

Chapter 1

Introduction to Semiconductor Photodetectors

1.1. Brief overview of semiconductor materials

A semiconductor material is a continuous crystalline medium characterized by an energy band structure corresponding, in the case of an infinite crystal, to a *continuum* of states (which, in practice, means that the characteristic dimensions of the crystal are significantly larger than the lattice parameter of the crystal structure; this applies as long as the crystal dimensions are typically larger than a few dozen nanometers). In general terms, the energy structure of a semiconductor consists of a valence band corresponding to molecular bonding states and a conduction band representing the molecular anti-bonding states. The energy range lying between the top of the valence band and the bottom of the conduction band is known as the forbidden band, or more commonly the bandgap. An electron situated in the valence band is in a ground state and remains localized to a particular atom in the crystal structure, whereas an electron situated in the conduction band exists in an excited state, in a regime where it interacts very weakly with the crystalline structure. What

2 Optoelectronic Sensors

differentiates semiconductors from insulators is essentially the size of the bandgap: we refer to semiconductors where the bandgap of the material is typically less than or equal to 6 eV, and to insulators when the bandgap is more than 6 eV: above this, the solar spectrum arriving on the Earth's surface is unable to produce inter-band transitions of electrons situated in the valence band of the material. Semiconductor materials are mostly divided into two large classes: elemental semiconductors (group IV of the periodic table): silicon, germanium, diamond, etc. and compound semiconductors: IV-IV (SiC), III-V (GaAs, InP, InSb, GaN) and II-VI (CdTe, ZnSe, ZnS, etc.). Impurities can be introduced into the volume of the semiconductor material and can modify its electrical conduction properties, sometimes considerably. An impurity is known as a *donor* when it easily releases a free electron into the conduction band. The characteristic energy level of the impurity is therefore in the bandgap, slightly below the conduction band. For example, in the case of compound semiconductors in group IV of the periodic table such as silicon, the main donor impurities are those which, being from group V of the periodic table (arsenic, phosphorous, etc.), are substituted in place of a silicon atom in the crystal structure: since silicon is tetravalent, these atoms naturally form four covalent bonds with the silicon atoms around them, and also easily give up their surplus electron to the crystal structure. These electrons become free to move, subject to a weak activation energy provided by thermal agitation. In this case we refer to n-type doping. In the case of silicon, a group III element incorporated into the crystal structure of silicon naturally forms three covalent bonds around it, and then completes its own outer-shell electronic structure by capturing an electron from its fourth nearest-neighbor silicon atom, again subject to a weak thermal activation energy. Such an impurity is known as an *acceptor*, and doping with acceptors is known as p-type doping. A hole carrying a positive elementary charge and corresponding to a vacant energy state in the valence band is therefore left in the crystal structure of the silicon. In the case of III-V composites, the donors are mostly atoms from group IV (silicon) substituted in place of group III elements, or group VI elements (S, Se, Te) substituted in place of group V elements, and acceptors are group II (zinc, magnesium) substituted in place of group

III elements. In the case of II-VI composites, the most commonly-encountered donors belong to group VII (chlorine, etc.) substituted in place of group VI elements, and acceptors belong to either group I (lithium, etc.) or to group V (nitrogen, arsenic, phosphorous, etc). In this last case, the group V element is substituted in place of a group VI element in the semiconductor crystal structure, whereas group I acceptors are substituted in place of group II elements. The chemical potential, or Fermi energy, of an *intrinsic* semiconductor (i.e. one free from n and p impurities) is found in the middle of the bandgap of the material. When a moderate n-type doping is added, the Fermi level rises from the middle of the bandgap towards the conduction band, by an increasing amount as the level of doping rises. When the level of n-type doping becomes large, the Fermi level can cross the bottom of the conduction band and be found inside this band (Mott transition). The semiconductor then behaves like a metal and for this reason is called a semi-metal. In this case it is referred to as degenerate. In the case of p-type doping, the semiconductor is said to be degenerate when the Fermi level is below the top of the valence band.

