

SEMICONDUCTOR PHOTON DETECTORS

17.1 PROPERTIES OF SEMICONDUCTOR PHOTODETECTORS

- A. Quantum Efficiency
- B. Responsivity
- C. Response Time

17.2 PHOTOCONDUCTORS

17.3 PHOTODIODES

- A. The *p-n* Photodiode
- B. The *p-i-n* Photodiode
- C. Heterostructure Photodiodes
- D. Array Detectors

17.4 AVALANCHE PHOTODIODES

- A. Principles of Operation
- B. Gain and Responsivity
- C. Response Time

17.5 NOISE IN PHOTODETECTORS

- A. Photoelectron Noise
- B. Gain Noise
- C. Circuit Noise
- D. Signal-to-Noise Ratio and Receiver Sensitivity



Heinrich Hertz (1857–1894) discovered photoemission in 1887.



Siméon Poisson (1781–1840) developed the probability distribution that describes photodetector noise.

A photodetector is a device that measures photon flux or optical power by converting the energy of the absorbed photons into a measurable form. Photographic film is probably the most ubiquitous of photodetectors. Two principal classes of photodetectors are in common use: **thermal detectors** and **photoelectric detectors**:

- Thermal detectors operate by converting photon energy into heat. However, most thermal detectors are rather inefficient and relatively slow as a result of the time required to change their temperature. Consequently, they are not suitable for most applications in photonics.
- The operation of photoelectric detectors is based on the **photoeffect**, in which the absorption of photons by some materials results directly in an electronic transition to a higher energy level and the generation of mobile charge carriers. Under the effect of an electric field these carriers move and produce a measurable electric current.

We consider only photoelectric detectors in this chapter.

The photoeffect takes two forms: external and internal. The former process involves **photoelectric emission**, in which the photogenerated electrons escape from the material as free electrons. In the latter process, **photoconductivity**, the excited carriers remain within the material, usually a semiconductor, and serve to increase its conductivity.

The External Photoeffect: Photoelectron Emission

If the energy of a photon illuminating the surface of a material in vacuum is sufficiently large, the excited electron can escape over the potential barrier of the material surface and be liberated into the vacuum as a free electron. The process, called **photoelectron emission**, is illustrated in Fig. 17.0-1(a). A photon of energy $h\nu$ incident on a metal releases a free electron from within the partially filled conduction band. Energy conservation requires that electrons emitted from below the Fermi level, where they

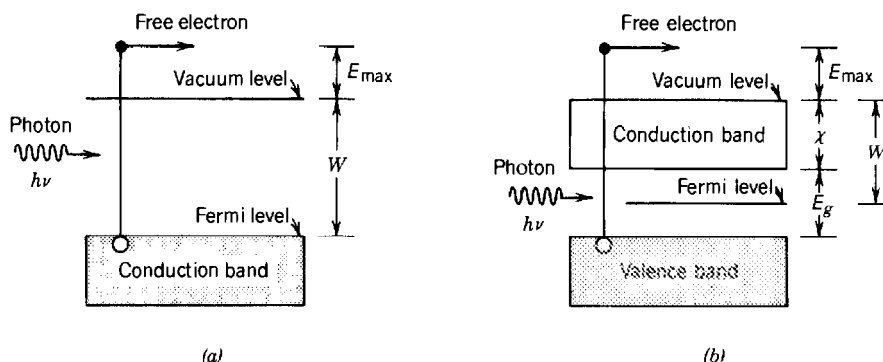


Figure 17.0-1 Photoelectric emission from (a) a metal and (b) a semiconductor.

are plentiful, have a maximum kinetic energy

$$E_{\max} = h\nu - W, \quad (17.0-1)$$

where the work function W is the energy difference between the vacuum level and the Fermi level of the material. Equation (17.0-1) is known as Einstein's photoemission equation. Only if the electron initially lies at the Fermi level can it receive the maximum kinetic energy specified in (17.0-1); the removal of a deeper-lying electron requires additional energy to transport it to the Fermi level, thereby reducing the kinetic energy of the liberated electron. The lowest work function for a metal (Cs) is about 2 eV, so that optical detectors based on the external photoeffect from pure metals are useful only in the visible and ultraviolet regions of the spectrum.

Photoelectric emission from a semiconductor is shown schematically in Fig. 17.0-1(b). Photoelectrons are usually released from the valence band, where electrons are plentiful. The formula analogous to Eq. (17.0-1) is

$$E_{\max} = h\nu - (E_g + \chi), \quad (17.0-2)$$

where E_g is the energy gap and χ is the electron affinity of the material (the energy difference between the vacuum level and the bottom of the conduction band). The energy $E_g + \chi$ can be as low as 1.4 eV for certain materials (e.g., NaKCsSb, which forms the basis for the S-20 photocathode), so that semiconductor photoemissive detectors can operate in the near infrared, as well as in the visible and ultraviolet. Furthermore, negative-electron-affinity semiconductors have been developed in which the bottom of the conduction band lies above the vacuum level in the bulk of the material, so that $h\nu$ need only exceed E_g for photoemission to occur (at the surface of the material the bands bend so that the conduction band does indeed lie below the vacuum level). These detectors are therefore responsive to slightly longer wavelengths in the near infrared, and exhibit improved quantum efficiency and reduced dark current. Photocathodes constructed of multiple layers or inhomogeneous materials, such as the S-1 photocathode, can also be used in the near infrared.

Photodetectors based on photoelectric emission usually take the form of vacuum tubes called **phototubes**. Electrons are emitted from the surface of a photoemissive material (cathode) and travel to an electrode (anode), which is maintained at a higher electric potential [Fig. 17.0-2(a)]. As a result of the electron transport between the cathode and anode, an electric current proportional to the photon flux is created in the circuit. The photoemitted electrons may also impact other specially placed metal or semiconductor surfaces in the tube, called dynodes, from which a cascade of electrons is emitted by the process of **secondary emission**. The result is an amplification of the generated electric current by a factor as high as 10^7 . This device, illustrated in Fig. 17.0-2(b), is known as a **photomultiplier tube**.

A modern imaging device that makes use of this principle is the **microchannel plate**. It consists of an array of millions of capillaries (of internal diameter $\approx 10 \mu\text{m}$) in a glass plate of thickness $\approx 1 \text{ mm}$. Both faces of the plate are coated with thin metal films that act as electrodes and a voltage is applied across them [Fig. 17.0-2(c)]. The interior walls of each capillary are coated with a secondary-electron-emissive material and behave as a continuous dynode, multiplying the photoelectron current emitted at that position [Fig. 17.0-2(d)]. The local photon flux in an image can therefore be rapidly converted into a substantial electron flux that can be measured directly. Furthermore, the electron flux can be reconverted into an (amplified) optical image by using a phosphor coating as the rear electrode to provide electroluminescence; this combination provides an **image intensifier**.

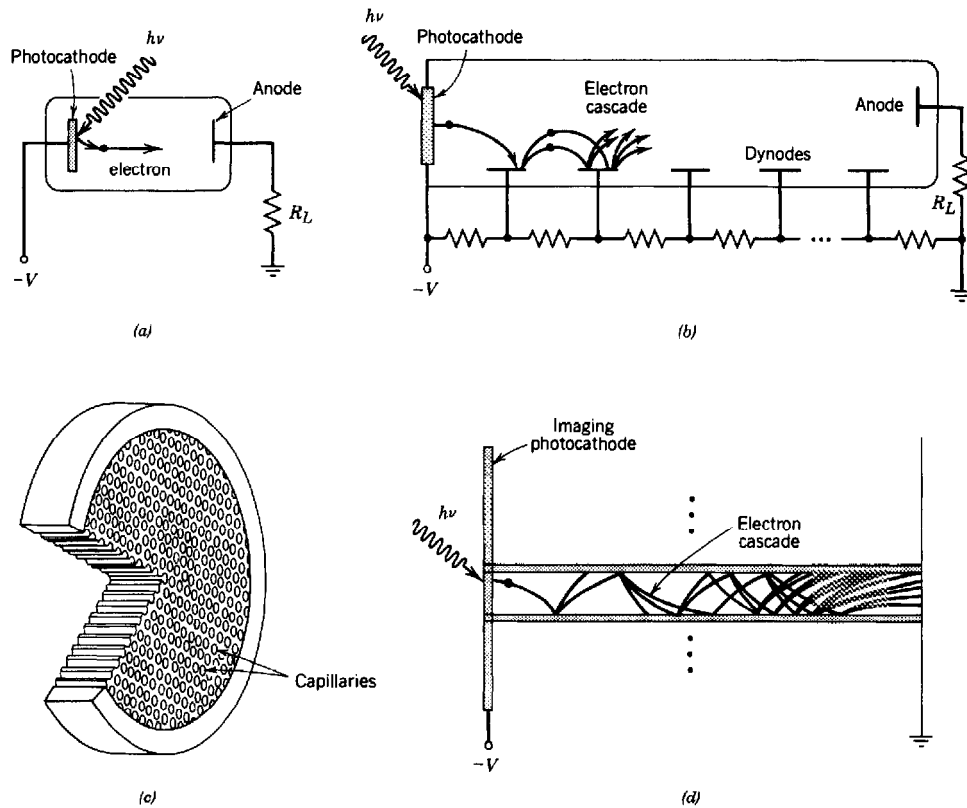


Figure 17.0-2 (a) Phototube. (b) Photomultiplier tube with semitransparent photocathode. (c) Cutaway view of microchannel plate. (d) Single capillary in a microchannel plate.

The Internal Photoeffect

Many modern photodetectors operate on the basis of the internal photoeffect, in which the photoexcited carriers (electrons and holes) remain within the sample. The most important of the internal photoeffects is **photoconductivity**. Photoconductor detectors rely directly on the light-induced increase in the electrical conductivity, which is exhibited by almost all semiconductor materials. The absorption of a photon by an intrinsic photoconductor results in the *generation* of a free electron excited from the valence band to the conduction band (Fig. 17.0-3). Concurrently, a hole is generated in the valence band. The application of an electric field to the material results in the *transport* of both electrons and holes through the material and the consequent production of an electric current in the electrical circuit of the detector.

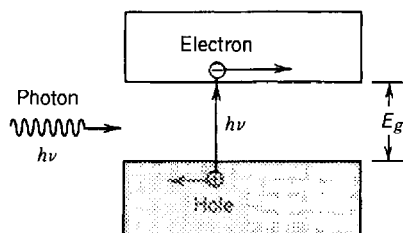


Figure 17.0-3 Electron-hole photogeneration in a semiconductor.

The semiconductor **photodiode** detector is a p - n junction structure that is also based on the internal photoeffect. Photons absorbed in the depletion layer *generate* electrons and holes which are subjected to the local electric field within that layer. The two carriers drift in opposite directions. Such a *transport* process induces an electric current in the external circuit.

Some photodetectors incorporate internal gain mechanisms so that the photoelectron current can be physically amplified within the detector and thus make the signal more easily detectable. If the depletion-layer electric field in a photodiode is increased by applying a sufficiently large reverse bias across the junction, the electrons and holes generated may acquire sufficient energy to liberate more electrons and holes within this layer by a process of impact ionization. Devices in which this internal *amplification* process occurs are known as **avalanche photodiodes** (APDs). Such detectors can be used as an alternative to (or in conjunction with) a laser amplifier (see Chaps. 13 and 16), in which the optical signal is amplified before detection. Each of these amplification mechanisms introduces its own form of noise, however.

In brief, semiconductor photoelectric detectors with gain involve the following three basic processes:

- *Generation:* Absorbed photons generate free carriers.
- *Transport:* An applied electric field induces these carriers to move, which results in a circuit current.
- *Amplification:* In APDs, large electric fields impart sufficient energy to the carriers so that they, in turn, free additional carriers by impact ionization. This internal amplification process enhances the responsivity of the detector.

This chapter is devoted to three types of semiconductor photodetectors: photoconductors, photodiodes, and avalanche photodiodes. All of these rely on the internal photoeffect as the generation mechanism. In Sec. 17.1 several important general properties of these detectors are discussed, including quantum efficiency, responsivity, and response time. The properties of photoconductor detectors are addressed in Sec. 17.2. The operation of photodiodes and avalanche photodiodes are considered in Secs. 17.3 and 17.4, respectively.

To assess the performance of semiconductor photodetectors in various applications, it is important to understand their noise properties, and these are set forth in Sec. 17.5. Noise in the output circuit of a photoelectric detector arises from several sources: the photon character of the light itself (photon noise), the conversion of photons to photocarriers (photoelectron noise), the generation of secondary carriers by internal amplification (gain noise), as well as receiver circuit noise. A brief discussion of the performance of an optical receiver is provided; we return to this topic in Sec. 22.4 in connection with the performance of fiber-optic communication systems.

17.1 PROPERTIES OF SEMICONDUCTOR PHOTODETECTORS

Certain fundamental rules govern all semiconductor photodetectors. Before studying details of the particular detectors of interest, we examine the quantum efficiency, responsivity, and response time of photoelectric detectors from a general point of view.

Semiconductor photodetectors and semiconductor photon sources are inverse devices. Detectors convert an input photon flux to an output electric current; sources achieve the opposite. The same materials are often used to make devices for both. The performance measures discussed in this section all have their counterparts in sources, as has been discussed in Chap. 16.

A. Quantum Efficiency

The **quantum efficiency** η ($0 \leq \eta \leq 1$) of a photodetector is defined as the probability that a single photon incident on the device generates a photocarrier pair that contributes to the detector current. When many photons are incident, as is almost always the case, η is the ratio of the flux of generated electron–hole pairs that contribute to the detector current to the flux of incident photons. Not all incident photons produce electron–hole pairs because not all incident photons are absorbed. This is illustrated in Fig. 17.1-1. Some photons simply fail to be absorbed because of the probabilistic nature of the absorption process (the rate of photon absorption in a semiconductor material was derived in Sec. 15.2B). Others may be reflected at the surface of the detector, thereby reducing the quantum efficiency further. Furthermore, some electron–hole pairs produced near the surface of the detector quickly recombine because of the abundance of recombination centers there and are therefore unable to contribute to the detector current. Finally, if the light is not properly focused onto the active area of the detector, some photons will be lost. This effect is not included in the definition of the quantum efficiency, however, because it is associated with the use of the device rather than with its intrinsic properties.

The quantum efficiency can therefore be written as

$$\eta = (1 - \mathcal{R})\zeta[1 - \exp(-\alpha d)], \quad (17.1-1)$$

Quantum Efficiency

where \mathcal{R} is the optical power reflectance at the surface, ζ the fraction of electron–hole pairs that contribute successfully to the detector current, α the absorption coefficient of the material (cm^{-1}) discussed in Sec. 15.2B, and d the photodetector depth. Equation (17.1-1) is a product of three factors:

- The first factor $(1 - \mathcal{R})$ represents the effect of reflection at the surface of the device. Reflection can be reduced by the use of antireflection coatings.
- The second factor ζ is the fraction of electron–hole pairs that successfully avoid recombination at the material surface and contribute to the useful photocurrent. Surface recombination can be reduced by careful material growth.
- The third factor, $\int_0^d e^{-\alpha x} dx / \int_0^\infty e^{-\alpha x} dx = [1 - \exp(-\alpha d)]$, represents the fraction of the photon flux absorbed in the bulk of the material. The device should have a sufficiently large value of d to maximize this factor.

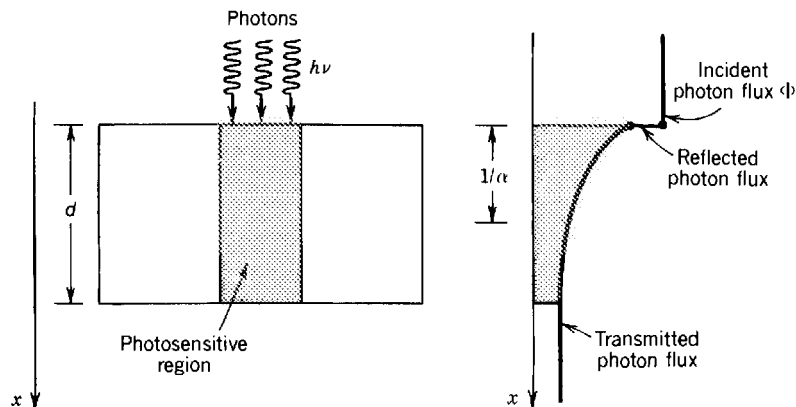


Figure 17.1-1 Effect of absorption on the quantum efficiency η .

It should be noted that some definitions of the quantum efficiency η exclude reflection at the surface, which must then be considered separately.

Dependence of η on Wavelength

The quantum efficiency η is a function of wavelength, principally because the absorption coefficient α depends on wavelength (see Fig. 15.2-2). For photodetector materials of interest, η is large within a spectral window that is determined by the characteristics of the material. For sufficiently large λ_o , η becomes small because absorption cannot occur when $\lambda_o \geq \lambda_g = hc_o/E_g$ (the photon energy is then insufficient to overcome the bandgap). The bandgap wavelength λ_g is the **long-wavelength limit** of the semiconductor material (see Chap. 15). Representative values of E_g and λ_g are shown in Figs. 15.1-5 and 15.1-6 (see also Table 15.1-3) for selected intrinsic semiconductor materials. For sufficiently small values of λ_o , η also decreases, because most photons are then absorbed near the surface of the device (e.g., for $\alpha = 10^4 \text{ cm}^{-1}$, most of the light is absorbed within a distance $1/\alpha = 1 \text{ } \mu\text{m}$). The recombination lifetime is quite short near the surface, so that the photocarriers recombine before being collected.

B. Responsivity

The **responsivity** relates the electric current flowing in the device to the incident optical power. If every photon were to generate a single photoelectron, a photon flux Φ (photons per second) would produce an electron flux Φ , corresponding to a short-circuit electric current $i_p = e\Phi$. An optical power $P = h\nu\Phi$ (watts) at frequency ν would then give rise to an electric current $i_p = eP/h\nu$. Since the fraction of photons producing detected photoelectrons is η rather than unity, the electric current is

$$i_p = \eta e\Phi = \frac{\eta eP}{h\nu} = \Re P. \quad (17.1-2)$$

The proportionality factor \Re , between the electric current and the optical power, is defined as the responsivity \Re of the device. $\Re = i_p/P$ has units of A/W and is given by

$$\Re = \frac{\eta e}{h\nu} = \eta \frac{\lambda_o}{1.24}. \quad (17.1-3)$$

Photodetector
Responsivity (A/W)
(λ_o in μm)

\Re increases with λ_o because photoelectric detectors are responsive to the photon flux rather than to the optical power. As λ_o increases, a given optical power is carried by more photons, which, in turn, produce more electrons. The region over which \Re increases with λ_o is limited, however, since the wavelength dependence of η comes into play for both long and short wavelengths. It is important to distinguish the detector responsivity defined here (A/W) from the light-emitting-diode responsivity (W/A) defined in (16.1-25).

The responsivity can be degraded if the detector is presented with an excessively large optical power. This condition, which is called detector saturation, limits the detector's **linear dynamic range**, which is the range over which it responds linearly with the incident optical power.

