ARMY RESEARCH LABORATORY

Computing the Effects of Strain on Electronic States: A Survey of Methods and Issues

by James J. Ramsey and Peter W. Chung

ARL-TR-6287 December 2012

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5067

ARL-TR-6287 **December 2012**

Computing the Effects of Strain on Electronic States: A Survey of Methods and Issues

James J. Ramsey and Peter W. Chung
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
December 2012	Final	1 September 2011–30 September 2012

4. TITLE AND SUBTITLE
Computing the Effects of Strain on Electronic States: A Survey of Methods and Issues

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
James J. Ramsey and Peter W. Chung

5d. PROJECT NUMBER
61102A H44

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: RDRL-CIH-C
Aberdeen Proving Ground, MD 21005-5067

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-6287

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We present a cursory survey of existing computational methods and known challenges in estimating accurate electronic band structures of materials suited for electronic device concepts. Both *ab initio* and empirical approaches are described, but greater emphasis is on atomistic and continuum-based empirical methods. Empirical pseudopotential, Slater-Koster tight-binding, k·p, envelope function, and effective mass methods are discussed and compared. Remarkable variability is found in the approaches for incorporating the effects of strain. Despite these differences, however, all methods are suited for handling infinitesimal strain. The results within this context show good agreement among the methods.

15. SUBJECT TERMS
computational methods, electronic structure, strain, deformation, semiconductor

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			James J. Ramsey
Unclassified	Unclassified	Unclassified	UU	116	19b. TELEPHONE NUMBER (*Include area code*) 410-278-6027

Contents

List of Figures

List of Tables

Acknowledgments

This work was supported in part by the U.S. Army Research Laboratory (ARL) through Program Element 61102A H44, the ARL Department of Defense Supercomputing Resource Center, and the Computational Sciences Division of the Computational and Information Sciences Directorate. A helpful technical review of the manuscript by Dr. Sasha Batyrev (ARL) and the use of both the NRL-TB code (from Dr. Michael Mehl, U.S. Naval Research Laboratory (NRL) and the NEMO5 code (from Drs. Gerhard Klimeck and James Fonseca, Purdue University) to help in our understanding of the Slater-Koster tight binding method are also gratefully acknowledged.

1. Introduction

Since the 1950s, strain has been known to control the electronic band structure of semiconductors (*1*). This mechanism was used as a parameter to design novel semiconductor material technologies recently in complementary metal-oxide-semiconductor (CMOS) technologies to control the mobility of electrons in short field-effect transistor channels. In doing so, it enabled Intel Corporation to move from 130- to 90-nm feature sizes in 2003 (*2*), which effectively enabled the continued realization of Moore's Law and still does so today.

Modeling, particularly with regard to the advent of strained silicon, played a significant role in the development of early science and breakthrough technologies. Indeed, the basic mechanisms of strained silicon, and their theoretical underpinnings, were known as early as the 1950s in many of the historic papers that led the way for CMOS and p-type metal-oxide-semiconductor (PMOS) transistors (*1*). Computations thus had a theoretical and mathematical framework from which to draw models and computable concepts.

Today, the majority of relevant scientific papers fall into one of two basic problem groups. Roughly speaking, the first focuses on the determination of the electronic band structure and the second focuses on the time-varying nature of electron transport. Computational research can likewise fit within this loose classification. On their own, each represents large but separate areas of scientific endeavor. But together, they provide a basis for robust design of semiconductor device materials.

With greater access now sought for more chemical species in the periodic table of elements to be considered for doping, substrate design, or even the main semiconductor material, the many computational approaches that provide an early testing platform for materials designs demand a reexamination for their transferability. Of particular connection to the Army, strain is a potential issue in many optoelectronic devices that are composed of multiple materials, such as semiconductor lasers or light-emitting diodes (LEDs). Such materials and devices are not uncommon in infrared, ultraviolet, and microwave detectors and sources either sought by or under consideration as new concepts by the U.S. Army.

As a specific motivation, consider the optical gain medium in quantum dot lasers composed of an array of indium arsenide (InAs) inclusions embedded in a gallium arsenide (GaAs) matrix that have relatively recently been commercialized (*3*–*5*). The mismatch between the lattice constants of these two materials is substantial: 7% (*6*). Strain may affect the states of optically active electrons through piezoelectric effects, especially for nitride semiconductor optoelectronics (*7*, *8*). However, even where such effects are absent or negligible, strain still has important effects on the electronic properties of a device, such as the minimum energy needed to promote an electron from the ground state to an excited state (*9*, *10*).

In this report, we perform a brief survey of various existing methods used to estimate the energies and quantum-mechanical wave functions of systems of electrons, how these methods take strain into account, and where they intersect with concepts from solid mechanics. As such, we consider only the first part of the two-part problem—the accurate computational determination of electronic band structure in strained materials and structures. The survey will predominantly cover approximate empirical approaches. However, in order to make clear what impact the approximations in these approaches have had, the survey will include a brief discussion of the computationally expensive quantum-mechanical approaches known as *ab initio* methods, which will be contrasted with the empirical electronic structure methods. The multiscale aspects of these empirical methods will be highlighted.

The intended audience is the community of researchers with interest in accurately computing the electronic structure properties and, in particular, the excited state properties of materials. However, considering the intersection with concepts of strain, the intended community would be more familiar with concepts of elasticity and deformability of solids. Thus, section 2 starts with a brief review of electronic structure theory, providing the mathematical starting point to which later assumptions are ascribed for the sake of computational feasibility. It then briefly describes *ab initio* methods and tersely lays the foundations for first-principles approaches, including discussion of spin-orbit coupling. Section 3 presents an overview of empirical atomistic approaches that covers the empirical pseudopotential and the Slater-Koster tight-binding methods. Section 4 describes the so-called continuum methods, specifically the interrelated $k \cdot p$ envelope function, and effective mass methods. A potential issue for computing electronic structures of heterogeneous domains comprised of multiple materials—the valence band offset—is described in section 5 within the context of the empirical approaches covered in this report. In section 6, we show computed comparisons of the different methods using models of GaAs, InAs, and aluminum arsenide (AlAs). Section 7 completes the report with conclusions.

2. Background: Electronic Structure Theory

In principle, the electronic structure of a system of n electrons and M atomic nuclei, whether it be a bulk crystal, molecule, nanostructure, etc., is the solution of a many-body quantum mechanical differential equation. For a nonrelativistic system, this equation is the (time-independent) many-electron Schrödinger equation (*11*):

$$\widehat{H}_{\text{many e}^-}\Psi_k(\{\mathbf{r}\}_n) = E^n_{\text{Tot},k}\Psi_k(\{\mathbf{r}\}_n)\,, \tag{1}$$

where $\{\mathbf{r}\}_n$ is shorthand for all of the electron coordinates $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$ and their spins $s_1, s_2, \ldots, s_n$. $\Psi_k(\{\mathbf{r}\}_n)$ is the k^{th} many-electron wavefunction, and for this wavefunction, $E^n_{\text{Tot},k}$ is the total energy of the n electrons in the system. $\widehat{H}_{\text{many e}^-}$ is a differential operator called the many-electron Hamiltonian. Here, the notation is such that $E^n_{\text{Tot},0} \leq E^n_{\text{Tot},1} \leq E^n_{\text{Tot},2}$, etc., where

$E^n_{\mathrm{Tot},0}$ is the lowest possible energy of the electrons in the system, that is, the ground-state energy, and $\Psi_0(\{\mathbf{r}\}_n)$ denotes a ground-state many-electron wavefunction. (If, for example, $E^n_{\mathrm{Tot},0} = E^n_{\mathrm{Tot},1}$, then $\Psi_1(\{\mathbf{r}\}_n)$ is also a ground-state wavefunction.) Wavefunctions with energies higher than $E^n_{\mathrm{Tot},0}$ correspond to excited states of the system. The physical meaning of Ψ_k is such that

$$P = \sum_{s_1,s_2,\ldots,s_n} \int_{\Omega_1} \int_{\Omega_2} \ldots \int_{\Omega_n} |\Psi_k(\{\mathbf{r}\}_n)|^2 \mathrm{d}^3\mathbf{r}_1 \mathrm{d}^3\mathbf{r}_2 \ldots \mathrm{d}^3\mathbf{r}_n \tag{2}$$

is the probability that electron 1 is in the region Ω_1, electron 2 is in the region Ω_2, etc., provided that Ψ_i is normalized so that $P = 1$ if $\Omega_1, \Omega_2, \ldots, \Omega_n \to \infty$. The summation in equation 2 is over all possible spin values of the electrons. The many-electron wavefunction must satisfy the Pauli exclusion principle, which states that two electrons cannot share the same state. The principle is satisfied if, and only if, this wavefunction changes sign when two electronic coordinates are exchanged, that is,

$$\begin{aligned}\Psi_{\mathrm{i}}(\mathrm{r}_1, \mathrm{s}_1, \mathrm{r}_2, \mathrm{s}_2, \ldots, \mathrm{r}_{\mathrm{i}}, \mathrm{s}_{\mathrm{i}}, \mathrm{r}_{\mathrm{i+1}}, \mathrm{s}_{\mathrm{i+1}} \ldots, \mathrm{r}_{\mathrm{n}}, \mathrm{s}_{\mathrm{n}}) \\ = -\Psi_{\mathrm{i}}(\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_{\mathrm{i+1}}, \mathrm{s}_{\mathrm{i+1}}, \mathrm{r}_{\mathrm{i}}, \mathrm{s}_{\mathrm{i}}, \ldots, \mathrm{r}_{\mathrm{n}}, \mathrm{s}_{\mathrm{n}}) \,.\end{aligned} \tag{3}$$

If two electrons were to share the same state, i.e., if $(\mathbf{r}_i, s_i) = (\mathbf{r}_{i+1}, s_{i+1})$, then Ψ_i would have to equal zero for equation 3 to hold.

If the nuclei are treated as classical point particles, then $\widehat{H}_{\text{many e}^-}$ takes the following form (*11*):

$$\widehat{H}_{\text{many e}^-} = \sum_{i=1}^{n} \left[\frac{\hat{\mathbf{p}}^i \cdot \hat{\mathbf{p}}^i}{2m_e} + V_{\text{ext}}(\mathbf{r}_i, \{\mathbf{R}\}_M)\right] + V_{\text{inter}}(\{\mathbf{r}\}) \,. \tag{4}$$

Here, $\hat{\mathbf{p}}^i = -\mathrm{i}\hbar\nabla_i$ is the momentum operator, $\mathrm{i} = \sqrt{-1}$, $\hbar = h/2\pi$ is the reduced Planck constant, ∇_i is the del operator defined with respect to electron coordinate $\mathbf{r}_i$, m_e is the mass of an electron, and

$$V_{\text{ext}}(\mathbf{r}, \{\mathbf{R}\}_M) = -\frac{1}{4\pi\epsilon_0} \sum_{i=1}^{M} \frac{Z_i q^2}{|\mathbf{r} - \mathbf{R}_i|} \,. \tag{5}$$

$$V_{\text{inter}}(\{\mathbf{r}\}_n) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{q^2}{|\mathbf{r}_i - \mathbf{r}_j|} \,, \tag{6}$$

where q is the magnitude of the charge of an electron, ϵ_0 is the permittivity of free space, and $\{\mathbf{R}\}_M$ is shorthand for the coordinates of all M nuclei of the system $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M$ and their respective atomic numbers $Z_1, Z_2, \ldots, Z_M$. $\hat{\mathbf{p}}^i \cdot \hat{\mathbf{p}}^i / 2m_e$ is the kinetic energy operator for electron i. The sum of $\sum_{i=1}^{n} V_{\text{ext}}(\mathbf{r}_i, \{\mathbf{R}\}_M)$ and $V_{\text{inter}}(\{\mathbf{r}\}_n)$ is the total electronic potential energy. $V_{\text{ext}}(\mathbf{r}, \{\mathbf{R}\}_M)$ is called an external potential and accounts for the attractive electrostatic interactions between each electron and the nuclei, while $V_{\text{inter}}(\{\mathbf{r}\}_n)$ is an interaction potential accounting for the repulsions between the electrons.

The presence of $V_{\text{inter}}(\{\mathbf{r}\}_n)$ prevents the n-electron Schrödinger equation from being separated into n individual equations, which makes an exact solution of the many-electron Schrödinger equation intractable for $n > 1$. Instead, the so-called independent electron approximation is adopted, where each electron is nominally treated as if it were independent, but the actual interaction of each electron with all other electrons in the system is taken into account through an effective external potential (*12*). This leads to a one-electron Schrödinger equation, which is

$$\hat{H}_{1\text{e}^-}\psi_i(\boldsymbol{r}, s) = \left[\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\text{ext,eff}}(\mathbf{r}, s, \{\mathbf{R}\}_M)\right]\psi_i(\mathbf{r}, s) = E_i\psi_i(\mathbf{r}, s) \tag{7}$$

for an isolated system, or

$$\hat{H}_{1\text{e}^-}\psi_{i\mathbf{k}}(\boldsymbol{r}, s) = \left[\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\text{ext,eff}}(\mathbf{r}, s, \{\mathbf{R}\}_M)\right]\psi_{i\mathbf{k}}(\mathbf{r}, s) = E_{i\mathbf{k}}\psi_{i\mathbf{k}}(\mathbf{r}, s) \tag{8}$$

for a system that is periodic along one or more dimensions. Here, $\hat{H}_{1\text{e}^-}$ is the one-electron Hamiltonian, $\mathbf{r}$ and s are the position and spin of the electron, $\hat{\mathbf{p}} = -\text{i}\hbar\nabla$ is the momentum operator, $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}/2m_e$ is the kinetic energy operator, $\hat{V}_{\text{ext,eff}}(\mathbf{r}, s, \{\mathbf{R}\})$ is an effective external potential operator, and $\psi_i(\mathbf{r}, s)$, E_i [or $\psi_{i\mathbf{k}}(\mathbf{r}, s)$, and $E_{i\mathbf{k}}$] are the one-electron wavefunction and corresponding eigenenergy of electron i (or $i\mathbf{k}$). The quantity $\mathbf{k}$ is called a Bloch wavevector, which will be defined and discussed later. The effective external potential may be decomposed as

$$\hat{V}_{\text{ext,eff}}(\mathbf{r}, s, \{\mathbf{R}\}_M) = V_{\text{ext}}(\mathbf{r}, \{\mathbf{R}\}_M) + \hat{V}_{\text{e}-\text{e}}(\mathbf{r}, s)\,. \tag{9}$$

Here, $\hat{V}_{\text{e}-\text{e}}(\mathbf{r}, s)$ is an operator that takes into account the interaction between an electron at $\mathbf{r}$ and all the other electrons in the system. The relationship between the one-electron wavefunction $\psi_i(\mathbf{r}, s)$, or $\psi_{i\mathbf{k}}(\mathbf{r}, s)$, and the full many-body wavefunction $\Psi_l(\{\mathbf{r}\}_n)$ depends on $\hat{V}_{\text{e}-\text{e}}(\mathbf{r}, s)$, so approximations to $\hat{V}_{\text{e}-\text{e}}(\mathbf{r}, s)$ lead to certain approximations to this relationship, as seen in section 2.1 on *ab initio* methods. The eigenenergies E_i, or $E_{i\mathbf{k}}$, can be viewed as a set of energy levels (*11*). Examples of this are pictured in figure 1, which shows a schematic of these levels for an isolated atom with a nuclear charge of n in four different configurations: (1) ground state of the atom when it is neutral, i.e., the number of electrons equals the nuclear charge n; (2) ground state of the atom with an extra negative charge; (3) ground state of the atom with an extra positive charge; and (4) a neutral excited state of the atom. The total electronic energies of each state are shown. Because of the Pauli exclusion principle, only two electrons (one spin-up and one spin-down) can occupy each energy level. If n is even, all the levels up to $E_{\text{ceil}(n/2)}$ are occupied in the ground state of a neutral atom. Otherwise, one of the energy levels is partially occupied, even in the ground state. [The ceil( )function returns the smallest integer greater than or equal to its argument.] Here, energy levels above the highest ground-state energy level, $E_{\text{ceil}(n/2)}$, are energies required to add an electron to the system at a certain level, so $E_{\text{ceil}(n/2)+1} = E_{\text{Tot},0}^{n+1} - E_{\text{Tot},0}^{n}$, $E_{\text{ceil}(n/2)+2} = E_{\text{Tot},1}^{n+1} - E_{\text{Tot},0}^{n}$, etc. Energies equal to or below the highest ground-state energy level are energies required to remove an electron from the system, so $E_{\text{ceil}(n/2)} = E_{\text{Tot},0}^{n} - E_{\text{Tot},0}^{n-1}$, $E_{\text{ceil}(n/2)-1} = E_{\text{Tot},0}^{n} - E_{\text{Tot},1}^{n-1}$, etc. (*13*). To a first approximation, the available

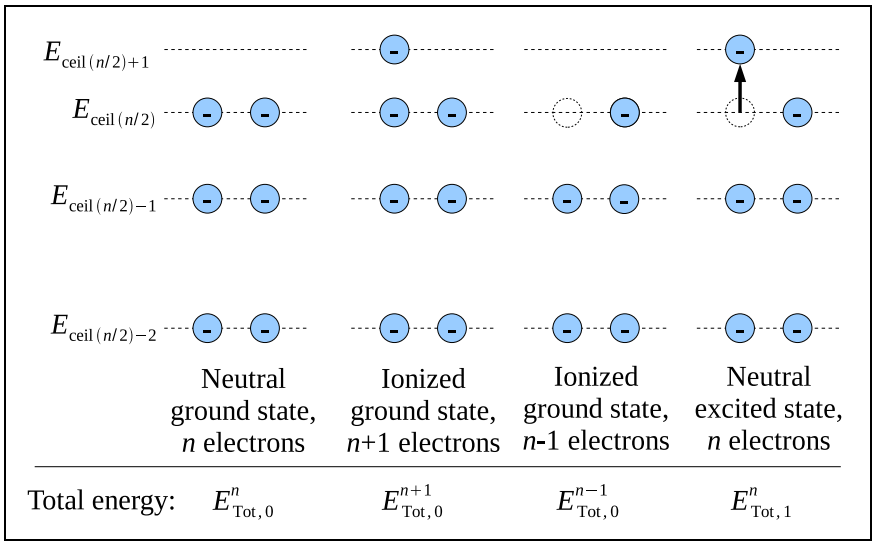

Figure 1. Energy levels of an isolated atom with a nuclear charge of n in four different configurations: (1) ground state of the atom when it is neutral, i.e., the number of electrons equals the nuclear charge n, (2) ground state of the atom with an extra negative charge, (3) ground state of the atom with an extra positive charge, and (4) a neutral excited state of the atom, where an electron has been excited from level ceil(n/2) to ceil(n/2)+1. The total electronic energies of each state are shown. The ceil() function, which returns the smallest integer greater than or equal to its argument, accounts for the case where n is odd. In the diagram shown, however, n is even.

neutral excited states may be estimated from these energy levels that—strictly speaking—pertain to addition and removal of electrons. For example, a neutral excited state may be treated as if it were due to an addition of an electron at a higher energy level followed by the removal of an electron at a lower level. Such a state is shown in figure 1, where an electron is excited to the next higher energy level and increased in energy by $E_{\mathrm{ceil}(n/2)+1} - E_{\mathrm{ceil}(n/2)}$. An electron in such an excited state will emit a photon with that amount of energy when it returns to the ground state, a process that is important in devices such as lasers and LEDs (*14*). Naturally, however, this approximation is best suited for the addition or removal of electrons in the system, as opposed to elementary excitations such as bound states or electron-hole pairs (*13*, *15*, *16*). Additional energy levels may be introduced by the attraction between an excited electron and the vacancy or "hole" it leaves behind in its former energy level. These new levels cannot be predicted from the addition and removal energies of the independent-electron approximation. For simplicity, the energy levels in the figure have been presented as if they were distinct, but this is often not the case. Rather, different one-electron wavefunctions, e.g., ψ_l and ψ_{l+1}, can share the same energy eigenvalue; that is, $E_l = E_{l+1}$. Such energy levels are called *degenerate*.

For periodic systems, the one-electron eigenenergy $E_{i\mathbf{k}}$ depends on the quantity $\mathbf{k}$ in equation 8, which comes from the Bloch theorem. This theorem states that if the potential $\hat{V}_{\mathrm{ext,eff}}$ is periodic, the one-electron wavefunction has the form (*12*)

$$\psi_{i\mathbf{k}}(\mathbf{r}, s) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} u_{i\mathbf{k}}(\mathbf{r}, s) \,, \tag{10}$$

where $\mathbf{k}$ is called a Bloch wavevector, and $u_{i\mathbf{k}}(\mathbf{r}, s)$ has the same periodicity as $\hat{V}_{\text{ext,eff}}$. The one-electron wavefunction, then, has the form of a traveling wave propagating in the direction of $\mathbf{k}$. The periodicity of the potential can be expressed in terms of a lattice vector $\overline{\mathbf{R}} = \sum_{i=1}^{N_p} m_i \mathbf{a}_i$, where m_i and N_p are integers and $\mathbf{a}_i$ is a primitive lattice vector. Accordingly,

$$\hat{V}_{\text{ext,eff}}(\mathbf{r}, s, \{\mathbf{R}\}_M) = \hat{V}_{\text{ext,eff}}(\mathbf{r} + \overline{\mathbf{R}}, s, \{\mathbf{R}\}_M) \,,$$

and (11)

$$u_{i\mathbf{k}}(\mathbf{r}, s) = u_{i\mathbf{k}}(\mathbf{r} + \overline{\mathbf{R}}, s) \,.$$

If the system is periodic in all three directions, then $N_p = 3$ and the three lattice vectors form a unit cell, as shown in figure 2. The Bloch wave vector can then point along any direction in space. If the periodicity is confined to a plane, then $N_p = 2$ and the wave vector $\mathbf{k}$ is also confined within the plane defined by the two lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. If there is only periodicity along one dimension, then $N_p = 1$ and $\mathbf{k}$ is parallel or antiparallel to $\mathbf{a}_1$ (*10*). For any value of N_p, a reciprocal space can be defined such that its primitive vectors satisfy the relationship $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$, where δ_{ij} is the Kronecker delta. A general vector in reciprocal space is $\mathbf{q} = \sum_{i=1}^{3} c_i \mathbf{b}_i$, where c_i may be any real number; a general reciprocal lattice vector, then, is $\mathbf{G} = \sum_{i=1}^{3} n_i \mathbf{b}_i$, where n_i is an integer. In general, $e^{\mathrm{i}\overline{\mathbf{R}}\cdot\mathbf{G}} = 1$. The Bloch wave vector $\mathbf{k}$ exists within the subset of reciprocal space called either the first Brillouin zone or simply *the Brillouin zone*, which consists of the points in reciprocal space closer to $\mathbf{G} = 0$ than to any other reciprocal lattice point (*12*).

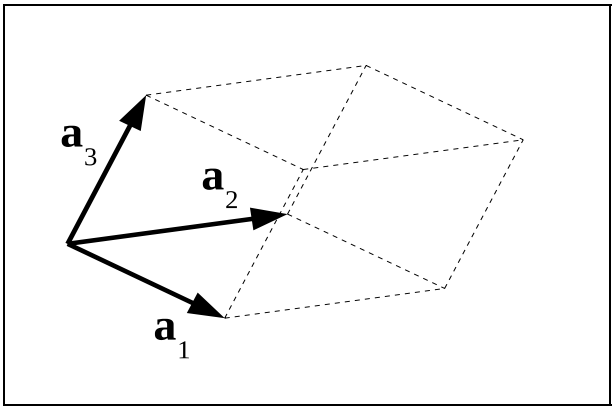

Figure 2. The three primitive lattice vectors of 3-D crystalline unit cell, which is shown in dashed lines as a parallelepiped. The origin of the primitive lattice vectors is shown as located at a corner of a unit cell of a crystal, but the origin is arbitrary and can be taken to be, for example, the center of a unit cell.

Figure 3d shows the Brillouin zone of a bulk crystal with a zincblende or diamond-type structure. These two structures are common in semiconductors such as GaAs and silicon (Si). The conventional cubic unit cells of such crystal structures are shown in figures 3a and b, while the primitive unit cells (i.e., the smallest possible unit cells needed to specify the crystal structure) are shown in figure 3c. Certain high-symmetry points in the Brillouin zone are given special labels. For example, the center of the zone, where $\mathbf{k} = 0$, is denoted σ. The labels for other high-symmetry points are shown in figure 3d. Typically, $E_{i\mathbf{k}}$ is plotted for values of $\mathbf{k}$ that trace a path connecting several of these high-symmetry points in the Brillouin zone (*17*), and such a plot is shown in figure 4, where the path traced is from L to σ to X to K and back to σ. These plots are diagrams of the *band structure* of the crystal.

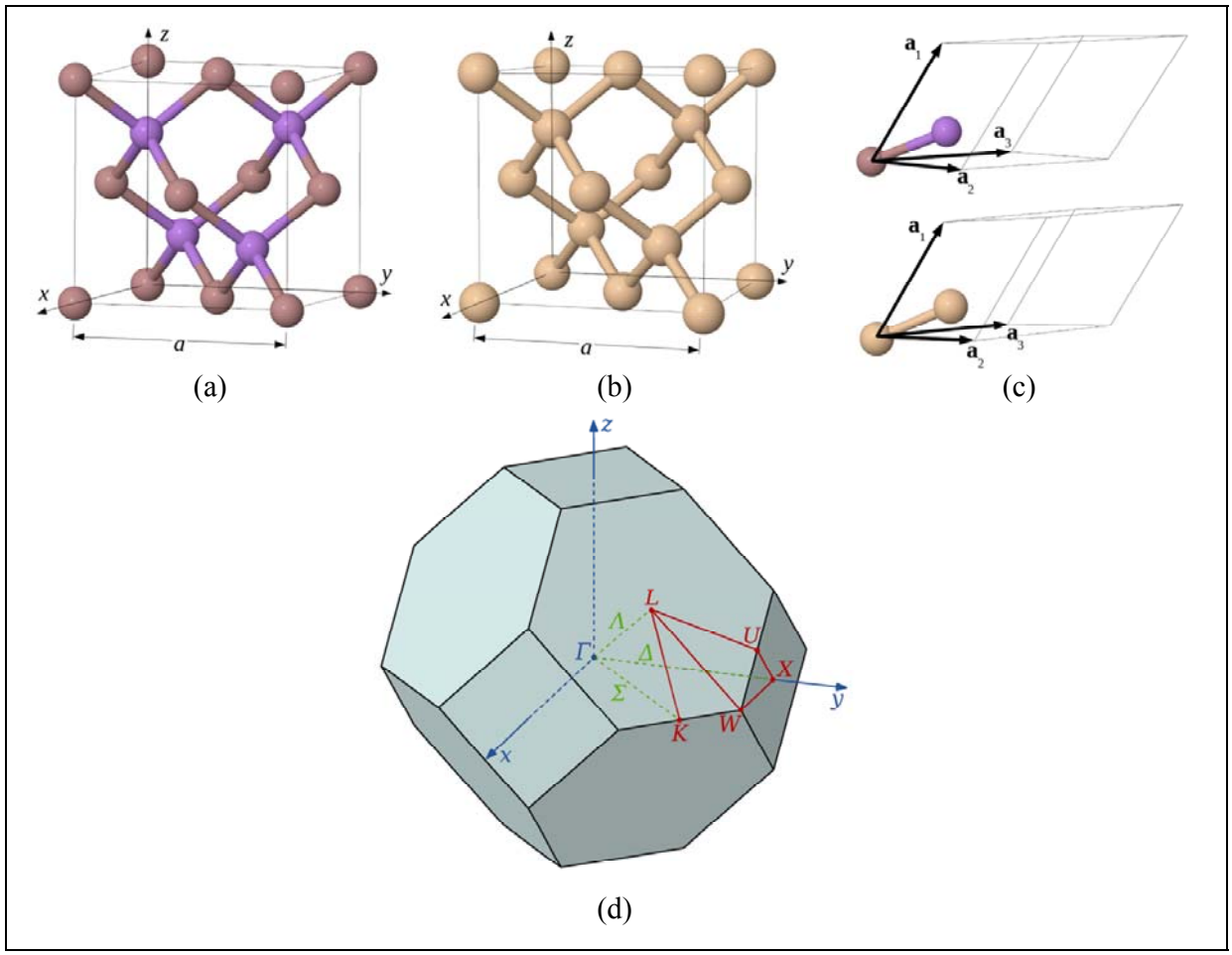

Figure 3. (a) Conventional unit cell of a material with a zincblende crystal structure and lattice constant a. (b) Conventional unit cell of a material with a diamond crystal structure and a lattice constant a. The only difference between the zincblende and diamond structures is that the atoms in the latter are all of the same type. (c) The primitive unit cells for zincblende- and diamond-type crystals. (d) Brillouin zone corresponding to the primitive cells shown in subfigure (c). The primitive lattice vectors are taken to be $\mathbf{a}_1 = (a/2)(\mathbf{e}_y + \mathbf{e}_z)$, $\mathbf{a}_2 = (a/2)(\mathbf{e}_x + \mathbf{e}_z)$, and $\mathbf{a}_3 = (a/2)(\mathbf{e}_x + \mathbf{e}_y)$, where $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are unit vectors pointing along the x-, y-, and z-directions shown in subfigures (a) and (b) (*12*). Special high-symmetry points in this Brillouin zone are denoted by σ, K, W, X, U, and L. The line segments connecting points σ and K, σ and X, and σ and L are denoted as Σ, τ, and $\ddot{\nu}$, respectively (*17*).

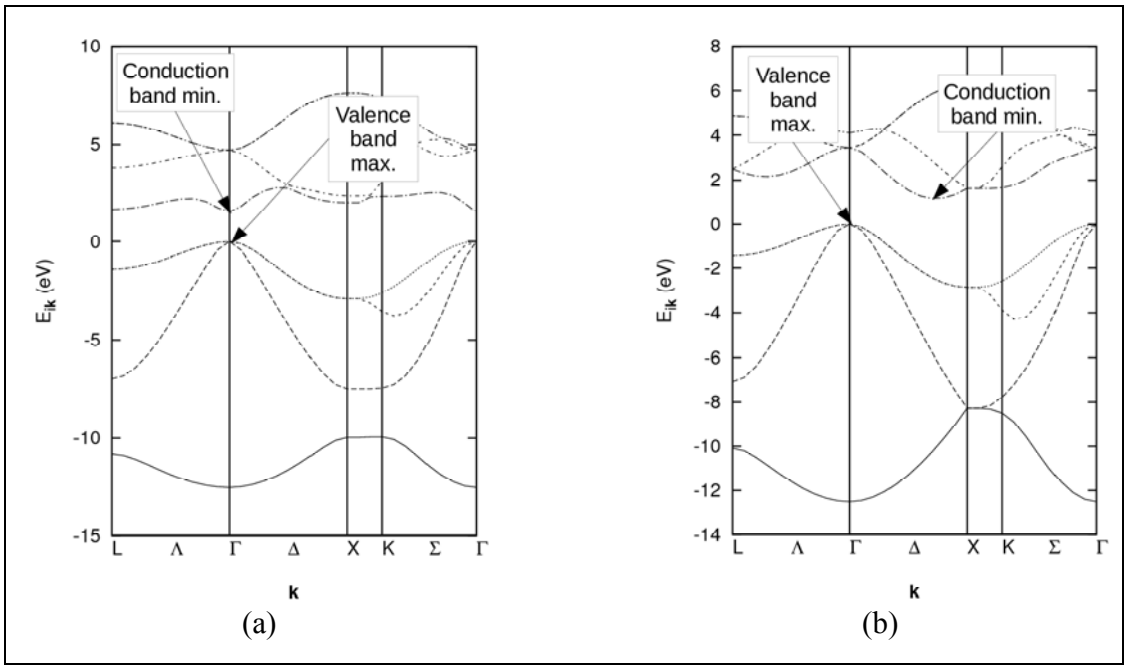

Figure 4. Band structures of (a) GaAs and (b) Si, as estimated by the method of Vogl et al. (*18*). In each diagram, the locations of the valence band maximum and conduction band minimum are shown. The zero-energy datum is taken to be the valence band maximum.

As mentioned before, the energy levels of an isolated n-electron atom in its ground state are occupied only up to a certain maximum level $E_{\mathrm{ceil}(n/2)}$. Similarly, in a crystal in its ground state, only bands with energies below a certain maximum level are occupied. For semiconductors and insulators, this maximum level is the valence band maximum. Figure 4 shows the valence band maxima for Si and GaAs. For an electron in these materials to be excited to the next highest band, called the conduction band, its energy must become at least equal to the conduction band minimum, which is also shown in figure 4. The difference between the valence band maximum and conduction band minimum is called the *band gap energy* and is denoted here as E_g; it governs the frequency of the photon emitted by an excited electron as it returns to the ground state. In the band structure for GaAs shown in figure 4a, the conduction band minimum and valence band maximum occur for the same Bloch wave vector $\mathbf{k} = 0$, so GaAs is called a *direct band gap* material. All that is needed, then, for an electron with such a wave vector to be promoted from the valence to the conduction band is to absorb an photon with energy E_g. The band structure for Si shown in figure 4b shows the conduction band minimum and valence band maximum occurring at different wave vector values. In such a case, to promote an electron from the valence band to the conduction band it is not enough for an electron to simply absorb a photon with an energy equal to E_g. The electron must also have momentum imparted to it by thermal vibrations of the crystal nuclei in order to change its wave vector to that of the conduction band minimum. For this reason, Si is called an *indirect band gap* material (*12*).

Metals, unlike semiconductors and insulators, have no band gap at all, as seen in the band structure of silver (Ag) in figure 5, which is why metals readily conduct. This illustrates the importance of band structure in determining electronic properties.

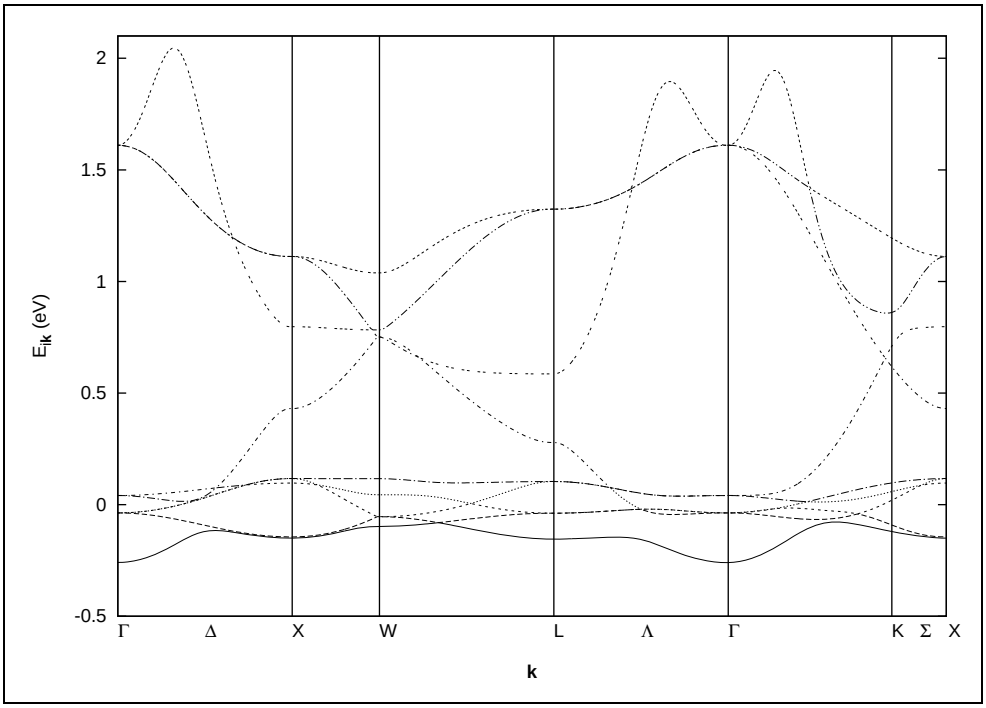

Figure 5. Band structure of Ag, as determined via the Naval Research Laboratory (NRL) tight-binding code (*19*, *20*). Since the lattice vectors of the primitive cells of the diamond and zincblende structures are the same as that of the primitive cell of the crystal structure of Ag (face-centered cubic, or fcc), the special points along the horizontal axis of this diagram are the same as those in figure 3d. The primitive cell of an fcc lattice has only one unit cell.

To summarize, while the electronic structure of a system is in principle the solution of an intractable many-body equation, in practice it can be characterized by a picture where the electrons are nominally treated as independent and each have their own eigenenergies and one-electron wavefunctions. For nonperiodic systems, these eigenenergies are discrete values E_i, but periodicity in one or more directions introduces a wavevector $\mathbf{k}$ on which the eigenenergies, now denoted as $E_{i\mathbf{k}}$, also depend. Examination of these eigenenergies can shed light on various physical properties of the system, as shown for relatively simple examples such as isolated atoms and bulk crystals. How one may obtain an effective one-electron equation from the many-electron equation, and the relationship of the one-electron wavefunctions to the many-body wavefunction, will be discussed in the following section on *ab initio* methods.

2.1 *Ab Initio* Methods

Ab initio methods amount to ways of either approximately solving the many-electron Schrödinger equation or determining quantities of interest, such as $E^n_{\mathrm{Tot},0}$, without directly

solving the many-electron equation itself. Often, these methods involve solving an effective one-electron Schrödinger equation. Some methods are characterized by the approximations involved in formulating the operator $\hat{V}_{\mathrm{e-e}}$ that take into account the interaction between electrons. Other formulations may be exact in principle but still require approximations to obtain solutions in practice. For the sake of simplicity, dependencies on the wave vector **k** or spin are suppressed in the following briefly outlined methods.

One of the earliest *ab initio* approaches, the Hartree approximation, makes calculations tractable through the assumption that the electrons of the system can be treated as if they were smeared out into a cloud with charge density $-q\rho(\mathbf{r})$, where $\rho(\mathbf{r}) = \sum_{i=1}^{n} |\psi_i(\mathbf{r})|^2$. The interaction potential between an electron at **r** and all other electrons is then approximated as an integral over $\rho(\mathbf{r})$ (*11*, *12*).

$$\frac{-q}{4\pi\epsilon_0} \sum_{j=1}^{n} \frac{-q}{|\mathbf{r} - \mathbf{r}_j|} \approx \frac{q^2}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}^3\mathbf{r}' = \hat{V}_{\mathrm{e-e}}(\mathbf{r}) \,. \tag{12}$$

The integration is over all possible values of $\mathbf{r}'$, that is, over all space.

With this approximation, the one-electron Schrödinger equation may be solved iteratively, as illustrated by the pseudocode in figure 6. One begins with an initial guess for $\psi_i(\mathbf{r})$, calculates $\rho(\mathbf{r})$ and then $\hat{V}_{\mathrm{e-e}}$ and $\hat{V}_{\mathrm{ext,eff}}$ from that guess, solves the one-electron Schrödinger equation using the estimated $\hat{V}_{\mathrm{ext,eff}}$, and then checks if the value of $\psi_i(\mathbf{r})$ and the initial guess approximately match, given a certain tolerance. If not, the more recently calculated estimate of $\psi_i(\mathbf{r})$ becomes the starting guess for the next iteration. (This iterative process could be refined to make it more numerically stable by, for example, using a mixture of previous guesses as the guess for the next iteration [*21*].) This is called a *self-consistent* approach.

The pseudocode shown in figure 6 does not specify the means of solving the eigenproblem in each iteration. It is possible, for example, to discretize the problem via a finite-difference scheme (*22*) or by finite elements (*23*). More commonly, though, the eigenproblem is solved through the means shown in figure 7, where the one-electron wavefunction is expanded in a linear combination of basis functions, such as plane waves, atomic orbitals, or Gaussians (*21*), and a matrix equation is obtained by substituting the expansion into the eigenproblem, multiplying through by the complex conjugate of one of the basis functions and integrating. The resulting matrix equation is a generalized eigenvalue problem that may be solved numerically. When $\hat{V}_{\mathrm{ext,eff}}$ depends on the one-electron wavefunctions, the latter's basis function expansions are substituted into it as well.

for $i = 1$ to n **do**
  Make initial guess $\psi_i^0(\mathbf{r})$ for wavefunction $\psi_i(\mathbf{r})$.
end for
Let $m = 0$.
Set tolerance to tol.

repeat
  Increment m by 1.
  Let $\rho^m(\mathbf{r}) = \sum_{i=1}^{n} |\psi_i^{m-1}(\mathbf{r})|^2$.
  Let $\hat{V}_{\mathrm{e-e}}^m(\mathbf{r}) = \dfrac{q^2}{4\pi\varepsilon_0} \int \dfrac{\rho^m(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}'$.
  Let $\hat{V}_{\mathrm{ext,eff}}^m(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + \hat{V}_{\mathrm{e-e}}^m(\mathbf{r})$.
  for $i = 1$ to n **do**
    Solve $\left(\dfrac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\mathrm{ext,eff}}^m(\mathbf{r}) \right) \psi_i^m(\mathbf{r}) = E_i \psi_i^m(\mathbf{r})$.
  end for

until $\mathrm{err}(m, m-1) \leq tol$

where $\mathrm{err}(m, m-1)$ is $|\psi_i^m(\mathbf{r}) - \psi_i^{m-1}(\mathbf{r})|$, $|\rho^m(\mathbf{r}) - \rho^{m-1}(\mathbf{r})|$,
  or $|V^m(\mathbf{r}) - V^{m-1}(\mathbf{r})|$ for all $\mathbf{r}$, i

Assume $\psi_i(\mathbf{r}) \approx \psi_i^m(\mathbf{r})$, $\rho(\mathbf{r}) \approx \rho^m(\mathbf{r})$, etc.

Figure 6. Example pseudocode for iteratively solving the one-electron Schrödinger equation using the Hartree approximation.

Expand $\psi_i^m(\mathbf{r})$ as basis function expansion, i.e. $\psi_i^m(\mathbf{r}) = \sum_j C_j^{im} \chi_j^m(\mathbf{r})$.

