4 Optoelectronics

4.1 Executive summary

All land line communication systems currently use optical fibres as the channel due to their very low attenuation, and deployed systems operate up to 40 Gigabits per second (Gbps), with experimental systems at 160 Gbps and higher. These high speed systems use very stable sources, usually the distributed feedback lasers, with external modulators and very fast detectors. Direct modulation of lasers leads to chirp, caused by the laser frequency varying when modulated, and therefore is not used in these systems. However, interest remains in high speed lasers for several other applications. This chapter discusses high speed optoelectronic devices which include light-emitting diodes (LEDs), semiconductor lasers, photoconductors, p–i–n diode photodetectors, avalanche photodetectors, metal–semiconductor–metal photodetectors, travelling wave photodetectors and briefly the phototransistor. The physics of the devices are outlined and then the parameters that make these devices fast are discussed with specific examples from the literature.

4.2 Optical sources

Two types of sources are widely used in optical communication systems, and they are the LED and the semiconductor laser. LEDs are inherently slow; their response times are determined by the lifetime of the carriers in the active region, and in most LEDs this is between 2 ns and 10 ns. Thus, small-signal response of LEDs is of the order of 100 MHz, and the large-signal response is smaller than this. Lasers are complex devices, and the response time is determined by the so-called *relaxation oscillation frequency*, and this is due to the interaction between the photons which have a finite lifetime, which is a measure of the cavity quality factor, Q, the differential gain of the structure, the lifetime of the carriers and other parameters [1]. Most lasers, when directly modulated, suffer from spectral broadening, and therefore direct modulation is not used for long haul fibre-optic systems; but when this broadening is not important, direct modulation may be used.

In this section, the fundamental aspects of LEDs and lasers are discussed, followed by an outline of the different types of lasers. To calculate the response of lasers, the rate equations are introduced and these are solved analytically for the relaxation frequency f_r. A short discussion of the noise mechanisms of lasers follows. A further section is included on the very high f_r lasers.

4.2.1 Preliminaries

The generation of non-thermal light in a semiconductor requires the creation of hole–electron pairs which recombine radiatively to emit photons at the bandgap energy, or impurity level to valence band energy. The more recent source of light is the intra-band relaxation of electrons from a higher energy level to a lower energy level in the conduction band of a quantum well, to emit photons, and this has resulted in the quantum cascade lasers. In both these sources of light, a means of generation of the excess electron–hole pairs or excitation of the carriers to a higher energy state requires the expenditure of energy. If this energy is from an electrical source, then this is called *electrical pumping* leading to electroluminescent emission. Alternatively, optical pumping with photons of a higher energy than the emitted photons is also a possibility. In this chapter, only electroluminescent sources are considered, and quantum cascade lasers are not discussed.

4.2.2 Light-emitting diodes

The ubiquitous green and red LEDs are fabricated in GaP which is an indirect gap semiconductor. The impurities of zinc oxide and nitrogen form isoelectronic bands which give rise to the red and green emissions, respectively. Other LEDs include the blue GaN device, the near infrared GaAs 870 nm device and the InGaAsP 1300 nm and 1550 nm devices.

The mechanisms that create the LED are based on p–n junctions. A forward-biased p–n junction injects minority carriers to both sides of the junction, and these diffuse away from the junction to recombine radiatively and non-radiatively. The radiative recombination results in emission of photons, with energy $h\nu$ approximately equal to the bandgap energy E_g. The recombination is spontaneous, which implies that emission has random phase, and the linewidth is of the order of a few k_BT, where k_B is Boltzmann's constant, and T is the temperature in kelvin, and the emission is incoherent. The internal quantum efficiency is a measure of how efficiently the injected carriers produce light, since some of them recombine non-radiatively. Define the total recombination rate R_{total}, in recombination events per second, as the sum of the radiative recombination rate R_{rr} and the non-radiative recombination rate R_{nr}. Then, the internal quantum efficiency is given by the ratio of the radiative recombination rate to the total rate:

$$\eta_i = \frac{R_{\text{rr}}}{R_{\text{total}}} = \frac{R_{\text{rr}}}{R_{\text{rr}} + R_{\text{nr}}}. \tag{4.1}$$

However, lifetimes are more easily measured or inferred, and the quantum efficiency is recast in terms of the lifetimes. The radiative recombination rate is defined as the ratio of the non-equilibrium carrier density N to the radiative lifetime τ_{rr}, and similarly, the non-radiative recombination rate is the ratio of the carrier density N to the non-radiative lifetime τ_{nr}, and thus $R_{\text{rr}} = N/\tau_{\text{rr}}$ and $R_{\text{nr}} = N/\tau_{\text{nr}}$. It follows that the total lifetime τ_{total} is given by

$$\frac{1}{\tau_{\text{total}}} = \frac{1}{\tau_{\text{rr}}} + \frac{1}{\tau_{\text{nr}}}. \tag{4.2}$$

Hence, substituting in the Equation (4.1), the internal quantum efficiency is given by

$$\eta_i = \frac{\tau_{nr}}{\tau_{nr} + \tau_{rr}}. \tag{4.3}$$

The total lifetime is obtained from the definition

$$R_{nr} + R_{rr} = R_{total} = \frac{N}{\tau_{total}}, \tag{4.4}$$

and the total recombination rate R_{total} is given by

$$R_{total} = A_{nr}N + B_r N^2, \tag{4.5}$$

where A_{nr} is the non-radiative recombination coefficient and B_r is the radiative recombination coefficient. It follows that

$$\tau_{total} = (A_{nr} + B_r N)^{-1}. \tag{4.6}$$

In direct gap semiconductors, the radiative lifetime is comparable to the non-radiative lifetime, whereas in indirect gap semiconductors, the non-radiative lifetime is much shorter than the radiative lifetime. Thus, the internal quantum efficiency is about 0.5 for direct gap semiconductors and this improves very considerably for lasers to almost unity when the radiative lifetime becomes very small compared to the non-radiative lifetime. In indirect gap semiconductors, silicon and germanium for example, the internal quantum efficiency is of the order of 10^{-5}. Surface recombination is also a problem as this is non-radiative and decreases the internal quantum efficiency. To reduce this, a heterojunction layer of a higher band gap may be deposited above the emission layer. The other problem is the absorption of the emitted photons in the generating layer, and there is little that can be done about this. The higher bandgap layer has very low absorption, and this heterolayer may help marginally.

Consider the structure of the LED which is usually a p–n or p^+–n or n^+–p junction. It is necessary to have the junction close to the surface so that the emitted light is able to escape from the material into the air. Initially, consider the carrier injection process in a p–n junction when it is forward-biased. A p–n junction when forward-biased injects holes into the n region and electrons in the p region. With forward bias of V_f, it can be shown that the current density in the p–n junction is given by

$$J = q\left(\frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n}\right)(e^{qV_f/k_B T} - 1), \tag{4.7}$$

where D_p and D_n are the minority hole and minority electron diffusion constants in the n-type and p-type material respectively adjacent to the junction; L_p and L_n are the diffusion lengths of the minority holes and minority electrons in the n- and p-type material respectively adjacent to the junction; p_{n0} and n_{p0} are the minority hole and electron densities in the n- and p-type material; q is the charge magnitude of an electron; k_B is Boltzmann's constant; T is the temperature in kelvin and $L_{p,n} = \sqrt{\tau_{p,n} D_{p,n}}$. In some LEDs, the junction is n^+–p, in which case p_{n0} in the n^+ region is very small, as it is equal to n_i^2/n^+, and in many III–V materials n_i, the intrinsic density, is very low, for example in GaAs it is about $10^6\ cm^{-3}$. The hole injection term which is the first term in

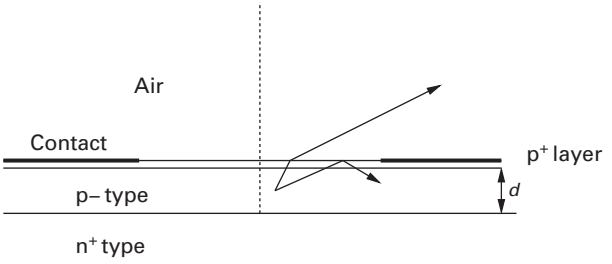

Fig. 4.1 Sketch of the n^+–p junction, with a very thin p^+ and the contacts of the LED. Also shown is the effect of total internal reflection in the escape of light from the p^+ layer.

Equation (4.7) is small, and the injection is largely from the n^+ region into the p region. Thus, the injection efficiency, which is the ratio of the injected current into the p region to the total injected current is almost unity due to the fact that the injected current into the n^+ region is much smaller than that injected into the p region. This is the case for the p^+–n junction as well, since the minority carrier density in the highly doped region is always much smaller than that in the lower doped region, and the injection efficiency is assumed to be unity. However, as indicated above, the internal quantum efficiency is of the order of 50% for direct gap semiconductors. As the junction temperature rises, the non-radiative recombination increases and so this figure is typical for room temperature devices.

Assuming that the injection is primarily from the n^+ side of the n^+–p junction into the p region. The minority electron distribution in the p layer is of the form $\Delta n e^{-z/L_n}$, where the z direction is normal to the surface in Figure 4.1, and z is the distance from the junction towards the surface. This minority carrier distribution may be shown to be equivalent to $n_{p0}(e^{qV/k_BT} - 1)e^{-z/L_p}$. Thus, the radiative recombination is over this distribution, and for all practical purposes it can be shown that this is equivalent to an uniform distribution over the distance L_n. Assume that the thickness of the active layer is d, which is assumed to be much larger than the diffusion length L_n. Then, integrating this distribution

$$\int_0^\infty \Delta n e^{-z/L_n} dz = \Delta n L_n, \tag{4.8}$$

which shows that Δn may be considered uniformly distributed over L_n. The choice of the upper limit of ∞ is because $d \gg L_n$. The lifetimes in GaAs and InGaAsP vary from 2 ns to about 10 ns, and mobility of minority carriers of 1000 cm^2 $(V.s)^{-1}$ gives L_n values of less than 0.5 μm, and the active layer thickness may exceed this value. In the following sections, the tacit assumption is that the minority carrier density is uniform over L_n.

In Figure 4.1, the top p layer has the injected minority carriers recombining to create the photon source. The spontaneous emission is isotropic and the trajectories of the photons at the surface to the air region result in total internal reflection when the incidence angle exceeds the critical angle. Thus, extraction efficiency of the emission needs to take into account this total internal reflection, when the incidence angle exceeds the critical angle given by

$$\theta_c = \sin^{-1} \frac{n_0}{n_2} \tag{4.9}$$

where n_0 is the index of air equal to 1 and n_2 is the index of the top layer of the LED through which the emitted photons escape. For normal incidence of the photons, the reflection coefficient Γ is given by

$$\Gamma = \frac{n_0 - n_2}{n_0 + n_2}. \tag{4.10}$$

The fractional transmitted power or the transmissivity $(1 - \Gamma^2)$ for normal incidence, and is given by

$$T(0) = \frac{4n_0 n_2}{(n_0 + n_2)^2}. \tag{4.11}$$

As the photon angle of incidence varies from the normal to θ_c, the reflection coefficient changes, depending on the polarisation, which averages out between the two perpendicular and parallel cases for the spontaneous emission, and the transmissivity also changes [7]. The external quantum efficiency is obtained from

$$\eta_{ext} = \frac{1}{4\pi} \int_0^{\theta_c} T(\theta) 2\pi \sin\theta d\theta. \tag{4.12}$$

Since the expressions for the transmissivity vary with θ, and are difficult to integrate, $T(\theta)$ is replaced by $T(0)$. Substituting for $T(0)$ from Equation (4.11) and using Equation (4.9), the external quantum efficiency becomes

$$\eta_{ext} = \frac{1}{n_2(n_0 + n_2)^2}. \tag{4.13}$$

For a value of n_2 of 3.5 and n_0 of unity, η_{ext} is of the order of 1.4%, which suggests that most of the light generated is trapped inside the device. The presence of a heterojunction layer on the top surface, discussed below, complicates this expression as the index of this top layer is lower than that of the active layer. An anti-reflection coating on the surface helps to improve this factor considerably. The obvious method of extracting light from the p-layer is to place a hemispherical lens, index identical to the p-layer, on the surface. If there is a hetero-layer, then this needs to modified further. In the case of the p layer, the extraction efficiency becomes very high, but the problem is getting a suitable lens with the same index.

The optical power emitted by the LED is given by

$$P_{opt} = \eta_{int} \eta_{ext} (h\nu) \frac{I}{q}, \tag{4.14}$$

where I/q is the number of electrons or holes that are injected into the active region per second; the internal quantum efficiency η_{int} defines the fraction of them that recombine radiatively; and the external quantum efficiency η_{ext} the fraction of the generated photons that escape from the active layer. The total quantum efficiency of the LED is a measure of its performance and is the ratio of the output optical power to the input

electrical power, which is given as $P_0 = V_0 I$, where V_0 is the voltage drop across the device and the current is given by I. Thus, substituting from Equation (4.14)

$$\eta_{\text{total}} = \eta_{\text{int}} \eta_{\text{ext}} \frac{h\nu}{qV_0}. \tag{4.15}$$

Neglecting contact resistance drops, $h\nu \approx E_g \approx qV_0$, where E_g is the band gap in eV, which then makes the total quantum efficiency, which is also the external power efficiency or the wall plug efficiency:

$$\eta_{\text{total}} \approx \eta_{\text{int}} \eta_{\text{ext}}. \tag{4.16}$$

This is usually less than a few percent, unless other techniques are used to extract the light more efficiently. Edge-emitting LEDs, discussed below, are therefore much more efficient. For visible LEDs the luminosity is also an issue [3], but this is not considered here.

The responsivity R of the LED is defined as the ratio of the emitted optical power to the current, and substituting from Equation (4.14), is given by

$$R = \eta_{\text{int}} \eta_{\text{ext}} \frac{h\nu}{q}. \tag{4.17}$$

This is of the order of $0.01\,\text{W A}^{-1}$ for the above values of η unless this becomes much larger.

The spectral width of the emission $\Delta\nu$ of LEDs is approximately defined by $k_B T/qh$ in Hz and peaks at $(E_g + k_B T/2q)/h$ [1]. The full width half maximum (FWHM) is $\sim 1.8 k_B T/qh$, and at room temperature ($T = 300\,\text{K}$) is about 11 THz. The spectral width in wavelength $\Delta\lambda$ varies as $\Delta\lambda = \Delta\nu(\lambda^2/c)$, and so varies from about 30 nm to 90 nm.

Modulation response

The LED is effectively a n^+–p junction, and therefore the usual diode current relationship holds, as discussed above:

$$I = I_0(e^{qV/k_B T} - 1) \tag{4.18}$$

To calculate the modulation response, the small-signal behaviour is obtained by initially biasing the LED at some bias point, where the current is given by I_b. The carrier lifetime in the active layer determines the modulation rate of these LEDs. If N is the density of carriers and I is the current that flows into the LED, the following rate equation determines the carrier dynamics:

$$\frac{dN}{dt} = \frac{I}{q\text{Vol}} - \frac{N}{\tau_{\text{total}}} \tag{4.19}$$

where Vol is the volume of the active region. Under steady-state conditions, the time dependence is zero and at a bias current of I_b, then

$$N_b = \frac{I_b \tau_{\text{total}}}{q\text{Vol}}, \tag{4.20}$$

where N_b is the carrier density at bias current I_b. To obtain the small-signal response, let the current have an ac component given by $I_m e^{j\omega_m t}$, and similarly the carrier density also has a small-signal term $N_m e^{j\omega_m t}$. Note that it is assumed that $I_m \ll I_b$ the bias current, and similarly $N_m \ll N_b$. Thus

$$I(t) = I_b + I_m e^{j\omega_m t} \tag{4.21}$$

$$N(t) = N_b + N_m e^{j\omega_m t}. \tag{4.22}$$

Substituting in Equation (4.19) two component equations arise; the steady-state equation is satisfied by the result in Equation (4.20). The time varying equation gives rise to the following solution

$$N_m(\omega_m) = \frac{\dfrac{I_m \tau_{total}}{(q\,\mathrm{Vol})}}{1 + j\omega_m \tau_{total}}. \tag{4.23}$$

The expression for optical power output given in Equation (4.14) varies as I/q, which is proportional to the carrier density. Thus, the modulated optical power $P_m(\omega)$ varies as N_m. It follows that

$$P_m(\omega_m) = P_{opt} \frac{1}{\left[1 + (\omega_m^2 \tau_{total}^2)\right]^{1/2}}, \tag{4.24}$$

where P_{opt} is the zero frequency steady-state output power. The frequency at which half power is obtained by setting the denominator of the above equation to 2, which leads to

$$f_{3dB} = \sqrt{3} \frac{1}{2\pi\,\tau_{total}}. \tag{4.25}$$

This equation confirms that the bandwidth is inversely proportional to the carrier lifetime.

LED structures

The basic planar LED is shown in Figure 4.2. In general, it is necessary to have a thin highly doped contact layer to reduce the contact resistance, and the emerging light has to pass through it. The absorption creates additional loss. If the active layer is intrinsic or undoped, with p^+ and n^+ layers on either side, then the injection into this layer is from both junctions. In practice, the intrinsic layer is always unintentionally doped as either p^- or n^-, where one of the junctions is a p–n junction and the other is a high–low junction; but injection takes place from both junctions.

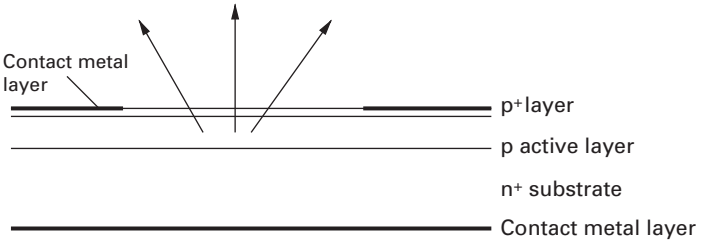

Fig. 4.2 Sketch of the n^+–p junction LED with a p^+ contact layer.

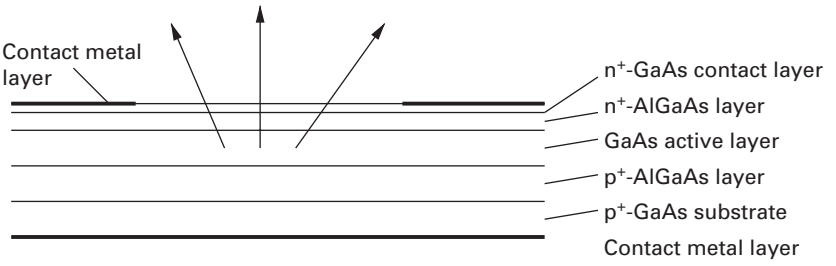

Fig. 4.3 Schematic diagram of the heterojunction LED.

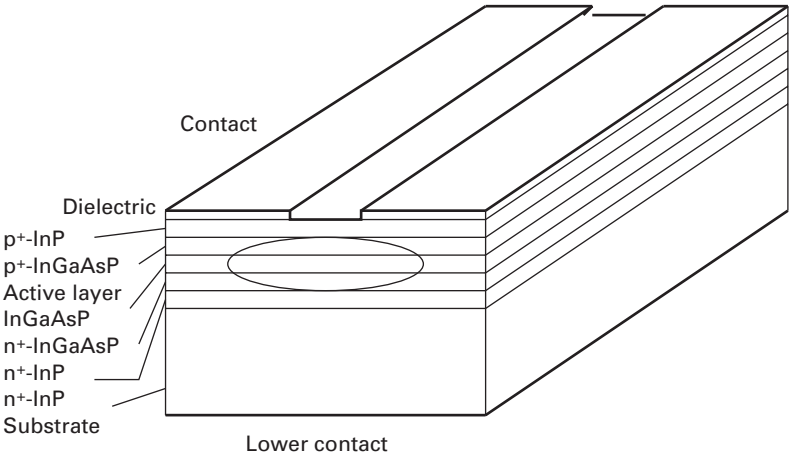

Fig. 4.4 Schematic diagram of the edge-emitting LED.

A variation of this structure is to have higher bandgap heterostructure layers both above and below the active layer, appropriately doped. The diffusion of the injected electrons and holes injected from the active layer is blocked by the heterojunctions that are now formed on both sides of the active layer, as shown in Figure 4.3. These heterolayers prevent the diffusion of the minority carriers to the surface, and therefore prevent surface recombination. Most LEDs designed at the current time use the heterolayers where available.

A further variation on this is the placement of dielectric mirrors below the active layer to reflect the light emitted towards the substrate. A second variant is the edge-emitting LED shown in Figure 4.4, in which the heterolayers above and below the active layers together with the active layer act as a waveguide. A high reflectivity mirror at one facet makes this a superluminescent diode; with mirrors on both ends and with adequate gain, the LED may operate as a laser. The external quantum efficiency of the edge-emitting LED is very much higher than the surface-emitting devices because the transmissivity at the edge facet is of the order of 0.7 for the waveguide index n_2 of 3.5. With a high reflectivity facet coating at one end, the external quantum efficiency is also about 0.7 for n_2 of 3.5. Anti-reflection coating on transmitting facet would increase the transmissivity to almost unity, and the external quantum efficiency also becomes nearly unity. Thus, the

total quantum efficiency or the wall plug efficiency of edge-emitting LEDs is determined largely by η_{int}, which is of the order of 0.5 in direct gap semiconductors.

The normal emission of the surface-emitting LED is Lambertian, which implies that the emission intensity at angle θ from the normal is $\cos\theta$, which means that the beam width is 120°. For the edge-emitting LED, the emission is elliptic in form, and the beam width is about 30° in the horizontal plane but remains 120° in the vertical plane.

LEDs are used in optical fibre communication systems that operate with multimode fibres, as the broad emission spectrum prevents modal noise being a problem [17]. A diode proposed by Burrus [4] has the fibre bonded into the face of the surface-emitting LED.

4.2.3 Semiconductor lasers

While LEDs utilise spontaneous emission for the emitted light, lasers operate on stimulated emission-generated light. Lasers are optical oscillators in which the gain medium is in a cavity; the light acquires gain in the medium between reflections from the ends of the cavity until steady state is reached, when the gain becomes saturated. The simplest version is a Fabry Perot cavity with the gain medium between two mirrors.

In this section, semiconductor lasers are discussed, as they may be designed to be high-speed and high frequency lasers. While solid state, fibre and other types of lasers may produce extremely short pulses, known as ultrafast lasers, their repetition rates are typically 80 MHz, to a maximum of a few GHz, and are not considered 'high-speed' or 'high frequency' lasers, and therefore not discussed here. The semiconductor laser threshold condition is first considered, then waveguides used in these lasers are outlined, and different types of lasers are discussed. This is followed by the derivation of the rate equation for these lasers, and the solutions for various conditions. Discussion of noise in semiconductor lasers is also included. Subsequently, quantum well, quantum dot lasers and vertical cavity lasers are briefly discussed.

