When $k = \pm\frac{n\pi}{a}$, the free electron energy is given by Eqn (8.34)

$$\text{i.e., } E = \frac{\hbar^2 n^2 \pi^2}{2ma^2} = \frac{n^2 h^2}{8ma^2}. \tag{8.35}$$

This is the same energy as that for a free particle in a potential well of width a.

If the size of the well is replaced by the distance between the atoms, say 0.3 nm, then from eqn (8.35),

$$E = 4n^2 \text{ eV}. \tag{8.36}$$

This shows that beyond a few discontinuities in the $(E-k)$ curve, $E \to \infty$. The regions beyond the discontinuities are called *Brillouin zones*.

8.7 Brillouin Zones in Two Dimensions

The condition for any discontinuity is

$$k = \pm\frac{n\pi}{a}. \tag{8.37}$$

For *two-dimensional case*, the corresponding conditions is

$$k_x n_1 + k_y n_2 = \frac{\pi}{2}\left(n_1^2 + n_2^2\right),$$

where n_1 and n_2 are integers relating to x- and y-axis respectively. To sketch the first zone n_1 and n_2 are made in turn to ±1 or 0. The borders of the first zone are

$$n_1 = +1, \quad n_2 = 0, \quad \text{so,} \quad k_x = \pm\frac{\pi}{a}$$

$$n_1 = 0, \quad n_2 = \pm1, \quad \text{so,} \quad k_y = \pm\frac{\pi}{a}.$$

So, the *first Brillouin zone* is a square passing through the points A, B, C and D [Fig. 8.13].
For the *2nd Brillouin zone*, we have

$$n_1 = +1, \quad n_2 = +1, \quad \text{giving} \quad k_x + k_y = \frac{2\pi}{a}$$

$$n_1 = -1, \quad n_2 = +1, \quad \text{giving} \quad -k_x + k_y = \frac{2\pi}{a}$$

$$n_1 = +1, \quad n_2 = -1, \quad \text{giving} \quad k_x - k_y = \frac{2\pi}{a}$$

$$n_1 = -1, \quad n_2 = -1, \quad \text{giving} \quad -k_x - k_y = \frac{2\pi}{a}.$$

The above equations describe a set of four lines at 45° to the k_x and k_y axes and passing through the points P, Q, R and S [Fig. 8.13]. Thus, the *second Brillouin zone* is the area between P, Q, R, S and A, B, C, D. Putting n_1 and $n_2 = 0, \pm1$ and ±2, the *third Brillouin zone* may be obtained [Fig. 8.13].
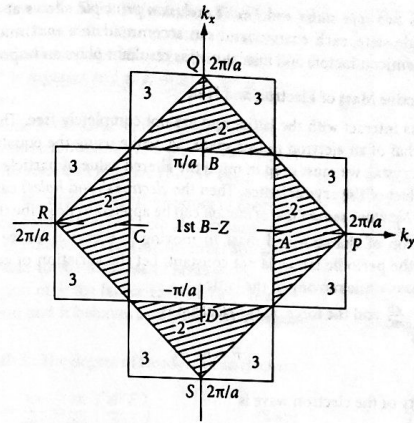
Fig. 8.13: Brillouin zones for a square lattice

8.7.1 Number of Wave Functions in a Band

We have seen that in an infinitely long one-dimensional crystal there are certain allowed energy ranges which have a continuous distribution of energy.

For a linear crystals of length L, the periodic boundary condition is $\psi(x+L) = \psi(x)$.

As the wave function are *Bloch functions*, we have $e^{iK(x+L)} u_k(x+L) = e^{iKx} u_k(x)$.

Since $u_k(x+L) = u_k(x)$.

$\therefore e^{iK(x+L)} = e^{iKx}$ or, $e^{iKL} = 1$.

$$\therefore k = \frac{2\pi n}{L} \tag{8.38}$$

with $n = \pm1, \pm2, \ldots$

$$\text{or, } k = \pm\frac{2\pi}{L}, \pm\frac{4\pi}{L}, \ldots. \tag{8.39}$$

So, the number of possible wave function in range dk is

$$dk = \frac{2\pi}{L} dn, \quad \text{[from eqn (8.38)]}$$

$$\text{or, } dn = \frac{L}{2\pi} dk. \tag{8.40}$$

In the first Brillouin zone, the maximum value of k is π/a, where a is the length of the *primitive cell*. Again, if in a crystal of length L, the number of primitive cells is N, then $a = L/N$. So, the maximum value of k in the *first Brillouin zone* is $N\pi/L$. This means that the series in eqn (8.39) terminates at $N\pi/L$. Thus, the total number of allowed k-values in the *first Brillouin zone* is N. Hence, the total number of possible states in any energy band is equal to the number of primitive unit cells N.

Since an electron has *two spin states* and *Pauli's exclusion principle* allows at best two electrons with opposite spins in a single state, each energy band can accommodate a maximum of $2N$ electrons. For distinguishing metals, semiconductors and insulators this resultant plays an important role.

8.7.2 Concept of Effective Mass of Electron and Hole

In a crystal the electrons interact with the lattice and are not completely free. The wave particle motion, therefore, differs from that of an electron in free space. So, while using the equations of electrodynamics for charge carriers in a crystal, we must keep in mind the altered value of particle mass called the *effective mass* (m^*) due to the effect of the crystal lattice. Then the *electrons* (and *holes*) can be regarded as almost free charge carriers and *Newton's second law of motion* can be applied to describe their dynamic behaviour.

We consider an electron of charge e and mass m moving in a crystal lattice of electric field E. The acceleration $a = \frac{eE}{m}$ in the periodic lattice is not constant. Let the variation of acceleration is due to the variation of electron's mass while moving in the lattice.

$\therefore$ the acceleration $a = \frac{eE}{m^*}$ and the force on the electron is

$$F = m^* a. \qquad (8.41)$$

Now the group velocity of the electron wave is

$$v_g = \frac{d\omega}{dk} = 2\pi \frac{dv}{dk} = \frac{2\pi}{h}\frac{dE}{dk} = \frac{1}{\hbar}\frac{dE}{dk}, \qquad (8.42)$$

where $E = hv$ and $\hbar = h/2\pi$.

Acceleration

$$a = \frac{dv_g}{dt} = \frac{1}{\hbar}\frac{d^2E}{dk\,dt} = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt}.$$

But $\hbar k = p$ and $\frac{dp}{dt} = F$.

$$\therefore \frac{dk}{dt} = \frac{F}{\hbar}.$$

So

$$a = \frac{1}{\hbar^2}\frac{d^2E}{dk^2}F \text{ or, } F = \frac{\hbar^2}{\frac{d^2E}{dk^2}}a. \qquad (8.43)$$

Comparing eqns (8.41) and (8.43), we get

$$m^* = \frac{\hbar^2}{\frac{d^2E}{dk^2}}. \qquad (8.44)$$

Thus, determining the value of $\frac{d^2E}{dk^2}$, we can calculate the value of the *effective mass* m^*.

(i) Variation of E with k: The variation of E with k for the first allowed band is shown in Fig. 8.14 (a). Using this type of variation of E with k, the group velocity $v_g = v$ can be calculated using eqn (8.42).

(ii) Variation of v with k: Fig. 8.14 (b) shows the variation of v with k. For $k = 0$, $v = 0$ and with the increase of k (i.e., energy), v increases and reaches a maximum value at $k = k_0$, where k_0 is the value of k at the point of inflection of the $(E - k)$ curve. Beyond k_0, v begins to decrease and finally assumes the zero value at $k = \pi/a$. The negative value of k shows a similar behaviour.

(iii) The variation of m^* with k: Fig. 8.14 (c) shows the variation of m^* with k. Around $k = 0$, m^* increases, reaches its maximum value at the point of inflection of the $(E - k)$ curve. Above the point of inflection, m^* is *negative* and as $k \rightarrow \pi/a$, m^* reduces to a small negative value.

The *negative mass* is due to the force on the electron (caused by *Bragg reflection*) in one direction leading to a gain in momentum in opposite direction. This results in *negative effective mass*.

The equation $E = \frac{\hbar^2 k^2}{2m}$ is satisfied near the bottom of the band and hence m^* has a constant value. But with the increase of k, m^* increases. Fig. 8.14 (b) shows that beyond the inflection point k_0, velocity decreases due to the decrease of acceleration. This implies that in this region of k the lattice exerts a large retarding force on the electron and it behaves as a positively charged particle called *hole*.

(iv) Variation of f_k with k: The degree of freedom of an electron is defined as

$$f_k = \frac{m}{m^*} = \frac{m}{\hbar^2}\left(\frac{d^2E}{dk^2}\right). \qquad (8.45)$$

f_k gives a measure of extent to which the electron is free. When m^* is large, f_k is small, the particle, then behaves as a *heavy particle*. If $f_k = 1$, *the electron behaves as a free electron*.
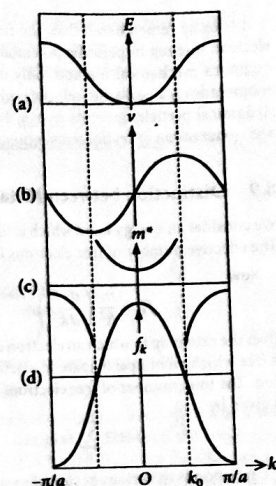

Fig. 8.14: E, m^* and f_k as a function of k

8.8 Crystal Momentum

The momentum associated with the motion of a free electron is $\hbar \vec{k}$, where $\vec{k}$ is the wave vector. This momentum of free electron is called the *true momentum* which helps to understand the dynamical behaviour of it. For example, the energy of the free electron is

$$E = \frac{\hbar k}{2m} = \left(\frac{1}{2m}\right)\left(\frac{h^2}{4\pi^2}\right)\left(\frac{4\pi^2}{\lambda^2}\right) = \left(\frac{1}{2m}\right)\left(\frac{h^2}{\lambda^2}\right) = \frac{1}{2m}p^2. \qquad (8.46)$$

This $\vec{k}$ was first introduced by *Sommerfeld* while formulating the theory of free electrons.

However, when an electron moves in a *periodic potential*, $\hbar\vec{k}$ does not represent the true momentum as the energy does not vary with $\vec{k}$ in the manner as it does in case of a free electron and as such known as the *crystal momentum*.

For an electron moving in a periodic potential, $\hbar\vec{k}$ is not the true electron momentum. This is also confirmed by the fact that, the Bloch wave functions which are eigenstate of the Hamiltonian, are not simultaneously eigenstates of the momentum operator

$$\vec{p}\psi_k(\vec{r}) = \frac{\hbar}{i}\nabla\psi_k(\vec{r}) = \frac{\hbar}{i}\nabla\left[e^{i\vec{k}\cdot\vec{r}}u_k(\vec{r})\right] = \hbar k\psi_k(\vec{r}) + e^{i\vec{k}\cdot\vec{r}}\frac{\hbar}{i}\nabla u_k(\vec{r}). \qquad (8.47)$$

$\therefore \vec{p}\psi_k(\vec{r})$ is not simply $\hbar k\psi_k(\vec{r})$ still $\hbar\vec{k}$ is an extension of momentum $\vec{p}$ to the periodic potential of the crystal. This is why $\hbar\vec{k}$ for a lattice is known as the *crystal momentum*, $\vec{k}$ being the wave vector introduced in *Bloch theorem*.

It is to be remembered that, the free electron momentum $\vec{p} = \hbar \vec{k}$ is a constant of motion but for an electron moving in periodic potential $\hbar \vec{k}$ is not constant of motion and cannot be calculated directly by quantum mechanical method, only its average value can be calculated. On the other hand, the crystal momentum is a perfectly well-defined constant for a state of given energy just as the true momentum of a dynamical particles gives its energy. For this reason, *in dealing with interaction of electron with lattice, we use conservation of crystal momentum instead of true electron momentum.*

8.9 Distinction between Metals, Insulators and Semiconductors

We consider an energy band which is filled with electrons up to a value k_1 $(k_1 < \pi/a)$ as shown in Fig. 8.15. The effective number of free electrons in the band is responsible to the conductivity of this band.

Now

$$f_k = \frac{m}{\hbar^2}\left(\frac{d^2E}{dk^2}\right)$$

gives the extent up to which an electron in the k state is free which would participate in electron conduction. The total number of free electrons in the band is given by

$$N_{\text{eff}} = \sum_k f_k, \qquad (8.48)$$

$\sum_k$ corresponds to various occupied energy states in the band.

Fig. 8.15: Energy band filled up to the state k_1

Now the number of possible states in k to $k + dk$ for a one-dimensional lattice of length L is

$$dn = \frac{L}{2\pi}\,dk.$$

∴ the total number of free electrons present within the limit $-k_1$ to $+k_1$ is

$$N_{\text{eff}} = 2 \times \frac{1}{2\pi}\int_{-k_1}^{+k_1} f_k\, dk = \frac{L}{\pi}\int_{-k_1}^{k_1} f_k\, dk. \qquad (8.49)$$

The factor 2 on the right side of eqn (8.49) is due to the fact that each state can occupy a maximum of two electrons with opposite spin.

$$\therefore N_{\text{eff}} = \frac{L}{\pi}\int_{-k_1}^{k_1} \frac{m}{\hbar^2}\left(\frac{d^2E}{dk^2}\right)dk, \quad \left[\because f_k = \frac{m}{\hbar^2}\left(\frac{d^2E}{dk^2}\right)\right]$$

$$= \frac{2L}{\pi}\frac{m}{\hbar^2}\int_0^{k_1} \frac{d^2E}{dk^2}\,dk = \frac{2Lm}{\pi\hbar^2}\left(\frac{dE}{dk}\right)_{k=k_1}. \qquad (8.50)$$

The behaviour of solids as *insulator, semiconductor* and *conductor* can be explained on the basis of *Band Theory* as follows:
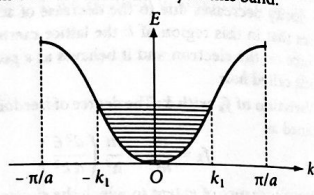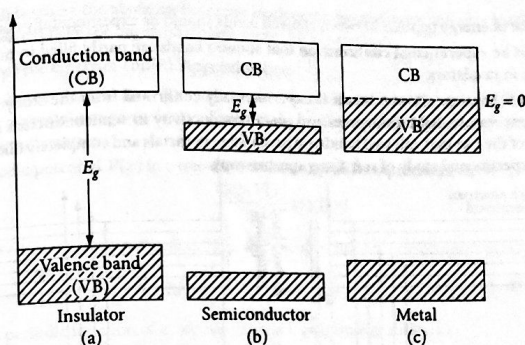
Fig. 8.16: Distribution of electrons in insulator, semiconductor and metal. The shaded areas are occupied by electrons

(a) **Insulators:** From the $(E$-$k)$ curve, it is clear that, at the top (and bottom) of the band $\frac{dE}{dk} = 0$ as $k = 0$, π/a. So from eqn (8.50), $N_{\text{eff}} = 0$. Thus when the valence band is completely filled (and the conduction band is completely empty) the solid behaves as a *perfect insulator* [Fig. 8.16(a)]. There is a *forbidden energy gap or band gap* $\left(E_g\right)$ between the valence band and conduction. For example, diamond is a good insulator having band gap $E_g \sim 6\,\text{eV}$. As the electrons in the valence band ordinarily do not possess this amount of energy they cannot transit to the conduction band and participate in the electrical conduction process. As a result, the solid behaves as an insulator. *All solids with large band gap* $\left(E_g\right)$ *are good electrical insulators.*

(b) **Semiconductors:** There is a class of solids like Germanium (Ge) and Silicon (Si), etc. for which the band gap E_g is relatively small $\sim 1\,\text{eV}$. It is, therefore, possible for some electrons at room temperature to have thermal energy to jump across the forbidden gap into the empty conduction band. In presence of an electric field, these electrons constitute a limited current. The conductivity of such materials lies between good conductor and insulator [Fig. 8.16(b)].

(c) **Conductors (Metals):** Since the value of $\frac{dE}{dk}$ is the maximum at the point of inflection of $(E - k)$ curve Fig. 8.15, the effective number of electrons N_{eff} is the maximum. So, in a material with completely filled valence band, there is a large number of the electrons available for conduction. Such materials are conductors like all metals. The valence band and conduction band in conductors partially overlap so that the band gap $E_g = 0$, Fig. 8.16 (c).

8.9.1 Effect of Temperature

In case of *good conductors* (metals) with the increase in temperature, more and more phonons are excited which scatter electrons and reduce their mobility. This, in turns, leads to the increase in resistivity and decrease the conductivity of metals. But in case of semiconductors, increase in temperature causes more and more electrons in the valence band to have sufficient energy to jump into the conduction band. Hence, *with the rise of temperature the conductivity of the semiconductors increases. This is an important distinction between metals and semiconductors.*

8.10 Direct Experimental Evidence for Band Structure in Solids

The essential requirement for the direct evidence of the band structure of the energy levels in solids are:

(i) The existence of energy gap ΔE between allowed bands should be experimentally confirmed.

