Chapter 1

ENERGY BAND THEORY

1.1. Electron in a crystal

This Section describes the behavior of an electron in a crystal. It will be demonstrated that the electron can have only discrete values of energy, and the concept of "energy bands" will be introduced. This concept is a key element for the understanding of the electrical properties of semiconductors.

1.1.1. Two examples of electron behavior

An electron behaves differently whether it is in a vacuum, in an atom, or in a crystal. In order to comprehend the dynamics of the electron in a semiconductor crystal, it is worthwhile to first understand how an electron behaves in a simpler environment. We will, therefore, study the "classical" cases of the electron in a vacuum (free electron) and the electron confined in a box-like potential well (particle-in-a-box).

1.1.1.1. Free electron

The free electron model can be applied to an electron which does not interact with its environment. In other words, the electron is not submitted to the attraction of the atoms in a crystal; it travels in a medium where the potential is constant. Such an electron is called a free electron. For a one-dimensional crystal, which is the simplest possible structure imaginable, the time-independent Schrödinger equation can be written for a constant potential V using Relationship A3.12 from Annex 3. Since the reference for potential is arbitrary the potential can be set equal to zero ($V = 0$) without losing generality.[1] The time-independent Schrödinger equation can, therefore, be written as:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\Psi(x) = E\,\Psi(x) \qquad (1.1.1)$$

where E is the electron energy, and m is its mass. The solution to Equation 1.1.1 is :

$$\Psi(x) = C_1\,exp(jkx) + C_2\,exp(-jkx) \qquad (1.1.2)$$

where:

$$k = \sqrt{\frac{2mE}{\hbar^2}} \quad \text{or} \quad E = \frac{\hbar^2 k^2}{2m} \qquad (1.1.3)$$

Equation 1.1.2 represents two waves traveling in opposite directions. $C_1\,exp(jkx)$ represents the motion of the electron in the $+x$ direction, while $C_2\,exp(-jkx)$ represents the motion of the electron in the $-x$ direction.

What is the meaning of the variable k? At first it can be observed that the unit in which k is expressed is m^{-1} or cm^{-1}; k is thus a vector belonging to the reciprocal space. In a one-dimensional crystal, however, k can be considered as a scalar number for all practical purposes. The momentum operator, p_x, of the electron, given by relationship A3.2, is:

$$p_x = \frac{\hbar}{j}\frac{\partial}{\partial x}$$

Considering an electron moving along the $+x$ direction in a one-dimensional sample and applying the momentum operator to the wave function $\Psi(x) = C_1\,exp(jkx)$ we obtain:

$$p_x\,\Psi(x) = \frac{\hbar}{j}\frac{d\Psi(x)}{dx} = C_1\,\hbar k\,exp(jkx) = \hbar k\,\Psi(x)$$

The eigenvalues of the operator p_x are thus given by:

$$p_x = \hbar k \qquad (1.1.4)$$

Hence, we can conclude that the number k, called the wave number, is equal to the momentum of the electron, within a multiplication factor $\hbar$. In classical mechanics the speed of the electron is equal to $v=p/m$, which yields $v = \hbar k/m$. We can thus relate the expression of the electron energy, given by Expression 1.1.3, to that derived from classical mechanics:

$$v = \hbar k/m \Rightarrow E = \frac{\hbar^2 k^2}{2m} = \frac{1}{2}mv^2 \qquad (1.1.5)$$

The energy of the free electron is a parabolic function of its momentum k, as shown in Figure 1.1.This result is identical to what is expected from classical mechanics considerations: the "free" electron can take any value of energy in a continuous manner. It is worthwhile noting that electrons

with momentum k or $-k$ have the same energy. These electrons have the same momentum but travel in opposite directions.

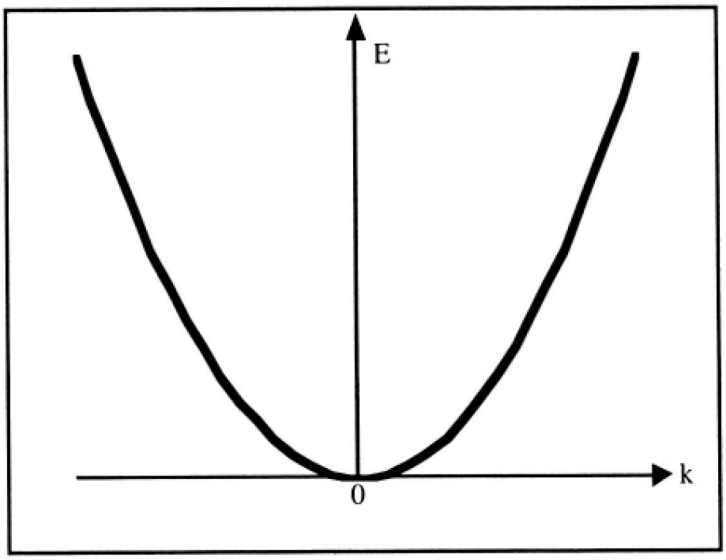

Figure 1.1: Energy vs. k for a free electron.

Another interpretation can be given to k. If we now consider a three-dimensional crystal, $\boldsymbol{k}$ is a vector of the reciprocal space. It is the called the wave vector. Indeed, the expression $exp(j\boldsymbol{kr})$, where $\boldsymbol{r}=(x,y,z)$ is the position of the electron, and represents a plane spatial wave moving in the direction of $\boldsymbol{k}$. The spatial frequency of the wave is equal to $\boldsymbol{k}$, and its spatial wavelength is equal to $\lambda = \frac{2\pi}{|\boldsymbol{k}|}$.

1.1.1.2. The particle-in-a-box approach

After studying the case of a free electron, it is worthwhile to consider a situation where the electron is confined within a small region of space. The confinement can be realized by placing the electron in an infinitely deep potential well from which it cannot escape. In some way the electron can be considered as contained within a box or a well surrounded by infinitely high walls (Figure 1.2). To some limited extent, the particle-in-a-box problem resembles that of electrons in an atom, where the attraction from the positively charge nucleus creates a potential well that "traps" the electrons.

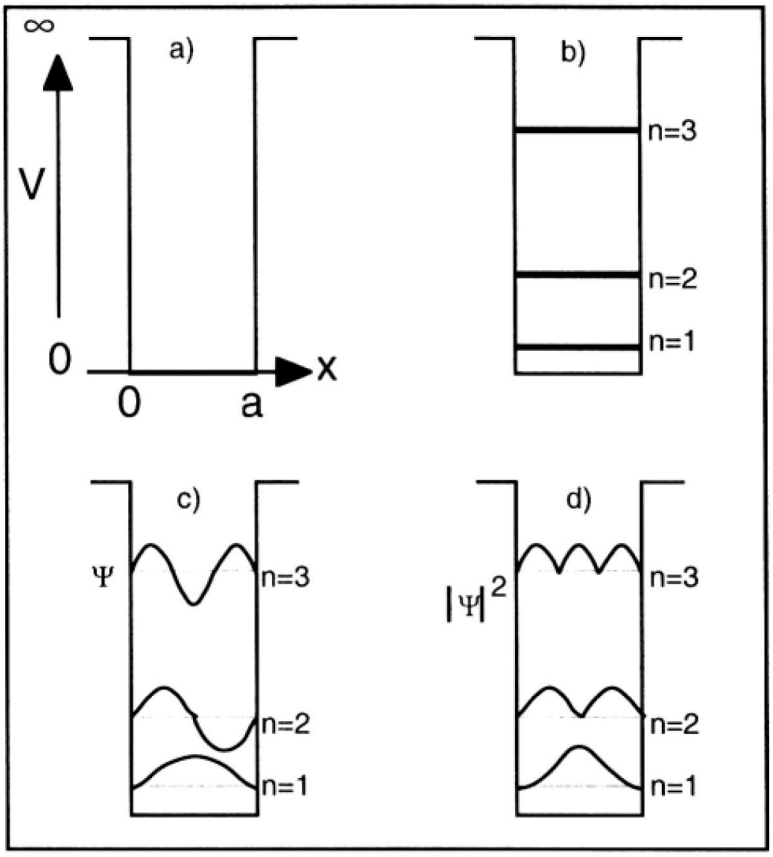

Figure 1.2: Particle in a box: a) Geometry of potential well; b) Energy levels; c) Wave functions; d) Probability density for n=1,2, and 3.

By definition the electron is confined inside the potential well and therefore, the wave function vanishes at the well edges: thus the boundary conditions to our problem are: $\Psi(x \leq 0) = \Psi(x \geq a) = 0$. Within the potential well ($0 \leq x \leq a$), where $V = 0$, the time-independent Schrödinger equation can be written as:

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} = E\,\Psi(x) \qquad (1.1.6)$$

which can be rewritten in the following form:

$$\frac{d^2\Psi(x)}{dx^2} + k^2\Psi(x) = 0 \quad with\ k = \sqrt{2mE/\hbar^2} \quad or \quad E = \frac{\hbar^2 k^2}{2m} \qquad (1.1.7)$$

The solution to this homogenous, second-order differential equation is:

$$\Psi(x) = A\,sin(kx) + B\,cos(kx) \qquad (1.1.8)$$

Using the first boundary condition $\Psi(x=0) = 0$ we obtain $B = 0$. Using the second boundary condition $\Psi(a) = 0$ we obtain $A\,sin(ka) = 0$ and therefore:

$$k = \frac{n\pi}{a} \quad \text{with} \quad n = 1,2,3,... \tag{1.1.9}$$

The wave function is thus given by: $\Psi_n(x) = A_n \sin(\frac{n\pi x}{a})$ (1.1.10)

and the energy of the electron is: $E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$ (1.1.11)

This result is quite similar to that obtained for a free electron, in both cases the energy is a function of the squared momentum. The difference resides in the fact that in the case of a free electron, the wave number k and the energy E can take any value, while in the case of the particle-in-a-box problem, k and E can only take discrete values (replacing k by $n\pi/a$ in Expression 1.1.3 yields Equation 1.1.11). These values are fixed by the geometry of the potential well. Intuitively, it is interesting to note that if the width of the potential well becomes very large ($a \rightarrow \infty$) the different values of k become very close to one another, such that they are no longer discrete values but rather form a continuum, as in the case for the free electron.

Which values can k take in a finite crystal of macroscopic dimensions? Let us consider the example of a one-dimensional linear crystal having a length L (Figure 1.3). If we impose $\Psi(x=0) = 0$ and $\Psi(x=L) = 0$ as in the case of the particle-in-the-box approach, Relationships 1.1.9 and 1.1.11 tell us that the permitted values for the momentum and for the energy of the electron will depend on the length of the crystal. This is clearly unacceptable for we know from experience that the electrical properties of a macroscopic sample do not depend on its dimensions.

Much better results are obtained using the Born-von Karman boundary conditions, referred to as cyclic boundary conditions. To obtain these conditions, let us bend the crystal such that $x=0$ and $x=L$ become coincident. From the newly obtained geometry it becomes evident that for any value of x, we have the cyclical boundary conditions: $\Psi(x+L) = \Psi(x)$. Using the free-electron wave function (Expression 1.1.2), and taking into account the periodic nature of the problem, we can write:

$$\Psi(x+L) = A\, exp(jk(x+L)) = A\, exp(jkx)\, exp(jkL) = A\, exp(jkx) = \Psi(x)$$

which imposes:

$$exp(jkL) = 1 \quad \Rightarrow \quad k = \frac{2\pi n}{L} \tag{1.1.12}$$

where n is an integer number.

In the case of a three-dimensional crystal with dimensions (L_x, L_y ,L_z), the Born-von Karman boundary conditions can be written as follows:

$$k_x = \frac{2\pi n_x}{L_x} \;,\; k_y = \frac{2\pi n_y}{L_y} \quad \text{and} \quad k_z = \frac{2\pi n_z}{L_z} \tag{1.1.13}$$

where n_x, n_y, n_z are integer numbers.

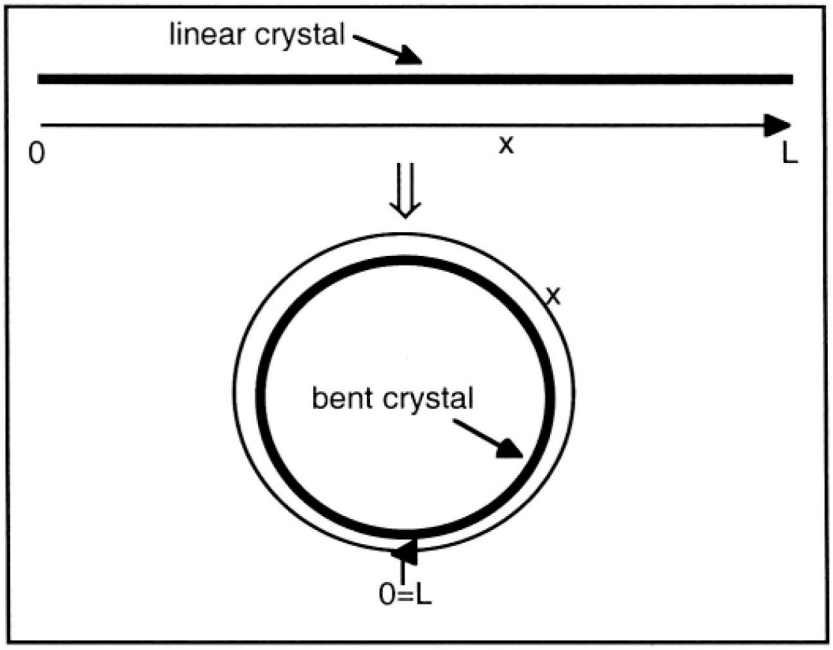

Figure 1.3: Bending of a crystal; Born-von Karman boundary conditions.

1.1.2. Energy bands of a crystal (intuitive approach)

In a single atom, electrons occupy discrete energy levels. What happens when a large number of atoms are brought together to form a crystal? Let us take the example of a relatively simple element with low atomic number, such as lithium (Z=3). In a lithium atom, two electrons of opposite spin occupy the lowest energy level (1s level), and the remaining third electron occupies the second energy level (2s level). The electronic configuration is thus $1s^2\ 2s^1$. All lithium atoms have exactly the same electronic configuration with identical energy levels. If an hypothetical molecule containing two lithium atoms is formed, we are now in the presence of a system in which four electrons "wish" to have an energy equal to that of the 1s level. But because of the Pauli exclusion principle, which states that only two electrons of opposite spins can occupy the same energy level, only two of the four 1s electrons can occupy the 1s level. This clearly poses a problem for the molecule. The problem is solved by splitting the 1s level into two levels having very close, but nevertheless different energies (Figure 1.4).

