Photoconductivity

⊛ Equilibrium and non-equilibrium carrier concentration :

⊙ The equilibrium carrier concentration can be achived only by means of heat. When ~~light~~ heat is supplied to the material, first it interacts with the lattice, so lattice vibration is increased. Hence electron associated with the bond are slowly realesed. At each step equilibrium is maintained. (Heat → lattice vibration → electron release).

⊙ If the carriers move on to the conduction band by means of any energy other than heat, then carriers are non-equilibrium carrier. The energy may be supplied by means of light, high energy particle etc. In the case of light, the photon directly interact with the bound electrons and transfer it to the conduction band.

Lattice has no role in this case.

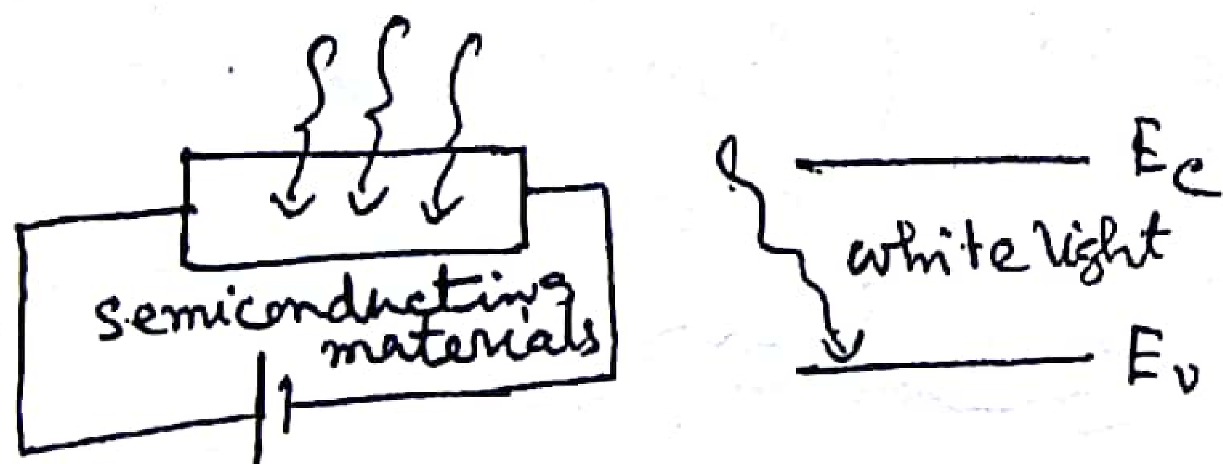
The non-equilibrium carrier concentration is

$$n = n_0 + \Delta n$$

where n_0 is equilibrium carrier concentration

and Δn is ~~non-equilibrium carrier~~ excess carrier concentration.

In case of conductivity due to the incident photon, the average energy of white light, nearly equal to 1.45 eV.



If the band gap is less than that 1.45 eV then there is a possibility of transfer of electron from V.B. to C.B.

In dark conductivity or photoconductivity, the external electric field should be applied.

** The equilibrium carrier concentration is designated simultaneously considering two process:

1. Generation (G)
2. Recombination (R)

When there is non-equilibrium carriers in C.B., the recombination rate is high.

Considering the generation process and recombination at a particular temperature in a semiconducting material we get,

$$\frac{\delta n}{\delta t} = G - R$$

G → generation of electron
R → recombination

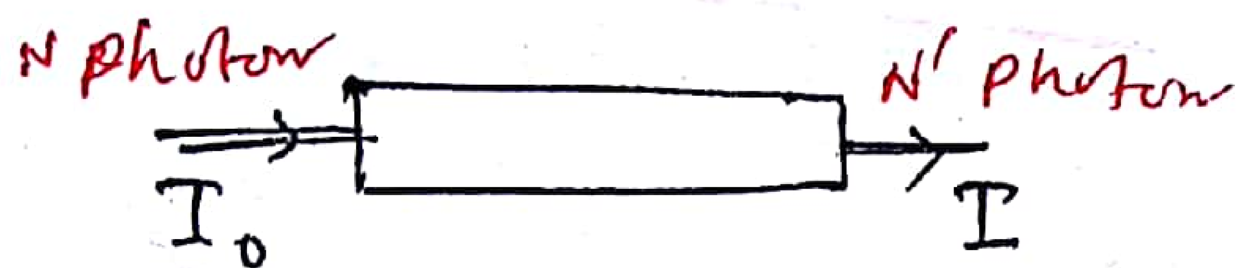
where n is non-equilibrium carrier concentration, and equal to $n = n_0 + \Delta n$ where n_0 is equilibrium carrier concentration.

So, $$\frac{\delta(\Delta n)}{\delta t} = G - R$$

If there is photoconductivity in the material, that indicates that the light which is falling on the material must be absorbed. The material should have high optical absorption co-efficient.

$$I = I_0 e^{-\alpha x},$$

[I_0 → primary beam intensity.
I → transmitted beam intensity
α → optical absorption co-efficient]



Energy absorbed by the material,

$$\Delta I = I_0 - I = I_0(1 - e^{-\alpha x}),$$ where, $\alpha x << 1$. $e^{-\alpha x} = (1 - \alpha x)$

$$= I_0(1 - 1 + \alpha x)$$

$$= I_0 \alpha x$$

or, $$\Delta N . h\nu = I_0 \alpha x,$$

$$\Rightarrow \frac{\Delta N . h\nu}{x} = I_0 \alpha$$

[$I_0 = N . h\nu$, N → no. of photon / volume
$h\nu$ → energy of one photon.
$I = N' . h\nu$
N' → no. of photon/volume transmitted after absorption
$\Delta I = I_0 - I$
$= (N - N')h\nu$
$= \Delta N . h\nu$]

No. of photon absorbed per unit volume per unit time,

$$= \frac{\Delta N}{x} = \frac{I_0 \alpha}{h\nu}$$

Now, quantum efficiency (β) = $\dfrac{\text{no. of pair produced per unit volume per unit time}}{\text{no. of photon absorbed per unit volume per unit time}}$

No. of electron-hole pairs produced per unit volume per unit time = No. of photons absorbed per unit volume per unit time × β

$$= \frac{\Delta N}{x} \times \beta$$

$$= \frac{I_0 \alpha \beta}{h\nu} = K I_0 \beta, \quad \text{where, } k = \frac{\alpha}{h\nu}$$

This is known as generation. $= K I_0 \beta$

Now, Recombination (R) is proportional to,

$$R \propto (np - n_0 p_0)$$

$$\propto [(n_0 + \Delta n)(p_0 + \Delta p) - n_0 p_0]$$

$$\propto [n_0 p_0 + n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p - n_0 p_0]$$

$$\propto [n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p]$$

If the charge neutrality is maintained,

$$\Delta n = \Delta p.$$

So,

$$R \propto [(n_0 + p_0)\Delta n + (\Delta n)^2]$$

Case I: If the sample is n-type, then $n_0 >> p_0$.

and the excess carrier concentration $\Delta n << n_0$, hence $(\Delta n)^2$ is negligible. [when illumination is very weak]

when the intensity of light is very weak, then $\Delta n << n_0$.

$$\therefore \boxed{R \propto n_0 \Delta n}$$

This shows that the recombination is depends on excess carrier concentration Δn.

now, $R = \gamma_n n_0 \Delta n$, γ_n is the transition probability constant.

now, life time of the carrier, $\tau = \frac{1}{\gamma_n n_0}$

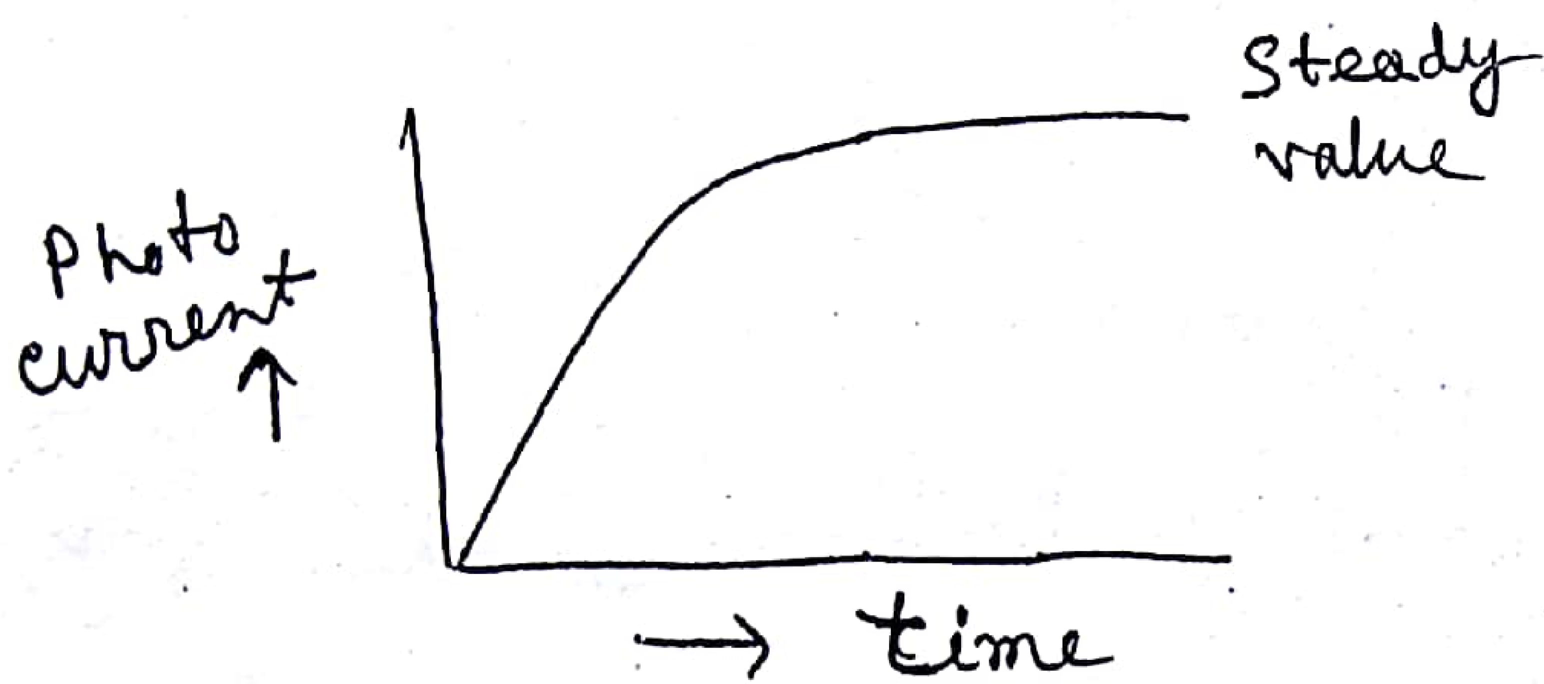
i.e., $\boxed{R = \frac{\Delta n}{\tau}}$

This shows that the recombination rate is high if the excess

carrier concentration is large and is also ~~high provided~~ the life time of the carrier is small.

τ is also known as mean life time of the non-equilibrium carrier, i.e, each non-equilibrium carrier appearing in a semiconductor live that limitted time before it recombination.

$\frac{1}{\tau}$ is the probability of recombination which is achieved by means of scattering. This is ~~only~~ mainly based by Mathesen's rule.



Hence, we get in case of non-steady state for weak intensity of light,

the eqn of excess charge carrier is given by:

$$\frac{\delta(\Delta n)}{\delta t} = G - R = kI_0\beta - \frac{\Delta n}{\tau}$$

$$= \frac{kI_0\beta\tau - \Delta n}{\tau}$$

Solution of Δn (excess carrier concentration)

Rearranging and integrating, we get.

$$\int \frac{\delta(\Delta n)}{kI_0\beta\tau - \Delta n} = -\int \frac{\delta t}{\tau}$$

$$\Rightarrow \int \frac{\delta(kI_0\beta\tau - \Delta n)}{kI_0\beta\tau - \Delta n} = \int -\frac{\delta t}{\tau}$$

$$\Rightarrow \ln(kI_0\beta\tau - \Delta n) = -\frac{t}{\tau} + \text{Const.}$$

Since, at $t=0, \Delta n = 0$; So, $\ln(kI_0\beta\tau) = \text{Constant}$.

So, $\ln(kI_0\tau\beta - \Delta n) = -\frac{t}{\tau} + \ln(kI_0\beta\tau)$

or, $\frac{kI_0\beta\tau - \Delta n}{kI_0\beta\tau} = e^{-t/\tau}$

$\Rightarrow \Delta n = kI_0\beta\tau(1-e^{-t/\tau}) = G\tau(1-e^{-t/\tau})$

This is the equation for the growth of carriers under illumination.

The dark current, i.e, when no light is applied and constant voltage is applied across the sample,

$$J_D = \sigma E = q(n_0 \mu_n + p_0 \mu_p) E$$

where n_0, p_0 are equilibrium carrier concentration and μ_n, μ_p is the mobility of the electron and hole respectively.

The photocurrent is given by :

$$J_{ph} = q\{(n_0 + \Delta n)\mu_n + (p_0 + \Delta p)\mu_p\} E$$

The change in photo current at any instant of time rather than steady value is

$$\Delta J_{ph}\Big|_t = J_{ph} - J_D = q \Delta n (\mu_n + \mu_p) E.$$

$\because$ $\Delta n = \Delta p$, when to satisfy the neutrality condition.