(ii) There should be experimental confirmation that allowed bands are partly filled in metals and completely filled in insulators.

The existence ΔE between allowed bands is experimentally confirmed from the study of variation of resistivity with temperature, light absorption and photo-conductivity in semiconductors and insulators. The confirmation of the fact that, allowed bands are partly filled in metals and completely filled in insulators comes from the experimental study of soft X-ray spectroscopy.
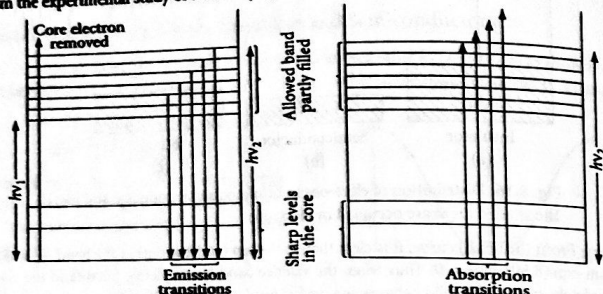

Fig. 8.17: Various transitions in soft X-ray spectrocopy

In Fig. 8.17, A represents the energy level of a core electron, say, $1s$ (K-shell) state. This type of core levels are sharp. Let the bombardment of a suitable cathode ray electron to the target atom, the core electron is removed and a vacancy is created there. This vacancy in the $1s$ state can be filled by the transition of an electron in the higher allowed band which is partly filled [Fig. 8.17 (a)].

Since there are no energy levels in higher band from where transition would take place to v_2, there will be no emission, if $v > v_2$. Sudden emissions of all transitions will occur behaving as emission edge within $v_1 < v < v_2$. For $v < v_1$, no emission will occur due to the absence of transition levels. The observed emission bandwidth is a measure of the extent up to which the allowed band is filled.

Fig. 8.17 (a) shows the various absorption transitions of X-rays in the same specimen.

For $v < v_2$, there is no absorption of energy but for $v = v_2$, there is an absorption edge. For $v > v_2$, an absorption band arises. Hence, the existence of the unfilled levels in the allowed band is evidenced by the absorption of energy. In case of metals, the absorption and emission edges coincide in frequency.

The emission bands for non-metals are different from those of metals and have no sharp edges. In case of metals, the transitions from filled to unfilled levels occur over a region proportional to kT at the Fermi levels and hence, are not perfectly sharp. This is why the width of the emission edge is temperature dependent.

8.11 Limiting Cases of Periodic Potential

In the foregoing discussion of *Kronig-Penny model*, Sommerfeld's free electron theory is used which is poorly related to the actual form of potential. The effect of ion cores cannot be totally ignored. Therefore, two limiting cases arise where the periodic potential is either very weak or very strong. In case of very strong periodic potential each electron is almost bound to the minimum of the potential and the rest of the lattice is as being a perturbation to this minimum. This is *tight binding approximation*.

In case of solids, where the above methods are inadequate, we use more complex methods like *orthogonalised plane wave* (OPW) or *augmented plane wave* (APW) *methods*.

8.11.1 Nearly Free Electron (NFE) Approximation

Here the crystal potential is assumed to be very weak in comparison to the electron KE E, and the effect of the periodic potential as a *perturbation*. Under these conditions the allowed bands will be broad and the forbidden gaps quite narrow.

Let the periodic potential $V(x)$ in a one-dimensional crystal be represented as

$$-\frac{2m V(x)}{\hbar^2} = \gamma f(x).$$ (8.51)

$f(x)$ represents the periodicity of lattice. Then the *one-dimensional Schrödinger's equation* may be written as

$$\frac{\partial^2 \psi}{\partial x^2} + \left[\frac{2m}{\hbar^2}E + \gamma f(x)\right]\psi(x) = 0 \text{ or, } \frac{\partial^2 \psi}{\partial x^2} + \left[K_0^2 + \gamma f(x)\right]\psi(x) = 0.$$ (8.52)

As $f(x)$ is a periodic function of x, we can express it as a Fourier series, i.e.,

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{-\frac{2\pi i n x}{a}}$$ (8.53)

with

$$C_n = \frac{1}{a}\int_0^a f(x) e^{\frac{2\pi i n x}{a}} dx,$$ (8.54)

where a is the lattice constant of the crystal.

If V_n be the Fourier expansion coefficient of $V(x)$, then according to eqn (8.51)

$$V_n = -\frac{\hbar^2}{2m}\gamma C_n.$$ (8.55)

The wave function $\psi(x)$ must have the Bloch form

$$\psi(x) = e^{iKx} U(x) = e^{iKx} \sum_{n=-\infty}^{\infty} b_n e^{-\frac{2\pi i n x}{a}}.$$ (8.56)

where $U(x)$ is a periodic function of a, it can be expressed as a Fourier series as

$$U_k(x) = \sum_{-a}^{\infty} b_n e^{-\frac{2\pi i n x}{a}}.$$ (8.57)

In case the electron is completely free, $\gamma = 0$ and $n = 0$, when

$$\psi(x) = b_0 e^{iKx}.$$ (8.58)

In the limit $U_k(x) \to b_0$ and $K \to K_0$. So, as $x \to 0$ all b_n (except b_0) approach zero. Then the approximate expression for the wave function is of the form:

$$\psi(x) = b_0 e^{iKx} + \gamma \left[e^{iKx}\sum_{n\neq 0} b_n e^{-\frac{2\pi i n x}{a}}\right]$$ (8.59)

Eqn (8.59) is valid for small values of γ

$$\text{or, } \psi(x) = \left[b_0 + \gamma\sum_{n\neq 0} b_n e^{-\frac{2\pi i n x}{a}}\right]e^{iKx} = U(x) e^{iKx}.$$ (8.60)

where

$$U(x) = b_0 + \gamma \sum_{n \neq 0} b_n e^{-\frac{2\pi n i x}{a}}.$$

Substituting eqn (8.59) in eqn (8.52), we get

$$b_0 \left(K_a^2 - K^2\right) e^{iKx} + \gamma \sum_{n \neq 0} \left[(K_0^2 - K_a^2) b_n + b_0 c_n\right] e^{iK_n x} + \gamma^2 \sum_{n \neq 0} \sum_{n' \neq 0} b'_n C_n e^{i\left(K - \frac{2\pi n}{a} - \frac{2\pi n'}{a}\right)x} = 0, \quad (8.61)$$

where

$$K_n = -K - \frac{2\pi n}{a}. \quad (8.62)$$

As a first approximation, we neglect γ^2 term in eqn (8.61). Then multiplying the equation by $e^{-iK_m x}$ and integrating over unit cell from $x = 0$ to $x = a$, we have

$$b_0 \left(K_0^2 - K^2\right) \int_0^a e^{\frac{2\pi i n x}{a}} dx + \gamma \sum_{n \neq 0} \left[\left\{(K_0^2 - K_n^2) b_n + b_0 c_n\right\} \times \int_0^{a_0} e^{\frac{2\pi i (m-n) x}{a}} dx\right] = 0. \quad (8.63)$$

For $m = 0$,

$$\int_0^a e^{\frac{2\pi i (m-n) x}{a}} dx = 0.$$

for all values of n in the summation. Then

$$b_0 \left(K_0^2 - K^2\right) = 0, \text{ or, } K = K_a. \quad (8.64)$$

Again, for $m \neq 0$,

$$\int_0^a e^{\frac{2\pi i m x}{a}} dx = 0$$

and the second integral in eqn (8.61) gives zero except $m = n$.
So

$$\gamma \left[(K_0^2 - K_m^2) b_m + b_0 C_m\right] a = 0 \text{ or, } b_m = \frac{b_0 C_m}{K_m^2 - K_0^2} = \frac{b_0 C_m}{K_m^2 - K^2}. \quad (8.65)$$

$\therefore$ from eqn (8.64),

$$K^2 = K_0^2, \text{ or, } \frac{2mE}{\hbar^2} = K_0^2 \text{ or, } E = \frac{K_0^2 \hbar^2}{2m} \quad (8.66)$$

The relation is the same as that for a free particle.
Now substituting the value of b_m from eqn (8.65) into eqn (8.60), we get

$$\psi(x) = b_0 e^{iKx} \left[1 + \gamma \sum_{n \neq 0} \frac{C_n}{K_n^2 - K^2} e^{-\frac{2\pi i n x}{a}}\right]. \quad (8.67)$$

Eqn (8.64) shows that the contribution of the first order correction to the energy of a free particle due to the periodic potential is zero. To obtain second order energy correction, we multiply eqn (8.61) by e^{-iKx} and then integrate from $x = 0$ to $x = a$,

$$ab_0 \left(K_0^2 - K^2\right) + \gamma \sum_{n \neq 0} \left[(K_0^2 - K_n^2) b_n + b_0 C_n\right] \int_0^a e^{-\frac{2\pi i n x}{a}} dx + \gamma^2 \sum_{n \neq 0} \sum_{n' \neq 0} b'_n C_n \int_0^a e^{-\frac{2\pi i (n+n') x}{a}} dx = 0. \quad (8.68)$$

Now

$$\int_0^a e^{-\frac{2\pi i n x}{a}} dx = 0,$$

for all allowed values of n and

$$\int_0^a e^{-\frac{2\pi i (n+n') x}{a}} dx = 0, \text{ if } n \neq -n'.$$

Thus, eqn (8.68) reduces to

$$b_0 \left(K_0^2 - K^2\right) a + \gamma^2 \sum_{n' \neq 0} b'_n C'_n a = 0. \quad (8.69)$$

Again, replacing n by $-n$ in eqn (8.53), we get

$$f(x) = \sum_{n=-\infty}^{\infty} C_{-n} e^{\frac{2\pi n i x}{a}}. \quad (8.70)$$

The complex conjugate of eqn (8.53) is

$$f^*(x) = \sum_{n=-\infty}^{\infty} C_n^* e^{\frac{2\pi n i x}{a}} \text{ or, } f(x) = \sum_{n=-\infty}^{n=\infty} C_n^* e^{\frac{2\pi n i x}{a}}. \quad (8.71)$$

Since $f^*(x) = f(x)$. Comparing eqns (8.70) and (8.71), we get

$$C_{-n} = C_n^*. \quad (8.72)$$

Now using eqns (8.65) and (8.72) in eqn (8.69), we get

$$K_0^2 = K^2 + \gamma^2 \sum_{n \neq 0} \frac{C_n^* C_n}{K^2 - K_n^2} = K^2 + \gamma^2 \sum_{n \neq 0} \frac{C^* C_n}{K^2 - \left(K - \frac{2\pi n}{a}\right)^2} \quad \left[\because K_n = K - \frac{2\pi n}{a}\right]. \quad (8.73)$$

The above relation can be expressed in terms of $E \left(= \frac{\hbar^2 K^2}{2m}\right)$ as follows:

$$E = \frac{\hbar^2 K^2}{2m} + \sum_{n \neq 0} \frac{|V_n|^2}{\left(\frac{\hbar^2 K^2}{2m}\right) - \frac{\hbar^2}{2m}\left(K - \frac{2\pi n}{a}\right)^2}. \quad (8.74)$$

If $K^2 \approx K_n^2$ for a particular value of n, then one of the quantities

$$\left(K^2 - K_n^2\right) \text{ in } \sum_{n \neq 0} \frac{C_n}{K_n^2 - K^2} e^{-\frac{2\pi n i x}{a}}$$

of eqn (8.67) tends to zero and the corresponding term will tend to infinity despite γ is very small. Then the eqn (8.59) is not a suitable form of the wave function $\psi(x)$. Further, if $K^2 \approx K_n^2$, then

$$K = \pm K_n = \pm K \mp \frac{2\pi n}{a}. \quad (8.75)$$

Since, for the upper sign, this equation gives $n = 0$ and hence excluded from eqn (8.67). For the lower sign, we get

$$K = \frac{n\pi}{a} \quad (8.76)$$

Eqn (8.62) refers to all the band edge points as n may assume both positive as well as negative values. Now let $K \approx n\pi/a$, whence $K_n = K - 2\pi n/a \simeq -\frac{n\pi}{a}$. Then $K_n^2 \simeq K^2$ and b_n will not be very large. Under this condition γb_n will not be very small. The wave function in this case can be written as

$$\psi(x) = b_0 e^{iKx} + \gamma b_n e^{iKx} e^{-\frac{2\pi inx}{a}} = b_0 e^{iKx} + \gamma b_n e^{iK_n x}. \qquad (8.77)$$

As $K_n = -K$ at the band edge, $\psi(x)$ in eqn (8.77) in the result of superposition of a wave travelling along the positive x-direction and the other along the negative x-direction, i.e., $\psi(x)$ represents a standing wave.

Now putting

$$f(x) = \sum_{n=-\infty}^{+\infty} C_n e^{-\frac{2\pi inx}{a}}$$

and $\psi(x)$ as given by eqn (8.77) in eqn (8.52), we get

$$b_0^2 \left(K_0^2 - K^2\right) e^{iKx} + \gamma b_n \left(K_0^2 - K^2\right) e^{i\left(K - \frac{2\pi n}{a}\right)x} + \gamma b_0 \sum_{n' \neq 0} C_n e^{i(K - 2\pi n'/a)x}$$

$$+ \gamma^2 b_n \sum_{n' \neq 0} C_n^* e^{i\left(K - \frac{2\pi n}{a} - \frac{2\pi n'}{a}\right)x} = 0. \qquad (8.78)$$

Multiplying eqn (8.78) by e^{-iKx} on both sides and then integrating within limits, $x = 0$ to $x = a$, we find that the contributions of the second and third terms become zero and that from the fourth term is obtained only for $n' = n$. Finally,

$$\left(K_0^2 - K^2\right) b_0 + \gamma^2 C_n^* b_n = 0. \qquad (8.79)$$

Similarly, multiplying both sides of eqn (8.78) by

$$e^{-iK_n x} = e^{-i(k - 2\pi n/a)x}$$

and then integrating from $x = 0$ to $x = a$, we get

$$C_n b_0 + \left(K_0^2 - K_n^2\right) b_n = 0. \qquad (8.80)$$

The eqns (8.79) and (8.80) are a set of homogeneous equations whose solutions other than $b_0 = b_n = 0$ exist, if

$$\begin{vmatrix} \left(K_0^2 - K^2\right) & \gamma^2 C_n^* \\ C_n & K_0^2 - K_n^2 \end{vmatrix} = 0 \text{ or, } \left(K_0^2 - K^2\right)\left(K_0^2 - K_n^2\right) - \gamma^2 C_n C_n^* = 0. \qquad (8.81)$$

This is a quadratic equation having solution

$$K_0^2 = \frac{1}{2}\left[\left(K^2 + K_n^2\right) \pm \sqrt{\left(K^2 - K_n^2\right) + 4\gamma^2 C_n^* C_n}\right] \qquad (8.82)$$

Using the relations, $E = \frac{\hbar^2 K_0^2}{2m}$.

Eqn (8.82) can be expressed as

$$E(k) = \frac{\hbar^2}{4m}\left[K + \left(K - \frac{2\pi n}{a}\right)^2 \pm \sqrt{\left\{K^2 - \left(K - \frac{2\pi n}{a}\right)^2\right\}^2 + \left\{\frac{4m|V_n|^2}{\hbar^2}\right\}}\right]. \qquad (8.83)$$

At the band edge, $K - K_n = \frac{n\pi}{a}$, where eqn (8.83) reduces to

$$E = E_n \pm |V_n| \qquad (8.84)$$

with

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{a}\right)^2 \qquad (8.85)$$

is the energy of a free particle at the band edge points.

The results of the foregoing discussion can be summarized as below:

(i) There exists an energy gap or discontinuity in the E versus K curve due to the internal Bragg reflection at the band edge points $K = \pm\frac{n\pi}{a}$. Eqn (8.81) shows that the energy gap for the forbidden band is of width $2|V_n|$, V_n being the nth Fourier coefficient in the Fourier expansion of the lattice potential.

(ii) For $|K| > \frac{n\pi}{a}$, $E(K)$ is greater than the value given by eqn (8.64) where $\left(K^2 - K_n^2\right)$ is sufficient to reduce γb_n sufficiently small. Similarly, for $|K| < \frac{n\pi}{a}$, $E(K)$ must be less than the value given by eqn (8.84). These conditions will be satisfied, if the positive sign and negative sign be chosen for $|K| > \frac{n\pi}{a}$ and $|K| < \frac{n\pi}{a}$ respectively in eqn (8.84).
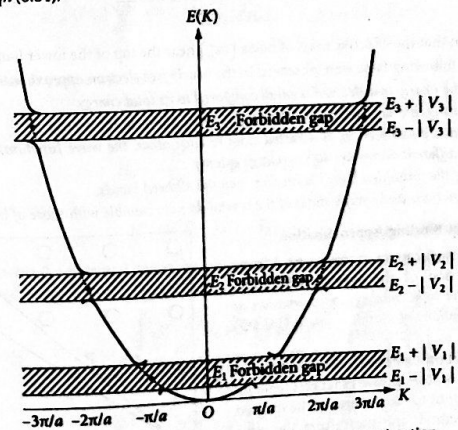

Fig. 8.18: E versus K curve for free electron approximation

(iii) The schematic representation of variation of $E(K)$ with K for an electron is shown in Fig. 8.18 as obtained from eqns (8.71) and (8.80). Fig. 8.18 shows that the E versus K curves are nearly parabolic near the band edge points.

