Photonics - Riassuntini Valentini

Giorgio Calandra

Politecnico di Milano - Physics Engineering - A.A 2023-2024

The following summary has been made by Giorgio Calandra in preparation for the exam of Photonics for what concerns Prof. Valentini's part. This summary does not want to be complete or clear and mostly relies on the topic frequently asked during the exam. I want to thank Giulia Campagna who passed me a hand-made version of this summary from which I have taken most of the structure and contents. May the force be with all of us!

1 Geomtrical Optics

A lens is made of two spherical surfaces close to each other delimiting a medium with a refractive index different from the one outside form a spherical lens. Assuming that the external medium on both sides of the lens is air ($n = 1$), a thin lens is characterised by a focal length f, which is given by the *Lensmaker's formula*:

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{1}$$

where R_1 and R_2 are the curvature radii of the lens and n is the refractive index of the material the lens is made of. For any optical system holds:

$$\frac{1}{f} = \frac{1}{s_1} + \frac{1}{s_2} \tag{2}$$

where s_1 and s_2 are defined in Fig.1. The lateral magnification M_L of an object is;

$$M_L = -\frac{|h_2|}{|h_1|} \tag{3}$$

If object and image are in media with different refractive indices the general expression of the lens law becomes:

$$\frac{n_1}{s_1} + \frac{n_2}{s_2} = \frac{n_1}{f_1} = \frac{n_2}{f_2} \tag{4}$$

where f_1 and f_2 are the front and back focal lengths, respectively. For an optical system, some important features which limit the amount of light emitted by the object and directed towards the image are:

- **Aperture stop:** limits the cone angle of the ray bundle from the object to the image
- **Field stop:** limits the field of view in the object plane
- **Entrance pupil:** is the image of the aperture stop in the object space

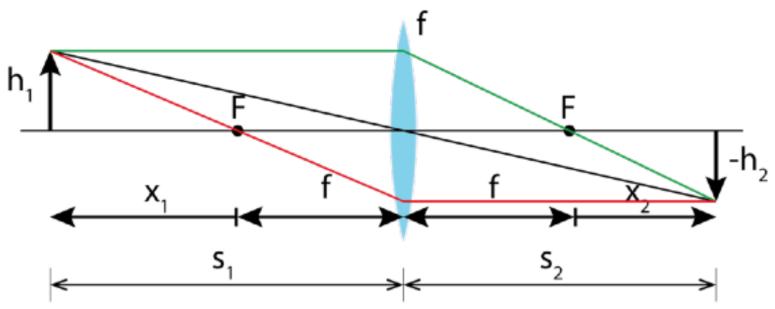

Figure 1: Image of an object made by a thin lens.

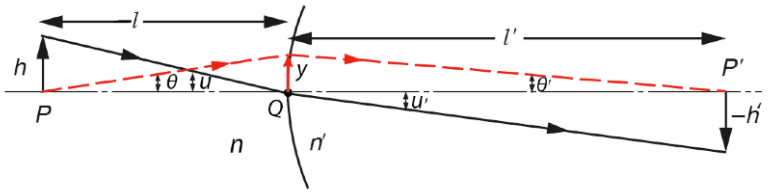

Figure 2: Lagrange Theorem.

- **Exit pupil:** is the image of the aperture stop in the image space
- **Entrance window:** is the image of the field stop in the object space
- **Exit window:** is the image of the field stop in the image space

In *paraxial approximation* optical rays form a small angle with the optical axis so that the sine of the angle can be approximated with the angle itself. Images are assumed to be exact replicas of objects since aberrations are negligible. In these conditions, the **Lagrange Theorem** holds:

$$nh\theta = n'h'\theta' \tag{5}$$

where all the quantities are defined in Fig.2. The product $nh\theta$ is called the Lagrange invariant. It allows us to write the magnification as:

$$|M_L| = \frac{h'}{h} = \frac{n\theta}{n'\theta'} \tag{6}$$

An optical system is said to be stigmatic for two axial points P and P' if P' is a perfect replica of P. Images made by a stigmatic system might still be affected by coma. This aberration affects off-axis points and is named from the shape of the blur occurring at such points, which resembles a comet tail. For a stigmatic system to be free of coma the **Abbe's sine condition** must be satisfied:

$$nh\sin\theta = n'h'\sin\theta' \tag{7}$$

2 Radiometric Quantities

The most basic quantity is the **Radiant Flux** Φ measured in Watt $[W]$ which is the power carried by a radiation field. All the other radiometric quantities are derived from this one:

- **Radiant Exitance** $M(x, y)$ is the emitted radiant flux per unit area $[W/m^2]$:

$$M(x, y) = \frac{d\Phi}{dA} \tag{8}$$

- **Irradiance** $E(x, y)$ is the recived radiant flux per unit area $[W/m^2]$:

$$E(x, y) = \frac{d\Phi}{dA} \tag{9}$$

- **Radiant Intensity** $I(\theta, \varphi)$ is the emitted radiant flux per unit of solid angle $[W/sr]$

$$I(\theta, \varphi) = \frac{d\Phi}{d\Omega} \tag{10}$$

- **Radiance** $L(x, y, \theta, \varphi)$ is the emitted/received radiant flux per unit of solid angle and projected area $[W/(m^2 \cdot sr)]$:

$$L(x, y, \theta, \varphi) = \frac{d\Phi}{dA_{proj}d\Omega} = \frac{d\Phi}{dAd\Omega\cos\theta} \tag{11}$$

  where θ is the angle between the normal to the surface and the direction of sight, as depicted in Fig.3(d).

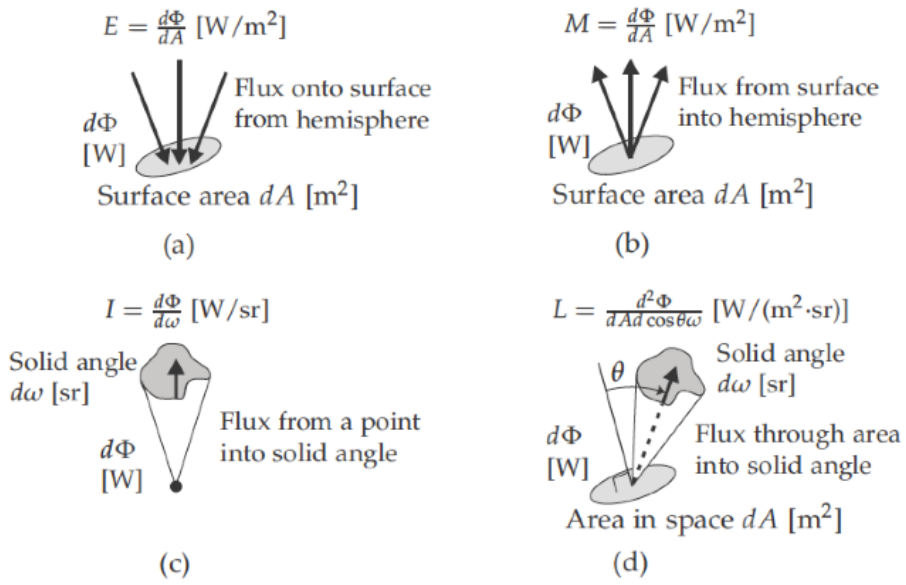

Figure 3: Radiometric Quantities.

2.0.1 Conservation of the Radiance

The radiance is one of the most important quantities of radiometry because it is conserved in the free space propagation. Let's assume that a radiation field propagates in a homogeneous, non-absorbing medium, from a source to a receiver. The demonstration has to be known (asked at the exam) but it is not reported here. The result is:

$$L_0 = L_1 \tag{12}$$

where L_0 and L_1 are the radiance of the source and the receiver respectively. If we are not in the free space propagation anymore but we are at the interface between two media with refraction indices n_1 and n_2 the radiance is not a conserved quantity. However, the so-called **Basic Radiance** L/n^2 is conserved:

$$\frac{L_1}{n_1^2} = \frac{L_2}{n_2^2} \tag{13}$$

For a generic sequence of media with refraction indices $n_1, n_2, \ldots, n_n$ it holds:

$$\frac{L_1}{n_1^2} = \frac{L_2}{n_2^2} = \cdots = \frac{L_n}{n_n^2} \tag{14}$$

2.0.2 Radiance on an image

The operating principle of electronic imaging devices is always the same: an optical system forms a replica of the object into a plane where a pixelated detector converts the photon flux into electron bunches that form the image. For these devices, the most important quantity to be considered in the final stage of image recording is the irradiance $E(x, y)$ in the detector plane. Let's consider the image of a small element dA_0 of an object, made by an aplanatic, i.e. aberration-free, optical system. The demonstration has to be known (asked at the exam) but it is not reported here. The result is:

$$\frac{L_1(\theta_1, \varphi)}{n_1^2} = \frac{L_0((\theta_0, \varphi))}{n_0^2} \tag{15}$$

which means that the basic radiance L/n^2 of the image is equal to that of the object.

