

# Momentum-space formalism for the total energy of solids

J Ihm, Alex Zunger and Marvin L Cohen†

Department of Physics, University of California and Materials and Molecular Research Division, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA

Received 8 May 1979

**Abstract.** A momentum-space formalism for calculating the total energy of solids is derived. This formalism is designed particularly for application with the self-consistent pseudopotential method. In the present formalism, the total energy is obtained through band-structure calculations without involving additional integrations. The Hellman–Feynman theorem is derived, as is a modified virial relation for the pseudopotential Hamiltonian which provides an alternative way of calculating forces and total energies.

## 1. Introduction

The calculation of the total energy of solids and related derivatives with respect to structural degrees of freedom has been an ongoing problem since the early days of solid state physics (Wigner and Seitz 1933, 1934, Fuchs 1935). Quantum-mechanical calculations on molecules suggest that correlation effects might sometimes be responsible for most of their binding energy (Schaefer 1972). The solid state approaches have concentrated on efforts to include most of these effects through an effective potential  $V_{\text{corr}}[\rho(r, r')]$  (Hohenberg and Kohn 1964, Kohn and Sham 1965), rather than by complicated wave-function-related configuration interactions or many-electron perturbation techniques. Besides the problem of considering correlation effects, the self-consistent solution of the Schrödinger equation within a desired accuracy is quite difficult; typically, the experimental binding energy of elemental solids is  $10^{-4}$ – $10^{-5}$  times the total energy. These difficulties have inspired a large set of total-energy calculations that circumvents the complete solution of the Schrödinger equation (Harrison 1966, Heine and Weaire 1970). It is based on various approximations to the nearly-free-electron representation and may include the effect of more localised electrons (e.g. d states in transition metals) through specific interaction models (Moriarty 1974, 1977). As the variational self-consistent charge density remains unspecified in this approach, various forms of linear dielectric screening of the basic Coulomb interactions are introduced (Harrison 1966, Heine and Weaire 1970).

Another set of total-energy calculations is based on direct solutions of the Schrödinger equation within a given interaction model: for example, Hartree–Fock (Harris and Monkhorst 1971, Wepfer *et al* 1974), density-functional (Ching and Callaway 1974,

† Guggenheim Fellow 1978–9.

Zunger and Freeman 1977a, b, c, 1978) or model potentials (Wendel and Martin 1978). Unlike the situation in molecular physics, we have in this case to deal with divergent terms. Moreover, the calculation of the total energy poses a practical difficulty associated with the need to compute a large number of six-dimensional integrals for the electron-electron interaction term. In linear band structure approaches (e.g. linear combination of atomic orbitals), this leads to a large number of multicentre integrals (Schaefer 1972, Harris and Monkhorst 1971, Wepfer *et al* 1974).

This major difficulty has led to a number of 'shape approximations' in which the charge density  $\rho(\mathbf{r})$  is replaced by a radially scalar quantity such as the muffin-tin (DeCicco 1965, Averill 1972, Snow 1973, Janak 1974) or cellular approaches (Wigner and Seitz 1933, 1934; Fuchs 1935). As the angular parts are readily treated, the calculation reduces to essentially one-dimensional integrals.

Total-energy calculations within the muffin-tin approximation have been reported by Averill (DeCicco 1965, Averill 1972, Janak 1974) on alkali metals and rare-gas solids, by Snow (1973) on copper, by Sabin *et al* (1975) on Ne, by Janak *et al* (Janak *et al* 1975, Janak and Williams 1976) on Li, Be, Na, Al, Ar, K, Ca and Cu and recently on a series of transition metals. These calculations have reproduced both the cohesive energies and the equilibrium zero-pressure lattice constants and have revealed interesting regularities in their properties. This approach is, however, usually specialised to the study of close-packed materials for which the muffin-tin approximation to the charge density seems valid. Extensions of the muffin-tin approximation to open structures like molecules, tetrahedrally bonded insulators and semiconductors have revealed serious errors both in the one-electron energies and in the cohesive energy. Generalisation of this method to include non-muffin-tin corrections seems to be extremely complicated at present (Danese 1974). Usually the calculation of the binding energy is complicated by the need to treat accurately core energies which constitute a large part of the total energy (e.g. 98% in Fe) but introduce a negligible lattice-constant dependence other than to screen the nuclear charge. If the muffin-tin approximation is assumed, it is possible to use a direct algebraic cancellation (DeCicco 1965, Averill 1972, Janak 1974) of these large energies by subtracting the atomic core energies. However, when the charge density is not approximated, all-electron calculations become very complicated.

Despite these difficulties, self-consistent all-electron calculations of the total energy without muffin-tin approximations have been performed on a number of systems. These include the Hartree-Fock calculations on metallic hydrogen, diamond, LiF and BN (Harris and Monkhorst 1971, Wepfer *et al* 1974) and the density functional calculations on diamond (Goroff and Kleinman 1970) and on LiF, BN, TiS<sub>2</sub> and diamond (Zunger and Freeman 1977a, b, c, 1978). Although excellent results have been obtained for these cases, calculations become extremely time consuming when extended to systems containing other than first-row elements.

