6.5. The State Density of Electrons and Holes

In order to determine the electron concentration

$$n = \int_0^{\infty} g(E) f(E) dE. \tag{6.18}$$

Besides the electron distribution function, it is necessary to know the density of states *g(E).*

$$g(E) = \frac{dN}{dE}. \tag{6.19}$$

The quantity dN is the number of states in the interval from E up E+dE.
Let dV_k be the volume the surfaces corresponding the E and $E + dE$.

$$g(k) = \frac{dN}{d\tau_k}, \tag{6.20}$$

The quantity $g\ (k)$ is the state density (in unit volume) in k–space.

$$g(E) = \frac{dN}{dE} = \frac{g(k) dV_k}{dE} \tag{6.21}$$

$d\tau_k$ – an elementary volume of k-space equal to $dk_x dk_y dk_z$.
It is known that the quantization condition can be written as follows:

$$k_i = \frac{2\pi n_i}{L_i}; \quad n_i = 0, \pm 1, \pm 2, \ldots$$

The number of allowed states in an elementary interval of k-space:

$$dn_i = \frac{L_i}{2\pi} dk_i; \tag{6.22}$$

$$dN_i = dn_x dn_y dn_z = \frac{L_x L_y L_z}{8\pi^3} dk_x dk_y dk_z = \frac{V}{8\pi^3} d\tau_k, \tag{6.23}$$

V is the crystal volume. The number of states per unit volume of crystal per unit volume of k–space :

$$g(k) = \frac{1}{V} \frac{dN_i}{d\tau_k} = \frac{1}{8\pi^3}. \tag{6.24}$$

That quantity must be multiplied by the factor of two in order take into consideration two projections of spin (+1/2 and –1/2).
The dispersion law in small vicinity of E_c

$$E(k) = E_c + \frac{\hbar^2 k^2}{2m_n^*}, \tag{6.25}$$

m_n^* - the electron effective mass (scalar quantity). That dispersion law holds in conduction zone of certain compounds of $A^{III}B^{V}$ and $A^{II}B^{VI}$ type. The energy surface in k – space is a sphere of the radius $a = \sqrt{\frac{1}{\hbar} 2m_n^*(E - E_c)}$. Its volume:

$$V_k = \frac{4}{3}\pi a^3 = \frac{4}{3}\pi \left[\frac{2m_n^*(E-E_c)}{\hbar^2}\right]^{3/2}. \tag{6.26}$$

The increment dV_k in interval dE:

$$dV_k = \frac{\partial V_k}{\partial E} dE = 2\pi \left(\frac{2m_n^*}{\hbar^2}\right)^{3/2} (E - E_c)^{1/2} dE. \tag{6.27}$$

The number of states dN in the volume dV_k :

$$dN = g(k) dV_k = \frac{1}{8\pi^3} 2\pi \left(\frac{2m_n^*}{\hbar^2}\right)^{3/2} (E - E_c)^{1/2} dE; \tag{6.28}$$

$$g(E) = \frac{dN}{dE} = 2\pi \left(\frac{2m_n^*}{h^2}\right)^{3/2} (E - E_c)^{1/2}, \tag{6.29}$$

$h=2\pi\hbar$.

For a simple spherical valence zone with extreme E_v at the point $k = 0$, the dispersion law in the small vicinity of that point can be written as follows:

$$E(k) = E_v + \frac{\hbar^2 k^2}{2m_p^*}, \tag{6.30}$$

For simple zone, analogue speculations lead to:

$$g(E) = 2\pi \left(\frac{2m_p^*}{h^2}\right)^{3/2} (E_v - E)^{1/2}. \tag{6.31}$$

6.6 The Electron Concentration in the Conduction Zone and the Concentration of Holes in the Valence Zone

The concentration of electrons in the conduction zone

$$n = \int_{E_{c\min}}^{E_{c\max}} g(E) f(E) dE. \tag{6.32}$$

When the energy is greater than that of Fermi, the function $f(E)$ tends rapidly to zero. Introducing g(E) (6.29) in (6.32) and taking into account the spin factor 2 we get:

$$n = 2\int_{E_c}^{\infty} \frac{2\pi\left(\frac{2m_{dn}^*}{h^2}\right)^{3/2} (E - E_c)^{1/2} dE}{e^{\frac{E-E_F}{k_BT}} + 1}. \tag{6.33}$$

Introduce new quantities:

$$\frac{E - E_c}{k_BT} = \varepsilon; \quad \frac{E_F - E_c}{k_BT} = \eta. \tag{6.33a}$$

The quantities ε and η are expressed in terms of k_BT. The electron energy in the condution zone and the Fermi energy are measured from the bottom of the conduction zone E_c. Thus:

$$n = 2\left(\frac{2\pi m_{dn}^* k_BT}{h^2}\right)^{3/2} \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon-\eta} + 1} = N_c F_{1/2}(\eta), \tag{6.34}$$

$$N_c = 2\left(\frac{2\pi m_{dn}^* k_BT}{h^2}\right)^{3/2}, \tag{6.35}$$

The function $F_{1/2}(\eta)$ is called the **Fermi integral** (of order of 1/2):

$$F_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon^{1/2} d\varepsilon}{e^{\varepsilon-\eta} + 1}. \tag{6.36}$$

Generally the Fermi integral is a function:

$$F_{1/2}(\eta) = \frac{2}{\Gamma(i + 1)} \int_0^{\infty} \frac{\varepsilon^{j} d\varepsilon}{e^{\varepsilon-\eta} + 1}, \tag{6.37}$$

$\Gamma\ (i + 1)$ is the Eiler gamma-function.

The Fermi integrals can not be calculated analytically. The numerical calculations are tabulated. The asymptotic evaluation for integrals (of the order 1/2):

$$F_{1/2}(\eta) = \begin{cases} e^{\varepsilon} & for \quad -\infty < \eta < \text{-}1 \quad (a) \\ \dfrac{e^{\eta}}{0.27e^{\eta} + 1} & for \quad -\infty < \eta < 5 \quad (b) \\ \dfrac{4}{3\sqrt{\pi}}\eta^{3/2} & for \quad 5 < \eta < \infty \quad (c) \end{cases} \tag{6.38}$$

Taking into account the limits in formula (6.38) and condition (6.33a) we can stay that formula (6.38a) is valid for not degenerated semiconductors and formula (6.38c) is valid for strongly degenerated semiconductors and metals. Formula (6.38b) is called Erenberg-Blackmore approximation and is used for treatment of not degenerated

semiconductors. That formula is more accurate than formula (6.38a) and can be of use when the degeneration is weak. In accordance with formula (6.34) and (6.36a) the electron concentration in the conduction zone for not degenerated semiconductor:

$$n = N_c e^{-\frac{E_c - E_F}{k_B T}} \tag{6.39}$$

Equation (6.35) can be written as follows:

$$N_c = 4.831 \cdot 10^{15} \left(\frac{m_{dn}^*}{m} \right)^{3/2} T^{3/2}. \tag{6.40}$$