Again, the photo current at the steady value

$$\Delta J_{ph}\Big|_s = J_{ph} - J_s = q \Delta n_s (\mu_n + \mu_p) E$$

$$\frac{\Delta J_{ph}|_s}{\Delta J_{ph}|_t} = \frac{\Delta n_s}{\Delta n} = e^{t/\tau}.$$

Photoconductivity, $\Delta\sigma = q \Delta n_s (\mu_n + \mu_p) = qG(\tau_n \mu_n + \tau_p \mu_p)$ (*)

$= \left(\frac{q \beta \alpha I_0}{h\nu}\right)(\mu_n \tau_n + \tau_p \mu_p)$

$\Delta n_s \rightarrow$ the excess carrier concentration at the steady value. $= G\tau$.

$$(\Delta n)_s = G\tau = K I_0 \beta \tau = \left(\frac{\beta \alpha I_0}{h\nu}\right)\tau$$

For electron we can write,

$$(\Delta n)_s = \left(\frac{\beta \alpha I_0}{h\nu}\right)\tau_n$$

Steady state value

$\beta \rightarrow$ quantum efficiency.
$\alpha \rightarrow$ absorption co-efficient
$\tau_n \rightarrow$ life time of electron

(*) when τ_n and τ_p are different. But for simple recombination, $\tau_p = \tau_n$

{ 2003 ⇒ 1 (iv)
2004 ⇒ 5. (a) }

Case II : when the intensity of light is strong.

then, $n_0 << \Delta n$ and $p_0 << \Delta p$.

now, the recombination rate,

$\gamma \rightarrow$ transmission Probability constant

$$R = \gamma [(n_0 + p_0) + \Delta n](\Delta n)$$, assuming $\Delta n = \Delta p$.

Since, under the strong illumination, $\Delta n >> n_0$,

$$R = \gamma (\Delta n)^2$$

now, the eqn of change in carrier concentration

$$\frac{\delta(\Delta n)}{\delta t} = G - R = G - \gamma (\Delta n)^2$$

Soln of Δn :

$$\Delta n(t) = \Delta n_1 \tanh \sqrt{G\gamma}\, t$$
$$= \sqrt{\frac{G}{\gamma}} \tanh \sqrt{G\gamma}\, t$$

Photoconductivity, $\Delta \sigma = q \Delta n (\mu_n + \mu_p)$.

2004 ⇒ 5.(b)

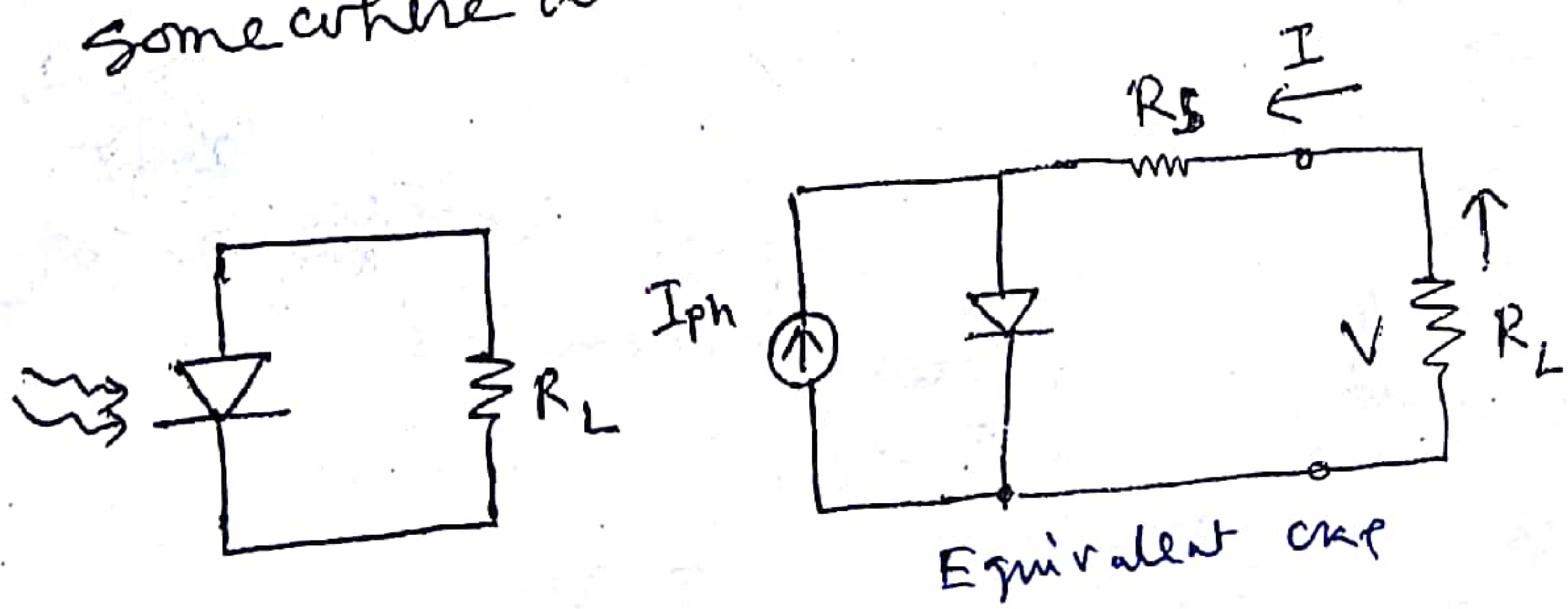
Scanned by CamScanner

Solar cell

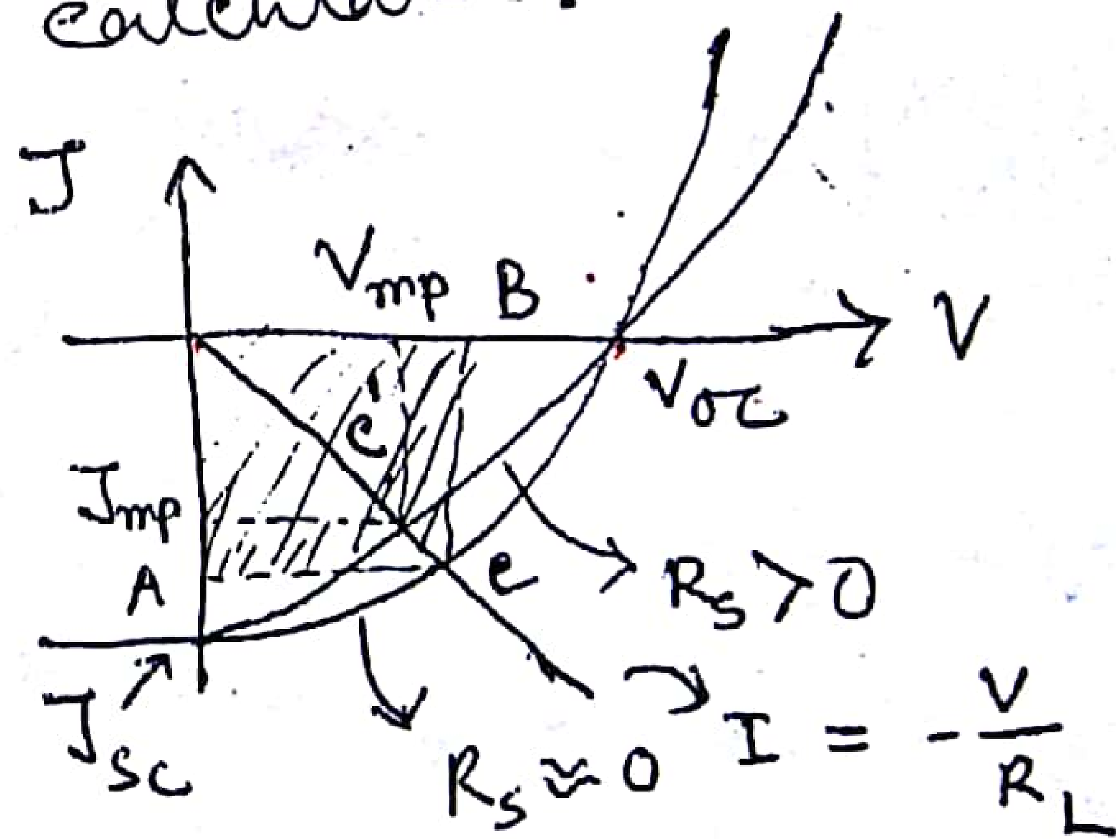
The most important use of photovoltaic effect today is for direct conversion of solar energy to electricity. Diodes made for that specific purpose are called solar cell or solar battery. It is seen that in the fourth quadrant of the I-V characteristics of a photodiode, the power is negative (positive voltage and negative current). This means that the device will deliver power from the junction instead of absorbing it from the power supply. Hence they are appropriately called cells.

Only Si cells are available commercially today, though GaAs, because of it's higher E_g, yields higher voltage and efficiency.

To be used as an energy source the solar cell is connected to a load R_L as in the fig below. The load characteristics or load line is $I = V/R_L$. To get maximum current, the output voltage of the cell will be zero (marked by A) and will deliver no power. Similarly to ge maximum voltage V_{oc} from the cell the output current will be zero (marked by B), again delivering no power. Hence to get maximum output power out of the solar cell the operating point should be somewhere at C which we shall now calculate.



Equivalent ckt



When the cell is exposed to the solar spectrum, a photon with energy less than E_g makes no contribution to the cell output (neglecting photon-assisted absorption). A photon with energy greater than E_g contributes an energy E_g to the cell output, and the excess over E_g is wasted as heat. To derive the ideal conversion efficiency, we shall consider the energy band diagram of p-n junction under solar radiation (see before)

$h\nu > E_g$

The solar cell assumed to have ideal I-V characteristic. The equivalent circuit is shown in fig above, where a constant-current source is in parallel with the junction. The source current I_{ph} results from the excitation of excess carriers by solar radiation; I_s is the diode saturation current and R_L is the load resistance.

The I-V characteristics of such a device are given by

$$I = I_s\left[\exp\left(\frac{qV}{\eta kT}\right) - 1\right] - I_{ph} \quad \text{---- (1)}$$

and $$J_s = I_s/A = qN_cN_v\left(\frac{1}{N_A}\sqrt{\frac{D_n}{\tau_n}} + \frac{1}{N_D}\sqrt{\frac{D_p}{\tau_p}}\right)e^{-E_g/kT} \quad \text{---- (2)}$$

where A is the device area.

By properly choosing a load, a close to 80% of product $I_{sc}V_{oc}$ can be extracted (I_{sc} is the short-circuited current and V_{oc} is the opencircuit voltage of the cell, the shaded area is the maximum power rectangle.). The I-V curve is more generally represented by ~~the shaded reg~~ the fig above.

The quantity I_m and V_m that corresponds to the current and voltage, respectively for maximum power output $P_m (= I_mV_m)$.

For the open-circuit voltage ($I=0$):

$$V_{oc} = \frac{\eta kT}{q}\ln\left(\frac{I_{ph}}{I_s} + 1\right) \approx \frac{\eta kT}{q}\ln\left(\frac{I_{ph}}{I_s}\right)$$

Hence for a given I_L, the open-circuit voltage increases logarithmatically with decreasing saturation current I_s.

Output power is given by

Power = $$P = IV = I_sV\left[\exp\left(\frac{qV}{\eta kT}\right) - 1\right] - I_{ph}V$$

The condition for maximum power can be obtained when $\frac{dP}{dV} = 0$.

now, $$\frac{dP}{dV} = \frac{d}{dV}\left[I_sV\left[\exp(\beta V) - 1\right] - I_{ph}V\right]$$

where $\beta = \frac{q}{\eta kT}$.

$$\Rightarrow \quad I_s(e^{\beta V}-1) - I_{ph} + I_s V \beta e^{\beta V} = 0$$

at $\frac{dP}{dV} = 0$, the maximum power point

$$I = I_m, \quad V = V_m$$

So,

$$I_s(e^{\beta V_m}-1) - I_{ph} + I_s V_m \beta e^{\beta V_m} = 0$$

$$\Rightarrow \quad e^{\beta V_m}(I_s + I_s V_m \beta) = I_s + I_{ph}$$

$$\Rightarrow \quad e^{\beta V_m} = \frac{1 + I_{ph}/I_s}{1 + \beta V_m}$$

$$\Rightarrow \quad V_m = \frac{1}{\beta} \ln\left(\frac{I_{ph}/I_s + 1}{1 + \beta V_m}\right) = \frac{1}{\beta}\ln(I_{ph}/I_s + 1) - \frac{1}{\beta}\ln(1+\beta V_m)$$

from eqn (1)

$$\Rightarrow \quad V_m = V_{oc} - \frac{1}{\beta}\ln(1+\beta V_m)$$

$$\beta = \frac{q}{\eta k T}$$

Again,

$$I_m = I_s[e^{\beta V_m} - 1] - I_{ph}$$

$$= -I_s \beta V_m e^{\beta V_m}$$

$$= -I_s \beta V_m \times \frac{1 + I_{ph}/I_s}{1 + \beta V_m}$$

~~(as $I_{ph} \gg I_s$, $1 + \frac{I_{ph}}{I_s} \approx I_{ph}/I_s$)~~

~~$= -I_s \beta V_m \times \frac{I_{ph}}{I_s}(1+\beta V_m)^{-1} = -\beta V_m \cdot I_{ph}$~~

$$= -\frac{I_s \beta V_m (1 + I_{ph}/I_s)}{1 + \beta V_m} = -\frac{(I_{ph} + I_s)\beta V_m}{1 + \beta V_m} = I_m$$

$h\nu > E_g$

Hence, <u>Maximum Power output</u> P_m is then

$$P_{max} = I_m \times V_m = -\frac{(I_{ph} + I_s)\beta V_m}{1 + \beta V_m} \times V_m$$

The area $P_m = I_m V_m$ is called <u>the maximum power rectangle.</u>

The maximum obtainable power :

$$P_m = I_m V_m = -\frac{(I_{ph} + I_s)\beta V_m}{1+\beta V_m} \cdot V_m$$

where, $I_{ph} \rightarrow$ the source current of the cell.

