# 1.01 Electrons in Semiconductors: Empirical and *ab initio* Pseudopotential Theories

**J R Chelikowsky**, University of Texas, Austin, TX, USA

© 2011 Elsevier B.V. All rights reserved.



## 1.01.1 Introduction

### 1.01.1.1 Early History

A long-sought goal of materials physics is to predict the properties of materials solely from a knowledge of the atomic constituents. Before the turn of the twentieth century, such a goal was not in the realm of possibility. Quantum mechanics had not been invented, much less applied to materials. Moreover, as the rudimentary quantum theory evolved, for example, the Bohr theory of the hydrogen atom, the arguments of the time were often focused on the structure of an individual atom and did not dream of addressing problems of materials. However, as quantum mechanics became accepted for atoms, interest in applying quantum theory to materials grew.

Indeed, once the predictive power of quantum mechanics for atoms and simple molecules became clear, the notion of predicting materials properties by solving the quantum mechanical behavior of the system of interest did not seem so outlandish. One of the first visionaries was Dirac (1929). In a very famous quote circa 1930 he stated:

> The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

The implication of this statement was clear. A solution of the quantum mechanical problem would yield the 'whole of chemistry.' One could predict properties of the chemical bond and the corresponding materials properties of matter without resort to laboratory work. Chemistry and materials properties could be predicted by calculations. Of course,

'knowing' and 'doing' are two different issues. The equations in question were much too complicated to be solved when Dirac made this statement. In fact, they are much too complicated to be solved today, even with state-of-the-art computational platforms.

As Dirac anticipated, the key to progress in this area is to develop 'approximate practical methods'. Today, we have a number of such practical methods that allow us to access some properties in a manner Dirac suggested. The first successes appeared some 30 or 40 years after Dirac's statement and were led by the development of the pseudopotential concept. The pseudopotential model of a solid provided a practical model to address the quantum mechanical problem and advances in computers provided the means to overcome the numerical hurdles.

A key attribute of a pseudopotential is that it allows us to address directly the electronically or chemically active valence states of an atom and remove from consideration the electronically inert core states. This decomposition between active and inactive states is readily done for most atoms, for example, in a silicon atom the $1s^2 2s^2 2p^6$ core states are tightly bound compared to the $3s^2 3p^2$ valence states. In **Figure 1**, we illustrate the pseudopotential model for a crystalline solid. In this case, the core

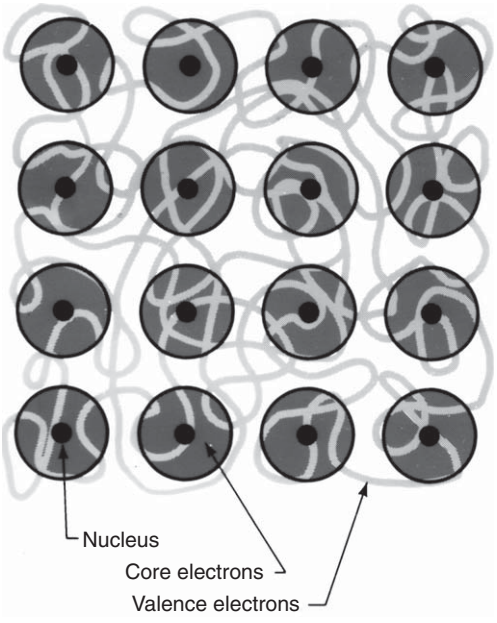


**Figure 1** Pseudopotential concept of a solid. The ion cores consisting of the core electrons and valence electrons are inert. The chemically active valence electrons move within this array of ion cores.

electrons and nucleus of an atom are treated as an inert ion core; the valence electrons are treated as an electron sea moving against the periodic background of the ion cores. This pseudopotential picture sets a length and energy scale to the problem that is determined by only the valence states. A profound consequence of this length and energy scale is that all atoms of the periodic table can be treated on an equal footing. Consider the situation without the pseudopotential approximation; in this case, an atom of hydrogen and an atom of cesium would require all the electronic states to be placed on an equal footing. The energy and length scales between the core and valance states would be different by orders of magnitude and would enormously complicate the problem.

The pseudopotential idea is not new. Several key elements were recognized after the advent of quantum mechanics. Both Enrico Fermi and Hans Hellmann contributed seminal papers in the mid-1930s (Fermi, 1934; Hellman, 1935). Interestingly, they recognized somewhat different, but important elements of pseudopotential theory. Fermi's focus was on determining the phase shift in wave functions of high-lying alkali atoms subject to perturbations from foreign atoms. He correctly recognized that if one were interested in the long-range behavior of the wave function, it was not necessary to get the details correct near the nucleus. It should be possible to replace the true potential near the nucleus with a simple potential, or pseudopotential, that yields a similar value for the long-range part. Fermi also established the importance of a scattering length within this paper.

Hellmann's work is probably closer to our contemporary picture (Schwarz *et al.*, 1999a, 1999b). Hellmann recognized that the valence orbitals for many-electron atoms contained pronounced oscillations in the region where the atomic core electrons remain. Determining an accurate description of the oscillatory behavior near the core and the long-range behavior in the chemically relevant region would render the problem computationally complex. Hellmann also recognized these core electrons did not play an important role in determining the chemical bond; after all this was clear from the periodic table. Elements such as Si and Ge have different ion-core configurations, but the valence electron configuration is the same, and their chemical properties are similar.

In today's language, we would argue that the valence electrons experience an additional kinetic

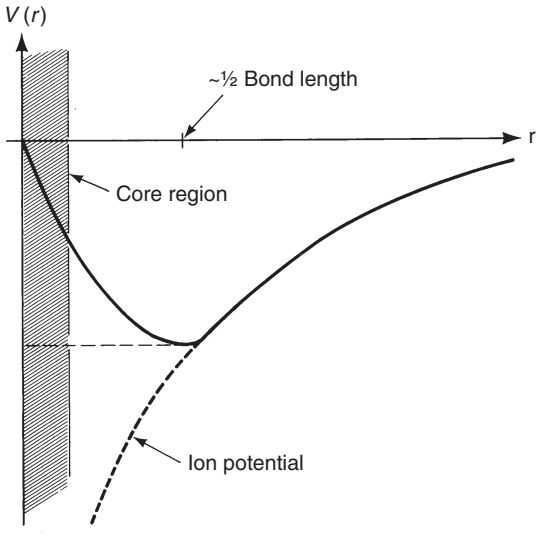


**Figure 2** Schematic of an ion-core pseudopotential. The all-electron potential and the pseudopotential ion-core potential are similar outside of the core region.

energy contribution in the core region because of the orthogonality requirement, that is, the valence state of an atom is orthogonal to a core state by nodal structure in the core region. In this 1935 paper, Hellmann proposed that this kinetic energy term could be quantified using the Thomas–Fermi expression for the kinetic energy of a free electron gas, which depends only on the electron density. When this term was added to the attractive Coulomb part of the valence electron interacting with the ion core, he argued that the resulting potential was 'weak and constant'. An example of a pseudopotential similar to those Hellmann suggested is schematically illustrated in **Figure 2**.

### 1.01.1.2 The Phillips–Kleinman Cancellation Theorem

While the work of Hellmann and Fermi established the elementary ideas associated with pseudopotentials, a rigorous transformation of an all-electron potential to a pseudopotential was lacking. One of the earliest such transformation is based on the ideas of Phillips and Kleinman (Kleinman and Phillips, 1960a, 1960b; Phillips and Kleinman, 1959; Phillips and Kleinman, 1962). Although their conclusions are more far reaching, we will focus on solving the electronic structure problem for an isolated atom. We will make the one-electron approximation, that is, the Schrödinger equation for the atomic orbitals can be written as

$$H\Psi_n(\mathbf{r}) = \left[\frac{-\hbar^2\nabla^2}{2m} + V(\mathbf{r})\right]\Psi_n(\mathbf{r}) = E_n\Psi_n(\mathbf{r}) \qquad (1)$$

where $m$ is the electron mass, $\hbar$ is Planck's constant divided by $2\pi$, $V$ is all-electron potential, and ($E_n$, $\Psi_n$) are the eigen pairs: the energy levels and corresponding wave function. Here, we assume both core and valence states exist, that is, at this point we do not consider an element such as H that has no core states. We express the wave function for the valence state, $\Psi_v$, as

$$\Psi_v(\mathbf{r}) = \Phi_v^p(\mathbf{r}) + \sum_c a_{v,c}\Psi_c(\mathbf{r}) \qquad (2)$$

The sum is over the core states, $\Psi_c$, and $\Phi_v^p$ represents the pseudopotential wave function. This form of the wave function recognizes that near the nucleus the valence wave function should appear atomic like and contain elements similar to the core states. Away from the nucleus, the core states will have little amplitude and we expect: $\Psi_v(\mathbf{r}) \approx \Phi_v^p(\mathbf{r})$: Moreover, we expect $\Phi_v^p(\mathbf{r})$ to be smoothly varying over all space. Outside of the core, the potential will vary slowly. Within the core, we expect its contributions will be small compared to the core-like functions. To determine the admixture of the core with the pseudo-wave-function, we demand the following condition:

$$\int \Psi_c^*(\mathbf{r})\Psi_v(\mathbf{r})\mathrm{d}^3r = \langle\Psi_c \mid \Psi_v\rangle = 0 \qquad (3)$$

where we use the 'bra–ket' notation. The valence state must be orthogonal to the core state, owing to the nature of the eigenvalue problem. The orthogonality requirement yields

$$a_{v,c} = -\langle\Psi_c \mid \Phi_v^p\rangle \qquad (4)$$

This form of the wave function was first proposed by Conyers Herring (Herring, 1940) in a different context. He proposed describing electronic states in crystals using orthogonalized plane waves. In this model, the valence states in a solid were replicated using plane waves that were orthgonalized to the core levels.

If we apply the Hamiltonian in Equation (1) to the valence wave function, we obtain

$$H\Phi_v^p + \sum_c \langle\Psi_c \mid \Phi_v^p\rangle(E_v - E_c)\Psi_c = E_v\Phi_v^p \qquad (5)$$

We have recognized that $H\Psi_c = E_c\Psi_c$. We can define a pseudopotential Hamiltonian, $H^p$, as

$$H^p = H + \mathcal{V}_R \tag{6}$$

where we define a potential operator, $\mathcal{V}_R$, as

$$\mathcal{V}_R = \sum_c \langle \Psi_c \mid \ldots \rangle (E_v - E_c)\Psi_c \tag{7}$$

The empty 'ket' given by $|\ldots>$ requires an operation wherein the wave function is projected on $\Psi_c$ and the functional dependence now corresponds to $\Psi_c$. If we define a pseudopotential as $V_p = V + \mathcal{V}_R$, we have now transformed the one-electron Schrödinger equation to

$$H_p\Phi_v^p = \left[\frac{-\hbar^2\nabla^2}{2m} + V_p\right]\Phi_v^p = E_v\Phi_v^p \tag{8}$$

This description of the atom has a number of advantages. The repulsive part of the potential, $\mathcal{V}_R$, largely cancels the attractive part of the all-electron potential in the core region. The pseudo-wave-function is smooth as the core oscillations have been removed and the eigen-value, $E_v$, is identical to the all-electron potential. While the wave function is now amenable to a simple basis, the pseudopotential within this construction is more complex than the all-electron potential. Although the potential is weak and only binds the valence state, this potential is energy dependent, state dependent and involves a nonlocal, non-Hermitian operator. As such, the Phillips–Kleinman potential is rarely, if ever, used for calculating pseudopotentials. However, the cancellation theorem is useful in demonstrating the essential features of a pseudopotential.

## 1.01.2 Defining Pseudopotentials

One could attempt to form a potential for an elemental crystal by writing

$$V(\mathbf{r}) = \sum_{\mathbf{R},\boldsymbol{\tau}} V_a(\mathbf{r} - \mathbf{R} - \boldsymbol{\tau}) \tag{9}$$

where $\mathbf{R}$ is a lattice vector, $\boldsymbol{\tau}$ a basis vector, and $V_a$ a potential that we associate with the atom. For a crystal, we can explicitly incorporate the periodicity of the crystal by expressing the potential in a Fourier series in three dimensions. We write

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_a(G)S(\mathbf{G})\exp(i\mathbf{G}\cdot\mathbf{r}) \tag{10}$$

where $\mathbf{G}$ are reciprocal lattice vectors. $S(\mathbf{G})$ is a structure factor given by

$$S(\mathbf{G}) = \frac{1}{N_a}\sum_{\tau}\exp(i\mathbf{G}\cdot\boldsymbol{\tau}) \tag{11}$$

where $N_a$ is the number of atoms in the basis. $V_a(G)$ is an atomic form factor given by

$$V_a(G) = \frac{1}{\Omega_a}\int V_a(r)\exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r})\mathrm{d}^3r \tag{12}$$

Here $\Omega_a$ is the volume per atom and $V_a$ is the atomic potential that we take to be spherically symmetric. The structure factor contains information on the crystal structure and the form factors contain information on the electronic interactions. The reciprocal lattice vectors are designed to have the property: $\exp(\mathrm{i}\mathbf{G}\cdot\mathbf{R}) = 1$ (Kittel, 2005). This insures that the potential is periodic:

$$\begin{aligned} V(\mathbf{r}+\mathbf{R}) &= \sum_{\mathbf{G}} V_a(G)S(\mathbf{G})\exp(\mathrm{i}\mathbf{G}\cdot(\mathbf{r}+\mathbf{R})) \\ &= \sum_{\mathbf{G}} V_a(G)S(\mathbf{G})\exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r}) = \mathrm{V}(\mathbf{r}) \end{aligned} \tag{13}$$

While this plane-wave expansion of the potential reflects the translational symmetry, the expansion is only helpful if the number of waves (or reciprocal lattice vectors) is manageable. This will not be the case for the all-electron potential owing to the singularity of the Coulomb potential at the nucleus. The pseudopotential avoids this issue by removing this singularity. What remains is to define a procedure to construct the potential.

### 1.01.2.1 Model Potentials

A realistic and powerful approach to find the required form factors is based on experiment. For example, suppose we consider a potential similar to what Hellmann proposed. Let us consider the following model, which only involves one parameter; the core radius, $r_c$:

$$V_c^p(r) = \begin{cases} 0 & r \le r_c \\ -Z_v e^2/r & r > r_c \end{cases} \tag{14}$$

This defines a potential where the cancellation occurs within the ion core ($r < r_c$). $Z_v$ is the valence charge of the ion core. This is an ion-core pseudopotential. It represents the interaction of the valence electron with the ion core (the nucleus plus core electrons) sans the role of any valence–valence electron interactions.

The next step in defining the atomic form factor is to employ Equation (12):

$$V_c^p(G) = \frac{-4\pi Z_v e^2}{\Omega_a G^2}\cos(Gr_c) \tag{15}$$

The long-range nature of the coulomb tail requires special handling in executing this integral (Kittel, 2005). The form factor 'rings', that is, the form factor oscillates as a cosine function owing to the discontinuity in the potential at $r_c$. What remains is to screen the ion-core pseudopotential by including a dielectric screening function that replicates the role of the valence electron–electron interactions. Often Thomas–Fermi screening is used for this purpose (Kittel, 2005):

$$V_c^p(G) = \frac{V_c^p(G)}{\epsilon_{TF}(G)} \tag{16}$$

Thomas–Fermi screening is given by

$$\epsilon_{TF}(q) = 1 + \frac{K_s^2}{q^2} \tag{17}$$

where $1/K_s$ is the Thomas–Fermi screening length: $K_s^2 = 6\pi n e^2/E_F$, where $n$ is the valence electron density ($n = Z_v/\Omega_a$) and $E_F$ is the Fermi energy. The atomic form factor is now given by

$$V_a^p(G) = \frac{-4\pi e^2 n}{G^2 + K_s^2}\cos(Gr_c) \tag{18}$$

A common practice is to set the potential to zero outside of the second node to remove the oscillatory or ringing behavior of the $\cos(Gr_c)$ term. An interesting consequence of Thomas–Fermi screening is that the potential now has a well-defined value when $G = 0$:

$$V_a^p(0) = \frac{-4\pi e^2 n}{K_s^2} = -\frac{2}{3}E_F \tag{19}$$

The $G = 0$ term corresponds to setting a reference energy for the solid as it corresponds to the average of the potential. This limiting case from Thomas–Fermi screening is not a very good approximation for a semiconductor such as silicon and is more appropriate for a metal. We illustrate a schematic pseudopotential in **Figure 3**.

The Phillips–Kleinman cancellation theorem yields a state-dependent pseudopotential; the potential depends on the nature of the core states. For example, in carbon there is no 1p core state and the valence 2p is not required to be orthogonal to the core. This is not true for silicon, in which both s and p states exist in the core; both the s and p valence states see a repulsive term. Historically, this difference has been used to explain why silicon and carbon chemistry are different. The previous model potential can be modified to reflect the state dependence:

$$V_{c,l}^p(r) = \begin{cases} A_l & r \le r_{c,l} \\ -Z_v e^2/r & r > r_{c,l} \end{cases} \tag{20}$$

This potential is not a simple function of position, but rather acts on the $l$-component present in the wave function (Cohen and Chelikowsky, 1982, 1989; Chelikowsky and Cohen, 1992). This can be accomplished by using a projection operator as will be discussed later. Setting this technical issue aside,

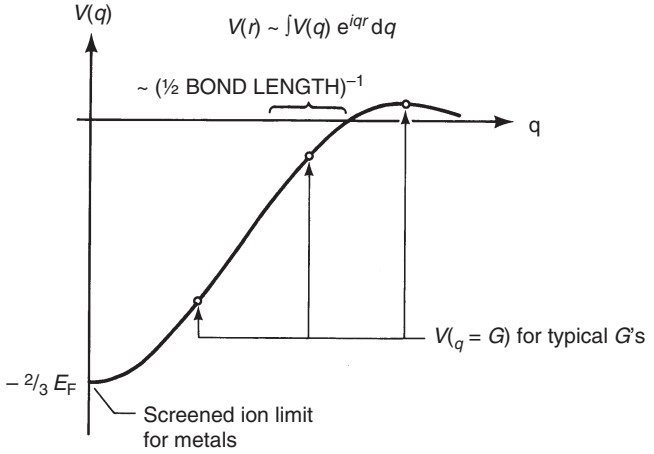


**Figure 3** Model of an atomic pseudopotential. The required form factors can be extracted from this model. Note the limiting case of $G = 0$ for a metallic system.

there are now additional parameters to be determined as the well depth and size must be established for each $l$-component.

A common procedure in the early days of pseudopotential involved fitting the well depths and sizes ($A_l$, $r_{c,l}$) to replicate optical data of ionized atoms. For example, to construct an ion-core pseudopotential for an Al atom, one needs to consider the optical excitations for an $Al^{+2}$ ion. In some cases, this is relatively easy, for example, treating the Na atom does not involve an ionized atom. However, an O atom would require examining a $C^{5+}$ ion. Experimentally, it is a difficult task to multiply ionize such an atom and measure the atomic levels. Often the potentials were estimated by extrapolation from lighter atoms to heavier ones. Another issue concerns improved dielectric functions for screening the potentials. Both issues were the subject of a number of studies, but without a definitive resolution. By the mid-1960s, an inventory of such model potentials existed (Animalu and Heine, 1965; Cohen and Heine, 1970) with some success, especially for metals.

## 1.01.3 Solving the Electronic Structure Problem

If we are given the crystalline potential, we can solve, the corresponding eigenvalue problem (Equation (1)) using a variety of approaches. Since the pseudopotential is weak compared to the all-electron potential, we can write the wave function in terms of a plane wave basis:

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} \alpha_n(\mathbf{k},\mathbf{G})\exp(\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot r) \qquad (21)$$

where $n$ is a band index and $\mathbf{k}$ is the wavevector. The sum is over all reciprocal lattice vectors, $\{\mathbf{G}\}$, but, in principle, the sum is truncated for $|\mathbf{k}+\mathbf{G}| > G_{\max}$. This form of the wave function is consistent with the Bloch form of the wave function for a periodic potential. The Bloch form of the wave function allows one to express the corresponding periodicity of the wave function, save a phase factor determined by the wavevector:

$$\psi_{n,\mathbf{k}}(\mathbf{r}+\mathbf{R}) = \exp(\mathrm{i}(\mathbf{k}\cdot\mathbf{R})\exp\psi_{n,\mathbf{k}}(\mathbf{r}) \qquad (22)$$

Inserting the crystalline potential (Equation (10)) and the plane wave basis (Equation (21)) in the one-electron Schrödinger equation yields the following:

$$\sum_{\mathbf{G}'}\left\{\left[\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 - E_n(\mathbf{k})\right]\delta_{\mathbf{G}\mathbf{G}'} + V_a^p(\mathbf{G}'-\mathbf{G})S(\mathbf{G}'-\mathbf{G})\right\}\alpha_n(\mathbf{k},\mathbf{G}') = 0 \qquad (23)$$

This corresponds to a set of linear equations, one for each $\mathbf{G}$ vector in the set. For a nontrivial solution, we require the following:

$$\det\left|\left[\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 - E_n(\mathbf{k})\right]\delta_{\mathbf{G}\mathbf{G}'} + V_a^p(\mathbf{G}'-\mathbf{G})S(\mathbf{G}'-\mathbf{G})\right| = 0 \qquad (24)$$

The diagonal elements of this determinant contain the kinetic energy contribution; the off-diagonal elements contain the form factors for the pseudopotential and the structure factor. In this example, we consider only elemental crystals so that the form factors can be separated out from the structure factor.

This eigenvalue problem is a standard mathematical problem, and is easily handled, save for highly complex systems. Typically, for common semiconductors a few hundred plane waves are required to obtain a converged result. This operation can often be performed on a laptop computer for crystalline semiconductors.

## 1.01.4 The Empirical Pseudopotential Method

One of the most significant scientific advances of the twentieth century was the development of the empirical pseudopotential method (EPM). The intuitive picture of Fermi and Hellmann was greatly advanced by Phillips and Kleinman; however, a practical method of predicting accurate energy band structures was not achieved until the EPM. This method, developed in the late 1960s, was based on fixing an energy band so that an accurate response function, as measured, was reproduced. To accomplish this feat, one had to develop a framework of translating an energy band solution to a response function.

### 1.01.4.1 Optical Properties of Semiconductors

The optical properties of a semiconductor can be characterized by an understanding of response functions such as the complex index of refraction, $N(\omega) = n(\omega) + \mathrm{i}k(\omega)$, or the complex dielectric

function, $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$ (Cohen and Chelikowsky, 1989). They are related to each other as follows:

$$\begin{aligned} N^2 &= \varepsilon_1 + i\varepsilon_2 \\ \varepsilon_1 &= n^2 - k^2 \\ \varepsilon_2 &= 2nk \end{aligned} \tag{25}$$

The normal incident reflectivity of a semiconductor is given by the response functions:

$$R = \left|\frac{N-1}{N+1}\right|^2 = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \tag{26}$$

Hence, a knowledge of the dielectric function, $\varepsilon(\omega)$, is sufficient to yield the complex index of refraction and the reflectivity. In fact, only the real (or imaginary) part of the dielectric function is required owing to the Kramers–Kronig relation:

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi}\mathcal{P}\int_0^\infty \frac{\omega'\varepsilon_2(\omega')}{\omega'^2 - \omega^2}\mathrm{d}\omega' \tag{27}$$

The principal part of the integral is required as is a knowledge of $\varepsilon_2$ over all frequencies, although, in practice, this is not a stringent requirement.