Substitute expansion into one-electron Schrödinger equation:

$$\sum_j \left(\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\text{ext,eff}}^m(\mathbf{r}) \right) C_j^{im} \chi_j^m(\mathbf{r}) = E_i \sum_j C_j^{im} \chi_j^m(\mathbf{r})$$

Multiply both sides of equation by $[\chi_{j'}^m(\mathbf{r})]^\dagger$ and integrate, so that:

$$\sum_j H_{j'j}^m C_j^{im} = E_i S_{j'j}^m C_j^{im}$$

where

$$H_{j'j}^m = \int [\chi_{j'}^m(\mathbf{r})]^\dagger \left(\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\text{ext,eff}}^m(\mathbf{r}) \right) \chi_j^m(\mathbf{r}) \mathrm{d}^3\mathbf{r}$$

and

$$S_{j'j}^m = \int [\chi_{j'}^m(\mathbf{r})]^\dagger \chi_j^m(\mathbf{r}) \mathrm{d}^3\mathbf{r}$$

Figure 7. Pseudocode for transforming a general one-electron Schrödinger equation into a matrix equation that can be solved numerically. The superscript "†" indicates the complex conjugate.

The approximation in equation 12 is fairly drastic because it implies that $\Psi_0(\{\mathbf{r}\}_n) = \prod_{i=1}^n \psi_i(\mathbf{r}_i)$, which violates the antisymmetry condition for the many-electron wavefunction in equation 3. The Hartree-Fock method (*24*) corrects this particular problem by assuming that $\Psi_0(\{\mathbf{r}\}_n)$ is a Slater determinant of the one-electron wavefunctions.

$$\Psi_0(\{\mathbf{r}\}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \cdots & \psi_1(\mathbf{r}_n) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \cdots & \psi_2(\mathbf{r}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_n(\mathbf{r}_1) & \psi_n(\mathbf{r}_2) & \cdots & \psi_n(\mathbf{r}_n) \end{vmatrix}. \tag{13}$$

An exchange of two coordinates implies a switching of two columns of the determinant, which causes a change in the sign of $\Psi_0(\{\mathbf{r}\}_n)$. As a consequence of this form of the many-electron wavefunction, $\hat{V}_{\mathrm{e-e}}(\mathbf{r})$ becomes (*12*)

$$\hat{V}_{\mathrm{e-e}}(\mathbf{r})\psi_i(\mathbf{r}) = \left[\frac{q^2}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}' \right] \psi_i(\mathbf{r}) - \frac{q^2}{4\pi\epsilon_0} \sum_{j=1}^n \int \frac{\psi_j^\dagger(\mathbf{r})\psi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}' \, \psi_j(\mathbf{r}) \delta_{s_i s_j}, \tag{14}$$

where the superscript "†" indicates the complex conjugate and $\delta_{s_i s_j}$ equals one if $s_i = s_j$, and zero otherwise where s_i and s_j are the spins of electrons i and j. Pseudocode for the Hartree-Fock approximation is shown in figure 3. One may solve the resulting eigenproblems from this algorithm through the method shown in figure 8. While in the original Hartree approximation,

for $i = 1$ to n **do**
  Make initial guess $\psi_i^0(\mathbf{r})$ for wavefunction $\psi_i(\mathbf{r})$.
end for
Let $m = 0$.
Set tolerance to tol.

repeat
  Increment m by 1.
  Let $\rho^m(\mathbf{r}) = \sum_{i=1}^{n} |\psi_i^{m-1}(\mathbf{r})|^2$.
  Let $\hat{V}_{\mathrm{e-e}}^m(\mathbf{r}) = \frac{q^2}{4\pi\varepsilon_0} \int \frac{\rho^m(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}'$.
  for $i = 1$ to n **do**
    Add $-\frac{q^2}{4\pi\varepsilon_0} \sum_{j=1}^{n} \int \frac{[\psi_j^{m-1}(\mathbf{r})]^\dagger \psi_i^{m-1}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}' \psi_j^{m-1}(\mathbf{r}) \delta_{s_i s_j}$ to $\hat{V}_{\mathrm{e-e}}^m(\mathbf{r}) \psi_i^m(\mathbf{r})$
    Let $\hat{V}_{\mathrm{ext,eff}}^m(\mathbf{r}) \psi_i^m(\mathbf{r}) = [V_{\mathrm{ext}}(\mathbf{r}) + \hat{V}_{\mathrm{e-e}}^m(\mathbf{r})] \psi_i^m(\mathbf{r})$.
    Solve $\left(\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\mathrm{ext,eff}}^m(\mathbf{r}) \right) \psi_i^m(\mathbf{r}) = E_i \psi_i^m(\mathbf{r})$.
  end for

until $\mathrm{err}(m, m-1) \leq tol$

where $\mathrm{err}(m, m-1)$ is $|\psi_i^m(\mathbf{r}) - \psi_i^{m-1}(\mathbf{r})|$, $|\rho^m(\mathbf{r}) - \rho^{m-1}(\mathbf{r})|$,
  or $|V^m(\mathbf{r}) - V^{m-1}(\mathbf{r})|$ for all $\mathbf{r}$, i

Assume $\psi_i(\mathbf{r}) \approx \psi_i^m(\mathbf{r})$, $\rho(\mathbf{r}) \approx \rho^m(\mathbf{r})$, etc.

Figure 8. Example pseudocode for iteratively solving the one-electron Schrödinger equation using the Hartree-Fock approximation.

$\hat{V}_{\mathrm{e-e}}$ is simply a function of $\mathbf{r}$ that is multiplied by $\psi_i(\mathbf{r})$, in the Hartree-Fock approximation, $\hat{V}_{\mathrm{e-e}}$ becomes an integral operator, and one that involves n integrations for each one-electron wavefunction $\psi_i(\mathbf{r})$. Taking the Pauli principle into account thus entails increased computational expense.

The Kohn-Sham equations of density functional theory (DFT) (*25*) are different from the previously mentioned, effective one-electron Schrödinger equations because they are essentially a mathematical tool to find the total energy and electron density at the ground state. In principle, the total energy of the system is a functional of the electron density $E[\rho]$. The value of $\rho(\mathbf{r})$ that minimizes $E[\rho]$ is the ground-state electron density, and the minimum value of $E[\rho]$ is the ground-state total energy $E_{\mathrm{Tot},0}^n$. However, $E[\rho]$ cannot, in general, be expressed as an explicit, analytic functional of the density that can be directly minimized (*21*). To work around this, Kohn and Sham (*26*) posited a fictitious auxiliary system of noninteracting electrons whose ground-state charge density and total energy are identical to that of the real n-electron system. For this auxiliary system, $\rho(\mathbf{r}) = \sum_{i=1}^{n} |\psi_i(\mathbf{r})|^2$, and

$$\hat{V}_{\mathrm{e-e}}(\mathbf{r}) = \frac{q^2}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3\mathbf{r}' + V_{XC}[\rho]\,, \tag{15}$$

where $V_{XC}[\rho]$ is a functional derivative with respect to ρ of the exchange-correlation energy $E_{XC}[\rho]$, which takes into account the part of the electron-electron interactions not taken into account by the first term of $\hat{V}_{\mathrm{e-e}}(\mathbf{r})$. The Kohn-Sham equations, then, are solved iteratively with the self-consistent approach used to solve the one-electron Schrödinger equations in the Hartree and Hartree-Fock approximations. Once the equations are solved, the ground-state total energy may be determined as follows (*25*, *26*):

$$E^n_{\mathrm{Tot},0} = \sum_{i=1}^{n} E_i - \frac{q^2}{4\pi\epsilon_0}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}' - \int \rho(\mathbf{r})V_{XC}\mathrm{d}^3\mathbf{r} + E_{XC}[\rho]\,. \tag{16}$$

In principle, the Kohn-Sham scheme can yield the exact values of the ground-state density and total energy, though in practice E_{XC} is not known exactly and must be approximated. However, because the Kohn-Sham equations are one-electron Schrödinger equations for a fictive, auxiliary system of noninteracting electrons, *there is no rigorous justification to treat the eigenvalues E_i that come from solving these equations as addition or removal energies of electrons of the true n-electron system*. This means, for example, that $E_{k+1} - E_k$ may not even be a good approximation for the energy needed for an electron to be excited from energy level k to $k+1$. Often, the band structure estimated from treating Kohn-Sham eigenvalues as if they were one-electron energies will capture qualitative trends but underestimate band gaps (*13*).

There are several other *ab initio* methods, which will be mentioned only briefly here since they are largely outside the scope of this survey. The configuration interaction method builds upon the Hartree-Fock method and expands the many-electron wavefunction into a linear combination of Slater determinants rather than a single one as in equation 1 (*11*). Both GW and Bethe-Salpeter are variations of many-body perturbation theories. The GW method is a generalization of the Hartree-Fock approximation that accounts for dynamic Coulomb screening and can represent charged excitations to yield band structures that are a closer match to experiment (*13*). The Bethe-Salpeter equation is derived from the so-called four-point Dyson equation, which therefore can represent excitonic effects and may be suited for accurate estimations of photoabsorption spectra (*13*, *15*). And, unlike the original DFT for time-independent systems, *time-dependent* DFT is capable of yielding excitation energies and photoabsorption spectra (*15*, *27*). However, like regular DFT, it is a reformulation of the many-body quantum mechanical equation in terms of electron densities and is dependent on an available approximation for the exchange-correlation energy E_{XC}. This can lead, for example, to difficulties in determining the excited states of nonmetallic bulk solids (*27*).

The examples of *ab initio* methods briefly outlined have several points in common. First, they are free of empirical parameters aside from fundamental constants such as $\hbar$ or m_e. Second, the operator that each uses to account for interactions between electrons, $\hat{V}_{\mathrm{e-e}}(\mathbf{r})$,

depends on the very one-particle wavefunctions $\psi_i(\mathbf{r})$ one is seeking to solve. In principle, one can solve the effective one-particle Schrödinger equation *self-consistently*, that is, start from trial values of $\psi_i(\mathbf{r})$ and iterate. In practice, the Hartree-Fock and Kohn-Sham equations are usually solved this way. Third, any effects of strain can be taken into account entirely through changes in the positions of the atomic nuclei $\{\mathbf{R}\}_M$, which manifest through changes in the external potential V_{ext}. Finally, these methods are all very computationally expensive and are suitable primarily for very small systems. For example, as a rule of thumb, given the current state of the art, DFT is usually feasible for systems of up to a few hundred atoms (*28*), though there are some numerical algorithms that introduce approximations that allow DFT to be applied to systems with thousands of atoms (*21*). The GW method is even more expensive and generally suitable for systems with no more than a few 10s of atoms (*29*, *30*). This starkly contrasts with the scalability of the empirical methods discussed in the following sections.

2.2 Spin-Orbit Interaction

In much of this survey, the effects of spin will be ignored for the sake of simplicity. However, if need be, to account for spin in the one-electron picture, the wavefunction may be written as a spinor, where (*31*)

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \psi_{i\mathbf{k}}^{\uparrow}(\mathbf{r}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \psi_{i\mathbf{k}}^{\downarrow}(\mathbf{r}) \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{17}$$

where the symbols "↑" and "↓" denote the spin-up and spin-down states of an electron. $\psi_{i\mathbf{k}}^{\dagger}(\mathbf{r})$ is, then, the Hermitian conjugate of $\psi_{i\mathbf{k}}(\mathbf{r})$. The one-electron Hamiltonian with a spin-orbit correction term is (*32*)

$$\widehat{H}_{1\text{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) = \frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_e} + \hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_\infty) + \frac{\hbar}{4m_e^2 c^2} \left[\nabla \hat{V}_{\text{ext,eff}} \times \hat{\boldsymbol{\sigma}}\right], \tag{18}$$

where c is the speed of light and $\hat{\boldsymbol{\sigma}}$ is a vector operator whose elements are the Pauli spin matrices (*31*, *32*)

$$\hat{\sigma}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_2 = \begin{bmatrix} 0 & -\text{i} \\ -\text{i} & 0 \end{bmatrix}, \text{ and } \hat{\sigma}_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{19}$$

The effect of spin-orbit interaction on the band structure of GaAs is shown in figure 9. This interaction introduces the split-off band, so called because it splits away from the valence band maximum. This effect occurs not only in GaAs but is common in semiconductors (*32*).

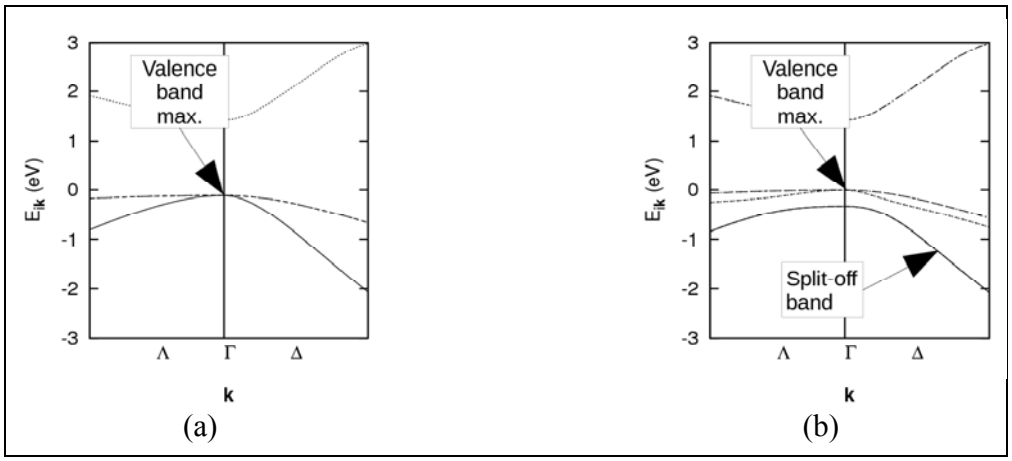

Figure 9. The calculated band structure of GaAs near the Γ point (a) without and (b) with spin-orbit correction. The band structures were estimated using the method and parameters of Boykin et al. (*33*).

3. Atomistic Methods

Empirical approaches tend to be less computationally expensive and thus scalable to relatively large systems, and this holds true even for the empirical methods that retain atomic-level resolution, such as the empirical pseudopotential (*34*) and tight-binding (*35*) methods. Simulations with these methods can be feasible for systems with thousands or even millions of atoms. One reason for their lack of computational expense as compared to *ab initio* methods is that they do not require iterating to self-consistency. While these methods can involve determining the eigenvalues of large matrices, these values only need to be determined once in the course of a calculation. Another reason is that the calculation of the elements of these matrices is comparatively less involved than those of the matrices formed in the implementation of the *ab initio* methods. For example, there is no need to integrate a functional of the electron density $\rho(\mathbf{r})$, which in turn requires a sum over n one-electron wavefunctions. There is a cost to this relative ease of computational expense, though. First, these methods require parameters that are fit to experiment and/or previous *ab initio* results, and the process of fitting these parameters is not trivial. As an example, Boykin et al. (*33*) use a genetic algorithm to fit over 30 parameters. Second, these methods involve various approximations, such as neglecting certain terms or assuming that various quantities may be expressed in certain functional forms. Even the lack of iterating until self-consistency entails an approximation, one that prevents the changes in the positions of the atomic nuclei $\{\mathbf{R}\}_M$ from fully taking the effects of strain into account, a matter that which will be discussed in more detail in section 3.3. Nonetheless, these empirical approaches are useful, especially where *ab initio* methods are impractical.

3.1 Empirical Pseudopotential Method

The empirical pseudopotential method (*34*) is an approach originally used to determine the electronic band structure of bulk crystals (*36*, *37*) where the effective potential is periodic. The effective potential may be split up as (*38*)

$$\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty) = V^{\text{loc}}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty) + \hat{V}_{NL}\,, \tag{20}$$

where $\{\mathbf{R}\}_\infty$ contains the set of coordinates of all the atoms of the infinite bulk crystal. The term $V^{\text{loc}}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)$ is a function of position and is here called the *local* part of the potential, while the operator $\hat{V}_{NL}$ is a correction term that accounts for $\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)$ not being a mere function of position and is called the *nonlocal* part. Both terms of $\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)$ are periodic, and $V^{\text{loc}}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)$ can be expanded as a complex Fourier series,

$$V^{\text{loc}}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty) = \textstyle\sum_{\mathbf{G}} \tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G})\mathrm{e}^{\mathrm{i}\mathbf{G}\cdot\mathbf{r}}\,, \tag{21}$$

where (*10*, *39*)

$$\tilde{\boldsymbol{V}}^{\mathbf{loc}}_{\mathbf{ext,eff}}(\mathbf{G}) = \frac{\mathbf{1}}{\boldsymbol{\Omega_c}}\int_{\boldsymbol{\Omega_c}} \boldsymbol{V}^{\mathbf{loc}}_{\mathbf{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)\,\mathbf{e}^{-\mathbf{i}\mathbf{G}\cdot\mathbf{r}}\mathbf{d}^{\mathbf{3}}\mathbf{r}\,. \tag{22}$$

Here, Ω_c is the volume of the unit cell of the bulk crystal and the quantity $\mathbf{G}$ is a reciprocal lattice vector. Because the bulk crystal is periodic, the Bloch theorem applies, i.e., $\psi_{i\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}u_{i\mathbf{k}}(\mathbf{r})$. Because of the periodicity of $u_{i\mathbf{k}}(\mathbf{r})$, it can be expanded in a complex Fourier series, so that

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \tfrac{1}{\sqrt{\Omega_c}}\textstyle\sum_{\mathbf{G}} \tilde{u}_{i\mathbf{k}}(\mathbf{G})\,\mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\,, \tag{23}$$

where $1/\sqrt{\Omega_c}$ is a normalization prefactor. Alternatively, this may be described as expanding $\psi_{i\mathbf{k}}(\mathbf{r})$ in terms of plane waves, each with a wavevector $\mathbf{k}+\mathbf{G}$. If one (1) substitutes equation 23 into equation 7, the one-electron Schrödinger equation, (2) multiplies both sides of the equation by $\mathrm{e}^{-\mathrm{i}(\mathbf{k}+\mathbf{G}')\cdot\mathbf{r}}/\sqrt{\Omega_c}$, (3) integrates, and (4) substitutes equation 22, one may obtain the following matrix equation (*10*):

$$\sum_{\mathbf{G}}\left[\frac{\hbar^2}{2m_e}|\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{G}\mathbf{G}'} + \tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G}'-\mathbf{G}) + \tilde{V}_{NL}(\mathbf{k},\mathbf{G}',\mathbf{G})\right]\tilde{u}_{i\mathbf{k}}(\mathbf{G}) = E_{i\mathbf{k}}\tilde{u}_{i\mathbf{k}}(\mathbf{G}')\,, \tag{24}$$

where

$$\tilde{V}_{NL}(\mathbf{k},\mathbf{G}',\mathbf{G}) = \tfrac{1}{\Omega_c}\textstyle\int_{\Omega_c} \mathrm{e}^{-\mathrm{i}(\mathbf{k}+\mathbf{G}')\cdot\mathbf{r}}\hat{V}_{NL}\mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\mathrm{d}^3\mathbf{r} \tag{25}$$

and

$$\delta_{\mathbf{G}\mathbf{G}'} = \frac{1}{\Omega_*}\int_{\Omega_*} \mathrm{e}^{\mathrm{i}(\mathbf{G}'-\mathbf{G})\cdot\mathbf{r}}\mathrm{d}^3\mathbf{r} = \begin{cases} 1, & \mathbf{G} = \mathbf{G}' \\ 0, & \mathbf{G} \neq \mathbf{G}' \end{cases}. \tag{26}$$

Here, Ω_* is an arbitrary volume, which for the previous derivation of the matrix equation is Ω_c. Given a value of the Bloch wave vector **k,** equation 24 can then be solved for the eigenvalues $E_{i\mathbf{k}}$. In principle, there are infinitely many values of the reciprocal lattice vector **G**, so to solve the matrix equation numerically, a subset of the values of this vector are used that satisfy the condition

$$\frac{\hbar^2}{2m_e}|\mathbf{k}+\mathbf{G}|^2 \leq E_{\mathrm{cut}}\ , \tag{27}$$

where E_{cut} is an energy cutoff (*37*). Algorithms for finding reciprocal lattice vectors within the energy cutoff are in appendix A.

If spin-orbit coupling is taken into account, for each wave vector **k** and reciprocal lattice vector **G** there are two plane waves, one for a spin-up electron and one for a spin-down electron. Equation 24 is then modified to become

$$\sum_{\mathbf{G}}\sum_s\left[\widetilde{H}(\mathbf{k},\mathbf{G}',\mathbf{G})\delta_{ss'} + \widetilde{H}_{SO}(\mathbf{k},\mathbf{G}',\mathbf{G},s',s)\right]\tilde{u}_{i\mathbf{k}}(\mathbf{G},s) = E_{i\mathbf{k}}\tilde{u}_{i\mathbf{k}}(\mathbf{G}',s')\ , \tag{28}$$

where $s, s' \in \{\uparrow,\downarrow\}$,

$$\widetilde{H}(\mathbf{k},\mathbf{G}',\mathbf{G}) = \frac{\hbar^2}{2m_e}|\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{G}\mathbf{G}'} + \tilde{V}^{\mathrm{loc}}_{\mathrm{ext,eff}}(\mathbf{G}'-\mathbf{G}) + \tilde{V}_{NL}(\mathbf{k},\mathbf{G}',\mathbf{G})\ , \tag{29}$$

$$\delta_{ss'} = \begin{cases} 1, & s = s' \\ 0, & \text{otherwise} \end{cases}, \tag{30}$$

and $\widetilde{H}_{SO}(\mathbf{k},\mathbf{G}',\mathbf{G},s',s)$ is a correction term for spin-orbit coupling. In some formulations, this term is expressed analytically as a function of **k**, **G**, and **G'** (*38*), while in others, the spin-orbit correction is calculated in real space and then numerically converted into Fourier space (*40*).

Both the nonlocal and spin-orbit correction terms may be neglected, as was done, for example, in the formulation by Cohen and Bergstresser (*37*). Alternatively, the spin-orbit correction may be kept while neglecting any other nonlocal corrections, as was done by Williamson and Zunger (*41*). Williamson et al. (*42*) accounted for these other nonlocal corrections not through the term $\hat{V}_{NL}$, but by multiplying the kinetic energy operator $\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}/2m_e$ (or equivalently, $\hbar^2|\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{G}\mathbf{G}'}/2m_e$, in Fourier space), by a fitting parameter. For the sake of simplicity, neither the nonlocal correction term nor the spin-orbit terms in the empirical pseudopotential method will be discussed further.

The principal approximation in the empirical pseudopotential method is to express $\tilde{V}_{\text{ext,eff}}(\mathbf{G}' - \mathbf{G})$ as a function of empirically determined parameters. The local part of the effective potential $V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r}, \{\mathbf{R}\}_\infty)$ is taken to be a sum of contributions from individual atoms (*34*),

$$V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r}, \{\mathbf{R}\}_\infty) = \sum_{i=1}^{\infty} \sum_{j=1}^{N_{\text{per cell}}} v_j\left(\mathbf{r} - \overline{\mathbf{R}}_i - \mathbf{d}_j\right), \tag{31}$$

where $\overline{\mathbf{R}}_i$ is the lattice vector of the i^{th} unit cell of the bulk crystal, and $\overline{\mathbf{R}}_i + \mathbf{d}_j$ is the position vector of the j^{th} atom within this cell. $N_{\text{per cell}}$ is the number of atoms per unit cell. One can rewrite the summation as

$$V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r}, \{\mathbf{R}\}_\infty) = \sum_{j=1}^{N_{\text{per cell}}} V_j\left(\mathbf{r} - \mathbf{d}_j\right), \tag{32}$$

where

$$V_j(\mathbf{r}) = \sum_{i=1}^{\infty} v_j(\mathbf{r} - \overline{\mathbf{R}}_i)\,. \tag{33}$$

Both $V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r}, \{\mathbf{R}\}_\infty)$ and $V_j(\mathbf{r})$ are periodic and can also be expressed as complex Fourier series. The former has already been shown in equations 21 and 22. For the latter, one may write

$$V_j(\mathbf{r}) = \sum_{\mathbf{G}} \tilde{V}_j(\mathbf{G}) \mathrm{e}^{\mathrm{i}\mathbf{G}\cdot\mathbf{r}}\,,\ \tilde{V}_j(\mathbf{G}) = \frac{1}{\Omega_c}\int_{\Omega_c} V_j(\mathbf{r})\, \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{r}} \mathrm{d}^3\mathbf{r}\,. \tag{34}$$

Substituting equation 32 into equation 22 yields

$$\begin{aligned}
\tilde{V}_{\text{ext,eff}}^{\text{loc}}(\mathbf{G}) &= \sum_{j=1}^{N_{\text{per cell}}} \frac{1}{\Omega_c}\int_{\Omega_c} V_j\left(\mathbf{r} - \mathbf{d}_j\right) \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{r}} \mathrm{d}^3\mathbf{r}\ , \\
&= \sum_{j=1}^{N_{\text{per cell}}} \frac{1}{\Omega_c}\int_{\Omega_c} V_j(\mathbf{r})\, \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot(\mathbf{r}+\mathbf{d}_j)} \mathrm{d}^3\mathbf{r}\,, \\
&= \sum_{j=1}^{N_{\text{per cell}}} \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{d}_j} \frac{1}{\Omega_c}\int_{\Omega_c} V_j(\mathbf{r})\, \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{r}} \mathrm{d}^3\mathbf{r}\,,
\end{aligned}$$

and

$$= \sum_{j=1}^{N_{\text{per cell}}} \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{d}_j}\, \tilde{V}_j(\mathbf{G})\,. \tag{35}$$

$\tilde{V}_j(\mathbf{G})$ is an *atomic form factor*, and in the empirical pseudopotential method, it is either an empirical parameter (*36*, *38*, *43*) or is expressed in terms of empirically determined parameters (*10*, *34*, *37*). Pseudocode for an implementation of the empirical pseudopotential method is shown in figure 10. This implementation assumes that the one-electron wavefunctions have been expanded in terms of plane waves, as in equation 23. For diamond and zincblende crystals, equation 35 is often rewritten in terms of symmetric and antisymmetric form factors $V_S(\mathbf{G})$ and $V_A(\mathbf{G})$, rather than directly in terms of the atomic form factors. For such crystals, $N_{\text{per cell}} = 2$, and if the lattice vector $\overline{\mathbf{R}}$ is taken to point to the center of the unit cell, then $\mathbf{d}_1 = -\mathbf{d}$ and $\mathbf{d}_2 = \mathbf{d}$, where $\mathbf{d} = (\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)/8$, with $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ being the lattice vectors shown in figure 3c. Given Euler's formula $\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$, one may rewrite equation 35 as

$$\begin{aligned}\tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G}) &= \left[\widetilde{V}_1(\mathbf{G}) + \widetilde{V}_2(\mathbf{G})\right]\cos(\mathbf{G}\cdot\mathbf{d}) + \mathrm{i}\left[\widetilde{V}_1(\mathbf{G}) - \widetilde{V}_2(\mathbf{G})\right]\sin(\mathbf{G}\cdot\mathbf{d}) \\ &= V_S(\mathbf{G})\cos(\mathbf{G}\cdot\mathbf{d}) + \mathrm{i}V_A(\mathbf{G})\sin(\mathbf{G}\cdot\mathbf{d})\,.\end{aligned} \tag{36}$$

The symmetric and antisymmetric form factors then become fitting parameters instead of the atomic form factors themselves. The method shown in figure 10 is then modified by taking $V_S(\mathbf{G})$ and $V_A(\mathbf{G})$ as givens rather than the atomic form factors, and by calculating $\tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G})$ through equation 36 instead of 35. Cohen and Bergstresser (37) treated $V_S(\mathbf{G})$ and $V_A(\mathbf{G})$ as functions of the magnitude of $\mathbf{G}$ in units of $\frac{2\pi}{a}$, where a is the lattice constant of the crystal, so that $V_S(\mathbf{G}) = V_S\left(\frac{a|\mathbf{G}|}{2\pi}\right)$ and $V_A(\mathbf{G}) = V_A\left(\frac{a|\mathbf{G}|}{2\pi}\right)$. The values of their form factors for particular values of $|\mathbf{G}|$ were fit to reflectivity and photoemission experiments. Examples of some of these parameters for bulk materials Si and GaAs are shown in table 1. For Si, $\widetilde{V}_1(\mathbf{G}) = \widetilde{V}_2(\mathbf{G})$, which means $V_A(\mathbf{G}) = 0$ for this material. If the unit cell changes size or shape while the relationship $\mathbf{d} = (\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)/8$ still holds, then

$$\mathbf{G}\cdot\mathbf{d} = \tfrac{1}{8}\textstyle\sum_{i=1}^{3} n_i\mathbf{b}_i\cdot(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3) = \tfrac{\pi}{4}\sum_{i=1}^{3} n_i\,;\ \text{and } n_i \text{ is an integer}\,. \tag{37}$$

Table 1. Example symmetric and antisymmetric form factors $V_S(a|\mathbf{G}|/2\pi)$ and $V_A(a|\mathbf{G}|/2\pi)$ from Cohen and Bergstresser (*37*), where a is the lattice constant of the material and the form factors are in Rydbergs.

Material	$V_S(\sqrt{3})$	$V_S(2)$	$V_S(\sqrt{8})$	$V_A(\sqrt{3})$	$V_A(2)$	$V_A(\sqrt{8})$
Si	–0.21	0	0.04	0	0	0
GaAs	–0.23	0	0.01	0.07	0.05	0

The quantity $\mathbf{G}\cdot\mathbf{d}$, then, is independent of strains that preserve the relative locations of the atoms within the unit cell, such as hydrostatic strains. However, the form factors are not independent of strain. For example, Aouina et al. (*43*) have recalculated the symmetric and antisymmetric form factors for AlAs under a range of pressures, as shown in table 2.

Atomic form factors $\tilde{V}_1(\mathbf{G})$, $\tilde{V}_2(\mathbf{G})$, ..., $\tilde{V}_{N_{\text{per cell}}}(\mathbf{G})$ are given.
Vectors $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{N_{\text{per cell}}}$ are given.
Wavevectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ are given.
E_{cut} is given.

Nonlocal potential part $\tilde{V}_{NL}(\mathbf{k}, \mathbf{G}', \mathbf{G})$ is assumed to be given.
Spin-orbit correction $\tilde{H}_{SO}(\mathbf{k}, \mathbf{G}', \mathbf{G}, s', s)$ is assumed to be given.
Boolean variable *noSpin* is true if there is no spin, and false otherwise.

for $\mathbf{k}$ in $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ **do**
  Find the $N_\mathbf{G}$ reciprocal lattice vectors $\mathbf{G}_1$, $\mathbf{G}_2$, ..., that satisfy
    $(\hbar^2/2m_e)|\mathbf{k}+\mathbf{G}_i|^2 \leq E_{\text{cut}}$.

  if *noSpin* **then**
    Let matrix $[H_{IJ}]$ be of size $N_\mathbf{G} \times N_\mathbf{G}$.
  else
    Let matrix $[H_{IJ}]$ be of size $2N_\mathbf{G} \times 2N_\mathbf{G}$.
  end if

  for $i = 1$ to $N_\mathbf{G}$ **do**
    $\mathbf{G}' = \mathbf{G}_i$
    for $j = 1$ to $N_\mathbf{G}$ **do**
      $\mathbf{G} = \mathbf{G}_j$

$$\tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G}'-\mathbf{G}) = \sum_{m=1}^{N_{\text{per cell}}} \mathrm{e}^{-\mathrm{i}(\mathbf{G}'-\mathbf{G})\cdot\mathbf{d}_m} \tilde{V}_m(\mathbf{G}'-\mathbf{G})$$

$$\tilde{H}(\mathbf{k}, \mathbf{G}', \mathbf{G}) = \frac{\hbar^2}{2m_e}|\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + \tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G}'-\mathbf{G}) + \tilde{V}_{NL}(\mathbf{k}, \mathbf{G}', \mathbf{G})$$

      if *noSpin* **then**
        Let $I = i$, $J = j$, and $H_{IJ} = \tilde{H}(\mathbf{k}, \mathbf{G}', \mathbf{G})$.
      else
        for s' in $\uparrow, \downarrow$ **do**
          for s in $\uparrow, \downarrow$ **do**
            Let $I = 2i-1$ if $s' = \uparrow$, and let $I = 2i$ otherwise.
            Let $J = 2j-1$ if $s = \uparrow$, and let $J = 2j$ otherwise.

            $H_{IJ} = \tilde{H}(\mathbf{k}, \mathbf{G}', \mathbf{G})\delta_{ss'} + \tilde{H}_{SO}(\mathbf{k}, \mathbf{G}', \mathbf{G}, s', s)$
          end for
        end for
      end if

    end for
  end for

  Find eigenvalues $E_{i\mathbf{k}}$ and eigenvectors $\tilde{u}_{i\mathbf{k}}$ of matrix $[H_{IJ}]$.
end for

Figure 10. Pseudocode for empirical pseudopotential method where the one-electron wavefunctions are expanded in a series of plane waves.

Table 2. Symmetric and antisymmetric form factors $V_S(a|\mathbf{G}|/2\pi)$ and $V_A(a|\mathbf{G}|/2\pi)$ for AlAs from Aouina et al. (*43*), calculated for several pressures. Pressure is in kilobars. Lattice constant a is in angstroms and the form factors are in Rydberg.

Pressure	**0**	**30**	**60**	**90**	**120**
$V_S(\sqrt{3})$	–0.212694	–0.212931	–0.212725	–0.211967	–0.210679
$V_S(\sqrt{8})$	0	0	0	0	0
$V_S(\sqrt{11})$	0.09275	0.099816	0.106973	0.114209	0.121637
$V_A(\sqrt{3})$	0.068833	0.068074	0.066888	0.065781	0.064803
$V_A(2)$	0.05	0.05	0.05	0.05	0.05
$V_A(\sqrt{11})$	–0.0075	–0.0075	–0.0075	–0.0075	–0.0075
Lattice constant $\boldsymbol{a}$	5.6611	5.6012	5.5406	5.4875	5.4403

Caruthers and Lin-Chung (*44*) work with the atomic form factors directly in their work on a GaAs-AlAs superlattice consisting of monoatomically thin layers of Ga, Al, and As. The unit cell of this superlattice is shown in figure 11 and consists of four atoms. For this case, equation 35 evaluates to

$$\tilde{V}^{\text{loc}}_{\text{ext,eff}}(\mathbf{G}) = \tilde{V}_{\text{Ga}}(\mathbf{G})\mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{d}_{\text{Ga}}} + \tilde{V}_{\text{Al}}(\mathbf{G})\mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{d}_{\text{Al}}} + \tilde{V}_{\text{As}}(\mathbf{G})\left[\mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{d}_{\text{As1}}} + \mathrm{e}^{-\mathrm{i}\mathbf{G}\cdot\mathbf{d}_{\text{As2}}}\right], \tag{38}$$

where $\mathbf{d}_{\text{Ga}}$, $\mathbf{d}_{\text{Al}}$, $\mathbf{d}_{\text{As1}}$, and $\mathbf{d}_{\text{As2}}$ are the vectors indicating the positions of the Ga, Al, and two As atoms within the unit cell. Caruthers and Lin-Chung take $\mathbf{d}_{\text{Ga}}$ to be zero. Like Cohen and Bergstresser, they take the atomic form factors to be functions of $a|\mathbf{G}|/2\pi$ and treat them as empirical parameters. A sample of these is shown in table 3. The values of these form factors are fitted in order to reproduce the electronic band structures of GaAs and AlAs. These form factors are no less strain-dependent than those of Cohen and Bergstresser or Aouina et al. However, in principle, there is nothing in the scheme of Caruthers and Lin-Chung that limits $\mathbf{d}_{\text{Al}}$, $\mathbf{d}_{\text{As1}}$, and $\mathbf{d}_{\text{As2}}$ to their equilibrium values, and they may change in response due to distortions from strain. When applying the empirical pseudopotential to more general nanostructures, the atoms of the structure may be regarded as being within a large unit cell, with $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \ldots$ as the positions of the atoms forming the nanostructure.

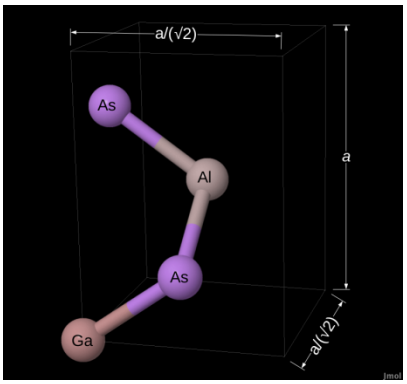

Figure 11. Unit cell of GaAs-A1As superlattice from Caruthers and Lin-Chung (*43*). a is the lattice constant of GaAs. The figure was created with Jmol (*45*).

Table 3. Example empirical parameters from Caruthers and Lin-Chung (*44*), where a is the lattice constant of the material, and the atomic form factors are in Rydberg.

$a\|\mathbf{G}\|/2\pi$	$\tilde{V}_{\mathrm{Ga}}(a\|\mathbf{G}\|/2\pi)$	$\tilde{V}_{\mathrm{Al}}(a\|\mathbf{G}\|/2\pi)$	$\tilde{V}_{\mathrm{As}}(a\|\mathbf{G}\|/2\pi)$
0	–0.1114	–0.1056	–0.1619
1	–0.0750	–0.0930	–0.1250
$\sqrt{2}$	–0.0560	–0.0760	–0.0970

Rather than use the values of atomic form factors at particular values of $a|\mathbf{G}|/2\pi$ as fitting parameters, one may instead express the form factors as continuous functions that contain empirical parameters. For example, Wang and Zunger (*46*) take the form factors to be

$$\tilde{V}_j(G) = \frac{a_{j1}(G^2 - a_{j2})}{a_{j3}\mathrm{e}^{a_{j4}G^2} - 1}, \text{ and } G = |\mathbf{G}|, \tag{39}$$

where a_{j1} through a_{j4} are fitting parameters. Similarly, Mäder and Zunger (*47*) express the form factors with the following function,

$$\tilde{V}_j(G) = \Omega_j\left[1 + f_j\mathrm{e}^{-\beta_j G^2}\right]\sum_{i=1}^{4} a_{ij}\mathrm{e}^{-c_{ij}\left(G - b_{ij}\right)^2}, \text{and } G = |\mathbf{G}|, \tag{40}$$

where the fitting parameters are Ω_j, f_j, β_j, a_{ij}, b_{ij}, and c_{ij}. They also fit these parameters for a range of unit cell volumes, so that the fitting parameters themselves have the desired property of being independent of at least the hydrostatic strain. Here, the reciprocal lattice vector $\mathbf{G}$ is not normalized by a factor of $a/2\pi$, so changes in a due to strain are reflected in G. Also, since these form factors are continuous functions of G, the allowed values of $\mathbf{G}$ are, at least in principle, arbitrary.

The large cell containing a nanostructure (or a representative unit cell of it) may be regarded as a supercell composed of smaller cells at least approximately resembling those of a bulk crystal. $V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r},\{\mathbf{R}\}_\infty)$, then, may be written as

$$V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r},\{\mathbf{R}\}_\infty) = \sum_{i=1}^{\infty}\sum_{j=1}^{N_{\text{cell}}}\sum_{\kappa=1}^{N_{\text{per cell}}(j)} v_{j,\kappa}\left(\mathbf{r}-\bar{\mathbf{R}}_i-\bar{\bar{\mathbf{R}}}_j-\bar{\mathbf{d}}_\kappa(j)\right), \tag{41}$$

where $\bar{\mathbf{R}}_i$ is the lattice vector pointing to the i^{th} supercell, $\bar{\mathbf{R}}_i+\bar{\bar{\mathbf{R}}}_j$ is the vector pointing to small cell j within supercell i, and $\bar{\mathbf{d}}_\kappa(j)$ is the relative position of atom κ within small cell j. As long as the wavefunction is expanded in terms of periodic functions—even if those functions are not plane waves—a summation over all the supercells tiled out to infinity is needed. The unit cells are not all alike, and may have different kinds of atoms within them. This is why $N_{\text{per cell}}$, $\bar{\mathbf{d}}_\kappa(j)$, and $v_{j,\kappa}$ are denoted here as depending on the index for each small cell j. These cells may even be distorted due to strain. The preceding summation may also be rewritten as

$$V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r},\{\mathbf{R}\}_\infty) = \sum_{i=1}^{\infty}\sum_{j=1}^{N_{\text{cell}}}\sum_{\alpha} W_\alpha(j) v_\alpha\left(\mathbf{r}-\bar{\mathbf{R}}_i-\bar{\bar{\mathbf{R}}}_j-\bar{\mathbf{d}}_\alpha\right), \tag{42}$$

where α is one of the atom types in the system, e.g., Ga, In, As, etc., and $\bar{\mathbf{d}}_\alpha$ is now the relative position of an atom of type α within one of the small cells of the system. Since a given atom type may not be in a particular unit cell, the summation includes a weighting factor $W_\alpha(j)$ that is 0 if atom α is not within small cell j (*48*, *49*). If there is no strain in the system, then it is 1 if there is an atom at $\bar{\mathbf{d}}_\alpha$ (*49*). Substituting equation 33 into the previous summation yields the decomposition of $\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)$ found in Wang et al. (*48*) and Wang and Zunger (*49*):

$$V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r},\{\mathbf{R}\}_\infty) = \sum_{j=1}^{N_{\text{cell}}}\sum_{\alpha} W_\alpha(j) V_\alpha\left(\mathbf{r}-\bar{\bar{\mathbf{R}}}_j-\bar{\mathbf{d}}_\alpha\right). \tag{43}$$

In the formulations by Kim et al. (*50*) and Williamson et al. (*42*), the weighting factor from the previous equations, $W_\alpha(j)$, becomes strain-dependent and has the assumed form

$$W_\alpha(j) = 1+\gamma_\alpha \text{Tr}[\boldsymbol{\epsilon}(j)], \tag{44}$$

where γ_α is a fitting parameter and $\text{Tr}[\boldsymbol{\epsilon}(j)]$ is the trace of the strain tensor at cell j. Williamson and Zunger (*41*) have found that without an explicitly strain-dependent pseudopotential, changes in the band gap E_g due to strain may be fit to experiment, but not the changes in the conduction band minima and valence band maxima themselves.