3 Lambertian Sources

All light sources, either primary or secondary, that appear alike under any observation angle are called **Lambertian** and share the same important property: their radiance does not depend on the angles θ and φ:

$$L(x, y, \theta, \varphi) = L(x, y) \tag{16}$$

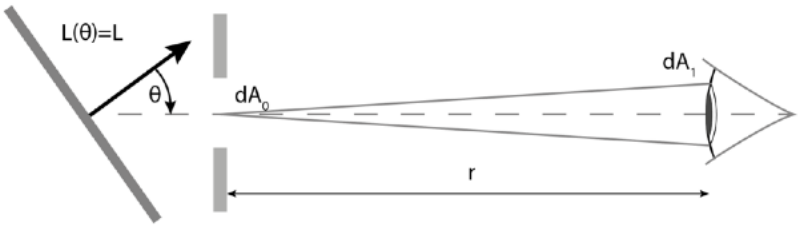

Figure 4: Independence on the angle for the radiance of a Lambertian source.

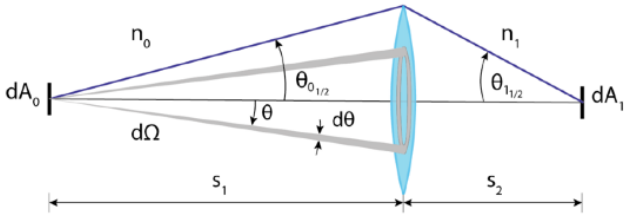

Figure 5: Irradiance of an image.

The fact that the radiance does not depend on θ and φ comes from the independence of the radiant flux Φ from θ and φ. Since by definition $L(\theta,\varphi) = d\Phi/dA_{proj}d\Omega$, the element of flux $d\Phi$ emitted by the source and received by the detector is:

$$d\Phi = L(\theta,\varphi)dA_{proj}d\Omega = L(\theta,\varphi)dA_{proj}\frac{dA_1}{r^2} \tag{17}$$

where dA_1 is the area of the detector and r is its distance from the source. By referring to Fig.4 we notice that the projection of the area of the source is always equal to dA_0 regardless of the value of θ. Therefore we have:

$$d\Phi = L(\theta,\varphi)dA_0\frac{dA_1}{r^2} \tag{18}$$

Therefore, being Φ constant over θ and φ the radiance L will be independent of these angles as well. Moreover, the radiant intensity I obeys the Lambert's cosine law:

$$I(\theta) = LA_0\cos\theta = I_0\cos\theta \tag{19}$$

The most important relation to keep in mind is:

$$\Phi = MA = \pi LA = \pi I_0 \tag{20}$$

where $M = \pi L$ is the radiance excitance $M = d\Phi/dA$.

3.0.1 Irradiance of the image of a Lambertian source

Let's consider a small element dA_0 at the source, a lens, which may represent a generic optical system, and a detector element of area dA_1 in the conjugate plan (see Fig.5). The irradiance in dA_1, i.e. at the detector, is equal to:

$$E = \pi L_0\frac{n_1^2}{n_0^2}\sin^2\theta_{1_{1/2}} \tag{21}$$

where the angle $\theta_{1_{1/2}}$ can be retrieved geometrically by using the Lens law. In the case in which s_1 is much larger than f and therefore $s_2 = f$ we get:

$$\tan\theta_{1_{1/2}} = \frac{D/2}{s_2} = \frac{D/2}{f} \Rightarrow f_\# = \frac{f}{D} = \frac{1}{2\tan\theta_{1_{1/2}}} \tag{22}$$

where D is the entrance pupil, i.e. the diameter of the lens, and $f_\#$ is called the Focal Ratio or f-number.

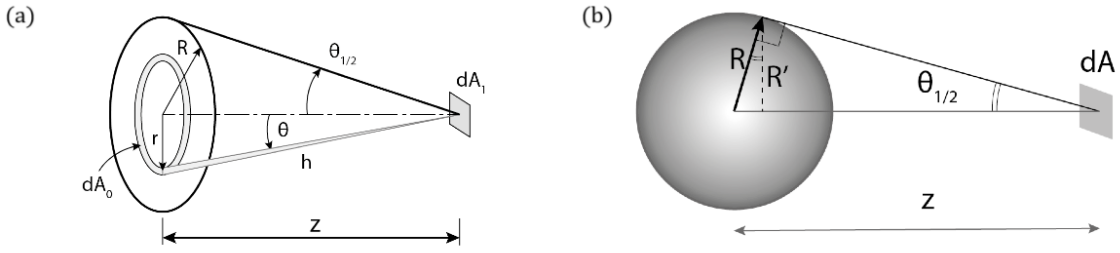

Figure 6: Lambertian disk (a) and sphere (b) with uniform radiances.

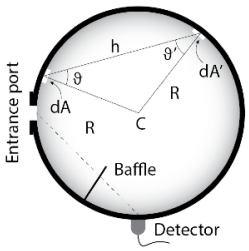

Figure 7: Schematic of an Integrating Sphere.

3.1 Lambertian Sources

Since most of the light sources we are likely to deal with in the real world are Lambertian, it is important to determine the irradiance produced by a Lambertian source on a distant object.

3.1.1 Lambertian Disk

Let's consider a disk of radius R and uniform radiance L and a detector placed at a distance z from the axis of the disk (see Fig.6)(a). It is possible to demonstrate:

$$E = \frac{d\Phi}{dA_1} = \pi L \sin^2 \theta_{1/2} = \pi L \frac{R^2}{R^2 + z^2} \tag{23}$$

3.1.2 Lambertian Sphere

Let's consider a sphere of radius R and uniform radiance L and a detector placed at a distance z from the axis of the disk (see Fig.6(b)). It is possible to demonstrate:

$$\Phi = MA = M4\pi R^2 = 4\pi^2 R^2 L \tag{24}$$

$$E = \frac{\pi R^2 L}{z^2} \tag{25}$$

$$I = \frac{\Phi}{4\pi} = \pi R^2 L \tag{26}$$

In this case we have $\sin \theta_{1/2} = R/z$.

3.2 Integrating Sphere

To face the common problem of measuring the total flux of an incoherent beam the best option consists in using an integrating sphere. This is a hollow sphere with a small aperture (indeed much smaller than its diameter) that transforms the flux entering the sphere from any direction into a uniform irradiance all over its internal surface, through multiple reflections. Integrating spheres have their internal surface covered with non-absorbing materials with a very high diffuse reflectance coefficient r. It is possible to demonstrate that the irradiance on any portion of the sphere (and so even on a detector placed inside the sphere is):

$$E = \frac{\Phi}{4\pi R^2} \frac{r}{1-r} \tag{27}$$

where R is the radius of the sphere.

4 Photodetectros

4.1 Noise in Photodetectors

Noise is an unwanted disturbance in an electrical signal. Noise is dramatically important in photodetection since the ultimate sensitivity of any detector is set by noise figures of merit. The description of noise in the time domain can be done by considering its statistical properties and can be expressed either through the current or the voltage root mean square. There are two fundamental sources of noise in photodetection:

- **Thermal Noise**: it is due to the random fluctuation of free charges. It is the electronic noise generated by the thermal agitation of the charge carriers inside an electrical conductor, i.e. a resistor element. Since a load resistor R_L must be always present in the detection circuit, this source of noise is unavailable. It is possible to demonstrate:

$$i_{th}^{RMS} = \sqrt{\frac{4k_B T}{R_L}\Delta f} \tag{28}$$

  where Δf is the bandwidth of the detection circuit.

- **Shot Noise**: it is caused by the quantisation of the charge and cannot be avoided due to its fundamental origin. In general, shot noise arises from the random occurrence of discrete events. In this case, it is possible to demonstrate:

$$i_{shot}^{RMS} = \sqrt{2qI\Delta f} \tag{29}$$

  where I is the mean value of the current flowing in the circuit.