There are useful sum rules that can be applied to test the accuracy of response functions. One obvious example is to use the $\omega \to 0$ of the Kramers–Kronig relationship:

$$\varepsilon_1(0) = 1 + \frac{2}{\pi}\int_0^\infty \frac{\varepsilon_2(\omega)}{\omega}\mathrm{d}\omega \tag{28}$$

Another sum rule is

$$\frac{\pi}{2}\omega_2^p = \int_0^\infty \omega\varepsilon_2(\omega)\mathrm{d}\omega \tag{29}$$

where $\omega_p$ is the plasma frequency given by ${\omega_p}^2 = 4\pi ne^2/m$.

The connection between the macroscopic response function and the atomistic world can be made by a computation of the imaginary part of the dielectric function. A realistic, but somewhat simplified, expression comes from Ehrenreich and Cohen (EC59). The dielectric function involves transitions between the valence band ($v$) to the conduction band ($c$):

$$\begin{aligned} \varepsilon_2(\omega) = {} & \frac{4\pi e^2\hbar}{3m^2\omega^2}\sum_{v,c}\frac{2}{(2\pi)^3} \\ & \times \int_{BZ} \delta(\omega_{vc}(\mathbf{k}) - \omega)|M_{vc}(\mathbf{k})|^2\mathrm{d}^3k \end{aligned} \tag{30}$$

The integral is over all states in the Brillouin zone where the dipole matrix element is given by

$$|M_{vc}(\mathbf{k})|^2 = |\langle u_v(\mathbf{k})|\nabla|u_c(\mathbf{k})\rangle|^2 \tag{31}$$

The functions, $u_n(\mathbf{k})$, are the periodic part of the Bloch wave functions:

$$u_n(\mathbf{k}) = \sum_{\mathbf{G}} \alpha_n(\mathbf{k}, \mathbf{G})\exp(i\mathbf{G}\cdot\mathbf{r}) \tag{32}$$

where $n$ is the band index. The periodic part obeys $u_n(\mathbf{k}, \mathbf{r} + \mathbf{R}) = u_n(\mathbf{k}, \mathbf{r})$. Implicit in this expression is the existence of a bandgap that delineates the valence and conduction bands. This form also assumes cubic symmetry. The physical content of the dielectric function is clear. The delta function insures energy conservation and the dipole matrix element accounts for symmetry, that is, we assume dipole transitions. Another implicit factor is that the transitions from filled bands to empty bands are direct; the wavevector of the initial and final states are equal. This is appropriate for optical transitions where the wavelength of light is orders of magnitude larger than the atomic length scale. Also, the Ehrenreich–Cohen form does not include excitonic-(electron–hole) interactions, local field corrections, or self-energy effects, which can be included (Rohlfing and Louie, 2000) albeit at some considerable computational cost. Moreover, the qualitative features of the optical spectra are correctly obtained from the Ehenreich–Cohen expression.

Once the imaginary part of the dielectric function is known, it is straightforward using the Kramers–Kronig transform to find the real part of the dielectric function and the reflectivity, which can be directly compared to experiment. To implement the Ehenreich–Cohen expression, we need to know the energy bands, $E_n(\mathbf{k})$ and the corresponding wave functions $\psi_{n,\mathbf{k}}$. This can be obtained once the pseudopotential is defined.

Usually, the dipole matrix elements in Equation (31) are smoothly varying with $\mathbf{k}$ and can be assumed constant, except at high symmetry points. In this situation, it is useful to remove the matrix elements and isolate that part of $\varepsilon_2(\omega)$ that produces the major structure. The resulting function is called the 'joint density of states':

$$\mathcal{J}_{vc}(\omega) = \frac{2}{(2\pi)^3}\int_{BZ}\delta[\omega_{vc}(\mathbf{k}) - \omega]\mathrm{d}^3k \tag{33}$$

This integral can be recast as

$$\mathcal{J}_{vc}(\omega) = \frac{2}{(2\pi)^3}\int_{\omega=\omega_{vc}}\frac{\mathrm{d}s}{|\nabla_{\mathbf{k}}\omega_{vc}(\mathbf{k})|} \tag{34}$$

where $ds$ is a surface element in wavevector space defined by $\omega = \omega_{vc}(\mathbf{k})$. This is similar to the density of states, save here we replace the band energy by the transition frequency between the valence and conduction bands. The structure in this joint density of states occurs at critical points, just as in the density of states (Cohen and Chelikowsky, 1989). These critical points occur when

$$\left|\nabla_{\mathbf{k}}\omega_{vc}(\mathbf{k}_{cp})\right| = 0 \tag{35}$$

Physically, this occurs when the group velocity of the hole in the valence band equals that of the electron in the conduction band. Four topologically distinct critical points occur: $M_0$, $M_1$, $M_2$, and $M_3$, which represent a local minimum, two saddle points, and a local maximum. The structure associated with these critical points is presented in **Figure 4**. Structure in either the imaginary part of the dielectric function or the density states can be identified with these critical points (Cohen and Chelikowsky, 1989).

### 1.01.4.2 The Structure and Form Factors of Semiconductors

This section focuses on cubic semiconductors that occur in either the zinc–blende or diamond structures. These two structures are isostructural; they have the same atomic arrangement, but the constituent atoms can be different. Semiconductors such as Si, Ge, GaAs, InSb, ZnSe, and CdTe occur in these structures. **Figure 5** illustrates the diamond structure. The diamond structure has cubic symmetry and can be conceptualized by considering two interpenetrating face-centered cubic crystals.

The lattice vectors for diamond are given by

$$\mathbf{R}_{n_1,n_2,n_3} = n_1\hat{a}_1 + n_2\hat{a}_2 + n_3\hat{a}_3 \tag{36}$$

where the lattice basis vectors are given by

$$\hat{a}_1 = a(\hat{y}+\hat{z})/2; \hat{a}_2 = a(\hat{x}+\hat{z})/2; \hat{a}_3 = a(\hat{x}+\hat{y})/2$$

The diamond and zinc–blende crystals possess a primitive cell given by these vectors, with the cell volume given by

$$\Omega_c = |\mathbf{a}_1\cdot\mathbf{a}_2 \times \mathbf{a}_3| = a^3/4 \tag{37}$$

There are two atoms in the basis located at $\boldsymbol{\tau} = \pm(1,1,1)a/8$, so the atomic volume, $\Omega_a$ is given by $a^3/8$.

The reciprocal lattice vectors are given by

$$\mathbf{G}_{m_1,m_2,m_3} = m_1\hat{b_1} + m_2\hat{b_2} + m_3\hat{b_3} \tag{38}$$

where the reciprocal lattice basis vectors are given by

$$\hat{b_1} = (2\pi/a)(-\hat{x}+\hat{y}+\hat{z})/2; \hat{b_2} = (2\pi/a)(\hat{x}-\hat{y}+\hat{z})/2$$
$$\hat{b_3} = (2\pi/a)(\hat{x}+\hat{y}-\hat{z})/2$$

This set of vectors has the desired property that $\mathbf{G}\cdot\mathbf{R}$ is an integral multiple of $2\pi$. In some disciplines such as crystallography, the factor of $2\pi$ is omitted from the definition of reciprocal lattice vectors and $\mathbf{G}\cdot\mathbf{R}$ is an integer.

For diamond, the structure factor is given by

$$S(\mathbf{G}) = \cos(\mathbf{G}\cdot\boldsymbol{\tau}) \tag{39}$$

The structure factor for the zinc–blende crystal cannot be separated from the potential terms. For a zinc–blende crystal, $\mathcal{AB}$, where $\mathcal{A}$ is the cation and $\mathcal{B}$ is the anion, it is easy to show that the product of the form factors and structure factor in Equations (23) and (24) can be replaced by

$$V_{\mathcal{A}}(G)\exp(i\mathbf{G}\cdot\boldsymbol{\tau}) + V_{\mathcal{B}}(G)\exp(-i\mathbf{G}\cdot\boldsymbol{\tau})/2 \tag{40}$$

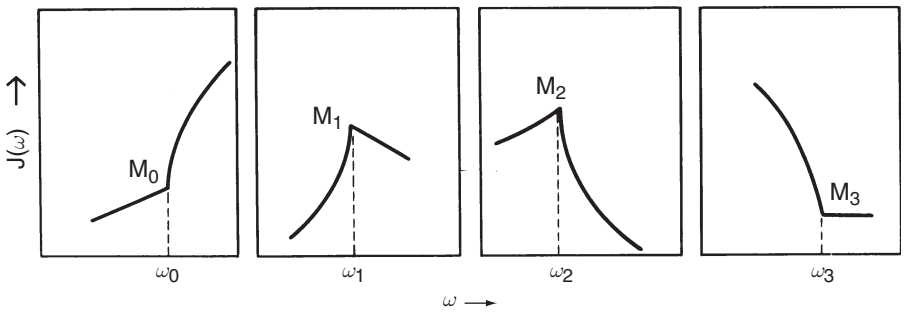


**Figure 4** Structure associated with the four types of critical points.

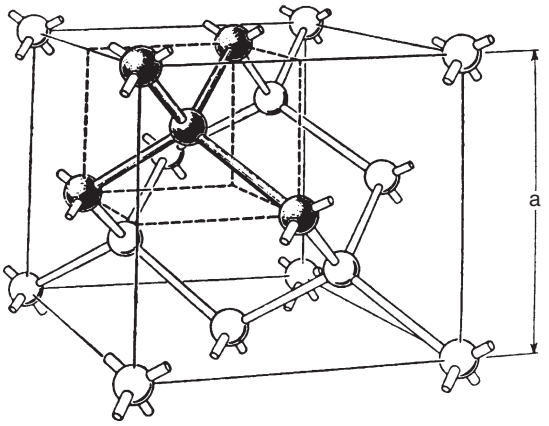


**Figure 5** Ball and stick model for the diamond crystal structure. The inset shows the tetrahedral coordination. The zinc–blende structure is isostructural, save the atomic species alternate.

It is often convenient to introduce a symmetric form factor, $V_S$, and an antisymmetric form factor, $V_A$, in which case this term can be regrouped as

$$V_S(G)\cos(\mathbf{G}\cdot\boldsymbol{\tau}) + \mathrm{i}V_A(G)\sin(\mathbf{G}\cdot\boldsymbol{\tau}) \qquad (41)$$

where

$$V_S(G) = (V_A(G) + V_B(G))/2$$
$$V_A(G) = (V_A(G) - V_B(G))/2$$

In the case of the diamond structure the antisymmetric form factor vanishes and the symmetric one corresponds to the constituent element of the diamond crystal. The form factor for the zinc–blende crystal is complex because the zinc–blende structure lacks inversion. The plane wave expansion is also complex, but this does not complicate the calculations significantly. The Hamiltonian remains hermitian.

There are two important consequences of this formalism. First, owing to the crystal symmetry, the structure factor can vanish for some reciprocal lattice vectors, meaning that the form factor for this vector is irrelevant. Second, the number of form factors required can be quite small if the pseudopotential is weak and converges quickly in reciprocal space. In the case of the diamond crystal, this can be as few as three distinct form factors. In **Table 1**, we summarize the needed form factors and structure factors. The form factors depend only on the magnitude of the reciprocal vector and are often tabulated by the value of $G^2 \times (a/2\pi)^2$, that is, the form factors are listed for $G^2 \times (a/2\pi)^2 = 0, 3, 4, 8, \ldots$ and are denoted by $V(0)$, $V(3)$, $V(4)$, $V(8), \ldots$. The structure factors are given by $\cos(\mathbf{G}\cdot\tau) = \cos(\pi(n_1m_1 + n_2m_2 + n_3m_3)/4)$; $\sin(\mathbf{G}\cdot\tau) = \sin(\pi(n_1m_1 + n_2m_2 + n_3m_3)/4)$.

**Table 1** Structure factors for diamond and zinc–blende crystals

| $G^2$ | $\cos(G\cdot\tau)$ | $\sin(G\cdot\tau)$ | *Form factor* |
|---|---|---|---|
| 0 | 1 | 0 | $V_S$ serves as a reference energy |
| 3 | $\pm1/\sqrt{2}$ | $\pm1/\sqrt{2}$ | $V_S$, $V_A$ both required |
| 4 | 0 | $\pm1$ | $V_A$ required |
| 8 | $\pm1$ | 0 | $V_S$ required |
| 11 | $\pm1/\sqrt{2}$ | $\pm1/\sqrt{2}$ | $V_S$, $V_A$ both required |

If we focus on the diamond structure only the symmetric form factors ($V(0)$, $V(3)$, $V(8)$, $V(11), \ldots$) are meaningful. $V(0)$ represents the average potential and appears only on the diagonal of the Hamiltonian matrix. It will displace the energy bands uniformly and is not important for spectroscopy. In the remainder of this section, we will omit this term and focus on spectroscopic issues. If the form factors rapidly converge, it is possible to define the potential by only three parameters corresponding to $V(G^2 \leq 11)$. Since the form factors are not linearly independent, the number of parameters are less than three. Moreover, we can get a rough estimate for these form factors from model potentials (Cohen and Heine, 1970). At a first pass, it might appear that a total of six parameters are needed for zinc–blende structures: $V_S(3)$, $V_A(3)$, $V_A(4)$, $V_S(8)$, $V_S(11)$, and $V_A(11)$. However, we can constrain the symmetric form factors in the following fashion. Suppose we are given the form factors for Ge and wish to find the form factor of GaAs. We can make the following approximation:

$$V_S(\mathrm{GaAs}) = (V(\mathrm{Ga}) + V(\mathrm{As}))/2 \approx V(\mathrm{Ge})$$

We express the symmetric form factors, the average of the Ga and As form factors to be the same as the Ge form factors. With this constraint, we need to fix only three parameters ($V_A(3)$, $V_A(4)$, and $V_A(11)$) to obtain the GaAs potential.

The EPM method was highly successful in establishing realistic form factors and energy band structures for a wide variety of materials. In **Figure 6**, we illustrate a block diagram for the EPM. Initially a set of form factors are chosen. The pseudopotential for the crystal and the Hamiltonian is constructed from this set of form factors and the structure factor. The one-electron Schrödinger is then solved and the band structure ($E_n(\mathbf{k})$) and wave functions ($\psi_{n,\mathbf{k}}(\mathbf{r})$) are obtained. From the band structure

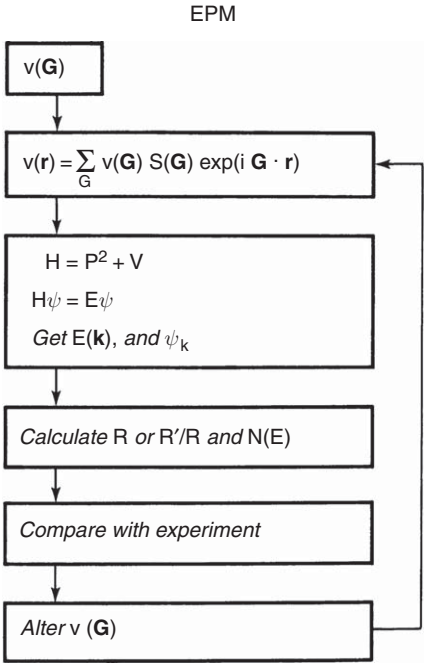


**Figure 6** Flow diagram for the empirical pseudopotential method (EPM).

the imaginary part of the dielectric function can be obtained and the corresponding reflectivity, which is then compared to experiment. (It is also possible to obtain the density of states and compare to the photoemission spectrum.) If the agreement between the calculated and measured properties is not good, the form factors can be altered and the calculations repeated. When the agreement is satisfactory, the resulting form factors and the energy band structure are deemed correct, at least within this framework.

One of the first successful implementations of the EPM was performed by Cohen and Bergstresser (1965). They examined 14 semi-conductors in the diamond and zinc–blende structures. The Cohen–Bergstresser form factors were fit to band structure features such as the fundamental gap. The resulting form factors and lattice constants are given in Table 2. The overall agreement is quite good; the bandgaps were replicated to within ~0.1–0.2 eV. Other features agreed to within ~0.5 eV over a ~10 eV range.

### 1.01.4.3 The Role of Nonlocality in Pseudopotentials

The Cohen–Bergstresser approach treats the pseudopotential as a simple function and neglects the role of the state dependence expected from the Phillips–Kleinman cancellation theorem. This is often the case for tetrahedral semiconductors, AB, where the constituent elements are not transition metals (or rare earths) or from the first row of the periodic table. However, important semiconductors such as GaN do contain elements from the first row, and the state dependence or angular momentum components need to be explicitly considered.

The nonlocal character of the ion-core pseudopotential for an atom can be expressed as

**Table 2** Pseudopotential form factors (Ry) and lattice constants for selected diamond and zinc–blende semiconductors

| | *a*(Å) | $V_S(3)$ | $V_S(8)$ | $V_S(11)$ | $V_A(3)$ | $V_A(4)$ | $V_A(11)$ |
|---|---|---|---|---|---|---|---|
| Si | 5.43 | −0.21 | 0.04 | 0.08 | | | |
| Ge | 5.66 | −0.23 | 0.01 | 0.06 | | | |
| Sn | 6.49 | −0.20 | 0.0 | 0.05 | | | |
| GaP | 5.44 | −0.22 | 0.03 | 0.07 | 0.12 | 0.07 | 0.02 |
| GaAs | 5.66 | −0.23 | 0.01 | 0.06 | 0.07 | 0.05 | 0.01 |
| AlSb | 6.13 | −0.21 | 0.02 | 0.06 | 0.06 | 0.04 | 0.02 |
| InP | 5.86 | −0.23 | 0.02 | 0.06 | 0.07 | 0.05 | 0.01 |
| GaSb | 6.12 | −0.22 | 0.00 | 0.05 | 0.06 | 0.05 | 0.01 |
| InAs | 6.04 | −0.22 | 0.00 | 0.05 | 0.08 | 0.05 | 0.03 |
| InSb | 6.48 | −0.20 | 0.00 | 0.04 | 0.06 | 0.05 | 0.01 |
| ZnS | 5.41 | −0.22 | 0.03 | 0.07 | 0.24 | 0.14 | 0.04 |
| ZnSe | 5.65 | −0.23 | 0.01 | 0.06 | 0.18 | 0.12 | 0.03 |
| ZnTe | 6.07 | −0.22 | 0.00 | 0.05 | 0.13 | 0.10 | 0.01 |
| CdTe | 6.41 | −0.20 | 0.00 | 0.04 | 0.15 | 0.09 | 0.04 |

From Cohen ML and Bergstresser TK (1965) Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures. *Physical Review* 141: 789.

$$V_c^p(\mathbf{r}) = \sum_{l=0}^{\infty} \mathcal{P}_l^\dagger V_{l,c}^p(r)\mathcal{P}_l = \mathcal{P}_s^\dagger V_{s,c}^p(r)\mathcal{P}_s + \mathcal{P}_p^\dagger V_{p,c}^p(r)\mathcal{P}_p + \mathcal{P}_d^\dagger V_{d,c}^p(r)\mathcal{P}_d + \cdots \quad (42)$$

$\mathcal{P}_l$ is an operator that projects out the $l$th component. The ion core pseudopotential is semilocal. The potential is radially symmetric, but angular dependent.

For many semiconductors only the s, p and d components of the wave functions are significant. This allows us to write

$$\mathcal{P}_s + \mathcal{P}_p + \mathcal{P}_d \approx 1 \quad (43)$$

If two of the components are similar, it is often possible to select one component and subsume the other terms into a 'local potential'. As a specific example, consider the case of GaAs. The energy band structure can be dramatically improved by considering a local potential and altering it with a potential that acts on only d waves:

$$V_c^p(\mathbf{r}) = V_{c,local}^p(r) + \mathcal{P}_d^\dagger \left[ V_d(r) - V_{c,local}^p(r) \right] \mathcal{P}_d^\dagger \quad (44)$$

A plane wave expedites the use of angular projections. It can be decomposed as follows:

$$\exp(i\mathbf{K}\cdot\mathbf{r}) = \sum_{i=0}^{\infty} (2l+1) i^l P_l(\cos(\gamma)) j_l(|K|r) \quad (45)$$

where $P_l(x)$ are Legendre polynomials: $P_0 = 1$, $P_1 = x$, $P_2 = (3x^2 - 1)/2$, $\gamma$ is the angle between $\mathbf{k}$ and $\mathbf{r}$ and $j_l(x)$ are spherical Bessel functions:

$$j_0(x) = \frac{\sin(x)}{x}$$

$$j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x}$$

$$j_2(x) = \left( \frac{3}{x^2} - 1 \right) \frac{\sin(x)}{x} - \frac{3\cos(x)}{x^2}$$

Using this decomposition allows one to write the required matrix element for a nonlocal potential as

$$\left\langle \mathbf{k}+\mathbf{G} \middle| V_{c,l}^p \middle| \mathbf{k}+\mathbf{G}' \right\rangle = \frac{4\pi}{\Omega_a} (2l+1) P_l(\cos\theta_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}'}) \times \int_0^\infty dr\, r^2 V_{c,l}^p(r) j_l(|\mathbf{k}+\mathbf{G}|r) j_l(|\mathbf{k}+\mathbf{G}'|r) \quad (46)$$

The off-diagonal matrix elements now depend on the wavevector, $\mathbf{k}$. Typically, a model potential is used for the nonlocal component. For example, one might consider a Gaussian potential: $V_{c,l}^p = A_l \exp(-(r/r_l)$. In this case the matrix elements in Equation (46) can be evaluated analytically (Cohen and Chelikowsky, 1989). Nonlocal components for tetrahedral semiconductors are available in the literature (Chelikowsky and Cohen, 1976). The agreement between experiment and nonlocal pseudopotential can be considerably better than for local pseudopotentials, albeit more parameters are utilized, typically two per element.

### 1.01.4.4 The EPM Applied to Diamond Structure Semiconductors

Group IV elements, save Pb, can form in the diamond structure (**Figure 5**). All elements of this group have a valence electron configuration of $s^2p^2$. When bonds are formed, we view this process as the promotion of an electron in the s-state to the p-state to form $sp^3$ hybrids. This explanation accounts for the nature of the diamond structure as $sp^3$ hybrid orbitals form tetrahedral bonding patterns. Moreover, the hybridization energy increases as one descends down the group IV column (Phillips, 1973). As such, we expect the diamond structure to be less stable versus metallic structures such as the face-centered cubic structure of Pb, or the gray on structure for Sn. The general increase in metallicity down the column is also rationalized by the lack of formation of $sp^3$ hybrid orbitals for the heavy elements.

