

Ashcroft & Mermin

Chapter 11

Other Methods for Calculating Band Structure

Part I

In Chapters 9 and 10 we explored approximate solutions to the one-electron Schrödinger equation in the limiting cases of nearly free electrons, and tight binding. In most cases of interest the tight-binding approximation (at least in the simple form outlined in Chapter 10) is suitable only for the representation of bands arising from the ion core levels, while the nearly free electron approximation cannot be directly applied to any real solid.¹ The purpose of this chapter is therefore to describe some of the more common methods actually used in the calculation of real band structures.

We remarked in Chapter 8 that in merely writing down a separate Schrödinger equation²

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right) \psi_{\mathbf{k}}(\mathbf{r}) = \varepsilon(\mathbf{k}) \psi_{\mathbf{k}}(\mathbf{r}) \quad (11.1)$$

for each electron we are already enormously simplifying the actual problem of many *interacting* electrons in a periodic potential. In an exact treatment each electron cannot be described by a wave function determined by a single-particle Schrödinger equation, independent of all the others.

The independent electron approximation does not in fact entirely neglect electron-electron interactions. Rather it assumes that most of their important effects can be taken into account with a sufficiently clever choice for the periodic potential $U(\mathbf{r})$ appearing in the one-electron Schrödinger equation. Thus $U(\mathbf{r})$ contains not only the periodic potential due to the ions alone, but also periodic effects due to the interaction of the electron (whose wave function appears in (11.1)) with all the other electrons. The latter interaction depends on the configuration of the other electrons; i.e., it depends on *their* individual wave functions, which are also determined by a Schrödinger equation of the form (11.1). Thus to know the potential appearing in (11.1), one must first know all the solutions to (11.1). Since, however, to know the solutions one must know the potential, one is in for some difficult mathematical efforts.

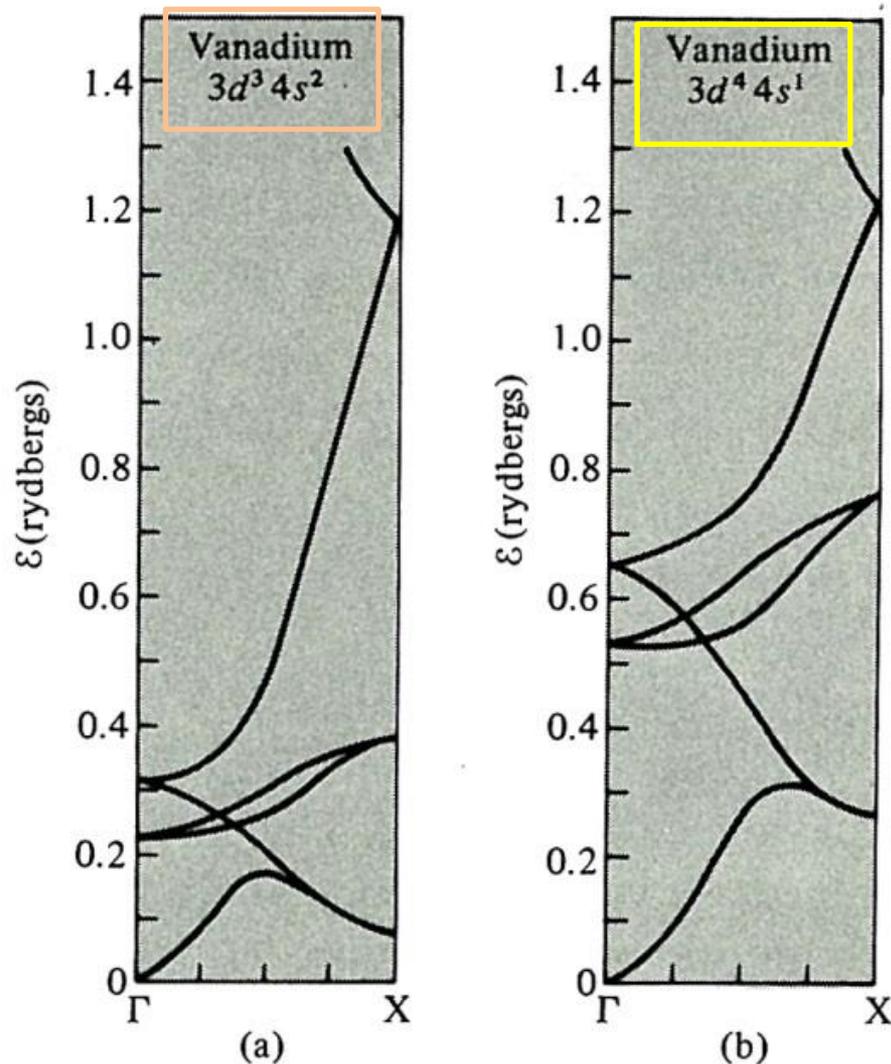
The simplest (and often the most practical) procedure is to start with a shrewd guess, $U_0(\mathbf{r})$, for $U(\mathbf{r})$, calculate from (11.1) the wave functions for the occupied electronic levels, and from these recompute $U(\mathbf{r})$. If the new potential, $U_1(\mathbf{r})$ is the same as (or very close to) $U_0(\mathbf{r})$, one says that *self-consistency* has been achieved and takes $U = U_1$ for the actual potential. If U_1 differs from U_0 , one repeats the procedure starting with U_1 , taking U_2 as the actual potential if it is very close to U_1 , and otherwise continuing on to the calculation of U_3 . The hope is that this procedure will converge, eventually yielding a self-consistent potential that reproduces itself.³