$I_s \rightarrow$ the diode saturation current of cell.

$V_m \rightarrow$ optimum voltage at the cell output at maximum power point

$\beta \rightarrow \frac{q}{\eta kT}$, $\eta \rightarrow$ ideality factor

$$P_m = -\frac{(I_{ph} + I_s)\beta V_m^2}{1+\beta V_m}$$

The maximum power rectangle must be always less than $I_{sc} V_{oc}$; i.e, $P_m < I_{sc} V_{oc}$.

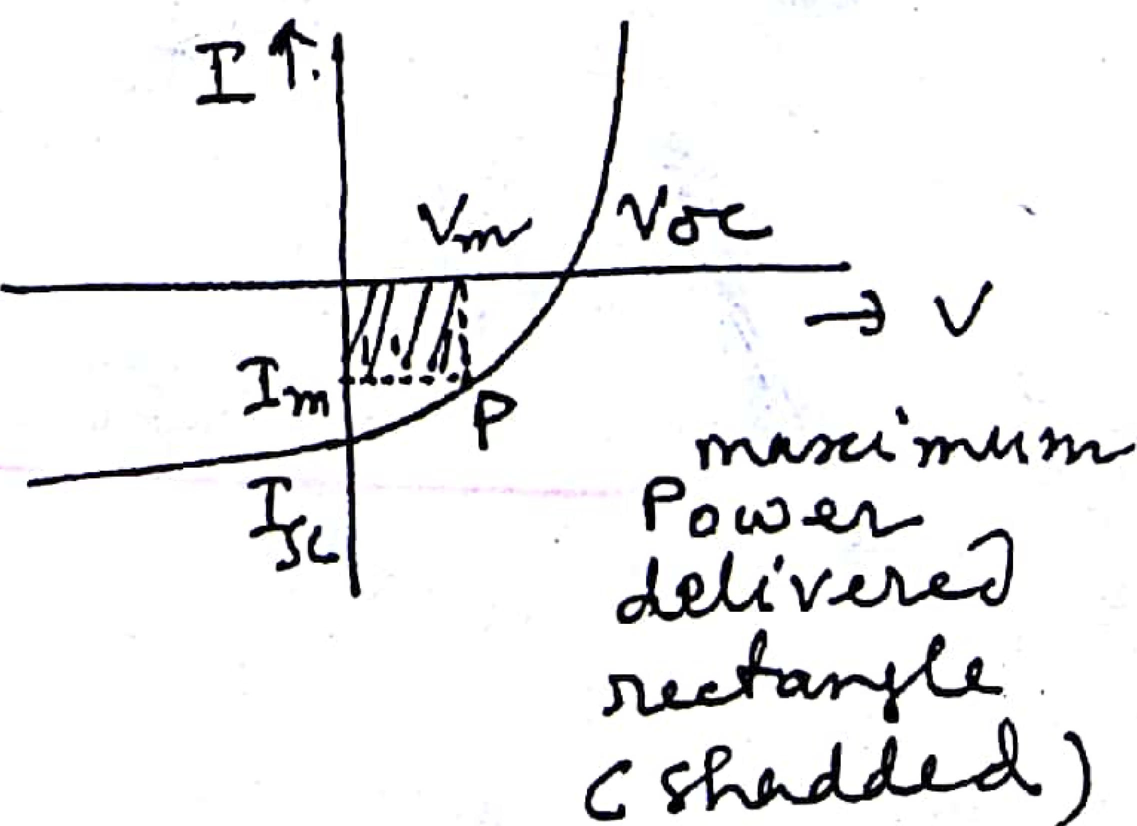
$I_{sc} \rightarrow$ short circuited current of the cell.

$V_{oc} \rightarrow$ open circuited voltage of the cell.

Fill factor of a solar cell :-

The Fill Factor (FF) of a solar cell is defined by a ratio of the maximum power output to the power that can be obtained by the multiplication of open circuit voltage and short circuited current of a given solar cell.

$$F.F. = \frac{V_m I_m}{V_{oc} I_{sc}} < 1$$



it is always less than 1.

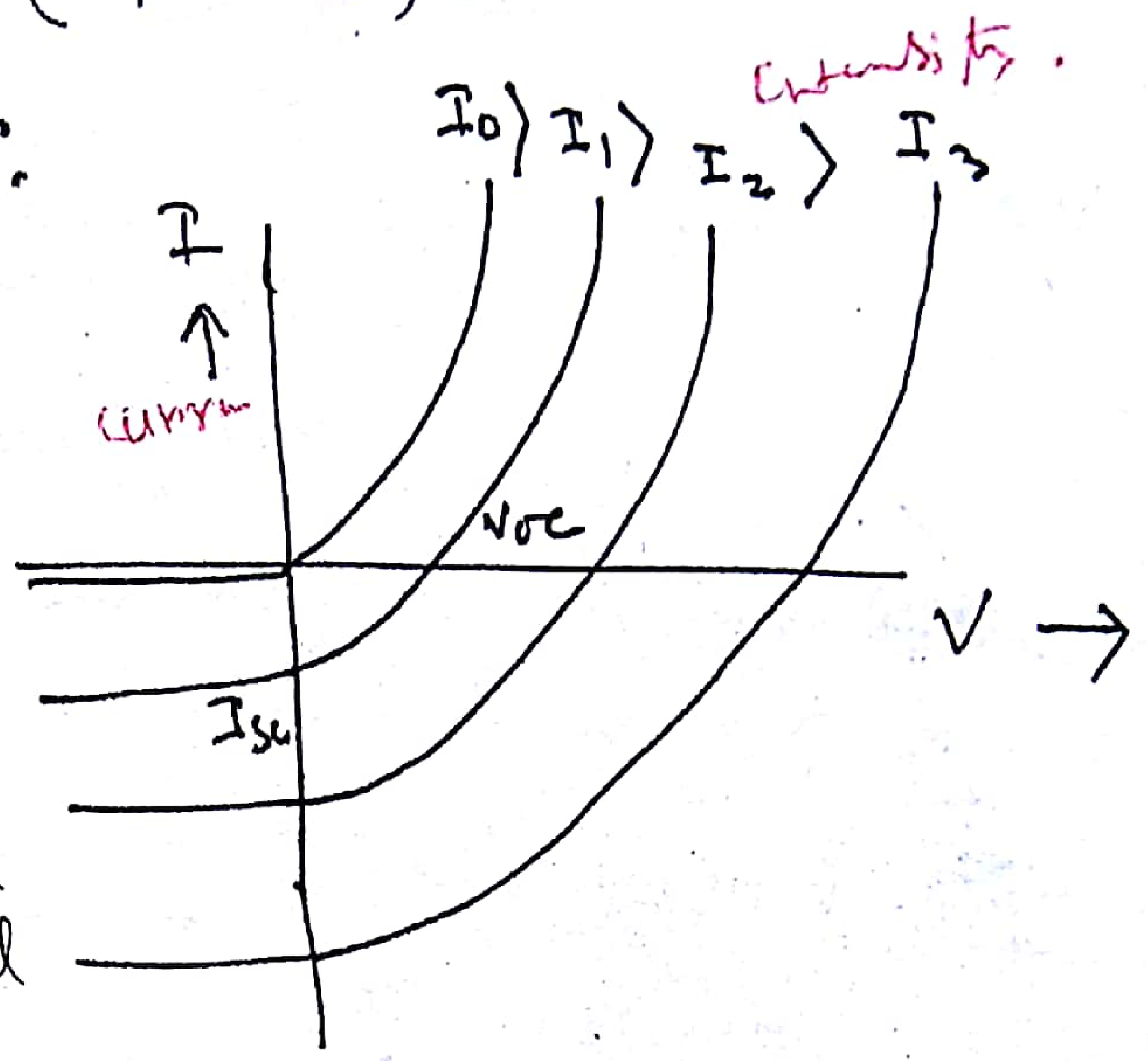
[Figure of merit for solar cell]

The optimum load impedance :-

$$R_m = \frac{V_m}{I_m} = -\frac{V_{oc} - \frac{1}{\beta}\ln(1+\beta V_m)}{\frac{(I_{ph}+I_s)\beta V_m}{1+\beta V_m}}$$

$$= -\left[\frac{V_{oc} - \frac{1}{\beta}\ln(1+\beta V_m)}{(I_{ph}+I_s)\beta V_m}\right]\cdot(1+\beta V_m)$$

Efficiency of a Solar Cell :



$I \to$ intensity of light

$I_0 \to$ dark condition

** The power is negative thermodynamically this means that solar cell can delivered power to an external circuit.

Effeciency of a solar energy converter is

$$\eta = \frac{\text{Power output}}{\text{Power input}} = \frac{P_{out}}{P_{in}} = \frac{P_m}{N_{ph}\, \rho_{av}}$$

$N_{ph} \Rightarrow$ Total no. of photon in the spectrum

$\rho_{av} \Rightarrow$ Average density of photon

$$\eta_{max} = \frac{\text{maximum power output}}{\text{Power input}}$$

Practical efficiency reach $\eta \to 13\%$ to 14%.

** P_{in} is the incident power of sun-falling on the solar material. This is arround $1.35\ KW/m^2$ on a normal suny day.

For good solar cell the requirements are :-

(1) Single crystal

(2) Solar grade material

(3) Life time of the carrier must be as large as possible.

(4) It should be high optical absorption co-efficient.

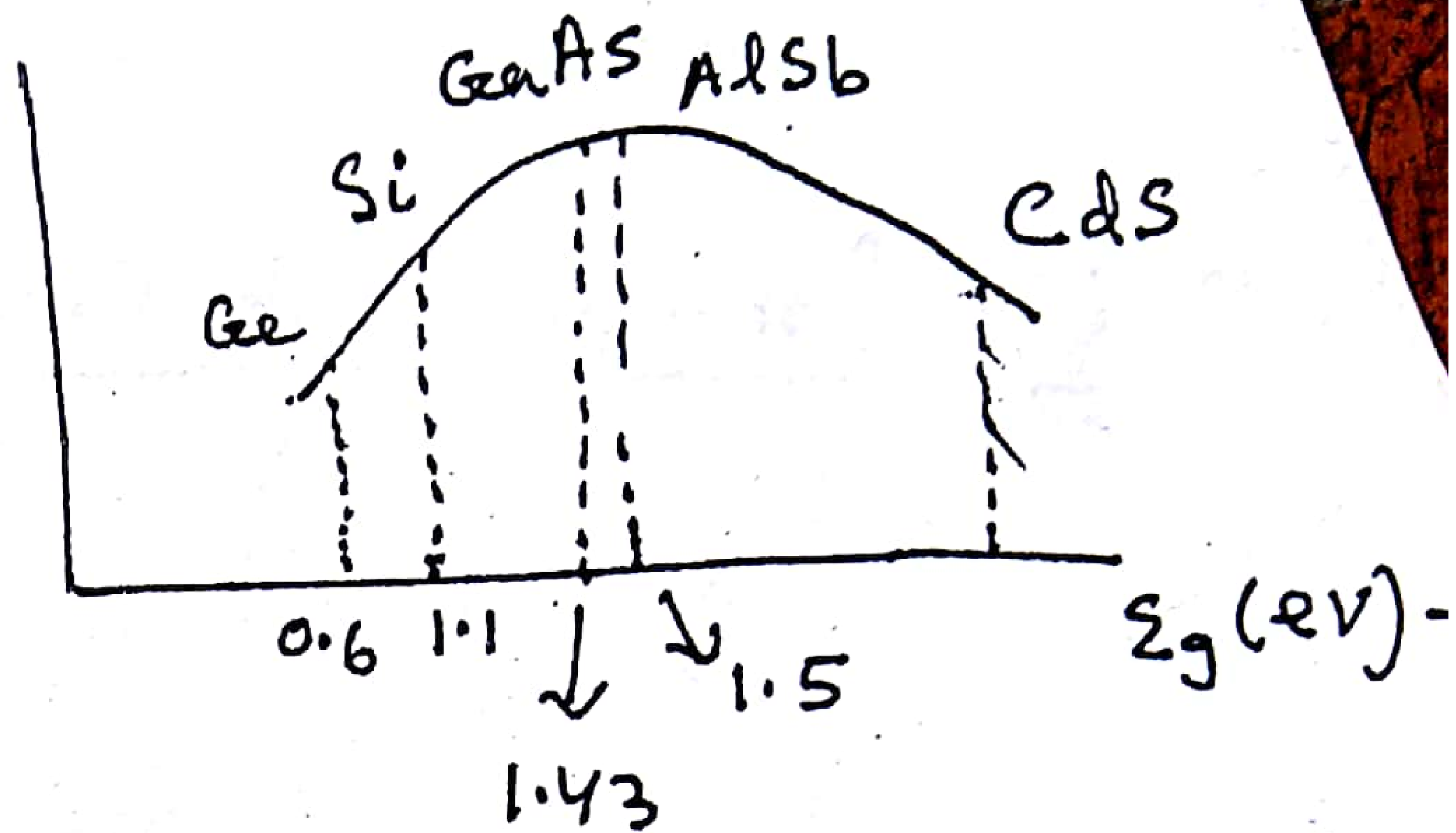
can convert solar energy with an efficiency of 40%.

Scanned by CamScanner

The solar cell assumed to have ideal I-V characteristic

**

V_m increase with increasing gap width of the semiconductor and I_{ph} decreases with increasing gap width. As a result, η_{max} first increases with increasing gap width, goes through a maximum at a gap width $E_g \sim 1.5$ eV and decreases for higher gap widths as shown in fig. For the same reason, R_m ~~[illegible]~~ increases steadily with E_g.



efficiency vs. energy gap.

