Non-equilibrium Electrical Properties of Semiconductors

8

8.1 Introduction

In the previous chapter, we established the basic relations and formalism for the distribution of electrons in the conduction band and holes in the valence band at thermal equilibrium.

Although the equilibrium state for electrons and holes in a semiconductor is the result of interactions between carriers or between carriers and phonons, it does not depend on the way this state is reached. The knowledge of the equilibrium properties is therefore not sufficient, and this is all the more true since semiconductor devices usually work under non-equilibrium conditions. In this chapter, we will thus discuss the dynamics of electrons and holes, including electrical conductivity, Hall effect, diffusion, as well as recombination mechanisms.

8.2 Electrical Conductivity

8.2.1 Ohm's Law in Solids

Because electrons and holes are charged particles, they can move in an orderly manner in a semiconductor under the influence of an electric field, for example. This motion generates an electrical current, called drift current, which is at the origin of the electrical conductivity phenomenon of certain solids. The magnitude of this current determines whether a solid is a "good" or a "bad" conductor and is directly related to the density of mobile electrical charge carriers in the solid. In this section, we will try to model the electrical conductivity in solids starting from the Drude model, which is a general model and is valid for any solid which contains mobile charge carriers. This model is based on the kinetic theory of gases which was briefly mentioned in Sect. 6.11.

© Springer International Publishing AG, part of Springer Nature 2019

M. Razeghi, *Fundamentals of Solid State Engineering*,
https://doi.org/10.1007/978-3-319-75708-7_8

In this model, an electron from the gas of electrons is considered as (i) a free moving particle in space with a momentum and an energy, (ii) which is subject to instantaneous collision events (e.g., with other particles such as electrons or atom cores or with irregularities in the crystal), (iii) the probability for a collision to occur during an interval of time dt is proportional to dt, (iv) and the particles reach their thermal equilibrium only through these collisions (see the Monte Carlo method in Appendix A.8).

Let us start by conceptually considering an electron with an electrical charge $-q$ in a uniform electric field strength $\vec{E}$. The force exerted on this electron is constant and equal to $-q\,\vec{E}$ ($q > 0$). Newton's action mass law is such that:

$$m\frac{d\,\vec{v}}{dt} = \vec{F} = -q\,\vec{E} \tag{8.1}$$

where $\vec{v}$ is the velocity of the electron and m is its mass (in a semiconductor $m = m_e$, the effective mass). This relation means that the acceleration of the electron is constant and therefore that its velocity increases linearly with time. In practice the velocity does not increase indefinitely, because collisions, which change the energy and or scatter the momentum, prevent the electron velocity from reaching extremely high values.

The current density vector $\vec{J}$ is a vector which is parallel to the flow of charge and whose magnitude is equal to the amount of electrical charge (in Coulomb) that passes per unit time through a unit area surface perpendicular to the flow of charges, as shown in Fig. 8.1a. The current density is expressed in units of $\mathrm{A{\cdot}cm^{-2}}$.

The current density can be determined by calculating the number of electrons which will traverse the surface S, during a time interval dt. Such electrons are in fact located in the volume defined between the surfaces denoted by S and S' in Fig. 8.1b. This volume is equal to $A\left|\vec{v}\right|dt$, where A is the area of the surface S.

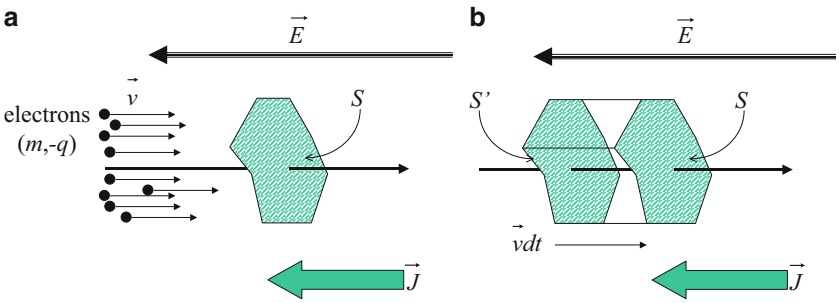


Fig. 8.1 Schematic diagrams showing (**a**) the flow of electrons and current density vector in a uniform electric field, (**b**) the displacement of the surface area A after a time dt at a velocity equal to that of the flowing electrons

Assuming that there is a concentration n of electrons in this region of space and that all of them have a velocity $\vec{v}$, the total amount of electrical charge traversing the surface S with area A, during a time interval dt, is:

$$nqA\left|\vec{v}\right|dt \tag{8.2}$$

The magnitude of the current density is thus the expression in Eq. (8.2) divided by the area and the time interval. Because the current density vector is parallel and in opposite direction to the flow of electrons, we obtain:

$$\vec{J} = -nq\,\vec{v} \tag{8.3}$$

In reality, the electrons are subject to collisions and do not all have the same velocity $\vec{v}$ individually, but they can be considered to have the same averaged velocity, and the expression in Eq. (8.3) remains valid by considering that $\vec{v}$ is the average velocity of the electron gas as a whole. Indeed, if there were no electric field, because collisions are a statistical process, the electrons are as likely to move in one direction in space as in another after a collision. The average velocity vector of the electron gas is thus zero, and there would be no electrical current, as expected (see the Monte Carlo method in Appendix A.8).

In order to calculate the average velocity of the electron gas that results from the electric field, we have to introduce, as was done earlier in Chap. 6, a characteristic time called electron relaxation time τ, which is the average duration between two consecutive collisions or scattering events. Such durations typically range on the order of 10^{-12}–10^{-14} s for electrons in metals. The probability of a collision to occur is in fact proportional to $\frac{1}{\tau}$. The average velocity is then called drift velocity and is denoted $\vec{v}^{\text{drift}}$. This quantity can be estimated by integrating Eq. (8.1) over time from $t = 0$ and $t = \tau$:

$$m\,\vec{v}^{\text{drift}} = -q\tau\,\vec{E} \text{ or } \vec{v}^{\text{drift}} = -\frac{q\tau}{m}\,\vec{E} \tag{8.4}$$

We see that the drift velocity is proportional to the electric field strength and this proportionality factor is called the mobility of electrons in the solid:

$$\begin{cases} \vec{v}^{\text{drift}} = -\mu\,\vec{E} \\ \mu = \dfrac{q\tau}{m} \end{cases} \tag{8.5}$$

This quantity is expressed in units of $\text{cm}^2\cdot\text{V}^{-1}\cdot\text{s}^{-1}$, and it represents the velocity that an electron gains per unit electric field strength (velocity ($\text{cm}\cdot\text{s}^{-1}$) divided by electric field strength ($\text{V}\cdot\text{cm}^{-1}$)). This parameter is not used often in metals but will be most useful to characterize semiconductors. The drift current density, which results from the drift of electrons in the electric field, can then be written using Eq. (8.3):

$$\vec{v}^{\text{drift}} = -nq\ \vec{v}^{\text{drift}} = nq\mu\ \vec{E} = \frac{nq^2\tau}{m}\vec{E} \tag{8.6}$$

A "drift" superscript has been added to emphasize that this is the drift current density. Here again, we see that the current density is proportional to the electric field strength. This proportionality factor is called the conductivity, is denoted σ, and is expressed in units of $\text{S}\cdot\text{cm}^{-1}$ (Siemens per cm) or inverse ($\Omega\cdot$cm):

$$\begin{cases} \vec{J}^{\text{drift}} = \sigma\vec{E} \\ \sigma = nq\mu = \dfrac{nq^2\tau}{m} \end{cases} \tag{8.7}$$

It is also a common practice to consider the inverse of the conductivity which is called the resistivity of the material:

$$\rho = \frac{1}{\sigma} = \frac{1}{nq\mu} \tag{8.8}$$

The linear relation in Eq. (8.7) is called Ohm's law. In strong electric fields, deviations from this linear dependence may occur, but one can keep the general expression for the current density in Eq. (8.7) by considering a field-dependent conductivity σ. In this case, the relation is called the generalized Ohm's law.

Example

Q Estimate the electron mobility in Cu.

A The charge carriers in Cu are electrons, and their mobility μ is related to the resistivity ρ of Cu through $\mu = \frac{1}{nq\rho}$, where n is the electron concentration participating in the conduction. Since there are two electrons in the valence shell of copper, this concentration can be determined by the concentration of Cu atoms or the density of Cu ($d = 8.92\ \text{g}\cdot\text{cm}^{-3}$) through $n = 2 \times \frac{d}{m_{\text{Cu}}}$, where m_{Cu} is the mass of a Cu atom. Assuming the resistivity of Cu is about $\rho = 1.7 \times 10^{-6}\ \Omega\cdot$cm, we get the mobility:

$$\begin{aligned} \mu &= \frac{m_{\text{Cu}}}{2dq\rho} \\ &= \frac{63.55 \times 1.67264 \times 10^{-27}}{2 \times \left(8.92 \times 10^{-3}\right) \times \left(1.60218 \times 10^{-19}\right) \times \left(1.7 \times 10^{-6}\right)} \\ &= 21.9\,\text{cm}^2/\text{Vs} \end{aligned}$$

For many, Ohm's law is more commonly recognized through the relation "$I = \frac{V}{R}$," where I is the current, V the voltage, and R the resistance of an electrical component. Indeed, let us consider a parallelepiped-shaped solid, as depicted in Fig. 8.2. We assume the electric field in the solid is uniform and that the electrical current flows perpendicularly to a side of the parallelepiped with surface area WH, as shown in Fig. 8.2.

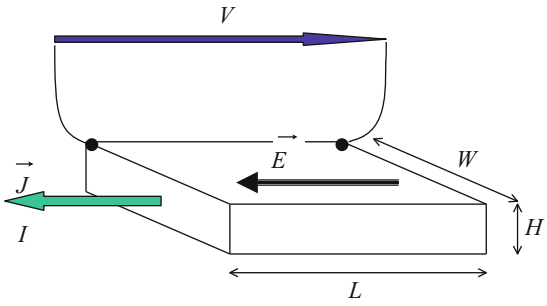


Fig. 8.2 Schematic diagram illustrating the geometry used to illustrate Ohm's law. A voltage is applied across two opposite faces of a rectangular solid and separated by a distance L. This results in an electric field and a current density perpendicular to these two faces

In this configuration, the electrical current I is equal to the magnitude of the current density multiplied by the area WH, i.e., $I = \mathrm{WHJ_{drift}}$. The voltage V is equal to the magnitude of the electric field strength $\left|\vec{E}\right|$ multiplied by the length L of solid considered, i.e., $V = L\left|\vec{E}\right|$. We therefore get successively:

$$\begin{aligned} I &= \mathrm{WHJ^{drift}} = \mathrm{WH}\sigma\left|\vec{E}\right| \\ &= \frac{\mathrm{WH}}{L}\sigma L\left|\vec{E}\right| = \frac{\mathrm{WH}}{L}\sigma V \end{aligned} \tag{8.9}$$

We thus recognize the relation:

$$I = \frac{V}{R} \tag{8.10}$$

where:

$$R = \frac{L}{\mathrm{WH}}\frac{1}{\sigma} \mathrm{or} R = \frac{L}{A}\rho \tag{8.11}$$

and $A = \mathrm{WH}$ is the area of the surface perpendicular to and traversed by the electrical current flow. The quantity R is called the resistance of slab of solid considered. This expression relates a macroscopic quantity (resistance) to an internal property of the solid (resistivity).

8.2.2 The Case of Semiconductors

So far, the discussion has been general and valid for any solid that contains mobile charge carriers. In the case of semiconductors, a few modifications to the previous results need to be made.

A semiconductor has two types of charge carriers which can contribute to the electrical conduction: electrons in the conduction band and holes in the valence band. There are thus two separate contributions to the drift current:

$$\vec{J}^{\text{drift}} = \vec{J}_{\text{e}}^{\text{drift}} + \vec{J}_{\text{h}}^{\text{drift}}$$

where each of the $\vec{J}_{\text{e}}^{\text{drift}}$ and $\vec{J}_{\text{h}}^{\text{drift}}$ is expressed through Eq. (8.7) using the carrier concentrations n and p, mobilities μ_{e} and μ_{h}, and effective masses m_{e} and m_{h} of the electron and the hole, respectively, in the semiconductor considered. Note that, unlike electrons, the holes flow in the same direction as the electric field, because of their positive charge. We thus obtain:

$$\begin{cases} \vec{v}_{\text{e}}^{\text{drift}} = -\mu_{\text{e}} \vec{E} \\ \vec{v}_{\text{h}}^{\text{drift}} = +\mu_{\text{h}} \vec{E} \end{cases} \text{and} \begin{cases} \vec{J}_{\text{e}}^{\text{drift}} = -nq\, \vec{v}_{\text{e}}^{\text{drift}} \\ \vec{J}_{\text{h}}^{\text{drift}} = +pq\, \vec{v}_{\text{h}}^{\text{drift}} \end{cases}. \tag{8.12}$$

The total drift current density can then be written as:

$$\begin{cases} \vec{J}^{\text{drift}} = \sigma \vec{E} \\ \sigma = q(n\mu_{\text{e}} + p\mu_{\text{h}}) \end{cases} \tag{8.13}$$

The typical room temperature conductivity in metals is $(0.1 \sim 3) \times 10^4$ S·cm^{-1}, while the conductivity in semiconductors depends on the carrier concentrations and therefore the doping level, as discussed in Chap. 7.