#### *1.01.4.4.1 The electronic structure of silicon*

The band structure of silicon serves as the hydrogen atom of solid-state physics. Virtually any new band structure method is tested on crystalline silicon. There is general accord on the basic energy band features of silicon, but this was not the case until the end of the 1970s. In **Figure 7**, we present the band structure of silicon as calculated using two different pseudopotentials. One of the potentials contains nonlocal corrections and yields a slightly wider valence band. The energy bands are plotted along directions between high symmetry points in the Brillouin zone as indicated in **Figure 8**.

Silicon is different in a fundamental way from other diamond or zinc–blende semiconductors. The ordering of the conduction bands of silicon is different from that of germanium or tin. In Si, the $\Gamma_{15}$ band is lower than the $\Gamma_{2'}$ band. This difference was the source of some controversy in the early days of energy band calculations. For example, some workers claimed that the first pseudopotential calculations were in strong disagreement with less empirical

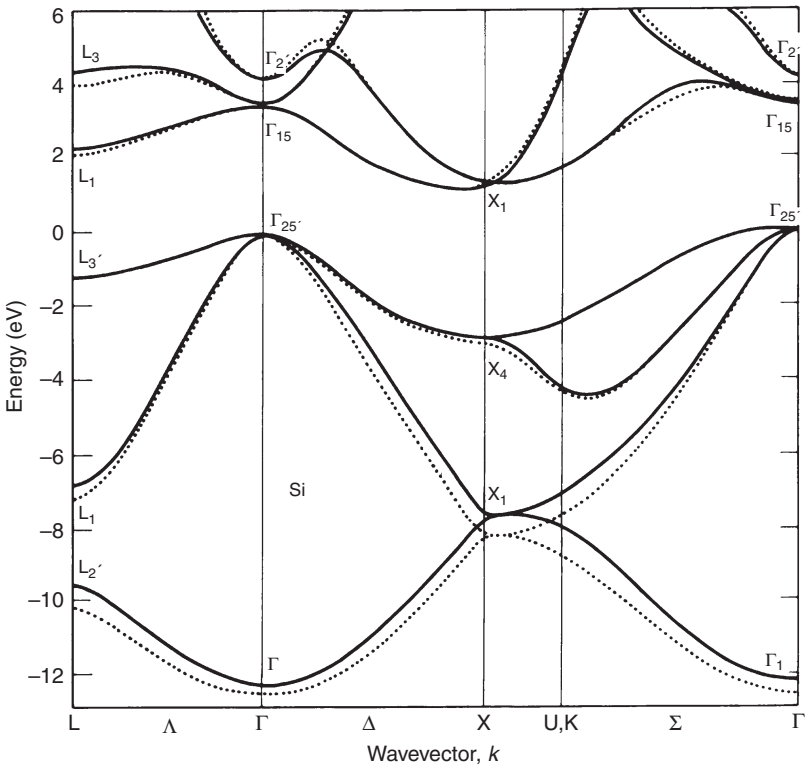


**Figure 7** Band structure for silicon.

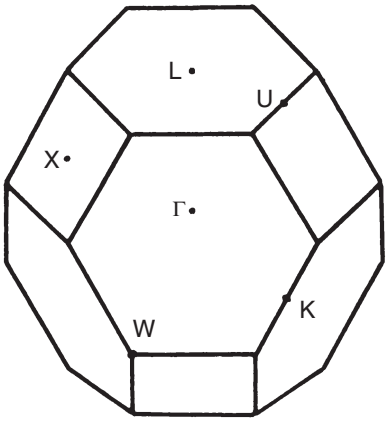


**Figure 8** Brillouin zone labeling high symmetry points.

methods for determining energy bands and with the available experimental data (Kunz, 1971). However, optical measurements of silicon–germanium alloys strongly suggested that silicon and germanium were different (Kline *et al.*, 1968). The issue was decided by the work of Aspnes and Studna (1972). They used low-field electroreflectance to resolve transitions from $\Gamma_{25}$ to $\Gamma_{2'}$, and showed unequivocally that the conduction band ordering is in complete agreement with the pseudopotential predictions.

Another feature of the silicon band structure, related to the lowest conduction band configuration, is the nearly parallel dispersion of the lowest conduction band and the uppermost valence band. Specifically, consider the dispersion along the A direction. Grover and Handler (1974) proposed on the basis of their electroreflectance data that the critical point at $L$ is essentially two dimensional, that is, the experimentally measured transverse band mass virtually vanishes. The value they obtained for $m_1$ was approximately 0.02 m. The corresponding mass calculated by pseudopotential methods is closer to 0. 1 m, but even this value results in a nearly two-dimensional $M_0$ critical point at the $L$ point (Cohen and Chelikowsky, 1989). Moreover, the dispersion for the valence and conduction bands is such that the energy difference between these bands is less than 0.01 eV over half the distance from $L$ to $\Gamma$.

This valence–conduction band configuration favors excitonic behavior. It has been traditionally suggested that the rather large discrepancies between the theoretical and experimental optical constants in this energy range arise from such an effect. To illustrate the point, we have displayed the measured and calculated imaginary part of the dielectric constant in

Figure 9. The measured dielectric function is nearly a factor of 2 larger near 3.4 eV than the calculated value for the energy region in question. This discrepancy is also observed in the real part of the dielectric function, which is displayed in Figure 10.

While the Ehrenreich–Cohen formalism (Ehrenreich and Cohen, 1959) replicates the essential features of the optical constants, it omits local fields and electron–hole interactions. Local-field corrections were addressed by Adler (1962) and Wiser (1963). Local fields allow for the microscopic variation of the field within the unit cell. Unfortunately, the corrections for local fields do little to improve the dielectric function in this region as indicated in Figure 9. The role of the electron–hole interactions in semiconductors was addressed by a number of groups. One of the first realistic attempts is from Hanke and Sham (1974). They examined the optical properties of diamond, including both local fields and exchange interactions. More recent work by Louie and collaborators (Hybertsen and Louie, 1986; Rohlfing and Louie, 1998, 2000) has unequivocally demonstrated the role of excitonic contributions to the optical response functions.

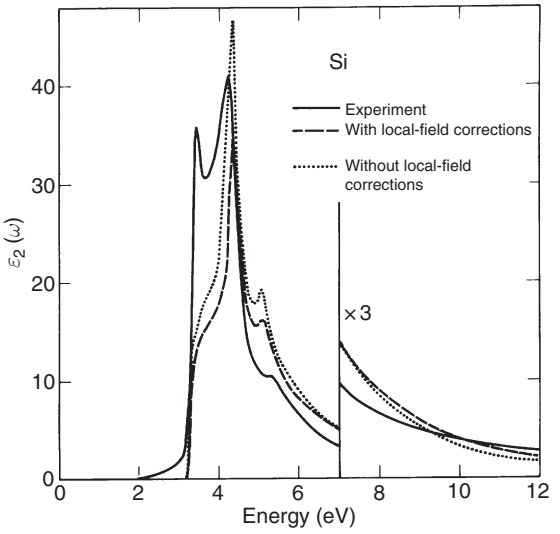


**Figure 9** Imaginary part of the dielectric function for Si with and without local fields.

In Figure 11, we exhibit the calculated and measured reflectivity spectrum for silicon. As for most solid-state spectra, the spectrum is not highly structured, and on first inspection appears to contain little information. Such an assessment would be incorrect. Details of the spectrum can be enhanced by numerically differentiating the spectrum. In Figure 12, we present the calculated and measured logarithmic derivative of the reflectivity spectrum. The derivative spectrum does not exhibit the three broad peaks of the undifferentiated spectrum but numerous sharp spectral features. These features are classified according to a standard nomenclature. The lowest energy structural feature is denoted by $E_0$ and corresponds to structure from the smallest direct transition. The next higher energy structural feature is denoted by $E_1$. This feature denotes structure associated with the $\Lambda$ direction. In the unique case of silicon, the $E_0$ and the $E_1$ structural features occur in

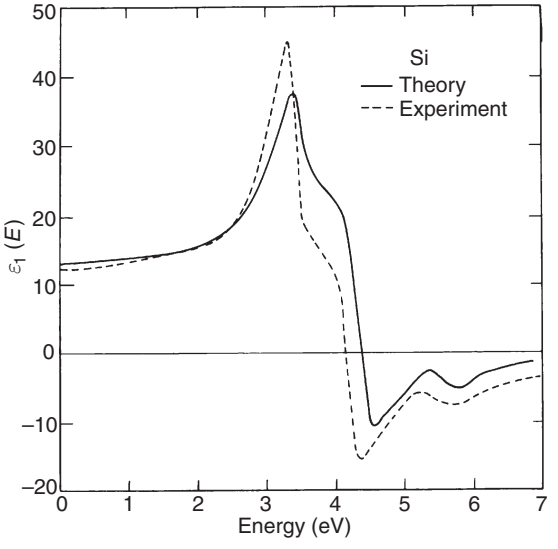


**Figure 10** Real part of the dielectric function for Si.

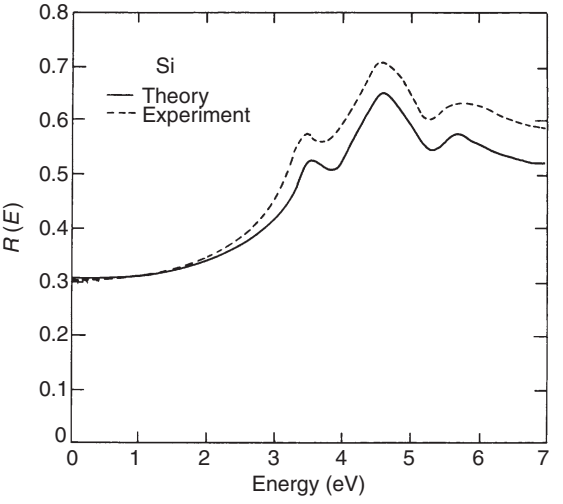


**Figure 11** Measure and calculated normal incident reflectivity for Si.

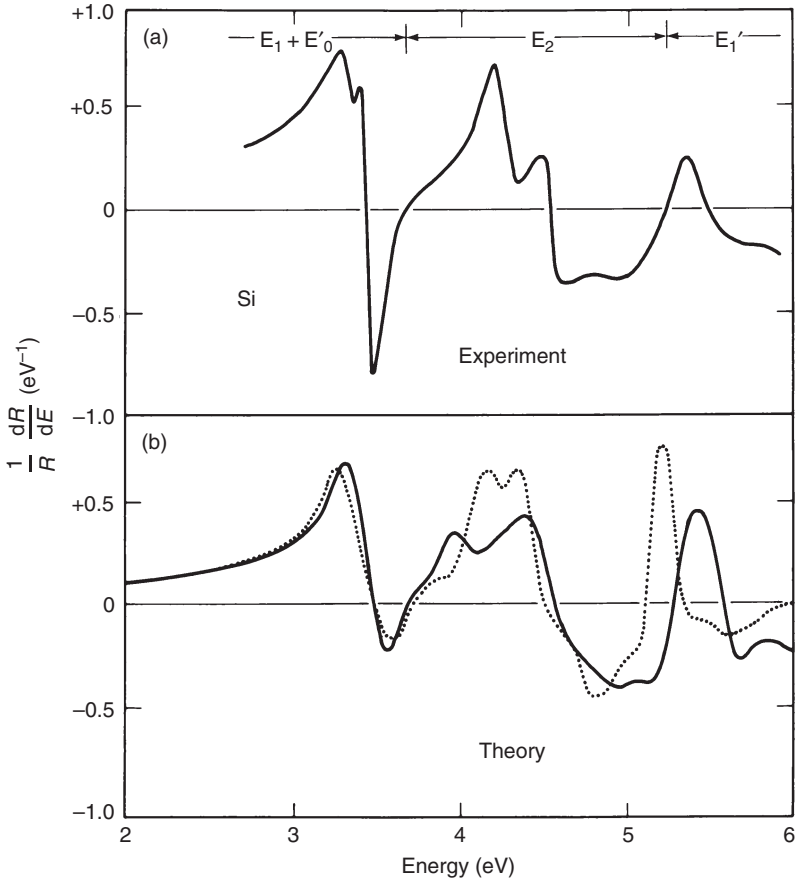


**Figure 12** Measured and calculated normal incident modulated reflectivity for Si.

the same energy region, that is, the lowest energy direct transitions at $L$ and at $\Gamma$ have nearly the same energy. In fact, a good deal of work has been peformed to determine precisely where in the Brillouin zone this transition takes place. Piezoelectric experiments (Gobeli and Kane, 1965), chemical shifts in Ge–Si alloys (Tauc and Abraham, 1961), electroreflectance (Pollak and Cardona, 1968), and some wavelength-modulation techniques (Koo *et al.*, 1971) have suggested that the first reflectivity peak arises near the zone center, perhaps along the $\Delta$ direction. However, other work (Grover and Handler, 1974) suggested that this transition occurs along the $\Lambda$ direction. Unfortunately, the pseudopotential work has not resolved this issue since the reflectivity structure in question arises from contributions near $\Gamma$ and along both the $\Lambda$ and $\Delta$ directions. The dominant transition in the calculations appears to be along the $\Lambda$ direction, but no firm conclusion can be drawn. What is clear is that the complexity of this structure does suggest a multiplicity of critical points within this energy region. This conclusion has been drawn by several authors (Welkowsky and Braunstein, 1972). The $E_2$ structure dominates the reflectivity spectrum and occurs at about 4.5 eV. This structure arises from large regions of the Brillouin zone.

While optical properties served as a centerpiece for the first EPM calculations, photoemission measurements provide information inaccessible to interband transitions. In particular, photoemission can provide the absolute position of bands relative to the vacuum. The valence band density of states for silicon calculated using pseudopotentials is compared with experiment in **Figure 13**.

Under certain conditions, it is possible to compare directly the theoretical density of states with photoemission measurements. The density of states, $D(E)$, for crystalline matter, can be defined as

$$D(E) = \frac{2\Omega}{(2\pi)^3}\sum_n \int \mathrm{d}^3k\,\delta(E - E_n(\mathbf{k})) \qquad (47)$$

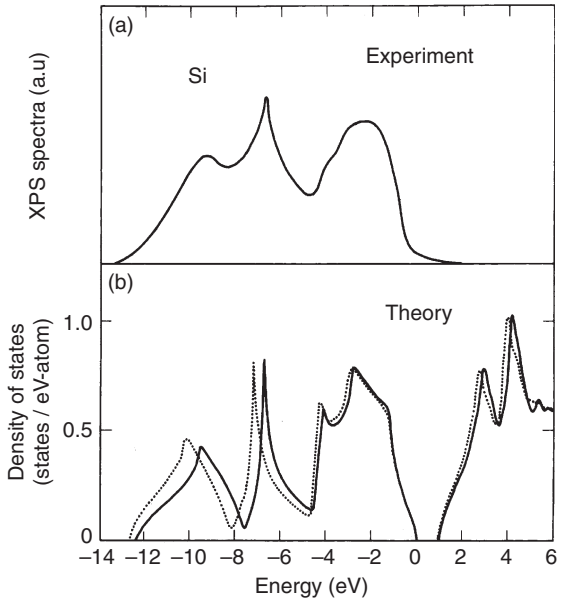


**Figure 13** Density of states for crystalline silicon. The top panel is from X-ray photoemission spectroscopy (Ley *et al.*, 1972; Cavell *et al.*, 1973).

where Ω is the cell volume. Typically, the density of states, which is an extensive property, is normalized in terms of states per atom. If matrix element effects are ignored, the density of states can be extracted from X-ray photoemission spectroscopy or ultraviolet photoemission measurements. For the testing of empirical pseudopotential band structures, the photoemission spectra have proven invaluable for the determination of accurate potentials. Photoemission measurements have indicated that the local pseudopotentials cannot provide band pictures where both the valence and conduction bands are simultaneously reproduced accurately when compared to experimental band pictures.

For diamond-structure semiconductors, the valence band density of states may be divided into three general regions as in **Figure 13**. Using the top of the valence band as our zero of energy, the region of −13 to −8 eV is predominantly of s-like character stemming from the atomic 3s-states of Si. The region from −8 to −4 eV is a transition region with contributions from both s- and p-states. The region from −5 to 0 eV is predominantly p-like. The feature which delineates these spectral regions is a sharp reduction or 'dip' in the density of states compared to the average density. The nonuniformity of the density of states in contrast to simple metals arises from hybridization among the atomic orbitals. Any disruption of the crystalline bonding environment will introduce states into the regions of reduced state density. For example, in amorphous materials, where there presumably exist bond angle and bond length deviations from the ideal crystalline state, these regions are strongly affected. The dip around −8 eV fills in, and the upper region of the density of states tends to narrow and shift to higher energies. For surface states or vacancy states where broken bonds are clearly involved, the defect localized states tend to fill in these regions, including the fundamental gap.

#### *1.01.4.4.2 The electronic structure of germanium, gallium arsenide, and zinc selenide*

The triad of semiconductors Ge, GaAs and ZnSe are special. They are isostructural with a bond length that hardly changes among the three. As such, the difference between the three can be attributed to potential differences. In pseudopotential theory, the symmetric form factors do not significantly change; the antisymmetric form factors (the difference between the cation and anion pseudopotentials) control the band structure, and the optical and photoemission spectra.

*1.01.4.4.2(i) Germanium* The valence electron configurations of germanium and silicon are both $s^2 p^2$; however, significant differences exist for the band structures. The chief differences arise from the core-charge configuration. While the nuclear charge increase is exactly balanced by an increase in core electron number, the larger number of core electrons is more effective in screening the nuclear charge. The valence levels in germanium are less tightly bound compared to silicon, and this is reflected in the more metallic character of the germanium band structure. Another significant feature that distinguishes germanium from silicon is the presence of 3d states within the core. While it is clear that these d-states do not significantly participate in the cohesive process, they can still influence the conduction band configuration. Specifically, the unoccupied bands, which may have 4d character, are influenced by an orthogonality requirement not present for the 3d states in silicon. This reasoning is reinforced by empirical evidence that suggests that the presence of *d*-electrons, that is, nonlocal pseudopotentials, which incorporate d-orbital corrections, appear crucial in describing the germanium conduction bands (Chelikowsky and Cohen, 1973; Phillips and Pandey, 1973).

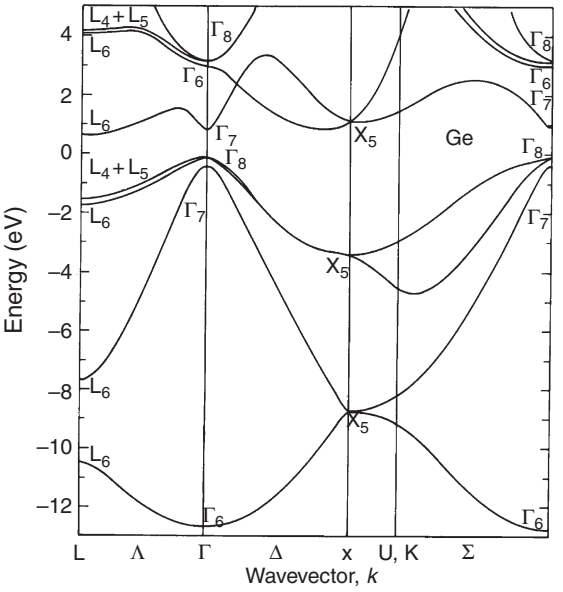


**Figure 14** Band structure for crystalline germanium. The top of the valence band is considered as the energy zero.

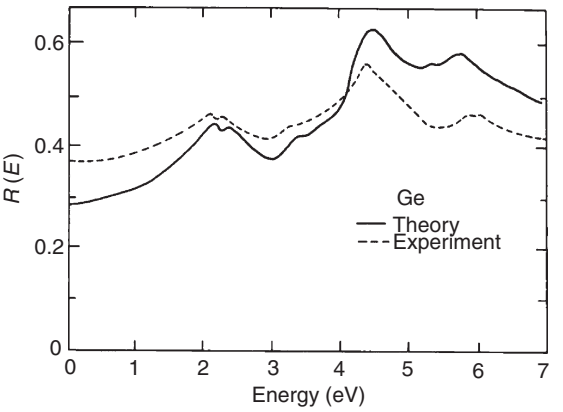


**Figure 15** Theoretical and experimental reflectivity for crystalline germanium. The experimental data is from Philipp and Ehrenreich, (1963).

In **Figure 14**, we present the band structure of germanium as determined from empirical pseudopotentials. This calculation includes spin–orbit interactions, which are most significant for heavier elements (Cohen and Chelikowsky, 1989). Unlike the case of silicon, most band structures for germanium are in reasonably good agreement. The band structure of germanium differs from silicon most significantly in the conduction band arrangement. The $\Gamma_{15}$ and $\Gamma_{2'}$ bands, using the notation in **Figure 7**, reverse their ordering. Moreover, while germanium and silicon are both indirect semiconductors, the conduction band minimum in germanium occurs at the $L$-point as opposed to a point near $X$ in silicon. As noted earlier, the valence band configuration for diamond semiconductors is insensitivie to the constituents: silicon, germanium, and tin have almost identical valence band configurations, but they have quite different conduction band configurations.

In **Figure 15**, we present the reflectivity spectrum of germanium as calculated from pseudopotential band theory and as determined by experiment (Grover and Handler, 1974). Compared to silicon, germanium exhibits a richer spectrum. The additional structure arises from the larger spin–orbit coupling occurring in germanium. For example, in the $E_1$ region near 2.2 eV, a distinct doublet structure occurs. This doublet arises from the spin–orbit splitting in the $L_{3'}$ level.

In **Figure 16**, the logarithmic derivative reflectivity spectrum is displayed. The derivative spectrum vividly illustrates the rich reflectivity structure available using modulated techniques. Perhaps the most interesting feature of the germanium reflectivity spectrum is the $E_2$ peak at 4.5 eV. Initially, the origin of this spectral feature appeared to be located at the $X$-point and to involve transitions between the highest valence band and the lowest conduction band. The energy of this transition is near the observed $E_2$ peak and the dipole matrix elements for these transitions are large. However, the phase space associated with this critical point is too small to yield such a prominent structure. The observation of a well-defined interband reduced mass for the E2 structure from electroreflectance (Aspnes, 1973) suggests a relatively well-localized critical-point origin for this spectral feature. Ruling out transitions near $X$, the critical point must reside elsewhere within the Brillouin zone. Empirical pseudopotential calculations reinforce this interpretation; these calculations (Chelikowsky and Cohen, 1973) yield a well-defined critical point near the special **k** point (Chadi and Cohen, 1973).

