Prof. **PIETRO PATIMISCO**

Dipartimento INTERATENO di Fisica
II PIANO – Stanza 233
pietro.patimisco@uniba.it

OFFICE HOURS:
on request, by email

COURSE PROGRAM
LECTURE SLIDES

http://polysense.poliba.it/index.php/optoelectronics-and-nanotechnology/

RECOMMENDED TEXTS

Govind P. Agrawal, Niloy K. Dutta – **Semiconductor Lasers**, AT&T Bell Laboratories, Murray Hill, New Jersey.
E. L. Dereniak, D. G. Crowe – **Optical radiation detectors**, Wiley, 1984.

SYLLABUS

1. Semiconductor structures

1.1. Silicon, Germanium and Gallium Arsenide. 1.1.1. Covalent Bonding. 1.1.2. Crystal Structure. 1.1.3. Energy Bands. 1.1.4. Band Gap. **1.2. Carrier Concentration at Thermal Equilibrium.** 1.2.1. Intrinsic Semiconductor. 1.2.2. Donors and Acceptors. 1.2.3. Extrinsic Carriers Concentration. **1.3. Carrier Transport Phenomena.** 1.3.1. Drift and Mobility. 1.3.2. Diffusion Process. 1.3.3. Current Density Equations. 1.3.4. Direct Recombination. 1.3.5. Continuity Equation. **1.4. p-n Junction.** 1.4.1. Thermal Equilibrium Condition. 1.4.2. Depletion Region. 1.4.3. Current-Voltage Characteristics. **1.5. Light-Emitting Diodes.** 1.5.1. Radiative Transitions and Optical Absorption. 1.5.2. Structure of a LED. 1.5.3. Emission Spectrum. 1.5.4. Quantum Efficiency. 1.5.5. The Light Escape Cone.

2. Diode Lasers

2.1. Conditions for laser emission. 2.1.1. Population inversion. 2.1.2. Two-level laser scheme. 2.1.3. Three-level laser scheme. 2.1.4. Optical Resonator. 2.1.5. Population inversion in a semiconductor. **2.2. Basic Concepts.** 2.2.1. Maxwell's Equations. 2.2.2. Longitudinal Modes. 2.2.3. Threshold Conditions. 2.2.4. Waveguide modes with Effective Index Approximation. **2.3. Emission Characteristics.** 2.3.1. Light-Current Characteristics. 2.3.2. Spectral Characteristics. **2.4. Distributed-Feedback Semiconductor Lasers.** 2.4.1. DFB Laser Structure. 2.4.2. Coupled-Wave Equations. 2.4.3. Continuous-Wave Operation. 2.4.4. Tunable DFB Diode lasers. 2.5. External-Cavity Diode Lasers. 2.5.1. Two-cavity Systems. 2.5.2. The Littrow and Littman-Metcalf configurations.

3. Quantum Cascade Lasers

3.1. 2D Quantum Heterostructures. 3.1.1. Finite Quantum Well. 3.1.2. Quantized Energy Levels. 3.1.3. Density of States. 3.1.4. Influence of Effective Mass. **3.2. Quantum Cascade Laser.** 3.2.1. The infrared spectral region. 3.2.2. The Auger Recombination. 3.2.3. The concept of quantum cascade. 3.2.4. Electron-phonon scattering in a subband. 3.2.5. Population inversion between two subbands. 3.2.6. Quantum Cascade Laser with Three-level active region. 3.2.7. Threshold current. 3.2.8. Advantages of Quantum Cascade Lasers.

SYLLABUS

4. Infrared Photodetectors

4.1. Photovoltaic Detector. 4.1.1. Photovoltaic Effect. 4.1.2. Photocurrent Generation. 4.1.3. Responsivity and Quantum Efficiency. 4.1.4. Temporal Frequency Response. 4.1.5. Noise Sources. 4.1.6. Figures of Merit. **4.2. Photoconductive Detector.** 4.2.1. Theory of Photoconductors. 4.2.2. Temporal Frequency Response. 4.2.3. Noise Sources. **4.3. Thermal Detector.** 4.3.1. Principle of Detection. 4.3.2. Heat Balance Equation. 4.3.3. Seebeck Effect and Thermocouples. 4.3.4. Thermopiles. 4.3.5. Pyroelectric Detectors.

5. Low Dimensional Structures

5.1. Introduction. 5.2. 2D Graphene. 5.2.1. Crystal Structure. 5.2.2. Brillouin Zone. 5.2.3. Energy Bands: Tight-Binding approach. 5.2.4. Energy Bands. 5.2.5. Density of States. **5.3. Quantum Wire.** 5.3.1. Energy Bands. 5.3.2. Density of States. 5.3.3. GaAs Nanowire: Subbands and Probability Density. **5.4. Quantum dot.** 5.4.1. Density of States. 5.4.2. Energy Levels in Spherical Potential Well. 5.4.3. Thermal vs Nonthermal Distribution. 5.4.4. Population Statistics: Rate Equations vs Random Population. **5.5. Phosphorene and Black Phosphorus.** 5.5.1. Crystal Structure. 5.5.2. Primitive Cell and Brillouin Zone. 5.5.3. Energy Bands and Density of States. 5.5.4. Field-Effect Transistors. 5.5.5. Photodetectors.

CHAPTER 1

SEMICONDUCTOR STRUCTURES

1.1 INTRODUCTION

The term semiconductor was introduced by **Alessandro Volta** in 1782 to indicate those materials that have intermediate properties between conductors and insulators.

The first experimental observation of a typical behavior of semiconductors, however, is due to **Michael Faraday**: in 1833 he noticed that some non-metallic compounds, such as silver sulfide (Ag_2S), increase their electrical conductivity as the temperature increases. This behavior is opposite to that of metals, in which conductivity decreases when the temperature increases.

From a chemical point of view, semiconductors are divided into two groups:

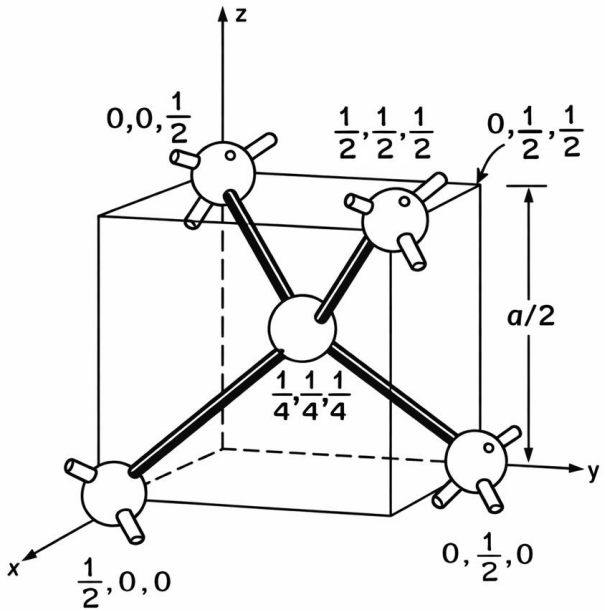
- **Elementary semiconductors.** Materials consisting of a single chemical element belong to this category. The most important are silicon and germanium, both tetravalent: this means that they have four electrons in the valence band
- **Compound semiconductors.** They are indicated by a chemical formula of the type AB, where A and B are different elements. Based on the value of A and B, distinct groups are identified:
  - **Group III–V:** A is trivalent and B is pentavalent. Examples: InSb (indium antimonide) and GaAs (gallium arsenide).
  - **Group II–VI:** A is divalent and B is hexavalent. Examples: ZnS (zinc sulfide) and CdS (cadmium sulfide).

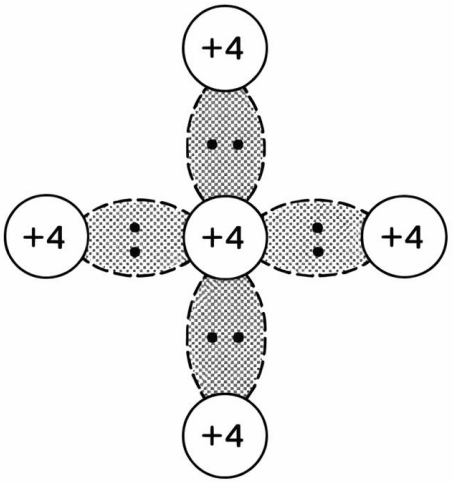
1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.1 Covalent bond

Silicon and **germanium** are referred to as **diamond**-type semiconductors, because they have the same crystal structure as diamond.

This structure belongs to the family of cubic crystals and consists of the superposition of two face-centered cubic lattices staggered along the diagonal of the cube by a quarter of its length, i.e. by $a\sqrt{3}/4$.





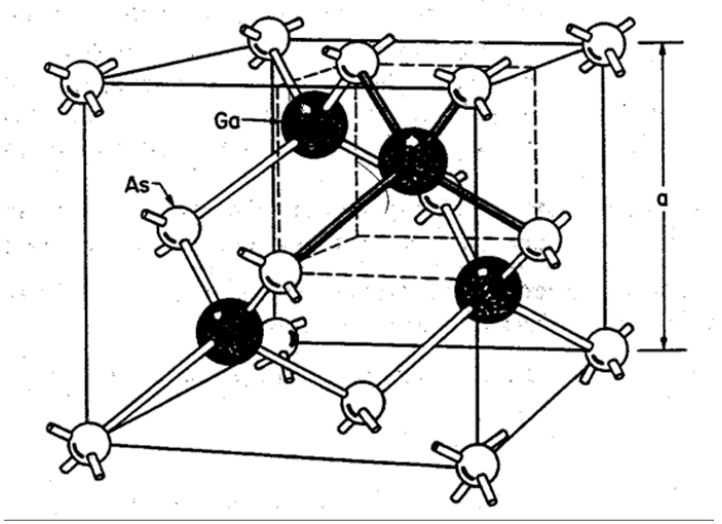
In a lattice of this type, all the atoms are identical, and each one is surrounded by four equidistant neighboring prime atoms, arranged at the vertices of a tetrahedron, as shown in the Figure.

1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.1 Covalent bond

Most of the compound semiconductors of group III–V, such as **GaAs**, have the structure called **zincblende**, which is substantially identical to that of diamond, with the only difference that one of the face-centered lattices is formed by the atoms of Group III (e.g. Ga), while the other is made up of the atoms of Group V (e.g. As).

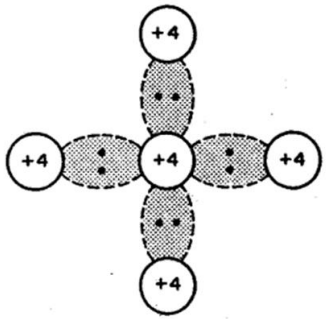
In both structures (diamond and zincblende) each atom has valence electrons that it shares with adjacent atoms forming **covalent bonds**; each pair of shared electrons is in fact a covalent bond.



In the zincblende lattice, the force that holds the crystal together is mainly due to covalent bonds, although there is a **weak ionic component** related to the electrostatic interaction between Ga^{-} and As^{+} ions, or vice versa.

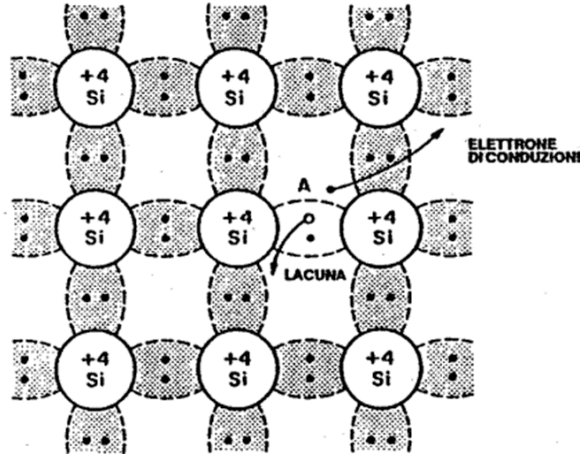
1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.1 Covalent bond



At low temperature the electrons remain bound within their tetrahedral covalent bonds and are not available for electrical conduction.

When the temperature rises, **thermal vibrations can break some covalent bonds**: in this case an electron is freed and can contribute to charge transport.



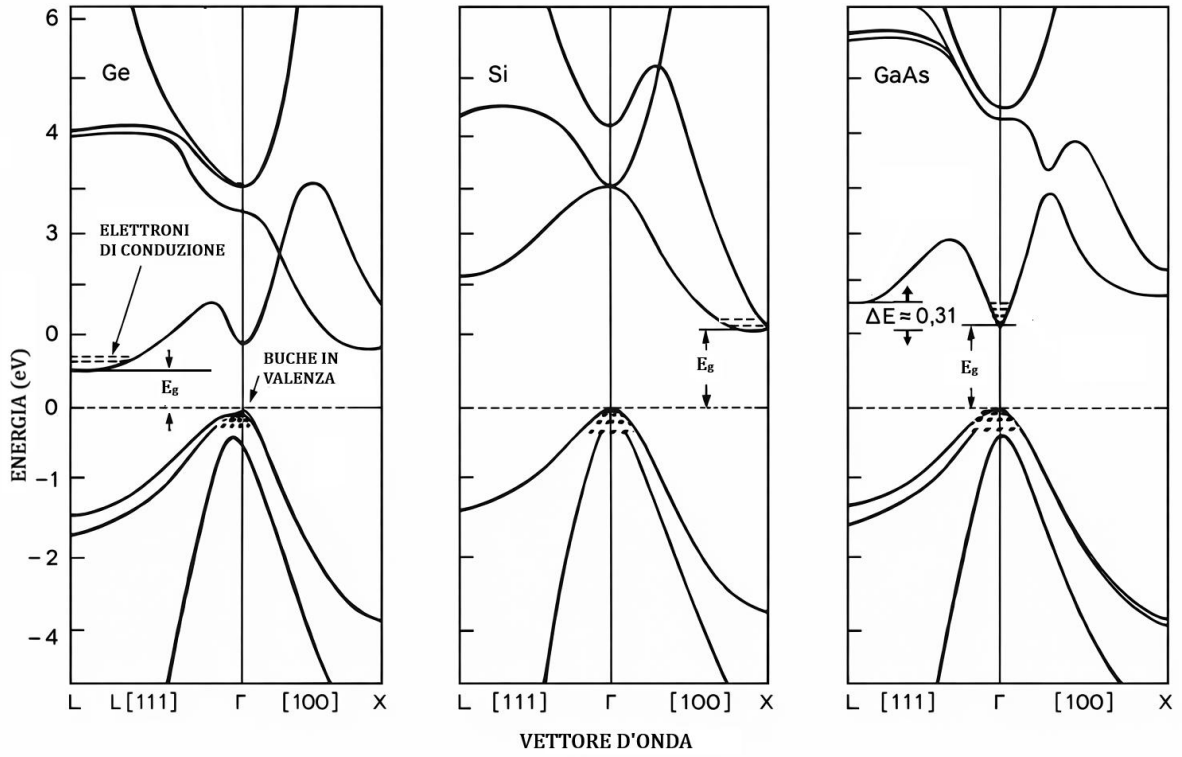
Breaking the bond also produces a lack of electron, called a **hole**, which can be filled by an electron belonging to a neighboring bond. This process involves moving the hole from one location on the lattice to another.

The hole behaves like a **positively charged quasi-particle** and, in the presence of an electric field, moves in the opposite direction to the electrons, contributing in turn to electrical conduction.

1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.3 Band structure

The band structure of **Germanium**, **Silicon** and **Gallium Arsenide** is shown in the Figure.



In all three cases there is a **band gap** of amplitude E_g between the lower edge of the conduction band and the upper edge of the valence band.

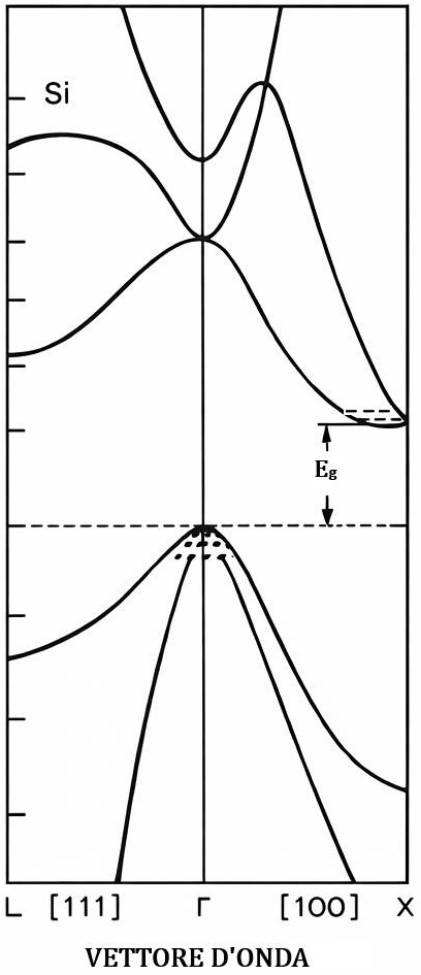
1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.3 Band structure

In **silicon**, the maximum of the valence band is located at point Γ, i.e. for $\vec{k} = 0$, while the minimum of the conduction band is located in the direction [100], at point X, so $\vec{k} \neq 0$.

This highlights an important difference between the **momentum of the particle** and the **crystal momentum**: in fact, the momentum of an electron is zero when its kinetic energy is zero, but in silicon an electron that is in the minimum of the conduction band can have zero kinetic energy while maintaining a crystal momentum that is different from zero.

As a result, to make a transition from the valence band to the conduction band, an electron in silicon must change not only its energy (by at least E_g), but also its crystalline impulse.



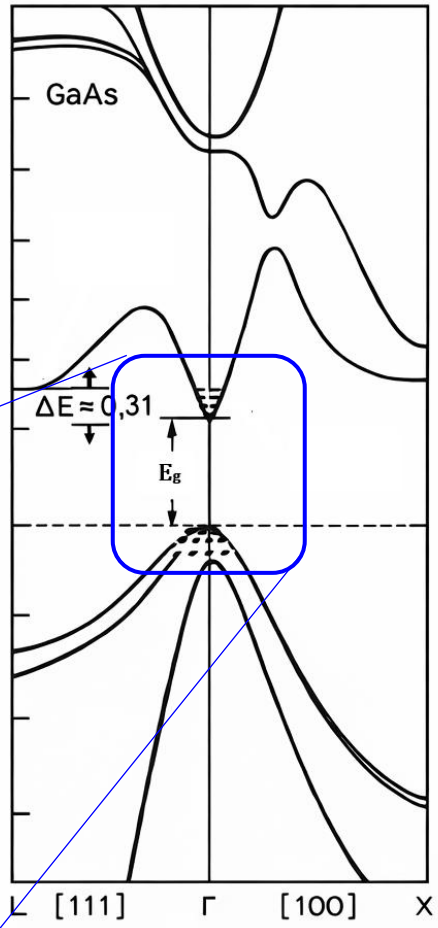
1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.3 Band structure

In **GaAs (gallium arsenide)**, on the other hand, the maximum of the valence band and the minimum of the conduction band are both found for $\vec{k} = 0$. This means that an electron can pass from the valence band to the conduction band without changing its crystalline pulse.

For this reason, GaAs is called a **direct semiconductor**, while silicon is classified as an **indirect semiconductor**, since the transitions between the valence band and the conduction band require a change in the crystal momentum.

The **bandgap** is defined as the energy difference between the minimum of the conduction band and the maximum of the valence band.

