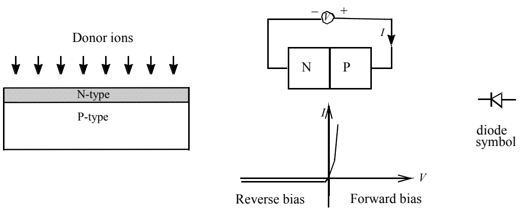
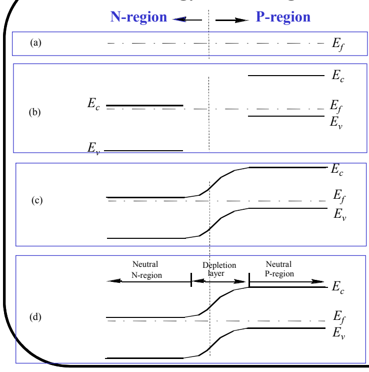
Chapter 4 PN Junctions

4.1 Building Blocks of the PN Junction Theory



PN junction is present in perhaps every semiconductor device.

4.1.1 Energy Band Diagram of a PN Junction



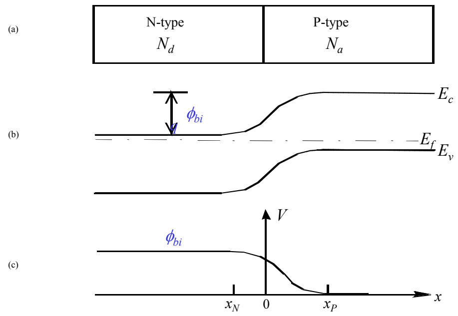
E_f is constant at equilibrium

E_c and E_v are known relative to E_f

E_c and E_v are smooth, the exact shape to be determined.

A depletion layer exists at the PN junction where $n \approx 0$ and $p \approx 0$.

4.1.2 Built-in Potential



Can the built-in potential be measured with a voltmeter?

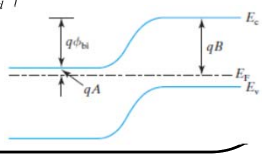
4.1.2 Built-in Potential

N-region $\quad n = N_d = N_c e^{-qA/kT} \Rightarrow A = \frac{kT}{q}\ln\frac{N_c}{N_d}$

P-region $\quad n = \frac{n_i^2}{N_a} = N_c e^{-qB/kT} \Rightarrow B = \frac{kT}{q}\ln\frac{N_c N_a}{n_i^2}$

$$\phi_{bi} = B - A = \frac{kT}{q}\left(\ln\frac{N_c N_a}{n_i^2} - \ln\frac{N_c}{N_d}\right)$$

$$\boxed{\phi_{bi} = \frac{kT}{q}\ln\frac{N_d N_a}{n_i^2}}$$



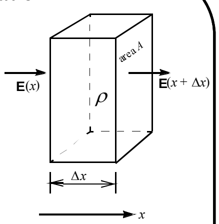
4.1.3 Poisson's Equation

Gauss's Law:

$$\varepsilon_s \mathcal{E}(x+\Delta x)A - \varepsilon_s \mathcal{E}(x)A = \rho \Delta x A$$

ε_s: permittivity (~$12\varepsilon_o$ for Si)

ρ: charge density (C/cm^3)



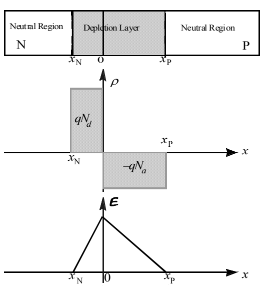
$$\frac{\mathcal{E}(x+\Delta x) - \mathcal{E}(x)}{\Delta x} = \frac{\rho}{\varepsilon_s}$$

$$\frac{d\mathcal{E}}{dx} = \frac{\rho}{\varepsilon_s}$$

$$\boxed{\frac{d^2 V}{dx^2} = -\frac{d\mathcal{E}}{dx} = -\frac{\rho}{\varepsilon_s}}$$ ***Poisson's equation***

4.2 Depletion-Layer Model

4.2.1 Field and Potential in the Depletion Layer



On the ***P-side*** of the depletion layer, $\rho = -qN_a$

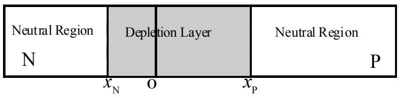
$$\frac{d\mathcal{E}}{dx} = -\frac{qN_a}{\varepsilon_s}$$

$$\mathcal{E}(x) = -\frac{qN_a}{\varepsilon_s}x + C_1 = \boxed{\frac{qN_a}{\varepsilon_s}(x_P - x)}$$

On the ***N-side***, $\rho = qN_d$

$$\boxed{\mathcal{E}(x) = \frac{qN_d}{\varepsilon_s}(x - x_N)}$$

4.2.1 Field and Potential in the Depletion Layer



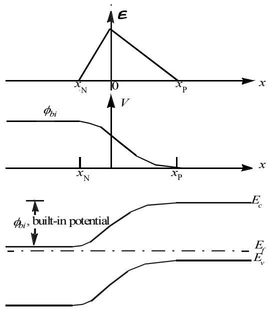
The electric field is continuous at $x = 0$.

$$\boxed{N_a|x_P| = N_d|x_N|}$$

Which side of the junction is depleted more?