We shall assume in this chapter (as in Chapters 8–10) that the potential $U(\mathbf{r})$ is a given function; i.e., that we are either engaged in the first step of this iterative procedure or, by a fortunate guess, are able to work with a reasonably self-consistent $U(\mathbf{r})$ from the start. The reliability of the methods we are about to describe is limited not only by the accuracy of the computed solutions to (11.1), which can be quite high, but also by the accuracy with which we have been able to estimate the potential $U(\mathbf{r})$. The resulting $\varepsilon_n(\mathbf{k})$ display a disconcerting sensitivity to errors in the construction of the potential, and it is often the case that the final accuracy of the computed band structure is limited more by the problem of finding the potential than by the difficulties in solving the Schrödinger equation (11.1) for a given U . This is strikingly illustrated in Figure 11.1.

Another point to emphasize at the start is that none of the methods we shall describe can be carried through analytically, except in the simplest one-dimensional examples. All require modern, high-speed computers for their execution. Progress in the theoretical calculation of energy bands has kept close pace with development of larger and faster computers, and the kinds of approximations one is likely to consider are influenced by available computational techniques.⁴

Figure 11.1

Energy bands for vanadium, calculated for two possible choices of crystal potential $U(\mathbf{r})$. Vanadium is body-centered cubic and the bands are plotted along the $[100]$ direction from the origin to the Brillouin zone boundary. The atomic structure of vanadium is five electrons around a closed-shell argon configuration. The bands displayed are the $3d$ and $4s$ derived bands (and higher bands). (a) The bands are shown as calculated in a $U(\mathbf{r})$ derived from an assumed $3d^3 4s^2$ configuration for atomic vanadium. (b) The bands are shown based on an assumed $3d^4 4s^1$ atomic configuration. (From L. F. Matheiss, *Phys. Rev. A* **970** 134, (1964).)



GENERAL FEATURES OF VALENCE-BAND WAVE FUNCTIONS

Since the low-lying core levels are well described by tight-binding wave functions, calculational methods aim at the higher-lying bands (which may be either filled, partially filled, or empty). These bands are referred to in this context, in contrast to the tight-binding core bands, as the valence bands.⁵ The valence bands determine the electronic behavior of a solid in a variety of circumstances, electrons in the core levels being inert for many purposes.

The essential difficulty in practical calculations of the valence-band wave functions and energies is revealed when one asks why the nearly free electron approximation of Chapter 9 cannot be applied to the valence bands in an actual solid. A simple, but superficial, reason is that the potential is not small. Very roughly we might estimate that, at least well within the ion core, $U(\mathbf{r})$ has the coulombic form

$$\frac{-Z_a e^2}{r}, \quad (11.2)$$

where Z_a is the atomic number. The contribution of (11.2) to the Fourier components $U_{\mathbf{k}}$ in Eq. (9.2) will be (see p. 167 and Eq. (17.73)):

$$U_{\mathbf{K}} \approx - \left(\frac{4\pi Z_a e^2}{K^2} \right) \frac{1}{v}. \quad (11.3)$$

If we write this as

$$|U_{\mathbf{K}}| \approx \frac{e^2}{2a_0} \left(\frac{a_0^3}{v} \right) \frac{1}{(a_0 K)^2} 8\pi Z_a, \quad \frac{e^2}{2a_0} = 13.6 \text{ eV}, \quad (11.4)$$

we see that $U_{\mathbf{K}}$ can be of the order of several electron volts for a very large number of reciprocal lattice vectors \mathbf{K} and is therefore comparable to the kinetic energies appearing in Eq. (9.2). Thus the assumption that $U_{\mathbf{K}}$ is small compared to these kinetic energies is not permissible.

A deeper insight into this failure is afforded by considering the nature of the core and valence wave functions. The core wave functions are appreciable only within the immediate vicinity of the ion, where they have the characteristic oscillatory form of atomic wave functions (Figure 11.2a). These oscillations are a manifestation of the high electronic kinetic energy within the core,⁶ which, in combination with the high negative potential energy, produces the total energy of the core levels. Since valence levels have higher total energies than core levels, within the core region, where they experience the same large and negative potential energy as the core electrons, the valence electrons must have even higher kinetic energies. Thus within the core region the valence wave functions must be even more oscillatory than the core wave functions.