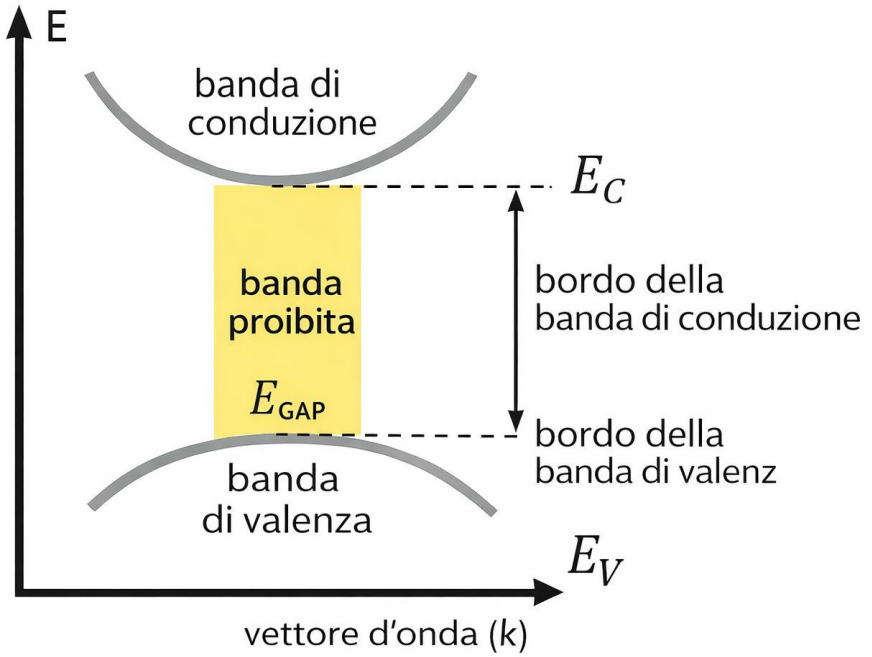


1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.4 Band Gap

The lowest energy point of the conduction band is called the **edge of the conduction band**, while the highest point of the valence band is called the **edge of the valence band**.

The lower edge of the conduction band is indicated by E_C : it represents the potential energy of a conduction electron when it is at rest, and the kinetic energy of this electron is measured above E_C .



Similarly, the upper edge of the valence band is denoted by E_V and corresponds to the potential energy of a hole; in this case the kinetic energy of the hole is measured below E_V.

1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.4 Band Gap

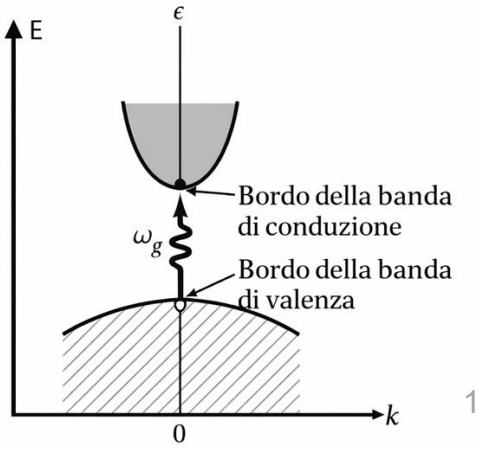
Both electrons in the conduction band and holes in the valence band contribute to electrical conduction. The dependence of the bandwidth of the bandwidth on temperature for silicon and GaAs can be described by the following relationships (energies in eV and temperature in K):

$$E_g(T) = 1.17 - \frac{4.73 \cdot 10^{-4} T^2}{(T + 636)} \quad silicon$$

$$E_g(T) = 1.52 - \frac{5.4 \cdot 10^{-4} T^2}{(T + 204)} \quad GaAs$$

In both cases the temperature coefficient $\boldsymbol{dE_g/dT}$ **is negative**, which means that the band gap decreases as the temperature increases.

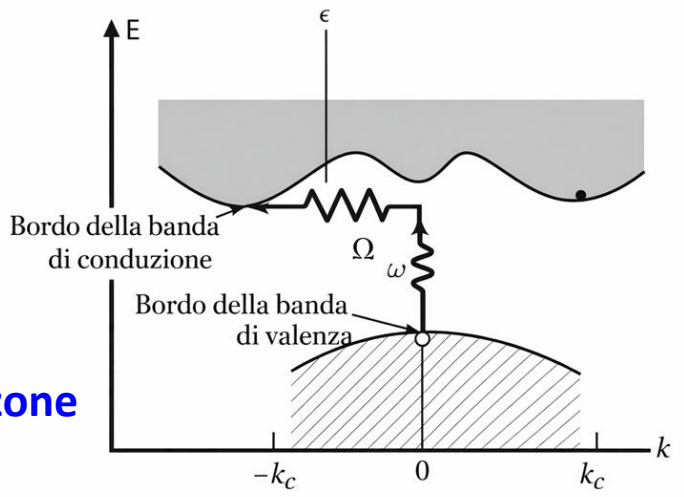
In a **direct absorption process**, a photon is absorbed by the crystal, and its energy is used to create an electron–hole pair. In this case, the optical absorption threshold at ω_g gives a direct measure of the band gap, given by $E_g = \hbar\omega_g$,



1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.4 Band Gap

In the **indirect absorption process**, on the other hand, the minimum energy gap of the band structure involves a valence band state and a conduction band state separated by a $\vec{k}_c$ wave vector, whose modulus must be compared with that of the $\vec{k}_f$ photon wave vector.



$|\vec{k}_c|$ coincides with the boundary of the **first Brillouin zone** and, in the case of silicon, is given by:

$$|\vec{k}_c| = \frac{\pi}{a} \sim 5.8 \cdot 10^7 \; cm^{-1}$$

with $a = 0.54\, nm$ the silicon lattice constant.

The **modulus of the photon's wave vector** can be estimated by the relation

$$|\vec{k}_f| = \frac{E_g}{\hbar c} \sim 5.1 \cdot 10^4 \; cm^{-1}$$

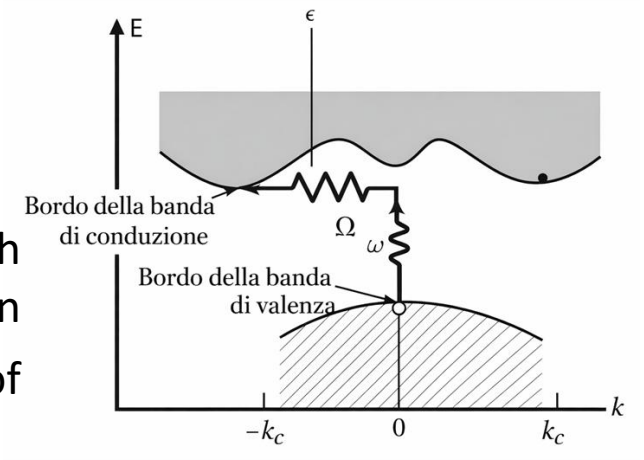
with $\hbar = 6.6 \cdot 10^{-16}\, eV \cdot s$ and $c \simeq 3.0 \cdot 10^{10}\, \mathrm{cm/s}$, and having assumed $E_g \approx 1\, eV$, a value valid for silicon at room temperature.

1.1 STRUCTURES OF SILICON, GERMANIUM AND GALLIUM ARSENIDE

1.1.4 Band Gap

$$|\vec{k}_c| = \frac{\pi}{a} \sim 5.8 \cdot 10^7\ cm^{-1}$$

$$|\vec{k}_f| = \frac{E_g}{\hbar c} \sim 5.1 \cdot 10^4\ cm^{-1}$$



It is therefore evident that a **direct photon transition** with energy equal to E_g cannot satisfy the law of conservation of the wave vector, since $|\vec{k}_f| \ll |\vec{k}_c|$, in the range of energies of interest.

However, the conservation of the crystalline momentum can be restored if, in addition to the photon, a phonon of a wave vector $\vec{K}$ and frequency Ω intervenes in the process. In this case, the laws of conservation of crystalline and photon momentum require respectively:

$$\overrightarrow{k_f} = \overrightarrow{k_c} + \vec{K}$$

$$\hbar\omega = E_g + \hbar\Omega$$

The energy of the phonon $\hbar\Omega$ is typically much smaller than E_g : even phonons with large wave vectors possess relatively low energies. For this reason, phonons represent an easily accessible source of crystalline momentum for indirect processes.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

From a microscopic point of view, at any temperature above absolute zero, thermal agitation provides energy to the electrons of the crystal to cause the **breaking of a covalent** bond between two adjacent atoms: one of the electrons involved in the bond acquires enough energy to free itself from the lattice and become a conduction electron.

From an energy point of view, this process can be schematized as the transition of an electron from the **valence band** to the **conduction band**. Each electron that is promoted in the conduction band leaves a hole in the valence band: for this reason, in an ideal intrinsic semiconductor, **the number of free electrons is always equal to the number of holes generated**.

An **intrinsic semiconductor** is defined as a material in which the concentration of impurities is negligible compared to the number of carriers created by thermal effect. Under such conditions, the electrical properties depend solely on thermally generated carriers.

To determine the concentration of electrons in the conduction band, it is necessary to evaluate, for each value of energy E, how many states are available and how likely they are to be occupied.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

The concentration of electrons included in an infinitesimal range of energy dE is given by the product:

$$n(E) = N(E)f(E)dE$$

where $N(E)$ represents the **density of allowable energy states per unit volume** and $f(E)$ is the **probability of occupation** of an energy state E.

The **total concentration of electrons in the conduction band** is obtained by integrating this expression from the lower edge of the conduction band E_C (which for convenience can be set equal to zero) to the upper limit of the band

$$n = \int_0^{E_{\text{sup}}} N(E)\, f(E)\, dE$$

The probability of occupation $f(E)$ is described by the **Fermi–Dirac distribution**,

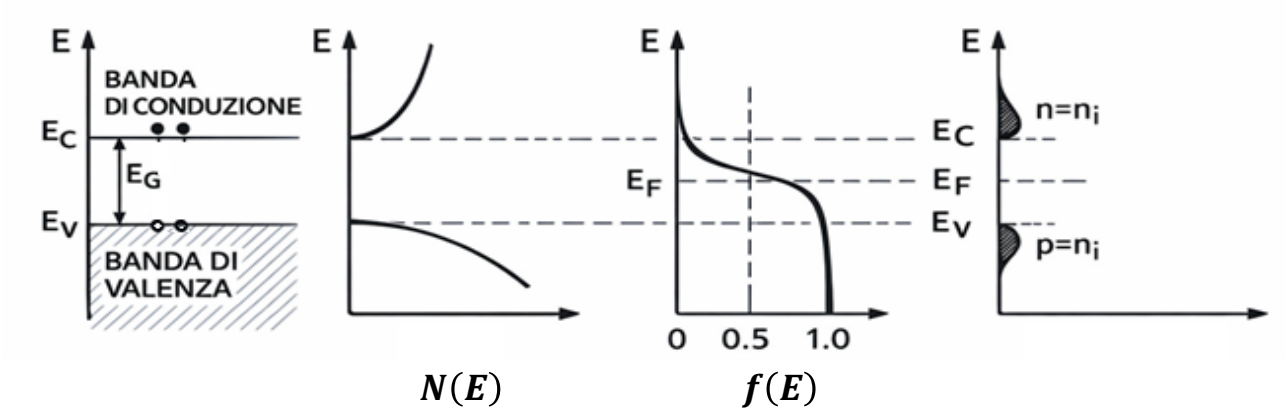
$$f(E) = \frac{1}{1 + e^{(E-E_F)/(kT)}}$$

where E_F is the Fermi level, k is the Boltzmann constant and T is the absolute temperature.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

The concentration of carriers can also be determined graphically, representing the density of the states and the distribution function and evaluating the area of the $N(E)f(E)$ product in the region of the conduction band



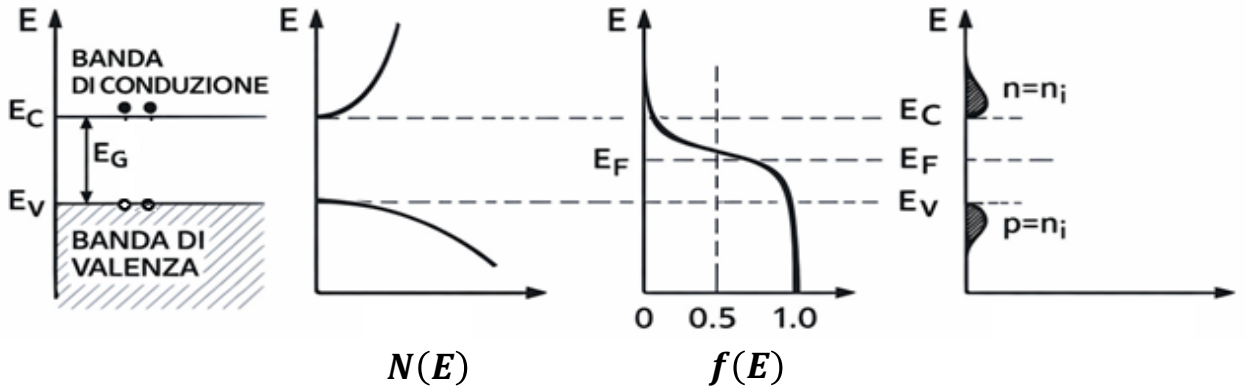
$N(E)$ $f(E)$

A very large number of permissible energy states are available in the **conduction band**. However, in an intrinsic semiconductor, the number of electrons present in this band is small, since the probability of occupying such states is very small.

In the **valence band**, on the contrary, there is also a large density of allowed states, but most of them are occupied by electrons: the probability of occupation is in fact close to unity. Consequently, in the valence band the number of empty states, i.e. holes, is limited.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers



Since the Fermi–Dirac distribution function $f(E)$ decreases exponentially as energy increases, the upper bound of the integral that gives the electron concentration can be extended to infinity without introducing significant errors.

Furthermore, in an intrinsic semiconductor the Fermi level E_F is approximately at the center of the bandgap; under such conditions, for energies sufficiently greater than E_F, the Fermi distribution can be approximated by the Boltzmann exponential form:

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \simeq e^{-(E-E_F)/kT}$$

1.2 CONCENTRAZIONE DEI PORTATORI ALL'EQUILIBRIO TERMICO

1.2.1 Portatori intrinseci

$$n = \int_0^{E_{sup}} N(E) f(E) dE$$

$$f(E) \simeq e^{-(E-E_F)/kT}$$

$$N(\mathrm{E}) = 4\pi \left(\frac{2m_e}{\hbar^2}\right)^{3/2} \sqrt{E}$$

Combining these expressions:

$$n = 4\pi \left(\frac{2m_e}{h^2}\right)^{3/2} \int_0^{\infty} \sqrt{E}\, e^{-\frac{E-E_F}{KT}} dE$$

with m_e effective mass of the electron.

Setting $x = \frac{E}{KT}$ T, the integral becomes:

$$n = 4\pi \left(\frac{2m_e}{h^2} KT\right)^{3/2} e^{\frac{E_F}{KT}} \int_0^{\infty} \sqrt{x}\, e^{-x} dx$$

The integral: $\int_0^{\infty} \sqrt{x}\, e^{-x} dx = \frac{\sqrt{\pi}}{2}$ and then:

$$n = 2 \left(\frac{2\pi m_e KT}{h^2}\right)^{3/2} e^{\frac{E_F}{KT}}$$

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

Referring now to the origin of the energy at the lower edge of the conduction band E_C, the concentration of electrons in the conduction band is:

$$n = 2\left(\frac{2\pi m_e KT}{h^2}\right)^{3/2} e^{-\frac{E_C - E_F}{KT}} = N_C e^{-\frac{E_C - E_F}{KT}}$$

$$n = 2\left(\frac{2\pi m_e KT}{h^2}\right)^{3/2} e^{\frac{E_F}{KT}}$$

where $N_C = 2\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2}$ represents the **effective density of the states in the conduction band**.

At room temperature ($T \approx 300$ K) you get the typical values: $N_C = 2.8 \times 10^{19}$ cm^{-3} for silicon and $N_C = 4.7 \times 10^{17}$ cm^{-3} for gallium arsenide.

The concentration p of the holes in the valence band is determined by a procedure similar to that used for electrons in the conduction band:

$$p = 2\left(\frac{2\pi m_h KT}{h^2}\right)^{3/2} e^{-\frac{E_F - E_V}{KT}} = N_V e^{-\frac{E_F - E_V}{KT}}$$

with m_h effective mass of holes and $N_V = 2\left(\frac{2\pi m_h KT}{h^2}\right)^{3/2}$ is the **effective density of the states in the valence band**.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

At room temperature ($T \approx 300$ K) you get the typical values: $N_V = 1.04 \times 10^{19}$ cm^{-3} for silicon and $N_V = 7.0 \times 10^{18}$ cm^{-3} for gallium arsenide.

In an **intrinsic semiconductor**, by definition, the number of electrons in the conduction band is equal to the number of holes in the valence band. The fundamental relationship therefore applies:

$$n = p = n_i$$

where n_i is the concentration of intrinsic carriers.

So equalizing the two expressions for n and p, we have:

$$N_C e^{-\frac{E_C - E_F}{KT}} = N_V e^{-\frac{E_F - E_V}{KT}}$$

$$n = N_C e^{-\frac{E_C - E_F}{KT}}$$

$$p = N_V e^{-\frac{E_F - E_V}{KT}}$$

From this equation we derive the position of the **intrinsic Fermi level**:

$$E_F = \frac{E_V + E_C}{2} + \frac{kT}{2} \ln \frac{N_V}{N_C}$$

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

$$E_F = \frac{E_V + E_C}{2} + \frac{KT}{2} ln \frac{N_V}{N_C}$$

$$N_C = 2\left(\frac{2\pi m_e KT}{h^2}\right)^{3/2}$$

$$N_V = 2\left(\frac{2\pi m_h KT}{h^2}\right)^{3/2}$$

By substituting the expressions previously determined for N_C and N_V as a function of the effective masses:

$$E_F = \frac{E_V + E_C}{2} + \frac{3KT}{4} ln \frac{m_h}{m_e}$$

Since, at room temperature, the second term is much smaller than the first, the intrinsic Fermi level E_i is placed practically in **the middle of the bandgap**.

The concentration of intrinsic carriers is finally determined by the relationship

$$n = p = n_i$$

$$np = n_i^2$$

leading to:

$$n_i = \sqrt{N_C N_V}\, e^{-E_g/2kT}$$

where $E_g = E_C - E_v$.

$$n = N_C e^{-\frac{E_C - E_F}{KT}}$$

$$p = N_V e^{-\frac{E_F - E_V}{KT}}$$

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.1 Intrinsic carriers

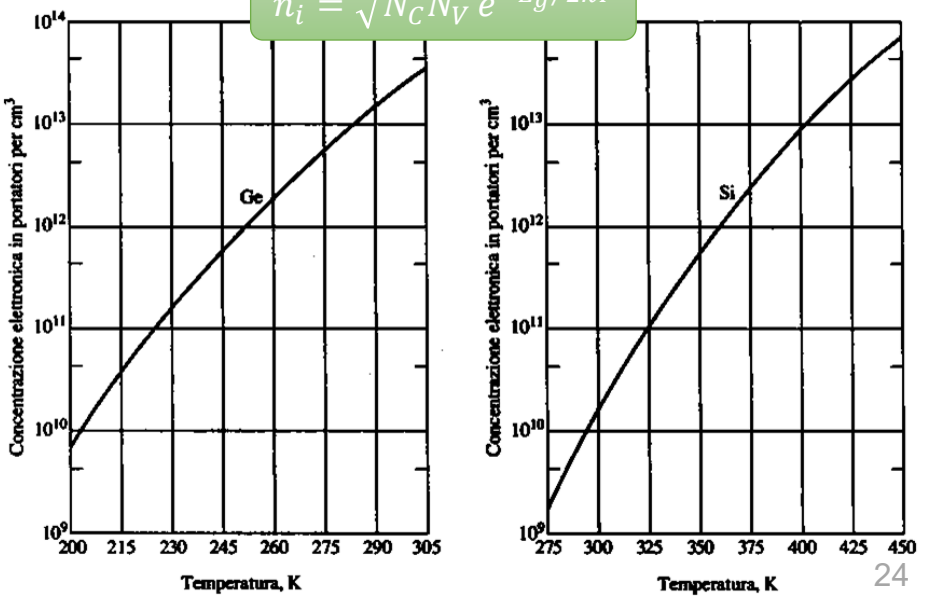
$$np = n_i^2$$

This relationship is called the **mass action law** and is valid under thermodynamic equilibrium conditions for both intrinsic and extrinsic semiconductors.

In a doped semiconductor, increasing the concentration of one type of carrier results in a decrease in the concentration of the other type due to recombination processes, but **the np product remains constant for a given temperature**.

