11 Detection of Optical Radiation

11.0 INTRODUCTION

The detection of optical radiation is often accomplished by converting the radiant energy into an electric signal whose intensity is measured by conventional techniques. Some of the physical mechanisms that may be involved in this conversion include

1. The generation of mobile charge carriers in solid-state photoconductive detectors
2. Changing through absorption the temperature of thermocouples, thus causing a change in the junction voltage
3. The release by the photoelectric effect of free electrons from photoemissive surfaces

In this chapter we consider in some detail the operation of four of the most important detectors:

1. The photomultiplier
2. The photoconductive detector
3. The photodiode
4. The avalanche photodiode

The limiting sensitivity of each is discussed and compared to the theoretical limit. We will find that by use of the heterodyne mode of detection the theoretical limit of sensitivity may be approached.

11.1 OPTICALLY INDUCED TRANSITION RATES

A common feature of all the optical detection schemes discussed in this chapter is that the electric signal is proportional to the rate at which electrons are excited by the optical field. This excitation involves a transition of the electron from some initial bound state, say a, to a final state (or a group of states) b in which it is free to move and contribute to the current flow. For example, in an n-type photoconductive detector, state a corresponds to electrons in the filled valence band or localized donor impurity atoms, while state b corresponds to electrons in the conduction band. The two levels involved are shown schematically in Figure 11-1. A photon of energy $h\nu$ is absorbed in the process of exciting an electron from a "bound" state a to a "free" state b in which the electron can contribute to the current flow.

An important point to understand before proceeding with the analysis of different detection schemes is the manner of relating the transition rate per electron from state a to b to the intensity of the optical field. This rate is derived by quantum mechanical considerations.[1] In our case it can be stated in the following form: Given a nearly sinusoidal optical field[2]

$$e(t) = \tfrac{1}{2}[E(t)e^{i\omega_0 t} + E^*(t)e^{-i\omega_0 t}] \equiv \mathrm{Re}[V(t)] \tag{11.1-1}$$

where $V(t) = E(t)\exp(i\omega_0 t)$,[3] the transition rate per electron induced by this field is proportional to $V(t)V^*(t)$. Denoting the transition rate as $W_{a\to b}$, we have

$$W_{a\to b} \propto V(t)V^*(t) \tag{11.1-2}$$

---

[1] More specifically, from first order time-dependent perturbation theory; see, for example, Reference [1].

[2] By "nearly sinusoidal" we mean a field where $E(t)$ varies slowly compared to $\exp(i\omega_0 t)$ or, equivalently, where the Fourier spectrum of $E(t)$ occupies a bandwidth that is small compared to ω_0. Under these conditions the variation of the amplitude $E(t)$ during a few optical periods can be neglected.

[3] $V(t)$ is referred to as the "analytic signal" of $e(t)$. See Problem 1.1.

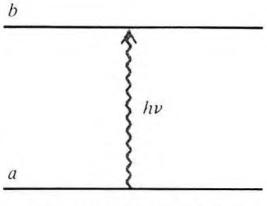


Figure 11-1 Most high-speed optical detectors depend on absorption of photons of energy $h\nu$ accompanied by a simultaneous transition of an electron (or hole) from a quantum state of low mobility (a) to one of higher mobility (b).

We can easily show that $V(t)V^*(t)$ is equal to twice the average value of $e^2(t)$, where the averaging is performed over a few optical periods.

To illustrate the power of this seemingly simple result, consider the problem of determining the transition rate due to a field

$$e(t) = E_0 \cos(\omega_0 t + \phi_0) + E_1 \cos(\omega_1 t + \phi_1) \tag{11.1-3}$$

taking $\omega_1 - \omega_0 \equiv \omega \ll \omega_0$. We can rewrite (11.1-3) as

$$\begin{aligned} e(t) &= \mathrm{Re}(E_0\, e^{i(\omega_0 t+\phi_0)} + E_1\, e^{i(\omega_1 t+\phi_1)}) \\ &= \mathrm{Re}[(E_0\, e^{i\phi_0} + E_1\, e^{i(\omega t+\phi_1)})e^{i\omega_0 t}] \end{aligned} \tag{11.1-4}$$

and, using (11.1-1), identify $V(t)$ as

$$V(t) = [E_0\, e^{i\phi_0} + E_1\, e^{i(\omega t+\phi_1)}]\, e^{i\omega_0 t}$$

thus, using (11.1-2), we obtain

$$\begin{aligned} W_{a\to b} &\propto (E_0\, e^{i\phi_0} + E_1\, e^{i(\omega t+\phi_1)})(E_0\, e^{-i\phi_0} + E_1\, e^{-i(\omega t+\phi_1)}) \\ &= E_0^2 + E_1^2 + 2E_0E_1 \cos(\omega t + \phi_1 - \phi_0) \end{aligned} \tag{11.1-5}$$

This shows that the transition rate has, in addition to a constant term $E_0^2 + E_1^2$, a component oscillating at the difference frequency ω with a phase equal to the difference of the two original phases. This coherent "beating" effect forms the basis of the heterodyne detection scheme, which is discussed in detail in Section 11.4.

11.2 PHOTOMULTIPLIER

The photomultiplier, one of the most common optical detectors, is used to measure radiation in the near ultraviolet, visible, and near infrared regions of the spectrum. Because of its inherent high current amplification and low noise, the photomultiplier is one of the most sensitive instruments devised by man and under optimal operation—which involves long integration time, cooling of the photocathode, and pulse-height discrimination—has been used to detect power levels as low as about 10^{-19} watt [2].

A schematic diagram of a conventional photomultiplier is shown in Figure 11-2. It consists of a photocathode (C) and a series of electrodes, called dynodes, that are labeled 1 through 8. The dynodes are kept at progressively higher potentials with respect to the cathode, with a typical potential difference between adjacent dynodes of 100 volts. The last electrode (A), the anode, is used to collect the electrons. The whole assembly is contained within a vacuum envelope in order to reduce the possibility of electronic collisions with gas molecules.

The photocathode is the most crucial part of the photomultiplier, since it converts the incident optical radiation to electronic current and thus determines the wavelength-response characteristics of the detector and, as will

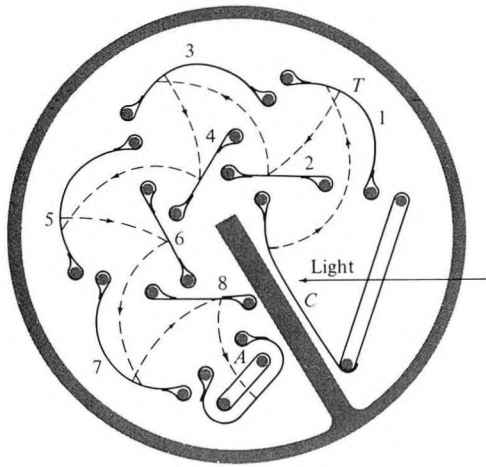


Figure 11-2 Photocathode and focusing dynode configuration of a typical commercial photomultiplier. C = cathode; 1–8 = secondary-emission dynodes; A = collecting anode. (After Reference [3].)

be seen, its limiting sensitivity. The photocathode consists of materials with low surface work functions. Compounds involving Ag-O-Cs and Sb-Cs are often used; see References [2, 3]. These compounds possess work functions as low as 1.5 eV, as compared to 4.5 eV in typical metals. As can be seen in Figure 11-3, this makes it possible to detect photons with longer wavelengths. It follows from the figure that the low-frequency detection limit corresponds to $h\nu = \phi$. At present the lowest-work-function materials make possible photoemission at wavelengths as long as 1–1.1 μm.

Spectral response curves of a number of commercial photocathodes are shown in Figure 11-4. The quantum efficiency (or quantum yield as it is often called) is defined as the number of electrons released per incident photon.

The electrons that are emitted from the photocathode are focused electrostatically and accelerated toward the first dynode, arriving with a kinetic

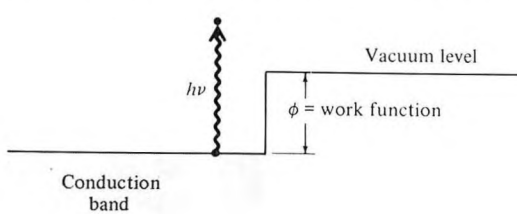


Figure 11-3 Photomultiplier photocathode. The vacuum level corresponds to the energy of an electron at rest an infinite distance from the cathode. The work function ϕ is the minimum energy required to lift an electron from the metal into the vacuum level, so only photons with $h\nu > \phi$ can be detected.

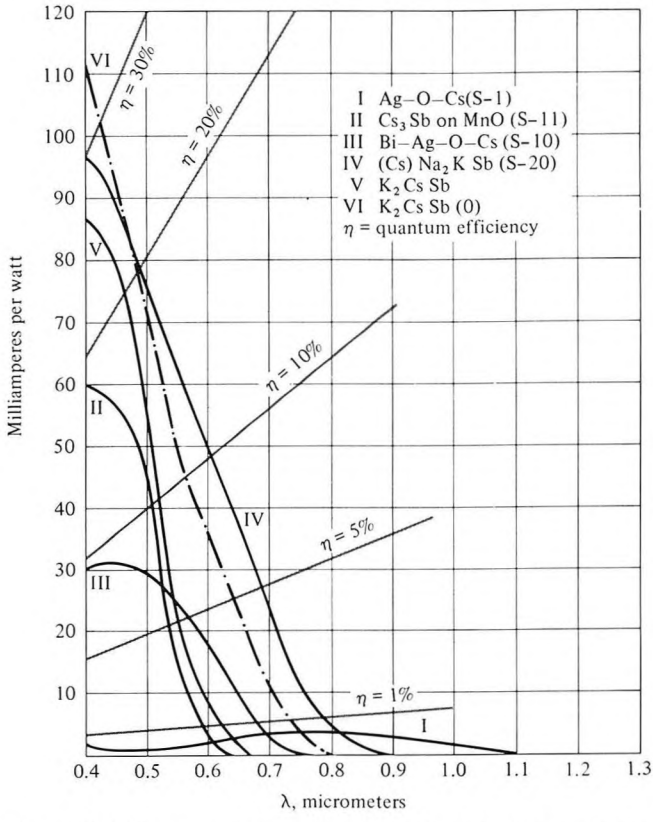


Figure 11-4 Photoresponse versus wavelength characteristics and quantum efficiency of a number of commercial photocathodes. (After Reference [3], p. 228.)

energy of, typically, about 100 eV. Secondary emission from dynode surfaces causes a multiplication of the initial current. This process repeats itself at each dynode until the initial current emitted by the photocathode is amplified by a very large factor. If the average secondary emission multiplication at each dynode is δ (that is, δ secondary electrons for each incident one) and the number of dynodes is N, the total current multiplication between the cathode and anode is

$$G = \delta^N$$

which, for typical values[4] of $\delta = 5$ and $N = 9$, gives $G \simeq 2 \times 10^6$.

[4]The value of δ depends on the voltage V between dynodes, and values of $\delta \simeq 10$ can be obtained (for $V \simeq 400$ volts). In commercial tubes, values of $\delta \simeq 5$, achievable with $V \simeq 100$ volts, are commonly used.

11.3 NOISE MECHANISMS IN PHOTOMULTIPLIERS

The random fluctuations observed in the photomultiplier output are due to

1. Cathode shot noise, given according to (10.4-9) by

$$\overline{(i_{N_1}^2)} = G^2 2\, e(\bar{i}_c + i_d)\Delta\nu \tag{11.3-1}$$

where $\bar{i}_c$ is the average current emitted by the photocathode due to the signal power that is incident on it. The current i_d is the so-called dark current, which is due to random thermal excitation of electrons from the surface as well as to excitation by cosmic rays and radioactive bombardment.

2. Dynode shot noise, which is the shot noise due to the random nature of the secondary emission process at the dynodes. Since current originating at a dynode does not exercise the full gain of the tube, the contribution of all the dynodes to the total shot noise output is smaller by a factor of $\sim\delta^{-1}$ than that of the cathode; since $\delta \simeq 5$ it amounts to a small correction and will be ignored in the following.
3. Johnson noise, which is the thermal noise associated with the output resistance R connected across the anode. Its magnitude is given by (10.5-9) as

$$\overline{(i_{N_2}^2)} = \frac{4kT\Delta\nu}{R} \tag{11.3-2}$$

Minimum Detectable Power in Photomultipliers—Video Detection

Photomultipliers are used primarily in one of two ways. In the first, the optical wave to be detected is modulated at some low frequency ω_m before impinging on the photocathode. The signal consists then, of an output current oscillating at ω_m, which, as will be shown below, has an amplitude proportional to the optical intensity. This mode of operation is known as *video*, or straight, detection.

In the second mode of operation, the signal to be detected, whose optical frequency is ω_s, is combined at the photocathode with a much stronger optical wave of frequency $\omega_s + \omega$. The output signal is then a current at the offset frequency ω. This scheme, known as *heterodyne* detection, will be considered in detail in Section 11-4.

The optical signal in the case of video detection may be taken as

$$\begin{aligned} e_s(t) &= E_s(1 + m\cos\omega_m t)\cos\omega_s t \\ &= \mathrm{Re}[E_s(1 + m\cos\omega_m t)e^{i\omega_s t}] \end{aligned} \tag{11.3-3}$$

where the factor $(1 + m\cos\omega_m t)$ represents amplitude modulation of the

carrier.[5] The photocathode current is given, according to (11.1-2), by

$$i_c(t) \propto [E_s(1 + m \cos \omega_m t)]^2$$
$$= E_s^2 \left[\left(1 + \frac{m^2}{2}\right) + 2m \cos \omega_m t + \frac{m^2}{2} \cos 2\omega_m t \right] \quad (11.3\text{-}4)$$

To determine the proportionality constant involved in (11.3-4), consider the case of $m = 0$. The average photocathode current due to the signal is then[6]

$$\bar{i}_c = \frac{P e\eta}{h\nu_s} \quad (11.3\text{-}5)$$

where $\nu_s = \omega_s/2\pi$, P is the average optical power, and η (the quantum efficiency) is the average number of electrons emitted from the photocathode per incident photon. This number depends on the photon frequency, the photocathode surface, and in practice (see Figure 11-4) is found to approach 0.3. Using (11.3-5), we rewrite (11.3-4) as

$$i_c(t) = \frac{P e\eta}{h\nu_s} \left[\left(1 + \frac{m^2}{2}\right) + 2m \cos \omega_m t + \frac{m^2}{2} \cos 2\omega_m t \right] \quad (11.3\text{-}6)$$

The signal output current at ω_m is

$$i_s = \frac{GP e\eta}{h\nu_s} (2m) \cos \omega_m t \quad (11.3\text{-}7)$$

If the output of the detector is limited by filtering to a bandwidth $\Delta\nu$ centered on ω_m, it contains a shot-noise current, which, according to (11.3-1), has a mean-squared amplitude

$$\overline{(i_{N_1}^2)} = 2G^2 e(\bar{i}_c + i_d)\, \Delta\nu \quad (11.3\text{-}8)$$

where $\bar{i}_c$ is the average signal current and i_d is the dark current.

The noise and signal equivalent circuit is shown in Figure 11-5, where for the sake of definiteness we took the modulation index $m = 1$. R represents the output load of the photomultiplier. T_e is chosen so that the term $4kT_e\Delta\nu/R$ accounts for the thermal noise of R as well as for the noise generated by the amplifier that follows the photomultiplier.

The signal-to-noise power ratio at the output is thus

$$\frac{S}{N} = \frac{\overline{i_s^2}}{\overline{(i_{N_1}^2)} + \overline{(i_{N_2}^2)}}$$
$$= \frac{2(Pe\eta/h\nu_s)^2 G^2}{2G^2 e(\bar{i}_c + i_d)\Delta\nu + (4kT_e\Delta\nu/R)} \quad (11.3\text{-}9)$$

---

[5]The amplitude modulation can be due to the information carried by the optical wave or, as an example, to chopping before detection.

[6]$P/h\nu_s$ is the rate of photon incidence on the photocathode; thus, if it takes $1/\eta$ photons to generate one electron, the average current is given by (11.3-5).

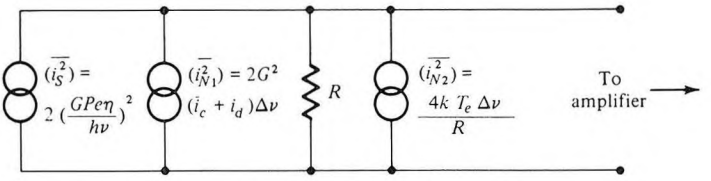


Figure 11-5 Equivalent circuit of a photomultiplier.

Due to the large current gain ($G \simeq 10^6$), the first term in the denominator of (11.3-9), which represents amplified cathode shot noise, is much larger than the thermal and amplifier noise term $4kT_e\Delta\nu/R$. Neglecting the term $4kT_e\Delta\nu/R$, assuming $i_d \gg \overline{i_c}$, and setting $S/N = 1$, we can solve for the minimum detectable optical power.

$$P_{\min} = \frac{h\nu_s(i_d\Delta\nu)^{1/2}}{\eta e^{1/2}} \tag{11.3-10}$$

Numerical Example: Sensitivity of Photomultiplier

Consider a typical case of detecting an optical signal under the following conditions:

$\nu_s = 6 \times 10^{14}$ Hz ($\lambda = 0.5\ \mu$m)

$\eta = 10$ percent

$\Delta\nu = 1$ Hz

$i_d = 10^{-15}$ ampere (a typical value of the dark photocathode current)

Substitution in (11.3-10) gives

$$P_{\min} = 3 \times 10^{-16} \text{ watt}$$

The corresponding cathode signal current is $\overline{i_c} \sim 10^{-17}$ ampere, so the assumption $i_d \gg \overline{i_c}$ is justified.

Signal-Limited Shot Noise

If one could, somehow, eliminate the Johnson noise and the dark current altogether, so that the only contribution to the average photocathode current is $\overline{i_c}$, which is due to the optical signal, then, using (11.3-5) and (11.3-9) to solve self-consistently for $P_{\min}$,

$$P_{\min} \simeq \frac{h\nu_s\Delta\nu}{\eta} \tag{11.3-11}$$

This corresponds to the quantum limit of optical detection. Its significance will be discussed in the next section. The practical achievement of this limit in video detection is nearly impossible since it depends on near total suppression of the dark current and other extraneous noise sources such as background radiation reaching the photocathode and causing shot noise.

The quantum detection limit (11.3-11) can, however, be achieved in the heterodyne mode of optical detection. This is discussed in the next section.