4.2 Figures of Merit of Photodetectors

Photodetectors are used to measure the energy carried by an electromagnetic field. There are three parameters to describe the performances of a photodetector:

- **Responsivity**: is the ratio between the output signal and the input optical power. In general, it depends on the wavelength of the impinging radiation and it can be equivalently expressed as:

$$R_I(\lambda) = \frac{I_{out}[A]}{P_{in}[W]} \qquad R_V(\lambda) = \frac{V_{out}[V]}{P_{in}[W]} \tag{30}$$

- **Frequency Response**: is the capability of a photodetector to correctly recover the time profile of the optical signal. In many cases, this can be represented by the means of a cutoff frequency f_c which is the frequency at which the output amplitude of a sinusoidally modulated optical signal reduces by a factor $\sqrt{2}$, with respect to a CW signal of equal average power.

- **Noise Equivalent Power (NEP)**: It is the optical power in a sinusoidal regime that would produce a modulated signal in a noise-free detector equal to the root mean square (r.m.s.) noise measured from the real detector without signal. In fact, the lowest optical power that can be detected is the one that produces an output signal comparable to the noise, so that the *Signal to Noise Ratio* (SNR or S/N) is equal to 1. In general, the r.m.s. of the noise amplitude is proportional to $\sqrt{A\Delta f}$, where A is the sensitive area of the detector and Δf the passband:

$$N_{rms} \propto \sqrt{A\Delta f} \tag{31}$$

  But since the NEP is given by the input power P required to give $S/N = 1$, and being the responsivity defined as $R_\lambda = S_{out}/P_{in} \Rightarrow P_{in} = S_{out}/R_\lambda$ we get:

$$NEP_\lambda = P_{\left(\frac{S}{N}=1\right)} = \frac{S_{rms}}{R_\lambda} = \frac{N_{rms}}{R_\lambda} \propto \sqrt{A\Delta f} \tag{32}$$

  We can define another parameter to make the NEP independent from the detector size and bandwidth:

$$NEP_\lambda^* = \frac{NEP_\lambda}{\sqrt{A\Delta f}} \tag{33}$$

  The last quantity to define is the *Detectivity* D which is the inverse of the NEP:

$$D_\lambda = \frac{1}{NEP_\lambda} \qquad D_\lambda^* = \frac{1}{NEP_\lambda^*} \tag{34}$$

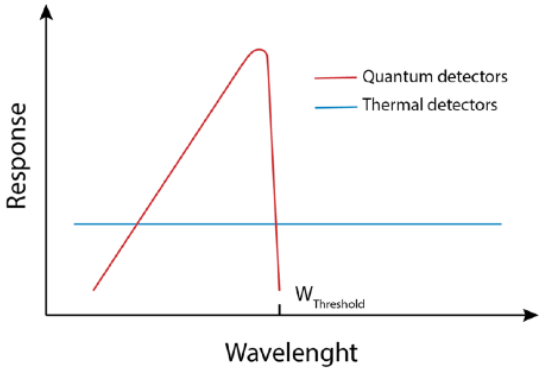

Figure 8: Typical output signal from Quantum and Thermal detectors

4.3 General Characteristic of Photodetectors

Radiation detectors can be roughly divided into two categories:

- **Thermal detectors**: the energy of the absorbed photons is converted into heat, which increases the temperature of any material. This temperature increase can result in a measurable physical effect like the change in electric resistance (e.g. bolometers) or the generation of an output voltage (e.g. thermoelectric detectors), which is proportional to the optical power impinging on the detector.
- **Quantum detectors**: the interaction of photons with an active element induces one-to-one events (e.g. emission of electrons, electron-hole pair creation) that can be counted by an appropriate circuitry that provides an output signal proportional to the photon flux.

Tab.1 compares some of the main features of these two types of detectors. What is told in the first

Table 1: Comparison between Thermal and Quantum Detectors

	Thermal Detectors	Quantum Detectors
Response	Practically flat over a wide spectral range	Highly frequency dependent, threshold frequency
Measurement	Simple and direct	Non-trivial because of the frequency dependence
Sensitivity	Lower than Quantum Detectors	Major strength, can reach single photons

line of the table is also understandable by looking at Fig.8

4.4 Thermal Detectors

These detectors are always made by an absorbing element with thermal capacitance C_{th} (that should be as small as possible), which is connected, through a thermal link with heat conductance G_{th}, to a massive heat sink with thermal capacitance $C_{HS} >> C_{th}$ which is kept at the equilibrium temperature T_{eq}. Light absorption delivers a power W into the detector which leads to an increase of temperature ΔT with respect to T_{eq}. The energy balance is:

$$W dt - G_{th} \Delta T dt = C_{th} d(\Delta T) \Rightarrow \frac{d(\Delta T)}{dt} = \frac{W}{C_{th}} - \frac{G_{th}}{C_{th}} \Delta T \tag{35}$$

The solution of this equation depends on the type of optical signal impinging onto the detector:

- Constant optical signal $W(t) = W_0$:

$$\Delta T(t) = \frac{W_0}{G_{th}} \left[1 - \exp\left(-\frac{t}{\tau_{th}} \right) \right], \quad \tau_{th} = \frac{C_{th}}{G_{th}} \tag{36}$$

- Sinusoidal optical signal $W(t) = W_0 + W_m \cos(2\pi f t)$:

$$\Delta T(t) = \frac{W_0}{G_{th}} + \frac{W_m \cos(2\pi f t)}{G_{th} \sqrt{1 + 4\pi^2 f^2 \tau_{th}^2}}, \quad \tau_{th} = \frac{C_{th}}{G_{th}} \tag{37}$$

The responsivity is proprortional to the temperature increase ΔT which in turns is proportional to $1/G_{th}$:

$$R_\lambda \propto \Delta T \propto \frac{1}{G_{th}} \tag{38}$$

Unfortunately, a decrease in G_{th} would also increase the time constant $\tau_{th} = C_{th}/G_{th}$ resulting in a worse temporal response. In general, the following symbolic equation holds:

$$Gain \times Bandwidth = cost \tag{39}$$

Let's now list some of the most relevant thermal detectors.

4.4.1 Thermoelectric Detectors

In Thermoelectric detectors, the absorbing element is made by a metallic surface (sometimes covered with a broadband absorbing layer) or a thick glass disk in the case in which a stronger absorber is needed (e.g. to resist a powerful laser pulse). The heat sink is made of a metallic body with cooling features (air convention or water flow cooling depending on the power). The temperature difference between the active element and heat sink is measured by means of *thermocouples*. Thermocouples are made by two joints between metallic wires made of different metals and exploit the Seebek effect to convert a temperature gradient into a voltage:

$$\nabla V = -Q \nabla T \tag{40}$$

where Q is called thermopower or Seebeck coefficient. It is possible to demonstrate that:

$$Q = -\frac{c_v}{3ne} = -\frac{\pi^2}{6}\frac{k_B}{e}\left(\frac{k_B T}{\epsilon_F}\right) \tag{41}$$

where c_v is the specific heat of the metal, n the electron density, e the electron charge, k_B the Boltzmann cosntat and ϵ_F the Fermi energy of the metal. Considering a thermocouple made of metals 1 and 2 with thermopowers Q_1 and Q_2, respectively, the voltage difference across the terminals of the thermocouple is given by:

$$\Delta V = (Q_1 - Q_2)\Delta T \tag{42}$$

with $\Delta T = T_1 - T_2$. Usually, the signal measured by a thermocouple is quite weak and N thermocouples can be placed in series to increase the signal. In this case, we have:

$$\Delta V = N(Q_1 - Q_2)\Delta T \tag{43}$$

4.4.2 Bolometers

Bolometers are based on a thin slab of material whose electric resistance is highly sensitive to temperature change. Around the working point of the detector, the resistance R of a bolometer can be assumed a linear function of the temperature T:

$$R = R_0\left[1 + \alpha(T - T_0)\right], \quad \alpha(T) = \frac{1}{\rho}\frac{d\rho}{dT} \tag{44}$$

where α is the temperature coefficient and ρ the resistivity of the material. Semiconductors are often used since they have a very large and negative temperature coefficient. These detectors are particularly important for measuring far infrared and THz waves.

4.4.3 Photoacustic Detectors (Golay Cell)

The Golay cell is a pneumatic detector, i.e. it relies on pressure changes. The radiation absorber is made of a metal film which transfers the heat to a small volume of gas in a sealed chamber. Because of the heat transfer the pressure in the chamber increases and leads to a deformation of a flexible mirror placed at the back of the chamber. An optical system is used to detect the deformation of the mirror. An optical beam is focused on the flexible mirror and the reflected beam is collected and focused onto a four-quadrant photodiode. The unbalance in the signal current exiting the quadrants of the photodiode allows for measuring the deformation of the mirror and, in turn, the power released inside the cell. Ideally, the absorbing metallic film should be the only element experiencing the heat exchange, all the other elements (gas, pneumatic chamber, etc.) should be perfect heat insulators.