An appreciation for the order of magnitude of the responsivity is gained by setting $\eta = 1$ in (17.1-3), whereupon $\Re = 1 \text{ A/W}$, i.e., $1 \text{ nW} \rightarrow 1 \text{ nA}$, at $\lambda_o = 1.24 \text{ } \mu\text{m}$. The linear increase of the responsivity with wavelength, for a given fixed value of η , is

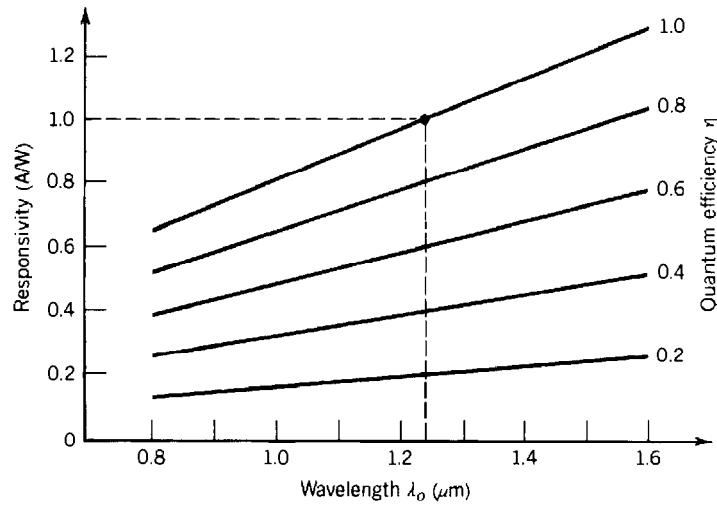


Figure 17.1-2 Responsivity \mathfrak{R} (A/W) versus wavelength λ_o with the quantum efficiency η as a parameter. $\mathfrak{R} = 1$ A/W at $\lambda_o = 1.24 \mu\text{m}$ when $\eta = 1$.

illustrated in Fig. 17.1-2. \mathfrak{R} is also seen to increase linearly with η if λ_o is fixed. For thermal detectors \mathfrak{R} is independent of λ_o because they respond directly to optical power rather than to the photon flux.

Devices with Gain

The formulas presented above are predicated on the assumption that each carrier produces a charge e in the detector circuit. However, many devices produce a charge q in the circuit that differs from e . Such devices are said to exhibit gain. The gain G is the average number of circuit electrons generated per photocarrier pair. G should be distinguished from η , which is the probability that an incident photon produces a detectable photocarrier pair. The gain, which is defined as

$$G = \frac{q}{e}, \quad (17.1-4)$$

can be either greater than or less than unity, as will be seen subsequently. Therefore, more general expressions for the photocurrent and responsivity are

$$i_p = \eta q \Phi = G \eta e \Phi = \frac{G \eta e P}{h\nu} \quad (17.1-5)$$

Photocurrent

and

$$\mathfrak{R} = \frac{G \eta e}{h\nu} = G \eta \frac{\lambda_o}{1.24}, \quad (17.1-6)$$

Responsivity in the Presence
of Gain (A/W)
(λ_o in μm)

respectively.

Other useful measures of photodetector behavior, such as signal-to-noise ratio and receiver sensitivity, must await a discussion of the detector noise properties presented in Sec. 17.5.

C. Response Time

One might be inclined to argue that the charge generated in an external circuit should be $2e$ when a photon generates an electron-hole pair in a photodetector material, since there are two charge carriers. In fact, the charge generated is e , as we will show below. Furthermore, the charge delivered to the external circuit by carrier motion in the photodetector material is not provided instantaneously but rather occupies an extended time. It is as if the motion of the charged carriers in the material draws charge slowly from the wire on one side of the device and pushes it slowly into the wire at the other side so that each charge passing through the external circuit is spread out in time. This phenomenon is known as **transit-time spread**. It is an important limiting factor for the speed of operation of all semiconductor photodetectors.

Consider an electron-hole pair generated (by photon absorption, for example) at an arbitrary position x in a semiconductor material of width w to which a voltage V is applied, as shown in Fig. 17.1-3(a). We restrict our attention to motion in the x direction. A carrier of charge Q (a hole of charge $Q = e$ or an electron of charge $Q = -e$) moving with a velocity $v(t)$ in the x direction creates a current in the external circuit given by

$$i(t) = -\frac{Q}{w}v(t).$$

(17.1-7)

Ramo's Theorem

This important formula, known as **Ramo's theorem**, can be proved with the help of an

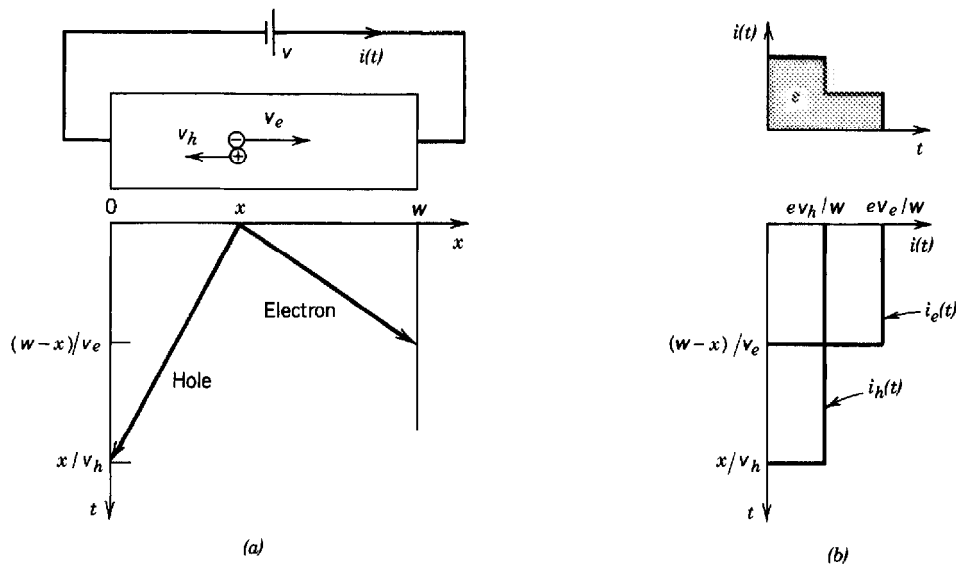


Figure 17.1-3 (a) An electron-hole pair is generated at the position x . The hole moves to the left with velocity v_h and the electron moves to the right with velocity v_e . The process terminates when the carriers reach the edge of the material. (b) Hole current $i_h(t)$, electron current $i_e(t)$, and total current $i(t)$ induced in the circuit. The total charge induced in the circuit is e .

energy argument. If the charge moves a distance dx in the time dt , under the influence of an electric field of magnitude $E = V/w$, the work done is $-QE dx = -Q(V/w) dx$. This work must equal the energy provided by the external circuit, $i(t)V dt$. Thus $i(t)V dt = -Q(V/w) dx$ from which $i(t) = -(Q/w)(dx/dt) = -(Q/w)v(t)$, as promised.

In the presence of a uniform charge density ρ , instead of a single point charge Q , the total charge is ρAw , where A is the cross-sectional area, so that (17.1-7) gives $i(t) = -(\rho Aw/w)v(t) = -\rho Av(t)$ from which the current density in the x direction $J(t) = -i(t)/A = \rho v(t)$.

In the presence of an electric field E , a charge carrier in a semiconductor will drift at a mean velocity

$$v = \mu E, \quad (17.1-8)$$

where μ is the carrier mobility. Thus, $J = \sigma E$, where $\sigma = \mu \rho$ is the conductivity.

Assuming that the hole moves with constant velocity v_h to the left, and the electron moves with constant velocity v_e to the right, (17.1-7) tells us that the hole current $i_h = -e(-v_h)/w$ and the electron current $i_e = -(-e)v_e/w$, as illustrated in Fig. 17.1-3(b). Each carrier contributes to the current as long as it is moving. If the carriers continue their motion until they reach the edge of the material, the hole moves for a time x/v_h and the electron moves for a time $(w-x)/v_e$ [see Fig. 17.1-3(a)]. In semiconductors, v_e is generally larger than v_h so that the full width of the transit-time spread is x/v_h .

The total charge q induced in the external circuit is the sum of the areas under i_e and i_h

$$q = e \frac{v_h}{w} \frac{x}{v_h} + e \frac{v_e}{w} \frac{w-x}{v_e} = e \left(\frac{x}{w} + \frac{w-x}{w} \right) = e,$$

as promised. The result is independent of the position x at which the electron-hole pair was created.

The transit-time spread is even more severe if the electron-hole pairs are generated uniformly throughout the material, as shown in Fig. 17.1-4. For $v_h < v_e$, the full width of the transit-time spread is then w/v_h rather than x/v_h . This occurs because uniform illumination produces carrier pairs everywhere, including at $x = w$, which is the point at which the holes have the farthest to travel before being able to recombine at $x = 0$.

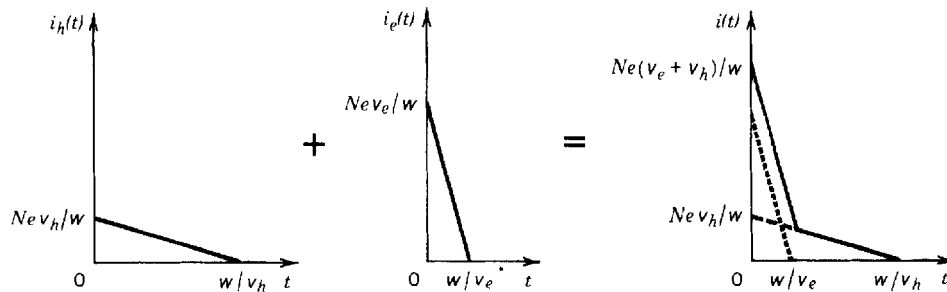


Figure 17.1-4 Hole current $i_h(t)$, electron current $i_e(t)$, and total current $i(t)$ induced in the circuit for electron-hole generation by N photons uniformly distributed between 0 and w (see Problem 17.1-4). The tail in the total current results from the motion of the holes. $i(t)$ can be viewed as the impulse-response function (see Appendix B) for a uniformly illuminated detector subject to transit-time spread.

Another response-time limit of semiconductor detectors is the RC time constant formed by the resistance R and capacitance C of the photodetector and its circuitry. The combination of resistance and capacitance serves to integrate the current at the output of the detector, and thereby to lengthen the impulse-response function. The impulse-response function in the presence of transit-time *and* simple RC time-constant spread is determined by convolving $i(t)$ in Fig. 17.1-4 with the exponential function $(1/RC) \exp(-t/RC)$ (see Appendix B, Sec. B.1). Photodetectors of different types have other specific limitations on their speed of response; these are considered at the appropriate point.

As a final point, we mention that photodetectors of a given material and structure often exhibit a fixed gain–bandwidth product. Increasing the gain results in a decrease of the bandwidth, and vice versa. This trade-off between sensitivity and frequency response is associated with the time required for the gain process to take place.

17.2 PHOTOCONDUCTORS

When photons are absorbed by a semiconductor material, mobile charge carriers are *generated* (an electron–hole pair for every absorbed photon). The electrical conductivity of the material increases in proportion to the photon flux. An electric field applied to the material by an external voltage source causes the electrons and holes to be *transported*. This results in a measurable electric current in the circuit, as shown in Fig. 17.2-1. **Photoconductor detectors** operate by registering either the photocurrent i_p , which is proportional to the photon flux Φ , or the voltage drop across a load resistor R placed in series with the circuit.

The semiconducting material may take the form of a slab or a thin film. The anode and cathode contacts are often placed on the same surface of the material, interdigitating with each other to maximize the light transmission while minimizing the transit time (see Fig. 17.2-1). Light can also be admitted from the bottom of the device if the substrate has a sufficiently large bandgap (so that it is not absorptive).

The increase in conductivity arising from a photon flux Φ (photons per second) illuminating a semiconductor volume WA (see Fig. 17.2-1) may be calculated as follows. A fraction η of the incident photon flux is absorbed and gives rise to excess

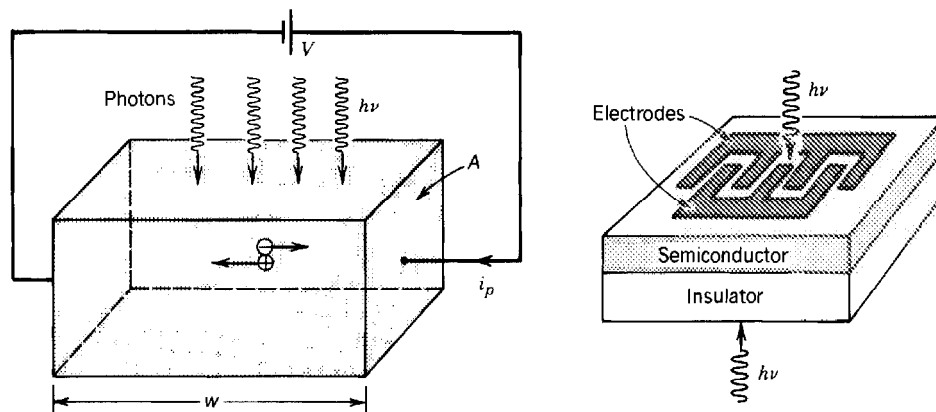


Figure 17.2-1 The photoconductor detector. Photogenerated carrier pairs move in response to the applied voltage V , generating a photocurrent i_p proportional to the incident photon flux. The interdigitated electrode structure shown is designed to maximize both the light reaching the semiconductor and the device bandwidth (by minimizing the carrier transit time).

electron-hole pairs. The pair-production rate R (per unit volume) is therefore $R = \eta\Phi/wA$. If τ is the excess-carrier recombination lifetime, electrons are lost at the rate $\Delta n/\tau$ where Δn is the photoelectron concentration (see Chap. 15). Under steady-state conditions both rates are equal ($R = \Delta n/\tau$) so that $\Delta n = \eta\tau\Phi/wA$. The increase in the charge carrier concentration therefore results in an increase in the conductivity given by

$$\Delta\sigma = e\Delta n(\mu_e + \mu_h) = \frac{e\eta\tau(\mu_e + \mu_h)}{wA}\Phi, \quad (17.2-1)$$

where μ_e and μ_h are the electron and hole mobilities. Thus the increase in conductivity is proportional to the photon flux.

Since the current density $J_p = \Delta\sigma E$ and $v_e = \mu_e E$ and $v_h = \mu_h E$ where E is the electric field, (17.2-1) gives $J_p = [e\eta\tau(v_e + v_h)/wA]\Phi$ corresponding to an electric current $i_p = AJ_p = [e\eta\tau(v_e + v_h)/w]\Phi$. If $v_h \ll v_e$ and $\tau_e = w/v_e$,

$$i_p \approx e\eta \frac{\tau}{\tau_e} \Phi. \quad (17.2-2)$$

In accordance with (17.1-5), the ratio τ/τ_e in (17.2-2) corresponds to the detector gain $G = \tau/\tau_e$, as explained subsequently.

Gain

The responsivity of a photoconductor is given by (17.1-6). The device exhibits an internal gain which, simply viewed, comes about because the recombination lifetime and transit time generally differ. Suppose that electrons travel faster than holes (see Fig. 17.2-1) and that the recombination lifetime is very long. As the electron and hole are transported to opposite sides of the photoconductor, the electron completes its trip sooner than the hole. The requirement of current continuity forces the external circuit to provide another electron immediately, which enters the device from the wire at the left. This new electron moves quickly toward the right, again completing its trip before the hole reaches the left edge. This process continues until the electron recombines with the hole. A single photon absorption can therefore result in an electron passing through the external circuit many times. The expected number of trips that the electron makes before the process terminates is

$$G = \frac{\tau}{\tau_e}, \quad (17.2-3)$$

where τ is the excess-carrier recombination lifetime and $\tau_e = w/v_e$ is the electron transit time across the sample. The charge delivered to the circuit by a single electron-hole pair in this case is $q = Ge > e$ so that the device exhibits gain.

However, the recombination lifetime may be sufficiently short such that the carriers recombine before reaching the edge of the material. This can occur provided that there is a ready availability of carriers of the opposite type for recombination. In that case $\tau < \tau_e$ and the gain is less than unity so that, on average, the carriers contribute only a fraction of the electronic charge e to the circuit. Charge is, of course, conserved and the many carrier pairs present deliver an integral number of electronic charges to the circuit.

The photoconductor gain $G = \tau/\tau_e$ can be interpreted as the fraction of the sample length traversed by the average excited carrier before it undergoes recombination. The transit time τ_e depends on the dimensions of the device and the applied voltage via (17.1-8); typical values of $w = 1$ mm and $v_e = 10^7$ cm/s give $\tau_e \approx 10^{-8}$ s. The

TABLE 17.2-1 Selected Extrinsic Semiconductor Materials with Their Activation Energy and Long-Wavelength Limit

Semiconductor:Dopant	E_A (eV)	λ_A (μm)
Ge:Hg	0.088	14
Ge:Cu	0.041	30
Ge:Zn	0.033	38
Ge:B	0.010	124
Si:B	0.044	28

recombination lifetime τ can range from 10^{-13} s to many seconds, depending on the photoconductor material and doping [see (15.1-17)]. Thus G can assume a broad range of values, both below unity and above unity, depending on the parameters of the material, the size of the device, and the applied voltage. The gain of a photoconductor cannot generally exceed 10^6 , however, because of the restrictions imposed by space-charge-limited current flow, impact ionization, and dielectric breakdown.

Spectral Response

The spectral sensitivity of photoconductors is governed principally by the wavelength dependence of η , as discussed in Sec. 17.1A. Different intrinsic semiconductors have different long-wavelength limits, as indicated in Chap. 15. Ternary and quaternary compound semiconductors are also used. Photoconductor detectors (unlike photoemissive detectors) can operate well into the infrared region on band-to-band transitions. However, operation at wavelengths beyond about $2\ \mu\text{m}$ requires that the devices be cooled to minimize the thermal excitation of electrons into the conduction band in these low-gap materials.

At even longer wavelengths extrinsic photoconductors can be used as detectors. **Extrinsic photoconductivity** operates on transitions involving forbidden-gap energy levels. It takes place when the photon interacts with a bound electron at a donor site, producing a free electron and a bound hole [or conversely, when it interacts with a bound hole at an acceptor site, producing a free hole and a bound electron as shown in Fig. 15.2-1(b)]. Donor and acceptor levels in the bandgap of doped semiconductor materials can have very low activation energies E_A . In this case the long-wavelength limit is $\lambda_A = hc_0/E_A$. These detectors must be cooled to avoid thermal excitation; liquid He at 4 K is often used. Representative values of E_A and λ_A are provided in Table 17.2-1 for selected extrinsic semiconductor materials.

The spectral responses of several extrinsic photoconductor detectors are shown in Fig. 17.2-2. The responsivity increases approximately linearly with λ_o , in accordance

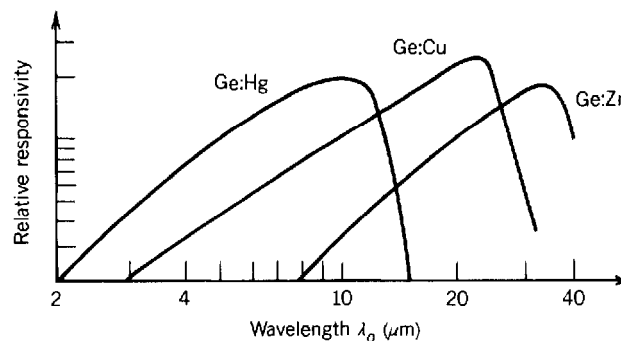


Figure 17.2-2 Relative responsivity versus wavelength λ_o (μm) for three doped-Ge extrinsic infrared photoconductor detectors.

with (17.1-6), peaks slightly below the long-wavelength limit λ_A and falls off beyond it. The quantum efficiency for these detectors can be quite high (e.g., $\eta \approx 0.5$ for Ge:Cu), although the gain may be low under usual operating conditions (e.g., $G \approx 0.03$ for Ge:Hg).

Response Time

The response time of photoconductor detectors is, of course, constrained by the transit-time and RC time-constant considerations presented in Sec. 17.1C. The carrier-transport response time is approximately equal to the recombination time τ , so that the carrier-transport bandwidth B is inversely proportional to τ . Since the gain G is proportional to τ in accordance with (17.2-3), increasing τ increases the gain, which is desirable, but it also decreases the bandwidth, which is undesirable. Thus the gain-bandwidth product GB is roughly independent of τ . Typical values of GB extend up to $\approx 10^9$.

17.3 PHOTODIODES

A. The p - n Photodiode

As with photoconductors, photodiode detectors rely on photogenerated charge carriers for their operation. A photodiode is a p - n junction (see Sec. 15.1E) whose reverse current increases when it absorbs photons. Although p - n and p - i - n photodiodes are generally faster than photoconductors, they do not exhibit gain.

Consider a reverse-biased p - n junction under illumination, as depicted in Fig. 17.3-1. Photons are absorbed everywhere with absorption coefficient α . Whenever a photon is absorbed, an electron-hole pair is *generated*. But only where an electric field is present can the charge carriers be *transported* in a particular direction. Since a p - n junction can support an electric field only in the depletion layer, this is the region in which it is desirable to generate photocarriers.