The Figure shows the trend of n_i as a function of T for Ge and Si. In both cases, n_i grows rapidly with T.

$$n_i = \sqrt{N_C N_V}\, e^{-E_g/2kT}$$



At the same temperature, Ge has significantly higher n_i values than Si, as a result of its smaller energy gap

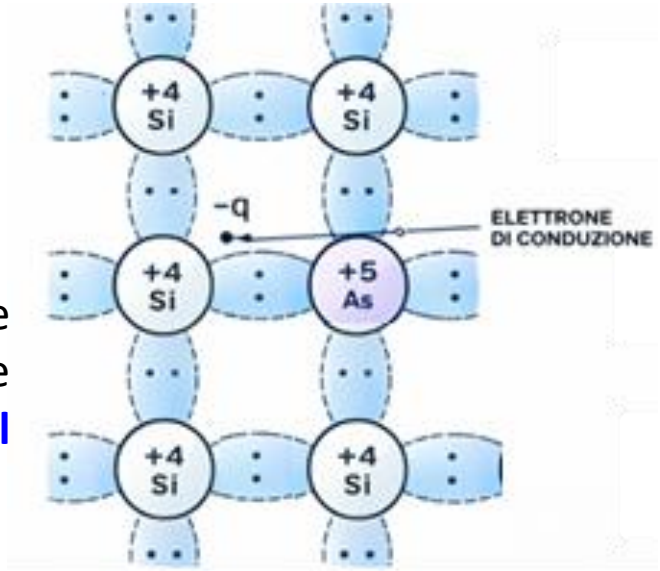
1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.2 Donors and acceptors

When small amounts of impurities are intentionally introduced into a pure semiconductor, the material becomes an **extrinsic**, or **doped semiconductor**.

The main effect of doping is the introduction of **energy levels of impurities within the bandgap**, which can significantly change the concentration of charge carriers.

First consider the case in which a silicon atom is replaced by an **arsenic atom**, which has five valence electrons. The arsenic atom forms four covalent bonds with adjacent silicon atoms, using four of its valence electrons.



The fifth electron is weakly bound to the impurity atom and can be easily promoted in the conduction band. This generates an **additional free electron in the conduction band**.

The semiconductor obtained is called **n-type**, since the majority carriers are electrons, and the arsenic atom is called a **donor**.

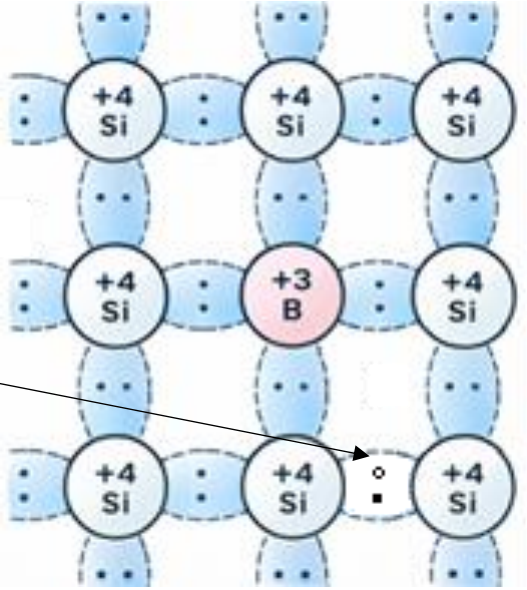
1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.2 Donors and acceptors

Similarly, if a silicon atom is replaced by a **boron** atom, which has only three valence electrons, an electron from a neighboring atom is "accepted" to complete the four covalent bonds required.

This process leaves a hole in the valence band, which behaves as a carrier of positive charge.

The resulting semiconductor is called **p-type**, and the boron atom is called an **acceptor**.



A simple way to estimate the energy levels associated with impurities can be obtained using the **Bohr hydrogen atom model**.

$$E_H = -\frac{mZ^2e^4}{(4\pi\varepsilon_0)^2 2\hbar^2}\frac{1}{n^2}$$

In the semiconductor, the electron bound to the donor moves in a medium characterized by an effective mass m_e and a dielectric permittivity ε_s.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.2 Donors and acceptors

Substituting these parameters for those of the hydrogen atom, we obtain for the ionization energy of the donor

$$E_H = -\frac{mZ^2e^4}{(4\pi\varepsilon_0)^2 2\hbar^2}\frac{1}{n^2}$$

$$E_D = \left(\frac{\varepsilon_0}{\varepsilon_s}\right)^2 \left(\frac{m_e}{m_0}\right) E_H$$

For shallow donors, the ionization energy measured from the edge of the conduction band is about 0.025 eV in silicon and 0.007 eV in GaAs.

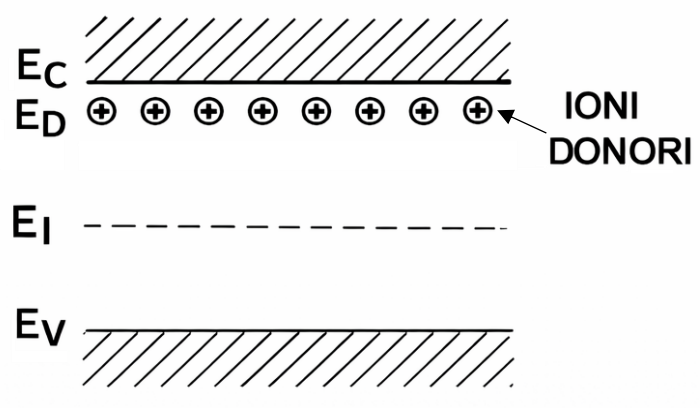
In the case of **acceptor atoms**, the procedure is quite similar. One can interpret the near-full valence band as a complete band plus a hole subject to the Coulomb field of negatively charged impurity. The ionization energy thus obtained, measured from the edge of the valence band, is about 0.05 eV in silicon.

For shallow donor and acceptor atoms in Si and GaAs, the thermal energy at room temperature is sufficient to ionize virtually all impurities. Under such conditions, we refer to **complete ionization**: each donor atom provides an electron to the conduction band, and each acceptor atom generates a hole in the valence band, directly determining the concentration of carriers in the doped semiconductor.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

Under conditions of **complete ionization**, all donor atoms are ionized and provide an electron to the conduction band. In this case, the concentration of electrons can be approximated simply by:



$$n = N_D$$

where N_D is the **concentration of donors**.

Substituting the expression for the concentration of electrons, we get:

$$n = N_C e^{-\frac{E_C - E_F}{KT}}$$

$$E_C - E_F = KT ln\left(\frac{N_C}{N_D}\right)$$

This relationship shows that, as the concentration of the donors increases, the energy difference $E_C - E_F$ decreases and the **Fermi level shifts towards the lower edge of the conduction band**.

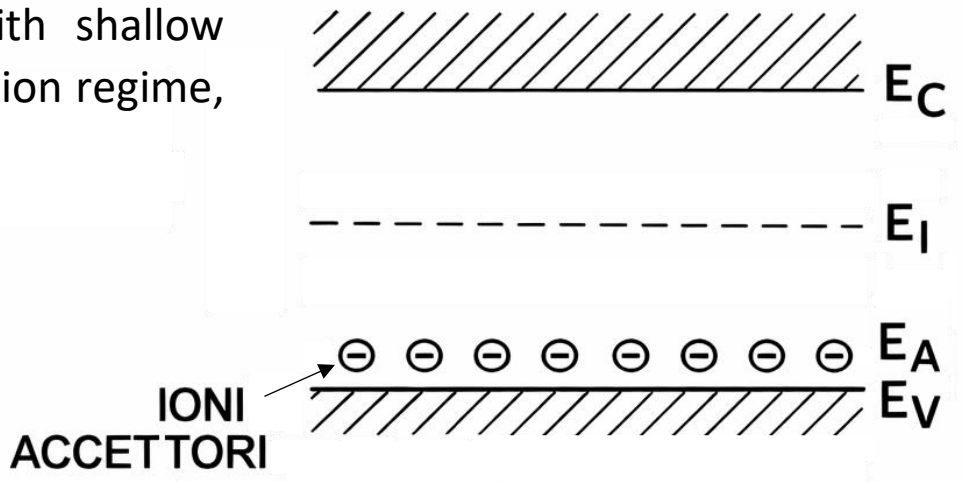
1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

Similarly, for a p-type semiconductor with shallow acceptor atoms and in the complete ionization regime, the concentration of the holes is:

$$p = N_A$$

where N_A is the concentration of acceptors.



Similarly as before, substituting the expression for the concentration of holes:

$$p = N_V e^{-\frac{E_F - E_V}{KT}}$$

$$E_F - E_V = KT ln\left(\frac{N_V}{N_A}\right)$$

Therefore, the higher the density of acceptor atoms, the closer **the Fermi level is to the upper edge of the valence band**.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

According to the **mass action law**, under conditions of thermodynamic equilibrium the product of the concentration of electrons n and that of holes p must coincide with the value that occurs in the intrinsic semiconductor.

For this reason, it is often convenient to express the concentrations of electrons and holes as a function of the **intrinsic concentration** n_i and the **intrinsic Fermi level** E_i.

$$np = n_i^2$$

Starting from the expression found for an extrinsic semiconductor, we add and subtract E_i to the numerator of the exponential :

$$n = N_C e^{-\frac{E_C - E_F}{KT}}$$

$$n_i = N_C e^{-\frac{E_C - E_i}{KT}}$$

$$n = N_C e^{-\frac{E_C - E_i}{KT}} e^{\frac{E_F - E_i}{KT}} = n_i e^{\frac{E_F - E_i}{KT}}$$

Similarly for holes:

$$p = N_V e^{-\frac{E_F - E_V}{KT}}$$

$$p_i = N_V e^{-\frac{E_i - E_V}{KT}}$$

$$p = N_V e^{-\frac{E_F - E_V}{KT}} = n_i e^{\frac{E_i - E_F}{KT}}$$

These expressions highlight how the concentrations of the carriers depend directly on the shift of the Fermi level with respect to its intrinsic value.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

When donor and acceptor impurities are present in a semiconductor at the same time, the type of conductivity is determined by the type of impurities present in the highest concentration.

The Fermi level is arranged in such a way as to guarantee the **global electrical neutrality** of the material: the total negative charge, due to the free electrons and ionized acceptors, must equal the total positive charge, associated with the holes and ionized donors. This condition is expressed by the equation

$$n + N_A = p + N_D$$

$$np = n_i^2$$

Combining it with the mass action law , we obtain the general expressions for carrier concentrations in an n-type semiconductor:

$$n_n = \frac{1}{2}\left[N_D - N_A + \sqrt{(N_D - N_A)^2 + 4n_i^2}\right], \quad p_n = \frac{n_i^2}{n_n}$$

The subscript n indicates that it refers to an n-type semiconductor. In this case, **electrons** are the **majority carriers**, while **holes** are the **minority carriers**.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

In a very similar way, for a p-type semiconductor we obtain:

$$p_p = \frac{1}{2}\left[N_A - N_D + \sqrt{(N_A - N_D)^2 + 4n_i^2}\right], \quad n_p = \frac{n_i^2}{p_p}$$

where the subscript p indicates the semiconductor of type p. In this case, the **holes** are the **majority carriers**, while the **electrons** are the **minority carriers**.

In most practical applications, the net concentration of impurities $|N_A - N_D|$ is much greater than the intrinsic concentration n_i. Under these conditions, the expressions for n_n and p_p are greatly simplified.

$$n_n \simeq N_D - N_A \qquad se \;\; N_D > N_A$$
$$p_p \simeq N_A - N_D \qquad se \;\; N_A > N_D$$

$$n_n = \frac{1}{2}\left[N_D - N_A + \sqrt{(N_D - N_A)^2 + 4n_i^2}\right]$$

By combining these expressions found for n_n and p_p, with those found previously, it is possible to determine the **position of the Fermi level** as a function of **temperature** for a **given concentration of donors or acceptors**.

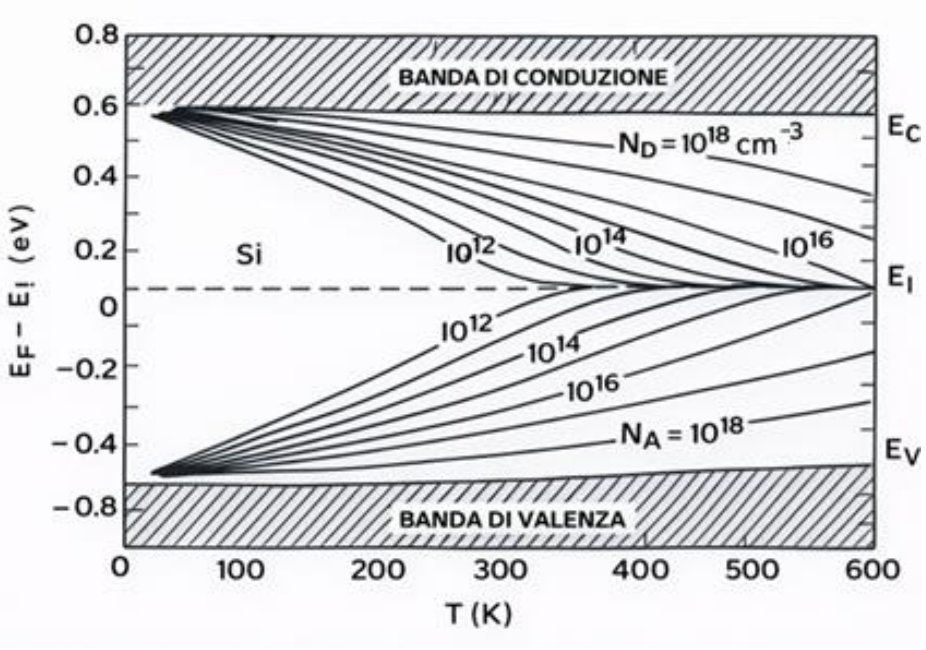
$$n_n = n_i e^{\frac{E_F - E_i}{KT}}$$
$$p_p = n_i e^{\frac{E_i - E_F}{KT}}$$

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

The Figure shows the trend of the position of the Fermi E_F level in silicon as a function of temperature for different concentrations of donor N_D e acceptor impurities N_A.

The vertical axis shows the difference $E_F - E_i$, i.e. the shift of the Fermi level with respect to the intrinsic E_i level, while the horizontal axis shows the temperature.



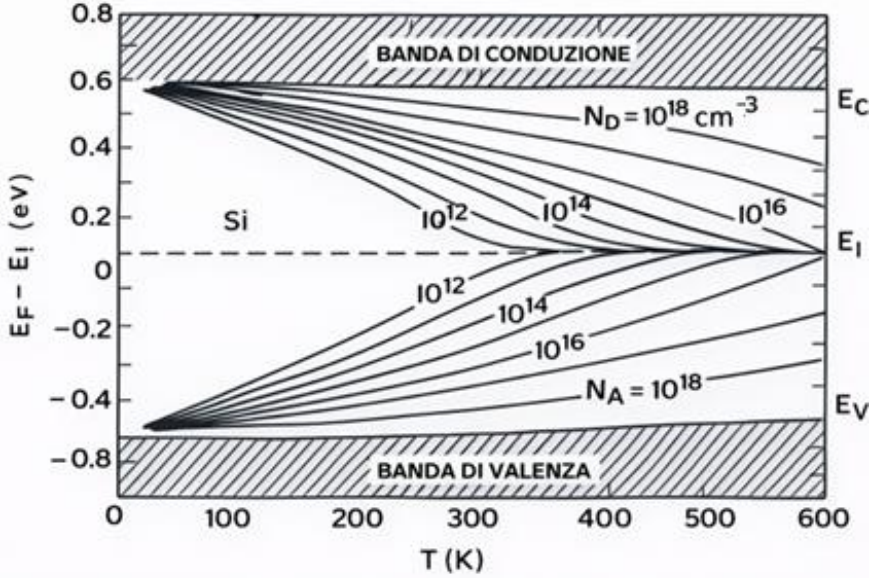
The dashed regions at the top and bottom of the Figure represent the conduction band and the valence band, respectively, bounded by the E_C and E_V edges. Their trend as the temperature varies highlights the slight dependence of the width of the band gap on temperature.

The central horizontal line corresponds to the intrinsic Fermi level E_i.

1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

At **low temperatures**, only a fraction of the impurities are ionized, and the Fermi level is close to the impurity levels. In the case of n-type semiconductors, E_F is near the lower edge of the conduction band, while in the case of p-type it is near the upper edge of the valence band.



As the temperature increases, the extrinsic regime is entered, in which the complete ionization of impurities takes place: the position of the Fermi level depends essentially on the doping concentration: the higher the N_D (or N_A), the closer E_F it approaches E_C (or E_V).

The different curves labeled with 10^{12}, 10^{14}, 10^{16}, 10^{18} cm^{-3} clearly show how an increase in the concentration of impurities produces an increasingly marked shift of the Fermi level towards the band corresponding to the majority carriers.

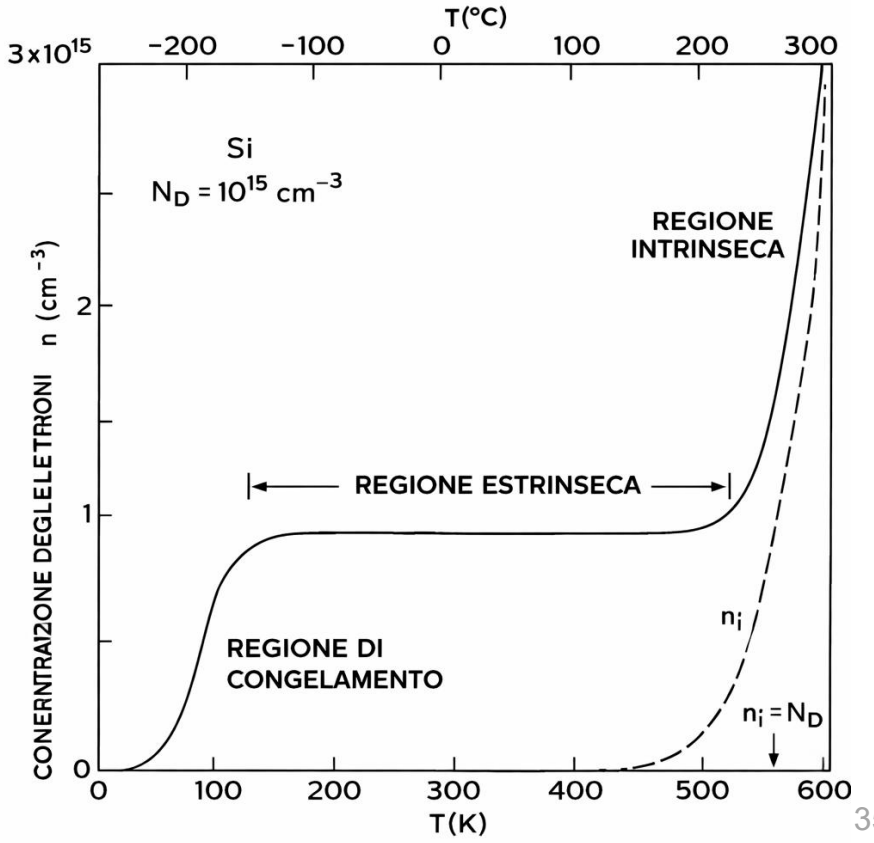
1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

The purely thermal behavior of an n-type semiconductor is illustrated in the Figure, which shows the trend of the **electron concentration as a function of temperature** for a silicon crystal doped with a donor concentration equal to $N_D = 10^{15}\ cm^{-3}$.

At **low temperatures**, the thermal energy available is not sufficient to completely ionize the donor impurities: a part of the electrons remains bound to the levels of the donors and the concentration of free electrons is therefore lower than the density of the donors.

This regime is called the **freezing region**.

