BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES
Vol. 52, No. 1, 2004

Detection of optical radiation

A. ROGALSKI[1]* and Z. BIELECKI[2]

[1]Institute of Applied Physics, [2]Institute of Optoelectronics, Military University of Technology, 2 Kaliskiego St., 00-908 Warsaw, Poland

Abstract. An overview of the important techniques for detection of optical radiation from the ultraviolet, through visible to infrared spectral regions is presented. At the beginning single-point devices are considered. Next, different application circuits used in direct detection systems together with elucidation of the design of front-end circuits and discussion of their performance are presented. Third part of the paper is devoted to advanced techniques including coherent detection. Finally, the updated information devoted to readout of signals from detector arrays and focal plane arrays is included. It is shown that detector focal plane technology has revolutionized many kinds of imaging in the past 25 years.

Keywords: photon and thermal detectors, direct detection systems, coherent and homodyne detection, focal plane arrays, monolithic and hybrid structures, CCD and CMOS readouts, multicolour detectors.

1. Introduction

The birth of photodetectors can be dated back to 1873 when Smith discovered photoconductivity in selenium. Progress was slow until 1905, when Einstein explained the newly observed photoelectric effect in metals, and Planck solved the blackbody emission puzzle by introducing the quanta hypothesis. Applications and new devices soon flourished, pushed by the dawning technology of vacuum tube sensors developed in the 1920s and 1930s culminating in the advent of television. Zworykin and Morton, the celebrated fathers of videonics, on the last page of their legendary book *Television* (1939) concluded that: "*when rockets will fly to the moon and to other celestial bodies, the first images we will see of them will be those taken by camera tubes, which will open to mankind new horizons.*" Their foresight became a reality with the Apollo and Explorer missions. Photolithography enabled the fabrication of silicon monolithic imaging focal planes for the visible spectrum beginning in the early 1960s. Some of these early developments were intended for a picturephone, other efforts were for television cameras, satellite surveillance, and digital imaging. Infrared imaging has been vigorously pursed in parallel with visible imaging because of its utility in military applications. More recently (1997), the CCD camera aboard the Hubble space telescope delivered a deep-space picture, a result of 10 day's integration, featuring galaxies of the 30^{th} magnitude — an unimaginable figure even for astronomers of our generation. Probably, the next effort will be in the big-band age. Thus, photodetectors continue to open to mankind the most amazing new horizons.

Before proceeding to detailed description of detection of optical radiation, it is now appropriate to digress on system considerations concerning photodetection. We would like to determine how good performance is in view of fundamental limits of sensitivity and speed of response, irrespective of the actual type of detector used. Next, different application circuits used in direct detection systems together with elucidation of the design of front-end circuits and discussion of their performance are presented. Third part of the paper is devoted to advanced techniques in photodetection including coherent detection and demonstrating how photodetection is far from being a completely explored field. In the last part, the updated information devoted readout of signals from focal plane arrays is included. It is shown that detector focal plane technology has revolutionized many kinds of imaging in the past 25 years.

2. Detection regimes and figures of merit

The common problem of any type of photon detector (yielding emitted electrons or internal electron-hole pairs as a response to incoming photons) is how to terminate the photodetector with a suitable load resistor, and to trade off the performance between bandwidth and signal-to-noise ratio. This is necessary for a wide family of detectors, including phototubes, photoconductors, photodiodes, CCD's, vidicon targets, etc., all of which are described by a current generator I_{ph} with a stray capacitance C across it.

Let us consider the equivalent circuit of a photodetector ending on a load resistor R_L, as shown in Fig. 1. We indicate current noise generator with diamond shape and asterisk and voltage noise generator with a circle and asterisk. This is the basic circuit for detection, but, as will be shown later, it is often not the best one to give a good compromise between bandwidth and noise. The output signal in voltage $V = IR_L$ or in current I, has a bandwidth, or 3-dB high-frequency cutoff given by

$$\Delta f = \frac{1}{2\pi R_L C}. \tag{1}$$

Two noise contributions are added to the signal. One is the Johnson (or thermal) noise of the resistance R_L, with a quadratic mean value

$$I_{nR}^2 = \frac{4kT\Delta f}{R_L}, \tag{2}$$

* e-mail: rogan@wat.edu.pl

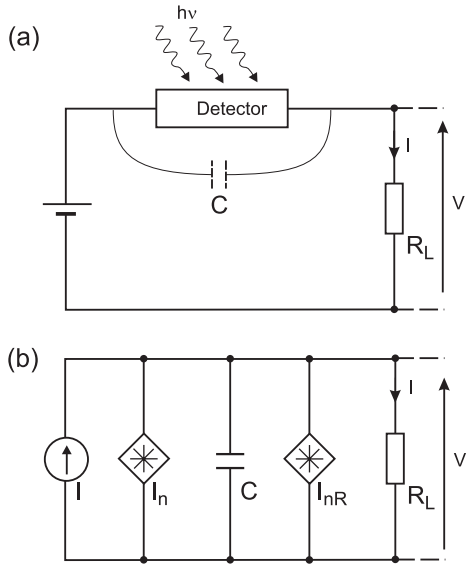

Fig. 1. General circuit (a) of a photodetector and its equivalent circuit (b) including noise generators (noise generators shown with asterisk)

where k is the Boltzmann constant and T is the absolute temperature.

The total current $I = I_{ph} + I_d$ is the sum of the signal current and the dark current. With this current is associated the quantum (or shot) noise arising from the discrete nature of electrons and photoelectrons. Its quadratic mean value is given by

$$I_n^2 = 2q(I_{ph} + I_d)\Delta f, \tag{3}$$

where q is the electron charge and Δf is the observation bandwidth, as in Eq. (1).

A general noise equivalent circuit for a photodetector is shown in Fig. 1(b). The above two fluctuations are added to the useful signal and the corresponding noise generators are placed across the device terminals. Since the two noises are statistically independent, it is necessary to combine their quadratic mean values to give the total fluctuation as

$$I_n^2 = 2q(I_{ph} + I_d)\Delta f + \frac{4kT\Delta f}{R_L}. \tag{4}$$

From Eqs. (1) and (4) it can be seen that bandwidth and noise optimisation impose opposite requirements on the value of R_L. To maximise Δf one should use the smallest possible R_L, whilst to minimise I_n^2 the largest possible R_L is required. A photodetector can have a good sensitivity using very high load resistances (up to GΩ's), but then only modest bandwidths ($\approx$ kHz or less), or it can be made fast by using low load resistances (e.g. $R_L = 50\ \Omega$), but at the expense of sensitivity.

Using Eq. (4), we can evaluate the relative weight of the two terms in total noise current. In general, the best possible sensitivity performance is achieved when the shot noise is dominant compared to the Johnson noise, i.e.

$$2q(I_{ph} + I_d)\Delta f \geqslant \frac{4kT\Delta f}{R_L},$$

which implies the condition

$$R_{L\,min} \geqslant \frac{2kT/q}{I_{ph} + I_d}. \tag{5}$$

From this equation at room temperature it is obtained

$$R_{L\,min} \geqslant \frac{50mV}{I_{ph} + I_d}.$$

So, at low signal levels (for $I_{ph} \ll I_d$), very high values of resistance are required; for example, for $I_d = 5$ pA, not an unusually low dark current, then $R_{L\,min} = 10$ GΩ. However, if we are using a resistor termination value $R_L < R_{L\,min}$ then, the total noise can be written as

$$I_n^2 = 2q(I_{ph} + I_d)\Delta f \left(1 + \frac{R_{L\,min}}{R_L}\right). \tag{6}$$

This means that the noise performance is degraded by factor $R_{L\,min}/R_L$, compared to the intrinsic limit allowed by the dark current level. Thus, using $R_L < R_{L\,min}$ means that the shot-noise performance is reached at a level of current not less than

$$I_{ph} + I_d \geqslant \frac{2kT/q}{R_L}. \tag{7}$$

From this equation, we see that, for a fast photodiode with a 50 Ω load at $T = 300$K, this current has very large value of 1 mA.

The above considerations are valid for photodetectors without any internal gain, G. Let us now extend the calculation of the signal-to-noise ratio (S/N) to photodetectors having internal gain. In this case, the shot noise can be expressed in the form

$$I_n^2 = 2q(I_{ph} + I_d)\Delta f G^2 F, \tag{8}$$

where F is the excess noise factor to account for the extra noise introduced by the amplification process. Of course, in a non-amplified detector, $F = 1$ and $G = 1$. The total noise is a sum of shot noise and thermal noise

$$N = \left[2q(I_{ph} + I_d)\Delta f G^2 F + \frac{4kT\Delta f}{R_L}\right]^{1/2}, \tag{9}$$

and then we obtain a S/N ratio

$$\frac{S}{N} = \frac{I_{ph}}{[2q(I_{ph} + I_d)\Delta f F + (4kT\Delta f/R_L G^2)]^{1/2}}. \tag{10}$$

If we now introduce a critical value I_{ph0}, called the threshold of quantum regime

$$I_{ph0} = I_d + \frac{2kT/q}{R_L F G^2},$$

then, Eq. (10) becomes

$$\frac{S}{N} = \frac{I_{ph}}{[2q(I_{ph} + I_{ph0})\Delta f F]^{1/2}}. \tag{11}$$

Analysing Eq. (11), two detection regimes can be found, according to whether the signal I_{ph} is larger or smaller than I_{ph0}. For the signals, $I_{ph} > I_{ph0}$, and $F = 1$

$$\frac{S}{N} = \left(\frac{I_{ph}}{2q\Delta f}\right)^{1/2}. \tag{12}$$

This S/N is called the quantum noise limit of detection. This limitation cannot be overcome by any detection system, whether operating on coherent or incoherent radiation. In fact, Eq. (12) is a direct consequence of the quantitative nature of light and the Poisson photon arrival statistics.

In the small signal regime, $I_{ph} < I_{ph0}$, we have

$$\frac{S}{N} = \frac{I_{ph}}{(2qI_{ph0}\Delta f)^{1/2}}. \tag{13}$$

That is, the S/N ratio is proportional to the signal, and the noise has a constant value, primarily given by the load resistance. This is the thermal regime of detection.

Figure 2 shows the trend of the S/N ratio [standardized to $(2q\Delta f)^{1/2}$], as a function of the signal amplitude I_{ph}/I_{ph0}. We can notice that, in the thermal regime, the slope is 20 dB/decade up to the threshold $I_{ph}/I_{ph0} = 1$, and from here onward the slope becomes 10 dB/decade in the quantum regime. The effect of an excess noise factor F is also shown in Fig. 2.

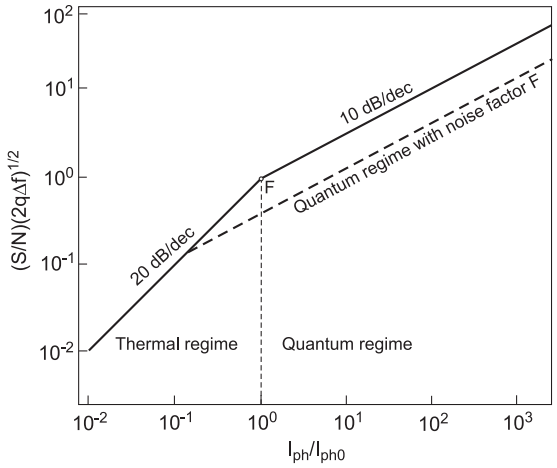

Fig. 2. The S/N ratio of a photodetector, as a function of the input signal, in the thermal and quantum regimes of detection. Reproduced from [1]

The threshold of the quantum regime $I_{ph} = I_{ph0} = I_d + (2kT/q)/R_L F G^2$ is the signal level that corresponds to the break point between thermal and quantum regimes. When the dark current is very small, the second term is the dominant one; any eventual internal gain greatly helps because it scales the load resistance as G^2.

In all cases, I_{ph0} can be interpreted as the equivalent dark current level of the photodetector and has value of

$$I_{ph0} = \frac{2kT/q}{R_{eq}G^2}, \tag{14}$$

with

$$\frac{1}{R_{eq}} = \frac{1}{R_F} + \frac{q}{2kT}G^2 I_d.$$

We define R_{eq} as the noise-equivalent load resistance of the photodetector.

To provide ease of comparison between detectors, certain figures of merit, computed from the measured data, have been defined.

The voltage (or analogous current) responsivity is given by

$$R = \frac{Q_u}{P}, \tag{15}$$

where Q_u is the output quantity supplied by the detector (e.g., a current I_u, a voltage V_u, or any other physical quantity) and P is the incident radiant power.

At equal responsivity, the detector with the smallest output noise Q_u on the useful signal is the most sensitive. Therefore, the first figure of merit for a detector is the NEP — noise equivalent power — defined as the ratio of output noise to responsivity:

$$NEP = \frac{g_n}{R}. \tag{16}$$

So, the NEP represents the input power that gives a unity signal to noise ratio, $S/N = 1$ at the output; that is, a marginal condition of detection.

The better the detector performance is, since the smaller the NEP is. Therefore, it is more convenient to define its inverse as a merit figure. In addition, it should be taken into consideration that whatever the noise source is, it can be expected that the noise quadratic total value is proportional to observation bandwidth Δf and detector area A. Thus is even better to take, as the intrinsic noise parameter of a detector, the ratio $NEP/(A\Delta f)^{1/2}$ normalized to unit area and bandwidth. In order to simplify the comparison of different detectors and to have a parameter that increases as the performance improves, the detectivity D^* (called *D-star*) is defined as:

$$D^* = \frac{(A\Delta f)^{1/2}}{NEP}. \tag{17}$$

This is the fundamental figure of merit used for detectors. It can be transformed to the following equation

$$D^* = \frac{(A\Delta f)^{1/2}}{P}\frac{S}{N}. \tag{18}$$

D^* is defined as the rms signal-to-noise ratio (S/N) in a 1 Hz bandwidth, per unit rms incident radiation power, per square root of detector area. D^* is expressed in $\mathrm{cmHz^{1/2}W^{-1}}$, a unit which has recently been called a "Jones".

As mentioned already, the ultimate performance of detectors is reached when the detector and amplifier noise are low compared to the photon noise. The photon noise is fundamental, in the sense that it arises not from any imperfection in the detector or its associated electronics but rather from the detection process itself, as a result of the discrete nature of the radiation field. The radiation falling on the detector is a combination of that from the target (signal) and that from the background. The practical operating limit for most infrared detectors is the background fluctuation limit, also known as the background limited infrared photodetector (BLIP) limit.

Figure 3 compares typical D^*s of different detectors as a function of wavelength. Also the BLIP and the dark current limits are indicated.

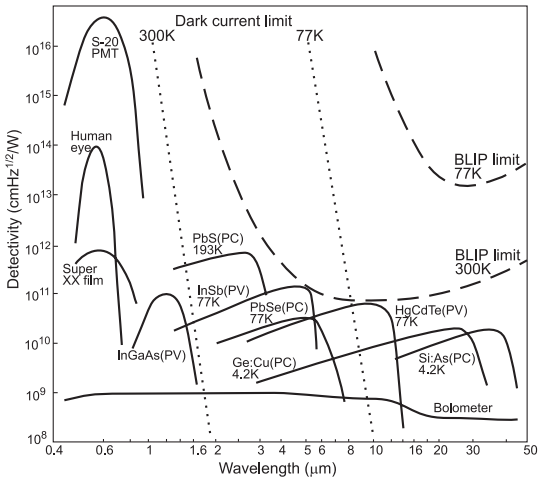

Fig. 3. Detectivity as a function of wavelength for a number of different photodetectors. The BLIP and the dark current limits are indicated. PC — photoconductive detector, PV — photovoltaic detector, and PMT — photomultiplier tube

3. Direct detection systems

3.1. Introduction. A receiver of optical radiation consists of a photodetector, preamplifier, and signal processing circuit (Fig. 4). In a photodetector, the optical signal is converted to an electrical one, which is amplified before further processing. The sensitivity of an optical detection system depends primarily on the first stage of a photoreceiver, i.e., the photodetector and preamplifier.