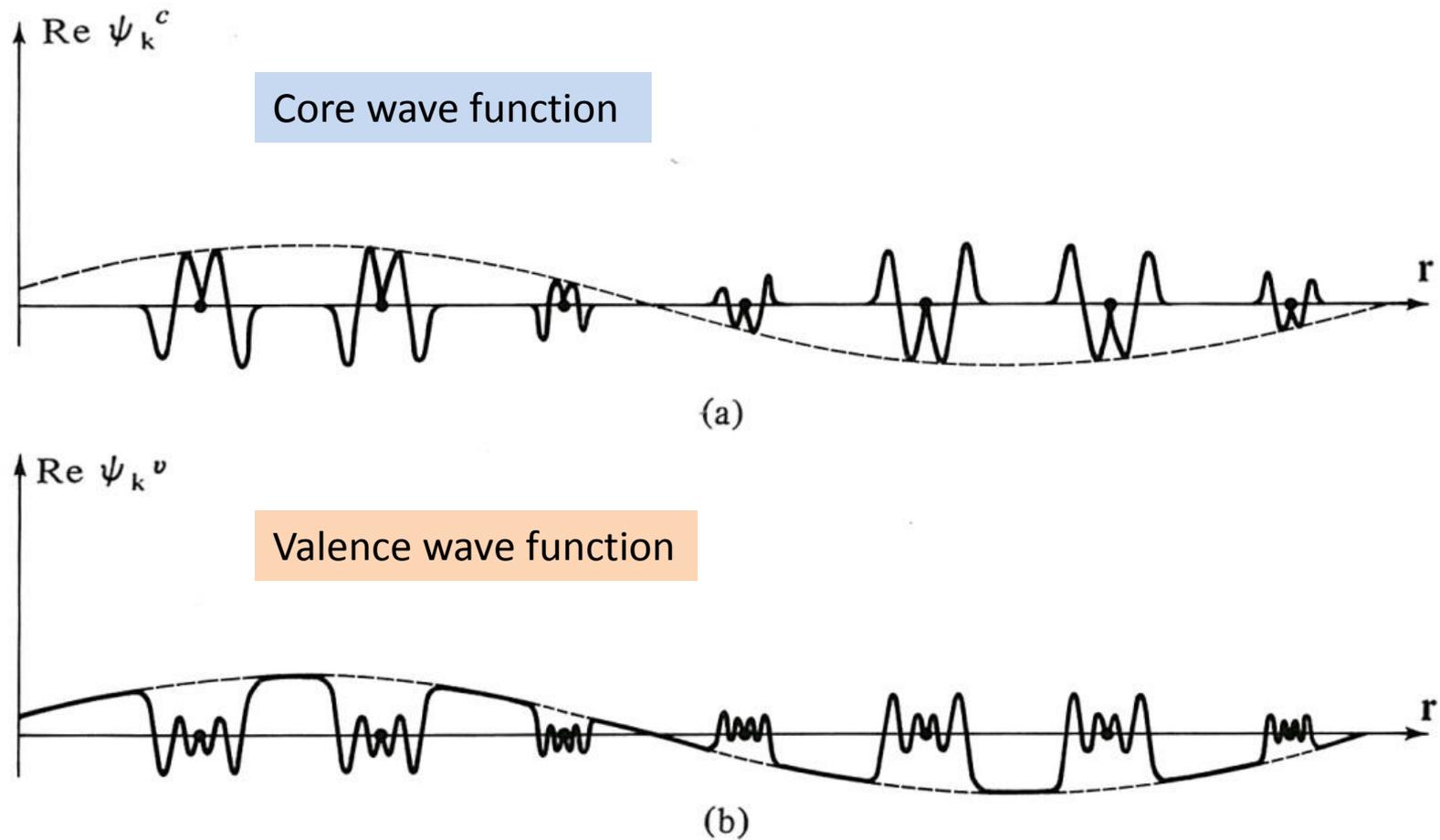


Figure 11.2

(a) Characteristic spatial dependence of a core wave function $\psi_k^c(\mathbf{r})$. The curve shows $\text{Re } \psi$ against position along a line of ions. Note the characteristic atomic oscillations in the vicinity of each ion. The dashed envelope of the atomic parts is sinusoidal, with wavelength $\lambda = 2\pi/k$. Between lattice sites the wave function is negligibly small. (b) Characteristic spatial dependence of a valence wave function $\psi_k^v(\mathbf{r})$. The atomic oscillations are still present in the core region. The wave function need not be at all small between lattice sites, but it is likely to be slowly varying and plane-wavelike there.

This conclusion can also be reached by an apparently different argument:

Eigenstates of the same Hamiltonian with different eigenvalues must be orthogonal. In particular any valence wave function $\psi_{\mathbf{k}}^v(\mathbf{r})$ and any core wave function $\psi_{\mathbf{k}}^c(\mathbf{r})$ must satisfy:

$$0 = \int d\mathbf{r} \psi_{\mathbf{k}}^c(\mathbf{r})^* \psi_{\mathbf{k}}^v(\mathbf{r}). \quad (11.5)$$

Core wave functions are appreciable only within the immediate vicinity of the ion, so the main contribution of this integral must come from the core region. It is enough to consider the contribution to (11.5) from the core region of a single ion, since Bloch's theorem ((8.3)) requires the integrand to be the same from cell to cell. Within this core region $\psi_{\mathbf{k}}^v(\mathbf{r})$ must have oscillations that carefully interlace with those of all the $\psi_{\mathbf{k}}^c(\mathbf{r})$ in order to cause the integrals (11.5) to vanish for all core levels.