Basic concepts

When the gain medium is in a cavity formed by two mirrors, shown schematically in Figure 4.5, light in the form of an electromagnetic wave, electric field amplitude $\mathcal{E}_0$, travels from one end of the cavity to the other end, where it is reflected by the end

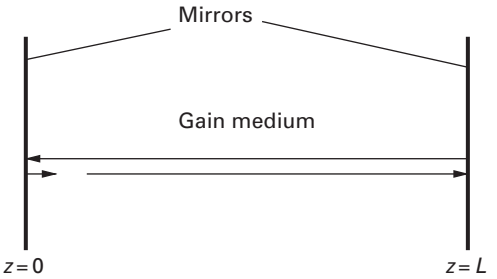

Fig. 4.5 Schematic diagram of gain medium in a cavity formed by two end mirrors.

mirror, and then propagates back, and is then reflected by the second mirror to return to its starting point. Suppose that the length of the cavity is L, this determines the distance travelled by the wave between reflections. Assume that the gain of the medium is given by $g/2$ Nepers per unit length, and hence the intensity gain is g per cm. Suppose the medium internal loss is $\alpha_{\text{int}}/2$ Nepers per unit length, or the intensity loss is α per cm, the medium index is n, and the reflection coefficients of the mirrors are r_1 and r_2. For the structure to commence oscillation, the loop gain should be unity or larger, which implies that the field at the starting point $\mathcal{E}_0$ has to undergo these reflections and is also subjected to gain to become after one round trip

$$\mathcal{E}_0 = \mathcal{E}_0(2L) = \mathcal{E}_0 e^{-j2k_0nL} e^{(g-\alpha_{\text{int}})L} |r_1||r_2|. \tag{4.26}$$

The real part of this equation gives

$$g = \alpha_{\text{int}} + \frac{1}{L} \ln \frac{1}{(|r_1||r_2|)}. \tag{4.27}$$

The mirror reflectivity is defined as $R_1 = |r_1|^2$, and similarly $R_2 = |r_2|^2$. Substituting in the above equation,

$$g = \alpha_{\text{int}} + \frac{1}{2L} \ln \frac{1}{(R_1R_2)} = \alpha_{\text{int}} + \alpha_{\text{mir}}. \tag{4.28}$$

The first term on the right-hand side is the medium loss without pumping and the second term is the mirror loss term, usually denoted by α_{mir}. For the gain medium of GaAs, the index is about 3.45, and the mirror is assumed to be formed by cleaved facets which result in plane mirrors, parallel to each other. The reflection coefficient of each facet mirror is

$$r_{1,2} = \frac{3.45 - 1}{3.45 + 1} = 0.551. \tag{4.29}$$

The facet mirror reflectivity is 0.303, and hence the mirror loss term α_{mir} is 1.194 cm^{-1}. The internal loss term is usually between 10 and 20 cm^{-1}, and therefore the gain needs to be 11.194–21.194 cm^{-1} for the round trip gain to be unity, and generally the gain needs to be higher than this value.

In practice, only part of the wave obtains gain from the active region, and this is defined by the confinement factor Γ, which is discussed later. Thus, the oscillation condition for the laser in Equation (4.28) becomes

$$\Gamma g = \alpha_{\text{int}} + \alpha_{\text{mir}}. \tag{4.30}$$

Also the complex propagation constant in the gain medium now is given by

$$\beta = nk_0 - j\left(\frac{\alpha}{2}\right), \tag{4.31}$$

where n is the guide effective index and the loss term α is the net loss, and given by

$$\alpha = \alpha_{\text{int}} + \alpha_{\text{mir}} - \Gamma g. \tag{4.32}$$

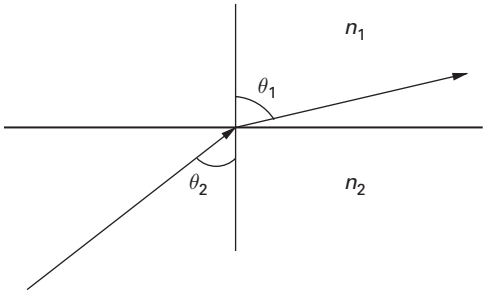

Fig. 4.6 Plane interface between medium 1, index n_1, and medium 2, index n_2, and the incident and transmitted light directions given by θ_1 and θ_2 respectively to the normal.

The imaginary part of Equation (4.26) determines the phase requirement:

$$2k_0nL = 2m\pi \quad \text{or} \quad \nu_{\text{m}} = \frac{mc}{(2nL)}, \tag{4.33}$$

where m is the longitudinal mode number, which may take values of $1, 2, 3, \ldots$, but cannot be zero, as the solution is then trivial. Note that the mode spacing in frequency is given by $c/(2nL)$. All longitudinal modes, with the values of m, satisfy this equation but only some of them are valid for a device. Since the medium gain is band-limited, the modes within this gain region are all excited at threshold when the gain is just larger than the losses. With increasing gain, generated by current pumping, the mode competition results in those close to the gain peak growing at the expense of the other modes.

Optical waveguides in semiconductors lasers

The light in the gain medium discussed above needs to be confined and guided, and this requires that the medium is in the form of an optical waveguide. An optical waveguide utilises the concept of total internal reflection that occurs when light emerges from a higher index medium to a lower index medium. In Figure 4.6, light in the form of an electromagnetic plane wave in a medium of index n_2 is incident on the plane interface between the media, at an angle θ_2 to the normal. The second medium index is given by n_1, with $n_2 > n_1$, and the light emerges into the second medium at an angle θ_2, obtained from Snell's law:

$$n_2 \sin\theta_2 = n_1 \sin\theta_1 \tag{4.34}$$

Since $n_2 > n_1$, it follows that $\theta_1 > \theta_2$. At the critical angle of θ_{2c}, the value of θ_1 becomes $\pi/2$, which implies that the emerging light travels along the interface. For incident angles greater than θ_{2c}, total internal reflection occurs, and with a second interface below the first, similar total internal reflection occurs so that the light remains confined to the high index region as shown in Figure 4.7. A similar confinement may occur in the plane normal to this to obtain two-dimensional guiding with appropriate layers to provide for an index guiding structure.

The solution of the wave equation for this three-layer guide used in a laser results in both even and odd modes. The even mode is sketched in Figure 4.8. Semiconductor lasers are classified as gain-guided lasers or index-guided lasers. In gain-guided

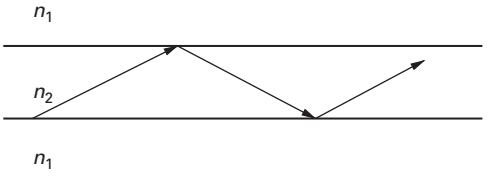

Fig. 4.7 Slab waveguide which provides confinement in the transverse plane.

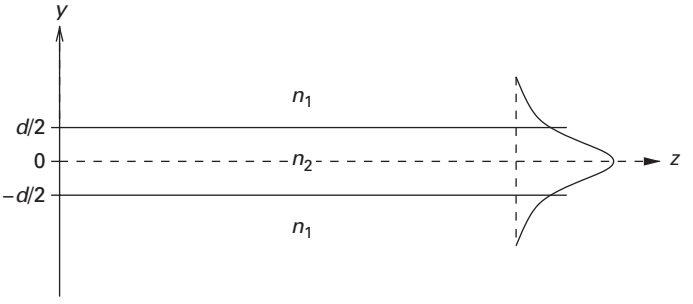

Fig. 4.8 Three-layer slab symmetric guide, which is typical for laser structures: the upper layer is p-type, the guide layer of higher index, is generally undoped, and the lower layer is n-type. The guide layer thickness is d. The even mode field distribution is sketched .

lasers, the mode confinement in the lateral direction, the horizontal direction, is not designed into the structure, and the mode is guided by the gain region of the structure. In index-guided lasers, the waveguide is well defined and the mode is confined in both the transverse direction, the vertical direction normal to the plane of the wafer and the lateral direction. A very popular structure among research scientists is the ridge laser, which is weakly index-guided. In all these lasers, the transverse confinement is obtained by the design of the heterolayers. Gain-guided, index-guided and weakly index-guided lasers are shown schematically in Figure 4.9.

Optical waveguides are analysed using Maxwell's equations, and the solution of the two-dimensional guide problem may be performed by a variety of methods. The results of this analysis enables the active region which acts as the guide to be designed. These guides are generally designed to be single mode in the $x - y$ plane. The width of the guide is usually designated by w, its thickness by d and the length of the laser is L, leading to an active volume of Lwd. The current is assumed to flow in the contact over the length of laser and the width of w, which results in a current flow area of Lw. The analysis of the guide is generally performed, assuming that there is no loss or gain, and these terms are added as perturbations. The analysis determines the propagation constant of the guide β, and also its effective relative permittivity ϵ_{reff} and effective index n_{eff}. The phase and group velocities are different in optical guides, and the corresponding effective indices are also obtained. An important parameter is the confinement factor Γ. This defines the fraction of the power contained in the active guide region to the total power in the particular mode of the guide.

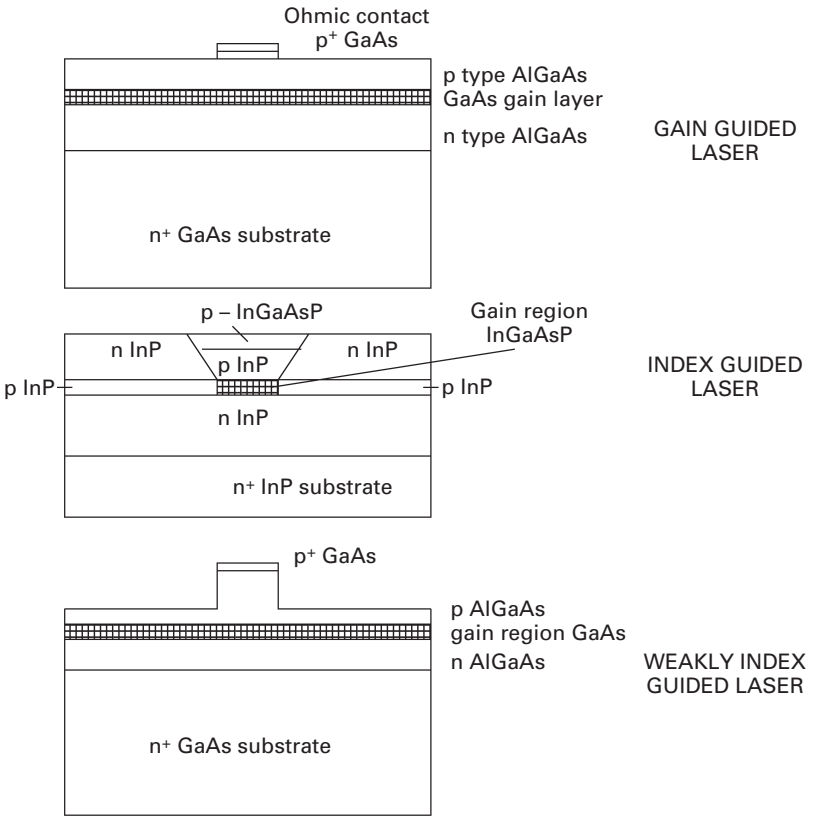

Fig. 4.9 Schematic diagram of the gain-guided laser, the index-guided laser and the ridge laser which is a weakly index-guided structure.

The effective index method is discussed next; it leads to approximate analytic expressions for the design of the guide layer. In this method, the guiding in the transverse, vertical, direction is analysed, and effective indices are calculated for each of the different transverse regions as they vary in the lateral direction. With the effective indices known for these transverse regions in the lateral, horizontal, direction, the one-dimensional guide in the lateral direction may be analysed, to obtain the solution of the entire guide.

Maxwell's equations lead to the Helmholtz's equation, assuming that the cross-section of the waveguide remains constant in the z direction and the propagation constant does not vary with z. Thus, the wave equation is of the form:

$$\nabla^2 \boldsymbol{\mathcal{E}} + \epsilon_{\mathrm{r}}(x, y) k_0^2 \boldsymbol{\mathcal{E}} = 0, \tag{4.35}$$

where $\epsilon_{\mathrm{r}}(x, y)$ is the structure permittivity which varies with both x and y directions but does not vary with the z direction. The dielectric constant also varies with pumping, typically the increase in carriers results in a small decrease in index, and vice versa. The dielectric constant is complex due to absorption or gain, and these effects are added as perturbations. Ignoring these effects enables the modes of the guiding structure to be

obtained, and using the effective index approach allows both index-guided structures and gain-guided laser structures to be analysed.

In the effective index approach, the assumption is that the solution may be separated into a y-varying, transverse varying, component in the form of a slab waveguide, and a similar effective index varying guide in the x direction. Assume that the solution of the Equation (4.35) obtained by the separation of variables is of the form:

$$\boldsymbol{\mathcal{E}} = \mathbf{a}\xi(y; x)\psi(x)e^{-\mathrm{j}\beta \mathrm{z}}, \tag{4.36}$$

where β is the propagation constant and $\mathbf{a}$ is unit vector in the direction of the $\mathcal{E}$-field, $\boldsymbol{\mathcal{E}}$, which defines the mode polarisation. Substituting in Helmholtz's Equation (4.35),

$$\frac{1}{\psi}\frac{d^2\psi}{dx^2} + \frac{1}{\xi}\frac{d^2\xi}{dy^2} + \left[k_0^2\epsilon_{\mathrm{r}}(x, y) - \beta^2\right] = 0, \tag{4.37}$$

where $k_0 = \omega\sqrt{(\epsilon_0\mu_0)}$, which defines the propagation constant in free space.

The next step is to solve the transverse, y directed, field distribution and with it the effective propagation constant $\beta_{\mathrm{eff}}(x)$ for a fixed value of x. Using the transverse part of the Equation (4.37)

$$\frac{d^2\xi}{dy^2} + \left[k_0^2\epsilon_{\mathrm{r}}(x, y) - \beta_{\mathrm{eff}}^2(x)\right]\xi = 0. \tag{4.38}$$

The lateral, x directed, field distribution and the propagation constant β are then obtained from the equation:

$$\frac{d^2\psi}{dx^2} + \left[\beta_{\mathrm{eff}}^2 - \beta^2\right]\psi = 0. \tag{4.39}$$

Consider the transverse modes of a typical laser structure, which may have as many as four or five layers in the slab guide form. However, the principle of the design is obtained from the three-layer guide shown in Figure 4.8. The solution of Equation (4.38) for the even TE mode ($\boldsymbol{\mathcal{E}}$ is directed along the x, lateral, direction) is of the form:

$$\xi = A_{\mathrm{e}}\cos(\kappa y) \qquad \text{for } |y| \le d/2 \tag{4.40}$$

$$= B_{\mathrm{e}}e^{-\gamma(|y|-d/2)} \qquad \text{for } |y| \ge d/2, \tag{4.41}$$

where

$$\kappa = k_0\left(n_2^2 - n_{\mathrm{eff}}^2\right)^{0.5} \tag{4.42}$$

$$\gamma = k_0\left(n_{\mathrm{eff}}^2 - n_1^2\right)^{0.5}. \tag{4.43}$$

Note that n_1 and n_2 are the refractive indices of the cladding and the guide layer respectively, and $n_2 > n_1$ for guiding.

In the TE mode case, the only components of field present are $\mathcal{E}_{\mathrm{x}}$, $\mathcal{H}_{\mathrm{y}}$ and $\mathcal{H}_{\mathrm{z}}$. The boundary conditions require the continuity of ξ and $d\xi/dy$ at $|y| = d/2$, and these correspond to the continuity of the $\mathcal{E}_{\mathrm{x}}$ component across the interfaces, and the continuity

of the $\mathcal{H}_z$ component across the interfaces, respectively. This leads to the following two equations:

$$A_e \cos\left(\frac{\kappa d}{2}\right) = B_e \tag{4.44}$$

and

$$\kappa A_e \sin\left(\frac{\kappa d}{2}\right) = \gamma B_e. \tag{4.45}$$

Dividing the above two equations,

$$\kappa \tan\left(\frac{\kappa d}{2}\right) = \gamma. \tag{4.46}$$

The solution of this equation gives the values of β_{eff} from which the effective index is obtained through the relationship $\beta_{\text{eff}} = k_0 n_{\text{eff}}$.

For the odd TE modes, it may be shown that the dispersion relationship becomes

$$-\kappa \cot\left(\frac{\kappa d}{2}\right) = \gamma. \tag{4.47}$$

This is obtained by setting the initial solution for ξ in Equation (4.40) as $\sin(\kappa y)$ instead of $\cos(\kappa y)$.

For the TM modes, the $\boldsymbol{\mathcal{E}}$ vector is along the y direction, normal to the interfaces in Figure 4.8. In this case, the only components of fields present are $\mathcal{E}_y$, $\mathcal{H}_x$ and $\mathcal{E}_z$. From the Maxwell curl equations, $\mathcal{E}_z$ is proportional to $d\mathcal{E}_y/dy$. Thus, the boundary conditions are the continuity of the y component of electric flux across the interfaces and the continuity of $\mathcal{E}_z$ across the interfaces. These lead to the following equations for the even TM modes:

$$n_2^2 A_e \cos\left(\frac{\kappa d}{2}\right) = n_1^2 B_e \tag{4.48}$$

and

$$\kappa A_e \sin\left(\frac{\kappa d}{2}\right) = \gamma B_e. \tag{4.49}$$

Dividing the above two equations gives the dispersion relationship:

$$\kappa n_1^2 \tan\left(\frac{\kappa d}{2}\right) = n_2^2 \gamma. \tag{4.50}$$

For the odd TM mode, the dispersion relationship is given by

$$-\kappa n_1^2 \cot\left(\frac{\kappa d}{2}\right) = n_1^2 \gamma. \tag{4.51}$$

Consider the TE mode solutions, squaring Equations (4.42) and (4.43), and adding results in

$$\kappa^2 + \gamma^2 = k_0^2 \left(n_2^2 - n_1^2\right). \tag{4.52}$$

This is the equation of a circle in the $\kappa - \gamma$ plane, and the intersection of the circle with the curves defined by the Equations (4.46) and (4.47) provides the modal solutions for

these equations. Note that multiple solutions are likely to occur as both tan and cot are periodic functions, and also depending on the parameters of the guide, defined by n_1, n_2, d and the wavelength which defines k_0.

At the cutoff of the guide which implies that the guide is no longer guiding, then $\gamma = 0$. Note that γ may not become negative, as it would imply exponential growth in the cladding region which is non-physical, and therefore the smallest value γ takes zero. When $\gamma = 0$, then from Equations (4.46) and (4.47)

$$\kappa d = p\pi, \tag{4.53}$$

where p is an integer whose even and odd values satisfy Equations (4.46) and (4.47) respectively, corresponding to the even and odd modes. Now let

$$D = p\pi = \kappa d. \tag{4.54}$$

Now for $\gamma = 0$, Equation (4.52) becomes

$$D = k_0 \left(n_2^2 - n_1^2\right)^{0.5} d, \tag{4.55}$$

where D is the normalized guide layer thickness. For a single transverse TE mode guide, this requires $D < \pi$. The layer thickness d for a single transverse mode guide is obtained as follows. Setting $k_0 = 2\pi/\lambda$ in the above Equations gives

$$d < \frac{\lambda}{2}\left(n_2^2 - n_1^2\right)^{-0.5}. \tag{4.56}$$

For GaAs/AlGaAs at 870 nm wavelength, with indices of 3.45 and 3.41, this suggests the guide thickness of less than 830 nm. The usual thickness is of the order of 0.2–0.5 μm. If the AlGaAs layer has a higher value of index, closer to that of GaAs, then the thickness of this layer may be larger. For InGaAsP lasers [2], the layer thickness is also of the order of 0.2 μm, and in this case the layer thickness needs to be less than 0.48 μm. According to [2], this relationship holds for lasers in the wavelength range 1.1–1.65 μm. The confinement factor in the transverse direction Γ_T is a measure of how much of the mode power lies in the guide active region, and is calculated from the equation:

$$\Gamma_T = \frac{\int_{-d/2}^{d/2} \xi^2(y)dy}{\int_{-\infty}^{\infty} \xi^2(y)dy}. \tag{4.57}$$

Performing the integration and simplifying, [2] derives this as

$$\Gamma_T \cong \frac{D^2}{(2 + D^2)}. \tag{4.58}$$

The effective index of the guide in the transverse direction is given as [2]:

$$n_{\text{eff}}^2 \cong n_1^2 + \Gamma_T\left(n_2^2 - n_1^2\right). \tag{4.59}$$

The lateral modes are next evaluated. The loss of laser structures may be as high as 5–10 Np cm^{-1}, and therefore in principle these losses need to be taken into account. However, when pumped, the gain which is typically of the order of 50 Np cm^{-1} and larger, allows the guide to become transparent, which implies that the loss is cancelled

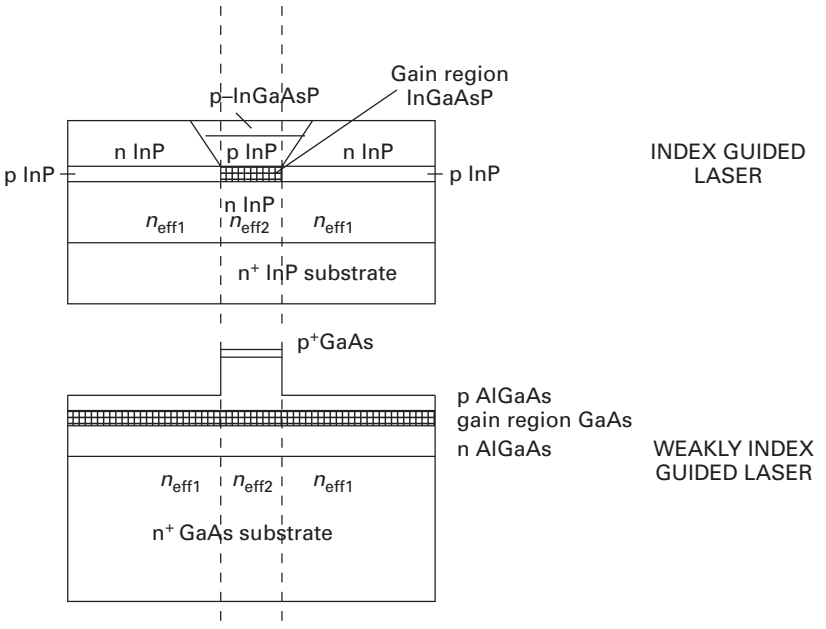

Fig. 4.10 Three-layer slab symmetric guide in which the slabs are specified by the effective indices n_{eff1} and n_{eff2} vertically, with $n_{\text{eff2}} > n_{\text{eff1}}$ for the index-guided laser and the weakly guiding ridge laser.

by the gain. However, if the losses need to be accounted for, this is usually performed by the perturbation technique. Thus, the lateral modes are calculated in a similar manner as the transverse modes, assuming the guide is lossless. In this case, the wave equation to be solved is given by Equation (4.39) and repeated here for convenience:

$$\frac{d^2\psi}{dx^2} + \left[\beta_{\text{eff}}^2(x) - \beta^2\right]\psi = 0. \tag{4.60}$$

Note that $\beta_{\text{eff}}(x) = k_0 n_{\text{eff}}(x)$ which is obtained from Equation (4.59), and thus the propagation constant of the total guide β is obtained. In the case of the index-guided laser in Figure 4.9, the transverse effective index in the main guide region has the largest effective index, and the regions outside have lower effective indices. As shown in Figure 4.10, the symmetric slab guide in the vertical direction with the effective indices calculated from the transverse slab guide equation. Here, $n_{\text{eff2}} > n_{\text{eff1}}$ so that the guide acts with the centre region as the larger index.

In this case, the TE mode from the transverse slab guide becomes a TM mode and the TM mode becomes a TE mode. Based on the earlier approach, the width of the central region is set as w, and the normalised width as W. Then, it follows that

$$W = k_0 \left(n_{\text{eff2}}^2 - n_{\text{eff1}}^2 \right)^{0.5} w \tag{4.61}$$

and

$$W = q\pi. \tag{4.62}$$

For the lowest order mode

$$w < \frac{\lambda}{2} \left(n_{\text{eff2}}^2 - n_{\text{eff1}}^2 \right)^{-0.5}. \tag{4.63}$$

In this case, the effective indices take different values, but this provides an indication to the width of the guide which usually is in the 3 μm region.