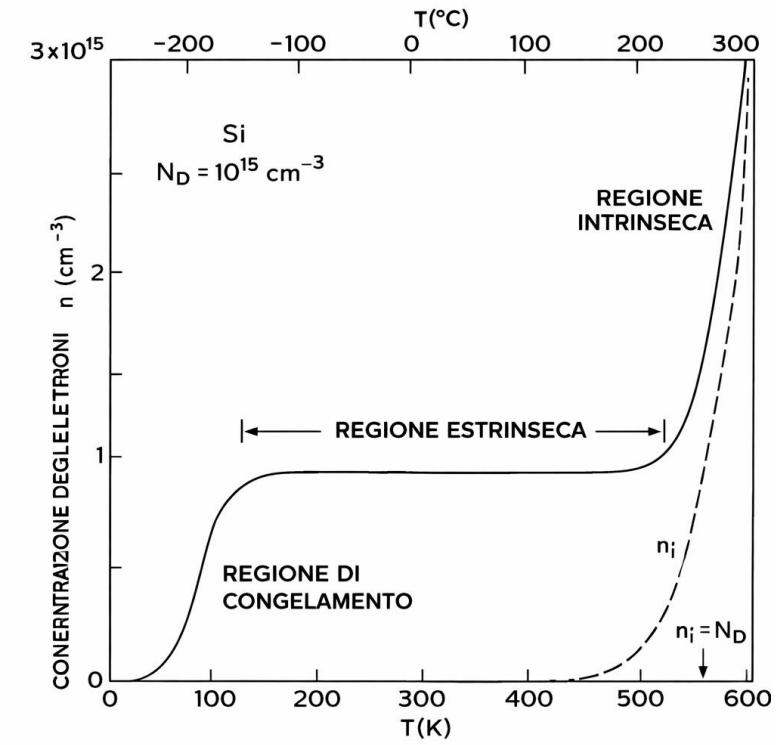


1.2 CONCENTRATION OF CARRIERS AT THERMAL EQUILIBRIUM

1.2.3 Extrinsic carrier concentration

As the temperature increases, the crystal enters the **extrinsic region**, in which almost all donors are ionized. In this temperature range, the concentration of electrons remains almost constant and is approximately equal to N_D.

By further increasing the temperature, the thermal generation of electron-hole pairs becomes progressively more relevant, as can be seen from the dashed curve of n_i as a function of temperature.



Finally, a temperature is reached for which the concentration of the intrinsic carriers n_i becomes comparable with the concentration of the donors: beyond this value, the behavior of the material is dominated by thermally generated carriers, and the semiconductor enters the **intrinsic regime**.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.1 Drift and mobility

Let us consider an n-type semiconductor sample, characterized by a uniform concentration of donors and under condition of **thermal equilibrium**.

As already discussed, the electrons in the conduction band can be treated, in good approximation, as **quasi-free particles**. The effect of the crystal lattice is in fact included in the **effective mass** m_n, which is slightly different from the mass of the free electron.

At thermal equilibrium, the average energy associated to the motion of electrons is determined by the **energy equipartition theorem**, according to which each degree of freedom corresponds to an average energy equal to $\frac{1}{2}kT$, where k is Boltzmann's constant and T is the absolute temperature.

Since the motion of electrons in space is three-dimensional, they possess three translational degrees of freedom; consequently, their average kinetic energy is:

$$\frac{1}{2}m_n v_{th}^2 = \frac{3}{2}kT$$

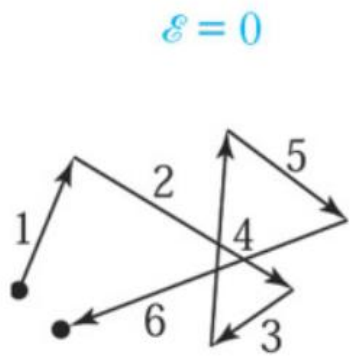
m_n is the effective mass of the electrons and v_{th} is the average thermal velocity.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.1 Drift and mobility

At **room temperature** ($T \simeq 300$ K), the thermal velocity of electrons is of the order of 10^7 cm/s in both silicon and gallium arsenide.

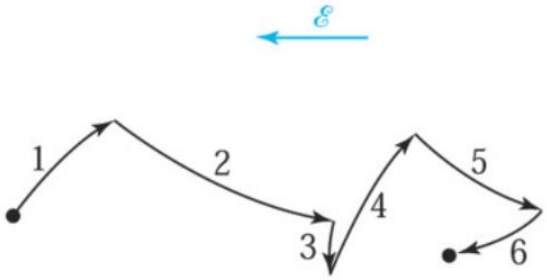
This means that, even in the absence of external fields, electrons move very quickly and in a disorderly manner, in all directions. The motion of a single electron can be described as a succession of **random collisions** with the atoms of the lattice, with the atoms of impurities and with other centers of dispersion, as in the Figure.



Because of this random motion, the average displacement of an electron, over sufficiently long time intervals, is zero.

The average distance traveled between two successive collisions is called the **mean free path**, while the average time between two collisions is called the **mean free time** τ_c, typically of the order of the picosecond.

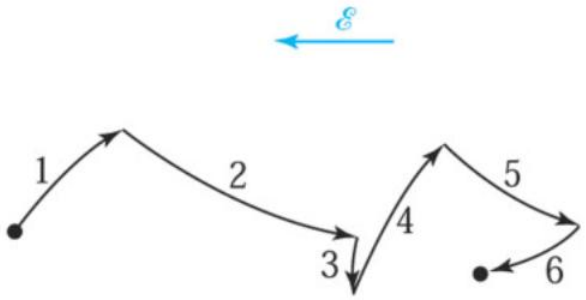
When an electric field $\vec{\mathcal{E}}$ is applied to the semiconductor, each electron is subjected to the force $-q\vec{\mathcal{E}}$ and is accelerated, between one collision and the next, along the direction of the field but in the opposite direction.



1.3 CARRIER TRANSPORT PHENOMENA

1.3.1 Drift and mobility

At this point, a small oriented velocity component, called **drift velocity** v_n , is superimposed on the random thermal motion. The overall motion of the electron thus results from the combination of the disordered thermal motion and the drift velocity, producing an **average net displacement in the opposite direction to the applied electric field**.



The drift velocity v_n can be determined by comparing the momentum acquired by the electron under the action of the force $-q\vec{\mathcal{E}}$ during the mean free time with the momentum gained by the electron over the same period, according to the **impulse-momentum theorem**, which states that the change in the momentum of a point mass in a time interval is equal to the momentum of the resultant force acting on it in the same interval.

This is justified by the fact that, in steady state, the accumulated momentum between two collisions is completely transferred to the lattice during each collision.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.1 Drift and mobility

The impulse given to an electron by the electric field during the mean free time τ_c is equal to $-q\vec{\mathcal{E}}\tau_c$ while the corresponding change in momentum is $m_n\vec{v}_n$.

Applying the impulse-momentum theorem, we then obtain the relation

$$-q\vec{\mathcal{E}}\tau_c = m_n\overrightarrow{v_n}$$

from which:

$$\overrightarrow{v_n} = -\left(\frac{q\tau_c}{m_n}\right)\vec{\mathcal{E}}$$

This expression shows that the drift velocity of electrons is directly proportional to the applied electric field. The proportionality coefficient

$$\mu_n = \frac{q\tau_c}{m_n}$$

is called **electron mobility** and is determined by the mean free time between collisions and the effective mass of the electron.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.1 Drift and mobility

$$\overrightarrow{v_n} = -\left(\frac{q\tau_c}{m_n}\right)\vec{\mathcal{E}} \qquad \mu_n = \frac{q\tau_c}{m_n}$$

In a compact form:

$$\overrightarrow{v_n} = -\mu_n \vec{\mathcal{E}}$$

which corresponds to a current density, in one-dimensional case, equal to:

$$J_n = qnv_n = q\mu_n n\mathcal{E}$$

In a similar way, for the holes in the valence band, we obtain

$$\overrightarrow{v_p} = \mu_p \vec{\mathcal{E}}$$

where μ_p is the **hole mobility**.

Carrier mobility is closely linked to the mean free time between collisions, which in turn depends on the different scattering mechanisms present in the semiconductor.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.1 Drift and mobility

The two main mechanisms are **lattice scattering** and **impurity scattering**.

Lattice scattering is due to the **thermal vibrations of the lattice atoms**, present at any temperature above absolute zero. These vibrations perturb the periodic potential of the crystal and allow the transfer of energy between the carriers and the lattice. Since the amplitude of vibrations increases with temperature, this mechanism becomes dominant at high temperatures. As a result, the mobility of the carriers decreases as the temperature increases.

Diffusion by impurities, on the other hand, occurs when a carrier passes in proximity of an ionized doping atom, donor or acceptor. In this case, the motion of the carrier is deflected by the **Coulomb interaction** with the impurity. The probability of this type of scattering depends on the total concentration of ionized impurities.

Unlike lattice scattering, impurity scattering becomes less effective at high temperatures: in fact, at high temperatures the carriers move more quickly, remain for a shorter time near the impurity and are therefore less probability to be scattered.

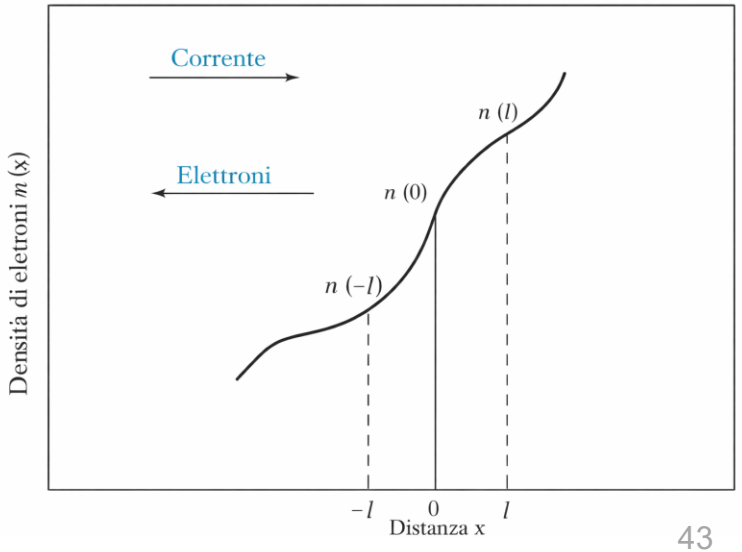
1.3 CARRIER TRANSPORT PHENOMENA

1.3.2 Diffusion processes

In the previous section, drift current, i.e. the transport of carriers due to the application of an electric field, was analyzed.

However, there is another important current component that occurs even in the absence of external fields, when there is a **spatial variation in the concentration of carriers** in the semiconductor. In these conditions, carriers tend to spontaneously move from regions with higher concentrations to those with lower concentrations: this phenomenon is called **diffusion**, and the associated current is called **diffusion current**.

To describe the process, let us suppose that the electron density varies along the x -direction, as in the Figure.



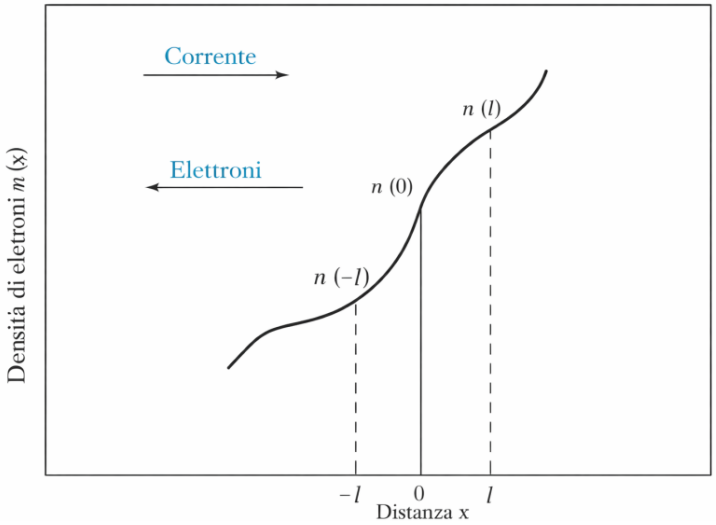
The semiconductor is kept at a uniform temperature, so that the average thermal energy of the electrons does not depend on x; instead, only the density $n(x)$ varies.

Consider the number of electrons crossing the plane at $x = 0$ per unit time and per unit area.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.2 Diffusion processes

Due to thermal motion, electrons move randomly with a thermal velocity v_{th} and travel on average a distance $l = v_{th}\tau_c$ between two successive collisions, where τ_c is the average free time.



Electrons in the $x = -l$ region have the same probability of moving to the right or to the left; therefore, in a time interval τ_c , about half of them cross the plane $x = 0$.

The average F_1 flux of electrons coming from the left is therefore

$$F_1 = \frac{\frac{1}{2}n(-l)l}{\tau_c} = \frac{1}{2}n(-l)v_{th}$$

Similarly, the average F_2 flow of electrons crossing the plane from the right is:

$$F_2 = \frac{1}{2}n(l)v_{th}$$

1.3 CARRIER TRANSPORT PHENOMENA

1.3.2 Diffusion processes

$$F_1 = \frac{1}{2} n(-l) v_{th} \qquad F_2 = \frac{1}{2} n(l) v_{th}$$

The net flow F of carriers from left to right will be:

$$F = F_1 - F_2 = \frac{1}{2} v_{th} [n(-l) - n(l)]$$

Expanding $n(\pm l)$ in a Taylor series around $x = 0$ and keeping the first terms, we have:

$$F = \frac{1}{2} v_{th} \left\{ \left[n(0) - l \frac{dn}{dx} \right] - \left[n(0) + l \frac{dn}{dx} \right] \right\}$$

from which follows:

$$F = -v_{th} l \frac{dn}{dx} = -D_n \frac{dn}{dx}$$

where the quantity $D_n = v_{th} l$ is the **diffusion coefficient**, or **diffusivity**, of the electrons.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.2 Diffusion processes

$$F = -v_{th} l \frac{dn}{dx} = -D_n \frac{dn}{dx}$$

$$D_n = v_{th} l$$

Since each electron has charge $-q$, the flow of particles is associated with a **diffusion current density** J_n :

$$J_n = -qF = qD_n \frac{dn}{dx}$$

This expression shows that the diffusion current is **proportional to the spatial gradient of the electron density** and originates from the random thermal motion of the carriers in the presence of a concentration gradient.

If the electron density increases with x, the dn/dx gradient is positive and the electrons diffuse in the negative direction along the x-axis. The conventional current is therefore positive and flows in the opposite direction to that of the motion of the electrons.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.3 Equation for Current Density

$$J_n = qD_n \frac{dn}{dx}$$

$$D_n = v_{th} l$$

$$\mu_n = \frac{q\tau_c}{m_n}$$

$$\frac{1}{2} m_n v_{th}^2 = \frac{1}{2} kT$$

The expression of the electron current density J_n can be further reformulated using the **energy equipartition theorem** in the one-dimensional case.

In this way the diffusivity of D_n electrons can be related to their mobility μ_n :

$$D_n = v_{th} l = v_{th}(v_{th}\tau_c) = v_{th}^2 \left(\frac{\mu_n m_n}{q}\right) = \left(\frac{kT}{m_n}\right)\left(\frac{\mu_n m_n}{q}\right)$$

from which

$$D_n = \left(\frac{kT}{q}\right)\mu_n$$

This important relationship is known as the **Einstein relation** and relates two fundamental quantities of carrier transport: diffusivity, which governs diffusion transport, and mobility, which characterizes drift transport.

A similar relationship applies to holes.

$$D_p = \frac{kT}{q}\,\mu_p$$

1.3 CARRIER TRANSPORT PHENOMENA

1.3.3 Equation for Current Density

When both an **electric field** and a **concentration gradient** are present in a semiconductor at the same time, both the drift current and the diffusion current contribute to the transport. The total electron current density is therefore the sum of the two contributions:

$$J_n = q\mu_n n\mathcal{E} + qD_n \frac{dn}{dx}$$

$J_n = q\mu_n n\mathcal{E}$

$J_n = qD_n \frac{dn}{dx}$

A similar expression also applies to holes:

$$J_p = q\mu_p p\mathcal{E} - qD_p \frac{dp}{dx}$$

The negative sign in the diffusive term of the holes derives from the fact that, in the presence of a concentration gradient, the holes diffuse towards regions of lower density, i.e. in the negative direction of the x-axis; the associated conventional current then also flows in the same direction.

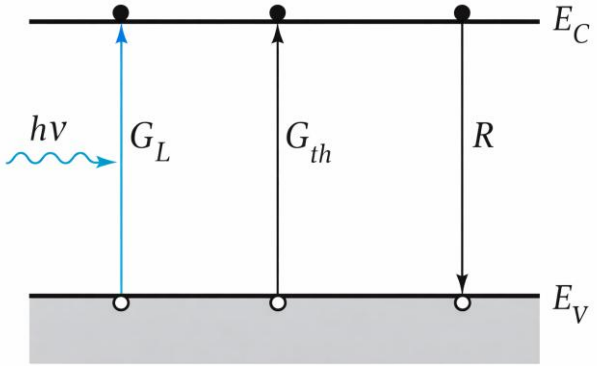
Finally, the **total conduction current density** in the semiconductor is given by the sum of the electron currents and holes:

$$J_{cond} = J_n + J_p$$

1.3 CARRIER TRANSPORT PHENOMENA

1.3.4 Direct recombination

Consider a direct-band semiconductor, under conditions of thermal equilibrium. This means that thermal energy can provide an electron in the valence band with the energy needed to make a transition to the conduction band, leaving a hole in the valence band.



This process is called **carrier generation** and is described by the **thermal generation rate** G_{th}, defined as the number of electron-hole pairs generated per unit volume and per unit time.

The reverse process occurs when an electron of the conduction band relaxes into the valence band, annihilating an electron-hole pair. This phenomenon is called **recombination** and is characterized by the **rate of thermal recombination** R_{th}.

Under conditions of thermal equilibrium, the concentration of the carriers must remain constant over time. This requires that the generation rate and the recombination rate balance exactly, i.e.:

$$G_{th} = R_{th}$$

so that the mass action law is always satisfied.

$$np = n_i^2$$

1.3 CARRIER TRANSPORT PHENOMENA

1.3.4 Direct recombination

When **excess carriers** are introduced into a direct-band semiconductor, the probability of electrons and holes recombining directly increases.

From the phenomenological point of view, the rate of direct recombination R is proportional to the product of the concentration of electrons in the conduction band and that of holes in the valence band:

$$R = \beta np$$

with β constant of proportionality called the **recombination coefficient**.

At thermal equilibrium, the rate of recombination must be equal to the rate of generation. For an n-type semiconductor one can therefore write

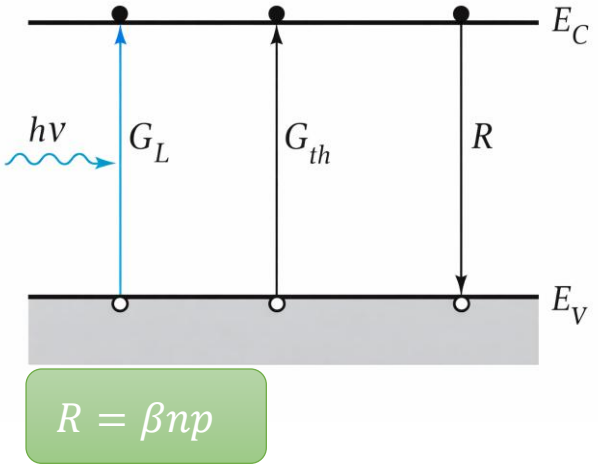
$$G_{th} = R_{th} = \beta n_{no} p_{no}$$

In this notation, the first subscript indicates the type of semiconductor, while the subscript o identifies quantities evaluated under equilibrium conditions. The quantities n_{no} and p_{no} are the electron and hole densities in an n-type semiconductor at thermal equilibrium, respectively.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.4 Direct recombination

When a semiconductor is illuminated, electron-hole pairs are generated at a G_L rate, so carrier concentrations are higher than equilibrium values.



The generation and recombination rates become:

$$R = \beta n_n p_n = \beta (n_{no} + \Delta n)(p_{no} + \Delta p)$$

$$G = G_L + G_{th}$$

where Δn and Δp represent the concentrations of excess carriers. To maintain **charge neutrality**, the following relation must be $\Delta n = \Delta p$.

