optical fiber local area networks (LANs) [Ref. 68]. This connector plug will mate directly with connectorized optical LAN components (i.e. transmitters and receivers). A duplex fiber connector for use with the Fiber Distributed Data Interface (see Section 14.7.1) is now commercially available. It comprises two ST ferrules housed in a protective moulded shroud and exhibits a typical insertion loss of 0.6 dB. Hence, it is clear that such duplex connectors will be employed more extensively in the future.

Multiple fiber connection is obviously advantageous when interconnecting a large number of fibers. Both cylindrical and biconical ferrule connectors (see Sections 5.4.1 and 5.4.2) can be assembled in housings to form multiple fiber configurations

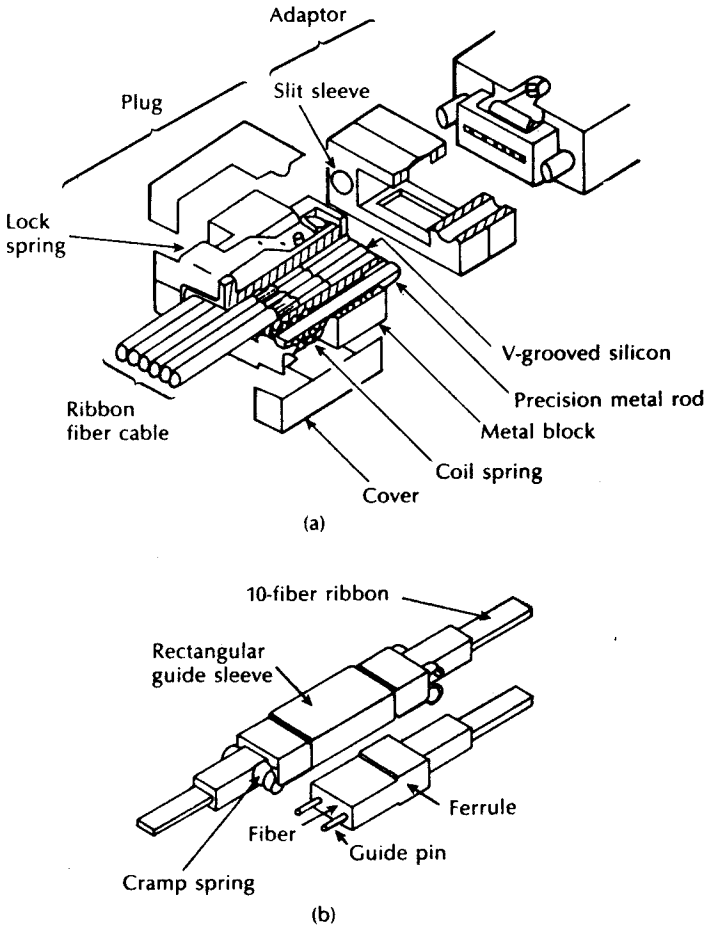


Figure 5.20 Multiple fiber connectors: (a) fiber ribbon connector using V-grooved silicon chips [Ref. 70]; (b) single-mode ten fiber connector [Ref. 73].

[Ref. 31]. Single ferrule connectors generally allow the alignment sleeve to float within the housing, thus removing any requirement for high tolerance on ferrule positioning within multiple ferrule versions. However, the force needed to insert multiple cylindrical ferrules can be large when many ferrules are involved. In this case multiple biconical ferrule connectors prove advantageous due to the low insertion force of the biconic configuration.

In addition to assembling a number of single fiber connectors to form a multiple fiber connector, other examples of multiple fiber connector exist in the literature. Silicon chip arrays have been suggested for the jointing of fiber ribbon cable for many years [Ref. 69]. However, difficulties were experienced in the design of an appropriate coupler for the two arrays. These problems have been largely overcome by the multiple connector design shown in Figure 5.20(a) which utilizes V-grooved silicon chips [Ref. 70]. In this connector, ribbon fibers are mounted and bonded into the V-grooves in order to form a plug together with precision metal guiding rods and coil springs. The fiber connections are then accomplished by butt jointing the two pairs of guiding rods in the slitted sleeves located in the adaptor, also illustrated in Figure 5.20(a). This multiple fiber connector has exhibited average insertion losses of 0.8 dB which were reduced to 0.4 dB by the use of index matching fluid. Improved loss characteristics were obtained with a more recent five fiber moulded connector, also used with fiber ribbons [Ref. 71]. In this case the mean loss and standard deviation without index matching were only 0.45 dB and 0.12 dB, respectively, when terminating 50 μm core multimode fibers [Ref. 72].

The structure of a small plastic moulded single-mode ten fiber connector is shown in Figure 5.20(b) [Ref. 73]. It comprises two moulded ferrules with ten fiber ribbon cables which are accurately aligned by guide pins, then held stable with a rectangular guide sleeve and a cramp spring. This compact multifiber connector which has dimensions of only 6×4 mm exhibited an average connection loss of 0.43 dB when used with single-mode fibers having a spot size (ω_0) of 5 μm.

5.5 Expanded beam connectors

An alternative to connection via direct butt joints between optical fibers is offered by the principle of the expanded beam. Fiber connection utilizing this principle is illustrated in Figure 5.21, which shows a connector consisting of two lenses for collimating and refocusing the light from one fiber into the other. The use of this interposed optics makes the achievement of lateral alignment much less critical than with a butt jointed fiber connector. Also, the longitudinal separation between the two mated halves of the connector ceases to be critical. However, this is achieved at the expense of more stringent angular alignment. Nevertheless, expanded beam connectors are useful for multifiber connection and edge connection for printed circuit boards where lateral and longitudinal alignment are frequently difficult to achieve.

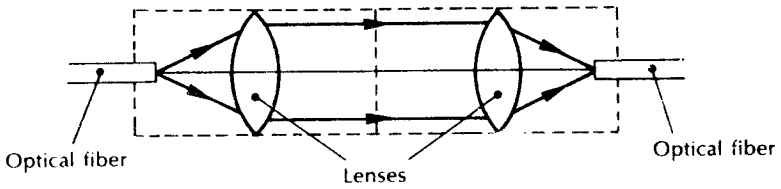


Figure 5.21 Schematic illustration of an expanded beam connector showing the principle of operation.

Two examples of lens coupled expanded beam connectors are illustrated in Figure 5.22. The connector shown in Figure 5.22(a) [Ref. 74] utilized spherical microlenses for beam expansion and reduction. It exhibited average losses of 1 dB which were reduced to 0.7 dB with the application of an antireflection coating on the lenses and the use of 50 μm core diameter graded index fiber.

A similar configuration has been used for single-mode fiber connection in which the lenses have a 2.5 mm diameter [Ref. 75]. Again with antireflection coated

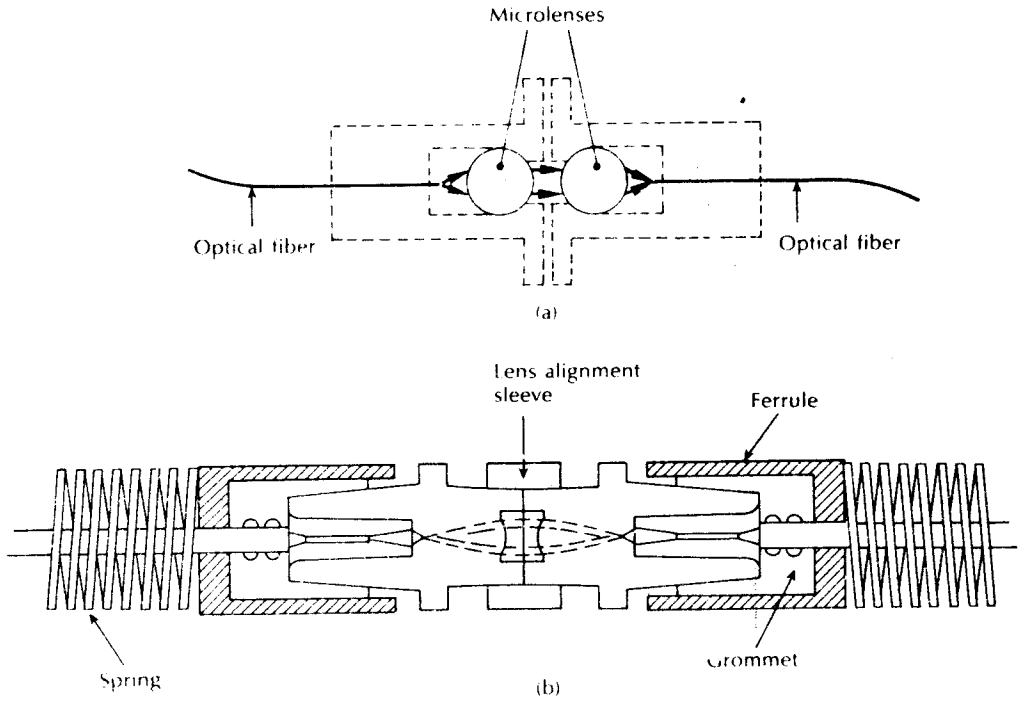


Figure 5.22 Lens coupled expanded beam connectors: (a) schematic diagram of a connector with two microlenses making a 1:1 image of the emitting fiber upon the receiving one [Ref. 74]; (b) moulded plastic lens connector assembly [Ref. 76].

lenses, average losses around 0.7 dB were obtained using 8 μm core diameter single-mode fibers. Furthermore, successful single-mode fiber connection has been achieved with a much smaller (250 μm diameter) sapphire ball lens expanded beam design [Ref. 31]. In this case losses in the range 0.4 to 0.7 dB were demonstrated over 1000 connections.

Figure 5.22(b) shows an expanded beam connector which employs a moulded spherical lens [Ref. 76]. The fiber is positioned approximately at the focal length of the lens in order to obtain a collimated beam and hence minimize lens to lens longitudinal misalignment effects. A lens alignment sleeve is used to minimize the effects of angular misalignment which, together with a ferrule, grommet, spring and external housing, provides the complete connector structure. The repeatability of this relatively straightforward lens design was found to be good, incurring losses of around 0.7 dB.

5.5.1 GRIN-rod lenses

An alternative lens geometry to facilitate efficient beam expansion and collimation within expanded beam connectors is that of the graded index (GRIN) rod lens [Refs. 77, 78]. In addition the focusing properties of such microlens devices have enabled them to find application within both fiber couplers (see Section 5.6) and source to fiber coupling (see Section 6.8).

The GRIN-rod lens which arose from developments on graded index fiber waveguides [Ref. 79] comprises a cylindrical glass rod typically 0.5 to 2 mm in diameter which exhibits a parabolic refractive index profile with a maximum at the axis similar to graded index fiber. Light propagation through the lens is determined by the lens dimensions and, because refractive index is a wavelength dependent parameter, by the wavelength of the light. The GRIN-rod lens can produce a collimated output beam with a divergent angle α of between 1° and 5° from a light source situated on, or near to, the opposite lens face, as illustrated in Figure 5.23. Conversely, it can focus an incoming light beam on to a small area located at the centre of the opposite lens face. Typically, light launched from a 50 μm diameter

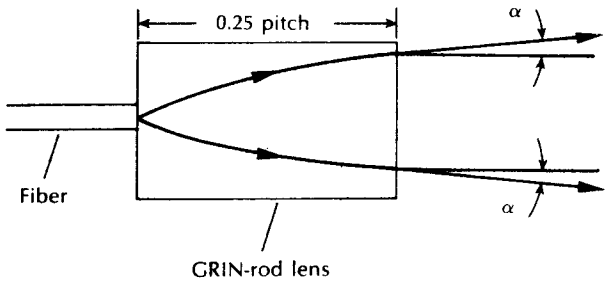


Figure 5.23 Formation of a collimated output beam from a GRIN-rod lens.

fiber core using a GRIN-rod lens results in a collimated output beam of between 0.5 and 1 mm.

Ray propagation through the GRIN-rod lens medium is approximately governed by the paraxial ray equation:

$$\frac{d^2r}{dz^2} = \frac{1}{n}\frac{dn}{dr} \tag{5.18}$$

where r is the radial coordinate, z is the distance along the optical axis and n is the refractive index at a point.

Furthermore, the refractive index at r following Eq. (2.75) distance r from the optical axis in a gradient index medium may be expressed as [Ref. 80]:

$$n(r) = n_1\left(1 - \frac{Ar^2}{2}\right) \tag{5.19}$$

where n_1 is the refractive index on the optical axis and A is a positive constant.

Using Eqs. (5.18) and (5.19), the position r of the ray is given by:

$$\frac{d^2r}{dz^2} = -Ar \tag{5.20}$$

Following Miller [Ref. 81], the general solution of Eq. (5.20) becomes:

$$r = K_1 \cos A^{\frac{1}{2}}r + K_2 \sin A^{\frac{1}{2}}r \tag{5.21}$$

where K_1 and K_2 are constants.

The refractive index variation with radius therefore causes all the input rays to follow a sinusoidal path through the lens medium. The traversion of one sinusoidal period is termed one full pitch and GRIN-rod lenses are manufactured with several pitch lengths. Three major pitch lengths are as follows:

1. The quarter pitch (0.25 pitch) lens, which produces a perfectly collimated output beam when the input light emanates from a point source on the opposite lens face. Conversely, the lens focuses an incoming light beam to a point at the centre of the opposite lens face (Figure 5.24(a)). Thus the focal point of the quarter pitch GRIN-rod lens is coincident with the lens faces, thus providing efficient direct butted connection to optical fiber.
2. The 0.23 pitch lens is designed such that its focal point lies outside the lens when a collimated beam is projected on the opposite lens face. It is often employed to convert the diverging beam from a fiber or laser diode into a collimated beam, as illustrated in Figure 5.24(b) [Ref. 82].
3. The 0.29 pitch lens is designed such that both focal points lie just outside the lens end faces. It is frequently used to convert a diverging beam from a laser diode into a converging beam. Hence, it proves useful for coupling the output from a laser diode into an optical fiber (Figure 5.24(c)), or alternatively for coupling the output from an optical fiber into a photodetector.

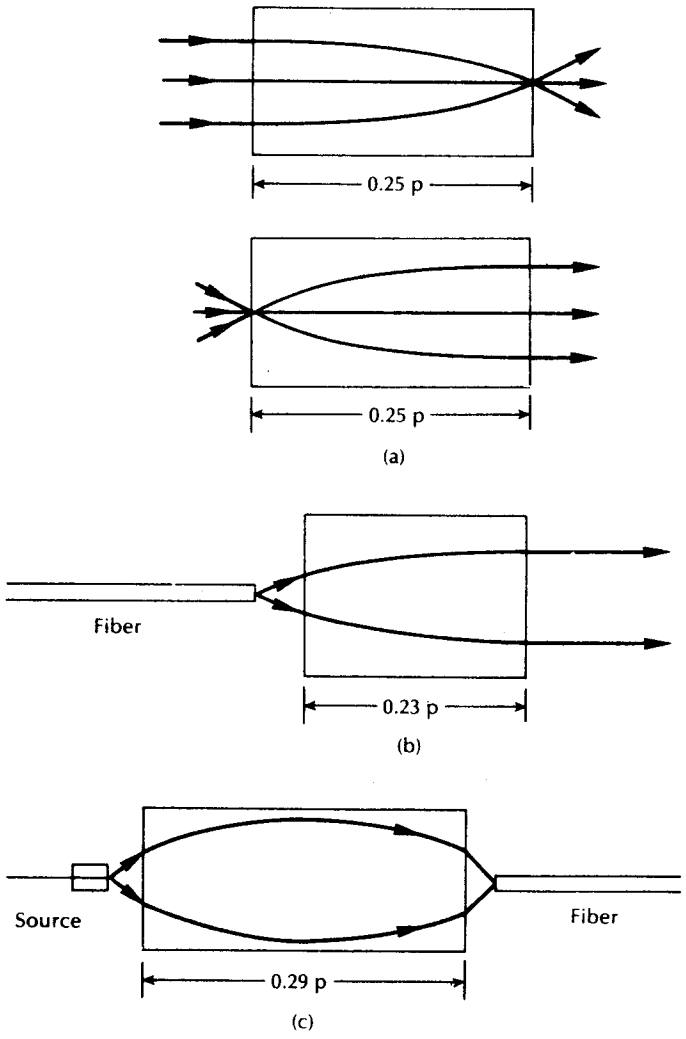


Figure 5.24 Operation of various GRIN-rod lenses: (a) the quarter pitch lens; (b) the 0.23 pitch lens; (c) the 0.29 pitch lens.

The majority of GRIN-rod lenses which have diameters in the range 0.5 and 2 mm may be employed with either single-mode or multimode (step or graded index) fiber. Various fractional pitch lenses, including those above as well as 0.5 p and 0.75 p, may be obtained from Nippon Sheet Glass Co. Ltd under the trade name SELFOC. They are available with numerical apertures of 0.37, 0.46 and 0.6.

A number of factors can cause divergence of the collimated beam from a GRIN-rod lens. These include errors in the lens cut length, the finite size of the fiber core

and chromatic aberration. As indicated previously, divergence angles as small as 1° may be obtained which yield expanded beam connector losses of around 1 dB [Ref. 31]. Furthermore, in contrast to butt jointed multimode fiber connectors, GRIN-rod lens connectors have demonstrated loss characteristics which are independent of the modal power distribution in the fiber [Ref. 83].

5.6 Fiber couplers

An optical fiber coupler is a device that distributes light from a main fiber into one or more branch fibers.* The latter case is more normal and such devices are known as multiport fiber couplers. More recently, interest has grown in these devices to divide or combine optical signals for application within optical fiber information distribution systems including data buses, local area networks, computer networks and telecommunication access networks (see Chapter 14).

Optical fiber couplers are often passive devices in which the power transfer takes place either:

(a) through the fiber core cross section by butt jointing the fibers or by using some form of imaging optics between the fibers (core interaction-type); or
(b) through the fiber surface and normal to its axis by converting the guided core modes to both cladding and refracted modes which then enable the power sharing mechanism (surface interaction type).

The mechanisms associated with these two broad categories are illustrated in Figure 5.25.

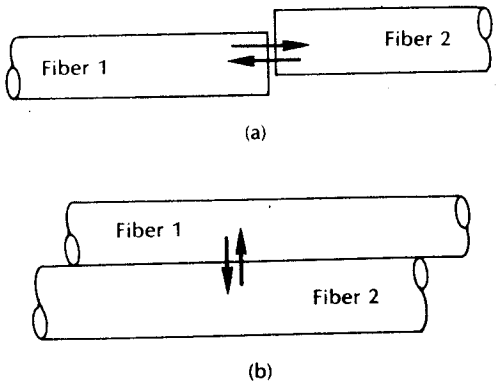


Figure 5.25 Classification of optical fiber couplers: (a) core interaction type; (b) surface interaction type.

* Devices of this type are also referred to as directional couplers.

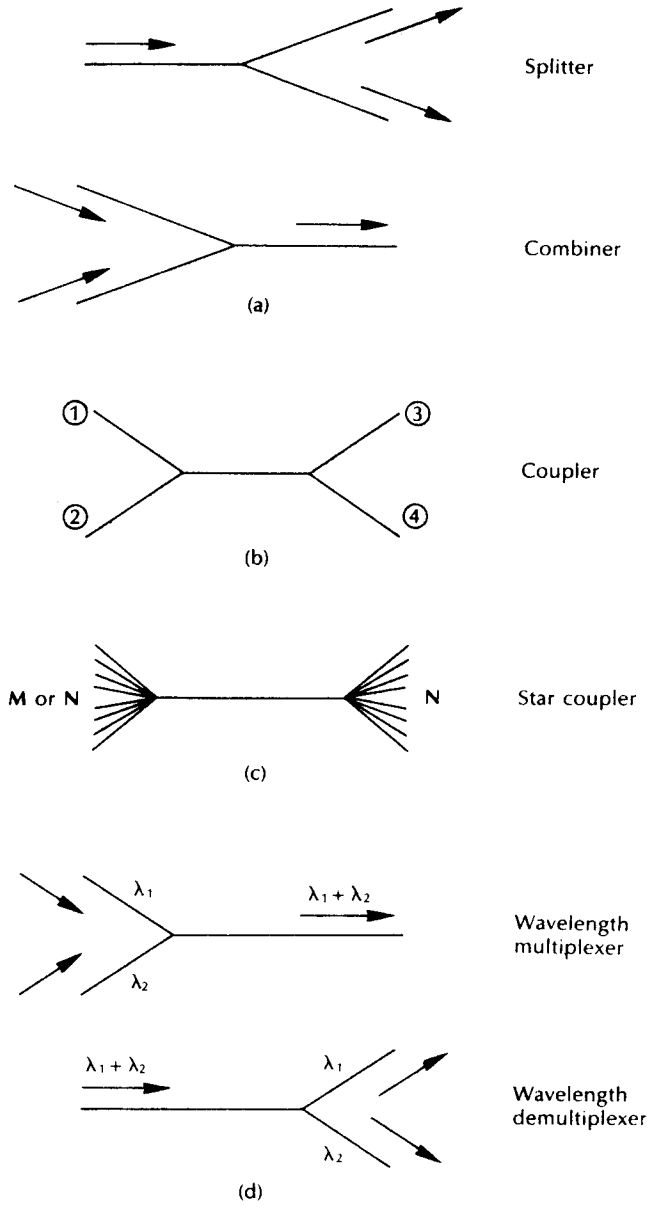


Figure 5.26 Optical fiber coupler types and functions: (a) three port couplers; (b) four port coupler; (c) star coupler; (d) wavelength division multiplexing and demultiplexing couplers.

Active waveguide directional couplers are also available which are realized using integrated optical fabrication techniques. Such device types, however, are dealt with in Section 10.6.1 and thus in this section the discussion is restricted to the above passive coupling strategies.

Multiport optical fiber couplers can also be subdivided into the following three main groups [Ref. 84], as illustrated in Figure 5.26:

1. Three and four port* couplers which are used for signal splitting, distribution and combining.
2. Star couplers which are generally used for distributing a single input signal to multiple outputs.
3. Wavelength division multiplexing (WDM) devices which are a specialized form of coupler designed to permit a number of different peak wavelength optical signals to be transmitted in parallel on a single fiber (see Section 11.9.3). In this context WDM couplers either combine the different wavelength optical signal on to the fiber (i.e. multiplex) or separate the different wavelength optical signals output from the fiber (i.e. demultiplex).

Ideal fiber couplers should distribute light among the branch fibers with no scattering loss† or the generation of noise, and they should function with complete insensitivity to factors including the distribution of light between the fiber modes, as well as the state of polarization of the light. Unfortunately, in practice passive fiber couplers do not display all of the above properties and hence the characteristics of the devices affect the performance of optical fiber networks. In particular, the finite scattering loss at the coupler limits the number of terminals that can be connected, or alternatively the span of the network, whereas the generation of noise and modal effects can cause problems in the specification of the network performance. Hence, couplers in a network cannot usually be treated as individual components with known parameters, a factor which necessitates certain compromises in their application. In this section, therefore, a selection of the more common fiber coupler types is described in relation to the coupling mechanisms, their performance and limitations.

5.6.1 Three and four port couplers

Several methods are employed to fabricate three and four port optical fiber couplers [Refs. 84 to 86]. The lateral offset method, illustrated in Figure 5.27(a) relies on the overlapping of the fiber end faces. Light from the input fiber is coupled to the output fibers according to the degree of overlap. Hence the input power can be distributed in a well defined proportion by appropriate control of the amount of lateral offset between the fibers. This technique, which can provide a bidirectional coupling capability, is well suited for use with multimode step index fibers but

* Four port couplers may also be referred to as 2×2 star couplers.
† The scattering loss through the coupler is often referred to as the excess loss.

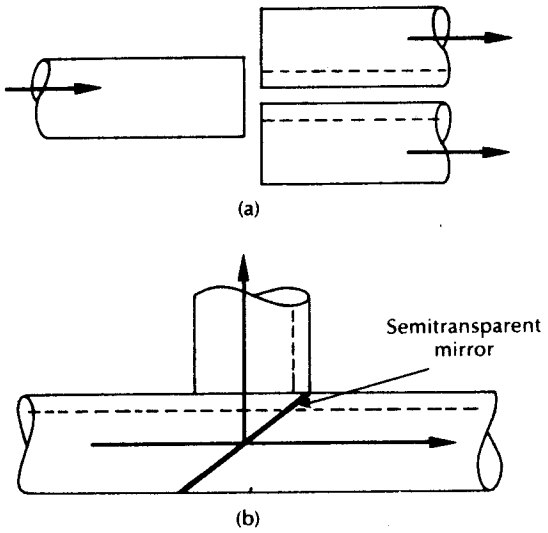


Figure 5.27 Fabrication techniques for three port fiber couplers: (a) the lateral offset method; (b) the semitransparent mirror method.

may incur higher excess losses than other methods as all the input light cannot be coupled into the output fibers.

Another coupling technique is to incorporate a beam splitter element between the fibers. The semitransparent mirror method provides an ingenious way to accomplish such a fiber coupler, as shown in Figure 5.27(b). A partially reflecting surface can be applied directly to the fiber end face cut at an angle of 45° to form a thin film beam splitter. The input power may be split in any desired ratio between the reflected and transmitted beams depending upon the properties of the intervening mirror, and typical excess losses for the device lie in the range 1 to 2 dB. Using this technology both three and four port couplers with both multimode and single-mode fibers have been fabricated [Ref. 86]. In addition, with suitable wavelength selective interference coatings this coupler type can form a wavelength division multiplexing device (see Section 5.6.3).

A fast growing category of optical fiber coupler is based on the use of microoptic components. In particular, a complete range of couplers has been developed which utilize the beam expansion and collimation properties of the graded index (GRIN) rod lens (see Section 5.5.1) combined with spherical retro-reflecting mirrors [Ref. 87]. These devices, two of which are displayed in Figure 5.28, are miniature optical assemblies of compact construction which generally exhibit low insertion loss (typically less than 1 dB) and are insensitive to modal power distribution.

Figure 5.28(a) shows the structure of a parallel surface type GRIN-rod lens three port coupler which comprises two quarter pitch lenses with a semitransparent mirror in between. Light rays from the input fiber F_1, collimate in the first lens

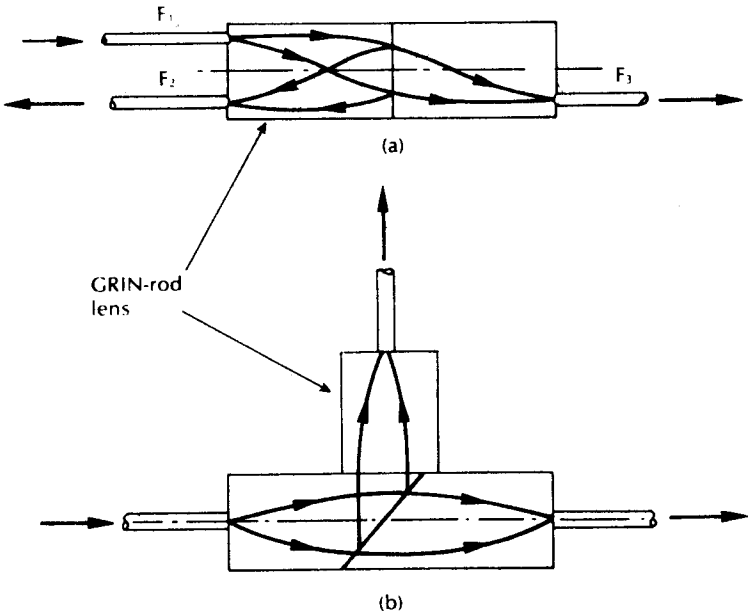


Figure 5.28 GRIN-rod lens based microoptic fiber couplers: (a) parallel surface type; (b) slant surface type.

before they are incident on the mirror. A portion of the incident beam is reflected back and is coupled to fiber F_2, whilst the transmitted light is focused in the second lens and then coupled to fiber F_3. The slant surface version of the similar coupler is shown in Figure 5.28(b). The parallel surface type, however, is the most attractive due to its ease of fabrication, compactness, simplicity and relatively low insertion loss. Finally, the substitution of the mirror by an interference filter* offers application of these devices to WDM (see Section 5.6.3).

Perhaps the most common method for manufacturing couplers is the fused biconical taper (FBT) technique, the basic structure and principle of operation of which is illustrated in Figure 5.29. In this method the fibers are generally twisted together and then spot fused under tension such that the fused section is elongated to form a biconical taper structure. A three port coupler is formed by removing one of the input fibers. Optical power launched into the input fiber propagates in the form of guided core modes. The higher order modes, however, leave the fiber core because of its reduced size in the tapered-down region and are therefore guided as cladding modes. These modes transfer back to guided core modes in the tapered-up region of the output fiber with an approximately even distribution between the two fibers.

* Such a dichroic device transmits only a certain wavelength band and reflects all other shorter or longer wavelengths.

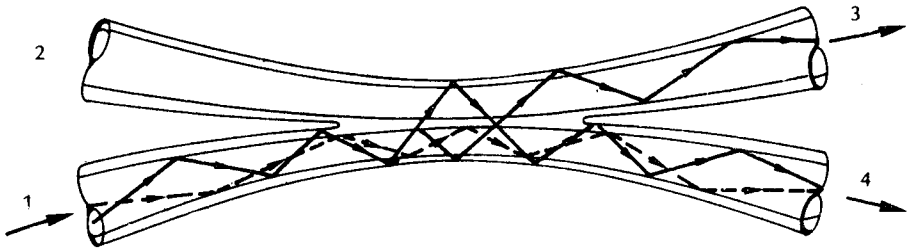


Figure 5.29 Structure and principle of operation for the fiber fused biconical taper coupler.

Often only a portion of the total power is coupled between the two fibers because only the higher order modes take part in the process, the lower order modes generally remaining within the main fiber. In this case a mode dependent (and therefore wavelength dependent) coupling ratio is obtained. However, when the waist of the taper is made sufficiently narrow, then the entire mode volume can be encouraged to participate in the coupling process and a larger proportion of input power can be shared between the output fibers. This strategy gives an improvement in both the power and modal uniformity of the coupler.

The various loss parameters associated with four port couplers may be written down with reference to Figure 5.29. Hence, the excess loss which is defined as the ratio of power input to power output is given by:

$$\textit{Excess loss} \text{ (four port coupler)} = 10 \log_{10} \frac{P_1}{(P_3 + P_4)} \text{ (dB)} \tag{5.22}$$

The insertion loss, however, is generally defined as the loss obtained for a particular port to port optical path.* Therefore, considering Figure 5.29:

$$\textit{Insertion loss} \text{ (ports 1 to 4)} = 10 \log_{10} \frac{P_1}{P_4} \text{ (dB)} \tag{5.23}$$

The crosstalk which provides a measure of the directional isolation† achieved by the device is the ratio of the backscattered power received at the second input port to the input power which may be written as:

$$\textit{Crosstalk} \text{ (four port coupler)} = 10 \log_{10} \frac{P_2}{P_1} \text{ (dB)} \tag{5.24}$$

* It should be noted that there is some confusion in the literature between coupler insertion loss and excess loss. Insertion loss is sometimes referred to when the value quoted is actually the excess loss. However, the author has not noticed the opposite where excess loss is used in place of insertion loss.

† The directional isolation and the crosstalk associated with a coupler are the same value in decibels but the former parameter is normally given as a positive value whereas the latter is negative value. Sometimes the directional isolation is referred to as the insertion loss between the two particular ports of the coupler which would be ports 1 to 2 in Figure 5.29.

Finally, the splitting or coupling ratio indicates the percentage division of optical power between the output ports. Again referring to Figure 5.29:

$$\textit{Split ratio} = \left[\frac{P_3}{(P_3 + P_4)}\right] \times 100\% \tag{5.25}$$

$$= \left[1 - \frac{P_4}{(P_3 + P_4)}\right] \times 100\% \tag{5.26}$$

Example 5.7

A four port multimode fiber FBT coupler has 60 μW optical power launched into port 1. The measured output powers at ports 2, 3 and 4 are 0.004, 26.0 and 27.5 μW respectively. Determine the excess loss, the insertion losses between the input and output ports, the crosstalk and the split ratio for the device.

Solution: The excess loss for the coupler may be obtained from Eq. (5.22) where:

$$\textit{Excess loss} = 10 \log_{10} \frac{P_1}{(P_3 + P_4)} = 10 \log_{10} \frac{60}{53.5}$$

$$= 0.5 \text{ dB}$$

The insertion loss is provided by Eq. (5.23) as

$$\textit{Insertion loss} \text{ (ports 1 to 3)} = 10 \log_{10} \frac{P_1}{P_3} = 10 \log_{10} \frac{60}{26}$$

$$= 3.63 \text{ dB}$$

$$\textit{Insertion loss} \text{ (ports 1 to 4)} = 10 \log_{10} \frac{60}{27.5} = 3.39 \text{ dB}$$

Crosstalk is given by Eq. (5.24) where:

$$\textit{Crosstalk} = 10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{0.004}{60}$$

$$= -41.8 \text{ dB}$$

Finally, the split ratio can be obtained from Eq. (5.25) as:

$$\textit{Split ratio} = \left[\frac{P_3}{P_3 + P_4} \times 100\right] = \frac{26}{53.5} \times 100$$

$$= 48.6\%$$

The split ratio for the FBT coupler is determined by the difference in the relative cross sections of the fibers, and the mode coupling mechanism is observed in both

multimode and single-mode fibers [Ref. 85]. An advantage of the FBT structure is its relatively low excess loss which is typically less than 0.5 dB,* with low crosstalk being usually better than −50 dB. A further advantage is the capability to fabricate FBT couplers with almost any fiber and geometry. Hence, they can be tailored to meet the specific requirements of a system or network. A major disadvantage, however, concerns the modal basis of the coupling action. The mode dependent splitting can result in differing losses through the coupler, a wavelength dependent performance, as well as the generation of modal noise when coherent light sources are employed [Ref. 88].