Effective Mass

The effective mass of the electron $\left(m_e^*\right)$ can be calculated by taking

$$K = \frac{n\pi}{a} + K',$$

where $K' \ll \pi/a$ in eqn (8.83).

$$\therefore E(K) = \frac{\hbar^2}{2m}\left[\left(\frac{n\pi}{a}\right)^2 + K'^2 + m\frac{\Delta E}{\hbar^2}\sqrt{1 - 4K'^2\left(\frac{n\pi}{a}\right)^2\left(\frac{\hbar^2}{m\Delta E}\right)^2}\right],$$ (8.86)

where

$$\Delta E = 2|V_n|.$$ (8.87)

For points near the bottom of the upper band K' is very small. From eqn (8.86) by expanding in binomial theorem and neglecting higher order terms we get

$$E(K) = E_n + \frac{1}{2}\Delta E + \frac{\hbar^2 K'^2}{2m}\left(1 + \frac{4E_n}{\Delta E}\right).$$ (8.88)

Differentiating twice, we get

$$\frac{d^2E}{d^2k'} = \frac{\hbar^2}{m}\left(1 + \frac{\Delta E_n}{\Delta E_n}\right).$$

$$\therefore m_e^* = \frac{\hbar^2}{\frac{d^2E}{dK^2}} = \frac{\hbar^2}{\frac{d^2E}{dK^2}} = \frac{m}{1 + \frac{4E_n}{\Delta E}}.$$ (8.89)

It can be shown that the effective mass of holes $\left(m_h^*\right)$ near the top of the lower band is the same as m_e^*.
In conclusion, following facts were observed in the *nearly free electron approximation method*:

(i) *The PE of the nearly free electron is small compared to its total energy.*
(ii) *There is a strong interaction between the neighbouring atoms in the crystal.*
(iii) *As the atoms are assumed to be situated close to each other, the wave functions of the neighbouring electrons in different atoms overlap to a large extent.*
(iv) *The width of the forbidden bands is smaller than the allowed bands.*
(v) *The wave functions and energy states of the crystal do not resemble with those of individual atoms.*

8.11.2 The Tight Binding Approximation

Bloch in 1928 introduced an entirely new scheme for the construction of the crystal wave function known as the *tight binding approximation or the linear combinations of atomic orbitals* (LCAO) method.

This approximation deals with the situations in which the overlap of atomic wave function is enough to require corrections to the pictures of the isolated atoms. In this method, one starts from the wave functions for an electron in a free atom and then constructs a crystal orbital which just equals the LACO's and describes the electron in the periodic field of crystal as a whole.

Fig. 8.19: Two-dimensional square lattice showing the positions of the nth atom and an electron

We consider a two-dimensional square array of atoms in a crystal where $\vec{r}$ and $\vec{r}_n$ represent respectively the position vector of the electron and nth atom respectively as shown in Fig. 8.19.
Fig. 8.20 shows the potentials of an electron in a free atom denoted by $V(\vec{r})$ and that in a crystal by $V_0(\vec{r} - \vec{r}_n)$.
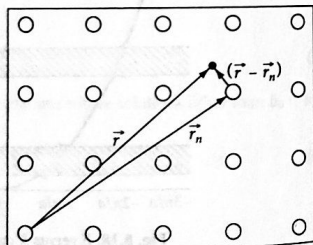
In this method the electron wave function is assumed to be influenced by the nearest atoms and not by the other atoms in the crystal. If $\psi_k(\vec{r} - \vec{r}_n)$ be the wave function for an electron in a free atom and $\psi_k(\vec{r})$ the wave function of the electron in a crystal, then $\psi_k(\vec{r})$ may be expressed as a linear combination as

$$\psi_k(\vec{r}) = \sum_n e^{(i\vec{k}\cdot\vec{r}_n)}\psi_0(\vec{r} - \vec{r}_n).$$ (8.90)

The summation extends over all the atoms in the crystal. It can be shown, the wave function $\psi_k(\vec{r})$ is a *Bloch function*, since

Fig. 8.20: Potential of an electron in an atom $[V(\vec{r})]$ and in a crystal $[V_0(\vec{r} - \vec{r}_n)]$

$$\psi_k(\vec{r} + \vec{r}_m) = \sum_m e^{i(\vec{k}\cdot\vec{r}_n)}\psi_0(\vec{r} + \vec{r}_m - \vec{r}_n) = \sum_n e^{i(\vec{k}\cdot\vec{r}_n)}\psi_0(\vec{r} - \vec{r}_n + \vec{r}_m)$$

$$= e^{i(\vec{k}\cdot\vec{r}_m)}\sum_n e^{i[\vec{k}\cdot(\vec{r}_n - \vec{r}_m)]}\psi_0[\vec{r} - (\vec{r}_n - \vec{r}_m)] = e^{i(\vec{k}\cdot\vec{r}_m)}\psi_k(\vec{r}).$$ (8.91)

The crystal *Hamiltonian* H can be written as

$$H = H_0 + H'$$ (8.92)

with

$$H_0 = -\frac{\hbar^2}{2m}\nabla^2 + V_0(\vec{r} - \vec{r}_n)$$ (8.93)

and

$$H' = V(\vec{r}) - V_0(\vec{r} - \vec{r}_n),$$ (8.94)

where $V_0(\vec{r} - \vec{r}_n)$ represents the potential due to the nth ground state atom on the electron at $\vec{r}$.
The electron energy is

$$\vec{E}(k) = \frac{\int \psi_k^*(\vec{r})\left[H_0 + H'\right]\psi_k(\vec{r})\,d\vec{r}}{\int \psi_k^*(\vec{r})\,\psi_k(\vec{r})\,d\vec{r}}.$$ (8.95)

Now

$$\int \psi_k^*(\vec{r})\,\psi_k(\vec{r})\,d\vec{r} = \sum_m\sum_n e^{i[\vec{k}\cdot(\vec{r} - \vec{r}_m)]} \times \int \psi_0^*(\vec{r} - \vec{r}_m)\,\psi_0(\vec{r} - \vec{r}_n)\,d\vec{r}.$$ (8.96)

In the overlap of the electron wave function between the neighbouring atoms be neglected, then

$$\int \psi_0^*(\vec{r} - \vec{r}_m)\,\psi_{0(\vec{r} - \vec{r}_n)}\,d\vec{r} = \delta_{mn}$$ (8.97)

and

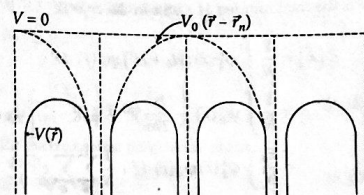$$\int \psi_k^*(\vec{r})\,\psi_k(\vec{r})\,d\vec{r} = N,$$ (8.98)

N is the total number of atoms in the crystal.

$$\therefore E(\vec{k}) = \frac{1}{N}\int \psi_k^*(\vec{r})(H_0 + H')\psi_k(\vec{r})\,d\vec{r} \qquad (8.99)$$

$$= \frac{1}{N}\int \psi_k^*(\vec{r})\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(\vec{r}-\vec{r}_n)\right]\psi_k(\vec{r})\,d\vec{r} + \frac{1}{N}\int \psi_k^*(\vec{r})[V(\vec{r}) - V_0(\vec{r}-\vec{r}_n)]\psi_k(\vec{r})\,d\vec{r}$$

$$= \frac{E_0}{N}\int \psi_k^*(\vec{r})\psi_k(\vec{r})\,d\vec{r} + \frac{1}{N}\sum_m\sum_n e^{i[\vec{k}\cdot(\vec{r}_n-\vec{r}_m)]} \times \int \psi_0^*(\vec{r}-\vec{r}_m)[V(\vec{r}) - V_0(\vec{r}-\vec{r}_n)]\psi_0(\vec{r}-\vec{r}_n)\,d\vec{r}.$$

$$(8.100)$$

Every term in the summation from $n=0$ to $n=N-1$ contains N identical terms which may be found out by putting $r_n = 0$. Then

$$E(\vec{k}) = E_0 + \sum_m e^{-i(\vec{k}\cdot\vec{r}_m)}\int \psi_0^*(\vec{r}-\vec{r}_m)[V(\vec{r}) - V_0(\vec{r})]\psi_0(\vec{r})\,d\vec{r}. \qquad (8.101)$$

If ψ_0 is spherically symmetric (like the ground state of alkali metal), then the contribution due to all the nearest neighbour may be assumed to be the same. The above integral for $\vec{r}_m = 0$ is

$$\int \psi_0^*(\vec{r})[V(\vec{r}) - V_0(\vec{r})]\psi(\vec{r})\,d\vec{r} = -\alpha \qquad (8.102)$$

and for the nearest neighbour atoms

$$\int \psi_0^*(\vec{r}-\vec{r}_n)[V(\vec{r}) - V_0(\vec{r})]\psi_0(\vec{r})\,d\vec{r} = -\beta \qquad (8.103)$$

where $\vec{r}_m$ is a vector connecting the atom at the origin with a nearest neighbour. α and β are constants called the *overlap integrals*. Thus, eqn (8.101) reduces to the form

$$E(\vec{k}) = E_0 - \alpha - \beta\sum_m e^{i(\vec{k}\cdot\vec{r}_m)}. \qquad (8.104)$$

Eqn (8.104) shows that the energy of the electron in the crystal differs from that of the free atom by a constant α plus a term which depends on the wave vector $\vec{k}$. *It is thus the last part which transforms the discrete atomic levels into an energy band in the solid.*

The summation in eqn (8.104) extends only over the nearest neighbour. Since free ion potential $V_0(\vec{r})$ is more negative than the lattice potential $V(\vec{r})$, $[V(\vec{r}) - V_0(\vec{r})]$ is a negative quantity and as such α and β are always positive. Thus, $E < E_0$.

8.11.3 Application of Tight Binding Approximation to Cubic Bravais Lattices

(a) The Simple Cubic (SC) Lattice: In this case there are six neighbours for any given ion at

$$\vec{r}_m = (\pm a, 0, 0),\ (0, \pm a, 0),\ (0, 0, \pm a), \qquad (8.105)$$

where a is the lattice constant.

Therefore

$$E = (E_0 - \alpha) - 2\beta\left(\cos k_x a + \cos k_y a + \cos k_z a\right). \qquad (8.106)$$

(b) The Body-Centred Cubic (BCC) Lattice: The eight nearest neighbours are at

$$r_m = (\pm a/2,\ \pm a/2,\ \pm a/2).$$

Therefore

$$E = (E_0 - \alpha) - 8\beta\cos\left(\frac{k_x a}{2}\right)\cos\left(\frac{k_y a}{2}\right)\cos\left(\frac{k_z a}{2}\right). \qquad (8.107)$$

(c) The Face-Centred Cubic (FCC) Lattice: Here the twelve nearest neighbours are at

$$r_m = \left(0, \pm\frac{a}{2}, \pm\frac{a}{2}\right),\ \left(\pm\frac{a}{2}, 0, \pm\frac{a}{2}\right),\ \left(\pm\frac{a}{2}, \pm\frac{a}{2}, 0\right) \qquad (8.108)$$

and

$$E = (E_0 - \alpha) - 4\beta\left[\cos\left(\frac{k_y a}{2}\right)\cos\left(\frac{k_z a}{2}\right) + \cos\left(\frac{k_z a}{2}\right)\cos\left(\frac{k_x a}{2}\right) + \cos\left(\frac{k_x a}{2}\right)\cos\left(\frac{k_y a}{2}\right)\right]. \qquad (8.109)$$

Thus, the s-band is described by a constant energy term $(E_0 - \alpha)$ and a term depending on the electron wave vector $\vec{k}$.

From eqn (8.106), E is minimum for $k_x = k_y = k_z = 0$.

Thus, the bottom of the energy band is given by

$$E_{bottom} = E_0 - \alpha - 6\beta. \qquad (8.110)$$

The energy maximum corresponds to $k_x = k_y = k_z = \pm\pi/a$ and top of the energy band is given by

$$E_{top} = E_0 - \alpha + 6\beta\ (\text{as } k_x = k_y = k_z = \pm\pi/a). \qquad (8.111)$$

Thus, the width of the energy band is

$$E_{top} - E_{bottom} = (E_0 - \alpha + 6\beta) - (E_0 - \alpha - 6\beta) = 12\beta. \qquad (8.112)$$

This corresponds to the s state of the isolated atom. Eqn (8.112) shows that, the width of energy band increases with β which represents the overlap of the wave functions on the neighbouring atoms.

The inner electronic levels of free atoms gives rise to *narrow bands* in the solid whereas the outer shell electrons give rise to *wider bands*.

Similarly, from eqn (8.107) for the BCC lattice $E_{bottom} = E_0 - \alpha - 8\beta$ and $E_{top} = E_0 - \alpha + 8\beta$.

$\therefore$ energy bandwidth is

$$E_{top} - E_{bottom} = 16\beta \qquad (8.113)$$

and for FCC lattice from eqn (8.109), $E_{bottom} = E_0 - \alpha - 12\beta$ and $E_{top} = E_0 - \alpha + 12\beta$.

$\therefore$ bandwidth

$$E_{top} - E_{bottom} = 24\beta. \qquad (8.114)$$

The variation of energy bandwidth in SC, BCC, FCC lattice is shown in Fig. 8.21.
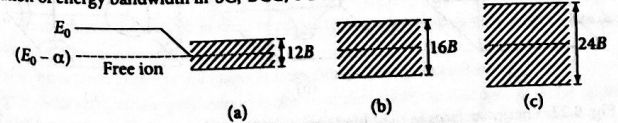

Fig. 8.21: s-bands for (a) SC (b) BCC (c) FCC lattice in tight binding approximation

8.11.4 Energy Surfaces

In tight binding approximation, the PE of the valence electrons in the field of their parent and neighbouring ions is larger than their KE. This indicates the separation between the ions is small compared with the electron in de Broglie wavelength, i.e., $|\vec{k}|a \ll 1$.

$$\therefore \cos ka \simeq 1 - \frac{k^2 a^2}{2}. \tag{8.115}$$

Then

$$E(\vec{k}) \simeq (E_0 - \alpha) - 2\beta\left[\left(1 - \frac{k_x^2 a^2}{2}\right) + \left(1 - \frac{k_y^2 a^2}{2}\right) + \left(1 - \frac{k_z^2 a^2}{2}\right)\right]$$

$$= (E_0 - \alpha) - 6\beta + \beta a^2\left(k_x^2 + k_y^2 + k_z^2\right)$$

or, $E(\vec{k}) = (E_0 - \alpha) - 6\beta + \beta a^2 k^2 + \cdots$ (for SC lattice). $\tag{8.116}$

Similarly, for BCC lattice,

$$E(\vec{k}) = (E_0 - \alpha) - 8\beta + \beta a^2 k^2 + \cdots \tag{8.117}$$

and for FCC lattice,

$$E(\vec{k}) = (E_0 - \alpha) - 12\beta + \beta a^2 k^2 + \cdots. \tag{8.118}$$

8.11.5 Effective Mass of the Electron

The effective mass of the electron is

$$m^* = \frac{\hbar^2}{\partial^2 E/\partial k^2} = \frac{\hbar^2}{2\beta a^2}\left[\because \frac{\partial^2 E}{\partial k^2} = 2\beta a^2 \text{ (for SC lattice)}\right]. \tag{8.119}$$

Thus, $m^* \propto 1/\beta$, this implies that *for a narrow energy band the electron behaves as a very heavy particle in SC lattice.*

From eqn (8.119), the width of the energy band $\left(E_{\text{top}} - E_{\text{bottom}}\right) = 12\beta$ which refers to the bottom region of the band and corresponds to the constant spherical energy surfaces (i.e., $k = 0$) as shown in Fig. 8.22 (a). This means that the lowest energy state is at the centre of the zone and the higher energy states are at the corners [Fig. 8.22 (c)]. With the increase of energy ($k > 0$), the energy surfaces bulge more and more towards the centre of the zone faces, until they intersect the zone boundary. Since the normal derivative of $E(\vec{k})$ is zero, the energy surfaces intersect the zone boundary orthogonally.
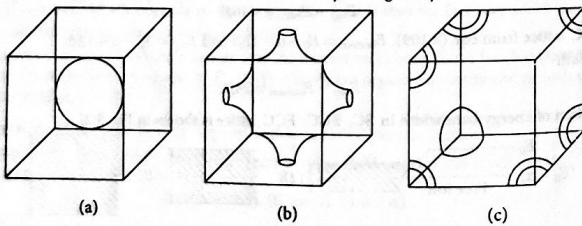

Fig. 8.22: Energy surfaces in tight binding approximation for SC lattice. (a) near the bottom of the band (b) half-way up the band and (c) near the top of the band.