At room temperature $N_c \approx 2{,}5 \cdot 10^{19}$ см$^{-3}$. The number of states depends on the effective mass and temperature. In semiconductors of small effective mass (for example, in certain narrow zone compounds of $A^{III}B^{V}$ и $A^{II}B^{VI}$), the number of states is not great. The filling of the conduction zone increases strongly with temperature and the semiconductor may become degenerated. When the degeneration is great (6.38c), the electron concentration in the conduction zone is:

$$n = \frac{4N_c}{3\sqrt{\pi}} \left(\frac{E_F - E_c}{k_B T} \right) = \frac{8\pi}{3} \left(\frac{2m_{dn}^*}{h^2} \right)^{3/2} \left(E_F - E_c \right)^{3/2}. \tag{6.41}$$

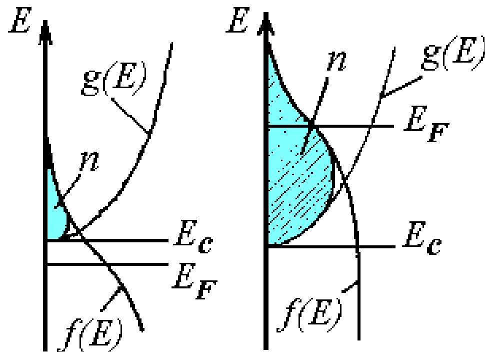

Fig.6.6 The electron concentration: a) non-degenerated semiconductor of n-type; b) degenerated semiconductor of p-type

When the degeneration is strong the concentration does not depend on temperature.

The electron concentration is shown in the Fig..6.6. When the semiconductor is not degenerated, the Fermi level is lower than the bottom of the conduction zone (at least by $k_B T$). While degenerating the Fermi level is higher at least of several $k_B T$. The dashed area of the conduction zone is proportional to the electron concentration. It is seen that the semiconductor is not degenerated when the concentration of electrons in the conduction zone is not great.

The concentration of holes in valence zone can be expressed analogously formula (6,18).

$$p = \int_{E_{v\min}}^{E_{v\max}} g(E) f_p(E) dE. \tag{6.42}$$

Introducing $f_p(E) = 1 - f_n(E)$ and $g(E)$ for the valence zone, changing the low limit of the integral by $-\infty$, and taking into account the rapid decrease of the distribution function $f_p(E)$ we get:

$$p = 4\pi\left(\frac{2m_{dp}^*}{h^2}\right)^{3/2} \int_{-\infty}^{E_v} \frac{(E_v - E)^{1/2}\, dE}{e^{\frac{E_F - E}{k_B T}} + 1}. \qquad (6.43)$$

Using the designations (6.33) and designating

$$\varepsilon = \frac{E - E_c}{k_B T}; \qquad (6.44)$$

$$\Delta\varepsilon = \frac{E_c - E_v}{k_B T}; \qquad (6.4a)$$

$$\varepsilon_p = \frac{E_v - E}{k_B T} = -\varepsilon - \Delta\varepsilon. \qquad (6.44b)$$

we arrive at

$$p = N_v \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\varepsilon_p^{1/2} d\varepsilon_p}{e^{(\varepsilon_p + \Delta\varepsilon + \eta)} + 1} = N_v F_{1/2}(-\Delta\varepsilon - \eta), \qquad (6.45)$$

$$N_v = 2\left(\frac{2\pi m_{dp}^* k_B T}{h^2}\right)^{3/2} -$$

is a number of states in the valence zone. The energy is measured in terms of $k_B T$

$$-\Delta\varepsilon - \eta = \frac{-E_c + E_v - E_F + E_c}{k_B T} = \frac{E_v - E_F}{k_B T} \qquad (6.46)$$

Equation (6.45) for the concentration of holes is analogue (6.34) for the concentration of electrons in the conduction zone.

The asymptotic forms of Fermi integral $F_{1/2}(-\Delta\varepsilon - \eta) = F_{1/2}\left(\frac{E_v - E_F}{k_B T}\right)$ are analogue the forms of the integral $F_{1/2}(\eta)$, [see (6.38)].

$$F_{1/2}(-\Delta\varepsilon - \eta) = \begin{cases} e^{-\Delta\varepsilon - \eta} & for \quad -\infty < (-\Delta\varepsilon - \eta) < \text{-1} \quad (a) \\ \dfrac{e^{-\Delta\varepsilon - \eta}}{0.27 e^{-\Delta\varepsilon - \eta} + 1} & for \quad -\infty < (-\Delta\varepsilon - \eta) < 5 \quad (b) \\ \dfrac{4}{3\sqrt{\pi}}(-\Delta\varepsilon - \eta)^{3/2} & \text{for} \quad 5 < (-\Delta\varepsilon - \eta) < \infty \quad (c) \end{cases} \qquad (6.47)$$

Thus, for non-degenerated semiconductor, the formula (6.47a) is true. The concentratin of the holes is:

$$p = N_v\, e^{-\Delta\varepsilon - \eta} = N_v e^{-\frac{E_v - E_F}{k_B T}}, \qquad (6.48)$$

For strongly degenerated semiconductoirs:

$$p = \frac{4N_v}{3\sqrt{\pi}}\left(\frac{E_v - E_F}{k_B T}\right)^{3/2} = \frac{8\pi}{3}\left(\frac{2m_{dp}^*}{h^2}\right)^{3/2}(E_v - E_F)^{3/2}. \qquad (6.49)$$

Analizing that equation we arive to the conclusion that by strong degeneration, the concentration of holes (Fig.6.6) becomes great and does not depend on temperature analogeu to the electron concentration (6.41) in conduction zone.

6.7 The Concentration of Electrons and Holes in an Intrinsic Semiconductor

For an intrinsic semiconductor, $n = p = n_i$. Taking into account equation (6.39) and (6.48) we get:

$$np = n_i^2 = N_v N_c e^{-\frac{E_v - E_c}{k_B T}} = N_v N_c e^{-\frac{E_g}{k_B T}} \qquad (6.50)$$

The intrinsic concentration (n_i) in a non-degenerated semiconductor:

$$n = p = n_i = \sqrt{N_v N_c}\, e^{-\frac{E_g}{2k_B T}} \qquad (6.51)$$

The concentration of electrons and holes in proper semiconductor dos not depend on the Fermi levels position and increases with energy in accordance with an exponential law. The factor 2 in the denominator of the exponent factor is due to the fact that the activation energy is needed to generate the *pair* of electric carriers (electron and hole).

The semiconductor as a whole is electrically neutral:

$$n = p \qquad (6.52)$$

For a non-degenerated semiconductor:

$$N_v e^{-\frac{E_c - E_F}{k_B T}} = N_v N_c e^{-\frac{E_F - E_c}{k_B T}} \qquad (6.53)$$

After simple transformation we get:

$$E_F = \frac{E_c + E_v}{2} + \frac{k_B T}{2}\ln\frac{N_v}{N_c} \qquad (6.54)$$

When the Fermi energy is measured from the ceiling of the valence zone:

$$E_F = E_v + \frac{E_g}{2} + \frac{k_B T}{2}\ln\frac{N_v}{N_c} \qquad (6.55)$$

Introducing the quantity N_c and N_v in (6.55) , we get:

$$E_F = E_v + \frac{E_g}{2} + \frac{3}{4} k_B T \ln \frac{m^*_{dp}}{m^*_{dn}} \tag{6.56}$$

A glance at equation (6.55) and (6.56) shows that at $T = 0$, the Fermi level in a intrinsic semiconductor is located at the middle of the forbidden zone. When the temperature increases the Fermi level moves to the zone where the number of states (hence the effective mass) is smaller.

