FRIEDRICH-SCHILLER-UNIVERSITÄT JENA
PHYSIKALISCH-ASTRONOMISCHE-FAKULTÄT

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

SOMMERSEMESTER 2021

Optoelectronics

FRANK SCHMIDL

LaTeX-Satz und Design von Martin Beyer

Contents

1 Introduction

Optoelectronics is a sub field of electronics or photonics. We can better say that optoelectronic devices are optical to electrical transducers which can *source, detect* and/or *control* light. Light in this context includes a wide frequency range of irradiation or electromagnetic waves.

Modern optoelectronics is based on the quantum mechanical effects of light on semiconducting or superconducting materials. We will/should discuss in which way light can interact with a solid. As an example we see here 1 the absortpion spectrum of a dielectric material.

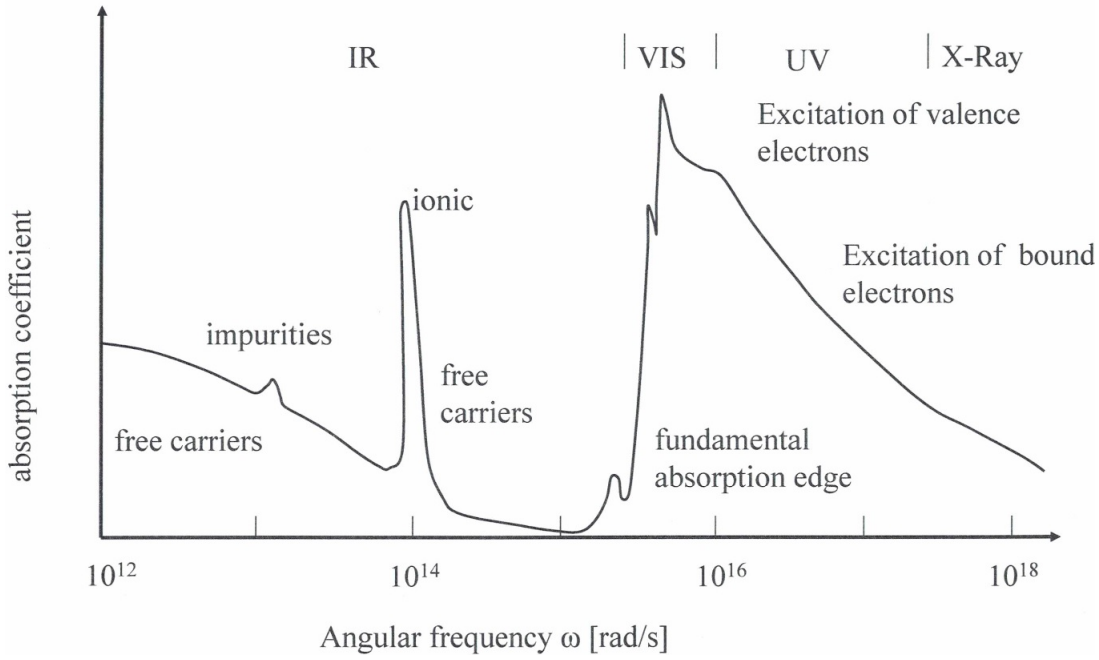


Figure 1: Absorption spectrum of a dielectric material.

Possible processes are

- phonon interactions which lead to piezoelectric behaviour and heating/cooling
- electron interactions leading to heating/cooling, interband/intraband excitation and polarization of ions and electrons.

What is a useful way to realize a transducer to transduce light into an electrical signal? A light sensing resistance R is a possible transducer between optical and electronical signals which can be done by metals, semiconductors or superconductors. However, we mainly focus on semiconducting materials.

2 Basics

2.1 Electrical properties of semiconducting materials

The first property is the *electrical resistivity* ρ which lies between metals and insulators. Some examples are

$$\mathrm{Cu}: \rho = 1{,}7 \cdot 10^{-8}\,\Omega\,\mathrm{cm} \tag{2.1}$$

$$\mathrm{SiO_2}: \rho \geq 10^{21}\,\Omega\,\mathrm{cm}. \tag{2.2}$$

There is a wide range of resistivity $10^{-5} \ldots 100\,\Omega\,\mathrm{cm}$ for semiconductors.

The second property is the conductivity $\sigma = \frac{1}{\rho}$ which can be manipulated by *doping*. We distinguish two types of conductivity called *n-type* and *p-type* which can be identified with Hall-measurements. The conductivity increases with temperature and shows a strong interaction with light. We can understand these properties via the bandstructure.

Before we will discuss the band structure we will give a quick reminder of the electrodynamic classical description of quasi-free particles:

Drude model

In the Drude model the electrons are described as an ideal gas and we introduce the conductivity as

$$\sigma = \frac{1}{\rho} = e \cdot n \cdot \mu, \tag{2.3}$$

where n ist the number of carriers or carrier concentration and μ the electron mobility. However, it is better to write

$$\sigma = \frac{j}{E} = \frac{\text{current density}}{\text{accerlerating electric field}}. \tag{2.4}$$

The current density $\boldsymbol{j}$ is defined as the product of the carrier velocity $\boldsymbol{v}$ and the charge density $\boldsymbol{j} = e \cdot n \cdot \boldsymbol{v}$. The electric field accelerates the (free) carriers and the proportionality factor is called the *mobility* $\boldsymbol{v} = \mu \boldsymbol{E}$.

How is the mobility related to the properties of the carriers? First we look at the force accelerating the carriers

$$\boldsymbol{F} = m^* \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = q \cdot \boldsymbol{E} \quad \Rightarrow \quad \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \frac{q \cdot \boldsymbol{E}}{m^*}. \tag{2.5}$$

In a solid the carrier is scattered after a short time and its velocity is set back to zero. We can describe this by a negative acceleration

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{v}{\tau} \overset{(2.5)}{=} \frac{q \cdot \boldsymbol{E}}{m^*}, \tag{2.6}$$

where τ is called scattering time. Now we can read off the mobility as

$$\mu = \frac{q \cdot \tau}{m^*} \tag{2.7}$$

and observe that it is proportional to the scattering time τ and inversly proportional to the effective mass m^*. This means the less scattering events take place and the lighter the carrier, the better is the mobility.

There are different mechanisms that determine the scattering time τ. They depend on temperature and other properties, e. g. impurities. The temperature dependence of the electrical resistance is displayed in figure 2.

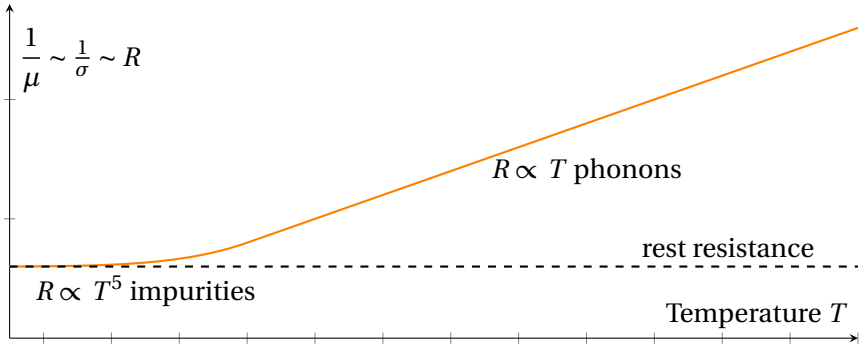


Figure 2: Schematic representation of the temperature dependence of electrical resistance for a classical metal.

We can now use MATTHIESEN's rule to write down an effective resistance

$$\rho_{\text{total}} = \rho_{\text{phonon}} + \rho_{\text{impurity}} \quad \text{or} \quad \frac{1}{\tau_{\text{total}}} = \frac{1}{\tau_{\text{phonon}}} + \frac{1}{\tau_{\text{impurity}}}. \tag{2.8}$$

For pure metals when can describe the electrical resistance with the BLOCH-GRÜNEISEN formula

$$\begin{aligned} T \leq \Theta_D \quad & \rho_{\text{BG}}(T) = \frac{124.4 \cdot C_2}{\Theta_D} \left(\frac{T}{\Theta_D} \right)^5 \\ T \gg \Theta_D \quad & \rho_{\text{BG}}(T) = \frac{C_2}{4\Theta_D} \frac{T}{\Theta_D}, \end{aligned} \tag{2.9}$$

which corresponds to the observed behaviour in figure 2. We also note that the impurity scattering is temperature independent.

2.2 Naive band diagramm

In a solid, the overlapping atomic orbits of the consisting atoms form bands.

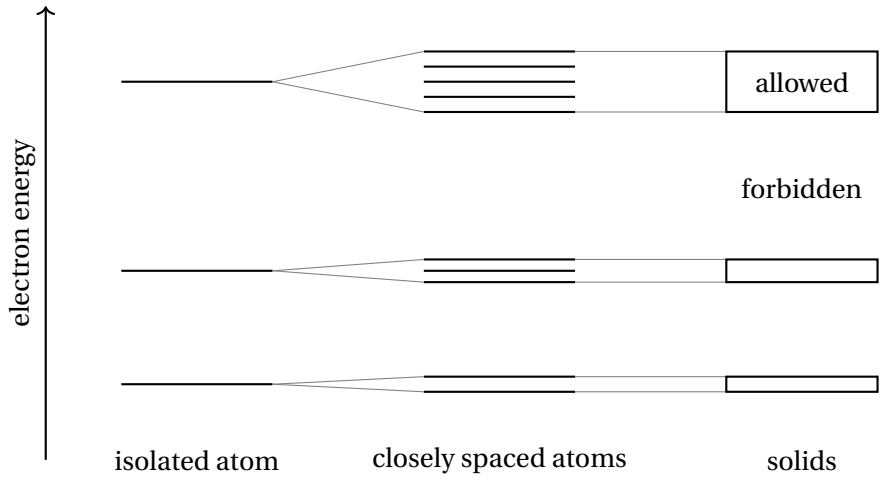


Figure 3: Diagram of single atoms, multi atom energy systems and solids with a band structure.

We only list the most important facts about the band structure:

- *Valence band*: highest occupied band, formed by valence electrons
- *Conduction band*: lowest empty band; delocalized, (fairly) freely moveable electrons which are not longer bound to one atom
- full bands do not contribute to electrical current
- partially occupied bands enable electrical current

In such a way we can identify the type of the conductor by its bandstructure:

- **Insulator**: VB completely filled, CB completely empty
  ⇒ no net electrical current conduction
- **Metals**: Number of conduction electrons is roughly the same as the number of atoms. Applied electrical fields will result in a rearrangement of the carriers in a band leading to a net electrical current conduction.
  - Valence band only partly filled (e. g. elements of groups I and III of the periodic system of elements like Na or Al)
  - the highest bands overlap (e. g. elements of group II of PSE like Zn or Mg)
- **semiconductor**: looks/behaves like an insulator; distance between valence and conduction band is fairly small (exceptions e. g. diamond)
  - Electrons can be excited (e. g. thermally) from VB into the CB ⇒ temperature specific dependence of electrical conductivity ⇒ carrier concentration is a function of temperature
  - 12 elemental semiconductors: Si, Ge (IV group of PSE), Se, Te
  - Compound semiconductor e. g. GaAs

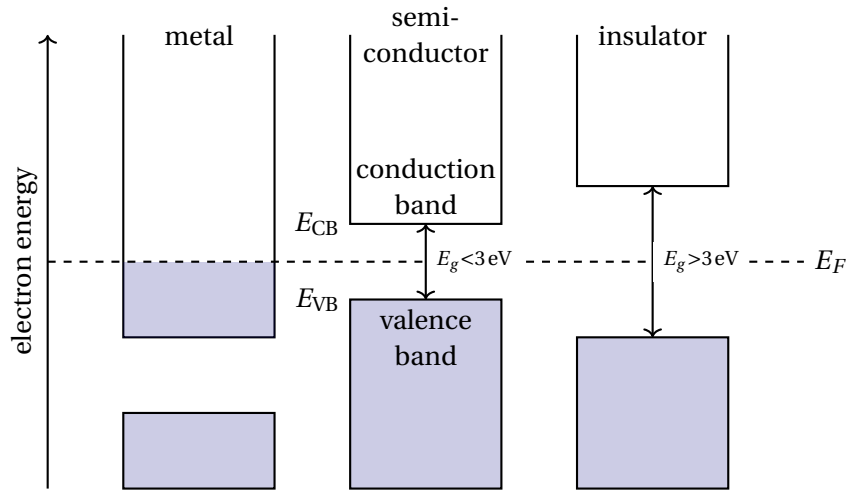


Figure 4: band structure of metals, semiconductors and insulators. In blue are shown the states occupied at $T = 0\,\mathrm{K}$.

2.3 Crystal structure

The crystal structure is responsible for many basic material properties. We define a crystal as a material with translation symmetry. We can represent the crystal in the following way:

$$\boldsymbol{r}' = \boldsymbol{r} + n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3, \tag{2.10}$$

where n_i are integers and $\boldsymbol{a}_i$ linear independent vectors. These vectors form the *crystal structure* which is constituated as *lattice* + *basis*. That means on any lattice point sits a basis group of atoms. The number of atoms (in a basis) depends on the specific crystal. It may be one (metals, nobel gases), several (compound semiconductors) or thousands (organic molecule cryxtals).

Remark: We define the primitive lattice as the lattice points which occupy the corners of the smallest (primitive) parallelipiped. This is also called the primitive unit cell.

We may use symmetry operations to classify the primitive lattice:

- Rotation by whole parts of 2π (e. g. $\frac{2\pi}{2}, \frac{2\pi}{3}, \frac{2\pi}{6}$)
- Reflection
- Inversion (reflection and rotation)

From these symmetry properties/characteristics, we can deduce 14 different point lattices, which are presented in table 1.

Notice: Many crystal structures are a consequence of the most dense package of bowls. There are two possible sequences: -ABC-ABC- or -AB-AB-AB-.

The cubic crystal symmetry shows a -ABC-ABC- structure. Examples are Si, Ge, GaAs. Notice that the cube used as the unit cell is *not* a primitive cell, e. g. it contains four lattice points for a face centred cubic crystal.

There are also cubic structures with more than one atom in a basis:

Table 1: The 7 crystal classes forming 14 Bravais lattices.

lattice	vectors	lattice angles
triclinic	$a_0 \neq b_0 \neq c_0 \neq a_0$	$\alpha \neq \beta \neq \gamma \neq \alpha$
monoclinic	$a_0 \neq b_0 \neq c_0 \neq a_0$	$\alpha = \beta = 90°, \gamma \neq 90°$
orthorhombic	$a_0 \neq b_0 \neq c_0 \neq a_0$	$\alpha = \beta = \gamma = 90°$
hexagonal	$a_0 = b_0 \neq c_0$	$\alpha = \beta = 90°, \gamma = 120°$
trigonal	$a_0 = b_0 = c_0$	$\alpha = \beta = \gamma \neq 90°$
tetragonal	$a_0 = b_0 \neq c_0$	$\alpha = \beta = \gamma = 90°$
cubic	$a_0 = b_0 = c_0$	$\alpha = \beta = \gamma = 90°$

- Diamond structure: both atoms are the same e. g. C, Si, Ge
- Zincblende structure: two different atoms e. g. GaAs

The hexagonal crystal symmetry shows a -AB-AB- structure. Examples are GaN, SiC, ZnO.

2.4 Binding type

The binding type is the reason for the special crystal structure of semiconductor materials. Silicium is a semiconductor in the IV-group of the PSE. It has four valence electrons and an electronic configuration of $1s^2 2s^2 2p^6 3s^2 3p^2$. Multielectron atoms can form hybrid orbitals, in the case of silicium sp^3-orbitals which form 4 covalent bonds. This hybridization of atomic orbitals leads to a tetraedrical shape of the sp^3 hybrid junction leading to a diamond structure.

2.5 Crystal planes, reciprocal lattice, Brillouin zone

Crystal planes play an important role in the characterization (e. g. electrical behaviour) of semiconductors. They are identified by *Miller indices*.

Construction of Miller indices:

1. Find intercepts of plane with the three basic axes in terms of the lattice constant
2. Take the reciprocals of these numbers
3. reduce them to smallest three integers with same ratio

Notice that the indices are enclosed in *brackets* (h k l) as integers. Negative indices (when the plane intercepts on the negative side) are indicated by a bar above the number $(0\,\bar{1}\,1)$. Planes of equivalent symmetry (like $(100), (010)$ or $(00\bar{1})$ in cubic crystals) are enclosed in *curly brackets* $\{100\}$.