A cross section through the energy surfaces at equal intervals is depicted in Fig. 8.23. Fig. 8.23(a) shows that for small k the energy surfaces are spherical around the zone centre. The energy surfaces are also spherical with corners at the covers of the zone. Fig. 8.23 (b) shows that the energy surfaces are spherical up to large value of k (i.e., nearly free electron approximation).
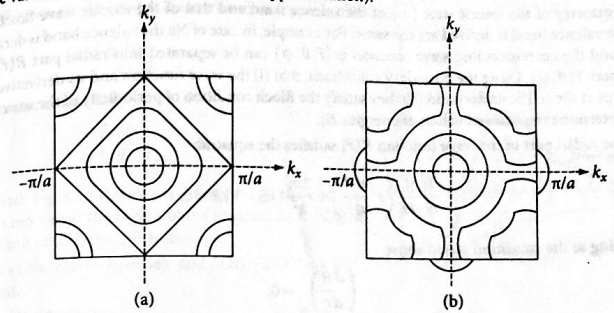

Fig. 8.23: Electron energy contours for SC lattice. (a) for tight binding approximation (b) for nearly free electron approximation.

Fundamental difference in the assumptions of nearly free electron approximation and tight binding approximation

Nearly free electron approximation (NFA)	Tight binding approximation (TBA)
1. The PE of the electron is small compared to its total energy.	1. The PE of the electron is nearly equal to its total energy.
2. The width of the forbidden bands is smaller compared to its allowed bands.	2. The width of the forbidden bands is larger compared to allowed energy bands.
3. As the atoms of the crystal are assumed to be closed spaced, the electron wave functions of neighbouring atoms overlap.	3. The atoms of the crystal are widely separated so that the electron wave functions of neighbouring atoms do not overlap to appreciable extent.
4. The interaction between the neighbouring atoms is strong.	4. The interaction between the neighbouring atoms is relatively weak.
5. The wave functions and the energy states in the crystal do not resemble with those of an individual atom.	5. The wave functions and energy as a whole are closely related to those of an individual atom.

8.12 Wigner-Seitz Cellular Method

Slater in 1934 introduced the *cellular method* to calculate band structure using *Wigner-Seitz's method.* The following points are assumed:

1. The solid is monatomic such that all atoms in it are translationally equivalent.
2. The Wigner-Seitz cell (*cellular polyhedron*) is chosen as the unit cell so that each atom lies at the centre of the cell and hence shows the symmetry properties of the crystal.
3. As the electron wave functions obey Bloch theorem it is necessary to find out the wave functions in one Wigner-Seitz cell of the structure.

4. *The method aims to make maximum use of spherical symmetry of the atom and the translational symmetry of the crystal in the integration of the wave function within each Wigner-Seitz cell.*

5. *The Wigner-Seitz cell is replaced by a sphere with the same volume as of the cell and the potential $V(r)$ at any point inside the cell is spherically symmetrical, while outside the sphere $V(r)$ is assumed to be negligible.*

The symmetry of the lowest state $(\vec{k})$ of the valence band and that of the atomic wave function from which the valence band is derived are the same. For example, in case of Na the valence band is derived from 3s state and the corresponding wave function $\psi(\vec{r}, \theta, \phi)$ can be separated into radial part $R(\vec{r})$ and the angular part $Y(\theta, \phi)$. Using the boundary conditions, that (i) the wave function and its derivative must be continuous at the cell boundary and (ii) they satisfy the Bloch condition of periodicity of the wave function we can determine the allowed values of energies E_0.

Now the radial part of the wave function $R(\vec{r})$ satisfies the equation

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{2m}{\hbar^2}(E_0 - V)R = 0. \qquad (8.120)$$

According to the condition stated above

$$\left(\frac{dR}{dr}\right)_{r_s} = 0. \qquad (8.121)$$

As $R(r)$ must resemble 3s state near the ion, it must have two spherical nodes. The allowed values of E_0 corresponding to different values of r_s may be calculated. The energy obtained from the condition

$$\left(\frac{dR}{dr}\right)_{r_s} = 0$$

is found to be much smaller than that corresponding to a free atom and this leads to main source of binding energy in metals like Na.

According to standard perturbation theory, in a cubic crystal, the energies of higher bands having no degeneracy at $\vec{k} = 0$ are given by

$$E_k = E_0 + \frac{\hbar^2 k^2}{2m^*}, \qquad (8.122)$$

where m^* is the effective mass and

$$m^* = \frac{1}{m}\left[1 + \frac{2}{m}\sum_\alpha \frac{|\langle 0|p_x|\alpha\rangle|^2}{E_0 - E_\alpha}\right]. \qquad (8.123)$$

In eqn (8.123), $\langle 0|p_x|\alpha\rangle$ represents the matrix element of the x-component of momentum operator between the states $\vec{k} = 0$ in the band O and another band α having energy E_α.

If E_k be the energy of an electron in the field of its own ion core, then the crystal energy is ΣE_k. The PE of all other electrons inside the cell is neglected as there is only one electron inside the cell at any moment. As the electron-ion, electron-electron and ion-ion interactions cancel in this approximation, the average energy of an electron is given by

$$E = E_0 + E(B), \qquad (8.124)$$

where $E(B)$ is the average KE of the electron and $E(B) = \frac{3}{5}E_F$; E_F, the Fermi energy.

Then the *binding energy* is given by

$$E_0 = E(A) - [E_0 + E(B)] = [E(A) - E_0] - E(B), \qquad (8.125)$$

where $E(A)$ is the energy of a free atom.

The difference $[E(A) - E_0]$ is due to the boundary conditions

$$\left(\frac{d\psi_0}{dr}\right)_{r=\infty} = 0$$

in the free atom and

$$\left(\frac{d\psi}{dr}\right)_{r=r_s} = 0$$

in the metal. Fig. 8.24 shows this difference. As the electron density is more near the region of the nucleus, $E_0 < E(A)$ in the solid and as such the PE is lowered.

In case of Na, $E(A) - E_0 \approx 3\,\text{eV}$ and $E(B) \sim 2\,\text{eV}$ for a half-filled band.

So, the binding energy

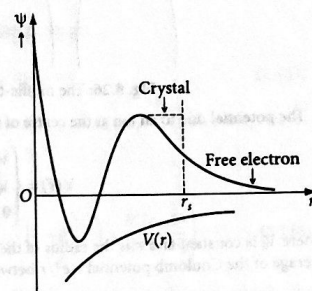$$E_0 = 3\,\text{eV} - 2\,\text{eV} = 1\,\text{eV} \text{ [from Fig. 8.24]}$$

Fig. 8.24: Renormalisation of atomic wave function in Wigner-Seitz model

It is to be remembered that the Wigner-Seitz method does not depend upon the structure of the crystal but depends on the atomic volume.

In conclusion, the cellular method is greatly over-simplified and nowadays it has little use. The main disadvantage lies in the fact that when we replace the Wigner-Seitz cell by a sphere, we ignore the crystal structure and all anisotropic effects are completely masked out.

8.13 The Augmented Plane Wave (APW) Method

Slater in 1937 used the results of cellular method to overcome the following two major difficulties, viz.,

(i) The potential used in cellular method has a discontinuous derivative whenever the boundary between two cells is crossed [Fig. 8.25], but actually the potential is quite flat in these regions.

(ii) The method requires lot of numerical calculations to match the wave functions at the boundary of the cells (specially for $\vec{k} = 0$ states).
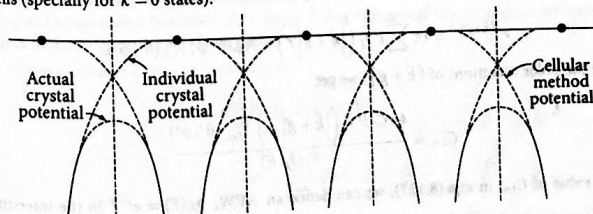

Fig. 8.25: The cellular method potential has a discontinuous derivative midway between two lattice points where the actual potential is flat.
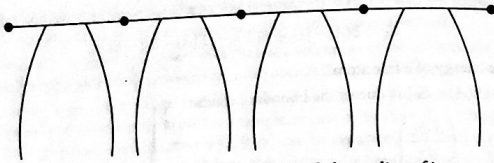
Fig. 8.26: The muffle-tin potential plotted along a line of ions

The potential due to an ion at the centre of the unit cell is given by

$$V(\vec{r}) = \begin{cases} V_i(\vec{r}) - V_0 & \text{for } r \le r_i \\ V_0 & \text{for } r_i < r < r_s \\ 0 & \text{for } r \ge r_s \end{cases} \qquad (8.126)$$

where V_0 is constant and r is the radius of the sphere inscribed in the Wigner-Seitz cell. V_0 is actually the average of the Coulomb potential $-e^2/r$ between the spheres of radii r_i and r_s, i.e.,

$$V_0 = -\frac{e^2 \int_{r_i}^{r_s} \frac{4\pi r^2}{r} dr}{\int_{r_i}^{r_s} 4\pi r^2 dr} = \frac{3e^2 \left(r_s^2 - r_i^2\right)}{2\left(r_s^3 - r_i^3\right)}.$$

From the two solutions in the region $r > r_i$ and $r < r_i$, a solution of the *Schrödinger's equation* can be written as

$$\sum_{l,m} C_{l,m} R_l(\vec{r}, E) Y_{l,m}(\theta, \phi), \ r < r_i$$
$$\phi_{k.g(r)} = C_0 e^{i(\vec{k}+\vec{g})\cdot\vec{r}}. \qquad (8.127)$$

Here $R_l(\vec{r}, E)$ is a solution of the radial equation

$$-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial R_l}{\partial r}\right) + \left[\frac{l(l+1)}{r^2} + \frac{2mV(r)}{\hbar^2}\right] R_l = \left(\frac{2mE}{\hbar^2}\right) R_l \qquad (8.128)$$

at a given energy E. $C_{l,m}$, C_0 are constants so that the two solutions for $r > r_i$ and $r < r_i$ are continuous at r_i. $g(\vec{r})$ is a *reciprocal lattice vector* and $Y(\theta, \phi)$ is a *spherical harmonic*. Now

$$e^{i(\vec{k}+\vec{g})\cdot\vec{r}} = 4\pi \sum i^l j_l\left(\left|\vec{k}+\vec{g}\right|\vec{r}\right) \times Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi'). \qquad (8.129)$$

θ', ϕ' define the polar directions of $\left(\vec{k}+\vec{g}\right)$, we get

$$C_{lm} = \frac{4\pi C_0 i^l j_l\left(\left|\vec{k}+\vec{g}\right|r_i\right) Y_{lm}^*(\theta', \phi')}{R_l(\vec{r}_i, E)}. \qquad (8.130)$$

Using this value of C_{lm} in eqn (8.127), we can define an APW, $\phi_k(\vec{r}) = e^{i\vec{k}\cdot\vec{r}}$ in the interstitial space satisfying

$$\phi_k(\vec{r}) = e^{i\vec{k}\cdot\vec{l}} \phi_k\left(\vec{r} - \vec{l}\right) \qquad (8.131)$$

inside the sphere with centre at $\vec{l}$.

The exact wave function $\psi_k(\vec{r})$ can be obtained by the superposition of individual $\phi_k(\vec{r})$, i.e.,

$$\psi_k(\vec{r}) = \sum_g \alpha_k - g\phi_k - g(\vec{r}). \qquad (8.132)$$

Using $\psi_k(\vec{r})$ as a trial function in a variational estimate of energy, we get

$$\left[\left(\vec{k}-\vec{g}\right)^2 - E\right]\alpha_{k-g} + \sum \beta_{gg'}\alpha_{k-g'} = 0, \qquad (8.133)$$

where

$$\beta_{gg'} = \frac{4\pi r_i^2}{\Omega_0}\left\{-\left[\left(\vec{k}-\vec{g}\right)\cdot\left(\vec{k}-\vec{g}\right) - E\right]\frac{j_l\left(\left|\vec{g}-\vec{g}'\right|r_i\right)}{\left|\vec{g}-\vec{g}'\right|}\right.$$
$$\left. + \sum_{l=0}^{\infty}(2l+1)P_l\left(\cos\theta_{gg'}\right) j_l\left(\left|\vec{k}-\vec{g}\right|r_i\right) j_l\left(\left|\vec{k}-\vec{g}'\right|r_i\right) \times \frac{R_l'(r_i, E)}{R_l(r_i, E)}\right\}, \qquad (8.134)$$

where $\theta_{gg'}$ is the angle between $\left(\vec{k}-\vec{g}\right)$ and $\left(\vec{k}-\vec{g}'\right)$, $R_l'(r_i, E)$ is the derivative of $R_l(r_i, E)$ and Ω_0, the volume of the unit cell.

Now equating the secular determinant of the coefficient to zero, the band structure can be determined, e.g.,

$$\left|(k-g)^2 - E(k)\delta_{gg'} + \beta_{gg'}\right| = 0. \qquad (8.135)$$

To have more accurate $E(k)$ value, a larger secular determinant is needed. APW method is good for calculating the *band structure in metals*.

8.14 Orthogonalised Plane Wave (OPW) and Pseudo Potential Method

Herring in 1940 introduced the OPW method to overcome the limitations in the NFE and LCAO methods in the calculations of energy band in solids.

The core electron wave functions are localised about the the lattice sites and the valence electron wave functions have a rapid oscillatory behaviour having a tail. The OPW method uses the fact that the valence wave functions are orthogonal to the core wave functions. Hence, instead of expanding the wave functions in terms of simple waves, *Herring* used an expansion of the wave functions in terms of waves which are already orthogonal to core wave functions. As a result of this orthogonalisation, *wiggles* are introduced in the plane waves in the neigbourhood of ion-sites.

An OPW of wave number $\vec{k}$ may be expressed as

$$|OPW_k\rangle = (1 - P)|\vec{k}\rangle \qquad (8.136)$$

with

$$|P\rangle = \sum_\alpha |\alpha\rangle\langle\alpha| \qquad (8.137)$$

is the *projection operator* and $|\alpha\rangle = \psi_a(r)$ is the core electron wave function. In eqn (8.137) $|\vec{k}\rangle$ and $P|\vec{k}\rangle$ represent respectively a *plane wave* and a *core function* so that, OPW = plane wave − core function as represented in Fig. 8.27.
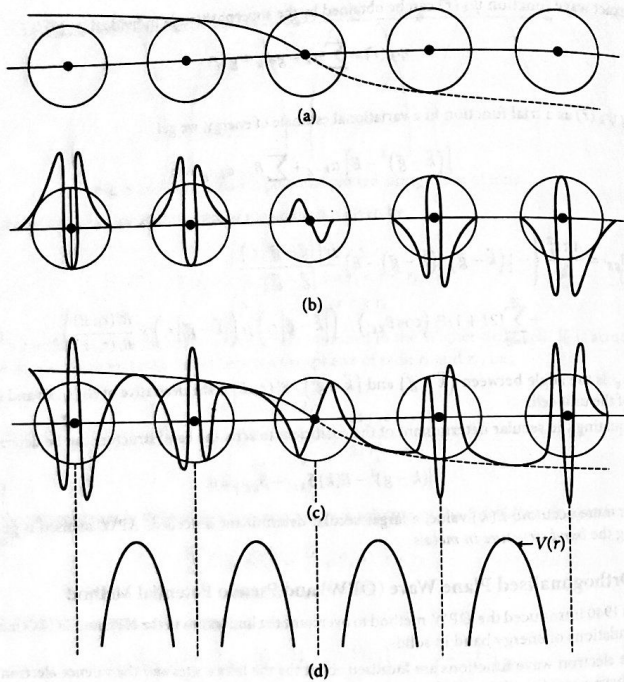
Fig. 8.27: (a) Plane wave (b) Rapid oscillation of Bloch function ψ_k near the atomic sites (c) Synthesis of OPW.