11.4 HETERODYNE DETECTION WITH PHOTOMULTIPLIERS

In the heterodyne mode of optical detection, the signal to be detected $E_s \cos \omega_s t$ is combined with a second optical field, referred to as the local-oscillator field, $E_L \cos(\omega_s + \omega)t$, shifted in frequency by $\omega(\omega \ll \omega_s)$. The total field incident on the photocathode is therefore given by

$$e(t) = \mathrm{Re}[E_L\, e^{i(\omega_s+\omega)t} + E_s\, e^{i\omega_s t}] \equiv \mathrm{Re}[V(t)] \tag{11.4-1}$$

The local-oscillator field originates usually at a laser at the receiving end, so that it can be made very large compared to the signal to be detected. In the following we will assume that

$$E_L \gg E_s \tag{11.4-2}$$

A schematic diagram of a heterodyne detection scheme is shown in Figure 11-6. The current emitted by the photocathode is given, according to (11.1-2) and (11.4-1), by

$$i_c(t) \propto V(t)V^*(t) = E_L^2 + E_s^2 + 2E_L E_s \cos \omega t$$

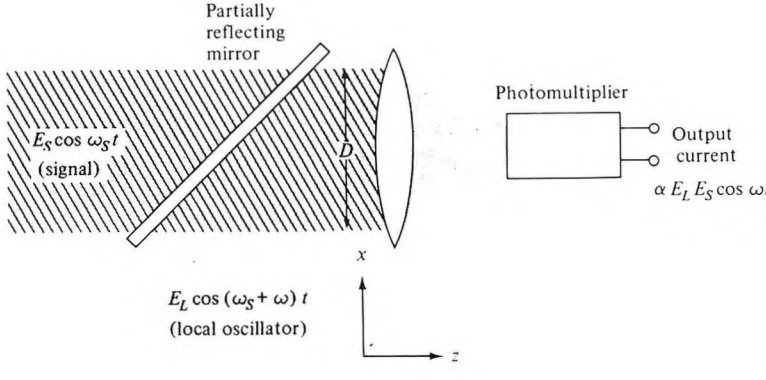


Figure 11-6 Schematic diagram of a heterodyne detector using a photomultiplier.

which, using (11.4-2) can be written as

$$i_c(t) = aE_L^2\left(1 + \frac{2E_s}{E_L}\cos\omega t\right) = aE_L^2\left(1 + 2\sqrt{\frac{P_s}{P_L}}\cos\omega t\right) \qquad (11.4\text{-}3)$$

where P_s and P_L are the signal and local-oscillator powers, respectively. The proportionality constant a in (11.4-3) can be determined as in (11.3-6) by requiring that when $E_s = 0$ the direct current be related to the local-oscillator power P_L by $\overline{i_c} = P_L\eta e/h\nu_L$,[7] so taking $\nu \approx \nu_L$

$$i_c(t) = \frac{P_L e\eta}{h\nu}\left(1 + 2\sqrt{\frac{P_s}{P_L}}\cos\omega t\right) \qquad (11.4\text{-}4)$$

The total cathode shot noise is thus

$$\overline{(i_{N_1}^2)} = 2e\left(i_d + \frac{P_L e\eta}{h\nu}\right)\Delta\nu \qquad (11.4\text{-}5)$$

where i_d is the average dark current while $P_L e\eta/h\nu$ is the dc cathode current due to the strong local-oscillator field. The shot-noise current is amplified by G, resulting in an output noise

$$\overline{(i_N^2)}_{\text{anode}} = G^2 2e\left(i_d + \frac{P_L e\eta}{h\nu}\right)\Delta\nu \qquad (11.4\text{-}6)$$

The mean-square signal current at the output is, according to (11.4-4),

$$\overline{(i_s^2)}_{\text{anode}} = 2G^2\left(\frac{P_s}{P_L}\right)\left(\frac{P_L e\eta}{h\nu}\right)^2 \qquad (11.4\text{-}7)$$

The signal-to-noise power ratio at the output is given by

$$\frac{S}{N} = \frac{2G^2(P_s P_L)(e\eta/h\nu)^2}{[G^2 2e(i_d + P_L e\eta/h\nu) + 4kT_e/R]\Delta\nu} \qquad (11.4\text{-}8)$$

where, as in (11.3-9), the last term in the denominator represents the Johnson (thermal) noise generated in the output load, plus the effective input noise of the amplifier following the photomultiplier. The big advantage of the heterodyne detection scheme is now apparent. By increasing P_L the S/N ratio increases until the denominator is dominated by the term $G^2 2eP_L e\eta/h\nu$. This corresponds to the point at which the *shot noise produced by the local oscillator current dwarfs all the other noise contributions*. When this state of affairs prevails, we have, according to (11.4-8),

$$\frac{S}{N} \simeq \frac{P_s}{h\nu\,\Delta\nu/\eta} \qquad (11.4\text{-}9)$$

---

[7]This is just a statement of the fact that each incident photon has a probability η of releasing an electron.

which corresponds to the quantum-limited detection limit. The minimum detectable signal—that is, the signal input power leading to an output signal-to-noise ratio of 1—is thus

$$(P_s)_{\min} = \frac{h\nu\,\Delta\nu}{\eta} \tag{11.4-10}$$

This power corresponds for $\eta = 1$ to a flux at a rate of one photon per $(\Delta\nu)^{-1}$ seconds—that is, one photon per resolution time of the system.[8]

Numerical Example: Minimum Detectable Power with a Heterodyne System

It is interesting to compare the minimum detectable power for the heterodyne system as given by (11.4-10) with that calculated in the example of Section 11.3 for the video system. Using the same data,

$$\nu = 6 \times 10^{14}\ \text{Hz}(\lambda = 0.5\ \mu\text{m})$$

$$\eta = 10\ \text{percent}$$

$$\Delta\nu = 1\ \text{Hz}$$

we obtain

$$(P_s)_{\min} \simeq 4 \times 10^{-18}\ \text{watt}$$

to be compared with $P_{\min} \simeq 3 \times 10^{-16}$ watt in the video case.

Limiting Sensitivity as a Result of the Particle Nature of Light

The quantum limit to optical detection sensitivity is given by (11.4-10) as

$$(P_s)_{\min} = \frac{h\nu\,\Delta\nu}{\eta} \tag{11.4-11}$$

This limit was shown to be due to the shot noise of the photoemitted current. We may alternatively attribute this noise to the granularity—that is, the particle nature—of light, according to which the minimum energy increment of an electromagnetic wave at frequency ν is $h\nu$. The power average P of an optical wave can be written as

$$P = \bar{N}h\nu \tag{11.4-12}$$

where $\bar{N}$ is the average number of photons arriving at the photocathode per second. Next assume a hypothetical noiseless photomultiplier in which *ex-*

[8] A detection system that is limited in bandwidth to $\Delta\nu$ cannot resolve events in time that are separated by less than $\sim(\Delta\nu)^{-1}$ second. Thus $(\Delta\nu)^{-1}$ is the resolution time of the system.

actly one electron is produced for each η^{-1} incident photon. The measurement of P is performed by counting the number of electrons produced during an observation period T and then averaging the result over a large number of similar observations.

The average number of electrons emitted per observation period T is

$$\bar{N}_e = \bar{N}T\eta \tag{11.4-13}$$

If the photons arrive in a perfectly random manner, then the number of photons arriving during the fixed observation period obeys Poissonian statistics[9]. Since in our ideal example, the electrons that are emitted mimic the arriving photons, they obey the same statistical distribution law. This leads to a fluctuation

$$\overline{(\Delta N_e)^2} \equiv \overline{(N_e - \bar{N}_e)^2} = \bar{N}_e = \bar{N}T\eta$$

Defining the minimum detectable number of quanta as that for which the rms fluctuation in the number of emitted photoelectrons equals the average value, we get

$$(\bar{N}_{\min} T\eta)^{1/2} = \bar{N}_{\min} T\eta$$

or

$$(\bar{N})_{\min} = \frac{1}{T\eta} \tag{11.4-14}$$

If we convert the last result to power by multiplying it by $h\nu$ and recall that $T^{-1} \simeq \Delta\nu$, where $\Delta\nu$ is the bandwidth of the system, we get

$$(P_s)_{\min} = \frac{h\nu\,\Delta\nu}{\eta} \tag{11.4-15}$$

in agreement with (11.4-10).

The above discussion points to the fact that the noise (fluctuation) in the photo current can be blamed on the physical process that introduces the

---

[9]This follows from the assumption that the photon arrival is perfectly random, so the probability of having N photons arriving in a given time interval is given by the Poisson law

$$p(N) = \frac{(\bar{N})^N e^{-\bar{N}}}{N!}$$

The mean-square fluctuation is given by

$$\overline{(\Delta N)^2} = \sum_{N=0}^{\infty} p(N)(N - \bar{N})^2 = \bar{N}$$

where

$$\bar{N} = \sum_{0}^{\infty} Np(N)$$

is the average N.

randomness. In the case of Poissonian photon arrival statistics (as is the case with ordinary lasers) and perfect photon emission ($\eta = 1$), the fluctuations are due to the photons. The opposite, hypothetical, case of no photon fluctuations but random photoemission ($\eta < 1$) corresponds to pure shot noise. The electrical measurement of noise power will yield the same result in either case and cannot distinguish between them.

11.5 PHOTOCONDUCTIVE DETECTORS

The operation of photoconductive detectors is illustrated in Figure 11-7. A semiconductor crystal is connected in series with a resistance R and a supply voltage V. The optical field to be detected is incident on and absorbed in the crystal, thereby exciting electrons into the conduction band (or, in p-type semiconductors, holes into the valence band). Such excitation results in a lowering of the resistance R_d of the semiconductor crystal and hence in an increase in the voltage drop across R, which, for $\Delta R_d/R_d \ll 1$, is proportional to the incident optical intensity.

To be specific, we show the energy levels involved in one of the more popular semiconductive detectors—mercury-doped germanium [7]. Mercury atoms enter germanium as acceptors with an ionization energy of 0.09 eV. It follows that it takes a photon energy of at least 0.09 eV (that is, a photon with a wavelength shorter than 14 μm) to lift an electron from the top of the valence band and have it trapped by the Hg (acceptor) atom. Usually the germanium crystal contains a smaller density N_D of donor atoms, which at low temperatures find it energetically profitable to lose their valence electrons to one of the far more numerous Hg acceptor atoms, thereby becoming positively ionized and ionizing (negatively) an equal number of acceptors.

Since the acceptor density $N_A \gg N_D$, most of the acceptor atoms remain neutrally charged.

An incident photon is absorbed and lifts an electron from the valence band onto an acceptor atom, as shown in process A in Figure 11-8. The electronic deficiency (that is, the hole) thus created is acted upon by the electric field, and its drift along the field direction gives rise to the signal

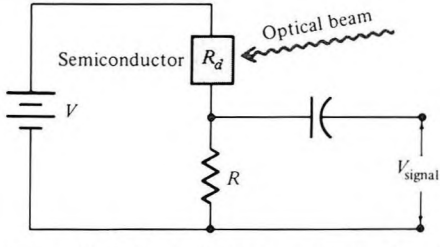


Figure 11-7 Typical biasing circuit of a photoconductive detector.

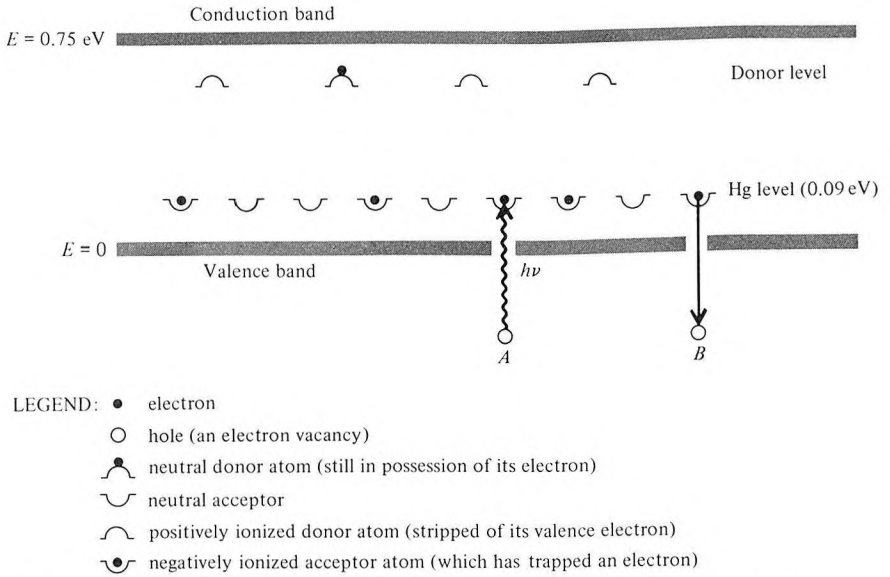


Figure 11-8 Donor and acceptor impurity levels involved in photoconductive semiconductors.

current. The contribution of a given hole to the current ends when an electron drops from an ionized acceptor level back into the valence band, thus eliminating the hole as in *B*. This process is referred to as electron–hole recombination or trapping of a hole by an ionized acceptor atom.

By choosing impurities with lower ionization energies, even lower-energy photons can be detected, and, indeed, photoconductive detectors commonly operate at wavelengths up to $\lambda = 50\ \mu$m. Cu, as an example, enters into Ge as an acceptor with an ionization energy of 0.04 eV, which would correspond to long-wavelength detection cutoff of $\lambda \simeq 32\ \mu$m. The response of a number of commercial photoconductive detectors is shown in Figure 11-9.

It is clear from this discussion that the main advantage of photoconductors compared to photomultipliers is their ability to detect long-wavelength radiation, since the creation of mobile carriers does not involve overcoming the large surface potential barrier. On the debit side we find the lack of current multiplication and the need to cool the semiconductor so that photoexcitation of carriers will not be masked by thermal excitation.

Consider an optical beam, of power P and frequency ν, that is incident on a photoconductive detector. Taking the probability for excitation of a carrier by an incident photon—the so-called quantum efficiency—as η, the carrier generation rate is $G = P\eta/h\nu$. If the carriers last on the average τ_0 seconds before recombining, the average number of carriers N_c is found by

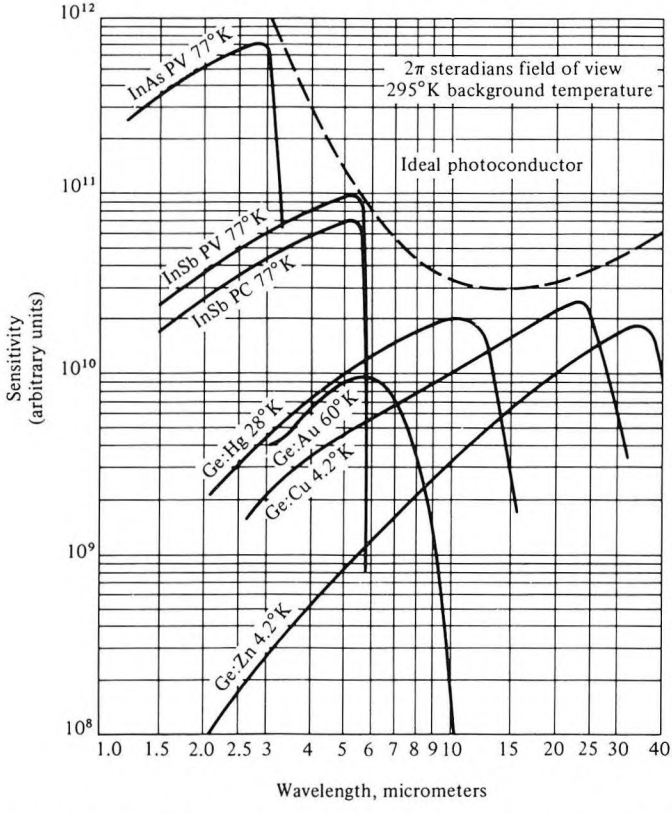


Figure 11-9 Relative sensitivity of a number of commercial photoconductors. (Courtesy Santa Barbara Research Corp.)

equating the generation rate to the recombination rate (N_c/τ_0), so

$$N_c = G\tau_0 = \frac{P\eta\tau_0}{h\nu} \tag{11.5-1}$$

Each one of these carriers drifts under the electric field influence[10] at a velocity $\bar{v}$ giving rise, according to (10.4-1), to a current in the external circuit of $i_e = e\bar{v}/d$, where d is the length (between electrodes) of the semiconductor crystal. The total current is thus the product of i_e and the number of carriers present, or, using (11.5-1),

$$\bar{i} = N_c i_e = \frac{P\eta\tau_0 e\bar{v}}{\mathrm{h}\nu\mathrm{d}} = \frac{e\eta}{h\nu}\left(\frac{\tau_0}{\tau_d}\right)P \tag{11.5-2}$$

[10]The drift velocity is equal to μE, where μ is the mobility and E is the electric field.

where $\tau_d = d/\bar{v}$ is the drift time for a carrier across the length d. The factor (τ_0/τ_d) is thus the fraction of the crystal length drifted by the average excited carrier before recombining.

Equation (11.5-2) describes the response of a photoconductive detector to a constant optical flux. Our main interest, however, is in the heterodyne mode of photoconductive detection, which, as has been shown in Section 11.4, allows detection sensitivities approaching the quantum limit. In order to determine the limiting sensitivity of photoconductive detectors, we need first to understand the noise contribution in these devices.

Generation Recombination Noise in Photoconductive Detectors

The principal noise mechanism in cooled photoconductive detectors reflects the randomness inherent in current flow. Even if the incident optical flux were constant in time, the generation of individual carriers by the flux would constitute a random process. This is exactly the type of randomness involved in photoemission, and we may expect, likewise, that the resulting noise will be shot noise. This is almost true except for the fact that in a photoconductive detector a photoexcited carrier lasts τ seconds[11] (its recombination lifetime) before being captured by an ionized impurity. The contribution of the carrier to the charge flow in the external circuit is thus $e(\tau/\tau_d)$, as is evident from inspection of (11.5-2). Since the lifetime τ is not a constant, but must be described statistically, another element of randomness is introduced into the current flow.

Consider a carrier excited by a photon absorption and lasting τ seconds. Its contribution to the external current is, according to (10.4-1)

$$i_e(t) = \begin{cases} \dfrac{e\bar{v}}{d} & 0 \le t \le \tau \\ 0 & \text{otherwise} \end{cases} \tag{11.5-3}$$

which has a Fourier transform

$$I_e(\omega, \tau) = \frac{e\bar{v}}{2\pi d} \int_0^\tau e^{-i\omega t}\, dt = \frac{-e\bar{v}}{d} [1 - e^{-i\omega\tau}] \tag{11.5-4}$$

so that

$$|I_e(\omega, \tau)|^2 = \frac{e^2\bar{v}^2}{4\pi^2\omega^2 d^2} [2 - e^{-i\omega\tau} - e^{i\omega\tau}] \tag{11.5-5}$$

According to (10.3-10) we need to average $|I_e(\omega, \tau)|^2$ over τ. This is done in a manner similar to the procedure used in Section 10.5. Taking the probability

[11] The parameter τ_0 appearing in (11.5-2) is the value of τ averaged over a large number of carriers.

function[12] $g(\tau) = \tau_0^{-1} \exp(-\tau/\tau_0)$, we average (11.5-5) over all the possible values of τ according to

$$\overline{|I_e(\omega)|^2} = \int_0^\infty |I_e(\omega, \tau)|^2 g(\tau)\, d\tau$$

$$= \frac{2e^2\bar{v}^2\tau_0^2}{4\pi^2 d^2(1 + \omega^2\tau_0^2)} \tag{11.5-6}$$

The spectral density function of the current fluctuations is obtained using Carson's theorem (10.3-10) as

$$S(\nu) = 2\bar{N}\frac{2e^2(\tau_0^2/\tau_d^2)}{1 + \omega^2\tau_0^2} \tag{11.5-7}$$

where we used $\tau_d = d/\bar{v}$ and where $\bar{N}$, the average number of carriers generated per second, can be expressed in terms of the average current $\bar{I}$ by use of the relation[13]

$$\bar{I} = \bar{N}e\frac{\tau_0}{\tau_d} \tag{11.5-8}$$

leading to

$$S(\nu) = \frac{4e\bar{I}(\tau_0/\tau_d)}{1 + 4\pi^2\nu^2\tau_0^2}$$

Therefore, the mean-square current representing the noise power in a frequency interval ν to $\nu + \Delta\nu$ is

$$\overline{i_N^2} \equiv S(\nu)\,\Delta\nu = \frac{4e\bar{I}(\tau_0/\tau_d)\Delta\nu}{1 + 4\pi^2\nu^2\tau_0^2} \tag{11.5-9}$$

which is the basic result for generation–recombination noise.

Numerical Example: Generation Recombination Noise in Hg Doped Germanium Photoconductive Detector

To better appreciate the kind of numbers involved in the expression for $\overline{i_N^2}$ we may consider a typical mercury-doped germanium detector operating at 20 K with the following characteristics:

$$d = 10^{-1}\ \text{cm}$$

$$\tau_0 = 10^{-9}\ \text{s}$$

[12] $g(\tau)\, d\tau$ is the probability that a carrier lasts between τ and $\tau + d\tau$ seconds before recombining.

