right)^{-1}$$

So the “effective mass” of an electron is related to the local inverse curvature of the E-k diagram



E–k diagram for a free electron with mass *m* (solid line) and a smaller mass, *m**. The parabolic *E–k* diagram leads to a linear *v* versus *k* relation and a constant mass.

How do electrons and holes populate the bands?

Probability of Occupation (Fermi Function) Concept

Now that we know the number of available states at each energy, how do the electrons occupy these states?

We need to know how the electrons are "distributed in energy".

Again, Quantum Mechanics tells us that the electrons follow the "Fermi-distribution function".

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \quad \text{where } k \equiv \text{Boltzman cons tan } t,\ T \equiv \text{Temperature in Kelvin}$$

and $E_F \equiv$ *Fermi energy* ($\sim$ *average energy in the crystal*)

f(E) is the probability that a state at energy E is *occupied*

1-f(E) is the probability that a state at energy E is *unoccupied*

Fermi-Dirac Function and Fermi Level

The Fermi function $f(E)$ specifies how many of the existing states at the energy E will be filled with electrons. The function $f(E)$ specifies, under equilibrium conditions, the probability that an available state at an energy E will be occupied by an electron.

It is a probability distribution function.

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \qquad (2.7)$$

E_F = Fermi energy or Fermi level

k = Boltzmann constant = 1.38×10^{-23} J/K
= 8.6×10^{-5} eV/K

T = absolute temperature in K

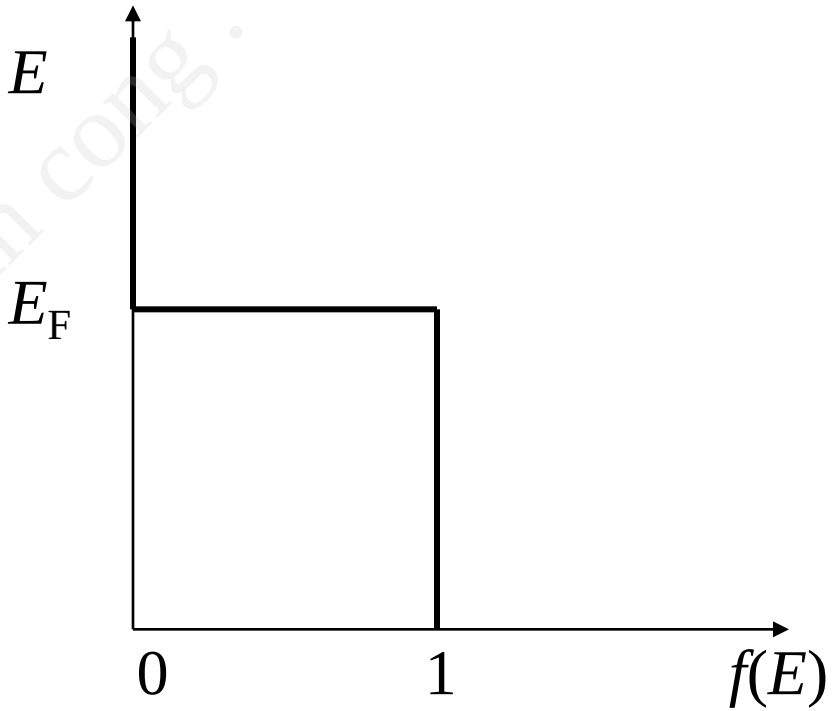
Fermi-Dirac distribution: **Consider $T \to 0$ K**

For $E > E_F$:

$$f(E > E_F) = \frac{1}{1 + \exp(+\infty)} = 0$$

For $E < E_F$:

$$f(E < E_F) = \frac{1}{1 + \exp(-\infty)} = 1$$



Fermi-Dirac distribution: **Consider $T > 0$ K**

If $E = E_F$ then $f(E_F) = ½$

If $E \geq E_F + 3kT$ then $\exp\left(\frac{E - E_F}{kT}\right) >> 1$

Thus the following approximation is valid: $f(E) = \exp\left(\frac{-(E - E_F)}{kT}\right)$

i.e., most states at energies $3kT$ above E_F are empty.

If $E \leq E_F - 3kT$ then $\exp\left(\frac{E - E_F}{kT}\right) << 1$

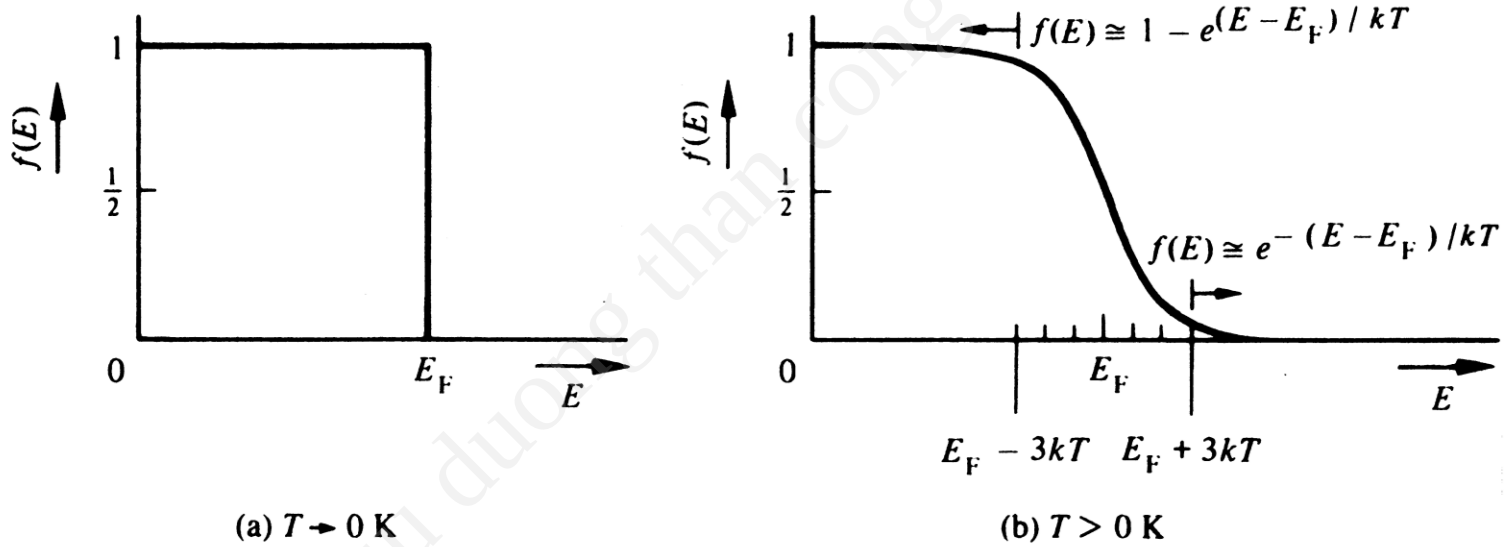
Thus the following approximation is valid: $f(E) = 1 - \exp\left(\frac{E - E_F}{kT}\right)$

So, $1 - f(E)$ = Probability that a state is empty, decays to zero.

So, most states will be filled.

kT (at 300 K) = 0.025eV, E_g(Si) = 1.1eV, so $3kT$ is very small in comparison.

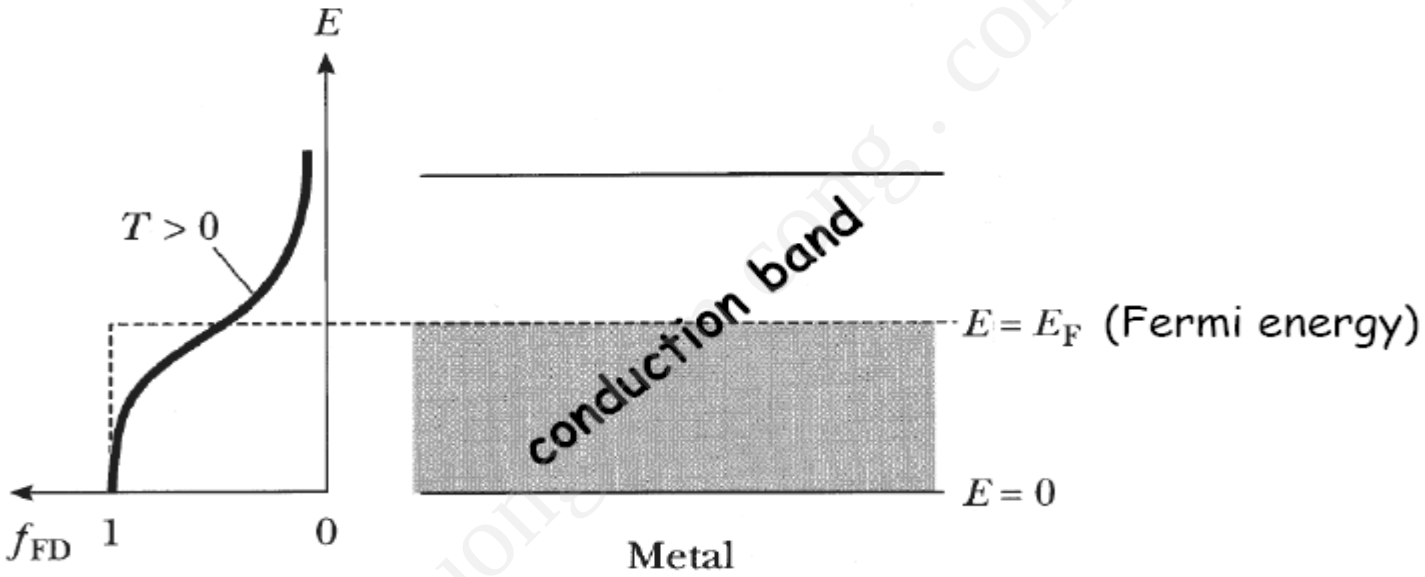
Temperature dependence of Fermi-Dirac distribution



(a) $T \to 0$ K

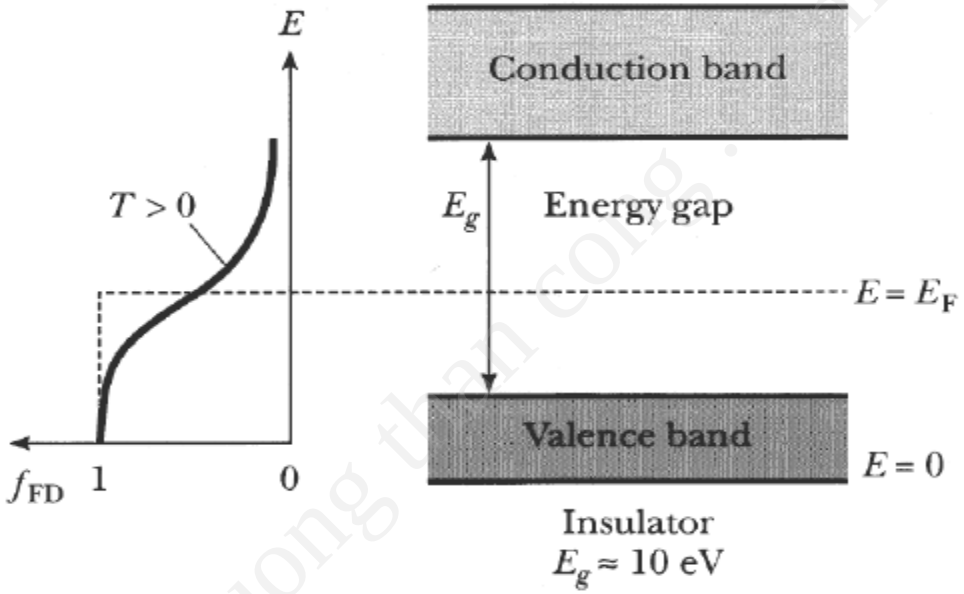
(b) $T > 0$ K

Energy bands in metals



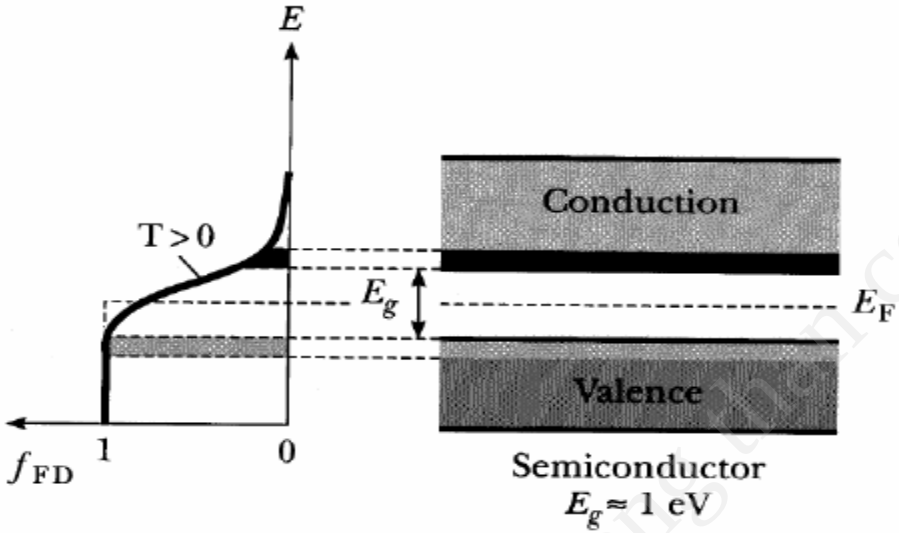
- In metals the conduction band is partially filled with (conduction) electrons with $E<E_F$
- States with $E>E_F$ in the conduction band are empty and can be easily occupied (at the cost of a tiny amount of energy) by the electrons near the Fermi surface, i.e., conduction electrons can move freely in a perfect metallic crystal

Energy bands in insulataors



- The valence (conduction) band is completely filled (empty)
- The conduction and valence bands are separated by and energy gap E_g~10eV
- The Fermi energy (chemical potential) falls inside the energy gap
- A valence electron requires $\Delta E > E_g$ to become a conduction electron, i.e., the density of conduction electrons is ~ $\exp(-E_g/k_B T)$

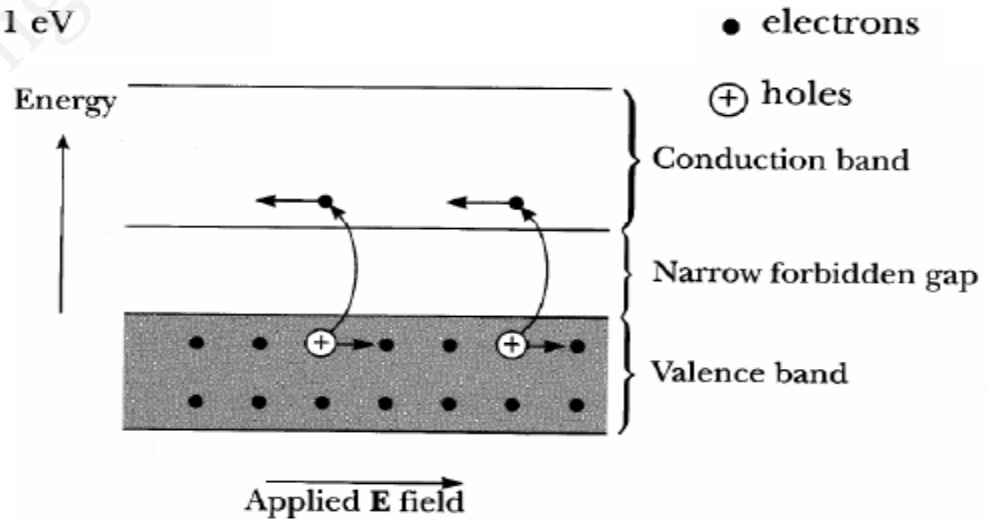
Energy bands in semiconductors



- similar band structure to insulators but with much smaller energy gap (E_g~1eV)
- poor (good) conductor (insulator) at T=0
- conductivity increases rapidly with temperature

There are two types of charge carriers in an intrinsic semiconductor (i.e., #electrons = #holes):

electrons and *holes*



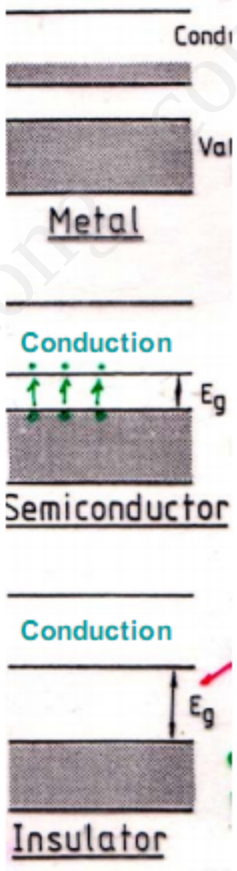
Band Structure in metals, insulators & semiconductors

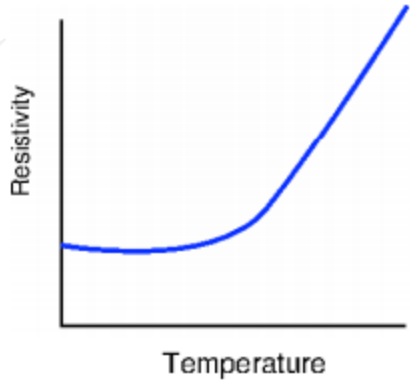
Band structure determines electrical properties:

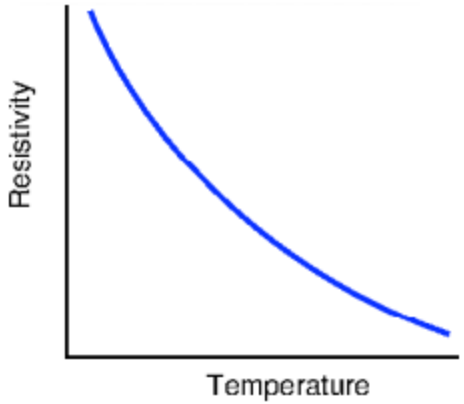
Metals → a partially filled conduction band. Electrons have sufficient energy to move.

Insulators → a completely filled highest valence band & large energy gap (2-10 eV) to the next (conduction) band.

Semiconductors → between metals and insulators → valence band is completely filled, but the gap is narrow (<2 eV). The gap allows e_s thermally excited to the conduction band → controlled conductivity.







Carrier Movement Within the Crystal

- Electron is a quasi-particle that behaves as a "wave" due to quantum mechanical effects.
- The electron "wavelength" is perturbed by the crystals periodic potential.

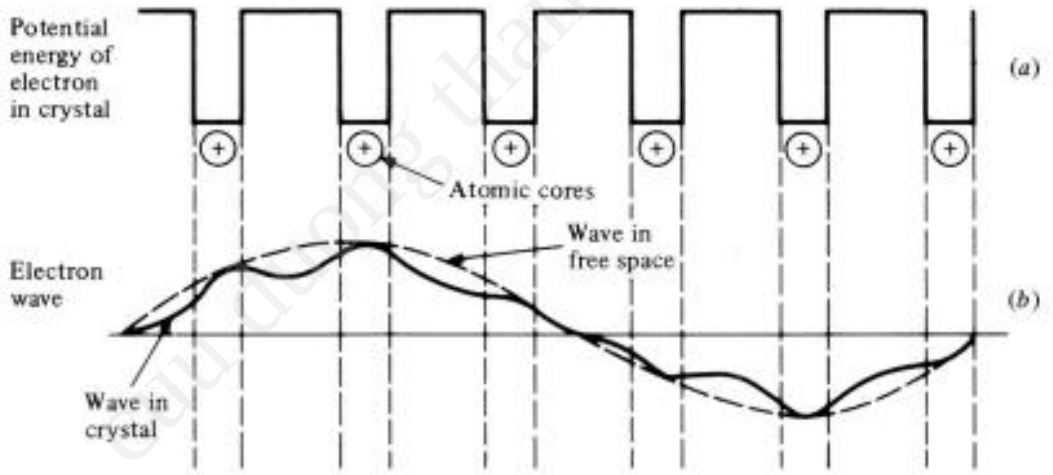


FIGURE 1-11
Representation of motion of electron wave in crystal potential. (*After Wolfendale [3].*)

Carrier Movement Within the Crystal

$$F = -qE = m_n^* \frac{dv}{dt}$$

$F \equiv force, v \equiv velocity, t \equiv time,$

$q \equiv electronic\ charge,$

$m_n^* \equiv electron\ effective\ mass$

$$F = qE = m_p^* \frac{dv}{dt}$$

$F \equiv force, v \equiv velocity, t \equiv time,$

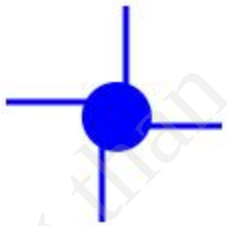
$q \equiv electronic\ charge,$

$m_p^* \equiv hole\ effective\ mass$

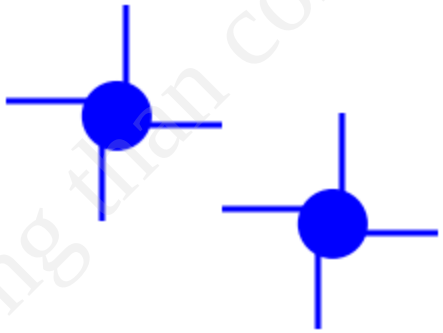
Table 2.1 Density of States Effective Masses at 300 K.

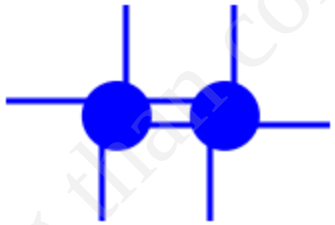
Material	m_n^*/m_0	m_p^*/m_0
Si	1.18	0.81
Ge	0.55	0.36
GaAs	0.066	0.52

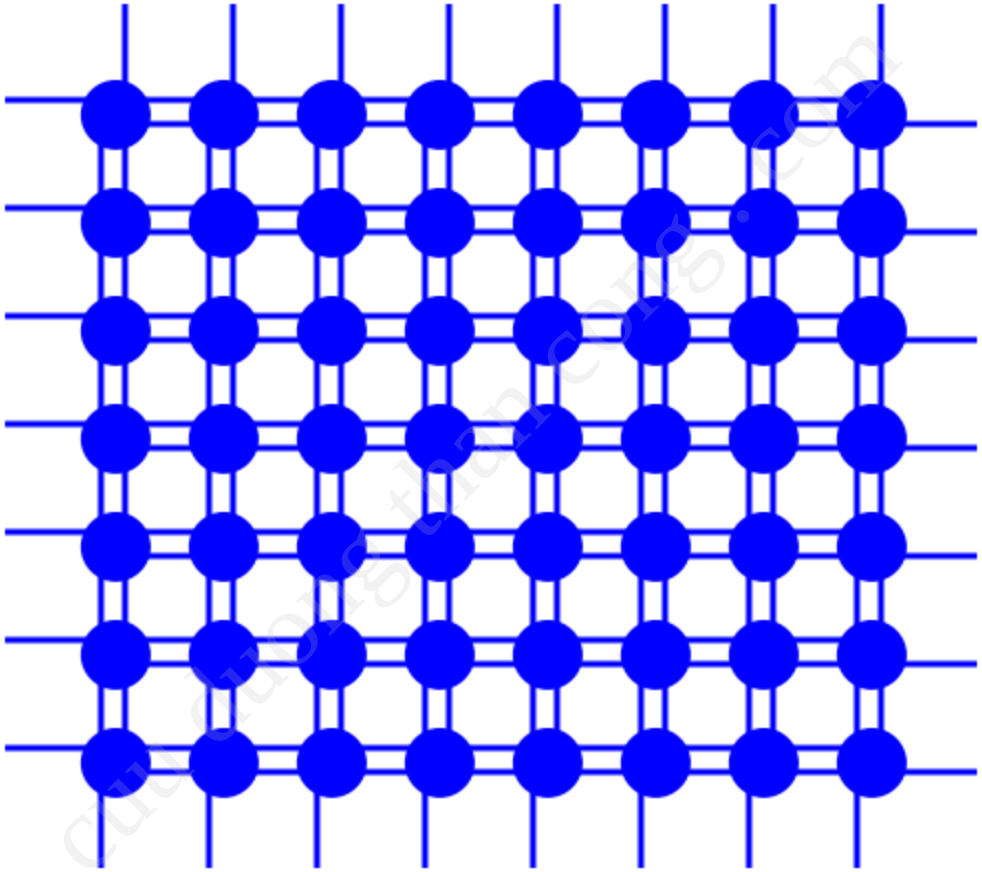
Ge and GaAs have "lighter electrons" than Si which results in faster devices



4 electrons available for sharing (covalent bonding) in outer shell of atoms

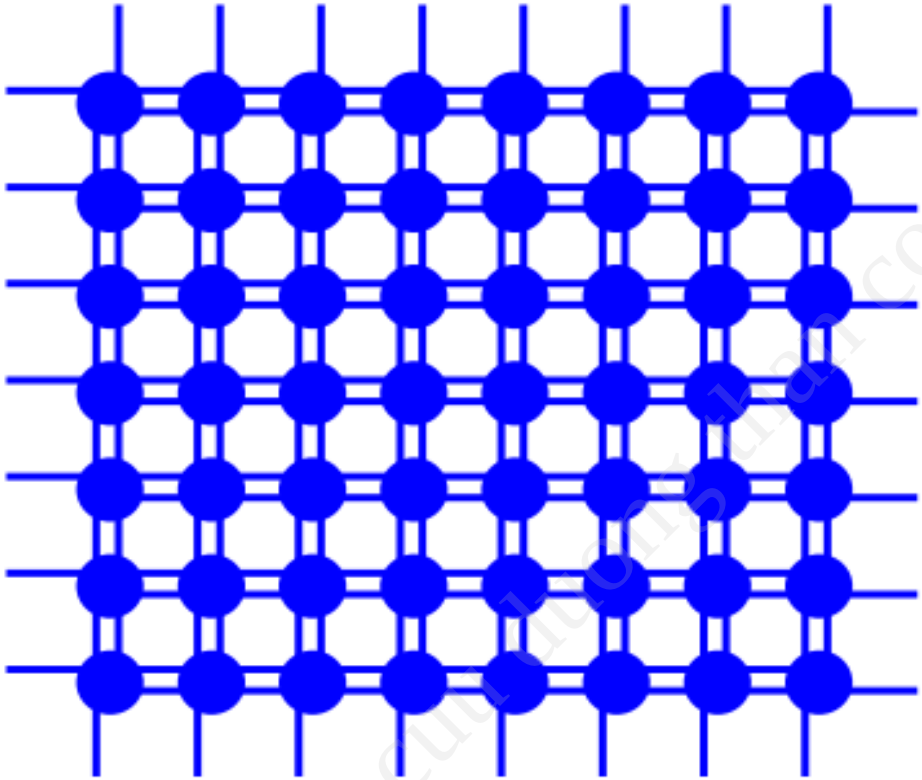






Band Occupation at Low Temperature (0 Kelvin)

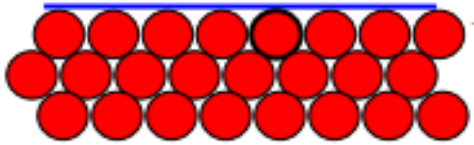
For ($E_{thermal}$=kT)=0



No electrons in conduction band means no electron conduction is possible

Ec

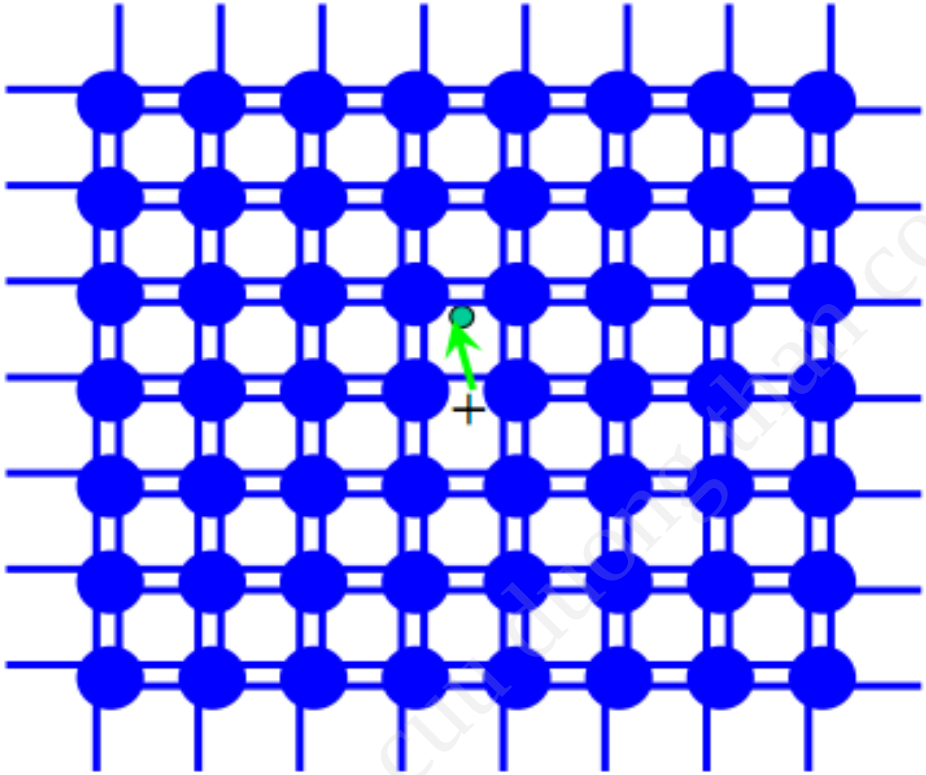
Ev



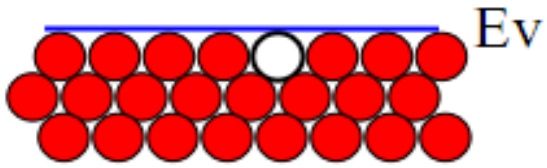
No “Holes” valence band means no “hole” conduction is possible

Band Occupation at Higher Temperature (T>0 Kelvin)

For ($E_{thermal}$=kT)>0



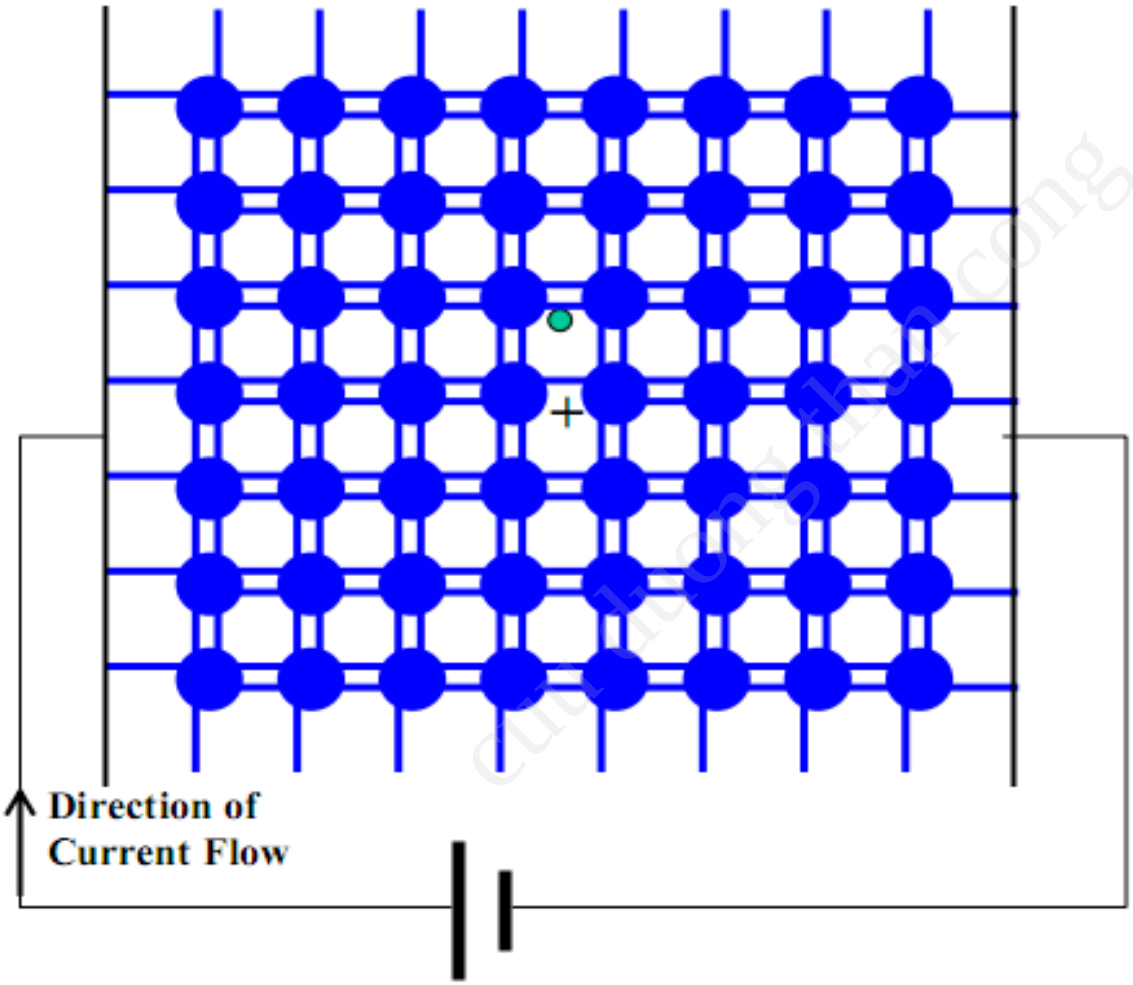
Electron free to move in conduction band

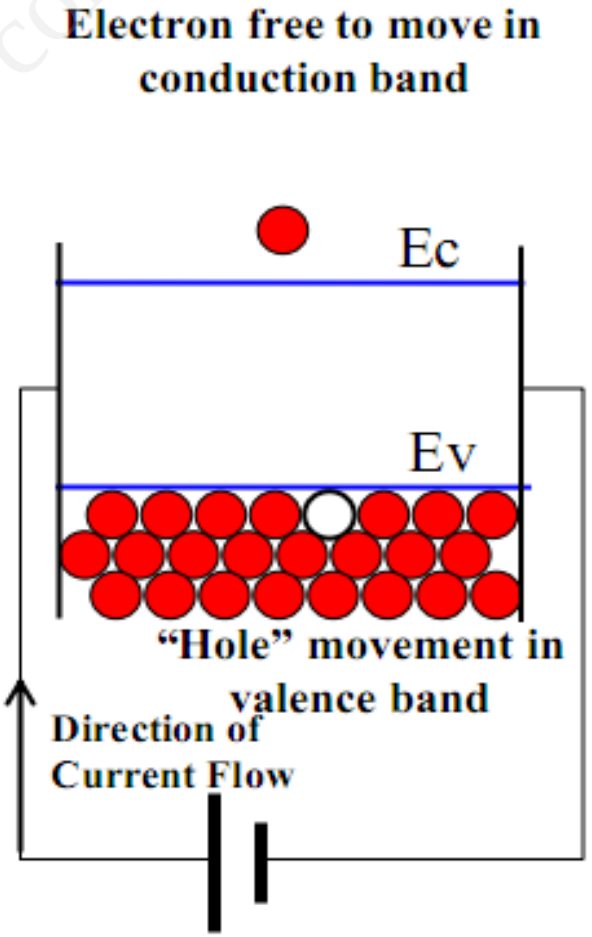


"Hole" free to move in valence band

Carrier Movement Under Bias

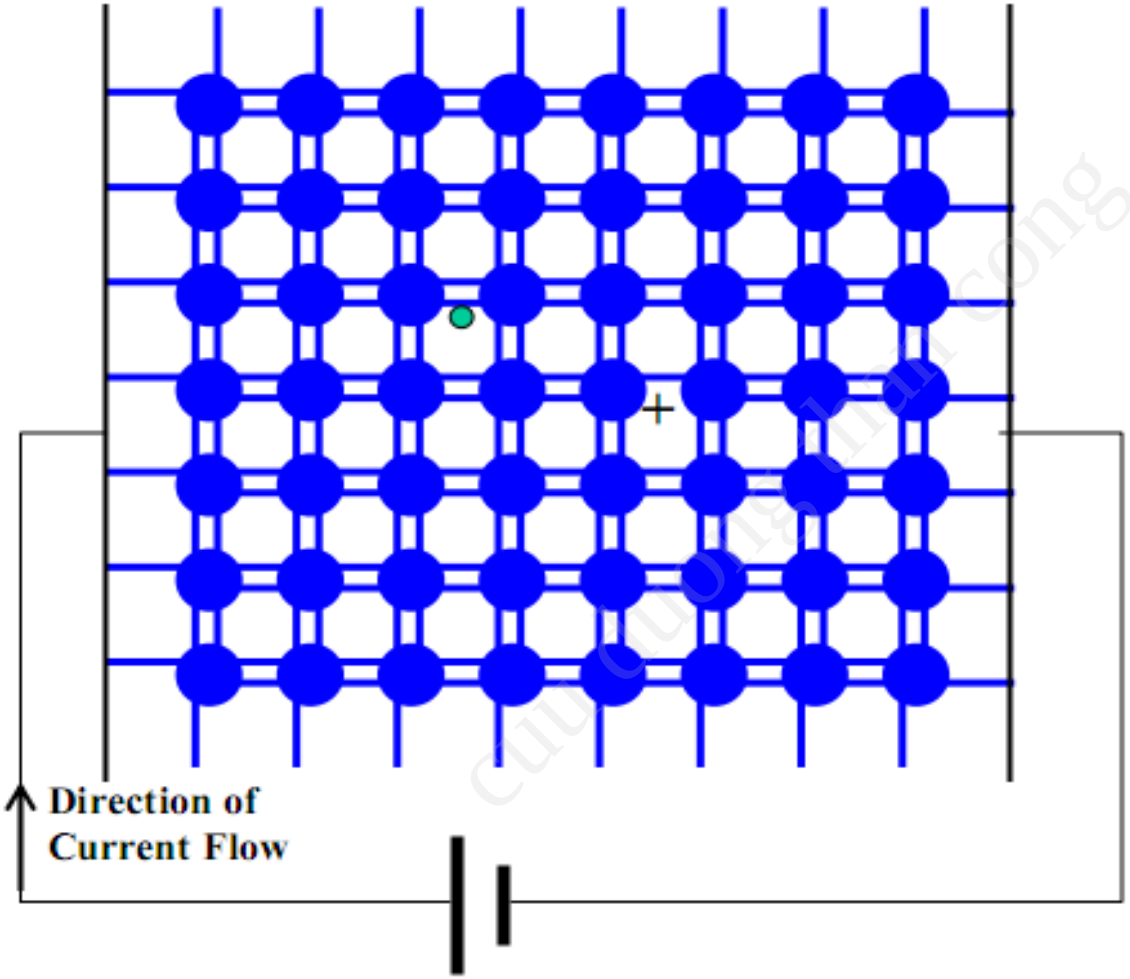
For ($E_{thermal}$=kT)>0

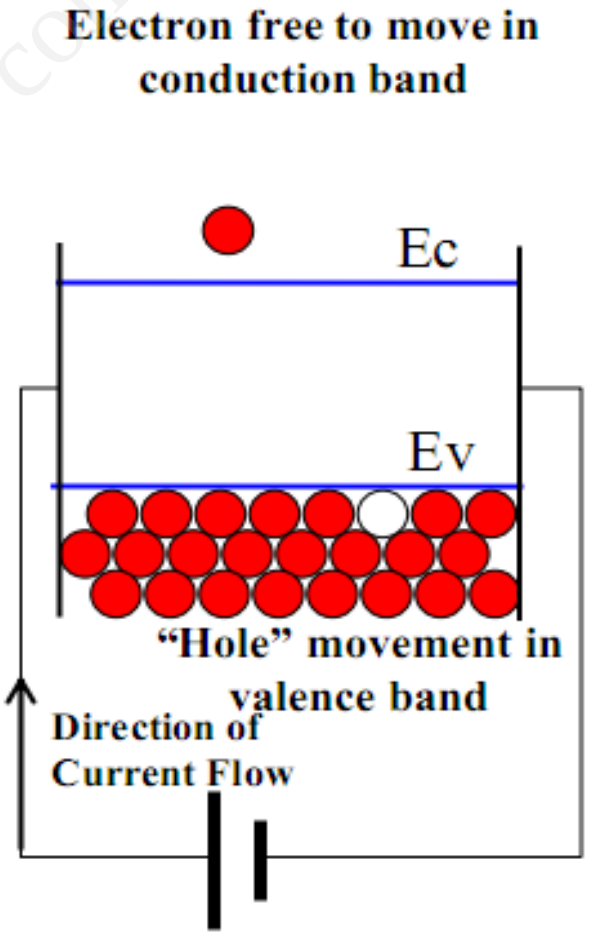




Carrier Movement Under Bias

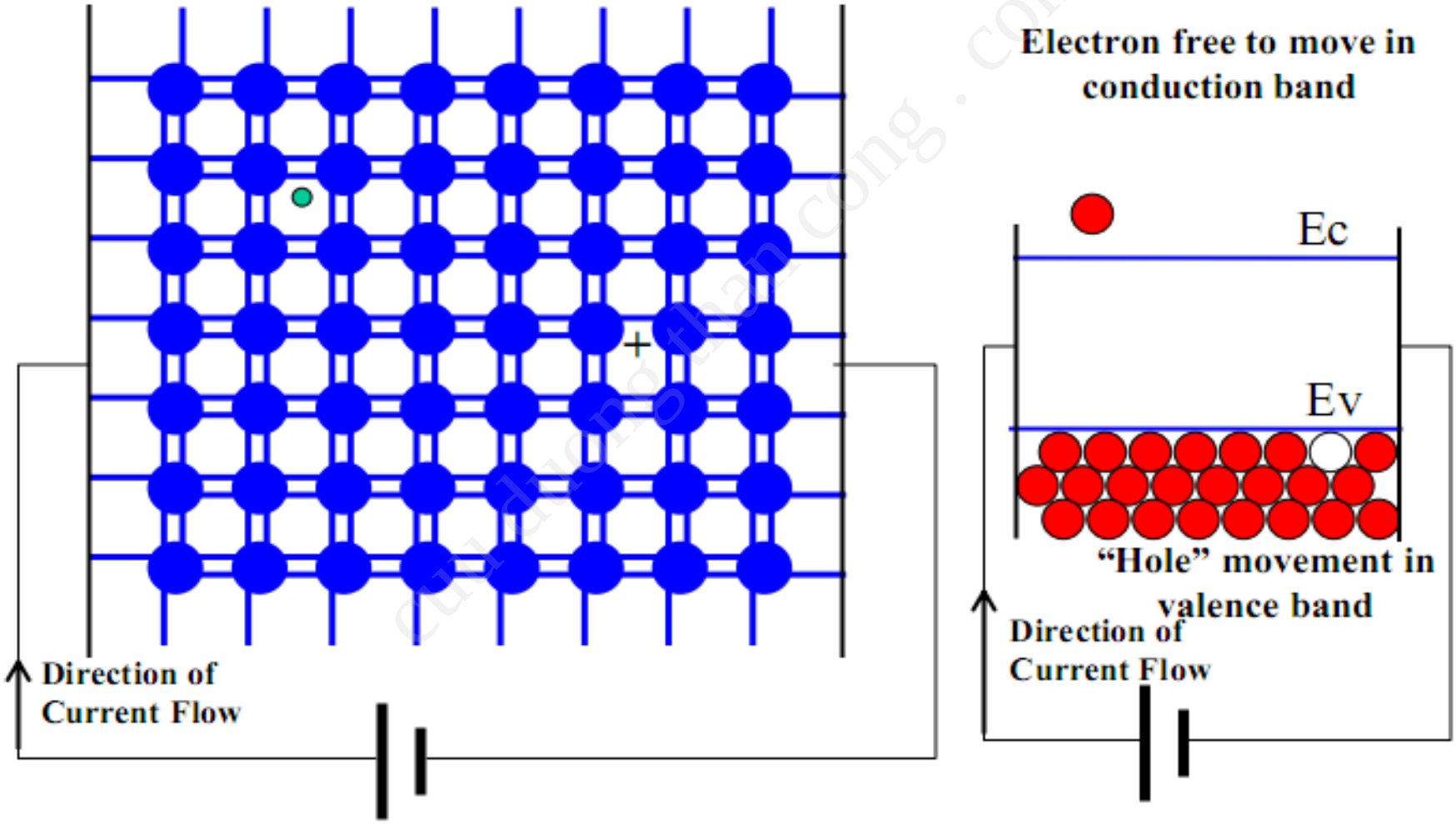
For $(E_{thermal}=kT)>0$



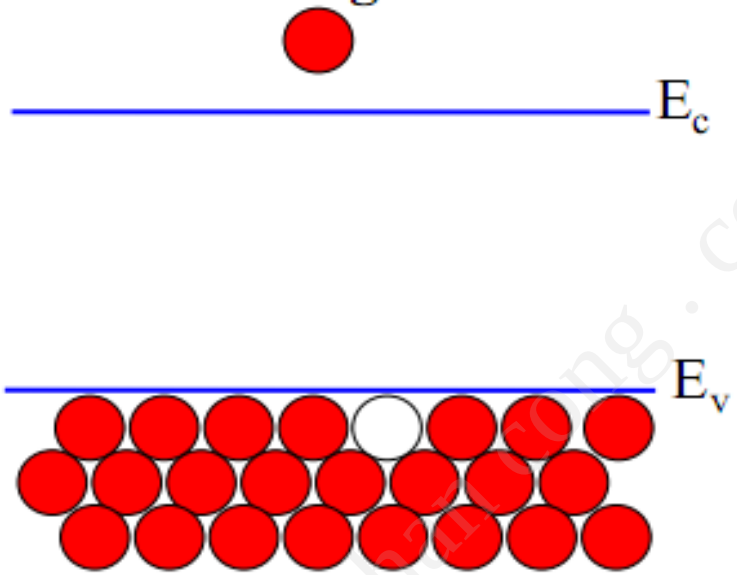


Carrier Movement Under Bias

For ($E_{thermal}$=kT)>0



Clarification of confusing issues:"Holes" and Electrons



The valance band may have ~4e22 cm^{-3} valence electrons "participating in the bonding processes holding the crystal together.

The valance band might only have ~1e6 to 1e19 cm^{-3} "holes" in the valence band (missing valence electrons). Thus, it is easier to account for the influence of the holes by counting the holes directly as apposed to counting very small changes in the valence electron concentrations.

Example: If there are 1e 22 cm-3 atoms in a crystal with each atom having 4 valence electrons. What is the difference in valence electron concentration for 1e12 holes verses 1e13 cm-3 holes?