There are, however, three possible locations where electron-hole pairs can be generated:

- Electrons and holes generated in the depletion layer (region 1) quickly drift in opposite directions under the influence of the strong electric field. Since the

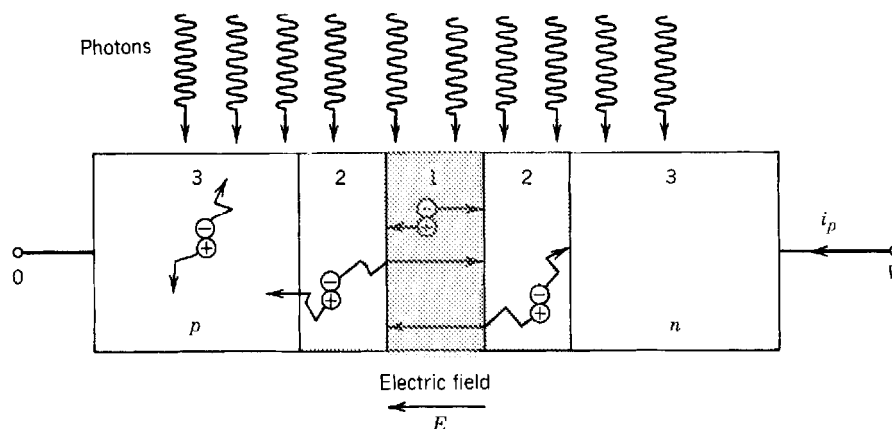


Figure 17.3-1 Photons illuminating an idealized reverse-biased p - n photodiode detector. The drift and diffusion regions are indicated by 1 and 2, respectively.

electric field always points in the n - p direction, electrons move to the n side and holes to the p side. As a result, the photocurrent created in the external circuit is always in the reverse direction (from the n to the p region). Each carrier pair generates in the external circuit an electric current pulse of area e ($G = 1$) since recombination does not take place in the depleted region.

- Electrons and holes generated away from the depletion layer (region 3) cannot be transported because of the absence of an electric field. They wander randomly until they are annihilated by recombination. They do not contribute a signal to the external electric current.
- Electron-hole pairs generated outside the depletion layer, but in its vicinity (region 2), have a chance of entering the depletion layer by random diffusion. An electron coming from the p side is quickly transported across the junction and therefore contributes a charge e to the external circuit. A hole coming from the n side has a similar effect.

Photodiodes have been fabricated from many of the semiconductor materials listed in Table 15.1-3, as well as from ternary and quaternary compound semiconductors such as InGaAs and InGaAsP. Devices are often constructed in such a way that the light impinges normally on the p - n junction instead of parallel to it. In that case the additional carrier diffusion current in the depletion region acts to enhance η , but this is counterbalanced by the decreased thickness of the material which acts to reduce η .

Response Time

The transit time of carriers drifting across the depletion layer (w_d/v_e for electrons and w_d/v_h for holes) and the RC time response play a role in the response time of photodiode detectors, as discussed in Sec. 17.1C. The resulting circuit current is shown in Fig. 17.1-3(b) for an electron-hole pair generated at the position x , and in Fig. 17.1-4 for uniform electron-hole pair generation.

In photodiodes there is an additional contribution to the response time arising from diffusion. Carriers generated outside the depletion layer, but sufficiently close to it, take time to diffuse into it. This is a relatively slow process in comparison with drift. The maximum times allowed for this process are, of course, the carrier lifetimes (τ_p for electrons in the p region and τ_n for holes in the n region). The effect of diffusion time can be decreased by using a p - i - n diode, as will be seen subsequently.

Nevertheless, photodiodes are generally faster than photoconductors because the strong field in the depletion region imparts a large velocity to the photogenerated carriers. Furthermore, photodiodes are not affected by many of the trapping effects associated with photoconductors.

Bias

As an electronic device, the photodiode has an i - V relation given by

$$i = i_s \left[\exp \left(\frac{eV}{k_B T} \right) - 1 \right] - i_p,$$

illustrated in Fig. 17.3-2. This is the usual i - V relation of a p - n junction [see (15.1-24)] with an added photocurrent $-i_p$ proportional to the photon flux.

There are three classical modes of photodiode operation: open circuit (photovoltaic), short-circuit, and reverse biased (photoconductive). In the open-circuit mode (Fig. 17.3-3), the light generates electron-hole pairs in the depletion region. The additional electrons freed on the n side of the layer recombine with holes on the p side, and vice versa. The net result is an increase in the electric field, which produces a photovoltage

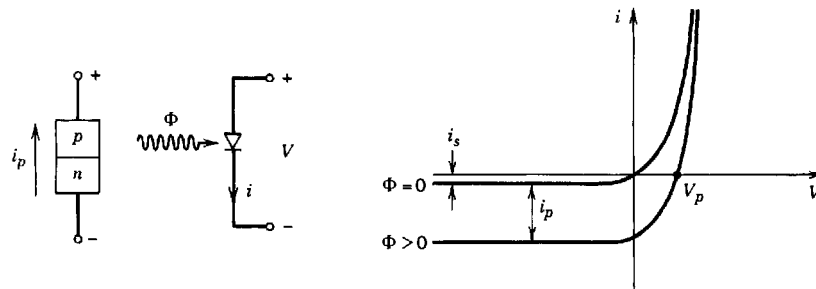


Figure 17.3-2 Generic photodiode and its i - V relation.

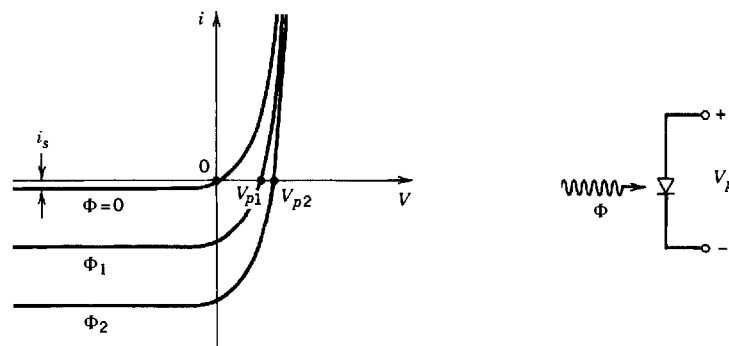


Figure 17.3-3 Photovoltaic operation of a photodiode.

V_p across the device that increases with increasing photon flux. This mode of operation is used, for example, in solar cells. The responsivity of a photovoltaic photodiode is measured in V/W rather than in A/W. The short-circuit ($V = 0$) mode is illustrated in Fig. 17.3-4. The short-circuit current is then simply the photocurrent i_p . Finally, a photodiode may be operated in its reverse-biased or “photoconductive” mode, as shown in Fig. 17.3-5(a). If a series-load resistor is inserted in the circuit, the operating conditions are those illustrated in Fig. 17.3-5(b).

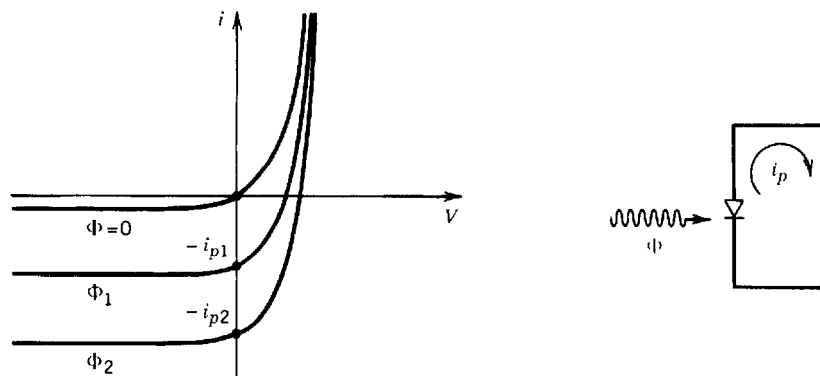


Figure 17.3-4 Short-circuit operation of a photodiode.

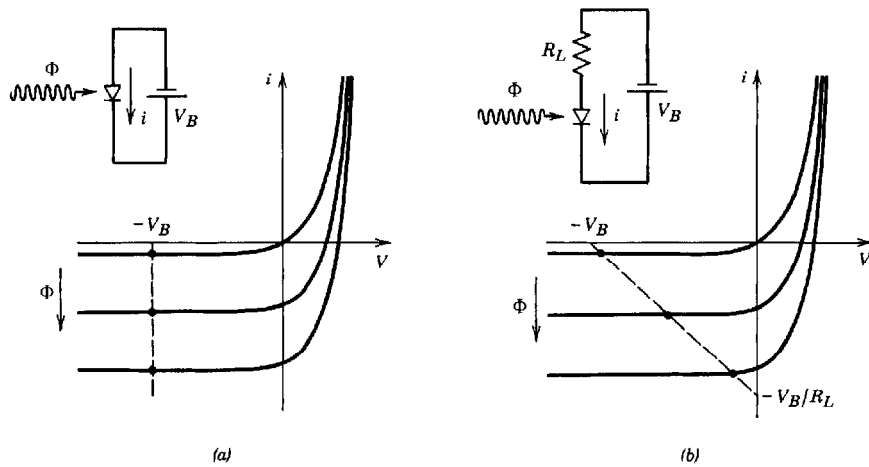


Figure 17.3-5 (a) Reverse-biased operation of a photodiode without a load resistor and (b) with a load resistor. The operating point lies on the dashed line.

Photodiodes are usually operated in the strongly reverse-biased mode for the following reasons:

- A strong reverse bias creates a strong electric field in the junction which increases the drift velocity of the carriers, thereby reducing transit time.
- A strong reverse bias increases the width of the depletion layer, thereby reducing the junction capacitance and improving the response time.
- The increased width of the depletion layer leads to a larger photosensitive area, making it easier to collect more light.

B. The *p-i-n* Photodiode

As a detector, the *p-i-n* photodiode has a number of advantages over the *p-n* photodiode. A *p-i-n* diode is a *p-n* junction with an intrinsic (usually lightly doped) layer sandwiched between the *p* and *n* layers (see Sec. 15.1E). It may be operated under the variety of bias conditions discussed in the preceding section. The energy-band diagram, charge distribution, and electric field distribution for a reverse-biased *p-i-n* diode are illustrated in Fig. 17.3-6. This structure serves to extend the width of the region supporting an electric field, in effect widening the depletion layer.

Photodiodes with the *p-i-n* structure offer the following advantages:

- Increasing the width of the depletion layer of the device (where the generated carriers can be transported by drift) increases the area available for capturing light.
- Increasing the width of the depletion layer reduces the junction capacitance and thereby the *RC* time constant. On the other hand, the transit time increases with the width of the depletion layer.
- Reducing the ratio between the diffusion length and the drift length of the device results in a greater proportion of the generated current being carried by the faster drift process.

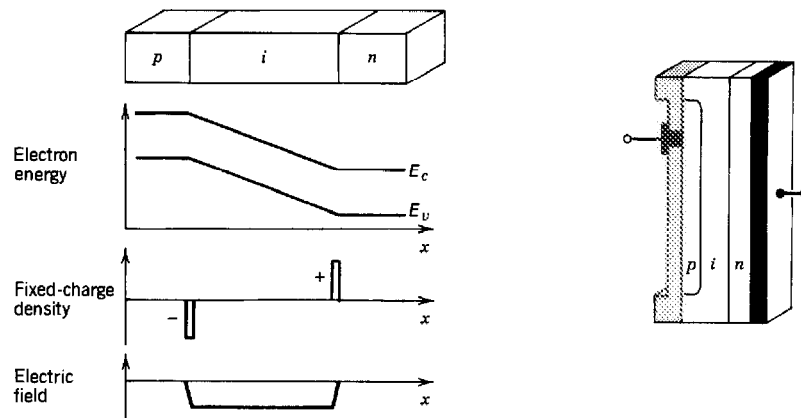


Figure 17.3-6 The p - i - n photodiode structure, energy diagram, charge distribution, and electric field distribution. The device can be illuminated either perpendicularly or parallel to the junction.

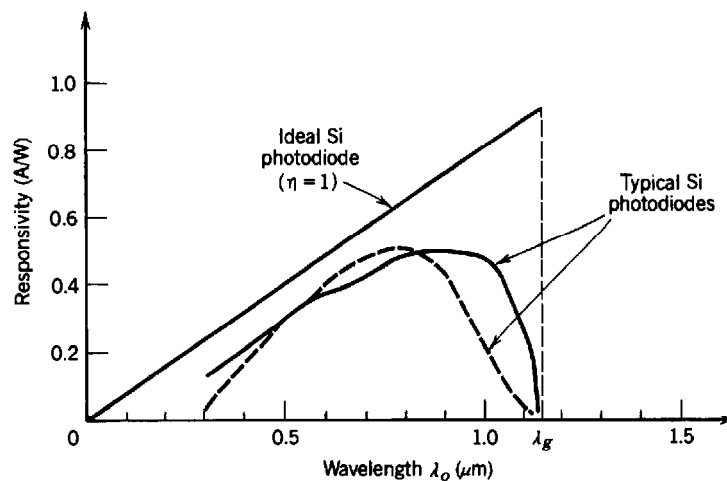


Figure 17.3-7 Responsivity versus wavelength (μm) for ideal and commercially available silicon p - i - n photodiodes.

Response times in the tens of ps, corresponding to bandwidths ≈ 50 GHz, are achievable. The responsivity of two commercially available silicon p - i - n photodiodes is compared with that of an ideal device in Fig. 17.3-7. It is interesting to note that the responsivity maximum occurs for wavelengths substantially shorter than the bandgap wavelength. This is because Si is an indirect-gap material. The photon-absorption transitions therefore typically take place from the valence-band to conduction-band states that typically lie well above the conduction-band edge (see Fig. 15.2-8).

C. Heterostructure Photodiodes

Heterostructure photodiodes, formed from two semiconductors of different bandgaps, can exhibit advantages over p - n junctions fabricated from a single material. A hetero-

junction comprising a large-bandgap material ($E_g > h\nu$), for example, can make use of its transparency to minimize optical absorption outside the depletion region. The large-bandgap material is then called a **window layer**. The use of different materials can also provide devices with a great deal of flexibility. Several material systems are of particular interest (see Figs. 15.1-5 and 15.1-6):

- $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ (AlGaAs lattice matched to a GaAs substrate) is useful in the wavelength range 0.7 to 0.87 μm .
- $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{InP}$ operates at 1.65 μm in the near infrared ($E_g = 0.75$ eV). Typical values for the responsivity and quantum efficiency of detectors fabricated from these materials are $\mathfrak{R} \approx 0.7$ A/W and $\eta \approx 0.75$. The gap wavelength can be compositionally tuned over the range of interest for fiber-optic communication, 1.3–1.6 μm .
- $\text{Hg}_x\text{Cd}_{1-x}\text{Te}/\text{CdTe}$ is a material that is highly useful in the middle-infrared region of the spectrum. This is because HgTe and CdTe have nearly the same lattice parameter and can therefore be lattice matched at nearly all compositions. This material provides a compositionally tunable bandgap that operates in the wavelength range between 3 and 17 μm .
- Quaternary materials, such as $\text{In}_{1-x}\text{Ga}_x\text{As}_{1-y}\text{P}_y/\text{InP}$ and $\text{Ga}_{1-x}\text{Al}_x\text{As}_y\text{Sb}_{1-y}/\text{GaSb}$, which are useful over the range 0.92 to 1.7 μm , are of particular interest because the fourth element provides an additional degree of freedom that allows lattice matching to be achieved for different compositionally determined values of E_g .

Schottky-Barrier Photodiodes

Metal–semiconductor photodiodes (also called **Schottky-barrier photodiodes**) are formed from metal–semiconductor heterojunctions. A thin semitransparent metallic film is used in place of the p -type (or n -type) layer in the p - n junction photodiode. The thin film is sometimes made of a metal–semiconductor alloy that behaves like a metal. The Schottky-barrier structure and its energy-band diagram are shown schematically in Fig. 17.3-8.

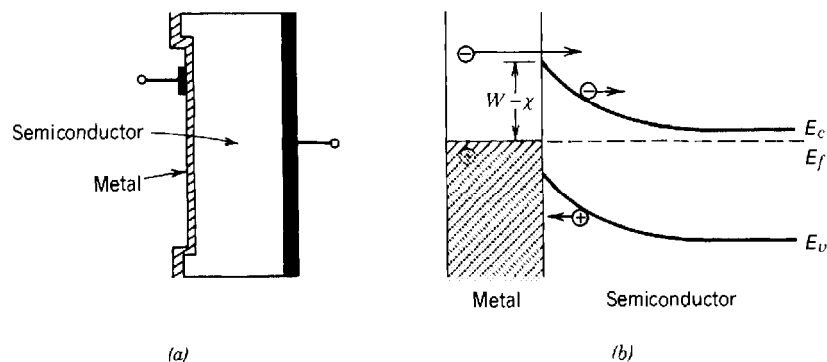


Figure 17.3-8 (a) Structure and (b) energy-band diagram of a Schottky-barrier photodiode formed by depositing a metal on an n -type semiconductor. These photodetectors are responsive to photon energies greater than the Schottky barrier height, $h\nu > W - \chi$. Schottky photodiodes can be fabricated from many materials, such as Au on n -type Si (which operates in the visible) and platinum silicide (PtSi) on p -type Si (which operates over a range of wavelengths stretching from the near ultraviolet to the infrared).

There are a number of reasons why Schottky-barrier photodiodes are useful:

- Not all semiconductors can be prepared in both p -type and n -type forms; Schottky devices are of particular interest in these materials.
- Semiconductors used for the detection of visible and ultraviolet light with photon energies well above the bandgap energies have a large absorption coefficient. This gives rise to substantial surface recombination and a reduction of the quantum efficiency. The metal–semiconductor junction has a depletion layer present immediately at the surface, thus eliminating surface recombination.
- The response speed of p - n and p - i - n junction photodiodes is in part limited by the slow diffusion current associated with photocarriers generated close to, but outside of, the depletion layer. One way of decreasing this unwanted absorption is to decrease the thickness of one of the junction layers. However, this should be achieved without substantially increasing the series resistance of the device because such an increase has the undesired effect of reducing the speed by increasing the RC time constant. The Schottky-barrier structure achieves this because of the low resistance of the metal. Furthermore Schottky barrier structures are majority-carrier devices and therefore have inherently fast responses and large operating bandwidths. Response times in the picosecond regime, corresponding to bandwidths ≈ 100 GHz, are readily available.

Representative quantum efficiencies for Schottky-barrier and p - i - n photodiode detectors are shown in Fig. 17.3-9; η can approach unity for carefully constructed Si devices that include antireflection coatings.

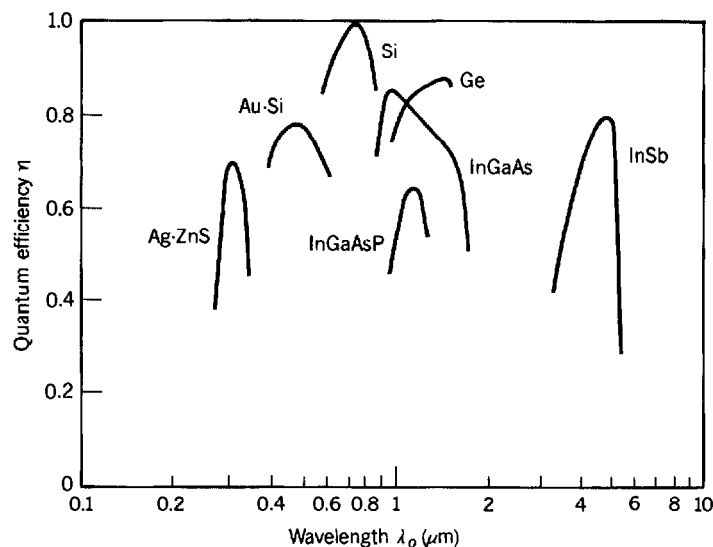


Figure 17.3-9 Quantum efficiency η versus wavelength λ_0 (μm) for various photodiodes. Si p - i - n photodiodes can be fabricated with nearly unity quantum efficiency if an antireflection coating is applied to the surface of the device. The optimal response wavelength of ternary and quaternary p - i - n photodetectors is compositionally tunable (the quantum efficiency for a range of wavelengths is shown for InGaAs). Long-wavelength photodetectors (e.g., InSb) must be cooled to minimize thermal excitation. (Adapted from S. M. Sze, *Physics of Semiconductor Devices*, Wiley, New York, 2nd ed. 1981.)

