A look at photon detectors

Whether photomultipliers, photodiodes or photoconductors are "best" in a given case depends on signal frequency, on the kinds of noise present and on the coherence of the radiation.

Robert J. Keyes and Robert H. Kingston

In the best of all possible worlds, we would have the ideal photon detector, a device that caught a photon, gave an unambiguous meter reading and kept count of the number of events. In the real world, these ideal devices do not exist; competing events both outside and inside the detector confuse the true measure of the photons that we are trying to monitor. Phenomena such as quantum noise, "dark" current and background radiation interfere to a degree that depends on the intensity of the signal being measured and on the photon frequency, to mention only a few of the experimental parameters.

Here we shall describe the processes that are important when keeping track of low-intensity photon "beams," at frequencies ranging from the ultraviolet through the infrared, because this is the situation for which most of us need a photon detector. We shall look at what "detecting a photon" really means in operational terms, and see what the physical bases of the competing noise processes are. We shall also look at the special problem of coherent detection, important to those of us who want to analyze laser signals. Finally, we shall briefly review what all this theory means in terms of the sensitivity of real detectors throughout the spectrum.

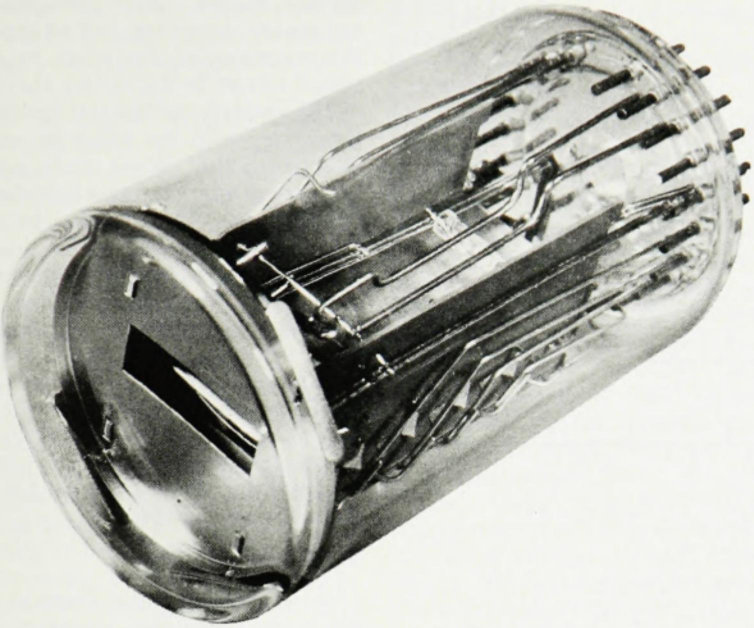
Robert Keyes, formerly head of the Optics and Infrared Group at the Massachusetts Institute of Technology Lincoln Laboratory, is a private consultant; Robert Kingston heads the Laboratory's Optics Division.