The conductivity in semiconductors depends much more strongly on the temperature than that in metals. This is because in semiconductors, at a temperature of 0 K, the Fermi energy lies within the forbidden gap (Fig. 5.11) and there is no electron in the conduction band (and thus no hole in the valence band) as the Fermi-Dirac distribution is strictly equal to zero there (Fig. 5.12). Remember that a full band does not carry current. By increasing the temperature, it is therefore possible to increase the concentrations of electrons in the conduction band, holes in the valence band, and enhance electrical conductivity as the Fermi-Dirac distribution is not strictly equal to zero any more. By contrast, in metals, the Fermi energy lies within the conduction band which is thus partially filled (Fig. 5.11), and an increase in temperature will not significantly affect the concentration of electrons in it.

8.3 Carrier Mobility in Solids

The mobility of electrons is controlled by two physical parameters: one is the effective mass and the other is the relaxation time. In Chap. 5, we have seen what determines the effective mass of a charge. Let us now consider the momentum lifetime. The scattering processes which determine the momentum lifetime of solids can be classified into two categories: (a) elastic scattering processes and (b) inelastic scattering processes. In category (a), the carrier changes its momentum but not its energy. Any break in the translational symmetry of the solid will give rise to elastic scattering, and in particular this includes the presence of impurity potentials, defects

interfaces, and dislocations, but there are also the deviations from periodic order caused by lattice vibrations: the electron-phonon interactions. The former contribute to category (a), and the latter involve energy exchange with the lattice and are in category (b). In category (b) the carrier changes both momentum and energy. An inelastic phonon-induced scattering process is allowed if it satisfies both the momentum and energy conservation conditions which are, respectively:

$$\vec{k'} = \vec{k} \pm \vec{q}$$
$$E\left(\vec{k'}\right) = E\left(\vec{k}\right) \pm \hbar\omega\left(\vec{q}\right)$$

where $E\left(\vec{k'}\right)$ is the energy of the particle after the scattering process and $\hbar\omega\left(\vec{q}\right)$is the energy of the phonon absorbed or emitted.

We have seen in Chap. 6 that a solid will in general have two types of phonons, so there are also two types of electron-phonon scattering processes. These are the electron-acoustic and electron-optical phonon scattering processes. The acoustic scattering occurs in all solids, but optic phonon scattering can only take place when there are optic modes in the system. The strength of the electron-acoustic and electron-optical coupling determines the efficiency, or the rate at which a carrier with a given momentum $\vec{k}$ is scattered into a momentum state $\vec{k'}$ via a phonon. In III–V semiconductors with polar modes, the electron optic coupling is an efficient process and is the most important mechanism by which hot carriers relax their excess energy when they have enough energy to emit an optic phonon. An electron can also absorb an optic phonon, but this is only possible if a sufficient number is thermally excited. The rate of optic phonon absorption increases therefore with temperature, following essentially the Bose-Einstein distribution law of phonon occupation. When more than one scattering process is contributing, the sum must be taken. This is done by summing the lifetimes in parallel so that the shortest time dominates. The total lifetime τ is thus given by the sum $\frac{1}{\tau} = \frac{1}{\tau_{\text{el}}} + \frac{1}{\tau_{\text{op}}} + \frac{1}{\tau_{\text{ac}}}$where the terms denote the inverse of the elastic, optic, and acoustic scattering lifetimes, respectively. The temperature dependence of the mobility in different materials is not simple to summarize, and the reader is referred to the specialized textbooks by Ridley and Sze. The physics of the situation however is as follows: at very low temperatures, the phonon modes freeze out and thermal velocities are low, the inelastic lifetimes therefore increase as we go down in temperature, and eventually elastic processes dominate. Elastic scattering processes can however be weakly dependent on temperature and will remain finite even at zero temperature creating a finite resistance unless the material becomes a superconductor at some stage. Elastic scattering can take place from neutral defects, and most effectively also from charged ionized defects and impurities. The state of ionization of an impurity will in general be a function of temperature, as we saw when we discussed doped semiconductors (see Sect. 7.6). This means that elastic scattering processes in doped semiconductors will in general have both strong temperature-dependent and weak temperature-dependent components. Here are a few typical measured bulk values

(see also Appendix A4) of the room temperature ($T = 300$ K) mobilities of some important semiconductors: Si electrons, 1500 cm^2/Vs; Si holes, 450 cm^2/Vs; GaAs electrons, 8500 cm^2/Vs; GaAs holes, 400 cm^2/Vs; InAs electrons, 33,000 cm^2/Vs; and InAs holes, 460 cm^2/Vs. From the example in the text, we calculated the mobility of Cu, which is a good metal, to be ~20 cm^2/Vs. This is typical for good metals and interestingly lower than for many semiconductors.

8.4 Hall Effect

At the end of the nineteenth century, physicists knew that if a metal wire carrying an electrical current was placed in a magnetic field, it experienced a force. The origin of this force was not known. In 1879, E.H. Hall tried to prove that this force was exerted only on the mobile charges (electrons) in the wire. By doing so, he conducted an experiment where an electrical current was run through a fixed conductor perpendicularly to a magnetic field.

Let us consider the Hall effect experiment geometry illustrated in Fig. 8.3. An electrical current, with current density $\vec{J}$ in the x-direction, is run through a parallelepiped-shaped solid. A magnetic induction or flux density $\vec{B}$ is directed perpendicularly to the current, in the z-direction. The movement of holes and electrons is shown in Fig. 8.3 as well.

8.4.1 *P*-Type Semiconductor

Let us now assume that the solid only contains one type of charge carriers and that they are holes. With the electrical current in the *(+x)*-direction, a hole moves also in the x-direction with a velocity $\vec{v}_h$, as shown in Fig. 8.3 . At the same time, it is subject to the Lorentz force equal to:

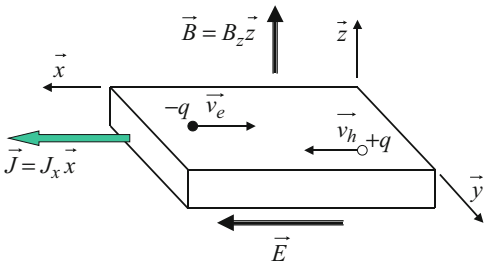


Fig. 8.3 Geometry used for a Hall effect experiment. A uniform electric field strength is applied inside a solid in the x-direction (e.g., by applying a voltage across the solid), which results in an electric current in the same direction. The movement of holes and electrons in the solid is shown. The solid is immersed in a magnetic induction which is directed in the z-direction, perpendicularly to this electric field

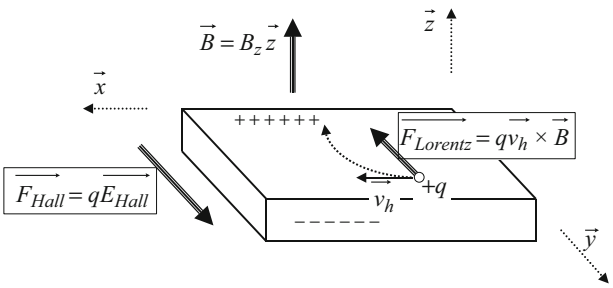


Fig. 8.4 Motion of a hole in the Hall effect experiment. Under the influence of the Lorentz force, the motion of holes is deviated in the *y*-direction toward one side of the solid which then becomes positively charged through the accumulation of holes. The opposite side of the solid therefore becomes negatively charged. This gives rise to an additional electric field which is directed in the y-direction

$$\vec{F}_{\text{Lorentz}} = q\,\vec{v}_{\text{h}} \times \vec{B} \tag{8.14}$$

which is in the *y*-direction. If the sample was without limits, the hole would exhibit a cyclotron (circular) motion around an axis parallel to $\vec{B}$. In the case of a finite size solid as the one shown in Fig. 8.4, holes would accumulate on one of its sides to create a surplus of positive charges. At the same time, negative charges would appear on the opposite side from the deficiency of holes there. This separation of charges results in an electric field strength $\vec{E}_{\text{Hall}}$, called Hall electric field and shown in Fig. 8.4, which drives holes in the *y*-direction and is opposite to the Lorentz force.

At equilibrium, the Lorentz force and the force due to the Hall electric field must balance each other. This can be expressed mathematically as:

$$\vec{0} = \vec{F}_{\text{h, Lorentz}} + \vec{F}_{\text{Hall}} = q\,\vec{v}_{\text{h}} \times \vec{B} + q\,\vec{E}_{\text{Hall}} \tag{8.15}$$

The Hall electric field strength is thus:

$$\vec{E}_{\text{Hall}} = -\,\vec{v}_{\text{h}} \times \vec{B} \tag{8.16}$$

The component of the Hall electric field strength in the *y*-direction (i.e., $\vec{E}_{\text{Hall}} = (E_{\text{Hall}})_y\,\vec{y}$) in the geometry shown in Fig. 8.4 is:

$$(E_{\text{Hall}})_y = (v_{\text{h}})_x B_z > 0 \tag{8.17}$$

From Eq. (8.12), we get:

$$J_x = pq(v_{\text{h}})_x$$

where p is the hole concentration in the solid, and we can rewrite Eq. (8.18) as:

$$(E_{\text{Hall}})_y = \frac{J_x}{pq} B_z \tag{8.18}$$

This expression contains macroscopic quantities which are characteristic of the material (p), parameters of the experiments (J and B), and quantities which are experimentally measured (E_{Hall}). Through this relation, we can easily extract properties characteristic of the materials from experiments. It is a common practice to introduce the Hall constant given by:

$$R_{\text{H}} = \frac{(E_{\text{Hall}})_y}{J_x B_z} = \frac{1}{pq} > 0 \tag{8.19}$$

The Hall constant therefore yields a direct measure of the hole concentration in the solid. We can define a hole Hall mobility as:

$$\mu_{\text{H,h}} = \sigma R_{\text{H}} \tag{8.20}$$

This Hall mobility has the same units as the drift mobility encountered in Eq. (8.12) in Sect. 8.2, i.e., $\text{cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$. However, it differs from the drift mobility by a factor, called the Hall factor, which is determined by the temperature and the types of scattering involving the charge carriers. Experimentally, this factor is taken to be equal to unity and only one mobility is considered. This can be illustrated by the fact that one can arrive at Eq. (8.20) from Eq. (8.19) by using the expression in Eq. (8.13) applied to holes only.

8.4.2 *N*-Type Semiconductor

In the case of a solid which contains only electrons as the mobile charge carriers, a similar analysis can be conducted. The motion of an electron in the Hall effect experiment is shown in Fig. 8.5. We can see that the electrons are deflected in the same direction as the holes in Fig. 8.6.

However, because electrons have a negative charge, the Hall electric field is in the opposite direction in comparison to the one from holes:

$$\vec{0} = \vec{F}_{\text{e,Lorentz}} + \vec{F}_{\text{Hall}} = -q\,\vec{v}_{\text{e}} \times \vec{B} - q\,\vec{E}_{\text{Hall}} \tag{8.21}$$

The Hall electric field strength is thus:

$$\vec{E}_{\text{Hall}} = -\vec{v}_{\text{e}} \times \vec{B} \tag{8.22}$$

The component of the Hall electric field strength in the y-direction (i.e., $\vec{E}_{\text{Hall}} = (E_{\text{Hall}})_y\,\vec{y}$) in the geometry shown in Fig. 8.5 is:

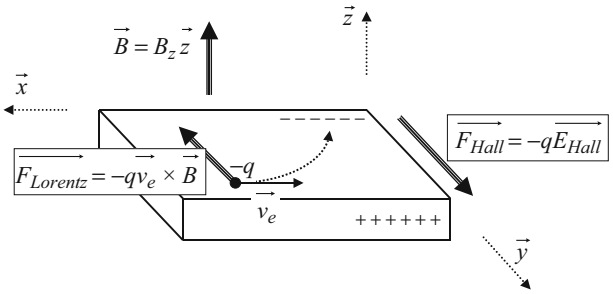


Fig. 8.5 Motion of an electron in the Hall effect experiment. Under the influence of the Lorentz force, the motion of electrons is deviated in the y-direction toward one side of the solid which then becomes negatively charged through the accumulation of electrons. The opposite side of the solid therefore becomes positively charged. This gives rise to an additional electric field which is directed in the y-direction

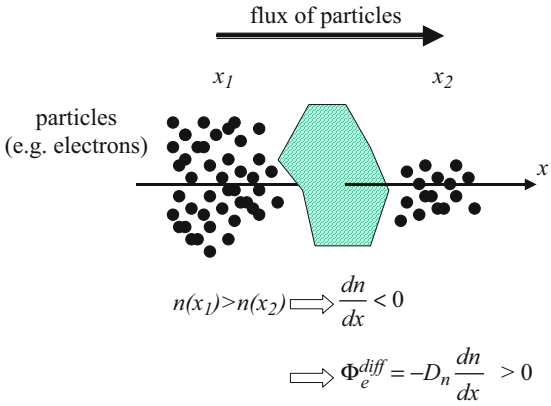


Fig. 8.6 Diffusion of particles (e.g., electrons) in a one-dimensional model. An imaginary surface with a unit area is considered, such that the concentration of particles on one side is larger than on the other side. The diffusion process is characterized by the flux of particles spontaneously passing through the imaginary surface per unit time

$$(E_{\text{Hall}})_y = (v_e)_x B_z < 0 \tag{8.23}$$

because $(v_e)_x < 0$. From Eq. (8.12), we have:

$$J_x = -nq(v_e)_x$$

and we can rewrite Eq. (8.23) as:

$$(E_{\text{Hall}})_y = -\frac{J_x}{nq} B_z \tag{8.24}$$

This expression is similar to Eq. (8.18), and the Hall constant defined in Eq. (8.19) becomes:

$$R_{\mathrm{H}} = \frac{(E_{\mathrm{Hall}})_y}{J_x B_z} = -\frac{1}{nq} < 0 \tag{8.25}$$

Here again, we see that the Hall constant yields the electron concentration in the solid. Moreover, it is negative, whereas it was positive when holes were the only charge carrier. The Hall constant is therefore a good method to determine if a semiconductor is *p*-type or *n*-type. The electron Hall mobility given by Eq. (8.20) is now transformed into:

$$\mu_{\mathrm{H,e}} = \sigma |R_{\mathrm{H}}| > 0 \tag{8.26}$$

Similar to the previous case, the electron Hall mobility is usually taken equal to the electron drift mobility.