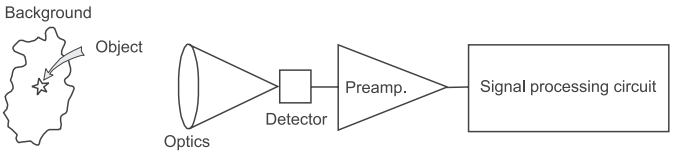

Fig. 4. Block diagram of an optical radiation receiver

A preamplifier should have low noise and a sufficiently wide bandwidth to ensure faithful reproduction of the temporal shape of an input signal. Figure 4 shows, the so-called, direct detection system in which the detector converts the incident radiation into a photosignal that is then processed electronically. It is necessary to minimise the noises from various sources, i.e., background noise, photodetector noise, biasing resistors noise, and any additional noises of signal processing. If further minimisation of noise of the first photoreceiver stages is not possible, advanced methods of optical signal detection can sometimes be used to still recover information carried by optical radiation signals of extremely low power. Heterodyne and homodyne detection can be used to reduce the effects of amplifier noise. Post detection methods are: phase-sensitive detection and synchronous integration of a signal.

3.2. Selection of active amplifying elements. Several types of discrete devices or integrated circuits (ICs) are suitable for the active element in preamplifiers: bipolar (BJT) or field-effect transistor (FET) or an integrated circuit with an input bipolar, FET or MOSFET transistor can be used.

The most important parameter of each receiving device is its signal-to-noise ratio — S/N. Because low-level signals reach the photoreceiver, the noise optimisation of a system — i.e., to obtain maximum S/N is a very important problem [2]. Optimum design of a preamplifier can be obtained by analysis of particular noise sources in a detector-preamplifier circuit. The equivalent input noise V_{ni} will be used to represent all noise sources. A scheme of detector-preamplifier noise circuit is shown in Fig. 5. A level of equivalent noise at the input of this circuit is determined by the detector noise V_{nd}, the background noise V_{nb}, and the preamplifier noise. The preamplifier noise is represented completely by the zero impedance voltage generator V_n in series with the input port, and the infinite impedance current generator I_n in parallel with the input. Typically, each of these terms is frequency dependent. For non-correlated noise components, the equivalent noise at the input of a photodetector-preamplifier circuit is described by the formula

$$V_{ni}^2 = V_{nd}^2 + V_{nb}^2 + V_n^2 + I_n^2 R_d^2, \tag{19}$$

where R_d is the detector resistance.

This single noise source, located at V_s, can be substituted for all sources of the system noise. Note that V_{ni}^2 is independent of the amplifier gain and its input impedance. Thus, V_{ni}^2 is the most useful index against which the noise characteristics of various amplifiers and devices can be compared.

The first transistor of a preamplifier is a dominant noise contributor of the input noise of signal processing circuits (Fig. 5). The effective contribution of this transistor noise also depends on the detector impedance. If

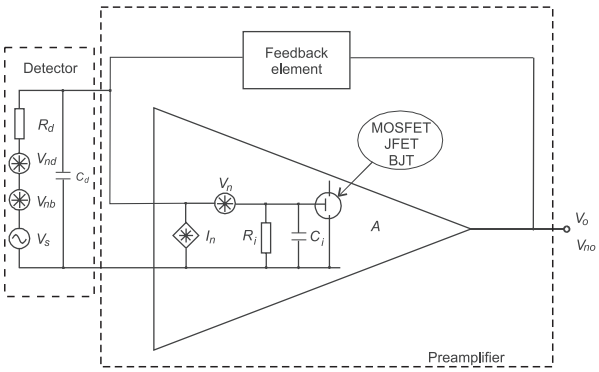

Fig. 5. Noise equivalent diagram of a photodetector-preamplifier circuit; R_i and C_i is the input resistance and the capacity of a preamplifier, respectively, A is the voltage gain of the preamplifier

the transistor of the input stage has a high noise current it will be a bad choice for use with a high-resistance (current source) detector. A low-resistance detector operates best with a preamplifier of low noise voltage.

Figure 6 shows the dependence of the ratio of preamplifier (or input transistor) noise to detector thermal noise as a function of detector resistance, for bipolar, JFET, and MOSFET amplifiers, in a common emitter or common source configuration [3]. It can be noticed that there are some ranges of detector resistance for which the preamplifier noise is lower than the detector thermal noise. For any given detector, the values of the thermal noise voltage V_t and the resistance R_d are determined by the detector type (they cannot be changed). However, it is possible to change the parameters V_n and I_n of the designed preamplifier. Thus, minimisation of the input noise of the circuit is possible. Changes in the values V_n and I_n are done by choosing adequate elements in the preamplifier circuit. For low-resistance detectors (from tens Ω to 1 kΩ), circuits with bipolar transistors at the input are usually used as they have low values of V_n. Sometimes, in order to decrease the value of optimal resistance R_o, a parallel connection of several active elements or even parallel connection of all amplifiers is advantageous [4]. Within the range of average detector resistances from 1 kΩ to 1 MΩ, the preamplifiers FET input stages can be used. However, for connection with high-resistance detectors (above 1 MΩ) especially recommended are the transistors of low I_n values. Such requirements fulfil JFET and MOSFET transistors.

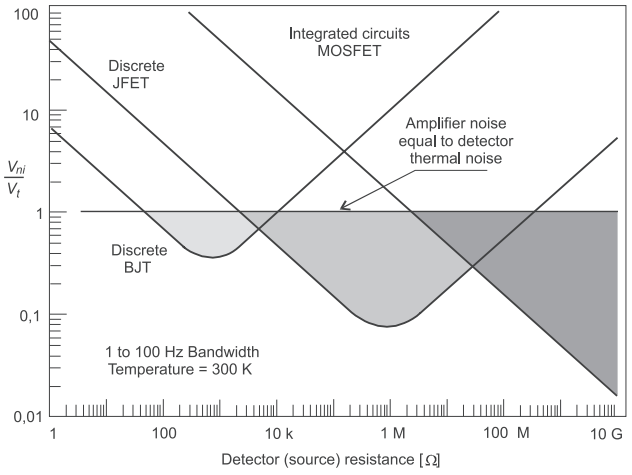

Fig. 6. Dependence of the ratio of preamplifier input voltage noise to detector thermal noise, as a function of detector resistance. Reproduced from [3]

3.3. First stages of photoreceivers. There are two general types of photon detectors without internal gain: photoconductive and junction devices (photovoltaic ones). These photodetectors are used with many types of preamplifiers. The choice of circuit configuration for the preamplifier is largely dependent on the system application. The two basic preamplifier structures — voltage amplifiers type and transimpedance (current in — voltage out type) is discussed. The voltage preamplifiers can be either low impedance or high impedance ones. A simplified circuit diagram is shown in Fig. 7. Bias voltage is supplied by V_b. The detector signal is developed as a voltage drop across R_b. The variable resistance component of the detector R_d is represented by the incremental resistance ΔR_d. The load resistance R_L provides a bias path for the amplifier input.

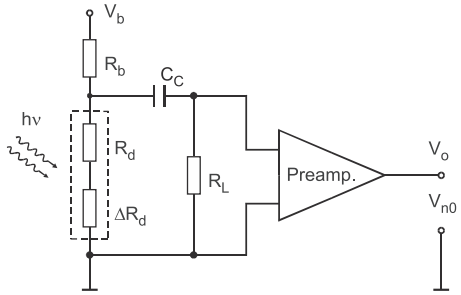

Fig. 7. A simple circuit of voltage mode preamplifier

The simplest preamplifier structure is the low input impedance voltage preamplifier. This design is usually implemented using a bipolar transistor. Either common emitter or grounded emitter input stages may be designed with a reasonably low input impedance. In the low input impedance preamplifiers, the signal source is loaded with a low impedance (e.g., 50 Ω) input stage. The time constant of the temporal response is determined by the combined load resistance and input capacitance of the detector and preamplifier and this determines the detection bandwidth. Preamplifiers of low input resistance can provide high bandwidth but not very sensitive photoreceiver. Photoreceivers with preamplifiers of low input resistance are therefore usually used when wide transmission band is required.

The high impedance preamplifier gives a significant improvement in sensitivity over the low impedance preamplifier, but it requires considerable electronic frequency equalization to compensate for its high-frequency roll-off. The preamplifier has problems of limited dynamic range, being easily saturated at higher input power levels. When a highly sensitive photoreceiver with a low dynamic range is needed, preamplifier of high input impedance is recommended. However, for this, there is a better configuration, that we shall now discuss. This is called the transimpedance receiver.

The transimpedance preamplifier finds many applications in optical signals detection. A schematic of this is shown in Fig. 8. In this circuit, R_d is the detector resistance, and it can be a photodiode or a photoresistor. Depending on detector type and required application, the detector can be biased from V_b or connected directly across the input without bias. The current produced by the detector flows through the resistor R_f located in a feedback loop. The optional potentiometer R_b is used

for setting the zero value of the output voltage without detector illumination [5]. It can be used for compensation of a dc level at the output of the circuit resulting from background radiation. Usually, the input bias current is negligible, so $I_{ph} = I_f$.

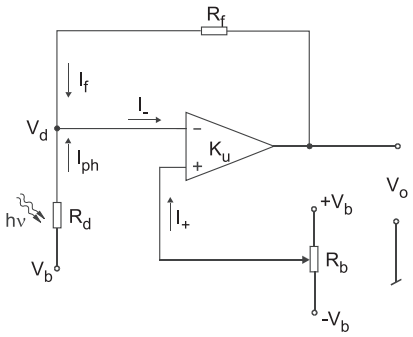

Fig. 8. Schematic diagram of transimpedance amplifier

The voltage at the preamplifier output is given by the formula

$$V_o = -\frac{R_f}{R_d} V_b = -I_{ph} R_f, \tag{20}$$

proper for low frequencies. The great advantage of this configuration is that the preamplifier gives low noise performance without the severe limitation on bandwidth imposed by high input impedance preamplifier. The bandwidth is much higher because the effective load resistance is very low, and has value $R_f/(A+1)$, where A is the gain of the amplifier at the appropriate frequency. It also provides greater dynamic range than the high input impedance structure. It is possible to linearly detect and process optical signals of many orders of powers magnitude.

At low frequencies, using a photodiode in unbiased configuration, across the input of transimpedance preamplifier, avoids the noise and dc offset problems that dark current might otherwise cause (noninverting input is on the ground). In this configuration, there is the small price of increased photodiode capacitance, which would be far more serious at high frequencies. For InGaAs photodiodes can have substantially lower quantum efficiencies if unbiased (due to much thinner depletion region) although the benefit to avoid dark current can be more useful here.

The effective input impedance R_i at the amplifier input as a result of the effects of the feedback circuit is given simply by

$$R_i = \frac{R_f}{A(f)+1} \approx \frac{R_f}{A(f)}, \tag{21}$$

where $A(f)$ is the frequency-dependent gain of the preamplifier. As this input impedance is effectively in parallel with the capacitance $(C_d + C_i)$, the 3 dB breakpoint f_{-3dB} in the frequency response is given by

$$f_{-3dB} = \frac{1}{2\pi R_i (C_d + C_i)} = \frac{A(f)}{2\pi R_f (C_d + C_i)}. \tag{22}$$

If $A(f)$ is described in term of gain-bandwidth product $A(f)\Delta f$ (or GBP) such that $GBP = \Delta f A(f)$, then

$$f_{-3dB} = \frac{GBP/f_{-3dB}}{2\pi R_f (C_d + C_i)}, \tag{23}$$

therefore

$$f_{-3dB} = [GBP/2\pi R_f (C_d + C_i)]^{1/2}. \tag{24}$$

A practical problem with many transimpedance design is to achieve a sufficiency high open-loop gain product, to keep the f_{-3dB} value high, yet still maintain a sufficiently low open-loop phase shift to maintain stability. It often proves necessary, in practice to use a value of R_L, than desired on low-noise grounds, in order to achieve the necessary f_{-3dB}. This gives rise to the widely-held belief that transimpedance designs are inherently noisier than high impedance input design. However, this belief is not valid with good high-gain bandwidth designs. Secondly, even with sub-optimum amplifier design, achieving poorer gain-bandwidth, it is always possible to perform an engineering compromise, by choosing a high value of R_L, to maintain low-noise, and use small amount of equalisation after the transimpedance amplifier to compensate for the resulting roll off.

The preamplifiers describe earlier are used with semiconductor detectors. Now, we would like to present exemplary preamplifiers to the photomultiplier tubes (PMTs).

The operating principles, construction and characteristics of a photomultiplier are given in Refs. 1 and 6. Now, we would like to analyse matching the preamplifier to the photomultiplier. Different pulse processing techniques are typically employed, depending on whether the arrival time or the amplitude (energy) of the detected event must be measured. Three basic types of preamplifiers are available: the current-sensitive preamplifier, the parasitic-capacitance preamplifier, and the charge-sensitive preamplifier. The simplified schematic of the current-sensitive preamplifier is shown in Fig. 9. The $R = 50\ \Omega$ input impedance of the current-sensitive preamplifier provides proper termination of the 50 Ω coaxial cable, and converts the current pulse from the detector to a voltage pulse. The amplitude of the voltage pulse at the preamplifier output will be

$$V_o = RI_iA, \tag{25}$$

where I_i is the amplitude of the current pulse from the detector.

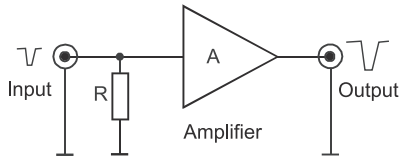

Fig. 9. A simplified schematic of the current-sensitive preamplifier

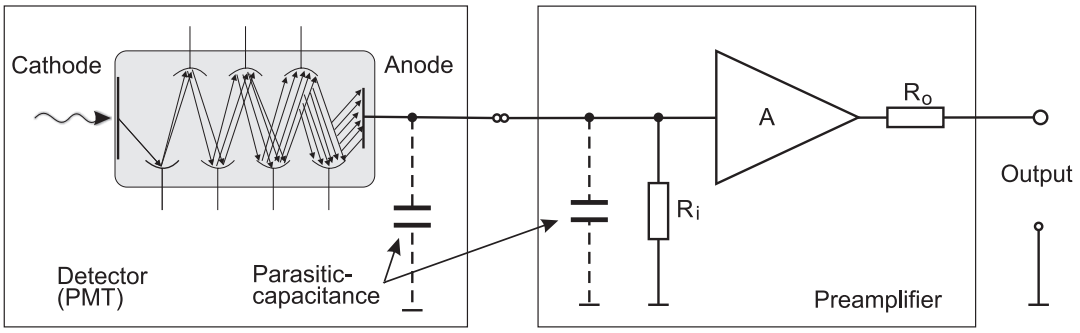

Fig. 10. A simplified diagram of the parasitic-capacitance preamplifier

For counting applications this signal can be fed to a fast discriminator, the output of which is recorded by a counter-timer. For timing application the dominant limitation on timing resolution with photomultiplier tubes is fluctuation in the transit times of the electrons. This causes a jitter in the arrival time of the pulse at the detector output. If the detector signals are small enough to require a current-sensitive preamplifier, the effect of a preamplifier input noise on time resolution must also be considered. It is important to choose a current-sensitive preamplifier which rise time is much faster than detector rise time. Unfortunately, the faster preamplifier does contribute extra noise, because of the unnecessarily wide bandwidth. The excess noise will increase the timing jitter. Choosing a preamplifier rise time that is much slower than the detector rise time reduces the preamplifier noise contribution but causes degradation in pulse rise time and its amplitude. Consequently, the timing jitter becomes worse. The optimum choice depends on the rise time and amplitude of the detector signal. Most current-sensitive preamplifiers designed for timing applications have ac-coupling.