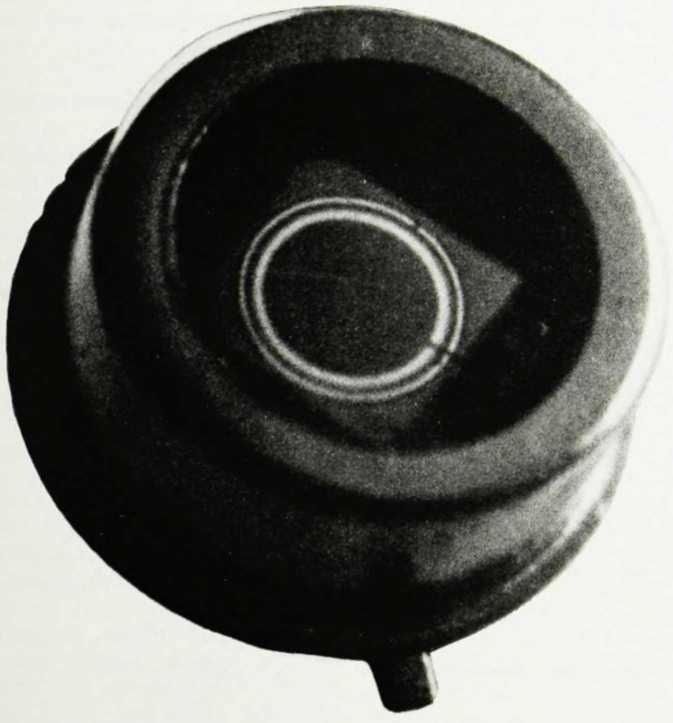
Photoelectric processes

The oldest forms of photon detector are the eye and photographic film, but these two are not usually suitable for detecting low-intensity signals; we shall limit the discussion here to those devices with a single electrical output that measures the photon stream at some point in space. What we want is clearly a photoelectric event—the production of an excited electronic state by a photon. The best way to proceed depends largely on the photon frequency: In general, vacuum-tube detectors are more responsive to the ultraviolet and visible regions of the spectrum, and semiconductors work better in the infrared (see figure 1). Thermal detectors, which operate as heat engines, are not sufficiently sensitive to be used for low-intensity signals.

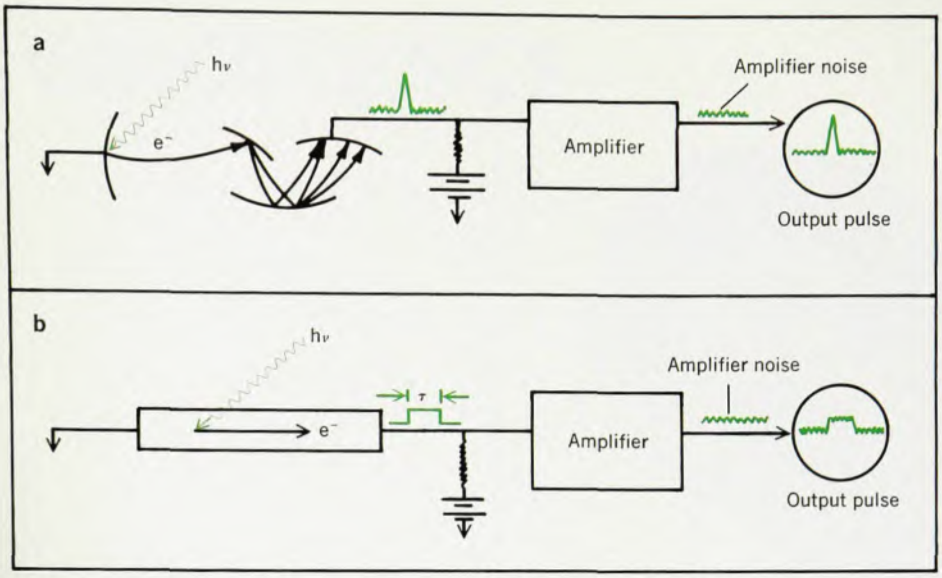
In the visible and ultraviolet region of the spectrum, we are most apt to use the original photoelectric effect, photoemission, in which radiation incident on a metallic surface causes the metal to emit an electron into the surrounding vacuum. Once the electron is emitted, our problem is to detect it. Fortunately, the electron multiplier, which consists of a photoemissive cathode, a series of electrodes at successively higher positive potentials (dynodes) and a collector plate at the highest positive potential, allows us to amplify the initial single-electron pulse, as in figure 2a, so that the measured current pulse is greater than the noise in the measuring amplifier. The dominant source of noise in photomultipliers is the thermal generation of electrons in the photosurface. Because thermal generation is a steep exponential function of temperature, only moderate cooling is needed to eliminate this source of noise.

