10
The Tight-Binding Method

Linear Combinations of Atomic Orbitals

Application to Bands from *s*-Levels

General Features of Tight-Binding Levels

Wannier Functions

In Chapter 9 we calculated electronic levels in a metal by viewing it as a gas of nearly free conduction electrons, only weakly perturbed by the periodic potential of the ions. We can also take a very different point of view, regarding a solid (metal or insulator) as a collection of weakly interacting neutral atoms. As an extreme example of this, imagine assembling a group of sodium atoms into a body-centered cubic array with a lattice constant of the order of centimeters rather than angstroms. All electrons would then be in atomic levels localized at lattice sites, bearing no resemblance to the linear combinations of a few plane waves described in Chapter 9.

If we were to shrink the artificially large lattice constant of our array of sodium atoms, at some point before the actual lattice constant of metallic sodium was reached we would have to modify our identification of the electronic levels of the array with the atomic levels of isolated sodium atoms. This would become necessary for a particular atomic level, when the interatomic spacing became comparable to the spatial extent of its wave function, for an electron in that level would then feel the presence of the neighboring atoms.

The actual state of affairs for the $1s$, $2s$, $2p$ and $3s$ levels of atomic sodium is shown in Figure 10.1. The atomic wave functions for these levels are drawn about two nuclei separated by 3.7 Å, the nearest-neighbor distance in metallic sodium. The overlap of the $1s$ wave functions centered on the two sites is utterly negligible, indicating that these atomic levels are essentially unaltered in metallic sodium. The overlap of the $2s$- and $2p$-levels is exceedingly small, and one might hope to find levels in the metal very closely related to these. However, the overlap of the $3s$-levels (which hold the atomic valence electrons) is substantial, and there is no reason to expect the actual electronic levels of the metal to resemble these atomic levels.

The *tight-binding approximation* deals with the case in which the overlap of atomic wave functions is enough to require corrections to the picture of isolated atoms, but not so much as to render the atomic description completely irrelevant. The approximation is most useful for describing the energy bands that arise from the partially filled d-shells of transition metal atoms and for describing the electronic structure of insulators.

Quite apart from its practical utility, the tight-binding approximation provides an instructive way of viewing Bloch levels complementary to that of the nearly free electron picture, permitting a reconciliation between the apparently contradictory features of localized atomic levels on the one hand, and free electron-like plane-wave levels on the other.

GENERAL FORMULATION

In developing the tight-binding approximation, we assume that in the vicinity of each lattice point the full periodic crystal Hamiltonian, H, can be approximated by the Hamiltonian, H_{at}, of a single atom located at the lattice point. We also assume that the bound levels of the atomic Hamiltonian are well localized; i.e., if ψ_n is a bound level of H_{at} for an atom at the origin,

$$H_{at}\psi_n = E_n\psi_n, \tag{10.1}$$

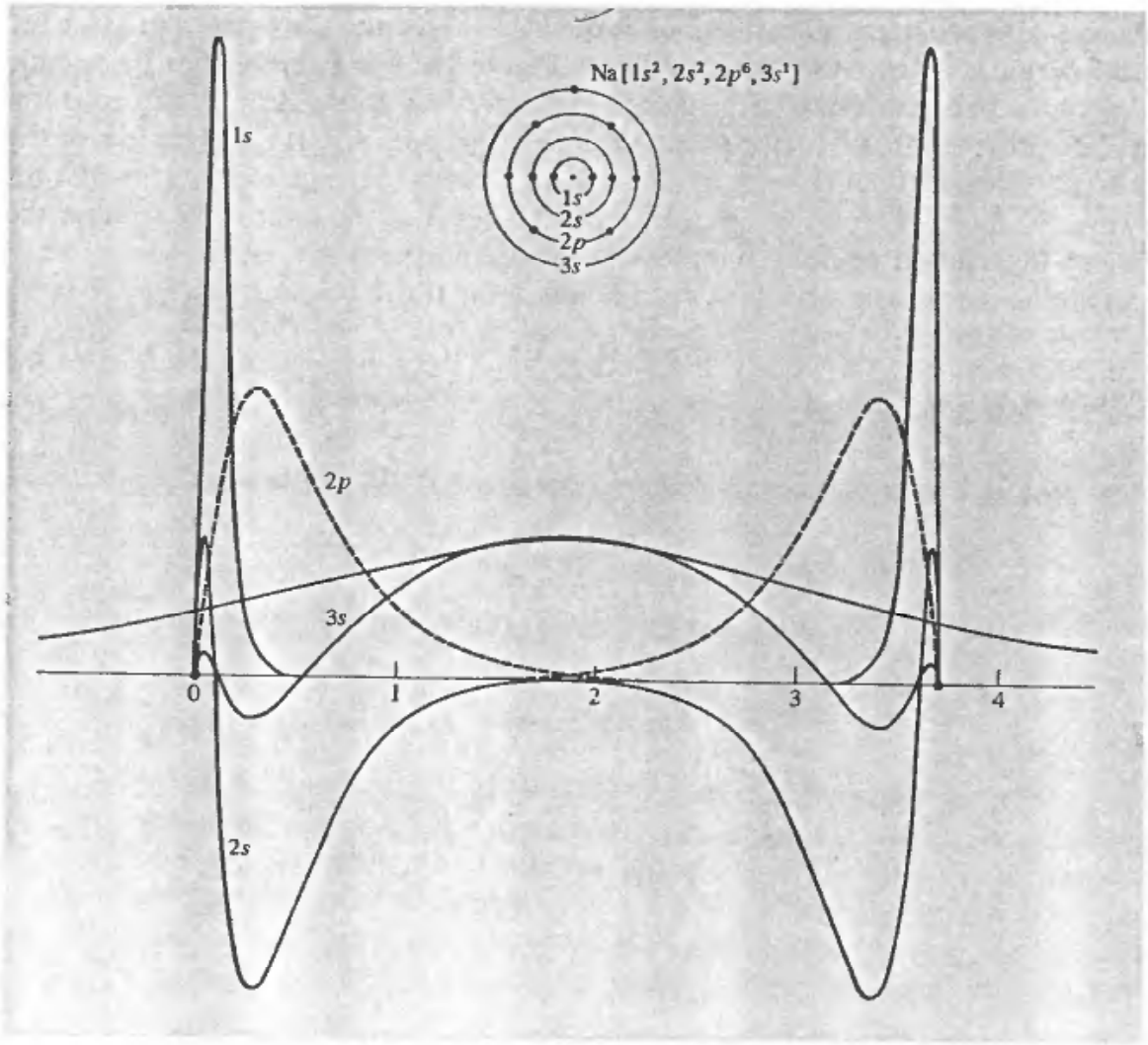

Figure 10.1
Calculated electron wave functions for the levels of atomic sodium, plotted about two nuclei separated by the nearest-neighbor distance in metallic sodium, 3.7 Å. The solid curves are $r\psi(r)$ for the 1*s*, 2*s*, and 3*s* levels. The dashed curve is r times the radial wave function for the 2*p* levels. Note how the 3*s* curves overlap extensively, the 2*s* and 2*p* curves overlap only a little, and the 1*s* curves have essentially no overlap. The curves are taken from calculations by D. R. Hartree and W. Hartree, *Proc. Roy. Soc.* **A193**, 299 (1948). The scale on the r-axis is in angstroms.

then we require that $\psi_n(\mathbf{r})$ be very small when r exceeds a distance of the order of the lattice constant, which we shall refer to as the "range" of ψ_n.