The most effective preamplifier for these detectors is the parasitic-capacitance preamplifier shown in Fig. 10. It has a high input impedance (above 1 MΩ).

The parasitic capacitance is presented by the detector and the preamplifier input (typically 10 to 50 pF). The resulting signal is a voltage pulse having an amplitude proportional to the total charge in the detector pulse. This type of preamplifier is sensitive to small changes in the parasitic capacitance. The rise time is equal to the duration of the detector current pulse. A resistor connected in parallel with the input capacitance causes an exponential decay of the pulse. An amplifier-follower is included as a buffer to drive the low impedance of a coaxial cable at the output. These preamplifiers are highly recommended for photomultiplier tubes, microchannel plate MPTs, and scintillation detectors.

For most of energy spectroscopy applications charge-sensitive preamplifier is preferred (Fig. 11). This preamplifier integrates the charge on the feedback capacitor. Its gain is not sensitive to a change in detector capacitance. The output voltage from the preamplifier V_o, and the decay time constant τ_f, are given respectively by

$$V_o = \frac{Q_d}{Q_f} = \frac{Eq}{\varepsilon C_f} 10^6 \text{ [mV]} \quad \text{and} \quad \tau_f = R_f C_f \tag{26}$$

where Q_d is the charge of the detector, C_f is the feedback capacitor, E is the energy in MeV of the incident radiation, q is the charge of an electron, ε is the amount of energy (eV) required to produce an electron-hole pair in the detector, and 10^6 converts MeV to eV. The resistor R_f should be made as large as possible consistent with the signal energy-rate product and the detector leakage current. The input capacitance must be much greater than the other sources of capacitance connected to the preamplifier input in order for the preamplifier sensitivity to be unaffected by external capacitance changes. The stability of the preamplifier sensitivity is dependent on the stability of the feedback capacitor (C_f is selected for good temperature stability) and the preamplifier open loop gain. The open loop gain will be very large so that small changes in the C_f can be neglected.

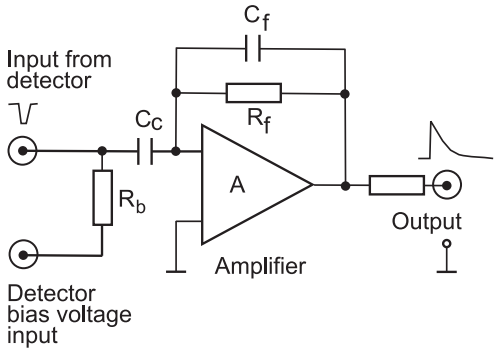

Fig. 11. A simplified diagram of the charge-sensitive preamplifier

The rise time of the output pulse of the preamplifier, in the ideal case, is equal to the charge collection time of the detector. When detectors with very fast collection times or large capacitances are used, the preamplifier itself may limit the rise time of V_o.

3.4. Photon counting techniques. Photomultiplier tubes are also used as a photon counting. Photon counting is one effective way to use a photomultiplier tube for measuring very low light (e.g. astronomical photometry and

fluorescence spectroscopy). A number of photons enters the photomultiplier tube creates an output pulse signal [Fig. 12(a)]. The actual output pulse obtained by the measurement circuit is a dc with fluctuation [Fig. 12(b)].

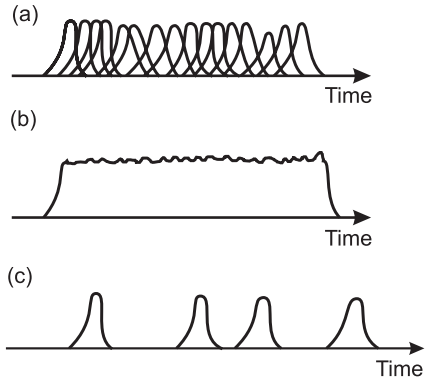

Fig. 12. A pulse signals at the output of a photomultiplier tube

When the light intensity becomes so low that the incident photons are separated as shown in Fig. 12(c), this condition is called single photon event. The number of output pulses is in direct proportion to the amount of incident light and this pulses counting method is advantageous for signal-to-noise ratio and stability over the dc method averaging all the pulses. This counting technique is called the photon counting method.

Since the photomultiplier tube output contains a variety of noise pulses in addition to signal pulses representing photoelectrons, simply counting of the pulses, without some form of noise elimination, will not result in an accurate measurement. The most effective approach to noise elimination is to investigate the height of the output pulses. Figure 13 shows the output pulse and discriminator level.

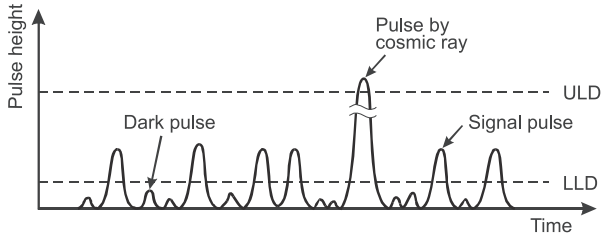

Fig. 13. Output pulse and discriminator level

A typical pulse height distribution (PHD) of the output of the photomultiplier tubes is shown in Fig. 14. In this PHD method, the low level discrimination (LLD) is set at the valley and the upper level discrimination (ULD) at the foot. Most pulses smaller than the LLD are noises and pulses larger then ULD result from interference (e.g. cosmic rays). Therefore, by counting the pulses between the LLD and UPD, accurate light measurements are made possible. In the pulse height distribution, H_m is the mean height of pulses. It is recommended that the LLD be set at 1/3 of H_m and the ULD a triple H_m.

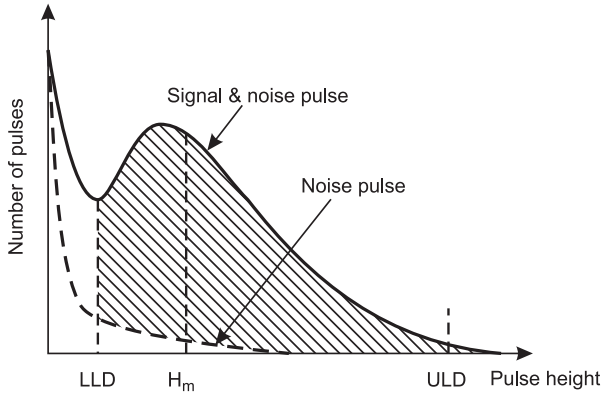

Fig. 14. Typical pulse height distribution

Also the avalanche photodiodes (APDs) can get an output pulses for each detected photons and thus a potentially very high sensitivities, comparable to that of photomultipliers. They are selected to have extremely low noise and low bulk dark-current. They are intended for ultra-low light level application (optical power less than 1 pW) and can be used in either their normal linear mode (polarization voltage $V_R <$ breakdown voltage V_{BR}), or as photon counters in the "Geiger" mode ($V_R > V_{BR}$), where a single photoelectron may trigger an avalanche pulse of about 10^8 carriers.

In the linear mode operation, the APD is well suited for application which requires high sensitivity and fast response time; for example: laser rangefinders, fast receiver modules, lidar, and ultrasensitive spectroscopy.

When biased above the breakdown voltage, an avalanche photodiode will normally conduct a large current. However, if this current is limited to less than the APD's "latching" current, there is a strong statistical probability that the current will fluctuate to zero in the multiplication region, and the APD will then remain in the "off" state until an avalanche pulse is triggered by either a bulk or photo-generated carrier. If the number of bulk carrier generated pulses is low, the APD can therefore be used to count individual current pulses from incident photons. The value of the bulk dark current is therefore a significant parameter in selecting an APD for photocounting, and can be reduced exponentially by cooling.

The APDs can be used in the Geiger mode using either "passive" or "active" pulse quenching circuits.

A passive quench takes the current from the diode and passes it through a load resistor and a series resistor, causing the bias voltage to drop. An active quench uses a transistor for lower and rise the bias voltage. Passive quench circuits, although simple, limit the rate at which photons can be counted, because time has to be allowed for the quench and for the return of the voltage to break down before the next photon arrives. This characteristic time is basically given by RC, where R is the series resistance and C is device capacitance. Because R must

be sufficiently large to trigger the quench, the time is limited by device capacitance to typically hundreds of nanoseconds, and time resolution is at best about 400 ps.

The simplest, and in many cases a perfectly adequate method of quenching a breakdown pulse, is through the use of a current-limiting load resistor. An example of such a "passive" quenching circuit is shown in Fig. 15(a). The load-line of the circuit is shown in Fig. 15(b).

To be in the conducting state at V_{BR}, two conditions must be met:

- the avalanche must have been triggered by either a photoelectron or a bulk-generated electron entering the avalanche region of the APD, to continue to be in the conducting state, and
- a sufficiently large current, called latching current must be passing through the device so there is always an electron or hole in the avalanche region.

For the currents $(V_R - V_{BR})/R_L$ much greater than I_{latch}, the diode remain conducting. If the current $(V_R - V_{BR})/R_L$ is much less than I_{latch}, the diode switches almost immediately to the non-conducting state. When R_L is large, the photodiode is non-conducting, and the operating point is at $V_R - I_d R_L$ in the non-conducting state. Following an avalanche breakdown, the device recharges to the voltage $V_R - I_d R_L$ with the time constant $R_L C$ where C is the total device capacitance including stray capacitance.

The rise-time is fast (e.g. 5 ns to 50 ns), decreases as $V_R - V_{BR}$ increases, and is very dependent on the capacitances of the load resistors, leads, etc. The jitter is typically the same order of magnitude as the rise-time.

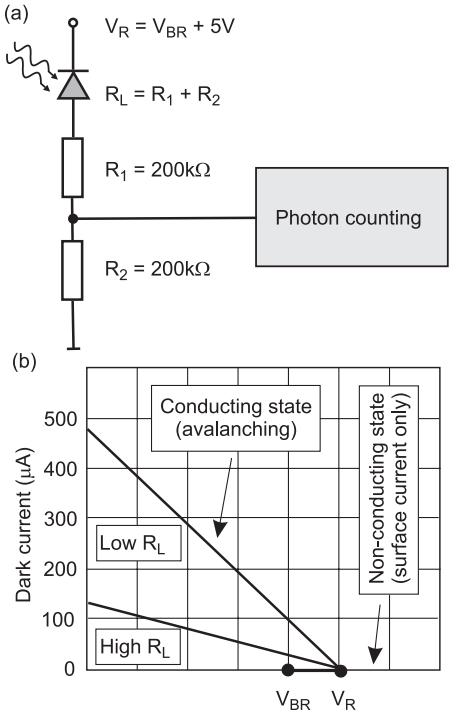

Fig. 15. Passive quenched circuit (a) and the load-line of the circuit (b)

To avoid an excessive dead-time when operating at large voltage above V_{BR}, an "actively quenched" circuit can be used. Active quenching can greatly increase this performance and has become commercially available over the last few years. Now, recharging can be very rapid through a small load resistor. Alternatively, the bias voltage can be maintained but the load resistor is replaced by a transistor which is kept off for a short time after an avalanche, and turned on for a period sufficient to recharge the photodiode. The response time is limited by the transistor switching rather than an RC circuit and has been reduced to as low as 50 ns, and the timing of photons can be made with resolution as high as 20 ps. In this mode, no amplifiers are necessary and single-photon detection probabilities of up to approximately 50% are possible.

Photon-counting is also advantageous where gating coincidence techniques are employed for signal retrieval. Other application in which APD operates in this mode is used in: lidar, astronomical observation, optical fiber test and fault location, optical rangefinding, ultrasensitive fluorescence, etc.

3.5. High-speed photoreceivers. A variety of different applications and measurement techniques have been developed for high-speed detection Most of them are described in Ref. 7. Here we focus mainly on the speed limitation of the response of p-i-n junction and Schottky barrier — two types of photodiodes which have the highest speed response.

In general, in comparison with Schottky barriers, the p-n junction photodiodes indicate some important advantages. The thermionic emission process in Schottky barrier is much more efficient than the diffusion process and therefore for a given built-in voltage, the saturation current in a Schottky diode is usually several orders of magnitude higher than in the p-n junction. In addition, the built-in voltage of a Schottky diode is smaller than that of a p-n junction with the same semiconductor. However, high-frequency operation of p-n junction photodiodes is limited by the minority-carrier storage problem. In other words, the minimum time required to dissipate the carriers injected by the forward bias is dictated by the recombination lifetime. In a Schottky barrier, electrons are injected from the semiconductor into the metal under forward bias if the semiconductor is n-type. Next they thermalize very rapidly ($\approx 10^{-14}$ s) by carrier-carrier collisions, and this time is negligible compared to the minority-carrier recombination lifetime.

In p-i-n photodiode an undoped i-region (p^- or n^-, depending on the method of junction formation) is sandwiched between p^+ and n^+ regions. Because of the very low density of free carriers in the i-region and its high resistivity, any applied bias drops entirely across the i-region, which is fully depleted at zero bias or very low value of reverse bias.

The response speed of p-i-n photodiode is ultimately limited either:

- by transit time across i-region,
- or by circuit parameters (RC time constant).

Influence of diffusion time (τ_d) of carriers to the depletion region, which is inherently a relatively slow process, can be neglected since the generation of carriers occurs mainly in high-field i-region. If the photodetector is not fully depleted, this time is taken for collection of carriers generated outside the depletion layer. Carrier diffusion may result in a tail (a deviation between input and output signal) in the response time characteristics.

The transit time of the p-i-n photodiode is shorter than that obtained in a p-n photodiode even though the depletion region is longer than in the p-n photodiode case due to carriers travel at near their saturation velocity virtually the entire time they are in the depletion region (in p-n junction the electric field peaked at the p-n interface and then rapidly diminished).

The transit time of carriers across i-layer depends on its width and the carrier velocity. Usually, even for moderate reverse biases that carriers drift across the i-layer with saturation velocity. The transit time can be reduced by reducing the i-layer thickness. The fast photodiode must be thin to minimize the time that it takes for the photogenerated electrons and holes to traverse the depletion region. A thin photodiode has high capacitance per unit area and therefore must also be small in diameter to reduce the RC time constant.