Answer: 4 x 1e22 cm^{-3} -1e12 cm^{-3} = 3.9999999999e22cm^{-3} verses

4 x 1e22 cm^{-3} -1e13 cm^{-3} = 3.999999999e22cm^{-3}

Material Classification based on Size of Bandgap:

Ease of achieving thermal population of conduction band determines whether a material is an insulator, semiconductor, or metal

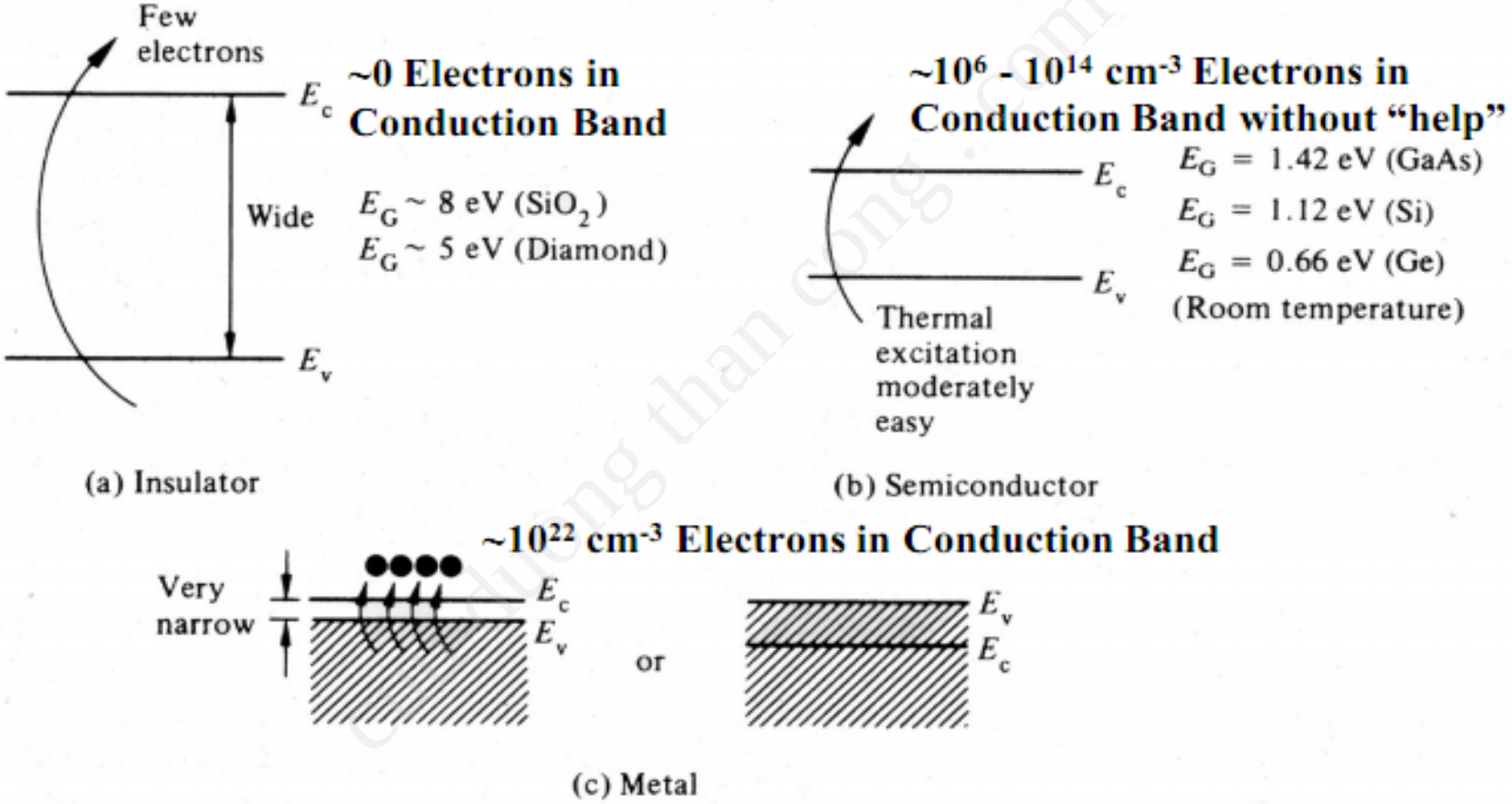


Figure 2.8 Explanation of the distinction between (a) insulators, (b) semiconductors, and (c) metals using the energy band model.

Intrinsic Carrier Concentration

- For each electron promoted to the conduction band, one hole is left in the valence band. Thus, the number of electrons in the conduction band is equal to the number of holes in the valence band unless there is "help" to change the relative populations in each band.

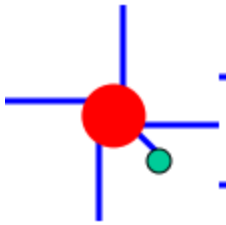
- Intrinsic carrier concentration is the number of electron (=holes) per cubic centimeter populating the conduction band (or valence band) is called the intrinsic carrier concentration, n_i

- n_i = f(T) that increases with increasing T (more thermal energy)

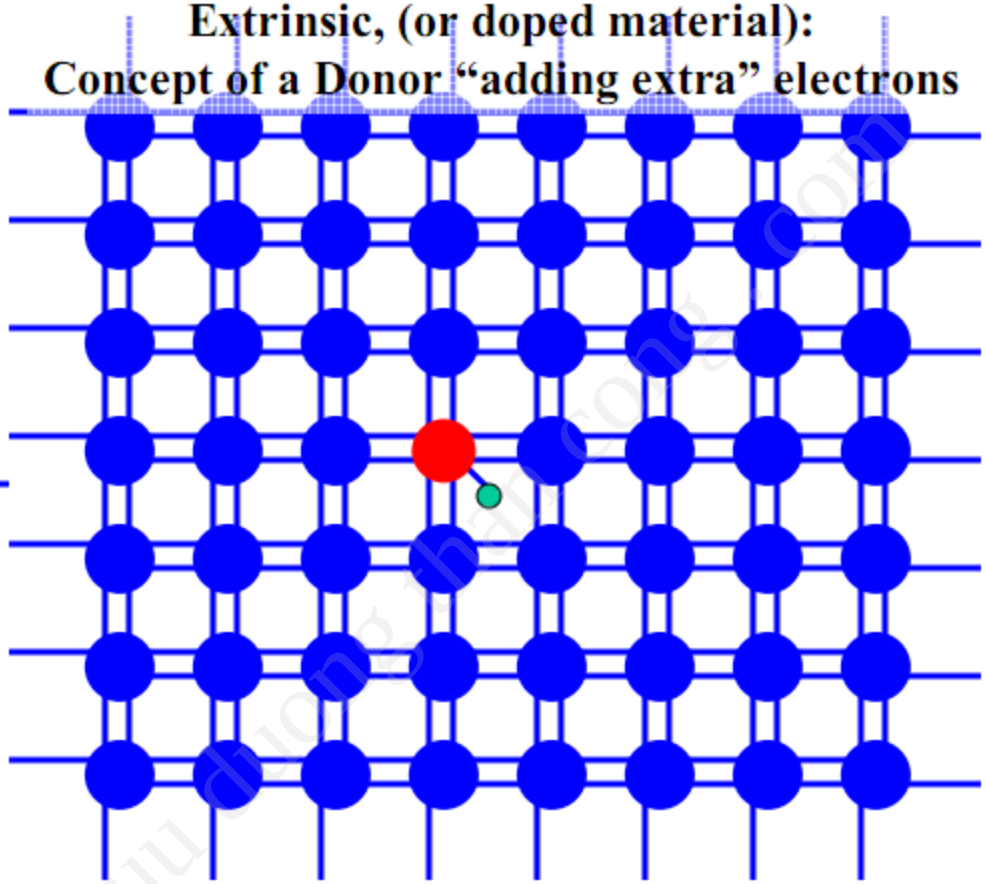
<u>At Room Temperature (T=300 K)</u>

n_i~2e6 cm^{-3} for GaAs with Eg=1.42 eV,
n_i~1e10 cm^{-3} for Si with Eg=1.1 eV,
n_i~2e13 cm^{-3} for Ge with Eg=0.66 eV,
n_i~1e-14 cm^{-3} for GaN with Eg=3.4 eV

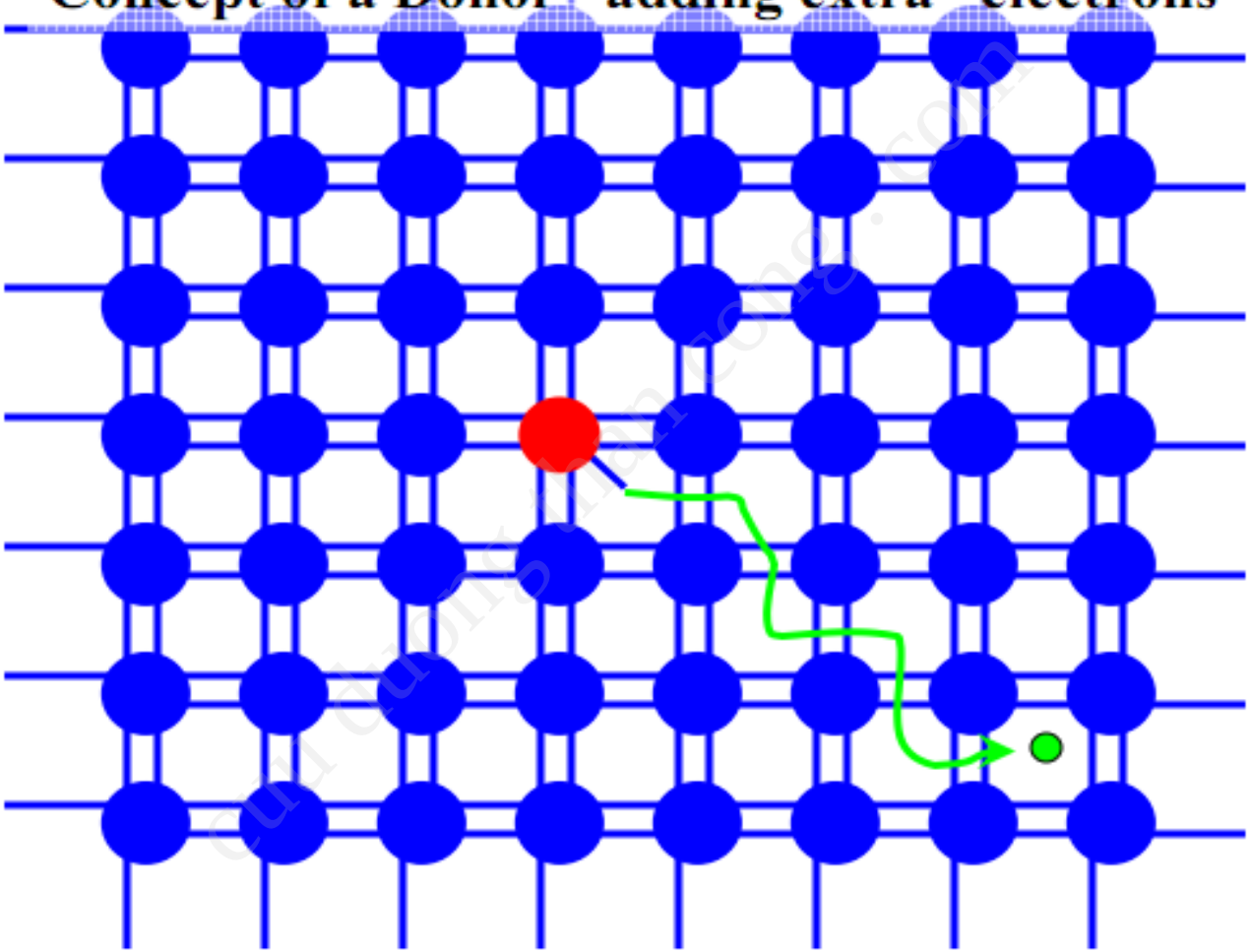
Extrinsic, (or doped material):
Concept of a Donor "adding extra" electrons



Example:
P, As, Sb
in Si

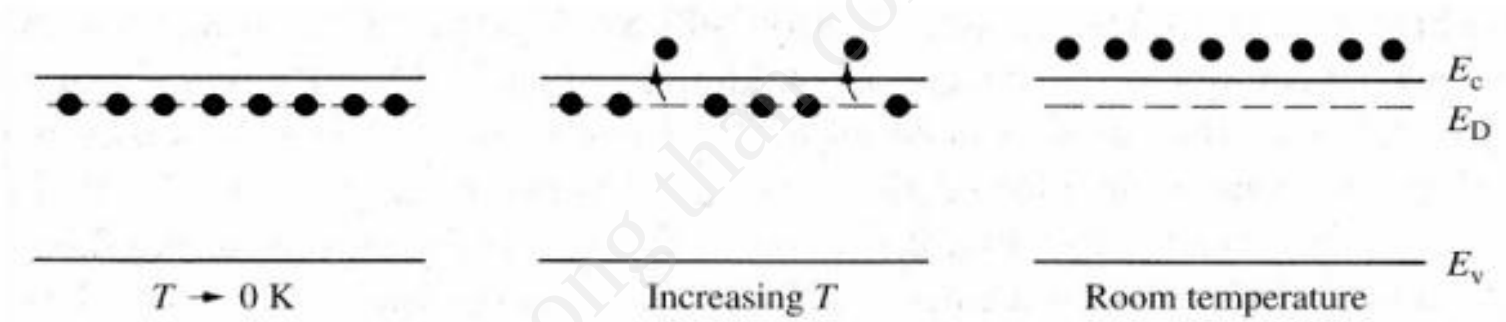


Extrinsic, (or doped material): Concept of a Donor "adding extra" electrons

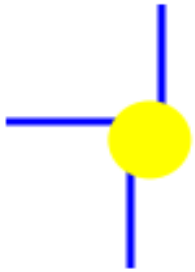


Concept of a Donor "adding extra" electrons:

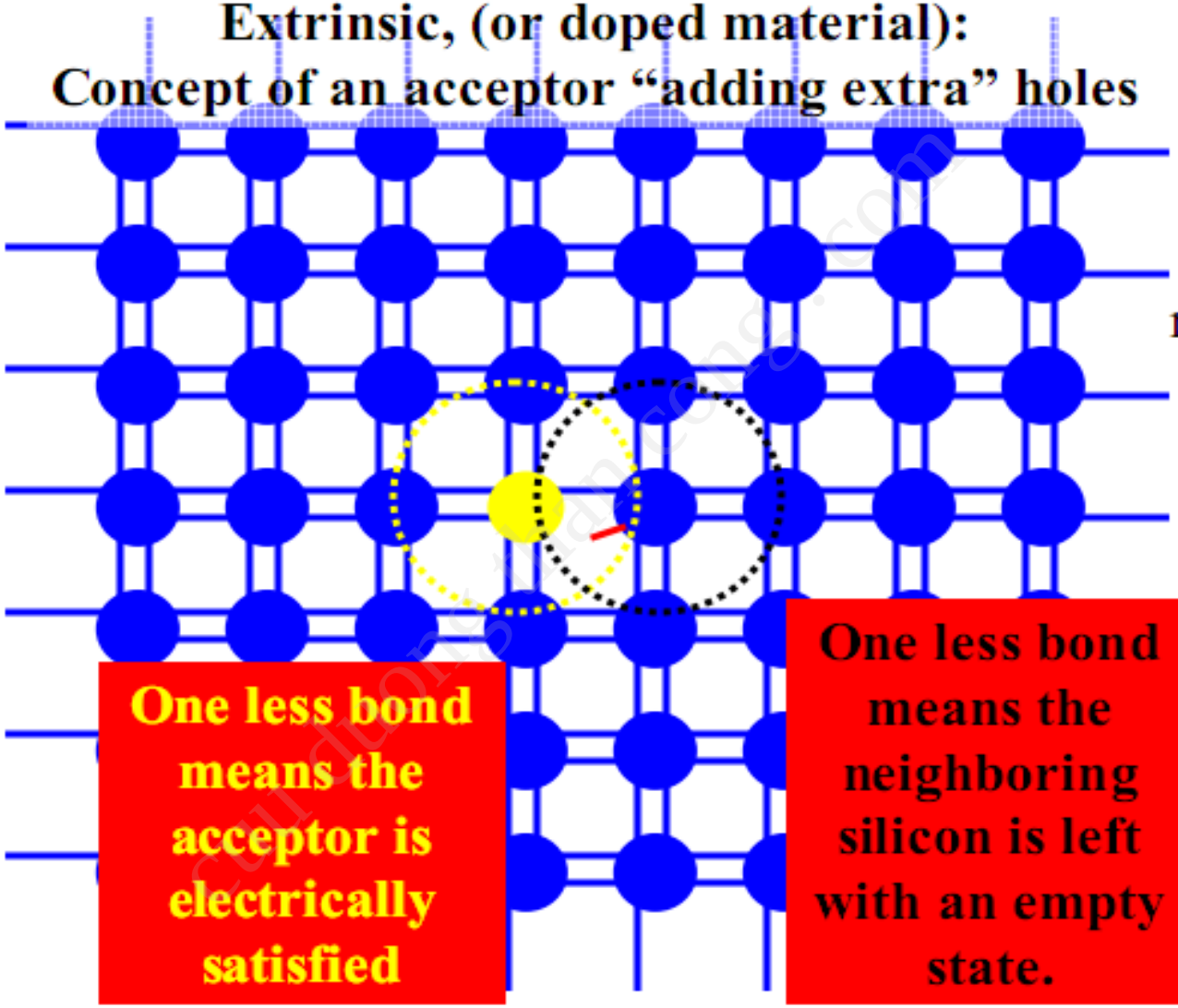
Band diagram equivalent view



Extrinsic, (or doped material): Concept of an acceptor "adding extra" holes



Example: B, Al, In in Si



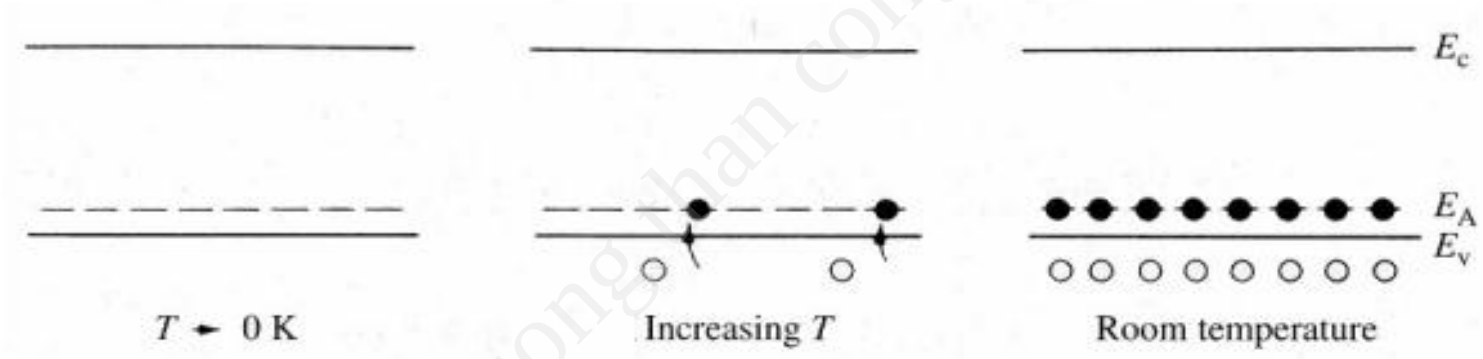
All regions of material are neutrally charged.

One less bond means the acceptor is electrically satisfied

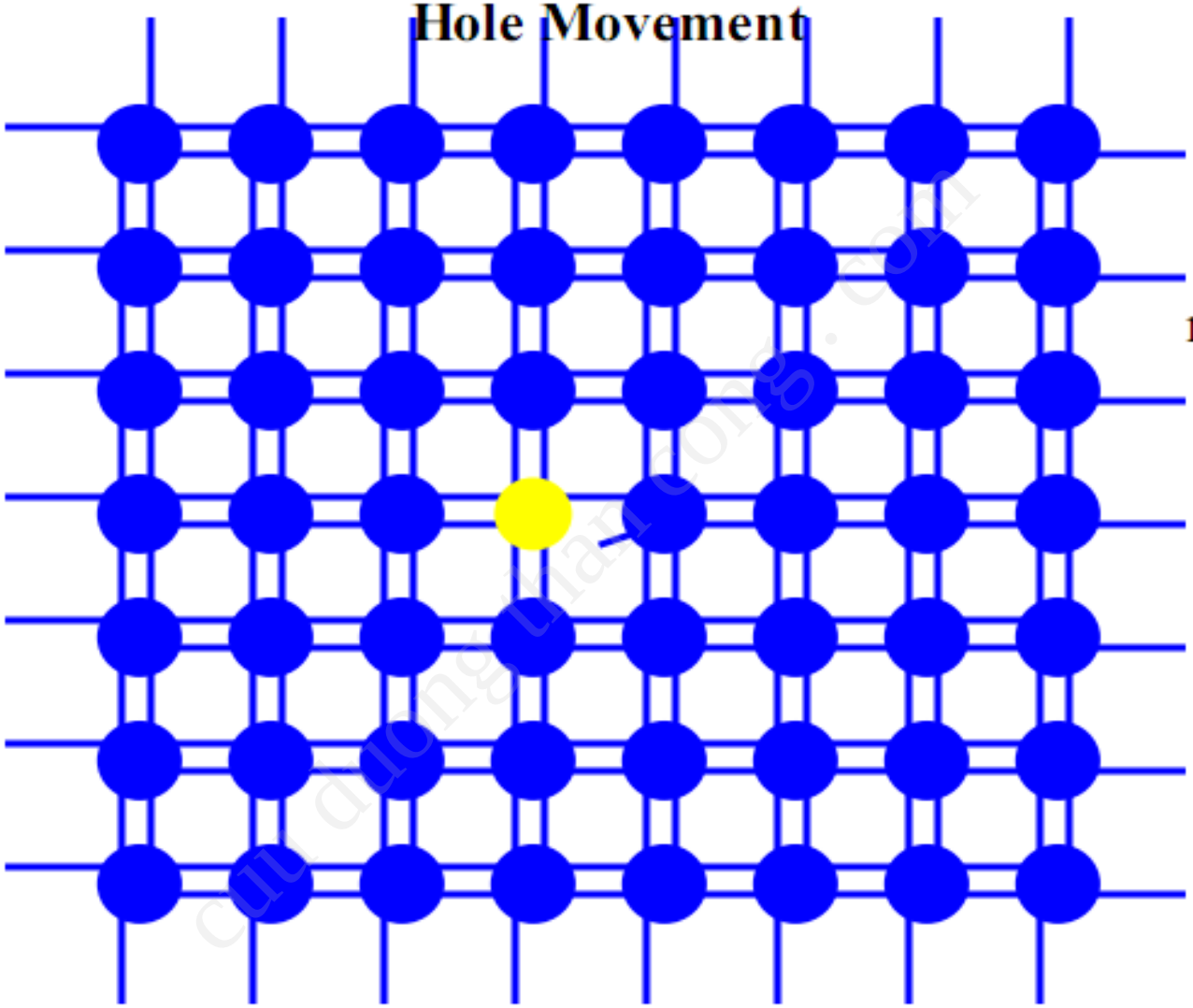
One less bond means the neighboring silicon is left with an empty state.

Concept of an Acceptor"adding extra hole":

Band diagram equivalent view



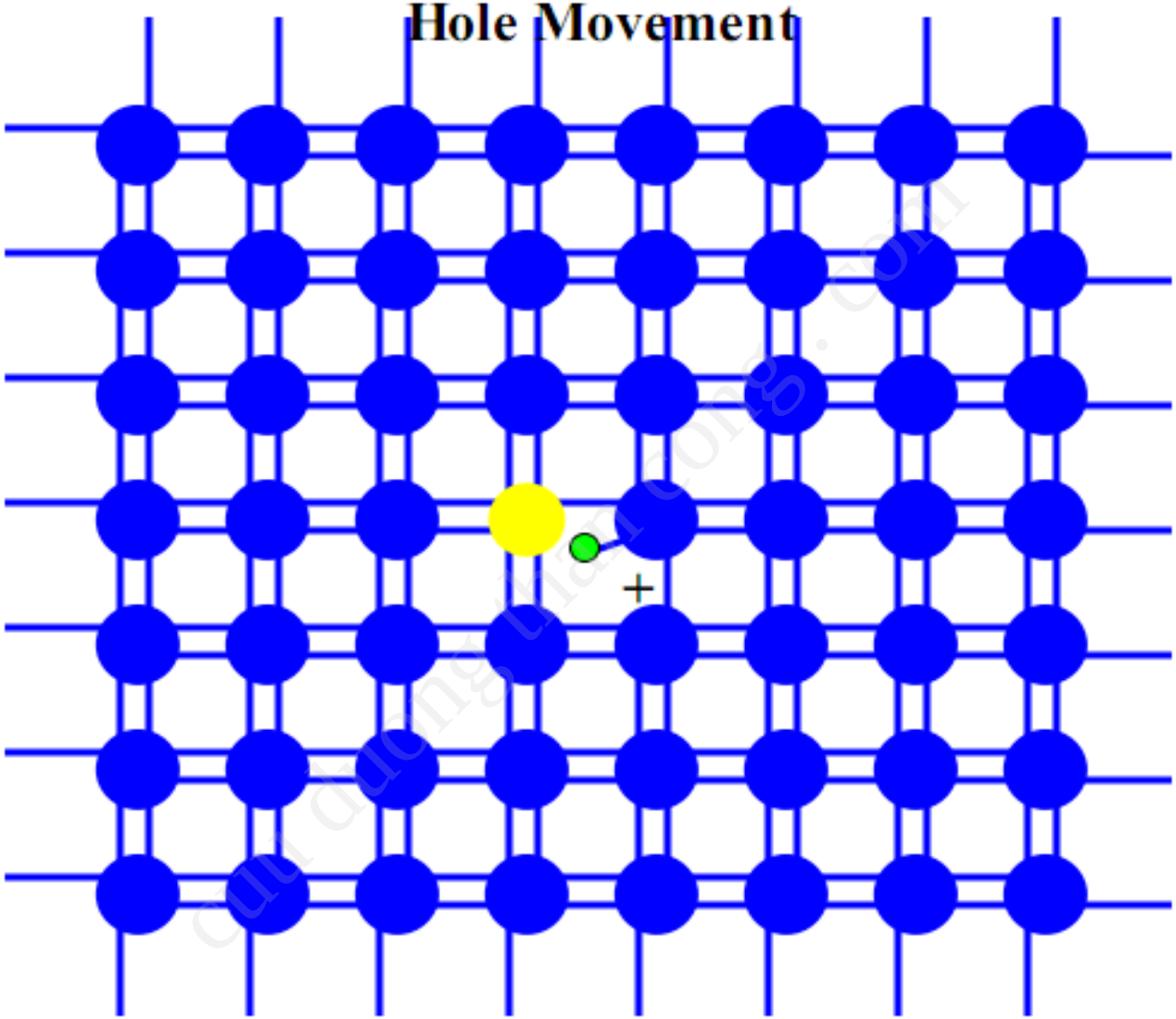
Hole Movement



All regions of material are neutrally charged.

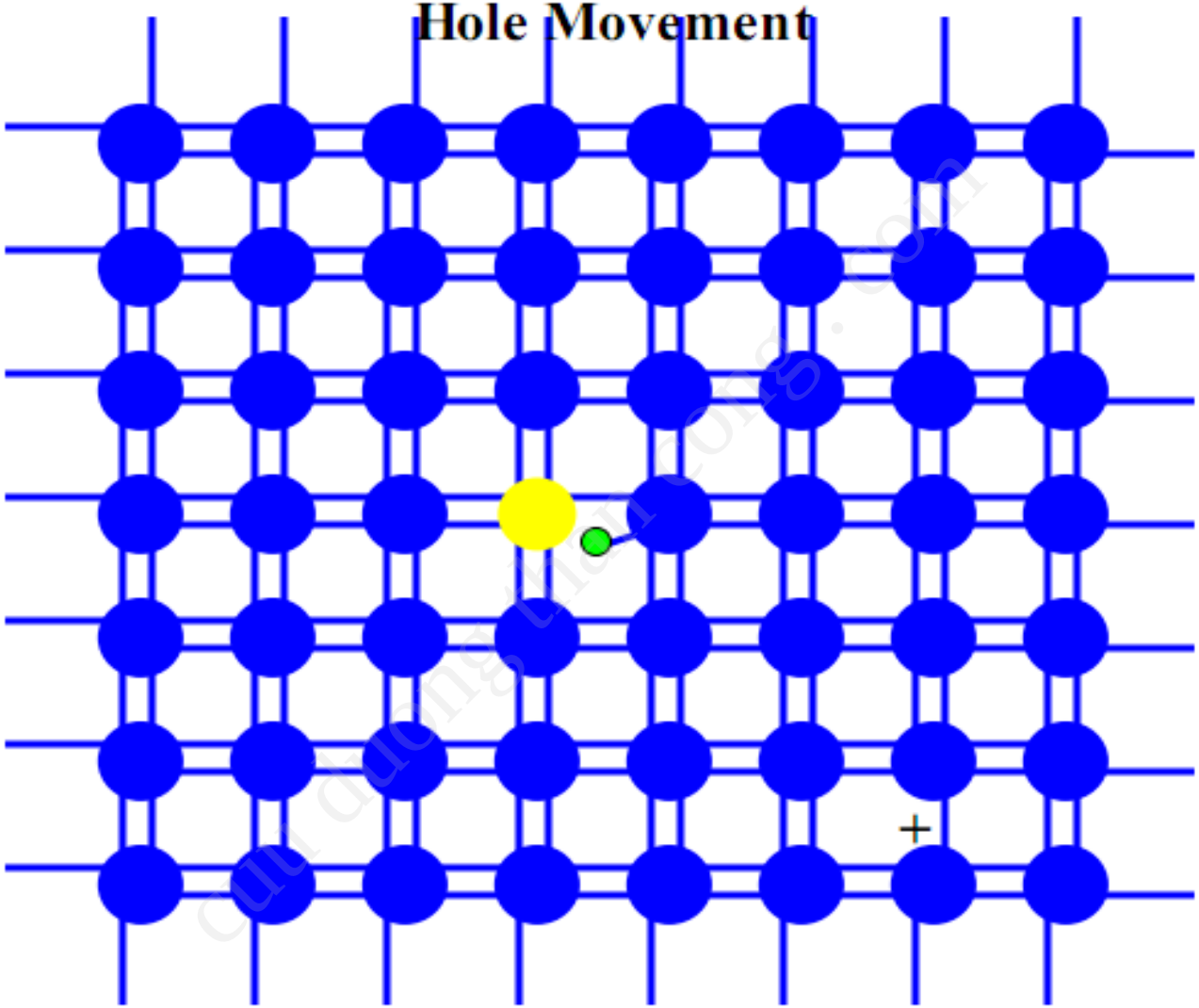
Empty state is located next to the Acceptor

Hole Movement



Another valence electron can fill the empty state located next to the Acceptor leaving behind a positively charged “hole”.

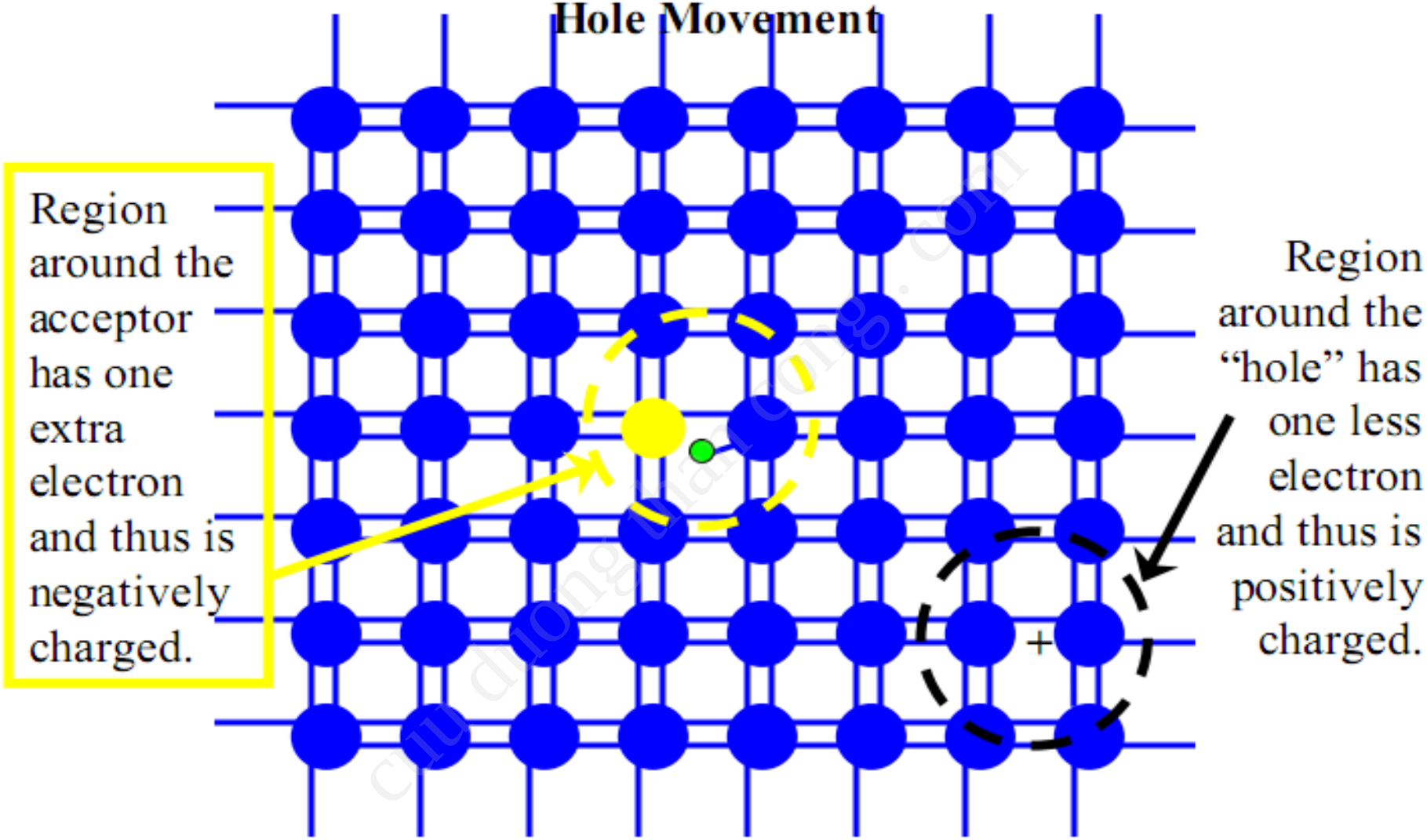
Hole Movement



The positively charged "hole" can move throughout the crystal

(really it is the valance electrons jumping from atom to atom that creates the hole motion).

Hole Movement



The positively charged "hole" can move throughout the crystal

(really it is the valance electrons jumping from atom to atom that creates the hole motion).

Extrinsic Semiconductor

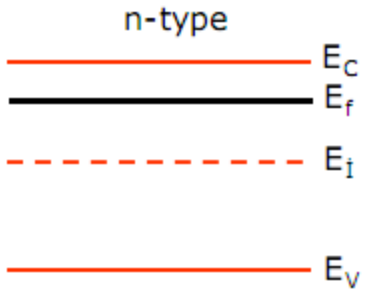
- Since the concentrations of free electrons and holes is small in an intrinsic semiconductor, only small currents are possible.
- Impurities can be added to the semiconductor to increase the concentration of free electrons and holes.
- An impurity would have one less or one more electron in the valance shell than silicon.
- Impurities for group 4 type atoms (silicon) would come from group 3 or group 5 elements.

Extrinsic Semiconductor

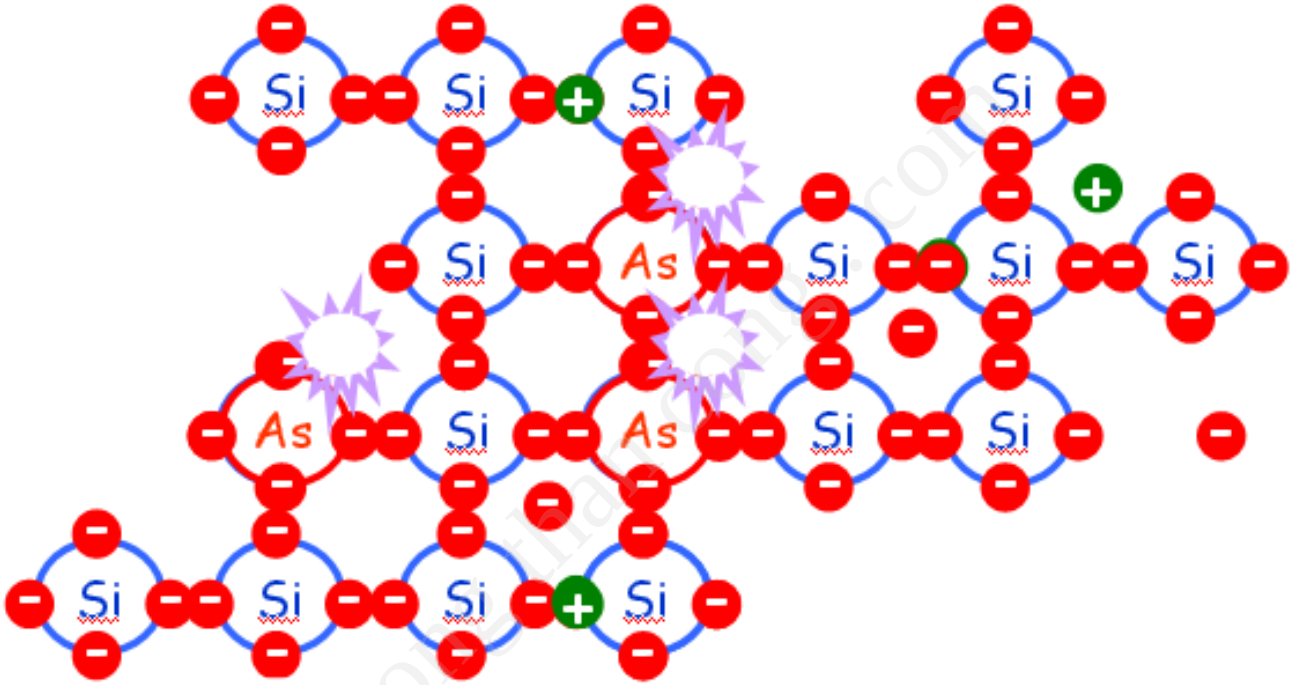
- The most common group 5 elements are phosphorous and arsenic.
- Group 5 elements have 5 electrons in the valence shell.
- Four of the electrons fill the covalent bonds in the silicon crystal structure.
- The 5th electron is loosely bound to the impurity atom and is a free electron at room temperature.

- The group 5 atom is called a **donor** impurity since it donates a free electron.
- The group 5 atom has a net positive charge that is fixed in the crystal lattice and cannot move.
- With a donor impurity, free electrons are created without adding holes.

- Adding impurities is called doping.
- A semiconductor doped with donor impurities has excess free electron and is called an n-type semiconductor.



Increasing conductivity by doping

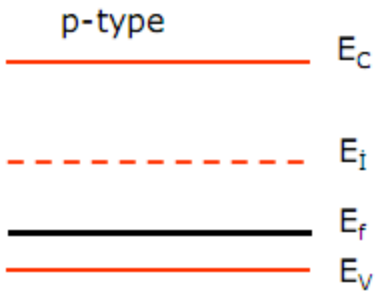


- Inject Arsenic into the crystal with an implant step.
- Arsenic is Group5 element with 5 electrons in its outer shell, (one more than silicon).
- This introduces extra electrons into the lattice which can be released through the application of heat and so produces and electron current
- The result here is an N-type semiconductor (n for negative current carrier)

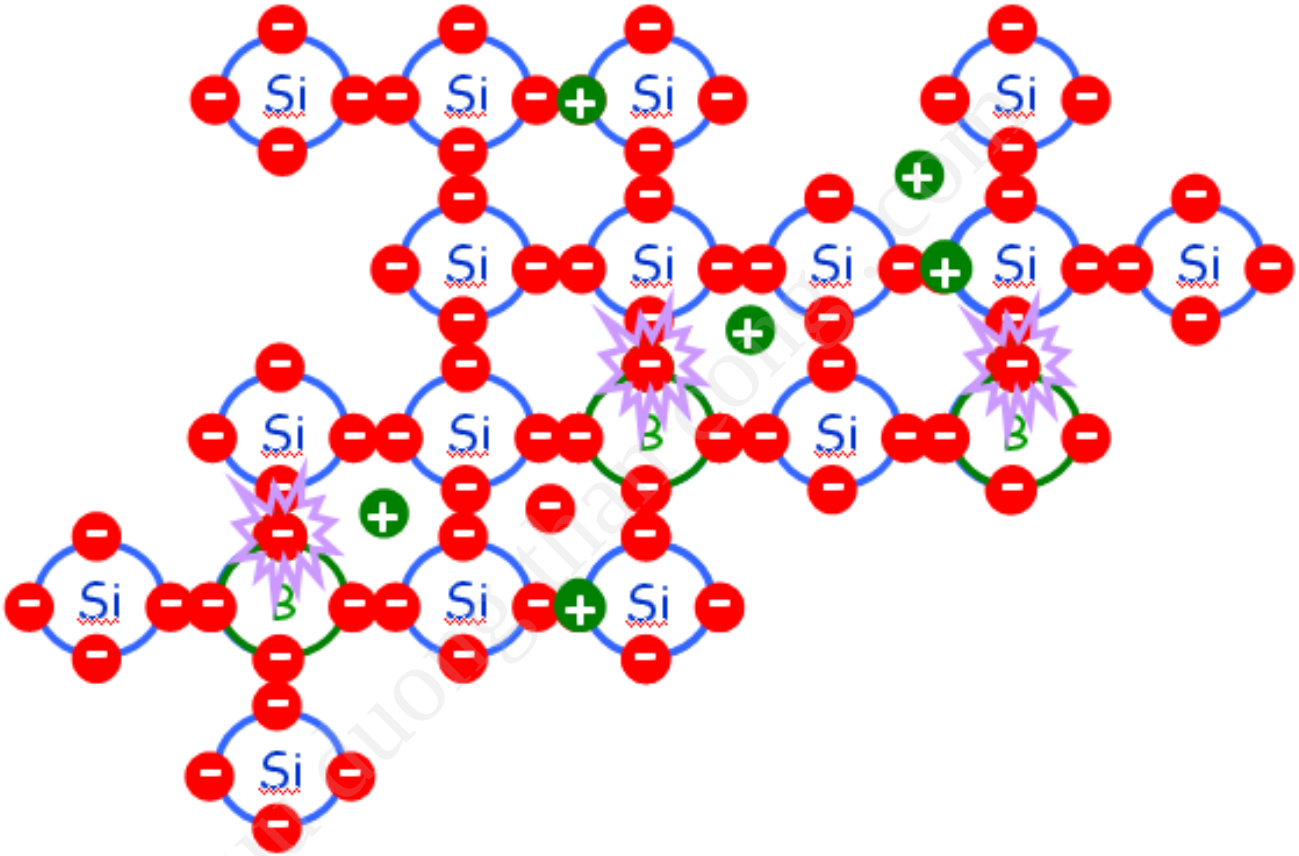
Extrinsic Semiconductor

- The most common group 3 impurity is boron which has 3 valence electrons.
- Since boron has only 3 valence electrons, the boron atom can only bond with three of its neighbors leaving one open bond position.
- At room temperature, silicon has free electrons that will fill the open bond position, creating a hole in the silicon atom whence it came.
- The boron atom has a net negative charge because of the extra electron, but the boron atom cannot move.

- Since boron accepts a valence electron, it is called an acceptor impurity.
- Acceptor impurities create excess holes but do not create free electrons.
- A semiconductor doped with an acceptor impurity has extra holes and is called a p-type semiconductor.



Increasing conductivity by doping



- Inject Boron into the crystal with an implant step.
- Boron is Group3 element is has 3 electrons in its outer shell (one less than silicon)
- This introduces holes into the lattice which can be made mobile by applying heat. This gives us a hole current
- The result is a P-type semiconductor (p for positive current carrier)

Summary of Important terms and symbols

Bandgap Energy: Energy required to remove a valence electron and allow it to freely conduct.
Intrinsic Semiconductor: A "native semiconductor" with no dopants. Electrons in the conduction band equal holes in the valence band. The concentration of electrons (=holes) is the intrinsic concentration, $\mathbf{n_i}$.
Extrinsic Semiconductor: A doped semiconductor. Many electrical properties controlled by the dopants, not the intrinsic semiconductor.
Donor: An impurity added to a semiconductor that adds an additional electron not found in the native semiconductor.
Acceptor: An impurity added to a semiconductor that adds an additional hole not found in the native semiconductor.
Dopant: Either an acceptor or donor.
N-type material: When electron concentrations (**n=number of electrons/cm^3**) exceed the hole concentration (normally through doping with donors).
P-type material: When hole concentrations (**p=number of holes/cm^3**) exceed the electron concentration (normally through doping with acceptors).
Majority carrier: The carrier that exists in higher population (ie n if n>p, p if p>n)
Minority carrier: The carrier that exists in lower population (ie n if n<p, p if p<n)
Other important terms (among others): Insulator, semiconductor, metal, amorphous, polycrystalline, crystalline (or single crystal), lattice, unit cell, primitive unit cell, zincblende, lattice constant, elemental semiconductor, compound semiconductor, binary, ternary, quaternary, atomic density, Miller indices, various notations, etc...

Developing the mathematical model for electrons and holes

The density of electrons is:

$$n = \int_{E_{Bottom\ of\ conduction\ band}}^{E_{Top\ of\ conduction\ band}} g_c(E) f(E) dE$$

Probability the state is filled

Number of states per cm^{-3} in energy range dE

The density of holes is:

$$p = \int_{E_{Bottom\ of\ valence\ band}}^{E_{Top\ of\ valence\ band}} g_v(E)[1 - f(E)]dE$$

Probability the state is empty

Number of states per cm^{-3} in energy range dE

Note: units of n and p are #/cm^3

Developing the mathematical model for electrons and holes

We can further define:

$$N_c = 2\left[\frac{m_n^*(kT)}{2\pi\hbar^2}\right]^{3/2}$$ *the effective density of states in the conduction band*

and

$$N_v = 2\left[\frac{m_p^*(kT)}{2\pi\hbar^2}\right]^{3/2}$$ *the effective density of states in the valence band*

This is a general relationship holding for all materials and results in:

$$N_c = 2.51x10^{19}\left(\frac{m_n^*}{m_o}\right)^{3/2} cm^{-3} \quad at\ 300K$$

$$N_v = 2.51x10^{19}\left(\frac{m_v^*}{m_o}\right)^{3/2} cm^{-3} \quad at\ 300K$$

Developing the mathematical model for electrons and holes

When $n=n_i$, $E_f=E_i$ (the intrinsic energy), then

$$n_i = N_c e^{(E_i - E_c)/kT} \quad or \quad N_c = n_i e^{(E_c - E_i)/kT}$$

and

$$n_i = N_v e^{(E_v - E_i)/kT} \quad or \quad N_v = n_i e^{(E_i - E_v)/kT}$$

$$n = n_i e^{(E_f - E_i)/kT}$$

and

$$p = n_i e^{(E_i - E_f)/kT}$$

Developing the mathematical model for electrons and holes

Other useful Relationships: n - p product

$$n_i = N_c e^{(E_i - E_c)/kT} \quad and \quad n_i = N_v e^{(E_v - E_i)/kT}$$

$$n_i^2 = N_c N_v e^{-(E_c - E_v)/kT} = N_c N_v e^{-E_G/kT}$$

$$\boxed{n_i = \sqrt{N_c N_v}\, e^{-E_G/2kT}}$$

Developing the mathematical model for electrons and holes

Other useful Relationships: n - p product

$$Since \quad n = n_i e^{(E_f - E_i)/kT} \qquad and \qquad p = n_i e^{(E_i - E_f)/kT}$$

$$\boxed{np = n_i^2}$$

Known as the ***Law of mass Action***

The number density, i.e., the number of electrons available for conduction in CB is

$$n = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \exp-(\frac{E_C - E_F}{kT})$$

$$n = N_C \exp-(\frac{E_C - E_F}{kT}) \qquad n = n_i \exp(\frac{E_F - E_i}{kT})$$

Extrinsic Carrier Concentrations

- Since $n_o p_o = n_i^2$ for any semiconductor in thermal equilibrium, and
- For an n-type semiconductor, $n_o \approx N_d$

$$p_o = \frac{n_i^2}{N_d}$$

- Where p_o is the concentration of holes in the n-type semiconductor.

The number density, i.e., the number of holes available for conduction in VB is

$$p = 2\left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2} \exp-(\frac{E_F - E_V}{kT})$$

$$p = N_V \exp-(\frac{E_F - E_V}{kT}) \qquad p = n_i \exp(\frac{E_i - E_F}{kT})$$

Extrinsic Carrier Concentrations

- Since $n_o p_o = n_i^2$ for any semiconductor in thermal equilibrium, and
- For a p-type semiconductor, $p_o \approx N_a$

$$n_o = \frac{n_i^2}{N_a}$$

- Where n_o is the concentration of free electrons in the p-type semiconductor.

Extrinsic Carrier Concentrations

- For a p-type semiconductor with acceptor impurities, the concentration of acceptor impurities is N_a with units #/cm^3.
- If $N_a >> n_i$, then the concentration of holes in the p-type semiconductor is approximately $p_o \approx N_a$.

Extrinsic Carrier Concentrations

- For an n-type semiconductor with donor impurities, the concentration of donor impurities is N_d with units #/cm^3.
- If $N_d >> n_i$, then the concentration of free electrons in the n-type semiconductor is approximately $n_o \approx N_d$.