D. Array Detectors

An individual photodetector registers the photon flux striking it as a function of time. In contrast, an array containing a large number of photodetectors can simultaneously register the photon fluxes (as functions of time) from many spatial points. Such detectors therefore permit electronic versions of optical images to be formed. One type of array detector, the microchannel plate [see Fig. 17.0-2(c)], has already been discussed.

Modern microelectronics technology permits other types of arrays containing large numbers of individual semiconductor photodetectors (called pixels) to be fabricated. One example of current interest, illustrated in Fig. 17.3-10, makes use of an array of nearly 40,000 tiny Schottky-barrier photodiodes of PtSi on *p*-type Si. The device is sensitive to a broad band of wavelengths stretching from the near ultraviolet to about 6 μm in the infrared, which corresponds to the Schottky barrier height of about 0.2 eV.

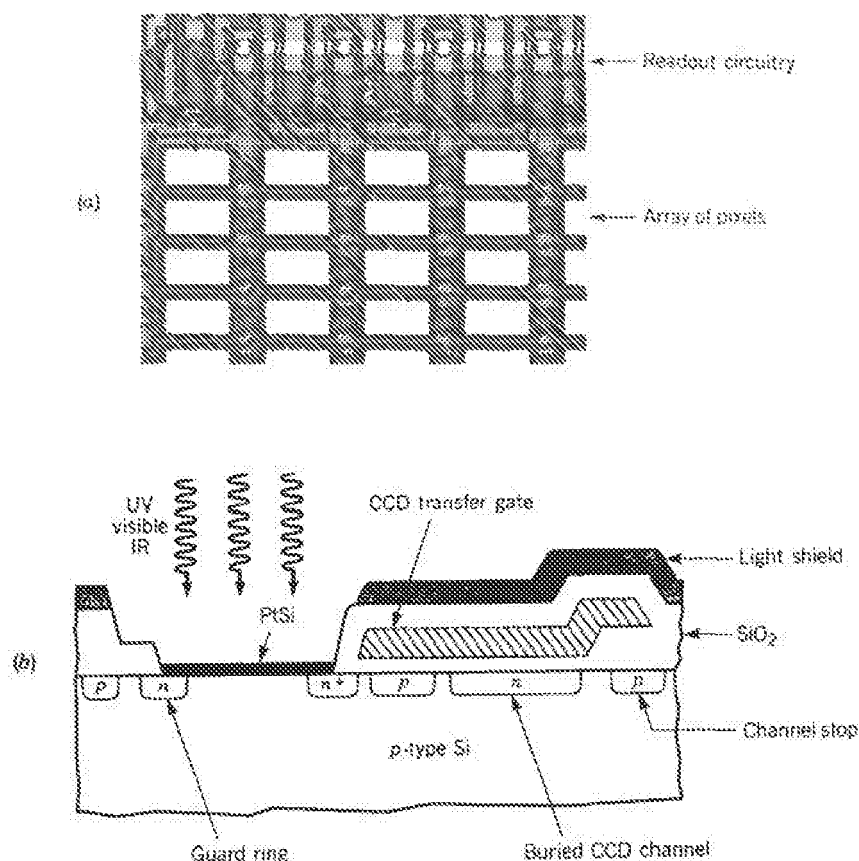


Figure 17.3-10 (a) Corner of an array of 160×244 PtSi/Si Schottky-barrier photodiodes. Each pixel is $40 \mu\text{m} \times 80 \mu\text{m}$ in size. Portions of the readout circuitry are visible. (Courtesy of W. F. Kosonocky.) (b) Cross section of a single pixel in the CCD array. The light shield prevents the generation of photocarriers in the CCD transfer gate and buried channel. The guard ring minimizes dark-current spikes and the channel stop confines the signal charge in the lateral direction. (Adapted from B.-Y. Tsaur, C. K. Chen, and J. P. Mattia, PtSi Schottky-Barrier Focal Plane Arrays for Multispectral Imaging in Ultraviolet, Visible and Infrared Spectral Bands, *IEEE Electron Device Letters*, vol. 11, pp. 162-164, 1990, copyright © IEEE.)

The quantum efficiency η ranges between 35% and 60% in the ultraviolet and visible regions (from $\lambda_0 = 290$ nm to about 900 nm) where the photon energy exceeds the bandgap of Si. At these wavelengths, the light transmitted through the PtSi film generates copious numbers of electron-hole pairs in the Si substrate [this is illustrated in Fig. 17.3-8(b) for a Schottky barrier with an n -type semiconductor]. At longer wavelengths, corresponding to photon energies below the bandgap of Si, the photogenerated carriers are produced by absorption in the PtSi film and η slowly decreases from about 3% at $1.5\text{ }\mu\text{m}$ to about 0.02% at $6\text{ }\mu\text{m}$. At all wavelengths, the array must be cooled to 77 K because of the low Schottky barrier height. However, similar devices have recently been fabricated from PtSi on n -type Si; these have a higher barrier height and can therefore be operated without cooling but they are only sensitive in the ultraviolet and visible. IrSi devices are also regularly used.

When illuminated, carriers with sufficient energy (holes in the p -type case) climb the Schottky barrier and enter the Si. This leaves a residue of negative charge (proportional to the number of photons absorbed by the pixel) to accumulate on the PtSi electrode. The electronic portion of the detection process is accomplished by transferring the negative charge from the PtSi electrode into a **charge-coupled device (CCD)** readout structure. The CCD transfer gate [Fig. 17.3-10(b)] permits the charge to be transferred to a buried CCD channel at a specified time. Many different kinds of electrode structures and clocking schemes have been developed for periodically reading out the charge accumulated by each pixel and thereby generating an electronic data stream representing the image.

The multispectral imaging capabilities of a Schottky-barrier-diode CCD array detector such as that described above is clearly illustrated in Fig. 17.3-11. The images are the radiation from a coffee mug that was partially filled with warm water and focused on the array by a lens. The left portion of the figure represents the infrared image (in the wavelength range from 3 to $5\text{ }\mu\text{m}$) whereas the right portion is the visible image obtained in room light using wavelengths shorter than $\approx 2\text{ }\mu\text{m}$. The infrared image clearly shows that the top of the mug and its handle are cooler than the rest. Of course, the photosensitive elements in a CCD array need not be Schottky-barrier diodes; p - n photodiodes are used as well.

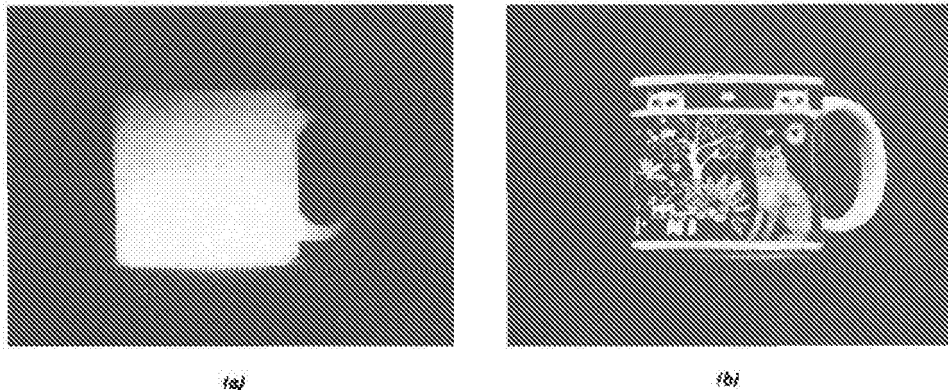


Figure 17.3-11 (a) Infrared and (b) visible images of a coffee mug partially filled with warm water obtained from a 160×244 -element PtSi/Si Schottky-barrier-diode focal-plane CCD array detector operated at 77 K. (After B.-Y. Tsaur, C. K. Chen, and J. P. Mattia, PtSi Schottky-Barrier Focal Plane Arrays for Multispectral Imaging in Ultraviolet, Visible and Infrared Spectral Bands, *IEEE Electron Device Letters*, vol. 11, pp. 162-164, 1990, copyright © IEEE.)

17.4 AVALANCHE PHOTODIODES

An **avalanche photodiode (APD)** operates by converting each detected photon into a *cascade* of moving carrier pairs. Weak light can then produce a current that is sufficient to be readily detected by the electronics following the APD. The device is a strongly reverse-biased photodiode in which the junction electric field is large; the charge carriers therefore accelerate, acquiring enough energy to excite new carriers by the process of **impact ionization**.

A. Principles of Operation

The history of a typical electron-hole pair in the depletion region of an APD is depicted in Fig. 17.4-1. A photon is absorbed at point 1, creating an electron-hole pair (an electron in the conduction band and a hole in the valence band). The electron accelerates under the effect of the strong electric field, thereby increasing its energy with respect to the bottom of the conduction band. The acceleration process is constantly interrupted by random collisions with the lattice in which the electron loses some of its acquired energy. These competing processes cause the electron to reach an average saturation velocity. Should the electron be lucky and acquire an energy larger than E_g at any time during the process, it has an opportunity to generate a second electron-hole pair by impact ionization (say at point 2). The two electrons then accelerate under the effect of the field, and each of them may be the source for a further impact ionization. The holes generated at points 1 and 2 also accelerate, moving toward the left. Each of these also has a chance of impact ionizing should they acquire sufficient energy, thereby generating a hole-initiated electron-hole pair (e.g., at point 3).

Ionization Coefficients

The abilities of electrons and holes to impact ionize are characterized by the **ionization coefficients** α_e and α_h . These quantities represent ionization probabilities per unit length (rates of ionization, cm^{-1}); the inverse coefficients, $1/\alpha_e$ and $1/\alpha_h$, represent

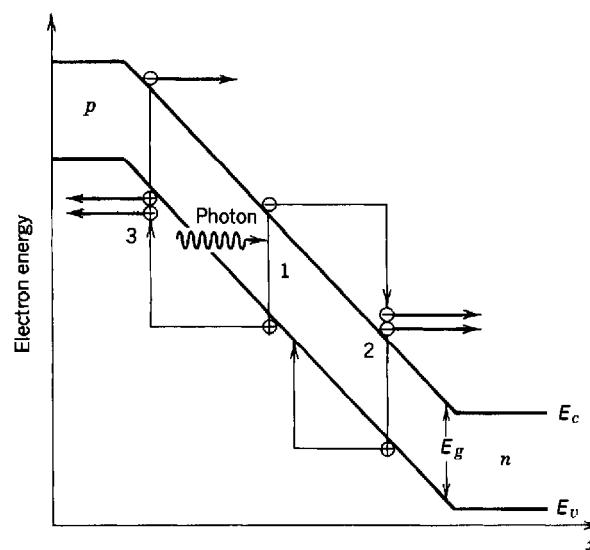


Figure 17.4-1 Schematic representation of the multiplication process in an APD.

the average distances between consecutive ionizations. The ionization coefficients increase with the depletion-layer electric field (since it provides the acceleration) and decrease with increasing device temperature. The latter occurs because increasing temperature causes an increase in the frequency of collisions, diminishing the opportunity a carrier has of gaining sufficient energy to ionize. The simple theory considered here assumes that α_e and α_h are constants that are independent of position and carrier history.

An important parameter for characterizing the performance of an APD is the **ionization ratio**

$$k = \frac{\alpha_h}{\alpha_e}.$$

When holes do not ionize appreciably [i.e., when $\alpha_h \ll \alpha_e$ ($k \ll 1$)], most of the ionization is achieved by electrons. The avalanching process then proceeds principally from left to right (i.e., from the p side to the n side) in Fig. 17.4-1. It terminates some time later when all the electrons arrive at the n side of the depletion layer. If electrons and holes both ionize appreciably ($k \approx 1$), on the other hand, those holes moving to the left create electrons that move to the right, which, in turn, generate further holes moving to the left, in a possibly unending circulation. Although this feedback process increases the gain of the device (i.e., the total generated charge in the circuit per photocarrier pair q/e), it is nevertheless undesirable for several reasons:

- It is time consuming and therefore reduces the device bandwidth.
- It is random and therefore increases the device noise.
- It can be unstable, thereby causing avalanche breakdown.

It is therefore desirable to fabricate APDs from materials that permit only one type of carrier (either electrons or holes) to impact ionize. If electrons have the higher ionization coefficient, for example, optimal behavior is achieved by injecting the electron of a photocarrier pair at the p edge of the depletion layer and by using a material whose value of k is as low as possible. If holes are injected, the hole of a photocarrier pair should be injected at the n edge of the depletion layer and k should be as large as possible. The ideal case of single-carrier multiplication is achieved when $k = 0$ or ∞ .

Design

As with any photodiode, the geometry of the APD should maximize photon absorption, for example by assuming the form of a $p-i-n$ structure. On the other hand, the multiplication region should be thin to minimize the possibility of localized uncontrolled avalanches (instabilities or microplasmas) being produced by the strong electric field. Greater electric-field uniformity can be achieved in a thin region.

These two conflicting requirements call for an APD design in which the absorption and multiplication regions are separate [**separate-absorption-multiplication (SAM) APD**]. Its operation is most readily understood by considering a device with $k \approx 0$ (e.g., Si). Photons are absorbed in a large intrinsic or lightly doped region. The photoelectrons drift across it under the influence of a moderate electric field, and finally enter a thin multiplication layer with a strong electric field where avalanching occurs. The reach-through $p^+-\pi-p-n^+$ APD structure illustrated in Fig. 17.4-2 accomplishes this. Photon absorption occurs in the wide π region (very lightly doped p region). Electrons drift through the π region into a thin $p-n^+$ junction, where they experience a sufficiently strong electric field to cause avalanching. The reverse bias applied across

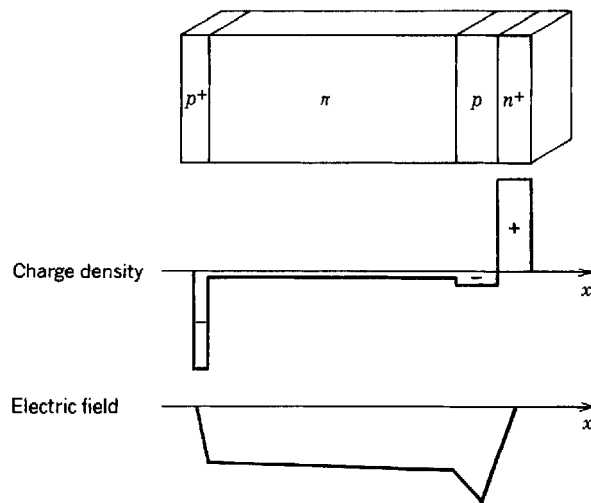


Figure 17.4-2 Reach-through $p^+-\pi-p-n^+$ APD structure.

the device is large enough for the depletion layer to reach through the p and π regions into the p^+ contact layer.

*Multilayer Devices

The noise inherent in the APD multiplication process can be reduced, at least in principle, by use of a multilayer avalanche photodiode. One such structure, called the staircase APD, has an energy-band diagram as shown in Fig. 17.4-3. A three-stage device is illustrated in both unbiased and reverse-biased conditions. The bandgap is

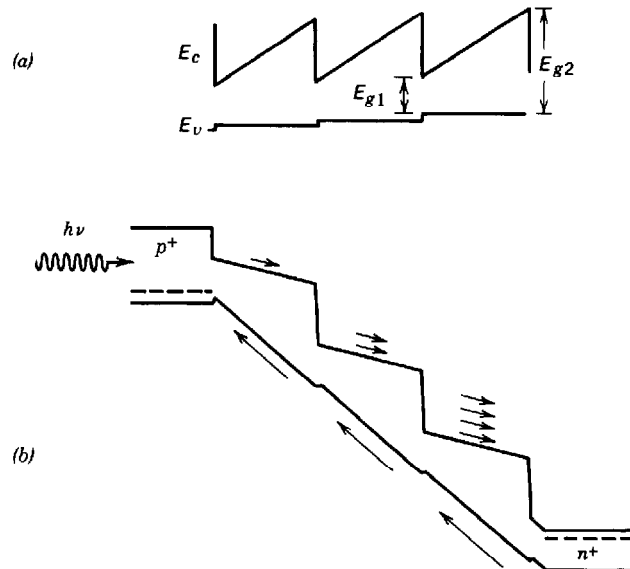


Figure 17.4-3 Energy-band diagram of a staircase APD under (a) unbiased and (b) reverse-biased conditions. The conduction-band steps encourage electron ionizations at discrete locations. (After F. Capasso, W. T. Tsang, and G. F. Williams, Staircase Solid-State Photomultipliers and Avalanche Photodiodes with Enhanced Ionization Rates Ratio, *IEEE Transactions on Electron Devices*, vol. ED-30, pp. 381–390, 1983, copyright © IEEE.)

compositionally graded (over a distance ≈ 10 nm), from a low value E_{g1} (e.g., GaAs) to a high value E_{g2} (e.g., AlGaAs). Because of the material properties, hole-induced ionizations are discouraged, thereby reducing the value of the ionization ratio k . Other potential advantages of such devices include the discrete locations of the multiplications (at the jumps in the conduction band edge), the low operating voltage, which minimizes tunneling, and the fast time response resulting from the reduced avalanche buildup time. Graded-gap devices of this kind are, however, difficult to fabricate.

B. Gain and Responsivity

As a prelude to determining the gain of an APD in which both kinds of carriers cause multiplication, the simpler problem of single-carrier (electron) multiplication ($\alpha_h = 0$, $k = 0$) is addressed first. Let $J_e(x)$ be the electric current density carried by electrons at location x , as shown in Fig. 17.4-4. Within a distance dx , on the average, the current is incremented by the factor

$$dJ_e(x) = \alpha_e J_e(x) dx,$$

from which we obtain the differential equation

$$\frac{dJ_e}{dx} = \alpha_e J_e(x),$$

whose solution is the exponential function $J_e(x) = J_e(0) \exp(\alpha_e x)$. The gain $G = J_e(w)/J_e(0)$ is therefore

$$G = \exp(\alpha_e w). \quad (17.4-1)$$

The electric current density increases exponentially with the product of the ionization coefficient α_e and the multiplication layer width w .

The double-carrier multiplication problem requires knowledge of both the electron current density $J_e(x)$ and the hole current density $J_h(x)$. It is assumed that only electrons are injected into the multiplication region. Since hole ionizations also produce electrons, however, the growth of $J_e(x)$ is governed by the differential equation

$$\frac{dJ_e}{dx} = \alpha_e J_e(x) + \alpha_h J_h(x). \quad (17.4-2)$$

As a result of charge neutrality, $dJ_e/dx = -dJ_h/dx$, so that the sum $J_e(x) + J_h(x)$ must remain constant for all x under steady-state conditions. This is clear from Fig.

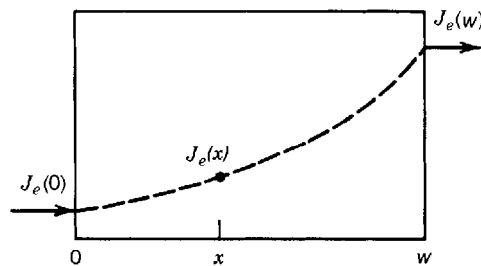


Figure 17.4-4 Exponential growth of the electric current density in a single-carrier APD.