The second component of the response time (t_r) is dependent mainly on the photodetector capacitance and input resistance of the preamplifier. The detector capacitance is a function of the area (A), the zero bias junction potential V_o, dielectric constant ($\varepsilon_o\varepsilon_s$), impurity concentrations of acceptors (N_a) and donors (N_d), and the bias voltage (V_b). If we assume that the external bias V_b is large compared to the V_o, and we have p^+-n abrupt junction, then the diode has a capacitance given by

$$C_d = \frac{A}{2}\left(2q\varepsilon_s\varepsilon_o N_d\right)^{1/2} V_b^{-1/2}. \tag{27}$$

In this case, lower RC time constant, and therefore improved bandwidth, can be achieved with the use of smaller detectors area and higher bias voltage. In practice, junctions are rarely abrupt; however, it still remains true that the capacitance decreases with increasing reverse bias.

The time constant is given by

$$\tau_{RC} = \frac{(R_L + R_s)\,R_{sh}}{R_L + R_s + R_{sh}} C = R_{eq}C, \tag{28}$$

where R_s is the series resistance of the photodiode, C is the sum of the photodiode and input preamplifier capacitances, R_{eq} is the equivalent resistance of the photodiode and load resistance.

The rise time depended on time constant is described by the formula

$$t_r = 2R_{eq}C. \tag{29}$$

The p-i-n photodiode has a "controlled" depletion layer width, which can be tailored to meet the requirements of photoresponse and bandwidth. A trade-off is necessary between response speed and quantum efficiency. For high response speed, the depletion layer width should be small but for high quantum efficiency (or responsivity) the width should be large.

Increases in bias voltage will usually increase carrier velocities and therefore reduce transit times but result in higher dark current and noise.

The detector response time is of the form

$$t_T = \left(t_r^2 + t_t^2 + t_d^2\right)^{1/2}. \tag{30}$$

If we take equation (29), the 3 dB cut-off frequency is given by

$$f_{-3dB} = \left(2\pi R_{eq}C\right)^{-1}. \tag{31}$$

Usually, a designer can increase the bandwidth only by using smaller detectors and/or by reducing the amplifiers input resistance.

A good figure of merit for comparing sensitivity to high-speed signals of similar photodiodes is the *Response factor* × *Bandwidth product*. Response factor is the photodiode responsivity multiplied by the impedance seen by the photodiode. Assuming that two photodiodes have single pole high frequency roll-off, the one with the highest *Response factor* × *Bandwidth product* will provide the greatest response to an ultra fast pulse.

3.6. Noise models of first stages of photoreceivers. A noise equivalent circuit of the first stage of a photoreceiver with a voltage type preamplifier is shown in Fig. 16. The signal current generator I_{ph} represents the detected signal. Noises in a detector (photodiode) are represented by three noise generators: I_{nph} is the shot noise originating from a photocurrent, I_{nd} is the shot noise of a dark current, I_{nb} is the shot noise from a background current. If the input resistance of a preamplifier is high, the value of the load resistance depends on the chosen bias resistor. The load (bias) resistor R_L affects both the level of the detector signal and its noise. The noise current generator I_{nR} is the thermal noise current and excess noise of the load resistance R_L. Since the thermal noise of I_{nR} is inversely related to the square root of the resistance, R_L must be large. For the lowest-noise system, at very low frequency, the detector will be the dominant noise source, but at higher frequencies amplifier noise becomes increasingly important.

Expression for signal-to-noise ratio at the first stage of a photoreceiver with voltage type preamplifier results from Fig. 16:

$$\frac{S}{N} = \frac{I_{ph}}{\left[I_{nph}^2 + I_{nd}^2 + I_{nb}^2 + I_n^2 + \frac{4kT\Delta f}{R_L} + \left(\frac{V_n}{R_L}\right)^2\right]^{1/2}}. \tag{32}$$

The numerator represents the photocurrent and the denominator represents the equivalent input noise of photodetector-preamplifier circuit. The first three components determine the noise originating from the photocurrent, the dark current and the background, the fourth component is the current noise of the preamplifier, the fifth is the thermal noise of the resistance R_L and the last one is the voltage noise contribution of the preamplifier.

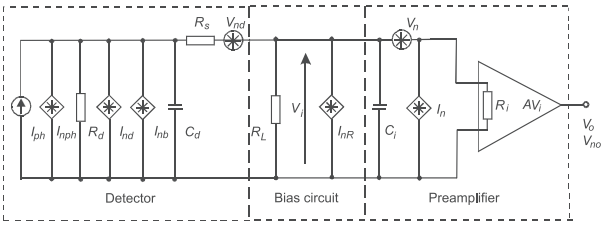

Fig. 16. Equivalent scheme of photodetector-voltage type preamplifier circuit (C_d is the detector capacity, R_i and C_i are the input resistance and capacitance of the preamplifier, respectively)

The noise in electrical circuits is often a function of frequency. For high frequencies, the noise equivalent signal current, $I_{n\,total}$, is given by

$$I_{n\,total}^2 = I_{nph}^2 + I_{nd}^2 + I_{nb}^2 + V_{nd}^2\omega^2 C_d^2 + I_n^2 + \frac{4kT\Delta f}{R_L} + V_n^2\omega^2(C_d + C_i)^2, \tag{33}$$

where V_{nd} is the voltage noise of the serial resistance, R_s.

The input capacitance of the preamplifier, C_i, may be considered to lie across the amplifier input on the photodiode side of the noise generator, V_n. Thus, with this assumption, the noise generator, V_n, is not a true input noise generator, as generally understood, as it should normally lie on the input side of all components (except, possibly, the noise current generator, which would have the same effect whichever side V_n it were to be connected to). The justification for this approach is firstly, that the amplifier capacitance is conveniently grouped with C_d, but secondly it enables V_n to be a "white" noise generator.

Re-writing Eq. (33) we obtain

$$I_{n\,total}^2 = \left[I_{nph}^2 + I_{nd}^2 + I_{nb}^2 + I_n^2 + \frac{4kT\Delta f}{R_L}\right] + \omega^2\left[V_{nd}^2 C_d^2 + V_n^2\left(C_d + C_i\right)^2\right]. \tag{34}$$

There are thus two terms, a "white" noise term in the first setoff square brackets and a second term which gives a noise current increasing in proportional to frequency. Although a capacitor does not add noise, the detector noise voltage (V_{nd}) and preamplifier noise voltage (V_n) is increased by the C_d and the $C_d + C_i$ respectively, as is evident from the coefficient of that term in Eq. 34. Analysing Eq. 34, we see, that for matching an amplifier to a detector, it is important to minimize the sum of $I_n + V_n^2\omega^2(C_d + C_i)^2$.

The sensitivity of an optical receiver is most conveniently expressed in terms of its noise-equivalent power (NEP). This is defined as the optical power necessary to make the signal current, I_{ph}, equal to the noise current, $I_{n\,total}$, i.e.

$$NEP = \left(I_{n\,total}^2\right)^{1/2}\frac{h\nu}{\eta q}. \tag{35}$$

Many data sheets show the NEP figure for a detector, unfortunately. If we connected a preamplifier the performance will often depend on the amplifier noise source, critically combined with parasitic features of the photodiode, particularly its capacitance and parallel resistance [see Eq. (33)].

It is possible to achieve a high value of the signal-to-noise ratio at the first stage of a photoreceiver when using a voltage type preamplifier by using a high resistance value for R_L, and ensuring low current noise I_n, and low voltage noise V_n of the preamplifier. Of course, high resistance value of a R_L causes narrowing of a photoreceiver bandwidth.

Figure 17 presents the noise equivalent scheme of the first stage of a photoreceiver using a transimpedance preamplifier. In this circuit, a noise sources of a detector are identical as for the case shown in Fig. 16, where R_{sh} is the shunt resistance of a detector. Preamplifier noise is represented by the voltage source V_n and the current source I_n. Thermal noises from a feedback resistor are represented by the current source I_{nf}.

From the arrangement in Fig. 17 it can be shown that the equivalent input noise is the square root of the sum of squares of noise components from: the photocurrent I_{ph}, the dark current of a detector I_d, the background current I_b, thermal noise of the resistor R_f, the current I_n, and the voltage V_n noise from a preamplifier. Thus, the signal-to-noise ratio is of the form

$$\frac{S}{N} = \frac{I_{ph}}{\left[I_{nph}^2 + I_{nd}^2 + I_{nb}^2 + I_n^2 + \frac{4kT\Delta f}{R_f} + \left(\frac{V_n}{R_f}\right)^2\right]^{1/2}}. \tag{36}$$

For high frequencies, the last term of the denominator should contain parallel combination of all impedances across the input of the preamplifier, e.g., R_f, R_d, $(\omega C_d)^{-1}$, and $(\omega C_i)^{-1}$.

In a photoreceiver using transimpedance preamplifier, an identical bandwidth can be obtained by choosing a feedback resistance R_f much higher than the resistance $R = R_d \| R_L \| R_i$ which would be possible in a simpler photoreceiver with a voltage type preamplifier. Thus, comparing formulae (32) and (36), it can be noticed that for the same bandwidth the signal-to-noise ratio is higher for transimpedance preamplifier than for a classic one. In practice, it means that transimpedance amplifiers can have wider bandwidths yet retain the low noise characteristics of high-impedance preamplifiers.

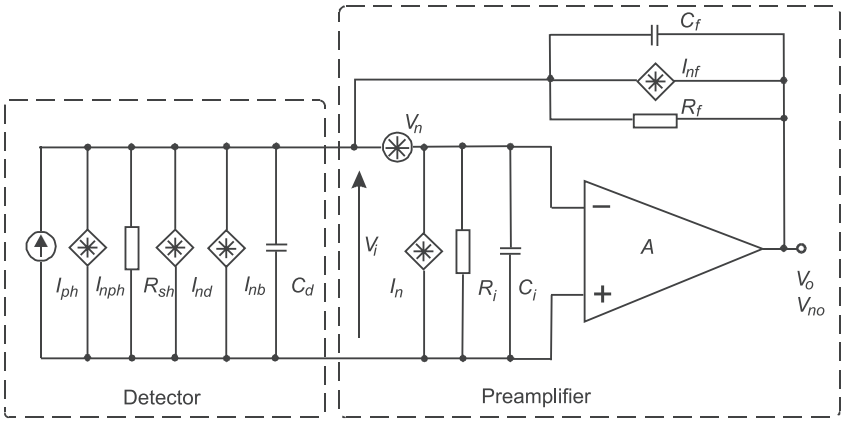

Fig. 17. Equivalent scheme of the first stage of a photoreceiver with transimpedance preamplifier. Reproduced from [6]

In p-n and p-i-n photodiodes, the basic source of noise is shot noise originating from the photocurrent I_{ph}, the dark current I_d, and the background radiation current I_b. In these photodetectors, the thermal noises of detector resistance and noises of active elements of a preamplifier can also play a significant role.

All the preamplifier noises V_n and I_n can be substituted for one equivalent current noise:

$$I_a^2 = \frac{1}{\Delta f}\int_0^{\Delta f}(I_n^2 + V_n^2|Y|^2)df, \tag{37}$$

where Y is the input admittance of an amplifier.

As stated earlier, photodiodes can operate with either voltage or transimpedance preamplifiers. For the equivalent schemes shown in Figs. 16 and 17, the signal-to-noise ratio is of the form:

$$\frac{S}{N} = \frac{I_{ph}}{\left[2q\left(I_{ph}+I_d+I_b\right)\Delta f + \frac{4kT\Delta f}{R_L} + I_a^2\right]^{1/2}}. \tag{38}$$

The first term represents the shot noise component of the photocurrent, the dark current and the background, whereas the second term is the thermal noise of load resistance of a photodetector, and the third term is the preamplifier noise. For the circuit shown in Fig. 17, the resistance R_L is the feedback resistor, R_f.

Let's consider a few special cases. Assuming that the signal current is higher than the dark current, I_d can be omitted in Eq. (38). This is valid when the dark current is insignificant or if the received optical power is high. The assumption can also be made that the shot noise significantly exceeds thermal noise when the optical power has a high level. For low frequency FET transimpedance receivers, the shot noise limit is usually obtained when the output signal exceeds 50 mV. It means that the term $4kT\Delta f/R_L$ can be omitted. If additionally $I_a^2 \ll 2q\Delta f(I_{ph}+I_d)$, the expression for signal-to-noise ratio can be simplified to the form

$$\frac{S}{N} = \left(\frac{I_{ph}}{2q\Delta f}\right)^{1/2} = \left(\frac{\eta\Phi_e\lambda}{2hc\Delta f}\right)^{1/2} = \left(\frac{\eta A E_e\lambda}{2hc\Delta f}\right)^{1/2}, \tag{39}$$

where λ is the wavelength of incident radiation, Φ_e is the incident radiant flux (in W) and E_e is the detector's irradiance. A photoreceiver, the signal-to noise ratio which is described by formula (39), is limited only by a shot noise. This noise is also called quantum limited one — see Eq. 12. Unfortunately, it is not always the case that the high optical power reaches a photoreceiver. If the power of an optical signal is low, the shot noise is negligible in relation to the thermal noise, then

$$\frac{S}{N} = \frac{\eta q\Phi_e}{h\nu}\left(\frac{R_L}{4kT\Delta f}\right)^{1/2}. \tag{40}$$

It is evident that when a photoreceiver is limited by the thermal noise, it is thermally dependent (see Eq. 13).

When thermal noise is limited, it can be seen by analysing equation (40) that the signal-to-noise ratio increases directly in proportion to the received optical power. Thus, in the range of a thermally dependent photoreceiver, small changes, e.g., of path transmission efficiency will cause significant differences in the signal-to-noise ratio of the received signal. In the quantum limited systems, an increase in the optical power by $\Delta\Phi_e$ [dB] gives an improvement in the signal-to-noise ratio of only half the change $\Delta\Phi_e$ when expressed in dB.

One way of overcoming the preamplifier noise is to use a detector having a high degree of internal gain, prior to connection to this device. There are two main types of detector that can be used; the photomultiplier, which is rather large and inconvenient, and the avalanche photodiode. The main advantages are apparent at high frequency, when amplifier noise contribution is more troublesome, because of stray capacitances lowering the impedance across the amplifier input.

Let's consider the signal-to-noise ratio in the first stage of photoreceiver with an avalanche photodiode. The high sensitivities of avalanche photodiodes can be obtained due to a phenomenon of avalanche multiplication, which significantly increases the current signal generated at the output of the detector and improves the signal-to-noise ratio. Of course, it does not influence on the noises from the load resistance and the amplifier noises, but increases

the signal as usually only noise originating from the signal current and dark current (quantum noise) are dominant. Unfortunately, at high gain levels, the random mechanism of carrier multiplication (M) introduces excess noise, in form of higher shot noise, which eventually exceeds the noises level resulting only from primary generation of unequilibrium carriers.

At moderate gain levels, the dominant source of the noise in an avalanche photodiode is the signal and dark current shot noises, which are multiplied. However, if the signal power is increases by a factor M^2, the noise power increases by a factor M^{2+x}. The factor x is typically between 0.3 and 0.5 for silicon APDs and between 0.7 and 1.0 for germanium and III-V alloy APDs.

