

3.15
Carrier Fundamentals
C.A. Ross, Department of Materials Science and Engineering

Reference: Pierret, chapters 1-2.

Electron and hole charge: $e = 1.6 \cdot 10^{-19} \text{ C}$

Effective mass: m^* , rest mass m_o

$$F = -eE = m_o \frac{dv}{dt} \quad \text{in vacuum}$$

$$F = -eE = m^* \frac{dv}{dt} \quad \text{in solid}$$

in Si, $m_n^*/m_o = 1.18$, $m_h^*/m_o = 0.81$ at 300K.

Intrinsic properties

in Si, $n = p = 10^{10} \text{ cm}^{-3}$ at 300K
 $N = 5 \cdot 10^{22} \text{ atoms cm}^{-3}$

Extrinsic properties

Donors – group V

Acceptors – group III

B	C	N	O	
Al	Si	P	S	
Ga	Ge	As	Se	
In	Sn	Sb	Te	
Tl	Pb	Bi	Po	

Band diagrams: E_c = conduction band edge, E_v = valence band edge

band gaps: Si 1.12 eV

diamond 5.4 eV

silica 5 eV

energies of dopant levels, in meV, in silicon (kT = 26 meV @ 300K)

P	45	B	45
---	----	---	----

As	54	Al	67
----	----	----	----

Ga	72
----	----

Carrier distributions (intrinsic)

$g(E) dE$ = density of electron states cm^{-3} in the interval $(E, E+dE)$,
units $\#/ \text{eV.cm}^{-3}$

$$g_c(E) dE = m_n^* \sqrt{(2m_n^*(E - E_c)) / (\pi^2 \bar{h}^3)} dE$$

$$g_v(E) dE = m_p^* \sqrt{(2m_p^*(E_v - E)) / (\pi^2 \bar{h}^3)} dE$$

In these states, the electrons distribute according to Fermi function

$$f(E) = 1 / \{1 + \exp(E - E_f)/kT\}$$

Number of electrons in the interval $(E, E+dE)$ is therefore $f(E)g(E)dE$.

In a doped semiconductor, the position of E_f with respect to the band gap determines whether there are more electrons or holes.

Total number of electrons: by integrating $f(E)g(E)dE$

$$n = n_i \exp(E_f - E_i)/kT$$

$$p = n_i \exp(E_i - E_f)/kT$$

where

$$n_i = N_c \exp(E_i - E_c)/kT$$

$N_c = 2\{2\pi m_n^* kT/\hbar^2\}^{3/2}$ = 'effective density of conduction band states'

E_i is the position of the Fermi level in the intrinsic case.

Similarly for N_v .

Hence

$$np = n_i^2 \text{ at equilibrium}$$

$$n_i^2 = N_c N_v \exp(E_v - E_c)/kT = N_c N_v \exp(-E_g)/kT$$

Intrinsic case:

$$E_i = (E_v + E_c)/2 + 3/4 kT \ln(m_p^*/m_n^*)$$

In a doped material, where $n \sim N_D$ or $p \sim N_A$

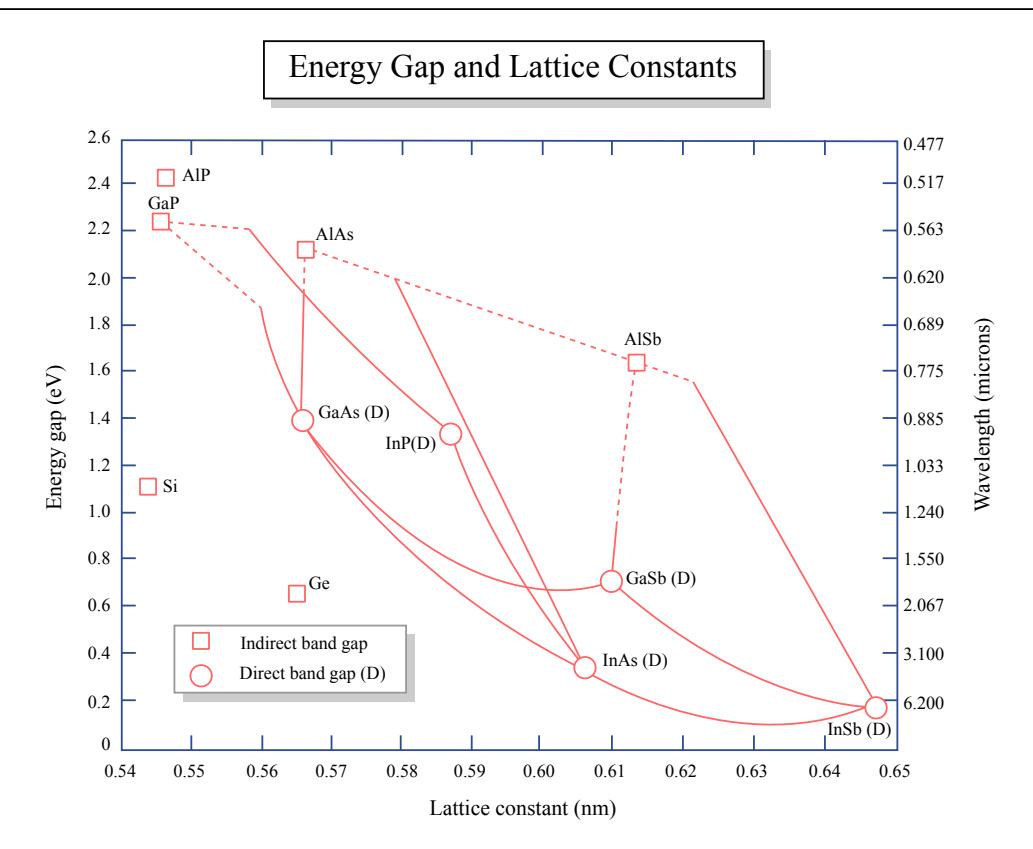
$$\begin{aligned} E_f - E_i &= kT \ln(n/n_i) = -kT \ln(p/n_i) \\ &\sim kT \ln(N_D/n_i) \quad \text{or} \quad -kT \ln(N_A/n_i) \\ \text{n-type} & & \text{p-type} \end{aligned}$$

Properties	Si	GaAs	SiO ₂	Ge
Atoms/cm ³ , molecules/cm ³ × 10 ²²	5.0	4.42	2.27 ^a	
Structure	diamond	zincblende	amorphous	
Lattice constant (nm)	0.543	0.565		
Density (g/cm ³)	2.33	5.32	2.27 ^a	
Relative dielectric constant, ϵ_r	11.9	13.1	3.9	
Permittivity, $\epsilon = \epsilon_r \epsilon_0$ (farad/cm) × 10 ⁻¹²	1.05	1.16	0.34	
Expansion coefficient (dL/LdT) × (10 ⁻⁶ K)	2.6	6.86	0.5	
Specific Heat (joule/g K)	0.7	0.35	1.0	
Thermal conductivity (watt/cm K)	1.48	0.46	0.014	
Thermal diffusivity (cm ² /sec)	0.9	0.44	0.006	
Energy Gap (eV)	1.12	1.424	~9	0.67
Drift mobility (cm ² /volt-sec)				
Electrons	1500	8500		
Holes	450	400		
Effective density of states (cm ⁻³) × 10 ¹⁹				
Conduction band	2.8	0.047		
Valence band	1.04	0.7		
Intrinsic carrier concentration (cm ⁻³)	1.45 × 10 ¹⁰	1.79 × 10 ⁶		

Mayer and
Lau,
Electronic
Materials
Science

Properties of Si, GaAs, SiO₂, and Ge at 300 K

Figure by MIT OCW.



PHYSICAL CONSTANTS, CONVERSIONS, AND USEFUL COMBINATIONS

Physical Constants

Avogadro constant	$N_A = 6.022 \times 10^{23}$ particles/mole
Boltzmann constant	$k = 8.617 \times 10^{-5}$ eV/K = 1.38×10^{-23} J/K
Elementary charge	$e = 1.602 \times 10^{-19}$ coulomb
Planck constant	$h = 4.136 \times 10^{-15}$ eV · s $= 6.626 \times 10^{-34}$ joule · s
Speed of light	$c = 2.998 \times 10^{10}$ cm/s
Permittivity (free space)	$\epsilon_0 = 8.85 \times 10^{-14}$ farad/cm
Electron mass	$m = 9.1095 \times 10^{-31}$ kg
Coulomb constant	$k_c = 8.988 \times 10^9$ newton-m ² /(coulomb) ²
Atomic mass unit	$u = 1.6606 \times 10^{-27}$ kg

Useful Combinations

Thermal energy (300 K)	$kT = 0.0258$ eV ≈ 1 eV/40
Photon energy	$E = 1.24$ eV at $\lambda = \mu\text{m}$
Coulomb constant	$k_c e^2 = 1.44$ eV · nm
Permittivity (Si)	$\epsilon = \epsilon_r \epsilon_0 = 1.05 \times 10^{-12}$ farad/cm
Permittivity (free space)	$\epsilon_0 = 55.3 \text{e/V} \cdot \mu\text{m}$

Prefixes

k = kilo = 10^3 ; M = mega = 10^6 ; G = giga = 10^9 ; T = tera = 10^{12}
 m = milli = 10^{-3} ; μ = micro = 10^{-6} ; n = nano = 10^{-9} ; p = pica = 10^{-12}

Symbols for Units

Ampere (A), Coulomb (C), Farad (F), Gram (g), Joule (J), Kelvin (K)
Meter (m), Newton (N), Ohm (Ω), Second (s), Siemen (S), Tesla (T)
Volt (V), Watt (W), Weber (Wb)

Conversions

$1 \text{ nm} = 10^{-9} \text{ m} = 10 \text{ \AA} = 10^{-7} \text{ cm}$; $1 \text{ eV} = 1.602 \times 10^{-19}$ Joule = 1.602×10^{-12} erg;
 $1 \text{ eV/particle} = 23.06 \text{ kcal/mol}$; $1 \text{ newton} = 0.102 \text{ kg}_{\text{force}}$;
 $10^6 \text{ newton/m}^2 = 146 \text{ psi} = 10^7 \text{ dyn/cm}^2$; $1 \mu\text{m} = 10^{-4} \text{ cm}$ $0.001 \text{ inch} = 1 \text{ mil} = 25.4 \mu\text{m}$;
 $1 \text{ bar} = 10^6 \text{ dyn/cm}^2 = 10^5 \text{ N/m}^2$; $1 \text{ weber/m}^2 = 10^4 \text{ gauss} = 1 \text{ tesla}$;
 $1 \text{ pascal} = 1 \text{ N/m}^2 = 7.5 \times 10^{-3} \text{ torr}$; $1 \text{ erg} = 10^{-7} \text{ joule} = 1 \text{ dyn-cm}$

Carrier Properties

Reference: Handout 1; Pierret Ch. 1-2

Charge

$$e \quad -1.6 \times 10^{-19} \text{ C}$$

$$h \quad +1.6 \times 10^{-19} \text{ C}$$

Effective Mass m^*

$$F = -eE = m^* \frac{dv}{dt}$$

300 K

s.

$$m_n^*/m_0 = 1.18$$

$$m_p^*/m_0 = 0.81$$

Ge

-55

GaAs

-0.66

+0.52



Intrinsic Properties

$n = p = n_i = 10^{10} / \text{cm}^3$ at RT in Si

$5 \cdot 10^{22} \text{ atoms/cm}^3$

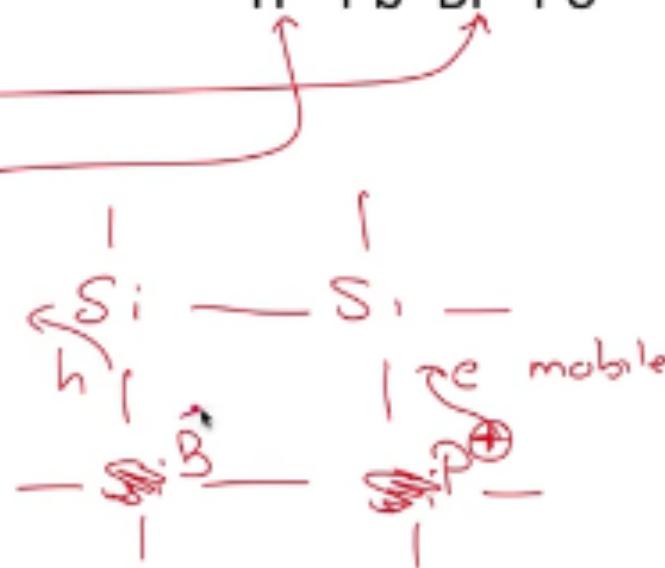
GaAs $2 \cdot 10^6 / \text{cm}^3$

B	C	N	O
Al	Si	P	S
Ga	Ge	As	Se
In	Sn	Sb	Te
Tl	Pb	Bi	Po

Extrinsic properties

Donors \rightarrow donate e

Acceptors \rightarrow donate h



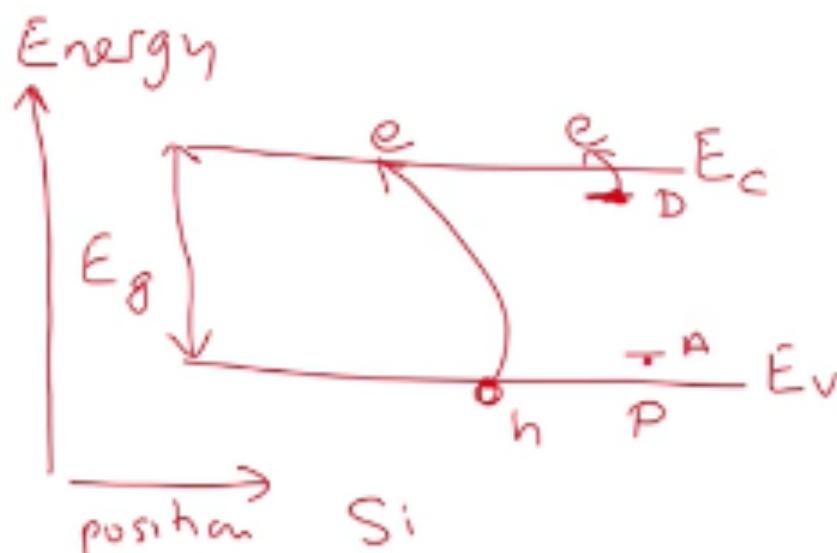
Intrinsic Properties

$$n = p = n_i$$

B	C	N	O
Al	Si	P	S
Ga	Ge	As	Se
In	Sn	Sb	Te
Tl	Pb	Bi	Po

Extrinsic properties

Donors
Acceptors

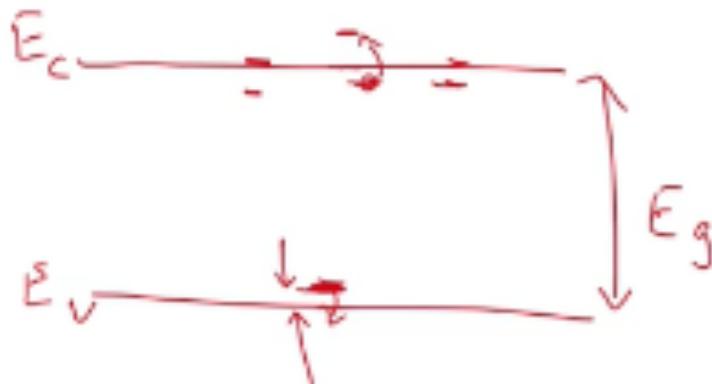


A band diagram...

band gaps: Si 1.12 eV

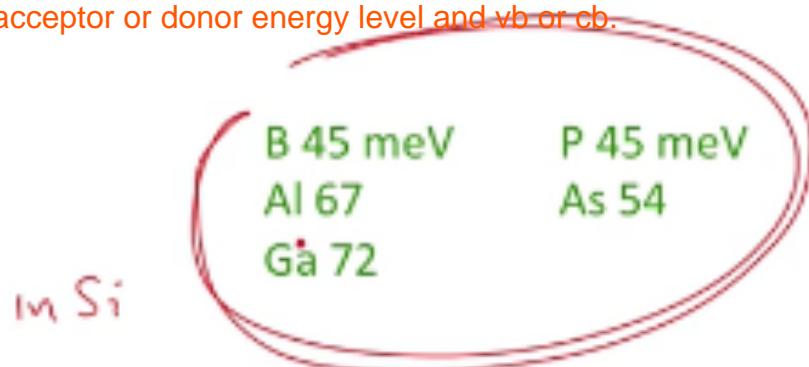
diamond 5.4 eV

silica 5 eV



...with dopants

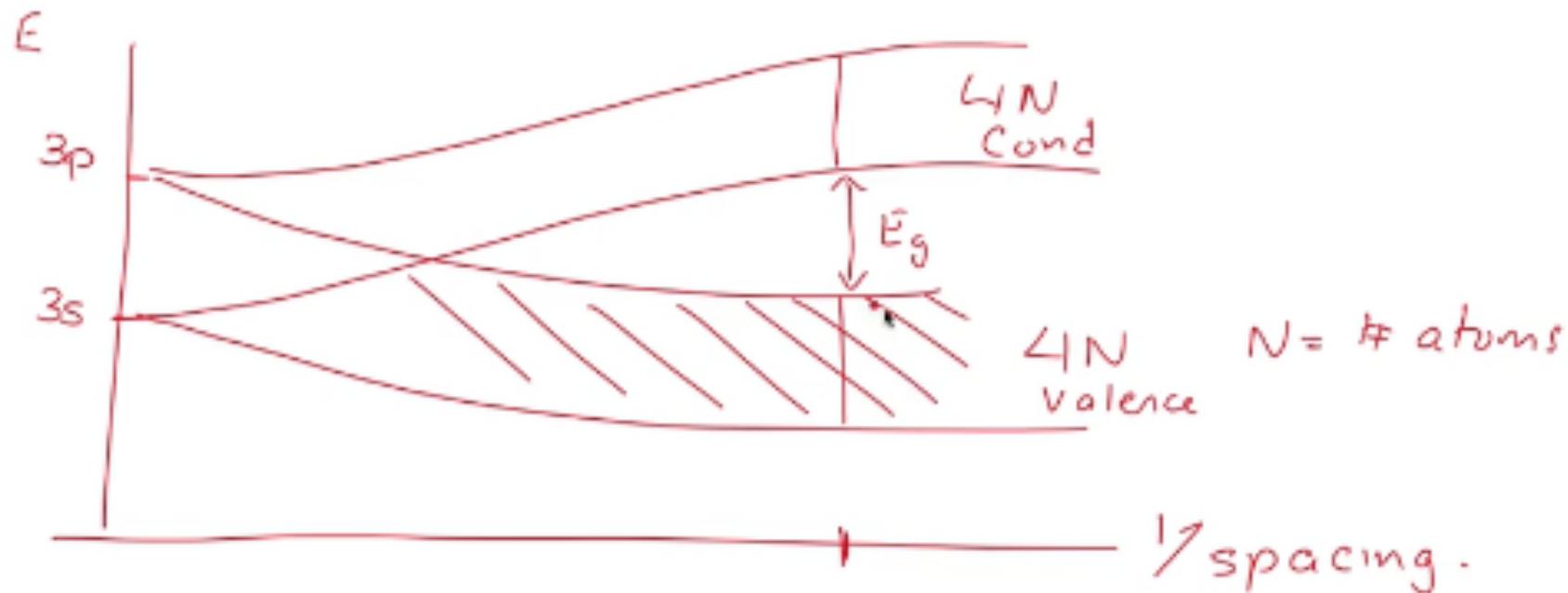
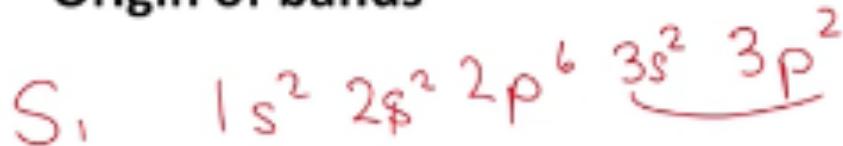
distance bw acceptor or donor energy level and vb or cb.



$$kT \sim 25 \text{ meV}$$



Origin of bands



Carrier Distribution

$g(E) dE$ = density of electron states cm^{-3} in the interval $(E, E+dE)$,
units #/ eV.cm^{-3}

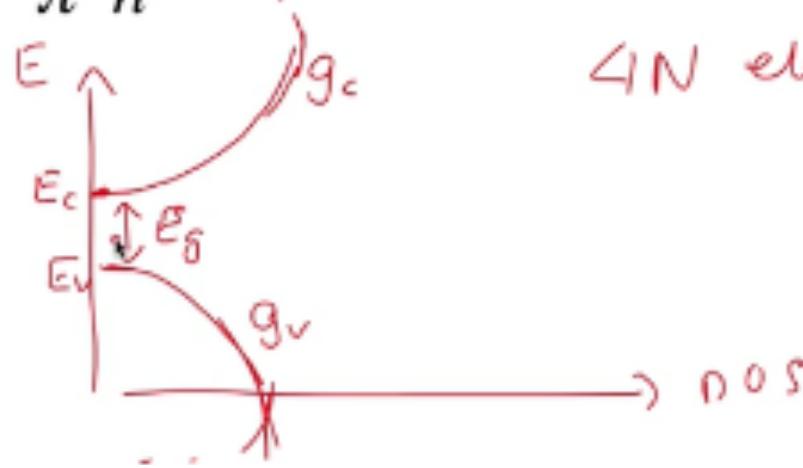
$f(E) dE$ = probability a state of energy E is occupied

$$\left\{ \begin{array}{l} g_c(E)dE = m_n^* \frac{\sqrt{2m_n^*(E - E_c)}}{\pi^2 \hbar^3} dE \\ g_v(E)dE = m_p^* \frac{\sqrt{2m_p^*(E_v - E)}}{\pi^2 \hbar^3} dE \end{array} \right.$$

DOS for
cond Band

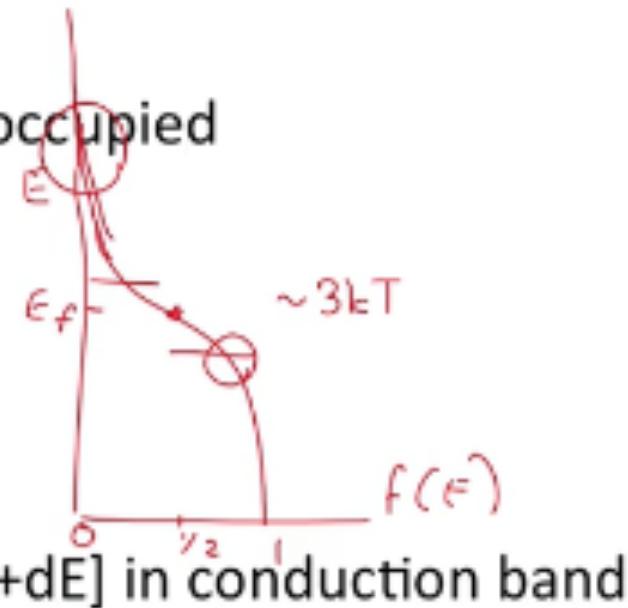
v B.

4N electrons

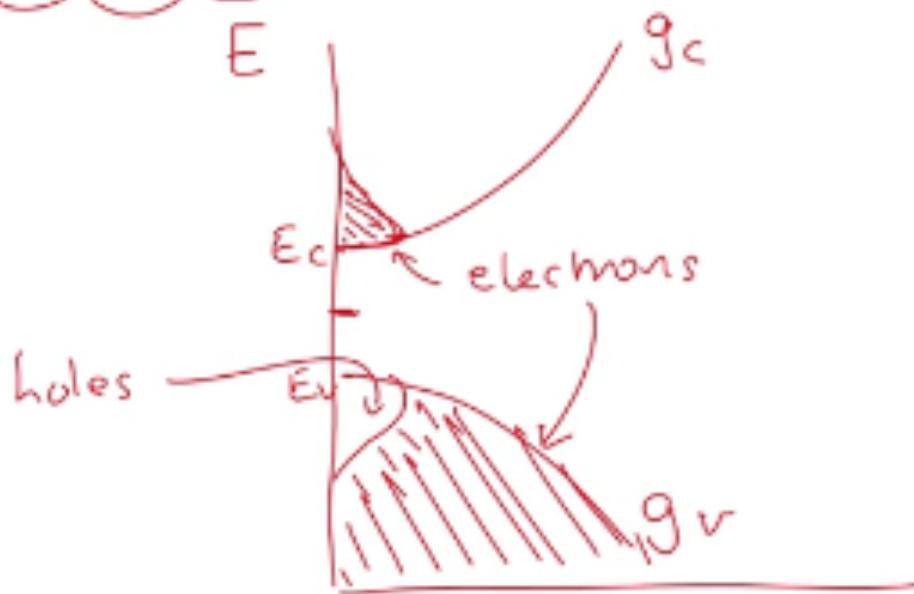


$f(E) dE$ = probability a state of energy E is occupied

$$f(E) = 1 / (1 + \exp(E - E_f) / kT)$$



$f(E) g_c(E) dE$ = number of electrons in $[E, E+dE]$ in conduction band



holes in VB
 $(1 - f(E)) g_v(E) dE$

Intrinsic

$$n = p$$



E_f is midgap
if $m_n^* = m_p^*$



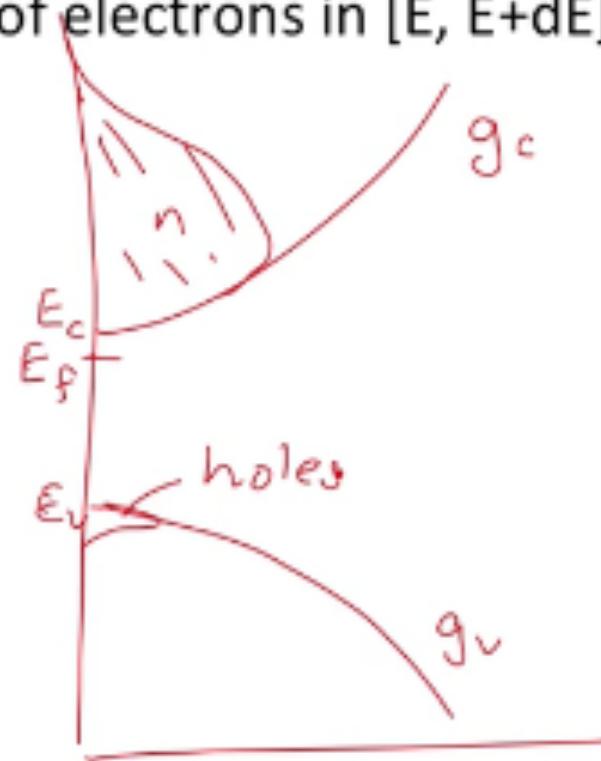
$f(E) dE$ = probability a state of energy E is occupied

$$f(E) = 1 / (1 + \exp(E - E_f) / kT)$$

$f(E)g_c(E) dE$ = number of electrons in $[E, E+dE]$ in conduction band

Doped

$n \gg p$



Total number of electrons and holes

$$n = \int_{E_{bottom}}^{\infty} g_c(E) f(E) dE$$

$$p = \int_{E_{bottom}}^{\infty} g_v(E) (1 - f(E)) dE$$

$$n = n_i \exp(E_f - E_i)/kT$$

intrinsic fermi level
 E_f in intrinsic material

$$n_i = N_c \exp(E_i - E_c)/kT$$

n_i = intrinsic carrier conc
 $\sim 10^{10}/\text{cm}^3$ for Si @ 300K

$$N_c = 2(2\pi m_n^* kT / h^2)^{3/2}$$

Effective DOS in CB

$$n_i = N_v \exp(E_v - E_i)/kT^*$$



Some important relations

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

AT EQM

$$n_i^2 = N_c N_v \exp(E_v - E_c) / kT$$

$$E_i = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln\left(\frac{m_p^*}{m_n^*}\right)$$

mid gap

$$\Rightarrow n_i \propto \exp^{-E_S/2kT}$$

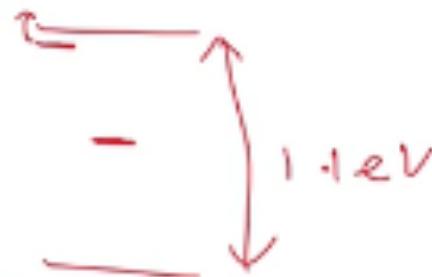
$E_g \uparrow n_i \downarrow$

Si: 73 meV away from mid gap

Doped

$$n \approx N_D \text{ or } p \approx N_A$$

donors/cm³ # acceptors



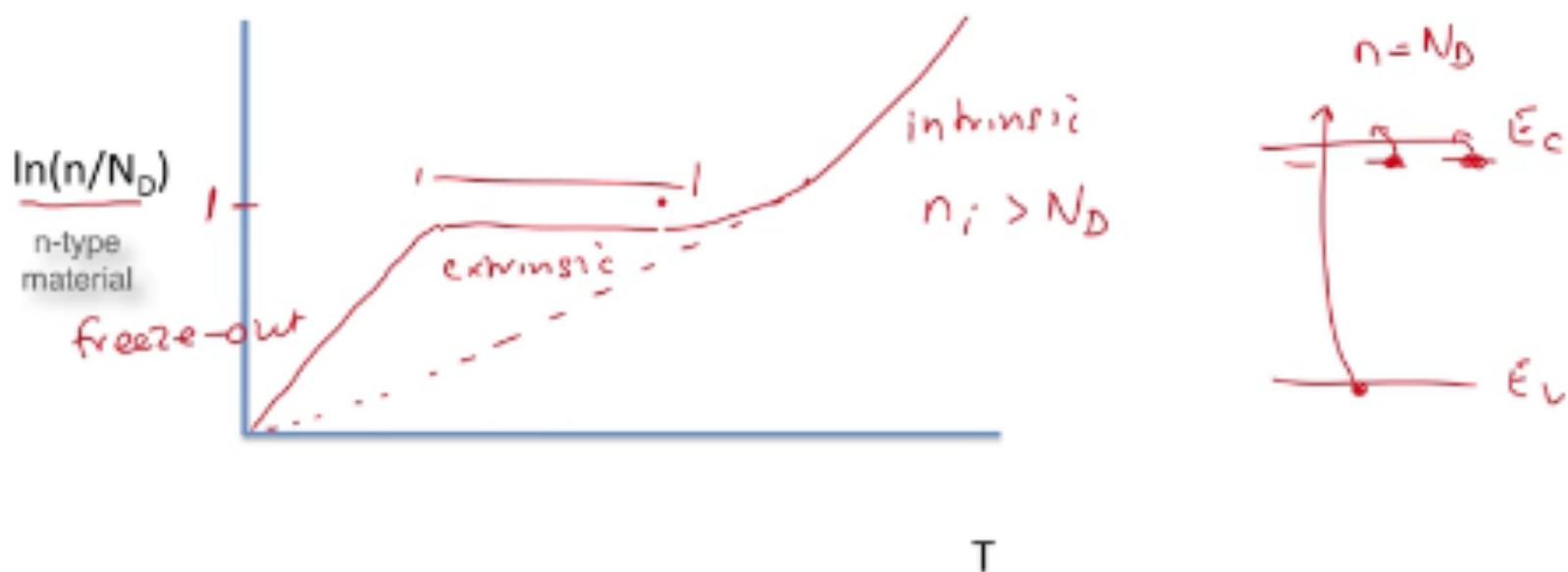
when fully ionized @ RT

$$\Rightarrow p = n_i^2 / N_D \text{ in n-type}$$

$$n = n_i^2 / N_{N_A}$$



Temperature dependence of carrier density



$$\begin{aligned} n &= N_D \\ (E_F - E_i) &= kT \ln\left(\frac{n}{n_i}\right) = -kT \ln\left(\frac{P}{n_i}\right) \\ &\approx kT \ln\left(\frac{N_D}{n_i}\right) \end{aligned}$$

Typical doping $10^{16} - 10^{19} / \text{cm}^3$

$$\begin{aligned} (n_i &= 10^{10} / \text{cm}^3) \\ N \text{ of Si} / \text{cm}^3 &= 10^{22} / \text{cm}^3 \end{aligned}$$

Summary

Carrier distributions

- Density of states $g(E)$ shows how many electrons can be present in each energy range, dE
- Fermi function $f(E)$ tells us the probability they are occupied
- Near the band edge, the density of states increases with $E^{0.5}$
- The product $f(E)g(E)dE$ can be integrated to give total number of electrons (or holes)
- Fermi level E_f is where $f(E) = 0.5$
- At equilibrium, $np = \text{constant} = n_i^2$. Increased temperature raises np . Donor doping raises n and lowers p , and shifts E_f up.
- Carrier concentration, and hence conductivity, varies with T .