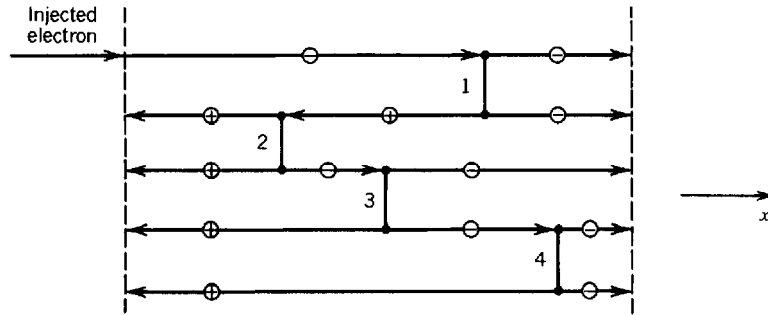


Figure 17.4-5 Constancy of the sum of the electron and hole current densities across a plane at any x .

17.4-5; the total number of charge carriers crossing any plane is the same regardless of position (four impact ionizations and five electrons-plus-holes crossing every plane are shown by way of illustration).

Since it is assumed that no holes are injected at $x = w$, $J_h(w) = 0$, so that

$$J_e(x) + J_h(x) = J_e(w), \quad (17.4-3)$$

as shown in Fig. 17.4-6. $J_h(x)$ can therefore be eliminated in (17.4-2) to obtain

$$\frac{dJ_e}{dx} = (\alpha_e - \alpha_h)J_e(x) + \alpha_h J_e(w). \quad (17.4-4)$$

This first-order differential equation is readily solved for the gain $G = J_e(w)/J_e(0)$. For $\alpha_e \neq \alpha_h$, the result is $G = (\alpha_e - \alpha_h) / \{\alpha_e \exp[-(\alpha_e - \alpha_h)w] - \alpha_h\}$, from which

$$G = \frac{1 - k}{\exp[-(1 - k)\alpha_e w] - k}. \quad (17.4-5)$$

APD Gain

The single-carrier multiplication result (17.4-1), with its exponential growth, is recovered when $k = 0$. When $k = \infty$, the gain remains unity since only electrons are injected and electrons do not multiply. For $k = 1$, (17.4-5) is indeterminate and the gain must be obtained directly from (17.4-4); the result is then $G = 1/(1 - \alpha_e w)$. An instability is reached when $\alpha_e w = 1$. The dependence of the gain on $\alpha_e w$ for several values of the

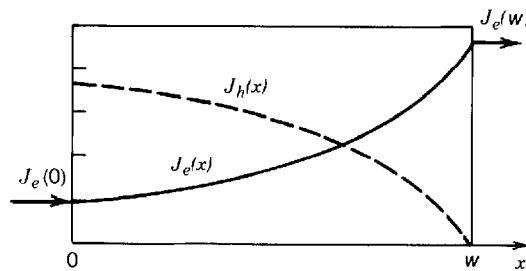


Figure 17.4-6 Growth of the electron and hole currents as a result of avalanche multiplication.

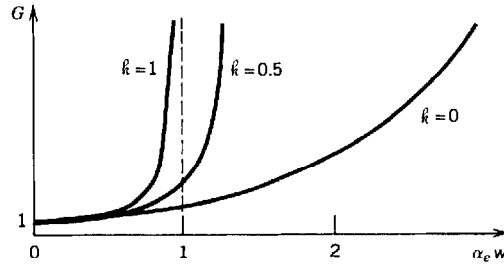


Figure 17.4-7 Growth of the gain G with multiplication-layer width for several values of the ionization ratio k , assuming pure electron injection.

ionization ratio k is illustrated in Fig. 17.4-7. The responsivity \mathfrak{R} is obtained by using (17.4-5) in the general relation (17.1-6).

The materials of interest are the same as those for photodiodes, with the additional proviso that they should have the lowest (or highest) possible value of ionization ratio k . APDs with values of k as low as 0.006 have been fabricated from silicon, providing excellent performance in the wavelength region 0.7 to 0.9 μm .

C. Response Time

Aside from the usual transit, diffusion, and RC effects that govern the response time of photodiodes, APDs suffer from an additional multiplication time called the **avalanche buildup time**. The response time of a two-carrier-multiplication APD is illustrated in Fig. 17.4-8 by following the history of a photoelectron generated at the edge of the absorption region (point 1). The electron drifts with a saturation velocity v_e , reaching the multiplication region (point 2) after a transit time w_d/v_e . Within the multiplication region the electron also travels with a velocity v_e . Through impact ionization it creates electron-hole pairs, say at points 3 and 4, generating two additional electron-hole pairs. The holes travel in the opposite direction with their saturation velocity v_h . The holes can also cause impact ionizations resulting in electron-hole pairs as shown, for example, at points 5 and 6. The resulting carriers can themselves cause impact ionizations, sustaining the feedback loop. The process is terminated when the last hole leaves the multiplication region (at point 7) and crosses the drift region to point 8. The total time τ required for the entire process (between points 1 and 8) is the sum of the transit times (from 1 to 2 and from 7 to 8) and the multiplication time denoted τ_m ,

$$\tau = \frac{w_d}{v_e} + \frac{w_d}{v_h} + \tau_m. \quad (17.4-6)$$

Because of the randomness of the multiplication process, the multiplication time τ_m is random. In the special case $k = 0$ (no hole multiplication) the maximum value of τ_m is readily seen from Fig. 17.4-8 to be

$$\tau_m = \frac{w_m}{v_e} + \frac{w_m}{v_h}. \quad (17.4-7)$$

For a large gain G , and for electron injection with $0 < k < 1$, an order of magnitude of

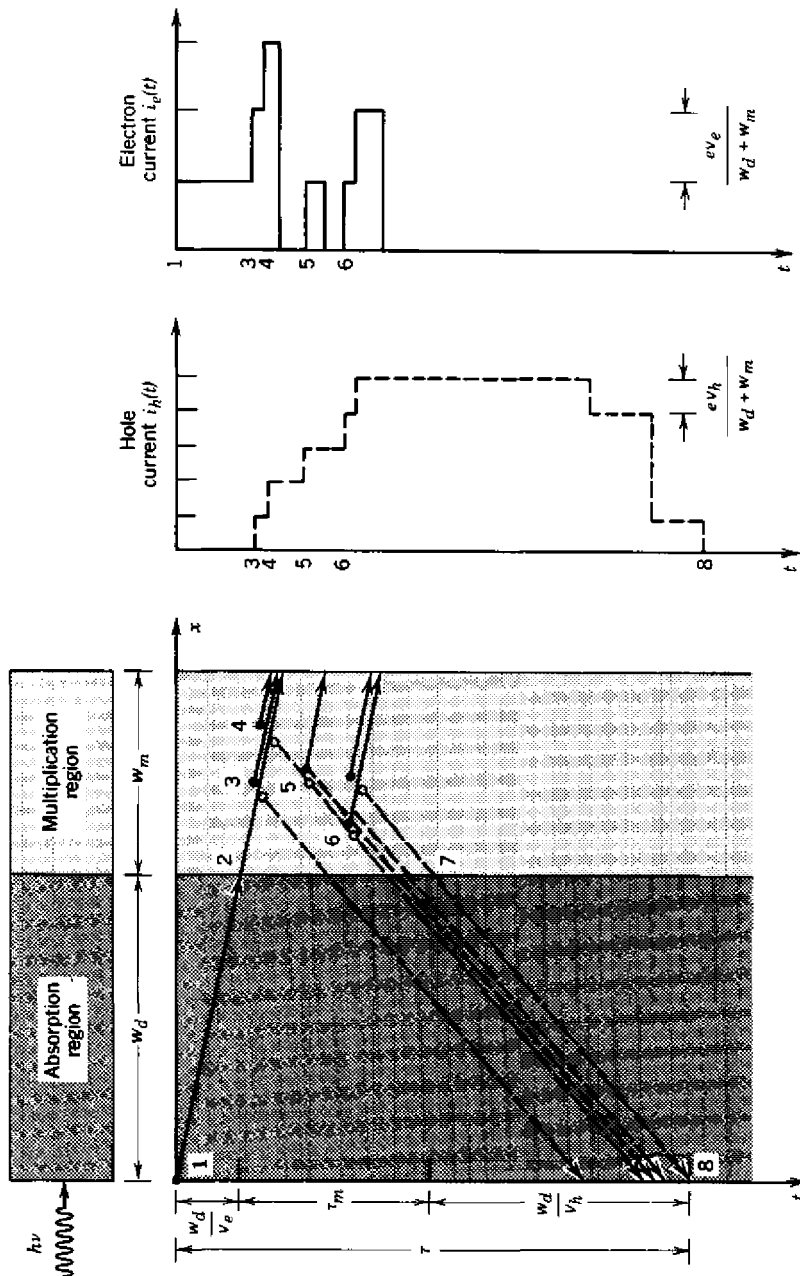


Figure 17.4-8 (a) Tracing the course of the avalanche buildup time in an APD with the help of a position-time graph. The solid lines represent electrons, and the dashed lines represent holes. Electrons move to the right with velocity v_e and holes move to the left with velocity v_h . Electron-hole pairs are produced in the multiplication region. The carriers cease moving when they reach the edge of the material. (b) Hole current $i_h(t)$ and electron current $i_e(t)$ induced in the circuit. Each carrier pair induces a charge e in the circuit. The total induced charge q , which is the area under the $i_e(t) + i_h(t)$ versus t curve, is Ge . This figure is a generalization of Fig. 17.1-3, which applies for a single electron-hole pair.

the average value of τ_m is obtained by increasing the first term of (17.4-7) by the factor Gk ,

$$\tau_m \approx \frac{Gk w_m}{v_e} + \frac{w_m}{v_h}. \quad (17.4-8)$$

A more accurate theory is rather complex.

EXAMPLE 17.4-1. Avalanche Buildup Time in a Si APD. Consider a Si APD with $w_d = 50 \mu\text{m}$, $w_m = 0.5 \mu\text{m}$, $v_e = 10^7 \text{ cm/s}$, $v_h = 5 \times 10^6 \text{ cm/s}$, $G = 100$, and $k = 0.1$. Equation (17.4-7) yields $\tau_m = 5 + 10 = 15 \text{ ps}$, whereupon (17.4-6) gives $\tau = 1020 \text{ ps} = 1.02 \text{ ns}$. On the other hand, (17.4-8) yields $\tau_m = 60 \text{ ps}$, so that (17.4-6) provides $\tau = 1065 \text{ ps} = 1.07 \text{ ns}$. For a p - i - n photodiode with the same values of w_d , v_e , and v_h , the transit time is $w_d/v_e + w_d/v_h \approx 1 \text{ ns}$. These results do not differ greatly because τ_m is quite low in a silicon device.

17.5 NOISE IN PHOTODETECTORS

The photodetector is a device that measures photon flux (or optical power). Ideally, it responds to a photon flux Φ (optical power $P = h\nu\Phi$) by generating a proportional electric current $i_p = \eta e\Phi = \Re P$ [see (17.1-2)]. In actuality, the device generates a *random* electric current i whose value fluctuates above and below its average, $\bar{i} \equiv i_p = \eta e\Phi = \Re P$. These random fluctuations, which are regarded as noise, are characterized by the standard deviation σ_i , where $\sigma_i^2 = \langle (i - \bar{i})^2 \rangle$. For a current of zero mean ($\bar{i} = 0$), the standard deviation is the same as the root-mean-square (rms) value of the current, i.e., $\sigma_i = \langle i^2 \rangle^{1/2}$.

Several sources of noise are inherent in the process of photon detection:

- **Photon Noise.** The most fundamental source of noise is associated with the random arrivals of the photons themselves (which are usually described by Poisson statistics), as discussed in Sec. 11.2.
- **Photoelectron Noise.** For a photon detector with quantum efficiency $\eta < 1$, a single photon generates a photoelectron-hole pair with probability η but fails to do so with probability $1 - \eta$. Because of the inherent randomness in this process of carrier generation, it serves as a source of noise.
- **Gain Noise.** The amplification process that provides internal gain in some photodetectors (such as APDs) is random. Each detected photon generates a random number G of carriers with an average value \bar{G} but with an uncertainty that is dependent on the nature of the amplification mechanism.
- **Receiver Circuit Noise.** The various components in the electrical circuitry of an optical receiver, such as resistors and transistors, contribute to the receiver circuit noise.

These four sources of noise are illustrated schematically in Fig. 17.5-1. The signal entering the detector (input signal) has an intrinsic photon noise. The photoeffect converts the photons into photoelectrons. In the process, the mean signal decreases by the factor η . The noise also decreases but by a lesser amount than the signal; thus the

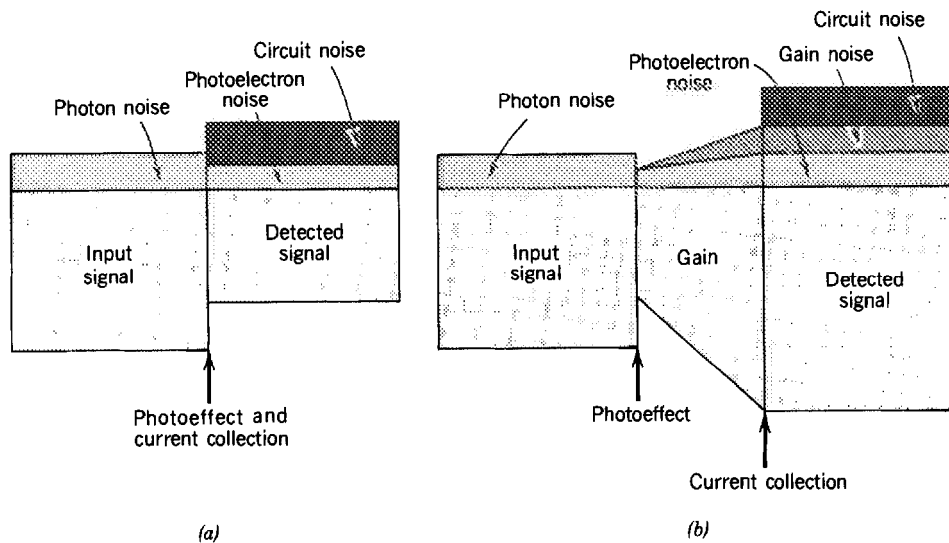


Figure 17.5-1 Signal and various noise sources for (a) a photodetector without gain (e.g., a *p-i-n* photodiode) and (b) a photodetector with gain (e.g., an APD).

signal-to-noise ratio of the photoelectrons is lower than that of the incident photons. If a photodetector gain mechanism is present, it amplifies both the signal and the photoelectron noise, and introduces its own gain noise as well. Finally, circuit noise enters at the point of current collection.

An optical receiver as a component in an information transmission system can be characterized by the following performance measures:

- The **signal-to-noise ratio (SNR)**. The SNR of a random variable is defined as $\text{SNR} = (\text{mean})^2 / \text{variance}$; thus the SNR of the current i is $\text{SNR} = \bar{i}^2 / \sigma_i^2$, whereas the SNR of the photon number is $\text{SNR} = \bar{n}^2 / \sigma_n^2$.
- The **minimum-detectable signal**, which is defined as the mean signal that yields $\text{SNR} = 1$.
- The **receiver sensitivity**, which, like the minimum-detectable signal, is defined as the signal corresponding to a prescribed $\text{SNR} = \text{SNR}_0$. Rather than selecting $\text{SNR}_0 = 1$, however, a higher value is usually chosen to ensure a good level of accuracy (e.g., $\text{SNR}_0 = 10$ to 10^3 , corresponding to 10 to 30 dB).

We proceed to derive expressions for the signal-to-noise ratio (SNR) for optical detectors with these sources of noise. Other sources of noise that are not explicitly considered here include background noise and dark-current noise. **Background noise** is the photon noise associated with light reaching the detector from extraneous sources (i.e., from sources other than the signal of interest, such as sunlight and starlight). Background noise is particularly deleterious in middle- and far-infrared detection systems because objects at room temperature emit copious thermal radiation in this region (see Fig. 12.3-4). Photodetection devices also generate **dark-current noise**, which, as the name implies, is present even in the absence of light. Dark-current noise results from random electron-hole pairs generated thermally or by tunneling. Also ignored are leakage current and $1/f$ noise.

A. Photoelectron Noise

Photon Noise

As described in Sec. 11.2, the photon flux associated with a fixed optical power P is inherently uncertain. The mean photon flux $\Phi = P/h\nu$ (photons/s), but it fluctuates randomly in accordance with a probability law dependent on the nature of the light source. The number of photons n counted in a time interval T is random with mean $\bar{n} = \Phi T$. The photon number for light from an ideal laser, or from a thermal source of spectral width much greater than $1/T$, obeys the Poisson probability distribution, for which $\sigma_n^2 = \bar{n}$. Thus the fluctuations associated with an average of 100 photons cause the actual number of photons to lie approximately within the range 100 ± 10 .

The photon-number signal-to-noise ratio \bar{n}^2/σ_n^2 is therefore

$$\text{SNR} = \bar{n}$$

(17.5-1)

Photon-Number
Signal-to-Noise Ratio

and the minimum detectable photon number is $\bar{n} = 1$ photon. If the observation time $T = 1 \mu\text{s}$ and the wavelength $\lambda_o = 1.24 \mu\text{m}$, this is equivalent to a minimum detectable power of 0.16 pW. The receiver sensitivity, the signal required to attain $\text{SNR} = 10^3$ (30 dB), is 1000 photons. If the time interval $T = 10 \text{ ns}$, this is equivalent to a sensitivity of 10^{11} photons/s or an optical power sensitivity of 16 nW (at $\lambda_o = 1.24 \mu\text{m}$).

Photoelectron Noise

A photon incident on a photodetector of quantum efficiency η either generates a photoevent (i.e., liberates a photoelectron or creates a photoelectron-hole pair) with probability η , or fails to do so with probability $1 - \eta$. Photoevents are assumed to be selected at random from the photon stream so that an incident mean photon flux Φ (photons/s) results in a mean photoelectron flux $\eta\Phi$ (photoelectrons/s). The number of photoelectrons detected in the time interval T is a random number m with mean

$$\bar{m} = \eta\bar{n}, \quad (17.5-2)$$

where $\bar{n} = \Phi T$ is the mean number of incident photons in the same time interval T . If the photon number is distributed in Poisson fashion, so is the photoelectron number, as can be ascertained by using an argument similar to that developed in Sec. 11.2D. It follows that the photoelectron-number variance is then precisely equal to \bar{m} , so that

$$\sigma_m^2 = \bar{m} = \eta\bar{n}. \quad (17.5-3)$$

It is apparent from this relationship that the photoelectron noise and the photon noise are nonadditive.

The underlying randomness inherent in the photon number, which constitutes a fundamental source of noise with which we must contend when using light to transmit a signal, therefore results in a photoelectron-number signal-to-noise ratio

$$\text{SNR} = \bar{m} = \eta\bar{n}.$$

(17.5-4)

Photoelectron-Number
Signal-to-Noise Ratio

The minimum-detectable photoelectron number for $\text{SNR} = 1$ corresponds to $\bar{m} = \eta\bar{n} = 1$ (i.e., one photoelectron or $1/\eta$ photons). The receiver sensitivity for $\text{SNR} = 10^3$ is 1000 photoelectrons or $1000/\eta$ photons.

Photocurrent Noise

We now examine the properties of the electric current $i(t)$ induced in a circuit by a random photoelectron flux with mean $\eta\Phi$. The treatment we provide includes the effects of photon noise, photoelectron noise, and the characteristic time response of the detector and circuitry (filtering). Every photoelectron-hole pair generates a pulse of electric current of charge (area) e and time duration τ_p in the external circuit of the photodetector (Fig. 17.5-2). A photon stream incident on a photodetector therefore results in a stream of electrical pulses which add together to constitute an electric current $i(t)$. The randomness of the photon stream is transformed into a fluctuating electric current. If the incident photons are Poisson distributed, these fluctuations are known as shot noise. More generally, for detectors with gain G , the generated charge in each pulse is $q = Ge$.