Knowing the total noise, a signal-to-noise ratio can be determined as

$$\frac{S}{N} = \frac{MI_{ph}}{\left\{2q\Delta f[I_{ph}M^{2+x} + I_s + (I_b + I_{db})M^{2+x}] + \frac{4kT\Delta f}{R_L}F_n\right\}^{1/2}}. \tag{41}$$

The numerator of this equation determines the photocurrent and denominator noises. The first term of the denominator is a shot noise term and the second one represents the thermal noise of load resistance with a preamplifier noise (F_n is the noise factor of preamplifier, I_{db} is the bulk leakage current component of primary dark current, and I_s is the surface leakage current of a dark current). Shot noise components except the surface leakage component of the dark current are multiplied, so the photocurrent, current from a background, and bulk leakage current components of the primary dark current are multiplied. When M is large, the thermal and amplifier noise term becomes insignificant and the S/R ratio decreases with increasing M. Therefore an optimum value of the multiplication factor M_{op} exists which maximises the S/N ratio.

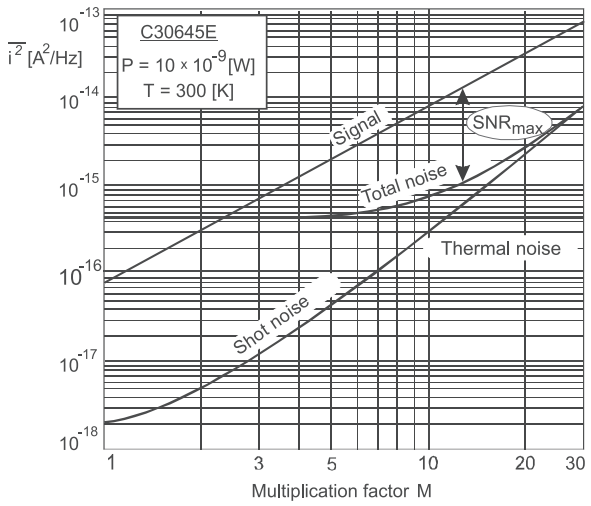

Fig. 18. Dependence of signal, thermal noise, shot noise and total noise on avalanche multiplication factor

Figure 18 presents the typical dependence of signal current, thermal noise, shot noise and total noise on the avalanche multiplication factor [8]. For low values of multiplication factor, the signal amplitude is usually lower than the total noise amplitude. In this range, thermal noises are dominant at the low optical input levels usually used with APDs. For high multiplication values, thermal noises become less important but shot noises are then significant. There is a supply voltage for which a distance between line representing the signal and the curve representing the total noise is the largest, corresponding to when the S/N ratio is maximal. Of course, these parameters will change with temperature, because the dark current of a photodiode and the avalanche multiplication factor are both strongly thermally dependent.

4. Coherent detection

So far, we have considered systems that were based on the modulation of the light intensity in a transmitter and direct detection of this in a photoreceiver. It was not essential for the modulated light wave to be a coherent wave and its spectrum could be wide. These systems are simple and cheap but they have constrains on their transmission possibilities. The photoreceivers decoded only the information connected with the intensity or with the square of the electromagnetic field amplitude whereas information can be carried also by its phase and frequency. The possible photoreceiver sensitivity results from the basic noise limits, as the noise of the photodetector, preamplifier, and background.

To improve the signal-to-noise ratio, it would be an advantage to increase the photocurrent at the detector output. We have already noted that using photoreceivers with avalanche photodiodes had constraints resulting from additional multiplication noises. To avoid problems with direct detection, coherent detection, a method of receiving based on interference of two beams of coherent laser radiation can be used [9–11]. Figure 19 illustrates the differences between coherent detection and direct detection.

Figure 19(a) presents a direct detection system. To narrow the received optical bandwidth we have applied a filter to limit the spectral range of radiation reaching the detector — Fig. 19(b). For example, optical filters based on the Bragg effect can have 5-nm bandwidth for 1.56 μm wavelengths, corresponding to a detection band, Δf, equal to 600 GHz [12]. This wide bandwidth, of what is a reasonably narrow optical filter illustrates that application of an optical filter cannot easily ensure narrow wavelength selection of a photoreceiver. A widely-spaced Fabry-Perot filter can achieve bandwidth of several hundred MHz, but is costly and difficult to stabilise, so impractical for most systems. However, the simple addition of a beam-splitter and an additional coherent light source, a so-called, local oscillator, provides a coherent

detection scheme, that can use either heterodyne and homodyne detection systems [see Fig. 19(c)].

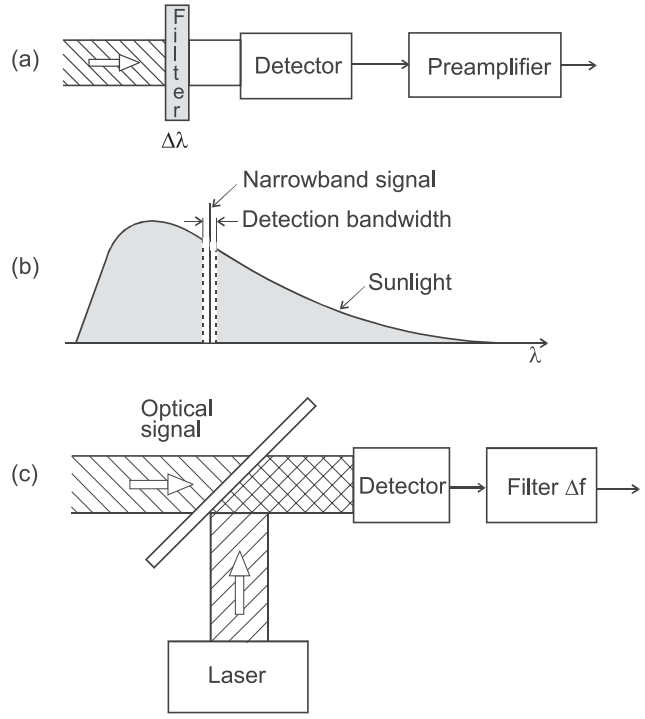

Fig. 19. Comparison of coherent versus incoherent optical detection

Heterodyne detection technique has been commonly used for many years in commercial and domestic radioreceivers and also for the microwave range of an electromagnetic spectrum. The main virtues of this method detection are higher sensitivity, higher and more easily obtained selectivity, plus the possibility of detection of all types of modulation and easier tuning over wide range [13]. Coherent optical detection has been developed since 1962, but compact and stable production of this system is more difficult, and the system is more expensive and troublesome than its radio-technique equivalent. The basic block diagram of heterodyne optical receiver is shown in Fig. 20.

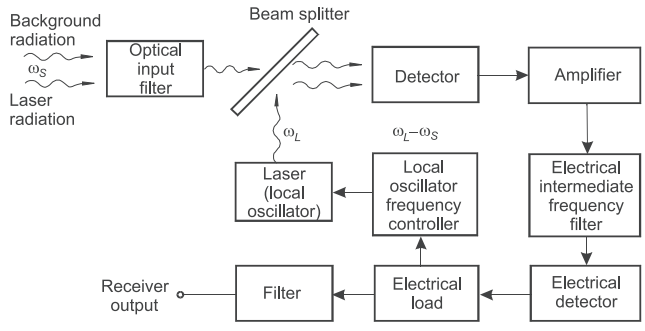

Fig. 20. Block diagram of heterodyne detection optical receiver

Laser radiation containing information, is after being passed through an input optical filter and beam splitter, arranged to coherently combine or "mix" with a light beam of a local oscillator at the detector surface. A beam-splitter can be made, in many ways, the simplest being a glass plate with adequate refraction coefficient. In a general case, a device fulfilling such a role is called a direction coupler, as an analogy to microwave or radio devices. A detector used for signal mixing has to have a square-law characteristic to detecting electronic field of the light, but this is conveniently typical of most optical detectors (photodiode, photoconductor, photomultiplier, APD, etc.). This signal is next amplified. An electrical filter of intermediate frequency extracts the desired difference component of the signal which next undergoes a demodulation process. The design and operation principle of the subsequent electrical detector depends on the nature of the modulation of a signal. The signal from a load resistance passes through an output filter to a receiver output and by means of a local oscillator frequency controller it controls a laser. A frequency control loop is used for the local oscillator laser to maintain a constant frequency difference $\omega_L - \omega_s = \omega_p$ with the input signal. An indispensable condition for efficient coherent detection is to match the polarisation, and to the shape of both waveforms of both beams to match the profile of the detector surface.

Expression for the signal-to-noise ratio at the heterodyne detection system with APD for $P_L \ll P_S$ is given by

$$\frac{S}{N} = \frac{\sqrt{2}R_i M\sqrt{P_S P_L}}{\left(2q\Delta f M^{2+x} R_i P_L + \frac{4kT\Delta fF}{R_L}\right)^{1/2}} \approx \left(\frac{R_i P_S}{q\Delta f M^x}\right)^{1/2}. \tag{42}$$

Assuming a photodiode responsivity $R_i = \eta q/h\nu$, we have

$$\frac{S}{N} = \frac{\eta P_S}{h\nu\Delta f M^x} = 2\left(\frac{S}{N}\right)_{quanta}. \tag{43}$$

Figure 21 shows a comparison of coherent detection sensitivity (solid lines) with the sensitivity of direct p-i-n photodiode detection ($M = 1$) for the same values of signal-to-noise ratio. Significant improvement in sensitivity can be observed for weak signals. Higher sensitivity of a detector ensures qualitatively better detection as increased information bit rate or can permit longer communications links to be used between each regenerator circuit. In long-distance fibre telecommunications, however, the use of optical fibre amplifier has taken much of the impetus from development of the coherent receiver, although the latter still have the unique advantage of highly selective narrowband detection.

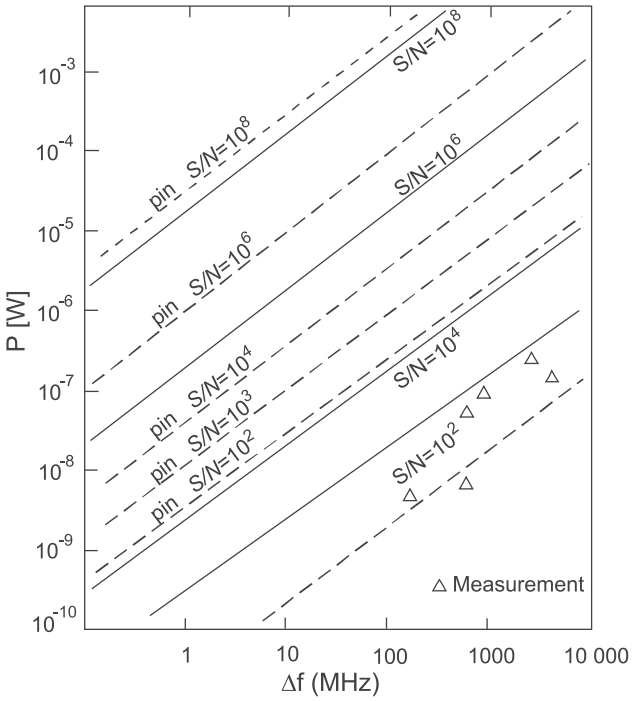

Fig. 21. Sensitivity of the coherent photoreceiver (dashed line) and p-i-n photodiode (solid line). Reproduced from [14]

In heterodyne detection, the spectrum of laser modulation was shifted into an intermediate frequency (IF) range, so selectivity of a photoreceiver depends on the bandwidth of the IF amplifier. This is arranged electronically, so it can easily be sufficiently narrow. Having narrow IF circuit bandwidth is especially important for detection of multichannel signals.

In practice, the technique of heterodyne detection is used for construction of Doppler velocimeters and laser rangefinders and as well as in spectroscopy (particular LIDAR systems). It may yet find application in more telecommunications systems.

If a signal frequency is equal to the frequency of a local oscillator, the IF frequency equals zero. It is a special case of coherent detection, so-called, homodyne detection.

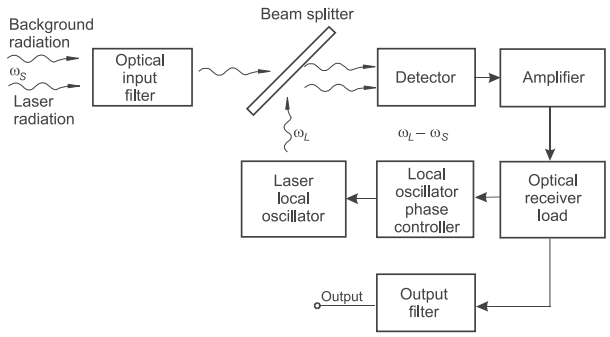

Fig. 22. Block diagram of homodyne detection optical receiver

In a homodyne detection optical receiver (Fig. 22), the incoming laser carrier is again combined with a reference wave from a local laser on a photodiode surface, but in this case both frequencies are the same. It does not contain two blocks, filter of IF frequency and demodulator which were in heterodyne receiver.

The photodetector current in a homodyne receiver is given by:

$$I_{hom} = R_i M(P_S + P_L) + 2R_i M(P_S P_L)^{1/2} \cos\phi_p(t). \quad (44)$$

The first component is a direct-current component but the second one contains the useful information regarding the optical signal. The current at the detector output increases with increase in local oscillator power and with the optical receiver responsivity.

If the local oscillator power is high, the shot noise originating from a signal current, thermal noise and dark-current noise can be omitted. For amplitude modulation we have

$$\frac{S}{N} = \frac{\left(2R_i M(P_S P_L)^{1/2}\right)^2 R_L}{\left(2q\Delta f M^{2+x} R_i P_L + 4kT\Delta f F/R_L\right) R_L} \approx \frac{2R_i P_S}{q\Delta f M^x} = \frac{2\eta P_S}{h\nu\Delta f M^x}. \quad (45)$$

As can be seen, the signal-to-noise ratio for homodyne detection is twice as high as heterodyne detection. This is basically because the homodyne detector allows direct addition or subtraction of the electrical fields, depending on whether the signal and local oscillator are in phase or 180° out of phase. With heterodyne detection, the relative phases change linearity with time, and mixing of signals is not effective when signals have 90° or 270° phase difference.

As it results from Eq. (44), homodyne detection derives the baseband modulation signal carrying the information directly. Thus, further electronic demodulation is not required.

Homodyne receivers are used in the most sensitive coherent systems. In practice, construction of such receivers is difficult because

- the local oscillator must be locked to keep a constant zero phase difference to the incoming optical signal, and this require excellent spectral purity,
- power fluctuation of the local laser must be eliminated.

Constant difference of both laser phases can be achieved using an optical phase-locked loop.

5. Focal plane arrays

Many materials have been investigated to fabricate photodetectors [15–19]. Figure 23 shows the quantum efficiency of some of the detector materials used to fabricate arrays of ultraviolet (UV), visible and infrared detectors. AlGaN detectors are being developed in the UV region. Silicon p-i-n diodes are shown with and without antireflection coating. Lead salts (PbS and PbSe) have intermediate quantum efficiencies, while PtSi Schottky barrier types and quantum well infrared photodetectors (QWIPs) have low values. InSb can respond from the near

UV out to 5.5 μm at 80 K. A suitable detector material for near-IR (1.0–1.7-μm) spectral range is InGaAs lattice matched to the InP. Various HgCdTe alloys, in both photovoltaic and photoconductive configurations, cover from 0.7 μm to over 20 μm. Impurity-doped (Sb, As, and Ga) silicon impurity-blocked conduction (IBC) detectors operating at 10 K have a spectral response cut-off in the range of 16 to 30 μm. Impurity-doped Ge detectors can extend the response out to 100–200 μm.