[13] This relation follows from the fact that the average charge per carrier flowing through the external circuit is $e(\tau_0/\tau_d)$, which, when multiplied by the generation rate $\bar{N}$, gives the current.

$$V \text{ (across the length } d) = 10 \text{ volts} \Rightarrow E = 10^2 \text{ V/cm}$$

$$\mu = 3 \times 10^4 \text{ cm}^2/\text{V-s}$$

The drift velocity is $\bar{v} = \mu E = 3 \times 10^6$ cm/s and $\tau_d = d/\bar{v} \simeq 3.3 \times 10^{-8}$ second, and therefore $\tau_0/\tau_d = 3 \times 10^{-2}$. Thus, on the average, a carrier traverses only 3 percent of the length ($d = 1$ mm) of the sample before recombining. Comparing (11.5-9) to the shot-noise result (10.4-9), we find that for a given average current $\bar{I}$ the generation recombination noise is reduced from the shot-noise value by a factor

$$\frac{\overline{(i_N^2)}_{\text{generation-recombination}}}{\overline{(i_N^2)}_{\text{shot noise}}} \underset{\omega\tau_0 \ll 1}{=} 2\left(\frac{\tau_0}{\tau_d}\right) \tag{11.5-10}$$

which, in the foregoing example, has a value of about 1/15. Unfortunately, as will be shown subsequently, the reduced noise is accompanied by a reduction by a factor of (τ_0/τ_d) in the magnitude of the signal power, which wipes out the advantage of the lower noise.

Heterodyne Detection in Photoconductors

The situation here is similar to that described by Figure 11-6 in connection with heterodyne detection using photomultipliers. The signal field

$$e_s(t) = E_s \cos \omega_s t$$

is combined with a strong local-oscillator field

$$e_L(t) = E_L \cos(\omega + \omega_s)t \qquad E_L \gg E_s$$

so the total field incident on the photoconductor is

$$e(t) = \text{Re}(E_s\, e^{i\omega_s t} + E_L\, e^{i(\omega_s+\omega)t}) \equiv \text{Re}[V(t)] \tag{11.5-11}$$

The rate at which carriers are generated is taken, following (11.1-2), as $aV(t)V^*(t)$ where a is a constant to be determined. The equation describing the number of excited carriers N_c is thus

$$\frac{dN_c}{dt} = aVV^* - \frac{N_c}{\tau_0} \tag{11.5-12}$$

where τ_0 is the average carrier lifetime, so N_c/τ_0 corresponds to the carrier's decay rate. We assume a solution for $N_c(t)$ that consists of the sum of dc and a sinusoidal component in the form of

$$N_c(t) = N_0 + (N_1\, e^{i\omega t} + \text{c.c.}) \tag{11.5-13}$$

where c.c. stands for "complex conjugate."

Substitution in (11.5-12) gives

$$N_c(t) = a\tau_0(E_s^2 + E_L^2) + a\tau_0\left(\frac{E_s E_L\, e^{i\omega t}}{1 + i\omega\tau_0} + \text{c.c.}\right) \tag{11.5-14}$$

where we took E_s and E_L as real. The current through the sample is given by the number of carriers per unit length N_c/d times $e\bar{v}$, where $\bar{v}$ is the drift velocity

$$i(t) = \frac{N_c(t)e\bar{v}}{d} \tag{11.5-15}$$

which, using (11.5-14), gives

$$i(t) = \frac{e\bar{v}a\tau_0}{d}\left(E_s^2 + E_L^2 + \frac{2E_sE_L\cos(\omega t - \phi)}{\sqrt{1 + \omega^2\tau_0^2}}\right) \tag{11.5-16}$$

where $\phi = \tan^{-1}(\omega\tau_0)$.

The current is thus seen to contain a signal component that oscillates at ω and is proportional to E_s. The constant a in (11.5-16) can be determined by requiring that, when $P_s = 0$, the expression for the direct current predicted by (11.5-16) agree with (11.5-2). This condition is satisfied if we rewrite (11.5-16) as

$$i(t) = \frac{e\eta}{h\nu}\left(\frac{\tau_0}{\tau_d}\right)\left[P_s + P_L + \frac{2\sqrt{P_sP_L}}{\sqrt{1 + \omega^2\tau_0^2}}\cos(\omega t - \phi)\right] \tag{11.5-17}$$

where P_s and P_L refer, respectively, to the incident-signal and local-oscillator powers and η, the quantum efficiency, is the number of carriers excited per incident photon. The signal current is thus

$$i_s(t) = \frac{2e\eta}{h\nu}\left(\frac{\tau_0}{\tau_d}\right)\frac{\sqrt{P_sP_L}}{\sqrt{1 + \omega^2\tau_0^2}}\cos(\omega t - \phi) \tag{11.5-18}$$

while the dc (average) current is

$$\bar{I} = \frac{e\eta}{h\nu}\left(\frac{\tau_0}{\tau_d}\right)(P_s + P_L) \tag{11.5-19}$$

Since the average current $\bar{I}$ appearing in the expression (11.5-9) for the generation recombination noise is given in this case by

$$\bar{I} = \left(\frac{e\eta}{h\nu}\right)\left(\frac{\tau_0}{\tau_d}\right)P_L \qquad P_L \gg P_S$$

we can, by increasing P_L, increase the noise power $\overline{i_N^2}$ and at the same time, according to (11.5-18), the signal $\overline{i_s^2}$ until the generation recombination noise (11.5-9) is by far the largest contribution to the total output noise. When this condition is satisfied, the signal-to-noise ratio can be written, using (11.5-9), (11.5-18), and (11.5-19) and taking $P_L \gg P_s$, as

$$\frac{S}{N} = \frac{\overline{i_s^2}}{\overline{i_N^2}} = \frac{2(e\eta\tau_0/h\nu\tau_d)^2P_sP_L/(1 + \omega^2\tau_0^2)}{4e^2\eta(\tau_0/\tau_d)^2P_L\,\Delta\nu/(1 + \omega^2\tau_0^2)h\nu} = \frac{P_s\eta}{2h\nu\,\Delta\nu} \tag{11.5-20}$$

The minimum detectable signal—that which leads to a signal-to-noise ratio

of unity—is found by setting the left side of (11.5-20) equal to unity and solving for P_s. It is

$$(P_s)_{\min} = \frac{2h\nu\,\Delta\nu}{\eta} \tag{11.5-21}$$

which, for the same η, is twice that of the photomultiplier heterodyne detection as given by (11.4-10). In practice, however, η in photoconductive detectors can approach unity, whereas in the best photomultipliers $\eta \simeq 30$ percent.

Numerical Example: Minimum Detectable Power of a Heterodyne Receiver Using a Photoconductor at 10.6 μm

Assume the following:

$$\lambda = 10.6\ \mu\text{m}$$

$$\Delta\nu = 1\ \text{Hz}$$

$$\eta \simeq 1$$

Substitution in (11.5-21) gives a minimum detectable power of

$$(P_s)_{\min} \simeq 10^{-19}\ \text{watt}$$

Experiments ([8, 9]) have demonstrated that the theoretical signal-to-noise ratio as given by (11.5-20) can be realized quite closely in practice; see Figure 11-10.

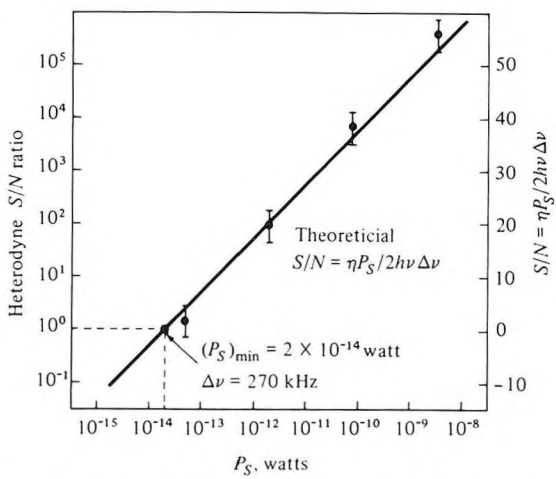


Figure 11-10 Signal-to-noise ratio of heterodyne signal to Ge:Cu detector at a heterodyne frequency of 70 MHz. Data points represent observed values. (After Reference [8].)

11.6 THE *p-n* JUNCTION

Before embarking on a description of the *p-n* diode detector, we need to understand the operation of the semiconductor *p-n* junction. Consider the junction illustrated in Figure 11-11. It consists of an abrupt transition from a donor-doped (that is, *n*-type) region of a semiconductor, where the charge carriers are predominantly electrons, to an acceptor-doped (*p*-type) region, where the carriers are holes. The doping profile—that is, the density of excess donor (in the *n* region) atoms or acceptor atoms (in the *p* region)—is shown in Figure 11-11(a). This abrupt transition results usually from diffusing suitable impurity atoms into a substrate of a semiconductor with the opposite type of conductivity. In our slightly idealized abrupt junction we assume that the *n* region ($x > 0$) has a constant (net) donor density N_D and the *p* region ($x < 0$) has a constant acceptor density N_A.

The energy-band diagram at zero applied bias is shown in Figure 11-11(b). The top (or bottom) curve can be taken to represent the potential energy of an electron as a function of position x, so the minimum energy needed to take an electron from the *n* to the *p* side of the junction is eV_d. Taking the separations of the Fermi level from the respective band edges as ϕ_n and ϕ_p as shown, we have

$$eV_d = E_g - (\phi_n + \phi_p)$$

V_d is referred to as the "built-in" junction potential.

Figure 11-11(c) shows the potential distribution in the junction with an applied reverse bias of magnitude V_a. This leads to a separation of eV_a between the Fermi levels in the *p* and *n* regions and causes the potential barrier across the junction to increase from eV_d to $e(V_d + V_a)$. The change of potential between the *p* and *n* regions is due to a sweeping of the mobile charge carriers from the region $-l_p < x < l_n$, giving rise to a charge double layer of stationary (ionized) impurity atoms, as shown in Figure 11-11(d).

In the analytical treatment of the problem we assume that in the depletion layer ($-l_p < x < l_n$) the excess impurity atoms are fully ionized and thus, using $\nabla \cdot \mathbf{E} = \rho/\epsilon$ and $\mathbf{E} = -\nabla V$, where V is the potential, we have

$$\frac{d^2V}{dx^2} = \frac{eN_A}{\epsilon} \qquad \text{for } -l_p < x < 0 \tag{11.6-1}$$

and

$$\frac{d^2V}{dx^2} = -\frac{eN_D}{\epsilon} \qquad 0 < x < l_n \tag{11.6-2}$$

where the charge of the electron is $-e$ and the dielectric constant is ϵ. The boundary conditions are

$$E = -\frac{dV}{dx} = 0 \text{ at } x = -l_p \quad \text{and } x = +l_n \tag{11.6-3}$$

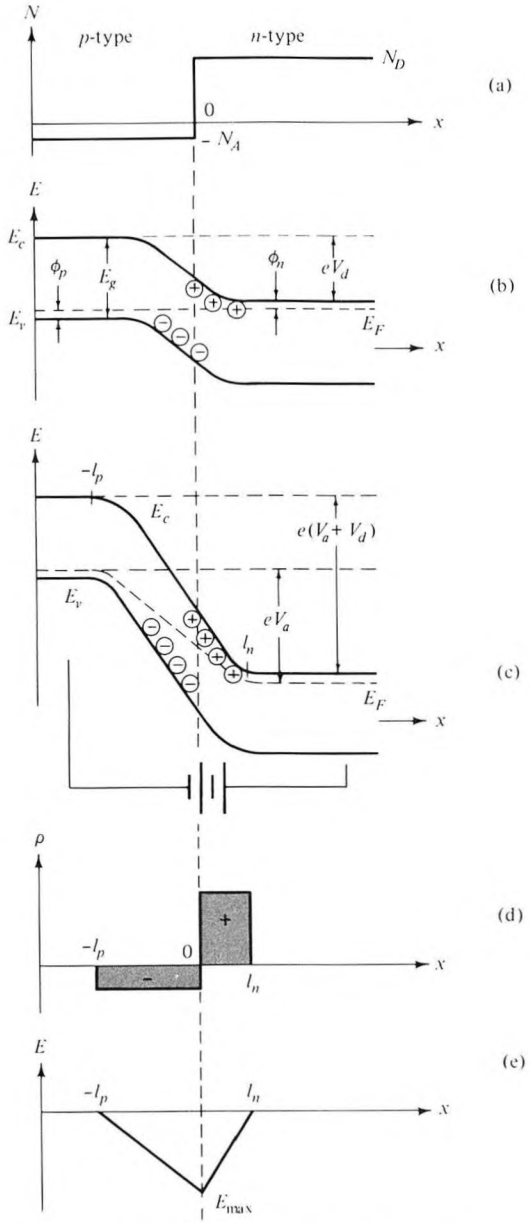


Figure 11-11 The abrupt *p-n* junction. (a) Impurity profile. (b) Energy-band diagram with zero applied bias. (c) Energy-band diagram with reverse applied bias. (d) Net charge density in the depletion layer. (e) The electric field. The circles in (b) and (c) represent ionized impurity atoms in the depletion layer.

$$V \text{ and } \frac{dV}{dx} \text{ are continuous at } x = 0 \tag{11.6-4}$$

$$V(l_n) - V(-l_p) = V_d + V_a \tag{11.6-5}$$

The solutions of (11.6-1) and (11.6-2) conforming with the arbitrary choice of $V(0) = 0$ are

$$V = \frac{e}{2\epsilon} N_A(x^2 + 2l_p x) \qquad \text{for } -l_p < x < 0 \tag{11.6-6}$$

$$V = -\frac{e}{2\epsilon} N_D(x^2 - 2l_n x) \qquad 0 < x < l_n \tag{11.6-7}$$

which, using (11.6-4), gives

$$N_A l_p = N_D l_n \tag{11.6-8}$$

so the double layer contains an equal amount of positive and negative charge.

Condition (11.6-5) gives

$$V_d + V_a = \frac{e}{2\epsilon} (N_D l_n^2 + N_A l_p^2) \tag{11.6-9}$$

which, together with (11.6-8) leads to

$$l_p = (V_d + V_a)^{1/2} \left(\frac{2\epsilon}{e}\right)^{1/2} \left(\frac{N_D}{N_A(N_A + N_D)}\right)^{1/2} \tag{11.6-10}$$

$$l_n = (V_d + V_a)^{1/2} \left(\frac{2\epsilon}{e}\right)^{1/2} \left(\frac{N_A}{N_D(N_A + N_D)}\right)^{1/2} \tag{11.6-11}$$

and, therefore, as before,

$$\frac{l_p}{l_n} = \frac{N_D}{N_A} \tag{11.6-12}$$

Differentiation of (11.6-6) and (11.6-7) yields

$$E = -\frac{e}{\epsilon} N_A(x + l_p) \qquad \text{for } -l_p < x < 0$$

$$E = -\frac{e}{\epsilon} N_D(l_n - x) \qquad 0 < x < l_n \tag{11.6-13}$$

The field distribution of (11.6-13) is shown in Figure 11-11(e). The maximum field occurs at $x = 0$ and is given by

$$E_{\max} = -2(V_d + V_a)^{1/2} \left(\frac{e}{2\epsilon}\right)^{1/2} \left(\frac{N_D N_A}{N_A + N_D}\right)^{1/2}$$

$$= -\frac{2(V_d + V_a)}{l_p + l_n} \tag{11.6-14}$$

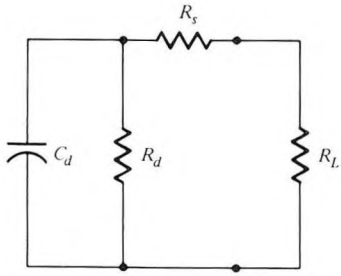


Figure 11-12 Equivalent circuit of a *p-n* junction. In typical back-biased diodes, $R_d \gg R_s$ and R_L, and $R_L \gg R_s$, so the resistance across the junction can be taken as equal to the load resistance R_L.

The presence of a charge $Q = -eN_A l_p$ per unit junction area on the p side and an equal and negative charge on the n side leads to a junction capacitance. The reason is that l_p and l_n depend, according to (11.6-10) and (11.6-11), on the applied voltage V_a, so a change in voltage leads to a change in the charge $eN_A l_p = eN_D l_n$ and hence to a differential capacitance per unit area,[14] given by

$$\frac{C_d}{\text{area}} \equiv \frac{dQ}{dV_a} = eN_A \frac{dl_p}{dV_a}$$

$$= \left(\frac{\epsilon e}{2}\right)^{1/2} \left(\frac{N_A N_D}{N_A + N_D}\right)^{1/2} \left(\frac{1}{V_a + V_d}\right)^{1/2} \qquad \text{(11.6-15)}$$

which, using (11.6-10) and (11.6-11), can be shown to be equal to

$$\frac{C_d}{\text{area}} = \frac{\epsilon}{l_p + l_n} \qquad \text{(11.6-16)}$$

as appropriate to a parallel-plate capacitance of separation $l = l_p + l_n$. The equivalent circuit of a *p-n* junction is shown in Figure 11-12. The capacitance C_d was discussed above. The diode shunt resistance R_d in back-biased junctions is usually very large ($>10^6$ ohms) compared to the load impedance R_L and can be neglected. The resistance R_s represents ohmic losses in the bulk p and n regions adjacent to the junction.

11.7 SEMICONDUCTOR PHOTODIODES

Semiconductor *p-n* junctions are used widely for optical detection: see References [10–12]. In this role they are referred to as junction photodiodes. The main physical mechanisms involved in junction photodetection are il-

[14]The capacitance is defined by $C = Q/V_a$, whereas the differential capacitance $C_d = dQ/dV_a$ is the capacitance "seen" by a small ac voltage when the applied bias is V_a.

lustrated in Figure 11-13. At A, an incoming photon is absorbed in the p side creating a hole and a free electron. If this takes place within a diffusion length (the distance in which an excess minority concentration is reduced to e^{-1} of its peak value, or in physical terms, the average distance a minority carrier traverses before recombining with a carrier of the opposite type) of the depletion layer, the electron will, with high probability, reach the layer boundary and will drift under the field influence across it. An electron traversing the junction contributes a charge e to the current flow in the external circuit, as described in Section 10.4. If the photon is absorbed near the n side of the depletion layer, as shown at C, the resulting hole will diffuse to the junction and then drift across it again, giving rise to a flow of charge e in the external load. The photon may also be absorbed in the depletion layer as at B, in which case both the hole and electron that are created drift (in opposite directions) under the field until they reach the p and n sides, respectively. Since in this case each carrier traverses a distance that is less than the full junction width, the contribution of this process to charge flow in the external circuit is, according to (10.4-1) and (10.4-7), e. In practice this last process is the most desirable, since each absorption gives rise to a charge e, and delayed current response caused by finite diffusion time is avoided. As a result, photodiodes often use a p-i-n structure in which an intrinsic high resistivity (i) layer is sandwiched between the p and n regions. The potential drop occurs mostly across this layer, which can be made long enough to ensure that most of the incident photons are absorbed within it. Typical construction of a p-i-n photodiode is shown in Figure 11-14.

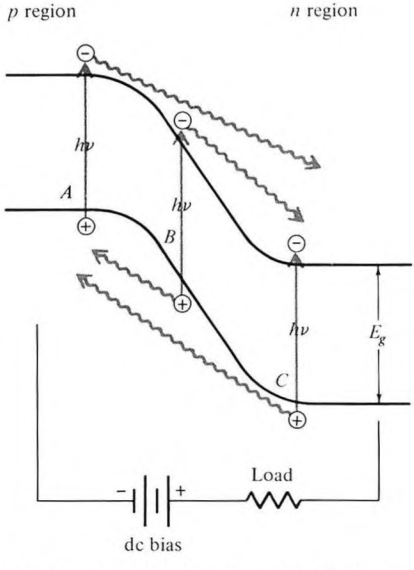


Figure 11-13 The three types of electron–hole pair creation by absorbed photons that contribute to current flow in a p-n photodiode.