The density of states for germanium is illustrated in **Figure 17** and compared to photoemission and inverse photoemission. Photoemission methods such as X-ray photoelectron spectroscopy (XPS) and ultraviolet photoemission spectroscopy (UPS) measure only properties of occupied states. Inverse photoemission spectroscopy allows one to detect the properties of empty states. Electrons with a well-defined energy are directed at a sample. These electrons decay to the lowest unoccupied states,

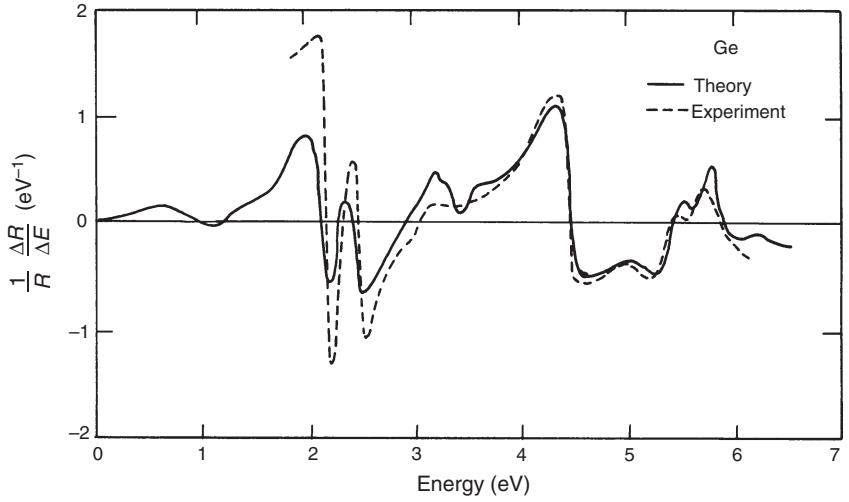


**Figure 16** Theoretical and experimental modulated reflectivity for crystalline germanium. The experimental data is from Zucca and Shen (1970).

with some of the decay transitions being radiative. Photons emitted in this process are detected and a photon counts versus incident electron energy spectrum is measured. Like photoemission, the inverse photoemission can be related to the empty state density of states, ignoring any matrix element effects. One can obtain a direct measure of the 'quasi-particle' gap, that is, the energy required to obtain a noninteracting electron–hole pair, by a combination of photoemission and inverse photoemission spectra.

***1.01.4.4.2(ii) Gallium arsenide*** We can approach the electronic structure of gallium aresenide from the perspective of perturbing the germanium potential. Intuitively, we expect

$$V_S(G) = [V_{Ga}(G) + V_{As}(G)]/2 \approx V_{Ge}(G) \qquad (48)$$

That is, the symmetric part of the potential should resemble the form factors of germanium. The antisymmetric form factor,

$$V_A(G) = [V_{Ga}(G) - V_{As}(G)]/2 \qquad (49)$$

which is a measure of the difference between the gallium and arsenic potentials, should contain the essential information which distinguishes germanium from gallium arsenide. Formally, it is the deviation of $V_A(G)$ from zero that distinguishes zinc–blende from diamond semiconductors. In this sense, we expect a smooth transition from a purely covalent semiconductor $V_A(G) = 0$, to a highly ionic semiconductor where $V_A(G)$ deviates notably from zero.

We expect that many properties of semiconductors, for example, optical gaps, cohesive energies dielectric constants, bond charges, etc., might be determined simply with an ionicity factor based on $V_A(G)$ or a similar measure. Actually, this is the case, and although $V_A(G)$ is not explicitly used, a more sophisticated measure, based on similar concepts, of ionic versus covalent contributions to crystal binding in diamond and zinc–blende semiconductors has been devised by Phillips (1973).

In **Figure 18**, we illustrate the calculated band structure for gallium arsenide using the EPM. The nonlocal pseudopotential EPM approach has yielded one of the most accurate band structures for GaAs to date. In fact, as a result of the extensive experimental studies in this compound, this band structure may be the most accurate band structure over a large energy range available for any material. By compiling the most recent experimental data, we estimate an error of about 0.1 eV over an energy span of nearly 20 eV. Thus, we have an accuracy of roughly 1% or better; this is almost an order of magnitude better than other band structures.

As we introduce antisymmetric form factors for the gallium arsenide crystal potential, several significant alterations from a germanium-like band structure may be observed. First, the bottom valence band, which now has atomic s-like character and is localized on the anion, splits off from the the rest of the valence band to form the antisymmetric gap. This gap grows with increasing ionicity of the crystal or charge transfer between the cation and anion.

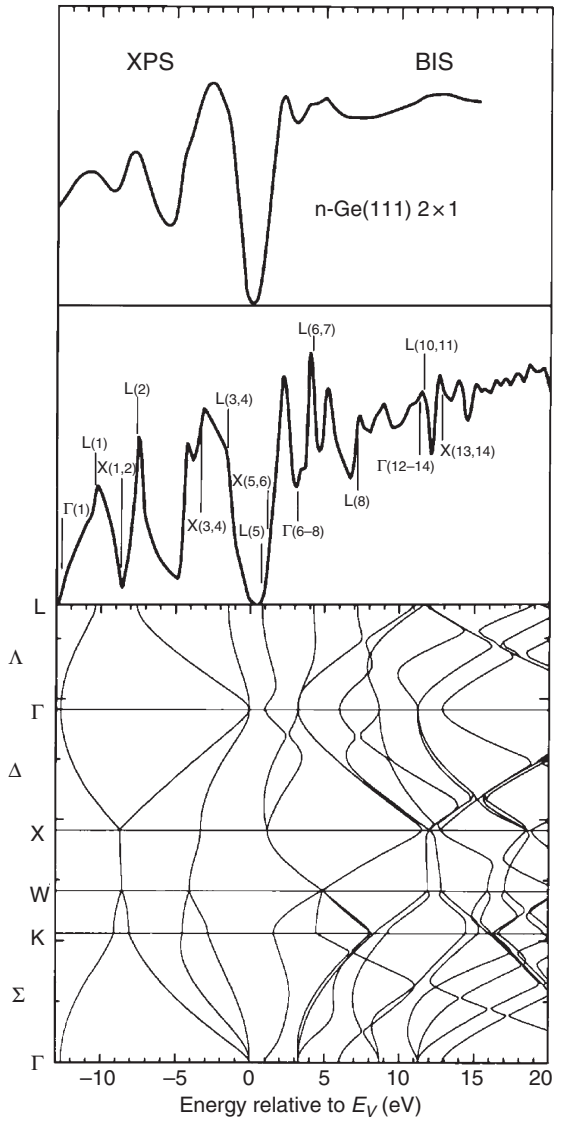


**Figure 17** Experimental X-ray photoemission spectroscopy (XPS) and Bremsstrahlung isochromat spectroscopy (BIS) or inverse photoemission spectra for crystalline germanium (Chelikowsky *et al.*, 1989). The critical point features of the theoretical density of states are also identified with band structure features.

Another smaller gap arises from the antisymmetric part of the potential and is located between the first and second conduction bands along the Δ-direction. This smaller gap is important for transport properties of zinc–blende semiconductors and has been observed to have subtle effects on the reflectivity spectrum of the zinc–blendes. We also note that the band gaps in gallium arsenide tend to be larger than in germanium.

In general, as a semiconductor becomes more ionic, the valence bandwidths narrow and the optical gap grows in size. Part of this effect arises from dehybridization accompanying the change from covalent to ionic bonding.

Perhaps the most important conclusion from our experience with GaAs is not simply that the pseudopotential approach can yield accurate results. Consider the parametrized nature of the potential. One might naively expect that the band structure could be adjusted at will with parametric nature of the form factors. This is not the case for two reasons. First, the form factors are not linearly independent, for example, the form factors $V(8)$ and $V(11)$ characterize the size of the ion core and tend to move the bands in similar fashion. In a fitting sense, we have only two or three parameters to fix all the energy bands. Second, the form factors are chosen to be close to those expected from model potentials and are not arbitrary.

In GaAs, this issue is important. The band ordering of the conduction band minima was first believed to be an issue with the EPM. In order to explain the Gunn effect (Gunn, 1964), it was thought that the correct ordering is $\Gamma_6^c - X_6^c - L_6^c$. (The Gunn effect occurs in III–V and results in negative differential resistance.) If it were possible to obtain this ordering by a simple adjustments of the form factors, the bands would have been fit to give ordering suggested by the Gunn effect. In fact, pseudopotential band calculations place the $L_6^c$ band below $X_6^c$, resulting in a $\Gamma_6^c - L_6^c - X_6^c$. Since experiment is the arbiter of such issues, the pseudopotential method was deemed to be in error.

However, in the late 1970s careful experiments showed the ordering employed to explain that the Gunn effect was not correct. In **Figure 19**, we illustrate the interpretation by Aspnes (1976). Using electroreflectance for core to conduction band transitions, he was able to show the ordering from the EPM was correct.

In **Figures 20** and **21**, we illustrate the computed optical response functions, that is, real and imaginary parts of the dielectric function, for GaAs. We also compare these functions to experiment (Philipp and Ehrenreich, 1963). The information content of the response functions is considerable. It is for this reason that Phillips (1973) based his bonding theories of semiconductors on spectroscopic data rather than following the approach of Pauling (1960), who used thermochemical data to analyze molecular bonding trends. The general shape of the real part of the dielectric function is that expected for a harmonic

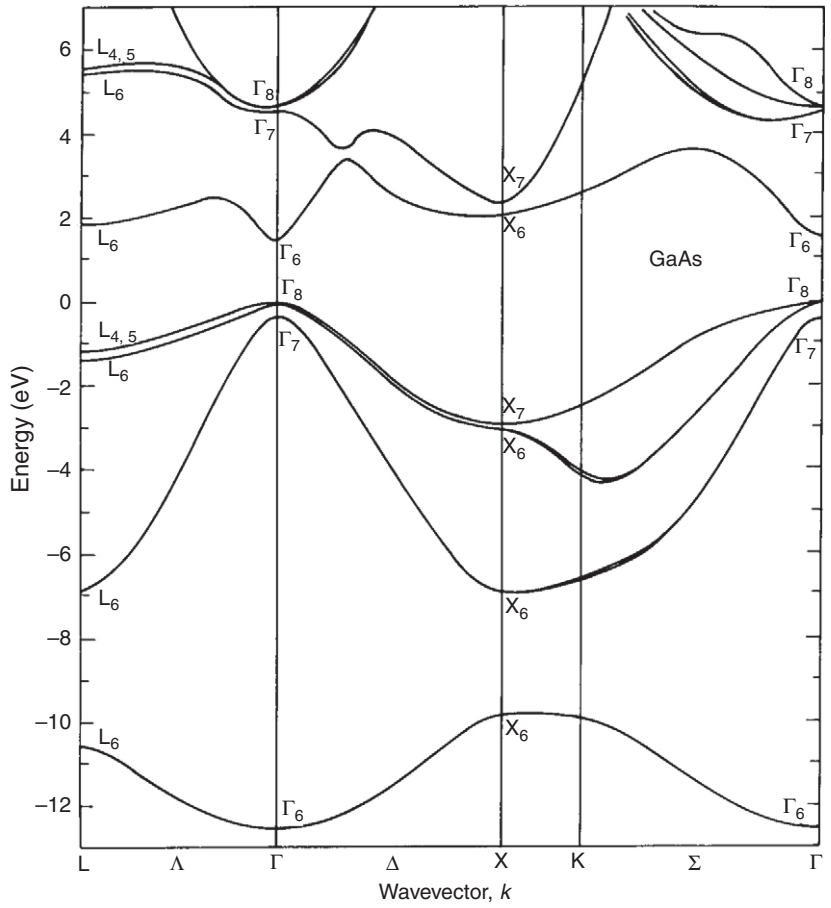


**Figure 18** Energy bands for crystalline gallium arsenide. The top of the valence band is taken to be the zero energy reference.

oscillator with a resonant frequency at about 5 eV. We can think of this resonant frequency as a fundamental property of GaAs that represents the average bonding–antibonding energy level separation. Phillips divides up this average bonding–antibonding gap into a part that is ionic and a part that is covalent. One can extract from the dielectric function relevant parameters for formulating ionicity scales (Phillips, 1973). Such scales are useful for making predictions of a variety of semiconductor properties. With a knowledge of the real and imaginary parts of the dielectric function, we can derive the reflectivity spectrum and reflectivity derivative spectrum.

While the structure of the calculated imaginary part of the dielectric function is similar, the magnitude in the ∼2–4 eV region is lower by almost a factor of 2. This discrepancy was initially attributed to a failure of the Ehrenreich–Cohen dielectric function (Ehrenreich and Cohen, 1959) to include electron–hole 'excitonic' interactions. Recent work based on a many-body formulation of the Bethe–Salpeter equation confirms this interpretation as indicted in **Figure 22**. The Bethe–Salpeter equation strongly enhances the magnitude of the imaginary part of the dielectric function, but does not shift the critical points.

The EPM-calculated reflectivity is shown in **Figure 23**. As for the silicon and germanium semiconductors, we may divide up the zinc–blende reflectivity structure into five distinct regions. For GaAs, the $E_0$ region extends from 1 to 2 eV, the $E_1$ region 2-4 eV, the $E'_0$ region 4-5 eV, the $E_2$ region 5-6 eV, and the $E'_2$ region 6-7 eV. The lowest energy region $E_0$ is dominated by structure originating from the fundamental gap at Γ. Spin–orbit interactions split the upper valence bands of gallium arsenide by about 0.3 eV; for heavier metal constituents, the splittings can be quite large, for example, 1 eV or more.

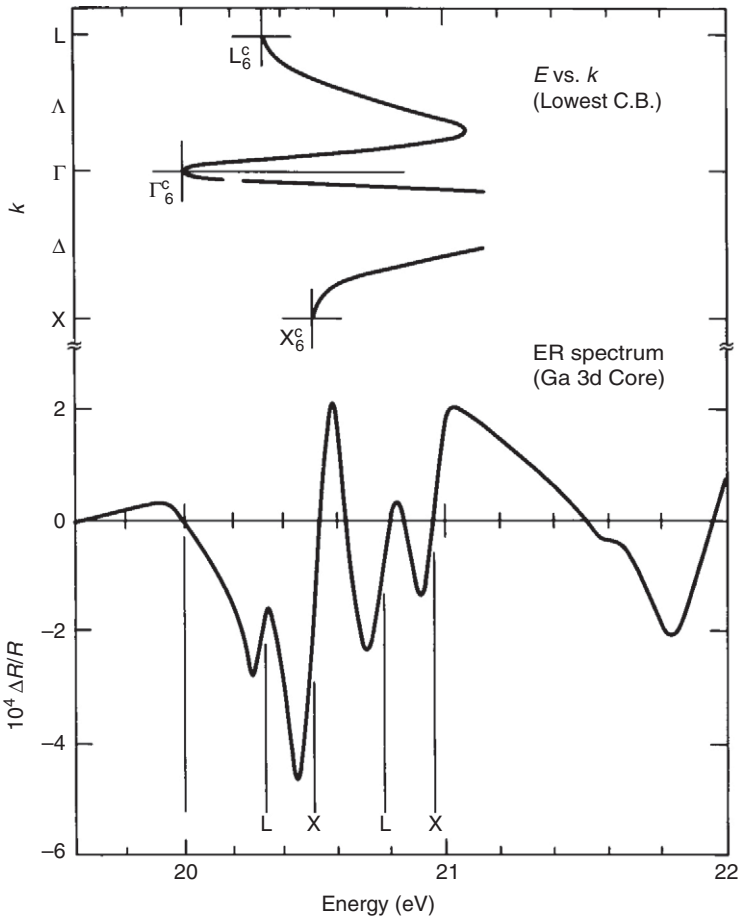


**Figure 19** Fine structure in the electroreflectance spectrum for GaAs (Aspnes, 1976). Assignment to the conduction band minima is indicated.

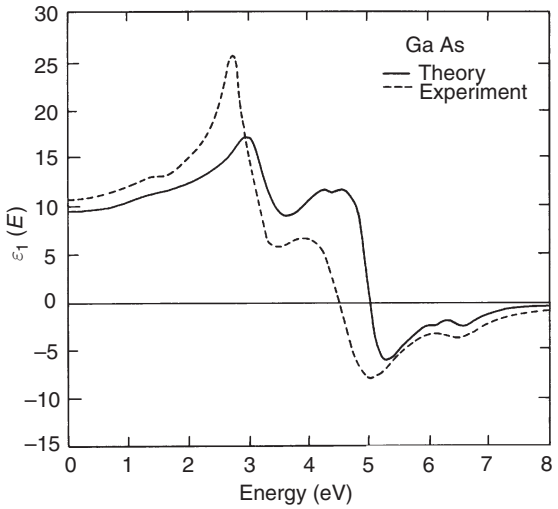


**Figure 20** Real part of the dielectric function for gallium arsenide compared to experiment (Philipp and Ehrenreich, 1963).

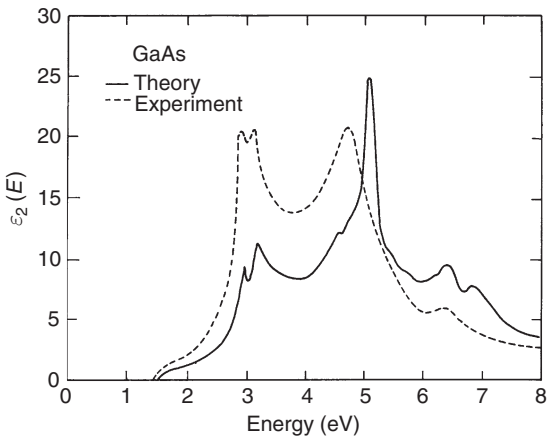


**Figure 21** Imaginary part of the dielectric function for gallium arsenide compared to experiment (Philipp and Ehrenreich, 1963).

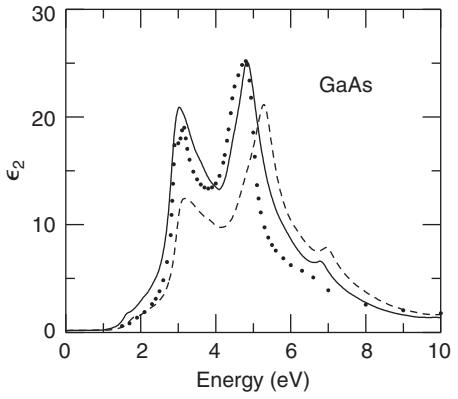


**Figure 22** Calculated optical absorption spectrum of GaAs with (solid lines) and without (dashed lines) electron–hole interaction from Rohlfing and Louie (1998).

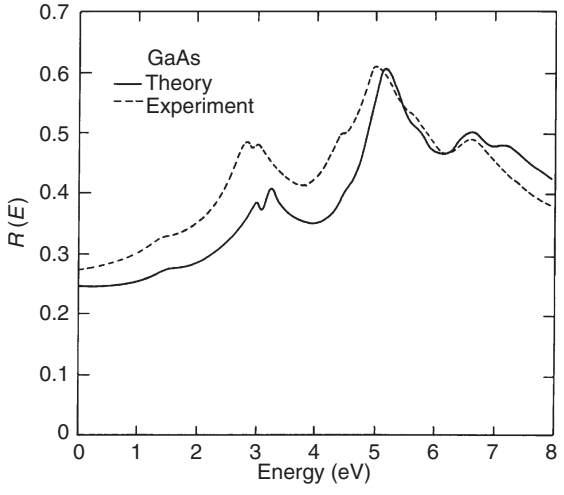


**Figure 23** Reflectivity spectrum for gallium arsenide. The experiment is from Philipp and Ehrenreich (1963).

Thus, the $E_0$ structure actually has a doublet character. The $E_1$ reflectivity peak originates from transitions near the zone boundary at $L$. Again, spin–orbit interactions can become important in this region by splitting the upper valence bands. We note that older, local pseudopotential calculations gave a different critical point symmetry for this region. The local potential results put an $M_0$ critical point at $L$ and an $M_1$ critical point along the $\Lambda$ direction. The nonlocal results move the $M_1$ critical point to the zone boundary and eliminate the $M_0$ point altogether. This latter configuration is in better accord with experiment.

The $E'_0$ structure has been somewhat controversial. It can arise from either of two regions: near the zone center or along the $\Delta$ direction. If it were to occur at the zone center, we would expect spin–orbit splitting to be important in both the valence and conduction bands. On this basis, recent work by Aspnes and Studna (1973) has given support to the zone-center assignment. The traditional argument against this assignment has been the small phase space associated with the $\Gamma$ point; however, excitonic effects may very well enhance its importance.

The $E_2$ peak, which dominates the absorption spectrum, represents the average bonding–antibonding transitions. This peak arises from transitions from the uppermost valence bands to the lowest conduction bands near the points $(2\pi/a)(3/4, 1/4, 1/4)$ in the Brillouin zone. This assignment is similarly found in the diamond structure semiconductors. The $E_2$ peak dominates the spectrum because of the large phase space which contributes to interband transitions in this region and because of large interband dipole matrix elements. At energies above the $E_2$ peak, reflectivity structure from transitions along the $\Delta$ direction.

The $E'_1$ structure arises from transitions from the top valence band to the second lowest conduction band at or near the zone boundary at $L$.

By examining the energy gradients and dipole matrix elements throughout the Brillouin zone, it is possible to determine the origin of structure of the imaginary part of the dielectric function. In such a manner, we analyzed the contribution to the $E'_0$ reflectivity structure. This structure is complex and has been somewhat controversial. In **Figure 24**, we label this structure as $A$, $B$, and $C$.