In the extreme case in which the crystal Hamiltonian begins to differ from H_{at} (for an atom whose lattice point we take as the origin) only at distances from $\mathbf{r} = \mathbf{0}$ that exceed the range of $\psi_n(\mathbf{r})$, the wave function $\psi_n(\mathbf{r})$ will be an excellent approximation to a stationary-state wave function for the full Hamiltonian, with eigenvalue E_n. So also will the wave functions $\psi_n(\mathbf{r} - \mathbf{R})$ for all $\mathbf{R}$ in the Bravais lattice, since H has the periodicity of the lattice.

To calculate corrections to this extreme case, we write the crystal Hamiltonian H as

$$H = H_{at} + \Delta U(\mathbf{r}), \tag{10.2}$$

where $\Delta U(\mathbf{r})$ contains all corrections to the atomic potential required to produce the full periodic potential of the crystal (see Figure 10.2). If $\psi_n(\mathbf{r})$ satisfies the atomic Schrödinger equation (10.1), then it will also satisfy the crystal Schrödinger equation (10.2), provided that $\Delta U(\mathbf{r})$ vanishes wherever $\psi_n(\mathbf{r})$ does not. If this were indeed the case, then each atomic level $\psi_n(\mathbf{r})$ would yield N levels in the periodic potential, with wave functions $\psi_n(\mathbf{r} - \mathbf{R})$, for each of the N sites $\mathbf{R}$ in the lattice. To preserve the Bloch description we must find the N linear combinations of these degenerate wave functions that satisfy the Bloch condition (see Eq. (8.6)):

$$\psi(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}\psi(\mathbf{r}). \tag{10.3}$$

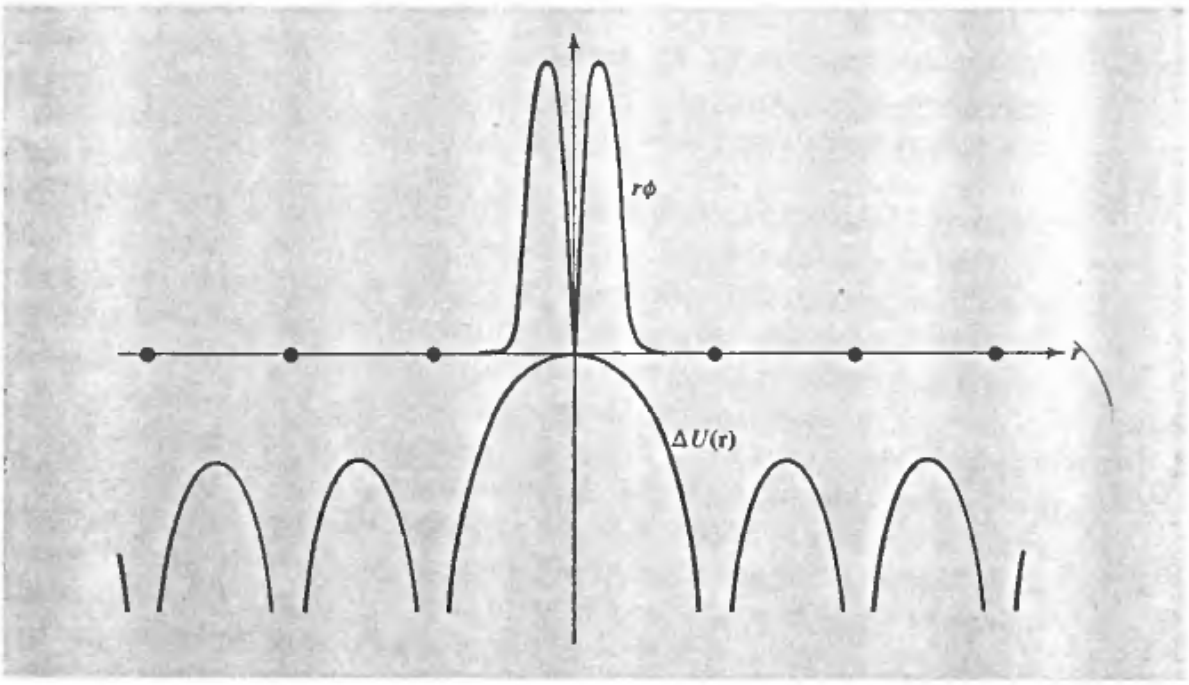

Figure 10.2
The lower curve depicts the function $\Delta U(\mathbf{r})$ drawn along a line of atomic sites. When $\Delta U(\mathbf{r})$ is added to a single atomic potential localized at the origin, the full periodic potential $U(\mathbf{r})$ is recovered. The upper curve represents r times an atomic wave function localized at the origin. When $r\phi(\mathbf{r})$ is large, $\Delta U(\mathbf{r})$ is small, and vice versa.

The N linear combinations we require are

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}\psi_n(\mathbf{r} - \mathbf{R}), \tag{10.4}$$

where $\mathbf{k}$ ranges through the N values in the first Brillouin zone consistent with the Born-von Karman periodic boundary condition.[1] The Bloch condition (10.3) is verified for the wave functions (10.4) by noting that

[1] Except when explicitly studying surface effects, one should avoid the temptation to treat a finite crystal by restricting the summation on $\mathbf{R}$ in (10.4) to the sites of a finite portion of the Bravais lattice. It is far more convenient to sum over an infinite Bravais lattice (the sum converging rapidly because of the short range of the atomic wave function ψ_n) and to represent the finite crystal with the usual Born-von Karman boundary condition, which places the standard restriction (8.27) on $\mathbf{k}$, when the Bloch condition holds. With the sum taken over all sites, for example, it is permissible to make the crucial replacement of the summation variable $\mathbf{R}'$ by $\bar{\mathbf{R}} = \mathbf{R}' - \mathbf{R}$, in the second to last line of Eq. (10.5).

$$\begin{aligned}\psi(\mathbf{r}+\mathbf{R}) &= \sum_{\mathbf{R}'} e^{i\mathbf{k}\cdot\mathbf{R}'}\psi_n(\mathbf{r}+\mathbf{R}-\mathbf{R}') \\ &= e^{i\mathbf{k}\cdot\mathbf{R}}\left[\sum_{\mathbf{R}'} e^{i\mathbf{k}\cdot(\mathbf{R}'-\mathbf{R})}\psi_n(\mathbf{r}-(\mathbf{R}'-\mathbf{R}))\right] \\ &= e^{i\mathbf{k}\cdot\mathbf{R}}\left[\sum_{\bar{\mathbf{R}}} e^{i\mathbf{k}\cdot\bar{\mathbf{R}}}\psi_n(\mathbf{r}-\bar{\mathbf{R}})\right] \\ &= e^{i\mathbf{k}\cdot\mathbf{R}}\psi(\mathbf{r}). \end{aligned} \tag{10.5}$$