**

If the series resistance R_s is taken into account I_{ph} and V_{oc} do not change, but (I-V) characteristic for $R_s > 0$ will lie below the characteristic for $R_s = 0$. Consequently, the maximum power rectangle decreases in area. The series resistance thus reduces the efficiency of the device and should be kept as small as possible.

**. Input power: 70-80 mW/cm² at sea-level

For Single crystal Si ≃ $\eta \sim 15\%$, $V_{oc} \sim 0.6$ V } area $\sim 40\ cm^2$

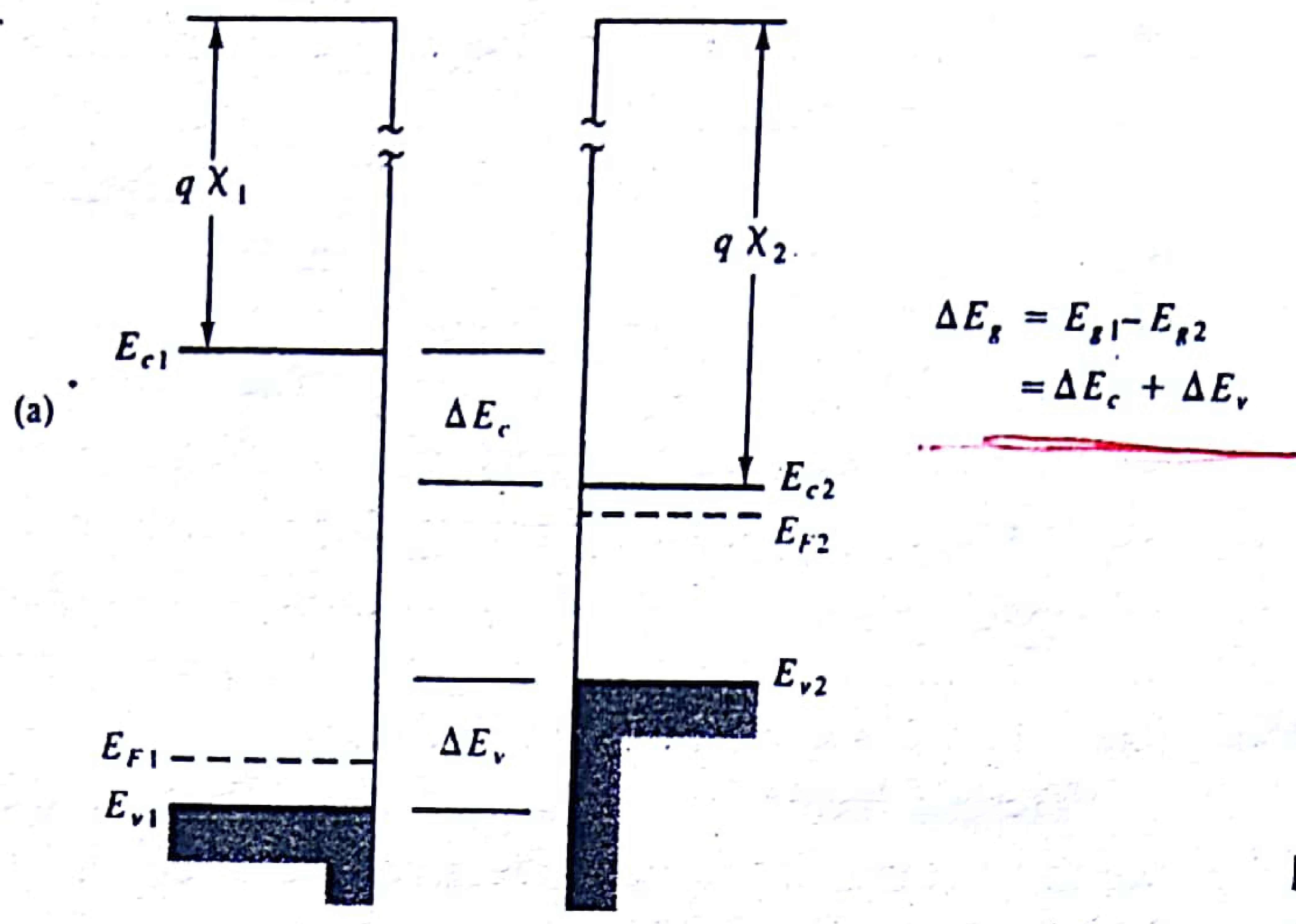
$R_s = 5\Omega$, $R_{sh} = 100\Omega$ }

HETEROJUNCTIONS

Thus far we have discussed p-n junctions formed within a single semiconductor (*homojunctions*) and junctions between a metal and a semiconductor. The third important class of junctions consists of those between two lattice-matched semiconductors with different band gaps (*heterojunctions*). We discussed lattice-matching in Section 1.4.1. The interface between two such semiconductors may be virtually free of defects, and continuous crystals containing single or multiple heterojunctions can be formed. The availability of heterojunctions and multilayer structures in compound semiconductors opens a broad range of possibilities for device development. We will discuss many of these applications in later chapters, including heterojunction bipolar transistors, field-effect transistors, and semiconductor lasers.

When semiconductors of different band gaps and electron affinities are brought together to form a junction, we expect discontinuities in the energy bands as the Fermi levels line up at equilibrium (Fig. 5-36). The discontinuities in the conduction band ΔE_c and the valence band ΔE_v accommodate the difference in band gap between the two semiconductors ΔE_g. In an ideal case, ΔE_c would be the difference in electron affinities $q(\chi_2 - \chi_1)$, and ΔE_v would be found from $\Delta E_g - \Delta E_c$. In practice, the band discontinuities are found experimentally for particular semiconductor pairs. For example, in the commonly used system GaAs–AlGaAs (see Figs. 3-6 and 3-13), the direct band gap difference ΔE_g^Γ between the wider band gap AlGaAs and the narrower band gap GaAs is apportioned approximately $\frac{2}{3}$ in the conduction band and $\frac{1}{3}$ in the valence band for the heterojunction. The built-in contact potential is divided between the two semiconductors as required to align the Fermi levels at equilibrium. The resulting depletion region on each side of the heterojunction and the amount of built-in potential on each side (making up the contact potential V_0) are found by solving Poisson's equation with the boundary condition of

Scanned by CamScanner



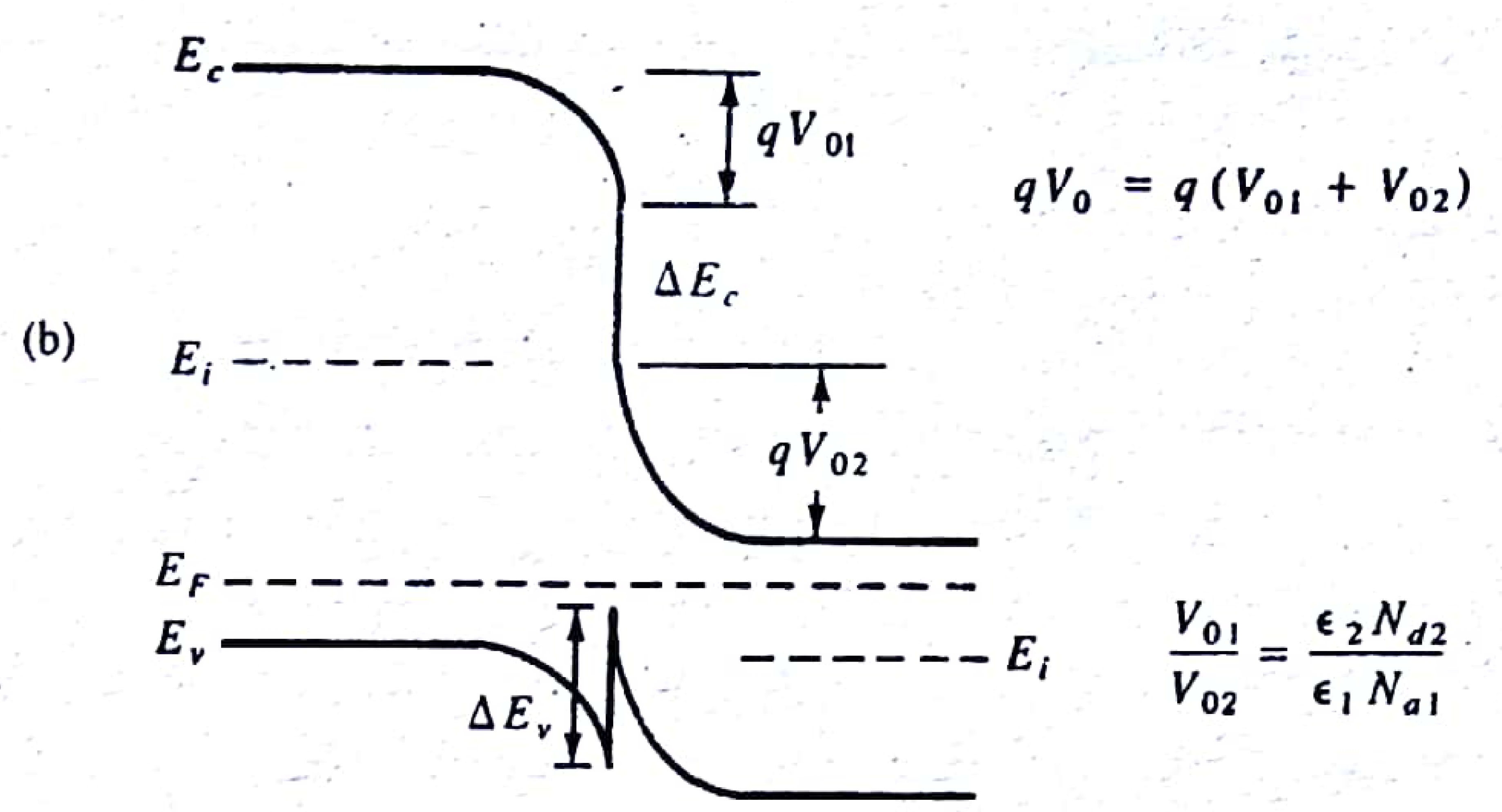
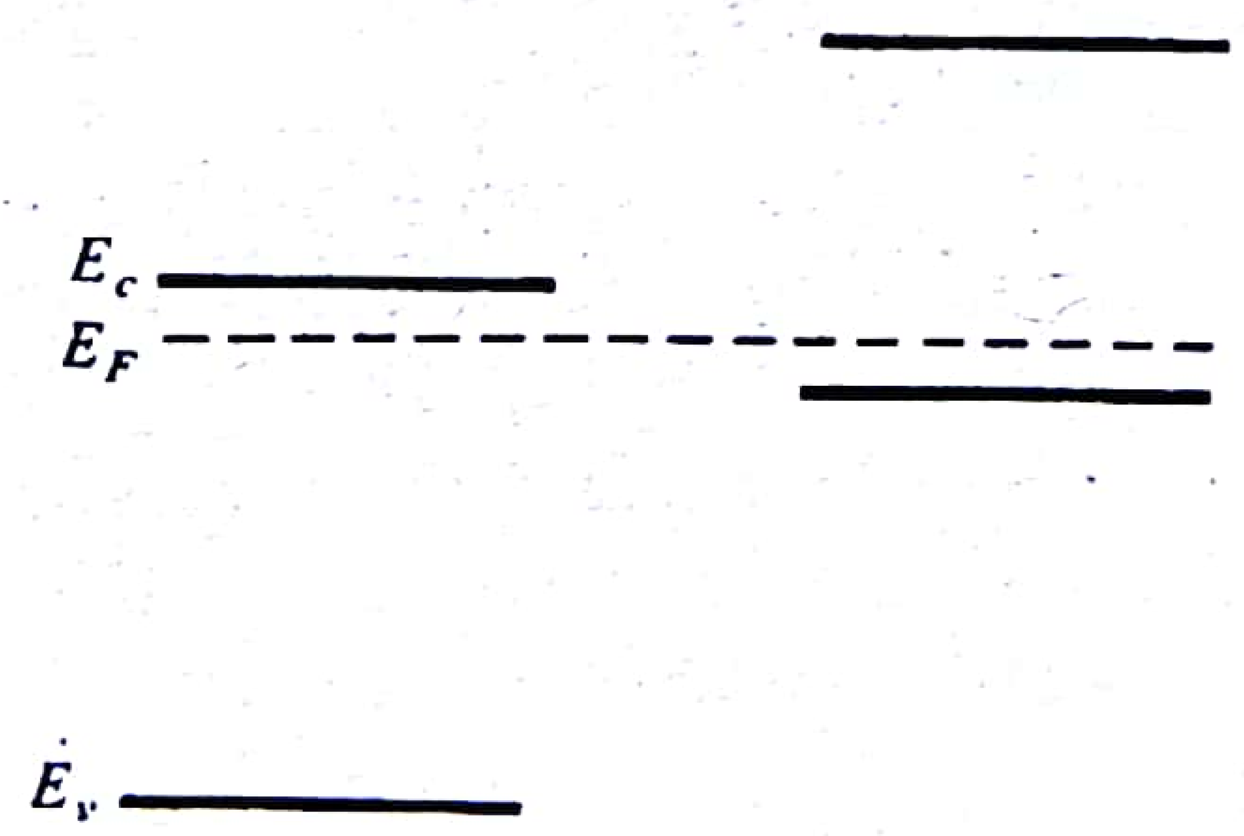


Figure 5-36 An ideal heterojunction between a p-type, wide band gap semiconductor and an n-type narrower band gap semiconductor: (a) band diagrams before joining; (b) band discontinuities and band bending at equilibrium.

continuous electric flux density, $\epsilon_1\mathscr{E}_1 = \epsilon_2\mathscr{E}_2$ at the junction. The barrier that electrons must overcome in moving from the n side to the p side may be quite different from the barrier for holes moving from p to n. The depletion region on each side is analogous to that described in Eq. (5-23), except that we must account for the different dielectric constants in the two semiconductors.