Rehn and Kyser (1972), using electroreflectance, observed only a $\Lambda$ symmetry for this structure and attributed the structure to the pseudocrossing of the $\Delta_5$ conduction bands. However, Aspnes and Studna (1973) have pointed out that this interpretation conflicts with band structure calculations where some $\Gamma$ symmetry structure is predicted. They further proposed that the $\Lambda$ symmetry structure arises from a pair of $M_1$ critical points approximately one-tenth way from $\Gamma$ to $X$. EPM calculations agree with their interpretation (Cohen and Chelikowsky, 1989). The relevant transitions are indicated in **Figure 25**. Aspnes and Studna also noted the possibility of the pseudocrossing producing some very weak structure at 4.4 eV. The 4.4 eV ($A$) transition corresponds to this type of transition. It should be noted, however, that there also exists a companion $M_0$ critical point owing to the spin–orbit splitting of the $\Delta_5$ valence band. Since this companion critical point occurs

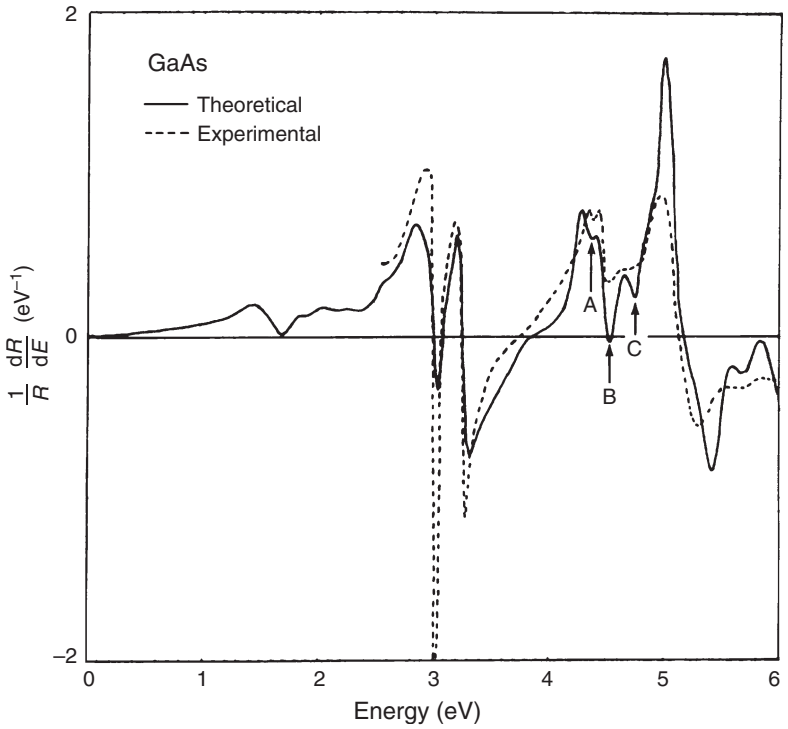


**Figure 24** Reflectivity derivative spectrum of gallium arsenide. See the text for a discussion of the structures: *A*, *B*, and *C*. Experiment is from Zucca and Shen (1970).

about 0.1 eV higher in energy, it is nearly degenerate with the $E'_0$ structure from $\Gamma$ and $\Delta$ at 4.5 eV. In the calculated derivative spectrum, this structure is masked by the stronger $M_1$ critical points, and this may be the case in the electroreflectance measurements.

In **Figure 26**, we illustrate the calculated valence band density of states for GaAs, and the results of XPS measurements and inverse photoemission (Chelikowsky *et al.*, 1989). As for the diamond structure semiconductors, we may divide the density of states into three general regions. The first region is the most tightly bound energy band. Electron states corresponding to this band are strongly localized on the anion and are descendants of the atomic As 4s states. The next region of note is a peak arising from the onset of the second valence band. This band shows almost no energy variation along the *X*–*U* symmetry direction; in fact, it is very flat over the entire square face of the Brillouin zone. This energy band configuration results in a sharp onset of states above the antisymmetric gap. The character of states associated with the second valence band changes from predominantly cation s-like states at the band edge to predominantly anion p-like states at the band maximum. The third region of interest in the density of states extends from the onset of the third valence band (at about 4 eV below the valence band maximum) to the valence band maximum. This region encompasses the top two valence bands and is predominantly p-like and is associated with anion states.

***1.01.4.4.2(iii) Zinc selenide*** The prototypical II–VI semiconductor is zinc selenide. However, unlike the prototypical III–V semiconductor (GaAs) or our prototypical diamond semiconductor (Si or Ge), we do not have as detailed a picture for this class of semiconductors. There are several reasons for this situation. First, extensive experimental information is lacking for ZnSe as compared to GaAs. The larger bandgap in ZnSe requires higher photon energies than GaAs for reflectivity measurements. Photon sources at these higher energies have not been routinely available. Also, some of the powerful optical techniques such as Schottky barrier electroreflectance cannot be easily applied to ZnSe owing to fabrication problems. Second, the Zn 3d-level resides close to the valence band. Traditionally, empirical pseudopotential band calculations do not explicitly

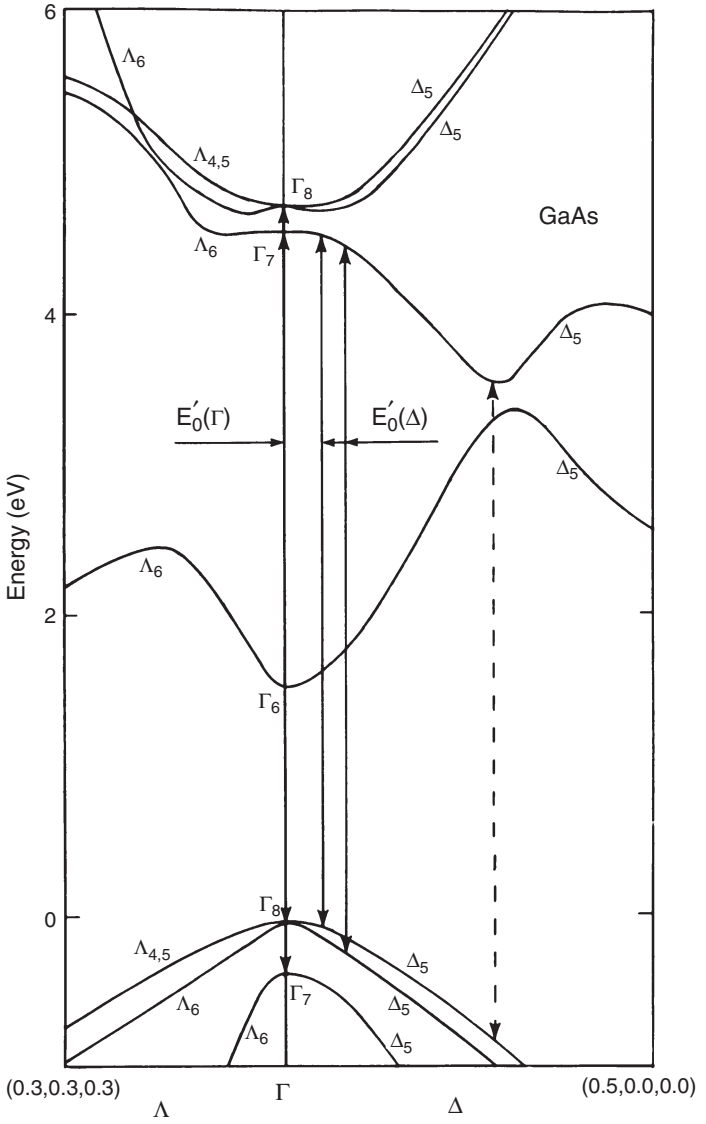


**Figure 25** Band structure of gallium arsenide near $\Gamma$ showing the critical point locations for the $E_0'$ structure. The indicated transitions give rise to the structure labeled *A*, *B*, and *C* in **Figure 24**.

include this level in ZnSe. The 3d cation level in gallium arsenide is significantly below the valence band and in Ge the 3d level is more than $\sim$30 eV below the top of the valence band.

In **Figure 27**, we illustrate the band structure of ZnSe. Most features of this band structure can be accounted for by extrapolating from Ge to GaAs. Compared to Ge and GaAs, the bandgaps in ZnSe, that is, the optical gap and the antisymmetric gap, have increased considerably in size. Moreover, the valence bands have narrowed and show less dispersion.

In **Figure 28**, we display the calculated and measured reflectivity spectrum for ZnSe (Freeouf, 1973; Walter *et al.*, 1970). Some work using derivative spectroscopy is available (Theis, 1977). Owing to the increased ionic component of the bonding in ZnSe, the reflectivity spectrum is shifted to higher energy as compared to Ge or GaAs. It also exhibits much more structure than Ge or GaAs because of the larger separation between bands.

For the most part, the reflectivity spectrum can be analyzed by analogy with GaAs. The $E_0$ reflectivity

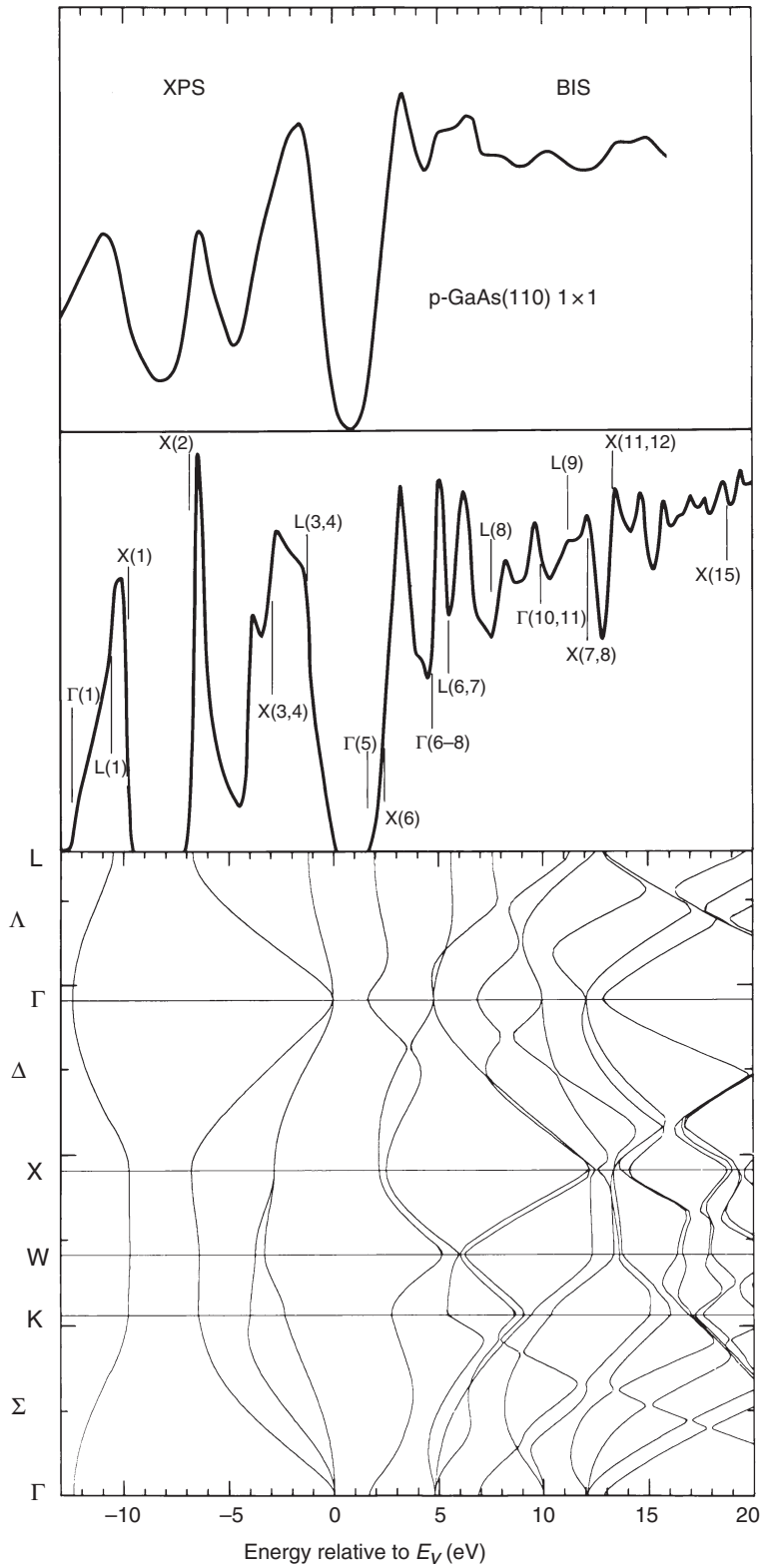


**Figure 26** Experimental X-ray photoemission spectroscopy (XPS) and Bremsstrahlung isochromat spectroscopy (BIS) or inverse photoemission spectra for crystalline gallium arsenide (Chelikowsky *et al.*, 1989). The critical point features of the theoretical density of states are also identified with band structure features.

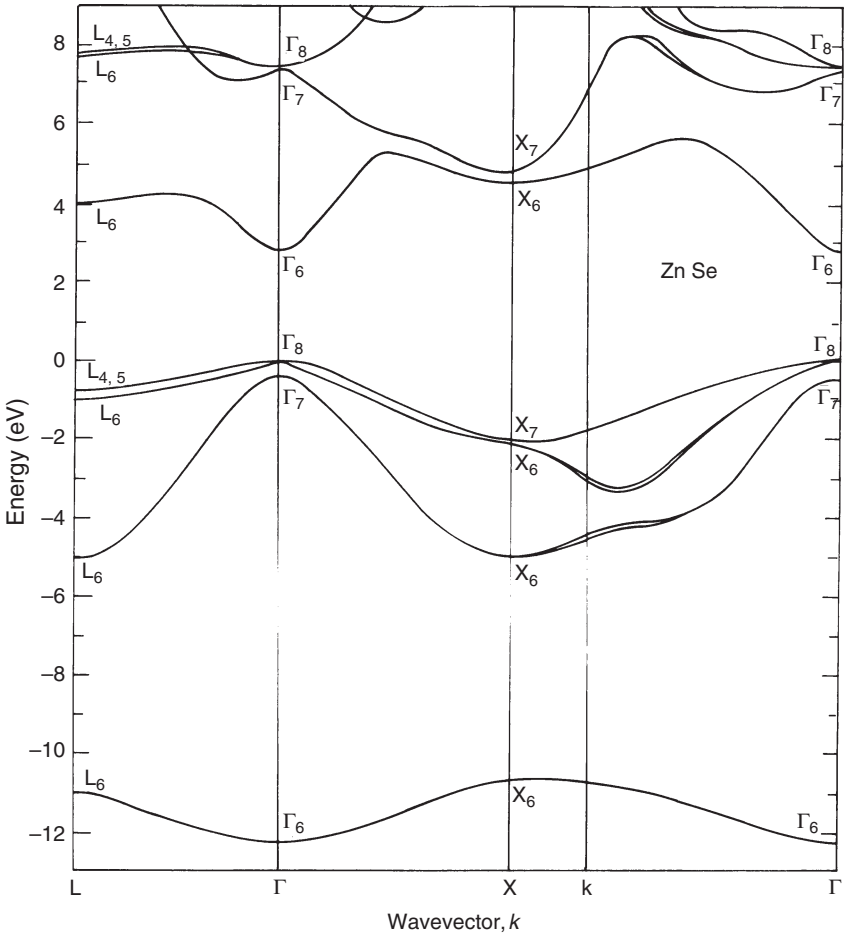


**Figure 27** Energy bands for crystalline zinc selenide. The top of the valence band is taken to be the zero energy reference.

structure lies between 2 and 4 eV and corresponds to the minimum energy gap at Γ. The $E_1$ structure at 4–5 eV arises from transitions near to or at the $L$ point. The structure corresponding to the $E_2$ peak arises from a localized region near the special point $(2\pi/a)(3/4, 1/4, 1/4)$. The $E_0'$ structure at 7–8 eV may be an exception having no analogy in GaAs. It lies above the $E_2$ structure; thus, the origin of structure is reversed in ZnSe as compared to GaAs. However, the origin of the $E'_0$ structure appears to be the same, that from near Γ and along the Δ direction. The highest energy reflectivity structure, $E'_1$, occurs at 8–9 eV and corresponds to transitions near the $L$ point.

With the exception of the line shape near the $E_2$ peak, the pseudopotential results are consistent with experiment. It is possible that the line-shape discrepancy is an artifact of the sampling scheme used. We note that at energies above 10 eV or so, the structure in the reflectivity might arise from transitions involving the Zn 3d-states.

It should not be surprising that the density of states of ZnSe may also be interpreted in terms of an extrapolation from the Ge and GaAs band structures. In **Figure 29**, we compare the calculated density of states for ZnSe with photoemission measurements and inverse photoemission (Chelikowsky *et al.*, 1989). Some of the early pseudopotential band structures, which had potentials based on optical data alone, were in mediocre agreement with the photoemission results. Specifically, local pseudopotentials chosen to reproduce optical gaps tend to overestimate the ionicity of the II–VI semiconductors. As a consequence, the valence bands become quite narrow compared to experiment. Nonlocal pseudopotentials do not suffer from this malady; they can be used to fit optical gaps with no corresponding increase in ionicity.

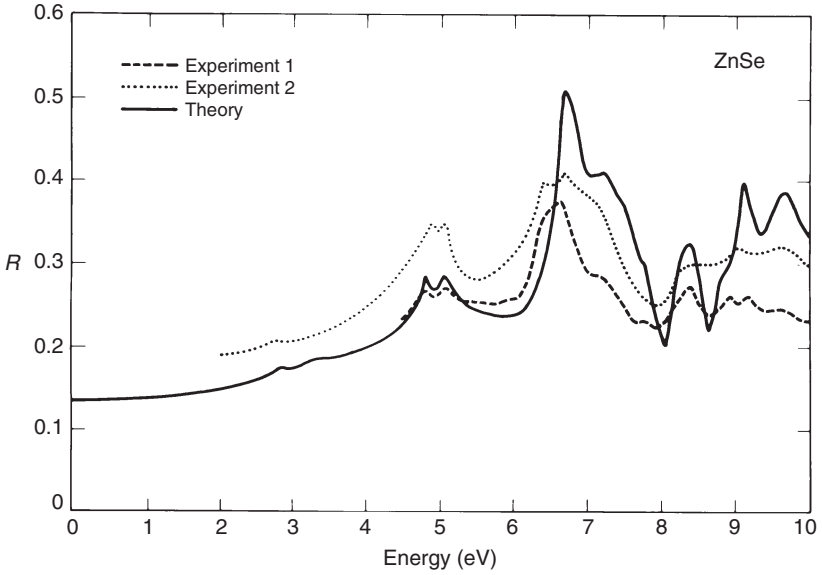


**Figure 28** Reflectivity spectrum for zinc selenide. Two experimental spectra are presented: the dashed line is from Walter *et al.* (1970) and the dotted line from Freeouf (1973).

## 1.01.5 The *Ab Initio* Pseudopotential Method

A deficiency of the empirical pseudopotential method is that the potential is 'biased' by the experimental data to which it was fit. For example, suppose we fit the optical properties of GaSb and extract a Ga potential, we must be careful in trying to use this fit Ga potential in another crystal. This potential might work fine in GaAs, which has a similar bonding configuration when compared to GaSb, but the potential may fail to describe GaN, which is considerably more ionic than GaSb. A related problem is that the form factors are fit only to a subset of reciprocal lattice vectors. If the crystal volume or structure changes, one must extrapolate to a new subset. Another problem can occur at a surface. In crystalline GaAs, each Ga atom is surrounded by four As atoms. For the cleavage plane the (110) surface, the Ga atom is bonded to three As atoms. Given the coordination change, the surface Ga atom cannot be expected to retain a 'bulk-like' screening potential. There is no reason to be confident that the Ga pseudopotential extracted from the crystalline environment will be very accurate at the surface. Of course, this problem is made worse if one wants to examine a liquid containing Ga atoms where the coordination may continuously change.

We can consider the form factors to be composed of two interactions: the ion core–valence electron interaction and the valence electron–valence electron interaction. A fundamental postulate of the pseudopotential method is that the ion-core pseudopotential is not dependent on the chemical environment. We assume that this part of the potential can be transferred with no loss in accuracy. The key problem is to determine the total pseudopotential by determining the potentials from the valence electrons interacting among themselves.

A common procedure is to construct a self-consistent potential. The wave functions obtained in an electronic structure calculation can be used to construct a new screening potential, which in turn can be used with the ion-core potential to compute a new total potential and new wave functions. When no changes occur in a feedback loop of this kind, the solution is considered to be self-consistent.

Let us assume a one-electron Hamiltonian, which can be based on density functional theory (Hohenberg and Kohn, 1964; Kohn and Sham, 1965):

$$\left[\frac{-\hbar^2\nabla^2}{2m} + V_{ion}^{p}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r})\right]\Psi_n(\mathbf{r}) = E_n\Psi_n(\mathbf{r}) \tag{50}$$

The ion-core pseudopotential, $V_{ion}^{p}$, can be taken as a linear superposition of ion-core atomic potentials. Determining the ionic potential can be accomplished by resorting to atomic structure

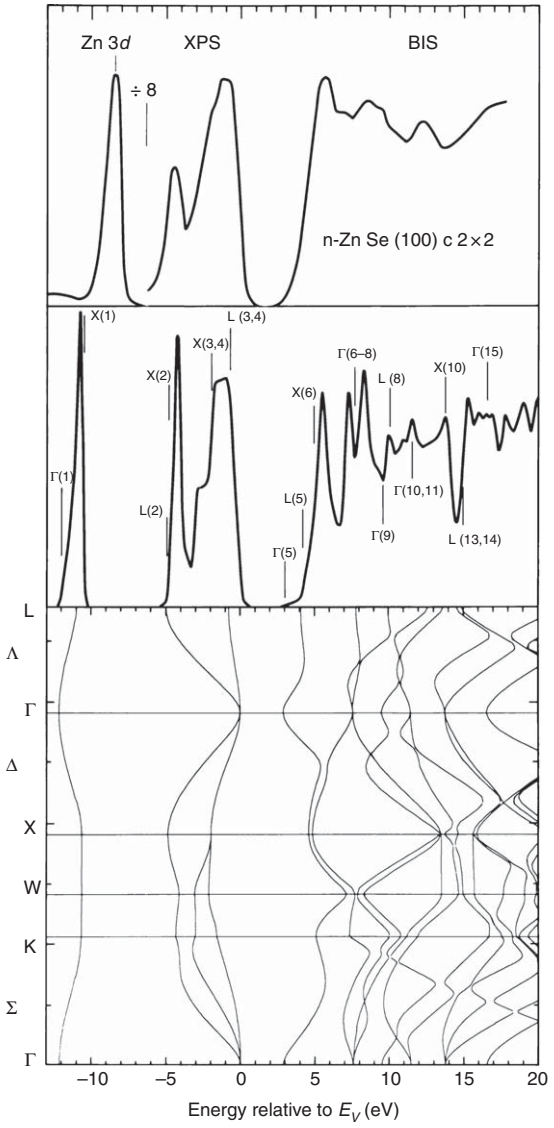


**Figure 29** Experimental X-ray photoemission spectroscopy (XPS) and Bremsstrahlung isochromat spectroscopy (BIS) or inverse photoemission spectra for crystalline zinc selenide (Chelikowsky *et al.*, 1989). The critical point features of the theoretical density of states are also identified with band structure features.

calculations, as discussed in the following section. The potential arising from the valence-electron interactions can be divided into two parts. One part represents the classical electrostatic terms, the Hartree or Coulomb potential:

$$\nabla^2 V_H(\mathbf{r}) = -4\pi e\rho(\mathbf{r})$$
$$\rho(\mathbf{r}) = e \sum_{n,occup} |\Psi_n(\mathbf{r})|^2 \qquad (51)$$

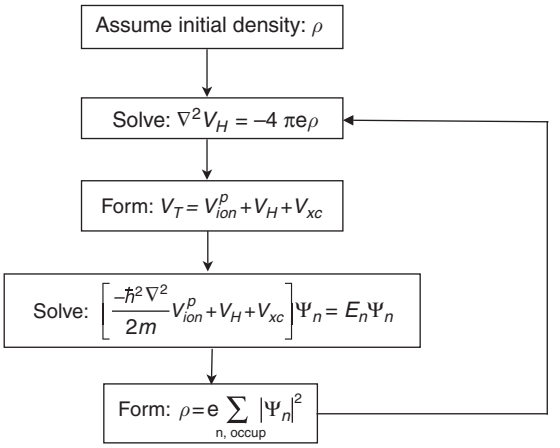


**Figure 30** Self-consistent field loop. The loop is repeated until the input and output charge densities are equal to within some specified tolerance.

where $\rho$ is the valence electron charge density. The second part of the screening potential, the exchange-correlation part of the potential, $V_{xc}$, is quantum mechanical in nature. A common approximation for this part of the potential arises from the local density approximation, that is, the potential depends only on the charge density at the point of interest, $V_{xc}(\mathbf{r}) = V_{xc}[\rho(\mathbf{r})]$. In principle, density functional theory is exact, provided one can obtain an exact functional for $V_{xc}$. This is an outstanding problem. It is commonly assumed that the functional extracted for a homogeneous electron gas (Ceperley and Alder, 1980) is universal and can be applied to the inhomogeneous gas problem. The procedure for generating a self-consistent field (SCF) potential is given in **Figure 30**. The SCF cycle is initiated with a potential constructed by a super-position of atomic densities. These densities are then used to solve a Poisson equation for the Hartree potential, and a density functional is used to obtain the exchange-correlation potential. A screening potential composed of the Hartree and exchange-correlation potentials is then added to the fixed ion-core pseudopotential, after which the one-electron Schrödinger equation, or Kohn–Sham equation, is solved. The resulting wave functions from this solution are then employed to construct a new potential and the cycle is repeated. In practice, the output and input potentials are mixed using a scheme that accounts for the history of the previous iterations (Broyden, 1965; Chelikowsky and Cohen, 1992).