The quantum efficiency for photoemission can approach unity in the ultraviolet through visible regions of the spectrum, but falls off rapidly to a few percent in the near infrared (one micron). In contrast, the efficiency of electron or hole creation in semiconductors can approach unity throughout the visible and infrared regions, if the semiconductor is sufficiently cooled. The cooling ensures that collision of the photons with already free carriers are negligible compared with the "across the gap" or "bound-free" excitations that we wish to measure; for this as-

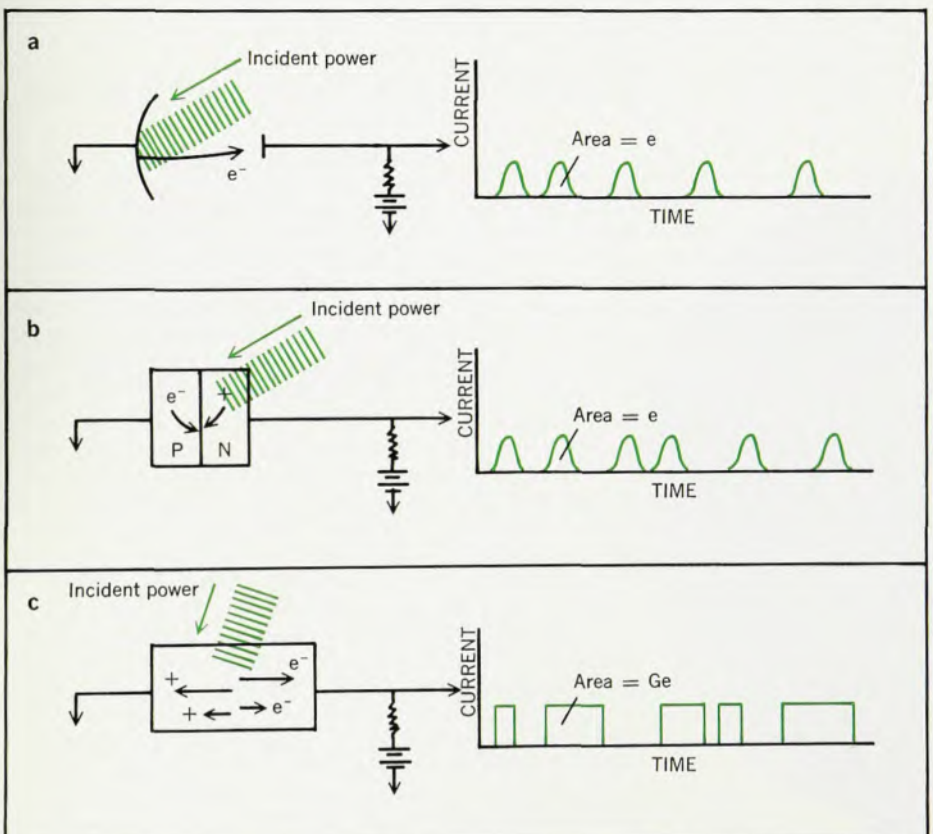




Photomultiplier and photodiode. The eleven-stage "Quantacon" photomultiplier (top) has a gallium-arsenide photocathode and is the type used in the Apollo lunar-ranging experiment. Silicon photodiode (bottom) is sensitive in the infrared region of the spectrum. Avalanche multiplication techniques are making silicon photodiodes competitive with photomultipliers in the one-micron region. (Photos by RCA.)



An electron multiplier amplifies the single-electron pulse resulting from photoemission, so that the measured current pulse is greater than amplifier noise (a). In the infrared region of the spectrum, semiconductors (b) must substitute for photomultipliers. The signal to noise ratio is not as great, but "avalanche" multiplication offers some improvement. Figure 2



Pulses from a photomultiplier (a) and a photodiode (b) follow a Poisson distribution and have the same unvarying charge content. A photoconductor, on the other hand, emits pulses with charge content that varies from pulse to pulse (c) and depends on the lifetime of the excited carriers. Figure 3

surance, the temperature must be much less than $h\nu/k$.

If the semiconductor is thick enough, the photon will eventually produce an excited carrier. But production does not ensure detection, and unfortunately the recording of this event by the external circuit is limited by the lifetime of the excited carrier and by the competing noise in the following amplifier (see figure 2b). Some of the noise can be overcome by biasing a semiconductor photodiode so that "avalanche" multiplication of excited carriers occurs. In any case, we see that detection of *individual photons,* as one might picture the constituents of a very weak signal (for example, from the laser-ranging reflector left on the moon by the Apollo astronauts) is very difficult except at short visible wavelengths. But, as we shall see, more serious problems usually limit the use of a photon detector in any practical system.