Either of these arguments leads to the conclusion that a valence wave function should have the form pictured in Figure 11.2b. If, however, the valence wave functions have an oscillatory structure on the scale of the core region, a Fourier expansion such as (9.1) must contain many short wavelength plane waves, i.e., many terms with large wave vectors. Thus the nearly free electron method, which leads to an approximate wave function composed of a very small number of plane waves, must be untenable.

In one way or another, all of the calculational methods now in use are attempts to come to grips with the necessity for reproducing this detailed, atomic-like structure of the valence wave functions in the core region, while facing the fact that the valence levels are not of the tight-binding type, and therefore have appreciable wave functions in the interstitial regions.

THE CELLULAR METHOD

The first serious attempt to calculate band structure (aside from Bloch's original use of the tight-binding method) was the cellular method of Wigner and Seitz.⁷ The method begins by observing that because of the Bloch relation (8.6):

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k} \cdot \mathbf{R}} \psi_{\mathbf{k}}(\mathbf{r}), \quad (11.6)$$

it is enough to solve the Schrödinger equation (11.1) within a single primitive cell C_0 . The wave function can then be determined via (11.6) in any other primitive cell from its values in C_0 .

However, not every solution to (11.1) within C_0 leads in this way to an acceptable wave function for the entire crystal, since $\psi(\mathbf{r})$ and $\nabla\psi(\mathbf{r})$ must be continuous as \mathbf{r} crosses the primitive cell boundary.⁸ Because of (11.6), this condition can be phrased entirely in terms of the values of ψ within and on the surface of C_0 . It is this boundary condition that introduces the wave vector \mathbf{k} into the cellular solution, and eliminates all solutions except those for a discrete set of energies, which are just the band energies $\varepsilon = \varepsilon_n(\mathbf{k})$.

Boundary conditions within C_0 are

$$\psi(\mathbf{r}) = e^{-i\mathbf{k} \cdot \mathbf{R}} \psi(\mathbf{r} + \mathbf{R}), \quad (11.7)$$

and

$$\hat{\mathbf{n}}(\mathbf{r}) \cdot \nabla \psi(\mathbf{r}) = -e^{-i\mathbf{k} \cdot \mathbf{R}} \hat{\mathbf{n}}(\mathbf{r} + \mathbf{R}) \cdot \nabla \psi(\mathbf{r} + \mathbf{R}), \quad (11.8)$$

where \mathbf{r} and $\mathbf{r} + \mathbf{R}$ are both points on the surface of the cell and $\hat{\mathbf{n}}$ is an outward normal (see Problem 1).

The analytical problem is therefore to solve (11.1) within the primitive cell C_0 subject to these boundary conditions. To preserve the symmetry of the crystal, one takes the primitive cell C_0 to be the Wigner-Seitz primitive cell (Chapter 4) centered on the lattice point $\mathbf{R} = 0$.

The foregoing is an exact restatement of the problem. The first approximation of the cellular method is the replacement of the periodic potential $U(\mathbf{r})$ within the Wigner-Seitz primitive cell by a potential $V(r)$ with spherical symmetry about the origin (see Figure 11.3). One might, for example, choose $V(r)$ to be the potential of a single ion at the origin, ignoring the fact that the neighbors of the origin will also contribute to $U(\mathbf{r})$ within C_0 , especially near its boundaries. This approximation is made entirely for practical reasons, to render a difficult computational problem more manageable.

