material, they are collected somewhat slowly by the diffusion process. This has the effect of producing a long 'diffusion tail' on a short optical pulse. When the APD is fully depleted by employing electric fields in excess of 10^4 V m^{-1}, all the carriers drift at saturation-limited velocities. In this case the response time for the device is limited by three factors. These are:

(a) the transit time of the carriers across the absorption region (i.e. the depletion width);
(b) the time taken by the carriers to perform the avalanche multiplication process; and
(c) the RC time constant incurred by the junction capacitance of the diode and its load.

At low gain the transit time and RC effects dominate giving a definitive response time and hence constant bandwidth for the device. However, at high gain the avalanche build-up time dominates and therefore the device bandwidth decreases proportionately with increasing gain. Such APD operation is distinguished by a constant gain–bandwidth product.

Often an asymmetric pulse shape is obtained from the APD which results from a relatively fast rise time as the electrons are collected and a fall time dictated by the transit time of the holes travelling at a slower speed. Hence, although the use of suitable materials and structures may give rise times between 150 and 200 ps, fall times of 1 ns or more are quite common which limit the overall response of the device.

8.9.2 Silicon reach through avalanche photodiodes

To ensure carrier multiplication without excess noise for a specific thickness of multiplication region within the APD it is necessary to reduce the ratio of the ionization coefficients for electrons and holes k (see Section 9.3.4). In silicon this ratio is a strong function of the electric field varying from around 0.1 at 3×10^5 V m^{-1} to 0.5 at 6×10^5 V m^{-1}. Hence for minimum noise, the electric field at avalanche breakdown must be as low as possible and the impact ionization should be initiated by electrons. To this end a 'reach through' structure has been implemented with the silicon avalanche photodiode. The silicon 'reach through' APD(RAPD) consists of p^+–π–p–n^+ layers as shown in Figure 8.13(a). As may be seen from the corresponding field plot in Figure 8.13(b), the high field region where the avalanche multiplication takes place is relatively narrow and centred on the p–n^+ junction. Thus under low reverse bias most of the voltage is dropped across the p–n^+ junction.

When the reverse bias voltage is increased the depletion layer widens across the p region until it 'reaches through' to the nearly intrinsic (lightly doped) π region. Since the π region is much wider than the p region the field in the π region is much lower than that at the p–n^+ junction (see Figure 8.13(b)). This has the effect of removing some of the excess applied voltage from the multiplication region to the

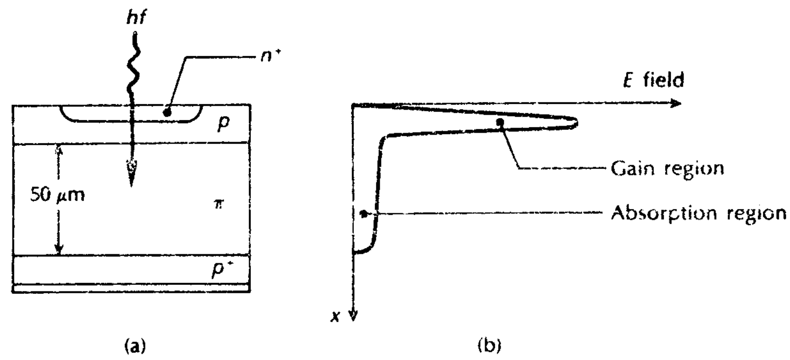


Figure 8.13 (a) Structure of a silicon RAPD. (b) The field distribution in the RAPD showing the gain region across the p–n^+ junction.

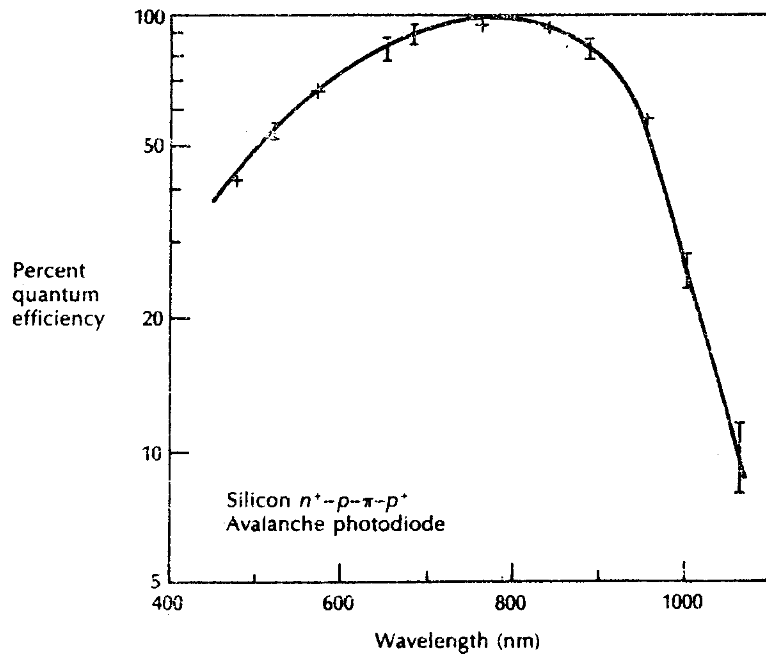


Figure 8.14 Measurement of quantum efficiency against wavelength for a silicon RAPD. After Ref. 16. Reprinted with permission from *The Bell System Technical Journal*. © 1978, AT&T.

π region giving a relatively slow increase in multiplication factor with applied voltage. Although the field in the π region is lower than in the multiplication region it is high enough (2×10^4 V cm^{-1}) when the photodiode is operating to sweep the carriers through to the multiplication region at their scattering limited velocity (10^7 cm s^{-1}). This limits the transit time and ensures a fast response (as short as 0.5 ns).

Measurements [Ref. 16] for a silicon RAPD for optical fiber communication applications at a wavelength of 0.825 μm have shown a quantum efficiency (without avalanche gain) of nearly 100% in the working region, as may be seen in Figure 8.14. The dark currents for this photodiode are also low and depend only slightly on bias voltage.

8.9.3 Germanium avalanche photodiodes

The elemental semiconductor germanium has been used to fabricate relatively sensitive and fast APDs that may be used over almost the entire wavelength range of primary interest at present (0.8–1.6 μm). However, it was clear from an early stage that higher dark currents together with larger excess noise factors (see Section 9.3.3) than those in silicon APDs were a problem with these devices. The large dark currents were associated with edge and surface effects resulting from difficulties in passivating germanium, and were also a direct consequence of the small energy bandgap as mentioned earlier in Section 8.4.2.

In the late 1970s when interest increased in the fabrication of detectors for longer wavelength operation (1.1–1.6 μm), germanium APDs using a conventional n^+p structure similar to the silicon APD shown in Figure 8.12 were produced [Ref. 18]. However, such devices exhibited dark currents near breakdown of between 100 nA and 300 nA which were very sensitive to temperature variations [Ref. 6]. Furthermore, unlike the situation with silicon APDs, these dark currents had significant components of both bulk (multiplied) and surface (unmultiplied) current. It was the multiplied component (typically 100 nA for the n^+p structure) which needed to be reduced (to around 1 nA) in order to provide low noise operation. In addition, large excess noise factors associated with the avalanche multiplication process were obtained as a result of electrons rather than holes (which have a higher impact ionization coefficient in germanium) initiating the multiplication process. One advantage, however, of such germanium APDs over their silicon counterparts is that because of the relatively high absorption coefficient exhibited by germanium at 1.3 μm, avalanche breakdown voltages are quite low (typically 25 V).

Germanium APD structures have been fabricated to provide multiplication initiated by holes thus to reduce the excess noise factor in the longer wavelength region. For example, a n^+np structure has been demonstrated [Ref. 19] which goes some way to achieving this performance by reducing the factor by some 30% on that obtained in n^+p devices. However, multiplied dark current around 1 μA was obtained when operating at a wavelength of 1.3 μm and a multiplication factor of

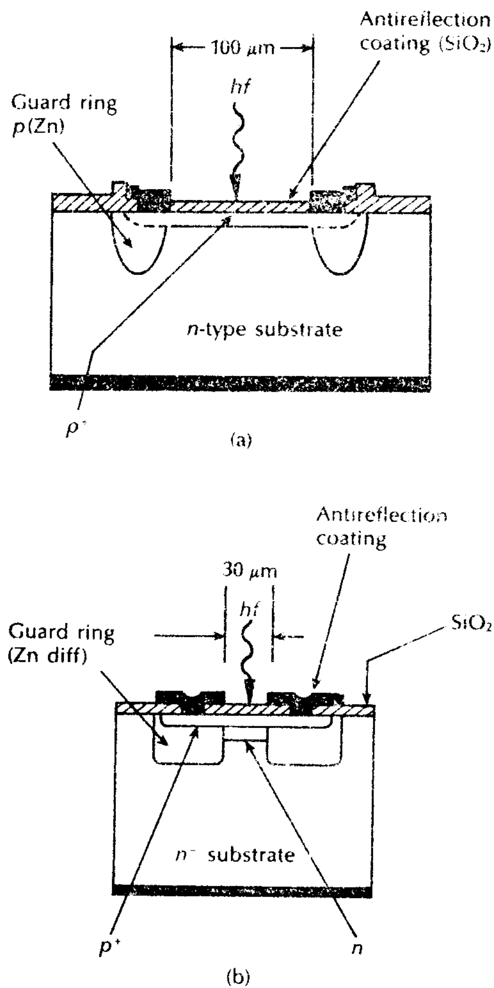


Figure 8.15 Germanium APDs: (a) p^+n structure [Ref. 20]; (b) Hi-Lo (p^+nn^-) structure [Ref. 23].

ten. An alternative device providing similar results utilizes the p^+n structure [Ref. 20] shown in Figure 8.15(a). In this case dark currents were reduced to between 150 and 250 nA by using an ion implanted technology [Ref. 21] and subsequently to around 5 nA by reducing the device sensitive area from 100 μm to 30 μm [Ref. 22].

Unfortunately, the speed of the p^+n structure at a wavelength of 1.5 μm is poor because most of the absorption in germanium at this wavelength takes place outside the depletion region.* This has led to the development of the p^+nn^- structure

* The absorption length in germanium at a wavelength of 1.5 μm is 10 μm.

shown in Figure 8.15(b) [Ref. 23] which resembles the reach through structure used for silicon APDs (see Section 8.9.2). It is known as a Hi–Lo structure as it combines high bandwidth (700 MHz) with low multiplied dark current (33 nA) and good excess noise performance. However, the breakdown voltage is higher at +85 V and the unmultiplied dark currents are around 1 μA. Nevertheless, these Hi–Lo devices appear to be among the highest performance germanium APDs for longer wavelength operation and are only eclipsed by the emerging III–V alloy APDs which do not exhibit quite the same fundamental material limitations.

8.9.4 III–V alloy avalanche photodiodes

Due to the drawbacks with germanium APDs for longer wavelength operation much effort has been expended in the study of III–V semiconductor alloys for the fabrication of APDs. In particular, the ternary InGaAs/InP and quaternary InGaAsP/InP material systems have been successfully employed. In common with the silicon reach through APD (see Section 8.9.2) separate absorption and multiplication regions are provided, as illustrated in Figure 8.16. This defines the so-called SAM (separate absorption and multiplication) APD which is a heterostructure device designed so that the multiplication takes place in the InP *p–n* junction [Ref. 24]. The performance of such long wavelength APDs is limited, however, by the fundamental properties of the material systems.

A first limitation is related to the large tunnelling currents associated with the narrow bandgap required for longer wavelength optical absorption. The band to band or defect assisted tunnelling currents become large before the electric field is high enough to obtain significant avalanche gain. This problem is substantially reduced using a separate absorption and multiplication region with the gain

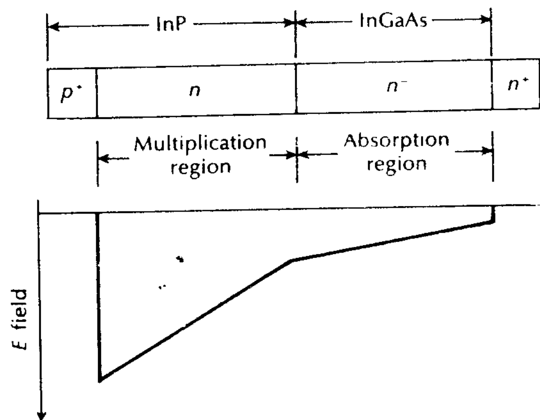


Figure 8.16 Separate absorption and multiplication (SAM) APD layer composition and electric field profile.

occurring at the InP p–n junction where the tunnelling is much less [Ref. 6]. However, control over the doping and thickness of the n type InP layer is critical in order to avoid excessive leakage current. Nevertheless, it is possible to obtain low dark currents of less than 10 nA (unmultiplied) together with quantum efficiencies of 80%, capacitance of approximately 0.5 pF and an operating voltage of around 100 V.

A second limitation associated with SAM APDs concerns the trapping of holes in the valence band discontinuity at the InGaAs/InP heterointerface, as illustrated in Figure 8.17(a) [Ref. 25]. This factor results in a slow component of the photoresponse which causes a speed limitation. However, the problem can be alleviated by incorporating a thin grading layer of InGaAsP (whose bandgap is intermediate between InGaAs and InP) between these two layers (Figure 8.17(b)) to smooth out the discontinuity and thus provide improved speed performance [Ref. 26]. Nevertheless, the gain–bandwidth products for such devices are still only between 10 and 20 GHz, not quite sufficient for high bit rate systems in the gigabit s^{-1} region [Ref. 6]. For example, with a gain around ten such devices will only provide operation to between 1 and 2 GHz.

An improved technique for increasing the speed of response of the device is to provide several (two or three) InGaAsP buffer layers to create compositional grading at the heterojunction interface [Refs. 12, 24]. This may be achieved by interposing a thin multiquantum-well (MQW) structure between the narrow and wideband gap layer. The configuration of a recent back illuminated mesa-structure separate absorption, grading and multiplication (SAGM) InGaAs APD is shown in Figure 8.18. This device type has displayed a gain–bandwidth product of up to 70 GHz [Ref. 27] thus allowing operation at bandwidths of 5 GHz or higher.

Overall, advanced photodiode developments are targeted at devices with improved sensitivities for operation at very high bandwidths, together with the

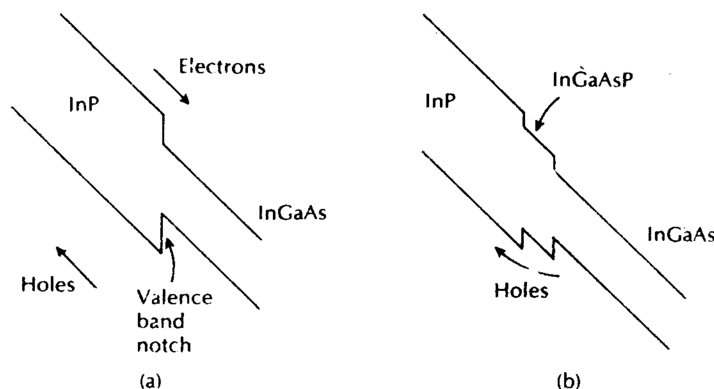


Figure 8.17 Energy band diagrams for SAM APDs: (a) InGaAs/InP heterojunction illustrating the notch in which holes may be trapped; (b) similar heterojunction to (a) with InGaAsP layer to reduce the effect of the notch.

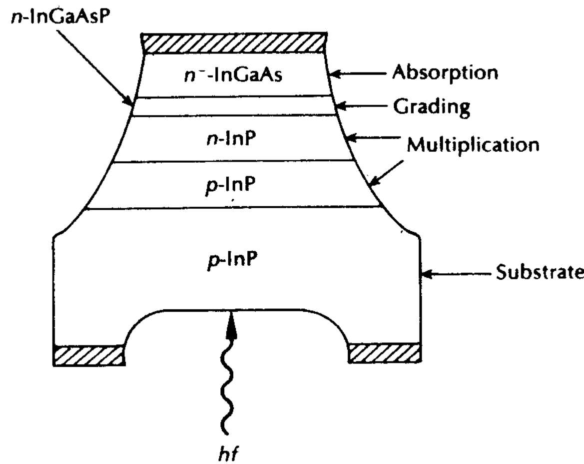


Figure 8.18 Separate absorption, grading and multiplication (SAGM) APD structure.

fabrication of structures with improved functionality at low cost [Ref. 12]. More recently, efforts to improve sensitivity have focused upon semiconductor superlattices in the form of MQW structures [Ref. 29] and staircase APDs [Ref. 30]. Both of these APD structures have gain regions comprising multiquantum wells formed by alternately growing thin layers of wide and narrow bandgap materials such as AlGaAs and GaAs respectively. By using materials exhibiting these properties the conduction and valence band discontinuities differ significantly, resulting in different ionization coefficients for electrons and holes. This factor should therefore give improvements in the noise performance of such III–V alloy APDs by reducing the ratio of the ionization coefficients for electrons and holes (k value, see Section 9.3.4) because the ionization coefficients of the two carrier types are normally approximately equal. In addition the other major advantage in using MQW APDs results from their improved bandwidth capabilities caused by the reduction in avalanche build-up time provided by the multilayer structures.

The step-like MQW energy band structure where the discontinuity in the conduction band is greater than that in the valence band is shown in Figure 8.19. The structure, which is illustrated both unbiased and biased, could comprise about 100 layers of alternate wide and narrow bandgap semiconductors. Although such devices have been fabricated using the AlGaAs/GaAs material system by molecular beam epitaxy (MBE), the structure does not provide the same favourable k value reduction when using InP-based alloys for longer wavelength operation. In this case problems with tunnelling in the high field regions containing the narrow bandgap layers tend to destroy the sensitivity improvement provided by the multiquantum wells [Ref. 12].

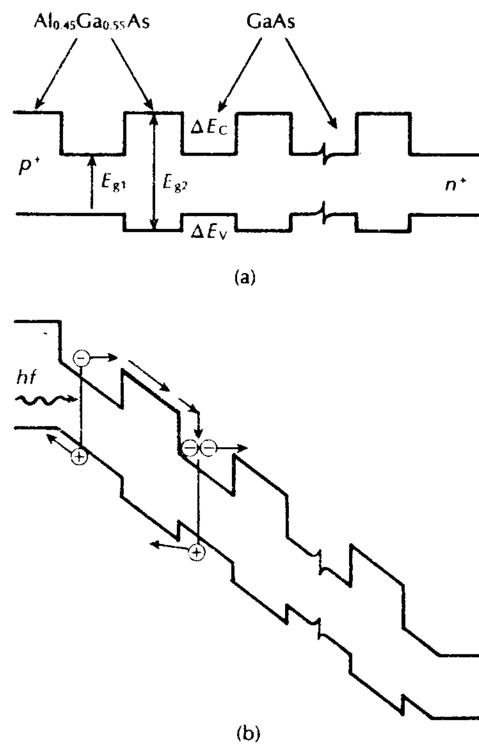


Figure 8.19 Energy band diagrams for MQW superlattice APD structure: (a) unbiased showing alternate layers of wide and narrow bandgap semiconductors; (b) the biased device [Ref. 29].

The energy band structure of a more complex scheme known as a staircase APD is shown in Figure 8.20. In this technique a narrow bandgap region is compositionally graded over a distance of 10 to 20 nm into a material with a minimum of twice the bandgap at the narrow end of the step. Again, the composition is abruptly changed to obtain the narrow bandgap as a second step is formed. The primary advantage of this staircase structure is that carrier multiplication caused by carrier transitions from the wide to the narrow bandgap material can occur at much lower electric field densities than that required with MQW devices.

In principle this APD could operate with a very low bias voltage, thereby removing the possibility of tunnelling in the narrow bandgap material. However, the structure and grading presents substantial fabricational problems and therefore has not, as yet, been realized. Moreover, it is suggested that the only suitable candidate material with the ability to grade continuously in alloy composition is GaAlAsSb [Ref. 31].

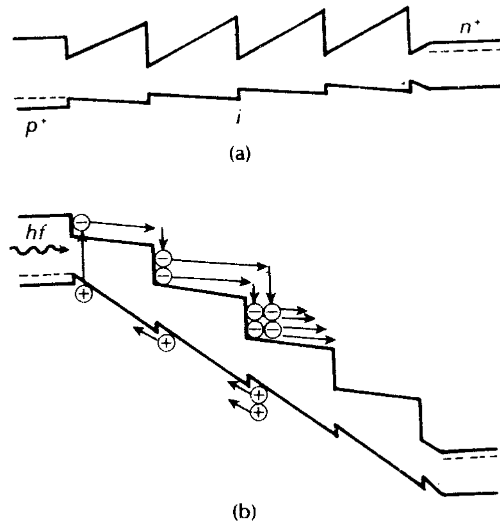


Figure 8.20 Energy band diagrams for the staircase APD: (a) the unbiased device: (b) the biased device under normal operation [Ref. 30].

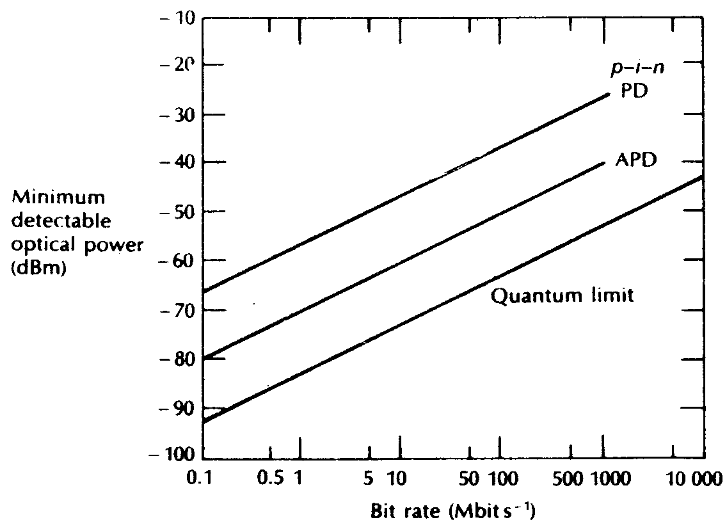
8.9.5 Benefits and drawbacks with the avalanche photodiode

APDs have a distinct advantage over photodiodes without internal gain for the detection of the very low light levels often encountered in optical fiber communications. They generally provide an increase in sensitivity of between 5 and 15 dB over *p–i–n* photodiodes whilst often giving a wider dynamic range as a result of their gain variation with response time and reverse bias.

The optimum sensitivity improvement of APD receivers over *p–i–n* photodiode devices is illustrated in the characteristics shown in Figure 8.21. The characteristics display the minimum detectable optical power for direct detection (see Section 7.5) versus the transmitted bit rate in order to maintain a bit error rate (BER) of 10^{-9} (see Section 11.6.3) in the shorter and longer wavelength regions. Figure 8.21(a) compares silicon photodiodes operating at a wavelength of 0.82 μm where the APD is able to approach within 10 to 13 dB of the quantum limit. In addition, it may be observed that the *p–i–n* photodiode receiver has a sensitivity around 15 dB below this level. InGaAs photodiodes operating at a wavelength of 1.55 μm are compared in Figure 8.21(b). In this case the APD requires around 20 dB more power than the quantum limit, whereas the *p–i–n* photodiode receiver is some 10 to 12 dB less sensitive than the APD.

APDs, however, also have several drawbacks which include:

(a) fabrication difficulties due to their more complex structure and hence increased cost;



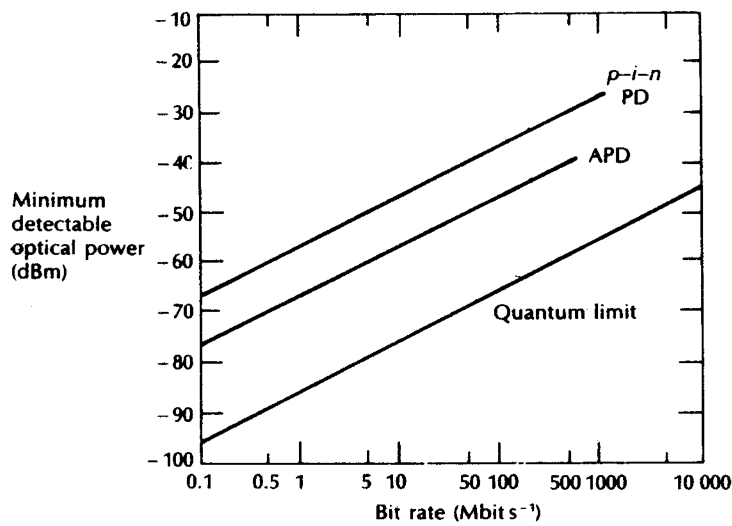


Figure 8.21 Receiver sensitivity comparison of *p–i–n* photodiode and APD devices at BER of 10^{-9}: (a) using silicon detectors operating at a wavelength of 0.82 μm; (b) using InGaAs detectors operating at a wavelength of 1.55 μm.

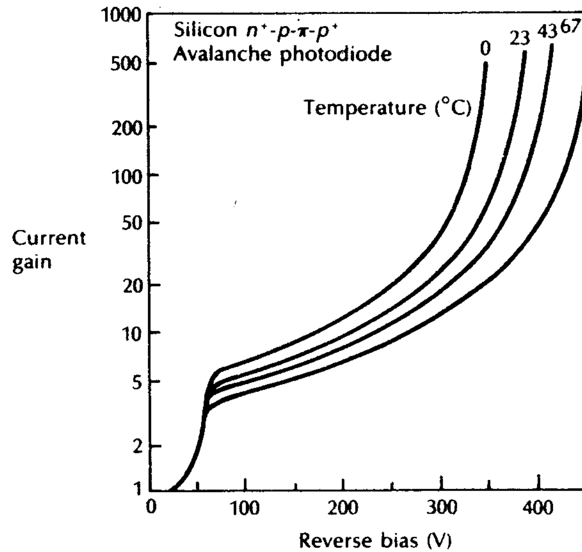


Figure 8.22 Current gain against reverse bias for a silicon RAPD operating at a wavelength of 0.825 μm. After Ref. 16. Reprinted with permission from *The Bell System Technical Journal*. © 1978, AT&T.

(b) the random nature of the gain mechanism which gives an additional noise contribution (see Section 9.3.3);

(c) the often high bias voltages required (50 to 400 V) which are wavelength dependent;

(d) the variation of the gain (multiplication factor) with temperature as shown in Figure 8.22 for a silicon RAPD [Ref. 16]; thus temperature compensation is necessary to stabilize the operation of the device.

8.9.6 Multiplication factor

The multiplication factor M is a measure of the internal gain provided by the APD. It is defined as:

$$M = \frac{I}{I_p} \tag{8.30}$$

where I is the total output current at the operating voltage (i.e. where carrier multiplication occurs) and I_p is the initial or primary photocurrent (i.e. before carrier multiplication occurs).

Example 8.6

The quantum efficiency of a particular silicon RAPD is 80% for the detection of radiation at a wavelength of 0.9 μm. When the incident optical power is 0.5 μW, the output current from the device (after avalanche gain) is 11 μA. Determine the multiplication factor of the photodiode under these conditions.

Solution: From Eq. (8.11), the responsivity

$$R = \frac{\eta e \lambda}{hc} = \frac{0.8 \times 1.602 \times 10^{-19} \times 0.9 \times 10^{-6}}{6.626 \times 10^{-34} \times 2.998 \times 10^{8}}$$

$$= 0.581 \text{ A W}^{-1}$$

Also from Eq. (8.4), the photocurrent

$$I_p = P_o R$$

$$= 0.5 \times 10^{-6} \times 0.581$$

$$= 0.291 \ \mu\text{A}$$

Finally, using Eq. (8.30):

$$M = \frac{I}{I_p} = \frac{11 \times 10^{-6}}{0.291 \times 10^{-6}}$$

$$= 37.8$$

The multiplication factor of the photodiode is approximately 38.

8.10 Mid-infrared photodiodes

Developments of photodiodes for mid-infrared transmission systems are at a relatively early stage; however, several potential devices have been demonstrated over recent years. Obtaining suitable lattice matching for III–V alloy materials is a problem when operating at wavelengths greater than 2.0 μm. A lattice matched InGaAsSb/GaSb material system has been utilized in a *p–i–n* photodiode for high speed operation at wavelengths up to 2.3 μm [Ref. 32].

An alternative approach which has achieved some success is the use of indium alloys that, due to the high indium content for operation above 2 μm, are mismatched with respect to the InP substrate causing inherent problems of dislocation-induced junction leakage and low quantum efficiency [Ref. 33]. However, these problems have been reduced by utilizing a compositionally graded buffer layer to accommodate the lattice mismatch. One technique has involved the replacement of the conventional *p–i–n* homojunction with an InGaAs/AlInAs heterojunction in which a wider bandgap *p*-type AlInAs layer acted as a transparent window at long wavelengths to ensure that optical absorption occurred in a lightly doped *n* type region of the device. This device, which exhibited a useful response

out to a wavelength of 2.4 μm, displayed a quantum efficiency as high as 95% over the wavelength range 1.3 to 2.25 μm with dark currents as low as 35 nA [Ref. 33]. A similar approach has been demonstrated with the ternary alloys $In_xGa_{1-x}As$/$InAs_yP_{1-y}$ to produce mesa structure photodiodes which operate at a wavelength of 2.55 μm [Ref. 34]. A compositionally graded region of InGaAs or InAsP accommodates the lattice mismatch between the ternary layers and the InP substrate. These devices which have $In_{0.85}Ga_{0.15}As$ absorbing layers and lattice matched $InAs_{0.68}P_{0.32}$ capping layers have displayed a quantum efficiency of 52% at a wavelength of 2.55 μm but with dark currents of between 10 and 20 μA. It is suggested [Ref. 34] that these high dark currents are a result of electrically active defects associated with misfit dislocations in the active regions of the diode which can be reduced by improved materials growth.

In common with injection lasers, the HgCdTe material system has been utilized to fabricate long wavelength photodiodes (see Section 6.11) $Hg_{1-x}Cd_xTe$ ternary alloys form a continuous family of semiconductors whose bandgap energy variation with x enables optical detection from 0.8 μm to the far-infrared. Furthermore, with this material system the hole ionization coefficient exhibits a resonant characteristic which is a function of the alloy composition [Ref. 31]. This phenomenon is a band structure effect which is also displayed by the $Ga_{1-x}Al_xSb$ alloys [Ref. 34]. Resonant impact ionization processes in such materials yield a high ionization coefficient ratio* providing enhanced sensitivity at particular operating wavelengths. However, for the HgCdTe material system this occurs for a composition in the vicinity of $x = 0.7$, which is suitable for detection at a wavelength of 1.3 μm. Hence both HgCdTe *p–i–n* photodiodes and APDs that exhibit low sensitivity and high speed response (500 MHz) have been produced for operation at this wavelength [Ref. 36]. In addition, $Hg_{0.4}Cd_{0.6}Te$ APDs have been fabricated for detection at a wavelength of 1.55 μm [Ref. 37]. With this material system, the potential remains, however, to provide even longer wavelength operation into the mid-infrared region, although this would be greatly assisted by improvements in the material quality to facilitate higher performance devices [Ref. 31].