A one-sided junction is called a ***N⁺P junction*** or ***P⁺N junction***

4.2.1 Field and Potential in the Depletion Layer



On the P-side,

$$V(x) = \frac{qN_a}{2\varepsilon_s}(x_P - x)^2$$

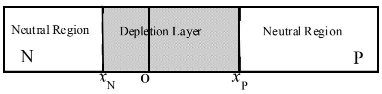
Arbitrarily choose the voltage at $x = x_P$ as $V = 0$.

On the N-side,

$$V(x) = D - \frac{qN_d}{2\varepsilon_s}(x - x_N)^2$$

$$= \phi_{bi} - \frac{qN_d}{2\varepsilon_s}(x - x_N)^2$$

4.2.2 Depletion-Layer Width



V is continuous at $x = 0$ ⟶

$$x_P - x_N = W_{dep} = \sqrt{\frac{2\varepsilon_s \phi_{bi}}{q}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)}$$

If $N_a >> N_d$, as in a P⁺N junction,

$$\boxed{W_{dep} = \sqrt{\frac{2\varepsilon_s \phi_{bi}}{qN_d}} \approx |x_N|} \qquad \boxed{|x_P| = |x_N| N_d / N_a \cong 0}$$

What about a N⁺P junction?

$$W_{dep} = \sqrt{2\varepsilon_s \phi_{bi}/qN} \quad \text{where} \quad \frac{1}{N} = \frac{1}{N_d} + \frac{1}{N_a} \approx \frac{1}{\text{lighter dopant density}}$$

EXAMPLE: *A P⁺N junction has $N_a = 10^{20}$ cm⁻³ and $N_d = 10^{17}$cm⁻³. What is a) its built in potential, b) W_{dep} , c) x_N, and d) x_P ?*

Solution:

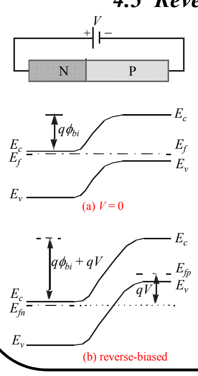
a) $$\phi_{bi} = \frac{kT}{q} \ln \frac{N_d N_a}{n_i^2} = 0.026\,\text{V} \ln \frac{10^{20} \times 10^{17}\,\text{cm}^{-6}}{10^{20}\,\text{cm}^{-6}} \approx 1\,\text{V}$$

b) $$W_{dep} \approx \sqrt{\frac{2\varepsilon_s \phi_{bi}}{qN_d}} = \left(\frac{2 \times 12 \times 8.85 \times 10^{-14} \times 1}{1.6 \times 10^{-19} \times 10^{17}}\right)^{1/2} = 0.12\,\mu\text{m}$$

c) $|x_N| \approx W_{dep} = 0.12\,\mu\text{m}$

d) $|x_P| = |x_N| N_d / N_a = 1.2 \times 10^{-4}\,\mu\text{m} = 1.2\,\text{Å} \approx 0$

4.3 Reverse-Biased PN Junction

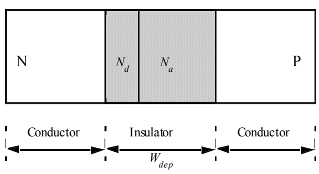


$$\boxed{W_{dep} = \sqrt{\frac{2\varepsilon_s(\phi_{bi} + |V_r|)}{qN}} = \sqrt{\frac{2\varepsilon_s \cdot \text{potential barrier}}{qN}}}$$

$$\frac{1}{N} = \frac{1}{N_d} + \frac{1}{N_a} \approx \frac{1}{\text{lighter dopant density}}$$

- ***Does the depletion layer widen or shrink with increasing reverse bias?***

4.4 Capacitance-Voltage Characteristics



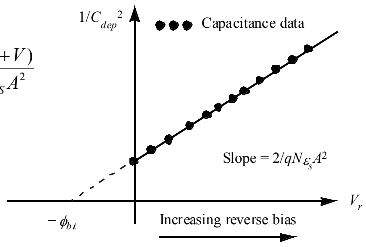
Reverse biased PN junction is a capacitor.

$$\boxed{C_{dep} = A\frac{\varepsilon_s}{W_{dep}}}$$

- Is C_{dep} a good thing?
- How to minimize junction capacitance?

4.4 Capacitance-Voltage Characteristics

$$\frac{1}{C_{dep}^{\ 2}} = \frac{W_{dep}^{\ 2}}{A^2\varepsilon_s^{\ 2}} = \frac{2(\phi_{bi}+V)}{qN\varepsilon_s A^2}$$



- From this C-V data can N_a and N_d be determined?

Slide 4-88

EXAMPLE: *If the slope of the line in the previous slide is $2x10^{23}$ F^{-2} V^{-1}, the intercept is 0.84V, and A is 1 μm^2, find the lighter and heavier doping concentrations N_l and N_h.*

Solution:

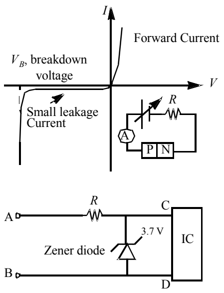
$$N_l = 2/(slope \times q\varepsilon_s A^2)$$
$$= 2/(2\times10^{23}\times1.6\times10^{-19}\times12\times8.85\times10^{-14}\times10^{-8}\,\text{cm}^2)$$
$$= 6\times10^{15}\ \text{cm}^{-3}$$

$$\phi_{bi} = \frac{kT}{q}\ln\frac{N_h N_l}{n_i^{\ 2}} \Rightarrow N_h = \frac{n_i^{\ 2}}{N_l}e^{\frac{q\phi_{bi}}{kT}} = \frac{10^{20}}{6\times10^{15}}e^{\frac{0.84}{0.026}} = 1.8\times10^{18}\ \text{cm}^{-3}$$

- *Is this an accurate way to determine N_l? N_h?*

Slide 4-89

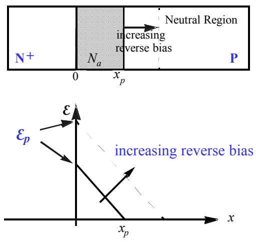
4.5 Junction Breakdown



A ***Zener diode*** is designed to operate in the breakdown mode.