If a crystal of lithium containing N number of atoms is now formed, the system will contain N number of 1s energy levels. The same consideration is valid for the 2s level. The number of atoms in a cubic centimeter of a crystal is on the order of 5×10^{22}. As a result, each energy level is split into 5×10^{22} distinct energy levels which extend throughout the crystal. Each of these levels can be occupied by two electrons by virtue of the Pauli exclusion principle. In practice, the energy difference between the highest and the lowest energy value resulting from this process of splitting an energy level is on the order of a few electron-volts; therefore, the energy difference between two neighboring energy levels is on the order of 10^{-22} eV. This value is so small that one can consider that the energy levels are no longer discrete, but form a continuum of permitted energy values for the electron. This introduces the concept of energy bands in a crystal. Between the energy bands (between the 1s and the 2s energy bands in Figure 1.4) there may be a range of energy values which are not permitted. In that case, a forbidden energy gap is produced between permitted energy bands. The energy levels and the energy bands extend throughout the entire crystal. Because of the potential wells generated by the atom nuclei, however, some electrons (those occupying the 1s levels) are confined to the immediate neighborhood of the nucleus they are bound to. The electrons of the 2s band, on the other hand, can overcome nucleus attraction and move throughout the crystal.

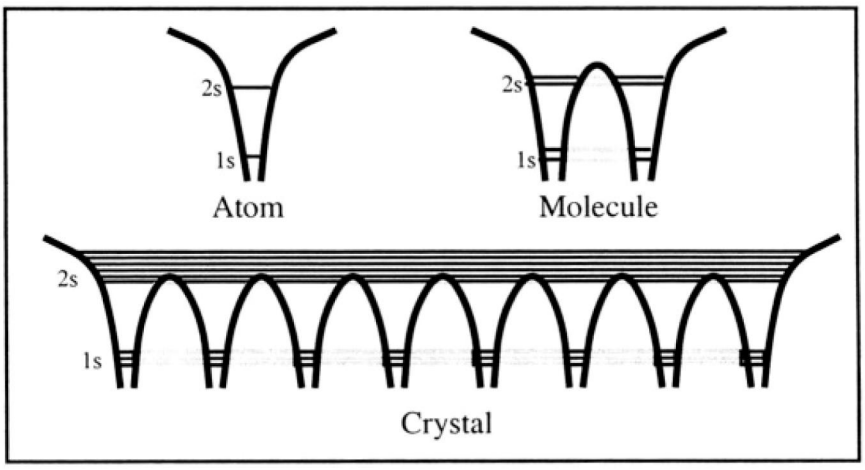

Figure 1.4: Permitted energy levels an atom, an hypothetical molecule, and a crystal of lithium.

1.1.3. Krönig-Penney model

Semiconductors, like metals and some insulators, are crystalline materials. This implies that atoms are placed in an orderly and periodic manner in the material (see Annex A4). While most usual crystalline materials are polycrystalline, semiconductor materials used in the

electronics industry are single-crystal. These single crystals are almost perfect and defect-free, and their size is much greater than any of the microscopic physical dimensions which we are going to deal with in this chapter.

In a crystal each atom of the crystal creates a local potential well which attracts electrons, just like in the lithium crystal described in Figure 1.4. The potential energy of the electron depends on its distance from the atom nucleus. Electrostatics provides us with a relationship establishing the potential energy resulting from the interaction between an electron carrying a charge $-q$ and a nucleus bearing a charge $+qZ$, where Z is the atomic number of the atom and is equal to the number of protons in the nucleus:

$$V(x) = \frac{-Zq^2}{4\pi\varepsilon|x|} \qquad (1.1.14)$$

In this relationship x is the distance between the electron and the nucleus, $V(x)$ is the potential energy and ε is the permittivity of the material under consideration. Equation 1.1.14 ignores the presence of other electrons, such as core electrons "orbiting" around the nucleus. These electrons actually induce a screening effect between the nucleus and outer shell electrons, which reduces the attraction between the nucleus and higher-energy electrons. The energy of the electron as a function of its distance from the nucleus is sketched in Figure 1.5.

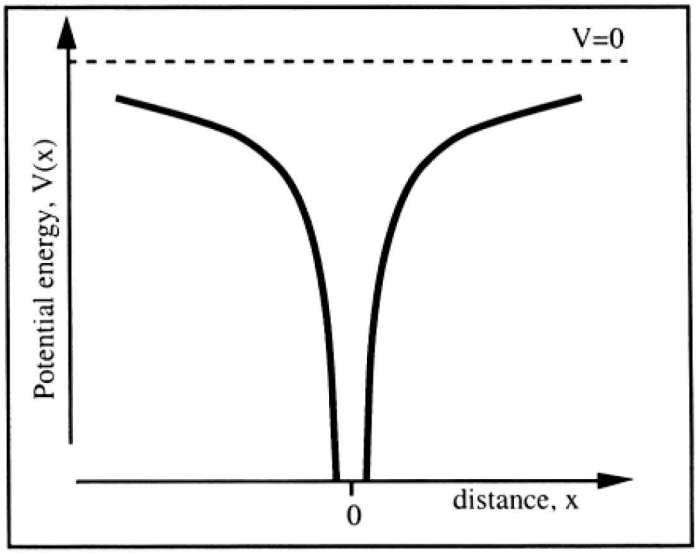

Figure 1.5: Energy of an electron as a function of its distance from the atom nucleus ($V=0$ when $x=\infty$). [2]

How will an electron behave in a crystal? In order to simplify the problem, we will suppose that the crystal is merely an infinite, one-

dimensional chain of atoms. This assumption may seem rather coarse, but it preserves a key feature of the crystal: the periodic nature of the position of the atoms in the crystal. In mathematical terms, the expression of the periodic nature of the atom-generated potential wells can be written as:

$$V(x+a+b) = V(x) \tag{1.1.15}$$

where $a+b$ is the distance between two atoms in the x-direction (Figure 1.6).

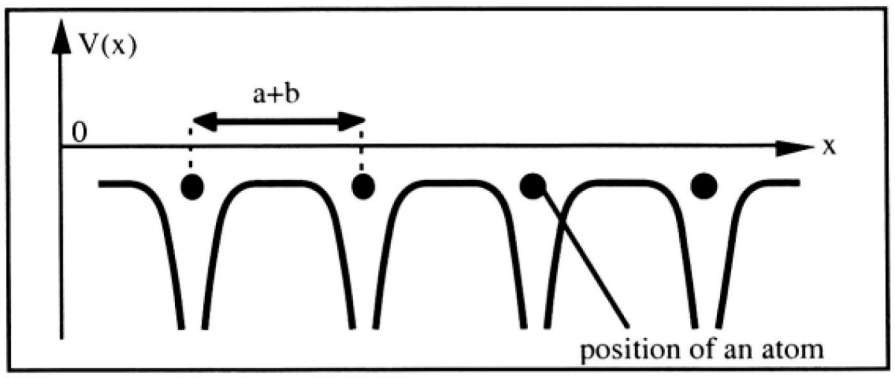

Figure 1.6: Periodic potential in a one-dimensional crystal.

The periodic nature of the potential has a profound influence on the wave function of the electron. In particular, the electron wave function must satisfy the time-independent Schrödinger equation whenever $x+a+b$ is substituted for x in the operators that act on $\Psi(x)$.[3] This condition is obtained if the wave function satisfies the Bloch theorem, which can be formulated as follows:

If $V(x)$ is periodic such that $V(x+a+b) = V(x)$,
then $\Psi(x+a+b) = \Psi(x)\, e^{jk(a+b)}$ (1.1.16)

A second formulation of the theorem is:
If $V(x)$ is periodic such that $V(x+a+b) = V(x)$,
then $\Psi(x) = u(x)\, e^{jkx}$ with $u(x+a+b) = u(x)$.

These two formulations are equivalent since

$$\Psi(x+a+b) = u(x+a+b)\, e^{jk(x+a+b)} = u(x)\, e^{jkx}\, e^{jk(a+b)} = \Psi(x)\, e^{jk(a+b)}$$

Since the potential in the crystal, $V(x)$, is a rather complicated function of x, we will use the approximation made by Krönig and Penney in 1931, in which $V(x)$ is replaced by a periodic sequence of rectangular potential wells.[4] This approximation may appear rather crude, but it preserves the periodic nature of the potential variation in the crystal while allowing a closed-form solution for $\Psi(x)$. The resulting potential is depicted in

Figure 1.7, and the following notations will be used: the inter-atomic distance is $a+b$, the potential energy near an atom is V_1, and the potential energy between atoms is V_0. Both V_1 and V_0 are negative with respect to an arbitrary reference energy, $V=0$, taken outside the crystal. We will study the behavior of an electron with an energy E lying between V_1 and V_0 ($V_0 > E > V_1$). This case is similar to a 1s electron previously shown for lithium.

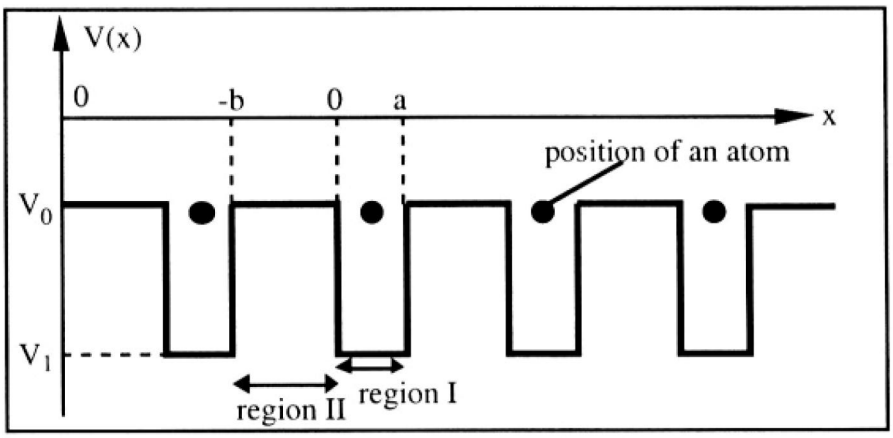

Figure 1.7: Periodic potential of the Krönig and Penney model.

In region I ($0<x<a$), the potential energy is $V(x)=V_1$, and the time-independent Schrödinger equation can be written as:

$$\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\Psi(x) + [E - V_1]\,\Psi(x) = 0 \qquad (1.1.17)$$

In region II ($-b<x<0$), the potential energy is $V(x)=V_0$, and the time-independent Schrödinger equation becomes:

$$\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\Psi(x) + [E - V_0]\,\Psi(x) = 0 \qquad (1.1.18)$$

The solution to these homogenous second-order differential equations are:

$$\Psi_I(x) = A\,exp(j\beta x) + B\,exp(-j\beta x) \quad \text{with } \beta = \sqrt{\frac{2m(E-V_1)}{\hbar^2}} \qquad (1.1.19)$$

and

$$\Psi_{II}(x) = C\,exp(\alpha x) + D\,exp(-\alpha x) \quad \text{with } \alpha = \sqrt{\frac{2m(V_0-E)}{\hbar^2}} \qquad (1.1.20)$$

Note that α and β are real numbers. The periodic nature of the crystal lattice suggests that the wave function satisfies the Bloch theorem (1.1.16) and can be written in the following form:

$$\Psi(x) = u_k(x)\,exp(jkx)$$

where $u_k(x)$ is a periodic function with period $a+b$, which imposes $u_k(x+n(a+b)) = u_k(x)$. One can thus write:

$$\Psi_I(x+n(a+b)) = \Psi_I(x)\ exp(jnk(a+b)) \quad (1.1.21)$$

and

$$\Psi_{II}(x+n(a+b)) = \Psi_{II}(x)\ exp(jnk(a+b)) \quad (1.1.22)$$

Boundary conditions must be used to calculate the four integration constants A, B, C and D of Equations 1.1.19 and 1.1.20. This can be done by imposing the condition that the wave function, $\Psi(x)$, and its first derivative, $d\Psi(x)/dx$, are continuous at $x=0$ and $x=a$. By doing so one obtains the following equations:

◊ $\Psi(x)$ is continuous at $x=0$. Thus $\Psi_I(0) = \Psi_{II}(0)$, which yields:

$$A+B=C+D \quad (1.1.23)$$

◊ $d\Psi(x)/dx$ is continuous at $x=0$. Therefore, $d\Psi_I(0)/dx = d\Psi_{II}(0)/dx$:

$$j\beta(A-B)=\alpha(C-D) \quad (1.1.24)$$

◊ $\Psi(x)$ is continuous at $x=a$ giving $\Psi_I(a) = \Psi_{II}(a)$. Using the Bloch theorem (Equation 1.1.16) at $x=a$ we have $\Psi_I(a) = \Psi_I(-b)\ exp(jk(a+b))$, which yields:

$$exp(jk(a+b))\ [A\ exp(-j\beta b) + B\ exp(j\beta b)] = C\ exp(\alpha a) + D\ exp(-\alpha a) \quad (1.1.25)$$

◊ $d\Psi(x)/dx$ is continuous at $x=a$ giving $d\Psi_I(a)/dx = d\Psi_{II}(a)/dx$. Using Bloch's theorem: $\Psi_I(a) = \Psi_I(-b)\ exp(jk(a+b))$ we obtain:

$$exp(jk(a+b))\ j\beta[A\ exp(-j\beta b) - B\ exp(j\beta b)] = \alpha\ [C\ exp(\alpha a) - D\ exp(-\alpha a)] \quad (1.1.26)$$

Equations (1.1.23) to (1.1.26) form a system of four equations with four unknowns: A, B, C and D. This system can be written in a matrix form:

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ j\beta & -j\beta & -\alpha & \alpha \\ exp(jk(a+b))exp(-j\beta b) & exp(jk(a+b))exp(j\beta b) & -exp(\alpha a) & -exp(-\alpha a) \\ exp(jk(a+b))j\beta exp(-j\beta b) & -exp(jk(a+b))j\beta exp(j\beta b) & -\alpha exp(\alpha a) & \alpha exp(-\alpha a) \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (1.1.27)$$

In order to obtain a non-trivial solution for A, B, C and D, *i.e.* a solution different from $A=B=C=D=0$, the determinant of the 4×4 matrix must be equal to zero, which is equivalent to writing (see Problem 1.5):

$$\frac{\alpha^2 - \beta^2}{2\alpha\beta}\ sinh(\alpha a)\ sin(\beta b) + cosh(\alpha a)\ cos(\beta b) = cos(k(a+b)) \quad (1.1.28)$$

The right-hand term of this equation depends only on *E,* through α and β (Expressions 1.1.19 and 1.1.20). Let us call this term *P(E)* and rewrite Expression 1.1.28 in the following form:

$$P(E) = \cos(k(a+b)) \qquad (1.1.29)$$

The right-hand side of Equation 1.1.29 is sketched as a function of energy in Figure 1.8. Because the argument in the exponential term of (1.1.16) must be imaginary, *k* must be real. Therefore, simultaneous solution of both left- and right-hand side of Equation 1.1.29 imposes that $-1 \leq P(E) \leq 1$. This defines permitted values of energy forming the energy bands, and forbidden values of energy constituting forbidden energy bands. This important result is the same to that intuitively unveiled in Section 1.1.2: in a crystal there are bands of permitted energy values separated by bands of forbidden energy values.

