Lecture note on **Physics of Semiconductors (3)**

18th April (2021)     Shingo Katsumoto, Institute for Solid State Physics, University of Tokyo

### 2.3.1 Effective mass approximation

Let us consider the effect of spatially non-uniform perturbation. For that we add the perturbation potential $U(\boldsymbol{r})$ to the Schrödinger equation in crystals to obtain

$$\left[-\frac{\hbar^2\nabla^2}{2m}+V(\boldsymbol{r})+U(\boldsymbol{r})\right]\zeta(\boldsymbol{r})=[\hat{H}_0+U(\boldsymbol{r})]\zeta(\boldsymbol{r})=E\zeta(\boldsymbol{r}). \tag{2.90}$$

$\zeta(\boldsymbol{r})$ can be expanded with the Bloch functions $\psi_{n\boldsymbol{k}}$, which are the eigenstates of $\hat{H}_0$ as

$$\zeta(\boldsymbol{r})=\sum_{n,\boldsymbol{k}}f(n,\boldsymbol{k})\psi_{n\boldsymbol{k}}(\boldsymbol{r})=\sum_{n,\boldsymbol{k}}f(n,\boldsymbol{k})u_{n\boldsymbol{k}}(\boldsymbol{r})e^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{2.91}$$

With taking the inner product with $\psi_{n'\boldsymbol{k}'}$ after substitution of (2.91) to (2.90),

$$[E_0(n',\boldsymbol{k}')-E]f(n',\boldsymbol{k}')+\sum_{n,\boldsymbol{k}}\langle n',\boldsymbol{k}'|U|n,\boldsymbol{k}\rangle f(n,\boldsymbol{k})=0, \tag{2.92}$$

where $\psi_{n\boldsymbol{k}}$ is written as $|n,\boldsymbol{k}\rangle$. The second term, the transition mediated by $U$ (we write it as $U_{n'\boldsymbol{k}',n\boldsymbol{k}}$), represents the scattering from $|n,\boldsymbol{k}\rangle$ to $|n',\boldsymbol{k}'\rangle$. $U$ and $u^*_{n'\boldsymbol{k}'}u_{n\boldsymbol{k}}$ are Fourier transformed into

$$U(\boldsymbol{r})=\int d\boldsymbol{q}U_{\boldsymbol{q}}e^{-i\boldsymbol{q}\cdot\boldsymbol{r}},\quad u_{n'\boldsymbol{k}'}(\boldsymbol{r})u_{n\boldsymbol{k}}(\boldsymbol{r})=\sum_{\boldsymbol{G}}b_{n'\boldsymbol{k}'n\boldsymbol{k}}(\boldsymbol{G})e^{i\boldsymbol{G}\cdot\boldsymbol{r}}.$$

The transformation of $u^*_{n'\boldsymbol{k}'}u_{n\boldsymbol{k}}$ is a Fourier series on the reciprocal lattice because the term has the lattice periodicity. The coefficients $b_{n'\boldsymbol{k}'n\boldsymbol{k}}$ are written with the unit cell space $\Omega_0$, the unit cell volume $v_0$ as

$$b_{n'\boldsymbol{k}'n\boldsymbol{k}}(\boldsymbol{G})=\int_{\Omega_0}\frac{d\boldsymbol{r}}{v_0}e^{-i\boldsymbol{G}\cdot\boldsymbol{r}}u^*_{n'\boldsymbol{k}'}(\boldsymbol{r})u_{n\boldsymbol{k}}(\boldsymbol{r}).$$

$$\therefore U_{n'\boldsymbol{k}',n\boldsymbol{k}}=\int d\boldsymbol{q}U_{\boldsymbol{q}}\sum_{\boldsymbol{G}}b_{n'\boldsymbol{k}'n\boldsymbol{k}}(\boldsymbol{G})\int d\boldsymbol{r}e^{i(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{q}+\boldsymbol{G})\cdot\boldsymbol{r}}.$$

The last integral can be performed to be $(2\pi)^3\delta(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{q}+\boldsymbol{G})$, and the integration over $\boldsymbol{q}$ gives

$$U_{n'\boldsymbol{k}',n\boldsymbol{k}}=(2\pi)^3\sum_{\boldsymbol{G}}U_{\boldsymbol{k}'-\boldsymbol{k}-\boldsymbol{G}}\,b_{n'\boldsymbol{k}'n\boldsymbol{k}}(\boldsymbol{G}). \tag{2.93}$$

$U(\boldsymbol{r})$ is assumed to have much slower spatial variation than the lattice potential. Then as $U_{\boldsymbol{q}}$, it is enough to restrict ourseleved to $|q|\ll\pi/a$, *i.e.* much smaller values than that of the Brillouin zone edge. The approximation corresponds to $\boldsymbol{k}'-\boldsymbol{k}\sim\boldsymbol{G}$. We further assume that $U$ does not cause strong scattering that drives the state to the zone edge, then $\boldsymbol{G}$ takes only $\vec{0}$. Also $|U|$ is smaller than the band gap, then there is no interband scattering via $U$, *i.e.* there is no matrix element for $n\neq n'$. Then we can approximate

$$U_{n'\boldsymbol{k}',n\boldsymbol{k}}\approx U_{\boldsymbol{k}'-\boldsymbol{k}}\delta_{n'n}. \tag{2.94}$$

(2.92) is written as

$$[E_0(\boldsymbol{k}')-E]f(n,\boldsymbol{k}')+\sum_{\boldsymbol{k}}U_{\boldsymbol{k}'-\boldsymbol{k}}f(n,\boldsymbol{k})=0. \tag{2.95}$$

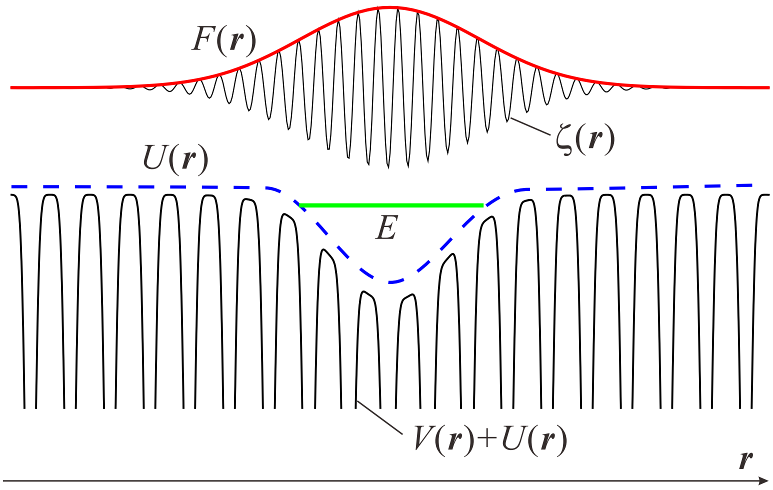


**Fig.** 2.16 A perturbation pontetial $U(\boldsymbol{r})$ is superposed on the crystal potential $V(\boldsymbol{r})$. The figure illustrates the potential and the wavefunction $\zeta(\boldsymbol{r})$, the envelope function $F(\boldsymbol{r})$ for the system with the crystal potential $V(\boldsymbol{r})$ the slowly varing perturbation potential $U(\boldsymbol{r})$.

Nest we consider the expansion in (2.91). In the present approximation, only the region $\boldsymbol{k} \sim 0$ is considered for $u_{n\boldsymbol{k}}$, and $u$ is almost constant for $\boldsymbol{k}(\approx u_{n0})$. Then we take it out from the sum over $\boldsymbol{k}$.

$$\zeta_n(\boldsymbol{r}) = u_{n0}\sum_{\boldsymbol{k}} f(n,\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = u_{n0}F_n(\boldsymbol{r}), \tag{2.96}$$

where the index $n$ is attached to $\zeta$ with ignoring the intermixing of the bands. Here, $F_n(\boldsymbol{r})$ defined as

Envelope function

$$F_n(\boldsymbol{r}) \equiv \sum_{\boldsymbol{k}} f(n,\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{2.97}$$

is the inverse Fourier transformation of $f(n,\boldsymbol{k})$, and called **envelope function**. $F_n(\boldsymbol{r})$ should be a slowly varing function over the scale of lattice constant(Fig. 2.16).