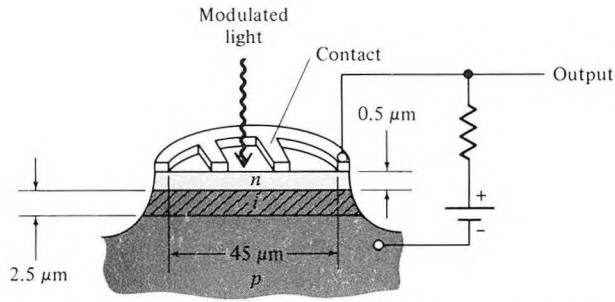


Figure 11-14 A *p-i-n* photodiode. (After Reference [13].)

It is clear from Figure 11-13 that a photodiode is capable of detecting only radiation with photon energy $h\nu > E_g$, where E_g is the energy gap of the semiconductor. If, on the other hand, $h\nu \gg E_g$, the absorption, which in a semiconductor increases strongly with frequency, will take place entirely near the input face (in the n region of Figure 11-14) and the minority carriers generated by absorbed photons will recombine with majority carriers before diffusing to the depletion layer. This event does not contribute to the current flow and, as far as the signal is concerned, is wasted. This is why the photoresponse of diodes drops off when $h\nu > E_g$. Typical frequency response curves of photodiodes are shown in Figure 11-15. The number of carriers flowing in the external circuit per incident photon, the so-called quantum efficiency, is seen to approach 50 percent in Ge.

Frequency Response of Photodiodes

One of the major considerations in optical detectors is their frequency response—that is, the ability to respond to variations in the incident intensity such as those caused by high-frequency modulation. The three main mechanisms limiting the frequency response in photodiodes are:

1. The finite diffusion time of carriers produced in the p and n regions. This factor was described in the last section, and its effect can be minimized by a proper choice of the length of the depletion layer.
2. The shunting effect of the signal current by the junction capacitance C_d shown in Figure 11-12. This places an upper limit of

$$\omega_m \simeq \frac{1}{R_e C_d} \tag{11.7-1}$$

on the intensity modulation frequency where R_e is the equivalent resistance in parallel with the capacitance C_d.
3. The finite transit time of the carriers drifting across the depletion layer.

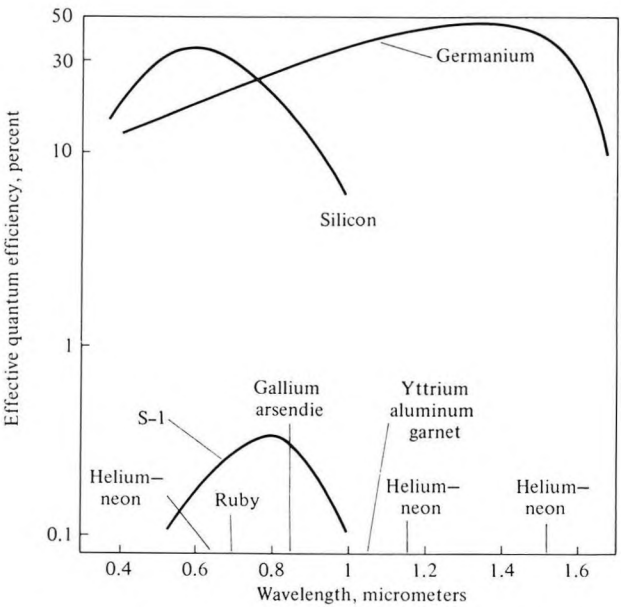


Figure 11-15 Quantum efficiencies for silicon and germanium photodiodes compared with the efficiency of the S-1 photodiode used in a photomultiplier tube. Emission wavelengths for various lasers are also indicated. (After Reference [13].)

To analyze first the limitation due to transit time, we assume the slightly idealized case in which the carriers are generated in a *single* plane, say point A in Figure 11-13, and then drift the full width of the depletion layer at a constant velocity v. For high enough electric fields, the drift velocity of carriers in semiconductors tends to saturate, so the constant velocity assumption is not very far from reality even for a nonuniform field distribution, such as that shown in Figure 11-11(e), provided the field exceeds its saturation value over most of the depletion layer length. The saturation of the whole velocity in germanium, as an example, is illustrated by the data of Figure 11-16.

The incident optical field is taken as

$$\begin{aligned} e(t) &= E_s(1 + m \cos \omega_m t) \cos \omega t \\ &\equiv \text{Re}[V(t)] \end{aligned} \tag{11.7-2}$$

where

$$V(t) \equiv E_s(1 + m \cos \omega_m t)\, e^{i\omega t} \tag{11.7-3}$$

Thus, the amplitude is modulated at a frequency $\omega_m/2\pi$. Following the discussion of Section 11.1 we take the generation rate $G(t)$; that is, the number of carriers generated per second, as proportional to the average of $e^2(t)$ over a time long compared to the optical period $2\pi/\omega$. This average is equal to

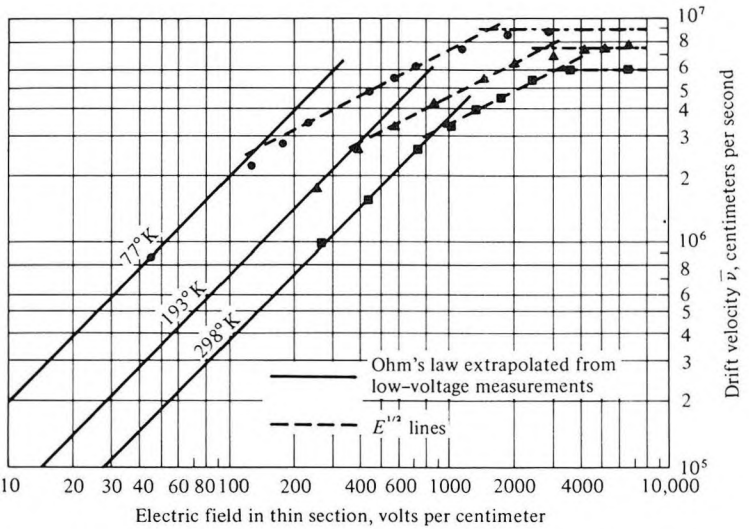


Figure 11-16 Experimental data showing the saturation of the drift velocity of holes in germanium at high electric fields. (After Reference [14].)

$\frac{1}{2}V(t)V^*(t)$, so the generation rate is taken as

$$G(t) = aE_s^2\left[\left(1 + \frac{m^2}{2}\right) + 2m\cos\omega_m t + \frac{m^2}{2}\cos 2\omega_m t\right] \tag{11.7-4}$$

where a is a proportionality constant to be determined. Dropping the term involving $\cos 2\omega_m t$ and using complex notation, we rewrite $G(t)$ as

$$G(t) = aE_s^2\left[1 + \frac{m^2}{2} + 2me^{i\omega_m t}\right] \tag{11.7-5}$$

A single carrier drifting at a velocity $\bar{v}$ contributes, according to (10.4-1), an instantaneous current

$$i = \frac{e\bar{v}}{d} \tag{11.7-6}$$

to the external circuit, where d is the width of the depletion layer. The current due to carriers generated between t' and $t' + dt'$ is $(e\bar{v}/d)G(t')\,dt'$ but, since each carrier spends a time $\tau_d = d/\bar{v}$ in transit, the instantaneous current at time t is the sum of contributions of carriers generated between t and $t - \tau_d$

$$i(t) = \frac{e\bar{v}}{d}\int_{t-\tau_d}^{t} G(t')\,dt' = \frac{e\bar{v}aE_s^2}{d}\int_{t-\tau_d}^{t}\left(1 + \frac{m^2}{2} + 2m\,e^{i\omega_m t'}\right)dt'$$

and, after integration,

$$i(t) = \left(1 + \frac{m^2}{2}\right) eaE_s^2 + 2meaE_s^2 \left(\frac{1 - e^{-i\omega_m \tau_d}}{i\omega_m \tau_d}\right) e^{i\omega_m t} \qquad (11.7\text{-}7)$$

The factor $(1 - e^{-i\omega_m \tau_d})/i\omega_m \tau_d$ represents the phase lag as well as the reduction in signal current due to the finite drift time τ_d. If the drift time is short compared to the modulation period, so $\omega_m \tau_d \ll 1$, it has its maximum value of unity, and the signal is maximum. This factor is plotted in Figure 11-17 as a function of the transit phase angle $\omega_m \tau_d$. We can determine the value of the constant a in (11.7-7) by requiring that (11.7-7) agree with the experimental observation according to which in the absence of modulation, $m = 0$, each incident photon will create η carriers. Thus the dc (average) current is

$$\bar{I} = \frac{Pe\eta}{h\nu} \qquad (11.7\text{-}8)$$

where P is the optical (signal) power when $m = 0$. Using (11.7-8), we can rewrite (11.7-7) as

$$i(t) = \frac{Pe\eta}{h\nu}\left(1 + \frac{m^2}{2}\right) + \frac{Pe\eta}{h\nu} 2m \left(\frac{i - e^{-i\omega_m \tau_d}}{i\omega_m \tau_d}\right) e^{i\omega_m t} \qquad (11.7\text{-}9)$$

To evaluate the effect of the other limiting factors on the modulation frequency response of a photodiode, we refer to the diode equivalent ac circuit in Figure 11-18. Here R_d is the diode incremental (ac) resistance, C_d the junction capacitance, R_s represents the contact and series resistance, L_p the

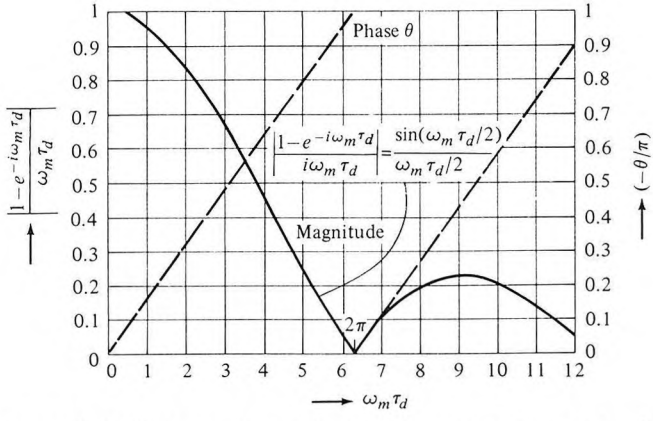


Figure 11-17 Phase and magnitude of the transit-time reduction factor $(1 - e^{-i\omega_m \tau_d})/i\omega_m \tau_d$.

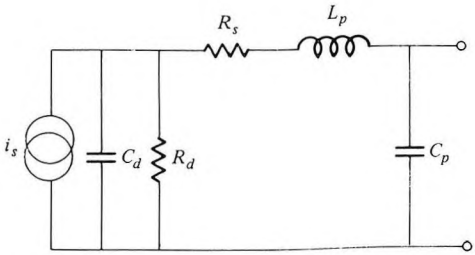


Figure 11-18 The equivalent high-frequency circuit of a semiconductor photodiode.

parasitic inductance associated mostly with the contact leads, and C_p the parasitic capacitance due to the contact leads and the contact pads.

Recent advances [20–22, 24] have resulted in metal–GaAs (Schottky) diodes with frequency response extending up to 10^{11} Hz. Figure 11-19 shows a schematic diagram of such a diode. This high-frequency limit was achieved by using a very small area (5 μm × 5 μm) that minimizes C_d, by using extremely short contact leads to reduce R_s and L_p, by fabricating the diode on semi-insulating GaAs substrate [20] to reduce C_p, and by using a thin (0.3 μm) n^-GaAs drift region to reduce the transit time. The resulting measured frequency response is shown in Figure 11-20. The measurement of the frequency response up to 100 GHz is by itself a considerable achievement. This was accomplished by first obtaining the impulse response of the photodiode by exciting it with picosecond pulses (which, for the range of fre-

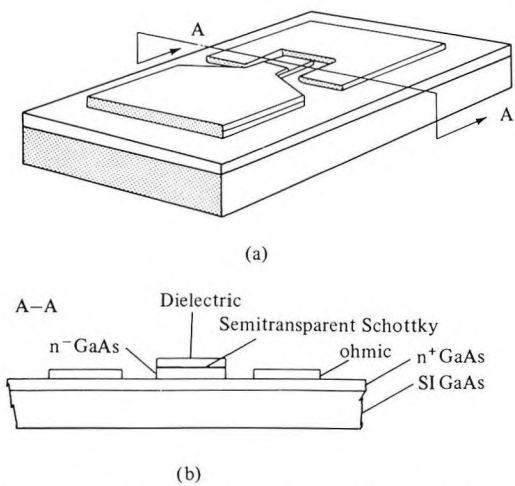


Figure 11-19 (a) Planar GaAs Schottky photodiode. (b) Cross section along A–A. The n^-GaAs layer (0.3 μm thick) and the n^+GaAs (0.4 μm thick) are grown by liquid-phase epitaxy on semi-insulating GaAs substrate. The semitransparent Schottky consists of 100 Å of Pt (After Reference [22].)

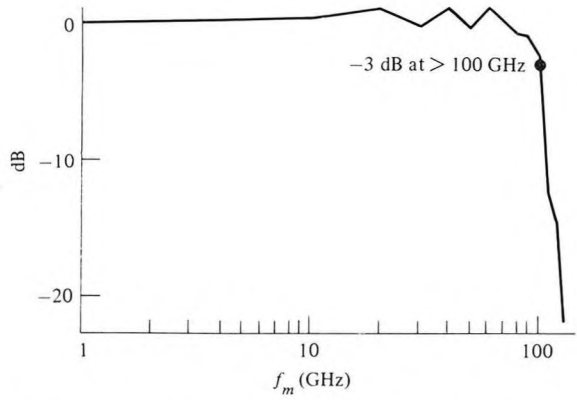


Figure 11-20 The modulation frequency response of the Schottky photodiode shown in Figure 11-19. (After Reference [22].)

quencies of interest, may be considered as delta functions) from a mode-locked laser [21]. The diode response, which is only a few picoseconds long, is measured by a new electrooptic sampling technique [23, 24]. The frequency response, as plotted in Figure 11-20, is obtained by taking the Fourier transform of the measured impulse response.

Numerical Example: Modulation Response of a GaAs *p-n* Junction Photodiode

Let us calculate the upper limit on the frequency response of the diode shown in Figure 11-19. The following data apply.

$$\text{Area} = 5\ \mu\text{m} \times 5\ \mu\text{m}$$

$$\epsilon = 12.25\epsilon_0$$

$$d = 0.3\ \mu\text{m}\ (= \text{thickness of drift region})$$

$$\bar{v} = 10^7\ \text{cm/s (saturation velocity of electron in GaAs)}$$

$$R_s \approx 10\ \text{ohms}$$

The transit-time limit f_m is obtained from the condition $2\pi f_m \tau_d = 2$. This, according to Figure 11-17, is the frequency where the response is down to 84 percent of its maximum (zero-frequency) value. The result is

$$f_m \sim \frac{\bar{v}}{\pi d} \sim 1.06 \times 10^{11}\ \text{Hz}$$

The junction capacitance, based on the above data, is $\sim 10^{-14}$ farad. The

parasitic capacitance can be kept in this case to $\sim 10^{-13}$ farad. Since the resistance R_d of the reverse-biased junction is very large, it is usually neglected.

The circuit limit to the frequency response is $f_m \sim 1/(2\pi R_s C_p) = 1.59 \times 10^{11}$ Hz. Since this value is larger than the transit-time limit, we conclude that the frequency response is transit-time limited to a value $\sim 10^{11}$ Hz, which is in agreement with the value obtained from Figure 11-20.

Detection Sensitivity of Photodiodes

We assume that the modulation frequency of the light to be detected is low enough that the transit time factor is unity and that the condition

$$\omega_m \ll \frac{1}{R_e C_d} \tag{11.7-10}$$

is fulfilled and, therefore, according to (11.7-1), the shunting of signal current by the diode capacitance C_d can be neglected. The diode current is given by (11.7-9) as

$$i(t) = \frac{Pe\eta}{h\nu}\left(1 + \frac{m^2}{2}\right) + \frac{Pe\eta}{h\nu} 2m\, e^{i\omega_m t} \tag{11.7-11}$$

The noise equivalent circuit of a diode connected to a load resistance R_L is shown in Figure 11-21. The signal power is proportional to the mean-square value of the sinusoidal current component, which, for $m = 1$, is

$$\overline{i_s^2} = 2\left(\frac{Pe\eta}{h\nu}\right)^2 \tag{11.7-12}$$

Two noise sources are shown. The first is the shot noise associated with the random generation of carriers. Using (10.4-9), this is represented by a noise generator $\overline{i_{N_1}^2} = 2e\overline{I}\,\Delta\nu$, where $\overline{I}$ is the average current as given by the first

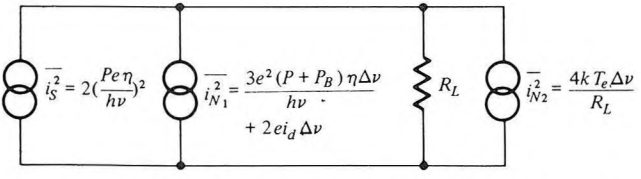


Figure 11-21 Noise equivalent circuit of a photodiode operating in the direct (video) mode. The modulation index m is taken as unity, and it is assumed that the modulation frequency is low enough that the junction capacitance and transit-time effects can be neglected. The resistance R_L is assumed to be much smaller than the shunt resistance R_d of the diode, so the latter is neglected. Also neglected is the series diode resistance, which is assumed small compared with R_L.

term on the right side of (11.7-11). Taking $m = 1$, we obtain

$$\overline{i_{N_1}^2} = \frac{3e^2(P + P_B)\eta\,\Delta\nu}{h\nu} + 2ei_d\Delta\nu \tag{11.7-13}$$

where P_B is the background optical power entering the detector (in addition to the signal power) and i_d is the "dark" direct current that exists even when $P_s = P_b = 0$. The second noise contribution is the thermal (Johnson noise) generated by the output load, which, using (10.5-9), is given by

$$\overline{i_{N_2}^2} = \frac{4kT_e\Delta\nu}{R_L} \tag{11.7-14}$$

where T_e is chosen to include the equivalent input noise power of the amplifier following the diode.[15] The signal-to-noise power ratio at the amplifier output is thus

$$\frac{S}{N} = \frac{\overline{i_s^2}}{\overline{i_{N_1}^2} + \overline{i_{N_2}^2}} = \frac{2(Pe\eta/h\nu)^2}{3e^2(P + P_B)\eta\Delta\nu/h\nu + 2ei_d\Delta\nu + 4kT_e\Delta\nu/R_L} \tag{11.7-15}$$

In most practical systems the need to satisfy Equation (11.7-10) forces one to use small values of load resistance R_L. Under these conditions and for values of P that are near the detectability limit ($S/N = 1$), the noise term (11.7-14) is much larger than the shot noise (11.7-13) and the detector is consequently not operating near its quantum limit. Under these conditions we have

$$\frac{S}{N} \simeq \frac{2(Pe\eta/h\nu)^2}{4kT_e\Delta\nu/R_L} \tag{11.7-16}$$

---

[15] In practice it is imperative that the signal-to-noise ratio take account of the noise power contributed by the amplifier. This is done by characterizing the "noisiness" of the amplifier by an effective input noise "temperature" T_A. The amplifier noise power measured at its output is taken as $GkT_A\Delta\nu$, where G is the power gain. (A hypothetical noiseless amplifier will thus be characterized by $T_A = 0$.) This power can be referred to the input by dividing by G, thus becoming $kT_A\Delta\nu$. The total effective noise power at the amplifier input is the sum of this power and the Johnson noise $kT\Delta\nu$ due to the diode load resistance; that is, $k(T + T_A)\Delta\nu \equiv kT_e\Delta\nu$. The amplifier noise temperature T_A is related to its "noise figure" F by the definition

$$F = 1 + \frac{T_A}{290}$$

It follows that the noise power generated within the amplifier and measured at its output is

$$N_A = GkT_A\Delta\nu = G(F - 1)kT_0\Delta\nu$$

where $T_0 = 290$. The ratio of the signal-to-noise power ratio at the input of the amplifier to the same ratio at the output is thus

$$\frac{(S/N)_{\text{in}}}{(S/N)_{\text{out}}} = \frac{S_{\text{in}}[G(F - 1)kT_0\Delta\nu + GkT\Delta\nu]}{kT\Delta\nu GS_{\text{in}}}$$

This ratio becomes equal to the "noise figure" F when the temperature T of the detector output load is equal to T_0. (Note that the choice $T_0 = 290$ is a matter of, universal, convention.)