8.4.3 Compensated Semiconductor

In a compensated semiconductor, both types of dopants are simultaneously present in the material. Since the electrons and holes released by the doping can recombine, a decrease of the free carriers' concentration can be observed. Adding *p*-type impurities to an *n*-doped system will therefore reduce the electron concentration and vice versa. The charged impurities are still there, having transferred the charge to each other (donor to acceptor) rather than to the bands. It is possible in this way to increase the resistance of doped systems by adding the opposite type of dopant. This can be very useful when ion implantation is used to dope a material, because with ions, one can in principle achieve a high degree of spatial resolution and select the depth of implantation. The ion beam can also be focused to compensate the local doping and thus produce submicron devices.

8.4.4 Hall Effect with Both Types of Charge Carriers

When both electrons and holes are contributing to the transport process, the calculation of the Hall coefficient is somewhat more complicated. Both types of carriers will contribute to the Hall effect in an intrinsic material, for example, or when light is photo-exciting pairs, or when electrons and holes are injected using different types of source drain electrode materials. The derivation of R_{H} is however straightforward and can be done by using the Newton law with the Lorentz force for both carriers:

$$\begin{aligned} m_{\mathrm{e}} \frac{dv_x}{dt} + m_{\mathrm{e}} v_x \frac{1}{\tau} &= -qE_x - qv_y B \\ m_{\mathrm{e}} \frac{dv_y}{dt} + m_{\mathrm{e}} v_y \frac{1}{\tau} &= -qE_y - qv_x B \end{aligned} \tag{8.27}$$

in the presence of electric fields, E_x, E_y, and a magnetic field B. Similar equation can be written down for holes, except that $q \rightarrow -q$. The steady-state velocities are obtained by assuming that the velocity no longer changes with time, i.e., by putting the acceleration term equal to zero. The products μB are very small under typical measurement conditions, so if we ignore even smaller terms of order B^2, we can write Eq. (8.27) as:

$$\begin{aligned} v_y^{\mathrm{e}} &= -\mu_{\mathrm{e}} E_y - \mu_{\mathrm{e}}^2 B E_x \\ v_y^{\mathrm{h}} &= \mu_{\mathrm{h}} E_y - \mu_{\mathrm{h}}^2 B E_x \end{aligned} \tag{8.28}$$

The above equations can be related to the total current J_y giving:

$$J_y = nq\mu_{\mathrm{e}}\left(E_y + \mu_{\mathrm{e}} E_x B\right) + pq\mu_{\mathrm{h}}\left(E_y - \mu_{\mathrm{h}} B E_x\right) \tag{8.29}$$

Under equilibrium condition, i.e., when the current $J_y = 0$, the ratio of the components of the electric field is such that:

$$\frac{E_y}{E_x} = \left\{\frac{p\mu_h^2 - n\mu_e^2}{n\mu_e + p\mu_h}\right\} B \tag{8.30}$$

and the Hall constant is now given by:

$$R_H = \frac{1}{q}\,\frac{p\mu_h^2 - n\mu_e^2}{\left(p\mu_h + n\mu_e\right)^2} \tag{8.31}$$

where p and n are the hole and electron concentrations and μ_{h} and μ_{e} are the hole and electron mobilities, all of which are positive parameters. The Hall mobility is the combination of the mobilities of the electrons and holes and given by:

$$\mu_{\mathrm{H}} = \sigma |R_{\mathrm{H}}| = \left|\frac{p\mu_{\mathrm{h}}^2 - n\mu_{\mathrm{e}}^2}{p\mu_{\mathrm{h}} + n\mu_{\mathrm{e}}}\right| \tag{8.32}$$

8.5 Charge Carrier Diffusion

In an inhomogeneous solid, certain regions may exhibit more electrons or holes than other regions. These will then migrate from the high concentration areas to the low concentration areas. This is a universal and natural phenomenon, called diffusion. This process is due to an imbalance in the thermodynamic chemical potential. One may picture the diffusion process as a drop of ink in a glass of clear water which slowly spreads in the entire volume of water. Because electrons and holes are charge carriers, their diffusion generates an electrical current, which is very important in many semiconductor devices.

8.5.1 Diffusion Currents

In this section, we will describe a simple one-dimensional model for the diffusion of electrons and holes in a semiconductor. Let us assume the electron concentration $n(x)$ is not uniform in the x-direction, as schematically illustrated in Fig. 8.6 .

The diffusion process is mathematically described by Fick's first law of diffusion which says that the flux, i.e., the number of electrons passing per unit time a unit area surface perpendicular to the x-direction, is given by:

$$\Phi_{\mathrm{e}}^{\mathrm{diff}} = -D_{\mathrm{n}}\frac{dn}{dx} \tag{8.33}$$

where D_{n} is called the diffusion coefficient or diffusivity and has the units of $\mathrm{cm}^2 \cdot \mathrm{s}^{-1}$. We use the subscript "n" to identify that this is the diffusivity for electrons. The negative sign in this expression means that the flux of electrons is in the direction opposite to the gradient (or slope) of concentration, as illustrated in Fig. 8.6.

Using a similar approach as for the electrical drift process in Sect. 8.2 to count the number of electrons that pass the unit area surface in Fig. 8.6 per unit time, we can extract the electron diffusion velocity $v_{\mathrm{h}}^{\mathrm{diff}}$:

$$\Phi_{\mathrm{e}}^{\mathrm{diff}} = n v_{\mathrm{e}}^{\mathrm{diff}} \tag{8.34}$$

which leads to the relation:

$$v_{\mathrm{e}}^{\mathrm{diff}} = -D_{\mathrm{n}}\frac{1}{n}\frac{dn}{dx} \tag{8.35}$$

The movement of these electrons creates an electrical current. The diffusion current density of electrons is then determined from Eq. (8.12):

$$J_{\mathrm{e}}^{\mathrm{diff}} = -nqv_{\mathrm{e}}^{\mathrm{diff}} = +qD_{\mathrm{n}}\frac{dn}{dx} \tag{8.36}$$

Similar relations to Eqs. (8.35) and (8.36) can be obtained for the diffusion of holes:

$$v_{\mathrm{h}}^{\mathrm{diff}} = -D_{\mathrm{p}}\frac{1}{p}\frac{dp}{dx} \tag{8.37}$$

$$J_{\mathrm{h}}^{\mathrm{diff}} = -pqv_{\mathrm{h}}^{\mathrm{diff}} = -qD_{\mathrm{p}}\frac{dp}{dx} \tag{8.38}$$

where p is the concentration of holes. Note that there is a sign change from Eqs. (8.36), (8.37), and (8.38) which is due to the positive charge of the hole. There is no such sign change from Eqs. (8.35), (8.36), and (8.37), because the origin of the diffusion process is not dependent on the electrical charge.

8.5.2 Einstein Relations

The drift and the diffusion of electrons and holes are intimately related processes, because both contribute to the observed electrical current in a semiconductor.

Let us continue on our simple one-dimensional model and consider a finite size solid onto which a uniform external electric field of strength $\vec{E} = E\,\vec{x}$ is applied. As a result, the electrons will be drifting to one side of the solid, and a concentration gradient will be achieved. These electrons will then start to diffuse in the direction opposite to this electrical drift until a balance is reached.

The drift current density is given by Eq. (8.12) $J_{\mathrm{e}}^{\mathrm{drift}} = nq\mu_{\mathrm{e}}E$, while the electrical diffusion current density is given by Eq. (8.36) $J_{\mathrm{e}}^{\mathrm{drff}} = +qD_{\mathrm{n}}\dfrac{dn}{dx}$. At the thermal equilibrium of this system, the sum of these two current densities:

$$J_{\mathrm{e}}^{\mathrm{drift}} + J_{\mathrm{e}}^{\mathrm{diff}} = nq\mu_{\mathrm{e}}E + qD_{\mathrm{n}}\frac{dn}{dx} \tag{8.39}$$

must be equal to zero, i.e.:

$$nq\mu_e E + qD_n\frac{dn}{dx} = 0 \tag{8.40}$$

This first-order differential equation can be rewritten as:

$$\frac{dn}{dx} + \frac{\mu_{\mathrm{e}}E}{D_{\mathrm{n}}}n = 0 \tag{8.41}$$

which leads to the solution:

$$n(x) = n(0)\exp\left(-\frac{\mu_{\mathrm{e}}Ex}{D_{\mathrm{n}}}\right) \tag{8.42}$$

where $n(0)$ is the electron concentration at $x = 0$. We see that we obtain an exponential-like distribution for this concentration. However, at thermal equilibrium, this quantity also obeys Boltzmann statistics, which is analogous to the Boltzmann probability distribution we encountered in Chap. 5. For a nondegenerate semiconductor, the electron concentration according to Boltzmann statistics should be given by:

$$n(x) = n(0)\exp\left(-\frac{qEx}{k_{\mathrm{b}}T}\right) \tag{8.43}$$

because qEx is the potential energy of the electron in an electric field strength of magnitude E. Comparing Eqs. (8.42) and (8.43), we obtain the relation:

$$\frac{\mu_e E}{D_n} = \frac{qE}{k_b T}$$

or:

$$\frac{D_n}{\mu_e} = \frac{k_b T}{q} \tag{8.44}$$

A similar relation can be obtained for holes:

$$\frac{D_p}{\mu_h} = \frac{k_b T}{q} \tag{8.45}$$

Equations (8.44) and (8.45) are called the Einstein relations and are valid only for nondegenerate semiconductors. For degenerate semiconductors, we first need to specify the amount of charge in the bands, and a factor involving the Fermi-Dirac integral Eq. (7.13) needs to be included in the above expressions. These relations are important because they provide a mathematical link between the drift and diffusion processes. They are however not always valid. They apply only when there is a small amount of charge in the band edges, which is the most interesting situation in semiconductor technology.

8.5.3 Diffusion Lengths

In the diffusion model considered so far, an electron or a hole can diffuse indefinitely in space. However, in most real case situations, the diffusion range is much more limited.

Let us consider the diffusion of electrons in a one-dimensional semiconductor model, where excess carriers are continuously generated at $x = 0$ and are then allowed to diffuse toward $x \to \infty$. By the term "excess carriers," we mean that an amount of electrons in addition to the thermal equilibrium concentration n_0 is injected into the semiconductor. The mechanisms by which this is achieved will be discussed later in the text. We will denote:

$$\Delta n(x) = n(x) - n_0 \tag{8.46}$$

the excess electron concentration which is a function of position. A possible shape for $\Delta n(x)$ is shown in Fig. 8.7.

During the diffusion process, an electron will experience recombination, i.e., they will not travel in space indefinitely but will be stopped, for example, when it encounters a hole (remember that a hole is an allowed state vacated by an electron) or when it gets trapped by a defect in the semiconductor crystal (e.g., an ionized donor which is positively charged).

The recombination mechanisms are numerous and diverse. However, it is possible to mathematically express their effects in a simple manner. For this, we introduce

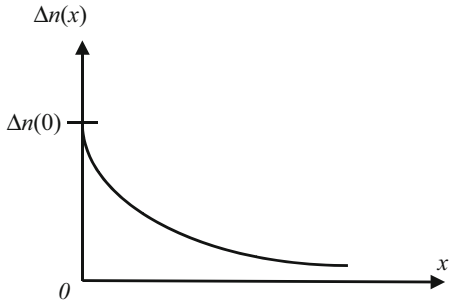


Fig. 8.7 Excess electron concentration in a one-dimensional model. The excess concentration decreases, as it gets deeper into the material as a result of recombination. The decrease has an exponential dependence

a characteristic time, τ_n, called the electron recombination lifetime, such that the recombination rate of an electron at a location where there is an excess $\Delta n(x)$ of electrons is given by:

$$R(x) = \frac{\Delta n(x)}{\tau_n} \tag{8.47}$$

This quantity has the units of $\mathrm{cm}^{-3}\cdot\mathrm{s}^{-1}$ and expresses the change in the excess carrier concentration per unit time.