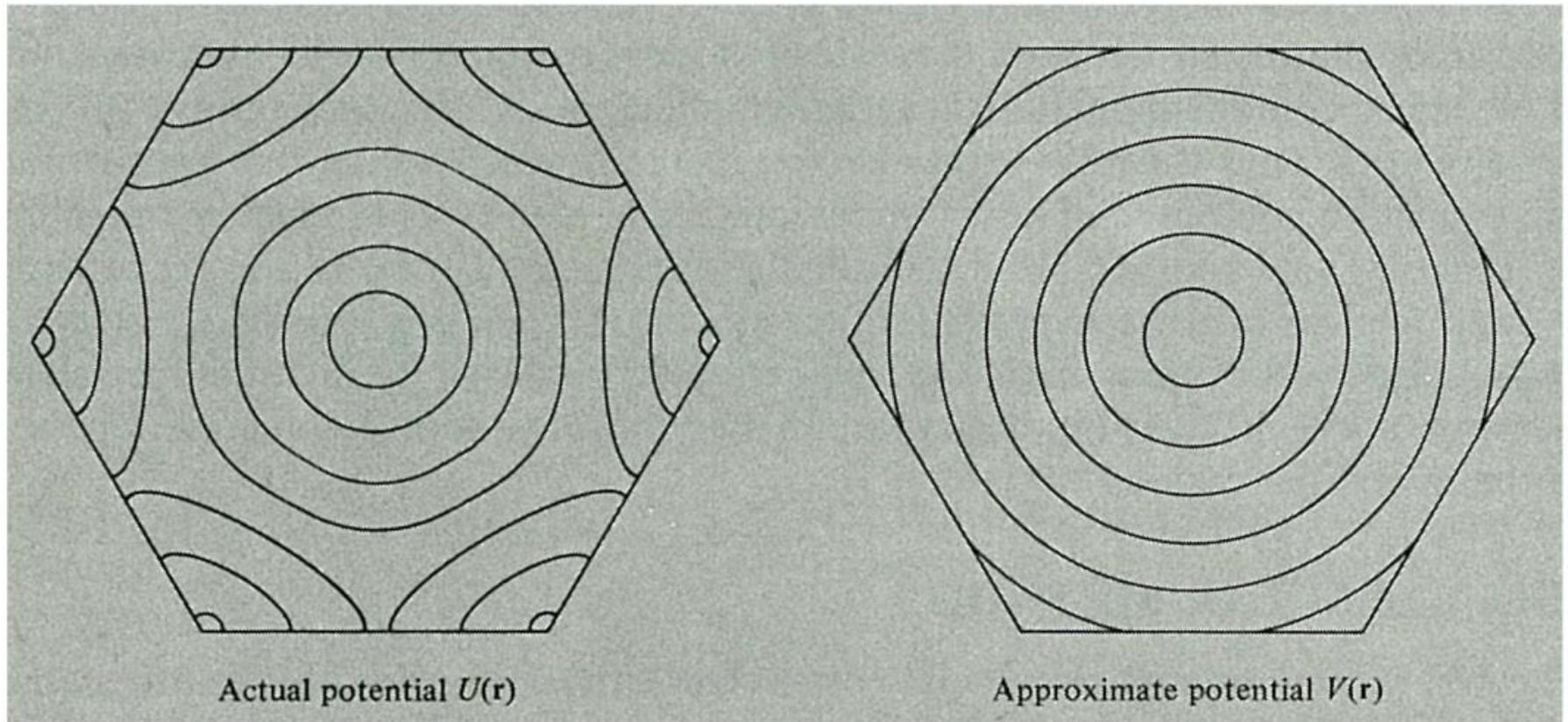


Figure 11.3

Equipotentials (i.e., curves of constant $U(\mathbf{r})$) within a primitive cell. For the actual crystal potential these will have spherical symmetry near the center of the cell where the potential is dominated by the contribution from the central ion. However near the boundary of the cell the potential will deviate substantially from spherical symmetry. The cellular method approximates the potential by a spherically symmetric one everywhere within the cell, with equipotentials as shown on the right.

Once a potential has been chosen spherically symmetric inside C_0 , then within the primitive cell a complete set of solutions to the Schrödinger equation (11.1) can be found of the form⁹

$$\psi_{lm}(\mathbf{r}) = Y_{lm}(\theta, \phi)\chi_l(r), \quad (11.9)$$

where $Y_{lm}(\theta, \phi)$ are spherical harmonics and $\chi_l(r)$ satisfies the ordinary differential equation

$$\chi_l''(r) + \frac{2}{r}\chi_l'(r) + \frac{2m}{\hbar^2}\left(\varepsilon - V(r) - \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right)\chi_l(r) = 0. \quad (11.10)$$

Given the potential $V(r)$ and given *any* value of ε , there is a unique $\chi_{l,\varepsilon}$ that solves (11.10) and is regular at the origin.¹⁰ These $\chi_{l,\varepsilon}$ can be calculated numerically, ordinary differential equations being easy to handle on machines. Since any linear combination of solutions to Schrödinger's equation with the same energy is itself a solution,

$$\psi(\mathbf{r}, \varepsilon) = \sum_{lm} A_{lm} Y_{lm}(\theta, \phi)\chi_{l,\varepsilon}(r) \quad (11.11)$$

will solve (11.1) at energy ε for arbitrary coefficients A_{lm} . However, (11.11) will only yield an acceptable wave function for the crystal if it satisfies the boundary conditions (11.7) and (11.8). It is in the imposition of these boundary conditions that the cellular method makes its next major approximation.

To begin with, one takes only as many terms in the expansion (11.11) as it is computationally convenient to handle.¹¹ Since there is only a finite number of coefficients in the expansion, we can, for a general cell, fit the boundary condition only at a finite set of points on its surface. The imposition of this finite set of boundary conditions (chosen to be as many as there are unknown coefficients) leads to a set of \mathbf{k} -dependent linear homogeneous equations for the A_{lm} and the values of ε for which the determinant of these equations vanishes are the required energies $\varepsilon_n(\mathbf{k})$.

In this way one can search for the eigenvalues $\varepsilon_n(\mathbf{k})$ for each fixed \mathbf{k} . Alternatively, one can fix ε , do a single numerical integration of (11.10), and then search for values of \mathbf{k} for which the determinant vanishes. Provided that one has not been so unfortunate as to choose ε in an energy gap, such values of \mathbf{k} can always be found, and in this way the constant-energy surfaces can be mapped out.