Note: In the case when the electron energy is greater than V_0, $E-V_0$ has a positive value and Equation 1.1.20 becomes:

$$\Psi_{II}(x) = C\exp(j\alpha x) + D\exp(-j\alpha x) \quad \text{with } \alpha = \sqrt{\frac{2m(V_0-E)}{\hbar^2}}$$

In that case the Krönig-Penney model yields an equation different from Relationship 1.1.28; however, the same general conclusion can be drawn, *i.e.,* the existence of permitted and for bidden energy bands.

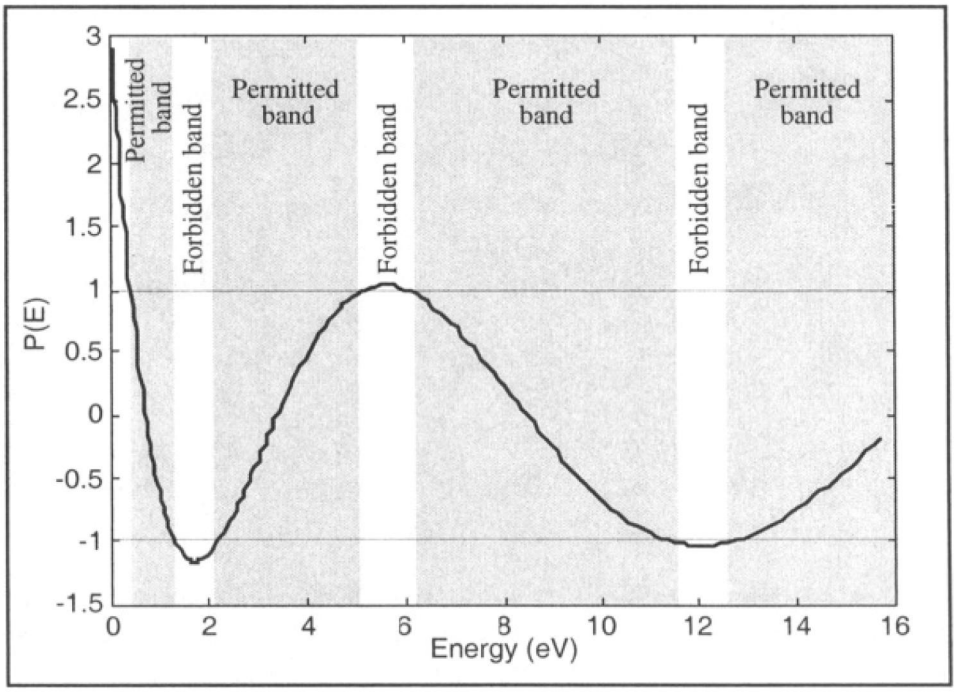

Figure 1.8: P(E) as a function of the electron energy, E, for silicon. The shaded areas correspond to the permitted energy bands, where there is a solution to Equation 1.1.29.

Using Expression 1.1.28 the *E(k)* diagram can be plotted as well. Figure 1.9 presents the energy of the electron as a function of the wave number *k*. The *E(k)* diagram for a free electron is also shown. It can be observed that the energy of the electron in a crystal coarsely represents the same dependence on k as that of a free electron. The main differences reside in the existence of forbidden energy values and curvatures of each segment of the *E(k)* curves.

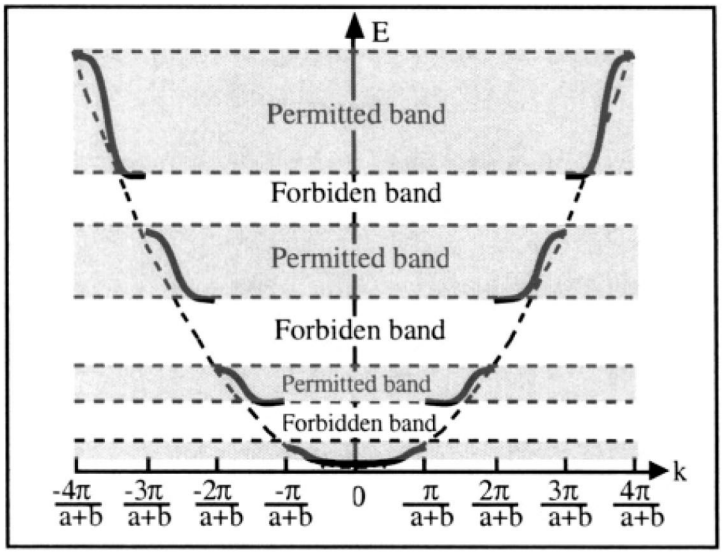

Figure 1.9: Energy versus k in a one-dimension crystal. The dotted line parabola represents the E(k) relationship for a free electron (from Figure 1.1).

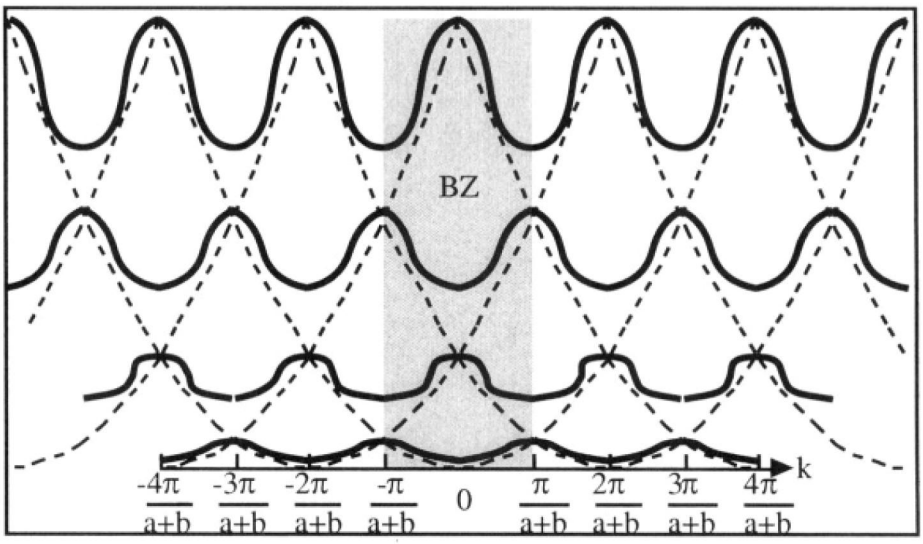

Figure 1.10: E(k) diagram of Figure 1.9, repeated with a $2\pi/(a+b)$ period. The shaded area highlights the first Brillouin zone (BZ).[5]

Because of the periodicity of the crystal lattice (period = $a+b$), the periodicity of the reciprocal lattice (k-space) is $\frac{2\pi}{a+b}$. The $E(k)$ curve can be extended from $k = -\infty$ to $k = +\infty$ with a periodicity of $\frac{2\pi}{a+b}$, which yields the permitted energy values for the entire one-dimensional crystal (Figure 1.10).

The $E(k)$ curves shown in Figure 1.10 can be limited to k-values ranging from $\frac{-\pi}{a+b}$ to $\frac{\pi}{a+b}$ without any loss of information. This particular region of the k-space is called the *first Brillouin zone.* The second Brillouin zone extends from $\frac{-2\pi}{a+b}$ to $\frac{-\pi}{a+b}$ and from $\frac{\pi}{a+b}$ to $\frac{2\pi}{a+b}$, the third zone extends from $\frac{-3\pi}{a+b}$ to $\frac{-2\pi}{a+b}$ and from $\frac{2\pi}{a+b}$ to $\frac{3\pi}{a+b}$, etc.

Applying the Born-von Karman boundary conditions (Expression 1.1.12) to the one-dimensional crystal yields the values for k:

$$exp\ (jkN(a+b)) = 1 \Rightarrow k = \frac{2\pi n}{N(a+b)} \quad (n=0,\pm1,\pm2,\pm3,...) \qquad (1.1.30)$$

where N is the number of lattice cells in the crystal (or the number of atoms in the case of a one-dimension crystal). The length of the crystal is equal to $N(a+b)$. Since we limit our study to the first Brillouin zone, the k-values which have to be considered are given by the following relationship: $\frac{-\pi}{a+b} \leq k < \frac{\pi}{a+b}$ (the value $k = \frac{\pi}{a+b}$ is excluded because it is a duplicate of the $k = \frac{-\pi}{a+b}$ wave number). The corresponding values for n range from $-N/2$ to $(N/2-1)$. Therefore, the values of k to consider are:

$$k = \frac{2\pi n}{N(a+b)} \quad (n=0,\pm1,\pm2,\pm3,...,\pm(\frac{N}{2}-1), -N/2) \qquad (1.1.31)$$

There are thus N wave numbers in the first Brillouin zone, which corresponds to the number of elementary lattice cells. For every wave number there is a permitted energy value in *each* energy band. By virtue of the Pauli exclusion principle, each energy band can thus contain a maximum of $2N$ electrons.

The one-dimensional volume of the first Brillouin zone is equal to $2\pi/(a+b)$. Since it contains N k-values, the density of k-values in the first Brillouin zone is given by:

$$n(k) = density\ of\ k = \frac{number\ of\ k\text{-}values}{volume\ of\ the\ zone}$$

$$= \frac{N}{2\pi/(a+b)} = \frac{N(a+b)}{2\pi} = \frac{L}{2\pi} \qquad (1.1.32)$$

In the case of a three-dimensional crystal, energy band calculations are, of course, much more complicated, but the essential results obtained from the one-dimensional calculation still hold. In particular, there exist permitted energy bands separated by forbidden energy gaps. The 3-D volume of the first Brillouin zone is $8\pi^3 N/V$, where V is the volume of the crystal, the number of wave vectors is equal to the number of elementary crystal lattice cells, N. The density of wave vectors is given by:

$$n(\mathbf{k}) = density\ of\ \mathbf{k} = \frac{number\ of\ \mathbf{k}\text{-}vectors}{volume\ of\ the\ zone} = \frac{NV}{8\pi^3 N} = \frac{V}{8\pi^3} \qquad (1.1.33)$$

1.1.4. Valence band and conduction band

Chemical reactions originate from the exchange of electrons from the outer electronic shell of atoms. Electrons from the most inner shells do not participate in chemical reactions because of the high electrostatic attraction to the nucleus. Likewise, the bonds between atoms in a crystal, as well as electric transport phenomena, are due to electrons from the outermost shell. In terms of energy bands, the electrons responsible for forming bonds between atoms are found in the last occupied band, where electrons have the highest energy levels for the ground-state atoms. However, there is an infinite number of energy bands. The first (lowest) bands contain core electrons such as the 1s electrons which are tightly bound to the atoms. The highest bands contain no electrons. The last ground-state band which contains electrons is called the *valence band,* because it contains the electrons that form the -often covalent- bonds between atoms.

The permitted energy band directly above the valence band is called the *conduction band.* In a semiconductor this band is empty of electrons at low temperature ($T=0K$). At higher temperatures, some electrons have enough thermal energy to quit their function of forming a bond between atoms and circulate in the crystal. These electrons "jump" from the valence band into the conduction band, where they are free to move. The energy difference between the bottom of the conduction band and the top of the valence band is called "forbidden gap" or "bandgap" and is noted E_g.

In a more general sense, the following situations can occur depending on the location of the atom in the periodic table (Figure 1.11):

A: The last (valence) energy band is only partially filled with electrons, even at $T=0K$.

B: The last (valence) energy band is completely filled with electrons at $T=0K$, but the next (empty) energy band overlaps with it (*i.e.:* an empty energy band shares a range of common energy values; $E_g < 0$).

C: The last (valence) energy band is completely filled with electrons and no empty band overlaps with it ($E_g > 0$).

In cases A and B, electrons with the highest energies can easily acquire an infinitesimal amount of energy and jump to a slightly higher permitted energy level, and move through the crystal. In other words, electrons can leave the atom and move in the crystal without receiving any energy. A material with such a property is a *metal.* In case C, a significant amount of energy (equal to E_g or higher) has to be transferred to an electron in order for it to "jump" from the valence band into a permitted energy level of the conduction band. This means that an electron must receive a significant amount of energy before leaving an atom and moving "freely" in the crystal. A material with such properties is either an *insulator* or a *semiconductor.*

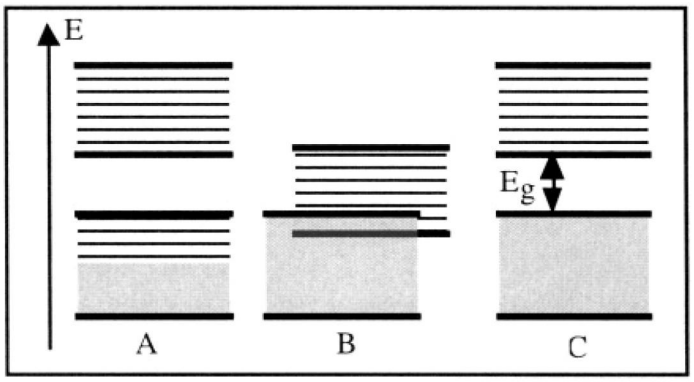

Figure 1.11: Valence band (bottom) and conduction band in a metal (A and B) and in a semiconductor or an insulator (C).[6]

The distinction between an insulator and a semiconductor is purely quantitative and is based on the value of the energy gap. In a semiconductor E_g is typically smaller than 2 eV and room-temperature thermal energy or excitation from visible-light photons can give electrons enough energy for "jumping" from the valence into the conduction band. The energy gap of the most common semiconductors are: 1.12 eV (silicon), 0.67 eV (germanium), and 1.42 eV (gallium arsenide). Insulators have significantly wider energy bandgaps: 9.0 eV (SiO_2), 5.47 eV (diamond), and 5.0 eV (Si_3N_4). In these materials room-temperature thermal energy is not large enough to place electrons in the conduction band.