While the empirical pseudopotential method has traditionally used a Fourier expansion of the one-electron wavefunction, other basis set expansions can be used. For example, in the linear combination of bulk bands (LCBB) form of the method (*48*, *49*),

$$\psi_i(\mathbf{r}) = \sum_\sigma\sum_m\sum_{\mathbf{k}} C_{m\mathbf{k}\sigma}^i \psi_{m\mathbf{k}}^\sigma(\mathbf{r}), \text{ and } \psi_{m\mathbf{k}}^\sigma(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} u_{m\mathbf{k}}^\sigma(\mathbf{r}), \tag{45}$$

where $C^i_{m\mathbf{k}\sigma}$ is an expansion coefficient, $\psi^\sigma_{m\mathbf{k}}(\mathbf{r})$ is the one-electron wavefunction with band index m for a bulk material σ (e.g., GaAs or InAs), and the values of $\mathbf{k}$ are a sampling of points within the Brillouin zone *of the bulk material* σ. Even though the one-electron wavefunction is no longer directly expanded in a basis of plane waves, the previously shown atomic form factors expressed in Fourier space, such as those in table 3 or equations 39 and 40, may still be used. This is because $u^\sigma_{m\mathbf{k}}(\mathbf{r})$ is still expanded in a plane-wave basis, so that

$$\psi^\sigma_{m\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega^\sigma_c}} \sum_{\mathbf{G}} \tilde{u}^\sigma_{m\mathbf{k}}(\mathbf{G})\, \mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\,, \tag{46}$$

where Ω^σ_c is the volume of the primitive cell of bulk material σ. The expansion coefficients $\tilde{u}^\sigma_{m\mathbf{k}}(\mathbf{G})$ may be determined through applying the traditional plane-wave based empirical pseudopotential method to a unit cell of material σ for a chosen value of E_{cut}, following the method illustrated in figure 10. For the case where the nonlocal part of the potential is zero and there is no spin-orbit correction, the matrix equation to be solved is

$$\sum_\sigma \sum_m \sum_{\mathbf{k}} H^{\sigma'\sigma}_{m'\mathbf{k}'m\mathbf{k}} C^i_{m\mathbf{k}\sigma} = E_i C^i_{m'\mathbf{k}'\sigma'}\,, \tag{47}$$

where

$$\begin{aligned} H^{\sigma'\sigma}_{m'\mathbf{k}'m\mathbf{k}} = {} & \frac{\hbar^2\Omega}{2m_e} \sum_{\mathbf{G},\mathbf{G}'} \tilde{u}^{\sigma'}_{m'\mathbf{k}'}(\mathbf{G}')\tilde{u}^{\sigma}_{m\mathbf{k}}(\mathbf{G})|\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{k}\mathbf{k}'}\delta_{\mathbf{G}\mathbf{G}'} \\ & + \sum_{\mathbf{G},\mathbf{G}'} \tilde{u}^{\sigma'}_{m'\mathbf{k}'}(\mathbf{G}')\tilde{u}^{\sigma}_{m\mathbf{k}}(\mathbf{G}) \\ & \times \sum_\alpha \tilde{V}_\alpha(\mathbf{k}'+\mathbf{G}'-\mathbf{k}-\mathbf{G}) \mathrm{e}^{-\mathrm{i}\bar{\mathbf{d}}_\alpha\cdot(\mathbf{k}+\mathbf{G}-\mathbf{k}'-\mathbf{G}')} \sum_{j=1}^{N_{\mathrm{cell}}} W_\alpha(j) \mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\bar{\mathbf{R}}_j}. \end{aligned} \tag{48}$$

$\mathbb{b}$ is the volume of the simulation domain, $\tilde{V}_\alpha$ is the atomic form factor expressed in Fourier space, and $W_\alpha(j)$ is the weighting factor of equation 43. There is also a variant of the LCBB pseudopotential method called the strained linear combination of bulk bands (SLCBB), where the basis function $\psi^\sigma_{m\mathbf{k}}(\mathbf{r})$ is $\sqrt{J(\mathbf{r})}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}u^\sigma_{m\mathbf{k}}\left(\boldsymbol{\chi}^{-1}(\mathbf{r})\right)$, where $\boldsymbol{\chi}^{-1}$ is an inverse deformation map that maps coordinate $\mathbf{r}$ in the deformed system onto its counterpart $\mathbf{r}_0$ in the undeformed system, and $\mathrm{d}^3\mathbf{r}^0 = J(\mathbf{r})\mathrm{d}^3\mathbf{r}$. In a sense, the LCBB and SLCBB methods can be described as *indirect* plane wave expansions. A one-electron wavefunction of the whole system, $\psi_i(\mathbf{r})$, is expanded in terms of the one-electron wavefunction of its constituent bulk materials, $\{\psi^\sigma_{m\mathbf{k}}(\mathbf{r})\}$, which are then expanded into plane waves. This indirect expansion, though, means that the dimensions of the matrix $\left[H^{\sigma'\sigma}_{m'\mathbf{k}'m\mathbf{k}}\right]$ are independent of the number of plane waves used to expand $\psi^\sigma_{m\mathbf{k}}(\mathbf{r})$.

The overall speed of the empirical pseudopotential method can be illustrated from a couple of examples. A parallel code implementing the traditional plane-wave version of the method ran on

a Cray-T3E900 cluster. In this implementation, the one-electron wavefunction was taken to have the plane wave expansion shown in equation 23. The atomic form factors were from Wang et al. (*51*), and the energy cutoff was set to $E_{\text{cut}} = 5$ Ry. The code was used to simulate a system composed of an InAs quantum dot embedded in a GaAs matrix. For a system containing 250,000 atoms, this code took about 20 h to run on 128 processors of the cluster (*52*, *51*). A serial code implementing the SLCBB version of the empirical pseudopotential method (*53*) was able to simulate the same 250,000-atom system in about 10 h on an IBM 595 workstation, and the eigenvalues determined by this code were within 0.8% of those calculated by the parallel plane-wave code. More details of the SLCBB simulation, such as the sampling of the Brillouin zone, the bulk bands chosen, etc., are in Wang and Zunger (*49*). Both codes used the same method to determine eigenvalues of a matrix, the folded-spectrum method (*46*), which scales linearly with the dimension of the matrix. The scalability of the empirical pseudopotential method is also partially determined by the choice of basis functions. Now, both the plane-wave and LCBB/SLCBB forms of the empirical pseudopotential method require a sum over the atoms in the simulation for each matrix element, at least if spin is not taken into account. In the pseudocode shown in figure 4 for the plane-wave method, this is apparent in the calculation of $V_{\text{ext,eff}}^{\text{loc}}(\mathbf{r}, \{\mathbf{R}\}_{\infty})$, where there is a summation over all the atoms in the unit cell, and it is implied in equation 48 for the LCBB method, where there is a summation over the N_{cell} small cells that compose the supercell of the system and a summation over the atom types α in the system.

However, in the traditional plane-wave empirical pseudopotential method, another scalability concern is that the size of the matrix $[H_{ij}]$ whose eigenvalues are to be found is determined by the number of plane waves $N_{\mathbf{G}}$, whose reciprocal lattice vectors satisfy an energy cutoff criterion (i.e., equation 27). This number may be roughly estimated from equations A-1 and A-2 in appendix A to be

$$N_{\mathbf{G}} \approx (n_1^{\max})(n_2^{\max})(n_3^{\max}) = \frac{1}{|\mathbf{b}_1||\mathbf{b}_2||\mathbf{b}_3|}\left(\sqrt{\frac{2m_e E_{\text{cut}}}{\hbar^2}}\right)^3. \tag{49}$$

This, in turn, is proportional to the volume Ω of the unit cell of the simulated system. Since $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$, the magnitude of a primitive reciprocal lattice vector is $|\mathbf{b}_i| = 2\pi/(|\mathbf{a}_i|\cos\theta_i)$, where θ_i is the angle between $\mathbf{b}_i$ and its corresponding primitive lattice vector $\mathbf{a}_i$. Therefore,

$$N_{\mathbf{G}} \approx \left(\prod_{i=1}^{3} \frac{|\mathbf{a}_i|\cos\theta_i}{2\pi}\right)\left(\sqrt{\frac{2m_e E_{\text{cut}}}{\hbar^2}}\right)^3 \propto \mathbb{b}\ , \tag{50}$$

because $\mathbb{b} \propto \prod_{i=1}^{3}|\mathbf{a}_i|$. This scaling with the volume holds regardless of whether the volume is filled with atoms, so it would apply if the plane-wave pseudopotential method were applied to, for example, a cluster of atoms at the center of an otherwise empty unit cell that was made large enough to prevent the cluster from interacting with its periodic images. In the LCBB and SLCBB forms of the empirical pseudopotential method, though, the number of plane waves used to expand each basis function $\psi_{m\mathbf{k}}^{\sigma}(\mathbf{r})$ is independent of the volume or even the number of atoms in the simulation. The LCBB and SLCBB methods are clearly more scalable than the plane-wave

version of the empirical pseudopotential method, but the matrix elements in the former two methods are more computationally expensive to calculate, so such methods may not be ideal for smaller systems. Also, application of the LCBB and SLCBB methods requires a careful choice of the bulk band indices and Brillouin zone sampling (i.e., m and $\mathbf{k}$ in equations 45–48). In short, there are tradeoffs between these two forms of the empirical pseudopotential method.

3.2 Slater-Koster Tight-Binding Method

In tight-binding methods (*54–59*), the one-electron wavefunction is expanded in terms of either atomic orbitals or functions with the same symmetries as atomic orbitals. For a finite system of M atoms, one may write (56, 57)

$$\psi_i(\mathbf{r}) = \sum_{m=1}^{M} \sum_o C_{om}^{(i)} \phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m) \,, \tag{51}$$

where o is a type of orbital symmetry, i.e., $o \in \{s, p_x, p_y, p_z, d_{xy}, \ldots\}$, $C_{om}^{(i)}$ is an expansion coefficient, and $\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)$ is the atomic orbital or orbital-like function centered at position vector $\mathbf{R}_m$, with symmetry o for an atom whose species is indicated by atomic number Z_m.

$\phi_{oZ_m}(\mathbf{r})$ is real, so $\phi_{oZ_m}^{\dagger}(\mathbf{r}) = \phi_{oZ_m}(\mathbf{r})$. Also, $\phi_{oZ_m}(\mathbf{r})$ is *localized* and therefore decays to zero as $|\mathbf{r}| \to \infty$. For an infinite periodic system that satisfies the Bloch theorem, the wavefunction is expanded instead as (*54*, *57*, *60*)

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sum_{m=1}^{N_{\text{per cell}}} \sum_o C_{om}^{(i\mathbf{k})} \phi_{o\mathbf{k}m}(\mathbf{r}) \,, \tag{52}$$

where $N_{\text{per cell}}$ is the number of atoms per unit cell, $C_{om}^{(i\mathbf{k})}$ is an expansion coefficient, and $\phi_{o\mathbf{k}m}(\mathbf{r})$ is a Bloch sum, defined such that

$$\phi_{o\mathbf{k}m}(\mathbf{r}) = \lim_{N_{\text{cell}} \to \infty} \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{j=1}^{N_{\text{cell}}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\overline{\mathbf{R}}_j + \mathbf{d}_m)} \phi_{oZ_m}\left(\mathbf{r} - \overline{\mathbf{R}}_j - \mathbf{d}_m\right) . \tag{53}$$

Here, N_{cell} is the number of unit cells in the system, $\overline{\mathbf{R}}_j$ is the lattice vector pointing to unit cell j, and $\overline{\mathbf{R}}_j + \mathbf{d}_m$ is the position of atom m in cell j. If one substitutes the tight-binding expansion for a finite system, i.e., equation 51, into the one-electron Schrödinger equation 7, multiplies both sides of the equation by $\phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{R}_{m'})$, and integrates over all space, then one obtains the following matrix equation:

$$\sum_{m=1}^{M} \sum_o H_{o'm'om} C_{om}^{(i)} = E_i \sum_{m=1}^{M} \sum_o S_{o'm'om} C_{om}^{(i)} \,, \tag{54}$$

where

$$H_{o'm'om} = \int \phi_{o'Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'})\,\hat{H}_{1\mathrm{e}}\text{-}\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\mathrm{d}^3\mathbf{r} \tag{55}$$

and

$$S_{o'm'om} = \int \phi_{o'Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'})\,\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\mathrm{d}^3\mathbf{r}\,. \tag{56}$$

Matrix equation 54 is a generalized eigenvalue problem that may be solved for E_i. $S_{o'm'om}$ is called the overlap matrix, and if it were diagonal, that is, if $S_{o'm'om} = \delta_{m'm}\delta_{o'o}$, then the generalized eigenvalue equation would become an ordinary matrix eigenvalue equation, which is less computationally expensive to solve than the generalized equation. However, if $\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)$ is a true atomic orbital, that is, a one-electron wavefunction with symmetry o of an atom with atomic number Z_m, then, in general, $S_{o'm'om} \neq \delta_{m'm}\delta_{o'o}$. Since orbitals centered by the same atom are orthogonal to one another, $S_{o'mom} = \delta_{o'o}$ still holds. Also, if atoms m' and m are sufficiently far apart, then $S_{o'm'om} \approx 0$ because the orbitals are localized. If $\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)$ is a Löwdin orbital, that is, given a set of true atomic orbitals $\{\check{\phi}_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\}$, $\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)$ is such that

$$\begin{aligned}\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m) &= \sum_{o'}\sum_{m'} \check{S}^{-1/2}_{o'm'om}\check{\phi}_{o'Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'})\,,\\ &\approx \check{\phi}_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\\ &\quad -\frac{1}{2}\sum_{o'}\sum_{m'}(\check{S}_{o'm'om}-\delta_{o'o}\delta_{m'm})\check{\phi}_{o'Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'})\,,\end{aligned} \tag{57}$$

and

$$\check{S}_{o'm'om} = \int \check{\phi}_{o'Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'})\,\check{\phi}_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\mathrm{d}^3\mathbf{r}\,, \tag{58}$$

where a superscript of $-1/2$ indicates the inverse square root of a matrix, then it can be shown that $S_{o'm'om}$ becomes diagonal and equals $\delta_{o'o}\delta_{m'm}$ (*61*).

The tight-binding matrix equation for an infinite periodic system is similar to that for a finite system, but with $N_{\text{per cell}}$ replacing M, and

$$\sum_{m=1}^{N_{\text{per cell}}}\sum_{o} H_{o'm'om\mathbf{k}}C^{(i\mathbf{k})}_{om} = E_{i\mathbf{k}}\sum_{m=1}^{N_{\text{per cell}}}\sum_{o} S_{o'm'om\mathbf{k}}C^{(i\mathbf{k})}_{om}\,, \tag{59}$$

where

$$H_{o'm'om\mathbf{k}} = \int \phi_{o'\mathbf{k}m'}(\mathbf{r})\hat{H}_{1\mathrm{e}}\text{-}\phi_{o\mathbf{k}m}(\mathbf{r})\mathrm{d}^3\mathbf{r} \tag{60}$$

and

$$S_{o'm'om\mathbf{k}} = \int \phi_{o'\mathbf{k}m'}(\mathbf{r})\phi_{o\mathbf{k}m}(\mathbf{r})\mathrm{d}^3\mathbf{r}\,. \tag{61}$$

The integration here is over a unit cell of the periodic system. $S_{o'm'om\mathbf{k}}$ is an overlap matrix similar to the one for the finite system $S_{o'm'om}$. In general, $S_{o'momk} = \delta_{o'o}$, and if the functions $\phi_{oZ_m}(\mathbf{r})$ in the Bloch sum $\phi_{o\mathbf{k}m}(\mathbf{r})$ are orthogonalized rather than true atomic orbitals, then $S_{o'm'om\mathbf{k}} = \delta_{m'm}\delta_{o'o}$.

In the Slater-Koster tight-binding method, the Hamiltonian matrix elements, either $H_{o'm'om}$ or $H_{o'm'om\mathbf{k}}$, are decomposed into sums of integrals that are neglected, become empirical parameters, or become linear combinations of special two-center integrals that are functions only of the distance between their two centers. *These special integrals are generally not evaluated directly through integration, but rather are estimated through simple closed-form functions with empirical parameters*. The decomposition begins by decomposing the effective potential into a sum of per-atom contributions, such as

$$\hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_M) = \sum_{k=1}^{M} v_{Z_k}(\mathbf{r} - \mathbf{R}_k) \tag{62}$$

for a finite system with M atoms, or

$$\hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_\infty) = \sum_{k=1}^{N_{\text{per cell}}} \sum_{j=1}^{\infty} v_{Z_k}\left(\mathbf{r} - \overline{\mathbf{R}}_j - \mathbf{d}_k\right) \tag{63}$$

for an infinite periodic system. (At this point, spin-orbit coupling is ignored.) If $m \neq m'$, then after substituting equations 7 and 62 in equation 55, one may write the Hamiltonian matrix element as (*62*, *63*)

$$\begin{aligned} H_{o'm'om} = & \frac{1}{2}\int \phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{R}_{m'})\left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e} + v_{Z_m}(\mathbf{r} - \mathbf{R}_m)\right]\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)\mathrm{d}^3\mathbf{r} \\ & + \frac{1}{2}\int \phi_{oZ_m}(\mathbf{r} - \mathbf{R}_{m'})\left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e} + v_{Z_m}(\mathbf{r} - \mathbf{R}_m)\right]\phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{R}_m)\mathrm{d}^3\mathbf{r} \\ & + \int \phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{R}_{m'})\,\frac{v_{Z_{m'}}(\mathbf{r} - \mathbf{R}_{m'}) + v_{Z_m}(\mathbf{r} - \mathbf{R}_m)}{2}\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)\mathrm{d}^3\mathbf{r} \\ & + \sum_{k\neq m\neq m'}\int \phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{R}_{m'})\,v_{Z_k}(\mathbf{r} - \mathbf{R}_k)\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)\mathrm{d}^3\mathbf{r}\,. \end{aligned} \tag{64}$$

For the special case where $m = m'$, the Hamiltonian matrix element is instead

$$\begin{aligned} H_{o'mom} = & \int \phi_{o'Z_m}(\mathbf{r} - \mathbf{R}_m)\left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e} + v_{Z_m}(\mathbf{r} - \mathbf{R}_m)\right]\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)\mathrm{d}^3\mathbf{r} \\ & + \sum_{k\neq m}\int \phi_{o'Z_m}(\mathbf{r} - \mathbf{R}_m)\,v_{Z_k}(\mathbf{r} - \mathbf{R}_k)\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)\mathrm{d}^3\mathbf{r}\,. \end{aligned} \tag{65}$$

At this point, one may begin simplifying. First, the three-center integrals, that is, the integrals containing three factors, each centered around a different atomic site (i.e., m', k, and m), are treated as if they were negligible (*54*). This removes the integrals in the summation over $k \neq m \neq m'$ in equation 64. The integrals whose potential is centered at site k, while the orbitals are centered around m, are neglected much as the three-center integrals are (*64*), which removes the integrals in the summation over $k \neq m$ in equation 65. This leaves the two-center integrals, whose factors are centered around m' or m, and the onsite integrals, all of whose factors are centered around m. As Slater and Koster themselves pointed out, the integrals that are neglected here are not necessarily negligible in comparison to the two-center integrals, but they are smaller than the two-center integrals, and their neglect will reduce the number of fitting parameters needed (*54*). Another simplification comes from noting that an isolated atom at $\mathbf{R}_m$ with the effective potential $v_{Z_m}(\mathbf{r}-\mathbf{R}_m)$ has the following one-electron Schrödinger equation:

$$\left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e} + v_{Z_m}(\mathbf{r}-\mathbf{R}_m)\right]\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m) = E_{oZ_m}\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\,, \tag{66}$$

where E_{oZ_m} is the eigenvalue corresponding to the orbital symmetry o for an atom with atomic number Z_m. If one denotes the two-center integral containing the average of two atomic potentials as

$$V_{o'm'om} = \int \phi_{o'Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'})\,\frac{v_{Z_{m'}}(\mathbf{r}-\mathbf{R}_{m'}) + v_{Z_m}(\mathbf{r}-\mathbf{R}_m)}{2}\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_m)\mathrm{d}^3\mathbf{r}\,, \tag{67}$$

then, given the simplifications, the Hamiltonian matrix elements in equations 64 and 65 become

$$\begin{aligned} H_{o'm'om} &\approx \frac{E_{oZ_m}+E_{oZ_{m'}}}{2}S_{o'm'om} + V_{o'm'om}\,. \\ H_{o'mom} &\approx E_{oZ_m}\delta_{o'o} \end{aligned} \tag{68}$$

The first expression indicates the Hamiltonian matrix element for two arbitrary orbitals o' and o for two different atoms m' and m, respectively. The second expresses the matrix element for the case where two different orbitals o' and o are centered about the same atom m. The atomic eigenvalues are fitting parameters. Thus, the energies must be provided from another source, generally empirical data or through more accurate electronic structure *ab initio* approach. However, $V_{o'm'om}$ is not a fitting parameter. Rather, provided that atoms m' and m are not so far apart as to make it negligible, $V_{o'm'om}$ is decomposed further into sums of special two-center integrals that depend only on the distance $R_{mm'}$ between atomic sites m and m'. To do this, a change of coordinates $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{R}_{m'}$ is applied to $V_{o'm'om}$ to make explicit that it depends on $\mathbf{R}_{mm'} = \mathbf{R}_m - \mathbf{R}_{m'}$, that is,

$$V_{o'm'om} = \int \phi_{o'Z_{m'}}(\mathbf{r})\,\frac{v_{Z_{m'}}(\mathbf{r})+v_{Z_m}(\mathbf{r}-\mathbf{R}_{mm'})}{2}\phi_{oZ_m}(\mathbf{r}-\mathbf{R}_{mm'})\mathrm{d}^3\mathbf{r}\,. \tag{69}$$

For brevity, the average of the two potentials at sites m' and m will be denoted as

$$\frac{v_{Z_{m'}}(\mathbf{r}) + v_{Z_m}(\mathbf{r} - \mathbf{R}_{mm'})}{2} = v_{\text{avg}}^{m'm}(\mathbf{r})\,. \tag{70}$$

One of the special two-center integrals centered around the two nearest neighboring sites m and m' whose orbitals both have s symmetry is denoted here as $V_{sm',sm,\sigma}$,

$$V_{sm',sm,\sigma} = \int \phi_{s,Z_{m'}}(\mathbf{r})\; v_{\text{avg}}^{m'm}(\mathbf{r})\phi_{s,Z_m}(\mathbf{r} - \mathbf{R}_{mm'})\mathrm{d}^3\mathbf{r}\,. \tag{71}$$

Here, because of the symmetry of the s-orbitals, the direction of $\mathbf{R}_{mm'}$ does not matter; only its magnitude matters. Another special two-center integral involves the overlap of s- and p-orbitals, as shown in figure 12. These two-center integrals for $\mathbf{R}_{mm'}$ along the x-, y-, and z-directions are

$$V_{sm',pm,\sigma} = \int \phi_{sZ_{m'}}(\mathbf{r})v_{\text{avg}}^{m'm}(\mathbf{r})\phi_{p_xZ_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x)\mathrm{d}^3\mathbf{r}\;,$$

$$= \int \phi_{sZ_{m'}}(\mathbf{r})v_{\text{avg}}^{m'm}(\mathbf{r})\phi_{p_yZ_m}\big(\mathbf{r} - R_{mm'}\mathbf{e}_y\big)\mathrm{d}^3\mathbf{r}\;,$$

and

$$= \int \phi_{sZ_{m'}}(\mathbf{r})v_{\text{avg}}^{m'm}(\mathbf{r})\phi_{p_zZ_m}(\mathbf{r} - R_{mm'}\mathbf{e}_z)\mathrm{d}^3\mathbf{r}\,. \tag{72}$$

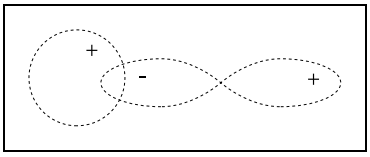

Figure 12. Overlapping between s- and p-orbitals. The plus and minus signs indicate where an orbital has a positive or negative value.

Other special two-center integrals involve pairs of overlapping p-orbitals. Figure 13 shows the two types of p-orbital overlap. Suppose that there are two p_x-orbitals with σ-type overlap, and the vector connecting their centers points along the x-direction, that is, $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}_x$, where $\mathbf{e}_x$ is a unit vector. Alternatively, suppose that the two orbitals with σ overlap are of p_y-type with $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}_y$, or of p_z-type with $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}_z$, where $\mathbf{e}_y$ and $\mathbf{e}_z$ are unit vectors pointing along the y- and z-directions. By symmetry, the two-center integral in all of these cases has the same value (*55*):

$$V_{pm',pm,\sigma} = \int \phi_{p_x Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_x Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r}\,,$$
$$= \int \phi_{p_y Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_y Z_m}\big(\mathbf{r} - R_{mm'}\mathbf{e}_y\big) \mathrm{d}^3\mathbf{r}\,, \tag{73}$$

and

$$= \int \phi_{p_z Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_z Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_z) \mathrm{d}^3\mathbf{r}\,.$$

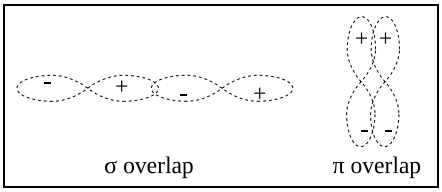

Figure 13. Types of overlap between p-orbitals. The plus and minus signs indicate where an orbital has a positive or negative value (*55*).

If the two p-orbitals have π-type overlap and $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}_x$, the orbitals must both be either p_y or p_z, that is,

$$V_{pm',pm,\pi} = \int \phi_{p_y Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_y Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r}\,,$$

and

$$= \int \phi_{p_z Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_z Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r}\,. \tag{74}$$

Similarly, if $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}_y$, then the orbitals in $V_{pm',pm,\pi}$ must both be of p_x or p_z type, and if $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}_z$, then the orbitals in $V_{pm',pm,\pi}$ must both be of p_x or p_y type.

Many two-center integrals turn out to be zero. For example,

$$\left.\begin{aligned} &\int \phi_{sZ_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_y Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r} \\ &\int \phi_{sZ_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_z Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r} \end{aligned}\right\} = 0\,. \tag{75}$$

The reason can be seen in figure 14. The interaction between the s-orbital and the positive lobe of the p-orbital is exactly canceled by a corresponding interaction between the s-orbital and the p-orbital's negative lobe. For similar reasons, the two-center integrals associated with the kinds of interaction between p-orbitals shown in figure 15 are also zero (*65*), e.g.,

$$\left.\begin{aligned} &\int \phi_{p_x Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_y Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r} \\ &\int \phi_{p_y Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_z Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x) \mathrm{d}^3\mathbf{r} \end{aligned}\right\} = 0\,. \tag{76}$$

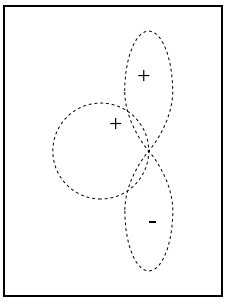

Figure 14. s- and p-orbitals whose net overlap is zero. The plus and minus signs indicate where an orbital has a positive or negative value. The overlap contribution from the positive lobe of the p-orbital is exactly canceled by that in the negative lobe.

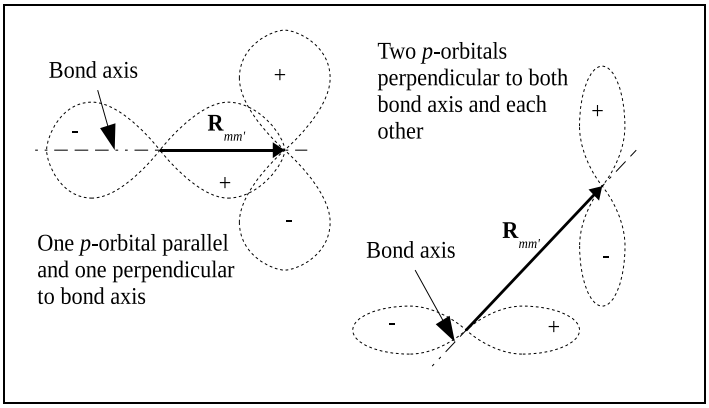

Figure 15. Pairs of p-orbitals whose net overlap is zero. The plus and minus signs indicate where an orbital has a positive or negative value. The bond axis runs through the points connected by the vector $\mathbf{R}_{mm'}$.

With the special two-center integrals $V_{sm',pm,\sigma}$, $V_{pm',pm,\sigma}$, and $V_{pm',pm,\pi}$ identified, the decomposition of a general two-center integral $V_{o'm'om}$, where $o', o \in \{s, p_x, p_y, p_z\}$, can be shown (*54*, *55*). Let there be a coordinate system $x'y'z'$ such that $\mathbf{R}_{mm'} = R_{mm'}\mathbf{e}'_x$, and the transformation between from unprimed to primed coordinates is (*66*)

$$x = (\mathbf{e}_x \cdot \mathbf{e}'_x)x' + (\mathbf{e}_x \cdot \mathbf{e}'_y)y' + (\mathbf{e}_x \cdot \mathbf{e}'_z)z',$$
$$y = (\mathbf{e}_y \cdot \mathbf{e}'_x)x' + (\mathbf{e}_y \cdot \mathbf{e}'_y)y' + (\mathbf{e}_y \cdot \mathbf{e}'_z)z',$$

and

$$z = (\mathbf{e}_z \cdot \mathbf{e}'_x)x' + (\mathbf{e}_z \cdot \mathbf{e}'_y)y' + (\mathbf{e}_z \cdot \mathbf{e}'_z)z'. \quad (77)$$

The p-orbitals can be factored as

$$\phi_{p_x Z_m}(\mathbf{r}) = R_{Z_m}(|\mathbf{r}|)x\,,$$
$$\phi_{p_y Z_m}(\mathbf{r}) = R_{Z_m}(|\mathbf{r}|)y\,,$$

and

$$\phi_{p_z Z_m}(\mathbf{r}) = R_{Z_m}(|\mathbf{r}|)z\,, \tag{78}$$

where $R_{Z_m}(|\mathbf{r}|)$ is a function dependent on the magnitude of $\mathbf{r}$ (*11*). A p_x-orbital, then, can be expressed in $x'y'z'$ coordinates as

$$\phi_{p_x Z_m}(\boldsymbol{r}) = R_{Z_m}(|\boldsymbol{r}|)\left[(\boldsymbol{e}_x \cdot \boldsymbol{e}'_x)x' + \left(\boldsymbol{e}_x \cdot \boldsymbol{e}'_y\right)y' + (\boldsymbol{e}_x \cdot \boldsymbol{e}'_z)z'\right], \tag{79}$$
$$\phi_{p_x Z_m}(\boldsymbol{r}) = R_{Z_m}(|\boldsymbol{r}|)\left[(\boldsymbol{e}_x \cdot \boldsymbol{e}'_x)x' + \left(\boldsymbol{e}_x \cdot \boldsymbol{e}'_y\right)y' + (\boldsymbol{e}_x \cdot \boldsymbol{e}'_z)z'\right],$$

and

$$= \phi_{p'_x Z_m}(\mathbf{r})(\mathbf{e}_x \cdot \mathbf{e}'_x) + \phi_{p'_y Z_m}(\mathbf{r})\left(\mathbf{e}_x \cdot \mathbf{e}'_y\right) + \phi_{p'_z Z_m}(\mathbf{r})(\mathbf{e}_x \cdot \mathbf{e}'_z)\,, \tag{79}$$

and similarly for the p_y- and p_z-orbitals. Accordingly, a general two-center integral between an s-orbital and a p_x-orbital is

$$\begin{aligned} V_{sm'p_x m} &= \int \phi_{sZ_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_x Z_m}(\mathbf{r} - \mathbf{R}_{mm'})\mathrm{d}^3\mathbf{r} \\ &= \int \phi_{sZ_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p'_x Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}'_x)\mathrm{d}^3\mathbf{r}\,(\mathbf{e}_x \cdot \mathbf{e}'_x) \\ &\quad + \int \phi_{sZ_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p'_y Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}'_x)\mathrm{d}^3\mathbf{r}\left(\mathbf{e}_x \cdot \mathbf{e}'_y\right) \\ &\quad + \int \phi_s Z_{m'}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p'_z Z_m}(\mathbf{r} - R_{mm'}\mathbf{e}'_x)\mathrm{d}^3\mathbf{r}\,(\mathbf{e}_x \cdot \mathbf{e}'_z) \\ &= V_{sm',pm,\sigma}(\mathbf{e}_x \cdot \mathbf{e}'_x)\,. \end{aligned} \tag{80}$$

A general two-center integral between two p-orbitals can be expressed as a linear combination of $V_{pm',pm,\sigma}$ and $V_{pm',pm,\pi}$. For example,

$$V_{p_x m' p_x m} = \int \phi_{p_x Z_{m'}}(\mathbf{r}) v_{\text{avg}}^{m'm}(\mathbf{r}) \phi_{p_x Z_m}(\mathbf{r} - \mathbf{R}_{mm'})\mathrm{d}^3\mathbf{r}\,,$$

and

$$= V_{pm',pm,\sigma}\,(\mathbf{e}_x \cdot \mathbf{e}'_x)^2 + V_{pm',pm,\pi}\left[\left(\mathbf{e}_x \cdot \mathbf{e}'_y\right)^2 + (\mathbf{e}_x \cdot \mathbf{e}'_z)^2\right]. \tag{81}$$

The directions of the y' and z' axes are arbitrary except for the constraint that they must be orthogonal to x' and each other, and, accordingly, the previous two-center integral can be

rewritten to indicate that it does not depend on $\mathbf{e}'_y$ and $\mathbf{e}'_z$ but rather on the direction cosines of $\mathbf{R}_{mm'}$, $l_R = (\mathbf{e}_x \cdot \mathbf{e}'_x)$, $m_R = (\mathbf{e}_y \cdot \mathbf{e}'_x)$, and $n_R = (\mathbf{e}_z \cdot \mathbf{e}'_x)$. Because of the orthogonality of primed and unprimed coordinates, one may write (*66*)

$$(\mathbf{e}_x \cdot \mathbf{e}'_y)^2 + (\mathbf{e}_x \cdot \mathbf{e}'_z)^2 = 1 - (\mathbf{e}_x \cdot \mathbf{e}'_x)^2 = 1 - l_R^2 \,. \tag{82}$$

Expressions of general two-center integrals in terms of l_R, m_R, n_R, and the special two-center integrals have been tabulated in the classic paper of Slater and Koster (*54*). Implementations of the tight-binding method can determine the decomposition of general two-center integrals through a look-up table based on this tabulation.

The elements of the overlap matrix $S_{o'm'om}$ in equation 56 are similar in form to those of the general two-center integral $V_{o'm'om}$ of equation 69, except that $[v_{Z_{m'}}(\mathbf{r}) + v_{Z_m}(\mathbf{r} - \mathbf{R}_{mm'})]/2$ is replaced by 1. Therefore, it is possible to decompose $S_{o'm'om}$ into a linear combination of integrals analogous to $V_{sm',sm,\sigma}$, $V_{sm',pm,\sigma}$, $V_{pm',pm,\sigma}$, $V_{pm',pm,\pi}$, etc. That is, one may define (*19*, *67*)

$$S_{sm',sm,\sigma} = \int \phi_{s,Z_{m'}}(\mathbf{r})\ \phi_{s,Z_m}(\mathbf{r} - \mathbf{R}_{mm'})\mathrm{d}^3\mathbf{r} \,, \tag{83}$$

and

$$\begin{aligned} S_{sm',pm,\sigma} &= \int \phi_{sZ_{m'}}(\mathbf{r})\phi_{p_xZ_m}(\mathbf{r} - R_{mm'}\mathbf{e}_x)\mathrm{d}^3\mathbf{r} \,, \\ &= \int \phi_{sZ_{m'}}(\mathbf{r})\phi_{p_yZ_m}(\mathbf{r} - R_{mm'}\mathbf{e}_y)\mathrm{d}^3\mathbf{r} \,, \\ &= \int \phi_{sZ_{m'}}(\mathbf{r})\phi_{p_yZ_m}(\mathbf{r} - R_{mm'}\mathbf{e}_y)\mathrm{d}^3\mathbf{r} \,, \end{aligned} \tag{84}$$

and so on. Integrals $S_{pm',pm,\sigma}$ and $S_{pm',pm,\pi}$ may be defined by replacing $v_{\mathrm{avg}}^{m'm}(\mathbf{r})$ with 1 in the two-center integrals $V_{pm',pm,\sigma}$ and $V_{pm',pm,\pi}$. Once these integrals are defined, decompositions such as the following may be done:

$$S_{sm'p_xm} = S_{sm',pm,\sigma}(\mathbf{e}_x \cdot \mathbf{e}'_x) = l_R S_{sm',pm,\sigma} \,, \tag{85}$$

and

$$S_{p_xm'p_xm} = S_{pm',pm,\sigma}\, l_R^2 + S_{pm',pm,\pi}(1 - l_R^2) \,. \tag{86}$$

These example decompositions are analogous to the decompositions of $V_{sm'p_xm}$ and $V_{p_xm'p_xm}$ in equations 80 and 81, respectively. As pointed out before, if atoms m and m' are sufficiently far apart, then the localization of the orbitals $\phi_{o'Z_{m'}}(\mathbf{r})$ and $\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_{mm'})$ entails that $S_{o'm'om} \approx 0$. This applies to the integrals $S_{sm',pm,\sigma}$, $S_{pm',pm,\sigma}$, etc., as well, and the localization of the orbitals also entails that the two-center integrals $V_{sm',sm,\sigma}$, $V_{sm',pm,\sigma}$, etc. decay as $R_{mm'} \to \infty$.

Accounting for spin-orbit coupling means that the number of orbital types is doubled, e.g., instead of $o', o \in \{s, p_x, p_y, \ldots\}$, one has $o', o \in \{s\uparrow, p_x\uparrow, p_y\uparrow, \ldots, s\downarrow p_x\downarrow, p_y\downarrow, \ldots\}$, where again the symbols "↑" and "↓" denote the spin-up and spin-down states of an electron. Let

$o = \bar{o}S$, where $\bar{o} \in \{s, p_x, p_y, \ldots\}$, and $S \in \{\uparrow, \downarrow\}$. An element of a tight-binding Hamiltonian matrix with the coupling, then, may be expressed as (*68*)

$$H_{o'mom} = H_{\bar{o}'m\bar{o}m}\delta_{S'S} + H^{\mathrm{SO}}_{o'mom}\,, \tag{87}$$

where $H_{\bar{o}'m\bar{o}m}$ is the corresponding Hamiltonian matrix element without spin-orbit correction. In the scheme of Chadi (*69*), the correction term $H^{\mathrm{SO}}_{o'mom}$ is zero except for the following matrix elements (*70*, *71*):

$$H^{\mathrm{SO}}_{p_x\downarrow mp_y\downarrow m} = \left(H^{\mathrm{SO}}_{p_y\downarrow mp_x\downarrow m}\right)^{\dagger} = -H^{\mathrm{SO}}_{p_x\uparrow mp_y\uparrow m} = -\left(H^{\mathrm{SO}}_{p_y\uparrow mp_x\uparrow m}\right)^{\dagger} = \mathrm{i}\lambda_m\,,$$

$$H^{\mathrm{SO}}_{p_z\downarrow mp_x\uparrow m} = \left(H^{\mathrm{SO}}_{p_x\uparrow mp_z\downarrow m}\right)^{\dagger} = -H^{\mathrm{SO}}_{p_z\uparrow mp_x\downarrow m} = -\left(H^{\mathrm{SO}}_{p_x\downarrow mp_z\uparrow m}\right)^{\dagger} = \lambda_m\,,$$

and

$$H^{\mathrm{SO}}_{p_y\downarrow mp_z\uparrow m} = \left(H^{\mathrm{SO}}_{p_z\uparrow mp_y\downarrow m}\right)^{\dagger} = H^{\mathrm{SO}}_{p_y\uparrow mp_z\downarrow m} = \left(H^{\mathrm{SO}}_{p_z\downarrow mp_x\uparrow m}\right)^{\dagger} = -\mathrm{i}\lambda_m\,, \tag{88}$$

where λ_m is an empirical parameter.

Figure 16 shows pseudocode for the tight-binding method applied to finite system. It is assumed in this algorithm that the means of empirically estimating the special two-center integrals $V_{sm',sm,\sigma}$, $V_{sm',pm,\sigma}$, $V_{pm',pm,\sigma}$, $V_{pm',pm,\pi}$, etc., is already available. So that the algorithm shown would fit on a single page, an optimization was left out that would be used in a more realistic code, which would take advantage of the fact that both $V_{o'm'om}$ and $S_{o'm'om}$ decay to zero for large $R_{mm'}$. If atoms m' and m are sufficiently far apart (e.g., not nearest or, possibly, next-nearest neighbors), then $H_{\bar{o}'m\bar{o}m}$ is negligible and can be immediately set to zero without an explicit determination of $V_{o'm'om}$ and $S_{o'm'om}$, which would be approximately zero.

The orbitals s, p_x, p_y, etc. for the basis set are given. The number of orbitals in the set is N_{orb}.
The number of atoms N_{atoms} is given.
Values of two-center integrals $V_{sm',sm,\sigma}$, $V_{pm',pm,\sigma}$, $V_{pm',pm,\pi}$, etc. are given.
Boolean variable *noSpin* is true if there is no spin, and false otherwise. $H^{\text{SO}}_{o'm'om}$ is given.

if *noSpin* **then**
  Let matrices $[H_{IJ}]$ and $[S_{IJ}]$ be of size $N_{\text{orb}}N_{\text{atoms}} \times N_{\text{orb}}N_{\text{atoms}}$.
else
  Let matrices $[H_{IJ}]$ and $[S_{IJ}]$ be of size $2N_{\text{orb}}N_{\text{atoms}} \times 2N_{\text{orb}}N_{\text{atoms}}$.
end if

Let $i = 0$.
for $\bar{o}'$ in $s, p_x, p_y, \ldots$ **do**
  for $m' = 1$ to N_{atoms} **do**
    Increment i by 1 and set $j = 0$.

    for $\bar{o}$ in $s, p_x, p_y, \ldots$ **do**
      for $m = 1$ to N_{atoms} **do**
        Increment j by 1.