Various ingenious techniques have been used to minimize the mismatch of the wave function at the boundaries due to the fact that the boundary conditions can only be imposed at a finite number of points; such cleverness, and the ability of computers to handle large determinants, have led to cellular calculations of very high accuracy,¹² producing band structures in substantial agreement with some of the other methods we shall describe.

The most famous application of the cellular method is the original calculation by Wigner and Seitz of the lowest energy level in the valence band of sodium metal. Since the bottom of the band is at $\mathbf{k} = \mathbf{0}$, the exponential factor disappears from the boundary conditions (11.7) and (11.8). Wigner and Seitz made the further approximation of replacing the Wigner-Seitz primitive cell by a sphere of radius r_0 with the same volume, thereby achieving a boundary condition with the same spherical symmetry as the potential $V(r)$. They could then consistently demand that the solution $\psi(\mathbf{r})$ itself have spherical symmetry, which requires that only the single term $l = 0, m = 0$ be retained in (11.11). Under these conditions the boundary conditions reduce to

$$\chi_0'(r_0) = 0. \quad (11.12)$$

Thus the solutions to the single equation (11.10) for $l = 0$, subject to the boundary condition (11.12), give the spherically symmetric cellular wave functions and energies.

Note that the problem has the same form as an atomic problem except that the atomic boundary condition—that the wave function vanish at infinity—is replaced by the cellular boundary condition—that the wave function have a vanishing radial derivative at r_0 . The $3s^1$ atomic and cellular wave functions are plotted together in Figure 11.4. Note that the cellular wave function is larger than the atomic one in the interstitial region, but differs from it very little in the core region.

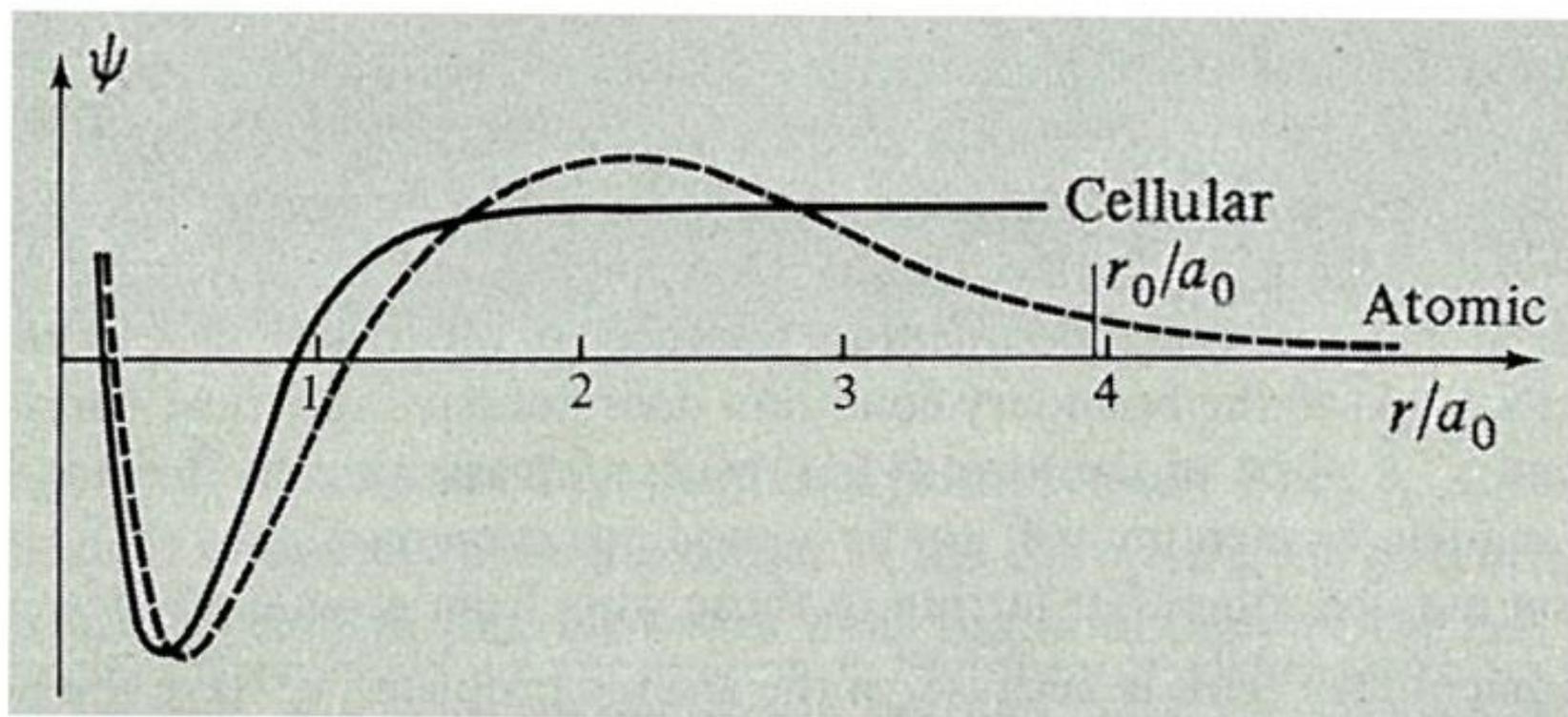


Figure 11.4

Comparison of $3s^1$ cellular (solid curve) and atomic (dashed curve) wave functions for sodium.