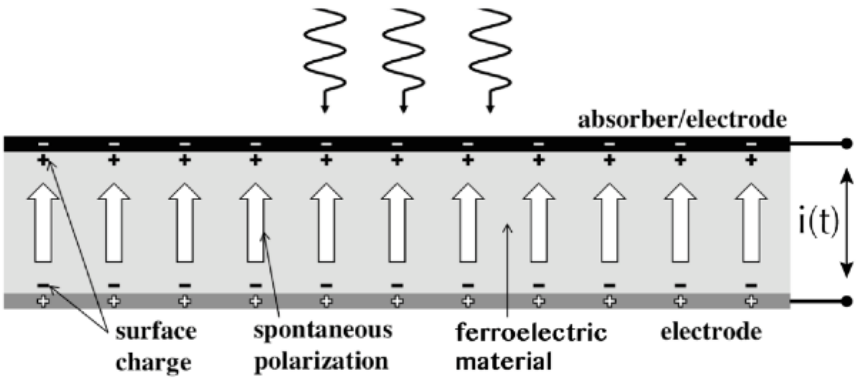

Figure 9: Schematic of a pyroelectric detector.

4.5 Pyroelectric Detectors

Pyroelectric detectors rely on a crystal that shows the ferroelectric effect which always also exhibits the pyroelectric effect, i.e. all ferroelectric materials are pyroelectric (and not vice versa). In these crystals, there is an electric polarisation which changes with temperature even when no field is applied to the system. This is due to an unbalance in the centre of the positive and negative charge distribution in the crystal. The fact that ferroelectric crystals undergo a variation in spontaneous polarisation when the temperature is changed can be used to make a light detector. The schematic representation of a pyroelectric detector is shown in Fig.9. A thin slab of ferroelectric crystal suitably poled and oriented is sandwiched between electrodes. At thermal equilibrium, the spontaneous polarisation vector $\mathbf{P}$ causes the presence of a bipolar polarisation charge of density $\pm\sigma_P$. To grant charge neutralities at equilibrium, the charge densities $\pm\sigma_P$ must be compensated by equal and opposite densities of free charge $\pm\sigma$ on the conductive electrodes. This makes the detector equivalent to a capacitor that stores a charge Q:

$$Q = A\sigma = A\sigma_P, \qquad \sigma_P = \mathbf{P}\cdot\mathbf{u}_n \tag{45}$$

where A is the area of the pyroelectric slab. The increase in temperature induces a variation of the polarisation charge that in turn affects the free charge Q. This results in a current $i(t)$ given by the time derivative of Q:

$$i(t) = \frac{dQ(t)}{dt} = A\frac{d\sigma_P}{dt} = A\frac{d\sigma}{dt} = A\frac{d\sigma}{dT}\frac{dT}{dt} \equiv pA\frac{dT(t)}{dt} \tag{46}$$

where $p = d\left|\mathbf{P}\right|/dT$ is the pyroelectric coefficient. If the detector receives a sinusoidally modulated optical power $P(t) = W_0 + W_m\cos 2\pi ft$ we have from Eq.(37):

$$\frac{dT(t)}{dt} = \frac{d\Delta T(t)}{dt} = W_m\cos\left(2\pi ft\right)\frac{2\pi f}{G_{th}\sqrt{1+4\pi^2 f^2\tau_{th}^2}}, \qquad \tau_{th} = \frac{C_{th}}{G_{th}} \tag{47}$$

where C_{th} is the heat capacitance of the pyroelectric material. Then, one can compute the responsivity for a current output and a voltage output:

$$R_I^{RMS} = \frac{i^{RMS}}{P_{in}^{RMS}} = pA\frac{2\pi f}{G_{th}\sqrt{1+4\pi^2 f^2\tau_{th}^2}} \tag{48}$$

$$R_V^{RMS} = \frac{v^{RMS}}{P_{in}^{RMS}} = 2\pi f\frac{pA}{C_{th}C_e}\frac{\tau_{th}}{\sqrt{1+4\pi^2 f^2\tau_{th}^2}}\frac{\tau_{RC}}{\sqrt{1+4\pi^2 f^2\tau_{RC}^2}} \tag{49}$$

where, in the voltage case, it was necessary to include in the measuring circuit a load resistor R_L ($V(t) = R_L i(t)$) and the capacitance of the detector C_e. Moreover, $\tau_{RC} = R_L C_e$ is the time constant of the electric circuit. Such a circuit has two poles: a thermal pole at $f_{th} = 1/\tau_{th}$ and an electric one at $f_{RC} = 1/\tau_{RC}$.

4.6 Quantum Detectors

Quantum detectors measure the photon flux impinging onto a particular region, called the sensitive area, of the detector. In these detectors, photons in the UV-IR spectral region have a one-to-one interaction with the valence electrons inside the active material (typically a semiconductor)

which results in an output current. Such a current is proportional to the photon flux through the quantum efficiency η which is a measure of the likelihood of collecting an output electron per absorbed photon. Often, an amplification process is required to increase the sensitivity. Thanks to this, quantum detectors can reach a single photon sensitivity. However, amplification is not free. Extra noise is always added to the insitric noise of the signal because of the amplification stage. This drawback is measured by a characteristic parameter of any amplifier called *Noise Figure*. It is the ratio between the SNR before and after the amplification stage. We can distinguish two main categories of photodetector:

- Phtoemissive Detectors
- Junction Photodectors

4.6.1 Photoemissive Detectors

Photoemissive detectors rely on electrons emitted, through *photoelectric effect*, by a *photocathode* which is made by a thin layer of a sensitive material.

- **Photoelectric Effect**. When a radiation of wavelength λ impinges on a metal, electron emission takes place provided that the photon energy is greater than the work function Φ of the metal. The work function is the energy difference between the fermi level and the vacuum level, which is the energy of an electron at rest far away from any material. Hence, the work function is the energy that electrons must acquire from an electromagnetic field or any other source of energy, to become free particles. This sets a wavelength threshold λ_{th} for electron emission, given by:
$$\lambda_{th} < \frac{hc}{e\Phi} \tag{50}$$
The energy in excess provided by the photon becomes kinetic energy of the emitted electrons. However, when excited electrons collide with the lattice and/or with thermal electrons, they lose their energy. Only a small fraction of electrons that do not undergo any collision can escape the metal with the maximum kinetic energy:
$$E_{k_{max}} = h\nu - \Phi \tag{51}$$
Most of the electrons exit the metal with a kinetic energy smaller than this.

- **Photocathode Materials**. Photocathodes can be theoretically made by metals but nowadays they are mostly made by semiconductors because of their higher spectral sensitivity. Moreover, semiconductors have a higher quantum efficiency. In fact, since they have a very low free electron density the collisions of the emitted electron are mostly due to lattice collisions, i.e. phonon collisions. In these types of collisions, the energy loss is very low resulting in a higher quantum efficiency. The thickness of the photocathode can be set by using the Lamber-Beer law:
$$I(x) = I_0 e^{-\mu_a x} \tag{52}$$
where μ_a is the attenuation coefficient. In the case in which a photocathode is made by a semiconductor, the energy required by an electron to be emitted, i.e. the energy provided by the electromagnetic field, must be $h\nu > E_G + \chi$ where χ is the electron affinity of the material (see Fig.10). The corresponding threshold wavelength is:
$$\lambda_{th} < \frac{hc}{e(E_G + \chi)} \tag{53}$$
To improve the spectral sensitivity and the quantum efficiency, N**egative Electron Affinity (NEA) photocathodes** have been developed. NEA photocathodes are made of a p-doped semiconductor, for example, GaAs, activated by forming a layer of Cesium or Cesium oxide (Cs_2O) on the surface of the crystal. This deposition behaves as an n-type material (donor) resulting in a layer of ionised acceptors at the very surface of the p-type semiconductor, since holes undergo electrostatic repulsion. In NEA photocathodes, once electrons have been promoted from the valence band to the conduction band by photon absorption, they are allowed to escape even after thermalisation since the bottom of the conduction band is above the vacuum level (see Fig.11). In this case, the threshold wavelength is given by:
$$\lambda_{th} < \frac{hc}{eE_G} \tag{54}$$

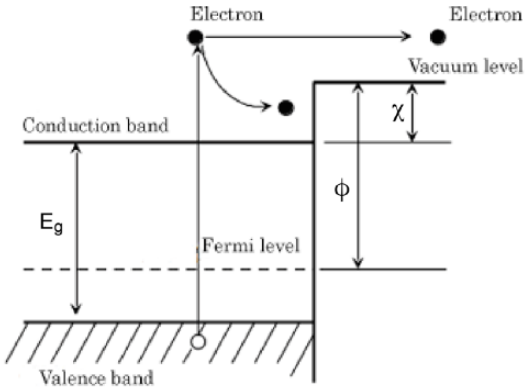

Figure 10: Schematic of the energy levels in a semiconductor photocathode.