Thus the wave functions (10.4) satisfy the Bloch condition with wave vector $\mathbf{k}$, while continuing to display the atomic character of the levels. The energy bands arrived at in this way, however, have little structure, $\mathcal{E}_n(\mathbf{k})$ being simply the energy of the atomic level, E_n, regardless of the value of $\mathbf{k}$. To remedy this deficiency we must recognize that a more realistic assumption is that $\psi_n(\mathbf{r})$ becomes small, but not precisely zero, before $\Delta U(\mathbf{r})$ becomes appreciable (see Figure 10.2). This suggests that we seek a solution to the full crystal Schrödinger equation that retains the general form of (10.4):[2]

$$\psi(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}\phi(\mathbf{r}-\mathbf{R}), \tag{10.6}$$

but with the function $\phi(\mathbf{r})$ not necessarily an exact atomic stationary-state wave function, but one to be determined by further calculation. If the product $\Delta U(\mathbf{r})\psi_n(\mathbf{r})$, though nonzero, is exceedingly small, we might expect the function $\phi(\mathbf{r})$ to be quite close to the atomic wave function $\psi_n(\mathbf{r})$ or to wave functions with which $\psi_n(\mathbf{r})$ is degenerate. Based on this expectation, one seeks a $\phi(\mathbf{r})$ that can be expanded in a relatively small number of localized atomic wave functions:[3,4]

$$\phi(\mathbf{r}) = \sum_n b_n\psi_n(\mathbf{r}). \tag{10.7}$$

If we multiply the crystal Schrödinger equation

$$H\psi(\mathbf{r}) = (H_{\text{at}} + \Delta U(\mathbf{r}))\psi(\mathbf{r}) = \mathcal{E}(\mathbf{k})\psi(\mathbf{r}) \tag{10.8}$$

by the atomic wave function $\psi_m^*(\mathbf{r})$, integrate over all $\mathbf{r}$, and use the fact that

$$\int \psi_m^*(\mathbf{r})H_{\text{at}}\psi(\mathbf{r})\,d\mathbf{r} = \int (H_{\text{at}}\psi_m(\mathbf{r}))^*\psi(\mathbf{r})\,d\mathbf{r} = E_m\int \psi_m^*(\mathbf{r})\psi(\mathbf{r})\,d\mathbf{r}, \tag{10.9}$$

we find that

$$(\mathcal{E}(\mathbf{k}) - E_m)\int \psi_m^*(\mathbf{r})\psi(\mathbf{r})\,d\mathbf{r} = \int \psi_m^*(\mathbf{r})\,\Delta U(\mathbf{r})\psi(\mathbf{r})\,d\mathbf{r}. \tag{10.10}$$

---

[2] It turns out (see p. 187) that any Bloch function can be written in the form (10.6), the function ϕ being known as a *Wannier function*, so no generality is lost in this assumption.

[3] By including only localized (i.e., bound) atomic wave functions in (10.7) we make our first serious approximation. A complete set of atomic levels includes the ionized ones as well. This is the point at which the method ceases to be applicable to levels well described by the almost free electron approximation.

[4] Because of this method of approximating ϕ, the tight-binding method is sometimes known as the method of the *linear combination of atomic orbitals* (or LCAO).

Placing (10.6) and (10.7) into (10.10) and using the orthonormality of the atomic wave functions,

$$\int \psi_m^*(\mathbf{r})\psi_n(\mathbf{r})\, d\mathbf{r} = \delta_{nm}, \tag{10.11}$$

we arrive at an eigenvalue equation that determines the coefficients $b_n(\mathbf{k})$ and the Bloch energies $\mathcal{E}(\mathbf{k})$:

$$\begin{aligned}(\mathcal{E}(\mathbf{k}) - E_m)b_m = &-(\mathcal{E}(\mathbf{k}) - E_m)\sum_n \left(\sum_{\mathbf{R}\neq 0} \int \psi_m^*(\mathbf{r})\psi_n(\mathbf{r} - \mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}\, d\mathbf{r}\right) b_n \\ &+ \sum_n \left(\int \psi_m^*(\mathbf{r})\, \Delta U(\mathbf{r})\psi_n(\mathbf{r})\, d\mathbf{r}\right) b_n \\ &+ \sum_n \left(\sum_{\mathbf{R}\neq 0} \int \psi_m^*(\mathbf{r})\, \Delta U(\mathbf{r})\psi_n(\mathbf{r} - \mathbf{R})\, e^{i\mathbf{k}\cdot\mathbf{R}}\, d\mathbf{r}\right) b_n. \end{aligned} \tag{10.12}$$

The first term on the right of Eq. (10.12) contains integrals of the form[5]

$$\int d\mathbf{r}\ \psi_m^*(\mathbf{r})\psi_n(\mathbf{r} - \mathbf{R}). \tag{10.13}$$

We interpret our assumption of well-localized atomic levels to mean that (10.13) is small compared to unity. We assume that the integrals in the third term on the right of Eq. (10.12) are small, since they also contain the product of two atomic wave functions centered at different sites. Finally, we assume that the second term on the right of (10.12) is small because we expect the atomic wave functions to become small at distances large enough for the periodic potential to deviate appreciably from the atomic one.[6]

Consequently, the right-hand side of (10.13) (and therefore $(\mathcal{E}(\mathbf{k}) - E_m)b_m$) is always small. This is possible if $\mathcal{E}(\mathbf{k}) - E_m$ is small whenever b_m is not (and vice versa). Thus $\mathcal{E}(\mathbf{k})$ must be close to an atomic level, say E_0, and all the b_m except those going with that level and levels degenerate with (or close to) it in energy must be small:[7]

$$\mathcal{E}(\mathbf{k}) \approx E_0, \quad b_m \approx 0 \text{ unless } E_m \approx E_0. \tag{10.14}$$

If the estimates in (10.14) were strict equalities, we would be back to the extreme case in which the crystal levels were identical to the atomic ones. Now, however, we

---

[5] Integrals whose integrands contain a product of wave functions centered on different lattice sites are known as *overlap integrals*. The tight-binding approximation exploits the smallness of such overlap integrals. They also play an important role in the theory of magnetism (Chapter 32).

[6] This last assumption is on somewhat shakier ground than the others, since the ionic potentials need not fall off as rapidly as the atomic wave functions. However, it is also less critical in determining the conclusions we shall reach, since the term in question does not depend on $\mathbf{k}$. In a sense this term simply plays the role of correcting the atomic potentials within each cell to include the fields of the ions outside the cell; it could be made as small as the other two terms by a judicious redefinition of the "atomic" Hamiltonian and levels.