Beside elemental semiconductors such as silicon and germanium, compound semiconductors can be synthesized by combining elements from column IV of the periodic table (SiC and SiGe) or by combining elements from columns III and V (GaAs, GaN, InP, AlGaAs, AlSb, GaP, AlP and AlAs). Elements from other columns can sometimes be used as well (HgCdTe, CdS,...). Diamond exhibits semiconducting properties at high temperature, and tin (right below germanium in column IV of the periodic table) becomes a semiconductor at low temperatures. About 98% of all semiconductor devices are fabricated from single-crystal silicon, such as integrated circuits, microprocessors and memory chips. The remaining 2% make use of III-V compounds, such as light-emitting diodes, laser diodes and some microwave-frequency components.

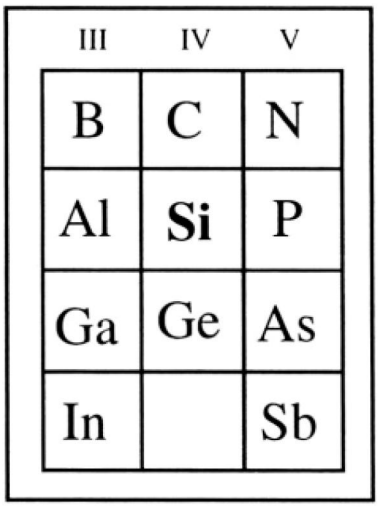

Figure 1.12: Main elements used in semiconductor technology (elemental semiconductors such as Si, and compound semiconductors such as GaAs).

It is worthwhile mentioning that it is possible for non-crystalline materials to exhibit semiconducting properties. Some materials, such as amorphous silicon, where the distance between atoms varies in a random fashion, can behave as semiconductors. The mechanisms for the transport of electric charges in these materials are, however, quite different from those in crystalline semiconductors.[7].

It is convenient to represent energy bands in real space instead of *k*-space. By doing so one obtains a diagram such as that of Figure 1.13, where the x-axis defines a physical distance in the crystal. The maximum energy of the valence band is noted E_V, the minimum energy of the conduction band is noted E_C, and the width of the energy bandgap is E_g.

It is also appropriate to introduce the concept of a Fermi level. The Fermi level, E_F, represents the maximum energy of an electron in the

material at zero degree Kelvin (*0 K*). At that temperature, all the allowed energy levels below the Fermi level are occupied, and all the energy levels above it are empty. Alternatively, the Fermi level is defined as an energy level that has a 50% probability of being filled with electrons, even though it may reside in the bandgap. In an insulator or a semiconductor, we know that the valence band is full of electrons, and the conduction band is empty at *0 K*. Therefore, the Fermi level lies somewhere in the bandgap, between E_V and E_C. In a metal, the Fermi level lies within an energy band.

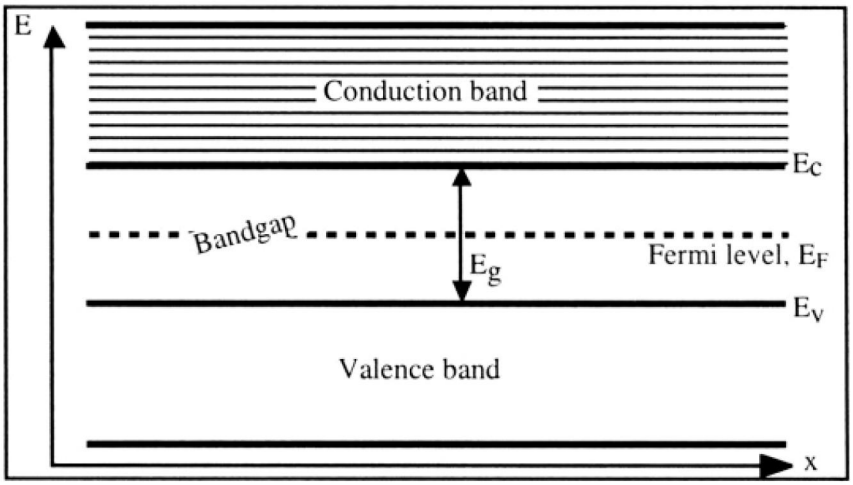

Figure 1.13: Valence and conduction band in real space.

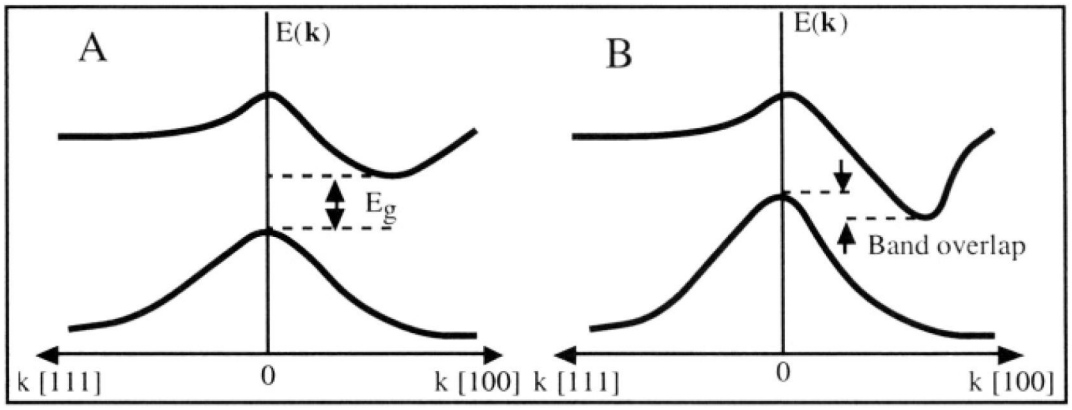

Figure 1.14: Examples of energy band extrema (minimum of the conduction band and maximum of the valence band in two crystals). In crystal A, E_g is the bandgap energy. There is no bandgap in crystal B because the conduction and the valence bands overlap.

It is impossible to represent the energy bands as a function of $\boldsymbol{k} = \boldsymbol{k}(k_x,k_y,k_z)$ for a three-dimensional crystal in a drawing made on a two-dimensional sheet of paper. One can, however, represent *E(k)* along main crystal directions in *k*-space and place them on a single graph. For

example, Figure 1.14 represents the maximum of the valence band and the minimum of the conduction band as function of *k* in the [100] and the [111] directions for two crystals. Crystal A is an insulator or a semiconductor $(E_g > 0)$; crystal B is a metal $(E_g < 0)$.

The energy band diagrams, plotted along the main crystal directions, allow us to analyze some properties of semiconductors. For instance, in Figure 1.15.B the minimum energy in the conduction band and the maximum energy in the valence band occur at the same *k*-values (k=0). A semiconductor exhibiting this property is called a direct-band semiconductor. Examples of direct-bandgap semiconductors include most compound elements such as gallium arsenide (GaAs). In such a semiconductor, an electron can "fall" from the conduction band into the valence band without violating the conservation of momentum law, *i.e.* an electron can fall from the conduction band to the valence band without a change in momentum. This process has a high probability of occurrence and the energy lost in that "jump" can be emitted in the form of a photon with an energy $h\nu=E_g$. In Figure 1.15.A, the minimum energy in the conduction band and the maximum energy in the valence band occur at different *k*-values. A semiconductor exhibiting this property is called an indirect bandgap semiconductor. Silicon and germanium are indirect-bandgap semiconductors. In such a semiconductor, an electron cannot "fall" from the conduction band into the valence band without a change in momentum. This tremendously reduces the probability of a direct "fall" of an electron from the conduction band into the valence band, as will be discussed in Chapter 3.

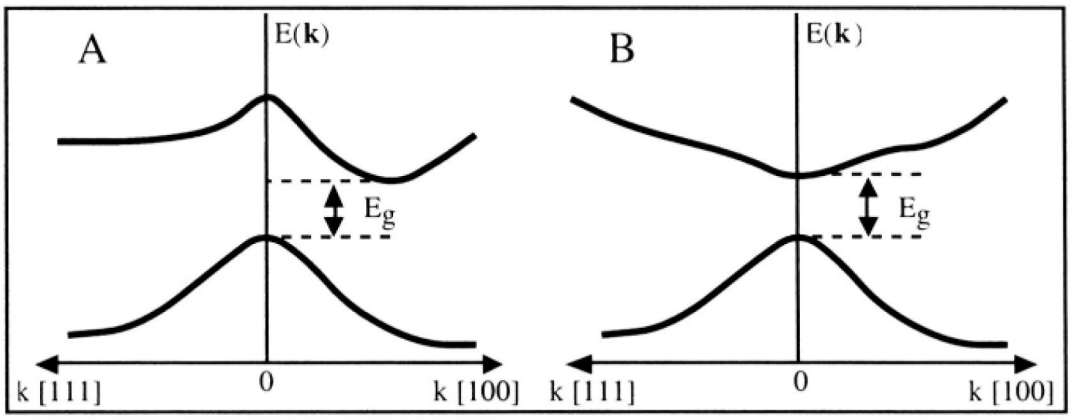

Figure 1.15: A: Indirect bandgap semiconductor, B: Direct bandgap semiconductor. [8]

1.1.5. Parabolic band approximation

For electrical phenomena, only the electrons located near the maximum of the valence band and the minimum of the conduction band

are of interest. These are the energy levels where free moving electrons and missing valence electrons are found. In that case, as can be seen in Figure 1.15, the energy dependence on momentum can be approximated by a square parabolic function. Near the minimum of the conduction band one can thus write:

$$E(k)=E_{min} + A(k-k_{min})^2 \qquad (1.1.34a)$$

Near the maximum of the valence band one can write:

$$E(k)=E_{max} - B(k-k_{max})^2 \qquad (1.1.34b)$$

with A and B being constants. This approximation is called the "parabolic band approximation" and resembles the $E(k)$ relationship found for the free electron model.

1.1.6. Concept of a hole

To facilitate the understanding of electrical conduction in a solid one can make a comparison between the flow of electrical charge in the energy bands and the movement of water drops in a pipe. Let us consider (Figure 1.16.A) two pipes which are sealed at both ends. The bottom pipe is completely filled with water and the top pipe contains no water (it is filled with air). In our analogy between electricity and water, each drop of water corresponds to an electron, and the bottom and top pipes correspond to the valence and the conduction band, respectively.[9] Tilting the pipes corresponds to the application of an electric field to the semiconductor. When the filled or empty pipes are tilted, no movement or flow of water is observed, *i.e.:* there is no electric current flow in the semiconductor. Thus the semiconductor behaves as an insulator (Figure 1.16.A).

Let us now remove a drop of water from the bottom pipe and place it in the top pipe, which corresponds to "moving" an electron from the valence to the conduction band. If the pipes are now tilted, a net flow of liquid will be observed, which correspond to an electrical current flow in the semiconductor (Figure 1.16.B).

The water flow in the top pipe (conduction band) is due to the movement of the water drop (electron). In addition, there is also water flow in the bottom pipe (valence band) since drops of water can occupy the space left behind as the air bubble moves. It is, however, easier to visualize the motion of the bubble itself instead of the movement of the "valence" water.

If, in this water analogy, an electron is represented by a drop of water, a bubble or absence of water in the "valence" pipe represents what is called a *hole.* Hence, a hole is equivalent to a missing electron in the crystal valence band. A hole is not a particle and it does not exist by itself. It draws its existence from the absence of an electron in the crystal, just like a bubble in a pipe exists only because of a lack of water. Holes can move in the crystal through successive "filling" of the empty space left by a missing electron. The hole carries a positive charge $+q$, as the electron carries a negative charge $-q$ ($q = 1.6 \times 10^{-19}$ Coulomb).

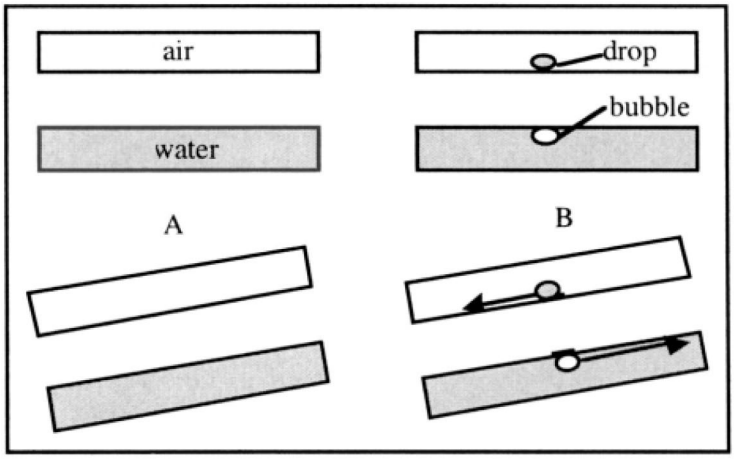

Figure 1.16: Energy bands and electrical conduction: water analogy.

1.1.7. Effective mass of the electron in a crystal

The mass m of an electron can be defined by the relationship $\boldsymbol{F}=m\boldsymbol{a}$ where $\boldsymbol{a}$ is the acceleration the electron undergoes under the influence of an external applied force $\boldsymbol{F}$. The fact that the electron is in a crystal will influence its response to an applied force. As a result, the apparent, "effective" mass of the electron in a crystal will be different from that of an electron in a vacuum.

In the case of a free electron Relationship 1.1.3 can be used to find the mass of the electron [10]:

$$E = \frac{\hbar^2 k^2}{2m} \Rightarrow m = \frac{\hbar^2}{d^2E/dk^2} \qquad (1.1.35)$$

where $m = m_o = 9.11\text{x}10^{-28}$ gram is the mass of the electron in a vacuum. The mass is a constant since E is a square function of k.

Using the rightmost term of 1.1.35 as the definition of the electron mass and using Equations 1.1.28 and 1.1.29 which defines the relationship

between E and k in a one-dimensional crystal, the mass of an electron within an energy band can be calculated:

$$P(E) = \cos k(a+b) \quad \text{and} \quad m^* = \frac{\hbar^2}{d^2E/dk^2} \tag{1.1.36}$$

where m^* is called the "effective mass" of the electron in a crystal. Unlike the case of a free electron the effective mass of the electron in a crystal is not constant, but it varies as a function of k (Figure 1.17).