Supposing $|\beta\rangle$ to be one of the core states

$$\langle\beta|OPW_k\rangle = \langle\beta|(1-P)|\vec{k}\rangle = \langle\beta|\vec{k}\rangle - \langle\beta|P|\vec{k}\rangle$$

$$= \langle\beta|\vec{k}\rangle - \sum_a \langle\beta|a\rangle\langle a|\vec{k}\rangle = \langle\beta|\vec{k}\rangle - \sum_a \langle a|\vec{k}\rangle\,\delta_{a,\beta} = 0. \quad (8.138)$$

Here the *orthogonalised character of the core states* has been used;

$$\langle\beta|a\rangle = \delta_{a,\beta}. \quad (8.139)$$

Assuming T to be the KE operator and $V(\vec{r})$, the electron-ion interaction, $\vec{r}$ begin the position of the electron relative to the ion, then the *Hamiltonian* of the system is

$$H = T + V(\vec{r}) = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}). \quad (8.140)$$

Now the conduction band state may be written in terms of general linear combination of OPW's as

$$\psi_k = \sum_q a_q(\vec{k})\,(1-P)\,|\vec{k}+\vec{q}\rangle. \quad (8.141)$$

Substituting ψ_k in *Schrödinger equation*,

$$H\psi_k = E(\vec{k})\,\psi_k, \quad (8.142)$$

we get

$$\sum_q a_q(\vec{k})\,H(1-P)\,|\vec{k}+\vec{q}\rangle = E(\vec{k})\sum_q a_q(\vec{k})\,(1-P)\,|\vec{k}+\vec{q}\rangle. \quad (8.143)$$

Let

$$\psi_k = \sum_q a_q(\vec{k})\,|\vec{k}+\vec{q}\rangle, \quad (8.144)$$

then

$$\psi_k = (1-P)\phi_k. \quad (8.145)$$

Substituting this value of ψ_k in eqn (8.143), we get

$$[T+V(r)]\phi_k - HP\psi_k = E(\vec{k})\,\phi_k - E(\vec{k})\,PQ_k$$
$$\text{or, } T\phi_k + [V(r) - HP + E(\vec{k})\,P]\,\phi_k = E(\vec{k})\,\phi_k$$
$$\text{or, } T\phi_k + W(\vec{r})\phi_k = E(\vec{k})\,\phi_k, \quad (8.146)$$

where

$$W(\vec{r}) = V(\vec{r}) + [E(\vec{k}) - H]\,P = V(\vec{r}) + \sum_a [E(\vec{k}) - E_a]\,|a\rangle\langle a|. \quad (8.147)$$

$W(\vec{r})$ is known as the pseudo-potential and $\phi_k(\vec{r})$ is the pseudo-wave function. Here $H|a\rangle = E_a|a\rangle$.

Eqn (8.146) shows that ψ_k and $W(\vec{r})$ lead to the same set of values as their real parts.

Since, $E(\vec{k}) > E_a$ and $V(\vec{r})$ is negative, eqn (8.147) shows that $|W(\vec{r})| < V(\vec{r})$, this means by properly choosing the term $[E(k) - H]\,P$ we can make $W(\vec{r})$ as small as we please. Thus, for calculating the eigenvalues, the electron-ion potential in the low order perturbation theory can be successfully used.

According to the second order perturbation theory

$$E(\vec{k}) = \frac{\hbar^2 k^2}{2m} + \langle\vec{k}|W|\vec{k}\rangle + \sum_q' \frac{\left|\langle\vec{k}+\vec{q}|W|\vec{k}\rangle\right|^2}{\frac{\hbar^2}{2m}\left[k^2 - |\vec{k}+\vec{q}|^2\right]}. \quad (8.148)$$

Here prime indicates that $q \neq 0$.

The term

$$E_{bs}(\vec{k}) = \sum_q' \frac{\left|\langle\vec{k}+\vec{q}|W|\vec{k}\rangle\right|^2}{\frac{\hbar^2}{2m}\left[k^2 - |\vec{k}+\vec{q}|^2\right]} \quad (8.149)$$

in eqn (8.148) represents the *band structure energy* of the $\vec{k}$ state.

If $\vec{r}$ be the position of an electron and $\vec{R}_j$ that of an ion, then the individual ionic potential is given by $w\left(|\vec{r}-\vec{R}|\right)$.

Thus
$$W(\vec{r}) = \sum_j w\left(|\vec{r}-\vec{R}|\right). \tag{8.150}$$

$$\therefore \langle \vec{k}+\vec{q}\,|W|\vec{k}\rangle = S(\vec{q})\,\langle \vec{k}+\vec{q}\,|w|\vec{k}\rangle \tag{8.151}$$

where
$$S(\vec{q}) = \frac{1}{N}\sum_j e^{-i\vec{q}\cdot\vec{R}_j} \tag{8.152}$$

is the *structure factor*. When the ions occupy precisely the lattice position at equilibrium, $S(\vec{q})$ becomes unity. This method has been successfully used to calculate the band structure of metals (Be, Na, K) and semiconductors (Gi, Si, etc.)

8.15 Answers of Some Important Questions

8.1. *Discuss how energy bands are formed?*

Ans. See Art. 8.2.

8.2. *State Bloch Theorem. What is the outcome of the theory?*

Ans. Bloch theorem states that the solution of the Schrödinger equation for a periodic potential must be of the form
$$\psi(x) = e^{\pm ikx}\,u_k(x),$$
where $u_k(x) = u_k(x+a)$, a being the period. Thus, the solutions are plane waves of the type $e^{\pm ikx}$ modulated by the function $u_k(x)$ having the same periodicity as the lattice.

The *outcome of Bloch theorem* is that the energy spectrum of an electron moving in a periodic potential consists of allowed and forbidden energy regions (or bands).

8.3. *What is the concept of effective mass of the electron?*

Ans. In a crystal the electrons interact with the lattice and are not completely free. The wave particle motion, therefore, differs from that of an electron in free space. So, while using the equations of electrodynamics for charge carriers in a crystal we must keep in mind the altered value of the particle mass called the *effective mass* (m^*) due to the effect of the crystal lattice. Then the *electrons* (and *holes*) can be regarded as free charge carriers and *Newton's second law of motion* can be applied to describe their dynamic behaviour.

Statement: *When an electron in a periodic potential is accelerated by an electric field or magnetic field, the mass of the electron varies and this mass is called the effective mass of the electron.*

If E be the externally applied electric field accelerating the electron in the lattice, then the effective mass of the electron is given by
$$m^* = \frac{\hbar^2}{(d^2E/dk^2)},$$
where $\hbar = h/2\pi$ and k is the wave vector.

8.4. *What do you mean by valence band?*

Ans. In a solid the highest filled energy band which includes electrons shared in *covalent bands* or electrons transferred in *ionic bonds* is known as *valence band*.

Valence band includes electrons which are transferred in case of *ionic bond* or shared in case of covalent bonds. These electrons are not free to gain energy from an external electric field and hence are unable to move through the solid. The sharp atomic levels are spread out in solids and all the energy states up to a forbidden energy region are occupied (at least at 0 K) and there is no more free electrons.

8.5. *What is a conduction band?*

Ans. The band of energy of valence electrons consisting of closely spaced energy levels is called conduction band. The bottom of the conduction band has energy E_c.

The electrons in the valence orbit of atoms are loosely bound to the nucleus and are free to move in the solid forming *free electron gas*. There is a change in energy of these electrons due to mutual interaction. The value of energy of these electrons varies from zero to a maximum value called the *Fermi energy* E_f. These electrons participate in the electric conduction process in presence of an external electric field. The energy band of these electrons is called the *conduction band*.

8.6. *What is forbidden gap?*

Ans. The energy gap between the valence band and conduction band is called *forbidden energy gap* or *band gap*. Electrons cannot have any energy in the forbidden energy gap. The energy of the forbidden energy is called *Band Gap* $\left(E_g\right)$. If E_v and E_c represent respectively the energy of the topmost level of the valence band and bottom level of conduction band, then $E_g = E_c - E_v$.

If E_g is very small, thermal excitation may excite the electrons in the valence band to the conduction band, so that the solid, although an insulator at 0 K, may become a conductor at temperature $T > 0$ K.

8.7. *What are the sources of electrical resistance in solids?*

Ans. The electrical conductivity depends mainly on two factors, viz., (a) temperature and (b) defects (e.g., impurities).

According to *Mathiessen's rule*
$$\rho_{total} = \rho_{pure} + \rho_{impurity},$$
where ρ_{pure} is the temperature dependent of resistivity and $\rho_{impurity}$ is the resistivity due to scattering of electrons by impurity atoms.

The *temperature dependence* of resistivity arises due to the scattering effect of thermal vibrations of the atoms on the electrons. If A_0 be the amplitude of the thermal vibrations, then the *mean free path* λ of the electron is
$$\lambda \propto 1/A_0^2.$$

Again, $A_0^2 \propto kT$. $\therefore \lambda_0^2 \propto 1/T$.
$$\therefore \rho \propto 1/\lambda \propto T,$$
i.e., the resistivity is directly proportional to the temperature T of the solid.

The *defect dependence* of resistivity is due to the additional scattering of electrons by the *defect or impurities*. Assigning two relaxation time, we can write,
$$\frac{1}{\tau} = \frac{1}{\tau_0} + \frac{1}{\tau'},$$
where τ_0 is independent of temperature while τ' depends on temperature.

$$\therefore \rho = \frac{m}{ne^2\tau} = \frac{m}{ne^2}\left(\frac{1}{\tau_0}+\frac{1}{\tau'}\right) = \frac{m}{ne^2\tau_0} + \frac{m}{ne^2\tau'} = \rho_0 + \rho_{pure}(T) \ \text{[Mathiessen's rule]}$$

Note that the symmetry properties discussed above refer to the same band. They hold for every band separately, but do not relate one band to another.

Let us turn now to the proofs of the above properties. We shall only outline these proofs here, leaving you to pursue the details in some of the advanced references listed at the end of the chapter. Consider first the translational property (i): The Bloch function at the point $k + G$ may be written as

$$\psi_{k+G} = e^{i(k+G)\cdot r}u_{k+G} = e^{ik\cdot r}(e^{iG\cdot r}u_{k+G}). \qquad (5.12)$$

Note that the factor inside the brackets of the last expression, which may be denoted by $v(r)$, is periodic in the r-space with a period equal to the lattice vector. That is,

$$v(r + R) = e^{iG\cdot(r+R)}u_{k+G}(r + R) = e^{iG\cdot r}u_{k+G}(r) = (v)r.$$

This follows from the fact that u_{k+G} is periodic, and $e^{iG\cdot R} = 1$, since $G\cdot R = n2\pi$, where n is some integer. The expression in the brackets in (5.12) has, therefore, the same behavior as $u_k(r)$ in Eq. (5.3). We have thus shown that the state function ψ_{k+G} has the same form as ψ_k, and consequently the two functions have the same energy, since there is no physical basis for distinguishing between them.

Property (ii) may be established by noting that the Schrödinger equation analogous to (5.6), which corresponds to the point $-k$, is the same as the equation obtained by writing the complex conjugate equation of (5.6). This means that the corresponding eigenvalues are equal, that is, that $E_n(-k) = E_n^*(k)$. Since the energy $E_n(k)$ is a real number, however, it follows that $E_n(-k) = E_n(k)$, which is property (ii).

Property (iii) is derived by noting that if the real lattice is rotated by a symmetry operation, the potential $V(r)$ remains unchanged, and hence the new state function obtained must have the same energy as the original state function. One may show further that these new states correspond to rotations in k-space, and this leads to the desired property.

5.5 NUMBER OF STATES IN THE BAND

We denoted the Bloch function by $\psi_{n,k}$, which indicates that each value of the band index n and the vector k specifies an electron state, or orbital. We shall now show that *the number of orbitals in a band inside the first zone is equal to the number of unit cells in the crystal*. This is much the same as the statement made in connection with the number of lattice vibrational modes (Section 3.3), and is proved in a like manner, by appealing to the boundary conditions.

Consider first the one-dimensional case, in which the Bloch function has the form

$$\psi_k(x) = e^{ikx}u_k(x). \qquad (5.13)$$

If we impose the periodic boundary condition on this function, it follows that the only allowed values of k are given by

$$k = n\frac{2\pi}{L}, \qquad (5.14)$$

where $n = 0$, ± 1, ± 2, etc. [Note that $u_k(x)$ is intrinsically periodic, so the condition $u_k(x + L) = u_k(x)$ is automatically satisfied.] As in Section 3.3, the allowed values of k form a uniform mesh whose unit spacing is $2\pi/L$. The number of states inside the first zone, whose length is $2\pi/a$, is therefore equal to

$$(2\pi/a)/(2\pi/L) = L/a = N,$$

where N is the number of unit cells, in agreement with the assertion made earlier.

A similar argument may be used to establish the validity of the statement in two- and three-dimensional lattices.

It has been shown that each band has N states inside the first zone. Since each such state can accommodate at most two electrons, of opposite spins, in accordance with the Pauli exclusion principle, it follows that the maximum number of electrons that may occupy a single band is $2N$. This result is significant, as it will be used in a later section to establish the criterion for predicting whether a solid is going to behave as a metal or an insulator.

5.6 THE NEARLY-FREE-ELECTRON MODEL

In Section 5.3 and 5.4 we studied the general properties of the state functions, and of the energies of an electron moving in a crystalline solid. To obtain explicit results, however, we must solve the Schrödinger equation (5.1) for the actual potential $V(r)$ in the particular solid of interest. But the process of solving the Schrödinger equation for any but the simplest potentials is an arduous and time-consuming task, inundated with mathematical details. Although this is essential for obtaining results that may be compared with experiments, it is preferable to start the discussion of explicit solutions by using rather simplified potentials. The advantage is that we can solve the Schrödinger equation with only minimal mathematical effort and thus concentrate on the new physical concepts involved.

In the present section we shall treat the *nearly-free-electron* (NFE) *model*, in which it is assumed that the crystal potential is so weak that the electron behaves essentially like a free particle. The effects of the potential are then treated by the use of perturbation methods, which should be valid inasmuch as the potential is weak. This model should serve as a rough approximation to the valence bands in the simple metals, that is, Na, K, Al, etc.

In the following section, we shall treat the *tight-binding model*, in which the atomic potentials are so strong that the electron moves essentially around a single atom, except for a small interaction with neighboring atoms, which may then be treated as a perturbation. This model lies at the opposite end from the NFE model in terms of the strength of crystal potential involved, and should serve as a rough approximation to the narrow, inner bands in solids, e.g., the 3d band in transition metals.

The empty-lattice model

The starting point for the NFE model is the solution of the Schrödinger equation for the case in which the potential is exactly zero, i.e., the electron is entirely free. However, we also require that the solutions satisfy the symmetry properties of Section 5.4, which are imposed by the translational symmetry of the real lattice. This leads to the so-called *empty-lattice model*.
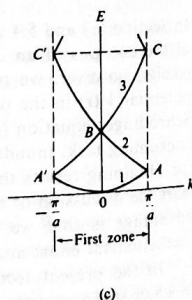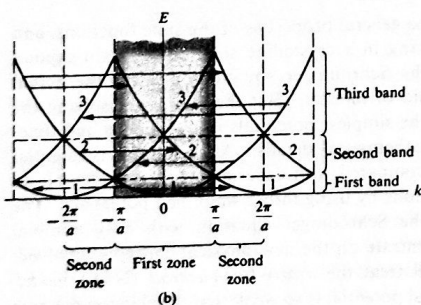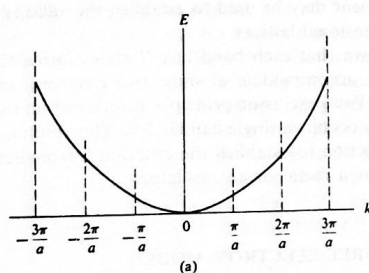

Fig. 5.10 (a) The familiar parabola representing the dispersion curve for a free particle, $E = \hbar^2 k^2/2m_0$. (b) The dispersion curves for the same particle in the empty-lattice model, showing translational symmetry and the various bands. (c) Dispersion curves in the empty-lattice model (first zone only).

For a one-dimensional lattice, the state functions and energies for the empty-lattice model are

$$\psi_k^{(0)} = \frac{1}{L^{1/2}} e^{ikx}, \tag{5.15}$$

and

$$E_{(k)}^{(0)} = \frac{\hbar^2 k^2}{2m_0}, \tag{5.16}$$

where the superscript 0 indicates that the solutions refer to the unperturbed state (Section A.7). The energy $E_{(k)}^{(0)}$ which is plotted versus k in Fig. 5.10(a) exhibits a curve in the familiar *parabolic* shape. Figure 5.10(b) shows the result of imposing the symmetry property (i) of Section 5.4. Segments of the parabola of Fig. 5.10(a) are cut at the edges of the various zones, and are translated by multiples of $G = 2\pi/a$ in order to ensure that the energy is the same at any two equivalent points. Figure 5.10(c) displays the shape of the energy spectrum when we confine our consideration to the first Brillouin zone only. [Conversely, Fig. 5.10(b) may be viewed as the result of translating Fig. 5.10(c) by multiples of G.]

The type of representation used in Fig. 5.10(c) is referred to as the *reduced-zone scheme*. Because it specifies all the needed information, it is the one we shall find most convenient. The representation of Fig. 5.10(a), known as the *extended-zone scheme*, is convenient when we wish to emphasize the close connection between a crystalline and a free electron. However, Fig. 5.10(b) employs the *periodic-zone scheme*, and is sometimes useful in topological considerations involving the k-space. All these representations are strictly equivalent; the use of any particular one is dictated by convenience, and not by any intrinsic advantages it has over the others.

The nearly-free-electron model

How is the energy spectrum of Fig. 5.10(c) altered when the crystal potential is taken into account, or "turned on?" Figure 5.11(a) shows this. The first and second bands, which previously touched at the point A (and A') in Fig. 5.10(c) are now split, so that an energy gap is created at the boundary of the Brillouin zone. A similar gap is created at the center of the zone, where bands 2 and 3 previously intersected (point B in Fig. 5.10c) and also at point C, where bands 3 and 4 previously intersected. Thus, in general, in the empty-lattice model, energy gaps are created in k-space wherever bands intersect, which occurs either at the center or the boundaries of the BZ. At these points the shape of the spectrum is strongly modified by the crystal potential, weak as this may be. (In effect, what the crystal potential has accomplished is to smooth over the sharp "corners" present in the band structure of the empty lattice.)