The temperature dependence of the Fermi level in intrinsic non-degenerated semiconductor is shown in Fig.6.7.

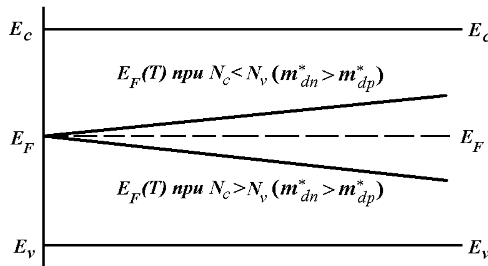

Fig.6.7. The Fermi level in an intrinsic semiconductor

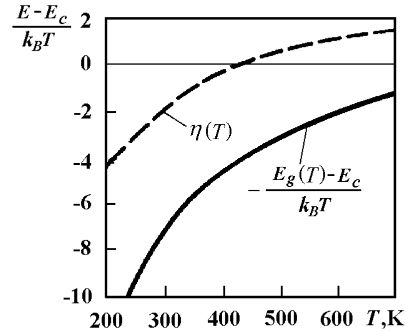

Fig.6.8. The temperature dependence of Fermi level and the width of forbidden zone in an intrinsic semiconductor

In semiconductors with the narrow zones, when the temperature increases the Fermi level can approach the bottom of the conduction zone (for compounds $A^{III}B^{V}$ and $A^{II}B^{IV}$ $m^*_{dn} < m^*_{dp}$) and the equation (6.55) and (6.56) will not be valid.

Degeneration in intrinsic semiconductor InSb when the temperature is increasing is shown in Fig.6.8. The energy is measured from the bottom of the conduction zone. When temperature increases the width of the forbidden zone follows the law $E_g \approx \left(0{,}26 - 2{,}7 \cdot 10^{-4} T\right)$eV. The Fermi level E_F approaches the bottom of the conduction zone (calculations are made for $m^*_{dp} = 10 m^*_{dn}$). At the temperature ~ 400K, the distance between the level and bottom is less then $k_B T$.

Analysis shows that for InSb at temperature T $\geq$ 440K, the Fermi level is inside of the conduction zone. The fast temperature degeneration in the region of the intrinsic conduction is due to the small width of the forbidden zone and the great difference between the effective masses of electrons and holes, which leads to increasing of Fermi level with temperature.

It should be noted that for the imaginary semiconductor with identical effective masses of electron and holes, there would not be the temperature dependence of Fermi level and it would be located at the middle of the forbidden zone.

6.8 The Width of the Forbidden Zone in a Intrinsic Semiconductor

In an intrinsic non-degenerated semiconductor the concentration of elections and holes is given by the equation (6.51) $n = p = n_i = \sqrt{N_v N_c}\, e^{-\frac{E_g}{2k_B T}}$. If the concentration of electrons (or holes) in the region of intrinsic conduction is known (for example by measuring with the Hall effect), the width of the forbidden zone of a semiconductor can be found. Indeed;:

$$\ln n_i = \ln\sqrt{N_v N_c} - \frac{E_g}{2k_B T}. \qquad (6.57)$$

The function $\ln(n_i) = f(1/T)$ is a straight line, which cuts a segment of the ordinate axis equal to $\ln\sqrt{N_v N_c}$. The inclination tangent of the line is $\frac{E_g}{2k_B}$ (fig.6.9). The width of the forbidden zone is given by:

$$E_g = -2k_B \mathrm{tg}\varphi \qquad (6.58)$$

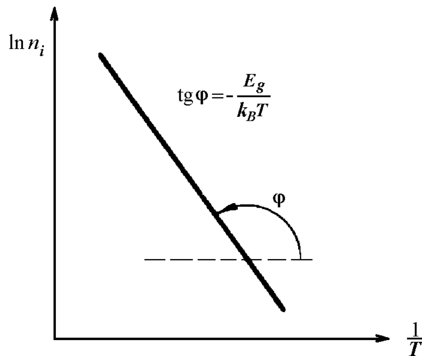

Fig.6.9. The charge carrier concentration in an intrinsic semiconductor as function of inverse temperature

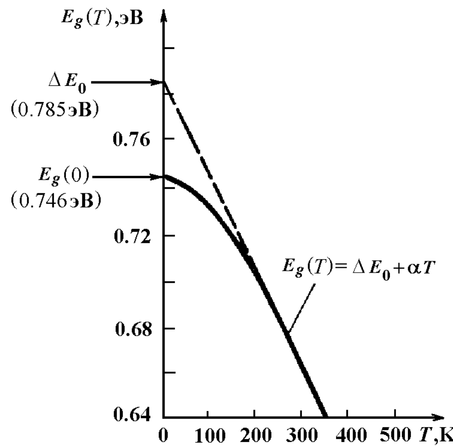

Fig.6.10. The temperature dependence of Eg in germanium and its linear extrapolation (dashed line)

In equation (6.57) we neglected the concentration temperature dependence.Let us take into consideration that dependence in the intrinsic semiconductor.

The quantities N_c *and* N_v depend on temperature as ~ $T^{3/2}$, the width of the forbidden zone depends on temperature too, $E_g = f(T)$. The last dependence is due to many static and dynamic factors and is very difficult for theoretical estimation. That dependence is found experimentally. For a great number of semiconductors, the width of the forbidden zone decreases with temperature. In a certain temperature interval that dependence is linear.

$$E_g(T) = \Delta E_0 + \alpha T, \qquad (6.59)$$

ΔE_0 is linear extrapolated width of the forbidden zone at zero temperature, α is a temperature coefficient (mostly negative).

The temperature dependence of germanium (optical measurements) and its linearization is shown in Fig.6.10. That dependence can be described as follows:

$$E_g(T) = (0{,}785 - 0{,}0004T) \qquad (6.60)$$

$\Delta E_0 = 0{,}785$ эВ и α = -4·10^{-4} eVdegree^{-1}. It follows from the figure that the quantity $\Delta E_0 = 0{,}785$ эВ is not the width of the forbidden zone of germanium at any temperature but just the segment of the ordinate axis cut at the zero temperature.