Let us now consider an infinitesimal region of space, located between x_0 and $x_0 + dx$, as illustrated in Fig. 8.8. This region experiences an influx and an outflux of electrons, denoted, respectively, $\left(\Phi_e^{diff}\right)_{in}$ and $\left(\Phi_e^{diff}\right)_{out}$ and shown in Fig. 8.8.

If $\left(\Phi_e^{diff}\right)_{in} > \left(\Phi_e^{diff}\right)_{out}$, there is a net influx or accumulation of electrons, but if $\left(\Phi_e^{diff}\right)_{out} > \left(\Phi_e^{diff}\right)_{in}$, there is a net outflux or depletion of electrons in this region. Under steady-state conditions, there must not be a never-ending accumulation or depletion of electrons. The influx of electrons must therefore be equal to the sum of the outflux of electrons and the number of electrons recombining within this region. The later quantity is equal to $R(x_0)$ multiplied by the width of the region dx, because we can assume that the function $R(x)$ does not vary too much over a narrow width dx around the point x_0. Numerically, this translates into:

$$\left(\Phi_e^{diff}\right)_{in} = \left(\Phi_e^{diff}\right)_{out} + R(x_0)dx \tag{8.48}$$

From Eq. (8.33), we can write:

$$\left(\Phi_e^{diff}\right)_{in} = -D_n\left(\frac{dn}{dx}\right)_{x=x_0} \text{ and } \left(\Phi_e^{diff}\right)_{out} = -D_n\left(\frac{dn}{dx}\right)_{x=x_0+dx}$$

But, from Eq. (8.46), we easily see that $\frac{dn}{dx} = \frac{d(\Delta n)}{dx}$ and therefore:

$$\begin{cases} \left(\Phi_e^{diff}\right)_{in} = -D_n\left(\frac{d(\Delta n)}{dx}\right)_{x=x_0} \\ \left(\Phi_e^{diff}\right)_{out} = -D_n\left(\frac{d(\Delta n)}{dx}\right)_{x=x_0+dx} \end{cases} \tag{8.49}$$

Equation (8.48) becomes then:

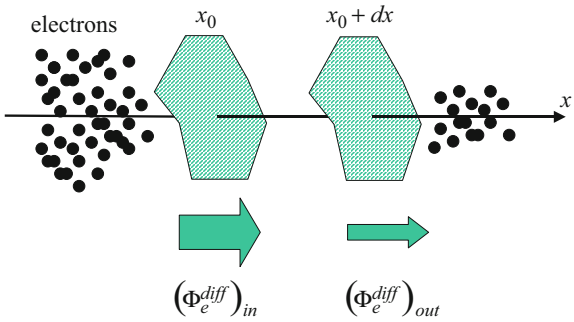


Fig. 8.8 Schematic of the influx and outflux of electrons in a region of space, in a one-dimensional model. In this experiment, the region between the two surfaces located at x_0 and $x_0 + dx$ is considered. This experiment is aimed at determining the net change in carrier concentration in it as a result of the diffusion of particles and their recombination

$$-D_{\mathrm{n}}\left(\frac{d(\Delta_{\mathrm{n}})}{dx}\right)_{x=x_0} = -D_{\mathrm{n}}\left(\frac{d(\Delta n)}{dx}\right)_{x=x_0+dx} + R(x_0)dx$$

which can be rewritten as:

$$D_{\mathrm{n}}\frac{\left(\frac{d(\Delta n)}{dx}\right)_{x=x_0+dx} - \left(\frac{d(\Delta n)}{dx}\right)_{x=x_0}}{dx} = R(x_0)$$

At the limit of $dx \to 0$, i.e., an infinitesimal region, the left-hand side expression becomes the derivative of $\frac{d(\Delta n)}{dx}$ evaluated at $x = x_0$, i.e.:

$$D_{\mathrm{n}}\left(\frac{d^2(\Delta n)}{dx^2}\right)_{x=x_0} = R(x_0) \tag{8.50}$$

This relation is valid for any arbitrarily chosen position x_0, which means that the following equation must be satisfied:

$$D_{\mathrm{n}}\frac{d^2(\Delta n)}{dx^2} = R(x)$$

Equating to Eq. (8.47), we get the differential equation that governs the shape of the excess electron concentration $\Delta n(x)$:

$$D_{\mathrm{n}}\frac{d^2(\Delta n)}{dx^2} = \frac{\Delta n}{\tau_{\mathrm{n}}} \tag{8.51}$$

This equation can be rewritten as:

$$\frac{d^2(\Delta n)}{dx^2} - \frac{\Delta n}{D_{\mathrm{n}}\tau_{\mathrm{n}}} = 0 \tag{8.52}$$

From this expression, we can easily see that the quantity $D_n\tau_n$ has the same dimension as the square of a distance. We can then define a distance L_n, called diffusion length, for electrons, given by:

$$L_n = \sqrt{D_n\tau_n} \tag{8.53}$$

The solution to Eq. (8.52) then has the general form:

$$\Delta n(x) = Ae^{\frac{x}{L_n}} + Be^{\frac{-x}{L_n}} \tag{8.54}$$

Here A and B are constants and are determined from boundary conditions. For example, let us assume the sample is delimited by $x = 0$ and $x \to \infty$, and that is thick enough so that all the excess electrons have been recombined before they reach its limit: $\Delta n \to 0$ when $x \to \infty$ as shown in Fig. 8.7. We thus have:

$$\Delta n(x) = \Delta n(0)e^{-\frac{x}{L_n}} \tag{8.55}$$

From this expression, we see the significance of the diffusion length in determining the spatial distribution of the electrons in the diffusion process as the characteristic length of path that a particle travels before recombining.

A similar diffusion length can be determined for holes and is given by:

$$L_p = \sqrt{D_p\tau_p} \tag{8.56}$$

where τ_p is the hole recombination lifetime.

Example

Q Assuming that in n-type silicon the characteristic time for the minority carriers (holes) is $\tau_p = 2 \times 10^{-10}$ s, estimate the diffusion length of these minority carriers at 300 K.

A The diffusion length is given by $L_p = \sqrt{D_p\tau_p}$. From the Einstein relations, we can determine the diffusion coefficient $D_p = \frac{k_b T\mu_h}{q}$. With the hole mobility in silicon being about $\mu_h = 450$ cm^2/Vs, we get:

$$\begin{aligned} L_p &= \sqrt{\frac{k_b T\mu_h}{q}\tau_p} \\ &= \sqrt{\frac{\left(1.38066 \times 10^{-23}\right) \times 300 \times \left(450 \times 10^{-4}\right)}{1.60218 \times 10^{-19}} \times 2 \times 10^{-10}} \\ &= 0.48\,\mu m \end{aligned}$$

8.6 Carrier Generation and Recombination Mechanisms

In the previous section, we briefly talked about excess carriers and their recombination. We also introduced a single recombination lifetime τ in order to avoid a detailed description of all the recombination processes.

Excess of carriers can exist when the semiconductor is not in its equilibrium state, as a result of additional energy that it received from phonons (heat), photons (light), or an electric field, for example. In a recombination process, the amount of excess carriers is reduced, and the excess energy is transferred or released.

In this section, we will discuss the four most important recombination mechanisms encountered in semiconductors, including direct band-to-band, Shockley-Read-Hall, Auger, and surface recombination. We will also attempt to express the recombination lifetime in each case in terms of known semiconductor parameters.

We will denote by:

$$\begin{cases} \Delta n(t) = n(t) - n_0 \\ \Delta p(t) = p(t) - p_0 \end{cases} \tag{8.57}$$

the excess electron and hole concentrations, respectively, where n_0 and p_0 are the equilibrium electron and hole concentrations.

It is important, at this time, to clearly distinguish equilibrium state from steady state. A system is said to be under equilibrium if it is not subject to external fields or forces. A system under the influence of external fields or forces is under steady state if the parameters that describe it (e.g., carrier concentrations) do not vary with time.

8.6.1 Carrier Generation

Before discussing the various recombination mechanisms, we must first review how carriers are generated in the first place. There are essentially two major types of generation.

The first one corresponds to the thermal generation of carriers and exists under all conditions, whether in equilibrium or non-equilibrium. The thermal generation rate will be denoted $G_t(T)$ and is expressed in units of $\text{cm}^{-3}\cdot\text{s}^{-1}$.

The other type is the generation resulting from external factors, such as optical absorption, electrical injection, etc. This process occurs only in non-equilibrium situations, and the associated generation rate, denoted G, is called the excess generation rate.

For each generation mechanism, there exists a recombination mechanism which is its counterpart. The generation and recombination of carriers are inverse processes to each other.

8.6.2 Direct Band-to-Band Recombination

In this type of recombination, an electron from the conduction band recombines with a hole in the valence band. This process is best pictured in the *E-k* diagram shown in Fig. 8.9 .

This recombination can be equivalently viewed as an electron which goes from a state in the conduction band to an allowed state in the valence band. This seems natural if we remember that a hole in the valence band is in fact an allowed electronic state that has been *vacated* by an electron. The energy that the electron thus loses is most often released in the form of a photon or light as shown in Fig. 8.9. We say that this is a radiative recombination.

This process is most likely to occur between the minimum of the conduction band and the maximum of the valence band and at the center of the first Brillouin zone where the momenta of the recombining electron and hole are both zero. Direct band-to-band radiative recombination is therefore most likely to occur in direct bandgap semiconductors, such as GaAs.

Let us look at this recombination mechanism in more detail. In the present case, the recombination rate, first introduced in Eq. (8.47), is proportional to both the concentration of electrons in the conduction band n and that of holes in the valence band p because these are the particles that are recombining. We can then write:

$$R = r(T)n(t)p(t) \tag{8.58}$$

where $r(T)$ is the recombination coefficient, which is expressed in units of $\text{cm}^3 \cdot \text{s}$, and T is the temperature.

In a non-equilibrium situation when the excess generation rate is nonzero, the net change in the electron and hole densities is given by:

$$-\frac{dn}{dt} = -\frac{d(\Delta n)}{dt} = R - G - G_t \tag{8.59}$$

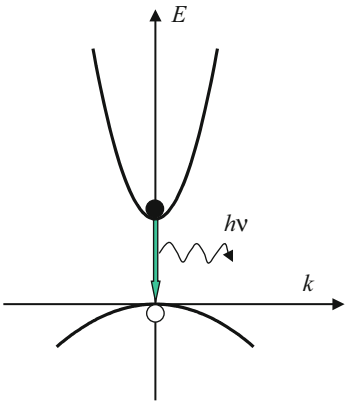


Fig. 8.9 Schematic E-k diagram of a direct band-to-band recombination process. The recombining electron and hole have the same wavevector

where we used the fact that the equilibrium concentration n_0 does not vary with time. At equilibrium, the excess generation rate G is equal to zero; thus the recombination rate must balance the thermal generation rate: $R = G_t$. Since at equilibrium we have $n = n_0$ and $p = p_0$, we can write from Eq. (8.58):

$$G_t = r(T)n_0p_0 \quad \text{or simply} \quad G_t = r(T)n_i^2 \tag{8.60}$$

where n_i is the intrinsic carrier concentration given in Eq. (7.31). From now, we will also omit the temperature dependence of $r(T)$ to simplify the equations.

Let us now consider the relaxation process, which occurs after the external source of generation is removed ($G = 0$). Taking into account Eqs. (8.58) and (8.60), Eq. (8.59) becomes:

$$-\frac{d(\Delta n)}{dt} = r\left[np - n_i^2\right] \tag{8.61}$$

Using Eq. (8.57), we can expand this expression into:

$$-\frac{d(\Delta n)}{dt} = r\left[(n_0 + \Delta n)(p_0 + \Delta p) - n_i^2\right]$$

i.e.:

$$-\frac{d(\Delta n)}{dt} = r\left[n_0p_0 + n_0\Delta p + p_0\Delta n + \Delta n\Delta p - n_i^2\right] \tag{8.62}$$

One obvious simplification can be immediately made in the previous expression as $n_0p_0 = n_i^2$ from Eq. (7.31). For further simplicity, we can assume $\Delta n = \Delta p$, i.e., the concentration of excess electrons is equal to the concentration of excess holes, which seems natural in order to ensure charge neutrality locally in the semiconductor at all times. Equation (8.62) then becomes:

$$-\frac{d(\Delta p)}{dt} = -\frac{d(\Delta n)}{dt} = r[(n_0 + p_0) + \Delta n]\Delta n \tag{8.63}$$

We can successively transform Eq. (8.63) into:

$$-\frac{{}^{d(\Delta n)}/_{dt}}{[(n_0 + p_0) + \Delta n]\Delta n} = r$$

$$\frac{1}{(n_0 + p_0)}\left(\frac{{}^{d(\Delta n)}/_{dt}}{(n_0 + p_0) + \Delta n} - \frac{{}^{d(\Delta n)}/_{dt}}{\Delta n}\right) = r$$

Each of the terms in the left-hand side is a logarithmic derivative. By integrating with respect to time from 0 to t, we get successively:

$$\frac{1}{(n_0+p_0)}\left[\ln\left((n_0+p_0)+\Delta n\right)-\ln\left(\Delta n\right)\right]_0^t = rt$$

$$\frac{1}{(n_0+p_0)}\left[\ln\left(\frac{(n_0+p_0)+\Delta n}{\Delta n}\right)\right]_0^t = rt$$

$$\ln\left(\frac{(n_0+p_0)+\Delta n(t)}{\Delta n(t)}\right)-\ln\left(\frac{(n_0+p_0)+\Delta n(0)}{\Delta n(0)}\right)=r(n_0+p_0)t$$

Taking the exponential on both sides of this last equation, we obtain:

$$\frac{(n_0+p_0)+\Delta n(t)}{\Delta n(t)}=\frac{(n_0+p_0)+\Delta n(0)}{\Delta n(0)}\exp[r(n_0+p_0)t]$$

and solving for $\Delta n(t)$, we get:

$$\Delta p(t)=\Delta n(t)=\frac{(n_0+p_0)\Delta n(0)}{[(n_0+p_0)+\Delta n(0)]\exp[r(n_0+p_0)t]-\Delta n(0)} \tag{8.64}$$

This shows the general form for the change in the excess electron concentration as a function of time. The only parameters of the variation are the equilibrium concentrations n_0 and p_0, the initial excess electron concentration $\Delta n(0)$, and the recombination coefficient $r(T)$. This complicated expression can be drastically simplified in some cases.