In the present paper, we follow a different approach to the calculation of the total energy of solids. This involves a direct calculation of the total energy from the variationally determined valence-electron eigenvalues and charge densities and does not use any shape approximation to  $\rho(\mathbf{r})$ . The large and geometry-insensitive core contributions are explicitly projected out of the total-energy expression by using a pseudopotential formalism. The local density formalism for the exchange-correlation potential is self-consistently employed in the calculation. Most importantly, a drastic simplification of the total-energy expression is made by formulating the relevant expression for the total energy in momentum space. This expression is particularly designed to be applied with the pseudopotential method and a plane-wave basis set. The present formalism, however,

is readily applicable to calculations with mixed basis sets (e.g. plane wave plus Gaussian) as well. This enables us to extend our calculations to the case of transition metals.

The virtue of the present method rests in its computational simplicity; once we perform the band structure calculation, the total energy is automatically obtained as a sum of a few previously calculated terms without multicentre integrations. The only input is the core pseudopotential, and no other adjustable parameters are necessary. Application of the present method to the study of bulk and surface Si (Ihm and Cohen 1979) and the transition metals Mo and W (Zunger and Cohen 1979) will appear in separate papers. In addition to the momentum-space expression for the total energy, we derive the Hellmann–Feynman theorem and the virial relation for the pseudopotential Hamiltonian to provide an alternative way of calculating forces and total energies.

In §2, the momentum-space expression for the total energy in the pseudopotential formalism is derived. The Hellmann–Feynman theorem and the modified virial theorem for the pseudopotential Hamiltonian are proved in §3 and 4 respectively.

## 2. Momentum-space formalism for total-energy calculations

Following the conventional density functional formalism (Hohenberg and Kohn 1964, Kohn and Sham 1965) in a pseudopotential framework (Phillips and Kleinman 1959), the total crystal energy (defined as the total energy difference between the solids and isolated cores) is given by

$$E_{\text{total}} = T + V + \int E_{\text{xc}}(r) d^3r, \quad (1)$$

where the total kinetic energy,  $T$  is

$$T = \sum_i \psi_i^*(r) (-\nabla^2) \psi_i(r) d^3r, \quad (2)$$

and the electrostatic potential energy,  $V$  is

$$V = \sum_{i, \mu, l} \int \psi_i^*(r) U_{\text{ps}, l}(r - R_\mu) \hat{P}_l \psi_i(r) d^3r \\ + \frac{1}{2} \iint \frac{2\rho(r)\rho(r')}{|r - r'|} d^3r d^3r' + \frac{1}{2} \sum_{\substack{\mu, v \\ \mu \neq v}} \frac{2Z^2}{|R_\mu - R_v|}. \quad (3)$$

Rydberg units are used throughout the paper.  $\psi_i(r)$  is the (pseudo)wavefunction of the valence electron where the index  $i$  denotes both the wavevector  $k_i$  and the band index  $n$  and runs over all occupied valence states.  $\rho(r) \equiv \sum_i \psi_i^*(r)\psi_i(r)$  is the (pseudo)valence-electron density,  $Z$  is the valence of the ion and the  $R_\mu$  is the lattice vector. Although we confine ourselves here to one kind of ion for notational convenience, generalisation to many kinds of ions is straightforward. The first term in equation (3) is the core–valence interaction energy for angular-momentum-dependent pseudopotentials ( $\sum_l U_{\text{ps}, l}(r)\hat{P}_l$ , where  $\hat{P}_l$  is the projection operator on angular momentum  $l$ ). The second term is the (valence) electron–electron Coulomb energy, and the last term is the lattice (ion–ion) energy. The ion–ion interactions can be replaced rigorously by point-ion interactions as long as the ions are spherically symmetric and non-overlapping. The last term in

equation (1) is the density functional exchange–correlation contribution to the total energy (Hohenberg and Kohn 1964, Kohn and Sham 1965).

The corresponding one-electron Schrödinger equation derived variationally from equation (1) is

$$\left( -\nabla^2 + \sum_{\mu,l} U_{ps,l}(\mathbf{r} - \mathbf{R}_\mu) \hat{P}_l + \int \frac{2\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' + \mu_{xc}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (4)$$

where

$$\mu_{xc}(\mathbf{r}) \equiv \partial E_{xc}(\mathbf{r}) / \partial \rho(\mathbf{r}).$$

The general form for  $E_{xc}$  is not known. Employing the  $X_\alpha$  method (Slater 1974), we obtain

$$\mu_{xc}(\mathbf{r}) = -\alpha(3/\pi) (3\pi^2)^{1/3} (\rho(\mathbf{r}))^{1/3}, \quad (5)$$

and

$$E_{xc}(\mathbf{r}) = \int \mu_{xc}(\mathbf{r}) d\rho(\mathbf{r}) = \frac{3}{4} \rho(\mathbf{r}) \mu_{xc}(\mathbf{r}). \quad (6)$$

For explicitness, equations (5) and (6) will be used to calculate the exchange–correlation in the present paper. It should be straightforward, however, to modify the expression in compliance with particular approximations employed for the exchange–correlation (Zunger and Cohen 1979).