In the remainder of the zone, however, the shape of the spectrum is affected very little by the crystal potential, since this is assumed to be weak. In that region
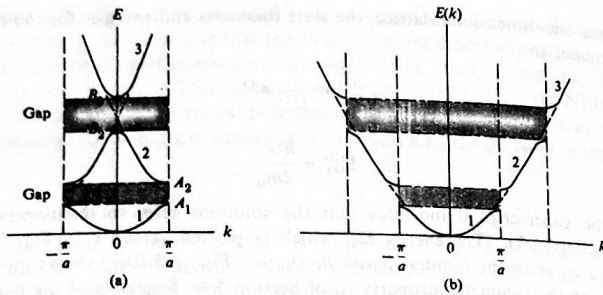
Fig. 5.11 (a) Dispersion curves in the nearly-free-electron model, in the reduced-zone scheme. (b) The same dispersion curves in the extended-zone scheme.

of the k-space the bands essentially retain their parabolic shape inherited from the empty-lattice model of Fig. 5.10(c), and the electron there behaves essentially like a free electron.

By comparing Fig. 5.10(c) and Fig. 5.11(a), one notes that a hint of a band structure is almost present even in the empty-lattice model, except that the gaps there vanish, since the bands touch at the zone boundaries. This vanishing is foreseen, of course, since no energy gaps are expected to appear in the spectrum of a free particle. The point is that even a weak potential leads to the creation of gaps, in agreement with the results of Sections 5.2 and 5.3.

Figure 5.11(b) shows the band structure for the NFE model, represented according to the extended-zone scheme, which should be compared with Fig. 5.10(a). Note that, except at the zone boundaries at which gaps are created, the dispersion curve is essentially the same as the free-electron curve.

We made the above assertions without proofs; we shall now outline proofs on the basis of the perturbation method of Section A.7. Suppose, for instance, that we seek to find the influence of the crystal potential on the first band in Fig. 5.10(c). When we treat the potential $V(x)$ as a perturbation, the perturbed energy $E_1(k)$ up to the second order of the potential is given by

$$E_1(k) = E_1^{(0)}(k) + \langle \psi_{1,k}^{(0)}|V|\psi_{1,k}^{(0)}\rangle + \sum_{k',n}{}' \frac{|\langle n, k'|V|1, k\rangle|^2}{E_1^{(0)}(k) - E_n^{(0)}(k')}. \quad (5.17)$$

Here the subscript 1 refers to the first band, which is the one of interest, and the superscript 0 refers to the empty-lattice model of Eqs. (5.15) and (5.16). The second term on the right side of (5.17), which is the first-order correction, is the average value of the potential. The third term, giving the second-order correction, involves summing over all states n, k, except where these indices are equal to the state $1, k$ under investigation.

First we note that the first-order correction is equal to

$$\langle \psi_{1,k}^{(0)}|V|\psi_{1,k}^{(0)}\rangle = \frac{1}{L}\int e^{-ikx}V(x)e^{ikx}\,dx = \frac{1}{L}\int V(x)\,dx,$$

which is the average value of the potential over the entire lattice. It is independent of k, and hence it is merely a constant. Its effect on the spectrum of Fig. 5.10(c) is simply to displace it rigidly by a constant amount, without causing any change in the shape of the energy spectrum. Since this term does not lead to anything of interest to us here, it will be set equal to zero, which can be accomplished by shifting the zero energy level.

We must therefore consider the second-order correction in Eq. (5.17). We first assert that the quantity $\langle n, k'|V|1, k\rangle$ can be shown to vanish except when $k' = k$, where both k and k' are restricted to the first zone. That is, the only states which are coupled to the $1, k$ state by the perturbation are those lying directly above this state, as shown in Fig. 5.12.
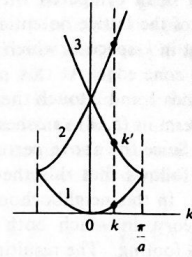

Fig. 5.12 Only those states lying directly above the state $\psi_{1,k}^{(0)}$ in k-space are coupled to it by the perturbation.

This assertion rests on the translational symmetry of the crystal potential $V(x)$.[†] Furthermore, since the energy difference in the denominator of the third

---

[†] An arbitrary potential $V(x)$ can always be expanded as a Fourier series

$$V(x) = \sum_k V_k e^{ikx},$$

where the summation is over all the allowed k's. The Fourier coefficient V_k is given by

$$V_k = (1/L)\int_0^L V(x)e^{-ikx}\,dx.$$

But if $V(x)$ is periodic, as is the case in a crystal, then only the values $k = G$ contribute to the above summation; that is, $V_k = 0$ for $k \neq G$. A periodic potential therefore has the expansion

$$V(x) = \sum_G V_G e^{iGx}.$$

It can be shown that the bracket in the numerator of (5.17) is the Fourier coefficient $V_{k'-k}$, and hence this bracket vanishes except for $k' - k = G$.

term in (5.17) increases rapidly as the band n rises, the major effect on band 1 arises from its coupling to band 2. We may therefore write

$$E_1(k) \simeq E_1^{(0)}(k) + \frac{|V_{-2\pi/a}|^2}{E_1^{(0)}(k) - E_2^{(0)}(k)},$$
(5.18)

where $V_{-2\pi/a}$ is the Fourier component of the potential, that is,

$$V_{-2\pi/a} = \frac{1}{L} \int V(x) e^{+i(2\pi/a)x} \, dx.$$

An explicit expression for $E_1(k)$ can be obtained by substituting the values for $E_1^{(0)}(k)$ and $E_2^{(0)}(k)$: namely $E_1^{(0)}(k) = \hbar^2 k^2/2m$, and $E_2^{(0)}(k) = \hbar^2(k - 2\pi/a)^2/2m$. [Note that if $0 < k < \pi/a$, then the second band is obtained by translating that part of the free-electron curve lying in the interval $-2\pi/a < k < -\pi/a$, as seen in Fig. 5.10(b), and hence the above expression for $E_2^{(0)}(k)$.] But this is not really necessary, because if the potential is weak, then $|V_{-2\pi/a}|^2$ is very small, and the second term in (5.18) is negligibly small compared with the first. In other words, $E_1(k) \simeq E_1^{(0)}(k)$, and the effect of the lattice potential is negligible.

There is, however, one point in k-space at which the above conclusion breaks down: the point $k = \pi/a$ at the zone edge. At this point the energies $E_1^{(0)}(k)$ and $E_2^{(0)}(k)$ are equal [recall that bands 1 and 2 touch there; see Fig. 5.10(c)], the denominator of the perturbation term in (5.18) vanishes, and hence the perturbation correction becomes very large. Since the above perturbation theory presumes the smallness of the correction, it follows that this theory cannot hold true in the neighborhood of the zone edge. In this neighborhood, one should instead invoke the *degenerate* perturbation theory, in which both bands 1 and 2 are treated simultaneously, and on an equal footing. The resulting energy values are (Ziman, 1963),

$$E_\pm(k) = \tfrac{1}{2}\{E_1^{(0)}(k) + E_2^{(0)}(k) \pm [(E_2^{(0)}(k) - E_1^{(0)}(k))^2 + 4|V_{-2\pi/a}|^2]^{1/2}\}, \quad (5.19)$$

where the plus sign corresponds to the deformed upper band—i.e., band 2—near the edge of the zone, and the minus sign refers to the deformed lower band—i.e., band 1.

Now let us substitute the values of $E_1^{(0)}(k)$ and $E_2^{(0)}(k)$ into (5.19) and plot $E_+(k)$ and $E_-(k)$ in the neighborhood of the zone edge. We obtain the spectrum shown in Fig. 5.11(a). In particular, the energy gap E_g is equal to the difference $E_+(k) - E_-(k)$ evaluated at the point $k = \pi/a$. Using (5.19), we readily find that

$$E_g = 2|V_{-2\pi/a}|.$$
(5.20)

That is, the energy gap is equal to twice the Fourier component of the crystal potential. In effect, band 1 has been depressed by an amount equal to $|V_{-2\pi/a}|$ and band 2 has been raised by the same amount, leading to an energy gap given by (5.20).

The same formula (5.19) may also be used to find the energy gap that arises at the center of the zone, at the intersection between bands 2 and 3, except that we now replace $E_1^{(0)}(k)$, $E_2^{(0)}(k)$ by $E_2^{(0)}(k)$ and $E_3^{(0)}(k)$, respectively. We also replace the potential term by $V_{-4\pi/a}$. This leads to the splitting of bands 2 and 3, as shown in Fig. 5.11(a), with an energy gap of $2|V_{-4\pi/a}|$. Obviously the procedure can be used to find both the splitting of the bands and the corresponding gaps at all appropriate points.

In addition to the above results, two qualitative conclusions emerge from the analysis. First, the higher the band, the greater its width; this is evident from referring back to the empty lattice model in Fig. 5.10(a), since the energy there increases as k^2. Second, the higher the energy, the narrower the gap; this follows from the fact that the gap is proportional to a certain Fourier component of the crystal potential, but note that the order of the component increases as the energy rises (from $V_{-2\pi/a}$ to $V_{-4\pi/a}$ in our discussion above). Since the potential is assumed to be well behaved, the components decrease rapidly as the order increases, and this leads to a decrease in the energy gap. It follows therefore that, as we move up the energy scale, the bands become wider and the gaps narrower; i.e., the electron behaves more and more like a free particle. This agrees with the qualitative picture drawn in Section 5.2.

Since the greatest effect of the crystal potential takes place near the points in k-space at which two bands touch, let us examine the behavior there more closely. If one applies the degenerate perturbation formula (5.17) to the splitting of bands 2 and 3 at the center of the zone, one finds that, for small k ($k \ll \pi/a$),

$$E_3(k) = E_B + |V_{-4\pi/a}| + \frac{\hbar^2}{2m_0}\alpha k^2,$$
(5.21)

and

$$E_2(k) = E_B - |V_{-4\pi/a}| - \frac{\hbar^2}{2m_0}\alpha k^2,$$
(5.22)

where the parameter α is given by

$$\alpha = 1 + \frac{4E_B}{E_g}$$
(5.23)

and $E_B = \hbar^2(2\pi/a)^2/2m_0$ is the energy of point B in Fig. 5.10(c). These results are very interesting for several reasons.

a) Equation (5.21) shows that, for an electron near the bottom of the third band, $E \sim k^2$ (ignoring the first two terms on the right, since they are simply constants), which is similar to the dispersion relation of a free electron. In other words, the electron there behaves like a free electron, with an effective mass m^* given by

$$m^* = m_0/\alpha,$$

which is different from the free mass. Referring to (5.23), one sees that the effective mass increases as the energy gap E_g increases. Such a relationship between m^* and E_g is familiar in the study of semiconductors.

b) Equation (5.22) shows that, for an electron near the top of the second band, $E \sim -k^2$, which is like a free electron, except for the surprising fact that the effective mass is negative. Such behavior is very unlike that of a free electron, and its cause lies, of course, in the crystal potential. The phenomenon of a negative effective mass near the top of the band is a frequent occurrence in solids, particularly in semiconductors, as we shall see later (Chapter 6).

We have thus far confined ourselves to a one-dimensional lattice, but we may extend this treatment to two- and three-dimensional lattices in a straightforward fashion. We find again, as expected, that starting with the empty-lattice model, the "turning on" of the crystal potential leads to the creation of energy gaps. Furthermore, these gaps occur at the boundaries of the Brillouin zone.

5.7 THE ENERGY GAP AND THE BRAGG REFLECTION

In discussing the NFE model, we focused on energy values. But perturbation also modifies state functions, and we shall now study this modification. If we apply the perturbation theory to the one-dimensional empty lattice, we find that the state function of the first band in Fig. 5.11(a) is given by

$$\psi_{1,k} = \psi_{1,k}^{(0)} + \frac{V_{-2\pi/a}}{E_1^{(0)}(k) - E_2^{(0)}(k)}\psi_{2,k}^{(0)}, \qquad (5.24)$$

where—again because of the form of the potential and also the energy difference in the denominator—the perturbation summation has been reduced to one term only, involving the state function of the second band $\psi_{2,k}^{(0)}$.

The state functions $\psi_{1,k}^{(0)}$ and $\psi_{2,k}^{(0)}$ refer to a free electron; $\psi_{2,k}^{(0)} \sim e^{ikx}$ represents a wave traveling to the right, while $\psi_{2,k}^{(0)} \sim e^{i(k-2\pi/a)x}$ represents a wave traveling to the left (note that $|k| < \pi/a$). The effect of the lattice potential is then to introduce a new left-traveling wave in addition to the incident free wave. This new wave is generated by the scattering of the electron by the crystal potential. If k is not close to the zone edge, however, the coefficient of $\psi_{2,k}^{(0)}$ in (5.24) is negligible. That is,

$$\psi_{1,k} \simeq \psi_{1,k}^{(0)} = \frac{1}{L^{1/2}}e^{ikx}, \qquad (5.25)$$

and the electron behaves like a free electron. The effects of the potential are negligible there, which is in agreement with the conclusions reached in Section 5.6.

Near the zone edge, however, the energy denominator in the correction term in (5.24) becomes very small, and the perturbation term large, which means that

the form (5.24) becomes invalid. As stated in Section 5.6, one must then use the degenerate perturbation theory, in which the state functions $\psi_{1,k}^{(0)}$ and $\psi_{2,k}^{(0)}$ are treated on an equal footing. One finds that, at the zone edge itself,

$$\psi_{\pm}(x) = \frac{1}{\sqrt{2}}[\psi_{1,\pi/a}^{(0)}(x) \pm \psi_{2,\pi/a}^{(0)}(x)] = \frac{1}{\sqrt{2L}}(e^{i(\pi/a)x} \pm e^{-i(\pi/a)x}). \qquad (5.26)$$

The function $\psi_+(x) \sim \cos(\pi/a)x$, and hence the probability is proportional to $|\psi_+(x)|^2 \sim \cos^2(\pi/a)x$. Such a state function distributes the electron so that it is piled predominantly at the nuclei (recall that the origin $x = 0$ is at the center of an ion) [see Fig. 5.13], and since the potential is most negative there, this
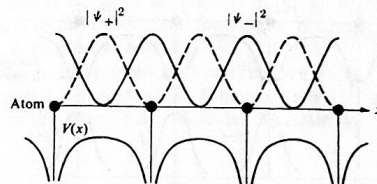

Fig. 5.13 Spatial distributions of electrons described by the functions ψ_+ and ψ_-.

distribution has a low energy. The function $\psi_+(x)$ therefore corresponds to the energy at the top of band 1, that is, point A_1 in Fig. 5.11(a).

By contrast, the function $\psi_-(x) \sim \sin \pi/ax$, depositing its electron mostly between the ions (as shown in Fig. 5.13), corresponds to the bottom of band 2 in Fig. 5.11(a), that is, point A_2. The gap arises, therefore, because of the two different distributions for the same value $k = \pi/a$, the distributions having different energies.

Scrutinizing (5.26) from the viewpoint of scattering, we see that at the zone edge, $k = \pi/a$, the scattering is so strong that the reflected wave has the same amplitude as the incident wave. As found above, the electron is represented there by a standing wave, $\cos \pi/ax$ or $\sin \pi/ax$, very unlike a free particle. An interesting result of this is that the electron, as a standing wave, has a zero velocity at $k = \pi/a$. This is a general result which is valid at all zone boundaries, and one which we shall encounter often in the following sections.

We have seen that the periodic potential causes strong scattering at $k = \pi/a$. Recall from Section 3.6 on lattice vibrations that this strong scattering arises as a result of the Bragg diffraction at the zone edge. In the present situation, the wave diffracted is the electron wave, whose wavelength is $\lambda = 2\pi/k$.

In higher-dimension lattices, the Bragg condition is satisfied along all boundaries of the Brillouin zone, as discussed in Section 2.6, and this results in the creation of energy gaps along these boundaries, in agreement with the conclusions of the last section.

5.8 THE TIGHT-BINDING MODEL

In the tight-binding model, it is assumed that the crystal potential is strong, which is the same as saying that the ionic potentials are strong. It follows, therefore, that when an electron is captured by an ion during its motion through the lattice, the electron remains there for a long time before leaking, or tunneling, to the next ion [see Fig. 5.14(a), which also shows that the energy of the electron is appreciably
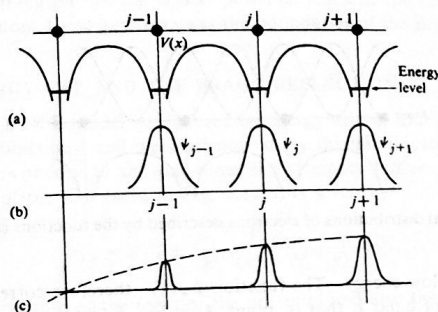

Fig. 5.14 The tight-binding model. (a) The crystal potential. (b) The atomic wave functions. (c) The corresponding Bloch function.

lower than the top of the potential barrier]. During the capture interval, the electron orbits primarily around a single ion, i.e., its state function is essentially that of an atomic orbital, uninfluenced by other atoms. Most of the time the electron is tightly bound to its own atom. The mathematical analysis to be developed must reflect this important fact.