[7] Note the similarity of this reasoning to that employed on pages 152 to 156. There, however, we concluded that the wave function was a linear combination of only a small number of plane waves, whose free electron energies were very close together. Here, we conclude that the wave function can be represented, through (10.7) and (10.6), by only a small number of atomic wave functions, whose atomic energies are very close together.

can determine the levels in the crystal more accurately, exploiting (10.14) to estimate the right-hand side of (10.12) by letting the sum over n run only through those levels with energies either degenerate with or very close to E_0. If the atomic level 0 is non-degenerate,[8] i.e., an s-level, then in this approximation (10.12) reduces to a single equation giving an explicit expression for the energy of the band arising from this s-level (generally referred to as an "s-band"). If we are interested in bands arising from an atomic p-level, which is triply degenerate, then (10.12) would give a set of three homogeneous equations, whose eigenvalues would give the $\mathcal{E}(\mathbf{k})$ for the three p-bands, and whose solutions $b(\mathbf{k})$ would give the appropriate linear combinations of atomic p-levels making up ϕ at the various $\mathbf{k}$'s in the Brillouin zone. To get a d-band from atomic d-levels, we should have to solve a 5×5 secular problem, etc.

Should the resulting $\mathcal{E}(\mathbf{k})$ stray sufficiently far from the atomic values at certain $\mathbf{k}$, it would be necessary to repeat the procedure, adding to the expansion (10.7) of ϕ those additional atomic levels whose energies the $\mathcal{E}(\mathbf{k})$ are approaching. In practice, for example, one generally solves a 6×6 secular problem that includes both d- and s-levels in computing the band structure of the transition metals, which have in the atomic state an outer s-shell and a partially filled d-shell. This procedure goes under the name of "s-d mixing" or "hybridization."

Often the atomic wave functions have so short a range that only nearest-neighbor terms in the sums over $\mathbf{R}$ in (10.12) need be retained, which very much simplifies subsequent analysis. We briefly illustrate the band structure that emerges in the simplest case.[9]

APPLICATION TO AN s-BAND ARISING FROM A SINGLE ATOMIC s-LEVEL

If all the coefficients b in (10.12) are zero except that for a single atomic s-level, then (10.12) gives directly the band structure of the corresponding s-band:

$$\mathcal{E}(\mathbf{k}) = E_s - \frac{\beta + \Sigma\gamma(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}}{1 + \Sigma\alpha(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}}, \tag{10.15}$$

where E_s is the energy of the atomic s-level, and

$$\beta = -\int d\mathbf{r}\, \Delta U(\mathbf{r})|\phi(\mathbf{r})|^2, \tag{10.16}$$

$$\alpha(\mathbf{R}) = \int d\mathbf{r}\, \phi^*(\mathbf{r})\phi(\mathbf{r} - \mathbf{R}), \tag{10.17}$$

and

$$\gamma(\mathbf{R}) = -\int d\mathbf{r}\, \phi^*(\mathbf{r})\, \Delta U(\mathbf{r})\phi(\mathbf{r} - \mathbf{R}). \tag{10.18}$$

---

[8] For the moment we ignore spin-orbit coupling. We can therefore concentrate entirely on the orbital parts of the levels. Spin can then be included by simply multiplying the orbital wave functions by the appropriate spinors, and doubling the degeneracy of each of the orbital levels.

[9] The simplest case is that of an s-band. The next most complicated case, a p-band, is discussed in Problem 2.

The coefficients (10.16) to (10.18) may be simplified by appealing to certain symmetries. Since ϕ is an s-level, $\phi(\mathbf{r})$ is real and depends only on the magnitude r. From this it follows that $\alpha(-\mathbf{R}) = \alpha(\mathbf{R})$. This and the inversion symmetry of the Bravais lattice, which requires that $\Delta U(-\mathbf{r}) = \Delta U(\mathbf{r})$, also imply that $\gamma(-\mathbf{R}) = \gamma(\mathbf{R})$. We ignore the terms in α in the denominator of (10.15), since they give small corrections to the numerator. A final simplification comes from assuming that only nearest-neighbor separations give appreciable overlap integrals.

Putting these observations together, we may simplify (10.15) to

$$\varepsilon(\mathbf{k}) = E_s - \beta - \sum_{\text{n.n.}} \gamma(\mathbf{R}) \cos \mathbf{k} \cdot \mathbf{R}, \tag{10.19}$$

where the sum runs only over those $\mathbf{R}$ in the Bravais lattice that connect the origin to its nearest neighbors.

To be explicit, let us apply (10.19) to a face-centered cubic crystal. The 12 nearest neighbors of the origin (see Figure 10.3) are at

$$\mathbf{R} = \frac{a}{2}(\pm 1, \pm 1, 0), \quad \frac{a}{2}(\pm 1, 0, \pm 1), \quad \frac{a}{2}(0, \pm 1, \pm 1). \tag{10.20}$$

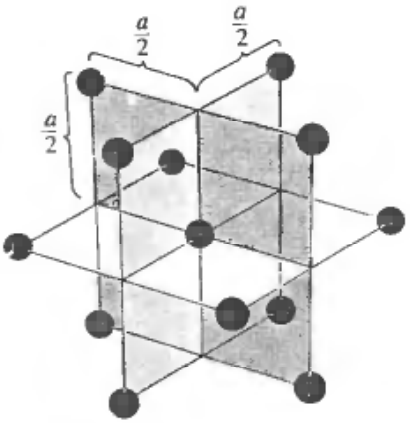

Figure 10.3
The 12 nearest neighbors of the origin in a face-centered cubic lattice with conventional cubic cell of side a.

If $\mathbf{k} = (k_x, k_y, k_z)$, then the corresponding 12 values of $\mathbf{k} \cdot \mathbf{R}$ are

$$\mathbf{k} \cdot \mathbf{R} = \frac{a}{2}(\pm k_i, \pm k_j), \qquad i, j = x, y; y, z; z, x. \tag{10.21}$$

Now $\Delta U(\mathbf{r}) = \Delta U(x, y, z)$ has the full cubic symmetry of the lattice, and is therefore unchanged by permutations of its arguments or changes in their signs. This, together with the fact that the s-level wave function $\phi(\mathbf{r})$ depends only on the magnitude of $\mathbf{r}$, implies that $\gamma(\mathbf{R})$ is the same constant γ for all 12 of the vectors (10.20). Consequently, the sum in (10.19) gives, with the aid of (10.21),

$$\varepsilon(\mathbf{k}) = E_s - \beta - 4\gamma(\cos \tfrac{1}{2}k_x a \cos \tfrac{1}{2}k_y a + \cos \tfrac{1}{2}k_y a \cos \tfrac{1}{2}k_z a + \cos \tfrac{1}{2}k_z a \cos \tfrac{1}{2}k_x a), \tag{10.22}$$

where

$$\gamma = -\int d\mathbf{r}\, \phi^*(x, y, z)\, \Delta U(x, y, z)\, \phi(x - \tfrac{1}{2}a, y - \tfrac{1}{2}a, z). \tag{10.23}$$

Equation (10.22) reveals the characteristic feature of tight-binding energy bands: The bandwidth—i.e., the spread between the minimum and maximum energies in the band—is proportional to the small overlap integral γ. Thus the tight-binding

bands are narrow bands, and the smaller the overlap, the narrower the band. In the limit of vanishing overlap the bandwidth also vanishes, and the band becomes N-fold degenerate, corresponding to the extreme case in which the electron simply resides on any one of the N isolated atoms. The dependence of bandwidth on overlap integral is illustrated in Figure 10.4.