To draw the band diagram for a heterojunction accurately, we must not only use the proper values for the band discontinuities but also account for the band bending in the junction. To do this, we must solve Poisson's equation across the heterojunction, taking into account the details of doping and space charge, which generally requires a computer solution. We can, however, sketch an approximate diagram without a detailed calculation. Given the experimental band offsets ΔE_v and ΔE_c, we can proceed as follows:

1. Align the Fermi level with the two semiconductor bands separated. Leave space for the transition region.



Scanned by CamScanner

10.4 SEMICONDUCTOR LASERS

The laser became an important part of semiconductor device technology in 1962 when the first p-n junction lasers were built in GaAs (infrared)[†] and GaAsP (visible).[‡] We have already discussed the incoherent light emission from p-n junctions (LEDs), generated by the spontaneous recombination of electrons and holes injected across the junction. In this section we shall concentrate on the requirements for population inversion due to these injected carriers and the nature of the coherent light from p-n junction lasers. These devices differ from the solid, gas, and liquid lasers discussed previously in several important respects. Junction lasers are remarkably small (typically on the order of $0.1 \times 0.1 \times 0.3$ mm), they exhibit high efficiency, and the laser output is easily modulated by controlling the junction current. Semiconductor lasers operate at low power compared, for example, with ruby or CO_2 lasers; on the other hand, these junction lasers compete with He–Ne lasers in power output. Thus the function of the semiconductor laser is to provide a portable and easily controlled source of low-power coherent radiation. They are particularly suitable for fiber optic communication systems (Section 6.4.2).

10.4.1 Population Inversion at a Junction

If a p-n junction is formed between degenerate materials, the bands under forward bias appear as shown in Fig. 10-11. If the bias (and thus the current) is large enough, electrons and holes are injected into and across the transition region in considerable concentrations. As a result, the region about the junction is far from being depleted of carriers. This region contains a large concentration of electrons within the conduction band and a large concentration of holes within the valence band. If these population densities are high enough, a con-

Scanned by CamScanner

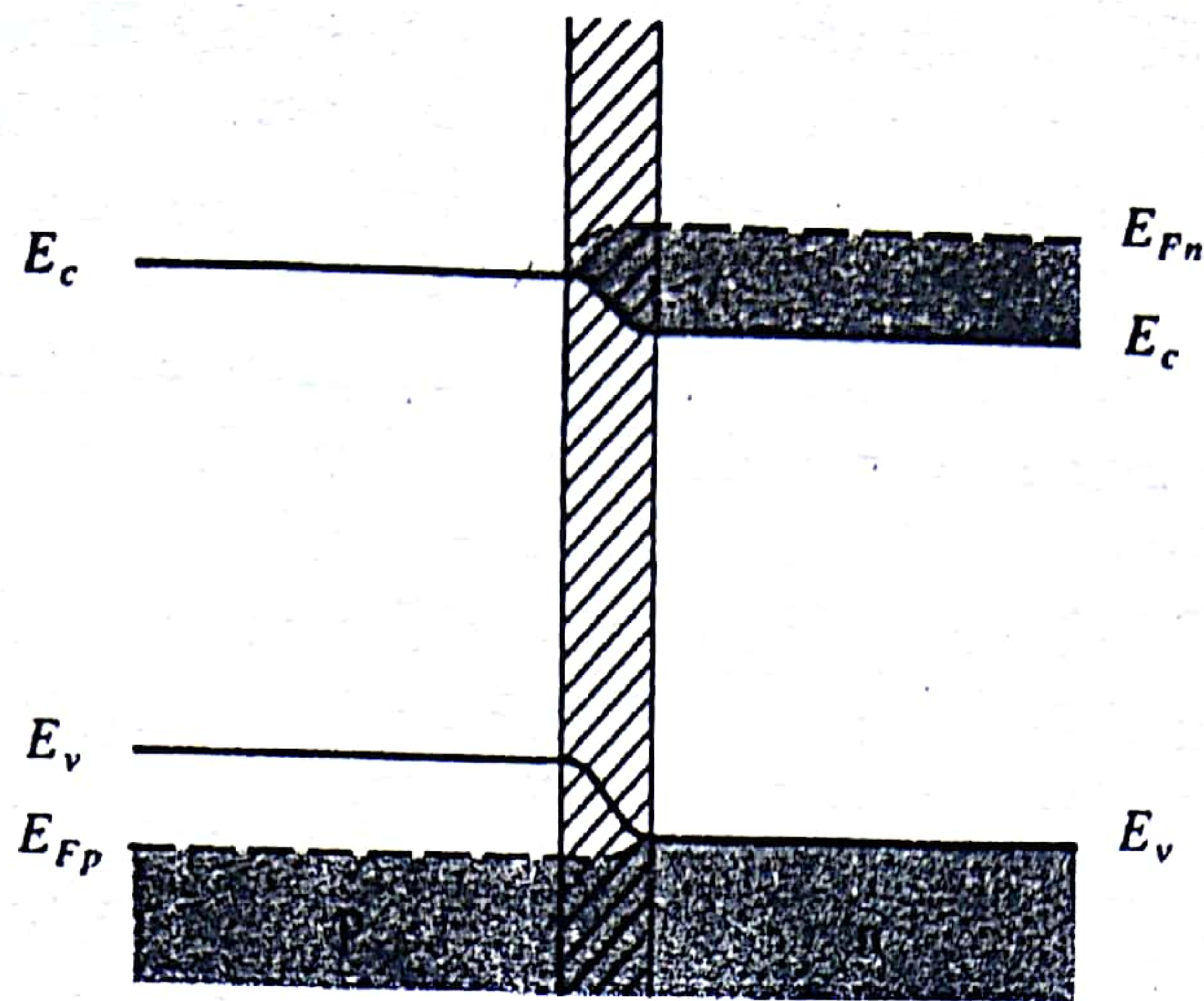


Figure 10-11 Band diagram of a p-n junction laser under forward bias. The cross-hatched region indicates the inversion region at the junction.

dition of population inversion results, and the region about the junction over which it occurs is called an *inversion region*.†

Population inversion at a junction is best described by the use of the concept of *quasi-Fermi levels* (Section 4.3.3). Since the forward-biased condition of Fig. 10-11 is a distinctly nonequilibrium state, the equilibrium equations defining the Fermi level are not applicable. In particular, the concentration of electrons in the inversion region (and for several diffusion lengths into the p material) is larger than equilibrium statistics would imply; the same is also true for the injected holes in the n material. We can use Eqs. (4-15) to describe the carrier concentrations in terms of the quasi-Fermi levels for electrons and holes in steady state. Thus

$$n = N_c e^{-(E_c - F_n)/kT} = n_i e^{(F_n - E_i)/kT} \tag{10-7a}$$

$$p = N_v e^{-(F_p - E_v)/kT} = n_i e^{(E_i - F_p)/kT} \tag{10-7b}$$

Using Eqs. (10-7a) and (10-7b), we can draw F_n and F_p on any band diagram for which we know the electron and hole distributions. For example, in Fig. 10-12, F_n in the neutral n region is essentially the same as the equilibrium Fermi level E_{Fn}. This is true to the extent that the electron concentration on the n side is equal to its equilibrium value. However, since large numbers of electrons are injected across the junction, the electron concentration begins at a

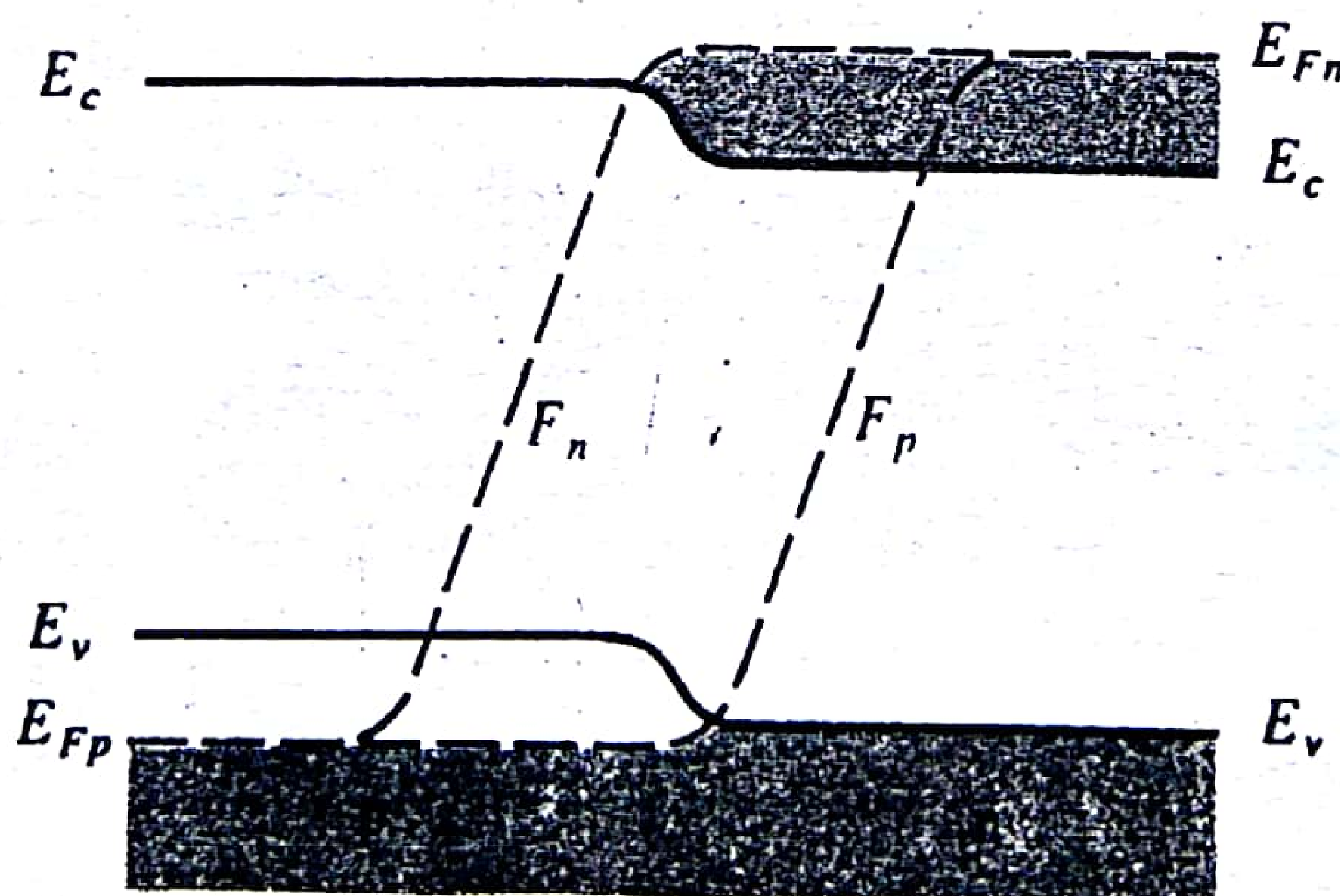


Figure 10-12 Quasi-Fermi levels in a laser junction under forward bias.

†This is a different meaning of the term from that used in reference to MOS transistors.

high value near the junction and decays exponentially to its equilibrium value n_p deep in the p material. Therefore, F_n drops from E_{Fn} as shown in Fig. 10-12. We notice that, deep in the neutral regions, the quasi-Fermi levels are essentially equal. The separation of F_n and F_p at any point is a measure of the departure from equilibrium at that point. Obviously, this departure is considerable in the inversion region, since F_n and F_p are separated by an energy greater than the band gap (Fig. 10-13).