Signal versus noise

We have posed the problem of the minimum detectable power or energy seen by a photon detector, and we should realize that our definition of a photon is a practical rather than a theoretical one. To measure its existence, we simply state the measurement rule: The average rate of production of photoelectrons (or holes in semiconductors, for that matter) is equal to the product of the quantum efficiency η and the ratio of the incident power P to the quantum energy $h\nu$ corresponding to the photon frequency, so that

$$i_{av} = e\eta P/h\nu$$

where i_{av} is the photocurrent and e is the electron charge.

An essential feature of this law is that the time behavior of the photoelectric events obeys Poisson statistics (random occurrence, see figure 3a), provided that the incident power remains constant. The random time distribution of electron (excited-carrier) production leads to a current fluctuation or noise in the presence of any photocurrent i such that

$$\delta i = e\delta n/T = eN^{1/2}/T = e(iT/e)^{1/2}/T = (ei/T)^{1/2}$$

because, for a given measurement time T, the fluctuation δn in the number of events is proportional to the square root of the total number N during the time T.

In a photoemitter, we measure the electron events directly, and the noise current (also known as "shot noise") may be shown to be

$$i_{rms} = (2eiB)^{1/2} \qquad (1)$$

where B is the bandwidth of the detector circuit.

We see then that even a "perfect" photon counter would produce a "noisy" output in the presence of a constant input signal, and, in fact, this quantum noise contribution, along with the thermal noise generated in the detector amplifier, are the two limiting factors in the sensitivity of any photon detector, no matter how high the quantum efficiency.

If only the desired signal were present, we could easily calculate the minimum power detectable by any given photoemissive device, assuming we knew the quantum efficiency and the noise-current contribution of the amplifier. But in many applications the sensitivity is more severely limited by the quantum noise generated by sources of radiation or electron production other than the true "signal." In particular, the "dark" current or thermionic-emission component in a photoemitter adds its share of current fluctuation, and even more important is the unwanted background radiation that emanates from the vicinity of the signal source or from the intervening path or optical components. This background radiation may be perfectly constant in time, yet it produces a quantum or "photon" noise that competes with the signal current. We can make the dark current negligible by cooling the photocathode, but the background-induced noise is strictly controlled by the environment of the signal source and the design of the optics and detector system.

Noise in semiconductors

Shot noise in a photoemitter detector is, as we have seen, a fundamental limitation. What is the equivalent noise mechanism in semiconductor devices? Here we must distinguish between the photodiode and the photoconductor. The photodiode is almost an exact analog of the photoemitter, except that minority carriers (rather than electrons) are produced near a reverse-biased junction and are "collected" after they diffuse to the barrier (see figure 3b), just as in a transistor. As each carrier is collected, it produces an electron-sized pulse at the output, and the noise current is the same as that in a photoemitter.

The photoconductor, in contrast, is a bulk device whose conductance is controlled by the photoinduced generation and eventual recombination of free carriers, either holes or electrons. Photoconductors may be either "intrinsic" or "extrinsic." In an intrinsic photoconductor, incident radiation produces an "across the gap" transition, creating an electron–hole pair by giving a bound electron enough energy to cross to a conduction band. In an extrinsic (impurity-doped) device, a free carrier is excited from a bound state. As we have noted, the device must be cool enough that kT is much lower than the photon energy—otherwise the fluctuations of thermally excited carriers will overwhelm the contribution of the photoinduced events.

When we apply a voltage to a photoconductor, we get a series of current pulses in the presence of radiation, as we did with the other devices. Here, however, the pulses, although Poisson distributed, will each have an effective charge that is neither equal to the charge on an electron nor constant (see figure 3c). Instead, as each free carrier is randomly produced, it will drift in the internal field of the photoconductor, producing a current at the output electrodes for a finite time until it recombines. Recombination is a probabilistic process, so that there will be a distribution of pulsewidths. The distribution will be exponential, with a mean width equal to the average lifetime of the carrier.

The average charge per pulse may be simply related to the ratio of the carrier lifetime to the time needed for the carrier to travel completely across the photoconductor. If, for example, the applied voltage is such that the carrier drifts *completely* across the device between the time it is excited and the time it recombines, the charge transfer would be one electron. This average charge induced per excited carrier is defined as the gain G, and it may be greater or less than unity, depending on the lifetime and drift velocity of the carriers.