Developing the mathematical model for electrons and holes

Charge Neutrality

- If excess charge existed within the semiconductor, random motion of charge would imply net (AC) current flow. ===> Not possible!
- Thus, all charges within the semiconductor must cancel.

$$q\left[\left(p - N_A^-\right) + \left(N_d^+ - n\right)\right] = 0$$

Mobile + charge

Immobile - charge

Immobile + charge

Mobile - charge

Developing the mathematical model for electrons and holes

Charge Neutrality: Total Ionization Case

N^-_A = Concentration of "ionized" acceptors ~ = N_A

N^+_D = Concentration of "ionized" donors ~ = N_D

$$\left(p - N_A\right) + \left(N_D - n\right) = 0$$

Developing the mathematical model for electrons and holes

Charge Neutrality: Total Ionization Case

$$(p - N_A) + (N_D - n) = 0$$

$$\left(\frac{n_i^2}{n} - N_A\right) + (N_D - n) = 0$$

$$n^2 - n(N_D - N_A) - n_i^2 = 0$$

$$n = \frac{N_D - N_A}{2} + \sqrt{\left(\frac{N_D - N_A}{2}\right)^2 + n_i^2} \quad or \quad p = \frac{N_A - N_D}{2} + \sqrt{\left(\frac{N_A - N_D}{2}\right)^2 + n_i^2}$$

and

$$pn = n_i^2$$

Developing the mathematical model for electrons and holes

If $N_D >> N_A$ and $N_D >> n_i$

$$n \cong N_D \quad and \quad p \cong \frac{n_i^2}{N_D}$$

If $N_A >> N_D$ and $N_A >> n_i$

$$p \cong N_A \quad and \quad n \cong \frac{n_i^2}{N_A}$$

Developing the mathematical model for electrons and holes

Example:

An intrinsic Silicon wafer has 1e10 cm^{-3} holes. When 1e18 cm^{-3} donors are added, what is the new hole concentration?

$$n \cong N_D = 10^{18}\ cm^{-3}$$

$$p = \frac{n_i^2}{n} = \frac{\left(10^{10}\right)^2}{10^{18}}\ cm^{-3} = 100\ cm^{-3}$$

Developing the mathematical model for electrons and holes

Example:

An intrinsic Silicon wafer has 1e10 cm^{-3} holes. When 1e18 cm^{-3} acceptors and 8e17 cm^{-3} donors are added, what is the new hole concentration?

$$p = \frac{1x10^{18} - 8x10^{17}}{2} + \sqrt{\left(\frac{1x10^{18} - 8x10^{17}}{2}\right)^2 + \left(1x10^{10}\right)^2}$$

$$p = 2x10^{17}\ cm^{-3} = N_A - N_D$$

Developing the mathematical model for electrons and holes

Example:

An intrinsic Silicon wafer at 470K has 1e14 cm^{-3} holes. When 1e14 cm^{-3} acceptors are added, what is the new electron and hole concentrations?

$N_D=0$

$$p = \frac{1x10^{14}}{2} + \sqrt{\left(\frac{1x10^{14}}{2}\right)^2 + \left(1x10^{14}\right)^2}$$

$$p = 1.62x10^{14}\ cm^{-3} \neq N_A - N_D$$

$$n = \frac{\left(1x10^{14}\right)^2}{1.62x10^{14}} = 6.2x10^{13}\ cm^{-3}$$

Developing the mathematical model for electrons and holes

Example:

An intrinsic Silicon wafer at 600K has 4e15 cm^{-3} holes. When 1e14 cm^{-3} acceptors are added, what is the new electron and hole concentrations?

N_D=0

$$p = \frac{1x10^{14}}{2} + \sqrt{\left(\frac{1x10^{14}}{2}\right)^2 + \left(4x10^{15}\right)^2}$$

$$p = 4x10^{15}\,cm^{-3} = n_i \neq N_A - N_D$$

$$n = \frac{\left(4x10^{15}\right)^2}{4x10^{15}} = 4x10^{15}\,cm^{-3} = n_i$$

$$\Downarrow$$

Intrinsic Material at High Temperature

Where is E_i?

Since we started with descriptions of intrinsic materials then it makes sense to reference energies from the intrinsic energy, E_i.

Intrinsic Material:

$$n = N_c e^{(E_f - E_c)/kT} = N_v e^{(E_v - E_f)/kT} = p$$

$$N_c e^{(E_i - E_c)/kT} = N_v e^{(E_v - E_i)/kT}$$

$$E_i = \frac{E_c + E_v}{2} + \frac{kT}{2} \ln\left(\frac{N_v}{N_c}\right)$$

Where is E_i?

Intrinsic Material:

But,

$$\frac{N_v}{N_c} = \left(\frac{m_p^*}{m_n^*}\right)^{3/2}$$

$$E_i = \underbrace{\frac{E_c + E_v}{2}}_{\text{Letting } E_v=0,\ \text{this is } E_g/2 \text{ or ``Midgap''}} + \underbrace{\frac{3kT}{4}\ln\left(\frac{m_p^*}{m_n^*}\right)}_{\text{-0.007 eV for Si @ 300K (0.6\% of } E_G \text{)}}$$

Where is E_i?

Extrinsic Material:

$$n = n_i e^{(E_f - E_i)/kT} \qquad p = n_i e^{(E_i - E_f)/kT}$$

Solving for $(E_f - E_i)$

$$E_f - E_i = kT \ln\left(\frac{n}{n_i}\right) = -kT \ln\left(\frac{p}{n_i}\right)$$

or for $N_D >> N_A$ *and* $N_D >> n_i$

$$E_f - E_i = kT \ln\left(\frac{N_D}{n_i}\right)$$

or for $N_A >> N_D$ *and* $N_A >> n_i$

$$E_f - E_i = -kT \ln\left(\frac{N_A}{n_i}\right)$$

Where is E_i?

Extrinsic Material:

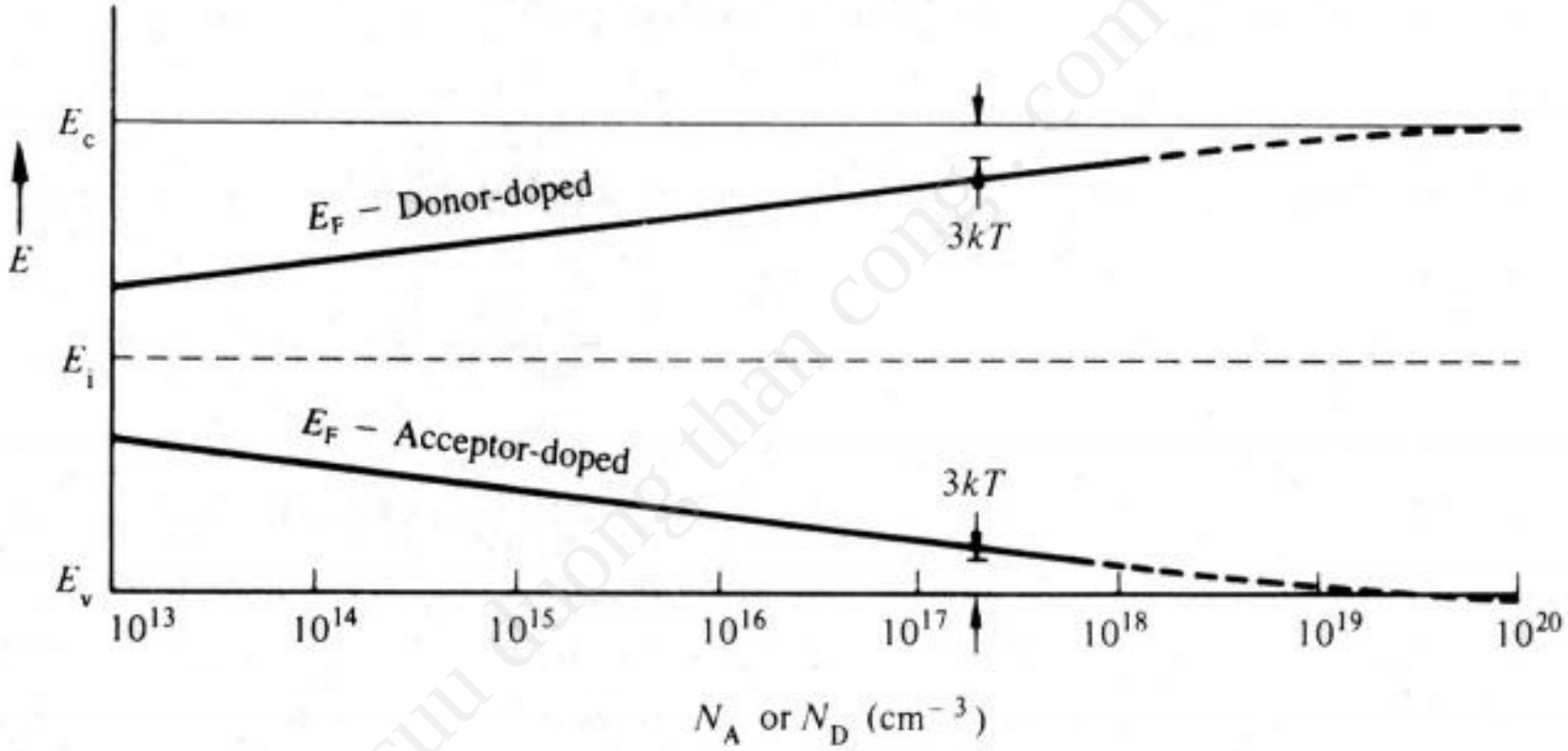


Figure 2.21 Fermi level positioning in Si at 300 K as a function of the doping concentration. The solid E_F lines were established using Eq. (2.38a) for donor-doped material and Eq. (2.38b) for acceptor-doped material ($kT = 0.0259$ eV, and $n_i = 10^{10}/cm^3$).

Majority carriers determine the conductivity of semiconductors

Extrinsic Semicon.	Majority Carrier	Minority Carrier	Conductivity σ
n-type	n_n density of electrons	p_n density of holes	$\sigma = n_n e \mu_n$
p-type	p_p density of holes	n_p density of electrons	$\sigma = p_p e \mu_p$

Current in Semiconductors

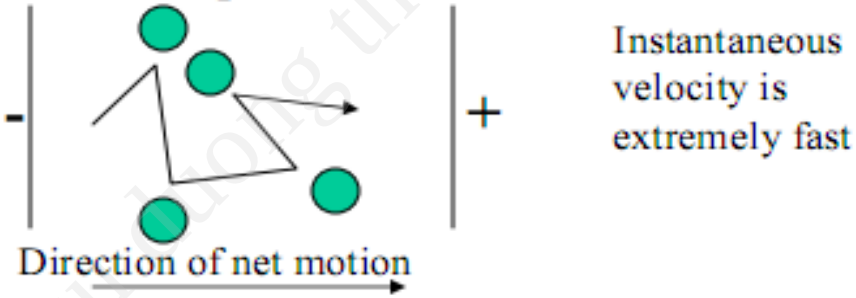
- The two processes that cause free electrons and holes to move in a semiconductor are drift and diffusion.
- Drift – the movement of holes and electrons due to an electric field
- Diffusion – the movement of holes and electrons due to variations in concentrations.

Ways Carriers (electrons and holes) can change concentrations

- Current Flow:
  - Drift: charged particle motion in response to an electric field.
  - Diffusion: Particles tend to spread out or redistribute from areas of high concentration to areas of lower concentration
- Recombination: Local annihilation of electron-hole pairs
- Generation: Local creation of electron-hole pairs

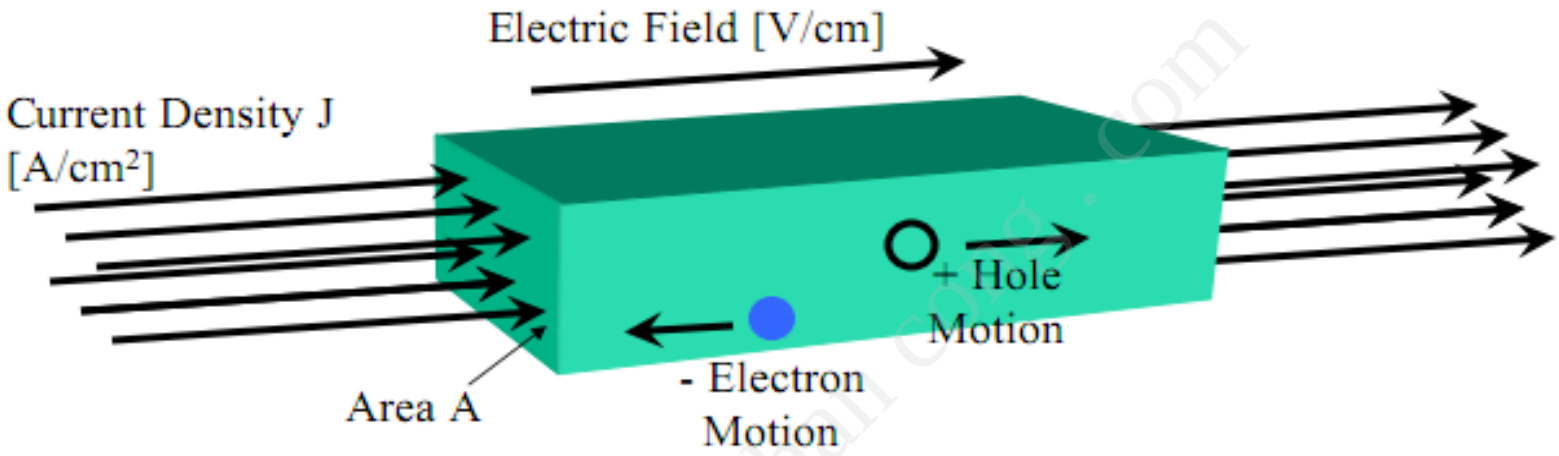
Drift

- Direction of motion:
  - Holes move in the direction of the electric field (from + to -)
  - Electrons move in the opposite direction of the electric field (from - to +)
- Motion is highly non-directional on a local scale, but has a net direction on a macroscopic scale



- Average net motion is described by the drift velocity, v_d with units cm/second
- Net motion of charged particles gives rise to a current

Drift



Given current density J (I=J x Area) flowing in a semiconductor block with face area A under the influence of electric field E, the component of J due to drift of carriers is:

$$J_p|_{Drift} = q\ p\ v_d \quad \text{and} \quad J_n|_{Drift} = q\ n\ v_d$$

Hole Drift current density

Electron Drift current density

Drift

At low electric field values,

$$J_p = qp\mu_p E \quad \text{and} \quad J_n = qn\mu_n E$$

μ is the "mobility" of the semiconductor and measures the ease with which carriers can move through the crystal. $[\mu]=cm^2/V\text{-Second}$

Thus, the drift velocity increases with increasing applied electric field.

More generally, for Silicon and Similar Materials the drift velocity can be empirically given as:

$$v_d = \frac{\mu_o E}{\left[1+\left(\frac{\mu_o E}{v_{sat}}\right)^{\beta}\right]^{1/\beta}} \cong \begin{cases} \mu_o E & when\ E \to 0 \\ v_{sat} & when\ E \to \infty \end{cases}$$

where v_{sat} is the saturation velocity

Drift

Mobility

μis the "mobility" of the semiconductor and measures the ease with which carriers can move through the crystal. [μ]= cm^2/V-Second

μ_n~1360 cm^2/V-Second for Silicon @ 300K

μ_p~460 cm^2/V-Second for Silicon @ 300K

μ_n~8000 cm^2/V-Second for GaAs @ 300K

μ_p~400 cm^2/V-Second for GaAs @ 300K

$$\mu_{n,p} = \frac{q\langle\tau\rangle}{m^*_{n,p}}$$

Where <τ> is the average time between "particle" collisions in the semiconductor.

Collisions can occur with lattice atoms, charged dopant atoms, or with other carriers.

Resistivity and Conductivity

Ohms Law States: $J=\sigma E=E/\rho$

where σ =conductivity [1/ohm-cm] and ρ=resistivity [ohm-cm]

Adding the electron and hole drift currents (at low electric fields),

$$J = J_p|_{Drift} + J_n|_{Drift} = q(\mu_n n + \mu_p p)E$$

Thus,

$$\sigma = q(\mu_n n + \mu_p p) \text{ and } \rho = 1/[q(\mu_n n + \mu_p p)]$$

But since μ_n and μ_p change very little and n and p change several orders of magnitude:

$\sigma \sim= q\mu_n n$ for n-type with n>>p

$\sigma \sim= q\,\mu_p p$ for p-type with p>>n

$$\sigma_i = n_i e(\mu_e + \mu_h) = C \cdot e(\mu_e + \mu_h) \cdot T^{\frac{3}{2}} \exp\left(-\frac{E_G}{2kT}\right)$$

- σ_i depends strongly on the temperature and the charge carrier densities
- extrinsic conductivity depends additionaly on excitation of dopants into the conduction band.

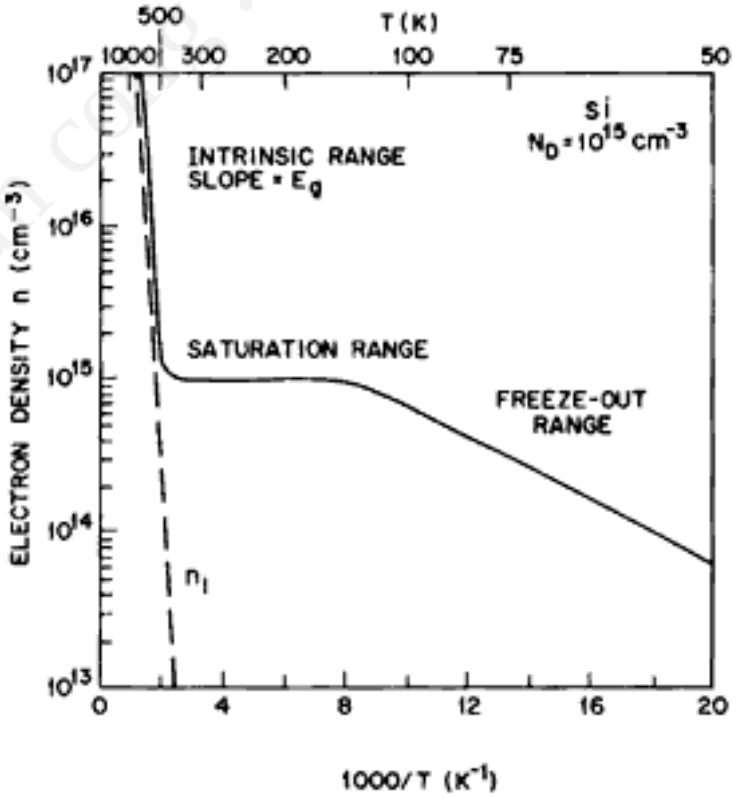
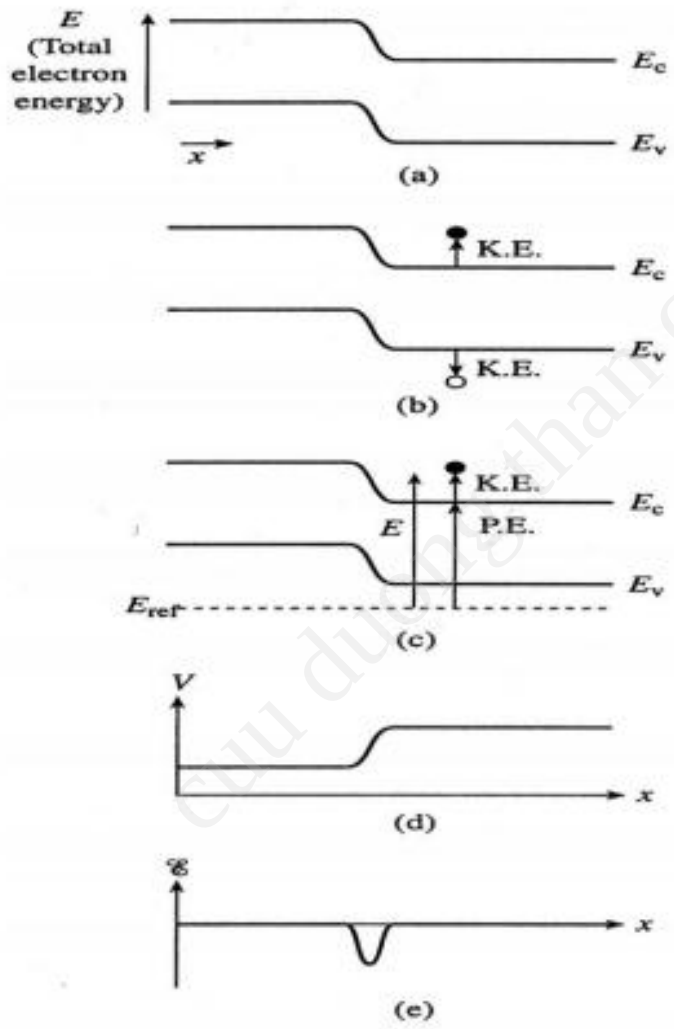


Fig. 16 Electron density as a function of temperature for a Si sample with donor impurity concentration of 10^{15} cm^{-3}. (After Smith, Ref. 5.)

Energy Band Bending under Application of an Electric Field

Band Bending



(a) Sample energy band diagram;

(b) Carrier kinetic energies;

(c) Electronic potential energy;

(d) Potenital;

(e) Electric field.

Diffusion

Nature attempts to reduce concentration gradients to zero.

Example: a bad odor in a room.

In semiconductors, this "flow of carriers" from one region of higher concentration to lower concentration results in a "diffusion current".

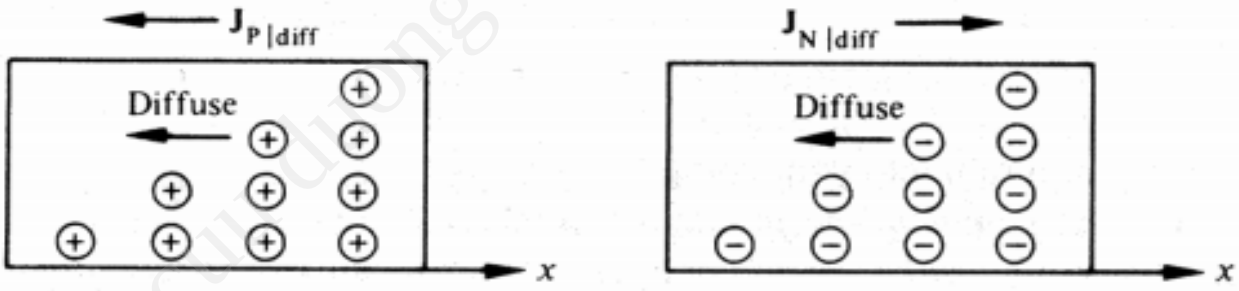


Figure 3.12 Visualization of electron and hole diffusion on a macroscopic scale.

Diffusion

Ficks law describes diffusion as the flux, F, (of particles in our case) is proportional to the gradient in concentration.

$$F = -D\nabla\eta$$

where η is the concentration and D is the diffusion coefficient

Derivation of Ficks Law at http://users.ece.gatech.edu/~gmay/ece3040 lecture #8

For electrons and holes, the diffusion current density (flux of particles times -/+q) can thus, be written as,

$$J_p\big|_{Diffusion} = -qD_p\nabla p \quad or \quad J_n\big|_{Diffusion} = qD_n\nabla n$$

Note in this case, the opposite sign for electrons and holes

Total current

In general, current can flow by drift and diffusion separately. Total current:

$$J_n = J_n^{drift} + J_n^{diff} = qn\mu_n E + qD_n \frac{dn}{dx}$$

$$J_p = J_p^{drift} + J_p^{diff} = qp\mu_p E - qD_p \frac{dp}{dx}$$

And

$$J_{total} = J_n + J_p$$

Einstein relation

Einstein relation

$$\frac{D}{\mu} = \frac{kT}{q}$$

In semiconductors:

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{q}$$

$\frac{kT}{q} \equiv thermal\ voltage\ [V]$

At 300 K: $\frac{kT}{q} \simeq 25\ mV$

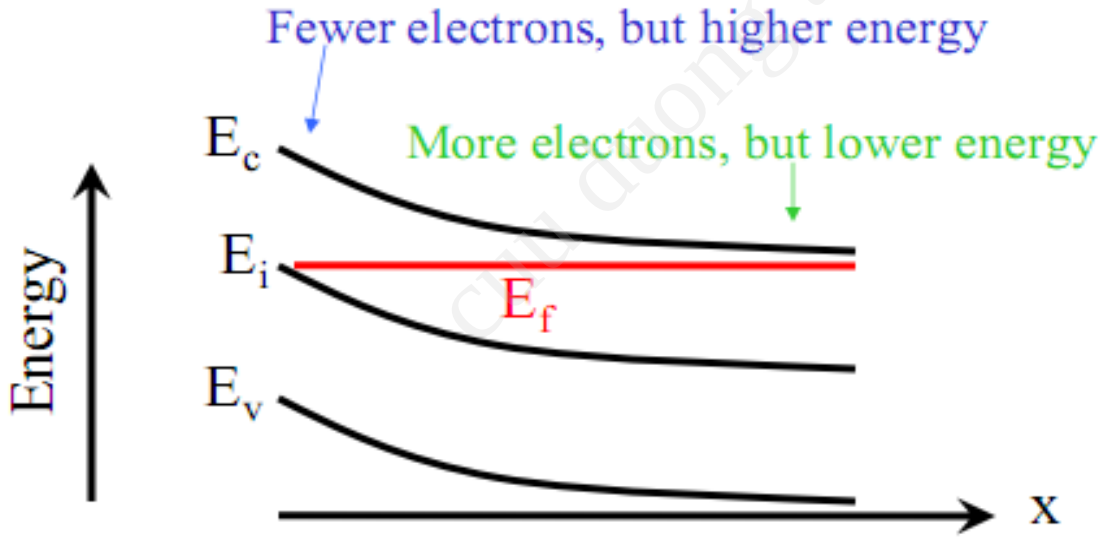
For example: for $N_d = 3 \times 10^{16}\ cm^{-3}$:

$$\mu_n \simeq 1000\ cm^2/V \cdot s \rightarrow D_n \simeq 25\ cm^2/s$$
$$\mu_p \simeq 400\ cm^2/V \cdot s \rightarrow D_p \simeq 10\ cm^2/s$$

Equilibrium Concept

Consider a non-uniformly doped semiconductor.

E_c-E_f varies with position

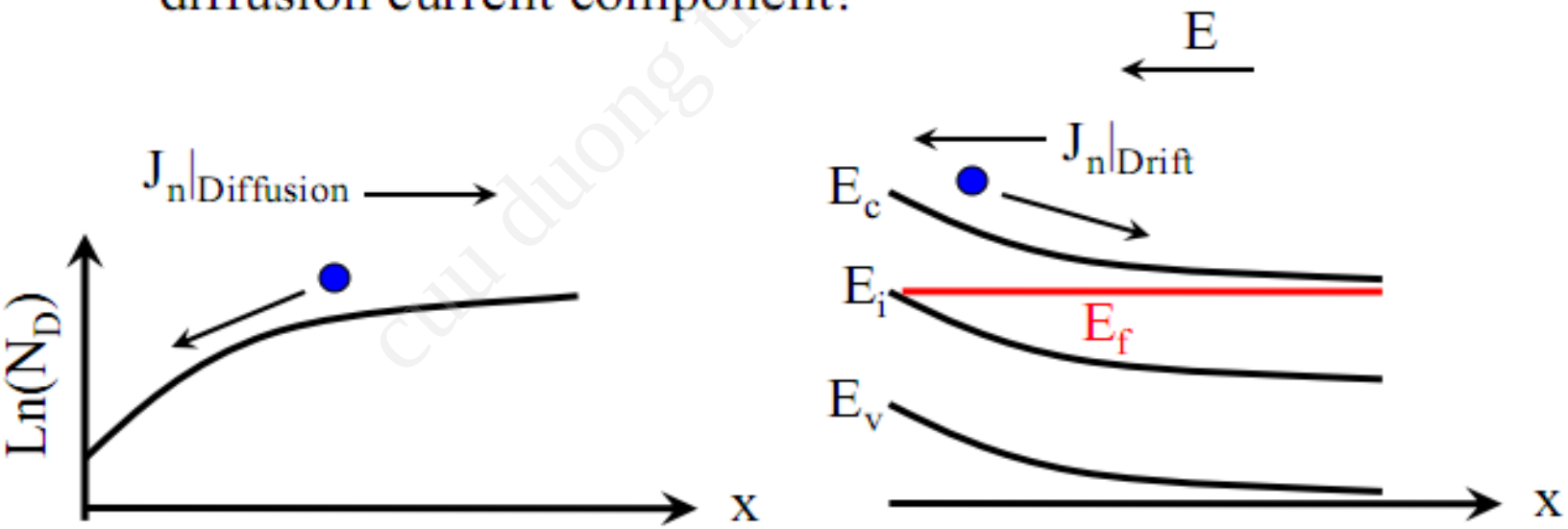


Since the electrons (or holes) are free to move anywhere in the material, the average energy of the electrons can not change. If the average energy did change from one position to another, there would be a net motion of electrons from high energy toward low energy.

E_f must be constant when no current flows!

Equilibrium Concept

- Remember:
  - No net current can flow otherwise we have a "perpetual motion machine".
  - But dEc/dx is nonzero so we have a drift current component.
  - The drift current component MUST be balanced by a diffusion current component!



Equilibrium Concept

Additionally, since electrons and holes operate "independently of each other",

$$J_n|_{Diffusion} + J_n|_{Drift} = 0 \quad \text{and} \quad J_p|_{Diffusion} + J_p|_{Drift} = 0$$

- Thus, for non-uniform doping in equilibrium, we have:
  - E_f is constant
  - No net current
  - Carrier Concentration gradients that result in a diffusion current component.
  - A "Built in" electric field that result in a drift current component.
  - BOTH electron and hole components must sum to zero. I.E. $J_n = J_p = 0$

Equilibrium Concept

Consider the case for electrons:

$$J_{n|Drift} + J_{n|Diffusion} = q\mu_n nE + qD_n \frac{dn}{dx} = 0 \qquad (*)$$

$$but \quad E = \frac{1}{q}\frac{dE_i}{dx} \quad and \quad n = n_i e^{(Ef-Ei)/kT} \quad and \quad \frac{dE_f}{dx} = 0$$

Thus, taking the derivative of n,

$$\frac{dn}{dx} = \frac{n_i}{kT} e^{(Ef-Ei)/kT}\left(\frac{dE_f}{dx} - \frac{dE_i}{dx}\right) = \frac{dn}{dx} = -\frac{n_i}{kT} e^{(Ef-Ei)/kT}\left(\frac{dE_i}{dx}\right)$$

$$= -\frac{q}{kT} nE$$

Thus (*) becomes,

$$\mu_n (qnE) - (qnE) D_n \frac{q}{kT} = 0$$

or,

$$\frac{D_n}{\mu_n} = \frac{kT}{q} \qquad Likewise\ for\ holes, \qquad \frac{D_p}{\mu_p} = \frac{kT}{q}$$

Einstein Relationship

Equilibrium Concept

Other "need to knows"

kT = Energy (thermal energy)

$= (8.617 \times 10^{-5}\ eV/K)\ (T \text{ in } K)$ [eV]

$= (8.617 \times 10^{-5}\ eV/K)\ (1.6 \times 10^{-19}\ J/eV)(T \text{ in } K)$ [J]

kT/q = Voltage (thermal voltage)

= J/coulomb

=J/(J/V)=volts

D_n = Diffusion coefficient [cm^2/second]

Example:

For Si, $\mu_n \sim 1358$ @ 27 C ====>

$D_n = (0.0259\ V)\ (1358\ cm^2/V\text{-second}) = 35.2\ cm^2/Second$

Carrier Mobility

Macroscopic understanding

$$\mu = \frac{V_d}{E}$$

In a perfect Crystal

$$\rho = 0$$

$$\sigma \rightarrow \infty$$

It is a superconductor

Microscopic understanding

$$\mu = \frac{q\tau}{m^*}$$

$$m_e^* < m_h^* \quad in\ general$$

$$m_e^*;\ n-type$$

$$m_h^*;\ p-type$$

- A perfect crystal:

→ a perfect periodicity

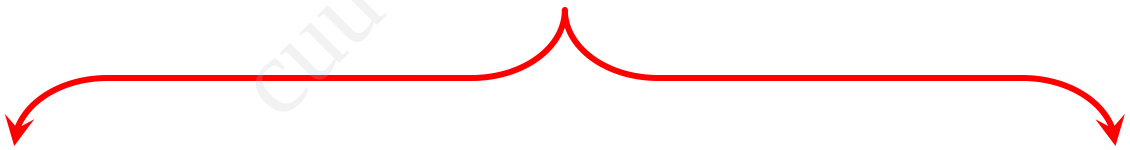
→ the potential seen by a carrier is completely periodic.

→ the crystal has no resistance to current flow and behaves as a superconductor. The perfect periodic potential does not impede the movement of the charge carriers.

- The presence of impurities, interstitials, subtitionals, temperature

→ creates a resistance to current flow.

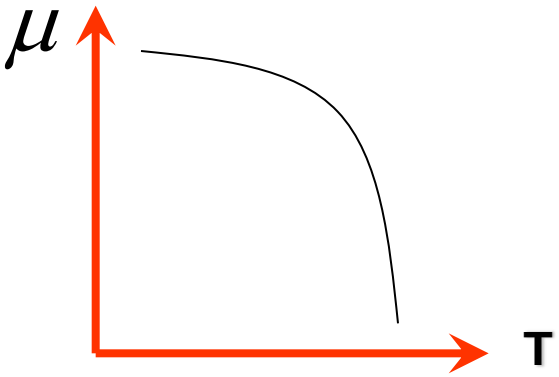
The mobility has two component



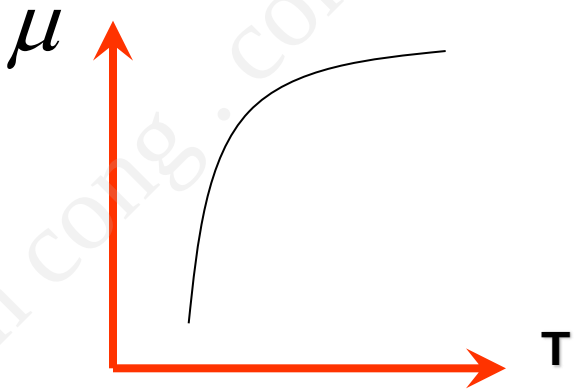
Lattice interaction component

Impurity interaction component

Mobility variation with temperature



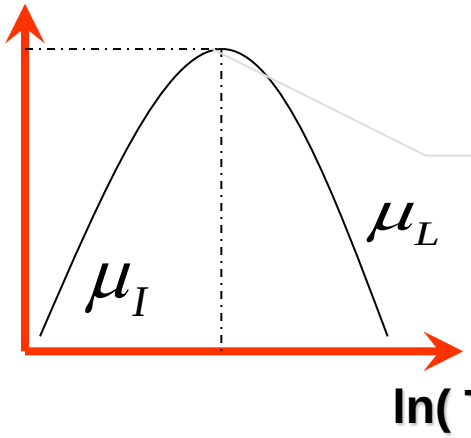
High temperature



Low temperature

$$\frac{1}{\mu_T} = \frac{1}{\mu_L} + \frac{1}{\mu_I}$$

This equation is called as Mattheisen's rule.



Peak depends on the density of impurities

Mobility variation with temperature

- Electron-phonon:
  - Very temperature dependent
  - Phonons are lattice vibrations
  - At low temperatures, lattice is “perfectly still”
- Impurity scattering
  - Temperature independent
  - Depends on impurity concentration

$$\frac{1}{\tau_{total}} = \frac{1}{\tau_{electron-phonon}} + \frac{1}{\tau_{impurity}}$$

Mobility variation with temperature

At high temperature (the lattice warms up) → μ_L component becomes significant.

μ_L decreases when temperature increases.

$$\mu_L = C_1 \times T^{-\frac{3}{2}} \Rightarrow T^{-\frac{3}{2}}$$

C_1 is a constant.

It is called as a $T^{-1.5}$ *power law.*

→ *Carriers are more likely scattered by the lattice atoms*.

Mobility variation with temperature

At low temperatures ⇒ μ_I component is significant.

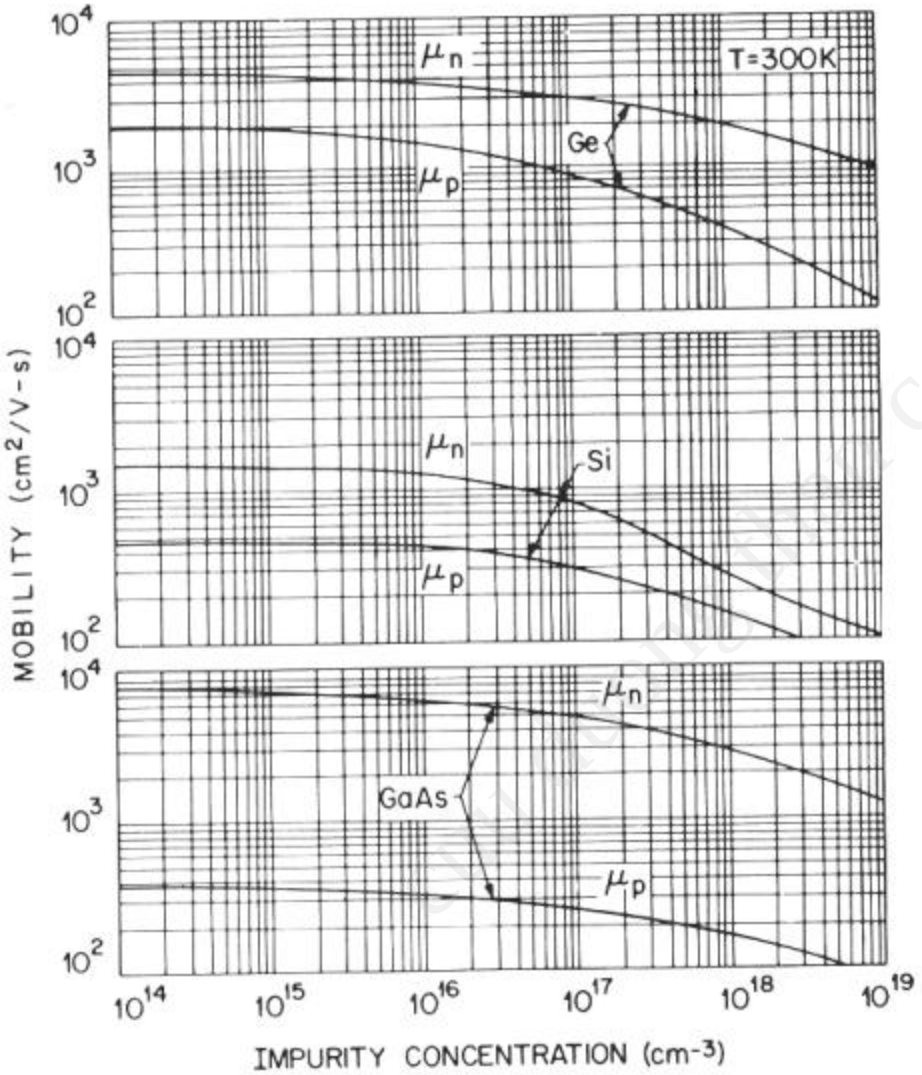
μ_I decreases when temperature decreases.

$$\mu_I = C_2 \times T^{\frac{3}{2}}$$

C_2 is a constant.

→ *Carriers are more likely scattered by ionized impurities*.

- The peak of the mobility curve depends on the number density of ionized impurities.
- Highly doped samples will therefore cause more scattering, and have a lower mobility, than low doped samples.



At low impurity concentration, electron-phonon scattering dominates

At high impurity concentration, impurity scattering dominates.

$$\frac{1}{\tau_{total}} = \frac{1}{\tau_{electron-phonon}} + \frac{1}{\tau_{impurity}}$$

General Classification	Semiconductor Symbol	Semiconductor Name
(1) Elemental	Si	Silicon
	Ge	Germanium
(2) Compounds		
(a) IV–IV	SiC	Silicon carbide
(b) III–V	AlP	Aluminum phosphide
	AlAs	Aluminum arsenide
	AlSb	Aluminum antimonide
	GaN	Gallium nitride
	GaP	Gallium phosphide
	GaAs	Gallium arsenide
	GaSb	Gallium antimonide
	InP	Indium phosphide
	InAs	Indium arsenide
	InSb	Indium antimonide
(c) II–VI	ZnO	Zinc oxide
	ZnS	Zinc sulfide
	ZnSe	Zinc selenide
	ZnTe	Zinc telluride
	CdS	Cadmium sulfide
	CdSe	Cadmium selenide
	CdTe	Cadmium telluride
	HgS	Mercury sulfide
(d) IV–VI	PbS	Lead sulfide
	PbSe	Lead selenide
	PbTe	Lead telluride

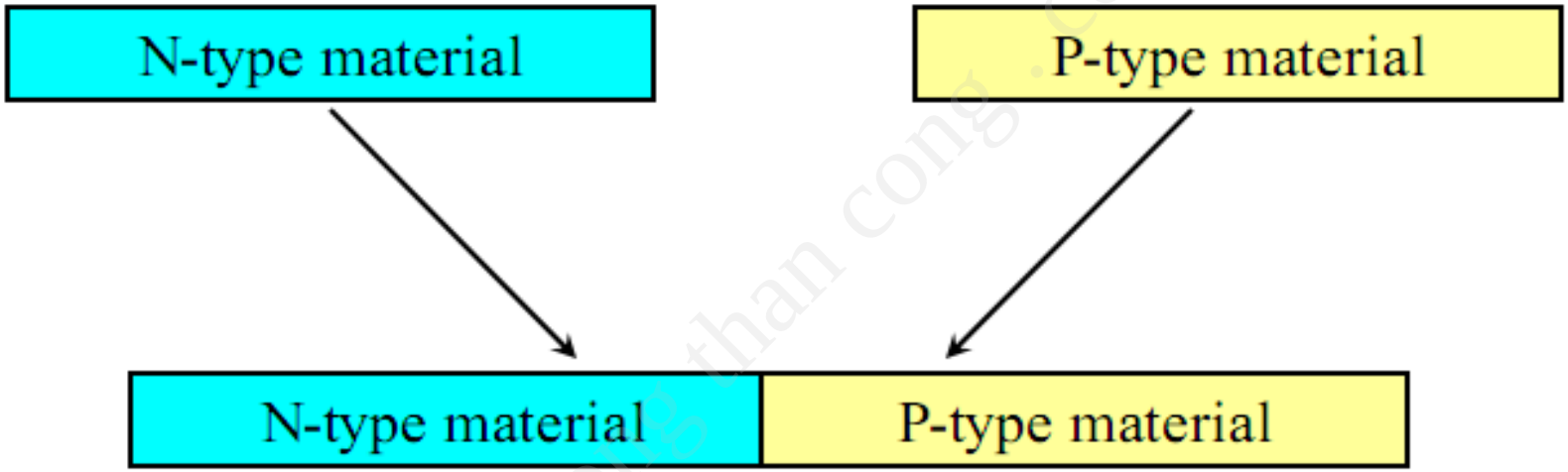
(3) Alloys

(a) Binary	$Si_{1-x}Ge_x$	
(b) Ternary	$Al_xGa_{1-x}As$	(or $Ga_{1-x}Al_xAs$)
	$Al_xIn_{1-x}As$	(or $In_{1-x}Al_xAs$)
	$Cd_{1-x}Mn_xTe$	
	$GaAs_{1-x}P_x$	
	$Ga_xIn_{1-x}As$	(or $In_{1-x}Ga_xAs$)
	$Ga_xIn_{1-x}P$	(or $In_{1-x}Ga_xP$)
	$Hg_{1-x}Cd_xTe$	
(c) Quaternary	$Al_xGa_{1-x}As_ySb_{1-y}$	
	$Ga_xIn_{1-x}As_{1-y}P_y$	

Mobility - Aside

- Note that $\mu_n > \mu_p$.
- Electrons are faster than holes.
- P-type and n-type devices operate the same. However, n-type devices are faster.

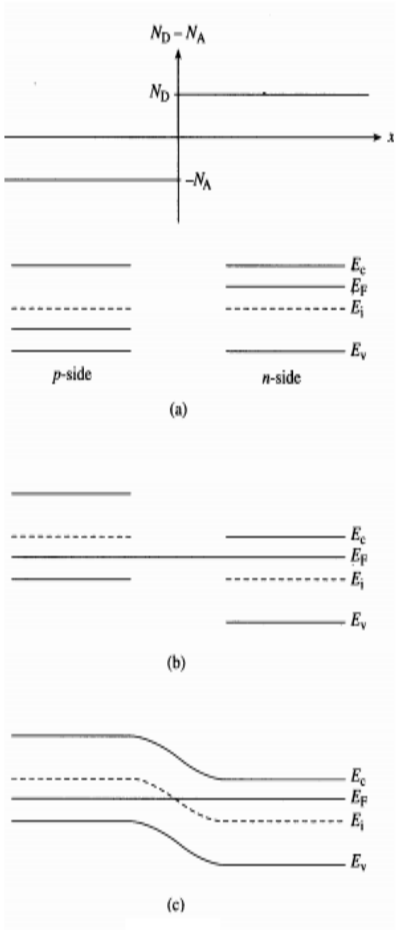
Our First Device: p-n Junction Diode



A p-n junction diode is made by forming a p-type region of material directly next to a n-type region.

Bringing p- and n-Materials Together

Equilibrium energy band diagram for the pn junction



$$n = n_i \exp\left(\frac{E_F - E_i}{kT}\right)$$

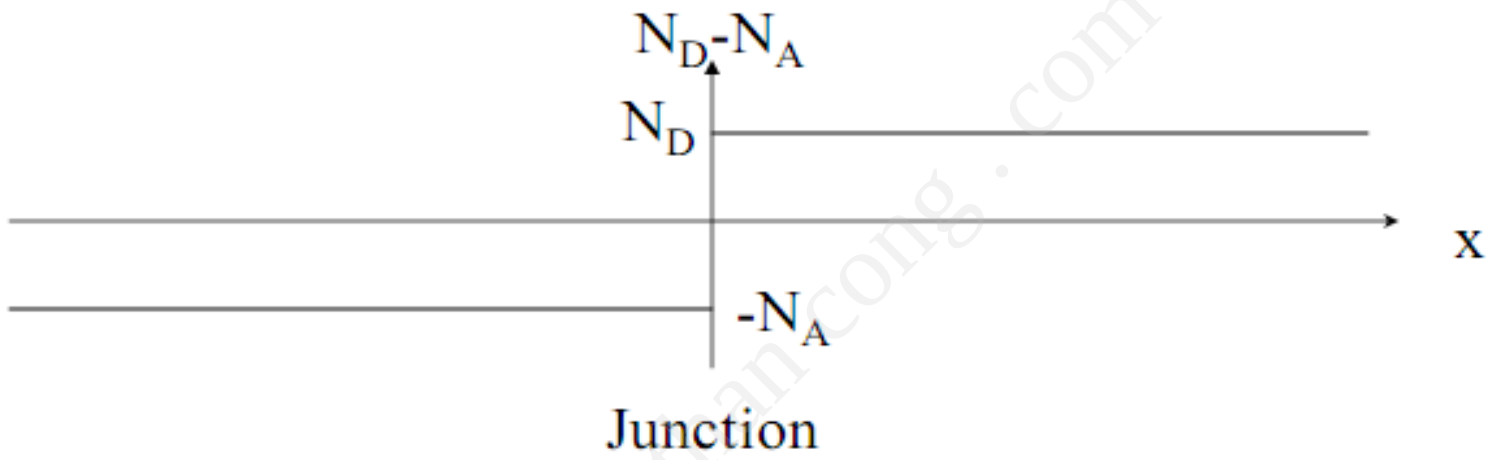
$$p = n_i \exp\left(\frac{E_i - E_F}{kT}\right)$$

E_F = same everywhere under equilibrium

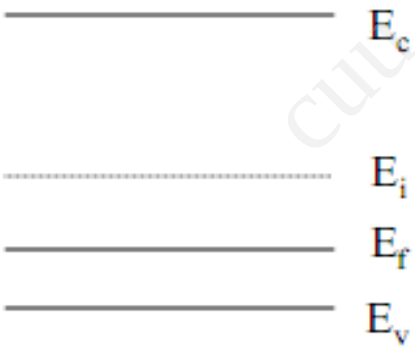
Join the two sides of the band by a smooth curve.