The "minimum detectable optical power" is by definition that yielding $S/N = 1$ and is, from (11.7-16),

$$(P)_{\min} = \frac{h\nu}{e\eta}\sqrt{\frac{2kT_e\Delta\nu}{R_L}} \tag{11.7-17}$$

which is to be compared to the theoretical limit of $h\nu\,\Delta\nu/\eta$, which, according to (11.3-11), obtains when the signal shot-noise term predominates. In practice, the value of R_L is related to the desired modulation bandwidth $\Delta\nu$ and the junction capacitance C_d by

$$\Delta\nu \simeq \frac{1}{2\pi R_L C_d} \tag{11.7-18}$$

which, when used in (11.7-16), gives

$$P_{\min} \simeq 2\sqrt{\pi}\,\frac{h\nu\Delta\nu}{e\eta}\sqrt{kT_eC_d} \tag{11.7-19}$$

This shows that sensitive detection requires the use of small area junctions so that C_d will be at a minimum.

Numerical Example: Minimum Detectable Power in the Case of Amplifier Limited Detection

Assume a typical Ge photodiode operating at $\lambda = 1.4\ \mu$m with $C_d = 1$ pF, $\Delta\nu = 1$ GHz, and $\eta = 50$ percent. Let the amplifier following the diode have an effective noise temperature $T_e = 1200 + 290 = 1490$ K (see footnote 15) [14–15]. Substitution in (11.7-19) gives

$$P_{\min} \simeq 3.34 \times 10^{-7}\ \text{watt}$$

for the minimum detectable signal power.

11.8 THE AVALANCHE PHOTODIODE

By increasing the reverse bias across a p-n junction, the field in the depletion layer can increase to a point at which carriers (electrons or holes) that are accelerated across the depletion layer can gain enough kinetic energy to "kick" new electrons from the valence to the conduction band, while still traversing the layer. This process, illustrated in Figure 11-22, is referred to as avalanche multiplication. An absorbed photon (A) creates an electron–hole pair. The electron is accelerated until at point C it has gained sufficient energy to excite an electron from the valence to the conduction

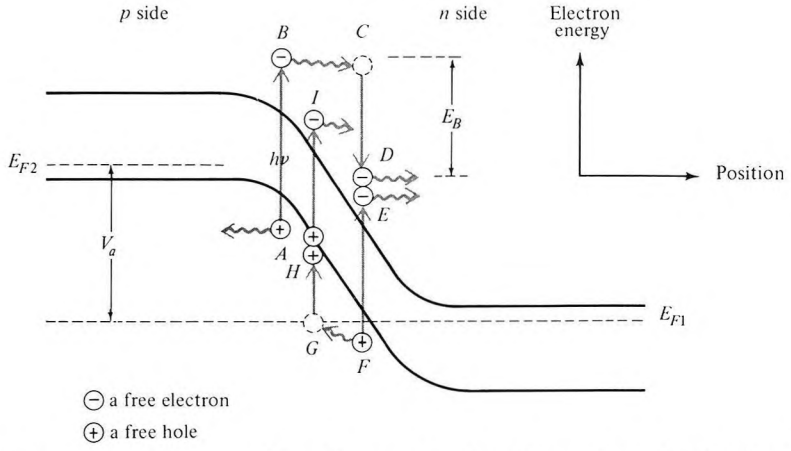


Figure 11-22 Energy-position diagram showing the carrier multiplication following a photon absorption in a reverse-biased avalanche photodiode.

band, thus creating a new electron–hole pair. The newly generated carriers drift in turn in opposite directions. The hole (F) can also cause carrier multiplication as in G. The result is a dramatic increase (avalanche) in junction current that sets in when the electric field becomes high enough. This effect, discovered first in gaseous plasmas and more recently in p-n junctions (References [15, 16]), gives rise to a multiplication of the current over its value in an ordinary (nonavalanching) photodiode. An experimental plot of the current gain M as a function of the junction field is shown in Figure 11-23.[16]

Avalanche photodiodes are similar in their construction to ordinary photodiodes except that, because of the steep dependence of M on the applied field in the avalanche region, special care must be exercised to obtain very uniform junctions. A sketch of an avalanche photodiode is shown in Figure 11-24.

Since an avalanche photodiode is basically similar to a photodiode, its equivalent circuit elements are given by expressions similar to those given above for the photodiode. Its frequency response is similarly limited by diffusion, drift across the depletion layer, and capacitive loading, as discussed in Section 11.7.

A multiplication by a factor M of the photocurrent leads to an increase by M^2 of the signal power S over that which is available from a photodiode

[16]If the probability that a photo-excited electron–hole pair will create another pair during its drift is denoted by p, the current multiplication is

$$M = (1 + p + p^2 + p^3 + \cdots) = \frac{1}{1 - p}$$

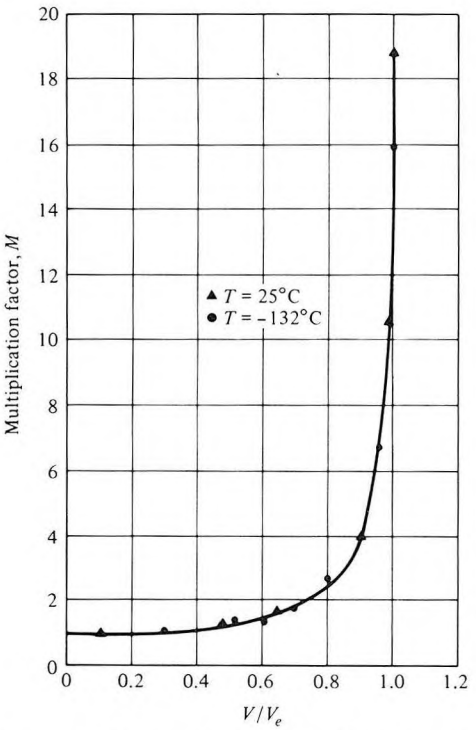


Figure 11-23 Current multiplication factor in an avalanche diode as a function of the electric field. (After Reference [16].)

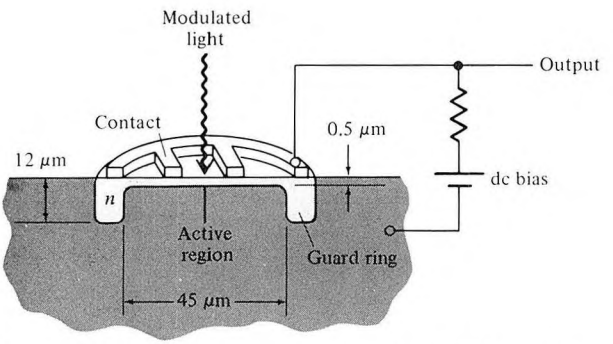


Figure 11-24 Planar avalanche photodiode. (After Reference [13].)

so that, using (11.7-12), we get

$$S \propto \overline{i_s^2} = 2M^2\left(\frac{Pe\eta}{h\nu}\right)^2 \tag{11.8-1}$$

where P is the optical power incident on the diode. This result is reminiscent of the signal power from a photomultiplier as given by the numerator of (11.3-9), where the avalanche gain M plays the role of the secondary electron multiplication gain G. We may expect that, similarly, the shot-noise power will also increase by M^2. The shot noise, however, is observed to increase as M^n, where $2 < n < 3$.[17] Experimental observation of a near ideal $M^{2.1}$ behavior is shown in Figure 11-25.

The signal-to-noise power ratio at the output of the diode is thus given, following (11.7-15), by

$$\frac{S}{N} = \frac{2M^2(Pe\eta/h\nu)^2}{[3e^2(P + P_B)\eta\Delta\nu/h\nu]M^n + 2ei_d\Delta\nu M^n + 4kT_e\Delta\nu/R_L} \tag{11.8-2}$$

The advantage of using an avalanche photodiode over an ordinary photodiode is now apparent. When $M = 1$, the situation is identical to that at the photodiode as described by (11.7-15). Under these conditions the thermal term $4kT_e\Delta\nu/R_L$ in the denominator of (11.8-2) is typically much larger than

[17] A theoretical study by McIntyre [17] predicts that if the multiplication is due to either holes or electrons, $n = 2$, whereas if both carriers are equally effective in producing electron–hole pairs, $n = 3$.

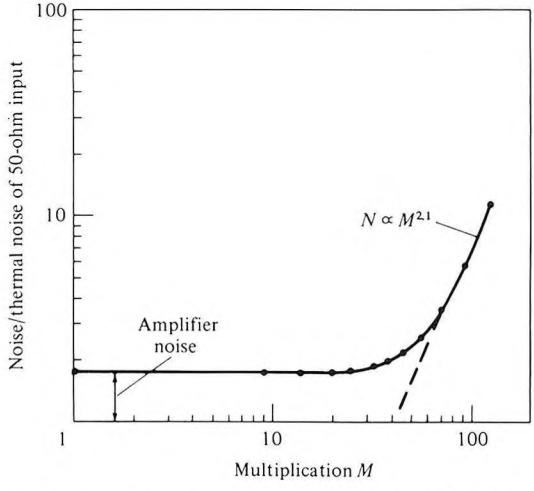


Figure 11-25 Noise power (measured at 30 MHz) as a function of photocurrent multiplication for an avalanche Schottky-barrier photodiode. (After Reference [18].)

the shot-noise terms. This causes S/N to increase with M. This improvement continues until the shot-noise terms become comparable with $4kT_e\Delta\nu/R_L$. Further increases in M result in a reduction of S/N since $n > 2$, and the denominator of (11.8-2) grows faster than the numerator. If we assume that M is adjusted optimally so that the denominator of (11.8-2) is equal to twice the thermal term $4kT_e\Delta\nu/R_L$, we can solve for the minimum detectable power (that is, the power input for which $S/N = 1$) obtaining

$$P_{\min} = \frac{2h\nu}{M'e\eta}\sqrt{\frac{kT_e\Delta\nu}{R_L}} \tag{11.8-3}$$

where M' is the optimum value of M as discussed previously. The improvement in sensitivity over the photodiode result (11.7-16) is thus approximately M'. Values of M' between 30 and 100 are commonly employed, so the use of avalanche photodiodes affords considerable improvement in sensitivity over that available from photodiodes.

11.9 POWER FLUCTUATION NOISE IN LASERS

The power output from lasers is ever fluctuating. This fluctuation may be due to temperature variations, acoustic vibrations, and other man-made causes. Even if all of these extraneous effects are eliminated, there remains a basic (quantum mechanical) contribution that is due to spontaneous emission of radiation into the laser mode by atoms dropping from the upper transition level into the lower levels. The field due to this spontaneous emission is not coherent with that of the laser mode, thus causing phase and amplitude fluctuations [1]. Since these fluctuations are random, they are described and quantified in terms of the statistical noise tools developed earlier in this chapter and in Chapter 10.

Let the power output of the laser be

$$P(t) = P_0 + \Delta P(t) \tag{11.9-1}$$

where the time-averaging value of the fluctuation is zero.

$$\overline{\Delta P(t)} = 0$$

so that P_0 is the average optical power. Using (10.2-6) and (10.2-8) we characterize the "power" of the fluctuation via the mean of the squared deviation[18]

$$\overline{(P(t) - P_0)^2} \equiv \overline{(\Delta P(t))^2} = \int_0^\infty S_{\Delta P}(f)\,df \tag{11.9-2}$$

$S_{\Delta P}(f)$ is related to the spectral density function $S_{\Delta P}(\omega)$, defined by (10.2-8) and (10.2-14), by

$$S_{\Delta P}(f) = 2\pi S_{\Delta P}(\omega) \qquad (\omega = 2\pi f) \tag{11.9-3}$$

[18]In this section, we will use f to denote "low" (rf) frequencies and ν for optical frequencies.

If an optical field at frequency ν with a power $P(t)$ is incident on a detector whose quantum efficiency (electrons per photon) is η, the output current is

$$i(t) = \frac{e\eta P(t)}{h\nu}$$

so that according to Equation (11.9-1) the optical power fluctuation $\Delta P(t)$ causes a fluctuating current component $\Delta i(t) = e\eta\Delta P(t)/h\nu$ with a mean square

$$\overline{i_{NL}^2}(t) \equiv \overline{[\Delta i(t)]^2} = \frac{e^2\eta^2}{(h\nu)^2}\,\overline{[\Delta P(t)]^2} = \frac{e^2\eta^2}{(h\nu)^2}\,S_{\Delta P}(f)\Delta f \qquad (11.9\text{-}4)$$

where Δf is the bandwidth of the electronic detection circuit.

The relative intensity noise (RIN), is defined as the relative fluctuation "power" in a $\Delta f = 1\text{Hz}$ bandwidth

$$\text{RIN} \equiv \frac{S_{\Delta P}\Delta f(=1\text{Hz})}{P_0^2} \qquad (11.9\text{-}5)$$

A single-mode semiconductor laser might possess a value of RIN $\approx 10^{-16}$ (or -160 db). Assuming that the detector circuit has a bandwidth of, say, $\Delta f = 10^9$ Hz, the relative mean-squared fluctuation in the detected current is

$$\frac{\overline{(\Delta i_d)^2}}{i_{d0}^2} = \frac{\overline{(\Delta P)^2}}{P_0^2} = \frac{S_{\Delta P}(f)\,\Delta f}{P_0^2} = 10^{-16} \times 10^9 = 10^{-7}$$

The RMS value of the power fluctuation is thus

$$\frac{\{\overline{[\Delta P(t)]^2}\}^{1/2}}{P_0} = 3.16 \times 10^{-4}$$

The mean-squared noise current in the output of the detector due to these fluctuations is given by (11.9-4)

$$\overline{i_{NL}^2}(t) = \frac{e^2\eta^2}{(h\nu)^2}\,(\text{RIN})P_0^2\Delta f \qquad (11.9\text{-}6)$$

Assuming as an example that $\lambda = 1.3\ \mu$m, $P_0 = 3$ mW, RIN $= 10^{-16}\ \text{Hz}^{-1}$, $\Delta f = 10^9$ Hz, and $\eta = 0.6$, we obtain

$$(\overline{i_{NL}^2})^{1/2} = 5.95 \times 10^{-7}\ \text{A}$$

Example: Optical Fiber Link Design

Our task here is to determine the maximum allowed repeater spacing for an optical fiber communication link. We will assume that the optical source is a 1.3 μm GaInAsP laser ($\nu = c/\lambda = 2.31 \times 10^{14}$ Hz) and that the fiber

possesses an attentuation of 0.3 db/km (corresponding to an attenuation constant $\alpha = 0.3/4.343 = 0.0691$ (km)$^{-1}$). The optical power launched into the fiber is $P_0 = 3$ mW. The channel is to transmit 10^9 bits/s so that the bandwidth of the detector circuit is taken as $\Delta f = 1/\text{period} = 10^9$ Hz. The system considerations dictate that the bit error probability at the detector output not exceed 10^{-10}. The detector output impedance is $R_L = 1{,}000\ \Omega$, and the amplifier (following the detector) noise figure is 6 db, i.e., $F = 4$ (see footnote 15).

From Figure 10-20 we determine that the signal-to-noise power ratio at the amplifier output must exceed 22 db to assure a bit error probability upon detection that is smaller than 10^{-10}. Our task is thus to calculate the signal power $\overline{i_s^2}$ and the total noise power $\overline{i_N^2}$ at the output of the detector as a function of the length L of the link.

The signal power is obtained from (11.7-11), assuming a modulation index $m = 0.5$, as

$$\overline{i_s^2} = \frac{e^2\eta^2 P_0^2 e^{-2\alpha L}}{2(h\nu)^2} \tag{11.9-7}$$

The total noise power at the output of the amplifier referred to its input is

$$\overline{i_N^2} = \overline{i_{NL}^2} + \overline{i_{NS}^2} + \overline{i_{NA}^2} \underset{\text{(in order)}}{=} \frac{\eta^2 e^2}{(h\nu)^2}(\text{RIN})P_0^2 e^{-2\alpha L}\Delta f + \frac{2\eta e^2}{h\nu} P_0 e^{-\alpha L}\Delta f + \frac{4kT_e}{R_L}\Delta f \tag{11.9-8}$$

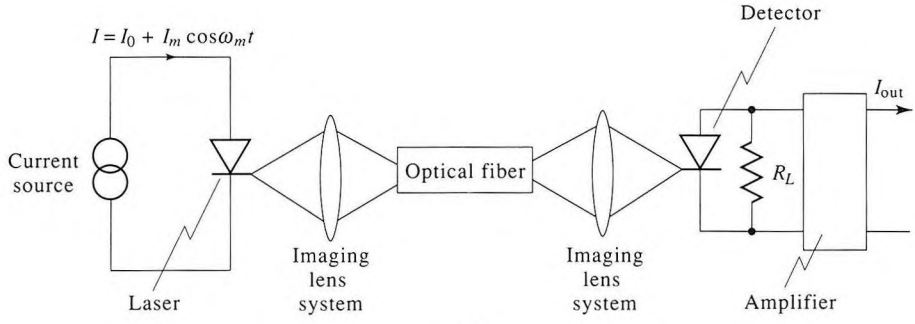
The first noise term is that due to power fluctuation (11.9-6); the second is the shot noise associated with the average current at the output of the detector $I_{d0} = \eta P_0 e \exp(-\alpha L)/(h\nu)$. The third term represents, as in (11.7-15), both the Johnson noise of the output resistor R_L as well as the amplifier output noise power (referred to its input, see footnote 15). If the temperature of the output resistor R_L is $T = 290$ K, $T_E = T + (F - 1)290 = 1160$ K.

Figure 11-26(a) shows the main elements of an optical fiber link. Figure 11-26(b) shows a plot of $\overline{i_s^2}$, $\overline{i_{NL}^2}$, $\overline{i_{NS}^2}$, and $\overline{i_{NA}^2}$ as well as the total noise power

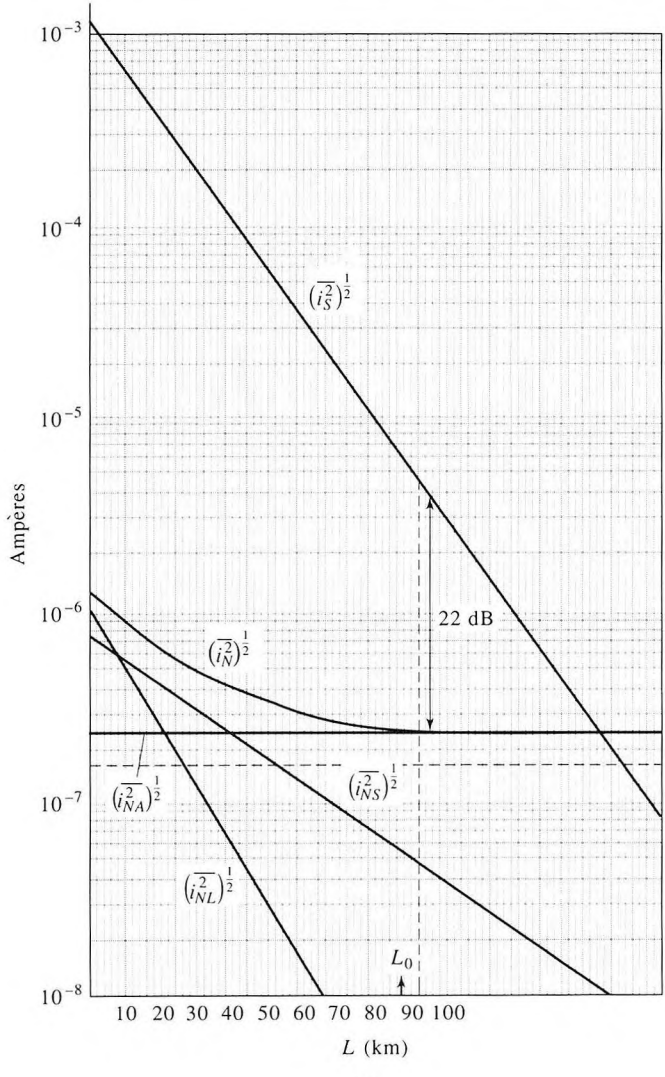
Figure 11-26 (a) An optical fiber communication link consisting of a laser, an optical coupling system c, a fiber L/(km long), a detector D, an output resistance R_L and an amplifier A with a current gain G and a noise figure F. (b) The signal ($\overline{i_s^2}$) laser fluctuation ($\overline{i_{NL}^2}$) detector shot noise ($\overline{i_{NS}^2}$), combined Johnson-amplifier noise ($\overline{i_{NA}^2}$), and the total noise

$$\overline{i_N^2} = \overline{i_{NL}^2} + \overline{i_{NS}^2} + \overline{i_{NA}^2}$$

currents as a function of the link length L. The currents are referred to the amplifier input plane S, i.e., they correspond to output currents divided by the current gain G of the output amplifier.