Slide 4-90

4.5.1 Peak Electric Field

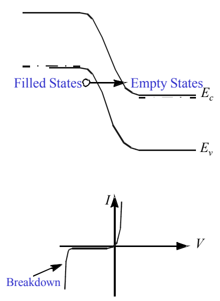


$$\mathcal{E}_p = \mathcal{E}(0) = \left[\frac{2qN}{\varepsilon_s}(\phi_{bi}+|V_r|)\right]^{1/2}$$

$$V_B = \frac{\varepsilon_s \mathcal{E}_{crit}^2}{2qN} - \phi_{bi}$$

Slide 4-91

4.5.2 Tunneling Breakdown



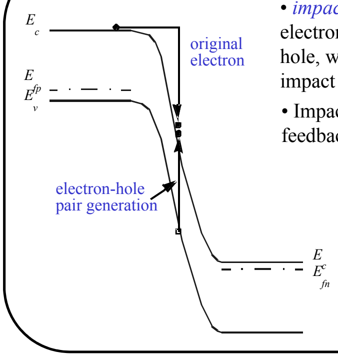
Dominant if both sides of a junction are very heavily doped.

$$J = Ge^{-H/\mathcal{E}_p}$$

$$\mathcal{E}_p = \mathcal{E}_{crit} \approx 10^6\ \text{V/cm}$$

Slide 4-92

4.5.3 Avalanche Breakdown



- *impact ionization*: an energetic electron generating electron and hole, which can also cause impact ionization.
- Impact ionization + positive feedback→*avalanche breakdown*

$$V_B = \frac{\varepsilon_s \mathcal{E}_{crit}^2}{2qN}$$

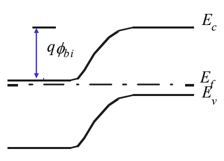
$$\boxed{V_B \propto \frac{1}{N}} = \frac{1}{N_a} + \frac{1}{N_d}$$

Slide 4-93

4.6 Forward Bias – Carrier Injection

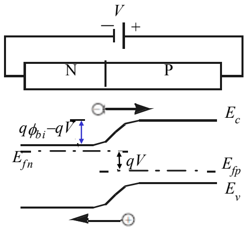
V=0

I=0



Drift and diffusion cancel out

Forward biased

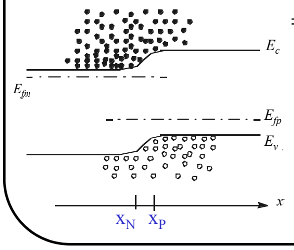


Minority carrier injection

4.6 Forward Bias – Quasi-equilibrium Boundary Condition

$$n(x_P) = N_c e^{-(E_c-E_{fn})/kT} = N_c e^{-(E_c-E_{fp})/kT} e^{(E_{fn}-E_{fp})/kT}$$

$$= n_{P0} e^{(E_{fn}-E_{fp})/kT} = n_{P0} e^{qV/kT}$$



- The minority carrier densities are raised by $e^{qV/kT}$
- Which side gets more carrier injection?

4.6 Carrier Injection Under Forward Bias– Quasi-equilibrium Boundary Condition

$$n(x_P) = n_{P0} e^{qV/kT} = \frac{n_i^2}{N_a} e^{qV/kT}$$

$$p(x_N) = p_{N0} e^{qV/kT} = \frac{n_i^2}{N_d} e^{qV/kT}$$

$$n'(x_P) \equiv n(x_P) - n_{P0} = n_{P0}(e^{qV/kT}-1)$$

$$p'(x_N) \equiv p(x_N) - p_{N0} = p_{N0}(e^{qV/kT}-1)$$

EXAMPLE: Carrier Injection

A PN junction has $N_a=10^{19}$cm^{-3} and $N_d=10^{16}$cm^{-3}. The applied voltage is 0.6 V.

***Question**: What are the minority carrier concentrations at the depletion-region edges?*

***Solution**:* $n(x_P) = n_{P0} e^{qV/kT} = 10 \times e^{0.6/0.026} = 10^{11}\ \text{cm}^{-3}$

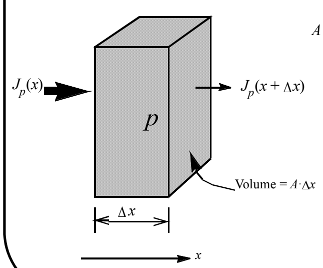
$p(x_N) = p_{N0} e^{qV/kT} = 10^4 \times e^{0.6/0.026} = 10^{14}\ \text{cm}^{-3}$