Let us make an assumption that the linear law (6.59) is true. Taking into account that

$$\sqrt{N_v N_c} = 4{,}831 \cdot 10^{15} \left(\frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} T^{3/2}$$

We can write:

$$n(T) = 4{,}831 \cdot 10^{15} \left(\frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} T^{3/2} e^{-\frac{\Delta E_0 - \alpha T}{2k_B T}} \qquad (6.61)$$

Taking the logarithms of (6.61) divided by $T^{3/2}$ we get:

$$\ln \frac{n}{T^{3/2}} = \ln \left[4{,}831 \cdot 10^{15} \left(\frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} \right] - \frac{\Delta E_0}{2k_B} \frac{1}{T} - \frac{\alpha}{2k_B}. \qquad (6.62)$$

The segment cut by the line $\ln(n) = f(1/T)$ is:

$$[\ln n]_0 = \ln \left[4{,}831 \cdot 10^{15} \left(\frac{m_{dp}^* m_{dn}^*}{m_0^2} \right)^{3/4} \right] - \frac{\alpha}{2k_B}, \qquad (6.63)$$

The inclination tangent:

$$\text{tg}\varphi = -\Delta E_0 / 2k_B \qquad (6.64)$$

Thus, $\Delta E_0 = -2k_B \text{tg}\varphi$. The width of the forbidden zone $E_g(T)$ for a temperature (not less than 200K) can be estimated. You are to use (6.63) at given quantities m_{dn}^* and m_{dp}^* to find the linear coefficient and put it in (6.59).

6.9 Statistics of Impurity States in Semiconductors

The donor and acceptor impurities change the charge carriers concentration and the Fermi level position. When analyzing the electric neutrality of impurity semiconductors we should take into account not only the positive and negative charge bound with the charge carriers in zones as it was done for a intrinsic semiconductor. The positive and negative charges due to ionized donor and acceptor atoms are to be taken into consideration.

The electric neutrality equation in a impurity semiconductor with donor and acceptor concentration N_d and N_a can be written as follows:

$$n + N_a^- = p + N_d^+, \tag{6.65}$$

N_a^- and N_d^+ are concentrations of ionized acceptors and donors.
When ionized the acceptors atoms attaches electrons:

$$N_a^- = n_a, \tag{6.66}$$

n_a is the concentration of electrons, which occupy the acceptor states in the forbidden zone. Analogues for the ionized donors:

$$N_d^+ = p_d, \tag{6.67}$$

p_d is the concentration of holes, which occupy the donor states. We remind our readers that the transition of electrons from donor levels into a conduction zone by ionization of donor atoms is equivalent the transition of holes from the conduction zone into the donor level. The equation (6.65) takes the form:

$$n + n_a = p + p_d. \tag{6.68}$$

Concentration of impurity atoms (not ionized):

$$N_d^0 = n_d; \tag{6.69}$$

$$N_a^0 = p_a, \tag{6.70}$$

N_d^0 and N_a^0 is concentration of non-ionized donor and acceptor atoms; n_d and p_a is concentration of electrons at donor levels and concentration of holes at acceptor levels. Obviously

$$N_d = N_d^0 + N_d^+; \tag{6.71}$$

$$N_a = N_a^- + N_a^0. \tag{6.72}$$

Then:

$$n_d + p_d = N_d; \tag{6.73}$$

$$n_a + p_a = N_a. \tag{6.74}$$

The electric neutrality condition can be written as follows:

$$n + n_d - p - p_a = N_d - N_a. \tag{6.75}$$

Thus, to solve the electric neutrality equation we must know the concentration of electrons and holes in impurity states. It should be noted that the statistics of impurity states differs from that one of allowed energy zones. Indeed, in accordance with exclusion principle at a energy level there can be two electrons (or holes) with the opposite spin. One electron (or hole) can occupy the one fold ionized impurity level.
The impurity sate system of electrons and holes is the system with **alternating number of particles.** In accordance with quantum statistics [11], the distribution function of electrons in impurity states is:

$$f_n = \frac{1}{\frac{1}{g} e^{\frac{E_t - E_F}{k_B T}} + 1}, \tag{6.76}$$

E_t is the impurity level, g –degeneration factor.
Taking into account he spin we can state that an energy level is twice degenerated. Thus factor g for electron donor states or acceptor hole states is 2. Factor g for holes of donor states or electrons of acceptor states can be evaluated if we remind that any energy state can be occupied either by electron or by a hole: $f_n + f_p = 1$.

$$f_n^d + f_p^d = 1; \tag{6.77}$$

$$f_n^a + f_p^a = 1 \tag{6.78}$$

Thus g-factor for holes on donor levels and electrons on acceptor levels is 1/2.

If the other mechanisms of degeneration are taken into consideration g-factor can be greater than 2. For example if the impurity state splits from an allowed zone of complex structure, the degeneration in extreme points of the allowed zone is transmitted to the impurity states. The degeneration factor of the acceptor states produced in semiconductors of the fourth group by introducing the impurity atoms of the third group is 4 (the extreme of the valence zone at **k**=0 is twice degenerated).

To find the quantities n_d, p_a, p_d, and n_a besides the energy distribution function, the impurity state density function is needed. The number of states with energy E_d per a unit volume of a crystal is equal to the donor atom concentration N_d. The number of states with energy E_a per a unit volume of a crystal is equal to the acceptor concentration N_a . Thus, we can write:

$$n_d = N_d f_n^d = \frac{N_d}{\frac{1}{2} e^{\frac{E_d - E_F}{k_B T}} + 1}; \tag{6.79}$$

$$p_a = N_a f_p^a = \frac{N_a}{\frac{1}{2} e^{\frac{E_F - E_a}{k_B T}} + 1}; \tag{6.80}$$

$$p_d = N_d - n_d = \frac{N_d}{2 e^{\frac{E_F - E_d}{k_B T}} + 1}; \tag{6.81}$$

$$n_a = N_a - p_a = \frac{N_a}{2 e^{\frac{E_a - E_F}{k_B T}} + 1}. \tag{6.82}$$

Introducing in equation (6.75) the quantities n, p, n_d, p_a, we get the general equation for finding the Fermi level in a impurity semiconductor:

$$N_c F_{1/2}\left(\frac{E_c - E_F}{k_B T}\right) + \frac{N_d}{2e^{\frac{E_d - E_F}{k_B T}} + 1} - N_v F_{1/2}\left(\frac{E_F - E_v}{k_B T}\right) + \\ + \frac{N_a}{2e^{\frac{E_a - E_F}{k_B T}} + 1} = N_d - N_a. \tag{6.83}$$

For a non-degenerated semiconductor, we can use the asymptotic interpretation of the Fermi integral:

$$F_{1/2}\left(-\frac{E_c - E_F}{k_B T}\right) = e^{-\frac{E_c - E_F}{k_B T}}; \tag{6.84}$$

$$F_{1/2}\left(-\frac{E_F - E_v}{k_B T}\right) = e^{-\frac{E_F - E_v}{k_B T}}. \tag{6.85}$$

The mathematics analysis of equation (6.83) is rather simplified because the Fermi integrals are substituted by exponents [see (6.38a) and (6.47a)], but remains all the same complicated, which is due to two types of impurity. So let us investigate the electric neutrality equation for non-degenerated semiconductor with a single type of impurity.