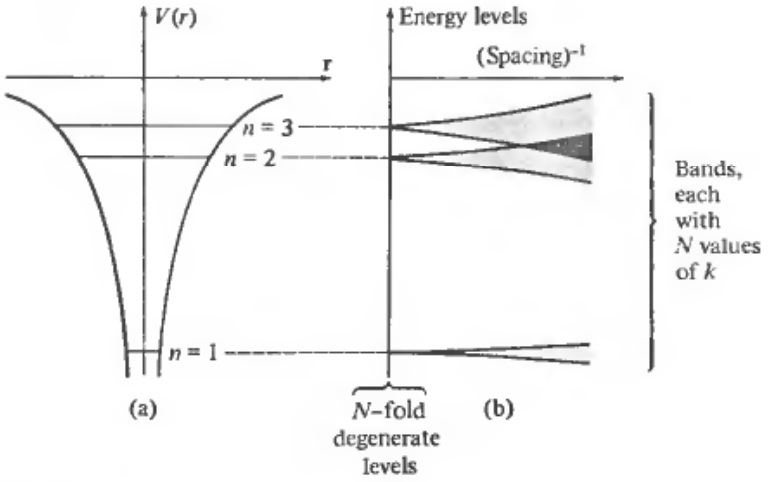

Figure 10.4
(a) Schematic representation of nondegenerate electronic levels in an atomic potential. (b) The energy levels for N such atoms in a periodic array, plotted as a function of mean inverse interatomic spacing. When the atoms are far apart (small overlap integrals) the levels are nearly degenerate, but when the atoms are closer together (larger overlap integrals), the levels broaden into bands.

In addition to displaying the effect of overlap on bandwidth, Eq. (10.22) illustrates several general features of the band structure of a face-centered cubic crystal that are not peculiar to the tight-binding case. Typical of these are the following:

1. In the limit of small ka, (10.22) reduces to:

$$\mathcal{E}(\mathbf{k}) = E_s - \beta - 12\gamma + \gamma k^2 a^2. \tag{10.24}$$

   This is independent of the direction of $\mathbf{k}$—i.e., the constant-energy surfaces in the neighbourhood of $\mathbf{k} = \mathbf{0}$ are spherical.[10]
2. If $\mathcal{E}$ is plotted along any line perpendicular to one of the square faces of the first Brillouin zone (Figure 10.5), it will cross the square face with vanishing slope (Problem 1).

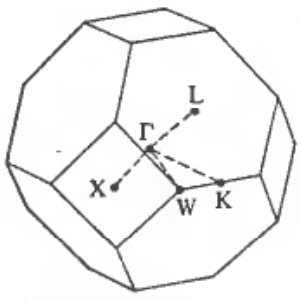

Figure 10.5
The first Brillouin zone for face-centered cubic crystals. The point Γ is at the center of the zone. The names K, L, W, and X are widely used for the points of high symmetry on the zone boundary.

[10] This can be deduced quite generally for any nondegenerate band in a crystal with cubic symmetry.

3. If $\mathcal{E}$ is plotted along any line perpendicular to one of the hexagonal faces of the first Brillouin zone (Figure 10.5), it need not, in general, cross the face with vanishing slope (Problem 1).[11]

GENERAL REMARKS ON THE TIGHT-BINDING METHOD

1. In cases of practical interest more than one atomic level appears in the expansion (10.7), leading to a 3×3 secular problem in the case of three p-levels, a 5×5 secular problem for five d-levels, etc. Figure 10.6, for example, shows the band structure that emerges from a tight-binding calculation based on the 5-fold degenerate atomic 3-d levels in nickel. The bands are plotted for three directions of symmetry in the zone, each of which has its characteristic set of degeneracies.[12]

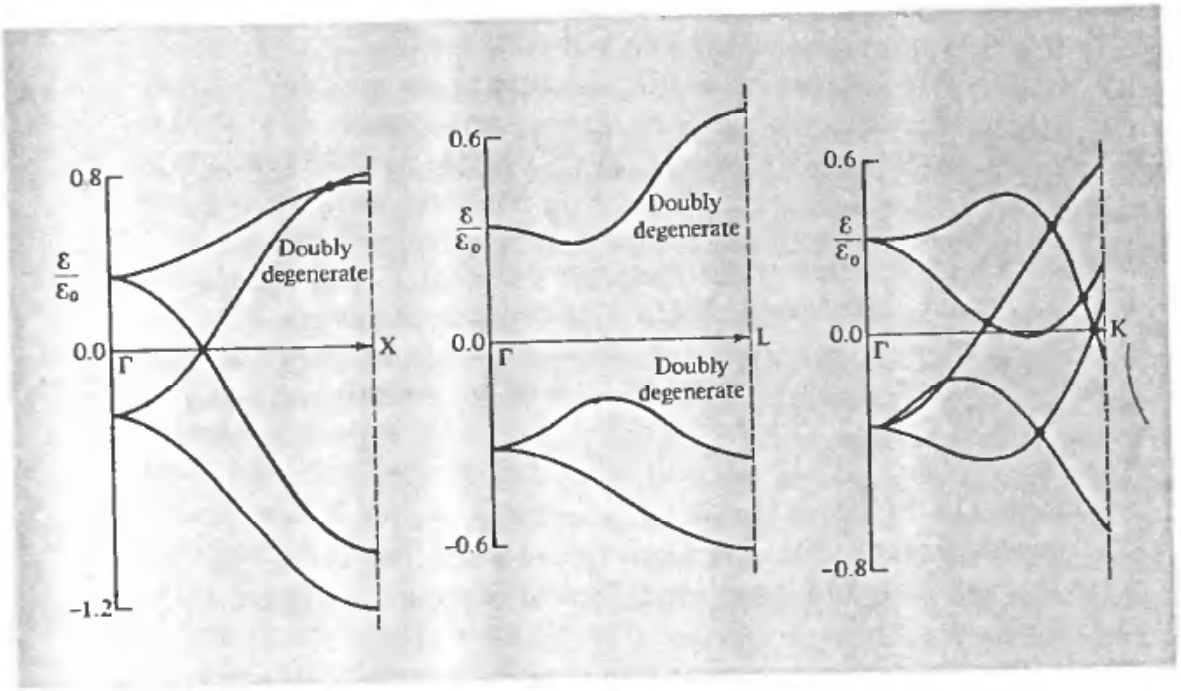

Figure 10.6
A tight-binding calculation of the 3d bands of nickel. (G. C. Fletcher, *Proc. Phys. Soc.* **A65**, 192 (1952).) Energies are given in units of $\mathcal{E}_0 = 1.349$ eV, so the bands are about 2.7 volts wide. The lines along which $\mathcal{E}$ is plotted are shown in Figure 10.5. Note the characteristic degeneracies along ΓX and ΓL, and the absence of degeneracy along ΓK. The great width of the bands indicates the inadequacy of so elementary a treatment.