The precise spectral behaviour of FBT couplers is quite complex. It depends upon the dimensions and the geometry of the fused cross section, and on whether the fusing process produces a coupling region where the two cores are close (strongly fused) or relatively far apart (weakly fused) [Ref. 89]. It can also depend upon the refractive index of the surrounding medium [Ref. 90] and, in coherent systems, on the state of polarization of the optical field. Theoretical considerations [Ref. 89] show that for a single-mode FBT coupler, a minimum wavelength dependence on the splitting ratio is achieved for small cladding radii and strong fusing (i.e. the fiber cores placed close together). In order to obtain such performance it is necessary to taper the fibers down to a radius of around 15 μm or less, and to ensure that the rate of taper is such that the major proportion of the coupling occurs in the neck region. The wavelength dependent behaviour associated with single-mode FBT couplers follows an approximately sinusoidal pattern over the wavelength range 0.8 μm to 1.5 μm as a result of the single-mode coupling length between the two fibers [Ref. 91]. This mechanism has been used in the manufacture of WDM multiplexer/demultiplexer couplers (see Section 5.6.3).

More recently, single-mode fiber couplers have been fabricated from polarization maintaining fiber (so-called hi-birefringence couplers) which preserve the polarization of the input signals (see Section 3.13.2). Moreover, using polarization maintaining fiber, it is possible to fabricate polarization sensitive couplers, which effectively function as polarizing beamsplitters [Ref. 92].

An alternative technology to either fiber joint couplers, microoptic lensed devices or fused fiber couplers is the optical waveguide coupler. Corning have demonstrated [Ref. 93] the way in which such passive optical waveguide coupling components compatible with both multimode and single-mode fiber can be fabricated. Their production involves two basic processes. Firstly, a mask of the desired branching function is deposited on to a glass substrate using a photolithographic process. The substrate is then subjected to a two stage ion exchange [Ref. 94], which creates virtually circular waveguides embedded within the surface of the substrate on which the mask was deposited. An example of a three port integrated waveguide coupler fabricated using the above process is shown schematically in Figure 5.30. Multimode fibers are bonded to the structure using etched V-grooves. Excess

* Environmentally stable single-mode fused fiber couplers with excess losses less than 0.1 dB are commercially available.

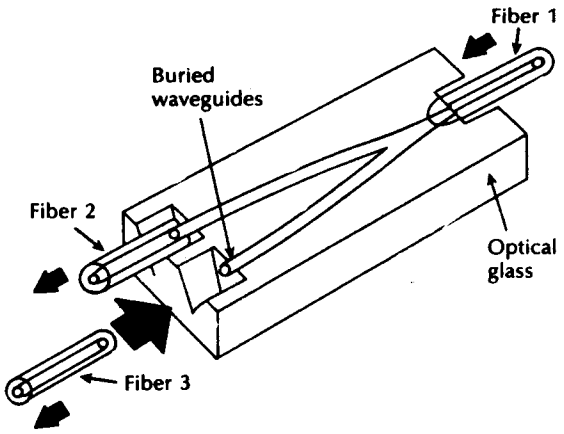


Figure 5.30 The Corning™ multimode fiber integrated waveguide three port coupler [Ref. 93].

losses were measured at 0.5 dB for the three port coupler and at 0.8 dB for the 1 × 8 star coupler [Ref. 93]. Clearly, this type of waveguide coupler is attractive because of the flexibility it allows at the masking stage. Furthermore, the same technique has been employed to fabricate WDM multiplexing and demultiplexing devices (see Section 5.6.3).

Finally, directional couplers have been produced which use the mode coupling that takes place between the guided and radiation modes when a periodic deformation is applied to the fiber. The principle of operation for this microbend* type coupler is illustrated in Figure 5.31 [Ref. 86]. Mode coupling between the guided and radiation modes may be obtained by pressing the fiber in close contact with a transparent mechanical grating. The radiated optical power can be collected by a lens or a shaped, curved glass plate. Interesting features of such devices are their variable coupling ratios which may be controlled over a wide range by altering the pressure on the fiber. In this context low light levels can be extracted from the fiber with very little excess loss (e.g. estimated at 0.05 dB [Ref. 95]).

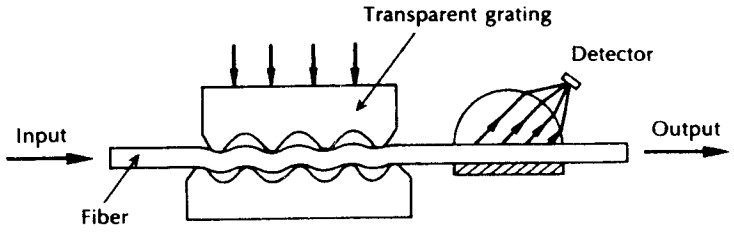


Figure 5.31 Schematic diagram of a microbend type coupler.

* This coupler operates in a similar manner to the microbend sensor described in Section 14.5.2.

5.6.2 Star couplers

Star couplers distribute an optical signal from a single input fiber to multiple output fibers, as may be observed in Figure 5.26. The two principal manufacturing techniques for producing multimode fiber star couplers are the mixer-rod and the fused biconical taper (FBT) methods. In the mixer-rod method illustrated in Figure 5.32 a thin platelet of glass is employed, which effectively mixes the light from one fiber, dividing it among the outgoing fibers. This method can be used to produce a transmissive star coupler or a reflective star coupler, as displayed in Figure 5.32. The typical insertion loss for an 8×8 mixer-rod transmissive star coupler with fiber pigtails is 12.5 dB with port to port uniformity of ± 0.7 dB [Ref. 84].

The manufacturing process for the fused biconical taper star coupler is similar to that discussed in Section 5.6.1 for the three and four port FBT coupler. Thus the fibers which constitute the star coupler are bundled, twisted, heated and pulled, to form the device illustrated in Figure 5.33. With multimode fiber this method

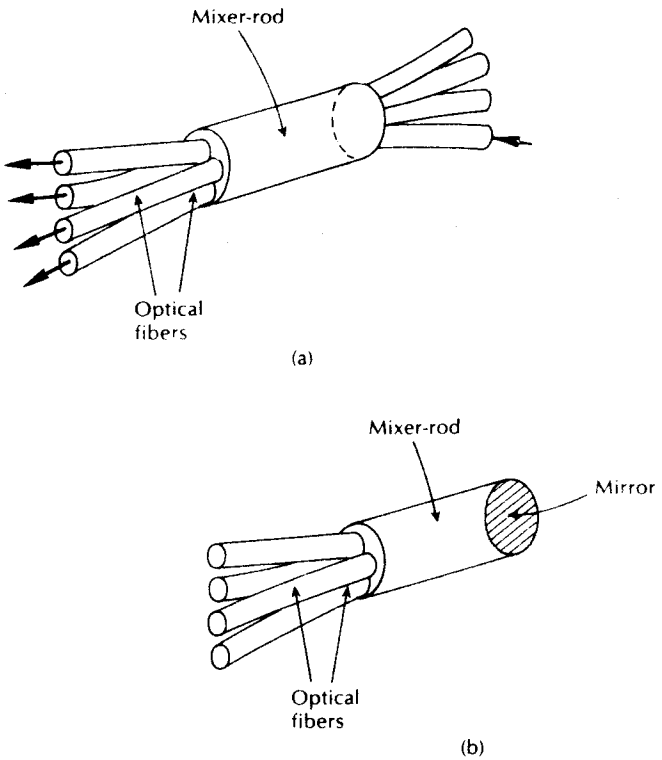


Figure 5.32 Fiber star couplers using the mixer-rod technique: (a) transmissive star coupler; (b) reflective star coupler.

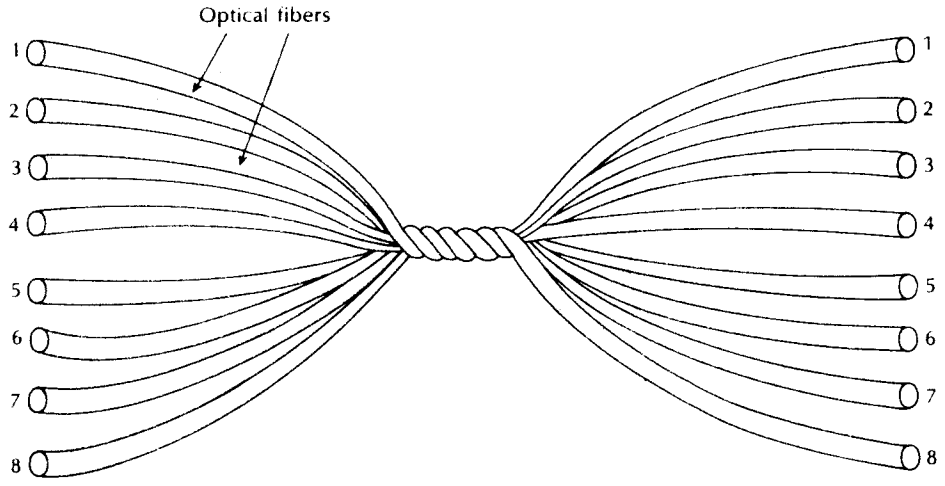


Figure 5.33 Fiber fused biconical taper 8 × 8 port star coupler.

relies upon the coupling of higher order modes between the different fibers. It is therefore highly mode dependent, which results in a relatively wide port to port output variation in comparison with star couplers based on the mixer-rod technique [Ref. 84].

In an ideal star coupler the optical power from any input fiber is evenly distributed among the output fibers. The total loss associated with the star coupler comprises its theoretical splitting loss together with the excess loss. The splitting loss is related to the number of output ports N following:

$$\textit{Splitting loss}\ (\text{star coupler}) = 10 \log_{10} N\ (\text{dB}) \tag{5.27}$$

It should be noted that for a reflective star coupler N is equal to the total number of ports (both input and output combined).

For a single input port and multiple output ports where $j = 1, N$, then the excess loss is given by:

$$\textit{Excess loss}\ (\text{star coupler}) = 10 \log_{10}\left[P_i \Big/ \sum_{1}^{N} P_j\right]\ (\text{dB}) \tag{5.28}$$

The insertion loss between any two ports on the star coupler may be obtained in a similar manner to the four port coupler using Eq. (5.23). Similarly, the crosstalk between any two input ports is given by Eq. (5.24).

Example 5.8

A 32 × 32 port multimode fiber transmissive star coupler has 1 mW of optical power launched into a single input port. The average measured optical power at

each output port is 14 μW. Calculate the total loss incurred by the star coupler and the average insertion loss through the device.

Solution: The total loss incurred by the star coupler comprises the splitting loss and the excess loss through the device. The splitting loss is given by Eq. (5.27) as:

$$\begin{aligned} \textit{Splitting loss} &= 10 \log_{10} N = 10 \log_{10} 32 \\ &= 15.05 \text{ dB} \end{aligned}$$

The excess loss may be obtained from Eq. (5.28) where:

$$\textit{Excess loss} = 10 \log_{10}\left[P_i \Big/ \sum_{1}^{N} P_j\right] = 10 \log_{10}[10^3/32 \times 14] = 3.49 \text{ dB}$$

Hence the total loss for the star coupler:

$$\begin{aligned} \textit{Total loss} = \textit{splitting loss} + \textit{excess loss} &= 15.05 + 3.49 \\ &= 18.54 \text{ dB} \end{aligned}$$

The average insertion loss from the input port to an output port is provided by Eq. (5.23) as:

$$\textit{Insertion loss} = 10 \log_{10} \frac{10^3}{14} = 18.54 \text{ dB}$$

Therefore, as may have been anticipated, the total loss incurred by the star coupler is equivalent to the average insertion loss through the device. This result occurs because the total loss is the loss incurred on a single (average) optical path through the coupler which effectively defines the average insertion loss for the device.

An alternative strategy for the realization of a star coupler is to construct a ladder coupler, as illustrated in Figure 5.34. The ladder coupler generally comprises a number of cascaded stages, each incorporating three or four port FBT couplers in order to obtain a multiport output. Hence, the example shown in Figure 5.34 consists of three stages, which gives eight output ports. It must be noted, however, that when three port couplers are used such devices do not form symmetrical star couplers* in that they provide a $1 \times N$ rather than a $N \times N$ configuration. Nevertheless, the ladder coupler presents a useful device to achieve a multiport output with relatively low insertion loss. Furthermore, when four port couplers are employed, then a true $N \times N$ star coupler may be obtained. It may be deduced from Figure 5.34 that the number of output ports N obtained with an M stage ladder coupler is 2^M. These devices have found relatively widespread application for the production of single-mode fiber star couplers.

* Such devices are sometimes referred to as tree couplers.

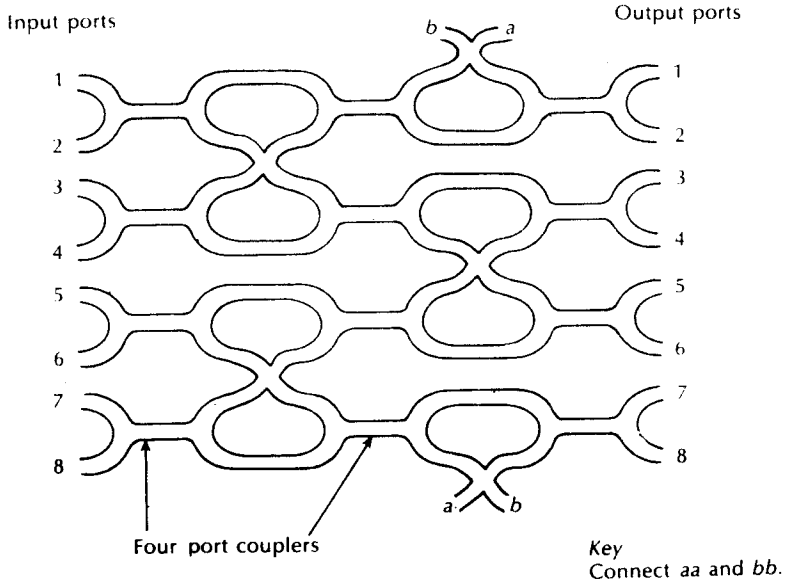


Figure 5.34 8×8 star coupler formed by cascading 12 four port couplers (ladder coupler). This strategy is often used to produce low loss single-mode fiber star or tree couplers.

Example 5.9

A number of three port single-mode fiber couplers are utilized in the fabrication of a tree (ladder) coupler with sixteen output ports. The three port couplers each have an excess loss of 0.2 dB with a split ratio of 50%. In addition, there is a splice loss of 0.1 dB at the interconnection of each stage. Determine the insertion loss associated with one optical path through the device.

Solution: The number of stages M within the ladder design is given by $2^M = 16$. Hence $M = 4$. Thus the excess loss through four stages of the coupler with three splices is:

$$\textit{Excess loss} = (4 \times 0.2) + (3 \times 0.1) = 1.1 \text{ dB}$$

Assuming a 50% split ratio at each stage, the splitting loss for the coupler may be obtained using Eq. (5.27) as:

$$\textit{Splitting loss} = 10 \log_{10} 16 = 12.04 \text{ dB}$$

Hence the insertion loss for the coupler which is equivalent to the total loss for one optical path though the device is:

$$\textit{Insertion loss} = \textit{splitting loss} + \textit{excess loss} \text{ (four stages)}$$
$$= 12.04 + 1.1 = 13.14 \text{ dB}$$

Significantly lower excess losses than that indicated in Example 5.9 have been achieved with single-mode fiber ladder couplers. In particular, a mean excess loss of only 0.13 dB for an 8×8 star coupler constructed using this technique has been reported [Ref. 96]. Four port FBT couplers with mean excess losses of 0.05 dB were used in this device. Alternatively, 3×3 single-mode fiber FBT couplers have been employed as a basis for ladder couplers. For example, a 9×9 star coupler with an excess loss of 1.46 dB and output port power uniformity of ± 1.50 dB has been demonstrated [Ref. 97].

5.6.3 Wavelength division multiplexing couplers

It was indicated in Section 5.6 that wavelength division multiplexing (WDM) devices are a specialized coupler type which enable light from two or more optical sources of differing nominal peak optical wavelength to be launched in parallel into a single optical fiber. Hence such couplers perform as either wavelength multiplexers or wavelength demultiplexers (see Section 11.9.3). The spectral performance characteristic for a typical five channel WDM device is shown in Figure 5.35. The important optical parameters associated with the WDM coupler are the attenuation of the light over a particular wavelength band, the interband isolation and the wavelength band or channel separation. Ideally, the device should have a low loss transmission window for each wavelength band, giving a low

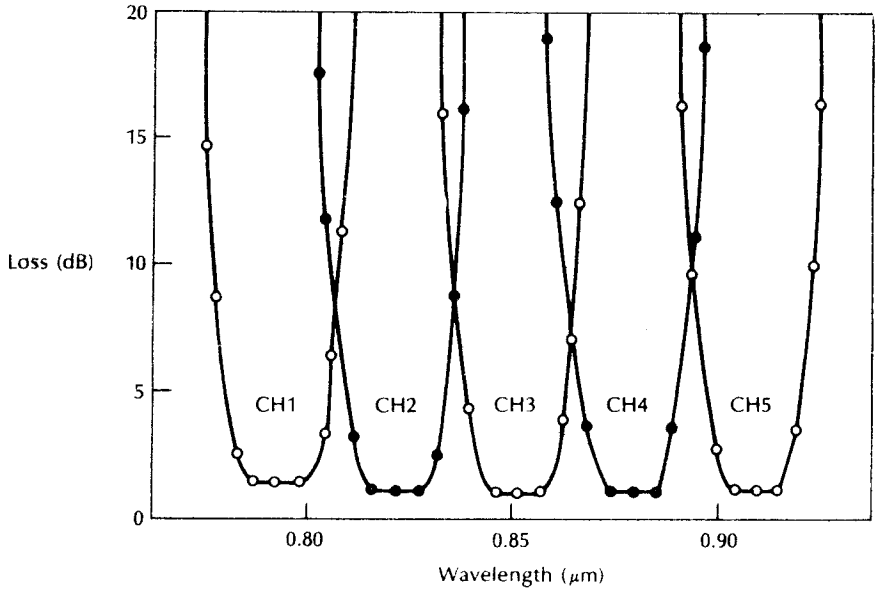


Figure 5.35 Typical flat passband spectral output characteristic for a WDM demultiplexer device (diffraction grating type) [Ref. 99].

insertion loss.* In addition, the device should exhibit high interband isolation, thus minimizing crosstalk. However, in practice, high interchannel isolation is only required at the receiver (demultiplexer) end of the link or at both ends in a bidirectional system. Finally, the channel separation should be as small as may be permitted by light source availability and stability together with crosstalk considerations.

Numerous techniques have been developed for the implementation of WDM couplers. Passive devices, however, may be classified into three major categories [Ref. 98], two of which are core interaction types: namely, angularly dispersive (usually diffraction grating) and filter, whilst the other is a surface interaction type which may be employed with single-mode fiber in the form of a directional coupler. Any other implementations tend to be hybrid combinations of the two core interaction types.

Although a glass prism may be utilized as an angularly dispersive element to facilitate wavelength multiplexing and demultiplexing, the principal angularly dispersive element used in this context is the diffraction grating. Any arrangement which is equivalent in its action to a number of parallel equidistant slits of the same width may be referred to as a diffraction grating. A common form of diffraction grating comprises an epoxy layer deposited on a glass substrate, on which lines are blazed. There are two main types of blazed grating. The first is produced by conventional mechanical techniques, whilst the other is fabricated by the anisotropic etching of single crystal silicon [Ref. 99] and hence is called a silicon grating. The silicon grating has been found to be superior to the conventional mechanically ruled device, since it provides greater design freedom in the choice of blazing angle θ_B (see Figure 5.36) and grating constant (number of lines per unit length). It is also highly efficient and produces a more environmentally stable surface.

A diffraction grating reflects light in particular directions according to the grating constant, the angle at which the light is incident on the grating and the optical wavelength. Two main structural types are used in the manufacture of WDM couplers: the Littrow device which employs a single lens and a separate plane grating, and the concave grating which does not utilize a lens since both focusing and diffraction functions are performed by the grating.

In a Littrow mounted grating, the blaze angle of the grating is such that the incident and reflected light beams follow virtually the same path, as illustrated in Figure 5.36, thereby maximizing the grating efficiency and minimizing lens astigmatism. For a given centre wavelength λ, the blaze angle is set such that [Ref. 100]

$$\theta_B = \sin^{-1}\left(\frac{\lambda}{2x}\right) \tag{5.29}$$

* In the case of the WDM coupler the device loss is specified by the insertion loss associated with a particular wavelength band. The use of excess loss as in the case of other fiber couplers is inappropriate because the optical signals are separated into different wavelength bands.

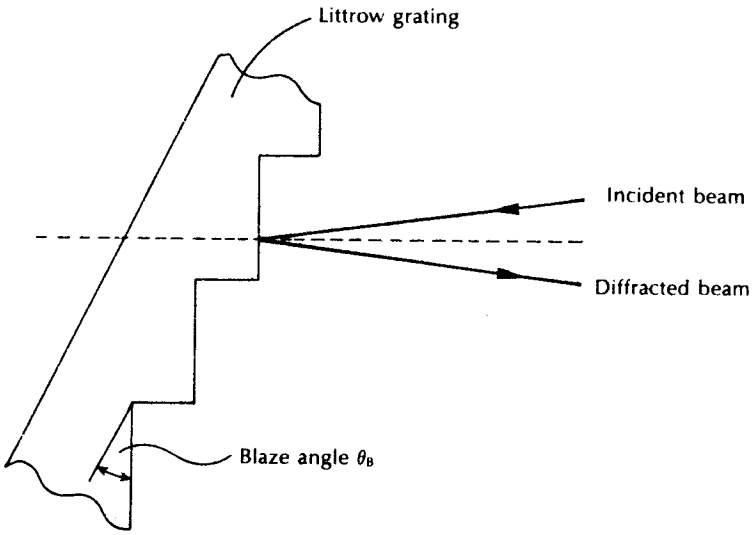


Figure 5.36 Littrow mounted diffraction grating.

where x is the line spacing on the grating. Schematic diagrams of Littrow type grating demultiplexers employing a conventional lens [Ref. 99] and a GRIN-rod lens [Ref. 101] are shown in Figure 5.37. The use of a spherical ball microlens has also been reported [Ref. 102]. Although all the lens type devices exhibit similar operating mechanisms and hence performance, the GRIN-rod lens configuration proves advantageous for its compactness and ease of alignment. Therefore the operation of a GRIN-rod lens type demultiplexer is considered in greater detail.

Referring to Figure 5.37(b), the single input fiber and multiple output fibers are arranged on the focal plane of the lens, which, for a quarter pitch GRIN-rod lens, is coincident with the fiber end face (see Section 5.5.1). The input wavelength multiplexed optical beam is collimated by the lens and hence transmitted to the diffraction grating, which is offset at the blaze angle so that the incoming light is incident virtually normal to the groove faces. The required offset angle can be produced by interposing a prism (glass wedge) between the lens and the grating, as illustrated in Figure 5.37(b) or, alternatively, by cutting and polishing the GRIN-rod lens and by mounting the grating on its end face. The former method gives superior performance since the optical properties of the GRIN-rod lens are not altered [Ref. 102]. On reflection from the grating, the diffraction process causes the light to be angularly dispersed according to the optical wavelength. Finally, the different optical wavelengths pass through the lens and are focused onto the different collecting output fibers. Devices of this type have demonstrated channel insertion losses of less than 2 dB and channel spacings of 18 nm with low crosstalk [Ref. 103].

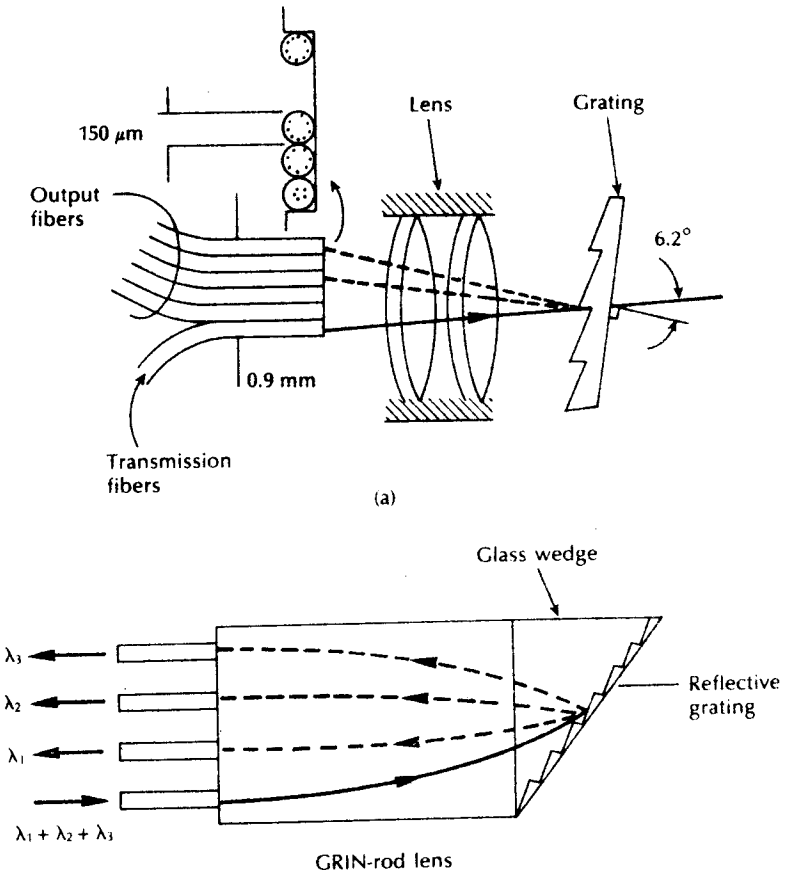


Figure 5.37 Littrow type grating demultiplexers: (a) using a conventional lens [Ref. 99]; (b) using a GRIN-rod lens [Ref. 101].

Finally, single-mode wavelength multiplexer and demultiplexer pairs based on a planar diffraction grating and a lithium niobate strip waveguide structure have also been reported [Ref. 104]. Six wavelength multiplexed channels were demonstrated, three over the wavelength region from 1275 nm to 1335 nm and three over the wavelength range from 1510 nm to 1570 nm. Crosstalk levels were less than −25 dB, with insertion losses for the multiplexer and demultiplexer of 5 to 8 dB and 1 to 2.2 dB respectively.

The other major core interaction type WDM devices employ optical filter technology. Optical spectral filters fall into two main categories: namely, interference filters and absorption filters. Dielectric thin film (DTF) interference

filters can be constructed from alternate layers of high refractive index (e.g. zinc sulphide) and low refractive index (e.g. magnesium fluoride) materials, each of which is one quarter wavelength thick [Ref. 107]. In this structure, shown schematically in Figure 5.38, light which is reflected within the high index layers does not suffer any phase shift on reflection, while those optical beams reflected within the low index layers undergo a phase shift of 180°. Thus the successive reflected beams recombine constructively at the filter front face, producing a high reflectance over a limited wavelength region which is dependent upon the ratio between the high and low refractive indices. Outside this high reflectance region, the reflectance changes abruptly to a low value. Consequently, the quarter wave stack can be used either as a high pass filter, a low pass filter, or as a high reflectance coating.

Absorption filters comprise a thin film of material (e.g. germanium) which exhibits an absorption edge at a specific wavelength. Absorption filters usually display very high rejection in the cutoff region. However, as their operation is dependent upon the fundamental optical properties of the material structure, they tend to be inflexible because the edge positions are fixed. Nevertheless, by fabricating interference filters on to an absorption layer substrate, a filter can be obtained which combines the sharp rejection of the absorption filter together with the flexibility of the interference filter. Such combined structures can be used as high performance edge filters.

Specific filter WDM coupler designs are now considered in further detail. Firstly, edge filters are generally used in devices which require the separation of two wavelengths (generally reasonably widely separated by 10% or more of median wavelength). A configuration which has been adopted [Ref. 106] is one in which the fiber is cleaved at a specific angle and then an edge filter is interposed between

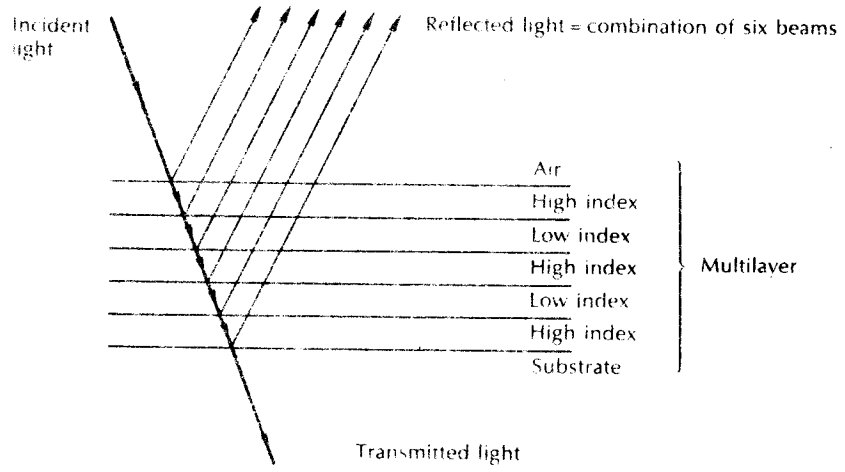


Figure 5.38 Multilayer interference filter structure.

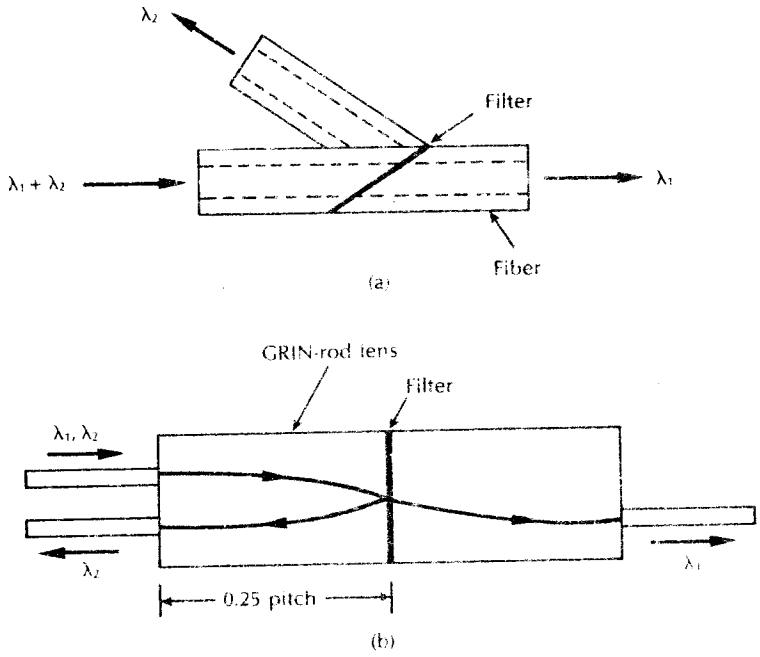


Figure 5.39 Two wavelength interference filter demultiplexers: (a) fiber-end device; (b) GRIN-rod lens device.

the two fiber ends, as illustrated in Figure 5.39(a). In a demultiplexing structure light at one wavelength is reflected by the filter and collected by a suitably positioned receive fiber, whilst the other optical wavelength is transmitted through the filter and then propagates down the cleaved fiber. Such a device, which has been tested with LED sources emitting at centre wavelengths of 755 nm and 825 nm, exhibited insertion losses of 2 to 3 dB with crosstalk levels less than −60 dB [Ref. 106]. An alternative two wavelength WDM device employing a cascaded BPF sandwiched between two GRIN-rod lenses is shown in Figure 5.39(b). A practical two channel (operating at wavelengths of 1.2 μm and 1.3 μm) multiplex/demultiplex system which is capable of operation in both directions using this WDM design has been reported [Ref. 107] to exhibit low insertion losses of around 1.5 dB with crosstalk levels less than −58 dB. This device also displayed acceptable environmental stability with insertion loss variations of less than 0.3 dB throughout a range of tests (i.e. vibration, temperature cycling and damp/heat tests).

Multiple wavelength multiplexer/demultiplexer devices employing DTF interference filters may be constructed from a suitably aligned series of bandpass filters with different passband wavelength regions, cascaded in such a way that each filter transmits a particular wavelength, but reflects all others. Such a multiple reflection demultiplexing device is illustrated in Figure 5.40. This structure has the

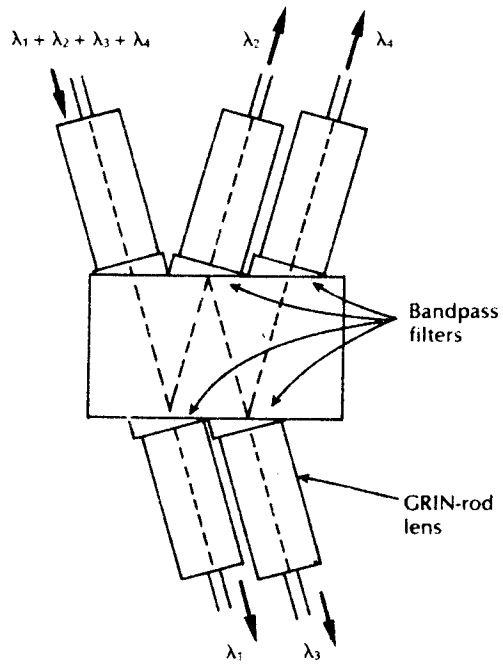


Figure 5.40 GRIN-rod lensed bandpass demultiplexer.

disadvantage that the insertion losses increase linearly with the number of multiplexed channels since losses are incurred at each successive reflection due to filter imperfections and the difficulties of maintaining good alignment [Ref. 98].