6.10 Concentration of Electrons in Semiconductor with a Single Type of Impurity

The electric neutrality equation (6.75) takes a form:

$$n + n_d - p = N_d. \tag{6.86}$$

Let us rewrite equation (6.86) in the form:

$$n = p - p_d. \tag{6.87}$$

The meaning of the last equation is obvious. In conduction zone the free electrons are generated by intrinsic ionization (simultaneously in valence zone, equivalent number of holes is generated) and impurity ionization (simultaneously at donor levels, the same number of bound holes is left). When the temperature is increasing at first the impurity atoms are ionized, at higher temperature the intrinsic ionization will begin. Because $\Delta E_d << \Delta E$, we can not to take into consideration the intrinsic ionization at low temperature. Thus we can write:

$$n = p_d. \tag{6.88}$$

For non-degenerated semiconductor:

$$n = N_c e^{-\frac{E_c - E_F}{k_B T}} = \frac{N_d}{g e^{\frac{E_F - E_a}{k_B T}} + 1}. \tag{6.89}$$

Introducing $\frac{E_c - E_d}{k_B T} = \varepsilon_d$, we get:

$$n = \frac{N_d}{g e^{\varepsilon_d + \eta} + 1}. \tag{6.90}$$

Substituting e^{η} by n/N_c in accordance with equations (6.38a) and (3.39), we get the square equati0n relative n:

$$n\left[g\frac{n}{N_c}e^{\varepsilon_d} + 1\right] = N_d. \tag{6.91}$$

Its solution:

$$n = \frac{1}{2}g^{-1}N_c e^{-\varepsilon_d}\left\{\left[1 + 4gN_d N_c^{-1}e^{\varepsilon_d}\right]^{1/2} - 1\right\} \tag{6.92}$$

Taking into account (6.90) we can write equation (6.92) as follows:

$$n = \frac{2N_d}{1 + \sqrt{1 + \frac{4gN_d}{N_c}e^{\varepsilon_d}}}. \tag{6.93}$$

If the following condition holds:

$$\frac{4gN_d}{N_c}e^{\varepsilon_d} >> 1,$$

$$n = \sqrt{\frac{N_c N_d}{g}}e^{-\frac{\Delta E_d}{2k_B T}}. \tag{6.94}$$

The condition (6.94) signifies that at low temperature in the conduction zone, the number of electrons is small in comparison with N_c. A glance at equation (6.94) shows that in the region of impurity ionization, the energy of heat activation of ionization process is $\Delta E_d/2$. Hence at low temperature, the Fermi level is in the middle between E_c and E_d. Indeed introducing (6.94) in (6.37), we get:

$$N_c e^{-\frac{E_c - E_F}{k_B T}} = \sqrt{\frac{N_c N_d}{g}}e^{-\frac{E_c - E_d}{2k_B T}}; \tag{6.95}$$

$$e^{\frac{E_F}{k_B T}} = \sqrt{\frac{N_d}{gN_c}}e^{\frac{E_c + E_d}{2k_B T}}; \tag{6.96}$$

$$E_F = \frac{E_c + E_d}{2} + \frac{k_B T}{2}\ln\frac{N_d}{gN_c}; \tag{6.97}$$

$$E_F = E_c - \frac{\Delta E_d}{2} + \frac{k_B T}{2}\ln\frac{N_d}{gN_c}. \tag{6.98}$$

The Fermi level at zero temperature coincides with the middle of the interval between E_c and E_d. When the temperature increases the Fermi level approaches E_c (while $N_c < N_d/g$), passes through maximum and removes from E_c (when $N_c > N_d/g$).

At higher temperature [see (6.93)] if

$$\frac{4gN_d}{N_c}e^{\varepsilon_d} << 1, \tag{6.99}$$

Then:

$$n = N_d. \qquad (6.100)$$

In that temperature interval, the electron concentration does not depend on temperature. It corresponds to the region of exhausting. Introducing (6.100) in (6.37) we get:

$$E_F - E_c = k_B T \ln \frac{n}{N_c} = k_B T \ln \frac{N_d}{N_c}; \qquad (6.101)$$

$$E_F = E_c - k_B T \ln \frac{N_d}{N_c}; \qquad (6.102)$$

The criteria (6.99) leads to producing a donor impurity in the exhausting region. The electron concentration becomes constant. At high impurity N_d or great ionization energy ΔE_d that criteria holds only rather high temperature when the process of the intrinsic ionization is strong. The exhausting donor impurity region is no more seen and the temperature interval with constant electron concentration is absent.

Discussing the electric neutrality equation (6.87) at low temperature we assumed that the concentration of holes in valence zone is small in comparison with p_d. At higher temperature, the intrinsic ionization is greater then the impurity one: $p >> p_d$..The equation (6.87) transforms into equation (6.52) from which the temperature dependence (6.51) and (6.54) follows.

$$\ln n = \ln \sqrt{\frac{N_c N_d}{2}} - \frac{\Delta E_d}{2k_B T}; \qquad (6.103)$$

$$\Delta E_d = -2k_B tg\varphi_1. \qquad (6.104)$$

That definition of ΔE_d can lead to a certain mistake if $\Delta E_d > 2k_B/T$. The hidden temperature dependence of N_c.should be taken into account.

The experimental dependence should be built in coordinates $\ln n/T^{3/4} = f(1/T)$. The tangent of the inclination angle more accurately defines ΔE_d [see (6.104)] the analysis in the coordinates $\ln n = f(1/T)$. The temperature interval of the exhausted region becomes narrower and is not seen at great concentration.

The impurity ionization region continuously transforms into region of the intrinsic ionization. The inclination of the segment I of the dependence $\ln n = f(1/T)$ does not change with impurity concentration.

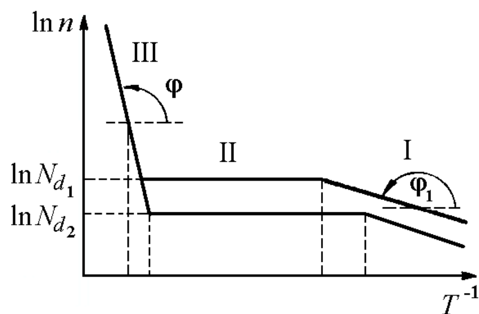

Fig.6.11. The electron concentration as function of inverse temperature for two concentration of the donor impurity: I-the region where the impurity conduction increases; II-the region of exhausting; III-the region of the intrinsic conduction

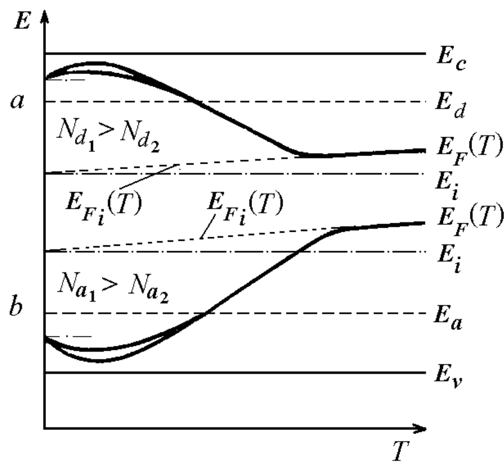

Fig.6.12. The temperature dependence of Fermi level of a semiconductor with a single type of the donor impurity (a) and a single type of the acceptor impurity (b) for two different concentration Dashed line - the Fermi level of a non-alloyed semiconductor

The temperature dependence of Fermi level for the semiconductors with the donor impurity of a single type is shown in Fig.6.12. When the donor concentration increases, the Fermi level approaches the conduction zone which is due to to the increasing of electron concentration and growing population of the conduction zone. A glance at Fig.6.11 shows that when the impurity concentration increases the exhausted region appears at higher temperature.