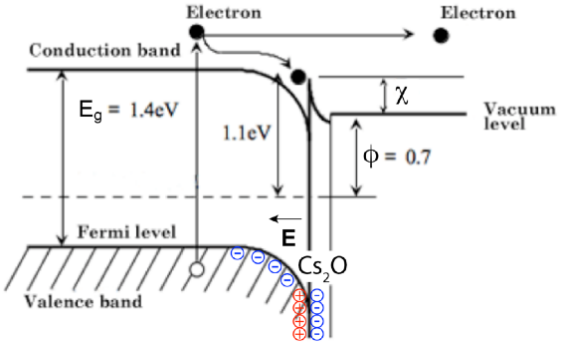

Figure 11: Schematic of the energy levels in a NEA photocathode.

4.6.2 Vacuum Photodiodes

The vacuum photodiode is the simplest photoemissive detector. It is made by a photocathode and an anode placed into a vacuum tube. A voltage ΔV is applied between the anode and cathode to collect at the anode all the electrons emitted by the photocathode, upon illumination (see Fig.12). The output current in a photodiode is given by:

$$I_\lambda = \Phi_\lambda \eta e = \frac{P_\lambda}{h\nu}\eta e = \frac{P_\lambda \lambda}{hc}\eta e \tag{55}$$

and the responsivity in terms of current or voltage is:

$$R^I_\lambda = \frac{I_\lambda}{P_\lambda} = \frac{\lambda}{hc}\eta e, \qquad R^V_\lambda = \frac{V_\lambda}{P_\lambda} = \frac{\lambda}{hc} R_L \eta e \tag{56}$$

Vacuum photodiodes have low sensitivity, fast temporal response and typically require an external amplification. There are three main sources of noise in vacuum photodiodes:

- **Dark current of the photocathode** due to thermionic emission. The current density exiting any metal for thermionic effect is given by the Richardson-Dushman equation:

$$J_T = \left[\frac{k_B^2 m_e e}{2\pi^2 \hbar^3}\right] T^2 \exp -\frac{\Phi}{k_B T} \tag{57}$$

  This effect can be reduced by lowering the temperature, but below certain values cooling might decrease the quantum efficiency.

- Even in vacuum, some ions are still present in the tube. These are attracted towards the cathode and when they collide they can induce electronic emission.

- **Shot noise** due to the charge quantization of the original current (the one of the photoemitted electrons). From the theory of shot noise we have:

$$i^{RMS}_{shot} = \sqrt{2eI_\lambda \Delta f} \tag{58}$$

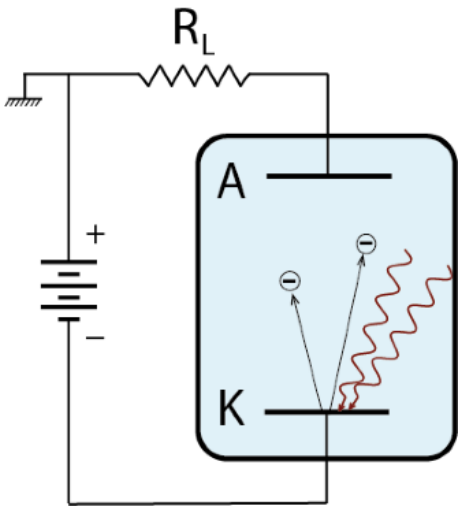

Figure 12: Schematic of a vacuum photodiode.

There are three different ways to express the Signal to Noise ratio, in terms of amplitude, power and decibel:

$$\left(\frac{S}{N}\right)_A = \frac{I_\lambda}{i_{Shot}^{RMS}} = \sqrt{\frac{P_\lambda \eta \lambda}{2hc\Delta f}} \tag{59}$$

$$\left(\frac{S}{N}\right)_P = \frac{I_\lambda^2}{(i_{Shot}^{RMS})^2} = \frac{P_\lambda \eta \lambda}{2hc\Delta f} \tag{60}$$

$$\left(\frac{S}{N}\right)_{dB} = 10\log_{10}\left(\frac{P_\lambda \eta \lambda}{2hc\Delta f}\right) \tag{61}$$

Overall, we have three contributions to the total current: the signal current I_s, the dark current I_d and the background current I_b due to other possible effects. So the total current is $I_t = I_s + I_d + I_b$. Moreover, we have the thermal noise of the load resistor R_L:

$$i_{th}^{RMS} = \sqrt{\frac{4k_B T}{R_L}\Delta f} \tag{62}$$

If we neglect the background effect ($I_b = 0$) and without any input radiation ($I_s = 0$), the current esclusively given by noise sources i_n^{RMS} is:

$$i_n^{RMS} = \sqrt{2eI_d\Delta f + \frac{4k_B T}{R_L}\Delta f} \tag{63}$$

Finally, we can compute the NEP by considering a sinusoidal input $P_{in}(t) = P_0 + P_0 \sin(\omega t)$. In this case, the r.m.s. of the minimum detectable current signal is:

$$i_s^{RMS} = P^{RMS} R_\lambda^I = \frac{P_0}{\sqrt{2}} R_\lambda^I = \frac{P_0}{\sqrt{2}} \frac{\lambda \eta e}{hc} \tag{64}$$

And the NEP:

$$NEP = \frac{i_n^{RMS}}{R_\lambda^I} = \frac{hc}{\lambda \eta e}\sqrt{2eI_d\Delta f + \frac{4k_B T}{R_L}\Delta f} \tag{65}$$

4.6.3 Photomultiplier Tubes

A photomultiplier tube (PMT) is a photoemissive detector consisting of a light input window, a photocathode, an electron multiplier, made by a sequence of dynodes, and an anode, assembled into a vacuum container (see Fig.13). When light enters the window, photoelectrons are emitted by the photocathode and are then accelerated and focused to strike on the first dynode, where electron multiplication takes place by secondary emission. This secondary emission is repeated at each of the subsequent dynodes. Dynodes are made of materials prone to emit secondary electrons upon receiving the kinetic energy of incoming electrons accelerated by an electric field. An external

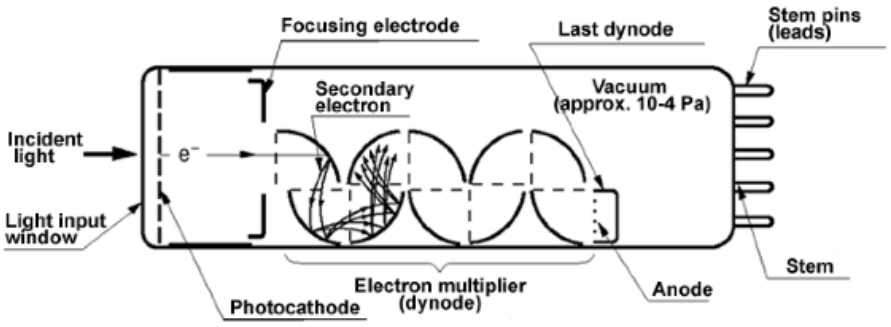

Figure 13: Schematic of a photomultiplier tube.