A two channel slab waveguide version of a filter WDM device has recently been introduced by Corning, which is based on the same technology as their optical waveguide coupler (see Section 5.6.1). The wavelength separation is accomplished within the waveguide using a diochroic filter which intersects the path of the incoming light beam. Longer wavelengths are transmitted and shorter wavelengths reflected. The multiplexer/demultiplexer device reported [Ref. 108] is compatible with both 50/125 μm and 85/125 μm graded index fibers. It combines/separates optical wavelength regions between 0.8 to 0.9 μm and 1.2 to 1.4 μm with an insertion loss lower than 1.5 dB and crosstalk levels less than -25 dB.

The wavelength dependent characteristics of single-mode fiber directional couplers were mentioned in Section 5.6.1. Both single-mode ground fiber and fused biconical taper fiber couplers can be fabricated to provide the complete transfer of optical power between the two fibers. However, since the optical power coupling characteristic of such single-mode fiber couplers is highly wavelength dependent they can be used to fabricate WDM devices.

Optical power transfer within multimode fiber couplers is a mode dependent phenomenon which, in general, takes place between the higher order modes propagating in the outer reaches of the fiber cores as well as in the cladding regions.

These higher order modes couple more freely when the fibers are in close proximity. Consequently, the spectral dependence of light transfer within multimode fiber couplers is far less pronounced and predictable than that exhibited by single-mode fiber structures. Therefore, multimode fiber WDM devices cannot readily be fabricated using the fused biconical taper or ground fiber techniques.

Optical power is coupled between two single-mode fibers by bringing the fiber cores close together over a region known as the interaction length. One or two methods are generally used to perform this function. The first technique [Ref.109] necessitates bending and fixing the two fibers into two blocks of suitable material (e.g. quartz) prior to grinding the two blocks down so that a proportion of each of the fiber cladding regions is worn away. Finally, the two blocks are brought together, as shown in Figure 5.41.

Parallel single-mode fiber waveguides exchange energy with a spatial period (coupling length) $L = 2\pi/k$, where k is the coupling coefficient (units of inverse length) for the two interacting waveguide modes [Ref. 110]. This result can be extended to curved regions where spacing between the waveguides over the interaction length is no longer fixed [Ref. 111]. Thus, for a pair of fibers curved against each other, the coupling coefficient k is a nonlinear function of the interaction length (which in turn is proportional to the square root of the radius of curvature R), the minimum spacing between the fiber cores, the refractive index of the intervening material, the fiber parameters and the wavelength of the light. The wavelength dependent properties of a single-mode ground fiber coupler can therefore be altered by adjusting several different parameters.

An early demonstration of such a two channel ground fiber directional coupler was made from two identical 2 μm core diameter single-mode fibers [Ref. 112]. This device, with a radius of curvature $R_1 = R_2 = 70$ cm and a minimum core separation of 4.5 μm gave a measured coupling ratio which followed the typical sinusoidal pattern, with approximately two periods over the 0.45 μm to 0.9 μm

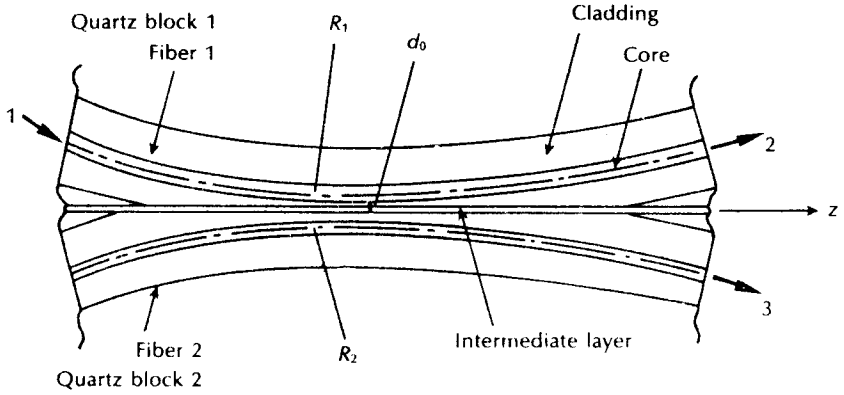


Figure 5.41 Schematic diagram showing ground (polished) single-mode fiber coupler.

wavelength region. By offsetting the cores laterally (i.e. in the direction z indicated in Figure 5.41), and effectively altering d_0, the spectral characteristics of the coupler were altered. The sinusoidal response curve was shifted by around 400 nm with a lateral offset of 5 μm. Interchannel wavelength spacings (wavelength separation between minimum and maximum on the sinusoidal pattern) were about 140 nm for this structure, which exhibited insertion losses as low as 0.1 dB using suitable index matching.

Wavelength selective ground fiber directional couplers constructed from single-mode fibers of different core diameters and refractive indices exhibit propagation constants which are matched at only one wavelength and hence can be used to produce true bandpass filters. The centre wavelength and spectral bandwidth of these couplers are essentially determined by the fiber parameters [Ref. 113]. Measured insertion losses for such devices fabricated for operation in the longer wavelength region were between 0.5 and 0.6 dB with crosstalk levels less than -22 dB [Ref. 109].

The second method of fabricating a single-mode fiber WDM coupler is the fused biconical taper (FBT) technique [Ref. 85]. Carefully fabricated fused couplers display very low insertion losses and provide a high degree of environmental stability. The manufacturing process requires the single-mode fibers to be fused together at around 1500 °C before being pulled whilst heat is still applied. The pulling process decreases the fiber core size causing the evanescent field of the transmitted optical signal to spread out further from the fiber core, which enables light to couple into the adjacent fiber. In practice this manufacturing process necessitates the monitoring of the optical power output from the two fibers, the process being halted when the required coupling ratio is reached [Ref. 114].

In common with the ground fiber coupler constructed by using similar fibers, the optical power transferred between the two fibers (or the coupling ratio) in a FBT coupler as a function of wavelength is sinusoidal with a period dependent on the dimensions and the geometry of the fused cross section, and on the refractive index of the surrounding medium [Ref. 90]. It can also depend upon whether the fusing process produces a coupling region where the two fiber cores are close (strongly fused) or relatively far apart (weakly fused) [Ref. 89]. The most popular method of varying the periodic coupling function in such fused WDM couplers is to extend the interaction length by continuing the stretching process during fusing. An increase in the interaction length has the effect of increasing the coupling ratio period. Such two channel devices have displayed insertion losses of 0.25 and 0.37 dB with crosstalk levels less than -22 dB [Ref. 115]. It should be noted, however, that a limitation with these WDM couplers is that they are not well suited for the provision of closely spaced or multiple channels.

In summary, the generalized performance characteristics of the various passive wavelength division multiplexing coupler generic types are provided in Table 5.1 [Ref. 98]. It should be stressed, however, that the entries in the table signify the characteristics of typical WDM devices of each type. In this context they are intended to act as a guide to the reader and are not a definitive statement of absolute device performance.

Table 5.1 Generalized characteristics of passive wavelength division multiplexer/demultiplexer device types [Ref. 98]

Device type	Mechanism	Implementation	Channel spectral bandwidth	Channel separation	No. of channels	Interchannel crosstalk levels	Insertion loss	Environmental stability	Drawbacks
Prism	Angularly dispersive	Bulk optic: difficult/ expensive	10–20 nm	Around 30 nm	Multi		Less than 10 dB	Acceptable	Optical material constraints produce low values of angular dispersion and bulkly devices
Diffraction grating	Angularly dispersive	Bulk optic: Straightforward compact	Dependent on fiber type/ dimensions and grating line constant but typically around 10 nm minimum	Dependent on fiber type/ dimensions and grating line constant but generally less than 40 nm	Multi, but generally in one wavelength region (e.g. 800–900 nm)	Better than −30 dB	Independent of no. of channels and typically less than 5 dB	Acceptable	High insertion loss for equal input fibers. Generally must be designed for use with particular fibers
Filter	Interference of absorption	Bulk optic: Straightforward compact	Minimum of 2.5 nm at 850 nm 4 nm at 1300 nm	May be used for widely separated wavelengths but also as low as 10 nm	Two for single filter. Multi requires cascaded filters	Better than −30 dB	Increases with number of channels generally greater than 1 dB	Generally acceptable apart from moisture absorption	Cascaded filters give increased insertion losses
Fiber directional coupler	Wavelength dependent power transfer between adjacent fibers	Fused biconical Ground fiber Straightforward compact	Generally periodic but large at around 40 nm	Generally widely separated around 100 nm	Two for single coupler Multi requires cascaded couplers	Around −20 dB	Low (0.2 dB typical)	Good	Only single-mode fiber devices. Ground fiber type requires careful fabrication
Integrated waveguide	Slab, multimode Planar, single mode	Integrated. Very compact Do not require lens	100 nm available	Wide – 850/1300 nm available Grating types around 5 nm	Two Multi, Ten demonstrated	Better than −25 dB Better than −20 dB	Low – less than 1.5 dB Medium – high (2–5 dB)	Potentially good	Commerically available Commerical devices not generally available

Problems

5.1 State the two major categories of fiber–fiber joint, indicating the differences between them. Briefly discuss the problem of Fresnel reflection at all types of optical fiber joint, and indicate how it may be avoided.

A silica multimode step index fiber has a core refractive index of 1.46. Determine the optical loss in decibels due to Fresnel reflection at a fiber joint with:

(a) a small air gap;
(b) an index matching epoxy which has a refractive index of 1.40.

It may be assumed that the fiber axes and end faces are perfectly aligned at the joint.

5.2 The Fresnel reflection at a butt joint with an air gap in a multimode step index fiber is 0.46 dB. Determine the refractive index of the fiber core.

5.3 Describe the three types of fiber misalignment which may contribute to insertion loss at an optical fiber joint.

A step index fiber with a 200 μm core diameter is butt jointed. The joint which is index matched has a lateral offset of 10 μm but no longitudinal or angular misalignment. Using two methods, estimate the insertion loss at the joint assuming the uniform illumination of all guided modes.

5.4 A graded index fiber has a characteristic refractive index profile (α) of 1.85, and a core diameter of 60 μm. Estimate the insertion loss due to a 5 μm lateral offset at an index matched fiber joint assuming the uniform illumination of all guided modes.

5.5 A graded index fiber with a parabolic refractive index profile ($\alpha = 2$) has a core diameter of 40 μm. Determine the difference in the estimated insertion losses at an index matched fiber joint with a lateral offset of 1 μm (no longitudinal or angular misalignment). When performing the calculation assume (a) the uniform illumination of only the guided modes and (b) the uniform illumination of both guided and leaky modes.

5.6 A graded index fiber with a 50 μm core diameter has a characteristic refractive index profile (α) of 2.25. The fiber is jointed with index matching and the connection exhibits an optical loss of 0.62 dB. This is found to be solely due to a lateral offset of the fiber ends. Estimate the magnitude of the lateral offset assuming the uniform illumination of all guided modes in the fiber core.

5.7 A step index fiber has a core refractive index of 1.47, a relative refractive index difference of 2% and a core diameter of 80 μm. The fiber is jointed with a lateral offset of 2 μm, an angular misalignment of the core axes of 3° and a small air gap (no longitudinal misalignment). Estimate the total insertion loss at the joint which may be assumed to comprise the sum of the misalignment losses.

5.8 Briefly outline the factors which cause intrinsic losses of fiber–fiber joints.

(a) Plot the loss resulting from a mismatch in multimode fiber core diameters or numerical apertures over a mismatch range 0 to 50%.

(b) An optical source is packaged with a fiber pigtail comprising 62.5/125 μm graded index fiber with a numerical aperture of 0.28 and a profile parameter of 2.1. The fiber pigtail is spliced to a main transmission fiber which is 50/125 μm graded index fiber with a numerical aperture of 0.22 and a profile parameter of 1.9. When the fiber axes are aligned without either a gap, radial or angular misalignment, calculate the insertion loss at the splice.

5.9 Describe what is meant by the fusion splicing of optical fibers. Discuss the advantages and drawbacks of this jointing technique.

A multimode step index fiber with a core refractive index of 1.52 is fusion spliced. The splice exhibits an insertion loss of 0.8 dB. This insertion loss is found to be entirely due to the angular misalignment of the fiber core axes which is 7°. Determine the numerical aperture of the fiber.

5.10 Describe, with the aid of suitable diagrams, three common techniques used for the mechanical splicing of optical fibers.

A mechanical splice in a multimode step index fiber has a lateral offset of 16% of the fiber core radius. The fiber core has a refractive index of 1.49, and an index matching fluid with a refractive index of 1.45 is inserted in the splice between the butt jointed fiber ends. Assuming no longitudinal or angular misalignment, estimate the insertion loss of the splice.

5.11 Discuss the principles of operation of the two major categories of demountable optical fiber connector. Describe in detail a common technique for achieving a butt jointed fiber connector.

A butt jointed fiber connector used on a multimode step index fiber with a core refractive index of 1.42 and a relative refractive index difference of 1% has an angular misalignment of 9°. There is no longitudinal or lateral misalignment but there is a small air gap between the fibers in the connector. Estimate the insertion loss of the connector.

5.12 Briefly describe the types of demountable connector that may be used with single-mode fibers. Further, indicate the problems involved with the connection of single-mode fibers.

A single-mode fiber connector is used with a 6 μm core diameter silica (refractive index 1.46) step index fiber which has a normalized frequency of 2.2 and a numerical aperture of 0.9. The connector has a lateral offset of 0.7 μm and an angular misalignment of 0.8°. Estimate the total insertion loss of the connector assuming that the joint is index matched and that there is no longitudinal misalignment.

5.13 A 10 μm core diameter single-mode fiber has a normalized frequency of 2.0. A fusion splice at a point along its length exhibits an insertion loss of 0.15 dB. Assuming only lateral misalignment contributes to the splice insertion loss, estimate the magnitude of the lateral misalignment.

5.14 A 5 μm core diameter single-mode step index fiber has a normalized frequency of 1.7, a core refractive index of 1.48 and a numerical aperture of 0.14. The loss in decibels due to angular misalignment at a fusion splice with a lateral offset of 0.4 μm is twice that due to the lateral offset. Estimate the magnitude in degrees of the angular misalignment.

5.15 Given the following parameters for a single-mode step index fiber with a fusion splice estimate (a) the fiber core diameter: and (b) the numerical aperture for the fiber.

Fiber normalized frequency = 1.9
Fiber core refractive index = 1.46
Splice lateral offset = 0.5 μm
Splice lateral offset loss = 0.05 dB
Splice angular misalignment = 0.3°
Splice angular misalignment loss = 0.04 dB

5.16 Two single-mode fibers have mode-field diameters of 9 μm and 11 μm. Assuming that there are no extrinsic losses calculate the coupling loss between the fibers as a result of the mode-field diameter mismatch. Comment on the result in relation to the direction of transmission of the optical signal between the two fibers.

Determine the loss if the mode-field diameter mismatch between the fibers is increased to 30%.

5.17 With the aid of simple sketches outline the major categories of multiport optical fiber coupler.

Describe two common methods used in the fabrication of three and four port fiber couplers.

5.18 A four port FBT coupler is shown in Figure 5.29. In addition a section of a tapered multimode step index fiber from such a coupler may be observed in Figure 5.42. A meridional ray propagating along the taper (characterized by the taper angle γ) is shown to undergo an increase in its propagation angle (i.e. the angle formed with the fiber axis). However, as long as the angle of incidence remains larger than the critical angle, then the ray is still guided and it emerges from the taper region forming an angle θ_0 with the fiber axis. When the taper is smooth and the number of reflections is high, then, in Figure 5.42, $\sin\theta_0 = R_1/R_2 \sin\theta_i$ where R_1 and R_2 are the core radii before and after the taper respectively. Show that the numerical aperture for the tapered fiber NA_T is given by:

$$NA_T = \frac{R_2}{R_1}(n_1^2 - n_2^2)^{\frac{1}{2}}$$

where n_1 and n_2 are the refractive indices of the fiber core and cladding respectively. Comment on this result when considering the modes of the light launched into the coupler.

5.19 The measured optical output powers from ports 3 and 4 of a multimode fiber FBT coupler are 47.0 μW and 52.0 μW respectively. If the excess loss specified for the device is 0.7 dB, calculate the amount of optical power that is launched into port 1 in order to obtain these output power levels. Hence, determine the insertion losses between the input and two output ports, as well as the split ratio for the device.

When the specified crosstalk for the coupler is -45 dB, calculate the optical output power level that would be measured at port 2 when the above input power level is maintained.

5.20 Indicate the distinction between fiber star and tree couplers.

Discuss the major techniques used in the fabrication of multimode fiber star couplers and describe how this differs from the strategy that tends to be adopted to produce single-mode fiber star couplers.

5.21 A 64 × 64 port transmissive star coupler has 1.6 mW of optical power launched into a single input port. If the device exhibits an excess loss of 3.90 dB, determine the total loss through the device and the average optical power level that would be expected at each output port.

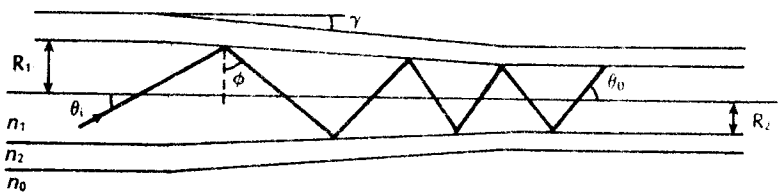


Figure 5.42 Section of a tapered multimode step index fiber for Problem 5.18.

5.22 An 8×8 port multimode fiber reflective star coupler has -8.0 dBm of optical power launched into a single port. The average measured optical power at each output port is -22.8 dBm. Obtain the excess loss for the device and hence the total loss experienced by an optical signal in transmission through the coupler. Check the result.

5.23 A number of four port single-mode fiber couplers are employed in the fabrication of a 32×32 port star coupler. Each four port coupler has a split ratio of 50% and when an optical input power level of -6 dBm is launched into port 1, the output power level from port 3 is found to be 122 μW. Furthermore, there is a splice loss of 0.06 dB at the interconnection of each stage within the ladder design. Calculate the optical power emitted from each of the output ports when the -6 dBm power level is launched into any one of the input ports. Check the result.

5.24 Outline the three major categories of passive wavelength division multiplexing coupler. Describe in detail one implementation of each category. Comment on the relative merits and drawbacks associated with each of the WDM devices you have described.

Answers to numerical problems

5.1 (a) 0.31 dB; (b) 3.8×10^{-4} dB
5.2 1.59
5.3 0.29 dB
5.4 0.67 dB
5.5 (a) 0.19 dB; (b) 0.17 dB; difference 0.02 dB
5.6 4.0 μm
5.7 0.71 dB
5.8 4.25 dB
5.9 0.35
5.10 0.47 dB
5.11 1.51 dB
5.12 0.54 dB
5.13 1.2 μm
5.14 0.65°
5.15 (a) 7.0 μm; (b) 0.10
5.16 0.17 dB, 0.54 dB
5.19 116.3 μm, 3.93 dB, 3.50 dB, 47.5%, 3.7 nW
5.21 21.96 dB, 10.18 μW
5.22 5.77 dB, 14.80 dB
5.23 6.40 μW

References

[1] S. R. Nagel, 'Fiber materials and fabrication methods', in S. E. Miller and I. P. Kaminov (Eds.), *Optical Fiber Telecommunications II*, Academic Press, pp. 121–215, 1988.

[2] T. R. Rowbotham, 'Submarine telecommunications', *Br. Telecom Technol. J.*, **5**(1), pp. 5–24, 1987.

[3] K. Nakagawa and K. Nosu, 'An overview of very high capacity transmission technology for NTT networks', *J. Lightwave Technol.*, **LT-5**(10), pp. 1498–1504, 1987.

[4] P. Cockrane and M. Brain, 'Future optical fiber transmission technology and networks', *IEEE Commun. Mag.*, pp. 45–60, November, 1988.

[5] D. C. Gloge and I. Jacobs, 'Terrestrial intercity transmission systems', in S. E. Miller and I. P. Kaminow (Eds.) *Optical Fiber Telecommunications II*, Academic Press, pp. 855–878, 1988.

[6] A. Stevenson, S. L. Arambepola, G. L. Blau, T. S. Brown, I. C. Catchpole and A. J. Flavin, 'A 2.4 Gbit/s long-reach optical transmission system', *Br. Telecom. Tech. J.*, 7(1), pp. 92–99, 1989.

[7] S. Fujita, M. Fitamura, T. Torikai, N. Henmi, H. Yamada, T. Suzaki, I. Takano and M. Shikada, '10 Gbit/s, 100 km optical fiber transmission experiment using high-speed MQW DFB-LD and back-illuminated GaInAs APD', *Electron. Lett.*, **25**(11), pp. 702–703, 1989.
[8] M. Born and W. Wolf, *Principles of Optics*, (6th edn), Pergamon Press, 1980.
[9] P. Mossman, 'Connectors for optical fibre systems', *Radio Electron. Eng. (J. IERE)*, **51**(7/8), pp. 333–340, 1981.
[10] J. S. Leach, M. A. Matthews and E. Dalgoutte, 'Optical fibre cable connections', in C. P. Sandbank (Ed.), *Optical Fibre Communication Systems*, pp. 86–105, John Wiley, 1980.
[11] F. L. Thiel and R. M. Hawk, 'Optical waveguide cable connection', *Appl. Opt.*, **15**(11), pp. 2785–2791, 1976.
[12] K. Miyazaki *et al.*, 'Theoretical and experimental considerations of optical fiber connector', OSA Topical Meeting on Opt. Fiber Trans. Williamsburg, Va, paper WA 4-1, 1975.
[13] H. Tsuchiya, H. Nakagome, N. Shimizu and S. Ohara, 'Double eccentric connectors for optical fibers', *Appl. Opt.*, **16**(5), pp. 1323–1331, 1977.
[14] K. J. Fenton and R. L. McCartney, 'Connecting the thread of light', *Electronic Connector Study Group Symposium, 9th Annual Symposium Proc.*, p. 63, Cherry Hill, NJ, 1976.
[15] C. M. Miller, Transmission vs transverse offset for parabolic-profile fiber splices with unequal core diameters', *Bell Syst. Tech. J.*, 55(7), pp. 917–927, 1976.
[16] D. Gloge, 'Offset and tilt loss in optical fiber splices', *Bell Syst. Tech. J.*, **55**(7), pp. 905–916, 1976.
[17] C. M. Miller and S. C. Mettler, 'A loss model for parabolic profile fiber splices', *Bell Syst. Tech. J.*, **57**(9), pp. 3167–3180, 1978.
[18] J. J. Esposito, 'Optical connectors, couplers and switches', in H. F. Wolf (Ed.), *Handbook of Fiber Optics, Theory and Applications*, pp. 241–303, Granada, 1979.
[19] J. F. Dalgleish, 'Connections', *Electronics*, pp. 96–98, 5 Aug. 1976.
[20] W. van Etten and J. van Der Platts, *Fundamentals of Optical Fiber Communications*, Prentice Hall International, 1991.
[21] D. Marcuse, 'Loss analysis of single-mode fiber splices', *Bell Syst. Tech. J.*, **56**(5), pp. 703–718, 1977.
[22] W. A. Gambling, H. Matsumura and A. G. Cowley, 'Jointing loss in single-mode fibres', *Electron. Lett.*, **14**(3), pp. 54–55, 1978.
[23] W. A. Gambling, H. Matsumura and C. M. Ragdale, 'Joint loss in single-mode fibres', *Electron. Lett.*, **14**(15), pp. 491–493, 1978.
[24] D. Botez and G. J. Herskowitz, 'Components for optical communications systems: a review', *Proc. IEEE*, **68**(6), pp. 689–731, 1980.
[25] G. Coppa and P. Di Vita, 'Length dependence of joint losses in multimode optical fibres', *Electron. Lett.*, **18**(2), pp. 84–85, 1982.
[26] S. Nemoto and T. Makimoto, 'Analysis of splice loss in single-mode fibers using a Gaussian field approximation', *Opt. Quantum Electron.*, **11**, pp. 447–457, 1979.
[27] W. C. Young and D. R. Frey, 'Fiber connectors', in S. E. Miller and I. P. Kaminow (Eds.) *Optical Fiber Telecommunications II*, Academic Press, pp. 301–326, 1988.
[28] Y. Ushui, T. Ohshima, Y. Toda, Y. Kato and M. Tateda, 'Exact splice loss prediction for single-mode fiber', *IEEE J. Quantum Electron.*, **QE-18**(4), pp. 755–757, 1982.

[29] K. Petermann, 'Nonlinear distortions due to fibre connectors', *Proceedings of 6th European Conference on Optical Communication* (UK), pp. 80–83, 1980.

[30] K. Petermann, 'Wavelength-dependent transmission at fibre connectors', *Electron Lett.*, **15**(22), pp. 706–708, 1979.

[31] C. M. Miller, S. C. Mettler and I. A. White, *Optical Fiber Splices and Connectors: Theory and methods*, Marcel Dekker, 1986.

[32] M. Ikeda, Y. Murakami and K. Kitayama, 'Mode scrambler for optical fibers', *Appl. Opt.*, **16**(4), pp. 1045–1049, 1977.

[33] N. Nashima and N. Uchida, 'Relation between splice loss and mode conversion in a graded-index optical fibre', *Electron. Lett.*, **15**(12), pp. 336–338, 1979.

[34] A. H. Cherin and J. F. Dalgleish, 'Splices and connectors for optical fibre communications', *Telecommun. J. (Eng. Ed.) Switzerland*, **48**(11), pp. 657–665, 1981.

[35] J. E. Midwinter, *Optical Fibers for Transmission*, John Wiley, 1979.

[36] E. A. Lacy, *Fiber Optics*, Prentice Hall, 1982.

[37] R. Jocteur and A. Tardy, 'Optical fiber splicing with plasma torch and oxyhydric microburner', *2nd European Conference on Optical Fibre Communication* (Paris), 1976.

[38] I. Hatakeyama and H. Tsuchiya, 'Fusion splices for single-mode optical fibers', *IEEE J. Quantum Electron.*, **QE-14**(8), pp. 614–619, 1978.

[39] M. Hirai and N. Uchida, 'Melt splice of multimode optical fibre with an electric arc', *Electron. Lett.*, **13**(5), pp. 123–125, 1977.

[40] M. Tsuchiya and I. Hatakeyama, 'Fusion splices for single-mode optical fibres', *Optical Fiber Transmission II*, Williamsburg, pp. PD1, 1–4, Feb. 1977.

[41] F. Esposto and E. Vezzoni, 'Connecting and splicing techniques', *Optical Fibre Communication*, by Technical Staff of CSELT, pp. 541–643, McGraw-Hill, 1981.

[42] D. B. Payne, D. J. McCartney and P. Healey, 'Fusion splicing of a 31.6 km monomode optical fibre system', *Electron. Lett.*, **18**(2), pp. 82–84, 1982.

[43] O. Kawata, K. Hoshino, Y. Miyajima, M. Ohnishi and K. Ishihara, 'A splicing end inspection technique for single-mode fibers using direct core monitoring', *J. of Lightwave Technol.*, **LT-2**, pp. 185–190, 1984.

[44] D. R. Briggs and L. M. Jayne, 'Splice losses in fusion-spliced optical waveguide fibers with different core diameters and numerical apertures', *Proceedings of 27th International Wire and Cable Symposium*, pp. 356–361, 1978.

[45] I. Hatakeyama, M. Tachikura and H. Tsuchiya, 'Mechanical strength of fusion-spliced optical fibres', *Electron. Lett.*, **14**(19), pp. 613–614, 1978.

[46] C. K. Pacey and J. F. Dalgleish, 'Fusion splicing of optical fibres', *Electron. Lett.*, **15**(1), pp. 32–34, 1978.

[47] T. G. Giallorenzi, 'Optical communications research and technology', *Proc. IEEE*, **66**(7), pp. 744–780, 1978.

[48] K. Nawata, Y. Iwahara and N. Suzuki, 'Ceramic capillary splices for optical fibres', *Electron. Lett.*, **15**(15), pp. 470–472, 1979.

[49] J. G. Woods, 'Fiber optic splices', *Proc. SPIE, Fiber Optic Commun. Technol.*, **512**, pp. 44–56, 1984.

[50] C. M. Miller, 'Loose tube splice for optical fibres', *Bell Syst. Tech. J.*, **54**(7), pp. 1215–1225, 1975.

[51] D. Gloge, A. H. Cherin, C. M. Miller and P. W. Smith, 'Fiber splicing', in S. E. Miller (Ed.), *Optical Fiber Telecommunications*, pp. 455–482, Academic Press, 1979.

[52] P. Hensel, J. C. North and J. H. Stewart, 'Connecting optical fibers', *Electron. Power*, **23**(2), pp. 133–135, 1977.
[53] A. R. Tynes and R. M. Derosier, 'Low-loss splices for single-mode fibres', *Electron. Lett.*, **13**(22), pp. 673–674, 1977.
[54] D. N. Knecht, W. J. Carlsen and P. Melman, 'Fiber optic field splice', *Proc. SPIE Int. Soc. Opt. Eng. (USA)* pp. 44–50, 1982.
[55] G. Cocito, B. Costa, S. Longoni, L. Michetti, L. Silvestri, D. Tribone and F. Tosco, 'COS 2 experiment in Turin: field test on an optical cable in ducts', *IEEE Trans. on Commun.*, **COM-26**(7), pp. 1028–1036, 1978.
[56] J. A. Aberson and K. M. Yasinski, 'Multimode mechanical splices', *Proc. Tenth ECOC* (Germany), p. 182, 1984.
[57] C. M. Miller, G. F. DeVeau and M. Y. Smith, 'Simple high-performance mechanical splice for single mode fibers', *Proc. Opt. Fiber Commun. Conf., OFC '85* (USA) paper MI2, 1985.
[58] M. Kawase, M. Tachikura, F. Nihei and H. Murata, 'Mass fusion splices for high density optical fiber units', *Proc. Eighth ECOC* (France), paper AX-5, 1982.
[59] Y. Katsuyama, S. Hatano, K. Hogari, T. Matsumoto and T. Kokubun, 'Single mode optical fibre ribbon cable', *Electron. Lett.*, **21**, pp. 134–135, 1985.
[60] H. Murata, *Handbook of Optical Fibers and Cables*, Marcel Dekker, 1988.
[61] E. L. Chinnock, D. Gloge, D. L. Bisbee and P. W. Smith, 'Preparation of optical fiber ends for low-loss tape splices', *Bell Syst. Tech. J.*, **54**, pp. 471–477, 1975.
[62] R. Delebecque, E. Chazelas and D. Boscher, 'Flat mass splicing process for cylindrical V-groved cables', *Proc. IWCS '82*, Cherry Hill, NJ (USA), pp. 184–187, 1982.
[63] N. E. Hardwick and S. T. Davies, 'Rapid ribbon splice for multimode fiber splicing', *Proc. Opt. Fiber Commun. Conf., OFC '85* (USA) paper TUQ27, 1985.
[64] T. W. Tamulevich, 'Fiber optic ceramic capillary connectors', *Photonics Spectra*, pp. 65–70, October, 1984.
[65] M. D. Drake, 'A critical review of fiber optic connectors', *Proc. SPIE, Fiber Optic Commun. Technol.*, **512**, pp. 57–69, 1984.
[66] G. Kotelly, 'Special report: fiber optic connectors', *Lightwave, J. of Fiber Optics*, pp. 1, 32–39, April 1989.
[67] W. C. Young, P. Kaiser, N. K. Cheung, L. Curtis, R. E. Wagner and D. M. Folkes, 'A transfer molded biconic connector with insertion losses below 0.3 dB without index match', *Proceedings of 6th European Conference on Optical Communication*, pp. 310–313, 1980.
[68] T. King, 'Fibre optic components for the fibre distributed data interface (FDDI) 100 Mbit/s local area network', *Proc. SPIE, Fibre Optics '88*, **949**, pp. 2–13, 1988.
[69] P. W. Smith, D. L. Bisbee, D. Gloge and E. L. Chinnock, 'A moulded-plastic technique for connecting and splicing optical fiber tapes and cables', *Bell Syst. Tech. J.*, **54**(6), pp. 971–984, 1975.
[70] Y. Fujii, J. Minowa and N. Suzuki, 'Demountable multiple connector with precise V-grooved silicon', *Electron. Lett.*, **15**(14), pp. 424–425, 1979.
[71] M. Oda, M. Ogai, A. Ohtake, S. Tachigami, S. Ohkubo, F. Nihei and N. Kashima, 'Nylon extruded fiber ribbon and its connection', *Proc. Opt. Fiber Commun. Conf., OFC '82* (USA), p. 46, 1982.
[72] S. Tachigami, A. Ohtake, T. Hayashi, T. Iso and T. Shirasawa, 'Fabrication and evaluation of high density multi-fiber plastic connector', *Proc. IWCS*, Cherry Hill, N. J. (USA), pp. 70–75, 1983.