As we said in Section 5.6, the TB (tight-binding) model is primarily suited to the description of low-lying narrow bands for which the shell radius is much smaller than the lattice constant. Here the atomic orbital is modified only slightly by the other atoms in the solid. An example is the 3d band, so important in transition metals.

Let us begin, then, with an atomic orbital, $\phi_v(x)$, whose energy in a free atom is E_v. We wish to examine the effects of the presence of other atoms in the solid. The index v characterizes the atomic orbital (for the atomic shell of interest).

First, the one-dimensional case: It is necessary to choose a suitable Bloch function, and while the choice is not unique, the following offers a reasonable form.

$$\psi_k(x) = \frac{1}{N^{1/2}} \sum_{j=1}^{N} e^{ikX_j} \phi_v(x - X_j), \tag{5.27}$$

where the summation extends over all the atoms in the lattice. The coordinate X_j specifies the position of the jth atom. That is, $X_j = ja$, where a is the lattice constant. The function $\phi_v(x - X_j)$ is the atomic orbital centered around the jth atom; it is large in the neighborhood of X_j, but decays rapidly away from this point, as shown in Fig. 5.14(b). By the time the neighboring site at X_{j+1} (or X_{j-1}) is reached, the function $\phi_v(x - X_j)$ has decayed so much that it has become almost negligible. In other words, there is only a little overlap between neighboring atomic orbitals. This is the basic assumption of the TB model. The factor $N^{1/2}$ is included in (5.27) to ensure that the function ψ_k is normalized to unity (if the atomic orbital ϕ_v is so normalized).

Let us turn now to the properties of the function $\psi_k(x)$, as defined by (5.27). First, it is necessary to ascertain that this function is a Bloch function, namely, that it can be written in the form (5.3). This can be established by rewriting (5.27) in the form

$$\psi_k(x) = \frac{1}{N^{1/2}} e^{ikx} \sum_{j=1}^{N} e^{-ik(x-X_j)} \phi_v(x - X_j),$$

where it is now readily recognized that the factor defined by the summation is periodic, with a period equal to the lattice constant a. Thus the function $\psi_k(x)$ has indeed the desired Bloch form, i.e., it describes a propagating electron wave, as shown in Fig 5.14(c).

Note also that near the center of the jth ion, the function $\psi_k(x)$ reduces to

$$\psi_k(x) \simeq e^{ikX_j} \phi_v(x - X_j) \sim \phi_v(x - X_j). \tag{5.28}$$

That is, the Bloch function is proportional to the atomic orbital. Thus in the neighborhood of the jth ion, the crystal orbital behaves much like an atomic orbital, in agreement with the basic physical assumption of the TB model.

The function $\psi_k(x)$ therefore satisfies both the mathematical requirement of the Bloch theorem and the basic assumption of the TB model, and as such is a suitable crystal orbital. It will be used now to calculate the energy of the band.

The energy of the electron described by ψ_k is given, according to quantum mechanics, by

$$E(k) = \langle \psi_k | H | \psi_k \rangle, \tag{5.29}$$

where H is the Hamiltonian of the electron[†]. Substituting for ψ_k from (5.27), one has

$$E(k) = \frac{1}{N} \sum_{j,j'} e^{ik(X_j - X_{j'})} \langle \phi_v(x - X_{j'}) | H | \phi_v(x - X_j) \rangle, \qquad (5.30)$$

where the double summation over j and j' extends over all the atoms in the lattice. Note that each term in the summation is a function of the difference $X_j - X_{j'}$, and not of X_j and $X_{j'}$ individually. Therefore, for each particular choice of j', the sum over j yields the same result, and since j' can take N different values, one obtains N equal terms, which thus leads to

$$E(k) = \sum_{j=-N/2}^{(N-1)/2} e^{ikX_j} \langle \phi_v(x) | H | \phi_v(x - X_j) \rangle, \qquad (5.31)$$

where we have arbitrarily put $X_{j'} = 0$ in (5.30). By splitting the term $j = 0$ from the others, one may write the above expression as

$$E(k) = \langle \phi_v(x) | H | \phi_v(x) \rangle + \sum_j{}' e^{ikX_j} \langle \phi_v(x) | H | \phi_v(x - X_j) \rangle, \qquad (5.32)$$

The first term gives the energy the electron would have if it were indeed entirely localized around the atom $j = 0$, while the second term includes the effects of the electron tunneling to the various other atoms. The terms in the summation are expected to be appreciable only for nearest neighbors—that is, $j = 1$ and $j = -1$— because as j increases beyond that point, the overlap between the corresponding functions and the state function at the origin becomes negligible (Fig. 5.14b). Note also that, since the property of electron delocalization is included entirely in the second term of (5.32), it is this term which is responsible for the band structure, and as such is of particular interest to us here.

To proceed with the evaluation of $E(k)$, according to (5.32), we need to examine the Hamiltonian H more closely. The expression for this quantity is given by

$$H = -\frac{\hbar^2}{2m_0} \frac{d^2}{dx^2} + V(x), \qquad (5.33)$$

where $V(x)$ is the crystal potential. Writing this potential as a sum of atomic potentials, one has

$$V(x) = \sum_j v(x - X_j). \qquad (5.34)$$

[†] The Hamiltonian H is simply the quantum operator which represents the total energy of the particle. Thus $H = -(\hbar^2/2m_0) \nabla^2 + V(\mathbf{r})$, where the first term on the right represents kinetic energy and the second term potential energy. The expression (5.29) for the energy is very plausible, since the term on the right is the average value of the energy in quantum mechanics.

In using this to evaluate the first term in Eq. (5.32), we shall find it convenient to split $V(x)$ into a sum of two terms

$$V(x) = v(x) + V'(x), \qquad (5.35)$$

where $v(x)$ is the atomic potential due to the atom at the origin and $V'(x)$ is that due to all the other atoms. These potentials are plotted in Figs. 5.15(a) and (b),
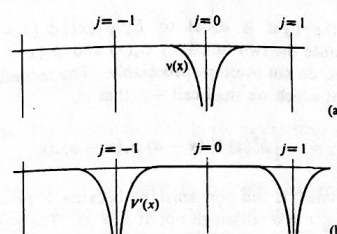

Fig. 5.15 The splitting of the crystal potential into (a) an atomic potential and (b) the remainder of the crystal potential.

respectively. Note in particular that $V'(x)$ is small in the neighborhood of the origin. The first term in (5.32) may now be written as

$$\langle \phi_v(x) | H | \phi_v(x) \rangle = \left\langle \phi_v(x) \left| \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + v(x) \right] \right| \phi_v(x) \right\rangle$$
$$+ \langle \phi_v(x) | V'(x) | \phi_v(x) \rangle. \qquad (5.36)$$

The first term on the right is equal to E_v, the atomic energy, since the operator involved is the Hamiltonian for a free atom. The second term is an integral which can be evaluated, and will be denoted by the constant $-\beta$. Explicitly,

$$\beta = -\int \phi_v^*(x) \, V'(x) \phi_v(x) \, dx, \qquad (5.37)$$

where the minus sign is introduced so that β is a positive number.[†] Note that β is a small quantity, since the function $\phi_v(x)$ is appreciable only near the origin, whereas $V'(x)$ is small there. Collecting the two terms above, we have

$$\langle \phi_v(x) | H | \phi_v(x) \rangle = E_v - \beta. \qquad (5.38)$$

[†] The integral in (5.37) is negative because $V'(x)$ is negative (Fig. 5.15b).

Let us now turn to the interaction term, i.e., the summation in (5.32). The term involving interaction with the nearest neighbor at $X_1 = a$ involves an integral which may be written as

$$\langle \phi_v(x) | H | \phi_v(x - a) \rangle = \langle \phi_v(x) | - \frac{\hbar^2}{2m_0} \frac{d^2}{dx^2}$$

$$+ v(x - a) | \phi_v(x - a) \rangle + \langle \phi_v(x) | V'(x - a) | \phi_v(x - a) \rangle. \quad (5.39)$$

The first term on the right is equal to $E_v \langle \phi_v(x) | \phi_v(x - a) \rangle$, which is a negligible quantity, since the two functions $\phi_v(x)$ and $\phi_v(x - a)$, being centered at two different atoms, do not overlap appreciably. The second term on the right of (5.39) is a constant which we shall call $-\gamma$, that is,

$$\gamma = - \int \phi_v^*(x) V'(x - a) \phi_v(x - a) \, dx. \quad (5.40)$$

Note that γ, though small, is still nonvanishing because $V'(x - a)$ is appreciable near the origin, that is, $x = 0$ (although not at $x = a$). The parameter γ is called the *overlap integral*, since it is dependent on the overlap between orbitals centered at two neighboring atoms.

The integral arising from the term $j = -1$ in the sum in (5.32), which is due to the atom on the left side of the origin, yields the same result as (5.39) because the atomic functions are symmetric.

Substituting the above results into (5.32), and restricting the sum to nearest neighbors only, one finds

$$E(k) = E_v - \beta - \gamma \sum_{j=-1}^{1}{}' e^{ikX_j}, \quad (5.41)$$

which may thus be written as

$$E(k) = E_v - \beta - 2\gamma \cos ka. \quad (5.42)$$

This is the expression we have been seeking. It gives band energy as a function of k in terms of well-defined parameters which we can evaluate from our knowledge of atomic energy and atomic orbitals.

Equation (5.42) may be rewritten more conveniently as

$$E(k) = E_0 + 4\gamma \sin^2 \left(\frac{ka}{2} \right), \quad (5.43)$$

where

$$E_0 = E_v - \beta - 2\gamma. \quad (5.44)$$

The energy $E(k)$ is plotted versus k in Fig. 5.16, where k is restricted to the first zone [although $E(k)$ is obviously periodic in k, in agreement with property (i)

of Section 5.4]. We see, as expected, that the original atomic level E_v has broadened into an energy band. The bottom of the band, located at $k = 0$, is equal to E_0, and its *width* is equal to 4γ.
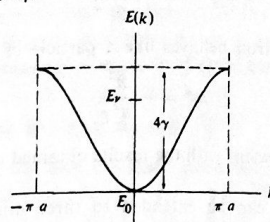

Fig. 5.16 The dispersion curve in the tight-binding model.

Note that the bottom of the band E_0 is lower than the atomic energy E_v, which is to be expected, since one effect of the presence of the other atom is to depress the potential throughout the system (refer to Fig. 5.14a). In addition to E_0, the electron has an amount of energy given by the second term in (5.43). This is a *kinetic energy*, arising from the fact that the electron is now able to move through the crystal.

Note also that the bandwidth, 4γ, is proportional to the overlap integral. This is reasonable, because, as we saw in Section 5.2, the greater the overlap the stronger the interaction, and consequently the wider the band.

When the electron is near the bottom of the band, where k is small, one may make the approximation $\sin (ka/2) \sim ka/2$, and hence

$$E(k) - E_0 = \gamma a^2 k^2, \quad (5.45)$$

which is of the same form as the dispersion relation of a free electron. An electron in that region of k-space behaves like a free electron with an effective mass

$$m^* = \frac{\hbar^2}{2a^2} \frac{1}{\gamma}. \quad (5.46)$$

It is seen that the effective mass is inversely proportional to the overlap integral γ. This is intuitively reasonable, since the greater the overlap the easier it is for the electron to tunnel from one atomic site to another, and hence the smaller is the inertia (or mass) of the electron. Conversely, a small overlap leads to a large mass, i.e., a sluggish electron. Of course, in the TB model, the overlap is supposed to be small, implying a large effective mass.

Note, however, that an electron near the top of the band shows unusual behavior. If we define $k' = \pi/a - k$, and expand the energy $E(k)$ near the

maximum point, using (5.43), we arrive at

$$E(k') - E_{max} = -\frac{a^2}{2}\gamma k'^2,\tag{5.47}$$

which shows that the electron behaves like a particle of *negative* effective mass

$$m^* = -\frac{\hbar^2}{a^2\gamma}.\tag{5.48}$$

This, you recall, is in agreement with the results obtained on the basis of the NFE model.

The above treatment can be extended to three dimensions in a straightforward manner. Thus for a sc lattice, the band energy is given by

$$E(k) = E_0' + 4\gamma\left[\sin^2\left(\frac{k_xa}{2}\right) + \sin^2\left(\frac{k_ya}{2}\right) + \sin^2\left(\frac{k_za}{2}\right)\right].\tag{5.49}$$

where E_0' is the energy at the bottom of the band. The energy contours for this band, in the $k_x - k_y$ plane, are shown in Fig. 5.17(a), and the dispersion curves along the [100] and [111] directions are shown in Fig. 5.17(b). The bottom of the band is at the origin $k = 0$, and the electron there behaves as a free particle with an effective mass given by (5.46). The top of the band is located at the corner of the zone along the [111] direction, that is, at $[\pi/a, \pi/a, \pi/a]$; the electron there has a negative effective mass given by (5.48). The width of the band is equal to 12γ.
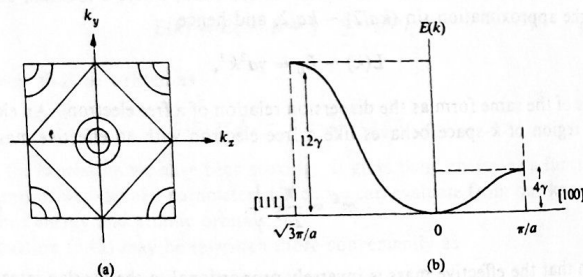

Fig. 5.17 (a) Energy contours for an sc lattice in the tight-binding model. (b) Dispersion curves along the [100] and [111] directions for an sc lattice in the TB model.

In this treatment of the TB model, we have seen how an atomic level *broadens* into a band as a result of the interaction between atoms in the solid. In this manner, each atomic level leads to its own corresponding band, and each band reflects the character of the atomic level from which it has originated.

In conclusion, we see that both the NFE and TB models lead to the same qualitative results, although the models start from opposite points of view. The principal results arrived at in both models are: (a) Energy gaps appear at zone boundaries. (b) An electron near the bottom of the band behaves like a free particle with a positive effective mass. (c) An electron near the top of the band behaves like a free particle with a negative effective mass.

5.9 CALCULATIONS OF ENERGY BANDS

In the last few sections we have discussed some methods of calculating energy bands. However, these methods—the NFE and TB models—are too crude to be useful in calculations of actual bands which are to be compared with experimental results. In this section we shall consider therefore some of the common methods employed in calculations of actual bands. Because this subject is an advanced one, requiring a considerable background in quantum mechanics, as well as meticulous attention to almost endless mathematical details, our discussion will be brief, primarily qualitative, and somewhat superficial. We shall nevertheless try to give the reader a glimpse of this fundamental subject in the hope that he may pursue it further, if he so desires, by referring to books listed in the bibliography at the end of the chapter.

Several different schemes for calculating energy bands have been used. Let us now discuss them individually.

The cellular method

The cellular method was the earliest method employed in band calculations (Wigner and Seitz, 1935). It was applied with success to the alkali metals, particularly Na and K; we shall use Na as an example.

The Schrödinger equation whose solution we seek is

$$\left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r})\right]\psi_k = E(k)\psi_k,\tag{5.50}$$

where $V(\mathbf{r})$ is the crystal potential and ψ_k the Bloch function. Here we are interested only in the 3s band. It is at once evident that this equation cannot be solved analytically. We must therefore use an approximation procedure.

When we use the cellular method, we divide the crystal into unit cells; each atom is centered at the middle of its cell, as shown in Fig. 5.18. Such a cell, known as the *Wigner–Seitz* (WS) *cell*, is constructed by drawing bisecting planes normal to the lines connecting an atom A, say, to its neighbors, and "picking out" the volume enclosed by these planes. (The procedure for constructing the WS cell, you may note, is analogous to that used in constructing the Brillouin zone in k-space.) For Na, which has a bcc structure, the WS cell has the shape of a regular dodecahedron (similar to Fig. 5.8b, but in real space).

In order to solve (5.50), we now assume that the electron, when in a particular cell, say A, is influenced by the potential of the ion in that cell only. The ions in other cells have a negligible effect on the electron in cell A because each of these cells is occupied, on the average, by another conduction electron which tends to screen the ion, thereby reducing its potential drastically. To ensure that the function ψ_k satisfies the Bloch form, it is necessary that u_k—where $\psi_k = e^{ik \cdot r} u_k$ —be periodic, that is, u_k has the same points on opposite faces of the cell, e.g., points P_1 and P_2 in Fig. 5.18(a).
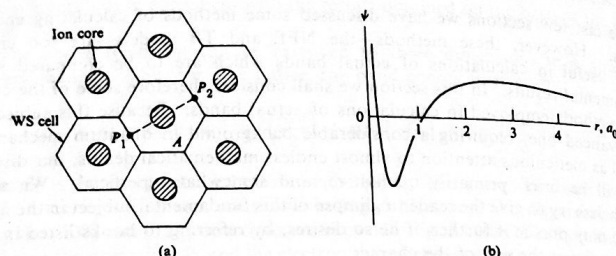

Fig. 5.18 (a) The WS cell. (b) The wave function ψ_0 at the bottom of the 3s band in Na versus the radial distance, in units of the Bohr radius.