### 1.01.5.1 Constructing Pseudopotentials from Density Functional Theory

The construction of ion-core pseudopotentials has become an active area of electronic structure theory. Methods for constructing such potentials have centered, on *ab initio* or first-principles pseudopotentials; that is, the informational base on which these potentials are based does not involve any experimental input.

The first step in the construction process is to consider an electronic structure calculation for a free atom. For example, in the case of a silicon atom, the Kohn–Sham equation (Kohn and Sham, 1965) can be solved for the eigenvalues and wave functions. Knowing the valence wave functions, that is, $3s^2$ and $3p^2$, and corresponding eigenvalues, the pseudowave functions can be constructed. This is an easy numerical calculation as the atomic densities are assumed to possess spherical symmetry and the problem reduces to a one-dimensional radial integration.

Once we know the solution for an all-electron potential, we can invert the Kohn–Sham equation and find the total pseudopotential. We can unscreen the total potential and extract the ion-core pseudopotential. This ion-core potential, which arises from tightly bound core electrons and the nuclear charge, is not expected to change from one environment to another. It should be transferable from the atom to a molecular state or to a solid state or liquid state. The issue of this transferability is one which must be addressed according to the system of interest. The immediate issue here is how to define pseudo-wavefunctions which can be used to define the corresponding pseudopotential.

Suppose we insist that the pseudo-wave-function be identical to the all-electron wave function outside of the core region. For example, let us consider the 3s state for a silicon atom. We want the pseudo-wavefunction to be identical to the all-electron state outside the core region:

$$\phi^p_{3s}(r) = \psi_{3s}(r) \quad r > r_c \tag{52}$$

where $\phi^p_{3s}$ is a pseudo-wave-function and $r_c$ defines the core size.

This assignment will guarantee that the pseudo-wave-function will possess properties identical to the all-electron wave function, $\psi_{3s}$ in the region away from the ion core.

For $r < r_c$, we alter the all-electron wave function. We are free to do this as we do not expect the valence wave function within the core region to affect the chemical properties of the system. We choose to make the pseudo-wave-function smooth and nodeless in the core region. This will provide rapid convergence with simple basis functions. One other criterion is mandated. Namely, the integral of the pseudocharge density within the core should be equal to the integral of the all-electron charge density. Without this condition, the pseudo-wavefunction differs by a scaling factor from the all-electron wave function. Pseudopotentials constructed with this constraint are called norm conserving (Hamann *et al.*, 1979). Since we expect the bonding in a solid to be highly dependent on the tails of the valence wave functions, it is imperative that the normalized pseudo-wave-function be identical to the all-electron wave functions.

There are many ways of constructing norm conserving pseudopotentials as within the core the pseudo-wave-function is not unique. One of the most straight-forward construction procedures is from Kerker (1980) and later extended by Troullier and Martins (1991):

$$\phi^p_l(r) = \begin{cases} r^l \exp(p(r)) & r \le r_c \\ \psi_l(r) & r > r_c \end{cases} \tag{53}$$

$p(r)$ is taken to be a polynomial of the form:

$$p(r) = c_0 + \sum_{n=1}^{6} c_{2n} r^{2n} \tag{54}$$

This form assures us that the pseudo-wave-function is nodeless and by taking even powers there is no cusp associated with the pseudo-wave-function. The parameters, $c_{2n}$, are fixed by the following criteria: (1) The all-electron and pseudo-wavefunctions have the same valence eigenvalue. (2) The pseudo-wave-function is nodeless and is identical to the all-electron wave function for $r > r_c$. (3) The pseudo-wave-function must be continuous as well as the first four derivatives of the wave function at $r_c$. (4) The pseudopotential has zero curvature at the origin. This construction is easy to implement and extend to include other constraints. An example of an atomic pseudo-wavefunction for Si is given in **Figure 31** where it is compared to an all-electron wave function. Unlike the 3s all-electron wave function, the pseudo-wavefunction has no nodes. The pseudo-wave-function is much easier to express as a Fourier transform or a combination of Gaussian orbitals than the all-electron wave function.

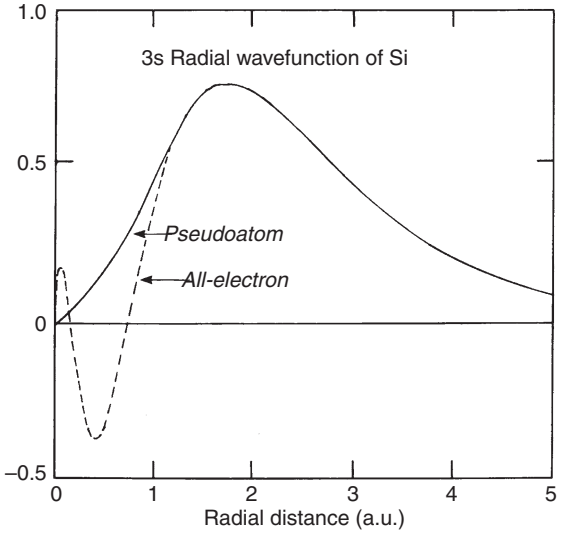


**Figure 31** An all-electron and a pseudo-wave function for the silicon 3s radial wave function.

Once the pseudo-wave-function is constructed, then the Kohn–Sham equation can be inverted to arrive at the ion-core pseudopotential:

$$V^p_{ion,l}(r) = \frac{\hbar^2\nabla^2\phi^p_l}{2m\phi^p_l} - E_{n,l} - V_H(r) - V_{xc}[\rho(r)] \qquad (55)$$

The ion-core pseudopotential is well behaved as $\phi^p$, has no nodes and does not vanish. The ion-core potential appears to be both state dependent and energy dependent. The energy dependence is usually weak. For example, the 4s state in silicon computed by the pseudopotential constructed from the 3s state is usually accurate. Physically, this happens because the 4s state is extended and experiences the potential in a region where the ion-core, potential has assumed a simple $-Z_v e^2/r$ behavior. However, the state dependence through $l$ is an issue; the difference between a potential generated via a 3s state and a 3p can be an issue. In particular, for first row elements such as C or O, the nonlocality is quite large as there are no p states within the core region. For the first row transition elements for such as Fe or Cu, this is also an issue as again there are no d-states within the core. This state dependence can be addressed in similar fashion as for a nonlocal empirical pseudopotential.

An additional advantage of the norm conserving potential concerns the logarithmic derivative of the pseudo-wave-function (Bachelet *et al.*, 1982). An identity exists:

$$-2\pi\left((r\phi)^2\frac{d^2\ln\phi}{dE\,dr}\right)_R = 4\pi\int_0^R \phi^2 r^2\,dr = \underline{Q}(R) \qquad (56)$$

The energy derivative of the logarithmic derivative of the pseudo-wave-function is fixed by the amount of charge within a radius, $R$. The radial derivative of the wave function $\phi$ is related to the scattering phase shift from elementary quantum mechanics. For a norm-conserving pseudopotential, the scattering phase shift at $R = r_c$ and at the eigenvalue of interest is identical to the all-electron case as $\underline{Q}_{all\ elect}(r_c) = \underline{Q}_{pseudo}(r_c)$. The scattering properties of the pseudopotential and the all-electron potential have the same energy variation to first order when transferred to other systems.

There is some flexibility in constructing pseudopotentials. The nonuniqueness of the pseudo-wave-function was recognized early in its inception. For example, within the Phillips–Kleinman formulation, one can always add a function, $f$, to the pseudo-wave-functions without altering the pseudopotential provided $f$ is orthogonal to the core states. Consider the matrix element in Equation (7). If one changes $\Phi^p_v$ to $\Phi^p_v + f$, then $\langle\Psi_c|\Phi^p_v + f\rangle = \langle\Psi_c|\Phi^p_v\rangle + \langle\Psi_c|f\rangle = \langle\Psi_c|\Phi^p_v\rangle$. Nothing is changed in the Phillips–Kleinman pseudopotential by this addition.

The nonuniqueness of the pseudopotential can be exploited to optimize the convergence of the pseudopotentials for the basis of interest. Much effort has been made to construct soft pseudopotentials. By soft, one means a rapidly convergent calculation using plane waves as a basis. Typically, soft potentials are characterized by a large core size, that is, a larger value for $r_c$. However, as the core becomes larger, the goodness of the pseudo-wave-function can be compromised as the transferability of the pseudopotential becomes more limited. In **Figure 32**, the ion-core pseudopotential for carbon is plotted in real space and in reciprocal space. Although the potentials look quite different in the core region, they all give reliable electronic structure properties for carbon.

As for the nonlocal terms that occur in empirical potentials (Section 1.01.4.3), we can handle nonlocality using a plane wave basis in a straightforward fashion. If we consider a reference potential such that $\Delta V_l = V_{local} - V_l$, then we need to determine matrix elements as in Equation (46) where $V^p_{c,l}$ is replaced by $\Delta V_l$. The reference potential can be chosen to be a local potential. In general, the nonlocal elements are very short ranged in real space as the

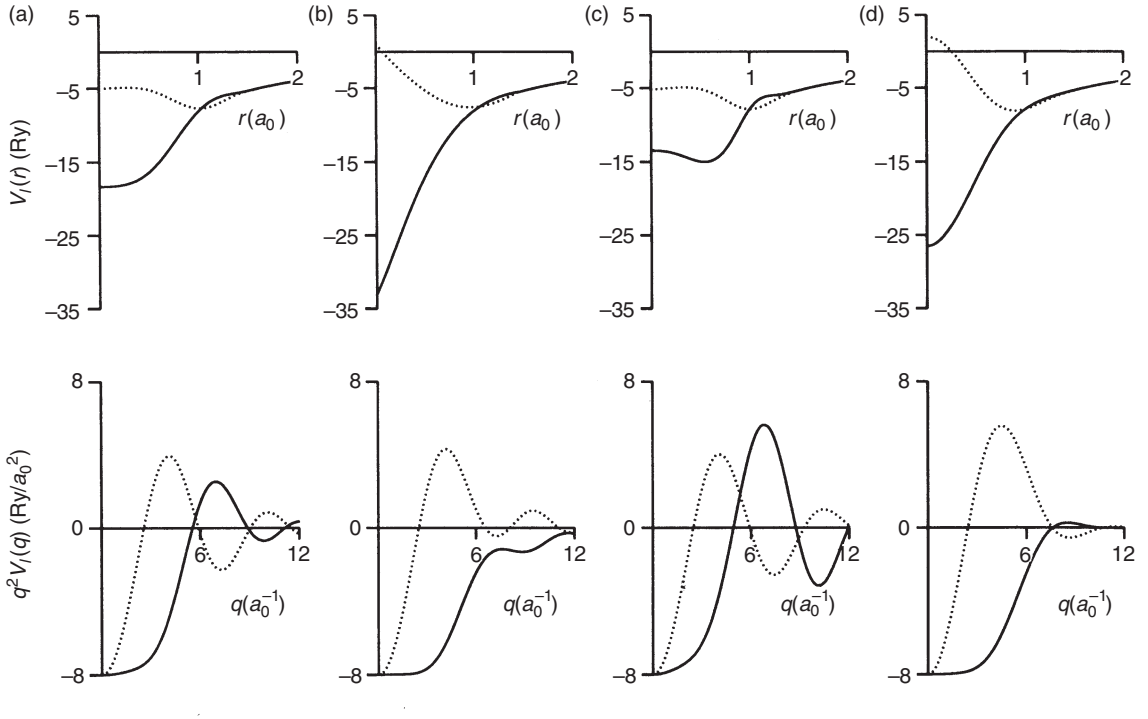


**Figure 32** Ion-core pseudopotentials for carbon generated by four different methods in real and reciprocal space; (a) Troullier and Martins (1991), (b) Kerker (1980), (c) Hamann *et al.* (1979), and (d) Vanderbilt (1985). The dotted and solid lines correspond to the s and p pseudopotentials, respectively.

size of the core is determined by $r_c$. As such, the upper limit of the real space integral is slightly more than $r_c$ in practice.

A major strength of *ab initio* pseudopotentials is that they can be employed for systems with many atoms of low symmetry, for example, clusters, liquids, and surfaces, without having to fit any form factors. However, a plane wave basis will require a large cutoff for such systems. In this case, the matrix elements in Equation (46) can be difficult to evaluate as elements depend on both $\mathbf{k}+\mathbf{G}$ and $\mathbf{k}+\mathbf{G}'$. Kleinman and Bylander (1982) suggested an alternate form for the required nonlocal matrix elements:

$$\Delta V_l^{KB}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') = \frac{\sum_m Y_{lm}^*(\widehat{\mathbf{k}+\mathbf{G}})\,Y_{lm}(\widehat{\mathbf{k}+\mathbf{G}'})}{4\pi\Omega_a \int \Phi_l^p(r)\Delta V_l(r)\Phi_l^p(r)dr} \times \int \Phi_l^p(r)\Delta V_l(r) j_l\big(|\mathbf{k}+\mathbf{G}'|r\big)r^2 dr \times \int \Phi_l^p(r)\Delta V_l(r) j_l\big(|\mathbf{k}+\mathbf{G}|r\big)r^2 dr \quad (57)$$

where $\Phi_l^p$ is the atomic reference pseudo-wave-function and $Y_{lm}^*$ is a spherical harmonic with the angles determined by the unit vector: $\widehat{\mathbf{k}+\mathbf{G}}$. This matrix element form has a great advantage as the integrals are separable in that they are solely a function of $\mathbf{k}+\mathbf{G}$ or a function of $\mathbf{k}+\mathbf{G}'$. The individual integrals can be stored and retrieved as required in setting up the matrix.

It is also possible to solve the Kohn–Sham problem directly in real space (Chelikowsky *et al.*, 1994). In this case, the Kleinman–Bylander form can be cast as

$$\Delta V_l^{kB}(x,y,z)\phi^p(x,y,z) = \sum_{lm} G_{lm} u_{lm}^p(x,y,z)\Delta V_l(x,y,z)$$

$$G_{lm} = \frac{\int u_{lm}^p \Delta V_l \phi^p\, dxdydz}{\int u_{lm}^p \Delta V_l u_{lm}^p\, dxdydz} \quad (58)$$

where $u_{lm}^p$ are the reference atomic pseudo-wave-functions. The nonlocal nature of the pseudopotential is apparent from the definition of $G_{lm}$; the value of these coefficients are dependent on the pseudo-wave-function, $\phi^p$, acted on by the operator $\Delta V_l$. This is very similar in spirit to the pseudopotential defined by Phillips and Kleinman (1959). While we have focused on a simple plane wave basis, there are other bases that can be employed, for example, one

can combine pseudopotentials with Gaussians (Chan *et al.*, 1986) or a uniform grid (Chelikowsky *et al.*, 1994). These real space methods can be easily implemented with real space *ab initio* pseudopotentials (Kronik *et al.*, 2006).

While our focus has been on norm conserving pseudopotentials, it is possible to generalize the pseudopotential method to include systems that are not based on such potentials. It is possible to make the pseudopotentials even weaker by relaxing this condition. Vanderbilt (1990) proposed such a method for constructing ultrasoft pseudopotentials, which are constructed as a generalized eigenvalue problem. The norm conservation constraint is relaxed, and the charge density within the ion core is not explicitly considered as part of the pseudo-wave-function. The relaxation of the norm conservation constraint allows one to consider a much larger core radius and a much softer potential.

One issue which is relevant for pseudopotential constructions, regardless of whether the potential is intended for use with a plane wave basis or not, concerns the issue of unbound, or weakly bound, atomic states. If an atom does not bind a state of interest, then the atomic wave function corresponding to this state is clearly not normalizable. Nonetheless, the pseudopotential corresponding to this state might be of some interest, for example, in a crystal such diverging wave functions are captured by the potentials of neighboring atoms. For example, Ba has a strong f-component resonance. However, these f-states are not bound for neutral Ba atom. In order to bind such states, one must consider highly ionized atomic states, which result in very strong pseudopotentials. Sometimes these potentials are so strong as to be useless for a plane wave basis, or so far removed from the chemical environment of interest that their transferability may be suspect. Hamann (1989) has suggested a method for handling such cases by integrating out the Kohn–Sham equation to a large distance and at that point terminating the pseudo-wave-function. The corresponding terminated wave function is then used to generate a pseudopotential for the component of interest.

## 1.01.5.2 Structural Properties of Semiconductor Crystals

### *1.01.5.2.1 Total electronic energy from pseudopotential–density functional theory*

Empirical pseudopotentials have been one of the most effective tools in understanding the optical and dielectric properties of semiconductor crystals (Cohen and Chelikowsky, 1989); however, to understand structural properties, we need to employ a different approach. Since the form factors in the empirical pseudopotential have been fit to optical transitions, there is no reason to believe that they will be very accurate for structural properties such as the phase stability, the equilibrium bond length, the bulk modulus, and the phonon spectrum of a crystal. *Ab initio* pseudopotentials can be used for this purpose and, in this section, we illustrate their application to the phase stability of tetrahedral semiconductors such as silicon.

To evaluate the total energy of a crystal can be a difficult task. The total energy of the system contains terms that individually diverge, for example, the repulsive Coulomb terms between the ion cores and the electron–electron interactions. These terms can be handled in Fourier or momentum space using a plane wave basis (Ihm *et al.*, 1979). The total energy of the system can be written as

$$E_T = E_{kin} + E_{ec} + E_H + E_{xc} + E_{cc} \tag{59}$$

$E_T$ is the total electronic energy, $E_{kin}$ represents the electronic kinetic energy, the electron–ion-core interaction energy is given by $E_{ec}$, the electrostatic Coulomb energy is given by the Hartree energy, $E_H$, the nonclassical exchange-correlation energy is given by $E_{xc}$, and the ion core–ion-core classical term is given by $E_{cc}$. The momentum-space expressions for the electronic energy terms are as follows:

$$E_{kin} = \frac{1}{N}\sum_{n,\mathbf{k},\mathbf{G}} \left|\phi_n^p(\mathbf{k}+\mathbf{G})\right|^2 \frac{\hbar^2|\mathbf{k}+\mathbf{G}|^2}{2m} \tag{60}$$

$$E_{ec} = \frac{1}{N}\sum_{n,\mathbf{k},\mathbf{G}} \phi_n^{p^*}(\mathbf{k}+\mathbf{G})\phi_n^p(\mathbf{k}+\mathbf{G}') \times\left(V_c^p(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') + \delta_{\mathbf{G},\mathbf{G}'}\frac{1}{\Omega_c}\int \frac{Ze^2}{r}d^3r\right) \tag{61}$$

$$E_H = \frac{\Omega_c}{2}\sum_{\mathbf{G},G\neq 0}\frac{4\pi e^2}{G^2}\left|\rho(\mathbf{G})\right|^2 \tag{62}$$

$$E_{xc} = \frac{\Omega_c}{2}\sum_{\mathbf{G}}\rho^*(\mathbf{G})\varepsilon_{xc}(\mathbf{G}) \tag{63}$$

$$E_{cc} = \frac{e^2}{2}\sum_{S,S'} Z_{v,s}Z_{v,s'}\left\{\frac{4\pi}{\Omega_c}\sum_{\mathbf{G},G\neq 0}\left[\frac{1}{G^2}\cos(\mathbf{G}\cdot(\boldsymbol{\tau}_S-\boldsymbol{\tau}_{S'}))\times\exp\left(\frac{-G^2}{4\eta^2}\right)\right] - \frac{\pi}{\Omega_c\eta^2} + \sum_{\mathbf{R},R\neq 0}\left[\frac{\operatorname{erfc}(\eta|\mathbf{R}-\boldsymbol{\tau}_S+\boldsymbol{\tau}_{S'}|)}{|\mathbf{R}-\boldsymbol{\tau}_S+\boldsymbol{\tau}_{S'}|}\right] - \frac{2\eta}{\sqrt{\pi}}\delta_{S,S'}\right\} \tag{64}$$

Each term yields the energy per cell; $N$ is the total number of cells, $\Omega_c$ is the cell volume, $\mathbf{R}$ is a lattice vector, $\mathbf{G}$ is a reciprocal lattice vector, $Z$ is the total core charge, which is a sum of the individual core charges $Z_v$, $\boldsymbol{\tau}$ is the basis vector, $\mathbf{k}$ is the crystal momentum, and $\phi_n^p$ is the wave function for state $n$. $\rho(\mathbf{G})V_c^p(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}')$, and $\phi_n^p(\mathbf{k}+\mathbf{G})$ represent the Fourier components of the charge, ion-core potential and the pseudo-wave-function. The sums involving the pseudo-wave-functions run only over the occupied states, $\eta$ is a parameter that controls the convergence of the Ewald summation in real space versus momentum space (Ihm *et al.*, 1979).

In practice, it is easier to make use of the eigenvalue explicitly and subtract off the double counting terms, that is, one can write the total energy as

$$E_T = \frac{1}{N}\sum_{n,\mathbf{k}} E_n(\mathbf{k}) - E_H + \Delta E_{xc} + E_{cc} \tag{65}$$

The sum is over all occupied states. The eigenvalue term contains the exchange-correlation term, whereas the total energy should contain the exchange-correlation energy. This term can be written as

$$\Delta E_{xc} = \Omega_c \sum_{\mathbf{G}} \rho^*(\mathbf{G})[\varepsilon_{xc}(\mathbf{G}) - V_{xc}(\mathbf{G})] \tag{66}$$

where $\varepsilon_{xc}$ is related to the exchange correlation via a functional derivative:

$$\Delta V_{xc} = \frac{\delta(\rho\,\varepsilon_{xc}[\rho])}{\delta\rho} \tag{67}$$

where, in the local density approximation,

$$E_{xc}[\rho] = \int \rho(\mathbf{r})\varepsilon_{xc}[\rho(\mathbf{r})]d^3r \tag{68}$$

In the simplest formalism, $\varepsilon_{xc}$ is extracted from a homogeneous electron gas (Kohn and Sham, 1965).