        Estimate $S_{\bar{o}'m'\bar{o}m}$.
        if $m' = m$ **then**
          Let $H_{\bar{o}'m'\bar{o}m} = E_{oZ_m}\delta_{\bar{o}'\bar{o}}$
        else
          Calculate $V_{\bar{o}'m'\bar{o}m}$ as a linear combination of $V_{sm',sm,\sigma}, V_{pm',pm,\sigma}$, etc.
          Let $H_{\bar{o}'m'\bar{o}m} = \frac{1}{2}(E_{\bar{o}Z_m} + E_{\bar{o}Z'_m})S_{\bar{o}'m'\bar{o}m} + V_{\bar{o}'m'\bar{o}m}$.
        end if

        if *noSpin* **then**
          Let $I = i$, $J = j$, $H_{IJ} = H_{\bar{o}'m'\bar{o}m}$, and $S_{IJ} = S_{\bar{o}'m'\bar{o}m}$.
        else
          for S' in $\uparrow, \downarrow$ **do**
            for S in $\uparrow, \downarrow$ **do**
              Let $I = 2i - 1$ if $S' = \uparrow$, and let $I = 2i$ otherwise.
              Let $J = 2j - 1$ if $S = \uparrow$, and let $J = 2j$ otherwise.
              Let $o' = \bar{o}'S'$ and $o = \bar{o}S$.

              Let $H_{IJ} = H_{\bar{o}'m'\bar{o}m}\delta_{SS'} + H^{\text{SO}}_{o'm'om}$ and $S_{IJ} = S_{\bar{o}'m'\bar{o}m}\delta_{SS'}$
            end for
          end for
        end if

      end for
    end for
  end for
end for

Solve generalized eigenvalue problem $[H_{IJ}]\mathbf{C}^i = E_i[S_{IJ}]\mathbf{C}^i$.

Figure 16. Tight-binding method for a finite system.

Like $H_{o'm'om}$, $H_{o'm'om\mathbf{k}}$ can be broken down into terms containing two-center integrals such as $V_{sm',sm,\sigma}$, $V_{pm',pm,\sigma}$, $V_{pm',pm,\pi}$, and $V_{sm',pm,\sigma}$. One begins by expressing it explicitly as a double Bloch sum (*54*):

$$H_{o'm'om\mathbf{k}} = \lim_{N_{\text{cell}}\to\infty} \frac{1}{N_{\text{cell}}} \sum_{j=1}^{N_{\text{cell}}} \sum_{j'=1}^{N_{\text{cell}}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\left(\overline{\mathbf{R}}_j + \mathbf{d}_m - \overline{\mathbf{R}}_{j'} - \mathbf{d}_{m'}\right)} \times \int \phi_{o'Z_{m'}}\left(\mathbf{r} - \overline{\mathbf{R}}_{j'} - \mathbf{d}_{m'}\right) \widehat{H}_{1\mathrm{e}^-} \phi_{oZ_m}\left(\mathbf{r} - \overline{\mathbf{R}}_j - \mathbf{d}_m\right) \mathrm{d}^3\mathbf{r}\,,$$

and

$$= \sum_{j=1}^{\infty} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\left(\overline{\mathbf{R}}_j + \mathbf{d}_m - \mathbf{d}_{m'}\right)} \int \phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{d}_{m'}) \widehat{H}_{1\mathrm{e}^-} \phi_{oZ_m}\left(\mathbf{r} - \overline{\mathbf{R}}_j - \mathbf{d}_m\right) \mathrm{d}^3\mathbf{r}\,. \tag{89}$$

The integral in equation 89 is essentially identical to the integral that defines $H_{o'm'om}$ in equation 55, with $\mathbf{R}_{m'} = \mathbf{d}_{m'}$ and $\mathbf{R}_m = \overline{\mathbf{R}}_j + \mathbf{d}_m$, and the process for decomposing this integral into two- and three-center integrals is the same as that for $H_{o'm'om}$. As with finite systems, all of the three-center integrals are neglected. Most of the two-center integrals are as well, especially the ones for which $\left|\overline{\mathbf{R}}_j + \mathbf{d}_m - \mathbf{d}_{m'}\right| \gg 0$. This reduces the infinite summation in $H_{o'm'om\mathbf{k}}$ to a finite one. For example, in a crystal with the zincblende structure, the vectors connecting atom $m = 1$ to its four nearest neighbors are

$$\begin{aligned}
\overline{\mathbf{R}}_1 + \mathbf{d}_1 - \mathbf{d}_2 &= a\left(+\mathbf{e}_x + \mathbf{e}_y + \mathbf{e}_z\right)/4\,, \\
\overline{\mathbf{R}}_2 + \mathbf{d}_1 - \mathbf{d}_2 &= a\left(+\mathbf{e}_x - \mathbf{e}_y - \mathbf{e}_z\right)/4, \\
\overline{\mathbf{R}}_3 + \mathbf{d}_1 - \mathbf{d}_2 &= a\left(-\mathbf{e}_x + \mathbf{e}_y - \mathbf{e}_z\right)/4,
\end{aligned}$$

and

$$\overline{\mathbf{R}}_4 + \mathbf{d}_1 - \mathbf{d}_2 = a\left(-\mathbf{e}_x - \mathbf{e}_y + \mathbf{e}_z\right)/4, \tag{90}$$

where a is the lattice constant of the crystal. If $S_{s1s2} = 0$, then if $\mathbf{k} = k_x\mathbf{e}_x + k_y\mathbf{e}_y + k_z\mathbf{e}_z$ and all but the nearest neighbors are neglected (*55*, *72*):

$$H_{s1s2\mathbf{k}} = V_{s1,s2,\sigma}\left[\mathrm{e}^{\mathrm{i}(k_x + k_y + k_z)a/4} + \mathrm{e}^{\mathrm{i}(k_x - k_y - k_z)a/4} + \mathrm{e}^{\mathrm{i}(-k_x + k_y - k_z)a/4} + \mathrm{e}^{\mathrm{i}(-k_x - k_y + k_z)a/4}\right]. \tag{91}$$

Also, if only nearest neighbors are taken into account, diagonal elements of $H_{o'm'om\mathbf{k}}$ are the atomic eigenvalues for the orbital on atom m, i.e., $H_{omom\mathbf{k}} = E_{oZ_m}$. The overlap matrix $S_{o'm'om\mathbf{k}}$ is determined much as $H_{o'm'om\mathbf{k}}$, i.e.,

$$S_{o'm'om\mathbf{k}} = \lim_{N_{\text{cell}}\to\infty} \frac{1}{N_{\text{cell}}} \sum_{j=1}^{N_{\text{cell}}} \sum_{j'=1}^{N_{\text{cell}}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\left(\bar{\mathbf{R}}_j + \mathbf{d}_m - \bar{\mathbf{R}}_{j'} - \mathbf{d}_{m'}\right)} \times \int \phi_{o'Z_{m'}}\left(\mathbf{r} - \bar{\mathbf{R}}_{j'} - \mathbf{d}_{m'}\right) \phi_{oZ_m}\left(\mathbf{r} - \bar{\mathbf{R}}_j - \mathbf{d}_m\right) \mathrm{d}^3\mathbf{r} \ ,$$

and

$$= \sum_{j=1}^{\infty} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\left(\bar{\mathbf{R}}_j + \mathbf{d}_m - \mathbf{d}_{m'}\right)} \int \phi_{o'Z_{m'}}(\mathbf{r} - \mathbf{d}_{m'}) \phi_{oZ_m}\left(\mathbf{r} - \bar{\mathbf{R}}_j - \mathbf{d}_m\right) \mathrm{d}^3\mathbf{r} \, . \tag{92}$$

The integral in equation 92 is essentially identical to the integral that defines $S_{o'm'om}$ in equation 56, with $\mathbf{R}_{m'} = \mathbf{d}_{m'}$ and $\mathbf{R}_m = \bar{\mathbf{R}}_j + \mathbf{d}_m$. Figure 17 shows a pseudocode implementation of the tight-binding method for an infinite crystal.

The spin-orbitals $s\uparrow, s\downarrow p_x\uparrow, p_x\downarrow$, etc. for the basis set are given. The number of spin-orbitals in the set is N_{orb}.
The number of atoms per unit cell N_{atoms} is given.
Values of two-center integrals $V_{sm',sm,\sigma}$, $V_{pm',pm,\sigma}$, $V_{pm',pm,\pi}$, etc. are given.
Wavevectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_{\mathbf{k}}}$ are given.

for $\mathbf{k}$ in $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_{\mathbf{k}}}$ **do**
  Let matrices $[H_{IJ}]$ and $[S_{IJ}]$ be of size $N_{\text{orb}}N_{\text{atoms}} \times N_{\text{orb}}N_{\text{atoms}}$.

  Let $I = 0$.
  for o' in $s\uparrow, s\downarrow, p_x\uparrow, p_x\downarrow, \ldots$ **do**
    for $m' = 1$ to N_{atoms} **do**
      Increment I by 1 and set $J = 0$.

      for o in $s\uparrow, s\downarrow, p_x\uparrow, p_x\downarrow, \ldots$ **do**
        for $m = 1$ to N_{atoms} **do**
          Increment J by 1.

          Let $H_{o'm'om\mathbf{k}} = 0$ and $S_{o'm'om\mathbf{k}} = 0$.

          for $\bar{\mathbf{R}}_j$ in $\bar{\mathbf{R}}_1, \bar{\mathbf{R}}_2, \ldots, \bar{\mathbf{R}}_{\max}$ **do** {where $|\bar{\mathbf{R}}_1| \le |\bar{\mathbf{R}}_2| \le \ldots \le \bar{\mathbf{R}}_{\max}$}
            Let atom M be the atom in the crystal located at $\bar{\mathbf{R}}_j + \mathbf{d}_m$.
            Determine $H_{o'm'oM}$ and $S_{o'm'oM}$ as one would for a finite system.

            Add $e^{i\mathbf{k}\cdot(\bar{\mathbf{R}}_j+\mathbf{d}_m-\mathbf{d}_{m'})}H_{o'm'oM}$ to $H_{o'm'om\mathbf{k}}$.
            Add $e^{i\mathbf{k}\cdot(\bar{\mathbf{R}}_j+\mathbf{d}_m-\mathbf{d}_{m'})}S_{o'm'oM}$ to $S_{o'm'om\mathbf{k}}$.
          end for

          Let $H_{IJ} = H_{o'm'om\mathbf{k}}$, and $S_{IJ} = S_{o'm'om\mathbf{k}}$.

        end for
      end for
    end for
  end for

  Solve generalized eigenvalue problem $[H_{IJ}]\mathbf{C}^{i\mathbf{k}} = E_{i\mathbf{k}}[S_{IJ}]\mathbf{C}^{i\mathbf{k}}$.
end for

Figure 17. Tight-binding method for an infinite crystal.

Here are examples of how the special two-center integrals $V_{sm',sm,\sigma}$, $V_{pm',pm,\sigma}$, etc., may be estimated through closed-form empirical formulas. One such formula (*56*) is

$$V_{o'm',om,\lambda} = \left(\frac{R^{(0)}_{mm'}}{R_{mm'}}\right)^{\eta_{o'o\lambda}} V^{(0)}_{o'm',om,\lambda}\,, \tag{93}$$

where $o \in \{s, p, \ldots\}$, $\lambda \in \{\sigma, \pi\}$, $R^{(0)}_{mm'}$ is the distance between atoms m' and m in the absence of strain, $V^{(0)}_{o'm',om,\lambda}$ is the strain-free value of $V_{o'm',om,\lambda}$, that is, the value of $V_{o'm',om,\lambda}$ evaluated at

$R^{(0)}_{mm'}$, and $\eta_{o'o\lambda}$ is an empirical parameter that is typically in the range of 1 to 4. $V^{(0)}_{o'm',om,\lambda}$ is an empirical parameter itself. Another formula, from Mehl and Papaconstantopoulos (*19*), is

$$V_{o'm',om,\lambda} = \left(e_{o'o\lambda} + f_{o'o\lambda}R_{mm'} + \bar{f}_{o'o\lambda}R^2_{mm'}\right)\mathrm{e}^{-g^2_{o'o\lambda}R_{mm'}}F_c(R_{mm'}) \,, \tag{94}$$

where $e_{o'o\lambda}$, $f_{o'o\lambda}$, $\bar{f}_{o'o\lambda}$, and $g_{o'o\lambda}$ are fitting parameters, and $F_c(R_{mm'})$ is a cutoff function,

$$F_c(R_{mm'}) = \frac{1}{1 + \exp\left[\left(R_{mm'} - R^{(0)}_{mm'}\right)/L\right]} . \tag{95}$$

Mehl and Papaconstantopoulos typically take L to be half the Bohr radius.

Off-diagonal elements of the overlap matrix may be determined through closed-form empirical formulas as well. For example, Papaconstantopoulos et al. (*67*) formulate this dependence as follows:

$$\begin{aligned} S_{o'm',om,\lambda} = {} & \left(\delta_{o'o} + a_{o'o\lambda}R_{mm'} + b_{o'o\lambda}R^2_{mm'} + c_{o'o\lambda}R^3_{mm'}\right) \\ & \times \mathrm{e}^{-d^2_{o'o\lambda}R_{mm'}}F_c(R_{mm'}) \,, \end{aligned} \tag{96}$$

where $a_{o'o\lambda}$. $b_{o'o\lambda}$, $c_{o'o\lambda}$, and $d_{o'o\lambda}$ are fitting parameters. The Kronecker delta $\delta_{o'o}$ ensures that the preceding expression is consistent with the condition $S_{o'mom} = \delta_{o'o}$ (where $m' = m$), which the overlap matrix must satisfy.

Such formulas are a part of how the tight-binding method takes strain into account, since they approximately account for changes in interatomic distance due to strain. Strain also leads to changes in the orientations of atoms relative to one another (i.e., the directions of $\mathbf{R}_{mm'}$ for pairs of atoms m' and m), and this is taken into account through the direction cosines l_R, m_R, and n_R used in the decomposition of $V_{o'm'om}$ into linear combinations of $V_{sm',sm,\sigma}$, $V_{sm',pm,\sigma}$, etc. (and also into any decomposition of $S_{o'm'om}$ into linear combinations of $S_{sm',pm,\sigma}$, $S_{pm',pm,\sigma}$, etc.).

It can be shown that the diagonal elements of $H_{o'm'om}$ or $H_{o'm'om\mathbf{k}}$ also depend on the strain. According to Boykin et al. (*33*), if $\phi_{oZ_m}(\mathbf{r} - \mathbf{R}_m)$ is a Löwdin orbital as defined in equation 57, then

$$E_{oZ_m} = E^{(0)}_{oZ_m} + \sum_{o'}\sum_{m'} C_{o'm'om}\frac{\left(V^{(0)}_{o'm'om}\right)^2 - \left(V_{o'm'om}\right)^2}{E^{(0)}_{oZ_m} + E^{(0)}_{o'Z_{m'}}} \,, \tag{97}$$

where the superscript "(0)" again denotes the strain-free version of a quantity, and $C_{o'm'om}$ is an empirical parameter. This is not the only scheme for accounting for the effects of strain on diagonal elements. For example, Mehl and Papaconstantopoulos (*19*) take the diagonal elements to be

$$H_{omom} = a_{oZ_m} + b_{oZ_m}\rho_m^{2/3} + c_{oZ_m}\rho_m^{4/3} + d_{oZ_m}\rho_m^2\,, \tag{98}$$

where a_{oZ_m}, b_{oZ_m}, c_{oZ_m}, $and\ d_{oZ_m}$ are fitting parameters, and

$$\rho_m = \sum_{m'} \exp\left(-\lambda_{Z_m Z_{m'}} R_{mm'}\right) F_c(R_{mm'})\,. \tag{99}$$

The summation with index m' is over the neighbors of atom m, and $\lambda_{Z_m Z_{m'}}$ is a fitting parameter. (Mehl and Papaconstantopoulos (*19*) fit their parameters to results from *ab initio* calculations.) Other methods for accounting for strain on diagonal elements have been discussed by Jancu and Voisin (*73*) and Niquet et al. (*74*).

The Slater-Koster tight-binding method has been used on systems with about a million atoms. For example, it has been used on a dome-shaped $In_{0.6}Ga_{0.4}As$ quantum dot embedded in a GaAs matrix. The dot was 30 nm in diameter and 5 nm in height, and the simulation domain was 40 × 40 × 15 nm. The dot itself contained about 718,000 atoms, with the GaAs matrix constituting the remainder of the million atoms in the simulation. The distribution of the In and Ga cations within the dot was random. For each cation (Ga or In) and anion (As), the set of basis functions contained three p-orbitals (p_x, p_y, and p_z) and two orbitals with s-symmetry, denoted s and s^*. Since spin was taken into account, the number of basis functions per atom doubled from 5 to 10. For a given distribution of Ga and In atoms within the dot, determination of the energy eigenvalues of the system took about 25 min on 31 processors of a cluster of Pentium III 933-MHz CPUs. If the corresponding eigenvectors had been computed, the computation time would have doubled (*56*). As seen from the pseudocode of the tight-binding implementations shown in figures 5 and 6, the computational resources needed for the method scales with the product of the number of orbitals in the basis set and the number of atoms in the system. Whether it scales linearly or as $O(N^3)$ depends on the choice of eigenvalue solver (*21*, *46*).

3.3 How Non-Self-Consistency Affects How Strain Is Taken Into Account

The atomistic methods previously described are not self-consistent, that is, their approximations of the effective potential of the one-electron Schrödinger equation do not depend on the charge density or the one-electron wavefunctions. Computationally, this is a great advantage, since it means that the Schrödinger equation does not have to be solved iteratively. However, as will be explained shortly, this entails a restriction on the physics that the effective potential $\hat{V}_{\text{ext,eff}}$ may take into account. For a system of M atoms (where M may be infinite), $\hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_M)$ may be written generally as

$$\hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_M) = -\frac{1}{4\pi\epsilon_0}\sum_{i=1}^{M}\frac{Z_i q^2}{|\mathbf{r}-\mathbf{R}_i|} + \frac{q^2}{4\pi\epsilon_0}\int\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3\mathbf{r}' + \hat{V}_{\text{other}}\,, \tag{100}$$

where the operator $\hat{V}_{\text{other}}$ is determined by one's choice of *ab initio* method. This can be contrasted with the decompositions of $\hat{V}_{\text{ext,eff}}$ into a sum of per-atom contributions in equation 31

for the empirical pseudopotential method and equations 62 and 63 for the Slater-Koster tight-binding method. Despite the contrasting details in the earlier descriptions of these two methods, we remark that they both follow similar algorithmic and fundamental theoretical structures. Namely, in the generalized effective potential in equation 100, only the first term, the contribution from the M nuclei of the system, naturally decomposes into a sum of per-atom contributions. One could decompose the second and third terms in $\hat{V}_{\text{ext,eff}}$ by assuming the following:

$$\rho(\mathbf{r}) = \sum_{i=1}^{M} \rho_i(\mathbf{r}) \ , \tag{101}$$

and

$$\hat{V}_{\text{other}} = \textstyle\sum_{i=1}^{M} \hat{v}_{i,\text{other}} \, , \tag{102}$$

such that

$$\hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_M) = \left\{ \sum_{i=1}^{M} \frac{1}{4\pi\epsilon_0} \left[\frac{Z_i q^2}{|\mathbf{r} - \mathbf{R}_i|} + q^2 \int \frac{\rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}^3\mathbf{r}' \right] + \hat{v}_{i,\text{other}} \right\}. \tag{103}$$

Barring an accounting of some nonlocal effects such as spin-orbit coupling (e.g., Chelikowsky and Cohen, [*38*]), this is largely what the atomistic methods described are effectively doing. Furthermore, in these methods, $\rho_i(\mathbf{r})$ is generally dependent on the species of atom i (i.e., whether it is Si, Ga, nitrogen [N], etc.) rather than its location in the system $\mathbf{R}_i$. This is a problem, for example, in piezoelectric systems, where an electric field due to strain may lead to the transfer of charge from atoms near one region of the system to atoms near another region, even if the atoms in each region are of the same types.

This problem can be overcome by treating this field as if it were external and adding the potential due to this field, $qV^{\text{stat}}(\mathbf{r}, \boldsymbol{\epsilon})$, where $\boldsymbol{\epsilon}$ is the strain if the field is induced from piezoelectricity, to the effective potential *(34*, *75–78*), that is,

$$\begin{aligned} \hat{V}_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_M) = & \sum_{i=1}^{M} \frac{1}{4\pi\epsilon_0} \left[\frac{Z_i q^2}{|\mathbf{r} - \mathbf{R}_i|} + q^2 \int \frac{\rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}^3\mathbf{r}' \right] \\ & + \sum_{i=1}^{M} \hat{v}_{i,\text{other}} + qV^{\text{stat}}(\mathbf{r}, \boldsymbol{\epsilon}) \, . \end{aligned} \tag{104}$$

$V^{\text{stat}}(\mathbf{r}, \boldsymbol{\epsilon})$ can be determined from a continuum mechanical calculation (*75*, *79*, *80*), provided that the strain $\boldsymbol{\epsilon}$ in the system has already been determined from the changes in atomic positions. Possible methods to determine this strain are in appendix B of this report.

4. The $k \cdot p$ and Envelope Function Methods

While the atomistic empirical methods are far less computationally expensive than the *ab initio* methods, their calculation time still scales with the number of atoms in the simulated system. This is not an issue for the $k \cdot p$ and envelope function methods, which do not take atomic positions as input at all. The $k \cdot p$ method (*81*) was originally developed to estimate the electronic band structure of bulk periodic semiconductor crystals, especially the band structure near the Γ point. The envelope function approximation, however, extends the method to systems that are not periodic in all dimensions, such as quantum wells, wires, and dots. In this extension, the effective potential within a given material is still assumed to behave approximately as the effective potential would in the bulk form of the material. Instead of taking strain into account through atomic positions, it is taken into account through an operator added to the effective Hamiltonian, an operator that takes the small strain tensor from continuum mechanics as input. These approximations mean that the envelope function approximation is essentially a continuum approach rather than an atomistic one.

4.1 General Formulation of the $k \cdot p$ Method for Bulk Crystal

The $k \cdot p$ method for bulk crystal employs the Bloch theorem, so, accordingly, the one-electron wavefunction is expressed as $\psi_{i\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} u_{i\mathbf{k}}(\mathbf{r})$, where $u_{i\mathbf{k}}(\mathbf{r})$ has the same periodicity as the crystal. Substituting this expression for $\psi_{i\mathbf{k}}(\mathbf{r})$ into the one-electron Schrödinger, equation 8 yields (*81*)

$$\left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e} + \hat{V}_{\mathrm{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_\infty) + \frac{\hbar}{m_e}\mathbf{k}\cdot\hat{\mathbf{p}} + \frac{\hbar^2 k^2}{2m_e}\right] u_{i\mathbf{k}}(\mathbf{r}) = E_{i\mathbf{k}} u_{i\mathbf{k}}(\mathbf{r}) \,, \tag{105}$$

where $k^2 = |\mathbf{k}|^2$, and $\{\mathbf{R}\}_\infty$ is the set of coordinates and atomic numbers of the nuclei of the infinite bulk crystal. Alternately, equation 105 may be written as

$$\hat{H}_{1\mathrm{e}^-}^{k\cdot p} u_{i\mathbf{k}}(\mathbf{r}) = \left[\hat{H}_{1\mathrm{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) + \frac{\hbar}{m_e}\mathbf{k}\cdot\hat{\mathbf{p}} + \frac{\hbar^2 k^2}{2m_e}\right] u_{i\mathbf{k}}(\mathbf{r}) = E_{i\mathbf{k}} u_{i\mathbf{k}}(\mathbf{r}) \,, \tag{106}$$

where $\hat{H}_{1\mathrm{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty)$ is the one-electron Hamiltonian for an infinite bulk crystal. For small $\mathbf{k}$, the operator $\hat{H}_{1\mathrm{e}^-}^{k\cdot p}$ is "close" to $\hat{H}_{1\mathrm{e}^-}$, so the eigenstates of $\hat{H}_{1\mathrm{e}^-}^{k\cdot p}$ may be expanded in terms of the eigenstates of $\hat{H}_{1\mathrm{e}^-}$ at $\mathbf{k} = 0$:

$$u_{i\mathbf{k}}(\mathbf{r}) = \sum_m c_m^{(i\mathbf{k})} u_{m0}(\mathbf{r}) \,. \tag{107}$$

These eigenstates are orthogonal, that is,

$$\int_{\Omega_c} u_{m'0}(\mathbf{r}) u_{m0}(\mathbf{r}) \mathrm{d}^3\mathbf{r} = \delta_{m'm}\,, \tag{108}$$

where Ω_c is the volume of the primitive unit cell of the crystal. Alternatively, $u_{i\mathbf{k}}(\mathbf{r})$ may be expanded not directly in terms of the eigenstates at the σ point but rather in terms of normalized orthogonal linear combinations of these eigenstates, i.e.,

$$u_{i\mathbf{k}}(\mathbf{r}) = \sum_m C_m^{(i\mathbf{k})} U_m(\mathbf{r})\,,\ U_i(\mathbf{r}) = \sum_m c_m^i u_{m0}(\mathbf{r})\,. \tag{109}$$

If one substitutes equation 109 into equation 106, multiplies the latter equation by $U_{m'}^{\dagger}(\mathbf{r})$, and integrates over a unit cell of the crystal, then one may obtain a matrix equation

$$\sum_m H_{m'm}^{k\cdot p}(\{\mathbf{R}\}_\infty) C_m^{(i\mathbf{k})} = E_{i\mathbf{k}} C_{m'}^{(i\mathbf{k})}, \tag{110}$$

where

$$H_{m'm}^{k\cdot p}(\{\mathbf{R}\}_\infty) = H_{m'm}^{1\mathrm{e}^-}(\{\mathbf{R}\}_\infty) + \frac{\hbar}{m_e}\mathbf{k}\cdot\mathbf{p}_{m'm} + \delta_{m'm}\frac{\hbar^2k^2}{2m_e}\,, \tag{111}$$

$$H_{m'm}^{1\mathrm{e}^-}(\{\mathbf{R}\}_\infty) = \int_{\Omega_c} U_{m'}^{\dagger}(\mathbf{r}) \widehat{H}_{1\mathrm{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) U_m(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \tag{112}$$

and

$$\mathbf{p}_{m'm} = \int_{\Omega_c} U_{m'}^{\dagger}(\mathbf{r}) \widehat{\mathbf{p}} U_m(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,. \tag{113}$$

Alternately, $\hbar^2k^2/2m_e$ may be moved to the right-hand side of the matrix equation, so that

$$\sum_m H_{m'm}^{k\cdot p*}(\{\mathbf{R}\}_\infty) C_m^{(i\mathbf{k})} = \left(E_{i\mathbf{k}} - \frac{\hbar^2k^2}{2m_e}\right) C_{m'}^{(i\mathbf{k})}\,, \tag{114}$$

where

$$H_{m'm}^{k\cdot p*}(\{\mathbf{R}\}_\infty) = H_{m'm}^{1\mathrm{e}^-} + \frac{\hbar}{m_e}\mathbf{k}\cdot\mathbf{p}_{m'm}\,. \tag{115}$$

$U_i(\mathbf{r})$ is often chosen such that $\widehat{H}_{1\mathrm{e}^-}U_i(\mathbf{r}) = E_{i0}U_i(\mathbf{r})$, so that $H_{m'm}^{1\mathrm{e}^-} = E_{m0}\delta_{m'm}$ (*32, 81–83*).

If the effect of spin-orbit is taken into account, one may then replace $\widehat{\mathbf{p}}$ in equations 106 and 113 with $\widehat{\boldsymbol{\pi}}$, where (*32*, *82*)

$$\widehat{\boldsymbol{\pi}} = \widehat{\mathbf{p}} + \frac{\hbar}{4m_ec^2}\left[\widehat{\boldsymbol{\sigma}} \times \nabla V_{\mathrm{ext,eff}}\right]. \tag{116}$$

When $\mathbf{k}$ is small, $\mathbf{k}\cdot\widehat{\boldsymbol{\pi}} \approx \mathbf{k}\cdot\widehat{\mathbf{p}}$. (*83*).

4.2 Example $k \cdot p$ Formulations for Bulk Crystals

In principle, the summation in the expansion of $u_{i\mathbf{k}}(\mathbf{r})$ in equations 107 and 109 is infinite and, accordingly, so is the matrix $H_{m'm}^{k\cdot p}$ (or $H_{m'm}^{k\cdot p*}$). In practice, $H_{m'm}^{k\cdot p}$ is transformed into a finite $k \cdot p$ Hamiltonian matrix, usually one with fairly small dimensions, e.g., 6 × 6 or 8 × 8. Two examples of how this may be done, the Kane and Luttinger-Kohn formulations, are shown in the following equations. In the formulation by Kane (*83*), the transformation is done simply by truncating the expansion of $u_{i\mathbf{k}}(\mathbf{r})$, including only the terms corresponding to the one-electron wavefunctions for the three valence bands and the lowest conduction band of a semiconductor with a diamond or zincblende crystal structure. When spin is taken into account, consideration of these four bands leads to an eight-term expansion of $u_{i\mathbf{k}}(\mathbf{r})$—four terms for spin-up and four for spin-down—and thus an 8 × 8 matrix. In the Luttinger-Kohn formulation (*82*), the expansion of $u_{i\mathbf{k}}(\mathbf{r})$ is infinite in principle, but the six terms pertaining to the three valence bands are assumed to be the dominant ones in the expansion, and the infinite matrix $H_{m'm}^{k\cdot p*}$ is accordingly transformed into a 6 × 6 matrix where the effects of the nondominant terms in the expansion of $u_{i\mathbf{k}}(\mathbf{r})$ are indirectly taken into account through material constants called the Luttinger parameters.

Kane (*83*) studied the band structure of indium antimonide (InSb), whose electronic band structure resembles the schematic shown in figure 18. Δ is the difference in energy between the maximum of the split-off valence band and the common maximum of the other two valence bands, labeled as "heavy hole" and "light hole." Those labels refer to the effective masses (discussed in detail in section 4.4) of the electrons or holes with energies in the corresponding bands. Holes are vacancies left behind when an electron is promoted from a valence to the conduction band, and they act like positive charge carriers. The expansion in equation 109 is truncated to include only the eight states at $\mathbf{k} = 0$ corresponding to the bands shown in the figure. The two states corresponding to the conduction band minimum, one for spin-down and one for spin-up, have the same radial symmetry as atomic s-orbitals, and these states are

$$U_1(\mathbf{r}) = \mathrm{i}S(\mathbf{r}) \downarrow, \text{ and } U_5(\mathbf{r}) = \mathrm{i}S(\mathbf{r}) \uparrow, \tag{117}$$

where

$$\uparrow = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{and } \downarrow = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{118}$$

and $S(\mathbf{r})$ is spherically symmetric and thus depends only on the magnitude of $\mathbf{r}$. Accordingly, it has the following parity properties:

$$S\big((r_1, r_2, r_3)\big) = S\big((|r_1|, |r_2|, |r_3|)\big). \tag{119}$$

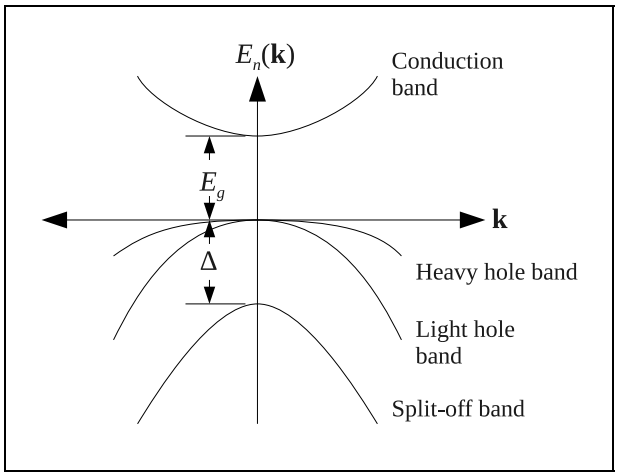

Figure 18. Schematic band structure near $\mathbf{k} = 0$ of a typical semiconductor with a diamond or zincblende crystal structure. E_g is the band gap energy. Δ is the difference in energy between the maximum of the split-off band and the common maximum of heavy and light hole bands. The left and right halves of the horizontal axis indicate the magnitude of $\mathbf{k}$-values pointing along certain crystal directions. (In this qualitative schematic, the actual directions are not important.) While the heavy hole, light hole, and split-off bands are typical for such a semiconductor and generally have their maxima at $\mathbf{k} = 0$ as shown, the actual minimum of the conduction band may be different from what is shown in this schematic (*17*).

The six valence states at $\mathbf{k} = 0$, which have the same symmetries as atomic p-orbitals, are

$$U_2(\mathbf{r}) = \frac{X(\mathbf{r}) - \mathrm{i}Y(\mathbf{r})}{\sqrt{2}} \uparrow,\ U_6(\mathbf{r}) = -\frac{X(\mathbf{r}) + \mathrm{i}Y(\mathbf{r})}{\sqrt{2}} \downarrow, \tag{120}$$

$$U_3(\mathbf{r}) = Z(\mathbf{r}) \downarrow,\ U_7(\mathbf{r}) = Z(\mathbf{r}) \uparrow, \tag{121}$$

$$U_4(\mathbf{r}) = \frac{X(\mathbf{r}) + \mathrm{i}Y(\mathbf{r})}{\sqrt{2}} \uparrow,\ U_8(\mathbf{r}) = \frac{X(\mathbf{r}) - \mathrm{i}Y(\mathbf{r})}{\sqrt{2}} \downarrow, \tag{122}$$

where the functions X, Y, and Z have the following parity properties:

$$X\big((r_1, r_2, r_3)\big) = \mathrm{sgn}(r_1) X\big((|r_1|, |r_2|, |r_3|)\big), \tag{123}$$

$$Y\big((r_1, r_2, r_3)\big) = \mathrm{sgn}(r_2) Y\big((|r_1|, |r_2|, |r_3|)\big), \tag{124}$$

and

$$Z\big((r_1, r_2, r_3)\big) = \mathrm{sgn}(r_3) Z\big((|r_1|, |r_2|, |r_3|)\big), \tag{125}$$

where $\mathrm{sgn}(x)$ is the sign of x, i.e., $\mathrm{sgn}(\pm x) = \pm 1$. Kane assumed that $\mathbf{k} \cdot \hat{\boldsymbol{\pi}} \approx \mathbf{k} \cdot \hat{\mathbf{p}}$, since $\mathbf{k}$ is taken to be small. Kane initially restricted the Bloch wave vector $\mathbf{k}$ to points along the z-direction (i.e., the r_3 axis) so that the matrix $H_{m'm}^{k \cdot p*}$ would simplify to the following block diagonal format:

$$\left[H_{m'm}^{k\cdot p*}\right] = \mathbf{H}_{\text{Kane}} = \begin{bmatrix} \mathbf{H}_{\text{Kane},4\times4} & 0 \\ 0 & \mathbf{H}_{\text{Kane},4\times4} \end{bmatrix}, \tag{126}$$

where

$$\mathbf{H}_{\text{Kane},4\times4} = \begin{bmatrix} E_s & 0 & kP & 0 \\ 0 & E_p - \Delta^\circ/3 & \sqrt{2}\Delta^\circ 3 & 0 \\ kP & \sqrt{2}\Delta^\circ 3 & E_p & 0 \\ 0 & 0 & 0 & E_p + \Delta^\circ/3 \end{bmatrix}, \tag{127}$$

$$E_s = \int_{\Omega_c} S^\dagger(\mathbf{r}) \widehat{H}_{1\text{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) S(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \tag{128}$$

$$\begin{aligned} E_p &= \int_{\Omega_c} X^\dagger(\mathbf{r}) \widehat{H}_{1\text{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) X(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \\ &= \int_{\Omega_c} Y^\dagger(\mathbf{r}) \widehat{H}_{1\text{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) Y(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \\ &= \int_{\Omega_c} Z^\dagger(\mathbf{r}) \widehat{H}_{1\text{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty) Z(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \end{aligned} \tag{129}$$

$$P_{\text{Kane}} = -\mathrm{i}\frac{\hbar}{m_e} \int_{\Omega_c} S^\dagger(\mathbf{r})\, \hat{p}_3 Z(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \tag{130}$$

and

$$\Delta^\circ = \frac{3h\mathrm{i}}{4m_e^2c^2} \int_{\Omega_c} X^\dagger(\mathbf{r}) \left[\frac{\partial V_{\text{ext,eff}}}{\partial r_1}\hat{p}_2 - \frac{\partial V_{\text{ext,eff}}}{\partial r_2}\hat{p}_1\right] Y(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,, \tag{131}$$

where $\hat{p}_i$ is a component of vector operator $\widehat{\mathbf{p}}$. P_{Kane} is called Kane's parameter (*32*). Here, $\mathbf{H}_{\text{Kane},4\times4}$ is written under the assumption that equation 114, the $k \cdot p$ matrix equation with the term $-\hbar^2k^2/2m_e$ on the right side, is being used. If $\mathbf{k}$ points in a general direction, a transformation can be applied to rotate the original coordinate system so that $\mathbf{k}$ points along the z-direction, which makes $\mathbf{H}_{\text{Kane}}$ block diagonal. Since $S(\mathbf{r})$ is spherically symmetric, it is unaffected by this rotation. The rotation causes $X(\mathbf{r})$, $Y(\mathbf{r})$, $Z(\mathbf{r})$, $\uparrow$, and $\downarrow$ to be replaced by $X'(\mathbf{r})$, $Y'(\mathbf{r})$, $Z'(\mathbf{r})$, $\uparrow'$, and $\downarrow'$ in equations 117, 120–122, 128, and 131. If $\mathbf{k}$ is expressed in the original coordinate system in spherical coordinates (k, θ, ϕ), the primed functions are, in terms of the original $X(\mathbf{r})$, $Y(\mathbf{r})$, and $Z(\mathbf{r})$,

$$X'(\mathbf{r}) = X(\mathbf{r}) \cos\theta \cos\phi + Y(\mathbf{r}) \cos\theta \sin\phi - Z(\mathbf{r}) \sin\theta\,,$$

$$Y'(\mathbf{r}) = -X(\mathbf{r}) \sin\phi + Y(\mathbf{r}) \cos\phi\,,$$

and

$$Z'(\mathbf{r}) = X(\mathbf{r}) \sin\theta \cos\phi + Y(\mathbf{r}) \sin\theta \sin\phi + Z(\mathbf{r}) \cos\theta\,. \tag{132}$$

The primed spinors are

$$\uparrow' = \begin{bmatrix} e^{-i\phi/2}\cos(\theta/2) \\ e^{i\phi/2}\sin(\theta/2) \end{bmatrix},$$

and

$$\downarrow' = \begin{bmatrix} -e^{-i\phi/2}\sin(\theta/2) \\ e^{i\phi/2}\cos(\theta/2) \end{bmatrix}. \tag{133}$$

$\mathbf{H}_{\mathrm{Kane},4\times4}$ is itself in a block diagonal format, with one block containing the first three rows and columns and the other block being a 1×1 "matrix" with $E_p + \Delta°/3$ as its sole element. Thus, one may define two eigenproblems, one for each block, as follows:

$$\mathbf{H}_{\mathrm{Kane},3\times3}\mathbf{C}^{(i\mathbf{k})} = \left(E_{i\mathbf{k}} - \frac{\hbar^2 k^2}{2m_e}\right)\mathbf{C}^{(i\mathbf{k})}, \tag{134}$$

where

$$\mathbf{H}_{\mathrm{Kane},3\times3} = \begin{bmatrix} E_s & 0 & kP \\ 0 & E_p - \Delta°/3 & \sqrt{2}\Delta°3 \\ kP & \sqrt{2}\Delta°3 & E_p \end{bmatrix}, \tag{135}$$

and

$$\left(E_p + \Delta°/3\right)C_4^{(i\mathbf{k})} = \left(E_{i\mathbf{k}} - \frac{\hbar^2 k^2}{2m_e}\right)C_4^{(i\mathbf{k})}. \tag{136}$$

In order for the highest valence band extremum to be zero, E_p is set to $-\Delta°/3$. This leads to the following solution for $E_{i\mathbf{k}}$:

$$E_{i\mathbf{k}} = E_{hh,\mathbf{k}} = \frac{\hbar^2}{2m_e}k^2. \tag{137}$$

This is the heavy hole band according to the Kane formulation. Note that this band curves upward as a function of k, which contradicts experiment (*32*, *83*) (and the schematic band structure in figure 18). This is a consequence of truncating the expansion of $u_{i\mathbf{k}}(\mathbf{r})$. Because of spin, this energy band is doubly degenerate, and the periodic parts of the two one-electron wavefunction that corresponds to this band are, for $\mathbf{k}$ pointing along a general direction,

$$u_{hh1\mathbf{k}}(\mathbf{r}) = U_4(\mathbf{r}) = \frac{X'(\mathbf{r}) + iY'(\mathbf{r})}{\sqrt{2}}\uparrow', \tag{138}$$

and

$$u_{hh2\mathbf{k}}(\mathbf{r}) = U_8(\mathbf{r}) = \frac{X'(\mathbf{r}) - iY'(\mathbf{r})}{\sqrt{2}}\downarrow',$$

where the primed quantities are defined in equations 132 and 133.