For the unperturbed dispersion relation, we apply that of a particle with the effective mass, and for simplicity the effective mass $m^*$ is assumed to be isotropic. Then $E_0(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m^*$ is substituted to (2.95) to give

$$\frac{\hbar^2 k\boldsymbol{k}'^2}{2m^*}f(\boldsymbol{k}) + \sum_{\boldsymbol{k}} U_{\boldsymbol{k}'-\boldsymbol{k}}f(\boldsymbol{k}) = Ef(\boldsymbol{k}'). \tag{2.98}$$

Here we omit writing $n$. The above can be inverse-Fourier transfomed with the care that the second term becomes convolution because it already has the summation over $\boldsymbol{k}$ to give

Effective mass equation

$$\left[\frac{\hbar^2\nabla^2}{2m^*} + U(\boldsymbol{r})\right]F(\boldsymbol{r}) = EF(\boldsymbol{r}). \tag{2.99}$$

The equation takes the form of the Schrödinger equation of the paticle with the mass $m^*$ and the potential $U(\boldsymbol{r})$. That is, for the envelop function, the problem is now a particle with the effective mass in the perturbation potential. This way of handling the problems on the level of envelope function is called **effective mass approximation**, and eq.(2.99) is called effective mass equation. In this sense, the Bloch states can be viewed as plane waves in the effective mass approximation.

The viewpoint is very usuful in designing various quantum systems in solids with semiconductor technologies. We test the approximation for the shallow impurity states in the next chapter. We also use it in many places in this lecture. We should be careful, however, that the envelope function is not the wavefunction itself. The difference becomes clear, particularly when the perturbation potential has a sharp spatial variation.

# Chaper 3 Carrier statistics and impurity doping

In this chapter we consider the energy distribution of **carriers** in semiconductors. We introduce the concept of carrier doping with very little amount of impurities, which brings drastic changes in the electric conduction.

## 3.1 Carrier statistics in intrinsic semiconductors

We call a pure semiconductor without any impurity as an **intrinsic semiconductor**. Of course this is just an idea, but *e.g.* non-doped Si for LSIs' can be considered as an intrinsic semiconductor. And under some conditions other semiconductors can also be treated as intrinsic semiconductors.

### 3.1.1 Density of states

We consider a simple lattice system which has a state per a unit cell with an edge length of $a$. We take the system size as $L = Na$ in one dimension. For an $n$-dimensional system, the volume $(2\pi/L)^n$ contains a single state in $k$-space(Fig. 3.1(a)). Given the kinetic energy as $E(k) = \hbar^2k^2/2m$, the number of states per volume between $E$ and $E + dE$ (Fig. 3.1(b)) devided by $dE$ is

$$\mathscr{D}(E) = \frac{1}{L^d}\left(\frac{L}{2\pi}\right)^d \frac{dV_d(k)}{dE} = \frac{1}{(2\pi)^d}\frac{dV_d(k)}{dk}\frac{dk}{dE} = \frac{1}{(2\pi)^d}\frac{m_0}{\hbar^2}\frac{dV_d(k)}{kdk}, \tag{3.1}$$

where $V_d(k)$ is the volume of $d$-dimensional sphere with the radius of $k$. This $\mathscr{D}(E)$ is called **energy density of state**. Because $V_1 = 2k,\ \ V_2 = \pi k^2,\ \ V_3 = 4\pi k^3/3$ (Fig. 3.2),

$$\boxed{\mathscr{D}^{(0)}_{d=1} = \frac{1}{\pi\hbar}\sqrt{\frac{2m_0}{E}}, \quad \mathscr{D}^{(0)}_{d=2} = \frac{m_0}{\pi\hbar^2}, \quad \mathscr{D}^{(0)}_{d=3} = \frac{\sqrt{2m_0^3}}{\pi^2\hbar^3}\sqrt{E},} \tag{3.2}$$

where the factor 2 comes from the freedom of spin.

In the case of electrons in crystals, the above expressions for density of states are applicable with replacing the mass with the effective mass where non-parabolicity of the band is ignorable, *e.g.*, around tops and bottoms of the bands. When we cannot apply the parabolic approximation, we need to go back to the definition of the density of states. For a three

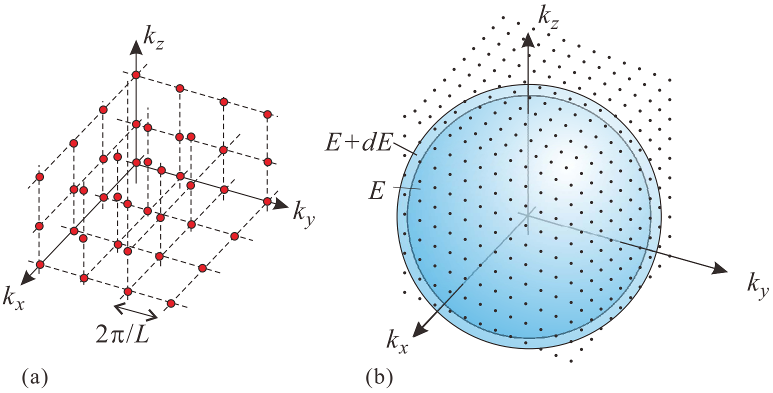


**Fig.** 3.1 (a) The red dots represent possible wavenumber in k-space in 3d empty lattice approximation for simple cubic. (b) Counts the number of dots in the spherical shell from $E$ to $E + dE$.

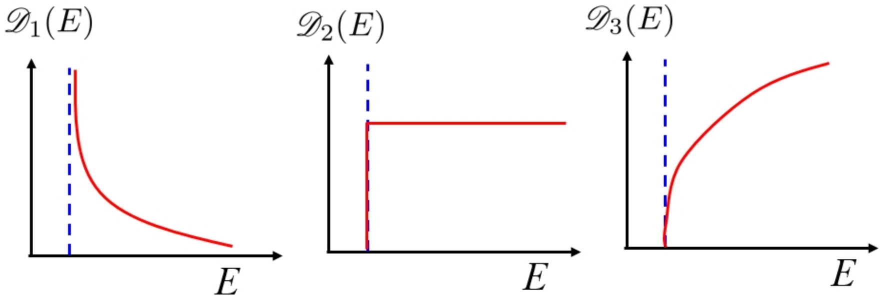


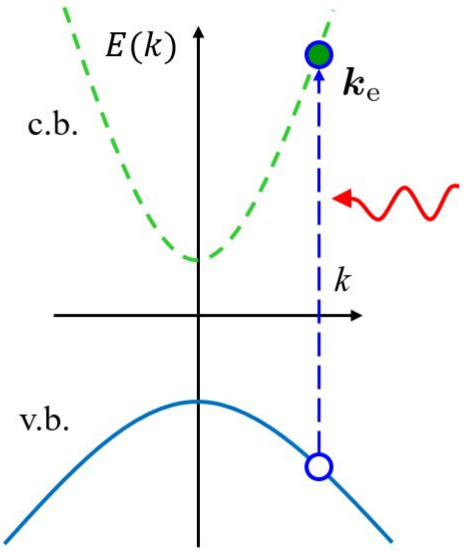
**Fig.** 3.2 Schematic diagrams of density of states for 1, 2, 3 dimentions in (3.2).

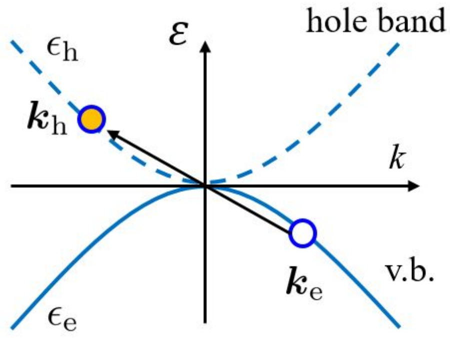
dimensional system it is given from

$$\mathscr{D}(E) = \int_{E(\boldsymbol{k})=E} \frac{dS_k}{(2\pi)^3} \frac{2}{\nabla_{\boldsymbol{k}} E(\boldsymbol{k})}. \tag{3.3}$$

The integral is over the equi-energy surface $E(\boldsymbol{k}) = E$ in $k$-space.