A gain greater than unity seems to imply that the excited carrier drifts out at one end of the semiconductor and reappears magically at the other terminal to drift again. In actuality, it does not, but charge neutrality requires that a new carrier appear at the opposite terminal to replace the old one, and the recombination probability remains the same, so that the effective lifetime may be many transit times.

As for the noise, we now have a series of random pulses with random charge content of average value Ge. It turns out that we can use the shot-noise formula of equation 1 by the simple artifice of replacing the electron by a G-sized electron and, to take into account the random fluctuation in size, multiplying by an extra factor of two. The resulting noise current is then

$$i_{\rm rms} = (4GeiB)^{1/2} = [4Ge(e\eta GP/h\nu)B]^{1/2} = Ge(4\eta PB/h\nu)^{1/2} \quad (2)$$

We see that the noise behavior of the entire family of photon detectors is markedly similar, if we are willing to deal with random-size electrons and to account for factors of two.

Do not conclude from equation 2 that a small G in a photoconductor is a fruitful way to decrease the noise; remember that the signal current is directly proportional to G so that trying to reduce G (by varying the semiconductor material, for example) is a losing game, because amplifier noise is the dominant limiting noise in photoconductors.

Limits to photodetection

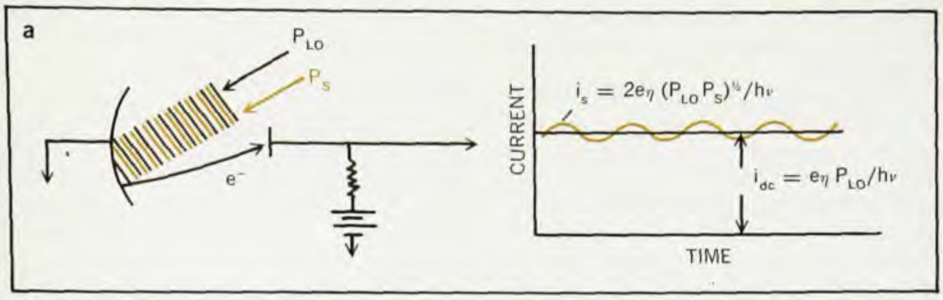
We are now prepared to look at the way the wavelength and the source of the radiation limit photon detection. Three categories of detection exist: photon-noise limited, amplifier-noise limited and background limited.

Photon noise is the limiting factor only in the visible and near-infrared regions of the spectrum. Here appropriate filtering and field-of-view restrictions can reduce background radiation, and electron-multiplier structure can overcome amplifier noise. The limit, then, is the quantum efficiency, which ranges from a high of about 30 percent in the blue region (4000Å) down to a few percent (in experimental tubes) at one micron in the infrared.[1]

A useful quantity, when discussing photon detection, is noise-equivalent power (NEP), which is defined as the signal power that produces a detector output voltage (or current) equal to the rms noise current. As a benchmark, photon-noise limited detection yields a noise-equivalent power equal to $2h\nu B/\eta$ for the photoemitter or semiconductor photodiodes and $4h\nu B/\eta$ for photoconductors. For a 1-Hz bandwidth, the normal reference for evaluating noise-equivalent power, this value corresponds to 8×10^{-19} watts for unity quantum efficiency at 5000Å.

Beyond one micron (10 000 Å) in the infrared, amplifier noise limits detection capabilities in the presence of low background, because vacuum electron multiplication is no longer possible (photoemission does not occur with these low-frequency photons), and we must resort to semiconductor detectors. Silicon photodiodes that use avalanche multiplication are becoming competitive in the one-micron region; their high quantum efficiency partly offsets the noisy internal amplification and compensates for the poor quantum efficiency of photoemitters in this region. For any semiconductor device, the noise-equivalent power increases precipitously, by several orders of magnitude, as the wavelength increases beyond about one micron. Either amplifier noise or photon noise caused by the high background radiation at the longer wavelengths produces an effective noise current proportional to the square root of the bandwidth. The noise-equivalent power then becomes proportional to the square root of the bandwidth, because the signal current needed to compete with the noise is directly proportional to the input signal power.