The eigenvalues of $\mathbf{H}_{\mathrm{Kane},3\times3}$ may be determined numerically. However, some insight into the physical meanings of E_s and $\Delta°$ may be obtained by estimating closed-form expressions from these eigenvalues, similar to the one for $E_{hh,\mathbf{k}}$. The characteristic polynomial of $\mathbf{H}_{\mathrm{Kane},3\times3}$ is

$$E'_{i\mathbf{k}}(E'_{i\mathbf{k}} - E_s)(E'_{i\mathbf{k}} + \Delta°) - k^2P^2(E'_{i\mathbf{k}} - 2\Delta°/3) = 0, \tag{139}$$

where $E'_{i\mathbf{k}} = E_{i\mathbf{k}} - \hbar^2 k^2/2m_e$, and the equation $E_p = -\Delta°/3$ has already been substituted into $\mathbf{H}_{\mathrm{Kane},3\times3}$. If $k = 0$, the solutions of the characteristic polynomial are $E'_{i0} = E_s, 0, -\Delta°$. These solutions are presumed to be extrema of the energy bands. For simplicity, these bands are also presumed to be parabolic in the neighborhood of $k = 0$. Accordingly, trial solutions for this polynomial for $k \neq 0$ are taken to be of the form $E'_{i\mathbf{k}} = E_s + e_i k^2, 0 + e_i k^2, -\Delta° + e_i k^2$, where e_i is a parameter to be determined (*32*). If these trial solutions are substituted into the characteristic polynomial and, then, any resulting terms containing powers of k higher than 2 are ignored, then for $i = c, lh, so$,

$$e_c = \frac{P_{\mathrm{Kane}}^2(E_s + 2\Delta°/3)}{E_s(E_s + \Delta°)}, \quad e_{lh} = -\frac{2P_{\mathrm{Kane}}^2}{3E_s}, \quad e_{so} = \frac{P_{\mathrm{Kane}}^2}{3(E_s + \Delta°)}, \tag{140}$$

and the energy bands are approximately (*32*, *83*)

$$E_{c\mathbf{k}} = E_s + \frac{\hbar^2}{2}\left[\frac{1}{m_e} + \frac{2P_{\mathrm{Kane}}^2(E_s + 2\Delta°/3)}{\hbar^2 E_s(E_s + \Delta°)}\right]k^2,$$

$$E_{lh,\mathbf{k}} = \frac{\hbar^2}{2}\left(\frac{1}{m_e} - \frac{4P_{\mathrm{Kane}}^2}{3\hbar^2 E_s}\right)k^2,$$

and

$$E_{so,\mathbf{k}} = -\Delta° + \frac{\hbar^2}{2}\left[\frac{1}{m_e} - \frac{2P_{\mathrm{Kane}}^2}{3\hbar^2(E_s + \Delta°)}\right]k^2, \tag{141}$$

where $E_{c\mathbf{k}}$, $E_{lh,\mathbf{k}}$, and $E_{so,\mathbf{k}}$ are the energies for a given $\mathbf{k}$ in the conduction, light hole, and split-off bands, respectively. By comparing these results to the qualitative band structure in figure 18, one can confirm that $E_s = E_g$ and $\Delta° = \Delta$. One may also note from equation 131 that if the spin-orbit coupling were nonexistent, parameter Δ would not exist either. E_g, Δ, and P_{Kane} may be determined empirically, and a compilation of these parameters, along with some discussion as to how they have been obtained, may be found in a survey by Vurgaftman et al. (*84*).

Because of spin, the energy bands $E_{c\mathbf{k}}$, $E_{lh,\mathbf{k}}$, and $E_{so,\mathbf{k}}$ are doubly degenerate. For $\mathbf{k}$ pointing along a general direction, the periodic parts of each one-electron wavefunction $u_{i\mathbf{k}}(\mathbf{r})$ are determined to be

$$u_{j1\mathbf{k}}(\mathbf{r}) = a_j U_1(\mathbf{r}) + b_j U_2(\mathbf{r}) + c_j U_3(\mathbf{r}),$$

$$= a_j \mathrm{i}S(\mathbf{r})\downarrow' + b_j \frac{X'(\mathbf{r}) - \mathrm{i}Y'(\mathbf{r})}{\sqrt{2}}\uparrow + c_j Z'(\mathbf{r})\downarrow',$$

$$u_{j2\mathbf{k}}(\mathbf{r}) = a_j U_5(\mathbf{r}) + b_j U_6(\mathbf{r}) + c_j U_7(\mathbf{r}),$$

and

$$a_j \mathrm{i}S(\mathbf{r})\uparrow' - b_j \frac{X'(\mathbf{r}) - \mathrm{i}Y'(\mathbf{r})}{\sqrt{2}}\downarrow' + c_j Z'(\mathbf{r})\uparrow', \tag{142}$$

where (a_j, b_j, c_j) is an eigenvector of $\mathbf{H}_{\text{Kane},3\times3}$, and $j = c, lh, so$. When $\mathbf{k} = (0,0,0)$, then

$$\begin{aligned} a_c &= 1, \quad b_c = 0, \qquad c_c = 0\,, \\ a_{lh} &= 0, \quad b_{lh} = \sqrt{1/3}\,, \quad c_{lh} = \sqrt{2/3}\,, \\ a_{so} &= 0, \quad b_{so} = \sqrt{2/3}\,, \quad c_{so} = -\sqrt{1/3}\,\cdot \end{aligned} \tag{143}$$

Pseudocode of the Kane formulation of the $k \cdot p$ method is shown in figure 19.

Parameters E_g, Δ and P_{Kane} are given.
Wavevectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ are given, where $\mathbf{k}_i$ is reasonably close to Γ point.

$\mathbf{H}_{\text{Kane},3\times3}(k, E_s, E_p, \Delta, P_{\text{Kane}})$ is given.

for $\mathbf{k}$ in $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ **do**
  Let $k = |\mathbf{k}|$. {$\mathbf{k}$ in spherical coordinates is (k, θ, ϕ).}

  Let $X'(\mathbf{r}) = X(\mathbf{r})\cos\theta\cos\phi + Y(\mathbf{r})\cos\theta\sin\phi - Z(\mathbf{r})\sin\theta$.
  Let $Y'(\mathbf{r}) = -X(\mathbf{r})\sin\phi + Y(\mathbf{r})\cos\phi$.
  Let $Z'(\mathbf{r}) = X(\mathbf{r})\sin\theta\cos\phi + Y(\mathbf{r})\sin\theta\sin\phi - Z(\mathbf{r})\cos\theta$.
  {$S(\mathbf{r})$ has spherical symmetry and isn't changed by a coordinate rotation.}

  Let $\uparrow' = \begin{bmatrix} e^{-i\phi/2}\cos(\theta/2) \\ e^{i\phi/2}\sin(\theta/2) \end{bmatrix}$. {Defining spin-up in spherical coordinates.}
  Let $\downarrow' = \begin{bmatrix} -e^{-i\phi/2}\sin(\theta/2) \\ e^{i\phi/2}\cos(\theta/2) \end{bmatrix}$. {Defining spin-down in spherical coordinates.}

  {For the conduction, light hole and split-off bands:}
  Let $\mathbf{H} = \mathbf{H}_{\text{Kane},3\times3}(k, E_g, -\Delta/3, \Delta, P_{\text{Kane}})$.
  Find eigenvalues $E'_{i\mathbf{k}}$ and eigenvectors $\mathbf{C}^{i\mathbf{k}}$ of $\mathbf{H}$.
  Let $a_i = C_1^{i\mathbf{k}}$, $b_i = C_2^{i\mathbf{k}}$, and $c_i = C_3^{i\mathbf{k}}$.

  for i in c, lh, so **do**
    Let $E_{i\mathbf{k}} = E'_{i\mathbf{k}} + \hbar^2 k^2 / 2m_e$.
    Let $u_{i1\mathbf{k}}(\mathbf{r}) = a_i \mathrm{i} S(\mathbf{r}) \downarrow' + b_i\{[X'(\mathbf{r}) - \mathrm{i}Y'(\mathbf{r})]/\sqrt{2}\} \uparrow' + c_i Z'(\mathbf{r}) \downarrow'$.
    Let $u_{i2\mathbf{k}}(\mathbf{r}) = a_i \mathrm{i} S(\mathbf{r}) \uparrow' - b_i\{[X'(\mathbf{r}) + \mathrm{i}Y'(\mathbf{r})]/\sqrt{2}\} \downarrow' + c_i Z'(\mathbf{r}) \uparrow'$.
  end for

  {For the heavy hole band:}
  $E_{hh\mathbf{k}} = \hbar^2 k^2 / 2m_e$.
  $u_{hh1\mathbf{k}}(\mathbf{r}) = \{[X'(\mathbf{r}) + \mathrm{i}Y'(\mathbf{r})]/\sqrt{2}\} \uparrow'$.
  $u_{hh1\mathbf{k}}(\mathbf{r}) = \{[X'(\mathbf{r}) - \mathrm{i}Y'(\mathbf{r})]/\sqrt{2}\} \downarrow'$.
end for

Figure 19. Pseudocode for Kane formulation of the $k \cdot p$ method.

The determination of the Hamiltonian matrix $\mathbf{H}_{\text{Kane}}$ does not require a detailed knowledge of S, X, Y, or Z; rather, only the parity properties shown in equations 119 and 123–125 are required. This causes most of the calculated matrix elements to be zero. The integrals that are nonzero become functions of empirical parameters. This is a common feature of $k \cdot p$ formulations.

While Kane (*83*) simply truncated the expansion in equation 107 to obtain a finite matrix for his $k \cdot p$ formulation, Luttinger and Kohn (*82*) took a different approach that, in principle, accounted for all the terms in the expansion. They broke the expansion into two sums:

$$u_{i\mathbf{k}}(\mathbf{r}) = \sum_{m}^{A} c_m^{(i\mathbf{k})} u_{m0}(\mathbf{r}) + \sum_{l}^{B} c_l^{(i\mathbf{k})} u_{l0}(\mathbf{r})\,. \tag{144}$$

The first summation is over the class *A* eigenstates; that is, the few states whose eigenenergies E_{m0} are near the bottom of the band gap. The second summation is over the class *B* eigenstates, which are the remaining states. Here, $U_m = u_{m0}$, so equations 107 and 109 are equivalent and $c_m^{(i\mathbf{k})} = C_m^{(i\mathbf{k})}$. The class *A* states may be expressed as (*32*)

$$\begin{aligned}
u_{10}(\mathbf{r}) &= -\frac{X(\mathbf{r}) + \mathrm{i}Y(\mathbf{r})}{\sqrt{2}}\begin{bmatrix}1\\0\end{bmatrix},\\
u_{20}(\mathbf{r}) &= -\left(\sqrt{\frac{1}{3}}\right)\frac{X(\mathbf{r}) + \mathrm{i}Y(\mathbf{r})}{\sqrt{2}}\begin{bmatrix}0\\1\end{bmatrix} + \sqrt{\frac{2}{3}}Z(\mathbf{r})\begin{bmatrix}1\\0\end{bmatrix},\\
u_{30}(\mathbf{r}) &= \left(\sqrt{\frac{1}{3}}\right)\frac{X(\mathbf{r}) - \mathrm{i}Y(\mathbf{r})}{\sqrt{2}}\begin{bmatrix}1\\0\end{bmatrix} + \sqrt{\frac{2}{3}}Z(\mathbf{r})\begin{bmatrix}0\\1\end{bmatrix},\\
u_{40}(\mathbf{r}) &= \frac{X(\mathbf{r}) - \mathrm{i}Y(\mathbf{r})}{\sqrt{2}}\begin{bmatrix}0\\1\end{bmatrix},\\
u_{50}(\mathbf{r}) &= \left(\sqrt{\frac{2}{3}}\right)\frac{X(\mathbf{r}) + \mathrm{i}Y(\mathbf{r})}{\sqrt{2}}\begin{bmatrix}0\\1\end{bmatrix} + \sqrt{\frac{1}{3}}Z(\mathbf{r})\begin{bmatrix}1\\0\end{bmatrix},
\end{aligned}$$

and

$$u_{60}(\mathbf{r}) = \left(\sqrt{\frac{2}{3}}\right)\frac{X(\mathbf{r}) - \mathrm{i}Y(\mathbf{r})}{\sqrt{2}}\begin{bmatrix}1\\0\end{bmatrix} - \sqrt{\frac{1}{3}}Z(\mathbf{r})\begin{bmatrix}0\\1\end{bmatrix}. \tag{145}$$

X, Y, and Z have the same meanings they do in the formulation by Kane (*83*) previously discussed. These states are also the periodic parts of the one-electron wavefunctions corresponding to the heavy hole, light hole, and split-off bands at $\mathbf{k} = 0$, just as in the Kane model. (See equations 138, 142, and 143 for comparison.) They have the same symmetry as the p-eigenstates of an atom where spin-orbit coupling is taken into account (*85*, pp. 99, 197). The states in equation 145 all belong to the valence bands. In the Luttinger-Kohn formulation, all conduction band states are within class *B*. Because of the parities of X, Y, and Z, $\mathbf{p}_{m'm} = 0$ for the class *A* states.

Equation 114 is rewritten in matrix form as

$$(\mathbf{H^0} + \mathbf{H^1})\mathbf{C}^{(i\mathbf{k})} = \left(E_{i\mathbf{k}} - \frac{\hbar^2 k^2}{2m_e}\right)\mathbf{C}^{(i\mathbf{k})}, \tag{146}$$

where the elements of $\mathbf{H^0}$ and $\mathbf{H^1}$ are such that

$$H^0_{m'm} = E_{m0}\delta_{m'm}\,,$$

and

$$H^1_{m'm} = \frac{\hbar}{m_e}\mathbf{k}\cdot\mathbf{p}_{m'm}. \tag{147}$$

Luttinger and Kohn (*82*), here, apply a canonical transformation, where

$$\mathbf{C}^{(i\mathbf{k})} = \mathrm{e}^{\mathbf{S}}\mathbf{B}^{(i\mathbf{k})}, \tag{148}$$

and then

$$\mathrm{e}^{-\mathbf{S}}(\mathbf{H^0} + \mathbf{H^1})\mathrm{e}^{\mathbf{S}}\mathbf{B}^{(i\mathbf{k})} = \left(E_{i\mathbf{k}} - \frac{\hbar^2 k^2}{2m_e}\right)\mathbf{B}^{(i\mathbf{k})}. \tag{149}$$

A new matrix, $\mathbf{H}^{\mathrm{reduced}}$, is defined as

$$\mathbf{H}^{\mathrm{reduced}} \equiv \mathrm{e}^{-\mathbf{S}}(\mathbf{H^0} + \mathbf{H^1})\mathrm{e}^{\mathbf{S}} + \frac{\hbar^2 k^2}{2m_e}\mathbf{I}\,, \tag{150}$$

where $\mathbf{I}$ is the identity matrix and

$$\mathbf{H}^{\mathrm{reduced}}\mathbf{B}^{(i\mathbf{k})} = E_{i\mathbf{k}}\mathbf{B}^{(i\mathbf{k})}. \tag{151}$$

The first term of $\mathbf{H}^{\mathrm{reduced}}$ may be expanded as

$$\begin{aligned}\mathrm{e}^{-\mathbf{S}}(\mathbf{H^0} + \mathbf{H^1})\mathrm{e}^{\mathbf{S}} = \mathbf{H^0} + \mathbf{H^1} + [\mathbf{H^0}, \mathbf{S}] + [\mathbf{H^1}, \mathbf{S}] \\ + \frac{1}{2}\big[[\mathbf{H^0}, \mathbf{S}], \mathbf{S}\big] + \frac{1}{2}\big[[\mathbf{H^0}, \mathbf{S}], \mathbf{S}\big] + \cdots,\end{aligned} \tag{152}$$

where the commutator $[\mathbf{A}_1, \mathbf{A}_2] = \mathbf{A}_1\mathbf{A}_2 - \mathbf{A}_2\mathbf{A}_1$. The matrix $\mathbf{S}$ is chosen so that

$$\mathbf{H^1} + [\mathbf{H^0}, \mathbf{S}] = 0\,, \tag{153}$$

and

$$\mathbf{H}^{\mathrm{reduced}} \approx \mathbf{H^0} + \tfrac{1}{2}[\mathbf{H^1}, \mathbf{S}] + \tfrac{\hbar^2 k^2}{2m_e}\mathbf{I}\,. \tag{154}$$

This removes coupling terms between classes *A* and *B* that are first order in $\mathbf{k}$. Given the definitions of $\mathbf{H^0}$ and $\mathbf{H^1}$, equation 153 becomes

$$\frac{\hbar}{m_e}\mathbf{k}\cdot\mathbf{p}_{m'm} + (E_{m'0} - E_{m0})S_{m'm} = 0\,, \tag{155}$$

and the elements of $\mathbf{H}^{\mathrm{reduced}}$ become (*81*), for indices m', m ranging from 1 to 6, the index values pertaining to class *A*,

$$H_{m'm}^{\text{reduced}} = \left(E_{m0} + \frac{\hbar^2 k^2}{2m_e}\right)\delta_{m'm} + \frac{\hbar^2}{2m_e}\sum_{l}^{B}\left[\frac{(\mathbf{k}\cdot\mathbf{p}_{m'l})(\mathbf{k}\cdot\mathbf{p}_{lm})}{E_{m0}-E_{l0}} + \frac{(\mathbf{k}\cdot\mathbf{p}_{m'l})(\mathbf{k}\cdot\mathbf{p}_{lm})}{E_{m'0}-E_{l0}}\right]. \quad (156)$$

The summation index l only ranges over the class B states because $\mathbf{p}_{m'm}$ vanishes for states in class A. If it is assumed that the states in class A all share at least approximately the same eigenvalue, i.e., $E_{m0} \approx E_{m'0} = E_0$, then the previous equation simplifies to (*32*)

$$H_{m'm}^{\text{LK}} \approx \left(E_{m0} + \frac{\hbar^2 k^2}{2m_e}\right)\delta_{m'm} + \frac{\hbar^2}{2m_e}\sum_{l}^{B}\frac{(\mathbf{k}\cdot\mathbf{p}_{m'l})(\mathbf{k}\cdot\mathbf{p}_{lm})}{E_0-E_{l0}}, \quad (157)$$

where $H_{m'm}^{\text{LK}}$ is called a Luttinger-Kohn Hamiltonian matrix, and

$$u_{i\mathbf{k}}(\mathbf{r}) \approx \sum_{m} B_m^{(i\mathbf{k})} u_{m0}(\mathbf{r})\,. \quad (158)$$

The summation in equation 157 containing matrix elements involving class B states is expressed in terms of the empirically determined Luttinger parameters γ_1, γ_2, and γ_3. These can be expressed as

$$\gamma_1 = -\frac{2m_e}{3\hbar^2}(A_0 + 2B_0)\,,$$
$$\gamma_2 = -\frac{m_e}{3\hbar^2}(A_0 - B_0)\,,$$
$$\gamma_3 = -\frac{m_e}{3\hbar^2}C_0\,, \quad (159)$$

where

$$A_0 = \frac{\hbar^2}{2m_e} + \frac{\hbar^2}{m_e^2}\sum_{l}^{B}\frac{p_{1l}^{(1)}p_{l1}^{(1)}}{E_0-E_{l0}}\,,$$
$$B_0 = \frac{\hbar^2}{2m_e} + \frac{\hbar^2}{m_e^2}\sum_{l}^{B}\frac{p_{1l}^{(2)}p_{l1}^{(2)}}{E_0-E_{l0}}\,,$$
$$C_0 = \sum_{l}^{B}\frac{p_{1l}^{(1)}p_{l2}^{(2)} + p_{2l}^{(2)}p_{l1}^{(1)}}{E_0-E_{l0}}\,, \quad (160)$$

and $p_{lm}^{(i)}$ is the i^{th} component of the vector $\mathbf{p}_{lm}$.

Chuang (*32*) has expressed a Luttinger-Kohn Hamiltonian matrix as follows:

$$\mathbf{H}^{\mathrm{LK}} = \begin{bmatrix} P+Q & -S & R & 0 & -S/\sqrt{2} & \sqrt{2}R \\ -S^{\dagger} & P-Q & 0 & R & -\sqrt{2}Q & \sqrt{3/2}\,S \\ R^{\dagger} & 0 & P-Q & S & \sqrt{3/2}\,S^{\dagger} & \sqrt{2}Q \\ 0 & R^{\dagger} & S^{\dagger} & P+Q & -\sqrt{2}R^{\dagger} & -S^{\dagger}/\sqrt{2} \\ -S^{\dagger}/\sqrt{2} & -\sqrt{2}Q^{\dagger} & \sqrt{3/2}\,S & -\sqrt{2}R & P-\Delta & 0 \\ \sqrt{2}R^{\dagger} & \sqrt{3/2}\,S^{\dagger} & \sqrt{2}Q^{\dagger} & -S/\sqrt{2} & 0 & P-\Delta \end{bmatrix}, \tag{161}$$

where

$$P = -\frac{\hbar^2\gamma_1}{2m_e}(k_1^2 + k_2^2 + k_3^2)\,,$$

$$Q = -\frac{\hbar^2\gamma_2}{2m_e}(k_1^2 + k_2^2 - 2k_3^2)\,,$$

$$R = -\frac{\hbar^2}{2m_e}\left(-\sqrt{3}\gamma_2(k_1^2 - k_2^2) + \mathrm{i}2\sqrt{3}\gamma_3 k_1 k_2\right),$$

and

$$S = -\frac{\hbar^2\gamma_3}{m_e}\sqrt{3}(k_1 - \mathrm{i}k_2)k_3\,. \tag{162}$$

Unlike the Kane formulation, the Luttinger-Kohn $k \cdot p$ formulation is written for Bloch wave vectors $\mathbf{k}$ of arbitrary direction, so the pseudocode for it, shown in figure 20, is trivially simple.

Luttinger parameters $\gamma_1, \gamma_2, \gamma_3$ are given.
Δ is given.
Wavevectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ are given.

Luttinger-Kohn Hamiltonian matrix $\mathbf{H}(\mathbf{k}, \gamma_1, \gamma_2, \gamma_3, \Delta)$ is given.

for $\mathbf{k}$ in $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ **do**
  Find eigenvalues $E_{i\mathbf{k}}$ and eigenvectors $\mathbf{B}^{i\mathbf{k}}$ of $\mathbf{H}(\mathbf{k}, \gamma_1, \gamma_2, \gamma_3, \Delta)$.
end for

Figure 20. Pseudocode for Luttinger-Kohn formulation of the $k \cdot p$ method.

4.3 The Envelope Function Approximation

If one substitutes the basis function expansion in equation 109 into the Bloch theorem expression for the one-electron wavefunction in equation 10, one obtains the following expression for the one-electron wavefunction:

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sum_m \left[C_m^{(i\mathbf{k})} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}\right] U_m(\mathbf{r}) = \sum_m f_m^{(i\mathbf{k})}(\mathbf{r}) U_m(\mathbf{r})\,. \tag{163}$$

Here, $\psi_{i\mathbf{k}}(\mathbf{r})$ is a sum of modulated periodic functions. The functions $U_m(\mathbf{r})$ are periodic functions with the same periodicity; that is, $U_m(\mathbf{r} + \overline{\mathbf{R}}) = U_m(\mathbf{r})$ where $\overline{\mathbf{R}}$ is a lattice vector of

the crystal. The functions $f_m^{(i\mathbf{k})}(\mathbf{r})$ that modulate the amplitudes of the functions $U_m(\mathbf{r})$ are envelope functions. In general, they do not share the periodicity of the $U_m(\mathbf{r})$, i.e., $f_m^{(i\mathbf{k})}(\mathbf{r}) \neq f_m^{(i\mathbf{k})}(\mathbf{r}+\overline{\mathbf{R}})$, and since $\mathbf{k}$ is restricted to the first Brillouin zone, $f_m^{(i\mathbf{k})}(\mathbf{r}) = C_m^{(i\mathbf{k})}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$ is a long-wavelength function that varies far more slowly than $U_m(\mathbf{r})$. The expansion for the one-electron wavefunction relies on Bloch theorem and thus only applies to systems that are periodic in all three spatial dimensions. The general idea behind the envelope function approximation (*86*, *87*), though, is that even for systems that lack periodicity along one or more directions, the one-electron wavefunction may still be approximately expressed as a sum of modulated periodic functions; that is,

$$\psi_\alpha(\mathbf{r}) = \sum_{m'} f_{m'}^{(\alpha)}(\mathbf{r}) U_{m'}(\mathbf{r})\,, \tag{164}$$

where α is a general label for the function. For a system that has two-dimensional periodicity, such as a system composed of a layer of one material sandwiched between two layers of a different material, the Bloch theorem still holds for planes normal to the layer, so for a normal direction along r_3, the envelope function is $f_m^{(\alpha)}(\mathbf{r}) = f_m^{(i\mathbf{k}_\parallel)}(\mathbf{r}) = f_m^{(i)}(r_3)\mathrm{e}^{\mathrm{i}\mathbf{k}_\parallel\cdot\mathbf{r}_\parallel}$ where $\mathbf{k}_\parallel = (k_1, k_2, 0)$ and $\mathbf{r}_\parallel = (r_1, r_2, 0)$ (*86*). For a wire pointing along r_3, $f_m^{(\alpha)}(\mathbf{r}) = f_m^{(ik_3)}(\mathbf{r}) = f_m^{(i)}(r_1, r_2)\mathrm{e}^{\mathrm{i}k_3 r_3}$ (*88*). Otherwise, there is no Bloch wave vector at all, and the label α is simply the index i. The envelope function is taken to be slowly varying in comparison to U_m. With the envelope function equation in place, a matrix differential equation analogous to the matrix equation 110 for the $k \cdot p$ method for bulk crystals may be obtained:

$$\sum_m \widehat{H}_{m'm}^{\mathrm{EFA,tot}}\left(\mathbf{r}, \{\mathbf{R}\}_{\mathrm{sys}}\right) f_m^{(\alpha)}(\mathbf{r}) \approx E_\alpha f_{m'}^{(\alpha)}(\mathbf{r})\,, \tag{165}$$

where

$$\widehat{H}_{m'm}^{\mathrm{EFA,tot}}\left(\mathbf{r}, \{\mathbf{R}\}_{\mathrm{sys}}\right) = \sum_l \Theta_l(\mathbf{r}) \widehat{H}_{m'm}^{\mathrm{EFA}}(\{\mathbf{R}\}_\infty^l)\,, \tag{166}$$

$$\widehat{H}_{m'm}^{\mathrm{EFA}}(\{\mathbf{R}\}_\infty^l) = H_{m'm}^{1\mathrm{e}^-}(\{\mathbf{R}\}_\infty^l) + \frac{\hbar}{m_e}\hat{\mathbf{k}}\cdot\mathbf{p}_{m'm} + \delta_{m'm}\frac{\hbar^2\hat{k}^2}{2m_e}\,, \tag{167}$$

and

$$H_{m'm}^{1\mathrm{e}^-}(\{\mathbf{R}\}_\infty^l) = \int_\Omega U_{m'}^\dagger(\mathbf{r}) \widehat{H}_{1\mathrm{e}^-}(\mathbf{r}, \{\mathbf{R}\}_\infty^l) U_m(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,. \tag{168}$$

Here, $\hat{\mathbf{k}} = \frac{\hat{\mathbf{p}}}{\hbar}$ and $\hat{k}^2 = \hat{\mathbf{k}}\cdot\hat{\mathbf{k}}$. For a direction i along which the system is periodic, $\hat{k}_i f_m^{(\alpha)} = k_i f_m^{(\alpha)}$ where k_i is an element of the Bloch wavevector. $\{\mathbf{R}\}_{\mathrm{sys}}$ is shorthand for the positions and species of the atoms in the system (which is, in general, *not* equal to $\{\mathbf{R}\}_{\infty)}^l$, the positions, and species of the atoms in a bulk crystal of type l. $\Theta_l(\mathbf{r})$ is a step function that is 1 if point $\mathbf{r}$ is within a material of type l and is zero otherwise. Away from an interface, for a point $\mathbf{r}$ within a material of type l, $\widehat{H}_{m'm}^{\mathrm{EFA,tot}}\left(\mathbf{r}, \{\mathbf{R}\}_{\mathrm{sys}}\right) = \widehat{H}_{m'm}^{\mathrm{EFA}}(\{\mathbf{R}\}_\infty^l)$.

The approximations involved in equations 165–167 may be made clear through a brief outline of their derivation by Burt (*87*). If one substitutes the envelope function expansion, equation 164, into the one-electron Schrödinger equation, one may obtain the following equation:

$$\sum_{m'}\left\{\left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e}U_{m'}(\mathbf{r})\right]+\frac{\hbar}{m_e}[\hat{\mathbf{p}}U_{m'}(\mathbf{r})]\cdot\hat{\mathbf{k}}+U_{m'}(\mathbf{r})\frac{\hbar^2\hat{k}^2}{2m_e}\right.$$
$$\left.+\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_{\text{sys}})U_{m'}(\mathbf{r})-E_\alpha U_{m'}(\mathbf{r})\right\}f_{m'}^{(\alpha)}(\mathbf{r})=0\,. \tag{169}$$

Since derivatives of $U_m(\mathbf{r})$ have the same periodicity as $U_m(\mathbf{r})$ itself, one may write, without approximation (*81*, *87*),

$$\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e}U_{m'}(\mathbf{r})=\sum_m\left[\int_{\Omega_c}U_{m'}^{\dagger}(\mathbf{r})\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e}U_m(\mathbf{r})\mathrm{d}^3\mathbf{r}\right]U_m(\mathbf{r})\,, \tag{170}$$

and

$$\hat{\mathbf{p}}U_{m'}(\mathbf{r})=\sum_m\mathbf{p}_{m'm}U_m(\mathbf{r})\ , \tag{171}$$

where the sums may be infinite. If the effective potential can be approximated as

$$\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_{\text{sys}})\approx\sum_l\Theta_l(\mathbf{r})V_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty^l)\ , \tag{172}$$

then (*81*, *87*)

$$\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_{\text{sys}})\psi_\alpha(\mathbf{r})\approx$$
$$\sum_l\Theta_l(\mathbf{r})\sum_{m'}\sum_m\left[\int_{\Omega_c}U_{m'}^{\dagger}(\mathbf{r})V_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty^l)U_m(\mathbf{r})\mathrm{d}^3\mathbf{r}\right]f_{m'}^{(\alpha)}(\mathbf{r})\,. \tag{173}$$

Once one substitutes equations170, 171, and 173 into equation 169, the matrix differential equation 165 readily follows.

The approximation of effective potential $\hat{V}_{\text{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_{\text{sys}})$ in equation 172 assumes the potential field is due to atoms in a regular lattice array. This is reasonable for a point $\mathbf{r}$ in an environment that locally resembles a bulk crystal, such as, for example, in the interior of a layer of a superlattice composed of alternating single-crystal layers, provided that the layers are sufficiently thick. However, for a point $\mathbf{r}$ near an interface, this approximation is rather crude. In actuality, the interface itself can be considered a region of small but finite thickness, as illustrated in figure 21, which shows an interface between GaAs and AlAs. The yellow lines mark planes of Al and Ga atoms on either side of the interface, with a plane of As atoms in between. Across the region between the planes of Ga and Al atoms, the true potential felt by the electron (i.e., equations 5 and 6) is generally continuous, though it may rise or fall steeply across this interface.

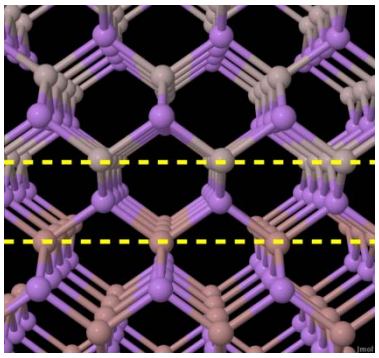

Figure 21. Interface between AlAs at top (grey and purple atoms) and GaAs at bottom (pink and purple atoms). The two yellow dotted lines indicate the planes of Al and Ga atoms on opposite sides of the interface. The figure was created with Jmol (*45*).

Since the envelope function expansion contains, in principle, an infinite number of terms, the matrix differential operator $\widehat{H}^{\mathrm{EFA,tot}}_{m'm}\left(\mathbf{r},\{\mathbf{R}\}_{\mathrm{sys}}\right)$ likewise has an infinite number of rows and columns. One way of obtaining a finite matrix operator from $\widehat{H}^{\mathrm{EFA,tot}}_{m'm}$ is to note how $\widehat{H}^{\mathrm{EFA}}_{m'm}(\{\mathbf{R}\}^{l}_{\infty})$ in equation 167 and $H^{k\cdot p}_{m'm}(\{\mathbf{R}\}^{l}_{\infty})$ in equation 111 are analogous, with the former obtainable from the latter by making the replacement $\mathbf{k}\rightarrow\hat{\mathbf{k}}$. This replacement, then, may be made for a previously formulated $k\cdot p$ Hamiltonian matrix for a bulk crystal (*86*). However, there are problems with this approach. Material parameters are functions of position across an interface, so, for example, the Luttinger parameters would become $\gamma_m(\mathbf{r})=\sum_l \Theta_l(\mathbf{r})\gamma^l_m$ where γ^l_m is the m^{th} Luttinger parameter for material l. While the scalar components of the Bloch wave vector $\mathbf{k}$ commute with such parameters, components of the differential operator $\hat{\mathbf{k}}$ do not (*89*). A work-around for this issue has been to impose a symmetrization scheme on $\widehat{H}^{\mathrm{EFA,tot}}_{m'm}$ (*90*), so that, for example, operators within it of the forms

$$A(x)\frac{\partial^2}{\partial x^2}\ \text{and}\ B(x)\frac{\partial}{\partial x}\ , \tag{174}$$

are replaced, respectively, by

$$\frac{\partial}{\partial x}A(x)\frac{\partial}{\partial x}\ \ \text{and}\ \frac{1}{2}\left(B(x)\frac{\partial}{\partial x}+\frac{\partial}{\partial x}B(x)\right). \tag{175}$$

This operator replacement scheme, however, is ad hoc (*91*) and can lead to spurious solutions (*92*).

Instead of directly substituting $\hat{\mathbf{k}}$ for $\mathbf{k}$ in a previously formulated finite-sized $k\cdot p$ Hamiltonian matrix, matrix differential equation 165 can itself be transformed into a matrix equation with a

finite number of terms, following the method of Burt (*87*). In this method, the expansion in equation 164 is split into two sums,

$$\psi_\alpha(\mathbf{r}) = \sum_s^A f_s^{(\alpha)}(\mathbf{r}) U_s(\mathbf{r}) + \sum_r^B f_r^{(\alpha)}(\mathbf{r}) U_r(\mathbf{r}) \,. \tag{176}$$

Following the notation given previously for the derivation of the Luttinger-Kohn Hamiltonian matrix, the states of the first sum are in class A, and the states in the second sum are in class B. The class A states are not necessarily the exact same ones used by Luttinger and Kohn, but they are the states that contribute the most to $\psi_\alpha(\mathbf{r})$. The index s refers to envelope functions and terms corresponding to class A states; index r refers to envelope functions and terms corresponding to class B states, and index m refers to terms in either sum. With this notation, the envelope function matrix equations may be expressed as

$$\sum_m \left[H_{rm}^{(0)}(\mathbf{r}) + \frac{\hbar}{m_e} \hat{\mathbf{k}} \cdot \mathbf{p}_{rm} + \delta_{rm} \frac{\hbar^2 \hat{k}^2}{2m_e} \right] f_m^{(\alpha)}(\mathbf{r}) = E_\alpha f_r^{(\alpha)}(\mathbf{r}) \,, \tag{177}$$

and

$$\sum_m \left[H_{sm}^{(0)}(\mathbf{r}) + \frac{\hbar}{m_e} \hat{\mathbf{k}} \cdot \mathbf{p}_{sm} + \delta_{sm} \frac{\hbar^2 \hat{k}^2}{2m_e} \right] f_m^{(\alpha)}(\mathbf{r}) = E_\alpha f_s^{(\alpha)}(\mathbf{r}) \,, \tag{178}$$

where

$$H_{m'm}^{(0)}(\mathbf{r}) = \sum_l \Theta_l(\mathbf{r}) \, H_{m'm}^{1e^-}(\{\mathbf{R}\}_\infty^l) \,. \tag{179}$$

Burt neglects the curvature of the class B envelope functions $\hat{k}^2 f_m^{(\alpha)}(\mathbf{r})$ is, so equation 177 becomes approximately (*81*)

$$\sum_{r'}^B H_{rr'}^{(0)}(\mathbf{r}) f_{r'}^{(\alpha)}(\mathbf{r}) + \sum_{l,s'}^A \left[H_{rs'}^{(0)}(\mathbf{r}) + \frac{\hbar}{m_e} \hat{\mathbf{k}} \cdot \mathbf{p}_{rs'} \right] f_{s'}^{(\alpha)}(\mathbf{r}) \approx E_\alpha f_r^{(\alpha)}(\mathbf{r}) \,, \tag{180}$$

where the sum over m has been split into two sums over r' and s', where r' and s' correspond to classes A and B, and the terms $\hbar \, \hat{\mathbf{k}} \cdot \mathbf{p}_{rr'}/m_e$ are also assumed to be negligibly small (*81*). If the class B periodic functions $U_r(\mathbf{r})$ are such that $H_{rr'}^{(0)} = H_{rr}^{(0)} \delta_{rr'}$, then the class B envelope functions are approximately (*81*, *87*)

$$f_r^{(\alpha)}(\mathbf{r}) \approx \frac{1}{E_\alpha - H_{rr}^{(0)}(\mathbf{r})} \sum_{s'}^A \left[H_{rs'}^{(0)}(\mathbf{r}) + \frac{\hbar}{m_e} \hat{\mathbf{k}} \cdot \mathbf{p}_{rs'} \right] f_{s'}^{(\alpha)}(\mathbf{r}) \,. \tag{181}$$

Equation 178 can be rewritten to show its dependence on $f_r^{(\alpha)}(\mathbf{r})$:

$$\sum_{s'}^{A}\left[H_{ss'}^{(0)}(\mathbf{r})+\frac{\hbar}{m_e}\hat{\mathbf{k}}\cdot\mathbf{p}_{ss'}+\delta_{ss'}\frac{\hbar^2\hat{k}^2}{2m_e}\right]f_{s'}^{(\alpha)}(\mathbf{r})$$
$$+\sum_{r}^{B}\left[H_{sr}^{(0)}(\mathbf{r})+\frac{\hbar}{m_e}\hat{\mathbf{k}}\cdot\mathbf{p}_{sr}\right]f_{r}^{(\alpha)}(\mathbf{r})=E_\alpha f_s^{(\alpha)}(\mathbf{r})\,. \tag{182}$$

Equation 181 is then substituted into equation 182, but several of the resulting terms are either zero due to symmetries in class *A* states or are neglected because they are only significant near an interface. (Terms containing derivatives of matrix element H_{rs} are an example of the latter, since these elements are constant within a bulk material, and regions away from the interface are presumed to be bulk-like.) This yields the following (*81*, *87*):

$$\textstyle\sum_{s'}^{A}\widehat{H}_{ss'}^{\mathrm{BF}}\left(\mathbf{r},\{\mathbf{R}\}_{\mathrm{sys}}\right)f_{s'}^{(\alpha)}(\mathbf{r})\approx E_\alpha f_s^{(\alpha)}(\mathbf{r})\,, \tag{183}$$

$$\widehat{H}_{ss'}^{\mathrm{BF}}\left(\mathbf{r},\{\mathbf{R}\}_{\mathrm{sys}}\right)=H_{ss'}^{(2)}(\mathbf{r})+\delta_{ss'}\frac{\hbar^2\hat{k}^2}{2m_e}+\frac{\hbar^2}{m_e^2}\sum_{r}^{B}\hat{\mathbf{k}}\cdot\mathbf{p}_{sr}\frac{1}{E_\alpha-H_{rr}^{(0)}(\mathbf{r})}\mathbf{p}_{rs'}\cdot\hat{\mathbf{k}}\,, \tag{184}$$

and

$$H_{ss'}^{(2)}(\mathbf{r})=H_{ss'}^{(0)}(\mathbf{r})+\textstyle\sum_{r}^{B}\frac{H_{sr}^{(0)}(\mathbf{r})H_{rs'}^{(0)}(\mathbf{r})}{E_\alpha-H_{rr}^{(0)}(\mathbf{r})}\approx H_{ss'}^{(0)}(\mathbf{r})\,, \tag{185}$$

where $\widehat{H}_{ss'}^{\mathrm{BF}}\left(\mathbf{r},\{\mathbf{R}\}_{\mathrm{sys}}\right)$ is called a Burt-Foreman Hamiltonian. Here, $\mathbf{p}_{sr}$ is position-independent and commutes with $\hat{\mathbf{k}}$. $H_{rr}^{(0)}(\mathbf{r})$, however, is clearly not position-independent. A Hamiltonian matrix derived from equations 183–185 should have the proper ordering between the operator $\hat{\mathbf{k}}$ and the empirical parameters formed from the sums over r (class *B*). One can see an example of such a Hamiltonian from Foreman (*93*). If $E_\alpha \approx E_0$ and $H_{m'm}^{(0)} = E_{m0}\delta_{m'm}$, then for a bulk material, $\widehat{H}_{m'm}^{\mathrm{BF}}$ resembles the Luttinger-Kohn Hamiltonian shown in equation 157.

Figure 22 is pseudocode that shows the outline of how an implementation of the envelope function approximation may be done. In this pseudocode, discretization refers to the use of either finite difference or finite-element methods to transform a matrix differential eigenequation $\widehat{\mathbf{H}}\mathbf{f}^{(\alpha)}(\mathbf{r})=E_{i\alpha}\mathbf{f}^{(\alpha)}(\mathbf{r})$, where the elements of $\widehat{\mathbf{H}}$ are differential operators and the elements of $\mathbf{f}^{\alpha}(\mathbf{r})$ are functions, into a matrix eigenvalue problem, $\mathbf{H}\mathbf{F}^{(\alpha)}=E_{i\alpha}\mathbf{F}^{(\alpha)}$, where the elements of $\mathbf{H}$ and $\mathbf{F}^{(\alpha)}$ are numerical values.

Set of empirical parameters *params* is given.
Matrix differential operator $\hat{\mathbf{H}}(\hat{k}_1, \hat{k}_2, \hat{k}_3, params)$ is given.
Number of dimensions along which electrons are confined, N_D, is given.

if $N_D = 1$ **then** {Electrons are free to move within a plane.}

Let differential operators $\hat{k}_1$ and $\hat{k}_2$ become the scalars k_1 and k_2.