It is to be noted that the pseudopotential method based on equation (4) has been very successful in calculating the band structure of solids and is at present a well-established technique. However, the total (crystal)-energy equation (equation 1) has not been tested as much because of the lack of a practical scheme of evaluating various integrals appearing in the equation. Moreover, it is not obvious whether the pseudo-valence electrons alone can give a reasonable result for the cohesive energy because calculations using other methods show that extreme accuracy is required in the calculation of the cohesive energy. The separation between valence and core electrons is a first-order approximation of the non-linear exchange–correlation function. Indeed, a total-energy calculation provides a far more stringent test of the pseudopotential method than a band-structure calculation does. It suffices here to refer to the successful applications of the present formalism for bulk and surface Si (Ihm and Cohen 1979) and bulk Mo and W (Zunger and Cohen 1979).

Thus far, we have not made use of the periodicity of the system. To simplify equation (1), each quantity will be expanded in terms of plane waves. In contrast to the all-electron potential, the smooth pseudopotential may permit rapid convergence of the plane-wave expansion. We can also cope with more localised features by generalising the present method to the mixed-basis (plane waves plus localised orbitals) formalism. The ensuing modifications in the mixed-basis set formalism are briefly outlined in the Appendix. Let the momentum-space representations of the wavefunction, the charge density, the inter-electronic Coulomb potential and the exchange–correlation potential be denoted by  $\psi(\mathbf{k}_i + \mathbf{G})$ ,  $\rho(\mathbf{G})$ ,  $V_{\text{Coul}}(\mathbf{G})$  and  $\mu_{xc}(\mathbf{G})$ , respectively, where the  $\mathbf{G}$  are reciprocal lattice vectors. The Fourier component of the interelectronic Coulomb potential,  $V_{\text{Coul}}(\mathbf{G})$  is

$$V_{\text{Coul}}(\mathbf{G}) = 8\pi\rho(\mathbf{G})/G^2, \quad (7)$$

from the Poisson equation. Let  $\Omega$  denote the total volume of the system and  $N$  denote the total number of atoms; then  $\Omega_{\text{at}} = \Omega/N$ . Using this Fourier representation, the

Coulomb repulsion energy becomes:

$$\frac{1}{2} \iint \frac{2\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r d^3r' = \frac{1}{2} \Omega \sum_{\mathbf{G}} V_{\text{Coul}}(\mathbf{G}) \rho(\mathbf{G}). \quad (8)$$

Using the translational invariance and the decomposition of the plane waves into spherical harmonics and Bessel functions, the pseudopotential energy can be written as:

$$\begin{aligned} \sum_{i,\mu,l} \int \psi_i^*(\mathbf{r}) U_{\text{ps},l}(\mathbf{r}-\mathbf{R}_\mu) \hat{P}_l \psi_i(\mathbf{r}) d^3r \\ = \Omega \sum_{i,l,\mathbf{G},\mathbf{G}'} \psi^*(\mathbf{k}_i+\mathbf{G}) \psi(\mathbf{k}_i+\mathbf{G}') \sum_{\mu} \frac{\exp[i(\mathbf{G}'-\mathbf{G})\cdot\mathbf{R}_\mu]}{N} \\ \times (1/\Omega_{\text{at}}) \int \exp[-i(\mathbf{k}_i+\mathbf{G})\cdot\mathbf{r}] U_{\text{ps},l}(\mathbf{r}) \hat{P}_l \exp[i(\mathbf{k}_i+\mathbf{G}')\cdot\mathbf{r}] d^3r \\ = \Omega \sum_{i,l,\mathbf{G},\mathbf{G}'} \psi^*(\mathbf{k}_i+\mathbf{G}) \psi(\mathbf{k}_i+\mathbf{G}') S(\mathbf{G}'-\mathbf{G}) U_{\text{ps},l,\mathbf{k}_i+\mathbf{G},\mathbf{k}_i+\mathbf{G}'}, \end{aligned} \quad (9)$$

where  $S(\mathbf{G}'-\mathbf{G})$  is the structure factor and the generalised non-local form factor is:

$$\begin{aligned} U_{\text{ps},l,\mathbf{k}_i+\mathbf{G},\mathbf{k}_i+\mathbf{G}'} &\equiv (1/\Omega_{\text{at}}) \int \exp[-i(\mathbf{k}_i+\mathbf{G})\cdot\mathbf{r}] U_{\text{ps},l}(\mathbf{r}) \hat{P}_l \exp[i(\mathbf{k}_i+\mathbf{G}')\cdot\mathbf{r}] d^3r \\ &= (1/\Omega_{\text{at}}) (2l+1) 4\pi \int U_{\text{ps},l}(r) j_l(|\mathbf{k}_i+\mathbf{G}|r) j_l(|\mathbf{k}_i+\mathbf{G}'|r) \\ &\quad \times r^2 dr P_l(\cos \gamma). \end{aligned} \quad (10)$$

In the last line of equation (10), the spherical symmetry of the ionic pseudopotential of the isolated atom is assumed, and the  $j_l$  and  $P_l$  are spherical Bessel functions and Legendre polynomials, respectively, with

$$\cos \gamma = [(\mathbf{k}_i+\mathbf{G})\cdot(\mathbf{k}_i+\mathbf{G}')]/(|\mathbf{k}_i+\mathbf{G}||\mathbf{k}_i+\mathbf{G}'|).$$

It is by virtue of the plane-wave expansion that we have a simple analytic expression for the non-local pseudopotentials in equation (10).