Next up: **Carrier Action**

3.15

Carrier Drift, Diffusion and R&G

C.A. Ross, Department of Materials Science and Engineering

Reference: Pierret, chapter 3.

Electron and holes can drift, diffuse, and undergo generation and recombination (R&G).

Drift:

$$\begin{aligned} \text{thermal velocity} & \quad 1/2 m v_{\text{thermal}}^2 = 3/2 kT \\ \text{drift velocity, } v_d & = \mu E \quad (\mu = \text{mobility}, E = \text{field}) \end{aligned}$$

$$\begin{aligned} \text{Current density (electrons)} & \quad J = n e v_d \\ \text{Current density (electrons \& holes)} & \quad J = e (n \mu_n + p \mu_h) E \\ \text{Conductivity} & \quad \sigma = J/E = e (n \mu_n + p \mu_h) \end{aligned}$$

Magnitude of mobility (cm^2/Vs)

	μ_n	μ_h	
Si	1500	450	
Ge	3900	1900	
Ag	50	-	
GaAs	8500	400	

Time between collisions is τ $\mu = e\tau/m^*$

Distance between collisions is l $l = \tau v_{\text{thermal}}$

Diffusion

$$J = e D_n \nabla n + e D_p \nabla p$$

Derivation of the Einstein relation: $D_n/\mu_n = kT/e$
typical D_n in Si is $40 \text{ cm}^2/\text{s}$

Carrier R&G

Mechanisms: band-to-band (direct)
RG centers or traps (indirect)

Thermal R and G at equilibrium: $R = G$
expect $R = G = r n p = r n_i^2$ $r = \text{rate constant}$

Shining light, etc. on the semiconductor causes additional R. These excess carriers n_l and p_l ($n_l = p_l$) decay once the light is turned off.

Illuminated: n-type material

$$n = N_D + n_l \quad \sim N_D$$

$$p = n_i^2/N_D + p_l \quad \sim p_l$$

$$\text{net rate of change of carriers} = R - G = rnp - r n_i^2$$

the rate of recombination of the minority carriers is

$$-dp/dt = r(N_D p - n_i^2) \quad \text{but } n_i^2 = N_D(p - p_l)$$

$$-dp/dt = rN_D(p - p + p_l) = rN_D p_l$$

This has a solution $p_l = p_{l,t=0} \exp(-t/\tau_p)$, where $\tau_p = 1/rN_D$ = minority carrier lifetime.

Example of Carrier Action – Formal solution

A piece of p-type Si is illuminated at one end; how does the carrier concentration vary with depth x ?

$$\begin{aligned} dn/dt &= dn/dt_{\text{drift}} + dn/dt_{\text{diffn}} + dn/dt_{\text{thermal RG}} + dn/dt_{\text{other RG}} \\ &= 0 \text{ at steady state} \end{aligned}$$

$$n = n_p + n_l \quad \text{where } n_p = n_i^2/N_A$$

Inside the material there is only thermal R&G:

$$G_{\text{thermal}} = rn_i^2 = r n_p N_A$$

$$R_{\text{thermal}} = rnp \sim r n_l N_A$$

$$R - G = r N_A(n_l - n_p) \sim r N_A n_l = n_l / \tau_n$$

In the steady state,

$$dn/dt = dn/dt_{\text{diffn}} - (R - G) = 0$$

$$dn/dt = 1/e \nabla J_{\text{diffn}} - (R - G) = 0$$

$$d^2n_l/dx^2 = n_l / \tau_n D_n$$

(since $dn/dt_{\text{diffn}} = 1/e \nabla J_{\text{diffn}} = D_n d^2n/dx^2$ from Fick's law)

$$\text{solution: } n_l = n_{l,x=0} \exp(-x/\sqrt{\tau_n D_n})$$

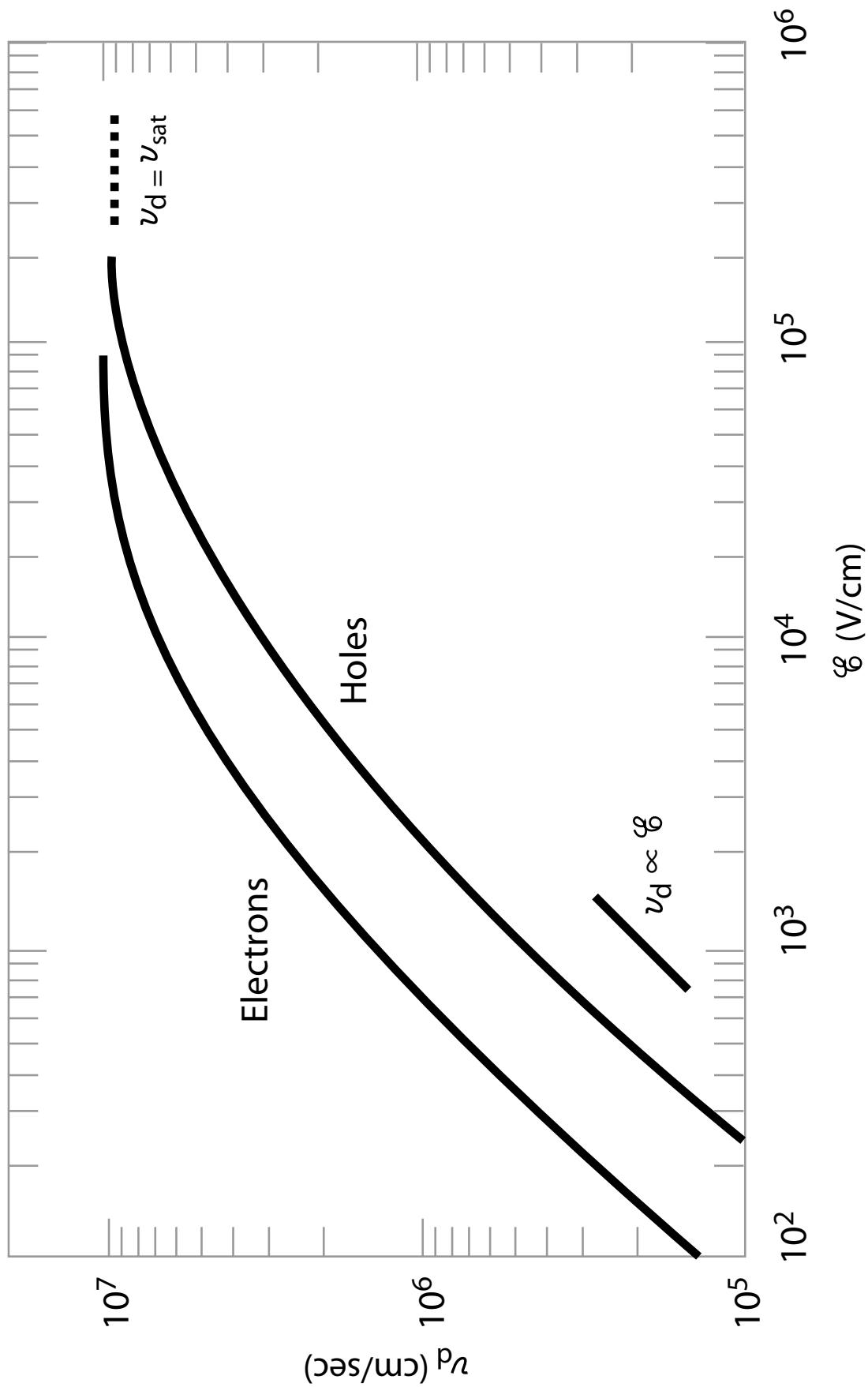
Generally, if excess carrier concentrations are written as n_l or p_l , then

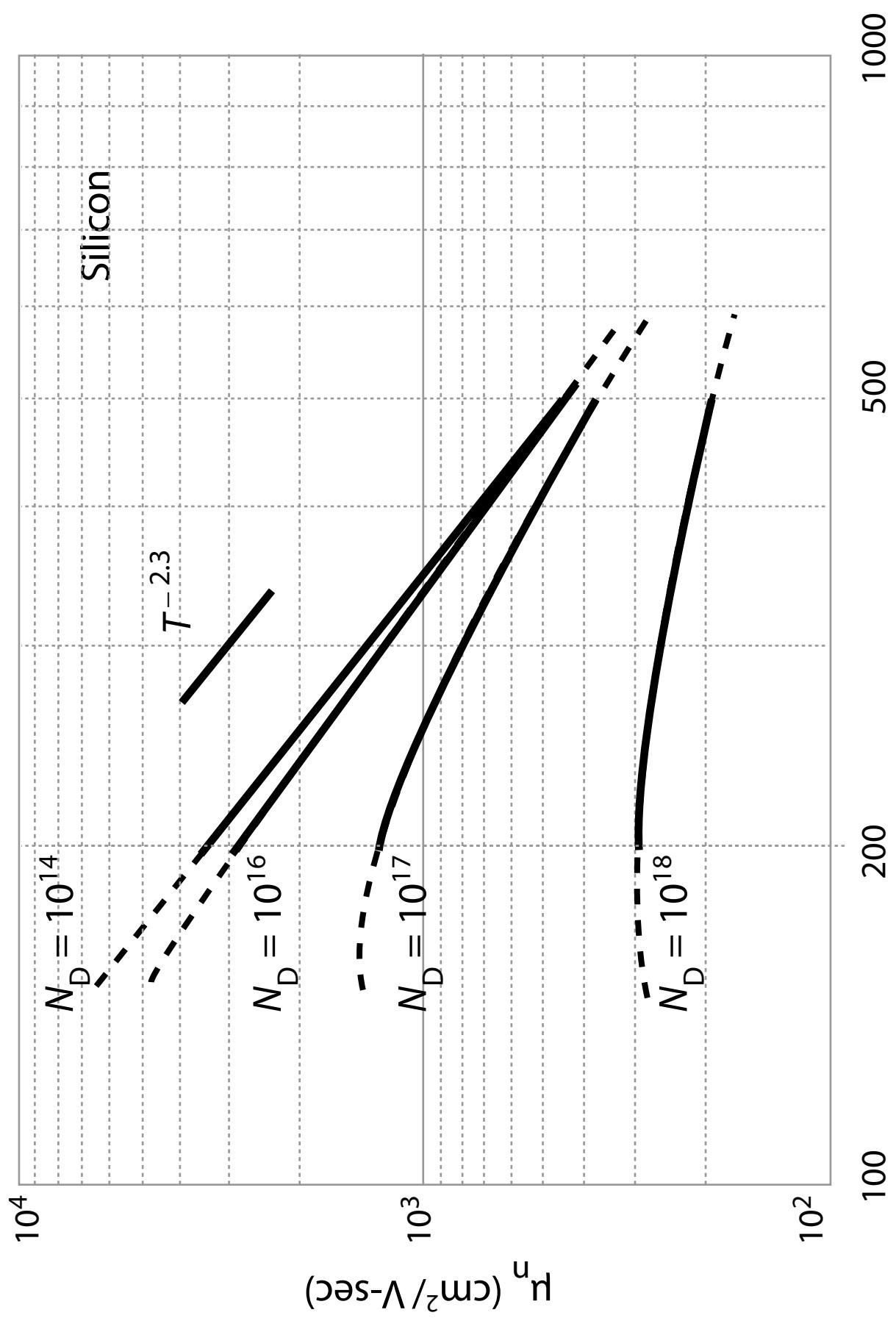
$$dn/dt_{\text{thermal}} = -n_l / \tau_n \quad \text{or} \quad dp/dt_{\text{thermal}} = -p_l / \tau_p$$

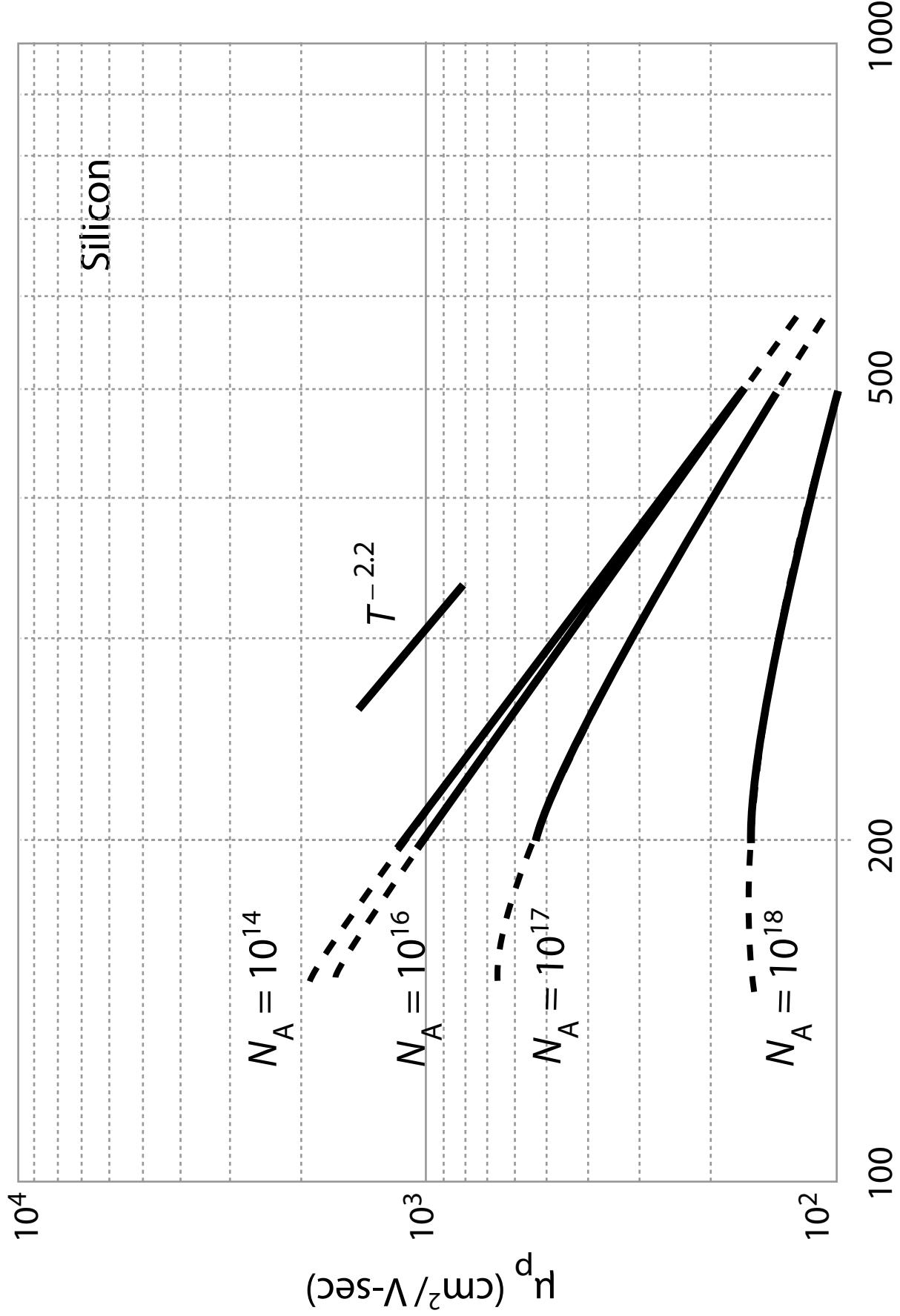
Minority carrier lifetimes are $\tau_n = 1/rN_A$, or $\tau_p = 1/rN_D$,

Minority carrier diffusion lengths are $\lambda_n = \sqrt{\tau_n D_n}$, or $\lambda_p = \sqrt{\tau_p D_p}$.

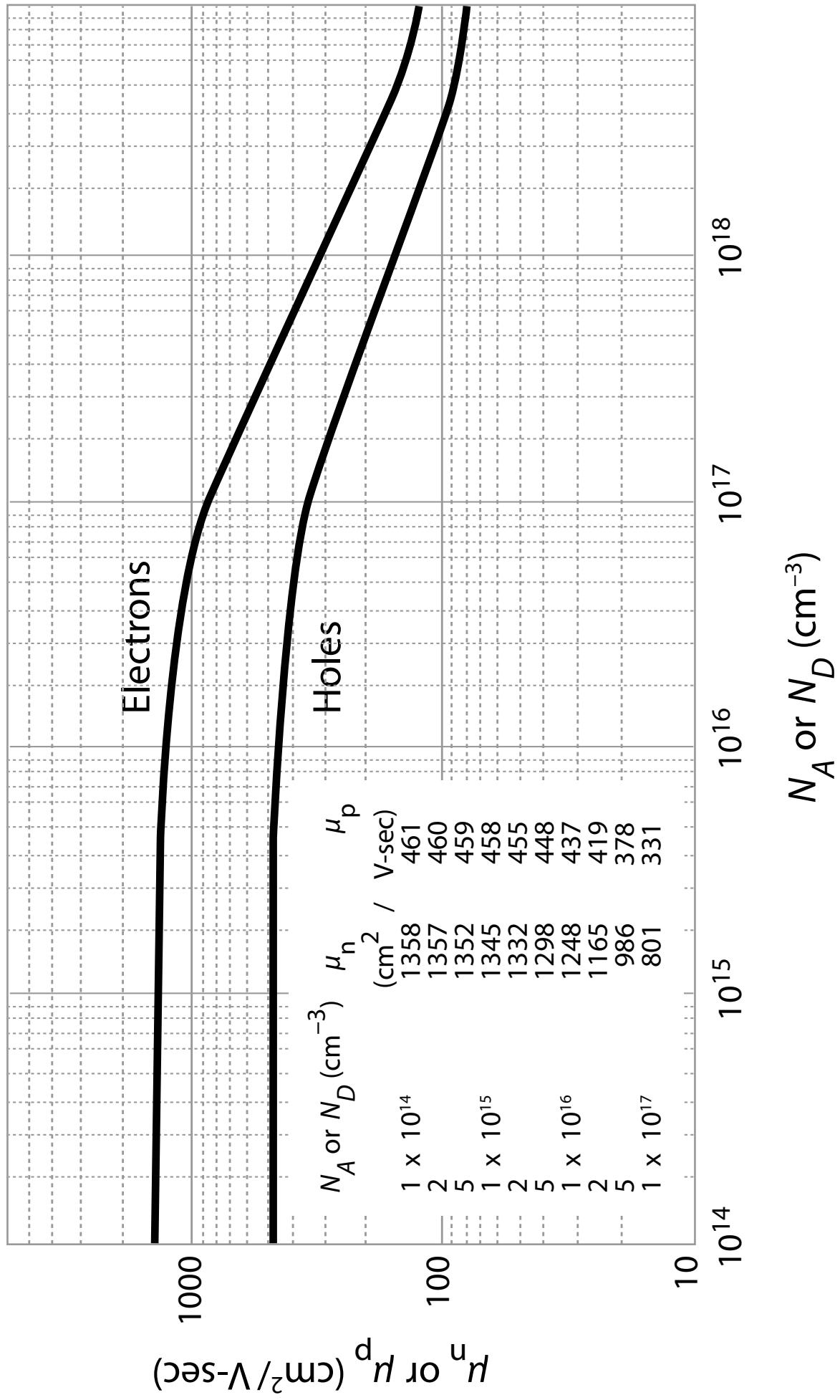
If traps dominate the recombination, then $\tau = 1/r_2 N_T$ where $r_2 \gg r$



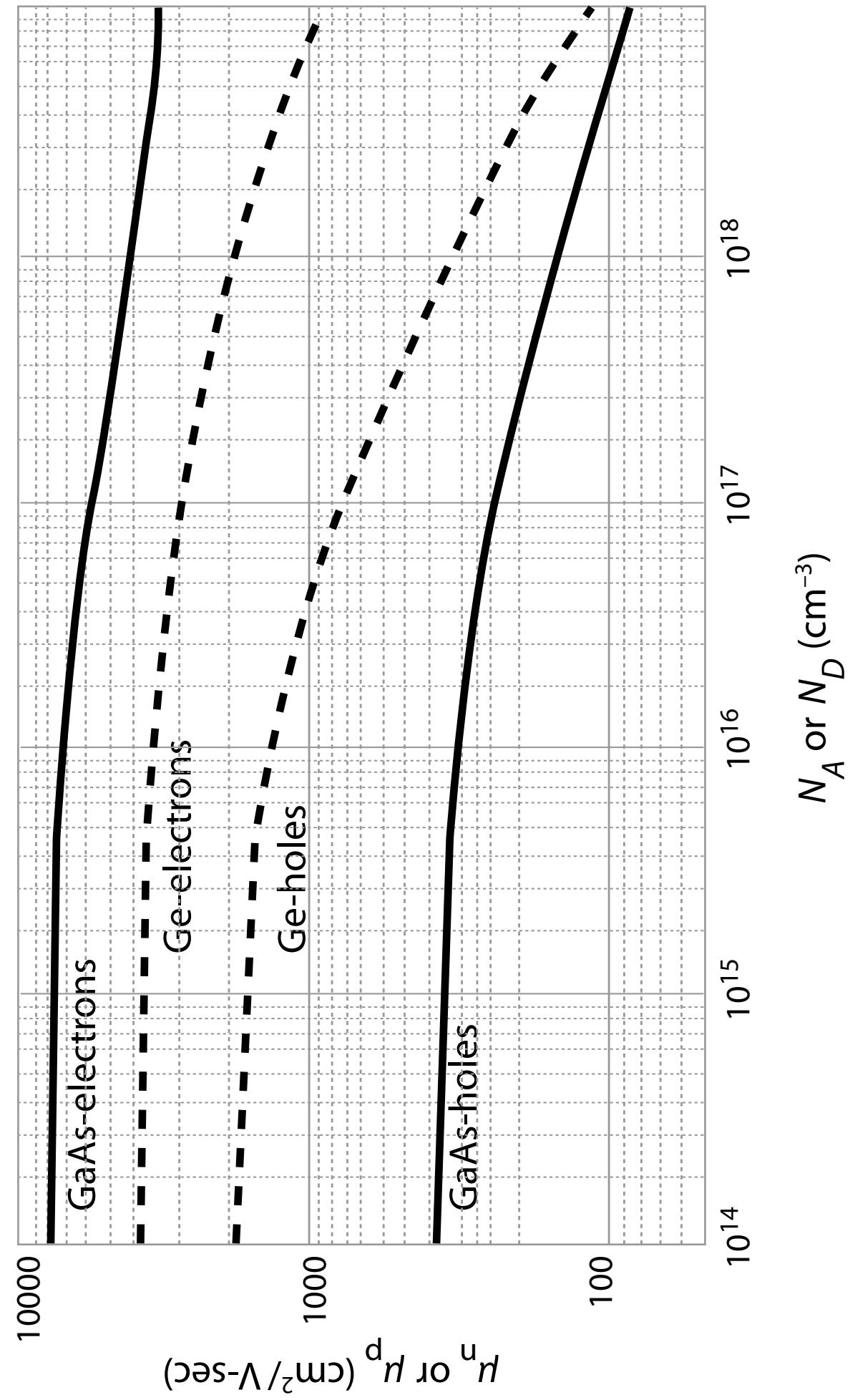




Silicon
 $T = 300\text{ K}$



$T = 300 \text{ K}$



USEFUL CARRIER MODELING EQUATIONS

Density of States and Fermi Function:

$$g_c(E) = \frac{m_n^* \sqrt{2m_n^*(E - E_c)}}{\pi^2 \hbar^3}, E \geq E_c$$

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

$$g_v(E) = \frac{m_p^* \sqrt{2m_p^*(E_v - E)}}{\pi^2 \hbar^3}, E \leq E_v$$

Carrier Concentration Relationships:

$$n = N_c e^{(E_F - E_C)/kT}$$

$$N_c = 2 \left[\frac{m_n^* kT}{2\pi \hbar^2} \right]^{3/2}$$

$$p = N_v e^{(E_v - E_F)/kT}$$

$$N_v = 2 \left[\frac{m_p^* kT}{2\pi \hbar^2} \right]^{3/2}$$

n, np-Product, and Charge Neutrality:

$$n_i = \sqrt{N_c N_v} e^{-E_G/2kT} \quad np = n_i^2 \quad p - n + N_D - N_A = 0$$

n, p, and Fermi Level Relationships:

$$n = \frac{N_D - N_A}{2} + \left[\left(\frac{N_D - N_A}{2} \right)^2 + n_i^2 \right]^{1/2}$$

$$E_i = \frac{E_C + E_v}{2} + \frac{3}{4} kT \ln \left(\frac{m_p^*}{m_n^*} \right)$$

$$n \approx N_D \quad N_D \gg N_A, N_D \gg n_i \quad E_F - E_i = kT \ln(n / n_i) = -kT \ln(p / n_i)$$

$$p \approx n_i^2 / N_D$$

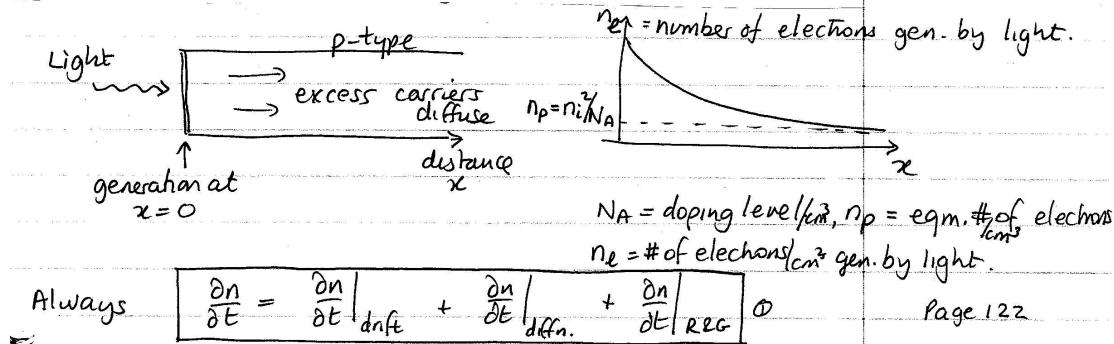
$$p \approx N_A \quad N_A \gg N_D, N_A \gg n_i \quad E_F - E_i = kT \ln(N_D / n_i) \quad N_D \gg N_A, N_D \gg n_i$$

$$n \approx n_i^2 / N_A \quad E_i - E_F = kT \ln(N_A / n_i)$$

Recombination and Generation Example

- At eqm, $R = rnp = rn_i^2$ for direct recombination. Therefore $G = R = rn_i^2$.
 - Away from eqm, $R = rnp$, but n and p are non-eqm. concentrations.
- For G , if the light is shining then $G \propto$ light intensity.
When the light is turned off, or inside the material, $G = n_i^2$

Example



At steady state $\frac{\partial n}{\partial t} = 0$. In this case, all the G occurs at the surface. Inside the material, $G = G_{\text{Thermal}} = rn_i^2 = rn_p N_A$
 and $R = rnp \approx rn_e N_A$ (since $n = n_p + n_e \approx n_e$)

$$\text{hence } (R - G) = r(n_e N_A - n_p N_A)$$

$$\approx rn_e N_A = n_e / \tau_n \quad \text{where } \tau_n = 1/rN_A$$

Also, $E = 0$ inside material.

$$\text{Hence } \frac{\partial n}{\partial t} = 0 = \left. \frac{\partial n}{\partial t} \right|_{\text{diffn.}} + -n_e / \tau_n$$

$$\text{recall that } J_{\text{diffusion}} = e D_n \frac{dn}{dx} \quad \text{A/cm}^2$$

$$\text{and } \left. \frac{\partial n}{\partial t} \right|_{\text{diffn.}} = \frac{1}{e} \nabla J_{\text{diff}} = D_n \frac{d^2 n}{dx^2}$$

$$\text{So ① becomes } 0 = D_n \frac{d^2 n}{dx^2} - n_e / \tau_n \quad (\text{with } n = n_p + n_e)$$

$$\text{which has a solution } n_e = n_e(x=0) \exp^{-\frac{x}{\sqrt{D_n \tau_n}}} \quad ②$$

General expressions for minority carriers

$$\text{Minority carrier lifetime } \tau_p = 1/rN_b \quad \text{in n-type}$$

$$\tau_n = 1/rN_A \quad \text{in p-type}$$

$n_p, -p_e$ are excess carrier

concentrations (often written $\Delta n, \Delta p$)

$$\left. \frac{\partial n}{\partial t} \right|_{\text{thermal}} = -\frac{n_e}{\tau_n}$$

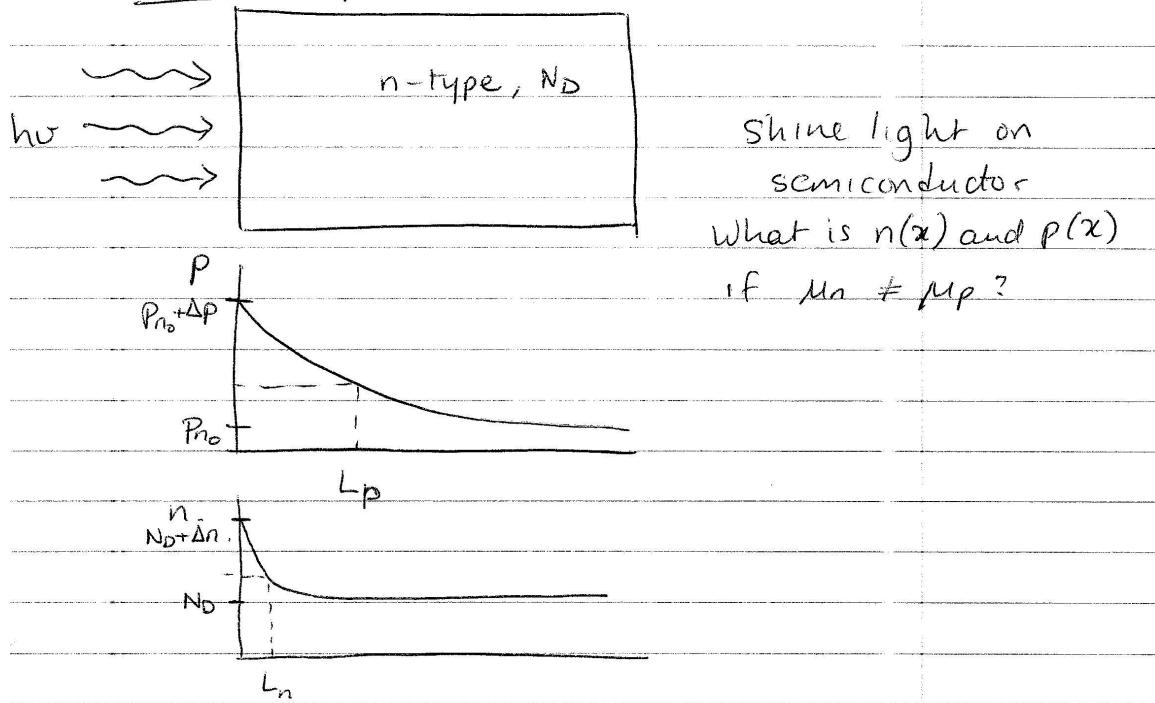
$$\left. \frac{\partial p}{\partial t} \right|_{\text{thermal}} = -\frac{p_e}{\tau_p}$$

Diffusion lengths (before recombination occurs) are $\sqrt{D_n \tau_n}$ or $\sqrt{D_p \tau_p}$.

For recombination dominated by traps, use $\tau = 1/r_2 N_T$, where

$r_2 \gg r$. N_T is the trap density. (Indirect recombination). Page 112.