[73] T. Sakake, N. Kashima and M. Oki, 'Very small single-mode ten-fiber connector', *J. Lightwave Technol.*, **6**(2), pp. 269–272, 1988.
[74] A. Nicia, 'Practical low-loss lens connector for optical fibers', *Electron. Lett.*, **14**(16), pp. 511–512, 1978.
[75] A. Nicia and A. Tholen, 'High efficiency ball-lens connector and related functional devices for single-mode fibers', *Proc. Seventh ECOC*, (Denmark) paper 7.5, 1981.
[76] D. M. Knecht and W. J. Carlsen, 'Expanded beam fiber optic connectors', *Proc. SPIE* (USA), pp. 44–50, 1983.
[77] K. Kobayashi, R. Ishikawa, K. Minemura and S. Sugimoto, 'Micro-optic devices for fiber-optic communications', *Fiber and Integrated Optics*, **2**, pp. 1–17, 1979.
[78] W. J. Tomlinson, 'Applications of GRIN rod lenses in optical fiber communication systems', *Appl. Opt.*, **19**, pp. 1127–1138, 1980.
[79] T. Uchida, M. Furukawa, I. Kitano, K. Koizumi and H. Matsomura, 'Optical characteristics of a light focusing guide and its application', *IEEE J. of Quantum Electron.*, **QE-6**, pp. 606–612, 1970.
[80] D. Marcuse and S. E. Miller, 'Analysis of a tubular gas lens', *Bell Syst. Tech. J.*, **43**, pp. 1159–1782, 1965.
[81] S. E. Miller, 'Light Propagation in generalized lenslike media', *Bell Syst. Tech. J.*, **44**, pp. 2017–2064, 1965.
[82] K. Sono, 'Graded index rod lenses', *Laser Focus*, **17**, pp. 70–74, 1981.
[83] J. M. Senior, S. D. Cusworth, N. G. Burrow and A. D. Muirhead, 'Misalignment losses at multimode graded-index fiber splices and GRIN rod lens couplers', *Appl. Opt.*, **24**, pp. 977–982, 1985.
[84] S. van Dorn, 'Fiber optic couplers', *Proc. SPIE, 574, Fiber Optic Couplers, Connectors and Splice Technology II*, pp. 2–8, 1985.
[85] K. O. Hill, D. C. Johnson and R. G. Lamont, 'Optical fiber directional couplers: biconical taper technology and device applications', *Proc. SPIE, Fiber Optic Couplers, Connectors and Splice Technology II*, **574**, pp. 92–99, 1985.
[86] A. K. Agarwal, 'Review of optical fiber couplers', *Fiber Integr. Opt.*, **6**(1), pp. 27–53, 1987.
[87] K. Kobayashi, R. Ishikawa, K. Minemura and S. Sugimoto, 'Micro-optic devices for fiber-optic communications', *Fiber Integr. Opt.*, **2**(1), pp. 1–17, 1979.
[88] B. S. Kawasaki, K. O. Hill and Y. Tremblay, 'Modal-noise generation in biconical taper couplers', *Opt. Lett.*, **6**, p. 499, 1981.
[89] J. V. Wright, 'Wavelength dependence of fused couplers', *Electron. Lett.*, **22**, pp. 329–331, 1986.
[90] F. P. Payne, 'Dependence of fused taper couplers on external refractive index', *Electron. Lett.*, **22**, pp. 1207–1208, 1986.
[91] D. T. Cassidy, D. C. Johnson and K. O. Hill, 'Wavelength dependent transmission of monomode optical fiber tapers', *Appl. Opt.*, **24**, pp. 945–950, 1985.
[92] I. Yokohama, K. Okamoto and J. Noda, 'Polarization-independent optical circulator consisting of two fiber-optic polarizing beamsplitters and two YIG spherical lenses', *Electron. Lett.*, **22**, pp. 370–372, 1985.
[93] E. Paillard, 'Recent developments in integrated optics', *Proc. SPIE, Fiber Optics '87*, **734**, pp. 131–136, 1987.
[94] T. Findalky, 'Glass waveguides by ion exchange', *Opt. Eng.*, **24**, pp. 244–250, 1985.
[95] J. P. Dakin, M. G. Holliday and S. W. Hickling, 'Non invasive optical bus for video distribution', *Proc. SPIE, Fibre Optics '88*, **949**, pp. 36–40, 1988.
[96] G. D. Khoe and H. Lydtin, 'European optical fibers and passive components:

status and trends', *IEEE J. Selected Areas in Commun.*, **SAC-4**(4), pp. 457–471, 1986.

[97] C. C. Wang, W. K. Burns and C. A. Villaruel, '9 × 9 single-mode fiber optic star couplers', *Opt. Lett.*, **10**(1), pp. 49–51, 1985.

[98] J. M. Senior and S. D. Cusworth, 'Devices for wavelength multiplexing and demultiplexing', *IEE Proc., Pt J*, **136**(3), pp. 183–202, 1989.

[99] Y. Fujii, K. Aoyama and J. Minowa, 'Optical demultiplexer using a silicon echette grating', *IEEE J. Quantum Electron.*, **QE-16**, pp. 165–169, 1980.

[100] F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*, Prentice Hall, 1987.

[101] R. Erdmann, 'Prism gratings for fiber optic multiplexing', *Proc. SPIE, Fiber Optics Multiplexing and Modulation*, **417**, pp. 12–17, 1983.

[102] A. Nicia, 'Wavelength multiplexing and demultiplexing systems for single mode and multimode fibers', *Seventh European Conf. on Opt. Commun. (ECOC 81)*, pp. 8.1–7, September 1981.

[103] J. Lipson, C. A. Young, P. D. Yeates, J. C. Masland, S. A. Wartonick, G. T. Harvey and P. H. Read, 'A four channel lightwave subsystem using wavelength multiplexing', *J. Lightwave Technol.*, **LT-3**, pp. 16–20, 1985.

[104] J. Lispon, W. J. Minford, E. J. Murphy, T. C. Rice, R. A. Linke and G. T. Harvey, 'A six-channel wavelength multiplexer and demultiplexer for single mode systems', *J. Lightwave Technol.*, **LT-3**, pp. 1159–1161, 1985.

[105] H. A. Macleod, *Thin Film Optical Filters* (2nd edn), Adam Hilger Ltd, 1986.

[106] G. Winzer, H. F. Mahlein and A. Reichelt, 'Single-mode and multimode all-fiber directional couplers for WDM', *Appl. Opt.*, **20**, pp. 3128–3135, 1981.

[107] Y. Fujii, J. Minowa and H. Tanada, 'Practical two-wavelength multiplexer and demultiplexer: design and performance', *Appl. Opt.*, **22**, pp. 3090–3097, 1983.

[108] M. McCourt and J. L. Malinge, 'Application of ion exchange techniques to the fabrication of multimode wavelength division multiplexers', *Proc. SPIE, Fiber Optics '88*, **949**, pp. 131–137, 1988.

[109] R. Zengerle and O. G. Leminger, 'Wavelength-selective directional coupler made of non-identical single-mode fibers', *J. Lightwave Technol.*, **LT-4**, pp. 823–826, 1986.

[110] D. Marcuse, 'Coupling of degenerative modes in two parallel dielectric waveguides', *Bell Syst. Tech. J.*, **50**, pp. 1791–1816, 1971.

[111] B. S. Kawasaki and K. O. Hill, 'Low loss access coupler for multimode optical fiber distribution networks', *Appl. Opt.*, **16**, pp. 327–328, 1977.

[112] M. J. F. Digonnet and H. J. Shaw, 'Wavelength multiplexing in single mode fiber couplers', *Appl. Opt.*, **22**, pp. 484–492, 1983.

[113] O. Leminger and R. Zengerle, 'Bandwidth of directional-coupler wavelength filters made of dissimilar optical fibres', *Electron. Lett.*, **23**, pp. 241–242, 1987.

[114] R. Zengerle and O. Leminger, 'Narrow band wavelength selective directional-coupler made of dissimilar optical fibres', *J. of Lightwave Technol.*, **LT-5**, pp. 1196–1198, 1987.

[115] H. A. Roberts, 'Single-mode fused wavelength division multiplexer', *Proc. SPIE*, **574**, pp. 100–104, 1985.

[116] J. B. Straws and B. Kawasaki, 'Passive optical components', in E. E. Basch (Ed.), *Optical-Fiber Transmission*, H. W. Sams & Co., pp. 241–264, 1987.

[117] W. J. Tomlinson, 'Passive and low-speed active optical components for fiber systems', in S. E. Miller and I. P. Kaminow (Eds.), *Optical Fiber Telecommunications II*, Academic Press, pp. 369–419, 1988.

6

Optical sources 1: the laser

6.1 Introduction

The optical source is often considered to be the active component in an optical fiber communication system. Its fundamental function is to convert electrical energy in the form of a current into optical energy (light) in an efficient manner which allows the light output to be effectively launched or coupled into the optical fiber. Three

main types of optical light source are available. These are:

(a) wideband 'continuous spectra' sources (incandescent lamps);
(b) monochromatic incoherent sources (light emitting diodes, LEDs);
(c) monochromatic coherent sources (lasers).

To aid consideration of the sources currently in major use the historical aspect must be mentioned. In the early stages of optical fiber communication the most powerful narrowband coherent light sources were necessary due to severe attenuation and dispersion in the fibers. Therefore, gas lasers (helium–neon) were utilized initially. However, the development of the semiconductor injection laser and the LED, together with the substantial improvement in the properties of optical fibers, has given prominence to these two specific sources.

To a large extent these two sources fulfil the major requirements for an optical fiber emitter which are outlined below:

1. A size and configuration compatible with launching light into an optical fiber. Ideally, the light output should be highly directional.
2. Must accurately track the electrical input signal to minimize distortion and noise. Ideally, the source should be linear.
3. Should emit light at wavelengths where the fiber has low losses and low dispersion and where the detectors are efficient.
4. Preferably capable of simple signal modulation (i.e. direct – see Section 7.5) over a wide bandwidth extending from audio frequencies to beyond the gigahertz range.
5. Must couple sufficient optical power to overcome attenuation in the fiber plus additional connector losses and leave adequate power to drive the detector.
6. Should have a very narrow spectral bandwidth (linewidth) in order to minimize dispersion in the fiber.
7. Must be capable of maintaining a stable optical output which is largely unaffected by changes in ambient conditions (e.g. temperature).
8. It is essential that the source is comparatively cheap and highly reliable in order to compete with conventional transmission techniques.

In order to form some comparison between these two types of light source the historical aspect must be enlarged upon. The first generation optical communication sources were designed to operate between 0.8 and 0.9 μm (ideally around 0.85 μm) because initially the properties of the semiconductor materials used lent themselves to emission at this wavelength. Also, as suggested in (3) this wavelength avoided the loss incurred in many fibers near 0.9 μm due to the OH ion (see Section 3.3.2). These early systems utilized multimode step index fibers which required the superior performance of semiconductor lasers for links of reasonable bandwidth (tens of megahertz) and distances (several kilometres). The LED (being a lower power source generally exhibiting little spatial or temporal coherence) was not suitable for long distance wideband transmission, although it found use in more moderate applications.

However, the role of the LED as a source for optical fiber communications was enhanced following the development of multimode graded index fiber. The substantial reduction in intermodal dispersion provided by this fiber type over multimode step index fiber allowed incoherent LEDs emitting in the 0.8 to 0.9 μm wavelength band to be utilized for applications requiring wider bandwidths. This position was further consolidated with the development of second generation optical fiber sources operating at wavelengths between 1.1 and 1.6 μm where both material losses and dispersion are greatly reduced. In this wavelength region, wideband graded index fiber systems utilizing LED sources may be operated over long distances without the need for intermediate repeaters. Furthermore, LEDs offer the advantages of relatively simple construction and operation with the inherent effects of these factors on cost and extended, trouble-free life.

In parallel with these later developments in multimode optical propagation came advances in single-mode fiber construction. This has stimulated the development of single-mode laser sources to take advantage of the extremely low dispersion offered by single-mode fibers. These systems are ideally suited to extra wideband, very long-haul applications and are currently under intensive investigation for long distance telecommunications. On the other hand, light is usually emitted from the LED in many spatial modes which cannot be as efficiently focused and coupled into single-mode fiber. Nevertheless, recently advanced LED sources have been developed that will allow moderate optical power levels to be launched into single-mode fiber (see Chapter 7). However, to date the LED has been utilized primarily as a multimode source giving acceptable coupling efficiencies into multimode fiber. Moreover, in this capacity the LED remains the major multimode source which is extensively used for increasingly wider bandwidth, longer-haul applications. Therefore at present the LED is chosen for many applications using multimode fibers and the injection laser diode (ILD) tends to find more use as a single-mode device in single-mode fiber systems. Although other laser types (e.g. Nd : YAG and glass fiber laser, Section 6.9), as well as the injection laser, may eventually find significant use in optical fiber communications, this chapter and the following one will deal primarily with major structures and configurations of semiconductor sources (ILD and LED), taking into account recent developments and possible future advances.

We begin by describing in Section 6.2 the basic principles of laser operation which may be applied to all laser types. Immediately following, in Section 6.3, is a discussion of optical emission from semiconductors in which we concentrate on the fundamental operating principles, the structure and the materials for the semiconductor laser. Aspects concerning practical semiconductor lasers are then considered in Section 6.4 prior to a more specific discussion of the structure and operation of some common injection laser types in Section 6.5. Following, in Section 6.6, the major single frequency injection laser structures which provide single-mode operation, primarily in the longer wavelength region (1.1 to 1.6 μm), are then described. In Section 6.7 we consider the operating characteristics which are common to all injection laser types, before a short discussion of injection laser to optical fiber coupling is presented in Section 6.8. Major nonsemiconductor laser

devices which have found use in optical fiber communications (the neodymium-doped yttrium–aluminium–garnet (Nd : YAG) laser and the glass fiber laser) are then outlined in Section 6.9. This is followed in Section 6.10 with a discussion of advanced linewidth narrowed and wavelength tunable laser types. Finally, in Section 6.11, developments in laser sources for transmission in the mid-infrared wavelength region (2 to 5 μm) are briefly considered to give an insight into this potentially important future area.

6.2 Basic concepts

To gain an understanding of the light-generating mechanisms within the major optical sources used in optical fiber communications it is necessary to consider both the fundamental atomic concepts and the device structure. In this context the requirements for the laser source are far more stringent than those for the LED. Unlike the LED, strictly speaking, the laser is a device which amplifies light. Hence the derivation of the term LASER as an acronym for Light Amplification by Stimulated Emission of Radiation. Lasers, however, are seldom used as amplifiers since there are practical difficulties in relation to the achievement of high gain whilst avoiding oscillation from the required energy feedback. Thus the practical realization of the laser is as an optical oscillator. The operation of the device may be described by the formation of an electromagnetic standing wave within a cavity (or optical resonator) which provides an output of monochromatic highly coherent radiation. By contrast the LED provides optical emission without an inherent gain mechanism. This results in incoherent light output.

In this section we elaborate on the basic principles which govern the operation of both these optical sources. It is clear, however, that the operation of the laser must be discussed in some detail in order to provide an appreciation of the way it functions as an optical source. Hence we concentrate first on the general principles of laser action.

6.2.1 Absorption and emission of radiation

The interaction of light with matter takes place in discrete packets of energy or quanta, called photons. Furthermore, the quantum theory suggests that atoms exist only in certain discrete energy states such that absorption and emission of light causes them to make a transition from one discrete energy state to another. The frequency of the absorbed or emitted radiation f is related to the difference in energy E between the higher energy state E_2 and the lower energy state E_1 by the expression:

$$E = E_2 - E_1 = hf \tag{6.1}$$

where $h = 6.626 \times 10^{-34}$ J s is Planck's constant. These discrete energy states for the atom may be considered to correspond to electrons occurring in particular energy

levels relative to the nucleus. Hence, different energy states for the atom correspond to different electron configurations, and a single electron transition between two energy levels within the atom will provide a change in energy suitable for the absorption or emission of a photon. It must be noted, however, that modern quantum theory [Ref. 1] gives a probabilistic description which specifies the energy levels in which electrons are most likely to be found. Nevertheless, the concept of stable atomic energy states and electron transitions between energy levels is still valid.

Figure 6.1(a) illustrates a two energy state or level atomic system where an atom is initially in the lower energy state E_1. When a photon with energy $(E_2 - E_1)$ is incident on the atom it may be excited into the higher energy state E_2 through absorption of the photon. This process is sometimes referred to as stimulated absorption. Alternatively, when the atom is initially in the higher energy state E_2 it can make a transition to the lower energy state E_1 providing the emission of a photon at a frequency corresponding to Eq. (6.1). This emission process can occur

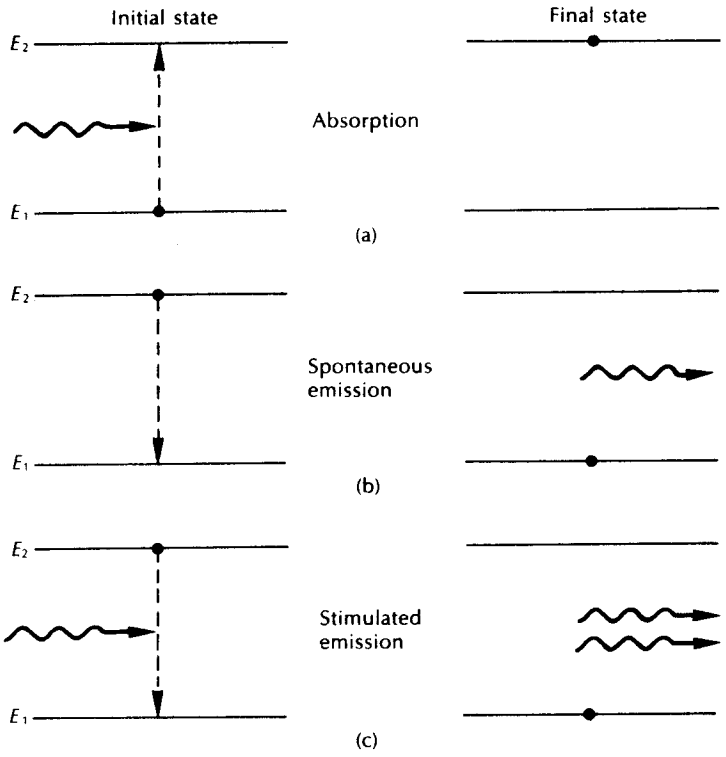


Figure 6.1 Energy state diagram showing: (a) absorption; (b) spontaneous emission; (c) stimulated emission. The black dot indicates the state of the atom before and after a transition takes place.

in two ways:

(a) by spontaneous emission in which the atom returns to the lower energy state in an entirely random manner;
(b) by stimulated emission when a photon having an energy equal to the energy difference between the two states ($E_2 - E_1$) interacts with the atom in the upper energy state causing it to return to the lower state with the creation of a second photon.

These two emission processes are illustrated in Figure 6.1(b) and (c) respectively. The random nature of the spontaneous emission process where light is emitted by electronic transitions from a large number of atoms gives incoherent radiation. A similar emission process in semiconductors provides the basic mechanism for light generation within the LED (see Section 6.3.2).

It is the stimulated emission process, however, which gives the laser its special properties as an optical source. Firstly, the photon produced by stimulated emission is generally* of an identical energy to the one which caused it and hence the light associated with them is of the same frequency. Secondly, the light associated with the stimulating and stimulated photon is in phase and has the same polarization. Therefore, in contrast to spontaneous emission, coherent radiation is obtained. Furthermore, this means that when an atom is stimulated to emit light energy by an incident wave, the liberated energy can add to the wave in a constructive manner, providing amplification.

6.2.2 The Einstein relations

Prior to discussion of laser action in semiconductors it is useful to consider optical amplification in the two level atomic system shown in Figure 6.1. In 1917 Einstein [Ref. 2] demonstrated that the rates of the three transition processes of absorption, spontaneous emission and stimulated emission were related mathematically. He achieved this by considering the atomic system to be in thermal equilibrium such that the rate of the upward transitions must equal the rate of the downward transitions. The population of the two energy levels of such a system are described by Boltzmann statistics which give:

$$\frac{N_1}{N_2} = \frac{g_1 \exp(-E_1/KT)}{g_2 \exp(-E_2/KT)} = \frac{g_1}{g_2} \exp(E_2 - E_1/KT)$$

$$= \frac{g_1}{g_2} \exp(hf/KT) \tag{6.2}$$

where N_1 and N_2 represent the density of atoms in energy levels E_1 and E_2,

* A photon with energy hf will not necessarily always stimulate another photon with energy hf. Photons may be stimulated over a small range of energies around hf providing an emission which has a finite frequency or wavelength spread (linewidth).

respectively, with g_1 and g_2 being the corresponding degeneracies* of the levels, K is Boltzmann's constant and T is the absolute temperature.

As the density of atoms in the lower or ground energy state E_1 is N_1, the rate of upward transition or absorption is proportional to both N_1 and the spectral density ρ_f of the radiation energy at the transition frequency f. Hence, the upward transition rate R_{12} (indicating an electron transition from level 1 to level 2) may be written as:

$$R_{12} = N_1 \rho_f B_{12} \tag{6.3}$$

where the constant of proportionality B_{12} is known as the Einstein coefficient of absorption.

By contrast atoms in the higher or excited energy state can undergo electron transitions from level 2 to level 1 either spontaneously or through stimulation by the radiation field. For spontaneous emission the average time an electron exists in the excited state before a transition occurs is known as the spontaneous lifetime τ_{21}. If the density of atoms within the system with energy E_2, is N_2, then the spontaneous emission rate is given by the product of N_2 and $1/\tau_2$. This may be written as $N_2 A_{21}$ where A_{21}, the Einstein coefficient of spontaneous emission, is equal to the reciprocal of the spontaneous lifetime.

The rate of stimulated downward transition of an electron from level 2 to level 1 may be obtained in a similar manner to the rate of stimulated upward transition. Hence the rate of stimulated emission is given by $N_2 \rho_f B_{21}$, where B_{21} is the Einstein coefficient of stimulated emission. The total transition rates from level 2 to level 1, R_{21}, is the sum of the spontaneous and stimulated contributions. Hence:

$$R_{21} = N_2 A_{21} + N_2 \rho_f B_{21} \tag{6.4}$$

For a system in thermal equilibrium, the upward and downward transition rates must be equal and therefore $R_{12} = R_{21}$, or

$$N_1 \rho_f B_{12} = N_2 A_{21} + N_2 \rho_f B_{21} \tag{6.5}$$

It follows that:

$$\rho_f = \frac{N_2 A_{21}}{N_1 B_{12} - N_2 B_{21}}$$

and

$$\rho_f = \frac{A_{21}/B_{21}}{(B_{12} N_1 / B_{21} N_2) - 1} \tag{6.6}$$

* In many cases the atom has several sublevels of equal energy within an energy level which is then said to be degenerate. The degeneracy parameters g_1 and g_2 indicate the number of sublevels within the energy levels E_1 and E_2 respectively. If the system is not degenerate, then g_1 and g_2 may be set to unity [Ref. 1].

Substituting Eq. (6.2) into Eq. (6.6) gives

$$\rho_f = \frac{A_{21}/B_{21}}{[(g_1B_{12}/g_2B_{21})\exp(hf/KT)] - 1} \tag{6.7}$$

However, since the atomic system under consideration is in thermal equilibrium it produces a radiation density which is identical to black body radiation. Planck showed that the radiation spectral density for a black body radiating within a frequency range f to $f + \mathrm{d}f$ is given by [Ref.3]:

$$\rho_f = \frac{8\pi hf^3}{c^3}\left(\frac{1}{\exp(hf/KT) - 1}\right) \tag{6.8}$$

Comparing Eq. (6.8) with Eq. (6.7) we obtain the Einstein relations:

$$B_{12} = \left(\frac{g_2}{g_1}\right)B_{21} \tag{6.9}$$

and

$$\frac{A_{21}}{B_{21}} = \frac{8\pi hf^3}{c^3} \tag{6.10}$$

It may be observed from Eq. (6.9) that when the degeneracies of the two levels are equal ($g_1 = g_2$), then the probabilities of absorption and stimulated emission are equal. Furthermore, the ratio of the stimulated emission rate to the spontaneous emission rate is given by:

$$\frac{\textit{Stimulated emission rate}}{\textit{Spontaneous emission rate}} = \frac{B_{21}\rho_f}{A_{21}} = \frac{1}{\exp(hf/KT) - 1} \tag{6.11}$$

Example 6.1

Calculate the ratio of the stimulated emission rate to the spontaneous emission rate for an incandescent lamp operating at a temperature of 1000 K. It may be assumed that the average operating wavelength is 0.5 μm.

Solution: The average operating frequency is given by:

$$f = \frac{c}{\lambda} = \frac{2.998 \times 10^8}{0.5 \times 10^{-6}} \simeq 6.0 \times 10^{14}\ \text{Hz}$$

Using Eq. (6.11) the ratio is:

$$\frac{\textit{Stimulated emission rate}}{\textit{Spontaneous emission rate}} = \frac{1}{\exp\left(\dfrac{6.626 \times 10^{-34} \times 6 \times 10^{14}}{1.381 \times 10^{-23} \times 1000}\right)}$$

$$= \exp(-28.8)$$

$$= 3.1 \times 10^{-13}$$

The result obtained in Example 6.1 indicates that for systems in thermal equilibrium spontaneous emission is by far the dominant mechanism. Furthermore, it illustrates that the radiation emitted from ordinary optical sources in the visible spectrum occurs in a random manner, proving these sources are incoherent.

It is apparent that in order to produce a coherent optical source and amplification of a light beam the rate of stimulated emission must be increased far above the level indicated by Example 6.1. From consideration of Eq. (6.5) it may be noted that for stimulated emission to dominate over absorption and spontaneous emission in a two level system, both the radiation density and the population density of the upper energy level N_2 must be increased in relation to the population density of the lower energy level N_1.

6.2.3 Population inversion

Under the conditions of thermal equilibrium given by the Boltzmann distribution (Eq. (6.2)) the lower energy level E_1 of the two level atomic system contains more atoms than the upper energy level E_2. This situation, which is normal for structures at room temperature, is illustrated in Figure 6.2(a). However, to achieve optical amplification it is necessary to create a nonequilibrium distribution of atoms such that the population of the upper energy level is greater than that of the lower energy level (i.e. $N_2 > N_1$). This condition, which is known as population inversion, is illustrated in Figure 6.2(b).

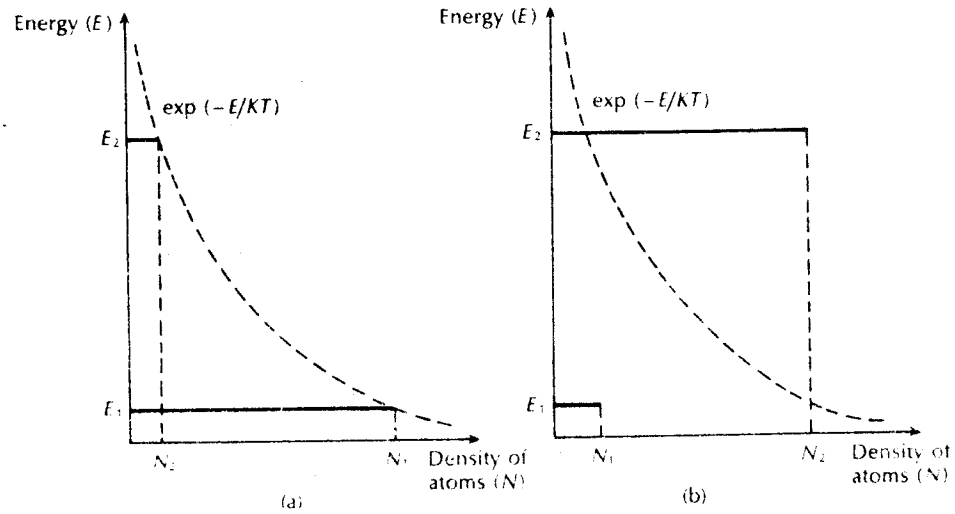


Figure 6.2 Populations in a two energy level system: (a) Boltzmann distribution for a system in thermal equilibrium; (b) a nonequilibrium distribution showing population inversion.

In order to achieve population inversion it is necessary to excite atoms into the upper energy level E_2 and hence obtain a nonequilibrium distribution. This process is achieved using an external energy source and is referred to as 'pumping'. A common method used for pumping involves the application of intense radiation (e.g. from an optical flash tube or high frequency radio field). In the former case atoms are excited into the higher energy state through stimulated absorption. However, the two level system discussed above does not lend itself to suitable population inversion. Referring to Eq. (6.9), when the two levels are equally degenerate (or not degenerate), then $B_{12} = B_{21}$. Thus the probabilities of absorption and stimulated emission are equal, providing at best equal populations in the two levels.

Population inversion, however, may be obtained in systems with three or four energy levels. The energy level diagrams for two such systems, which correspond to two nonsemiconductor lasers, are illustrated in Figure 6.3. To aid attainment of population inversion both systems display a central metastable state in which the atoms spend an unusually long time. It is from this metastable level that the stimulated emission or lasing takes place. The three level system (Figure 6.3(a)) consists of a ground level E_0, a metastable level E_1 and a third level above the metastable level E_2. Initially, the atomic distribution will follow Boltzmann's law. However, with suitable pumping the electrons in some of the atoms may be excited from the ground state into the higher level E_2. Since E_2 is a normal level the electrons will rapidly decay by nonradiative processes to either E_1 or directly to E_0. Hence empty states will always be provided in E_2. The metastable level E_1 exhibits

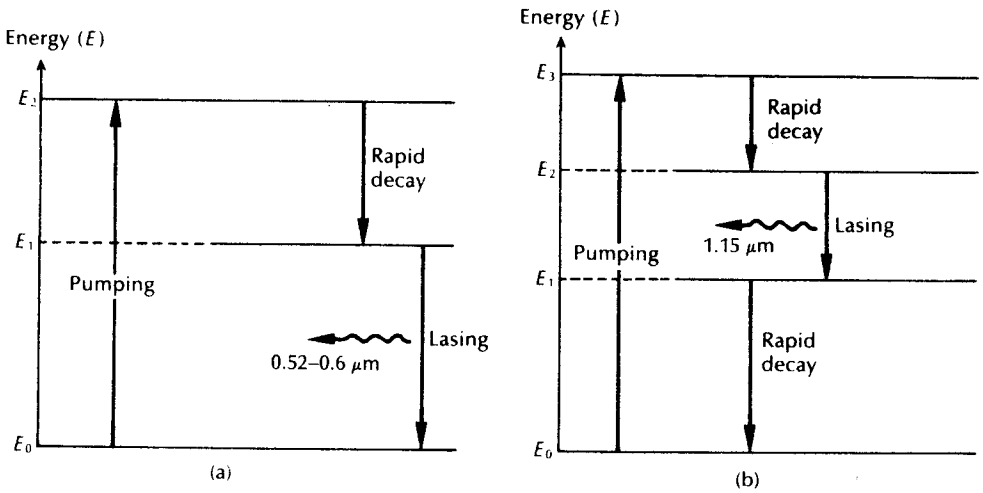


Figure 6.3 Energy level diagrams showing population inversion and lasing for two nonsemiconductor lasers: (a) three level system – ruby (crystal) laser; (b) four level system –.He–Ne (gas) laser.

a much longer lifetime than E_2 which allows a large number of atoms to accumulate at E_1. Over a period the density of atoms in the metastable state N_1 increases above those in the ground state N_0 and a population inversion is obtained between these two levels. Stimulated emission and hence lasing can then occur creating radiative electron transitions between levels E_1 and E_0. A drawback with the three level system such as the ruby laser is that it generally requires very high pump powers because the terminal state of the laser transition is the ground state. Hence more than half the ground state atoms must be pumped into the metastable state to achieve population inversion.

By contrast, a four level system such as the He–Ne laser illustrated in Figure 6.3(b) is characterized by much lower pumping requirements. In this case the pumping excites the atoms from the ground state into energy level E_3 and they decay rapidly to the metastable level E_2. However, since the populations of E_3 and E_1 remain essentially unchanged a small increase in the number of atoms in energy level E_2 creates population inversion, and lasing takes place between this level and level E_1.

6.2.4 Optical feedback and laser oscillation

Light amplification in the laser occurs when a photon colliding with an atom in the excited energy state causes the stimulated emission of a second photon and then both these photons release two more. Continuation of this process effectively creates avalanche multiplication, and when the electromagnetic waves associated with these photons are in phase, amplified coherent emission is obtained. To achieve this laser action it is necessary to contain photons within the laser medium and maintain the conditions for coherence. This is accomplished by placing or forming mirrors (plane or curved) at either end of the amplifying medium, as illustrated in Figure 6.4. The optical cavity formed is more analogous to an oscillator than an amplifier as it provides positive feedback of the photons by reflection at the mirrors at either end of the cavity. Hence the optical signal is fed back many times whilst receiving amplification as it passes through the medium. The structure therefore acts as a Fabry–Perot resonator. Although the amplification of the signal from a single pass through the medium is quite small, after multiple passes the net gain can be large. Furthermore, if one mirror is made partially transmitting, useful radiation may escape from the cavity.

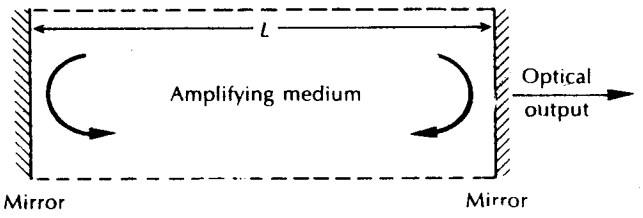


Figure 6.4 The basic laser structure incorporating plane mirrors

A stable output is obtained at saturation when the optical gain is exactly matched by the losses incurred in the amplifying medium. The major losses result from factors such as absorption and scattering in the amplifying medium, absorption, scattering and diffraction at the mirrors and nonuseful transmission through the mirrors.

Oscillations occur in the laser cavity over a small range of frequencies where the cavity gain is sufficient to overcome the above losses. Hence the device is not a perfectly monochromatic source but emits over a narrow spectral band. The central frequency of this spectral band is determined by the mean energy level difference of the stimulated emission transition. Other oscillation frequencies within the spectral band result from frequency variations due to the thermal motion of atoms within the amplifying medium (known as Doppler broadening*) and by atomic collisions†. Hence the amplification within the laser medium results in a broadened laser transition or gain curve over a finite spectral width, as illustrated in Figure 6.5. The spectral emission from the device therefore lies within the frequency range dictated by this gain curve.