Discretize the matrix differential equation $\hat{\mathbf{H}}(k_1, k_2, \hat{k}_3, params)\,\mathbf{f}^{(i\mathbf{k}_\parallel)}(\mathbf{r}) = E_{i\mathbf{k}_\parallel}\mathbf{f}^{(i\mathbf{k}_\parallel)}(\mathbf{r})$, where $\mathbf{k}_\parallel = (k_1, k_2, 0)$, along a line pointing along the r_3 axis (i.e. z-direction), forming a matrix eigenvalue problem $\mathbf{H}(\mathbf{k}_\parallel)\mathbf{F}^{(i\mathbf{k}_\parallel)} = E_{i\mathbf{k}_\parallel}\mathbf{F}^{(i\mathbf{k}_\parallel)}$.

for k_1 from $k_1^{\min}$ to $k_1^{\max}$ **do**
for k_2 from $k_2^{\min}$ to $k_2^{\max}$ **do**
Let $\mathbf{k}_\parallel = (k_1, k_2, 0)$.
Find eigenvalues $E_{i\mathbf{k}_\parallel}$ and eigenvectors $\mathbf{F}^{(i\mathbf{k}_\parallel)}$ of $\mathbf{H}(\mathbf{k}_\parallel)$.
end for
end for

else if $N_D = 2$ **then** {Electrons can only move along line.}

Let differential operator $\hat{k}_3$ become the scalar k_3.

Discretize the matrix differential equation $\hat{\mathbf{H}}(\hat{k}_1, \hat{k}_2, k_3, params)\,\mathbf{f}^{(ik_3)}(\mathbf{r}) = E_{ik_3}\mathbf{f}^{(ik_3)}(\mathbf{r})$ in the $r_2 r_3$ plane (i.e. xy plane), forming a matrix eigenvalue problem $\mathbf{H}(k_3)\mathbf{F}^{(ik_3)} = E_{ik_3}\mathbf{F}^{(ik_3)}$.

for k_3 from $k_3^{\min}$ to $k_3^{\max}$ **do**
Find eigenvalues E_{ik_3} and eigenvectors $\mathbf{F}^{(ik_3)}$ of $\mathbf{H}(k_3)$.
end for

else if $N_D = 3$ **then** {Electrons confined in all three dimensions.}

Discretize the matrix differential equation $\hat{\mathbf{H}}(\hat{k}_1, \hat{k}_2, \hat{k}_3, params)\,\mathbf{f}^{(i)}(\mathbf{r}) = E_i\mathbf{f}^{(i)}(\mathbf{r})$ in all three dimensions, forming a matrix $\mathbf{H}\mathbf{F}^{(i)} = E_i\mathbf{F}^{(i)}$.

Find eigenvalues E_i and eigenvectors $\mathbf{F}^{(i)}$ of $\mathbf{H}$.

end if

Figure 22. Pseudocode for implementing the envelope function approximation. The matrix differential operator $\hat{\mathbf{H}}(\hat{k}_1, \hat{k}_2, \hat{k}_3, params)$ may either be a symmetrized bulk $k \cdot p$ Hamiltonian matrix where the substitution $\mathbf{k} \to \hat{\mathbf{k}}$ has been made or a Burt-Foreman Hamiltonian.

The envelope function expansion of the one-electron wavefunction in equation 164 is not fully general. In a region of a structure that is locally periodic—that is, a region composed of unit cells that may be tiled throughout a finite space rather than to infinity—it is physically reasonable. The local periodicity is captured by the periodic functions $U_m(\mathbf{r})$, while the lack of full periodicity due to the presence of features such as material interfaces in the structure is accounted for by the envelope functions $f_m^{(\alpha)}$. However, regions of structures may lack even approximate local periodicity, for example, if they are sufficiently inhomogeneously strained. (Such regions may also be regions where the potential cannot be approximated as bulk-like.) The significance of this can be seen in figure 23, where $U_m(\mathbf{r})$ is superimposed over an undistorted lattice with the same periodicity as $U_m(\mathbf{r})$ and a distorted lattice that lacks periodicity. A slowly varying envelope function $f_m^{(\alpha)}$ can modify the amplitude of $U_m(\mathbf{r})$, but it cannot stretch or compress parts of it to match the distortions in the lattice. Also, unlike the periodic functions in Fourier series expansions, the periodic functions in the envelope function expansion of equation 164 all have the same periodicity over the whole domain Ω. Lattice mismatch in a structure, though, means that different regions within the structure may have different local periodicities, so this would limit the envelope function expansion—at least the one given in equation164—to lattice-matched structures, such as those made of GaAs and AlAs.

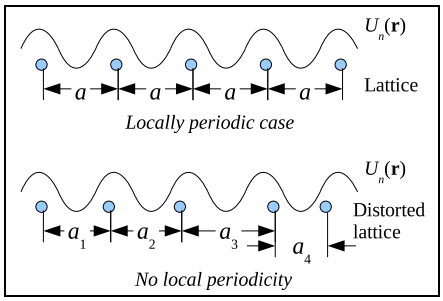

Figure 23. Schematic of a periodic function $U_n(\mathbf{r})$ superimposed over a lattice of evenly spaced atoms (top) and a distorted lattice where the spacing between neighboring atoms is no longer the same.

However, the envelope function approximation has, nonetheless, been applied to structures with significant lattice mismatch, such as a gallium nitride (GaN) dot embedded in an AlN matrix (*94*), or an InAs quantum dot embedded in a GaAs matrix (*89*). Work by Foreman (*95*) indicates why this appears to lead to reasonable results in practice. They show the one-electron wavefunction may be expanded as

$$\psi_\alpha(\mathbf{r}) = \sum_{m'} \sum_{i=1}^{N_{\text{cell}}} F_{m'}^{(\alpha)}(\overline{\mathbf{R}}_i)\delta_B(\mathbf{r} - \overline{\mathbf{R}}_i)U_{m'}(\mathbf{r}, \overline{\mathbf{R}}_i)\,, \tag{186}$$

where N_{cell} is the number of crystal unit cells in the system (which in general may not be identical when the system is composed of multiple materials), $\overline{\mathbf{R}}_i$ is the position vector pointing to unit cell i, and $\delta_B(\mathbf{r}-\overline{\mathbf{R}}_i)$ is defined as

$$\delta_B(\mathbf{r}-\mathbf{R}_i)=\frac{\Omega_i}{(2\pi)^3}\int_{\text{FBZ}}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}d^3\mathbf{r}\,. \tag{187}$$

Here, "FBZ" is the first Brillouin zone of a crystal unit cell at $\overline{\mathbf{R}}_i$, and Ω_i is the volume of this cell. This function is similar to the Dirac delta function in that it peaks sharply near $\overline{\mathbf{R}}_i$, and is zero outside a neighborhood of $\overline{\mathbf{R}}_i$, with a volume about equal to Ω_i. $U_{m'}(\mathbf{r},\overline{\mathbf{R}}_i)$ is the value of $U_{m'}(\mathbf{r})$ for a bulk crystal whose unit cell is the same as that of cell i. $F_{m'}^{(\alpha)}(\overline{\mathbf{R}}_i)$ is a discrete envelope function defined at points $\overline{\mathbf{R}}_i$, and its relationship to the continuous envelope function is

$$f_{m'}^{(\alpha)}(\mathbf{r})=\sum_{i=1}^{N_{\text{cell}}}F_{m'}^{(\alpha)}(\overline{\mathbf{R}}_i)\delta_B(\mathbf{r}-\overline{\mathbf{R}}_i)\,. \tag{188}$$

If $U_{m'}(\mathbf{r},\overline{\mathbf{R}}_i)$ is the same for all crystal unit cells in the system, then $U_{m'}(\mathbf{r},\overline{\mathbf{R}}_i)=U_{m'}(\mathbf{r})$, and the expansion in equation 186 reduces to that in equation 164. However, this expansion is more general than the original envelope function expansion and accounts for changes in the local periodicity. With this redefinition of $f_{m'}^{(\alpha)}(\mathbf{r})$, the matrix operator $\widehat{H}_{m'm}^{\text{EFA,tot}}\left(\mathbf{r},\{\mathbf{R}\}_{\text{sys}}\right)$ in equation 165 becomes

$$\begin{aligned}\widehat{H}_{m'm}^{\text{EFA,tot}}\left(\mathbf{r},\{\mathbf{R}\}_{\text{sys}}\right)=\overline{H}_{m'm}^{(0)}(\mathbf{r})&+\frac{\hbar}{2m_e}\left[\mathbf{p}_{m'm}(\mathbf{r})\cdot\hat{\mathbf{k}}+\hat{\mathbf{k}}\cdot\mathbf{p}_{m'm}(\mathbf{r})\right]\\&+\delta_{m'm}\frac{\hbar^2\hat{k}^2}{2m_e}+\frac{\hbar}{2m_e}\left[\mathbf{P}_{m'm}(\mathbf{r})\cdot\hat{\mathbf{k}}+\hat{\mathbf{k}}\cdot\mathbf{P}_{m'm}(\mathbf{r})\right]\\&+\frac{1}{2m_e}\sum_l\left[\mathbf{p}_{m'l}(\mathbf{r})\cdot\mathbf{P}_{lm}(\mathbf{r})+\mathbf{P}_{m'l}(\mathbf{r})\cdot\mathbf{p}_{lm}(\mathbf{r})\right]\\&+\frac{1}{2m_e}\sum_l\mathbf{P}_{m'l}(\mathbf{r})\cdot\mathbf{P}_{lm}(\mathbf{r})\,,\end{aligned} \tag{189}$$

where

$$\overline{H}_{m'm}^{(0)}(\mathbf{r})=\sum_i\overline{\Theta}_i(\mathbf{r})\,H_{m'm}^{1\mathrm{e}^-}\left(\{\mathbf{R}\}_\infty^i\right), \tag{190}$$

$$\mathbf{p}_{m'm}(\mathbf{r})=\sum_i\overline{\Theta}_i(\mathbf{r})\int_{\Omega_i}U_{m'}^\dagger(\mathbf{r},\overline{\mathbf{R}}_i)\hat{\mathbf{p}}U_m(\mathbf{r},\overline{\mathbf{R}}_i)\mathrm{d}^3\mathbf{r}\,, \tag{191}$$

$$\mathbf{P}_{m'm}(\mathbf{r})=\sum_i\overline{\Theta}_i(\mathbf{r})\int_{\Omega_i}U_{m'}^\dagger(\mathbf{r},\overline{\mathbf{R}}_i)\hat{\mathbf{p}}[\delta_B(\mathbf{r}-\overline{\mathbf{R}}_i)U_m(\mathbf{r},\overline{\mathbf{R}}_i)]\mathrm{d}^3\mathbf{r}\,, \tag{192}$$

and

$$\overline{\Theta}_i(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \text{ in unit cell } i \\ 0, & \text{otherwise} \end{cases} . \tag{193}$$

Here, $\{\mathbf{R}\}_{\infty}^{i}$ is the set of positions and species of the atoms of a bulk crystal whose unit cell is the same as that of cell i. The symmetry of the sum $\mathbf{p}_{m'm}(\mathbf{r}) \cdot \hat{\mathbf{k}} + \hat{\mathbf{k}} \cdot \mathbf{p}_{m'm}(\mathbf{r})$ in the second term of equation 189 is *not* due to an artificially imposed symmetrization scheme, such as the one discussed above and shown in equation 175. If the structure in question is composed of locally periodic crystal regions separated by material interfaces, then away from an interface, $\mathbf{p}_{m'm}$ is effectively independent of $\mathbf{r}$, $\mathbf{P}_{m'm}(\mathbf{r})$ is zero, and $\widehat{H}_{m'm}^{\mathrm{EFA,tot}}(\mathbf{r}, \{\mathbf{R}\}_{\mathrm{sys}})$ in equation 189 reduces to the $\widehat{H}_{m'm}^{\mathrm{EFA,tot}}(\mathbf{r}, \{\mathbf{R}\}_{\mathrm{sys}})$ in equation 167. The terms containing $\mathbf{P}_{m'm}(\mathbf{r})$, then, are interfacial terms, which are often negligible, especially in systems composed predominantly of regions of bulk crystal, e.g., superlattices with thick layers (*87*). The end result is that the envelope function matrix equation, i.e., equation 165, may be applied to structures that are not necessarily lattice-matched.

4.4 Effective Mass Approximation

Sometimes the term "effective mass approximation" is used as a synonym for the $k \cdot p$ and envelope function methods (*32*, *82*, *87*), but here it will be used to describe a particular simplification of these methods. One can see this approximation in the solutions to the Kane Hamiltonian matrix for small k, equations 141, which all have the following form:

$$E_{i\mathbf{k}} = E_i + \frac{\hbar^2 k^2}{2m_i} . \tag{194}$$

This is a solution to the Schrödinger equation for an independent particle with mass m_i in a potential with the constant value E_i,

$$\left[\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_i} + E_i\right] \psi_{i\mathbf{k}}(\mathbf{r}) = E_{i\mathbf{k}} \psi_{i\mathbf{k}}(\mathbf{r}) , \tag{195}$$

and because of this, m_i is called the effective mass. Accordingly, electrons with a small Bloch wave vector $\mathbf{k}$ behave approximately as if they were free electrons with this effective mass. Holes generally behave approximately as if they were free electrons with a negative effective mass.

In the envelope function method, the effective mass approximation amounts to an envelope function expansion with only one term, that is,

$$\psi_{\alpha}(\mathbf{r}) = f^{(\alpha)}(\mathbf{r}) U(\mathbf{r}) , \tag{196}$$

and the envelope function matrix differential equation reduces to the following one-particle equation (*10*, *17*),

$$\left[E_i + \frac{\hbar^2 \hat{k}^2}{2m_i}\right] f^{(\alpha)}(\mathbf{r}) = \left[\frac{\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}}{2m_i} + E_i\right] f^{(\alpha)}(\mathbf{r}) = E_{\alpha} f^{(\alpha)}(\mathbf{r}) . \tag{197}$$

Strictly speaking, equation 197 is valid only if m_i is constant. In a heterostructure, m_i is at most piecewise constant, taking on one value in a part of the structure composed of one material and a different value in a part composed of a different material. (Similarly, E_i can at most be only piecewise constant.) In such a structure, then, m_i varies with $\mathbf{r}$, so the operator $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}/2m_i$ is no longer Hermitian, so E_α is no longer guaranteed to be real. This problem can be avoided by modifying the equation to become the following (*10*):

$$\left[\hat{\mathbf{p}} \cdot \left(\frac{\hat{\mathbf{p}}}{2m_i(\mathbf{r})}\right) + E_i\right] f^{(\alpha)}(\mathbf{r}) = E_\alpha f^{(\alpha)}(\mathbf{r}) \,, \tag{198}$$

or equivalently,

$$\left[\hat{\mathbf{k}} \cdot \left(\frac{\hbar^2 \hat{\mathbf{k}}}{2m_i(\mathbf{r})}\right) + E_i\right] f^{(\alpha)}(\mathbf{r}) = E_\alpha f^{(\alpha)}(\mathbf{r}) \,. \tag{199}$$

Figure 24 shows a pseudocode outline of an implementation of the effective mass approximation for the case where the effective mass is either piecewise constant or, for the pure bulk case, constant. In this pseudocode, $\Theta_l(\mathbf{r})$ is a step function that is 1 if point $\mathbf{r}$ is within a material of type l and is zero otherwise. Here, discretization refers to the use of either finite difference or finite element methods to transform a differential eigenequation $\hat{H}\mathbf{f}^{(\alpha)}(\mathbf{r}) = E_\alpha \mathbf{f}^{(\alpha)}(\mathbf{r})$ into a numerical matrix eigenvalue problem $\mathbf{H}\mathbf{F}^{(\alpha)} = E_\alpha \mathbf{F}^{(\alpha)}$.

One use of the effective mass approximation is to supplement a $k \cdot p$ or envelope function implementation that only solves for the energies of valence bands. That is, if one uses a Luttinger-Kohn Hamiltonian to determine the heavy- and light-hole bands, one may then use the effective mass approximation to estimate the behavior of the conduction band near $\mathbf{k} = 0$.

4.5 Accounting for Strain in the k · p and Envelope Function Methods

In *ab initio* and empirical atomistic methods, the effects of strain are at least partially taken into account through the atomic positions $\{\mathbf{R}\}_M$ that are the input of the effective potential $V_{\text{ext,eff}}(\mathbf{r}, \{\mathbf{R}\}_M)$. However, in the $k \cdot p$ method, the effect of $V_{\text{ext,eff}}$ is taken into account indirectly through parameters such as E_g, Δ, the Kane parameter P_{Kane}, or the Luttinger parameters. Atomic positions are not an input to the method at all. Instead, a strain-dependent term is added to the effective potential operator, usually using the approach of Bir and Pikus (*1*).

For each material l, effective mass $m_i(l)$ is given.
For each material l, band extremum $E_i(l)$ is given.
Number of dimensions along which electrons are confined, N_D, is given.

if $N_D = 0$ **then** {Electrons are free to move in any direction.}

Let $m_i = m_i(l)$ and $E_i = E_i(l)$ {There is only one bulk material l here.}
Energy band as a function of wavevector $\mathbf{k}$ is $E_{i\mathbf{k}} \approx E_i + \hbar^2|\mathbf{k}|^2/2m_i$.

else

Let Hamiltonian operator $\hat{H}(\hat{k}_1,\hat{k}_2,\hat{k}_3) \approx \sum_l \Theta_l(\mathbf{r}) \left(\hat{\mathbf{k}} \cdot \frac{\hbar^2 \hat{\mathbf{k}}}{m_i(l)} + E_i(l) \right)$, where $\hat{\mathbf{k}} = (\hat{k}_1,\hat{k}_2,\hat{k}_3)$.

if $N_D = 1$ **then** {Electrons are free to move within a plane.}

Let differential operators $\hat{k}_1$ and $\hat{k}_2$ become the scalars k_1 and k_2.

Discretize the differential equation $\hat{H}(k_1,k_2,\hat{k}_3) f^{(i\mathbf{k}_\parallel)}(\mathbf{r}) = E_{i\mathbf{k}_\parallel} f^{(i\mathbf{k}_\parallel)}(\mathbf{r})$,
where $\mathbf{k}_\parallel = (k_1,k_2,0)$, along a line pointing along the r_3 axis (i.e. z-direction), forming a matrix eigenvalue problem $\mathbf{H}(\mathbf{k}_\parallel)\mathbf{F}^{(i\mathbf{k}_\parallel)} = E_{i\mathbf{k}_\parallel}\mathbf{F}^{(i\mathbf{k}_\parallel)}$.

for k_1 from $k_1^{\min}$ to $k_1^{\max}$ **do**
for k_2 from $k_2^{\min}$ to $k_2^{\max}$ **do**
Let $\mathbf{k}_\parallel = (k_1,k_2,0)$.
Find eigenvalues $E_{i\mathbf{k}_\parallel}$ and eigenvectors $\mathbf{F}^{(i\mathbf{k}_\parallel)}$ of $\mathbf{H}(\mathbf{k}_\parallel)$.
end for
end for

else if $N_D = 2$ **then** {Electrons can only move along line.}

Let differential operator $\hat{k}_3$ become the scalar k_3.

Discretize the differential equation $\hat{H}(\hat{k}_1,\hat{k}_2,k_3) f^{(ik_3)}(\mathbf{r}) = E_{ik_3} f^{(ik_3)}(\mathbf{r})$
in the r_2r_3 plane (i.e. xy plane), forming a matrix eigenvalue problem $\mathbf{H}(k_3)\mathbf{F}^{(ik_3)} = E_{ik_3}\mathbf{F}^{(ik_3)}$.

for k_3 from $k_3^{\min}$ to $k_3^{\max}$ **do**
Find eigenvalues E_{ik_3} and eigenvectors $\mathbf{F}^{(ik_3)}$ of $\mathbf{H}(k_3)$.
end for

else if $N_D = 3$ **then** {Electrons confined in all three dimensions.}

Discretize the differential equation $\hat{H}(\hat{k}_1,\hat{k}_2,\hat{k}_3) f^{(i)}(\mathbf{r}) = E_i f^{(i)}(\mathbf{r})$ in
all three dimensions, forming a matrix $\mathbf{H}\mathbf{F}^{(i)} = E_i\mathbf{F}^{(i)}$.

Find eigenvalues E_i and eigenvectors $\mathbf{F}^{(i)}$ of $\mathbf{H}$.

end if
end if

Figure 24. Pseudocode for outline of implementation of effective mass approximation for (piecewise) constant effective mass.

The positions and species of atoms in a strained bulk crystal will be denoted here as $\{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}$, and, accordingly, the one-electron Schrödinger equation for a strained bulk crystal where spin-orbit coupling is ignored is

$$\widehat{H}_{1\mathrm{e}^-}\left(\mathbf{r},\{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\right)\psi_{i\mathbf{k}}(\boldsymbol{r}) = \left[\frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m_e} + \hat{V}_{\mathrm{ext,eff}}\left(\mathbf{r},\{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\right)\right]\psi_{i\mathbf{k}}(\mathbf{r}) = E_{i\mathbf{k}}^{\boldsymbol{\chi}}\psi_{i\mathbf{k}}(\mathbf{r})\,, \tag{200}$$

where $E_{i\mathbf{k}}^{\boldsymbol{\chi}}$ denotes an eigenvalue for the strained system. The one-electron Hamiltonian for the strained crystal may be rewritten as

$$\widehat{H}_{1\mathrm{e}^-}\left(\mathbf{r},\{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\right) = \widehat{H}_{1\mathrm{e}^-}(\mathbf{r},\{\mathbf{R}\}_\infty) + \left[\hat{V}_{\mathrm{ext,eff}}\left(\mathbf{r},\{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\right) - \hat{V}_{\mathrm{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)\right]\,, \tag{201}$$

where $\{\mathbf{R}\}_\infty$ here denotes the positions and species of atoms in an *un*strained bulk crystal. The term $\hat{V}_{\mathrm{ext,eff}}\left(\mathbf{r},\{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\right) - \hat{V}_{\mathrm{ext,eff}}(\mathbf{r},\{\mathbf{R}\}_\infty)$, though, cannot be regarded as a small perturbation on the Hamiltonian for the unstrained system, $\widehat{H}_{1\mathrm{e}^-}(\mathbf{r},\{\mathbf{R}\}_\infty)$, because, as Bir and Pikus point out (*1*), this term is not generally small. This can be illustrated in figure 25, which shows model undeformed and deformed sinusoidal potentials of a fictive one-dimensional crystal, denoted v and v^χ, respectively, and the differences between them. A small uniform strain ϵ is introduced that stretches the period of v by a factor of $1+\epsilon$ and changes its amplitude slightly by $w\epsilon$, where w is a parameter. Although these model potentials have only slightly different periods, with $v(x) = v_0 \sin x$ and $v^\chi(x) = (v_0 + w\epsilon)\sin[x/(1+\epsilon)]$, the difference between them grows as x increases to become on the same order as the potentials themselves.

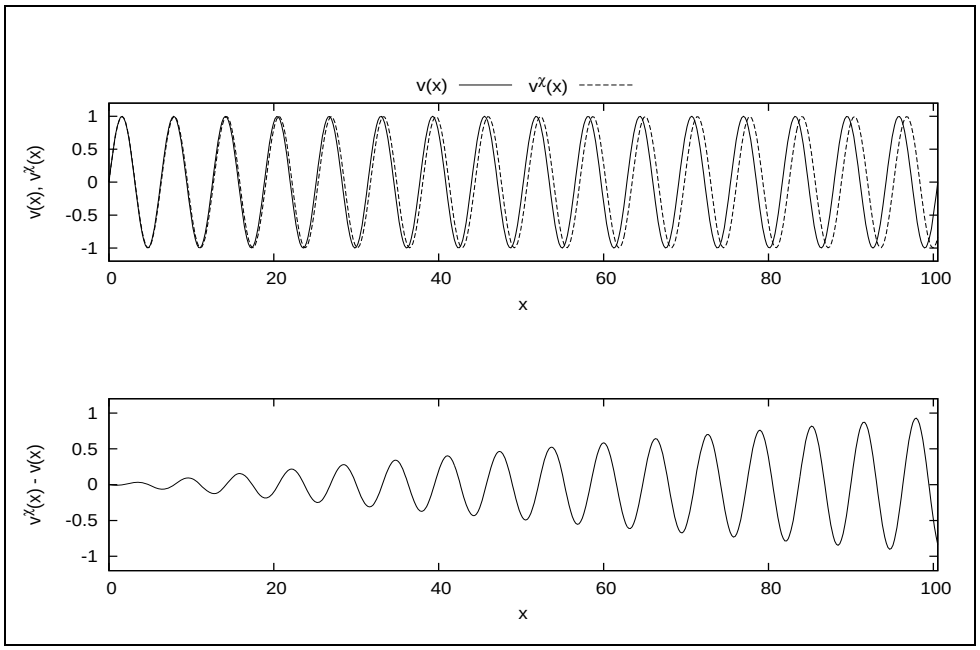

Figure 25. Undeformed and deformed sinusoidal potentials of a one-dimensional crystal, and the difference between them. $v(x) = v_0 \sin x$, and $v^\chi(x) = (v_0 + w\epsilon)\sin[x/(1+\epsilon)]$. For the graph above, $v_0 = 1$, $\epsilon = 1\%$, and $w = -0.3$.

In the previous one-dimensional example, one can specify a deformation map χ that relates the coordinate of the unstrained system, now denoted x^0, to the coordinate of the strained system; that is, $x = \chi(x^0) = (1+\epsilon)x^0$. One can also observe that while the difference $v^\chi(x) - v(x)$ is not small, $v^\chi\big(\chi(x^0)\big) - v(x^0) = w\epsilon$ is small, provided that the parameter w is not too large. Similarly, the difference

$$\hat{V}_{\text{ext,eff}}\big(\boldsymbol{\chi}(\mathbf{r}^0), \{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\big) - \hat{V}_{\text{ext,eff}}(\mathbf{r}^0, \{\mathbf{R}\}_\infty) \approx V_{ij}(\mathbf{r}^0, \{\mathbf{R}\}_\infty)\epsilon_{ij}\,, \tag{202}$$

where $\mathbf{r}^0$ is the coordinate vector for the unstrained system, is also small, provided that V_{ij} is not too large. Again, the convention of summing over repeated indices is used. For a bulk crystal subjected to homogeneous strain, the deformation map $\boldsymbol{\chi}(\mathbf{r}^0)$ equals $\mathbf{F}\mathbf{r}^0 = (\mathbf{I} + \nabla_0\mathbf{u})\mathbf{r}^0$, where $\mathbf{F}$ is the deformation gradient and $\nabla_0\mathbf{u}$ is the gradient of the displacement with respect to the coordinates of the unstrained system. If there is no rotation, the displacement gradient is symmetric and the small-strain tensor may be taken to be $\boldsymbol{\epsilon} = \nabla_0\mathbf{u}$ (*66*), and $\boldsymbol{\chi}(\mathbf{r}^0) = (\mathbf{I} + \boldsymbol{\epsilon})\mathbf{r}^0$. (Use of the small strain tensor, however, implies that $\nabla_0\mathbf{u} \approx \nabla\mathbf{u}$ where $\nabla\mathbf{u}$ is the gradient of the displacement with respect to the coordinates of the *strained* system [*96*].) The coordinate transformation $\mathbf{r} = \boldsymbol{\chi}(\mathbf{r}^0)$ is now introduced into the one-electron Schrödinger equation for a strained crystal:

$$\hat{H}_{1\text{e}^-}\big(\boldsymbol{\chi}(\mathbf{r}^0), \{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\big)\psi_{i\mathbf{k}}\big(\boldsymbol{\chi}(\mathbf{r}^0)\big) = E_{i\mathbf{k}}^{\boldsymbol{\chi}}\psi_{i\mathbf{k}}\big(\boldsymbol{\chi}(\mathbf{r}^0)\big)\,. \tag{203}$$

The momentum operator within $\hat{H}_{1\text{e}^-}\big(\boldsymbol{\chi}(\mathbf{r}^0), \{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\big)$ needs to be expressed in terms of the coordinates of the unstrained system (*32*). Applying the chain rule yields

$$\hat{p}_i = \frac{\partial r_j^0}{\partial r_i}(-\mathrm{i}\hbar)\frac{\partial}{\partial r_j^0} = \left[\frac{\partial}{\partial r_i}\left(\chi_j^{-1}(\mathbf{r})\right)\right](-\mathrm{i}\hbar)\frac{\partial}{\partial r_j^0} = \left[\frac{\partial}{\partial r_i}\Big(\chi_j^{-1}(\mathbf{r})\Big)\right]\hat{p}_j^0\,, \tag{204}$$

where $\hat{\mathbf{p}}^0 = (\hat{p}_1^0, \hat{p}_2^0, \hat{p}_3^0)$ is the momentum operator of the one-electron Hamiltonian of the unstrained system in terms of $\mathbf{r}^0$. For small strain, the inverse deformation map $\mathbf{r}^0 = \boldsymbol{\chi}^{-1}(\mathbf{r})$ is approximately (*32*, *1*)

$$\boldsymbol{\chi}^{-1}(\mathbf{r}) = (\mathbf{I} - \boldsymbol{\epsilon})\mathbf{r}\,, \tag{205}$$

and, accordingly,

$$\hat{p}_i \approx \big(\delta_{ij} - \epsilon_{ij}\big)\hat{p}_j^0\,. \tag{206}$$

Once the momentum operator is expressed in terms of unstrained coordinates, then one may write

$$\hat{\mathbf{p}} \cdot \hat{\mathbf{p}} = \hat{p}_i \hat{p}_i \approx \left(\delta_{ij} - \epsilon_{ij}\right)\hat{p}_j^0 (\delta_{ik} - \epsilon_{ik})\hat{p}_k^0 \,,$$
$$= \left(\delta_{jk} - 2\epsilon_{jk} + \epsilon_{ij}\epsilon_{jk}\right)\hat{p}_j^0 \hat{p}_k^0 \,,$$

and

$$\approx \left(\delta_{jk} - 2\epsilon_{jk}\right)\hat{p}_j^0 \hat{p}_k^0 \,, \tag{207}$$

where the small strain assumption means that the product of strain components is negligible. The one-electron Hamiltonian, then, becomes (*32*, *1*)

$$\hat{H}_{1\mathrm{e}^-}\left(\boldsymbol{\chi}(\mathbf{r}^0), \{\mathbf{R}\}_\infty^{\boldsymbol{\chi}}\right) = \hat{H}_{1\mathrm{e}^-}(\mathbf{r}^0, \{\mathbf{R}\}_\infty) + \epsilon_{jk}\hat{D}_{jk} \,, \tag{208}$$

where

$$\hat{H}_{1\mathrm{e}^-}(\mathbf{r}^0, \{\mathbf{R}\}_\infty) = \frac{\hat{\mathbf{p}}^0 \cdot \hat{\mathbf{p}}^0}{2m_e} + \hat{V}_{\mathrm{ext,eff}}(\mathbf{r}^0, \{\mathbf{R}\}_\infty) \,, \tag{209}$$

and

$$\hat{D}_{jk} = -\frac{\hat{p}_j^0 \hat{p}_k^0}{m_e} + V_{jk}(\mathbf{r}^0, \{\mathbf{R}\}_\infty) \,. \tag{210}$$

The one-electron wavefunction satisfies the Bloch theorem in both strained and unstrained coordinates. To show this, the Bloch wave vector $\mathbf{k}$ needs to be expressed in terms of unstrained coordinates. The primitive lattice vectors and reciprocal lattice vectors in both strained and unstrained coordinates still satisfy the following relationships (*12*),

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}, \text{and } \mathbf{b}_i^0 \cdot \mathbf{a}_j^0 = 2\pi\delta_{ij} \,, \tag{211}$$

where the superscript "0" indicates that a quantity pertains to the unstrained system. Since $\mathbf{a}_j = (\mathbf{I} + \boldsymbol{\epsilon})\mathbf{a}_j^0$, $\mathbf{b}_i = (\mathbf{I} + \boldsymbol{\epsilon})^{-1}\mathbf{b}_i^0$, and

$$\mathbf{k} = \sum_{i=1}^{3} m_i \mathbf{b}_i = (\mathbf{I} + \boldsymbol{\epsilon})^{-1} \sum_{i=1}^{3} m_i \mathbf{b}_i^0 = (\mathbf{I} + \boldsymbol{\epsilon})^{-1}\mathbf{k}^0 \,, \tag{212}$$

where m_i is a real number. Accordingly, $\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} = \mathrm{e}^{\mathrm{i}\mathbf{k}^0\cdot\mathbf{r}^0}$, and

$$\begin{aligned} \psi_{i\mathbf{k}}(\mathbf{r}) &= \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} u_{i\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}^0\cdot\mathbf{r}^0} u_{i(\mathbf{I}+\boldsymbol{\epsilon})^{-1}\mathbf{k}^0}\left((\mathbf{I} + \boldsymbol{\epsilon})\mathbf{r}^0\right) \\ &= \mathrm{e}^{\mathrm{i}\mathbf{k}^0\cdot\mathbf{r}^0} u_{i\mathbf{k}^0}^{\boldsymbol{\chi}}(\mathbf{r}^0) \equiv \psi_{i\mathbf{k}^0}^{\boldsymbol{\chi}}(\mathbf{r}^0) \,, \end{aligned} \tag{213}$$

where $u_{i\mathbf{k}^0}^{\boldsymbol{\chi}}(\mathbf{r}^0) \equiv u_{i(\mathbf{I}+\boldsymbol{\epsilon})^{-1}\mathbf{k}^0}\left((\mathbf{I} + \boldsymbol{\epsilon})\mathbf{r}^0\right)$. The function $u_{i\mathbf{k}^0}^{\boldsymbol{\chi}}(\mathbf{r}^0)$ has the periodicity of the unstrained lattice (*1*); that is,

$$
\begin{aligned}
u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0+\bar{\mathbf{R}}^0) &= u_{i(\mathbf{I}+\boldsymbol{\epsilon})^{-1}\mathbf{k}^0}\big((\mathbf{I}+\boldsymbol{\epsilon})(\mathbf{r}^0+\bar{\mathbf{R}}^0)\big)\,,\\
&= u_{i\mathbf{k}}\big((\mathbf{I}+\boldsymbol{\epsilon})(\mathbf{r}^0+\bar{\mathbf{R}}^0)\big)\,,\\
&= u_{i\mathbf{k}}\big((\mathbf{I}+\boldsymbol{\epsilon})\mathbf{r}^0+(\mathbf{I}+\boldsymbol{\epsilon})\bar{\mathbf{R}}^0\big)\,,\\
&= u_{i\mathbf{k}}(\mathbf{r}+\bar{\mathbf{R}}) = u_{i\mathbf{k}}(\mathbf{r})\,,
\end{aligned}
$$

and

$$
= u_{i(\mathbf{I}+\boldsymbol{\epsilon})^{-1}\mathbf{k}^0}\big((\mathbf{I}+\boldsymbol{\epsilon})\mathbf{r}^0\big) = u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0), \tag{214}
$$

where $\bar{\mathbf{R}}$ and $\bar{\mathbf{R}}^0$ are lattice vectors of the strained and unstrained crystal.

Applying the operator $\epsilon_{jk}\widehat{D}_{jk}$ to $\psi^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)$ yields

$$
\begin{aligned}
\epsilon_{jk}\widehat{D}_{jk}\psi^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0) &= -\frac{\mathrm{e}^{\mathrm{i}\mathbf{k}^0\cdot\mathbf{r}^0}}{m_e}\big[\epsilon_{jk}\hat{p}^0_j\hat{p}^0_k+\epsilon_{jk}\big(\hbar k^0_j\hat{p}^0_k+\hbar k^0_k\hat{p}^0_k\big)+\epsilon_{jk}\hbar^2k^0_jk^0_k\big]u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\\
&\qquad+\epsilon_{jk}V_{jk}(\mathbf{r}^0,\{\mathbf{R}\}_\infty)u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\,,\\
&= -\frac{\mathrm{e}^{\mathrm{i}\mathbf{k}^0\cdot\mathbf{r}^0}}{m_e}\big[\epsilon_{jk}\hat{p}^0_j\hat{p}^0_k+2\epsilon_{jk}\hbar k^0_j\hat{p}^0_k+\epsilon_{jk}\hbar^2k^0_jk^0_k\big]u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\\
&\quad+\epsilon_{jk}V_{jk}(\mathbf{r}^0,\{\mathbf{R}\}_\infty)u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\,,
\end{aligned}
$$

and

$$
= \mathrm{e}^{\mathrm{i}\mathbf{k}^0\cdot\mathbf{r}^0}\epsilon_{jk}\left[\widehat{D}_{jk}-\frac{2\hbar k^0_j\hat{p}^0_k}{m_e}-\frac{\hbar^2k^0_jk^0_k}{m_e}\right]u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\,. \tag{215}
$$

This leads to the following Schrödinger-like equation:

$$
\begin{aligned}
&\left[\widehat{H}_{1\mathrm{e}^-}(\mathbf{r}^0,\{\mathbf{R}\}_\infty)+\frac{\hbar}{m_e}\big(\delta_{jk}-2\epsilon_{jk}\big)k^0_j\hat{p}^0_k+\epsilon_{jk}\widehat{D}_{jk}\right]u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\\
&\quad=\left[E_{i\mathbf{k}^0}-\frac{\hbar^2}{2m_e}\big(\delta_{jk}-2\epsilon_{jk}\big)k^0_jk^0_k\right]u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0)\,.
\end{aligned} \tag{216}
$$

As before, the periodic part of the wavefunction is expanded in terms of the states where the Bloch wave vector is zero, i.e.,

$$
u^{\boldsymbol{\chi}}_{i\mathbf{k}^0}(\mathbf{r}^0) = \sum_m C^{(i\mathbf{k}^0)}_m U_m(\mathbf{r})\,, \tag{217}
$$

and a matrix equation is obtained:

$$
\sum_m H^{k\cdot p*}_{m'm}(\{\mathbf{R}\}_\infty,\boldsymbol{\epsilon})C^{(i\mathbf{k})}_m = \left[E^{\boldsymbol{\chi}}_{i\mathbf{k}^0}-\frac{\hbar^2}{2m_e}\big(\delta_{jk}-2\epsilon_{jk}\big)k^0_jk^0_k\right]C^{(i\mathbf{k})}_{m'}, \tag{218}
$$

$$
H^{k\cdot p*}_{m'm}(\{\mathbf{R}\}_\infty,\boldsymbol{\epsilon}) = H^{1\mathrm{e}^-}_{m'm}(\{\mathbf{R}\}_\infty)+\frac{\hbar}{m_e}\big(\delta_{jk}-2\epsilon_{jk}\big)k^0_jp^{0(k)}_{m'm}+\epsilon_{jk}D^{(jk)}_{m'm}\,, \tag{219}
$$

$$
p^{0(i)}_{m'm} = \int_{\Omega_c} U^\dagger_{m'}(\mathbf{r})\hat{p}^0_iU_m(\mathbf{r})\mathrm{d}^3\mathbf{r}\,, \tag{220}
$$

and

$$D_{m'm}^{(ij)} = \int_{\Omega_c} U_{m'}^{\dagger}(\mathbf{r})\widehat{D}_{ij}U_m(\mathbf{r})\mathrm{d}^3\mathbf{r}\,, \tag{221}$$

with Ω_c being the volume of the unit cell of the crystal. Since the strain here is taken to be homogeneous, ϵ_{ij} is a constant and can be moved outside of any integration.

When the basis functions $U_m(\mathbf{r}) = u_{m0}(\mathbf{r})$ are the Luttinger-Kohn basis functions in equation 145, then $\epsilon_{jk}D_{m'm}^{(jk)}$ is such that (*32*)

$$\left[\epsilon_{jk}D_{m'm}^{(jk)}\right] = \begin{bmatrix} P_\epsilon + Q_\epsilon & -S_\epsilon & R_\epsilon & 0 & -S_\epsilon/\sqrt{2} & \sqrt{2}R_\epsilon \\ -S_\epsilon^\dagger & P_\epsilon - Q_\epsilon & 0 & R_\epsilon & -\sqrt{2}Q_\epsilon & \sqrt{3/2}\,S_\epsilon \\ R_\epsilon^\dagger & 0 & P_\epsilon - Q_\epsilon & S_\epsilon & \sqrt{3/2}\,S_\epsilon^\dagger & \sqrt{2}Q_\epsilon \\ 0 & R_\epsilon^\dagger & S_\epsilon^\dagger & P_\epsilon + Q_\epsilon & -\sqrt{2}R_\epsilon^\dagger & -S_\epsilon^\dagger/\sqrt{2} \\ -S_\epsilon^\dagger/\sqrt{2} & -\sqrt{2}Q_\epsilon^\dagger & \sqrt{3/2}\,S_\epsilon & -\sqrt{2}R_\epsilon & P_\epsilon & 0 \\ \sqrt{2}R_\epsilon^\dagger & \sqrt{3/2}\,S_\epsilon^\dagger & \sqrt{2}Q_\epsilon^\dagger & -S_\epsilon/\sqrt{2} & 0 & P_\epsilon \end{bmatrix}, \tag{222}$$

where

$$\begin{aligned} P_\epsilon &= a_v \mathrm{Tr}\,\boldsymbol{\epsilon}\,, \\ Q_\epsilon &= \frac{b}{2}(\epsilon_{11} + \epsilon_{22} - 2\epsilon_{33})\,, \\ R_\epsilon &= -\frac{\sqrt{3}}{2}b(\epsilon_{11} - \epsilon_{22}) + \mathrm{i}d\epsilon_{12}\,, \end{aligned}$$

and

$$S_\epsilon = d(\epsilon_{13} - \mathrm{i}\epsilon_{23})\,. \tag{223}$$

The empirical parameters a_v, b, and d are linear combinations of the matrix elements $D_{m'm}^{(ij)}$, and they related to certain strain states (*81*). If the strain is purely hydrostatic, then the valence band maximum is shifted by

$$\Delta E_v = a_v \mathrm{Tr}\,\boldsymbol{\epsilon}\,. \tag{224}$$

If the strain is biaxial with $\epsilon_{11} = \epsilon_{22} \neq \epsilon_{33}$, the maxima of the heavy and light hole valence band are no longer the same, but are separated by

$$\Delta E_{hl} = 2|b\epsilon_{33}|\,. \tag{225}$$

If the strain is pure shear with $\epsilon_{ij} = e_s(1 - \delta_{ij})$, the separation becomes

$$\Delta E_{hl} = 2|de_s|\,. \tag{226}$$

The conduction band, which is not taken into account by the strain correction to the Luttinger-Kohn Hamiltonian matrix in equation 222, also shifts due to strain. For the case of hydrostatic strain, this shift is

$$\Delta E_c = a_c \mathrm{Tr}\, \boldsymbol{\epsilon}\,, \tag{227}$$

where (*32*)

$$a_c = \int_{\Omega_c} S^{\dagger}(\mathbf{r}) \widehat{D}_{ij} S(\mathbf{r}) \mathrm{d}^3\mathbf{r}\ , \tag{228}$$

and $S(\mathbf{r})$ has the same definition that it does in the formulation of the Kane $k \cdot p$ Hamiltonian matrix. A diagram illustrating the effects of strain on a bulk crystal band structure is shown in figure 26. In general, compressive strains tend to increase the size of the band gap, tensile strains tend to decrease it, and strains departing from a pure hydrostatic state tend to separate the heavy and light hole bands (*97*).