A direction in the crystal is defined by a vector which is in the units of the three base vectors. The direction is indicated by Miller indices, but in square brackets [010]. A set of symmetrically equivalent directions is given by angle brackets ⟨010⟩. In cubic crystals, directions are perpendicular to the planes with the same Miller indices!

Reciprocal lattice

For a better interpretation and numerical calculations it is helpful to transfer/transform the original lattice in the so called reciprocal lattice:

$$\boldsymbol{A} = 2\pi \frac{\boldsymbol{b}\times\boldsymbol{c}}{\boldsymbol{a}\cdot\boldsymbol{b}\times\boldsymbol{c}}, \qquad \boldsymbol{B} = 2\pi \frac{\boldsymbol{c}\times\boldsymbol{a}}{\boldsymbol{a}\cdot\boldsymbol{b}\times\boldsymbol{c}}, \qquad \boldsymbol{C} = 2\pi \frac{\boldsymbol{a}\times\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{b}\times\boldsymbol{c}}, \tag{2.11}$$

where $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are the reciprocal lattice vectors. This reciprocal lattice may be regarded as the *Fourier transform* of the original lattice

$$\boldsymbol{G} = h\boldsymbol{A} + k\boldsymbol{B} + l\boldsymbol{C}, \tag{2.12}$$

where $\boldsymbol{G}$ is a translation vector in the reciprocal space. h, k, l are the Laue indices (or Miller indices) and are of integer type.

Remark: The Brillouin zone is the Wigner-Seitz cell of the reciprocal lattice. We can construct it by forming a midplane connecting straight planes from one lattice point of the reciprocal lattice to all others.

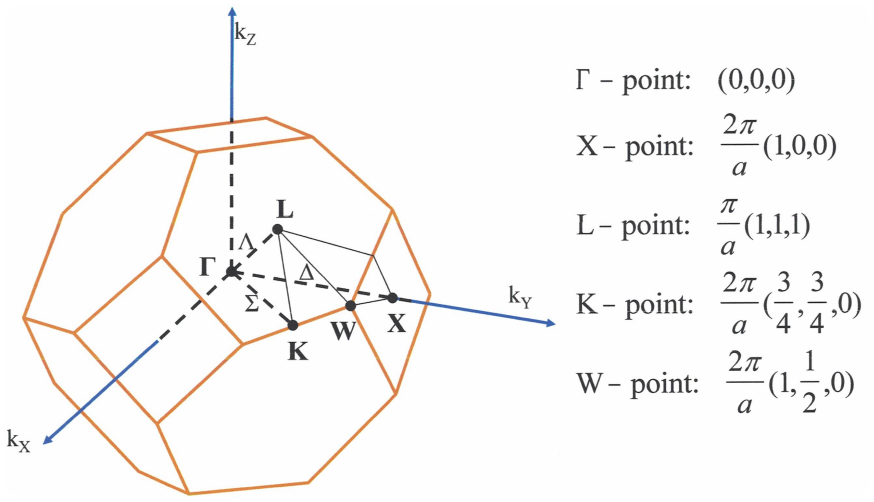


Figure 5: Brillouin zone and high symmetry points for the fcc lattice which forms the Bravais lattice for diamond and zinc-blende structures.

2.6 Dispersion relation

For a better understanding of their movement, electrons in a solid must be described by their qunatum mechanical wave function Ψ with the two perameters *energy* E and *momentum* $\boldsymbol{p}$. For a free moving electron we find a dispersion relation $E(\boldsymbol{k})$

$$E(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2). \tag{2.13}$$

However, in a solid, an electron is not completely free. It interacts with the atoms of the crystal which form a periodic potential

$$V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R}) \quad \text{where} \quad \boldsymbol{R} \text{ lattice vector.} \tag{2.14}$$

The particle wave function has to fulfill the respective Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{r})\right]\Psi_k(\boldsymbol{r}) = E_k \Psi_k(\boldsymbol{r}) \tag{2.15}$$

with periodic boundary conditions, due to the periodic structure of the crystal.

One possible solution is a standing wave which may be

$$\Psi_+(x) \sim \cos\left(\frac{\pi}{a}x\right) \qquad \Psi_-(x) \sim \sin\left(\frac{\pi}{a}x\right). \tag{2.16}$$

At $k = \pm n\frac{\pi}{a}$ an energy separation (band gap) occurs. The energy separation can be explained by the formation of energy bands from coupling of atomic levels or by standing waves in a crystal due to reflections for $\boldsymbol{K} = n\boldsymbol{G}$.

Effective mass

Within the band, the electrons can move as *quasi-free* particles. This concept is even closer to a really free electron, if all interactions with the lattice potential is taken into account by ascribing an effective mass m^* to this particle. The effective mass is defined as

$$m^* = \hbar^2\left(\frac{\mathrm{d}^2 E}{\mathrm{d}k^2}\right)^{-1}, \tag{2.17}$$

i. e. the effective mass is inversly proportional to the curvature of the dispersion relation. For small k (in the centre of the first Brillouin zone) the dispersion relation reads

$$E(k) = \frac{\hbar^2 k^2}{2m^*} \tag{2.18}$$

which is similar as for free electrons (2.13). In this picture the effective mass describes the influence of the lattice on the electron movement. Typical values are

$$m^* = 0.57\, m_e(\mathrm{Ge}) \qquad m^* = 0.063\, m_e(\mathrm{GaAs}). \tag{2.19}$$

The velocity of the electron depends on the effective mass in a manner similar to the way it does in a vacuum. The electron velocity - strictly, the group velocity of the wave packet for electron motion

$$v = -\frac{1}{\frac{h}{\pi}}\frac{\mathrm{d}E}{\mathrm{d}k} \tag{2.20}$$

is also given by the *first derivative* with respect to $\boldsymbol{k}$. These relations are summarized in figure 6.

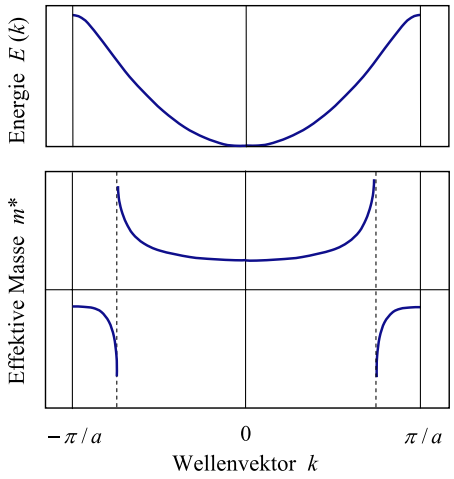


Figure 6: Dispersion relation $E(k)$ and effective mass m^*. We observe that the velocity (derivative of the energy) is zero at the band edge and its maximum is in the centre part of the energy band (poles of the mass). On the lower-energy side of the energy band the effective mass is positive and negative on the higher-energy side.

Movement of carriers in electrical fields: A quasi-free electron in a band (in an electric field) takes up some energy and thus changes its wave vector $\boldsymbol{k}$ by $\Delta\boldsymbol{k}$.

1. *Full band*: If all places are occupied (in a band all electrons change by $\Delta\boldsymbol{k}$), that means due to the *back scattering* at the edge of the Brillouin zone no macroscopic change does occur ⇒ no conduction.
2. *Partly filled band*: Centre of gravity of the electrons changes in an electric field. This is equivalent to some electrical conduction.

Excitation of an electron: Electrons may be excited from valence into the conduction band. In such a way the electron leaves a *vacancy* in the valence band. In an applied electric field the remaining electron ensemble changes again its momentum $\boldsymbol{p}$ or $\boldsymbol{k}$. However, this is equivalent to the movement of the vacancy in the *opposite* direction. Therefore the vacancy behaves like a positively charged particle a so called *hole* with positive mass. Its also called a positively charge quasi-particle.

2.7 Real band structures

We can distinguish between two types of semiconductors.

1. *Direct semiconductors*: The minimum of the conduction band and the maximum of the valence band are in the centre of the first BZ at $k = 0$ (Γ-point). They are typically isotrope, i. e. with spherical energy surface. They are compound semiconductors like GaAs, InP, ZnSe. They show strong interaction with light
2. *Indirect semiconductors*: The minimum of the conduction band is not at $k = 0$. Examples are Silicium and Germanium. They only show weak interaction with light.

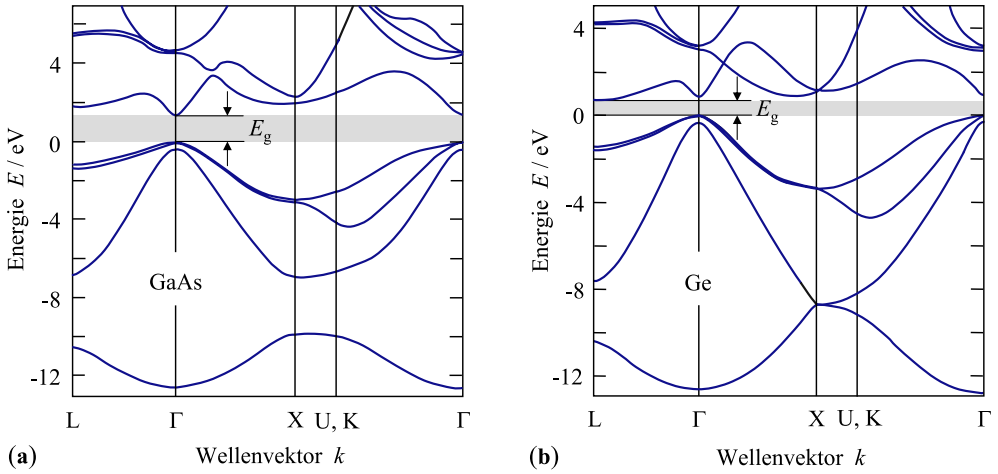


Figure 7: a.) Direct semiconductor: Band structure of GaAs. Minimum and Maximum of conduction and valence band oppose each other at the Γ-point.
b.) Indirect semiconductor: Band structure of Germanium. The minimum of the conduction band is not at the Γ-point.

Notice that the valence band structure is more complicated then the conduction band structure. Originally it degenerates at $k = 0$. The reason were the p-orbitals of the atoms. By spin-orbit interaction, one valence band is split off even at $k = 0$ (*split-off band*). The *two remaining* top-most bands have different curvature resulting in a heavy hole and a light hole band.

In a 3D crystal the dispersion relation may be different for all three directions $\boldsymbol{k}_x, \boldsymbol{k}_y, \boldsymbol{k}_z$. This leads to a more complex band structure.

2.8 Interaction with light

The excitation of carriers from the valence to the conduction band may be initiated by light. When the photon energy is absorbed the energy may be transferred to an electron which thus is excited from the VB to the CB. The excitation (or recombination) process can be regarded as a *collision process* of two particles.

There are three basic processes:

1. *Simple absorption* in a direct semiconductor.
2. *Spontaneous emission* of light by recombination of an excited electron characterized by an absorption coefficient α with $I = I_0 \exp(-\alpha d)$ (d-thickness).
3. *Stimulated emission* of a photon. Incoming photons of suitable energy trigger an excited electron to recombine. Then a second photon is created which is totally identical to the incoming photon.

For the collision of photon and excited electron two laws of conservation hold:

1. Conservation of energy: $E = \hbar\omega = \frac{kc}{\lambda}$

2. Conservation of momentum $\boldsymbol{p} = \hbar\boldsymbol{k} = \frac{\hbar}{\lambda} = \frac{E}{c}$.

For a reasonable energy (in the same range as the band gap in a semiconductor i. e. 1 eV) the photon's momentum is very small $k \approx 10^7 \frac{1}{\mathrm{m}}$. On the other hand, the typical momentum of the electron in a crystal is of the order of a reciprocal lattice vector:

$$\frac{\pi}{a} \rightarrow k \leq 6 \cdot 10^9 \frac{1}{\mathrm{m}} \tag{2.21}$$

which is much larger than the momentum of the photons

$$|\boldsymbol{k}|_{\text{Phot.}} \ll |\boldsymbol{k}|_{\text{Electr.}} \tag{2.22}$$

Thus, the photons momentum to be transferred to the electron is close to zero which corresponds to a *vertical transition* in the E-k-diagram.

Direct semiconductor

That means in a direct semiconductor an excited electron which falls back from its excited state at the conducting band minimum at $k = 0$ to the valence band maximum at $k = 0$ will create a photon of energy of about E_g.

Indirect semiconductor An electron in the conduction band at $k \neq 0$ *cannot* easily fall down to the valence band because the two electron states differ not only in energy, but also in momentum. The missing momentum must be put in by a third particle typically a *phonon*. The same rules hold for the inverse process.

This means we have a *three particle process*. An electron and phonon create a photon. The probability of such a process is much less than that of two particle processes. The resulting photon energy is then $E_{\text{Photon}} \leq E_g$ because the phonon also carries energy. Therefore the probability of radiative transition in an indirect semiconductor is very low. This explains the problems with application of Si and Ge as optoelectronic devices.

Last but not least and excited electron may *relax* into the VB by *non-irradiative processes* which may be multi phonon processes. This effect can be used e. g. for *laser ablation*.

Relaxation time τ

The recombination or relaxation is described by the relaxation time τ which tells us how long a particle stays in the excited state before it spontaneously relaxes or recombines

$$\frac{1}{\tau} = \frac{1}{\tau_{\text{radiative}}} + \frac{1}{\tau_{\text{non-radiative}}}. \tag{2.23}$$

Notice, compound semiconductors have a direct band structure. They are well suited for optoelectronic applications.

2.9 Carrier concentration

We have seen we must discuss the probability of excitation or relaxation. The basic question is: How many carriers are in the conduction and/or valence band at which temperature? This is controlled by the laws of statistics.

We define n as the concentration of electrons in the conduction band and correspondingly p the concentration of holes in the valence band.

Pure materials without any impurities, defects or dopings are called *intrinsic semiconductors* which only have thermally activated electrons.

Density of states

We can answer our basic question by introducing the density of states. However, we reformulate or question as: How large is the number of states which can be occupied and are these states occupied?

We start with the first point, the *number of states*. Every atom which forms the crystal delivers one state (which may be occupied by two electrons having opposite spin). The number of these states is equal to the number of solutions of the respective SCHRÖDINGER equation which are characterized by their momentum $\boldsymbol{k}$.

As we remember from solid state physics we assume periodic boundary conditions (periodicity L) in all three dimensions. Then the *eigenvectors* are

$$\boldsymbol{k} = (2n_x\frac{\pi}{L}; 2n_y\frac{\pi}{L}; 2n_z\frac{\pi}{L}) \qquad n \text{ - integer.} \tag{2.24}$$

All these states are in the first Brillouin zone. The *total number of states per volume* (up ton an energy E) is given by

$$N = \frac{1}{3\pi^2}\left(\frac{2mE}{\hbar^2}\right)^{3/2}. \tag{2.25}$$

Why? We must count the number of states in a sphere in k-space $V_E = \frac{4}{3}\pi k_E^3$ where we remember the energy of free electrons (2.13) and the volume $V = \left(\frac{2\pi}{L}\right)^3$. From this we can obtain the *density of states* $D(E)$ as

$$D(E) = \frac{\mathrm{d}N}{\mathrm{d}E} = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E} \tag{2.26}$$

$$\text{or } D_C(E)\,\mathrm{d}E = \frac{1}{2\pi^2}\left(\frac{2m_e^*}{\hbar^2}\right)^{3/2}\sqrt{E-E_C}\,\mathrm{d}E \quad \text{conduction band}$$

$$D_V(E)\,\mathrm{d}E = \frac{1}{2\pi^2}\left(\frac{2m_p^*}{\hbar^2}\right)^{3/2}\sqrt{E_V-E}\,\mathrm{d}E \quad \text{valence band}$$

The density of states describes how many states we can find in the energy interval between E and $E+\mathrm{d}E$.

Light interaction leads to absorption only if we find electrons in a state we can absorb light in such a way. The density of states give us the probability of absorption. The higher the energy, the larger the density of states, the higher the probability of absorption.