1.2. Photodetection with semiconductors: basic phenomena

Photodetection in semiconductors works on the general principle of the creation of electron-hole pairs under the action of light. When a semiconductor material is illuminated by photons of an energy greater than or equal to its bandgap, the absorbed photons promote electrons from the valence band into excited states in the conduction band, where they behave like free electrons able to travel long distances across the crystal structure under the influence of an intrinsic or externally-applied electric field. In addition, the positively-charged holes left in the valence band contribute to electrical conduction by moving from one atomic site to another under the effects of the electric field. In this way the separation of electron-hole pairs generated by the absorption of light gives rise to a *photocurrent*, which refers by definition to the fraction of the photogenerated free charge-carriers collected at the edges of the material by the electrodes of the photodetecting structure, and whose intensity at a given

4 Optoelectronic Sensors

wavelength is an increasing function of the incident light intensity. On this level we can distinguish between two large categories of photodetectors based on the nature of the electric field, which causes the charge separation of photogenerated electron-hole pairs: *photoconductors*, which consist of a simple layer of semiconductor simply with two ohmic contacts, where the electric field leading to the collection of the charge-carriers is provided by applying a bias voltage between the contacts at either end, and *photovoltaic* photodetectors, which use the internal electric field of a p-n or Schottky (metal-semiconductor) junction to achieve the charge separation. This last term covers p-n junction photodetectors (photovoltaic structures consisting of a simple p-n junction, and p-i-n photodetectors which include a thin layer of semiconductor material between the p and n region which is not deliberately doped), as well as all Schottky junction photodetectors (Schottky barrier photodiodes and metal-semiconductor-metal (MSM) photodiodes).

We will now briefly introduce the main physical concepts at the root of the operation of the different semiconductor photodetector families. Here the emphasis is placed on a phenomenological description of the working mechanisms of the devices in question; the corresponding formalism has been deliberately kept to an absolute minimum in the interests of clarity and concision.

1.3. Semiconductor devices

Photoconductors represent the simplest conceivable type of photodetector: they consist of a finite-length semiconductor layer with an ohmic contact at each end (Figure 1.1). A fixed voltage of magnitude V_B is applied between the two end contacts, in such a way that a bias current I_B flows through the semiconductor layer, simply following Ohm's law. The active optical surface is formed from the region between the two collection electrodes. When it is illuminated, the photogenerated charges produced under the effect of the applied electric field lead to a photocurrent I_{PH} which is added to the bias current, effectively increasing the conductivity of the device.

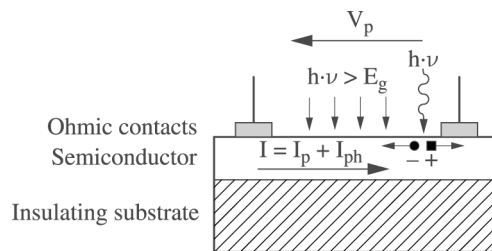


Figure 1.1. Diagram of a photoconducting device

The main point of interest in a photoconducting device is its increased gain, the response of photoconductors being typically several orders of magnitude greater than that of photovoltaic detectors for a given material. On the other hand, its other operational parameters (bandwidth, UV/visible contrast, infrared sensitivity) are generally below that of other types of photodetectors, which often greatly limits the scope of its potential applications (this is particularly the case for photoconductors based on III-V nitrides, as we will see later on).

1.4. p-n junctions and p-i-n structures

In p-n diodes, the metallurgical linkage of a region of a p-type doped semiconductor and a region of n-type doping forms a p-n junction, where the joining of the Fermi levels in equilibrium mostly occurs through a flow of charge between the n and p regions. In equilibrium we therefore find a region with no free charge carriers immediately around the junction, similar to a charged capacitor, where there are, on the n side, positively ionized donors and, on the p side, negatively ionized acceptors (this zone is known as the space charge region (SCR), where ionized donors and acceptors provide fixed charges). The presence of charged donors and acceptors produces an electric field in that region which curves the energy bands and, in equilibrium, forms an energy barrier between the two regions: the

bottom of the conduction band and the top of the valence band on the n side are below the corresponding levels on the p side (Figure 1.2).