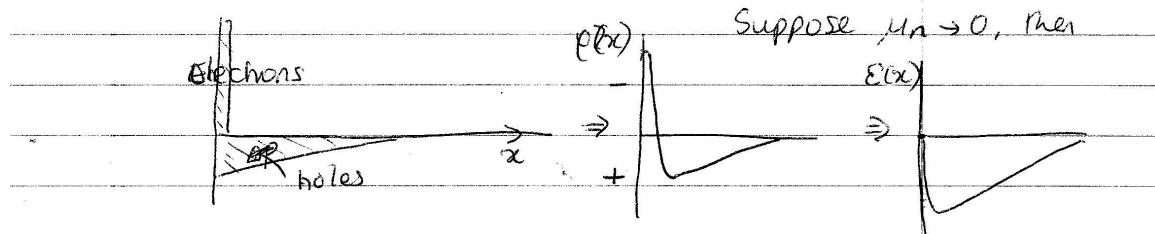
R & G example with different mobilities



$$\text{At } x=0 \quad \Delta n = \Delta p$$

Suppose $\mu_p \gg \mu_n$. $L_p \gg L_n$.

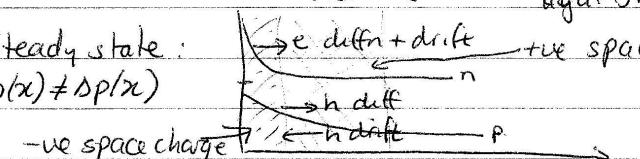
Expect holes to diffuse further before they recombine giving net -ve charge at small x and +ve at large x



Electric field pulls electrons \rightarrow same direction as diffn

holes \leftarrow against diffusion

Steady state: $\frac{d\ln(n)}{dx} \neq \frac{d\ln(p)}{dx}$



Carrier Action

Reference: Handout 2; Pierret Ch. 3

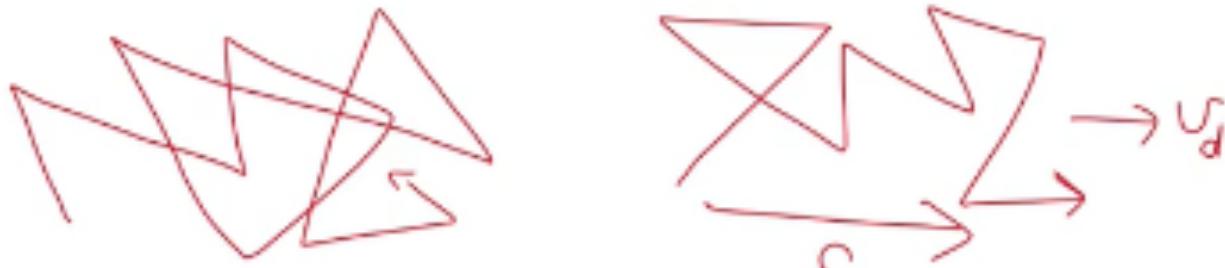
Carriers can:

- Drift in E
- Diffuse in ∇n or ∇p
- Undergo R&G (recombination and generation)

.



Drift



thermal velocity

$$\underline{1/2 m v_{\text{thermal}}^2 = 3/2 kT}$$

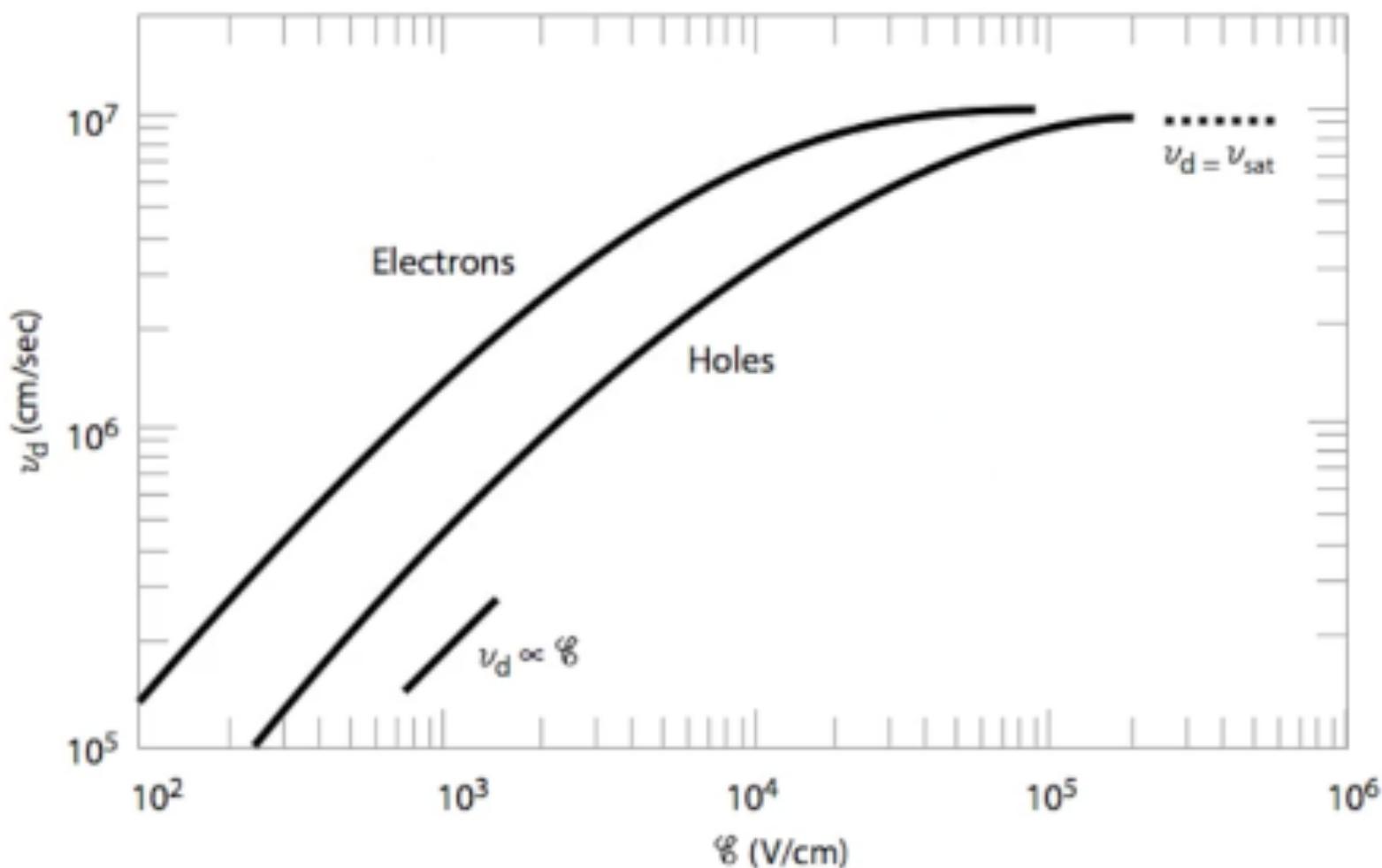
drift velocity

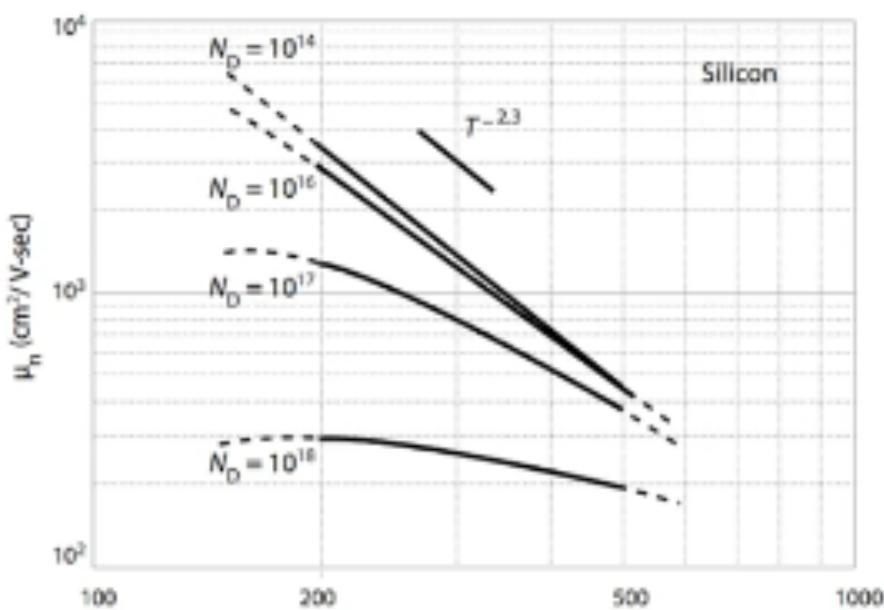
$$v_d = \mu E \quad (\mu = \text{mobility}, E = \text{field})$$

$\text{cm}^2/\text{Vs.}$



Data for Silicon

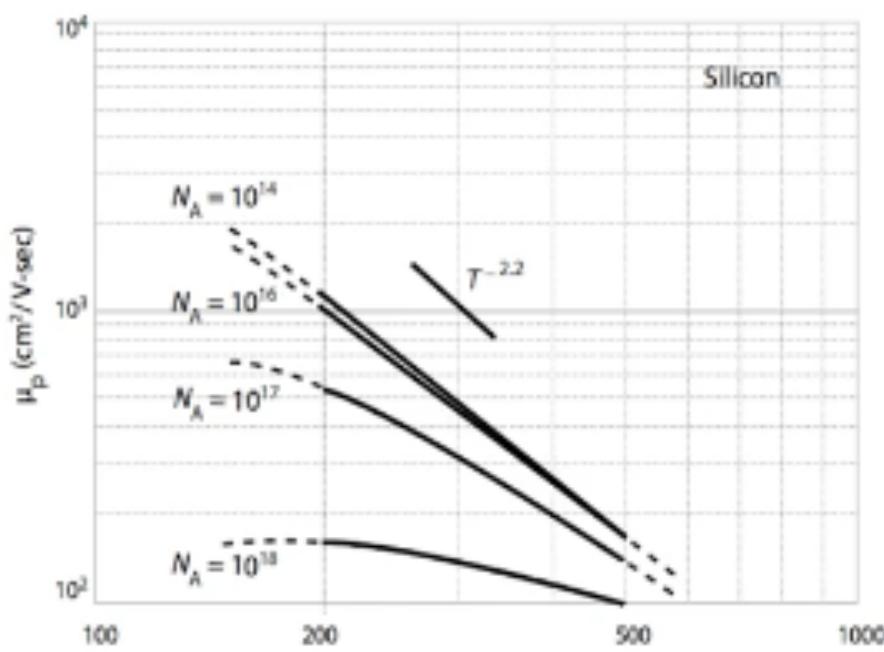


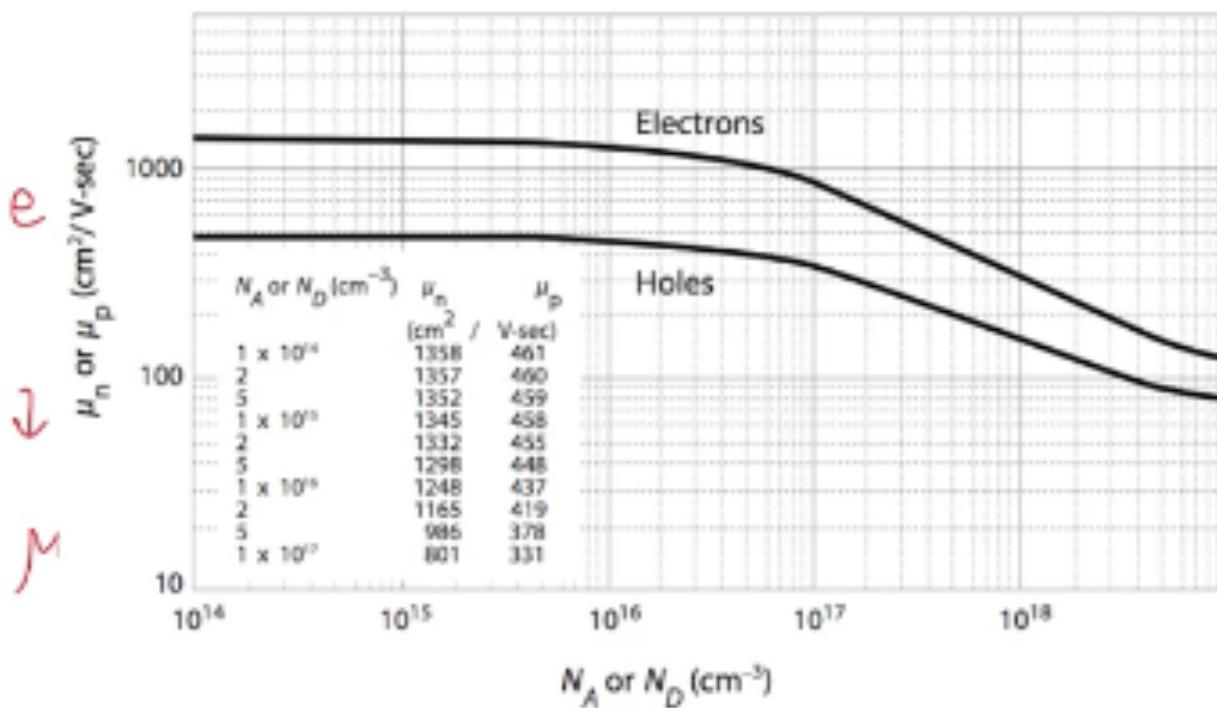


$T \uparrow \mu_e \downarrow$

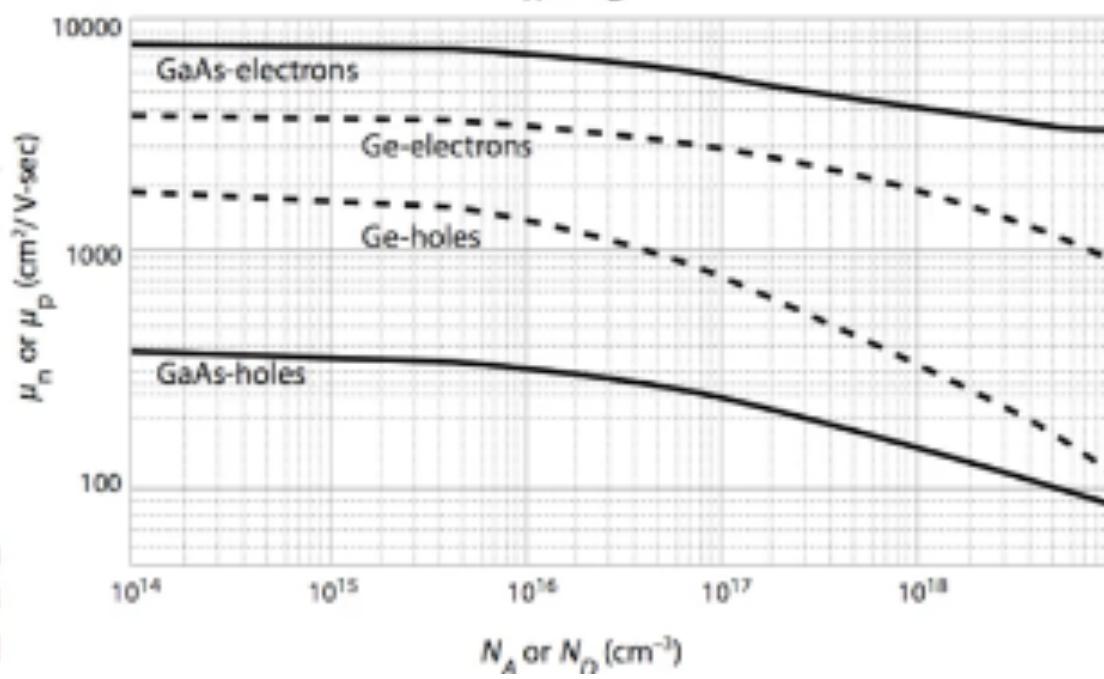
$N_D, N_A \uparrow \mu_h \downarrow$

h





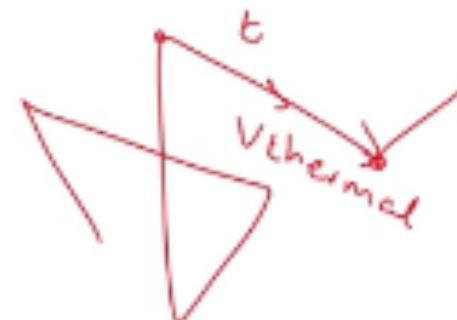
$T \uparrow \mu \downarrow$
 $N_D, N_A \uparrow \mu$



Mobility depends on temperature and doping

Time between collisions is t $\mu = et/m^*$

Distance between collisions is λ $\lambda = tv_{\text{thermal}}$



Conductivity

Current density (electrons)

$$J = n e v_d$$

just electrons

Current density (electrons & holes) $J = e (n \mu_n + p \mu_p) E$

Conductivity

$$\sigma = J/E = e (n \mu_n + p \mu_p)$$

n type $n \gg p$

$$\sigma \sim e n \mu_n$$

Magnitude of mobility (cm^2/Vs)

	μ_n	μ_p
Si	1500	450
Ge	3900	1900
Ag	50	-
GaAs	8500	400

1 e per atom
 $\sim 10^{22}$ carriers/ cm^3

N_D or N_A
 $\sim 10^{20}$

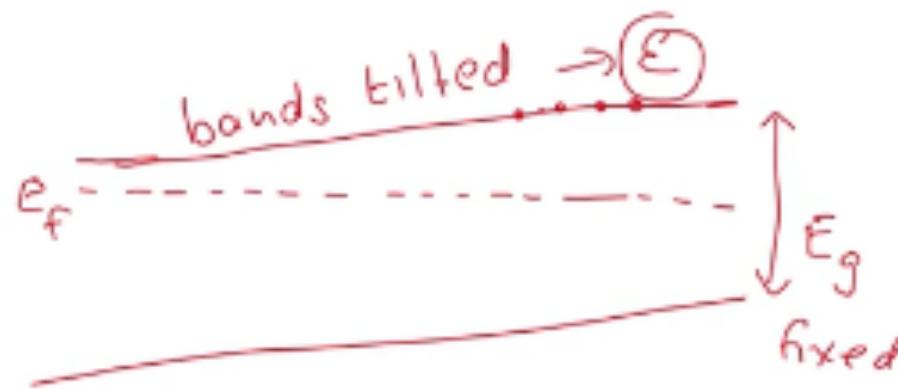
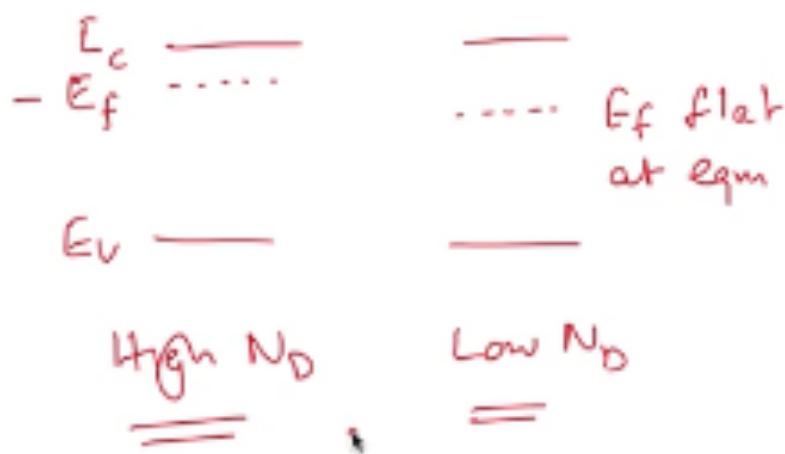
InSb 80,000 750 $E_g \sim 0.2 \text{ eV}$



Diffusion

current density $J = eD_n \nabla n + eD_p \nabla p$

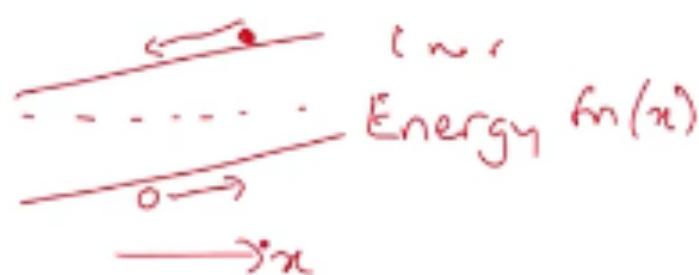
How do diffusivity and mobility relate? Nonuniformly doped semiconductor:



At equilibrium, what do we know about the Fermi level? The bands? The electric field?



At equilibrium



$$n = n_i \exp(E_f - E_i) / kT$$

$$dE_f / dx = 0 \quad \text{so} \quad \frac{dn}{dx} = \left\{ \frac{-n_i}{kT} \exp(E_f - E_i) / kT \right\} \frac{dE_i}{dx}$$

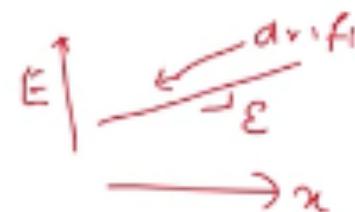
Now put $E = 1/e dE_i/dx$, then

$$\frac{dn}{dx} = \left\{ \frac{-n_i}{kT} \exp(E_f - E_i) / kT \right\} \frac{dE_i}{dx} = \frac{-n}{kT} eE$$



At equilibrium

$$\Sigma = \frac{1}{e} \frac{dE}{dx} \leftarrow \text{energy}$$



$$J_{\text{drift}} + J_{\text{d.f.}} = 0$$

$$e\mu_n(\bar{E}) + eD_n \left(\frac{dn}{dx} \right) = 0$$

High \bar{D}_n Low N_D

$$n = n_i \exp(E_f - E_i) / kT$$

$$\frac{dE_f}{dx} = 0$$



$$\underline{\frac{dE_f}{dx} = 0} \quad \text{so} \quad \frac{dn}{dx} = \left\{ \frac{-n_i}{kT} \exp(E_f - E_i) / kT \right\} \frac{dE_i}{dx}$$

Now put $E = 1/e dE_i/dx$, then

$$\frac{dn}{dx} = \left\{ \frac{-n_i}{kT} \exp(E_f - E_i) / kT \right\} \frac{dE_i}{dx} = \left(\frac{-n}{kT} eE \right) \text{electric field}$$



Substitute into $e\mu_n nE + eD_n \frac{dn}{dx} = 0$

$$e\mu_n nE + eD_n \left(-\frac{ne\epsilon}{kT}\right) = 0$$

$$\mu_n + -\frac{eD_n}{kT} = 0$$

Einstein relation: $D_n/\mu_n = kT/e$

typical D_n in Si is 40 cm²/s $\mu_n = 1500 \text{ cm}^2/\text{Vs}$
Large

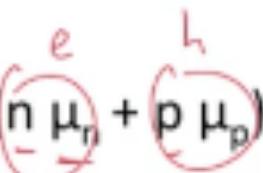
e.g. Na⁺ in glass $D \sim 10^{-16} \text{ cm}^2/\text{s}$



Summary

Drift: response of carriers to electric field, $v_d = \mu E$

- Mobility μ
- Conductivity $\sigma = J/E = e(n\mu_n + p\mu_p)$



Diffusion: response of carriers to concentration gradient

- Current density $J = eD_n \nabla n + eD_p \nabla p$

Mobility and diffusivity differ for electrons and holes

Mobility and diffusivity are related by the Einstein equation

- $D/\mu = kT/e$

Next we will consider R&G: recombination and generation

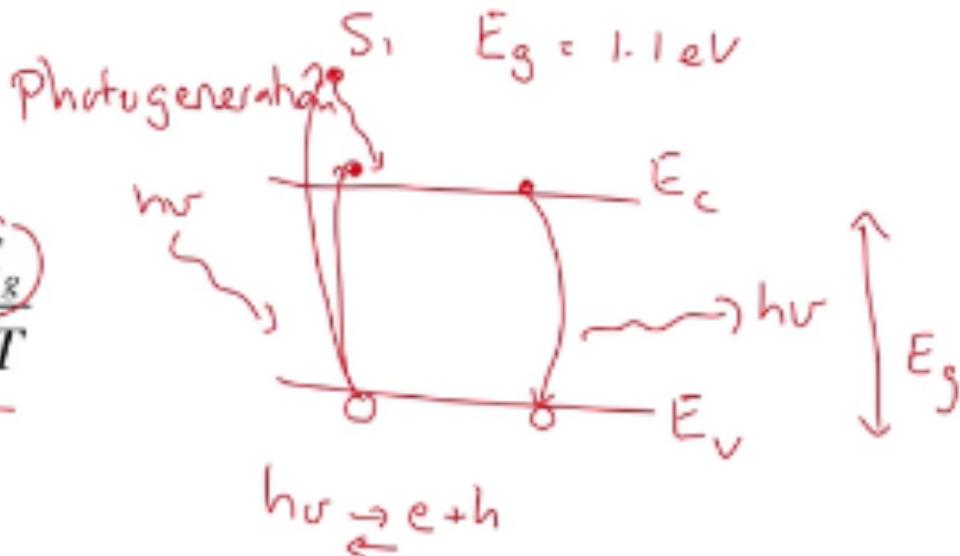
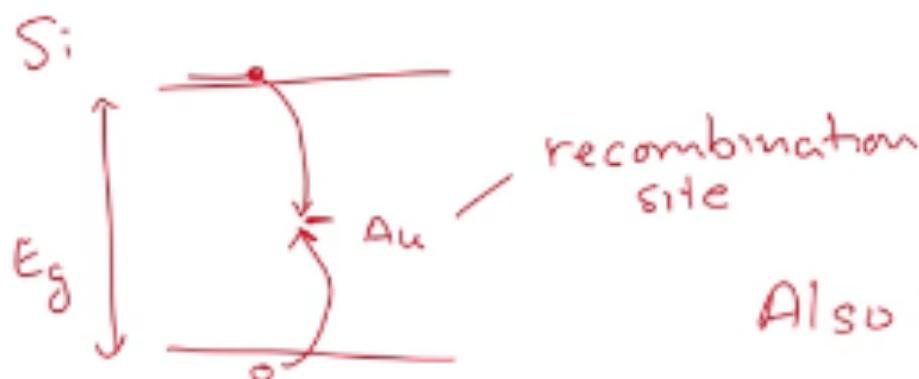


R & G

1) Direct R&G

$$\text{(recall } np = n_i^2 = \sqrt{N_c N_v} \exp \frac{-E_g}{2kT} \text{)}$$

2) via RG centers or traps



Thermal generation

$$kT$$

Also defects



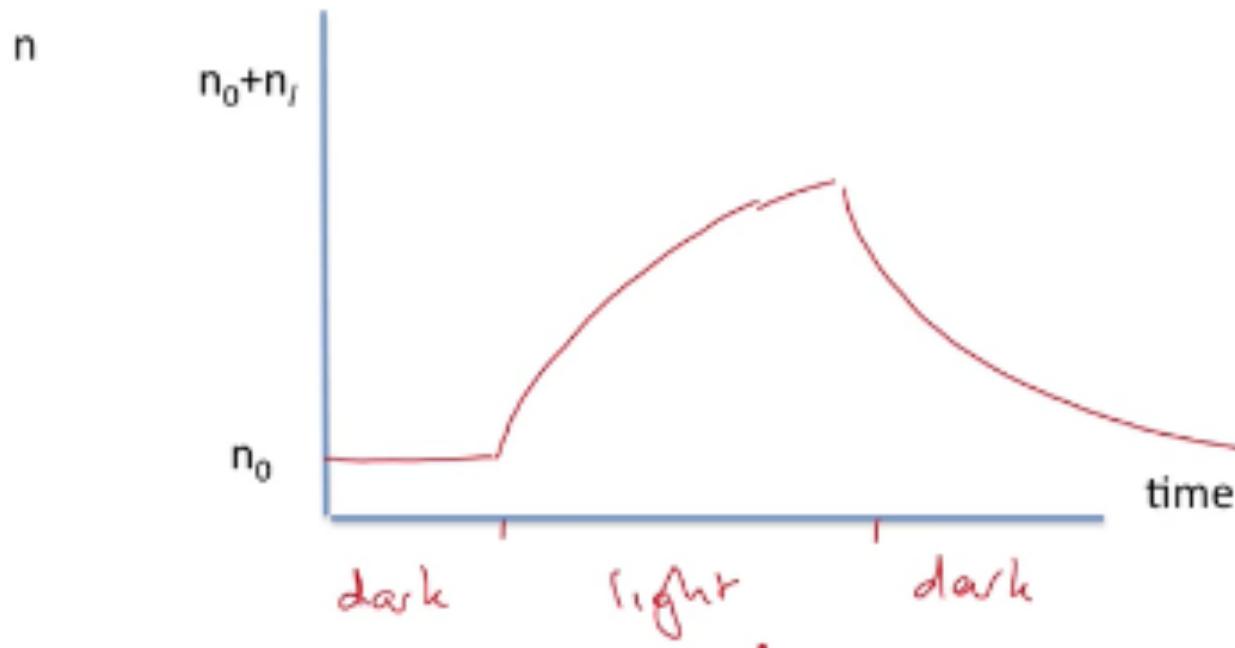
Recombination: $R \sim np = r np$ *rate const*

At equilibrium $R = G = r np = r n_i^2$ *recomb or gen rate at eqm*

Light causes additional G so that $n = n_0 + n_i$ *dynamical process of thermal R & G*
at eqm photogenerated