#### 1.01.5.2.2 Phase stability of crystals

The total energy, $E_T$, of a crystal structure can be calculated using Equation (65) as a function of the lattice constant and any internal structural parameters. One of the first applications of this work was carried out by Ihm and Cohen (1980), and Yin and Cohen (1980, 1982c, 1982b). This work represents a seminal point in condensed matter physics (Chelikowsky, 2000). It demonstrated a workable scheme for examining the structural properties of solids and its extension led to molecular dynamics using quantum forces (Car and Parrinello, 1985).

The first application involved the silicon crystal with a number of different phases: face-centered cubic, body-centered cubic, hexagonal close-packed, hexagonal diamond, cubic diamond, white tin and simple cubic. For each structure, the energy was minimized for a given volume by optimizing the internal coordinates, for example, for a hexagonal structure the $c/a$ ratio was optimized. Depending on the system studied the range of volumes considered is varied. In the case of silicon. Cohen, *et al.* considered a volume of one-half the ambient volume of the known diamond phase of silicon to an upper limit of about 20% larger than the ambient volume. For such a diverse set of crystal structures, it is important to make certain the convergence is similar to all phases. For structures which yield metallic systems, the $\mathbf{k}$-point sampling is sometimes critical because the variations in the Fermi surface make it difficult to assure that occupied states are being sampled. Once the total energy is computed for a finite number of volumes, the energy versus volume points are fit to an equation of state such as the Birch (1952) equation or the Murnaghan (1944) equation.

$$E_T(V) = E_T(V_0) + \frac{B_0 V}{B_0'}\left(\frac{(V_0/V)^{B_0'}}{B'_0 - 1} + 1\right) - \frac{B_0 V_0}{B_0' - 1} \tag{69}$$

where $B_0$ and $B_0'$ are the bulk modulus and its pressure derivative at the equilibrium volume $V_0$. Using Equation (69), the fit of the calculated points yields the equilibrium energy $E_T(V_0)$, the equilibrium volume, and the bulk modulus and its pressure derivative at equilibrium. Representative calculated and measured values for the lattice constant and bulk modulus results are given for Si in **Table 1**. The results are impressive when one considers that this is an *ab initio* calculation requiring only the atomic number and crystal structure as input. Generally, the pseudpotential–density functional method yields lattice constants and bulk moduli to an accuracy of about 1% and 5%, respectively. The cohesive energy can be calculated by comparing the total energy at the equilibrium lattice constant with the energy of the isolated atoms. Spin polarization effects (von Barth and Hedin 1972; Gunnarsson *et al.*, 1974) and zero point vibrational energy need to be considered. Early results computed within the local density approximation gave reasonable estimates of the cohesive energies; however, in general the generalized gradient approximation yields more accurate cohesive energies (Becke, 1992).

*1.01.5.2.2(i) Silicon and germanium phases* The total energy versus volume curves for seven crystal phases of silicon are shown in Figures 33 and 34. It is encouraging that the lowest energy state corresponds to the diamond structure, as this structure is observed experimentally under ambient conditions. (As with any such calculation, there is no guarantee that there is an untried structure with a lower energy state.) Experimental data for the structural properties of silicon and germanium are listed in Table 3.

For small volumes other structural phases are lower in energy than the diamond phase; hence, pressure-induced solid–solid structural phase transitions are predicted. Structural transitions occur between phases when the Gibbs free energy

$$G = U + PV - TS \qquad (70)$$

for the phases in question are equal. $U$ is the internal energy, $P$ is the pressure, and $S$ is the entropy. Although some attempts have been made to consider phase transitions at finite temperature (Sugino and Car, 1995), most phase stability studies are done at zero temperature. In this cases, $G = U + PV$, where the internal energy is given by the total electronic energy and the pressure is given by $P = -dE_T/dV$.

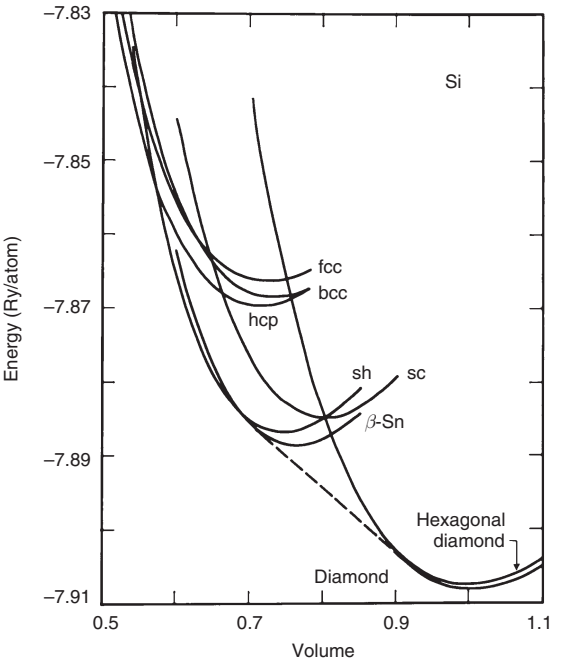


**Figure 33** Total electronic energy for seven phases of crystalline silicon as a function of the atomic volume. The volume has been normalized to the measured value of diamond. The dashed line is a common tangent of the energy curves for the diamond and $\beta$-tin phases. The calculations are from Yin and Cohen (1982c). The total energy reference is the energy per pseudo-Si atom. fcc, face-centered cubic; bcc, body-centered cubic; hcp, hexagonal close packing.

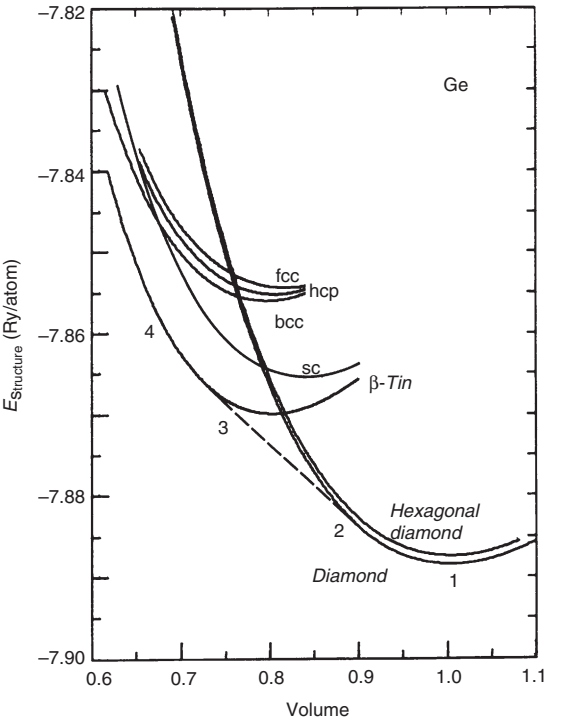


**Figure 34** Total electronic energy for seven phases of crystalline silicon as function of the atomic volume. The volume has been normalized to the measured value of diamond. The dashed line is a common tangent of the energy curves for the diamond and $\beta$-tin phases. The calculations are from Yin and Cohen (1982c). The total energy reference is the energy per pseudo-Ge atom. fcc, face-centered cubic; bcc, body-centered cubic; hcp, hexagonal close packing.

**Table 3** Comparison of calculated and measured static properties of silicon and germanium

| *Semiconductor* | *Lattice constant (Å)* | *Cohesive energy (eV/atom)* | *Bulk modulus (Mbar)* |
|---|---|---|---|
| Si | | | |
| Calculated | 5.45 | 4.84 | 0.98 |
| Measured | 5.43 | 4.63 | 0.99 |
| Ge | | | |
| Calculated | 5.66 | 4.26 | 0.73 |
| Measured | 5.65 | 3.85 | 0.77 |

The calculated values are from Yin and Cohen (1982c). Measured values are from Donohue (1974), Brewer (1977) and McSkimin and Andreatch Jr. (1963).

For $T=0$, the pressure-induced phase transformation can be found by a common tangent line between the $E_T(V)$ curves of the two phases considered. The negative of the slope of the tangent line is the transition pressure. Examples of common tangent constructions are given in **Figures 33** and **34** for the transitions from the diamond structure to the $\beta$-tin phase for Si and Ge. As hydrostatic pressure is applied the path indicated in the figures, $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$, illustrates the change from diamond at 1 to $\beta$-tin at 4, with the transformation occurring along the $2 \rightarrow 3$ segment of the path where both structures can coexist. The predicted phase transition pressure for the diamond $\rightarrow \beta$-tin phase is 99 kbar for Si and 96 kbar for Ge; the measured values are 125 and 100 kbar, respectively (Yin and Cohen, 1982c).

The successful prediction of high-pressure structural phases such as hexagonal forms of Si and Ge and face-centered cubic Si are one of the impressive results of the pseudopotential-density functional theory calculations (Chang and Cohen, 1984; Liu *et al.*, 1988; Vohra *et al.*, 1986). Moreover, these high-pressure phases are metallic and predicted to be superconductors (Chang *et al.*, 1985). Two of the predictions of superconductivity have been verified while the others remain untested.

The calculations for the group IV semiconductors serve as prototypes for the application of the total-energy pseudopotential method to study structural properties of solids.

***1.01.5.2.2(ii) III–V phases*** The first extensions to other solids were made for the III–V semiconductors: GaAs, AlAs, GaP, AlP, AlSb, GaSb, InP, InAs, and InSb (Froyen and Cohen, 1983a, 1983b; Zhang and Cohen, 1987). Total electronic energy calculations for these III–V semiconductors are displayed in **Figures 35** and **36**.

Although these studies for III–V semiconductors generally give good results for the lattice constants and bulk moduli, the agreement between the calculated and measured properties of the high-pressure phases and for phase transition pressures and volumes is not always satisfactory. As an example of the latter, the calculated transition pressures for the Al compounds are consistently lower by around 50% than the measured values (Zhang and Cohen; 1987). In contrast, the calculated lattice constants are within 0.4, 0.3, and 0.3% of the measured values for AlP, AlAs, and AlSb, respectively (Zhang and Cohen, 1987). It is unclear why the transition pressures of the Al compounds are underestimated.

### 1.01.5.2.3 Vibrational properties

An important application of the *ab initio* pseudopotential approach is calculating the vibrational properties of solids. These calculations are based on the change in the energy of atom as it moves from an equilibrium position. Once we have established that the energy of atom can be accurately calculated as a function of position, phonon and lattice vibrational modes can be calculated. The input needed for such calculations is minimal. In contrast to phenomenologic force constant models, which often require as many as 15 parameters to achieve reasonable fits to phonon dispersion curves, the total energy approach uses only the masses and atomic numbers of the constituent atoms. To calculate the phonon dispersion curve, $\omega(\mathbf{q})$, along a specific direction, the frozen-phonon approach is easy to implement (Chadi and Martin, 1976; Heine and Weaire, 1970; Wendel and Martin, 1979; Yin and Cohen, 1980). In this approach, a supercell is chosen, and the atomic cores are displaced to simulate a particular phonon mode. The change in the total energy arising from the distortion depends on the type of distortion assumed and the amplitude. Since a specific phonon wavevector, $\mathbf{q}$, is chosen, this method is limited to calculating $\omega(\mathbf{q})$ at specific points in the Brillouin zone. The phonon frequencies can also be evaluated by computing the Hellman–Feynman forces on the displaced atoms. Another approach that allows the full phonon dispersion to be computed is based on linear response theory (Baroni *et al.*, 2001). However, the frozen-phonon approach is the simplest method for computing $\omega(\mathbf{q})$ at symmetry points. Calculations at nonsymmetry points are possible but require larger unit cells. Standard calculations usually limit the computations to three or four $\mathbf{q}$-points along a symmetry direction.

As an example of a frozen phonon supercell calculation, we briefly examine the calculations of the phonon properties for the $\Gamma$ and $X$ points of the Brillouin zone for Si and Ge. The phonon polarizations for these points can be easily determined using group theory as illustrated in **Figure 37**, and the primitive cell for the distorted lattice contains two atoms for the $LTO(\Gamma)$ mode and four atoms for the phonon modes at $X$. The phonon-distorted lattice has inversion symmetry, which simplifies the calculation.

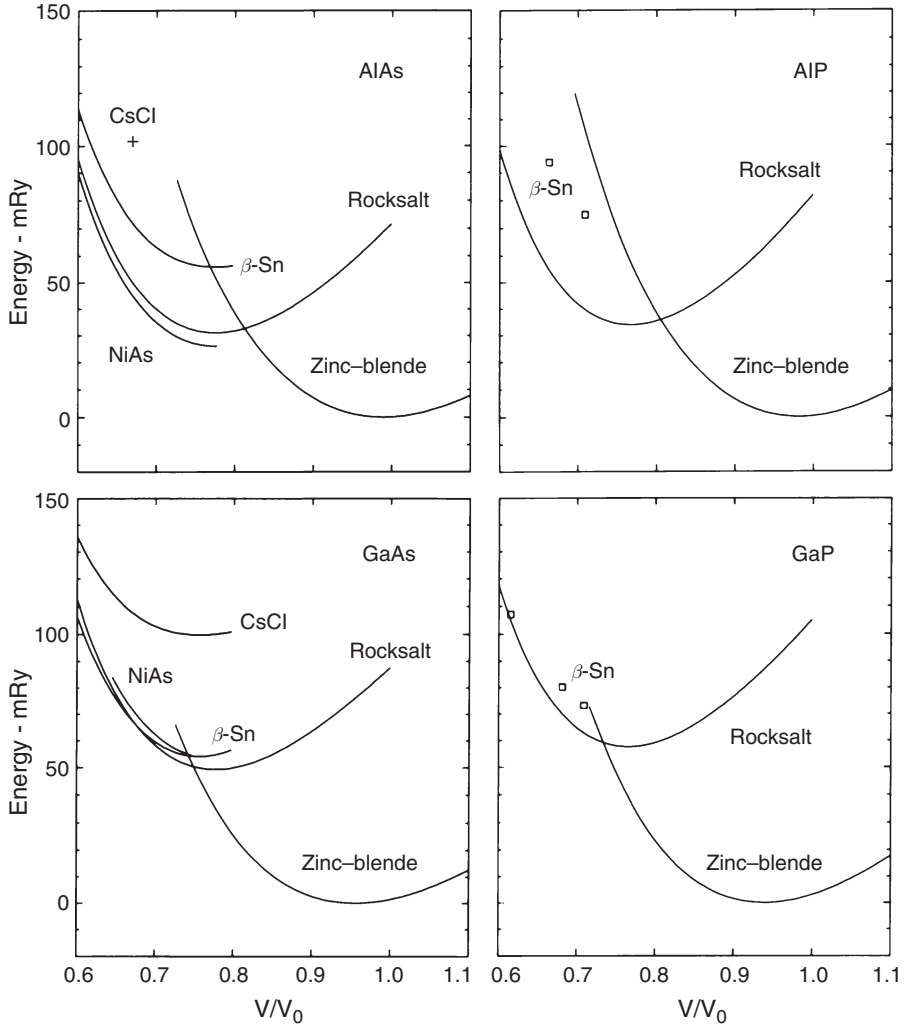


**Figure 35** Total electronic energy for some representative III–V structures as function of the atomic volume. The calculations are from Froyen and Cohen (1983a, 1983b).

For a given amplitude $u$, the change in the total energy arising from the distortion is $\Delta E_T(u)$. The force constant $k$ can be obtained from a second derivative of the $E_T(u)$ or a first derivative of the force $F(u)$:

$$k = \left(\frac{\partial^2 E_T}{\partial u^2}\right)_{u=0} \cong \frac{2\Delta E_T(u)}{u^2} \tag{71}$$

or

$$k = \left(\frac{\partial F}{\partial u}\right)_{u=0} \cong -\frac{F(u)}{u} \tag{72}$$

The phonon frequency is given by

$$\omega = 2\pi f = \sqrt{k/M} \tag{73}$$

where $M$ is the atomic mass. The calculation of the total energies or forces is done typically for five different amplitudes ranging from 0.01 to 0.1 Å. The total energies can fit by a quadratic function to about 1%. For some phonon mode calculations such as the $LTO(\Gamma)$ modes in Si and Ge, higher-order fits such as the third-order term are needed. These fits allow the evaluation of anharmonic terms. **Table 4** illustrates some typical results. An impressive achievement of these calculations is the good agreement with experiment obtained for the $TA(X)$ modes in Si and Ge. For empirical force constant model calculations, many-neighbor parameters are needed to obtain the correct values for these modes because of the important role played by long-ranged forces.

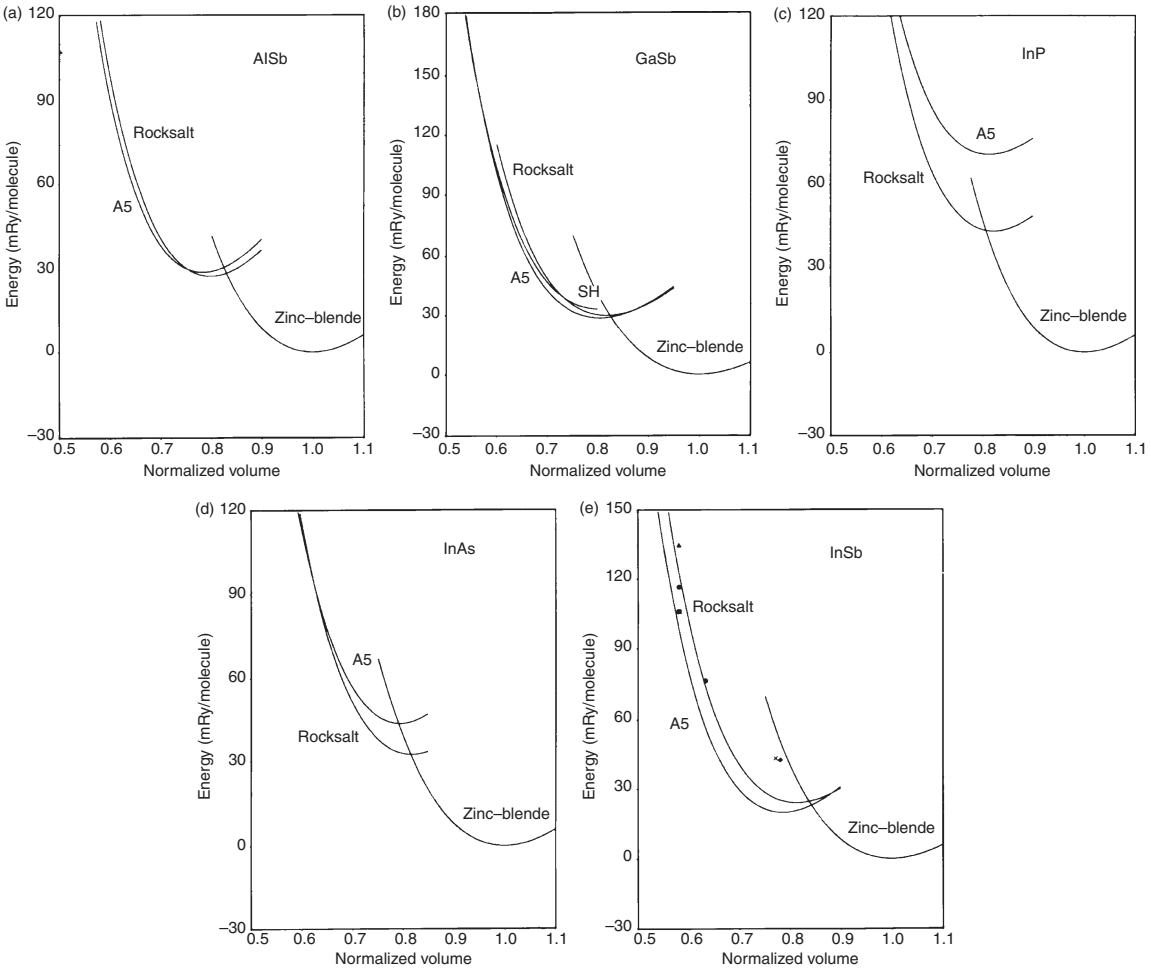


**Figure 36** Total electronic energy for some representative III–V structures as function of the atomic volume. The calculations are from Froyen and Cohen (1983a, 1983b).

The use of the total electronic energy or the forces via the Hellmann–Feynman theorem involves different approaches to calculate the phonon frequencies, but the results are essentially the same as indicated in **Table 4**. Additional information can be obtained from the force calculation. For example, for the $LOA(X)$ mode, the forces on atoms 2 and 3 differ for finite amplitude (see **Figure 37**). This arises from the effects of anharmonicity and illustrates that it is more difficult to compress a bond than to stretch it. For the phonon calculation, the average of the two force constants is used, and this quantity does not vary by more than 0.5% with the amplitude.

If the interlayer force constants are computed, then the full dispersion can be calculated as discussed earlier. This gives the value of the sound velocity. For example, the $q \to 0$ $TA$ velocity $v_{TA}$ [110] propagating along the [110] direction with $[1\bar{1}0]$ polarization is associated with the shear modulus $C_{11} - C_{12}$:

$$v_{TA}[110] = \left(\frac{C_{11} - C_{12}}{2\rho_M}\right)^{1/2} \tag{74}$$

where $\rho_M$ is the mass density. The calculated results for the velocity and shear modulus are in good agreement with experiment as shown in **Table 5**.