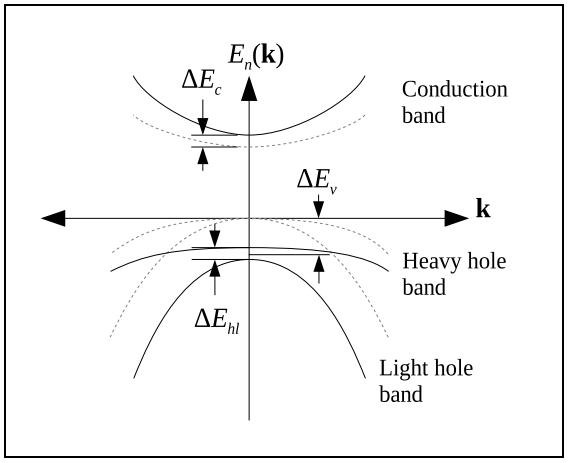

Figure 26. Schematic band structure of a typical bulk semiconductor with a diamond or zincblende crystal structure, where the solid lines indicate the band structure of a strained semiconductor, while the dotted lines indicate the original band structure before the strain is applied. The split-off band is not shown. ΔE_{hl} is the difference between the maxima of the heavy hole and light hole bands, which are not necessarily the same once the semiconductor is strained. ΔE_c is the shift of the conduction band minimum, while $\Delta E_v \pm \Delta E_{hl}/2$ is the downward shift of the light and heavy hole bands, respectively (*81*). The left and right halves of the horizontal axis indicates the magnitude of **k**-values pointing along certain crystal directions. (In this schematic, the actual directions are not important.)

At this point, a strain-dependent term has only been provided for the one-electron Schrödinger equation for an infinite periodic system. The empirical parameters a_c, a_v, b, and d all pertain to a bulk crystal. Furthermore, the strain is taken to be homogeneous, and the coordinate

transformation $\mathbf{r} = (\mathbf{I} + \boldsymbol{\epsilon})\mathbf{r}^0$ implies that the deformation has no rotational component. (See the discussion in appendix B of a similar approximation in the atomistic strain formulation of Pryor et al. [*98*].) Nonetheless, when the envelope function method is applied to strained structures, usually what is done is to take a preexisting $k \cdot p$ Hamiltonian matrix, substitute $\hat{\mathbf{k}}^0$ (i.e., $\hat{\mathbf{p}}^0/\hbar$) for $\mathbf{k}^0$ (possibly with an additional correction to account for Burt-Foreman operator ordering), and then add the same strain correction term that one would use for a bulk $k \cdot p$ matrix (*10*, *32*, *81*, *89*).

Blount (*99*) and Sham and Ziman (*100*) have shown how a strain-dependent term may be introduced into the one-electron Schrödinger equation through a coordinate transformation that does not imply homogeneity of the strain. The coordinate transformation is simply the relationship between the strained coordinate, the unstrained coordinate, and the displacement $\mathbf{u}$, i.e.,

$$\mathbf{r} = \boldsymbol{\chi}(\mathbf{r}^0) = \mathbf{r}^0 + \mathbf{u}\,. \tag{229}$$

The momentum operator can be expressed as (*99*)

$$\hat{p}_i = \frac{\partial r_j^0}{\partial r_i}(-\mathrm{i}\hbar)\frac{\partial}{\partial r_j^0} = \left[\frac{\partial}{\partial r_i}\left(r_j - u_j\right)\right]\hat{p}_j^0 = \left(\delta_{ij} - \frac{\partial u_j}{\partial r_i}\right)\hat{p}_j^0\,, \tag{230}$$

and so for small strain (*100*)

$$\hat{\mathbf{p}} \cdot \hat{\mathbf{p}} = \hat{p}_i\hat{p}_i \approx \left(\delta_{jk} - 2\frac{\partial u_j}{\partial r_k}\right)\hat{p}_j^0\hat{p}_k^0 - \left(\hat{p}_i^0\frac{\partial u_k}{\partial r_i}\right)\hat{p}_k^0\,. \tag{231}$$

Furthermore,

$$\begin{aligned} 2\frac{\partial u_j}{\partial r_k}\hat{p}_j^0\hat{p}_k^0 &= \frac{\partial u_j}{\partial r_k}\hat{p}_j^0\hat{p}_k^0 + \frac{\partial u_j}{\partial r_k}\hat{p}_j^0\hat{p}_k^0 = \frac{\partial u_j}{\partial r_k}\hat{p}_j^0\hat{p}_k^0 + \frac{\partial u_k}{\partial r_j}\hat{p}_k^0\hat{p}_j^0 \\ &= \frac{\partial u_j}{\partial r_k}\hat{p}_j^0\hat{p}_k^0 + \frac{\partial u_k}{\partial r_j}\hat{p}_j^0\hat{p}_k^0 = \left(\frac{\partial u_j}{\partial r_k} + \frac{\partial u_k}{\partial r_j}\right)\hat{p}_j^0\hat{p}_k^0 = 2\epsilon_{jk}\hat{p}_j^0\hat{p}_k^0\,, \end{aligned} \tag{232}$$

and

$$\frac{\partial u_k}{\partial r_i} = \epsilon_{ki} + \omega_{ki}\,, \tag{233}$$

where ω_{ki} is the infinitesimal rotation tensor (96). Therefore,

$$\hat{\mathbf{p}} \cdot \hat{\mathbf{p}} \approx \left(\delta_{jk} - 2\epsilon_{jk}\right)\hat{p}_j^0\hat{p}_k^0 - [\hat{p}_i^0(\epsilon_{ki} + \omega_{ki})]\hat{p}_k^0\,,$$

and

$$= \left(\delta_{jk} - 2\epsilon_{jk}\right)\hat{p}_j^0\hat{p}_k^0 - \left[\frac{\partial}{\partial r_i}(\epsilon_{ki} + \omega_{ki})\right](-\mathrm{i}\hbar)\hat{p}_k^0. \tag{234}$$

If both ϵ_{ki} and ω_{ki} are sufficiently slowly varying, then the second term in the above equation is negligible, and the remainder of the above equation coincides with the corresponding result from Bir and Pikus. Also, equations 232 through 234 depend on the assumption that the small strain

tensor is formed from the displacement gradient with respect to the coordinates of the strained system, $\partial u_k/\partial r_i$ or $\nabla\mathbf{u}$, while the formulation of Bir and Pikus is derived under the assumption that the small strain tensor is formed from the displacement gradient with respect to the coordinates of the unstrained system, $\partial u_k/\partial r_i^0$ or $\nabla_0\mathbf{u}$.

If the one-electron wavefunction is expanded according to the envelope function approximation in unstrained coordinates, then

$$\psi_\alpha\big(\boldsymbol{\chi}(\mathbf{r}^0)\big) \equiv \psi_\alpha^{\boldsymbol{\chi}}(\mathbf{r}^0) = \textstyle\sum_{m'} f_{m'}^{(\alpha)}(\mathbf{r}^0)U_{m'}(\mathbf{r}^0)\,. \tag{235}$$

Applying $\epsilon_{jk}\hat{p}_j^0\hat{p}_k^0$ to a term in the this envelope function expansion yields

$$\begin{aligned}\epsilon_{jk}\hat{p}_j^0\hat{p}_k^0 f_{m'}^{(\alpha)}(\mathbf{r}^0)U_{m'}(\mathbf{r}^0) &= \epsilon_{jk}\big[\hat{p}_j^0\hat{p}_k^0 U_{m'}(\mathbf{r}^0) + U_{m'}(\mathbf{r}^0)\hat{p}_j^0\hat{p}_k^0\big]f_{m'}^{(\alpha)}(\mathbf{r}^0)\\ &\quad + \epsilon_{jk}\big\{\big[\hat{p}_j^0 U_{m'}(\mathbf{r}^0)\big]\hat{p}_k^0 + \big[\hat{p}_k^0 U_{m'}(\mathbf{r}^0)\big]\hat{p}_j^0\big\}f_{m'}^{(\alpha)}(\mathbf{r}^0)\,,\\ &= \epsilon_{jk}\big[\hat{p}_j^0\hat{p}_k^0 U_{m'}(\mathbf{r}^0) + \hbar^2 U_{m'}(\mathbf{r}^0)\hat{k}_j^0\hat{k}_k^0\big]f_{m'}^{(\alpha)}(\mathbf{r}^0)\\ &\quad + 2\hbar\epsilon_{jk}\big[\hat{p}_j^0 U_{m'}(\mathbf{r}^0)\big]\hat{k}_k^0 f_{m'}^{(\alpha)}(\mathbf{r}^0)\,,\end{aligned} \tag{236}$$

where $\hat{k}_i^0 = \hat{p}_i^0/\hbar$. Following Burt (*87*), the following expansions may be made:

$$\hat{p}_j^0 U_{m'}(\mathbf{r}^0) = \textstyle\sum_m p_{m'm}^{0(j)} U_m(\mathbf{r}^0)\,, \tag{237}$$

and

$$\hat{p}_j^0\hat{p}_k^0 U_{m'}(\mathbf{r}^0) = \sum_m \left[\int_{\Omega_c} U_{m'}^\dagger(\mathbf{r}^0)\hat{p}_j^0\hat{p}_k^0 U_m(\mathbf{r}^0)\mathrm{d}^3\mathbf{r}\right] U_m(\mathbf{r}^0)\,. \tag{238}$$

The effective potential may be expanded as a Taylor series about $\epsilon_{jk} = 0$ and $\omega_{jk} = 0$.

$$\begin{aligned}&V_{\text{ext,eff}}\big(\boldsymbol{\chi}(\mathbf{r}^0), \{\mathbf{R}\}_{\text{sys}}^{\boldsymbol{\chi}}\big)\\ &\approx V_{\text{ext,eff}}\big(\mathbf{r}^0, \{\mathbf{R}\}_{\text{sys}}\big) + V_{ij}\big(\mathbf{r}^0, \{\mathbf{R}\}_{\text{sys}}\big)\epsilon_{ij} + V_{ij}^{\omega}\big(\mathbf{r}^0, \{\mathbf{R}\}_{\text{sys}}\big)\omega_{ij}\,,\end{aligned} \tag{239}$$

where

$$V_{ij} = \left.\frac{\partial V_{\text{ext,eff}}}{\partial \epsilon_{ij}}\right|_{\substack{\epsilon_{ij}=0\\ \omega_{ij}=0}} \text{ and } V_{ij}^{\omega} = \left.\frac{\partial V_{\text{ext,eff}}}{\partial \omega_{ij}}\right|_{\substack{\epsilon_{ij}=0\\ \omega_{ij}=0}}\,. \tag{240}$$

If the system is subject to only small rotations, then a stationary test charge is unlikely to experience much of a change in potential as the system rotates about it, and V_{ij}^{ω} is likely to be negligible. If one presumes that equation 172 still holds for the undeformed configuration (i.e., with $\mathbf{r}^0$ in place of $\mathbf{r}$), and that

$$V_{ij}\big(\mathbf{r}^0, \{\mathbf{R}\}_{\text{sys}}\big) \approx \sum_l \Theta_l(\mathbf{r}^0)V_{ij}(\mathbf{r}^0, \{\mathbf{R}\}_\infty^l)\,, \tag{241}$$

then one may treat $V_{ij}\big(\mathbf{r}^0, \{\mathbf{R}\}_{\text{sys}}\big)$ much as Burt (*87*) treats $V_{\text{ext,eff}}\big(\mathbf{r}^0, \{\mathbf{R}\}_{\text{sys}}\big)$ in equation 173; that is,

$$V_{ij}\left(\mathbf{r}^0,\{\mathbf{R}\}_{\mathrm{sys}}\right)\psi_\alpha^{\boldsymbol{\chi}}(\mathbf{r}^0)\approx$$
$$\sum_l \Theta_l(\mathbf{r}^0)\sum_{m'}\sum_m\left[\int_{\Omega_c} U_{m'}^\dagger(\mathbf{r}^0)V_{ij}(\mathbf{r}^0,\{\mathbf{R}\}_\infty^l)U_m(\mathbf{r}^0)\mathrm{d}^3\mathbf{r}\right]f_{m'}^{(\alpha)}(\mathbf{r}^0)\,. \tag{242}$$

Applying the envelope function approximation, then, leads to the following equations:

$$\sum_m \widehat{H}_{m'm}^{\mathrm{EFA,tot}}\left(\mathbf{r}^0,\{\mathbf{R}\}_{\mathrm{sys}},\boldsymbol{\epsilon}\right)f_m^{(\alpha)}(\mathbf{r}^0)\approx E_\alpha^{\boldsymbol{\chi}} f_{m'}^{(\alpha)}(\mathbf{r}^0)\,, \tag{243}$$

where

$$\widehat{H}_{m'm}^{\mathrm{EFA,tot}}\left(\mathbf{r}^0,\{\mathbf{R}\}_{\mathrm{sys}},\boldsymbol{\epsilon}\right)=\sum_l \Theta_l(\mathbf{r})\widehat{H}_{m'm}^{\mathrm{EFA}}(\{\mathbf{R}\}_\infty^l,\boldsymbol{\epsilon})\,, \tag{244}$$

and

$$\widehat{H}_{m'm}^{\mathrm{EFA}}(\{\mathbf{R}\}_\infty^l,\boldsymbol{\epsilon})=H_{m'm}^{1\mathrm{e}^-}(\{\mathbf{R}\}_\infty^l)+\frac{\hbar}{m_e}\left(\delta_{jk}-2\epsilon_{jk}\right)p_{m'm}^{0(k)}\hat{k}_j^0+\epsilon_{jk}D_{m'm}^{(jk)}$$
$$+\frac{\hbar^2}{2m_e}\left(\delta_{jk}-2\epsilon_{jk}\right)\hat{k}_j^0\hat{k}_k^0-\frac{1}{2}\left[\frac{\partial}{\partial r_j}\left(\epsilon_{kj}+\omega_{kj}\right)\right]\left(-\frac{\mathrm{i}\hbar}{m_e}p_{m'm}^{0(k)}\right). \tag{245}$$

$\widehat{H}_{m'm}^{\mathrm{EFA}}(\{\mathbf{R}\}_\infty^l,\boldsymbol{\epsilon})$ and $H_{m'm}^{k\cdot p^*}(\{\mathbf{R}\}_\infty^l,\boldsymbol{\epsilon})$ are almost analogous, with the former obtainable from the latter by making the replacement $\mathbf{k}^0\rightarrow\hat{\mathbf{k}}^0$ and adding a gradient term that is the last term in the above equation. This gradient term will be zero if $p_{m'm}^{0(k)}$ is zero, which happens when $U_m(\mathbf{r}^0)$ and $U_{m'}(\mathbf{r}^0)$ are linear combinations of $X(\mathbf{r}^0)$, $Y(\mathbf{r}^0)$, and $Z(\mathbf{r}^0)$, e.g., the class A basis functions of the Luttinger-Kohn formulation, or when $U_m(\mathbf{r}^0)$ and $U_{m'}(\mathbf{r}^0)$ are multiples of $S(\mathbf{r}^0)$. If $U_{m'}(\mathbf{r}^0)=S(\mathbf{r}^0)$, $U_m(\mathbf{r}^0)=Z(\mathbf{r}^0)$, and $k=3$, then $-\mathrm{i}\hbar\,p_{m'm}^{0(k)}/m_e$ is Kane's parameter, P_{Kane}, from equation 130. Usually, the gradient term is neglected, but the envelope function formulation by Zhang (*101*) has incorporated it.

Previously, it was mentioned that the envelope function approximation may not be valid in homogeneously strained regions, since local periodicity may be effectively lost in such regions. However, the envelope function expansion in equation 235 is in terms of the undeformed coordinates, and this corrects for loss of local periodicity. This envelope function expansion may be rewritten as

$$\psi_\alpha(\mathbf{r})=\sum_{m'} f_{m'}^{(\alpha)}\left(\boldsymbol{\chi}^{-1}(\mathbf{r})\right)U_{m'}\left(\boldsymbol{\chi}^{-1}(\mathbf{r})\right). \tag{246}$$

If the original undeformed system was composed of regions that are locally periodic, then the periodicity of $U_{m'}$ in the undeformed coordinate $\mathbf{r}^0$ reflects this. The function $U_{m'}\left(\boldsymbol{\chi}^{-1}(\mathbf{r})\right)$, though, is not in general periodic in $\mathbf{r}$, and the argument $\boldsymbol{\chi}^{-1}(\mathbf{r})$ allows this function to track departures from local periodicity. This is illustrated in figure 27.

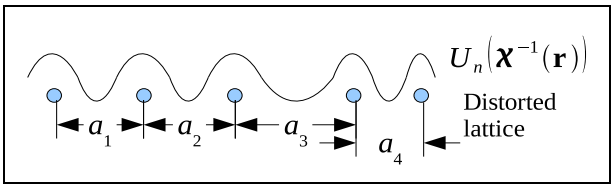

Figure 27. Schematic of a function $U_n(\boldsymbol{\chi}^{-1}(\mathbf{r}))$ superimposed over a distorted lattice where the spacing between neighboring atoms is no longer the same. The inverse deformation map $\boldsymbol{\chi}^{-1}$ allows the function to track the distortions in the lattice.

5. Valence Band Offset

If one is determining the band structure of a single material, the choice of zero-energy datum is largely immaterial. In the example band structures shown in figures 4, 9, and 18, the valence band maximum was arbitrarily chosen to be this zero-energy datum. However, if two materials are brought together, their one-electron energies need to be determined with respect to a common energy datum, and this is often expressed in terms of the valence band offset. That is, if the valence band offset between materials A and B is E_{VBO}, then the valence band maximum in bulk material B is set to be higher than the valence band maximum of A by E_{VBO}. This band offset may be determined by a variety of theoretical and experimental means (*102*), and once it is determined, it may be used to modify the parameters for the various methods. For example, if the fitting parameters for each material lead to band structures where the valence band maximum for each material is zero, then in the tight binding method, the diagonal elements H_{omom} of the Hamiltonian matrix for the atoms in material B are replaced by $H_{omom} + E_{\mathrm{VBO}}$. Similarly, for the envelope function method, diagonal elements $\widehat{H}^{\mathrm{EFA}}_{mm}(\{\mathbf{R}\}^{B}_{\infty})$ would be replaced by $\widehat{H}^{\mathrm{EFA}}_{mm}(\{\mathbf{R}\}^{B}_{\infty}) + E_{\mathrm{VBO}}$. This only has to be done, however, if the fitting parameters do not already take the valence band offset parameters into account. For example, the fitting parameters from Boykin et al. (*33*) for GaAs and InAs lead to valence band maxima of 0.0 and 0.22 for each material, respectively, so an additional offset should not be added to H_{omom} when these parameters are used.

6. Example Results and Discussion

With the help of available data in the literature, we next present results of comparisons among the empirical methods described previously for model systems of (1) bulk GaAs, (2) a slab of InAs, (3) an AlAs/GaAs/AlAs quantum well, and (4) an InAs quantum dot embedded in an InGaAs matrix. We use the plane-wave version of the empirical pseudopotential method of

section 3.1 and the Slater-Koster tight-binding method of section 3.2. The bulk crystal results pertain to the effects of strain on the electronic structure at the Γ point. These results are compared with the corresponding results predicted from a form of the $k \cdot p$ method that accounts for the effects of strain on the valence bands, as discussed in sections 4.2 and 4.5, combined with an effective mass method for bulk crystals, discussed in section 4.4, to account for the conduction band. The particular strain Hamiltonian used for the valence bands is equation 222, while the change in the conduction band minimum with strain is accounted through equation 227. The InAs slab is modeled through the tight-binding method under arbitrary homogeneous strains. And the strain-free GaAs quantum, well clad by two AlAs layers, is modeled by both the tight-binding and envelope-function methods. Finally, results from the tight-binding and envelope-function methods are shown for an InAs quantum dot embedded in a layer of $In_{0.4}Ga_{0.6}As$.

Figure 28 shows a comparison of results from the empirical pseudopotential method (EPM) by Mäder and Zunger (*47*) and a $k \cdot p$ formulation by Van de Walle (*103*), both applied to a bulk crystal of GaAs. The $k \cdot p$ formulation used here is equivalent to using the Chuang $k \cdot p$ Hamiltonian in equation 161 with the deformation potential values from Van de Walle, together with an effective mass approximation for the conduction band. The first subfigure shows the energies of conduction and valence electrons for wave vector $\mathbf{k} = 0$, that is, the band edges calculated by the two methods, while the second subfigure shows the band gaps determined from those energies. Agreement between the two methods is far better for the band gap results than for the electronic energies themselves, which illustrates the problem discussed by Williamson and Zunger (*41*), who found that without an explicitly strain-dependent pseudopotential, the changes in band gaps could be fit to experiment but not the changes in the band edges themselves. The EPM formulation of Mäder and Zunger does not contain any explicitly strain-dependent terms, unlike the formulation of Kim et al. (*50*). The variation with strain of electronic energies due to this latter formulation is shown in figure 29, along with the corresponding $k \cdot p$ results of Van de Walle. The latter formulation better captures the changes in electronic energies due to hydrostatic strain.

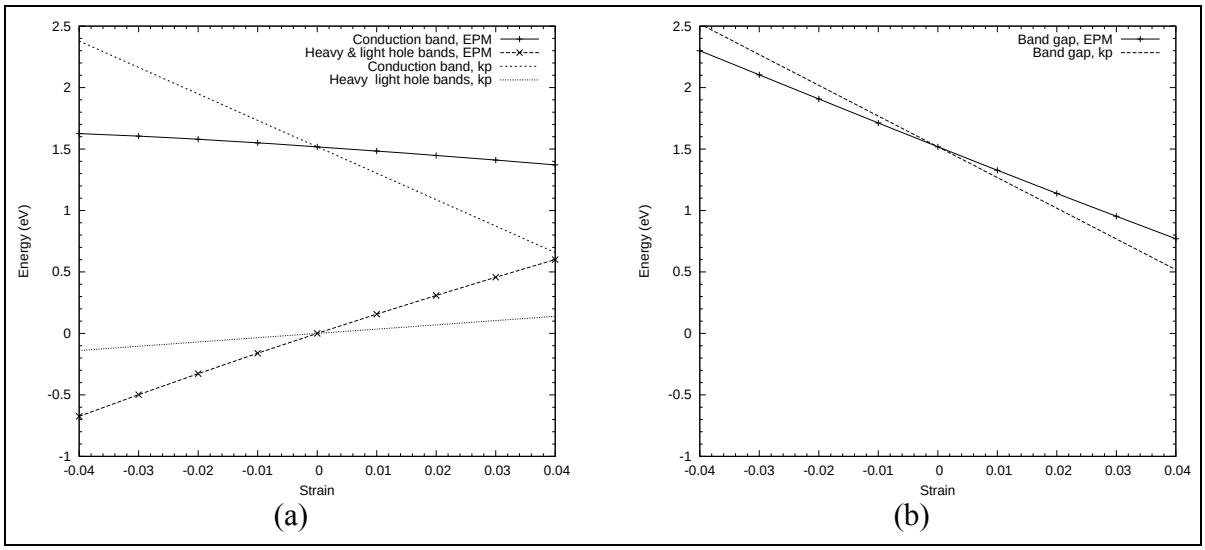

Figure 28. For bulk GaAs, the variation with hydrostatic strain of (a) the energies of conduction and valence electrons for wave vector $\mathbf{k} = 0$ and (b) the band gap, i.e., the *difference* between the energies of conduction and valence electrons, for wave vector $\mathbf{k} = 0$. The electronic energies are calculated using the empirical pseudopotential method (EPM) as formulated by Mäder and Zunger (*47*) and using a $k \cdot p$ formulation from Van de Walle (*103*).

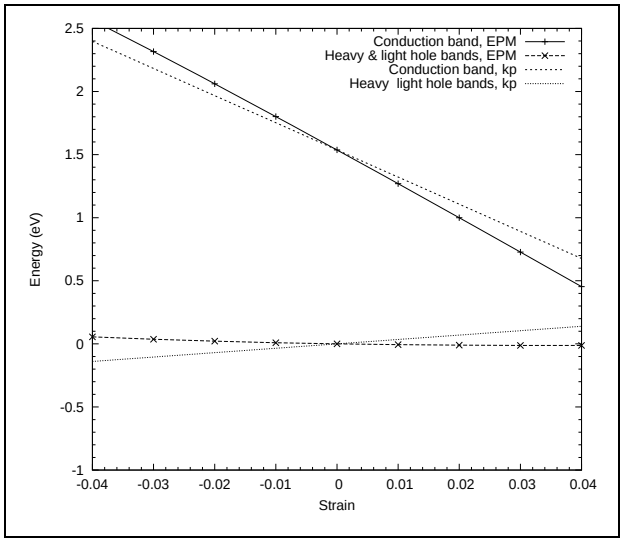

Figure 29. For bulk GaAs, the variation with hydrostatic strain of the energies of conduction and valence electrons for wave vector $\mathbf{k} = 0$, where the electronic energies are calculated using the EPM as formulated by Kim et al. (*50*) and using a $k \cdot p$ formulation from Van de Walle (*103*).

As shown in figure 30, the strain-dependent EPM formulation of Kim, Wang, and Zunger also tracks changes in electronic energies due to strain states besides hydrostatic strain, such as strain due to a biaxial stress state in the plane normal to the (001) crystal direction of GaAs. Here, the z-direction is taken along (001), and the normal stresses point along the x- and y-axes and are equal in magnitude. These stresses are either both tensile or both compressive, so that $\epsilon_{xx} = \epsilon_{yy}$, and the value of ϵ_{zz} is determined from the Poisson effect. The first subfigure compares the EPM results with a spin-free $k \cdot p$ formulation that accounts for the heavy and light-hole bands (*81*), along with the effective mass approximation for the conduction band used earlier. The second subfigure uses Van de Walle's $k \cdot p$ formulation, which includes spin effects. This particular strain-dependent EPM formulation does not account for spin-orbit coupling, and so it agrees better with the spin-free $k \cdot p$ formulation, especially with regard to the light-hole band. Without spin-orbit coupling, the variation of the light-hole band edge with strain is linear. The tight-binding formulation of Boykin et al. (*33*)—which does account for spin—does capture the nonlinearity in the variation from strain due to spin-orbit coupling.

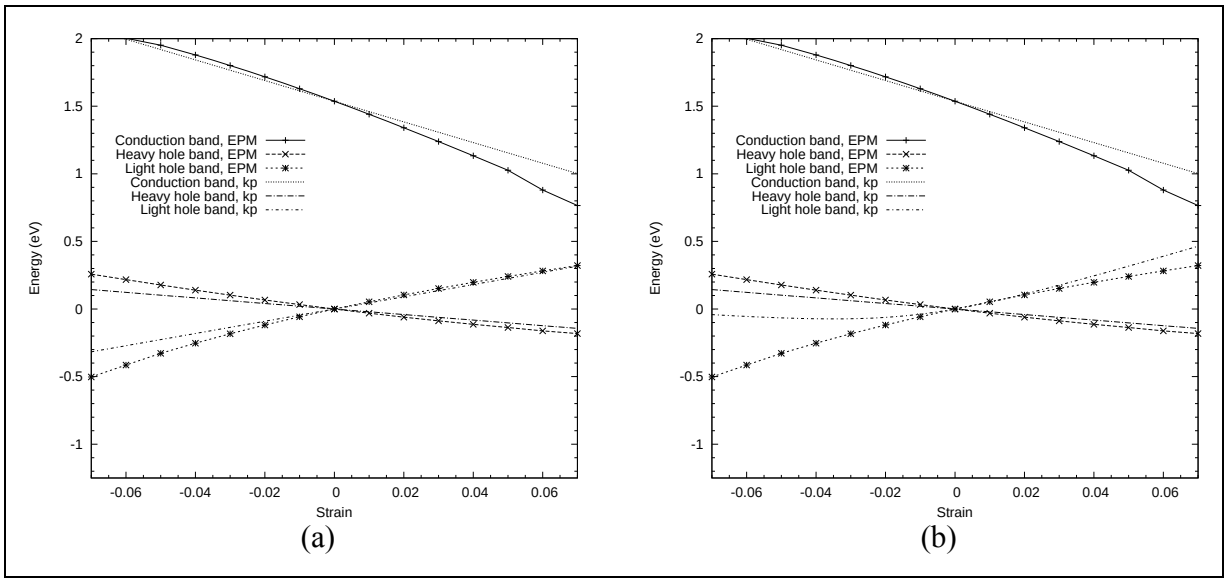

Figure 30. For bulk GaAs, the variation with biaxial stress-induced strain of the energies of conduction and valence electrons for wave vector $\mathbf{k} = 0$, where the electronic energies are calculated with the EPM as formulated by Kim et al. (*50*) and also (a) a spin-free $k \cdot p$ formulation and (b) a $k \cdot p$ formulation with spin. Both $k \cdot p$ formulations use the deformation potential values from Van de Walle (*103*). The strain along the horizontal axis is the strain in the plane normal to the (001) crystal direction.

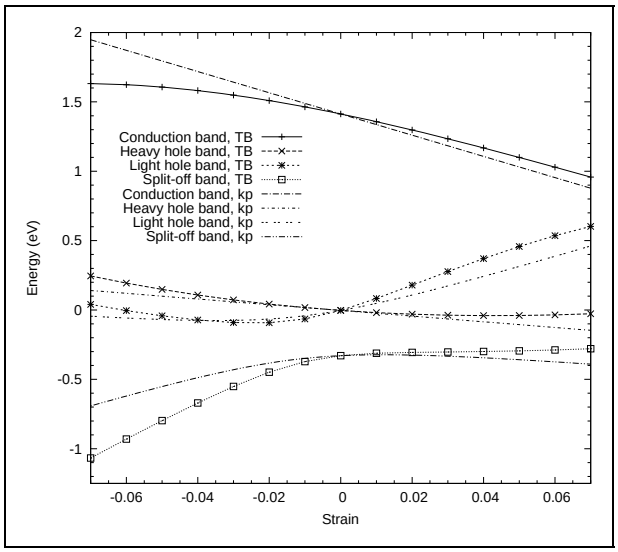

Figure 31. For bulk GaAs, the variation with biaxial stress-induced strain of the energies of conduction and valence electrons for wave vector $\mathbf{k} = 0$, where the electronic energies are calculated with the tight-binding formulation of Boykin et al. (*33*) and also a $k \cdot p$ formulation from Van de Walle (*103*). The strain along the horizontal axis is the strain in the plane normal to the (001) crystal direction.

To further illustrate the effect of strain on electronic structure results, the band structures of a slab of InAs about 1.2 nm thick subjected to in-plane stresses are shown in figure 32. The stresses are along the (100) and (010) directions, leading to a strain of $\epsilon_{\parallel}$ along these same directions. No stress is imposed along the (001) direction. One-particle energies are shown for electrons with Bloch wave vectors pointing along the (100) and (110) crystal directions. The band structure is determined through the tight-binding method. Dangling bonds are again terminated with hydrogens. Essentially, this is an idealized quantum well where electrons are confined not by semiconductor layers but by vacuum, and the strain is allowed to vary arbitrarily rather than be fixed by the lattice mismatch between the material of the well and its surrounding semiconductor. The band gap narrows with increasing in-plane tensile strain and widens with increasing in-plane compressive strain, and increasing in-plane tensile strain also leads to the valence bands bunching together and crossing over one another.

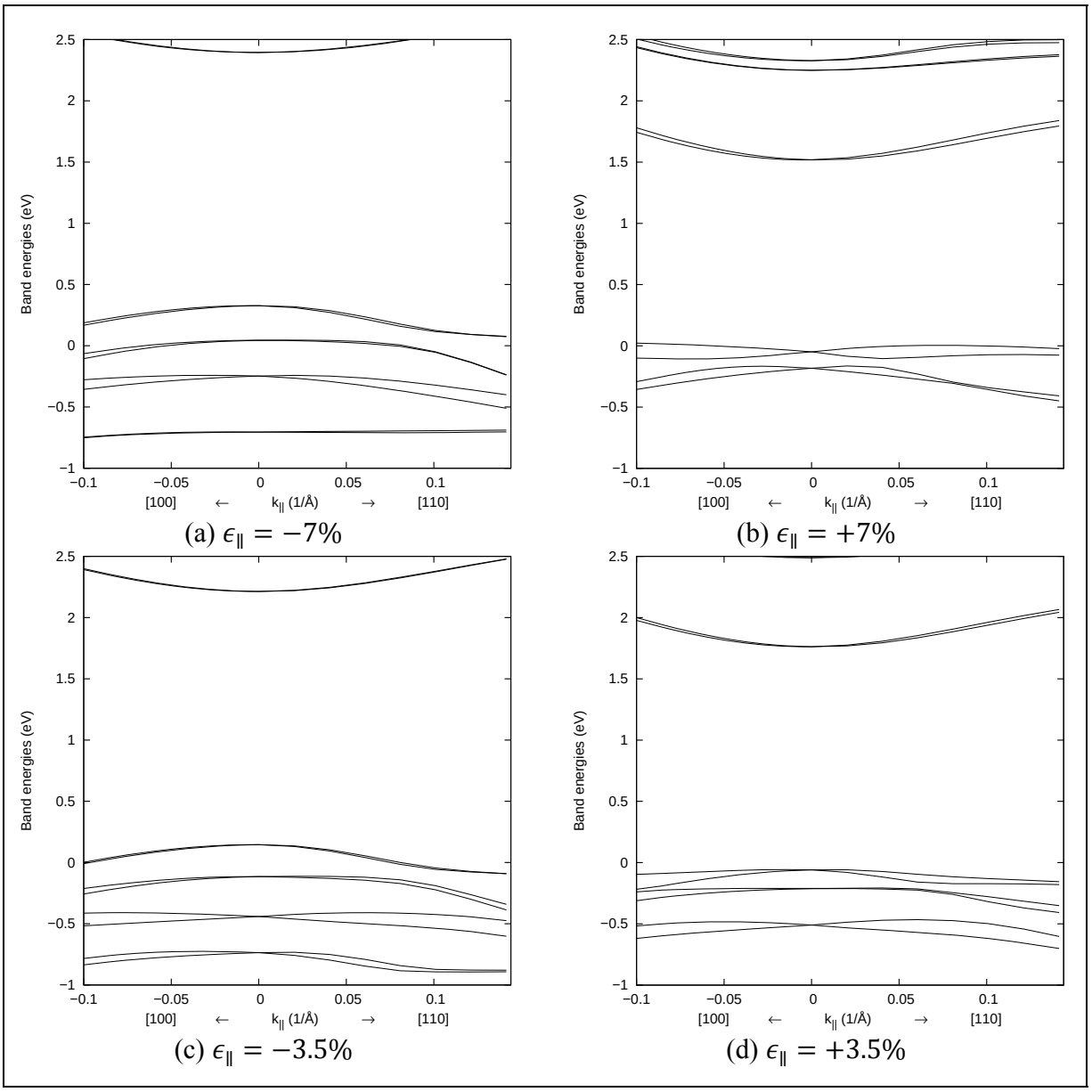

Figure 32. Band structure of a 1.2-nm thick InAs slab subject to stresses along the (100) and (010) directions, leading to a strain of $\epsilon_{\parallel}$ along these same directions. No stress is imposed along the (001) direction. The $k_{\parallel}$ values to the left of zero are the negative magnitudes of Bloch wavevectors pointing along crystal direction (100), while the $k_{\parallel}$ values to the right of zero are the magnitudes of Bloch wavevectors pointing along crystal direction (110).

A comparison of the results of the tight-binding and envelope function methods is shown in figure 33. These methods were used to find the one-particle energies of electrons with Bloch wave vectors pointing along the (100) and (110) crystal directions in a 5-nm-thick GaAs quantum well sandwiched between 30-nm-thick layers of AlAs. Two sets of tight-binding results were generated with NEMO5 modeling code (*35*), one using parameters from Jancu et al. (*104*) and one using parameters from the NEMO5 material database. The GaAs parameters from

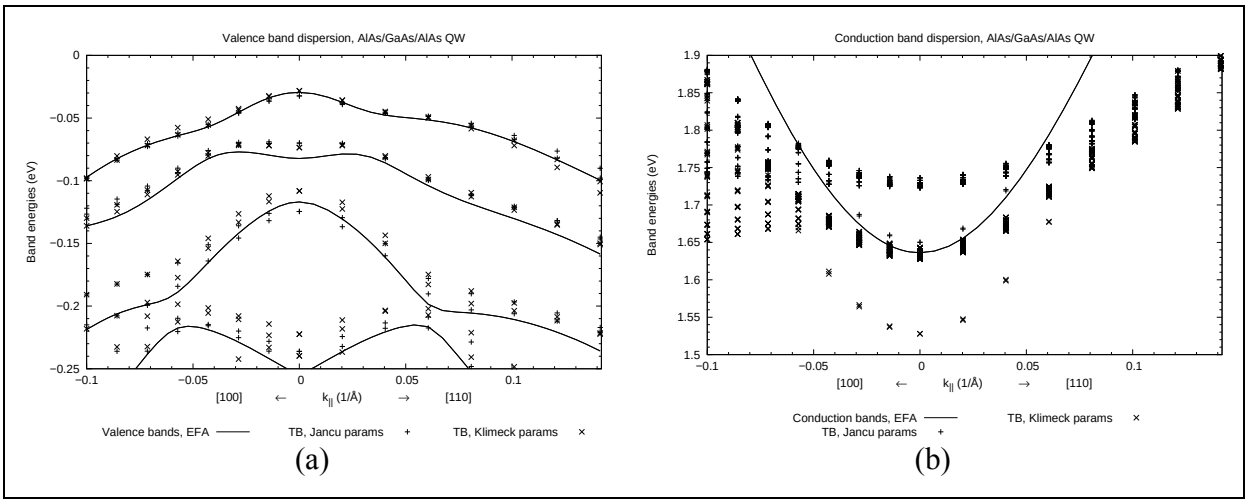

Figure 33. Comparison of results from tight-binding and envelope function methods for a 5-nm-thick GaAs quantum well sandwiched between 30-nm-thick layers of AlAs. Valence bands are shown in (a) and conduction bands in (b). Two sets of tight-binding parameters are used, one from Jancu et al. (*104*) and one from used in the material database of the tight-binding software code NEMO5 and attributed to one of its authors, G. Klimeck (*35*). The 8×8 $k \cdot p$ Hamiltonian used by the envelope function calculations may be found in the documentation of nextnano (*105*) or Andlauer (*106*). The $k_{\|}$ values to the left of zero are the negative magnitudes of Bloch wavevectors pointing along crystal direction (100), while the $k_{\|}$ values to the right of zero are the magnitudes of Bloch wavevectors pointing along crystal direction (110).

this database may be found in Boykin et al. (*33*), and the AlAs parameters from this database are shown in table 4. A valence band offset of between AlAs and GaAs was needed when the parameters from Jancu et al. were used, and the valence band maximum of AlAs was taken to be 0.5 eV lower than that of GaAs. However, the parameters from the NEMO5 database already incorporated a valence band offset. Boundary conditions were periodic only along the directions in the plane of the well, (100) and (010), and dangling bonds were terminated with hydrogen atoms. The envelope function results were generated from nextnano simulation software using an 8×8 $k \cdot p$ Hamiltonian matrix documented on the nextnano web site (*105*) and by Andlauer (*106*). Agreement between the two methods is good for the valence bands, especially the top two bands. For the conduction bands, the tight-binding results using the parameters of Jancu et al. only agree with the envelope function results for small wave vector values, and the tight-binding results using the other set of parameters lead to a conduction band minimum about 0.1 eV less than that predicted from the envelope function method. For the valence bands, a similar agreement between tight-binding and envelope function results has been demonstrated by de Franceschi et al. (*107*), who compared tight-binding results using the parameters of Jancu et al. with results from envelope function results using a 6×6 $k \cdot p$ Hamiltonian matrix.

Table 4. Tight-binding parameters for AlAs from the material database of NEMO5 (*35*). These parameters are values of $E^{(0)}_{oZ_m}$ and $V^{(0)}_{o'm',om,\lambda}$ from equations 93 and 97, and the spin parameter λ_m in equation 88 for $o', o \in \{s, p, d, s^*\}$ and $m', m \in \{\mathrm{Al}, \mathrm{As}\}$.