2. A quite general feature of the tight-binding method is the relation between bandwidth and the overlap integrals

$$\gamma_{ij}(\mathbf{R}) = -\int d\mathbf{r}\, \phi_i^*(\mathbf{r})\, \Delta U(\mathbf{r}) \phi_j(\mathbf{r} - \mathbf{R}). \tag{10.25}$$

---

[11] Compare the nearly free electron case (page 158), where the rate of change of $\mathcal{E}$ along a line normal to a Bragg plane was always found to vanish as the plane was crossed at points far from any other Bragg planes. The tight-binding result illustrates the more general possibility that arises because there is no plane of mirror symmetry parallel to the hexagonal face.

[12] The calculated bands are so wide as to cast doubt on the validity of the entire expansion. A more realistic calculation would have to include, at the very least, the effects of the 4s-level.

If the γ_{ij} are small, then the bandwidth is correspondingly small. As a rule of thumb, when the energy of a given atomic level increases (i.e., the binding energy decreases) so does the spatial extent of its wave function. Correspondingly, the low-lying bands in a solid are very narrow, but bandwidths increase with mean band energy. In metals the highest band (or bands) are very broad, since the spatial ranges of the highest atomic levels are comparable to a lattice constant, and the tight-binding approximation is then of doubtful validity.

3. Although the tight-binding wave function (10.6) is constructed out of localized atomic levels ϕ, an electron in a tight-binding level will be found, with equal probability, in any cell of the crystal, since its wave function (like any Bloch wave function) changes only by the phase factor $e^{i\mathbf{k}\cdot\mathbf{R}}$ as one moves from one cell to another a distance $\mathbf{R}$ away. Thus as $\mathbf{r}$ varies from cell to cell, there is superimposed on the atomic structure within each cell a sinusoidal variation in the amplitudes of Re ψ and Im ψ, as illustrated in Figure 10.7.

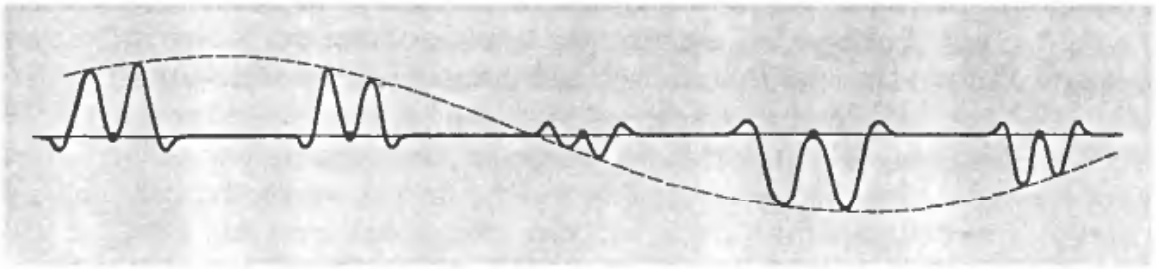

Figure 10.7
Characteristic spatial variation of the real (or imaginary) part of the tight-binding wave function (10.6).

A further indication that the tight-binding levels have a running wave or itinerant character comes from the theorem that the mean velocity of an electron in a Bloch level with wave vector $\mathbf{k}$ and energy $\varepsilon(\mathbf{k})$ is given by $\mathbf{v}(\mathbf{k}) = (1/\hbar)\,\partial\varepsilon/\partial\mathbf{k}$. (See Appendix E.) If ε is independent of $\mathbf{k}$, $\partial\varepsilon/\partial\mathbf{k}$ is zero, which is consistent with the fact that in genuinely isolated atomic levels (which lead to zero bandwidth) the electrons are indeed tied to individual atoms. If, however, there is any nonzero overlap in the atomic wave functions, then $\varepsilon(\mathbf{k})$ will not be constant throughout the zone. Since a small variation in ε implies a small nonzero value of $\partial\varepsilon/\partial\mathbf{k}$, and hence a small but nonzero mean velocity, as long as there is any overlap electrons will be able to move freely through the crystal! Decreasing the overlap only reduces the velocity; it does not eliminate the motion. One can view this motion as a quantum-mechanical tunneling from lattice site to lattice site. The less the overlap, the lower the tunneling probability, and hence the longer it takes to go a given distance.

4. In solids that are not monatomic Bravais lattices, the tight-binding approximation is more complicated. This problem arises in the hexagonal close-packed metals, which are simple hexagonal with a two-point basis. Formally, one can treat the two-point basis as a molecule, whose wave functions are assumed to be known, and proceed as above, using molecular instead of atomic wave functions. If the nearest-neighbor overlap remains small, then, in particular, it will be small in each "molecule," and an atomic s-level will give rise to two nearly degenerate molecular levels. Thus a single atomic s-level yields two tight-binding bands in the hexagonal close-packed structure.

Alternatively, one can proceed by continuing to construct linear combinations of atomic levels centered at the Bravais lattice points *and* at the basis points, generalizing (10.6) to

$$\psi(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}(a\phi(\mathbf{r} - \mathbf{R}) + b\phi(\mathbf{r} - \mathbf{d} - \mathbf{R})), \tag{10.26}$$

(where $\mathbf{d}$ is the separation of the two basis atoms). This can be viewed as essentially the first approach, in which, however, approximate molecular wave functions are used, the approximation to the molecular levels being combined with the tight-binding approximation to the levels of the entire crystal.[13]

5. In the heavier elements spin-orbit coupling (see page 169) is of great importance in determining the atomic levels, and should therefore be included in a tight-binding treatment of the broadening of these levels into bands in the solid. In principle the extension is straightforward. We simply include in $\Delta U(\mathbf{r})$ the interaction between the electron's spin and the electric field of all ions except the one at the origin, incorporating that interaction into the atomic Hamiltonian. Once this is done we can no longer use spin-independent linear combinations of atomic orbital wave functions, but must work with linear combinations of both orbital and spin levels. Thus the tight-binding theory of an *s*-level, when spin-orbit coupling is appreciable, would approximate ϕ not by a single atomic *s*-level but by a linear combination (with $\mathbf{k}$ dependent coefficients) of two levels with the same orbital wave functions and two opposite spins. The tight-binding theory of a *d*-band would go from a 5×5 determinantal problem to a 10×10 one, etc. As mentioned in Chapter 9, effects of spin-orbit coupling, though often small, can frequently be quite crucial, as when they eliminate degeneracies that would rigorously be present if such coupling were ignored.[14]

6. All the analysis of electronic levels in a periodic potential in this chapter (and the preceding two) has been done within the independent electron approximation, which either neglects the interaction between electrons, or, at best, includes it in some average way through the effective periodic potential experienced by each single electron. We shall see in Chapter 32 that the independent electron approximation can fail when it gives at least one *partially* filled band that derives from well-localized atomic levels with small overlap integrals. In many cases of interest (notably in insulators and for the very low-lying bands in metals) this problem does not arise, since the tight-binding bands are so low in energy as to be completely filled. However, the possibility of such a failure of the independent electron approximation must be kept in mind when narrow tight-binding bands are derived from partially filled atomic shells—in metals, generally the *d*- and *f*-shells. One should be particularly aware of this possibility in solids with a magnetic structure.

This failure of the independent electron approximation obscures the simple picture the tight-binding approximation suggests: that of a continuous transition from the

---

[13] The "approximate molecular wave functions" will thus be $\mathbf{k}$-dependent.