Since the structure forms a resonant cavity, when sufficient population inversion exists in the amplifying medium the radiation builds up and becomes established as standing waves between the mirrors. These standing waves exist only at frequencies for which the distance between the mirrors is an integral number of half wavelengths. Thus when the optical spacing between the mirrors is L the resonance condition along the axis of the cavity is given by [Ref. 4]:

$$L = \frac{\lambda q}{2n} \tag{6.12}$$

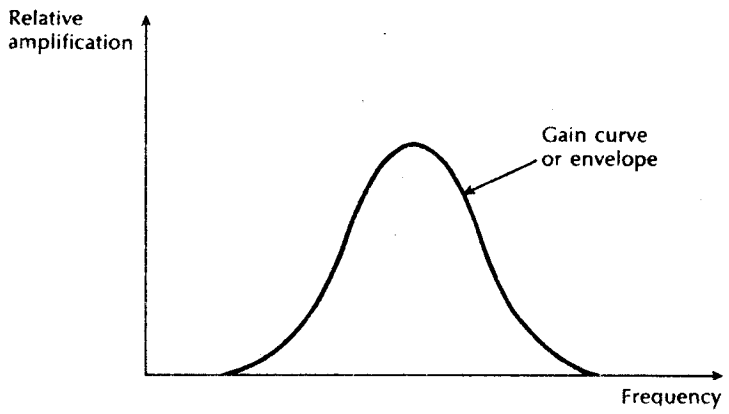


Figure 6.5 The relative amplification in the laser amplifying medium showing the broadened laser transition line or gain curve.

* Doppler broadening is referred to as an inhomogeneous broadening mechanism since individual groups of atoms in the collection have different apparent resonance frequencies.

† Atomic collisions provide homogeneous broadening as every atom in the collection has the same resonant frequency and spectral spread.

where λ is the emission wavelength, n is the refractive index of the amplifying medium and q is an integer. Alternatively, discrete emission frequencies f are defined by

$$f = \frac{qc}{2nL} \tag{6.13}$$

where c is the velocity of light. The different frequencies of oscillation within the laser cavity are determined by the various integer values of q and each constitutes a resonance or mode. Since Eqs. (6.12) and (6.13) apply for the case when L is along the longitudinal axis of the structure (Figure 6.4) the frequencies given by Eq. (6.13) are known as the longitudinal or axial modes. Furthermore, from Eq. (6.13) it may be observed that these modes are separated by a frequency interval δf where:

$$\delta f = \frac{c}{2nL} \tag{6.14}$$

The mode separation in terms of the free space wavelength, assuming $\delta f \ll f$ and as $f = c/\lambda$, is given by:

$$\delta\lambda = \frac{\lambda \delta f}{f} = \frac{\lambda^2}{c}\,\delta f \tag{6.15}$$

Hence substituting for δf from Eq. (6.14) gives:

$$\delta\lambda = \frac{\lambda^2}{2nL} \tag{6.16}$$

In addition it should be noted that Eq. (6.15) can be used to determine the device spectral linewidth as a function of wavelength when it is quoted in Hertz, or vice versa (see problem 6.4).

Example 6.2

A ruby laser contains a crystal length 4 cm with a refractive index of 1.78. The peak emission wavelength from the device is 0.55 μm. Determine the number of longitudinal modes and their frequency separation.

Solution: The number of longitudinal modes supported within the structure may be obtained from Eq. (6.12) where:

$$q = \frac{2nL}{\lambda} = \frac{2 \times 1.78 \times 0.04}{0.55 \times 10^{-6}} = 2.6 \times 10^5$$

Using Eq. (6.14) the frequency separation of the modes is:

$$\delta f = \frac{2.998 \times 10^8}{2 \times 1.78 \times 0.04} = 2.1 \text{ GHz}$$

Although the result of Example 6.2 indicates that a large number of modes may be generated within the laser cavity, the spectral output from the device is defined by the gain curve. Hence the laser emission will only include the longitudinal modes contained within the spectral width of the gain curve. This situation is illustrated in Figure 6.6 where several modes are shown to be present in the laser output. Such a device is said to be multimode.

Laser oscillation may also occur in a direction which is transverse to the axis of the cavity. This gives rise to resonant modes which are transverse to the direction of propagation. These transverse electromagnetic modes are designated in a similar manner to transverse modes in waveguides (Section 2.3.2) by TEM_{lm} where the integers l and m indicate the number of transverse modes (see Figure 6.7). Unlike the longitudinal modes which contribute only a single spot of light to the laser output, transverse modes may give rise to a pattern of spots at the output. This may be observed from the low order transverse mode patterns shown in Figure 6.7 on which the direction of the electric field is also indicated. In the case of the TEM_{00} mode all parts of the propagating wavefront are in phase. This is not so, however, with higher order modes (TEM_{10}, TEM_{11}, etc.) where phase reversals produce the

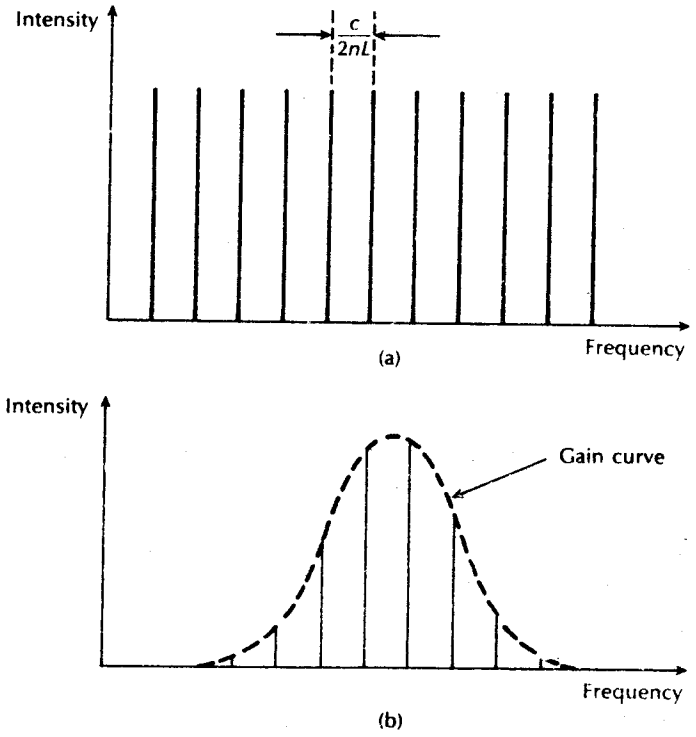


Figure 6.6 (a) The modes in the laser cavity. (b) The longitudinal modes in the laser output.

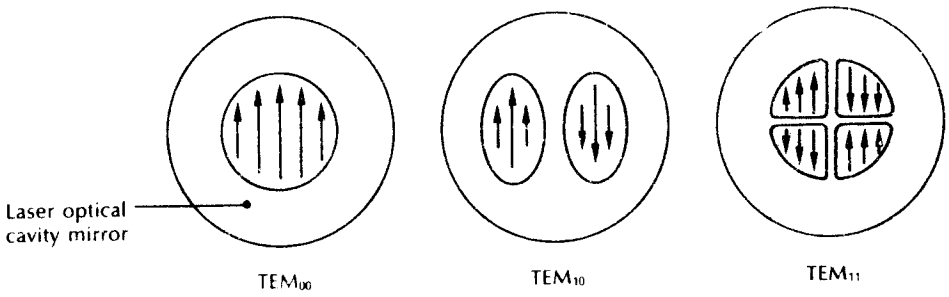


Figure 6.7 The lower order transverse modes of a laser.

various mode patterns. Thus the greatest degree of coherence, together with the highest level of spectral purity, may be obtained from a laser which operates in only the TEM_{00} mode. Higher order transverse modes only occur when the width of the cavity is sufficient for them to oscillate. Consequently, they may be eliminated by suitable narrowing of the laser cavity.

6.2.5 Threshold condition for laser oscillation

It has been indicated that steady state conditions for laser oscillation are achieved when the gain in the amplifying medium exactly balances the total losses.* Hence, although population inversion between the energy levels providing the laser transition is necessary for oscillation to be established, it is not alone sufficient for lasing to occur. In addition a minimum or threshold gain within the amplifying medium must be attained such that laser oscillations are initiated and sustained. This threshold gain may be determined by considering the change in energy of a light beam as it passes through the amplifying medium. For simplicity, all the losses except those due to transmission through the mirrors may be included in a single loss coefficient per unit length $\bar{\alpha}$ cm^{-1}. Again we assume the amplifying medium occupies a length L completely filling the region between the two mirrors which have reflectivities r_1 and r_2. On each round trip the beam passes through the medium twice. Hence the fractional loss incurred by the light beam is:

$$\textit{Fractional loss} = r_1 r_2 \exp(-2\bar{\alpha}L) \tag{6.17}$$

Furthermore, it is found that the increase in beam intensity resulting from stimulated emission is exponential [Ref. 4]. Therefore if the gain coefficient per unit length produced by stimulated emission is $\bar{g}$ cm^{-1}, the fractional round trip gain is given by

$$\textit{Fractional gain} = \exp(2\bar{g}L) \tag{6.18}$$

* This applies to a CW laser which gives a continuous output, rather than pulsed devices for which slightly different conditions exist. For oscillation to commence the fractional gain and loss must be matched.

Hence

$$\exp(2gL) \times r_1 r_2 \exp(-2\bar{\alpha}L) = 1$$

and

$$r_1 r_2 \exp 2(\bar{g} - \bar{\alpha})L = 1 \tag{6.19}$$

The threshold gain per unit lengtn may be obtained by rearranging the above expression to give:

$$\bar{g}_{th} = \bar{\alpha} + \frac{1}{2L} \ln \frac{1}{r_1 r_2} \tag{6.20}$$

The second term on the right hand side of Eq. (6.20) represents the transmission loss through the mirrors.*

For laser action to be easily achieved it is clear that a high threshold gain per unit length is required in order to balance the losses from the cavity. However, it must be noted that the parameters displayed in Eq. (6.20) are totally dependent on the laser type.

6.3 Optical emission from semiconductors

6.3.1 The *p–n* junction

To allow consideration of semiconductor optical sources it is necessary to review some of the properties of semiconductor materials, especially with regard to the *p–n* junction. A perfect semiconductor crystal containing no impurities or lattice defects is said to be intrinsic. The energy band structure [Ref. 1] of an intrinsic semiconductor is illustrated in Figure 6.8(a) which shows the valence and conduction bands separated by a forbidden energy gap or bandgap E_g, the width of which varies for different semiconductor materials.

Figure 6.8(a) shows the situation in the semiconductor at a temperature above absolute zero where thermal excitation raises some electrons from the valence band into the conduction band, leaving empty hole states in the valence band. These thermally excited electrons in the conduction band and the holes left in the valence band allow conduction through the material, and are called carriers.

For a semiconductor in thermal equilibrium the energy level occupation is described by the Fermi–Dirac distribution function (rather than the Boltzmann). Consequently, the probability $P(E)$ that an electron gains sufficient thermal energy at an absolute temperature T such that it will be found occupying a particular energy level E, is given by the Fermi–Dirac distribution [Ref. 1]:

$$P(E) = \frac{1}{1 + \exp(E - E_F)/KT} \tag{6.21}$$

* This term is sometimes expressed in the form $1/L \ln 1/r$, where r, the reflectivity of the mirrored ends, is equal to $\sqrt{(r_1 r_2)}$

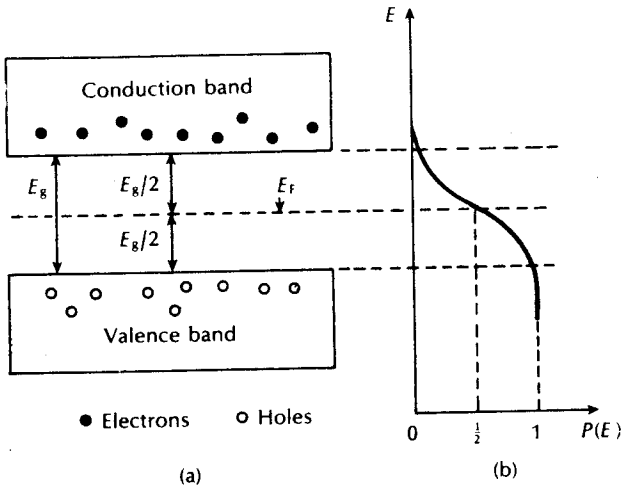


Figure 6.8 (a) The energy band structure of an intrinsic semiconductor at a temperature above absolute zero, showing an equal number of electrons and holes in the conduction band and the valence band respectively. (b) The Fermi–Dirac probability distribution corresponding to (a).

where K is Boltzmann's constant and E_F is known as the Fermi energy or Fermi level. The Fermi level is only a mathematical parameter but it gives an indication of the distribution of carriers within the material. This is shown in Figure 6.8(b) for the intrinsic semiconductor where the Fermi level is at the centre of the bandgap, indicating that there is a small probability of electrons occupying energy levels at the bottom of the conduction band and a corresponding number of holes occupying energy levels at the top of the valence band.

To create an extrinsic semiconductor the material is doped with impurity atoms which either create more free electrons (donor impurity) or holes (acceptor impurity). These two situations are shown in Figure 6.9 where the donor impurities form energy levels just below the conduction band whilst acceptor impurities form energy levels just above the valence band.

When donor impurities are added, thermally excited electrons from the donor levels are raised into the conduction band to create an excess of negative charge carriers and the semiconductor is said to be n type, with the majority carriers being electrons. The Fermi level corresponding to this carrier distribution is raised to a position above the centre of the bandgap, as illustrated in Figure 6.9(a). When acceptor impurities are added, as shown in Figure 6.9(b), thermally excited electrons are raised from the valence band to the acceptor impurity levels leaving an excess of positive charge carriers in the valence band and creating a p type semiconductor where the majority carriers are holes. In this case Fermi level is lowered below the centre of the bandgap.

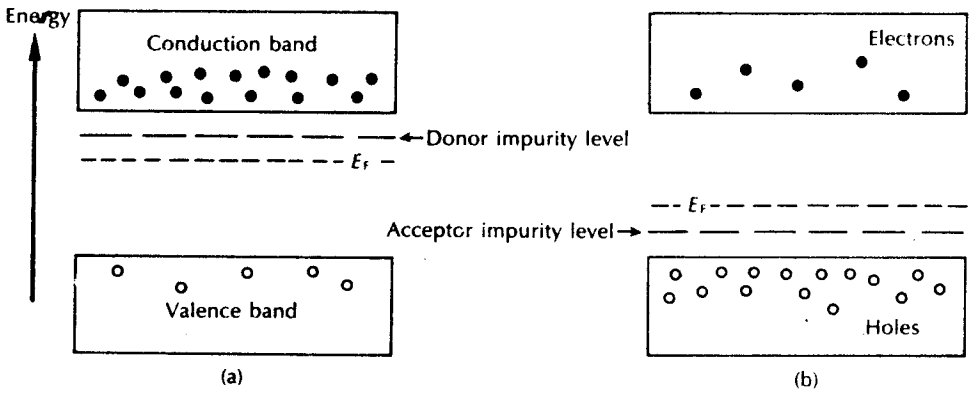


Figure 6.9 Energy band diagrams: (a) *n* type semiconductor; (b) *p* type semiconductor.

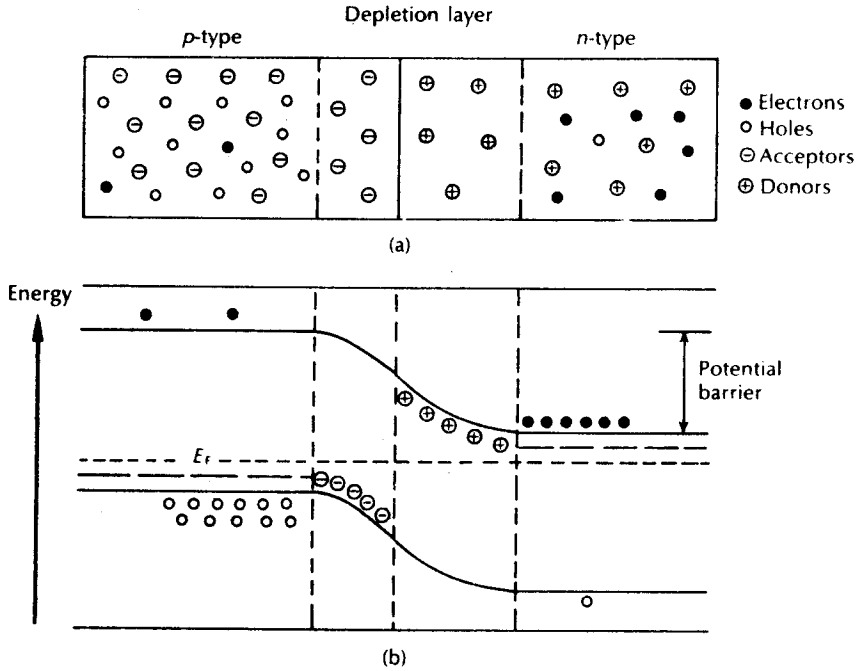


Figure 6.10 (a) The impurities and charge carriers at a *p–n* junction. (b) The energy band diagram corresponding to (a).

The p–n junction diode is formed by creating adjoining p and n type semiconductor layers in a single crystal, as shown in Figure 6.10(a). A thin depletion region or layer is formed at the junction through carrier recombination which effectively leaves it free of mobile charge carriers (both electrons and holes). This establishes a potential barrier between the p and n type regions which restricts the interdiffusion of majority carriers from their respective regions, as illustrated in Figure 6.10(b). In the absence of an externally applied voltage no current flows as the potential barrier prevents the net flow of carriers from one region to another. When the junction is in this equilibrium state the Fermi level for the p and n type semiconductor is the same as shown Figure 6.10(b).

The width of the depletion region and thus the magnitude of the potential barrier is dependent upon the carrier concentrations (doping) in the p and n type regions, and any external applied voltage. When an external positive voltage is applied to the p type region with respect to the n type, both the depletion region width and the resulting potential barrier are reduced and the diode is said to be forward biased. Electrons from the n type region and holes from the p type region can flow more readily across the junction into the opposite type region. These minority carriers are effectively injected across the junction by the application of the external voltage and form a current flow through the device as they continuously diffuse away from the interface. However, this situation in suitable semiconductor materials allows carrier recombination with the emission of light.

6.3.2 Spontaneous emission

The increased concentration of minority carriers in the opposite type region in the forward biased p–n diode leads to the recombination of carriers across the bandgap. This process is shown in Figure 6.11 for a direct bandgap (see Section

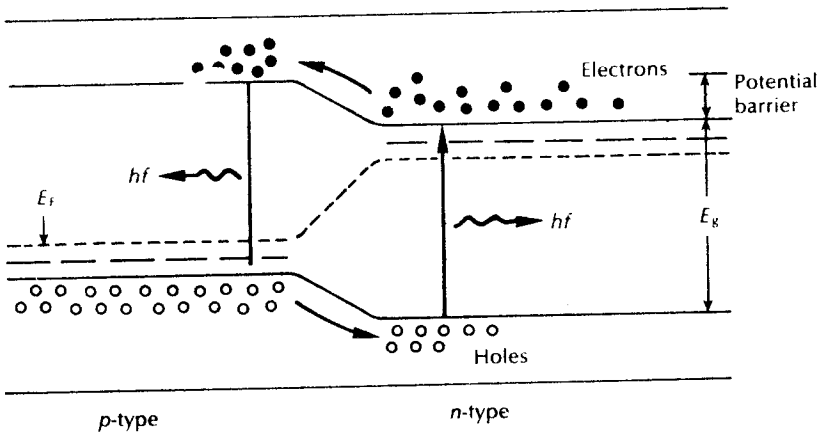


Figure 6.11 The p–n junction with forward bias giving spontaneous emission of photons.

6.3.3) semiconductor material where the normally empty electron states in the conduction band of the *p* type material and the normally empty hole states in the valence band of the *n* type material are populated by injected carriers which recombine across the bandgap. The energy released by this electron–hole recombination is approximately equal to the bandgap energy E_g. Excess carrier population is therefore decreased by recombination which may be radiative or nonradiative.

In nonradiative recombination the energy released is dissipated in the form of lattice vibrations and thus heat. However, in band to band radiative recombination the energy is released with the creation of a photon (see Figure 6.11) with a frequency following Eq. (6.1) where the energy is approximately equal to the bandgap energy E_g and therefore:

$$E_g = hf = \frac{hc}{\lambda} \tag{6.22}$$

where c is the velocity of light in a vacuum and λ is the optical wavelength. Substituting the appropriate values for h and c in Eq. (6.22) and rearranging gives:

$$\lambda = \frac{1.24}{E_g} \tag{6.23}$$

where λ is written in μm and E_g in eV.

This spontaneous emission of light from within the diode structure is known as electroluminescence.* The light is emitted at the site of carrier recombination which

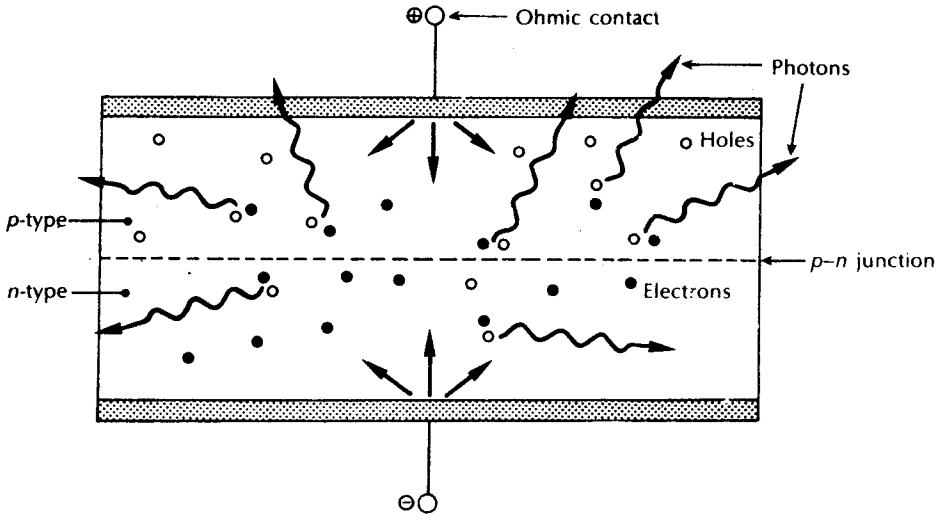


Figure 6.12 An illustration of carrier recombination giving spontaneous emission of light in a *p–n* junction diode.

* The term electroluminescence is used when the optical emission results from the application of an electric field.

is primarily close to the junction, although recombination may take place through hole diode structure as carriers diffuse away from the junction region (see Figure 6.12). However, the amount of radiative recombination and the emission area within the structure is dependent upon the semiconductor materials used and the fabrication of the device.

6.3.3 Carrier recombination

6.3.3.1 Direct and indirect bandgap semiconductors

In order to encourage electroluminescence it is necessary to select an appropriate semiconductor material. The most useful materials for this purpose are direct bandgap semiconductors in which electrons and holes on either side of the forbidden energy gap have the same value of crystal momentum and thus direct recombination is possible. This process is illustrated in Figure 6.13(a) with an energy-momentum diagram for a direct bandgap semiconductor. It may be observed that the energy maximum of the valence band occurs at the same (or very nearly the same) value of electron crystal momentum* as the energy minimum of the conduction band. Hence when electron–hole recombination occurs the

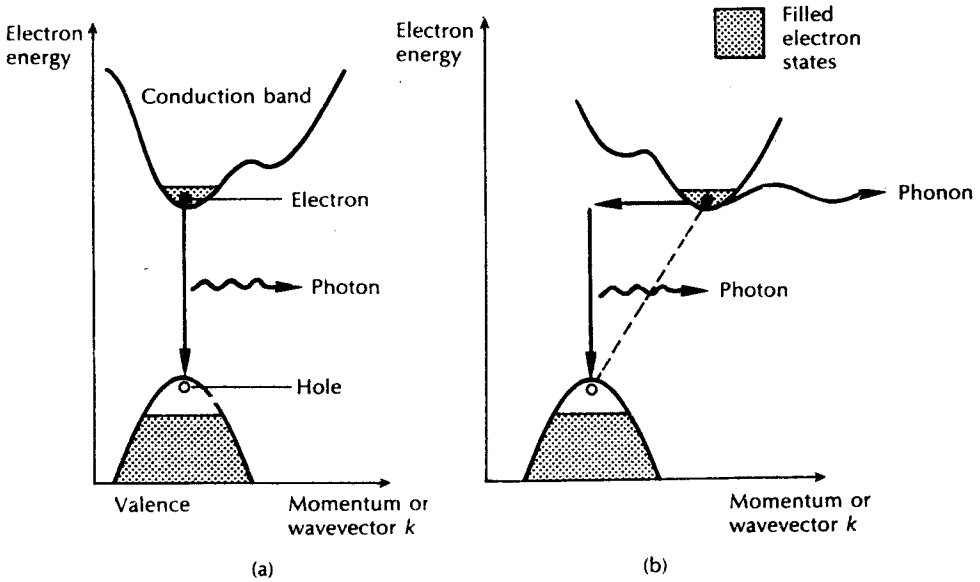


Figure 6.13 Energy–momentum diagrams showing the types of transition: (a) direct bandgap semiconductor; (b) indirect bandgap semiconductor.

* The crystal momentum p is related to the wavevector k for an electron in a crystal by $p = 2\pi\hbar k$, where $\hbar$ is Planck's constant [Ref. 1]. Hence the abscissa of Figure 6.13 is often shown as the electron wavevector rather than momentum.

Table 6.1 Some direct and indirect bandgap semiconductors with calculated recombination coefficients

Semiconductor material	Energy bandgap (eV)	Recombination coefficient B_r ($cm^3 s^{-1}$)
GaAs	Direct: 1.43	7.21×10^{-10}
GaSb	Direct: 0.73	2.39×10^{-10}
InAs	Direct: 0.35	8.5×10^{-11}
InSb	Direct: 0.18	4.58×10^{-11}
Si	Indirect: 1.12	1.79×10^{-15}
Ge	Indirect: 0.67	5.25×10^{-14}
GaP	Indirect: 2.26	5.37×10^{-14}

momentum of the electron remains virtually constant and the energy released, which corresponds to the bandgap energy E_g, may be emitted as light. This direct transition of an electron across the energy gap provides an efficient mechanism for photon emission and the average time the minority carrier remains in a free state before recombination (the minority carrier lifetime) is short (10^{-8} to 10^{-10} s). Some commonly used direct bandgap semiconductor materials are shown in Table 6.1 [Ref. 3].

In indirect bandgap semiconductors, however, the maximum and minimum energies occur at different values of crystal momentum (Figure 6.13(b)). For electron–hole recombination to take place it is essential that the electron loses momentum such that it has a value of momentum corresponding to the maximum energy of the valence band. The conservation of momentum requires the emission or absorption of a third particle, a photon. This three particle recombination process is far less probable than the two particle process exhibited by direct bandgap semiconductors. Hence, the recombination in indirect bandgap semiconductors is relatively slow (10^{-2} to 10^{-4} s). This is reflected by a much longer minority carrier lifetime, together with a greater probability of nonradiative transitions. The competing nonradiative recombination processes which involve lattice defects and impurities (e.g. precipitates of commonly used dopants) become more likely as they allow carrier recombination in a relatively short time in most materials. Thus the indirect bandgap emitters such as silicon and germanium shown in Table 6.1 give insignificant levels of electroluminescence. This disparity is further illustrated in Table 6.1 by the values of the recombination coefficient B_r given for both the direct and indirect bandgap recombination semiconductors shown.

The recombination coefficient is obtained from the measured absorption coefficient of the semiconductor, and for low injected minority carrier density relative to the majority carriers it is related approximately to the radiative minority carrier lifetime* τ_r by [Ref. 4]:

$$\tau_r = [B_r(N + P)]^{-1} \tag{6.24}$$

* The radiative minority carrier lifetime is defined as the average time a minority carrier can exist in a free state before radiative recombination takes place.

where N and P are the respective majority carrier concentrations in the n and p type regions. The significant difference between the recombination coefficients for the direct and indirect bandgap semiconductors shown, underlines the importance of the use of direct bandgap materials for electroluminescent sources. Direct bandgap semiconductor devices in general have a much higher internal quantum efficiency. This is the ratio of the number of radiative recombinations (photons produced within the structure) to the number of injected carriers which is often expressed as a percentage.

Example 6.3

Compare the approximate radiative minority carrier lifetimes in gallium arsenide and silicon when the minority carriers are electrons injected into the p type region which has a hole concentration of 10^{18} cm^{-3}. The injected electron density is small compared with the majority carrier density.

Solution: Equation (6.24) gives the radiative minority carrier lifetime τ_r as

$$\tau_r \simeq [B_r(N + P)]^{-1}$$

In the p type region the hole concentration determines the radiative carrier lifetime as $P \gg N$. Hence,

$$\tau_r \simeq [B_r N]^{-1}$$

Thus for gallium arsenide:

$$\begin{aligned}\tau_r &\simeq [7.21 \times 10^{-10} \times 10^{18}]^{-1} \\ &= 1.39 \times 10^{-9} \\ &= 1.39 \text{ ns}\end{aligned}$$

For silicon

$$\begin{aligned}\tau_r &\simeq [1.79 \times 10^{-15} \times 10^{18}]^{-1} \\ &= 5.58 \times 10^{-4} \\ &= 0.56 \text{ ms}\end{aligned}$$

Thus the direct bandgap gallium arsenide has a radiative carrier lifetime factor of around 2.5×10^{-6} less than the indirect bandgap silicon.

6.3.3.2 Other radiative recombination processes

In the preceding sections, only full bandgap transitions have been considered to give radiative recombination. However, energy levels may be introduced into the bandgap by impurities or lattice defects within the material structure which may greatly increase the electron–hole recombination (effectively reduce the carrier lifetime). The recombination process through such impurity or defect centres may be either radiative or nonradiative. Major radiative recombination processes at 300 K other than band to band transitions are shown in Figure 6.14. These are band

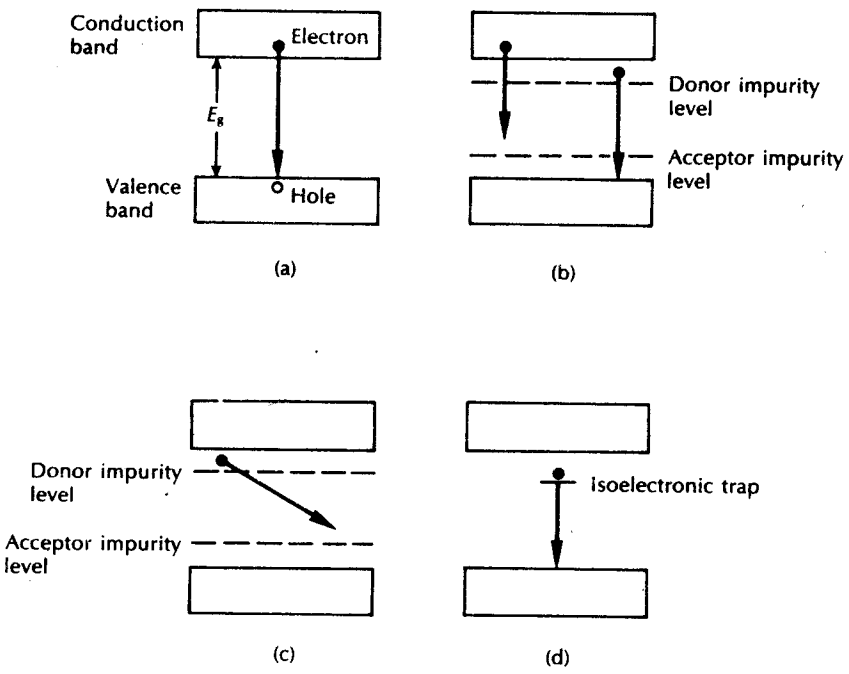


Figure 6.14 Major radiative recombination processes at 300 K (a) conduction to valence band (band to band) transition: (b) conduction band to acceptor impurity, and donor impurity to valence band transition; (c) donor impurity to acceptor impurity transition; (d) recombination from an isoelectronic impurity to the valence band.

to impurity centre or impurity centre to band, donor level to acceptor level and recombination involving isoelectronic impurities.

Hence, an indirect bandgap semiconductor may be made into a more useful electroluminescent material by the addition of impurity centres which will effectively convert it into a direct bandgap material. An example of this is the introduction of nitrogen as an impurity into gallium phosphide. In this case the nitrogen forms an isoelectronic impurity as it has the same number of valence (outer shell) electrons as phosphorus but with a different covalent radius and higher electronegativity [Ref. 1]. The nitrogen impurity centre thus captures an electron and acts as an isoelectronic trap which has a large spread of momentum. This trap then attracts the oppositely charged carrier (a hole) and a direct transition takes place between the impurity centre and the valence band. Hence gallium phosphide may become an efficient light emitter when nitrogen is incorporated. However, such conversion of indirect to direct bandgap transitions is only readily achieved in materials where the direct and indirect bandgaps have a small energy difference. This is the case with gallium phosphide but not with silicon or germanium.