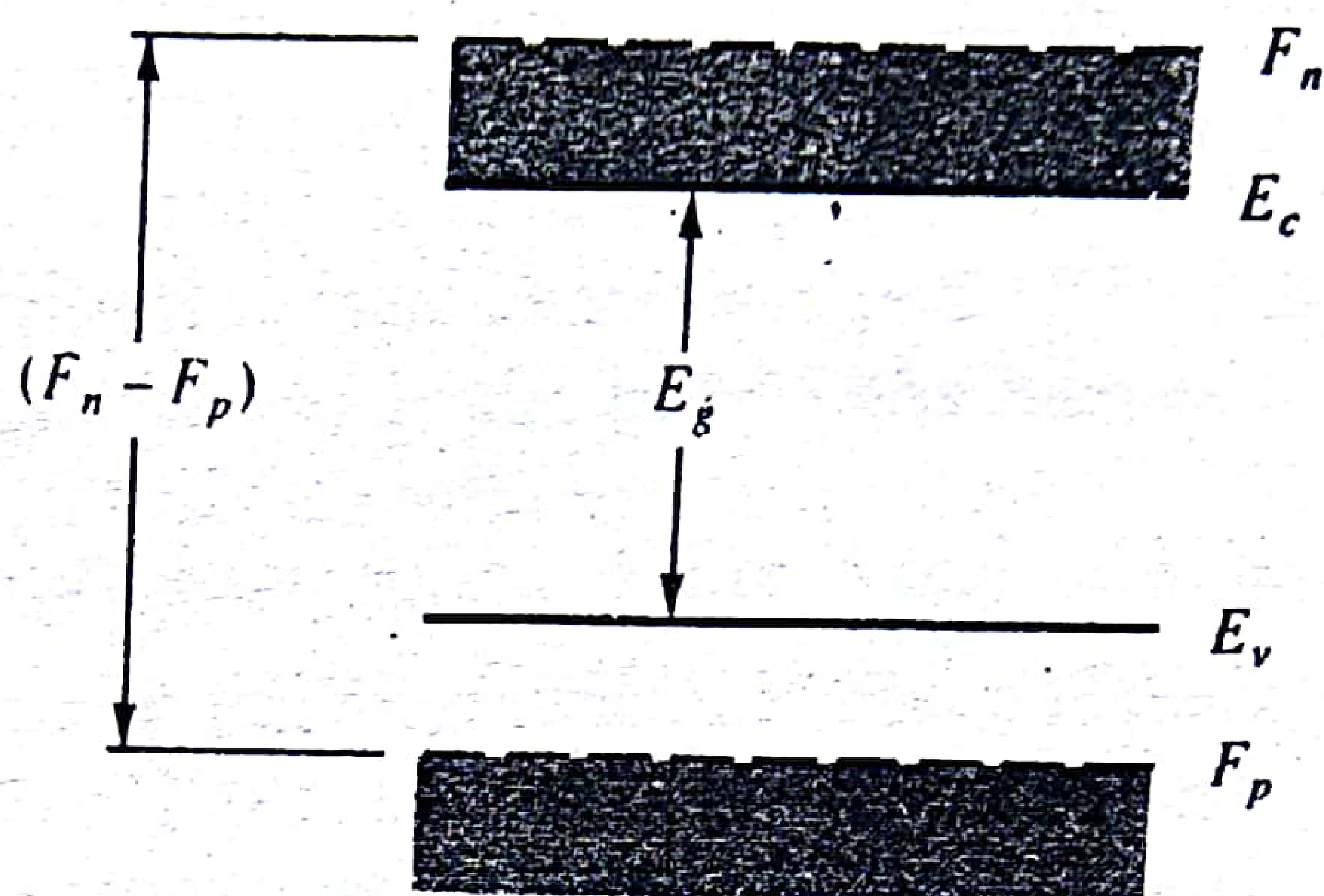


Figure 10-13 Expanded view of the inversion region.

Unlike the case of the two-level system discussed in Section 10.1, the condition for population inversion in semiconductors must take into account the distribution of energies available for transitions between the bands. The basic definition of population inversion holds—for dominance of stimulated emission between two energy levels separated by energy $h\nu$, the electron population of the upper level must be greater than that of the lower level. The unusual aspect of a semiconductor is that bands of levels are available for such transitions. Population inversion obviously exists for transitions between the bottom of the conduction band E_c and the top of the valence band E_v in Fig. 10-13. In fact, transitions between levels in the conduction band up to F_n and levels in the valence band down to F_p take place under conditions of population inversion. For any given transition energy $h\nu$ in a semiconductor, population inversion exists when

$$(F_n - F_p) > h\nu \qquad (10\text{-}8a)$$

For band-to-band transitions, the minimum requirement for population inversion occurs for photons with $h\nu = E_c - E_v = E_g$

$$(F_n - F_p) > E_g \qquad (10\text{-}8b)$$

When F_n and F_p lie within their respective bands (as in Fig. 10-13), stimulated emission can dominate over a range of transitions, from $h\nu = (F_n - F_p)$ to $h\nu = E_g$. As we shall see below, the dominant transitions for laser action are determined largely by the resonant cavity and the strong recombination radiation occurring near $h\nu = E_g$.

In choosing a material for junction laser fabrication, it is necessary that electron-hole recombination occur directly, rather than through trapping processes such as are dominant in Si or Ge. Gallium arsenide is an example of such a "direct" semiconductor. Furthermore, we must be able to dope the material n-type or p-type to form a junction. If an appropriate resonant cavity can

be constructed in the junction region, a laser results in which population inversion is accomplished by the bias current applied to the junction (Fig. 10-14).

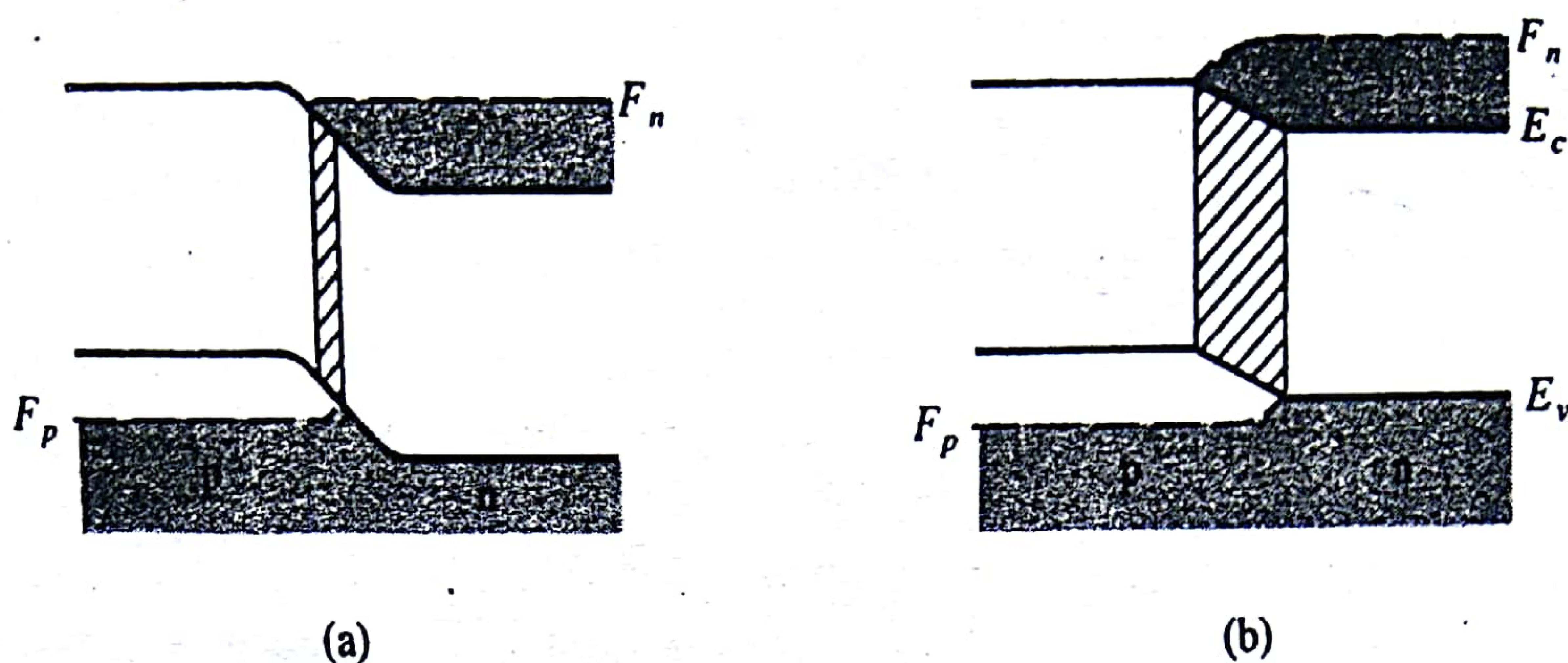


Figure 10-14 Variation of inversion region width with forward bias: $V(a) < V(b)$.

10.4.2 Emission Spectra for p-n Junction Lasers

Under forward bias, an inversion layer can be obtained along the plane of the junction, where a large population of electrons exists at the same location as a large hole population. A second look at Fig. 10-13 indicates that spontaneous emission of photons can occur due to direct recombination of electrons and holes, releasing energies ranging from approximately $F_n - F_p$ to E_g. That is, an electron can recombine over an energy from F_n to F_p, yielding a photon of energy $h\nu = F_n - F_p$, or an electron can recombine from the bottom of the conduction band to the top of the valence band, releasing a photon with $h\nu = E_c - E_v = E_g$. These two energies serve as the approximate outside limits of the laser spectra.

The photon wavelengths which participate in stimulated emission are determined by the length of the resonant cavity as in Eq. (10-5). Figure 10-15 illustrates a typical plot of emission intensity vs. photon energy for a semiconductor laser. At low current levels (Fig. 10-15a), a spontaneous emission spectrum containing energies in the range $E_g < h\nu < (F_n - F_p)$ is obtained. As the current is increased to the point that significant population inversion exists, stimulated emission occurs at frequencies corresponding to the cavity modes as shown in Fig. 10-15b. These modes correspond to successive numbers of integral half-wavelengths fitted within the cavity, as described by

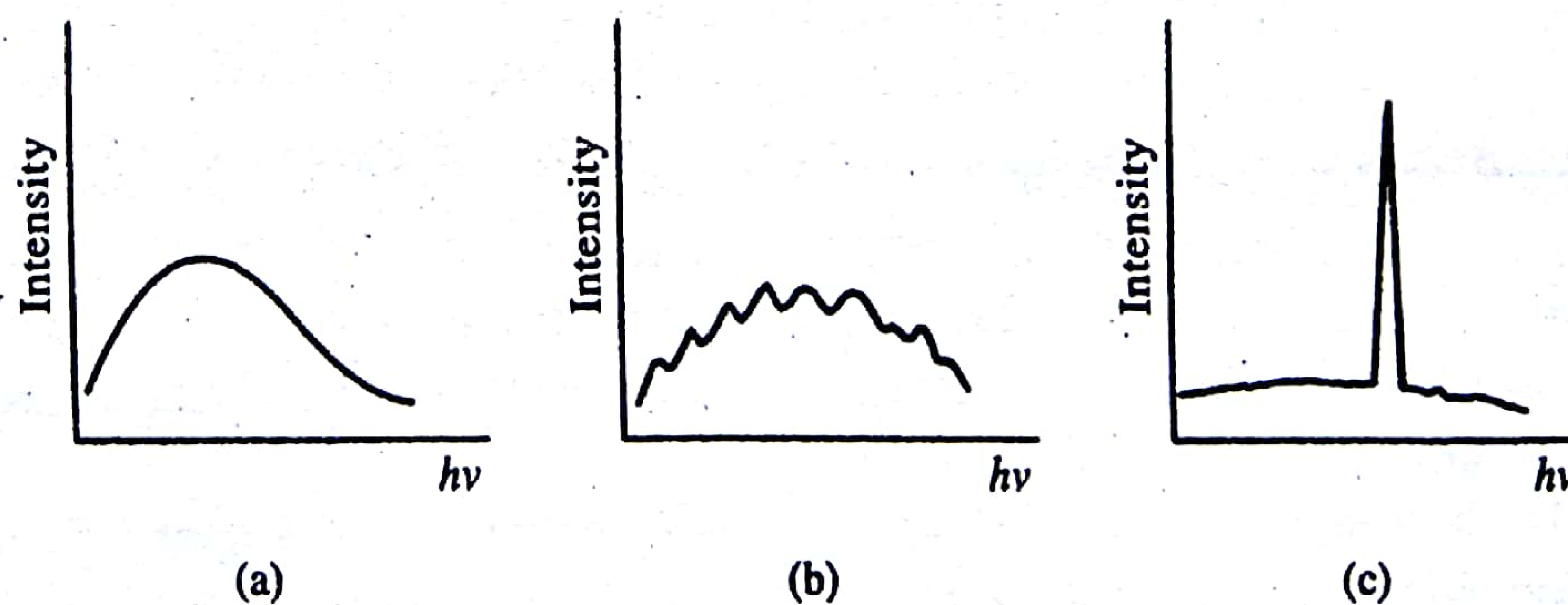


Figure 10-15 Light intensity vs. photon energy $h\nu$ for a junction laser: (a) incoherent emission below threshold; (b) laser modes at threshold; (c) dominant laser mode above threshold. The intensity scales are greatly compressed from (a) to (b) to (c).

Scanned by CamScanner

Eq. (10-5). Finally, at a still higher current level, a most preferred mode or set of modes will dominate the spectral output (Fig. 10-15c). This very intense mode represents the main laser output of the device; the output light will be composed of almost monochromatic radiation superimposed on a relatively weak radiation background, due primarily to spontaneous emission.

The separation of the modes in Fig. 10-15b is complicated by the fact that the index of refraction $\mathbf{n}$ for GaAs depends on wavelength λ. From Eq. (10-5) we have

$$\mathbf{m} = \frac{2L\mathbf{n}}{\lambda_0} \quad (10\text{-}9)$$

If $\mathbf{m}$ (the number of half-wavelengths in L) is large, we can use the derivative to find its rate of change with λ_0:

$$\frac{d\mathbf{m}}{d\lambda_0} = -\frac{2L\mathbf{n}}{\lambda_0^2} + \frac{2L}{\lambda_0}\frac{d\mathbf{n}}{d\lambda_0} \quad (10\text{-}10)$$

Now reverting to discrete changes in $\mathbf{m}$ and λ_0, we can write

$$-\Delta\lambda_0 = \frac{\lambda_0^2}{2L\mathbf{n}}\left(1 - \frac{\lambda_0}{\mathbf{n}}\frac{d\mathbf{n}}{d\lambda_0}\right)^{-1}\Delta\mathbf{m} \quad (10\text{-}11)$$

If we let $\Delta\mathbf{m} = -1$, we can calculate the change in wavelength $\Delta\lambda_0$ between adjacent modes (i.e., between modes $\mathbf{m}$ and $\mathbf{m} - 1$).