In many infrared-detection systems, the background radiation sets the limit, so that improved detectors would be of no advantage. Background radiation is particularly important beyond about



Coherent detection. The photocurrent in coherent detection has two components: a constant induced current, proportional to the incident power P_{LO} of the local oscillator, and a cross-product current that varies sinusoidally at the difference frequency between the local-oscillator field and the signal field.
Figure 4

five microns; in this region the emission from the atmosphere is enough to set the noise limit if one is detecting thermal radiation from a distant source. In this case of background-limited detection, the noise-equivalent power becomes equal to $(4P_B h\nu B/\eta)^{1/2}$ for a photoconductor, where P_B is the power of the background radiation striking the detector.

Detecting coherent signals

Our analysis so far bodes poorly for the detection of laser radiation in all but the visible and near-infrared regions of the spectrum: Although the narrow linewidth of laser beams allows optical designs with narrow spectral filters, which could eliminate all background and theoretically allow photon-noise limited detection far into the infrared, this extension is apparently not feasible because of amplifier noise. Fortunately, heterodyne or coherent detection helps solve this problem, with the help of a local laser source. Just as we do in radio or television reception, we shall mix the incoming weak signal wave with a strong oscillator beam at the detector.[2]

Because we want the signal and local oscillator waves to add coherently (that is, the electric fields should add constructively or destructively), we arrange that the detector only "sees" the effect of a signal wave that arrives within an angle less than λ/d of the oscillator beam, where d is the beam diameter. Otherwise, the two wavefronts will not remain in phase across the detector surface. The induced photocurrent is proportional to the incident power P_{LO} of the local oscillator, that is, to the square of the incident electric field, so that we have a constant current $e\eta P_{LO}/h\nu$ due to the local oscillator (see figure 4). In addition, there is a cross product of the local-oscillator field and the signal field that varies sinusoidally at the difference frequency between the two waves. This term is equal to $2e\eta(P_{LO}P_S)^{1/2}/h\nu$, where P_S is the signal power. A much smaller current component, due to the signal wave alone, may be ignored.

The noise current is now set by the "photon" noise of the local oscillator and is equal to $[2e^2(\eta P_{LO}/h\nu)B]^{1/2}$. We find that the noise-equivalent power becomes $h\nu B/\eta$ for the photoemitter or the semiconductor photodiode. For the photoconductor, we have $2h\nu B/\eta$. In either case we must make the local-oscillator power so large that its noise current, rather than amplifier noise, is the dominant contribution.

Particularly significant in this mode of detection is the coherent nature of the process: The detector–local-oscillator system only "sees" radiation coming from the diffraction-limited field of view of the illuminated surface, and the output wave from the detector is a frequency-shifted replica of the input waveform. The current output is proportional to the incident electric field. The detection process here, then, is linear, as contrasted with normal photon detection, in which the output current is proportional to input power. Because of this coherence, the bandwidth B required in the detector circuit is the true spectral bandwidth of the incoming signal. Heterodyne detection, then, is limited to narrow spectral lines, typically laser radiation, because detector bandwidths are seldom larger than 10^9 hertz. (A typical blackbody-radiation source is about 10^{13} hertz.) Under these restrictions of spectral purity and diffraction-limited field of view, we see that the minimum detectable signal is effectively the same as the ideal photon-noise limited case for incoherent detection. Coherent detection is effectively limited to laser systems in a more restrictive sense as well, because local-oscillator power of the required frequency purity and intensity is available only from lasers.