There are perhaps two major difficulties with the cellular method:

1. The computational difficulties involved in numerically satisfying a boundary condition over the surface of the Wigner-Seitz primitive cell, a fairly complex polyhedral structure.
2. The physically questionable point of whether a potential representing an isolated ion is the best approximation to the correct potential within the entire Wigner-Seitz primitive cell. In particular, the potential used in the cellular calculations has a discontinuous derivative whenever the boundary between two cells is crossed (Figure 11.5), whereas in actual fact the potential is quite flat in such regions.

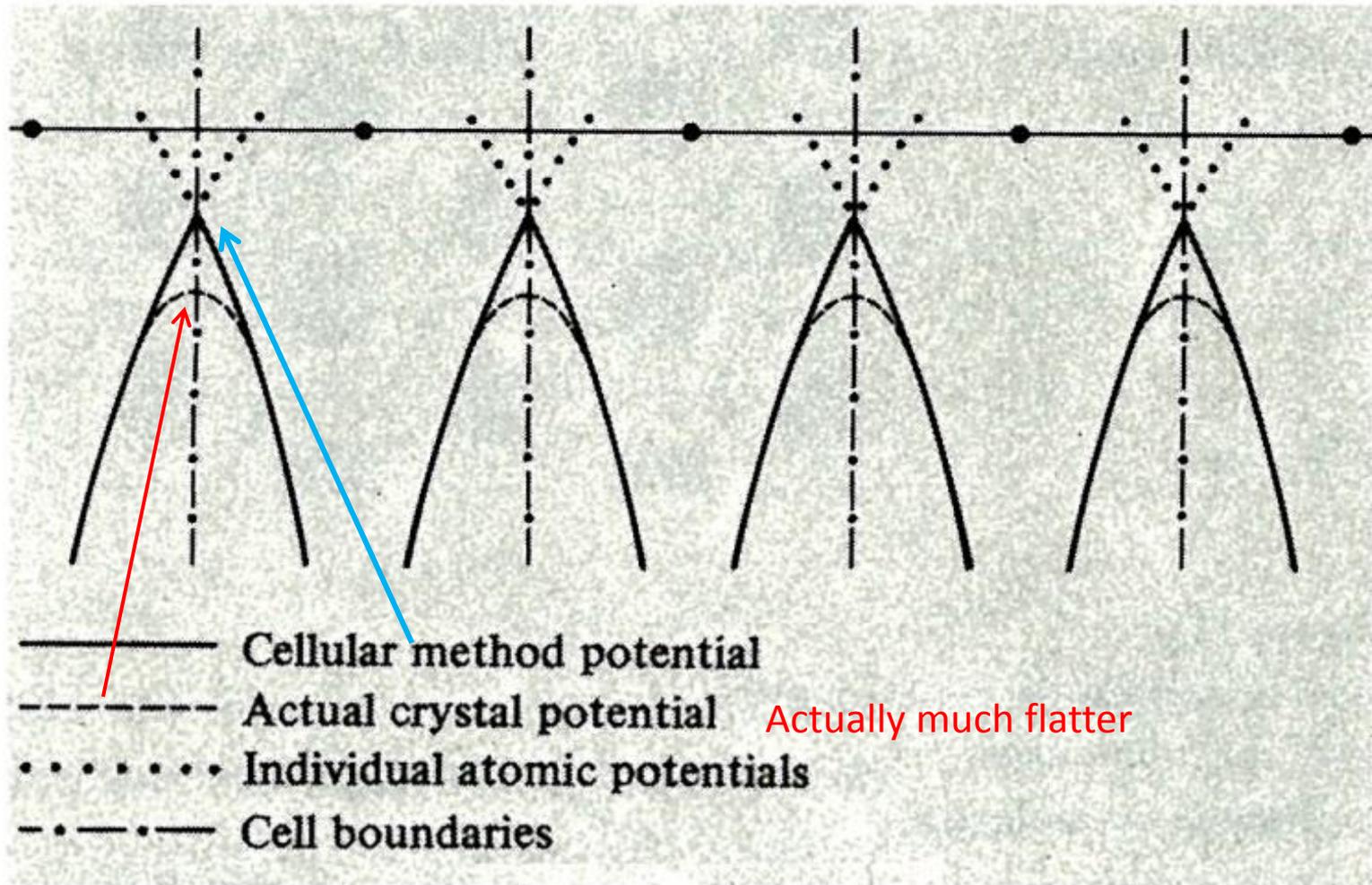


Figure 11.5

The cellular method potential has a discontinuous derivative midway between lattice points, but the actual potential is quite flat there.

A potential that overcomes both objections is the muffin-tin potential, which is taken to represent an isolated ion within a sphere of specified radius r_0 about each lattice point, and taken to be zero (i.e., constant) elsewhere (with r_0 chosen small enough that the spheres do not overlap). (See Figure 11.6.) The muffin-tin potential mitigates both problems, being flat in the interstitial regions, and leading to matching conditions on a spherical rather than a polyhedral surface.