For weak excitation levels, i.e., $\Delta n(0) << (n_0 + p_0)$, Eq. (8.64) becomes:

$$\begin{aligned}\Delta n(t) &\approx \frac{(n_0+p_0)\Delta n(0)}{(n_0+p_0)\exp[r(n_0+p_0)t]-\Delta n(0)}\\ &\approx \frac{(n_0+p_0)\Delta n(0)}{(n_0+p_0)\exp[r(n_0+p_0)t]}\end{aligned}$$

or simply:

$$\Delta n(t)\approx \Delta n(0)\exp[-r(n_0+p_0)t] \tag{8.65}$$

and similarly for $\Delta p(t)$:

$$\Delta p(t)\approx \Delta p(0)\exp[-r(n_0+p_0)t] \tag{8.66}$$

By defining a direct band-to-band recombination lifetime for electrons and holes as:

$$\tau_p=\tau_n=\frac{1}{r(n_0+p_0)} \tag{8.67}$$

we obtain:

$$\begin{cases} \Delta n(t) \approx \Delta n(0) e^{-\frac{t}{\tau_n}} \\ \Delta p(t) \approx \Delta p(0) e^{-\frac{t}{\tau_0}} \end{cases} \tag{8.68}$$

This is the same lifetime introduced in Eq. (8.47). Indeed, in the current conditions, we have by using Eqs. (8.59) and (8.68):

$$R - G_t = -\frac{d(\Delta n)}{dt} = \frac{1}{\tau_n} \Delta n(0) e^{-\frac{t}{\tau_n}}$$

or:

$$R - G_t = \frac{\Delta n(t)}{\tau_n} \tag{8.69}$$

which is analogous to Eq. (8.47).

8.6.3 Shockley-Read-Hall Recombination

The previous band-to-band recombination most often occurs in pure semiconductor. When defects or impurities are present in the crystal, which is nearly always the case to some extent, energy levels appear in the bandgap and may participate in the recombination mechanisms. These are called Shockley-Read-Hall (SRH) recombinations, and the energy is not released in the form of a photon but is rather given to the crystal lattice in the form of phonons. Such processes are also sometimes called band-to-impurity recombinations. This is therefore normally a non-radiative recombination step.

In the present model, we consider the steady-state generation and recombination of electrons and holes involving an impurity level, also called recombination center, with an energy E_T in the bandgap, as shown in Fig. 8.10 . Let us assume that electrons and holes are generated at a rate equal to G, which is the excess generation rate of Subsect. 8.6.1.

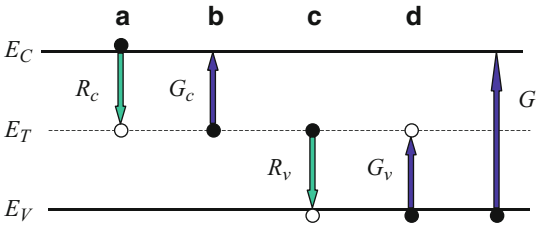


Fig. 8.10 The four possible transitions for an electron involving a recombination center in the bandgap: (**a**) capture of an electron from the conduction band by the center, (**b**) emission of an electron from the center into the conduction band, (**c**) emission of an electron from the center into a vacant state in the valence band, and (**d**) capture of an electron from the valence band by the center

There are four possible electron transitions which can involve this level: (a) the capture of an electron from the conduction band by the center, (b) the emission of an electron from the center into the conduction band, (c) the emission of an electron from the center into a vacant state in the valence band, and (d) the capture of an electron from the valence band by the center. The transition (c) can be equivalently viewed as the capture of a hole by the center and (d) as the emission of a hole from the center into the valence band. Each of these transitions is illustrated in Fig. 8.10.

The recombination of electrons or holes is enhanced by the presence of the impurity level if the probability of transitions (a) and (c) is higher than that of (b) and (d).

If the probability of (a) and (b) is higher than (c) and (d), the impurity level plays more the role of an electron recombination center. If the probability of (c) and (d) is higher than (a) and (b), the impurity level plays more the role of a hole recombination center.

Before analyzing each transition in more detail, let us first assume there is a density N_T of impurity-related states at an energy E_T. At thermal equilibrium, the density of the recombination center states which are occupied by electrons is then given by:

$$N_T f_e(E_T) = \frac{N_T}{\exp\left(\frac{E_T - E_F}{k_b T}\right) + 1} \tag{8.70}$$

where f_e is the Fermi-Dirac distribution given by Eq. (5.28). The density of the recombination center states which are empty of electrons at equilibrium is given by:

$$N_T[1 - f_e(E_T)] = \frac{N_T}{1 + \exp\left(-\frac{E_T - E_F}{k_b T}\right)} \tag{8.71}$$

However, when carriers are transiting through the recombination centers in Fig. 8.10, the density of occupied and empty center states is different from their equilibrium values. We thus introduce a non-equilibrium distribution function f such that the densities of occupied and empty center states are $N_T f$ and $N_T(1 - f)$, respectively. Knowledge of the exact value of this function is not important in analyzing each of the transitions illustrated in Fig. 8.10.

Transition Rates

Let us first discuss the transition (a), i.e., the capture of an electron from the conduction band by the center. The capture rate, or concentration of electrons captured by the center per unit time, is denoted R_c and is expressed in units of cm $^{-3}\cdot s^{-1}$. It must be proportional to the density of electrons in the conduction band n and the density of empty centers $N_T(1 - f)$.

In addition, R_c should also depend on a parameter which describes "how often an electron encounters the recombination center." This parameter is the product $\nu_{th}\sigma_n$ of two quantities: the electron thermal velocity ν_{th} (in units of cm$\cdot s^{-1}$) and the capture cross section σ_n of electrons for this particular recombination center (in units of cm^2).

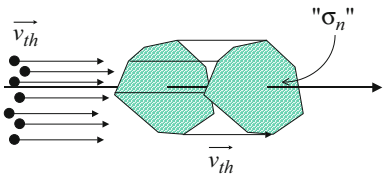


Fig. 8.11 Schematic illustration of the concepts of electron thermal velocity and capture cross section. Using ballistic terminology, the electrons moving with the thermal velocity which would collide with an object having a cross section equal to σ_n are located in the volume delimited by the two shaded surfaces in this figure

These two parameters can be better understood by considering the illustration in Fig. 8.11. It shows that the electrons which have a velocity ν_{th} and which will reach a surface of area σ_n are located in a volume equal to the product $\nu_{th}\sigma_n$ during a unit time.

The electron thermal velocity in a nondegenerate semiconductor is given by:

$$\nu_{th} = \sqrt{\frac{3k_b T}{m}} \tag{8.72}$$

where m is the mass of the electron. The thermal velocity is on the order of 10^7 cm·s^{-1} at room temperature.

The capture cross section of electrons for a recombination center characterizes the interaction between an electron and this center. It corresponds to the effective area around the center that an electron experiences when it is approaching the center. The cross section depends on the type of interaction involved between the electron and the center: the stronger the interaction is, the larger the influence of the capture cross section is. σ_n is usually determined empirically and is on the order of 10^{-15} cm^2. The capture rate R_c in the transition (a) is therefore equal to:

$$R_c = \nu_{th}\sigma_n n N_T (1 - f) \tag{8.73}$$

The emission of an electron from the center into the conduction band, corresponding to transition (b) in Fig. 8.10, is characterized by an emission rate G_c which has the same units as R_c. This quantity is equal to the density of occupied center states $N_T f$ multiplied by the electron emission probability e_n which is a parameter characteristic of the recombination center in the semiconductor:

$$G_c = e_n N_T f \tag{8.74}$$

Because the transitions (c) and (d) are analogous to (a) and (b) but involve holes instead of electrons, we can easily determine the hole capture rate R_v and the hole emission rate G_v from those for electrons Eqs. (8.73) and (8.74).

Indeed, R_v must be proportional to the density of holes in the valence band p, the density of centers which are occupied (by electrons) $N_T f$, the thermal velocity of

holes which is the same as that of electrons given in Eq. (8.72), and the capture cross section of holes σ_p for the center considered:

$$R_v = \nu_{th}\sigma_p p N_T f \tag{8.75}$$

G_v must be equal to the density of center states which are empty (of electrons) $N_T(1-f)$ multiplied by the hole emission probability e_p:

$$G_\nu = e_p N_T(1-f) \tag{8.76}$$

All these expressions for the recombination and emission rates are not independent but must satisfy a number of equations arising from the conservation of electrons and holes. The total number of electrons (or holes) recombined must be equal to the number of electrons (or holes) generated; thus, we can write:

$$\begin{cases} R_c = G_c + G \\ R_v = G_v + G \end{cases} \tag{8.77}$$

Emission Probabilities e_n and e_p

At equilibrium, the excess generation rate G is equal to zero. Moreover, the electron and hole densities are equal n_0 and p_0, respectively, and the distribution function f is equal to $f_e = f_e(E_T)$. All the other parameters remain unchanged. Therefore, by expressing Eq. (8.77) at equilibrium using Eqs. (8.73), (8.74), (8.75), and (8.76), we get:

$$\begin{cases} \nu_{th}\sigma_n n_0 N_T(1-f_e) = e_n N_T f_e \\ \nu_{th}\sigma_p p_0 N_T f_e = e_p N_T(1-f_e) \end{cases}$$

which allow us to extract the electron and hole emission probabilities:

$$\begin{cases} e_n = \nu_{th}\sigma_n n_0 \dfrac{1-f_e}{f_e} \\ e_p = \nu_{th}\sigma_p p_0 \dfrac{f_e}{1-f_e} \end{cases} \tag{8.78}$$

This last set of equations can be simplified by using the expression for the Fermi-Dirac distribution in Eq. (5.28) to obtain:

$$\frac{1-f_e}{f_e} = \exp\left(\frac{E_T - E_F}{k_b T}\right) \tag{8.79}$$

and by using the expressions of n_0 and p_0 given in Eqs. (7.21) and (7.29) for a nondegenerate semiconductor:

$$n_0 \frac{1-f_e}{f_e} = N_c \exp\left(\frac{E_F - E_C}{k_b T}\right) \exp\left(\frac{E_T - E_F}{k_b T}\right)$$
$$= N_c \exp\left(\frac{E_T - E_C}{k_b T}\right)$$

This last quantity can be denoted n_T and would correspond to the electron density in the conduction band if the Fermi energy was equal to the recombination center energy level ($E_F = E_T$):

$$n_T = n_0 \frac{1-f_e}{f_e} = N_c \exp\left(\frac{E_T - E_C}{k_b T}\right) \tag{8.80}$$

A similar expression can be derived for:

$$p_T = p_0 \frac{f_e}{1-f_e} = N_v \exp\left(\frac{E_V - E_T}{k_b T}\right) \tag{8.81}$$

Therefore, Eq. (8.78) is simplified into:

$$\begin{cases} e_n = \nu_{th} \sigma_n n_T \\ e_p = \nu_{th} \sigma_p p_T \end{cases} \tag{8.82}$$

The Non-equilibrium Distribution Function f

The non-equilibrium distribution function, included in the expressions of the transition rates in Eqs. (8.73), (8.74), (8.75), and (8.76), can be determined by eliminating the excess generation rate G in Eq. (8.77). For this, we first calculate the difference between the two expressions in Eq. (8.77):

$$R_c - R_v = G_c - G_v$$

which becomes:

$$\nu_{th} \sigma_n n N_T (1-f) - \nu_{th} \sigma_p p N_T f = e_n N_T f - e_p N_T (1-f)$$