Choosing a detector

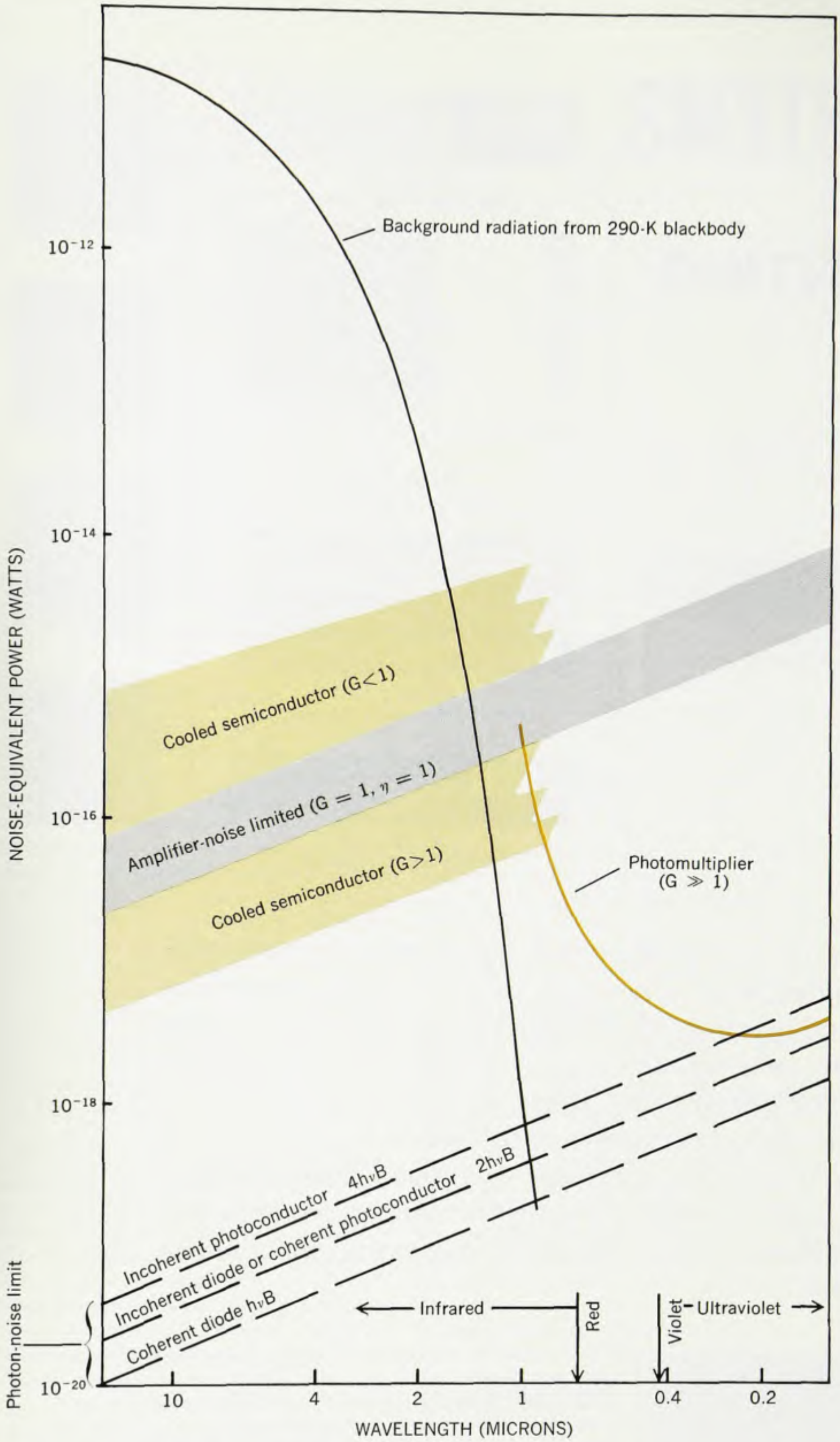
How does one select a detector from among the available types? A major factor in the decision is usually the photon wavelength. In figure 5 we see the reciprocal of the noise-equivalent power plotted against photon wavelength, under standard conditions—detector area equal to one cm^2 and electrical bandwidth of one hertz. Under these conditions, the reciprocal of noise-equivalent power is known as the "detectivity." Indicated on the figure are the photon-noise limited and coherent cases, the 290-K background limitations on a sensor exposed to a full hemisphere, the lower limits of noise-equivalent power for a semiconductor detector with insufficient gain to overcome preamplifier noise and the response curve for a photomultiplier.