The term "focal plane array" (FPA) refers to an assemblage of individual detector picture elements ("pixels") located at the focal plane of an imaging system. Although the definition could include one-dimensional ("linear") arrays as well as two-dimensional (2-D) arrays, it is frequently applied to the latter. Usually, the optics part of an optoelectronic images device is limited only to focusing of the image onto the detectors array. These so-called "staring arrays" are scanned electronically usually using circuits integrated with the arrays. The architecture of detector-readout assemblies has assumed a number of forms which are discussed below. The types of readout integrated circuits (ROICs) include the function of pixel deselecting, antiblooming on each pixel, subframe imaging, output preamplifiers, and may include yet other functions. Infrared imaging systems which use 2-D arrays belong to so-called "second generation" systems.

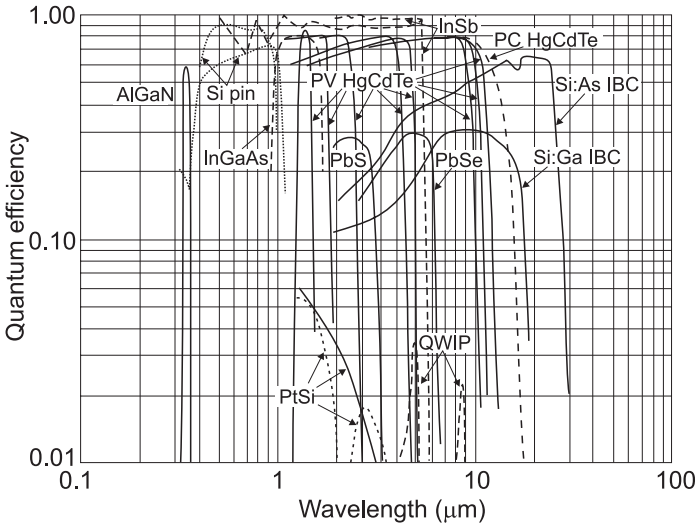

Fig. 23. Quantum efficiency of UV, visible, and infrared detector arrays

Development in detector FPA technology has revolutionized many kinds of imaging in the past twenty five years [20]. From γ rays to the infrared and even radio waves, the rate at which images can be acquired has increased by more than a factor of a million in many cases. Figure 24 illustrates the trend in array size over the past thirty years. Imaging FPAs have developed in proportion to the ability of silicon integrated circuit (ICs) technology to read and process the array signals, and with ability to display the resulting image. FPAs have nominally the same growth rate as dynamic random access memory (DRAM) ICs (which have had a doubling–rate period of approximately 18 months; it is a consequence of Moore's Law, which predicts the ability to double transistor integration on each IC about every 18 months) but lag behind in size by about 5–10 years. ROICs are somewhat analogous to DRAM–only readouts, but require a minimum of three transistors per pixel, compared to one for each memory cell. Readouts are also analogous in terms of an emphasis on low noise inputs and generally maximum charge storage capacity. Charge coupled devices (CCDs) with close to 100 M pixels offer the largest formats. PtSi, InSb and HgCdTe have been following the pace of DRAM. In the infrared, 4 M pixel arrays are now available for astronomy applications.

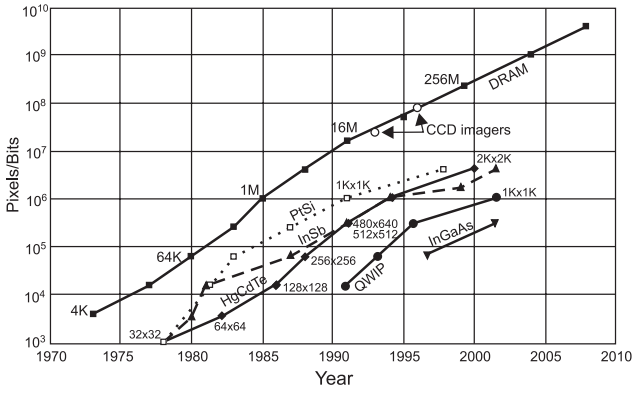

Fig. 24. Increase in array format size over the past thirty years. Reproduced from [20]

5.1. Monolithic arrays. In general, the architectures of FPAs may be classified as monolithic and hybrid. In the monolithic approach, some of the multiplexing is done in the detector material itself rather than in an external readout circuit. The basic element of a monolithic array is a metal-insulator-semiconductor (MIS) structure as shown in Fig. 25(c). Used as part of a charge transfer device, a MIS capacitor detects and integrates the generated photocurrent.

There are a few obvious advantages to the monolithic structure, principally in the simplicity and lower cost associated with a directly integrated structure. Common examples of these FPAs in the visible and near infrared (0.7–1.0 μm) are found in camcorders and digital cameras. Two generic types of silicon technology provide the bulk of devices in these markets: charge coupled devices (CCDs) and complementary metal-oxide-semiconductor (CMOS) imagers. CCD technology has achieved the highest pixel counts or largest formats with numbers approaching 10^8 (see Fig. 26). CMOS imagers are also rapidly moving to large formats and are expected to compete with CCDs for the large format applications within a few years.

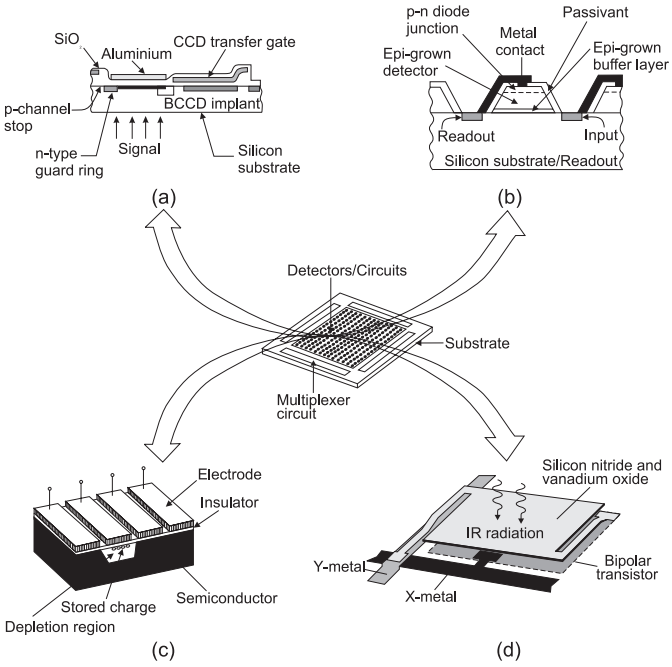

Fig. 25. Monolithic focal plane arrays: (a) all-silicon; (b) heteroepitaxy-on-silicon; (c) non-silicon (e.g., HgCdTe CCD); and (d) microbolometer

An example CCD array made by Dalsa is illustrated in Fig. 27. This array has approximately 25 M pixels, and can operate at 2.5 frames per second with 8 outputs. Although monolithic structures have been proposed for a wide variety of infrared detector materials over the past 30 years, only a few have been demonstrated. These include PtSi, and more recently PbS, PbTe, and uncooled microbolometers. Uncooled detector arrays are revolutionizing the infrared imaging community, by eliminating the need for expensive and high-maintenance coolers that are required by traditional infrared detectors. At present they can be built in formats as large as 480×640 array with 25×25 μm^2 pixels.

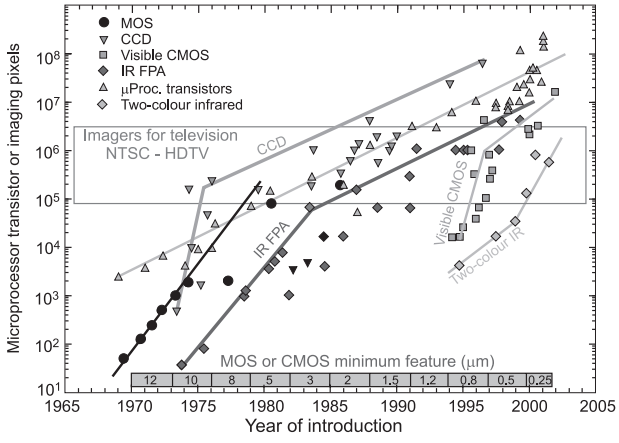

Fig. 26. Imaging array formats compared with the complexity of microprocessor technology as indicated by transistor count. The timeline design rule of MOS/CMOS features is shown at the bottom. Reproduced from [20]

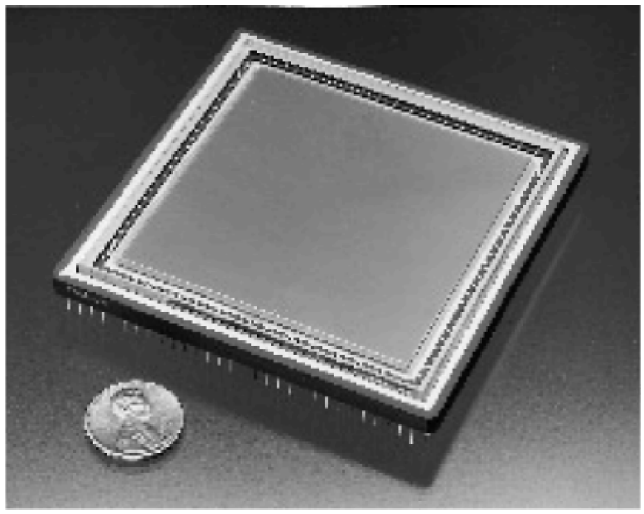

Fig. 27. Silicon CCD visible array with 5040×5040 pixels. Pixel size is 12×12 μm^2. The chip size is 60×60 μm^2. Quantum efficiency is $> 20\%$ at 900 nm. Reproduced from [21]

CCD technology is very mature in respect to both the fabrication yield and the attainment of near-theoretical sensitivity. Figure 28 shows the schematic circuit for a typical CCD imager. Incident radiation generates electron-hole pairs in the depletion region of the MIS structure. The photogenerated carriers are first integrated in an electronic well at the pixel and subsequently transferred to slow and fast CCD shift registers. At the end of the CCD register, a charge carrying information on the received signal can be readout and converted into a useful signal. More information about CCD operation can be found in Ref. 22.

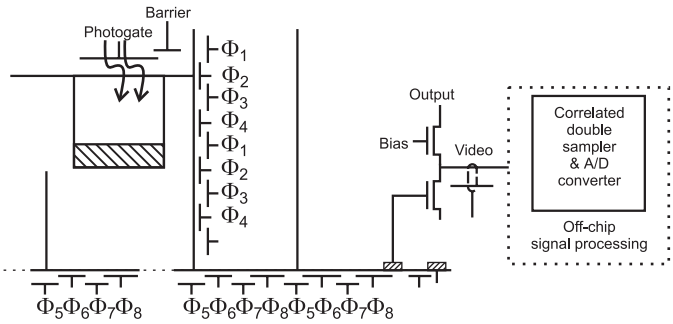

Fig. 28. Architecture of typical CCD image. Reproduced from [23]

At present, the following readout techniques are used in CCD devices:

- floating diffusion amplifier in each pixel,
- system with correlated double sampling (CDS),
- floating gate amplifier.

The floating diffusion amplifier, a typical CCD output preamplifier, can be implemented in each unit cell as shown in dotted box in Fig. 29. The unit cell consists of three transistors and the detector. Photocurrent is integrated onto the stray capacitance, which is the combined

capacitance presented by the gate of the source follower $T2$, the interconnection, and the detector capacitance. The capacitance is reset to the voltage level V_R by supplying the reset clock (Φ_R) between successive integration frames. Integration of the signal charge makes the potential of the source follower input node lower. The source follower is active only when the transistor $T3$ is clocked. The drain current of the source follower $T2$ flows through the enable transistor $T3$ and load resistor outside the array.

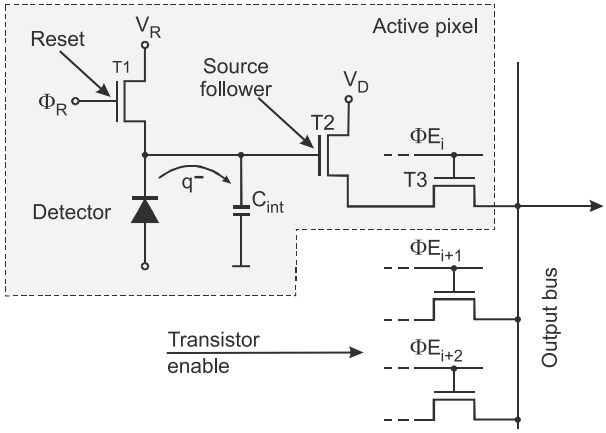

Fig. 29. Unit cell with floating diffusion amplifier. Reproduced from [10]

Figure 30 shows a preamplifier, in this example — the source follower per detector (SFD), the output of which is connected to a clamp circuit. The output signal is initially sampled across the clamp capacitor during the onset of photon integration (after the detector is reset). The action of the clamp switch and capacitor subtracts any initial offset voltage from the output waveform. Because the initial sample is made before significant photon charge has been integrated, by charging the capacitor, the final integrated photon signal swing is unaltered. However, any offset voltage or drift present at the beginning of integration is, by the action of the circuit subtracted from the final value. This process of sampling each pixel twice, once at the beginning of the frame and again at the end, and providing the difference is called correlated double sampling (CDS).

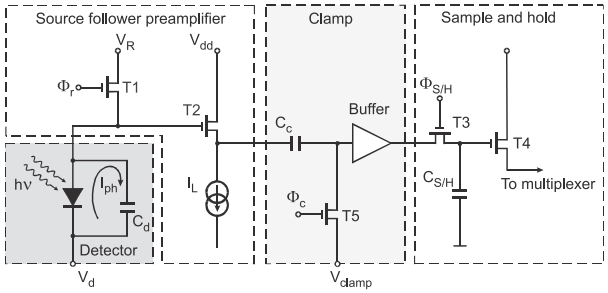

Fig. 30. Correlated double sampling circuit. Reproduced after [3]

The value of the initial CDS sample represents dc offsets, low frequency drift and $1/f$ noise, and high-frequency noise; this initial value is subtracted from the final value, which also includes dc offset, low-frequency drift, and high-frequency noise. Since the two samples occur within a short period of time, the dc and lower-frequency drift components of each sample do not change significantly; hence, these terms cancel in the subtraction process.

The floating gate amplifier configuration is shown in Fig. 31. It consists of two MOSFET transistors, the source follower $T2$ and the zeroing transistor $T1$. The floating gate (reading gate) is in the same row as the CCD transfer gates. If a moving charge is under the gate, it causes a change in the gate potential of the transistor of the gate $T2$. At the preamplifier output, a voltage signal appears. This manner of readout does not cause degradation or decay of a moving charge so, the charge can be detected at many places. An amplifier, in which the same charge is sampled with several floating gates is called a floating diffusion amplifier.