In the local pseudopotential approximation, the pseudopotential energy (equation 9) reduces to:

$$\sum_{i,\mu} \int \psi_i^*(\mathbf{r}) U_{\text{ps}}(\mathbf{r}-\mathbf{R}_\mu) \psi_i(\mathbf{r}) d^3r = \Omega \sum_{\mathbf{G}} S(\mathbf{G}) U_{\text{ps}}(\mathbf{G}) \rho(\mathbf{G}). \quad (11)$$

It will be shown later that the multiple sum in equation (9) need not be evaluated in the final expression for the total energy (equation 26). For practical purposes, it is very convenient to decompose the general non-local pseudopotentials into a purely local part and non-local parts, namely,

$$\sum_l U_{\text{ps},l}(\mathbf{r}) \hat{P}_l = U_{\text{ps}}(\mathbf{r}) + \sum_l^{\infty} (U_{\text{ps},l}(\mathbf{r}) - U_{\text{ps}}(\mathbf{r})) \hat{P}_l = U_{\text{ps}}(\mathbf{r}) + \sum_l U'_{\text{ps},l}(\mathbf{r}) \hat{P}_l, \quad (12)$$

where  $U'_{\text{ps},l}(\mathbf{r}) \equiv U_{\text{ps},l}(\mathbf{r}) - U_{\text{ps}}(\mathbf{r})$ , such that the purely local pseudopotential,  $U_{\text{ps}}(\mathbf{r})$ , takes care of the long-range interaction ( $U_{\text{ps}}(\mathbf{r}) \sim -2z/r$  for large  $r$ ) thereby making the non-local parts, the values of  $U'_{\text{ps},l}(\mathbf{r})$ , of short range. This procedure is always possible because at large  $r$ ,  $U_{\text{ps},l}(\mathbf{r}) \sim -2Z/r$  for any  $l$  (i.e. the angular-momentum dependence

disappears for large  $r$ ). We still have infinite degrees of freedom for behaviour of  $U_{\text{ps}}(r)$  at small values of  $r$ . If required,  $U_{\text{ps}}(r)$  can be chosen such that both  $U_{\text{ps}}(r)$  and  $U'_{\text{ps},l}(r)$  are as smooth as possible at small  $r$ . In the local approximation,  $U'_{\text{ps},l}(r) = 0$  identically. Note that the Fourier transforms of the functions  $U'_{\text{ps},l}(r)$  do not have singularity at  $\mathbf{G} = 0$ . Combining equations (9), (11) and (12), the core-valance interaction energy becomes:

$$\sum_{i, \mu, l} \int \psi_i^*(r) U_{\text{ps},l}(r - \mathbf{R}_\mu) \hat{P}_l \psi_i(r) d^3r = \Omega \left( \sum_{\mathbf{G}} S(\mathbf{G}) U_{\text{ps}}(\mathbf{G}) \rho(\mathbf{G}) + \sum_{i, l, \mathbf{G}, \mathbf{G}'} \psi^*(\mathbf{k}_i + \mathbf{G}) \psi(\mathbf{k}_i + \mathbf{G}') S(\mathbf{G}' - \mathbf{G}) U'_{\text{ps},l, \mathbf{k}_i + \mathbf{G}, \mathbf{k}_i + \mathbf{G}'} \right), \quad (13)$$

where the non-diagonal terms in the second sum represent non-local  $\mathbf{k}_i + \mathbf{G}$  to  $\mathbf{k}_i + \mathbf{G}'$  scattering events. The one-electron Schrödinger equation (equation 4) in the momentum representation becomes:

$$\sum_{\mathbf{G}'} [(\mathbf{k}_i + \mathbf{G}')^2 \delta_{\mathbf{G}\mathbf{G}'} + V_{\mathbf{G}\mathbf{G}'}^i] \psi(\mathbf{k}_i + \mathbf{G}') = \epsilon_i \psi(\mathbf{k}_i + \mathbf{G}) \quad (14)$$

where

$$V_{\mathbf{G}\mathbf{G}'}^i = V_{\text{Coul}}(\mathbf{G}' - \mathbf{G}) + \mu_{\text{xc}}(\mathbf{G}' - \mathbf{G}) + S(\mathbf{G}' - \mathbf{G}) [U_{\text{ps}}(\mathbf{G}' - \mathbf{G}) + \sum_l U'_{\text{ps},l, \mathbf{k}_i + \mathbf{G}, \mathbf{k}_i + \mathbf{G}'}]. \quad (15)$$