Using Eq. (8.82), we obtain:

$$\nu_{th} \sigma_n n N_T (1-f) - \nu_{th} \sigma_p p N_T f = \nu_{th} \sigma_n n_T N_T f - \nu_{th} \sigma_p p_T N_T (1-f)$$

and, after simplifying by ν_{th} and N_T:

$$\sigma_n n + \sigma_p p_T = f \left\lfloor \sigma_n n + \sigma_p p + \sigma_n n_T + \sigma_p p_T \right\rfloor$$

Thus finally we have:

$$f = \frac{\sigma_n n + \sigma_p p_T}{\sigma_n (n + n_T) + \sigma_p (p + p_T)} \tag{8.83}$$

Recombination Lifetimes

The net recombination rate of electrons from the conduction band is given by the difference between the recombination rate R_c and the generation rate G_c, i.e.:

$$-\frac{d(\Delta n)}{dt} = R_c - G_c \tag{8.84}$$

This quantity is also equal to the net recombination rate of holes from the valence band in view of Eq. (8.77):

$$-\frac{d(\Delta p)}{dt} = R_v - G_v \tag{8.85}$$

Using the non-equilibrium distribution function (Eq. 8.83) and the expressions for R_c, G_c, and e_n in Eqs. (8.73), (8.74) and (8.82), we can calculate successively:

$$\begin{aligned} R_c - G_c &= v_{th}\sigma_n n N_T(1-f) - e_N N_T f \\ &= v_{th}\sigma_n N_T[n-(n+n_T)f] \\ &= v_{th}\sigma_n N_T\left[n-(n+n_T)\frac{\sigma_n n+\sigma_p p_T}{\sigma_n(n+n_T)+\sigma_p(p+p_T)}\right] \\ &= \frac{v_{th}\sigma_n N_T}{\sigma_n(n+n_T)+\sigma_p(p+p_T)}\left[n\sigma_p(p+p_T)-(n+n_T)\sigma_p p_T\right] \\ &= \frac{v_{th}\sigma_n N_T}{\sigma_n(n+n_T)+\sigma_p(p+p_T)}\sigma_p[np-n_T p_T] \end{aligned}$$

From the definitions of n_T and p_T in Eqs. (8.80) and (8.81), we have $n_T p_T = n_i^2$ where n_i is the intrinsic carrier concentration. The previous equation can then be simplified into:

$$R_c - G_c = v_{th}\sigma_n\sigma_p N_T \frac{(np-n_i^2)}{\sigma_n(n+n_T)+\sigma_p(p+p_T)} \tag{8.86}$$

Introducing the excess carriers Δn and Δp as in Eq. (8.57), and still assuming $\Delta n = \Delta p$, we get:

$$R_c - G_c = v_{th}\sigma_n\sigma_p N_T \frac{(n_0+p_0+\Delta n)\Delta n}{\sigma_n(n_0+n_T+\Delta n)+\sigma_p(p_0+p_T+\Delta n)} \tag{8.87}$$

Here we have also used the relation $n_0 p_0 = n_i^2$. This expression can be further simplified by first considering two particular cases.

(i) For low excess carrier concentrations, i.e., weak excitation levels where $\Delta n << n_0$, p_0, and for an n-type semiconductor, where we can assume that n_0 is much higher than p_0, n_T, and p_T, Eq. (8.87) becomes:

$$R_c - G_c \approx v_{th}\sigma_n\sigma_p N_T \frac{(n_0)\Delta n}{\sigma_n(n_0)}$$

which can be rewritten, by taking into account Eq. (8.84):

$$-\frac{d(\Delta n)}{dt} = R_c - G_c \approx v_{th}\sigma_p N_T \Delta n \tag{8.88}$$

From this last expression, we can introduce a recombination lifetime τ_{p_0} such that:

$$-\frac{d(\Delta n)}{dt} \approx \frac{\Delta n}{\tau_{p_0}}$$

i.e.:

$$\tau_{p_0} = \frac{1}{v_{th}\sigma_p N_T} \tag{8.89}$$

Note that the subscript "p" has been used for this lifetime, because it depends on the capture cross section of holes. This corresponds to a lifetime of holes. Therefore, in an *n*-type semiconductor, the excess carrier lifetime approaches that of holes.

(ii) In the second case, still $\Delta n << n_0, p_0$; but for a *p*-type semiconductor this time, where we can assume that p_0 is much higher than n_0, n_T, and p_T, Eq. (8.87) becomes:

$$R_c - G_c \approx v_{th}\sigma_n N_T \Delta n$$

Here again, we can rewrite this as:

$$-\frac{d(\Delta n)}{dt} \approx \frac{\Delta n}{\tau_{n_0}}$$

with:

$$\tau_{n_0} = \frac{1}{v_{th}\sigma_n N_T} \tag{8.90}$$

Here, the suffix "*n*" has been used, because the lifetime depends on the capture cross section of electrons. This corresponds to a lifetime of electrons. Therefore, in a *p*-type semiconductor, the excess carrier lifetime approaches that of electrons. Using the expressions in Eqs. (8.89) and (8.90), we can simplify Eq. (8.87):

$$R_c - G_c = \frac{(n_0 + p_0 + \Delta n)\Delta n}{\tau_{p_0}(n_0 + n_T + \Delta n) + \tau_{n_0}(p_0 + p_T + \Delta n)} \tag{8.91}$$

From Eqs. (8.84) and (8.85), we can write:

$$-\frac{d(\Delta n)}{dt} = -\frac{d(\Delta p)}{dt} = \frac{(n_0 + p_0 + \Delta n)\Delta n}{\tau_{p_0}(n_0 + n_T + \Delta n) + \tau_{n_0}(p_0 + p_T + \Delta n)} \tag{8.92}$$

We can now introduce the Shockley-Read-Hall recombination lifetime $\tau_n = \tau_p$ such that:

$$-\frac{d(\Delta n)}{dt} = -\frac{d(\Delta p)}{dt} = \frac{\Delta p}{\tau_p} = \frac{\Delta n}{\tau_n}$$

i.e.:

$$\tau_n(t) = \tau_p(t) = \frac{\tau_{p_0}(n_0 + n_T + \Delta n) + \tau_{n_0}(p_0 + p_T + \Delta n)}{(n_0 + p_0 + \Delta n)} \tag{8.93}$$

which becomes independent of time for weak excitation levels $\Delta n << n_0, p_0$:

$$\tau_n = \tau_p = \frac{\tau_{p_0}(n_0 + n_T) + \tau_{n_0}(p_0 + p_T)}{(n_0 + p_0)} \tag{8.94}$$

From this relation, we can easily find the two previous particular cases, i.e., for an n-type semiconductor, $\tau_n = \tau_p = \tau_{p_0}$, and for a p-type semiconductor, $\tau_n = \tau_p = \tau_{n_0}$.

8.6.4 Auger Band-to-Band Recombination

Unlike the direct band-to-band or the SRH processes, in the Auger band-to-band, or simply Auger recombination, the energy that is released when an electron recombines with a hole is transferred to a third particle, an electron in the conduction band or a hole in the valence band. This carrier particle is called an Auger electron or Auger hole. The energy that this third particle acquires is subsequently released in the form of heat or phonons into the lattice. Auger recombination is an intrinsic non-radiative mechanism which is more effective at higher temperatures and for smaller bandgap semiconductors. This recombination mechanism occurs most often in doped direct bandgap semiconductors.

There are three possible Auger recombination mechanisms, depending on what type of Auger carrier is excited and where it is excited. These are illustrated in Fig. 8.12.

The first process, shown in Fig. 8.12a, is called a CHCC process to indicate that an electron from the conduction band (C) recombines with a hole in the valence band (H) to lead to the excitation of another electron which remains in the conduction

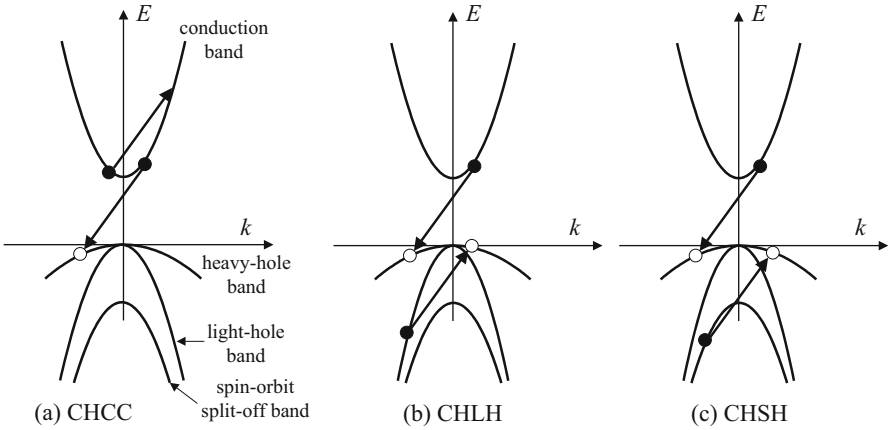


Fig. 8.12 Auger recombination process semiconductors. The energy released through the recombination of an electron in the conduction band and a hole in the valence band is yielded to: (**a**) another electron in the conduction band which is then excited to a higher state in the band, (**b**) an electron in the LH band which is excited to a vacant electronic state in the HH band, (**c**) an electron in the split-off band which is excited to a vacant electronic state in the HH band

band (CC). In the case of an Auger hole, the valence band structure is more complex than the conduction band, as we saw in Subsect. 5.4.3. We must then distinguish whether this hole is excited into the light-hole band (CHLH process, Fig. 8.12b) or the spin-orbit split-off band (CHSH process, Fig. 8.12c).

In all three cases, the total energy and the total momentum (i.e., $\hbar\,\vec{k}$) of the system constituted by the three particles must be conserved.

Similar to the direct band-to-band recombination, the Auger recombination rates are expressed in units of $\mathrm{cm}^{-3}\cdot\mathrm{s}^{-1}$ and are proportional, in all three processes, to the density of electrons in the conduction band n and that of holes in the valence band p, because these are the particles which are recombining.

In the CHCC case, this rate is also proportional to the density of electrons which are susceptible to be excited, i.e., n again. The recombination rate in the CHCC process is therefore given by:

$$R_{\mathrm{CHCC}} = r_1 n^2 p \tag{8.95}$$

where r_1 is the Auger recombination coefficient for this case and is expressed in units of cm^{-1}.

For the CHLH and CHSH processes, the same argument leads to a compounded recombination rate equal to:

$$R_{\mathrm{CHLH+CHSH}} = r_2 n p^2 \tag{8.96}$$

where r_2 is the Auger recombination coefficient when Auger holes are excited.

The total Auger recombination rate is therefore:

$$R = R_{\mathrm{CHCC}} + R_{\mathrm{CHLH+CHSH}} = r_1 n^2 p + r_2 n p^2 \tag{8.97}$$

We can now follow the same analysis as the one conducted for the direct band-to-band recombination in order to determine the Auger recombination lifetime. We start from the rate Eq. (8.59). At equilibrium, $\frac{dn}{dt} = 0$ and $G = 0$, and the thermal generation rate is thus equal to:

$$G_{\mathrm{t}} = R = r_1 n_0^2 p_0 + r_2 n_0 p_0^2 \tag{8.98}$$

Let us now consider the relaxation process, which occurs after the external source of generation is removed ($G = 0$). Taking into account Eqs. (8.97) and (8.98), Eq. (8.59) becomes:

$$-\frac{d(\Delta n)}{dt} = R - G_{\mathrm{t}} = r_1\left(n^2 p - n_0^2 p_0\right) + r_2\left(n p^2 - n_0 {p_0}^2\right) \tag{8.99}$$

where $\Delta n = \Delta p$ is the excess electron and hole concentrations defined in Eq. (8.61). This expression can be expanded using Eq. (8.61), and we obtain:

$$\begin{aligned} -\frac{d(\Delta n)}{dt} &= -r_1\left[n_0^2 p_0 - (n_0 + \Delta n)^2 (p_0 + \Delta n)\right] - r_2\left[n_0 p_0^2 - (n_0 + \Delta n)(p_0 + \Delta n)^2\right] \\ &= r_1\left[\left(n_0^2 + 2n_0 p_0\right)\Delta n + (2n_0 + p_0)(\Delta n)^2 + (\Delta n)^3\right] \\ &\quad + r_2\left[\left(p_0^2 + 2n_0 p_0\right)\Delta n + (2p_0 + n_0)(\Delta n)^2 + (\Delta n)^3\right] \end{aligned}$$

We can now introduce the Auger recombination lifetime $\tau_{\mathrm{n}} = \tau_{\mathrm{p}}$ such that:

$$-\frac{d(\Delta n)}{dt} = -\frac{d(\Delta p)}{dt} = \frac{\Delta p}{\tau_{\mathrm{p}}} = \frac{\Delta n}{\tau_{\mathrm{n}}}$$

$$\tau_{\mathrm{n}}(t) = \tau_{\mathrm{p}}(t) = 1 \begin{array}{l} r_1\left[\left(n_0^2 + 2n_0 p_0\right) + (2n_0 + p_0)\Delta n + (\Delta n)^2\right] \\ + r_2\left[\left(p_0^2 + 2n_0 p_0\right) + (2p_0 + n_0)\Delta n + (\Delta n)^2\right] \end{array} \tag{8.100}$$

which becomes independent of time for weak excitation levels $\Delta n << n_0, p_0$:

$$\tau_{\mathrm{n}} = \tau_{\mathrm{p}} = \frac{1}{r_1\left(n_0^2 + 2n_0 p_0\right) + r_2\left(p_0^2 + 2n_0 p_0\right)} \tag{8.101}$$

8.6.5 Surface Recombination

The surface of a semiconductor is a violation of the crystal periodicity and therefore gives rise to energy levels near the surface which lie within the bandgap. These

correspond to surface traps. However, unlike the previously discussed carrier recombination mechanisms which occur in the bulk solid, surface recombination occurs at the surface of the solid. Moreover, the surface recombination takes place even in pure materials. Such processes play an important role in semiconductor device technology.