8.11 Phototransistors

The problems encountered with APDs for use in the longer wavelength region stimulated a renewed interest in bipolar phototransistors in the late 1970s. Hence, although these devices have been investigated for a number of years, they have yet to find use in major optical fiber communication systems. In common with the APD the phototransistor provides internal gain of the photocurrent. This is achieved through transistor action rather than avalanche multiplication. A symbolic representation of the *n–p–n* bipolar phototransistor is shown in Figure 8.23(a). It differs from the conventional bipolar transistor in that the base is unconnected, the

* In this case the ratio of hole ionization coefficient to electron ionization coefficient.

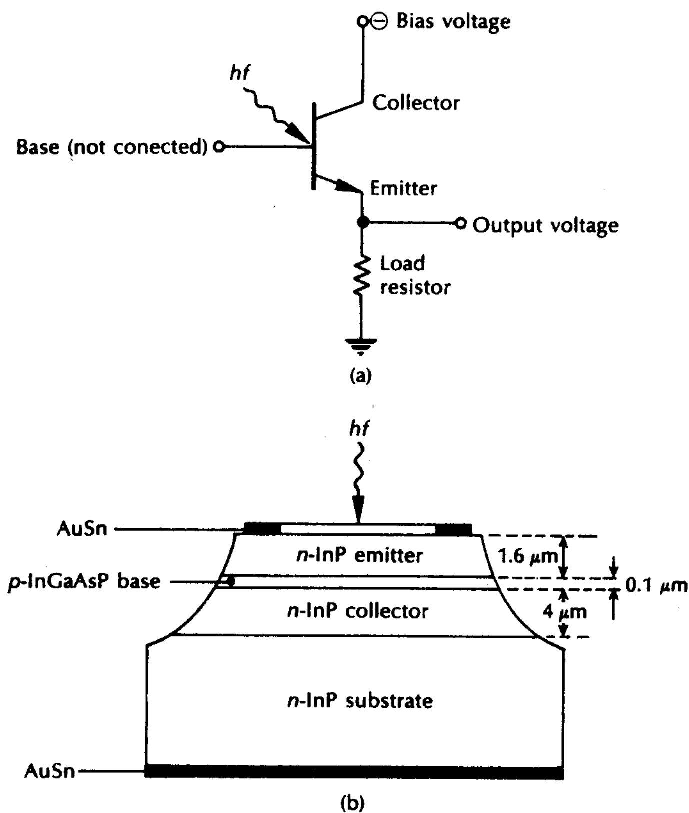


Figure 8.23 (a) Symbolic representation of the *n–p–n* phototransistor showing the external connections. (b) Cross section of an *n–p–n* InGaAsP/InP heterojunction phototransistor [Ref. 38].

base–collector junction being photosensitive to act as a light-gathering element. Thus absorbed light affects the base current giving multiplication of primary photocurrent through the device.

The structure of a *n–p–n* InGaAsP/InP heterojunction phototransistor is shown in Figure 8.23(b) [Ref. 38]. The three layer heterostructure (see Section 6.3.5) is grown on an InP substrate using liquid-phase epitaxy (LPE). It consists of an *n* type InP collector layer followed by a thin (0.1 μm)*p* type InGaAsP base layer. The third layer is a wide bandgap *n* type InP emitter layer. Radiation incident on the device passes unattenuated through the wide bandgap emitter and is absorbed in the base, base–collector depletion region and the collector. A large secondary photocurrent between the emitter and collector is obtained as the photogenerated holes are swept into the base, increasing the forward bias on the device. The use of the heterostructure permits low emitter–base and collector–base junction capacitances

together with low base resistance. This is achieved through low emitter and collector doping levels coupled with heavy doping of the base, and allows large current gain. In addition the potential barrier created by the heterojunction at the emitter–base junction effectively eliminates hole injection from the base when the junction is forward biased. This gives good emitter base injection efficiency. The optical gain G_o of the device is given approximately by [Ref. 38]:

$$G_o \simeq \eta h_{FE} = \frac{hf}{e}\frac{I_c}{P_o} \qquad (8.31)$$

where η is the quantum efficiency of the base–collector photodiode, h_{FE} is the common emitter current gain, I_s is the collector current, P_o is the incident optical power, e is the electronic charge and hf is the photon energy.

The phototransistor shown in Figure 8.23(b) is capable of operating over the 0.9 to 1.3 μm wavelength band giving optical gains in excess of one hundred, as demonstrated in Example 8.7. Moreover, the InGaAs/InP heterojunction phototransistor has also been the subject of significant development work but it is not anticipated that these devices will replace photodiodes within high sensitivity, high transmission rate applications [Ref. 31].

Example 8.7

The phototransistor of Figure 8.23(b) has a collector current of 15 mA when the incident optical power at a wavelength of 1.26 μm is 125 μW. Estimate:

(a) the optical gain of the device under the above operating conditions;
(b) the common emitter current gain if the quantum efficiency of the base–collector photodiode at a wavelength of 1.26 μm is 40%.

Solution: (a) Using Eq. (8.31), the optical gain is given by:

$$G_o \simeq \frac{hf}{e}\frac{I_c}{P_o} = \frac{hc}{\lambda e}\frac{I_c}{P_o}$$

$$= \frac{6.626 \times 10^{-34} \times 2.998 \times 10^{8} \times 15 \times 10^{-3}}{1.26 \times 10^{-6} \times 1.602 \times 10^{-19} \times 125 \times 10^{-6}}$$

$$= 118.1$$

(b) The common emitter current gain is:

$$h_{FE} = \frac{G_o}{\eta} = \frac{118.1}{0.4} = 295.3$$

In this example a common emitter current gain of 295 gives an optical gain of 118. It is therefore possible that this type of device will become an alternative to the APD for optical detection at wavelengths above 1.1 μm [Refs. 39–41].

8.12 Photoconductive detectors

The photoconductive detector or photoconductor, which provides what is conceptually the simplest form of semiconductor optical detection, has not until recently been considered as a serious contender for photodetection within optical fiber communications. Lately, however, there has been renewed interest in such devices, particularly for use in the longer wavelength region because of the suitability of III–V semiconductors for photoconductive detection applications [Ref. 12]. The basic detection process in a semiconductor discussed in Section 8.3 indicated that an electron may be raised from the valence band to the conduction band by the absorption of a photon provided that the photon energy was greater than the bandgap energy (i.e. $hf \geqslant E_g$). In this case, as long as the electron remains in the conduction band it will cause an increase in the conductivity of the semiconductor, a phenomenon referred to as photoconductivity. This forms the basic mechanism for the operation of photoconductive detectors.

A typical photoconductor device structure designed for operation in the longer wavelength region is shown in Figure 8.24 [Ref. 12]. The conducting channel comprises a thin layer (1 to 2 μm) of n type InGaAs which can absorb a significant amount of the incident light over the wavelength range 1.1 to 1.6 μm. In particular, good sensitivities at reasonably high bandwidths have been obtained with lightly doped (less than 5×10^{14} cm^{-3}) n type $In_{0.53}Ga_{0.47}As$ channel layers. Moreover, the composition of the InGaAs is arranged to be lattice-matched to the semi-insulating InP substrate to avoid the formation of dislocations and other crystalline imperfections in the epitaxial layer. Low resistance contacts are made to the conducting layer through the use of interdigital anodes and cathodes, as illustrated in Figure 8.24. In addition, these contacts are designed to maximize the coupling of light into the absorbing region by minimizing their obstruction of the active area whilst reducing the distance that photogenerated carriers have to travel prior to being collected at one of the electrodes. The optical coupling efficiency can also be improved by the application of an antireflection coating to the surface of the photoconductor facing the optical input.

In operation the incident light on the channel region is absorbed, thereby generating additional electron–hole pairs. These photogenerated carriers increase the channel conductivity which results in an increased current in the external circuit. The optical receiver must therefore be sensitive to very small changes in resistance induced by the incident light. Furthermore, once the carriers have been generated the electrons will be swept by the applied electric field towards the anode whilst the holes move towards the cathode. In general, however, the mobility of the holes is considerably smaller than that of the electrons in III–V alloys such as GaAs and InGaAs. Thus the electrons, being the fastest charge carriers, provide the minimum time for detection and hence the limitation on the speed of response of the photoconductive detector.

In addition, whilst the fast electrons are collected at the anode, the corresponding holes are still proceeding across the channel. This creates an absence of electrons

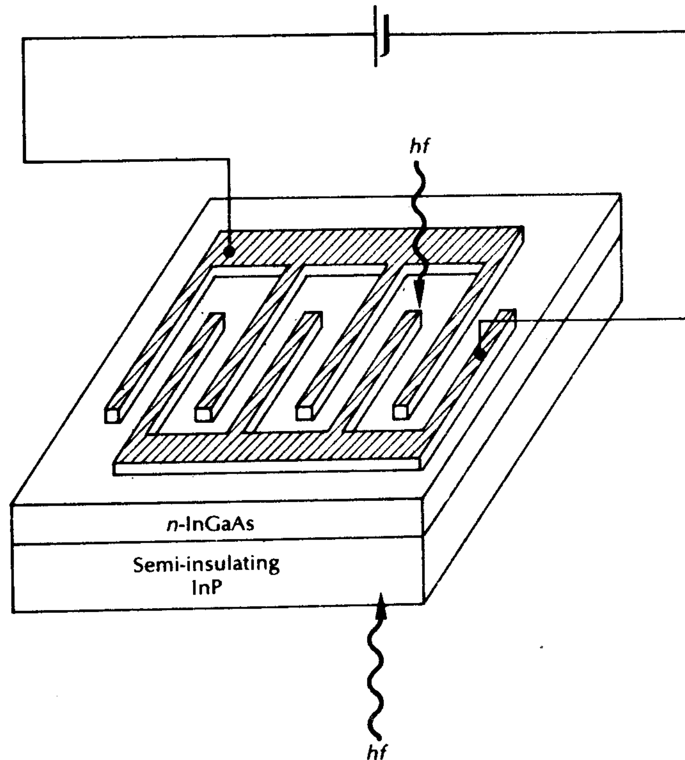


Figure 8.24 Photoconductive detector structure for operation in the 1.1 to 1.6 μm wavelength range.

and hence a net positive charge in the channel region. However, the excess charge is immediately compensated by the injection of further electrons from the cathode into the channel. Thus further electrons may be generated from the absorption of a single photon. This factor creates what is known as the photoconductive gain G which may be defined as the ratio of the slow carrier transit time (or lifetime) t_s to the fast carrier transit time t_f. Hence:

$$G = \frac{t_s}{t_f} \tag{8.32}$$

Moreover, the photocurrent I_p produced by the photoconductor following Eq. (8.8) can be written as [Ref. 42]:

$$I_p = \frac{\eta P_o e}{hf} G \tag{8.33}$$

where η is the device quantum efficiency which in a similar manner to the

absorption process (see Eq. (8.1)) follows an exponential distribution, P_o is the incident optical power and e is the charge on an electron.

Since the current response in the photoconductor remains for a time t_s after the end of an incident optical pulse and exhibits an exponential decay with a time constant equal to this slow carrier transit time (or lifetime), then the maximum 3 dB bandwidth B_m for the device is given by:

$$B_m = \frac{1}{2\pi t_s} \tag{8.34}$$

Equations (8.32) and (8.34) can be combined to provide an expression for the gain–bandwidth product of a photoconductor as:

$$G \cdot B_m = \frac{1}{2\pi t_f} \tag{8.35}$$

It may be noted that an implication of Eq. (8.35) is that when t_f is fixed, photoconductive gain can only be obtained at the expense of the maximum bandwidth permitted by the device. There is therefore a trade-off between gain (and hence sensitivity) and speed of response.

Example 8.8

The electron transit time in an InGaAs photoconductive detector is 5 ps. Determine the maximum 3 dB bandwidth permitted by the device when its photoconductive gain is 70.

Solution: Using Eq. (8.35) the maximum 3 dB bandwidth provided by the photoconductor may be written as:

$$B_m = \frac{1}{2\pi t_f G} = \frac{1}{2\pi \times 5 \times 10^{-12} \times 70}$$

$$= 454.7 \text{ MHz}$$

The result obtained in Example 8.8 typifies the bandwidth (i.e. less than 500 MHz) currently achievable with InGaAs photoconductors with gains in the range 50 to 100. By contrast, silicon photoconductors can provide very large gains of around 1000 but this is accompanied by a reduction in the 3 dB bandwidth obtained from these devices (response times typically in the range 1 μs to 1 ms).

The sensitivity of a photoconductive detector is limited by the noise generated within the device, even though the quantum efficiency may be quite high. For example, in an InGaAs photoconductor with a 2 μm channel the quantum efficiency is around 88%. However, there are numerous sources contributing to the noise

component within photoconductors [Ref. 43]. In particular, noise arises from two sources: Johnson noise associated with the thermal noise from the bulk resistance of the photoconductor slab, and generation–recombination noise caused by fluctuations in the generation and recombination rates of the photogenerated carrier pairs. The former noise source, which is often dominant, results in a finite dark conductivity for the device which generates a randomly varying background dark current. It can be shown [Ref. 12] that the signal to noise ratio of a photoconductor receiver increases with increasing channel resistance and gain. Therefore, a method for increasing the sensitivity of the photoconductor is to increase its photoconductive gain. Unfortunately, as indicated previously, this reduces the device response time.

Alternatively the photoconductor can be designed such that the dark current is reduced to a level such that the generation–recombination noise, which is fundamental, tends to dominate. Some success has been achieved in this direction using *p* type substrates [Ref. 44]. In this case, however, the speed of response for the device was limited by the electrode spacing required to improve the sensitivity. Improved performance has been obtained with a recent photoconductive-like detector which effectively combines a photoconductor and a *p–i–n* photodiode [Ref. 45]. This hybrid device type also illustrates another useful factor concerning photoconductive detectors in that they are very suitable for monolithic integration to provide optoelectronic integrated circuits (see Section 10.7).

Problems

8.1 Outline the reasons for the adoption of the materials and devices used for photodetection in optical fiber communications. Discuss in detail the *p–i–n* photodiode with regard to performance and compatibility requirements in photodetectors.

8.2 A *p–i–n* photodiode on average generates one electron–hole pair per three incident photons at a wavelength of 0.8 μm. Assuming all the electrons are collected calculate:

(a) the quantum efficiency of the device;
(b) its maximum possible bandgap energy;
(c) the mean output photocurrent when the received optical power is 10^{-7} W.

8.3 Explain detection process in the *p–n* photodiode. Compare this device with the *p–i–n* photodiode.

8.4 Define the quantum efficiency and the responsivity of a photodetector.

Derive an expression for the responsivity of an intrinsic photodetector in terms of the quantum efficiency of the device and the wavelength of the incident radiation.

Determine the wavelength at which the quantum efficiency and the responsivity are equal.

8.5 A *p–n* photodiode has a quantum efficiency of 50% at a wavelength of 0.9 μm. Calculate:

(a) its responsivity at 0.9 μm;
(b) the received optical power if the mean photocurrent is 10^{-6} A;
(c) the corresponding number of received photons at this wavelength.

8.6 When 800 photons per second are incident on a *p–i–n* photodiode operating at a wavelength of 1.3 μm they generate on average 550 electrons per second which are collected. Calculate the responsivity of the device.

8.7 Explain what is meant by the long wavelength cutoff point for an intrinsic photodetector, deriving any relevant expressions.

Considering the bandgap energies given in Table 8.1, calculate the long wavelength cutoff points for both direct and indirect optical transitions in silicon and germanium.

8.8 A *p–i–n* photodiode ceases to operate when photons with energy greater than 0.886 eV are incident upon it; of which material is it fabricated?

8.9 (a) The time taken for electrons to diffuse through a layer of *p* type silicon is 28.8 ns. If the minority carrier diffusion coefficient is 3.4×10^{-3} m^2 s^{-1}, determine the thickness of the silicon layer.

(b) Assuming the depletion layer width in a silicon photodiode corresponds to the layer thickness obtained in part (a) and that the maximum response time of the photodiode is 877 ps, estimate the carrier (hole) drift velocity.

8.10 A silicon *p–i–n* photodiode with an area of 1.5 mm^2 is to be used in conjunction with a load resistor of 100 Ω. If the requirement for the device is a fast response time, estimate the thickness of the intrinsic region that should be provided. It may be assumed that the permittivity for silicon is 1.04×10^{-10} F m^{-1} and that the electron saturation velocity is 10^7 m s^{-1}.

8.11 Define the noise equivalent power (NEP) for a photodetector. Commencing with Eq. (8.8) obtain an expression for the NEP of a photodiode in which the dark current noise dominates.

A silicon *p–i–n* photodiode with active dimensions 10 μm has a specific detectivity of 7×10^{10} m Hz$^{\frac{1}{2}}$ W^{-1} when operating at a wavelength of 0.85 μm. The device quantum efficiency at this wavelength is 64%. Assuming that it is the dominant noise source, calculate the dark current over a 1 Hz bandwidth in the device.

8.12 The specific detectivity of a wide area silicon photodiode at its operating wavelength is 10^{11} m Hz$^{\frac{1}{2}}$ W^{-1}. Estimate the smallest detectable signal power at this wavelength when the sensitive area of the device is 25 mm^2 and the signal bandwidth is 1 kHz.

8.13 Discuss the operation of the silicon RAPD, describing how it differs from the *p–i–n* photodiode.

Outline the advantages and drawbacks with the use of the RAPD as a detector for optical fiber communications.

8.14 Compare and contrast the structure and performance characteristics of germanium and III–V semiconductor alloy APDs for operation in the wavelength range 1.1 to 1.6 μm.

8.15 An APD with a multiplication factor of 20 operates at a wavelength of 1.5 μm. Calculate the quantum efficiency and the output photocurrent from the device if its responsivity at this wavelength is 0.6 A W^{-1} and 10^{10} photons of wavelength 1.5 μm are incident upon it per second.

8.16 Given that the following measurements were taken for an APD, calculate the multiplication factor for the device.

Received optical power at 1.35 μm = 0.2 μW
Corresponding output photocurrent = 4.9 μA
(after avalanche gain)
Quantum efficiency at 1.35 μm = 40%

8.17 An APD has a quantum efficiency of 45% at 0.85 μm. When illuminated with radiation of this wavelength it produces an output photocurrent of 10 μA after avalanche gain with a multiplication factor of 250. Calculate the received optical power to the device. How many photons per second does this correspond to?

8.18 When 10^{11} photons per second each with an energy of 1.28×10^{-19} J are incident on an ideal photodiode, calculate:

(a) the wavelength of the incident radiation;
(b) the output photocurrent;
(c) the output photocurrent if the device is an APD with a multiplication factor of 18.

8.19 A silicon RAPD has a multiplication factor of 10^3 when operating at a wavelength of 0.82 μm. At this operating point the quantum efficiency of the device is 90% and the dark current is 1 nA.

Determine the number of photons per second of wavelength 0.82 μm required in order to register a light input to the device corresponding to an output current (after avalanche gain) which is greater than the level of the dark current (i.e. $I > 1$ nA).

8.20 Indicate the material systems under investigation and discuss their application in the fabrication of photodiodes for use in the mid-infrared wavelength region.

8.21 An InGaAsP heterojunction phototransistor has a common emitter current gain of 170 when operating at a wavelength of 1.3 μm with an incident optical power of 80 μW. The base collector quantum efficiency at this wavelength is 65%. Estimate the collector current in the device.

8.22 Describe the basic detection process in a photoconductive detector.

The maximum 3 dB bandwidth allowed by an InGaAs photoconductive detector is 380 MHz when the electron transit time through the device is 7.6 ps. Calculate the photocurrent obtained from the device when 10 μW of optical power at a wavelength of 1.32 μm is incident upon it, and the device quantum efficiency is 75%.

Answers to numerical problems

8.2 (a) 33%; (b) 24.8×10^{-20} J; (c) 21.3 nW
8.4 1.24 μm
8.5 (a) 0.36 A W^{-1}; (b) 2.78 μW; (c) 1.26×10^{13} photon s^{-1}
8.6 0.72 A W^{-1}
8.7 0.3 μm, 1.09 μm, 1.53 μm, 1.85 μm
8.8 $In_{0.7}Ga_{0.3}As_{0.64}P_{0.36}$
8.9 (a) 14 μm; (b) 10^5 ms^{-1}
8.10 395 μm
8.11 1.23×10^{-14} A
8.12 1.58 pW
8.15 50%, 15.9 nA
8.16 24.1
8.17 77.8 nW, 3.33×10^{11} photon s^{-1}
8.18 (a) 1.55 μm; (b) 1.6 μA; (c) 28.8 μA
8.19 6.94×10^6 photon s^{-1}
8.21 9.3 mA
8.22 0.44 mA

References

[1] H. Melchior, M. B. Fisher and F. R. Arams, 'Photodetectors for optical communication systems', *Proc. IEEE*, **58**, pp. 1446–1486, 1970.
[2] H. Melchior, 'Detectors for lightwave communications', *Phys. Today*, **30**, pp. 32–39, 1977.

[3] S. D. Personick, 'Photodetectors for fiber systems', in M. K. Barnoski (Ed.), *Fundamentals of Optical Fiber Communications* (2nd edn), pp. 257–293, Academic Press, 1981.
[4] T. P. Lee and T. Li, 'Photodetectors', in S. E. Miller and A. G. Chynoweth (Eds.), *Optical Fiber Telecommunications*, pp. 593–626, Academic Press, 1979.
[5] S. M. Sze, *Physics of Semiconductor Devices* (2nd edn), John Wiley, 1981.
[6] D. H. Newman and S. Ritchie, 'Sources and detectors for optical fibre communications applications: the first 20 years', *IEE Proc., Pt. J*, **133**(3), pp. 213–229, 1986.
[7] R. W. Dixon and N. K. Dutta, 'Lightwave device technology', *AT&T Tech. J.*, **66**(1), pp.73–83, 1987.
[8] J. E. Bowers and C. A. Burrus, 'Ultrawide-band long-wave *p–i–n* photodetectors', *J. of Lightwave Technol.*, **LT–5**(10), pp. 1339–1350, 1987.
[9] A. W. Nelson, S. Wong, S. Ritchie and S. K. Sargood, 'GaInAs PIN photodiodes grown by atmospheric-pressure MOVPE', *Electron. Lett.*, **21**(19), pp. 838–840, 1985.
[10] S. Miura, H. Kuwatsuka, T. Mikawa and O. Wada, 'Planar embedded InP/GaInAs *p–i–n* photodiode for very high speed operation', *J. of Lightwave Technol.*, **LT-5** (10), pp. 1371–1376, 1987.
[11] G. Keiser, *Optical Fiber Communications* (2nd edn), McGraw-Hill, 1991.
[12] S. R. Forrest, 'Optical detectors for lightwave communication', in S. E. Miller and I. P. Kaminow (Eds.), *Optical Fiber Telecommunications II*, Academic Press, pp. 569–599, 1988.
[13] M. Schwartz, *Information Transmission, Modulation and Noise* (4th edn), McGraw-Hill, 1990.
[14] T. P. Lee, C. A. Burrus Jr and A. G. Dentai, 'InGaAsP/InP photodiodes microplasma-limited avalanche multiplication at 1–1.3 μm wavelength', *IEEE J. Quantum Electron.*, **QE-15**, pp. 30–35, 1979.
[15] P. P. Webb, R. J. McIntyre and J. Conradi, 'Properties of avalanche photodiodes', *RCA Rev.*, **35**, pp. 235–277, 1974.
[16] A. R. Hartman, H. Melchior, D. P. Schinke and T. E. Seidel, 'Planar epitaxial silicon avalanche photodiode', *Bell Sys. Tech. J.*, **57**, pp. 1791–1807, 1978.
[17] T. Pearsall, 'Photodetectors for communication by optical fibres', in M. J. Howes and D. V Morgan (Eds.), *Optical Fibre Communications*, pp. 107–165, John Wiley, 1980.
[18] H. Ando, H. Kanbe, T. Kimura, T. Yamaoka and T. Kaneda, 'Characteristics of germanium avalanche photodiodes in the wavelength region 1–1.6 μm, *IEEE J. Quantum Electron.*, **QE-14**(11), pp. 804–809, 1978.
[19] T. Mikawa, S. Kagawa, T. Kaneda, T. Sakwai, H. Ando and O. Mikami, 'A low-noise n^+np germanium avalanche photodiode', *IEEE J. Quantum Electron.* **QE-17**(2), pp. 210–216, 1981.
[20] O. Mikami, H. Ando, H. Kanbe, T. Mikawa, T. Kaneda and Y. Toyama, 'Improved germanium avalanche photodiodes', *IEEE J. Quantum Electron.*, **QE-16**(9), pp. 1002–1007, 1980.
[21] S. Kagawa, T. Kaneda, T. Mikawa, Y. Banba, Y. Toyama and O. Mikami, 'Fully ion-implanted p^+-n germanium avalanche photodiodes', *Appl. Phys. Lett.*, **38**(6), pp. 429–431, 1981.
[22] T. Mikawa, T. Kaneda, H. Nishimoto, M. Motegi and H. Okushima, 'Small-active-area germanium avalanche photodiode for single-mode fibre at 1.3 μm wavelength', *Electron. Lett.*, **19**(12), pp. 452–453, 1983.
[23] M. Niwa, Y. Tashiro, K. Minemura and H. Iwasaki, 'High-sensitivity Hi–Lo

germanium avalanche photodiode for 1.5 μm wavelength optical communication *Electron. Lett.*, **20**(13), pp. 552–553, 1984.

[24] G. E. Stillman, 'Detectors for optical-waveguide communications', in E. E. Basch (Ed.), *Optical-Fiber Transmission*, H. W. Sams & Co., pp. 335–374, 1987.

[25] S. R. Forrest, O. K. Kim and R. G. Smith, 'Optical response time in $In_{0.53}Ga_{0.47}$ As/InP avalanche photodiodes', *Appl. Phys. Lett.*, **41**, pp. 95–98, 1982.

[26] J. C. Campbell, A. G. Dentai, W. S. Holder and B. L. Kasper, 'High performance avalanche photodiode with separate absorption "grading" and multiplication regions', *Electron. Lett.*, **19**, pp. 818–820, 1983.

[27] B. L. Kasper and J. C. Campbell, 'Multigigabit-per-second avalanche photodiode lightwave receivers', *J. of Lightwave Technol.*, **LT-5**(10), pp. 1351–1364, 1987.

[28] C. Fujihashi, 'Dark-current multiplication noises in avalanche photodiodes and optimum gains', *J. of Lightwave Technol.*, **LT-5**(6), pp. 798–808, 1987.

[29] R. Chin, N. Holonyak, G. E. Stillman, J. Y. Tang and K. Hess, 'Impact ionisation in multi-layer heterojuction structures', *Electron. Lett.*, **16**(12), pp. 467–468, 1980.

[30] F. Capasso, 'Band-gap engineering via graded gap, superlattice and periodic doping structures: applications and novel photodetectors and other devices', *J. Vac. Sci. Technol. B*, **1**(2), pp. 457–461, 1983.

[31] B. T. Debney and A. C. Carter, 'Optical detectors and receivers', in J. Dakin and B. Culshaw (Eds.) *Optical Fiber Sensors: Principles and components*, Artech House, pp. 107–149, 1988.

[32] J. E. Bowers, A. K. Srivastrava, C. A. Burrus, J. C. Dewinter, M. A. Pollack, J. L. Zyskind, 'High-speed GaInAsSb/GaSb PIN photodetectors for wavelengths to 2.3 μm', *Electron. Lett.*, **22**(3), pp. 137–138, 1986.

[33] A. J. Moseley, M. D. Scott, A. H. Moore and R. H. Wallis, 'High-efficiency, low-leakage MOCVD-grown GaInAs/AlInAs heterojunction photodiodes for detection to 2.4 μm, *Electron Lett.*, **22**(22), pp. 1206–1207, 1986.

[34] R. U. Martinelli, T. J. Zamerowski and P. A. Longeway, '$In_xG_{1-x}As/InAs_yP_{1-y}$ lasers and photodiodes for 2.55 μm optical fiber communications', *Opt. Fiber. Commun. Conf., OFC '88* (USA), paper TUC6, January 1988.

[35] O. Hildebrand, W. Kuebart, K. W. Benz and M. H. Pilkuhn, 'GaAlSb avalanche photodiodes: resonant impact ionisation with very high ratio of ionisation coefficients', *IEEE J. Quantum Electron.*, **17**(2), pp. 284–288, 1981.

[36] G. Pichard, J. Meslage, T. Nguyen Duy and F. Raymond, '1.3 μm CdHgTe avalanche photodiodes for fibre-optic applications', in *Proc. 9th ECOC 83* (Switzerland) pp. 479–482, 1983.

[37] H. Haupt and O. Hildebrand, 'Lasers and photodetectors in Europe', *IEEE J. on Selected Areas in Commun.*, **SAC-4**(4), pp. 444–456, 1986.

[38] P. D. Wright, R. J. Nelson and T. Cella, 'High gain InGaAsP–InP heterojunction phototransistors'. *Appl. Phys. Lett.*, **37**(2), pp. 192–194, 1980.

[39] R. A. Milano, P. D. Dapkus and G. E. Stillman, 'Heterojunction phototransistors for fiber-optic communications', *Proc. SPIE Int. Soc. Opt. Eng.*, **272**, pp. 43–50, 1981.

[40] K. Tubatabaie-Alavi and C. G. Fonstad, 'Recent advances in InGaAs/InP phototransistors', *Proc. SPIE Int. Soc. Opt. Eng.*, **272**, pp. 38–42, 1981.

[41] G. E. Stillman, L. W. Cook, G. E. Bulman, N. Tabatabaie, R. Chin and P. D. Dapkus, 'Long-wavelength (1.3 to 1.6 μm) detectors for fiber-optical communications', *IEEE Trans. Electron. Dev.*, **ED-29**(9), pp. 1355–1371, 1982.

[42] S. M. Sze, *Semiconductor Devices: Physics and technology*, John Wiley 1985.
[43] S. R. Forrest, 'The sensitivity of photoconductor receivers for long-wavelength optical communications', *J. of Lightwave Technol.*, **LT-3**(2), pp. 347–360, 1985.
[44] C. Y. Chen, A. G. Dentai, B. L. Kasper and P. A. Garbinski, 'High-speed junction-depleted $Ga_{0.47}In_{0.53}As$ photoconductive detectors', *Appl. Phys. Lett*, **46**, pp. 1164–1166, 1985.
[45] T. Morita, M. Murata, K. Koike and K. Ono, 'High speed GaInAs photoconductive-like detectors for long wavelength optical communication', *Proc. 14th ECOC '88* (UK) pp. 424–427, September 1988.

9

Direct detection receiver performance considerations

9.1 Introduction
9.2 Noise
9.3 Receiver noise
9.4 Receiver structures
9.5 FET preamplifiers
9.6 High performance receivers
Problems
References

9.1 Introduction

The receiver in an intensity modulated/direct detection (IM/DD) optical fiber communication system (see Section 7.5) essentially consists of the photodetector plus an amplifier with possibly additional signal processing circuits. Therefore the receiver initially converts the optical signal incident on the detector into an electrical signal, which is then amplified before further processing to extract the information originally carried by the optical signal.

The importance of the detector in the overall system performance was stressed in Chapter 8. However, it is necessary to consider the properties of this device in the context of the associated circuitry combined in the receiver. It is essential that the detector performs efficiently with the following amplifying and signal processing circuits. Inherent to this process is the separation of the information originally

contained in the optical signal from the noise generated within the rest of the system and in the receiver itself, as well as any limitations on the detector response imposed by the associated circuits. These factors play a crucial role in determining the performance of the system.