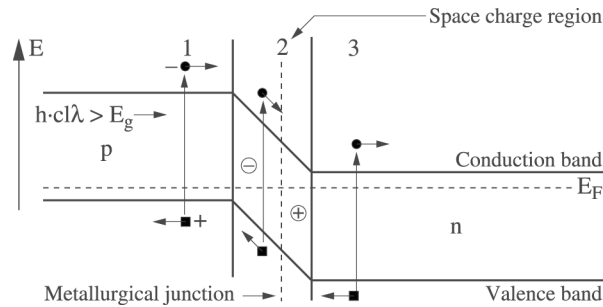


Figure 1.2. Curvature of the energy bands and mechanisms of photocurrent generation in a p-n junction

The width of the SCR is a decreasing function of the level of doping in the material, while the height of the energy barrier is an increasing function of it. An electron-hole pair produced in this SCR (situation 2 in Figure 1.2) is therefore separated by the effect of the internal electric field of the junction, and so does not recombine. These are the charge carriers which contribute to the photocurrent, to which we can add, to some extent, those generated at a distance from the junction less than or equal to the diffusion length (situations 1 and 3 in Figure 1.2). The band structure of the junction implies that the photocurrent will consist of minority charge carriers. For this reason, the photocurrent flows in the opposite direction to the bias on the diode, where the forward direction is defined as the direction of flow of the majority charge carriers (from the n to the p region in the case of electrons, and vice versa for holes). Moreover, the application of an opposing external electric field ($V_p - V_n < 0$) allows us to increase the height of the energy barrier in the vicinity of the junction, and also increase the spatial extent of the SCR, which significantly improves the efficiency of the separation of electron-hole pairs by increasing the electric field within the junction.

We note that when the doping level is moderate, the width of the SCR is important. This effect is beneficial in the case of p-n junction photodetectors, where in order to increase the photoresponse it is desirable to ensure that the mechanisms of electron-hole pair generation through incident light take place predominately inside the SCR. A simple means of increasing the spatial extent of the SCR is to introduce between the n and p regions a thin layer of intrinsic semiconductor material which is not intentionally doped: the structure is therefore referred to as p-i-n. Such a structure is interesting because it is possible to maintain high levels of doping in the n and p regions without significantly reducing the extent of the SCR, whose width is then largely determined by the thickness of the “i” layer. Additionally, increasing the width of the SCR reduces the capacitance of the structure, which makes p-i-n structures particularly well-suited for high-speed operation.

1.5. Avalanche effect in p-i-n structures

When the reverse-bias voltage established at the terminals of a p-i-n structure increases sufficiently that the electric field established in the junction reaches values close to the breakdown field (in structures of micron-scale thickness, this is generally the case when the bias voltage at the terminals reaches a few dozen volts), the photogenerated charge carriers in the SCR (which is effectively the region that is not intentionally doped) are accelerated enough to separate other secondary charge carriers from the atoms in the lattice that they impact in the course of their motion: this is the avalanche effect which results in a multiplication of the charge carriers in the SCR. The gain is therefore greater than 1 for the generation of charge carriers by light, and this gain can even typically reach 10 or 20 under favorable conditions. This effect is exploited in what are called avalanche photodiodes where the levels of n- and p-type doping are generally adjusted to high values above 10^{18} cm^{-3} to maximize the intrinsic electric field of the junction.

1.6. Schottky junction

A Schottky junction is formed by bringing a metal and a semiconductor into contact. The basic phenomena which lead to the formation of a Schottky junction with an n-type semiconductor are summarized in Figure 1.3.