When light is turned off, they decay and $n \rightarrow n_0$
by R





1) Solution for time dependence of carrier concentration (n-type)

n-type $n \gg p$ biggest change

Δn is measurable

$$n = \left(n_0 \right)_\text{eqm} + \left(n_\text{e} \right)_\text{eqm}$$
$$p = \left(p_0 \right)_\text{eqm} + \left(p_\text{e} \right)_\text{photogenerated}$$

p_n = eqm number of holes in n-type = p_0

$$-\frac{dp}{dt} = R - G = r(np - n^2)$$
$$= r(N_D p - p_n N_D)$$

$$\tau_p = 1 / rN_D$$

1) Solution for time dependence of carrier concentration (n-type)

n-type $n \gg p$ biggest change

Δn is immeasurable

$$n = \underbrace{(n_0)}_{\text{eqn carriers}} + \underbrace{(n_e)}_{\text{photogenerated}}$$

$$\begin{aligned}-\frac{dp}{dt} &= R - G = r(np - n_i^2) \\&= r(N_D p - p_n N_D) \\&= r N_D(p - p_n)\end{aligned}$$

p_n = eq. # holes

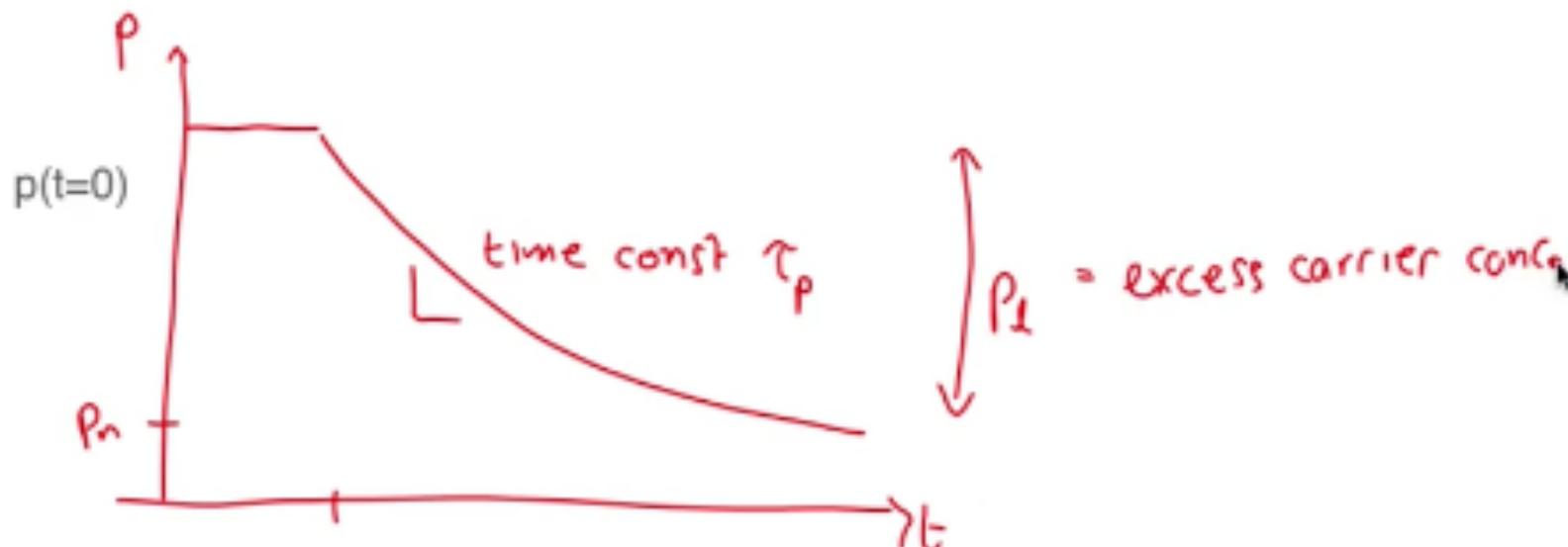
τ_p = minority carrier lifetime

$$\tau_p = 1/rN_D$$

$$\therefore (p - p_n) = (p - p_n)_{t=0} \exp(-t/\tau_p)$$



1) Solution for time dependence of carrier concentration (n-type)



$$\tau_p = 1 / rN_D$$

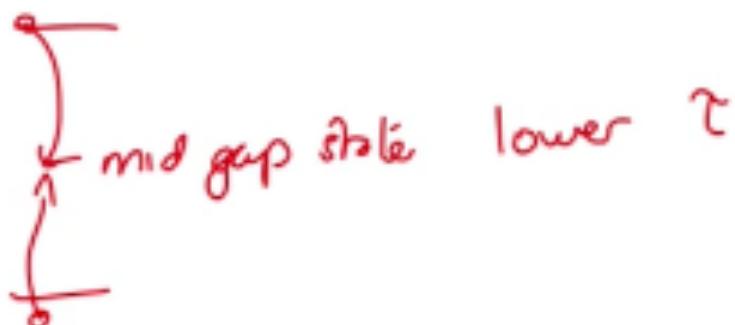


General solution $\frac{dp}{dt} = \Delta p / \tau_p$

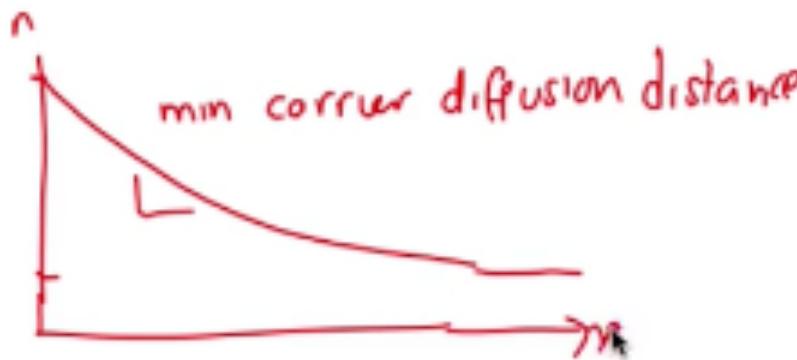
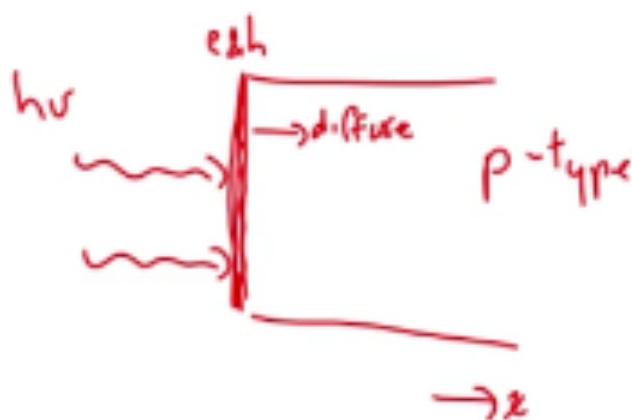
$$\tau \sim 1 \text{ ms} \quad \text{good quality Si}$$

$$D \sim 40 \text{ cm}^2/\text{s}$$

$$\rightarrow x = \sqrt{Dt} \sim \underline{60 \mu\text{m}}$$



2) Solution for position dependence of carrier concentration at steady state (p-type)



$$0 = \frac{\partial n}{\partial t} = \cancel{\left. \frac{\partial n}{\partial t} \right|_{drift}} + \left. \frac{\partial n}{\partial t} \right|_{diffusion} + \cancel{\left. \frac{\partial n}{\partial t} \right|_{thermalRG}} + \cancel{\left. \frac{\partial n}{\partial t} \right|_{otherRG}}$$

at steady state

no E ✓ ✓ =

*light at
surface only*

Where is R, G, drift and diffusion occurring?



$$n = n_p + n_i \text{ where } n_p = n_i^2 / N_A$$

Inside the material, $G_{\text{thermal}} = r n_i^2 = \underline{r n_p N_A}$

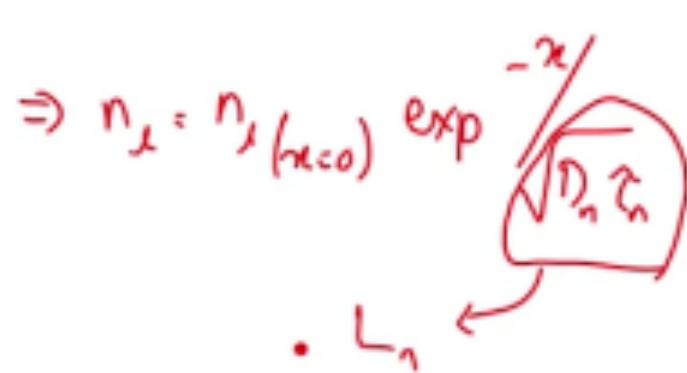
$$\text{and } R = r n_p \approx r n_i N_A = r (n_p + n_i) N_A$$

$$\Rightarrow R - G = r n_i N_A = n_i / \tau_n \quad \tau = 1 / r N_A = \text{min carrier lifetime}$$

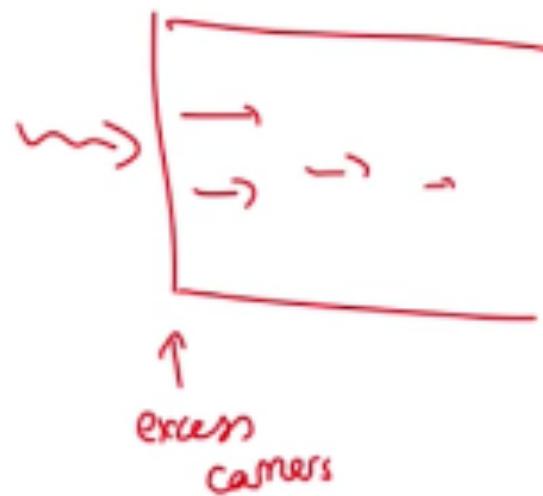
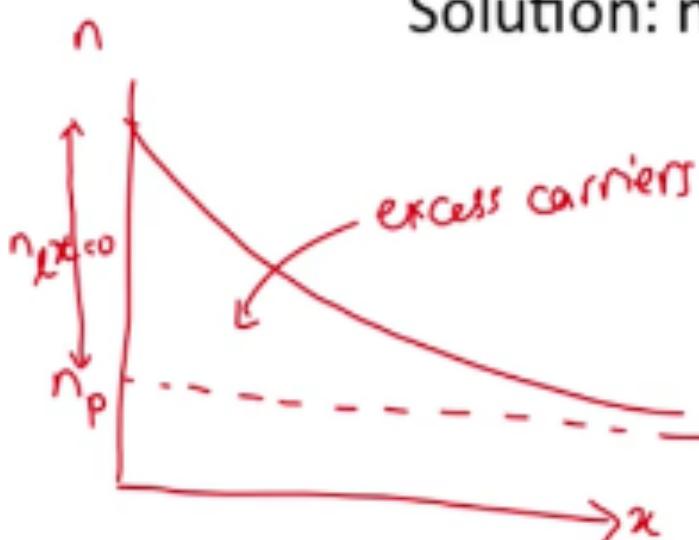
Steady state $\frac{dn}{dt} = \frac{1}{e} \nabla J_{dR} - (R - G) = 0$
 $= \frac{1}{e} \frac{d}{dx} \left(e D_n \frac{dn}{dx} \right) - \frac{n_i}{\tau_n} = 0$

Since $n = n_p + n_i$, $\frac{dn}{dx} = \frac{dn_i}{dx}$
 $\underset{\text{const}}{=}$

$$\Rightarrow \frac{d^2 n_i}{dx^2} = \frac{n_i}{\tau_n D_n}$$



$$\text{Solution: } n_I = n_{I,x=0} \exp(-x/\sqrt{\tau_n D_n})$$



$$L_n = \sqrt{\tau_n D_n}$$

So far, $n \ll p$

Undoped? Many defects?

$$R = \frac{r_m}{2} (N_T)$$

conc of traps

$r_i \gg r$ even if $N_T \ll N_A$

Can dominate



Summary

Carrier action: R&G

- carriers are created and annihilated
- thermal R&G refers to carriers being excited across bandgap at finite temperatures.
- other G processes include photogeneration

> Transient and steady state solutions for $n(t, x)$ or $p(t, x)$

- transient solution, from boundary conditions, typical exponential decay back to equilibrium concentrations
- at steady state solve continuity equation: net flux = 0

Minority carrier diffusion lengths and lifetimes



3.15

pn Junctions

C.A. Ross, Department of Materials Science and Engineering

Reference: Pierret, chapter 5-6.

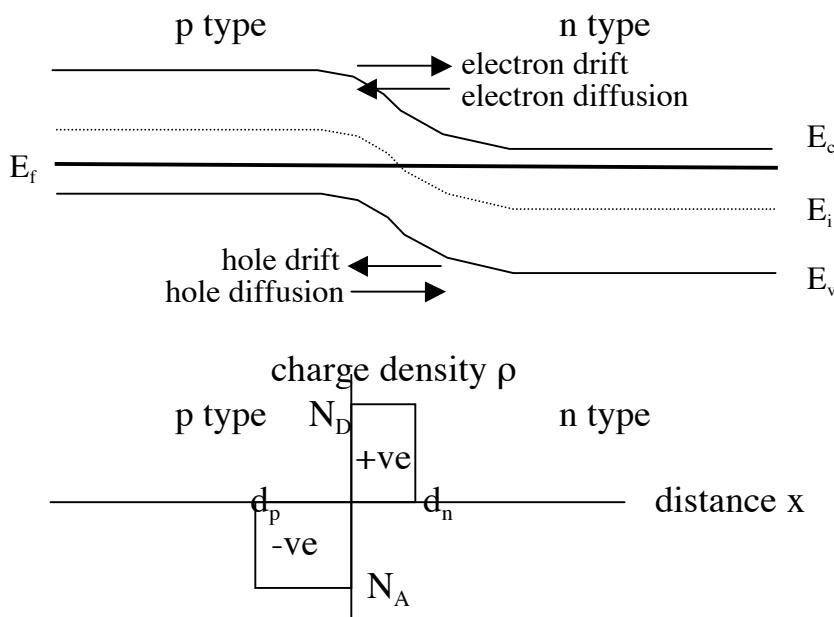
Unbiased (equilibrium) pn junction

Imagine an abrupt pn junction. The p side has a high hole concentration and the n side has a high electron concentration.

There is immediate **diffusion** of the carriers down the concentration gradient.

This leaves a space charge due to the ionized dopants. The resulting electric field leads to **drift** of carriers *in the opposite direction* compared to the diffusion flux.

At equilibrium the drift and diffusion currents are balanced.



$$\text{Gauss' law } \mathbf{E} = 1/\epsilon_0 \epsilon_r \int \rho(x) dx \quad \text{where } \rho = e(p - n + N_D - N_A)$$

Energy (i.e. position of energy bands) = eV; can be found from voltage vs distance; calculate from $\mathbf{E} = -dV/dx$

$$\text{Depletion region width } d = d_p + d_n \quad (\text{some books use } d = x_p + x_n)$$

Built-in voltage V_o : from earlier,

$$n = n_i \exp(E_f - E_i)/kT$$

$$p = n_i \exp(E_i - E_f)/kT$$

The Fermi level is flat across the junction:

$$eV_o = (E_f - E_i)_{n\text{-type}} - (E_f - E_i)_{p\text{-type}} \\ = kT \ln(n_n/n_p) \text{ or } kT \ln(N_A N_D / n_i^2)$$

Using the depletion approximation $\rho = -N_A e$ in the p-type and $N_D e$ in the n-type:

$$E = N_A e d_p / \epsilon_o \epsilon_r = N_D e d_p / \epsilon_o \epsilon_r \quad \text{at } x = 0 \\ V_o = (e / 2\epsilon_o \epsilon_r) (N_D d_n^2 + N_A d_p^2) \\ d_n = \sqrt{\{(2\epsilon_o \epsilon_r V_o / e) (N_A / (N_D (N_D + N_A)))\}} \\ d = d_p + d_n = \sqrt{\{(2\epsilon_o \epsilon_r V_o / e) (N_D + N_A) / N_A N_D\}}$$

Biased pn junction (apply voltage V_A)

Forward bias raises the n-side energy levels (or lowers the p-side)

by applying -ve to the n-side (or +ve to p-side)

This **reduces** the voltage barrier. The quasi-Fermi level is higher on the n-side.

The diffusion term changes because the number of carriers eligible to diffuse increases exponentially.

The drift term does not change.

Outside the depletion region there is a net diffusion current.

Reverse bias lowers the n-side energy levels.

Diffusion is reduced; drift is unchanged. Only a small reverse current flows.

Reverse bias increases the depletion width (V_A is -ve)

$$d = \sqrt{\{(2\epsilon_o \epsilon_r (V_o - V_A) / e) (N_D + N_A) / N_A N_D\}}$$

The ideal diode equation

In forward bias the diffusion flux increases because more carriers are able to diffuse. This comes from the Fermi function. When E_f is away from the band edge,

$$f(E) = 1 / \{1 + \exp(E - E_f)/kT\} \sim \exp(-(E - E_f)/kT)$$

If we shift the energy levels by V_A , we change the available number of carriers by a factor

$$\{\exp(-(e(V_o - V_A) - E_f)/kT\} / \{\exp(-(eV_o - E_f)/kT\} \\ = \exp eV_A/kT$$

Therefore diffusion flux $J_{\text{diff}} = J_o \exp eV_A/kT$

To evaluate J_o , we know that $J_o = -J_{drift} = J_{diff}$ at $V_A = 0$.

Consider an asymmetric junction with $N_A \gg N_D$, then the current is mainly holes, and their concentration decays in the n-type material (outside the depletion region) over a distance $\lambda_p = \sqrt{(\tau_p D_p)}$. The diffusion current

$$J_{diff} = eD_p \nabla p = eD_p (p_{n(x=0)} - p_{no}) / \lambda_p \quad (\text{where } p_{no} = n_i^2 / N_D)$$

$$\sim eD_p (p_{n(x=0)}) / \lambda_p$$

$$p_n = p_p \exp -eV_o/kT \text{ (unbiased)}$$

$$\text{and } p_n = p_p \exp -e(V_o - V_A)/kT \text{ (forward biased)}$$

$$\text{so } p_n = p_{no} \exp -eV_A/kT$$

$$\text{Hence } J_{diff} = \{eD_p n_i^2 / N_D \lambda_p\} \exp -eV_A/kT = J_o \exp eV_A/kT$$

Include both electron and hole terms: $J_o = en_i^2 \{D_p/N_D \lambda_p + D_n/N_A \lambda_n\}$

Also, $J_{drift} = J_o$ gives an expression for J_{drift}

The ideal diode equation is then

$$J = J_{diff} + J_{drift} = J_o \{\exp eV_A/kT - 1\}$$

What happens in reverse bias? The current reaches a reverse saturation value of J_o ($\sim 10^{-12} \text{ A cm}^{-2}$ in Si)

All minority carriers reaching the depletion region are sucked across (i.e. the junction ‘collects’ minority carriers). There is no diffusion flux across the depletion region. There is a diffusion flux outside the depletion region that supplies minority carriers to the junction: its value is just $-en_i^2 \{ p_{no} D_p / \lambda_p + n_{po} D_n / \lambda_n \} = -J_o$.

**Reverse bias pn junction collects minority carriers
Forward bias pn junction injects minority carriers**

Non-idealities:

- a) Reverse bias Zener breakdown, where carriers tunnel through a narrow depletion width
- b) Avalanche diode, where impact ionization generates more carriers in the depletion region.

pn junctions

Reference: Handout 3; Pierret Ch. 5-6

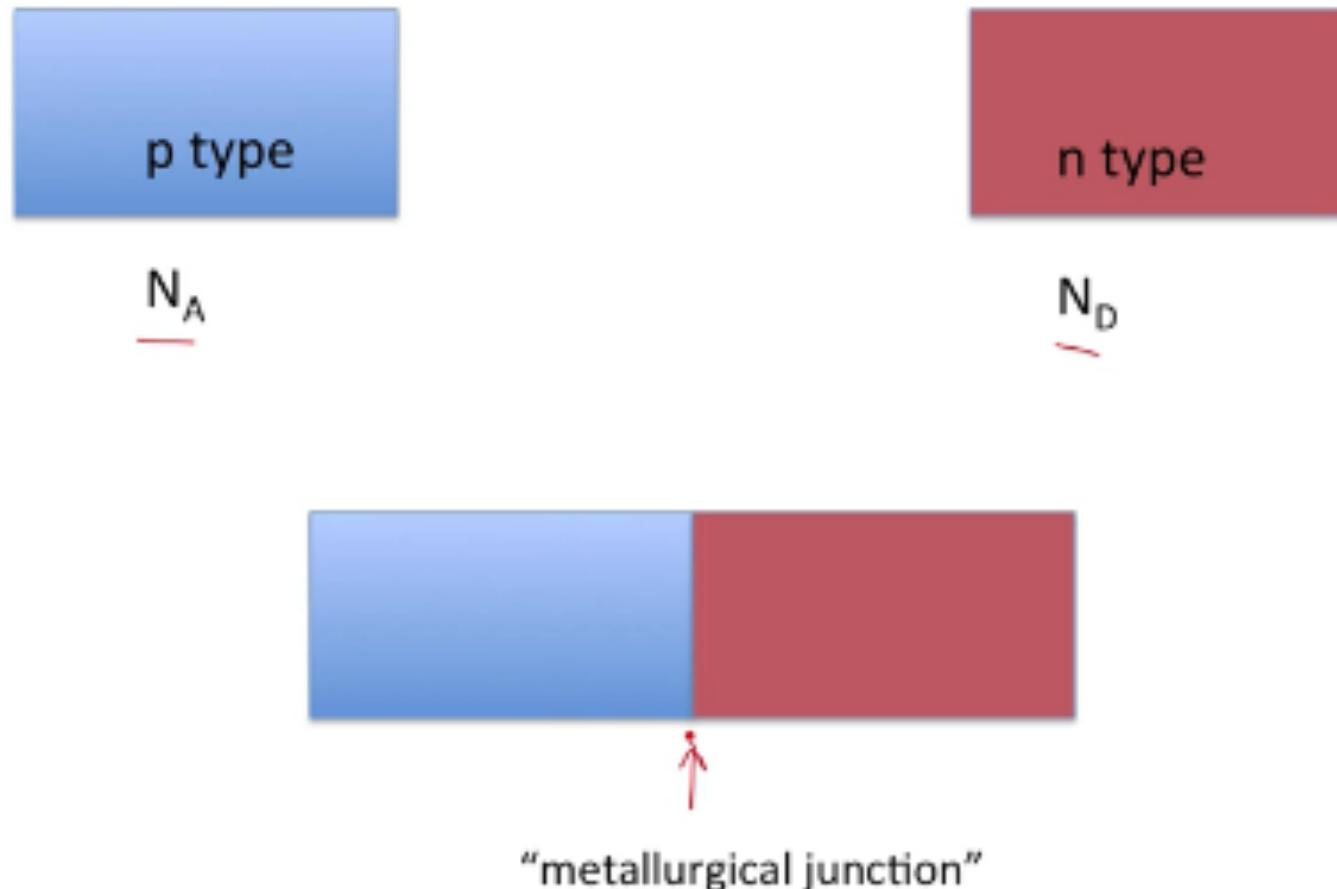
We will study

- • Electrostatics and band structure
- • The effect of bias
- • The ideal diode equation
- .— • Some examples of diodes

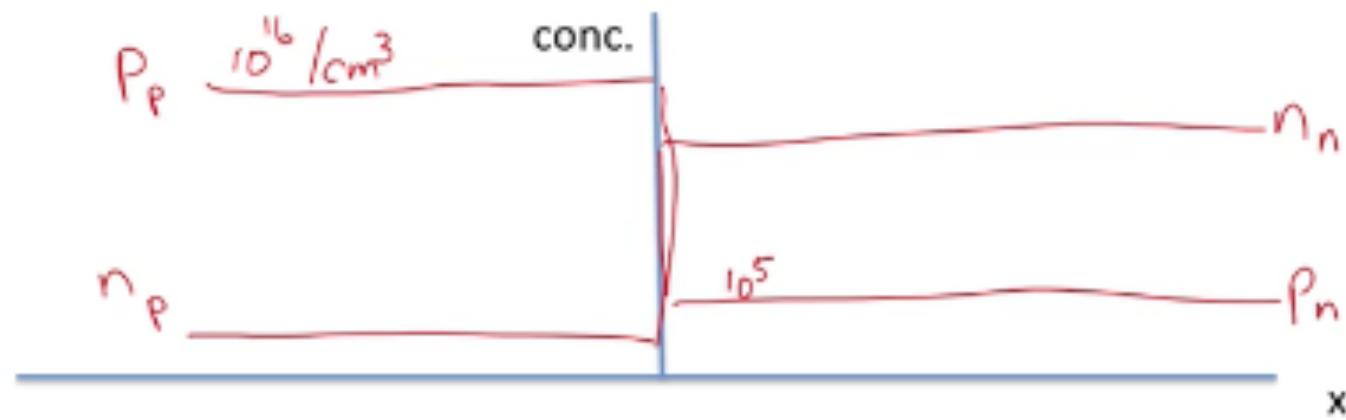


pn electrostatics

What diffusion and drift to you expect at a pn junction?



Carrier concentrations



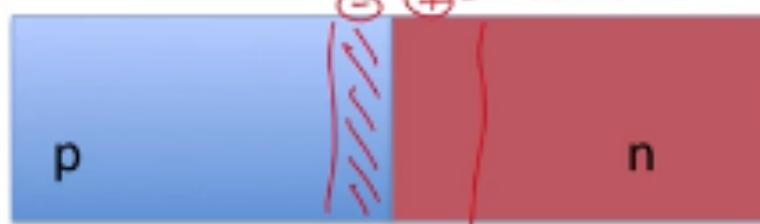
Eam.

drift + diffusion = 0

→ holes diffuse
← e diffuse

$$N_D < N_A$$

due to ionized dopants



↑ ↓ depleted electrons
depleted holes

←
Electric field
→ electrons & holes



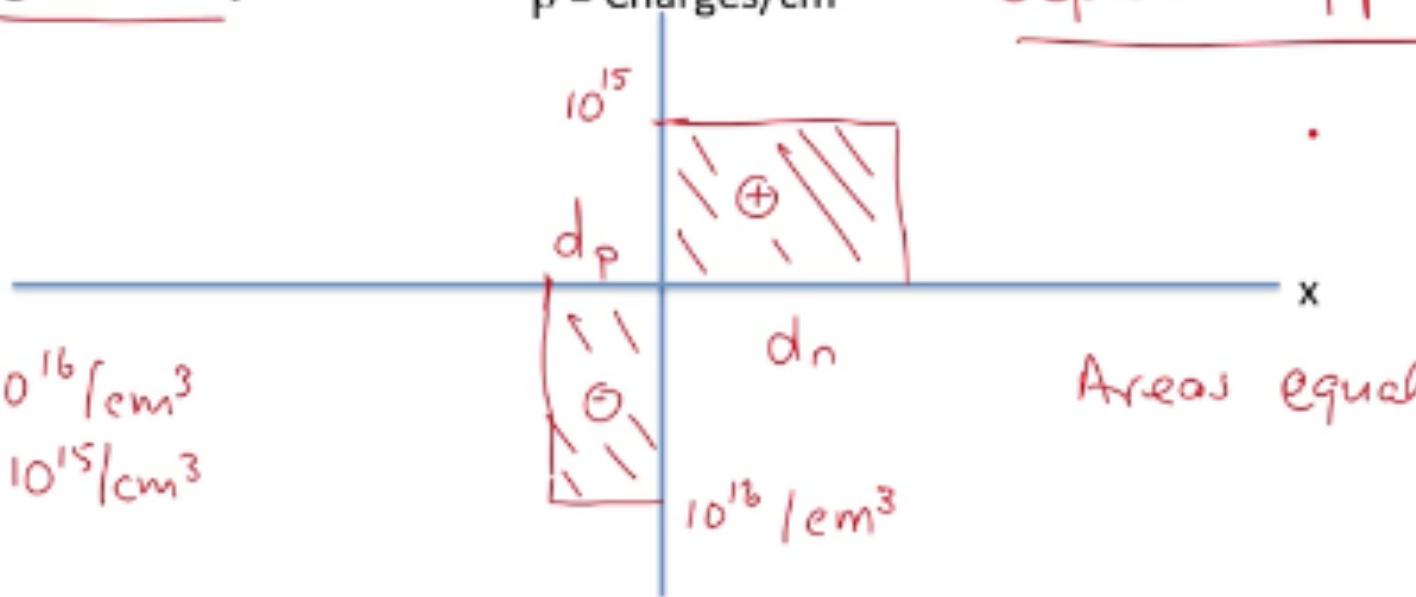
Charge density

$$N_A = 10^{16} / \text{cm}^3$$

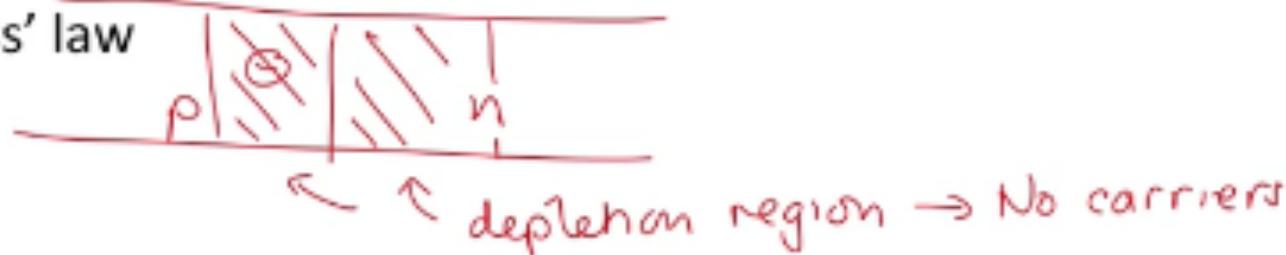
$$N_D = 10^{15} / \text{cm}^3$$

$$\rho = \text{Charges/cm}^3$$

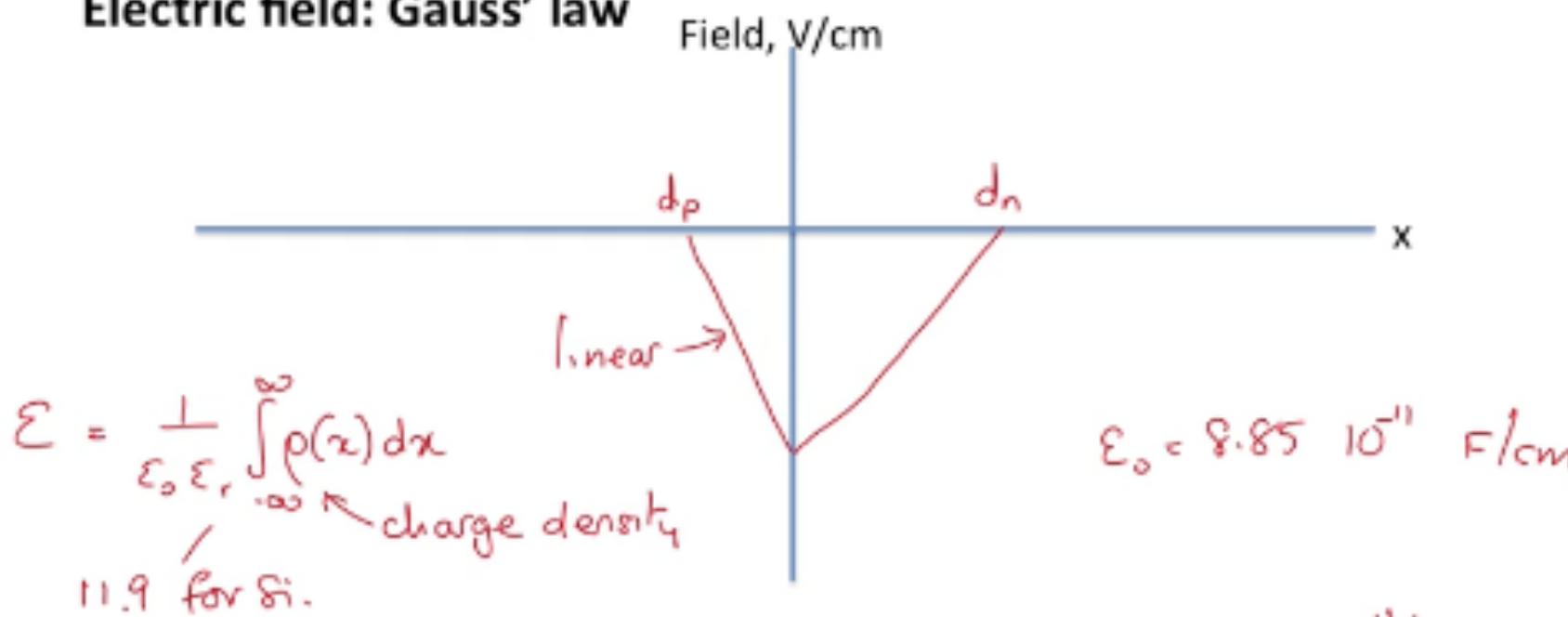
Depletion approximation



Electric field: Gauss' law

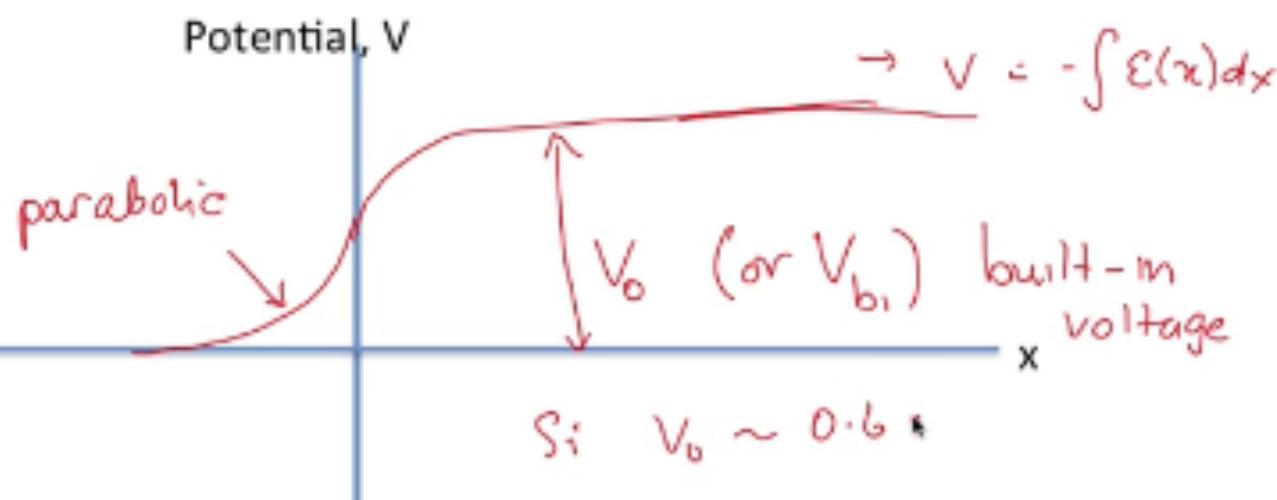


Electric field: Gauss' law



$$\epsilon_0 = 8.85 \cdot 10^{-12} \text{ F/cm}$$

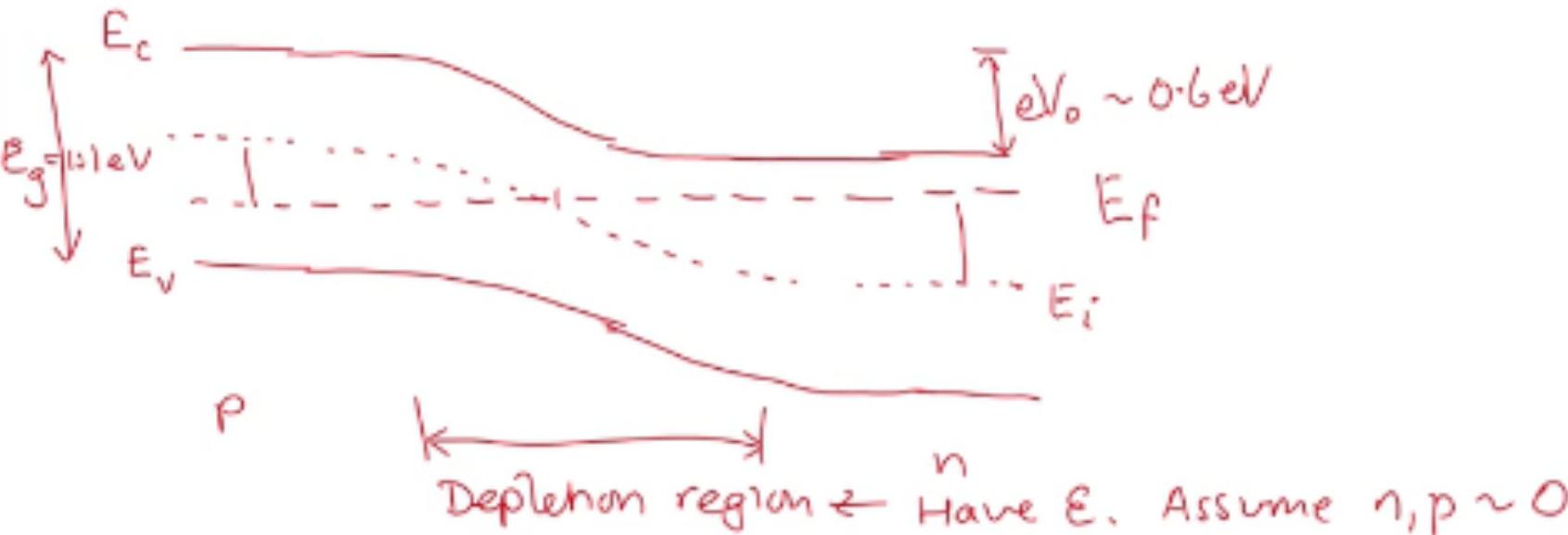
$$E = -\frac{dV}{dx}$$



Band diagram of a pn junction

at eqm $E_f = \text{const}$

Energy levels $\approx -\text{voltage}$



$$n_n = n_i \exp(E_f - E_i)/kT \quad \text{so } (E_f - E_i) = kT \ln(n_n/n_i)$$

e in n-type

n_n = number of electrons in n-type

$$\begin{aligned} eV_0 &= (E_f - E_i)_{\text{n-type}} - (E_f - E_i)_{\text{p-type}} \\ &= kT \ln(n_n p_p / n_i^2) \end{aligned}$$

$$\Rightarrow V_0 = \frac{kT/e \ln(N_A N_D / n_i^2)}{\sim 0.6 \text{ eV at RT in Si}}$$