In order to consider receiver design it is useful to regard the limit on the performance of the system set by the signal to noise ratio (SNR) at the receiver. It is therefore necessary to outline noise sources within optical fiber systems. The noise in these systems has different origins from that of copper-based systems. Both types of system have thermal noise generated in the receiver. However, although optical fiber systems exhibit little crosstalk the noise generated within the detector must be considered, as well as the noise properties associated with the electromagnetic carrier.

In Section 9.2 we therefore briefly review the major noise mechanisms which are present in direct detection optical fiber communication receivers prior to more detailed discussion of the limitations imposed by photon (or quantum) noise in both digital and analog transmission. This is followed in Section 9.3 with a more specific discussion of the noise associated with the two major receiver types (i.e. employing *p–i–n* and avalanche photodiode detectors). Expressions for the SNRs of these two receiver types are also developed in this section. Section 9.4 considers the noise and bandwidth performance of common preamplifier structures utilized in the design of optical fiber receivers. In Section 9.5 we present a brief account of low noise field effect transistor (FET) preamplifiers which find wide use within optical fiber communication receivers. This discussion also includes consideration of *p–i–n* photodiode/FET (PIN–FET) hybrid receiver circuits which have been developed for optical fiber communications. Finally, major high performance receiver design strategies to provide low noise and high bandwidth operation as well as wide dynamic range are described in Section 9.6.

9.2 Noise

Noise is a term generally used to refer to any spurious or undesired disturbances that mask the received signal in a communication system. In optical fiber communication systems we are generally concerned with noise due to spontaneous fluctuations rather than erratic disturbances which may be a feature of copper-based systems (due to electromagnetic interference, etc.).

There are three main types of noise due to spontaneous fluctuations in optical fiber communication systems: thermal noise, dark current noise and quantum noise.

9.2.1 Thermal noise

This is the spontaneous fluctuation due to thermal interaction between, say, the free electrons and the vibrating ions in a conducting medium, and it is especially prevalent in resistors at room temperature.

The thermal noise current i_t in a resistor R may be expressed by its mean square value [Ref. 1] and is given by:

$$\overline{i_t^2} = \frac{4KTB}{R} \tag{9.1}$$

where K is Boltzmann's constant, T is the absolute temperature and B is the post-detection (electrical) bandwidth of the system (assuming the resistor is in the optical receiver).

9.2.2 Dark current noise

When there is no optical power incident on the photodetector a small reverse leakage current still flows from the device terminals. This dark current (see Section 8.4.2) contributes to the total system noise and gives random fluctuations about the average particle flow of the photocurrent. It therefore manifests itself as shot noise [Ref. 1] on the photocurrent. Thus the dark current noise $\overline{i_d^2}$ is given by:

$$\overline{i_d^2} = 2eBI_d \tag{9.2}$$

where e is the charge on an electron and I_d is the dark current. It may be reduced by careful design and fabrication of the detector.

9.2.3 Quantum noise

The quantum nature of light was discussed in Section 6.2.1 and the equation for the energy of this quantum or photon was stated as $E = hf$. The quantum behaviour of electromagnetic radiation must be taken into account at optical frequencies since $hf > KT$ and quantum fluctuations dominate over thermal fluctuations.

The detection of light by a photodiode is a discrete process since the creation of an electron–hole pair results from the absorption of a photon, and the signal emerging from the detector is dictated by the statistics of photon arrivals. Hence the statistics for monochromatic coherent radiation arriving at a detector follows a discrete probability distribution which is independent of the number of photons previously detected.

It is found that the probability $P(z)$ of detecting z photons in time period τ when it is expected on average to detect z_m photons obeys the Poisson distribution [Ref. 2]:

$$P(z) = \frac{z_m^z \exp(-z_m)}{z!} \tag{9.3}$$

where z_m is equal to the variance of the probability distribution. This equality of the mean and the variance is typical of the Poisson distribution. From Eq. (8.7) the electron rate r_e generated by incident photons is $r_e = \eta P_o/hf$. The number of electrons generated in time τ is equal to the average number of photons detected

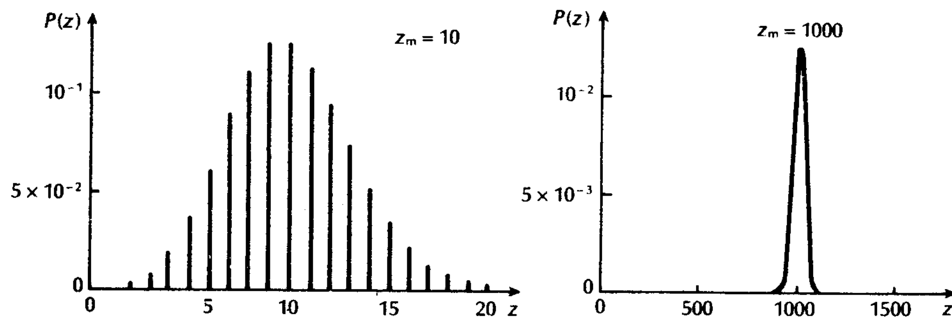


Figure 9.1 Poisson distributions for $z_m = 10$ and $z_m = 1000$.

over this time period z_m. Therefore:

$$z_m = \frac{\eta P_o \tau}{hf} \tag{9.4}$$

The Poisson distributions for $z_m = 10$ and $z_m = 1000$ are illustrated in Figure 9.1 and represent the detection process for monochromatic coherent light.

Incoherent light is emitted by independent atoms and therefore there is no phase relationship between the emitted photons. This property dictates exponential intensity distribution for incoherent light which if averaged over the Poisson distribution [Ref. 2] gives

$$P(z) = \frac{z_m^z}{(1 + z_m)^{z+1}} \tag{9.5}$$

Equation (9.5) is identical to the Bose–Einstein distribution [Ref. 3] which is used to describe the random statistics of light emitted in black body radiation (thermal light). The statistical fluctuations for incoherent light are illustrated by the probability distributions shown in Figure 9.2.

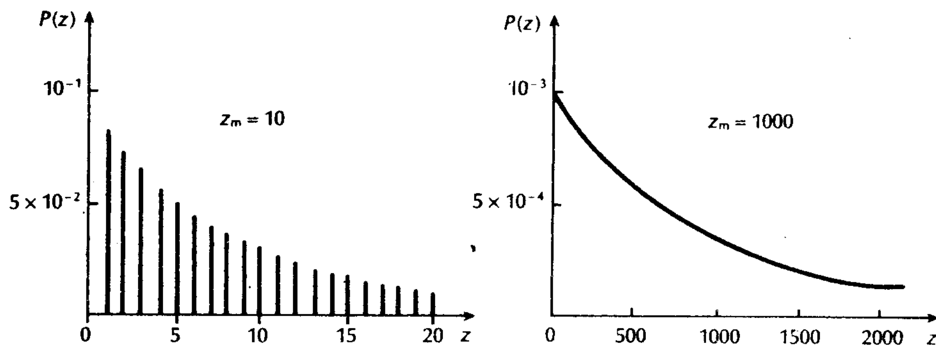


Figure 9.2 Probability distributions indicating the statistical fluctuations of incoherent light for $z_m = 10$ and $z_m = 1000$.

9.2.4 Digital signalling quantum noise

For digital optical fiber systems it is possible to calculate a fundamental lower limit to the energy that a pulse of light must contain in order to be detected with a given probability of error. The premise on which this analysis is based is that the ideal receiver has a sufficiently low amplifier noise to detect the displacement current of a single electron–hole pair generated within the detector (i.e. an individual photon may be detected). Thus in the absence of light, and neglecting dark current, no current will flow. Therefore the only way an error can occur is if a light pulse is present and no electron–hole pairs are generated. The probability of no pairs being generated when a light pulse is present may be obtained from Eq. (9.3) and is given by:

$$P(0/1) = \exp(-z_m) \tag{9.6}$$

Thus in the receiver described $P(0/1)$ represents the system error probability $P(e)$ and therefore:

$$P(e) = \exp(-z_m) \tag{9.7}$$

However, it must be noted that the above analysis assumes that the photodetector emits no electron–hole pairs in the absence of illumination. In this sense it is considered perfect. Equation (9.7) therefore represents an absolute receiver sensitivity and allows the determination of a fundamental limit in digital optical communications. This is the minimum pulse energy E_{min} required to maintain a given bit error rate (BER) which any practical receiver must satisfy and is known as the quantum limit.

Example 9.1

A digital optical fiber communication system operating at a wavelength of 1 μm requires a maximum bit error rate of 10^{-9}. Determine:

(a) the theoretical quantum limit at the receiver in terms of the quantum efficiency of the detector and the energy of an incident photon;

(b) the minimum incident optical power required at the detector in order to achieve the above bit error rate when the system is employing ideal binary signalling at 10 Mbits s^{-1}, and assuming the detector is ideal.

Solution: (a) From Eq. (9.7) the probability of error.

$$P(e) = \exp(-z_m) = 10^{-9}$$

and thus $z_m = 20.7$.

z_m corresponds to an average number of photons detected in a time period τ for a BER of 10^{-9}

From Eq. (9.4):

$$z_m = \frac{\eta P_o \tau}{hf} = 20.7$$

Hence the minimum pulse energy or quantum limit

$$E_{min} = P_o \tau = \frac{20.7\ hf}{\eta}$$

Thus the quantum limit at the receiver to maintain a maximum BER of 10^{-9} is

$$\frac{20.7\ hf}{\eta}$$

(b) From part (a) the minimum pulse energy:

$$P_o \tau = \frac{20.7\ hf}{\eta}$$

Therefore the average received optical power required to provide the minimum pulse energy is:

$$P_o = \frac{20.7\ hf}{\tau \eta}$$

However, for ideal binary signalling there are an equal number of ones and zeros (50% in the on state and 50% in the off state). Thus the average received optical power may be considered to arrive over two bit periods, and

$$P_o(\text{binary}) = \frac{20.7\ hf}{2\tau\eta} = \frac{20.7\ hf\ B_T}{2\eta}$$

where B_T is the bit rate. At a wavelength of 1 μm, $f = 2.998 \times 10^{14}$ Hz, and assuming an ideal detector, $\eta = 1$.

Hence

$$P_o\ (\text{binary}) = \frac{20.7 \times 6.626 \times 10^{-34} \times 2.998 \times 10^{14} \times 10^{7}}{2}$$

$$= 20.6\ \text{pW}$$

In decibels (dB)

$$P_o \text{ in dB} = 10 \log_{10} \frac{P_o}{P_r}$$

where P_r is a reference power level.

When the reference power level is one watt:

$$P_o = 10 \log_{10} P_o \qquad \text{where } P_o \text{ is expressed in watts}$$

$$= 10 \log_{10} 2.06 \times 10^{-11}$$

$$= 3.14 - 110$$

$$= -106.9\ \text{dBW}$$

When the reference power level is one milliwatt

$$P_o = 10 \log_{10} 2.06 \times 10^{-8}$$
$$= 3.14 - 80$$
$$= -76.9 \text{ dBm}$$

Therefore the minimum incident optical power required at the receiver to achieve an error rate of 10^{-9} with ideal binary signalling is 20.6 pW or −76.9 dBm.

The result of Example 9.1 is a theoretical limit and in practice receivers are generally found to be at least 10 dB less sensitive. Furthermore, although some 20.7 photons are required in order to detect a binary one with a bit error rate of 10^{-9}, it is clear that these photons can arrive at the receiver over two bit periods if an equal number of transmitted ones and zeros are assumed (i.e. there are no photons transmitted in the zero bit periods). Hence the 20.7 photon per pulse requirement can be considered as an average of around 10.4 photons per bit at the quantum limit.

9.2.5 Analog transmission quantum noise

In analog optical fiber systems quantum noise manifests itself as shot noise which also has Poisson statistics [Ref. 1]. The shot noise current i_s on the photocurrent I_p is given by:

$$\overline{i_s^2} = 2eBI_p \tag{9.8}$$

Neglecting other sources of noise the SNR at the receiver may be written as:

$$\frac{S}{N} = \frac{I_p^2}{\overline{i_s^2}} \tag{9.9}$$

Substituting for $\overline{i_s^2}$ from Eq. (9.8) gives:

$$\frac{S}{N} = \frac{I_p}{2eB} \tag{9.10}$$

The expression for the photocurrent I_p given in Eq. (8.8) allows the SNR to be obtained in terms of the incident optical power P_o.

$$\frac{S}{N} = \frac{\eta P_o e}{hf2eB} = \frac{\eta P_o}{2hfB} \tag{9.11}$$

Equation (9.11) allows calculation of the incident optical power required at the receiver in order to obtain a specified SNR when considering quantum noise in analog optical fiber systems.

Example 9.2

An analog optical fiber system operating at a wavelength of 1 μm has a post detection bandwidth of 5 MHz. Assuming an ideal detector and considering only quantum noise on the signal, calculate the incident optical power necessary to achieve an SNR of 50 dB at the receiver.

Solution: From Eq. (9.11), the SNR is

$$\frac{S}{N} = \frac{\eta P_o}{2hfB}$$

Hence

$$P_o = \left(\frac{S}{N}\right) \frac{2hfB}{\eta}$$

For $S/N = 50$ dB, when considering signal and noise powers:

$$10 \log_{10} \frac{S}{N} = 50$$

and therefore $S/N = 10^5$

At 1 μm, $f = 2.998 \times 10^{14}$ Hz. For an ideal detector $\eta = 1$ and, thus the incident optical power:

$$P_o = \frac{10^5 \times 2 \times 6.626 \times 10^{-34} \times 2.998 \times 10^{14} \times 5 \times 10^6}{1}$$

$$= 198.6 \text{ nW}$$

In dBm

$$P_o = 10 \log_{10} 198.6 \times 10^{-6}$$

$$= -40 + 2.98$$

$$= -37.0 \text{ dBm}$$

Therefore the incident optical power required to achieve an SNR of 50 dB at the receiver is 198.6 nW which is equivalent to -37.0 dBm.

In practice receivers are less sensitive than the Example 9.2 suggests and thus in terms of the absolute optical power requirements analog transmission compares unfavourably with digital signalling.

However, it should be noted that there is a substantial difference in information transmission capacity between the digital and analog cases (over similar bandwidths) considered in Examples 9.1 and 9.2. For example, a 10 Mbit s^{-1} digital optical fiber communication system would provide only about 150 speech channels using standard baseband digital transmission techniques (see Section 11.5). In contrast a 5 MHz analog system, again operating in the baseband, could provide

as many as 1250 similar bandwidth ($\simeq 3.4$ kHz) speech channels. A comparison of signal to quantum noise ratios between the two transmission methods, taking account of this information capacity aspect, yields less disparity although digital signalling still proves far superior. For instance, applying the figures quoted above within Examples 9.1 and 9.2, in order to compare two systems capable of transmitting the same number of speech channels (e.g. digital bandwidth of $10\ \text{Mbit s}^{-1}$ and analog bandwidth of 600 kHz) gives a difference in absolute sensitivity in favour of digital transmission of approximately 31 dB. This indicates a reduction of around 9 dB on the 40 dB difference obtained by simply comparing the results over similar bandwidths. Nevertheless, it is clear that digital signalling techniques still provide a significant benefit in relation to quantum noise when employed within optical fiber communications.

9.3 Receiver noise

In order to investigate the optical receiver in greater detail it is necessary to consider the relative importance and interplay of the various types of noise mentioned in the preceding section. This is dependent on both the method of demodulation and the type of device used for detection.

The conditions for coherent detection are not met in IM/DD optical fiber systems for the reasons outlined in Section 7.5. Thus heterodyne and homodyne detection, which are very sensitive techniques and provide excellent rejection of adjacent channels, are not used, as the optical signal arriving at the receiver tends to be incoherent. In practice the vast majority of installed optical fiber communication systems use incoherent or direct detection in which the variation of the optical power level is monitored and no information is carried in the phase or frequency content of the signal. Therefore, the noise considerations in this section are based on a receiver employing direct detection of the modulated optical carrier which gives the same signal to noise ratio as an unmodulated optical carrier. The substantial developments in coherent optical fiber transmission, however, which have taken place over recent years are described in Chapter 12. Nevertheless, the major performance parameters associated with direct detection receivers which are discussed in this section and the following ones also apply to coherent optical receivers.

Figure 9.3 shows a block schematic of the front end of an optical receiver and the various noise sources associated with it. The majority of the noise sources shown apply to both main types of optical detector (*p–i–n* and avalanche photodiode). The noise generated from background radiation, which is important in atmospheric propagation and some copper-based systems, is negligible in both types of optical fiber receiver, and thus is often ignored. Also the beat noise generated from the various spectral components of the incoherent optical carrier can be shown to be insignificant [Ref. 4] with multimode propagation and hence

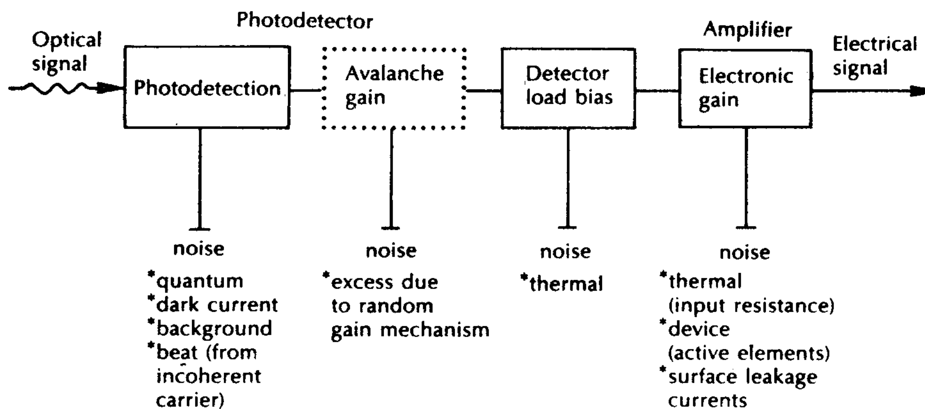


Figure 9.3 Block schematic of the front end of an optical receiver showing the various sources of noise.

will not be considered. It is necessary, however, to take into account the other sources of noise shown in Figure 9.3.

The avalanche photodiode receiver is the most complex case as it includes noise resulting from the random nature of the internal gain mechanism (dotted in Figure 9.3). It is therefore useful to consider noise in optical fiber receivers employing photodiodes without internal gain, before avalanche photodiode receivers are discussed.

9.3.1 *p–n* and *p–i–n* photodiode receiver

The two main sources of noise in photodiodes without internal gain are dark current noise and quantum noise, both of which may be regarded as shot noise on the photocurrent (i.e. effectively, analog quantum noise). When the expressions for these noise sources given in Eqs. (9.2) and (9.4) are combined the total shot noise $\overline{i_{TS}^2}$ is given by:

$$\overline{i_{TS}^2} = 2eB(I_p + I_d) \tag{9.12}$$

If it is necessary to take the noise due to the background radiation into account then the expression given in Eq. (9.12) may be expanded to include the background radiation induced photocurrent I_b giving

$$\overline{i_{TS}^2} = 2eB(I_p + I_d + I_b) \tag{9.13}$$

However, as I_b is usually negligible the expression given in Eq. (9.12) will be used in the further analysis.

When the photodiode is without internal avalanche gain, thermal noise from the detector load resistor and from active elements in the amplifier tends to dominate. This is especially the case for wideband systems operating in the 0.8 to 0.9 μm

wavelength band because the dark currents in well-designed silicon photodiodes can be made very small. The thermal noise $\overline{i_t^2}$ due to the load resistance R_L may be obtained from Eq. (9.1) and is given by:

$$\overline{i_t^2} = \frac{4KTB}{R_L} \tag{9.14}$$

The dominating effect of this thermal noise over the shot noise in photodiodes without internal gain may be observed in Example 9.3.

Example 9.3

A silicon *p–i–n* photodiode incorporated into an optical receiver has a quantum efficiency of 60% when operating at a wavelength of 0.9 μm. The dark current in the device at this operating point is 3 nA and the load resistance is 4 kΩ.

The incident optical power at this wavelength is 200 nW and the post detection bandwidth of the receiver is 5 MHz. Compare the shot noise generated in the photodiode with the thermal noise in the load resistor at a temperature of 20 °C.

Solution: From Eq. (8.8) the photocurrent is given by:

$$I_p = \frac{\eta P_o e}{hf} = \frac{\eta P_o e \lambda}{hc}$$

Therefore

$$I_p = \frac{0.6 \times 200 \times 10^{-9} \times 1.602 \times 10^{-19} \times 0.9 \times 10^{-6}}{6.626 \times 10^{-34} \times 2.998 \times 10^{8}}$$
$$= 87.1 \text{ nA}$$

From Eq. (9.12) the total shot noise is:

$$\overline{i_{TS}^2} = 2eB(I_d + I_p)$$
$$= 2 \times 1.602 \times 10^{-19} \times 5 \times 10^{6}[(3 + 87.1) \times 10^{-9})]$$
$$= 1.44 \times 10^{-19} \text{ A}^2$$

and the root mean square (rms) shot noise current is

$$(\overline{i_{TS}^2})^{\frac{1}{2}} = 3.79 \times 10^{-10} \text{ A}$$

The thermal noise in the load resistor is given by Eq. (9.14):

$$\overline{i_t^2} = \frac{4KTB}{R_L}$$
$$= \frac{4 \times 1.381 \times 10^{-23} \times 293 \times 5 \times 10^{6}}{4 \times 10^{3}}$$
$$= 2.02 \times 10^{-17} \text{ A}^2$$

($T = 20\,^\circ\text{C} = 293$ K)

Therefore the rms thermal noise current is

$$(\overline{i_t^2})^{\frac{1}{2}} = 4.49 \times 10^{-9}\ \text{A}$$

In this example the rms thermal noise current is a factor of 12 greater than the total rms shot noise current.

Example 9.3 does not include the noise sources within the amplifier, shown in Figure 9.3. These noise sources, associated with both the active and passive elements of the amplifier, can be represented by a series voltage noise source $\overline{v_a^2}$ and a shunt current noise source $\overline{i_a^2}$.

Thus the total noise associated with the amplifier $\overline{i_{amp}^2}$ is given by:

$$\overline{i_{amp}^2} = \int_0^B (\overline{i_a^2} + \overline{v_a^2}\,|Y|^2)\,\mathrm{d}f \tag{9.15}$$

where Y is the shunt admittance (combines the shunt capacitances and resistances) and f is the frequency. An equivalent circuit for the front end of the receiver, including the effective input capacitance C_a and resistance R_a of the amplifier is shown in Figure 9.4. The capacitance of the detector C_d is also shown and the noise resulting from C_d is usually included in the expression for $\overline{i_{amp}^2}$ given in Eq. (9.15).

The SNR for the p–n or p–i–n photodiode receiver may be obtained by summing the noise contributions from Eqs. (9.12), (9.14) and (9.15). It is given by:

$$\frac{S}{N} = \frac{I_p^2}{2eB(I_p + I_d) + \dfrac{4KTB}{R_L} + \overline{i_{amp}^2}} \tag{9.16}$$

The thermal noise contribution may be reduced by increasing the value of the load resistor R_L, although this reduction may be limited by bandwidth considerations which are discussed later. Also, the noise associated with the amplifier $\overline{i_{amp}^2}$ may be reduced with low detector and amplifier capacitance.

However, when the noise associated with the amplifier $\overline{i_{amp}^2}$ is referred to the load resistor R_L, the noise figure F_n [Ref. 1] for the amplifier may be obtained. This

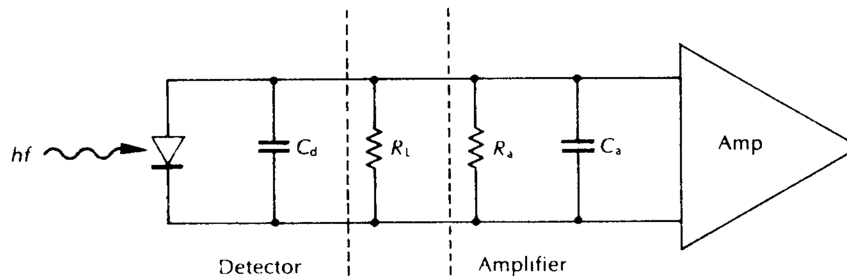


Figure 9.4 The equivalent circuit for the front end of an optical fiber receiver.

allows $\overline{i_{amp}^2}$ to be combined with the thermal noise from the load resistor $\overline{i_t^2}$ to give:

$$\overline{i_t^2} + \overline{i_{amp}^2} = \frac{4KTBF_n}{R_L} \quad (9.17)$$

The expression for the SNR given in Eq. (9.16) can now be written in the form:

$$\frac{S}{N} = \frac{I_p^2}{2eB(I_p + I_d) + \dfrac{4KTBF_n}{R_L}} \quad (9.18)$$

Thus if the noise figure F_n for the amplifier is known, Eq. (9.18) allows the SNR to be determined.

Example 9.4

The receiver in Example 9.3 has an amplifier with a noise figure of 3 dB. Determine the SNR at the output of the receiver under the same conditions as Example 9.3.

Solution: From Example 9.3:

$$I_p = 87.1 \times 10^{-9} \text{ A}$$
$$\overline{i_{TS}^2} = 1.44 \times 10^{-19} \text{ A}^2$$
$$\overline{i_t^2} = 2.02 \times 10^{-17} \text{ A}^2$$

The amplifier noise figure

$$F_n = 3 \text{ dB} = 10 \log_{10} 2$$

Thus F_n may be considered as $\times 2$.

In Eq. (9.18) the SNR is given by:

$$\frac{S}{N} = \frac{I_p^2}{2eB(I_p + I_d) + \dfrac{4KTBF_n}{R_L}}$$
$$= \frac{I_p^2}{\overline{i_{TS}^2} + (\overline{i_t^2} \times F_n)}$$
$$= \frac{(87.1 \times 10^{-9})^2}{(1.44 \times 10^{-19}) + (2.02 \times 10^{-17} \times 2)}$$
$$= 1.87 \times 10^2$$

SNR in dB $= 10 \log_{10} 1.87 \times 10^2 = 22.72$ dB.

Alternatively it is possible to conduct the calculation in dB if we neglect the shot noise (say, $\overline{i_{TS}^2} = 0$).

In dB:

$$I_p = 9.40 - 80 = -70.60$$

Hence

$$I_p^2 = -141.20 \text{ dB}$$

and

$$\overline{i_t^2} = 3.05 - 170 = -166.95 \text{ dB}.$$

The amplifier noise figure $F_n = 3$ dB.
Therefore the

$$SNR = -141.20 + 166.95 - 3$$
$$= 22.75 \text{ dB}$$

A slight difference in the final answer may be noted. This is due to the neglected shot noise term.

A quantity discussed in Section 8.8.3 which is often used in the specification of optical detectors (or detector–amplifier combinations) is the noise equivalent power (NEP). It is defined as the amount of incident optical power P_o per unit bandwidth required to produce an output power equal to the detector (or detector–amplifier combination) output noise power. The NEP is therefore the value of P_o which gives an output SNR of unity. Thus the lower the NEP for a particular detector (or detector–amplifier combination), the less optical power is needed to obtain a particular SNR.

9.3.2 Receiver capacitance and bandwidth

Considering the equivalent circuit shown in Figure 9.4, the total capacitance for the front end of an optical receiver C_T is given by:

$$C_T = C_d + C_a \tag{9.19}$$

where C_d is the detector capacitance and C_a is the amplifier input capacitance. It is important that this total capacitance is minimized not only from the noise considerations discussed previously but also from the bandwidth penalty which is incurred due to the time constant of C_T and the load resistance R_L. We assume here that R_L is the total loading on the detector and therefore have neglected the amplifier input resistance R_a. However, in practical receiver configurations R_a may have to be taken into account (see Section 9.4.1). The reciprocal of the time constant $2\pi R_L C_T$ must be greater than, or equal to, the post detection bandwidth B;

$$\frac{1}{2\pi R_L C_T} \geqslant B \tag{9.20}$$

When the equality exists in Eq. (9.20) it defines the maximum possible value of B for the straightforward termination indicated in Figure 9.4.

Assuming that the total capacitance may be minimized, then the other parameter which affects B is the load resistance R_L. To increase B it is necessary to reduce R_L. However, this introduces a thermal noise penalty as may be seen from Eq. (9.14) where both the increase in B and decrease in R_L contribute to an increase in the thermal noise. A trade-off therefore exists between the maximum bandwidth and the level of thermal noise which may be tolerated. This is especially important in receivers which are dominated by thermal noise.

Example 9.5

A photodiode has a capacitance of 6 pF. Calculate the maximum load resistance which allows an 8 MHz post detection bandwidth.

Determine the bandwidth penalty with the same load resistance when the following amplifier also has an input capacitance of 6 pF.

Solution: From Eq. (9.20) the maximum bandwidth is given by:

$$B = \frac{1}{2\pi R_L C_d}$$

Therefore the maximum load resistance

$$R_L(\text{max}) = \frac{1}{2\pi C_d B} = \frac{1}{2\pi \times 6 \times 10^{-12} \times 8 \times 10^6}$$

$$= 3.32 \text{ k}\Omega$$

Thus for an 8 MHz bandwidth the maximum load resistance is 3.32 kΩ.

Also, considering the amplifier capacitance, the maximum bandwidth

$$B = \frac{1}{2\pi R_L (C_d + C_a)} = \frac{1}{2\pi \times 3.32 \times 10^3 \times 12 \times 10^{-12}}$$

$$= 4 \text{ MHz}$$

As would be expected the maximum post detection bandwidth is halved.

9.3.3 Avalanche photodiode (APD) receiver

The internal gain mechanism in an APD increases the signal current into the amplifier and so improves the SNR because the load resistance and amplifier noise remain unaffected (i.e. the thermal noise and amplifier noise figure are unchanged). However, the dark current and quantum noise are increased by the multiplication process and may become a limiting factor. This is because the random gain mechanism introduces excess noise into the receiver in terms of increased shot noise

above the level that would result from amplifying only the primary shot noise. Thus if the photocurrent is increased by a factor M (mean avalanche multiplication factor), then the shot noise is also increased by an excess noise factor M^x, such that the total shot noise $\overline{i_{SA}^2}$ is now given by:

$$\overline{i_{SA}^2} = 2eB(I_p + I_d)M^{2+x} \tag{9.21}$$

where x is between 0.3 and 0.5 for silicon APDs and between 0.7 and 1.0 for germanium or III–V alloy APDs.

Equation (9.21) is often used as the total shot noise term in order to compute the SNR, although there is a small amount of shot noise current which is not multiplied through impact ionization. The shot noise current in the detector which is not multiplied is a device parameter and may be considered as an extra shot noise term. However, it tends to be insignificant in comparison with the multiplied shot noise and is therefore neglected in the further analysis (i.e. all shot noise is assumed to be multiplied).