### 3.1.2 Concept of holes





The total current $\boldsymbol{J}_{\text{v.b.}}$ carried by a full valence band is zero by canceling of counter-going electrons ($\boldsymbol{J}_{\text{v.b.}} = \sum_{\text{v.b.}}(-e)\boldsymbol{v}_{\boldsymbol{k}} = 0$). When a state with crystal momentum $\boldsymbol{k}$ is empty as seen in the left, the current is

$$\boldsymbol{J}_{\text{v.b.}}(\boldsymbol{k}) = \sum_{\text{v.b.}}(-e)\boldsymbol{v}_{\boldsymbol{k}'} - (-e)\boldsymbol{v}_{\boldsymbol{k}} = e\boldsymbol{v}_{\boldsymbol{k}}, \tag{3.4}$$

as if there is a particel with the charge $+e$ and the velocity $\boldsymbol{v}_{\boldsymbol{k}}$. Such a many-body state in valence band is called **hole**.

We write the wavenumber of hole as $\boldsymbol{k}_\mathrm{h}$, then it should be the variation of the total wavenumber (momentum) due to the creation of the hole.

$$\boldsymbol{k}_\mathrm{h} = \sum_{\text{v.b.}} \boldsymbol{k}'_\mathrm{e} - \boldsymbol{k}_\mathrm{e} = -\boldsymbol{k}_\mathrm{e}. \tag{3.5}$$

When an electric field $\boldsymbol{E}$ is applied, the valence electrons are accelerated and move in the k-space. The “hole" follows the movement, that is, the equation of motion for holes is the same as that for electrons. However, we define a hole has the charge $+e$, then the acceleration by the electric field should be opposite to that for electrons and for consistency the sign of the effective mass should be opposite.

$$m^* \frac{d\boldsymbol{v}}{dt} = (-e)\boldsymbol{E} \;\rightarrow\; (-m^*)\frac{d\boldsymbol{v}}{dt} = e\boldsymbol{E}.$$

The kinetic energy decreases with the creation of a hole and if we take the origin of energy at the top of valence band, we get

$$\left(\frac{1}{m_h^*}\right)_{ij} = -\left(\frac{1}{m_e^*}\right)_{ij}, \quad E_\mathrm{h}(\boldsymbol{k}_\mathrm{h}) = E_\mathrm{h}(-\boldsymbol{k}_\mathrm{e}) = -E_e(\boldsymbol{k}_\mathrm{e}). \tag{3.6}$$

Form (3.6), $m_h^*$ is positive around the valence band top and the dispersion is obtained with 180° rotation of the electron dispersion. The density of states $\mathscr{D}_h(E)$ is the same as $\mathscr{D}_e(E)$. The above definitions make it possible to treat the holes as positive charge carriers. The hole band drawn in the lower left is used to be consistent with the electron dispersion and the hole picture. We need to be careful that the frequently-used whilte hole picture (actually in the left) which is really an electron dispersion and the “white hole" is placed at $\boldsymbol{k}_\mathrm{e}$ which is $\boldsymbol{k}_\mathrm{h}$.

### 3.1.3 Carrier distribution in thermal equilibrium

Let us see how electrons and holes distribute in energy space at a finite temperature obeying the Fermi distribution function. For a while we treat general properties, which hold also for doped semiconductors. The effect of doping can be included in the position of the Fermi level $E_{\rm F}$. The numbers of electrons and holes which exist in $E \sim E + dE$ are

$$g_e(E)dE = \mathscr{D}_e(E)f(E)dE, \tag{3.7a}$$

$$g_h(E)dE = \mathscr{D}_h(E)[1-f(E)]dE \equiv \mathscr{D}_h(E)f_h(E)dE. \tag{3.7b}$$

Here we introduced the hole distribution function as (Fig. 3.3(c)),

$$f_h(E) = 1 - f(E) = \frac{1}{1+\exp(E_{\rm F}-E)/k_{\rm B}T)}. \tag{3.8}$$

For the density of states, we use those of particles with the effective masses. From (3.2),

$$\mathscr{D}_e(E) = \frac{\sqrt{2m_e^{*3}}}{\pi^2\hbar^3}\sqrt{E-E_c} \quad \text{(conduction band)}, \tag{3.9a}$$

$$\mathscr{D}_h(E) = \frac{\sqrt{2m_h^{*3}}}{\pi^2\hbar^3}\sqrt{E_v-E} \quad \text{(valence band)}. \tag{3.9b}$$

Here $E_c$, $E_v$ are the bottom of conduction band and the top of valence band respectively as in Fig. 3.3(a).

Hence the distributions of electrons and holes at a finite temperature should be as in Fig. 3.3(b), giving the electron concentration in the conduction band $n$, the hole concentration $p$ in the valence band as

$$n = \int_{E_c}^{\infty} g_e(E)dE = \frac{\sqrt{2m_e^{*3}}}{\pi^2\hbar^3}\int_{E_c}^{\infty}\frac{\sqrt{E-E_c}dE}{1+\exp(E-E_{\rm F})/k_{\rm B}T}, \tag{3.10a}$$

$$p = \int_{-\infty}^{E_v} g_h(E)dE = \frac{\sqrt{2m_h^{*3}}}{\pi^2\hbar^3}\int_{-\infty}^{E_v}\frac{\sqrt{E_v-E}dE}{1+\exp(E_{\rm F}-E)/k_{\rm B}T}. \tag{3.10b}$$

In the case of $f_{\rm F}(E) \ll 1 (E \geq E_c)$, $f_h(E) \ll 1 (E \leq E_v)$, the distribution can be approximated by Maxwellian as

$$f_{\rm F}(E) \sim \exp(E_{\rm F}-E)/k_{\rm B}T, \quad f_h(E) \sim \exp(E-E_{\rm F})/k_{\rm B}T. \tag{3.11}$$

We apply the identity

$$\int_0^\infty \sqrt{x}e^{-x}dx = \frac{\sqrt{\pi}}{2}$$

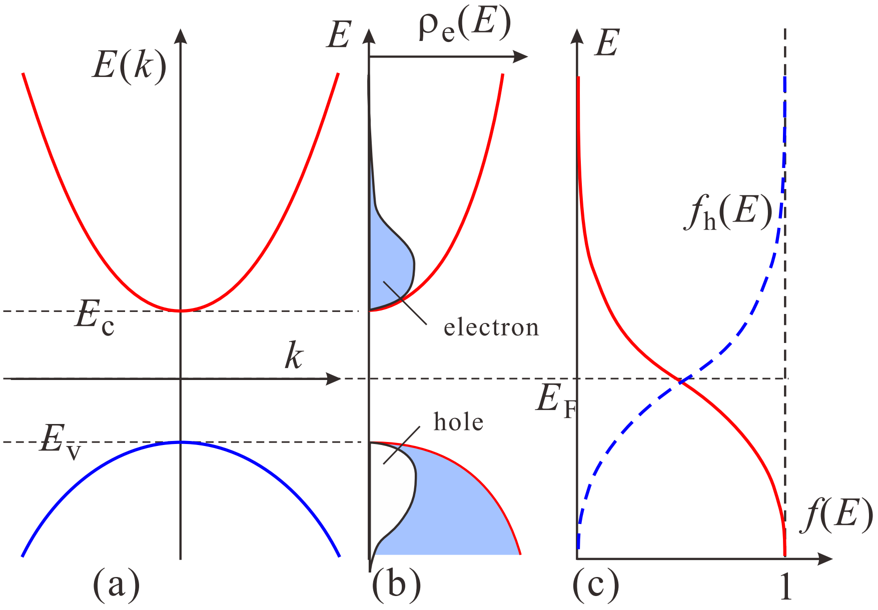


**Fig.** 3.3 (a) Schematic diagram of energy bands. (b) Density of states and the distributions of electrons $n(E)$(blue-gray), $p(E)$(white). (c) Electron distribution function $f(E)$(red solid line), and hole distribution function $f_h(E)$(blue broken line).