***Question**: What are the excess minority carrier concentrations?*

***Solution**:* $n'(x_P) = n(x_P) - n_{P0} = 10^{11} - 10 = 10^{11}\ \text{cm}^{-3}$

$p'(x_N) = p(x_N) - p_{N0} = 10^{14} - 10^4 = 10^{14}\ \text{cm}^{-3}$

4.7 Current Continuity Equation



$$A \cdot \frac{J_p(x)}{q} = A \cdot \frac{J_p(x+\Delta x)}{q} + A \cdot \Delta x \cdot \frac{p'}{\tau}$$

$$-\frac{J_p(x+\Delta x) - J_p(x)}{\Delta x} = q\frac{p'}{\tau}$$

$$-\frac{dJ_p}{dx} = q\frac{p'}{\tau}$$

4.7 Current Continuity Equation

$$-\frac{dJ_p}{dx} = q\frac{p'}{\tau}$$

Minority drift current is negligible;
∴ $J_p = -qD_p dp/dx$

$$qD_p \frac{d^2 p}{dx^2} = q\frac{p'}{\tau_p}$$

$$\frac{d^2 p'}{dx^2} = \frac{p'}{D_p \tau_p} = \frac{p'}{L_p^2}$$

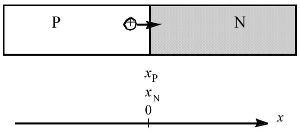
$$\frac{d^2 n'}{dx^2} = \frac{n'}{L_n^2}$$

L_p and L_n are the diffusion lengths

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

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

4.8 Forward Biased Junction-- Excess Carriers



$$\frac{d^2p'}{dx^2} = \frac{p'}{L_p^{\,2}}$$

$$p'(\infty) = 0$$

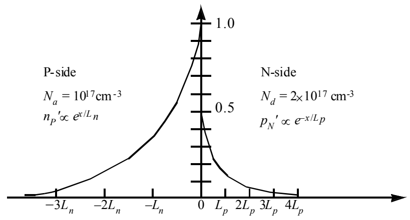
$$p'(x_N) = p_{N0}(e^{qV/kT} - 1)$$

$$p'(x) = Ae^{x/L_p} + Be^{-x/L_p}$$

$$p'(x) = p_{N0}(e^{qV/kT} - 1)e^{-(x-x_N)/L_p}, \quad x > x_N$$

Slide 4-100

4.8 Excess Carrier Distributions



$$p'(x) = p_{N0}(e^{qV/kT} - 1)e^{-(x-x_N)/L_p}, \quad x > x_N$$

$$n'(x) = n_{P0}(e^{qV/kT} - 1)e^{(x-x_P)/L_n}, \quad x < x_P$$

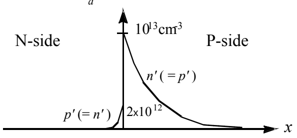
Slide 4-101

EXAMPLE: Carrier Distribution in Forward-biased PN Diode

N-type	P-type
$N_d = 5{\times}10^{17}$ cm^{-3} $D_p = 12$ cm^2/s $\tau_p = 1$ µs	$N_a = 10^{17}$ cm^{-3} $D_n = 36.4$ cm^2/s $\tau_n = 2$ µs

- *Sketch n'(x) on the P-side.*

$$n'(x_P) = n_{P0}(e^{qV/kT} - 1) = \frac{n_i^{\,2}}{N_a}(e^{qV/kT} - 1) = \frac{10^{20}}{10^{17}}e^{0.6/0.026} = 10^{13}\ \text{cm}^{-3}$$



Slide 4-102

EXAMPLE: Carrier Distribution in Forward-biased PN Diode

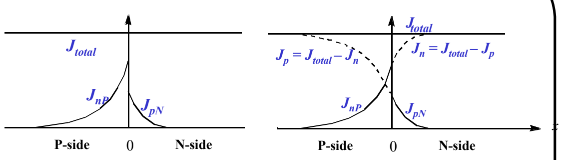
- *How does L_n compare with a typical device size?*

$$L_n = \sqrt{D_n\tau_n} = \sqrt{36\times2\times10^{-6}} = 85\ \mu\text{m}$$

- *What is p'(x) on the P-side?*

Slide 4-103

4.9 PN Diode I-V Characteristics



$$J_{pN} = -qD_p\frac{dp'(x)}{dx} = q\frac{D_p}{L_p}p_{N0}(e^{qV/kT} - 1)e^{-(x-x_N)/L_p}$$

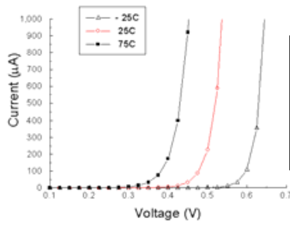
$$J_{nP} = qD_n\frac{dn'(x)}{dx} = q\frac{D_n}{L_n}n_{P0}(e^{qV/kT} - 1)e^{(x-x_P)/L_n}$$

$$\text{Total current} = J_{pN}(x_N) + J_{nP}(x_P) = \left(q\frac{D_p}{L_p}p_{N0} + q\frac{D_n}{L_n}n_{P0}\right)(e^{qV/kT} - 1)$$

$= J$ at all x

Slide 4-104

The PN Junction as a Temperature Sensor



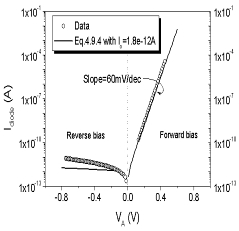
$$I = I_0(e^{qV/kT} - 1)$$

$$I_0 = Aqn_i^{\,2}\left(\frac{D_p}{L_pN_d} + \frac{D_n}{L_nN_a}\right)$$

What causes the IV curves to shift to lower V at higher T ?