The SNR for the avalanche photodiode may be obtained by summing the combined noise contribution from the load resistor and the amplifier given in Eq. (9.17), which remains unchanged, with the modified noise term given in Eq. (9.21). Hence the SNR for the APD is:

$$\frac{S}{N} = \frac{M^2 I_p^2}{2eB(I_p + I_d)M^{2+x} + \dfrac{4KTBF_n}{R_L}} \tag{9.22}$$

It is apparent from Eq. (9.22) that the relative significance of the combined thermal and amplifier noise term is reduced due to the avalanche multiplication of the shot noise term. When Eq. (9.22) is written in the form:

$$\frac{S}{N} = \frac{I_p^2}{2eB(I_p + I_d)M^x + \dfrac{4KTBF_n}{R_L} M^{-2}} \tag{9.23}$$

it may be seen that the first term in the denominator increases with increasing M whereas the second term decreases. For low M the combined thermal and amplifier noise term dominates and the total noise power is virtually unaffected when the signal level is increased, giving an improved SNR. However, when M is large, the thermal and amplifier noise term becomes insignificant and the SNR decreases with increasing M at the rate of M^x. An optimum value of the multiplication factor M_{op} therefore exists which maximizes the SNR. It is given by:

$$\frac{2eB(I_p + I_d)M_{op}^x}{(4KTBF_n/R_L)M_{op}^{-2}} = \frac{2}{x} \tag{9.24}$$

and therefore

$$M_{op}^{2+x} = \frac{4KTF_n}{xeR_L(I_p + I_d)} \tag{9.25}$$

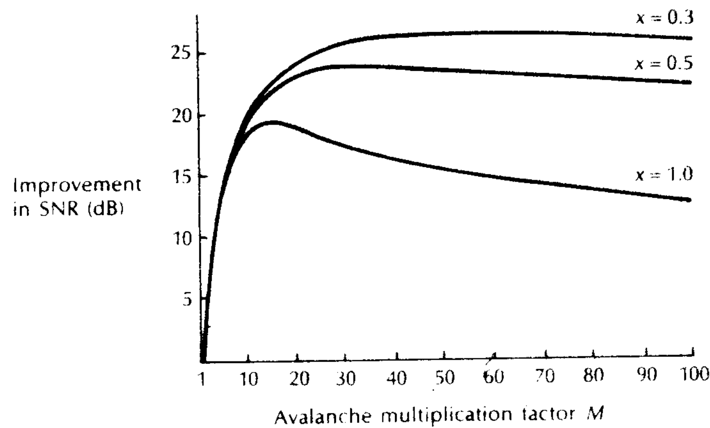


Figure 9.5 The improvement in SNR as a function of avalanche multiplication factor M for different excess noise factors M^x. Reproduced with permission from I. Garrett, *The Radio and Electron. Eng.*, **51**, p. 349, 1981.

The variation in M_{op} for both silicon and germanium APDs is illustrated in Figure 9.5 [Ref. 5]. This shows a plot of Eq. (9.22) with F_n equal to unity and neglecting the dark current. For good silicon APDs where x is 0.3, the optimum multiplication factor covers a wide range. In the case illustrated in Figure 9.5 M_{op} commences at about 40 where the possible improvement in SNR above a photodiode without internal gain is in excess of 25 dB. However, for germanium and III–V alloy APDs where x may be equal to unity it can be seen that less SNR improvement is possible (less than 19 dB). Moreover, the maximum is far sharper, occurring at a multiplication factor of about 12. Also it must be noted that Figure 9.5 demonstrates the variation of M_{op} with x for a specific case, and therefore only represents a general trend. It may be observed from Eq. (9.25) that M_{op} is dependent on a number of other variables apart from x.

Example 9.6

A good silicon APD ($x = 0.3$) has a capacitance of 5 pF, negligible dark current and is operating with a post detection bandwidth of 50 MHz. When the photocurrent before gain is 10^{-7} A and the temperature is 18 °C; determine the maximum SNR improvement between $M = 1$ and $M = M_{op}$ assuming all operating conditions are maintained.

Solution: Determine the maximum value of the load resistor from Eq. (9.20):

$$R_L = \frac{1}{2\pi C_d B} = \frac{1}{2\pi \times 5 \times 10^{-12} \times 50 \times 10^6}$$

$$= 635.5\ \Omega$$

When $M = 1$, the SNR is given by Eq. (9.22),

$$\frac{S}{N} = \frac{I_p^2}{2eBI_p + \dfrac{4KTB}{R_L}}$$

where $I_d = 0$ and $F_n = 1$

The shot noise is:

$$2eBI_p = 2 \times 1.602 \times 10^{-19} \times 50 \times 10^6 \times 10^{-7}$$
$$= 1.602 \times 10^{-18} \text{ A}^2$$

and the thermal noise is:

$$\frac{4KTB}{R_L} = \frac{4 \times 1.381 \times 10^{-23} \times 291 \times 50 \times 10^6}{636.5}$$
$$= 1.263 \times 10^{-15} \text{ A}^2$$

It may be noted that the thermal noise is dominating.

Therefore

$$\frac{S}{N} = \frac{10^{-14}}{1.602 \times 10^{-18} \times 1.263 \times 10^{-15}} = 7.91$$

and the SNR in dBs is:

$$\frac{S}{N} = 10 \log_{10} 7.91 = 8.98 \text{ dB}$$

Thus the SNR when $M = 1$ is 9.0 dB.

When $M = M_{op}$ and $x = 0.3$, from Eq. (9.25):

$$M_{op}^{2+x} = \frac{4KT}{xeR_L I_p}$$

where $I_d = 0$ and $F_n = 1$. Hence:

$$M_{op}^{2.3} = \frac{4 \times 1.381 \times 10^{-23} \times 291}{0.3 \times 1.602 \times 10^{-19} \times 636.5 \times 10^{-7}}$$

and

$$M_{op} = (5.255 \times 10^3)^{0.435}$$
$$= 41.54$$

The SNR at M_{op} may be obtained from Eq. (9.22):

$$\frac{S}{N} = \frac{M^2 I_p^2}{2eBI_p M^{2.3} + \dfrac{4KTB}{R_L}}$$

$$= \frac{(41.54)^2 \times 10^{-14}}{\{1.602 \times 10^{-18} \times (41.54)^{2.3}\} + 1.263 \times 10^{-15}}$$

$$= 1.78 \times 10^3$$

and the SNR in dBs is

$$\frac{S}{N} = 10 \log_{10} 1.78 \times 10^3 = 32.50\,\text{dB}$$

Therefore the SNR when $M = M_{op}$ is 32.5 dB and the SNR improvement over $M = 1$ is 23.5 dB.

Example 9.7

A germanium APD (with $x = 1$) is incorporated into an optical fiber receiver with a 10 kΩ load resistance. When operated at a temperature of 120 K, the minimum photocurrent required to give a SNR of 35 dB at the output of the receiver is found to be a factor of 10 greater than the dark current. If the noise figure of the following amplifier at this temperature is 1 dB and the post detection bandwidth is 10 MHz, determine the optimum avalanche multiplication factor.

Solution: From Eq. (9.22) with $x = 1$ and $M = M_{op}$ (i.e. minimum photocurrent specifies that $M = M_{op}$) the SNR is:

$$\frac{S}{N} = \frac{M_{op}^2 I_p^2}{2eB(I_p + I_d)M_{op}^3 + \dfrac{4KTB}{R_L}}$$

Also from Eq. (9.25)

$$M_{op}^3 = \frac{4KTF_n}{eR_L(I_p + I_d)}$$

Therefore

$$M_{op} = \left\{\frac{4KTF_n}{eR_L(I_p + I_d)}\right\}^{\frac{1}{3}}$$

Substituting into Eq. (9.22), this gives:

$$\frac{S}{N} = \frac{\left\{\dfrac{4KTF_n}{eR_L(I_p + I_d)}\right\}^{\frac{2}{3}} I_p^2}{\dfrac{8KTBF_n}{R_L} + \dfrac{4KTBF_n}{R_L}}$$

and as $I_d = 0.1\ I_p$ the SNR is:

$$\frac{S}{N} = \frac{\left(\dfrac{4KTF_n}{1.1eR_L}\right)^{\frac{2}{3}} I_p^{\frac{4}{3}}}{\dfrac{12KTBF_n}{R_L}}$$

Therefore the minimum photocurrent I_p:

$$I_p^{\frac{4}{3}} = \left(\frac{S}{N}\right) \frac{\dfrac{12KTBF_n}{R_L}}{\left(\dfrac{4KTF_n}{1.1\ eR_L}\right)^{\frac{2}{3}}}$$

where the SNR is:

$$\frac{S}{N} = 35 \text{ dB} = 3.16 \times 10^3$$

and as $F_n = 1$ dB which is equivalent to 1.26:

$$\frac{12KTBF_n}{R_L} = \frac{12 \times 1.381 \times 10^{-23} \times 120 \times 10^7 \times 1.26}{10^4}$$

$$= 2.51 \times 10^{-17}$$

Also

$$\left(\frac{4KTF_n}{1.1\ eR_L}\right)^{\frac{2}{3}} = \left(\frac{4 \times 1.381 \times 10^{-23} \times 120 \times 1.26}{1.1 \times 1.602 \times 10^{-19} \times 10^4}\right)^{\frac{2}{3}}$$

$$= 2.82 \times 10^{-4}$$

Therefore

$$I_p = \left(\frac{3.16 \times 10^3 \times 2.51 \times 10^{-17}}{2.82 \times 10^{-4}}\right)^{\frac{3}{4}}$$

$$= 6.87 \times 10^{-8} \text{ A}$$

To obtain the optimum avalanche multiplication factor we substitute back into Eq. (9.25), where:

$$M_{op} = \left(\frac{4 \times 1.381 \times 10^{-23} \times 120 \times 1.26}{1.602 \times 10^{-19} \times 10^3 \times 1.1 \times 6.87 \times 10^{-8}}\right)^{\frac{1}{3}}$$

$$= 8.84$$

In Example 9.7 the optimum multiplication factor for the germanium APD is found to be approximately 9. It shows the dependence of the optimum multiplication factor on the variables in Eq. (9.25), and although the example does not necessarily represent a practical receiver (some practical germanium APD receivers are cooled to reduce dark current), the optimum multiplication factor is influenced by device and system parameters as well as operating conditions.

9.3.4 Excess avalanche noise factor

The value of the excess avalanche noise factor is dependent upon the detector material, the shape of the electric field profile within the device and whether the avalanche is initiated by holes or electrons. It is often represented as $F(M)$ and in the preceding section we have considered one of the approximations for the excess

noise factor, where:

$$F(M) = M^x \tag{9.26}$$

and the resulting noise is assumed to be white with a Gaussian distribution.

However, a second and more exact relationship is given by [Ref. 6]:

$$F(M) = M\left[1 - (1 - k)\left(\frac{M-1}{M}\right)^2\right] \tag{9.27}$$

where the only carriers are injected electrons and k is the ratio of the ionization coefficients of holes and electrons. If the only carriers are injected holes:

$$F(M) = M\left[1 + \left(\frac{1-k}{k}\right)\left(\frac{M-1}{M}\right)^2\right] \tag{9.28}$$

The best performance is achieved when k is small, and for silicon APDs k is between 0.02 and 0.10, whereas for germanium and III–V alloy APDs k is between 0.3 and 1.0.

With electron injection in silicon photodiodes, the smaller values of k obtained correspond to a larger ionization rate for the electrons than for the holes. As k departs from unity, only the carrier with the larger ionization rate contributes to the impact ionization and the excess avalanche noise factor is reduced. When the impact ionization is initiated by electrons this corresponds to fewer ionizing collisions involving the hole current which is flowing in the opposite direction (i.e. less feedback). In this case the amplified signal contains less excess noise. The carrier ionization rates in germanium photodiodes are often nearly equal and hence k approaches unity, giving a high level of excess noise.

9.4 Receiver structures

A full equivalent circuit for the digital optical fiber receiver, in which the optical detector is represented as a current source i_{det}, is shown in Figure 9.6. The noise sources (i_t, i_{TS} and i_{amp}) and the immediately following amplifier and equalizer are also shown. Equalization [Ref. 7] compensates for distortion of the signal due to the combined transmitter, medium and receiver characteristics. The equalizer is

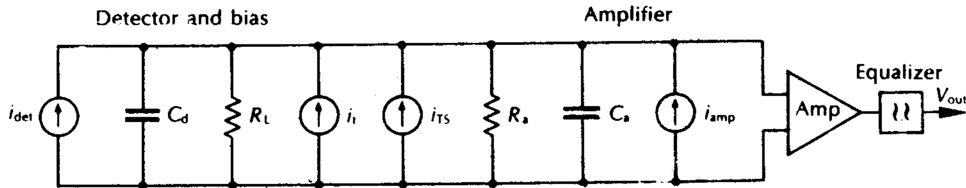


Figure 9.6 A full equivalent circuit for a digital optical fiber receiver including the various noise sources.

often a frequency shaping filter which has a frequency response that is the inverse of the overall system frequency response. In wideband systems this will normally boost the high frequency components to correct the overall amplitude of the frequency response. To acquire the desired spectral shape for digital systems (e.g. raised cosine, see Figure 11.37), in order to minimize intersymbol interference, it is important that the phase frequency response of the system is linear. Thus the equalizer may also apply selective phase shifts to particular frequency components.

However, the receiver structure immediately preceding the equalizer is the major concern of this section. In both digital and analog systems it is important to minimize the noise contributions from the sources shown in Figure 9.6 so as to maximize the receiver sensitivity whilst maintaining a suitable bandwidth. It is therefore useful to discuss various possible receiver structures with regard to these factors.

9.4.1 Low impedance front end

Three basic amplifier configurations are frequently used in optical fiber communication receivers. The simplest, and perhaps the most common, is the voltage amplifier with an effective input resistance R_a as shown in Figure 9.7.

In order to make suitable design choices, it is necessary to consider both bandwidth and noise. The bandwidth considerations in Section 9.3.2 are treated solely with regard to a detector load resistance R_L. However, in most practical receivers the detector is loaded with a bias resistor R_b and an amplifier (see Figure 9.7). The bandwidth is determined by the passive impedance which appears across the detector terminals which is taken as R_L in the bandwidth relationship given in Eq. (9.20).

However, R_L may be modified to incorporate the parallel resistance of the detector bias resistor R_b and the amplifier input resistance R_a. The modified total load resistance R_{TL} is therefore given by:

$$R_{TL} = \frac{R_b R_a}{R_b + R_a} \tag{9.29}$$

Considering the expressions given in Eqs. (9.20) and (9.29), to achieve an optimum bandwidth both R_b and R_a must be minimized. This leads to a low

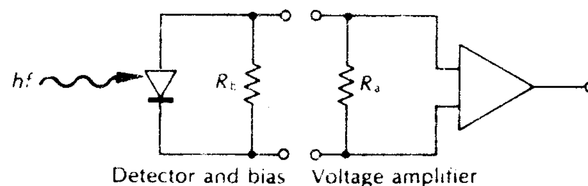


Figure 9.7 Low impedance front end optical fiber receiver with voltage amplifier.

impedance front end design for the receiver amplifier. Unfortunately this design allows thermal noise to dominate within the receiver (following Eq. (9.14)), which may severely limit its sensitivity. Therefore this structure demands a trade-off between bandwidth and sensitivity which tends to make it impractical for long-haul, wideband optical fiber communication systems.

9.4.2 High impedance (integrating) front end

The second configuration consists of a high input impedance amplifier together with a large detector bias resistor in order to reduce the effect of thermal noise. However, this structure tends to give a degraded frequency response as the bandwidth relationship given in Eq. (9.20) is not maintained for wideband operation. The detector output is effectively integrated over a large time constant and must be restored by differentiation. This may be performed by the correct equalization at a later stage [Ref. 8] as illustrated in Figure 9.8. Therefore the high impedance (integrating) front end structure gives a significant improvement in sensitivity over the low impedance front end design, but it creates a heavy demand for equalization and has problems of limited dynamic range (the ratio of maximum to minimum input signals).

The limitations on dynamic range result from the attenuation of the low frequency signal components by the equalization process which causes the amplifier to saturate at high signal levels. When the amplifier saturates before equalization has occurred the signal is heavily distorted. Thus the reduction in dynamic range is dependent upon the amount of integration and subsequent equalization employed.

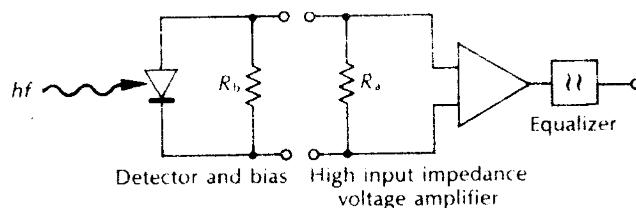


Figure 9.8 High impedance integrating front end optical fiber receiver with equalized voltage amplifier

9.4.3 The transimpedance front end

This configuration largely overcomes the drawbacks of the high impedance front end by utilizing a low noise, high input impedance amplifier with negative feedback. The device therefore operates as a current mode amplifier where the high input impedance is reduced by negative feedback. An equivalent circuit for an optical fiber receiver incorporating a transimpedance front end structure is shown in Figure 9.9. In this equivalent circuit the parallel resistances and capacitances are combined

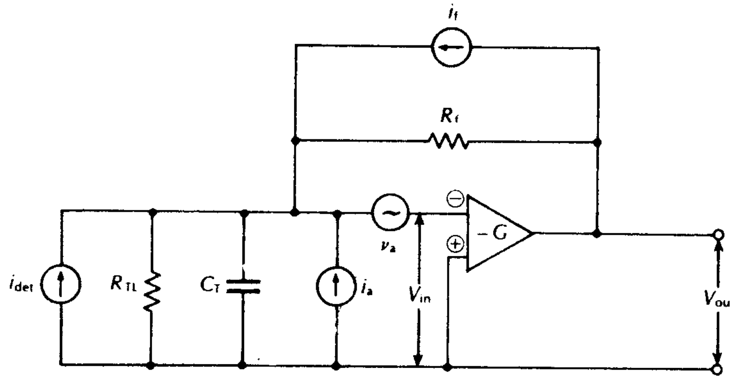


Figure 9.9 An equivalent circuit for the optical fiber receiver incorporating a transimpedance (current mode) preamplifier.

into R_{TL} and C_T respectively. The open loop current to voltage transfer function $H_{OL}(\omega)$ for this transimpedance configuration corresponds to the transfer function for the two structures described previously which do not employ feedback (i.e. the low and high impedance front ends). It may be written as:

$$H_{OL}(\omega) = -G\frac{V_{in}}{i_{det}} = -G\frac{R_{TL}\dfrac{1}{j\omega C_T}}{R_{TL}+\dfrac{1}{j\omega C_T}} = \frac{-GR_{TL}}{1+j\omega R_{TL}C_T}\ (\text{V A}^{-1}) \tag{9.30}$$

where G is the open loop voltage gain of the amplifier and ω is the angular frequency of the input signal.

In this case the bandwidth (without equalization) is constrained by the time constant given in Eq. (9.20).*

When the feedback is applied, the closed loop current to voltage transfer function $H_{CL}(\omega)$ for the transimpedance configuration is given by (see Appendix E)

$$H_{CL}(\omega) \simeq \frac{-R_f}{1+(j\omega R_f C_T/G)}\ (\text{V A}^{-1}) \tag{9.31}$$

where R_f is the value of the feedback resistor. In this case the permitted electric bandwidth B (without equalization) may be written as:

$$B \leqslant \frac{G}{2\pi R_f C_T} \tag{9.32}$$

Hence, comparing Eq. (9.32) with Eq. (9.20) it may be noted that the trans-

* The time constant can be obtained directly from Eq. (9.30) where the maximum bandwidth is defined by $\omega = 2\pi B = 1/R_{TL}C_T$.

impedance (or feedback) amplifier provides a much greater bandwidth than do the amplifiers without feedback. This is particularly pronounced when G is large.

Moreover, it is interesting to consider the thermal noise generated by the transimpedance front end. Using a referred impedance noise analysis it can be shown [Ref. 12] that to a good approximation the feedback resistance (or impedance) may be referred to the amplifier input in order to establish the noise performance of the configuration. Thus when $R_f \ll R_{TL}$, the major noise contribution is from thermal noise generated in R_f. The noise performance of this configuration is therefore improved when R_f is large, and it approaches the noise performance of the high impedance front end when $R_f = R_{TL}$. Unfortunately, the value of R_f cannot be increased indefinitely due to problems of stability with the closed loop design. Furthermore, it may be observed from Eq. (9.32) that increasing R_f reduces the bandwidth of the transimpedance configuration. This problem may be alleviated by making G as large as the stability of the closed loop will allow. Nevertheless, it is clear that the noise in the transimpedance amplifier will always exceed that incurred by the high impedance front end structure.

Example 9.8

A high input impedance amplifier which is employed in an optical fiber receiver has an effective input resistance of 4 MΩ which is matched to a detector bias resistor of the same value. Determine:

(a) The maximum bandwidth that may be obtained without equalisation if the total capacitance C_T is 6 pF.
(b) The mean square thermal noise current per unit bandwidth generated by this high input impedance amplifier configuration when it is operating at a temperature of 300 K.
(c) Compare the values calculated in (a) and (b) with those obtained when the high input impedance amplifier is replaced by a transimpedance amplifier with a 100 kΩ feedback resistor and an open loop gain of 400. It may be assumed that $R_f \ll R_{TL}$, and that the total capacitance remains 6 pF.

Solution: (a) Using Eq. (9.29), the total effective load resistance:

$$R_{TL} = \frac{(4 \times 10^6)^2}{8 \times 10^6} = 2\ \text{M}\Omega$$

Hence from Eq. (9.20) the maximum bandwidth is given by:

$$B = \frac{1}{2\pi R_{TL} C_T} = \frac{1}{2\pi \times 2 \times 10^6 \times 6 \times 10^{-12}}$$
$$= 1.33 \times 10^4\ \text{Hz}$$

The maximum bandwidth that may be obtained without equalization is 13.3 kHz.

(b) The mean square thermal noise current per unit bandwidth for the high impedance configuration following Eq. (9.14) is:

$$\overline{i_t^2} = \frac{4KT}{R_{TL}} = \frac{4 \times 1.381 \times 10^{-23} \times 300}{2 \times 10^6}$$

$$= 8.29 \times 10^{-27}\ \mathrm{A^2\,Hz^{-1}}$$

(c) The maximum bandwidth (without equalization) for the transimpedance configuration may be obtained using Eq. (9.32), where

$$B = \frac{G}{2\pi R_f C_T} = \frac{400}{2\pi \times 10^5 \times 6 \times 10^{-12}}$$

$$= 1.06 \times 10^8\ \mathrm{Hz}$$

Hence a bandwidth of 106 MHz is permitted by the transimpedance design.

Assuming $R_f \ll R_{TL}$, the mean square thermal noise current per unit bandwidth for the transimpedance configuration is given by:

$$\overline{i_t^2} = \frac{4KT}{R_f} = \frac{4 \times 1.381 \times 10^{-23} \times 300}{10^5}$$

$$= 1.66 \times 10^{-25}\ \mathrm{A^2\,Hz^{-1}}$$

The mean square thermal noise current in the transimpedance configuration is therefore a factor of 20 greater than that obtained with the high input impedance configuration.

The equivalent value in decibels of the ratio of these noise powers is:

$$\frac{\textit{Noise power in the transimpedance configuration}}{\textit{Noise power in the high input impedance configuration}} = 10 \log_{10} 20$$

$$= 13\ \mathrm{dB}$$

Thus the transimpedance front end in Example 9.8 provides a far greater bandwidth without equalization than the high impedance front end. However, this advantage is somewhat offset by the 13 dB noise penalty incurred with the transimpedance amplifier over that of the high input impedance configuration. Nevertheless it is apparent, even from this simple analysis, that transimpedance amplifiers may be optimized for noise performance, although this is usually obtained at the expense of bandwidth. This topic is pursued further in Ref. 13. However, wideband transimpedance designs generally give a significant improvement in noise performance over the low impedance front end structures using simple voltage amplifiers (see Problem 9.18). Finally it must be emphasized that the approach adopted in Example 9.8 is by no means rigorous and includes two important simplifications: firstly, that the thermal noise in the high impedance amplifier is assumed to be totally generated by the effective input resistance of the

device; and secondly, that the thermal noise in the transimpedance configuration is assumed to be totally generated by the feedback resistor when it is referred to the amplifier input. Both these assumptions are approximations, the accuracy of which is largely dependent on the parameters of the particular amplifier. For example, another factor which tends to reduce the bandwidth of the transimpedance amplifier is the stray capacitance C_f generally associated with the feedback resistor R_f. When C_f is taken into account the closed loop response of Eq. (9.31) becomes:

$$H_{CL}(\Omega) \simeq \frac{-R_f}{1 + j\omega R_f(C_T/G + C_f)} \tag{9.33}$$

However, the effects of C_f may be cancelled by employing a suitable compensating network [Ref. 14].

The other major advantage which the transimpedance configuration has over the high impedance front end is a greater dynamic range. This improvement in dynamic range obtained using the transimpedance amplifier is a result of the different attenuation mechanism for the low frequency components of the signal. The attenuation is accomplished in the transimpedance amplifier through the negative feedback and therefore the low frequency components are amplified by the closed loop rather than the open loop gain of the device. Hence for a particular amplifier the improvement in dynamic range is approximately equal to the ratio of the open loop to the closed loop gains. The transimpedance structure therefore overcomes some of the problems encountered with the other configurations and is often preferred for use in wideband optical fiber communication receivers [Ref. 15].

9.5 FET Preamplifiers

The lowest noise amplifier device which is widely available is the silicon field effect transistor (FET). Unlike the bipolar transistor, the FET operates by controlling the current flow with an electric field produced by an applied voltage on the gate of the device (see Figure 9.10) rather than with a base current. Thus the gate draws virtually no current, except for leakage, giving the device an extremely high input impedance (can be greater than 10^{14} ohms). This, coupled with its low noise and capacitance (no greater than a few picofarads), makes the silicon FET appear an ideal choice for the front end of the optical fiber receiver amplifier. However, the superior properties of the FET over the bipolar transistor are limited by its comparatively low transconductance g_m (no better than 5 millisiemens in comparison with at least 40 millisiemens for the bipolar). It can be shown [Ref. 13] that a figure of merit with regard to the noise performance of the FET amplifier is g_m/C_T^2. Hence the advantage of high transconductance together with low total capacitance C_T is apparent. Moreover, as $C_T = C_d + C_a$, it should be noted that the figure of merit is optimized when $C_a = C_d$. This requires FETs to be specifically matched to particular detectors, a procedure which device availability does not generally permit in current optical fiber receiver design. As indicated above, the gain of the FET is restricted.

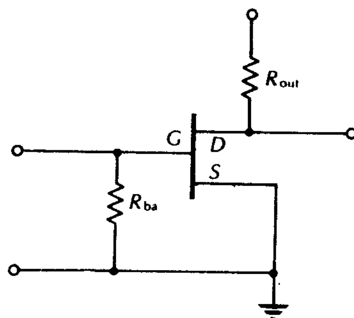


Figure 9.10 Grounded source FET configuration for the front end of an optical fiber receiver amplifier.

This is especially the case for silicon FETs at frequencies above 25 MHz where the current gain drops to values near unity as the transconductance is fixed with a decreasing input impedance. Therefore at frequencies above 25 MHz, the bipolar transistor is a more useful amplifying device.*

Figure 9.10 shows the grounded source FET configuration which increases the device input impedance especially if the amplifier bias resistor R_{ba} is large. A large bias resistor has the effect of reducing the thermal noise but it will also increase the low frequency impedance of the detector load which tends to integrate the signal (i.e. high impedance integrating front end). Thus compensation through equalization at a later stage is generally required.

9.5.1 Gallium arsenide MESFETs

Although silicon FETs have a limited useful bandwidth, much effort has been devoted to the development of high performance microwave FETs since the mid-1970s. These FETs are fabricated from gallium arsenide and, being Schottky barrier devices [Refs. 16 to 19], are called GaAs metal Schottky field effect transistors (MESFETs). They overcome the major disadvantage of silicon FETs in that they will operate with both low noise and high gain at microwave frequencies (GHz). Thus in optical fiber communication receiver design they present an alternative to bipolar transistors for wideband operation. These devices have therefore been incorporated into high performance receiver designs using both *p–i–n* and avalanche photodiode detectors [Refs. 21 to 32]. In particular, there has been much interest in hybrid integrated receiver circuits utilizing *p–i–n* photodiodes with GaAs MESFET amplifier front ends. The hybrid integration of a photodetector with a GaAs MESFET preamplifier having low leakage current,

* The figure of merit in relation to noise performance for the bipolar transistor amplifier may be shown [Ref. 13] to be $(h_{FE})^{\frac{1}{2}}/C_T$ where h_{FE} is the common emitter current gain of the device. Hence the noise performance of the bipolar amplifier may be optimized in a similar manner to that of the FET amplifier

low capacitance (less than 0.5 pF) and high transconductance (greater than 30 millisiemens) provides a strategy for low noise optical receiver design [Ref. 33].

9.5.2 PIN–FET hybrid receivers

The *p–i–n*/FET, or PIN–FET, hybrid receiver utilizes a high performance *p–i–n* photodiode followed by a low noise preamplifier often based on a GaAs MESFET, the whole of which is fabricated using thick film integrated circuit technology. This hybrid integration on a thick film substrate reduces the stray capacitance to negligible levels giving a total input capacitance which is very low (e.g. 0.4 pF). The MESFETs employed have a transconductance of approximately 15 millisiemens at the bandwidths required (e.g. 140 Mbit s^{-1}). Early work [Refs. 22 and 23] in the 0.8 to 0.9 μm wavelength band utilizing a silicon *p–i–n* detector showed the PIN–FET hybrid receiver to have a sensitivity of -45.8 dBm for a 10^{-9} bit error rate which is only 4 dB worse than current silicon RAPD receivers (see Section 8.9.2).

The work was subsequently extended into the longer wavelength band (1.1 to 1.6 μm) utilizing III–V alloy *p–i–n* photodiode detectors. An example of a PIN–FET hybrid high impedance (integrating) front end receiver for operation at a wavelength of 1.3 μm using an InGaAs *p–i–n* photodiode is shown in Figure 9.11 [Refs. 24 to 27]. This design, used by British Telecom, consists of a preamplifier with a GaAs MESFET and microwave bipolar transistor cascode followed by an emitter follower output buffer. The cascode circuit is chosen to ensure that sufficient gain is obtained from the first stage to give an overall gain of 18 dB. As

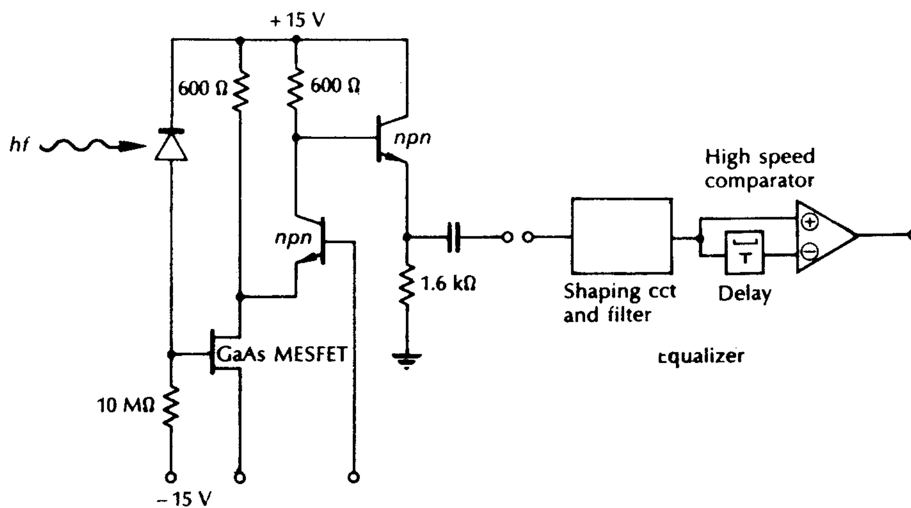


Figure 9.11 PIN–FET hybrid high impedance integrating front end receiver [Refs. 24 to 27]

the high impedance front end effectively integrates the signal, the following digital equalizer is necessary. The pulse shaping and noise filtering circuits comprise two passive filter sections to ensure that the pulse waveform shape is optimized and the noise is minimized. Equalization for the integration (i.e. differentiation) is performed by monitoring the change in the integrated waveform over one period with a subminiature coaxial delay line followed by a high speed low level comparator. The receiver is designed for use at a transmission rate of 140 Mbit s^{-1} where its performance is found to be comparable to germanium and III–V alloy APD receivers. For example, the receiver sensitivity at a bit error rate of 10^{-9} is −44.2 dBm.