At very great impurity concentration, the quantity ΔE_d decreases for a number of semiconductors. All speculations concerning the temperature dependence of concentration and the Fermi level of a donor semiconductor can be applied to an acceptor semiconductor with impurity of one type.

The function $\ln p = f(1/T)$ is analogues that one in Fig.6.11 (substituting ΔE_d by ΔE_a and N_d by N_a. The temperature dependence of the Fermi level of an acceptor semiconductor with a single acceptor is not symmetric relative the level E_i. The temperature dependence $E_F(T)$ for concentration $N_{a_1} > N_{a_2}$ is shown in Fig.6.12b.

6.11 The Temperature Dependence of the Charge Carrier Concentration of Real Semiconductors

In real semiconductors there are simultaneously donor and acceptor impurities. With the special rectification methods we can diminish the concentration of unwonted impurity, but it is impossible to exclude it totally.

The phenomenon of **impurity compensation** is observed in a semiconductor if there are in it at the same time two kinds of impurities. The behavior of the semiconductor in the region of impurity electric conduction in comparison with that one of a semiconductor of one type of impurity changes.

The phenomenon of reciprocal compensation of donors and acceptors is as follows. If the acceptor energy states E_a in the forbidden zone are lower than the donor

energy states E_d (such situation as a rule is for alloying impurities) then it is unnecessary to transmit the energy $\Delta E_d = E_c - E_d$. to ionize the donor atoms. The empty electron level of neutral acceptors E_a is lower then populated electron level of neutral donors E_d Thus at zero temperature, the electrons transmit from the donor levels to the lower acceptor levels.

Obviously if $N_d > N_a$, all acceptors will be ionized; a certain number of donors will not be ionized. The concentration of non-ionized donors $N_d' = N_d - N_a$ would act as effective donor concentration. In the process of the heat impurity ionization only those N_d' donors can give off electrons in the conduction zone.

Acceptors being totally ionized can not give holes in the valence zone. It is convenient to say about total compensation of acceptors and partial compensation of donors. Such semiconductor is called the **partially compensated donor semiconductor.** When $N_a > N_d$ the donors are totally compensated. The acceptors would be compensated partially. Such semiconductor is called the **partially compensated acceptor semiconductor.**

When $N_d = N_a$ all the donors and acceptors are compensated. Such semiconductor is called the **totally compensated semiconductor.** The thermal ionization of impurities in a totally compensated semiconductor when the impurity atoms get energy ΔE_d or ΔE_a is impossible.

The concentration of the free charges carriers i.e. the electrons in the conduction zone and the holes in the valence zone may increase. That increasing is due to to intrinsic ionization or thermal transitions of electrons from the filled acceptor levels into the conduction zone or from the valence zone into the free donor levels.

For alloyed impurities $\Delta E_{a,d} << E_g$. The last two processes are like the process of intrinsic ionization (the energy magnitude). Thus, the concentration of the free electric carriers of the totally compensated and intrinsic semiconductors is very close at those temperatures. But the mobility of charge carriers in the totally compensated semiconductor is considerably less then that one of the intrinsic semiconductor. The additional scattering of the carriers from charged ions of donors and acceptors causes it. Hence, the higher specific resistance of a semiconductor is not the evidence of its higher cleaning from impurities.

Let us discus the impurity compensation phenomenon using the simplest model of a semiconductor of a single type donors of concentration N_d , the energy levels in forbidden zone E_d, a single type of acceptors with concentration $N_{a,}$ and energy position of levels E_a.

Assume that $N_d > N_a$ (partly compensated donor semiconductor). In equation (6.75), the quantity $p_a = 0$, because the acceptors are totally compensated and there are no neutral acceptor atoms. Chose the temperature interval where the intrinsic ionization is small i.e. the impurity conduction region. We can not to take into account the concentration of holes and assume that $p < n$, n_d, N_d , N_a. Equation (6,74) can be written as follows:

$$n + n_d = N_d - N_a \tag{6.105}$$

or having in mind that $N_d - n_d = p_d$

$$n + N_a = p_d \tag{6.106}$$

Le us analyze that equation for a non-degenerated semiconductor when the formula (6.39) for the electron concentration is true. Taking the quantity p_d from (6.81), and substituting the factor two by more general degeneration factor g. we get:

$$N_c e^{-\frac{E_c-E_F}{k_BT}} + N_a = \frac{N_d}{ge^{\frac{E_F-E_d}{k_BT}}+1} \qquad (6.107)$$

Introducing non-dimensional quantity $\eta = \frac{E_c-E_F}{k_BT}$ we get:

$$\eta = \ln\left\{\left[2g^{-1}(N_d-N_a)\right]\Big/\left[g^{-1}N_c+N_a e^{\varepsilon_d}\right]+\left[\left(g^{-1}N_c+N_a e^{\varepsilon_d}\right)+ +4g^{-1}N_c(N_d-N_a)e^{\varepsilon_d}\right]^{1/2}\right\}. \qquad (6.108)$$

That expression is very complex. We will discus it in certain temperature diapasons. The quantity η can be found from $\eta = \ln n/N_c$ using (6.39). In accordance with equation (6.107):

$$n = 2(N_d-N_a)\Big/\left\{\left[1+(gN_a/N_c)e^{\varepsilon_d}\right]+\left\{\left[1+(gN_a/N_c)e^{\varepsilon_d}\right]^2+ +(4g/N_c)(N_d-N_a)e^{\varepsilon_d}\right\}^{1/2}\right\}. \qquad (6.109)$$

When the concentrations of compensate centers N_a tends to zero formula (6.109) transforms into (6.93), which describes the electron concentration in a pure donor semiconductor. Let us compare the expression (6.93) and (6.109). At high temperature when all impurities are ionized according (6.93) $n = N_d$ (exhausting of donor impurity), according (6.109) $n = N_d - N_a = N_d'$ (exhausting of effective ionization of donors). While lowering the temperature the factor $(4g/N_c)(N_d-N_a)e^{\varepsilon_d}$ and $(gN_a/N_c)e^{\varepsilon_d}$ in equation (6.109) at first are of the unit order and then become greater then the unit. If the degree of compensation is very small ($N_a \ll N_d$) there is such temperature interval where $N_a \ll n \ll N_d$. Thus:

$$(gN_a/N_c)e^{\varepsilon_d} \ll 1 \ll \left[4g(N_d-N_a)/N_c\right]e^{\varepsilon_d}. \qquad (6.110)$$

The electron concentration temperature dependence in that temperature interval is described by:

$$n = \left[g^{-1}N_c(N_d-N_a)e^{-\varepsilon_d}\right]^{1/2} \quad for \quad N_a \ll n \ll N_d \qquad (6.111)$$

Returning to dimensional quantities:

$$n = \sqrt{\frac{N_c(N_d-N_a)}{g}}\, e^{-\frac{\Delta E_d}{2k_BT}} \quad for \quad N_a \ll n \ll N_d \qquad (6.112)$$

That expression is analogues to (6.94) for the electron concentration in intrinsic donor semiconductor when the quantity N_d is substituted by N_d - N_a.