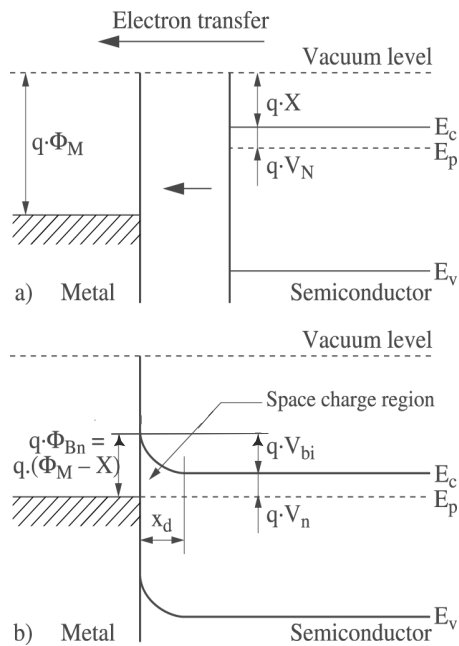


Figure 1.3. Formation of a Schottky junction (in an n-type semiconductor)

In thermal equilibrium, when the Fermi levels of the metal and the semiconductor are equalized, a transfer of electronic charge occurs from the semiconductor to the metal in the case where the work function $q \cdot \Phi_M$ of the metal (q being the elementary charge) is greater than the electron affinity X of the semiconductor, and a SCR appears at the edge of the semiconductor of width x_d next to the junction, where the only charges present are the positively-ionized donors. A curvature of the energy bands therefore occurs at the junction, which leads to the appearance of an energy barrier between the metal and the

semiconductor, called a *Schottky barrier*, whose height is given to first approximation by the expression:

$$q \cdot \Phi_{Bn} = q \cdot (\Phi_M - \chi) \quad [1.1]$$

In equilibrium, therefore, we find an intrinsic electric field immediately next to the metal-semiconductor junction which is comparable in form to that found in a p-n junction. Consequently, it is the phenomenon of photogeneration of charge carriers inside and near to the SCR which is responsible for the appearance of a photocurrent, with the electron-hole pairs being separated by the effect of the electric field in the Schottky junction. It is possible, as in the case of the p-n junction, to modify the intensity of the internal electric field in the junction by applying a bias voltage V between the semiconductor and the metal of the Schottky contact (Figure 1.4).

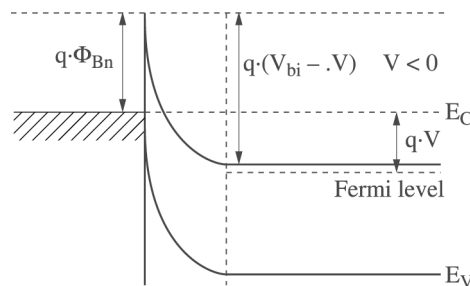


Figure 1.4. Reverse-bias of a Schottky junction (*n*-type semiconductor material)

In the case of an *n*-type semiconductor, the application of a negative voltage between the semiconductor and the metal electrode of the Schottky contact has the effect of reverse-biasing the Schottky junction, which leads to an increase in the height of the effective barrier, along with an increase in the width of the SCR. This last effect is of course favorable for photodetection. Indeed, it follows that the majority charge carriers (electrons) cannot flow towards the Schottky contact, and only the minority carriers (holes) generated by external excitation (in particular photogeneration) can reach the Schottky contact and hence produce an electric current: as in the case of the p-n

junction, we therefore find that the current flows in reverse through the Schottky junction, that is, from the semiconductor towards the Schottky contact. The illumination of Schottky photodiodes can occur through the front or rear face (often this second option is chosen in the case where the substrate material is transparent to the light to be detected, as is the case for example with sapphire). In the case of illumination through the front face, we resort to a semi-transparent Schottky contact, characterized by a very small thickness of metal (of the order of 100 Å) selected to ensure sufficient optical transmission: while a thin layer of gold of 100 Å thickness transmits up to 95% of the incident light in the infrared, the percentage transmitted in the ultraviolet is around 30% in the range 300-370 nm. The gain of p-i-n photodiodes (other than the specific case of avalanche photodiodes) and Schottky photodiodes is at most 1, which would be the case if all the photogenerated charge carriers were collected by the electrodes at the ends of the device.