The lateral confinement factor Γ_{L}, following the earlier derivation, is given by

$$\Gamma_{\text{L}} = \frac{W^2}{(2 + W^2)}. \tag{4.64}$$

The waveguide mode refractive index is now given by

$$n_{\text{eff}}^2 \cong n_{\text{eff1}}^2 + \Gamma_{\text{L}} (n_{\text{eff2}}^2 - n_{\text{eff1}}^2). \tag{4.65}$$

The laser confinement factor Γ is the product of Γ_{T} and Γ_{L} and is given by

$$\Gamma = \Gamma_{\text{T}} \Gamma_{\text{L}}. \tag{4.66}$$

The two most important parameters that are obtained from this analysis are the effective index of the guide and the confinement factor. From the effective index, the guide propagation constant may be obtained.

These results are for the index-guided structures, and the same technique may be used for the weakly guiding ridge structure in Figure 4.9. The analysis of the gain-guided laser is much more complex, and interested readers are referred to the paper by Nash [24].

Emission characteristics

The waveguide which guides the light has facet mirrors on both sides of the device. These may be coated to obtain high reflectivity or to reduce the reflectivity as desired, but generally these facets are uncoated, in which case the reflectivities are equal.

With current pumping, the light intensity against input current is plotted in the L–I characteristic as shown schematically in Figure 4.11. Note the different regions of this curve: the pre-threshold where the laser output is from spontaneous emission, the threshold current at which the device starts to lase, which is the linear region, followed by the saturation region due to gain saturation. The current into the laser is of the form:

$$I = wLJ, \tag{4.67}$$

where w is the width of the active region and L is the cavity length. In practice, index-guided lasers in Figure 4.9 have leakage current through the various reverse-biased junctions, and should be added to this current expression, but this is ignored in the present discussion.

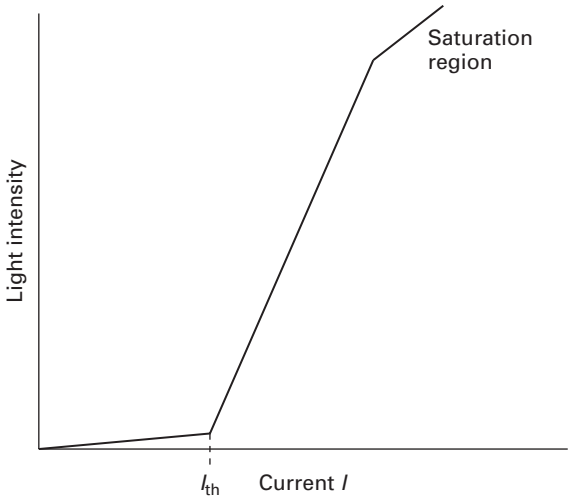

Fig. 4.11 Schematic diagram of the light intensity against current drive, the L–I characteristic.

Assume that the gain is given approximately by the expression

$$g = a(N - N_0), \tag{4.68}$$

where a is the gain coefficient, equal to $(\partial g/\partial N)$, N is the carrier density, and N_0 is the carrier density at which transparency is obtained, when population inversion occurs. The threshold carrier density is defined when the product of the confinement factor and the gain is equal to the loss, or

$$\Gamma a(N_{\text{th}} - N_0) = \alpha_{\text{mir}} + \alpha_{\text{int}}, \tag{4.69}$$

or

$$N_{\text{th}} = N_0 + (\alpha_{\text{mir}} + \alpha_{\text{int}})/(a\Gamma). \tag{4.70}$$

The number of carriers pumped into the active region per second is I/q. However, since the equations are in carrier density, the number of carriers injected per second per unit volume is $I/(q\text{Vol}) = I/(qwLd) = J/qd$. The loss of the carrier density is through recombination at the rate $R(N)$, which also includes the stimulated emission recombination, and this is defined as R_{total}. Thus, the rate equation for the carriers is given by

$$\frac{dN}{dt} = \frac{J}{qd} - R_{\text{total}} = \frac{J}{qd} - \frac{N}{\tau_{\text{e}}} - R_{\text{stim}}N_{\text{ph}}, \tag{4.71}$$

where τ_{e} is the carrier lifetime and the last term is the stimulated emission recombination. At and below threshold, this last term may be omitted as the contribution from stimulated emission is small. At steady state, the time variation is zero, and hence

$$J = qdR_{\text{total}}. \tag{4.72}$$

The stimulated emission recombination rate is $R_{\text{stim}}N_{\text{ph}}$, where N_{ph} is the photon density and R_{stim} is defined by

$$R_{\text{stim}} = \frac{c}{n}g(N), \tag{4.73}$$

where $g(N)$ is the gain defined in Equation (4.68), c/n is the group velocity.

From Equations (4.70) and (4.71), the threshold current density is obtained as

$$J_{\text{th}} = \frac{qdN_{\text{th}}}{\tau_{\text{e}}(N_{\text{th}})}, \tag{4.74}$$

where, neglecting stimulated emission recombination close to threshold,

$$\frac{1}{\tau_{\text{e}}(N)} = (A_{\text{nr}} + B_{\text{r}}N + CN^2), \tag{4.75}$$

where A_{nr} is non-radiative recombination coefficient, B_{r} is the radiative recombination coefficient and C is the Auger non-radiative recombination coefficient, which is important for long wavelength lasers.

When the laser operates beyond threshold, the carrier density is clamped at the threshold value, and further injection results in conversion into photons by stimulated emission. Thus, Equation (4.72) may be written as

$$J = qdR_{\text{total}} = qd\frac{N_{\text{th}}}{\tau_{\text{e}}} + qdR_{\text{stim}}N_{\text{ph}}. \tag{4.76}$$

Substituting from Equation (4.73), and replacing g by $\alpha_{\text{int}} + \alpha_{\text{mir}}$ at threshold,

$$J - J_{\text{th}} = qdv_{\text{g}}(\alpha_{\text{int}} + \alpha_{\text{mir}})N_{\text{ph}}. \tag{4.77}$$

The photon density in the cavity depends on the carriers injected into the cavity and the quantum efficiency of the material. However, the photons in the cavity travel at the group velocity of the medium v_{g}, and are either absorbed due to internal losses or escape from the facets. Thus, the photons have a finite lifetime in the cavity, and the photon lifetime τ_{p} is given by

$$\tau_{\text{p}} = \frac{1}{v_{\text{g}}(\alpha_{\text{mir}} + \alpha_{\text{int}})}. \tag{4.78}$$

The Equation (4.77) becomes

$$J - J_{\text{th}} = qd\frac{N_{\text{ph}}}{\tau_{\text{p}}} \tag{4.79}$$

The relationship between the injected carrier density per second and the photon density in the cavity is now obtained. The carriers recombine at a rate defined by the carrier lifetime, to produce photons, and this is accounted for by the quantum efficiency of the material from Equation (4.3), and the photon density also decays at the photon lifetime. Thus,

$$\eta_{\text{int}}\frac{(J - J_{\text{th}})}{qd} = \frac{N_{\text{ph}}}{\tau_{\text{p}}} \tag{4.80}$$

or

$$N_{\text{ph}} = \eta_{\text{int}} \tau_{\text{p}} \frac{(J - J_{\text{th}})}{qd}. \tag{4.81}$$

This equation shows that the photon density in the cavity increases linearly with current density above the threshold current density.

The output power per facet is obtained by the product of energy of each photon $h\nu$, the photon loss per facet $v_{\text{g}}\alpha_{\text{mir}}/2$, the active laser volume *Vol* and the photon density N_{ph}:

$$P_{\text{facet}} = \frac{1}{2} h\nu v_{\text{g}} \alpha_{\text{mir}} \text{Vol} N_{\text{ph}}, \tag{4.82}$$

where $\text{Vol} = Lwd$. Substituting for N_{ph}

$$P_{\text{facet}} = \frac{h\nu}{2q} \eta_{\text{int}} \frac{\alpha_{\text{mir}}}{\alpha_{\text{mir}} + \alpha_{\text{int}}} Lw(J - J_{\text{th}}). \tag{4.83}$$

Since $I = LwJ$, this equation becomes

$$P_{\text{facet}} = \frac{h\nu}{2q} \eta_{\text{int}} \frac{\alpha_{\text{mir}}}{\alpha_{\text{mir}} + \alpha_{\text{int}}} (I - I_{\text{th}}). \tag{4.84}$$

Note that the threshold current is obtained from Equation (4.74)

$$I_{\text{th}} = \frac{qLwdN_{\text{th}}}{\tau_{\text{e}}}. \tag{4.85}$$

The total output power $P_{\text{out}} = 2P_{\text{facet}}$ and the differential (external) quantum efficiency is given by

$$\eta_{\text{d}} = \frac{d\left(\dfrac{P_{\text{out}}}{h\nu}\right)}{d\dfrac{(I - I_{\text{th}})}{q}} = \eta_{\text{int}} \frac{\alpha_{\text{mir}}}{\alpha_{\text{mir}} + \alpha_{\text{int}}}. \tag{4.86}$$

This is proportional to the slope of the L–I curve. Note that in the laser the η_{int} needs to take into account the stimulated emission rate. Then, the total recombination rate becomes

$$R_{\text{total}} = A_{\text{nr}} N + B_{\text{r}} N^2 + C N^3 + R_{\text{stim}} N_{\text{ph}}. \tag{4.87}$$

The coefficients have been identified in Equation (4.75) except for R_{stim}, which is the stimulated emission rate and given in equation (4.73) Thus, the internal quantum efficiency now is given by

$$\eta_{\text{int}} = \frac{R_{\text{rad}}}{R_{\text{total}}} = \frac{B_{\text{r}} N^2 + R_{\text{stim}} N_{\text{ph}}}{A_{\text{nr}} N + B_{\text{r}} N^2 + C N^3 + R_{\text{stim}} N_{\text{ph}}} \tag{4.88}$$

Above threshold, the $R_{\text{stim}} N_{\text{ph}}$ dominates and is much larger than the other terms, which makes η_{int} almost unity.

The power efficiency of the laser is given by

$$\eta_P = \frac{P_{\text{out}}}{VI} = \frac{h\nu}{qV} \frac{\alpha_{\text{mir}}}{(\alpha_{\text{mir}} + \alpha_{\text{int}})} \frac{(I - I_{\text{th}})}{I}, \tag{4.89}$$

where V is the applied bias in volts.

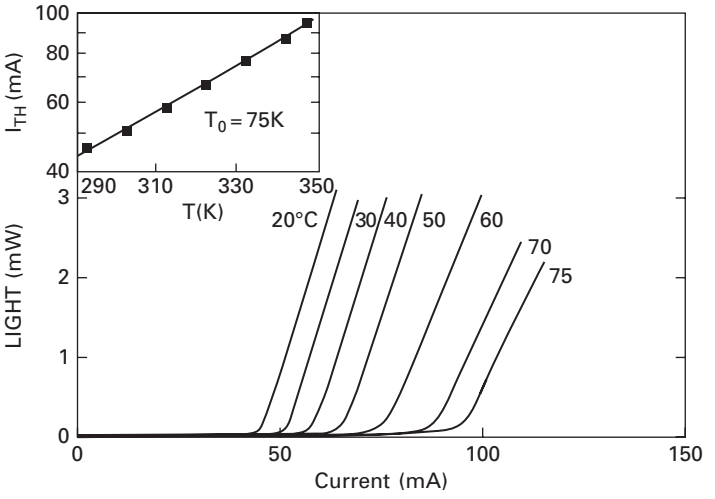

Fig. 4.12 The L–I curves for a buried heterostructure laser at different temperatures. The inset plots the threshold current against temperature to obtain T_0 (R. J. Nelson, R. B. Wilson, P. D. Wright, P. A. Barnes, N. K. Dutta, *IEEE Journal of Quantum Electronics*, Vol. 17, No. 2, pp. 202–207, 1981. ©1981 IEEE).

The temperature dependence of the threshold current density varies as

$$J_{\text{th}} = J_{\text{th}0} e^{T/T_0}. \tag{4.90}$$

A typical example of the variation of the L–I curves for different temperatures is shown in Figure 4.12 [25], and the inset in this figure is also the plot of threshold current with temperature to obtain T_0.

This figure shows that as the temperature rises, the threshold current increases as the recombination becomes increasingly non-radiative.

The edge-emitting laser has a large number of longitudinal modes within the gain region, and at threshold all these are excited. Figure 4.13 shows these modes, and at threshold only those modes that have enough gain to neutralise the loss finally emerge. In this figure, two modes are shown to have this property, and in general it is possible that only one of these modes is likely to be dominant and the other becomes a secondary mode as the current and hence the gain is increased.

In Figure 4.14, the spectrum of the emission is plotted relative to the current excitation along the L–I curve, and this is also from [25].

The spectrum narrows from several modes close to threshold to a dominant mode with a few subsidiary modes at higher drive currents. The ratio of the intensity of the dominant mode to the next highest mode expressed in dBs is a measure of the mode suppression ratio (MSR). Since there is no guarantee that the Fabry–Perot laser with facet mirrors will produce a device with a single dominant mode with a large MSR of at least 20 dB, other techniques for mode selection, such as gratings in the active region or outside, may be used.

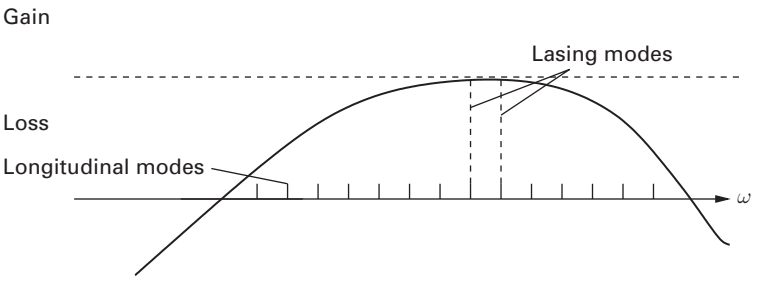

Fig. 4.13 The longitudinal modes of the laser within the gain region of the laser, showing the modes that are likely to lase when the gain is equal to the loss, including the mirror loss.

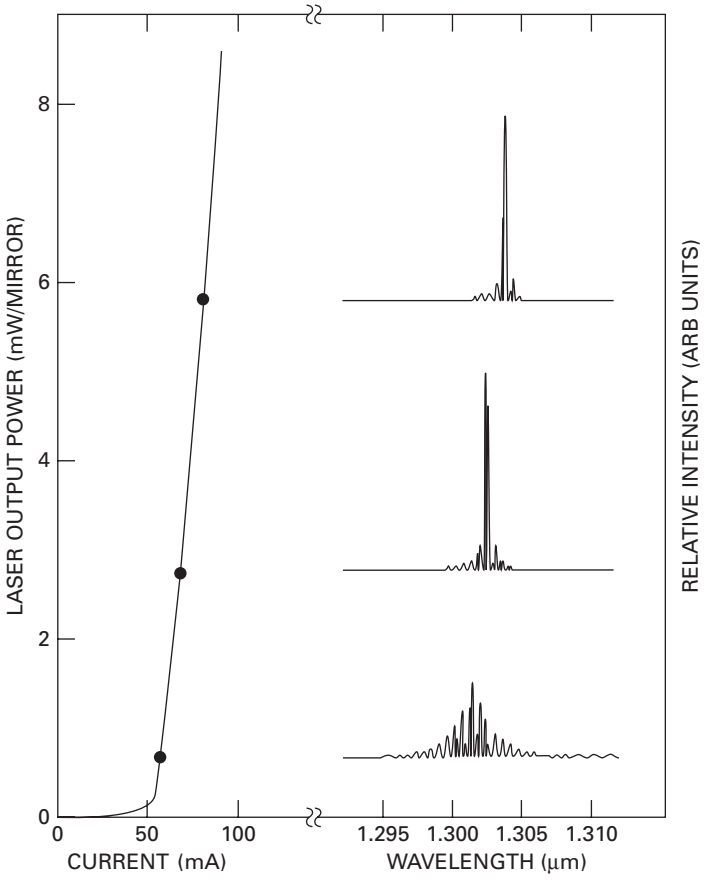

Fig. 4.14 The L–I curve of a buried heterostructure laser, together with the mode spectrum at different current levels of excitation (R. J. Nelson, R. B. Wilson, P. D. Wright, P. A. Barnes, N. K. Dutta, *IEEE Journal of Quantum Electronics*, Vol. 17, No. 2, pp. 202–207, 1981. ©1981 IEEE).

The laser illuminates the facets and these fields determine the near field pattern of the laser. It is usual to approximate the fields as a Gaussian distribution in the transverse direction, and a similar Gaussian distribution in the lateral direction (see [1]). The product of these Gaussians give rise to an elliptic distribution of the E-field on the facet. The far-field beam pattern is obtained using the usual methods by the spatial Fourier transform of the near field pattern [1].

Calculation of absorption, emission and gain

The calculation of the absorption, emission and gain in a semiconductor is a complex process, and will not be given here, as the method has been discussed in several textbooks, for example in [8]. Figure 4.15 shows the calculated gain/absorption spectra of InGaAsP for different levels of carrier injection, with the gain peak shifting as the carrier density increases [9].

Rate equations

Since lasers have a complex relationship between the injected carriers and the photons generated, the calculation of the dynamic response is more involved than that in LEDs. Essentially, the laser operates as an oscillator with a cavity, which is many wavelengths long, except in specific cases. Therefore, the cavity has many resonances corresponding to the expressions in Equation (4.33) in which the resonance frequencies are given by $\nu_{\mathrm{m}} = mc/2nL$, for different values of m. The corresponding radial frequency for the mth mode is assumed to be $\Omega_{\mathrm{m}} = 2\pi\nu_{\mathrm{m}}$, and the corresponding wave number k is given by

$$k_{\mathrm{m}} = \frac{n\Omega_{\mathrm{m}}}{c} = \frac{m\pi}{L}. \tag{4.91}$$

For the present, the subscript m is omitted for convenience. The laser radial frequency ω is undetermined, but nearly coincides with the cavity radial frequency Ω. The effective permittivity ϵ_{reff} is defined by the effective index of the laser guide, but the perturbation of gain and loss has to be added. However, the use of the effective permittivity or index reduces the wave equation to the one-dimensional form.

The injection of carriers into the laser causes a small change in index of the various layers, and for the active layer this takes the form:

$$\Delta n = bN, \tag{4.92}$$

where b is equal to $\partial n/\partial N$, and has a small negative value. Define

$$\beta_{\mathrm{c}} = -\frac{2k_0 b}{a} = -2k_0 \left(\frac{\frac{\partial n}{\partial N}}{\frac{\partial g}{\partial N}} \right), \tag{4.93}$$

which results in $\beta_{\mathrm{c}} \propto b$, but is dimensionless, and also positive. β_{c} has been described as the anti-guiding parameter, or the linewidth enhancement factor.

The effective index is now given by

$$n_{\mathrm{eff}} = n + \Gamma\Delta n. \tag{4.94}$$

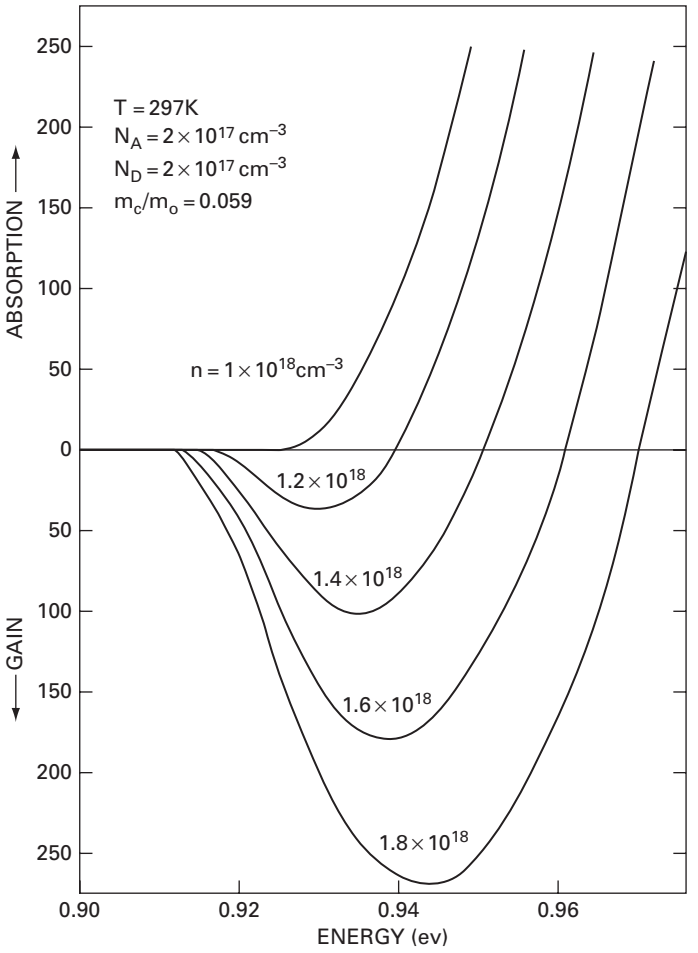

Fig. 4.15 The absorption/gain spectra for different levels of carrier injection. Reprinted with permission from N. K. Dutta, *Journal of Applied Physics*, Vol. 51, pp. 6095–6100, 1980. ©1980, American Institute of Physics.

This assumes that Δn is mainly in the active guide due to carriers, which is strictly not correct, as the carriers have to be injected from the top and bottom contacts. However, both the electrons and the holes end up in the active region to recombine and so this is an acceptable approximation. This assumption also simplifies equations to be derived below.

Also, the effective permittivity is given by

$$\epsilon_{\text{reff}} = (n + \Gamma \Delta n)^2 - j\frac{n\alpha}{k_0} \approx n^2 + 2\Gamma n \Delta n - \frac{jn\alpha}{k_0}. \tag{4.95}$$

The one-dimensional wave equation applied to a cavity requires a solution in the form $\sin k_m z$, with k_m defined in Equation (4.91), neglecting the effects of the end mirrors.

The electric field wave equation for the cavity is

$$\frac{d^2\mathcal{E}}{dz^2} - \frac{\epsilon_{\text{reff}}}{c^2}\frac{d^2\mathcal{E}}{dt^2} = 0, \tag{4.96}$$

and the solution is of the form:

$$\mathcal{E}(z,t) = \sum_i \sin(k_j z)\mathcal{A}(t)e^{j\omega t}, \tag{4.97}$$

where $\mathcal{A}(t)$ is assumed to be slowly varying in time compared to the light wave frequency. Assuming this is a single longitudinal mode laser, the summation and the subscripts i and j are dropped. Substituting this in the wave Equation (4.96) and neglecting the second time derivative of A

$$-k^2\mathcal{A} + \frac{\omega^2}{c^2}\epsilon_{\text{reff}}\mathcal{A} - \frac{2j\omega}{c^2}\epsilon_{\text{reff}}\frac{d\mathcal{A}}{dt} = 0. \tag{4.98}$$

Simplifying with $k^2 \simeq \Omega^2 n^2/c^2$:

$$\left(\frac{\omega^2}{c^2}\epsilon_{\text{reff}} - \frac{\Omega^2 n^2}{c^2}\right)\mathcal{A} - \frac{2j\omega}{c^2}n^2\frac{d\mathcal{A}}{dt} = 0. \tag{4.99}$$

However, since the laser frequency ω is very close to the cavity frequency Ω, then $(\omega^2 - \Omega^2) \approx 2\omega(\omega - \Omega)$. With further simplifications, substituting for ϵ_{reff} from Equation (4.95), this equation becomes

$$\frac{d\mathcal{A}}{dt} = -j\frac{n}{n_g}(\omega - \Omega)\mathcal{A} - \frac{j\omega}{n_g}\left(\Gamma\Delta n - j\frac{\alpha}{2k_0}\right)\mathcal{A}, \tag{4.100}$$

where n_g is the group index.