The procedure is now clear in principle: We attempt to solve (5.50) in a single cell, using for $V(r)$ the potential of a *free* ion, which can be found from atomic physics. In Na, for instance, $V(r)$ is the potential of the ion core Na$^+$. It is still very difficult, however, to impose the requirements of periodicity on the function for the actual shape of the cell (the truncated octahedron), and to overcome this hurdle Wigner and Seitz replaced the cell by a WS *sphere* of the same volume as the actual cell, i.e., one employs a *WS sphere*. Using these simplifying assumptions concerning the potential and the periodic conditions, one then solves the Schrödinger equation numerically, since an analytical solution cannot usually be found. The resulting wave function ψ_0 at the bottom of the band, $k = 0$, is shown in Fig. 5.18(b). The wave functions at other values of k near the bottom of the band may then be approximated by

$$\psi_k \simeq \frac{1}{V^{1/2}} e^{ik \cdot r} \psi_0, \tag{5.51}$$

which has the Bloch form.

The procedure is also capable of yielding the energy $E(k)$. The energy E_0 of the bottom of the band is obtained from the same calculations which give ψ_0,

and the energy at any other point k is obtained by using

$$E(k) = \left\langle \psi_k \left| -\frac{\hbar^2}{2m_0} \nabla^2 + V(r) \right| \psi_k \right\rangle, \tag{5.52}$$

where the wave function ψ_k is substituted from (5.51). The energy found in this manner was used by Wigner and Seitz to evaluate the cohesive energy, and the results are in satisfactory agreement with experiment.

One noteworthy feature of these results is the shape of the wave function in Fig. 5.18(b). The wave function oscillates at the ion core, but once outside the core the function is essentially a constant. This constancy of the wave function holds true for almost 90% of the cell volume. Thus the wave function behaves like a plane wave, as seen from (5.51), over most of the cell, and hence over most of the crystal. Looking at this in terms of the potential, we see that where the function is a plane wave, the potential must be a constant. Thus the *effective* potential acting on the electron is essentially a constant, except in the region at the ion core itself. Viewing the motion of the electron in the crystal as a whole, we conclude that the electron moves in a region of constant potential throughout most of the crystal; only at the cores themselves does the electron experience any appreciable potential. This surprising result explains why the conduction electrons in Na, for example, may be regarded as essentially free electrons. Mathematically, it is a consequence of the periodic conditions imposed on the wave function in the cell, and this is particularly apparent when one realizes that the wave function for the 3s electron in a free Na atom is very unlike ψ_0 outside the ion core. The flatness of ψ_0 is thus due to the imposition of the periodic conditions, and not to any special property of the ionic potential.[†] The effect of the periodic condition is to cancel out the ionic potential outside the core, and thus render the potential a constant. We shall find this result very useful in the development of other methods of band calculation.

Despite its usefulness, the cellular method is greatly oversimplified, and is not currently much in use. One of its chief disadvantages is that when one replaces the WS cell by a sphere, one ignores the crystal structure entirely. All anisotropic effects, for instance, are completely masked out.

The augmented-plane wave (APW) method

The APW method (Slater, 1937) uses the results of the cellular method, but is so formulated as to avoid its shortcomings. Since the effective crystal potential was found to be constant in most of the open spaces between the cores, the APW method begins by assuming such a potential (Fig. 5.19), which is referred to as the

[†] The boundary conditions require that the derivative of the function ψ_0 vanish at the surface of the WS sphere (why?). Thus the function is flat near the surface of this sphere, as shown in Fig. 5.18(b).
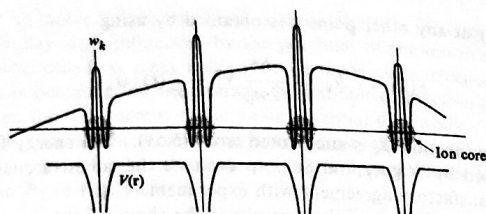
5.9

Fig. 5.19 The potential and wave function in the APW method.

muffin-tin potential. The potential is that of a free ion at the core, and is strictly constant outside the core. The wave function for the wave vector **k** is now taken to be

$$w_{\mathbf{k}} = \begin{cases} \dfrac{1}{V^{1/2}} e^{i\mathbf{k}\cdot\mathbf{r}}, & r > r_s, \\ \text{atomic function}, & r < r_s, \end{cases} \tag{5.53}$$

where r_s is the core radius. Outside the core the function is a plane wave because the potential is constant there. Inside the core the function is atom-like, and is found by solving the appropriate free-atom Schrödinger equation. Also, the atomic function in (5.53) is chosen such that it joins continuously to the plane wave at the surface of the sphere forming the core; this is the boundary condition here.

The function $w_{\mathbf{k}}$ does not have the Bloch form, but this can be remedied by forming the linear combination

$$\psi_{\mathbf{k}} = \sum_{\mathbf{G}} a_{\mathbf{k}+\mathbf{G}} \, w_{\mathbf{k}+\mathbf{G}}, \tag{5.54}$$

where the sum is over the reciprocal lattice vectors, which has the proper form. The coefficients $a_{\mathbf{k}+\mathbf{G}}$ are determined by requiring that $\psi_{\mathbf{k}}$ minimize the energy.[†] In practice the series in (5.54) converges quite rapidly, and only four or five terms—or even less—suffice to give the desired accuracy.

The APW method is a sound one for calculating the band structure in metals, and has been used a great deal in the past few years. It incorporates the essential features of the problem in a straightforward and natural fashion.

The pseudopotential method

Yet another method popular among solid-state physicists for calculating band structure in solids is the pseudopotential method, which is distinguished by the manner

---

† The "best" linear combination (5.54) is that which makes the energy as low as possible.

5.9

in which the wave function is chosen. We seek a function which oscillates rapidly inside the core, but runs smoothly as a plane wave in the remainder of the open space of the WS cell. Such a function was chosen in the APW method according to (5.53), but this is not the only choice possible. Suppose we take

$$w_{\mathbf{k}} = \phi_{\mathbf{k}} - \sum_i a_i v_i, \tag{5.55}$$

where $\phi_{\mathbf{k}}$ is a plane wave and v_i an atomic function. The sum over i extends over all the atomic shells which are occupied. For example, in Na, the sum extends over the 1s, 2s, and 2p shells. The coefficients a_i are chosen such that the function $w_{\mathbf{k}}$, representing a 3s electron, is orthogonal to the core function v_i.[†] By requiring this orthogonality, we ensure that the 3s electron, when at the core, does not occupy the other atomic orbitals already occupied. Thus we avoid violating the Pauli exclusion principle.

The function $w_{\mathbf{k}}$ has the features we are seeking: Away from the core, the atomic functions v_i are negligible, and thus $w_{\mathbf{k}} \simeq \phi_{\mathbf{k}}$, a plane wave. At the core, the atomic functions are appreciable, and act so as to induce rapid oscillations, as shown in Fig. 5.20.
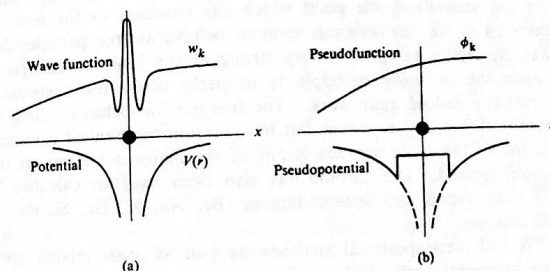

Fig. 5.20 The pseudopotential concept. (a) The actual potential and the corresponding wave function, as seen by the electron. (b) The corresponding pseudopotential and pseudofunction.

If one now substitutes $w_{\mathbf{k}}$ into the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_0} \nabla^2 + V \right] w_{\mathbf{k}} = E(\mathbf{k}) w_{\mathbf{k}}, \tag{5.56}$$

---

† Two functions ψ_1 and ψ_2 are said to be *orthogonal* if the integral $\int \psi_1{}^* \psi_2 \, d^3 r = 0$. This concept of orthogonality is very useful in quantum mechanics. The atomic functions in the various atomic shells are all mutually orthogonal.

and rearranges the terms, one finds that the equation may be written in the form

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V'\right]\phi_k = E(k)\phi_k,$$

(5.57)

where

$$V' = V - \sum_i b_i \langle v_i | V | v_i \rangle.$$

(5.58)

These results are very interesting: Equation (5.57) shows that the effective potential is given by V, while (5.58) shows that V' is weaker than V, because the second term on the right of (5.58) tends to cancel the first term. This cancellation of the crystal potential by the atomic functions is usually appreciable, often leading to a very weak potential V'. This is known as the *pseudopotential*. Since V' is so weak, the wave function as seen from (5.57) is almost a plane wave, given by ϕ_k, and is called the *pseudofunction*.

The pseudopotential and pseudofunction are illustrated graphically in Fig. 5.20(b). Note that the potential is quite weak, and, in particular, the singularity at the ion core is entirely removed. Correspondingly, the rapid "wiggles" in the wave function have been erased, so that there is a smooth plane-wave-like function.

Now we can understand one point which has troubled us for some time: why the electrons in Na, for instance, seem to behave as free particles despite the fact that the crystal potential is very strong at the ionic cores. Now we see that, when the exclusion principle is properly taken into account, the effective potential is indeed quite weak. The free-particle behavior, long taken to be an empirical fact, is now borne out by quantum-mechanical calculations. The explanation of this basic paradox is one of the major achievements of the pseudopotential method. This method has also been used to calculate band structure in many metals and semiconductors (Be, Na, K, Ge, Si, etc.) with considerable success.

The APW and pseudopotential methods, as well as other related systems, require much numerical work which can feasibly be carried out only by modern electronic computers. It often takes a whole year or more to develop the necessary program and perform the calculations for one substance on a large computer!

5.10 METALS, INSULATORS, AND SEMICONDUCTORS

Solids are divided into two major classes: *Metals* and *insulators*. A metal—or conductor—is a solid in which an electric current flows under the application of an electric field. By contrast, application of an electric field produces no current in an insulator. There is a simple criterion for distinguishing between the two classes on the basis of the energy-band theory. This criterion rests on the following statement: *A band which is completely full carries no electric current, even in the*

presence of an electric field. It follows therefore that a solid behaves as a metal only when some of the bands are partially occupied. The proof of this statement will be supplied later (Section 5.13), but we shall accept it for the time being as an established fact.

Let us now apply this statement to Na, for example. Since the inner bands 1s, 2s, 2p are all fully occupied, they do not contribute to the current. We may therefore concern ourselves only with the topmost occupied band, the *valence band*. In Na, this is the 3s band. As we saw in Section 5.5, this band can accommodate $2N_c$ electrons, where N_c is the total number of primitive unit cells. Now in Na, a Bravais bcc lattice, each cell has one atom, which contributes one valence (or 3s) electron. Therefore the total number of valence electrons is N_c, and as these electrons occupy the band, only half of it is filled, as shown in Fig. 5.21(a). Thus sodium behaves like a metal because its valence band is only partially filled.
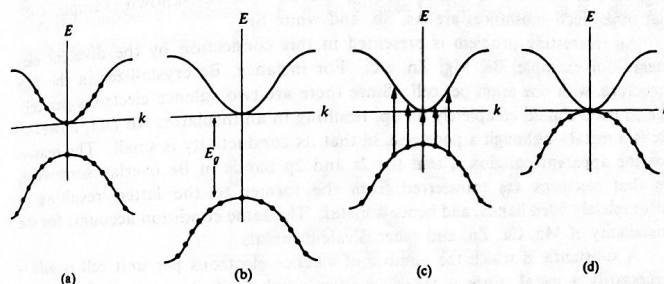

Fig. 5.21 The distribution of electrons in the bands of (a) a metal, (b) an insulator, (c) a semiconductor, and (d) a semimetal.

In a similar fashion, we conclude that the other alkalis, Li, K, etc., are also metals because their valence bands—the 2s, 4s, etc., respectively—are only partially full. The noble metals, Cu, Ag, Au, are likewise conductors for the same reason. Thus in Cu the valence band (the 4s band) is only half full, because each cell in its fcc structure contributes only one valence electron.

As an example of a good insulator, we mention diamond (carbon). Here the top band originates from a hybridization of the 2s and 2p atomic states (Section A.8), which gives rise to two bands split by an energy gap (Fig. 5.21b.) Since these bands arise from s and p states, and since the unit cell here contains two atoms, each of these bands can accommodate $8N_c$ electrons. Now in diamond each atom contributes 4 electrons, resulting in 8 valence electrons per cell. Thus the

valence band here is completely full, and the substance is an insulator, as stated above.[†]

There are substances which fall in an intermediate position between metals and insulators. If the gap between the valence band and the band immediately above it is small, then electrons are readily excitable thermally from the former to the latter band. Both bands become only partially filled and both contribute to the electric condition. Such a substance is known as a *semiconductor*. Examples are Si and Ge, in which the gaps are about 1 and 0.7 eV, respectively. By contrast, the gap in diamond is about 7 eV. Roughly speaking, a substance behaves as a semiconductor at room temperature whenever the gap is less than 2 eV.

The conductivity of a typical semiconductor is very small compared to that of a metal, but it is still many orders of magnitude larger than that of an insulator. It is justifiable, therefore, to classify semiconductors as a new class of substance, although they are, strictly speaking, insulators at very low temperatures.

In some substances the gap vanishes entirely, or the two bands even overlap slightly, and we speak of *semimetals* (Fig. 5.21d). The best-known example is Bi, but other such substances are As, Sb, and white Sn.

An interesting problem is presented in this connection by the divalent elements, for example, Be, Mg, Zn, etc. For instance, Be crystallizes in the hcp structure, with one atom per cell. Since there are two valence electrons per cell, the 2s band should completely fill up, resulting in an insulator. In fact, however, Be is a metal—although a poor one, in that its conductivity is small. The reason for the apparent paradox is that the 2s and 2p bands in Be overlap somewhat, so that electrons are transferred from the former to the latter, resulting in incompletely filled bands, and hence a metal. The same condition accounts for the metallicity of Mg, Ca, Zn, and other divalent metals.

A substance in which the number of valence electrons per unit cell is *odd* is necessarily a metal, since it takes an even number of electrons to fill a band completely. But when the number is even, the substance may be either an insulator or a metal, depending on whether the bands are disparate or overlapping.

[†] The case of hydrogen is of special interest. Although it is gaseous at atmospheric pressure, hydrogen solidifies at high pressure. But the familiar solid hydrogen is an insulator, having two atoms per unit cell, which causes the complete filling of the 1s band. Theory predicts, however, that at very high pressure ($\simeq 2$ megabars), solid hydrogen undergoes a crystal structure transformation and a concomitant change to a metallic state. Many experimenters are currently attempting to observe this transformation, and tentative successes have been reported, but definitive results are still lacking at the time of writing. Even diamond has been reported to undergo transition to the metallic state at high pressure ($\simeq 1.5$ megabars). Simultaneously a structural phase transformation to a body-centered tetragonal structure occurs. The decrease in the lattice constant caused by the pressure is about 17%.

5.11 DENSITY OF STATES

The *density of states* for electrons in a band yields the number of states in a certain energy range. This function is important in electronic processes, particularly in transport phenomena. When we denote the density-of-states function by $g(E)$, it is defined by the relation

$$g(E)\,dE = \text{number of electron states per unit volume in the energy range}$$
$$(E, E + dE). \tag{5.59}$$

This definition of $g(E)$ is analogous to that of the phonon density of states $g(\omega)$, so our discussion here parallels that presented in connection with $g(\omega)$. (See Sections 3.3 and 3.7; particularly 3.7.) To evaluate $g(E)$ one applies the definition (5.59): One draws a shell in k-space whose inner and outer surfaces are determined by the energy contours $E(\mathbf{k}) = E$ and $E(\mathbf{k}) = E + dE$, respectively, as shown in Fig. 5.22. The number of allowed k values lying inside this shell then gives the number of states which, when divided by the thickness of the shell dE, yields the desired function $g(E)$.

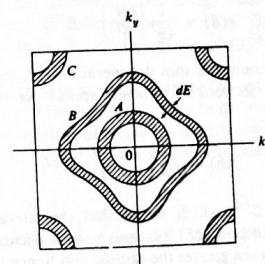
Fig. 5.22 Concentric shells in k-space used to evaluate the density of states $g(E)$.

It is evident that $g(E)$ is intimately related to the shape of the energy contours, and hence the band structure. The complexities of this structure are reflected in the form taken by $g(E)$. Let us first evaluate $g(E)$ for the case in which the dispersion relation for electron energy has the standard form

$$E = \frac{\hbar^2 k^2}{2m^*}. \tag{5.60}$$