$E^{(0)}_{sZ_{\mathrm{Al}}}$	0.79695	$V^{(0)}_{s\mathrm{Al},s\mathrm{As},\sigma}$, $V^{(0)}_{s\mathrm{As},s\mathrm{Al},\sigma}$	–1.64584
$E^{(0)}_{sZ_{\mathrm{As}}}$	–5.17012	$V^{(0)}_{p\mathrm{Al},p\mathrm{As},\sigma}$, $V^{(0)}_{p\mathrm{As},p\mathrm{Al},\sigma}$	4.53156
$E^{(0)}_{pZ_{\mathrm{Al}}}$	6.63291	$V^{(0)}_{p\mathrm{Al},p\mathrm{As},\pi}$, $V^{(0)}_{p\mathrm{As},p\mathrm{Al},\pi}$	–1.86816
$E^{(0)}_{pZ_{\mathrm{As}}}$	4.39708	$V^{(0)}_{d\mathrm{Al},d\mathrm{As},\sigma}$, $V^{(0)}_{d\mathrm{As},d\mathrm{Al},\sigma}$	–1.97058
$E^{(0)}_{dZ_{\mathrm{Al}}}$	12.92120	$V^{(0)}_{d\mathrm{Al},d\mathrm{As},\pi}$, $V^{(0)}_{d\mathrm{As},d\mathrm{Al},\pi}$	1.67733
$E^{(0)}_{dZ_{\mathrm{As}}}$	13.13880	$V^{(0)}_{d\mathrm{Al},d\mathrm{As},\delta}$, $V^{(0)}_{d\mathrm{As},d\mathrm{Al},\delta}$	–1.58868
$E^{(0)}_{s^*Z_{\mathrm{Al}}}$	24.16590	$V^{(0)}_{s^*\mathrm{Al},s^*\mathrm{As},\sigma}$, $V^{(0)}_{s^*\mathrm{As},s^*\mathrm{Al},\sigma}$	–2.84245
$E^{(0)}_{s^*Z_{\mathrm{As}}}$	19.80470	—	—
$V^{(0)}_{s\mathrm{Al},p\mathrm{As},\sigma}$	2.95309	$V^{(0)}_{s\mathrm{As},p\mathrm{Al},\sigma}$	3.02223
$V^{(0)}_{s\mathrm{Al},d\mathrm{As},\sigma}$	–2.64111	$V^{(0)}_{s\mathrm{As},d\mathrm{Al},\sigma}$	–3.03196
$V^{(0)}_{s\mathrm{Al},s^*\mathrm{As},\sigma}$	–1.88341	$V^{(0)}_{s\mathrm{As},s^*\mathrm{Al},\sigma}$	–2.78690
$V^{(0)}_{p\mathrm{Al},d\mathrm{As},\sigma}$	–1.02836	$V^{(0)}_{p\mathrm{As},d\mathrm{Al},\sigma}$	–2.47345
$V^{(0)}_{p\mathrm{Al},d\mathrm{As},\pi}$	2.86419	$V^{(0)}_{p\mathrm{As},d\mathrm{Al},\pi}$	2.52741
$V^{(0)}_{s^*\mathrm{Al},p\mathrm{As},\sigma}$	1.30469	$V^{(0)}_{s^*\mathrm{As},p\mathrm{Al},\sigma}$	1.92174
$V^{(0)}_{s^*\mathrm{Al},d\mathrm{As},\sigma}$	–1.73510	$V^{(0)}_{s^*\mathrm{As},d\mathrm{Al},\sigma}$	–1.84300
λ_{Al}	0.01586	λ_{As}	0.17386

A comparison of the results of the tight-binding and envelope function methods is shown in figure 33. These methods were used to find the one-particle energies of electrons with Bloch wave vectors pointing along the (100) and (110) crystal directions in a 5-nm-thick GaAs quantum well sandwiched between 30 nm-thick layers of AlAs. Two sets of tight-binding results were generated with NEMO5 (*35*), one using parameters from Jancu et al. (*104*) and one using parameters from the NEMO5 material database. The GaAs parameters from this database may be found in Boykin et al. (*33*), and the AlAs parameters from this database are shown in table 4. A valence band offset of between AlAs and GaAs was needed when the parameters from Jancu et al. were used, and the valence band maximum of AlAs was taken to be 0.5 eV lower than that of GaAs. However, the parameters from the NEMO5 database already incorporated a valence band offset. Boundary conditions were periodic only along the directions in the plane of the well, (100) and (010), and dangling bonds were terminated with hydrogen atoms. The envelope function results were generated from nextnano using an 8×8 $k \cdot p$ Hamiltonian matrix

documented on the nextnano web site (*105*) and by Andlauer (*106*). Agreement between the two methods is good for the valence bands, especially the top two bands. For the conduction bands, the tight-binding results using the parameters of Jancu et al. only agree with the envelope function results for small wave vector values, and the tight-binding results using the other set of parameters lead to a conduction band minimum about 0.1 eV less than that predicted from the envelope function method. For the valence bands, a similar agreement between tight-binding and envelope function results has been demonstrated by de Franceschi et al. (*107*), who compared tight-binding results using the parameters of Jancu et al. with results from envelope function results using a 6×6 $k \cdot p$ Hamiltonian matrix.

Comparisons between tight-binding and envelope function results have also been made for a dome-shaped InAs quantum dot by Sengupta et al. (*108*). The dot was 5 nm in height, 20 nm in diameter, and embedded in a 5-nm layer of $In_{0.4}Ga_{0.6}As$, which in turn was sandwiched between 30-nm layers of GaAs. The strain in the quantum dot was determined through two different atomistic valence force field (VFF) models, the harmonic Keating model (*109*) and an anharmonic model (*110*). Results from the tight-binding and envelope function methods are shown in table 5. The theoretical results were compared to experimental results for a similar quantum dot whose peak photoluminescence occurred at the wavelength $\lambda = 1.52$ μm (*111*), corresponding to a band gap of about $E_g \approx 2\pi\hbar c/\lambda \approx 0.82$ eV. Agreement between the two theoretical methods, tight-binding and the envelope function approximation, is good for both methods of determining the strain, but agreement with experiment is much better when the anharmonic VFF model is used to determine strain.

Table 5. Results for the quantum dot studied by Sengupta et al. (*108*). Dot is 5 nm in height, 20 nm in diameter, and embedded in a 5-nm layer of $In_{0.4}Ga_{0.6}As$, sandwiched between two 30-nm GaAs layers.

Model	Band Gap (eV)
Envelope function, harmonic VFF	1.063
Tight-binding, harmonic VFF	1.040
Envelope function, anharmonic VFF	0.885
Tight-binding, anharmonic VFF	0.828

7. Conclusion

Various methods for estimating electronic structure have been discussed, along with discussions of how strain is incorporated into these methods. We did not find—among the methods both included and not included in this survey—a generalized and explicit treatment of continuum deformability. Thus a gap exists between the methods suited to handle slowly varying elastic fields and fully atomistic approaches. Atomistic approaches offer the greatest generality but lack

an ability to treat deformation explicitly without recourse to a minimization of atomic scale forces and its attendant computational expense. They categorically require a full recalculation over all atomistic degrees of freedom based on each new configuration of nuclei, which therefore scales according to the number of atoms, electrons, or electronic wavefunctions regardless of the degree to which the deforming material exhibits continuum-like behavior. Continuum-based methods incorporate notions of stress, but rely exclusively on linear elasticity or infinitesimal strain theories. Elastic fields are, at best, mildly inhomogeneous. Namely, finite deformation effects such as near dislocation or vacancy cores, particularly for complex lattices, are not presently possible. Results for finite strain effects, therefore, require a separate atomistic calculation for verification—first to minimize the energy of the nuclear configuration in the deformed state and secondly, in some cases, to determine the resulting electronic structure. For systems whose sizes simultaneously need to be sufficiently large to capture the convergence of the elastic field solution to the bulk limit while also providing the required resolution near features of interest, such as at interfaces and defects, single point calculations alone may be computationally costly and calculations over multiple mechanically deformed states may be prohibitive.

In *ab initio* methods, strain is taken into account simply through the positions of the atoms in the system being simulated. However, in atomistic empirical methods, the positions of the atoms are not always sufficient to fully take the effects of strain into account. In the empirical pseudopotential method, a strain-dependent prefactor in the atomic form factors has been used to better account for the effects of strain on band edges (*41*). In the tight-binding method, atomic positions are usually sufficient to account for the effects of strain. However, if there is long-range charge transfer, such as that due to piezoelectric material properties, then the resulting electric field has to be treated like an external field in these atomistic methods, and in the case of piezoelectricity, this field is determined through continuum mechanical calculations (*34*, *78*).

The $k \cdot p$ and envelope function methods are empirical electronic structure methods that do not depend at all on the positions of the atoms in the simulation. Rather, the potential felt by an electron in a region of the system composed of some material is assumed to be the potential in a bulk crystal of the material. Because of this, they are less computationally expensive than the atomistic methods, but the positions of the atoms can no longer be used to take strain into account. Instead, a strain-dependent term is added to the one-electron Hamiltonian. This term was originally derived by Bir and Pikus (*1*) for an infinite bulk subjected to homogeneous strain. However, the term can be derived in an alternative fashion for more general states of strain, though the gradient of the strain still needs to be small (*100*).

Finally, examples of the results from the empirical electronic structure methods discussed have been shown. Generally, when the methods are applied to the same problem and each is examined within the assumptions unique to each method, the results are mostly in agreement.

8. References

1. Bir, G. L.; Pikus, G. E. *Symmetry and Strain-Induced Effects in Semiconductors*; Halstead Press: New York, 1974.
2. Bohr, M. *Invention of Uniaxial Strained Silicon Transistors at Intel*; Intel Corporation: Santa Clara, CA, 2007.
3. Ledentsov, N. N.; Grundmann, M.; Heinrichsdorff, F.; Bimberg, D.; Ustinov, V. M.; Zhukov, A. E.; Maximov, M. V.; Alferov, Z. I.; Lott, J. A. Quantum-Dot Heterostructure Lasers. *IEEE J. of Selected Topics in Quantum Electronics* **2000**, *6*, 439–451.
4. Henini, M.; Bugajski, M. Advances in Self-Assembled Semiconductor Quantum Dot Lasers. *Microelectronics Journal* **2005**, *36*, 950–956.
5. QD Laser, Inc. QDLaser. http://qdlaser.com (accessed 19 March 2012).
6. Adachi, S. *Handbook on Physical Properties of Semiconductors*, Vol. 2; Kluwer Academic Publishers: Norwell, MA, 2004.
7. Wang, J.; Jeon, J. B.; Sirenko, Y. M.; Kim, K. W. Piezoelectric Effect on Optical Properties of Pseudomorphically Strained Wurtzite GaN Quantum Wells. *IEEE Photonics Tech. Letters* **1997**, *9*, 728–730.
8. Park, S.-H.; Chuang S.-L. Piezoelectric Effects on Electrical and Optical Properties of Wurtzite GaN/AlGaN Quantum Well Lasers. *Applied Phys. Lett.* **1998**, *72*, 3103–3105.
9. Vurgaftman, I.; Hinckley, J. M.; Singh, J. A Comparison of Optoelectronic Properties of Lattice-Matched and Strained Quantum-Well and Quantum-Wire Structures. *IEEE J. of Quantum Elec.* **1994**, *30*, 75–84.
10. Harrison, P.; Ikonić, Z.; Jovanović, V.; Califano, M.; Evans, C. A.; Indjin, D. *Quantum Wells, Wires, and Dots*; 3rd ed.; John Wiley and Sons: Chichester, UK, 2009.
11. Levine, I. N. *Quantum Chemistry*; 4th ed.; Prentice-Hall: Englewood Cliffs, NJ, 1991.
12. Ashcroft, N. W.; Mermin, N. D. *Solid State Physics*; Brooks Cole: Stamford, CT, 1976.
13. Aulbur, W. G.; Jönsson, L.; Wilkins, J. W. Quasiparticle Calculations in Solids. *Solid State Phys.* **1999**, *58*, 1–218.
14. Fox, M. *Quantum Optics: An Introduction*; Oxford University Press: New York, 2006.
15. Onida, G.; Reining, L.; Rubio, A. Electronic Excitations: Density-Functional versus Many-body Green's Function Approaches. *Reviews of Modern Phys.* **2002**, *74*, 601–659.

16. ABINIT Group. Notes on the Bethe-Salpeter Equation. http://www.abinit.org /documentation/helpfiles/for-v6.12/tutorial/theory_bse.html (accessed 24 August 2012).

17. Davies, J. H. *The Physics of Low-Dimensional Semiconductors*; Cambridge University Press: New York, 1998.

18. Vogl, P.; Hjalmarson, H. P.; Dow, J. D. A Semi-Empirical Tight-Binding Theory of the Electronic Structure of Semiconductors. *J. of the Physical Chem. of Solids* **1983**, *44*, 365–378.

19. Mehl, M. J.; Papaconstantopoulos, D. A. Applications of a Tight-Binding Total-Energy Method for Transition and Noble Metals: Elastic Constants, Vacancies, and Surfaces of Monatomic Metals. *Physical Review B* **1996**, *54*, 4519–4530.

20. Mehl, M. STATIC: A Scalable Tight-Binding Total Energy Evaluation Code. http://cst-www.nrl.navy.mil/bind/static (accessed 18 July 2012).

21. Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods*; Cambridge University Press: Cambridge, UK, 2004.

22. Nogueira, F.; Castro, A.; Marques, M. A. L. A Tutorial on Density Functional Theory. In *A Primer in Density Functional Theory*; Springer-Verlag: New York, 2003; pp 218–256.

23. Ram-Mohan, L. R. *Finite Element and Boundary Element Applications in Quantum Mechanics*; Oxford University Press: New York, 2002.

24. Roothan, C. C. J. New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951**, *23*, 69–89.

25. Perdew, J. P.; Kurth, S. Density Functionals for Non-Relativistic Coulomb Systems in the New Century. In *A Primer in Density Functional Theory*; Springer-Verlag: Berlin, Heidelberg, 2003; pp 1–55.

26. Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965**, *140*, A1133–A1138.

27. Marques, M. A. L.; Gross, E. K. U. Time-Dependent Density Functional Theory. In *A Primer in Density Functional Theory*; Springer-Verlag: Berlin, Heidelberg, 2003; pp 144–184.

28. Sholl, D.; Steckel, J. A. *Density Functional Theory: A Practical Introduction*; John Wiley and Sons: Hoboken, NJ, 2009.

29. Lambrecht, W. R. L. Which Electronic Structure Method for the Study of Defects: A Commentary. *Physica Status Solidi (B)* **2011**, *248*, 1547–1558.

30. Nieminen, R. M. Supercell Methods for Defect Calculations. In *Theory of Defects in Semiconductors*; Springer-Verlag: Berlin, Heidelberg, 2007; pp 29–68.

31. Miller, D. A. B. *Quantum Mechanics for Scientists and Engineers*; Cambridge University Press: New York, 2008.

32. Chuang, S. L. *Physics of Optoelectronic Devices*; John Wiley and Sons: New York, 1995.

33. Boykin, T. B.; Klimeck, G.; Bowen, C.; Oyafuso, F. Diagonal Parameter Shifts Due to Nearest-Neighbor Displacements in Empirical Tight-Binding Theory. *Physical Review B* 2002, *66*, 125207.

34. Bester, G. Electronic Excitations in Nanostructures: An Empirical Pseudopotential Based Approach. *J. of Physics: Condensed Matter* **2009**, *21*, 023202.

35. Steiger, S.; Povolotskyi, M.; Park, H.-H.; Kubis, T.; Klimeck, G. NEMO5: Parallel Multiscale Nanoelectronics Modeling Tool. *IEEE Transactions on Nanotechnology* **2011**, *10*, 1464–1474.

36. Phillips, J. C. Energy-Band Interpolation Scheme Based on a Pseudopotential. *Physical Review* **1958**, *112*, 685–695.

37. Cohen, M. L.; Bergstresser, T. K. Band Structures and Pseudopotential Form Factors for Fourteen Semiconductors of the Diamond and Zinc-Blende Structures. *Physical Review* **1966**, *141*, 789–796.

38. Chelikowsky, J. R.; Cohen, M. L. Nonlocal Pseudopotential Calculations for the Electronic Structure of Eleven Diamond and Zinc-Blende Semiconductors. *Physical Review B* **1976**, *14*, 556–582.

39. Kreyszig, E. *Advanced Engineering Mathematics*; 7th ed.; John Wiley and Sons: New York, 1993.

40. Kim, K.; Kent, P. R. C.; Zunger, A.; Geller, C. B. Atomistic Description of the Electronic Structure of In(x)Ga(1-x)As Alloys and InAs/GaAs Superlattices. *Physical Review B* **2002**, *66*, 045208.

41. Williamson, A. J.; Zunger, A. InAs Quantum Dots: Predicted Electronic Structure of Free-Standing versus GaAs-Embedded Structures. *Physical Review B* **1999**, *59*, 15819–15824.

42. Williamson, A. J.; Wang, L. W.; Zunger, A. Theoretical Interpretation of the Experimental Electronic Structure of Lens-Shaped Self-Assembled InAs/GaAs Quantum Dots. *Physical Review B* **2000**, *62*, 12963–12977.

43. Aouina, N. Y.; Mezrag, F.; Boucenna, M.; El-Farra, M.; Bouarissa, N. High Pressure Electronic Properties and Elastic Stability Criteria of AlAs. *Materials Science and Engineering B* **2005**, *123*, 87–93.

44. Caruthers, E.; Lin-Chung, P. J. Pseudopotential Calculations for (GaAs)1-(AlAs)1 and Related Monolayer Heterostructures. *Physical Review B* **1978**, *17*, 2705–2717.

45. Jmol Project. Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org (accessed 23 March 2012).

46. Wang, L.-W.; Zunger, A. Electronic Structure Pseudopotential Calculations of Large (~1000 Atoms) Si Quantum Dots. *J. of Physical Chem.* **1994**, *98*, 2158–2165.

47. Mäder, K. A.; Zunger, A. Empirical Atomic Pseudopotentials for AlAs/GaAs Superlattices, Alloys, and Nanostructures. *Physical Review B* **1994**, *50*, 17393–17405.

48. Wang, L.-W.; Franceschetti, A.; Zunger, A. Million-Atom Pseudopotential Calculation of Γ-X Mixing in GaAs/AlAs Superlattices and Quantum Dots. *Physical Review Letters* **1997**, *78*, 2819–2822.

49. Wang, L.-W.; Zunger, A. Linear Combination of Bulk Bands Method for Large-Scale Electronic Structure Calculations on Strained Nanostructures. *Physical Review B* **1999**, *59*, 15806–15818.

50. Kim, J.; Wang, L.-W.; Zunger, A. Comparison of the Electronic Structure of InAs/GaAs Pyramidal Quantum Dots With Different Facet Orientations. *Physical Review B* **1998**, *57*, R9408–R9411.

51. Wang, L.-W.; Kim, J.; Zunger, A. Electronic Structures of [110]-Faceted Self-Assembled Pyramidal InAs/GaAs Quantum Dots. *Physical Review B* **1999**, *59*, 5678–5687.

52. Canning, A.; Wang, L. W.; Williamson, A.; Zunger, A. Parallel Empirical Pseudopotential Electronic Structure Calculations for Million Atom Systems. *J. of Computational Physics* **2000**, *160*, 29–41.

53. Wang, L.-W. SLCBB Webpage. https://hpcrd.lbl.gov/~linwang/SLCBB_soft.html (accessed 16 August 2012).

54. Slater, J. C.; Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Physical Review* **1954**, *94*, 1498–1524.

55. Harrison, W. A. *Electronic Structure and the Properties of Solids*; Dover Publications: New York, 1989.

56. Klimeck, G.; Oyafuso, F.; Boykin, T. B.; Bowen; R. C.; von Allmen, P. Development of a Nanoelectronic 3-D (NEMO 3-D) Simulator for Multimillion Atom Simulations and Its Application to Alloyed Quantum Dots. *CMES: Computer Modeling in Engineering and Sciences* **2002**, *3*, 601–642.

57. Pantelides, S. T.; Pollmann, J. Critique of the Empirical Tight-Binding Method for Semiconductor Surfaces and Interfaces. *J. of Vacuum Science and Tech.* **1979**, *16*, 1349–1358.

58. Liu, F. Self-Consistent Tight-Binding Method. *Physical Review B* **1995**, *52*, 10677–10680.

59. Esfarjani, K.; Kawazoe, Y. Self-Consistent Tight-Binding Formalism for Charged Systems. *J. of Physics: Condensed Matter* **1998**, *10*, 8257–8267.

60. Hamaguchi, C. *Basic Semiconductor Physics*; Springer-Verlag: Berlin, Heidelberg, 2010.

61. Löwdin, P.-O. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *J. of Chem. Physics* **1950**, *18*, 365–375.

62. Colombo, L. Tight-Binding Molecular Dynamics: A Primer. *Rivista del Nuovo Cimento* **2005**, *28*, 1–59.

63. Ercolessi, F. Lecture Notes on Tight-Binding Molecular Dynamics, and Tight-Binding Justification of Classical Potentials. http://www.ud.infn.it/~ercolessi/SA/tb.pdf (accessed 28 March 2012).

64. Papaconstantopoulos, D. A.; Mehl, M. J. The Slater–Koster tight-binding method: a computationally efficient and accurate approach. *J. of Physics: Condensed Matter* **2003**, *15*, R413–R440.

65. Sutton, A. P. *Electronic Structure of Materials*; Oxford University Press: New York, 1993.

66. Hjelmstad, K. D. *Fundamentals of Structural Mechanics*; Springer Science: New York, 2005.

67. Papaconstantopoulos, D. A.; Mehl, M. J.; Erwin, S. C.; Pederson, M. R. Tight-Binding Hamiltonians for Carbon and Silicon. In *Tight-Binding Approach to Computational Materials Science*; Turchi, P., Gonis A., Colombo, L., Eds.; Vol. 491; Materials Research Society: Warrendale, PA, 1998; pp 221–230.

68. Papaconstantopoulos, D. A.; Lach-hab, M.; Mehl, M. J. Tight-binding Hamiltonians for Realistic Electronic Structure Calculations. *Physica B* **2001**, *296*, 129–137.

69. Chadi, D. J. Spin-Orbit Splitting in Crystalline and Compositionally Disordered Semiconductors. *Physical Review B* **1977**, *16*, 790–796.

70. Hass, K. C.; Ehrenreich, H.; Velický, B. Electronic Structure of Hg(1-x)Cd(x)Te. *Physical Review B* **1983**, *27*, 1088–1100.

71. Di Carlo, A. Microscopic Theory of Nanostructured Semiconductor Devices: Beyond the Envelope-Function Approximation. *Semiconductor Science and Tech.* **2003**, *18*, R1–R31.

72. Chadi, D. J.; Cohen, M. L. Tight-Binding Calculations of the Valence Bands of Diamond and Zincblende Crystals. *Physica Status Solidi (B)* **1975**, *68*, 405–419.

73. Jancu, J. M.; Voisin, P. Tetragonal and Trigonal Deformations in Zinc-Blende Semiconductors: A Tight-Binding Point of View. *Physical Review B* **2007**, *76*, 115202.

74. Niquet, Y. M.; Rideau, D.; Tavernier, C.; Jaouen, H.; Blase, X. Onsite Matrix Elements of the Tight-Binding Hamiltonian of a Strained Crystal: Application to Silicon, Germanium, and Their Alloys. *Physical Review B* **2009**, *79*, 245201.

75. Fonoberov, V. A.; Balandin, A. A. Properties of GaN and ZnO Quantum Dots. In *Handbook of Semiconductor Nanostructures and Nanodevices*; American Scientific Publishers: Valencia, CA, 2006; pp 119–158.

76. Morkoç H.; Leach, J. Polarization in GaN Based Heterostructures and Polarization in GaN Based Heterostructures (HFETs). In *Polarization Effects in Semiconductors: From Ab Initio Theory to Device Applications*; Springer Science: New York, 2008; pp 373–466.

77. Bester, G.; Wu, X.; Vanderbilt, D.; Zunger, A. Importance of Second-Order Piezoelectric Effects in Zinc-Blende Semiconductors. *Physical Review Letters* **2006**, *96*, 187602.

78. Saito, T.; Arakawa, Y. Electronic Structure of Piezoelectric In(0.2)Ga(0.8)N Quantum Dots in GaN Calculated Using a Tight-Binding Method. *Physica E: Low-Dimensional Systems and Nanostructures* **2002**, *15*, 169–181.

79. Hong, K. B.; Kuo, M. K. Fully Coupled and Semi-Coupled Piezoelectric Models on the Optical Properties of InGaN Quantum Dots. *Semiconductor Science and Tech.* **2010**, *25*, 065005.

80. Fonoberov, V. A.; Balandin, A. A. Optical Properties of Wurtzite and Zinc-Blende GaN/AlN Quantum Dots. *J. of Vacuum Science and Tech. B* **2004**, *22*, 2190–2194.

81. Voon, L. C. L. Y.; Willatzen, M. *The k·p Method: Electronic Properties of Semiconductors*; Springer-Verlag: New York, 2009.

82. Luttinger, J. M.; Kohn, W. Motion of Electrons and Holes in Perturbed Periodic Fields. *Physical Review* **1955**, *97*, 869–883.

83. Kane, E. O. Band Structure of Indium Antimonide. *J. of Phys. and Chem. of Solids* **1957**, *1*, 249–261.

84. Vurgaftman, I. J.; Meyer, R.; Ram-Mohan, L. R. Band Parameters for III-V Compound Semiconductors and Their Alloys. *J. of Applied Physics* **2001**, *89*, 5815–5875.

85. Mahan, G. D. *Quantum Mechanics in a Nutshell*; Princeton University Press: Princeton, NJ, 2009.

86. Bastard, G. *Wave Mechanics Applied to Semiconductor Heterostructures*; Les Éditions de Physique: Les Ulis Cedex, 1988.

87. Burt, M. G. The Justification for Applying the Effective-Mass Approximation to Microstructures. *J. of Physics: Condensed Matter* **1992**, *4*, 6651–6690.

88. Stier, O.; Bimberg, D. Modeling of Strained Quantum Wires Using Eight-Band k·p Theory. *Physical Review B* **1997**, *55*, 7726–7732.

89. Veprek, R.; Steiger, S.; Witzigmann, B. Reliable k·p Band Structure Calculation for Nanostructures Using Finite Elements. *J. of Computational Electronics* **2008**, *7*, 521–529.

90. Chao, C. Y.-P.; Chuang, S. L. Spin-Orbit-Coupling Effects on the Valence-Band Structure of Strained Semiconductor Quantum Wells. *Physical Review B* **1992**, *46*, 4110–4122.

91. Lassen, B.; Voon, L. C. L. Y.; Willatzen, M.; Melnik, R. Exact Envelope-Function Theory Versus Symmetrized Hamiltonian for Quantum Wires: A Comparison. *Solid State Communications* **2004**, *132*, 141–149.

92. Foreman, B. A. Elimination of Spurious Solutions From Eight-Band k·p Theory. *Physical Review B* **1997**, *56*, R12748–R12751.

93. Foreman, B. A. Effective-Mass Hamiltonian and Boundary Conditions for the Valence Bands of Semiconductor Microstructures. *Physical Review B* **1993**, *48*, 4964–4967.

94. Marquardt, O.; Mourad, D.; Schulz, S.; Hickel, T.; Czycholl, G.; Neugebauer, J. Comparison of Atomistic and Continuum Theoretical Approaches to Determine Electronic Properties of GaN/AlN Quantum Dots. *Physical Review B* **2008**, *78*, 235302.

95. Foreman, B. A. Envelope-Function Formalism for Electrons in Abrupt Heterostructures with Material-Dependent Basis Functions, *Physical Review B* **1996**, *54*, 1909–1921.

96. Mase, G. T.; Mase, G. E. *Continuum Mechanics for Engineers*; 2nd ed., CRC Press: Boca Raton, FL, 1999.

97. Sweeney, S.; Adams, A. Optoelectronic Devices and Materials. In *Springer Handbook of Electronic and Photonic Materials*; Kasap, S., Capper, P., Koughia, C., Eds.; Springer Science: New York, 2006; pp 887–916.

98. Pryor, C.; Kim, J.; Wang, L. W.; Williamson, A. J.; Zunger, A. Comparison of Two Methods For Describing the Strain Profiles In Quantum Dots. *J. of Applied Physics* **1998**, *83*, 2548–2554.

99. Blount, E. I. Ultrasonic Attenuation by Electrons in Metals. *Physical Review* **1959**, *114*, 418–436.

100. Sham, L. J.; Ziman, J. M. The Electron-Phonon Interaction. *Solid State Physics* **1963**, *15*, 221–298.

101. Zhang, Y. Motion of Electrons In Semiconductors Under Inhomogeneous Strain with Application to Laterally Confined Quantum Wells. *Physical Review B* **1994**, *49*, 14352–14366.

102. Franciosi, A.; van de Walle, C. G. Heterojunction Band Offset Engineering. *Surface Science Reports* **1996**, *25*, 1–140.

103. Van de Walle, C. Band Lineups and Deformation Potentials in the Model-Solid Theory. *Physical Review B* **1989**, *39*, 1871–1883.

104. Jancu, J.-M.; Scholz, R.; Beltram, F.; Bassani, F. Empirical Spds* Tight-Binding Calculation for Cubic Semiconductors: General Method and Material Parameters. *Physical Review B* **1998**, *57*, 6493–6507.

105. Birner, S. Nextnano3: Next Generation 3D Nanodevice Simulator. http://www.nextnano .de /nextnano3 (accessed 20 June 2012).

106. Andlauer, T. *Optoelectronic and Spin-Related Properties of Semiconductor Nanostructures In Magnetic Fields*; Technischen Universität München: München, 2009.

107. de Franceschi, S.; Jancu, J.-M.; Beltram, F. Boundary Conditions in Multiband k·p Models: A Tight-Binding Test. *Physical Review B* **1999**, *59*, 9691–9694.

108. Sengupta, P.; Lee, S.; Steiger, S.; Ryu, H.; Klimeck, G. Multiscale Modeling of a Quantum Dot Heterostructure. *Material Research Society Symposium Proceedings*, 2011; Vol. 1370, pp 47–52.

109. Keating, P. N. Effect of Invariance Requirements on the Elastic Strain Energy of Crystals with Application to the Diamond Structure. *Physical Review* **1966**, *145*, 637–645.

110. Bernard, J. E.; Zunger, A. Strain Energy and Stability of Si-Ge Compounds, Alloys, and Superlattices. *Physical Review B* **1991**, *44*, 1663–1681.

111. Tatebayashi, J.; Nishioka, M.; Arakawa, Y. Over 1.5 μm Light Emission From InAs Quantum Dots Embedded In InGaAs Strain-Reducing Layer Grown By Metalorganic Chemical Vapor Deposition. *Applied Phys. Lett.* **2001**, *78*, 3469–3471.

INTENTIONALLY LEFT BLANK.

Appendix A. Finding Reciprocal Lattice Vectors Within a Given Energy Cutoff

If $\mathbf{k} = 0$, equation 27 from the body of this report may be rewritten as

$$|\mathbf{G}| \leq G_{\text{cut}} = \sqrt{\frac{2m_e E_{\text{cut}}}{\hbar^2}}\,, \tag{A-1}$$

where G_{cut} is the magnitude of the largest reciprocal lattice vector satisfying the cutoff criterion. Since $\mathbf{G} = \sum_{i=1}^{3} n_i\, \mathbf{b}_i$, the largest possible value of n_i is[1]

$$n_i^{\max} = \frac{G_{\text{cut}}}{|\mathbf{b}_i|}\,. \tag{A-2}$$

A method for finding the reciprocal lattice vectors satisfying the energy cutoff criterion for $\mathbf{k} = 0$ is shown in figure A-1. Once these vectors have been found, another method, shown in figure A-2, may be used to find reciprocal lattice vectors satisfying the energy cutoff criterion for nonzero $\mathbf{k}$.[2]

[1] Kohanoff, J. *Electronic Structure Calculations for Solids and Molecules*; Cambridge University Press: New York, 2006.

[2] Varga, K.; Driscoll, J. A. *Computational Nanoscience*; Cambridge University Press: New York, 2011.

E_{cut} is given.
Primitive reciprocal lattice vectors $\mathbf{b}_i$ are given.

$$G_{\text{cut}} = \sqrt{2m_e E_{\text{cut}}/\hbar^2}$$

for $i = 1$ to 3 **do**
  $n_i^{\max} = G_{\text{cut}}/|\mathbf{b}_i|$
end for

Initialize list *GvecList*.

for $n_1 = -n_1^{\max}$ to $n_1^{\max}$ **do**
  for $n_2 = -n_2^{\max}$ to $n_2^{\max}$ **do**
    for $n_3 = -n_3^{\max}$ to $n_3^{\max}$ **do**

      $\mathbf{G} = n_1\mathbf{b}_1 + n_2\mathbf{b}_2 + n_3\mathbf{b}_3$
      if $|\mathbf{G}| \leq G_{\text{cut}}$ **then**
        Add $\mathbf{G}$ to *GvecList*.
      end if

    end for
  end for
end for

Figure A-1. Pseudocode for finding reciprocal lattice vectors $\mathbf{G}$ that satisfy the energy cutoff criterion when $\mathbf{k} = 0$.

E_{cut} is given.
Wavevectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ are given.
Also given is the list *GvecList* of reciprocal lattice vectors whose magnitudes are less than $G_{\text{cut}} = \sqrt{2m_e E_{\text{cut}}/\hbar^2}$.

for $\mathbf{k}$ in $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{N_\mathbf{k}}$ **do**

  Initialize list *NewGvecList*($\mathbf{k}$).
  for $\mathbf{G}$ in *GvecList* **do**

    if $|\mathbf{k}+\mathbf{G}| \leq \sqrt{2m_e E_{\text{cut}}/\hbar^2}$ **then**
      Add $\mathbf{G}$ to *NewGvecList*($\mathbf{k}$).
    end if

  end for
end for

Figure A-2. Pseudocode for finding reciprocal lattice vectors $\mathbf{G}$ that satisfy the energy cutoff criterion for nonzero $\mathbf{k}$.

INTENTIONALLY LEFT BLANK.

Appendix B. Calculating Strain From Atomic Positions

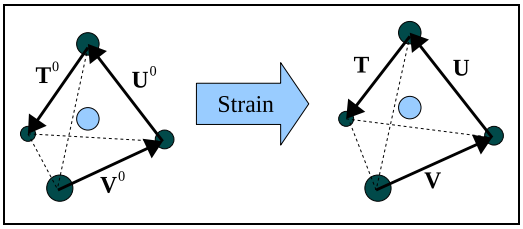

Figure B-1. Before the strain is applied, the atoms are taken to be in their ideal tetrahedral positions, with the lighter colored atom at the center of the tetrahedron. After the strain, the tetrahedron is distorted, and vectors $\mathbf{T}^0$, $\mathbf{U}^0$, and $\mathbf{V}^0$ become **T**, **U**, and **V**.

There is no unique method to determine the strain from atomic coordinates. Saito and Arakawa[1] and Steiger et al.[2] use the method of Pryor et al.[3] to determine the strain. Figure B-1 illustrates the positions of some atoms before and after strain. The atoms are tetrahedrally coordinated, with the atom shown in lighter color being at the center of the tetrahedron. After the strain is applied, this tetrahedron is distorted. The vectors $\mathbf{T}^0$, $\mathbf{U}^0$, and $\mathbf{V}^0$ are aligned to the edges of the ideal tetrahedron, while the vectors **T**, **U**, and **V** are aligned to the edges of the distorted one. The relationships between these sets of vectors and the strain ϵ_{ij} are taken to satisfy the matrix equation

$$\begin{bmatrix} T_1 & U_1 & V_1 \\ T_2 & U_2 & V_2 \\ T_3 & U_3 & V_3 \end{bmatrix} = \begin{bmatrix} 1+\epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & 1+\epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & 1+\epsilon_{33} \end{bmatrix} \begin{bmatrix} T_1^0 & U_1^0 & V_1^0 \\ T_2^0 & U_2^0 & V_2^0 \\ T_3^0 & U_3^0 & V_3^0 \end{bmatrix}, \tag{B-1}$$

which can be inverted to find the strain, so that

$$\begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{bmatrix} = \begin{bmatrix} T_1^0 & U_1^0 & V_1^0 \\ T_2^0 & U_2^0 & V_2^0 \\ T_3^0 & U_3^0 & V_3^0 \end{bmatrix} \left(\begin{bmatrix} T_1 & U_1 & V_1 \\ T_2 & U_2 & V_2 \\ T_3 & U_3 & V_3 \end{bmatrix} \right)^{-1} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{B-2}$$

Equation B-1 is equivalent to the following set of tensor equations:

[1] Saito, T.; Arakawa, Y. Electronic Structure of Piezoelectric In(0.2)Ga(0.8)N Quantum Dots in GaN Calculated Using a Tight-Binding Method. *Physica E: Low-Dimensional Systems and Nanostructures* **2002**, *15*, 169–181.

[2] Steiger, S.; Povolotskyi, M.; Park, H.-H.; Kubis T.; Klimeck, G. NEMO5: Parallel Multiscale Nanoelectronics Modeling Tool. *IEEE Transactions on Nanotechnology* **2011**, *10*, 1464–1474.

[3] Pryor, C.; Kim, J.; Wang, L. W.; Williamson, A. J.; Zunger, A. Comparison of Two Methods for Describing the Strain Profiles In Quantum Dots. *J. of Applied Physics* **1998**, *83*, 2548–2554.

$$T_i = (\delta_{ij} + \epsilon_{ij})T_j^0,$$
$$U_i = (\delta_{ij} + \epsilon_{ij})U_j^0,$$

and

$$V_i = (\delta_{ij} + \epsilon_{ij})V_j^0 \cdot \tag{B-3}$$

Here, a repeated index implies summation from 1 to 3 over that index. These equations are a more specific version of the Cauchy-Born rule,[4] where $\delta_{ij} + \epsilon_{ij}$ has replaced the deformation gradient F_{ij}. They can be interpreted as a version of the Cauchy-Born rule that neglects rotation, since[5]

$$F_{ij} = \delta_{ij} + \frac{\partial u_i}{\partial r_j^0} = \delta_{ij} + (\epsilon_{ij} + \omega_{ij}) \approx \delta_{ij} + \epsilon_{ij}\,, \tag{B-4}$$

where u_i is the displacement, r_j^0 is the position vector expressed in the coordinates of the undeformed system, and ϵ_{ij} and ω_{ij} are

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial r_j^0} + \frac{\partial u_j}{\partial r_i^0}\right),$$

and

$$\omega_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial r_j^0} - \frac{\partial u_j}{\partial r_i^0}\right). \tag{B-5}$$

Since ω_{ij} is neglected, it is implicitly assumed that F_{ij} is symmetric. If it is not, the values of ϵ_{ij} determined from method of Pryor et al.[3] will not be either.

Another method to determine strain from changes in atomic coordinates treats deviations from the Cauchy-Born rule as residuals to be minimized. Horstemeyer and Baskes[6] take the minimizing function at atom i to be

$$\phi_i(\mathbf{F}) = \sum_{j=1}^{N} \left|\mathbf{R}_{ij} - \mathbf{F}\mathbf{R}_{ij}^0\right| W(j)\,, \tag{B-6}$$

where $W(j)$ is a weighting function, $\mathbf{F}$ is the deformation gradient, and $\mathbf{R}_{ij}^0$ and $\mathbf{R}_{ij}$ are the vectors connecting atom i to each of its N neighbors in the undeformed and deformed configurations, as shown in figure B-2. If the Cauchy-Born rule holds exactly at atom i, then $\mathbf{R}_{ij} = \mathbf{F}\mathbf{R}_{ij}^0$ and $\phi_i(\mathbf{F}) = 0$. If $W(j) = 1$, the value of $\mathbf{F}$ that minimizes ϕ_i is[7,8]

---

[4]Steinmann, P.; Elizondo, A.; Sunyk, R. Studies of Validity of the Cauchy-Born Rule By Direct Comparison of Continuum and Atomistic Modelling. *Modelling and Simuation in Materials Science and Engineering* **2007**, *15*, S271–S281.

[5]Mase, G. T.; Mase, G. E. *Continuum Mechanics for Engineers*; 2nd ed., CRC Press: Boca Raton, FL, 1999.

[6]Horstemeyer, M. F.; Baskes, M. I. Strain Tensors at the Atomic Scale; In *Multiscale Phenomena in Materials—Experiments and Modeling*; Materials Research Society: Boston, 2000.

[7]Zimmermann, J. A. Continuum and Atomistic Modeling of Dislocation Nucleation at Crystal Surface Ledges. Ph.D. Thesis, Stanford University, Stanford, CA, 2000.

[8]Zimmermann, J. A.; Bammann, D. J.; Gao, H. Deformation Gradients for Continuum Mechanical Analysis of Atomistic Simulations. *International J. of Solids and Structures* **2009**, *46*, 238–253.

$$\mathbf{F} = \mathbf{W}_1 \mathbf{W}_2^{-1} , \tag{B-7}$$

where

$$\mathbf{W}_1 = \sum_{j=1}^{N} \left(\mathbf{R}_{ij} \otimes \mathbf{R}_{ij}^0 \right),$$

and

$$\mathbf{W}_2 = \sum_{j=1}^{N} \left(\mathbf{R}_{ij}^0 \otimes \mathbf{R}_{ij}^0 \right) . \tag{B-8}$$

Operator $\otimes$ is the dyad product.[5] From $\mathbf{F}$, several different strain tensors can be obtained.

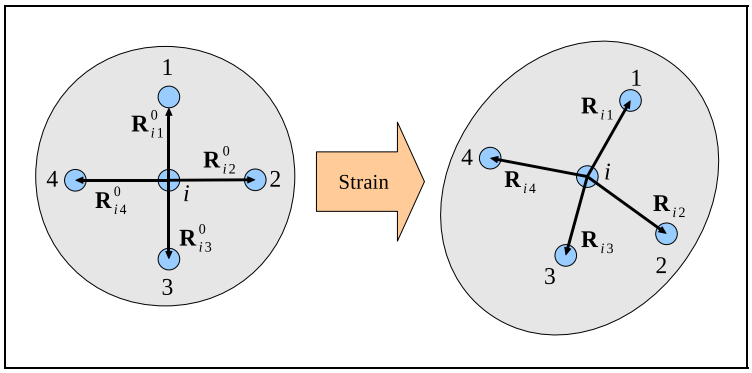

Figure B-2. Relative positions of the neighbors of atom i before and after the strain is applied.

NO. OF COPIES	ORGANIZATION
1 (PDF only)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA 8725 JOHN J KINGMAN RD STE 0944 FORT BELVOIR VA 22060-6218
1	DIRECTOR US ARMY RESEARCH LAB IMAL HRA 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR US ARMY RESEARCH LAB RDRL CIO LL 2800 POWDER MILL RD ADELPHI MD 20783-1197