The temporal variation of the concentration of holes in an n-type semiconductor is therefore described by the rate equation:

$$\frac{dp_n}{dt} = G - R = G_L + G_{th} - R$$

In steady-state conditions $\frac{dp_n}{dt} = 0$, and consequently

$$G_L = R - G_{th} \equiv U$$

where U represents the net recombination rate.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.4 Direct recombination

Substituting the expressions found before for R and G_{th} :

$$U = R - G_{th}$$

$$U = \beta(n_{no} + \Delta n)(p_{no} + \Delta p) - \beta n_{no} p_{no} = \beta(n_{no}\Delta p + p_{no}\Delta n + \Delta n \Delta p)$$

Since $\Delta n = \Delta p$

$$U = \beta(n_{no} + p_{no} + \Delta p)\Delta p$$

$$R = \beta(n_{no} + \Delta n)(p_{no} + \Delta p)$$

$$G_{th} = R_{th} = \beta n_{no} p_{no}$$

In the case of **low levels of injection** $p_{no} \ll n_{no}$, typical of n-type semiconductors, the equation is simplified:

$$U \approx \beta n_{no} \Delta p = \frac{p_n - p_{n0}}{\frac{1}{\beta n_{no}}}$$

It is therefore observed that the net rate of recombination is **proportional to the concentration of excess minority carriers**. At thermal equilibrium, when $\Delta p = 0$, $U = 0$ is obviously obtained.

The constant of proportionality $\tau_p = \frac{1}{\beta n_{no}}$ is called **lifetime of the excess minority carriers** and represents the characteristic time with which the excess holes are removed by recombination.

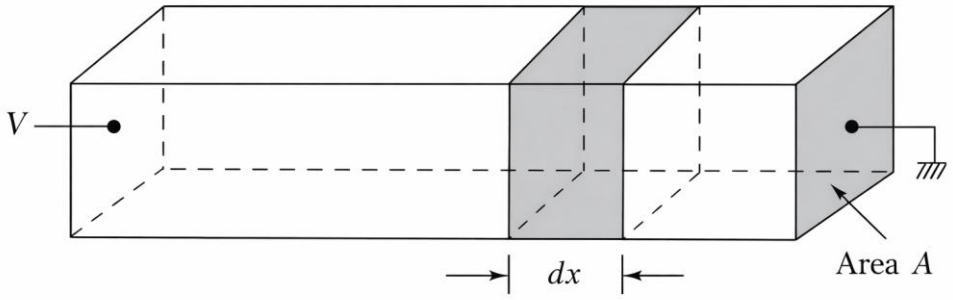
1.3 CARRIER TRANSPORT PHENOMENA

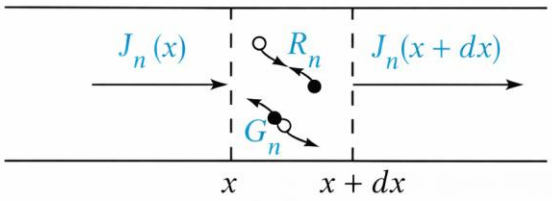
1.3.5 Continuity Equation

In the previous sections we have separately analyzed the main transport and transformation mechanisms of carriers: drift in the presence of an electric field, diffusion due to a concentration gradient and recombination.

Now consider the general case in which these phenomena act simultaneously within a semiconductor material. The equation that describes the preservation of carriers under such conditions is called the **continuity equation**.

To derive the one-dimensional continuity equation for electrons, consider an infinitesimal layer of thickness dx centered at the x position, as shown in Figure.



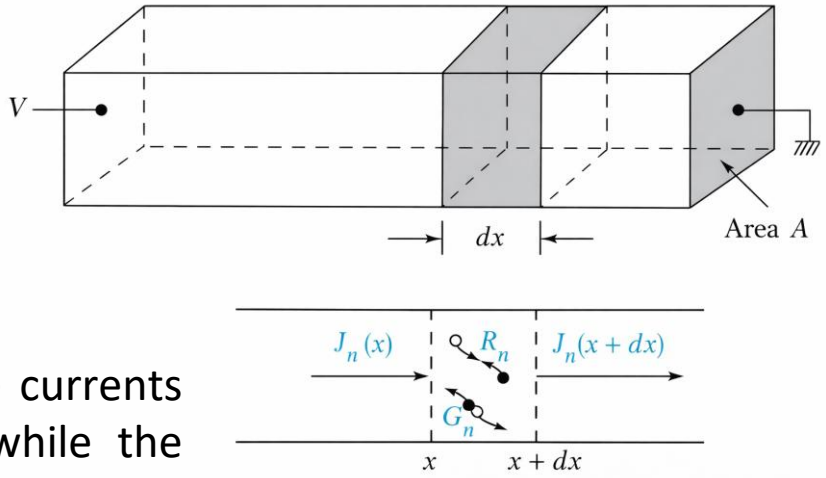


The number of electrons contained in this volume element can vary for two main reasons: due to the net flow of current that enters and exits the volume and due to the generation and recombination of carriers within the volume itself.

1.3 CARRIER TRANSPORT PHENOMENA

1.3.5 Continuity Equation

The overall change in the number of electrons is therefore given by the sum of four contributions: the number of electrons entering side x, minus the number exiting at side $x + dx$, plus the number of electrons generated, minus the number recombining in the thickness dx.



The first two terms are obtained by dividing the currents on the two sides by the elementary charge, while the generation and recombination rates are denoted by G_n and R_n, respectively.

Denoting the cross-section of the volume with A, we can write the **rate of change of the number of electrons** contained in the elementary volume of section A and thickness dx as:

$$\frac{\partial n}{\partial t} A dx = \left[\frac{J_n(x)A}{-q} - \frac{J_n(x + dx)A}{-q} \right] + (G_n - R_n) A dx$$

1.3 CARRIER TRANSPORT PHENOMENA

1.3.5 Continuity Equation

$$\frac{\partial n}{\partial t}Adx = \left[\frac{J_n(x)A}{-q} - \frac{J_n(x+dx)A}{-q}\right] + (G_n - R_n)Adx$$

By expanding the current density at the point $x + dx$ in Taylor series, $J_n(x+dx) = J_n(x) + \frac{\partial J_n}{\partial x}dx$, and simplifying, we obtain the continuity equation for electrons:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial x} + (G_n - R_n)$$

$$J_n = q\mu_n n\mathcal{E} + qD_n\frac{dn}{dx}$$

We then replace the expressions found above.

$$U = \frac{p_n - p_{n0}}{\tau_p}$$

Assuming $G_{th} \approx 0$, and then $U = R_n = \frac{p_n - p_{n0}}{\tau_p}$, the continuity equations for minority carriers are obtained.

$$G_L = R_n - G_{th} \equiv U$$

For electrons in a p-type semiconductor:

$$\frac{\partial n_p}{\partial t} = n_p\mu_n\frac{\partial \mathcal{E}}{\partial x} + \mu_n\mathcal{E}\frac{\partial n_p}{\partial x} + D_n\frac{\partial^2 n_p}{\partial x^2} + G_n - \frac{n_p - n_{p0}}{\tau_n}$$

1.3 CARRIER TRANSPORT PHENOMENA

1.3.5 Continuity Equation

$$\frac{\partial n_p}{\partial t} = n_p \mu_n \frac{\partial \mathcal{E}}{\partial x} + \mu_n \mathcal{E} \frac{\partial n_p}{\partial x} + D_n \frac{\partial^2 n_p}{\partial x^2} + G_n - \frac{n_p - n_{p0}}{\tau_n}$$

and for holes in an n-type semiconductor:

$$\frac{\partial p_n}{\partial t} = p_n \mu_p \frac{\partial \mathcal{E}}{\partial x} - \mu_p \mathcal{E} \frac{\partial p_n}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2} + G_p - \frac{p_n - p_{n0}}{\tau_p}$$

These equations describe the temporal and spatial evolution of carrier concentrations, including the effects of drift, diffusion, generation, and recombination.

The continuity equations must be satisfied together with the **Poisson equation**,

$$\frac{d^2\psi}{dx^2} = -\frac{\rho_s}{\varepsilon_s}$$

where ε_s is the permittivity of the semiconductor and ρ_s is the **spatial charge density**, given by the sum of carrier densities and concentrations of ionized impurities:

$$\rho_s = q(p - n + N_D^+ - N_A^-)$$

The system formed by the continuity equations and the Poisson equation, with the appropriate boundary conditions, admits a unique solution.

1.4 THE p-n JUNCTION

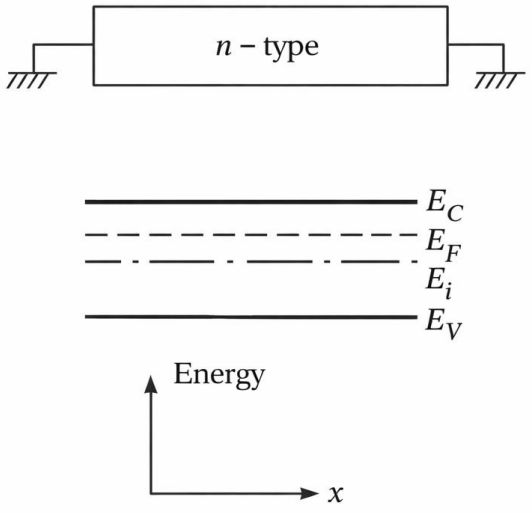
1.4.1 Thermal equilibrium condition

Consider a doped semiconductor, for example of n-type, and an electric field $\vec{\mathcal{E}}$ applied. Each conduction electron will be subjected to a force equal to $-q\,\vec{\mathcal{E}}$, that can be intended as the **negative gradient of potential energy**. In one-dimensional form:

$$-q\mathcal{E} = -\frac{dE_C}{dx}$$

where E_C is the lower edge of the conduction band, which represents the potential energy of an electron.

Since we are interested in the gradient of the potential energy, any energy level parallel to E_C can be used, such as the Fermi level E_F, the intrinsic level E_i, or the edge of the valence band E_V



In the case of a doped semiconductor, it is particularly convenient to refer to the **intrinsic Fermi level E_i**, obtaining

$$q\mathcal{E} = \frac{dE_C}{dx} = \frac{dE_i}{dx}$$

1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

$$q\mathcal{E} = \frac{dE_C}{dx} = \frac{dE_i}{dx}$$

Now let's introduce the **electrostatic potential** ψ, defined so that its negative gradient is the applied electric field:

$$\mathcal{E} \equiv -\frac{d\psi}{dx}$$

Let's compare the two expressions:

$$\psi = -\frac{E_i}{q}$$

which directly connects the **electrostatic potential** with the **potential energy** of the electrons in the semiconductor.

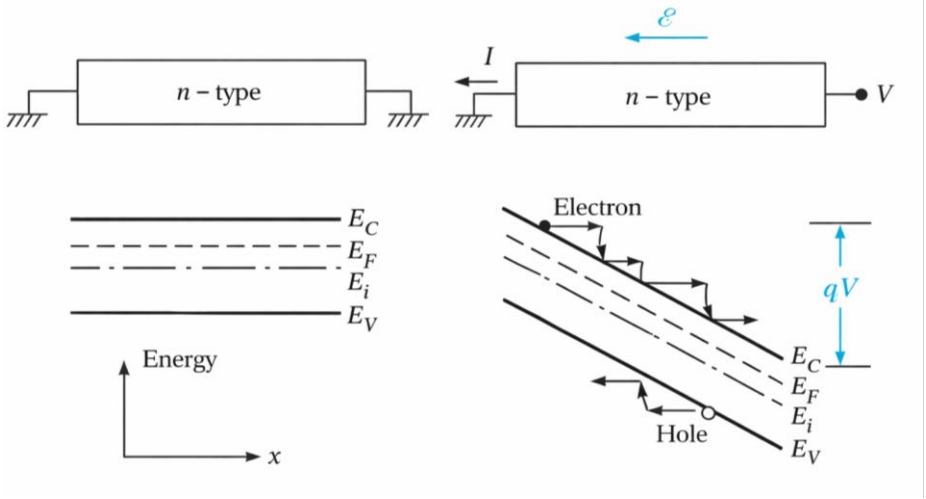
It follows that, if a constant electric field is applied between the two edges of a semiconductor sample, the potential ψ scales linearly with the spatial coordinate and, consequently, the edges of the energy bands also slope linearly.

1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

The inclination of the bands determines the **drift motion of the carriers**, which gives rise to the **electric current**. The electrons in the conduction band, under the action of $\vec{\mathcal{E}}$ in Figure, moves from left to the right.

The kinetic energy of electrons is measured with respect to the lower edge of the conduction band, i.e. with respect to E_C.



Between one collision and the next, the electron is accelerated by the field and increases its kinetic energy. When a collision with the lattice occurs, the electron gives up part or all of this energy to the lattice, returning to a condition close to thermal equilibrium, i.e. near the bottom of the conduction band. In the E–x diagram, this process corresponds to a **vertical transition to E_C**.

Once the collision is over, the process begins again: the electron is accelerated again, acquires energy, collides and loses it, in a continuous succession of cycles that produces a net drift shift.

A similar situation applies to holes in the valence band, where their drift motion occurs in the opposite direction to the electrons.

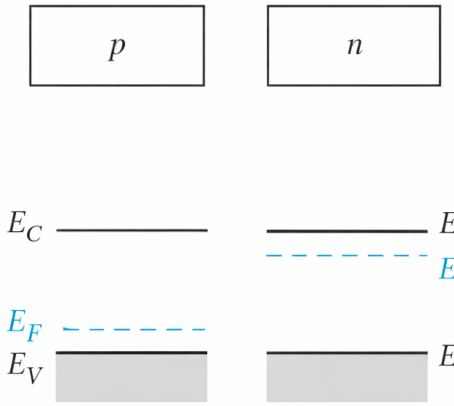
1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

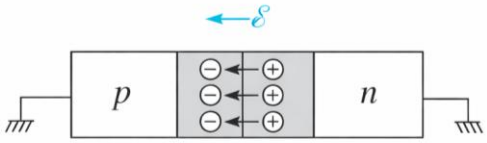
Now consider two semiconductors, one of p-type and the other n-type, both uniformly doped and initially separated.

In this condition the Fermi level E_F is located near the edge of the valence band in the p-type semiconductor and near the edge of the conduction band in the n-type semiconductor.

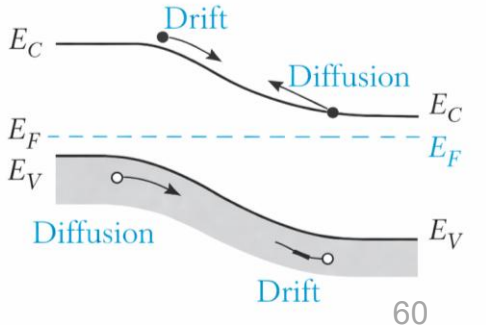
The p-type material is therefore characterized by a high concentration of holes and few electrons, while in the n-type material the opposite occurs.



When the two materials are brought into contact to form the **junction**, the strong concentration gradients of the carriers near the interface induce a **diffusion process**: the holes diffuse from the p-side to the n-side, while the electrons diffuse from the n-side to the p-side.

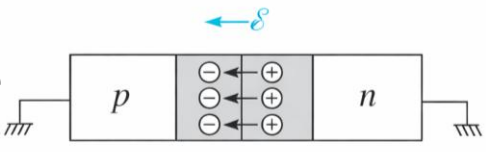


As the holes leave the p-side, the negative acceptor ions N_A^-, fixed in the lattice, remain uncompensated. Similarly, when the electrons leave the n-side, the positive donor ions N_D^+ remain unbalanced.

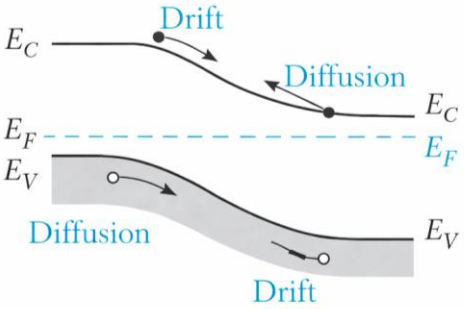


1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition



A **region of spatial charge** is thus formed: negative on the p-side and positive on the n-side. This charge distribution generates an **internal electric field** directed from the positive region to the negative one, generating also a **drift current** of holes and electrons.



At thermal equilibrium, i.e. in the absence of external excitations and at a constant temperature, the currents of electrons and holes through the junction compensate each other exactly, resulting in zero overall.

Therefore, for each type of carrier, the drift current induced by the electric field must exactly compensate for the diffusion current due to the concentration gradient. For holes we therefore have:

$$J_p = J_p^{(\text{drift})} + J_p^{(\text{diffusion})} = q\mu_p p\mathcal{E} - qD_p\frac{dp}{dx} = 0$$

Using the relations previously found:

$$J_p = q\mu_p p\left(\frac{1}{q}\frac{dE_i}{dx}\right) - kT\mu_p\frac{dp}{dx} = 0$$

$$J_n = q\mu_n n\mathcal{E} + qD_n\frac{dn}{dx}$$

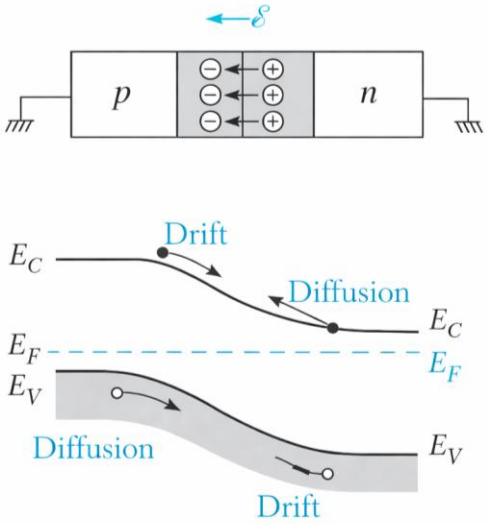
$$\mathcal{E} = \frac{1}{q}\frac{dE_i}{dx}$$

$$D_p = \left(\frac{kT}{q}\right)\mu_p$$

1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

$$J_p = q\mu_p p\left(\frac{1}{q}\frac{dE_i}{dx}\right) - kT\mu_p\frac{dp}{dx} = 0$$



We can substitute the expression for the concentration of holes and its derivative:

$$\frac{dp}{dx} = n_i e^{\frac{(E_i - E_F)}{kT}}\frac{1}{kT}\left(\frac{dE_i}{dx} - \frac{dE_F}{dx}\right) = \frac{p}{kT}\left(\frac{dE_i}{dx} - \frac{dE_F}{dx}\right)$$

to obtain:

$$J_p = \mu_p p\left(\frac{dE_i}{dx}\right) - \mu_p p\left(\frac{dE_i}{dx} - \frac{dE_F}{dx}\right) = \mu_p p\frac{dE_F}{dx} = 0$$

$$p = n_i e^{\frac{(E_i - E_F)}{kT}}$$

With a similar procedure for electrons, we find:

$$J_n = \mu_n n\frac{dE_F}{dx} = 0$$

From both relations it immediately follows that:

$$\frac{dE_F}{dx} = 0$$

1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

$$\frac{dE_F}{dx} = 0$$

Therefore, for the net current of electrons and holes to be zero, **the Fermi level must be constant, i.e. independent of the x-spatial coordinate**, throughout the sample.

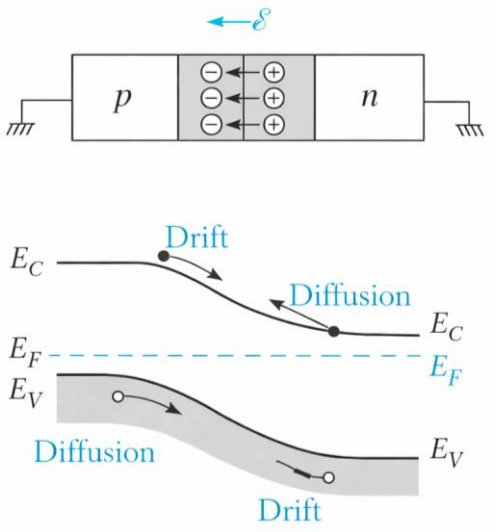
The spatial charge distribution and the corresponding electrostatic potential ψ are determined by the Poisson equation:

$$\frac{d^2\psi}{dx^2} \equiv -\frac{d\mathcal{E}}{dx} = -\frac{\rho_s}{\varepsilon_s} = -\frac{q}{\varepsilon_s}\left(p - n + N_D^+ - N_A^-\right)$$

We assume, as usual, that all donors and acceptors are completely ionized.