Separate this equation into its real and imaginary parts by substituting

$$\mathcal{A} = Ae^{j\phi} \tag{4.101}$$

into (4.100) to obtain for the real and imaginary parts:

$$\frac{dA}{dt} = -\frac{\alpha}{2k_0}A = \frac{1}{2}v_g(\Gamma g - \alpha_{\text{mir}} - \alpha_{\text{int}})A \tag{4.102}$$

and

$$\frac{d\phi}{dt} = -(\omega - \Omega) - \frac{\omega}{n_g}\Gamma\Delta n. \tag{4.103}$$

The rate equations may be written in terms of the photon density N_{ph} and carrier density N, or in terms of total number of photons and total number of carriers in the laser active volume, and both these approaches have been used in the literature. In the following derivations, the total number of photons and carriers in the laser volume are used. The total number of photons in the cavity is obtained from the equation:

$$S = \frac{\epsilon_0 n^2}{h\nu}\int \text{Vol}\,\mathcal{E}^2 dV, \tag{4.104}$$

and since $A^2 \propto S$, then multiplying Equation (4.102) by A, this equation becomes

$$\frac{dS}{dt} = (G - \gamma_p)S + R_{\text{sp}}, \tag{4.105}$$

where

$$G = \Gamma v_g g \tag{4.106}$$

is the normalised gain, which is the stimulated emission rate. The photon decay rate is given by

$$\gamma_p = v_g(\alpha_{int} + \alpha_{mir}) = \frac{1}{\tau_p}. \tag{4.107}$$

The term R_{sp} has been added to account for the spontaneous emission in the lasing process. The spontaneous emission takes place over the whole laser cavity but only a small fraction β_{sp} couples into the waveguide mode, which is the integral part of the laser. R_{sp} may be written as

$$R_{sp} = \beta_{sp}\eta_{sp}\gamma_e N_t, \tag{4.108}$$

where

$$\gamma_e = (A_{nr} + B_r N + CN^2) = \frac{1}{\tau_e} \tag{4.109}$$

and β_{sp} is the spontaneous emission factor, which is usually a fitting parameter with values of 10^{-4}–10^{-5}, according to Agrawal [2]. The term $\eta_{sp} = B_r N/\gamma_e$ is the internal quatum efficiency, and is the fraction of carriers that recombine to emit photons through spontaneous emission.

N_t is the total number of carriers in the active volume obtained by integrating the carrier density N over the volume:

$$N_t = \int N dV. \tag{4.110}$$

The phase Equation (4.103) needs further simplification. From Equations (4.92) and (4.93)

$$\Delta n = -\frac{\beta_c}{2k_0} aN \approx -\frac{\beta_c}{2k_0}\Delta g. \tag{4.111}$$

The last term in Equation (4.103) becomes

$$\frac{\omega}{n_g}\Gamma\Delta n = -\frac{1}{2}\beta_c \Gamma v_g \Delta g. \tag{4.112}$$

Now $\Delta G = \Gamma v_g \Delta g$, and replacing ΔG by $G - \gamma_p$, Equation (4.103) becomes

$$\frac{d\phi}{dt} = -(\omega - \omega_{th}) + \frac{1}{2}\beta_c(G - \gamma_p). \tag{4.113}$$

The first term on the right-hand side in the above equation has been expanded in terms of the threshold frequency, and at threshold the cavity frequency Ω is very close to the threshold frequency.

The carrier rate equation given above in Equation (4.71) is

$$\frac{dN}{dt} = \frac{J}{qd} - \frac{N}{\tau_e} - R_{stim}N_{ph}. \tag{4.114}$$

The expression for R_{stim} is given by

$$R_{\text{stim}} = \frac{c}{n_{\text{g}}} g(N) = v_{\text{g}} g(N). \tag{4.115}$$

Equation (4.75) defines the carrier lifetime τ_{e}. Since the volumetric values of N and N_{ph} are used in the photonic rate equations, integrating this equation over the active volume results in

$$\frac{dN_{\text{t}}}{dt} = \frac{I}{q} - \gamma_{\text{e}} N_{\text{t}} - GS, \tag{4.116}$$

where $\gamma_{\text{e}} = 1/\tau_{\text{e}}$. The confinement factor is introduced in the last term to convert the expression to G, and the total photon number S is obtained from the integration of the photon density N_{ph}.

In practice, Fabry–Perot lasers with facet mirrors are longitudinally multimode, and therefore the rate equations apply to each individual mode, and have to be solved simultaneously. Thus, the rate equations become

$$\dot{S}_{\text{m}} = (G_{\text{m}} - \gamma_{\text{p}}) S_{\text{m}} + R_{\text{sp}}(\omega_{\text{m}}) \tag{4.117}$$

$$\dot{N}_{\text{t}} = \frac{I}{q} - \gamma_{\text{e}} N_{\text{t}} - \sum_{\text{m}} G S_{\text{m}}. \tag{4.118}$$

The rate equations developed above allow the calculation of the $L-I$ curve, the longitudinal mode spectrum and the MSR, among other properties of the laser, provided the parameters are known. The dynamic behaviour which includes the turn-on delay, the modal behaviour, small- and large-signal modulation may also be calculated from these equations. The calculation of noise requires the addition of the noise sources to these equations, and these may also be obtained. In the following section, some aspects of the calculation methods for the steady-state and dynamic behaviour are discussed.

Steady-state and dynamic characteristics

Under steady-state conditions, for the single longitudinal mode laser in Equation (4.105), the time variation is zero. This makes this equation:

$$(G - \gamma_{\text{p}}) S + R_{\text{sp}}(\omega) = 0, \tag{4.119}$$

which becomes

$$S = \frac{R_{\text{sp}}}{\gamma_{\text{p}} - G}, \tag{4.120}$$

which states that the spontaneous emission photons in the cavity are created by the injected current. When the net stimulation emission rate G is nearly equal to the photon decay rate γ_{p}, then threshold is reached. The value of G is a little below the decay rate at threshold, and as the gain increases, G is asymptotic to γ_{p}, but the denominator should always remain positive.

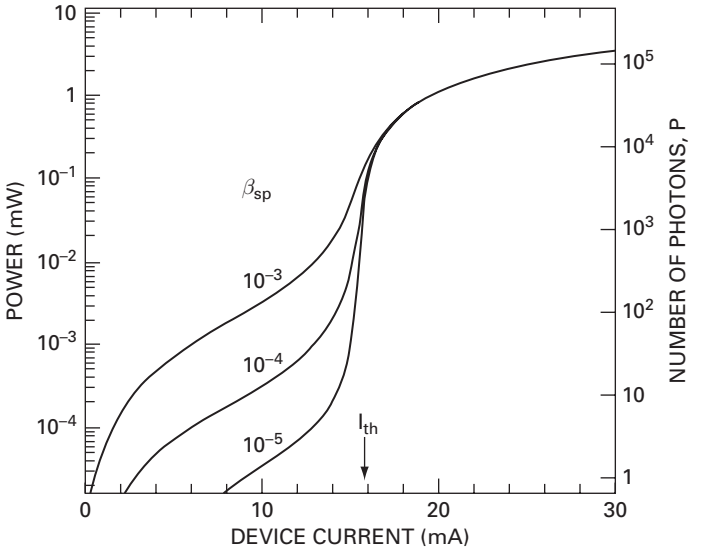

Fig. 4.16 Power output against current (G. P. Agrawal and N. K. Dutta, *Semiconductor Lasers*, Van Norstrand Rheinhold, 1993. ©Springer). With kind permission of Springer Science and Business Media.

Substituting in the carrier rate Equation (4.116)

$$\frac{I}{q} = \gamma_e N_t + R_{sp}\frac{G}{\gamma_p - G}. \tag{4.121}$$

This equation may be used to calculate the light intensity output against current, the L–I curve. Lee [19] has discussed the method of solution, as this is a non-linear equation. Figure 4.16 shows the results of the calculation of output power against current from [2] for the particular laser that has been modelled.

The time evolution of the build up in the carrier levels and the photon levels are shown in Figure 4.17 from Marcuse's paper [21] for a multimode laser when the current has been increased as a step function from 0 to 1.5 I_{th}. The rate equations are solved by numerical calculations outlined in [21], and the results are shown in the Figure 4.17.

The time delay of the build up to stimulated emission is typically of the order of 3 ns, and this is the reason why lasers are biased just below threshold current when they are pulsed on and off. The output power at this bias level is extremely low and so for most purposes may be regarded as negligible. Note the oscillations in the light power output at the relaxation oscillation frequency of the laser. This relaxation oscillation frequency f_r determines the maximum small-signal response of the laser, and this may be derived from the above rate equations. Also it may be shown that the relative intensity noise of the laser also peaks at the relaxation oscillation frequency, and thus modulation at or near this frequency needs to be avoided to improve the signal-to-noise ratio of the detected signal.

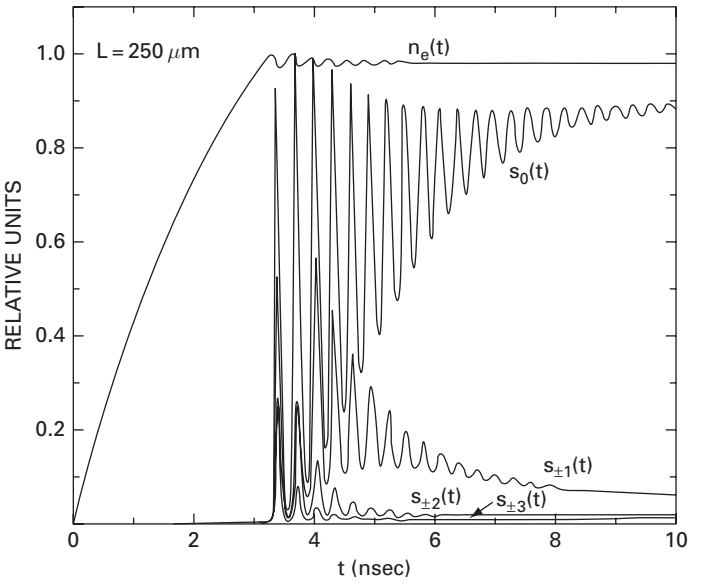

Fig. 4.17 Transient response for a laser in which the current is increased as a step function from 0 to $1.5\,I_{th}$, showing the carrier density and power in the fundamental mode and other modes (D. Marcuse, T. P. Lee, *IEEE Journal of Quantum Electronics* Vol. QE-19, No. 9, pp. 1397–1406, 1983. ©1983 IEEE).

This figure shows that the oscillations are damped and in the form of $e^{-(\Gamma_r \pm j\Omega_r)t}$, where Γ_r is the decay rate and Ω_r is the frequency of oscillation, which is the radial relaxation oscillation frequency in radians s^{-1}.

To calculate the relaxation oscillation frequency, the rate Equations (4.105) and (4.116), repeated here for convenience, are used.

$$\dot{S} = (G - \gamma_p)S + R_{sp} \tag{4.122}$$

$$\dot{N}_t = \frac{I}{q} - \gamma_e N_t - GS. \tag{4.123}$$

Consider perturbations to S and N_t given by δS and δN_t. Additionally,

$$G(N_t, S) \cong G + G_{N_t}\delta N_t + G_S \delta S, \tag{4.124}$$

where $G_{N_t} = \partial G/\partial N_t$ and $G_S = \partial G/\partial S$. Substituting in the rate equations, separating the perturbed terms, results in

$$\delta\dot{S} = -\Gamma_S \delta S + \left(G_{N_t} S + \frac{\partial R_{sp}}{\partial N_t}\right)\delta N_t \tag{4.125}$$

$$\delta\dot{N}_t = \Gamma_{N_t}\delta N_t - (G + G_S S)\delta S, \tag{4.126}$$

where

$$\Gamma_S = \frac{R_{sp}}{S} - G_S S, \tag{4.127}$$

and

$$\Gamma_{N_t} = \gamma_e + N_t \left(\frac{\partial \gamma_e}{\partial N_t} \right) + G_{N_t} S. \tag{4.128}$$

To solve these equations, the assumption is that these perturbations decay at the rate shown in the time evolution result in Figure 4.16.

Thus, let

$$\delta S = \delta S_0 e^{-ht} \tag{4.129}$$

$$\delta N_t = \delta N_{t0} e^{-ht}, \tag{4.130}$$

where δS_0 and δN_{t0} are the initial values of the perturbations, and

$$h = \Gamma_r \pm j\Omega_r. \tag{4.131}$$

The real part is the decay rate

$$\Gamma_r = \frac{1}{2}(\Gamma_{N_t} + \Gamma_S), \tag{4.132}$$

and the radial relaxation oscillation frequency, after some approximations, is given by

$$\Omega_r \approx (G G_{N_t} S)^{1/2}. \tag{4.133}$$

With further substitutions for

$$S = \frac{(I - I_{th})}{(qG)} \tag{4.134}$$

$$I_{th} = q\gamma_e N_{t-th} \tag{4.135}$$

$$G_{N_t} = \frac{\Gamma v_g a}{\text{Vol}}, \tag{4.136}$$

the expression for the relaxation frequency becomes

$$\Omega_r = \left[\frac{1 + \Gamma v_g a N_0 \tau_p}{\tau_p \tau_e} \left(\frac{I}{I_{th}} - 1 \right) \right]^{1/2}. \tag{4.137}$$

where N_0 is the transparency carrier density. Derivation of the small-signal response shows that Ω_r is the key parameter for high-speed lasers; the larger this is, the higher the laser response frequency. The term $\Gamma v_g a N_0 \tau_p$ evaluates to a number close to unity, and therefore may not be neglected. However, making τ_p smaller increases the value of Ω_r. This reduction in τ_p may be obtained by decreasing the facet mirror reflectivities by coatings, or alternatively by making the laser guide more lossy. The parameter τ_e which is the carrier recombination rate is initially determined by the material. After threshold this recombination rate becomes shorter, and little can be done to reduce it further. Reducing I_{th} and increasing the ratio I/I_{th} also increases the magnitude of Ω_r.

The small-signal modulation response calculation uses the rate equations, but in this case the phase Equation (4.103) is also featured. The derivation assumes that the laser is biased above threshold at some current I_b, and the small-signal current modulation term, modulation at ω_m, is given by $I_m \sin(\omega_m t)$. The expression for the total current is

$$I(t) = I_b + I_m \sin(\omega_m t), \tag{4.138}$$

with the assumption that $I_m \ll (I_b - I_{th})$, which implies small-signal modulation.

The total cavity photons may be written as

$$S(t) = S_b + S_m \sin(\omega_m t + \theta_m), \tag{4.139}$$

where S_b is the cavity photons at the bias current of I_b and the second term is the sinusoidal time varying component, with a phase term which usually lags the current component. The carriers in the cavity have a similar expression given by

$$N_t(t) = N_{t-b} + N_{t-m} \sin(\omega_m t + \xi_m), \tag{4.140}$$

with the phase delay ξ_m, which differs from the photon number delay.

Substituting in the rate equations, the solutions for the various small-signal terms may be obtained. The small-signal photons in the cavity, which is also a measure of the output power, may be shown to be

$$S_m = \frac{G_{N_t} S I_m / q}{[(\omega_m^2 - \Omega_r^2 - \Gamma_r^2)^2 + 4\omega_m^2 \Gamma_r^2]^{0.5}}, \tag{4.141}$$

and the phase lag term for the photons is given by

$$\theta_m = \tan^{-1}\left(\frac{2\Gamma_r \omega_m}{\omega_m^2 - \Omega_r^2 - \Gamma_r^2}\right). \tag{4.142}$$

The modulated cavity photons or equivalently the modulated output light power is almost constant when $\omega_m \ll \Omega_r$, then peaks to a maximum near Ω_r, and falls off for $\omega_m > \Omega_r$. Note that the laser is a forward-biased p–i–n diode, and therefore additional series resistance due to contact resistance and parasitic inductance and capacitance from the bond wire and pads results in further degradation of the response. This is shown [33] in Figure 4.18. It may be shown that the 3 dB bandwidth is given by [1], assuming $\Gamma_r \ll \Omega_r$:

$$f_{3dB} = \frac{\sqrt{3}\Omega_r}{2\pi}. \tag{4.143}$$

When the laser current is modulated, the increase and decrease in the current result in the laser effective index varying inversely as the current according to Equation (4.92). Thus, the longitudinal mode cavity changes its electrical length since the index changes with current injection. As the modulation takes place the frequency of emission of the laser keeps changing in synchronism with this modulation. Using the phase rate equation, it may be shown that the change in frequency results in a frequency chirp given by [1]:

$$\delta\nu(t) = \frac{1}{2\pi}\frac{d\phi}{dt} = \frac{\beta_c}{4\pi}\left[G_{N_t}(N_t - N_{t-0}) - \frac{1}{\tau_p}\right], \tag{4.144}$$

where β_c is the linewidth broadening factor of Equation (4.93) and

$$N_{t-0} = \int N_0 dV. \tag{4.145}$$

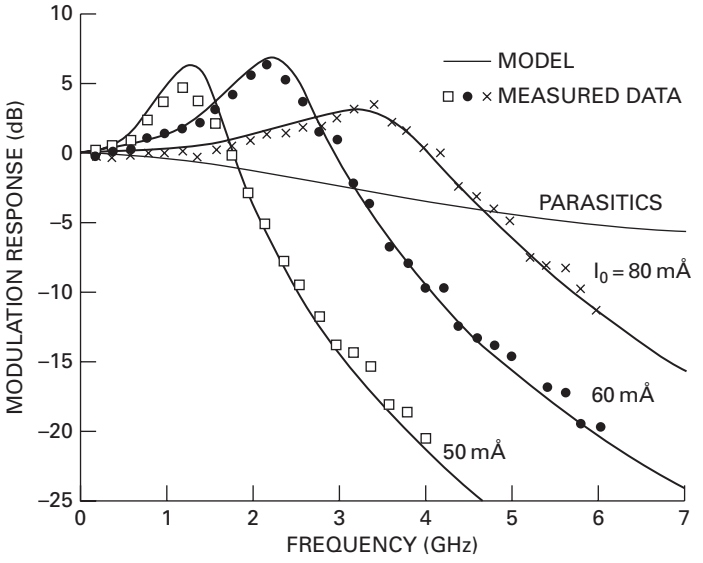

Fig. 4.18 Modulation response for different current drives, showing the effect of parasitics, from [33] with permission (R. S. Tucker and I. P. Kaminow, *Journal of Lightwave Technology*, Vol. 2, No. 4, pp. 385–393, 1984. ©1984 IEEE, OSA).

This linewidth broadening that occurs with chirp is unacceptable for long haul fibre optic systems, and therefore an external modulator is used with dc current into the laser.

The process of obtaining a single longitudinal mode requires a grating in the waveguide either in the whole of the gain section which makes it the distributed feedback (DFB) laser or at the ends as the grating acts as a frequency selective reflector which makes it a distributed Bragg reflector (DBR) laser. This does ensure that a single emission line emerges from the laser, and feeds into an external modulator.

Laser noise

The semiconductor laser has noise associated with the emission, and the primary source of noise is the spontaneous emission, followed by the carrier recombination noise, which is essentially shot noise. Since this is in a cavity, the oscillations are affected by an amplitude or intensity noise component called the *Relative intensity noise* (RIN), and a phase noise component which affects the linewidth of the emission. The method of analysis is to include the noise terms in the rate equations, and traditionally these are the Langevin noise components, which added to the rate equations. These noise terms are assumed to be Gaussian, with a zero mean, and the correlations are assumed to be Markovian, which simplifies the equations. Thus, the rate equations become

$$\frac{dS}{dt} = (G - \gamma_p)S + R_{spon} + F_S(t) \tag{4.146}$$

$$\frac{dN_t}{dt} = \frac{I}{q} - \gamma_e N_t - GS + F_{N_t}(t) \tag{4.147}$$

$$\frac{d\phi}{dt} = -(\omega - \omega_{th}) + \frac{1}{2}\beta_c(G - \gamma_p) + F_\phi(t), \tag{4.148}$$

where $F_S(t)$, F_{N_t} and $F_\phi(t)$ are the Langevin noise components. The mean of each of these components is zero, but their auto- and cross-correlation terms result in so-called *diffusion components* which have specific values. Thus,

$$< F_i(t) > = 0 \tag{4.149}$$

$$< F_i(t), F_j(t') > = 2D_{ij}\delta(t - t') \tag{4.150}$$

where i, j, k are S, N_t, ϕ, and it may be shown [2] that the dominant contributions are from the auto-correlation terms $D_{SS} = R_{sp}S$ and $D_{\phi\phi} = R_{sp}/4S$. The method of solution is by perturbations of all three variables; these lead to equations which give rise to the auto-correlation factors for the perturbations, and consequently noise spectral density is obtained through the Fourier transform. Thus, the spectral density of the photon number or the light intensity leads to the relative intensity noise. The intensity auto-correlation is given by A_{SS}, and thus

$$A_{SS}(\tau) = < \delta S(t)\delta S(t + \tau) > \bar{S}^2, \tag{4.151}$$

where $\bar{S}$ is the time average photon intensity and $\delta S = S - \bar{S}$, represents the fluctuations, and hence the noise. Fourier transform of A_{SS} gives the RIN

$$\text{RIN} = \int_{-\infty}^{\infty} A_{SS}(\tau)e^{j\omega t}dt, \tag{4.152}$$

and it may be shown that [2]

$$\text{RIN} = \frac{2R_{sp}[(\Gamma_{N_t}^2 + \omega^2) + G_{N_t}S^2(1 + \gamma_e N_t/R_{sp}S) - 2\Gamma_{N_t}G_{N_t}S]}{S\left[(\Omega_r - \omega)^2 + \Gamma_r^2\right]\left[(\Omega_r + \omega)^2 + \Gamma_r^2\right]}. \tag{4.153}$$

From this expression, the RIN may be plotted as a function of frequency as shown in Figure 4.19 [2]. The RIN peaks at the relaxation oscillation frequency, and therefore it follows that direct modulation of lasers should not be performed close to f_r.

It may also be shown that the signal-to-noise ratio is given by

$$\text{SNR} = \left(\frac{2\Gamma_r S}{R_{sp}}\right)^{0.5}. \tag{4.154}$$

Using the phase rate equation with the Langevin noise terms (Equation (4.148)), it may be shown [2] that the frequency noise spectral density is given by the expression in the following equation. The frequency noise is the phase noise integrated over time:

$$A_{ff} \cong \frac{R_{sp}}{2S}\left(1 + \frac{\beta_c^2\Omega_r^4}{\left[(\Omega_r^2 - \omega^2)^2 + (2\omega\Gamma_r)^2\right]}\right). \tag{4.155}$$

The above expression shows that the term is flat in the region $\omega \ll \Omega_r$, and similar to the RIN, it peaks at the relaxation oscillation frequency Ω_r, and then falls off rapidly.