10^{16} 10^{15} 10^{10}
 | | /



The depletion approximation

If we assume the depletion region really has no electrons or holes, we can derive the depletion width and V_0 :

$$\rightarrow \rho = -N_A e \text{ on p-side, } +N_D e \text{ on n-side}$$

$$E = N_A e d_p / \epsilon_0 \epsilon_r = N_D e d_n / \epsilon_0 \epsilon_r \quad \text{at } x = 0$$

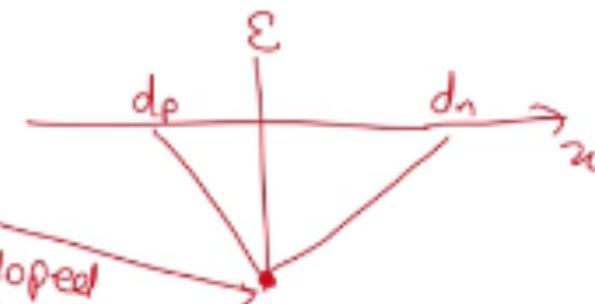
$\rightarrow d_p$ is smaller than d_n when $N_A > N_D$

Depletion width is smaller on more heavily doped side

$$V_0 = \frac{e}{2\epsilon_0 \epsilon_r} (N_D d_n^2 + N_A d_p^2) \quad \sim 0.6 \text{ eV}$$

$$d_n = \left\{ \frac{2\epsilon_0 \epsilon_r V_0}{e} \frac{N_A}{N_D (N_D + N_A)} \right\}^{0.5}$$

$$d = d_p + d_n = \left\{ \frac{2\epsilon_0 \epsilon_r V_0}{e} \frac{N_D + N_A}{N_D N_A} \right\}^{0.5}$$

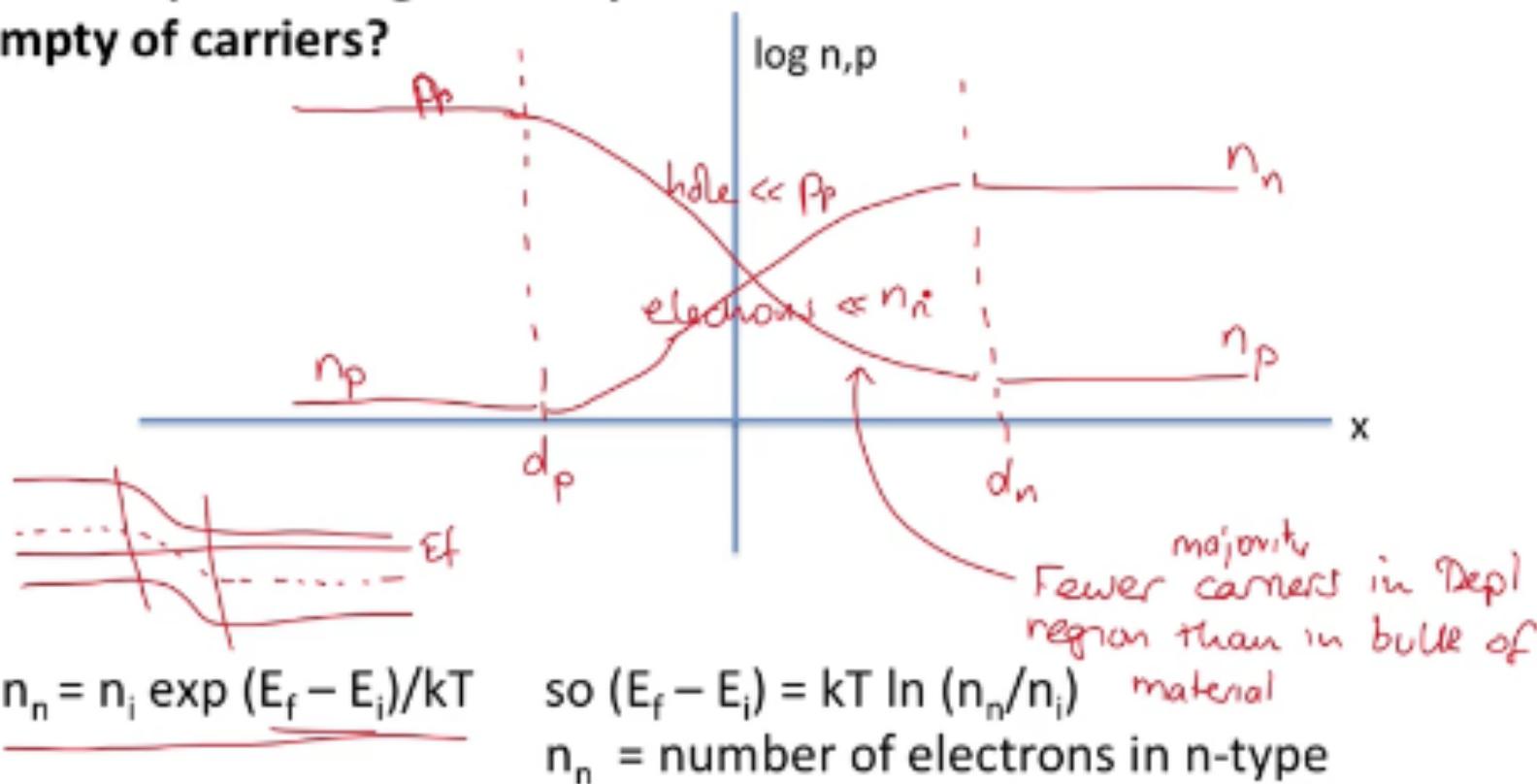


= total depletion width

1 - 3 μm



Is the depletion region really empty of carriers?



$$\begin{aligned}eV_0 &= (E_f - E_i)_{\text{n-type}} - (E_f - E_i)_{\text{p-type}} \\&= kT \ln(n_n p_p / n_i^2)\end{aligned}$$

Summary

Put p and n type materials in contact: diffusion of carriers leads to depletion of carriers near the interface and a space charge which causes an electric field.

The electric field causes drift that balances diffusion at equilibrium.

Assuming full depletion of carriers (the depletion approximation), from the space charge we integrate to get electric field, then integrate again to get voltage across the junction. The bands bend in the depletion region, indicating that an electric field is present.

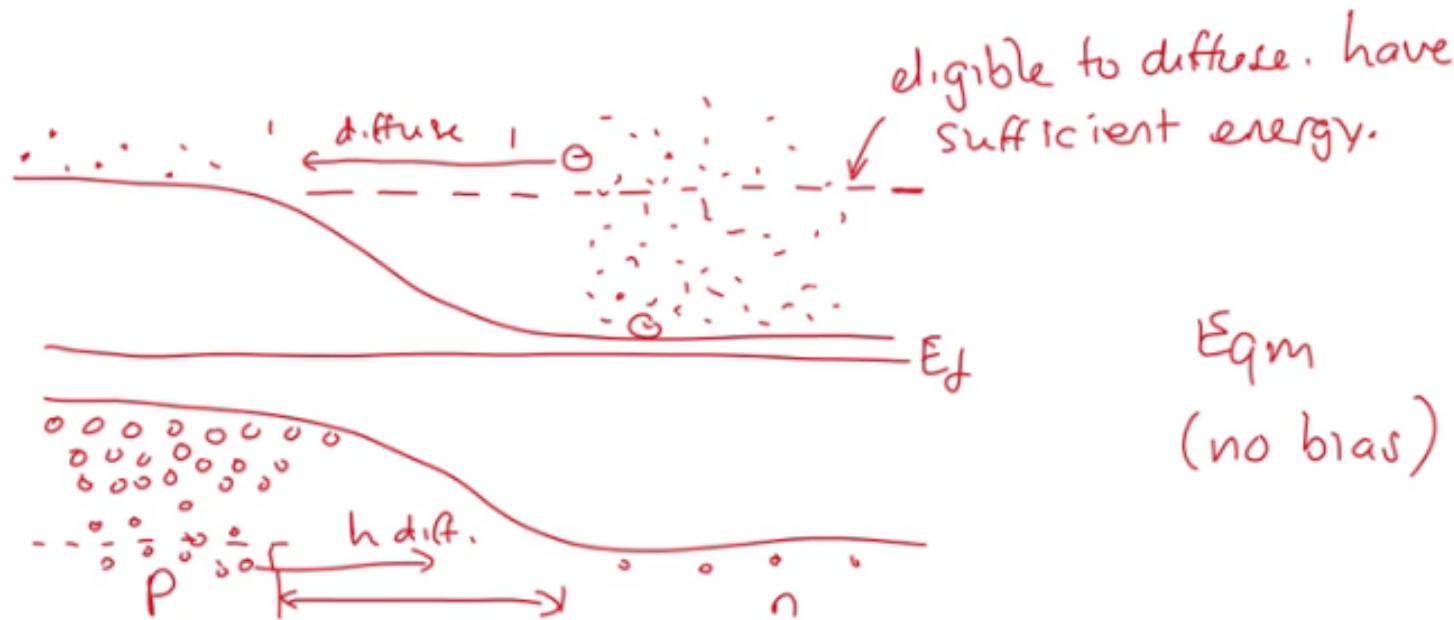
The built-in voltage is $V_o = kT/e \ln (N_D N_A / n_i^2)$

The depletion width is $d = d_p + d_n = \left\{ \frac{2\epsilon_o \epsilon_r V_o}{e} \frac{N_D + N_A}{N_D N_A} \right\}^{0.5}$

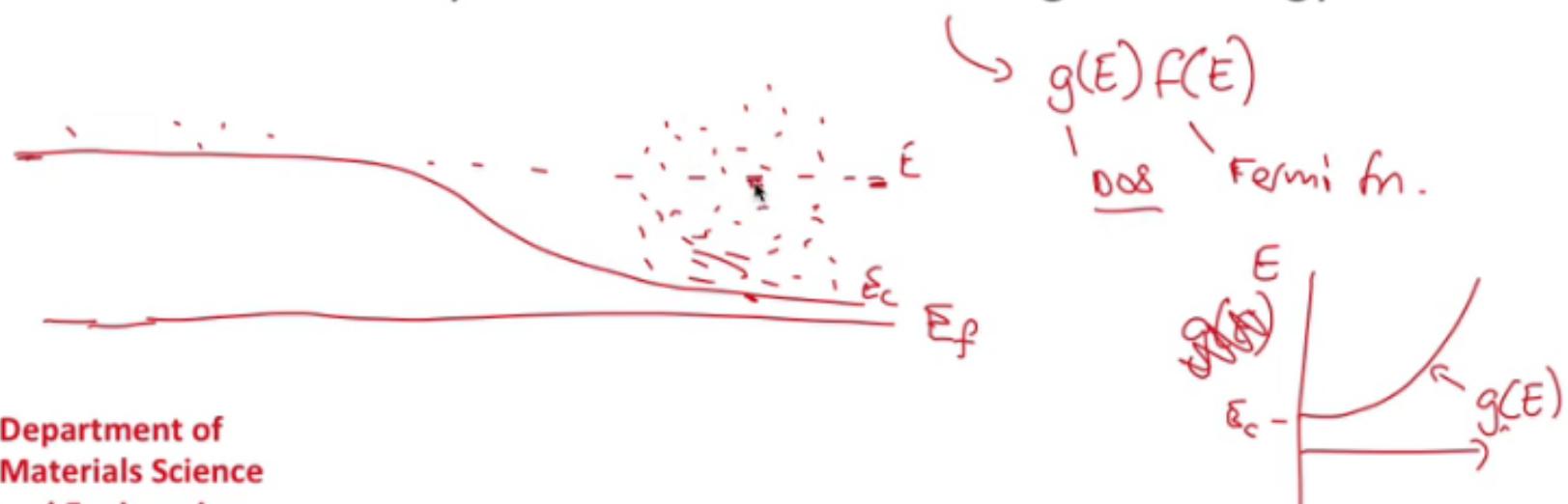
Next: What happens when we apply a voltage across the pn junction?

Effects of Bias on the pn junction

= applied voltage

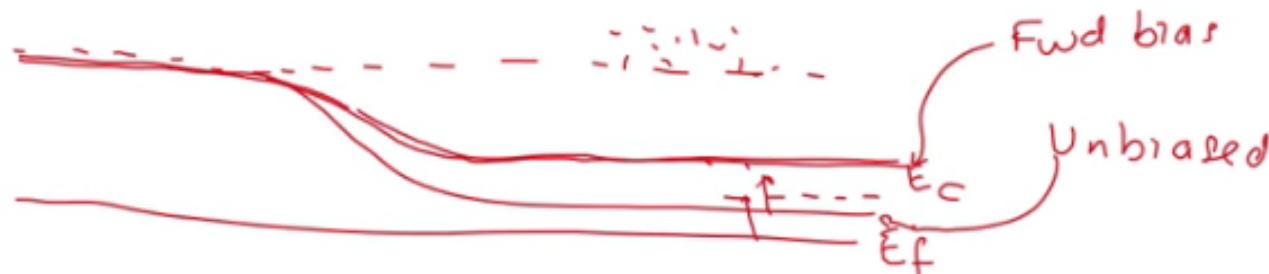


Unbiased case: how many electrons are there at a given energy:



Forward bias raises the n-side (or lowers the p-side)

Diffusion involves only carriers with sufficient energy to cross barrier.
Exponentially more of them are able to diffuse as V_A increases.

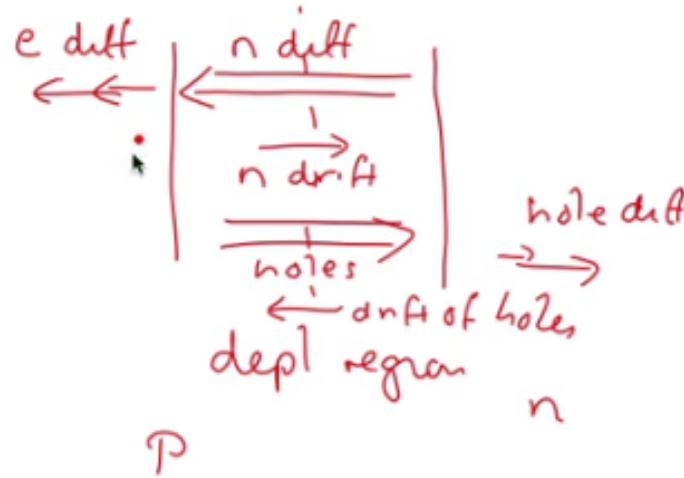
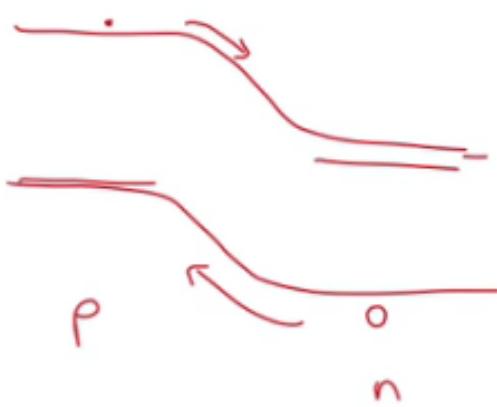


carriers that can diffuse $\propto g(E) \underline{f(E)}$



Forward bias raises the n-side (or lowers the p-side)

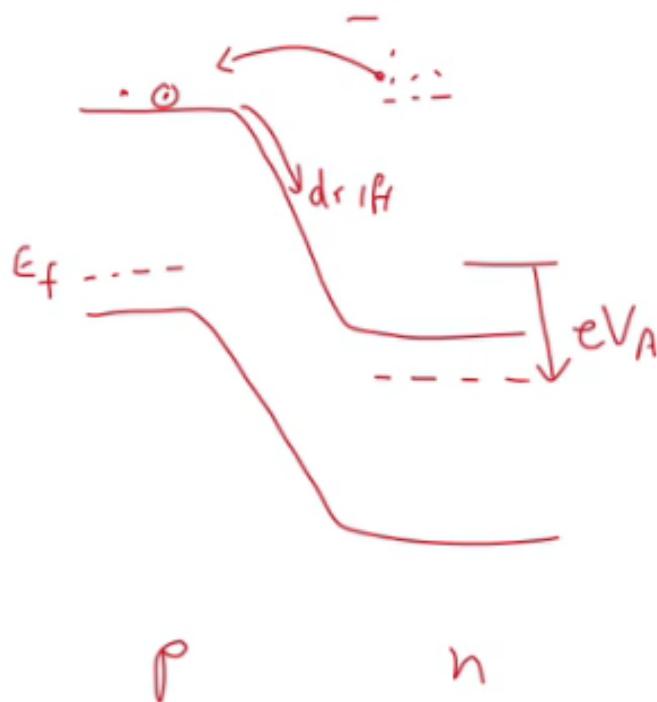
The drift term is unaffected to first order. Any minority carrier that reaches the depletion region will drift across it, irrespective of the magnitude of the electric field.



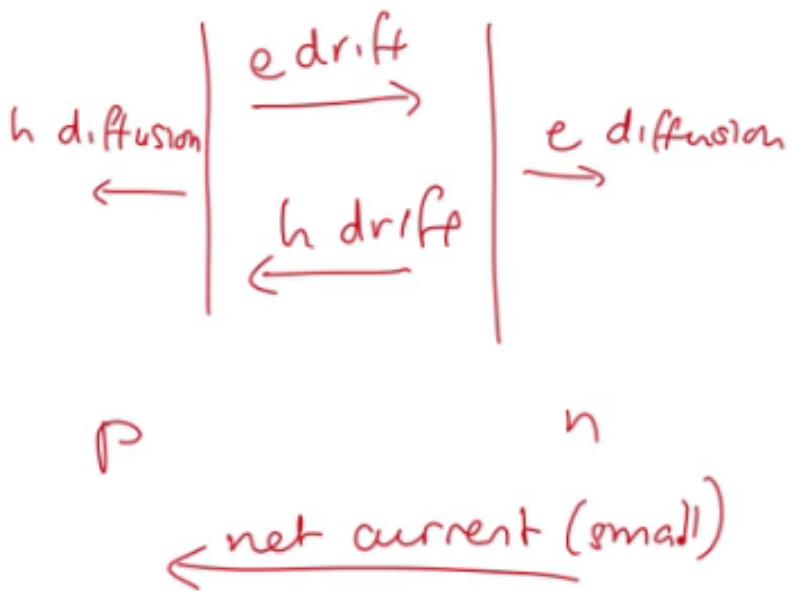
Outside the depletion region:



Reverse bias $-V_A$ lowers the n-side (or raises the p-side)



Diffusion fluxes reduced

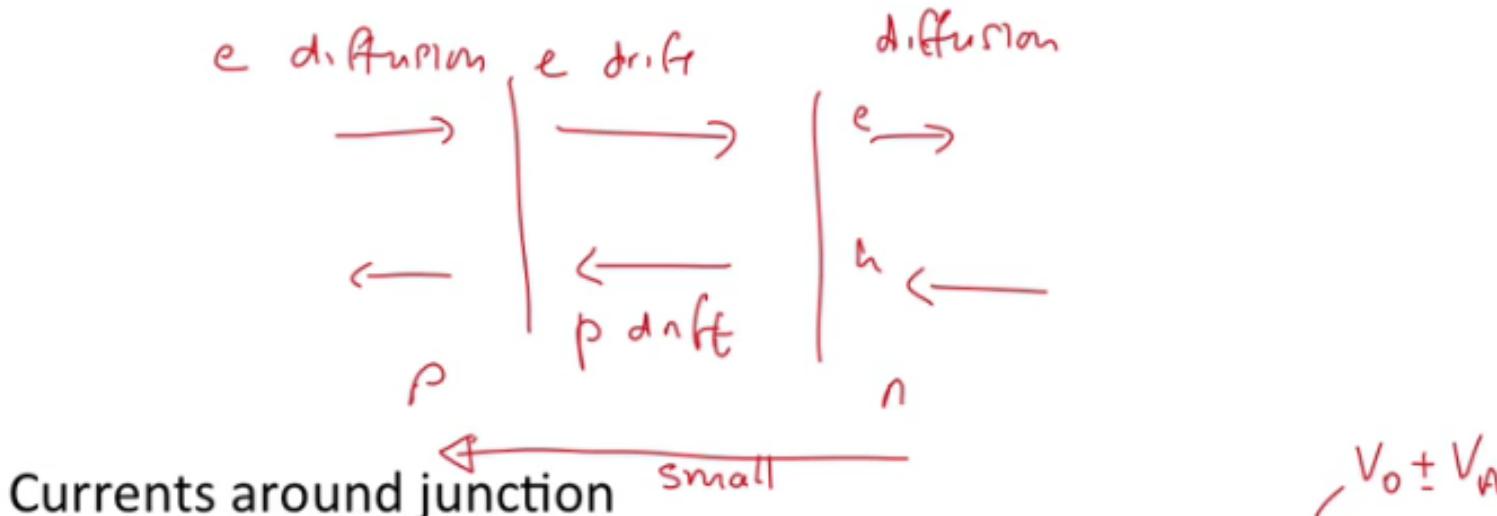


Increases the barrier.

Diffusion is shut off. Drift provides small reverse current.



Reverse bias $-V_A$ lowers the n-side (or raises the p-side)



Depletion width increases:

$$d = d_p + d_n = \left\{ \frac{2\epsilon_0\epsilon_r V_o}{e} \frac{N_D + N_A}{N_D N_A} \right\}^{0.5}$$

Reverse bias junctions **collect minority carriers.**



Summary

If we apply a voltage (bias) across a pn junction, we can visualise this as bending the bands.

Forward bias (positive voltage at p-side): V_A raises the n-side (or lowers the p-side). Barrier is reduced and exponentially more carriers are able to diffuse across it; large current flows.

Reverse bias (negative voltage at p-side): V_A lowers the n-side (or raise the p-side). Barrier is raised and diffusion is cut off; only a small reverse drift current is present.

Forward bias: Inject minority carriers (i.e. inject n into p-side)

Reverse bias: Collect minority carriers

Next: Calculate the I-V relation in the junction: the Ideal Diode Equation.

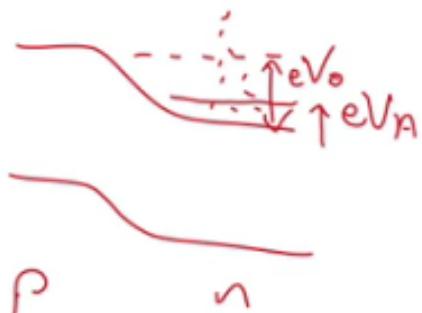


> The ideal diode equation

$I - V$

1. evaluate the diffusion flux vs. V_A .
2. evaluate the diffusion flux at $V_A = 0$
3. combine with drift term, assumed independent of V_A .

Diffusion flux



$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_f}{kT}\right)} \\ \approx \underbrace{\exp\left(-\frac{(E - E_f)}{kT}\right)}$$

eligible carriers
 $\propto \exp\left(eV_A/kT\right)$

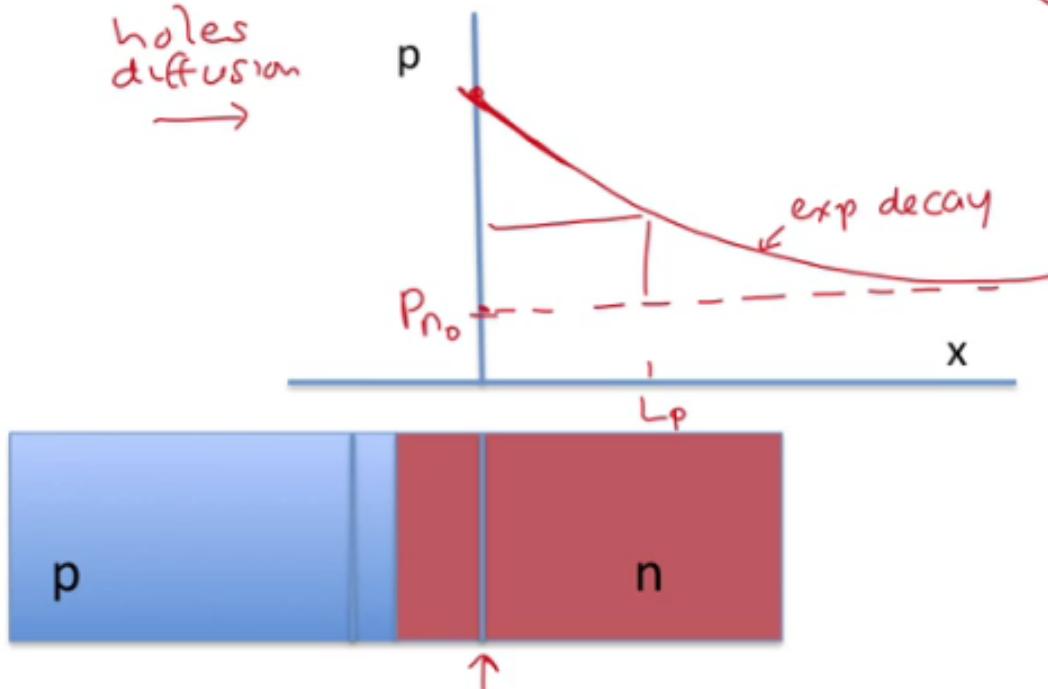
$$\frac{\text{Carriers at } V_o}{\text{Carriers at } (V_o - V_A)} = \frac{\exp\left(-(eV_o - E_f)/kT\right)}{\exp\left(-(eV_o - eV_A - E_f)/kT\right)} = \frac{\exp\left(-eV_A/kT\right)}{}$$



$$= \exp\left(-eV_A/kT\right)$$

Diffusion flux at the edge of the depletion region

Holes decay in n-side over a length $L_p = \sqrt{D_p \tau_p}$



$$\begin{aligned} J_{d,f} &= e \frac{\Delta P}{L_p} D_p \\ &= e(P_n(x=0) - \cancel{P_{n_0}}) D_p \\ &\simeq e P_n(x=0) \underline{D_p / L_p} \end{aligned}$$



We need $p_n(x=0)$:

> At equilibrium, $p_{n_0} = p_p \exp(-eV_o/kT)$

> Under bias, $p_n = p_p \exp(-e(V_o - V_A)/kT)$

$$P_n = P_{n_0} \exp(eV_A/kT)$$

$$\Rightarrow P_n(x=0) \text{ is just } P_{n_0} \exp(eV_A/kT)$$

Substitute into diffusion equation:

$$J_{\text{diff}} = J_o \exp(eV_A/kT)$$

Holes only

$$\text{where } J_o = e(D_p/L_p)p_{n_0} = (eD_p/L_p)(n_i^2/N_D)$$

electrons

$$\text{so } J_{\text{diff}} = (eD_p/L_p n_i^2/N_D + eD_n/L_n n_i^2/N_A) \exp(eV_A/kT)$$

J_s



Drift flux

$$\text{at } V_A = 0, J_{\text{diff}} + J_{\text{drift}} = 0$$

$$\text{so } J_{\text{drift}} = - \left(\frac{eD_p}{L_p} n_i^2 / N_D + \frac{eD_n}{L_n} n_i^2 / N_A \right) = -J_0$$