In regions far from the junction, charge neutrality is preserved, so the total spatial charge density is zero. In such neutral regions, the Poisson equation reduces to:

$$\frac{d^2\psi}{dx^2} = 0$$



$$\rho_s = q(p - n + N_D^+ - N_A^-)$$

$$\frac{d^2\psi}{dx^2} = -\frac{\rho_s}{\varepsilon_s}$$

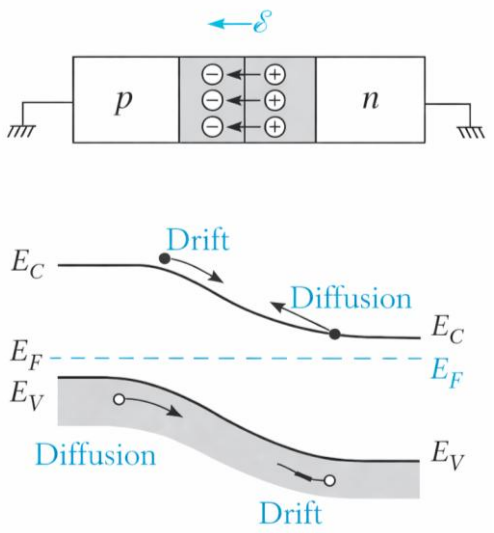
$$\psi = -\frac{E_i}{q}$$

$$\mathcal{E} = \frac{1}{q}\frac{dE_i}{dx}$$

1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

$$\frac{d^2\psi}{dx^2} = -\frac{q}{\varepsilon_s}\left(p - n + N_D^+ - N_A^-\right) = 0$$



from which follows the condition of local neutrality

$$p - n + N_D^+ - N_A^- = 0$$

Now consider the neutral region of **p-type**. With the condition $N_D^+ = 0$ and, since $p \gg n$, neglect the concentration of electrons.

The electrostatic potential ψ_p with respect to the Fermi level is obtained by imposing $p = N_A^-$ in the expression of the concentration of holes

$$N_A^- = n_i e^{\frac{(E_i - E_F)}{kT}}$$

$$p = n_i e^{\frac{(E_i - E_F)}{kT}}$$

$$\psi = -\frac{E_i}{q}$$

from which:

$$(E_i - E_F) = kT ln\left(\frac{N_A^-}{n_i}\right)$$

The potential in the p-type neutral region with respect to the Fermi level will be:

$$\psi_p \equiv -\frac{1}{q}(E_i - E_F)\Big|_{x \le -x_p} = -\frac{kT}{q} ln\left(\frac{N_A^-}{n_i}\right)$$

1.4 THE p-n JUNCTION

1.4.1 Thermal equilibrium condition

$$\psi_p \equiv -\frac{1}{q}(E_i - E_F)\Big|_{x \le -x_p} = -\frac{kT}{q} ln\left(\frac{N_A^-}{n_i}\right)$$

Similarly, for the neutral region of **n-type**, imposing $n = N_D^+$:

$$n = n_i e^{\frac{E_F - E_i}{KT}}$$

$$N_D^+ = n_i e^{\frac{(E_F - E_i)}{kT}}$$

$$\psi = -\frac{E_i}{q}$$

the potential ψ_n in the n-type neutral region with respect to the Fermi level is obtained:

$$\psi_n \equiv -\frac{1}{q}(E_i - E_F)\Big|_{x \ge x_n} = \frac{kT}{q} ln\left(\frac{N_D^+}{n_i}\right)$$

The difference in total electrostatic potential between the two neutral regions in thermal equilibrium is called the **built-in potential** V_{bi} :

$$V_{bi} = \psi_n - \psi_p = \frac{kT}{q} ln\left(\frac{N_A^- N_D^+}{n_i^2}\right)$$

1.4 THE p-n JUNCTION

1.4.2 The depletion zone

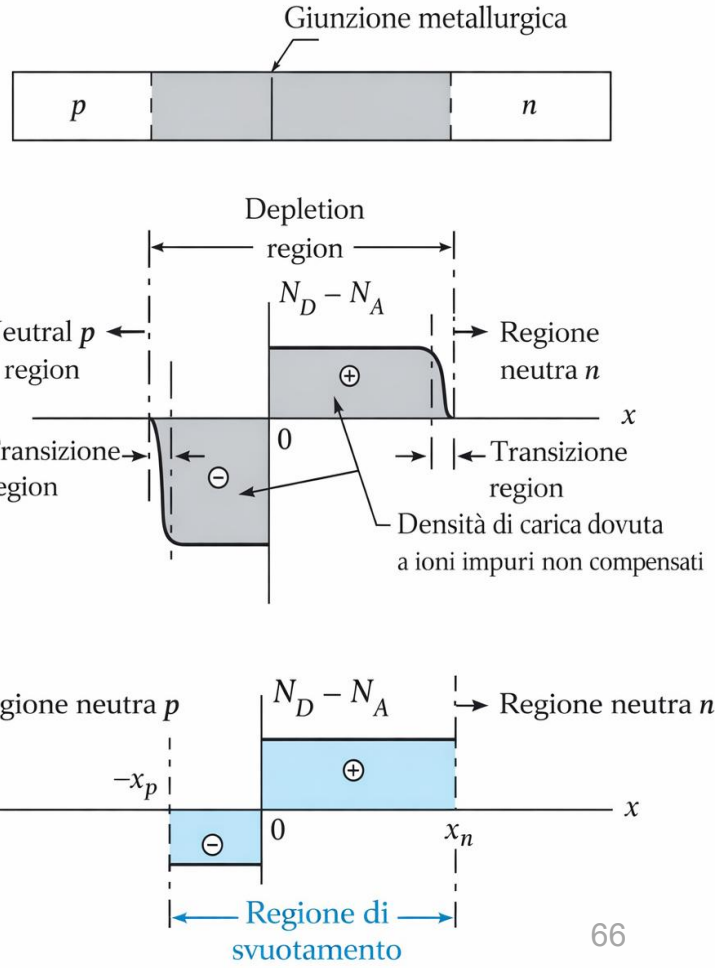
A representation of the distribution of charges at a p-n junction is shown in Figure.

Approaching from the neutral region towards the junction, we first find a **thin transition region**, in which the spatial charge of the impurity ions is only partially compensated by the mobile carriers.

Beyond this zone we enter the emptying region, where the densities of the carriers are negligible and we can set $p = n = 0$.

The width of the transition regions is much smaller than that of the depletion zone.

It is therefore a good approximation to neglect these zones and represent the **depletion region by a rectangular charge distribution**, where the extensions of the depletion zone on the p and n sides are x_p and x_n, respectively.

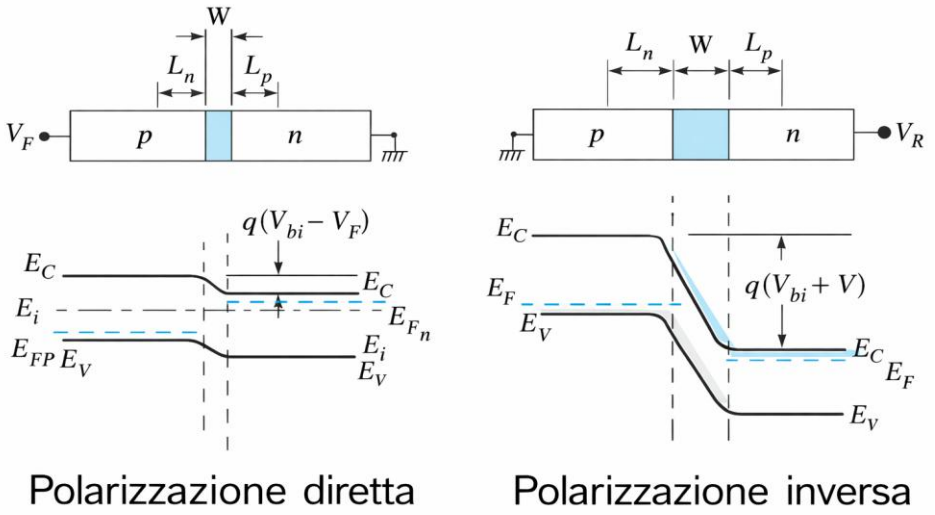


1.4 THE p-n JUNCTION

1.4.3 The voltage-current characteristic

Applying a voltage to a p–n junction **alters the balance** between the diffusion current and the drift current of the electrons and holes.

Forward polarization is when the applied voltage reduces the electrostatic potential across the depletion region.



Polarizzazione diretta Polarizzazione inversa

In this condition, a greater number of electrons present in the high-energy tail of the conduction band of the n-side have sufficient energy to overcome the reduced barrier and diffuse from the n-side to the p-side.

Similarly, holes in the p-side valence band can overcome the barrier and spread to the n-side. Thus, the injection of minority carriers occurs: electrons injected into the p-side and holes injected into the n-side.

Reverse polarization is when the applied voltage increases the electrostatic potential across the depletion region. In this case, the energy barrier increases and the diffusion currents are greatly reduced.

1.4 THE p-n JUNCTION

1.4.3 The voltage-current characteristic

To derive the **ideal voltage-current characteristic** of the junction, we adopt the following fundamental assumptions:

(a) the depletion region has **sharp boundaries** and, outside of them, the semiconductor can be considered electrically neutral

(b) the carrier densities at the edges of the depletion region are related by the **electrostatic potential difference** applied to the junction;

(c) low injection condition, i.e. the densities of the minority carriers injected are much smaller than the densities of the majority carriers, i.e. the concentrations of the majority carriers in the neutral regions remain unchanged as the voltage varies.

(d) in the depletion region there are **no generation or recombination currents**, and the currents are constant throughout the region.

At thermal equilibrium, the density of the majority carriers in the neutral regions essentially coincides with the doping concentration.

We will indicate with the subscripts n and p the type of semiconductor and with the subscript 0 the condition of thermal equilibrium.

Thus, n_{n0} and n_{p0} represent the electron densities in equilibrium on the n and p sides of the junction, respectively.

1.4 THE p-n JUNCTION

1.4.3 The voltage-current characteristic

With this notation and, we assume $N_A^- = p_{p0}$ and e $N_D^+ = n_{n0}$; The built-in potential then becomes:

$$V_{bi} = \frac{kT}{q} ln\left(\frac{p_{po} n_{no}}{n_i^2}\right)$$

$$V_{bi} = \psi_n - \psi_p = \frac{kT}{q} ln\left(\frac{N_A^- N_D^+}{n_i^2}\right)$$

$$p_{po} n_{po} = n_i^2$$

Using the mass action law for the p-side of the Junction

$$V_{bi} = \frac{kT}{q} ln\left(\frac{n_{no}}{n_{po}}\right)$$

leading to:

$$n_{no} = n_{po} e^{\frac{qV_{bi}}{kT}}$$

In a very similar way on the n-side of the junction we also obtain:

$$p_{po} = p_{no} e^{\frac{qV_{bi}}{kT}}$$

It is worth noting that the electron and hole densities at the two boundaries of the depletion region are linked to each other by the potential difference V_{bi} under conditions of thermal equilibrium.

1.4 THE p-n JUNCTION

1.4.3 The voltage-current characteristic

$$n_{no} = n_{po} e^{\frac{qV_{bi}}{kT}} \qquad p_{po} = p_{no} e^{\frac{qV_{bi}}{kT}}$$

Based on the second assumption, we expect the same relationship to remain valid even when the electrostatic potential difference is changed by an applied voltage.

(b) the carrier densities at the boundaries are related by the electrostatic potential difference across the junction

With a **forward polarization**, the potential barrier is reduced to $V_{bi} - V_F$; with an **inverse polarization**, it increases to $V_{bi} + V_R$.

As a result, the above equation is modified:

$$n_n = n_p e^{\frac{q(V_{bi}-V)}{kT}}$$

where n_n and n_p are the out-of-equilibrium electron densities at the boundaries of the depletion region at the n- and -side, respectively, and V is positive in forward polarization and negative in reverse polarization.

In the low-injection regime, the density of injected minority carriers is much smaller than the density of majority carriers; therefore, we can assume, $n_n \approx n_{no}$

1.4 LA GIUNZIONE p-n

1.4.3 La caratteristica tensione-corrente

$$n_{no} = n_{po} e^{\frac{qV_{bi}}{kT}}$$

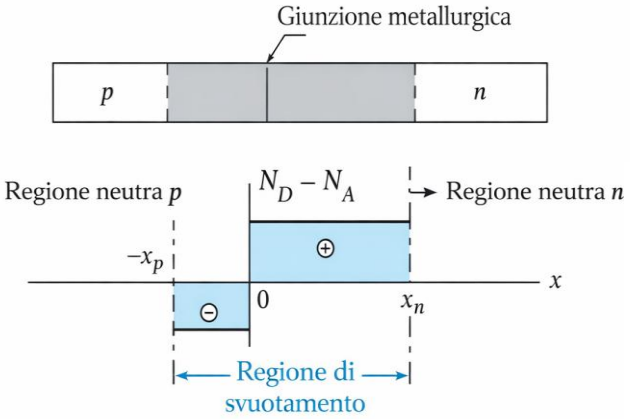
$$n_n \approx n_{no}$$

$$n_n = n_p e^{\frac{q(V_{bi}-V)}{kT}}$$

$$\frac{\partial p_n}{\partial t} = p_n \mu_p \frac{\partial \mathcal{E}}{\partial x} - \mu_p \mathcal{E} \frac{\partial p_n}{\partial x} + D_p \frac{\partial^2 p_n}{\partial x^2} + G_p - \frac{p_n - p_{n0}}{\tau_p}$$

Combining these expressions, we obtain for the electron density at the boundary of the depletion region on the p-type side ($x = -x_p$):

$$n_p = n_{po} e^{\frac{qV}{kT}}$$



Similarly, for the holes at the boundary of the n-type region ($x = x_n$) we have:

$$p_n = p_{no} e^{\frac{qV}{kT}}$$

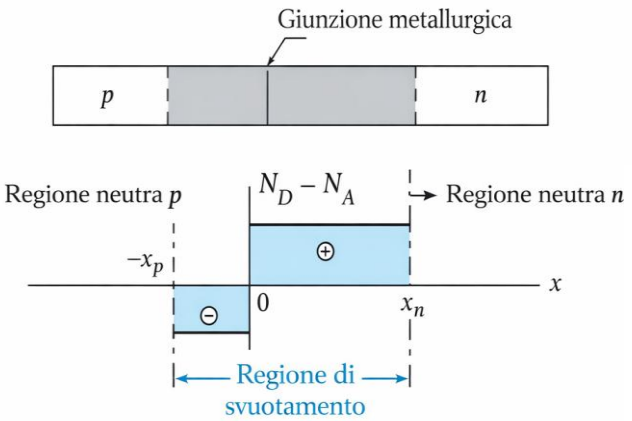
Under the assumptions, there is no current generation in the depletion region: all current components originate in the neutral regions. In the n-type neutral region, $\mathcal{E} = 0$; therefore, the continuity equation in steady-state conditions is simplified as:

$$\frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{D_p \tau_p} = 0$$

1.4 LA GIUNZIONE p-n

1.4.3 La caratteristica tensione-corrente

$$\frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{D_p \tau_p} = 0$$



The general solution has the form:

$$p_n - p_{n0} = A\, e^{x/L_p} + B\, e^{-x/L_p}$$

with $L_p = \sqrt{D_p \tau_p}$ is **the diffusion length of the holes** in the n-side.

Now let's apply the two boundary conditions discussed above.

The first condition requires that, sufficiently far from the junction, the thermal equilibrium of the carriers is restored. This is equivalent to imposing that $p_n(x \to \infty) = p_{no}$. Applying this condition to the general solution of the differential equation, we immediately obtain that the coefficient of the increasing exponential must cancel out, i.e. $A = 0$.

The second boundary condition is derived from the equation at the edges of the depletion region, evaluated at the point $x = x_n$. From it we obtain for the coefficient B :

$$B = \frac{p_{n0}\left(e^{\frac{qV}{kT}} - 1\right)}{e^{-x_n/L_p}}$$

$$p_n = p_{no} e^{\frac{qV}{kT}}$$

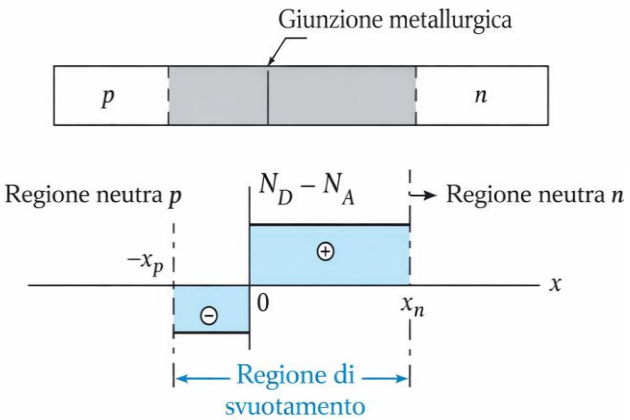
1.4 LA GIUNZIONE p-n

1.4.3 La caratteristica tensione-corrente

$$p_n - p_{n0} = A\,e^{x/L_p} + B\,e^{-x/L_p}$$

$$A = 0$$

$$B = \frac{p_{n0}\left(e^{\frac{qV}{kT}} - 1\right)}{e^{-x_n/L_p}}$$



We can have now the particular solution for the distribution of holes on the n-side:

$$p_n - p_{n0} = p_{n0}\left(e^{\frac{qV}{kT}} - 1\right)e^{-\frac{(x-x_n)}{L_p}}$$

We can now determine the diffusion current density of the holes in the n-type side by substituting the solution obtained in the general expression of the diffusion current and evaluating it at the point $x = x_n$, where the electric field is zero ($\mathcal{E} = 0$) :

$$J_p(x_n) = -qD_p \left.\frac{dp_n}{dx}\right|_{x_n} = \frac{qD_p p_{n0}}{L_p}\left(e^{\frac{qV}{kT}} - 1\right)$$

$$J_p = q\mu_p p_n \mathcal{E} - qD_p \frac{dp_n}{dx}$$

In a very similar way, in the p-type neutral region the distribution of minority carriers (electrons) is

$$n_p - n_{po} = n_{po}\left(e^{\frac{qV}{kT}} - 1\right)e^{\frac{(x+x_p)}{L_n}}$$

1.4 LA GIUNZIONE p-n

1.4.3 La caratteristica tensione-corrente

$$n_p - n_{po} = n_{po}\left(e^{\frac{qV}{kT}} - 1\right)e^{\frac{(x+x_p)}{L_n}}$$

$$J_n = q\mu_n n\mathcal{E} + qD_n\frac{dn}{dx}$$

from which the electron current density at the boundary $x = -x_p$ where the electric field is zero ($\mathcal{E} = 0$) is:

$$J_n(-x_p) = qD_n\left.\frac{dn_p}{dx}\right|_{x_n} = \frac{qD_n n_{po}}{L_n}\left(e^{\frac{qV}{kT}} - 1\right)$$

where $L_n = \sqrt{D_n\tau_n}$ is called the **diffusion length of the electrons** (minority carriers) in the p region.