When compared with the APD receiver the PIN–FET hybrid has both cost and operational advantages especially in the longer wavelength region. The low voltage operation (e.g. +15 and −15 V supply rails) coupled with good sensitivity and ease of fabrication makes the incorporation of this receiver into wideband optical fiber communication systems commercially attractive. A major drawback with the PIN–FET receiver is the possible lack of dynamic range. However, the configuration shown in Figure 9.11 gave adequate dynamic range via a control circuit which maintained the mean voltage at the gate at 0 V by applying a negative voltage proportional to the mean photocurrent to the MESFET bias resistor. With a −15 V supply rail an optical dynamic range of some 20 dB was obtained. This was increased to 27 dB by reducing the value of the MESFET bias resistor from 10 to 2 MΩ which gave a slight noise penalty of 0.5 dB. These figures compare favourably with practical APD receivers.

Transimpedance front end receivers have also been fabricated using the PIN–FET hybrid approach. An example of this type of circuit [Ref. 29] is shown

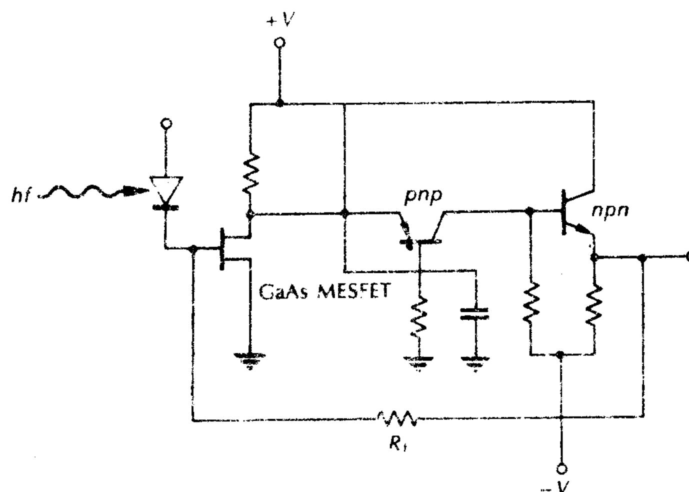


Figure 9.12 PIN–FET hybrid transimpedance front end receiver [Ref. 29].

in Figure 9.12. The amplifier consists of a GaAs MESFET followed by two complementary bipolar microwave transistors. A silicon *p–i–n* photodiode was utilized with the amplifier and the receiver was designed to accept data at a rate of 274 Mbits s^{-1}. In this case the effective input capacitance of the receiver was 4.5 pF giving a sensitivity around −35 dBm for a bit error rate of 10^{-9}.

These figures are somewhat worse than the high impedance front end design discussed previously. However, this design has the distinct advantage of a flat frequency response to a wider bandwidth which requires little, if any, equalization.

9.6 High performance receivers

It is clear from the discussions in Sections 9.3 to 9.5 that noise performance is a major design consideration providing a limitation to the sensitivity which may be obtained with a particular receiver structure and component mix. However, two other important receiver performance criteria were also outlined in the aforementioned sections, namely, bandwidth and dynamic range. Moreover, distinct tradeoffs exist between these three performance attributes such that an optimized design for one criterion may display a degradation in relation to one or both of the other criteria. Nevertheless, although high performance receiver design may seek to provide optimization for one particular attribute, attempts are generally made to minimize the degradations associated with the other performance parameters. In this section we describe further the strategies that have been adopted to produce high performance receivers for optical fiber communications, together with some of the performance results which have been obtained over the last few years.

As mentioned in Section 9.5.2, low noise performance combined with potential high speed operation has been a major pursuit in the hybrid integration of *p–i–n* photodiodes with GaAs MESFETs. In this context it is useful to compare the noise performance of various transistor preamplifiers over a range of bandwidths. A theoretical state-of-the-art performance comparison for the silicon junction FET and (JFET), the silicon metal oxide semiconductor FET (MOSFET) and the silicon bipolar transistor preamplifier with a GaAs MESFET device for transmission rates from 1 Mbit s^{-1} to 10 Gbit s^{-1} is shown in Figure 9.13 [Ref. 34]. It may be observed that at low speeds the three FET preamplifiers provide higher sensitivity than the Si bipolar device. In addition it is apparent that below 10 Mbit s^{-1} the Si MOSFET preamplifier provides a lower noise performance than the GaAs MESFET. Above 20 Mbit s^{-1}, however, the highest sensitivity is obtained with the GaAs MESFET device, even though at very high speeds the Si MOSFET and Si bipolar transistor preamplifiers exhibit a noise performance that is only slightly worse than the aforementioned device. Furthermore, it is clear that, as indicated in Section 9.5, the Si bipolar transistor preamplifier displays a noise improvement over the Si JFET, in this case at speeds above 50 Mbit s^{-1}.

The optimization of PIN–FET receiver designs for sensitivity and high speed operation have been investigated [Refs. 35, 36]. Also a wideband (10 GHz) low

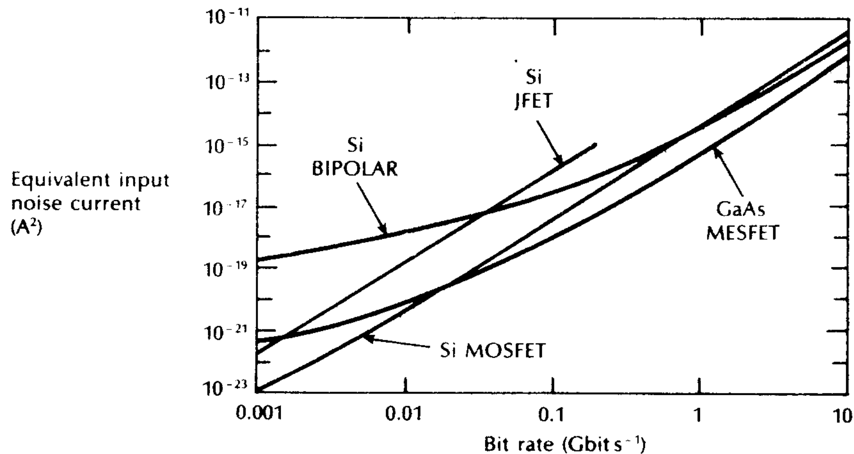
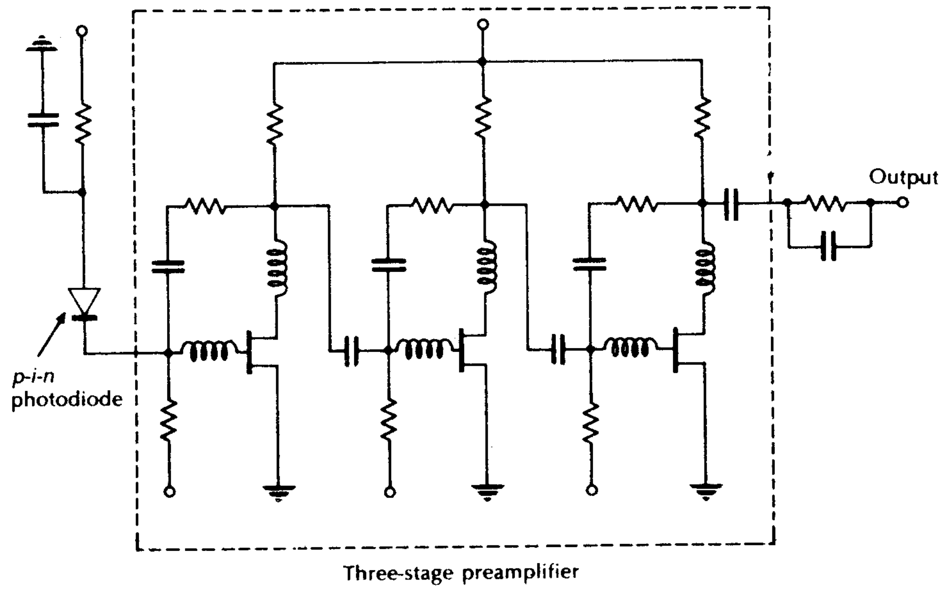


Figure 9.13 Noise characteristics for various optical receiver transistor preamplifiers. Reproduced with permission from B. L. Kasper, 'Receiver design', in S. E. Miller and I. P. Kaminow (Eds.), *Optical Fiber Telecommunications II*, Academic Press Inc., p. 689, 1988.

noise device using discrete commercial components has recently been reported [Ref. 37]. In addition, new high speed, low noise transistor types are under investigation for optical receiver preamplifiers. These devices include the heterojunction bipolar transistor (HBT) [Ref. 38] and high electron mobility transistor (HEMT) [Refs. 39, 40]. The latter device type comprises a selectively doped heterojunction FET which has displayed 3 dB bandwidths up to 20 GHz within the three stage optical preamplifier illustrated in Figure 9.14(a) [Ref. 39]. Each stage comprised a shunt feedback configuration containing a single HEMT with mutual conductance of 70 millisiemens and a gate to source capacitance of 0.36 pF (see Figure 9.14(b)). When operated with an InGaAs *p–i–n* photodiode the preamplifier exhibited a 21.5 dB gain with an averaged input equivalent noise current density of 7.6 pA $Hz^{-\frac{1}{2}}$ over the range 100 MHz to 18 GHz.

Although the above discussion has centred on *p–i–n* receiver preamplifier designs, high speed APD optical receivers are also under investigation [Refs. 41 to 43]. In particular, a high sensitivity APD–FET receiver designed to operate at speeds up to 8 Gbit s^{-1} and at wavelengths in the range 1.3 to 1.5 μm is shown in Figure 9.15 [Ref. 42]. The receiver employed a 60 GHz gain–bandwidth product InGaAs/InGaAsP/InP APD followed by a hybrid GaAs MESFET high impedance front-end. Moreover, a receiver sensitivity of −25.8 dBm was obtained for a bit error rate of 10^{-9}.

An additional strategy for the provision of wideband, low noise receivers, especially using the *p–i–n* photodiode detector, involves the monolithic integration of this device type with III–V semiconductor alloy FETs or HBTs [Refs. 44 to 48].



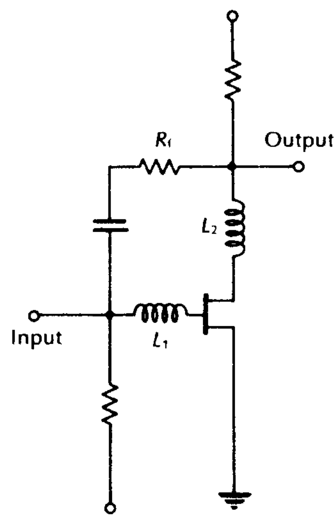


Figure 9.14 Circuit configuration for a high speed optical receiver using a HEMT preamplifier [Ref. 39]: (a) *p–i–n*–HEMT optical receiver; (b) single shunt feedback stage.

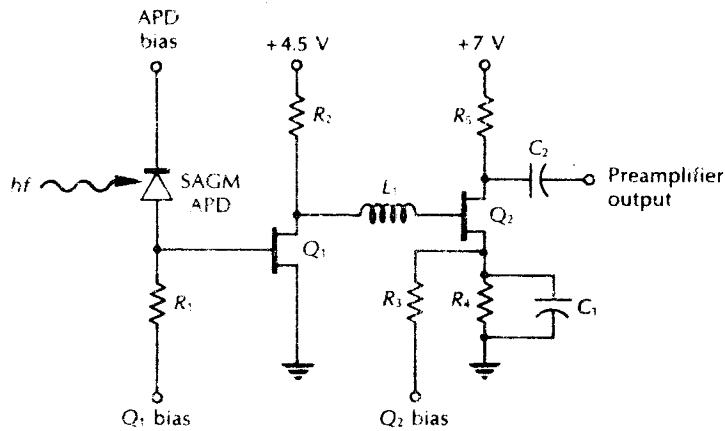


Figure 9.15 Circuit configuration for a high sensitivity APD–FET optical receiver [Ref. 42].

Such monolithic integrated receivers or optoelectronic integrated circuits (OEICs) are discussed further in Section 10.7. However, it should be noted that the major recent activities in this area have concerned devices for operation in the 1.1 to 1.6 μm wavelength range. An example of the circuit configuration of a monolithic PIN–FET receiver is illustrated in Figure 9.16 [Ref. 48]. The design comprises a voltage variable FET feedback resistor which produces active feedback as an input shunt automatic gain control (AGC) circuit which extends the dynamic range by diverting excess photocurrent away from the input of the basic receiver.

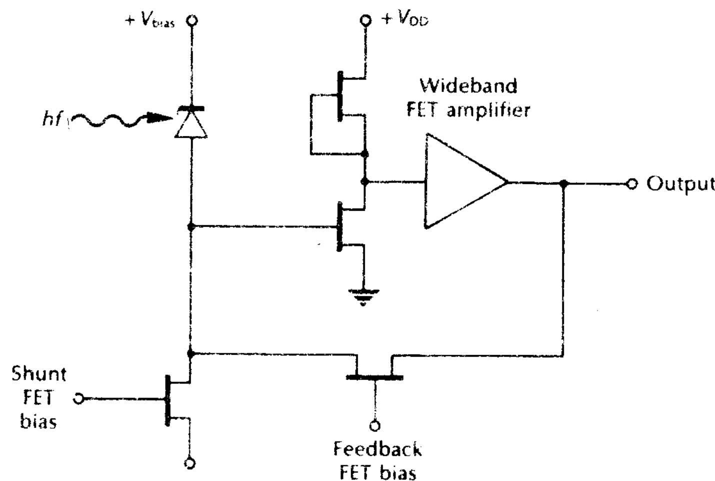


Figure 9.16 Monolithic PIN–FET optical receiver circuit configuration.

Furthermore, the shunt FET gives additional dynamic range extension through the mechanism of active receiver bias compensation, which is discussed further in relation to Figure 9.18.

The receiver dynamic range is an important performance parameter as it provides a measure of the difference between the device sensitivity and its saturation or overload level. A receiver saturation or overload level is largely determined by the value of the photodiode bias resistor or, alternatively, the feedback resistor in the transimpedance configuration. Because the photodiode bias resistor has a small value in the low impedence front end design, then the saturation level is high.* Similarly, the relatively low value of feedback resistor in the transimpedance configuration gives a high saturation level which combined with a high sensitivity, provides a wide dynamic range, as indicated in Section 9.4.3. By contrast the high value of photodiode bias resistor in the high impedance front end causes a low saturation level which, even taking account of the high sensitivity of the configuration, gives a relatively narrow dynamic range. The difference between the two latter receiver structures may be observed in the dynamic range and sensitivity characteristics shown in Figure 9.17. Although the sensitivity decreases in moving from the high impedance design (left hand side) to the transimpedance configuration (right hand side) as the value of the feedback resistor R_f is reduced, the saturation level increases at a faster rate, producing a significantly wider dynamic range for the transimpedance front end receiver.

The significance of the receiver dynamic range becomes apparent when the reader considers the ideal multipurpose use of such a device for operation with a variety of optical source powers over different fiber lengths. Moreover, when a high impedance receiver with a 1 MΩ bias resistor is utilized, the saturation level occurs at an input optical power of 0.5 μW or −33 dBm. Therefore, this device can only be employed in long-haul communication applications where the input power level is low. Corresponding figures for the transimpedance configuration (1 kΩ feedback resistor) and the low impedance front end (200 Ω bias resistor) are 0.5 mW (−3 dBm) and 2.5 mW (+4 dBm). In all cases the saturation level can be substantially improved by using active receiver bias compensation, as illustrated schematically in Figure 9.18. Hence, as the d.c. voltage at the input to the amplifier increases with the incident optical power, the control loop applies an equal but opposite shift in the voltage to the other side of the bias resistor. In this way the voltage at the input to the preamplifier becomes independent of the detected power level. However, in practice the feedback voltage in the control loop cannot be unbounded and therefore the technique has limitations. Nevertheless, saturation levels for high impedance front end receiver designs may be improved to around 20 μW, or 17 dBm using this technique.

Even when using bias compensation with a high impedance front end receiver to improve the saturation level, the overall dynamic range tends to be poor. For such

* Unfortunately, the sensitivity of the low impedance configuration is poor and hence the dynamic range is generally not large.

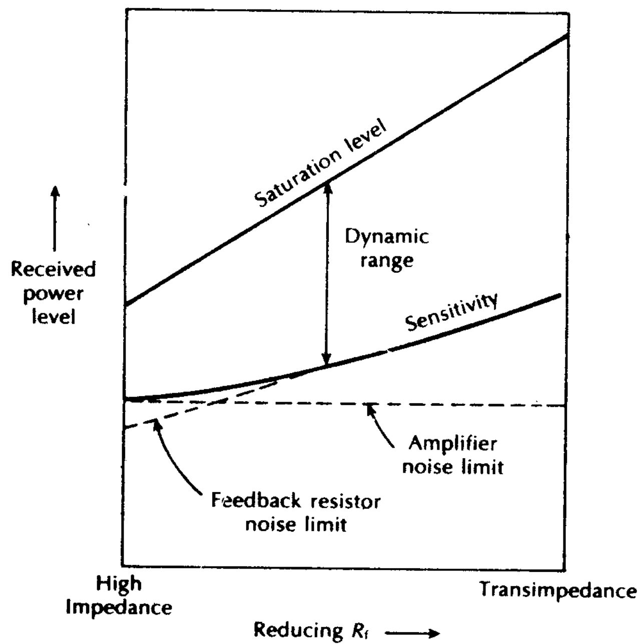


Figure 9.17 Characteristics illustrating the variation in received power level against the value of the feedback resistor R_f in the transimpedance front end receiver structure. The high impedance front end receiver corresponds to $R_f = \infty$.

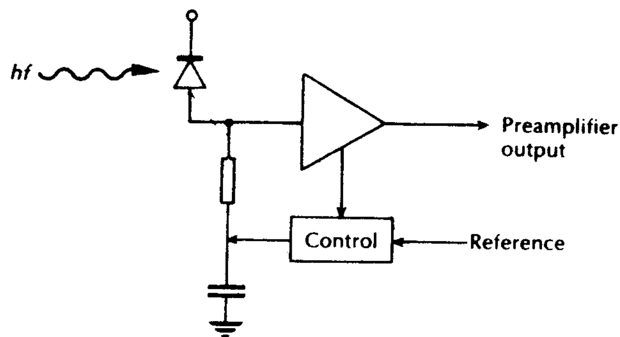


Figure 9.18 Active receiver bias compensation.

a receiver operating at a speed of 1 Gbit s^{-1} it is usually in the range 20 to 27 dB, whereas for a corresponding transimpedance receiver configuration without bias compensation, the dynamic range can be 30 to 39 dB.* Furthermore, in the latter case alternative design strategies have proved successful in increasing the receiver dynamic range. In particular the use of optically coupled feedback has demonstrated dynamic ranges of around 40 dB for *p–i–n* receivers operating at modest bit rates [Refs. 49, 50].

The optical feedback technique, which is shown schematically in Figure 9.19, eliminates the thermal noise associated with the feedback resistor in the transimpedance front end design. This strategy proves most useful at low transmission rates because in this case the feedback resistors employed are normally far smaller than the optimum value for low noise performance so as to maintain the resistor at a practical size (e.g. 1 MΩ). Moreover, large values of feedback resistor limit the dynamic range of the conventional transimpedance receiver structure, whilst also introducing parasitic shunt capacitance which can cause signal integration and hence restrict the bandwidth of the preamplifier. It may be observed from Figure 9.19 that the optical feedback signal is provided by an LED that is driven from the preamplifier output through a small resistor. This resistor acts as a load to generate an output voltage for the following amplifiers. The current feedback to the signal *p–i–n* photodetector is obtained from a second *p–i–n* photodiode which detects the optical feedback signal

The removal of the feedback resistor in the optical feedback technique allows low noise performance and hence high receiver sensitivity of the order of −64 dBm at transmission rates of 2 Mbit s^{-1} [Ref. 50]. In addition, as the feedback LED is a low impedance device that can be driven with a low output voltage, the problem

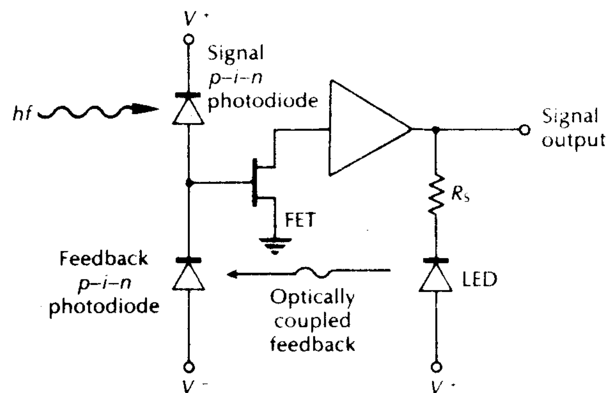


Figure 9.19 Optical feedback transimpedence receiver schematic.

* It should be noted that in both cases the bottom end of the range refers to *p–i–n* photodiode receivers whilst the top end of the range is only obtained with APD receivers.

associated with amplifier saturation is much reduced. Therefore this factor, combined with the high sensitivity produced by the strategy, enables wide dynamic range. It should be noted, however, that some penalties occur when employing this technique in that there is an increase in receiver input capacitance and also an increase in detector dark current noise (resulting from the use of two photodiodes) in comparison with the conventional transimpedance preamplifier structure. Nevertheless, it is suggested that the optical feedback receiver component costs can be comparable to resistive feedback designs [Ref. 50] whilst providing a significant performance improvement.

An alternative strategy for the realization of high sensitivity receivers, in this case for high speed operation, is to employ preamplification using an optical amplifier prior to the receiver [Refs. 51 to 55]. The two basic optical amplifier technological types, namely the semiconductor laser amplifier (SLA) and the fiber amplifier, are discussed in Sections 10.3 and 10.4 respectively. It is clear, however, that both device types may be utilized in this preamplification role which is illustrated schematically for an SLA device in Figure 9.20.* The SLA shown in Figure 9.20 operates as a near-travelling wave amplifier and therefore the output emissions are predominantly spontaneous creating a spectral bandwidth which is determined by the gain profile of the device. Because the typical spectral bandwidth is in the range 30 to 40 nm, a bandpass optical filter† is employed to reduce the intensity of the spontaneous emission reaching the optical detector. This has the effect of reducing the spontaneous noise products and thus improving the overall receiver sensitivity.

Although the sensitivity improvement introduced by the laser preamplifier is a function of the device internal gain, the coupling losses between the various elements and the bandwidth of the bandpass filter, it is typically in the range 10 to 15 dB when using an SLA. Moreover, it is interesting to observe from the sensitivity versus transmission rate characteristics shown in Figure 9.21 [Ref. 53] that the SLA preamplifier *p–i–n* photodiode configuration illustrated in Figure 9.20 displays a

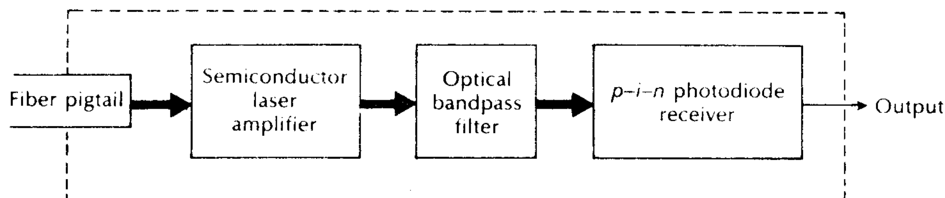


Figure 9.20 Block schematic of a SLA preamplified *p–i–n* photodiode receiver.

* It should be noted that the corresponding schematic showing a fiber amplifier fulfilling this role is provided in Figure 10.9(c).

† The optimum filter bandwidth is determined by a number of factors including the detector noise, the transmission rate, the transmitter chirp characteristics and the filter insertion loss but is typically in the range 0.5 to 3 nm.

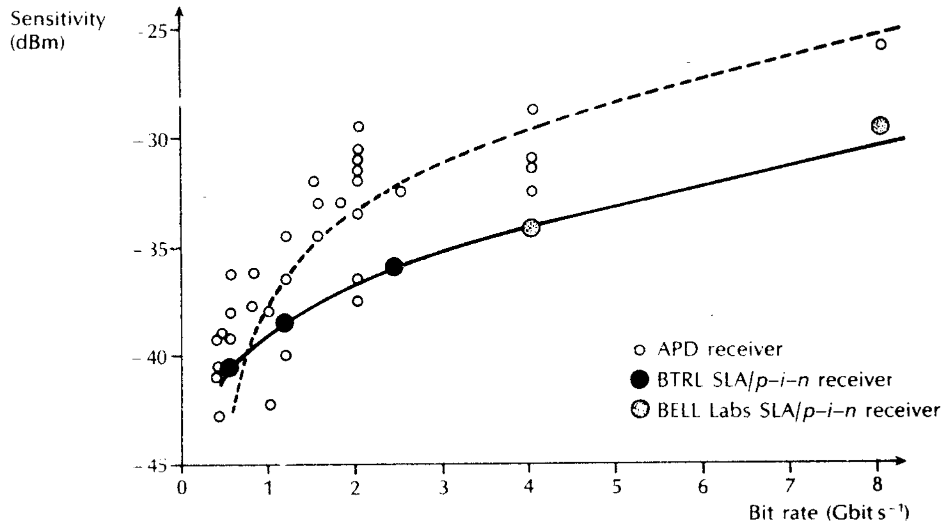


Figure 9.21 Characteristics showing receiver sensitivity against transmission bit rate for SLA-preamplified *p-i-n* photodiode receivers and APD receivers. Reproduced with permission from C. A. Hunter, L. N. Barker, D. J. T. Heatley, K. H. Cameron and R. L. Calton, 'The design and performance of semiconductor laser preamplified optical receivers', *IEE Colloq. Dig.* No. 1989/119, paper 9, 1989.

significant improvement over high performance APD receivers, particularly at speeds of 2.4 Gbit s^{-1} and above.

Improvements in overall receiver sensitivity have also been demonstrated using erbium doped fiber preamplifiers [Refs. 53, 54]. Such fiber amplifiers can provide much lower input and output coupling losses together with smaller amounts of spontaneous emission because of their narrower spectral bandwidths (4 to 10 nm) in comparison with SLAs. This latter attribute means that fiber preamplifiers can be used without the requirement for an optical bandpass filter between the device and the receiver. In this context the spontaneous noise bandwidth is simply determined by the gain characteristic of the fiber amplifier. A recent demonstration of an erbium doped fiber preamplified *p–i–n* photodiode receiver displayed an improvement in receiver sensitivity of 10.5 dB [Ref. 55].

Although in this chapter we have focused on receiver performance and design techniques for intensity modulated/direct detection optical fiber communication systems, many of the strategies discussed are also utilized within the generally more complex receiver structures required to enable coherent transmission. The various coherent demodulation schemes are discussed in some detail in Section 12.6 and the coherent receiver sensitivities are compared both with each other and with direct detection in Section 12.7. However, the specific preamplifier noise and technological considerations are not repeated as they apply equally to both detection techniques.

Problems

9.1 Briefly discuss the possible sources of noise in optical fiber receivers. Describe in detail what is meant by quantum noise. Consider this phenomenon with regard to:

(a) digital signaling;
(b) analog transmission,

giving any relevant mathematical formulae.

9.2 A silicon photodiode has a responsivity of 0.5 A W^{-1} at a wavelength of 0.85 μm. Determine the minimum incident optical power required at the photodiode at this wavelength in order to maintain a bit error rate of 10[illegible], when utilizing ideal binary signalling at a rate of 35 Mbit s^{-1}.

9.3 An analog optical fiber communication system requires an SNR of 40 dB at the detector with a post detection bandwidth of 30 MHz. Calculate the minimum optical power required at the detector if it is operating at a wavelength of 0.9 μm with a quantum efficiency of 70%. State any assumptions made.

9.4 A digital optical fiber link employing ideal binary signalling at a rate of 50 Mbit s^{-1} operates at a wavelength of 1.3 μm. The detector is a germanium photodiode which has a quantum efficiency of 45% at this wavelength. An alarm is activated at the receiver when the bit error rate drops below 10[illegible]. Calculate the theoretical minimum optical power required at the photodiode in order to keep the alarm inactivated. Comment briefly on the reasons why in practice the minimum incident optical power would need to be significantly greater than this value.

9.5 Discuss the implications of the load resistance on both thermal noise and post detection bandwidth in optical fiber communication receivers.

9.6 A silicon *p–i–n* photodiode has a quantum efficiency of 65% at a wavelength of 0.8 μm. Determine:

(a) the mean photocurrent when the detector is illuminated at a wavelength of 0.8 μm with 5 μW of optical power;
(b) the rms quantum noise current in a post detection bandwidth of 20 MHz;
(c) the SNR in dB, when the mean photocurrent is the signal.

9.7 The photodiode in Problem 9.6 has a capacitance of 8 pF. Calculate:

(a) the minimum load resistance corresponding to a post detection bandwidth of 20 MHz;
(b) the rms thermal noise current in the above resistor at a temperature of 25 °C;
(c) the SNR in dB resulting from the illumination in Problem 9.6 when the dark current in the device is 1 nA.

9.8 The photodiode in Problems 9.6 and 9.7 is used in a receiver where it drives an amplifier with a noise figure of 2 dB and an input capacitance of 7 pF. Determine:

(a) the maximum amplifier input resistance to maintain a post detection bandwidth of 20 MHz without equalization;
(b) the minimum incident optical power required to give an SNR of 50 dB.

9.9 A germanium photodiode incorporated into an optical fiber receiver working at a wavelength of 1.55 μm has a dark current of 500 nA at the operating temperature. When the incident optical power at this wavelength is 10^{-6} W and the responsivity of the device is 0.6 A W^{-1}, shot noise dominates in the receiver. Determine the SNR in dB at the receiver when the post detection bandwidth is 100 MHz

9.10 Discuss the expression for the SNR in an APD receiver given by:

$$\frac{S}{N} = \frac{M^2 I_p^2}{2eB(I_p + I_d)M^{2+x} + \dfrac{4KTBF_n}{R_L}}$$

with regard to the various sources of noise present in the receiver. How may this expression be modified to give the optimum avalanche multiplication factor?

9.11 A silicon RAPD has a quantum efficiency of 95% at a wavelength of 0.9 μm, has an excess avalanche noise factor of $M^{0.3}$ and a capacitance of 2 pF. It may be assumed that the post detection bandwidth (without equalization) is 25 MHz, and that the dark current in the device is negligible at the operating temperature of 290 K. Determine the minimum incident optical power which can yield an SNR of 23 dB.