6.3.4 Stimulated emission and lasing

The general concept of stimulated emission via population inversion was indicated in Section 6.2.3. Carrier population inversion is achieved in an intrinsic (undoped) semiconductor by the injection of electrons into the conduction band of the material. This is illustrated in Figure 6.15 where the electron energy and the corresponding filled states are shown. Figure 6.15(a) shows the situation at absolute zero when the conduction band contains no electrons. Electrons injected into the material fill the lower energy states in the conduction band up to the injection energy or the quasi-Fermi level for electrons. Since charge neutrality is conserved within the material an equal density of holes is created in the top of the valence band by the absence of electrons, as shown in Figure 6.15(b) [Ref. 5].

Incident photons with energy E_g but less than the separation energy of the quasi-Fermi levels $E_q = E_{Fc} - E_{Fv}$ cannot be absorbed because the necessary conduction band states are occupied. However, these photons can induce a downward transition of an electron from the filled conduction band states into the empty valence band states, thus stimulating the emission of another photon. The basic condition for stimulated emission is therefore dependent on the quasi-Fermi level separation energy as well as the bandgap energy and may be defined as:

$$E_{Fc} - E_{Fv} > hf > E_g \tag{6.25}$$

However, it must be noted that we have described an ideal situation whereas at normal operating temperatures the distribution of electrons and holes is less well defined but the condition for stimulated emission is largely maintained.

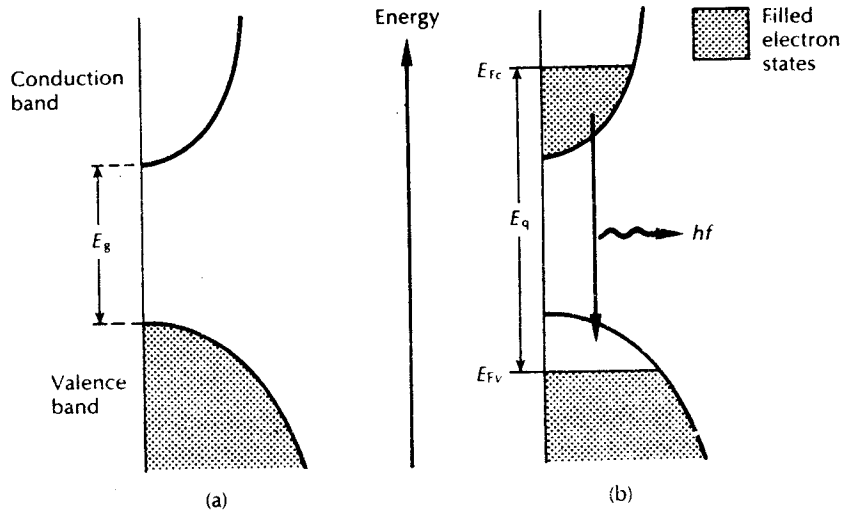


Figure 6.15 The filled electron states for an intrinsic direct bandgap semiconductor at absolute zero [Ref. 5]: (a) in equilibrium; (b) with high carrier injection.

Population inversion may be obtained at a p–n junction by heavy doping (degenerative doping) of both the p and n type material. Heavy p type doping with acceptor impurities causes a lowering of the Fermi level or boundary between the filled and empty states into the valence band. Similarly, degenerative n type doping causes the Fermi level to enter the conduction band of the material. Energy band diagrams of a degenerate p–n junction are shown in Figure 6.16. The position of the Fermi level and the electron occupation (shading) with no applied bias are shown in Figure 6.16(a). Since in this case the junction is in thermal equilibrium, the Fermi energy has the same value throughout the material. Figure 6.16(b) shows the p–n junction when a forward bias nearly equal to the bandgap voltage is applied and hence there is direct conduction. At high injection carrier density* in such a junction there exists an active region near the depletion layer that contains simultaneously degenerate populations of electrons and holes (sometimes termed doubly degenerate). For this region the condition for stimulated emission of Eq.

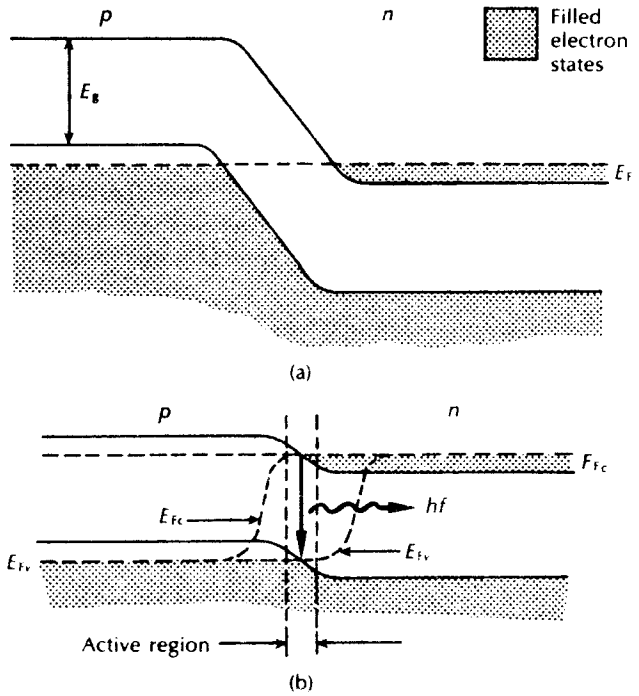


Figure 6.16 The degenerate p–n junction: (a) with no applied bias; (b) with strong forward bias such that the separation of the quasi-Fermi levels is higher than the electron–hole recombination energy hf in the narrow active region. Hence stimulated emission is obtained in this region.

* This may be largely considered to be electrons injected into the p–n region because of their greater mobility

(6.22) is satisfied for electromagnetic radiation of frequency $E_g/h < f < (E_{Fc} - E_{Fv})/h$. Therefore, any radiation of this frequency which is confined to the active region will be amplified. In general, the degenerative doping distinguishes a p–n junction which provides stimulated emission from one which gives only spontaneous emission as in the case of the LED.

Finally, it must be noted that high impurity concentration within a semiconductor causes differences in the energy bands in comparison with an intrinsic semiconductor. These differences are particularly apparent in the degeneratively doped p–n junctions used for semiconductor lasers. For instance at high donor level concentrations in gallium arsenide, the donor impurity levels form a band that merges with the conduction band. These energy states, sometimes referred to as 'bandtail' states [Ref. 7] extend into the forbidden energy gap. The laser transition may take place from one of these states. Furthermore the transitions may terminate on acceptor states which because of their high concentration also extend as a band into the energy gap. In this way the lasing transitions may occur at energies less than the bandgap energy E_g. When transitions of this type dominate, the lasing peak energy is less than the bandgap energy. Hence the effective lasing wavelength can be varied within the electroluminescent semiconductor used to fabricate the junction laser through variation of the impurity concentration. For example, the lasing wavelength of gallium arsenide may be varied between 0.85 and 0.95 μm, although the best performance is usually achieved in the 0.88 to 0.91 μm band (see Problem 6.6).

However, a further requirement of the junction diode is necessary to establish lasing. This involves the provision of optical feedback to give laser oscillation. It may be achieved by the formation of an optical cavity (Fabry–Perot cavity, see Section 6.2.4) within the structure by polishing the end faces of the junction diode to act as mirrors. Each end of the junction is polished or cleaved and the sides are roughened to prevent any unwanted light emission and hence wasted population inversion.

The behaviour of the semiconductor laser can be described by rate equations for electron and photon density in the active layer of the device. These equations assist in providing an understanding of the laser electrical and optical performance characteristics under direct current modulation as well as its potential limitations. The problems associated with high speed direct current modulation are unique to the semiconductor laser whose major application area is that of a source within optical fiber communications.

Although the rate equations may be approached with some rigour [Ref. 8] we adopt a simplified analysis which is valid within certain constraints [Ref. 9]. In particular, the equations represent an average behaviour for the active medium within the laser cavity and they are not applicable when the time period is short compared with the transit time of the optical wave in the laser cavity.* The two rate

* Thus performance characteristics derived from these rate equations become questionable when the time scale is less than 10 ps or the modulation bandwidth is greater than 100 GHz.

equations for electron density n, and photon density ϕ, are:

$$\frac{dn}{dt} = \frac{J}{ed} - \frac{n}{\tau_{sp}} - Cn\phi \qquad (m^{-3}s^{-1}) \tag{6.26}$$

and

$$\frac{d\phi}{dt} = Cn\phi + \delta\frac{n}{\tau_{sp}} - \frac{\phi}{\tau_{ph}} \qquad (m^{-3}s^{-1}) \tag{6.27}$$

where J is the current density, in amperes per square meter, e is the charge on an electron, d is the thickness of the recombination region, τ_{sp} is the spontaneous emission lifetime which is equivalent to τ_{21} in Section 6.2.2, C is a coefficient which incorporates the B coefficients in Section 6.2.2, δ is a small fractional value and τ_{ph} is the photon lifetime.

The rate equations given in Eqs. (6.26) and (6.27) may be balanced by taking into account all the factors which affect the numbers of electrons and holes in the laser structure. Hence, in Eq. (6.26), the first term indicates the increase in the electron concentration in the conduction band as the current flows into the junction diode. The electrons lost from the conduction band by spontaneous and stimulated transitions are provided by the second and third terms respectively. In Eq. (6.27) the first term depicts the stimulated emission as a source of photons. The fraction of photons produced by spontaneous emission which combine to the energy in the lasing mode is given by the second term. This term is often neglected, however, as δ is small. The final term represents the decay in the number of photons resulting from losses in the optical cavity.

Although these rate equations may be used to study both the transient and steady state behaviour of the semiconductor laser, we are particularly concerned with the steady state solutions. The steady state is characterized by the left hand side of Eqs. (6.26) and (6.27) being equal to zero, when n and ϕ have nonzero values. In addition, the fields in the optical cavity which are represented by ϕ must build up from small initial values, and hence $d\phi/dt$ must be positive when ϕ is small. Therefore, setting δ equal to zero in Eq. (6.27), it is clear that for any value of ϕ, $d\phi/dt$ will only be positive when:

$$Cn - \frac{1}{\tau_{ph}} \geqslant 0 \tag{6.28}$$

There is therefore a threshold value of n which satisfies the equality of Eq. (6.28). If n is larger than this threshold value, then ϕ can increase; however, when n is smaller it cannot. From Eq. (6.28) the threshold value for the electron density n_{th} is:

$$n_{th} = \frac{1}{C\tau_{ph}} \qquad (m^{-3}) \tag{6.29}$$

The threshold current written in terms of its current density J_{th}, required to maintain $n = n_{th}$ in the steady state when $\phi = 0$, may be obtained from Eq. (6.26) as:

$$\frac{J_{th}}{ed} = \frac{n_{th}}{\tau_{sp}} \qquad (m^{-3}s^{-1}) \tag{6.30}$$

Hence Eq. (6.30) defines the current required to sustain an excess electron density in the laser when spontaneous emission provides the only decay mechanism. The steady state photon density ϕ_s is provided by substituting Eq. (6.30) in Eq. (6.26) giving:

$$0 = \frac{(J - J_{th})}{ed} - Cn_{th}\phi_s.$$

Rearranging we obtain:

$$\phi_s = \frac{1}{Cn_{th}} \frac{(J - J_{th})}{ed} \qquad (m^{-3}) \tag{6.31}$$

Substituting for Cn_{th} from Eq. (6.29) we can write Eq. (6.31) in the form:

$$\phi_s = \frac{\tau_{ph}}{ed} (J - J_{th}) \qquad (m^{-3}) \tag{6.32}$$

The photon density ϕ_s cannot be a negative quantity as this is meaningless, and for ϕ_s to be greater than zero the current must exceed its threshold value. Moreover, ϕ_s is proportional to the amount by which J exceeds its threshold value. As each photon has energy hf it is possible to determine the optical power density in $W\,m^{-2}$ by assuming that half the photons are travelling in each of two directions.

An idealized optical output power against current characteristic (also called light output against current characteristic) for a semiconductor laser is illustrated in Figure 6.17. The solid line represents the laser characteristic, whereas the dashed line is a plot of Eq. (6.32) showing the current threshold. It may be observed that the device gives little light output in the region below the threshold current which corresponds to spontaneous emission only within the structure. However, after the

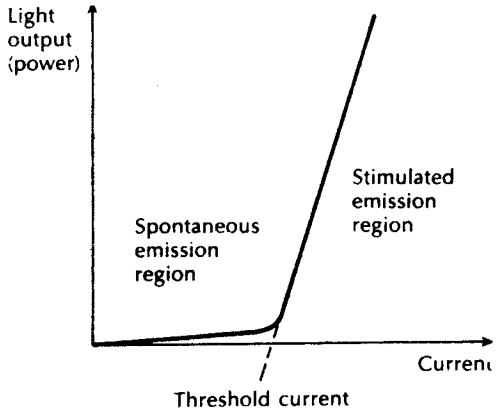


Figure 6.17 The ideal light output against current characteristic for an injection laser.

threshold current density is reached, the light output increases substantially for small increases in current through the device. This corresponds to the region of stimulated emission when the laser is acting as an amplifier of light.

In common with all other laser types a requirement for the initiation and maintenance of laser oscillation is that the optical gain matches the optical losses within the cavity (see Section 6.2.5). For the p–n junction or semiconductor laser this occurs at a particular photon energy within the spectrum of spontaneous emission (usually near the peak wavelength of spontaneous emission). Thus when extremely high currents are passed through the device (i.e. injection levels of around 10^{18} carriers cm^{-3}), spontaneous emission with a wide spectrum (linewidth) becomes lasing (when a current threshold is passed) and the linewidth subsequently narrows.

For strongly confined structures the threshold current density for stimulated emission J_{th} is to a fair approximation [Ref. 4] related to the threshold gain coefficient $\bar{g}_{th}$ for the laser cavity through:

$$\bar{g}_{th} = \bar{\beta} J_{th} \tag{6.33}$$

where the gain factor $\bar{\beta}$ is a constant appropriate to specific devices. Detailed discussion of the more exact relationship is given in Ref. 4.

Substituting for $\bar{g}_{th}$ from Eq. (6.18) and rearranging we obtain:

$$J_{th} = \frac{1}{\bar{\beta}} \left[\bar{\alpha} + \frac{1}{2L} \ln \frac{1}{r_1 r_2} \right] \tag{6.34}$$

Since for the semiconductor laser the mirrors are formed by a dielectric plane and are often uncoated, the mirror reflectivities r_1 and r_2 may be calculated using the Fresnel reflection relationship of Eq. (5.1).

Example 6.4

A GaAs injection laser has an optical cavity of length 250 μm and width 100 μm. At normal operating temperature the gain factor $\bar{\beta}$ is 21×10^{-3} A cm^{-3} and the loss coefficient $\bar{\alpha}$ per cm is 10. Determine the threshold current density and hence the threshold current for the device. It may be assumed that the cleaved mirrors are uncoated and that the current is restricted to the optical cavity. The refractive index of GaAs may be taken as 3.6.

Solution: The reflectivity for normal incidence of a plane wave on the GaAs–air interface may be obtained from Eq. (5.1) where:

$$r_1 = r_2 = r = \left(\frac{n-1}{n+1} \right)^2$$

$$= \left(\frac{3.6-1}{3.6+1} \right)^2 \simeq 0.32$$

The threshold current density may be obtained from Eq. (6.34) where:

$$J_{th} = \frac{1}{\bar{\beta}}\left[\bar{\alpha}\frac{1}{L}\ln\frac{1}{r}\right]$$

$$= \frac{1}{21\times10^{-3}}\left[10+\frac{1}{250\times10^{-4}}\ln\frac{1}{0.32}\right]$$

$$= 2.65\times10^{3}\ \text{A cm}^{-2}$$

The threshold current I_{th} is given by:

$$I_{th} = J_{th}\times\text{area of the optical cavity}$$
$$= 2.65\times10^{3}\times250\times100\times10^{-8}$$
$$= 663\ \text{mA}$$

Therefore the threshold current for this device is 663 mA if the current flow is restricted to the optical cavity.

As the stimulated emission minority carrier lifetime is much shorter (typically 10^{-11} s) than that due to spontaneous emission, further increases in input current above the threshold will result almost entirely in stimulated emission, giving a high internal quantum efficiency (50 to 100%). Also, whereas incoherent spontaneous emission has a linewidth of tens of nanometres, stimulated coherent emission has a linewidth of a nanometre or less.

6.3.5 Heterojunctions

The preceding sections have considered the photoemissive properties of a single p–n junction fabricated from a single crystal semiconductor material. This is known as a homojunction. However, the radiative properties of a junction diode may be improved by the use of heterojunctions. A heterojunction is an interface between two adjoining single crystal semiconductors with different bandgap energies. Devices which are fabricated with heterojunctions are said to have heterostructure.

Heterojunctions are classified into either an isotype (n–n or p–p) or an anisotype (p–n). The isotype heterojunction provides a potential barrier within the structure which is useful for the confinement of minority carriers to a small active region (carrier confinement). It effectively reduces the carrier diffusion length and thus the volume within the structure where radiative recombination may take place. This technique is widely used for the fabrication of injection lasers and high radiance LEDs. Isotype heterojunctions are also extensively used in LEDs to provide a transparent layer close to the active region which substantially reduces the absorption of light emitted from the structure.

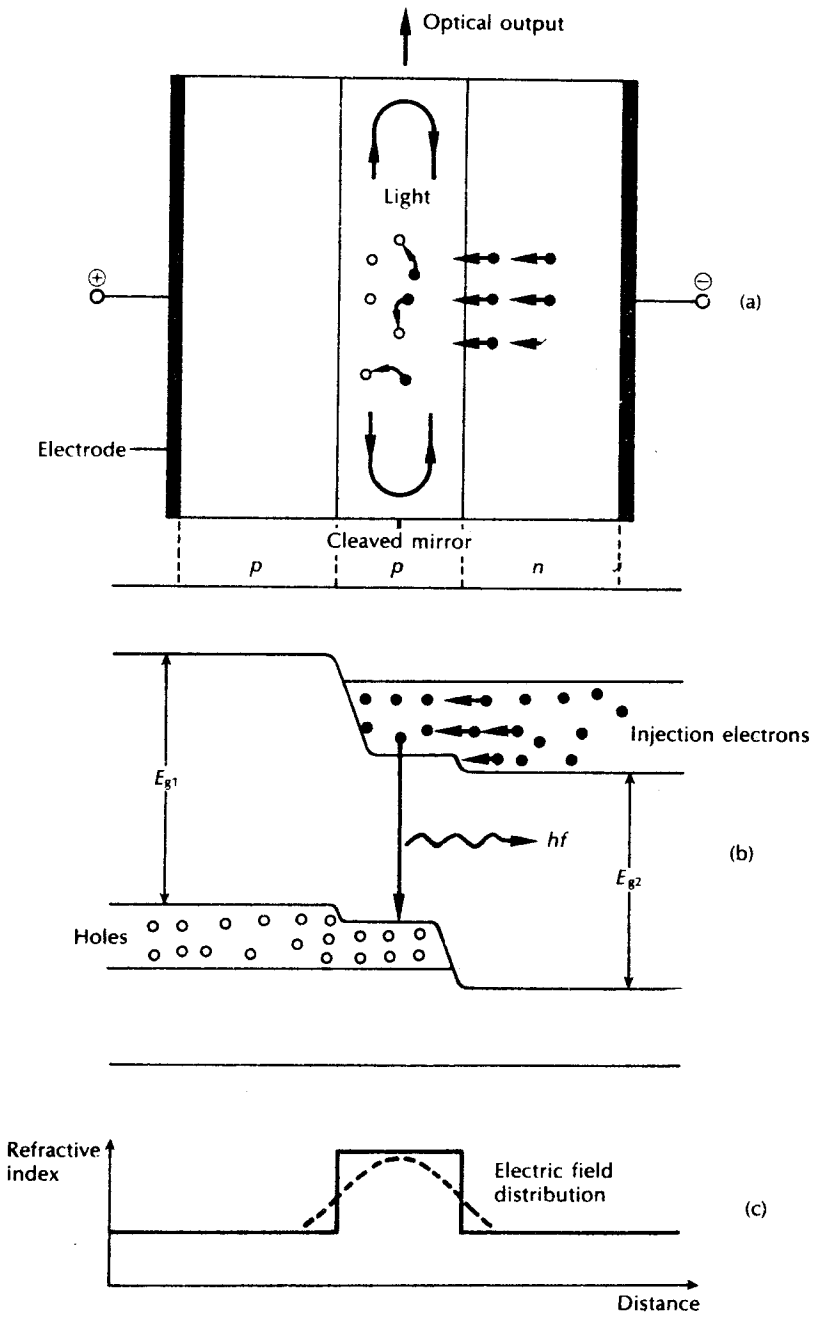


Figure 6.18 The double heterojunction injection laser: (a) the layer structure, shown with an applied forward bias; (b) energy band diagram indicating a *p–p* heterojunction on the left and a *p–n* heterojunction on the right; (c) the corresponding refractive index diagram and electric field distribution.

Alternatively, anisotype heterojunctions with sufficiently large bandgap differences improve the injection efficiency of either electrons or holes. Both types of heterojunction provide a dielectric step due to the different refractive indices at either side of the junction. This may be used to provide radiation confinement to the active region (i.e. the walls of an optical waveguide). The efficiency of the containment depends upon the magnitude of the step which is dictated by the difference in bandgap energies and the wavelength of the radiation.

It is useful to consider the application of heterojunctions in the fabrication of a particular device. They were first used to provide potential barriers in injection lasers. When a double heterojunction (DH) structure was implemented the resulting carrier and optical confinement reduced the threshold currents necessary for lasing by a factor of around 100. Thus stimulated emission was obtained with relatively small threshold currents (50 to 200 mA). The layer structure and an energy band diagram for a DH injection laser are illustrated in Figure 6.18. A heterojunction is shown either side of the active layer for laser oscillation. The forward bias is supplied by connecting a positive electrode of a supply to the p side of the structure and a negative electrode to the n side. When a voltage which corresponds to the bandgap energy of the active layer is applied, a large number of electrons (or holes) are injected into the active layer and laser oscillation commences. These carriers are confined to the active layer by the energy barriers provided by the heterojunctions which are placed within the diffusion length of the injected carriers. It may also be observed from Figure 6.18(c) that a refractive index step (usually a difference of 5 to 10%) at the heterojunctions provides radiation containment to the active layer. In effect the active layer forms the centre of a dielectric waveguide which strongly confines the electroluminescence within this region, as illustrated in Figure 6.18(c). The refractive index step shown is the same for each heterojunction which is desirable in order to prevent losses due to lack of waveguiding which can occur if the structure is not symmetrical.

Careful fabrication of the heterojunctions is also important in order to reduce defects at the interfaces such as misfit dislocations or inclusions which cause nonradiative recombination and thus reduce the internal quantum efficiency. Lattice matching is therefore an important criterion for the materials used to form the interface. Ideally, heterojunctions should have a very small lattice parameter mismatch of no greater than 0.1%. However, it is often not possible to obtain such good lattice parameter matching with the semiconductor materials required to give emission at the desired wavelength and therefore much higher lattice parameter mismatch is often tolerated ($\simeq 0.6\%$).

6.3.6 Semiconductor materials

The semiconductor materials used for optical sources must broadly fulfil several criteria. These are as follows:

1. p–n junction formation. The materials must lend themselves to the formation of p–n junctions with suitable characteristics for carrier injection.

2. Efficient electroluminescence. The devices fabricated must have a high probability of radiative transitions and therefore a high internal quantum efficiency. Hence the materials utilized must be either direct bandgap semiconductors or indirect bandgap semiconductors with appropriate impurity centres.
3. Useful emission wavelength. The materials must emit light at a suitable wavelength to be utilized with current optical fibers and detectors (0.8 to 1.7 μm). Ideally, they should allow bandgap variation with appropriate doping and fabrication in order that emission at a desired specific wavelength may be achieved.

Initial investigation of electroluminescent materials for LEDs in the early 1960s centred around the direct bandgap III–V alloy semiconductors including the binary compounds gallium arsenide (GaAs) and gallium phosphide (GaP) and the ternary gallium arsenide phosphide ($GaAs_xP_{1-x}$). Gallium arsenide gives efficient electroluminescence over an appropriate wavelength band (0.88 to 0.91 μm) and for the first generation optical fiber communication systems was the first material to be fabricated into homojunction semiconductor lasers operating at low temperature. It was quickly realized that improved devices could be fabricated with heterojunction structures which through carrier and radiation confinement would give enhanced light output for drastically reduced device currents. These heterostructure devices were first fabricated using liquid phase epitaxy (LPE) to produce $GaAs/Al_xGa_{1-x}As$ single heterojunction lasers. This process involves the precipitation of material from a cooling solution on to an underlying substrate. When the substrate consists of a single crystal and the lattice constant or parameter of the precipitating material is the same or very similar to that of the substrate (i.e. the unit cells within the two crystalline structures are of a similar dimension), the precipitating material forms an epitaxial layer on the substrate surface. Subsequently, the same technique was used to produce double heterojunctions consisting of $Al_xGa_{1-x}As/GaAs/Al_xGa_{1-x}As$ epitaxial layers, which gave continuous (CW) operation at room temperature [Refs. 10 and 11]. Some of the common material systems now utilized for double heterojunction device fabrication, together with their useful wavelength ranges, are shown in Table 6.2.

The GaAs/AlGaAs DH system is currently by far the best developed and is used for fabricating both lasers and LEDs for the shorter wavelength region. The bandgap in this material may be 'tailored' to span the entire 0.8 to 0.9 μm wavelength band by changing the AlGa composition. Also there is very little lattice mismatch (0.017%) between the AlGaAs epitaxial layer and the GaAs substrate which gives good internal quantum efficiency. In the longer wavelength region (1.1 to 1.6 μm) a number of III–V alloys have been investigated which are compatible with GaAs, InP and GaSb substrates. These include ternary alloys such as $GaAs_{1-x}Sb_x$ and $In_xGa_{1-x}As$ grown on GaAs.

However, although the ternary alloys allow bandgap tailoring they have a fixed lattice parameter. Therefore, quaternary alloys which allow both bandgap tailoring and control of the lattice parameter (i.e. a range of lattice parameters is available

Table 6.2 Some common material systems used in the fabrication of electroluminescent sources for optical fiber communications

Material systems active layer/confining layers	Useful wavelength range (μm)	Substrate
$GaAs/Al_xGa_{1-x}As$	0.8–0.9	GaAs
$GaAs/In_xGa_{1-x}P$	0.9	GaAs
$Al_yGa_{1-y}As/Al_xGa_{1-x}As$	0.65–0.9	GaAs
$In_yGa_{1-y}As/In_xGa_{1-x}P$	0.85–1.1	GaAs
$GaAs_{1-x}Sb_x/Ga_{1-y}Al_yAs_{1-x}Sb_x$	0.9–1.1	GaAs
$Ga_{1-y}Al_yAs_{1-x}Sb_x/GaSb$	1.0–1.7	GaSb
$In_{1-x}Ga_xAs_yP_{1-y}/InP$	0.92–1.7	InP

for each bandgap) appear to be of more use for the longer wavelength region. The most advanced are $In_{1-x}Ga_xAs_yP_{1-y}$ lattice matched to InP and $Ga_{1-y}Al_yAs_{1-x}Sb_x$ lattice matched to GaSb. Both these material systems allow emission over the entire 1.0 to 1.7 μm wavelength band. At present the InGaAsP/InP material system is the most favourable for both long wavelength light sources and detectors. This is due to the ease of fabrication with lattice matching on InP which is also a suitable material for the active region with a bandgap energy of 1.35 eV at 300 K. Hence, InP/InGaAsP (active/confining) devices may be fabricated. Conversely, GaSb is a low bandgap material (0.78 eV at 300 K) and the quaternary alloy must be used for the active region in the GaAlAsSb/GaSb system. Thus compositional control must be maintained for three layers in this system in order to minimize lattice mismatch in the active region, whereas it is only necessary for one layer in the InP/InGaAsP system.

6.4 The semiconductor injection laser

The electroluminescent properties of the forward biased *p–n* junction diode have been considered in the preceding sections. Stimulated emission by the recombination of the injected carriers is encouraged in the semiconductor injection laser (often called the injection laser diode (ILD) or simply the injection laser) by the provision of an optical cavity in the crystal structure in order to provide the feedback of photons. This gives the injection laser several major advantages over other semiconductor sources (e.g. LEDs) that may be used for optical communications. These are as follows:

1. High radiance due to the amplifying effect of stimulated emission. Injection lasers will generally supply milliwatts of optical output power.
2. Narrow linewidth of the order of 1 nm (10 Å) or less which is useful in minimizing the effects of material dispersion.

3. Modulation capabilities which at present extend up into the gigahertz range and will undoubtedly be improved upon.
4. Relative temporal coherence which is considered essential to allow heterodyne (coherent) detection in high capacity systems, but at present is primarily of use in single-mode systems.
5. Good spatial coherence which allows the output to be focused by a lens into a spot which has a greater intensity than the dispersed unfocused emission. This permits efficient coupling of the optical output power into the fiber even for fibers with low numerical aperture. The spatial fold matching to the optical fiber which may be obtained with the laser source is not possible with an incoherent emitter and, consequently, coupling efficiencies are much reduced.

These advantages, together with the compatibility of the injection laser with optical fibers (e.g. size) led to the early developments of the device in the 1960s. Early injection lasers had the form of a Fabry–Perot cavity often fabricated in gallium arsenide which was the major III–V compound semiconductor with electroluminescent properties at the appropriate wavelength for first generation systems. The basic structure of this homojunction device is shown in Figure 6.19, where the cleaved ends of the crystal act as partial mirrors in order to encourage stimulated emission in the cavity when electrons are injected into the *p* type region. However, as mentioned previously these devices had a high threshold current density (greater than 10^4 A cm^{-2}) due to their lack of carrier containment and proved inefficient light sources. The high current densities required dictated that these devices when operated at 300 K were largely utilized in a pulsed mode in order to minimize the junction temperature and thus avert damage.

Improved carrier containment and thus lower threshold current densities (around 10^3 A cm^{-2}) were achieved using heterojunction structures (see Section 6.3.5). The double heterojunction injection laser fabricated from lattice matched III–V alloys

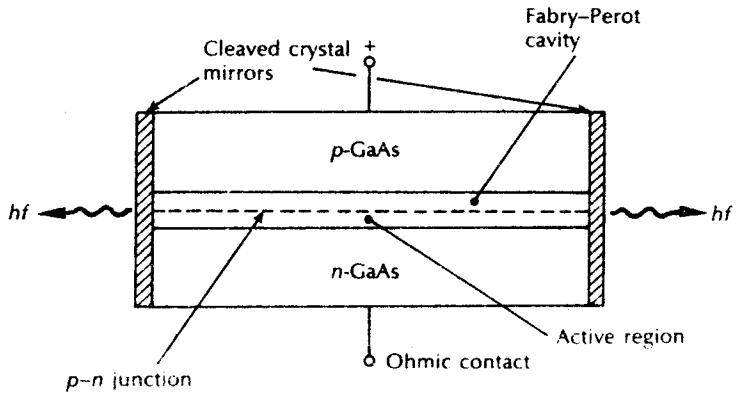


Figure 6.19 Schematic diagram of a GaAs homojunction injection laser with a Fabry–Perot cavity.

provided both carrier and optical confinement on both sides of the p–n junction, giving the injection laser a greatly enhanced performance. This enabled these devices with the appropriate heat sinking to be operated in a continuous wave (CW) mode at 300 K with obvious advantages for optical communications (e.g. analog transmission). However, in order to provide reliable CW operation of the DH injection laser it was necessary to provide further carrier and optical confinement which led to the introduction of stripe geometry DH laser configurations. Prior to discussion of this structure, however, it is useful to consider the efficiency of the semiconductor injection laser as an optical source.

6.4.1 Efficiency

There are a number of ways in which the operational efficiency of the semiconductor laser may be defined. A useful definition is that of the differential external quantum efficiency η_D which is the ratio of the increase in photon output rate for a given increase in the number of injected electrons. If P_e is the optical power emitted from the device, I is the current, e is the charge on an electron, and hf is the photon energy, then:

$$\eta_D = \frac{\mathrm{d}P_e/hf}{\mathrm{d}I/e} \simeq \frac{\mathrm{d}P_e}{\mathrm{d}I(E_g)} \tag{6.35}$$

where E_g is the bandgap energy expressed in electronvolts. It may be noted that η_D gives a measure of the rate of change of the optical output power with current and hence defines the slope of the output characteristic (Figure 6.17) in the lasing region for a particular device. Hence η_D is sometimes referred to as the slope quantum efficiency. For a CW semiconductor laser it usually has values in the range 40 to 60%. Alternatively, the internal quantum efficiency of the semiconductor laser η_i, which was defined in Section 6.3.3.1 as:

$$\eta_i = \frac{\textit{number of photons produced in the laser cavity}}{\textit{number of injected electrons}} \tag{6.36}$$

may be quite high with values usually in the range 50 to 100%. It is related to the differential external quantum efficiency by the expression [Ref. 4]:

$$\eta_D = \eta_i \left[\frac{1}{1 + (2\bar{\alpha}L/\ln(1/r_1 r_2))} \right] \tag{6.37}$$

where $\bar{\alpha}$ is the loss coefficient of the laser cavity, L is the length of the laser cavity and r_1, r_2 are the cleaved mirror reflectivities.

Another parameter is the total efficiency (external quantum efficiency) η_T which is efficiency defined as:

$$\eta_T = \frac{\textit{total number of output photons}}{\textit{total number of injected electrons}} \tag{6.38}$$

$$= \frac{P_e/hf}{I/e} \simeq \frac{P_e}{IE_g} \tag{6.39}$$

As the power emitted P_e changes linearly when the injection current I is greater than the threshold current I_{th}, then:

$$\eta_T \simeq \eta_D\left(1 - \frac{I_{th}}{I}\right) \tag{6.40}$$

For high injection current (e.g. $I = 5I_{th}$) then $\eta_T \simeq \eta_D$, whereas for lower currents ($I \simeq 2I_{th}$) the total efficiency is lower and around 15 to 25%.