Since there are no carrier generation in the depletion region, the total current is constant throughout the device and is the sum of the two contributions

$$J = J_p(x_n) + J_n(-x_p) = J_s\left(e^{\frac{qV}{kT}} - 1\right)$$

$$J_p(x_n) = \frac{qD_p p_{n0}}{L_p}\left(e^{\frac{qV}{kT}} - 1\right)$$

with $J_s = \frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n}$ called saturation current

1.4 LA GIUNZIONE p-n

1.4.3 La caratteristica tensione-corrente

$$J = J_s\left(e^{\frac{qV}{kT}} - 1\right)$$

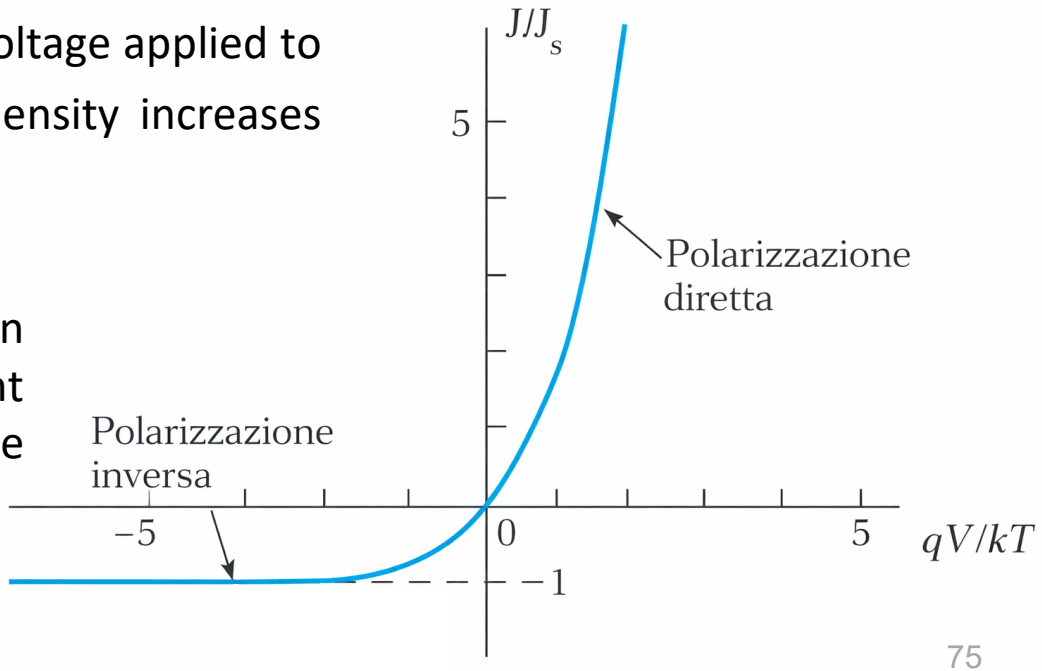
$$J_s = \frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{po}}{L_n}$$

The expression of the current density J is known as the **ideal diode equation**.

The resulting current-voltage characteristic is shown in Figure.

In **forward polarization**, with positive voltage applied to the p-side, for $V \gtrsim \frac{3kT}{q}$, the current density increases exponentially with the applied voltage.

In **reverse polarization**, on the other hand, the current density is saturated at the constant value $J \simeq -J_s$.



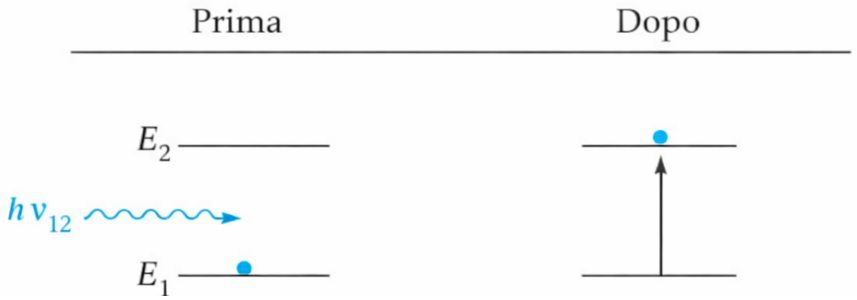
1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

To fully understand the three main processes of interaction between light and matter in a semiconductor, it is useful to start with simpler considerations on the **energy levels** rather than energy band.
There are three fundamental processes of interaction between a photon and an electron in a semiconductor: **absorption**, **spontaneous emission** and **stimulated emission**.

To introduce them clearly, let's consider a two-level energy system, consisting of two states E_1 and E_2 of an atom, where E_1 represents the ground state and E_2 the excited state. Any transition between these two levels involves the emission or absorption of a photon of frequency ν_{12} such that $h\nu_{12} = E_2 - E_1$.

At room temperature, most atoms are in the ground state E_1.

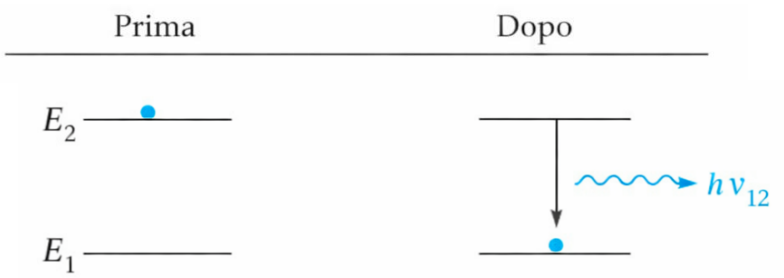


This situation changes when a photon of energy exactly equal to $h\nu_{12}$ interacts with this system: an atom in the E_1 state can absorb the photon and move to the excited E_1 state. This process is called **absorption**.

1.5 LIGHT-EMITTING DIODES

1.5.1 Transizioni radiative e assorbimento ottico

The excited state of the atom is unstable. After a short time interval, in the absence of external stimuli, the atom spontaneously returns to the ground state, emitting a photon of energy $h\upsilon_{12}$ This process is called **spontaneous emission**.



If, on the other hand, a photon of energy $h\upsilon_{12}$ hits an atom that is already in the excited state E_2, the atom can be induced to return to the state E_1 by emitting a second photon, with the same energy, direction and phase as the incident photon.

This process is called **stimulated emission**. Since the emitted photons are in phase, the resulting radiation is coherent.

From an application point of view, **spontaneous emission is the dominant mechanism in LEDs**, **stimulated emission is the basis of the operation of laser diodes**, while **absorption is the fundamental process in photoconductive and photovoltaic detector**.

1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

We indicate with n_1 and n_2 the instantaneous populations of the E_1 and E_2 level. Under conditions of thermal equilibrium and for $E_2 - E_1 > 3kT$, they ratio is ruled by the Boltzmann distribution:

$$\frac{n_2}{n_1} = e^{-\frac{(E_2-E_1)}{kT}} = e^{-\frac{h\nu_{12}}{kT}}$$

The negative sign of the exponent indicates that, in equilibrium, the population of the upper level is much smaller than that of the lower level.

In a steady state, in order to keep the n_1 and n_2 constant, the overall emission rate (stimulated and spontaneous) must balance the absorption rate:

$$\text{stimulated emission rate} + \text{spontaneous emission rate} = \text{absorption rate}$$

The **stimulated emission rate** is proportional to the energy density of the photon field $\rho(h\nu_{12})$ and to the population of the upper level n_2, and is therefore equal to $B_{21}n_2\rho(h\nu_{12})$, with B_{21} is the Einstein coefficient for stimulated emission.

The **spontaneous emission rate** depends only on the population of the upper level, and is equal to $A_{21}n_2$, with A_{21} is the Einstein coefficient for spontaneous emission.

The **absorption rate** is proportional to $n_1\rho(h\nu_{12})$ and is therefore equal $B_{12}n_1\rho(h\nu_{12})$, with B_{12} Einstein coefficient for absorption.

1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

stimulated emission rate + spontaneous emission rate = absorption rate

Therefore, in steady state, it results:

$$B_{21}n_2\rho(h\upsilon_{12}) + A_{21}n_2 = B_{12}n_1\rho(h\upsilon_{12})$$

The ratio between stimulated emission and spontaneous emission is:

$$\frac{\text{stimulated emission rate}}{\text{spontaneous emission rate}} = \frac{B_{21}}{A_{21}}\rho(h\upsilon_{12})$$

To make the stimulated emission dominant, it is therefore necessary to **increase the energy density of the photon** field $\rho(h\upsilon_{12})$, a condition obtained by using a resonant optical cavity to trap the photons in a finite space.

Finally, comparing stimulated emission and absorption:

$$\frac{\text{stimulated emission rate}}{\text{absorption rate}} = \frac{B_{21}}{B_{12}}\frac{n_2}{n_1}$$

In order for stimulated emission to prevail over absorption, the population of the upper level must exceed that of the lower level. This condition, opposite to that of equilibrium one, is called **population inversion**.

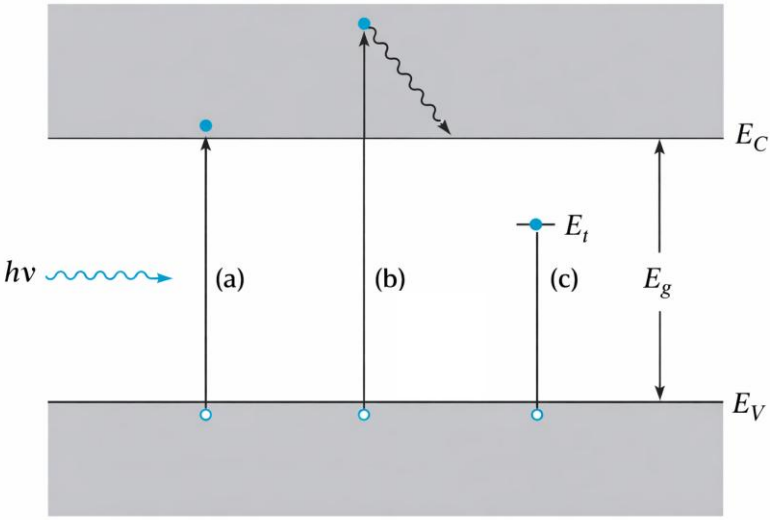
1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

After introducing the fundamental processes of absorption, spontaneous emission and stimulated emission starting from a two-level system, we can now apply these concepts to the specific case of semiconductors with energy bands.

In this context, the discrete energy levels of atoms are replaced by energy bands and bandgap, and optical transitions must consider the band structure of the material.

When the material is illuminated, photons can be absorbed generating electron-hole pairs. In the case in Figure(a), if the energy of the photon satisfies the condition $h\nu = E_g$, the absorption generates an electron-hole pair: an electron is promoted from the upper edge of the valence band to the lower edge of the conduction band.

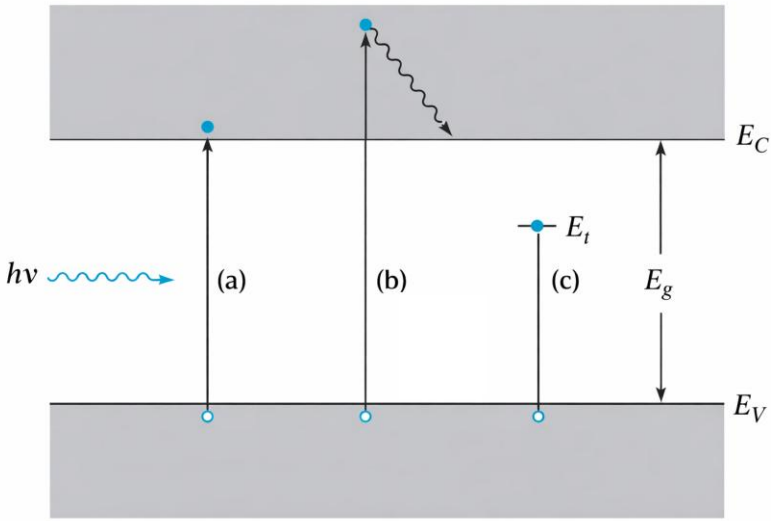


In the case of Figure(b), if the energy of the photon is greater than the gap, $h\nu > E_g$, an electron-hole pair is still generated, but the excess energy is quickly dissipated in the form of heat through interactions with the lattice.

1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

Both processes (a) and (b) are called **intrinsic transitions** or **band-to-band transitions**, because they directly involve the two fundamental bands of the semiconductor.



In the case of Figure(c), if $hv < E_g$, absorption is possible only in the presence of energy levels within the bandgap, introduced by chemical impurities or structural defects. In this case, the photon promotes an electron between one of these intermediate levels and one of the bands: this is called an **extrinsic transition**.

This discussion is also valid for the reverse situation.

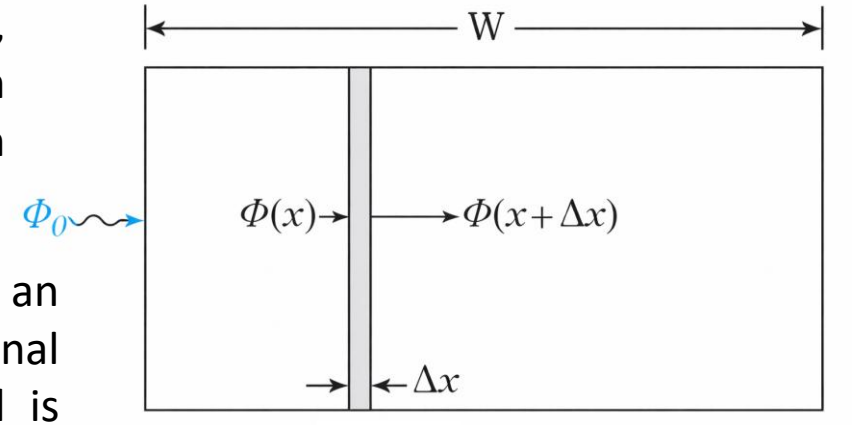
A similar classification applies to inverse emission processes: for example, the recombination of an electron at the edge of the conduction band with a hole at the edge of the valence band gives rise to the emission of a photon with energy equal to the gap E_g.

1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

Let us now consider quantitatively what happens when many photons induce band-to-band transitions in a semiconductor. Suppose that a semiconductor is illuminated by a light source with photon energy $h\nu > E_g$ and with an incident flux equal to Φ_0 photons per square centimeter per second.

As the flow of photons penetrates the material, a fraction of the photons are absorbed through transitions from valence states to conduction states.



The number of photons absorbed in an infinitesimal thickness dx is proportional to the local intensity of the flux and is given by $\alpha\ \Phi(x)\ dx$, with α the absorption coefficient.

It follows that the variation of the flux along the direction of propagation satisfies the relation:

$$\Phi(x+dx) - \Phi(x) = -\alpha\Phi(x)dx$$

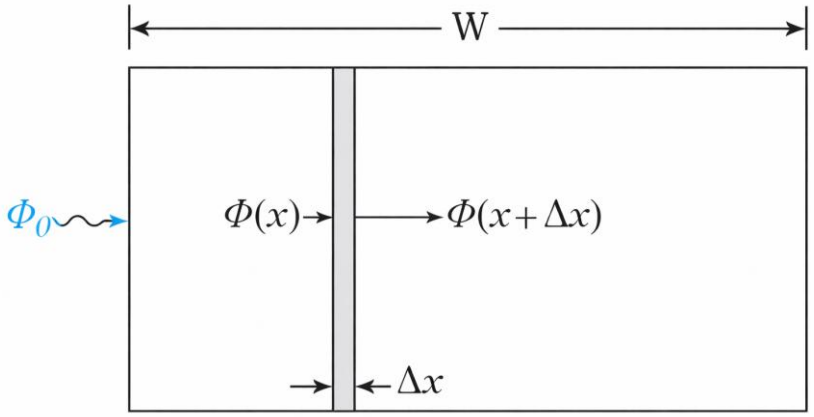
1.5 LIGHT-EMITTING DIODES

1.5.1 Radiative transitions and optical absorption

$$\Phi(x+dx) - \Phi(x) = -\alpha\Phi(x)dx$$

or, in the differential limit

$$\frac{d\Phi(x)}{dx} = -\alpha\Phi(x)$$



where the minus sign indicates the decrease in intensity due to absorption.

Imposing the boundary condition $\Phi(x) = \Phi_0$ for $x = 0$, the solution is:

$$\Phi(x) = \Phi_0 e^{-\alpha x}$$

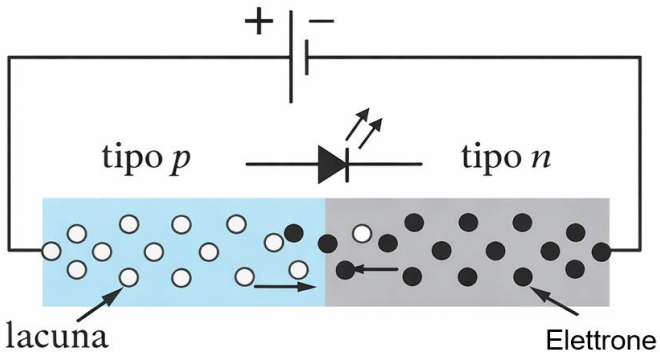
which shows how the flow of photons decreases exponentially within the semiconductor. The absorption coefficient α depends on the energy of the photon hv and is therefore directly related to the energy structure of the material.

1.5 LIGHT-EMITTING DIODES

1.5.2 Structure of an LED

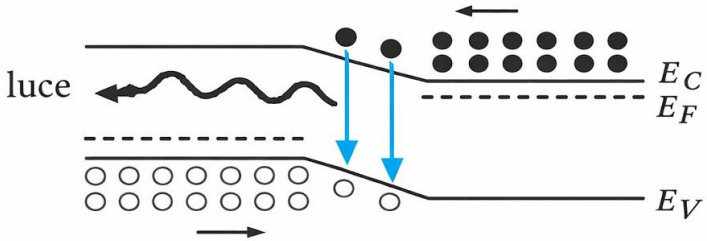
The fundamental structure of an LED is a **p–n junction**.

Under direct polarization conditions, electrons are injected from the n-type side and holes from the p-type side, as shown in Figure.



The application of V voltage reduces the built-in potential, allowing injected carriers to cross the depletion region and become excess minority carriers in the opposite regions.

Near the junction, the concentration of carriers exceeds the equilibrium value $(pn > n_i^2)$ and the process of radiative recombination, i.e. the spontaneous emission of photons, is triggered.

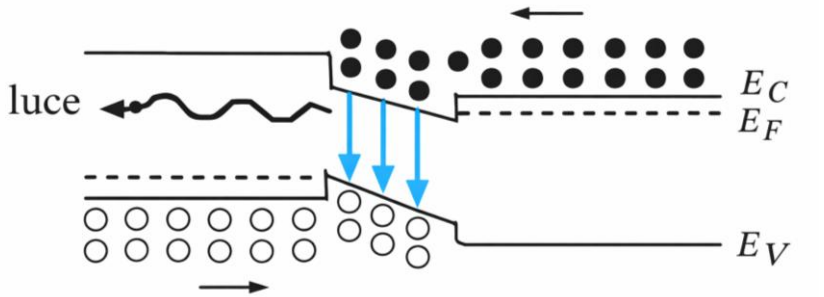


The efficiency of an LED can be significantly improved by using a **heterojunction**, i.e. a junction between different semiconductor materials.

1.5 LIGHT-EMITTING DIODES

1.5.2 Structure of an LED

By injecting carriers from a wide-bandgap semiconductor to an **active region characterized by a narrower bandgap semiconductor**, energy band discontinuities (band offsets) act as confinement barriers.



In this way, electrons and holes remain trapped in the active region, increasing the probability of radiative recombination.

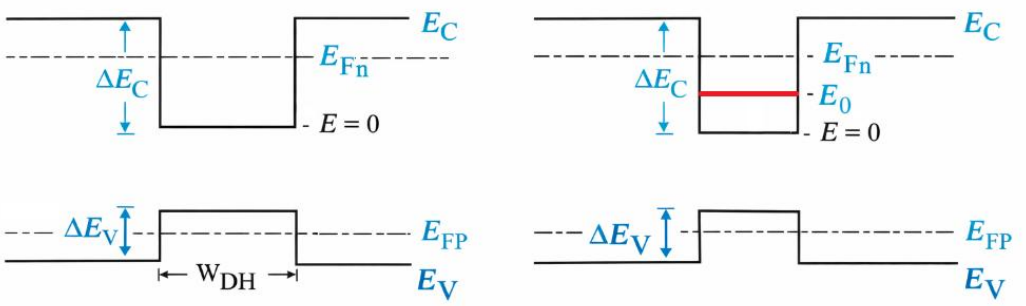
This configuration, known as **double heterostructure**, allows to obtain high concentrations of carriers in the active region, as schematically shown in Figure. In addition, the emitted photons are not reabsorbed in the wider bandgap confinement layers, since their energy is lower than the bandgap value of the barriers.

The thickness of the central active layer can be further reduced to a **quantum well**.

A quantum well is a double heterostructure in which the narrow bandgap layer is very small in thickness, typically less than 50 nm.