10.4.3 The Basic Semiconductor Laser

To build a p-n junction laser, we need to form a junction in a highly doped, direct semiconductor (GaAs, for example), construct a resonant cavity in the proper geometrical relationship to the junction, and make contact to the junction in a mounting which allows for efficient heat transfer. A simple fabrication technique is outlined in Fig. 10-16. Beginning with a degenerate n-type sample, a p region is formed on one side, for example by diffusing Zn into the

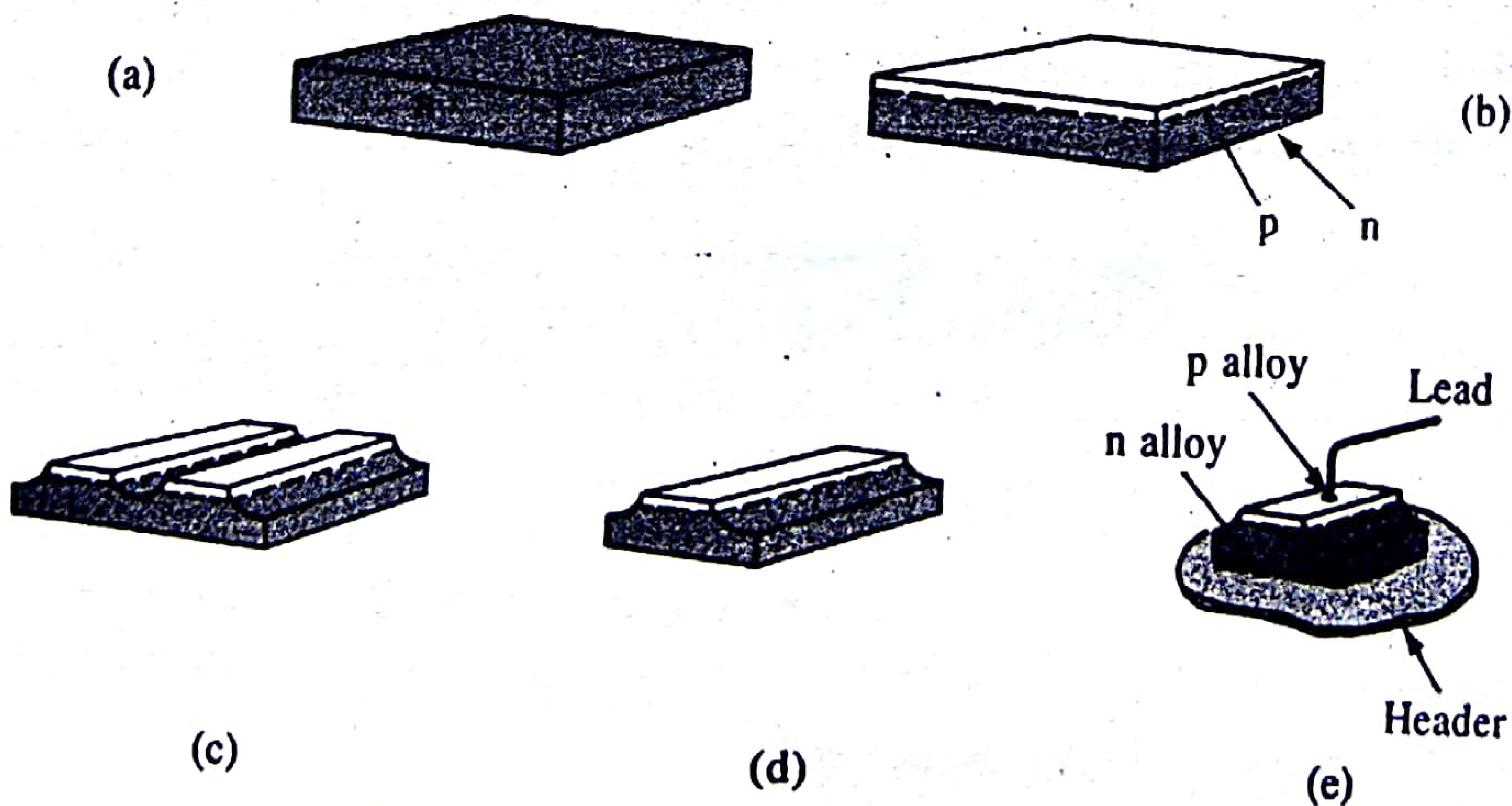


Figure 10-16 Fabrication of a simple junction laser: (a) degenerate n-type sample; (b) diffused p layer; (c) isolation of junctions by cutting or etching; (d) individual junction to be cut or cleaved into devices; (e) mounted laser structure.

Scanned by CamScanner

n-type GaAs. Since Zn is in column II of the periodic table and is introduced substitutionally on Ga sites, it serves as an acceptor in GaAs; therefore, the heavily doped Zn diffused layer forms a p^+ region (Fig. 10-16b). At this point we have a large-area planar p-n junction. Next, grooves are cut or etched along the length of the sample as in Fig. 10-16c, leaving a series of long p regions isolated from each other. These p-n junctions can be cut or broken apart (Fig. 10-16d) and then cleaved into devices of the desired length.

At this point in the fabrication process, the very important requirement of a resonant cavity must be considered. It is necessary that the front and back faces (Fig. 10-16e) be flat and parallel. This can be accomplished by cleaving. If the sample has been oriented so that the long junctions of Fig. 10-16d are perpendicular to a crystal plane of the material, it is possible to cleave the sample along this plane into laser devices, letting the crystal structure itself provide the parallel faces. The device is then mounted on a suitable header, and contact is made to the p region. Various techniques are used to provide adequate heat sinking of the device for large forward current levels.

10.4.4 Heterojunction Lasers

The device described above was the first type used in the early development of semiconductor lasers. Since the device contains only one junction in a single type of material, it is referred to as a *homojunction* laser. To obtain more efficient lasers, and particularly to build lasers that operate at room temperature, it is necessary to use multiple layers in the laser structure. Such devices, called *heterojunction lasers*, can be made to operate continuously at room temperature to satisfy the requirements of optical communications. An example of a heterojunction laser is shown in Fig. 10-17. In this structure the injected carriers are confined to a narrow region so that population inversion can be built up

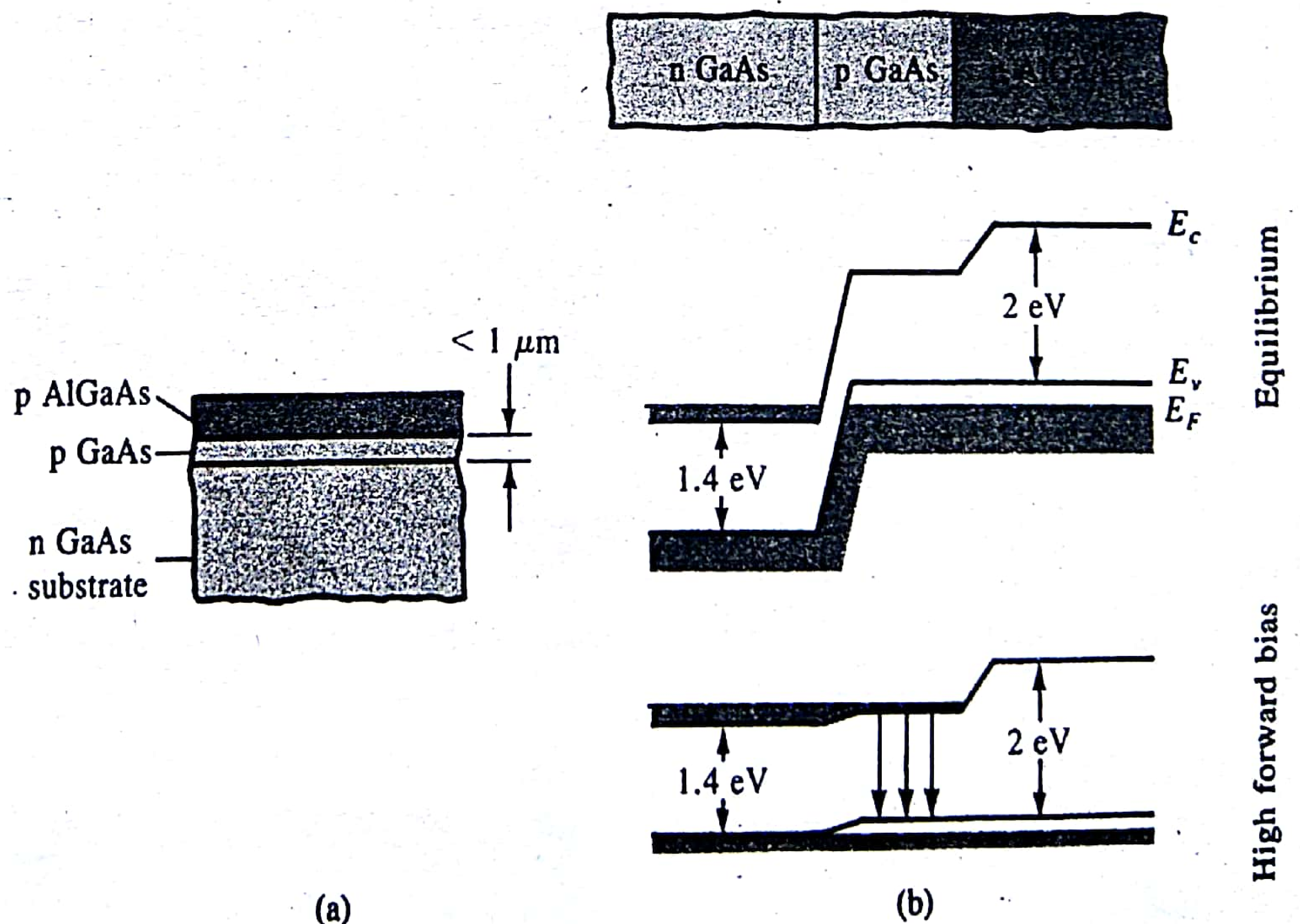


Figure 10-17 Use of a single heterojunction for carrier confinement in laser diodes: (a) AlGaAs heterojunction grown on the thin p-type GaAs layer; (b) band diagrams for the structure of (a), showing confinement of electrons to the thin p region under bias.

at lower current levels. The result is a lowering of the *threshold current* at which laser action begins. Carrier confinement is obtained in this single-heterojunction laser by the layer of AlGaAs grown epitaxially on the GaAs.

In GaAs the laser action occurs primarily on the p side of the junction due to a higher efficiency for electron injection than for hole injection. In a normal p-n junction the injected electrons diffuse into the p material such that population inversion occurs for only part of the electron distribution near the junction. However, if the p material is narrow and terminated in a barrier, the injected electrons can be confined near the junction. In Fig. 10-17a, an epitaxial layer of p-type AlGaAs ($E_g \simeq 2$ eV) is grown on top of the thin p-type GaAs region. The wider band gap of AlGaAs effectively terminates the p-type GaAs layer, since injected electrons do not surmount the barrier at the GaAs-AlGaAs heterojunction (Fig. 10-17b). As a result of the confinement of injected electrons, laser action begins at a substantially lower current than for simple p-n junctions. In addition to the effects of carrier confinement, the change of refractive index at the heterojunction provides a waveguide effect for optical confinement of the photons.

A further improvement can be obtained by sandwiching the active GaAs layer between two AlGaAs layers (Fig. 10-18). This *double-heterojunction* structure further confines injected carriers to the active region, and the change in refractive index at the GaAs–AlGaAs boundaries helps to confine the generated light waves. In the double-heterojunction laser shown in Fig. 10-18b the injected current is restricted to a narrow stripe along the lasing direction, to re-

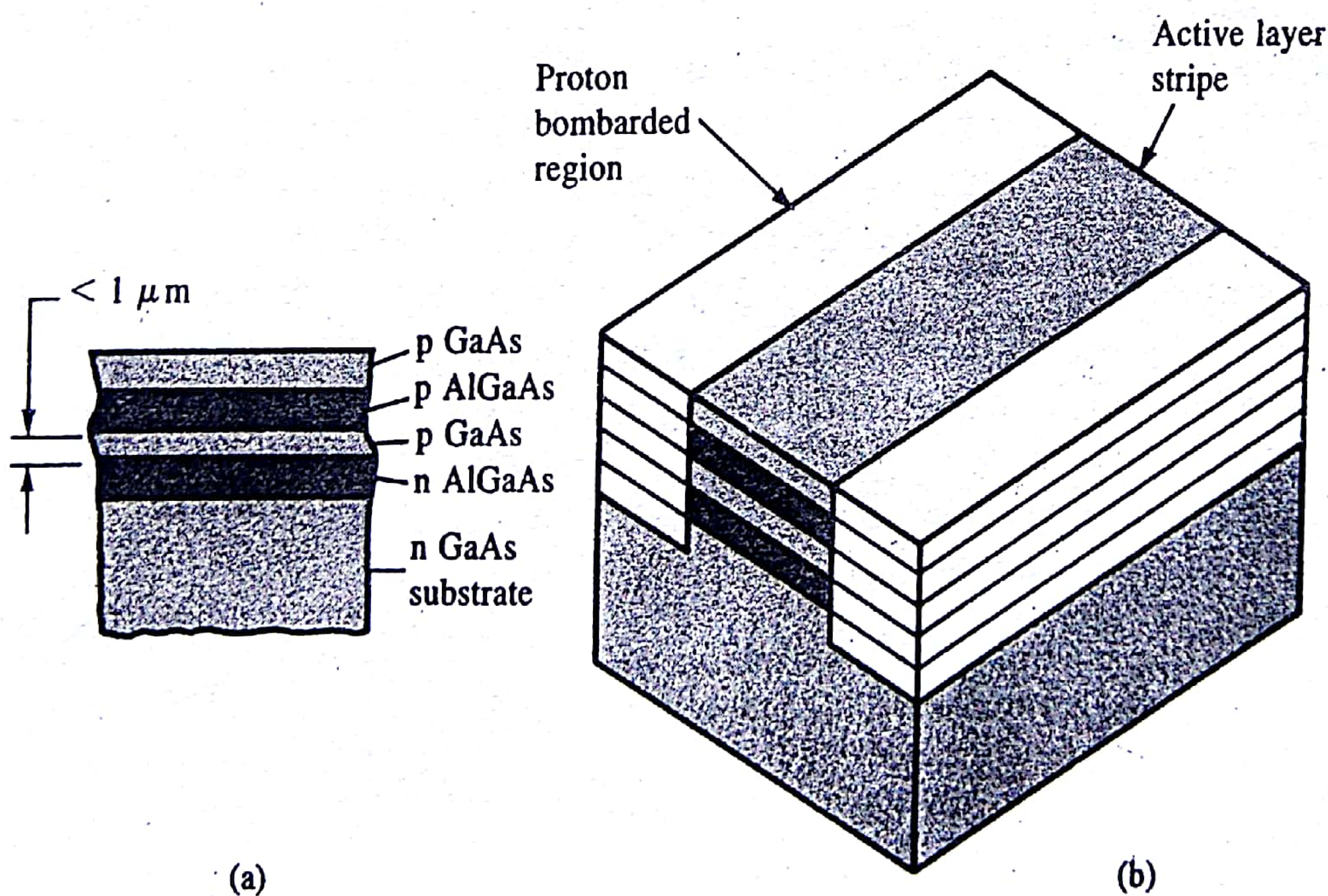


Figure 10-18

A double-heterojunction laser structure: (a) multiple layers used to confine injected carriers and provide waveguiding for the light; (b) a stripe geometry designed to restrict the current injection to a narrow stripe along the lasing direction. One of many methods for obtaining the stripe geometry, this example is obtained by proton bombardment of the shaded regions in (b), which converts the GaAs and AlGaAs to semi-insulating form.