voltage potential is applied between the anode and cathode to accelerate electrons while keeping the trajectory collimated. Special care has to be put into the optimisation of the trajectories from the photocathode to the first dynode because the time performance of PMTs mainly depends on the minimisation of this transit time. When a primary electron with energy E_k strikes the surface of a dynode, δ secondary electrons are emitted on average. In the simplest case of N equal dynodes, the current amplification factor or gain G is:

$$G = \delta^N \tag{66}$$

and the responsivity is G-times the one of a vacuum photodiode:

$$R_\lambda = G\frac{\lambda\eta e}{hc} \tag{67}$$

Moreover, the relation between δ and the applied voltage V_B is almost linear in some conditions: $\delta \propto V_B$. High-gain photomultiplier tubes are capable of detecting single photons. The noise sources of photomultiplier tubes are the same as vacuum photodiodes. The dark current emitted by the cathode is still given by the Richardson-Dushman equation. Yet, in the case of PMTs, thermally emitted electrons undergo multiplication, as much as signal electrons. An additional noise source, not present in phototubes, comes from the multiplication process, which has a stochastic uncertainty. In fact, the gain G is not a deterministic quantity but is affected by a variance $v = \sigma_G^2$, where σ is the standard deviation of the values of G measured in a sequence of experiments. What is most important is the relative variance (RV_G) of G, which is the ratio of the variance to the square of the mean gain. It is possible to prove:

$$RV_G = \frac{\sigma_G^2}{G_{mean}} \approx \frac{1}{\delta_1}\frac{\delta}{\delta - 1} \tag{68}$$

where σ_1 is the electron multiplication factor of the first dynode. Another important parameter is the Noise Factor F:

$$F = \frac{(S/N)_{in}^2}{(S/N)_{out}^2} = 1 + RV_G = \frac{\delta}{\delta - 1} \tag{69}$$

where the last equivalence holds when all the dynodes have the same multiplication factor. Finally, the r.m.s. noise current is similar to the vacuum photodiodes one:

$$i_n^{RMS} = \sqrt{i_{shot}^{RMS} + i_{th}^{RMS}} = \sqrt{2eG^2(I_s + I_d)F\Delta f + \frac{4k_BT}{R_L}\Delta f} \tag{70}$$

4.7 Junction Photodetectors

Light detectors based on a pn junction are the most performant and widespread detectors in the VIS-NIR spectral range. When a junction between semiconductors with opposite doping is formed, the diffusion of the majority carriers to the other side of the junction leads to the creation of a depletion layer deprived of majority carriers across the junction (see Fig.14).

4.7.1 PN and PIN Photodiodes

The schematic of a pn photodiode is shown in Fig.15(a). When a photon is absorbed within the

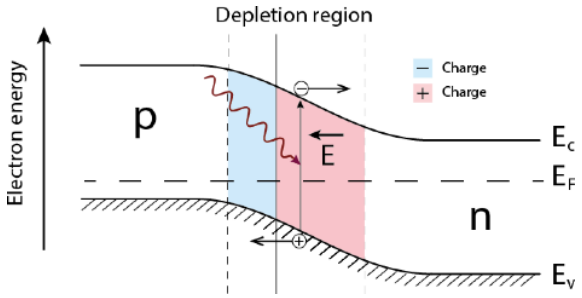

Figure 14: pn junction showing the charge depletion region.

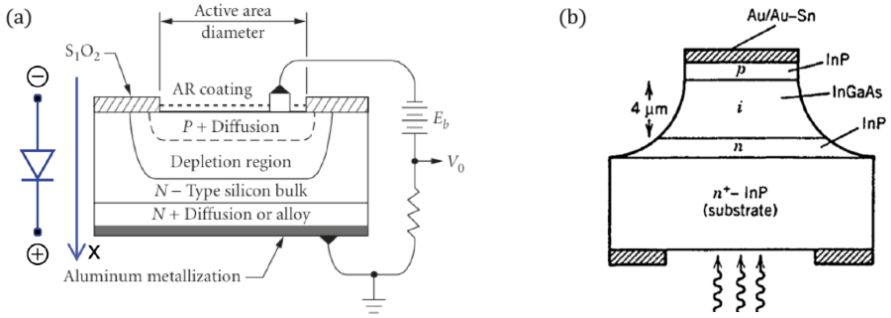

Figure 15: (a) PN photodiode and (b) PIN photodiode.

depletion layer, the generated electron and hole are separated and accelerated by the electric field. An electron is expelled from the n-type semiconductor into the external circuit and, at the same time, an electron enters the p-type semiconductor and recombines with a hole. Therefore, even if a couple of charges (e-h) was generated, a single electron charge flows in the external circuit per each absorbed photon. The quantum efficiency η depends on the reflectance r at the air-silicon interface:

$$\eta \approx (1-r)\left(1-e^{-\alpha w_d}\right) \tag{71}$$

where α is the absorption coefficient and w_d is the tickness of the depleation layer. The current generated is:

$$I_\lambda = \Phi_\lambda \eta e = \frac{P_{in}}{h\nu}\eta e = \frac{E_0 A \lambda}{hc}\eta e \tag{72}$$

where E_0 is the irradiance and A is the active area. The responsivity in terms of current is:

$$R_\lambda^I = \frac{I_\lambda}{P_{in}} = \frac{\lambda \eta e}{hc} \tag{73}$$

To widen the active layer, while keeping a low bias voltage an enhanced photodiode configuration, called PIN, has been devised (see Fig.15)(b). In PIN photodiodes a thick layer of intrinsic semiconductor is sandwiched between the p and n-type semiconductors. The intrinsic region is completely depleted with just a voltage drop of a few volts and the size of the active layer is almost independent of the bias voltage. Quantum efficiency is therefore:

$$\eta \approx (1-r)\left(1-e^{-\alpha w_i}\right) \tag{74}$$

where w_i is the width of the intrinsic region. In PIN photodiodes the noise sources are the usual ones. The r.m.s. of the noise current is:

$$i_n^{RMS} = \sqrt{2e(I_s+I_d)\Delta f + \frac{4k_BT}{R_L}\Delta f} \tag{75}$$

In general, at room temperature, the thermal noise is several orders of magnitude bigger than the shot noise, as in the vacuum photodiodes case. If we neglect the shot noise in the previous expression, the SNR can be easily found:

$$\left(\frac{S}{N}\right)_I = \frac{i_s^{RMS}}{i_n^{RMS}} = \underbrace{R_\lambda^I P_{in}}_{i_s^{RMS}} \underbrace{\sqrt{\frac{R_L}{4k_BT\Delta f}}}_{i_n^{RMS}} = P_{in}\sqrt{\frac{R_L}{4k_BT\Delta f}}\frac{\lambda\eta e}{hc} \tag{76}$$

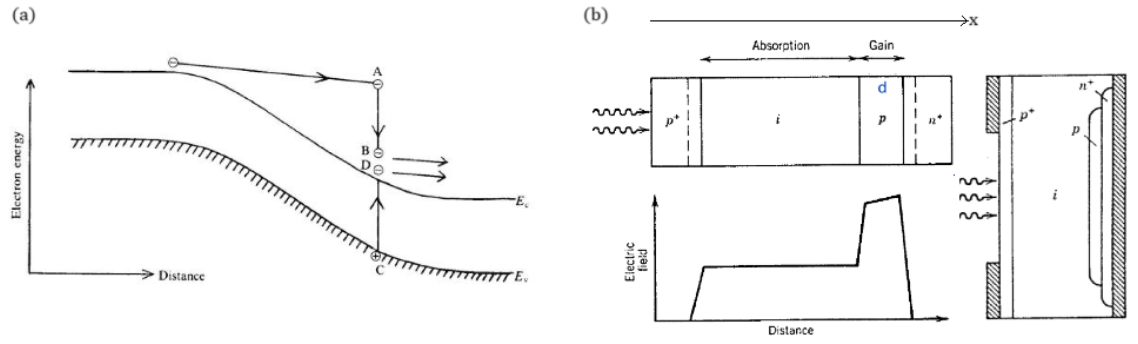

Figure 16: (a) Operating principle of avalanche photodiodes. (b) Schematic of an avalanche photodiode.

The NEP can be found through this expression since it is nothing more than the input power required to get $S/N = 1$. So that:

$$\left(\frac{S}{N}\right)_I = 1 \Rightarrow NEP = \frac{hc}{\lambda\eta e}\sqrt{\frac{4k_BT\Delta f}{R_L}} \tag{77}$$

and the NEP^* is simply:

$$NEP^* = \frac{NEP}{\sqrt{\Delta f}} = \frac{hc}{\lambda\eta e}\sqrt{\frac{4k_BT}{R_L}} \tag{78}$$

It is important to notice that all these expressions are valid in the case in which **the shot noise is neglectable**.