The linewidth fluctuation is obtained by considering the electric field fluctuations, and including both the amplitude and phase fluctuations. After considerable simplifications, the FWHM $\Delta\omega = 2\pi\Delta f$ of the line, assumed to be a Lorentzian, is given by [2]

$$\Delta f = \frac{R_{sp}(1 + \beta_c^2)}{4\pi S}. \tag{4.156}$$

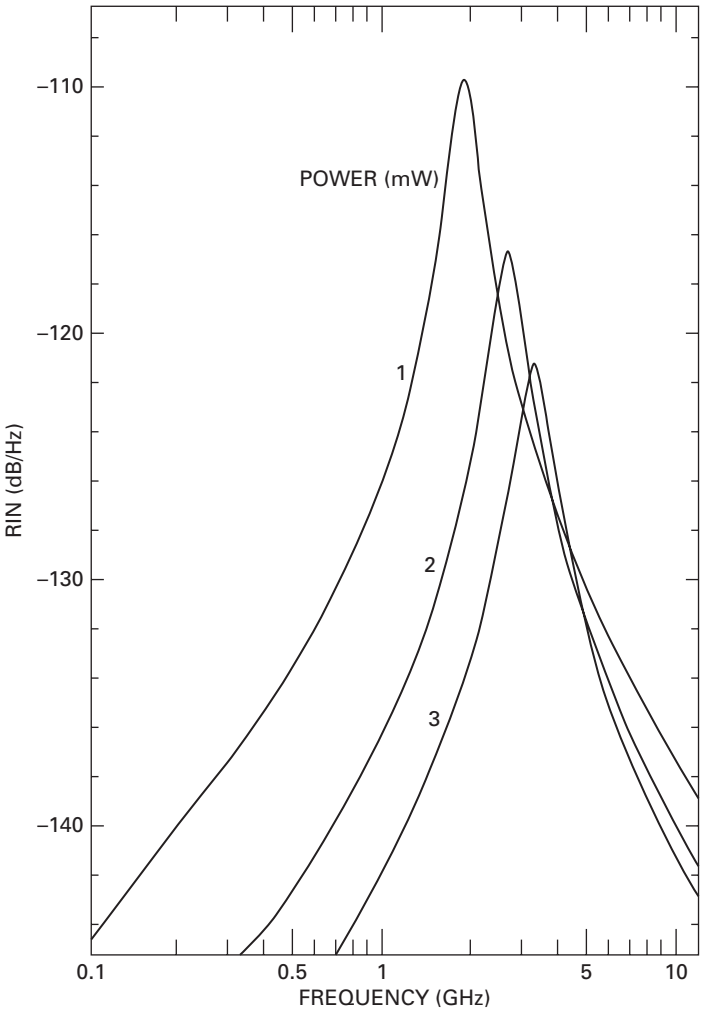

Fig. 4.19 Calculated RIN for a typical 1.3 μm InGaAsP laser at different output power levels (G. P. Agrawal and N. K. Dutta, *Semiconductor Lasers*, Van Norstrand Rheinhold, 1993. ©Springer). With kind permission of Springer Science and Business Media.

This may be written as

$$\Delta f = (1 + \beta_c^2)\Delta f_0, \tag{4.157}$$

where

$$\Delta f_0 = \frac{R_{sp}}{4\pi S}, \tag{4.158}$$

which is the unperturbed linewidth of the laser.

Quantum well and quantum dot lasers

The gain medium considered so far has been assumed to be bulk material, where gain is in the region of $100\,\mathrm{cm}^{-1}$. The use of quantum wells or quantum dots as the gain

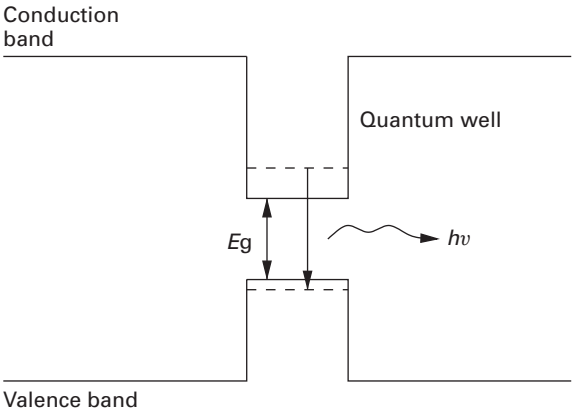

Fig. 4.20 A single quantum well, showing the lowest energy levels in the conduction and valence bands.

medium results in considerable reduction in the number of carriers injected to obtain population inversion and thus the threshold current.

In traditional double heterostructure lasers, the active region is between higher bandgap materials, and the holes and electrons recombine in this region to emit light. To obtain stimulated emission, the energy difference between quasi-Fermi levels for the holes and electrons in this active region needs to be greater than the band gap of the active region. Then population inversion takes place and the light output is dominated by stimulated emission. When the active region becomes very narrow, comparable to the de Broglie wavelength in the material, quantum confinement occurs, and this region becomes a quantum well.

In a quantum well the energy level in the direction across the well, the x direction in Figure 4.20, is quantised, and is a continuum in the other directions, y and z directions. The depth of the well in the conduction band is ΔE_c and in the valence band is ΔE_v, and the emission energy for the recombination from $E1c$ to $E1v$ is greater than (E_g). These levels are calculated by solving the time-independent Schrödinger equation of the well with finite barriers for the conduction band electrons, and for the valence band heavy holes and light holes. The bulk material degeneracy of the heavy hole and light hole bands is removed in the quantum well with the heavy hole band being uppermost, followed by the light hole band, and at a lower level, the split-off band as shown schematically in Figure 4.21. The effective masses of the heavy and light holes are different and therefore the levels are also different. The normal wavelength of the lasing level is from $E1c$ to $E1hh$ which is a TM mode wave.

The density of states per unit volume for the quantum well is given by

$$\rho_{\mathrm{qw-ci}} = \frac{\pi m_{\mathrm{ci}}}{h^2 L_{\mathrm{qw}}}, \tag{4.159}$$

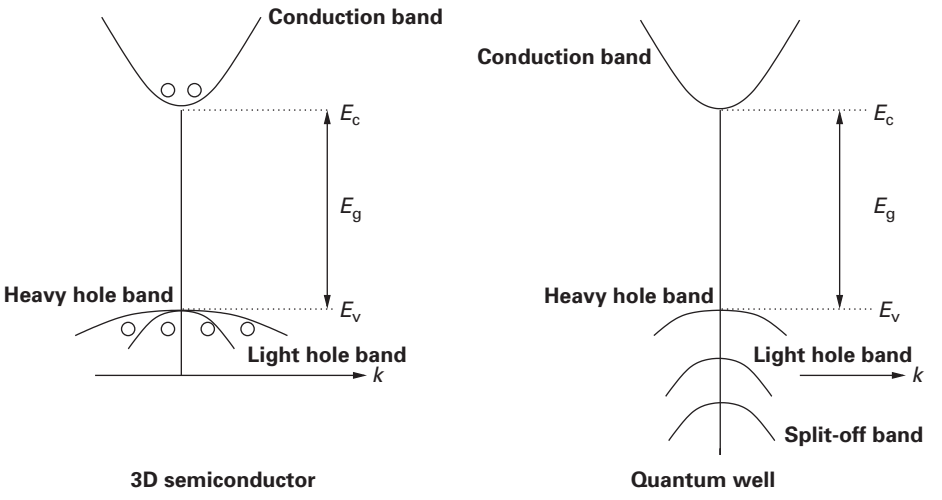

Fig. 4.21 Schematic band diagram of a bulk semiconductor and a quantum well.

where m_{ci} is the effective mass of the electrons in ith sub-band of the quantum well and L_{qw} is the width of the quantum well. The density of states for the valence band is identical except that the appropriate effective masses are used.

The density of states for the conduction band of bulk material, assuming parabolic bands, is of the form:

$$\rho_c(E) = 4\pi \left(\frac{2m_c}{h^2}\right)^{3/2} E^{1/2}. \tag{4.160}$$

Comparing the two, it is apparent that the density of states for the quantum well is independent of energy. Furthermore, it may be shown that the density of states for the quantum well is much smaller than that of the bulk material. Thus, the current density for threshold is generally much smaller than that for bulk material. The current state of the art structures have current densities of the order of $100\,\mathrm{A\cdot cm^{-2}}$ per quantum well, in contrast to bulk structures which have threshold current densities of almost ten times this figure. The maximum gain of quantum wells is estimated at between $1000\,\mathrm{cm^{-1}}$ and almost $10\,000\,\mathrm{cm^{-1}}$. However, the confinement factor of a single quantum well is of the order of 0.01–0.02, which makes the net gain of the order of $100\,\mathrm{cm^{-1}}$. In contrast, bulk structures have gains of the order of $100\,\mathrm{cm^{-1}}$, and confinement factors of about 0.20–0.40. In multiple quantum well structures, the spacing between the wells is chosen to be sufficiently large so that the wells are not coupled. Typical well widths are about 10 nm, and the barriers between them are also of similar widths.

Quantum well lasers come in different forms, from the single quantum well laser to the multi-quantum well laser, and these may be incorporated in the waveguides for lasers discussed earlier. Figure 4.22 sketches some of these structures from the graded index single quantum well laser to the separate confinement single and multiple quantum well lasers.

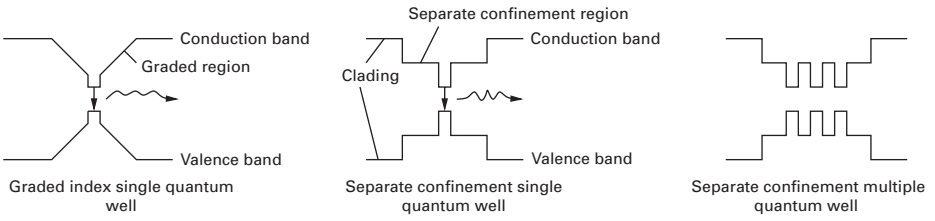

Fig. 4.22 Schematic band diagrams of the graded index single quantum well laser structure, and separate confinement single and multiple quantum well laser structures.

Strained quantum well lasers have the quantum wells with compressive or tensile strain, and sometimes the barriers between wells are also strained. The compressive strain increases the band gap, and also the separation between the heavy hole and light hole bands, and the emission is in TM mode. With tensile strain the band gap is reduced, and the gap between the light hole and heavy hole bands is reduced, and may even push the light hole band above the heavy hole band. With the light hole band above the heavy hole band, the lasers emit in the TE mode. Strained quantum well lasers have lower threshold current densities and are sometimes preferred over unstrained well structures.

Quantum well lasers show better performance compared to the bulk lasers with reduced threshold current densities, better linewidths and better chirp performance. With the distributed Bragg grating to obtain single longitudinal mode with further modifications, these have become the lasers of choice for fibre optical systems.

Quantum dot lasers are a class of lasers that use confinement to improve on the performance of the quantum well lasers. In these devices, the gain is through current injection into quantum dots. A quantum dot is a three-dimensional structure in which every dimension is less than the de Broglie wavelength in the material, and quantum confinement occurs. The density of states in this case is a delta function, and in principle the number of carriers to be injected is very small. Quantum dots are created by the Stranski–Krastinov growth technique [5], with further modifications, by allowing a strained layer to relax, and this results in self-organised dots. These dots need to have wetting layers to make contacts, and since single layer of dots does not provide sufficient gain this composite of dot layers and wetting layers need to be repeated many times. Lasers made of these active layers show reduced threshold current densities, linewidth reduction, and reduced chirp [11] compared to quantum well lasers. The major problem with these lasers is that the dots vary in size and so the distribution affects the linewidth.

Vertical cavity surface emitting lasers (VCSELs) are another class of lasers that use quantum wells for the gain medium. In this case, the cavity is usually one wavelength long, and the quantum well is placed in the middle so as to provide maximum gain to the standing wave within the cavity. Since the gain is so small, the mirror reflectivities need to be extremely high, as close to unity as possible. This is obtained by a multi-layer dielectric stack mirror with reflectivities of 0.997 for the top output mirror and 0.999 for the lower mirror. A schematic diagram of a VCSEL is shown in Figure 4.23. Currently, VCSELs have been built in the visible, at 840 nm, 980 nm, and also at the longer wavelengths of 1300 nm and 1550 nm. These devices may be built with sub-milliampere

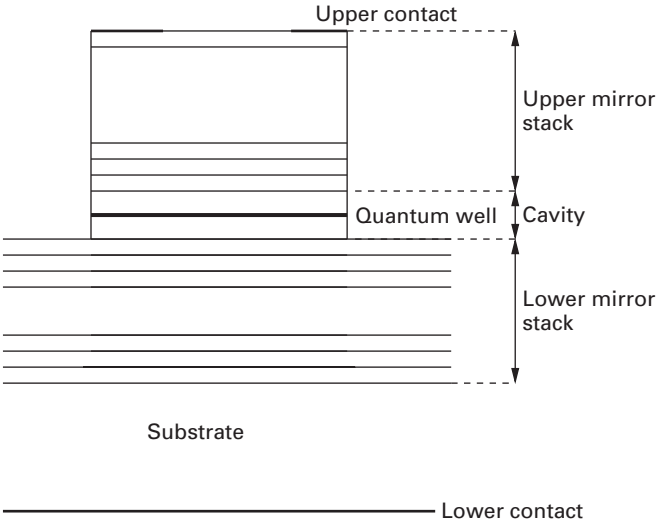

Fig. 4.23 Schematic diagram of a single quantum well vertical cavity laser.

threshold current. These lasers are also one of the most efficient devices with very high wall plug efficiencies. The 840 nm VCSELs are used in data-communication applications where data is transferred between computers, processors and storage media. The linewidth of these lasers is generally high, of the order of 0.1 nm, and direct modulation of these results in linewidth broadening. These lasers find application in coarse wavelength division multiplexed systems, where channel spacing may be as high as 40 nm, and in the 10 GHz Ethernet applications.

High-speed lasers

High-speed lasers need to have high relaxation frequencies, Ω_r or f_r. The expressions for Ω_r are repeated here:

$$\Omega_r \cong (G G_{N_t} S)^{1/2}. \tag{4.161}$$

With the substitution for S,

$$S = \frac{(I - I_{th})}{(qG)} \tag{4.162}$$

results in

$$\Omega_r = \left[\frac{G_{N_t}(I - I_{th})}{q}\right]^{0.5}. \tag{4.163}$$

Also note that

$$I_{th} = q\gamma_e N_{t-th} \tag{4.164}$$

$$G_{N_t} = \Gamma v_g a/\text{Vol}, \tag{4.165}$$

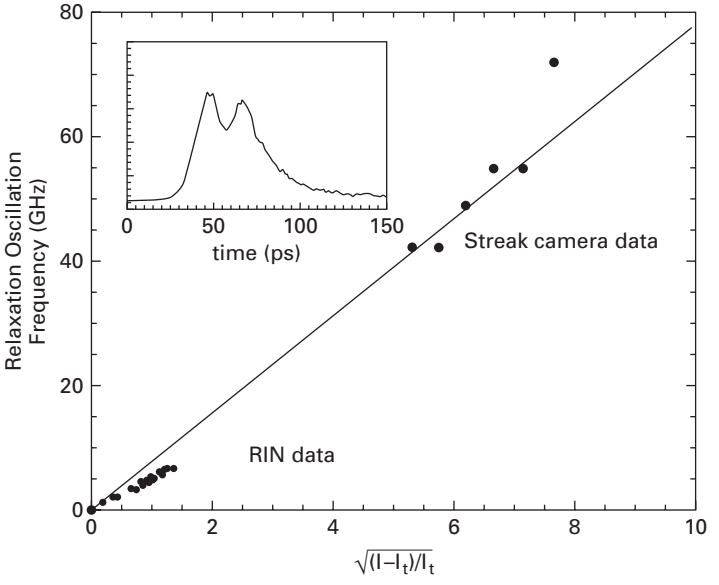

Fig. 4.24 Plot of relaxation oscillation frequency against square root of normalised current. Reprinted with permission from D. Tauber, G. Wang, R. S. Geels, J. E. Bowers, L. A. Coldren, *Applied Physics Letters*, Vol. 62, No. 4, 1993. Copyright 1993, American Institute of Physics.

the expression for the relaxation frequency becomes

$$\Omega_r = \left[\frac{1 + \Gamma v_g N_0 \tau_p}{\tau_p \tau_e}\left(\frac{I}{I_{th}} - 1\right)\right]^{1/2}. \tag{4.166}$$

Equation (4.161) suggests that high values of G_{N_t} are required, which is essentially that $\partial g/\partial N$ is high per unit volume, and S needs to be high, which implies that the photon density N_{ph} should be high in the cavity. From Equation (4.163), it follows that I_{th} should be as low as possible. The small-signal derivation in Equation (4.145) shows again that f_{3dB} is determined by Ω_r. Thus, this equation may be written as

$$f_{3dB} \approx \frac{\sqrt{3}\Omega_r}{2\pi}. \tag{4.167}$$

Ideally, the quantum dot lasers should have the highest relaxation oscillation frequency. However, the problem of the parasitics of the wetting layer and other issues [11] result in these lasers not being as fast as the quantum well tunneling injection lasers discussed below.

Of the lasers discussed above, the lowest threshold current device is the VCSEL, and early measurements have shown that the relaxation oscillation frequency measured by streak camera [31] was 84 GHz shown in Figure 4.24. However, the problem is that the parasitics of the device in its present form, with the current passing through the entire top and bottom mirrors, are large. Thus, the device parasitics should be made as low as possible, as well as the packaging parasitics, as otherwise the modulation signal will be attenuated severely at high frequencies. A high-speed VCSEL operating at 35 GHz

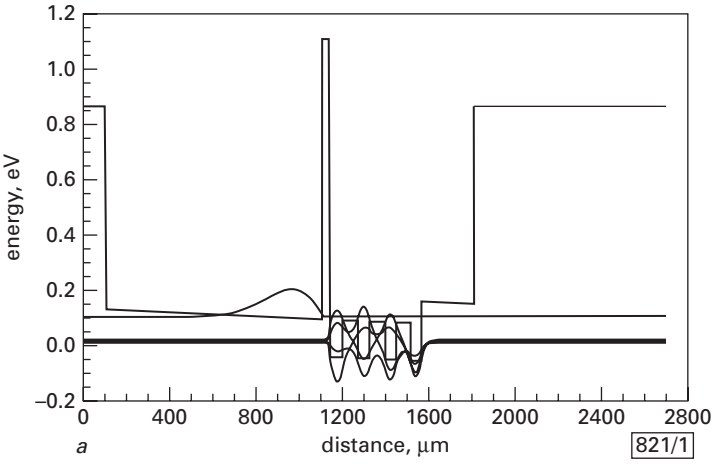

Fig. 4.25 Plot of the band diagram of the conduction band of the tunnelling injection laser (X. Zhang, A. L. Gutierrez-Aitkens, D. Klotzkin, P. Bhattacharya, C. Caneau, R. Bhatt, *Electronics Letters*, Vol. 32, No. 18, pp. 1715–1717, 1996. ©1996 IEEE).

[6] has the laser drive and modulation current pass through the top mirror but only through a small number of layers of the lower mirror to achieve this result.

The approach by Bhattacharya's group [35] has resulted in tunneling injection quantum well lasers that have the potential for very high modulation rates. The claim is that when the electrons travel from the cladding and fall into the separate confinement region, they gain energy and become hot. The electrons diffuse across the separate confinement region (SC), and when they fall into the quantum wells, they become even more hot. The holes having a large mass are not as mobile, and therefore are largely unaffected by these changes in potential. The approach is to have a reservoir of electrons in the SC region, and these tunnel out into the SC region, and also into the quantum wells with only small gains in energy, to allow high modulation rates. The theory discussed above for Ω_r does not account for the temperature of the carriers in determining its value. A more elaborate theory would do this, but only qualitative explanations are therefore available.

Figure 4.25 shows the conduction band diagram of the device. Careful design of the reservoir, its width and its level with respect to the SC region and the quantum wells are necessary. Again the device parasitics need to be low. The results are indeed impressive, 3 dB bandwidths of 100 GHz, and these are probably the fastest lasers built to date. Note that the structure used in high-speed quantum dot lasers has also used the tunnelling structure, but these are not as fast as the quantum lasers discussed above.

4.3 Photodetectors

Photodetectors are devices used to convert light signals into electrical versions. The performance of the different types of detectors is determined by their quantum efficiency,

their frequency response and responsivity. In this chapter only solid state detectors are considered, and the slower detectors are only briefly discussed. The photomultiplier, which is widely used as a sensitive and fast (of the order of 100 MHz) detector, will not be discussed here. Solid state detectors considered here are the extrinsic kind, in which the photon energy is close to the band gap of the semiconductor material used for the detector. Intrinsic photodetectors are used for detection of light with energies below the band gap, and depend on deep level traps, or with different energy levels in a quantum well, but these are not discussed here. In this chapter, the detectors considered are photoconductors, and the junction devices which include the p–i–n diode, the metal–semiconductor–metal (MSM) photodetector and the avalanche photodetector (APD). The photoconductor is a slow device and also noisy, but its simplicity is an attractive feature.

4.3.1 Preliminaries

Consider a semiconductor photodetector which absorbs photons with energies at or above E_g, the bandgap energy of the semiconductor. Suppose the incident optical power is given by P_{in}, and it is assumed that all the incident photons enter the semiconductor. Suppose the photocurrent generated as a result of this incident optical power is given by I_p, then the relationship between P_{in} and I_p is

$$I_p = R P_{in}, \tag{4.168}$$

where R is the responsivity of the photodetector in units of AW^{-1}.

The quantum efficiency of the detector η may be defined as ratio of the number of hole–electron pairs generated to the number of incident photons, and is given by

$$\eta = \frac{\dfrac{I_p}{q}}{\dfrac{P_{in}}{h\nu}} = \frac{h\nu}{q} R. \tag{4.169}$$

Thus, R may be written as

$$R = \frac{q\eta}{h\nu} = \frac{\eta\lambda q}{hc}. \tag{4.170}$$

Suppose the thickness of the semiconductor is w, and the absorption coefficient is α Np m^{-1}, then the transmitted optical power escaping from the semiconductor is given by

$$P_{tr} = P_{in} e^{-\alpha w}. \tag{4.171}$$

Thus, the absorbed power is given by

$$P_{abs} = P_{in} - P_{tr} = P_{in}(1 - e^{-\alpha w}). \tag{4.172}$$

Since every absorbed photon creates one hole–electron pair, the quantum efficiency is given by

$$\eta = \frac{P_{abs}}{P_{in}} = (1 - e^{-\alpha w}), \tag{4.173}$$

which assumes that all the incident photons enter the semiconductor with no reflection.

4.3.2 Photoconductor detectors

The photoconductor detector depends on the increase in conductivity of a semiconductor when illuminated with photons of energy above the band gap. The absorbed light creates hole–electron pairs, which increases the conductivity, and with an applied bias, the excess carriers drift to the appropriate electrodes, and constitute an increase in current. Holes and electrons created by the light may be swept out before they recombine by the applied bias field. Alternatively, they recombine as they drift towards the appropriate electrode. The electrons are swept out faster than the holes, and to maintain charge neutrality, more electrons are injected, and this constitutes gain.

A typical photoconductor detector takes the form of a slab of material of thickness a, width b and length L, with ohmic contacts at the sides as shown in Figure 4.26.