The energy levels introduced by the surface traps can be considered as a special case of recombination centers in Shockley-Read-Hall recombination mechanism. The same analysis as in Subsect. 8.6.4 can be conducted here for surface recombination, provided a surface density of recombination centers $(N_T)_s$ is used instead of the bulk density of centers N_T. All the other parameters would keep the same meaning.

The excess surface recombination rate is the number of electrons or holes which are recombined per unit area of the surface and per unit time. It is thus expressed in units of $cm^{-2} \cdot s^{-1}$ and can be obtained by analogy with the SRH recombination in Eq. (8.87):

$$(R - G_t)_s = v_{th}\sigma_n\sigma_p(N_T)_s \frac{(n_0 + p_0 + \Delta n)\Delta n}{\sigma_n(n_0 + n_T + \Delta n) + \sigma_p(p_0 + p_T + \Delta n)} \tag{8.102}$$

Here, Δn is the excess electron concentration near the surface considered. We can rewrite this relation as:

$$-\frac{d(\Delta n)}{dt} = (R - G_t)_s = S_n \Delta n \tag{8.103}$$

where:

$$S_n = v_{th}\sigma_n\sigma_p(N_T)_s \frac{(n_0 + p_0 + \Delta n)}{\sigma_n(n_0 + n_T + \Delta n) + \sigma_p(p_0 + p_T + \Delta n)} \tag{8.104}$$

This quantity is expressed in units of $cm \cdot s^{-1}$ and has thus the same dimension as a velocity. It is called the surface recombination velocity.

8.7 Quasi-Fermi Energy

In Sect. 7.5, we calculated the equilibrium electron concentration in the conduction band n_0 and the hole concentration in the valence band p_0 using the Fermi-Dirac distribution and arrived at Eqs. (7.18) and (7.27) in the general case and Eqs. (7.21) and (7.29) in the nondegenerate. For a given semiconductor material, these concentrations depended solely on a single parameter, the Fermi energy E_F.

Under non-equilibrium conditions, where the electron and hole concentrations in their respective bands are given by:

$$\begin{cases} n = n_0 + \Delta n \\ p = p_0 + \Delta p \end{cases} \tag{8.105}$$

the Fermi-Dirac distribution is not valid any more. However, it is convenient to maintain the mathematical formalism of the equations mentioned previously, and this is most often done for a nondegenerate semiconductor only.

Therefore, by analogy with Eq. (7.21), the non-equilibrium electron concentration in the conduction band is given by:

$$n = N_{\mathrm{c}} \exp\left(\frac{E_{\mathrm{F_n}} - E_{\mathrm{C}}}{k_{\mathrm{b}} T}\right) \tag{8.106}$$

where the quantity $E_{F_{\mathrm{n}}}$ is used instead of the Fermi energy E_{F}. This quantity is called the electron quasi-Fermi energy. Using this expression, Eqs. (7.21) and (8.105), we can write:

$$\frac{\Delta n}{n_0} = \frac{n}{n_0} - 1 = \exp\left(\frac{E_{\mathrm{F_n}} - E_{\mathrm{F}}}{k_{\mathrm{b}} T}\right) - 1 \tag{8.107}$$

Therefore, under non-equilibrium conditions, the difference between the quasi-Fermi level and the Fermi level determines the relative excess electron concentration with respect to the equilibrium concentrations.

Using this quasi-Fermi energy, it is possible to define a quasi-Fermi-Dirac distribution for electrons, which is analogous to Eq. (5.28) with E_{F} replaced by $E_{\mathrm{F_n}}$:

$$f_{e_{\mathrm{n}}}(E) = \frac{1}{\exp\left(\frac{E - E_{\mathrm{F_n}}}{k_{\mathrm{b}} T}\right) + 1} \tag{8.108}$$

A similar concept can be introduced for holes in the valence band. The hole quasi-Fermi energy $E_{F_{\mathrm{p}}}$ is defined such that:

$$p = N_{\mathrm{v}} \exp\left(\frac{E_{\mathrm{V}} - E_{\mathrm{F_p}}}{k_{\mathrm{b}} T}\right) \tag{8.109}$$

A quasi-Fermi-Dirac distribution for holes can also be defined by analogy with Eq. (7.23):

$$f_{h_{\mathrm{p}}}(E) = \frac{1}{\exp\left(\frac{E_{\mathrm{F_p}} - E}{k_{\mathrm{b}} T}\right) + 1} \tag{8.110}$$

The quasi-Fermi-Dirac distributions allow separate mathematical computations for electrons and holes in an easier manner. At equilibrium, the electron and hole quasi-Fermi energies are both equal to the Fermi energy, i.e., $E_{\mathrm{F_n}} = E_{\mathrm{F_p}} = E_{\mathrm{F}}$.

Example

Q Estimate the difference between the quasi-Fermi energies E_{Fn} and E_{Fp} and the Fermi energy E_F in an intrinsic semiconductor, given that the excess carrier concentration $\Delta n = \Delta p$ is 1% of n_0.

A The quasi-Fermi energies E_{Fn} and E_{Fp} are related to the excess carrier concentration through the expression $E_{F_n} - E_F = k_b T \ln\left(\frac{\Delta n}{n_0}\right)$ and $E_F - E_{F_p} = k_b T \ln\left(\frac{\Delta p}{p_0}\right)$, where n_0 and p_0 are the equilibrium electron and hole concentrations and are both equal to the intrinsic carrier concentration n_i since the semiconductor is assumed intrinsic at equilibrium. Therefore $\frac{\Delta n}{n_0} = \frac{\Delta p}{p_0} = 0.01$ and we obtain: $E_{F_n} - E_F = E_F - E_{F_p} = 0.0095 k_b T$.

8.8 Transport Theory: Beyond Drude

In this chapter we derived the electrical conductivity of materials using a very simple classical Newton's laws approach. We did this because the so-called Drude theory of conductivity is surprisingly powerful and useful. But it does not include the Pauli principle, for example, and does not incorporate the concept of the Fermi distribution and the Fermi level. There are many situations in which the simple Drude theory is not adequate. So we will here derive a more rigorous transport theory based on the work of Ludwig Eduard Boltzmann, and we will show how it differs from Drude, and in what special limits it reduces to the Drude theory.

8.8.1 The Boltzmann Equation

We must start with the concept of the distribution function of electrons $f_{\vec{k}}(\vec{r}, t)$. This quantity is the probability of an electron occupying the Bloch state $\boldsymbol{k}$ in the solid at position r at time t. Note that this is not the equilibrium Fermi distribution function $f^{\,0}_{\vec{k}}$ from Eq. (5.28) which only depends on the energy and Fermi energy. The new non-equilibrium distribution $f_{\vec{k}}(\vec{r}, t)$ tells us how many particles there are in this region of space at time t and with momentum k. In a steady-state situation, the total rate of change with time of the distribution function must be zero. Specifically, there are changes in the function $f_{\vec{k}}(\vec{r}, t)$ caused by specific processes. Thus the distribution changes because:

1. The particles in the material are diffusing in space.
2. Electric and magnetic fields are accelerating the particles.
3. There are scattering processes which change the momentum and energies of the particles. These processes include scattering from impurities, defects, phonons, etc., all processes which break the Bloch symmetry of the crystal.

The information we seek is in $f_{\vec{k}}(\vec{r},t)$. Knowing this function we can compute the current via

$$\vec{J} = -\int d\vec{k}\, q v_{\vec{k}} f(\vec{k},\vec{r},t) \tag{8.111}$$

where $v_{\vec{k}}$ is the velocity. To calculate the distribution function, we now examine each of the above processes in turn. First we note that because the particles diffuse in space, one source of time variation is "diffusion" which is described by the variation:

$$\left.\frac{\partial f}{\partial t}\right|_{\text{diffusion}} = -\vec{v}_{\vec{k}}.\nabla f_{\vec{k}} \rightarrow \frac{\partial f}{\partial t} = \frac{\partial \vec{r}}{\partial t}.\frac{\partial f_{\vec{k}}}{\partial \vec{r}} \tag{8.112}$$

Then there is the influence of external fields. To proceed we remember from Eqs. (8.1) and (8.14) that with Bloch states, the Newton laws act on the pseudo momentum parameter k as:

$$\frac{d\vec{k}}{dt} = -\frac{q}{\hbar}\left(\vec{E}_{\text{a}} + \vec{v}_{\vec{k}} \times \vec{B}\right) \rightarrow \text{MKS} \tag{8.113}$$

where E_a is the applied field. Thus it follows that the field variation is:

$$\left.\frac{\partial f}{\partial t}\right|_{\text{field}} = -\frac{d\vec{k}}{dt}.\frac{\partial f(\vec{k},\vec{r},t)}{\partial \vec{k}} = \frac{q}{\hbar}\left(\vec{E}_{\text{a}} + \vec{v} \times \vec{B}\right).\frac{\partial f(\vec{k},\vec{r},t)}{\partial \vec{k}} \tag{8.114}$$

$$v_{\vec{k}} = \frac{1}{\hbar}\vec{\nabla} E_{\vec{k}} \tag{8.115}$$

Finally the change due to collisions can be written in terms of a generalized rate equation:

$$\left.\frac{\partial f}{\partial t}\right|_{\text{collisons}} = \int \left\{f_{\vec{k}}\left(1 - f_{\vec{k}'}\right) - f_{\vec{k}'}\left(1 - f_{\vec{k}}\right)\right\} W\left(\vec{k},\vec{k}'\right) d\vec{k}' \tag{8.116}$$

where $W\left(\vec{k},\vec{k}'\right)$ is the rate at which electrons are scattered from k to k'.

In the steady state, the sum of all variations must add up to zero:

$$\left.\frac{\partial f_{\vec{k}}}{\partial t}\right|_{\text{field}} + \left.\frac{\partial f_{\vec{k}}}{\partial t}\right|_{\text{diffusion}} + \left.\frac{\partial f_{\vec{k}}}{\partial t}\right|_{\text{collisions}} = \frac{df_{\vec{k}}}{dt} = 0 \tag{8.117}$$

So we have:

$$\begin{aligned}\frac{\partial f}{\partial t} &= \int \left\{f_{\vec{k}}\left(1-f_{\vec{k'}}\right) - f_{\vec{k'}}\left(1-f_{\vec{k}}\right)\right\} W\left(\vec{k},\vec{k'}\right) d\,\vec{k'} \\ &+\frac{q}{\hbar}\left(\vec{E}_{\mathrm{a}} + \vec{v}\times\vec{B}\right).\frac{\partial f\left(\vec{k},\vec{r},t\right)}{\partial\,\vec{k}} - \vec{v}_{\vec{k}}.\,\vec{\nabla} f_{\vec{k}} = 0\end{aligned} \tag{8.118}$$

In principle if one knows the scattering rates, then one can compute the result by following the trajectory of the particles in space and time. The Boltzmann equation can also be solved numerically using the Monte Carlo Method. Let us consider the simple first-order solution in an electric field using the relaxation time approximation. So one says the following: the field produces a small change in the distribution function which we call:

$$f = f_0 + f_1\left(\vec{k},\vec{r},t\right) \tag{8.119}$$

This deviation from the steady state must return to zero when the system has had time to relax or reach a steady state, so we can write:

$$-\frac{\partial f_k}{\partial t} = \frac{1}{\tau} f_1\left(\vec{k},\vec{r},t\right) \tag{8.120}$$

$$f_1 = f_1\left(\vec{k},\vec{r},0\right)e^{-t/\tau} \tag{8.121}$$

It is also useful to write:

$$\frac{\partial f}{\partial\,\vec{k}} = \frac{\partial E_{\vec{k}}}{\partial\,\vec{k}}\frac{\partial f}{\partial E_{\vec{k}}} \tag{8.122}$$

Substituting back into the Boltzmann equation and including only the electric field term, we have:

$$-q\frac{df}{dE_{\vec{k}}}\vec{v}_{\vec{k}}.\vec{E}_{\mathrm{a}} = -\frac{f_1}{\tau} \tag{8.123}$$

$$f_1 = q\frac{df}{dE_{\vec{k}}}\left(\vec{v}_{\vec{k}}.\vec{E}_{\mathrm{a}}\right)\tau \tag{8.124}$$

This actually already gives us the first-order solution in an electric field.