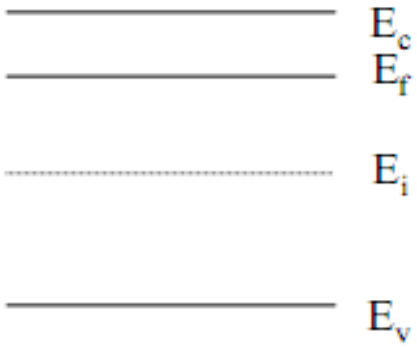
- Large concentration gradient at boundary
- Gradient causes electrons to diffuse from n to p and holes to diffuse from p to n
- p-type material is negatively charged with fixed acceptor ions in the vicinity of pn-boundary.
- n-type material is positively charged with fixed donor ions near the junction
- Electric field across the boundary directed from n to p-region
- Counteracts the diffusion of holes and electrons as it causes electrons to drift from p to n and holes to drift from n to p

Our First Device: p-n Junction Diode



In regions far away from the "junction" the band diagram looks like:





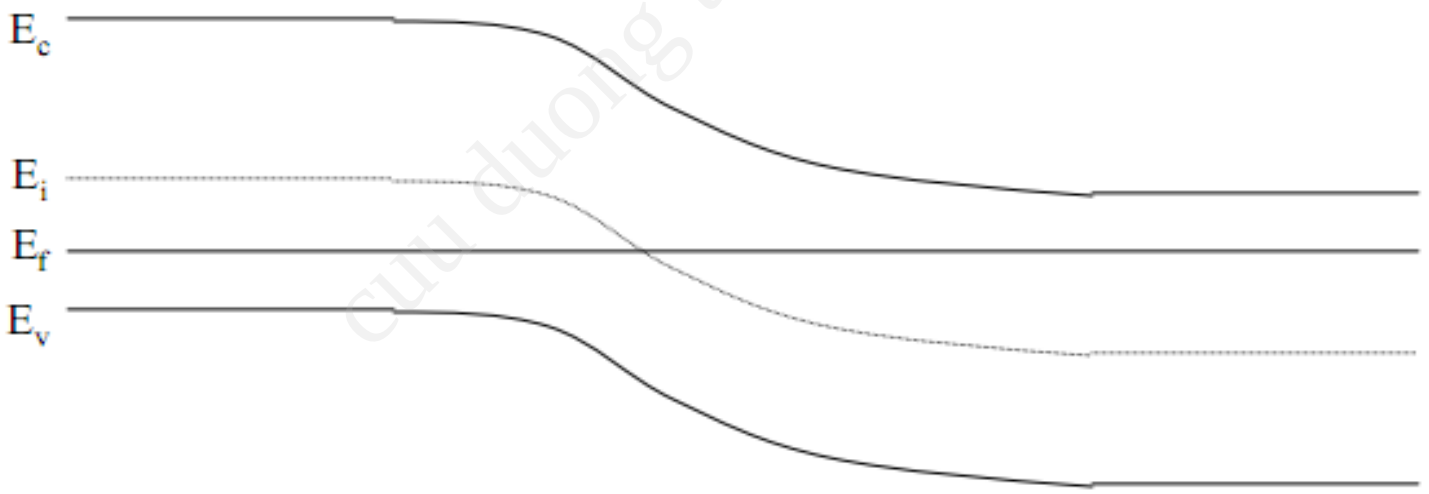
Our First Device: p-n Junction Diode

But when the device has no external applied forces, no current can flow. Thus, the fermi-level must be flat!

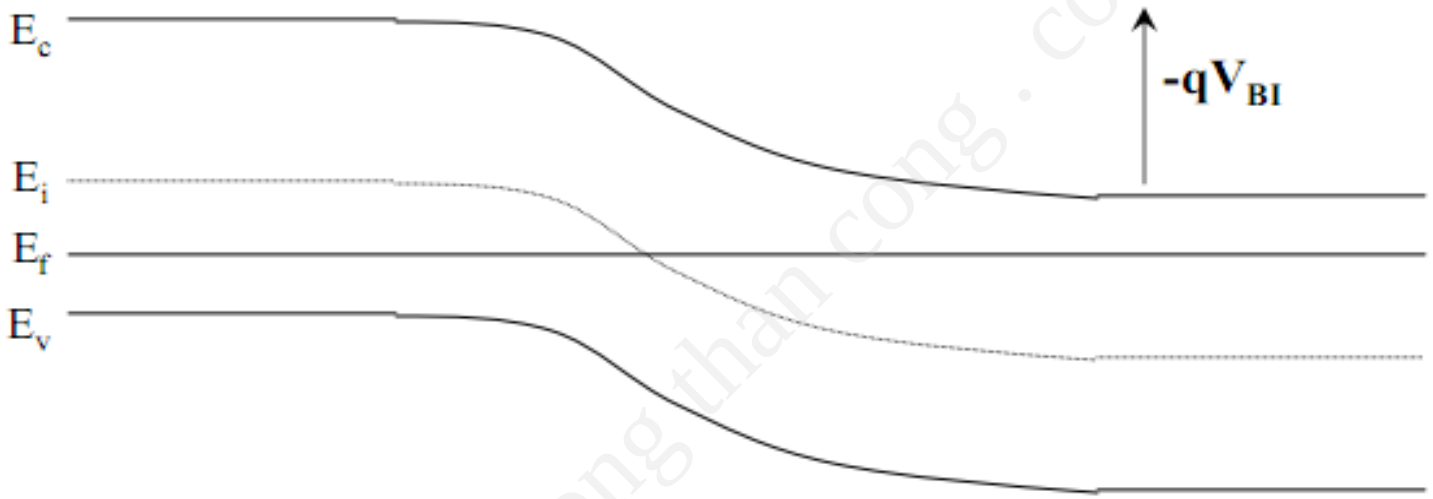
We can then fill in the junction region of the band diagram as:

or…

Our First Device: p-n Junction Diode

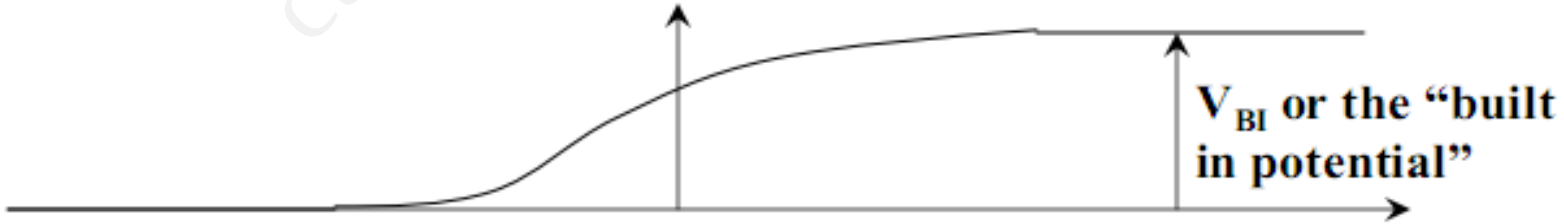


Our First Device: p-n Junction Diode

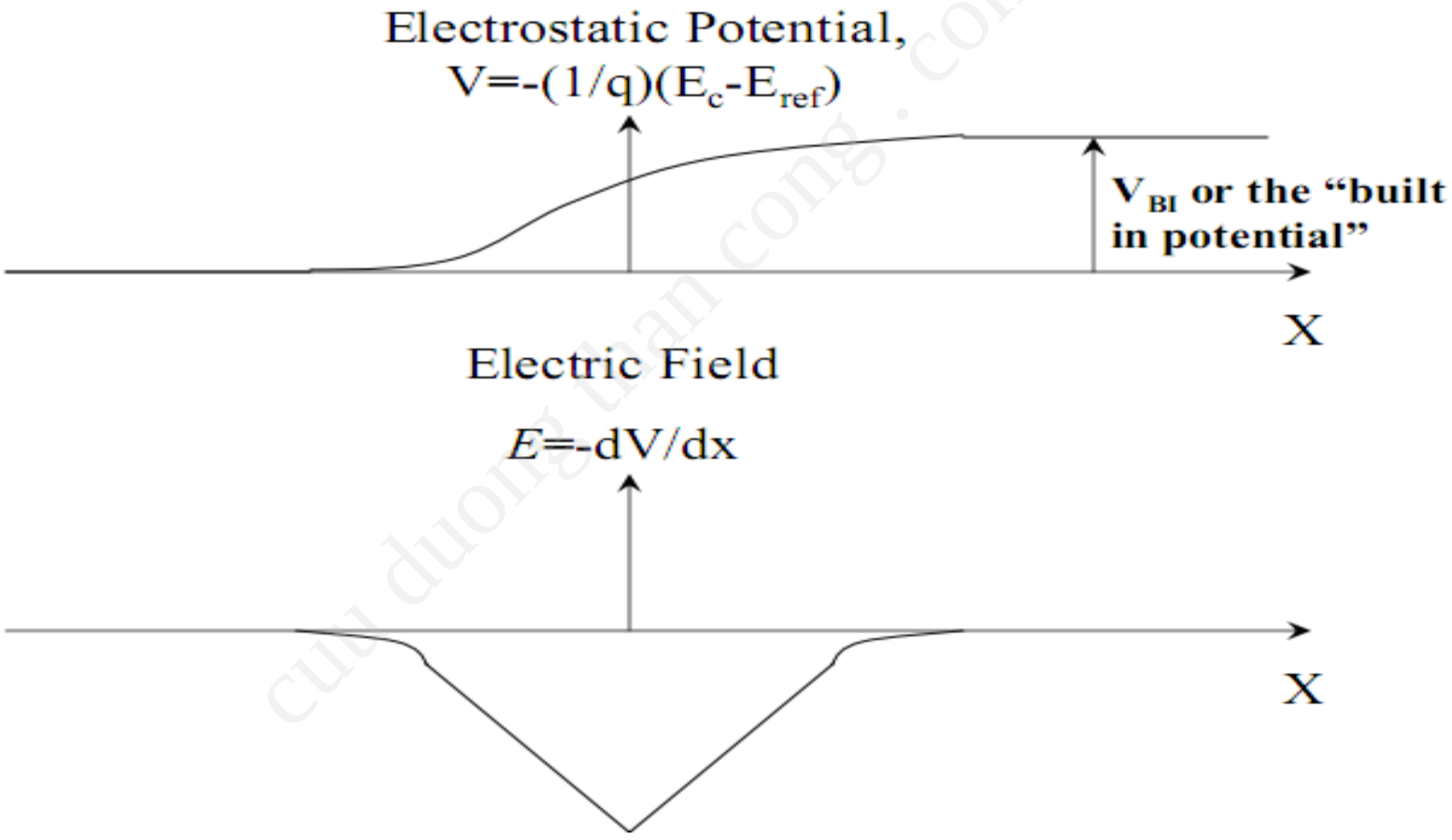


Electrostatic Potential,

$V = -(1/q)(E_c - E_{ref})$



Our First Device: p-n Junction Diode



Our First Device: p-n Junction Diode

Poisson's Equation:

Electric Field **Charge Density (NOT resistivity)**

$$\nabla \bullet E = \frac{\rho}{K_s \varepsilon_o} \quad or\ in\ 1D, \frac{dE}{dx} = \frac{\rho}{K_s \varepsilon_o}$$

Permittivity of free space

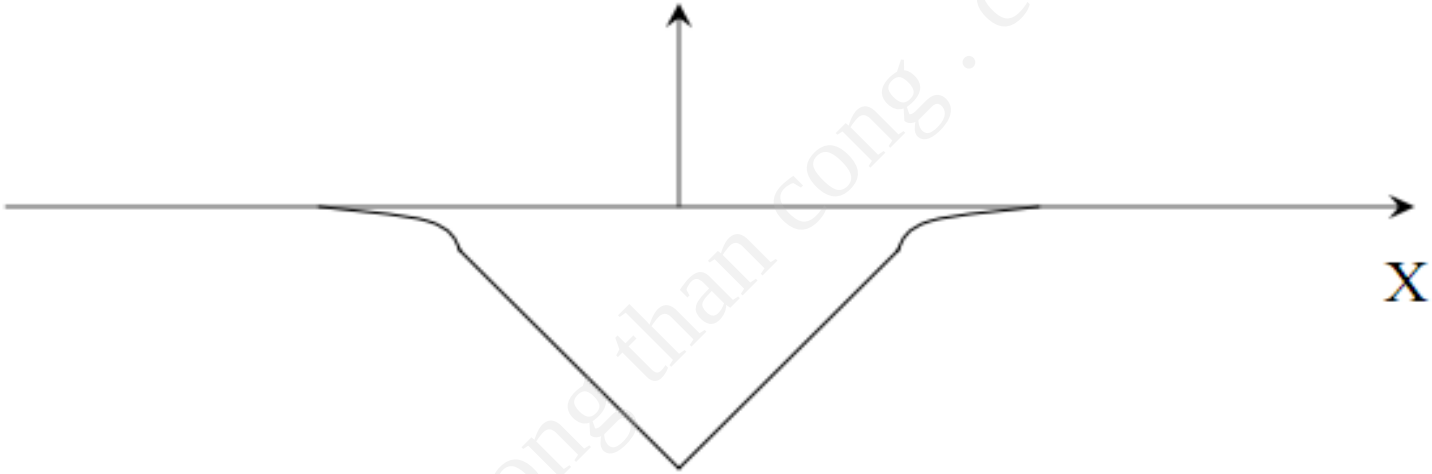
Relative Permittivity of Semiconductor

(previously referred to as ε_R)

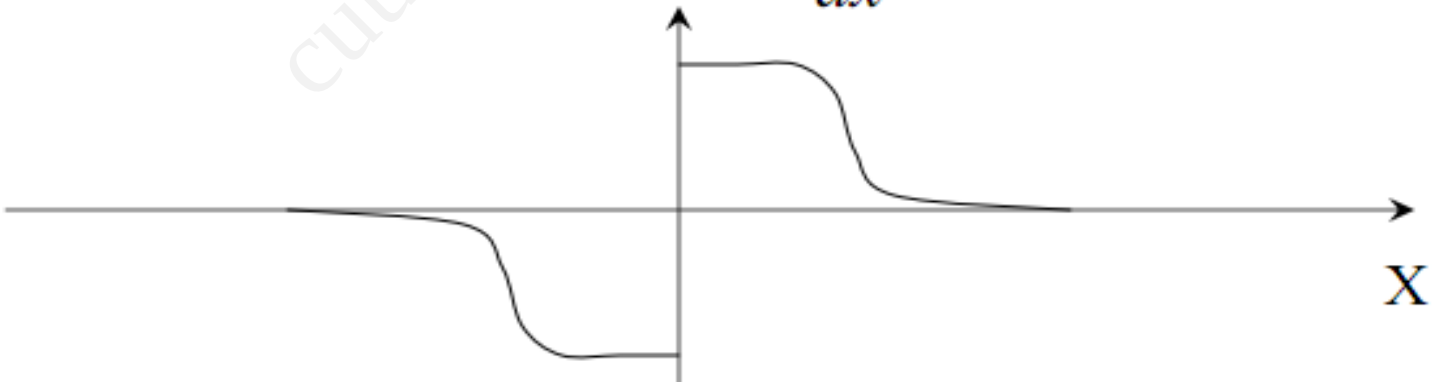
$$\rho = q(p - n + N_D - N_A)$$

Our First Device: p-n Junction Diode

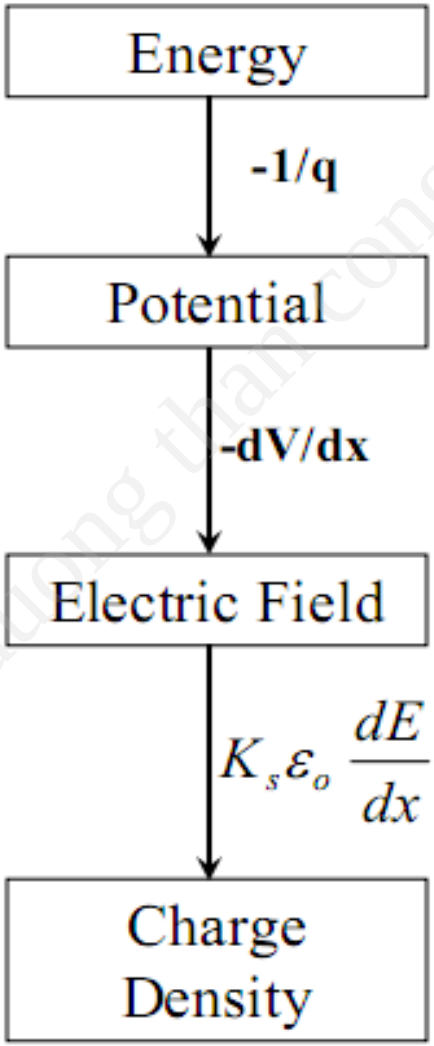
Electric Field, E=-dV/dx



$$\rho = K_s \varepsilon_o \frac{dE}{dx}$$

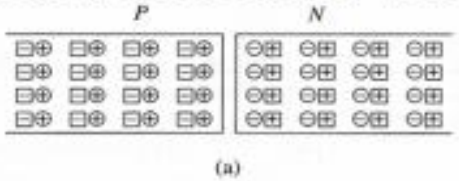


Our First Device: p-n Junction Diode

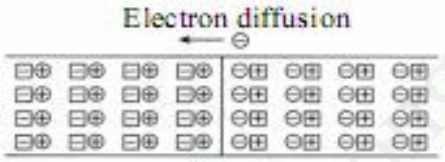


Movement of electrons and holes when forming the junction

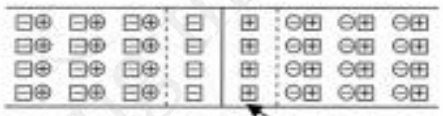
- ○ Circles are charges free to move (electrons and holes)
- □ Squares are charges NOT free to move (ionized donor or acceptor atoms)



(a)



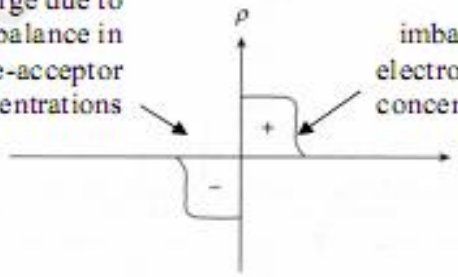
(b)



(c)

Local region of negative charge due to imbalance in hole-acceptor concentrations

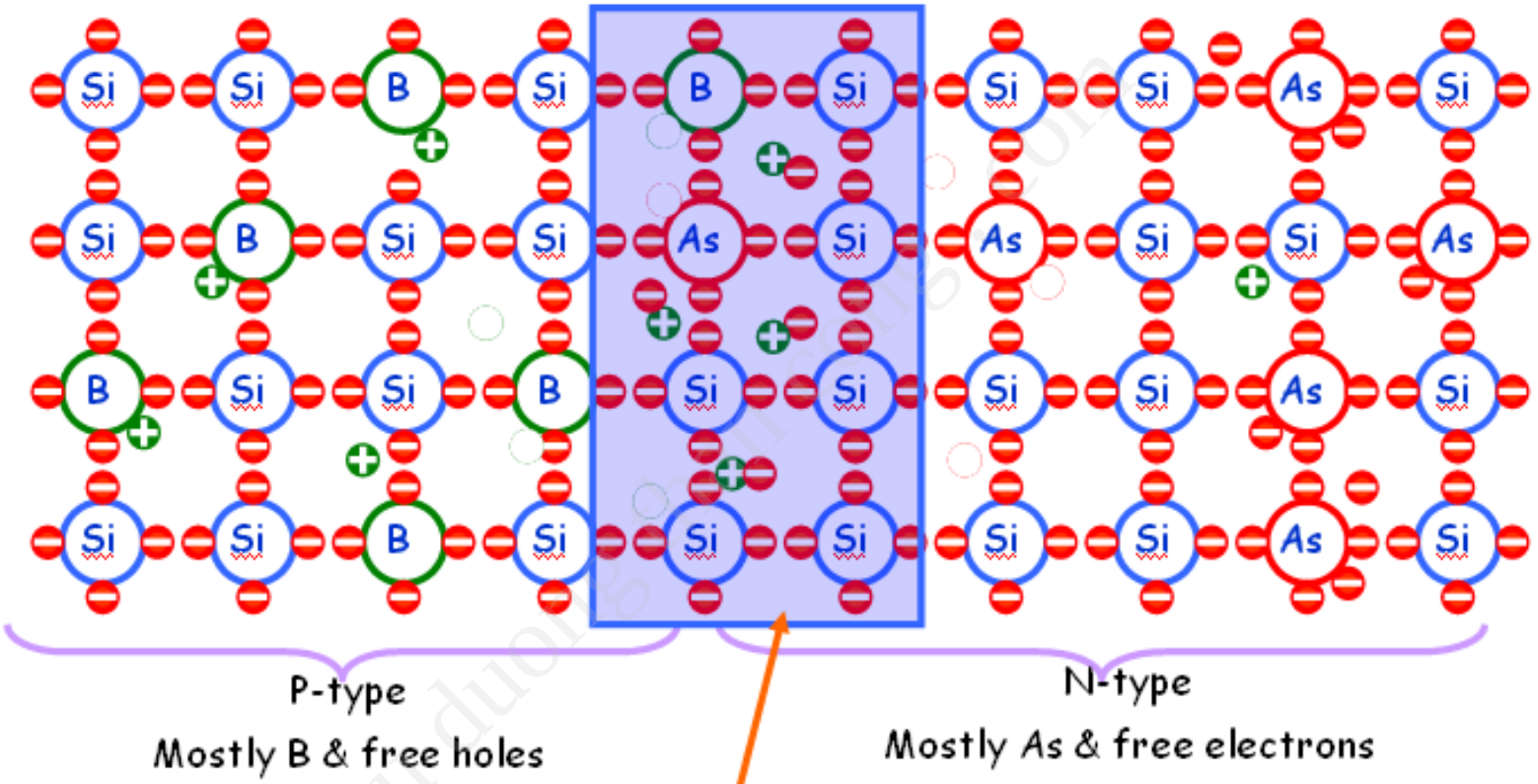
Local region of positive charge due to imbalance in electron-donor concentrations



(d)

Space Charge or Depletion Region

PN Junction: No electrical bias applied

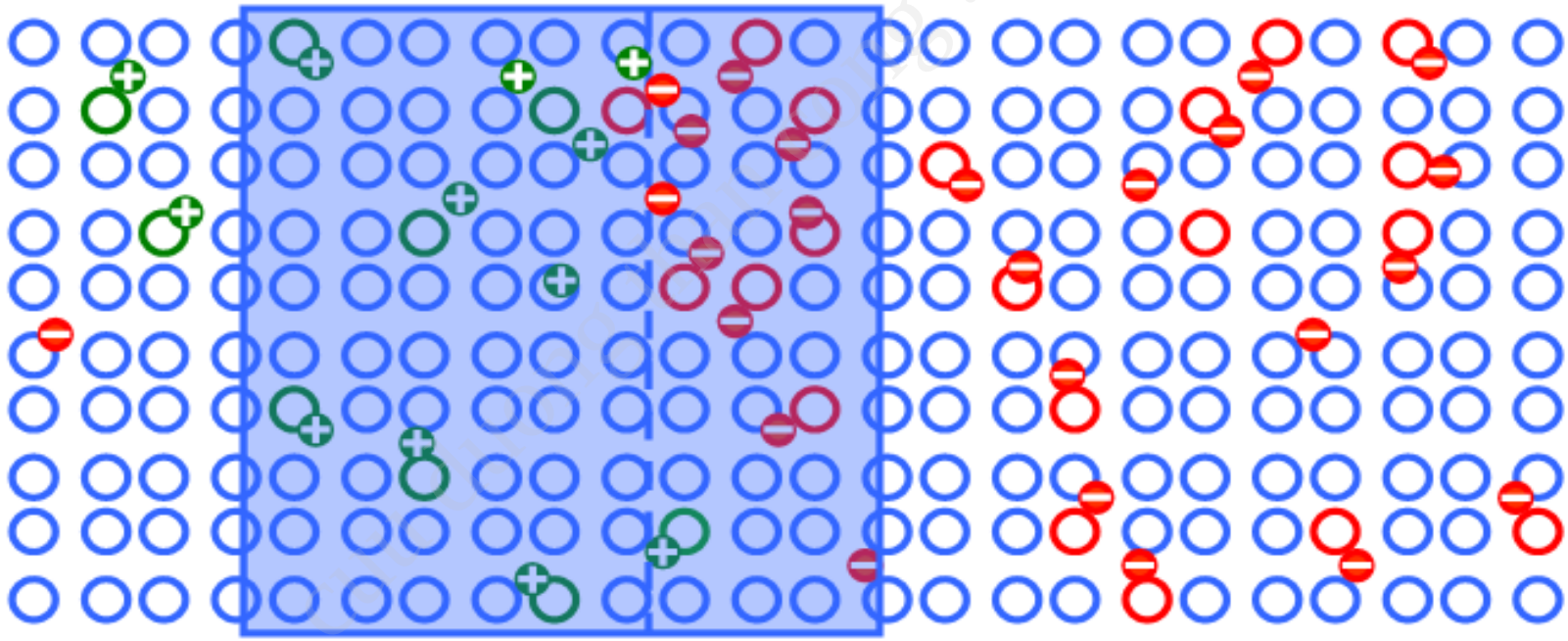


- Diffusion effects – The holes and electrons move from area of high concentration to areas of low concentration.
- Holes & electrons annihilate each other to form an area depleted of free charge. This is known as the depletion region and blocks any further flow of charge carriers across the junction

PN Junction: Asymmetric junction

Si Atom
As Atom
B atom
Free electron
Free hole

- An asymmetric junction occurs where the dopant concentration on one side is higher than on the other.
- Here, there are more As atoms than B atoms.
- Electrons & holes must annihilate in pairs



There are more electrons than holes near the junction. This means the electrons will have to travel deeper into the P-type region to create pairs reducing the width of the depletion region on that size

Movement of electrons and holes when forming the junction

$$E = -\,dV/dx$$

$$-E\,dx = dV$$

$$-\int_{-x_p}^{x_n} E\,dx = \int_{V(-x_p)}^{V(x_n)} dV \quad = \quad V(x_n) - V(-x_p) = V_{bi}$$

but...

$$J_N = q\mu_n nE + qD_N \frac{dn}{dx} = 0$$

← No net current flow in equilibrium

$$E = -\frac{D_N}{\mu_n}\frac{\frac{dn}{dx}}{n} = -\frac{kT}{q}\frac{\frac{dn}{dx}}{n}$$

thus...

$$V_{bi} = -\int_{-x_p}^{x_n} E\,dx = \frac{kT}{q}\int_{n(-x_p)}^{n(x_n)} \frac{\frac{dn}{dx}}{n}dx = \frac{kT}{q}\ln\left[\frac{n(x_n)}{n(-x_p)}\right]$$

Movement of electrons and holes when forming the junction

$$V_{bi} = \frac{kT}{q}\ln\left[\frac{n(x_n)}{n(-x_p)}\right] = \frac{kT}{q}\ln\left[\frac{N_D}{n_i^2 / N_A}\right]$$

$$\boxed{V_{bi} = \frac{kT}{q}\ln\left[\frac{N_A N_D}{n_i^2}\right]}$$

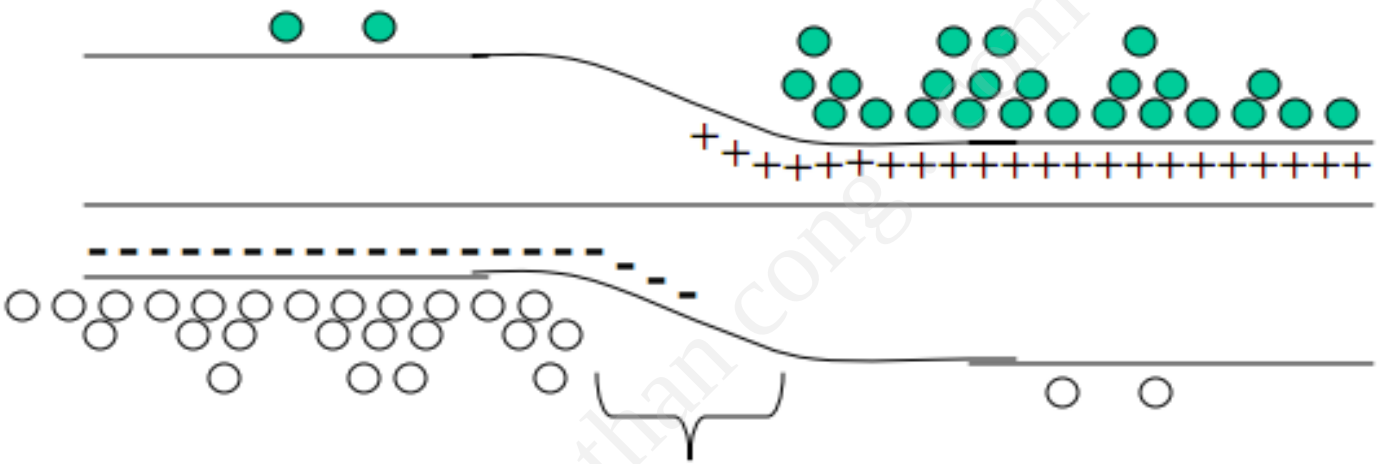
For $N_A=N_D=10^{15}/cm^{-3}$ in silicon at room temperature, V_{bi}~0.6 V*

For a non-degenerate semiconductor, $|-qV_{bi}|<|E_g|$

*Note to those familiar with a diode turn on voltage: This is not the diode turn on voltage! This is the voltage required to reach a flat band diagram and sets an upper limit (typically an overestimate) for the voltage that can be applied to a diode before it burns itself up.

Movement of electrons and holes when forming the junction

Depletion Region Approximation



Depletion Region Approximation states that approximately no free carriers exist in the space charge region and no net charge exists outside of the depletion region (known as the quasi-neutral region). Thus,

$$\frac{dE}{dx} = \frac{\rho}{K_S \varepsilon_o} = \frac{q}{K_S \varepsilon_o}(p - n + N_D - N_A) = 0 \quad within\ the\ quasi-neutral\ region$$

becomes...

$$\frac{dE}{dx} = \frac{q}{K_S \varepsilon_o}(N_D - N_A) \quad within\ the\ space\ charge\ region$$

Movement of electrons and holes when forming the junction

Depletion Region Approximation: Step Junction Solution

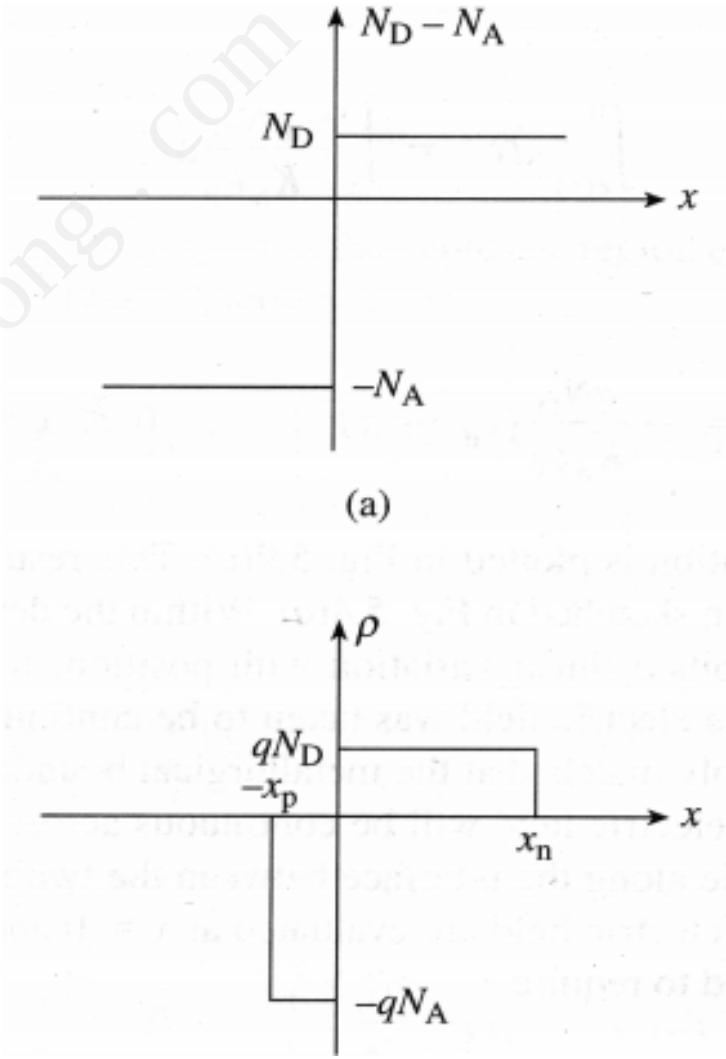
$$\rho = \begin{cases} -qN_A & for\ -x_p \le x \le 0 \\ qN_D & for\ 0 \le x \le x_n \\ 0 & for\ x \le -x_p\ and\ x \ge x_n \end{cases}$$

thus,

$$\frac{dE}{dx} = \begin{cases} \dfrac{-qN_A}{K_S\varepsilon_o} & for\ -x_p \le x \le 0 \\ \dfrac{qN_D}{K_S\varepsilon_o} & for\ 0 \le x \le x_n \\ 0 & for\ x \le -x_p\ and\ x \ge x_n \end{cases}$$

Where we have used:

$$\frac{dE}{dx} = \frac{\rho}{K_S\varepsilon_o}$$



Movement of electrons and holes when forming the junction

Depletion Region Approximation: Step Junction Solution

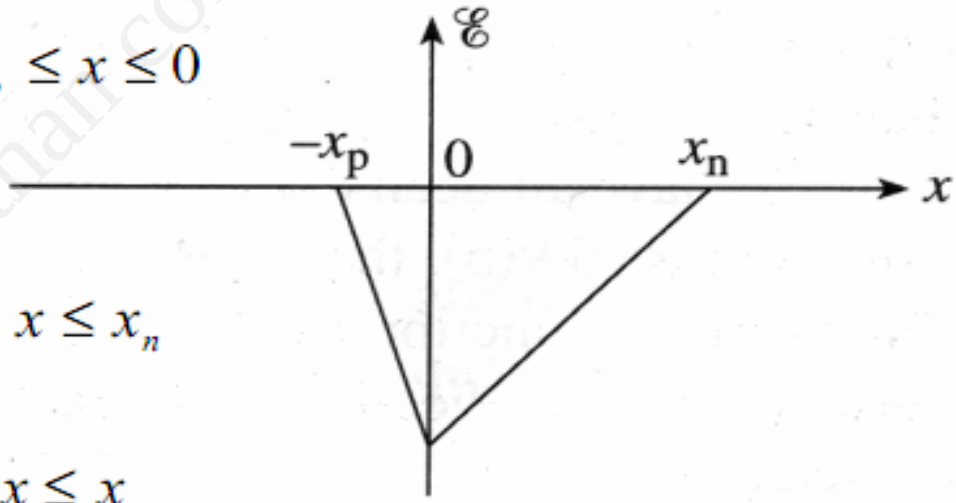
$$\int_0^{E(x)} dE' = \int_{-x_p}^{x} \frac{-qN_A}{K_S\varepsilon_o} dx' \quad for\ -x_p \le x \le 0$$

$$E(x) = \frac{-qN_A}{K_S\varepsilon_o}\left(x + x_p\right) \quad for\ -x_p \le x \le 0$$

and

$$\int_{E(x)}^{0} dE' = \int_{x}^{x_n} \frac{qN_D}{K_S\varepsilon_o} dx' \quad for\ 0 \le x \le x_n$$

$$E(x) = \frac{-qN_D}{K_S\varepsilon_o}\left(x_n - x\right) \quad for\ 0 \le x \le x_n$$



Since $E(x=0^-)=E(x=0^+)$

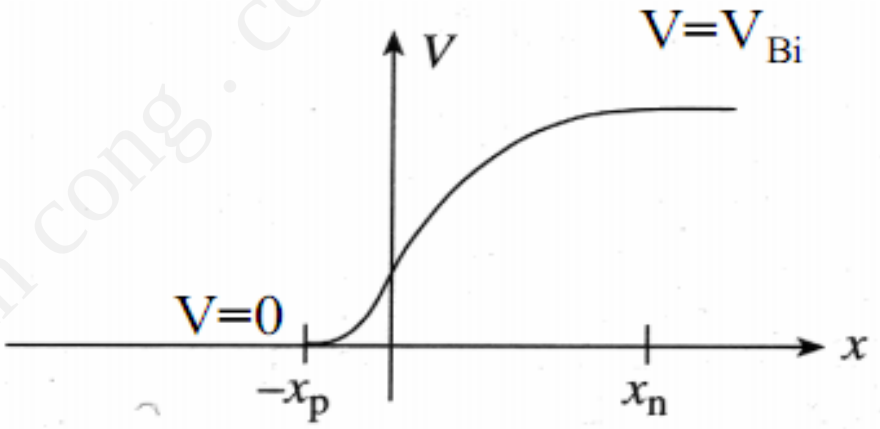
$$N_A x_p = N_D x_n$$

Movement of electrons and holes when forming the junction

Depletion Region Approximation: Step Junction Solution

$$E = -\frac{dV}{dx}$$

$$\frac{dV}{dx} = \begin{cases} \dfrac{qN_A}{K_S\varepsilon_o}(x_p + x) & for\ -x_p \le x \le 0 \\ \dfrac{qN_D}{K_S\varepsilon_o}(x_n - x) & for\ 0 \le x \le x_n \end{cases}$$



or,

$$\int_0^{V(x)} dV' = \int_{-x_p}^{x} \frac{qN_A}{K_S\varepsilon_o}(x_p + x')dx' \quad for\ -x_p \le x \le 0$$

$$\int_{V(x)}^{V_{Bi}} dV' = \int_{x}^{x_n} \frac{qN_D}{K_S\varepsilon_o}(x_n - x')dx' \quad for\ 0 \le x \le x_n$$

$$V(x) = \begin{cases} \dfrac{qN_A}{2K_S\varepsilon_o}(x_p + x)^2 & for\ -x_p \le x \le 0 \\ V_{bi} - \dfrac{qN_D}{2K_S\varepsilon_o}(x_n - x)^2 & for\ 0 \le x \le x_n \end{cases}$$

Movement of electrons and holes when forming the junction

Depletion Region Approximation: Step Junction Solution

At x=0,

$$\frac{qN_A}{2K_S\varepsilon_o}(x_p)^2 = V_{bi} - \frac{qN_D}{2K_S\varepsilon_o}(x_n)^2$$

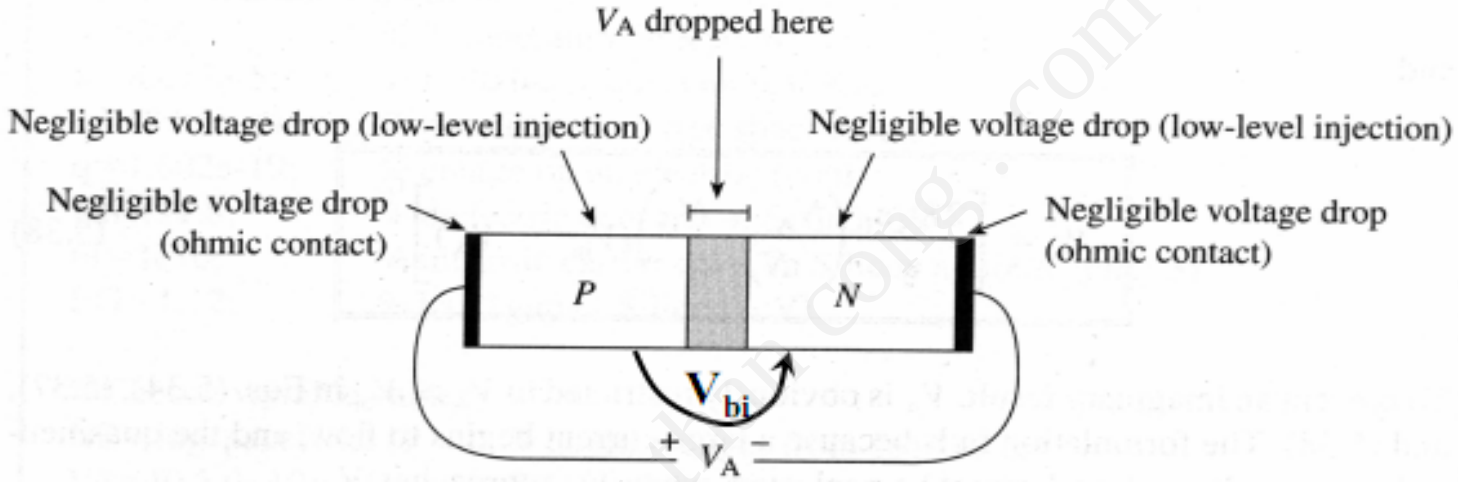
$$U\sin g,\ x_p = \left(x_n N_D\right)\Big/ N_A$$

$$x_n = \sqrt{\frac{2K_S\varepsilon_o}{q}\frac{N_A}{N_D(N_A+N_D)}V_{bi}} \quad and \quad x_p = \sqrt{\frac{2K_S\varepsilon_o}{q}\frac{N_D}{N_A(N_A+N_D)}V_{bi}}$$

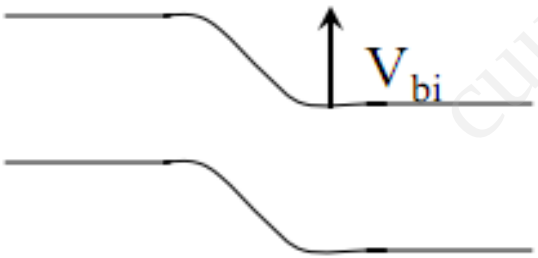
$$W = x_p + x_n = \sqrt{\frac{2K_S\varepsilon_o}{q}\frac{(N_A+N_D)}{N_A N_D}V_{bi}}$$

Movement of electrons and holes when forming the junction

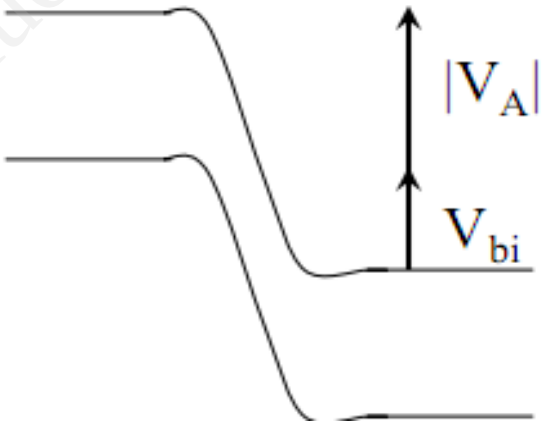
Depletion Region Approximation: Step Junction Solution



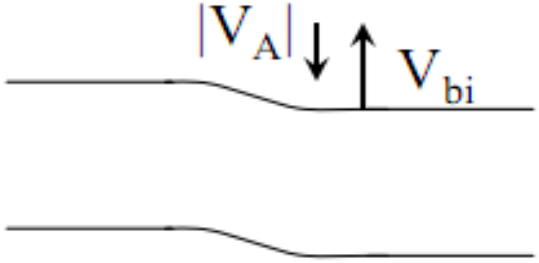
V_A=0 : No Bias



V_A<0 : Reverse Bias



V_A>0 : Forward Bias



Movement of electrons and holes when forming the junction

Depletion Region Approximation: Step Junction Solution

Thus, only the boundary conditions change resulting in direct replacement of V_{bi} with $(V_{bi}-V_A)$

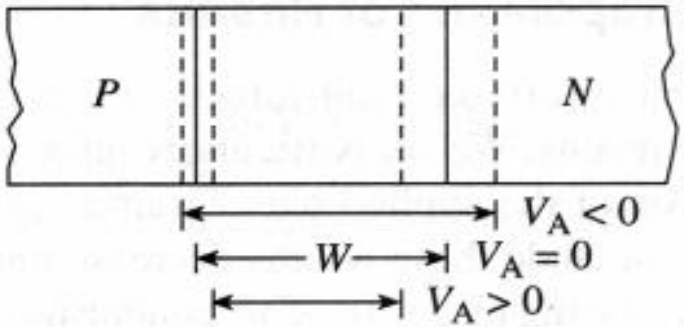
$$x_n = \sqrt{\frac{2K_S\varepsilon_o}{q}\frac{N_A}{N_D(N_A+N_D)}(V_{bi}-V_A)} \quad and \quad x_p = \sqrt{\frac{2K_S\varepsilon_o}{q}\frac{N_D}{N_A(N_A+N_D)}(V_{bi}-V_A)}$$

$$W = x_p + x_n = \sqrt{\frac{2K_S\varepsilon_o}{q}\frac{(N_A+N_D)}{N_AN_D}(V_{bi}-V_A)}$$

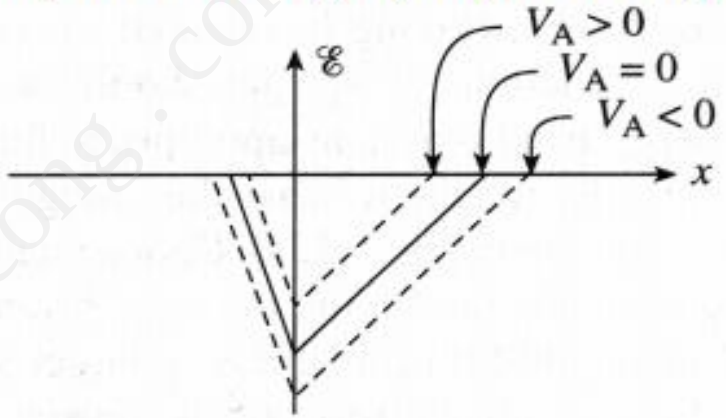
Movement of electrons and holes when forming the junction

Step Junction Solution: What does it mean?

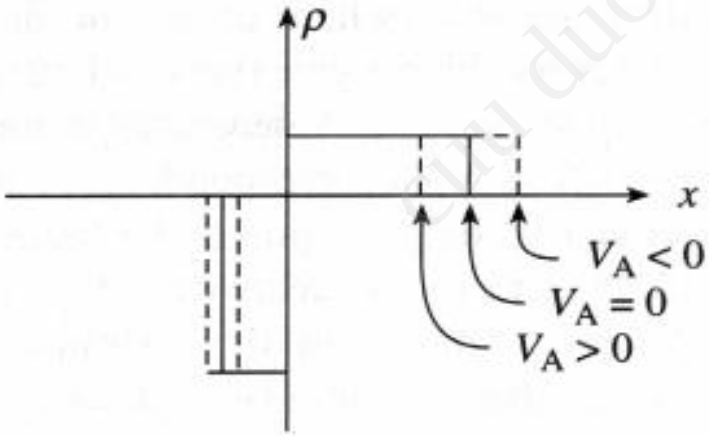
Consider a p^+ -n junction (heavily doped p-side, normal or lightly doped n side).

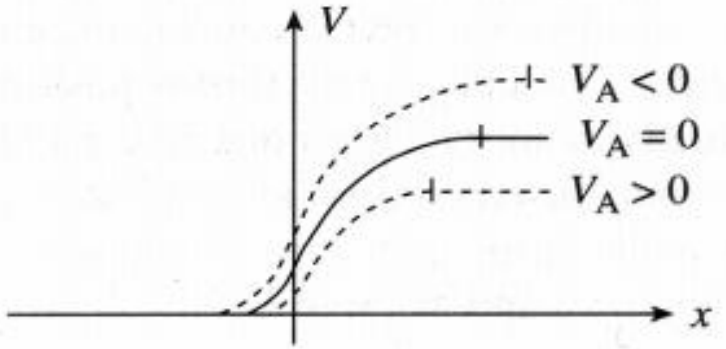


(a)



(c)

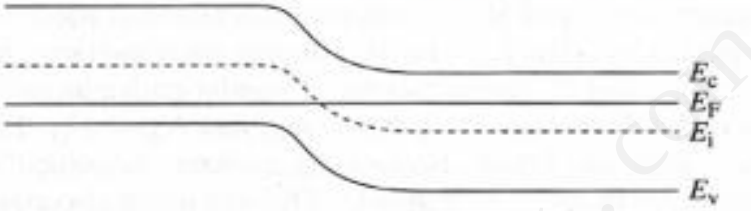




Movement of electrons and holes when forming the junction

Step Junction Solution: What does it mean?