Slide 4-105

4.9.1 Contributions from the Depletion Region



$n \approx p \approx n_i e^{qV/2kT}$

Net recombination (generation) rate:

$$\frac{n_i}{\tau_{dep}}(e^{qV/2kT}-1)$$

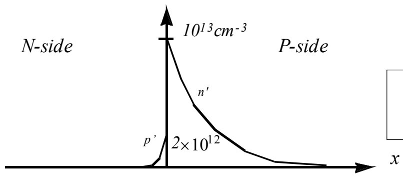
$$I = I_0(e^{qV/kT}-1) + A\frac{qn_iW_{dep}}{\tau_{dep}}(e^{qV/2kT}-1)$$

Space-Charge Region (SCR) current

$$I_{leakage} = I_0 + A\frac{qn_iW_{dep}}{\tau_{dep}}$$

Under forward bias, SCR current is an extra current with a slope 120mV/decade

4.10 Charge Storage

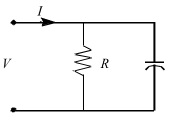


$Q \propto I$

$$I = Q/\tau_s$$
$$Q = I\tau_s$$

What is the relationship between τ_s (charge-storage time) and τ (carrier lifetime)?

4.11 Small-signal Model of the Diode



$$G \equiv \frac{1}{R} = \frac{dI}{dV} = \frac{d}{dV}I_0(e^{qV/kT}-1) \approx \frac{d}{dV}I_0e^{qV/kT}$$
$$= \frac{q}{kT}I_0(e^{qV/kT}) = I_{DC}/\frac{kT}{q}$$

What is G at 300K and I_{DC} = 1 mA?

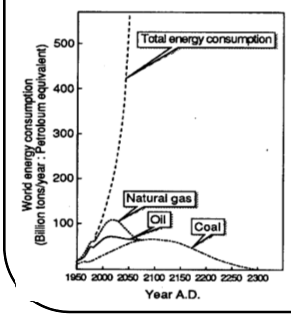
Diffusion Capacitance:

$$C = \frac{dQ}{dV} = \tau_s\frac{dI}{dV} = \tau_s G = \tau_s I_{DC} / \frac{kT}{q}$$

Which is larger, diffusion or depletion capacitance?

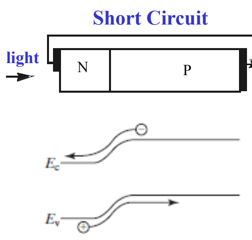
Part II: Application to Optoelectronic Devices

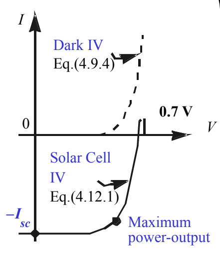
4.12 Solar Cells



- ***Solar Cells*** is also known as ***photovoltaic cells.***
- Converts sunlight to electricity with 10-30% conversion efficiency.
- 1 m^2 solar cell generate about 150 W peak or 25 W continuous power.
- Low cost and high efficiency are needed for wide deployment.

4.12.1 Solar Cell Basics

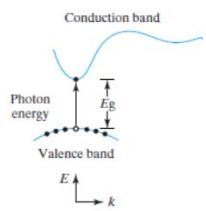




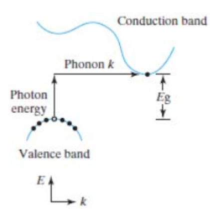
$$I = I_0(e^{qV/kT}-1) - I_{sc}$$

Direct-Gap and Indirect-Gap Semiconductors

- Electrons have both particle and wave properties.
- An electron has energy E and wave vector k.

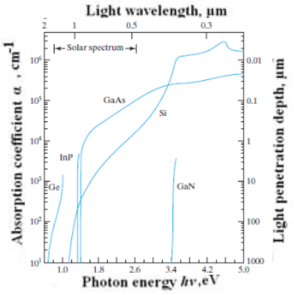


direct-gap semiconductor



indirect-gap semiconductor

4.12.2 Light Absorption



Light intensity (x) $\propto e^{-\alpha x}$

α(1/cm): absorption coefficient

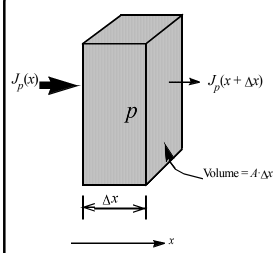
$1/\alpha$: light penetration depth

$$\text{Photon Energy (eV)} = \frac{hc}{\lambda} = \frac{1.24}{\lambda}(\mu m)$$

A thinner layer of direct-gap semiconductor can absorb most of solar radiation than indirect-gap semiconductor. But Si…

Slide 4-112

4.12.3 Short-Circuit Current and Open-Circuit Voltage



If light shines on the N-type semiconductor and generates holes (and electrons) at the rate of G $s^{-1}cm^{-3}$,

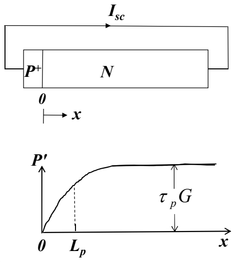
$$\frac{d^2 p'}{dx^2} = \frac{p'}{L_p^2} - \frac{G}{D_p}$$

If the sample is uniform (no PN junction), $d^2p'/dx^2 = 0 \rightarrow p' = GL_p^2/D_p = G\tau_p$

Slide 4-113

Solar Cell Short-Circuit Current, I_{sc}

Assume very thin P+ layer and carrier generation in N region only.



$$p'(\infty) = L_p^2 \frac{G}{D_p} = \tau_p G$$

$$p'(0) = 0$$

$$p'(x) = \tau_p G(1 - e^{-x/L_p})$$

$$J_p = -qD_p \frac{dp'(x)}{dx} = q\frac{D_p}{L_p}\tau_p G e^{-x/L_p}$$

$$I_{sc} = AJ_p(0) = AqL_p G$$

G is really not uniform. L_p needs be larger than the light penetration depth to collect most of the generated carriers.