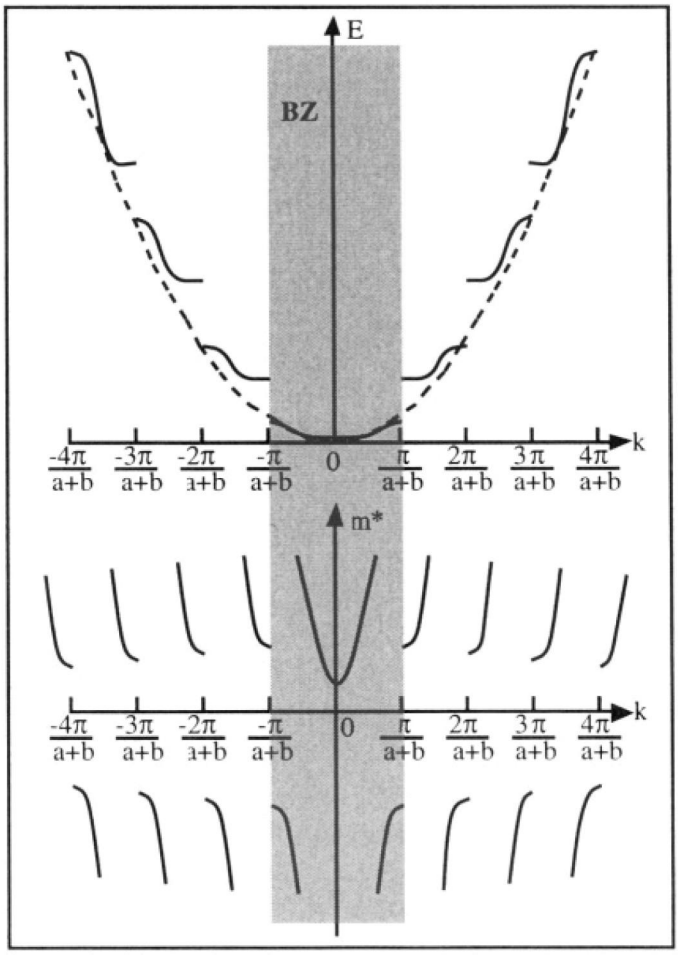

Figure 1.17: Electron energy and effective mass vs. k in a one-dimensional crystal. The first Brillouin zone (BZ) is shown in gray.

Additionally, the mass in the crystal will be different for differing energy bands. The following general observations can be made:

◊ if the electron is in the upper half of an energy band, its effective mass is negative

◊ if the electron is in the lower half of an energy band, its effective mass is positive

◊ if the electron is near the middle of an energy band, its effective mass tends to be infinite

The negative mass of electrons located in the top part of an energy band may come as a surprise, but can easily be explained using the concept of a hole. Let us consider the acceleration, a, given to an electron with charge $-q$ and negative mass, $-m^*$, by an electric field, $\mathcal{E}$. It is easy to realize that this acceleration corresponds to a hole with positive mass, $+m^*$, and positive charge ,+q, since:

$$a = \frac{F}{-m^*} = \frac{-q\mathcal{E}}{-m^*} = \frac{q\mathcal{E}}{m^*} \quad with \quad m^* > 0 \tag{1.1.37}$$

In the case of a three-dimensional crystal the expression of the effective mass is more complicated because the acceleration of an electron can be in a direction different from that of the applied force. In that case the effective mass is expressed by a 3×3 tensor:

$$m^* = \begin{bmatrix} m^*_{xx} & m^*_{xy} & m^*_{xz} \\ m^*_{yx} & m^*_{yy} & m^*_{yz} \\ m^*_{zx} & m^*_{zy} & m^*_{zz} \end{bmatrix}$$

$$\text{with } m^*_{xx} = \frac{\hbar^2}{\partial^2 E/\partial k_x^2}, \; m^*_{xy} = \frac{\hbar^2}{\partial^2 E/\partial k_x \partial k_y}, \text{ etc.} \tag{1.1.38}$$

Usually physics of semiconductor devices deals only with electrons situated near the minimum of the conduction band or holes located near the maximum of the valence band. In the case of silicon the mass of electrons near the minimum of the conduction band along the [100] k_x-direction is equal to $m^*_l = 0.97\, m_o$, and in the orthogonal directions it is $m^*_t = 0.19\, m_o$. m^*_l is called the longitudinal mass and m^*_t the transversal mass, while m_o is the mass of a free electron in a vacuum. These masses are related to the energy by the following relationship called "parabolic energy band approximation":

$$E(\mathbf{k}) = E_c(k_m) + \frac{\hbar^2}{2m^*_l}(k_x - k_{m,x})^2 + \frac{\hbar^2}{2m^*_t}(k_y^2 + k_z^2) \tag{1.1.39}$$

where $E_c(k_m)$ is the lowest energy state in the conduction band along the [100] or [-100] k_x-directions (Figure 1.18). In most practical cases, for the sake of simplicity, the effective mass is considered to be constant. In that case m^* is approximated by a scalar value.

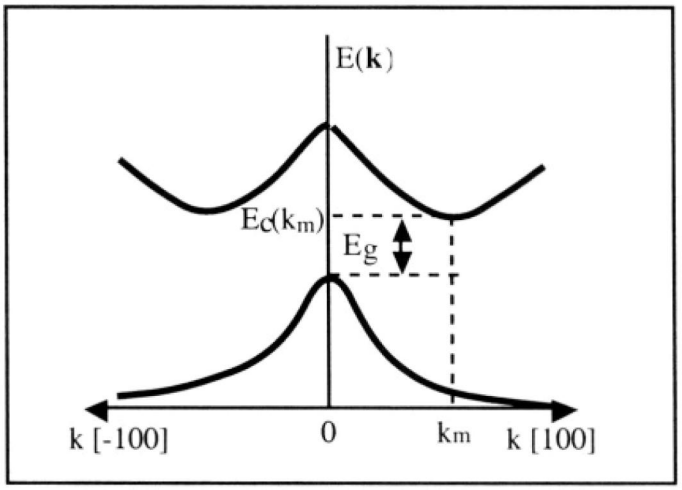

Figure 1.18: Energy bands $E(k)$ along the [100] and [-100] crystallographic directions in silicon. E_g = 1.12 eV.

In a one-dimensional case the square-law dependence of the energy on k, $E(k) = E_c(k_m) + \frac{\hbar^2}{2m_l^*}(k_x - k_{m,x})^2$ is illustrated by Figure 1.19.A There are two vectors $k_m + dk$ and $k_m - dk$ which correspond to a same energy value $E_c(k_m + dk)$. In a two-dimensional crystal (Figure 1.19.B) the locus of (k_x, k_y) values corresponding to the energy level $E_c(k_m + dk)$ is an ellipse in the (k_x, k_y) plane.

The three-dimensional case cannot be drawn on a sheet of paper, but extrapolating from the 1D and 2D cases it is easy to conceive that the k values corresponding to the energy level $E_c(k_m + dk)$ form ellipsoids in the (k_x, k_y, k_z) space (Figure 1.19.C). In a three-dimensional crystal such as silicon there are 6 equivalent crystal directions ([100], [-100], [010], [0-10], [001] and [00-1]) which present an energy minimum (conduction band minimum). The locus of k-values corresponding to a particular energy value is 6 ellipsoids (Figure 1.19.C). The center of these ellipsoids are the six k-values corresponding to the conduction band energy minima. For simplification the ellipsoids can be approximated by spheres (Figure 1.19.D), which is equivalent to equating the transverse and the longitudinal mass $(m_l^* = m_t^*)$. The energy in the vicinity of the maximum of the valence band is given by:

$$E(k) = E_v(0) - \frac{\hbar^2}{2m^*}(k_x^2 + k_y^2 + k_z^2) \qquad (1.1.40)$$

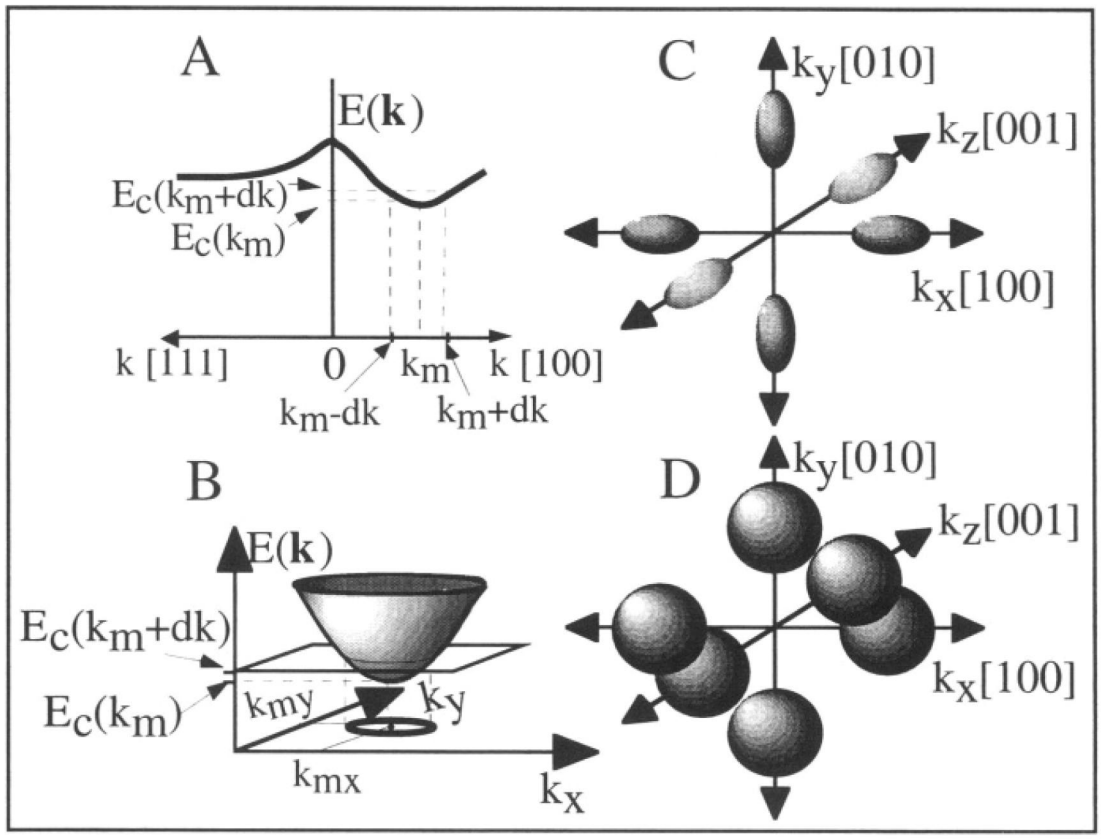

Figure 1.19: A: Values of k of equal energy. A: one-dimensional case; B: two-dimensional case; C: three-dimensional case (silicon); D: approximation of ellipsoids by spheres (silicon).

1.1.8. Density of states in energy bands

The density of permitted states in a three-dimensional crystal is given by (1.1.33). Its value is:

$$n(\boldsymbol{k}) = 1/8\pi^3 \tag{1.1.41}$$

per crystal unit volume. If we define $f(\boldsymbol{k})$ as the probability that these states are occupied, then the electron density, n, in an energy band $E_n(\boldsymbol{k})$ can be calculated by integrating the product of the density of states by the occupation probability over the first Brillouin zone:

$$n = \int_{BZ} n(\boldsymbol{k})\, f(\boldsymbol{k})\, d\boldsymbol{k} \tag{1.1.42}$$

Similarly, the density of holes within an energy band is given by:

$$p = \int_{BZ} n(\boldsymbol{k})\, [1-f(\boldsymbol{k})]\, d\boldsymbol{k} \tag{1.1.43}$$

The function $n(\boldsymbol{k})$ represents the density of permitted states in an energy band. The function $f(\boldsymbol{k})$ is a statistical distribution function which is a

function of the energy, $E_n(\mathbf{k})$. Under thermodynamic equilibrium conditions, $f(\mathbf{k})$ is the Fermi-Dirac distribution function defined as:[11]

Fermi-Dirac Distribution

$$f(\mathbf{k}) = \frac{1}{1 + \exp[(E_n(\mathbf{k})-E_F)/kT]} \qquad (1.1.44a)$$

or

$$f(E) = \frac{1}{1 + \exp[(E-E_F)/kT]} \qquad (1.1.44b)$$

where E_F is an energy value called the "Fermi level", k is the Boltzmann constant, and T is the temperature in Kelvin. The Fermi-Dirac function is plotted in Figure 1.1.20 for $T > 0K$. It is worthwhile noting that $f(E) = 0.5$ if $E = E_F$, regardless of temperature. Therefore, a second definition of the Fermi level is that it is the energy level which has a 50% probability of being occupied.

In order to integrate Expressions 1.1.42 or 1.1.43 easily, the dependency of n and f on $\mathbf{k}$ must be transformed into a dependency on the energy, E. To do this, let us consider a unit cell of the reciprocal crystal lattice where k_x, k_y and k_z are given by Relationship 1.1.13 with $n_x=n_y=n_z=1$; the volume of this cell is equal to $k_xk_yk_z=8\pi^3/L^3$. If the crystal has unit volume, then $L^3=1$ and the volume of a unit cell of a unit-volume crystal in k-space is equal to $8\pi^3$. In this crystal the volume of a spherical shell with a thickness dk in k-space is given by (volume of a shell of thickness dk in Figure 1.19.D):

$$\frac{4\pi}{3}\,[(\mathbf{k}+d\mathbf{k})^3 - \mathbf{k}^3] \cong 4\pi\,\mathbf{k}^2\,d\mathbf{k} \qquad (1.1.45)$$

The number of unit cells in that volume is given by the volume of the shell divided by the unit volume of the cell:

$$\frac{4\pi\,\mathbf{k}^2\,d\mathbf{k}}{8\pi^3} = \frac{\mathbf{k}^2}{2\pi^2}\,d\mathbf{k} \qquad (1.1.46)$$

The number of $\mathbf{k}$ vectors (and thus the number of energy levels, since there is an energy level for each k vector) is equal to the number of unit cells. Using the Pauli exclusion principle (which states that there can be only 2 electrons for each $\mathbf{k}$ vector), the number of electrons is given by:

$$n(\mathbf{k})\,d\mathbf{k} = \frac{\mathbf{k}^2}{\pi^2}\,d\mathbf{k} \qquad (1.1.47)$$

Using the parabolic band approximation, $E(\mathbf{k})=\hbar^2\mathbf{k}^2/2m^*$ and using a constant effective mass, one obtains:

$$n(E)\,dE = \frac{1}{2\pi^2\hbar^3}(2m^*)^{3/2}E^{1/2}\,dE \qquad (1.1.48)$$

This equation yields the density of states for a particle of mass m^* having an energy ranging between E and $E+dE$. In the case of electrons with a mass m_e^* located near the bottom of the conduction band, the energy is referenced to the minimum of the conduction band (E_C), which yields:

$$n(E)\,dE = \frac{1}{2\pi^2\hbar^3}(2m_e^*)^{3/2}(E-E_C)^{1/2}\,dE \qquad (1.1.49)$$