The dark current flowing in the slab is given by

$$I = qab(n\mu_n + p\mu_p)\frac{V}{L}, \tag{4.174}$$

where n and p are the free electron and hole number densities, μ_n and μ_p are the mobilities of the electrons and holes respectively, and V is the applied voltage across the slab. When illuminated, the conductivity increases due to the electron–hole pairs created by the photons. Thus, the current when the slab is illuminated is given by

$$(I + \Delta I) = qba[(n + \Delta n)\mu_n + (p + \Delta p)\mu_p]\frac{V}{L}, \tag{4.175}$$

where I is the dark current, assumed to be small. The increase in current due to the illumination is

$$\Delta I = qab(\Delta n\mu_n + \Delta p\mu_p)\frac{V}{L}. \tag{4.176}$$

Suppose the minority carriers are electrons, then the rate equation for the electrons takes the form:

$$\frac{d\Delta n}{dt} = \frac{\eta P}{hvabL} - \frac{\Delta n}{\tau}, \tag{4.177}$$

where η is the quantum efficiency, which determines the number of hole–electrons generated per photon, usually taken to be unity, and P is the optical power absorbed. Note that an equal number of holes as electrons are generated, but remain as the majority carrier. Under steady state, the time variation is set to zero, and then

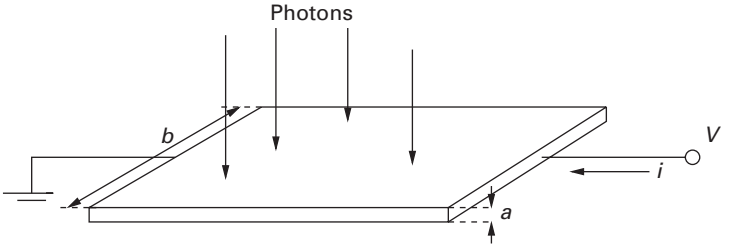

Fig. 4.26 Schematic diagram of a photoconductor slab: length between contacts L, width b and thickness a.

$$\Delta n = \frac{\eta P \tau}{h\nu abL}. \tag{4.178}$$

This assumes that the photon-generated hole-electron pairs are in the volume of the slab, and the volume is abL. The electron current due to these excess electrons is given by

$$\Delta I_{\mathrm{n}} = q\Delta n \mu_{\mathrm{n}} ba \frac{V}{L} = \frac{q\eta PG}{h\nu}, \tag{4.179}$$

where G is the gain of the device. Substituting for Δn from Equation (4.178), G is defined as

$$G = \frac{\mu_{\mathrm{n}} \tau V}{L^2} = \frac{\tau}{t_{\mathrm{tr,n}}}, \tag{4.180}$$

where the electron transit time $t_{\mathrm{tr,n}}$ is given by

$$t_{\mathrm{tr,n}} = \frac{L}{v_{\mathrm{n}}} = \frac{L}{\mu_{\mathrm{n}} \mathcal{E}} = \frac{L^2}{\mu_{\mathrm{n}} V}. \tag{4.181}$$

In addition to the electrons, adding the motion of the holes results in this equation becoming

$$G = \frac{(\mu_{\mathrm{n}} + \mu_{\mathrm{p}})\tau V}{L^2} = \frac{\mu_{\mathrm{n}} + \mu_{\mathrm{p}}}{\mu_{\mathrm{n}}} \frac{\tau}{t_{\mathrm{tr,n}}}. \tag{4.182}$$

Note that since $\mu_{\mathrm{h}} \ll \mu_{\mathrm{n}}$, the expression for gain is that given in Equation (4.180).

Making L as small as possible increases the gain and also reduces the response time. If the carriers are swept out before they recombine, then the electrons reach the ohmic contact before the holes. To maintain charge neutrality, extra electrons are injected into the photoconductor. If sweep out of the carriers occurs, then the lifetime is the transit time of the carriers and is given by

$$t_{\mathrm{tr,n,p}} = \frac{L}{\mu_{\mathrm{n,p}} \left(\frac{V}{L}\right)} = \frac{L^2}{\mu_{n,p} V}. \tag{4.183}$$

Since the mobility of holes is lower than that of electrons, the transit time of electrons is smaller. If $\mu_{\mathrm{p}} \ll \mu_{\mathrm{n}}$, then the gain term is dominated by the electron mobility term and in turn the transit time

$$G = \frac{\tau}{t_{\mathrm{tr,n}}}. \tag{4.184}$$

Thus, the photon-induced current from Equation (4.179) becomes

$$I_{\mathrm{ph}} = \Delta I_{\mathrm{n}} = \frac{\tau}{t_{\mathrm{tr,n}}} \frac{q\eta P}{h\nu}. \tag{4.185}$$

In Equation (4.177), the time variation of Δn may be as $e^{\mathrm{j}\omega \mathrm{t}}$, and in the small-signal case the response time varies inversely as $(1 + j\omega\tau)$, which defines the bandwidth. Thus, if the optical power is given by

$$P(\omega) = P_{\mathrm{opt}} + P_1 e^{\mathrm{j}\omega \mathrm{t}}, \tag{4.186}$$

then the ac current term is

$$\Delta I_{\mathrm{ac}} = \frac{q\eta P_1}{h\nu} \frac{\tau}{t_{\mathrm{tr,n}}} \frac{1}{1 + j\omega\tau}. \tag{4.187}$$

The rms current magnitude is given by

$$\Delta I_{\mathrm{ac,rms}} = \frac{q\eta P_{\mathrm{l}}}{\sqrt{2}h\nu}\frac{\tau}{t_{\mathrm{tr,n}}}\frac{1}{(1+\omega^2\tau^2)^{1/2}}. \tag{4.188}$$

Noise in these detectors is from several sources and these are thermal or Johnson noise, generation-recombination noise, and at low frequencies, the $1/f$ flicker noise. If the resistance of the device is R_{cond}, then the thermal noise current is given by

$$\left\langle |\imath_{\mathrm{th}}|^2 \right\rangle = \frac{4k_{\mathrm{B}}T\Delta f}{R_{\mathrm{cond}}}, \tag{4.189}$$

where Δf is the bandwidth. The generation-recombination noise term is given by

$$\left\langle |\imath_{\mathrm{GR}}|^2 \right\rangle = \frac{4qGI_{\mathrm{o}}\Delta f}{1+\omega^2\tau^2}. \tag{4.190}$$

where $I_{\mathrm{o}} = q\eta P_{\mathrm{opt}}G/\omega\tau$. Neglecting the contribution of 1/f noise at low frequencies, the signal-to-noise ratio is given by

$$\frac{S}{N} = \frac{\Delta I^2_{\mathrm{ac,rms}}}{\left\langle |\imath_{\mathrm{th}}|^2 \right\rangle + \left\langle |\imath_{\mathrm{GR}}|^2 \right\rangle} \tag{4.191}$$

$$= \frac{\eta P_{\mathrm{l}}^2}{8h\nu P_{\mathrm{opt}}\Delta f}\left[1 + \frac{k_{\mathrm{B}}T}{Gq}(1+\omega^2\tau^2)\frac{1}{R_{\mathrm{cond}}I_{\mathrm{o}}}\right]^{-1}. \tag{4.192}$$

The noise equivalent power (NEP) is defined as the incident rms optical power required to produce a signal-to-noise ratio of unity in a bandwidth of 1 Hz, [3], and the detectivity of a detector is the inverse of NEP or $\mathrm{D} = (\mathrm{NEP})^{-1}$. The normalised detectivity D^* is defined as

$$D^* = \frac{A^{1/2}(\Delta f)^{1/2}}{NEP} \quad \left(\mathrm{cm.Hz^{1/2}\,W^{-1}}\right), \tag{4.193}$$

where A is the area of the photoconductor on which light is incident. The bandwidth is usually set to 1 Hz, and the reference area is set to 1 cm^2. Note that D^* is usually expressed as $D^*(\lambda, f, 1)$, where λ is the wavelength of light, f is its frequency of modulation, and 1 is the bandwidth in Hz and may be obtained from Equation (4.192).

The structure of these devices may take various forms, including the interdigitated surface contact structure shown in Figure 4.27.

4.3.3 P–I–N diodes

Detection of photons with energies at or above the band gap of a semiconductor requires that they are absorbed and create hole–electron pairs, and a current be induced due to this absorption. The depletion layer of a reverse-biased p–n junction of the semiconductor causes the holes and electrons to separate and be collected by the appropriate contact/collection region. The photons entering this device, schematically shown in Figure 4.28, create hole–electron pairs as the light is absorbed, from the top contact region, through the depletion layer on both sides of the junction and possibly beyond, to the lower contact region. The fields in the depletion layer are high enough to separate the

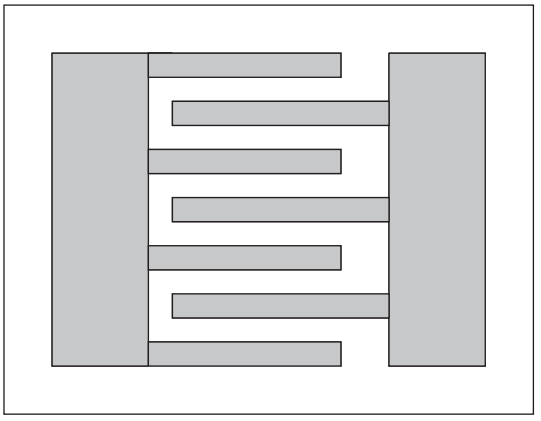

Fig. 4.27 Schematic diagram of a photoconductor detector with interdigitated fingers.

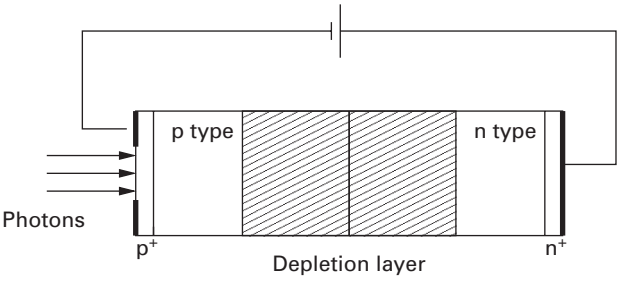

Fig. 4.28 Schematic diagram of a p–n junction photodiode.

holes and electrons, and send them to the respective majority carrier region, holes to the p region and electrons to the n region, because of the reverse bias. Holes and electrons generated in the p and n contact regions need to be considered differently. The minority carriers, electrons in the p region and holes in the n region, about one diffusion length from the depletion layer, diffuse towards the depletion layer and are accelerated to the appropriate majority carrier region. This diffusion is a slow process, which degrades the response of the diode detector.

The alternative to the simple p–n junction detector is to use a p–i–n structure, the 'i' region is either an 'i' layer or a p^- or n^- layer. In this device, the depletion layer extends through the whole of the 'i' region with no bias or with a negative bias. The usual top p^+ layer is made thin to ensure that absorption in the top contact layer is negligibly small. Most of the absorption is in the 'i' or n^- layer, with some in the lower n^+ layer, as shown in Figures 4.29 and 4.30.

Suppose the photon flux per unit area is given by ϕ_0, and the absorption coefficient is given by α, then the hole–electron generation rate is given by

$$Gen(z) = \phi_0 \alpha e^{-\alpha z}, \tag{4.194}$$

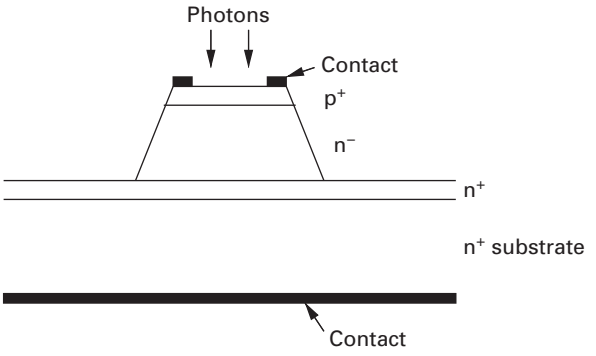

Fig. 4.29 Schematic diagram of a p–n junction photodiode in mesa form, with a ring top contact.

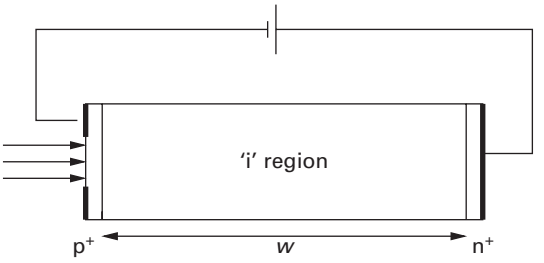

Fig. 4.30 Schematic diagram of a p–i–n junction photodiode.

and the photon flux density ϕ_0 is given by

$$\phi_0 = \frac{P_{\text{inc}}(1-R)}{Ahv} \tag{4.195}$$

where the incident optical power is given by P_{inc}, R is the reflectivity of the surface and A is the device area. Note that the extra α introduced in the generation term is to normalise the current density. The drift current density is given by

$$J_{\text{drift}} = -q\int_0^w Gen(z)dz = q\phi_0(1-e^{-\alpha w}). \tag{4.196}$$

The tail end of the optical power also enters the lower n^+ region, and generates hole–electron pairs. In this region, the hole, minority carrier, motion is determined by the diffusion equation

$$D_{\text{p}}\frac{\partial^2 p_{\text{n}}}{\partial z^2} - \frac{p_{\text{n}} - p_{\text{n0}}}{\tau_{\text{p}}} - Gen(z) = 0, \tag{4.197}$$

where D_{p} is the diffusion coefficient for holes, τ_{p} is the lifetime for holes beyond p_{n0}, the equilibrium hole density for the n^+ doped layer. The boundary conditions are $p_{\text{n}} = p_{\text{n0}}$ for $z = \infty$ and $p_{\text{n}} = 0$ for $z = w$; the former has some validity since the structure is generally grown on an n^+ substrate. The latter boundary condition has the excess hole density to be zero since those generated at the boundary are acted upon by the depletion layer and accelerated away. The solution is given by

$$p_{\mathrm{n}} = p_{\mathrm{n0}} - (p_{\mathrm{n0}} + Ce^{-\alpha \mathrm{w}})e^{(\mathrm{w}-\mathrm{z})/\mathrm{L_p}} + Ce^{-\alpha \mathrm{z}}, \tag{4.198}$$

where $L_{\mathrm{p}} = \sqrt{D_{\mathrm{p}}\tau_{\mathrm{p}}}$ and

$$C = \frac{\phi_0 \alpha L_{\mathrm{p}}^2}{D_{\mathrm{p}}\left(1 - \alpha^2 L_{\mathrm{p}}^2\right)}. \tag{4.199}$$

The diffusion current density is

$$J_{\mathrm{diff}} = -qD_{\mathrm{p}}\left(\frac{\partial p_{\mathrm{n}}}{\partial z}\right) \tag{4.200}$$

$$= q\phi_0 \frac{\alpha L_{\mathrm{p}}}{1 + \alpha L_{\mathrm{p}}} e^{-\alpha \mathrm{w}} + qp_{\mathrm{n0}}\frac{D_{\mathrm{p}}}{L_{\mathrm{p}}}, \tag{4.201}$$

and the total current density is the sum of the drift current density and the diffusion current density, and is given by

$$J_{\mathrm{tot}} = q\phi_0\left(1 - \frac{e^{-\alpha \mathrm{w}}}{1 + \alpha L_{\mathrm{p}}}\right) + qp_{\mathrm{n0}}\frac{D_{\mathrm{p}}}{L_{\mathrm{p}}}. \tag{4.202}$$

The value of p_{n0} is small in the n^+ region, and therefore under illumination, this last term is small and is usually omitted, and the current is proportional to the photon flux.

The quantum efficiency may be obtained from these expressions:

$$\eta_{\mathrm{ext}} = \frac{J_{\mathrm{tot}}A/q}{P_{\mathrm{opt}}/h\nu} = (1 - R_{\mathrm{r}})\left(1 - \frac{e^{-\alpha \mathrm{w}}}{1 + \alpha L_{\mathrm{p}}}\right), \tag{4.203}$$

where P_{opt} is the optical power absorbed and R_{r} is the reflectivity. To make the quantum efficiency high, the diode reflectivity needs to be made small using an anti-reflection coating, so that $R_{\mathrm{r}} \approx 0$. Also $\alpha w \gg 1$ makes η_{ext} closer to unity but at the expense of transit time delay, and hence the frequency response becomes small.

The frequency response of the photodiode is governed by several factors, and these are discussed below. The diffusion of the minority carriers generated outside the depletion region and their transit to the appropriate electrode region are major limitations to the response time of the diode. However, careful design may minimise this effect, including the use of heterostructures so that the p^+ and the n^+ regions are in higher bandgap material, which prevents the generation of hole–electron pairs in these contact regions. The RC time constant of the diode, where C is its capacitance and R is shunt resistance added to extract the signal, determines the circuit response, and is minimised by suitable choice of R, and the area of the diode. The carrier transit time across the depletion region is also a major factor, and this is analysed below.

Suppose the incident optical flux is modulated to have a time varying component of the form $\phi_1 e^{j\omega_{\mathrm{m}}t}$. At any point z in the depletion layer, the carriers generated have to drift to the appropriate electrode region at the saturation velocity v_{s}. Thus, the conduction current density due to the carriers generated at z has a phase delay of $e^{-\mathrm{j}\omega_{\mathrm{s}}\mathrm{z}/\mathrm{v_s}}$ and is given by

$$J_{\mathrm{cond}}(z) = q\phi_1 e^{\mathrm{j}\omega_{\mathrm{m}}\mathrm{t}} e^{-\mathrm{j}\omega_{\mathrm{m}}\mathrm{z}/\mathrm{v_s}}. \tag{4.204}$$

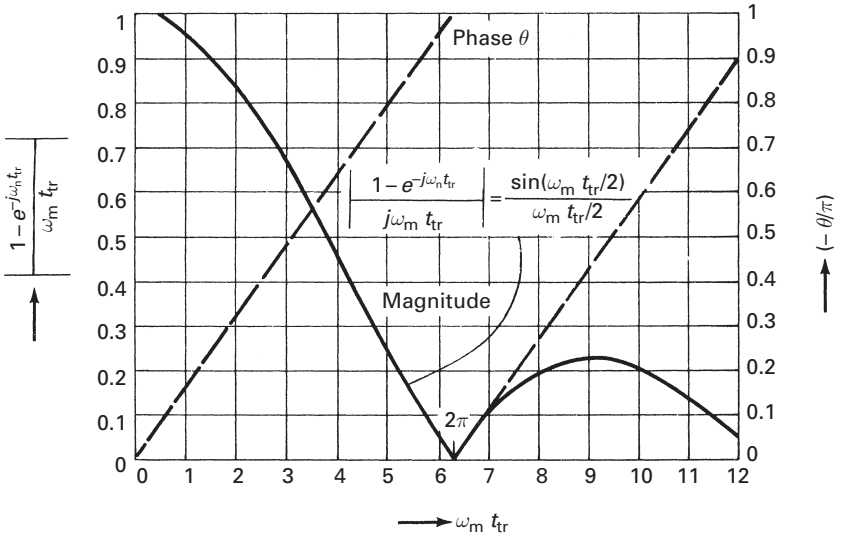

Fig. 4.31 Plot of the transit time factor in Equation (4.206) against the product of the transit time and the modulation frequency (A. Yariv, *Optical Electronics in Modern Communications*, Fifth Edition, Oxford University Press 1997. Figure 4.11–4.17. By permission of Oxford University Press, Inc.).

Considering the carriers generated over the entire depletion layer, neglecting the absorption coefficient, the total conduction current density is given by

$$J_{\text{cond}} = \frac{1}{w}\int_0^w q\phi_1 e^{\text{j}\omega_m t} e^{-\text{j}\omega z/v_s} dz \tag{4.205}$$

$$= q\phi_1 \left(\frac{1-e^{-\text{j}\omega_m t_{tr}}}{j\omega t_{tr}}\right) e^{\text{j}\omega_m t}, \tag{4.206}$$

where t_{tr} is the transit time equal to w/v_s. In addition to the conduction current density, the displacement current density is given by

$$J_{\text{disp}} = j\omega_m \epsilon \mathcal{E} = \frac{j\omega_m \epsilon V}{w}. \tag{4.207}$$

The total current density is the sum of these two terms. The conduction current density is reduced by the transit time factor as shown in Equation (4.206). Figure 4.31 shows a plot of this factor $(1 - e^{-\text{j}\omega_m t_{tr}})/(\omega_m t_{tr})$ against $\omega_m t_{tr}$. Note that when the denominator becomes 2π, this factor goes to zero.

The expression for the transit time effect in Equation (4.206) becomes $1/\sqrt{2}$ when the value of $\omega_m t_{tr}$ is 2.4, and also has a phase shift of $2\pi/5$. Thus, the 3 dB response is given by

$$f_{3\,\text{dB}} = \frac{2.4}{2\pi t_{tr}} = \frac{0.382 v_s}{w} = 0.382 v_s \alpha, \qquad \text{for } \alpha w = 1. \tag{4.208}$$

Note that including the effect of the absorption coefficient in the above derivations, Equations (4.204)–(4.208) in fact reduces the transit time factor further, and therefore these results are optimistic.

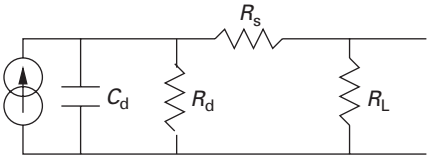

Fig. 4.32 Equivalent circuit of a p–i–n photodiode.

The equivalent circuit of the p–i–n photodiode is shown in Figure 4.32, and consists of a current generator, the diode incremental resistance R_d in parallel, the diode capacitance C also in parallel, the series resistance R_s, which includes the contact resistance and the load resistance R_L also in parallel. The inductance of the wire bond or contact line in series with R_s has been omitted. Note that R_L is generally small compared to R_d.

The analysis of this section shows that to make frequency response of the photodiodes high, the capacitance of the diode should be low so that the R_LC time constant is small and the 'i' layer thickness should be between $1/\alpha$ and $2/\alpha$, where α is the absorption coefficient, and this comes at the expense of responsivity. Using higher bandgap materials for the p^+ and n^+ regions allows higher fields to be used as the breakdown voltage is increased. The usual choice is to make R_LC time constant equal to the transit time across the depletion region. To reduce the capacitance, the area is made small, of the order of 50 μm^2 or less. High-speed p–i–n diodes have been designed and built for many years, and currently diodes with responses in the 50 GHz region are available [10]. To reduce the transit time effects, the absorbing layer is made as thin as possible, of the order of 0.15 μm, and this is at the expense of responsivity. To overcome the reduced absorption in the thin 'i' layer, the use of a dielectric mirror above the contact layer and an epitaxial mirror below the absorbing layer causes the light to have multiple passes, and improves the responsivity [10]. The problem of having this in the n^+ layer requires careful design of the doping of the heterostructure mirror, since current has to flow through it. Note that the lower mirror is similar to those used in VCSELS and there the mirror series resistance problem has been alleviated considerably.

Heterostructure diodes in which the p^+ and n^+ regions are of a larger bandgap material, with the absorption 'i' or n^- layer, have also been designed. The advantage of this structure is that the incoming light is not absorbed in the heavily doped regions, and the depletion fields may be higher. However, the heterojunctions may have to be graded to prevent carrier trapping.

The noise output of the p–i–n diode is dominated by the thermal or Johnson noise. Then the input signal photocurrent is given by

$$I_{ph} = \frac{q\eta_{int}P_{in}}{h\nu} \equiv RP_{in}, \tag{4.209}$$

where R is the diode responsivity in $A\,W^{-1}$. The shot noise is due to the background optical power generating a current given by I_B, the dark current I_d due to the reverse-biased p–i–n diode and the photocurrent I_{ph}, and is given by

$$\left\langle |\sigma_s|^2 \right\rangle = 2q(I_{ph} + I_B + I_d)\Delta f. \tag{4.210}$$

The diode is usually connected to a preamplifier, noise figure F_n, input resistance R_a, which is in parallel with the diode output resistance R_d and the load resistance across the diode R_L, and neglecting the series resistance R_s which is small, the equivalent resistance given by

$$\frac{1}{R_{eq}} = \left(\frac{1}{R_d} + \frac{1}{R_L} + \frac{1}{R_a}\right) \approx \frac{1}{R_L}, \tag{4.211}$$

and the corresponding Johnson or thermal noise current squared, including the preamplifier noise figure, is given by

$$\left\langle |\sigma_t|^2 \right\rangle = \frac{4k_b T F_n \Delta f}{R_{eq}}. \tag{4.212}$$

Thus, the signal-to-noise ratio is given by

$$\frac{S}{N} = \frac{\left(\frac{q\eta_{int} P_{in}}{h\nu}\right)^2}{2q(I_{ph} + I_B + I_d)\Delta f + \frac{4k_b T F_n \Delta f}{R_{eq}}}. \tag{4.213}$$

Usually, the shot noise current is small, R_{eq} is approximately equal to R_L, and the signal-to-noise ratio becomes

$$\frac{S}{N} = \frac{\left(\frac{q\eta_{int} P_{in}}{h\nu}\right)^2}{\frac{4k_b T \Delta f}{R_L}} \tag{4.214}$$

Thus, to make the signal-to-noise ratio large, R_L needs to be made as large as possible, but the problem here is that the RC time constant then becomes large and the response time becomes slow. Thus, the values of R_L are of the order of 50–100 Ω to ensure that the RC time constant approaches the transit time limit.