Slide 4-114

Open-Circuit Voltage

•Total current is I_{SC} plus the PV diode (dark) current:

$$I = Aq\frac{n_i^2}{N_d}\frac{D_p}{L_p}(e^{qV/kT} - 1) - AqL_p G$$

•Solve for the open-circuit voltage (V_{oc}) by setting $I=0$

(assuming $e^{qV_{oc}/kT} >> 1$)

$$0 = \frac{n_i^2}{N_d}\frac{D_p}{L_p}e^{qV_{oc}/kT} - L_p G$$

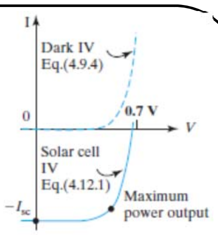
$$V_{oc} = \frac{kT}{q}\ln(\tau_p G N_d / n_i^2)$$

How to raise V_{oc} ?

Slide 4-115

4.12.4 Output Power

A particular operating point on the solar cell I-V curve maximizes the output power (I × V).



$$\text{Output Power} = I_{sc} \times V_{oc} \times FF$$

•Si solar cell with 15-20% efficiency dominates the market now

•Theoretically, the highest efficiency (~24%) can be obtained with 1.9eV >Eg>1.2eV. Larger Eg lead to too low Isc (low light absorption); smaller Eg leads to too low Voc.

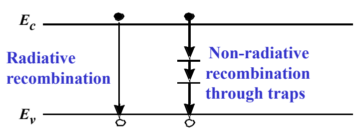
•***Tandem solar cells*** gets 35% efficiency using large ***and*** small Eg materials tailored to the short and long wavelength solar light.

Slide 4-116

4.13 Light Emitting Diodes and Solid-State Lighting

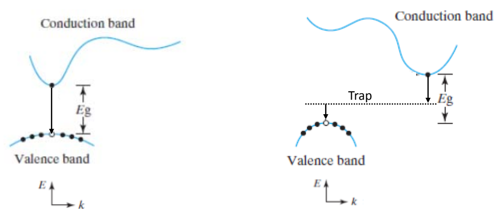
Light emitting diodes (LEDs)

- LEDs are made of compound semiconductors such as InP and GaN.
- Light is emitted when electron and hole undergo ***radiative recombination.***



Slide 4-117

Direct and Indirect Band Gap



Direct band gap
Example: GaAs

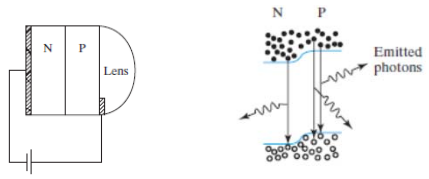
Direct recombination is efficient as k conservation is satisfied.

Indirect band gap
Example: Si

Direct recombination is rare as k conservation is not satisfied

Slide 4-118

4.13.1 LED Materials and Structure



$$\text{LED wavelength } (\mu\text{m}) = \frac{1.24}{\text{photon energy}} \approx \frac{1.24}{E_g(eV)}$$

Slide 4-119

4.13.1 LED Materials and Structure

	$E_g(eV)$	Wavelength (μm)	Color	Lattice constant (Å)
InAs	0.36	3.44		6.05
InN	0.65	1.91	infrared	3.45
InP	1.36	0.92		5.87
GaAs	1.42	0.87	Red Yellow	5.66
GaP	2.26	0.55	Green Blue	5.46
AlP	3.39	0.51		5.45
GaN	2.45	0.37		3.19
AlN	6.20	0.20	UV	3.11

Light-emitting diode materials

compound semiconductors

binary semiconductors:
- **Ex: GaAs, efficient emitter**

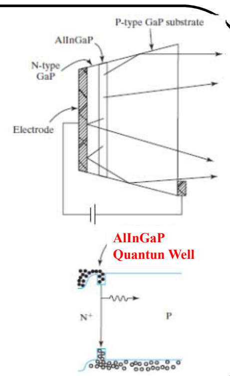
ternary semiconductor :
- **Ex: $GaAs_{1-x}P_x$, tunable E_g (to vary the color)**

quaternary semiconductors:
- **Ex: AlInGaP , tunable E_g and lattice constant (for growing high quality epitaxial films on inexpensive substrates)**

Slide 4-120

Common LEDs

Spectral range	Material System	Substrate	Example Applications
Infrared	InGaAsP	InP	Optical communication
Infrared -Red	GaAsP	GaAs	Indicator lamps. Remote control
Red-Yellow	AlInGaP	GaA or GaP	Optical communication. High-brightness traffic signal lights
Green-Blue	InGaN	GaN or sapphire	High brightness signal lights. Video billboards
Blue-UV	AlInGaN	GaN or sapphire	Solid-state lighting
Red-Blue	Organic semicon-ductors	glass	Displays



Slide 4-121

4.13.2 Solid-State Lighting

luminosity (lumen, lm): a measure of visible light energy normalized to the sensitivity of the human eye at different wavelengths

Incandescent lamp	Compact fluorescent lamp	Tube fluorescent lamp	**White LED**	Theoretical limit at peak of eye sensitivity (λ=555nm)	Theoretical limit (white light)
17	**60**	**50-100**	**90-?**	**683**	**~340**

Luminous efficacy of lamps in lumen/watt

Organic Light Emitting Diodes (OLED) :
has lower efficacy than nitride or aluminide based compound semiconductor LEDs.