1.7. Metal-semiconductor-metal (MSM) structures

An MSM structure consists of two Schottky electrodes, often interlinked in the form of a comb structure, leaving a free semiconductor surface between the two contacts which forms the active region in which light will be absorbed. A bias voltage can be applied between the two electrodes, in order to break the initial electrical symmetry of the contacts: one of the Schottky junctions is reverse-biased, producing a SCR of increased width, and the other junction is forward-biased.

The absorption of light near the reverse-biased junction creates electron-hole pairs which are separated under the effects of the electric field present in the SCR, thus creating the photocurrent. The other electrode, consisting of a forward-biased (and hence transmissive) Schottky junction, simply acts as a collection electrode. The band diagram of the device under increased bias voltage (V_B) is represented schematically in Figure 1.5, in which L is the distance between two adjacent contact fingers, Φ_0 is the height of the Schottky barrier and

I_{ph} is the photocurrent. MSM photodetectors normally use semiconductor materials which are not intentionally doped, are chemically very pure and electrically very resistive. The SCRs associated with Schottky junctions made of these materials are hence of significant width which, for a given bias voltage, allows the electric field of the junction to extend more easily into semiconductor regions some way from the contact. It follows that photogenerated electron-hole pairs are more easily separated and collected by the electrodes at either end.

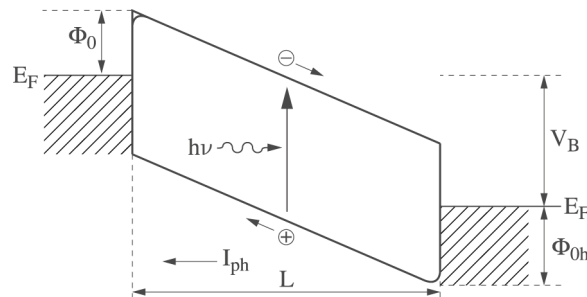


Figure 1.5. Energy band diagram for an MSM structure under electrical bias; effect of illumination

1.8. Operational parameters of photodetectors

The main parameters which define the behavior of an ultraviolet photodetector are respectively the response coefficient, the gain, the quantum efficiency, the bandwidth, the noise equivalent power (NEP) and the detectivity.

1.8.1. Response coefficient, gain and quantum efficiency

The response coefficient of a photodetector, R_i , links the photocurrent I_{ph} to the power of the incident light P_{opt} through the relationship:

$$I_{ph} = R_i \cdot P_{opt} \quad [1.2]$$

It is important to note in passing that the response coefficient is a quantity independent of the active optical surface of the photodetector structure: indeed, the photocurrent as well as the incident optical power are both, in the ideal case, proportional to the active optical surface. At a given wavelength λ , the flux Φ of photons arriving on the semiconductor surface, which is defined as the number of photons reaching the active surface per unit time, is given by:

$$\Phi = P_{opt} \cdot \lambda / (h \cdot c) \quad [1.3]$$

where h is the Planck constant and c is the speed of light.

The quantum efficiency η is defined as the probability of creating an electron-hole pair from an absorbed photon. Considering that all the incident light is absorbed in the semiconductor material, the rate G of electron-hole pair generation per unit time is thus given by:

$$G = \eta \cdot \Phi = \eta \cdot P_{opt} \cdot \lambda / (h \cdot c) \quad [1.4]$$