Before providing an analytical derivation, we first show in a simplified way that the photocurrent i in a circuit of bandwidth B , generated by a photon flux Φ , can be determined by considering a characteristic time interval $T = 1/2B$ (the resolution time of the circuit) and by relating the random number m of photoelectrons counted within that interval to the photocurrent $i(t)$, where t is the instant of time immediately at the end of the interval T . For rectangular current pulses of duration T , the current and photoelectron-number random variables are then related by $i = (e/T)m$, so that the mean and variance are given by

$$\bar{i} = \frac{e}{T} \bar{m}$$

$$\sigma_i^2 = \left(\frac{e}{T} \right)^2 \sigma_m^2,$$

where $\bar{m} = \eta\Phi T = \eta\Phi/2B$ is the number of photoelectrons collected in the time interval $T = 1/2B$. Substituting $\sigma_m^2 = \bar{m}$ for the Poisson law yields the photocurrent mean and variance:

$$\bar{i} = e\eta\Phi$$

(17.5-5)
Photocurrent Mean

$$\sigma_i^2 = 2e\bar{i}B.$$

(17.5-6)
Photocurrent Variance

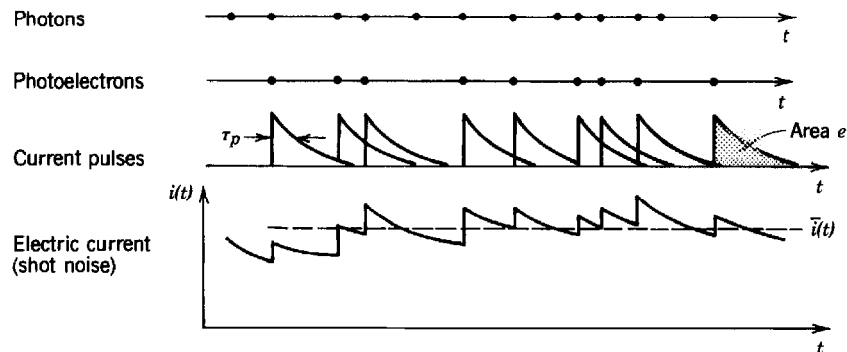


Figure 17.5-2 The electric current in a photodetector circuit comprises a superposition of electrical pulses, each associated with a detected photon. The individual pulses illustrated are exponential but they can assume an arbitrary shape (see, e.g., Figs. 17.1-3(b) and 17.1-4).

It follows that the signal-to-noise ratio of the photoelectric current, $\text{SNR} = \bar{i}^2 / \sigma_i^2$, is

$$\text{SNR} = \frac{\bar{i}}{2eB} = \frac{\eta\Phi}{2B} = \bar{m}. \quad (17.5-7)$$

Photocurrent
Signal-to-Noise Ratio

The SNR is directly proportional to the photon flux Φ and inversely proportional to the electrical bandwidth of the circuit B .

EXAMPLE 17.5-1. SNR and Receiver Sensitivity. For $\bar{i} = 10$ nA and $B = 100$ MHz, $\sigma_i \approx 0.57$ nA, corresponding to a signal-to-noise ratio $\text{SNR} = 310$ or 25 dB. An average of 310 photoelectrons are detected in every time interval $1/2B = 5$ ns. The minimum-detectable photon flux is $\Phi = 2B/\eta$, and the receiver sensitivity for $\text{SNR} = 10^3$ is $\Phi = 1000(2B/\eta) = 2 \times 10^{11}/\eta$ photons/s.

Equations (17.5-5) and (17.5-6) will now be proved for current pulses of arbitrary shape.

Derivation of the Photocurrent Mean and Variance

Assume that a photoevent generated at $t = 0$ produces an electric pulse $h(t)$, of area e , in the external circuit. A photoelectron generated at time t_1 then produces a displaced pulse $h(t - t_1)$. If the time axis is divided into incremental time intervals Δt , the probability p that a photoevent occurs within an interval is $p = \eta\Phi\Delta t$. The electric current i at time t is written as

$$i(t) = \sum_l X_l h(t - l\Delta t), \quad (17.5-8)$$

where X_l has either the value 1 with probability p , or 0 with probability $1 - p$. The variables $\{X_l\}$ are independent. The mean value of X_l is $0 \times (1 - p) + 1 \times p = p$. Its mean-square value is $\langle X_l^2 \rangle = 0^2 \times (1 - p) + 1^2 \times p = p$. The mean of the product $X_l X_k$ is p^2 if $l \neq k$, and p if $l = k$. The mean and mean-square values of $i(t)$ are now determined as follows:

$$\bar{i} = \langle i \rangle = \sum_l p h(t - l\Delta t) \quad (17.5-9)$$

$$\begin{aligned} \langle i^2 \rangle &= \sum_l \sum_k \langle X_l X_k \rangle h(t - l\Delta t) h(t - k\Delta t) \\ &= \sum_{l \neq k} p^2 h(t - l\Delta t) h(t - k\Delta t) + \sum_l p h^2(t - l\Delta t). \end{aligned} \quad (17.5-10)$$

Substituting $p = \eta\Phi\Delta t$, and taking the limit $\Delta t \rightarrow 0$, so that the summations become

integrals, (17.5-9) and (17.5-10) yield

$$\bar{i} = \eta\Phi \int_0^\infty h(t) dt = e\eta\Phi \quad (17.5-11)$$

$$\langle i^2 \rangle = (e\eta\Phi)^2 + \eta\Phi \int_0^\infty h^2(t) dt. \quad (17.5-12)$$

It follows that the variance of i is $\sigma_i^2 = \langle i^2 \rangle - \langle i \rangle^2$, or

$$\sigma_i^2 = \eta\Phi \int_0^\infty h^2(t) dt. \quad (17.5-13)$$

Defining

$$B = \frac{1}{2e^2} \int_0^\infty h^2(t) dt = \frac{1}{2} \frac{\int_0^\infty h^2(t) dt}{\left[\int_0^\infty h(t) dt \right]^2}, \quad (17.5-14)$$

we finally obtain (17.5-5) and (17.5-6).

The parameter B defined by (17.5-14) represents the device-circuit bandwidth. This is readily verified by noting that the Fourier transform of $h(t)$ is its transfer function $\mathcal{H}(\nu)$. The area under $h(t)$ is simply $\mathcal{H}(0) = e$. In accordance with Parseval's theorem [see (A.1-7) in Appendix A], the area under $h^2(t)$ is equal to the area under the symmetric function $|\mathcal{H}(\nu)|^2$, so that

$$B = \int_0^\infty \left| \frac{\mathcal{H}(\nu)}{\mathcal{H}(0)} \right|^2 d\nu. \quad (17.5-15)$$

The quantity B is therefore the power-equivalent spectral width of the function $|\mathcal{H}(\nu)|$ (i.e., the bandwidth of the device-circuit combination), in accordance with (A.2-10) of Appendix A. As an example, if $\mathcal{H}(\nu) = 1$ for $-\nu_c < \nu < \nu_c$ and 0 elsewhere, (17.5-15) yields $B = \nu_c$.

These relations are applicable for all photoelectric detection devices without gain (e.g., phototubes and junction photodiodes). Use of the formulas requires knowledge of the bandwidth of the device, biasing circuit, and amplifier; B is determined by inserting the transfer function of the overall system into (17.5-15).

B. Gain Noise

The photocurrent mean and variance for a device with fixed (deterministic) gain G is determined by replacing e with $q = Ge$ in (17.5-5) and (17.5-6):

$$\bar{i} = eG\eta\Phi = \frac{eG\eta P}{h\nu} \quad (17.5-16)$$

$$\sigma_i^2 = 2eGiB = 2e^2G^2\eta B\Phi. \quad (17.5-17)$$

The signal-to-noise ratio, which is given by

$$\text{SNR} = \frac{\bar{i}}{2eGB} = \frac{\eta\Phi}{2B} = \bar{m}, \quad (17.5-18)$$

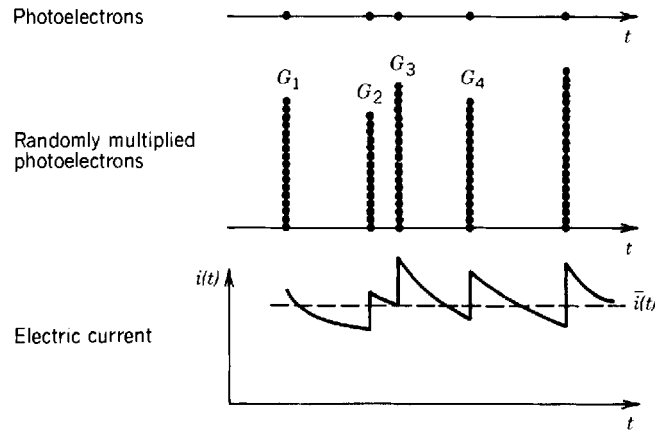


Figure 17.5-3 Each photoelectron-hole pair in a photodetector with gain generates a random number G_l of electron-hole pairs, each of which produces an electrical current pulse of area eG_l in the detector circuit. The total electric current $i(t)$ is the superposition of these pulses.

then turns out to be independent of G . This is because the mean current \bar{i} and its rms value σ_i are both increased by the same factor G as a result of the gain.

This simple result does not apply when the gain itself is random, as is the case in a photomultiplier tube, a photoconductor, and an avalanche photodiode. The derivation of the photocurrent mean and variance given in the previous section must be generalized to account for the randomness in G . The electric current (17.5-8) should then be written in the form

$$i(t) = \sum_l X_l G_l h(t - l \Delta t),$$

where, as before, X_l takes the value 1 with probability $p = \eta \Phi \Delta t$, and 0 with probability $1 - p$. Included now are the G_l , which are independent random numbers representing the gain imparted to a photoelectron-hole pair generated in the l th time slot. The process is illustrated in Fig. 17.5-3. If the random variable G_l has mean value $\langle G \rangle = \bar{G}$ and mean-square value $\langle G^2 \rangle$, an analysis similar to that provided in (17.5-8) through (17.5-14) leads to

$$\bar{i} = e \bar{G} \eta \Phi$$

(17.5-19)
Photocurrent Mean
(Detector with Random Gain)

$$\sigma_i^2 = 2e \bar{G} i B F,$$

(17.5-20)
Photocurrent Variance
(Detector with Random Gain)

where

$$F = \frac{\langle G^2 \rangle}{\langle G \rangle^2}$$

(17.5-21)
Excess Noise Factor

is called the **excess noise factor**.

The excess noise factor is related to the variance of the gain σ_G^2 by the relation $F = 1 + \sigma_G^2 / \langle G \rangle^2$. When the gain is deterministic $\sigma_G^2 = 0$ and $F = 1$ so that (17.5-20) properly reduces to (17.5-17). When the gain is random, $\sigma_G^2 > 0$ and $F > 1$; both increase with the severity of the gain fluctuations. The resulting electric current i is then more noisy than shot noise.

In the presence of random gain, the signal-to-noise ratio i^2 / σ_i^2 becomes

$$\text{SNR} = \frac{i}{2e\bar{G}BF} = \frac{\eta\Phi/2B}{F} = \frac{\bar{m}}{F}, \quad (17.5-22)$$

Signal-to-Noise Ratio
(Detector with
Random Gain)

where \bar{m} is the mean number of photoelectrons collected in the time $T = 1/2B$. This is smaller than the deterministic-gain expression by the factor F ; the reduction in the SNR arises from the randomness of the gain.

Excess Noise Factor for an APD

When photoelectrons are injected at the edge of a uniformly multiplying APD, the gain G of the device is given by (17.4-5). It depends on the electron ionization coefficient α_e and the ionization ratio $k = \alpha_h / \alpha_e$, as well as on the width of the multiplication region w . The use of a similar (but more complex) analysis, incorporating the randomness associated with the gain process, leads to an expression for the mean-square gain $\langle G^2 \rangle$, and therefore for the excess noise factor F . This more general derivation gives rise to an expression for the mean gain \bar{G} which is identical to that given by (17.4-5). The excess noise factor F turns out to be related to the mean gain and the ionization ratio by

$$F = k\bar{G} + (1 - k)\left(2 - \frac{1}{\bar{G}}\right). \quad (17.5-23)$$

Excess Noise Factor
for an APD

A plot of this result is presented in Fig. 17.5-4.

Equation (17.5-23) is valid when electrons are injected at the edge of the depletion layer, but both electrons and holes have the capability of initiating impact ionizations. If only holes are injected, the same expression applies, provided that k is replaced by $1/k$. Gain noise is minimized by injecting the carrier with the higher ionization coefficient, and by fabricating a structure with the lowest possible value of k if electrons are injected, or the highest possible value of k if holes are injected. Thus the ionization coefficients for the two carriers should be as different as possible.

Equation (17.5-23) is said to be valid under conditions of single-carrier-initiated double-carrier multiplication since both types of carrier have the capacity to impact ionize, even when only one type is injected. If electrons and holes are injected simultaneously, the overall result is the sum of the two partial results.

The gain noise introduced in a conventional APD arises from two sources: the randomness in the locations at which ionizations may occur and the feedback process associated with the fact that both kinds of carrier can produce impact ionizations. The first of these sources of noise is present even when only one kind of carrier can multiply; it gives rise to a minimum excess noise factor $F = 2$ at large values of the mean gain \bar{G} , as is apparent by setting $k = 0$ and letting \bar{G} become large in (17.5-23). The second source of noise (the feedback process) is potentially more detrimental since it can result in a far larger increase in F . In a photomultiplier tube, there is only one

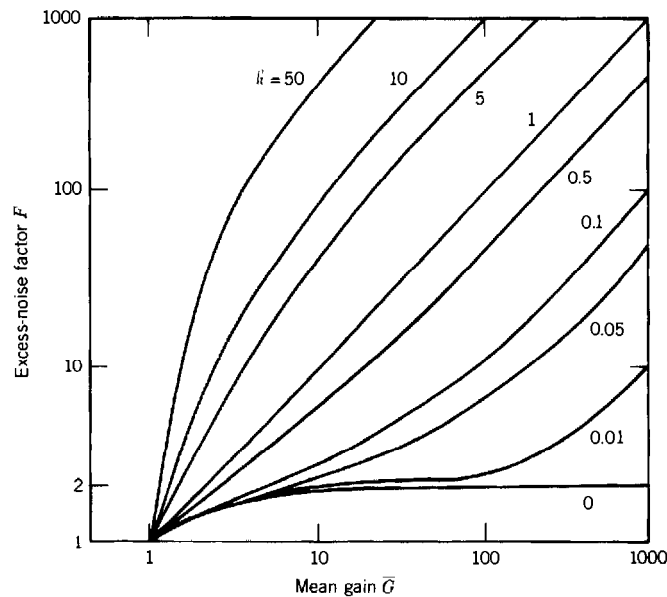


Figure 17.5-4 Excess noise factor F for an APD, under electron injection, as a function of the mean gain \bar{G} for different values of the ionization ratio k . For hole injection, $1/k$ replaces k .

kind of carrier (electrons) and ionizations always occur at the dynodes, so there is no location randomness. In general, therefore, $1 \lesssim F < 2$ for photomultiplier tubes.

EXAMPLE 17.5-2. Excess Noise Factor of a Si APD. A Si APD (for which $k \approx 0.1$) with $\bar{G} = 100$ (and electron injection) has an excess noise factor $F = 11.8$. Use of the APD therefore increases the mean value of the detected current by a factor of 100, while reducing the signal-to-noise ratio by a factor of 11.8. We show in the next section, however, that in the presence of circuit noise the use of an APD can increase the overall SNR.

*Excess Noise Factor for a Multilayer APD

In principle, both sources of gain noise (location randomness and feedback) can be eliminated by the use of a multilayer avalanche photodiode structure. One example of such a structure, the graded-gap staircase device, is illustrated in Fig. 17.4-3. An electron in the conduction band can gain sufficient energy to promote an impact ionization only at discrete locations (at the risers in the staircase), which eliminates the location randomness. The feedback noise is reduced by the valence-band-edge discontinuity, which is in the wrong direction for fostering impact ionizations and therefore should result in a small value of k . In theory, totally noise-free multiplication ($F = 1$) can be achieved. As indicated earlier, however, devices of this type are very difficult to fabricate.

C. Circuit Noise

Yet additional noise is introduced by the electronic circuitry associated with an optical receiver. Circuit noise results from the thermal motion of charged carriers in resistors

and other dissipative elements (thermal noise), and from fluctuations of charge carriers in transistors used in the receiver amplifier.

Thermal Noise

Thermal noise (also called **Johnson noise** or **Nyquist noise**) arises from the random motions of mobile carriers in resistive electrical materials at finite temperatures; these motions give rise to a random electric current $i(t)$ even in the absence of an external electrical power source. The thermal electric current in a resistance R is therefore a random function $i(t)$ whose mean value $\langle i(t) \rangle = 0$, i.e., it is equally likely to be in either direction. The variance of the current σ_i^2 (which is the same as the mean-square value since the mean vanishes) increases with the temperature T .

Using an argument based on statistical mechanics, presented subsequently, it can be shown that a resistance R at temperature T exhibits a random electric current $i(t)$ characterized by a power spectral density (defined in Sec. 10.1B)

$$S_i(f) = \frac{4}{R} \frac{hf}{\exp(hf/k_B T) - 1}, \quad (17.5-24)$$

where f is the frequency. In the region $f \ll k_B T/h$, which is of principal interest since $k_B T/h = 6.25$ THz at room temperature, $\exp(hf/k_B T) \approx 1 + hf/k_B T$ so that

$$S_i(f) \approx 4k_B T/R. \quad (17.5-25)$$

The variance of the electric current is the integral of the power spectral density over all frequencies within the bandwidth B of the circuit, i.e.,

$$\sigma_i^2 = \int_0^B S_i(f) df.$$

When $B \ll k_B T/h$, we obtain

$$\sigma_i^2 \approx 4k_B TB/R.$$

(17.5-26)

Thermal Noise
Current Variance
in a Resistance R

Thus, as shown in Fig. 17.5-5, a resistor R at temperature T in a circuit of bandwidth B behaves as a noiseless resistor in parallel with a source of noise current with zero mean and an rms value σ_i determined by (17.5-26).

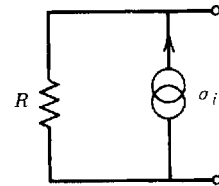


Figure 17.5-5 A resistance R at temperature T is equivalent to a noiseless resistor in parallel with a noise current source of variance $\sigma_i^2 = \langle i^2 \rangle \approx 4k_B TB/R$, where B is the circuit bandwidth.

EXAMPLE 17.5-3. Thermal Noise in a Resistor. A 1-k Ω resistor at $T = 300$ K, in a circuit of bandwidth $B = 100$ MHz, exhibits an rms thermal noise current $\sigma_i \approx 41$ nA.

*Derivation of the Power Spectral Density of Thermal Noise

We now derive (17.5-24) by showing that the electrical power associated with the thermal noise in a resistance is identical to the electromagnetic power radiated by a one-dimensional blackbody. The factor $hf/[\exp(hf/k_B T) - 1]$ in (17.5-24) is recognized as the mean energy \bar{E} of an electromagnetic mode of frequency f (the symbol ν is reserved for optical frequencies) in thermal equilibrium at temperature T [see (12.3-8)]. This equation may therefore be written as $S_i(f)R = 4\bar{E}$. The electrical power dissipated by a noise current i passing through a resistance R is $\langle i^2 \rangle R = \sigma_i^2 R$, so that the term $S_i(f)R$ represents the electrical power density (per Hz) dissipated by the noise current $i(t)$ through R . We proceed to demonstrate that $4\bar{E}$ is the power density radiated by a one-dimensional blackbody.

As discussed in Sec. 12.3B, an atomic system in thermal equilibrium with the electromagnetic modes in a cavity radiates a spectral energy density $\varrho(\nu) = M(\nu)\bar{E}$, where $M(\nu) = 8\pi\nu^2/c^3$ is the three-dimensional density of modes, and the spectral intensity density is $c\varrho(\nu)$. Although the charge carriers in a resistor move in all directions, only motion in the direction of the circuit current flow contributes. The density of modes in a single dimension is $M(f) = 4/c$ modes/m-Hz [see (9.1-7)] so that the corresponding energy density is $\varrho(f) = M(f)\bar{E} = 4\bar{E}/c$ and the radiated power density is $c\varrho(f) = 4\bar{E}$ as promised.