Fermi distribution

Now we want to answer the question how these states are occupied? The *probability* whether an available state is occupied is governed by the BOLTZMANN probability

$$p(E) \sim \exp\left(-\frac{E}{k_B T}\right). \tag{2.27}$$

Electrons are FERMI particles, i. e. they are governed by the Pauli principle: Once a state is occupied it cannot be occupied by another particle with exactly the same quantum numbers. Thus the electrons are distributed over the available states according to the FERMI *distribution*

$$f(E) = \frac{1}{1+\exp\left(\frac{E-E_F}{k_B T}\right)}, \qquad E_F \text{ - Fermi energy.} \tag{2.28}$$

The fermi energy E_F is the chemical potential of electrons gas in a semiconductor. It is the specific energy of the electron ensemble for introduction to or removal of one particle of the ensemble. For $T = 0\,\mathrm{K}$ the fermi energy is

$$E_F \sim \frac{E_C + E_V}{2}. \tag{2.29}$$

At $T = 0\,\mathrm{K}$ all states up to E_F are occupied and all others (higher energy) are empty. This means (in intrinsic semiconductors) the valence band is completely filled and the conduction band is completely empty.

Now we want to find the number of carriers in the conduction band at $T > 0$:

$$\begin{aligned} n(T) &= \int_0^\infty D(E)\cdot f(E)\,\mathrm{d}E = \int_{E_F}^\infty \frac{1}{2\pi^2}\left(\frac{2m_e^*}{\hbar^2}\right)^{3/2}\sqrt{E-E_C}\,\frac{1}{1+\exp\left(\frac{E-E_F}{k_B T}\right)}\,\mathrm{d}E \\ &= 2\left(\frac{m_e^* k_B T}{2\pi\hbar^2}\right)^{3/2}\exp\left(\frac{E_F - E_C}{k_B T}\right) = N_C \exp\left(-\frac{E_G}{2k_B T}\right), \end{aligned} \tag{2.30}$$

with $E_g = 2(E_C - E_F)$. We call N_C the *effective density of states* of the conduction band. We observe that the carrier concentration increases drastically with rising T and thus the conductivity increases.

The number of holes in the valence band can be deduced in the same way

$$p(T) = N_V \exp\left(\frac{E_V - E_F}{k_B T}\right). \tag{2.31}$$

From both equations we can deduce the so-called *Mass action law*

$$n \cdot p = 4\left(\frac{k_B T}{2\pi\hbar^2}\right)^3 (m_e^* m_p^*)^{3/2}\exp\left(-\frac{E_g}{k_B T}\right). \tag{2.32}$$

Notice that this product $n \cdot p$ does not depend on the FERMI energy E_F. By increasing E_F we can increase n by simultaneously decreasing p.

Remark: For intrinsic semiconductors $n = p$ which leads to the $\sqrt{\text{mass action law}}$

$$n_i = \sqrt{n \cdot p} = 2\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2} (m_e^* m_p^*)^{3/4} \exp\left(-\frac{E_g}{2k_B T}\right). \tag{2.33}$$

With $n = p$ we can assume

$$E_F = \frac{1}{2}(E_C + E_V) + \frac{3}{4} k_B T \ln\left(\frac{m_p^*}{m_e^*}\right) \tag{2.34}$$

which is the temperature dependence of the FERMI level. For $T = 0\,\text{K}$ the FERMI-energy E_F is in the middle of the band gap E_g. For increasing T the FERMI energy is closer to the conduction band.

Conductivity

Analogous to (2.3) we can express the conductivity as

$$\begin{aligned} \sigma &= n \cdot e \cdot \mu_e + p \cdot e \cdot \mu_p \\ &= \underbrace{e(\mu_e + \mu_p) \cdot (m_e^* \cdot m_p^*)^{3/4} 2\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2}}_{B(T)} \exp\left(-\frac{E_g}{2k_B T}\right) \\ \Rightarrow \ln\sigma &= \ln B(T) - \frac{1}{2}\frac{E_g}{k_B T}. \end{aligned} \tag{2.35}$$

This is called the exponential law of conductivity. The conductivity goes exponential with T, i. e. $\ln\sigma \sim \frac{1}{T}$. It also depends on the energy gap E_g and linearly on the mobility μ. The carrier mobility of electrons differs from the carrier mobility of holes $\mu_e \gg \mu_p$. At room temperature we find

$$\begin{aligned} \sigma_{\text{metal}} &\gg \sigma_{\text{semiconductor}} \\ n_{\text{metal}} \approx 10^{22}\,\frac{1}{\text{cm}^3} &\gg n_{\text{intrinsic sc}} \approx 10^{13}\,\frac{1}{\text{cm}^3}. \end{aligned} \tag{2.36}$$

2.10 Doping

An intrinsic semiconductor has a low carrier density at room temperature. How can we manipulate the carrier density n? The answer is to simply induce extra carriers which can be done by doping. Doping can be achieved by introducing *impurities* which provide extra carriers (electrons in CB or holes in VB).

In elemental semiconductors these may be elements of the group V (N,P,As,Sb,Bi) replacing group IV atoms in the lattice. They act as a *donor* providing extra electrons. On the other hand the doping materials may be elemetns of the group III (B,Al,Ga,In,Tl) which also replace group IV elements but act as an *acceptor*. They can accept and electron and create a hole.

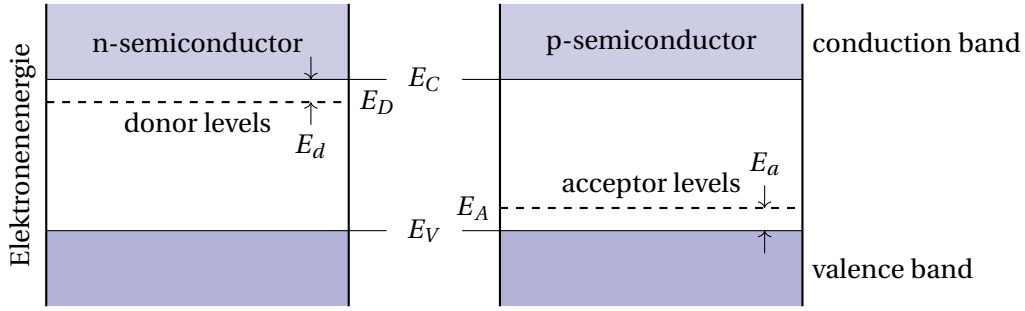


Figure 8: doping of a semiconductor - simple band structure model

Such a donor can easily release its extra electron to the conduction band. However, some activation or ionization energy is necessary. In the normal situation we have a low concentration of donor atoms

$$N_D \ll N_{\text{Crystal}}. \tag{2.37}$$

Donor atoms form *localized states* (like a single atom). The ionization energy of such a localized state can be calculated in the same way as the binding energy of an electron to a proton $E \sim \frac{E_0}{n^2}$. The donor introduces a new energy level slightly below the conduction band edge E_C with $\Delta E = E_D$.

During the doping the FERMI level E_F changes its position. At moderate doping concentrations we find

$$E_F(T) = \frac{1}{2}(E_D + E_C) + \frac{1}{2} k_B T \ln\left(\frac{N_D}{2n_0}\right) \tag{2.38}$$

$$\text{with } n = \frac{1}{2} N_D \exp\left(-\frac{E_F - E_D}{k_B T}\right) \quad \text{for } e^- \text{ from donor levels}$$

$$n = \frac{1}{2} N_D \exp\left(-\frac{E_C - E_F}{k_B T}\right) \quad \text{for } e^- \text{ from valence band.}$$

Discussion: At $T = 0$ the Fermi level is at the donor level. This means all electrons are still bound to their donors. Increasing the temperature will rise the electron density rather quickly (n-doped semiconductor) and the fermi level rises half-way between the donor level and the conduction band edge. A further increase of temperature decreases the fermi level

E_F below the donor level E_D and in the limit of high temperatures to the middle between conduction band and valence band.

In a homogeneous doped semiconductor is the concentration of electrons in the CB plus the concentration of ionized acceptor levels equal to the concentration of holes in the VB plus the concentration of ionized donor levels:

$$n + N_A^- = p + N_D^+. \tag{2.39}$$

This is referred to as the *new neutrality condition* for doped semiconductors. Remember in the undoped case we had $n = p$. The concentration of ionized donors is

$$N_D^+ = N_D - N_D^0 = \frac{N_D}{2\exp\left(\frac{E_F - E_D}{k_B T} + 1\right)} \quad \text{with } N_D^0 = \frac{N_D}{\frac{1}{2}\exp\left(\frac{E_D - E_F}{k_B T} + 1\right)}. \tag{2.40}$$

The only difference of the statistic description is the factor $\frac{1}{2}$. The reason is, that only *one* electron is localized in *one* donor level. In the conduction band two electrons (with opposite spin) can have t same energy level.

Similarly acceptors can be described by their respective energy E_A but the situation is more complicate because they owe to the *larger effective mass*, hence a smaller *binding radius* and a more complex valence band structure. Consequences of the more complex nature are

- E_A depends strongly on the respective acceptor element
- E_A is typically much larger than the donor ionization energy $E_A > E_D$ which is shown in table 2.

Table 2: Donor and acceptor ionization energy of various elements and dopants

	element	dopant	ionization energy [eV]
E_D	Ge	P	0.012
		As	0.013
		Sb	0.0098
E_D	Si	P	0.044
		As	0.049
		Sb	0.039
E_A	Ge	Al	0.01
		Ga	0.011
		In	0.011
E_A	Si	Al	0.057
		Ga	0.065
		In	0.16

Temperature dependence of a doped semiconductor

At low temperatures the donors get more and more ionized with increasing temperature. When all donors are ionized the carrier concentration saturates up to those temperatures. At higher temperatures thermal excitation of carriers from the valence band into the conduction band starts to dominate. The temperature dependence is shown in figure 9.

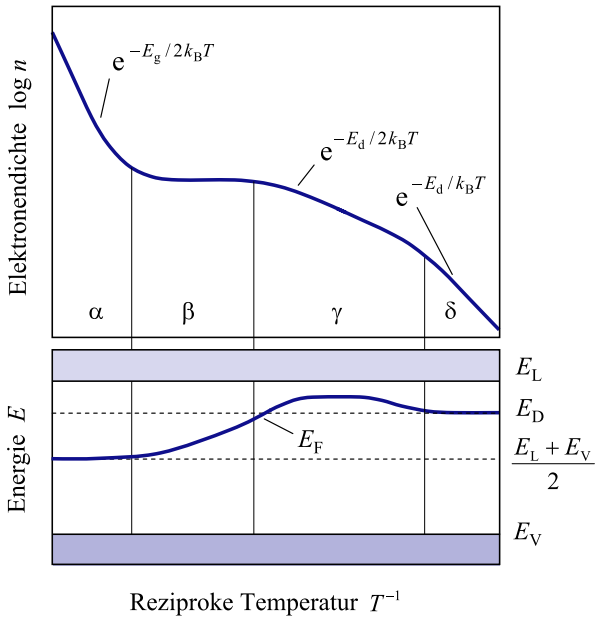
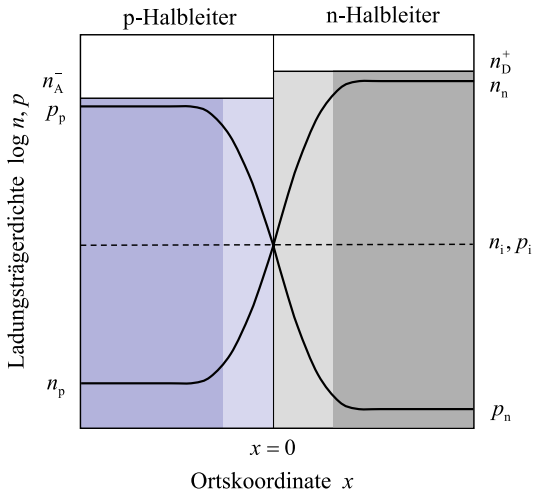


Figure 9: Electron density in the CB of a n-doped semiconductor (top) and fermi-level as a function of reciprocal temperature (bottom).

2.11 pn-junction

If we dope the semiconductor material with different doping-elements we can create a so called pn-junction (we discuss the ideal sharp doping profile).

We change the carrier concentration in the contact region between n and p-doped area. In this contact region we can observe different microscopic processes:



Due to the strong gradient carriers diffuse in the other doping region:

- electrons from n-doped region (n_n) to p-doped region (n_p)
- holes from p-doped region (p_p) to n-doped region (p_n)
- with $N_A^- = p_p + p_n$ and $N_D^+ = n_n + n_p$

We call p_p and n_n majority carriers and p_n and n_p minority carriers respectively. We can find both in the same region. The diffusion of majority carriers generates a *diffusion current* I_{Diff} across the boundary with

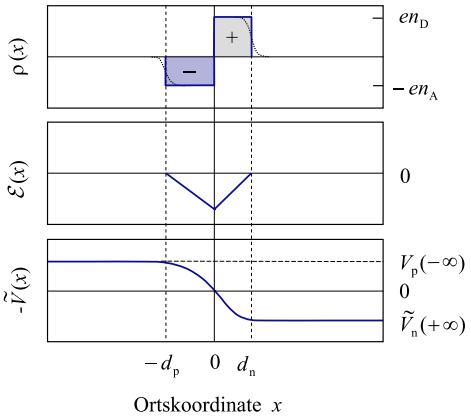
$$V_{\text{Diff}} = \frac{k_B T}{e} \ln \frac{N_D}{N_A}. \tag{2.41}$$

The minority carriers form a *space charged region* which generates an electric field $\boldsymbol{E}$.

The electrical field $\boldsymbol{E}$ creates a force resulting in a field current I_{field}. In the case of a thermal equilibrium we found

$$I_{\text{Diff}} = I_{\text{field}} \tag{2.42}$$

which is called *mass action law*. In such a way a space-charge region or depletion layer is formed. This depletion layer acts as a barrier between the two doped regions.



The direct contact area is *highly resistive* (due to recombination). This means the applied voltage (in an electronic device) falls *across the contact* not across the doped volume. The induced electrical field limiting the electron and hole diffusion. The *built-in electrical field* is the maximum at the pn-boundary and goes to zero at the edges of the depletion region.

The electric potential of the n-side region is higher than that of the p-side region by an amount v_D (built-in potential or diffusion potential) is the result of the *band-bending* v_{D_n} in the n-type region and v_{D_p} in the p-type region. The current I across such a junction is given by

$$j(v) = \left(\frac{e \cdot D_p}{L_p} p_n + \frac{e \cdot D_n}{L_n} n_p \right) \cdot \left[\exp\left(\frac{e \cdot V}{k_B T} - 1 \right) \right] \tag{2.43}$$

where L_n and L_p are the diffusion length in the respective regions and D_n, D_p the diffusion constant or coefficient.

We can distinguish two different conditions (operation mode):

- *reverse biasing*
- *forward biasing* (conduction direction)

Forward bias mode (conduction direction)

The positive pole is at the p-doped region and the negative pole at the n-doped region. We discuss this situation in the picture of carrier concentration

The positive (negative) pole attracts the electrons especially the majority carries in the n-doped (p-doped) region. Thus the thickness of the depletion layer decreases leading to conduction.

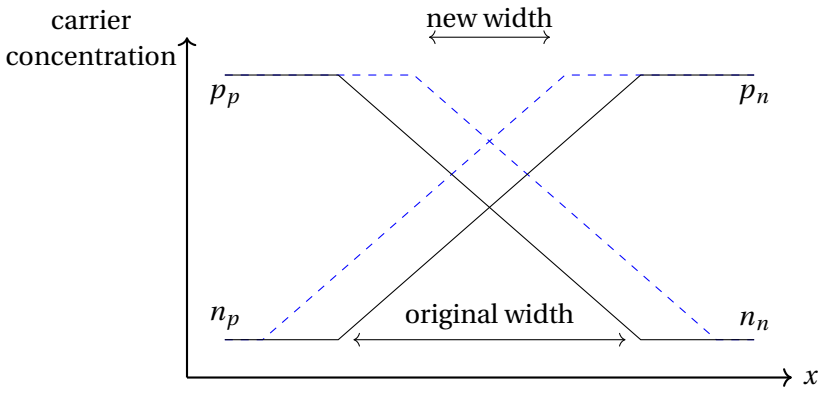


Figure 10: Change of the depletion layer width in forward bias mode.

Reverse bias mode

The negative pole is at the p-doped region and attracts the holes (majority carriers) in the p-doped region. This results in an increase of the depletion layer.