Formally, the muffin-tin potential can be defined (for all \mathbf{R}) by:

$$\begin{aligned} U(\mathbf{r}) &= V(|\mathbf{r} - \mathbf{R}|), & \text{when } |\mathbf{r} - \mathbf{R}| < r_0 & \text{ (the core or atomic region),} \\ &= V(r_0) = 0, & \text{when } |\mathbf{r} - \mathbf{R}| > r_0 & \text{ (the interstitial region),} \end{aligned} \quad (11.13)$$

where r_0 is less than half the nearest-neighbor distance.¹³

If we agree that the function $V(r)$ is zero when its argument exceeds r_0 , then we can write $U(\mathbf{r})$ very simply as

$$U(\mathbf{r}) = \sum_{\mathbf{R}} V(|\mathbf{r} - \mathbf{R}|). \quad (11.14)$$

Two methods are in wide use for computing the bands in a muffin-tin potential: the augmented plane-wave (APW) method and the method of Korringa, Kohn, and Rostoker (KKR).

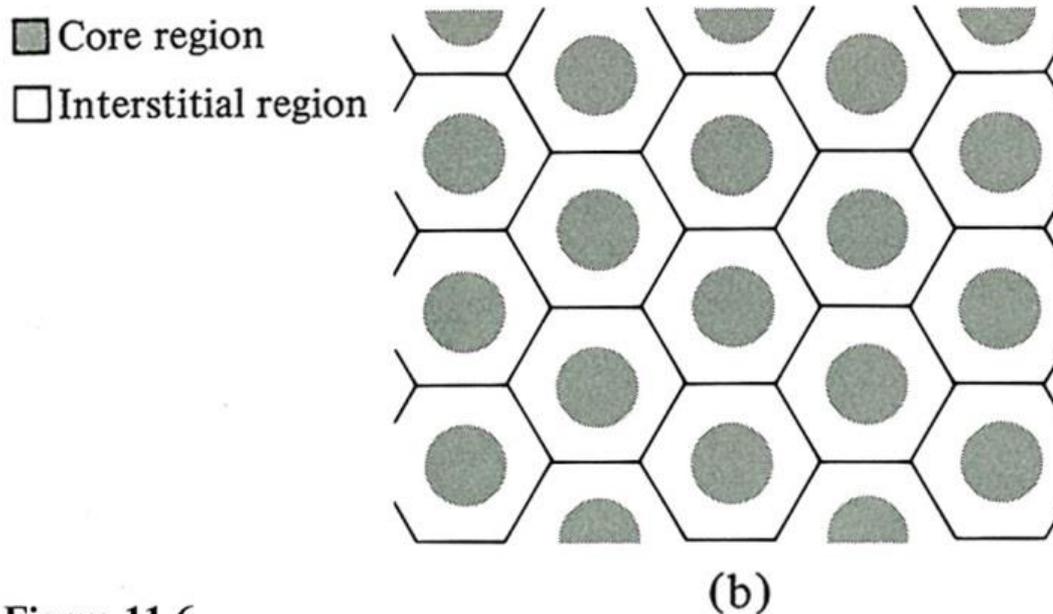
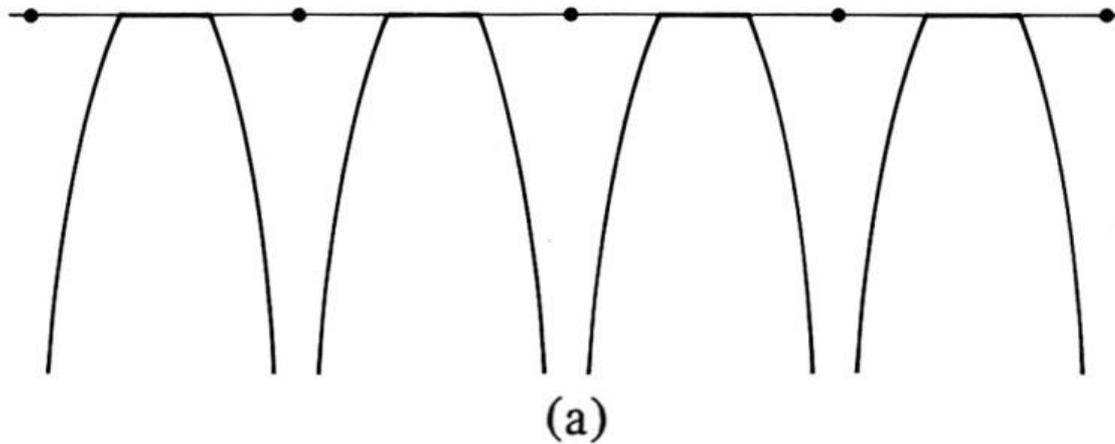


Figure 11.6

(a) The muffin-tin potential, plotted along a line of ions. (b) The muffin-tin potential is constant (zero) in the interstitial regions and represents an isolated ion in each core region.