Circuit-Noise Parameter: Resistance-Limited and Amplifier-Limited Optical Receivers

It is convenient to lump the various sources of circuit noise (thermal noise in resistors as well as noise in transistors and other circuit devices) into a single random current source i_r at the receiver input that produces the same total noise at the receiver output (Fig. 17.5-6). The mean value of i_r is zero and the variance σ_r^2 depends on temperature, receiver bandwidth, circuit parameters, and type of devices.

Furthermore, it is convenient to define a dimensionless circuit-noise parameter

$$\sigma_q = \frac{\sigma_r T}{e} = \frac{\sigma_r}{2Be}, \quad (17.5-27)$$

where B is the receiver bandwidth and $T = 1/2B$ is the receiver resolution time. Since

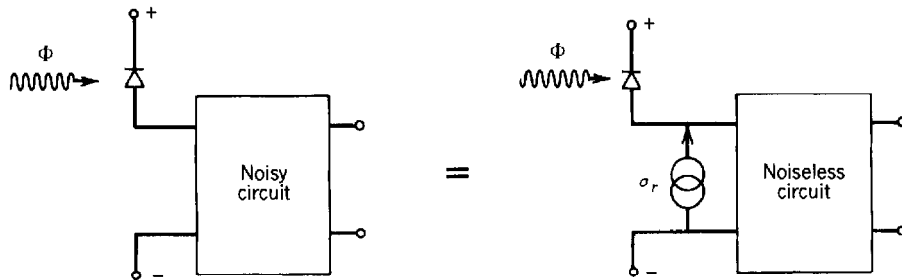


Figure 17.5-6 Noise in the receiver circuit can be replaced with a single random current source with rms value σ_r .

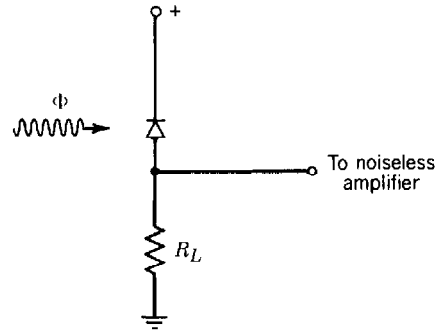


Figure 17.5-7 The resistance-limited optical receiver.

σ_r is the rms value of the electric noise current, σ_r/e is the rms electron flux (electrons/s) arising from circuit noise, and $\sigma_q = (\sigma_r/e)T$ therefore represents the rms number of circuit-noise electrons collected in the time T . It will become apparent in Sec. 17.5D that the circuit-noise parameter σ_q is a figure of merit that characterizes the quality of the optical receiver circuit.

An optical receiver comprising a photodiode, in series with a load resistor R_L followed by an amplifier, is illustrated in Fig. 17.5-7. This simple receiver is said to be **resistance limited** if the circuit-noise current arising from thermal noise in the load resistor substantially exceeds contributions from other sources. The amplifier may then be regarded as noiseless and the circuit-noise mean-square current is simply $\sigma_r^2 = 4k_B T B / R_L$. The circuit-noise parameter defined by (17.5-27) is therefore

$$\sigma_q = \left(\frac{k_B T}{e^2 R_L B} \right)^{1/2}, \quad (17.5-28)$$

which is inversely proportional to the square-root of the bandwidth B .

EXAMPLE 17.5-4. Circuit-Noise Parameter. At room temperature, a resistance $R_L = 50 \, \Omega$ in a circuit of bandwidth $B = 100 \, \text{MHz}$ generates a random current of rms value $\sigma_r = 0.18 \, \mu\text{A}$. This corresponds to a circuit-noise parameter $\sigma_q \approx 5700$.

Receivers using well-designed low-noise amplifiers can yield lower circuit-noise parameters than does the resistance-limited receiver. Consider a receiver using an *FET amplifier*. If the noise arising from the high input resistance of the amplifier can be neglected, the receiver is limited by thermal noise in the channel between the FET source and drain. With the use of an equalizer to boost the high frequencies attenuated by the capacitive input impedance of the circuit, the circuit-noise parameter at room temperature for typical circuit component values turns out to be

$$\sigma_q \approx \frac{B^{1/2}}{100} \quad (B \text{ in Hz}). \quad (17.5-29)$$

Circuit-Noise Parameter
(FET Amplifier Receiver)

For example, if $B = 100 \, \text{MHz}$, then $\sigma_q = 100$. This is significantly smaller than the circuit-noise parameter associated with a $50\text{-}\Omega$ resistance-limited amplifier of the same

bandwidth. The circuit-noise parameter σ_q increases with B because of the effect of the equalizer.[†]

Receivers using *bipolar transistor amplifiers*, in contrast, have a circuit-noise parameter σ_q that is independent of the bandwidth B over a wide range of frequencies.[†] For bandwidths between 100 MHz and 2 GHz, σ_q is typically ≈ 500 , provided that appropriate transistors are used and that they are optimally biased.

D. Signal-to-Noise Ratio and Receiver Sensitivity

The simplest measure of the quality of reception is the signal-to-noise ratio (SNR). The SNR of the current at the input to the noiseless circuit represented in Fig. 17.5-6 is the ratio of the square of the mean current to the sum of the variances of the constituent sources of noise, i.e.,

$$\text{SNR} = \frac{\bar{i}^2}{2e\bar{G}iBF + \sigma_r^2} = \frac{(e\bar{G}\eta\Phi)^2}{2e^2\bar{G}^2\eta B\Phi F + \sigma_r^2}. \quad (17.5-30)$$

Signal-to-Noise
Ratio of an
Optical Receiver

The first term in each of the denominators represents photoelectron and gain noise [see (17.5-20)], whereas the second term represents circuit noise. For a detector without gain, $\bar{G} = 1$ and $F = 1$. Even if it provides amplification, the noiseless circuit does not alter the signal-to-noise ratio.

EXERCISE 17.5-1

Signal-to-Noise Ratio of the Resistance-Limited Optical Receiver. Assume that the optical receiver shown in Fig. 17.5-7 uses an ideal *p-i-n* photodiode ($\eta = 1$) and the resistance R_L is 50 Ω at room temperature ($T = 300$ K). The bandwidth is $B = 100$ MHz. At what value of photon flux Φ is the photoelectron-noise current variance equal to the resistor thermal-noise current variance? What is the corresponding optical power at $\lambda_o = 1.55$ μm ?

It is useful to write the SNR in (17.5-30) in terms of the mean number of detected photons \bar{m} in the resolution time of the receiver $T = 1/2B$,

$$\bar{m} = \eta\Phi T = \frac{\eta\Phi}{2B}, \quad (17.5-31)$$

and the circuit noise-parameter $\sigma_q = \sigma_r/2Be$. The resulting expression is

$$\text{SNR} = \frac{\bar{G}^2\bar{m}^2}{\bar{G}^2F\bar{m} + \sigma_q^2}. \quad (17.5-32)$$

Signal-to-Noise Ratio
of an Optical Receiver

[†]For further details, see S. D. Personick, *Optical Fiber Transmission Systems*, Plenum Press, New York, 1981, Sec. 3.4; note that the parameter σ_q is equivalent to Z/λ in this reference.

Equation (17.5-32) has a simple interpretation. The numerator is the square of the mean number of multiplied photoelectrons detected in the receiver resolution time $T = 1/2B$. The denominator is the sum of the variances of the number of photoelectrons and the number of circuit-noise electrons collected in T . For a photodiode without gain $\bar{G} = F = 1$, so that (17.5-32) reduces to

$$\text{SNR} = \frac{\bar{m}^2}{\bar{m} + \sigma_q^2}.$$

(17.5-33)

Signal-to-Noise Ratio
of an Optical Receiver
in the Absence of Gain

The relative magnitudes of \bar{m} and σ_q^2 determine the relative importance of photoelectron noise and circuit noise. The manner in which the parameter σ_q characterizes the circuit's performance as an optical receiver is now apparent. For example, if $\sigma_q = 100$, then circuit noise dominates photoelectron noise as long as the mean number of photoelectrons recorded per resolution time lies below 10,000.

We proceed now to examine the dependence of the SNR on photon flux Φ , circuit bandwidth B , receiver circuit-noise parameter σ_q , and gain \bar{G} . This will allow us to determine when the use of an APD is beneficial and will permit us to select an appropriate preamplifier for a given photon flux. In undertaking this parametric study, we rely on the expressions for the SNR provided in (17.5-30), (17.5-32), and (17.5-33).

Dependence of the SNR on Photon Flux

The dependence of the SNR on $\bar{m} = \eta\Phi/2B$ provides an indication of how the SNR varies with the photon flux Φ . Consider first a photodiode without gain, in which case (17.5-33) applies. Two limiting cases are of interest:

- *Circuit-Noise Limit:* If Φ is sufficiently small, such that $\bar{m} \ll \sigma_q^2$ ($\Phi \ll 2B\sigma_q^2/\eta$), the photon noise is negligible and circuit noise dominates, yielding

$$\text{SNR} \approx \frac{\bar{m}^2}{\sigma_q^2}. \quad (17.5-34)$$

- *Photon-Noise Limit:* If the photon flux Φ is sufficiently large, such that $\bar{m} \gg \sigma_q^2$ ($\Phi \gg 2B\sigma_q^2/\eta$), the circuit-noise term can be neglected, whereupon

$$\text{SNR} \approx \bar{m}. \quad (17.5-35)$$

For small \bar{m} , therefore, the SNR is proportional to \bar{m}^2 and thereby to Φ^2 , whereas for large \bar{m} , it is proportional to \bar{m} and thereby to Φ , as illustrated in Fig. 17.5-8. For all levels of light the SNR increases with increasing incident photon flux Φ ; the presence of more light improves receiver performance.

When the Use of an APD Provides an Advantage

We now compare two receivers that are identical in all respects except that one exhibits no gain, whereas the other exhibits gain \bar{G} and excess noise factor F (e.g., an APD). For sufficiently small \bar{m} (or photon flux Φ), circuit noise dominates. Amplifying the photocurrent above the level of the circuit noise should then improve the SNR. The APD receiver would then be superior. For sufficiently large \bar{m} (or photon flux),

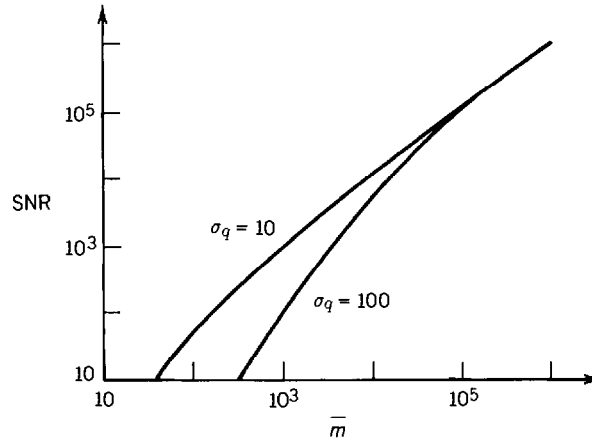


Figure 17.5-8 Signal-to-noise ratio (SNR) as a function of the mean number of photoelectrons per receiver resolution time, $\bar{m} = \eta\Phi/2B$, for a photodiode at two values of the circuit-noise parameter σ_q .

circuit noise is negligible. Amplifying the photocurrent then introduces gain noise, thereby reducing the SNR. The photodiode receiver would then be superior. Comparing (17.5-32) and (17.5-33) shows that the SNR of the APD receiver is greater than that of the photodiode receiver when $\bar{m} < \sigma_q^2(1 - 1/\bar{G}^2)/(F - 1)$. For $\bar{G} \gg 1$, the APD provides an advantage when $\bar{m} < \sigma_q^2/(F - 1)$. If this condition is not satisfied, the use of an APD compromises rather than enhances receiver performance. When σ_q is very small, for example, it is evident from (17.5-32) that the APD SNR = \bar{m}/F is inferior to the photodiode SNR = \bar{m} . The SNR is plotted as a function of \bar{m} for the two receivers in Fig. 17.5-9.

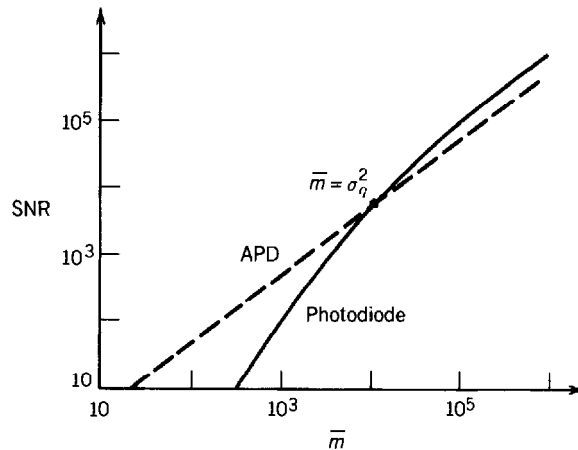


Figure 17.5-9 SNR versus $\bar{m} = \eta\Phi/2B$ for a photodiode receiver (solid curve) and for an APD receiver with mean gain $\bar{G} = 100$ and excess noise factor $F = 2$ (dashed curve) obtained from (17.5-32). The circuit noise parameter $\sigma_q = 100$ in both cases. For small photon flux (circuit-noise-limited case), the APD yields a higher SNR than the photodiode. For large photon flux (photon-noise limit), the photodiode receiver is superior to the APD receiver. The transition between the two regions occurs at $\bar{m} \approx \sigma_q^2/(F - 1) = 10^4$.

Dependence of the SNR on APD Gain

The use of an APD is beneficial for a sufficiently small photon flux such that $\bar{m} < \sigma_q^2/(F - 1)$. The optimal gain of an APD is now determined by making use of (17.5-32),

$$\text{SNR} = \frac{\bar{G}^2 \bar{m}}{\bar{G}^2 F + \sigma_q^2 / \bar{m}}. \quad (17.5-36)$$

For an APD, the excess noise factor F is itself a function of \bar{G} , in accordance with (17.5-23). Substitution yields

$$\text{SNR} = \frac{\bar{G}^2 \bar{m}}{\ell \bar{G}^3 + (1 - \ell)(2\bar{G}^2 - \bar{G}) + \sigma_q^2 / \bar{m}}, \quad (17.5-37)$$

where ℓ is the APD carrier ionization ratio. This expression is plotted in Fig. 17.5-10 for $\bar{m} = 1000$ and $\sigma_q = 500$. For the single-carrier multiplication ($\ell = 0$) APD, the SNR increases with gain and eventually saturates. For the double-carrier multiplication ($\ell > 0$) APD, the SNR also increases with increasing gain, but it reaches a maximum at an optimal value of the gain, beyond which it decreases as a result of the sharp increase in gain noise. In general, there is therefore an optimal choice of APD gain.

Dependence of the SNR on Receiver Bandwidth

The relation between the SNR and the bandwidth B is implicit in (17.5-30). It is governed by the dependence of the circuit-noise current variance σ_r^2 on B . Consider three receivers:

- The *resistance-limited receiver* exhibits $\sigma_r^2 \propto B$ [see (17.5-26)] so that

$$\text{SNR} \propto B^{-1}. \quad (17.5-38)$$

- The *FET amplifier* receiver obeys $\sigma_q \propto B^{1/2}$ [see (17.5-29)] so that $\sigma_r = 2eB\sigma_q \propto B^{3/2}$. This indicates that the dependence of the SNR on B in (17.5-30) assumes

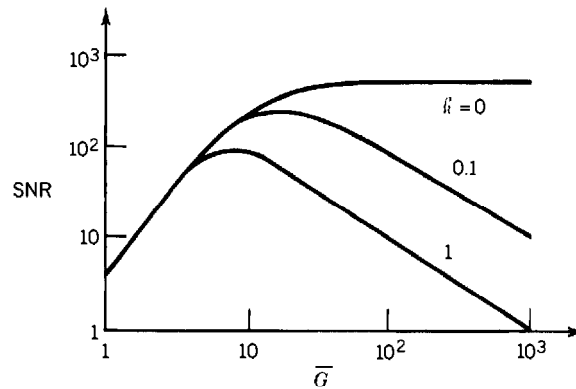


Figure 17.5-10 Dependence of the SNR on the APD mean gain \bar{G} for different ionization ratios ℓ when $\bar{m} = 1000$ and $\sigma_q = 500$.

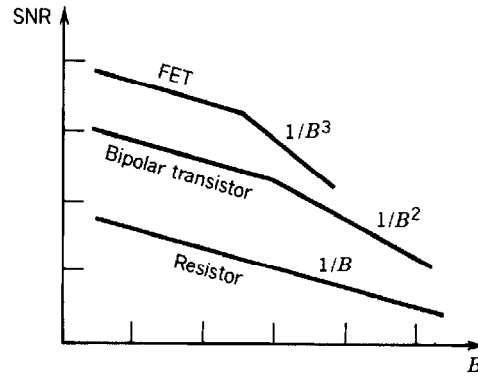


Figure 17.5-11 Double-logarithmic plot of the dependence of the SNR on the bandwidth B for three types of receivers.

the form

$$\text{SNR} \propto (B + sB^3)^{-1}, \quad (17.5-39)$$

where s is a constant.

- The *bipolar-transistor amplifier* has a circuit-noise parameter σ_q that is approximately independent of B . Thus $\sigma_r \propto B$, so that (17.5-38) take the form

$$\text{SNR} \propto (B + s'B^2)^{-1}, \quad (17.5-40)$$

where s' is a constant.

These relations are illustrated schematically in Fig. 17.5-11. The SNR always decreases with increasing B . For sufficiently small bandwidths, all of the receivers exhibit an SNR that varies as B^{-1} . For large bandwidths, the SNR of the FET and bipolar transistor-amplifier receivers declines more sharply with bandwidth.

Receiver Sensitivity

The receiver sensitivity is the minimum photon flux Φ_0 , with its corresponding optical power $P_0 = h\nu\Phi_0$ and corresponding mean number of photoelectrons $\bar{m}_0 = \eta\Phi_0/2B$, required to achieve a prescribed value of signal-to-noise ratio SNR_0 . The quantity \bar{m}_0 can be determined by solving (17.5-32) for $\text{SNR} = \text{SNR}_0$. We shall consider only the case of the unity-gain receiver, leaving the more general solution as an exercise.

Solving the quadratic equation (17.5-33) for \bar{m}_0 , we obtain

$$\bar{m}_0 = \frac{1}{2} \left[\text{SNR}_0 + (\text{SNR}_0^2 + 4\sigma_q^2 \text{SNR}_0)^{1/2} \right]. \quad (17.5-41)$$

Two limiting cases emerge:

$$\text{Photon-noise limit } \left(\sigma_q^2 \ll \frac{\text{SNR}_0}{4} \right): \quad \bar{m}_0 = \text{SNR}_0 \quad (17.5-42)$$

$$\text{Circuit-noise limit } \left(\sigma_q^2 \gg \frac{\text{SNR}_0}{4} \right): \quad \bar{m}_0 = (\text{SNR}_0)^{1/2} \sigma_q. \quad (17.5-43)$$

Receiver
Sensitivity

EXAMPLE 17.5-5. Receiver Sensitivity. We assume that $\text{SNR}_0 = 10^4$, which corresponds to an acceptable signal-to-noise ratio of 40 dB. If the receiver circuit-noise parameter $\sigma_q \ll 50$, the receiver is photon-noise limited and its sensitivity is $\bar{m}_0 = 10,000$ photoelectrons per receiver resolution time. In the more likely situation for which $\sigma_q \gg 50$, the receiver sensitivity $\approx 100\sigma_q$. If $\sigma_q = 500$, for example, the sensitivity is $\bar{m}_0 = 50,000$, which corresponds to $2B\bar{m}_0 = 10^5 B$ photoelectrons/s. The optical power sensitivity $P_0 = 2B\bar{m}_0 h\nu/\eta = 10^5 B h\nu/\eta$ is directly proportional to the bandwidth. If $B = 100$ MHz and $\eta = 0.8$, then at $\lambda_o = 1.55$ μm the receiver sensitivity is $P_0 \approx 1.6$ μW .