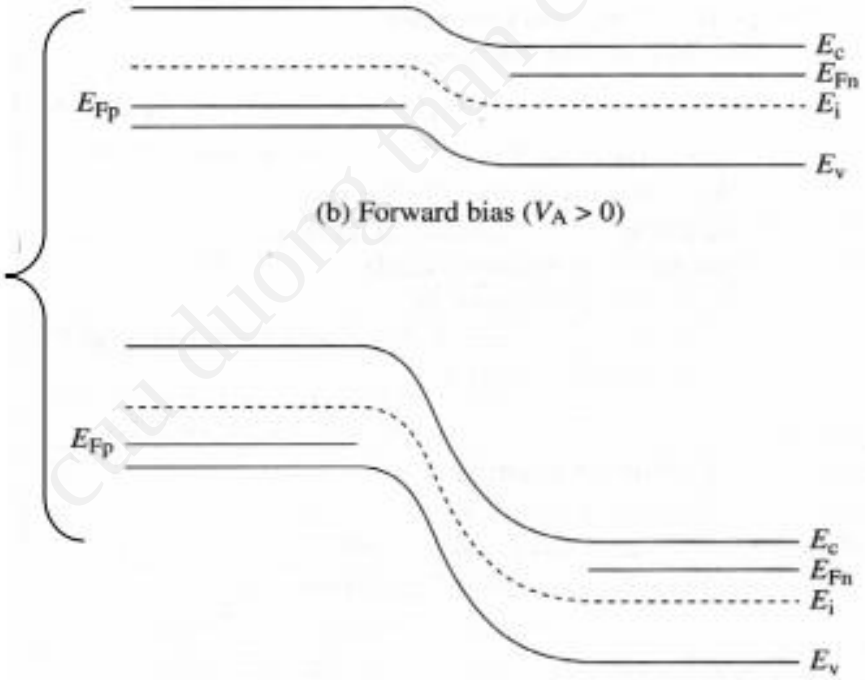
Fermi-level only applies to equilibrium (no current flowing)



(a) Equilibrium ($V_A = 0$)

Majority carrier Quasi-fermi levels

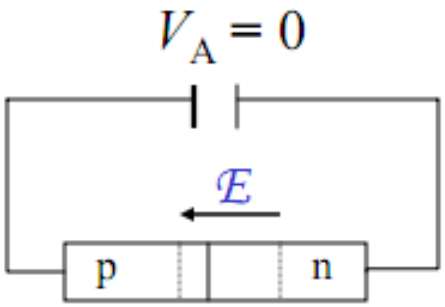
$$E_{fp}-E_{fn}=-qV_A$$

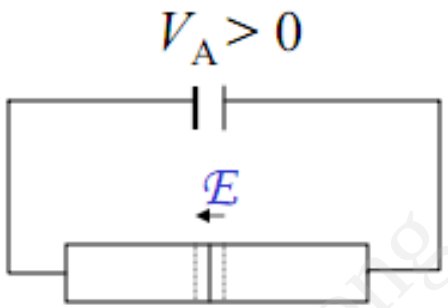


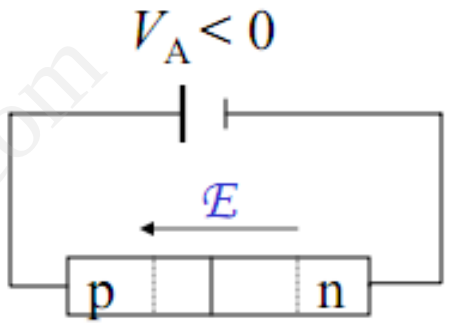
(b) Forward bias ($V_A > 0$)

(c) Reverse bias ($V_A < 0$)

PN junction under various bias conditions







Hole diffusion current

Hole drift current

Electron diffusion current

Electron drift current

Hole diffusion current

Hole drift current

Electron diffusion current

Electron drift current

Hole diffusion current

Hole drift current

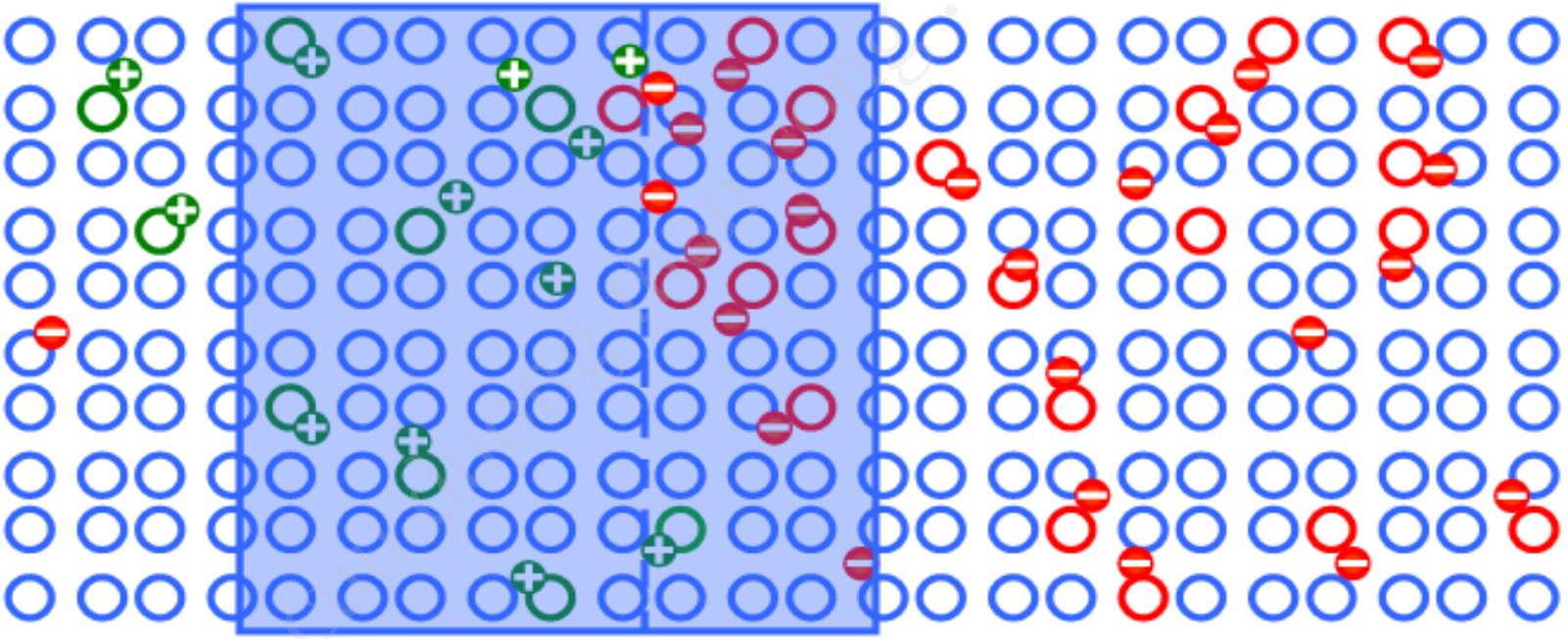
Electron diffusion current

Electron drift current

PN Junction: Asymmetric junction

- Si Atom
- As Atom
- B atom
- Free electron
- Free hole

- An asymmetric junction occurs where the dopant concentration on one side is higher than on the other.
- Here, there are more As atoms than B atoms.
- Electrons & holes must annihilate in pairs

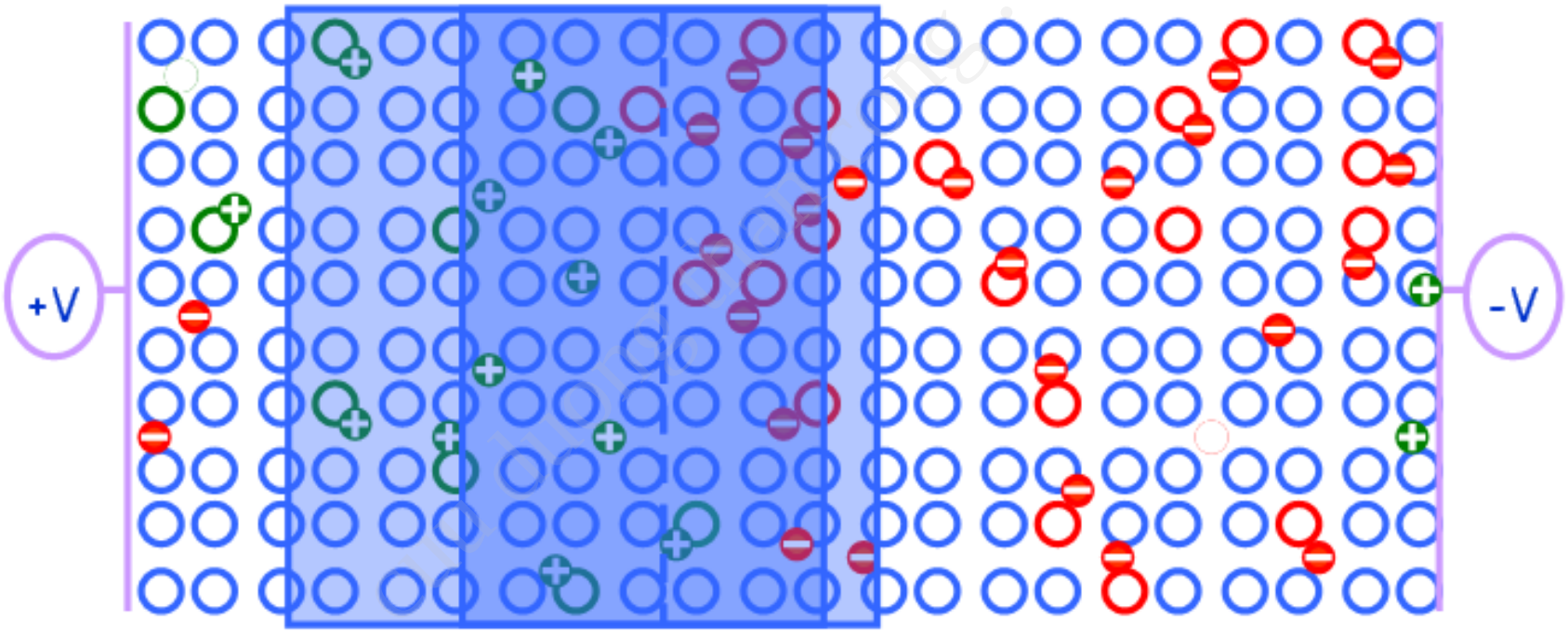


There are more electrons than holes near the junction. This means the electrons will have to travel deeper into the P-type region to create pairs reducing the width of the depletion region on that size

PN Junction: Forward Bias

- Si Atom
- As Atom
- B atom
- Free electron
- Free hole

- Apply an external forward bias
- → Positive voltage to P-type,
- → Negative voltage to N-type



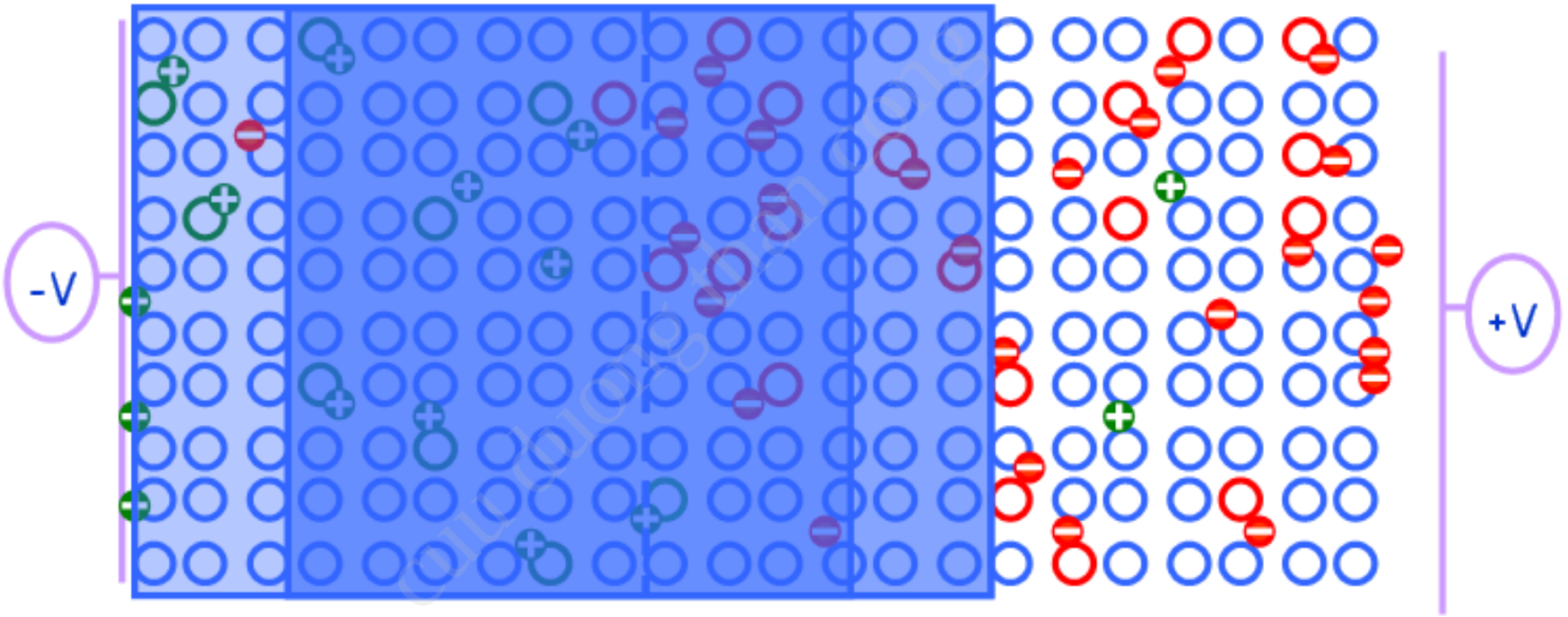
- -ve Voltage feeds in electrons,
- +ve voltage "pulls" electrons across barrier
- Vice Versa for holes

→ Depletion shrinks

PN Junction: Reverse Bias

- Si Atom
- As Atom
- B atom
- Free electron
- Free hole

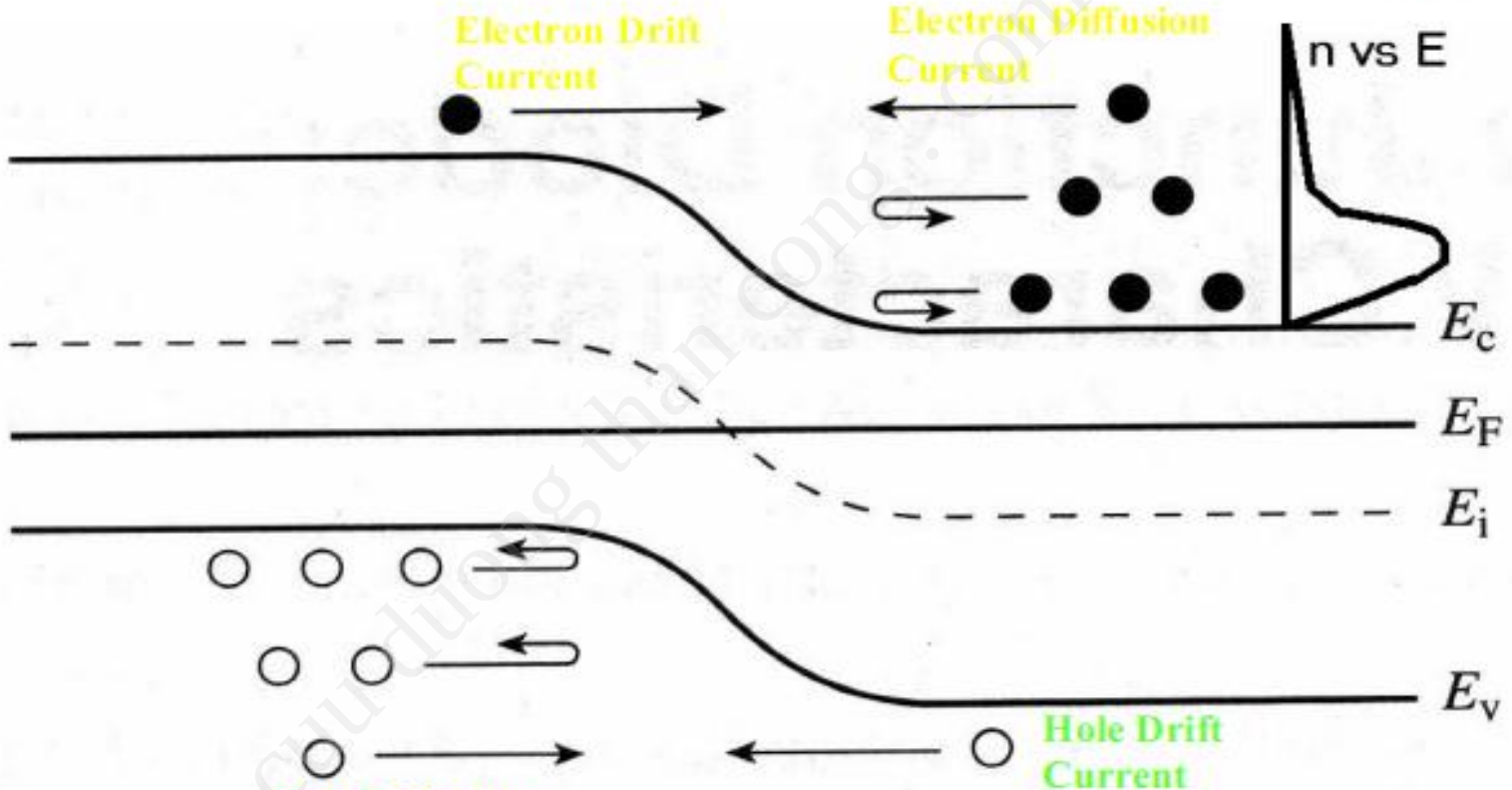
- Apply an external reverse bias
- → Negative voltage to P-type,
- → Positive voltage to N-type



- +ve Voltage pulls electrons from the junction
- This widens the barrier → harder to pull electrons over barrier
- Vice Versa for holes

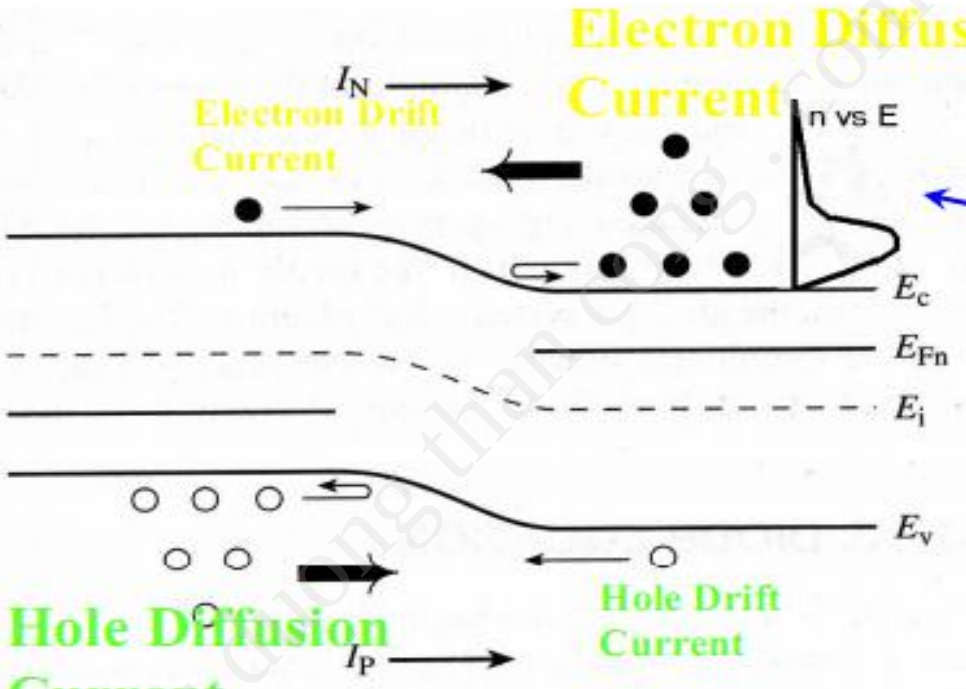
P-n Junction I-V Characteristics

In Equilibrium, the Total current balances due to the sum of the individual components



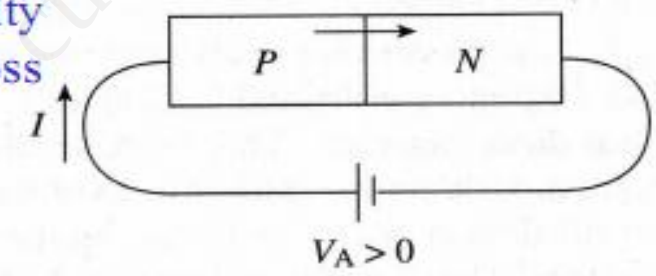
(a) Equilibrium ($V_A = 0$)

P-n Junction I-V Characteristics



Current flow is proportional to $e^{(Va/Vref)}$ due to the exponential decay of carriers into the majority carrier bands

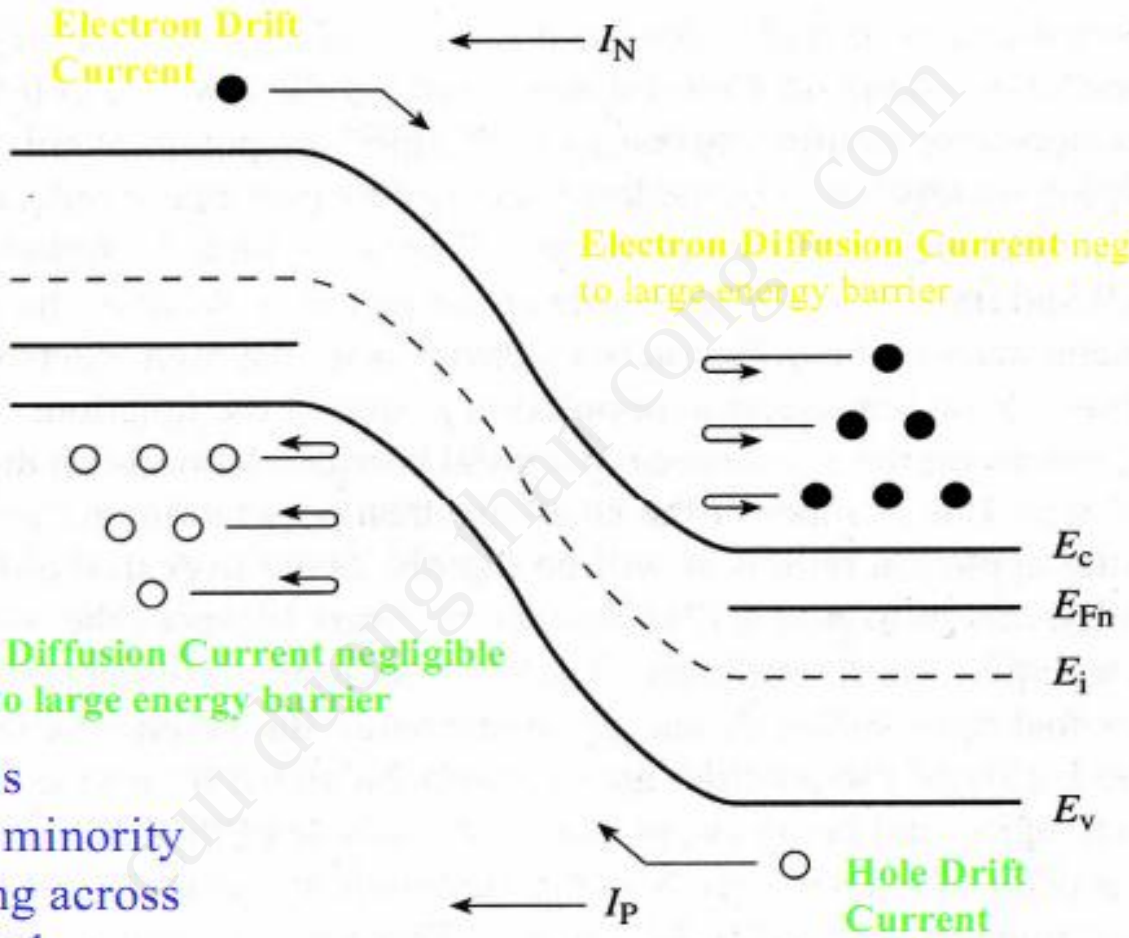
Current flow is dominated by majority carriers flowing across the junction and becoming minority carriers



(b) Forward bias ($V_A > 0$)

P-n Junction I-V Characteristics

Current flow is constant due to thermally generated carriers swept out by E-fields in the depletion region

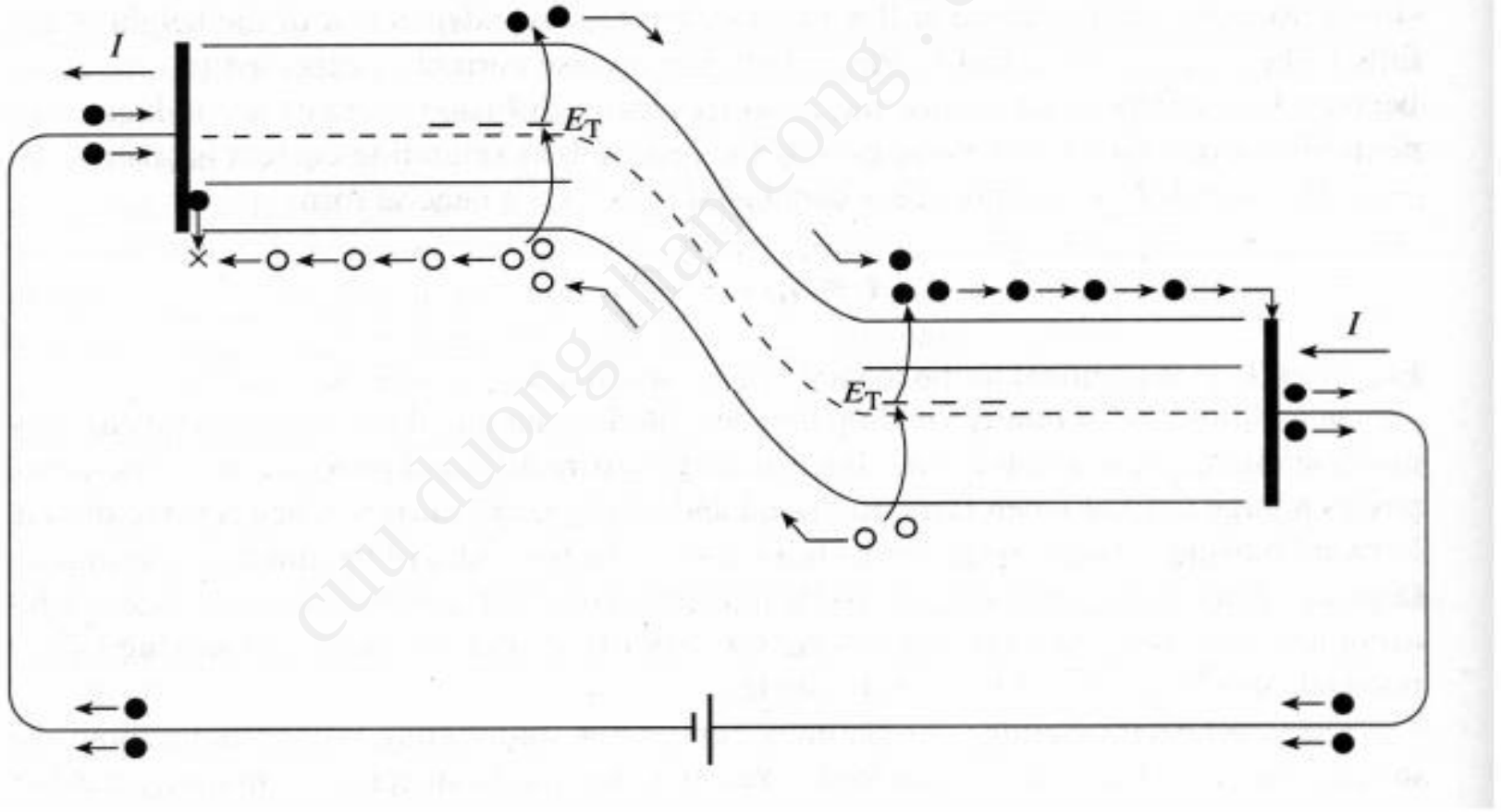


(c) Reverse bias ($V_A < 0$)

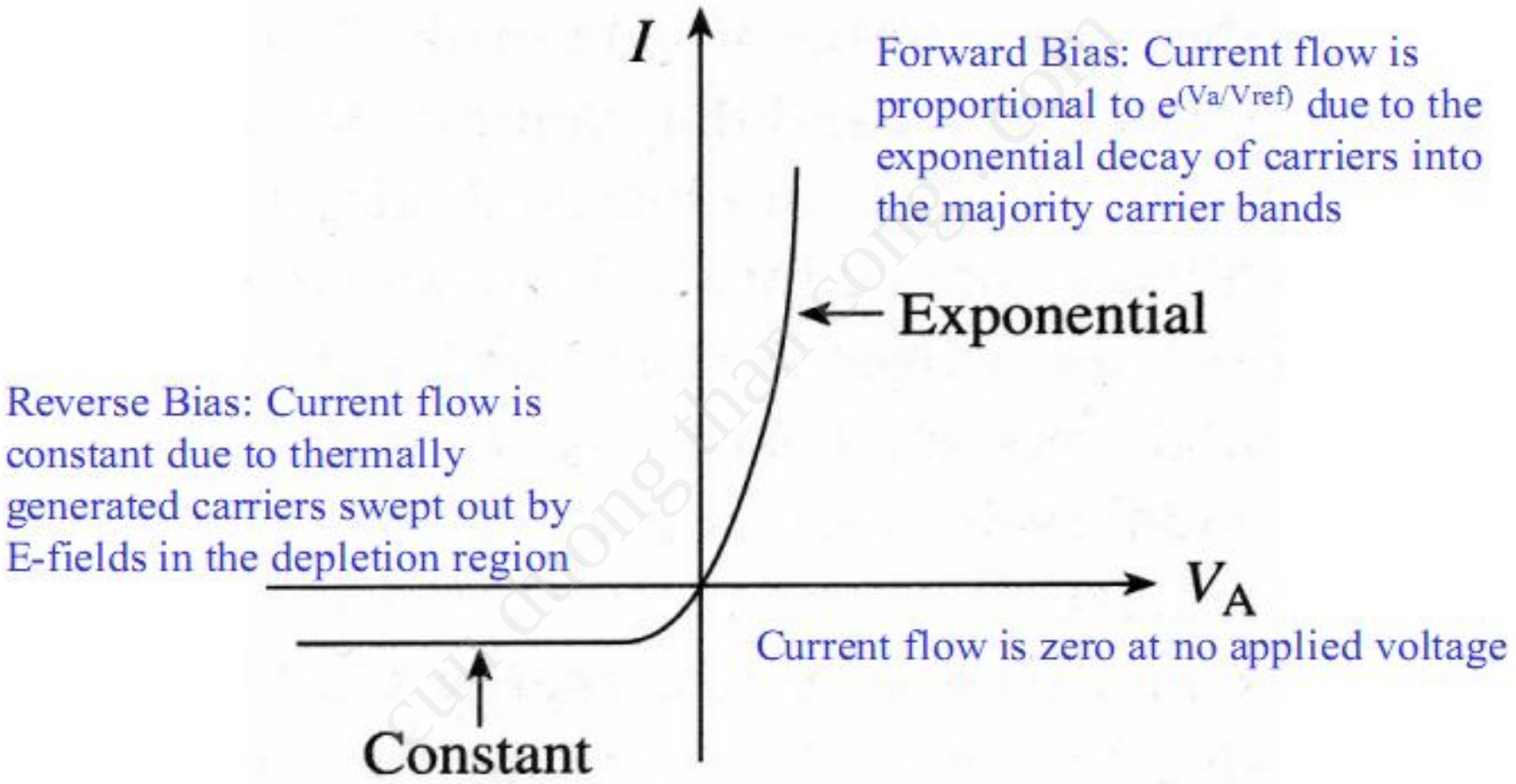
Current flow is dominated by minority carriers flowing across the junction and becoming majority carriers

P-n Junction I-V Characteristics

Where does the reverse bias current come from? Generation near the depletion region edges "replenishes" the current source.



P-n Junction I-V Characteristics
Putting it all together



$$I = I_o(e^{Va/Vref} - 1)$$

PN-junction diodes: Applications

Diode applications:

- Rectifiers
- Switching diodes
- Zener diodes
- Varactor diodes (Varactor = Variable reactance)

Photodiodes

- pn junction photodiodes
- p-i-n and avalanche photodiodes

Solar Cells

Light Emitting Diodes

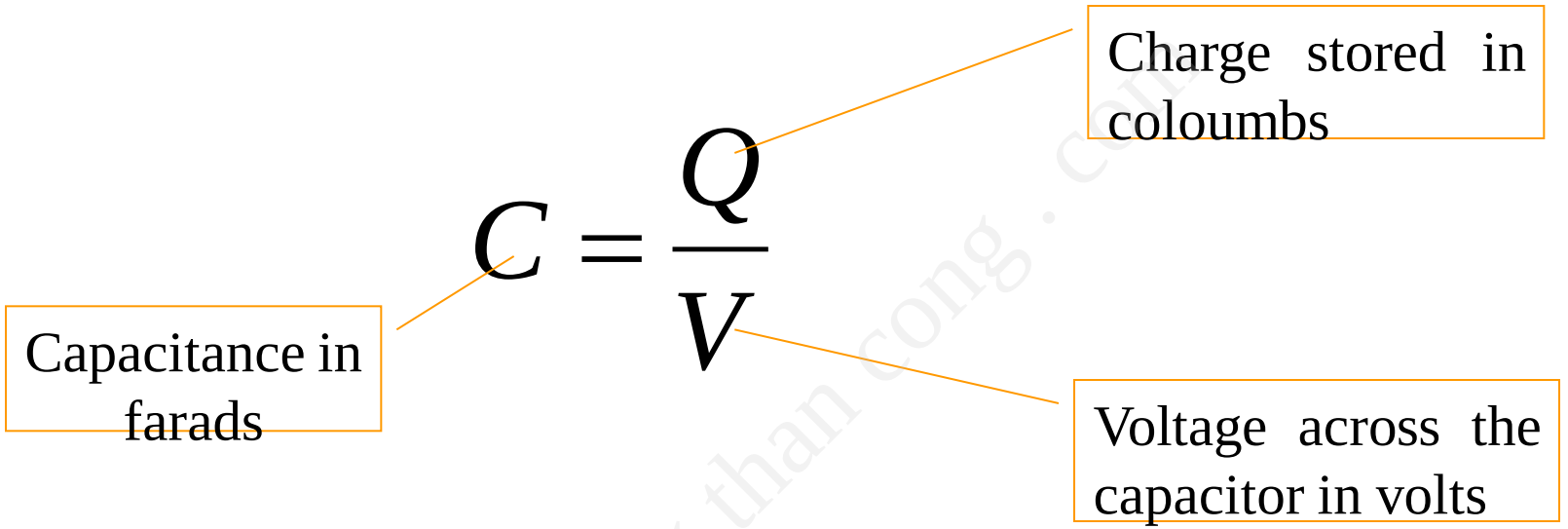
Lasers

Depletion Capacitance

- When a reverse bias is applied to p-n junction diode, the depletion region width, W, increases. This cause an increase in the number of the uncovered space charge in depletion region.

- Whereas when a forward bias is applied depletion region width of the p-n junction diode decreases so the amount of the uncovered space charge decreases as well.

- So the p-n junction diode behaves as a device in which the amount of charge in depletion region depends on the voltage across the device. So it looks like a capacitor with a capacitance.



- Capacitance of a diode varies with **W** (Depletion Region width)
- **W** (DR width varies width applied voltage **V**)

Capacitance per unit area of a diode ;

$$C_{DEP} = \frac{\varepsilon_{Si}}{W} \frac{F}{cm^2}$$

For one-sided abrupt junction; e.g. $N_A >> N_D \Rightarrow x_n >> \quad W \cong x_n$

$$N_A x_p = N_D x_n$$

$$C_{DEP} = \frac{\varepsilon_{Si}}{W} \cong \frac{\varepsilon_{Si}}{x_n} \Rightarrow x_n = \sqrt{\frac{2\varepsilon_{Si} V_{bi}}{qN_D}} \quad \text{for} \quad N_A >> N_D$$

The application of reverse bias ;

$$C_{DEP} = \frac{\varepsilon_{Si}}{\sqrt{\frac{2\varepsilon_{Si}(V_{bi} + V_R)}{qN_D}}} = \sqrt{\frac{q\varepsilon_{Si} N_D}{2(V_{bi} + V_R)}}$$

If one makes C - V measurements and draw $1/C^2$ against the voltage V_R; *obtain* built-in voltage and doping density of low-doped side of the diode from the intercept and slope.

$$\frac{1}{C^2} = \frac{2(V_{bi} + V_R)}{q\varepsilon_{Si} N_D}$$

$$slope = \frac{2}{q\varepsilon_{Si} N_D}$$

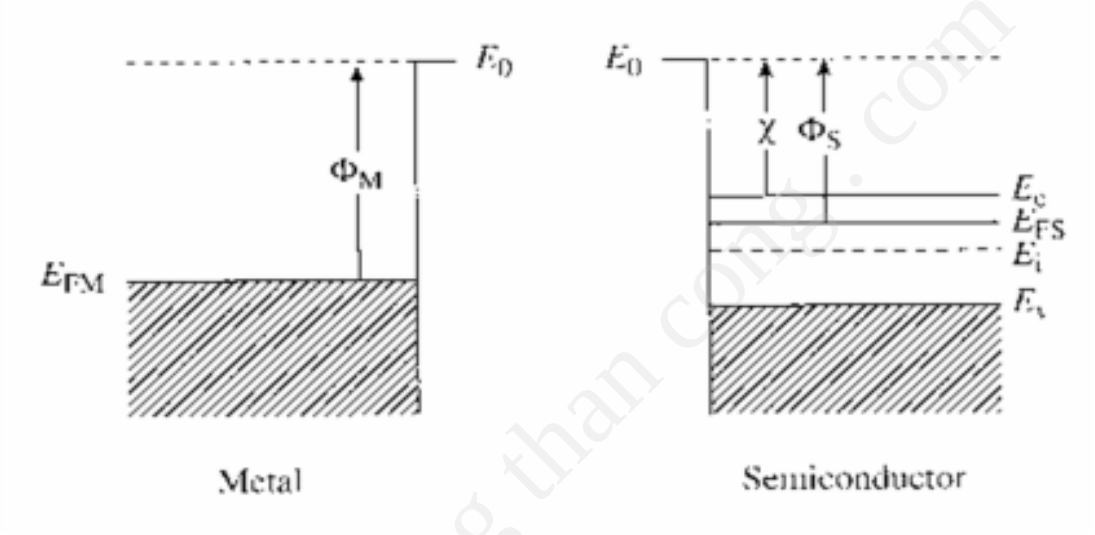
$$V_{bi} = \frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right)$$

- Why a Schottky?
  - Minority Carrier Charge storage in p-n junctions tends to limit the switching times of p-n junction diodes
    - Turn off times limited by minority carrier lifetimes
  - Schottky Diodes have little (no) minority carrier stored charge and thus have application in fast switching applications (motors, etc…)
  - Cheaper
- Disadvantages:
  - Generally higher leakage currents
  - Generally lower breakdown voltages

Metal-semiconductor (MS) junctions

- P-N junctions formed depletion regions by bringing together two materials with dissimilar fermi energies, allowing charge transfer and subsequent alignment of the energy bands.
- Several other combinations of such materials can also form "useful junctions".
  - Schottky Diodes (metal-semiconductor junction)
  - Ohmic contacts (metal-semiconductor junction)
  - Thermocouples (metal-metal junction)

Ideal Metal-Semiconductor Contacts



Assumptions - Ideal MS contacts

- Metal (M) and Semiconductor (S) are in intimate contact, on atomic scale
- No oxides or charges at the interface (very bad assumption in some cases – some interfaces are dominated by interfacial oxides or interface charge).
- No intermixing at the interface (in some cases, it is impossible to put a metal on a semiconductor without some exchange of atoms – intermixing-occurring)
- These assumptions require ultra-clean interfaces otherwise non-ideal behavior results (fermi-level pinning of III-V compounds is common for example)

Definitions

- **Vacuum level, E_0** - corresponds to energy of free electrons in vacuum.
- The difference between vacuum level and Fermi-level is called workfunction, Φ of materials.
  - **Workfunction, Φ_M** is an invariant property of a given metal. It is the minimum energy required to remove electrons from the metal. (Lowest value is 1.95eV for Cs, 3.66eV for Mg, 5.15eV for Ni, and highest value is 5.7eV for Pt, etc.). Electron density varies with crystallographic orientation so the work function varies with orientation as well.
- However, since the electron concentration depends on doping in a semiconductor, the semiconductor **workfunction, Φ_s,** depends on the doping.

$$\Phi_s = \chi + (E_C - E_F)_{FB}$$

where $\chi = (E_0 - E_C)|_{SURFACE}$ is a a fundamental property of the semiconductor. (Example: $\chi = 4.0$ eV, 4.03 eV and 4.07 eV for Ge, Si and GaAs respectively)

Energy band diagrams for ideal MS contacts

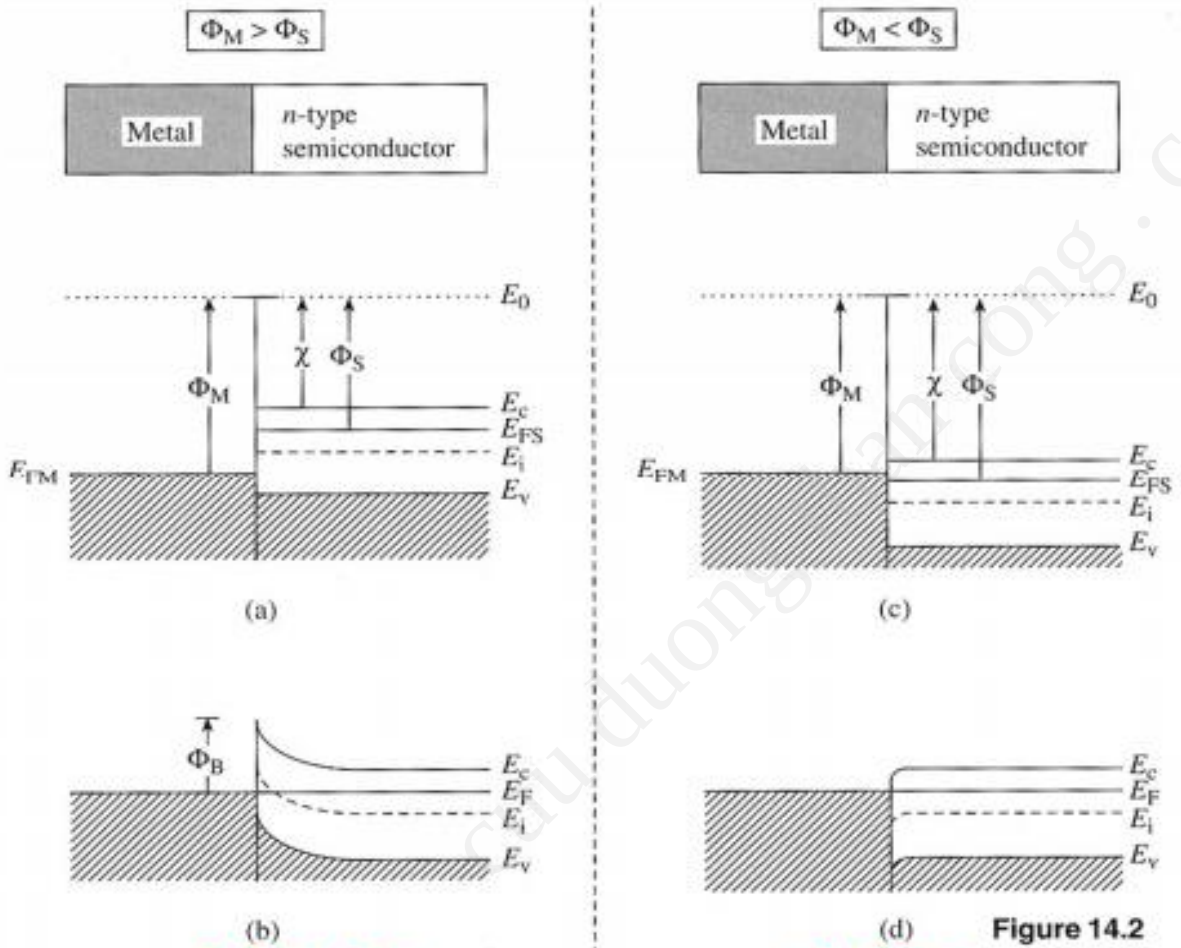


Figure 14.2

An instant after contact formation

Under equilibrium conditions

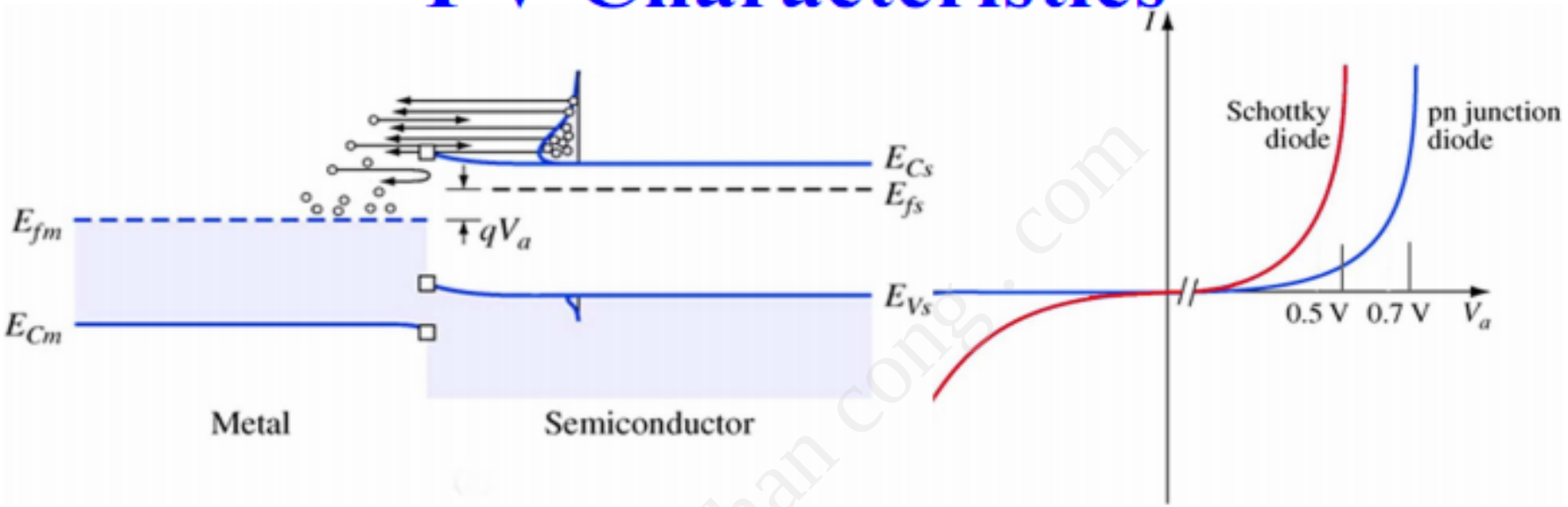
Schottky
$\Phi_M > \Phi_S$

Ohmic
$\Phi_M < \Phi_S$

MS (n-type) contact with $\Phi_M > \Phi_S$

- Soon after the contact formation, electrons will begin to flow from the semiconductor to the metal.
- The removal of electrons from the n-type material leaves behind uncompensated N_d^+ donors, creating a surface depletion layer, and hence a built-in electric field (similar to p^+-n junction).
- Under equilibrium, the Fermi-level will be constant and no energy transfer (current) flows
- A barrier Φ_B forms blocking electron flow from M to S.
- Based on the Electron Affinity Model (EAM), the simplest of models used to describe MS junctions, $\Phi_B = \Phi_M - \chi$ ... ideal MS (n-type) contact. Φ_B is called the "barrier height".
- Electrons in a semiconductor will encounter an energy barrier equal to $\Phi_M - \Phi_S$ while flowing from S to M.