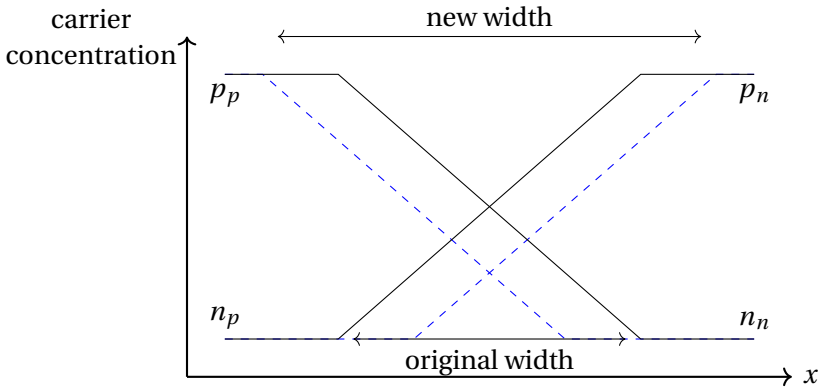


Figure 11: Change of the depletion layer width in forward bias mode.

The thickness of the depletion layer can be written as

$$d \approx \sqrt{\frac{2\varepsilon_r (V_D - V)}{e}} \cdot \sqrt{\frac{1}{N_A} + \frac{1}{N_D}}. \tag{2.44}$$

If we use values such as $V = 0\,\mathrm{V}, V_D = 700\,\mathrm{mV}, N_{A,D} = 10^{16}\,\frac{1}{\mathrm{cm}^3}$ we obtain a thickness of $d \approx 0{,}43\,\mu\mathrm{m}$ (for Si).

We also want to discuss the current-voltage characteristic of a pn-junction. We note that the characteristic is highly nonlinear and follows an exponential law:

$$I_D = I_S \exp\left(\frac{V_D}{V_T} - 1\right) \quad \text{with} \quad V_T = \frac{k_B T}{e}. \tag{2.45}$$

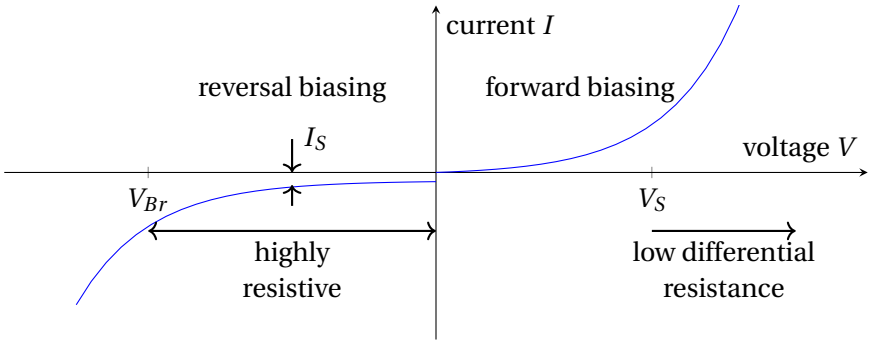


Figure 12: Current-voltage characteristic of a pn-junction.

We also want to present some experimental results:

- *forward*: threshold voltage: $V_S(\text{Si}) = 0{,}7\,\text{V}$, $V_S(\text{Ge}) = 0{,}3\,\text{V}$
  $V_S = f(T)$ depends on temperature: $\frac{\partial V_S}{\partial T} \approx -2\,\frac{\text{mV}}{\text{K}}$
  this gives rise to possible temperature sensors.
- *reverse*: leakage current $I_S(\text{Si}) < 10^{-6}\,\text{A}$, $I_S(\text{Ge}) < 10^{-5}\,\text{A}$

If the voltage V gets bigger than the breakthrough voltage V_{Br}, the pn-junction is destroyed by a large current density.

band structure description

How can we describe the pn-junction via a band structure?

We prepare a pn-junction by connecting two regions of p and n-doped semiconductors together creating a sharp doping profile as in figure 13. In the disconnected case the fermi levels at $T = 0$ lie between the doping levels and the band edge. If we join the different regions the fermi levels must align (*first rule of junction description*)[1].

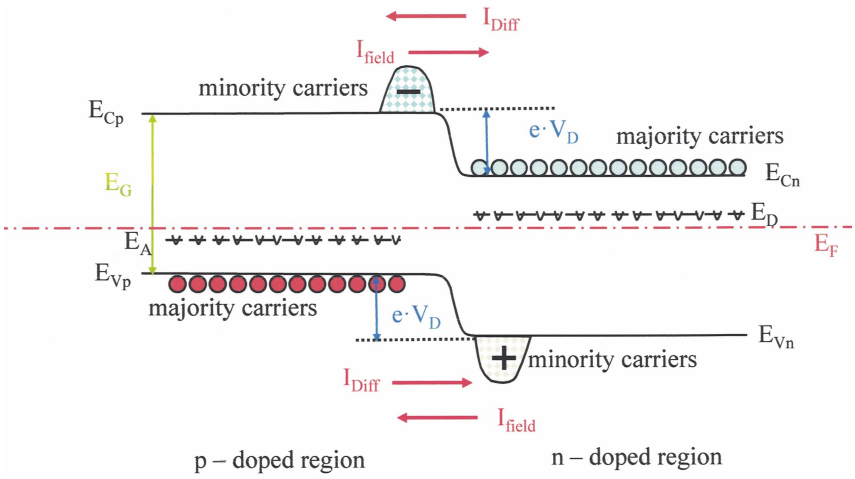


Figure 13: band structure of a pn-junction. If both regions are in contact the fermi levels will align which leads to band banding

[1] If the two fermi levels do not coincide, the electrons or holes diffuse until the levels are aligned.

The *built-in potential* V_D is given by the difference between the energies of the two fermi levels of the n- and p-doped regions or between the conduction/valence band edges in both regions

$$-qV_D = E_{C_n} - E_{C_p} = E_{v_n} - E_{v_p} = E_{F_n} - E_{F_p}. \tag{2.46}$$

We can also write down the carrier density of the majority and minority electrons in the respective regions

$$n_n = N_D \exp\left(\frac{E_{F_n} - E_{C_n}}{k_B T}\right) \tag{2.47}$$

$$n_p = N_o \exp\left(\frac{E_{F_p} - E_{C_p}}{k_B T}\right). \tag{2.48}$$

In thermal equilibrium we will find

$$\frac{n_p}{n_n} = \exp\left(\frac{E_{C_n} - E_{C_p}}{k_B T}\right) = \frac{p_n}{p_p}. \tag{2.49}$$

The built-in potential can then be expressed by the carrier densities of holes or electrons

$$V_D = \frac{k_B T}{q} \ln\left(\frac{n_n}{n_p}\right) = \frac{k_B T}{q} \ln\left(\frac{p_p}{p_n}\right) \tag{2.50}$$

which means that we can manipulate V_D via doping.

However, $e\,V_D$ is always smaller than the band gap energy E_g. We can increase V_D by increasing the donor and/or acceptor concentration. Furthermore we can change the width of the depletion layer (remember: $d \sim \sqrt{V_D + V}$).

We may also manipulate the capacitance of the pn-junction

$$C = \sqrt{\frac{\varepsilon_s \varepsilon_v \cdot q \cdot N_D \cdot N_A}{2(N_D + N_A)} (V_D + V)}, \tag{2.51}$$

which depends on the donor/acceptor concentration and the potential difference.

Forward bias conditions

For forward bias we need a positive biasing on the p-side region and a negative biasing on the n-side region respectively. Then the built-in potential barrier can be reduced from $q\,V_D$ to $q(V_D - V)$. Then current starts to flow. We define the *saturation current* as the sum of injected electron and hole currents.

The result is an electron diffusion from the n-side to the p-side region and a corresponding hole diffusion from the p-side to the n-side region. The density of minority carriers increases *exponentially* with increasing voltage. This process is called *injection*, where we increase the minority carrier density to a level above that at thermal equilibrium. This ist a fundamental process of optoelectronic devices.

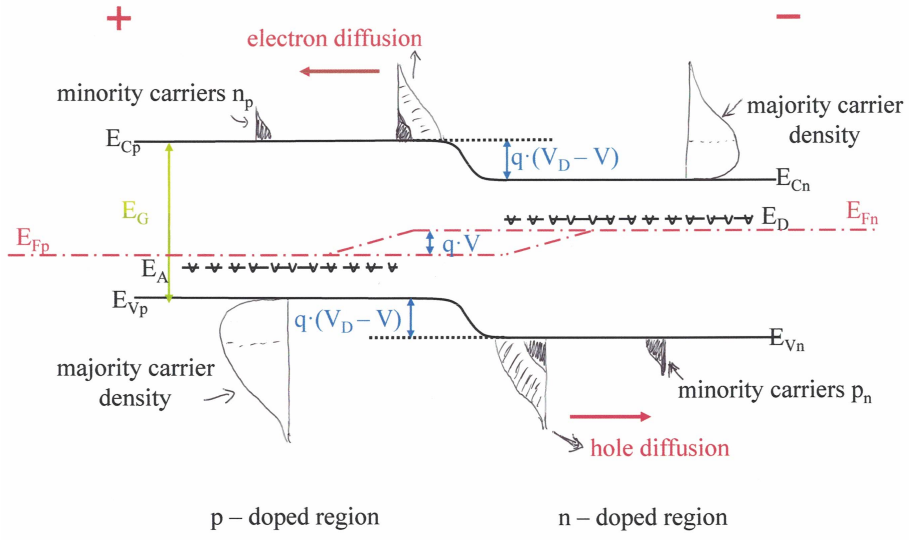


Figure 14: pn junction with external voltage V in forward bias mode.

The resulting forward current I_F is displayed in 15. We observe different recombination currents due to:

- Defects I_{NR} generation-recombination current (low biasing)
- surface I_{Sur} surface recombination current (low biasing)

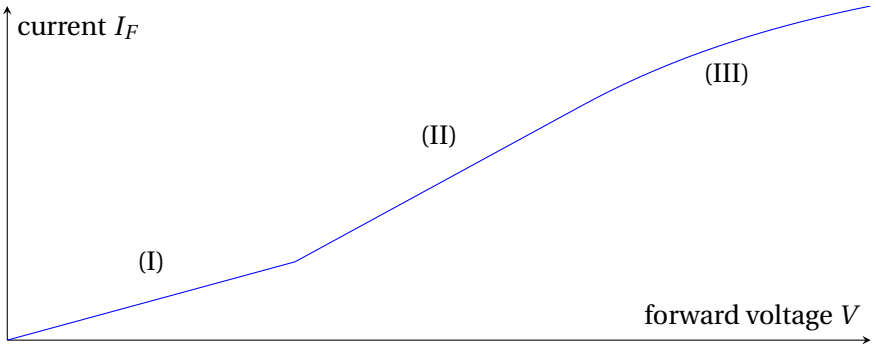


Figure 15: Forward current I_F in forward bias mode. In the first part the surface recombination current I_{Sur} is dominant for III-V semiconductors whereas the generation-recombination current I_{NR} is dominant in silica. In the second part diffusion current dominantes. In the last part effects of Joules heating take place saturating the current due to additional resistivity.

Reverse bias conditions

For reverse bias we need a negative biasing on the p-side region and a positive biasing on the n-side region respectively. Then the built-in potential barrier can be increased from $q\,V_D$ to $q(V_D + V)$. The result is called *extraction*, where the minority carrier density is decreased below that at thermal equilibrium. The reverse current I_{Sat} is nearly constant, but I_{NR} and I_{Sur} increase with the applied bias voltage.

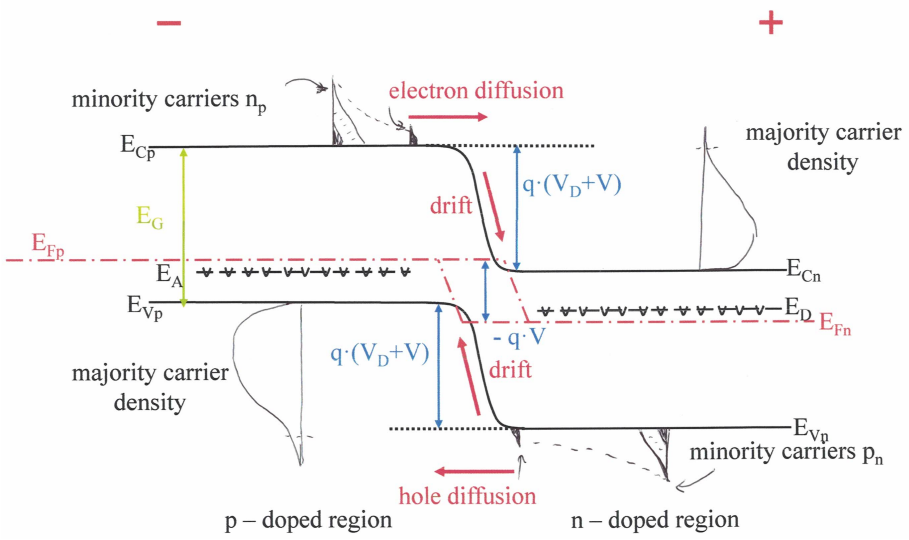


Figure 16: pn junction with external voltage V in reverse bias mode.

Breakthrough/Breakdown voltage

There exists a point in the I/V curve under reverse biasing at which the reverse current *abruptly* increases. This is called breakdown which is caused by *avalanche breakdown* or *Zener breakdown*:

a.) Avalanche breakdown

- In reverse bias mode a high electrical field with the depletion layer is created.
- Electrons are diffusing from the outside of the depletion layer and get accelerated with the depletion layer.
- They collide with lattice atoms and create new electrons and holes by *ionization*.
- This process amplifies itself with the new electrons thus increasing the number of electrons quickly (electron avalanche).

b.) Zener breakdown

- pn-junctions formed with semiconductors having a *high* impurity concentration leads to small depletion layers $d \sim \frac{1}{N_D} \sim \frac{1}{N_A}$
- In the thin depletion layer quantum mechanical tunnelling occurs. The tunnelling electrons increase the reverse current quickly.

For pn-junctions with *low/high* impurity concentrations *Avalanche/Zener* breakdown occurs.

2.12 Heterojunctions

A heterojunction is a junction formed by connecting two different materials. We distinguish two kinds of heterojunctions related to semiconductors:

- junctions with two kinds of semiconductors
- junctions with a semiconductor and a metal.

Semiconductor heterojunctions can be classified into two types from an electrical point of view:

- pn-heterojunctions: Two semiconductors having different conduction types (n-p)
- iso-heterojunction: Two semiconductors having the same conduction type (n-n or p-p)

Depending on the size and signature of the band offset different types of interfaces can form. Different band alignments at heterojunctions can be achieved by shifting the Fermi level (doping, temperature).

How can we realize such structures? Hetero structures can be synthesized by band gap or band structure engineering. If two different semiconductor layers are grown on top of each other they build a heterostructure. Then they are connected at the interface (c. f. thin film technology). At this interface the band structure of both materials must be connected in some way. Four different cases are possible as shown in figure 17.

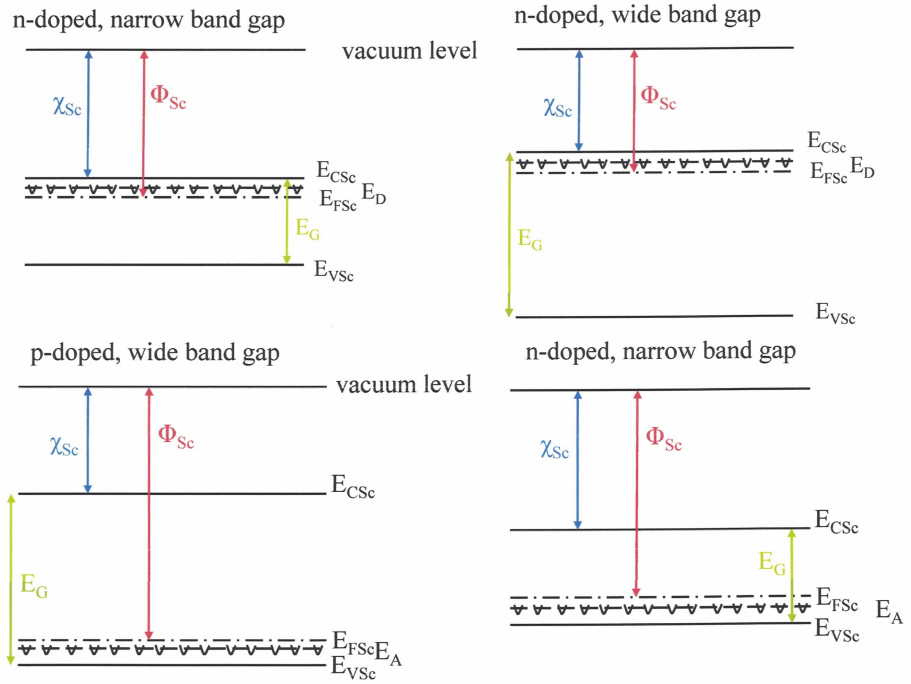


Figure 17: Different kinds of doped semiconductors. χ_{Sc} describes the electron affinity and Φ_{Sc} is the work function.