When interlayer force constants are calculated, the resulting phonon spectrum is in good agreement with experiment. For example, for Si the $\Gamma$ to $X$ phonon dispersion curve was computed by Yin and Cohen (1982a) using the calculated interlayer

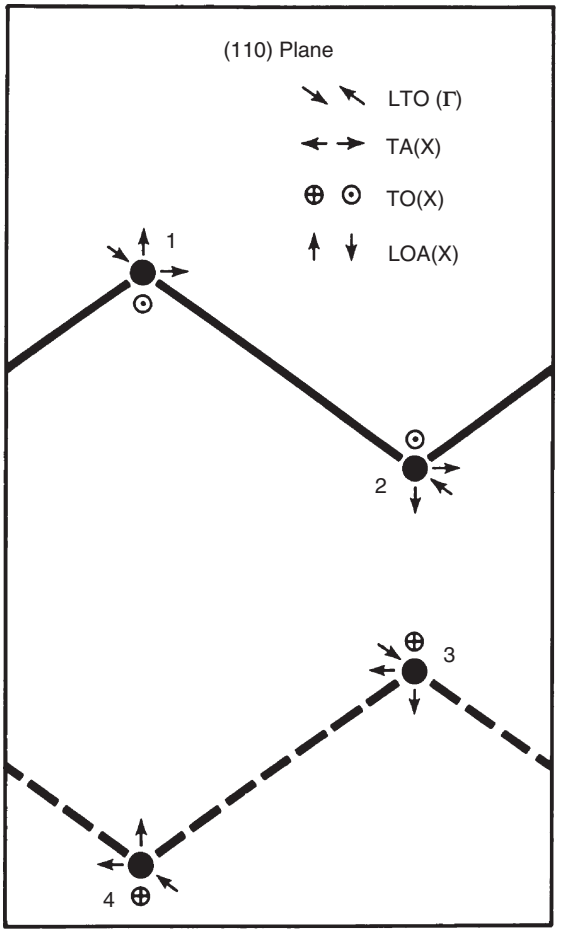


**Figure 37** Phonon polarization at Γ and *X* for the diamond structure. Atoms are numbered and denoted by black dots. The solid lines denote an atomic chain in a (110) plane, and the dashed lines denote the projection of an atomic chain a distance $\sqrt{a/2}$ away from that plane where *a* is the lattice constant (Yin and Cohen, 1982b).

force constants. The results are close to the experimentally determined points as indicated in Figure 38.

The frozen-phonon total-energy approach can also be used to extract nonharmonic contributions to the phonon couplings. By examining the $E_T(u)$ curves terms beyond the quadratic validity, higher order terms can be obtained. For example, Vanderbilt *et al.* (1986) performed a detailed study on third- and fourth-order anharmonic coupling constants for optical, phonons in C, Si, and Ge. These calculations determined bare phonon–phonon scattering amplitudes. By including virtual processes, renormalized multiphonon vertices and phonon self-energies were calculated. One specific application was the study of the renormalized four-phonon vertices. These were found to be negative, which is the wrong sign for allowing the formation of a proposed two-phonon bound state (Cohen and Ruvalds, 1969).

## 1.01.6 Summary and Conclusions

The pseudopotential concept has had a profound impact on our understanding of the electronic structure of semiconductors. In this chapter, both empirical and *ab initio* pseudopotential concepts were outlined and some central applications discussed.

Empirical pseudopotentials provided a means for understanding the optical and dielectric properties of semiconductors and the underlying energy band structures. It is sometimes stated that the

**Table 4** Comparison of the calculated phonon frequencies, *f* (in THz), of Si and Ge (Yin and Cohen, 1982b) at Γ and *X* with experiment (Dolling, 1963; Nilsson and Nelin, 1971, 1972)

| | *LTO(Γ)* | *LOA(X)* | *TO(X)* | *TA(X)* |
|---|---|---|---|---|
| Si | | | | |
| *f* (energy) | 15.16 (−2%) | 12.16 (−1%) | 13.48 (−3%) | 4.45 (−1%) |
| *f* (force) | 15.14 (−3%) | 11.98 (−3%) | 13.51 (−3%) | 4.37 (−3%) |
| *f* (expt) | 15.33 | 12.32 | 13.90 | 4.49 |
| Ge | | | | |
| *f* (energy) | 8.90 (−2%) | 7.01 (−3%) | 7.75 (−6%) | 2.44 (2%) |
| *f* (force) | 8.89 (−3%) | 6.96 (−3%) | 7.78 (−6%) | 2.45 (2%) |
| *f* (expt) | 9.12 | 7.21 | 8.26 | 2.40 |

The values of the frequencies were determined by energy and by force calculations. The deviation from experimental values are given in parentheses.

**Table 5** Comparison of calculated values (Yin and Cohen, 1982b) of $C_{11}$–$C_{12}$ and the *TA* velocity along the [110] (with [1$\bar{1}$0] polarization) for Si and Ge with experiment (McSkimin and Andreatch Jr., 1963).

| | | $C_{11}-C_{12}$ (Mbar) | $v_{TA}$ [110] ($10^5$ cm s$^{-1}$) |
|---|---|---|---|
| Si | | | |
| | Theory | 1.07 | 4.79 |
| | Expt. | 1.03 | 4.69 |
| Ge | | | |
| | Theory | 0.74 | 2.64 |
| | Expt. | 0.82 | 2.77 |

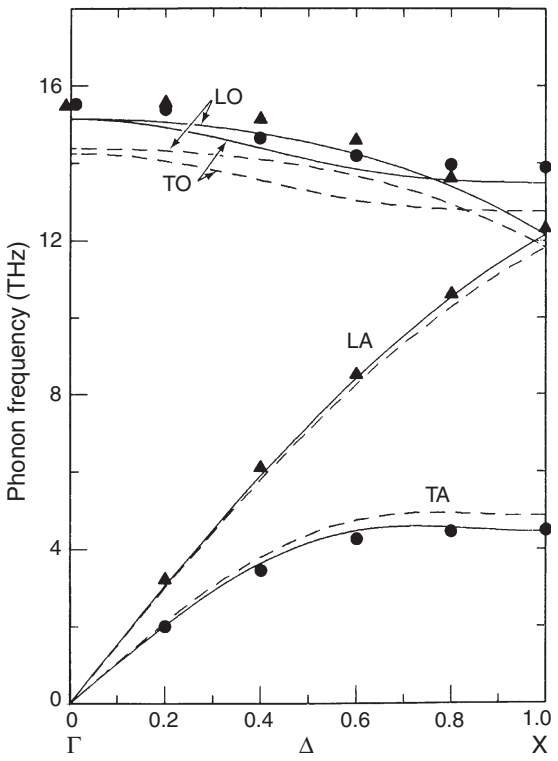


**Figure 38** Phonon branches from Γ to *X* along the Δ direction from Yin and Cohen (1982a). The dashed lines are calculated from the computed force constants. The solid lines are calculated using the computed third-nearest layer force constants and the frozen-phonon results at Γ and *X*. Experimental points are from Dolling (1963), Nilsson and Nelin (1972), and Sinha (1973) and denoted by dots for the transverse modes and triangles for the longitudinal modes.

empirical pseudopotential method (EPM) established the validity of the energy band concept for solids in general. There is much truth to this statement. In the 1950s, there was extended discussions about the validity of the one-electron picture and whether it could be applied to solid and specifically semiconductors such as silicon. Some workers thought that the electron–hole interactions for a optical excitation would be so strong as to obscure any attempt at understanding such excitations using energy band theory. Given that the EPM allowed adjustments to the one-electron potential, if the EPM had failed to work with the flexibility of choosing a potential, one would have questioned whether it was possible to construct a meaningful one-electron potential and corresponding energy band structure. The consequences of this failure would have been very damaging to prospects for using energy bands to interpret optical, dielectric, photoemission, and transport properties of semiconductors.

Fortunately, this was not the case; the EPM provided a simple and effective way to couple band structure theory to experimental work and its impact was immediate. While much of this effort took place in the 1960s and 1970s, the EPM still serves as an effective means to interpret optical data for semiconductors. While more sophisticated methods now exist to examine optical properties, few of the conclusions of the EPM have been overturned by these methods.

The coupling of density functional theory with pseudopotentials in the early 1980s resulted in another impressive advance for understanding the electronic structure of materials. The accuracy of density functional theory for predicting electronic and structural energies was problematic at that time. However, the direct numerical application of structural energies to problems involving bond lengths, compressibilities, phonon modes, and phase stabilities showed conclusively the applicability of pseudopotential methods to these problems. Over the last 10 years, more than 10 000 papers have appeared with pseudopotentials in the title or abstract and these papers have been cited over 100 000 times.

The vitality and future of the pseudopotential method is without question. The method is now used to examine new and exciting materials systems. These systems include nanoscale systems, amorphous solids, glasses, and liquids. Moreover, owing to strong advances in both hardware and software (algorithms), it is now possible to address systems with thousands of atoms on an almost routine basis. It is for this reason that the pseudopotential concept is said to be the standard model for condensed matter. (See Chapter 1.02).

## References


Adler SL (1962) Quantum theory of the dielectric constant in real solids. *Physical Review* 126: 413.

Animalu AOE and Heine V (1965) The screened model for 25 elements. Phil. Mag. 12, 1249.

Aspnes DE (1973) Interband masses of higher interband critical points in Ge. *Physical Review Letters* 31: 230.

Aspnes DE (1976) GaAs lower conduction-band minima: Ordering and properties. *Physical Review B* 14: 5331.

Aspnes DE and Studna AA (1972) Direct observation of $E_0$ and $E_0 + \Delta_0$ transitions in silicon. *Solid State Communications* 11: 1375.

Aspnes DE and Studna AA (1973) Schonky-barrier electroreflectance: Application to GaAs. *Physical Review B* 7: 4605.

Bachelet G, Hamann DR, and Schlüter M (1982) Pseudopotentials that work: From hydrogen to plutonium. *Physical Review B* 26: 4199.

Barom S, de Gironcoli S, Dal Corso A, and Giannozzi P (2001) Phonons and related crystal properties from density-functional perturbation theory. *Reviews of Modern Physics* 73(2): 515.

Becke AD (1992) Density-functional thermochemistry 1. The effect of the exchange-only gradient correction. *Journal of Chemical Physics* 96: 2155.

Birch F (1952) Elasticity and constitution of the earth's interior. *Journal of Geophysical Research* 57: 227.

Brewer L (1977) Tech. Rep. LBL-3720. Berkeley, CA: Lawrence Berkeley Laboratory.

Broyden CG (1965) A class of methods for solving nonlinear simultaneous equations. *Mathematics of Computation* 19: 577. http://www.jstor.org/stable/2003941 (accessed October 2009).

Car R and Parrinello M (1985) Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters* 55: 2471.

Cavell RG, Kowalczyk SP, Ley L, *et al.* (1973) X-ray photoemission cross-section modulation in diamond, silicon, germanium, methane, silane, and germane. *Physical Review B* 7: 5313.

Ceperley DM and Alder BJ (1980) Ground state of the electron gas by a stochastic method. *Physical Review Letters* 45: 566.

Chadi DJ and Cohen ML (1973) Special points in the Brillouin zone. *Physical Review B* 8: 5747.

Chadi DJ and Martin RM (1976) Calculation of lattice dynamical properties from electronic energies: Application to carbon, silicon and germanium. *Solid State Communications* 19: 643.

Chan CTD, Vanderbilt D, and Louie SG (1986) Application of a general self-consistency scheme in the linear combination of atomic orbitals formalism to the electronic and structural properties of silicon and tungsten. *Physical Review B* 33: 2455.

Chang KJ and Cohen ML (1984) Structural and electronic properties of the high-pressure hexagonal phases of silicon. *Physical Review B* 30: 5376.

Chang KJ, Dacorogna MM, Cohen ML, Mignot JM, Chouteau G, and Martinez G (1985) Superconductivity in high-pressure metallic phases of Si. *Physical Review Letters* 54: 2375.

Chelikowsky JR (2000) The origin of the pseudopotential density functional method. *Theoretical Chemistry Accounts* 103: 340.

Chelikowsky JR and Cohen ML (1973) High-resolution band structure and the $E_2$ peak in Ge. *Physical Review Letters* 31: 1582.

Chelikowsky JR and Cohen ML (1976) Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors of semiconductors. *Physical Review B* 14: 556.

Chelikowsky JR and Cohen ML (1992) *Ab initio* pseudopotentials for semiconductors. In: Moss TS and Landsberg PT (eds.) *Handbook of Semiconductors*, 2nd edn., p. 59. Amsterdam: Elsevier.

Chelikowsky JR, Troullier N, and Saad Y (1994) The finite-difference-pseudopotential method: Electronic structure calculations without a basis. *Physical Review Letters* 72: 1240.

Chelikowsky JR, Wagener TJ, Weaver JH, and Jin A (1989) Valence- and conduction-band densities of states for tetrahedral semiconductors: Theory and experiment. *Physical Review B* 40: 9644.

Cohen M and Chelikowsky J (1982) Pseudopotentials for semiconductors. In: Paul W (ed.) *Handbook on Semiconductors,* vol. 1, p 219. Amsterdam: North Holland.

Cohen MH and Ruvalds J (1969) Two-phonon bound states. *Physical Review Letters* 23: 1378.

Cohen ML and Bergstresser TK (1965) Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures. *Physical Review* 141: 789.

Cohen ML and Chelikowsky JR (1989) *Electronic Structure and Optical Properties of Semiconductors,* 2nd edn.; Berlin: Springer.

Cohen ML and Heine V (1970) The fitting of pseudopotentials to experimental data and their subsequent application. In: Ehrenreich H, Seitz F, and Turnbull D (eds.) *Solid State Physics,* vol. 24, p 37. New York: Academic Press.

Dirac PAM (1929) Quantum mechanics of many electron systems. *Proceedings of the Royal Society of London* A 123: 714.

Dolling G (1963) *Inelastic Scattering of Neutrons in Solids and Liquids*, Vol I and II, Vienna: IAEA.

Donohue J (1974) *The Structure of the Elements*. New York: Wiley.

Ehrenreich H and Cohen MH (1959) Self-consistent field approach to the many-electron problem. *Physical Review* 115: 786.

Fermi E (1934) Sullo spostamento per pressione dei termini elevati delle serie spettrali (on the pressure displacement of higher terms in spectral series). *Nuovo Cimento* 11: 157.

Freeouf JL (1973) Far-ultraviolet reflectance of II-VI compounds and correlation with the Penn-Phillips gap. *Physical Review B* 78: 3810.

Froyen S and Cohen ML (1983a) Static and structural properties of III-V zinc blende semiconductors. In: *Proceedings of the 16th International Conference on the Physics of Semiconductors*. Part I, p. 561. Amsterdam: North-Holland.

Froyen S and Cohen ML (1983b) Structural properties of III-V zinc-blende semiconductors under pressure. *Physical Review B* 28: 3258.

Gobeli GW and Kane EO (1965) Dependence of the optical constants of silicon on uniaxial stress. *Physical Review Letters* 15(4): 142.

Grover JW and Handler P (1974) Electroreflectance of silicon. *Physical Review B* 9(6): 2600.

Gunn JB (1964) Instabilities of current in III–V semiconductors. *IBM Journal of Research and Development* 8: 151.

Gunnarsson O, Lundqvist BI, and Wilkins JW (1974) Cohesive energy of simple metals. Spin-dependent effect. *Physical Review B* 10: 1319.

Hamann DR (1989) Generalized norm-conserving pseudopotentials. *Physical Review B* 49: 2980.

Hamann DR, Schlüter M, and Chiang C (1979) Norm-conserving pseudopotentials. *Physical Review Letters* 43: 1494.

Hanke W and Sham LJ (1974) Dielectric response in the wannier representation. Application to the optical spectrum of diamond. *Physical Review Letters* 33(10): 582.

Heine V and Weaire D (1970) Pseudopotential theory of cohesion and structure. *Solid State Physics* 24: 249.
Hellman H (1935) A new approximation method in the problem of many electrons. *Journal of Chemical Physics* 3: 61.
Herring C (1940) A new method for calculating wave functions in crystals. *Physical Review* 57: 1169.
Hohenberg P and Kohn W (1964) Inhomogeneous electron gas. *Physical Review* 136: B864.
Hybertsen MS and Louie SG (1986) Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Physical Review B* 34: 5390.
Ihm J and Cohen ML (1980) Calculation of structurally related properties of bulk and surface Si. *Physical Review B* 21: 1527.
Ihm J, Zunger A, and Cohen ML (1979) Momentum-space formalism for the total energy of solid. *Journal of Physics C* 12: 4409.
Kerker GP (1980) Nonsingular atomic pseudopotentials for solid state applications. *Journal of Physics C* 13: L189.
Kittel C (2005) *Introduction to Solid State Physics* 8th edn., New York: Wiley.
Kleinman L and Bylander DM (1982) Efficacious form for model pseudopotentials. *Physical Review Letters* 48: 1425.
Kleinman L and Phillips JC (1960a) Crystal potential and energy bands of semiconductors. II. Self-consistent calculations for cubic boron nitride. *Physical Review* 117: 460.
Kleinman L and Phillips JC (1960b) Crystal potential and energy bands of semiconductors. III. Self-consistent calculations for silicon. *Physical Review* 118: 1153.
Kline JS, Pollak FH, and Cardona M (1968) Electroreflectance in Ge-Si alloys. *Helvetica Physica Acta* 41: 968.
Kohn W and Sham LJ (1965) Self-consistent equations including exchange and correlation effects. *Physical Review* 140: A1133.
Koo J, Shen YR, and Zucca RRL (1971) Effects of uniaxial stress on $E'_0$-peak of silicon. *Solid State Communications* 9: 2229.
Kronik L, Makmal A, Tiago ML, *et al.* (2006) Parsec – the pseudopotential algorithm for real-space electronic structure calculations: Recent advances and novel applications to nano-structures. *Physica Status Solidi (b)* 243: 1063.
Kunz AB (1971) Energy bands and soft X-rays absorption in Si. *Physical Review Letters* 27: 567.
Ley L, Kowalczyk S, Pollak R, and Shirley DA (1972) X-ray photoemission spectra of crystalline and amorphous Si and Ge valence bands. *Physical Review Letters* 29: 1088.
Liu AY, Chang KJ, and Cohen ML (1988) Theory of electronic, vibrational, and superconducting properties of fcc silicon. *Physical Review B* 37: 6344.
McSkimin HJ and Andreatch P, Jr. (1963) Elastic moduli of germanium versus hydrostatic pressure at 25.0 and $-195.8^{\circ}$. *Journal of Applied Physics* 34: 651.
Murnaghan FD (1944) The compressibility of media under extreme pressures. *Proceedings of the National Academy of Sciences of the United States of America* 30: 244.
Nilsson G and Nelin G (1971) Phonon dispersion relations in germanium at 80 K. *Physical Review B* 3: 364.
Nilsson G and Nelin G (1972) Study of the homology between silicon and germanium by thermal-neutron spectrometry. *Physical Review B* 6: 3777.
Pauling L (1960) *The Nature of the Chemical Bond*. Ithaca, NY: Cornell University Press.
Philipp HR and Ehrenreich H (1963) Optical properties of semiconductors. *Physical Review* 129: 1550.
Phillips JC (1973) *Bonds and Bands in Semiconductors*. New York: Academic Press.
Phillips JC and Kleinman L (1959) New method for calculating wave functions in crystals and molecules. *Physical Review* 116: 287.
Phillips JC and Kleinman L (1962) Crystal potential and energy bands of semiconductors. IV. Exchange and correlation. *Physical Review* 128: 2098.
Phillips JC and Pandey KC (1973) Nonlocal pseudopotential for Ge. *Physical Review Letters* 30: 787.
Pollak FH and Cardona M (1968) Piezo-electroreflectance in Ge. GaAs, and Si. *Physical Review* 172(3): 816.
Rehn V and Kyser D (1972) Symmetry of the 4.5 eV optical interband threshold in gallium arsenide. *Physical Review Letters* 28: 494.
Rohlfing M and Louie S (2000) Electron-hole excitations and optical spectra from first principles. *Physical Review B* 62: 4927.
Rohlfing M and Louie SG (1998) Electron-hole excitations in semiconductors and insulators. *Physical Review Letters* 81: 2312.
Schwarz W, Andraea D, Arnold S, *et al.* (1999a) Hans G.A. Hellmann (1903–1938): Part I. A pioneer of quantum chemistry. *Bunsen - Magazin* 1: 10.
Schwarz W, Andraea D, Arnold S, *et al.* (1999b) Hans G.A. Hellmann (1903–1938): Part II. A German pioneer of quantum chemistry in Moscow. *Bunsen - Magazin* 2: 60.
Sinha SK (1973) Phonons in semiconductors. *CRC Critical Reviews in Solid State and Materials Sciences* 3: 273.
Sugino O and Car R (1995) *Ab-initio* molecular-dynamics study of first-order phase-transitions – melting of silicon. *Physical Review Letters* 74: 1823.
Tauc J and Abraham A (1961) Optical investigation of the band structure of Ge-Si alloys. *Journal of Physics and Chemistry of Solids* 20: 190.
Theis D (1977) Wavelength-modulated reflectivity spectra of ZnSe and ZnS from 2.5 to 8 eV. *Physica Status Solidi (b)* 79: 125.
Troullier N and Martins J (1991) Efficient pseudopotentials for planewave calculations. *Physical Review B* 43: 1993.
Vanderbilt D (1985) Optimally smooth norm-conserving pseudopotentials. *Physical Review B* 32: 8412.
Vanderbilt D (1990) Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B* 41: 7892.
Vanderbilt D, Louie SG, and Cohen ML (1986) Calculation of an harmonic phonon couplings in carbon, silicon, and germanium. *Physical Review B* 33: 8740.
Vohra YK, Brister KE, Desgreniers S, Ruoff AL, Chang KJ, and Cohen ML (1986) Phase-transition studies of germanium to 1.25 Mbar. *Physical Review Letters* 56: 19444.
von Barth U and Hedin L (1972) Local exchange-correlation potential for the spin-polarized case. I. *Journal of Physics C* 5: 1629.
Walter JP, Cohen ML, Petroff Y, and Balkanski M (1970) Calculated and measured reflectivity of zinc telluride and zinc selenide. *Physical Review B* 1: 2661.
Welkowsky M and Braunstein R (1972) Interband transitions and exciton effects in semiconductors. *Physical Review B* 5: 497.
Wendel H and Martin RM (1979) Theory of structural properties of covalent semiconductors. *Physical Review B* 19: 5251.
Wiser N (1963) Dielectric constant with local field effects included. *Physical Review* 129: 62.
Yin MT and Cohen ML (1980) Microscopic theory of the phase transformation and lattice dynamics of Si. *Physical Review Letters* 45: 1004.
Yin MT and Cohen ML (1982a) *Ab initio* calculation of the phonon dispersion relation: Application to silicon. *Physical Review B* 25: 4317.
Yin MT and Cohen ML (1982b) Theory of lattice-dynamical properties of solids: Application to Si and Ge. *Physical Review B* 26: 3259.
Yin MT and Cohen ML (1982c) Theory of static structural properties, crystal stability, and phase transformations. Application to Si and Ge. *Physical Review B* 26: 5668.
Zhang SB and Cohen ML (1987) High-pressure phases of III–V zinc-blende semiconductors. *Physical Review B* 35: 7604.
Zucca RRL and Shen YR (1970) Wavelength-modulation spectra of some semiconductors. *Physical Review B* 1: 2668.

## Further Reading

Cohen ML and Chelikowsky JR (1989) *Electronic Structure and Optical Properties of Semiconductors,* 2nd edn. Berlin: Springer.

Karixas E (2003) *Atomic and Electronic Structure of Solids*. Cambridge: Cambridge University Press.

Kittel C (2004) *Introduction to Solid State Physics,* 8th edn. New York: Wiley.

Marder MP (2000) *Condensed Matter Physics*. New York: Wiley.

Martin RM (2004) *Electronic Structure: Basic Theory and Practical Method*. Cambridge: Cambridge University Press.