I-V Characteristics

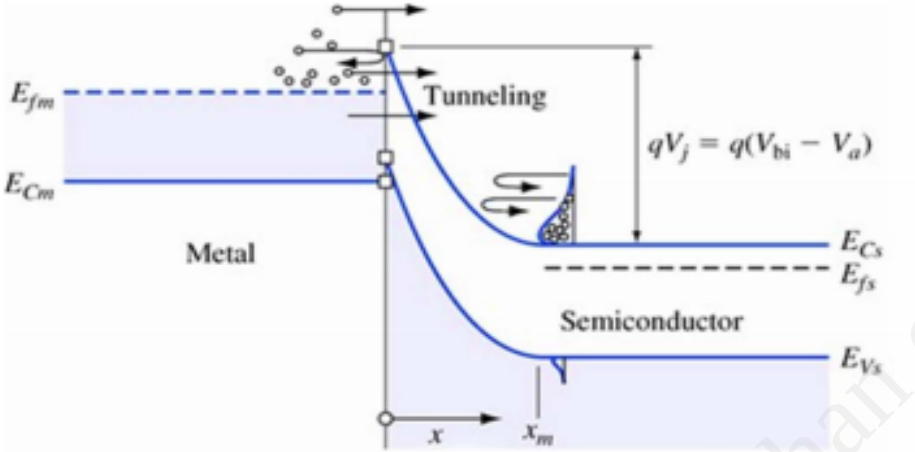


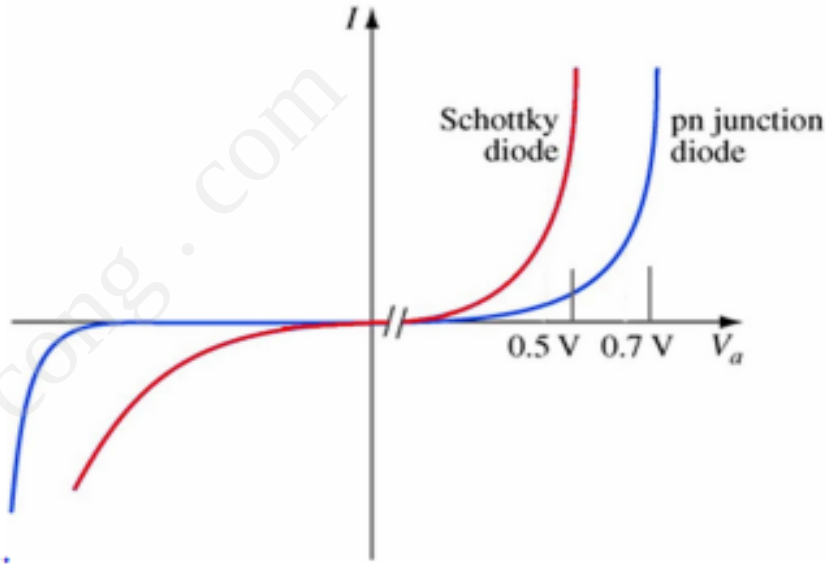
Since MS Schottky diode is a majority carrier device (i.e only majority carriers are injected from semiconductor to the metal) and thus has no minority carrier storage, the frequency response of the device is much higher than that of equivalent p^+ n diode.

The "turn on voltage" of a Schottky diode is typically smaller than a comparable p-n junction since the barrier to forward current flow (Φ_m-Φ_s) is typically small. This "turn on" voltage can be as small as 0.3 Volts in some Si Schottky diodes.

This makes a Schottky diode the best choice for power switch protection in inductive load applications (motors, solenoids, coils, etc…) and in high frequency rectification but not a good choice when low leakage or high breakdown voltage is required.

I-V Characteristics





Leakage in a Schottky diode is dominated by:

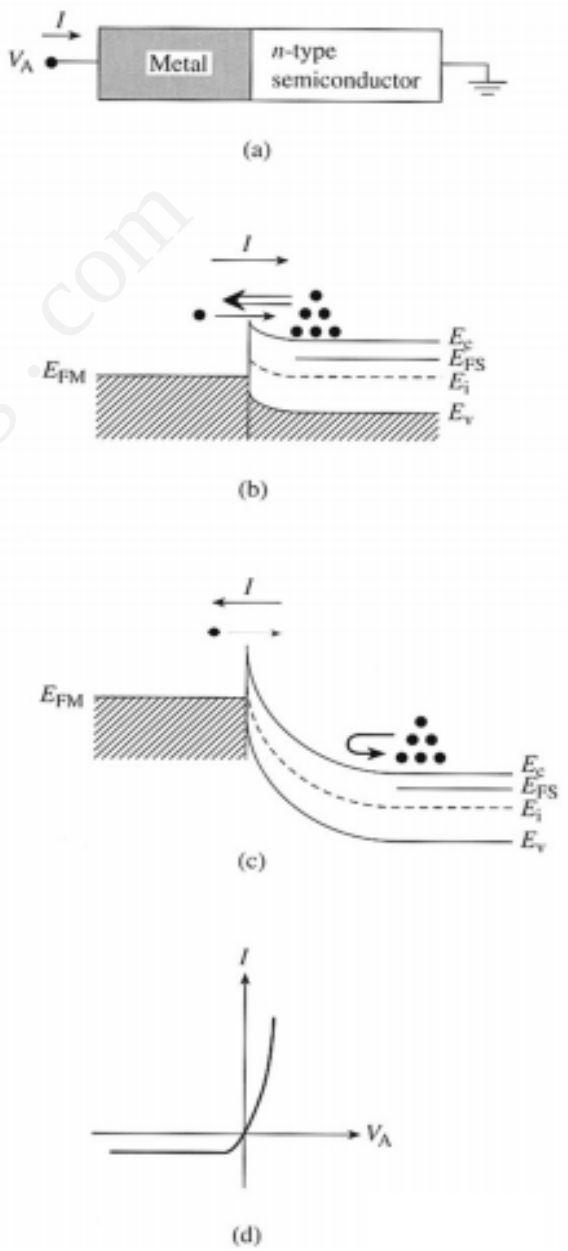
1) "Thermionic Emission" (metal electrons emitted over the barrier – not likely)
2) "Thermionic Field Emission" (metal electrons of higher energy tunneling through the barrier – more likely)
3) "Direct tunneling" (metal electrons tunneling through the barrier – most likely in higher doped semiconductors or very high electric fields).

Since generation does not require the entire bandgap energy to be surmounted, the reverse leakage current for a Schottky diode is generally much larger than that for a p^+n diode. Likewise, breakdown (for the same reason) is generally at smaller voltages.

MS (n-type) contact with $\Phi_M > \Phi_S$

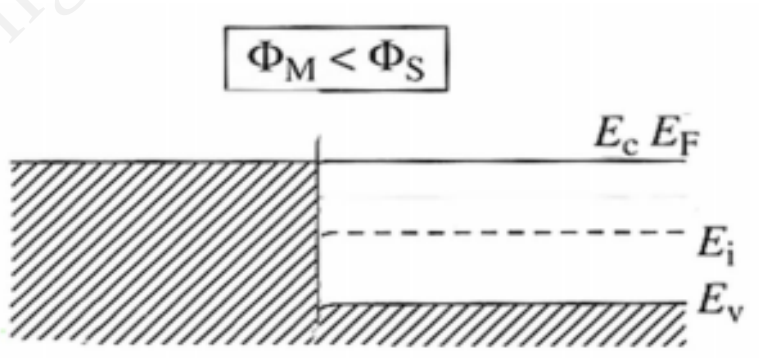
A forward bias will reduce the barrier height unbalancing the electron current flow, resulting in a huge forward current that increases exponentially with applied voltage

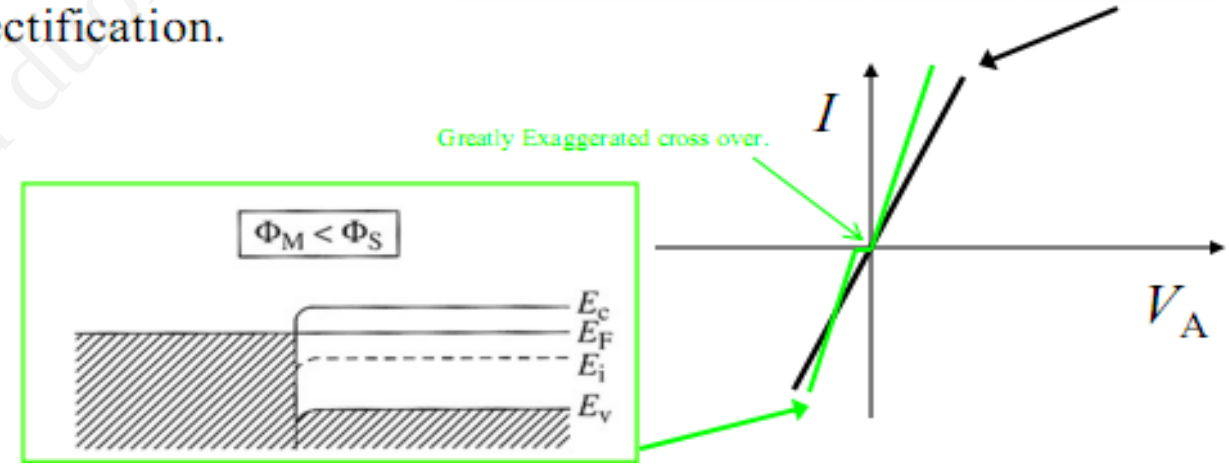
A reverse bias will increase the barrier height resulting in a small "reverse current" flow that will be dominated by tunneling currents for high doped semiconductors and/or thermally assisted field emission for moderate/low doped semiconductors.



Ohmic Contacts: MS (n-type) contact with $\Phi_M < \Phi_S$

- There is no barrier for electron flow from the semiconductor to the metal. So, even for a small $V_A > 0$ results in large current.
- The small barrier that exists for electron flow from metal to the semiconductor, but vanishes when $V_A < 0$ is applied to the metal. Large current flows when $V_A < 0$.
- The MS (n-type) contact when $\Phi_M < \Phi_S$ behaves like an ohmic contact.
- Lack of depletion (accumulation occurs) means (essentially) no rectification.





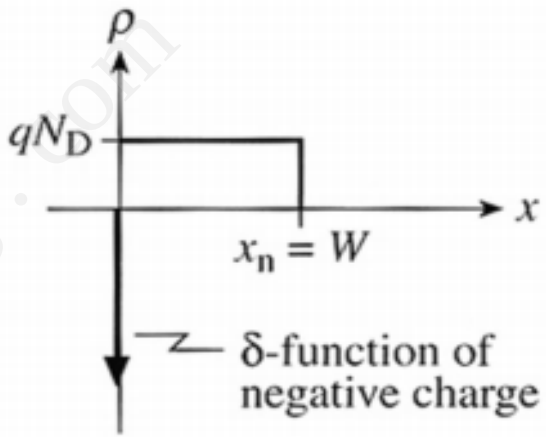
Generalization of Metal Semiconductor Contact Energy Relationships

	n-type	p-type
$\Phi_M > \Phi_S$	rectifying	ohmic
$\Phi_M < \Phi_S$	ohmic	rectifying

Schottky Diode Electrostatics

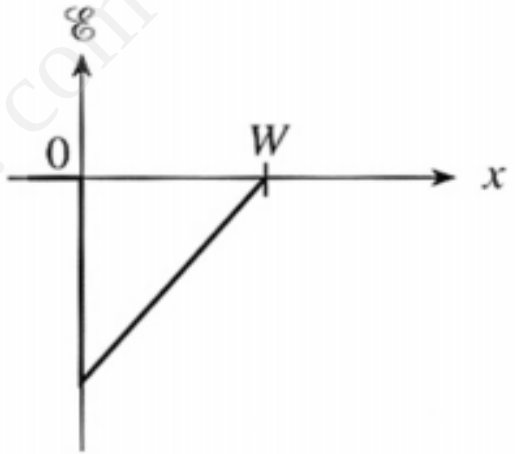
$$V_{bi} = \frac{1}{q}\left[\Phi_B - (E_C - E_F)_{FB}\right]$$

$$\rho \approx qN_D \quad \text{for} \quad 0 \le x \le W$$
$$\approx 0 \quad \text{for} \quad x > W$$



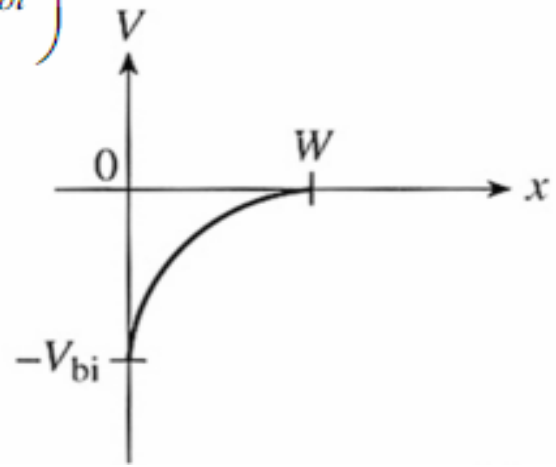
Schottky Diode Electrostatics

$$\frac{d\mathcal{E}}{dx} = \frac{\rho}{\varepsilon_{Si}} = \frac{qN_D}{\varepsilon_{Si}} \quad \text{for} \quad 0 \le x \le W$$



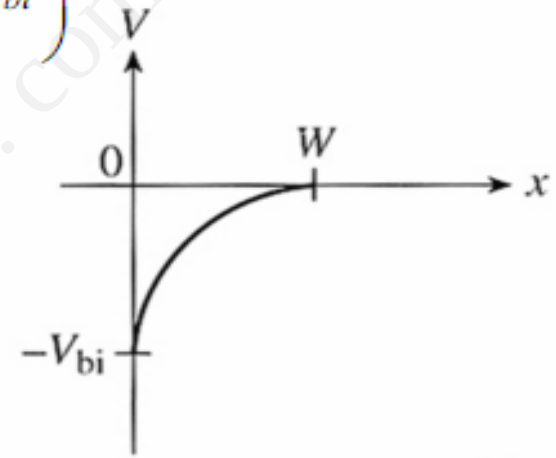
$$\mathbf{E} = \frac{qN_D}{\varepsilon_{Si}}(x - W) \quad \text{for} \quad 0 \le x \le W$$

$$V(x) = \frac{qN_D}{\varepsilon_{Si}}\left(Wx - \frac{1}{2}x^2\right) - \left(\left(\frac{(E_c - E_f)_{FB}}{q}\right) + V_{bi}\right) \quad \text{for} \quad 0 \le x \le W$$

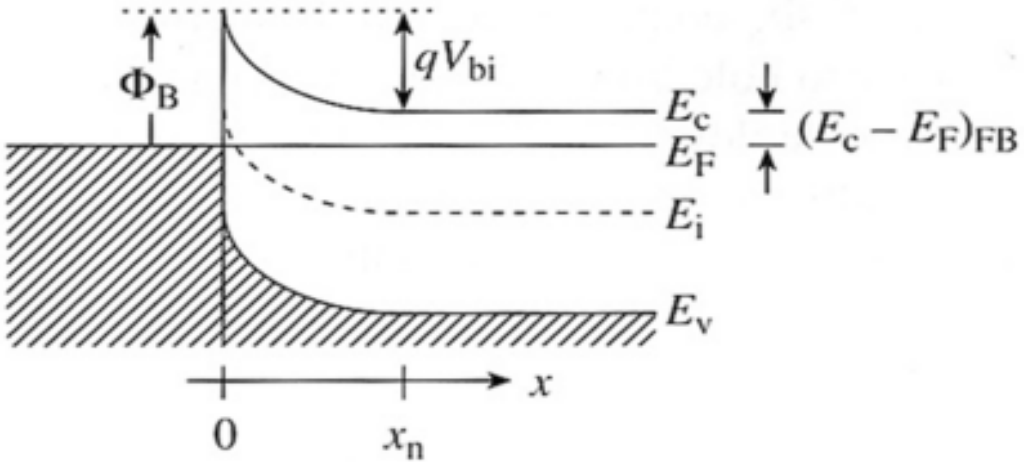


Schottky Diode Electrostatics

$$V(x) = \frac{qN_D}{\varepsilon_{Si}}\left(Wx - \frac{1}{2}x^2\right) - \left(\left(\frac{(E_c - E_f)_{FB}}{q}\right) + V_{bi}\right) \quad \text{for} \quad 0 \le x \le W$$



$$W = \sqrt{\frac{2\,\varepsilon_{Si}}{q\,N_D}\left(V_{bi} - V_A - \frac{kT}{q}\right)}$$



Example

Find barrier height, built-in voltage, maximum E-field, and the depletion layer width at equilibrium for W-Si (n-type) contact.
Given: $\Phi_M = 4.55\text{eV}$ for W; $\chi(\text{Si}) = 4.01\text{eV}$; Si doping = $10^{16}\,\text{cm}^{-3}$
Draw the band diagram at equilibrium.

Solution:

Find $E_F - E_i$ $\quad E_F - E_i = 0.357\text{eV}$

Find $E_C - E_F$ $\quad E_C - E_F = 0.193\text{eV}$

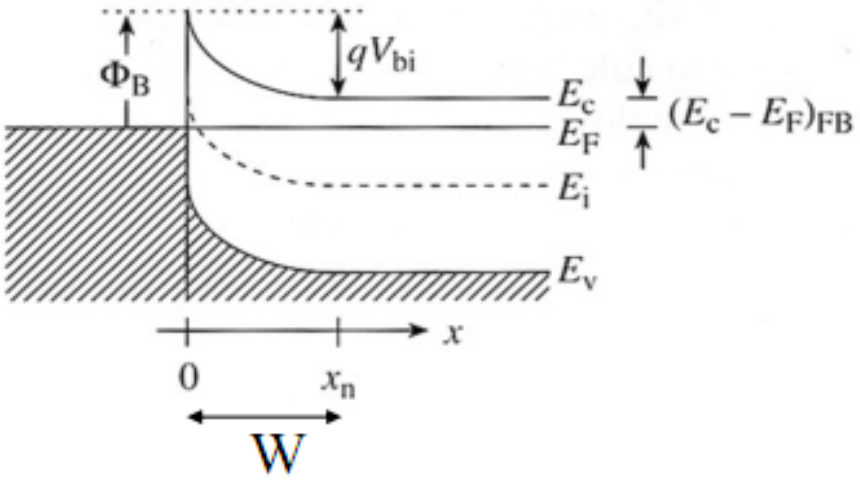
$$\Phi_B = \Phi_M - \chi = 0.54\text{eV}$$

$$\Phi_S = \chi + (E_C - E_F)_{FB} = 4.203\ \text{eV}$$

$$V_{bi} = 0.347\ \text{V}$$

$$W = 0.21\ \mu\text{m}$$

$$\mathcal{E}(x=0) = \mathcal{E}_{max} = 3.4 \times 10^4\ \text{V/cm}$$



I-V Characteristics

$$I = I_S\left(e^{\frac{qV_A}{kT}} - 1\right) \quad \text{where} \quad I_S = A\mathcal{A}^* T^2 e^{-\frac{\Phi_B}{kT}}$$

where Φ_B is Schottky barrier height, V_A is applied voltage, A is area, and $\mathcal{A}^*$ is Richardson's constant.

$$\text{where} \quad A^* = \frac{4\pi q m^* k^2}{h^3} = 120\left[\frac{A}{cm^2 K}\right]$$

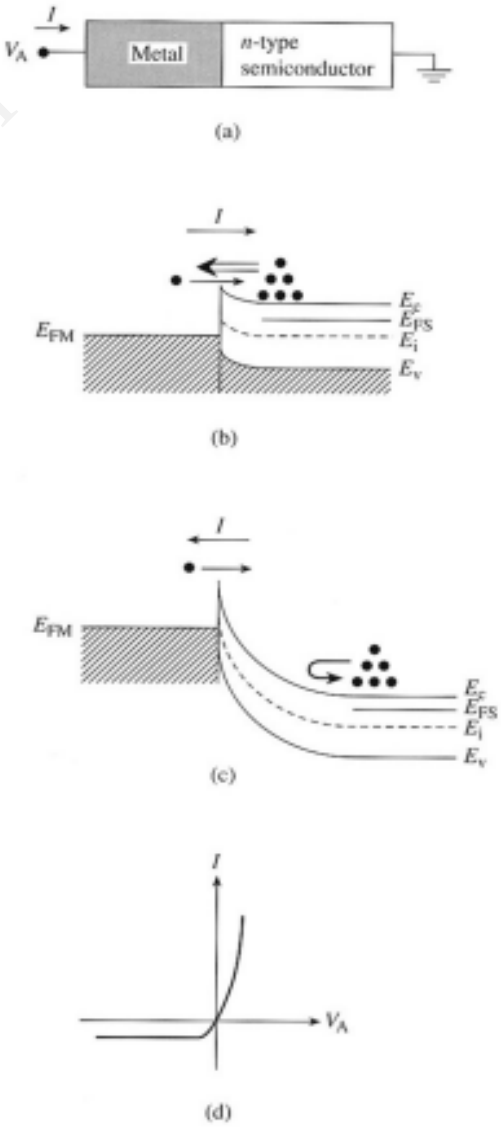
Note: $\mathcal{A}^*$ is Richardson's constant but often times, there is an extra prefactor λ which accounts for quantum mechanical reflection of the electrons approaching the potential barrier and to a lesser degree varies with the band structure of the emitting material.

I-V Characteristics

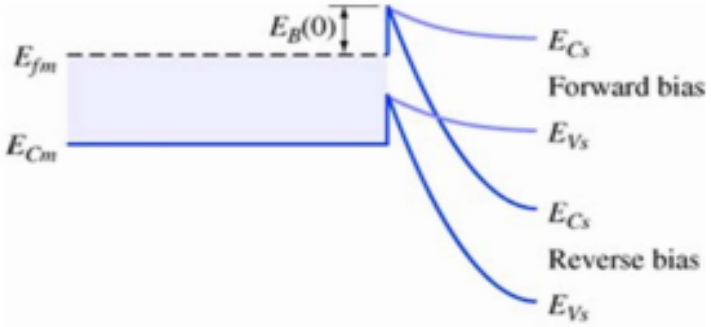
$$I = \left(A\mathbf{A}^{*}T^{2}\, e^{-\frac{\Phi_B}{kT}} \right) e^{\frac{qV_A}{kT}} - \left(A\mathbf{A}^{*}T^{2}\, e^{-\frac{\Phi_B}{kT}} \right)$$

Results from the emission of electrons from the semiconductor to the metal over the barrier, qV_{BI}. This must balance the leakage current at zero bias and increases exponentially due to diffusion current as the barrier is lowered in forward bias.

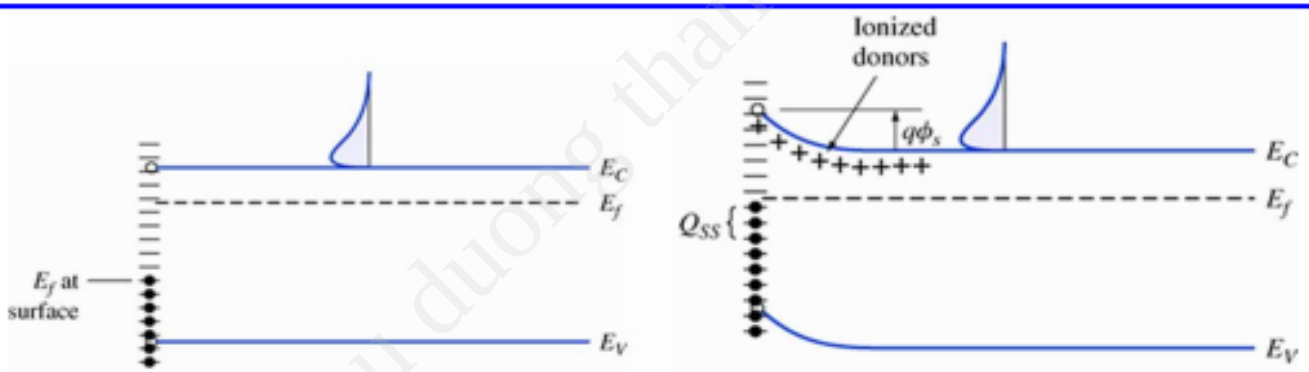
Results from the emission of electrons from the metal to the semiconductor over the barrier, Φ_B



Details of Schottky Behavior



Note that the barrier height is (mostly) independent of bias.

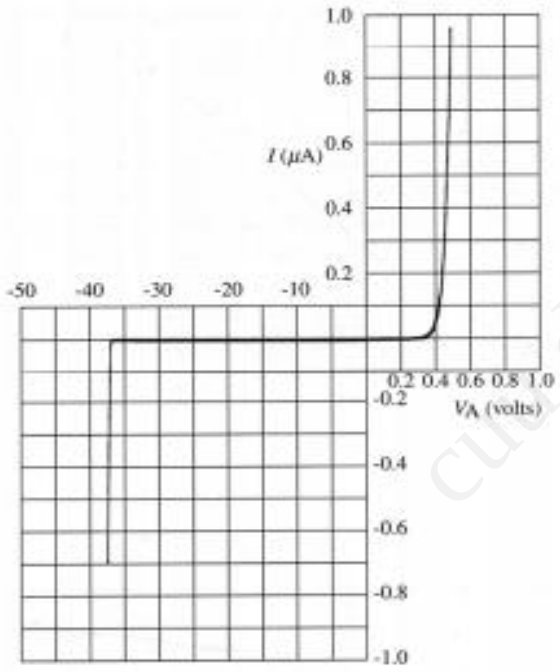


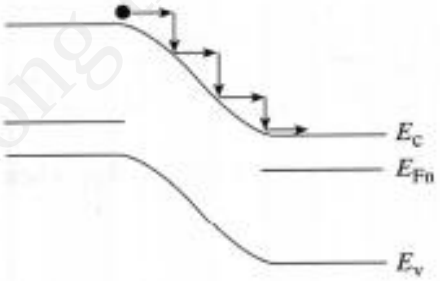
Surface charges resulting from broken bonds, surface contamination or even surface oxidation can dominate band alignment making the EAM invalid. One common case is the "fermi-level pinning" often found in III-As and III-P materials. In these materials, the large numbers of surface states force the surface fermi level to become fixed (pinned) at one energy position regardless of the metal used to contact the surface.

Breakdown Mechanisms

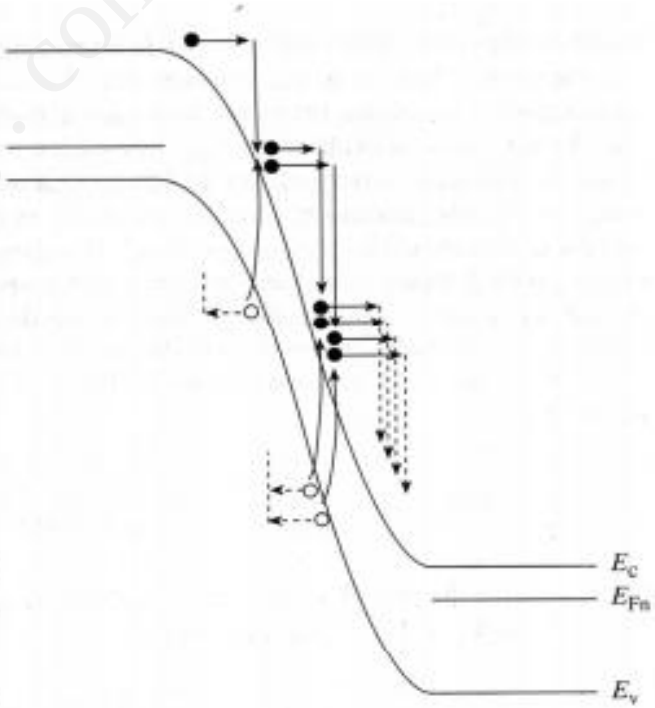
Avalanche Breakdown:

Excess current flows due to electron-hole pair multiplication due to impact ionization. This current rapidly increases with increasing reverse bias.





(a) Small reverse bias



(b) $|V_A| \rightarrow V_{BR}$

Breakdown Mechanisms

Zener Breakdown:

Excess current flows due to bonding electrons "tunneling" into empty conduction band states. The "tunneling barrier" must be sufficiently thin. This current rapidly increases with increasing reverse bias.

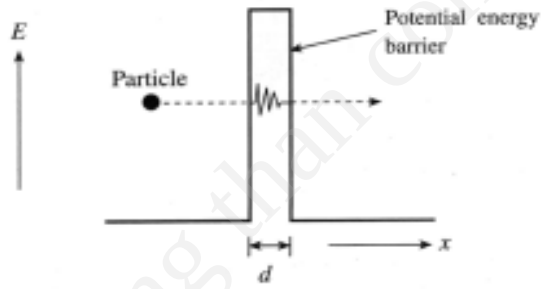
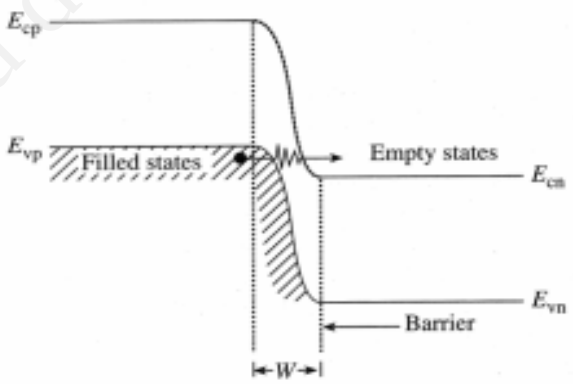


Figure 6.13 General visualization of tunneling.



Device

- Two main categories of transistors:
  - bipolar junction transistors (BJTs)
  - field effect transistors (FETs).

- Transistors have 3 terminals where the application of current (BJT) or voltage (FET) to the input terminal increases the amount of charge in the active region.

- The physics of "transistor action" is quite different for the BJT and FET.

- In analog circuits, transistors are used in amplifiers and linear regulated power supplies.

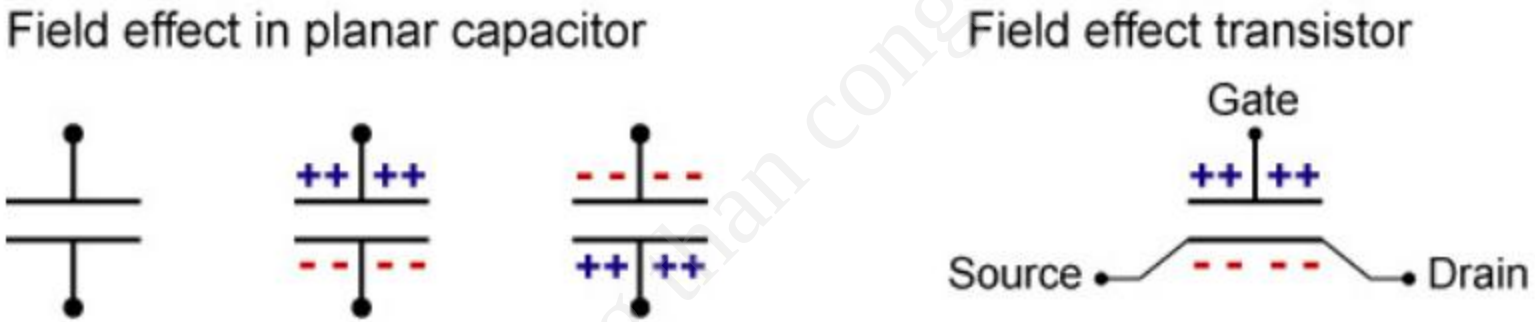
- Transistors may be grouped into two major divisions: *bipolar* and *field-effect*.

- Bipolar transistors: utilize a small current to control a large current.

- Field-effect transistor: utilize a small *voltage* to control current → and then focus on one particular type: the *junction* field-effect transistor.

- All field-effect transistors are *unipolar* rather than *bipolar* devices. That is, the main current through them is comprised either of electrons through an N-type semiconductor or holes through a P-type semiconductor.

Field effect: Induction of an electronic charge due to an electric field.



JFET

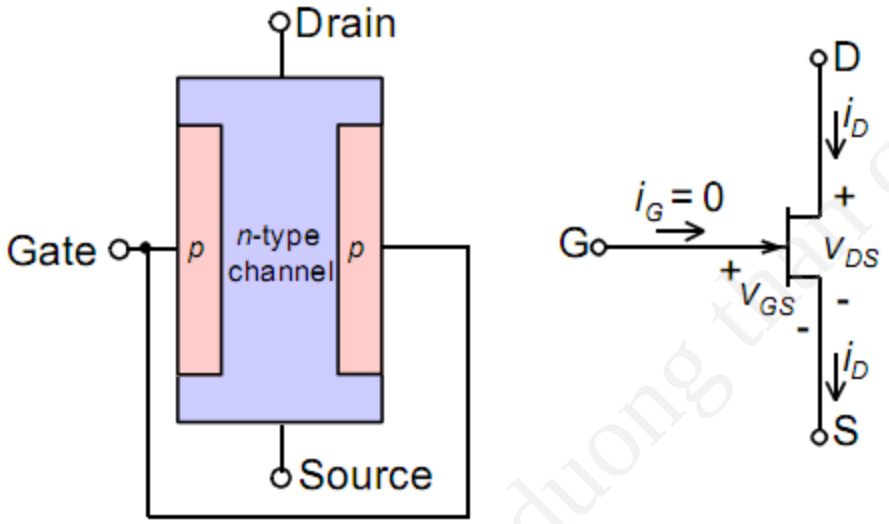


Fig. 127. The *n*-channel JFET representative physical structure (left) and schematic symbol (right).

The *p*-*n* junction is a typical diode . . .

Holes move from *p*-type into *n*-type . . .

Electrons move from *n*-type into *p*-type . . .

Region near the *p*-*n* junction is left without any available carriers - *depletion region*

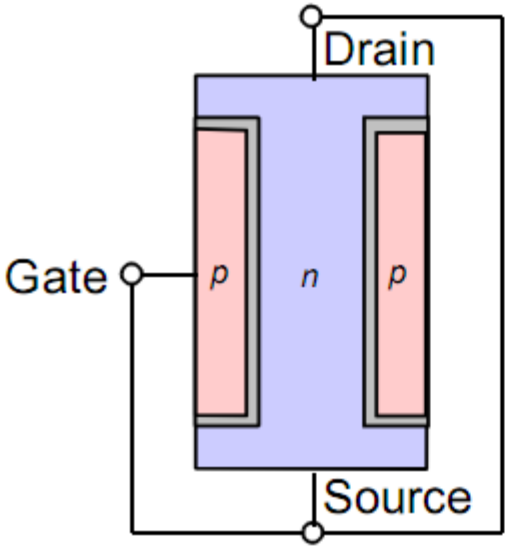


Fig. 128. Depletion region depicted for $v_{GS} = 0$, $v_{DS} = 0$.

The depletion region is shown at left for zero applied voltage (called *zero bias*). . .

Carriers are still present in the *n*-type channel . . .

Current could flow between drain and source (if $v_{DS} \neq 0$) . . .

Channel has relatively low resistance.

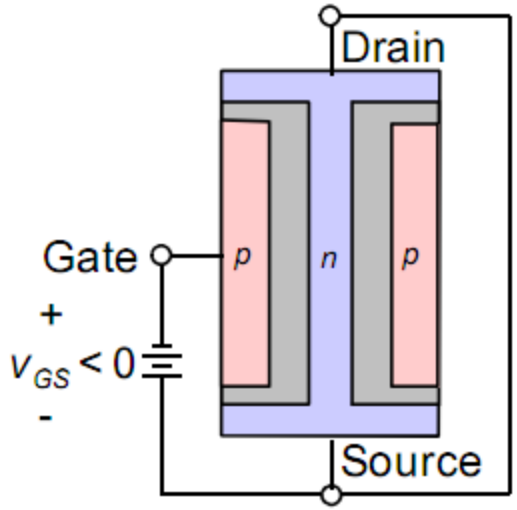


Fig. 129.
Depletion region for negative V_{GS} (reverse bias).

As the reverse bias increases across the *p-n* junction, the depletion region width increases,

Because negative voltage at the Gate pulls holes away from junction,

And positive voltage at the Source pulls electrons away from junction.

Thus, the channel becomes narrower, and the channel resistance increases.

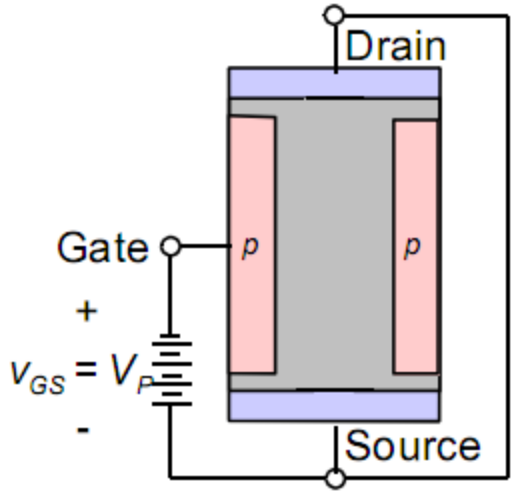


Fig. 130. Depletion region at pinch-off ($v_{GS} = V_P$).

With sufficient reverse bias the depletion region *pinches-off* the entire channel:

$$v_{GS} = V_P \text{, } \underline{\textit{pinch-off voltage}}$$

The channel resistance becomes infinite; current flow impossible for any v_{DS} (less than breakdown).

Typical values: $-5 < V_P < -2$

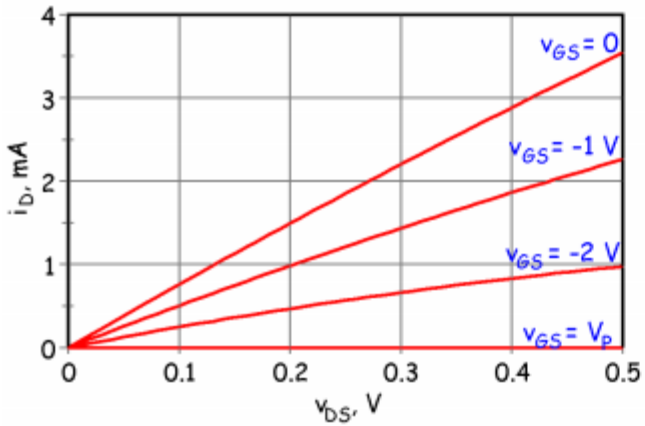


Fig. 131. FET *i-v* curves for small V_{DS}.

Thus, the FET looks like a *voltage-controlled resistance* at small values of v_{DS}.

This region of FET operation is called the *voltage-controlled resistance*, or *triode*, region.

Now, as v_{DS} increases, the depletion region becomes asymmetrical:

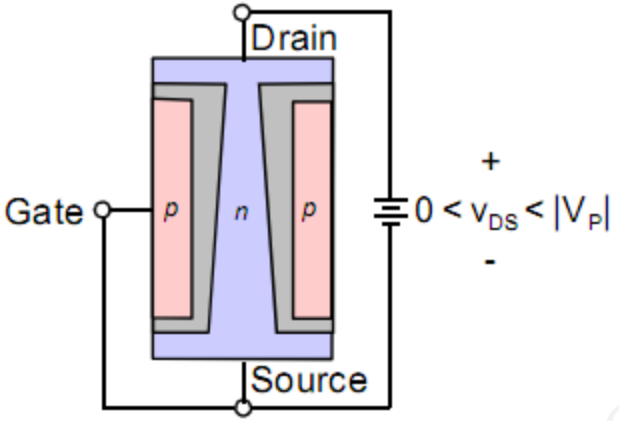


Fig. 132. Asymmetrical depletion region as v_{DS} increases.

Reverse bias is greater at the drain end, so the depletion region is greater at the drain end.

Thus the channel becomes more restricted and, for fixed v_{GS}, *i-v* curves become flatter (i.e., more horizontal).

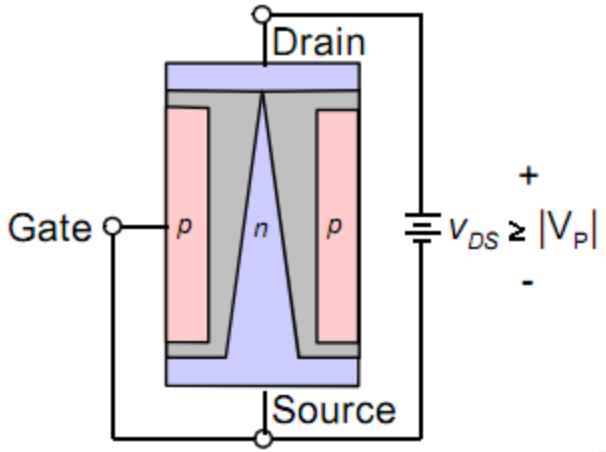


Fig. 133.Pinch-off at drain end for $v_{DS} = V_P$.

For $v_{DS} = |V_P|$ channel becomes pinched-off only at drain end.

Carriers drift across pinched-off region under influence of the *E* field.

The *rate* of drift, and therefore the drain current flow, is dependent on width of entire channel (i.e., on v_{GS}), but *independent of* v_{DS}*!!!*

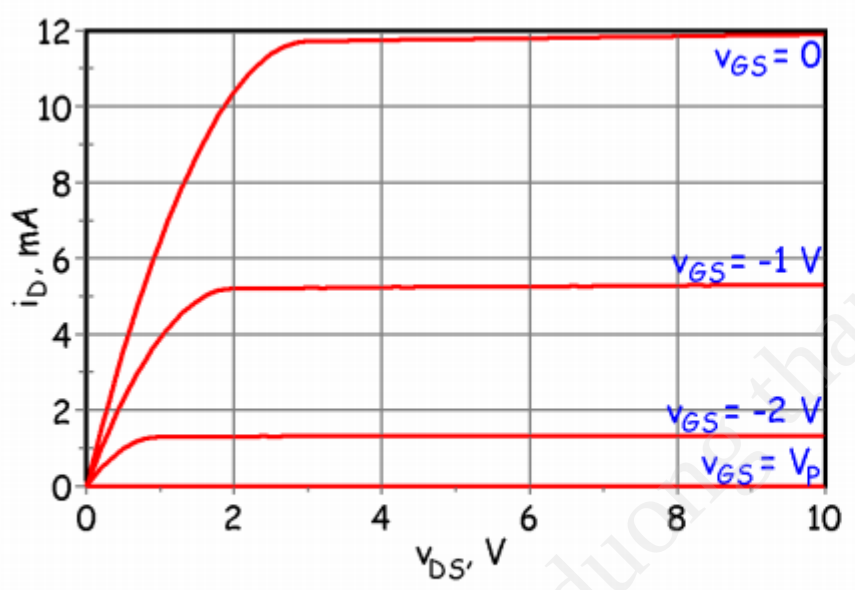


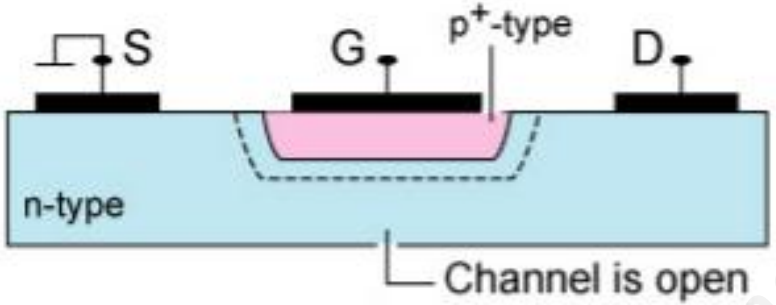
Fig. 134. *N*-channel JFET output characteristics (2N3819).

As v_{GS} changes, the curves become horizontal at different values of drain current.

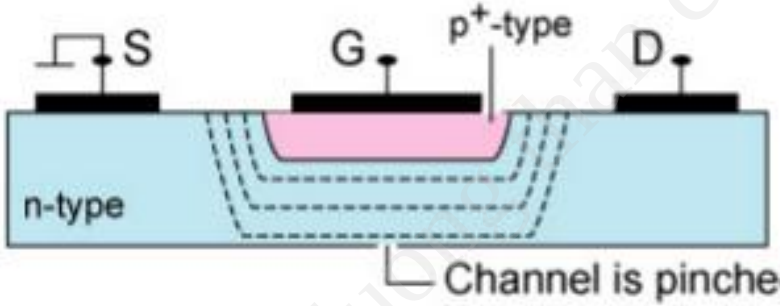
Thus, we have a device with the *output characteristics* at left.

Note that they are *very* similar to BJT curves, though the physical operation is *very* different.