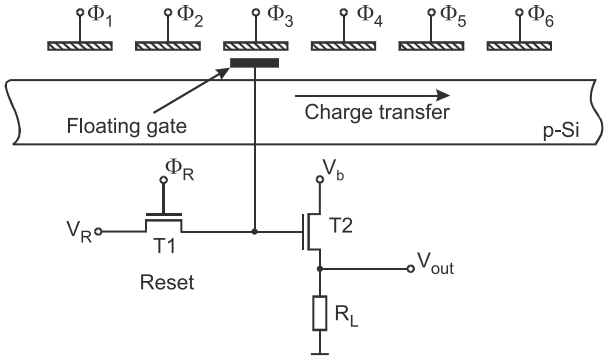

Fig. 31. Floating gain amplifier circuit. Reproduced after [24]

The configuration of CCD devices requires specialized processing, unlike CMOS imagers which can be built on fabrication lines designed for commercial microprocessors. CMOS have the advantage that existing foundries, intended for application specific integrated circuits (ASICs), can be readily used by adapting their design rules. Design rules of 0.18 μm are in production, with pre-production runs of 0.13 μm design rules already underway. As a result of such fine design rules, more functionality has been put into the unit cells of multiplexers and smaller unit cells, leading to large array sizes. Figure 26 shows the timelines for minimum circuit features and the resulting CCD, IR FPA and CMOS visible imager sizes with respect to imaging pixels. Along the horizontal axis there is also a scale depicting the general availability of various MOS and CMOS processes. The ongoing migration to even finer lithographies will thus enable the rapid development of CMOS-based imagers having even higher resolution, better image quality, higher levels of integration and lower overall imaging system cost than CCD-based solutions. At present, CMOS

having with minimum features of $\leqslant 0.5$ μm is also enabling monolithic visible CMOS imagers, because the denser photolithography allows low-noise signal extraction and high performance detection with the optical fill factor within each pixel. The silicon wafer production infrastructure which has put personal computers into many homes is now enabling CMOS-based imaging in consumer products such as digital still and video cameras.

A typical CMOS multiplexer architecture (see Fig. 32) consists of fast (column) and slow (row) shift registers at the edges of the active area, and pixels are addressed one by one through the selection of a slow register, while the fast register scans through a column, and so on. Each photodiode is connected in parallel to a storage capacitor located in the unit cell. A column of diodes and storage capacitors is selected one at a time by a digital horizontal scan register and a row bus is selected by the vertical scan register. Therefore, each pixel can be individually addressed.

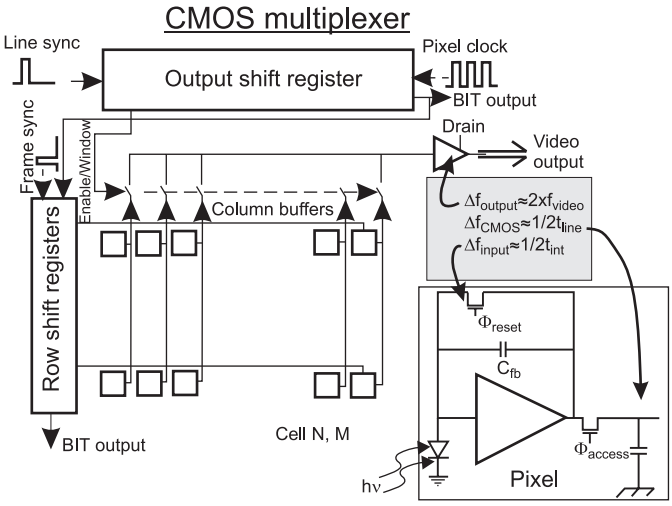

Fig. 32. CMOS multiplexing readout with CTIA detector interface. Reproduced from [25]

CMOS-based imagers use active or passive pixels [25–27] as shown, in simplified form, in Fig. 33. In comparison with passive pixel sensors (PPSs), active pixel sensors (APSs) apart from read functions exploit some form of amplification at each pixel. PPSs have simple pixels consisting of as few as two components (a photodiode and a MOSFET switch). As a result, circuit overhead is low and the optical collection efficiency [fill factor (FF)] is high even for monolithic devices. A large optical FF of up to 80% maximises signal selection and minimises fabrication cost by obviating the need for microlenses. Microlenses, typically used in CCD and CMOS APS imagers for visible application, concentrate the incoming light into the photosensitive region when they are accurately deposited over each pixel (see Fig. 34). When the FF is low and microlenses are not used, the light falling elsewhere is either lost or, in some cases, creates artifacts in the imagery by generating electrical currents in the active circuitry.

APSs incorporate transistors in each pixel to convert the photogenerated charge to a voltage, amplify the signal voltage and reduce noise. Adding these components, however, reduces the FF of monolithic imagers to about 30–50% in 0.5-μm processes at a 5–6-μm pixel pitch or in 0.25-μm processes at a 3.3–4.0-μm pixel pitch [25].

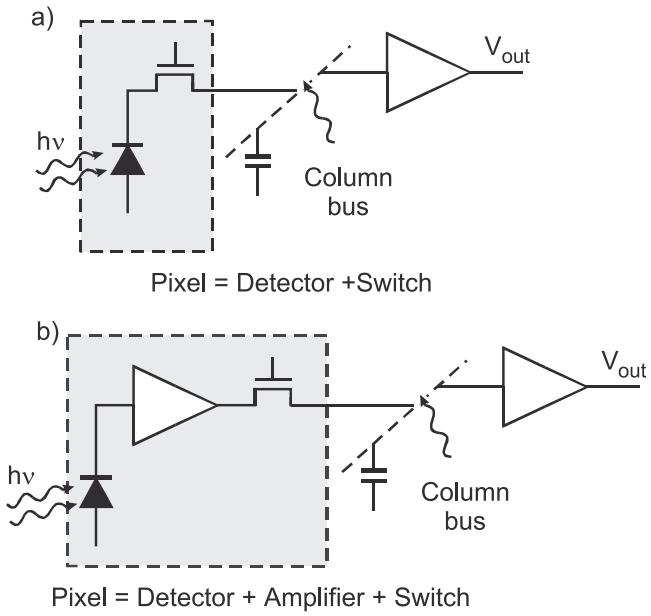

Fig. 33. Passive (a) and active (b) pixel sensor. Reproduced from [25]

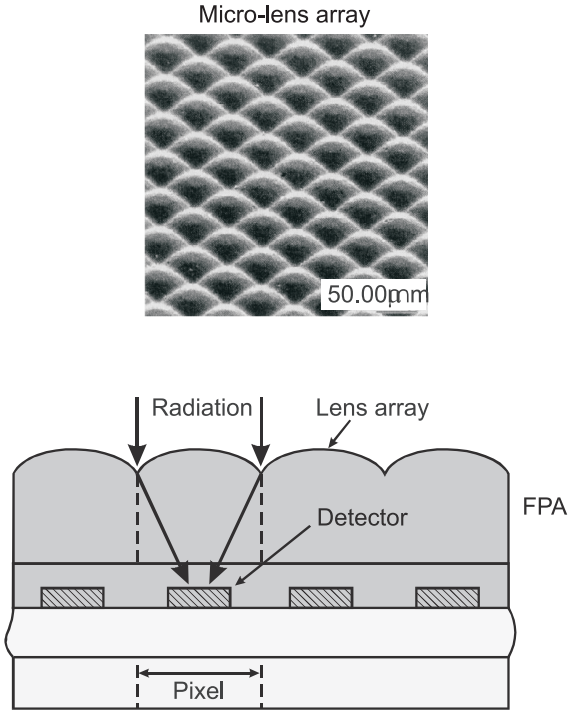

Fig. 34. Micrograph and cross-sectional drawing of microlensed hybrid FPAs. Reproduced from [25]

5.2. Hybrid arrays. Ultraviolet and infrared imagers are most commonly built with a hybrid structure. Visible hybrids have also been built for specific applications. Hybrid FPAs detectors and multiplexers are fabricated on different substrates and mated with each other by flip-chip bonding or loophole interconnection (see Fig. 35). In this case, we can optimise the detector material and multiplexer independently. Indium bump bonding of readout electronics, first demonstrated in the mid-1970s, provides for multiplexing the signals from thousands of pixels onto a few output lines, greatly simplifying the interface between the sensor and the system electronics.

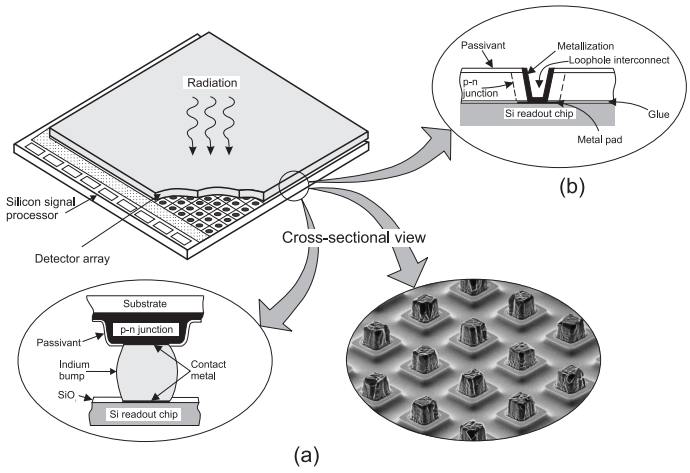

Fig. 35. Hybrid focal plane arrays with independently optimised signal detection and readout: (a) indium bump technique, (b) loophole technique

Key to the development of ROICs has been the evolution in input preamplifier technology. This evolution has been driven by increased performance requirements and silicon processing technology improvements. A brief discussion of the various circuits is given below.

The direct injection (DI) circuit was one of the first integrated readout preamplifiers and has been used as an input to CCDs and visible imagers for many years. This readout configuration requires the area at the unit cell ($< 20 \times 20$ μm) to be minimised. Photon current in DI circuits is injected, via the source of the input transistor, onto an integration capacitor (Fig. 36). As the photon current [Fig. 36(b)] charges the capacitor throughout the frame a simple charge integration takes place. Next a multiplexer reads out the final value and the capacitor voltage is reset prior to the beginning of the frame. To reduce detector noise, it is important that a uniform, near-zero-voltage bias be maintained across all the detectors.

Feedback enhanced direct injection (FEDI) is similar to direct injection except that inverting amplifier is provided between the detector node and the input MOSFET gate (Fig. 36 dashed line). The inverting gain provides feedback to yield better control over the detector bias at different photocurrent levels. It can maintain a constant detector bias at medium and high backgrounds. The amplifier reduces the input impedance of the DI and therefore increases the injection efficiency and bandwidth. The minimum operating photon flux range of the FEDI is an order of magnitude below that of the DI, thus the response is linear over a larger range than the DI circuit.

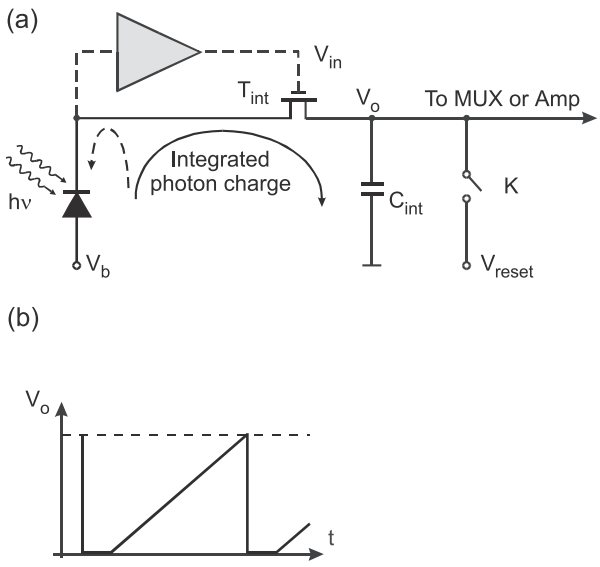

Fig. 36. Direct injection readout circuit. Reproduced from [28]

The combined source follower/detector (SFD) unit cell is shown in Fig. 37. The unit cell consists of an integration capacitance, a reset transistor ($T1$) operated as a switch, the source-follower transistor ($T2$), and selection transistor ($T3$). The integration capacitance may just be the detector capacitance and transistor $T2$ input capacitance. The integration capacitance is reset to a reference voltage (V_R) by pulsing the reset transistor. The photocurrent is then integrated on the capacitance during the integration period. The ramping input voltage of the SFD is buffered by the source follower and then multiplexed, via the $T3$ switch, to a common bus prior to the video output buffer. After the multiplexer read cycle, the input node is reset and the integration cycle begins again. The switch must have very low current leakage characteristics when in the open state, or this will add to the photocurrent signal. The dynamic range of the SFD is limited by the current voltage characteristics of the detector. As the signal is integrated, the detector bias changes with time and incident light level. The SFD has low noise for low bandwidth applications such as astronomy and still has acceptable signal-to-noise at very low backgrounds (e.g. a few photons per pixel per 100 ms). It is nonlinear at medium and high backgrounds, resulting in a limited dynamic range. The gain is set by the detector responsivity and the combined detector plus source-follower-input capacitance. The major noise sources are the kTC noise (resulting from resetting the detector), MOSFET channel thermal and MOSFET 1/f noise.

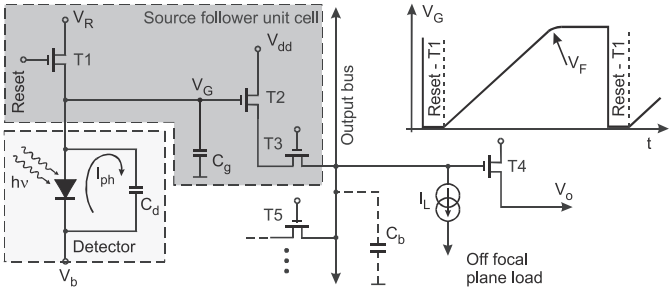

Fig. 37. Schematic of source-follower per detector unit cell

The capacitor-feedback transimpedance amplifier (CTIA) is a reset integrator and addresses broad range of detector interface and performance requirements across many applications. The CTIA consists of an inverting amplifier with a gain of A, the integration capacitance C_f placed in a feedback loop, and the reset switch K (Fig. 38). The photoelectron charge causes a slight change in a voltage at the inverting input node of an amplifier. The amplifier responds with a sharp reduction in output voltage. As the detector current accumulates over the "frame time", uniform illumination results in a linear ramp at the output. At the end of integration, the output voltage is sampled and multiplexed to the output bus. Since the input impedance of the amplifier is low, the integration capacitance can be made extremely small, yielding low noise performance. The feedback, or integration, capacitor sets the gain. The switch K is cyclically closed to achieve reset. The CTIA provides low input impedance, stable detector bias, high gain, high frequency response and a high photon current injection efficiency. It has very low noise from low to high backgrounds.

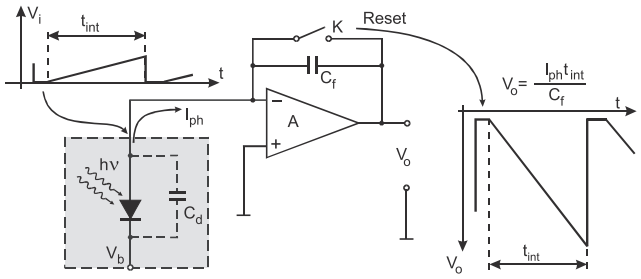

Fig. 38. Schematic of a capacitive transimpedance amplifier unit cell

The resistor load (RL) gate modulation circuit is shown in Fig. 39. It was introduced to extend the SFD performance advantages to high irradiance backgrounds and dark currents. This circuit uses the photocurrent to modulate the gate voltage and thereby induce an output current in the MOSFET. The drain current of the MOSFET transistor accumulates onto an integration capacitor. In high background irradiance, this circuit provides a design that can reject much of these background components, as when the background alone is present on the detector, the bias on the detector or the load resistor can be adjusted to give negligible drain current or integration of a charge. When the signal is then applied, the transistor drain current increases with photon current and thereby allows some level of background flux rejection. The load resistor is designed such that it has low $1/f$ noise, excellent temperature stability, and good cell to cell uniformity.