At further lowering of temperature, the electron concentration becomes close to N_a and even less. In denominator of (6.109) the both factors of the type

$\left(gN_a/N_c\right)e^{\varepsilon_d}$ dominant and at low temperature:

$$n=\left[g^{-1}N_c\left(N_d-N_a\right)/N_a\right]e^{-\varepsilon_d} \qquad (6.113)$$

On common terms:

$$n=\frac{N_c\left(N_d-N_a\right)}{gN_a}e^{-\frac{\Delta E_d}{2k_BT}};\; n<<N_a<<N_d \qquad (6.114)$$

Comparing equation (6.114) and (6.112) we see that at low temperature for a compensated semiconductor, the logarithms of the electron density as a function of the inverse temperature changes two times faster. Indeed in

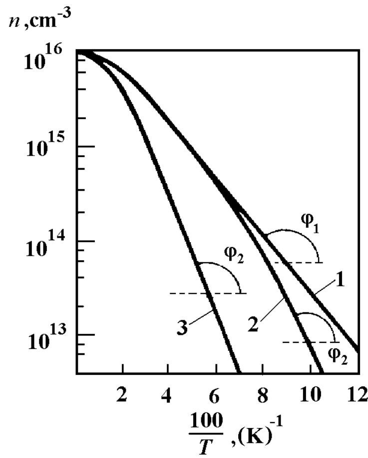

Fig.6.13. The temperature dependence of the electron concentration of a donor semiconductor

$N_d-N_a=10^{16}$cм$^{-3}$, $\Delta E_d=0.01$eV; $m_c^*=0.25m$; $y=2$; 1-$N_a=0$; 2- $N_a=10^{14}$cм$^{-3}$; 3 - $N_a=10^{16}$cм$^{-3}$

the region described by (6.114) the tangent of inclination angle of the linear segment of the logarithms function is $-\Delta E_d/k_B$, but according to (6.112) it is $-\Delta E_d/2k_B$.

In the transitional region [from (6.114) to (6.112)] $n\approx N_a$. At higher compensation degree i.e. at higher quantities of N_a, the temperature region described by (6.112) disappears.

Now let us discuss the region of the donor impurity concentration region where $n=N_d-N_a$ does not depend on temperature. The temperature dependence of electron concentration for different donor impurity compensation at the same donor concentration $N_d'=N_d-N_a$ is shown in Fig.6.13.

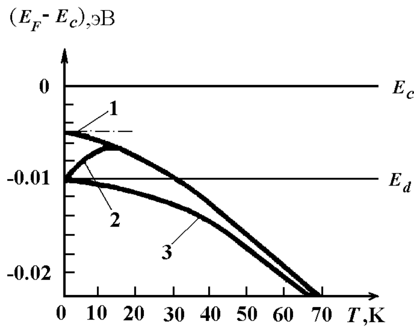

Fig.6.14. The temperature dependence of Fermi level of a donor semiconductor with different degree of donor compensation

The temperature dependence of the Fermi level for the region of the impurity electric conduction at concentrations, which correspond to those ones of Fig.6.13, is shown in Fig.6.14. The curves 1 of Fig. 6.13 and 6.14 correspond to intrinsic donor semiconductor. The inclination in the region of the impurity ionization [see Fig.6.13] is described by equation (6.104).

The curve 2 and 3 describe the partly compensated semiconductor. The compensation degree of a sample described by the curve 2 is not great. The region of impurity ionization (Fig.6.13) can be subdivided in two parts with different slopes: (low temperature) $\mathrm{tg}\varphi_2=-\Delta E_d/k_B$ and $\mathrm{tg}\varphi_1=-\Delta E_d/2k_B$ (higher temperature).

At sufficiently high compensation degree (curve 3) the region with the slope

$-\Delta E_d/2k_B$ disappears. The curves 2 and 3 (Fig.6.14) at low temperature are described by:

$$E_F = E_c - \Delta E_d - \ln\frac{gN_a}{N_d - N_a}, \qquad (6.115)$$

Introducing (6.114) in (6.37) we have got equation (6.115).

A glance at formula (6.115) shows that when $N_a > \frac{N_d}{1+g}$, the Fermi level is lowering from E_d (the curve 3). When $N_a < \frac{N_d}{1+g}$, the Fermi level is higher then E_d (the curve 2).

Even very small addition of the compensate impurity makes the Fermi level at zero temperature to transmit from E_c - $\Delta E_d/2$ to E_c - ΔE_d.. Obviously, the behavior of carriers of pertly compensated semiconductor in the intrinsic conduction region is analogue to that one of the intrinsic donor semiconductor with a single type of donors.

The analysis for partly compensated donor semiconductor is analogues. You are only to substitute the quantities E_d and N_d by E_a and N_a; and the qantities $(N_d - N_a)$ and N_d by $(N_a- N_d)$ and N_a . Of course it must be taken into account that the g-factor is different.

Discussing the energy diagram of the impurity semiconductor we assumed that donors occupied only one level in the forbidden zone E_d, and acceptors - E_a. Besides the basic states E_d and (E_a), there are the excited states. For example for shallow hydrogen impurity states there are excited states with the energy ratio of $1/n^2$ (n – the principal quantum number).

At higher temperature the electrons populate the excited impurity states and change the concentration of free charge carriers in the conduction zone. It is due to the fact that the activation energy and g-factor is different at higher temperature.

The transitions from the impurity excited states into conduction zone can change the concentration temperature dependence in the region described by (6.114), when the temperature is greater and the great number of donor electrons populate the excited states.

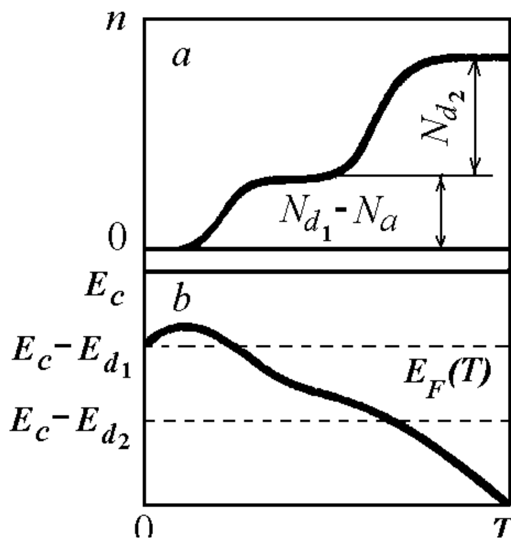

Fig.6.15. the temperature dependence of electron concentration (a) and the Fermi level (b) of a donor semiconductor with two types of partly compensated donor impurity

The imput of excited states in the temperature dependence of the charge carriers concenration in the region of impurity conduction leads to decreasing of the free carriers concentration. If it is not taken into account, the systematic error would be done when finding ΔE_d and g – factor using the function $n(T)$.