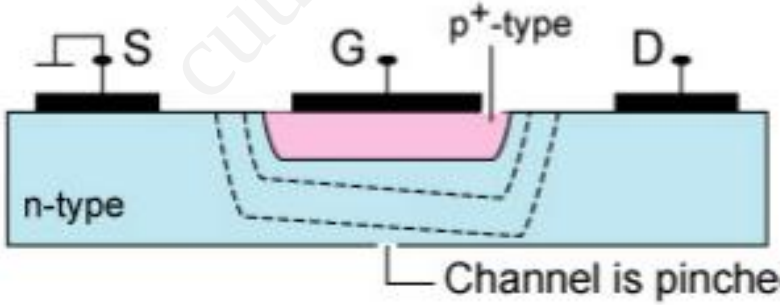
JFET



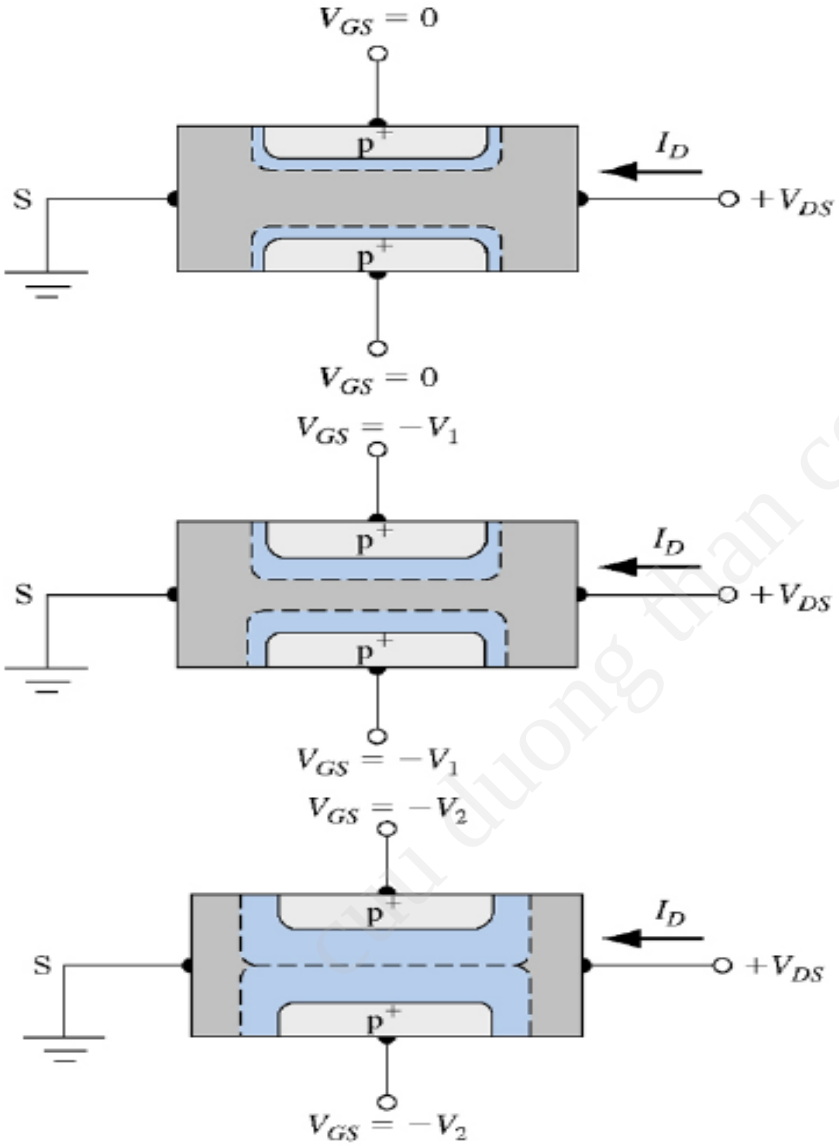
$V_{GS} = 0$
$V_{DS} = 0$

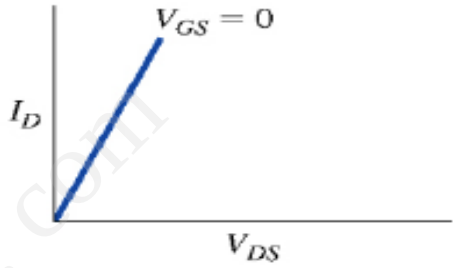


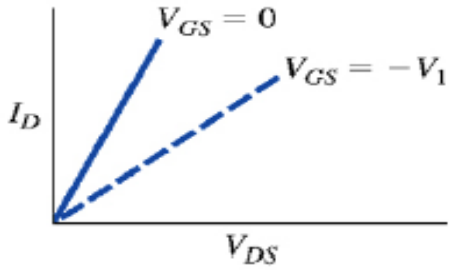
V_{GS} = reverse bias (negative)
$V_{DS} = 0$

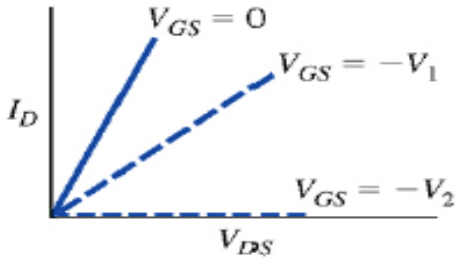


V_{GS} = reverse bias (negative)
$V_{DS} > 0$

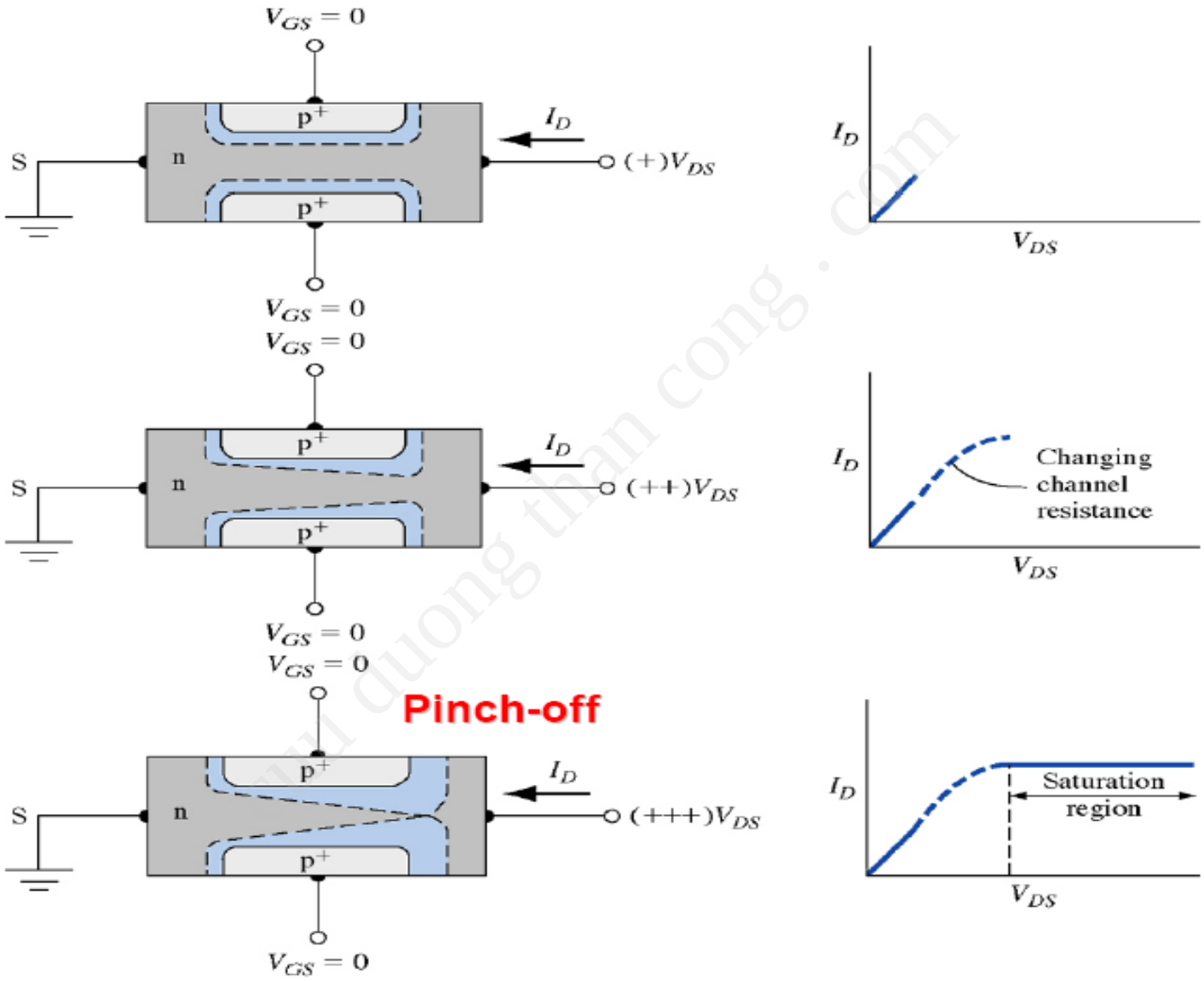








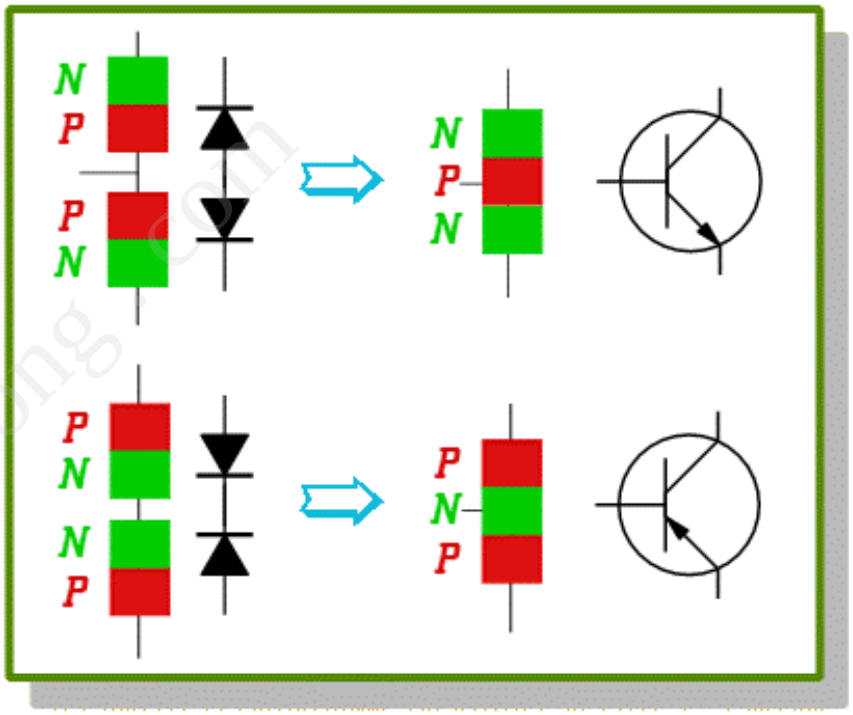
V_G controls the channel width → V_G control I_d

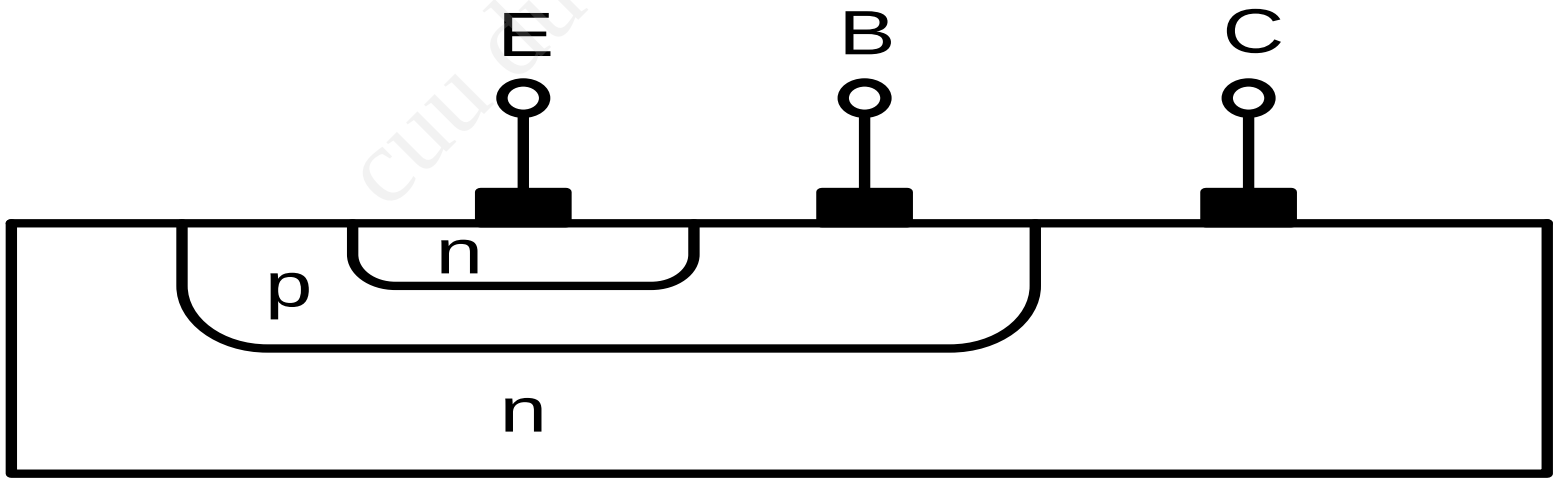


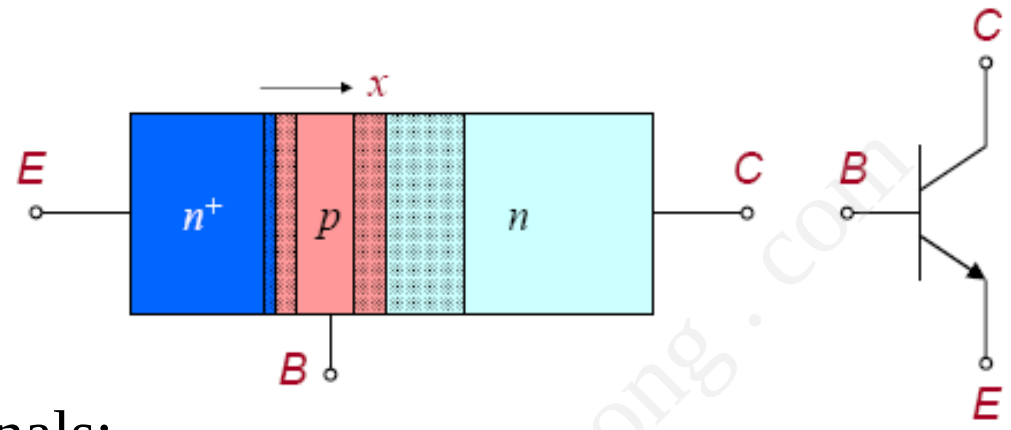
After pinch-off: $I_D \neq f(V_D)$; $I_D = f(V_G)$ - current source

BIPOLAR TRANSISTOR

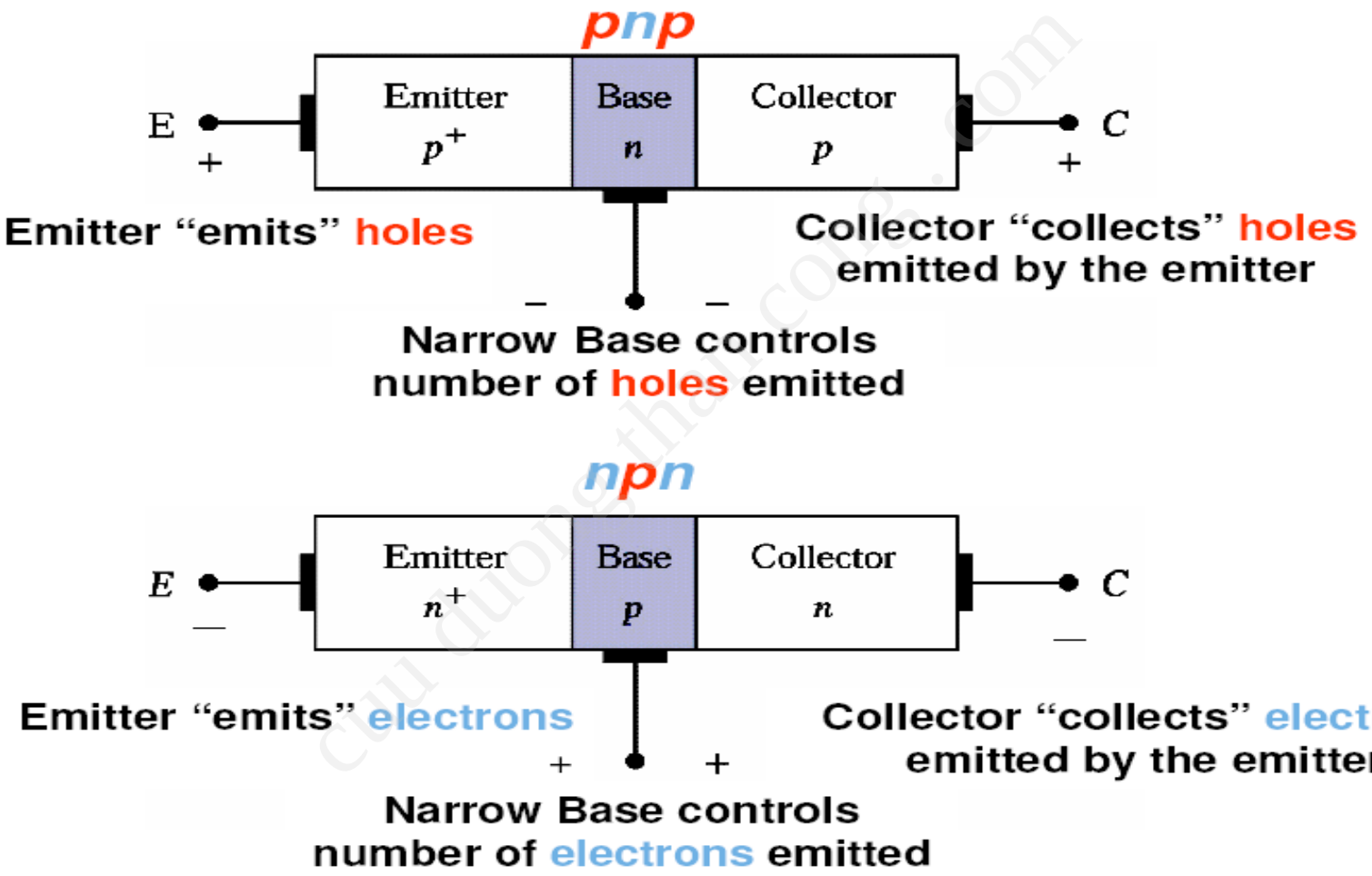
- A bipolar transistor essentially consists of a pair of PN Junction diodes that are joined back-to-back.
- There are therefore two kinds of BJT, the NPN and PNP varieties.
- The three layers of the sandwich are conventionally called the Collector, Base, Emitter.







- Three terminals:
  - Base (B): very thin and lightly doped central region (little recombination).
  - Emitter (E) and collector (C) are two outer regions sandwiching B.
- Normal operation (linear or active region):
  - B-E junction forward biased; B-C junction reverse biased.
  - The emitter emits (injects) majority charge into base region and because the base very thin, most will ultimately reach the collector.
  - The emitter is highly doped while the collector is lightly doped.
  - The collector is usually at higher voltage than the emitter.



Operation mode	Biasing polarity *B-E* junction	Biasing polarity *B-C* junction
Active	Forward	Reverse
Saturation	Forward	Forward
Cutoff	Reverse	Reverse

* Note: There is also a mode of operation called inverted (active), which is rarely used.

Active: Most widely encountered operation, e.g., as amplifiers.
Large signal gain, small signal distortion (i-v: flat region)

Saturation: Equivalent to an on state when BJT is used as a Switch
High current flow, Low voltage (in digital circuit "zero" logic level)

Cutoff: Equivalent to an off state when BJT is used as a Switch
Low current flow, High voltage (in digital circuit "one" logic level)

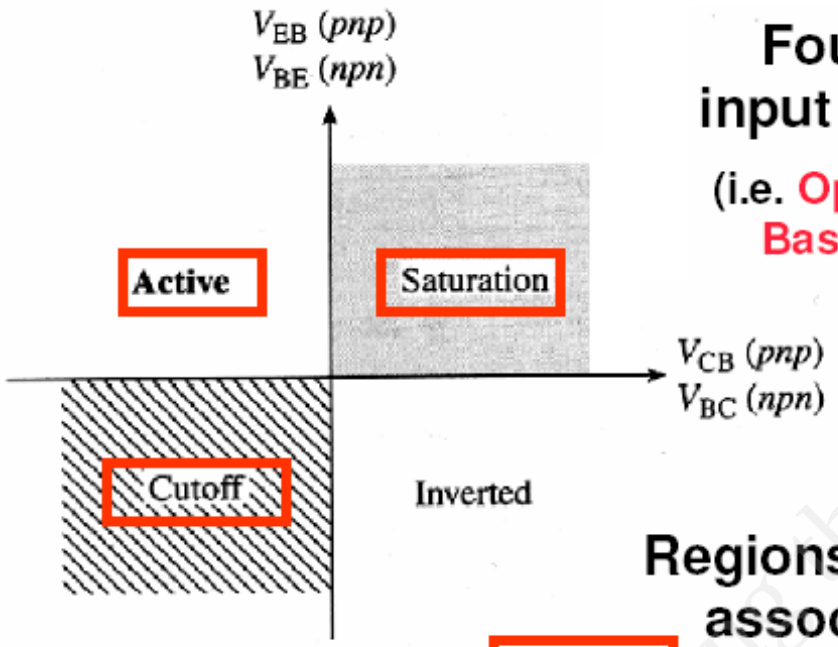
- Active:
  - Most importance mode, e.g. for amplifier operation.
  - The region where current curves are practically flat.

- Saturation:
  - Barrier potential of the junctions cancel each other out causing a virtual short.
  - Ideal transistor behaves like a closed switch.

- Cutoff:
  - Current reduced to zero
  - Ideal transistor behaves like an open switch.

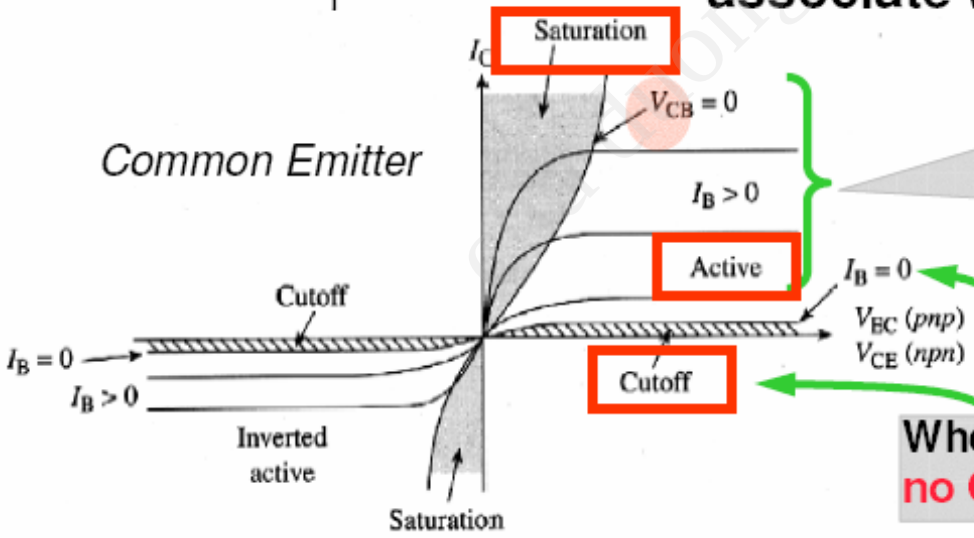
Four BJT operation modes vs. input and output voltage combinations

(i.e. Operational modes can be defined based on Base-Emitter voltage and Base-Collector voltage)



Note: the inverted (active) mode is rarely used.

Regions of the BJT CE output characteristics associate with the four operation modes.

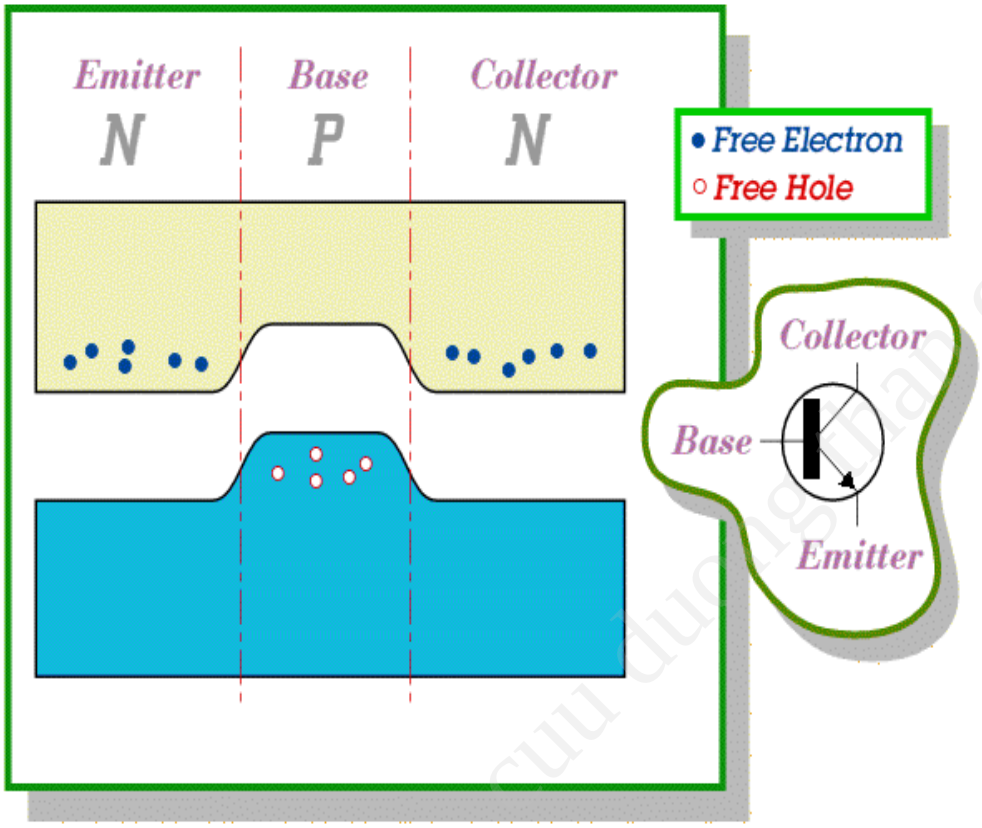


When Base current flows, a Collector current can flow.

The device is then a current controlled current device.

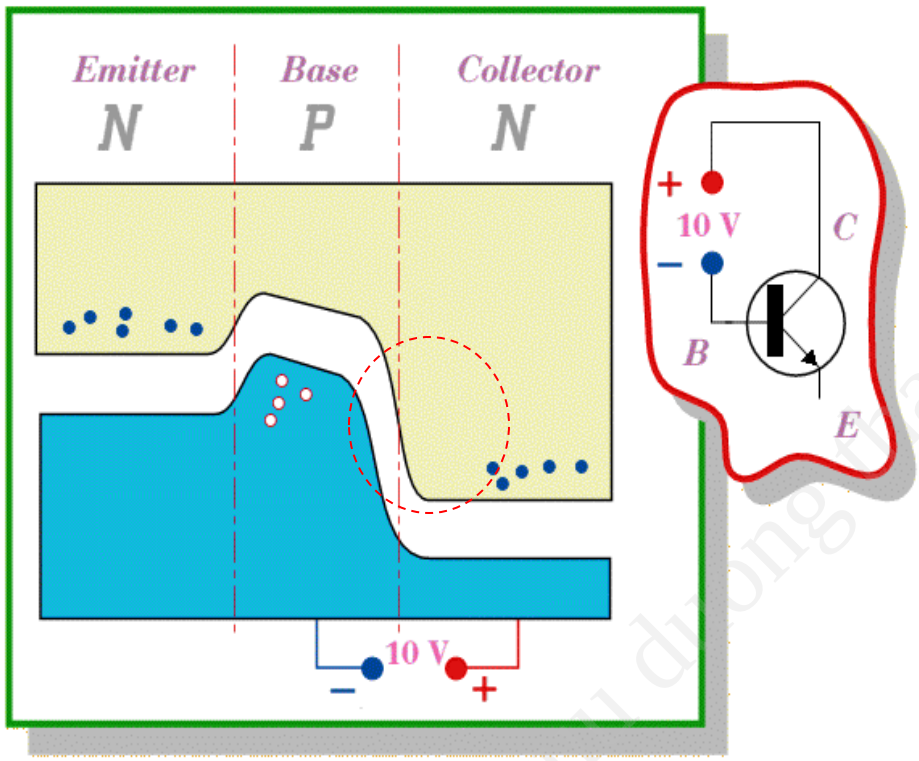
When there is no Base current, almost no Collector current flows.

How the BJT works



NPN Bipolar Transistor

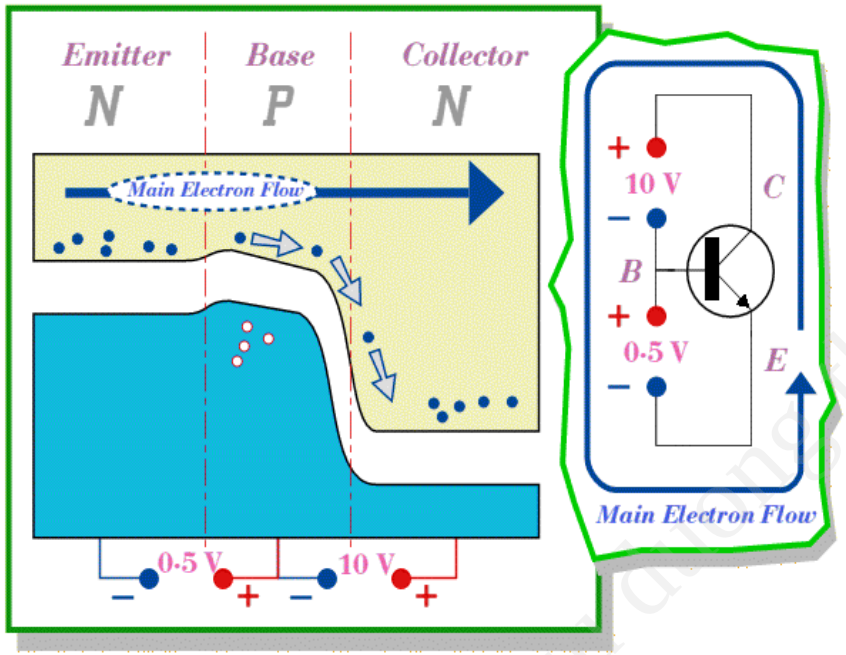
- Figure shows the energy levels in an NPN transistor under no externally applying voltages.
- In each of the N-type layers conduction can take place by the free movement of electrons in the conduction band.
- In the P-type (filling) layer conduction can take place by the movement of the free holes in the valence band.
- However, in the absence of any externally applied electric field, we find that depletion zones form at both PN-Junctions, so no charge wants to move from one layer to another.



Apply a Collector-Base voltage

- What happens when we apply a moderate voltage between the collector and base parts.
- The polarity of the applied voltage is chosen to increase the force pulling the N-type electrons and P-type holes apart.
- This widens the depletion zone between the collector and base and so no current will flow.
- In effect we have reverse-biassed the Base-Collector diode junction.

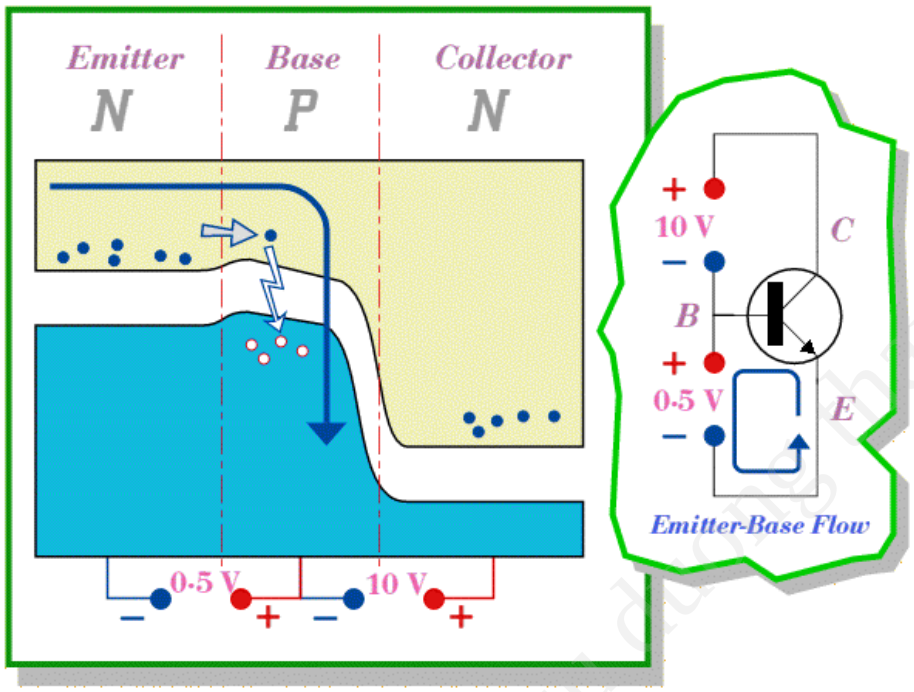
Charge Flow



Apply an Emitter-Base voltage

- What happens when we apply a relatively small Emitter-Base voltage whose polarity is designed to forward-bias the Emitter-Base junction.
- This 'pushes' electrons from the Emitter into the Base region and sets up a current flow across the Emitter-Base boundary.
- Once the electrons have managed to get into the Base region they can respond to the attractive force from the positively-biassed Collector region.
- As a result the electrons which get into the Base move swiftly towards the Collector and cross into the Collector region.
- Hence a Emitter-Collector current magnitude is set by the chosen Emitter-Base voltage applied.
- Hence an external current flowing in the circuit.

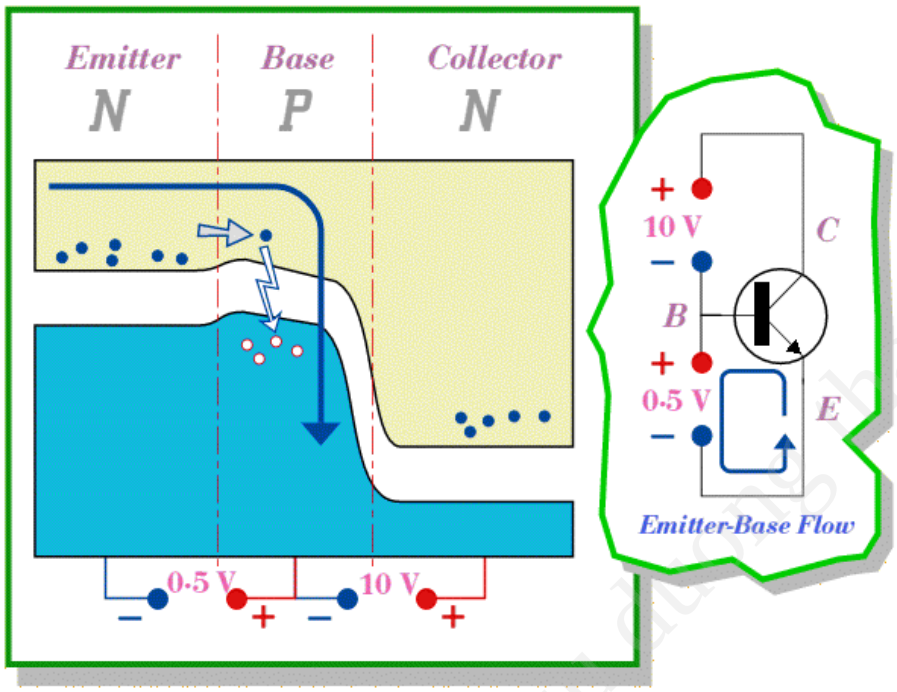
Charge Flow



Some electron fall into a hole

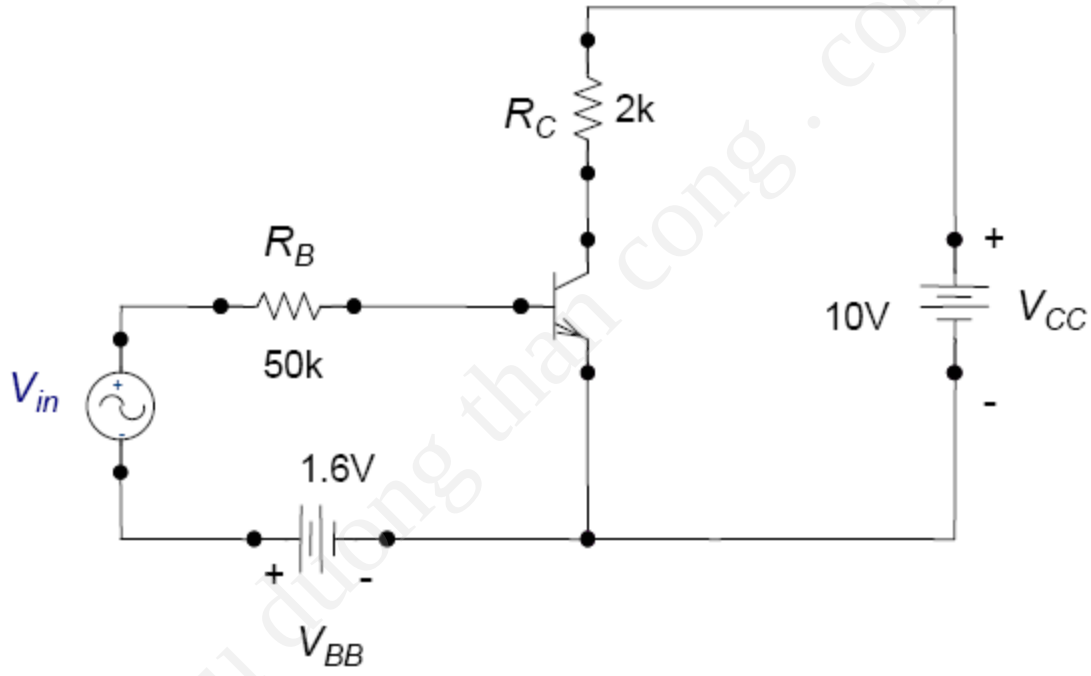
- Some of free electrons crossing the Base encounter a hole and 'drop into it'.
- As a result, the Base region loses one of its positive charges (holes).
- The Base potential would become more negative (because of the removal of the holes) until it was negative enough to repel any more electrons from crossing the Emitter-Base junction.
- The current flow would then stop.

Charge Flow

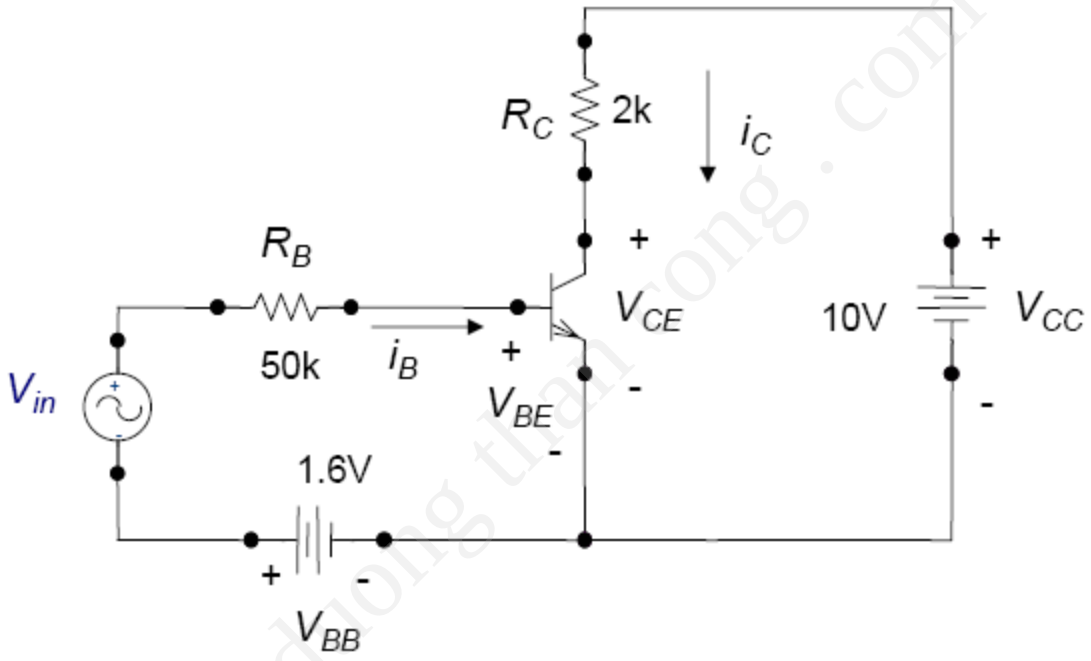


Some electron fall into a hole

- To prevent this happening we use the applied E-B voltage to remove the captured electrons from the base and maintain the number of holes.
- The effect, some of the electrons which enter the transistor via the Emitter emerging again from the Base rather than the Collector.
- For most practical BJT only about 1% of the free electrons which try to cross Base region get caught in this way.
- Hence a Base current, I_B, which is typically around one hundred times smaller than the Emitter current, I_E.



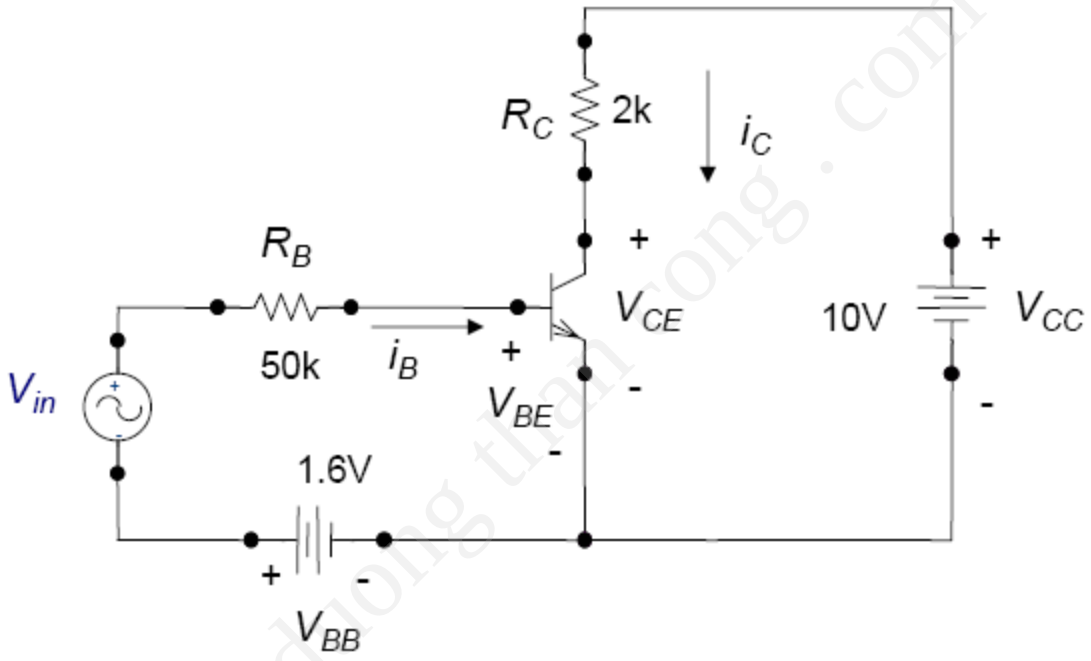
Input & Output



- We would want to know the collector current (i_C), collector-emitter voltage (V_{CE}), and the voltage across *RC*.

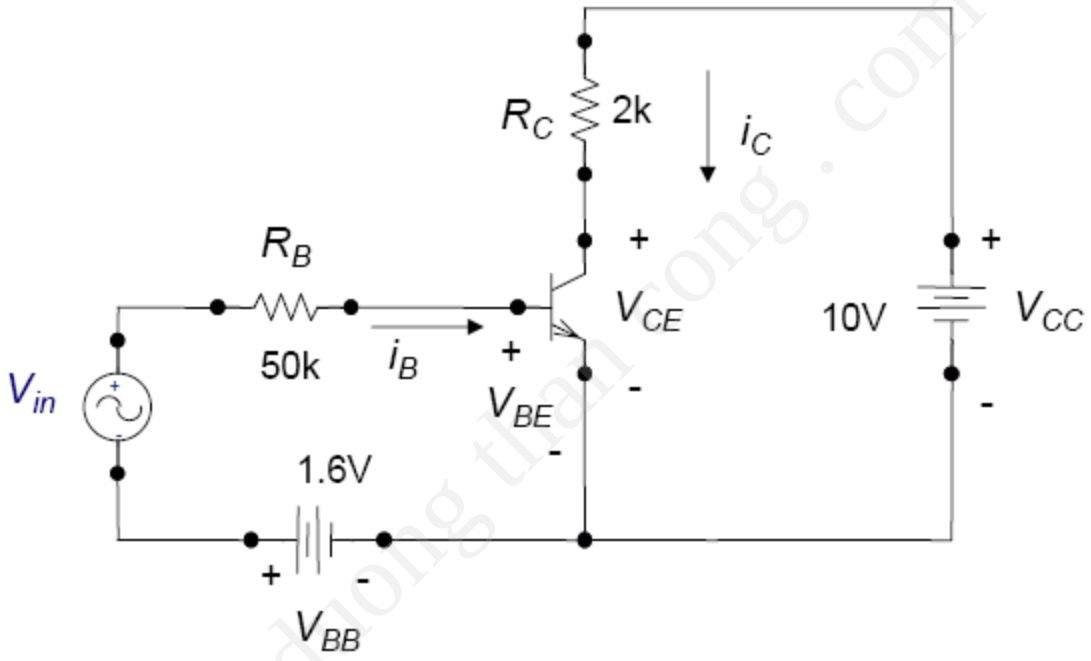
- To get this we need to fine the base current (i_B) and the base-emitter voltage (V_{BE}).

Input & Output



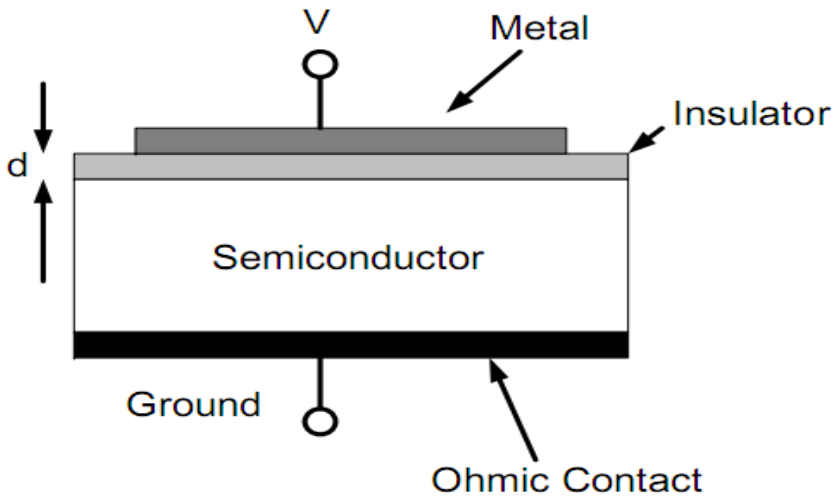
$$V_{in}(t) + V_{BB} = i_B(t)R_B + V_{BE}(t)$$

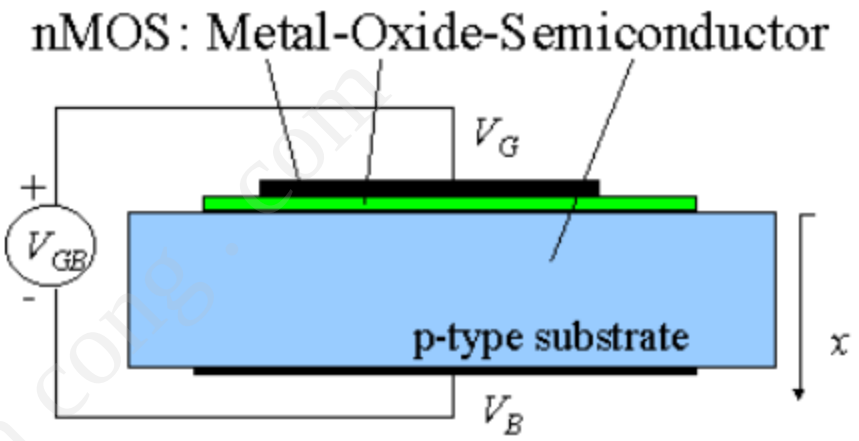
Input & Output



$$V_{CC} = i_C(t)R_C + V_{CE}(t)$$

MOSFET



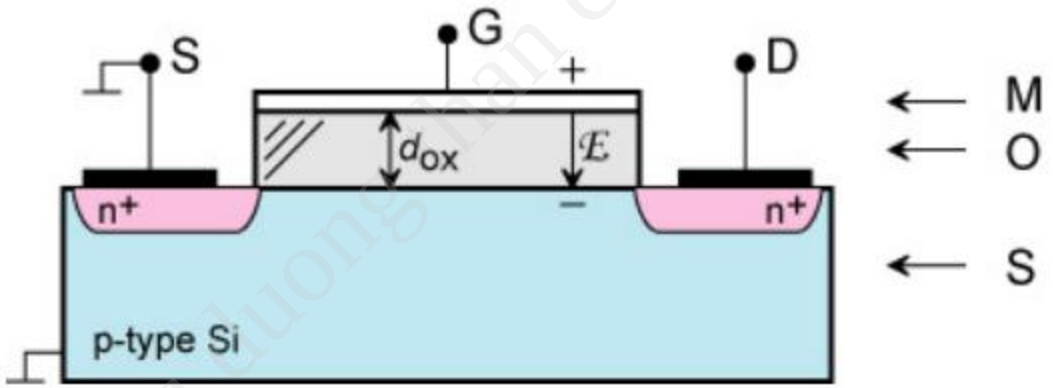


The MOS capacitor consists of a Metal-Oxide-Semiconductor structure:

- The semiconductor substrate as the bulk
- A thin oxide layer
- A top metal contact as the gate.
- A second metal layer forms an Ohmic contact to the back of the Semiconductor
- The oxide is characterized as a semiconductor with a very large bandgap, which blocks any flow of carriers between the semiconductor (p) and the gate metal.

Structure of an n-channel Si MOSFET

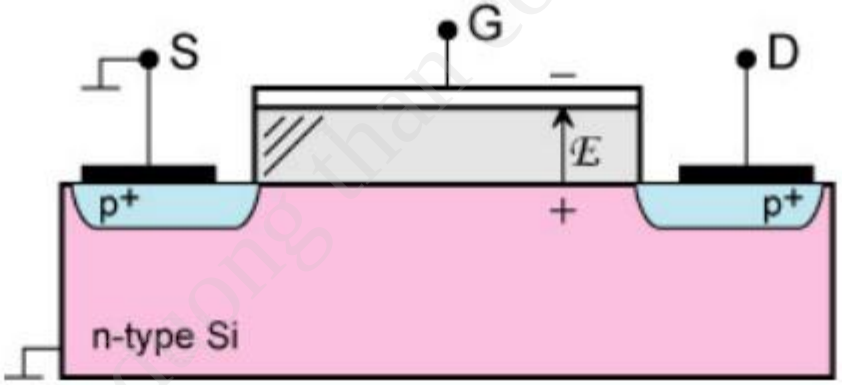
n-channel MOSFET



An electron (n-type) channel is induced in the p-type semiconductor by positive charges on the gate.

Structure of a p-channel Si MOSFET

p-channel MOSFET



A hole (p-type) channel is induced in the n-type semiconductor by negative charges on the gate.

The ideal MOS capacitor

MOS capacitor consists of

M Metal with work function Φ_M

O The oxide is SiO_2, also called silicon dioxide or silica. SiO_2 has a large gap E_g > 5eV. SiO_2 is transparent for all visible wavelengths. SiO_2 is a great insulator with a very high breakdown field.

S The semiconductor is Si. Work function Φ_{Semi}

We consider here p-type Si.