9.12 With the device and conditions given in Problem 9.11, calculate:

(a) the SNR obtained when the avalanche multiplication factor for the RAPD falls to half the optimum value calculated;

(b) the increased optical power necessary to restore the SNR to 23 dB with $M = 0.5M_{op}$.

9.13 What is meant by the excess avalanche noise factor $F(M)$? Give two possible ways of expressing this factor in analytical terms. Comment briefly on their relative merits.

9.14 A germanium APD (with $x = 1.0$) operates at a wavelength of 1.35 μm where its responsivity is 0.45 A W^{-1}. The dark current is 200 nA at the operating temperature of 250 K and the device capacitance is 3 pF. Determine the maximum possible SNR when the incident optical power is 8×10^{-7} W and the post detection bandwidth without equalization is 560 MHz.

9.15 The photodiode in Problem 9.14 drives an amplifier with a noise figure of 3 dB and an input capacitance of 3 pF. Determine the new maximum SNR when they are operated under the same conditions.

9.16 Discuss the three main amplifier configurations currently adopted for optical fiber communications. Comment on their relative merits and drawbacks.

A high impedance integrating front end amplifier is used in an optical fiber receiver in parallel with a detector bias resistor of 10 MΩ. The effective input resistance of the amplifier is 6 MΩ and the total capacitance (detector and amplifier) is 2 pF.

It is found that the detector bias resistor may be omitted when a transimpedance front end amplifier design is used with a 270 kΩ feedback resistor and an open loop gain of 100.

Compare the bandwidth and thermal noise implications of these two cases, assuming an operating temperature of 290 K.

9.17 A *p–i–n* photodiode operating at a wavelength of 0.83 μm has a quantum efficiency of 50% and a dark current of 0.5 nA at a temperature of 295 K. The device is unbiased but loaded with a current mode amplifier with a 50 kΩ feedback resistor and an open loop gain of 32. The capacitance of the photodiode is 1 pF and the input capacitance of the amplifier is 6 pF.

Determine the incident optical power required to maintain a SNR of 55 dB when the post detection bandwidth is 10 MHz. Is equalization necessary?

9.18 A voltage amplifier for an optical fiber receiver is designed with an effective input resistance of 200 Ω which is matched to the detector bias resistor of the same value. Determine:

(a) The maximum bandwidth that may be obtained without equalization if the total capacitance (C_T) is 10 pF.

(b) The rms thermal noise current generated in this configuration when it is operating over the bandwidth obtained in (a) and at a temperature of 290 K. The thermal noise generated by the voltage amplifier may be assumed to be from the effective input resistance to the device.

(c) Compare the values calculated in (a) and (b) with those obtained when the voltage amplifier is replaced by a transimpedance amplifier with a 10 kΩ feedback resistor and an open loop gain of 50. It may be assumed that the feedback resistor is also used to bias the detector, and the total capacitance remains 10 pF.

9.19 What is a PIN–FET hybrid receiver? Discuss in detail its merits and possible drawbacks in comparison with the APD receiver.

9.20 Identify the characteristics which are of greatest interest in the pursuit of high performance receivers.

Discuss the major techniques which have been adopted in order to produce such high performance receivers for use in long-haul optical fiber communications.

Answers to numerical problems

9.2 −70.4 dBm
9.3 −37.2 dBm
9.4 −70.1 dBm
9.6 (a) 2.01 μA; (b) 3.59 nA; (c) 55.0 dB
9.7 (a) 994.7 Ω; (b) 18.19 nA; (c) 39.3 dB
9.8 (a) 1.137 kΩ; (b) 19.58 μW
9.9 40.1 dB
9.11 −50.3 dBm
9.12 (a) 14.2 dB; (b) −49.6 dBm
9.14 23.9 dB
9.15 21.9 dB
9.16 High impedance front end: 21.22 kHz, 4.27×10^{-27} $A^2 Hz^{-1}$; Transimpedance front end: 29.47 MHz, 5.93×10^{-26} $A^2 Hz^{-1}$
9.17 −23.1 dBm, equalization is unnecessary
9.18 (a) 159.13 MHz; (b) 160 nA; (c) 79.56 MHz, 11.3 nA, noise power 23 dB down

References

[1] (a) M. Schwartz, *Information Transmission, Modulation and Noise* (4th edn), McGraw-Hill, 1990. (b) F. R. Conner, *Noise*, (2nd edn), Edward Arnold, 1982.

[2] P. Russer, 'Introduction to optical communications', in M. J. Howes and D. V. Morgan (Eds.), *Optical Fibre Communications*, pp. 1–26, John Wiley, 1980.

[3] M. Garbuny, *Optical Physics*, Academic Press, 1965.

[4] W. M. Hubbard, 'Efficient utilization of optical frequency carriers for low and moderate bit rate channels', *Bell Syst. Tech. J.*, **50**, pp. 713–718, 1973.

[5] I. Garrett, 'Receivers for optical fibre communications', *Electron. and Radio Eng.*, **51**(7/8), pp. 349–361, 1981.

[6] P. P. Webb, R. J. McIntyre and J. Conradi, 'Properties of avalanche photodiodes', *RCA Rev.*, **35**, pp. 234–278, 1974.

[7] W. R. Bennett and J. R. Davey, *Data Transmission*, McGraw-Hill, 1965.

[8] S. D. Personick, 'Receiver design for digital fiber optic communication systems (Part I and II)', *Bell Syst. Tech. J.*, **52**, pp. 843–886, 1973

[9] T. P. Lee and T. Li. 'Photodetectors', in. S. E. Miller and A. G. Chynoweth (Eds.) *Optical Fiber Telecommunications*, pp. 593–623. Academic Press, 1979.

[10] S. D. Personick, 'Receiver design', in S. E. Miller and A. G. Chynoweth (Eds.), *Optical Fiber Telecommunications*, pp. 627–651, Academic Press, 1979.

[11] J. E. Goell, 'Input amplifiers for optical PCM receivers', *Bell Syst. Tech. J.*, **54**, pp. 1771–1793. 1974.

[12] J. L. Hullett and T. V. Muoi, 'Referred imepedance noise analysis for feedback amplifiers', *Electron. Lett.*, **13**(13), pp. 387–389, 1977.

[13] R. G. Smith and S. D. Personick, 'Receiver design for optical fiber communication systems', in H. Kressel (Ed.), *Semiconductor Devices for Optical Communication* (2nd edn), Springer-Verlag, 1982.

[14] J. L. Hullett, 'Optical communication receivers', *Proc. IREE Australia*, **40**(4), pp. 127–136, 1979.

[15] J. L. Hullett and T. V. Muoi, 'A feedback receiver amplifier for optical transmission systems', *Trans. IEEE*, **COM 24**, pp. 1180–1185, 1976.

[16] J. S. Barrera, 'Microwave transistor review, Part 1. GaAs field-effect transistors'. *Microwave J.* (USA), **19**(2), pp. 28–31, 1976.

[17] B. S. Hewitt, H. M. Cox, H. Fukui, J. V. Dilorenzo, W. O. Scholesser and D. E. Iglesias, 'Low noise GaAs MESFETs', *Electron. Lett.*, **12**(12), pp. 309–310, 1976.

[18] D. V. Morgan, F. H. Eisen and A. Ezis, 'Prospects for ion bombardment and ion implantation in GaAs and InP device fabrication', *IEE Proc.*, **128**(1–4), pp. 109–129, 1981.

[19] J. Mun, J. A. Phillips and B. E. Barry, 'High-yield process for GaAs enhancement-mode MESFET integrated circuits', *IEE Proc.*, **128**(1–4), pp. 144–147, 1981.

[20] S. D. Personick, P. Balaban, J. H. Bobsin and P. R. Kumar, 'A detailed comparison of four approaches to the calculation of the sensitivity of optical fiber system receivers', *IEEE Trans. Commun.*, **COM-25**, pp. 541–549, 1977.

[21] S. D. Personick, 'Design of receivers and transmitters for fiber systems', in M. K. Barnoski (Ed.), *Fundamentals of Optical Fiber Communications* (2nd edn), Academic Press, 1981.

[22] D. R. Smith, R. C. Hooper and I. Garrett, 'Receivers for optical communications: A comparison of avalanche photodiodes with PIN–FET hybrids', *Opt. Quant. Electron.*, **10**, pp. 293–300, 1978.

[23] R. C. Hooper and D. R. Smith, 'Hybrid optical receivers using PIN photodiodes, *IEE (London) Colloquium on Broadband High Frequency Amplifiers*, pp. 9/1–9/5, 1979.

[24] K. Ahmad and A. W. Mabbitt, '$Ga_{1-x}In_xAs$ photodetectors for 1.3 micron PIN–FET receiver', *IEEE NY (USA) International Electronic Devices Meeting* (Washington, DC), pp. 646–649, 1978.

[25] D. R. Smith, R. C. Hooper and R. P. Webb, 'High performance digital optical receivers with PIN photodiodes', *IEEE (NY) Proceedings of the International Symposium on Circuits and Systems* (Tokyo), pp. 511–514, 1979.

[26] D. R. Smith, R. C. Hooper, K. Ahmad, D. Jenkins, A. W. Mabbitt and R. Nicklin, '*p–i–n*/FET hybrid optical receiver for longer wavelength optical communication systems', *Electron. Lett.*, **16**(2), pp. 69–71, 1980.

[27] R. C. Hooper, D. R. Smith and B. R. White, 'PIN–FET Hybrids for digital optical receivers', *IEEE NY (USA) 30th Electronic Components Conference*, San Francisco, pp. 258–260, 1980.

[28] S. Hata, Y. Sugeta, Y. Mizushima, K. Asatani and K. Nawata, 'Silicon *p–i–n* photodetectors with integrated transistor amplifiers', *IEEE Trans. Electron. Devices*, **ED-26**(6), pp. 989–991, 1979.

[29] K. Ogawa and E. L. Chinnock, 'GaAs FET transimpedance front-end design for a wideband optical receiver', *Electron. Lett*, **15**(20), pp. 650–652, 1979.

[30] S. M. Abbott and W. M. Muska, 'Low noise optical detection of a 1.1 Gb/s optical data stream, *Electron. Lett.*, **15**(9), pp. 250–251, 1979.

[31] L. A. Godfrey, 'Designing for the fastest response ever – ultra high speed photodetection', *Opt. Spectra (USA)*, **13**(10), pp. 43, 46, 1979.

[32] R. I. MacDonald, 'High gain optical detection with GaAs field effect transistors', *Appl. Opt. (USA)*, **20**(4), pp. 591–594, 1981.

[33] M. Brain and T. P. Lee, 'Optical receivers for lightwave communication systems', *J. of Lightwave Technol.*, **LT-3**(6), pp. 1281–1300, 1985.

[34] B. L. Kasper, 'Receiver design', in S. E. Miller and I. P. Kaminow (Eds.), *Optical Fiber Telecommunications II*, Academic Press, pp. 689–722, 1988.

[35] G. P. Vella-Coleiro, 'Optimization of optical sensitivity of *p–i–n* FET receivers', *IEEE Electron. Device Lett.*, **9**(6), pp. 269–271, 1988.

[36] R. A. Minasian, 'Optimum design of 4-Gbit/s GaAs MESFET optical preamplifier', *J. of Lightwave Technol.*, **LT-5**(3), pp. 373–379, 1987.

[37] M.A.R. Violas, D. J.-T. Heatley, A. M. O. Duarte and D. M. Beddow, '10 GHz bandwidth low noise optical receiver using discrete commercial devices', *Electron. Lett.*, **26**(1), pp. 35–36, 1990.

[38] C. W. Farley, M. F. Chang, P. M. Asbeck, N. H. Sheng, R. Pierson, G. J. Sullivan, K. C. Wang and R. B. Nubling, 'High-speed (f_t = 78 GHz) AlInAs/GaInAs single heterojunction HBT', *Electron. Lett.*, **25**(13), pp. 846–847, 1989.

[39] N. Ohkawa, '20 GHz bandwidth low-noise HEMT preamplifier for optical receivers', *Electron. Lett.*, **24**(7), pp. 1061–1062, 1988.

[40] S. D. Walker, L. C. Blank, R. A. Garnham and J. M. Boggis, 'High electron mobility transistor lightwave receiver for broadband optical transmission system applications', *J. of Lightwave Technol.*, **7**(3), pp. 454–458, 1989.

[41] B. L. Kasper and J. C. Campbell, 'Multigigabit-per-second avalanche photodiode lightwave receivers', *J. of Lightwave Technol.*, **LT-5**(10), pp. 1351–1364, 1987.

[42] B. L. Kasper, J. C. Campbell, J. R. Talman, A. H. Gnauck, J. E. Bowers and W. S. Holden, 'An APD/FET optical receiver operating at 8 Gbit/s', *J. of Lightwave Technol.*, **LT-5**(3), pp. 344–347, 1987.

[43] J. J. O'Reilly and R. S. Fyath, 'Performance of optical receivers employing ultralow noise avalanche photodiodes', *J. of Optical Communications*, **9**(3), pp. 82–84, 1988.

[44] M. J. N. Sibley, R. T. Unwin, D. R. Smith, B. A. Boxall and R. J. Hawkins, 'A monolithic common collector front-end optical preamplifier', *J. of Lightwave Technol.*, **LT-3**(1), pp. 13–15, 1985.

[45] Y. Archambault, D. Pavlidis and J. P. Guet, 'GaAs monolithic integrated optical preamplifier', *J. of Lightwave Technol.*, **LT-5**(3), pp. 355–366, 1987.

[46] W. T. Colleran and A. A. Abidi, 'Wideband monolithic GaAs amplifier using cascodes', *Electron. Lett.*, **23**(18), pp. 951–952, 1987.

[47] S. Miura, T. Mikawa, T. Fujii and O. Wada, 'High-speed monolithic GaInAs pinFET', Electron. Lett., **24**(7), pp. 394–395, 1988.

[48] G. F. Williams and H. P. Leblanc, 'Active feedback lightwave receivers', *J. of Lightwave Technol.*, **LT-4**(10), pp. 1502–1508, 1986.

[49] B. L. Kasper, A. R. McCormick, C. A. Burrus Jr and J. R. Talman, 'An optical-feedback transimpedance receiver for high sensitivity and wide dynamic range at low bit rates', *J. of Lightwave Technol.*, **6**(2), pp. 329–338, 1988.

[50] S. G. Methley, 'An optical feedback receiver, with high sensitivity', *Proc. SPIE Int. Soc. Opt. Eng., Fibre Optics' 88*, **949**, pp. 51–55, April 1988.

[51] I. W. Marshall and M. J. O'Mahony, '10 GHz optical receiver using a travelling wave semiconductor laser preamplifier', *Electron. Lett.*, **23**(20), p. 1052, 1987.

[52] N. A. Olsson and M. G. Oberg, 'Ultra low reflectivity 1.5 μm semiconductor laser preamplifier', *Electron. Lett.*, **24**(9), pp. 569–570, 1988.

[53] C. A. Hunter, L. N. Barker, D. J. T. Heatley, K. H. Cameron and R. L. Calton, 'The design and performance of semiconductor laser preamplifier optical receivers', *IEE Colloquium on Optical Amplifiers for Communication*. Dig.No. 1989/119, paper 9, 27 October, 1989.

[54] C. R. Giles, J. L. Desurvire, J. L. Zyskind and J. R. Simpson, 'Near quantum limited erbium doped fiber preamplifier with 215 photons/bit sensitivity at 1.8 Gbit/s', *IOOC 89*, Kobe, Japan, paper 20PDA-5, 1989.

[55] M. J. Pettitt, R. A. Baker and A. Hadifotiou, 'System performance of optical fibre preamplifier', *Electron. Lett.*, **25**(4), pp. 273–275, 1989.

10

Optical amplification and integrated optics

10.1 Introduction
10.2 Optical amplifiers
10.3 Semiconductor laser amplifiers
10.4 Fiber amplifiers
10.5 Integrated optics
10.6 Some integrated optical devices
10.7 Optoelectronic integration
10.8 Optical bistability and digital optics
10.9 Optical computation
Problems
References

10.1 Introduction

The preceding four chapters have been concerned with the devices employed to provide the electrical–optical interfaces within optical fiber communications. Optical sources were dealt with in Chapters 6 and 7 followed by optical detectors in Chapter 8 prior to consideration of optical receiver noise and its effect on receiver design in Chapter 9. These electrical–optical and optical–electrical conversion devices are crucial components for the realization of optical fiber communications, as may be observed in the following two chapters which discuss optical fiber systems. However, these devices are also, in a number of respects, a limiting factor

within the implementation of optical fiber systems. The conversion of the information signal from the electrical domain to the optical domain and vice-versa often provides a bottleneck within optical fiber communications which may restrict both the operating bandwidth and the quality of the transmitted signal. Performing operations on signals in the optical domain combined with the pursuit of more efficient mechanisms to provide the electrical–optical interfaces has therefore assumed increasing significance within optical fiber communications and its associated application areas.

The above considerations have stimulated a growing activity in the area of active devices and components which allow optical signals to be manipulated without returning them back to the electrical regime where such operations have normally been carried out in the past. Potentially, such devices alleviate the possible bottleneck associated with the interfaces as well as providing more efficient, and hence cost effective, methods for processing the optical signals. Moreover, in some cases the use of these devices and components may represent the only realistic solution for the implementation of particular optical fiber transmission techniques and systems.

This chapter therefore discusses the major developments in the area of active optical devices which may be utilized for a variety of functions within optical fiber communications. In addition it deals with the technologies associated with the integration of such optical and optoelectronic devices into circuits and subsystems (i.e. integrated optics and optoelectronic integration) which are becoming important areas within the design strategies for advanced optical fiber communication systems. A major development which was stimulated by the massive effort to produce laser sources for optical fiber communications is that of optical amplification. The technology associated with such active optical devices is now well established and their use within optical fiber communications is assured.

Section 10.2 introduces the concept of an optical amplifier and outlines the various generic types that are under investigation. Semiconductor laser amplifiers are then considered in some detail in Section 10.3. This is followed in Section 10.4 with a discussion of the various types of fiber amplifier which have evolved more recently, assisted by the activities which have led to the realization of fiber lasers (see Section 6.9.2).

The concepts and technology associated with integrated optics are then introduced in Section 10.5 prior to consideration of some common integrated optical devices in Section 10.6. In particular the latter section concentrates on directional couplers and switches, together with modulator devices which provide an alternative to direct current modulation for optical sources. Developments in the area of optoelectronic integration are then dealt with in Section 10.7, particular emphasis in this technology being concerned with the implementation of transmitters/receivers and switches on a single substrate.

Optical bistability and digital optics are then discussed in Section 10.8 to provide the reader with an insight into this important area, together with an understanding of how these phenomena may be utilized within optical fiber communications.

Finally, in Section 10.9 developments in the field of optical computation are outlined. Although it is clear that these latter developments do not, at present, significantly influence optical fiber communications, it is likely that in the future there will be a requirement for the combination of optical communication, optical switching and optical computational technologies, within what is fast becoming a predominantly optical fiber telecommunication network.

10.2 Optical amplifiers

Optical amplifiers, as their name implies, operate solely in the optical domain with no interconversion of photons to electrons. Therefore, instead of using regenerative repeaters which, as currently implemented, require optoelectronic devices for source and detector, together with substantial electronic circuitry for pulse slicing, retiming and shaping (see Section 11.6.1), optical amplifiers can be placed at intervals along a fiber link to provide linear amplification of the transmitted optical signal. The optical amplifier, in principle, provides a much simpler solution in that it is a single in-line component which can be used for any kind of modulation at virtually any transmission rate. Moreover, such a device can be bidirectional and if it is sufficiently linear it may allow multiplex operation of several signals at different optical wavelengths (i.e. wavelength division multiplexing). In particular with single-mode fiber systems, the effects of signal dispersion can be small and hence the major limitation on repeater spacing becomes attenuation due to fiber losses. Such systems do not require full regeneration of the transmitted digital signal at each repeater, and optical amplification of the signal proves sufficient. Hence over recent years optical amplifiers have emerged as promising network elements not just for use as linear repeaters but as optical gain blocks, optical receiver preamplifiers and, when used in a nonlinear mode, as optical gates, pulse shapers and routeing switches [Ref. 1].

The two main approaches to optical amplification to date have concentrated on semiconductor laser amplifiers which utilize stimulated emission from injected carriers and fiber amplifiers in which gain is provided by either stimulated Raman or Brillouin scattering* (see Sections 3.5 and 3.14), or by rare earth dopants (see Section 6.9.2). Both amplifier types (i.e. semiconductor and fiber; specifically rare earth and Raman) have the ability to provide high gain over wide spectral bandwidths, making them eminently suitable for future optical fiber system applications.

The typical gain profiles for various optical amplifier types based around the 1.5 μm wavelength region are illustrated in Figure 10.1 [Ref. 2]. It may be observed that the InGaAsP travelling wave semiconductor laser amplifier (TWSLA), the erbium doped fiber amplifier and the Raman fiber amplifier all provide wide

* Amplification from both stimulated Raman or Brillouin scattering can occur in undoped relatively long fiber lengths ($\simeq$ 10 km) or doped short lengths ($\simeq$ 10 m) of fiber [Ref. 2].

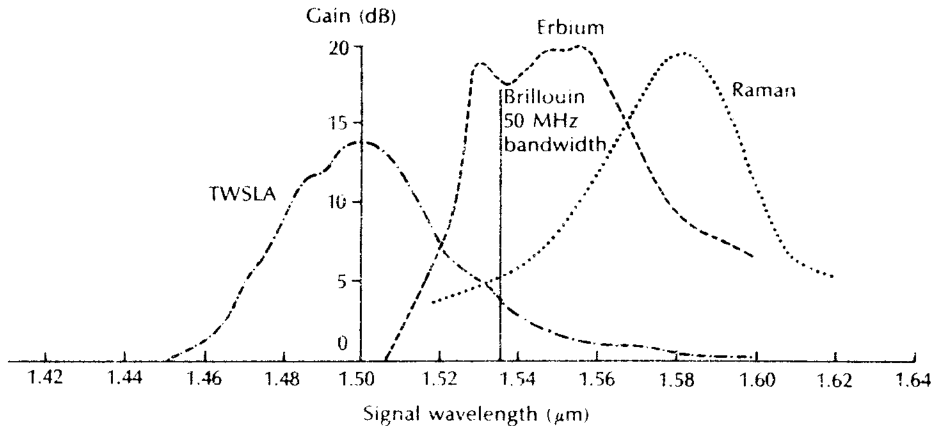


Figure 10.1 Optical amplifier gain characteristics. Reproduced with permission from P. Cochrane, *Br. Telecom. Technol. J.*, **8** (2), p. 5, 1990.

spectral bandwidths. Hence these optical amplifier types lend themselves to applications involving wavelength division multiplexing [Ref. 3]. By contrast, the Brillouin fiber amplifier has a very narrow spectral bandwidth, possibly around 50 MHz and therefore cannot be employed for wideband amplification. It could, however, be used for channel selection within a WDM system by allowing amplification of a particular channel without boosting other nearby channels.

Whereas semiconductor laser amplifiers exhibit low power consumption and their single-mode waveguide structures make them particularly appropriate for use with single-mode fiber, it is fiber amplifiers which present fewer problems of compatibility for in-line interconnection within optical fiber links [Ref. 4]. At present, semiconductor laser amplifiers are the most developed optical amplifier generic type but research into fiber amplifiers has also made rapid progress towards commercial products over the last few years.

10.3 Semiconductor laser amplifiers

The semiconductor laser amplifier (SLA) is based on the conventional semiconductor laser structure where the output facet reflectivities are between 30 and 35% [Refs. 4, 5]. SLAs can be used in both nonlinear and linear modes of operation [Ref. 1]. Various types of SLA may be distinguished including the resonant or Fabry–Perot amplifier which is an oscillator biased below oscillation threshold [Ref. 6], the travelling wave (TW) and the near travelling wave (NTW) amplifiers which are effectively single pass devices [Refs. 1, 7] and the injection locked laser amplifier, which is a laser oscillator designed to oscillate at the incident signal frequency [Ref. 8]. Such devices are capable of providing high internal gain

(15 to 35 dB) with low power consumption and their single-mode waveguide structure makes them particularly suitable for use with single-mode fiber. SLAs can, however, be classified into two main groups which are Fabry–Perot amplifiers (FPAs) and travelling wave amplifiers (TWAs)* [Refs. 1, 9], the difference between these groups being the facet reflectivities. A schematic diagram of an SLA is shown in Figure 10.2. It is based on the conventional semiconductor laser structure (gain- or index-guided) with an active region width w, thickness d and length L. When the input and output laser facet reflectivities denoted by R_1 and R_2 are each around 0.3, which depicts a normal semiconductor laser, then an FPA is obtained.† In this case, as the facet reflectivity is large, a highly resonant amplifier is formed and the transmission characteristic comprises very narrow passbands, as displayed in Figure 10.3. The mode zero corresponds to the peak gain wavelength and the mode

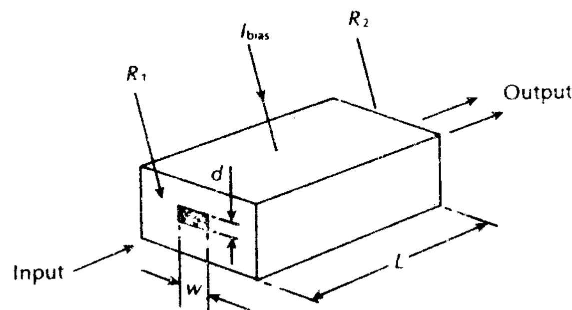


Figure 10.2 Schematic structure of the semiconductor laser amplifier.

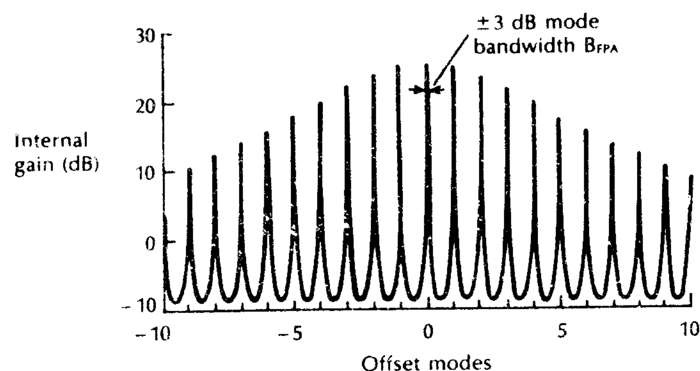


Figure 10.3 The Fabry–Perot amplifier passband where mode 0 corresponds to the peak gain wavelength [Ref. 1].

* The neumonics for the travelling wave semiconductor laser amplifiers are sometimes given as TWSLA [e.g. Refs. 2, 10].
† A FPA may be defined as an amplifier with facet reflectivities in the order 0.01 to 0.3 [Ref. 1].

spacing $\delta\lambda$ can be obtained from Eq. (6.16). For operation, the FPA is biased below the normal lasing threshold current, and light entering one facet appears amplified at the other facet together with inherent noise. In practice, the amplifier chip is bonded into a package with single-mode fiber pigtails which are used to guide light into and out of the amplifier. The inherent filtering of the FPA, although useful in certain applications, means the device is very sensitive to fluctuations in bias current, temperature and signal polarization. However, because of the resonant nature of FPAs, combined with their high internal fields, they are used within nonlinear applications: for example, to provide pulse shaping and bistable elements (see Section 10.8).

To form a travelling wave SLA antireflection coatings may be applied to the laser facets to reduce or eliminate the end reflectivities. This can be achieved by depositing a thin layer of silicon oxide or silicon nitride on the end facets such that the reflectivities are reduced to 1×10^{-3} or less. Such a device becomes a TWA operating in the single-pass amplification mode in which the Fabry–Perot resonance is suppressed by the reduction in facet reflectivity.* This has the effect of substantially increasing the amplifier spectral bandwidth and it makes the transmission characteristics less dependent upon fluctuations in bias current, temperature and input signal polarization. Hence the TWA proves superior to the FPA (particularly for linear applications) and also provides advantages in relation to both signal gain saturation and noise characteristics [Ref. 9]. Moreover, antireflection facet coatings have the effect of increasing the lasing current threshold, as illustrated in Figure 10.4 and so in practice such SLAs are operated at currents far beyond the normal lasing threshold current.

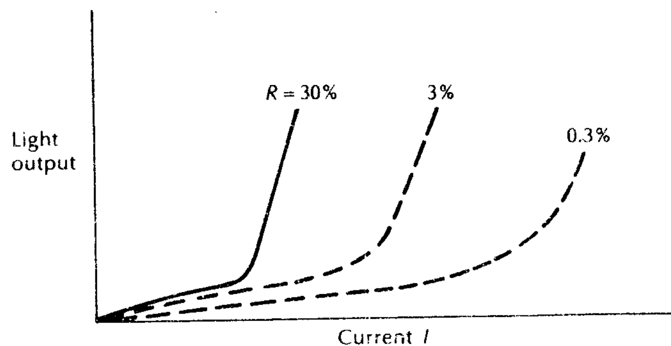


Figure 10.4 Light output against current characteristic for the semiconductor laser amplifier with different values of facet reflectivity R.

* In theory, a true TWA is the limiting case of a device with facets exhibiting zero reflectivity. However, in practice, even with the best antireflection coatings, some residual facet reflectivity remains (e.g. a low reflectivity of 1×10^{-4} has been obtained at a wavelength of 1.5 μm [Ref. 11]). Hence such devices are also referred to as near travelling wave amplifiers [Refs. 1, 12].

10.3.1 Theory

The general equation for the cavity gain G of a SLA as a function of signal frequency f takes the form [Refs. 1, 9, 10]:

$$G(f) = \frac{(1 - R_1)(1 - R_2)G_s}{(1 - \sqrt{R_1 R_2 G_s})^2 + 4\sqrt{R_1 R_2 G_s} \sin^2 \phi} \tag{10.1}$$

where R_1 and R_2 are the input and output facet reflectivities respectively, G_s is the single pass gain and ϕ is the single pass phase shift through the amplifier. It should be noted that Eq. (10.1) does not include coupling losses to and from the amplifier and that the phase shift ϕ may be written as [Ref. 9]:

$$\phi = \frac{\pi(f - f_o)}{\delta f} \tag{10.2}$$

where f_o is the Fabry–Perot resonant frequency and δf is the free spectral range of the SLA.

The 3 dB spectral bandwidth of an FPA, or essentially the ± 3 dB single longitudinal mode bandwidth defined by the FWHP points B_{FPA}, is shown in Figure 10.3. It may be observed that using Eqs. (10.1) and (10.2), B_{FPA} may be expressed as:

$$B_{FPA} = 2(f - f_o) = \frac{2\delta f}{\pi} \sin^{-1}\left[\frac{1 - \sqrt{R_1 R_2 G_s}}{2(\sqrt{R_1 R_2 G_s})^{\frac{1}{2}}}\right]$$

$$= \frac{c}{\pi n L} \sin^{-1}\left[\frac{1 - \sqrt{R_1 R_2 G_s}}{2(\sqrt{R_1 R_2 G_s})^{\frac{1}{2}}}\right] \tag{10.3}$$

where the mode separation frequency interval δf given by Eq. (6.14) combines the velocity of light c and the refractive index of the amplifier medium n with its length L. Alternatively the 3 dB spectral or optical bandwidth may be expressed as a function of the FPA cavity gain G following [Ref. 1]:

$$B_{FPA} = \frac{c}{\pi n L} \sin^{-1}\left[\frac{1}{2}\left(\frac{(1 - R_1)(1 - R_2)}{\sqrt{R_1 R_2} G}\right)\right] \tag{10.4}$$

Example 10.1

An uncoated FPA has facet reflectivities of 30% and a single pass gain of 4.8 dB. The amplifier has a 300 μm long active region, a mode spacing of 1 nm and a peak gain wavelength of 1.5 μm. Determine the refractive index of the active medium and the 3 dB spectral bandwidth of the device.