From equations (1), (8) and (13) it can be seen that the total energy (equation 1) reduces to

$$E_{\text{total}} = \Omega \left( \sum_{i, \mathbf{G}} |\psi(\mathbf{k}_i + \mathbf{G})|^2 (\mathbf{k}_i + \mathbf{G})^2 + \frac{1}{2} \sum_{\mathbf{G}} V_{\text{Coul}}(\mathbf{G}) \rho(\mathbf{G}) + \frac{3}{4} \sum_{\mathbf{G}} \mu_{\text{xc}}(\mathbf{G}) \rho(\mathbf{G}) + \sum_{\mathbf{G}} S(\mathbf{G}) U_{\text{ps}}(\mathbf{G}) \rho(\mathbf{G}) + \sum_{i, l, \mathbf{G}, \mathbf{G}'} \psi^*(\mathbf{k}_i + \mathbf{G}) \psi(\mathbf{k}_i + \mathbf{G}') S(\mathbf{G}' - \mathbf{G}) \times U'_{\text{ps},l, \mathbf{k}_i + \mathbf{G}, \mathbf{k}_i + \mathbf{G}'} \right) + \frac{1}{2} \sum_{\substack{\mu, \nu \\ \mu \neq \nu}} \frac{2Z^2}{|\mathbf{R}_\mu - \mathbf{R}_\nu|}. \quad (16)$$

To simplify equation (16), we multiply on the left of equation (4) by  $\psi_i^*(r)$ , integrate over  $r$  and sum over  $i$ , and substitute the result into equation (1):

$$E_{\text{total}} = \sum_i \epsilon_i - \frac{1}{2} \int \frac{2\rho(r)\rho(r')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' - \frac{1}{4} \int \mu_{\text{xc}}(r) \rho(r) d^3r + \frac{1}{2} \sum_{\substack{\mu, \nu \\ \mu \neq \nu}} \frac{2Z^2}{|\mathbf{R}_\mu - \mathbf{R}_\nu|} = \sum_i \epsilon_i - \Omega \left[ \frac{1}{2} \sum_{\mathbf{G}} V_{\text{Coul}}(\mathbf{G}) \rho(\mathbf{G}) + \frac{1}{4} \sum_{\mathbf{G}} \mu_{\text{xc}}(\mathbf{G}) \rho(\mathbf{G}) \right] + \frac{1}{2} \sum_{\substack{\mu, \nu \\ \mu \neq \nu}} \frac{2Z^2}{|\mathbf{R}_\mu - \mathbf{R}_\nu|} \quad (17)$$

In practice, some mathematical manipulations are necessary to calculate  $E_{\text{total}}$  from equation (17) because  $V_{\text{Coul}}(0)$ ,  $U_{\text{ps}}(0)$  and

$$\sum_{\substack{\mu, \nu \\ \mu \neq \nu}} \frac{2Z^2}{|\mathbf{R}_\mu - \mathbf{R}_\nu|}$$

are individually divergent quantities. First, we solve the band-structure eigenvalue problem, equation (14), with  $V_{\text{Coul}}(0)$  and  $U_{\text{ps}}(0)$  set equal to zero. This corresponds to a

constant shift of the potential (or redefinition of the vacuum level). This arbitrary shift is compensated for by the following procedure. For small  $|G|$ , the local potential is:

$$U_{\text{ps}}(G) = (-8\pi Z/\Omega_{\text{at}} G^2) + \alpha_1 + (\text{higher terms in } G), \quad (18)$$

where the constant term  $\alpha_1$  is given by:

$$\alpha_1 = \lim_{G \rightarrow 0} [U_{\text{ps}}(G) + (8\pi Z/\Omega_{\text{at}} G^2)] = (1/\Omega_{\text{at}}) \int [U_{\text{ps}}(r) + (2Z/r)] d^3r \quad (19)$$

measuring the degree of repulsiveness of the pseudopotential (equivalently, the degree of cancellation of kinetic energy in the core region). Note that equation (18) is no more than a formal expansion because  $G$  takes only discrete values in periodic systems. Also, notice that we do not have a term of order  $1/G$  in the expansion in equation (18). If  $U_{\text{ps}}(G)$  had a term like  $1/G$  in the expansion near  $G = 0$ , then  $\alpha_1$ , as defined by equation (19), would diverge. This, in turn, means that the total energy (equation 26) would diverge. One example of this behaviour is the Simons-Bloch pseudopotential (Simons and Bloch 1973)

$$U_{\text{SB}}(r) = -2Z/r + \sum_l B_l P_l/r^2, \quad (20)$$

where the  $B_l$  are constants depending on  $l$ .  $\alpha_1$  diverges with this potential (unless an artificial cut-off is introduced) because of the unphysically long-range character of the  $1/r^2$  potential added to  $-2Z/r$ . (In fact,  $\alpha_1 \sim 1/G$  rather than a constant.) Even in the presence of the ions neutralising the whole system, the total average potential felt by the electrons is infinite with this pseudopotential. The Simons-Bloch pseudopotential without a cut-off is therefore inadequate for application to extended systems even though it might give good results for localised systems such as atoms or molecules.

Another quantity relevant to the total energy is  $\beta$ , defined from the expansion at small  $G$  of the charge density:

$$\rho(G) = (Z/\Omega_{\text{at}}) + \beta G^2 + (\text{higher terms in } G). \quad (21)$$

From equation (7), we have

$$\beta = \lim_{G \rightarrow 0} \frac{\Omega_{\text{at}}}{8\pi} \left( V_{\text{Coul}}(G) - \frac{8\pi Z}{\Omega_{\text{at}} G^2} \right) = \frac{\Omega_{\text{at}}}{8\pi} \frac{1}{\Omega} \int \int \frac{2[\rho(r) - (Z/\Omega_{\text{at}})]}{|r - r'|} d^3r d^3r'. \quad (22)$$

$\beta$  is a rather complicated integral. Fortunately,  $\beta$  does not appear in the final expression for the total energy because both the electron-electron interaction and the ion-electron interaction contain  $\beta$  and cancellation occurs, as will be shown in equation (25).