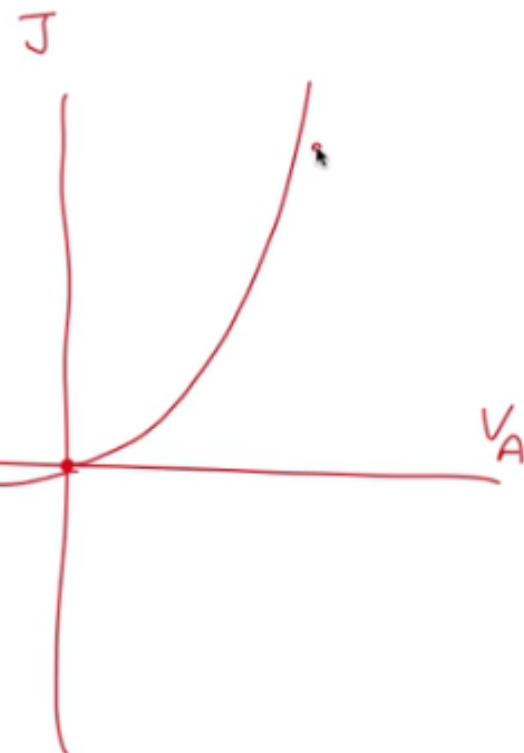
independent of V_A

$$J = J_{\text{diff}} + J_{\text{drift}}$$
$$= J_0 \exp \left(\frac{eV_A}{kT} \right) - J_0$$

$$\Rightarrow J = J_0 \left(\exp \frac{eV_A}{kT} - 1 \right)$$

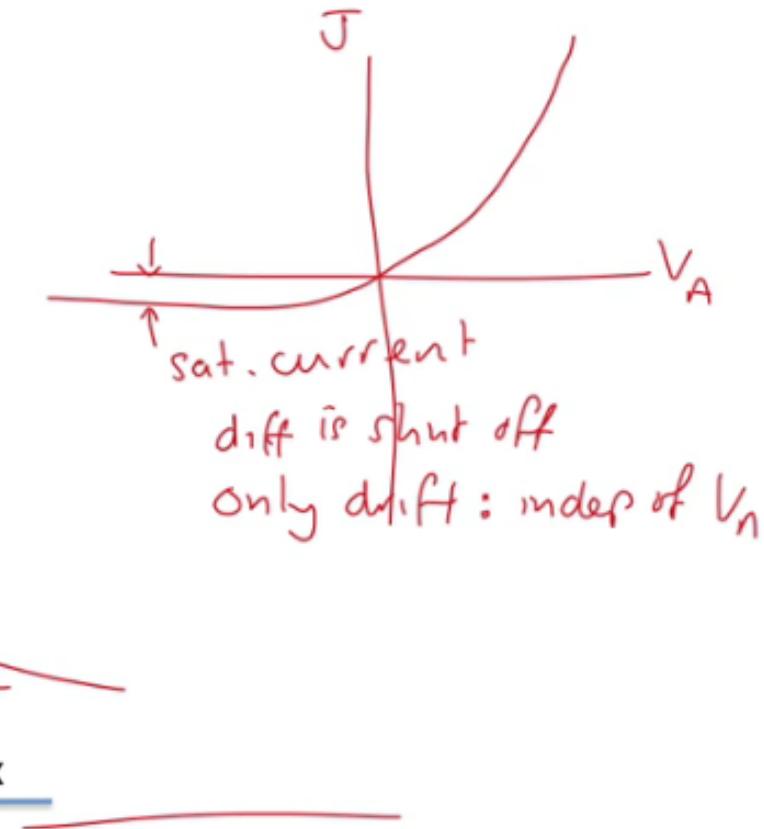
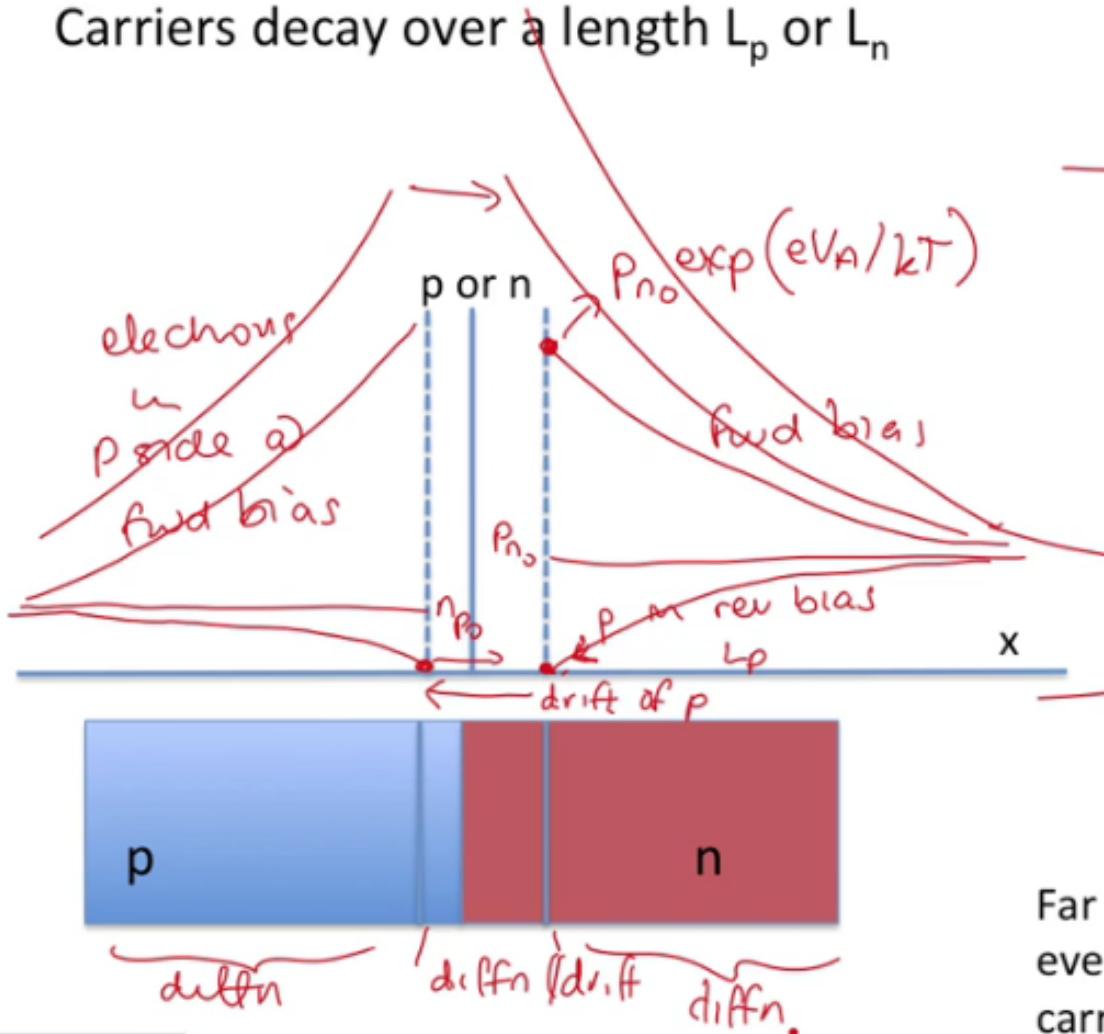
$$J_0 \sim 10^{-12} \text{ A/cm}^2 \text{ in Si}$$

\downarrow
drift term
reverse saturation



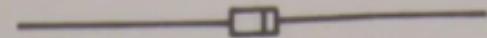
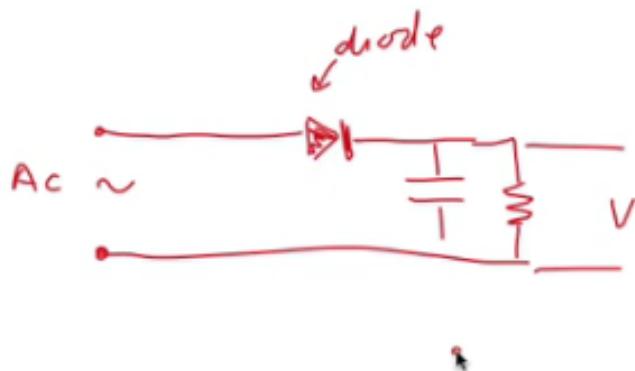
Carrier concentrations in biased junction

Carriers decay over a length L_p or L_n



Far from the junction, diffusion flux eventually becomes a majority carrier drift current, since electric field is not exactly zero.

Examples of diodes



Band Marks
Cathode End



Anode

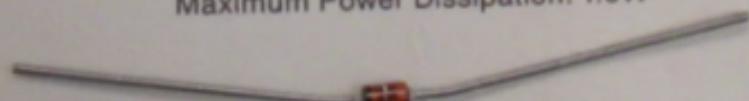
Cathode

Typical Electrical Characteristics

Voltage (V_Z): 6.2V

Current (I_Z): 41mA

Maximum Power Dissipation: 1.0W



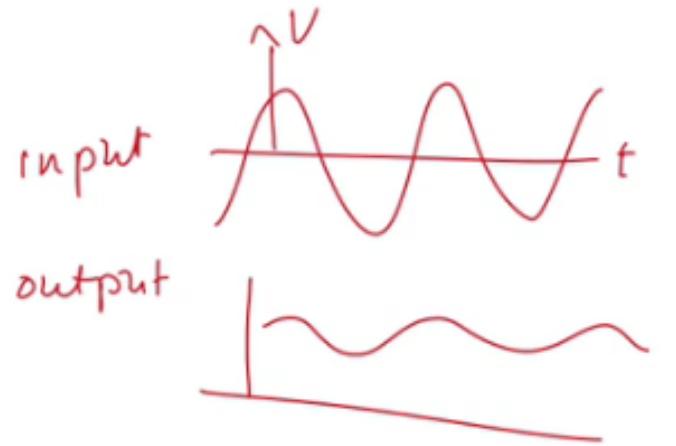
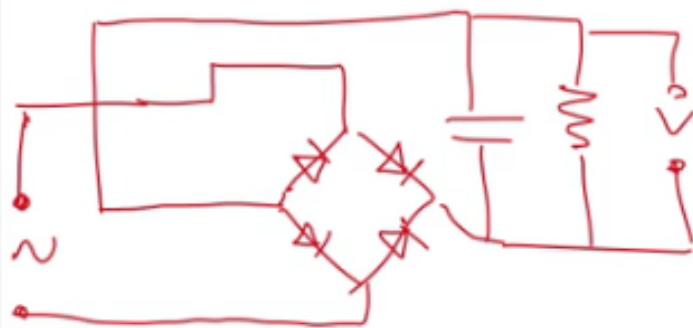
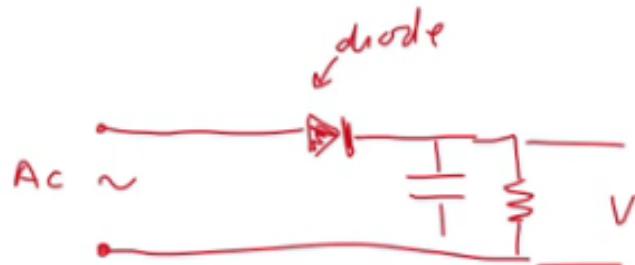
For Further Information
See Radio Shack Data Books

The crystal radio is a good example
of using a diode in a circuit to
obtain AM signal



Department of
Materials Science
and Engineering

Examples of diodes

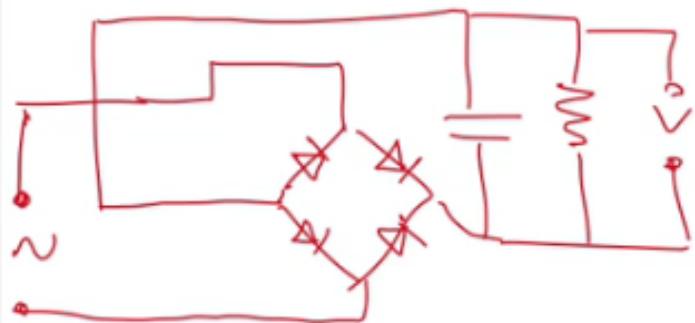
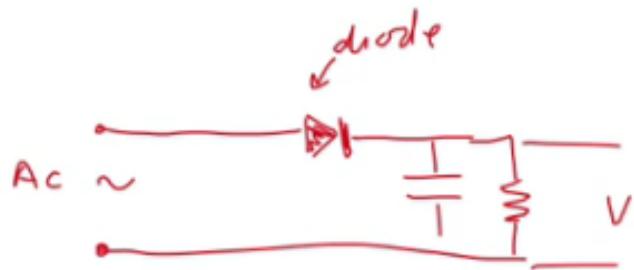


AC \rightarrow DC rectification

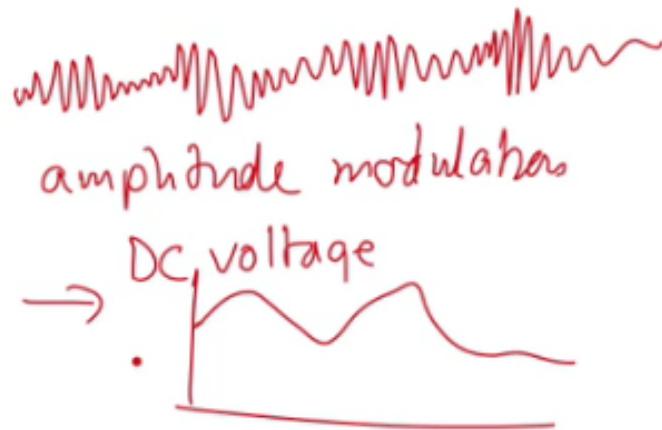
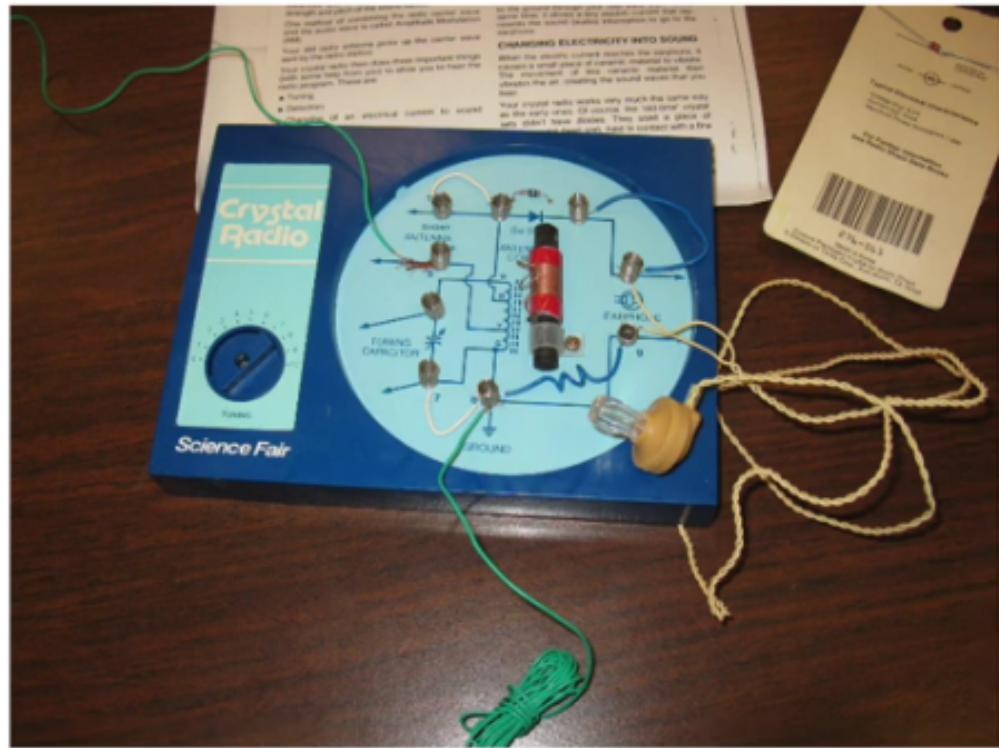
The crystal radio is a good example
of using a diode in a circuit to
obtain AM signal



Examples of diodes

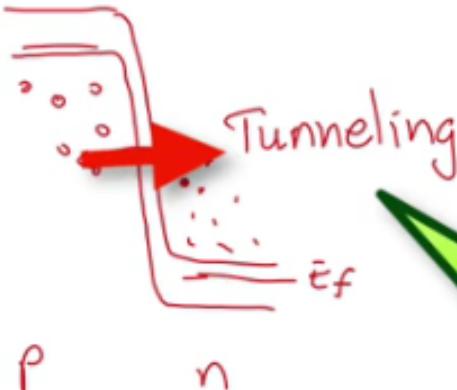


The crystal radio is a good example of using a diode in a circuit to obtain AM signal



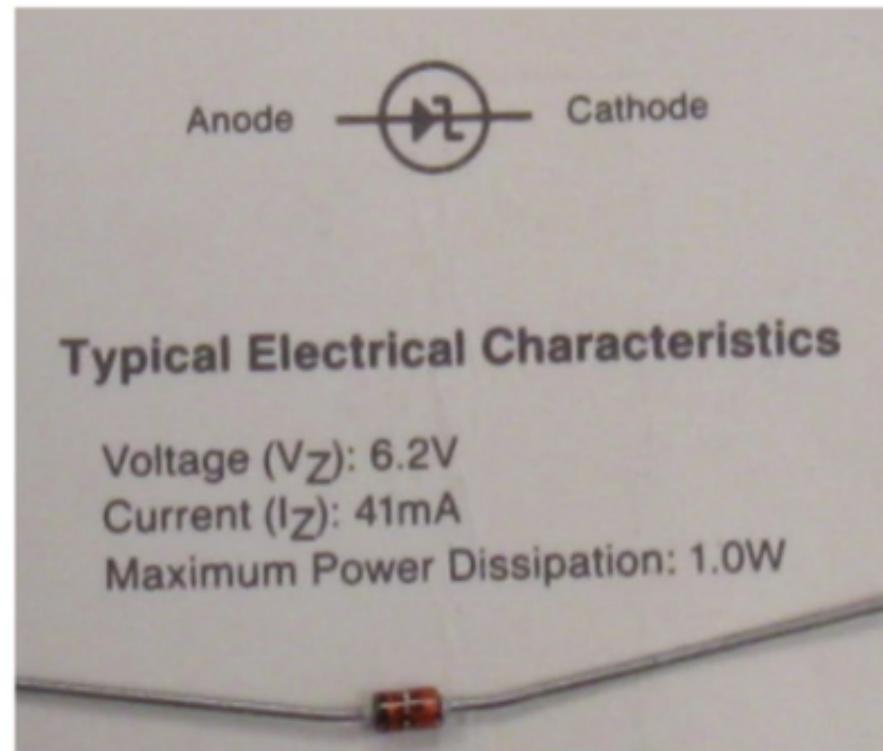
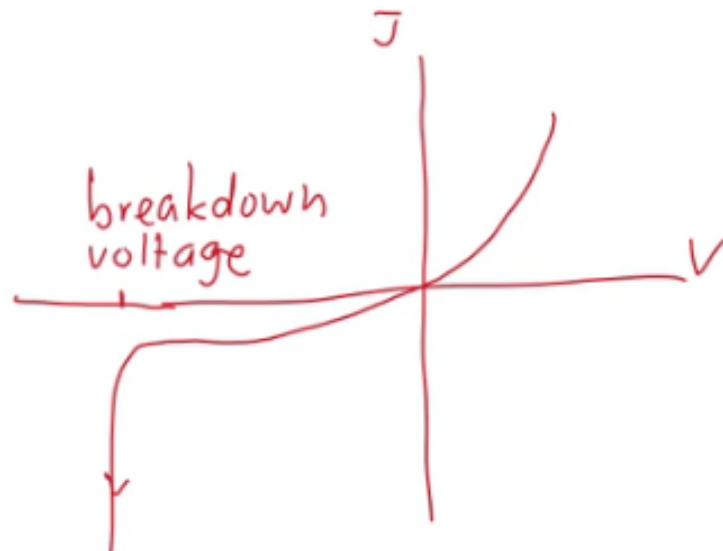
Examples of diodes

Zener diodes (tunneling)



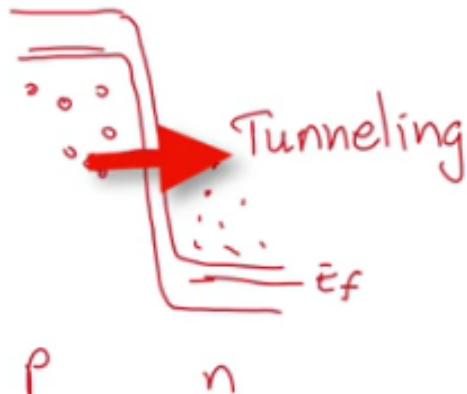
Electrons from the mostly full VB tunnel into states in the mostly empty CB making a reverse current.

Avalanche diodes



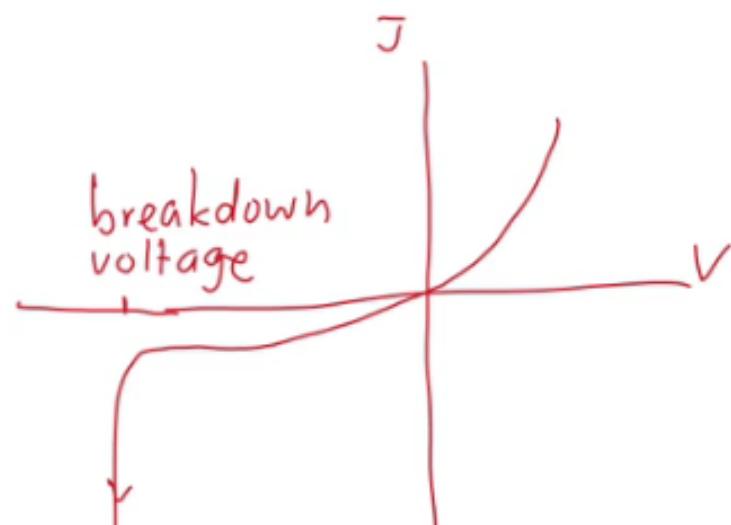
Examples of diodes

Zener diodes (tunneling)



→ highly doped

Occurs for $V_A \sim 10V_0$



Avalanche diodes



Typical Electrical Characteristics

Voltage (V_Z): 6.2V

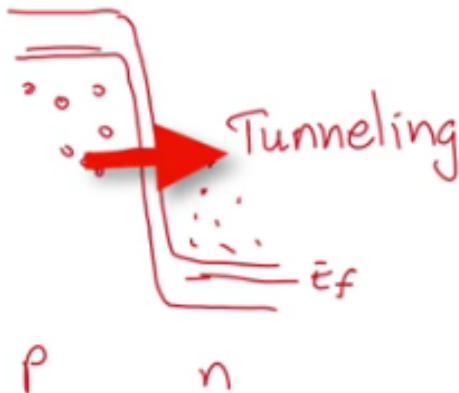
Current (I_Z): 41mA

Maximum Power Dissipation: 1.0W



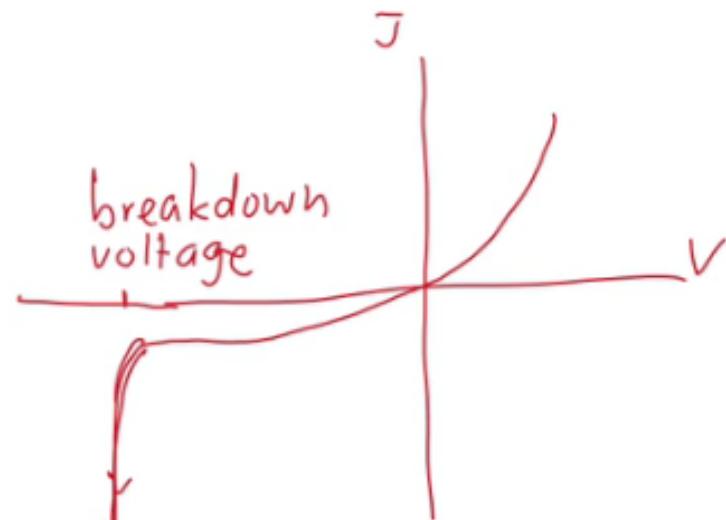
Examples of diodes

Zener diodes (tunneling)



→ highly doped

Occurs for $V_A \sim 10V_0$



Avalanche diodes



highly doped



Summary

- pn junction electrostatics
- pn: effect of bias
- ideal diode equation: evaluate diffusion term as a function of barrier height, and assume drift term is independent of bias.

$$J = (eD_p/L_p n_i^2/N_D + eD_n/L_n n_i^2/N_A) (\exp(eV_A/kT) - 1)$$

- Also consider the decay of excess carrier concentrations away from the depletion region.
- practical diodes (Zener: rely on tunneling across a narrow depletion width. Avalanche: rely on carrier multiplication in large electric field.)

Next up: **Bipolar junction transistor**



3.15

Transistors

C.A. Ross, Department of Materials Science and Engineering

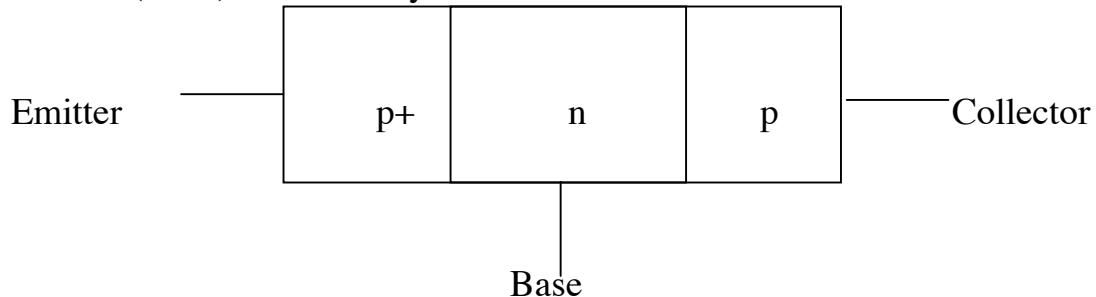
Reference: Pierret, chapter 10, 15.1-2, 16.1-2 and 17.1.

Transistors are three-terminal devices that use a small voltage (or current) applied to one contact to modulate (i.e. control) a large voltage (or current) between the other two contacts.

An analogy is the vacuum tube from the 1900s. A small voltage applied to the ‘grid’ modulates a large current between the anode and cathode.

The bipolar junction transistor

The BJT (1947) is a minority carrier device



Four modes (forward active, reverse active, saturation, cutoff)

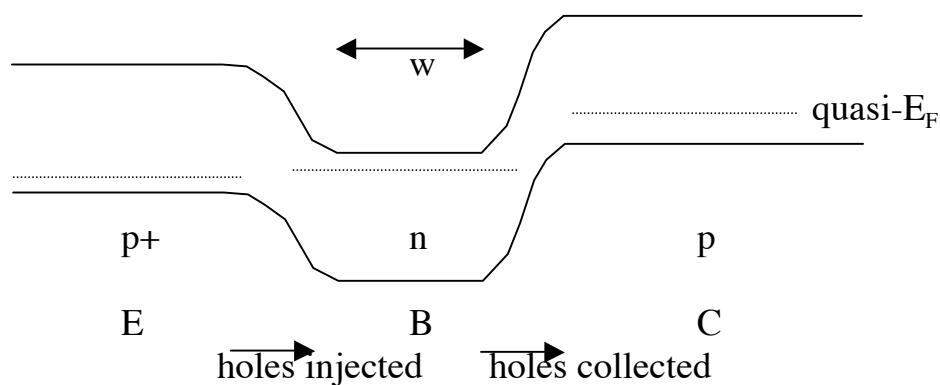
Forward active operation (pnp transistor)

The emitter E emits minority carriers into the base B: EB is forward biased
Holes survive their journey through the base

The collector C collects minority carriers from the base B: BC is reverse biased.

A small voltage (or current) at B has a large effect on the current I_E

$$I_B \ll I_E \sim I_C$$



Forward current I_E

At the edge of the depletion region at the left of the base,

$$p = (n_i^2/N_{D,B}) \exp(eV_{EB}/kT)$$

At the right of the base, $p = 0$

$$\text{Current across base } I_E = (eD_p/w) (n_i^2/N_{D,B}) \exp(eV_{EB}/kT)$$

Base current I_B

The base current \ll collector current, so most current goes straight through from E to C.

$$\text{Current gain } \alpha = I_C/I_E \sim 1, \beta = I_C/I_B \sim 100 - 1000$$

$$I_B = I_n \text{ (electrons going from B to E in forward bias)} \\ + I_{nC} \text{ (electrons going from C to B in reverse bias -small)} \\ + eR_B \text{ (recombination in base -small)}$$

The gain $\beta = I_C/I_B \sim I_E/I_B = N_{A,E}/N_{D,B}$ is determined by doping.

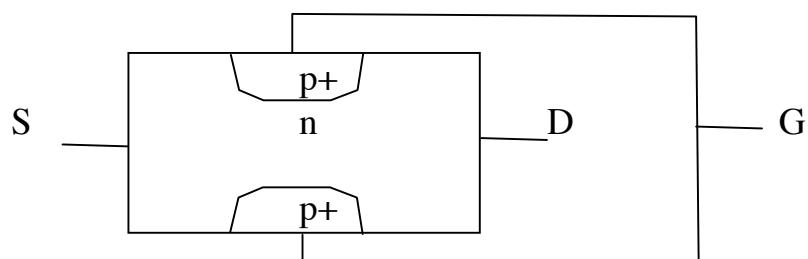
'Common base' circuit: by setting V_{EB} and I_E we control V_{CB} and I_C

'Common emitter' circuit: by setting V_{EB} and I_B we control V_{EC} and I_C

Digital logic: make the transistor act like a switch by running between saturation and cutoff.

Junction Field Effect Transistor

Apply a reverse voltage to gate G. This makes the depletion regions grow, alters the n-channel width and therefore alters its resistance, which changes the source-drain voltage (for constant current). This is a voltage amplifier and also a majority carrier device.



For $V_G = 0$ it is a linear resistor. However, applying V_{SD} causes pinch-off (i.e. the depletion regions touch) and no more current I_{SD} can flow.

Applying a negative V_G makes pinchoff occur earlier.

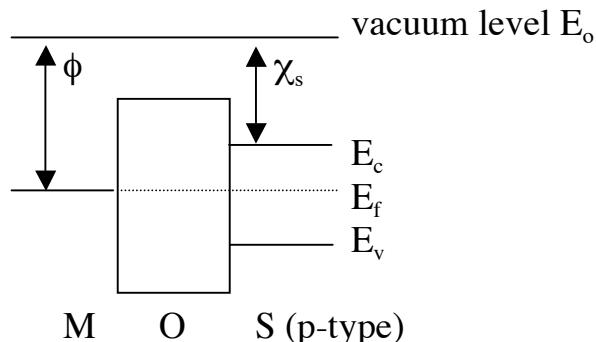
Beyond pinchoff, a small change in V_G causes a large change in V_{SD} .

For depletion width d, channel length t, and initial channel width t,

at pinchoff, $V_{SD, \text{sat}} = (eN_D t^2 / 8\epsilon_0 \epsilon_r) - (V_o + V_G)$

MOS devices (metal – oxide – semiconductor)

Ideal MOS bandstructure: where $\phi = \chi_s + (E_c - E_f)$, the Fermi level is flat without any band bending

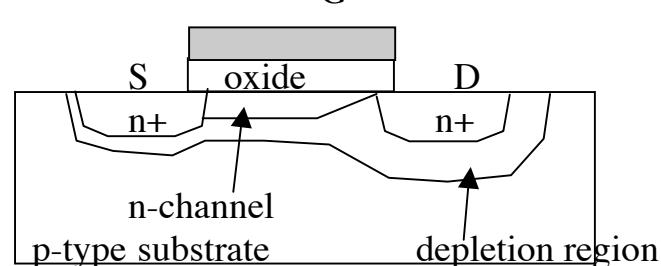


Negative voltage to the gate (i.e. to the metal) causes charge to accumulate (acts as a capacitor)

Small positive voltage to the gate causes charge to be depleted

Large positive voltage to the gate causes inversion: once all the mobile carriers are depleted, carriers of the opposite sign are attracted to the region of the semiconductor next to the oxide. This creates a channel of opposite type.

MOSFET



Source and substrate are grounded. Applying a large positive V_G creates an n-channel which conducts current between S and D.

If $V_D = 0$, channel has uniform width.

If $V_D > 0$, channel is thinner towards the D and may pinch-off; also the depletion width is larger.

At pinch-off, the current I_{SD} cannot increase any more.

Example application: a DRAM (dynamic random access memory) stores one bit in a cell consisting of a MOSFET plus a MOS-capacitor. Opening the channel of the MOSFET (by applying V_G) allows the capacitor to be charged to represent '1' or uncharged for '0'.

3.15: Transistors in ‘forward active’ mode

Common base circuit

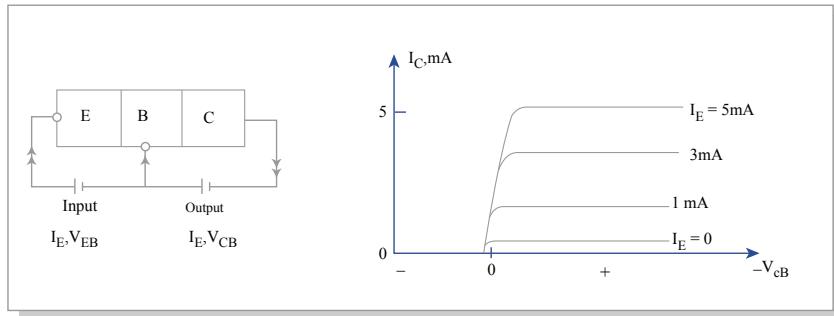


Figure by MIT OCW.