Initially we assume that

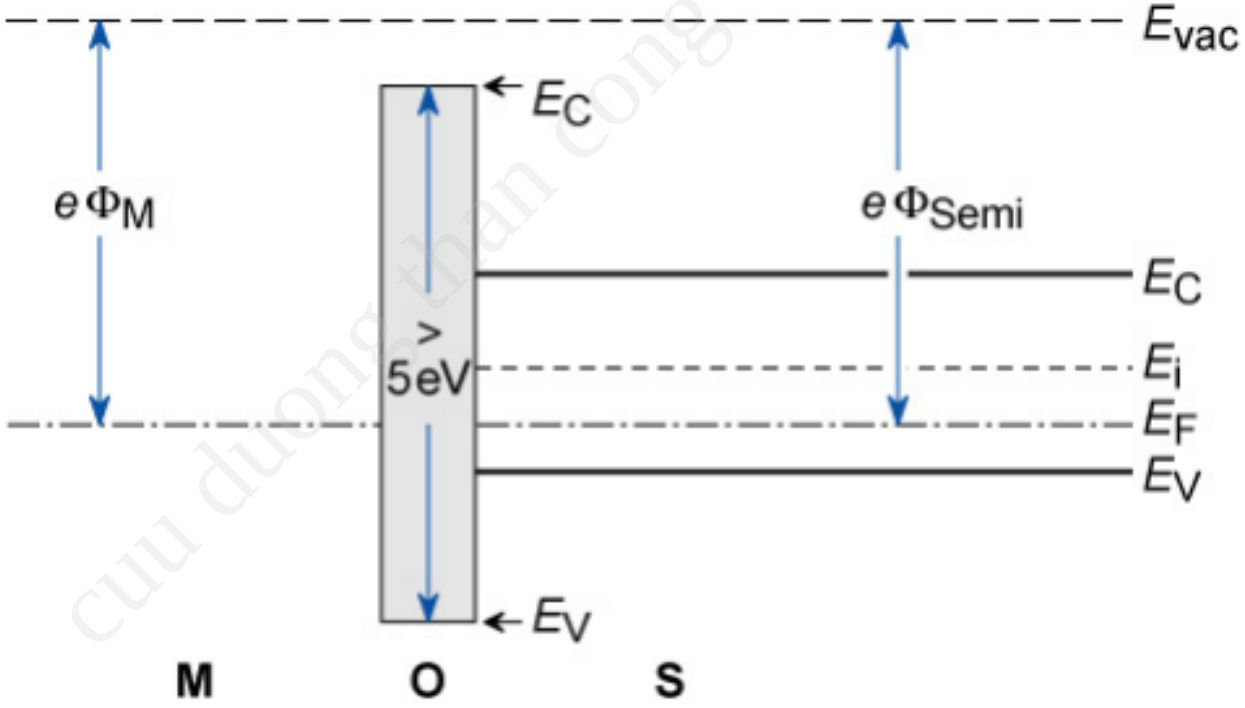
$$\Phi_M = \Phi_{Semi}$$

where Φ_M is the metal work function

and

Φ_{Semi} is the semiconductor work function.

(1) $V_G = 0$ (Equilibrium)
(E_F = constant throughout structure)



(2) $V_G > 0$ (Accumulation)

The gate bias is negative.

This means that E_F at the gate "goes up".

M and S have much higher conductivity than O.

➔ Voltage between gate and channel drops mostly across the oxide

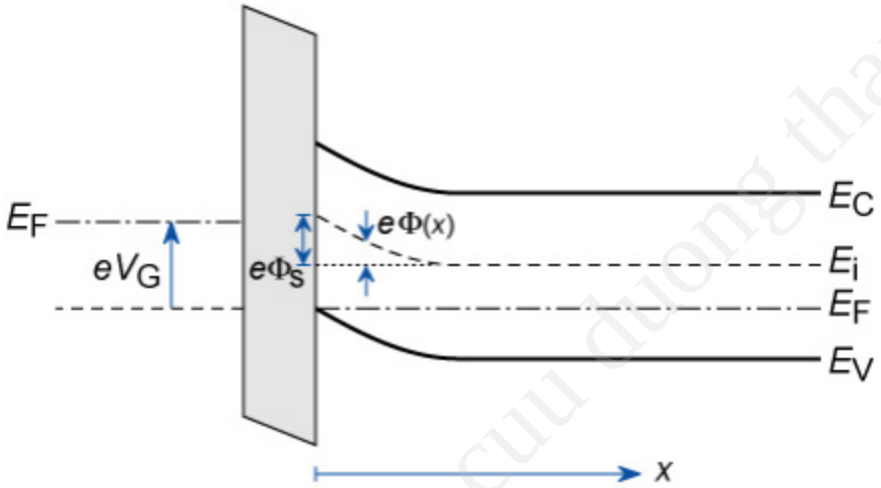
➔ An electric field is generated in oxide.

$$p(x) = n_i \, e^{(E_i - E_F)/kT} \qquad (1)$$

Definition of surface potential = Φ_S

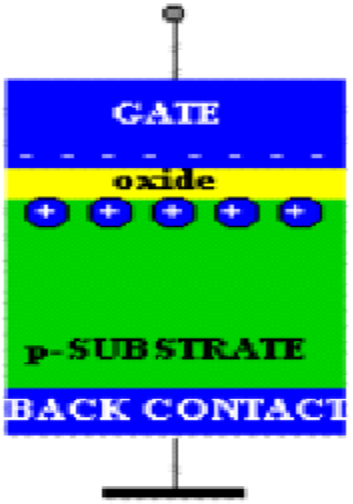
Fermi levels are different in M and S.

Fermi levels are constant within M and within S.



$p(x)$ increases near the surface ➔ **Accumulation** (*i. e.* we have an accumulation of holes near the surface)

Surface potential energy = $e\,\Phi_S$ = difference of bulk value of E_i and surface value of E_i



ACCUMULATION

When bands bend upwards:

$$E_i^{bulk} - E_i^{surface} = e\Phi_S < 0 \quad (2)$$

When bands bend downwards:

$$E_i^{bulk} - E_i^{surface} = e\Phi_S > 0 \quad (3)$$

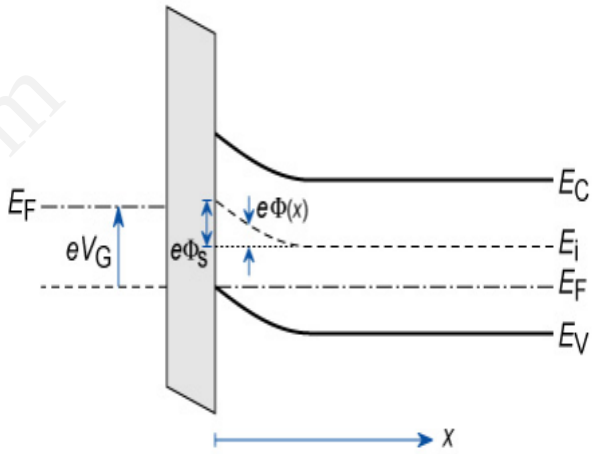
Under flatband conditions:

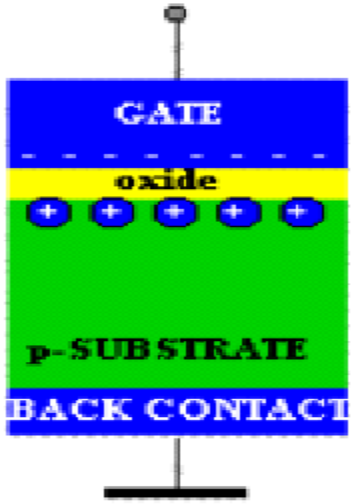
$$\Phi_S = 0 \quad (4)$$

Introducing dependence of the potential Φ on the position x:

$$e\,\Phi(x) = E_i^{bulk} - E_i(x) \quad (5)$$

$$\Phi_S = \Phi(x = 0) \quad (6)$$





ACCUMULATION

Using Eqs. (1) and (5), one obtains

$$p(x) = n_i \, e^{(E_i - E_F)/kT}$$

$$= n_i \, e^{[E_i^{bulk} - e\Phi(x) - E_F]/kT}$$

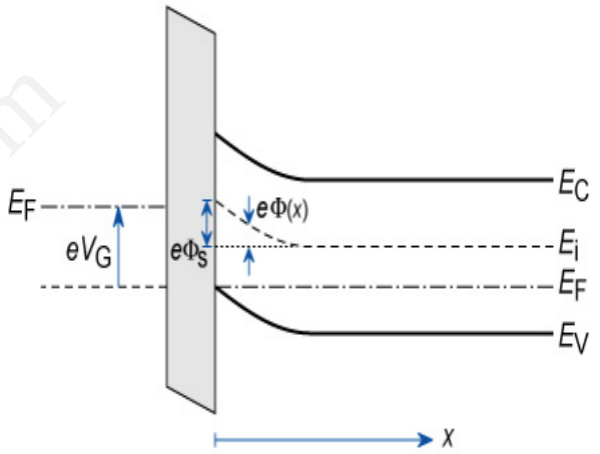
$$= n_i \, e^{(E_i^{bulk} - E_F)/kT} \, e^{-e\Phi(x)/kT}$$

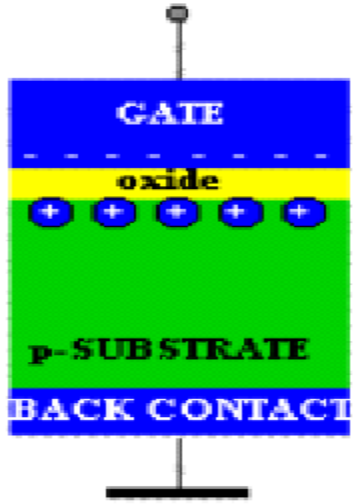
$$p(x) = p_0 \, e^{-e\Phi(x)/kT} \qquad (7)$$

Since $\Phi(x) < 0$, $p(x)$ increases close to the surface.

That is, we have accumulation.

That is the result of Eq. (7) is consistent with band diagram.





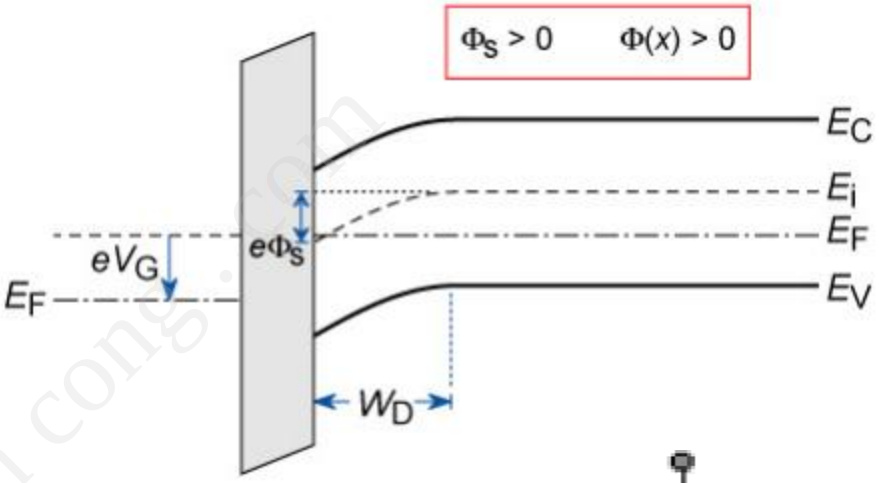
ACCUMULATION

(3) $V_G > 0$ (Depletion)

The gate bias is positive.

E_F "goes down" in the metal.

Semiconductor is depleted near surface.



The depletion layer thickness

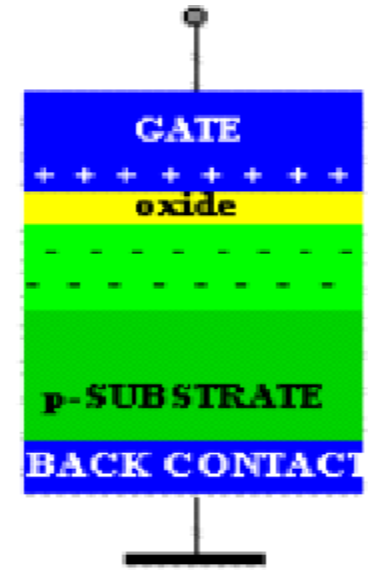
$$W_D = \sqrt{\frac{2\varepsilon}{eN_A}\Phi_S} \qquad (8)$$

E_F is near E_i at the surface.

➔ Semiconductor is practically intrinsic at the surface.

Recall Eq. (7): $p(x) = p_0 \, e^{-e\Phi(x)/kT}$

It is $\Phi(x) > 0$ ➔ $p < p_0$, that is, we have a depleted layer near the surface.



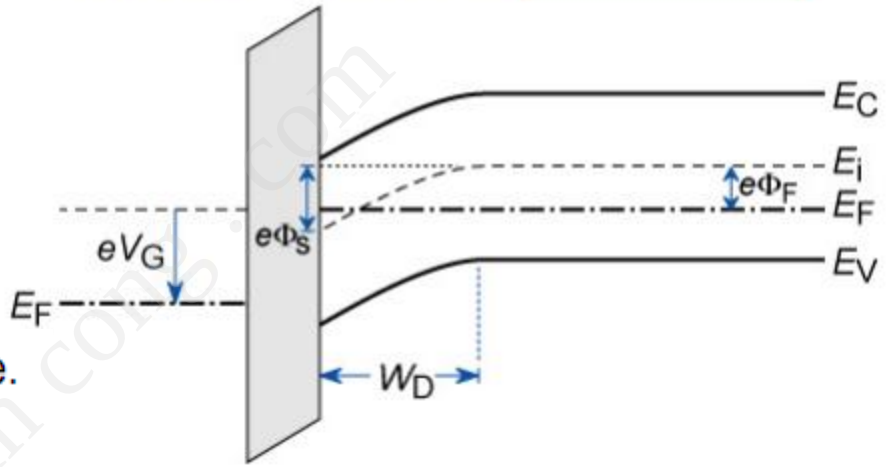
DEPLETION

(4) $V_G >> 0$. (Onset of strong inversion)

The gate bias is position.

E_F goes further down in metal.

Onset of strong inversion: $e\Phi_s = 2\, e\Phi_F$

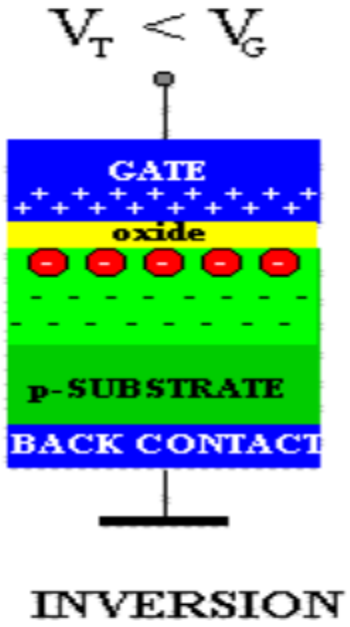


Semiconductor is depleted of holes near surface.

E_F is closer to E_C than to E_V at the surface.

➔ Semiconductor is n-type near surface

➔ Conductivity type of semiconductor is inverted.



Criterion for the onset of strong inversion:

$$e\Phi_S = 2\, e\Phi_F \qquad \text{(Onset of strong inversion)} \qquad (9)$$

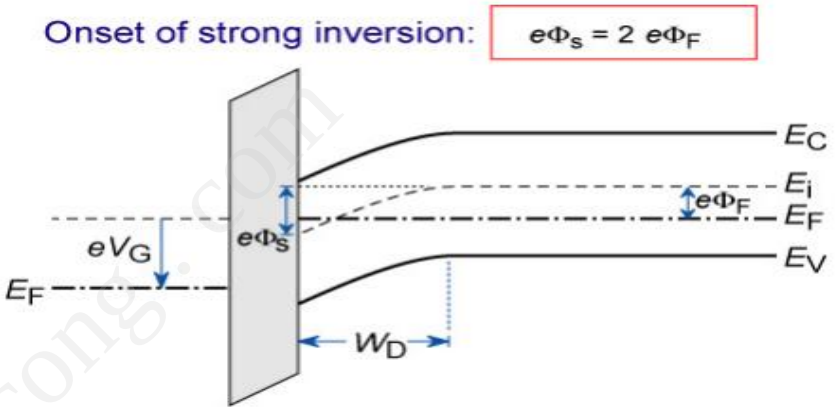
where

$$e\Phi_F = E_i^{bulk} - E_F^{bulk} \qquad (10)$$

(4) $V_G >> 0$. (Onset of strong inversio

The gate bias is position.

E_F goes further down in metal.



Onset of strong inversion means that the semiconductor is as strongly n-type at the surface as it is p-type in the bulk.

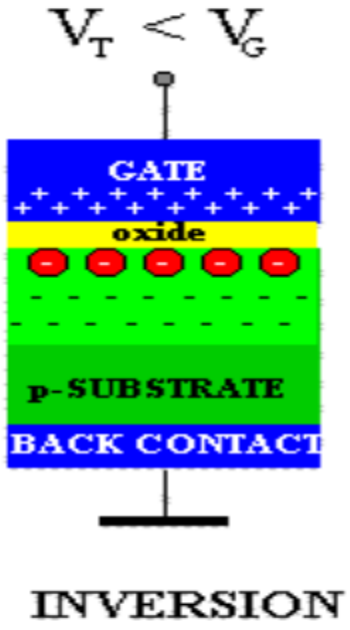
Using Boltzmann statistics

$$p = n_i \, e^{(E_i - E_F)/kT} \qquad (11)$$

and Eqs. (9) and (10), one obtains

$$e\Phi_S = 2e\Phi_F = 2kT \ln \frac{N_A}{n_i} \quad \text{(Onset of strong inversion)} \quad (12)$$

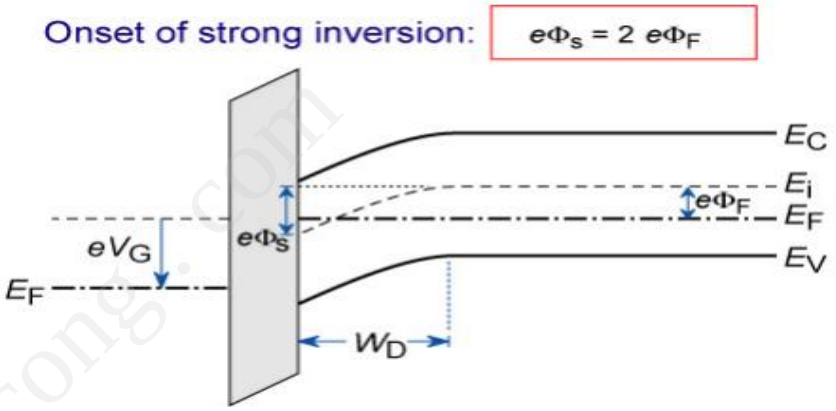
At the onset of strong inversion, an n-channel begins to be formed at the semiconductor surface.



(4) $V_G >> 0$. (Onset of strong inversion)

The gate bias is position.

E_F goes further down in metal.



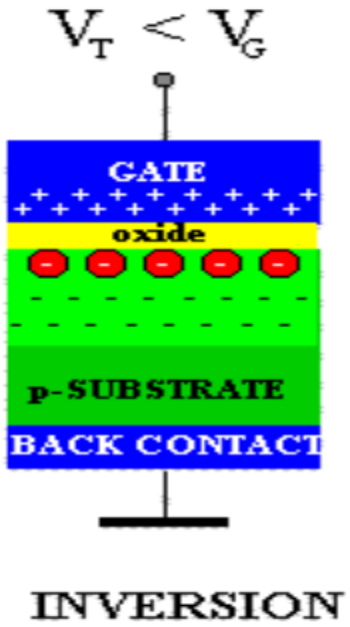
The depletion layer thickness at the onset of strong inversion is given by:

$$W_{D,\max} = \sqrt{\frac{2\varepsilon}{eN_A} 2\Phi_F} = \sqrt{\frac{2\varepsilon}{eN_A} 2 \frac{kT}{e} \ln \frac{N_A}{n_i}}$$

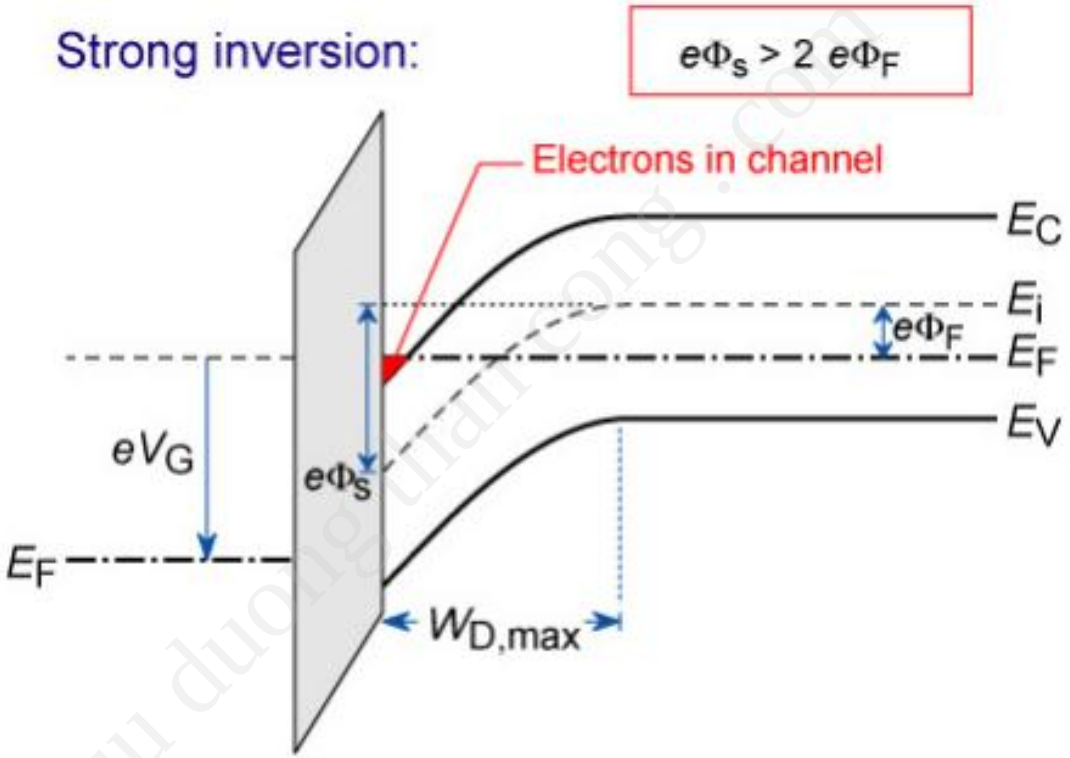
$$W_{D,\max} = 2\sqrt{\frac{\varepsilon\, kT}{e^2 N_A} \ln \frac{N_A}{n_i}} \qquad (13)$$

$W_{D,\max}$ is the maximum depletion layer thickness.

A further increase in V_G will result in more inversion rather than in more depletion.



(5) Let's go beyond the onset of strong inversion.



Beyond the onset of strong inversion, electrons are filled into the electron channel. The depletion layer thickness does not increase further, *i. e.* $W_D = W_{D,\,max}$.

Real MOS capacitor

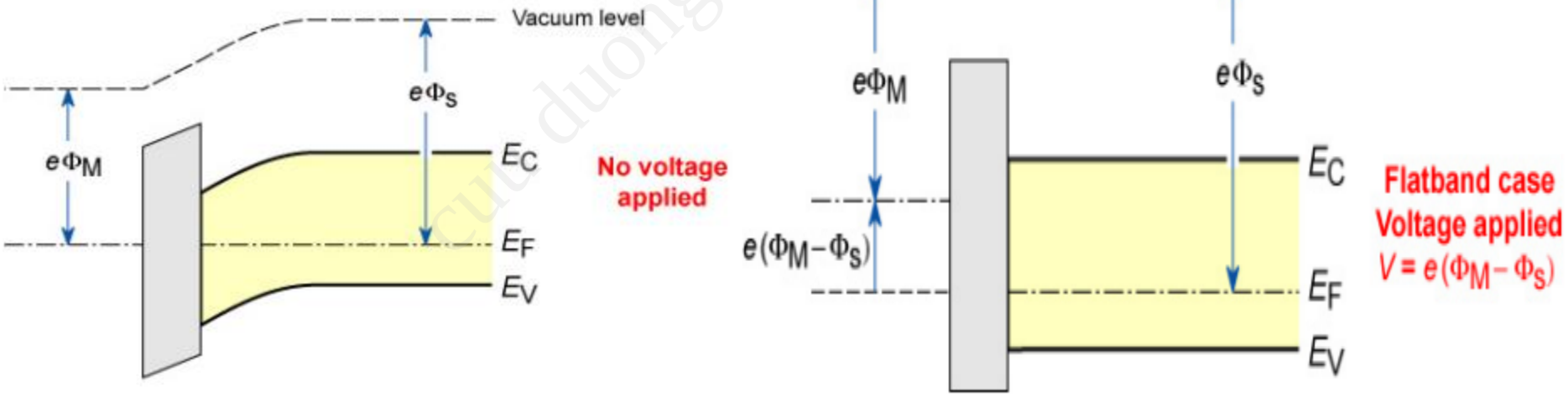
Generally, there is a work function difference between metal and semiconductor. That is

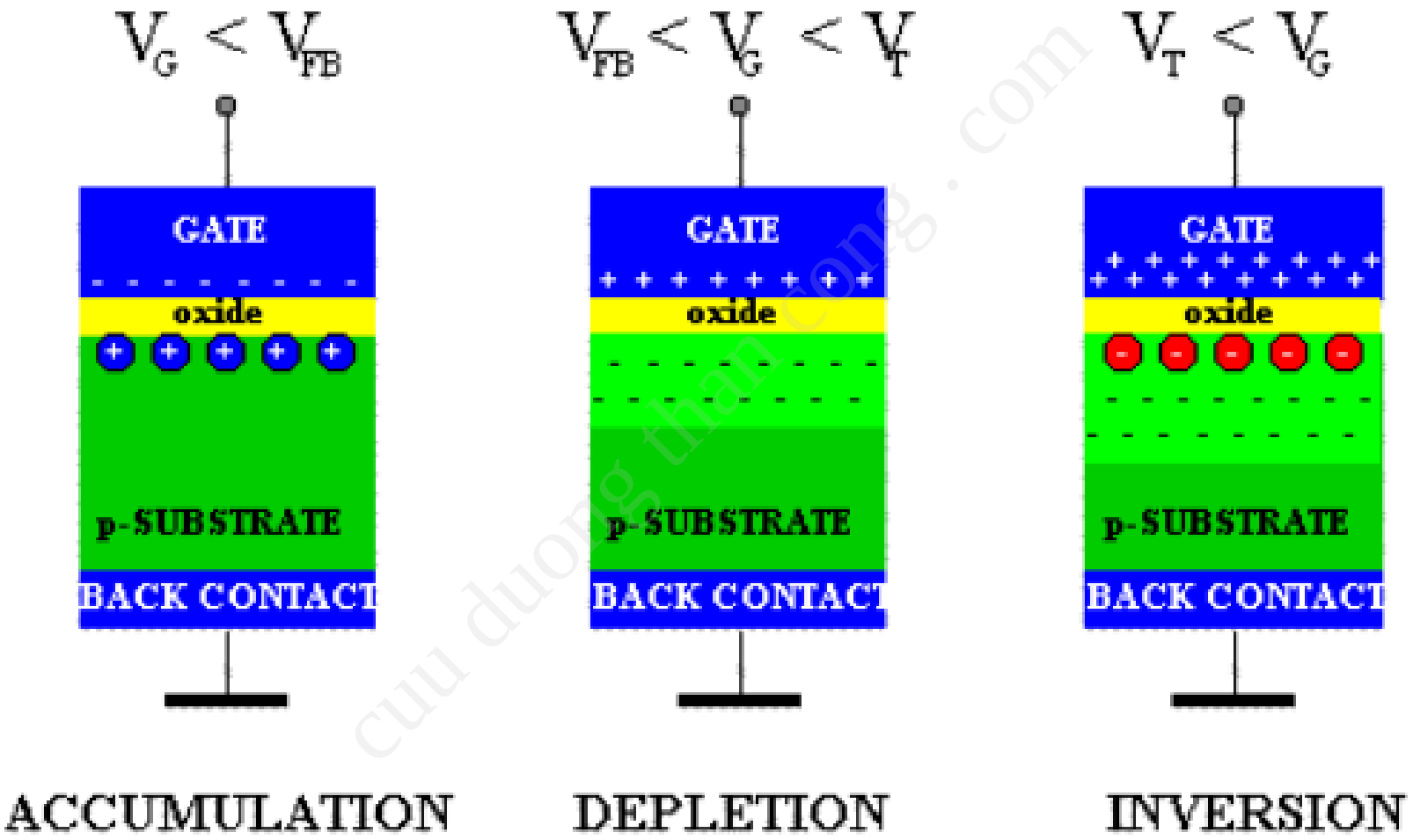
$$\Phi_M \neq \Phi_S$$

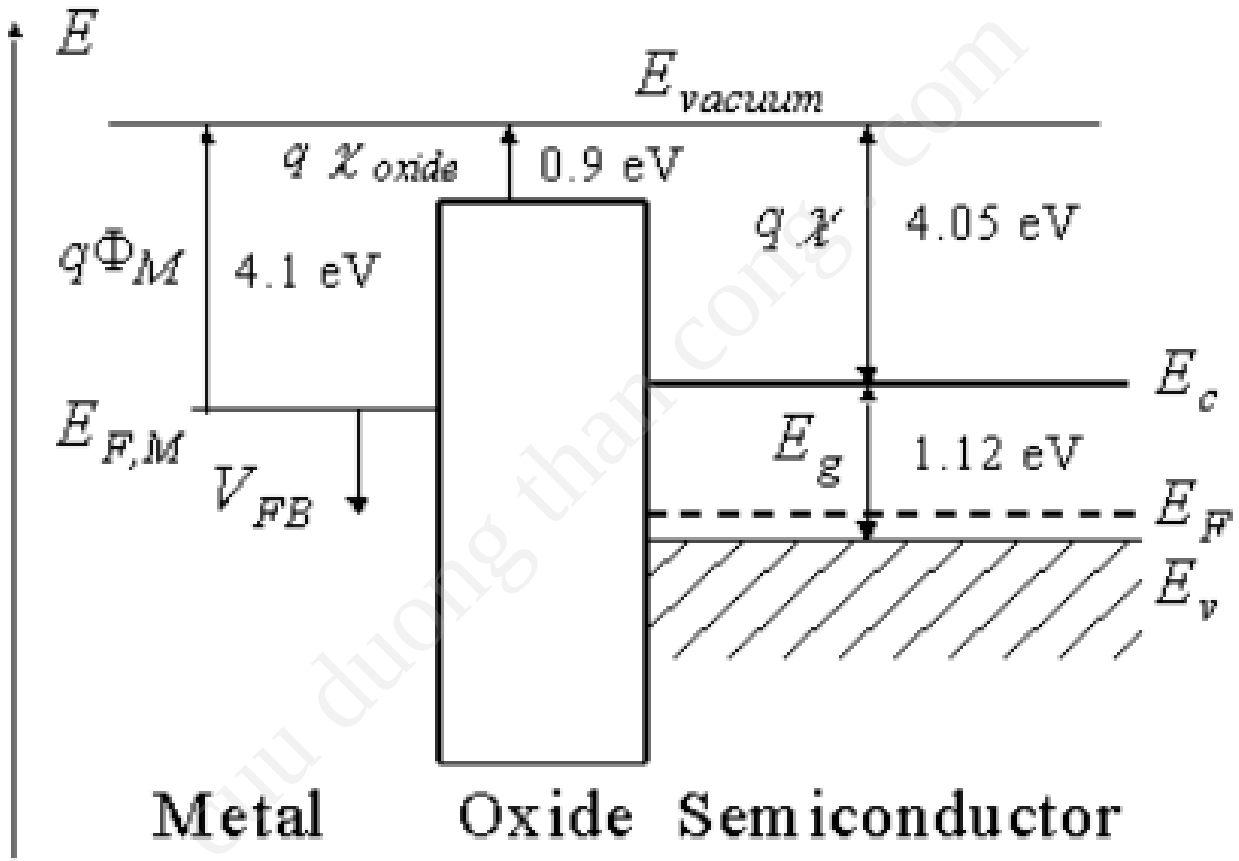
Work function difference

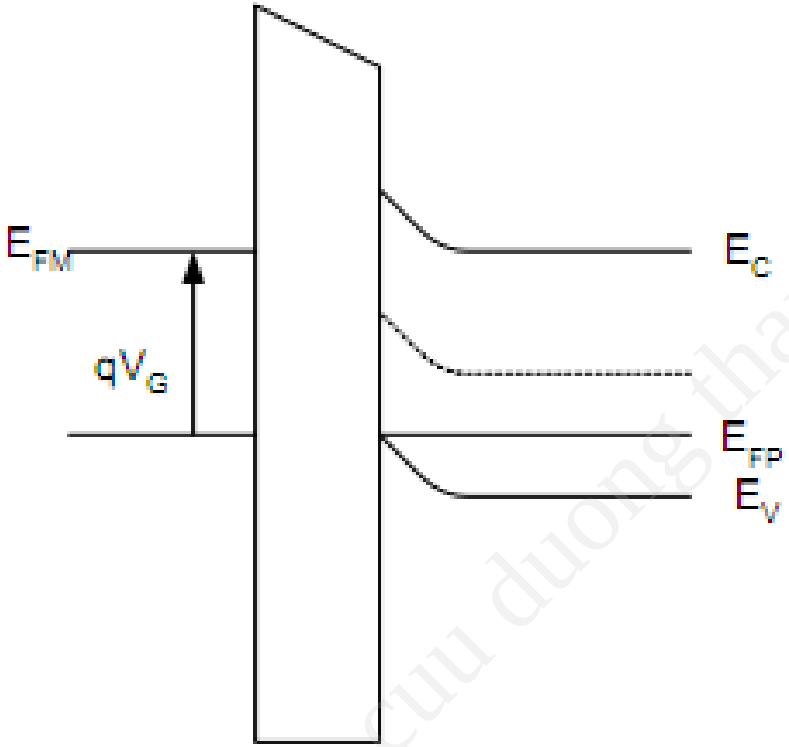
$$\Phi_{MS} = \Phi_M - \Phi_S$$

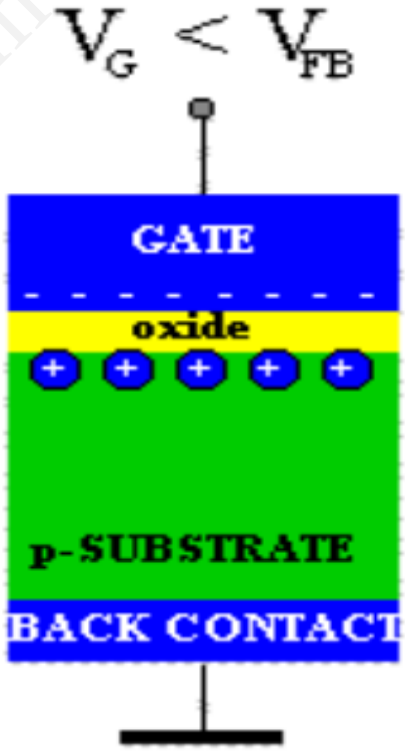
Band diagram (for $\Phi_{MS} \neq 0$)



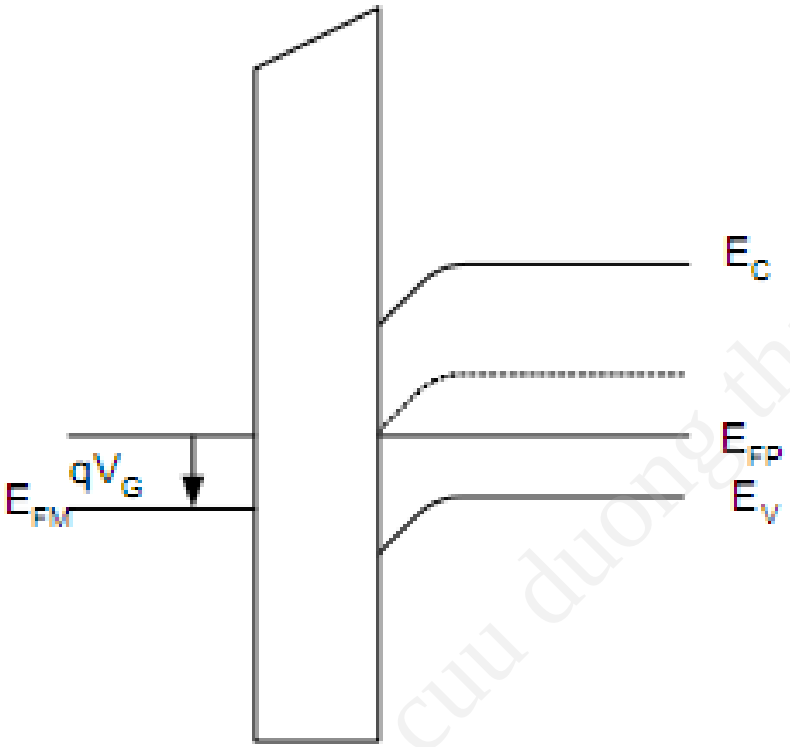


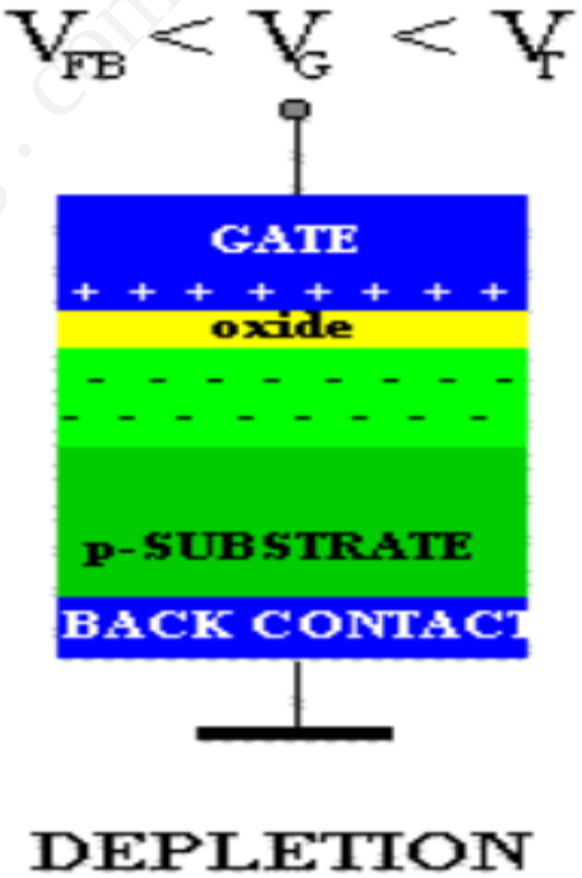




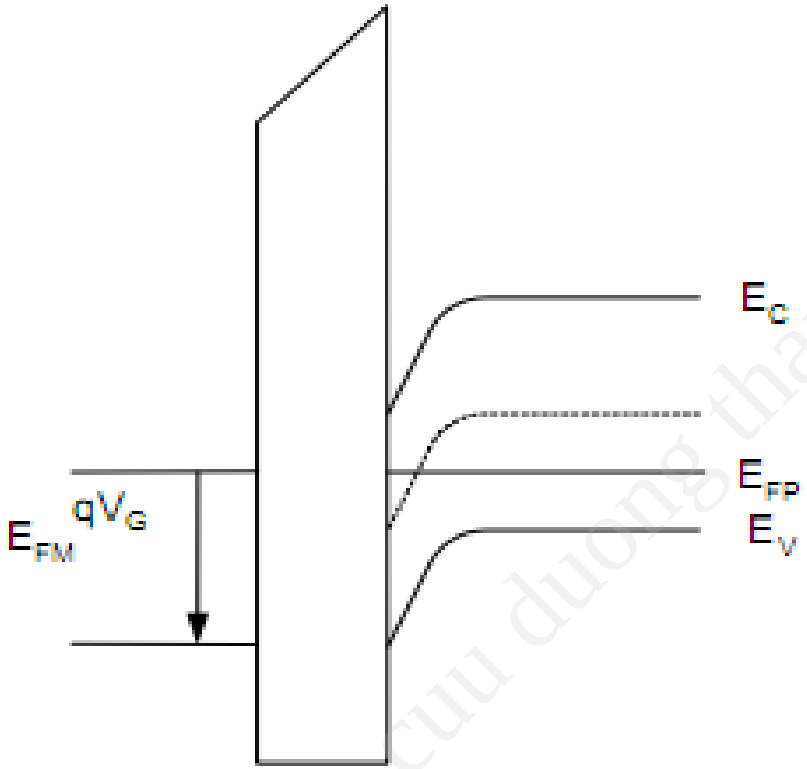


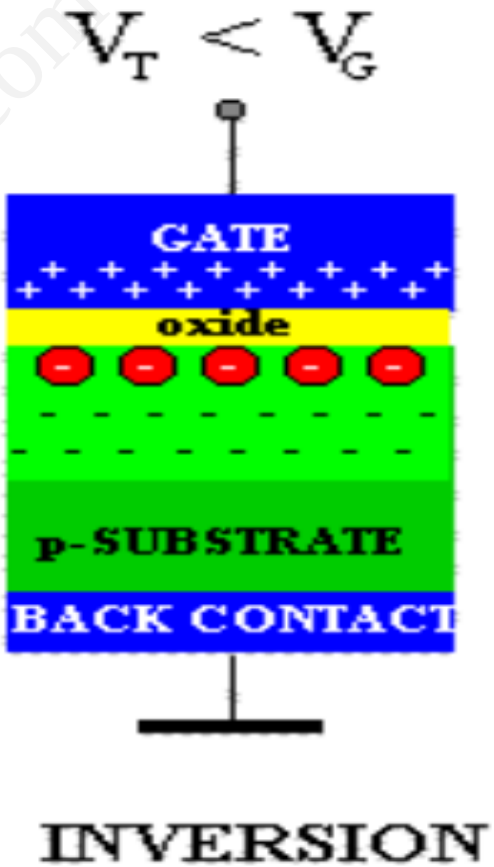
ACCUMULATION





DEPLETION





Basic MOSFET operation

- n-channel
  - A positive voltage is applied to the gate thus attracting electrons which form a channel between the two n^+ source and drain regions
- p-channel
  - A negative voltage is applied to the gate thus attracting hole which form a channel between the two p^+ source and drain regions

Basic MOSFET operation

- n-channel
  - Requires a gate voltage more positive than V_T to create an electron channel.
- p-channel
  - Requires a gate voltage more negative than V_T to create an electron channel.

Energy band diagrams and charge density diagrams MOSFET (n Si)

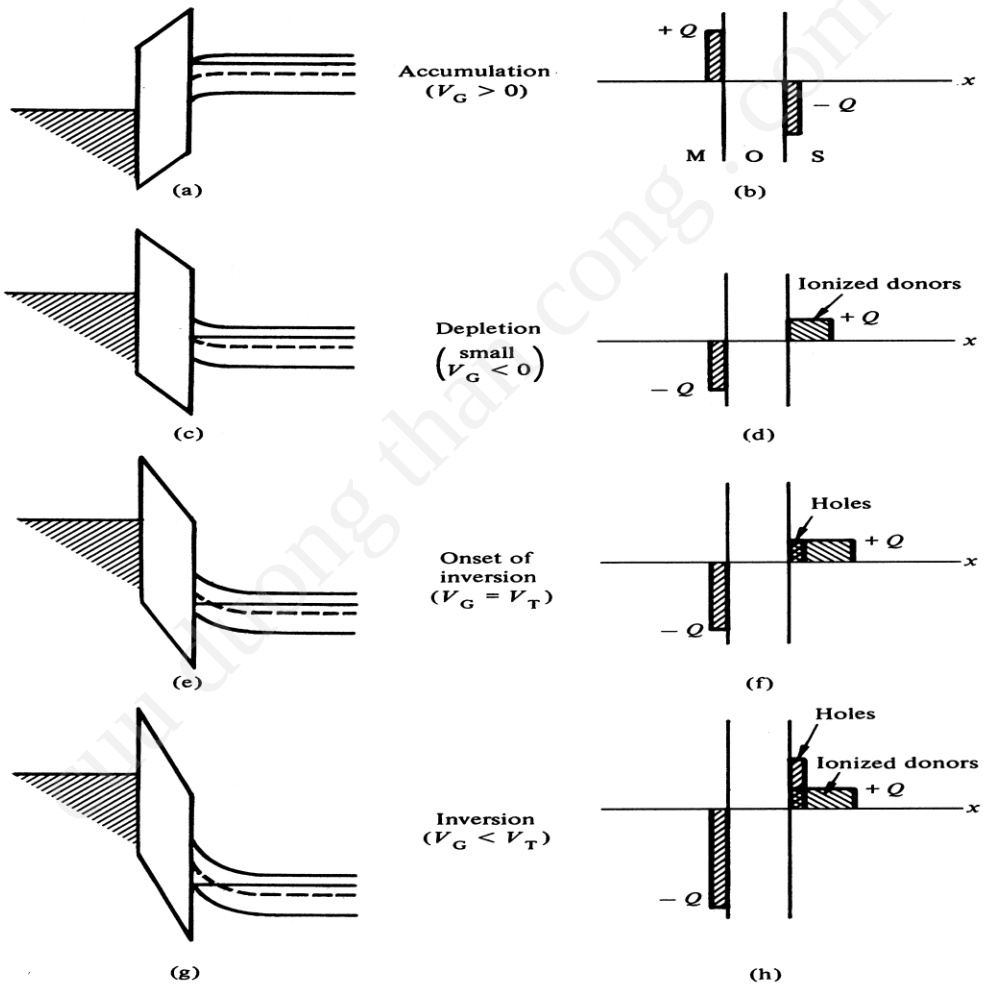


Figure 16.5

Energy band diagrams and charge density diagrams MOSFET (p Si)

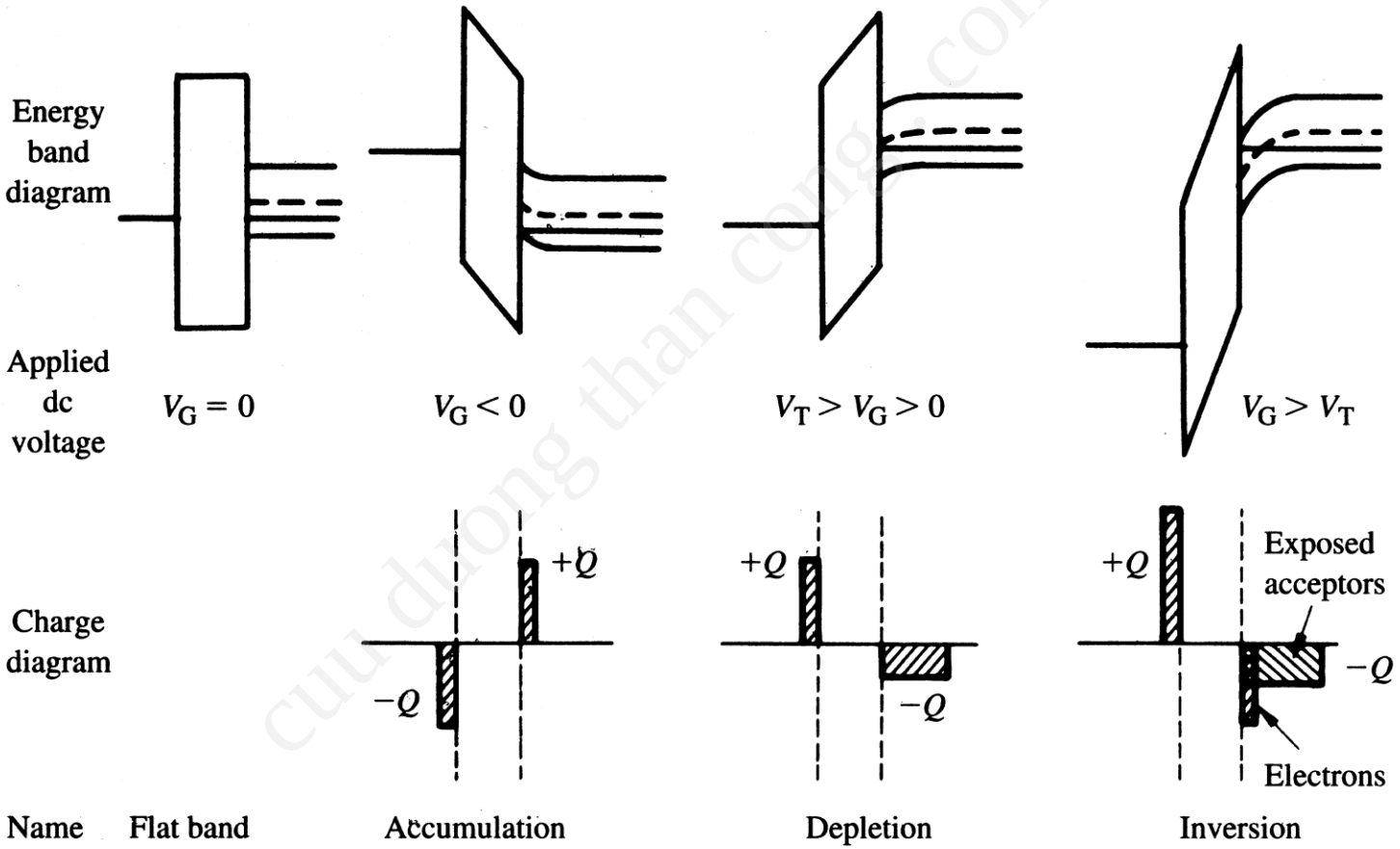


Figure 16.6

MOS Capacitor

Capacitor under bias Summary