Scanned by CamScanner

Indirect Recombination; Trapping

In column IV semiconductors and in certain compounds, the probability of direct electron-hole recombination is very small (Appendix III). There is some band gap light given off by materials such as Si and Ge during recombination, but this radiation is very weak and may be detected only by sensitive equipment. The vast majority of the recombination events in indirect materials occur via *recombination levels* within the band gap, and the resulting energy loss by recombining electrons is usually given up to the lattice as heat rather than by the emission of photons. Any impurity or lattice defect can serve as a recombination center if it is capable of receiving a carrier of one type and subsequently capturing the opposite type of carrier, thereby annihilating the pair. For example, Fig. 4-8 illustrates a recombination level E_r which is below E_F at equilibrium and therefore is substantially filled with electrons. When excess electrons and holes are created in this material, each EHP recombines at E_r in two steps: (a) hole capture and (b) electron capture.

Since the recombination centers in Fig. 4-8 are filled at equilibrium, the first event in the recombination process is hole capture. It is important to note that this event is equivalent to an electron at E_r falling to the valence band, leaving behind an empty state in the recombination level. Thus in hole capture, energy is *given up* as heat to the lattice. Similarly, energy is given up when a conduction band electron subsequently falls to the empty state in E_r. When both of these events have occurred, the recombination center is back to its original state (filled with an electron), but an EHP is missing. Thus one EHP recombination has taken place, and the center is ready to participate in another recombination event by capturing a hole.

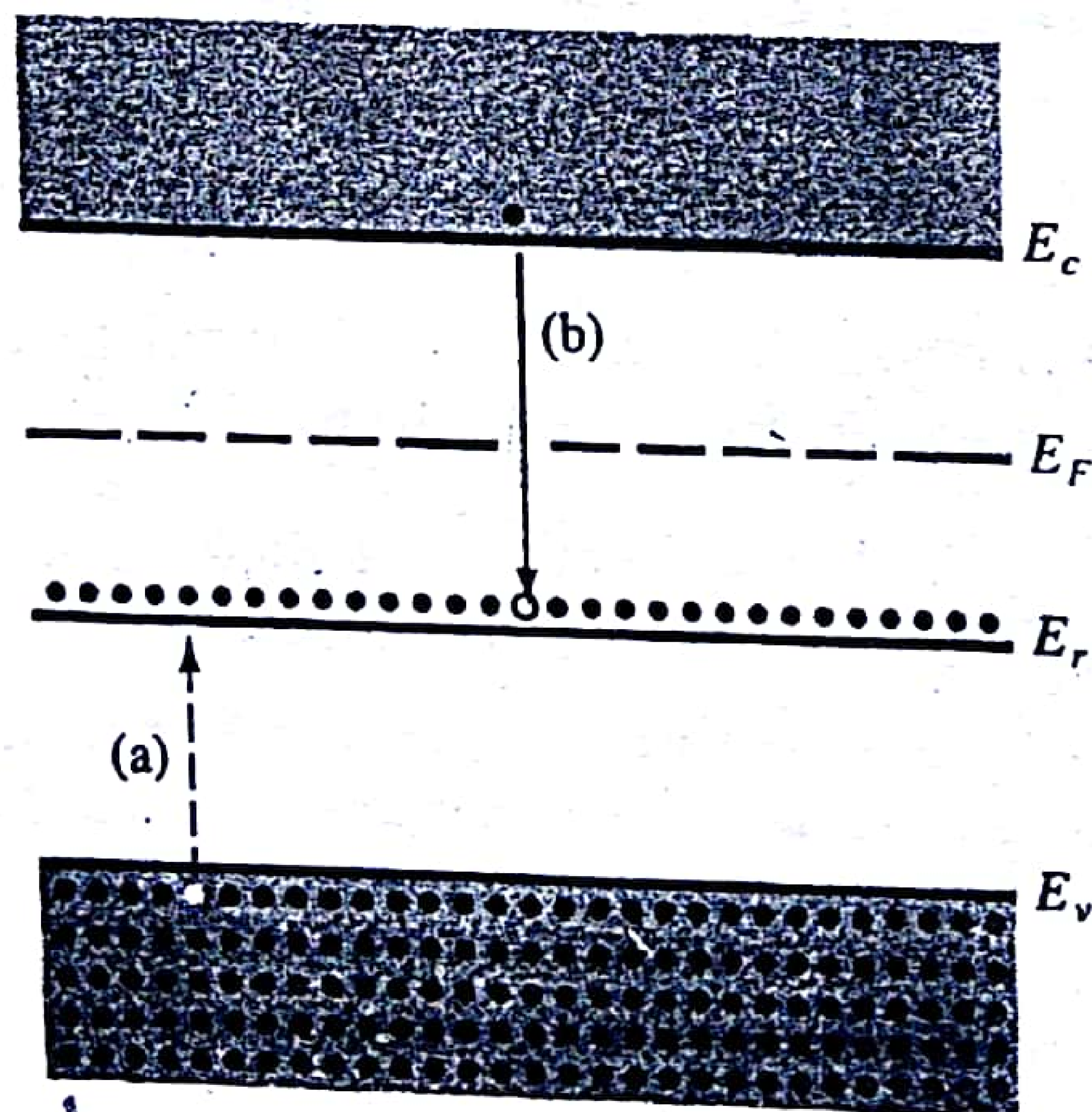
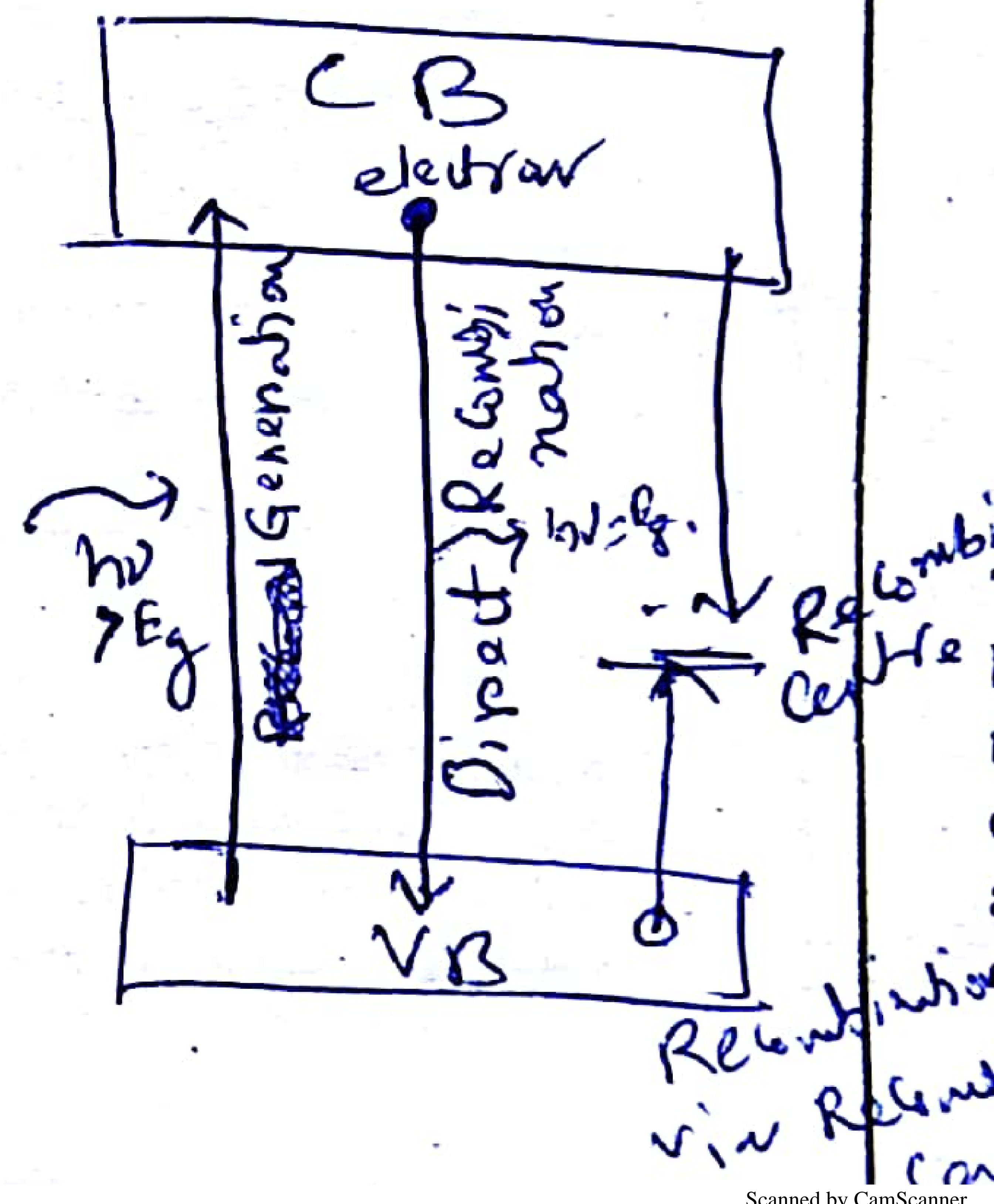


Figure 4-8 Capture processes at a recombination level: (a) hole capture at a filled recombination center; (b) electron capture at an empty center.



The carrier lifetime resulting from indirect recombination is somewhat more complicated than is the case for direct recombination, since it is necessary to account for unequal times required for capturing each type of carrier. In particular, recombination is often delayed by the tendency for a captured carrier to be thermally reexcited to its original band before capture of the opposite type of carrier can occur (Section 4.2.1). For example, if electron capture (b) does not follow immediately after hole capture (a) in Fig. 4-8, the hole may be thermally reexcited to the valence band. Energy is required for this process, which is equivalent to a valence band electron being raised to the empty state in the recombination level. This process delays the recombination, since the hole must be captured again before recombination can be completed.

When a carrier is trapped temporarily at a center and then is reexcited without recombination taking place, the process is often called *temporary trapping*. Although the nomenclature varies somewhat, it is common to refer to an impurity or defect center as a *trapping center* (or simply *trap*) if, after capture of one type of carrier, the most probable next event is reexcitation. If the most probable next event is capture of the opposite type of carrier, the center is predominately a recombination center. The recombination can be slow or fast, depending on the average time the first carrier is held before the second carrier is captured. In general, trapping levels located deep in the band gap are slower in releasing trapped carriers than are the levels located near one of the bands. This results from the fact that more energy is required, for example, to reexcite a trapped electron from a center near the middle of the gap to the conduction band than is required to reexcite an electron from a level closer to the conduction band.

As an example of impurity levels in semiconductors, Fig. 4-9 shows the energy level positions of various impurities in Si. In this diagram a superscript indicates whether the impurity is positive (donor) or negative (acceptor) when ionized. Some impurities introduce multiple levels in the band gap; for example, Zn introduces a level (Zn^-) located 0.31 eV above the valence band and a second level ($Zn^=$) near the middle of the gap. Each Zn impurity atom is capable of accepting two electrons from the semiconductor, one in the lower level and then one in the upper level.

The effects of recombination and trapping can be measured by a *photoconductive decay* experiment. As Fig. 4-7 shows, a population of excess electrons and holes disappears with a decay constant characteristic of the particular recombination process. The conductivity of the sample during the decay is

$$\sigma(t) = q[n(t)\mu_n + p(t)\mu_p] \tag{4-9}$$

Therefore, the time dependence of the carrier concentrations can be monitored by recording the sample resistance as a function of time. A typical experimental arrangement is shown schematically in Fig. 4-10. A source of short pulses of light is required, along with an oscilloscope for displaying the sample voltage as the resistance varies. Microsecond light pulses can be obtained by periodically discharging a capacitor through a flash tube containing a gas such as

Scanned by CamScanner