[14] The inclusion of spin-orbit coupling in the tight-binding method is discussed by J. Friedel, P. Lenghart, and G. Leman, *J. Phys. Chem. Solids* **25**, 781 (1964).

metallic to the atomic state as the interatomic distance is continuously increased.[15] If we took the tight-binding approximation at face value, then as the lattice constant in a metal increased, the overlap between all atomic levels would eventually become small, and all bands—even the partially filled conduction band (or bands)—would eventually become narrow tight-binding bands. As the conduction band narrowed, the velocity of the electrons in it would diminish and the conductivity of the metal would drop. Thus, we would expect a conductivity that dropped continuously to zero with the overlap integrals as the metal was expanded.

In fact, however, it is likely that a full calculation going beyond the independent electron approximation would predict that beyond a certain nearest-neighbor separation the conductivity should drop abruptly to zero, the material becoming an insulator (the so-called *Mott transition*).

The reason for this departure from the tight-binding prediction lies in the inability of the independent electron approximation to treat the very strong additional repulsion a second electron feels at a given atomic site when another electron is already there. We shall comment further on this in Chapter 32, but we mention the problem here because it is sometimes described as a failure of the tight-binding method.[16] This is somewhat misleading in that the failure occurs when the tight-binding approximation to the independent electron model is at its best; it is the independent electron approximation itself that fails.

WANNIER FUNCTIONS

We conclude this chapter with a demonstration that the Bloch functions for *any* band can always be written in the form (10.4) on which the tight-binding approximation is based. The functions ϕ that play the role of the atomic wave functions are known as *Wannier functions*. Such Wannier functions can be defined for any band, whether or not it is well described by the tight-binding approximation; but if the band is not a narrow tight-binding band, the Wannier functions will bear little resemblance to any of the electronic wave functions for the isolated atom.

To establish that any Bloch function $\psi_{n\mathbf{k}}(\mathbf{r})$ can be written in the form (10.4), we first note that considered as a function of $\mathbf{k}$ for fixed $\mathbf{r}$, $\psi_{n\mathbf{k}}(\mathbf{r})$ is periodic in the reciprocal lattice. It therefore has a Fourier series expansion in plane waves with wave vectors in the reciprocal of the reciprocal lattice, i.e., in the direct lattice. Thus for any fixed $\mathbf{r}$ we can write

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} f_n(\mathbf{R}, \mathbf{r}) e^{i\mathbf{R}\cdot\mathbf{k}}, \tag{10.27}$$

where the coefficients in the sum depend on $\mathbf{r}$ as well as on the "wave vectors" $\mathbf{R}$, since for each $\mathbf{r}$ it is a different function of $\mathbf{k}$ that is being expanded.

---

[15] A difficult procedure to realize in the laboratory, but a very tempting one to visualize theoretically, as an aid in understanding the nature of energy bands.

[16] See, for example, H. Jones, *The Theory of Brillouin Zones and Electron States in Crystals*, North-Holland, Amsterdam, 1960, p. 229.

The Fourier coefficients in (10.27) are given by the inversion formula[17]

$$f_n(\mathbf{R}, \mathbf{r}) = \frac{1}{v_0} \int d\mathbf{k}\, e^{-i\mathbf{R}\cdot\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}). \tag{10.28}$$

Equation (10.27) is of the form (10.4), provided that the function $f_n(\mathbf{R}, \mathbf{r})$ depends on $\mathbf{r}$ and $\mathbf{R}$ only through their difference, $\mathbf{r} - \mathbf{R}$. But if $\mathbf{r}$ and $\mathbf{R}$ are both shifted by the Bravais lattice vector $\mathbf{R}_0$, then f is indeed unchanged as a direct consequence of (10.28) and Bloch's theorem, in the form (8.5). Thus $f_n(\mathbf{R}, \mathbf{r})$ has the form:

$$f_n(\mathbf{R}, \mathbf{r}) = \phi_n(\mathbf{r} - \mathbf{R}) \tag{10.29}$$

Unlike tight-binding atomic functions $\phi(\mathbf{r})$, the Wannier functions $\phi_n(\mathbf{r} - \mathbf{R})$ at different sites (or with different band indices) are orthogonal (see Problem 3, Eq. (10.35)). Since the complete set of Bloch functions can be written as linear combinations of the Wannier functions, the Wannier functions $\phi_n(\mathbf{r} - \mathbf{R})$ for all n and $\mathbf{R}$ form a complete orthogonal set. They therefore offer an alternative basis for an exact description of the independent electron levels in a crystal potential.

The similarity in form of the Wannier functions to the tight-binding functions leads one to hope that the Wannier functions will also be localized—i.e., that when $\mathbf{r}$ is very much larger than some length on the atomic scale, $\phi_n(\mathbf{r})$ will be negligibly small. To the extent that this can be established, the Wannier functions offer an ideal tool for discussing phenomena in which the spatial localization of electrons plays an important role. Perhaps the most important areas of application are these:

1. Attempts to derive a transport theory for Bloch electrons. The analog of free electron wave packets, electronic levels in a crystal that are localized in both $\mathbf{r}$ and $\mathbf{k}$, are conveniently constructed with the use of Wannier functions. The theory of Wannier functions is closely related to the theory of when and how the semiclassical theory of transport by Bloch electrons (Chapters 12 and 13) breaks down.
2. Phenomena involving localized electronic levels, due, for example, to attractive impurities that bind an electron. A very important example is the theory of donor and acceptor levels in semiconductors (Chapter 28).
3. Magnetic phenomena, in which localized magnetic moments are found to exist at suitable impurity sites.

Theoretical discussions of the range of Wannier functions are in general quite subtle.[18] Roughly speaking, the range of the Wannier function decreases as the band gap increases (as one might expect from the tight-binding approximation, in which the bands become narrower as the range of the atomic wave functions decreases). The various "breakdown" and "breakthrough" phenomena we shall mention in

---

[17] Here v_0 is the volume in k-space of the first Brillouin zone, and the integral is over the zone. Equations (10.27) and (10.28) (with $\mathbf{r}$ regarded as a fixed parameter) are just Eqs. (D.1) and (D.2) of Appendix D, with direct and reciprocal space interchanged.

[18] A relatively simple argument, but only in one dimension, is given by W. Kohn, *Phys. Rev.* **115**, 809 (1959). A more general discussion can be found in E. I. Blount, *Solid State Physics*, Vol. 13, Academic Press, New York, 1962, p. 305.

Chapter 12 that occur when the band gap is small find their reflection in the fact that theories based on the localization of the Wannier functions become less reliable in this limit.