This is the easiest to visualize, though of limited usefulness. We use one power supply to put EB into forward bias (p side positive) and another one to put the BC into reverse bias (p side negative). The figure is drawn for a pnp transistor.

First, suppose we send a particular current I_E into the emitter (I_E is related of course to V_{EB}). This is shown as a family of curved lines, each corresponding to a different I_E . What happens as we vary V_{CB} ? We know that all of the I_E current will be collected by the CB junction, provided that CB is in reverse bias. So the output, I_C , is the same as the input I_E for any value of reverse bias on CB, shown as the positive side of zero on the horizontal axis. In fact, CB will collect all of I_E even if CB is unbiased, due to the built-in voltage. So the output does not start to drop until we start to put forward bias onto CB, which prevents the collection of I_E (shown at the negative side of zero on the horizontal axis).

Suppose now we fix the voltage V_{CB} (dotted vertical line). This puts a fixed reverse bias onto CB. CB collects all of I_E so the output I_C is just equal to the input I_E . This does not provide any amplification.

Common emitter circuit

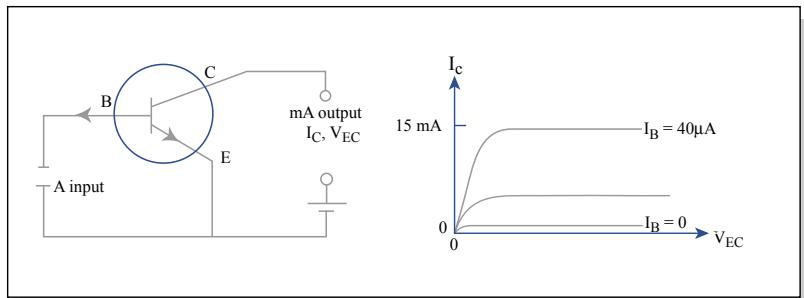


Figure by MIT OCW.

This is a more useful circuit because it gives amplification (the previous circuit did not amplify, because the output I_C was the same as the input I_E). Suppose we connect one input power supply that sends a current I_B between the base and the emitter. The polarity is chosen so that EB is in forward bias. We connect another power supply between E and C (this is chosen to make sure BC is in reverse bias).

The base current I_B is primarily composed of electrons that contribute to the forward current through EB. In the EB junction, there is a relation between the forward currents of holes and of electrons. These currents are in the ratios of the doping levels of the sides of the junction. If E is doped more heavily than B, a small electron current through EB implies that there is a large hole current through EB. This hole current is collected by the reverse biased BC junction, and flows through to make up the output current I_C . Therefore, a given I_B leads to a much larger I_C , with a gain of typically around 100 (that is, the ratio of doping levels in E and B). So to use this as a **current amplifier**, set a fixed V_{EC} (vertical dotted line), input I_B and you will produce an amplified output I_C .

Transistors

Reference: Handout 4; Pierret chapter 10, 15.1-2, 16.1-2 and 17.1.

Transistors are three-terminal devices that use a small voltage (or current) applied to one contact to modulate (i.e. control) a large voltage (or current) between the other two contacts.

An analogy is provided by the vacuum tube from the 1900s.
A small voltage applied to the 'grid' modulates a large current between the anode and cathode.

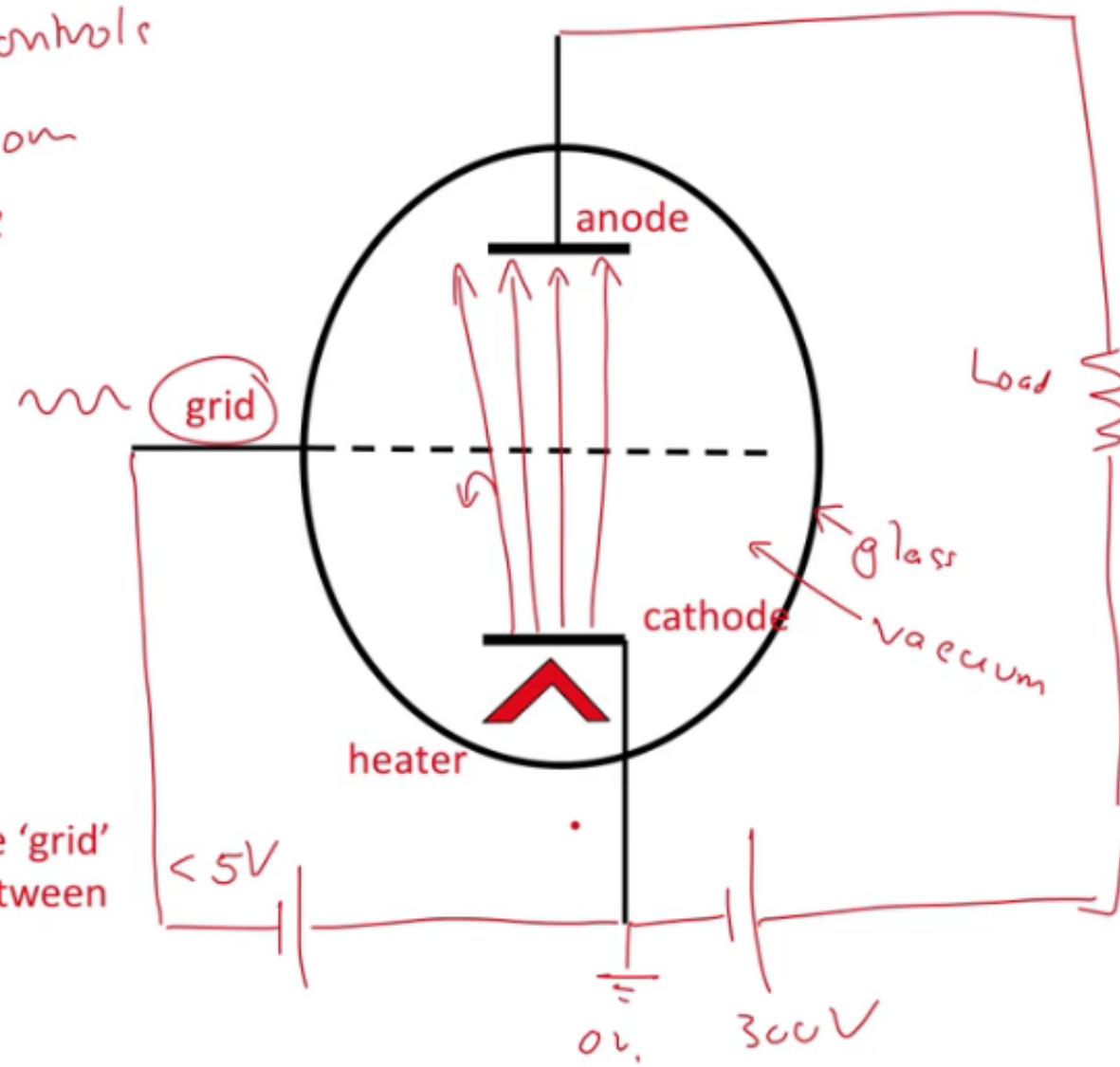
Image removed for
copyright considerations



Vacuum Tubes as Transistors

Small V_{grid} — controls

large current from
Anode to Cathode

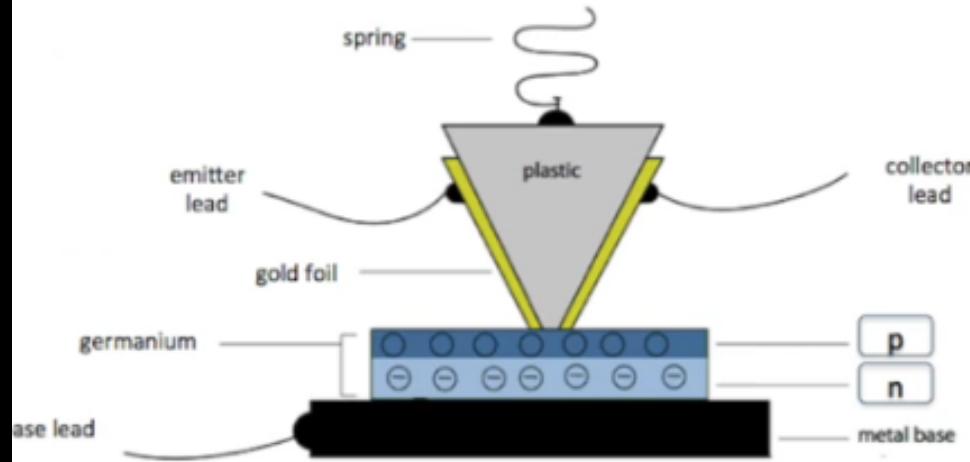


A small voltage applied to the 'grid' modulates a large current between the anode and cathode.



The first solid state transistor (1947, Bell Labs)

Image removed for
copyright considerations



Point contact transistor: Au contacts on n-type Ge

Current produced p-type layer: small change in current on one Au contact (emitter) caused large current between base and collector.

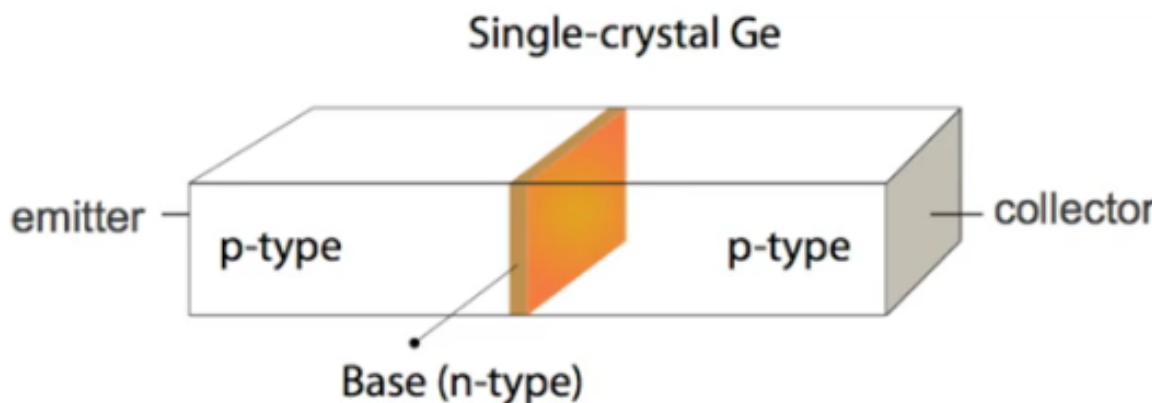
Bardeen, Brattain, Shockley, Bell Labs



The First Junction Transistor

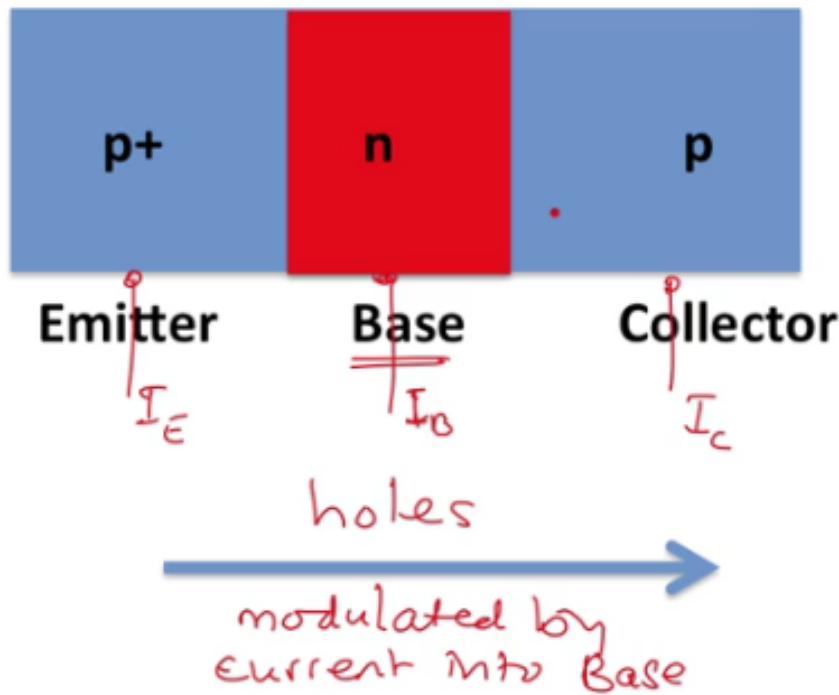
First transistor with diffused pn junctions by William Shockley
Bell Laboratories, Murray Hill, New Jersey (1949)

Image removed for
copyright considerations



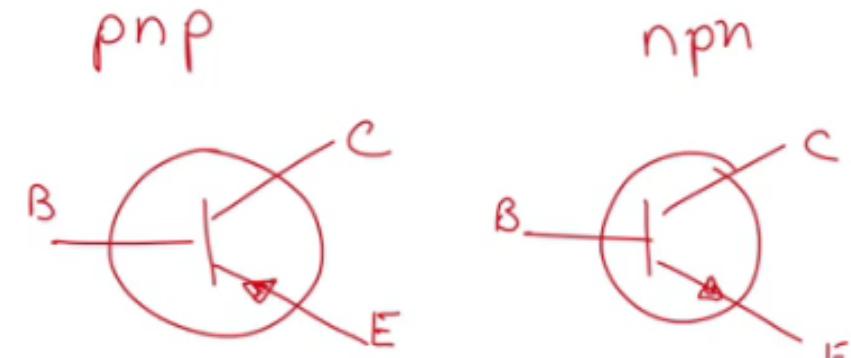
The bipolar junction transistor

BJT



E emits holes — EB jn is forward biased
C collect holes — BC reverse biased

Large hole current from E to C is
modulated by a small current at B.



is forward biased
reverse biased



The BJT

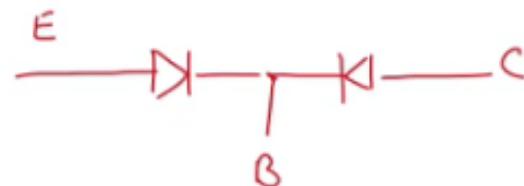
E emits carriers into B → EB fwd bias

C collects all the carriers from the depletion region → BC rev

Minority carrier device



The BJT differs from two diodes



How wide should the base be?

$w_B < L_p$ so that min. carriers survive
transit through base

$L_p \sim 40 \mu m$ in good Si

BJT gives current gain: $\underline{\underline{I_E}} \rightarrow \underline{\underline{\text{Small } I_B}} \rightarrow \underline{\underline{\text{Large } I_E \text{ (or } I_c\text{)}}}$

$$\underline{\underline{I_E \approx I_c \gg I_B}}$$

gain $\beta = \frac{I_c}{I_B}$



EB fwd bias

BC rev bias

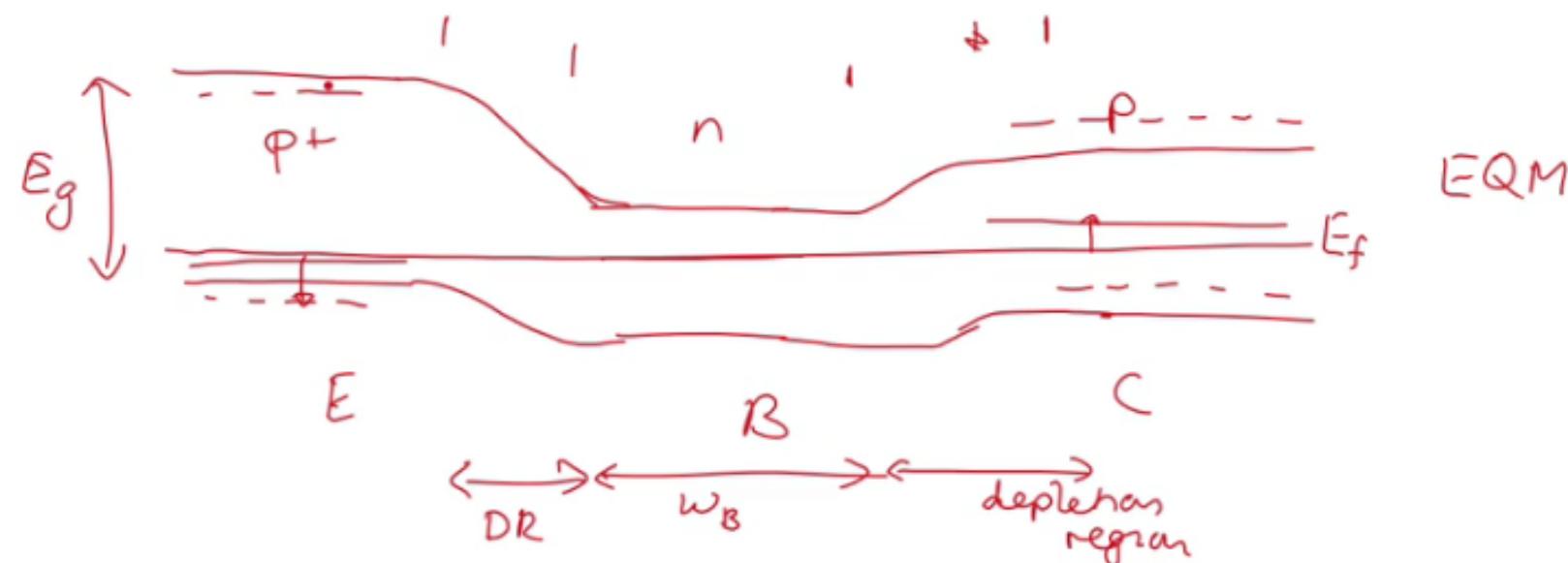
Forward Active Mode

Take $N_{A,E} > N_{D,B} > N_{A,C}$ and $w_B \ll L_p$

10^{19}

10^{17}

10^{15} cm^{-3}



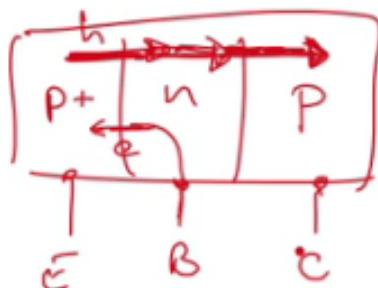
EQM



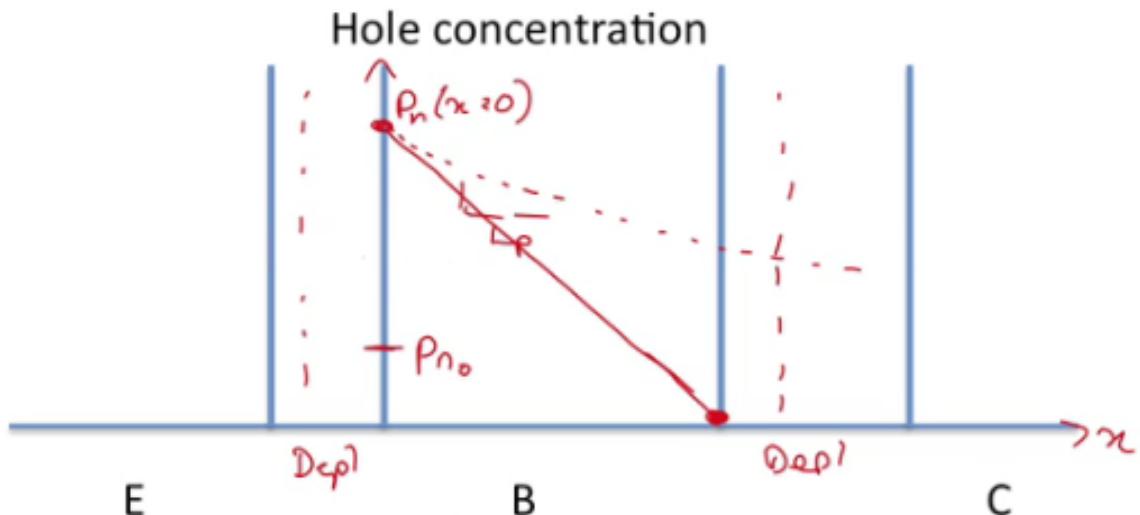
Forward Active Mode

1. EB is forward biased, *inject holes into B*
 $N_{A,E} > N_{D,B}$, current mostly holes
2. Holes travel through base if $w_B < L_p$
3. BC is reverse biased all holes are collected
current in C is large
4. Varying I_B affects I_E greatly.

I_B controls I_C or I_E current



The forward current I_E



$$J_{p, \text{base}} = \frac{e D_p \cdot p_n(x=0)}{\omega_B}$$

$$p_n(x=0) = p_{n0} \exp \frac{e V_{EB}}{kT}$$

\Rightarrow know hole flux through base $\approx I_c$



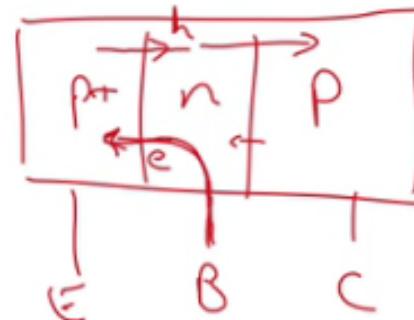
The base current I_B

$$w_B \ll L_p, I_E \approx I_c \gg I_B$$

Gain

$$\alpha = I_C/I_E \approx 1$$

$$\beta = I_C/I_B \approx 100 - 1000 = \alpha/(1-\alpha)$$



In the base current:

$I_B = I_n$ fwd current of e across EB junction

$$\frac{I_n}{I_p} = \frac{N_{D,B}}{N_{A,E}}$$

+ I_{nc} ~ reverse current of e coming from C
Small

+ eR_B — e consumed by recombining with
holes in base

Small if $L_p \gg w_B$



Forward Active Mode

You cannot have a large hole current from E to C without having a small electron current from B to E.

Their ratio is β , determined by doping

$$\frac{I_E}{I_B} = \frac{N_{A,E}}{N_{D,B}}$$

often $100 \sim 1000$

Transit time

$$\underline{\quad} \propto w_B$$

pnp vs npn

holes

electrons



Other operation modes

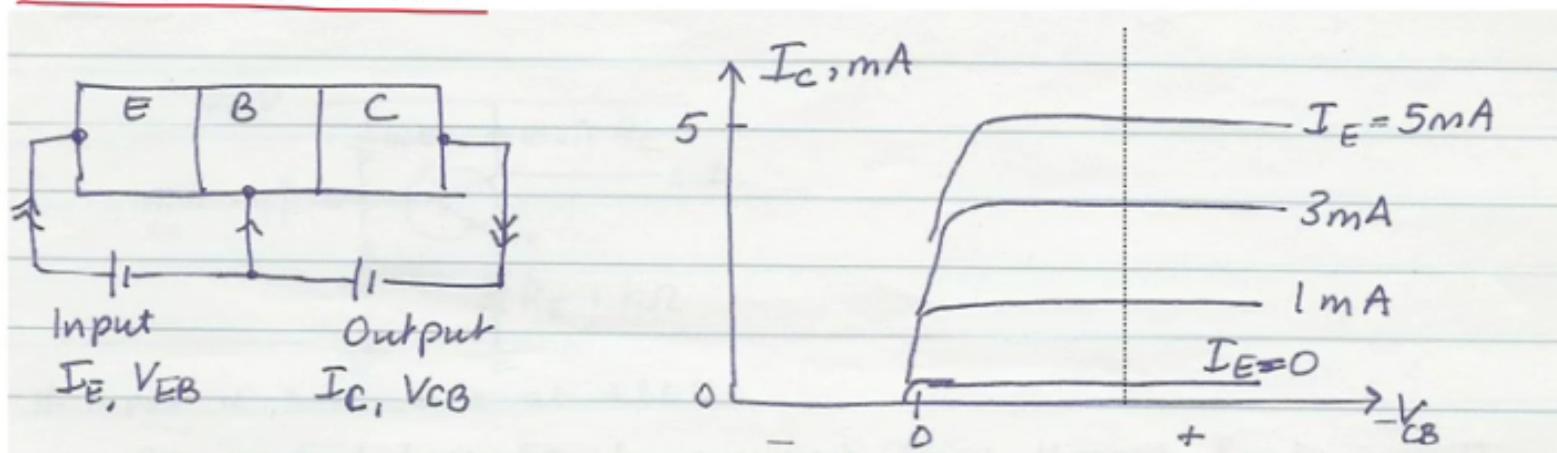
	EB jn	BC jn	
Forward active	F	R	gain 100-1000
Reverse active	R	F	gain $\ll 1$
Saturation	F	F	conductive
Cutoff	R	R	off - no current.

used for digital applications

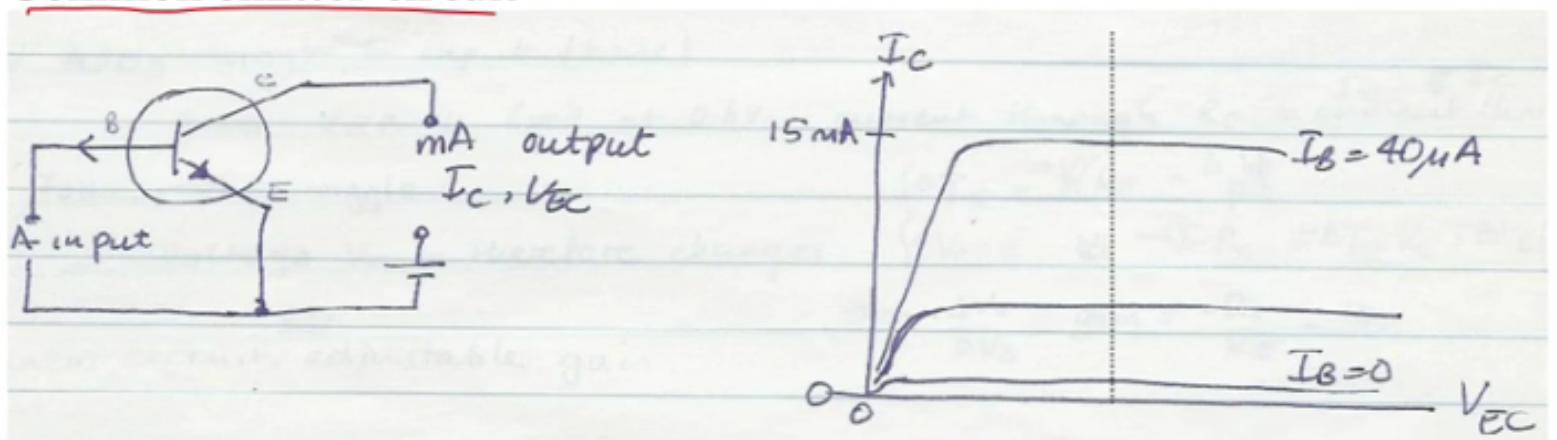


Some transistor circuits

Common base circuit



Common emitter circuit



Digital logic: act as switch

cutoff vs saturation



Summary – Bipolar junction transistor

A transistor is a 3-terminal device. A small current or voltage applied to one contact (base, or gate) controls a large current or voltage between the other two contacts (emitter and collector, or source and drain).

A **bipolar junction transistor** has emitter-base-collector regions which are doped n-p-n or p-n-p.

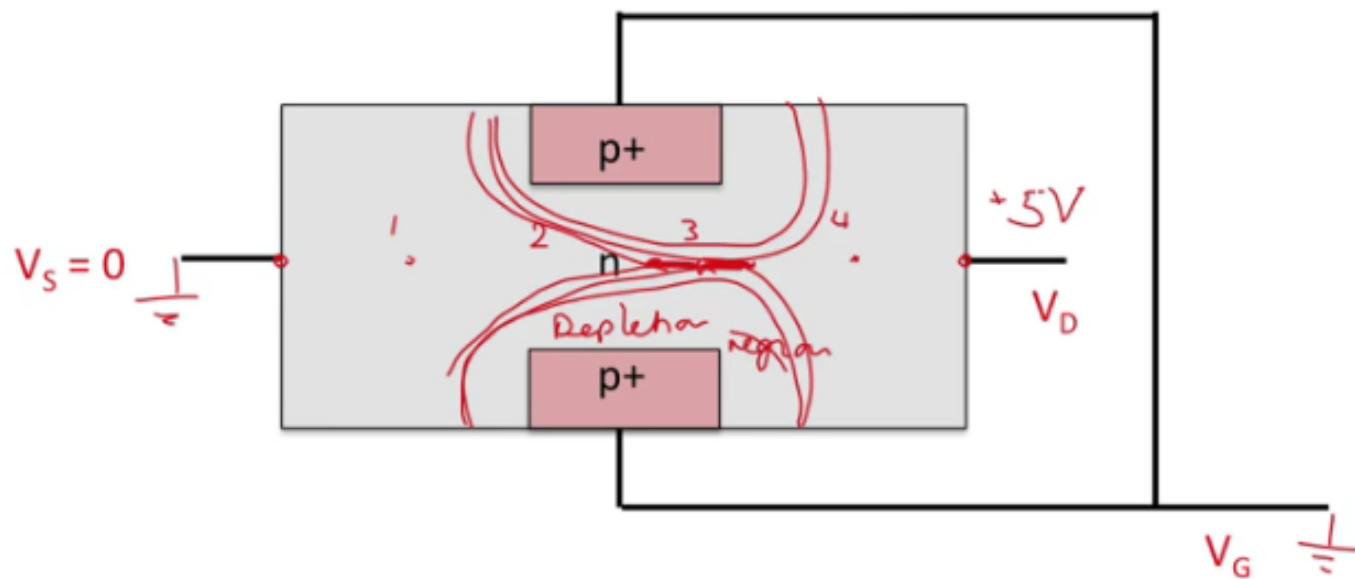
In forward active mode, EB is forward biased and injects minority carriers into B. BC is reverse biased and collects them. Small changes in V_B lead to large changes in I_C . The gain, $\beta = I_C/I_B \approx 100 - 1000$.

BJTs can also be operated between saturation and cutoff for digital logic.

3.15 Electrical, Optical, and Magnetic Materials and Devices

**Prof. Caroline A. Ross
Department of Materials Science and
Engineering, MIT**

Part 4: Field Effect Transistors

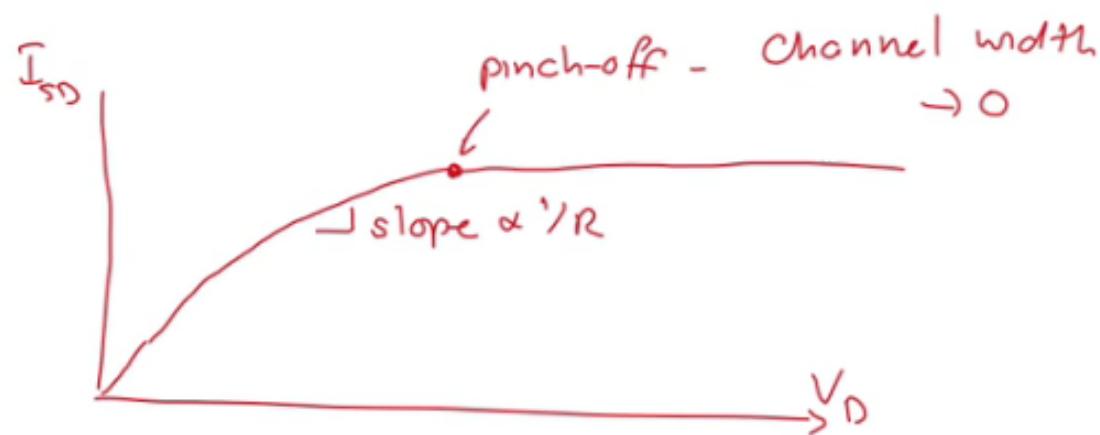


Qualitative JFET:

2) $V_G = 0$, V_D large

Increasing $V_D >$
pinchoff threshold

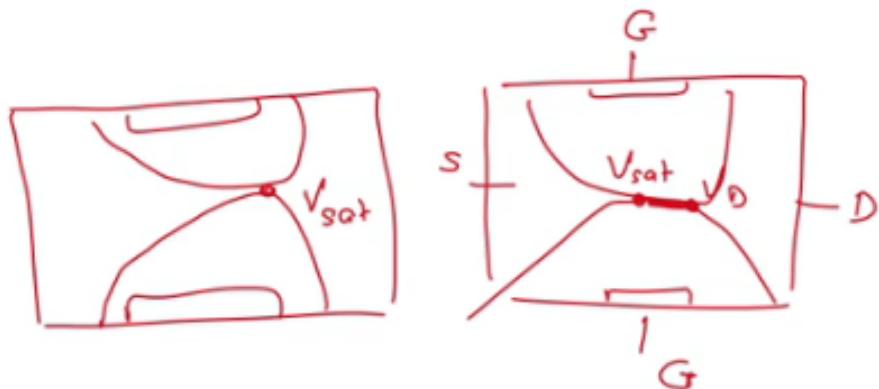
— just increase length of pinched-off channel!
Additional voltage extends the pinched off region



Pinch-off at $V_{SD,sat}$

Depletion regions touch. Current still flows through depletion region, though. *Carriers in depl. reg. are << N_D or N_A, but >> 0*

Practically all the voltage V_{SD} is dropped across the pinchoff region.