In the case of holes with a mass m_h^* located near the top of the valence band, the energy is referenced to the maximum of the valence band (E_V), and one obtains:

$$n(E)\,dE = \frac{1}{2\pi^2\hbar^3}(2m_h^*)^{3/2}(E_V-E)^{1/2}\,dE \qquad (1.1.50)$$

Integration of Equations 1.1.42 and 1.1.43 can now be performed. The integration can be further simplified by approximating the Fermi-Dirac (FD) distribution by the Maxwell-Boltzmann (MB) distribution. Both distributions are almost identical provided that $E-E_F$ is large enough, which is the case in typical semiconductors (*i.e.* $\frac{1}{1+\exp(u)} \cong \exp(-u)$ when $u >> 1$ (see Problem 1.10):

Maxwell-Boltzmann Distribution

$$f(E) = \underbrace{\frac{1}{1+\exp[(E-E_F)/kT]}}_{\text{Fermi-Dirac}} \cong \underbrace{\exp\left[-\frac{E-E_F}{kT}\right]}_{\text{Maxwell-Boltzmann}} \qquad (1.1.51)$$

To calculate the electron density, n, in the conduction band (CB) we replace the integral over k-values in Relationship 1.1.42 by an integral over energy:

$$n = \int_{BZ} n(\mathbf{k})f(\mathbf{k})d\mathbf{k} = \int_{CB} n(E)f(E)dE \quad (\text{cm-3})$$

$$= \frac{1}{2\pi^2\hbar^3}(2m_e^*)^{3/2}\int_{CB}(E-E_C)^{1/2}\exp\left[-\frac{E-E_F}{kT}\right]dE \qquad (1.1.52)$$

In a typical semiconductor the vast majority of the electrons in the conduction band have an energy close to E_C. Therefore, the lower and upper bound of the integral can thus be replaced by E_C and infinity,

respectively. To integrate, a change of variables can be used where $y = (E-E_C)/kT$, which yields:

$$n = \frac{1}{2\pi^2\hbar^3}(2m_e^*)^{3/2}\int_{E_C}^{\infty}(E-E_C)^{1/2}\exp\left[-\frac{E-E_F}{kT}\right]dE$$

$$= \frac{1}{2\pi^2\hbar^3}(2m_e^*kT)^{3/2}\exp\left[-\frac{E_C-E_F}{kT}\right]\int_0^{\infty}y^{1/2}\exp(-y)\,dy$$

$$= \frac{1}{2\pi^2\hbar^3}(2m_e^*kT)^{3/2}\exp\left[-\frac{E_C-E_F}{kT}\right]\frac{\sqrt{\pi}}{2} \qquad (1.1.53)$$

or: $$n=N_C\exp\left[-\frac{E_C-E_F}{kT}\right] \text{ with } N_C = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \text{ (cm-3)} \qquad (1.1.54)$$

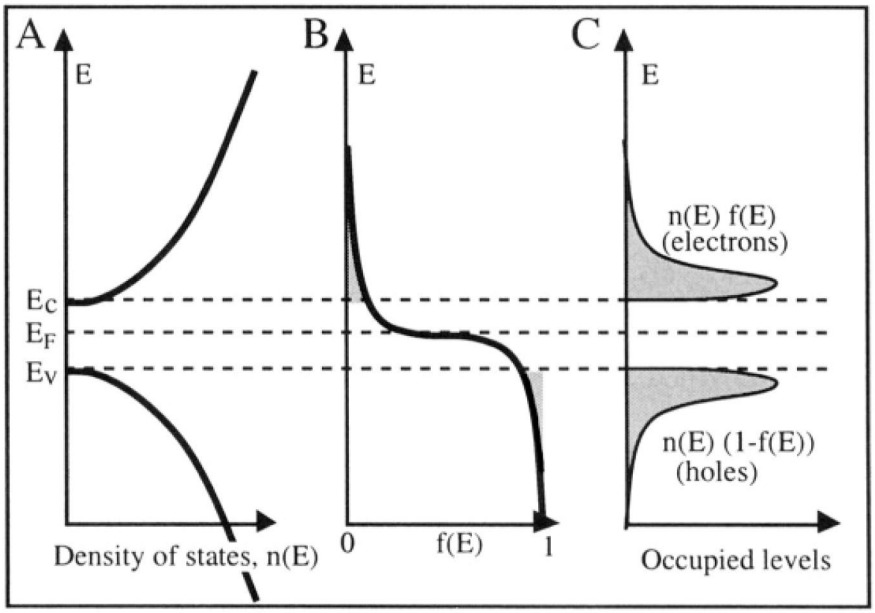

Figure 1.20: Density of states near the bottom of the conduction band and the top of the valence band (A), Fermi-Dirac function (B), and density of holes and electrons in the conduction and valence bands (C), for T≠0K. Note that at T=0K, f(E)=1 for $E<E_F$ and f(E)=0 for $E>E_F$. At T=0K the valence band is completely filled with electrons (empty of holes) and there are no electrons in the conduction band.[12]

N_C is called the "*effective density of states in the conduction band*". It represents the number of states having an energy equal to E_C which, when multiplied by the occupation probability at E_C, yields the number of electrons in the conduction band. Likewise the total number of holes in

the valence band can be calculated using this technique, based on Equation (1.1.43). The effective density of states for holes in the valence band is:

$$p=N_v \exp\left[-\frac{E_F-E_v}{kT}\right] \text{ with } N_v = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \text{ (cm-3)} \qquad (1.1.55)$$

The density of holes and electrons in the conduction and valence bands is shown in Figure 1.20.C for a Fermi level E_F at midpoint of E_c and E_v.

1.2. Intrinsic semiconductor

By virtue of Expressions 1.1.54 and 1.1.55 the product of the electron concentration and hole concentration in a semiconductor under thermodynamic equilibrium conditions is given by:

$$pn = N_c \exp\left[-\frac{E_c-E_F}{kT}\right] N_v \exp\left[-\frac{E_F-E_v}{kT}\right] = N_c N_v \exp(-E_g/kT)$$

$$= 32\left(\frac{\pi^2 k^2 m_e^* m_h^*}{h^4}\right)^{3/2} T^3 \exp(-E_g/kT) \equiv n_i^2 \qquad (1.2.1a)$$

where n_i is called the intrinsic carrier concentration.

pn Product under Thermodynamic Equilibrium

$$pn = n_i^2 \qquad (1.2.1b)$$

A semiconductor is said to be "intrinsic" if the vast majority of its free carriers (electrons and holes) originate from the semiconductor atoms themselves. In that case if an electron receives enough thermal energy to "jump" from the valence band to the conduction band, it leaves a hole behind in the valence band. Thus, every hole in the valence band corresponds to an electron in the conduction band, and the number of conduction electrons is exactly equal to the number of valence holes:

$$p = n = n_i \qquad (1.2.2)$$

and

$$E_F = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln\left(\frac{m_h^*}{m_e^*}\right) \equiv E_i \qquad (1.2.3)$$

or, if $m_e^* = m_h^*$ (simplifying approximation): where

$$E_i = \frac{E_c + E_v}{2} \qquad (1.2.4)$$

where E_i is called the intrinsic energy level. It is the energy of the Fermi level in an intrinsic semiconductor. One can generally consider that it lies right in the middle of the energy bandgap (Expression 1.2.4). n_i is the intrinsic carrier concentration (electrons or holes, $n_i=p_i$) and is a only a function of temperature and of the material through E_g. In silicon n_i is equal to 1.45×10^{10} cm^{-3} at T=300K. However, the variation of n_i with temperature is illustrated in Figure 1.21. The carrier concentration is equal to zero at *T=0K*. When temperature is raised an increasing number of electron gather sufficient thermal energy to leave the semiconductor atoms and become free to move in the conduction band. These electrons are called "free electrons". Since they can move in the crystal they can contribute to an electrical current. An equal number of "free holes" can move in the crystal and contribute to an electrical current as well.

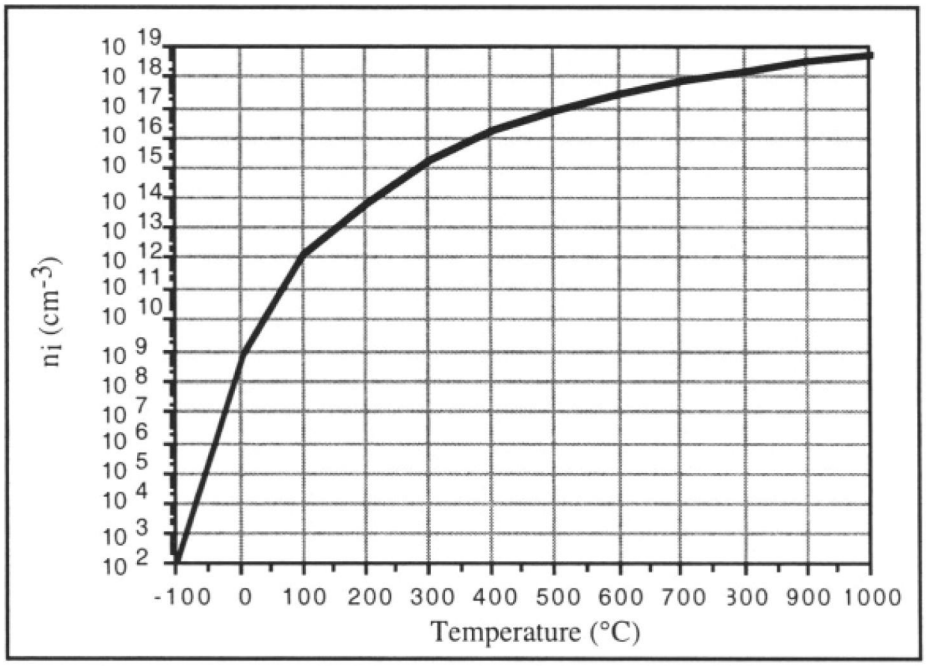

Figure 1.21: Evolution with temperature of the intrinsic carrier concentration, n_i, in silicon.

The conductivity of a material directly depends on the number of free carriers it contains (free electrons and free holes): the larger the number of carriers, the higher the conductivity. Thus, the conductivity of an intrinsic semiconductor increases with temperature (Figure 1.21).

Using equations 1.1.54 and 1.1.55 the intrinsic carrier concentration car be calculated:

$$n_i = N_c \exp\left[-\frac{E_c-E_i}{kT}\right] = N_v \exp\left[-\frac{E_F-E_v}{kT}\right] \qquad (1.2.5)$$

1.3. Extrinsic semiconductor

The silicon used in the semiconductor industry has a purity level of 99.9999999%. One can, however, intentionally introduce in silicon trace amounts of elements which are close to silicon in the periodic table, such as those located in columns III (boron) or V (phosphorus, arsenic). If, for instance, an atom of arsenic is substituted for a silicon atom, it will form four bonds by sharing four electrons with the neighboring silicon atoms (Figure 1.22). The thermal energy of the crystal at room temperature is large enough to remove the loosely held fifth electron from the arsenic's outer electronic shell, such that this electron will now reside in the conduction band where it is free to move in the crystal. Arsenic atoms in silicon are called *donor* atoms because each of these atoms "donates" an electron to the crystal. The free electron can contribute to electrical conduction.

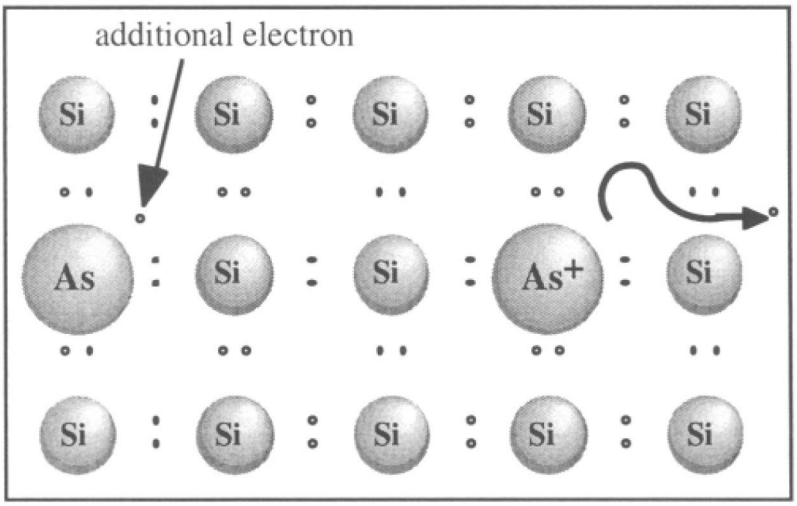

Figure 1.22: Donor impurity (arsenic in silicon). An arsenic atom introduces an extra electron in the crystal (left). An electron is released by an arsenic atom: the electron moves freely in the crystal and the arsenic atom carries a fixed positive charge (right). Note that while free electrons can move in the crystal, dopant atoms cannot.

Similarly, substituting a silicon atom with an atom from the third column of the periodic table, such as boron, will result in a missing electron (Figure 1.23). The boron atom can easily capture an electron to form a fourth bond with silicon atoms, thereby creating an immobile negatively charged boron atom. This releases a hole in the crystal, located in the valence band. This hole can move about in the crystal, thereby participating in electrical conduction. Because in silicon group III atoms create a hole which can be "filled" with an electron, these atoms are called *acceptor* atoms. Such atoms are usually introduced into the semiconductor

in very small amounts (1 atom of boron per 10^6 atoms of silicon, for instance). We will see later that the introduction of even minute amounts of these impurities dramatically modify the electrical properties of a semiconductor. Atoms possessing the property of releasing or capturing electrons in a semiconductor are indiscriminately called *doping impurities, doping atoms,* or *dopants.*

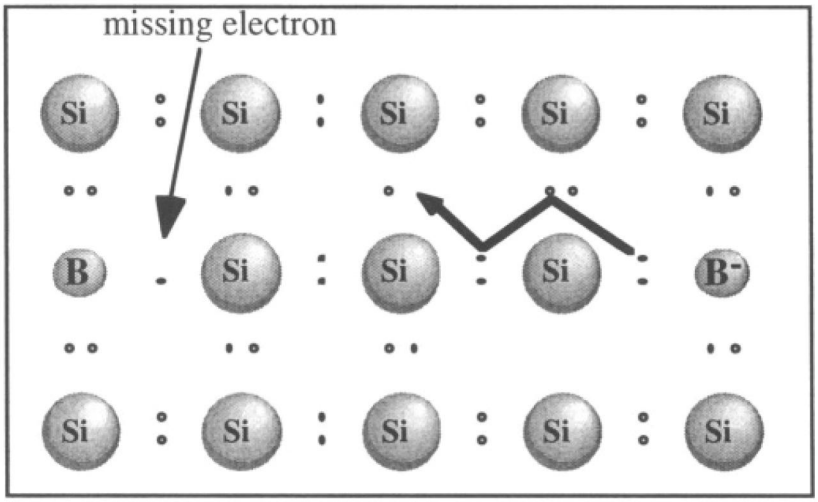

Figure 1.23: Acceptor impurity (boron in silicon). A boron atom introduces a missing electron in the crystal (left). A hole is released by a boron atom: the hole moves freely in the crystal and the boron atom carries a fixed negative charge (right). Note that while free holes can move in the crystal, dopant atoms cannot.