with $x = (E - E_{\rm F})/k_{\rm B}T$ to obtain

$$n = 2\left(\frac{m_e^* k_{\rm B}T}{2\pi\hbar}\right)^{3/2} \exp\left(\frac{E_{\rm F} - E_c}{k_{\rm B}T}\right) \equiv N_c \exp\left(\frac{E_{\rm F} - E_c}{k_{\rm B}T}\right), \tag{3.12a}$$

$$p = 2\left(\frac{m_h^* k_{\rm B}T}{2\pi\hbar}\right)^{3/2} \exp\left(\frac{E_v - E_{\rm F}}{k_{\rm B}T}\right) \equiv N_v \exp\left(\frac{E_v - E_{\rm F}}{k_{\rm B}T}\right). \tag{3.12b}$$

$N_c$ and $N_v$ are the coefficients which give $n$, $p$ respectively in the situation the energy states are concentrated at $E_c$ and $E_v$. They are called **effective density of states**. From (3.10a) and (3.10b) we obtain

Law of mass action

$$np = N_c N_v \exp\left(\frac{E_v - E_c}{k_{\rm B}T}\right) = N_c N_v \exp\left(-\frac{E_g}{k_{\rm B}T}\right) = n_{\rm i}^2 \tag{3.13}$$

Here the width of forbidden band $E_g \equiv E_c - E_v$ is called **energy gap**, and $n_{\rm i}$ は真性半導体の場合のキャリア濃度である. Equation (3.13) does not depend on the position of $E_{\rm F}$, which varies, *e.g.* with doping. In other words, the product $np$ in the thermal equilibrium is determined only by the temperature and the species of semiconductor.

In intrinsic semicondutors, there is no space charge and the charge neutral condition leads to $n = p$ hence is written as $n_{\rm i}$ in the above law of mass action. The relation $n = p$ in intrinsic semiconductors leads to

$$E_{\rm F} = \frac{E_c + E_v}{2} + \frac{k_{\rm B}T}{2} \ln \frac{N_v}{N_c} = \frac{E_c + E_v}{2} + \frac{3k_{\rm B}T}{4} \ln \frac{m_h}{m_e}, \tag{3.14}$$

which gives the position of $E_{\rm F}$. At low temperatures the second term gets small and $E_{\rm F}$ comes close to the middle of the band gap.

## 3.2 Impurity doping

In semiconductors, very small amount of impurities give drastic change in the material properties. Such addition of impurities is called **doping** [*1].

### 3.2.1 Donors and acceptors

As a typical example, the case of Si is shown schematically in Fig. 3.4. In Si pure crystal, as in (a), a Si atom has four nearest neighbor atoms, which have 4 covalent electrons. As a result each atom has eight electrons in the outmost shell, which fill up $3s$ and $3p$ orbits forming the closed shell geometry. When the center atom is replaced with an Sb (group-V) atom, there is an excess electron for the closed shell structure as in (b). On the other hand, the positive charge in the nucleous excesses the negative one of surrounding electrons by $+e$, which forms a Coulomb potential around the Sb atom. The excess electron is excited to the conduction band or loosely trapped in the bound state in the Coulomb potential. An impurity that emits electrons to the conduction band or the shallow levels is called **donor**.

When the center atom is replaced with a B atom (group-III), which is just the opposite of Sb, there are not enough electrons to form a closed shell structure. Therefore, holes are created in the valence band to supplement the electrons, but as a result, the electron charge becomes extra around the B atom, and a Coulomb potential of only $-e$ is generated. Impurities that emit holes into the valence band and shallow levels in this way are called **acceptors**(acceptor).

[*1] In so called studies in strongly correlated systems, which began with the studies of high-$T_{\rm c}$ superconductors, addtion of atoms wtih concentrations even far above 1% is called “doping" as far as the crystal structure is unchanged. Such regions are called “alloying" in the semiconductor fields. Furthermore, enhancement of carrier concentrations with application of strong electric field is sometimes called “electric field doping", and the addition of impurities is called “chemical doping." Here, however, I follow tranditional epxression in the field of semiconductors.

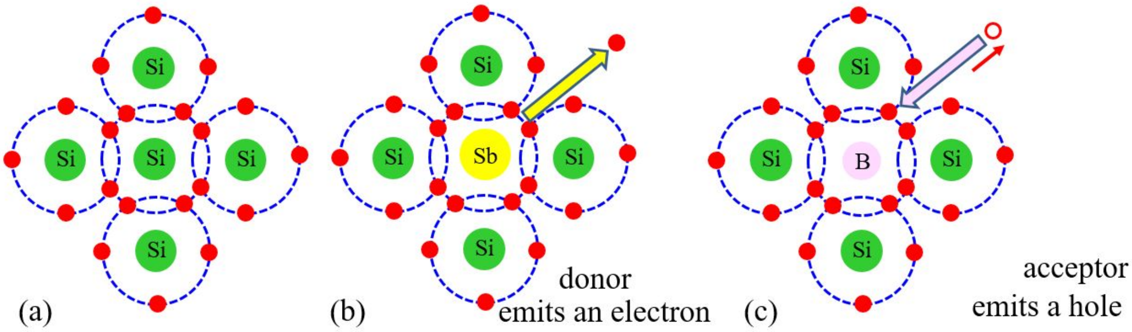


**Fig.** 3.4 (a) Illustration of electronic structure in the outmost shell in a Si atom in a Si crystal. (b) In the case of replacement with an Sb atom. There is an excess electron for Sb to form the closed shell structure. The excess electron should go out from the covalent system. (c) In the case of replacement with a B atom, which opens a "hole" in the valence band.

The situation is a little more complicated in the case of compound semiconductors than in Group IV elemental semiconductors. For example, when a group III-V semiconductor is doped with a group IV element, replacing the group III site becomes a donor, and replacing the group V site becomes an acceptor. Elements whose donor / acceptor changes depending on the doping method are called amphoteric.

### 3.2.2 Effective mass approximation for shallow hydrogen-like levels

Regarding the Coulomb potential formed by donors and acceptors, semiconductors generally have a relatively large permittivity due to the polarization of valence electrons, and this impurity attraction potential is considerably weaker than in vacuum. Therefore, the binding energy of the impurity bound state is smaller than that of the hydrogen atom, and in many cases, it spreads over several unit cells, and the effective mass approximation can be applied.

For the case of isotropic effective mass, taking the origin at the impurity potision we write the effective mass equation from $U(r) = -e^2/4\pi\epsilon_0\epsilon r$ as

$$\left[-\frac{\hbar^2\nabla^2}{2m^*} - \frac{e^2}{4\pi\epsilon_0\epsilon r}\right] F(\boldsymbol{r}) = EF(\boldsymbol{r}), \tag{3.15}$$

which as the same form as that of hydrogen atom other than the effective mass $m^*$ and the relative permittivity $\epsilon$. Hence we can readily apply the results for hydrogen atom. Here we write the **effective Rydberg constant** and the **effective Bohr radius** as

$$Ry^* = \frac{e^2 m^*}{2(4\pi\epsilon\epsilon_0)^2\hbar^2} = \frac{m^*}{m}\frac{1}{\epsilon^2}Ry, \quad a_{\rm B}^* = \frac{4\pi\epsilon\epsilon_0\hbar^2}{m^* e^2} = \frac{m}{m^*}\epsilon a_{\rm B} \tag{3.16}$$

respectively. The eigenenergy is then represented as

$$E_n = E_c - \frac{Ry^*}{n^2} \quad (n = 1, 2, \cdots) \tag{3.17}$$

and the wavefunction corresponding to $1s$ state is

$$\psi_{1s}(\boldsymbol{r}) = \sqrt{\frac{1}{\pi a_{\rm B}^{*3}}} \exp\left(-\frac{\boldsymbol{r}}{a_{\rm B}^*}\right). \tag{3.18}$$