The external power efficiency of the device (or device efficiency) η_{ep} in converting electrical input to optical output is given by:

$$\eta_{ep} = \frac{P_e}{P} \times 100 = \frac{P_e}{IV} \times 100\% \tag{6.41}$$

where $P = IV$ is the d.c. electrical input power.

Using Eq. (6.39) for the total efficiency we find:

$$\eta_{ep} = \eta_T\left(\frac{E_g}{V}\right) \times 100\% \tag{6.42}$$

Example 6.5

The total efficiency of an injection laser with a GaAs active region is 18%. The voltage applied to the device is 2.5 V and the bandgap energy for GaAs is 1.43 eV. Calculate the external power efficiency of the device.

Solution: Using Eq. (6.42), the external power efficiency is given by:

$$\eta_{ep} = 0.18\left(\frac{1.43}{2.5}\right) \times 100 \simeq 10\%$$

This result indicates the possibility of achieving high overall power efficiencies from semiconductor lasers which are much larger than for other laser types.

6.4.2 Stripe geometry

The DH laser structure provides optical confinement in the vertical direction through the refractive index step at the heterojunction interfaces, but lasing takes place across the whole width of the device. This situation is illustrated in Figure 6.20 which shows the broad area DH laser where the sides of the cavity are simply formed by roughening the edges of the device in order to reduce unwanted emission in these directions and limit the number of horizontal transverse modes. However, the broad emission area creates several problems including difficult heat sinking, lasing from multiple filaments in the relatively wide active area and unsuitable light output geometry for efficient coupling to the cylindrical fibers.

To overcome these problems whilst also reducing the required threshold current, laser structures in which the active region does not extend to the edges of the device

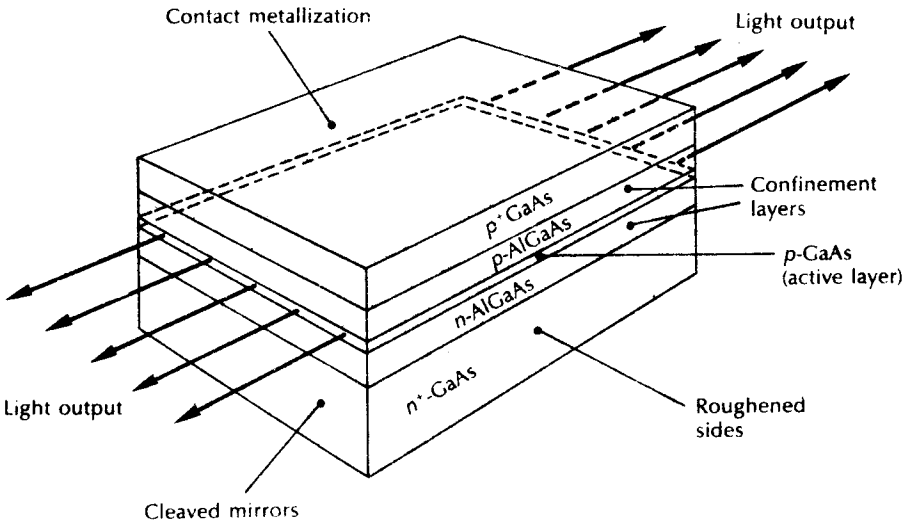


Figure 6.20 A broad area GaAs/AlGaAs DH injection laser.

were developed. A common technique involved the introduction of stripe geometry to the structure to provide optical containment in the horizontal plane. The structure of a DH stripe contact laser is shown in Figure 6.21 where the major current flow through the device and hence the active region is within the stripe. Generally, the stripe is formed by the creation of high resistance areas on either side by techniques such as proton bombardment [Ref. 10] or oxide isolation [Ref. 11]. The stripe therefore acts as a guiding mechanism which overcomes the major problems of the broad area device. However, although the active area width is reduced the light output is still not particularly well collimated due to isotropic emission from a small active region and diffraction within the structure. The optical output and far field emission pattern are also illustrated in Figure 6.21. The output beam divergence is typically 45° perpendicular to the plane of the junction and 9° parallel to it. Nevertheless, this is a substantial improvement on the broad area laser.

The stripe contact device also gives, with the correct balance of guiding, single transverse (in a direction parallel to the junction plane) mode operation, whereas the broad area device tends to allow multimode operation in this horizontal plane. Numerous stripe geometry laser structures have been investigated with stripe widths ranging from 2 to 65 μm, and the DH stripe geometry structure has been widely utilized for optical fiber communications. Such structures have active regions which are planar and continuous. Hence the stimulated emission characteristics of these injection lasers are determined by the carrier distribution (which provides optical gain) along the junction plane. The optical mode distribution along the junction plane is, however, decided by the optical gain and therefore these devices are said to be gain-guided laser structures (see Section 6.5.1).

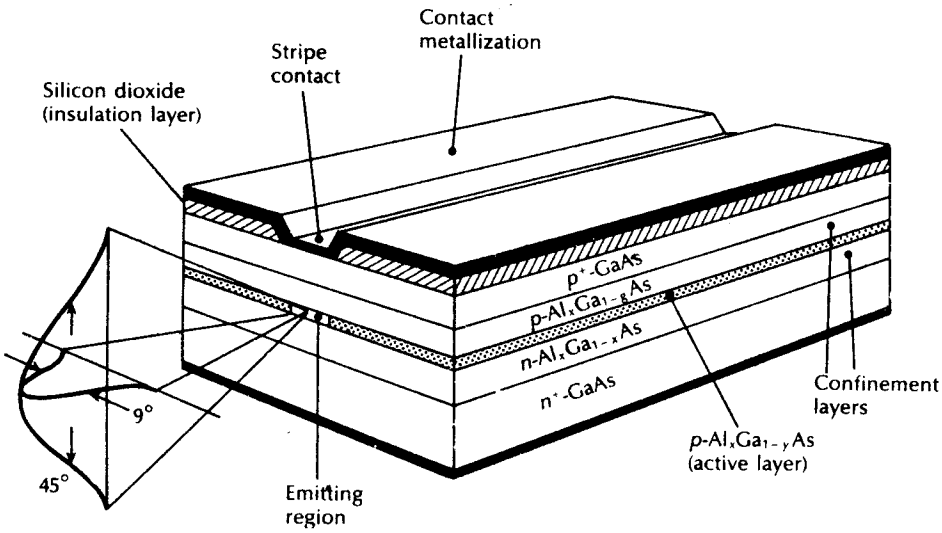


Figure 6.21 Schematic representation of an oxide stripe AlGaAs DH injection laser.

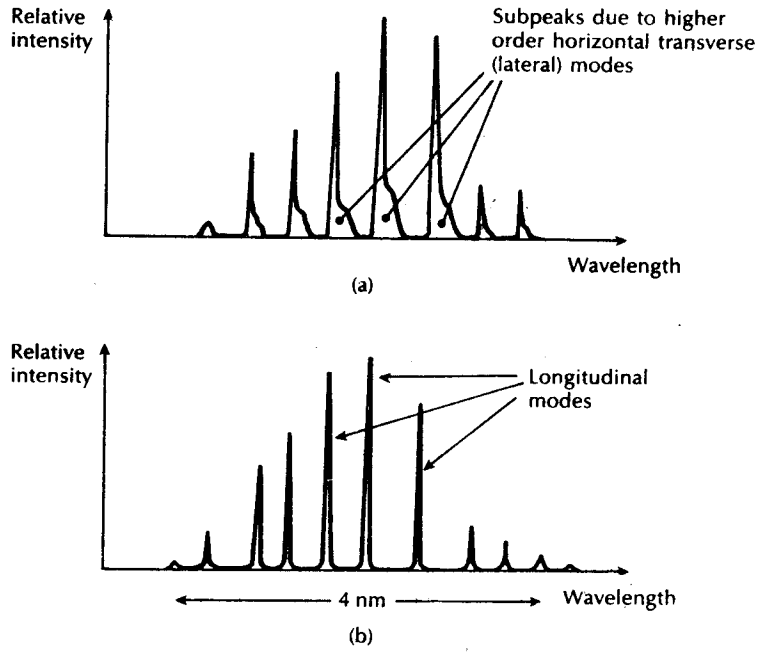


Figure 6.22 Output spectra for multimode injection lasers: (a) broad area device with multitransverse modes; (b) stripe geometry device with single transverse mode.

6.4.3 Laser modes

The typical output spectrum for a broad area injection laser is shown in Figure 6.22(a). It does not consist of a single wavelength output but a series of wavelength peaks corresponding to different longitudinal (in the plane of the junction, along the optical cavity) modes within the structure. As indicated in Section 6.2.4 the spacing of these modes is dependent on the optical cavity length as each one corresponds to an integral number of lengths. They are generally separated by a few tenths of a nanometre, and the laser is said to be a multimode device. However, Figure 6.22(a) also indicates some broadening of the longitudinal mode peaks due to subpeaks caused by higher order horizontal transverse modes.* These higher order lateral modes may exist in the broad area device due to the unrestricted width of the active region. The correct stripe geometry inhibits the occurrence of the higher order lateral modes by limiting the width of the optical cavity leaving only a single lateral mode which gives the output spectrum shown in Figure 6.22(b) where only the longitudinal modes may be observed. This represents the typical output spectrum for a good multimode injection laser.

6.4.4 Single-mode operation

For single-mode operation, the optical output from a laser must contain only a single longitudinal and single transverse mode. Hence the spectral width of the emission from the single-mode device is far smaller than the broadened transition linewidth discussed in Section 6.2.4. It was indicated that an inhomogeneously broadened laser can support a number of longitudinal and transverse modes simultaneously, giving a multimode output. Single transverse mode operation, however, may be obtained by reducing the aperture of the resonant cavity such that only the TEM_{00} mode is supported. To obtain single-mode operation it is then necessary to eliminate all but one of the longitudinal modes.

One method of achieving single longitudinal mode operation is to reduce the length L of the cavity until the frequency separation of the adjacent modes given by Eq. (6.14) as $\delta f = c/2nL$ is larger than the laser transition linewidth or gain curve. Then only the single mode which falls within the transition linewidth can oscillate within the laser cavity. However, it is clear that rigid control of the cavity parameters is essential to provide the mode stabilization necessary to achieve and maintain this single-mode operation.

The structures required to give mode stability are discussed with regard to the multimode injection laser in Section 6.5 and similar techniques can be employed to produce a laser emitting a single longitudinal and transverse mode. For example, the correct DH structure will restrict the vertical width of the waveguiding region to less than 0.4 μm allowing only the fundamental transverse mode to be supported

*Tranverse modes in the plane of the junction are often called lateral modes, transverse mode being reserved for modes perpendicular to the junction plane.

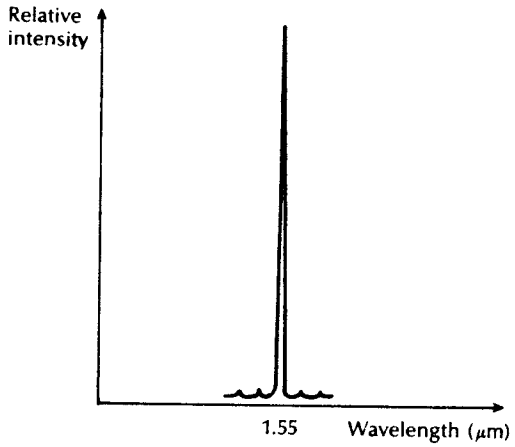


Figure 6.23 Typical single longitudinal mode output spectrum from a single-mode injection laser.

and removing any interference of the higher order transverse modes on the emitted longitudinal modes.

The lateral modes (in the plane of the junction) may be confined by the restrictions on the current flow provided by the stripe geometry. In general, only the lower order modes are excited which appear as satellites to each of the longitudinal modes. However, as will be discussed in Section 6.5.1, stripe contact devices often have instabilities and strong nonlinearities (e.g. kinks) in their light output against current characteristics. Tight current confinement as well as good waveguiding are therefore essential in order to achieve only the required longitudinal modes which form between the mirror facets in the plane of the junction. Finally, as indicated above, single-mode operation may be obtained through control of the optical cavity length such that only a single longitudinal mode falls within the gain bandwidth of the device. Figure 6.23 shows a typical output spectrum for a single-mode device.

However, injection lasers with short cavity lengths (around 50 μm) are difficult to handle and have not been particularly successful. Nevertheless, such devices, together with the major alternative structures which provide single-mode operation, are dealt with in Section 6.6 under the title of single frequency injection lasers.

6.5 Some injection laser structures

6.5.1 Gain-guided lasers

Fabrication of multimode injection lasers with a single or small number of lateral modes is achieved by the use of stripe geometry. These devices are often called gain-

guided lasers as indicated in Section 6.4.2. The constriction of the current flow to the stripe is realized in the structure either by implanting the regions outside the stripe with protons (proton isolated stripe) to make them highly resistive, or by oxide or *p–n* junction isolation. The structure for an aluminium gallium arsenide oxide isolated stripe DH laser is shown in Figure 6.21. It has an active region of gallium arsenide bounded on both sides by aluminium gallium arsenide regions. This technique has been widely applied, especially for multimode laser structures used in the shorter wavelength region. The current is confined by etching a narrow stripe in a silicon dioxide film.

Two other basic techniques for the fabrication of gain-guided laser structures are illustrated in Figure 6.24(a) and (b), which show the proton isolated stripe and the *p–n* junction isolated stripe structures respectively. In Figure 6.24(a) the resistive region formed by the proton bombardment gives better current confinement than the simple oxide stripe and has superior thermal properties due to the absence of the silicon dioxide layer; *p–n* junction isolation involves a selective diffusion through the *n* type surface region in order to reach the *p* type layers, as illustrated in Figure 6.24(b). None of these structures confines all the radiation and current to the stripe region and spreading occurs on both sides of the stripe. With stripe widths of 10 μm or less, such planar stripe lasers provide highly efficient coupling into multimode fibers, but significantly lower coupling efficiency is achieved into small core diameter single-mode fibers.

The optical output power against current characteristic for the ideal semiconductor laser was illustrated in Figure 6.17. However, with certain practical laser diodes the characteristic is not linear in the simulated emission region, but exhibits

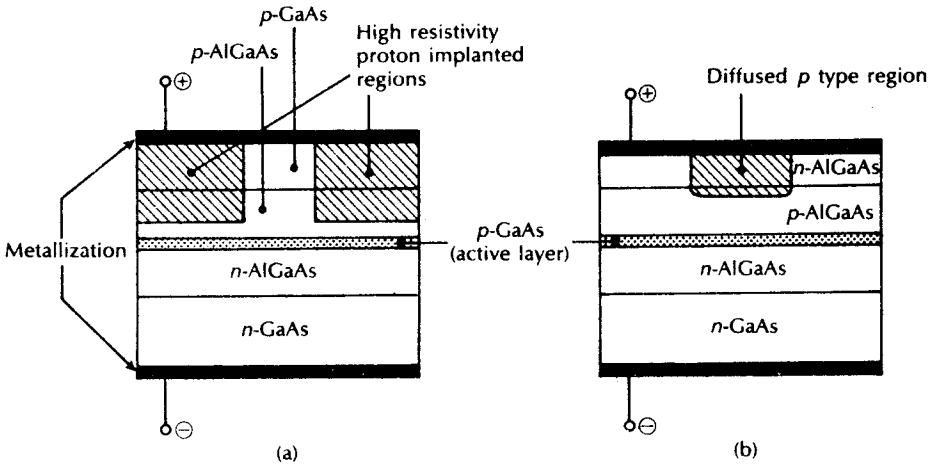


Figure 6.24 Schematic representation of structures for stripe geometry injection lasers: (a) proton isolated stripe GaAs/AlGaAs laser; (b) *p–n* junction isolated (diffused planar stripe) GaAs/AlGaAs laser.

kinks. This phenomenon is particularly prevalent with gain-guided injection laser devices. The kinks may be classified into two broad categories.

The first type of kink results from changes in the dominant lateral mode of the laser as the current is changed. The output characteristic or laser *A* in Figure 6.25(a) illustrates this type of kink where lasing from the device changes from the fundamental lateral mode to a higher order lateral mode (second order) in a current region corresponding to a change in slope. The second type of kink involves a 'spike', as observed for laser *B* of Figure 6.25(a). These spikes have been shown to be associated with filamentary behaviour within the active region of the device [Ref. 4]. The filaments result from defects within the crystal structure.

Both these mechanisms affect the near and far field intensity distributions (patterns) obtained from the laser. A typical near field intensity distribution corresponding to a single optical output power level in the plane of the junction is shown in Figure 6.25(b). As this distribution is in the lateral direction, it is determined by the nature of the lateral waveguide. The single intensity maximum shown indicates that the fundamental lateral mode is dominant. To maintain such a near field pattern the stripe geometry of the device is important. In general, relatively narrow stripe devices ($< 10\,\mu$m) formed by a planar process allow the fundamental lateral mode to dominate. This is especially the case at low power levels where near field patterns similar to Figure 6.25(b) may be obtained.

Although gain-guided lasers are commercially available for operation in both the shorter wavelength range (using GaAs active regions) and the longer wavelength

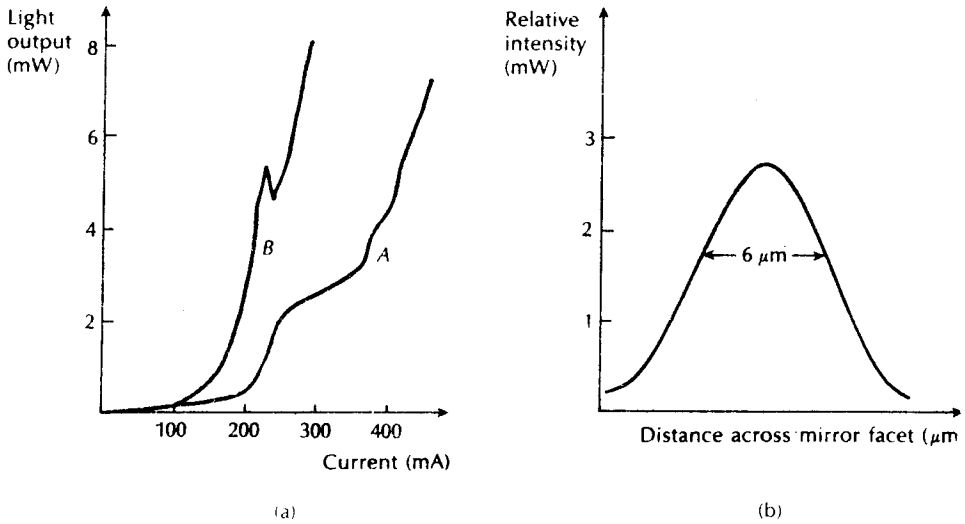


Figure 6.25 (a) The light output against current characteristic for an injection laser with nonlinearities or a kink in the stimulated emission region; (b) A typical near field intensity distribution (pattern) in the plane of the junction for an injection laser.

range (using InGaAsP active regions) they exhibit several undesirable characteristics. Apart from the nonlinearities in the light output versus current characteristics discussed above, gain-guided injection lasers have relatively high threshold currents (100 to 150 mA) as well as low differential quantum efficiency [Ref. 13]. These effects are primarily caused by the small carrier-induced refractive index reduction within the devices which results in the movement of the optical mode along the junction plane. The problems can be greatly reduced by introducing some real refractive index variation into the lateral structure of the laser such that the optical mode along the junction plane is essentially determined by the device structure.

6.5.2 Index-guided lasers

The drawbacks associated with the gain-guided laser structures were largely overcome through the development of index-guided injection lasers. In some such structures with weak index-guiding, the active region waveguide thickness is varied by growing it over a channel or ridge in the substrate. In the ridge waveguide laser shown in Figure 6.26(a), the ridge not only provides the loading for the weak index-guiding but also acts as a narrow current confining stripe [Ref. 14]. These devices have been fabricated to operate at various wavelengths with a single lateral mode, and room temperature CW threshold currents as low as 18 mA with output powers of 25 mW have been reported [Ref. 15]. More typically, the threshold currents for such weakly index-guided structures are in the range 40 to 60 mA, as illustrated in Figure 6.26(b) which compares a light output versus current characteristic for a ridge waveguide laser with that of an oxide stripe gain-guided device.

Alternatively, the application of a uniformly thick, planar active waveguide can be achieved through lateral variations in the confinement layer thickness or the refractive index. The inverted-rib waveguide device (sometimes called plano-convex waveguide) illustrated in Figure 6.26(c) is an example of this structure. However, room temperature CW threshold currents are between 70 and 90 mA with output powers of around 20 mW for InGaAsP devices operating at a wavelength of 1.3 μm [Ref. 16].

Strong index-guiding along the junction plane can provide improved transverse mode control in injection lasers. This can be achieved using a buried heterostructure (BH) device in which the active volume is completely buried in a material of wider bandgap and lower refractive index [Refs. 17, 18]. The structure of a BH laser is shown in Figure 6.27(a). The optical field is well confined both in the transverse and lateral directions within these lasers, providing strong index-guiding of the optical mode together with good carrier confinement. Confinement of the injected current to the active region is obtained through the reverse biased junctions of the higher bandgap material. It may be observed from Figure 6.27 that the higher bandgap, low refractive index confinement material is AlGaAs for GaAs lasers operating in the 0.8 to 0.9 μm wavelength range, whereas it is InP in InGaAsP devices operating in the 1.1 to 1.6 μm wavelength range.

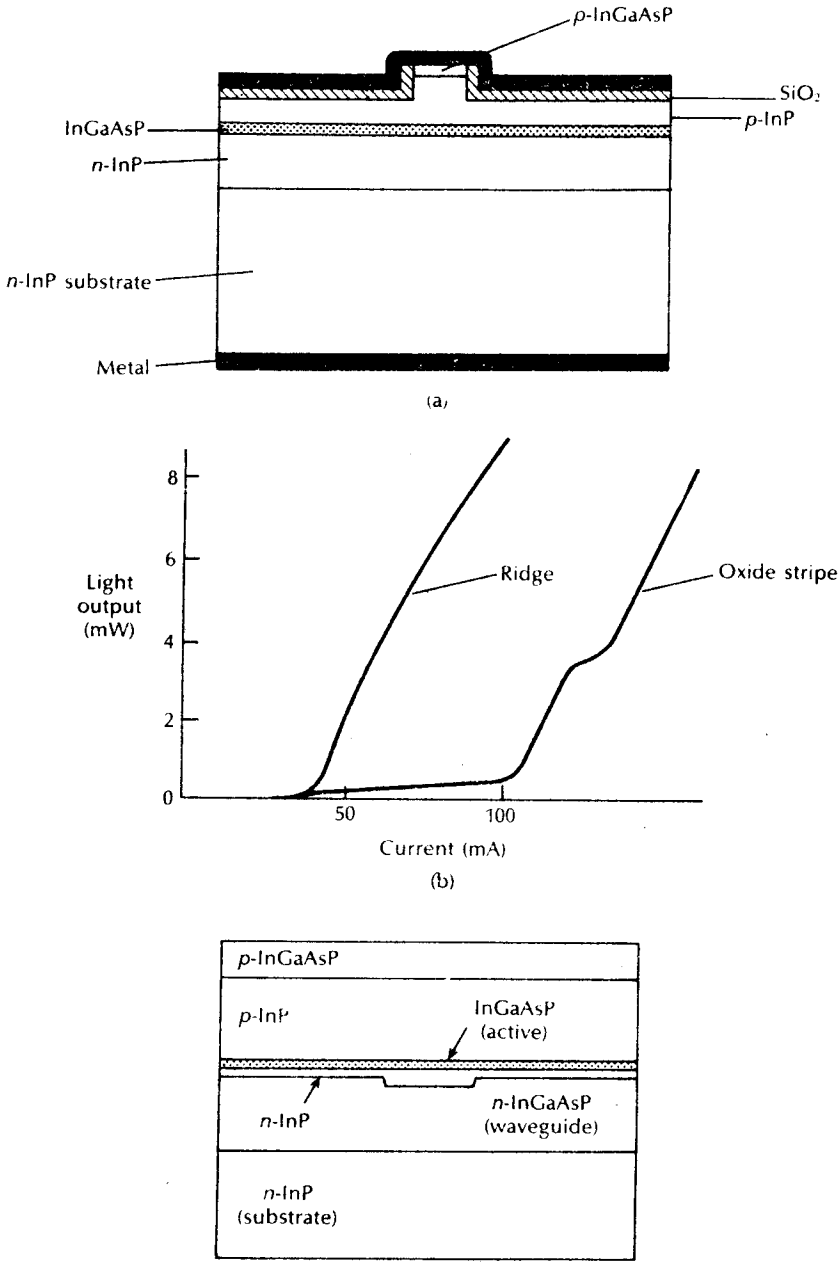


Figure 6.26 Index-guided lasers: (a) ridge waveguide injection laser structures; (b) light output versus current characteristic for (a) compared with that of an oxide stripe (gain-guided) laser; (c) rib (plano-convex) waveguide injection laser structure.

A wide variety of BH laser configurations are commercially available offering both multimode and single-mode operation. In general, the lateral current confinement provided by these devices leads to lower threshold currents (10 to 20 mA) than may be obtained with either weakly index-guided or gain-guided structures. A more complex structure called the double channel planar buried heterostructure (DCPBH) laser is illustrated in Figure 6.27(b) [Ref. 19]. This device which has a planar InGaAsP active region provides very high power operation with CW output powers up to 40 mW in the longer wavelength region. Room temperature threshold currents are in the range 15 to 20 mA for both 1.3 μm and 1.55 μm emitting devices [Ref. 20]. Lateral mode control may be achieved by reducing the dimension of the active region, with a cross sectional area of 0.3 μm^2 being required for fundamental mode operation [Ref. 13].

Parasitic capacitances resulting from the use of the reverse biased current confinement layers can reduce the high speed modulation capabilities of BH lasers. However, this problem has been overcome through either the regrowth of

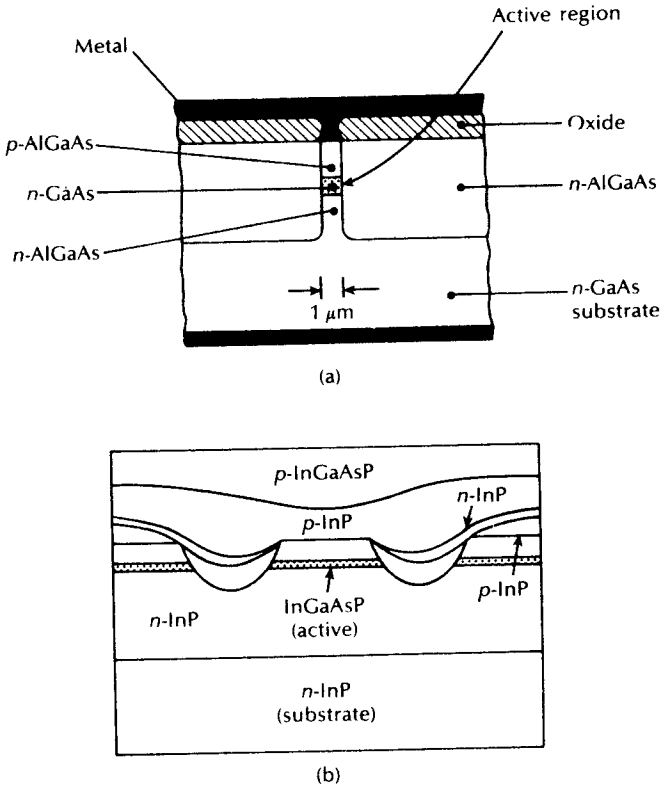


Figure 6.27 Buried heterostructure laser structures: (a) GaAs/AlGaAs BH device; (b) InGaAsP/InP double channel planar BH device.

semiinsulating material [Ref. 2] or the deposition of a dielectric material [Ref. 22]. Using these techniques, modulation speeds in excess of 20 GHz have been achieved which are limited by the active region rather than the parasitic capacitances [Ref. 18].

6.5.3 Quantum-well lasers

In recent years DH lasers have been fabricated with very thin active layer thicknesses of around 10 nm instead of the typical range for conventional DH

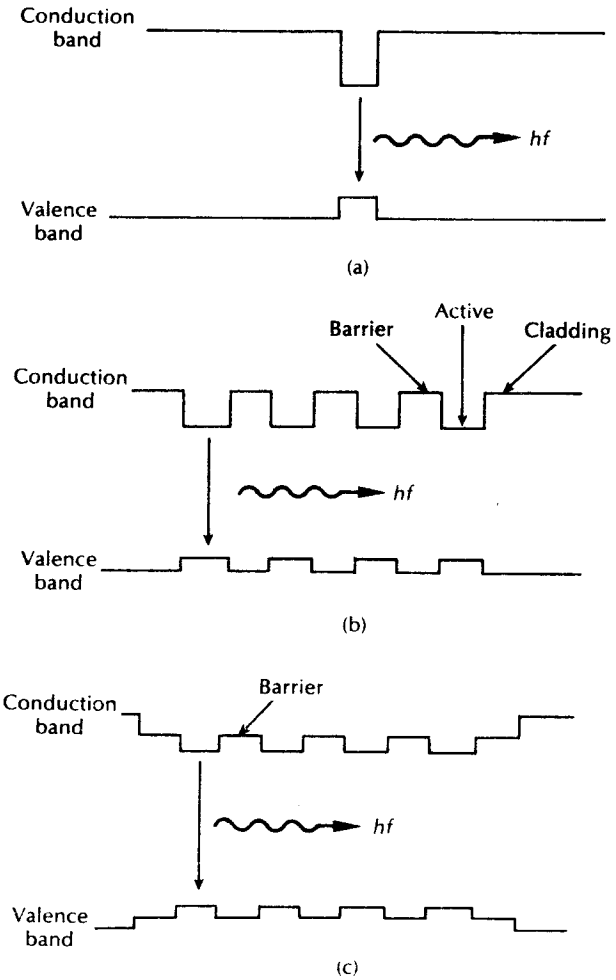


Figure 6.28 Energy band diagrams showing various types of quantum-well structure: (a) single quantum-well; (b) multiquantum-well; (c) modified multiquantum-well.

structures of 0.1 to 0.3 μm. The carrier motion normal to the active layer in these devices is restricted, resulting in a quantization of the kinetic energy into discrete energy levels for the carriers moving in that direction. This effect is similar to the well known quantum mechanical problem of a one dimensional potential well [Ref. 13] and therefore these devices are known as quantum-well lasers. In this structure the thin active layer causes drastic changes to the electronic and optical properties in comparison with a conventional DH laser. These changes are due to the quantized nature of the discrete energy levels with a step-like density of states which differs from the continuum normally obtained. Hence, quantum-well lasers exhibit an inherent advantage over conventional DH devices in that they allow high gain at low carrier density, thus providing the possibility of significantly lower threshold currents.

Both single quantum-well (SQW), corresponding to a single active region and multiquantum-well (MQW), corresponding to multiple active regions lasers have been fabricated [Ref. 13]. In the latter structure, the layers separating the active regions are called barrier layers. Energy band diagrams for the active regions of these structures are displayed in Figure 6.28. It may be observed in Figure 6.28(c) that when the bandgap energy of the barrier layer differs from the cladding layer in a MQW device it is usually referred to as a modified multiquantum-well laser [Ref. 23].

Better confinement of the optical mode is obtained in MQW lasers in comparison with SQW lasers, resulting in a lower threshold current density for these devices. A substantial amount of experimental work has been carried out on MQW lasers using the AlGaAs/GaAs material system. It has demonstrated the superior characteristics of MQW devices over conventional DH lasers in relation to lower threshold currents, narrower linewidths, higher modulation speeds, lower frequency chirp and less temperature dependence (see Section 6.7.1) [Ref. 18]. However, these potential performance advantages have not, as yet, been fully explored for longer wavelength MQW lasers using the InGaAsP/InP material system.

6.6 Single frequency injection lasers

Although the structures described in Section 6.5 provide control of the lateral modes of the laser, the Fabry–Perot cavity formed by the cleaved laser mirrors may allow several longitudinal modes to exist within the gain spectrum of the device (see Section 6.4.3). Nevertheless, such injection laser structures will provide single longitudinal mode operation, even though the mode discrimination obtained from the gain spectrum is often poor. However, improved longitudinal mode selectivity can be achieved using structures which give adequate loss discrimination between the desired mode and all of the unwanted modes of the laser resonator. It was indicated in Section 6.4.4 that such mode discrimination could be obtained by shortening the laser cavity This technique has met with only limited success [Ref.

18] and therefore alternative structures have been developed to give the necessary electrical and optical containment to allow stable single longitudinal mode operation.

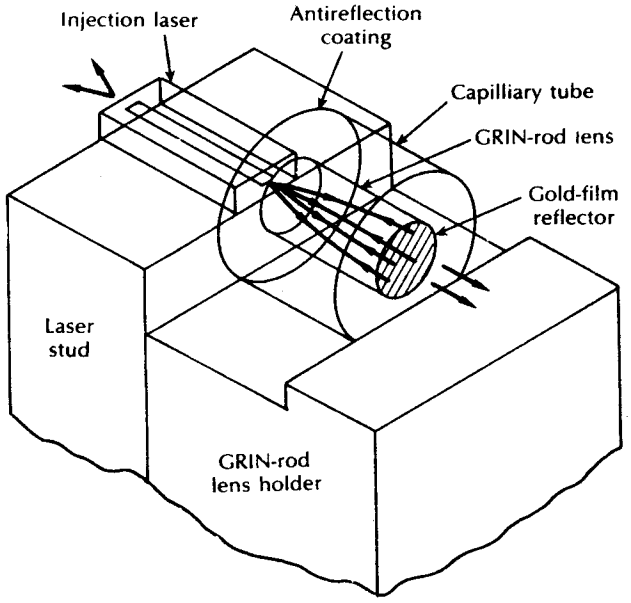
Longitudinal mode selectivity may be improved through the use of frequency-selective feedback so that the cavity loss is different for the various longitudinal modes. Devices which employ this technique to provide single longitudinal mode operation are often referred to as single frequency or dynamic single-mode (DSM) lasers [Refs. 24, 25]. Such lasers are of increasing interest not only to reduce fiber intramodal dispersion within high speed systems but also for the provision of suitable sources for coherent optical transmission (see Chapter 12). Strategies which have proved successful in relation to single frequency operation are the use of short cavity resonators, coupled cavity resonators and distributed feedback.