The introduction of a donor atom such as phosphorus (P) or arsenic (As) in silicon gives rise to a permitted energy level in the bandgap (E_d in Figure 1.24) . This level is located a few meV below the bottom of the conduction band, and at very low temperature contains the electrons which can be given by the impurity atoms to the crystal. At room temperature these electrons possess enough thermal energy (equal to kT/q = 25.6 meV) to break free from the impurity atoms and move freely in the crystal or, in other words, it can "jump" from the energy level E_d introduced by the impurity into the conduction band (Figure 1.24). When an electron moves away from a donor atom, such as arsenic (As), the atom becomes ionized (As^+) and carries a positive charge, $+q$, as shown in Figure 1.22.

Similarly, the introduction of an acceptor atom such as boron (B) in silicon gives rise to a permitted energy level in the bandgap. This level is located a few meV above the top of the valence band. At room temperature electrons in the top of the valence band possess enough thermal energy to "jump" into the energy levels created by the impurity atoms (or: valence electrons are "captured" by acceptor atoms), which

gives rise to holes in the valence band. These holes are free to move in the crystal. When an electron is captured by an acceptor atom, a hole is thus released in the crystal, and the acceptor atom (boron) becomes ionized (B^-) and carries a negative charge, $-q$, as shown in Figure 1.23.

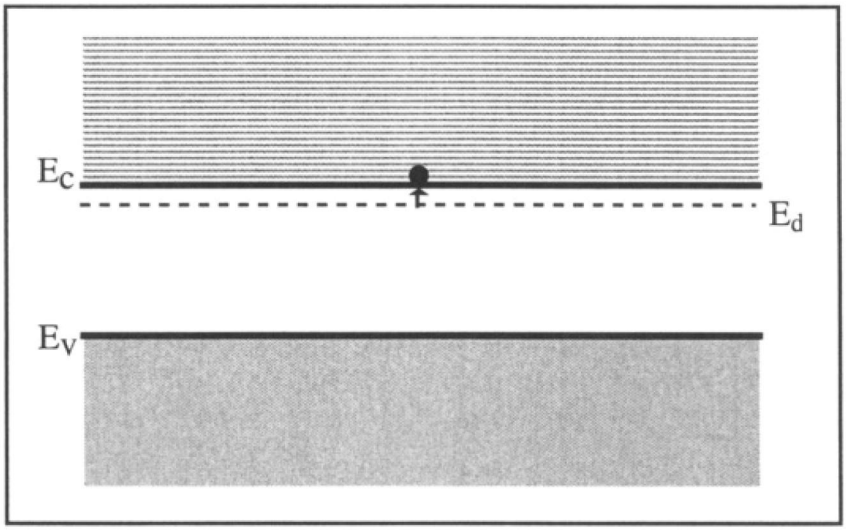

Figure 1.24: "Jump" of an electron from a donor level of energy E_d into the conduction band.

Donor and acceptor impurities are commonly introduced into semiconductors to increase electron or hole concentrations, which modifies the electrical properties of the material. The energy levels created in the bandgap by the presence of such impurities are situated close to the top of the valence band or the bottom of the conduction band. Other elements, such as gold, iron, copper and zinc introduce one or several energy levels in the bandgap of silicon. These levels are located closer to the center of the bandgap and are called "deep levels". The latter usually have a detrimental effect on semiconductors, which is why the semiconductor industry uses crystals having a very high degree of purity. The influence of deep levels on the properties of semiconductors will be discussed in Section 3.5, which is devoted to generation/recombination phenomena.

A semiconductor containing donor impurities is called an *N-type semiconductor,* since most of the carriers have a negative charge, and a semiconductor containing acceptor impurities is called a *P-type semiconductor,* since most of the carriers have a positive charge. The concentration of donor and acceptor atoms in the semiconductor are labeled N_d and N_a, respectively, and are expressed in atoms per cubic centimeters (cm^{-3}). Thus, an N-type semiconductor has more free electrons than holes, and vice-versa. However, the material itself is charge neutral due to the ionized impurities which carry a charge equal and opposite to that of the free carriers.

1.3.1. Ionization of impurity atoms

Whenever a donor (acceptor) impurity atom releases an electron (hole) it becomes ionized and carries a positive (negative) charge, $+q$ ($-q$). If a doping atom is not ionized, it does not release a free carrier in the crystal, and therefore, does not contribute to electrical conduction. Consider a donor impurity, such as arsenic in silicon. The ionization of the arsenic atom is a reversible process which can be written as:

$$As^0 \Leftrightarrow As^+ + e^- \tag{1.3.1}$$

where As^0 represents a non-ionized arsenic atom, and As^+ an ionized atom. Quite naturally the total impurity concentration is equal to the sum of the ionized and non-ionized impurity concentrations:

$$N_d = N_d^+ + N_d^0 \tag{1.3.2}$$

The probability of occupancy of the donor level, E_d, can be obtained by substituting E_d for E in the Fermi-Dirac distribution function. Previously (Equation 1.1.51), the Pauli exclusion principle was taken into account for determining the probability of filling energy states. In other words, each energy level could be populated with two electrons. In this case, however, an ionized arsenic atom can receive only one electron. A correction factor, called "degeneracy factor" equal to 1/2 must, therefore, be introduced in the Fermi-Dirac equation, which yields:

$$f(E_d) = \frac{1}{1 + \frac{1}{2}\exp[(E_d - E_F)/kT]} = \frac{N_d^0}{N_d} \tag{1.3.3}$$

The concentration of ionized donor atoms can be obtained using 1.3.2 and 1.3.3:

$$N_d^+ = N_d - N_d^0 = N_d\,(1 - f(E_d)) = N_d \frac{1}{1 + 2\exp\left(\frac{E_F - E_d}{kT}\right)} \tag{1.3.4}$$

The following example illustrate how one can determine how many donor atoms are ionized at room temperature.

Example:

Consider the following numerical example in silicon:

$E_d = E_C - 50$ meV

$E_F = (E_C + E_V)/2$ (assuming the doping concentration is very low)

kT/q = 0.0259V at room temperature (T=300K)

What is the ratio of ionized donor impurities to total impurities, N_d^+/N_d?

One finds readily that $E_g/2=0.56eV$ and $E_F - E_d = -0.5$ leV.

Therefore, using (1.3.4), $N_d^+/N_d=0.999999996$. Thus we can conclude from this example that at room temperature, virtually all donor atoms are ionized, or in mathematical terms, $N_d^+ \cong N_d$.

In the case of acceptor impurities (boron, for example), the reversible ionization reaction is:

$$B^0 + e^- \Leftrightarrow B^- \tag{1.3.5}$$

and we have:

$$N_a = N_a^- + N_a^0 \tag{1.3.6}$$

Using a calculation similar to that developed for donor atoms one finds:

$$f(E_a) = \frac{1}{1 + \frac{1}{2} exp[(E_F-E_a)/kT]} = \frac{N_a^0}{N_a} \tag{1.3.7}$$

and therefore, the probability of ionizing an acceptor is:

$$N_a^- = N_a - N_a^0 = N_a (1-f(E_a)) = N_a \frac{1}{1 + 2 exp\left(\frac{E_a-E_F}{kT}\right)} \tag{1.3.8}$$

At room temperature, virtually all acceptor atoms are ionized or, in mathematical terms, $N_a^+ \cong N_a$. Based on these derivations it is safe to assume that at room temperature every donor/acceptor atom contributes a free electron/hole to the semiconductor.

1.3.2. Electron-hole equilibrium

Consider a semiconductor crystal containing both N-type and P-type impurities. Because the crystal is charge neutral one can write:

Charge Neutrality under Thermodynamic Equilibrium

$$n + N_a^- = p + N_d^+ \tag{1.3.9a}$$

As we have seen in the previous Section all doping impurities are ionized at room temperature, therefore, $N_a^- = N_a$ and $N_d^+ = N_d$. Relationship 1.3.9a can thus be re-written in the following form:

$$n + N_a = p + N_d \Rightarrow n\text{-}p = N_d\text{-}N_a \qquad (1.3.9b)$$

Using elementary algebra one finds that $(p+n)^2 = (p-n)^2 + 4pn$. Relationship (1.3.9b) can be combined with $pn = n_i^2$ (Equation 1.2.1) to yield $(p+n)^2 = (N_d-N_a)^2 + 4n_i^2$. Since $(p+n)$ is a positive number one obtains:

$$p+n = \sqrt{(N_d-N_a)^2 + 4n_i^2} \qquad (1.3.10)$$

Combining 1.3.10 with Equation 1.3.9b one can write:

$$n = \frac{1}{2}\left[(N_d-N_a) + \sqrt{(N_d-N_a)^2 + 4n_i^2} \right] \qquad (1.3.11a)$$

and

$$p = \frac{1}{2}\left[(N_a-N_d) + \sqrt{(N_d-N_a)^2 + 4n_i^2} \right] \qquad (1.3.11b)$$

Using Relationships 1.3.11.a and 1.2.1 for an N-type semiconductor, where $N_d >> N_a$ and $N_d >> n_i$, we find that the electron and hole concentrations are given by:

Electron and Hole Concentration in N-type Semiconductor

$$n \cong N_d \quad and \quad p \cong \frac{n_i^2}{N_d} \qquad (1.3.12a)$$

Using Relationships 1.3.11.b and 1.2.1 for an P-type semiconductor, where $N_a >> N_d$ and $N_a >> n_i$, we find that the hole and electron concentrations are given by:

Electron and Hole Concentration in P-type Semiconductor

$$p \cong N_a \quad and \quad n \cong \frac{n_i^2}{N_a} \qquad (1.3.12b)$$

There are exceptions to Equations 1.3.12 a and b: at low temperatures not all impurities are ionized, and as a result, carrier freeze-out occurs: $n = N_d^+ < N_d$ and $p = N_a^- < N_a$. And at high temperature the intrinsic carrier concentration can become much larger than the concentration of carriers

released by doping impurities. In that case, $N_d << n = n_i = p_i = p >> N_a$ and the semiconductor is intrinsic even though it is doped. The influence of high and low temperatures on carrier concentration is illustrated by Problem 1.12.

1.3.3. Calculation of the Fermi Level

In the case of an N-type semiconductor, combining Relationships 1.1.54 and 1.3.12a yields:

$$n = N_d = N_c \exp\left[-\frac{E_c - E_F}{kT}\right] \tag{1.3.13a}$$

from which we find:

$$E_c - E_F = kT \ln\left(\frac{N_c}{N_d}\right) \tag{1.3.14a}$$

Using Expression (1.2.5):

$$n_i = N_c \exp\left[-\frac{E_c - E_i}{kT}\right] \Rightarrow E_i - E_c = kT \ln\left(\frac{n_i}{N_c}\right)$$

one finally obtains:

$$E_i - E_F = E_c - E_F + (E_i - E_c) = kT\left(\ln\left(\frac{N_c}{N_d}\right) + \ln\left(\frac{n_i}{N_c}\right)\right)$$

$$\Downarrow$$

$$E_F - E_i = kT \ln\left(\frac{N_d}{n_i}\right) \quad or \quad n = N_d = n_i \exp\left[\frac{E_F - E_i}{kT}\right] \tag{1.3.15a}$$

Hence the Fermi level, E_F, can be calculated from Equation 1.3.15a if the doping concentration is known. In an N-type semiconductor the Fermi level is located in the upper half of the bandgap, above the intrinsic energy level, E_i. The Fermi level increases logarithmically with the donor atom concentration, N_d. It is now possible to introduce a new variable, the Fermi potential, Φ_F (unit: volt). It is defined by the following relationship:

$$-q\Phi_F = E_F - E_i \tag{1.3.16}$$

Using Equation 1.3.15a the relationship between the electron concentration and the Fermi potential can be obtained:

Fermi Potential (N-Type Semiconductor)

$$n = n_i \exp\left[\frac{-q\Phi_F}{kT}\right] \quad \text{or} \quad \Phi_F = -\frac{kT}{q} \ln\left(\frac{n}{n_i}\right) = -\frac{kT}{q} \ln\left(\frac{Nd}{n_i}\right) \tag{1.3.17a}$$

For a P-type semiconductor equations 1.3.13a through 1.3.17a will use the same numbering system where the "a" is replaced by "b" in the equation. Combining Relationships 1.1.55 and 1.3.12b yields:

$$p = N_a = N_v \, exp\left[- \frac{E_F - E_v}{kT} \right] \tag{1.3.13b}$$

from which we find:

$$E_F - E_v = kT \, ln\left(\frac{N_v}{N_a}\right) \tag{1.3.14b}$$

Using Expression 1.2.5

$$n_i = N_v \, exp\left[- \frac{E_i - E_v}{kT} \right] \Rightarrow E_v - E_i = kT \, ln\left(\frac{n_i}{N_v}\right)$$

one finally obtains:

$$E_i - E_F = -(E_v - E_i) - (E_F - E_v) = - kT\left(ln\left(\frac{N_v}{N_a}\right) + ln\left(\frac{n_i}{N_v}\right)\right)$$

$$\Downarrow$$

$$E_i - E_F = kT \, ln\left(\frac{N_a}{n_i}\right) \quad or \quad p = N_a = n_i \, exp\left[\frac{E_i - E_F}{kT}\right] \tag{1.3.15b}$$

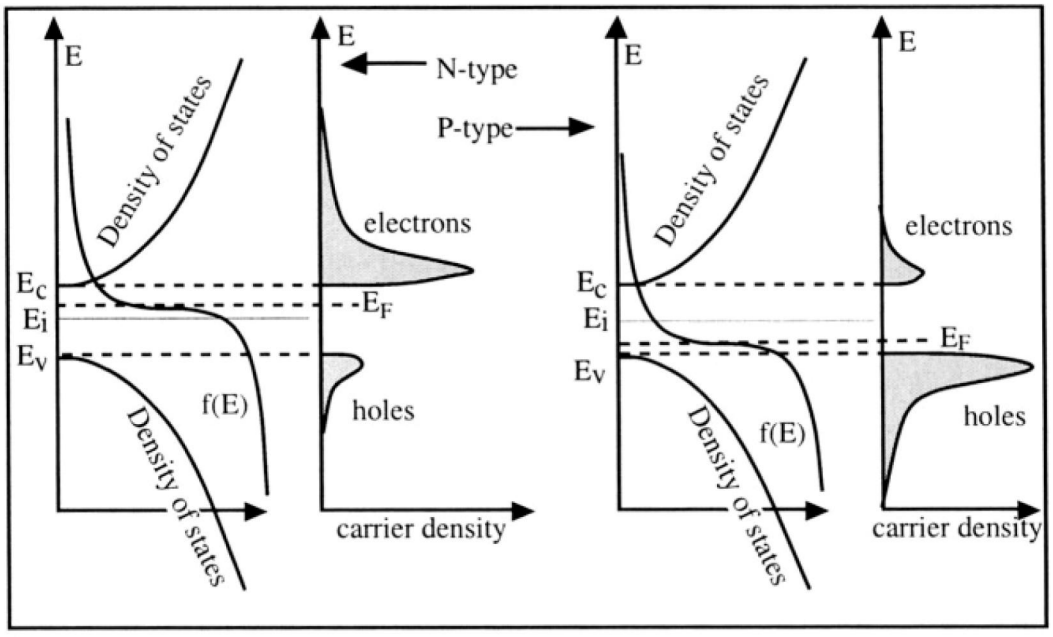

Figure 1.25: Location of the Fermi level, Fermi-Dirac distribution, f(E), and electron and hole concentration in an N-type and a P-type semiconductor. [13]

Equation 1.3.15b allows one to find the position of the Fermi level, E_F, in the bandgap. In a P-type semiconductor the Fermi level is located in the lower half of the bandgap, below the intrinsic energy level, E_i. The Fermi level decreases with increasing acceptor atom concentration, N_a.