An expample of semiconductor with such an isotropic effective mass is GaAs. At the conduction band minimum placed at Γ-point, $\epsilon \approx 11.5$, $m^* \approx 0.067m$. $a_{\rm B}^* = 172 a_{\rm B} = 91$ Åis sufficiently longer than the lattice constant 5.65 Åand guarantees the legitimacy of the effective mass approximation. From $Ry^* = 5.07 \times 10^{-4} Ry = 5.57 \times 10^3$ m$^{-1}$ the binding energy of $1s$ state is as small as 6.9 meV.

| Semiconductor | Calculated binding energy (meV) | Experimental binding energy (meV) |
|---|---|---|
| GaAs | 5.72 | $\mathrm{Si}_{\mathrm{Ga}}$(5.84); $\mathrm{Ge}_{\mathrm{Ga}}$(5.88)<br>$\mathrm{S}_{\mathrm{As}}$(5.87); |
| InP | 7.14 | 7.14 |
| InSb | 0.6 | $\mathrm{Te}_{\mathrm{Sb}}$ (0.6) |
| CdTe | 11.6 | $\mathrm{In}_{\mathrm{Cd}}$ (14); $\mathrm{Al}_{\mathrm{Cd}}$ (14) |
| ZnSe | 25.7 | $\mathrm{Al}_{Zn}$ (26.3); $\mathrm{Ga}_{\mathrm{Zn}}$ (27.9)<br>$\mathrm{F}_{\mathrm{Se}}$ (29.3); $\mathrm{Cl}_{Se}$ (26.9) |

**Tab.** 3.1 Effective mass approximation for hydrogen-like impurities and the measured value of binding energies.

Tab. 3.1 shows the comparison of the value given by (3.17) and experimentally measured values for isotropic effective mass condition. The agreement is satisfactory.

Then what if the effective mass is anisotropic and there are six conduction band valleys, as in Si? We consider the effective mass approximation for the valley along (0,0,1). The equation for the spheroidal surface is

$$E_1(\boldsymbol{k}) = \frac{\hbar^2}{2}\left[\frac{k_x^2+k_y^2}{m_t}+\frac{(k_z-k_0)^2}{m_l}\right]. \tag{3.19}$$

Then the effective mass equation is

$$\left[-\frac{\hbar^2}{2m_t}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)-\frac{\hbar^2}{2m_l}\frac{\partial^2}{\partial z^2}-\frac{e^2}{4\pi\epsilon_0\epsilon r}\right]F(\boldsymbol{r}) = EF(\boldsymbol{r}). \tag{3.20}$$

The eigenfunction can be approximated by the variational method assuming an anisotropic exponential function. As a trial function we take $a$ and $b$ as the paramters and write down as

$$F_{1s}(\boldsymbol{r}) = \sqrt{\frac{1}{\pi a^2 b}}\exp\left(-\sqrt{\frac{x^2+y^2}{a^2}+\frac{z^2}{b^2}}\right). \tag{3.21}$$

The stationary condition gives numerical solutions as $a = 2.5$ nm, $b = 1.42$ nm, $E = 29$ meV. However the experiments give 33 meV for Li (the shallowest case), 45 meV for P manifesting that the approximation is not appropriate. Further discussion will be given in Appendix 3B.

## 3.3 Carrier statistics in doped semiconductors

Let us consider the case we dope donors uniformly with the density $N_{\mathrm{D}}$. At absolute zero all the electrons emitted from the donors are bound to the donors. [*2] At finite temperatures some of them are excited to the conduction band and can carry electric charges. We call them “carriers" or “electrons". Let $n$ be the density of such electrons and $n_{\mathrm{D}}$ be the density of electrons bounded at the donors. From the charge neutrality condition we get $n + n_{\mathrm{D}} = N_{\mathrm{D}}$.

Now we estimate Helmholtz free energy $F = U - TS$ by considering the number of cases $W$ for assigning $n_{\mathrm{D}}$ electrons to $N_{\mathrm{D}}$ states. From $S = k_{\mathrm{B}} \ln W$,

$$F = E_{\mathrm{D}} n_{\mathrm{D}} - k_{\mathrm{B}} T \ln\left[2^{n_{\mathrm{D}}}\frac{N_{\mathrm{D}}!}{n_{\mathrm{D}}!(N_{\mathrm{D}}-n_{\mathrm{D}})!}\right].$$

$E_{\mathrm{D}}$ is the position of the bound state measured from the bottom of the conduction band and $2^{n_{\mathrm{D}}}$ is due to the spin degeneracy. We assume that the Coulomb repulsion prevents double occupation of a localized state with two electrons.

[*2] In so called degenerate semiconductors the following discussion does not hold.

According to Starling approximation $\ln N! \sim N \ln N - N$, the chemical potential (Fermi energy) is given as

$$\mu = E_\mathrm{F} = \frac{\partial F}{\partial n_\mathrm{D}} = E_\mathrm{D} - k_\mathrm{B} T \ln \left[ \frac{2(N_\mathrm{D} - n_\mathrm{D})}{n_\mathrm{D}} \right]. \tag{3.22}$$

And from this

$$n_\mathrm{D} = N_\mathrm{D} \left[ 1 + \frac{1}{2} \exp \left( \frac{E_\mathrm{D} - E_\mathrm{F}}{k_\mathrm{B} T} \right) \right]^{-1} \tag{3.23}$$

is obtained. The factor 1/2 on the exponential function is due to the spin degeneracy.

Similarly, for uniform doping of acceptors with density $N_\mathrm{A}$, the density of electrons bounded to the acceptors $n_\mathrm{A}$ is

$$n_\mathrm{A} = N_\mathrm{A} \left[ 1 + 2 \exp \left( \frac{E_\mathrm{A} - E_\mathrm{F}}{k_\mathrm{B} T} \right) \right]^{-1}. \tag{3.24}$$

Here we have a factor 2 instead of 1/2 but the density of holes bounded to the acceptors is $p_\mathrm{A} = N_\mathrm{A} - n_\mathrm{A}$ and symmetrical with $n_\mathrm{D}$ having a factor 1/2.

From (3.22), if we dope only “shallow" donors, for which the effective mass approximation holds, $E_\mathrm{F}$ comes to $E_\mathrm{D}$ at $T \to 0$. $E_\mathrm{D}$ should be much smaller than $E_\mathrm{g}$. Accordingly from (3.23), the electron concentration $n$ becomes much higher than that of the intrinsic semiconductor at finite temperatures. This type of semiconductors are called **n-type**. Similarly doping of acceptors enhances the hole concentration $p$. We call them **p-type**.

When donors and acceptors co-exist, the semiconductor becomes n-type for $N_\mathrm{D} \gg N_\mathrm{A}$ and p-type for $N_\mathrm{D} \ll N_\mathrm{A}$. In the former, some of the electrons emitted from donors are captured to acceptors and almost all the acceptors are ionized. In the latter, the other way around. In both cases we say such semiconductors are **compensated**.

Remember the semiconductor equation (3.13) then the product $np$ does not depend on the doping. If one of $n$ , $p$ increases with doping, then the other decreases. In the case of n-type semiconductor under $N_D \gg N_A$, $n$ is much higher than $p$ by many orders, and we call the electrons **majority carriers** and the holes **minority carriers**. The other way around in the case of p-type semiconductors.

Even in the presence of donors and acceptors eq.(3.12) hold and simultaneous satisfaction of them gives $n$, $p$ and $E_\mathrm{F}$. To obtain $E_\mathrm{F}$ with knowledge of $n$, $p$ approximate expressions

$$E_\mathrm{F} \approx E_C + k_\mathrm{B} T \left[ \ln \left( \frac{n}{N_C} \right) + 2^{-3/2} \left( \frac{n}{N_C} \right) \right], \tag{3.25a}$$

$$E_\mathrm{F} \approx E_V - k_\mathrm{B} T \left[ \ln \left( \frac{p}{N_V} \right) + 2^{-3/2} \left( \frac{p}{N_V} \right) \right] \tag{3.25b}$$

are convenient. In the region where (3.23), (3.24) hold the last term can be omitted.