The vacuum level also controls the band structure. In the case of classical pn-junctions with the same material on both sides this is not a problem. For heterojunctions we have two main parameters:

- χ - electron affinity: Corresponds to the energy required to take an electron from the conduction band edge to the vacuum level.
- Φ - work function: Required energy to take an electron from the Fermi level to the vacuum level. It depends on the carrier concentration and the conduction type (determine the position of E_F).

The properties of heterojunctions are basically similar to those of homojunctions, although some important differences appear. If we combine semiconductor based materials with different gap and/or electronic structure we have to take into considerations the different positions of the vacuum level in these materials. Due to the differences in the chemical potential we can find different types of band-bending.

For example we discuss the SC-metal junction where we can distinguish two types of such structures:

1. $\Phi_M < \Phi_{Sc}$ *ohmic contact*:
   the voltage dependent carrier transport in both directions is possible.
2. $\Phi_M \geq \Phi_{Sc}$ *Shottky contact*:
   It acts as an additional tunnel barrier. A typical application is a detector or an emitter.

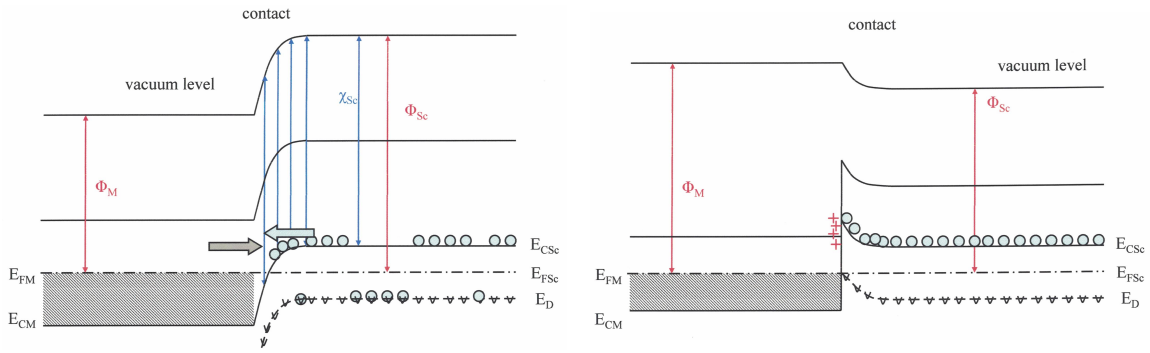


Figure 18: Left: Ohmic contact: $\Phi_M < \Phi_{Sc}$
Right: Shottky contact: $\Phi_M \geq \Phi_{Sc}$

When different semiconductors are in contact with each other the energy band is connected relative to the vacuum level (with respect to the electron affinity). this introduces discontinuities in the conduction band ΔE_C and valence band ΔE_V at the heterojunction. This leads to new properties compared to homojunctions, e. g. spikes and notches occur

$$\Delta E_C = \chi_P - \chi_N, \qquad \Delta E_V = \Delta E_g - \Delta E_C. \tag{2.52}$$

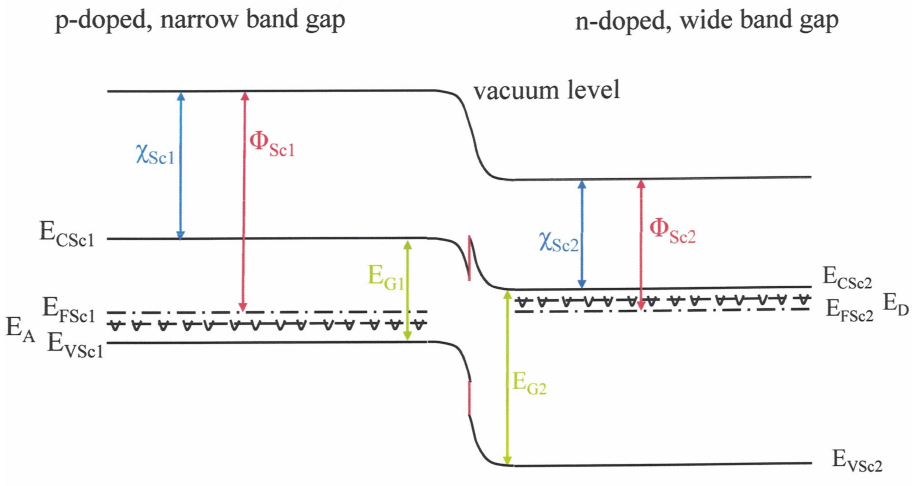


Figure 19: Band bending leads to spikes and notches in the conduction and valence band.

Remark: The same behaviour of band bending is also possible if one side of the heterojunction is not doped. Its nearly the same situation, however, the Fermi energy is fixed in the middle of the energy gab of the undoped semiconductor. In the region of the semiconductor with smaller energy gap free carriers can accumulate. This means we can create *free* electrons without impurities[2]. In a pn-heterojunction it is now possible to realize *high electron density* and *high mobility* at the same time, which is impossible with classic pn-junctions.

Forward bias mode

Under a forward bias, the built-in potential is reduced by the bias voltage. This reduction leads to the diffusion of electrons from the n-side to the p-side region. However, the diffusion of holes is prevented by the large potential barrier: $\Delta E_V + q(V_D - V)$. As in a pn-homojunction, the forward current in a heterojunction is also dominated mainly by the *diffusion* current, surface recombination current J_{Sur} and the generation-recombination current J_{NR}. The forward current can be expressed as

$$I_F = A_{\text{pn}}(J_{\text{diff}} + J_{\text{NR}}) + A_{\text{Sur}} J_{\text{Sur}} \sim \exp\left(\frac{qV_F}{a_f k_B T}\right), \tag{2.53}$$

where A_{pn} is the area of the pn-junction, A_{Sur} the area of the pn-junction surface and a_f a constant between 1 and 2. The main part of the diffusion current is an electron flow from the wide to the narrow band. As there are no holes involved we observe a different behaviour as in classic pn-junctions

$$J_{\text{Diff}} = J_{\text{sho}}\left[\exp\left(\frac{qV}{k_B T}\right) - 1\right] \quad \text{with} \quad J_{\text{sho}} = \frac{q\, D_{ep}\, n_p}{L_{ep}} \sim \exp\left(-\frac{E_{gp}}{k_B T}\right) \tag{2.54}$$

where J_{sho} ist the saturation current of the pn-heterojunction. D_{ep} is the diffusion constant, L_{ep} the diffusion lengths of electrons in the p-region, n_p the number of electrons in the p-region and E_{gp} the narrow band gap energy of the p-region.

[2] In the normal case, carriers exist only if we bring dopants inside the semiconductor material. Then a high density of electrons corresponds to high impurity levels. These higher doping levels lead to a smaller mobility and higher scattering rates.

After the applied forward voltage becomes larger than the narrow band-gap energy the electron diffusion current increases quickly

$$\text{if} \quad V_F > \frac{E_{gp}}{q} \quad \text{then} \quad I \quad \text{increases exponentially.} \tag{2.55}$$

Recombination processes occur at interfacial states which originate primarily from crystal defects (dangling bounds [3])

Isojuntion

Isojunctions are made by different kinds of semiconductors but with the same conduction type. When connected the bands in distance to the junction line up as governed by the Fermi level. This may result in some *band bending* at the interface. Depending on the positions of the Fermi levels and the chemical potentials γ_i a triangular quantum well may form at the interface.

We can also find discontinuities in the conduction band ΔE_C and the valence band ΔE_V

[3] A dangling bound is an unsatisfied valence of the atom.

3 Optical devices

3.1 Optical properties

The operation of semiconductor optoelectronic devices is based on the *optical properties* of the semiconductor as well as on their *electrical properties*. We should start with a more detailed discussion of emission and absorption which are the main optical properties.

3.1.1 Emission and absorption of light

In the thermal equilibrium state the occupation probability of electrons and holes in each band is given by the Fermi level and is expressed with a Fermi-Dirac distribution function. In an excited state (e. g. created by optical irradiation) a *large* number of electrons and holes are generated in each band. We can express the *occupation* probability of electrons in the conduction band and holes in the valence band by so called *quasi fermi levels* E_{F_C} and E_{F_V}.

For the equilibrium state each occupation probability is given by substitution of E_{F_C} and E_{F_V} in E_F

$$f_n(E) = \frac{1}{1+\exp\left(\frac{E-E_{F_C}}{k_B T}\right)}, \qquad f_p(E) = \frac{1}{1+\exp\left(\frac{E-E_{F_V}}{k_B T}\right)}. \tag{3.1}$$

The difference between E_{F_C} and E_{F_V} indicates the deviation from the equilibrium state $\sim E_{F_C} - E_{F_V}$. That means if $E_{F_C} = E_{F_V} = E_F$ the excitation ends. A non equilibrium state returns to equilibrium.

By using the quasi-Fermi levels we can also describe the carrier distribution and/or electron density (in the conduction band)

$$\text{electron density} \qquad n = N_C \cdot f(E_C) = N_C \exp\left(\frac{E_{F_C} - E_C}{k_B T}\right) \tag{3.2}$$

$$\text{hole density} \qquad p = N_V \cdot f(E_V) = N_V \exp\left(\frac{E_V - E_{F_V}}{k_B T}\right). \tag{3.3}$$

Emission

When the semiconductor is in thermal equilibrium state and is illuminated with light $E_\lambda \geq E_{\text{gap}}$ (band-gap energy)

1. the electrons are excited from valence band to the conduction band which is called *optical excitation* or *pumping*. The corresponding holes are left in the valence band.
2. The excited electrons stay in the conduction band for certain time (*life time*).
3. The electrons return to the valence band through *direct* or *indirect* recombination.

In the recombination processes the energy corresponding to the direct transition is emitted as a photon (*spontaneous emission*). The wavelength λ is given by

$$\lambda = \frac{h\,c}{E_g} \approx \frac{1.24}{E_g\,[\text{eV}]}\mu\text{m} \qquad \text{direct semiconductor} \tag{3.4}$$

$$\lambda = \frac{h\,c}{E_g - E_{\text{phon.}}} \approx \frac{1.24}{E_g - E_{\text{phon.}}}\mu\text{m} \qquad \text{indirect semiconductor.} \tag{3.5}$$

For a direct semiconductor the emitted energy is nearly equal to the band-gap energy E_g. The spectral width of the emitted light increases in the range of 10 meV...100 meV at room temperature because of the distribution of the electrons and holes in their respective bands. The phase of the emitted light is essentially random.

We can distinguish several possible processes of emission:

- band-to-band transition (1)
- band-to-impurity level transition (2) + (3)
- impurity level-to-impurity level transition (4). They occur in extrinsic (doped) semiconductors.

They are shown in figure 20.

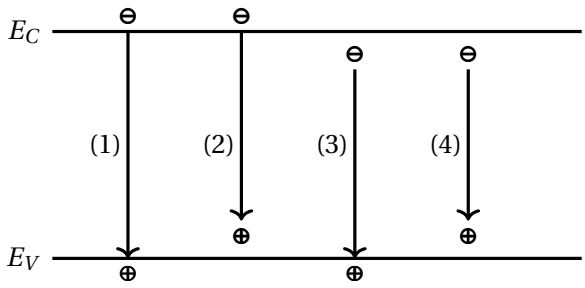


Figure 20: The different possible emission processes in a semiconductor.

For indirect semiconductors the radiative transition is related mainly to impurity levels such as excitonic transitions.

the radiative recombination rate increases in proportion to the product of the electron and hole densities (because the probability of recombination increases as electron and hole densities increase)

$$\text{spontaneous emission rate: } R_{\text{sp}} = B_{\text{sp}} \cdot n\,p, \tag{3.6}$$

where B_{sp} is the radiative recombination coefficient describing the transition probability of spontaneous emission. The value of B_{sp} depends on the band-gap energy and the type of band structure

$$\begin{aligned} B_{\text{sp}} &\approx 10^{-9} \quad \text{to} \quad 10^{-11}\,\frac{\text{cm}^3}{\text{s}}\,(\text{direct SC}) \\ B_{\text{sp}} &\approx 10^{-13} \quad \text{to} \quad 10^{-15}\,\frac{\text{cm}^3}{\text{s}}\,(\text{indirect SC}). \end{aligned} \tag{3.7}$$

The spontaneous emission rate consists of two parts: The radiative recombination rate in the thermal equilibrium state R_{sp0} and the recombination rate for the excited excess carriers R_r

$$\begin{aligned} R_{sp} &= R_{sp0} + R_r = B_{sp}(n_0 + \Delta n)(p_0 + \delta p) \\ &= B_{sp}\big[n_0 p_0 + n_0 \underbrace{\Delta p}_{\approx \Delta n} + \Delta n p_0 + \Delta n \underbrace{\Delta p}_{\approx \Delta n}\big] \\ &= B_{sp}\big[\underbrace{n_0 p_0}_{\sim R_{sp0}} + \underbrace{\Delta n(n_0 + p_0 + \Delta n)}_{\sim R_r s}\big], \end{aligned} \tag{3.8}$$

where Δn and Δp are the excited carrier densities. Under high excitation conditions $\Delta n \gg n_0$ or p_0, then

$$R_{sp} \approx R_r \approx B_{sp}\Delta n^2 \approx B_{sp} n^2. \tag{3.9}$$

Excitation increases the electron density in the conduction band and the hole density in the valence band. In such a way the intensity of spontaneous emission increases because $R_{sp} \sim n^2$. The emission peak energy becomes higher, because, as the carrier density increases, the electrons in the conduction band and holes in the valence band are filled from the band edges. As a result spontaneous emission with higher energies increases. This introduces a peak shift and is called *band-filling effect*.

Absorption

There are several optical absorption processes we want to discuss in the following:

a) Fundamental absorption

b) Absorption via energy levels in the band-gap

c) Exciton absorption

d) Intraband transition

e) Free carrier absorption.

The amount of absorption is expressed by *Lamberts law*

$$\mathrm{d}I = -\alpha I \,\mathrm{d}x \quad \text{or} \quad I = I_0 \mathrm{e}^{-\alpha x} \tag{3.10}$$

with α being the absorption coefficient, I the light intensity and x the distance in the material. The absorption coefficient is a function of wavelength/energy of the light. The absorption length (or penetration depth) is the inverse of α. At the absorption length, the light intensity has dropped to $1/\mathrm{e}$ of its initial value I_0.

a.) Fundamental absorption

Fundamental absorption corresponds to the excitation of electrons from the valence band to the conduction band. If the energy of the light is greater/lower than the band-gap energy, the light is absorbed/transmitted. The energy at which the fundamental absorption starts is called *absorption edge*. At energies near the absorption edge α depends on the energy difference

$$\begin{aligned} \alpha_{CV} &\sim \sqrt{h\nu - E_g} && \text{direct SC} \\ \alpha_{CV} &\sim \sqrt{h\nu - E_g \pm E_{\text{phon.}}} && \text{indirect SC} \end{aligned} \tag{3.11}$$

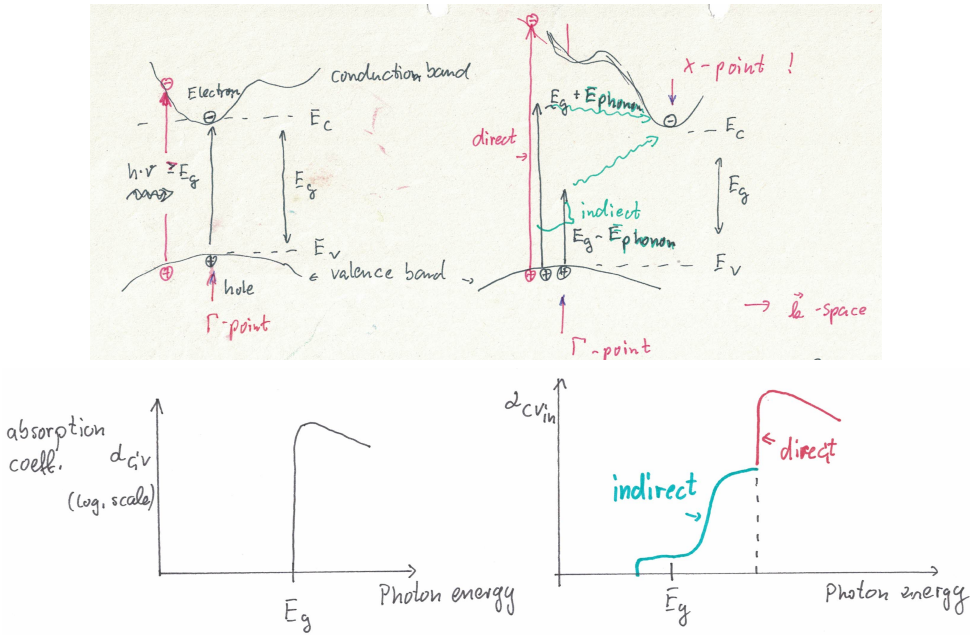


Figure 21: Illustration of the fundamental absorption processes for direct semiconductors (left) and indirect semiconductors (right).