Solution: The refractive index of the active medium at the peak gain wavelength

may be obtained by rearranging Eq. (6.16) such that:

$$n = \frac{\lambda^2}{2\delta\lambda L} = \frac{(1.5 \times 10^{-6})^2}{2 \times 1 \times 10^{-9} \times 300 \times 10^{-6}}$$

$$= 3.75$$

Using Eq. (10.3) the 3 dB spectral bandwidth is given by:

$$B_{\text{FPA}} = \frac{c}{\pi n L} \sin^{-1}\left[\frac{1 - \sqrt{R_1 R_2} G_s}{2(\sqrt{R_1 R_2} G_s)^{\frac{1}{2}}}\right]$$

$$= \frac{2.998 \times 10^8}{\pi \times 3.75 \times 300 \times 10^{-6}} \sin^{-1}\left[\frac{1 - \sqrt{0.09} \times 3.020}{2(\sqrt{0.09} \times 3.020)^{\frac{1}{2}}}\right]$$

$$= 8.482 \times 10^{10} \sin^{-1}\left[\frac{0.040}{1.904}\right]$$

$$= 8.482 \times 10^{10} \times 0.494 = 4.2 \text{ GHz}.$$

The above result demonstrates the narrow spectral bandwidth obtained with an uncoated FPA.

The single pass gain G_s defined in terms of the device parameters and the applied bias current following Eq. (6.18) for the semiconductor laser* is generally written in the form:

$$G_s = \exp[\bar{g}L] \tag{10.5}$$

where $\bar{g}$ is the nett gain coefficient per unit length and L is the amplifier active length. However, the nett gain per unit length $\bar{g}$ may be defined in terms of the material gain coefficient g_m, the optical confinement factor Γ, and the effective loss coefficient per unit length $\bar{\alpha}$ as [Ref. 1]:

$$\bar{g} = \Gamma g_m - \bar{\alpha} \tag{10.6}$$

Furthermore, the material gain coefficient g_m is related to the signal intensity I following [Ref. 1]:

$$g_m = \frac{g_o}{1 + I/I_s} \tag{10.7}$$

where g_o is the unsaturated material gain coefficient in the absence of the input signal and I_s is the saturation intensity. Hence substitution of Eqs. (10.6) and (10.7)

* The round trip gain of Eq. (6.18) includes a factor of 2 which is omitted for the single pass gain.

in Eq. (10.5) for the single pass gain gives:

$$G_s = \exp[(\Gamma g_m - \bar{\alpha})L] \tag{10.8}$$

$$= \exp\left[\left(\frac{\Gamma g_o}{1 + I/I_s} - \bar{\alpha}\right)L\right] \tag{10.9}$$

It may be observed from Eqs. (10.8) and (10.9) that the single pass gain decreases with increasing intensity and that the material gain coefficient is reduced by a factor of 2 when the internal signal intensity I is equal to the saturation intensity I_s.

The phase shift ϕ_s associated with the single pass amplifier includes the nominal phase shift ϕ_o and an additional component resulting from the change in carrier density from the nominal density in the absence of a signal. Hence the total phase shift is given by [Ref. 13]:

$$\phi_s = \phi_o + \frac{g_o bL}{2}\left(\frac{I}{I + I_s}\right) \tag{10.10}$$

where b is the linewidth broadening factor, and the nominal phase shift is:

$$\phi_o = \frac{2\pi nL}{\lambda} \tag{10.11}$$

where n is the material refractive index. Thus Eqs. (10.9) and (10.10) indicate that both single pass gain and phase are functions of optical intensity. It is clear that for a constant signal intensity (i.e. with frequency modulation) there is no inherent signal distortion; however, with a time-varying intensity the gain and phase may also change with time, causing signal distortion. Furthermore, as G_s and ϕ_s are functions of the input signal intensity, then the SLA will exhibit nonlinear and bistable characteristics at high input powers.

It may be noted that Figure 10.3 shows the general form of the gain versus wavelength for an SLA obtained from Eq. (10.1). Furthermore, Eqs. (10.3) and (10.4) give the 3 dB spectral bandwidth for an FPA as the bandwidth of one longitudinal mode. The 3 dB spectral bandwidth, however, of a TWA is determined by the full gain width of the amplifier medium itself, as illustrated in Figure 10.5(a) rather than the Fabry–Perot gain profile. Hence the 3 dB bandwidth of a TWA is three orders of magnitude larger than that of an FPA [Ref. 9]. Nevertheless, the passband comprises peaks and troughs whose relative amplitudes are determined by the facet reflectivities, the single pass gain (and hence the applied bias current) and the input intensity. This gain undulation or peak–trough ratio of the passband ripple ΔG, which is defined as the difference between the resonant and nonresonant signal gain, may be observed in Figure 10.5(b). It is given by [Refs. 1, 9]:

$$\Delta G = \left(\frac{1 + \sqrt{R_1 R_2 G_s}}{1 - \sqrt{R_1 R_2 G_s}}\right)^2 \tag{10.12}$$

For wideband operation the peak–trough ratio must be small and for convenience is normally considered to be less than 3 dB for TWAs over their signal-

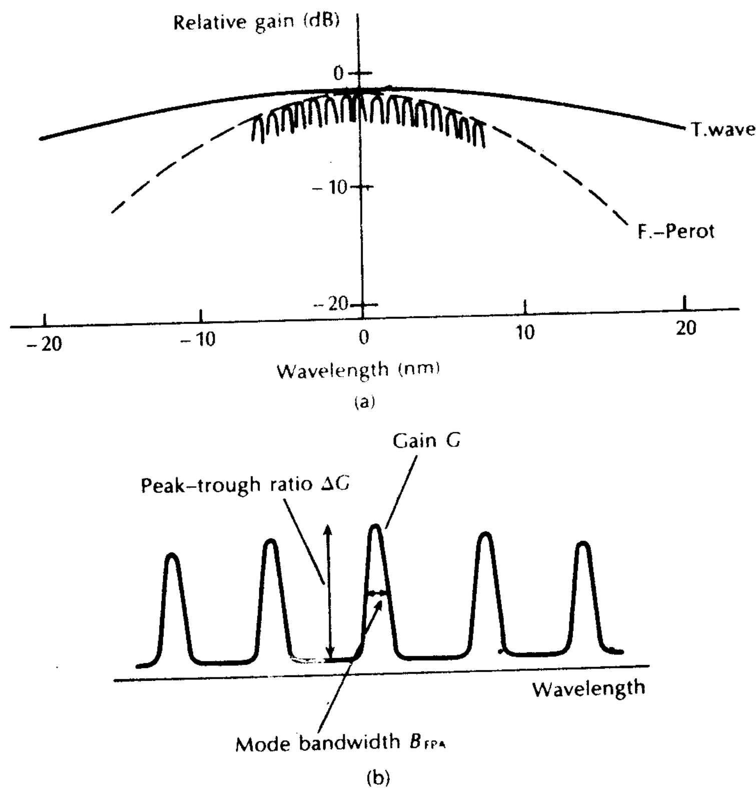


Figure 10.5 Passband characteristics for semiconductor laser amplifiers: (a) overall characteristics for both the travelling wave and Fabry–Perot amplifiers showing the large passband ripple in the latter case; (b) illustration of the peak–trough ratio of the passband ripple given in Eq. (10.12).

gain spectrum* [Refs. 1, 9]. Hence an amplifier whose gain ripple significantly exceeds 3 dB is usually categorized as an FPA.

Example 10.2

Derive an approximate expression for the cavity gain of a TWA in the limiting case of a 3 dB peak–trough ratio.

Solution: For a 3 dB peak–trough ratio Eq. (10.12) becomes:

$$\left(\frac{1 + \sqrt{R_1R_2G_s}}{1 - \sqrt{R_1R_2G_s}}\right)^2 = 0.5$$

* Sometimes this definition is said to apply to a near travelling wave amplifier as, in theory, a gain ripple of zero would correspond to a pure TWA.

Therefore,

$$1 + \sqrt{R_1R_2}G_s = 0.707\ (1 - \sqrt{R_1R_2}G_s)$$

$$\sqrt{R_1R_2}G_s = \frac{0.293}{1.707} = 0.172$$

For a TWA $R_1, R_2 \ll 1$ and, assuming a zero single pass phase shift, Eq. (10.1) becomes

$$G \simeq \frac{G_s}{(1 - \sqrt{R_1R_2}G_s)^2}$$

Substituting for $\sqrt{R_1R_2}G_s$ gives:

$$G \simeq \frac{G_s}{(1 - 0.172)^2} = \frac{0.172}{(1 - 0.172)^2\sqrt{R_1R_2}}$$

$$= \frac{0.25}{\sqrt{R_1R_2}}$$

The approximate expression for the cavity gain for a TWA in the limiting case is $0.25/\sqrt{R_1R_2}$. Thus for wide spectral bandwidth operation the available cavity gain is determined by the quality of the antireflection coatings on the device.

10.3.2 Performance characteristics

The wide spectral bandwidths that may be achieved using high quality antireflection facet coatings on TWAs are in the region of 50 to 70 nm. However, in comparison with FPAs, such devices require significantly higher bias currents for operation as may be observed from Figure 10.4. In addition, whereas the narrow spectral bandwidth of FPAs provides inherent noise filtering, it is not obtained with TWAs and therefore they are subject to increased levels of noise.

The residual facet reflectivity in TWAs introduces a further problem when considering the use of such amplifiers within optical fiber communication systems. This problem results from the effect of backward gain within the devices. The gain of the backward travelling signal G_b is defined as the ratio of the power in the backward travelling signal P_b to the input signal power P_{in} into the amplifier. Hence the gain of the backward travelling signal is given by [Ref. 14]:

$$G_b = \frac{P_b}{P_{in}} = \frac{(\sqrt{R_1} - \sqrt{R_2}G_s)^2 + 4\sqrt{R_1R_2}G_s \sin^2 \phi}{(1 - \sqrt{R_1R_2}G_s)^2 + 4\sqrt{R_1R_2}G_s \sin^2 \phi} \qquad (10.13)$$

A graphical representation of Eq. (10.13) in which $R_1 = R_2$ for $G_s = 25$ dB is shown in Figure 10.6. It may be observed that the backward gain is approaching the potential forward gain at high facet reflectivity. Moreover, even at low facet reflectivity (0.01%) the backward gain is still very significant (10 dB). In systems with cascaded amplifiers, optical isolators may therefore be required to avoid the

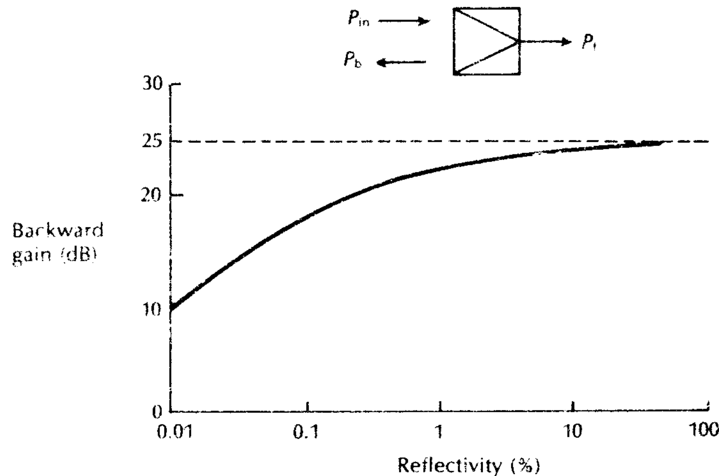


Figure 10.6 Backward gain against facet reflectivity for the travelling wave amplifier determined from Eq. (10.13). P_f is the power in the forward travelling signal.

interaction of backward signals between the devices, unless the backward wave amplitude can be made sufficiently small [Ref. 1].

Another very important characteristic of SLAs is noise since it largely determines the maximum number of devices which can be cascaded as linear repeaters within an optical fiber communication system [Ref. 9]. The overall noise generated by an SLA comprises signal spontaneous beat noise, spontaneous–spontaneous beat noise, spontaneous emission shot noise and amplified signal shot noise [Ref. 15]. The beat noise components occur between the signal and the spontaneously emitted photons as well as between the spontaneously emitted photons themselves. In addition the former noise component tends to dominate in the higher power operating region required for repeater applications [Ref. 9]. The theoretical treatment of the SLA noise characteristics is beyond the scope of this text but interested readers are directed to Ref. 15. It is clear, however, that unlike an electronic amplifier, the noise output from an SLA is a function of the signal intensity. The theoretical mean noise power as a function of facet reflectivity for an SLA, of length 500 μm with a fixed gain of 25 dB, is shown in Figure 10.7 [Ref. 1]. It may be observed that in the travelling wave region the noise power tends towards a fixed value of -3 dBm. This value is in good agreement with the results obtained from measurement [Ref. 1].

It was mentioned previously that the gain of the FPA was very sensitive to changes in temperature and signal polarization. A dependence on these two parameters is also observed in TWAs. For example, at an operating wavelength of 1.5 μm the gain decreases by around 3 dB when the temperature of a TWA of length 500 μm is increased by 5 °C [Ref. 1]. Moreover, although a decrease in temperature increases the device gain, it also increases the passband ripple when

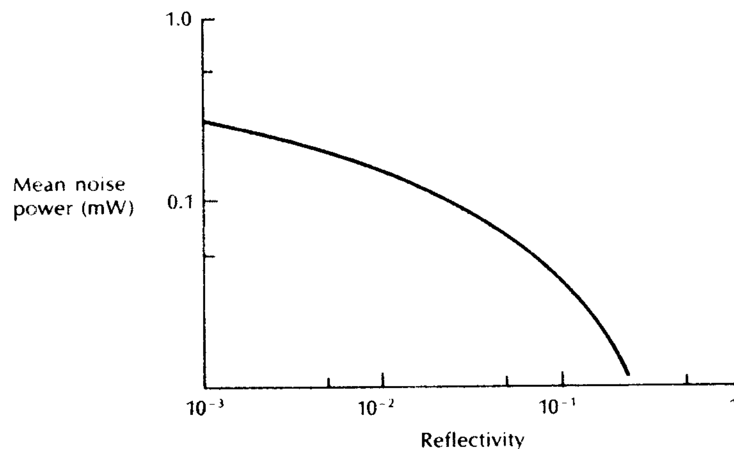


Figure 10.7 Theoretical mean noise power against facet reflectivity characteristic for a semiconductor laser amplifier of length 500 μm with a gain of 25 dB [Ref. 1].

there is residual reflectivity. With the FPA these effects are compounded by a shift in mode wavelength of approximately 10 GHz $^\circ C^{-1}$ caused by the variation in refractive index with temperature. Hence it is suggested that for high gain FPAs the temperature must be controlled to within 0.1 °C [Ref. 1].

The dependence of the gain on the polarization of the input signal results from the difference in the single pass gain for the TE and TM polarization modes. It is caused by a difference in their optical confinement factors (i.e. $\Gamma_{TE} \neq \Gamma_{TM}$). Furthermore, this effect is magnified in a resonant cavity because the mode propagation constants and the modal facet reflectivities are also polarization dependent. Hence the use of polarization controllers may be necessary when employing FPAs. The gain difference is minimized, however, with TWAs. For example, the gain difference for an FPA and a TWA both operating at a wavelength of 1.5 μm was found to be 10 dB and 2.5 dB respectively [Ref. 1].

10.4 Fiber amplifiers

System studies employing fiber based optical amplifiers have at present not progressed as far as those using semiconductor laser amplifiers [Ref. 2]. Nevertheless, investigations and demonstrations have indicated that such fiber devices offer significant potential within optical fiber communications [Refs. 17 to 19]. Hence it appears certain that fiber amplifiers will complement the growing device technology associated with SLAs. Although the various fiber amplifier types have significantly different performance characteristics, some of which do not match those obtained with SLAs, in all fiber amplifier devices the spectral bandwidths and centre

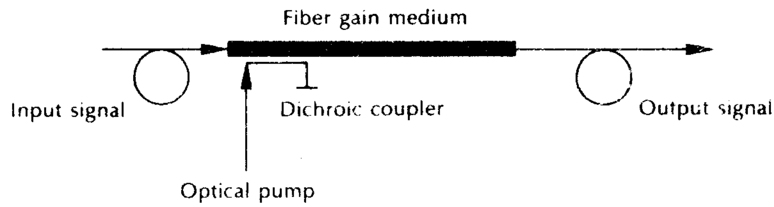


Figure 10.8 Schematic of a fiber amplifier.

wavelengths are largely defined by the atomic structure and not the mechanical geometry. Variations resulting from temperature changes, ageing and pump power are therefore less significant in fiber amplifiers than in SLAs.

A general representation of a fiber amplifier is shown in Figure 10.8. The gain medium normally comprises a length of single-mode fiber connected to a dichroic coupler (i.e. a wavelength division multiplexing coupler; see Section 5.6.3) which provides low insertion loss at both signal and pump wavelengths. Excitation occurs

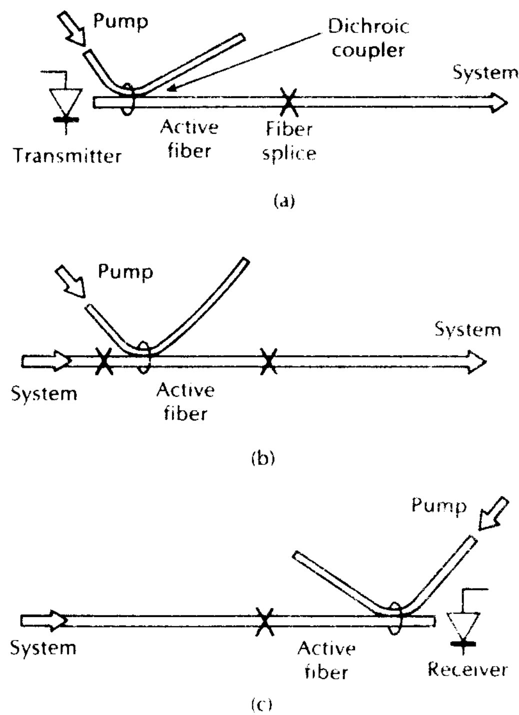


Figure 10.9 Some potential system applications for the fiber amplifier: (a) a power amplifier at the transmitter; (b) an optical repeater; (c) a preamplifier at the receiver

through optical pumping from a high power solid state or semiconductor laser which is combined with the optical input signal within the coupler. The amplified optical signal is therefore emitted from the other end of the gain medium.

The major options for implementing fiber amplifiers were mentioned in Section 10.2; namely, rare earth doped fiber amplifiers, Raman fiber amplifiers and Brillouin fiber amplifiers.* In particular, the former two device types, in common with travelling wave SLAs, are expected to be used across a broad range of system applications, some of which are illustrated in Figure 10.9. These include use as: a power amplifier at the transmitter, an in-line optical repeater amplifier and an optical preamplifier at the receiver.

10.4.1 Rare earth doped fiber amplifiers

Both neodymium and erbium doped fiber lasers were discused in Section 6.9.2. To date work on rare earth doped fiber amplifiers has concentrated on the erbium dopant, particularly in silica based single-mode fibers. High gains of between 30 and 40 dB with low noise have been demonstrated [Refs. 3, 20, 21] with optical pump powers in the range 50 to 100 mW. Such devices can be made to lase over the longer wavelength region (1.5 to 1.6 μm) of interest within optical fiber communications. Practical pump bands exist at wavelengths of 532 nm, 670 nm, 807 nm, 980 nm and 1480 nm [Refs. 2, 22]. However, the latter three wavelengths comprise the most important pump bands. The amplification is dependent on the material gain of a relatively short section (1 to 100 m) of the fiber. Aluminium codoping can be used to broaden the spectral bandwidth to around 40 nm (see Figure 10.1). It should be noted, however, that the spectral dependence on gain is not always as constant as that illustrated in Figure 10.1 (for example, see Figure 6.44) and hence the spectral bandwidth for erbium doped silica fibers may be restricted to around 300 GHz (2.4 nm) [Ref. 20].

A factor which limits the gain available from an erbium doped fiber amplifier is a phenomenon known as excited state absorption (ESA). This process is illustrated in the energy level diagrams for an erbium doped fiber system shown in Figure 10.10. Erbium provides a three level lasing scheme which is illustrated in Figure 10.10(a). However, in the erbium fiber amplifier photons at the pump wavelength tend to promote the electrons in the upper lasing level into a still higher state of excitation, as shown in Figure 10.10(b). These electrons then decay nonradiatively to intermediate levels, such as the pump bands, and then eventually back to the upper lasing level. Hence ESA reduces the pumping efficiency of the device and as a result it is necessary to pump at a higher power to obtain a specific gain.

The reduction of ESA in erbium doped fiber amplifiers is therefore being pursued [Refs. 19, 22]. This may be achieved by changing the location of the energy levels through codoping of the erbium–silica fiber amplifier with other compounds such as phosphorus pentoxide. Another technique is to pump the fiber amplifier at a

* In these devices the gain medium is often a standard single-mode fiber.

wavelength which does not cause the population of an excited state. Unfortunately, significant ESA is present at the favoured 807 nm pump band. Nevertheless, improved efficiency is obtained with the 980 and 1480 nm pump wavelengths. In particular the 980 nm wavelength displays high efficiency (twice the dB W^{-1} gain figure of the 1480 nm wavelength) but pump sources are not readily available, whereas operation at 1480 nm can be facilitated by both semiconductor and solid state laser sources.

An alternative solution to avoid ESA, however, is to change to another glass technology in place of silica. In this context the success that has been achieved in lasing with a fluorozirconate host glass (see Section 6.9.2) may provide a way forward. Signal amplification has already been obtained in an erbium doped

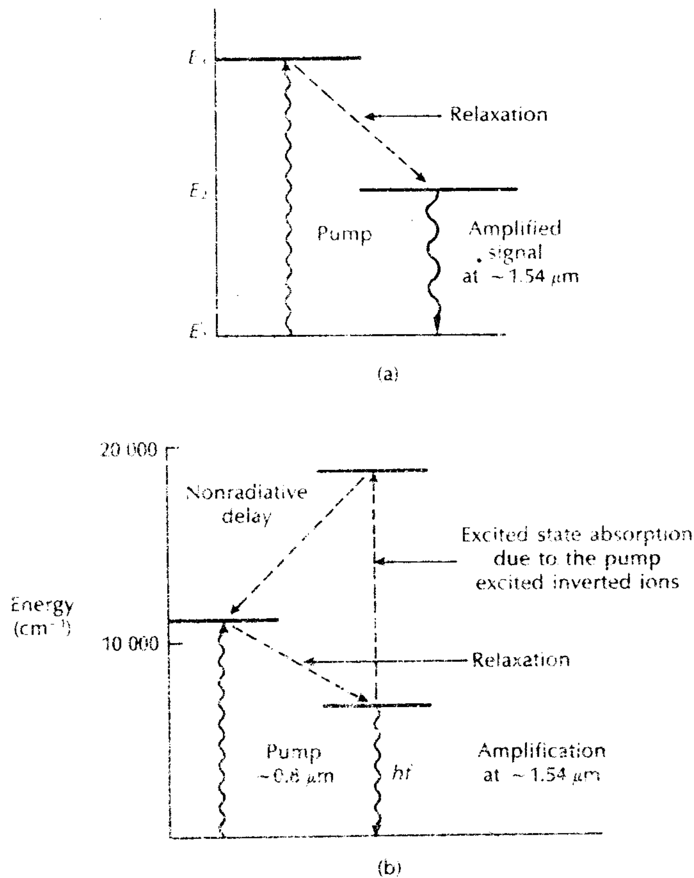


Figure 10.10 Energy level diagrams for erbium doped silica fiber laser: (a) the three level lasing scheme provided by Er^{3+} doping; (b) an illustration of excited state absorption, which is a major limitation.

multimode fluorozirconate fiber using a 488 nm pump wavelength to provide gain at a wavelength of 1.525 μm [Ref. 27].

10.4.2 Raman and Brillouin fiber amplifiers

Nonlinear effects within optical fiber may also be employed to provide optical amplification. Such amplification can be achieved by using stimulated Raman scattering, stimulated Brillouin scattering or stimulated four photon mixing, giving parametric gain (see Sections 3.5 and 3.14) by injecting a high power laser beam into undoped (or doped) optical fiber. Among these Raman amplification exhibits advantages of self phase matching between the pump and signal together with a broad gain–bandwidth or high speed response in comparison with the other nonlinear processes. In particular the broad gain–bandwidth associated with Raman amplification is attractive for application to possible future wavelength division multiplexed (WDM) systems [Ref. 18].

The pump signal optical wavelengths in Raman fiber amplifiers are typically 500 cm^{-1} higher in frequency than the signal to be amplified, and the pumping signal can propagate in either direction along the fiber. A schematic representation of both the forward and backward pumping capability of Raman fiber amplifiers is shown in Figure 10.11. Moreover, continuous-wave Raman gains exceeding 20 dB have been demonstrated experimentally in silica fiber [Ref. 24] which in principle exhibits a broad spectral bandwidth of up to 40 nm with suitable doping of the fiber [Ref. 2]. In addition, Raman gain in excess of 40 dB has been obtained using fluoride glass fiber in which the Raman shift is 590 cm^{-1} [Ref. 25]. More recently, Raman fiber amplifiers have been investigated for WDM system applications. For example, the simultaneous amplification using 60 mW pump power of three DFB

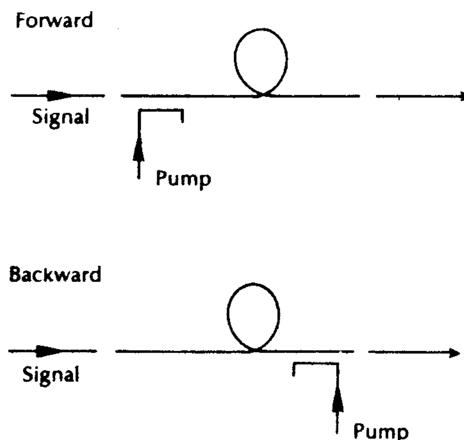


Figure 10.11 Illustrations of the forward and backward pumping capability associated with the fiber Raman amplifier.

laser diodes operating at wavelengths between 1570 nm and 1580 nm provided each channel with 5 dB gain [Ref. 26]. Furthermore, the gain–bandwidth of this fiber Raman amplifier was estimated to be in the range 20 to 30 nm.

The Raman gain G_R is dependent on a number of factors including the fiber length, the fiber attenuation and the fiber core diameter.* It may be expressed as a function of the optical pump power P_p as [Ref. 23]:

$$G_R = \exp\left(\frac{g_R P_p L_{eff}}{A_{eff} k}\right) \tag{10.14}$$

where g_R is the power Raman gain coefficient, and A_{eff} and L_{eff} are the effective fiber core area and length, respectively, and k is a numerical factor that accounts for polarization scrambling between the optical pump and signal [Ref. 28]. It should be noted that for complete polarization scrambling, as in conventional single-mode fiber, $k = 2$. The effective fiber core area and length are given by:

$$A_{eff} = \pi r_{eff}^2 \tag{10.15}$$

$$L_{eff} = \frac{1 - \exp(-\alpha_p L)}{\alpha_p} \tag{10.16}$$

where r_{eff} is the effective core radius, α_p is the fiber transmission loss at the pump wavelength and L is the actual fiber length.

The theoretical Raman gain characteristics as a function of fiber length for standard 10 μm core single-mode fibers with a pump input power of 1.6 W are shown in Figure 10.12 [Ref. 18]. It may be observed that the Raman gain becomes larger as the fiber lengths increase up to around 50 km where asymptotically it reaches a constant value. Moreover, it is clear that higher Raman gains can be obtained with lower loss fibers. Although not apparent from Figure 10.12, it is also the case that the Raman gain is increased as the fiber core diameter is decreased (see Eq. (10.14)). Nevertheless, in general the optical pump power required for Raman amplification tends to be high.

By contrast, stimulated Brillouin scattering is a very efficient nonlinear amplification mechanism that can provide high gains at modest optical pump powers of around 1 mW [Ref. 19]. However, it results from the scattering process in which the pump wavelength is often only around 20 GHz distance from the frequency of the optical signal to be amplified. Moreover, it is a narrow band process and the gain–bandwidth may only be in the range 15 to 20 MHz in silica fiber at a wavelength of 1.5 μm [Ref. 29]. The limitation on the spectral bandwidth in a pure silica fiber is around 50 MHz [Ref. 2] which fundamentally restricts the use of Brillouin amplifiers to relatively low speed communications. Although it is possible to extend the spectral bandwidth to 100 to 200 MHz with germanium doping of the fiber core, it does not significantly alleviate this problem.

* In standard single-mode fibers there are relatively low concentrations of germanium in the core which increases the peak Raman gain in comparison with pure silica core fiber.

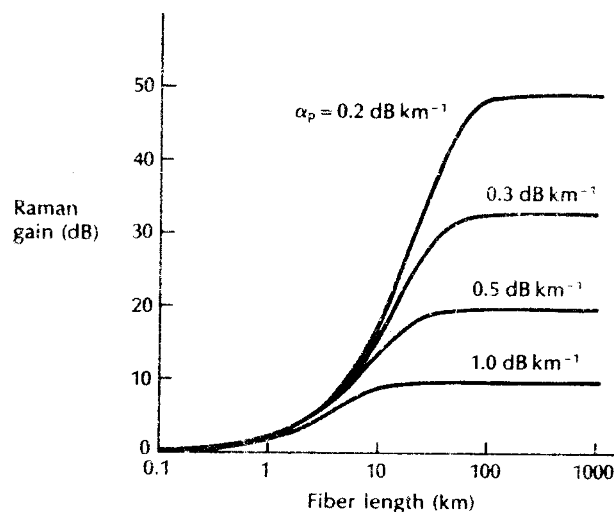


Figure 10.12 Raman gain dependence on fiber length and pump loss (α_p) for a pump input power of 1.6 W and a fiber core diameter of 10 μm. Reproduced with permission from Y. Aoki, 'Properties of fiber Raman amplifiers and their applicability to digital optical communication systems', *J. Lightwave Technol.*, 6, p. 1225, 1988. Copyright © 1988 IEEE.

Nevertheless, when the fiber is pumped with a CW laser at a power in the range 5 to 10 mW, gains in excess of 15 dB can be obtained [Ref. 2]. A very precise frequency difference of around 11 GHz, however, must be maintained between the optical pump and the signal to ensure that the Brillouin scattering phenomenon continues unabated. This fiber amplifier type is therefore perceived to have a rather restricted range of application. However, the narrowband process could be useful in the provision of tunable filters within WDM systems. It can provide channel selection by allowing amplification of a particular channel without boosting other nearby channels. For example, Brillouin amplification has been investigated for channel selection in densely packed, single-mode fiber systems [Ref. 30]. In this case data transmitted at a rate of 45 Mbit s^{-1} were detected without errors with an interfering channel spaced only 140 MHz away at an operating wavelength of 1.5 μm.