PROBLEMS

1. (a) Show that along the principal symmetry directions shown in Figure 10.5 the tight-binding expression (10.22) for the energies of an s-band in a face-centered cubic crystal reduces to the following:

(i) Along ΓX $(k_y = k_z = 0, \quad k_x = \mu\, 2\pi/a, \qquad 0 \leqslant \mu \leqslant 1)$

$$\varepsilon = E_s - \beta - 4\gamma(1 + 2\cos \mu\pi).$$

(ii) Along ΓL $(k_x = k_y = k_z = \mu\, 2\pi/a, \qquad 0 \leqslant \mu \leqslant \tfrac{1}{2})$

$$\varepsilon = E_s - \beta - 12\gamma \cos^2 \mu\pi.$$

(iii) Along ΓK $(k_z = 0, \quad k_x = k_y = \mu\, 2\pi/a, \qquad 0 \leqslant \mu \leqslant \tfrac{3}{4})$

$$\varepsilon = E_s - \beta - 4\gamma(\cos^2 \mu\pi + 2\cos \mu\pi).$$

(iv) Along ΓW $(k_z = 0, \quad k_x = \mu\, 2\pi/a, \quad k_y = \tfrac{1}{2}\mu\, 2\pi/a, \qquad 0 \leqslant \mu \leqslant 1)$

$$\varepsilon = E_s - \beta - 4\gamma(\cos \mu\pi + \cos \tfrac{1}{2}\mu\pi + \cos \mu\pi \cos \tfrac{1}{2}\mu\pi).$$

(b) Show that on the square faces of the zone the normal derivative of ε vanishes.

(c) Show that on the hexagonal faces of the zone, the normal derivative of ε vanishes only along lines joining the center of the hexagon to its vertices.

2. ***Tight-Binding p-Bands in Cubic Crystals***

In dealing with cubic crystals, the most convenient linear combinations of three degenerate atomic p-levels have the form $x\phi(r)$, $y\phi(r)$, and $z\phi(r)$, where the function ϕ depends only on the magnitude of the vector $\mathbf{r}$. The energies of the three corresponding p-bands are found from (10.12) by setting to zero the determinant

$$|(\varepsilon(\mathbf{k}) - E_p)\,\delta_{ij} + \beta_{ij} + \tilde{\gamma}_{ij}(\mathbf{k}) = 0, \tag{10.30}$$

where

$$\tilde{\gamma}_{ij}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}\gamma_{ij}(\mathbf{R}),$$

$$\gamma_{ij}(\mathbf{R}) = -\int d\mathbf{r}\, \psi_i^*(\mathbf{r})\psi_j(\mathbf{r} - \mathbf{R})\, \Delta U(\mathbf{r}),$$

$$\beta_{ij} = \gamma_{ij}(\mathbf{R} = 0). \tag{10.31}$$

(A term multiplying $\varepsilon(\mathbf{k}) - E_p$, which gives rise to very small corrections analogous to those given by the denominator of (10.15) in the s-band case, has been omitted from (10.30).)

(a) As a consequence of cubic symmetry, show that

$$\beta_{xx} = \beta_{yy} = \beta_{zz} = \beta,$$
$$\beta_{xy} = 0. \tag{10.32}$$

(b) Assuming that the $\gamma_{ij}(\mathbf{R})$ are negligible except for nearest-neighbor $\mathbf{R}$, show that $\tilde{\gamma}_{ij}(\mathbf{k})$ is diagonal for a simple cubic Bravais lattice, so that $x\phi(r)$, $y\phi(r)$, and $z\phi(r)$ each generate independent bands. (Note that this ceases to be the case if the $\gamma_{ij}(\mathbf{R})$ for next nearest-neighbor $\mathbf{R}$ are also retained.)

(c) For a face-centered cubic Bravais lattice with only nearest-neighbor γ_{ij} appreciable, show that the energy bands are given by the roots of

$$0 = \begin{vmatrix} \mathcal{E}(\mathbf{k}) - \mathcal{E}^0(\mathbf{k}) + 4\gamma_0 \cos\tfrac{1}{2}k_y a \cos\tfrac{1}{2}k_z a & -4\gamma_1 \sin\tfrac{1}{2}k_x a \sin\tfrac{1}{2}k_y a & -4\gamma_1 \sin\tfrac{1}{2}k_x a \sin\tfrac{1}{2}k_z a \\ -4\gamma_1 \sin\tfrac{1}{2}k_y a \sin\tfrac{1}{2}k_x a & \mathcal{E}(\mathbf{k}) - \mathcal{E}^0(\mathbf{k}) + 4\gamma_0 \cos\tfrac{1}{2}k_z a \cos\tfrac{1}{2}k_x a & -4\gamma_1 \sin\tfrac{1}{2}k_y a \sin\tfrac{1}{2}k_z a \\ -4\gamma_1 \sin\tfrac{1}{2}k_z a \sin\tfrac{1}{2}k_x a & -4\gamma_1 \sin\tfrac{1}{2}k_z a \sin\tfrac{1}{2}k_y a & \mathcal{E}(\mathbf{k}) - \mathcal{E}^0(\mathbf{k}) + 4\gamma_0 \cos\tfrac{1}{2}k_x a \cos\tfrac{1}{2}k_y a \end{vmatrix} \tag{10.33}$$

where

$$\begin{aligned} \mathcal{E}^0(\mathbf{k}) &= E_p - \beta \\ &\quad - 4\gamma_2(\cos\tfrac{1}{2}k_x a \cos\tfrac{1}{2}k_z a + \cos\tfrac{1}{2}k_x a \cos\tfrac{1}{2}k_y a + \cos\tfrac{1}{2}k_y a \cos\tfrac{1}{2}k_z a), \\ \gamma_0 &= -\int d\mathbf{r}\,[x^2 - y(y - \tfrac{1}{2}a)]\phi(r)\phi([x^2 + (y - \tfrac{1}{2}a)^2 + (z - \tfrac{1}{2}a)^2]^{1/2})\,\Delta U(\mathbf{r}), \\ \gamma_1 &= -\int d\mathbf{r}\, x(y - \tfrac{1}{2}a)\phi(r)\phi([(x - \tfrac{1}{2}a)^2 + (y - \tfrac{1}{2}a^2) + z^2]^{1/2})\,\Delta U(\mathbf{r}), \\ \gamma_2 &= -\int d\mathbf{r}\, x(x - \tfrac{1}{2}a)\phi(r)\phi([(x - \tfrac{1}{2}a)^2 + (y - \tfrac{1}{2}a)^2 + z^2]^{1/2})\,\Delta U(\mathbf{r}). \end{aligned} \tag{10.34}$$

(d) Show that all three bands are degenerate at $\mathbf{k} = \mathbf{0}$, and that when $\mathbf{k}$ is directed along either a cube axis (ΓX) or a cube diagonal (ΓL) there is a double degeneracy. Sketch the energy bands (in analogy to Figure 10.6) along these directions.

3. Prove that Wannier functions centered on different lattice sites are orthogonal,

$$\int \phi_n^*(\mathbf{r} - \mathbf{R})\phi_{n'}(\mathbf{r} - \mathbf{R}')\,d\mathbf{r} \propto \delta_{n,n'}\,\delta_{\mathbf{R},\mathbf{R}'}, \tag{10.35}$$

by appealing to the orthonormality of the Bloch functions and the identity (F.4) of Appendix F. Show also that

$$\int d\mathbf{r} \left|\phi_n(\mathbf{r})\right|^2 = 1 \tag{10.36}$$

if the integral of the $|\psi_{n\mathbf{k}}(\mathbf{r})|^2$ over a primitive cell is normalized to unity.