6.6.1 Short and coupled cavity lasers

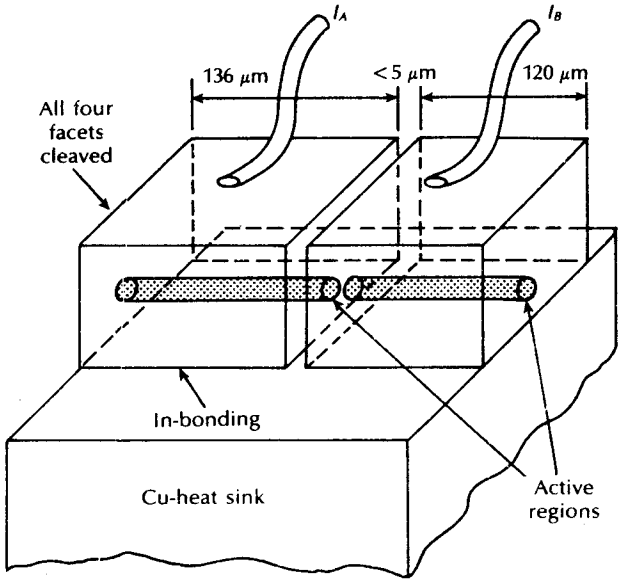
It was suggested in Section 6.4.4 that a straightforward method for increasing the longitudinal mode discrimination of an injection laser is to shorten the cavity length; shortening from, say, 250 to 25 μm will have the effect of increasing the mode spacing from 1 to 10 nm. The peak of the gain curve can then be adjusted to provide the desired single-mode operation. Conventional cleaved mirror structures are, however, difficult to fabricate with cavity lengths below 50 μm and therefore configurations employing resonators, either microcleaved [Ref. 26] or etched [Ref. 27], have been utilized. Such resonators form a short cavity of length 10 to 20 μm in a direction normal to the active region providing stable single frequency operation.

Multiple element resonators or resonators with distributed reflectors also give a loss mechanism with a frequency dependence which is strong enough to provide single frequency oscillation under most operating conditions. Mode selectivity in such a coupled cavity laser is obtained when the longitudinal modes of each Fabry–Perot cavity coincide and therefore constitute the longitudinal modes of the coupled system for which both cavities are in resonance. One example of a three mirror resonator shown in Figure 6.29(a) uses a graded index (GRIN) rod lens (see Section 5.5.1) to enhance the coupling to an external mirror [Ref. 28]

An alternative approach is illustrated in Figure 6.29(b) in which two active laser sections are separated by a gap of approximately a single wavelength. When the gap is obtained by recleaving a finished laser chip into two partially attached segments it yields the cleaved-coupled-cavity (C^3) laser [Ref. 29]. This four mirror resonator device has provided dynamic single-mode operation with side mode suppression ratios of several thousand being achieved through control of the magnitudes and the relative phases of the two injection currents, as well as the temperature [Ref. 30]. Another attribute of the C^3 device is that its single frequency emission can be tuned discretely over a range of some 26 nm by varying the current through one section [Ref. 13]. This tunability, which occurs through mode jumps of around 2 nm each, is discussed further in Section 6.10.



(a)



(b)

Figure 6.29 Coupled cavity lasers: (a) short external cavity laser using GRIN-rod lens; (b) cleaved-coupled-cavity laser.

6.6.2 Distributed feedback lasers

An elegant approach to single frequency operation which has recently found widespread application involves the use of distributed resonators, fabricated into the laser structure to give integrated wavelength selectivity. The structure which is employed is the distributed Bragg diffraction grating which provides periodic variation in refractive index in the laser heterostructure waveguide along the direction of wave propagation so that feedback of optical energy is obtained through Bragg reflection (see Section 10.6.3) rather than by the usual cleaved mirrors. Hence the corrugated grating structure shown in Figure 6.30(a) determines the wavelength of the longitudinal mode emission instead of the Fabry–Perot gain curve shown in Figure 6.30(b). When the period of the corrugation is equal to $l\lambda_B/2n_e$, where l is the integer order of the grating, λ_B is the Bragg wavelength and n_e is the effective refractive index of the waveguide; then only the mode near the Bragg wavelength λ_B is reflected constructively (i.e. Bragg reflection). Therefore, as may be observed in Figure 6.30(a), this particular mode will lase whilst the other modes exhibiting higher losses are suppressed from oscillation.

It should be noted that first order gratings (i.e. $l = 1$) provide the strongest coupling within the device. Nevertheless, second order gratings are sometimes used as their larger spatial period eases fabrication.

From the viewpoint of device operation, semiconductor lasers employing the distributed feedback mechanism can be classified into two broad categories, referred to as the distributed feedback (DFB) laser [Ref. 31] and the distributed Bragg reflector (DBR) laser [Ref. 32]. These two device structures are shown

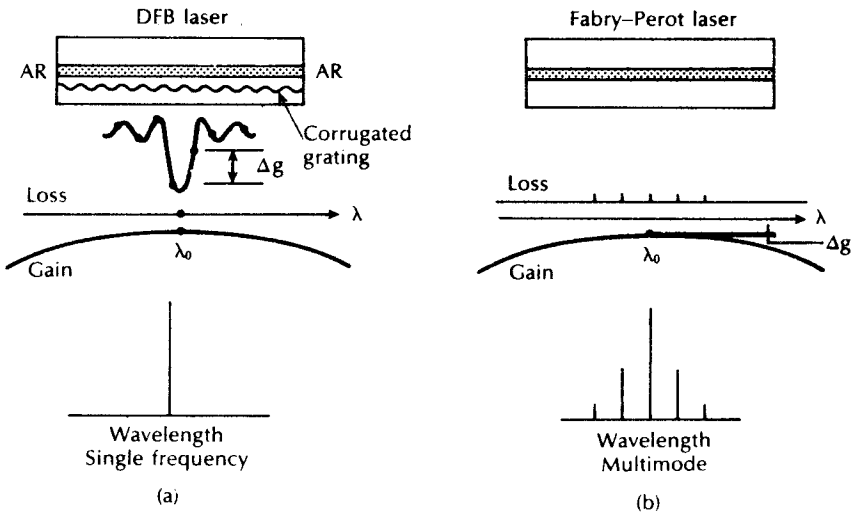


Figure 6.30 Illustration showing the single frequency operation of: (a) the distributed feedback (DFB) laser in comparison with; (b) the Fabry–Perot laser.

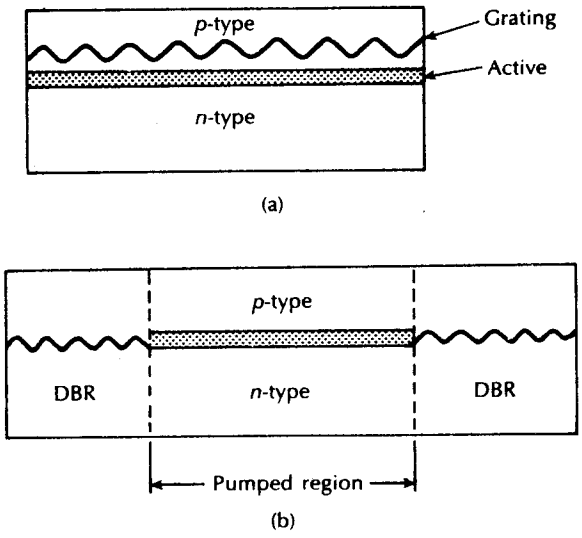


Figure 6.31 Schematic of distributed feedback lasers: (a) DFB laser; (b) DBR laser.

schematically in Figure 6.31. In the DFB laser the optical grating is usually applied over the entire active region which is pumped, whereas in the DBR laser the grating is etched only near the cavity ends and hence distributed feedback does not occur in the central active region. The unpumped corrugated end regions effectively act as mirrors whose reflectivity results from the distributed feedback mechanism which is therefore dependent on wavelength. In addition, this latter device displays the advantage of separating the perturbed regions from the active region but proves somewhat lossy due to optical absorption in the unpumped distributed reflectors. It should be noted that in Figure 6.31 the grating is shown in a passive waveguide layer adjacent to the active gain region for both device structures. This structure has evolved as a result of the performance deterioration with earlier devices (at temperatures above 80 K) in which the corrugations were applied directly to the active layer [Ref. 33].

At present, DBR lasers are less well developed than DFB lasers and it is the latter devices which are under intensive investigation for the provision of single frequency semiconductor optical sources. Any of the semiconductor laser structures discussed in Section 6.5 can be employed to fabricate a DFB laser after etching a grating into an appropriate cladding layer adjacent to the active layer. The grating period is determined by the desired emission frequency from the structure following the Bragg condition (see Section 10.6.3). In particular, buried heterostructure DFB lasers have been developed in many laboratories which exhibit low threshold currents (10 to 20 mA), high modulation speeds (several G bit s^{-1}) and output powers comparable with Fabry–Perot devices with similar BH geometries [Refs.

13, 18, 34]. The structure and the light output versus current characteristic for a low threshold current DFB–BH laser operating at a wavelength of 1.55 μm is displayed in Figure 6.32 [Ref. 35]. A substantial change in the output characteristic with increasing temperature may be observed for this separate confinement DH device (see Section 6.7.1).

In theory when considering a DFB laser with both end facets antireflection (AR) coated (see Figure 6.30(a)), then two modes located symmetrically on either side of the Bragg wavelength will experience the same lowest threshold gain within an ideal symmetrical structure and will therefore lase simultaneously. However, in practice, the randomness associated with the cleaving process creates different end phases, thus removing the degeneracy of the modal gain and providing only single-mode operation. Moreover, facet asymmetry can be increased by placing a high reflection coating on one end facet and a low reflection coating on the other (known as the hi–lo structure) in order to improve the power output for single frequency operation [Ref. 18].

Another technique to improve the performance of the DFB laser is to modify the grating at a central point to introduce an additional optical phase shift, typically a quarter wavelength or less [Refs. 36, 37]. Such a device is shown in Figure 6.33 which illustrates the structure of an InGaAsP/InP double channel planar buried

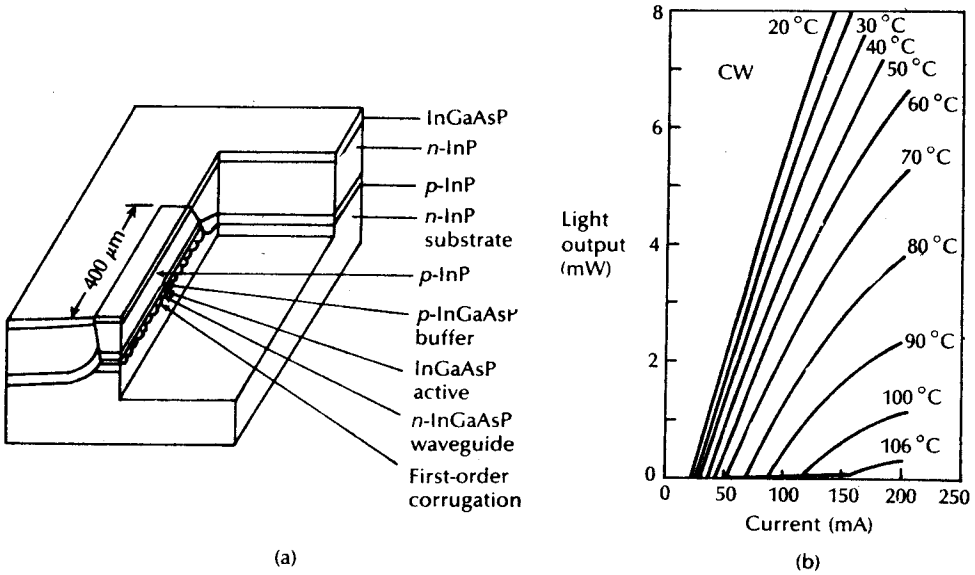


Figure 6.32 DFB buried heterostructure laser with a window structure: (a) structure; (b) light output versus current characteristics for various temperatures under CW operation. Reproduced with permission from S. Tsuji, A. Ohishi, H. Nakamura, M. Hirao, N. Chinone and H. Matsumura, 'Low threshold operation of 1.5 μm DFB laser diodes', *J. Lightwave Technol.*, **LT-5**, p. 822, 1987. Copyright © 1987 IEEE.

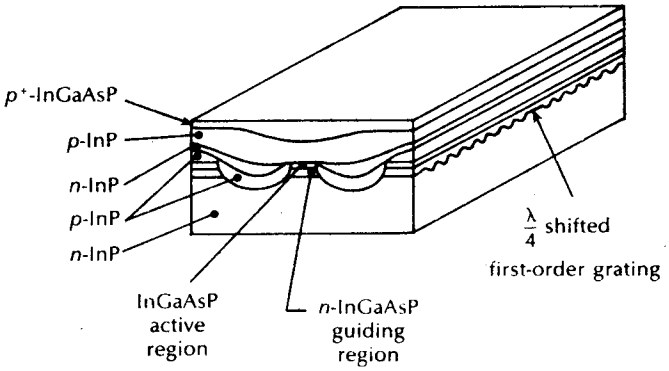


Figure 6.33 Quarter wavelength shifted double channel planar BH-DFB laser.

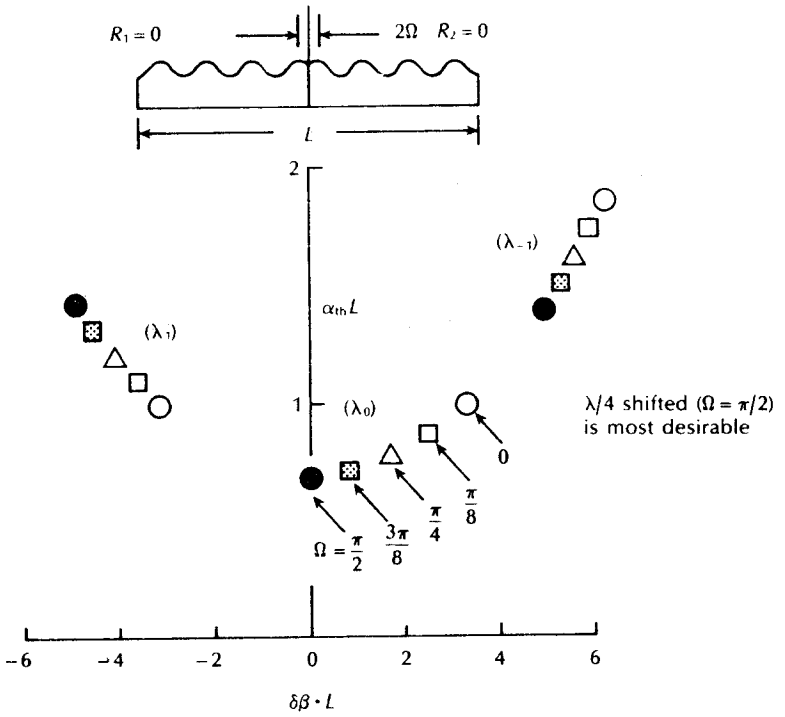


Figure 6.34 The threshold gain and mode frequency in a phase shifted DFB laser.

heterostructure DFB laser with a quarter wavelength shifted first order grating [Ref. 38]. This structure, which provides excellent, stable, single frequency operation, incorporates a $\pi/2$ phase shift (equivalent to one quarter wavelength) in the corrugation at the centre of the laser cavity with both end facets AR coated. The threshold gain and the mode frequency (relative to the Bragg wavelength) for the device as the phase shift is varied from 0 to $\pi/2$ is shown in Figure 6.34. It may be observed that the lowest threshold gain for the central mode (at λ_0) is obtained precisely at the Bragg wavelength when the phase shift is $\pi/2$. Furthermore, the gain difference between the central mode and the nearest side mode (at λ_1) has the largest value at this phase shift.

The performance of the quarter wavelength shifted DFB laser is superior to that of the conventional DFB structure because the large gain difference between the central mode and the side modes gives improved dynamic single-mode stability with negligible mode partition noise (see Section 6.7.4) at multigigabit s^{-1} modulation speeds [Ref. 38]. In addition, narrow linewidths of around 3 MHz ($\simeq 2 \times 10^{-5}$ nm) have been obtained under CW operation [Ref. 18], which is substantially less than the typical 100 MHz ($\simeq 6 \times 10^{-4}$ nm) linewidth associated with the Fabry–Perot injection laser. Linewidth narrowing is achieved within such DFB lasers by detuning the lasing wavelength towards the shorter wavelength side of the gain peak (i.e. towards λ_{-1} in Figure 6.34) in order to increase the differential gain between the central mode and the nearest side mode (λ_1 in Figure 6.34). This strategy is sometimes referred to as Bragg wavelength detuning.

6.7 Injection laser characteristics

When considering the use of the injection laser for optical fiber communications it is necessary to be aware of certain of its characteristics which may affect its efficient operation. The following sections outline the major operating characteristics of the device (the ones which have not been dealt with in detail previously) which generally apply to all the various materials and structures previously discussed, although there is substantial variation in behaviour between them.

6.7.1 Threshold current temperature dependence

Figure 6.35 shows the variation in threshold current with temperature for two gain-guided (oxide insulated stripe) injection lasers [Ref. 39]. Both devices had stripe widths of approximately 20 μm but were fabricated from different material systems for emission at wavelengths of 0.85 μm and 1.55 μm (AlGaAs and InGaAsP devices respectively).

In general terms the threshold current tends to increase with temperature, the temperature dependence of the threshold current density J_{th} being approximately exponential [Ref. 4] for most common structures. It is given by:

$$J_{th} \propto \exp \frac{T}{T_0} \tag{6.43}$$

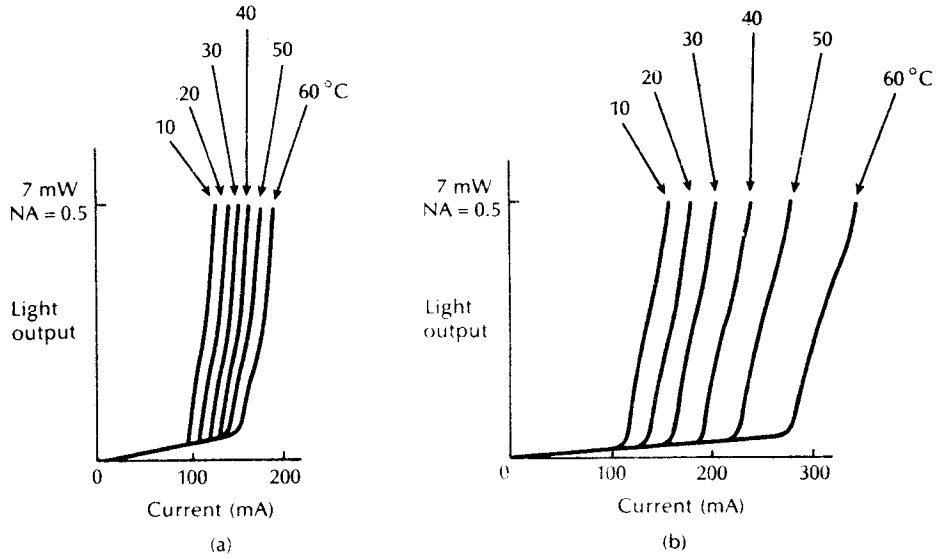


Figure 6.35 Variation in threshold current with temperature for gain-guided injection lasers: (a) AlGaAs device; (b) InGaAsP device. Reproduced with permission from P. A. Kirby, 'Semiconductor laser sources for optical communication', *The Radio and Electron. Eng.*, **51**, p. 362, 1981.

where T is the device absolute temperature and T_0 is the threshold temperature coefficient which is a characteristic temperature describing the quality of the material, but which is also affected by the structure of the device. For AlGaAs devices, T_0 is usually in the range 120 to 190 K, whereas for InGaAsP devices it is between 40 and 75 K [Ref. 40]. This emphasizes the stronger temperature dependence of InGaAsP structures which is illustrated in Figure 6.35 and Example 6.6. The increase in threshold current with temperature for AlGaAs devices can be accounted for with reasonable accuracy by consideration of the increasing energy spread of electrons and holes injected into the conduction and valence bands. It appears that the intrinsic physical properties of the InGaAsP material system may cause its higher temperature sensitivity; these include Auger recombination, intervalence band absorption and carrier leakage effects over the heterojunctions [Ref. 41]. The relative significance of these various mechanisms is, however, not clearly understood [Ref. 42].

Example 6.6

Compare the ratio of the threshold current densities at 20 °C and 80 °C for a AlGaAs injection laser with $T_0 = 160$ K and the similar ratio for an InGaAsP device with $T_0 = 55$ K.

Solution: From Eq. (6.43) the threshold current density:

$$J_{th} \propto \exp \frac{T}{T_0}$$

For the AlGaAs device:

$$J_{th}\ (20\,^{\circ}\mathrm{C}) \propto \exp \frac{293}{160} = 6.24$$

$$J_{th}\ (80\,^{\circ}\mathrm{C}) \propto \exp \frac{353}{160} = 9.08$$

Hence the ratio of the current densities:

$$\frac{J_{th}\ (80\,^{\circ}\mathrm{C})}{J_{th}\ (20\,^{\circ}\mathrm{C})} = \frac{9.08}{6.24} = 1.46$$

For the InGaAsP device:

$$J_{th}\ (20\,^{\circ}\mathrm{C}) \propto \exp \frac{293}{55} = 205.88$$

$$J_{th}\ (80\,^{\circ}\mathrm{C}) \propto \exp \frac{353}{55} = 612.89$$

Hence the ratio of the current densities:

$$\frac{J_{th}\ (80\,^{\circ}\mathrm{C})}{J_{th}\ (20\,^{\circ}\mathrm{C})} = \frac{612.89}{205.88} = 2.98$$

Thus in Example 6.6 the threshold current density for the AlGaAs device increases by a factor of 1.5 over the temperature range, whereas the threshold current density for the InGaAsP device increases by a factor of 3. Hence the stronger dependence of threshold current on temperature for InGaAsP structures is shown in this comparison of two average devices. It may also be noted that it is important to obtain high values of T_0 for the devices in order to minimize temperature dependence.

The increased temperature dependence for the InGaAsP/InP material system is also displayed by the more advanced, mode-stabilized device structures. Figure 6.36 provides the light output against current characteristic at various device temperatures for a strongly index-guided DCPBH injection laser (see Section 6.5.2) emitting at a wavelength of 1.55 μm [Ref. 13]. Moreover, the similar characteristic for a BH–DFB laser is shown in Figure 6.32(b). It is therefore necessary to pay substantial attention to thermal dissipation in order to provide efficient heat sinking arrangements (e.g. thermoelectric cooling, etc.) to achieve low operating currents. In addition, the need to minimize or eliminate the thermal resistance degradation

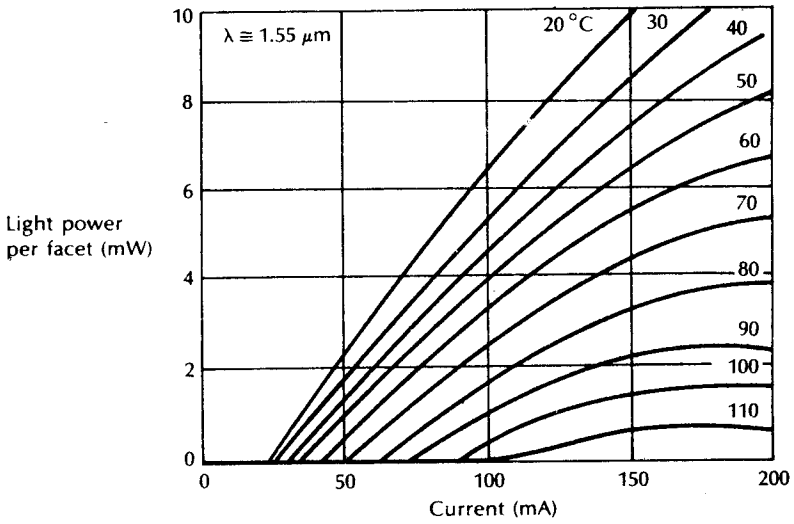


Figure 6.36 Light output versus current characteristics at various temperatures for a InGaAsP double channel planar BH laser emitting at a wavelength of 1.55 μm. Reproduced with permission from N. K. Dutta, 'Optical sources for lightwave system applications', in E. E. Basch (Ed.), *Optical-Fiber Transmission*, H. W. Sams & Co., p. 265, 1987.

associated with the solder bond on such devices (an effect which could, to a certain extent, be tolerated with GaAs injection lasers) has also become critically important [Ref. 42]. In all cases, however, adequate heat sinking along with consideration of the working environment are essential so that devices operate reliably over the anticipated current range.

6.7.2 Dynamic response

The dynamic behaviour of the injection laser is critical, especially when it is used in high bit rate (wideband) optical fiber communication systems. The application of a current step to the device results in a switch-on delay, often followed by high frequency (of the order of 10 GHz) damped oscillations known as relaxation oscillations (RO). These transient phenomena occur whilst the electron and photon populations within the structure come into equilibrium and are illustrated in Figure 6.37. In addition, when a current pulse reaches a laser which has significant parasitic capacitance after the initial delay time, the pulse will be broadened because the capacitance provides a source of current over the period that the photon density is high. Consequently, the injection laser output can comprise several pulses as the electron density is repetitively built up and quickly reduced, thus causing ROs. The switch-on delay t_d may last for 0.5 ns and the RO for perhaps twice that period. At data rates above 100 Mbit s^{-1} this behaviour can produce a serious deterioration

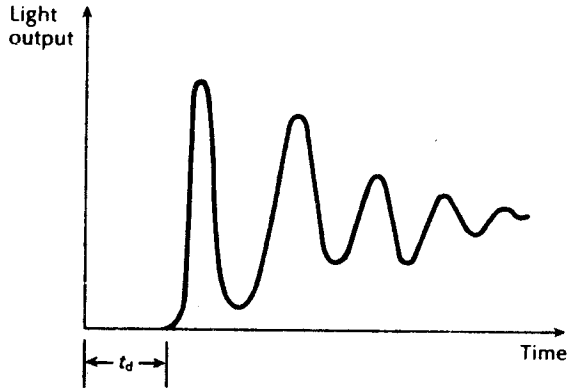


Figure 6.37 The possible dynamic behaviour of an injection laser showing the switch-on delay and relaxation oscillations.

in the pulse shape. Hence, reducing t_d and damping the relaxation oscillations is highly desirable.

The switch-on delay is caused by the initial build-up of photon density resulting from stimulated emission. It is related to the minority carrier lifetime and the current through the device [Ref. 7]. The current term, and hence the switch-on delay, may be reduced by biasing the laser near threshold (prebiasing). However, damping of the ROs is less straightforward. They are basic laser phenomena which vary with device structure and operating conditions; however, RO damping has been observed, and is believed to be due to several mechanisms including lateral carrier diffusion [Refs. 43 and 44], the feeding of the spontaneous emission into the lasing mode [Ref. 45] and gain nonlinearities [Ref. 46]. Narrow stripe geometry DH lasers and all the mode stabilized devices (see Sections 6.5 and 6.6) give RO damping, but it tends to coincide with a relatively slow increase in output power. This is thought to be the result of lateral carrier diffusion due to lack of lateral carrier confinement. However, it appears that RO damping and fast response may be obtained in BH structures with stripe widths less than the carrier diffusion length (i.e. less than 3 μm) [Ref. 47]. Moreover, this phenomenon has been employed within a digital transmission system by biasing the laser near threshold and then by using a single RO as a 'one' bit [Ref. 48].

6.7.3 Frequency chirp

The direct current modulation of a single longitudinal mode semiconductor laser can cause a dynamic shift of the peak wavelength emitted from the device [Ref. 49]. This phenomenon, which results in dynamic linewidth broadening under the direction modulation of the injection current, is referred to as frequency chirping. It arises from gain-induced variations in the laser refractive index due to the strong

coupling between the free carrier density and the index of refraction which is present in any semiconductor structure. Hence, even small changes in carrier density, apart from producing relaxation oscillations in the device output, will also result in a phase shift of the optical field, giving an associated change in the resonance frequency within both Fabry–Perot and DFB laser structures.

The laser linewidth broadening or chirping combined with the chromatic dispersion characteristics of single-mode fibers (see Section 3.9) can cause a significant performance degradation within high transmission rate systems [Ref. 50]. In particular, it may result in a shift in operating wavelength from the zero-dispersion wavelength of the fiber, which can ultimately limit the achievable system performance. For example, theoretical predictions [Ref. 51] of the wavelength shift that may occur with an InGaAsP laser under modulation of a few gigabit s^{-1} are around 0.05 nm (6.4 GHz frequency shift).

A number of techniques can be employed to reduce frequency chirp. One approach is to bias the laser sufficiently above threshold so that the modulation current does not drive the device below the threshold where the rate of change of optical output power varies rapidly with time. Unfortunately, this strategy gives an extinction ratio penalty (see Section 11.2.1.6) of the order of several decibels at the receiver. Another method involves the damping of the relaxation oscillations that can occur at turn-on and turn-off which result in large power fluctuations. This has been achieved, for instance, by shaping the electrical drive pulses [Ref. 52].

Certain device structures also prove advantageous for chirp reduction. In particular, quantum-well lasers (see Section 6.5.3), Bragg wavelength detuned DFB lasers (see Section 6.6.2) and multielectrode DFB lasers (see Section 6.10.2) provide improved performance under direct current modulation in relation to frequency chirping. Such lasers, however, require complex fabricational processes. An alternative technique which has proved effective in minimizing the effects of chirp is to allow the laser to emit continuously and to impress the data on to the optical carrier using an external modulator [Ref. 51]. Such devices, which may be separate lithium niobate-based components or can be monolithically integrated with the laser [Ref. 53], are described in Chapter 10.

6.7.4 Noise

Another important characteristic of injection laser operation involves the noise behaviour of the device. This is especially the case when considering analog transmission. The sources of noise are:

(a) phase or frequency noise;
(b) instabilities in operation such as kinks in the light output against current characteristic (see Section 6.5.1) and self pulsation;
(c) reflection of light back into the device;
(d) mode partition noise.

It is possible to reduce, if not remove (b), (c) and (d) by using mode stabilized

devices and optical isolators. Phase noise, however, is an intrinsic property of all laser types. It results from the discrete and random spontaneous or simulated transitions which cause intensity fluctuations in the optical emission and are an inevitable aspect of laser operation. Each event causes a sudden jump (of random magnitude and sign) in the phase of the electromagnetic field generated by the device. It has been observed that the spectral density of this phase or frequency noise has a characteristic represented by $1/f$ to $1/f^2$ up to a frequency (f) of around 1 MHz, as illustrated in Figure 6.38 [Ref. 55].

At frequencies above 1 MHz the noise spectrum is flat or white and is associated with quantum fluctuations (sometimes referred to as quantum noise, see Figure 6.38) which are a principal cause of linewidth broadening within semiconductor lasers [Ref. 56]. Although the low frequency components can easily be tracked and therefore are not a significant problem within optical fiber communications, this is not the case for the white noise component where as time elapses the phase executes a random walk away from the value it would have had in the absence of spontaneous emission.

For injection lasers operating at frequencies less than 100 MHz quantum noise levels are usually low (signal to noise ratios less than -80 dB) unless the device is biased within 10% of threshold. Over this region the noise spectrum is flat. However, for wideband systems when the laser is operating above threshold, quantum noise becomes more pronounced. This is especially the case with multimode devices (signal to noise ratios of around -60 dB). The higher noise level would appear to result from a peak in the noise spectrum due to a relaxation resonance which typically occurs between 200 MHz and 1 GHz [Ref. 7]. Single-mode lasers have demonstrated greater noise immunity by as much as 30 dB when the current is raised above threshold [Ref. 57]. Nevertheless, the wandering of the phase determines both the laser linewidth and the coherence time which are both major considerations, particularly within coherent optical fiber communications [Ref. 58].

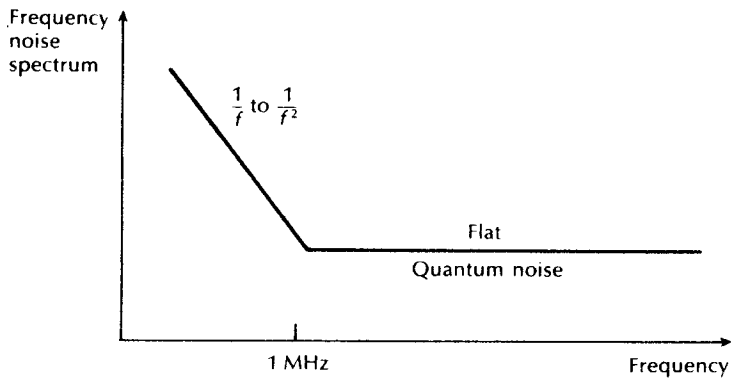


Figure 6.38 Spectral characteristic showing injection laser phase noise.