A possible limitation of Brillouin amplification, however, for bidirectional WDM applications results from crosstalk due to Brillouin gain if the frequency difference between the counterpropagating waves coincides with the Brillouin shift of around 20 GHz [Ref. 31]. The power level at which significant crosstalk can occur is only of the order of 100 μW [Ref. 32]. Fortunately, since the Brillouin gain–bandwidth is particularly narrow such crosstalk can generally be avoided by a correct choice of signal wavelengths, without restricting the channel packing density.

10.5 Integrated optics

The multitude of potential application areas for optical fiber communications coupled with the tremendous advances in the field have over recent years stimulated a resurgence of interest in the area of integrated optics (IO). The concept of IO involves the realization of optical and electro-optical elements which may be integrated in large numbers on to a single substrate. Hence, IO seeks to provide an alternative to the conversion of an optical signal back into the electrical regime prior to signal processing by allowing such processing to be performed on the optical signal. Thin transparent dielectric layers on planar substrates which act as optical waveguides are used in IO to produce miniature optical components and circuits.

The birth of IO may be traced back to basic ideas outlined by Anderson in 1966 [Ref. 33]. He suggested that a microfabrication technology could be developed for single-mode optical devices with semiconductor and dielectric materials in a similar manner to that which had taken place with electronic circuits. It was in 1969, however, after Miller [Ref. 34] had introduced the term 'integrated optics' whilst discussing the long term outlook in the area, that research began to gain momentum.

Developments in IO have now reached the stage where simple signal processing and logic junctions may be physically realized. Furthermore, such devices may form the building blocks for future digital optical computers. Nevertheless, at present, these advances are closely linked with developments in lightwave communication employing optical fibers.

A major factor in the development of integrated optics is that it is essentially based on single-mode optical waveguides and therefore tends to be incompatible with multimode fiber systems. Hence IO did not make a significant contribution to first and second generation optical fiber systems (see Section 14.1). The advent, however, of single-mode transmission technology has further stimulated work in IO in order to provide devices* and circuits for these more advanced third generation systems. It is apparent that the continued expansion of single-mode optical fiber communications will create a growing market for such IO components. Furthermore, it is predicted that the next generation of optical fiber communication systems employing coherent transmission will lean heavily on IO techniques for their implementation (see Chapter 12).

The proposals for IO devices and circuits which in many cases involve reinventions of electronic devices and circuits exhibits major advantages other than solely a compatibility with optical fiber communications. Electronic circuits have a practical limitation on speed of operation at a frequency of around 10^{10} Hz resulting from their use of metallic conductors to transport electronic charges and build-up signals. The large transmission bandwidths (over 1 GHz) currently under

* This is especially the case in relation to the fabrication of single-mode injection lasers (see Section 6.6).

investigation for optical fiber communications are already causing difficulties for electronic signal processing within the terminal equipment. The use of light with its property as an electromagnetic wave of extremely high frequency (10^{14} to 10^{15} Hz) offers the possibility of high speed operation around 10^4 times faster than that conceivable employing electronic circuits. Interaction of light with materials such as semiconductors or transparent dielectrics occurs at speeds in the range 10^{12} (pico) to approaching 10^{-15} (femto) seconds, thus providing a basis for subpicosecond optical switching.

The other major attribute provided by optical signals interacting within a responsive medium is the ability to utilize lightwaves of different frequencies (or wavelengths) within the same guided wave channel or device. Such frequency division multiplexing allows an information transfer capacity far superior to anything offered by electronics. Moreover, in signal processing terms it facilitates parallel access to information points within an optical system. This possibility for powerful parallel signal processing coupled with ultrahigh speed operation offers tremendous potential for applications within both communications and computing.

The devices of interest in IO are often the counterparts of microwave or bulk optical devices. These include junctions and directional couplers, switches and modulators, filters and wavelength multiplexers, lasers and amplifiers, detectors and bistable elements. It is envisaged that developments in this technology will provide the basis for the fourth generation systems mentioned in Section 14.1 where full monolithic integration may be achieved.

10.5.1 Planar waveguides

The use of circular dielectric waveguide structures for confining light is universally utilized within optical fiber communications. IO involves an extension of this guided wave optical technology through the use of planar optical waveguides to confine and guide the light in guided wave devices and circuits. The mechanism of optical confinement in symmetrical planar waveguides was discussed in Section 2.3 prior to investigation of circular structures. In fact the simplest dielectric waveguide structure is the planar slab guide shown in Figure 10.13. It comprises a planar film of refractive index n_1 sandwiched between a substrate of refractive index n_2 and a cover layer of refractive index n_3 where $n_1 > n_2 \geqslant n_3$. Often the cover layer consists of air where $n_3 = n_0 = 1$, and it exhibits a substantially lower refractive index than the other two layers. In this case the film has layers of different refractive index above and below the guiding layer and hence performs as an asymmetric waveguide.

In the discussions of optical waveguides given in Chapter 2 we were solely concerned with symmetrical structures. When the dimensions of the guide are reduced so are the number of propagating modes. Eventually the waveguide dimensions are such that only a single-mode propagates, and if the dimensions are reduced further this single-mode still continues to propagate. Hence there is no cutoff for the fundamental mode in a symmetric guide. This is not the case for an

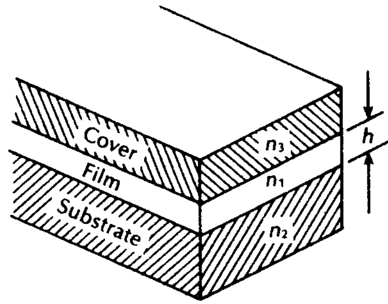


Figure 10.13 A planar slab waveguide. The film with high refractive index n_1 acts as the guiding layer and the cover layer is usually air where $n_3 = n_0 = 1$.

asymmetric guide where the dimensions may be reduced until the structure cannot support any modes and even the fundamental is cutoff. If the thickness or height of the guide layer of a planar asymmetric guide is h (see Figure 10.13), then the guide can support a mode of order m with a wavelength λ when [Ref. 35]:

$$h \geqslant \frac{(m + \frac{1}{2})\lambda}{2(n_1^2 - n_2^2)^{\frac{1}{2}}} \tag{10.17}$$

Equation (10.17) which assumes, $n_2 > n_3$ defines the limits of the single-mode region for h between values when $m = 0$ and $m = 1$. Hence for a typical thin film glass guide with $n_1 = 1.6$ and $n_2 = 1.5$, single-mode operation is maintained only when the guide has a thickness in the range $0.45\lambda \leqslant h \leqslant 1.35\lambda$.

An additional consideration of equal importance is the degree of confinement of the light to the guiding layer. The light is not exclusively confined to the guiding region and evanescent fields penetrate into the substrate and cover. An effective guide layer thickness h_{eff} may be expressed as:

$$h_{eff} = h + x_2 + x_3 \tag{10.18}$$

where x_2 and x_3 are the evanescent field penetration depths for the substrate and cover regions respectively. Furthermore, we can define a normalized effective thickness H for an asymmetric slab guide as:

$$H = kh_{eff}(n_1^2 - n_2^2)^{\frac{1}{2}} \tag{10.19}$$

where k is the free space propagation constant equal to $2\pi/\lambda$. The normalized frequency (sometimes called the normalized film thickness) for the planar slab guide following Eq. (2.68) is given by:

$$V = kh(n_1^2 - n_2^2)^{\frac{1}{2}} \tag{10.20}$$

An indication of the degree of confinement for the asymmetric slab waveguide may be observed by plotting the normalized effective thickness against the normalized frequencv for the TE modes. A series of such plots is shown in Figure

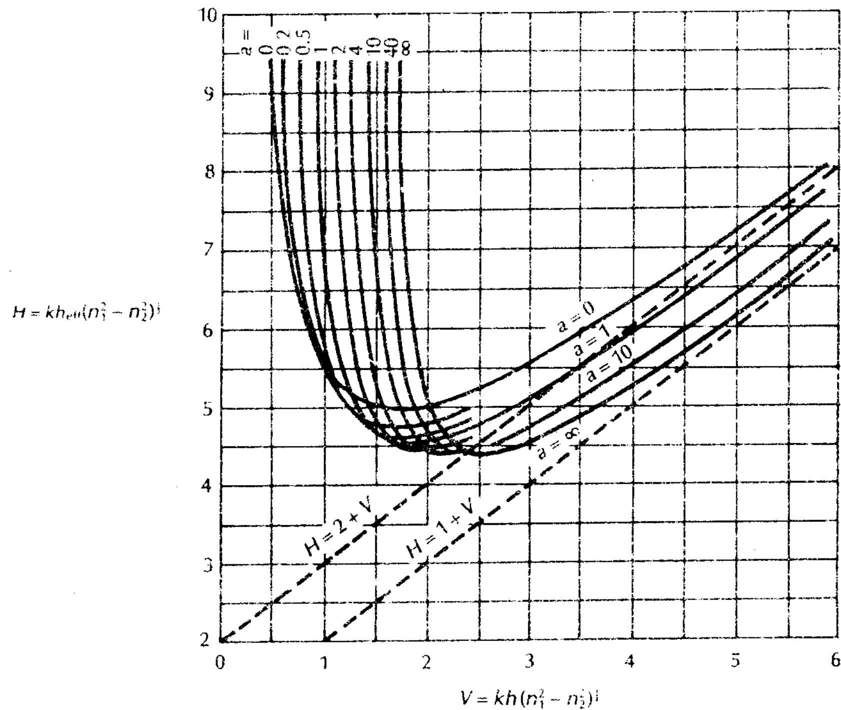


Figure 10.14 The normalized effective thickness H as a function of the normalized frequency V for a slab waveguide with various degrees of asymmetry. Reproduced with permission from H. Kogelnik and V. Ramaswamy, *Appl. Opt.*, **13**, p. 1857, 1974.

10.14 [Ref. 36] for various values of the parameter a which indicates the asymmetry of the guide, and is defined as:

$$a = \frac{n_2^2 - n_3^2}{n_1^2 - n_2^2} \tag{10.21}$$

It may be observed in Figure 10.14 that the confinement improves with decreasing film thickness only up to a point where $V = 2.5$. For example, the minimum effective thickness for a highly asymmetric guide ($a = \infty$) occurs when $H_{min} = 4.4$ at $V = 2.55$. Using Eq. (10.19) this gives a minimum effective thickness of:

$$(h_{eff})_{min} = \frac{4.4}{k}(n_1^2 - n_2^2)^{-\frac{1}{2}}$$

$$= 0.7\,\lambda(n_1^2 - n_2^2)^{-\frac{1}{2}} \tag{10.22}$$

Therefore considering a typical glass waveguide ($n_1 = 1.6$ and $n_2 = 1.5$), we obtain

a minimum effective thickness of:

$$(h_{\text{eff}})_{\min} \simeq 1.26\,\lambda \qquad (10.23)$$

Assuming a minimum operating wavelength to be 0.8 μm limits the effective thickness of the guide, and hence the confinement to around 1 μm. Therefore it appears there is a limit to possible fabrication with IO which is not present in other technologies* [Ref. 38]. At present there is still ample scope but confinement must be considered along with packing density and the avoidance of crosstalk.

The planar waveguides for IO may be fabricated from glasses and other isotropic materials such as silicon dioxide and polymers. Although these materials are used to produce the simplest integrated optical components, their properties cannot be controlled by external energy sources and hence they are of limited interest. In order to provide external control of the entrapped light to cause deflection, focusing, switching and modulation, active devices employing alternative materials must be utilized. A requirement for these materials is that they have the correct crystal symmetry to allow the local refractive index to be varied by the application of either electrical, magnetic or acoustic energy.†

To date, interest has centred on the exploitation of the electro-optic effect due to the ease of controlling electric fields through the use of electrodes together with the generally superior performance of electro-optic devices. Acousto-optic devices have, however, found a lesser role, primarily in the area of beam deflection. Magneto-optic devices [Ref. 39] utilizing the Faraday effect are not widely used, as in general, electric fields are easier to generate than magnetic fields.

A variety of electro-optic and acousto-optic materials have been employed in the fabrication of individual devices. Two basic groups can be distinguished by their refractive indices. These are materials with a refractive index near 2 ($LiNbO_3$, $LiTaO_3$, NbO_5, ZnS and ZnO) and materials with a refractive index greater than 3 (GaAs, InP and compounds of Ga and In with elements of Al, As and Sb).

Planar waveguide structures are produced using several different techniques which have in large part been derived from the microelectronics industry. For example, passive devices may be fabricated by radiofrequency sputtering to deposit thin films of glass onto glass substrates. Alternatively, active devices are often produced by titanium (Ti) diffusion into lithium niobate ($LiNbO_3$) or by ion implantation into gallium arsenide [Ref. 41].

The planar slab waveguide shown in Figure 10.13 confines light in only one direction, allowing it to spread across the guiding layer. In many instances it is useful to confine the light in two dimensions to a particular path on the surface of the substrate. This is achieved by defining the high index guiding region as a thin strip (strip guide) where total internal reflection will prevent the spread of the light beam across the substrate. In addition the strips can be curved or branched as

* The 1 μm barrier to confinement applies with all suitable waveguide materials. However, metal clad waveguides are not so limited but are plagued by high losses [Ref. 37].

† Using the electro-optic, magneto-optic or acousto-optic effects [Ref. 40].

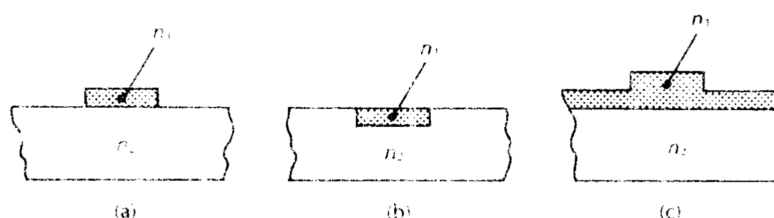


Figure 10.15 Cross section of some strip waveguide structures: (a) ridge guide; (b) diffused channel (embedded strip) guide; (c) rib guide.

required. Examples of such strip waveguide structures are shown in Figure 10.15. They may be formed as either a ridge on the surface of the substrate or by diffusion to provide a region of higher refractive index below the substrate, or as a rib of increased thickness within a thin planar slab. Techniques employed to obtain the strip pattern include electron and laser beam lithography as well as photolithography. The rectangular waveguide configurations illustrated in Figure 10.15 prove very suitable for use with electro-optic deflectors and modulators giving a reduction in the voltage required to achieve a particular field strength. In addition they allow a number of optical paths to be provided on a given substrate.

A trade-off also exists between the minimum radius of curvature which is required for high density integration and the ease of fabrication. It is clear from Eq. (10.22) that the waveguide dimensions are dependent upon the refractive index change. When the change is large, the dimensions of the waveguide may be reduced, even though the scattering losses become larger. As the maximum confinement of the single-mode guide occurs when it is operated near to the cutoff of the second order mode, then when the refractive index change is large, the radius of curvature of the waveguide can also be made very small. It is therefore necessary to find a compromise for the waveguide material used.

Titanium in-diffusion of $LiNbO_3$ gives rise to refractive index increases in the order of 0.01 to 0.02 which dictates a bend radius of the order of a few centimetres for negligible losses. It is, however, possible to use a proton exchange technique to increase the refractive index change up to 0.15 [Ref. 42]. By contrast, semiconductor III–V alloy waveguides based on compositional modification of the crystal give an index change of around 0.1 or more [Ref. 43]. Therefore, bend radii of the order of 1 mm or less may be obtained using these compounds. Moreover, although the effects of interest in IO are usually exhibited over short distances of around one wavelength, efficient devices require relatively long interaction lengths, the effects being cumulative. Hence, typical device lengths range from 0.5 to 10 mm.

Optical connections to and from waveguide devices are normally made by optical fibers. The overall insertion loss for such devices therefore comprises a waveguide-fiber coupling loss as well as the waveguide optical propagation loss. Careful

fabrication of Ti : $LiNbO_3$ waveguides with mode spot sizes well matched to that of typical single-mode fibers has yielded coupling losses in the range 0.5 to 1.0 dB per connection [Ref. 44]. In general, however, semiconductor waveguide devices exhibit larger fiber coupling losses because they operate with smaller spot sizes.

Propagation losses within both slab and strip waveguides are generally much greater than those obtained in single-mode optical fibers. However, more recently, propagation losses for Ti : $LiNbO_3$ waveguides have gone below 0.2 dB cm^{-1}, with excess bend losses being maintained below 0.1 dB per bend [Ref. 45]. By contrast propagation losses in semiconductor waveguides around 1 dB cm^{-1} are obtained when operating at wavelengths corresponding to the bandgap energy. Much lower losses of approximately 0.2 dB cm^{-1}, however, have to be achieved at operating wavelengths far below the bandgap energy [Ref. 46].

10.6 Some integrated optical devices

In this section some examples of various types of integrated optical devices together with their salient features are considered. However, the numerous developments in this field exclude any attempt to provide other than general examples in the major areas of investigation which are pertinent to optical fiber communications. The requirement for multichannel communication within the various systems considered in Chapters 11 and 12 demands the combination of information from separate channels, transmission of the combined signals over a single optical fiber link, and separation of the individual channels at the receiver prior to routeing to their individual destinations. Hence the application of IO in this area is to provide optical methods for multiplexing, modulation and routeing. These various functions may be performed with a combination of optical beam splitters, switches, modulators, filters, sources and detectors.

10.6.1 Beam splitters, directional couplers and switches

Beam splitters are a basic element of many optical fiber communication systems often providing a Y-junction by which signals from separate sources can be combined, or the received power divided between two or more channels. A passive Y-junction beam splitter fabricated from $LiNbO_3$ is shown in Figure 10.16. Unfortunately, the power transmission through such a splitter decreases sharply with increasing half angle γ, the power being radiated into the substrate. Hence the total power transmission depends critically upon γ which, for the example chosen, must not exceed 0.5° if an acceptable insertion loss is to be achieved [Ref. 47]. In order to provide effective separation of the output arms so that access to each is possible, the junction must be many times the width of the guide. For example, around 3000 wavelengths are required to give a separation of about 30 μm between the output arms. Therefore, for practical reasons, the device is relatively long.

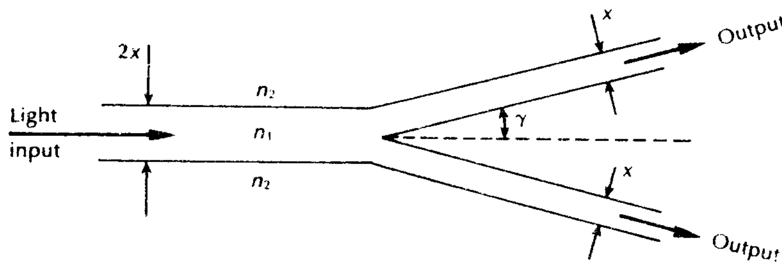


Figure 10.16 A passive Y-junction beam splitter.

The passive Y-junction beam splitter finds application where equal power division of the incident beam is required. However, the Y-junction is of wider interest when it is fabricated from an electro-optic material, in which case it may be used as a switch. Such materials exhibit a change in refractive index δn which is directly proportional to an applied electric field* E following,

$$\delta n = \pm \tfrac{1}{2} n_1^3 r E \qquad (10.24)$$

where n_1 is the original refractive index, and r is the electro-optic coefficient. Hence an active Y-junction may be fabricated from a single crystal electro-optic material as illustrated in Figure 10.17. Lithium niobate is often utilized as it combines relatively low loss with large values of electro-optic coefficients† (as high as 30.8×10^{-12} m V^{-1}). Metal electrodes are attached so that when biasing is applied, one side of the waveguide structure exhibits an increased refractive index whilst the value of refractive index on the other side is reduced. The light beam is therefore

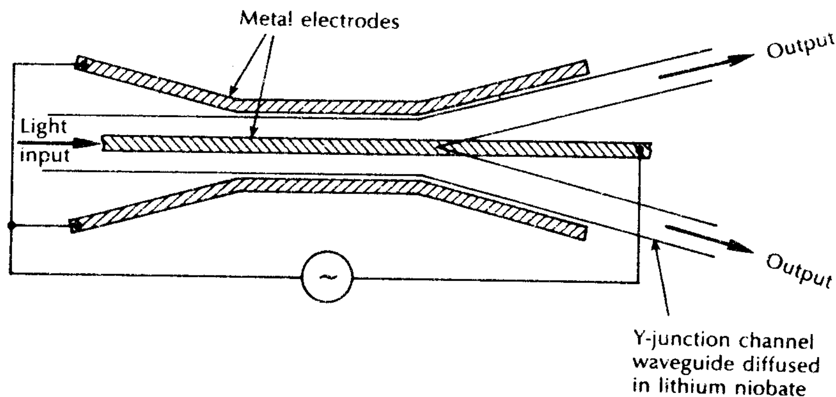


Figure 10.17 An electro-optic Y-junction switch.

* The linear variation of refractive index with the electric field is known as the Pockels effect [Ref. 40].
† The change in refractive index is related by the applied field via the linear and quadratic electro-optic coefficients [Ref. 39].

deflected towards the region of higher refractive index causing it to follow the corresponding output arm. Furthermore, the field is maintained in the electrodes which extend beyond the junction ensuring continuation of the process. With switching voltages around 30 V, these devices prove to be quite efficient allowing for larger junction angles to be tolerated than those of the passive Y-junction beam splitter. However, a physical length of several hundred wavelengths is still required for the switch. These devices therefore serve the function of optical signal routeing. In addition, high speed switches can be used to provide time division multiplexing of several lower bit rate channels onto a single-mode fiber link.

Switches may also be fabricated by placing two parallel strip waveguides in close proximity to each other as illustrated in Figure 10.18. The evanescent fields generated outside the guiding region allow transverse coupling between the guides. When the two waveguide modes have equal propagation constants β with amplitudes A and B (Figure 10.18), then the coupled mode equations may be written as [Ref. 48]:

$$\frac{\mathrm{d}A}{\mathrm{d}z} = \mathrm{j}\beta A + \mathrm{j}CB$$
$$\frac{\mathrm{d}B}{\mathrm{d}z} = \mathrm{j}\beta B + \mathrm{j}CA \tag{10.25}$$

where C is the coupling coefficient per unit length. In this case, assuming no losses all the energy from waveguide X will be transferred to waveguide Y over a coupling length l_0. Furthermore it can be shown [Ref. 49] that for this complete energy transfer l_0 is given by $\pi/2C$. If the waveguide modes have different propagation constants, however, only part of the energy from guide X will be coupled into guide Y, and this energy will be subsequently recoupled back into X.

It is also noted that when the propagation constants differ the coupling length l is reduced from the matched value l_0 and although less energy is transferred, the

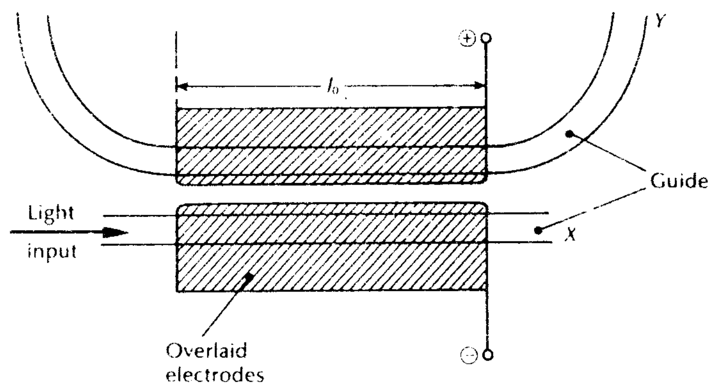


Figure 10.18 Electro-optically switched directional coupler. The COBRA configuration using two electrodes [Ref. 50].

exchange occurs more rapidly. This property may be utilized to good effect in the formation of an optical switch. The mismatch in propagation constants can be adjusted such that the coupling length l is reduced to $l_0/2$. In this case, energy coupled from one guide into the other over a distance $l_0/2$ will be recoupled into the original guide over a similar distance. Hence two distinct cases exist for a switch of length l_0, namely the matched case whereby all the energy is transferred from one guide to the other and the mismatched case when $l = l_0/2$ where over a distance l_0 the energy is recoupled into the original guide.

Optical switches of the above type use electrodes placed on the top of each matched waveguide (Figure 10.18) so that the refractive indices of the guides are differentially altered to produce the differing propagation constants for the mismatched case. A widely used switch utilizing this technique is called the COBRA (*Commutateur Optique Binaire Rapide*) [Ref. 50] and is normally formed from titanium diffused lithium niobate. Fabrication of the device, however, is critical in order to provide a coupling length which is exactly l_0 or an odd multiple of l_0. An electrode structure which avoids this problem by dividing the electrodes into halves with opposite polarities on each half is shown in Figure 10.19. With this device, which is called the stepped $\Delta\beta$ reversal coupler, it is always possible to obtain both the matched and mismatched cases described previously by applying suitable values of the reversed voltage. Hence the fabricated coupling length is no longer critical as the effective coupling length of the device may be adjusted electrically to achieve l_0.

The increasing deployment of optical fiber, particularly in the telecommunications network, has stimulated a great interest in optical or photonic switching in order to provide routeing in what is, at present, a circuit switched network [Refs. 51 to 56]. The technology discussed in Sections 10.6 to 10.8 provides the basic building blocks for such optical switching systems. Such switching systems can be classified in terms of their switching mechanism into space division switches, time division switches and wavelength or frequency* division switches [Ref. 51].

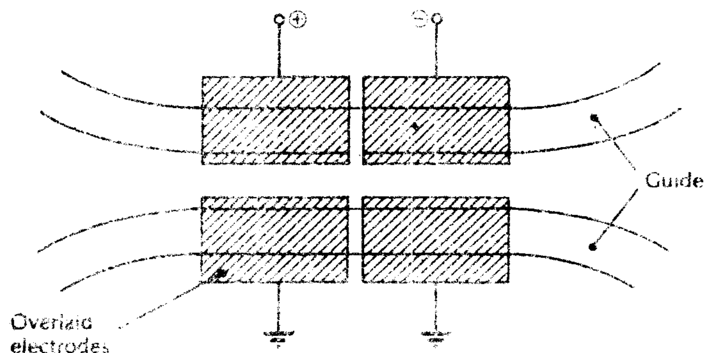


Figure 10.19 The stepped $\Delta\beta$ reversal coupler switch.

* In the optical domain these two terms are often used to indicate the same principle.

Although at a relatively early stage of development, optical switching matrices have been realized using IO technology. Optical space division switches incorporating electro-optically controlled directional couplers have been demonstrated. An example is illustrated in Figure 10.20 [Ref. 53]. The device is an 8 × 8 space switch comprising sixty-four directional couplers on a single lithium niobate substrate. It exhibited insertion losses in the range 6 to 8 dB and required a modulation voltage of 40 V. In addition a four channel time division switch using optical fiber delay lines combined with 4 × 4 lithium niobate optical switches has also been reported [Ref. 57]. Optical time division switching at 32 Mbit s^{-1} was obtained with this device.

Finally, an optical wavelength division switch for two channels using an acousto-optic deflector (see Section 10.6.2) with a photodiode and injection laser array has also been demonstrated [Ref. 58]. The WDM input signal is deflected by the acousto-optic deflector according to the frequency of an electrical signal applied to control transducer. Each electrical signal frequency corresponds to a deflection to

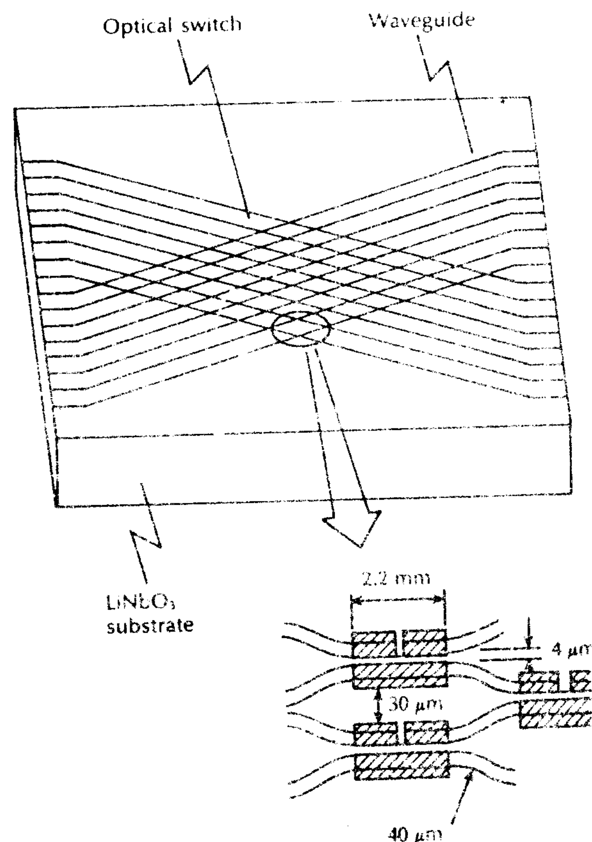


Figure 10.20 An 8 × 8 optical space division switch matrix.

a particular photodetector. Hence when the electrical signal frequency is altered, the optical signal with a wavelength corresponding to the desired frequency is deflected to the appropriate detector. Moreover, as each optical detector is connected to an individual injection laser emitting at a different wavelength to the one received, then the optical signal wavelength is converted to another wavelength. The device demonstrated the ability to wavelength switch two 400 Mbit s^{-1} optical signals.

10.6.2 Modulators

The limitations imposed by direct current modulation of semiconductor injection lasers currently restricts the maximum achievable modulation frequencies to a few gigahertz. Furthermore, with most injection lasers high speed current modulation also creates undesirable wavelength modulation which imposes problems for systems employing wavelength division multiplexing. Thus to extend the bandwidth capability of single-mode fiber systems there is a requirement for high speed modulation which can be provided by integrated optical waveguide intensity modulators. Simple on/off modulators may be based on the techniques utilized for the active beam splitters and switches described in Section 10.6.1. In addition a large variety of predominantly electro-optic modulators have been reported [Ref. 59] which exhibit good characteristics. For example, an important waveguide modulator is based upon a Y-branch interferometer which employs optical phase shifting produced by the electro-optic effect.

The change in refractive index exhibited by an electro-optic material with the application of an electric field given by Eq. (10.24) also provides a phase change for light propagating in the material. This phase change $\delta\phi$ is accumulative over a distance L within the material and is given by [Ref. 60]:

$$\delta\phi = \frac{2\pi}{\lambda}\,\delta n L \tag{10.26}$$

When the electric field is applied transversely to the direction of optical propagation we may substitute for δn from Eq. (10.24) giving:

$$\delta\phi = \frac{\pi}{\lambda}\, n_1^3 r E L \tag{10.27}$$

Furthermore taking E equal to V/d, where V is the applied voltage and d is the distance between electrodes gives:

$$\delta\phi = \frac{\pi}{\lambda}\, n_1^3 r \frac{VL}{d} \tag{10.28}$$

It may be noted from Eq. (10.28) that in order to reduce the applied voltage V required to provide a particular phase change, the ratio L/d must be made as large as possible.