Through the visible region and out to one micron, photomultiplier tubes most closely approach the ideal photon counter. The data plotted in figure 5 are for commercial tubes with alkali-metal cathodes. Recently developed GaAs–Cs_2O and InAsP–Cs_2O emitters could increase photomultiplier performance by a factor of 10 to 30 in the one-micron region; the improvement is a direct result of increased quantum efficiencies at these wavelengths.

Our present understanding of photoemitter band structures and vacuum work functions indicates that photoemission out to wavelengths of 2.4 microns is possible. A unique characteristic of photoemitters is their high-frequency capability: In the form of simple diodes, gigahertz responses are possible, and for photomultipliers, 100 MHz is the limit. The bandwidth limitations of photomultipliers results from transit-time spreads in the secondary-emission process.

In the spectral region from 1 to 40 microns, photoconductors have the highest sensitivity (about 10^{-15} watts) among the detectors.[3] Photoconductors characteristically have high quantum efficiencies, often as high as 0.6. When sufficiently cooled, their major limitation is preamplifier noise. To reach the photon-noise-limited $(h\nu B/\eta)$ performance, internal gains of from 10^3 to 10^4 would be needed. This goal is possible, but we must remember that high gain in photoconductors comes only at the cost of frequency response.

Although avalanche photodiodes do not appear in figure 5, they warrant brief discussion. When avalanche multiplication was first observed in silicon p-n junctions, hopes were raised that the solid-state equivalent of photomultipliers was emerging. Unfortunately, this hope has not been realized. The best avalanche silicon diodes are about equivalent to standard photomultipliers at one micron because of the



Noise-equivalent power for photon detectors varies with radiation wavelength. In the region above the colored curve, photomultipliers are more sensitive. Within the shaded area, semiconductor detectors are preferable. Note how background radiation begins to dominate at longer wavelengths. All data shown are for a 1-Hz bandwidth and a 1-cm^2 detector area, except that the photon-noise limit and photomultiplier curves are minimum values, independent of detector area.
Figure 5

extra noise produced in the avalanche process.

Remember that, in figure 5, an area of 1 cm^2 and an electrical bandwidth of 1Hz were assumed. When background or amplifier noise is the limiting factor, the noise-equivalent power is proportional to $(AB)^{1/2}$, where A is the detector area. For the case of background noise, the dependence is obvious, because background power is proportional to area. In the amplifier-limited situation, the dependence may be understood when we note that, for a given bandwidth requirement, the maximum usable load resistance R is inversely proportional to detector capacitance, which is proportional to area. The effective amplifier noise current is $(4kTB/R)^{1/2}$, so that the noise-equivalent power is again proportional to $(AB)^{1/2}$. In the limit, we can reduce the bandwidth and the detector area to such values that we approach the photon-noise limit but only at the expense of an extremely small detector area or an increased response time.

It should be apparent from our discussion that photon detectors in the infrared region are seriously limited in their sensitivity except for measurements of broadband thermal sources, where the background already limits the detection capability. For laser detection, we can resort to the coherent mode, but, just as in a radar receiver, the diffraction-limited behavior of the detection process means that in this situation large collection apertures will lead to extremely narrow fields of view for the system. In addition, the local oscillator must be tuned to within a detector bandwidth of the received signal frequency, so that precise frequency control and fine tuning of the laser are required.

Alternative photon-detection schemes are presently being pursued. One possibility is photon conversion to a shorter wavelength, as in the parametric upconverter, where the output photon may be detected more efficiently in the visible-frequency range. Extension of photoemitter response into the infrared also shows great promise, as does avalanche multiplication in semiconductors, in which new materials and techniques may allow "noiseless" amplification out to the ten-micron region.

* * *

This work was sponsored in part by the US Department of the Air Force.

References

1. R. L. Bell, W. E. Spicer, Proc. IEEE **58**, 1788 (1970).
2. M. C. Teich, Proc. IEEE **56**, 37 (1968).
3. R. J. Keyes, T. M. Tuist in *Semiconductors and Semimetals* vol. 5 (R. K. Willardson, A. C. Beer, eds.), Academic, New York (1970), chapter 8. □

Reprints of this special report on Electro-Optics are available from PHYSICS TODAY. See page 111 for information.