4.7.2 Avalanche Photodiodes

In low light conditions, the limit to the sensitivity (NEP) of junction photodetectors is set by the thermal noise of the load resistance. Therefore, to improve their sensitivity, internal amplification can be obtained by exploiting a charge multiplication mechanism similar to the one of photomultipliers. When a high reverse bias is applied to a pn junction, primary electrons and holes generated within the depletion layer can acquire enough kinetic energy from the electric field to generate secondary couples e-h by impact ionisation (see Fig.16). A sequence of impact ionisation events can lead to an avalanche process that results in a multiplication factor M. This increases the responsivity and reduces the NEP. In actual devices, the generation region is the intrinsic one while the multiplication region is either the p or the n-type one. The multiplication factor M can be defined as the ratio of the electron current exiting the gain region ($x = d$) to the electron current entering that region ($x = 0$). It is possible to demonstrate:

$$M = \frac{i_e(0)}{i_e(d)} = \frac{k_A - 1}{k_A - e^{\alpha_e d(k_A - 1)}} \tag{79}$$

where d is the width of the multiplication/gain region, α_e [α_h] is the electron [hole] impact ionisation coefficient and $k_A = \alpha_h/\alpha_e$. The responsivity is equal to the one of a standard photodiode multiplied for a multiplication factor dependent on the frequency $M(f)$:

$$R_{APD}(f) = M(f)R_\lambda = \frac{M_0}{\sqrt{1 + 4\pi^2 f^2 \tau_e^2 M_0^2}}\frac{\lambda\eta e}{hc} \tag{80}$$

where τ_e is the transit time which is equal to:

$$\tau_e = k_A\tau_{drift}, \qquad if \;\; k_A << 1 \tag{81}$$

and M_0 is the multiplication factor. As it was for photomultiplier, also in avalanche photodiodes we must take into account an additional noise source due to the variance of the multiplication factor M. The Noise Factor F_A strongly depend on k_A:

$$F_A = k_A M + (1 - k_A)\left(2 - \frac{1}{M}\right) \tag{82}$$

The r.m.s. noise current in APDs can be obtained by taking into account the multiplication process and the Noise Factor. Each primary electron undergoing an avalanche multiplication becomes a bunch of M electrons. Then, we need to consider that both the signal and the dark currents are multiplied by the factor M. Hence, the r.m.s. noise current is:

$$i_n^{RMS} = \sqrt{i_{shot}^{RMS} + i_{th}^{RMS}} = \sqrt{2eM^2(I_s + I_d)F_A\Delta f + \frac{4k_BT}{R_L}\Delta f} \tag{83}$$

which is the same expression of photomultiplier tubes. Then we have:

$$\left(\frac{S}{N}\right)_I = \frac{i_s^{RMS}}{i_n^{RMS}} = \frac{MR_\lambda P_{in}}{\sqrt{2eM^2(I_s + I_d)F_A\Delta f + \frac{4k_BT}{R_L}\Delta f}} \approx MR_\lambda P_{in}\sqrt{\frac{R_L}{4k_BT}\Delta f} \tag{84}$$

where the last equivalence holds if we neglect the shot noise. In this case, we also get:

$$NEP = \frac{1}{MR_\lambda}\sqrt{\frac{4k_BT}{R_L}\Delta f} \tag{85}$$

5 MOS Capacitors and CCD

The devices presently available for scientific imaging belong to two major families, namely Charge Coupled Devices (CCD) and Complementary Metal Oxide Semiconductors (CMOS). Even if great differences exist between the two technologies, they share an important common feature: they are both based on the charge storage capability provided by the depletion layer of a properly biased semiconductor substrate. The electrons generated by photons, proportionally to the light intensity absorbed by the semiconductor substrate are accumulated in Metal Oxide Semiconductor (MOS) capacitors that form the pixels of the sensor. Then, after the exposure of the detector to light, a charge transfer process takes place to bring the electrons accumulated in the matrix of sensitive elements to a readout circuitry where charge-to-voltage conversion takes place. The ideal MOS capacitor and its band diagram are depicted in Fig.17.

5.1 Operating Regimes

Depending on the gate bias voltage V_G there are three different regimes in which the MOS structure can be found. Before dealing with them just recall that the electron and hole density are given by:

$$n = n_i \exp\left(\frac{E_F - E_i}{k_BT}\right), \qquad p = n_i \exp\left(\frac{E_i - E_F}{k_BT}\right) \tag{86}$$

- **Accumulation Regime** $V_G < 0$: the Fermi level of the metal is pulled up by a quantity $|qV_G|$. The voltage drop must be reflected across the system but since the metal cannot sustain any electric field on its inside, this voltage drop must be established in both the semiconductor and the insulator. In particular, most of the voltage will drop on the insulator because of its smaller dielectric constant. The holes in the p-type semiconductor are attracted

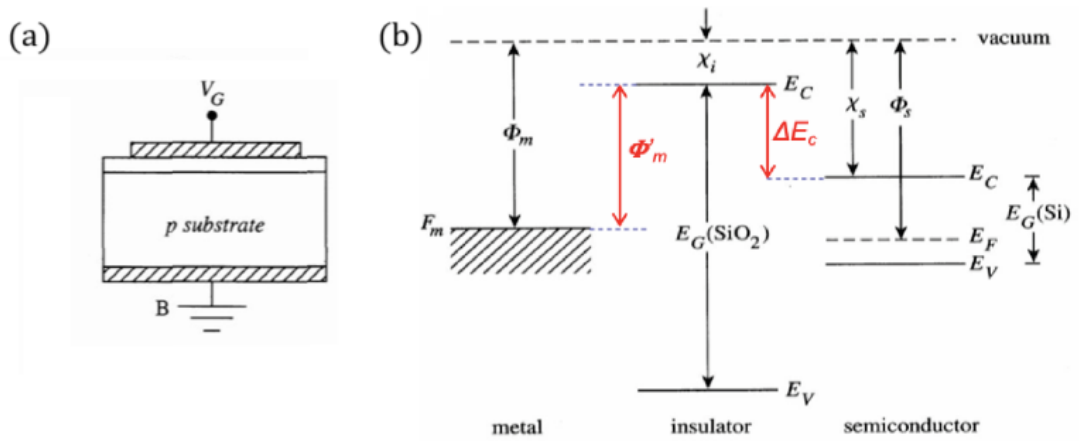

Figure 17: Schematic (a) and band diagram (b) of an ideal MOS structure.

by the electric field toward the interface with the insulator, where their density increases above the density in the bulk. Hence, an excess of positive charge accumulates in the semiconductor within a layer of small thickness. The excess charge density equals the modulus of the negative charge density on the metal and screens the bulk of the semiconductor from the surface field. This regime is not relevant for the use of MOS capacitors in CCD.

- **Depletion Regime** $V_G > 0$: the Fermi level of the metal is pulled down by $|qV_G|$. The insulator and semiconductor bands bend downward, and the electric field is now directed toward the bulk. The holes in the semiconductor are pushed away from the interface, where their density becomes negligible compared to the acceptor concentration. Hence, a depletion region is formed, which leaves a negative charge, with volumetric density qN_A, made by the ionised acceptors. The thickness W of the depletion layer increases with V_G and settles to the value required for electric neutrality.

- **Inversion Regime**: as the gate voltage keeps increasing and we give time to the MOS system to reach equilibrium, not only does the depletion width grow, but the band bending in the semiconductor also increases. The electron concentration grows as well and, eventually, it becomes equal to that of the holes in the bulk ($p_{bulk} = N_a$). When this is the case, we say that we are at the **threshold of strong inversion**. The strong inversion is defined by the condition:

$$n_{interface} = p_{bulk} = N_a \tag{87}$$

  The predominant charge carrier at the interface has changed from holes to electrons. The gate voltage at the threshold of strong inversion is denoted by V_T.

- **Deep Depletion Regime**: Let's now analyze a non-equilibrium state. Suppose we can sweep V_G rapidly from the accumulation regime to $V_G > V_T$. Immediately afterwards, the electron density at the interface is extremely small. So, the electron quasi-Fermi level is close to the valence band at the surface and keeps flat through the depletion region. The hole quasi-Fermi level, on the contrary, remains flat all over the semiconductor since holes are majority carriers and they take a short time to rearrange themselves in the substrate. So we will have the mass-action law not respected, which results in a net generation of electron-hole pairs. The very final conclusion of the process is the electron concentration rising at the interface. *Deep depletion is a very important non-equilibrium condition since image detectors operate in this regime to accumulate electrons generated by photons.*

5.2 Capacitance in different regimes

Let's now explore the behaviour of the capacitance C of a MOS capacitor in the different operating regimes. The apparatus to measure the capacitance is made by a constant voltage generator V_G (static gate voltage) and a small sinusoidal voltage generator v_s whose frequency can be switched from high frequency (HF) to low frequency (LF). In general, for such a circuit, the capacitance can be measured as:

$$C = \frac{i}{2\pi f v_s} \tag{88}$$

where i is the signal current flowing in the circuit. The behaviour of the capacitance as a function of V_G and the test circuit are shown in Fig.19.