The absorption coefficient increases more rapidly in direct SC than in indirect SC. First we want to explain the increase of the absorption coefficient which is explained by the dependency on the density of states in the conduction band. Remember that for higher energies the number of available states increases. Therefore α is higher.

Thus the absorption coefficient increases exponentially and tends to saturate (at higher energies). The transition propability (in the optical absorption process) is higher in direct SC than in indirect SC, because the indirect transition is only possible with a change in momentum via a phonon. This additional condition reduces the probability.

Remark: The doping concentration influences the fundamental absorption process in any way, e. g. in highly doped SC the increase of the doping concentration (n-type) is due to the increase of the fermi level ($E_F > E_C$). In such a degenerated situation (overdoped) states near the band edge in the conduction band are occupied. Thus the absorption edge shifts to higher energies.

Absorption via impurity levels

There are many possibilities of absorbing light via impurity levels. A simplified illustration of the absorption in a direct band-gap SC is shown in figure 22

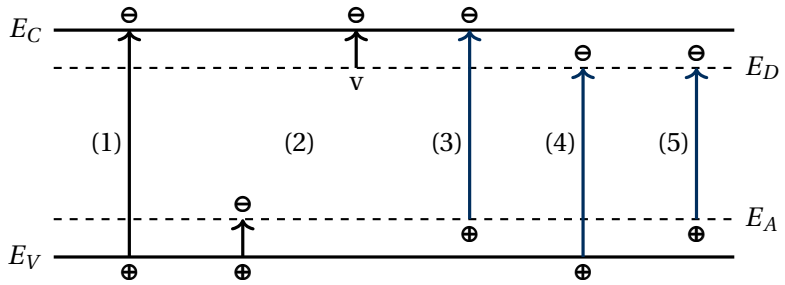


Figure 22: The different possible absorption processes in a semiconductor. The blue transitions require ionized levels. The probability of transition drops from left to right.
(1): fundamental absorption
(2): Donor/acceptor level to valence/conduction band
(3): Ionized acceptor level to conduction band.
(4): Valence band to ionized donor level.
(5): ionized acceptor level to ionized donor level.

For the blue transitions in figure 22 the energies are observed as shoulders on the low-energy side of the absorption edge (just like in figure 21). We found a lot of possibilities for absorption processes in extrinsic SC. The optical absorptions of the band-to-impurity levels occur at wavelengths in the far-infrared region corresponding to the energies of donor-level-to-conduction band or valence-band-to-acceptor level transition.

Exciton absorption

Excitons are electron-hole pairs which are held together by their mutual Coulomb interaction. We distinguish two kinds of excitons: *free* and *bound* excitons. The latter are localized in the vicinity of a donor, acceptor or neutral atom. Excitons are created only in *very pure* SC, because the Coulomb interaction is easily screened by free carriers. Exciton absorption typically occurs at a photon energy lower than the band-gap energy due to the exciton energy which is a few meV. Thus in direct SC at low temperatures a sharp line transition is usually observed at an energy slightly below the band-gap energy.

Intraband transition

Intraband transition occurs in both n-type and p-type SC. The most import transition is that in p-type SC. The reason for Intraband transition is the separation of the valence band into three sub-bands, the *light-hole* band, the *heavy-hole* band and the *split-off* band which are results of the spin-orbit interaction.

The possible interactions are shown in figure 23. The observed transitions depend on the doping level and temperature.

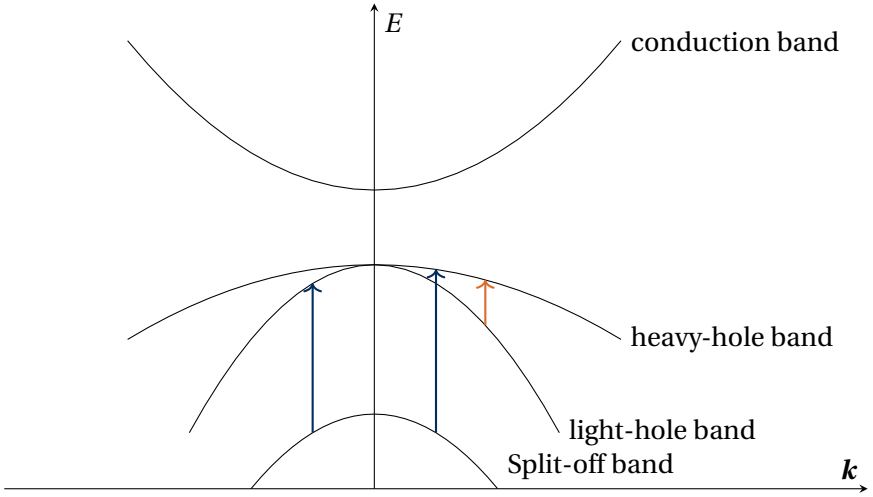


Figure 23: Intraband transitions in the valence sub-bands. We observe three different transitions from the split-off band to the light/heavy-hole band (blue) and from the light-hole band to the heavy-hole band (orange).

Free carrier absorption Free carriers within the conduction band can interact with light. Within the same valley the electron can be excited to higher energy levels. Since free states only exist at the lower edge of the conduction band these excitations require a change in the momentum of the carriers during the transition (phonons are necessary). The absorption coefficient described by this type of free-carrier absorption α_{fc} is proportional to the square of the wavelength

$$\alpha_{\text{fc}} \sim N \cdot \lambda^2, \tag{3.12}$$

where N is the free carrier concentration. Free-carrier absorption is also induced by the interband transition.

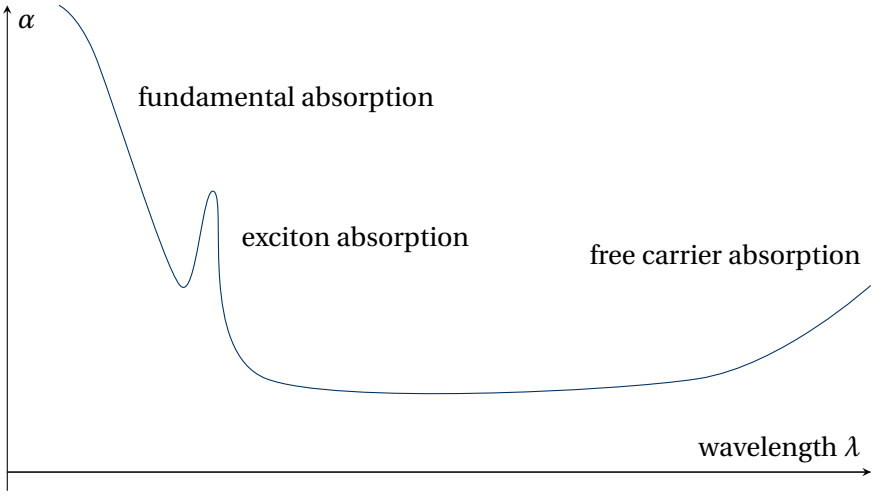


Figure 24: Total absorption coefficient as the sum of each absorption process in a logarithmic scale.

The total absorption coefficient is the sum of each coefficient determined by the different absorption processes. For high energies fundamental absorption is dominant. Slightly below

the band-gap, exciton absorption occurs and for larger wavelength free carrier absorption is dominant

$$\alpha = \alpha_{\text{fund.}} + \alpha_{\text{impurities}} + \alpha_{\text{excit.}} + \alpha_{\text{intrab.}} + \alpha_{\text{fc}}. \tag{3.13}$$

3.1.2 Stimulated emission

The second radiative recombination process (besides spontaneous emission) is stimulated emission. If we assume a two level system in our SC material with E_1 being the ground state and E_1 the excited state. Then in this case, if light with an energy $E_g = E_2 - E_1$ is incident to the excited state E_2 the excited electrons are *stimulated* and move in the same phase as the incident light. This leads to a coherent emission radiation with the same frequency. We can describe this situation by

$$\frac{n_2/\rho_2}{n_1/\rho_1} = \exp\left(-\frac{E_2 - E_1}{k_B T}\right), \tag{3.14}$$

where n_1, n_2 are the electron densities at E_1, E_2 and ρ_1, ρ_2 the density of states of electrons at E_1, E_2. This equation indicates that in thermal equilibrium the electron density at a higher energy level is less than at a lower level. Therefore only absorption is observed. However, if by excitation the incident light is amplified $n_2/\rho_2 > n_1/\rho_1$ the incident light is amplified leading to stimulated emission.

To be completed...

3.2 Scintillation counter

A scintillation counter is a combination of a scintillator (scintillation crystal) and a photomultiplier. Its basically a photon counter with a

- high gain/amplification $\rightarrow \frac{10^6 \text{ electrons}}{\text{photoelectron}}$
- high sensitivity $\sim \frac{\text{A}}{\text{mW}}$
- high frequency response
- low noise
- wavelength range: 100…1100 nm
- response time 1…20 ns

The quantum efficiency depends on the photo cathode and is about 10 to 25 %. The parts of the scintillation counter and the photomultiplier are shown in figure 25

The light arrives on the photoemission cathode and generates an electron with an energy corresponding to the difference of photon energy and the work function of the cathode. In an electrical field the electron is accelerated and hit another cathode (dynode) with much

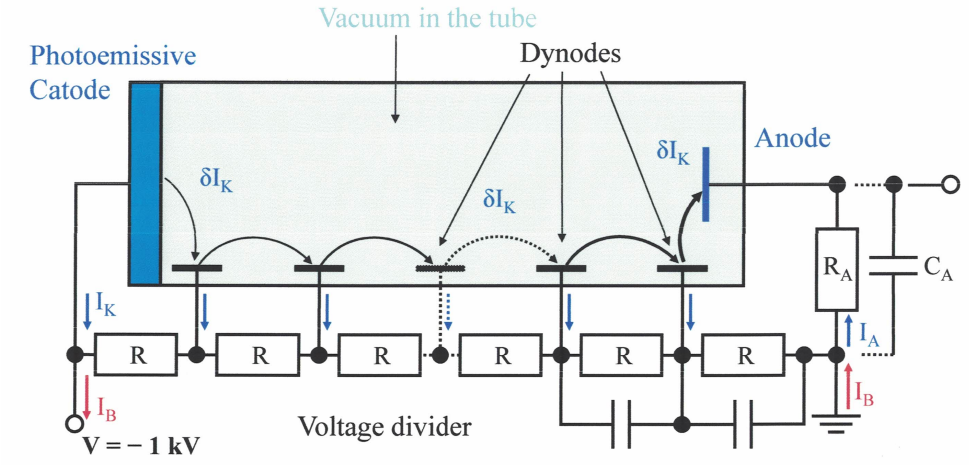


Figure 25: Scheme of a photomultiplier tube (PMT). It consists of several dynodes where the electrons are multiplied.

greater energy which releases several new electrons. They are again accelerated in the electrical field of the next dynode. The geometry of dynodes is such that a cascade occurs with an ever increasing number of electrons. Thus a strong amplification is achieved.

Remark: Typically uses voltages between dynodes are about 100 V and the gain per stage is about 10. The maximum anode current is limited to mA. Two different operation modes are possible: pulse counting (with C_A) and output-current-measurement (with R_A). We use alkali-metals as cathodes because they have a very low work function (Na, K, Sb, Cs).

The sensitivity of the scintillation device is influenced by *dark currents*. Even in total darkness you get a small anode current from a photomultiplier. This is caused by thermally excited electrons. This can be suppressed by cooling of the PMT.

Here are some safety instructions when working with PMT:

1. The applied voltage is lethal. Avoid contact during operation.
2. The tube of the PM is under vacuum. The photo cathode is very thin and might break easily. The fragments are usually toxic.
3. The PM is very sensitive. Too much light intensity (day light) destroys the dynodes or the anode leading to thermal evaporation.

3.3 Photoresistor

A photoresistor or photoconductor is a light dependent resistor (SC material). The resistance decreases with increasing incident light intensity.

Typically photoresistors are made of *high resistance* SC like intrinsic SC or slightly doped SC. The wavelength depends on the band-gap energy (intrinsic SC) being in the visible or infrared range and on the doping levels (doped SC) being in the IR or FIR range.

We can distinguish two operation modes of a photoresistor:

1. Generation of photo current: A current I_{ph} is flowing through the circuit and can be measured. An ampere meter is connected in series.
2. Change of voltage: A voltage V_{RL} drop across the load resistor (acts as a voltage divider) can be measured.

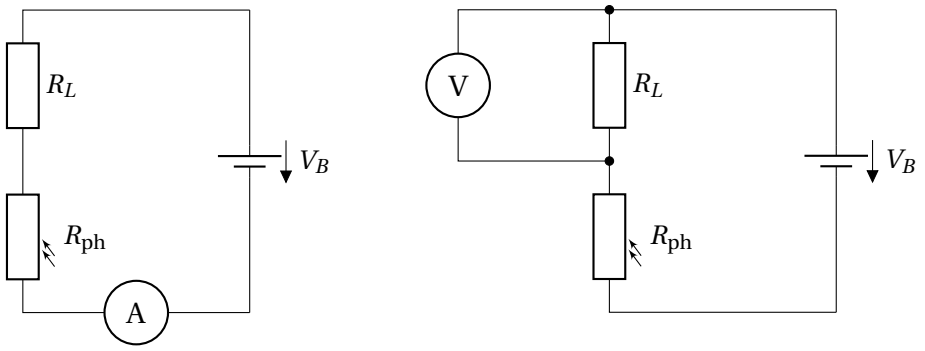


photo current mode voltage mode

Figure 26: Electronic circuit of a photoresistor measuring device with the two different operation modes. R_L is the load resistor and R_{ph} the photoresistor and V_B the bias voltage which is in any case necessary.

In both cases the resistance of R_{ph} decreases by irradiation with light. The generated photo current can be written as

$$I_{\text{ph}} = e \cdot A \cdot \Phi \cdot \eta \cdot \frac{\tau}{t}, \tag{3.15}$$

where A is the irradiation area, τ the life time, Φ the photon flux density and η quantum efficiency.

Possible materials for photoresistors are CdS (cadmium sulfide), PbS (lead sulfide) or Se (selenium). Typical resistance values of photoresistors are

$$\begin{aligned} &\text{dark resistance} && 10^6 \ldots 10^7\,\Omega \\ &\text{irradiated} && 100 \ldots 3500\,\Omega. \end{aligned} \tag{3.16}$$

The sensor consists of a thick film. Photoresistors have several advantages:

- Cheap and easy to produce, because they are prepared by thermal evaporation and sintering.
- film thickness: $10 \ldots 30\,\mu\text{m}$
- high sensitivity due to the high thickness.