In an n-type semiconductor with compensation, $p$,$n_\mathrm{A}$ can be ignored and the electrically neutral condition is

$$n + N_A = N_\mathrm{D} - n_\mathrm{D}. \tag{3.26}$$

Substitution of eq.(3.23) gives

$$\frac{n + N_\mathrm{A}}{N_\mathrm{D} - N_\mathrm{A} - n} = \frac{1}{2} \exp \left( \frac{E_\mathrm{D} - E_\mathrm{F}}{k_\mathrm{F} T} \right). \tag{3.27}$$

Equation (2.22) holds for the case of doped semiconductors with shifts of $E_\mathrm{F}$, multiplication of each side of the equation results in

$$\frac{n(n + N_\mathrm{A})}{N_\mathrm{D} - N_\mathrm{A} - n} = \frac{1}{2} N_c \exp \left( - \frac{\mathit{\Delta} E_\mathrm{D}}{k_\mathrm{B} T} \right), \quad \mathit{\Delta} E_\mathrm{D} \equiv E_c - E_\mathrm{D}. \tag{3.28}$$

The temperature dependence of carrier concentration $n$ described by eq.(3.28) has the following four characteristic regions:

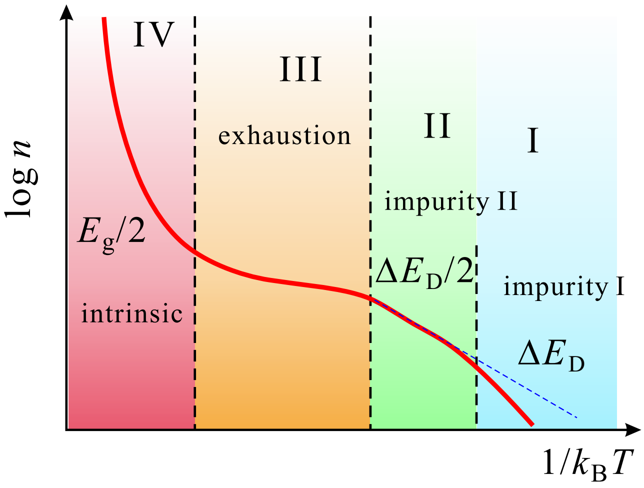


**Fig.** 3.5 Characteristic four temperature regions of an n-type semiconductor with compensation. Schematical temperature dependence of carrier concentration $n$ is plotted versus $1/T$ in semi-log scale.

I. Impurity (Freeze-out) region I: At very low temperatures and the case of $n \ll N_A \ll N_D$,

$$n \approx \frac{N_D N_c}{2N_A} \exp\left(-\frac{\Delta E_D}{k_B T}\right), \tag{3.29}$$

where $n$ decreases with lowering the temperature in an Arrhenius type with an activation energy of $\Delta E_D$.

II. Impurity (Freeze-out) region II: In middle temperature range, in the case of $N_A \ll n \ll N_D$,

$$n \approx \left(\frac{N_c N_D}{2}\right)^{1/2} \exp\left(-\frac{\Delta E_D}{2k_B T}\right), \tag{3.30}$$

where the temperature dependence shows again an Arrhenius type but with a different activation energy, which is a half of that in the impurity region I.

III. Exhaustion (Saturation) region: Temperature is higher than $\Delta E_D$ ($k_B T > \Delta E_D$). The exponential function in eq.(3.28) is now almost a constant ($\sim 1$) and

$$n \approx N_D - N_A. \tag{3.31}$$

Electrons once captured in donors are “exhaustively" excited to the conduction band and work as carriers.

IV. Intrinsic region: At higher temperatures where direct thermal excitation for the valence band to conduction band cannot be ignored in comparison with $N_D$, the temperature dependence of the carrier concentration asymptotically approaches to that in an intrinsic semiconductor described as eq.(3.10b), (3.14).

We show the behavior in Fig. 3.5 schematically. For semiconductor devices, the exhaustion region III is mostly used.

## Appendix 2B: Wannier functions and the effective mass approximation

There is a way to derive the effective mass approximation by using expansion with the **Wannier function**. Thouhg it is essentially the same as in Sec.2.3.1, Wannier functions have several convenient points and we may use them afterwards. In that case, I will introduce it again, but let’s take a quick look at what it is as an appendix.

## 2B.1 Wannier function

The Wannier function is defined as Fourier transform of Bloch function as follows.

$$w_n(\boldsymbol{r}-\boldsymbol{R}_j)=\frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}}\exp(-i\boldsymbol{k}\cdot\boldsymbol{R}_j)\psi_{n\boldsymbol{k}}(\boldsymbol{r}). \tag{2B.1}$$

The Bloch function is usually given in the coordinate representation but here the spatial coordinate is a parameter and it is now taken as a function of wavenumber $\boldsymbol{k}$. The summation over $\boldsymbol{k}$ is inside the Brillouin zone. The Bloch function is a product of a lattice periodic function and a plane wave and sperad over the space. On the other hand, the Wannier function has tendency to localized to the lattice point $\boldsymbol{R}_j$. This is understood by making the lattice-periodic function in the Bloch function a constant, which makes the Wanner function completely localized on $\boldsymbol{R}_j$. Equation (2B.1) can be seen as the expantion of the Wannier function with the Bloch function. Conversely the Bloch function can be expanded by the Wannier function as

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r})=\frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}}\exp(i\boldsymbol{k}\cdot\boldsymbol{R}_j)w_n(\boldsymbol{r}-\boldsymbol{R}_j). \tag{2B.2}$$

An advantage in the Wannier function is the orthgonality, that is

$$\langle w_{n'}^*(\boldsymbol{r}-\boldsymbol{R}_{j'})|w_n(\boldsymbol{r}-\boldsymbol{R}_j)\rangle=\delta_{jj'}\delta_{nn'}. \tag{2B.3}$$

We skip the proof but straightforwardly performed with using the summatoins on the lattice and the reciprocal lattice. The Wannier function is normalized if it is defined as (2B.1) and also forms a complete set.

## 2B.2 Derivation of effective mass approximation

The problem is the addition of perturbation potential $U(\boldsymbol{r})$ to the crystal Hamiltonian $\mathscr{H}_0$, that is

$$[\mathscr{H}_0+U(\boldsymbol{r})]\phi(\boldsymbol{r})=E\phi(\boldsymbol{r}). \tag{2B.4}$$

The derivation goes almost parallely as that with the Bloch funtion. First we expand the wavefunction with the Wannier functions as

$$\phi(\boldsymbol{r})=\sum_{n,j'}F_n(\boldsymbol{R}_{j'})w_n(\boldsymbol{r}-\boldsymbol{R}_{j'}). \tag{2B.5}$$

We assume that $\mathscr{H}_1$ does not have the elements for interband transition (the amplitude is too small) and we drop $n$ henceforth. Tne Wannier function $w(\boldsymbol{r}-\boldsymbol{R}_j)$ is simply written as $|j\rangle$. Equation (2B.5) is substituted into eq. (2B.4) and with takeing the inner product with $\langle j|$, the orthogonality (2B.3) leads to

$$\sum_{j'}\langle j|\mathscr{H}_0|j'\rangle F(\boldsymbol{R}_{j'})+\sum_{j'}\langle j|U(\boldsymbol{r})|j'\rangle F(\boldsymbol{R}_{j'})=EF(\boldsymbol{R}_j). \tag{2B.6}$$

As seen above $|j\rangle$ is localized to $\boldsymbol{R}_j$ and because $U(\boldsymbol{r})$ is slowly varying function in the scale of lattice constant, we can approximate as

$$\sum_{j'}\langle j|U(\boldsymbol{r})|j'\rangle\approx\sum_{j'}U(\boldsymbol{R}_{j'})\langle j|j'\rangle=U(\boldsymbol{R}_j). \tag{2B.7}$$