When using (17.5-41) to determine the receiver sensitivity, it should be kept in mind that the circuit-noise parameter σ_q is, in general, a function of the bandwidth B , in accordance with:

$$\text{Resistance-limited receiver:} \quad \sigma_q \propto B^{-1/2}$$

$$\text{FET amplifier:} \quad \sigma_q \propto B^{1/2}$$

$$\text{Bipolar-transistor amplifier:} \quad \sigma_q \text{ independent of } B$$

For these receivers, the sensitivity \bar{m}_0 depends on bandwidth B as illustrated in Fig. 17.5-12. The optimal choice of receiver therefore depends in part on the bandwidth B .

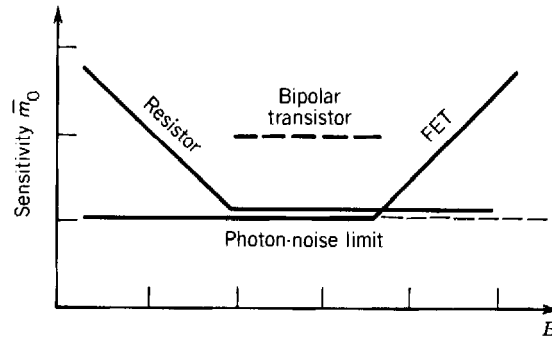


Figure 17.5-12 Double-logarithmic plot of receiver sensitivity \bar{m}_0 (the minimum mean number of photoelectrons per resolution time $T = 1/2B$ guaranteeing a minimum signal-to-noise ratio SNR_0) as a function of bandwidth B for three types of receivers. The curves approach the photon-noise limit at values of B for which $\sigma_q^2 \ll \text{SNR}_0/4$. In the photon-noise limit (i.e., when circuit noise is negligible), $\bar{m}_0 = \text{SNR}_0$ in all cases.

EXERCISE 17.5-2

Sensitivity of the APD Receiver. Derive an expression analogous to (17.5-41) for the sensitivity of a receiver incorporating an APD of gain \bar{G} and excess noise factor F . Show that in the limit of negligible circuit noise, the receiver sensitivity reduces to

$$\bar{m}_0 = F \cdot \text{SNR}_0.$$

READING LIST

Books

See also the reading list in Chapter 15.

- J. D. Vincent, *Fundamentals of Infrared Detector Operation and Testing*, Wiley, New York, 1990.
- N. V. Joshi, *Photoconductivity*, Marcel Dekker, New York, 1990.
- P. N. J. Dennis, *Photodetectors*, Plenum Press, New York, 1986.
- A. van der Ziel, *Noise in Solid State Devices and Circuits*, Wiley-Interscience, New York, 1986.
- R. K. Willardson and A. C. Beer, eds., *Semiconductors and Semimetals*, vol. 22, *Lightwave Communications Technology*, W. T. Tsang, ed., part D, *Photodetectors*, Academic Press, New York, 1985.
- E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, Wiley, New York, 1984.
- R. W. Boyd, *Radiometry and the Detection of Optical Radiation*, Wiley, New York, 1983.
- W. Budde, ed., *Physical Detectors of Optical Radiation*, Academic Press, New York, 1983.
- M. J. Buckingham, *Noise in Electron Devices and Systems*, Wiley, New York, 1983.
- R. J. Keyes, ed., *Optical and Infrared Detectors*, vol. 19, *Topics in Applied Physics*, Springer-Verlag, Berlin, 2nd ed. 1980.
- D. F. Barbe, ed., *Charge Coupled Devices*, vol. 39, *Topics in Applied Physics*, Springer-Verlag, Berlin, 1980.
- R. W. Engstrom, *RCA Photomultiplier Handbook (PMT-62)*, RCA Electro Optics and Devices, Lancaster, PA, 1980.
- B. O. Seraphin, ed., *Solar Energy Conversion: Solid-State Physics Aspects*, vol. 31, *Topics in Applied Physics*, Springer-Verlag, Berlin, 1979.
- R. H. Kingston, *Detection of Optical and Infrared Radiation*, Springer-Verlag, New York, 1978.
- B. Saleh, *Photoelectron Statistics*, Springer-Verlag, New York, 1978.
- A. Rose, *Concepts in Photoconductivity and Allied Problems*, Wiley-Interscience, New York, 1963; R. E. Krieger, Huntington, NY, 2nd ed. 1978.
- M. Cardona and L. Ley, eds., *Photoemission in Solids*, vol. 26, *Topics in Applied Physics*, Springer-Verlag, Berlin, 1978.
- R. K. Willardson and A. C. Beer, eds., *Semiconductors and Semimetals*, vol. 12, *Infrared Detectors II*, Academic Press, New York, 1977.
- J. Mort and D. M. Pai, eds., *Photoconductivity and Related Phenomena*, Elsevier, New York, 1976.
- R. K. Willardson and A. C. Beer, eds., *Semiconductors and Semimetals*, vol. 5, *Infrared Detectors*, R. J. Keyes, ed., Academic Press, New York, 1970.
- A. H. Sommer, *Photoemissive Materials*, Wiley, New York, 1968.

Special Journal Issues

- Special issue on quantum well heterostructures and superlattices, *IEEE Journal of Quantum Electronics*, vol. QE-24, no. 8, 1988.
- Special issue on semiconductor quantum wells and superlattices: physics and applications, *IEEE Journal of Quantum Electronics*, vol. QE-22, no. 9, 1986.
- Special issue on light emitting diodes and long-wavelength photodetectors, *IEEE Transactions on Electron Devices*, vol. ED-30, no. 4, 1983.
- Special issue on optoelectronic devices, *IEEE Transactions on Electron Devices*, vol. ED-29, no. 9, 1982.
- Special issue on light sources and detectors, *IEEE Transactions on Electron Devices*, vol. ED-28, no. 4, 1981.
- Special issue on quaternary compound semiconductor materials and devices—sources and detectors, *IEEE Journal of Quantum Electronics*, vol. QE-17, no. 2, 1981.
- Special joint issue on optoelectronic devices and circuits, *IEEE Transactions on Electron Devices*, vol. ED-25, no. 2, 1978.
- Special joint issue on optical electronics, *Proceedings of the IEEE*, vol. 54, no. 10, 1966.

Articles

- D. Parker, Optical Detectors: Research to Reality, *Physics World*, vol. 3, no. 3, pp. 52–54, 1990.
- S. R. Forrest, Optical Detectors for Lightwave Communication, in *Optical Fiber Telecommunications II*, S. E. Miller and I. P. Kaminow, eds., Academic Press, New York, 1988.
- G. Margaritondo, 100 Years of Photoemission, *Physics Today*, vol. 44, no. 4, pp. 66–72, 1988.
- F. Capasso, Band-Gap Engineering: From Physics and Materials to New Semiconductor Devices, *Science*, vol. 235, pp. 172–176, 1987.
- S. R. Forrest, Optical Detectors: Three Contenders, *IEEE Spectrum*, vol. 23, no. 5, pp. 76–84, 1986.
- M. C. Teich, K. Matsuo, and B. E. A. Saleh, Excess Noise Factors for Conventional and Superlattice Avalanche Photodiodes and Photomultiplier Tubes, *IEEE Journal of Quantum Electronics*, vol. QE-22, pp. 1184–1193, 1986.
- D. S. Chemla, Quantum Wells for Photonics, *Physics Today*, vol. 38, no. 5, pp. 56–64, 1985.
- F. Capasso, Multilayer Avalanche Photodiodes and Solid-State Photomultipliers, *Laser Focus/Electro-Optics*, vol. 20, no. 7, pp. 84–101, 1984.
- P. P. Webb and R. J. McIntyre, Recent Developments in Silicon Avalanche Photodiodes, *RCA Engineer*, vol. 27, pp. 96–102, 1982.
- H. Melchior, Detectors for Lightwave Communication, *Physics Today*, vol. 30, no. 11, pp. 32–39, 1977.
- P. P. Webb, R. J. McIntyre, and J. Conradi, Properties of Avalanche Photodiodes, *RCA Review*, vol. 35, pp. 234–278, 1974.
- R. J. Keyes and R. H. Kingston, A Look at Photon Detectors, *Physics Today*, vol. 25, no. 3, pp. 48–54, 1972.
- H. Melchior, Demodulation and Photodetection Techniques, in F. T. Arecchi and E. O. Schulz-Dubois, eds., *Laser Handbook*, vol. 1, North-Holland, Amsterdam, 1972, pp. 725–835.
- W. E. Spicer and F. Wooten, Photoemission and Photomultipliers, *Proceedings of the IEEE*, vol. 51, pp. 1119–1126, 1963.

PROBLEMS

- 17.1-1 **Effect of Reflectance on Quantum Efficiency.** Determine the factor $1 - \mathcal{R}$ in the expression for the quantum efficiency, under normal and 45° incidence, for an unpolarized light beam incident from air onto Si, GaAs, and InSb (see Sec. 6.2 and Table 15.2-1 on page 588).
- 17.1-2 **Responsivity.** Find the maximum responsivity of an ideal (unity quantum efficiency and unity gain) semiconductor photodetector made of (a) Si; (b) GaAs; (c) InSb.
- 17.1-3 **Transit Time.** Referring to Fig. 17.1-3, assume that a photon generates an electron–hole pair at the position $x = w/3$, that $v_e = 3v_h$ (in semiconductors v_e is generally larger than v_h), and that the carriers recombine at the contacts. For each carrier, find the magnitudes of the currents, i_h and i_e , and the durations of the currents, τ_h and τ_e . Express your results in terms of e , w , and v_e . Verify that the total charge induced in the circuit is e . For $v_e = 6 \times 10^7$ cm/s and $w = 10$ μ m, sketch the time course of the currents.
- 17.1-4 **Current Response with Uniform Illumination.** Consider a semiconductor material (as in Fig. 17.1-3) exposed to an impulse of light at $t = 0$ that generates N electron–hole pairs uniformly distributed between 0 and w . Let the electron and hole velocities in the material be v_e and v_h , respectively. Show that the hole

current can be written as

$$i_h(t) = \begin{cases} -\frac{Ne v_h^2}{w^2} t + \frac{Ne v_h}{w} & 0 \leq t \leq \frac{w}{v_h} \\ 0, & \text{elsewhere,} \end{cases}$$

while the electron current is

$$i_e(t) = \begin{cases} -\frac{Ne v_e^2}{w^2} t + \frac{Ne v_e}{w}, & 0 \leq t \leq \frac{w}{v_e} \\ 0, & \text{elsewhere,} \end{cases}$$

and that the total current is therefore

$$i(t) = \begin{cases} \frac{Ne}{w} \left[(v_h + v_e) - \frac{1}{w} (v_h^2 + v_e^2) t \right], & 0 \leq t \leq \frac{w}{v_e} \\ \frac{Ne v_h}{w} \left[1 - \frac{v_h}{w} t \right], & \frac{w}{v_e} \leq t \leq \frac{w}{v_h} \end{cases}$$

The various currents are illustrated in Fig. 17.1-4. Verify that the electrons and holes each contribute a charge $Ne/2$ to the external circuit so that the total charge generated is Ne .

- *17.1-5 **Two-Photon Detectors.** Consider a beam of photons of energy $h\nu$ and photon flux density ϕ (photons/cm²-s) incident on a semiconductor detector with bandgap $h\nu < E_g < 2h\nu$, such that one photon cannot provide sufficient energy to raise an electron from the valence band to the conduction band. Nevertheless, two photons can occasionally conspire to jointly give up their energy to the electron. Assume that the current *density* induced in such a detector is given by $J_p = \xi \phi^2$, where ξ is a constant. Show that the responsivity (A/W) is given by $\mathfrak{R} = [\xi/(hc_o)^2] \lambda_o^2 P/A$ for the two-photon detector, where P is the optical power and A is the detector area illuminated. Explain physically the proportionality to λ_o^2 and P/A .
- 17.2-1 **Photoconductor Circuit.** A photoconductor detector is often connected in series with a load resistor R and a dc voltage source V , and the voltage V_p across the load resistor is measured. If the conductance of the detector is proportional to the optical power P , sketch the dependence of V_p on P . Under what conditions is this dependence linear?
- 17.2-2 **Photoconductivity.** The concentration of charge carriers in a sample of intrinsic Si is $n_i = 1.5 \times 10^{10}$ cm⁻³ and the recombination lifetime $\tau = 10$ μ s. If the material is illuminated with light, and an optical power density of 1 mW/cm² at $\lambda_o = 1$ μ m is absorbed by the material, determine the percentage increase in its conductivity. The quantum efficiency $\eta = \frac{1}{2}$.
- 17.3-1 **Quantum Efficiency of a Photodiode Detector.** For a particular $p-i-n$ photodiode, a pulse of light containing 6×10^{12} incident photons at wavelength $\lambda_o = 1.55$ μ m gives rise to, on average, 2×10^{12} electrons collected at the terminals of the device. Determine the quantum efficiency η and the responsivity \mathfrak{R} of the photodiode at this wavelength.

- 17.4-1 **Quantum Efficiency of an APD.** A conventional APD with gain $\bar{G} = 20$ operates at a wavelength $\lambda_o = 1.55 \mu\text{m}$. If its responsivity at this wavelength is $\mathfrak{R} = 12 \text{ A/W}$, calculate its quantum efficiency η . What is the photocurrent at the output of the device if a photon flux $\Phi = 10^{10}$ photons/s, at this same wavelength, is incident on it?
- 17.4-2 **Gain of an APD.** Show that an APD with ionization ratio $k = 1$, such as germanium, has a gain given by $\bar{G} = 1/(1 - \alpha_e w)$, where α_e is the electron ionization coefficient and w is the width of the multiplication layer. [Note: Equation (17.4-5) does not give a proper answer for the gain when $k = 1$.]
- 17.5-1 **Excess Noise Factor for a Single-Carrier APD.** Show that an APD with pure electron injection and no hole multiplication ($k = 0$) has an excess noise factor $F \approx 2$ for all appreciable values of the gain. Use (17.4-5) to show that the mean gain is then $G = \exp(\alpha_e w)$. Calculate the responsivity of a Si APD for photons with energy equal to the bandgap energy E_g , assuming that the quantum efficiency $\eta = 0.8$ and the gain $\bar{G} = 70$. Find the excess noise factor for a double-carrier-multiplication Si APD when $k = 0.01$. Compare it with the value $F \approx 2$ obtained in the single-carrier-multiplication limit.
- *17.5-2 **Gain of a Multilayer APD.** Use the Bernoulli probability law to show that the mean gain of a single-carrier-multiplication multilayer APD is $\bar{G} = (1 + P)^l$, where P is the probability of impact ionization at each stage and l is the number of stages. Show that the result reduces to that of the conventional APD when $P \rightarrow 0$ and $l \rightarrow \infty$.
- *17.5-3 **Excess Noise Factor for a One-Stage Photomultiplier Tube.** Derive an expression for the excess noise factor F of a one-stage photomultiplier tube assuming that the number of secondary emission electrons per incident primary electron is Poisson distributed with mean δ .
- *17.5-4 **Excess Noise Factor for a Photoconductor Detector.** The gain of a photoconductor detector was shown in Sec. 17.2 to be $G = \tau/\tau_e$, where τ is the electron-hole recombination lifetime and τ_e is the electron transit time across the sample. Actually, G is random because τ can be thought of as random. Show that an exponential probability density function for the random recombination lifetime, $P(\tau) = (1/\bar{\tau}) \exp(-\tau/\bar{\tau})$, results in an excess noise factor $F = 2$, confirming that photoconductor **generation-recombination (GR) noise** degrades the SNR by a factor of 2.
- 17.5-5 **Bandwidth of an RC Circuit.** Using the definition of bandwidth provided in (17.5-14), show that a circuit of impulse response function $h(t) = (e/\tau) \exp(-t/\tau)$ has a bandwidth $B = 1/4\tau$. What is the bandwidth of an RC circuit? Determine the thermal noise current for a resistance $R = 1 \text{ k}\Omega$ at $T = 300 \text{ K}$ connected to a capacitance $C = 5 \text{ pF}$.
- 17.5-6 **Signal-to-Noise Ratio of an APD Receiver.** By what factor does the signal-to-noise ratio of a receiver using an APD of mean gain $\bar{G} = 100$ change if the ionization ratio k is increased from $k = 0.1$ to 0.2 . Assume that circuit noise is negligible. Show that if the mean gain $\bar{G} \gg 1$ and $\gg 2(1 - k)/k$, the SNR is approximately inversely proportional to \bar{G} .
- 17.5-7 **Noise in an APD Receiver.** An optical receiver using an APD has the following parameters: quantum efficiency $\eta = 0.8$; mean gain $\bar{G} = 100$; ionization ratio $k = 0.5$; load resistance $R_L = 1 \text{ k}\Omega$; bandwidth $B = 100 \text{ kHz}$; dark and leakage current $= 1 \text{ nA}$. An optical signal of power 10 nW at $\lambda_o = 0.87 \mu\text{m}$ is received. Determine the rms values of the different noise currents, and the SNR. Assume

that the dark and leakage current has a noise variance that obeys the same law as photocurrent noise and that the receiver is resistance limited.

- 17.5-8 **Optimal Gain in an APD.** A receiver using a $p-i-n$ photodiode has a ratio of circuit noise variance to photoelectron noise variance of 1000. If an APD with ionization ratio $k = 0.2$ is used instead, determine the optimal mean gain for maximizing the signal-to-noise ratio and the corresponding improvement in signal-to-noise ratio.
- 17.5-9 **Receiver Sensitivity.** Determine the receiver sensitivity (i.e., optical power required to achieve a $\text{SNR} = 10^3$) for a photodetector of quantum efficiency $\eta = 0.8$ at $\lambda_o = 1.3 \mu\text{m}$ in a circuit of bandwidth $B = 100 \text{ MHz}$ when there is no circuit noise. The receiver measures the electric current i .
- 17.5-10 **Noise Comparison of Three Photodetectors.** Consider three photodetectors in series with a $50\text{-}\Omega$ load resistor at 77 K (liquid nitrogen temperature) that are to be used with a $1\text{-}\mu\text{m}$ wavelength optical system with a bandwidth of 1 GHz : (1) a $p-i-n$ photodiode with quantum efficiency $\eta = 0.9$; (2) an APD with quantum efficiency $\eta = 0.6$, gain $\bar{G} = 100$, and ionization ratio $k = 0$; (3) a 10-stage photomultiplier tube (PMT) with quantum efficiency $\eta = 0.3$, overall mean gain $\bar{G} = 4^{10}$, and overall gain variance $\sigma_G^2 = \bar{G}^2/4$.
 (a) For each detector, find the photocurrent SNR when the detector is illuminated by a photon flux of 10^{10} s^{-1} .
 (b) By which devices is the signal detectable?
- *17.5-11 **A Single-Dynode Photomultiplier Tube.** Consider a photomultiplier tube with quantum efficiency $\eta = 1$ and only one dynode. Incident on the cathode is light from a hypothetical photon source that gives rise to a probability of observing n photons in the counting time $T = 1.3 \text{ ns}$, which is given by

$$p(n) = \begin{cases} \frac{1}{2}, & n = 0, 1 \\ 0, & \text{otherwise.} \end{cases}$$

When one electron strikes the dynode, either two or three secondary electrons are emitted and these proceed to the anode. The gain distribution $P(G)$ is given by

$$P(G) = \begin{cases} \frac{1}{3}, & G = 2 \\ \frac{2}{3}, & G = 3 \\ 0, & \text{otherwise.} \end{cases}$$

Thus it is twice as likely that three electrons are emitted as two.

- (a) Calculate the SNR of the input photon number and compare the result with that of a Poisson photon number of the same mean.
 (b) Find the probability distribution for the photoelectron number $p(m)$ and the SNR of the photoelectron number.
 (c) Find the mean gain $\langle G \rangle$ and the mean-square gain $\langle G^2 \rangle$.
 (d) Find the excess noise factor F .
 (e) Find the mean anode current \bar{i} in a circuit of bandwidth $B = 1/2T$.
 (f) Find the responsivity of this photomultiplier tube if the wavelength of the light is $\lambda_o = 1.5 \mu\text{m}$.
 (g) Explain why (17.5-20) for σ_i^2 is not applicable.