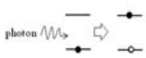
Terms: **luminosity** measured in **lumens**. **luminous efficacy,**

Slide 4-122

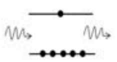
4.14 Diode Lasers

4.14.1 Light Amplification

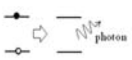
(a) Absorption



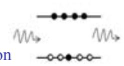
(d) Net Light Absorption



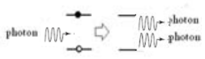
(b) Spontaneous Emission



(e) Net Light Amplification



(c) Stimulated Emission



Light amplification requires *population inversion*: electron occupation probability is larger for higher E states than lower E states.

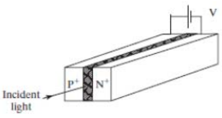
Stimulated emission: emitted photon has identical frequency and directionality as the stimulating photon; light wave is amplified.

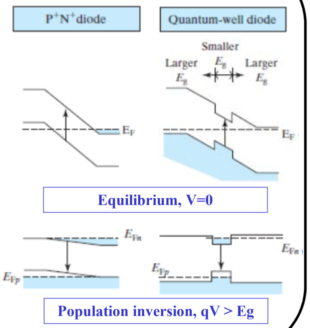
Slide 4-123

4.14.1 Light Amplification in PN Diode

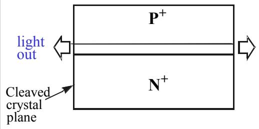
Population inversion is achieved when

$qV = E_{fn} - E_{fp} > E_g$





4.14.2 Optical Feedback and Laser



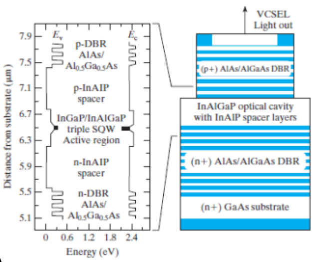
***Laser threshold** is reached (light intensity grows by feedback) when*

$$R_1 \times R_2 \times G \geq 1$$

- **R1, R2**: reflectivities of the two ends
- **G** : light amplification factor (gain) for a round-trip travel of the light through the diode

Light intensity grows until $R_1 \times R_2 \times G = 1$, when the light intensity is just large enough to stimulate carrier recombinations at the same rate the carriers are injected by the diode current.

4.14.2 Optical Feedback and Laser Diode



- ***Distributed Bragg reflector (DBR)*** **reflects light with multi-layers of semiconductors.**
- ***Vertical-cavity surface-emitting laser (VCSEL)*** **is shown on the left.**
- **Quantum-well laser has smaller threshold current because fewer carriers are needed to achieve population inversion in the small volume of the thin small-*Eg* well.**

4.14.3 Laser Applications

***Red** diode lasers: CD, DVD reader/writer*

***Blue** diode lasers: Blu-ray DVD (higher storage density)*

***1.55 µm infrared** diode lasers: Fiber-optic communication*

4.15 Photodiodes

***Photodiodes**:* Reverse biased PN diode. Detects photo-generated current (similar to Isc of solar cell) for optical communication, DVD reader, etc.

Avalanche photodiodes: Photodiodes operating near avalanche breakdown amplifies photocurrent by impact ionization.

Tunnel Diode

Several important devices for high-frequency applications use the instabilities that occur in semiconductors. An important type of instability involves *negative conductance*. Here we shall concentrate on three of the most commonly used negative conductance devices: *Esaki* or *tunnel* diodes, which depend on quantum-mechanical tunneling; transit-time diodes, which depend on a combination of carrier injection and transit-time effects; and *Gunn* diodes, which depend on the transfer of electrons from a high-mobility state to a low-mobility state. Each is a two-terminal device that can be operated in a negative conductance mode to provide amplification or oscillation at microwave frequencies in a proper circuit.

10.1 TUNNEL DIODES

The tunnel diode is a p-n junction device that operates in certain regions of its I–V characteristic by the quantum mechanical tunneling of electrons through the potential barrier of the junction. (See Sections 2.4.4 and 5.4.1.) The tunneling process for reverse current is essentially the Zener effect, although negligible reverse bias is needed to initiate the process in tunnel diodes. As we shall see in this section, the tunnel diode (often called the Esaki diode after L. Esaki, who received the Nobel Prize in 1973 for his

work on the effect) exhibits the important feature of *negative resistance* over a portion of its I–V characteristic.

10.1.1 Degenerate Semiconductors

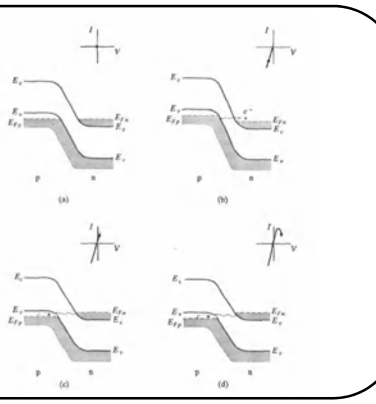
Thus far, we have discussed the properties of relatively pure semiconductors; any impurity doping represented a small fraction of the total atomic density of the material. Since the few impurity atoms were so widely spaced throughout the sample, we could be confident that no charge transport could take place within the donor or acceptor levels themselves. At high doping, the impurities are so close together that we can no longer consider the donor level as being composed of discrete, noninteracting energy states. Instead, the donor states form a band, which may overlap the bottom of the conduction band. If the conduction-band electron concentration n exceeds the effective density of states N_c, the Fermi level is no longer within the band gap, but lies within the conduction band. When this occurs, the material is called *degenerate* n-type. The analogous case of degenerate p-type material occurs when the acceptor concentration is very high and the Fermi level lies in the valence band. We recall that the energy states below E_F are mostly filled and states above E_F are empty, except for a small distribution dictated by the Fermi statistics. Thus, in a degenerate n-type sample, the region between E_c and E_F is for the most part filled with electrons, and in a degenerate p-type sample, the region between E_v and E_F is almost completely filled with holes.