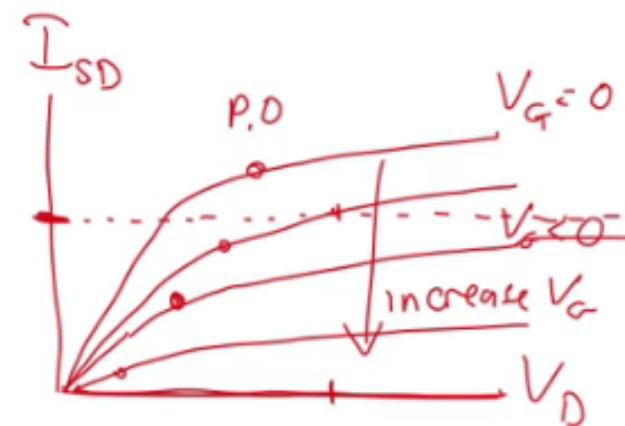
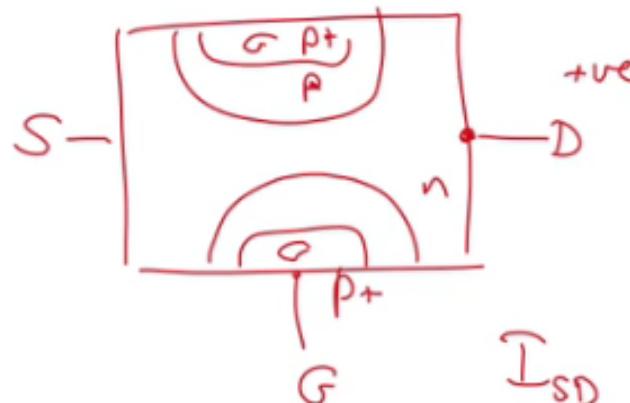


Qualitative JFET:

3) $V_G < 0$

This lowers $\underline{V_{SD,sat}}$

Already have pn junction in reverse bias,
so saturation occurs at lower V_D .



- > In the post-pinchoff regime (high V_D) a small change in V_G has a large change in $\underline{V_{SD}}$ for a fixed $\underline{I_{SD}}$.

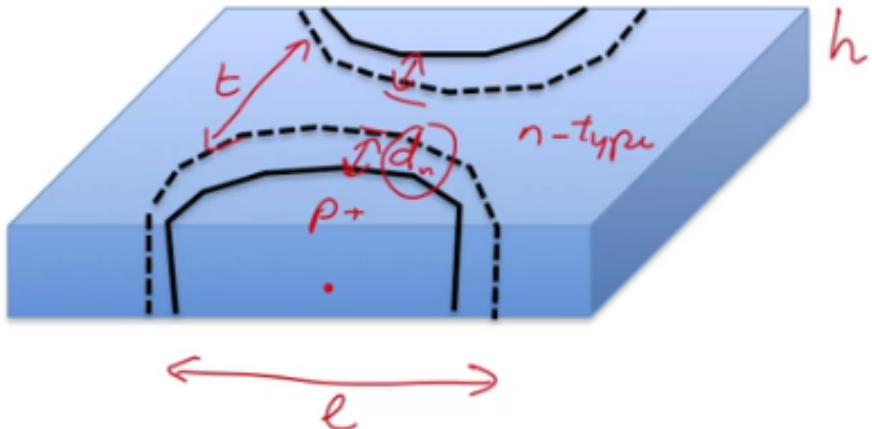
⇒ Voltage amplifier

Small $\Delta V_G \Rightarrow$ large ΔV_{SD} (if I_{SD} fixed)



Quantitative JFET:

$$I_{SD} - V_{SD}$$

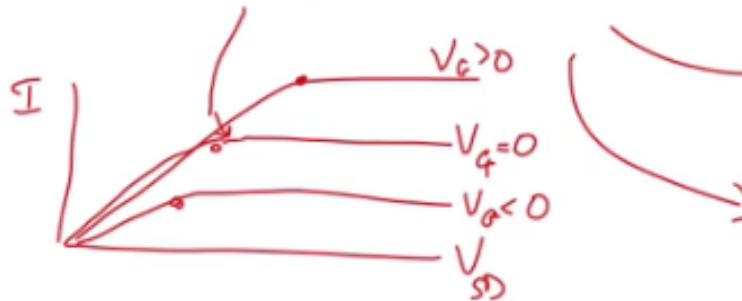


$$R = \frac{\rho l}{A} = \frac{l}{\sigma A}$$

$$= \frac{l}{e N_D \mu_n (t - 2d) h}$$

$$d_n = \left(\frac{2\epsilon_0 \epsilon_r (V_o + V_D)}{e N_D} \right)^{1/2}$$

At $V_{D,sat}$ $d = t/2$



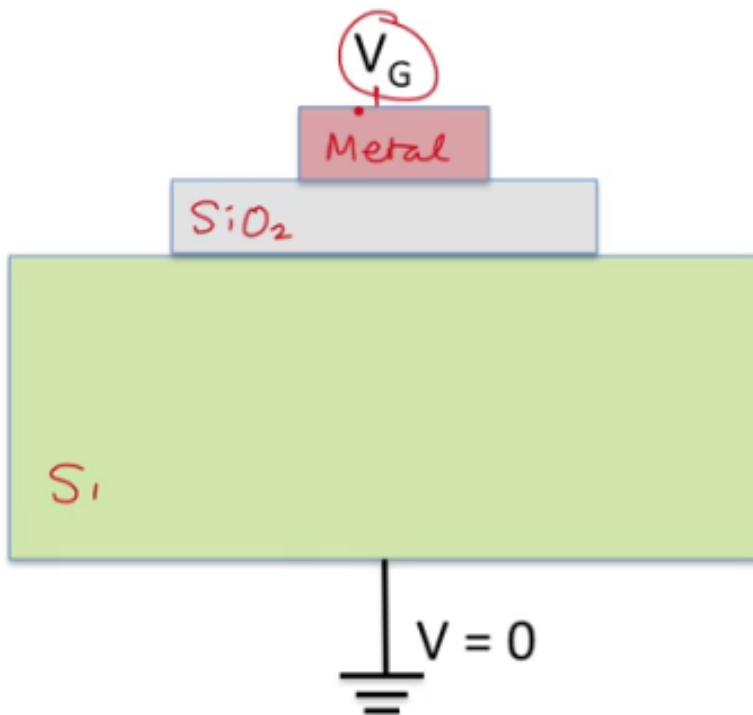
$$V_{D,sat} = \frac{e N_D t^2}{8 \epsilon_0 \epsilon_r} - V_o$$

$$V_{SD,sat} = (e N_D t^2 / 8 \epsilon_0 \epsilon_r) - (V_o + V_G)$$

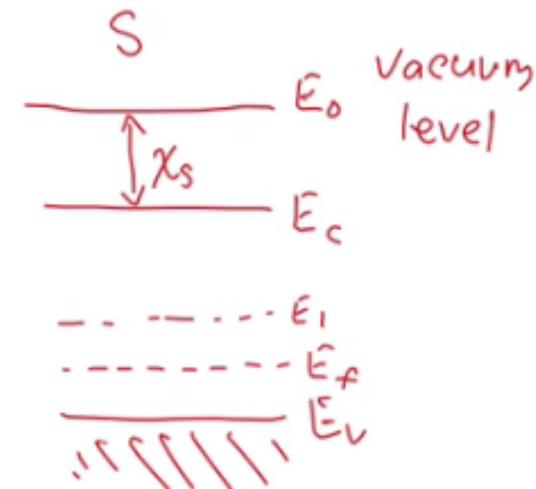
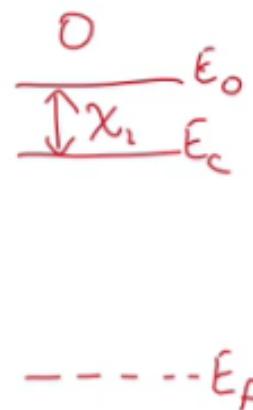
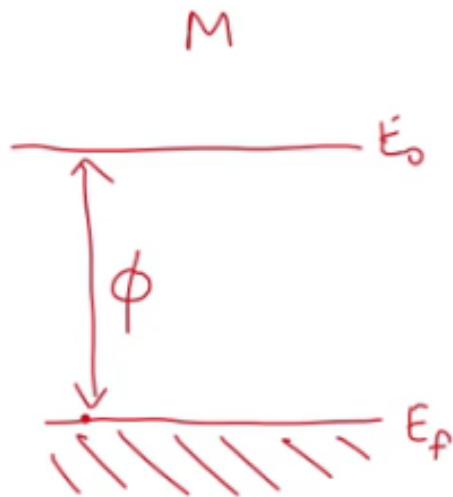
with V_G applied



Metal-Oxide-Semiconductor Devices



MOS Band structure



E_V

p-type

Insulator

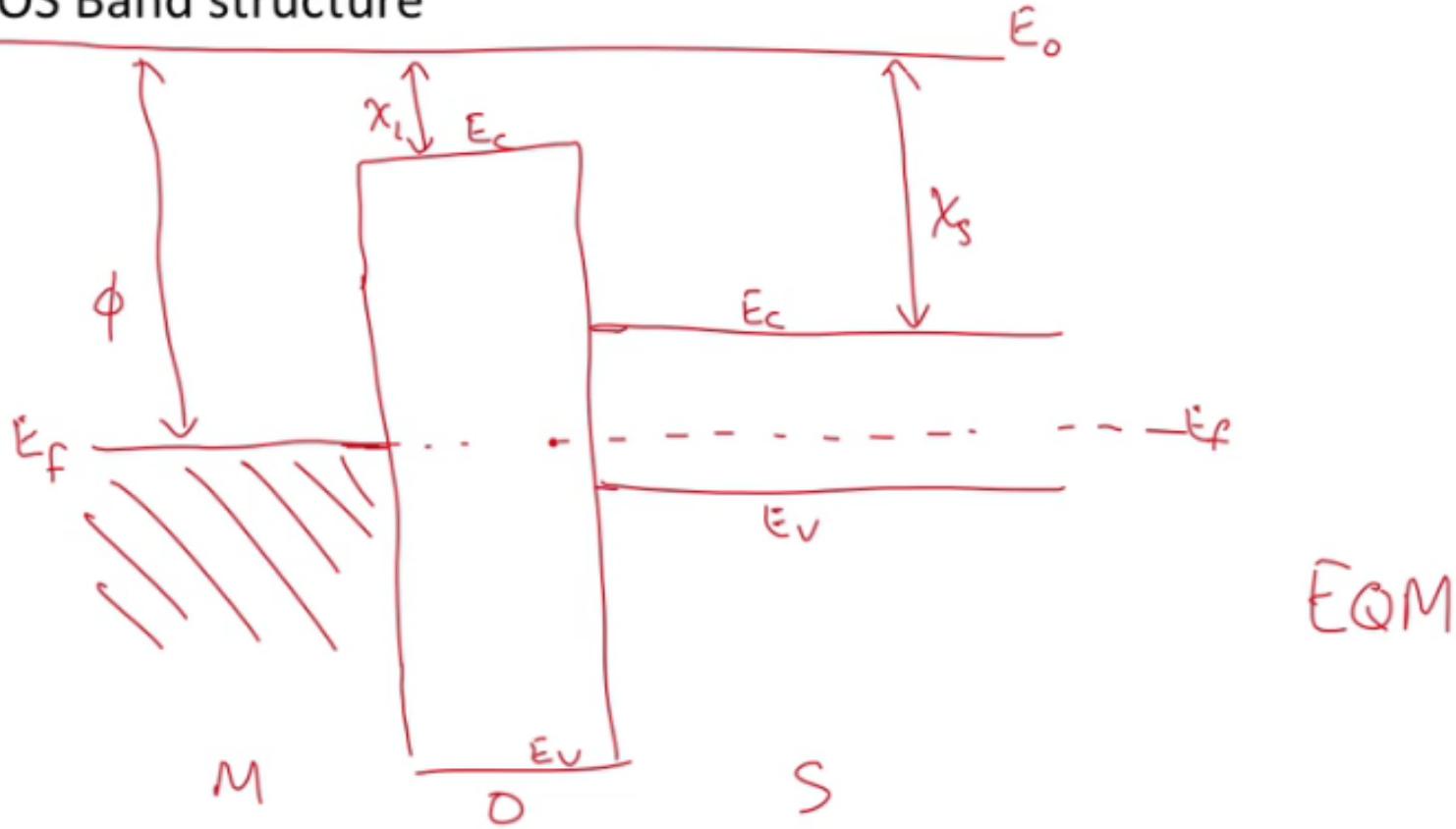
ϕ = work function
energy to take e-
from E_F to vacuum

χ = electron
affinity

Energy to remove e-
from E_c



MOS Band structure



Ideal MOS:

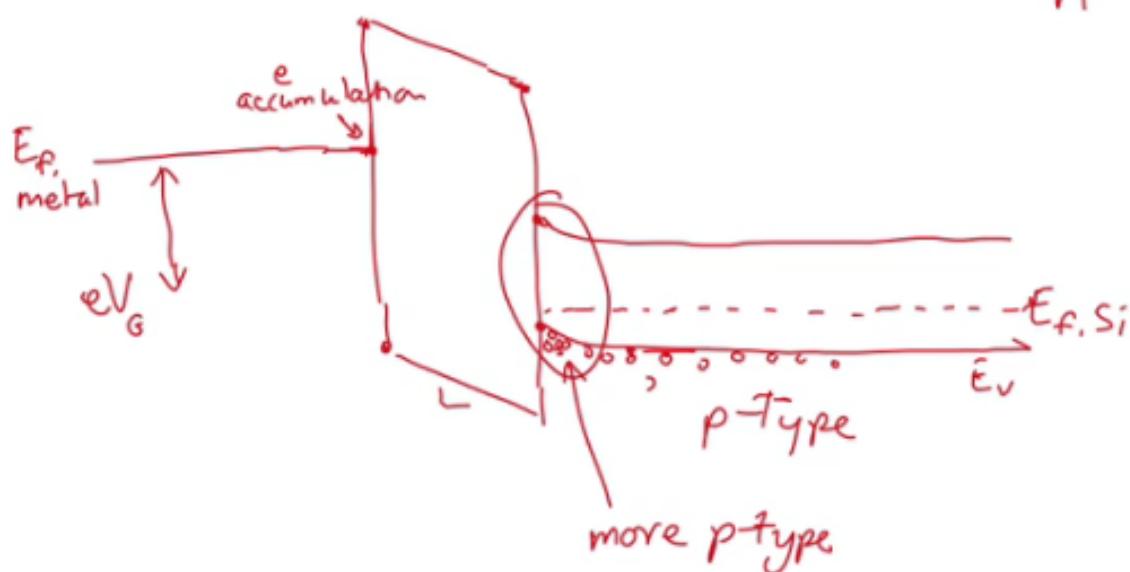
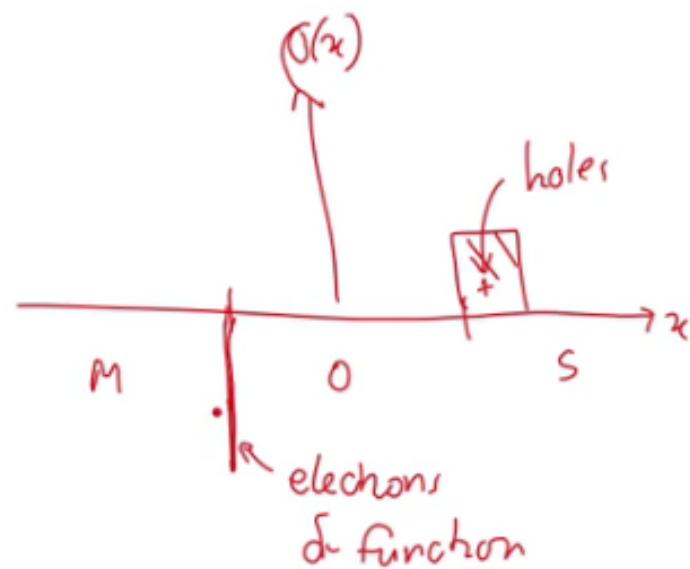
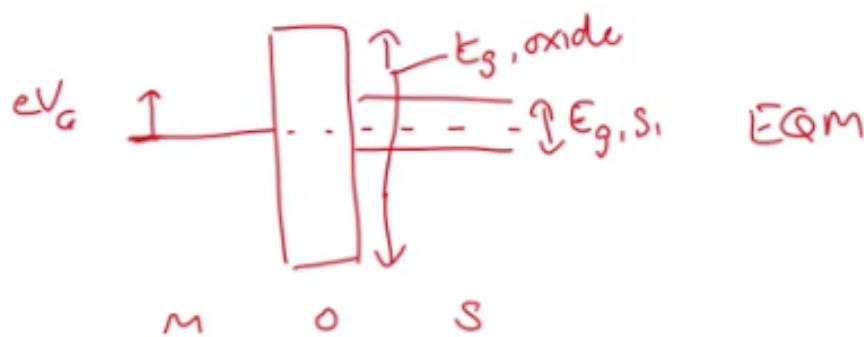
$$\phi = \chi + (E_C - E_F) \text{ for oxide and for Si}$$

E_O is flat when E_F is flat.

Non-ideal. E_O is not flat, get band bending at interfaces \Rightarrow charge accumulation

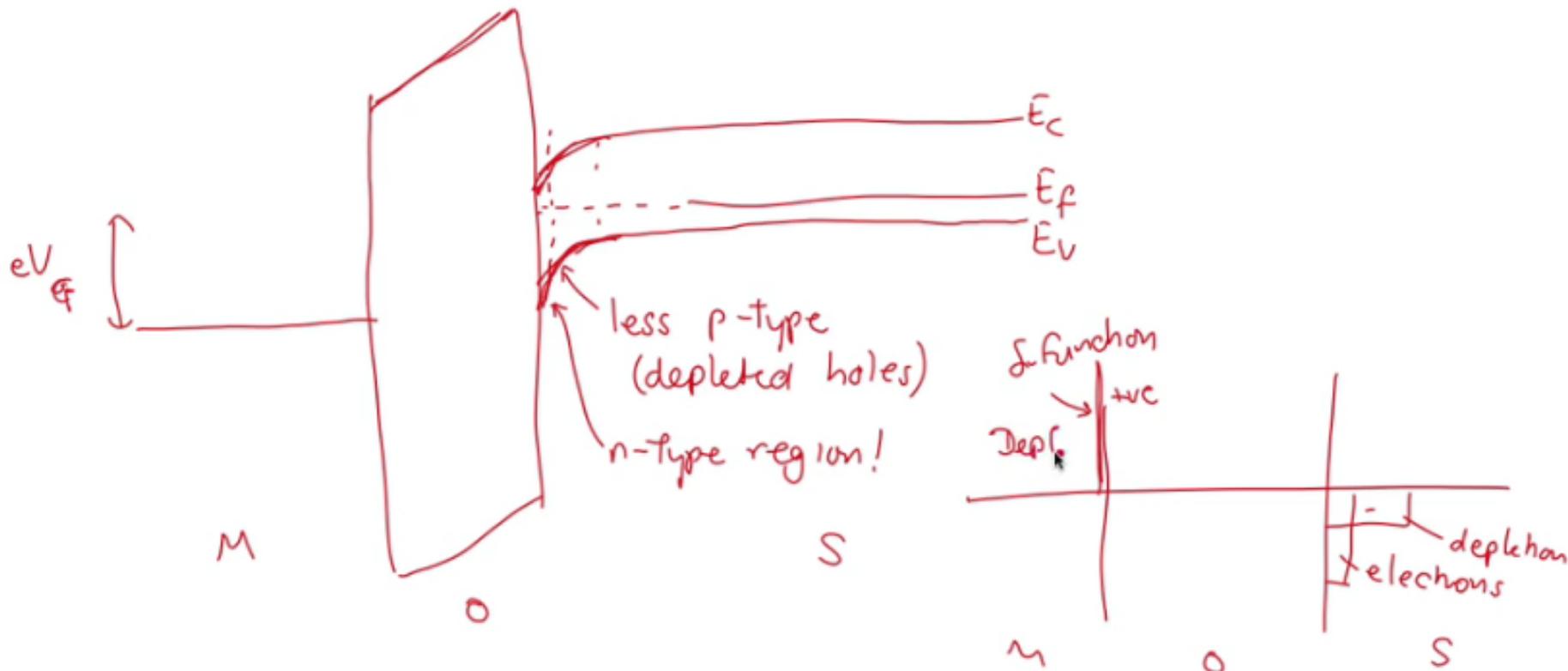


Apply negative voltage to G (metal), raising its energy levels



Accumulation (a capacitor)
of holes in s/c \Rightarrow interface

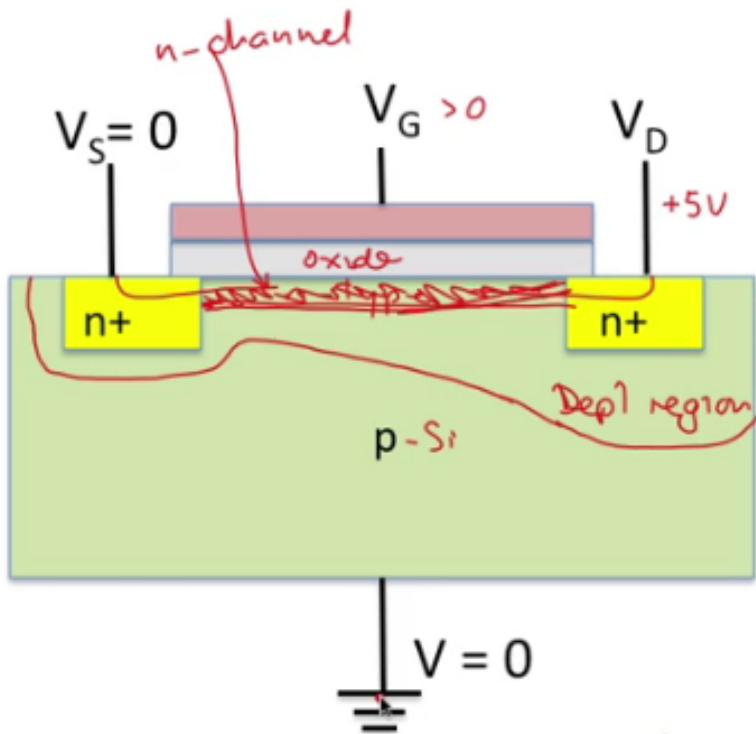
Apply positive voltage to G (metal), lowering its energy levels



Depletion then Inversion (we changed the majority carrier!!)



The MOS Field Effect Transistor



A large positive V_G forms an n-channel.

$$V_G > V_{\text{threshold}}$$

Then SD can conduct current.

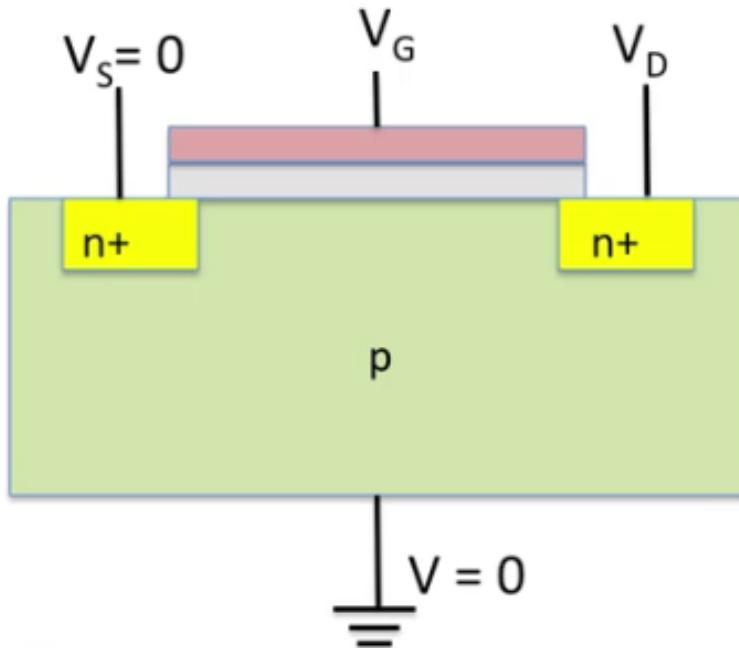
$V_D = 0$ channel has a uniform depth

$V_D > 0$ channel less deep near D side

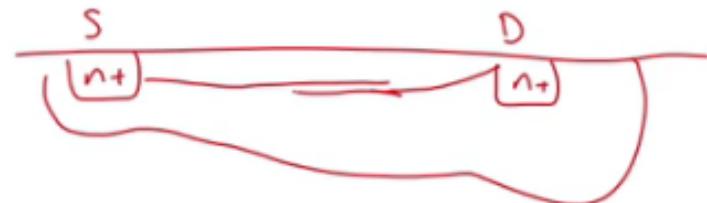
Channel will pinch off for high enough V_D .

Depletion region also wider near drain

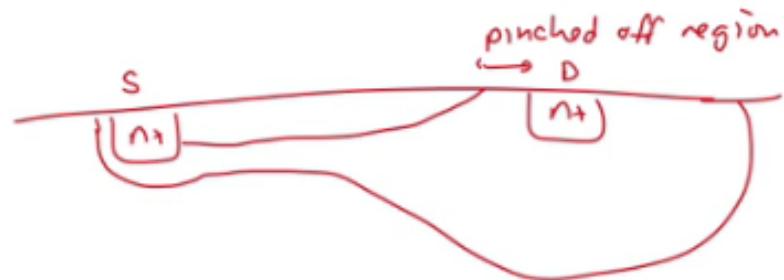
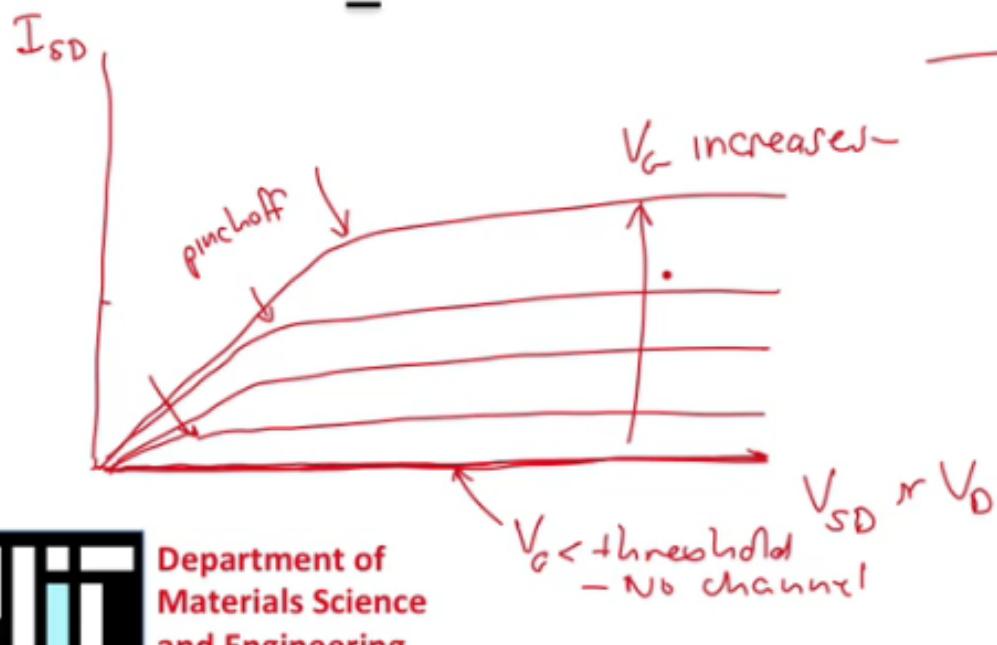




As V_D increases the n-channel pinches off.



Moderate V_D



Large V_D



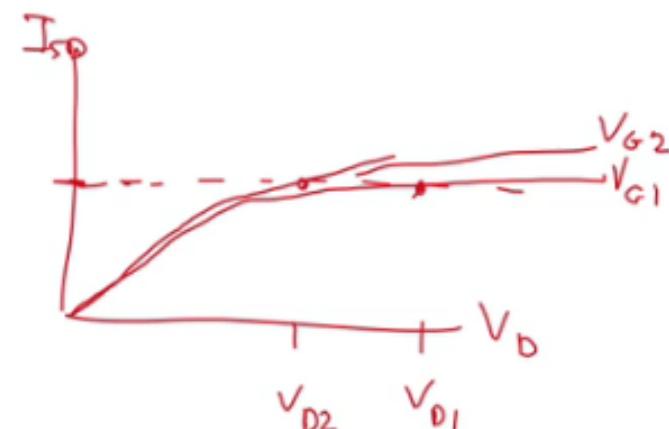
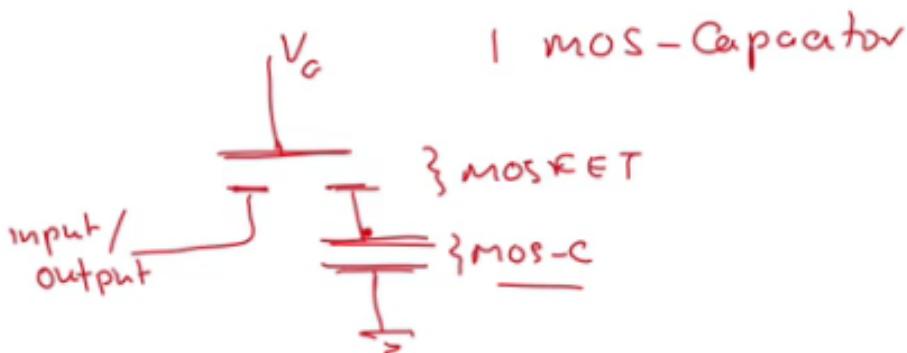
MOSFET applications

Channel can be n-type or p-type: NMOS, PMOS: CMOS logic uses both and is low power

\ complementary

- > On-off switch by changing V_G
above threshold - on
below - off
- > Voltage Amplifier with small constant I_{SD}

DRAM cell = 1 MOSFET +



New MOSFET designs with fins: FinFETs

Traditional Planar Transistor

22 nm Tri-Gate Transistor



Image removed for copyright
considerations

Intel Ivy Bridge Processor, 2012
1.4 billion transistors in 160 mm^2



Image removed for copyright
considerations



Department of
Materials Science
and Engineering

Summary – Field effect transistors and MOS

In a FET a voltage applied to a gate modulates the conductivity between a source and a drain contact. They are usually used to modulate voltage (whereas BJTs modulate current) and are majority carrier devices (BJTs are minority carrier devices).

In the JFET, V_G alters the width of a channel connecting S and D, by changing the depletion width. A large enough V_G will pinch off the channel – limiting how much current flows from S to D.

MOSFETs rely on the MOS structure, in which a voltage applied to the metal changes the carrier concentration in the region of Si closest to the oxide layer, and can even invert it to create a channel of opposite majority carrier type. In a MOSFET the channel connects S and D. The combination of V_G and V_D determines whether the channel pinches off and how much SD current flows. MOSFETs can be used as an on-off switch and are the essential components of microprocessors.