The lattice (ion-ion) energy per atom is usually expressed (Coldwell-Horsfall and Maradudin 1960) assuming the average electrostatic potential is zero, as

$$\frac{1}{2} \sum_v \frac{2Z^2}{|R_v|} = \gamma_{\text{Ewald}} + \frac{1}{2} \lim_{G \rightarrow 0} \frac{8\pi Z^2}{\Omega_{\text{at}} G^2}, \quad (23)$$

that is,

$$\gamma_{\text{Ewald}} \equiv \frac{1}{2} \sum_v \frac{2Z^2}{|R_v|} - \frac{1}{2} \lim_{G \rightarrow 0} \frac{8\pi Z^2}{\Omega_{\text{at}} G^2} = \frac{1}{2} \left( \sum_v \frac{2Z^2}{|R_v|} - \frac{1}{\Omega_{\text{at}}} \int \frac{2Z^2}{r} d^3r \right) \quad (24)$$

$$\langle \text{potential} \rangle = 0$$

where the prime means that  $\mathbf{R}_v = 0$  is excluded in the summation. Combining equations (7), (18), (21) and (23), the energy per atom coming from the three divergent terms (with  $S(0) = 1$ ) is:

$$\lim_{G \rightarrow 0} \Omega_{\text{at}} \left[ \frac{1}{2} V_{\text{Coul}}(\mathbf{G}) \rho(\mathbf{G}) + U_{\text{ps}}(\mathbf{G}) \rho(\mathbf{G}) \right] + \frac{1}{2} \sum_v' \frac{2Z^2}{|\mathbf{R}_v|} = \lim_{G \rightarrow 0} \left[ \frac{\Omega_{\text{at}}}{2} \frac{8\pi[(Z/\Omega_{\text{at}} + \beta G^2)]^2}{G^2} \right. \\ \left. + \Omega_{\text{at}} \left( \frac{-8\pi Z}{\Omega_{\text{at}} G^2} + \alpha_1 \right) \left( \frac{Z}{\Omega_{\text{at}}} + \beta G^2 \right) + \frac{1}{2} \frac{8\pi Z^2}{\Omega_{\text{at}} G^2} \right] + \gamma_{\text{Ewald}} = \alpha_1 Z + \gamma_{\text{Ewald}}. \quad (25)$$

In summary, we solve equation (14) with  $V_{\text{Coul}}(0)$  and  $U_{\text{ps}}(0)$  set equal to zero and then add  $(\alpha_1 Z + \gamma_{\text{Ewald}})$  to the total energy. The final expression for the total energy per atom is

$$E_{\text{total}}(\text{per atom}) = \frac{1}{N} \sum_i \epsilon_i - \frac{1}{2} \Omega_{\text{at}} \sum_{\mathbf{G} \neq 0} V_{\text{Coul}}(\mathbf{G}) \rho(\mathbf{G}) - \frac{1}{4} \Omega_{\text{at}} \sum_{\mathbf{G}} \mu_{\text{xc}}(\mathbf{G}) \rho(\mathbf{G}) + \alpha_1 Z \\ + \gamma_{\text{Ewald}}. \quad (26)$$

The first term on the right-hand side is the sum of the electron eigenvalues of the occupied states; the second and third terms correspond to the correction for overcounting of the electron-electron interaction. The fourth term is the correction coming from the 'pseudo' nature of the potential, and the last term is the ion-ion Coulomb energy. The non-local pseudopotential contribution (Equation 9) does not appear explicitly here; it enters equation (26) indirectly through band-structure eigenvalues, values of  $\epsilon_i$  and charge density,  $\rho(\mathbf{G})$ .

Because of the non-uniqueness of the pseudopotential,  $\alpha_1$  may depend on the choice of the pseudopotential. In order for the pseudopotential method to make sense,  $E_{\text{total}}$  should not be significantly affected by a change in  $\alpha_1 Z$ . In fact, the change in  $\alpha_1 Z$  is mostly neutralised by the corresponding change in the value  $\epsilon_1$ . Suppose we have two different pseudopotentials which give equally good band structures, but one has a stronger repulsive core than the other. The ionic pseudopotential with a stronger repulsive core may give a larger value of  $\alpha_1$ , but it also gives rise to a downward shift of the overall band energy  $\epsilon_i$ . This can be seen clearly when we note that the zero of the pseudopotential, by definition, coincides with the spatial average of the pseudopotential because  $U_{\text{ps}}(G = 0)$  is set equal to zero. A stronger repulsive core means a higher average potential, hence, lower eigenvalues relative to the average potential. (Equivalently, change in the pseudopotential close to the origin results in changes in high  $G$  components, which affects the relative band structure very little but causes a rigid shift of the band structure.)