If we now introduce the gain parameter g which corresponds to the number of charge carriers detected relative to the number of photogenerated electron-hole pairs, then the photocurrent is given by the equation:

$$I_{ph} = q \cdot G \cdot g = q \cdot \eta \cdot P_{opt} \cdot \lambda / (h \cdot c) \cdot g = (q \cdot \eta \cdot \lambda / (hc) \cdot g) \cdot P_{opt} \quad [1.5]$$

where q is the elementary charge (1.602×10^{-19} C), from which we obtain the expression for the response coefficient of the detector:

$$R_i = q \cdot g \cdot \eta \cdot \lambda / (hc) \quad [1.6]$$

1.8.2. Temporal response and bandwidth

The speed of response of a photodetector may be limited by capacitative effects, by the trapping of charge carriers or by the saturation speed of charge carriers in the semiconductor. These phenomena all lead to a reduction in the response of the photodetector

in the high-frequency domain. The cutoff frequency f_C of the photodetector is defined as the frequency of optical signal for which the response coefficient is half that for a continuous optical signal. The temporal response of a photodetector is characterized by the fall time τ_f (or the rise time τ_r), which is defined as the time needed for the photocurrent to fall from 90% to 10% of its maximum (or to rise from 10% to 90% of it). In the case of a transient exponential response with a time constant τ , the following relationship links the bandwidth BW and the temporal response of the photodetector:

$$BP = 1 / (2 \cdot \pi \cdot \tau) = 2.2 / (2 \cdot \pi \tau_m) = 2.2 / (2 \cdot \pi \cdot \tau_d) \quad [1.7]$$

1.8.3. Noise equivalent power

The NEP is defined as the incident optical power for which the signal-to-noise ratio is 1, and hence the photocurrent I_{ph} is equal to the noise current I_b . In other words, it is the smallest optical power which can be measured. It follows that the NEP parameter is given by the equation:

$$NEP = I_b / R_i \text{ (in } W) \quad [1.8]$$

In the case of white noise, the noise current I_b increases as the square root of the bandwidth of the photodetector device. It follows that it is preferable and customary to use the following expression for the NEP, normalized with respect to the bandwidth BW :

$$NEP^* = NEP \cdot (BW)^{-1/2} \text{ (in } W \cdot Hz^{-1/2}) \quad [1.9]$$

In semiconductors, there are five sources of noise:

- shot noise, mainly due to the random nature of the collisions of incident photons;
- thermal noise, due to random collisions of charge carriers with the atoms of the crystal lattice, in permanent vibration due to thermal motion;
- partition noise, caused by the separation of the electric current into two parts flowing across separate electrical contacts;

– generation-recombination noise, caused by the random generation and recombination of charge carriers, either band to band or via trapping levels situated in the bandgap;

– $1/f$ noise, associated with the presence of potential barriers at the level of the electrical contacts. This last type of noise dominates at low frequencies.

1.8.4. Detectivity

This figure of merit is defined by the equation:

$$D = (NEP)^{-1} = R_i / I_b \left(\text{in } W^{-1} \right) \quad [1.10]$$

In general terms, the photocurrent signal increases in proportion to the active optical area A_{opt} , and in addition the noise current increases with the square root of the product of the active optical area with the bandwidth BW . It follows that the preferred method of comparing between different photodetectors is to use an expression for the detectivity normalized with respect to these parameters, written:

$$D^* = D \cdot (A_{opt} \cdot BP)^{1/2} = (R_i / I_b) \cdot (A_{opt} \cdot BP)^{1/2} \left(\text{in } W^{-1} \cdot cm.Hz^{1/2} \right) \quad [1.11]$$

The normalized detectivity is the most important parameter for characterizing a photodetector because it allows direct comparison of the performance of photodetectors using technologies and methods of operation which are at first glance very different. It is clear from the preceding definitions that the determination of the NEP and the detectivity requires measurement of three parameters: the response coefficient, the bandwidth and the noise current of the photodetector device. The measurement of the noise current must be made in darkness. The device is biased using a very stable voltage source, and the entire measurement system must itself have an intrinsic noise level considerably lower than the intrinsic noise of the photodetector device.