However, photoresistors have one big disadvantage. They have a high time constant ranging between milliseconds and seconds depending on the intensity of irradiation. The life time of the excited carriers is increased due to traps in the band gap caused by disordered thin film (amorphous or polycrystalline materials with dangling bounds). We can achieve lower time constants with materials optimized for operation in the IR-range like PbS, InSb, $\text{Hg}_x\text{Cd}_{1-x}\text{Te}$ with time constants of $\mu\text{s} \ldots 100\,\mu\text{s}$.

3.4 Photo-diode

A classical photo-diode is made of a pn-junction or a heterojunction. It uses the photovoltaic effect that during the illumination of the SC ($h\nu \geq E_g$) electron-hole pairs are generated. If there is a pn-junction in the illuminated area, electrons and holes are separated by the electrical field at the pn-junction without any additional electric bias [4].

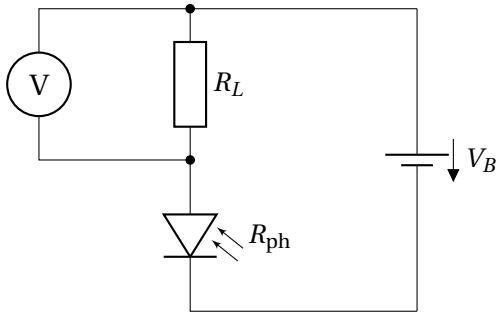


Figure 27: Electronic circuit of a Photo-diode. Here we apply a reverse bias condition and measure the voltage of the load resistor R_L.

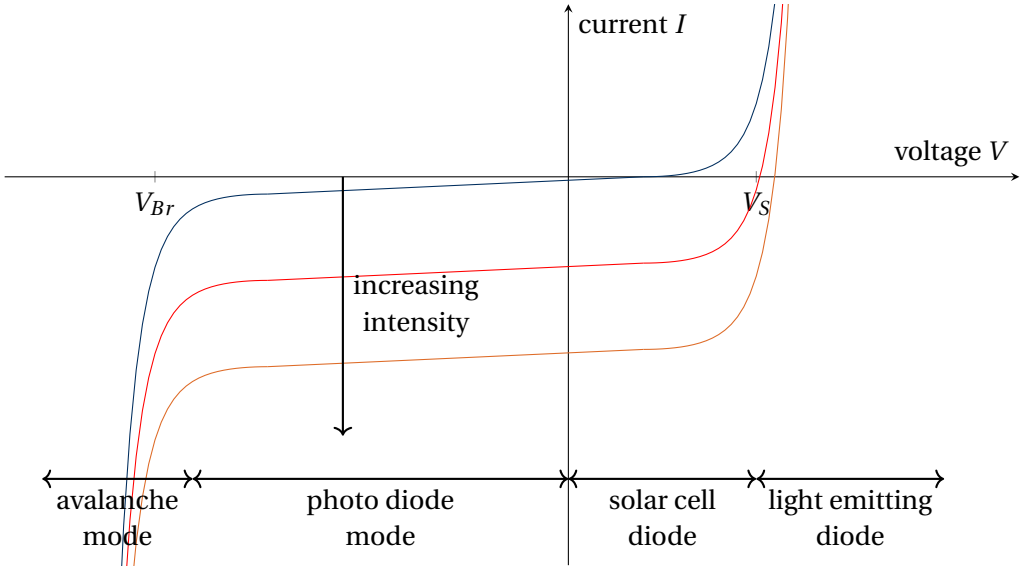


Figure 28: Current-voltage characteristic of a pn-junction with and without irradiation. We can identify the three operation modes of a pn-junction. In reverse bias mode we distinguish the avalanche mode and the photodiode mode and for forward bias we have the solar cell mode until the threshold voltage, when the photodiode becomes a light emitting diode.

We use the photo-diode in reverse bias mode where we have a certain *dark current* which consists of different parts

$$I_{\text{dark}} = I_d + I_S + I_{\text{gr}} + I_t, \tag{3.17}$$

where I_d is the diffusion current we already discussed for conventional pn-junctions. Ideally this is equal to the saturation current. I_S is a surface leakage current due to dangling bounds

[4] Remember that in the case of a photoresistor we need the bias voltage to separate the electron-hole pairs.

on the surface of the SC material. I_{gr} is the generation-recombination current, which is a leakage current in the depletion layer due to intrinsic impurities of the SC material. Finally we have the tunnel current I_t which is due to tunnelling and is dominant beyond the break-through voltage (but should also be avoided in the avalanche mode).

The diffusion current is given by equation (2.43)

$$j_d(v) = \left(\frac{e \cdot D_p}{L_p} p_n + \frac{e \cdot D_n}{L_n} n_p\right) \cdot \left[\exp\left(\frac{e \cdot V}{k_B T} - 1\right)\right]. \tag{3.18}$$

It depends on the band-gap energy and the effective density of states.
The *surface leakage* current I_S can be written as

$$I_S = \left(q \frac{A_{\text{Sur}}}{2}\right) n_i \cdot \sigma \cdot v_{\text{th}} \cdot N_S. \tag{3.19}$$

where A_{Sur} is the surface/interface are of the depletion layer, σ the carrier capture cross section, N_S the trap density and v_{th} the thermal velocity. The surface/interface states are formed at the surface/interface of the device. These states can easily trap carriers much more effectively as generation-recombination centres. This is shown in figure 29

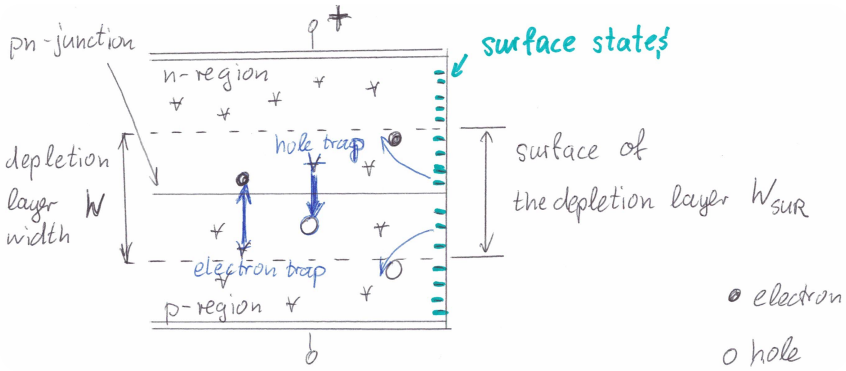


Figure 29: Trapping of electrons and holes in the depletion layer and free surface states due to dangling bounds.

The generation-recombination current I_{gr} is given as

$$I_{gr} = A_{\text{pn}} W q \sigma v_{\text{th}} N_T \sqrt{N_C \cdot N_V} \exp\left(-\frac{E_g}{2 k_B T}\right), \tag{3.20}$$

where A_{pn} is the area of the pn-junction. I_{gr} increases exponentially as the temperature increases and the band-gap energy decreases.

Lastly we consider the tunnelling current I_t

$$I_d = A_{\text{pn}} \left[\sqrt{2} q^3 F_{\max} \frac{|V_b|}{\pi h^2}\right] \sqrt{\frac{m_e}{E_g}} \exp\left(\frac{-\pi^2 \sqrt{m_e} E_g^{2/3}}{\sqrt{2} h q F_{\max}}\right), \tag{3.21}$$

where $F_{\max}$ is the maximum electrical field in the pn-junction and V_b is the reverse bias voltage. We observe that tunnelling probability increases with decreasing thickness of the potential barrier.

As seen before the dark current is temperature dependent. If T increases the dark current increases as well. In all cases the dominant current also depends on the operation mode. For photodiodes the generation-recombination current and the diffusion current are dominant whereas for avalanche photodiodes the tunnelling current strongly influences the I/V characteristic.

Quantum efficiency

The quantum efficiency η contributes to the photo-induced current and is defined by

$$\begin{aligned}\eta_{\text{ph}} &= \frac{\text{number of electron-hole pairs}}{\text{number of incident photons}} \\ &= \frac{\frac{I_{\text{Photo}}}{q}}{\frac{P_{\text{inc}}}{h\nu}} \cdot 100\ [\%]. \end{aligned} \tag{3.22}$$

Then we also define the responsivity S as

$$S = \frac{\text{photo-induced current}}{\text{incident optical power}} = \frac{I_{\text{photo}}}{P_{\text{inc}}} = \frac{q\eta_{\text{ph}}}{h\nu} = \eta_{\text{ph}}\frac{\lambda(\mu\text{m})}{1.24}. \tag{3.23}$$

The quantum efficiency describes the number electron-hole pairs in the depletion layer to the incoming photons. We can visualize this in figure 30.

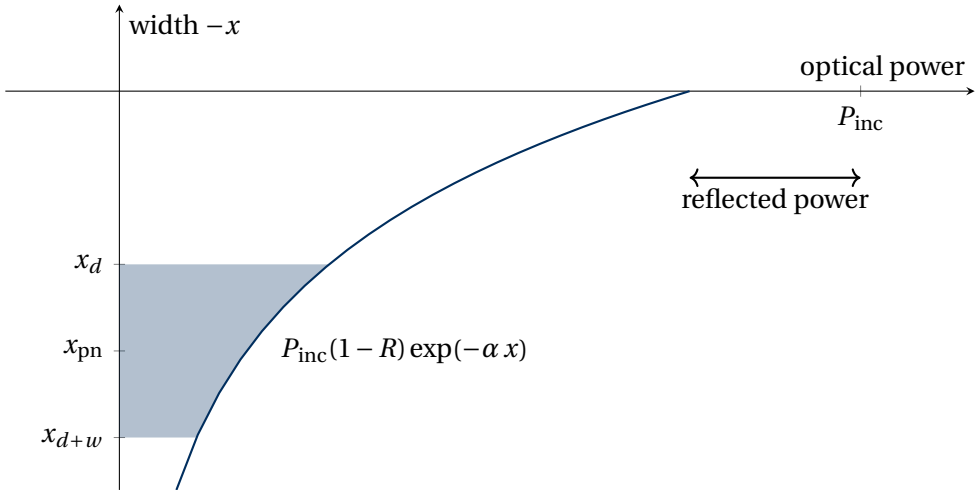


Figure 30: Optical power absorbed in the depletion layer of a pn-junction which follows Lamberts law. A part of the incoming power is lost due to reflections at the surface. Electron-hole pairs created outside the depletion layer do not contribute.

The quantum efficiency is influenced by:

- the reflectivity of incident light at the surface
- the recombination of photo-induced carriers at the surface and in the depletion layer

- the optical absorption outside the depletion layer.

The reflectivity of the incident light is typically 30 %. The reason is the refractive index of the SC material (about $\approx 3 \ldots 3.5$). In order to prevent reflection, the SC material is coated with antireflective dielectric films which can reduce the reflectivity to less than 1 %.

We can further enhance the photodiode by avoiding defects during the crystal growth and device fabrication. Often the light absorption layer is covered with a SC material with a larger gap for instance InP ($E_g \approx 1{,}35\,\text{eV}$) can act as cover or window layer for an InGaAs photodiode ($E_g \approx 0{,}75\,\text{eV}$). Figure 31 shows the quantum efficiency of different photo-diodes.

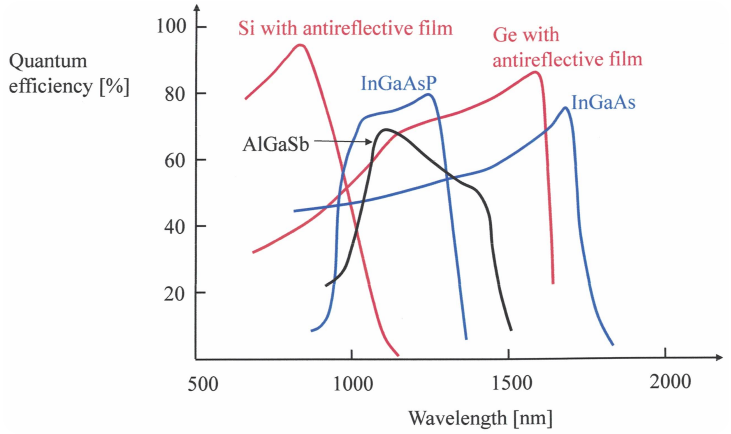


Figure 31: Quantum efficiency of pn-photodiodes with different materials.

Photo-diode modes

There are two possible operation modes for photo-diodes:

1. Measure the voltage across the load resistor R_L and apply an additional voltage V_B
2. Measure the photo current I without an additional bias voltage.

The operation modes depend on the measurement technique. We can measure very small current signals or we can measure very fast voltage signals with an oscilloscope. We want to look at the current-voltage characteristic of the voltage measurement in more detail (see figure 32).

With the load resistor the electronic circuit is a voltage divider. It is helpful to measure the voltage across the load resistor R_L, because than the voltage increases, when the diode is illuminated. Ideally, the load resistor should be chosen as high as possible in order to obtain the largest sensitivity of the measurement device. However, then we can only measure small currents.

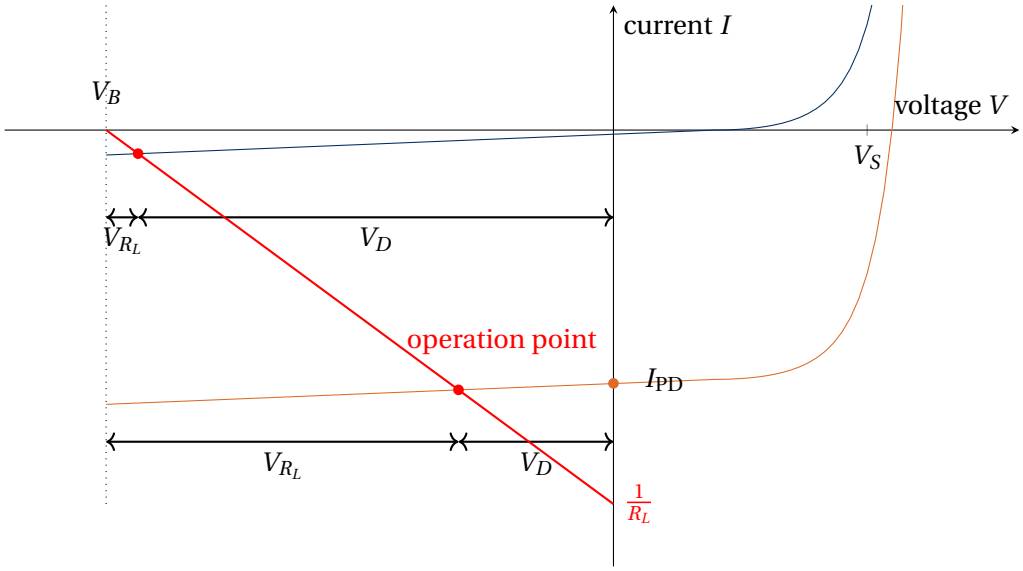


Figure 32: Current-voltage characteristic of a photo diode in the operation mode with load resistor and voltage measurement.

3.5 PIN photo diodes

The great disadvantage of our pn-junction is the width of the depletion layer, because only in this region electron-hole pairs can be used in the detector junction (see figure 30).

The main advantages of PIN photo-diodes is a *higher quantum efficiency*, a *low dark current* and a *low operation voltage*.

A PIN photo-diode consists of a:

1. heavily negative doped region (n)
2. undoped (or lightly doped) region (i)
   The undoped region is in the most cases inserted between a p-type and n-type region and is called *i-layer*. It determines the performance of the diode.
3. heavily positive doped region (p)
4. guard ring which avoids edge breakdown
5. dielectric layer which reduces reflection

In the i-layer we find a carrier concentration of $10^{13}\,\frac{1}{\mathrm{cm}^3}$...$10^{16}\,\frac{1}{\mathrm{cm}^3}$ which is several orders of magnitude lower than in the p/n-layer. In this way the depletion layer expands to the i-layer which leads to a larger thickness of 3 µm (GaAs) or few tens of µm (Si or Ge). If we apply a reverse bias voltage the i-layer is fully depleted and the width d of the depletion layer increases $d \sim \sqrt{V_B}$. When the *reach-through* voltage is reached the thickness of the depletion layer and the i-layer are equal. Note that the quantum efficiency increases in proportion to the increase in the depletion layer width.