The advantage of the pseudopotential scheme in the total-energy calculation is obvious when we compare the ratio of the cohesive energy to the total energy (in Rydberg units). For tungsten,  $E_{\text{cohesive}}/E_{\text{total}} \sim 0.6/30000$  with all-electron calculation while  $E_{\text{cohesive}}/E_{\text{total}}(\text{pseudo}) \sim 0.6/16$  with pseudopotential calculation. An extreme accuracy is required for the all-electron calculation whereas the condition is three orders of magnitude relaxed using the pseudopotential method.

On the other hand, the usefulness of the momentum representation approach to the total energy rests in part on the convergence rate of the reciprocal lattice sums in equation (26). These are of the form  $\sum_{\mathbf{G}} f(\mathbf{G}) \rho(\mathbf{G})$  where  $f(\mathbf{G})$  is  $V_{\text{Coul}}$ , the exchange potential  $V_{\text{x}}$  or the correlation potential  $V_{\text{corr}}$ . We have examined the convergence rate for the rather



extreme case of the transition metals Mo and W (Zunger and Cohen 1979) (characterised by localised d states). We have found that a convergence of 3 mRyd in the individual momentum sums can be reached by including about 500, 200 and 200 plane waves, respectively, for  $f = V_{\text{coul}}, V_x$  and  $V_{\text{corr}}$ . A convergence of 1 mRyd requires about 1000 plane waves. Note that for  $f = U_{\text{ps}}(\mathbf{G})$  (required if we use equation 16 rather than equation 17) a much higher cut-off is required (e.g. about 1500 plane waves for Mo) due to the localised nature of transition-metal pseudopotentials (Zunger and Cohen 1979). In contrast, only about 200 plane waves are required for a 1 mRyd accuracy for the Si total energy (Ihm and Cohen 1979) because the corresponding pseudopotential is much smoother.

### 3. Hellmann–Feynman theorem

We prove here the Hellmann–Feynman theorem for the pseudopotential Hamiltonian. The  $X\alpha$  approximation is assumed for the exchange–correlation potential. The result is independent of the approximation employed for the exchange–correlation contribution because this theorem has been proved (Hellmann 1937, Feynman 1939, Slater 1972) to hold for the all-electron case both with and without the  $X\alpha$  approximation. Particular care is required for the treatment of the non-local operator; otherwise, the proof exactly parallels Slater’s proof (Slater 1972) for the all-electron case.

There are two different sources contributing to  $\nabla_{\mathbf{R}_\mu} E_{\text{total}}$ . One comes from explicit dependence of the total energy on  $\mathbf{R}_\mu$  and the other from the implicit dependence through  $\psi_i(\mathbf{r})$ , as the solution of the one-electron Schrödinger equation,  $\psi_i$  depends on  $\mathbf{R}_\mu$ . The latter contribution is identically zero as in the all-electron case proven by Slater (Slater 1972) since the non-local operator satisfies the relation:

$$[U_{\text{ps},l}(\mathbf{r}) \hat{P}_l]^\dagger = [\hat{P}_l U_{\text{ps},l}(\mathbf{r}) \hat{P}_l]^\dagger = U_{\text{ps},l}(\mathbf{r}) \hat{P}_l. \quad (27)$$

The rest of the proof is identical to the all-electron case. Therefore, we will consider below only the former contribution.

In equation (3), there are two terms explicitly dependent on  $\mathbf{R}_\mu$ . The negative gradient of the last term in equation (3) with respect to  $\mathbf{R}_\mu$  is

$$-\nabla_{\mathbf{R}_\mu} \left( \frac{1}{2} \sum_{\substack{\mathbf{p}, \mathbf{v} \\ \mathbf{p} \neq \mathbf{v}}} \frac{2Z^2}{|\mathbf{R}_\mathbf{p} - \mathbf{R}_\mathbf{v}|} \right) = -\nabla_{\mathbf{R}_\mu} \sum_{\substack{\mathbf{v} \\ \mathbf{v} \neq \mu}} \frac{2Z^2}{|\mathbf{R}_\mu - \mathbf{R}_\mathbf{v}|} = 2Z^2 \sum_{\substack{\mathbf{v} \\ \mathbf{v} \neq \mu}} \frac{\mathbf{R}_\mu - \mathbf{R}_\mathbf{v}}{|\mathbf{R}_\mu - \mathbf{R}_\mathbf{v}|^3} = \mathbf{F}_1, \quad (28)$$

which gives the force exerted on the ion at  $\mathbf{R}_\mu$  by other ions. Note that we have assumed the ions are spherically symmetric and non-overlapping. The force exerted on the ion by the electrons is rather complicated and has a unique interpretation in the pseudopotential formalism. As we will see shortly, the force has a classical interpretation only if we regard  $-(\frac{1}{4}\pi) \nabla_r^2 U_{\text{ps}}(\mathbf{r})$  as the effective charge density of the ion. We define an angular-momentum-dependent effective charge density of the ion by

$$-(1/4\pi) \nabla^2 U_{\text{ps},l}(\mathbf{r}) \equiv \rho_{\text{ion},l}(\mathbf{r}). \quad (29)$$