$I = I_0 + I_m \cos\omega_m t$
Current source
Laser
Imaging lens system
Optical fiber
Imaging lens system
Detector
I_{out}
R_L
Amplifier

(a)


10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
Ampères
$(\overline{i_S^2})^{\frac{1}{2}}$
22 dB
$(\overline{i_N^2})^{\frac{1}{2}}$
$(\overline{i_{NA}^2})^{\frac{1}{2}}$
$(\overline{i_{NS}^2})^{\frac{1}{2}}$
$(\overline{i_{NL}^2})^{\frac{1}{2}}$
L_0
10 20 30 40 50 60 70 80 90 100
L (km)

(b)

as a function of the link length L. The important thing to note is the relative change of the various powers with distance. The distance L_0, where the detected signal-to-noise power ratio is down to 22 db, is read off as $L_0 = 87$ km.[19] This distance is thus chosen as the link length. Notice, as an example, that the dominant noise contribution at $L > 33$ km is the amplifier-detector noise $\overline{i_{NA}^2}$. If the latter were reduced by, say, 3 db, the link length could be increased by 5 km, as indicated by the dashed line.

The signal-to-noise power ratio of a *p-n* diode detector is given by (11.7-16) in the case where the dominant contributions to the noise power are the amplifier noise and the Johnson (thermal) noise of the load resistance R_L in the diode output circuit. The mean-square noise current is then

$$\overline{i_N^2} \approx \frac{4kT_e\Delta f}{R_L} \tag{11.9-9}$$

The signal peak current is given by (11.7-8) as

$$i_S = \frac{P_S e\eta}{h\nu} \tag{11.9-10}$$

where P_s is the peak pulsed optical power incident on the detector. The signal-to-noise current ratio at the amplifier output (see footnote 15) is thus

$$\frac{i_S}{(\overline{i_N^2})^{1/2}} = \frac{P_s e\eta/h\nu}{(4kT_e\Delta f/R_L)^{1/2}} \tag{11.9-11}$$

Our next problem is that of finding the minimum value of the signal power P_s so that $i_S/(\overline{i_N^2})^{1/2}$ in (11.9-11) exceeds the needed value of 12.59. We thus need to know T_e, R_L, and Δf. T_e is obtained from the given value of the amplifier noise figure ($F = 6$ dB). Taking $T = 290$ K, we obtain, using footnote 15, $T_e = 290 + (4 - 1)290 = 1160$. In order to achieve this bandwidth, the load resistance R_L must not exceed (see 11.7-18) the value

$$R_L = \frac{1}{2\pi\Delta f C} \tag{11.9-12}$$

where C is the total output capacitance given as $3 \times 10^{-12} f$. Using the above value of Δf and C, we obtain

$$R_L \leq 53 \text{ ohms}$$

We return now to (11.9-11), which, using $\eta = 0.5$, $\lambda = 1.35$ μm, $i_S/(\overline{i_N^2})^{1/2} = 12.59$, yields

$$P_S \cong 2.41 \times 10^{-7} \text{ watt}$$

for the minimum power input to the photodiode.

The total transmission loss in the 50 km fiber is 20 dB. We will assume

[19] That is, $10 \log (\overline{i_s^2}/\overline{i_N^2}) = 22$.

that an additional 4 dB loss is caused by coupling the laser output to the fiber and at the fiber output so that the total loss is 24 dB (that is, 251). The laser power output must thus exceed

$$P_{\text{laser}} = 6 \times 10^{-3} \text{ watt}$$

which is a reasonable power level for CW diode lasers.

If the fiber had been substantially lossier than in the above example, we could still have met our design specifications by using an avalanche photodiode.

11.10 INFRARED IMAGING AND BACKGROUND-LIMITED DETECTION [25–28]

Arrays of cooled infrared detectors based mostly on photoconductive semiconductors such as mercury cadmium telluride (HgCdTe) have become increasingly important elements in the fast developing technology of infrared imaging and detection. The application areas served by this new technology include tumor detection, the mapping of earth resources by orbiting satellites, "spy" satellites, and nighttime "seeing." We will not concern ourselves here with the system aspects of these applications but rather with the basic noise physics of a single element that is prerequisite to system considerations. The concepts involved here are the same as those we have encountered in the early sections of this chapter, but the operational considerations merit a dedicated treatment. To be specific, we will focus our discussion to doped, say, *n*-type, photoconductors, such as HgCdTe, in which the optical input field causes excitation of electrons to the conduction band so that the signal current is due to the drift of the excited carriers.

Consider the photoconductive detector shown in Figure 11-27. The detected radiation is incident on the "face" whose area is A. The thickness of the detector is t. The optical signal power input to the detector is P_s. This power gives rise to a signal current given according to (11.5-2) by

$$i_s = \frac{\eta P_s e}{h\nu_{\text{opt}}} \left(\frac{\tau_0}{\tau_d}\right) \tag{11.10-1}$$

where τ_0 is the lifetime of the photo-excited carriers, $\tau_d = \frac{t}{\bar{v}}$ is the transit time for a carrier with a velocity $\bar{v}$. ν_{opt} is the frequency of the optical beam. η is the fraction of the incoming photons that are usefully absorbed in the photoconductor. τ_0 is the average lifetime of a photo-excited carrier. In addition to the signal current, we have two other major sources of noise currents that are not related to the signal.

The first is the shot noise associated with the drift under the influence of the applied external field of thermally excited (minority) carriers, while the second is the shot noise due to carriers excited by the ever present incoming background optical radiation. Referring to Figure 11-27(b), we

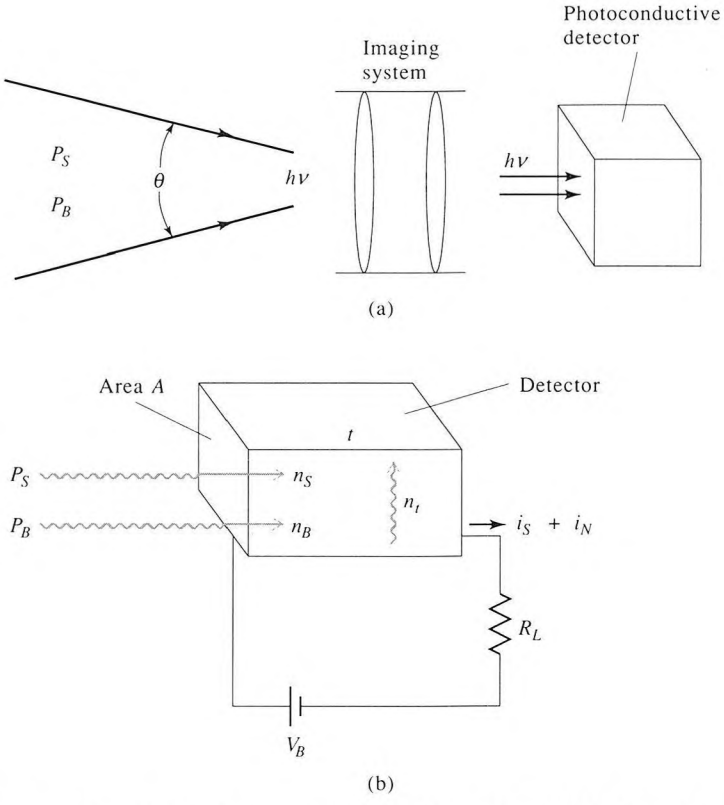


Figure 11-27 (a) A schematic diagram of an infrared detecting element intercepting radiation from an acceptance angle θ. (b) A detailed view of the photoconductor.

distinguish in our photoconductor three populations of carriers: n_s (cm^{-3}) due to the signal, n_t due to thermal excitation, and n_B due to incoming background radiation. The density n_B of carriers excited by the background radiation is

$$n_B = \frac{(P_B/A)\eta\tau_0}{h\nu_{\text{opt}}t} \tag{11.10-2}$$

This last expression can be obtained by setting the rate of photon absorption per unit volume $\dfrac{(P_B\eta)}{(h\nu_{\text{opt}}At)}$ equal to the rate $\dfrac{n_B}{\tau_0}$ of minority carrier recombination. The incident background power P_B is most often that of the background blackbody radiation, in which case

$$\mathcal{J}_B \equiv \frac{P_B}{A} = \frac{2\pi h\nu_{\text{opt}}^3\Delta\nu(\sin^2\theta/2)}{C^2(e^{h\nu_{\text{opt}}/kT_B}-1)} \tag{11.10-3}$$

where A is the cross-sectional area of the detector, $\Delta\nu$ is the optical bandwidth of the radiation allowed into the detector, $-T_B$ is the background temperature, while θ is the angle [see Figure 11-27(a)] within which radiation is accepted by the detector. In practice the acceptance angle θ and the background temperature T are dictated by the application. A reasonable strategy in such a case is to cool the detector to the point where $n_t < n_B$. This renders the contribution of n_t to the shot noise equal to that of n_B so that additional cooling will not materially improve the signal-to-noise ratio at the output. Since the signal-to-noise ratio under this condition is determined by the background radiation, it is referred to as background limited infrared performance (BLIP).

Let us assume that a BLIP condition has been achieved and calculated the resulting detector performance. From (11.5-9) the mean-squared output noise current is (in the limit $\nu\tau_0 \ll 1$)

$$\overline{i_{NB}^2} = 4e\bar{I}_B\left(\frac{\tau_0}{\tau_d}\right)\Delta f$$

$$= 4e(n_B e\bar{v}A)\frac{\tau_0}{\tau_d}\Delta f \qquad \text{(11.10-4)}$$

Where $\bar{I}_B = n_B e\bar{v}A$ is the average current due to the background radiation–excited carriers and Δf is the bandwidth of the (electronic) detection circuit. Substituting for n_B from (11.10-2) gives

$$\overline{i_{NB}^2} = \frac{4e^2P_B\eta}{h\nu_{\text{opt}}}\left(\frac{\tau_0}{\tau_d}\right)^2\Delta f \qquad \text{(11.10-5)}$$

The minimum detectable signal, also known as the noise equivalent power (NEP) of the detector, is that value of the signal power P_s for which

$$\overline{i_s^2} = \overline{i_N^2}$$

Using (11.10-1) and (11.10-5) we can solve for the minimum detectable power when the main noise contribution is due to the background radiation, i.e., $n_t < n_B$

$$(\text{NEP})_B = P_s(\overline{i_s^2} = \overline{i_{NB}^2}) = 2\sqrt{\frac{A\mathcal{J}_B h\nu_{\text{opt}}\Delta f}{\eta}} \qquad \text{(11.10-6)}$$

A common figure of merit used in the infrared imaging community to describe detector sensitivity is the specific peak detectivity D* (''Dee'' star) defined as

$$\text{D}^* = \frac{\sqrt{A\Delta f}}{\text{NEP}}$$

When the detector is cooled sufficiently so that it is background limited, D*

becomes

$$D_B^* = \frac{\sqrt{A\Delta f}}{(NEP)_B} = \frac{1}{2}\sqrt{\frac{\eta}{h\nu_{opt}\mathcal{P}_B}} \tag{11.10-6a}$$

where, to remind us, the B subscript stands for the background limited condition and Δf is bandwidth of the detection circuit including the photoconductive element.

In a detector limited by thermal excitation of carriers, i.e., one where $n_t > n_B$ we have

$$\overline{I}_t = n_t e \bar{v} A \tag{11.10-7}$$

$$\overline{i_{Nt}^2} = 4e\overline{I}_t\left(\frac{\tau_0}{\tau_d}\right)\Delta f = 4e(n_t e\bar{v}A)\left(\frac{\tau_0}{\tau_d}\right)\Delta f \tag{11.10-8}$$

Equating $\overline{i_{Nt}^2}$ to $\overline{i_s^2}$ as in (11.10-6), we obtain

$$(NEP)_t = \frac{2h\nu_{opt}}{\eta}\sqrt{\left(\frac{\tau_d}{\tau_0}\right) n_t \bar{v} A \Delta f}$$

and

$$D_t^* = \frac{\sqrt{A\Delta f}}{(NEP)_t} = \frac{\eta}{2h\nu_{opt}}\sqrt{\frac{\tau_0/\tau_d}{n_t\bar{v}}} = \frac{\eta}{2h\nu_{opt}}\sqrt{\frac{\tau_0}{n_t t}} \tag{11.10-9}$$

where, in the last expression, we used $\tau_d \equiv t/\bar{v}$.

It is obvious that the condition $n_t \leq n_B$ is equivalent to

$$(NEP)_t \leq (NEP)_B \quad \text{or} \quad D_t^* \geq D_B^* \tag{11.10-10}$$

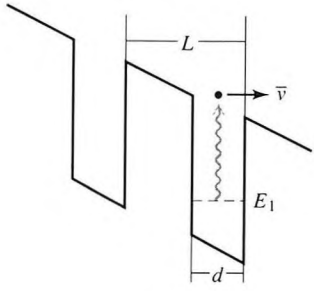
A key issue in infrared detection is to determine to what temperature a detector element need be cooled to be background limited. To answer this question we need to know n_B and the dependence of n_t on the material parameters and the temperature.

As an example consider an infrared detector as shown in Figure 11-28(a) in which the photoconductive medium is a GaAs/GaAlAs superlattice [29]. It is based on excitation of electrons from a confined "quantum-well" state (see Section 16.1) to continuum (unconfined) states where they are free to conduct. These wells consist of thin layers (~100Å) of crystalline GaAs layers sandwiched between higher energy gap $Ga_{1-x}Al_xAs$ crystalline layers.

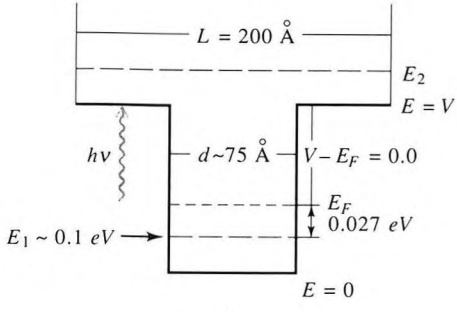
In this case [25],

$$n_t = \frac{m^* kT}{\pi\hbar^2 L} e^{-(V-E_F)/kT} \tag{11.10-11}$$

where L is the width of the unit cell in Figure 11-28(a), m^* is the carrier (electron) effective mass, E_F the Fermi energy of the material, and V the depth of the quantum well. Using the data of Figure 11-28(b) and $m^* =$



(a)



(b)

Figure 11-28 Schematic illustration of a typical quantum well detector structure (only two wells are shown) with the relavant energies. In the upper part of the figure (a), the quantum well detector under bias is shown. The second part of the figure (b) shows the relevant distances and energies. The sub-band levels are given by E_1 and E_2.

$0.067m_e$, $\frac{\tau_0}{\tau_d} = 0.5$, $\lambda = \frac{c}{\nu_{\text{opt}}} = 10\ \mu$m, we can use (11.10-9) to plot D_t^* vs. the detector temperature T. The result is shown in Figure 11-29(a).

To use this curve we need first to obtain a value for D_B^* using (11.10-3) and (11.10-6a). We then find the temperature in Figure 11-29(a) where $D_t^* = D_B^*$. As an example, given $D_B^* = 10^{12}$ we find from the figure that the quantum-well detector needs to be cooled to $T < 48$ K in order to become background limited.[20] Figure 11-29 curve (b) is a plot based on Equations (11.10-3, 11.10-6, 11.10-9) of the acceptance angle θ_B for which D_B^* is equal to D_t^*. Further increases of θ_B will thus cause the detector to be background limited. In the figure we used a background temperature $T_B = 300$ K,

[20]In practice cooling the detector below the liquid N_2 temperature, 77.7°K, is expensive and is reserved to very demanding applications.

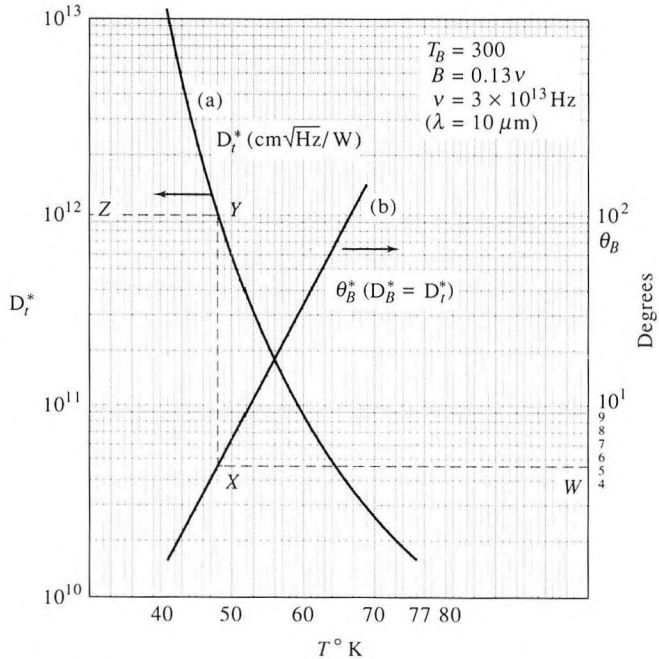


Figure 11-29 Curve (a) is a plot of the thermal-excitation-limited detectivity D_t^* at $\lambda = 10\ \mu m$ vs. temperature of the GaAs quantum well detector. Curve (b) is a plot of the acceptance angle θ_B which results in $D_B^* = D_t^*$ (BLIP) for the background conditions stated in the figure.

$\nu_{opt} = 3 \times 10^{13}$ ($\lambda = 10\ \mu m$), $\Delta\nu = 0.13\nu_{opt}$. As an example, if our detector has a D_t^* of 10^{12}, then the BLIP condition obtains when $\theta_B = 4.6°$.

We can also use the curves of Figure 11-29 in reverse to find D_B^* for a given θ_B. For instance, for $\theta_B = 4.6°$ we follow the sequence $W \rightarrow X \rightarrow Y \rightarrow Z$ to find $D_B^* = 10^{12}$; in the process we also learn that the detector temperature must be below 48 K in order to be background limited.

The most widely used material for infrared imaging near $\lambda = 10\ \mu m$ is the semiconductor HgCdTe [28] (MCT) whose composition can be adjusted to yield an energy of a photon with $\lambda \sim 10\ \mu m$. The photoconduction in this case is due to excitation across the energy gap of the semiconductor. Typical MCT photoconductive detectors use N-type material so that the carriers responsible for the signal (and noise) are the (minority) holes. Our theoretical discussion up to this point applies if we merely take n_B, n_t, and n_s, respectively, as the density of holes excited by the background radiation, thermal process, and the "signal" radiation.