Note that in Drude theory we say that in the steady state, the field force balances the frictional force. Here it is the population which reaches a steady state, not the individual particles. Another way of looking at it is to go back to Eq. (8.119) and allow the electron momentum to be increased by the field up to a relaxation time τ, after which it is interrupted and has to start accelerating again. Thus in the steady state, $\delta k_x = -q\tau E_{ax}/\hbar$ which gives a concomitant change in the energy to $E(k_y, k_z, k_x - qE_{ax}\tau/\hbar)$. The change in the energy gives rise to a steady-state change in the distribution function f as in Eq. (8.119), E_a is the applied field, and:

$$f_{\vec{k}} = f^{0}_{\vec{k}} + \frac{\partial f^{0}_{k}(E_{\vec{k}})}{\partial E_{\vec{k}}} \frac{\partial E(\vec{k})}{\partial \vec{k}} \cdot \frac{q\tau}{\hbar} \vec{E}_{a} \tag{8.125}$$

Substituting Eq. (8.125) back into the expression for the current, we have:

$$\vec{J} = -\frac{1}{4\pi^3} \int d\vec{k}\, q v_{\vec{k}} f(\vec{k}, \vec{r}, t) = -\frac{1}{4\pi^3} \int d\vec{k}\, q v_{\vec{k}} f_1(\vec{k}, \vec{r}, t) \tag{8.126}$$

$$\vec{J} = \frac{1}{4\pi^3} q^2 \int d\vec{k}\, \tau(E_{\vec{k}}) \vec{v}_{\vec{k}} (\vec{v}_{\vec{k}} . \vec{E}_{a}) \left(-\frac{\partial f^{0}_{\vec{k}}}{\partial E_{\vec{k}}} \right) \tag{8.127}$$

Write the volume integral as an integral over energy and surface of constant energy:

$$\begin{aligned} d\vec{k} &= k^2 dk \sin(\vartheta) d\theta d\phi \\ k^2 &= 2mE/\hbar^2 \\ d\vec{k} &= m^{3/2} E^{1/2} dE \sin(\vartheta) d\theta d\phi = dE dS_F / \hbar v_{\vec{k}} \end{aligned} \tag{8.128}$$

So that:

$$J = \frac{1}{4\pi^2} \int d\vec{k}\, q^2 \tau(E_{\vec{k}}) \vec{v}_{\vec{k}} (\vec{v}_{\vec{k}} \vec{E}_{a}) \left(-\frac{\partial f^{0}_{\vec{k}}}{\partial E_{\vec{k}}} \right) \tag{8.129}$$

$$J = 2q^2 \int dE_k g_V(E_k) \tau(E_k) \vec{v}_{\vec{k}} (\vec{v}_{\vec{k}} \vec{E}_{a}) \left(-\frac{\partial f^{0}_{\vec{k}}}{\partial E_k} \right) \tag{8.130}$$

$$E_{KE} = 1/2\, mv^2 \tag{8.131}$$

$$v_{\vec{k}} v_{\vec{k}} \vec{E}_{a} \rightarrow v_x^2 E_{a,x} = \frac{2E}{3m} E_{a,x} \tag{8.132}$$

where in order to avoid confusion, we use E_a for applied field and we use g_V to denote the density of states per unit volume. With the low-temperature form:

$$-\frac{\partial f}{\partial E} = \delta(E - E_F) \tag{8.133}$$

The current density reduces to the form (where $\bar{d}$ is the dimensionality of the system):

$$J = \frac{q^2\tau(E_F)}{\bar{d}} v_F^2 g_V(E_F) E_{a,x} \tag{8.134}$$

8.8.2 Connection to Drude Theory

In order to relate this expression to the familiar Drude result from Eq. (8.7), we consider three dimensions and also assume that the relaxation time is energy independent and that we are dealing with nearly free electrons:

$$J = \frac{q^2\tau}{3} v_F^2 g(E_F) E_{a,x} \sim \frac{1}{3} m^* v_F^2 g(E_F) \frac{q^2\tau}{m^*} \tag{8.135}$$

$$J = n\frac{q^2\tau}{m^*} E_{a,x} \rightarrow n = \frac{1}{3} m^* v_F^2 g(E_F) \tag{8.136}$$

$$\sigma = \frac{q^2\tau(E_F)}{3} v_F^2 g(E_F) \rightarrow n\frac{q^2\tau}{m^*} \tag{8.137}$$

Here n is the effective carrier density.

The connection is made and we see why Drude represents a serious approximation:

(i) The more correct i.e, Boltzmann form of the conductivity scales with the density of states at the Fermi level. So that if there is no free charge that responds to an applied field, there is no conduction.
(ii) The Boltzmann equation includes the Fermi distribution, so the Pauli principle is obeyed.
(iii) The Boltzmann equation result allows the relaxation time to be energy dependent. This energy dependence enters the result via the integral Eq. (8.130) which then also takes care of the temperature dependence.
(iv) The Boltzmann equation result allows the group velocity to deviate from the nearly free electron law.

In general we see therefore that the Boltzmann equation conductivity is far superior to the Drude theory and is really the right way to proceed in a Bloch solid.

8.9 Summary

In this chapter, we have covered a few important non-equilibrium transport phenomena involving charge carriers. Firstly, we discussed the electrical conductivity (Ohm's law) in the presence of an external electric field. There, we introduced the concepts of conductivity, resistivity, as well as carrier collision or scattering. Then, secondly we described the Hall effect for an *n*-type and then a *p*-type semiconductor in the presence of perpendicular electric and magnetic fields. There, we introduced the notion of carrier mobility. Thirdly, we discussed the diffusion of charge carriers in an inhomogeneous semiconductor, leading to the concepts of diffusion length and the Einstein relations.

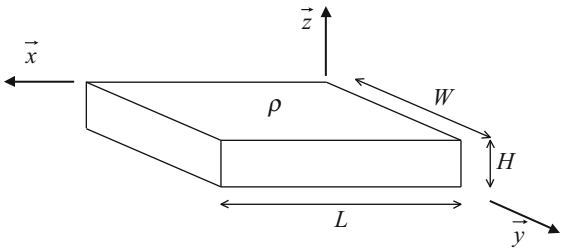
The recombination mechanisms of charge carriers in a semiconductor have been described, including the direct band-to-band, Shockley-Read-Hall, Auger, and surface recombination processes. The concepts of recombination lifetime and capture cross section were introduced.

We introduced the notion of quasi-Fermi energy to describe the electron and hole distribution under non-equilibrium conditions while at the same time maintaining the same mathematical formalism as under equilibrium conditions.

In the last part of the chapter, we introduced the reader to a more powerful transport description known as the Boltzmann equation approach. We derived a more general formula for the conductivity, and we showed why it is superior to the Drude method.

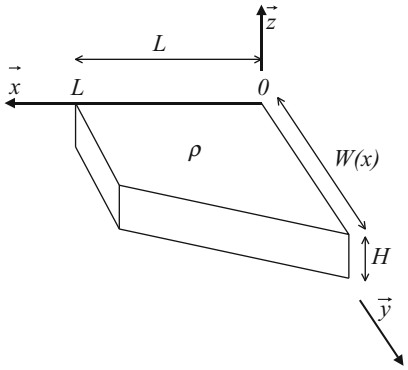
Problems

1. Consider the semiconductor slab shown in the figure below with dimensions $L = 1$ cm, $W = 0.2$ cm, and $H = 0.25$ cm and with a resistivity of 0.01 Ω·cm. What would be the resistance one would measure across opposite faces in all three directions (x, y, and z)? Knowing there is a uniform concentration $n = 10^{16}$ cm^{-3} of electrons in this semiconductor (and no holes), calculate the mobility of these electrons.



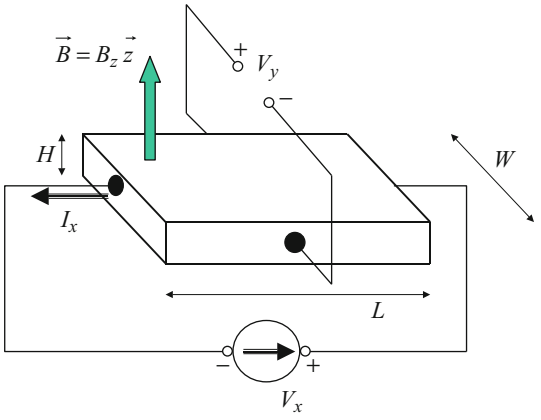
2. Consider the semiconductor block with a resistivity of 0.01 Ω·cm as shown in the figure below. The width of this block is constant but follows the relation $W = 1 + 2(L - x)$ cm when x is varied from 0 to L. The other dimensions are

$L = 1$ cm and $H = 0.25$ cm. Calculate the resistance in the x-direction. For this, you may consider the semiconductor block as a series of parallelepiped slabs next to one another.

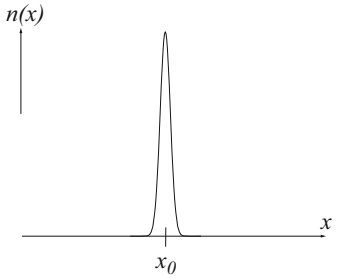


3. Do the same as in Problem 2, but in the y-direction.
4. Consider the Hall effect measurement experiment depicted in the figure below. The dimensions of the semiconductor slab are $L = 2$ mm, $W = 1$ mm, and $H = 2$ µm. Assume the current $I_x = 10$ mA, the voltages $V_x = 10$ V and $V_y = -4$ V, and a magnetic induction $B_z = 0.05$ T.

   Determine if the semiconductor is n-type or p-type, the Hall constant, the carrier concentration, the Hall mobility, the conductivity, and the resistivity of the semiconductor (assumed uniform).

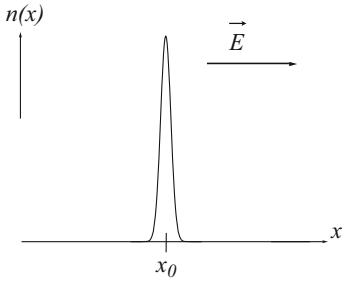


5. Consider an experiment where excess electrons are generated in a "burst" at $t = 0$ and $x = x_0$ in a semiconductor, resulting in the concentration profile $n(x)$ shown in the figure below.



Draw the shape of the concentration profile $n(x)$ as a result of the one-dimensional diffusion in the x-direction. No other external forces are present. Draw several shapes corresponding to several times after the initial "burst."

6. Do the same as in Problem 5, but consider, in addition, that there is an electric field strength $\vec{E}$ in the direction as shown in the figure below.



7. The electron mobility in a Ge crystal is experimentally found to be proportional to $T^{-1.66}$ (i.e., the mobility decreases with increasing temperature). Knowing that this mobility is 4000 cm^2/Vs at 300 K, determine the electron diffusion coefficient at 300 K and 77 K. Compare.
8. Consider an n-type Si semiconductor at room temperature with an excess electron concentration which decreases from 4×10^{16} cm^{-3} to 1 cm^{-3} (practically zero) over a distance of 1 mm. Determine the diffusion length of these electrons.
9. Assume a one-dimensional model in which holes are generated at a rate of $G(x,t)$. Let τ_p be the recombination lifetime for holes and p_0 be the equilibrium hole concentration. Give an expression for $\frac{\partial p(x,t)}{\partial t}$, i.e., the rate of change for the hole concentration at position x, as a function of the diffusion current $J_h^{diff}(x,t)$ and the parameters defined previously. This relation is called a continuity equation and states that the total number of holes must be accounted for. Using Eq. (8.42), rewrite this relation such that it involves the hole concentration $p(x,t)$ as the only unknown.

Further Reading

Anselm A (1981) Introduction to semiconductor theory. Prentice-Hall, Englewood Cliffs, NJ
Ashcroft NW, Mermin ND (1976) *Solid State Physics*, Holt. Rinehart and Winston, New York
Cohen MM (1972) Introduction to the quantum theory of semiconductors. Gordon and Breach, New York
Esaki L (ed) (1991) Highlights in condensed matter physics and future concepts, vol 285. Plenum Press, New YorkNATO Science Forum Series
Ferry DK (1991) Semiconductors. Macmillan, New York
Hummel RE (1986) Electronic properties of materials. Springer-Verlag, New York
Orton JW, Blood P (1990) The electrical characterization of semiconductors: measurement of minority carrier properties. Academic Press, San Diego
Pankove JI (1975) Optical processes in semiconductors. Dover, New York
Peyghambarian N, Koch SW, Mysyrowicz A (1993) Introduction to semiconductor optics. Prentice-Hall, Englewood Cliffs, NJ
Pierret RF (1989) Advanced semiconductor fundamentals. Addison-Wesley, Reading, MA
Pollock CR (1995) Fundamentals of optoelectronics. Irwin, Burr Ridge, IL
Ridley BK (1999) Quantum processes in semiconductors. Oxford University Press, New York
Rogalski A (1995) Infrared photon detectors, Bellingham, Washington
Streetman BG (1990) Solid state electronic devices. Prentice-Hall, Englewood Cliffs, NJ
Sze SM (1981) Physics of semiconductor devices. John Wiley & Sons, New York
Wang S (1989) Fundamentals of semiconductor theory and device physics. Prentice-Hall, Englewood Cliffs, NJ
Wolfe CM, Holonyak N Jr, Stillman GE (1989) Physical properties of semiconductors. Prentice-Hall, Englewood Cliffs, NJ