The term of crystal Hamiltonian can be written with shifting the spatial origin as

$$\langle j|\mathscr{H}_0|j'\rangle=\langle w(\boldsymbol{r})|\mathscr{H}_0|w(\boldsymbol{r}-(-\boldsymbol{R}_{j'}+\boldsymbol{R}_j))\rangle\equiv h_0(\boldsymbol{R}_j-\boldsymbol{R}_{j'}). \tag{2B.8}$$

The Bloch funcition $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ is the eigenstate of $\mathscr{H}_0$ and the application of the effective mass approximation to the eigenenergy gives

$$\langle\psi_{\boldsymbol{k}}(\boldsymbol{r})|\mathscr{H}_0|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle=E_0(\boldsymbol{k})=\frac{\hbar^2\boldsymbol{k}^2}{2m^*}. \tag{2B.9}$$

$\psi_{\boldsymbol{k}}(\boldsymbol{r})$ can be expanded as (2B.2) and leads to

$$E_0(\boldsymbol{k}) = \frac{1}{N}\sum_{j,j'}\exp[-i\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_{j'})]\langle j|\mathscr{H}_0|j'\rangle = \frac{1}{N}\sum_{j,j'}\exp[-i\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_{j'})]h_0(\boldsymbol{R}_j - \boldsymbol{R}_{j'})$$
$$= \sum_j \exp(-i\boldsymbol{k}\cdot\boldsymbol{R}_j)h_0(\boldsymbol{R}_j). \tag{2B.10}$$

The inverse transformation gives

$$h_0(\boldsymbol{R}_j) = \frac{1}{N}\sum_{\boldsymbol{k}} E_0(\boldsymbol{k})\exp(i\boldsymbol{k}\cdot\boldsymbol{R}_j). \tag{2B.11}$$

Though $F(\boldsymbol{R}_j)$ is just defined on the lattice point as in (2B.5), since the spatial variation in $U(\boldsymbol{r})$ is slow, the differences between the values of $F$ for the neighboring lattice point are small and smooth interpolation is possible. The operator of spatial shift by $\boldsymbol{a}$ is $\exp(-\boldsymbol{a}\cdot\nabla)$[*3], then we can write

$$F(\boldsymbol{r} - \boldsymbol{R}_j) = \exp(-\boldsymbol{R}_j\cdot\nabla)F(\boldsymbol{r})$$

to obtain

$$\sum_{j'} h_0(\boldsymbol{R}_{j'})F(\boldsymbol{r} - \boldsymbol{R}_{j'}) = \sum_{j'} h_0(\boldsymbol{R}_{j'})\exp(-\boldsymbol{R}_{j'}\cdot\nabla)F(\boldsymbol{r}). \tag{2B.12}$$

On the other hand for (2B.10),

$$E_0(\boldsymbol{k})F(\boldsymbol{r}) = \sum_{j'} h_0(\boldsymbol{R}_{j'})\exp(-i\boldsymbol{k}\cdot\boldsymbol{R}_{j'})F(\boldsymbol{r}). \tag{2B.13}$$

We formally inverse Fourier transform (2B.12) and (2B.13). Then these equations have the common right hand side if we make replacement of $\boldsymbol{k} \to -\nabla$. Therefore we can write

$$\sum_{j'} h_0(\boldsymbol{R}_{j'})F(\boldsymbol{r} - \boldsymbol{R}_{j'}) = E_0(-i\nabla)F(\boldsymbol{r}). \tag{2B.14}$$

All the above results are restored into (2B.6) and we replace $\boldsymbol{R}_j$ with a continuous variable $\boldsymbol{r}$. Then we obtain

$$\left[-\frac{\hbar^2}{2m^*}\nabla^2 + U(\boldsymbol{r})\right]F(\boldsymbol{r}) = EF(\boldsymbol{r}). \tag{2B.15}$$

Here I have introduce the proof in the textbook [1]. It has, though, a small jump in the logic from (2B.13)→(2B.14). Also the description of $\boldsymbol{R}_j \to \boldsymbol{r}$ is rather vague. The derivation in [2] is more strict but it needs the width of paper. For the transformation $\boldsymbol{R}_j \to \boldsymbol{r}$, clearly written as “to be strict, this should be done in variations."

# Appendix 3A: Methods for impurity doping

Various methods have been developed for impurity doping. A part of it is introduced in the following. Impurities to be doped are called dopants, and base crystal is called host.

## 3A.1 Mixing of impurities to the raw material

We have introduced the semiconductor crystal growth method, but especially in the method of growing bulk crystals from a material melt, if impurities are mixed in the raw material in advance, doping may be performed with relatively good uniformity. In many cases, for example, even in the Czochralski method, a concentration gradient is generated in the crystal growth direction by segregation. For this reason, various growth measures are taken, such as adding a dopant

[*3] This can be confirmed, for example by the Taylor expansion.

to the melt crucible in order to obtain a uniform doping concentration. In addition, in the case of amphoteric impurities, there is a possibility that some impurities will be compensated depending on the growth conditions.

In epitaxial thin film growth, by controlling the dopant and irradiating it on the growth surface, it is possible to start a stepwise distribution and create various concentration distributions with high non-equilibrium while maintaining crystallinity. This bf modulation doping method plays a major role in occupying an extremely important position in the semiconductor industry for semiconductor thin films.

## 3A.2 Thermal diffusion method

A method in which the host is kept at a high temperature in a state where the dopant is present in a high concentration on the surface of the host, and is mixed inside by heat diffusion. Methods for increasing the concentration of the surface include contacting with the vapor of the dopant and pre-depositing a thin film of the dopant on the host surface. In the figure below, the dopant and host wafer are simultaneously enclosed in a quartz tube and heated together so that the vapor of the dopant flows to the surface of the high-temperature wafer. In some cases, the whole is put into the furnace without creating a flow. In addition, if some of the constituent elements of the host have a high vapor pressure, it is necessary to suppress the separation from the surface by mixing the vapor of this element.

In the thermal diffusion method, the concentration is high near the surface and low as it goes inside. It is usually used to form a device near the surface. It is used for integrated circuit formation because the doping region can be patterned by masking the wafer when the vapor is applied.

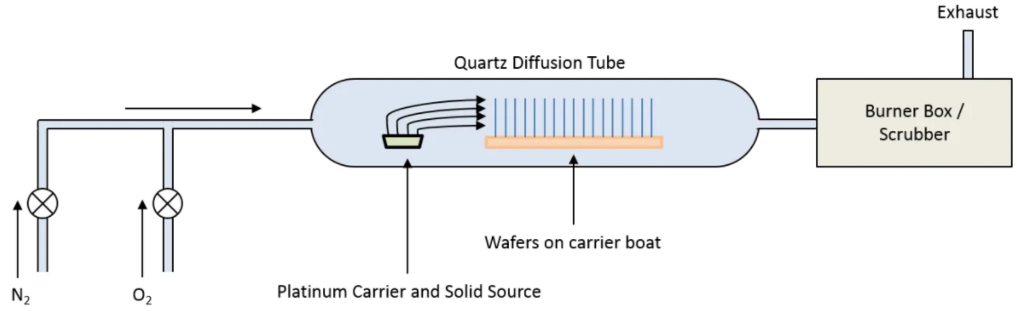


**Fig.** 3A.1 Schematic diagram of thermal diffution doping.

## 3A.3 Ion implantation

The ion implantation method is used not only for doping but also for cutting and oxidation of the inner layer. As in Fig. 3A.2(a), ions entering from the source are bent by a magnetic field, passed through a diaphragm for mass spectrometry, then narrowed down by a lens, irradiated on a wafer, and scanned by an XY voltage.

As in the imaginary figure in (b), Since the ions that reach the surface have high kinetic energy, they invade the crystal in a non-equilibrium manner and stop at a depth corresponding to the average kinetic energy. Since the crystallinity of the passing region decreases due to the collision of ions and the dopant is not always in the stable position, it is often annealed after implantation. The distribution of dopants is generally represented by the Gaussian distribution after annealing.

In addition to doping, it was once actively used as one of the Silicon on Insulator (SOI) techniques that form an oxide film inside by implanting oxygen ions and annealing as described above. At present, the seemingly primitive method of bonding after surface oxidation is mainly used.