Using Equation 1.1.16, the relationship between the Fermi potential, Φ_F, and the hole concentration can be obtained:

Fermi Potential (P-Type Semiconductor)

$$p = n_i \exp\left[\frac{q\Phi_F}{kT}\right] \quad \text{or} \quad \Phi_F = \frac{kT}{q} ln\left(\frac{p}{n_i}\right) = \frac{kT}{q} ln\left(\frac{N_a}{n_i}\right) \qquad (1.3.17b)$$

Note that Φ_F is positive in a P-type semiconductor and negative in an N-type semiconductor. A graphical representation of electron and hole concentrations for both N- and P-type semiconductors is shown in figure 1.25. Note the position of the Fermi level, E_F, and the asymmetry of carrier densities for both types.

1.3.4. Degenerate semiconductor

We have hitherto assumed that the introduction of doping impurities in a semiconductor does not affect certain intrinsic parameters of the crystal, such as the width of the energy bandgap. As we have seen before the presence of donor doping atoms such as phosphorus or arsenic introduces a permitted energy level, E_d, in the bandgap. Typical doping concentrations are in the 10^{15} to 10^{18} atoms/cm^3 range, which is small compared to the actual number of semiconductor atoms (5×10^{22} atoms/cm^3 in silicon).

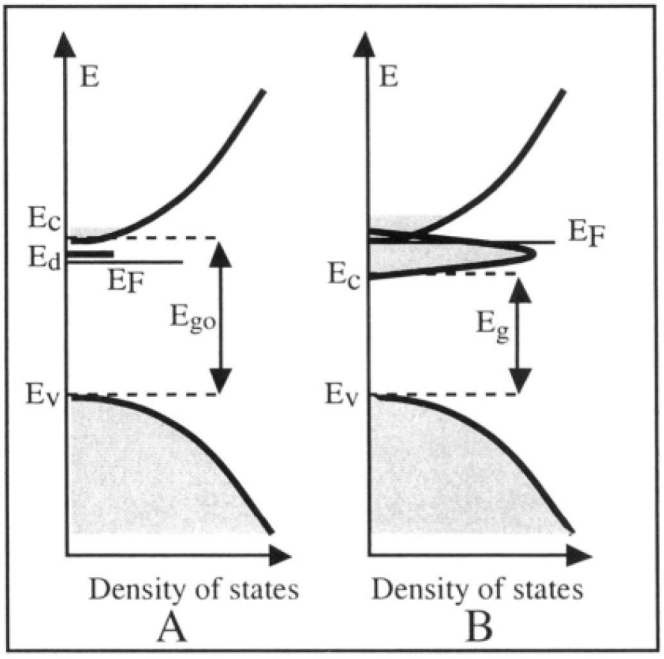

Figure 1.26: Density of states in a non-degenerate N-type semiconductor (A) and a degenerate N-type semiconductor (B). The gray areas correspond to states populated with electrons. [14]

If a very large concentration of impurities is introduced (e.g.: 10^{20} atoms/cm^3) the permitted level E_d spreads out and "degenerates" into a permitted band which overlaps with the conduction band. As a result the width of the bandgap is reduced (from E_{g0} to E_g in Figure 1.26) and the properties of the semiconductor are significantly modified. Such a semiconductor is called a "degenerate" semiconductor or a "degenerately doped" semiconductor. A degenerate semiconductor exhibits electrical properties similar to those of a metal.

1.4. Alignment of Fermi levels

Often, the doping concentration in a semiconductor is not one constant value throughout the material. Consider a piece of N-type semiconductor in which the doping concentration varies along one direction of space, *x*. The concentration of doping atoms is described by the function $N_d(x)$ shown in Figure 1.27.A.

Consider now that leftmost and rightmost parts of the sample are separated. According to Relationship 1.3.15a, $E_F(right) > E_F(left)$ because $N_d(right) > N_d(left)$ (Figure 1.27.B). Imagine a test energy level in the bandgap having an energy, E_T, located between $E_F(right)$ and $E_F(left)$. In the left part of the sample the test level has a low probability of being populated with an electron, because $E_T > E_F$. In the right part of the sample, on the other hand, the test level has a high probability of being populated with an electron, because $E_T < E_F$.

Let us now consider the entire sample, and in particular, focus on the middle region where the doping concentration changes abruptly. If the energy bands near $x=x_0$ stay as they are in the leftmost and rightmost parts of the sample, the test level E_T will have both a high and a low probability of being occupied by an electron, which is a contradiction in itself. The test level must have a single occupation probability. This condition can be satisfied only if E_F at the immediate left of x_0 is equal to E_F at the immediate right of x_0. And since this condition must be true for any arbitrary position along the x-axis, the Fermi level must be unique and constant throughout the sample. This is a very important property of the Fermi level, which can be enunciated the following way: *at thermodynamic equilibrium the Fermi level in a structure is unique and constant.* This property not only applies to non-homogeneously doped semiconductors, but to metal-semiconductor structures and contacts between different semiconductors. Because E_F is constant the conduction, valence, and intrinsic levels bend within a transition region around x_0 (Figure 1.27.C).

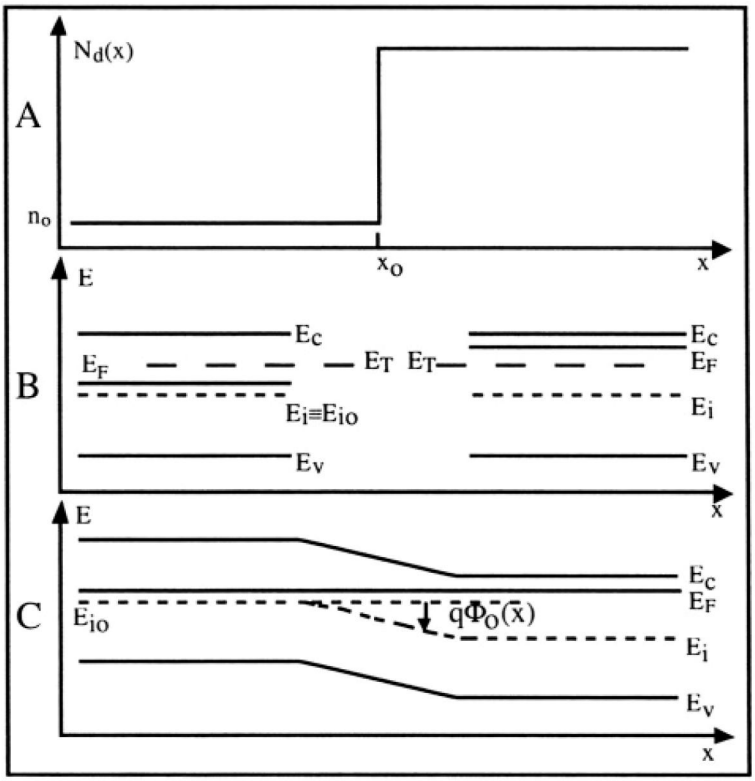

Figure 1.27: A: Inhomogenous doping profile.
B: Energy levels at the left and the right of the sample.
C: Band diagram in the complete structure.

Under thermodynamic equilibrium conditions electrons are transferred from the electron-rich right part of the sample (where the Fermi level is highest) into the electron-poor left part of the sample (where the Fermi level is lowest), through a diffusion process which will be discussed in Chapter 2. To make a comparison with fluid mechanics the alignment of the Fermi levels in the sample is similar to the alignment of the water levels in glasses of water connected together (Figure 1.28), where the transfer of electrons by a diffusion mechanism would find its equivalent in the transfer of water molecules due to a pressure differential. The diffusion process (electron transfer or water transfer) ceases when an equilibrium state is reached.

Since Relationships 1.3.13 a and b to 1.3.15 a and b are valid at any location along the x-axis, a constant Fermi level imposes a curvature of all energy bands and energy levels, E_c, E_v and E_i. However, all these levels remain parallel to one other, due to the fact that the bandgap energy is a constant of the material. The magnitude of this energy level bending reflects the presence of an internal potential, noted $\Phi_o(x)$ which, once multiplied by *-q,* is equal to the variation of the energy levels E_c, E_v and E_i between the left and the right of the sample (Figure 1.27.C). The internal potential is a real electrical potential variation due to the

appearance of an electric field in the semiconductor caused by the charge imbalance resulting from the diffusion of electrons from one part of the semiconductor to the other when thermodynamic equilibrium is established. Since the electron concentration is related to E_F-E_i by Relationship 1.3.15a, one can write:

$$n(x) = n_i \exp\left[\frac{E_F\text{-}E_i(x)}{kT}\right] \tag{1.3.18}$$

or, using the notations of Figure 1.27:

$$n(x) = n_i \exp\left[\frac{E_F\text{-}E_{io} + q\Phi_o(x)}{kT}\right] \tag{1.3.19}$$

or:

Boltzmann Relationship for Electrons

$$n(x) = n_i \exp\left[\frac{E_F\text{-}E_{io} + q\Phi_o(x)}{kT}\right] = n_o \exp\left[\frac{q\Phi_o(x)}{kT}\right] \tag{1.3.20a}$$

where n_o is the electron concentration in the left region of the sample, taken as reference. E_{io} is the midgap energy in the left part of the sample, also taken as reference. It is easy to show that an equivalent relationship can be derived for holes:

Boltzmann Relationship for Holes

$$p(x) = n_i \exp\left[-\frac{E_F\text{-}E_{io} + q\Phi_o(x)}{kT}\right] = p_o \exp\left[\frac{-q\Phi_o(x)}{kT}\right] \tag{1.3.20b}$$

Relationships 1.3.20a and 1.3.20b are called the "Boltzmann relationships". They will play an important role in the theory of the PN junction (Chapter 4).

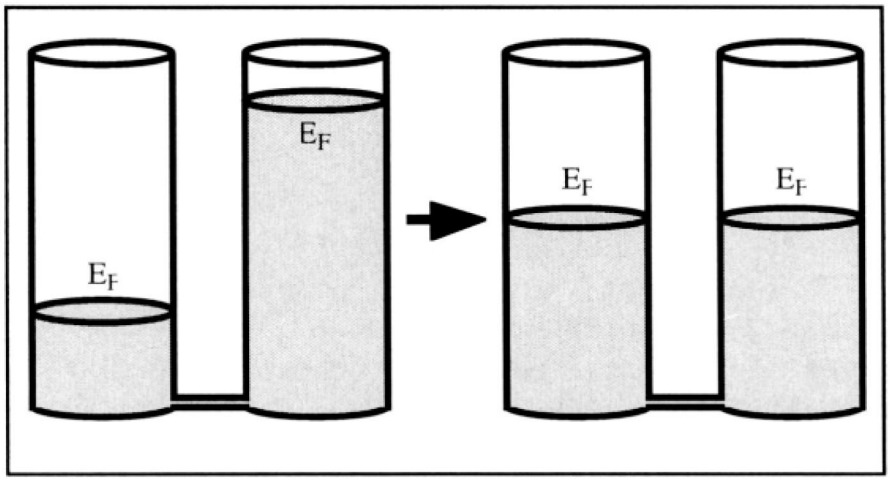

Figure 1.28: Bernoulli principle in fluids.

Important Equations

Fermi-Dirac Distribution

$$f(k) = \frac{1}{1 + exp[(E_n(k)-E_F)/kT]} \qquad (1.1.44a)$$

or

$$f(E) = \frac{1}{1 + exp[(E-E_F)/kT]} \qquad (1.1.44b)$$

Maxwell-Boltzmann Distribution

$$f(E) = \frac{1}{1 + exp[(E-E_F)/kT]} \cong exp\left[-\frac{E-E_F}{kT}\right] \qquad (1.1.51)$$

Fermi-Dirac *Maxwell-Boltzmann*

Free Carrier Concentration

$$n=N_C\, exp\left[-\frac{E_C-E_F}{kT}\right] \quad with\ N_C = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \qquad (1.1.54)$$

$$p=N_V\, exp\left[-\frac{E_F-E_V}{kT}\right] \quad with\ N_V = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \qquad (1.1.55)$$

pn Product under Thermodynamic Equilibrium

$$pn = n_i^2 \qquad (1.2.1b)$$

Intrinsic Carrier Concentration:

$$n_i = N_C\, exp\left[-\frac{E_C-E_i}{kT}\right] = N_V\, exp\left[-\frac{E_F-E_V}{kT}\right] \qquad (1.2.5)$$

Charge Neutrality under Thermodynamic Equilibrium

$$n + N_a^- = p + N_d^+ \qquad (1.3.9a)$$