The most important task that confronts the infrared detector scientist is to develop materials that enable background-limited performance at the *highest possible temperature*. The background limit condition $n_t = n_B$ can be

written using (11.10-2) as

$$\frac{\mathcal{J}_B \eta \tau_0}{h\nu_{\text{opt}} t} = n_t \tag{11.10-12}$$

Since n_t invariably increases with T, it follows from (11.10-12) that *the temperature for background-limited operation increases with the carrier lifetime* τ_0. In the example given above, the temperature T for background-limited detection condition $n_t = n_B$ is given, according to (11.10-11), by the condition $n_B = n_t$, i.e., by the value of T satisfying

$$\frac{\mathcal{J}_B \eta \tau_0}{h\nu_{\text{opt}} t} = \frac{m^* kT}{\pi \hbar^2 L} e^{-(V - E_F)/kT} \tag{11.10-13}$$

In a typical HgCdTe at 77 K, the carrier lifetime is $\tau_0 \sim 10^{-6}$ s while in our quantum well detector $\tau_0 \sim 10^{-11}$ s [the time for an excited carrier to drop in energy below the top of the well thus becoming immobile (trapped). This happens after the emission of only a few optical-branch phonons by the excited carrier.]

It follows that HgCdTe is background limited and thus has an NEP, described by (11.10-6), at a higher temperature than a GaAs/GaAlAs detector used in the above example. To illustrate this point, we show in Figure 11-30 a plot of the thermal generation current that is the rate of decay (per unit of incidence area) of thermally excited carriers[21]

$$I_{t-g} = \frac{n_t t}{\tau_0} \text{ [photons/(cm}^2 - \text{s)]} \tag{11.10-14}$$

where n_t is the density of thermally excited carriers (for HgCdTe it is the minority carrier density [26–28]). I_{t-g} is commonly used in the infrared imaging community to compare different materials since at the background limit (BLIP) it is equal to the rate $\eta \mathcal{J}_B / h\nu_{\text{opt}}$ of (absorbed) background photons (per cm^2) incident on the detector. Since the latter rate is determined by system considerations (see Equation 11.10-3 and the following discussion), given the background absorbed photon flux $\eta \mathcal{J}_B / h\nu_{\text{opt}}$ we can determine at a glance the temperature to which our detector needs to be cooled to achieve BLIP condition or, equivalently, the temperature to which the detector needs to be cooled for BLIP operation at a given background photon flux.

We note by comparing Figure 11-30(a) to 11-30(b) that at given T, I_{t-g} in HgCdTe is ~6 orders of magnitudes smaller than in the GaAs/GaAlAs detector, reflecting mainly the difference in carrier lifetime τ_0. It is $\sim 10^{-11}$ s in GaAs/GaAlAs [25] and $\sim 10^{-6}$ s in HgCdTe. As an example consider a system subject to a background photon flux near $\lambda = 10$ μm of 3 ×

[21]This rate is equal, at thermal equilibrium, to the rate at which the "thermal" carriers are generated.

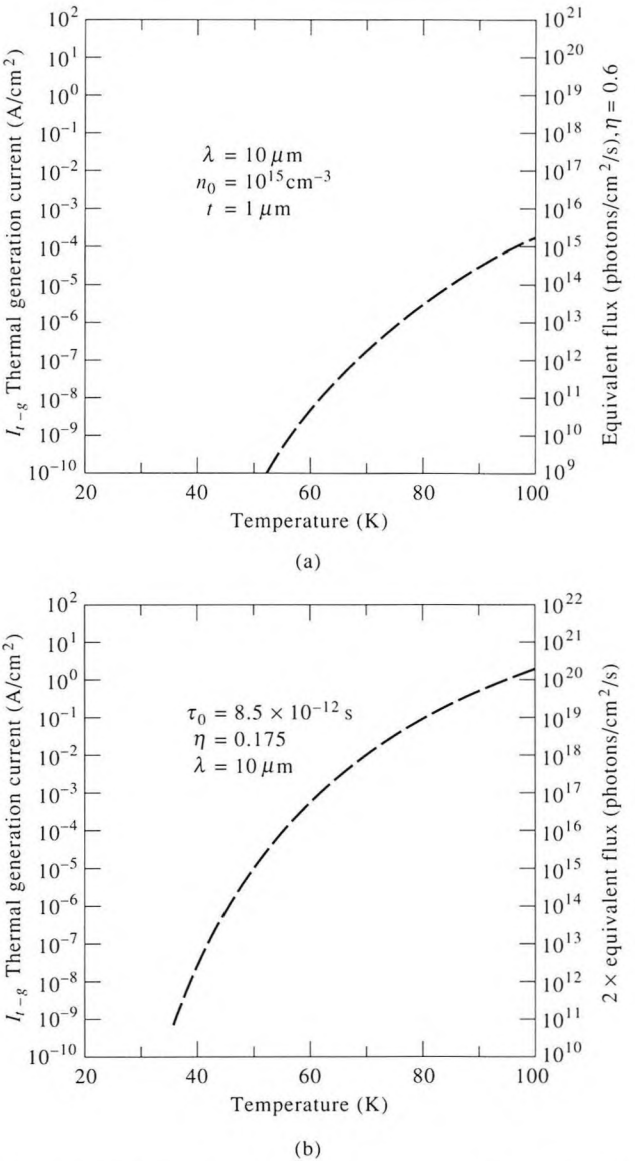


Figure 11-30 (a) Thermal equivalent current for HgCdTe IR devices. On the left ordinate it is expressed in A/cm^2 while on the right ordinate it is expressed in the equivalent arrival rate of (10 μm) photons/cm^2. (b) Thermal equivalent generation current for GaAs/AlGaAs.

10^{13} cm^{-2} – s^{-1}. The HgCdTe detector becomes, according to Figure 11-30(a), background limited at ~80 K while the GaAs/GaAlAs detector [Figure (b)] needs to be cooled to ~45 K. If we calculate the D_t^* corresponding to this incoming background flux, we obtain using, for example, Figure 11-29(a) at 45 K, $D_t^* = 2.2 \times 10^{13}$ cm-Hz$^{1/2}$-W^{-1}.

11.11 OPTICAL AMPLIFICATION IN FIBER LINKS

Optical amplification in fiber links [31–33] has recently been recognized as having major system implications for very long distance transmission of information (>1000 km) using optical fibers and for distribution systems involving a large number of subscribers. These purely optical repeaters may, in most cases, obviate the need for the repeater stations currently used that involve detection, electronic amplification, and remodulation of a (new) launched optical beam.

The raison d'être for the optical amplifiers is that they make it possible to maintain the optical power at sufficiently high levels along the path so that the signal-to-noise ratio (SNR) degradation due to signal shot noise and receiver noise is reduced to practical inconsequence.

A new and dominant noise source, amplified spontaneous emission, however, is introduced by the optical amplifier [35], and its effect on the signal-to-noise ratio (SNR) of the detected signal current will be considered below. Before doing so, we will briefly review the relevant physics of the amplifier.

The most common amplifier uses a transition at $\lambda = 1.535\ \mu$m in an E_r^{3+} ion introduced as a dopant into the present in a silica fiber [34, 35]. The pertinent energy levels are shown in Figure 11-31(a).

The laser transition can be pumped by radiation at $\lambda \sim 0.98\ \mu$m or $\lambda \sim 1.49\ \mu$m as shown. This pumping field is usually obtained from semiconductor lasers and is coupled into the amplifying fiber whose length is typically between a few meters and a few tens of meters. A schematic diagram of the amplifier configuration is shown in Figure 11-31(b). The fiber amplifier section can be spliced smoothly into the fiber transition. A plot of the gain vs. signal wavelength is shown in Figure 11-32.

The main effect of the optical amplifier on the SNR of the detected signal is to add, upon detection, a noise current component, at frequencies near that of the signal current. This noise is due to beating between the amplified (optical) spontaneous emission (ASE) power of the amplifier and the signal optical field. We first need to obtain an expression for the optical spontaneous emission power at the output of an optical amplifier. This topic is the subject of Appendix D. The main result, Equation D-8, is that the (amplified) spontaneous emission power in a *single* mode with a spectral bandwidth $\Delta\nu_{\text{opt}}$ at the output of an optical amplifier is [36]

$$F_0 = \mu h\nu\, \Delta\nu_{\text{opt}}(G - 1) \tag{11.11-1}$$

where G is the power gain of the optical amplifier and

$$\mu = \frac{N_2}{N_2 - N_1\dfrac{g_2}{g_1}} \tag{11.11-2}$$

is the atomic inversion factor of the transition. It accounts for the larger

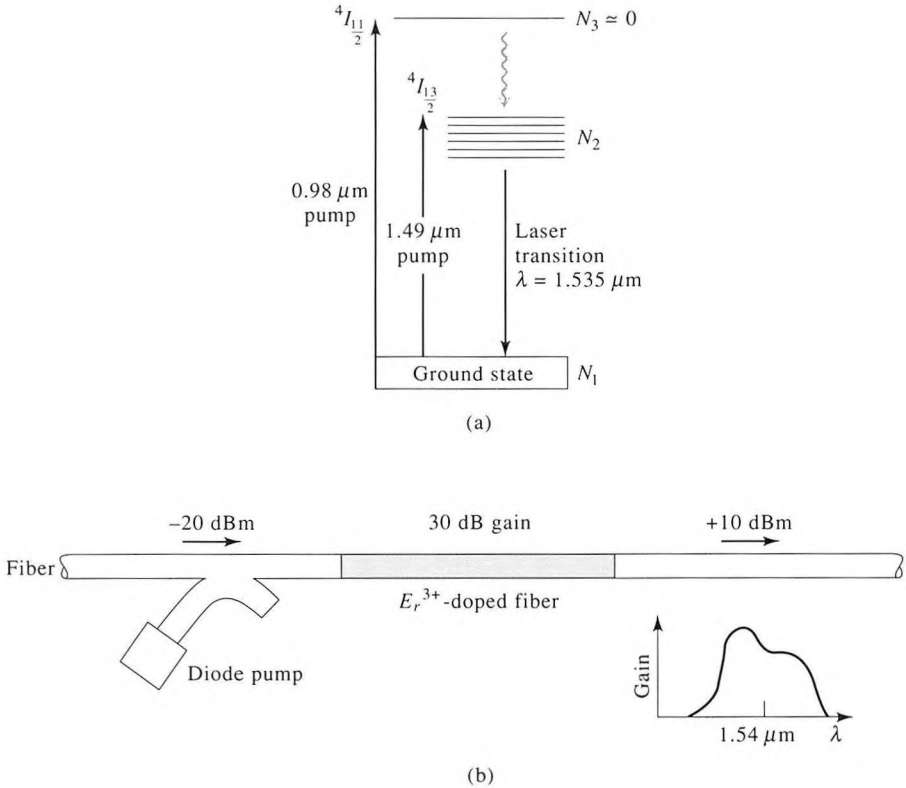


Figure 11-31 (a) The pertinent energy level diagram of E_r^{3+} in silica. For pumping at $\lambda = .98\ \mu$m (preferred). (b) A schematic diagram showing the amplifying fiber spliced into the transmission fiber and the method for coupling the pump radiation into the fiber.

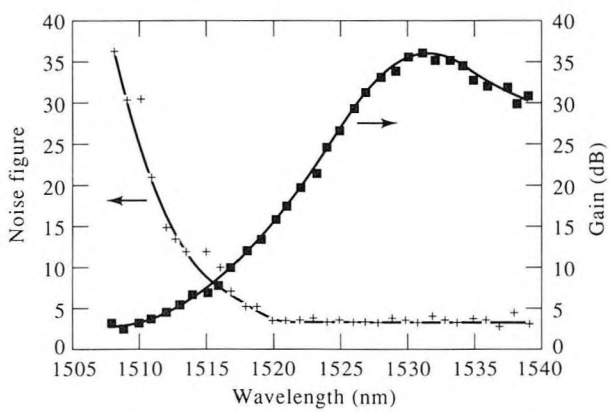


Figure 11-32 Noise factor and gain spectrum of the silica E_r^{3+} fiber amplifier for a constant pump power of 34.2 mW at 0.98 μm. (After Reference [35].)

value of N_2, and hence larger spontaneous emission power, in atomic (amplifier) systems in which $N_1 \neq 0$.[22]

If we denote the optical power at ω_2 as S and that of the spontaneous emission at ω_1 as F, then the beat current component with a frequency $\omega_m = \omega_2 - \omega_1$ is (see Equation 11.4-4)

$$i = \frac{Se\eta}{h\nu}\left[1 + \frac{F}{S} + 2\sqrt{\frac{F}{S}}\cos(\omega_m t + \Phi_{ASE} - \Phi_s)\right] \tag{11.11-3}$$

where ϕ_{ASE} and ϕ_S are the phases of the ASE field and the signal optical field, respectively. The mean-squared beat current is then

$$(\overline{i^2})_{\text{ASE-signal}} = 2FS\left(\frac{e\eta}{h\nu}\right)^2 \tag{11.11-4}$$

which, using (11.11-1) and putting $\Delta\nu_{opt} = 2\Delta f_{sig}$, yields[23]

$$(\overline{i^2})_{\text{ASE-signal}} = \frac{4e^2\eta^2 S(G-1)\mu\Delta f_{sig}}{h\nu} \tag{11.11-5}$$

In the remainder we will drop the subscript "sig" and use Δf only.

Consider an optical in-line amplifier as shown in Figure 11-33. The input signal power is S_0, and it enters the amplifier in a *single* transverse (usually the fundamental) fiber mode. The amplified output signal is GS_0, while F_0, as given by (11.11-1), represents the (optical) amplified spontaneous emission power at the output, which is generated within the amplifier in a band $\Delta\nu$. If we were to detect the signal at the input to the amplifier, the main noise contribution would, in an ideal case, i.e., a noiseless receiver, be that of the signal shot noise so that the signal-to-noise power ratio (SNR) at the input

[22] In a laser the gain per pass is given by $G = \exp[a(N_2 - N_1)L_{amp}]$ where L_{amp} is the length and a is a constant depending on the atoms. A large N_1 thus causes a larger N_2 for a given gain. The SE power is proportional to N_2.

[23] Two ASE frequency bands, each with a width $\Delta\nu_{sig}$, one above and one below the signal frequency contribute incoherently to the beat power so that the effective $\Delta\nu_{opt} = 2\Delta f_{sig}$.

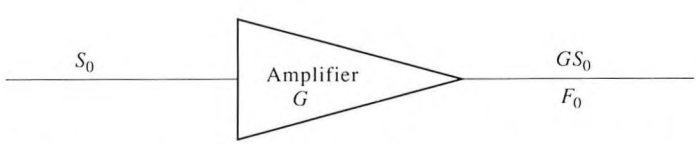


Figure 11-33 An optical amplifier with a power gain G and an input signal power S_0. F is the total power of the amplified spontaneous emission (ASE) at the output of the amplifier in the appropriate bandwidth $\Delta\nu$.

to the amplifier is

$$\mathrm{SNR}_{\mathrm{in}} = \frac{\left(\frac{S_0 e}{h\nu}\right)^2}{2e^2 \frac{S_0}{h\nu} \Delta f} = \frac{S_0}{2h\nu\Delta f} \tag{11.11-6}$$

The detected signal "power"[24] at the output is

$$(\overline{i^2})_{\mathrm{out}} = \left(\frac{GS_0 e}{h\nu}\right)^2 \tag{11.11-7}$$

while the noise power is that of the ASE-signal noise (11.11-5) and the shot noise

$$(\overline{i^2_{\mathrm{shot}}})_{\mathrm{out}} = \frac{2e^2 G S_0}{h\nu} \Delta f \tag{11.11-8}$$

The noise current component that is due to beating of ASE frequencies with themselves is proportional to F_0^2 and can be made to be negligible compared to the ASE-signal current if the signal power $S(z)$ is not allowed to drop too far and/or by optical filtering. We have neglected for similar reasons the shot noise due to the ASE. The (S/N) ratio at the output of the amplifier is thus

$$\mathrm{SNR}_{\mathrm{out}} = \frac{\left(\frac{GS_0 e}{h\nu}\right)^2}{\frac{2e^2 G S_0}{h\nu} \Delta f + \frac{4e^2 G(G-1) S_0 \mu \Delta f}{h\nu}} \tag{11.11-9}$$

where we assumed a 100 percent detector quantum efficiency. For large gain $G \gg 1$, the second term in the denominator of (11.11-9), dominates, and

$$\mathrm{SNR}_{\mathrm{out}} \approx \frac{S_0}{4\mu h\nu \Delta f} \tag{11.11-10}$$

The ratio of the input (SNR) to the output value is thus

$$\frac{\mathrm{SNR}_{\mathrm{in}}}{\mathrm{SNR}_{\mathrm{out}}} \approx 2\mu$$

which in an ideal, four-level ($N_1 = 0$, $\mu = 1$) amplifier is equal to 2. The single high-gain optical amplifier will thus degrade the SNR of the detected output by a factor of 2 (3 db). We recall that this degradation is tolerated only in order to save the signal from the, far worse, fate of succumbing, in its attenuated state, to the noise of the receiver. An experimental verification of the SNR result is shown in Figure 11-32.

---

[24]The "power" everywhere is taken as the mean square of the current. Since our final results involve only (signal-to-noise) power ratios, this procedure is justified.

In a very long (100 km) fiber link, we will need to amplify the signal a number of times. We will consequently develop in what follows a formalism for treating systematically cascades of amplifiers.

A generalization of the expression (11.11-9) for the SNR of the detected signal at an arbitrary point z along the link is to write

$$\mathrm{SNR}(z) = \frac{\left[\dfrac{eS(z)}{h\nu}\right]^2}{\dfrac{2e^2\mathrm{S}(z)\Delta f}{h\nu} + \dfrac{4e^2F(z)S(z)}{(h\nu)^2} + \dfrac{4kT_e\Delta f}{R}} \qquad \text{(11.11-11)}$$

where the last term in the denominator represents the mean-squared thermal noise current of the receiver (at point z) whose effective noise temperature is T_e. R is the output impedance of the detector including the receiver's input impedance. Equation (11.11-11) neglects, again, the shot noise due to the ASE, the ASE-ASE beat noise, and intensity fluctuation noise of the source laser. If the signal power $S(z)$ can be maintained above a certain level by repeated amplification, we can neglect the receiver noise term. Under these realistic circumstances, the SNR expression (11.11-11) becomes

$$\mathrm{SNR}(z) = \frac{S^2(z)}{2S(z)h\nu\Delta f + 4S(z)F(z)} \qquad \text{(11.11-12)}$$

$S(z)$ is the signal power at z, while $F(z)$ is the total ASE power at z originating in *all* the preceding amplifiers ($z' < z$).

Let us next consider the realistic scenario of a long fiber with amplifiers employed serially at fixed and equal intervals (z_0), as illustrated in Figure (11-34).

The signal power level $S(z)$ at the fiber input and at the output of each amplifier is S_0. The signal is attenuated by a factor of $L \equiv \exp(-\alpha z_0)$ in the distance z_0 between amplifiers and is boosted back up by the gain $G = L^{-1} = e^{\alpha z_0}$ at each amplifier to the initial level S_0. The spontaneous emission power $F(z)$ is attenuated by a factor L between two neighboring amplifiers and increases by an increment of F_0 at the output of each amplifier. We employ Equation 11.11-1 to calculate the SNR of the detected current at the output of the nth amplifier. Assuming $G \gg 1$, the result is

$$\mathrm{SNR}_n = \frac{S_0}{2h\nu\Delta f[1 + 2n\mu(e^{\alpha z_0} - 1)]} \qquad \text{(11.11-13)}$$

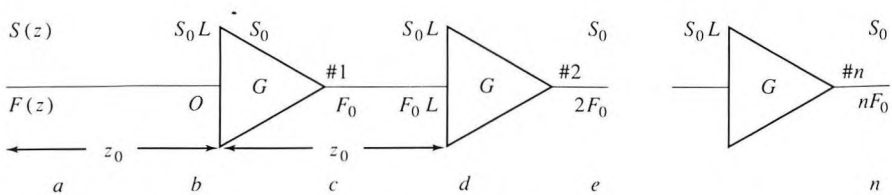


Figure 11-34 A fiber link with periodic amplification.

where, because of the high signal and ASE levels, we neglected the thermal receiver noise. When $\exp(\alpha z_0) = G \gg 1$, we find a z^{-1} (more exactly an n^{-1}) dependence of the SNR rather than the $\exp(-\alpha z)$ dependence of a fiber without amplification in which the main noise mechanism is shot noise. The physical reason for this difference is that the repeated amplification keeps the signal level high as well as the level of the signal-ASE beat noise. The latter is kept well above the signal shot noise. A fixed amount of beat noise power is thus added at each stage leading to the inverse distance dependence of the SNR.