Ionization of the donor levels, which are close to E_d., can lead to violating of $n(T)$ given by (6.112) and (6.114) in the region of impurity electric conduction

If the donor levels are sufficiently separated in the forbidden zone, there would be horizontal segments of the temperature dependence n(T) in the region

of the impurity electric conduction (Fig.6.15). When the transitions from the level E_{d_1} in the conduction zone stop, the transitions from the E_{d_2} begin. The imput of the transitions can be easily evaluated. If the levels are too close that is impossible.

6.12 Degenerated Impurity Semiconductors

Till now we used the equations of non-degenerated semiconductor. We have established that the Fermi level when the donor concentration of a intrinsic semiconductor N_d or the effective donor concentration N_d - N_a in compensated semiconductor is increasing approaches the bottom of the conduction zone.

At certain impurity concentration the criteria of a non-degenerated semiconductor E_c - $E_F >> k_BT$ can be violated and the semiconductor becomes degenerated. Then while solving the electric neutrality equation, the expression (6.34) and (6.45) should be used. The analysis show that the Fermi level when the concentration increases can approach the bottom of the conduction zone closer then k_BT (even enter the conduction zone).

According to (6.97) the quantity E_F (T) reaches a maximum $E_F(T_{\max}) = E_{F\max}$ at certain temperature $T_{\max.}$ Differentiating equation (6.97) with respect to temperature we find the extreme condition:

$$\frac{dE_F}{dT} = \frac{k_B}{2}\ln\frac{N_d}{gN_c} - \frac{k_BT}{2}\cdot\frac{N_d}{gN_c}\cdot\frac{2}{N_d}\cdot\frac{dN_c}{dT} = 0, \qquad (6.116)$$

Hence:

$$\ln\frac{N_d}{gN_c} - \frac{T}{N_c}\frac{dN_c}{dT} = 0, \qquad (6.117)$$

Taking into account that

$$\frac{dN_c}{dT} = \frac{3}{2}\frac{N_c}{T}, \qquad (6.118)$$

We get:

$$\ln\frac{N_d}{gN_c} + \frac{3}{2}N_{cT_{\max}} = \frac{N_d}{ge^{3/2}}, \qquad (6.119)$$

Having in mind formula (6.40) we arrive to:

$$T_{\max} = 8.15\left(\frac{m}{m_{dn}^*}\right)\left(\frac{N_d}{10^{18}}\right)^{2/3}. \qquad (6.120)$$

The temperature $T_{\max}$ increases with impurity concentration.

Introducing (6.120) in (6.97) (the Fermi level temperature dependence) we get:

$$E_{F\max} = \frac{E_c + E_d}{2} + \frac{k_BT_{\max}}{2}\ln\frac{N_d}{gN_c(T_{\max})} + \frac{E_c + E_d}{2} + $$
$$+\frac{3}{4}k_BT_{\max} = \frac{E_c + E_d}{2} + 5.3\cdot 10^{-4}\left(\frac{m}{m_{dn}^*}\right)\left(\frac{N_d}{10^{18}}\right)^{2/3}. \qquad (6.121)$$

The energy is expressed in electron volts. It is clear that $E_{F\max}$ depends on impurity concentration and effective mass for state density analogues $T_{\max.}$ The quantities $E_{F\max}$ and $T_{\max}$ increases with impurity concentration N_d and decreasing of the effective electron mass m_{dn}^* in the conduction zone

A impurity concentration when the Fermi level coincides with the conduction zone bottom E_c is called the **critical concentration of degeneration** $N_{d.cr.}$. Let us evaluate the error when calculating the critical concentration of degeneration by (6.121). We remind our readers that formula (6.121) is deduced for a non-degenerated semiconductor. Calculation shows that the error is about 20%.

According (6.121):

$$E_{F\max} = E_c = \frac{E_c + E_d}{2} + 5.3 \cdot 10^{-4} \left(\frac{m}{m_{dn}^*} \right) \left(\frac{N_{d.cr.}}{10^{18}} \right)^{2/3}. \qquad (6.122)$$

Hence with accuracy of 20%:

$$N_{d.cr.}\left(\text{cm}^{-3}\right) = 10^{22.5} \left(\frac{m_{dn}^*}{m} \right)^{3/2} \left[\Delta E_d \left(\text{eV}\right)\right]^{3/2}. \qquad (6.123)$$

It follows from (6.123) that critical degeneration concentration depends on effective mass and ionization energy of an impurity. For $m_{dn}^* = 3m$ and ΔE_d = 0,03eV $N_{d(\text{крит})} \approx 2 \cdot 10^{19}\,\text{cm}^{-3}$. When the quantity m_{dn}^* decreases to 0,001m and ΔE_d to 0,0001eV, the critical concentration is about 10^{12}cm^{-3}.

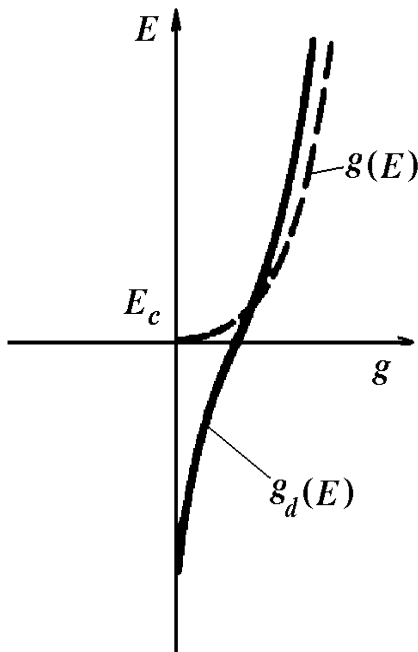

Fig.6.16. The state distribution function of the intrinsic semiconductor *g(E)* and the strongly alloyed one *g_d(E)*.

The dependence of the degeneration concentration on the effective mass of charge carriers can be explained. Indeed, the concentration of the free carriers increases with temperature. If the effective mass for the state density is small, the density of states in the conduction zone is small too. The zone with small density of states is filling rapidly by the free electrons and the degeneration becomes at a rather small impurity concentration.

The prominent Russian physicists V.L.Bonch-Bruevitch solved the three-dimensional problem of impurity chaotic distribution for extreme concentrations using the Green functions. It was shown that at high impurity concentration, the state density is not a zero in entire forbidden zone. Near the Fermi level the distribution is like that one of the ideal Fermi-gas. In the vicinity of extreme points of the allowed zones, that function it is rapidly decreasing with the depth of the forbidden zone.

The formation of the 'tail' of the function $g_d(E)$ at strong alloying is shown in Fig.6.16. The

distribution function $g(E)$ of non-alloyed semiconductor is shown by the dashed line.

In the depth of the forbidden zone, the function $g_d(E)$ strongly depends on the chosen potential of interaction of the electrons and impurity ions. The voltage-current characteristic of a tunnel diode allows to find the form of the 'tail' of the distribution function.