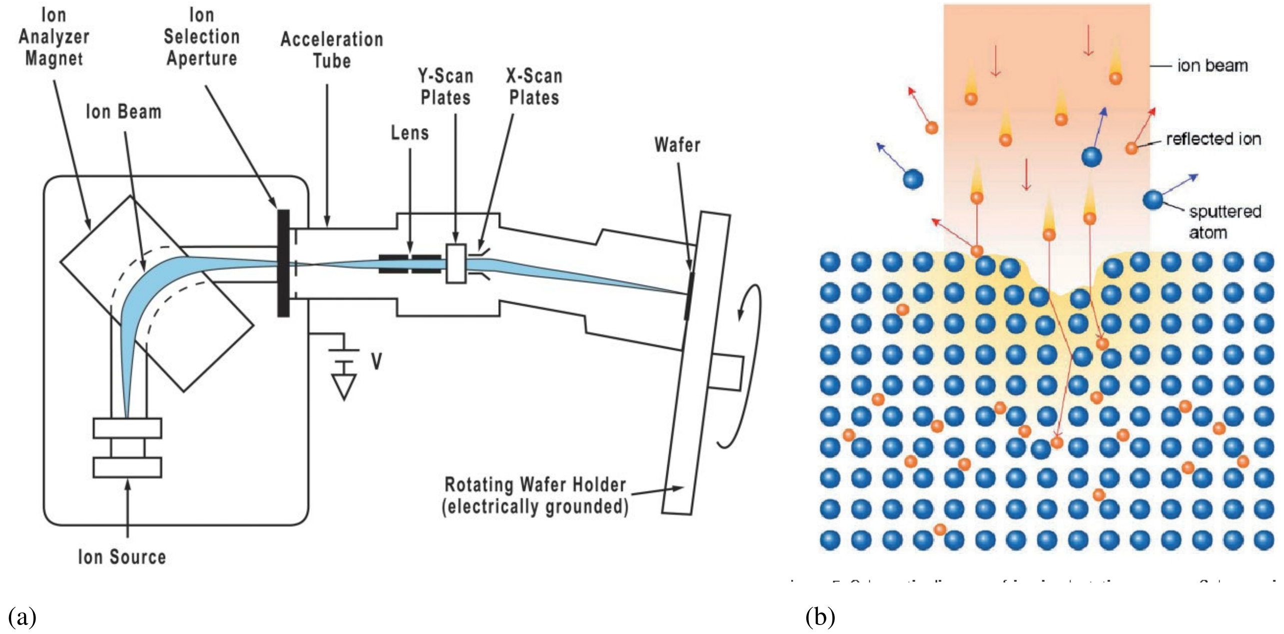


**Fig.** 3A.2 (a) Illustration of ion implantation doping. Ions coming out of the source are subjected to mass spectrometry using a magnetic field to sort them, and then the focused beam is scanned onto the wafer. (b) Imaginary illustration of the host surface during the ion implantation.

### 3A.4 Irradiation with neutrons

Currently, it is rarely used, but there is an interesting doping method that uses neutrons. For that the following nuclear reaction is used.

$$[n\ +\ ^{30}\mathrm{Si}]\ \to {}^{31}\mathrm{Si}\ \to\ [^{31}\mathrm{P}\ +\ \beta]$$

With reactor neutrons, extremely uniform doping can be performed without compromising crystallinity. However, due to problems such as throughput, this method remains at the research level.

## Appendix 3B: Shallow donors in Si

For the improvement of the effective mass approximation for the shallow donor levels in Si, we consider the effect of multiple (6) valleys. Since the impurity potential also has a significant magnitude and steepness near the center, it is possible that it has matrix elements between the eigenfunctions attached to the degenerate valleys. We consider the donor wavefunction $\chi(\boldsymbol{r}) = F(\boldsymbol{r})\psi(\boldsymbol{r})$, where $F(\boldsymbol{r})$ is the envelope function and $\psi(\boldsymbol{r})$, for each valley and obtain the donor function as a linear combination of them as

$$\phi^{(i)}(\boldsymbol{r}) = \sum_{j=1}^{6} \alpha_j^{(i)} \chi_j, \tag{3B.1}$$

where $j$ is the index of valley and $i$ is the index of symmetry reflecting that of surrounding atoms. Here we assume there is no mixing between the eigenstates with different quantum number in the trap potential.

While the isotropic potential approximation does not hold in the vicinity of impurities, the spatial symmetry of the **crystal field** created by the atoms around the impurity ions must be taken into consideration when taking a linear bond. It is convenient to use point group theory for the discussion, and I will use it here as well. Regarding the point group, if I have time, I would like to supplement the minimum knowledge in the appendix of lecture notes etc.[**?**, 3]. In the case of Si, the nearest neighbor atoms are at the apexes of a regular tetrahedron containing impurity ions. The point group corresponding to the symmetry is expressed as the symbol $T_\mathrm{d}$. The elements of point group have correspondences to the representations, which express symmetry operations. The independent elements have one-to-one correspondence to the reduced representations. The reduced representations in $T_\mathrm{d}$ group are $A_1$, $E$ and $T_1$, which have single, double, triple degeneracy respectively and there are six elements. The index $i$ in eq.(3B.1) correspons to these six elements. The coefficients for these elements are as in the following table.

|  |  | $j$ |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|
|  | normalization const. | 1 | 2 | 3 | 4 | 5 | 6 | expression |
| $\alpha_j^{(1)}$ | $1/\sqrt{6}$ | 1 | 1 | 1 | 1 | 1 | 1 | $A_1$ |
| $\alpha_j^{(2)}$ | 1/2 | 1 | 1 | $-1$ | $-1$ | 0 | 0 | $E$ |
| $\alpha_j^{(3)}$ | 1/2 | 1 | 1 | 0 | 0 | $-1$ | $-1$ | $E$ |
| $\alpha_j^{(4)}$ | $1/\sqrt{2}$ | 1 | $-1$ | 0 | 0 | 0 | 0 | $T_1$ |
| $\alpha_j^{(5)}$ | $1/\sqrt{2}$ | 0 | 0 | 1 | $-1$ | 0 | 0 | $T_1$ |
| $\alpha_j^{(6)}$ | $1/\sqrt{2}$ | 0 | 0 | 0 | 0 | 1 | $-1$ | $T_1$ |

**Tab.** 3.2 Linear combination coefficients for the donor states in Si.

Then we index the donor states with the quantum number of valley wavefunction $chi_j$ and the above reduced representation. Here the qunantum number (qn) of $\chi_j$ is determined by main qn $n$, directional qn $l$, magnetic qn $m$ but with anisotropy. Hence the indices are like $1s(A_1)$, $1s(E)$ etc.

From Tab. 3.2, we see that in the elements other than $A_1$, the wavefunctions are superposed in the inverse phase and the amplitude at the origin is small. In $A_1$, all the wavefunctions are superposed in phase and the amplitude at the origin is large. Hence the state of $1s(A_1)$ largely deviates from the effective mass approximation and that causes large decrease in the eigenenergy. The larger the positive charge in the nuclear the larger the binding energy. This eneryg splitting is called **valley-orbit splitting**.

| Effective mass theory | Li | P | As | Sb | Bi |
|---|---|---|---|---|---|
| 32 | 32.5 | 45 | 53.7 | 43 | 70.6 |

**Tab.** 3.3 Donor binding energy in Si (meV)

As in Tab. 3.3, such tendency really appears in $1s(A_1)$. Pantelides and Sah gave theoretical calculation of the valley-orbit splitting, which reproduces the experiments well[4].

## References

[1] C. Hamaguchi, “Basic Semiconductor Physics" 3rd ed. (Springer, 2017).

[2] P. Yu and M. Cardona, “Fundamentals of Semiconductors", (4th ed. Springer, 2010).

[3] M. Tinkham, “Group Theory and Quantum Mechanics" (Dover, 2003).

[4] S. Pantelides and C. T. Sah, Phys. Rev. B 10, 621-637 (1974); *ibid.* 638-658 (1974).