A p-n junction between two degenerate semiconductors is illustrated in terms of energy bands in Fig. 10–1a. This is the equilibrium condition, for which the Fermi level is constant throughout the junction. We notice that E_{Fp} lies below the valence-band edge on the p side, and E_{Fn} is above the conduction-band edge on the n side. Thus, the bands must overlap on the energy scale in order for E_F to be constant. This overlapping of bands is very important; it means that, with a small forward or reverse bias, filled states and empty states appear opposite each other, separated by essentially the width of the depletion region. If the metallurgical junction is sharp, the depletion region will be very narrow for such high-doping concentrations, and the electric field at the junction will be quite large. Hence, the conditions for electron tunneling are met: filled and empty states separated by a narrow potential barrier of finite height. In Fig. 10–1, the bands are shown filled to the Fermi level for convenience of illustration, with the understanding that a distribution is implied.

Since the bands overlap under equilibrium conditions, a small reverse bias (Fig. 10-1b) allows electron tunneling from the filled valence-band states below E_{Fp} to the empty conduction-band states above E_{Fn}. This condition is similar to the Zener effect, except that no bias is required to create the condition of overlapping bands. As the reverse bias is increased, E_{Fn} continues to move down the energy scale with respect to E_{Fp}, placing more filled states on the p side opposite empty states on the n side. Thus, the tunneling of electrons from p to n increases with increasing reverse bias. The resulting conventional current is opposite to the electron flow—that is, from n to p. At

Slide 1-130

Figure 10–1
Tunnel diode band diagrams and I–V characteristics for various biasing conditions: (a) equilibrium (zero bias) condition, no net tunneling; (b) small reverse bias, electron tunneling from p to n; (c) small forward bias, electron tunneling from n to p; (d) increased forward bias, electron tunneling from n to p decreases as bands pass by each other



Slide 1-131

equilibrium (Fig. 10-1a), there is equal tunneling from n to p and from p to n, given a zero net current.

When a small forward bias is applied (Fig. 10-1c), E_{Fn} moves up in energy with respect to E_{Fp} by the amount qV. Thus, electrons below E_{Fn} on the n side are placed opposite empty states above E_{Fp} on the p side. Electron tunneling occurs from n to p as shown, with the resulting conventional current from p to n. This forward-tunneling current continues to increase with increased bias as more filled states are placed opposite empty states. However, as E_{Fn} continues to move up with respect to E_{Fp}, a point is reached at which the bands begin to pass by each other. When this occurs, the number of filled states opposite empty states decreases. The resulting decrease in tunneling current is illustrated in Fig. 10-1d. This region of the I–V characteristic is important in that the *decrease* in tunneling current with *increased* bias produces

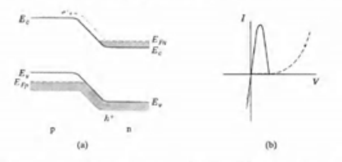


Figure 10–2
Band diagram (a) and I–V characteristic (b) for the tunnel diode beyond the tunnel current region. In (b), the tunneling component of current is shown by the solid curve and the diffusion current component is dashed.

a region of negative slope; that is, the *dynamic resistance dV/dI* is negative. This negative-resistance region is useful in oscillators.

Slide 1-132

If the forward bias is increased beyond the negative-resistance region, the current begins to increase again (Fig. 10-2). Once the bands have passed each other, the characteristic resembles that of a conventional diode. The forward current is now dominated by the diffusion current—electrons surmounting their potential barrier from n to p and holes surmounting their potential barrier from p to n. Of course, the diffusion current is present in the forward tunneling region, but it is negligible compared with the tunneling current.

The total tunnel diode characteristic (Fig. 10-3) has the general shape of an N (if a little imagination is applied); therefore, it is common to refer to this characteristic as exhibiting a *type-N negative resistance*. It is also called a *voltage-controlled negative resistance*, meaning that the current decreases rapidly at some critical voltage (in this case, the *peak voltage* V_p, taken at the point of maximum forward tunneling).

The values of *peak tunneling current* I_p and *valley current* I_v (Fig. 10-3) determine the magnitude of the negative-resistance slope for a diode of given material. For this reason, their ratio I_p/I_v is often used as a figure of merit for

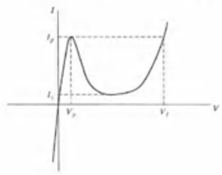


Figure 10–3
Total tunnel diode characteristic.

Slide 1-133

the tunnel diode. Similarly, the ratio V_p/V_f is a measure of the voltage spread between the two positive-resistance regions.

The negative resistance of the tunnel diode can be used in a number of ways to achieve oscillation and other circuit functions. The fact that the tunneling process does not present the time delays of drift and diffusion makes the tunnel diode a natural choice for certain high-speed circuits. However, the tunnel diode has not achieved widespread application, because of its relatively low current operation and competition from other devices.

Slide 1-134