- **Accumulation Regime** $V_G < 0$: the MOS capacitor is in the accumulation regime. This means that plenty of holes are available at the interface with the oxide. The system behaves like a plane capacitor of thickness t_i, filled with a dielectric of permittivity ϵ_i. Hence, its capacitance per unit area C is given by:

$$C = \frac{\epsilon_i}{t_i} = C_i \tag{89}$$

- **Depletion Regime** $V_G > 0$: free carriers in the semiconductor have moved away from the interface within a layer of thickness W. Now, the signal voltage modulates the electrons in the metal and the holes at the border of the depletion layer in a thin layer dW. The capacitance decreases and can be easily calculated considering the series of two capacitors of different thickness and electric permittivity: t_i, ϵ_i for the oxide and W, ϵ_s for the semiconductor. Therefore, the capacitance in the depletion regime is:

$$\frac{1}{C} = \frac{1}{C_i} + \frac{1}{C_s} = \frac{t_i}{\epsilon_i} + \frac{W}{\epsilon_s} \tag{90}$$

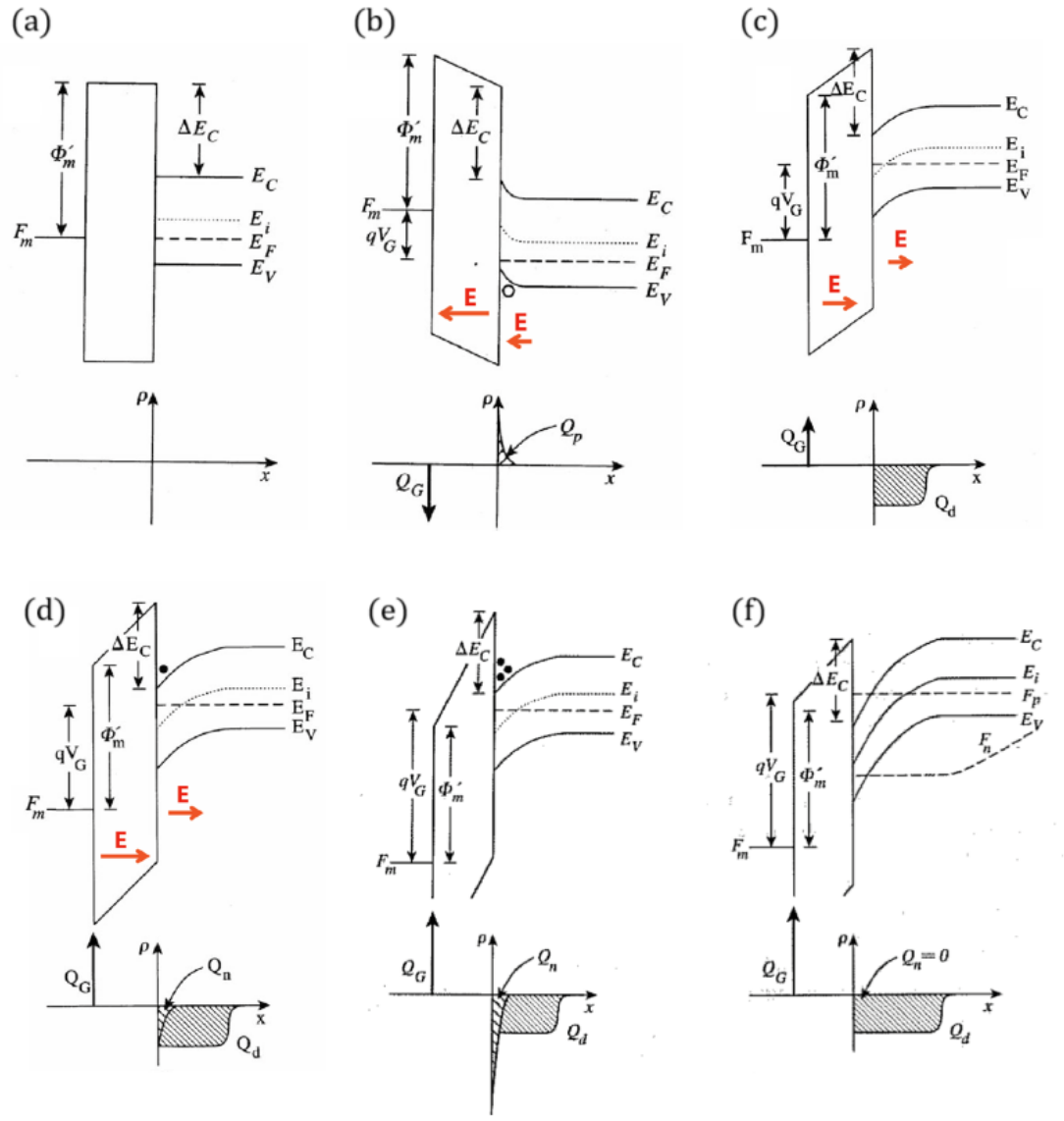

Figure 18: Band structure, charge density and electric field in (a) Ideal MOS capacitor $V_T = 0$ (or $V_T = V_{FB}$ in real case), (b) Accumulation regime, (c) Depletion regime, (d) Threshold for strong inversion $V_G = V_T$, (e) Strong inversion $V_G > V_T$ and (f) Deep Depletion.

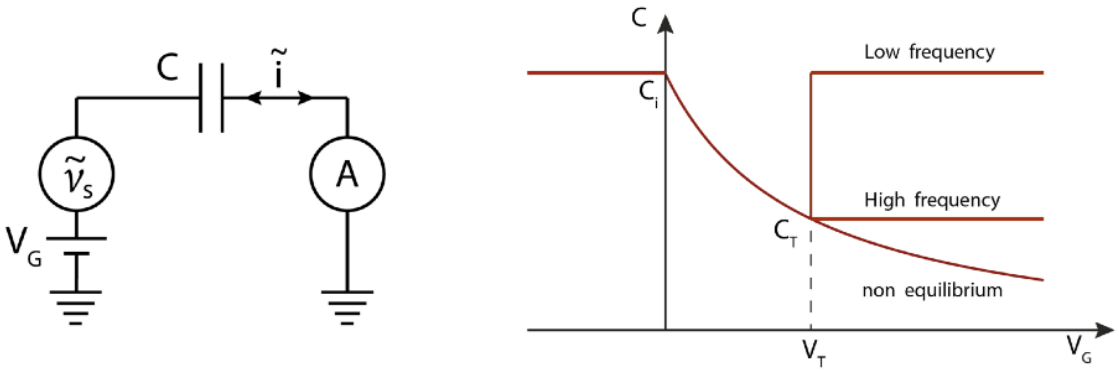

Figure 19: Test circuit and capacitance as a function of the gate voltage.

- **Inversion Regime**: by increasing V_G, the threshold for strong inversion is eventually reached. Here, we need to distinguish between LF and HF small signal inputs:
  - *LF modulation*: the density of electrons at the interface with the oxide equals that of holes in the bulk. Therefore, the signal voltage now modulates the electron density Q_n. Since the modulated charge is now at the interface, like it was in the accumulation regime, the capacitance quickly rises to the one of the insulator $C = C_i$.
  - *HF modulation*: the capacitance is clipped to the one reached at the threshold value, since the modulated charge is still made by holes at the edge of the depletion layer, whose thickness does not change with V_G.
- **Deep Depletion Regime**: the gate voltage is switched very quickly from zero to the final value $V_G > V_T$ and the analysis is made immediately afterwards. In this case, not enough time is given to Q_n to build up and the capacitance continues to decrease with VG.

5.3 Noise in CCDs

The noise that affects CCDs and other image sensors, like CMOS sensors, can be classified into the following four main sources:

- **Fixed Pattern Noise** N_f: is caused by variations in sensitivity between pixels. These variations may be due to:
  - Quantum efficiency differences.
  - Non-uniformities in the aperture area and in the thickness.
  - Instabilities in the electronics.
- **Shot Noise of the Signal** N_s: is generated by statistical variation in the photon flux incident on a CCD. It is expressed by the standard deviation of the number of photons collected per pixel during the exposure time (assumed to be Poissionian):

$$N_s = \sqrt{S} \tag{91}$$

  where S represents the Signal, namely the number of light-generated electrons in a pixel in the exposure time. This noise is unavoidable.
- **Dark Shot Noise** N_d: is caused by dark current and is proportional to the square root of the number of electrons generated per pixel in the dark state. Can be reduced by cooling.
- **Readout Noise** N_r: is mainly due to electrical fluctuations from thermal noise caused by the MOSFET used as the amplifier in the CCD output stage. It also comes from the readout circuit of the CCD. The Readout noise is the fundamental uncertainty in the output of a CCD.

Since these noise sources are independent, the total noise is given by:

$$N_t = \sqrt{N_f^2 + N_s^2 + N_d^2 + N_r^2} \tag{92}$$