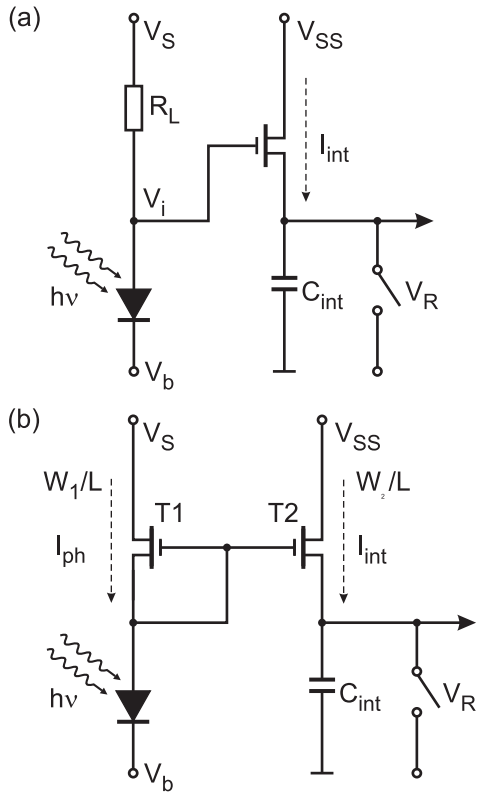

Fig. 39. Resistor load gate modulation (a) and current mirror gate modulation (b)

The current mirror (CM) gate modulation, see Fig. 39(b), extends readouts to very high background levels. In this current mirror preamplifier, the MOSFET replaces the resistor of the RL circuit. The photon current flowing into the drain of the first of two closely matched transistors includes a common gate to source voltage change in both transistors. This results in a similar current in the second transistor. If the source voltage, V_s and V_{ss}, of the two matched transistors are connected, both will have the same gate to source voltages which will induce a current in the output transistor identical to the detector current flowing through the input transistor. In this circuit, the integration current is a linear function of a detector current. This CM interfaces easily to direct access or CCD multiplexers and has low area requirements for the unit cell. The CM circuit requires gain and offset corrections for most applications. The advantages over the RL circuit include its better linearity and absence of a load resistor.

A wide variety of detector materials have been adapted to the hybrid format. UV, visible, and infrared arrays most commonly employ a photodiode structure. Photodiodes are preferred to photoconductors because of their

relatively high impedance, which matches directly into the high input impedance stage of an FET readout circuit and also allows lower power dissipation. Mesa photodiodes are used in AlGaN, InSb, and HgCdTe detectors, whereas planar photodiodes are used in Si, PtSi, Ge, HgCdTe, InGaAs, and InSb detectors. A third photodiode structure — used exclusively with HgCdTe detectors – is the high-density vertically-integrated photodiode, or loophole photodiode.

An alternative hybrid detector for the long wavelength IR region (8–14-μm) is the quantum well infrared photoconductor (QWIP). These high impedance detectors are built from alternating thin layers (superlattices) of GaAs and AlGaAs. A distinct feature of n-type QWIPs is that the optical absorption strength is proportional to the electric-field polarization component of an incident photon in a direction normal to the plane of the quantum wells. For imaging, it is necessary to couple light uniformly to 2-D arrays of these detectors, so a diffraction grating is incorporated on one side of the detectors to redirect a normally incident photon into propagation angles more favourable for absorption.

Also extrinsic silicon detectors can form high impedance photoconductors in a hybrid configuration and be operated out to about 30 μm. Shallow, hydrogen-like impurities, such as phosphorus, antimony, or arsenic, provide electrons which can be ionised with photon energies in the range of 30–50 meV, depending upon the dopant and concentration used. Arrays of QWIP as well as extrinsic silicon detectors in a 1024×1024 format have been demonstrated [20].

The largest hybrid arrays have been principally built for astronomy where dark currents as low as 0.02 electrons/sec are measured at 30 K. The most recent development is the 2×2 K format of InSb (1–5 μm, 25 μm pixels) and HgCdTe (1–3 μm, 18 μm pixels). Individual arrays can be arranged in groups to give larger format configurations. Two-side buttable 2052×2052 arrays can be arranged in a 4104×4104 format, as illustrated in Fig. 40. The three-side buttable arrays can be arranged in $2 \times n$ array configurations to build up large rectangular formats [29].

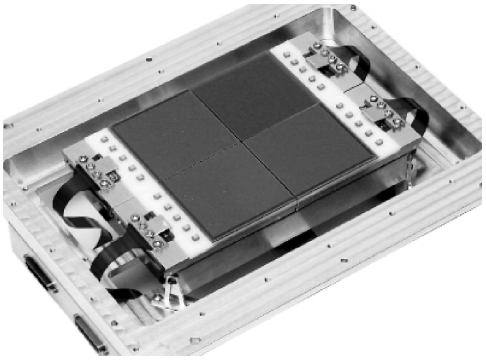

Fig. 40. Four two-side buttable 2052×2052 InSb arrays in a mosaic configuration 4104×4104 arrays pixels. Reproduced from [29]

In the infrared spectral region, third generation systems are now being developed. In this class of detector, two main competitors, HgCdTe photodiodes and QWIPs are considered. Third generation IR systems considered to be those that provide enhanced capabilities like larger number of pixels, higher frame rates, better thermal resolution as well as multicolour functionality and other on-chip functions. Multicolour capabilities are highly desirable for advance IR systems. Systems that gather data in separate IR spectral bands can discriminate both absolute temperature and unique signatures of objects in the scene. By providing this new dimension of contrast, multiband detection also offers advanced colour processing algorithms to further improve sensitivity compared to that of single-colour devices.

Two-colour array capability is based upon stacking materials with different spectral responses on top of each other. The shorter wavelength flux is absorbed in the first layer which then transmits the longer wavelength flux through the second layer. One such structure — a HgCdTe two-colour device, with two indium bumps per pixel — is illustrated in Fig. 41.

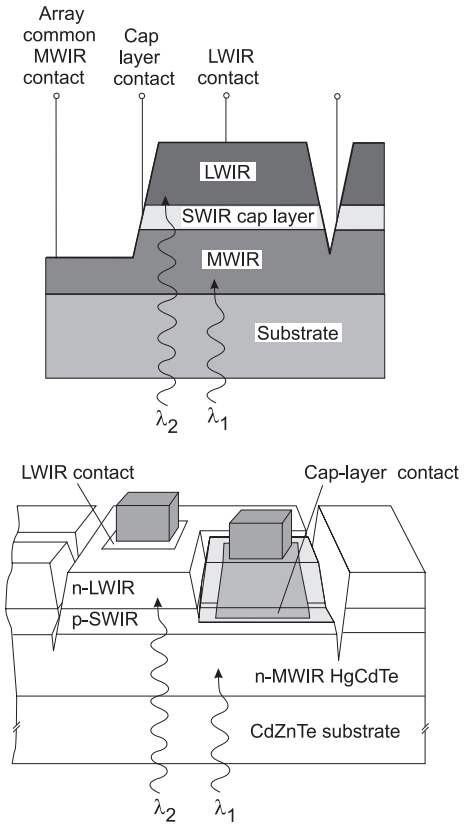

Fig. 41. Cross section of integrated two-colour HgCdTe detectors in an n-p-n layer structure for simultaneous operating mode. A thin p-type barrier separates the two absorbing bands

Two-colour QWIP detector structures have also been built [30]. For example, Fig. 42 shows the excellent imagery in each colour. Note the appearance of the front-held

optical filter and the vertically-held hot soldering iron in the two bands. At present, however, imaging systems using two-colour arrays are in limited use. Some considerations have suggested that three-colour FPAs would be more generally useful. Recently, a four-colour QWIP FPA has been demonstrated by stacking different multi-quantum well structures, which are sensitive in 4–5.5, 8.5–10, 10–12, and 13–15.5 μm bands. The 640×512 format FPA consists of four 640×128 pixel areas which are capable of acquiring images in these bands.

Fig. 42. Simultaneous images from 256×256 MWIR/LWIR QWIP FPAs. Note appearance of the front-held filter and the hot soldering iron in the two bands. Reproduced from [31]

It should be mentioned that hyperspectral arrays are distinguished from multispectral ones in typically having a hundred or more bands. HgCdTe, and other detector materials such as silicon and InSb, have been used in hyperspectral assemblies in the form of 2-D arrays with a closely-packed layout of rows and columns. A prism, grating, or a "wedged" filter is used to illuminate each row with a different wavelength. Figure 43 shows an example of a hyperspectral array.

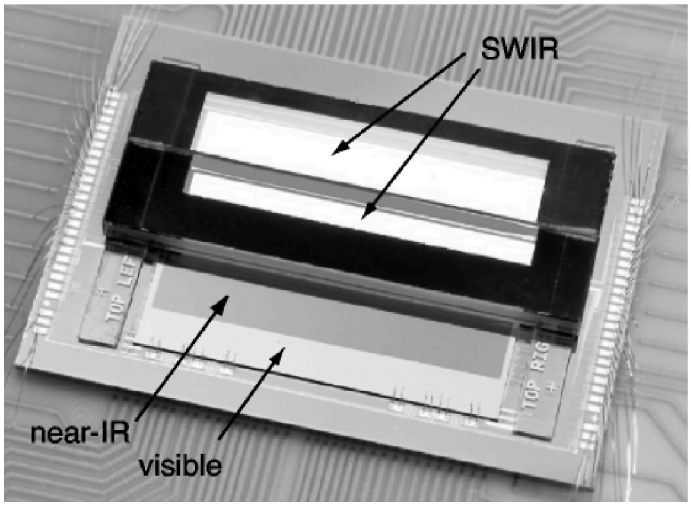

Fig. 43. Hyperspectral array with 300 bands made from silicon and HgCdTe together with four wedge filters. The array has 100 spectral bands in the range from 0.4 to 1.0 μm, and 200 from 1.0 to 2.5 μm. Reproduced from [20]

6. Conclusions

The paper provides an overview of the important techniques for detection of optical radiation from the ultraviolet, through visible to infrared spectral regions. At the beginning single-point devices are considered, next direct detector systems and advanced techniques including coherent detection, and finally image counterparts containing focal plane arrays. The reader should gain an understanding of the similarities and contrasts, the strengths and weaknesses of the great number of approaches that have been developed over a century of effort to improve our ability to sense photons. The emphasis is always upon the methods of operation and limits of different techniques. Also currently achieved performance levels are briefly described.

This paper offers a rather wide coverage of detection techniques. However, for a full understanding of the technical content, the paper requires as a prerequisite the basic courses in electronic devices and circuits, and the very fundamentals of semiconductors and noise.

References

[1] S. Donati, *Photodetectors. Devices, Circuits, and Applications*, New Jersey: Prentice Hall Inc., 1999.

[2] C. D. Motchenbacher and J. A. Connelly, *Low-noise Electronic System Design*, New York: John Wiley & Sons, Inc., 1995.

[3] J. L. Vampola, "Readout electronics for infrared sensors", in *The Infrared and Electro-Optical Systems Handbook*, Vol. 3, ed. by W. D. Rogatto, Bellingham: SPIE Press, 1993.

[4] Z. Bielecki, W. Kołosowski, R. Dufrene and M. Borejko, *Proceedings of 11th European Gallium Arsenide & other Compound Semiconductors Application Symposium*, Munich, 137–140 (2003).

[5] G. H. Rieke, *Detection of Light: From Ultraviolet to the Submillimeter*, Cambridge: Cambridge University Press, 1994.

[6] *Photomultiplier Tubes — Basic and Applications*, Editorial Committee, Hammamatsu Corporation, 1999.

[7] H. Kume, "Optical detectors and receivers," in *Handbook of Optoelectronics*, Bristol: IOP, 2004.

[8] Z. Bielecki, "Maximisation of signal to noise ratio in infrared radiation receivers," *Opto-Electron. Rev.* 10, 209–216 (2002).

[9] M. I. Skolnik, *Radar Handbook*, New York: Mc. Graw-Hill, 1990.

[10] *RCA Electro-Optics Handbook*, New Jersey: RCA, 1974.

[11] G. Keiser, *Optical Fiber Communications*, New York: McGraw-Hill, 1979.

[12] J. R. Barry and E. A. Lee, "Performance of coherent optical receivers", *Proc. IEEE* 78, 1369–1394 (1990).

[13] R. Brun, "Gallium arsenide eyesafe laser rangefinder," *Proc. SPIE* 1207, 172–181 (1990).

[14] R. J. Keyes, *Optical and Infrared Detectors*, Berlin: Springer, 1977.

[15] I. P. Csorba, *Image Tubes*, Howard W. Sams & Co., 1985.

[16] Z. Bielecki and A. Rogalski, *Detection of Optical Radiation*, Warsaw: WNT, 2001 (in Polish).

[17] A. Rogalski, *Infrared Detectors*, Amsterdam: Gordon and Breach Science Publishers, 2000.

[18] A. Rogalski, "Infrared detectors: Status and trends," *Progress in Quantum Electronics* 27, 59–210 (2003).

[19] A. Rogalski, "Photon detectors," in *Encyclopedia of Optical Engineering*, 1985–2036, ed. R. Driggers, New York: Marcel Dekker Inc., 2003.

[20] P. Norton, "Detector focal plane array technology," in *Encyclopedia of Optical Engineering*, 320–348, ed. R Driggers, New York: Marcel Dekker Inc., 2003.

[21] http://www.dalsa.com/

[22] J. Janesick, *Scientific Charge-Coupled Devices*, Bellingham: SPIE Press, 2001.

[23] L. J. Kozlowski, J. Montroy, K. Vural and W. E. Kleinhans, "Ultra-low noise infrared focal plane array status," *Proc. SPIE* 3436, 162–171 (1998).

[24] E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, New York: Wiley & Sons, 1984.

[25] L. J. Kozlowski, K. Vural, J. Luo, A. Tomasini, T. Liu and W. E. Kleinhans, *Opto-Electron. Rev.* 7, 259–269 (1999).

[26] E. R. Fossum, "Active pixel sensors: Are CCD's dinosaurs?," *Proc. SPIE* 1900, 2–14 (1993).

[27] E. R. Fossum and B. Pain, "Infrared readout electronics for space science sensors: State of the art and future directions", *Proc SPIE* 2020, 262–285 (1993).

[28] M. J. Hewitt, J. L. Vampola, S. H. Black and C. J. Nielsen, "Infrared readout electronics: a historical perspective", *Proc. SPIE* 2226, 108–119 (1994).

[29] P. J. Love, K. J. Ando, R. E. Bornfreund, E. Corrales, R. E. Mills, J. R. Cripe, N. A. Lum, J. P. Rosbeck and M. S. Smith, "Large-format infrared arrays for future space and ground-based astronomy applications", *Proc SPIE* 4486, 373–384 (2002).

[30] A. Rogalski, "Quantum well photoconductors in infrared detector technology", *J. Appl. Phys.* 93, 4355–4391 (2003).

[31] T. Whitaker, "Sanders' QWIPs detect two color at once," *Compound Semiconductors* 5(7), 48–51 (1999).