4.3.4 Avalanche photodiodes

The p–i–n diode has no gain, and adding gain enhances its performance. Increasing the reverse bias close to or at the breakdown of the diode, the electric field in the 'i' or n^- region becomes high, and results in the carriers being accelerated to a higher velocity before a collision with the lattice occurs. If the velocity is high, then this collision may be inelastic and causes ionisation, resulting in an additional electron and hole being generated. This additional electron and hole together with the original electron are also accelerated in turn to have further collisions to create additional hole–electron pairs. A schematic diagram of this process is shown in Figures 4.33 and 4.34 in which an electron is injected, at the start of the high field region of the depletion layer. This figure assumes that the ionisation coefficient for the electrons α_e is equal to that for the holes α_h. These coefficients are probabilistic, and are the reciprocal of the average distance that the carrier travels before an ionising collision occurs. These coefficients are a function of the electric field, and vary with the material parameters. In Silicon,

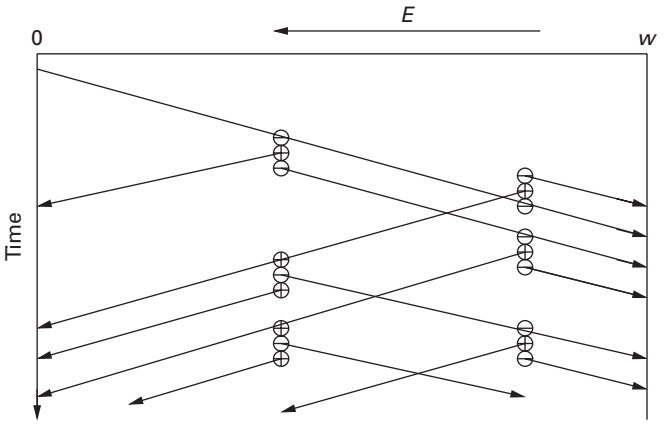

Fig. 4.33 Schematic diagram of the avalanche process with injection of an electron into the avalanche region, which after collision creates an additional hole–electron pair, which, together with the original electron, is also accelerated to have further ionising collisions, resulting in additional hole–electron pairs for each carrier. In this diagram, the ionisation coefficients of the electrons and holes are assumed to be nearly equal.

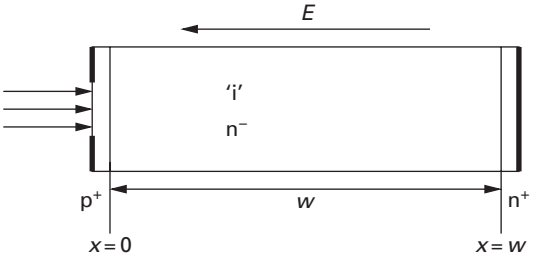

Fig. 4.34 Schematic diagram of the avalanche diode.

$\alpha_e > \alpha_h$; but in III–V material, GaAs and InP and others they are almost equal. The collisional increase in carriers results in an increase in the current which implies current gain.

The expression for the ionisation coefficient is given by [3]:

$$\alpha_{e,h} = \alpha_\infty \exp\left[-\left(\frac{b}{\mathcal{E}}\right)^m\right]. \tag{4.215}$$

In GaAs, α_∞ is $\sim 1.3 \times 10^6\,\mathrm{cm}^{-1}$, b is $2 \times 10^6\,\mathrm{V\,cm}^{-1}$ and m is 2. The multiplication coefficient or factor M is the ratio of the output current density to the input current density, and may be subdivided for the electron component and the hole component.

Suppose that the total current density in the diode is given by

$$J_{tot} = J_e(x) + J_h(x), \tag{4.216}$$

where $J_e(x)$ is the electron current density, $J_h(x)$ is the hole current density and J_{tot} is a constant at any plane of the diode. The multiplication factors are defined as

$$M_e = \frac{J_{e,out}}{J_{e,in}} = \frac{J_e(w)}{J_e(0)} \tag{4.217}$$

for the electron current density, and

$$M_h = \frac{J_{h,out}}{J_{h,in}} = \frac{J_h(0)}{J_h(w)} \tag{4.218}$$

for the hole current density.

Suppose a p–i–n diode with an 'i' region is considered, the field across the 'i' region is uniform. Also assume that the ionisation coefficients $\alpha_e = \alpha_h$, and are constant at the fixed electric field in the 'i' region, and the width of the avalanching region is w. The holes and electrons are accelerated in opposite directions, and the collisions result in ionisation and multiplication. Assume that at $x = 0$, there is only electron injection, hole injection is zero at $x = w$, and therefore the avalanching is initiated by electrons. In this case, the multiplication factor is given by

$$M = 1 + \alpha_e w + (\alpha_e w)^2 + (\alpha_e w)^3 + (\alpha_e w)^4 + \cdots \tag{4.219}$$

$$= \frac{1}{1 - \alpha_e w} \text{ for } \alpha_e w < 1. \tag{4.220}$$

This final result implies that the multiplication factor goes to ∞ when $\alpha_e w$ is unity, or that each carrier has one ionising collision over the distance w.

For the case of the p–i–n, with an n^- region instead of the 'i' region, the field increases linearly from the n^+ region to the p^+ region, and the above assumptions of constant ionisation coefficients no longer hold. The total current density flowing in the diode is the sum of the electron current density and the hole current density.

The current densities in the avalanching region satisfy the following equations:

$$\frac{dJ_e}{dx} = \alpha_e J_e + \alpha_h J_h + qGen(x) \tag{4.221}$$

$$-\frac{dJ_h}{dx} = \alpha_e J_e + \alpha_h J_h + qGen(x), \tag{4.222}$$

where $qGen(x)$ is the optical generation rate. Subtracting these two equations results in

$$\frac{d}{dx}[J_e(x) + J_h(x)] = 0, \tag{4.223}$$

or

$$J_e(x) + J_h(x) = \text{constant} = J_{tot}. \tag{4.224}$$

This implies that the current density is a constant at any plane in the diode, as claimed earlier in Equation (4.216). Substituting for J_h from Equation (4.216) by $J_{tot} - J_e$ in Equation (4.221)

$$\frac{dJ_e}{dx} - (\alpha_e - \alpha_h) J_e = \alpha_h J_{tot} + qGen(x). \tag{4.225}$$

Consider the case when $\alpha_e = \alpha_h$, neglecting the *Gen* term, this becomes

$$\frac{dJ_e}{dx} = \alpha_e J_{tot}, \tag{4.226}$$

and the solution is

$$J_e(w) = J_{tot} \int_0^w \alpha_e dx + J_e(0). \tag{4.227}$$

Assuming that there is no hole injection at $x = w$, then $J_e(w) = J_{tot}$, and this equation becomes

$$J_e(w) = J_e(w) \int_0^w \alpha_e dx + J_e(0). \tag{4.228}$$

Dividing by $J_e(0)$, this equation becomes

$$\frac{J_e(w)}{J_e(0)} = M_e = M_e \int_0^w \alpha_e dx + 1, \tag{4.229}$$

or

$$M_e = \frac{1}{1 - \int_0^w \alpha_e dx}. \tag{4.230}$$

For avalanching,

$$\int_0^w \alpha_e dx = 1. \tag{4.231}$$

When $\alpha_e \neq \alpha_h$, then Equation (4.225) needs to be solved. This is first-order differential equation of the form:

$$\frac{dy}{dx} + P(x)y = Q(x), \tag{4.232}$$

and the solution is given by

$$y = \frac{\int_0^x Q(x') e^{\int_0^{x'} P(x'')dx''} + y(0)}{e^{\int_0^x P(x')dx'}}. \tag{4.233}$$

Consider the case of electron injection at $x = 0$, with $Gen(x) = 0$ and no hole injection at $x = w$, which implies that $J_h(w) = 0$, and $J_{tot} = Je(w)$. Following the above procedure, with these boundary conditions, the solution of Equation (4.225) is obtained, and the electron current multiplication factor becomes [8]:

$$M_e = \frac{J_e(w)}{J_e(0)} = \frac{e^{\int_0^w (\alpha_e - \alpha_h)dx'}}{1 - \int_0^w dx' \alpha_h(x') e^{\int_{x'}^w (\alpha_e - \alpha_h)dx''}}. \tag{4.234}$$

When $\alpha_e = \alpha_h$, then this equation becomes, as shown in Equation (4.230)

$$M_e = \frac{1}{1 - \int_0^w \alpha_e dx}. \tag{4.235}$$

When the denominator of these multiplication factors is zero, the diode avalanches. A similar expression to that in Equation (4.234) may be derived for M_h, if the avalanche

is initiated by holes. For this case, it is assumed that at $x = w$, $J_e(w) = 0$, $Gen(x) = 0$ and $J_h(0) = J_{tot}$, and the variable is $J_h(x)$ to give

$$M_h = \frac{J_h(w)}{J_e(0)} = \frac{e^{-\int_0^w (\alpha_e - \alpha_h) dx'}}{1 - \int_0^w dx' \alpha_e(x') e^{-\int_0^{x'} (\alpha_e - \alpha_h) dx''}}, \tag{4.236}$$

and again the avalanching occurs when the denominator goes to zero. The avalanching is dependent on the field in the depletion layer, carrier population and the collision frequency, and not dependent on the carrier initiating the avalanche process. By careful control of the bias, the avalanche gain may be controlled to values as desired.

Notice that the generation term has been omitted in these expressions, but may be included as necessary. Silicon avalanche photodiodes (APDs) may use the absorption region as the avalanche region. In this case, the derivations above with the generation term would be the appropriate equations. Near infrared wavelengths have used GaAs/AlGaAs APDs, and for wavelengths in the telecommunications band, 1300 nm and 1550 nm bands, the absorption layer is generally InGaAs which is lattice-matched to InP. The avalanche region for these diodes is separate and usually InP, since large leakage currents occur due to tunneling with high reverse bias in InGaAs. This type of diode is termed the *separate absorption and multiplication* (SAM) APDs [26], and is currently the usual APDs at these wavelengths, as shown schematically in Figure 4.35. Accumulation of holes occurs in the InP–InGaAs valence band junction region, and a graded junction alleviates this problem [22].

The multiplication factor is a random variable and therefore the excess noise due to this avalanche process may be estimated by a noise factor F_M. This factor is a function of the multiplication factor M and k_A, where k_A is the ratio of the ionisation coefficients and k_A lies in the range $0 < k_A < 1$. Thus, $k_A = \alpha_h/\alpha_e$, provided $\alpha_h < \alpha_e$, or alternatively $k_A = \alpha_e/\alpha_h$, provided $\alpha_e < \alpha_h$. The expression for the noise factor is

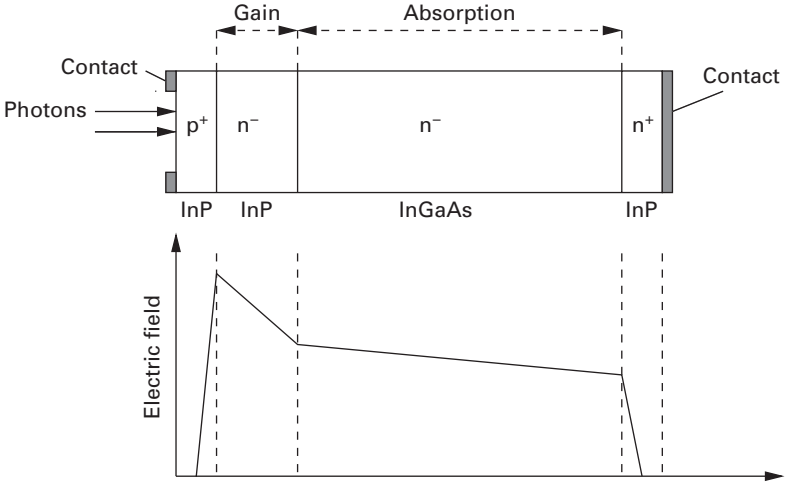

Fig. 4.35 Schematic diagram of the separate absorption and multiplication avalanche photodiode, also showing the field in the different regions.

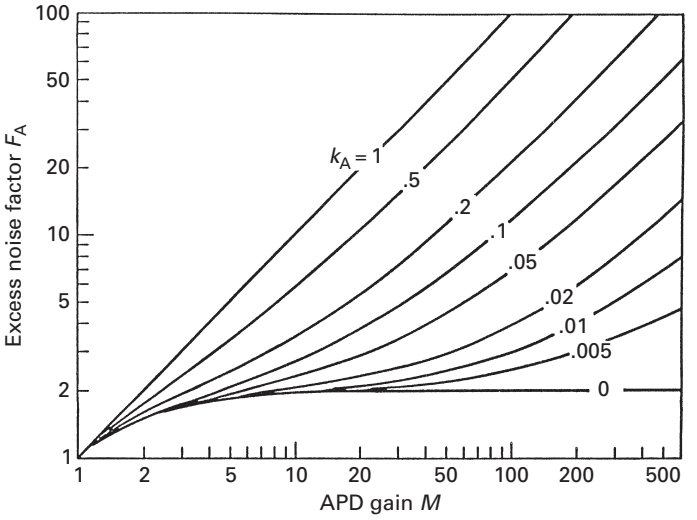

Fig. 4.36 Excess noise factor F_A as a function of the multiplication factor M and the ratio of the ionisation coefficients k_A (G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd edition, John Wiley & Sons, 2002. ©2002 John Wiley & Sons). Reprinted with permission of John Wiley & Sons, Inc.

given by [23]:

$$F_A = k_A M + (1 - k_A)\left(2 - \frac{1}{M}\right). \tag{4.237}$$

Figure 4.36 shows the excess noise factor F_A plotted against the multiplication factor M, as a function of k_A, from the Equation (4.237) [1].

The responsivity of the avalanche photodiode includes the multiplication factor M and is given by

$$R_{APD} = M\frac{\eta q}{h\nu}, \tag{4.238}$$

and therefore the input signal photocurrent is given by

$$I_{ph,APD} = R_{APD} P_{opt} = M\frac{\eta q}{h\nu} P_{opt}. \tag{4.239}$$

The shot noise current squared of the avalanche diode is also enhanced by the multiplication factor and the noise factor and is given by

$$\left\langle |\sigma_s|^2 \right\rangle = M^2 F_A (R_{APD} P_{in} + I_D + I_B)\Delta f, \tag{4.240}$$

where I_D is the dark current and I_B the background current. The thermal noise current squared is given by, as in Equation (4.212), including the amplifier noise figure F_n,

$$\left\langle |\sigma_t|^2 \right\rangle = \frac{4k_b T F_n \Delta f}{R_{eq}}. \tag{4.241}$$

Hence the signal-to-noise ratio is given by

$$\frac{S}{N} = \frac{\left(\frac{Mq\eta_{\text{int}}P_{\text{in}}}{h\nu}\right)^2}{2qM^2F_{\text{A}}(I_{\text{ph}} + I_{\text{B}} + I_{\text{d}})\Delta f + \frac{4k_{\text{b}}TF_{\text{n}}\Delta f}{R_{\text{eq}}}}. \tag{4.242}$$

In this case, the shot noise term dominates and therefore the signal-to-noise ratio becomes

$$\frac{S}{N} = \frac{\left(\frac{Mq\eta_{\text{int}}P_{\text{in}}}{h\nu}\right)^2}{2qM^2F_{\text{A}}(I_{\text{ph}} + I_{\text{B}} + I_{\text{d}})\Delta f}. \tag{4.243}$$

The APD is the preferred photodiode, but the introduction of the gain stage reduces the frequency response. However, recent results have shown that these devices have gain bandwidth products of over 300 GHz, with a response of 28 GHz [18]. Lower gain results show a higher response of over 30 GHz [29]. These devices are much more expensive than p–i–n diodes, although the performance is better because of the gain that arises due to avalanching.

4.3.5 Metal–semiconductor–metal detectors

A planar version of the p–i–n diode is the metal–semiconductor–metal (MSM) detector, in which the contacts are metal Schottky barrier diodes to a thin undoped semiconductor layer, and the region between the metal contacts, usually in the form of an interdigitated structure, shown in Figure 4.37, is completely depleted. The gap between the fingers is made small, and the transit time limitation is small. The capacitance has two components: one in parallel with the gap across the fingers, and the second is the capacitance to the ground of both electrodes. However, this latter may be made small so that the RC time constant of the diode is small. Consequently, these detectors are extremely fast,

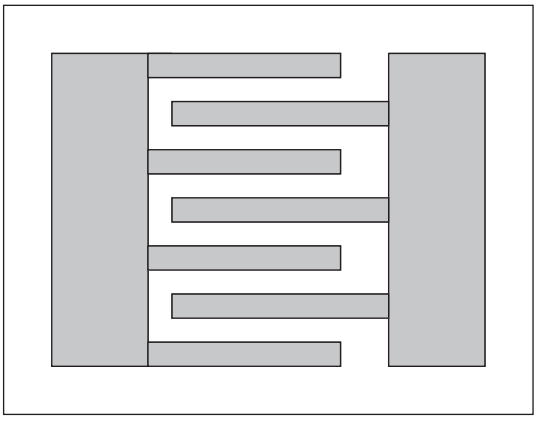

Fig. 4.37 Interdigitated form of the MSM detector, in which the contacts are Schottky diodes to the thin epitaxial absorption layer.

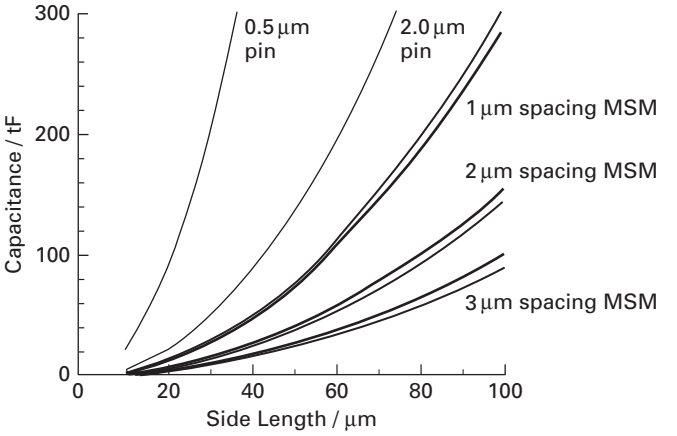

Fig. 4.38 Capacitance against the edge length of the fingers of the interdigitated MSM device for gaps of 1, 2 and 3 μm spacing, with finger widths of 0.5 and 1.0 μm. The parallel plate capacitance of the p–i–n diode is plotted for comparison, with the 'i' or n^- layer thickness of 0.5 and 1 μm. The electrode width for the p–i–n diode has not been given in the paper (J. B. D. Soole and H. Schumacher, *IEEE Journal of Quantum Electronics*, Vol. 27, No. 3, March 1991. ©1991 IEEE).

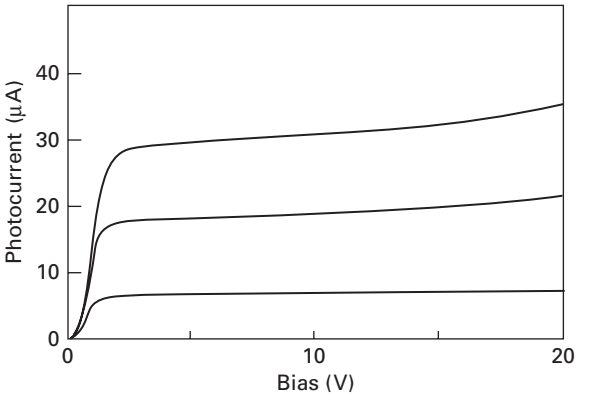

Fig. 4.39 Response of the MSM detector showing the detected photocurrent against the bias voltage for different values of light intensity for the 1 μm electrode width and a 2 μm interelectrode gap device (J. B. D. Soole and H. Schumacher, *IEEE Journal of Quantum Electronics*, Vol. 27, No. 3, March 1991. ©1991 IEEE).

but their responsivity is affected by the shadowing effect of the fingers. However, it was found that the quantum efficiency increases with increasing bias, which suggests that there is some gain that arises due to trap densities at the interfaces between layers, at electrode interfaces. However, [30] suggests that with careful growth this gain may be much reduced as seen in Figure 4.39. Note that one of the diodes is forward-biased and the other is reverse-biased, and again careful processing results in symmetric response and soft breakdown.

MSM photodetectors of InGaAs on an InP substrate [30] for operation at 1.3 μm and 1.5 μm wavelengths are discussed here. The Schottky barrier on InGaAs is of the

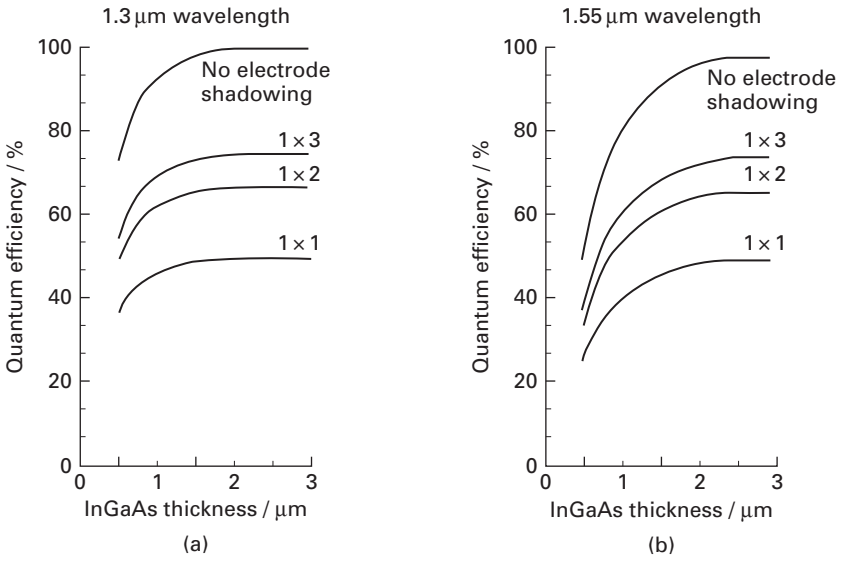

Fig. 4.40 Quantum efficiency of the InGaAs MSM device the electrode widths of 1 μ and different spacings from 1, 2, 3 μm inter-electrode spacing. The zero width electrode result is also plotted here (J. B. D. Soole and H. Schumacher, *IEEE Journal of Quantum Electronics*, Vol. 27, No. 3, March 1991. ©1991 IEEE).

order of 0.2 V, and is leaky in reverse bias, and therefore is enhanced by a layer of lattice-matched InAlAs under the electrodes. The thickness of the InGaAs layer was 1.3 μm, and the InAlAs layer was 80 nm. The capacitance between the fingers has been calculated by means of the usual Schwarz–Christofel transformation in this paper, and it is shown that the edge capacitance of the fingers is much less than the capacitance of comparable area mesa-type p–i–n diodes, as shown in Figure 4.38. Provided the series resistance is comparable, the RC time constant is much less for the MSM detector.

The device with a 1 μm wide electrode with the interelectrode gap of 2 μm between them has a response shown in Figure 4.39 of the detected photocurrent against the bias voltage across the electrodes for different values of light intensity on the device. Note that the response is not flat but rises, indicating that there is gain in the device; this may be in part due to the photoconductor effect and also may be at the higher bias due to avalanching near the electrodes. The quantum efficiency of the MSM detectors is hampered by the shadowing effect of the electrodes and this is plotted for this InGaAs device in Figure 4.40, which shows that this may rise to as high as about 70% for the 1 μm electrode width and a 2 μm interelectrode gap device. The bandwidth of the device is in the 20 GHz region, and other devices on GaAs show similar or better frequency response.