1.5 LIGHT-EMITTING DIODES

1.5.2 Structure of an LED



When the thickness of the active region becomes comparable with the De Broglie wavelength of the carriers (of the order of 10 nm), evident quantum effects emerge, with the appearance of discrete intra-band energy levels (e.g. the E_0 level in the Figure).

Since the position of E_0 is related to the thickness of the quantum well, it allows **the emission wavelength to be engineered** more precisely than in the case of the bulk semiconductor.

From an efficiency perspective, reducing the thickness of the quantum well results in a significant increase in the density of carriers in the active region.

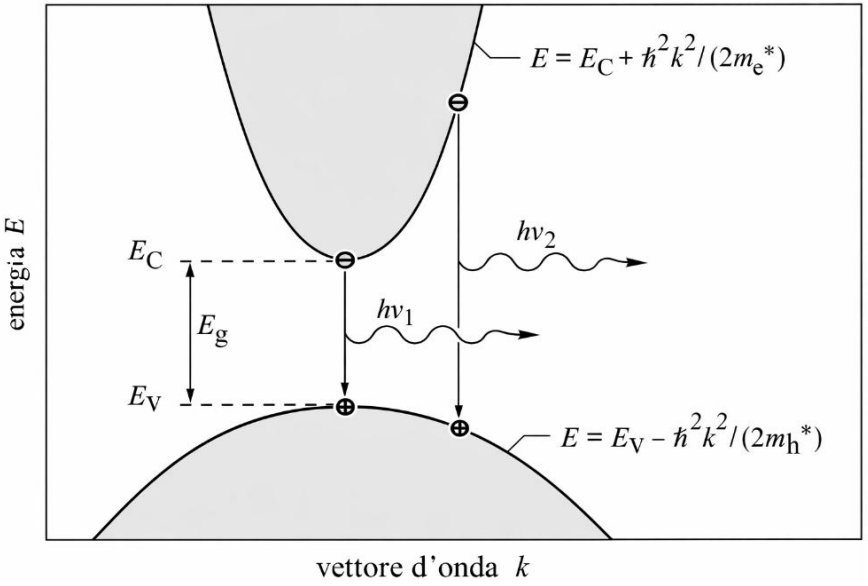
As a result, the average lifetime of carriers is reduced due to the increased probability of radiative recombination. Furthermore, thanks to the reduced thickness of the active region, the reabsorption of the emitted photons is strongly limited, further contributing to the improvement of the overall efficiency of the device.

1.5 LIGHT-EMITTING DIODES

1.5.3 Emission spectrum

The fundamental mechanism by which a semiconductor LED emits light is the **spontaneous radiative recombination** of electron-hole pairs, accompanied by the emission of a photon.

In order to quantitatively understand the spectral characteristics of the emitted radiation, it is necessary to describe the radiative transition process considering the dispersion relationships of the energy bands.



We assume that, near the band edges, the dispersion relations are parabolic, i.e.:

$$E = E_c + \frac{\hbar^2 k^2}{2m_e^*} \qquad \text{(for electrons)}$$

$$E = E_V - \frac{\hbar^2 k^2}{2m_h^*} \qquad \text{(for holes)}$$

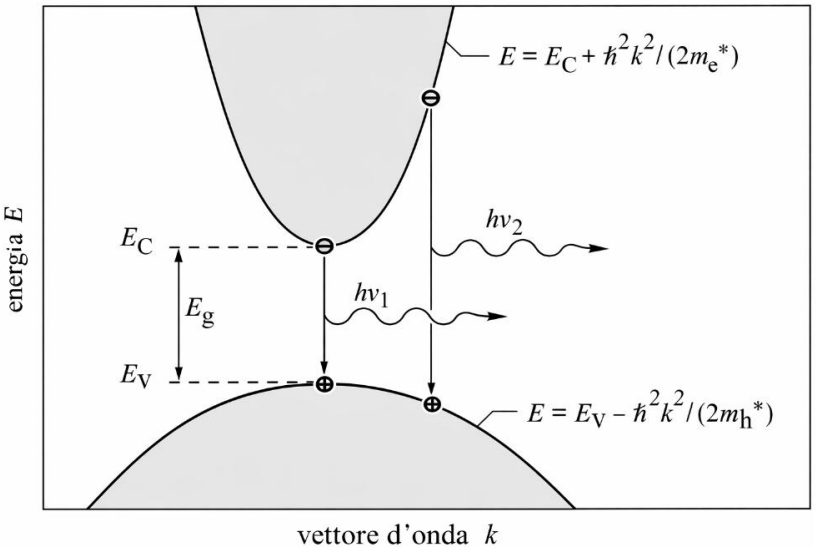
where m_e^* and m_h^* are the effective masses of electrons and holes, E_C is the lower edge of the conduction band and E_V the upper edge of the valence band.

1.5 LIGHT-EMITTING DIODES

1.5.3 Emission spectrum

The recombination process must satisfy the laws of conservation of energy and momentum.

From the Boltzmann distribution it follows that, at room temperature, electrons and holes have an average kinetic energy of the order of kT, smaller than the bandgap.



The conservation of energy imposes that the energy of the emitted photon is equal to the difference between the energy of the electron E_e and that of the hole E_h :

$$h\nu = E_e - E_h \approx E_g$$

that is, with a good approximation, the energy of the photon coincides with the energy of the band gap E_g. This results in a fundamental result for LED design: the emission wavelength is determined directly by the value of the band gap of the semiconductor material.

For example, GaAs has a bandwidth gap of $E_g = 1.42$ eV at room temperature; consequently, a GaAs LED emits in the infrared, with a wavelength of about $\lambda \simeq 870$ nm.

1.5 LIGHT-EMITTING DIODES

1.5.3 Emission spectrum

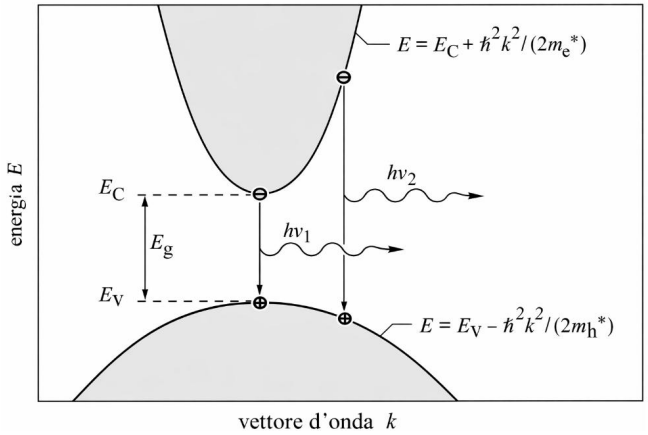
It is useful to compare the momentum of a photon with the extension of the first Brillouin zone in the one-dimensional case, for example for an atomic chain.

The momentum of a photon is derived from the de Broglie relation:

$$p = \hbar k = \frac{\hbar}{\lambda}$$

Since the wavelength λ is of the order of a few hundred nanometers, **the momentum of the photon is extremely small compared to the size of the Brillouin zone**, equal to 2π/a, with a typically of the order of a few Ångström.

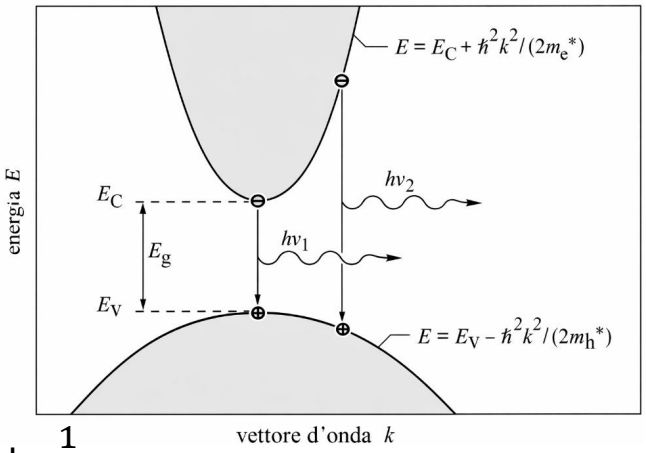
It follows that, during a radiative transition from the conduction band to the valence band, the momentum of the electron cannot vary appreciably. The transitions are therefore **vertical in the E-k diagram**, as shown in the Figure: the electrons recombine only with holes that have the same momentum.



1.5 LIGHT-EMITTING DIODES

1.5.3 Emission spectrum

By setting the conservation of momentum requirement, the energy of the emitted photon can be expressed via the **joint dispersion relation**.



$$h\upsilon = E_e - E_h = E_c + \frac{\hbar^2k^2}{2m_e^*} - E_V + \frac{\hbar^2k^2}{2m_h^*} = E_g + \frac{\hbar^2k^2}{2m_r^*}$$

where m_r^* is the reduced effective mass, defined by $\frac{1}{m_r^*} = \frac{1}{m_e^*} + \frac{1}{m_h^*}$.

The **joint density of states** is therefore

$$N(E) = \frac{1}{2\pi^2}\left(\frac{2m_r^*}{\hbar^2}\right)^{3/2}\sqrt{E - E_g}$$

Since the distribution of carriers in the permitted bands follows the Boltzmann distribution, the spectral intensity of emission $I(E)$ is proportional to the product of the joint density of state and the occupancy distribution:

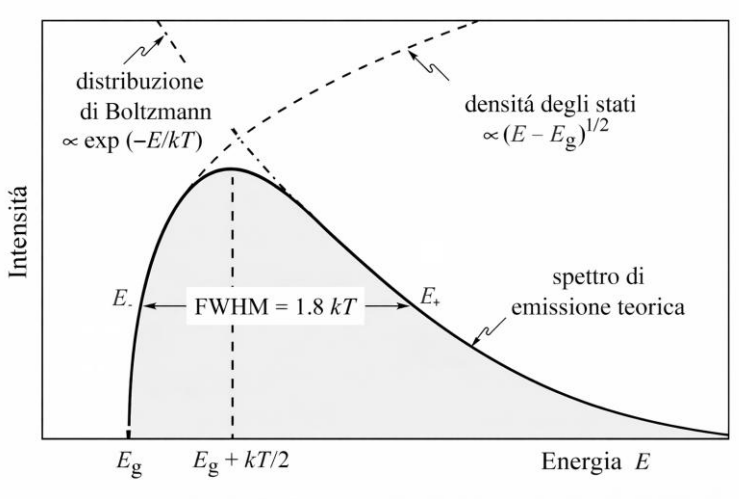
$$I(E) \propto N(E)\, f_B(E) \propto \sqrt{E - E_g}\, e^{-E/kT}$$

1.5 LIGHT-EMITTING DIODES

1.5.3 Emission spectrum

The spectral shape of the emission of an LED is therefore represented in the Figure as a product between the two opposite trends.

$$I(E) \propto \sqrt{E - E_g}\, e^{-\frac{E}{kT}}$$



We determine the two most relevant figures of merit for an emission spectrum: peak energy and full width at half maximum (FWHM) value

The peak energy can be determined directly by imposing zero the derivative of I(E)

$$\frac{\partial I(E)}{\partial E} = e^{-E/kT}\left[\frac{1}{2\sqrt{E - E_g}} - \frac{\sqrt{E - E_g}}{kT}\right] = 0$$

from which the solution for the peak energy value E_P

$$E_P = E_g + \frac{1}{2}kT$$

The FWHM value |ΔE| can be calculated by determining the two energy values $E_\pm$ for which the peak intensity is halved, i.e. $I(E_\pm) = \frac{1}{2}I(E_P)$.

1.5 LIGHT-EMITTING DIODES

1.5.3 Emission spectrum

This leads to the solution of a transcendent equation of which we report only the result:

$$|\Delta E| = |E_+ - E_-| = 1.8\,kT$$

Converting it in wavelength unit, using $|\Delta E| = \frac{hc}{\lambda^2}\Delta\lambda$, we get

$$\Delta\lambda = \frac{1.8kT\lambda^2}{hc}$$

$E = h\upsilon = h\frac{c}{\lambda}$

For example, for a GaAs LED emitting at $\lambda = 870$ nm at room temperature, you get $\Delta E \simeq 46$ meV and $\Delta\lambda \simeq 28$ nm.

The width of the spectral distribution of the output power of an LED emission plays an important role from an application point of view.

In particular, although it covers only a small portion of the visible spectrum, it is much narrower than the wavelength range associated with a single color perceived by the human eye.

For example, the color red extends from approximately 625 to 730 nm, which is much wider than the typical spectrum of an LED. For this reason, the emission of an LED is perceived as monochromatic by the human eye.

1.5 LIGHT-EMITTING DIODES

1.5.4 Quantum efficiency

The active region of an ideal LED emits one photon for each electron injected. Therefore, the ideal active region of an LED has a **unity quantum efficiency**.

Internal quantum efficiency η_{int} is defined as:

$$\eta_{int} = \frac{\text{number of photons emitted per second by the active region}}{\text{Number of electrons injected per second}} = \frac{\frac{P_{int}}{h\upsilon}}{\frac{I}{e}}$$

where P_{int} is the optical power emitted by the active region and I is the injection current.

In an ideal LED, all the photons emitted by the active region are also emitted in free space, and the extraction efficiency is equal to one.

In real devices, however, only a fraction of the optical power generated actually exit the semiconductor.

Several mechanisms contribute to losses: part of the radiation can be reabsorbed by the substrate, if it is absorbent at the same wavelength; a part can be absorbed by metal contacts. In addition, total internal reflection phenomena can trap light inside the device, reducing the emission power to the outside.

1.5 LIGHT-EMITTING DIODES

1.5.4 Quantum efficiency

Light extraction efficiency η_{ext} is defined as:

$$\eta_{ext} = \frac{\text{number of photons emitted outside per second}}{\text{number of photons emitted per second by the active region}} = \frac{\frac{P}{h\upsilon}}{\frac{P_{int}}{h\upsilon}}$$

where P is the optical power emitted to the outside.

Without sophisticated manufacturing techniques, it is difficult to exceed η_{ext} = 50%.

External quantum efficiency η is defined as:

$$\eta = \frac{\text{number of photons emitted outside per second}}{\text{Number of electrons injected per second}} = \frac{\frac{P}{h\upsilon}}{\frac{I}{e}} = \eta_{int}\eta_{ext}$$

and gives the ratio between the number of useful photons emitted and the number of electrons injected.

The **power efficiency** (or **wallplug efficiency**) is defined as the ratio between the optical power emitted and the electrical power absorbed IV by the LED when electrically polarized:

$$\eta_{power} = \frac{P}{IV}$$

This quantity measures the overall efficiency of conversion from electrical energy to light energy.

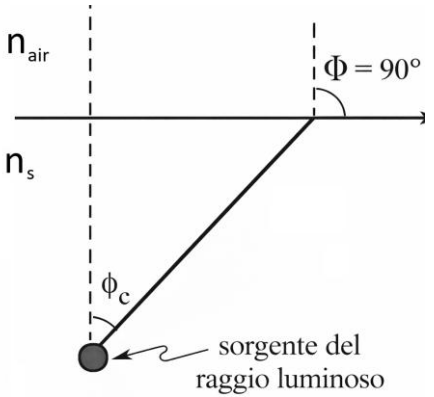
1.5 LIGHT-EMITTING DIODES

1.5.5 The Light Escape Cone

Light generated within a LED may fail to escape from the material due to total reflection within the semiconductor-air interface.

If a light ray strikes the surface almost normally, it can be refracted and emerge outside; on the contrary, for oblique or grazing incidences, total internal reflection may occur.

This phenomenon significantly reduces external efficiency (η_{ext}), especially in LEDs made of semiconductor materials with a high refractive index.



Let ϕ the angle of incidence within the semiconductor at the semiconductor-to-air interface. The angle of the refracted ray Φ is related to ϕ by Snell's law:

$$n_s \sin\phi = n_{air} \sin\Phi$$

where n_s is the refractive index of the semiconductor and n_{air} is that of air.

The condition for the total internal reflection is when the refracted ray is tangent to the surface, i.e. for $\Phi = 90°$, as shown in the Figure. In this case:

$$\sin\phi_c = \frac{n_{air}}{n_s} \sin 90° = \frac{n_{air}}{n_s}$$

1.5 LIGHT-EMITTING DIODES

1.5.5 The Light Escape Cone

$$\sin\phi_c = \frac{n_{air}}{n_s}\sin 90° = \frac{n_{air}}{n_s}$$

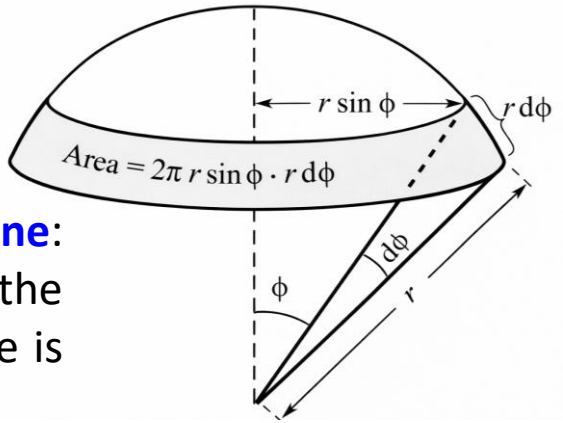
and therefore the **critical angle** is:

$$\phi_c = \sin^{-1}\left(\frac{n_{air}}{n_s}\right)$$

Because the refractive index of semiconductors is typically high, the critical angle is small. For example, for GaAs $n_s \simeq 3.4$, and therefore ϕ_c is on the order of a few degrees.

In this case, the approximation $\sin\phi_c \simeq \phi_c$ can be assumed, and:

$$\phi_c \approx \frac{n_{air}}{n_s}$$



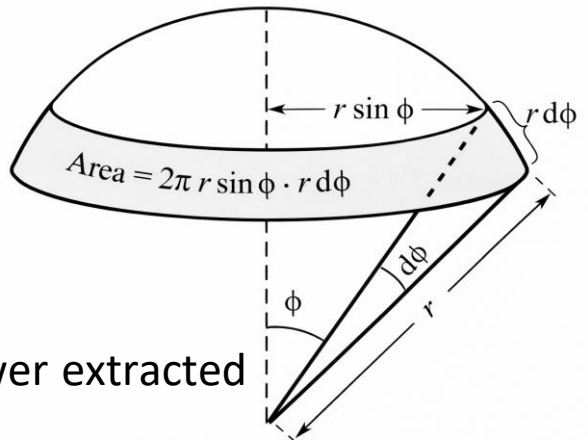
The critical angle defines the so-called **light escape cone**: only the radiation emitted inside this cone can exit the semiconductor, while all the radiation emitted outside is reflected internally.

To evaluate the fraction of power that can escape, consider a point source that isotropically emits a total power P_{source} inside the semiconductor.

1.5 LIGHT-EMITTING DIODES

1.5.5 The Light Escape Cone

The power that emerges on the outside is proportional to the area of the spherical cap delimited by the angle ϕ_c, as schematically represented in the Figure. This area is worth



$$A = \int_0^{\phi_c} 2\pi r sen\phi r d\phi = 2\pi r^2 (1 - cos\phi_c)$$

Since the total area of the sphere is $4\pi r^2$, the fraction of power extracted will be:

$$P_{escape} = P_{source} \frac{2\pi r^2 (1 - cos\phi_c)}{4\pi r^2} = \frac{1}{2} (1 - \cos\phi_c)$$

For small critical angles, the cosine can be expanded in Taylor series, $\cos\phi_c \simeq 1 - \frac{\phi_c^2}{2}$:

$$\frac{P_{escape}}{P_{source}} \approx \frac{1}{2}\left[1 - \left(1 - \frac{\phi_c^2}{2}\right)\right] = \frac{1}{4}\phi_c^2$$

leading to: $\frac{P_{escape}}{P_{source}} \approx \frac{1}{4}\frac{n_{air}^2}{n_s^2}$ $\quad \phi_c \approx \frac{n_{air}}{n_s}$

This relationship shows that, in semiconductors with high refractive index, only a small fraction (< 4% assuming $n_{\text{air}} = 1$ and $n_s > 2.5$) of the light generated can escape from a planar LED.