Equation (11.11-13) suggests that the SNR at z can be improved by reducing z_0, i.e., by using smaller intervals between the amplifiers which, of course, entails reducing the gain $G = \exp(\alpha z_0)$ of each. Let us take the limit of Equation (11.11-13) as $z_0 \to 0$, i.e., the separation between amplifiers tends to zero. In this limit the whole length of the fiber acts as a distributed amplifier with a gain constant $g = \alpha$, just enough to maintain the signal at a constant value. Since $S(z)$ is a constant, we need only evaluate the ASE optical power $F(z)$ in order to obtain, using (11.11-12), an expression for the SNR at z. To find how much noise power is added by the amplifying fiber, we consider a differential length dz. It may be viewed as a discrete amplifier with a gain of $\exp(gdz)$ so that its contribution to $F(z)$ is given by (11.11-1) as

$$dF = (e^{g(dz)} - 1)\mu h\nu \Delta f \tag{11.11-14}$$

or

$$\frac{dF}{dz} = g\mu h\nu \Delta f, \; F(z) = g\mu h\nu \Delta f z \tag{11.11-15}$$

where, since no spontaneous emission is present at the input, we used $F(0) = 0$. Using 11.11-15 in (11.11-11) and taking $S(z) = S_0$, $g = \alpha$ results in

$$\mathrm{SNR}(z) = \frac{S_0}{2[1 + 2\mu\alpha z]h\nu\Delta f} \tag{11.11-16}$$

We can also obtain (11.11-16) as the limit of (11.11-13) when $z_0 \to 0$. It is interesting to compare the (ideal) distributed amplifier to the discrete amplifier case of Equation (11.11-13)

$$(\mathrm{SNR})(z) = \frac{S_0}{2[1 + 2(z/z_0)\mu(e^{\alpha z_0} - 1)]\, h\nu\Delta f} \tag{11.11-17}$$

$$z_0 = \alpha^{-1}$$

where we used $G = \exp(\alpha z_0)$ and $n = z/z_0$.

Figure 11-35 shows plots of the ideal continuous amplification case described by Equation (11.11-16) as well as two cases of discrete amplifier cascades [Equation (11.11-13)]. The advantage of continuous amplification

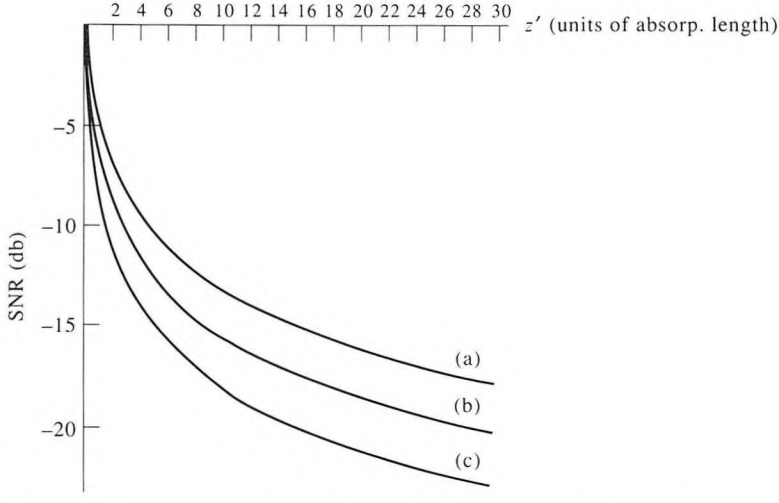


Figure 11-35 A universal plot of the degradation of the SNR compared to the initial ($z = 0$) value in the cases of (a) continuous amplification ($g = \alpha$), ($\mu = 1$); (b) periodic amplification every $z_0 = \alpha^{-1}$ ($z' = 1, 2, 3, \ldots$), ($\mu = 1$), (curve is to be read only at $z' = 1, 2, 3 \ldots$); and (c) periodic amplification every $z_0 = 2\alpha^{-1}$ ($z' = 2, 4, 6, \ldots$), ($\mu = 1$), (curve is to be read only at $z' = 2, 4, 6, \ldots$).

compared to, say, amplification every α^{-1} is seen to be less than 2 db so that the latter may be taken as a practical optimum configuration. In a low-loss optical fiber, say with $\alpha = 0.2$ db/km, the distance between amplifiers that are placed every α^{-1} km would be 21.7 km. Figure 11-36 shows the SNR of the detected signal along a realistic link for the case of (a) continuous amplification; (b) discrete amplifiers spaced by $z_0 = \alpha_0^{-1}$: and (c) for the case of no amplification at all. The launched power is $P_0 = \not{p}$ mW, $\lambda = 1.55$ μm, $\Delta f = 10^9$ Hz, and $\alpha = 0.2$ db/km. Curve (b) is to be read only at multiples of $z = \alpha^{-1} = 21.7$ km, which are the output planes of the optical amplifiers. Curve (c) assumes detection with a receiver with $T_e = 725$ K ($F = 4$ db) and an input impedance of 1000 Ω.

We note that if, for example, we need to maintain a SNR exceeding 50 db, we must use a fiber link shorter than 100 km if no amplifier is used, but if laser amplifiers are used every, say, $z_0 = \alpha^{-1}$ (=21.7 km), fiber length in excess of 1000 km can be employed.

Serious consideration has also been devoted to the use of semiconductor (SC) laser amplifiers [31]. These are identical in their construction to semiconductor laser oscillators, which are discussed in Chapter 15, except that the facets are coated with antireflection layers to reduce optical feedback and thus prevent oscillation from taking place. The main advantage is the possibility of very large gains, > 20 db in a short ($< \sim 400$ μm) semiconductor chip. The main disadvantages of the SC amplifier compared to the fiber

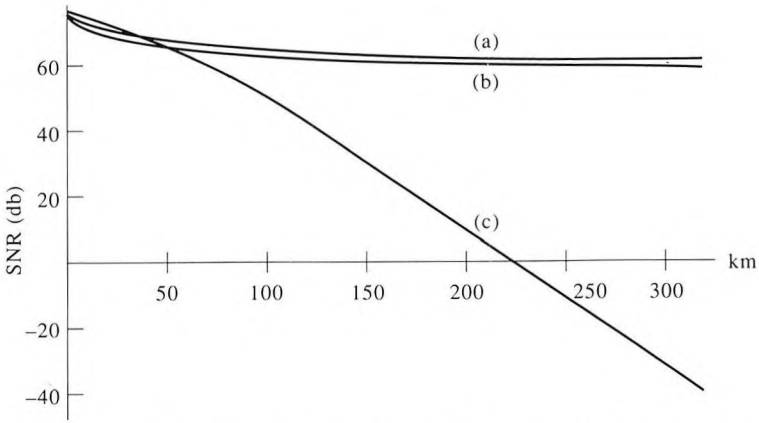


Figure 11-36 SNR of detected signal in a fiber link with (a) a continuous amplifier $g = \alpha$, ($\mu = 1$); (b) discrete amplifiers employed every absorption length $\alpha^{-1} =$ 21.7 km (0.2 db/km fiber loss), ($\mu = 1$) (curve is to be read only at multiples of 21.7 km); and (c) no optical amplification and detection with a receiver with a noise figure of 4 db. The power launched into the fiber is 5 mW, the fiber loss is 0.2 db/km, $\lambda = 1.55$ μm, the detection bandwidth is $\Delta f = 10^9$, and the detector load impedance is 1000 ohms.

amplifier is the presence of residual reflections and the resulting need for optical isolators. The presence of even minute reflection ($R < 10^{-5}$) can give rise to instabilities and excess noise in the source laser oscillator. Impressive results, however, have been demonstrated [36].

The above discussion centers on the use of optical amplification in long distance transmission of data. A second class of applications, no less important, is that of distribution systems with a very large number of subscribers. The use of optical amplifiers makes it possible to maintain the power arriving at a subscriber's premises at sufficiently high levels so as not to be degraded by the receiver noise. The number of subscribers that can thus be served by a single laser can be increased by anywhere from 1 to 3 orders of magnitude. This topic is the subject of Problem 11.13.

Problems

11.1 Show that the total output shot-noise power in a photomultiplier including that originating in the dynodes is given by

$$\overline{(i_N^2)} = G^2 2e(\bar{i}_c + i_d)\Delta f \frac{1 - \delta^{-N}}{1 - \delta^{-1}}$$

where δ is the secondary-emission multiplication factor and N is the number of stages.

11.2 Calculate the minimum power that can be detected by a photoconductor in the presence of a strong optical background power P_B. *Answer:*

$$(P_s)_{\min} = 2\left(\frac{P_B h\nu \Delta f}{\eta}\right)^{1/2}$$

11.3 Derive the expression for the minimum detectable power using a photoconductor in the video mode (that is, no local-oscillator power) and assuming that the main noise contribution is the generation–recombination noise. The optical field is given by $e(t) = E(1 + \cos \omega_m t) \cos \omega t$, and the signal is taken as the component of the photocurrent at ω_m.

11.4 Derive the minimum detectable power of a Ge:Hg detector with characteristics similar to those described in Section 11.7 when the average current is due mostly to blackbody radiation incident on the photocathode. Assume $T = 295$ K, an acceptance solid angle $\Omega = \pi$ and a photocathode area of 1 mm^2. Assume that the quantum yield η for blackbody radiation at $\lambda < 14$ μm is unity and that for $\lambda > 14$ μm, $\eta = 0$. [*Hint:* Find the flux of photons with wavelengths $14\ \mu\text{m} > \lambda > 0$ using blackbody radiation formulas or, more easily, tables or a blackbody "slide rule."]

11.5 Find the minimum detectable power in Problem 11.4 when the input field of view is at $T = 4.2$ K.

11.6 Derive Equations (11.6-15) and (11.6-16).

11.7 Show that the transit time reduction factor $(1 - e^{-i\omega_m \tau_d})/i\omega_m \tau_d$ in Equation (11.7-7) can be written as

$$\alpha - i\beta$$

where

$$\alpha = \frac{\sin \omega_m \tau_d}{\omega_m \tau_d} \qquad \beta = \frac{1 - \cos \omega_m \tau_d}{\omega_m \tau_d}$$

Plot α and β as a function of $\omega_m \tau_d$.

11.8 Derive the minimum detectable optical power for a photodiode operated in the heterodyne mode. (*Answer:* $P_{\min} = h\nu\, \Delta\nu/\eta$.)

11.9 Discuss the limiting sensitivity of an avalanche photodiode in which the noise increases as M^2. Compare it with that of a photomultiplier. What is the minimum detectable power in the limit of $M \gg 1$, and of zero background radiation and no dark current?

11.10 Derive an expression for the magnitude of the output current in a heterodyne detection scheme as a function of the angle θ between the signal and local-oscillator propagation directions. Taking the aperture diameter (see Figure 11-6) as D, show that if the output is to remain near its maximum ($\theta = 0°$) value, θ should not exceed λ/D. [*Hint:* You may replace the lens

in Figure 11-6 by the photoemissive surface.] Show that instead of Equation (11.4-4) the current from an element $dx\ dy$ of the detector is

$$di(x, t) = \frac{P_L e\eta}{h\nu(\pi D^2/4)} \left[1 + 2\sqrt{\frac{P_s}{P_L}} \cos(\omega t + kx \sin\theta) \right] dx\ dy$$

The propagation directions lie in the z-x plane. The contribution of $dx\ dy$ to the (complex) signal current is thus

$$dI_s(x, t) = \frac{2\sqrt{P_s P_L}}{h\nu(\pi D^2/4)} e^{ikx\sin\theta}\ dx\ dy$$

11.11 Show that for a Poisson distribution (footnote 9) $\overline{(\Delta N)^2} = \bar{N}$.

11.12 Calculate the smallest temperature micrometer that can be measured by an infrared detector "looking" at an object at $T = 350$ K with a background temperature of $T = 300$ K. The detector has a $D_\lambda^* = 10^{11}$ cm $(\text{Hz})^{1/2}/\text{W}$ and responds to $\Delta\lambda \sim 0.1\lambda$ centered on $\lambda = 10\ \mu$m. The output circuit bandwidth is $\Delta f = 10^3$ Hz.

11.13 Assume a fiber distribution network fed by a single semiconductor laser at $\lambda = 1.55\ \mu$m with a power output $P_0 = 10$ mW. The power is divided into N branches, amplified (in each branch) and then divided again into M branches.

Determine the maximum number of "subscribers" NM that can be serviced by the system assuming: $\Delta f = 10^9$ Hz; R (receiver input impedance) is 10^3 ohms, $T_e = 1000$ K; and a minimum SNR at the subscriber of 42 db. The maximum power level at the output of the amplifiers is 10 mW.

References

1. Yariv, A., *Quantum Electronics*, 3d ed. New York: Wiley, 1988, p. 54.
2. Engstrom, R. W., "Multiplier phototube characteristics: Application to low light levels," *J. Opt. Soc. Am.* 37:420, 1947.
3. Sommer, A. H., *Photo-Emissive Materials*. New York: Wiley, 1968.
4. Forrester, A. T., "Photoelectric mixing as a spectroscopic tool," *J. Opt. Soc. Am.* 51:253, 1961.
5. Siegman, A. E., S. E. Harris, and B. J. McMurtry, "Optical heterodyning and optical demodulation at microwave frequencies." In *Optical Masers*, J. Fox, ed. New York: Wiley, 1963, p. 511.
6. Mandel, L., "Heterodyne detection of a weak light beam," *J. Opt. Soc. Am.* 56:1200, 1966.
7. Chapman, R. A., and W. G. Hutchinson, "Excitation spectra and photoionization of neutral mercury centers in germanium," *Phys. Rev.* 157:615, 1967.
8. Teich, M. C., "Infrared heterodyne detection," *Proc. IEEE* 56:37, 1968.
9. Buczek, C., and G. Picus, "Heterodyne performance of mercury doped germanium," *Appl. Phys. Lett.* 11:125, 1967.

10. Lucovsky, G., M. E. Lasser, and R. B. Emmons, "Coherent light detection in solid-state photodiodes," *Proc. IEEE* 51:166, 1963.
11. Riesz, R. P., "High speed semiconductor photodiodes," *Rev. Sci. Instr.* 33:994, 1962.
12. Anderson, L. K., and B. J. McMurtry, "High speed photodetectors," *Appl. Opt.* 5:1573, 1966.
13. D'Asaro, L. A., and L. K. Anderson, "At the end of the laser beam, a more sensitive photodiode," *Electronics*, May 30, 1966, p. 94.
14. Shockley, W., "Hot electrons in germanium and Ohm's law," *Bell Syst. Tech. J.* 30:990, 1951.
15. McKay, K. G., and K. B. McAfee, "Electron multiplication in silicon and germanium," *Phys. Rev.* 91:1079, 1953.
16. McKay, K. G., "Avalanche breakdown in silicon," *Phys. Rev.* 94:877, 1954.
17. McIntyre, R., "Multiplication noise in uniform avalanche diodes," *IEEE Trans. Elect. Devices* ED-13:164, 1966.
18. Lindley, W. T., R. J. Phelan, C. M. Wolfe, and A. J. Foyt, "GaAs Schottky barrier avalanche photodiodes," *Appl. Phys. Lett.* 14:197, 1969.
19. Nahory, R. E., M. A. Pollack, E. D. Beebe, and J. C. DeWinter, "Continuous operation of a 1.0 μm wavelength $GaAs_{1-x}Sb_x/Al_yGa_{1-y}$-$As_{1-x}Sb_x$ double-heterostructure injection laser," *Appl. Phys. Lett.* 28:19, 1976.
20. Bar-Chaim, N., K. Y. Lau, I. Ury, and A. Yariv, "High speed GaAlAs/GaAs photodiode on a semi-insulating GaAs substrate," *Appl. Phys. Lett.* 43:261, 1983.
21. Wang, S. Y., D. M. Bloom, and D. M. Collins, "20-GHz bandwidth GaAs photodiode," *Appl. Phys. Lett.* 42:190, 1983.
22. Wang, S. Y., and D. M. Bloom, "100 GHz bandwidth planar GaAs Schottky photodiode," *Elec. Lett.* 19:554, 1983.
23. Valdmanis, J. A., G. Mourou, and C. W. Gabel, "Picosecond electro-optic sampling system," *Appl. Phys. Lett.* 41:211, 1982.
24. Kolner, B. H., D. M. Bloom, and P. S. Cross, "Characterization of high speed GaAs photodiodes using a 100-GHz electro-optic sampling system," 1983 Conference on Lasers and Electro-optics, paper ThGl.
25. Kinch, M. A. and A. Yariv, "Performance limitations of GaAs/GaAlAs infrared superlattices," *Appl. Phys. Lett.* 55:2093, 1990.
26. R. E. Burgess, "Fluctuations in the number of electrons and holes in semiconductors." *Proc. Phys. Soc. B* 68:661, 1955.
27. Van Der Ziel, A., *Fluctuation Phenomena in Semiconductors* Ch. 4, New York: Academic Press, 1959.
28. Long, D. "On generation-recombination noise in infrared detector materials," *Infrared Phys.* 7:167, Pergamon Press, 1967.
29. L. C. Chiu, J. S. Smith, S. Margalit, A. Yariv, and A. Y. Cho, "Application of internal photoemission from quantum well heterojunction superlattices to infrared detectors" *Infrared Phys.* 23(2):93, 1983.
30. B. F. Levine, C. G. Bethea, G. Hasnain, J. Walker, and R. J. Malek,

"High detectivity D* = 10^{10} cm$\sqrt{\text{Hz}}$/W GaAs/AlGaAs multiquantum well λ = 8.3 μm Infrared Detector," *Appl. Phys. Lett.* 53:2196, 1988.

31. Simon J. C., "Semiconductor laser amplifier for single mode optical fiber communications," *J. Opt. Commun.* 4:51, 1983.
32. Mears, R. J., L. Reekie, I. M. Jauncey, and D. N. Payne, "Low noise erbium-doped fiber amplifier operating at 1.54 mm," *Elec. Lett.* 23:1026, 1987.
33. Hagimoto, K., et al. "A 212 Km non-repeatered transmission experiment at 1.8 Gb/s using LD pumped Er^{3+}-doped fiber amplifiers in an Im/direct-detection repeater system," in Proceedings of the Optical Fiber Conference, Houston, TX, postdeadline Paper PD15, 1989.
34. Olshansky, R., "Noise figure for Er-doped optical fibre amplifiers," *Elec. Letts.* 24:1363, 1988.
35. Payne, David N., "Tutorial session abstracts," Optical Fiber Communication (OFC 1990) Conference, San Francisco 1990. Published by Opt. Soc. of Am., Washington, D.C.
36. See, for example, Eisenstein, G., U. Koren, G. Raybon, T. L. Koch, M. Wiesenfeld, M. Wegener, R. S. Tucker, and B. I. Miller, "Large-signal and small-signal gain characteristics of 1.5 mm quantum well optical amplifiers," *Appl. Phys. Lett.* 56:201, 1990.

Supplementary Reference

37. Boyd, R. W., *Radiometry and the Detection of Optical Radiation*. New York: Wiley, 1983.