4.3.6 Travelling wave p–i–n photodiodes

As discussed above, the frequency response limitation of p–i–n diodes arises from the transit time of photogenerated carriers reaching their respective electrode contact

regions, and also from the RC time constant of the device. Making the absorption layer small, 0.3 μm, results in poor responsivity; placing mirrors underneath to obtain a double pass improves the response, but the RC time constant rises. The MSM device reduces the RC time constant but the transit time limitation remains. The second limitation is the device saturation that occurs when the light intensity becomes high. The alternative is the waveguide photodetector (WPD), which is an edge-fed p–i–n diode that is made very narrow and long, and the absorption layer may be thin to overcome the transit time limitation, making it highly efficient. However, the RC time constant limits the device response to about 55 GHz bandwidth–efficiency value (product of bandwidth and quantum efficiency) [16].

To overcome the problem of the RC limitation of the waveguide detector, and also be able to handle high power detection, the travelling wave detector was proposed by [8, 32]. An optical waveguide with a thin absorption layer is the basis for this device. The absorption layer is made thin enough to only absorb a small fraction of the light per unit length so as not to saturate the device, thus high power signals may be detected without saturation. However, the limitation of the travelling wave detector arises from the velocity mismatch between the electrical wave on the electrode structure and the optical waveguide. The electrode on top of the guide with the accompanying ground electrodes on the sides form a coplanar waveguide, in which the detected signal travels in the form of a voltage/current wave, see Figure 4.41. The optical signal travels at the guide layer group velocity, usually at c/n_g, and n_g is the group index of the guide layer. The detected signal forms electrical forward and backward waves on the electrode structure. The backward wave reflects at the input of the electrode structure, if it is an open circuit at the start of the detector, or alternatively, the wave is absorbed in the load if matched. Thus, the electrical wave travelling on the electrodes is a combination of the forward wave and the reflected component of the backward wave travelling in the forward direction, and also the backward wave traveling in the reverse direction. The lack of velocity match between these two optical and electrical waves leads to walk off, and the frequency response drops [15]. However, bandwidth of 172 GHz and bandwidth–efficiency product of 76 GHz [12] and pulse response with transform bandwidth as high as 560 GHz [28] have been reported. Careful choice of the absorbing layer, absorbing coefficient of α_0, and its thickness and location in the waveguide

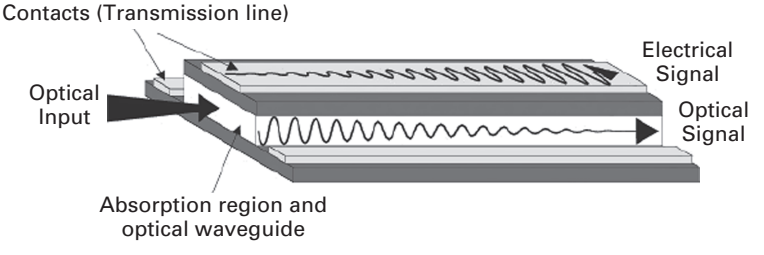

Fig. 4.41 Schematic diagram of a travelling wave photodiode, with coplanar electrodes (G. Rengel-Sharp, R. E. Miles, S. Iezekiel, *The Radio Science Bulletin*, Vol. 311, No. 12, pp. 55–64, December 2004. ©2004 URSI).

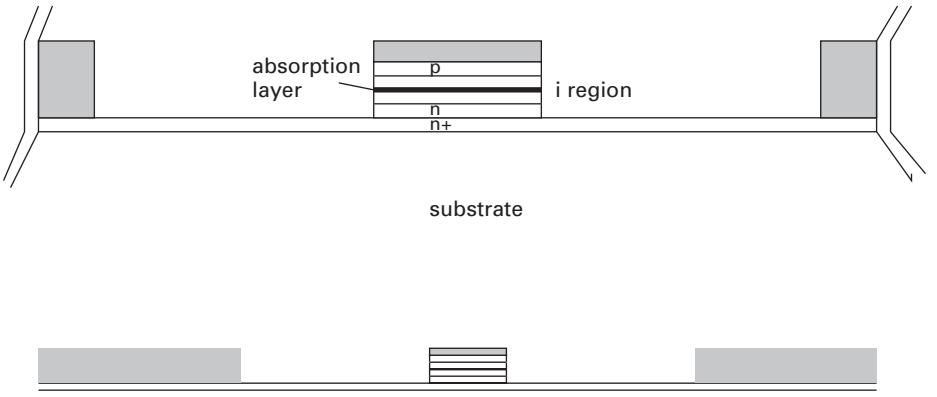

Fig. 4.42 Schematic diagram of a travelling wave photodiode, showing the p–i–n structure with the absorbing layer, and the electrodes.

allow the confinement factor Γ to be determined, and thus the decay factor becomes $e^{-\alpha_0 \Gamma z}$ for the optical wave intensity. These types of travelling wave devices are the fully distributed detectors. A second design has used a passive guide between MSM detectors [20], in which each detector only detects a small fraction of the power carried by the optical wave. Velocity matching allows the 3dB frequency response to rise to 49 GHz. Both these types of travelling wave devices are the fastest photodetectors built at the present time. The usual optical waveguides are of the ridge type, with the top electrode and adjacent ground electrodes of the coplanar waveguide (CPW) type, as shown in Figure 4.41 [27].

The theory of the distributed detector has been developed by [12] and [15]. Essentially, the p–i–n diode has to support the optical mode, and the electrode structure supports the electrical mode. Figure 4.42 shows a typical structure of this travelling wave photodetector.

The optical mode travels in the central p–i–n region, and the electrical mode is supported by the electrode structure transmission line shown in this figure. The equivalent circuit of the elemental section of the line is now the usual R–L in series and C–G in parallel, but now with the elemental p–i–n structure in parallel; the modified circuit is shown in Figure 4.43. Note that the usual capacitance is that of the depletion layer, and therefore is across the current source.

In this circuit, R_{cpw} is the series resistance of the central electrode and varies due to the skin effect, increasing as $\sqrt{f}$. The semiconductor series resistance R_{semi} is in series with the depletion layer, and is determined by the doping levels of the p^+ and n^+ layers, shown as p and n in Figure 4.42. The capacitance of the central electrode is dominated by the depletion layer capacitance of the p–i–n diode, which is much larger than the edge capacitance. The depletion layer capacitance gives rise to a slow wave effect, similar to that suggested by Hasegawa [14]. The inductance is the usual CPW value, and is unaffected by the presence of the p–i–n structure. The series resistance G_s is the top contact layer in parallel with the electrodes. The distributed current source is

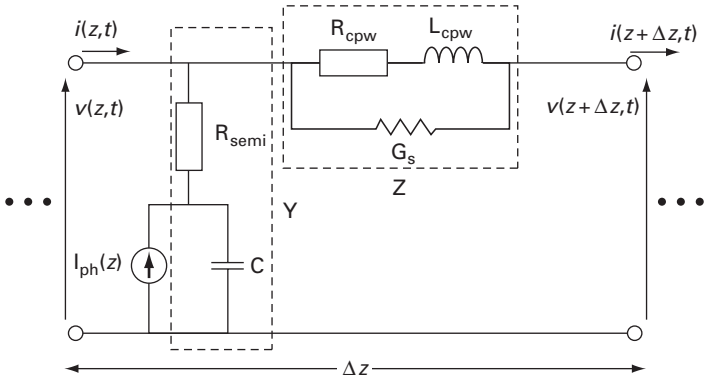

Fig. 4.43 Schematic diagram of the equivalent circuit of a travelling wave photodiode (G. Rengel-Sharp, R. E. Miles and S. Iezekiel, *The Radio Science Bulletin*, Vol. 311, No. 12, pp. 55–64, December 2004. ©2004 URSI).

related to the input optical power P_o, and is given by

$$I_{ph}(z) = P_o \frac{\eta q \lambda}{hc} \Gamma \alpha_o e^{-\Gamma \alpha_o} e^{-j\beta_o z}, \tag{4.244}$$

where λ is the wavelength of the optical power, α_o is the absorption layer coefficient, Γ is the confinement factor in this absorption layer, β_o is the optical propagation constant and η is the quantum efficiency of the distributed p–i–n diode. The terminations at the input may be an open circuit or a matched load, and at the output is matched to extract the signals on the electrical wave. This equivalent circuit may be used to analyse the performance of the travelling wave detector.

The major concern in the treatment of the travelling wave detector is the velocity mismatch between the optical wave and the electrical wave on the electrode structure. Giboney *et al.* [12] have analysed the transmission line using an impulse excitation, to account for the velocity mismatch. Performing the Fourier transform, the following expression for the normalised photocurrent is obtained for the case when $\Gamma \alpha_o \ell \ll 1$ for the frequency ω:

$$\iota_v(\omega) = \frac{1}{2}\left[\frac{\omega_f}{\omega_f - j\omega} + \gamma(\omega)\frac{\omega_r}{\omega_r + j\omega}\right] e^{-j\omega(\ell/v_e)}, \tag{4.245}$$

where

$$\omega_f = \frac{\Gamma \alpha_o v_e}{\left(\frac{1 - v_e}{v_o}\right)}$$

and

$$\omega_r = \frac{\Gamma \alpha_o v_e}{\left(\frac{1 + v_e}{v_o}\right)}.$$

γ is the electrical reflection coefficient at the device input, $\gamma = 1$ for an open circuit and $\gamma = 0$ for matched load. The term v_e is the electrical wave velocity and v_o is the

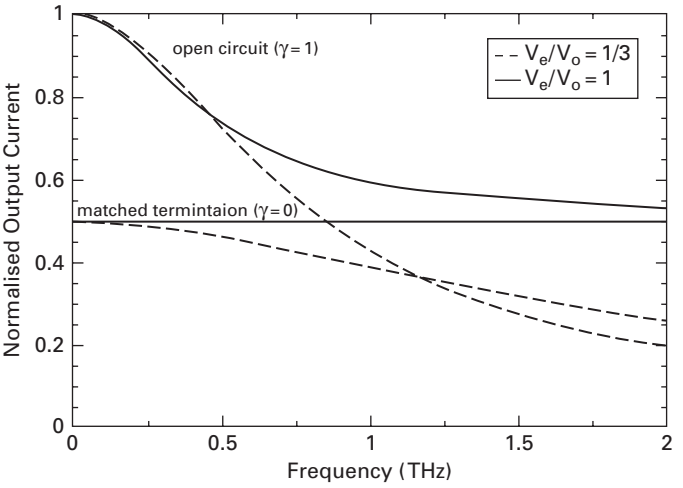

Fig. 4.44 Normalised current response for both the velocity-matched and -mismatched cases for both the open circuit and the matched terminations at the input (G. Rengel-Sharp, R. E. Miles and S. Iezekiel, *The Radio Science Bulletin*, Vol. 311, No. 12, pp. 55–64, December 2004. ©2004 URSI).

optical wave velocity. The results of this calculation are shown in Figure 4.44. Notice that the normalised current for the matched case has no frequency variation, but since half of it goes to the input-matched load, the output current is reduced by half. For the velocity-mismatched case, for $\gamma = 0$, again the frequency variation is lower than for the case when the input has an open circuit, when $\gamma = 1$. The open circuit case has considerable frequency variation, but the current is higher at all frequencies for the velocity-matched case, and becomes asymptotic to the $\gamma = 0$ case. For the velocity-mismatched case, the current falls off faster at the higher frequencies, and becomes asymptotic to the velocity-mismatched $\gamma = 0$ case.

For the velocity mismatched case with matched input termination, $\gamma = 0$, when $\ell\Gamma\alpha_o \gg 0$, the 3 dB bandwidth is given by [27]

$$f_{3dB} = \frac{\Gamma\alpha_o}{2\pi}\frac{v_o v_e}{(v_o - v_e)}, \tag{4.246}$$

and for the case when open circuit input termination, $\gamma = 1$, the bandwidth is

$$f_{3dB} \approx \frac{\Gamma\alpha_o v_e}{3\pi}. \tag{4.247}$$

This last result holds for velocity mismatch in the range $0 \leq v_e/v_o \leq 1.47$ [13]. Additional losses occur due to the optical guide scattering loss and microwave transmission line loss, and carrier transit time effects, and these will change the results and hence the performance of the photodetector. Since the optical waveguide is in the heterostructure form, carrier trapping may be a problem, but graded interface junctions would alleviate this.

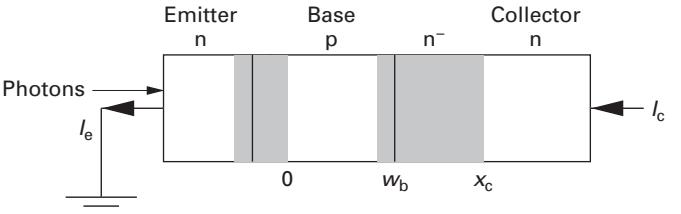

Fig. 4.45 Schematic diagram of a photo-HBT showing the depletion layers.

4.3.7 Heterostructure bipolar transistor photodetector

The bipolar transistor acts as excellent photodetector with gain. In the HBT, Figure 4.45, the absorption of light occurs solely in the base, and possibly in the subcollector region. In a double HBT in which the emitter and collector are of larger bandgap material, the absorption is confined to the base region, even though the base is not fully depleted. The usual modus operandi is to leave the base open circuit, as the light-generated carriers constitute the base current, though in some case to avoid carrier trapping contact is made to the base. The transit time of the electrons in the base of an n–p–n transistor determines the frequency limit of operation of the HBT, in addition to the base resistance effects. For a reasonable fraction of the incident to be absorbed, the base and the base–subcollector regions should be of the order of 1 μm thick, which makes these devices extremely slow, typically about 1 GHz bandwidth.

According to [3], the optical gain G of the single heterojunction phototransistor is the ratio of the number of carriers in the collector current due to the photogenerated–base and subcollector carriers to the number of photons incident on the base-subcollector:

$$G = \frac{h\nu I_{\text{ph-coll}}}{q P_{\text{inc}}}. \tag{4.248}$$

For a thick subcollector depletion width, which is greater than $1/\alpha$, then

$$G = \eta\beta_{\text{T}}, \tag{4.249}$$

where η is the quantum efficiency and β_{T} is optical current gain of the photo-transistor.

4.4 Problems

(1) A GaAs LED has a structure shown in Figure 4.1, and injection takes place from the n^+ layer into the p layer where the recombination takes place. Suppose the lifetime of the electrons in this layer is 10 ns, the electron mobility is 3000 cm^2 $(\text{V.s})^{-1}$ and hole mobility is 300 cm^2 $(\text{V.s})^{-1}$, determine the thickness of the p layer if it is to be one diffusion length.

(2) Calculate, numerically, the correct value of the external quantum efficiency of a LED radiating into air. Assume that the semiconductor index is 3.5, the polarisation random which implies that half the radiation is perpendicular polarisation and the other half is parallel polarisation. Assume that the transmission coefficient for the perpendicular polarisation from medium 1, index n_1, into medium 2, index n_2, is given by

$$\tau_{\perp} = \frac{2n_1 \cos\theta_1}{(n_1 \cos\theta_1 + n_2 \cos\theta_2)},$$

and the transmission coefficient for parallel polarisation is given by

$$\tau_{\parallel} = \frac{2n_1 \cos\theta_2}{(n_2 \cos\theta_1 + n_1 \cos\theta_2)}.$$

(3) A ridge laser is to operate at 800 nm, with a guide index of 3.41 and cladding index of 3.43. Determine the maximum thickness and width of the laser waveguide.

(4) An InGaAsP laser operating at 1300 nm has an effective guide index of 3.4. What is the reflectivity of its cleaved facets? Suppose the internal loss is $30\,\mathrm{cm}^{-1}$, what is the photon lifetime, assuming the group index is 3.5? Suppose the laser length is 300 μm, width 3 μm and thickness of 0.2 μm, the gain coefficient a is $2 \times 10^{-16}\,\mathrm{cm}^2$, the value of the transparency number density is $10^{18}\,\mathrm{cm}^{-3}$ and carrier lifetime is 1 ns. What is the threshold carrier density and the threshold current? What is the relaxation oscillation frequency at twice the threshold current?

(5) A p–i–n diode receiver at 1300 nm has a preamplifier noise figure of 2 dB. The signalling rate is 1 GHz, bandwidth of 100 MHz, the load resistance is 100 Ω, dark current is 1 nA, the diode quantum efficiency is 0.95 and a signal of 5 μW illuminates this diode. Determine the signal-to-noise ratio at the output.

(6) Suppose the p–i–n diode is replaced by an avalanche photodetector, multiplication factor of 10, noise factor of 3, dark current of 10 nA and background current of 1 nA. This APD is illuminated with 1 μW of optical power; determine the signal-to-noise ratio.

References

[1] Agrawal G. P. (2002). *Fiber-Optic Communication Systems*, 3rd edn. John Wiley & Sons.
[2] Agrawal G. P., Dutta N. K. (1993). *Semiconductor Lasers*. Van Norstrand Rheinhold.
[3] Bhattacharya P. (1997). *Semiconductor Optoelectronic Devices*, 2nd edn. Prentice Hall.
[4] Burrus C. A., Dawson R. W. (1970). Small-area high-current-density GaAs electroluminescent diodes and a method of operation for improved degradation characteristics. *Appl. Phys. Lett. 17*, 3, 97–99.
[5] Chang K. P., Yang S. L., Chuu D. S., Hsiao R. S., Chen J. F. (2005). Characterization of self-assembled InAs quantum dots with InAlAs/InGaAs strain-reduced layers by photoluminescence spectroscopy. *J. Appl. Phys. 97*, 83511–83514.

[6] Chang Y.-C., Wang C. S., Coldren L. A. (2007). High-efficiency, high-speed VCSELs with 35 Gbit/s error-free operation. *Electron. Lett. 43*, 19, 1022–1023.

[7] Cheng D. K. (1989). *Field and Wave Electromagnetics*, 2nd edn. Addison-Wesley.

[8] Chuang S. L. (1995). *Physics of Optoelectronic Devices*. John Wiley & Sons.

[9] Dutta N. K. (1980). Calculated absorption, emission, and gain in $In_{0.72}Ga_{0.28}As_{0.6}P_{0.4}$. *J. Appl. Phys. 52*, 6095–6100.

[10] Effendberger F. J., Joshi A. M. (1996). Ultrafast, dual-depletion region, InGaAs/InP p-i-n detector. *J. Lightw. Technol. 14*, 8 (August), 1859–1864.

[11] Fathpour K. P., Mia Z., Bhattacharya P. (2005). High speed quantum dot lasers. *J. Phys. D: Appl. Phys. 38*, 2103–2111.

[12] Giboney K. S., Nagarajan R. L., Reynolds T. E., *et al.* (1995). Traveling-wave photodetectors with 172-GHz bandwidth and 76-GHGz bandwidth-efficiency product. *IEEE Photonic Technol. Lett. 7,* 4 (April), 412–414.

[13] Giboney K. S., Rodwell M. J. W., Bowers J. E. (1997). Traveling-wave photodetector theory. *IEEE Trans. Microw. Theory Tech. 45*, 8 (August), 1310–1319.

[14] Hasegawa H., Furukawa M., Yani H. (1971). Properties of microstrip line on Si-SiO_2 system. *IEEE Trans. Microw. Theory Tech. 19*, 11 (November), 869–881.

[15] Hietala V., Vawter G. A., Brennam T. M., Hammons B. E. (1995). Traveling-wave photodetectors for high-power, large-bandwidth applications. *IEEE Trans. Microw. Theory Tech. 43*, 9 (September), 2291–2298.

[16] Kato K. (1999). Ultrawide band/high-frequency photodetectors. *IEEE Trans. Microw. Theory Tech. 47*, 7 (July), 1265–1281.

[17] Keiser G. (2000). *Optical Fiber Communications*, 2nd edn. McGraw-Hill.

[18] Kinsey G. S., Campbell J. C., Dentai A. G. (2001). Waveguide avalanche photodiode operating at 1.55 μm with a gain-bandwidth product of 320 GHz. *IEEE Photonic Technol. Lett. 13*, 8 (August), 842–844.

[19] Lee T. P., Burrus C. A., Copeland J. A., Dentai A. G., Marcuse D. (1982). Short cavity InGaAsP injection lasers: dependence of mode spectra and single-longitudinal-mode power on cavity length. *IEEE J. Quan. Electron. QE-18*, 7, 1101–1113.

[20] Lin L. Y., Wu M. C., Itoh T., *et al.* (1997). High-power high-speed photodetectors – design, analysis, and experimental demonstration. *IEEE Trans. Microw. Theory Tech. 45*, 8 (August), 1320–1331.

[21] Marcuse D., Lee T. P. (1983). On approximate analytical solutions of the rate equations for studying transient spectra of injection lasers. *IEEE J. Quan. Electron. QE-19*, 9, 1101–1113.

[22] Matsushima Y., Sasaki K., Noda Y. (1981). New type InGaAs/InP heterostructure avalanche photodiode with buffer layer. *IEEE Electron Device Lett. 2*, 7 (July), 179–181.

[23] McIntyre R. J. (1972). The distribution of gains in uniformly multiplying avalance photodiodes: theory. *IEEE Trans. Electron Devices 2*, 6 (June), 703–713.

[24] Nash F. R. (1973). Mode guidance of double-heterostructure GaAs lasers. *J. Appl. Phys. 44*, 10, 4696–4707.

[25] Nelson R. J., Wilson R. B., Wright P. D., Barnes P. A., Dutta N. K. (1981). CW electrooptical properties of InGaAsP (λ = 1.3 μm) buried-heterostructure lasers. *IEEE J. Quan. Electron. QE-17*, 2, 202–207.

[26] Nishida K., Taguchi K., Matsumo Y. (1979). InGaAsP heterostructure avalanche photodiodes with high avalanche gain. *Appl. Phys. Lett. 35*, 3 (August), 251–253.

[27] Rengel-Sharp G., Miles R. E., Iezekiel S. (2004). Traveling-wave photodetectors: a review. *The Radio Science Bulletin, URSI 311*, 12 (December), 54–64.

[28] Shi J.-W., Gan K.-G., Chiu Y.-J., *et al*. (2001). Metal-semiconductor-metal traveling-wave photodetectors. *IEEE Photonic Technol. Lett. 16*, 6 (June), 623–625.

[29] Shi J.-W., Liu Y.-H., Liu C.-W. (2004). Design and analysis of separate-absorption-transport-charge-multiplication traveling-wave avalanche photodetectors. *J. Lightw. Technol. 22*, 6 (June), 1583–1590.

[30] Soole J. B. D., Schumacher H. (1991). InGaAS metal-semiconductor-metal photodetector for long wavelength optical communications. *IEEE J. Quan. Electron. 27*, 3 (March), 737–752.

[31] Tauber D., Wang G., Geels R. S., Bowers J. E., Coldren L. A. (1993). Large and small signal dynamics of vertical cavity surface emitting lasers. *Appl. Phys. Lett. 62*, 4, 325–327.

[32] Taylor H. F., Ekonoyan O., Park C. S., Choi K. N., Chang K. (1990). Traveling-wave photodetectors. *SPIE Optoelectronics Signal Processing for Phased Array Antennas II*, 59–63.

[33] Tucker R. S., Kaminow I. P. (1984). High frequency characteristics of directly modulated InGaAsP ridge waveguide and buried heterostructure lasers. *J. Lightw. Technol. 2*, 4, 385–393.

[34] Yariv A. (1997). *Optical Electronics in Modern Communications*, 5th edn. Oxford University Press.

[35] Zhang X., Gutierrez-Aitkens A. L., Klotzkin D., Bhattacharya P., Caneau C., Bhatt R. (1996). 0.98 μm multiquantum well tunnelling injection lasers with ultra–high modulation bandwidths. *Electron. Lett. 32*, 18, 1715–1717.