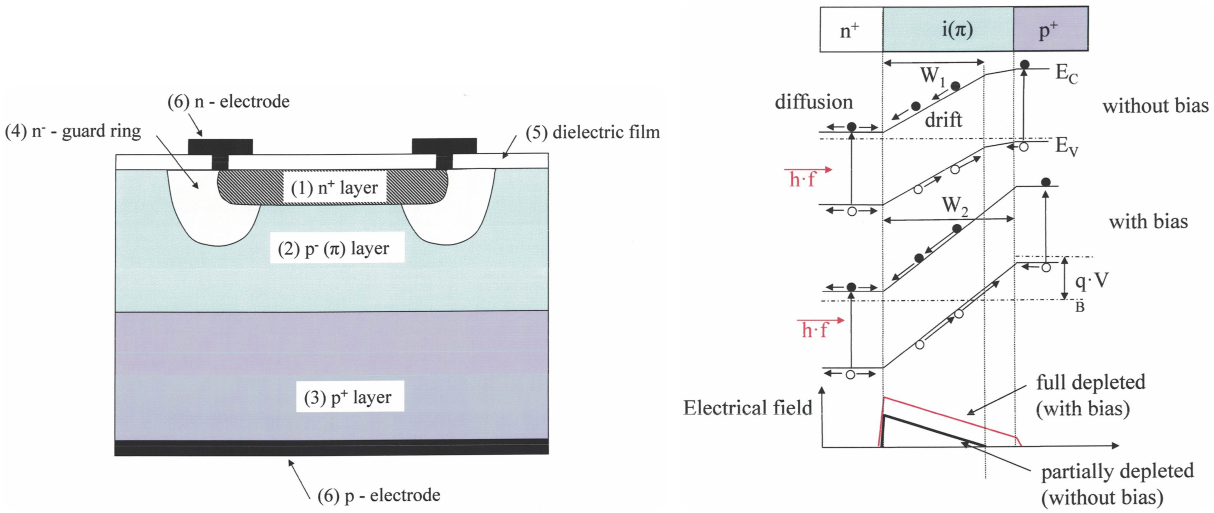


Figure 33: Components (left) and band diagram (right) of a PIN-photo-diode (n^+-p^--p^+ type).

The photo current I_{ph} again consists of several parts

$$I_{\text{ph}} = I_{\text{photo}} + I_{\text{Sat}} + I_{\text{bg}}, \tag{3.24}$$

where the largest part is the photo-induced current I_{photo}. However, we also observe leakage currents, e. g. a saturation current I_{sat} or background radiation I_{bg}. Under normal operating conditions we have $I_{\text{photo}} \ll I_{\text{sat}} + I_{\text{bg}}$. Then I_{ph} is nearly equal to the photo-induced current $I_{\text{ph}} \approx I_{\text{photo}}$. Then we can write the drift current density as

$$\frac{I_{\text{ph}}}{A_{\text{ph}}} = j_{\text{ph}} = j_{\text{ph-dr}} + j_{\text{ph-diff}}, \tag{3.25}$$

where we have a drift current density $j_{\text{ph-dr}}$ from the photo-induced carriers generated within the depletion layer and a diffusion current density $j_{\text{ph-diff}}$ from the photo-induced minority carriers. They are generated within their diffusion length from the edge of the depletion layer. The total current can thus be expressed as

$$J_{\text{ph}} = q(1-R)\left(\frac{P_{\text{inc}}}{A_{\text{ph}}}h\nu\right)\frac{1-\exp(-\alpha\, d)}{1+\alpha L_{\text{ep}}} + q n_{p0}\frac{D_{\text{ep}}}{L_{\text{ep}}}, \tag{3.26}$$

where D_{ep} is the diffusion constant for electrons and L_{ep} the minority carrier diffusion length, which is in the order of a few µm. Thus $\alpha\, L_{\text{ep}}$ is usually negligible. Then we can conclude an external quantum efficiency

$$\eta_{\text{ex}} = (1-R)\frac{1-\exp(-\alpha\, d)}{1+\alpha L_{\text{ep}}}\eta_{\text{in}} \tag{3.27}$$

as a function of the internal quantum efficiency η_{in} which is necessary, if not all photo-induced carriers contribute to the photo-induced current (defects at the surface/depletion region). The external quantum efficiency is in the order of about 70 %.

Frequency response

The $\tau = C \cdot R$ time constant is a common problem in each type of photo-diode. In a common-pn-junction diode the limited width of the depletion layer leads to a large capacitance thus determining the time constant. The same is true for PIN-diodes before reach-through. Basically the pn-doped regions act as capacitor ($C = \varepsilon_0 \varepsilon \frac{A}{d}$) plates with the depletion layer thickness being the distance of those plates. The residual part of the i-layer acts as a high electrical resistance because of the low carrier concentration. We can substitute the photo-diode with an equivalent circuit as shown in figure 34

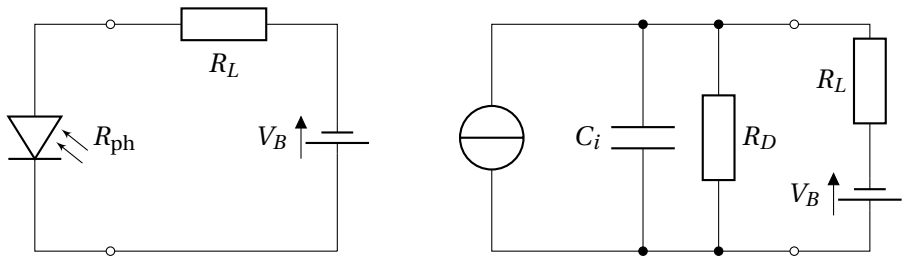


Figure 34: Electronic circuit of a photo-diode (left) and its equivalent circuit with an ideal current source, the capacitance C_i of the i-layer and the internal resistance R_D of the diode. In the ideal case the internal resistivity of the voltage source is zero.

The total resistivity contributing to the time constant depends also on the load resistance

$$\frac{1}{R_{\mathrm{DL}}} = \frac{1}{R_D} + \frac{1}{R_L} \quad \Rightarrow \quad R_{\mathrm{DL}} = \frac{R_D R_L}{R_D + R_L}. \tag{3.28}$$

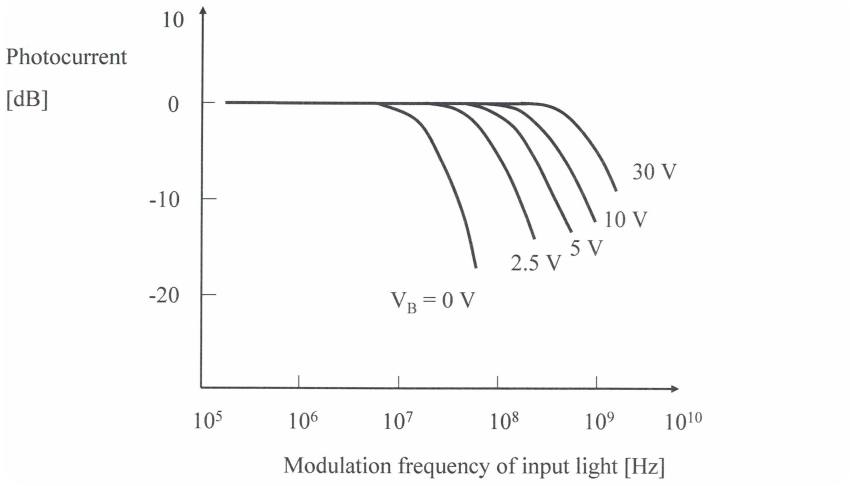


Figure 35: Time dependence of PIN-photo-diodes.

The *cut-off frequency* of the photo-diode is given by

$$f_C = \frac{1}{2\pi C_i R_{\mathrm{DL}}}. \tag{3.29}$$

We can decrease C_i by increasing the bias voltage. In this case the depletion layer widens $C \sim \frac{1}{\sqrt{V_B}}$. For a InGaAs PIN-diode typical values are $C_i = 100\,\mathrm{fF}, R_D = 100\,\mathrm{M\Omega}, R_L = 50\,\Omega$.

This results in a cut-off frequency of 30 GHz. We may also think about lowering the value of the load resistor R_L. However, as discussed earlier, we need a high resistivity for sensitive measurements. So we can either measure fast or sensitive. Generally, for high-frequency responsivity PIN-diodes are used at a high reverse bias after reach-through.

3.6 Avalanche photodiodes

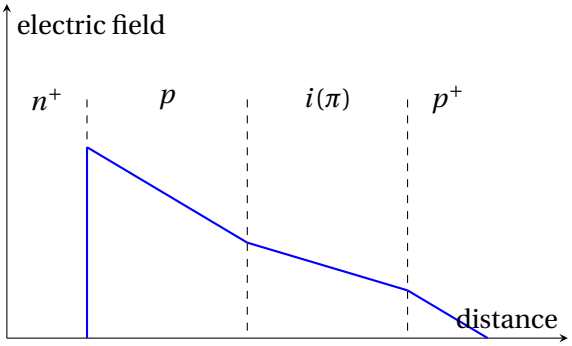
This type of photodiodes operates at nearly its breakdown voltage. Here we amplify the photo-induced current by avalanche multiplication of the carriers (electrons or holes). The APD consists of (see figure 36) a

1. Light-absorbing region (i-layer)
2. Avalanche region (p- or n-doped)
3. antireflective dielectric film (to reduce reflection at the surface)
4. guard ring around the diffusion region (to avoid edge breakdown)
5. Heavily doped n- and p-layers

The so called *mesa* structures (i-layer, avalanche region) can be used to reduce the junction capacitance for high frequency applications. The depletion layer mainly elongates into the p-layer and under normal operation conditions reaches the p^+-layer. Since the light absorbing region is separated from the avalanche region, the photodiodes are also called SAM-APD (separated absorption and multiplication APD). The doping levels of the avalanche region and i-layer are

$$\begin{aligned} \text{p-layer} \quad & \sim 10^{16}\,\frac{1}{\text{cm}^3} \\ \text{i-layer} \quad & \sim 10^{14}\,\frac{1}{\text{cm}^3}. \end{aligned} \tag{3.30}$$

Light is absorbed in the i-layer which generates electron-hole pairs. Only the electrons drift in the p-doped avalanche region. Since there is a high electric field (see figure on the right), the electrons are accelerated and ionize atoms in the avalanche region creating more electron-hole pairs. The large number of electrons then reach n^+ layer.



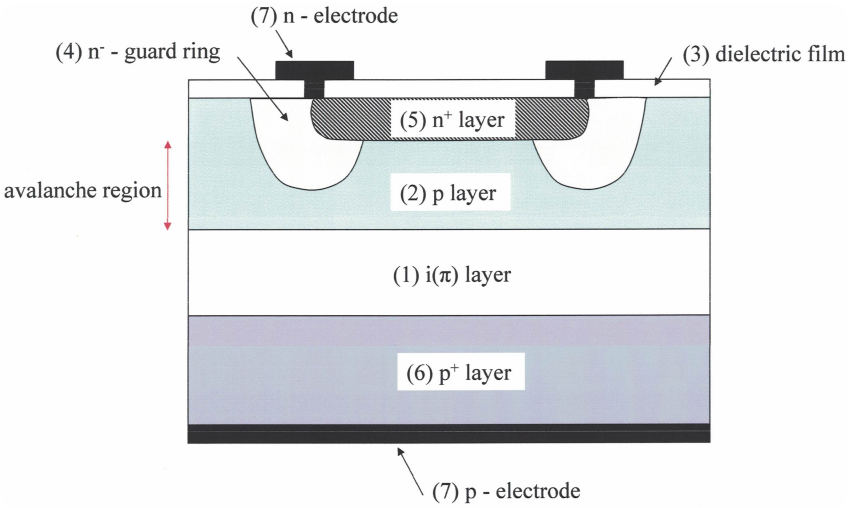


Figure 36: Components of an Avalanche photodiode.

Impact ionization

The main process for amplification is impact ionization. If the electrons are accelerated and a high-velocity carrier collides with the lattice *impact ionization* occurs. It is a three-carrier collision process in which a high-energy primary carrier produces an electron-hole pair. The ionization requires the energy for excitation from the valence band to the conduction band

$$E_{e^-} \leq \frac{3}{2} E_g. \tag{3.31}$$

The ionization coefficient α_{ion} for the electron is given by

$$\alpha_{\text{ion}} = \frac{q \cdot F_{\text{field}}}{E_{\text{e,th}}} \exp\left(-\frac{E_{\text{e,th}}}{q \cdot F_{\text{field}} \cdot L_{\text{e,mfp}}}\right), \tag{3.32}$$

where $L_{\text{e,mfp}}$ is the mean free path length of the electron between two collisions and $E_{\text{e,th}}$ the threshold ionization energy which is the minimum energy for impact ionization. The ionization coefficient for the hole is similarly

$$\alpha_{\text{ion}} = \frac{q \cdot F_{\text{field}}}{E_{\text{h,th}}} \exp\left(-\frac{E_{\text{h,th}}}{q \cdot F_{\text{field}} \cdot L_{\text{h,mfp}}}\right). \tag{3.33}$$

Finally we want to introduce the multiplication factor M

$$M = \frac{\text{photocrrent amplified at high bias in the avalanche mode}}{\text{photocurrent at a low bias photodiode mode}}$$
$$M_{\text{max}} = \sqrt{\frac{|V_{\text{Br}}|}{m_M I_{\text{Ph}} \cdot R_{\text{eq}}}}, \tag{3.34}$$

where V_{Br} is the avalanche breakdown voltage, m_M a constant depending on SC material, the doping profile and the wavelength. R_{eq} is the series resistance of the i-layer and the low doped layer (avalanche layer). We observe that the multiplication factor is proportional

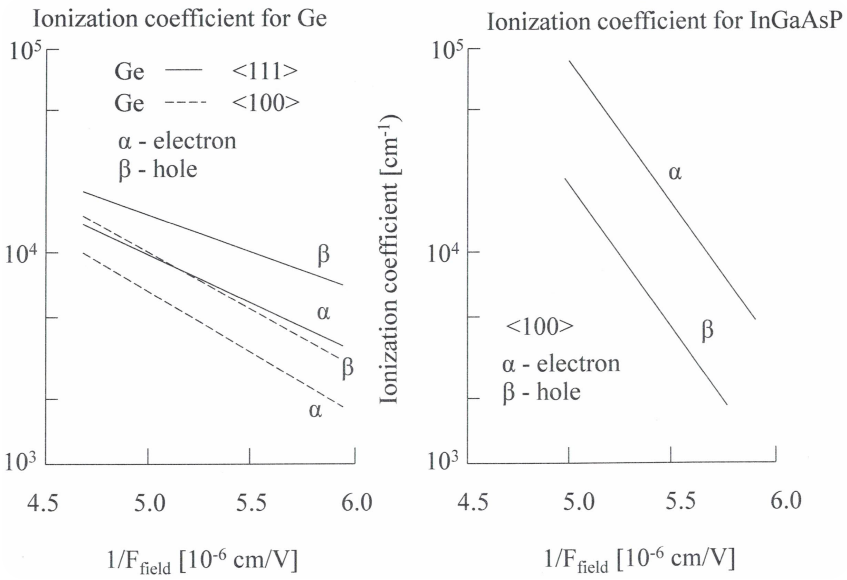


Figure 37: Ionization coefficient of APDs. It is a material property depending on the material itself and the crystal orientation.

to the square root of the breakthrough voltage. Also working with a higher bias voltage increases the multiplication factor rapidly. However, at higher voltages ($\geq 10^6 \frac{\mathrm{V}}{\mathrm{m}}$) Zener breakdown occurs and the multiplication of the photocurrent disappears. Therefore the doping concentration of the depletion layer should be as low as possible. Also the band-gap energy of the material used in the depletion layer should be as large as possible.

Lastly lets take a look at the frequency response of the multiplication factor which is given by

$$M(\omega) = \frac{M_0}{\sqrt{1 + (\omega^2 M_0 \tau)^2}} \quad \text{with} \quad \tau \sim \left(\frac{\beta_{\text{ion}}}{\alpha_{\text{ion}}}\right)\frac{w}{v}, \tag{3.35}$$

where τ is the mean drift time within the avalanche region. It is influenced by the ionization coefficients, the width of the avalanche region w and the saturation velocity $v = \mu \cdot F_{\text{field}} \approx 10^7 \frac{\mathrm{cm}}{\mathrm{s}}$. We observe that for increasing drift time and frequency the multiplication factor decreases. However, choosing a material with a high electron ionization coefficient compared to hole ionization will increase the multiplication factor. Larger saturation velocities (higher mobility μ) are also helpful.