$\rho_{\text{ion},l}(\mathbf{r})$  is not directly related to the real charge distribution of the core. Combined with the corresponding angular-momentum-dependent valence charge density, it gives the effective ion–electron interaction without calculating the core-electron contributions separately. The electronic counterpart to  $\rho_{\text{ion},l}(\mathbf{r})$  is the ‘projected electron density’

defined by

$$\sum_i [\hat{P}_i \psi_i(\mathbf{r})]^+ \hat{P}_i \psi_i(\mathbf{r}) = \sum_i \psi_i^*(\mathbf{r}) \hat{P}_i \psi_i(\mathbf{r}) \equiv \rho_{\text{elec},i}(\mathbf{r}). \quad (30)$$

Note that  $\rho_{\text{elec},i}(\mathbf{r})$  is not identical to  $\hat{P}_i \sum_i \psi_i^*(\mathbf{r}) \psi_i(\mathbf{r}) = \hat{P}_i \rho(\mathbf{r})$ . Now we take the gradient of the first term on the right-hand side of equation (3),

$$\begin{aligned} -\nabla_{\mathbf{R}_\mu} \sum_i \int \psi_i^*(\mathbf{r}) \sum_{l,v} U_{\text{ps},l}(\mathbf{r} - \mathbf{R}_v) \hat{P}_i \psi_i(\mathbf{r}) d^3r \\ = -\sum_i \int \psi_i^*(\mathbf{r}) \sum_l [\nabla_{\mathbf{R}_\mu} U_{\text{ps},l}(\mathbf{r} - \mathbf{R}_\mu)] \hat{P}_i \psi_i(\mathbf{r}) d^3r = \sum_i \int \psi_i^*(\mathbf{r}) \sum_l \\ \times [\nabla_{\mathbf{r}} U_{\text{ps},l}(\mathbf{r} - \mathbf{R}_\mu)] \hat{P}_i \psi_i(\mathbf{r}) d^3r. \end{aligned} \quad (31)$$

Using the identity

$$V(\mathbf{r}) = -\frac{1}{4\pi} \int \frac{\nabla_{\mathbf{r}'}^2 V(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r', \quad (32)$$

equation (31) becomes

$$\begin{aligned} \sum_i \int \psi_i^*(\mathbf{r}) \sum_l \nabla_{\mathbf{r}} \left( -\frac{1}{4\pi} \int \frac{\nabla_{\mathbf{r}'}^2 U_{\text{ps},l}(\mathbf{r}' - \mathbf{R}_\mu) d^3r'}{|\mathbf{r} - \mathbf{r}'|} \right) \hat{P}_i \psi_i(\mathbf{r}) d^3r \\ = \sum_i \int \psi_i^*(\mathbf{r}) \sum_l \int \left( -\frac{1}{4\pi} \nabla_{\mathbf{r}'}^2 U_{\text{ps},l}(\mathbf{r}' - \mathbf{R}_\mu) \right) \left( \nabla_{\mathbf{r}} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) d^3r' \hat{P}_i \psi_i(\mathbf{r}) d^3r \\ = \sum_l \int \left[ -\frac{1}{4\pi} \nabla_{\mathbf{r}'}^2 U_{\text{ps},l}(\mathbf{r}' - \mathbf{R}_\mu) \right] \left( -\nabla_{\mathbf{r}'} \int \left( \sum_i \psi_i^*(\mathbf{r}) \hat{P}_i \psi_i(\mathbf{r}) / |\mathbf{r} - \mathbf{r}'| \right) d^3r \right) \\ \times d^3r' = \sum_l \int \rho_{\text{ion},l}(\mathbf{r}' - \mathbf{R}_\mu) \mathbf{E}_l(\mathbf{r}') d^3r' = \mathbf{F}_2, \end{aligned} \quad (33)$$

where

$$\mathbf{E}_l(\mathbf{r}') = -\nabla_{\mathbf{r}'} \cdot \int \frac{\rho_{\text{elec},l}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d^3r \quad (34)$$

is the effective electric field produced by electrons and experienced by  $\rho_{\text{ion},l}$ . It reduces to a usual all-electron result if  $V(\mathbf{r}) = -2Z/r$  because  $\rho_{\text{ion}}$  becomes a delta function from equation (29).  $\mathbf{F}_2$  is the force integrated over the ion localised at  $\mathbf{R}_\mu$  and over the total valence electrons. Therefore,  $(\mathbf{F}_1 + \mathbf{F}_2)$  gives the total Coulombic force exerted on the ion, hence the Hellmann–Feynman theorem.

We now go on to derive the momentum-space expression for the Hellmann–Feynman theorem. Since  $\mathbf{F}_1$  can be calculated exactly using Ewald's method, we will concentrate on the simplification of  $\mathbf{F}_2$  in momentum space.

From equation (9), we have

$$\begin{aligned} \mathbf{F}_2 = -\nabla_{\mathbf{R}_\mu} \Omega \sum_{i,l,\mathbf{G},\mathbf{G}'} \psi^*(\mathbf{k}_i + \mathbf{G}) \psi(\mathbf{k}_i + \mathbf{G}') \sum_v \exp[i(\mathbf{G}' - \mathbf{G}) \cdot \mathbf{R}_v] N^{-1} U_{\text{ps},l,\mathbf{k}_i + \mathbf{G},\mathbf{k}_i + \mathbf{G}'} \\ = -i\Omega_{\text{at}} \sum_{i,l,\mathbf{G},\mathbf{G}'} (\mathbf{G}' - \mathbf{G}) \exp[i(\mathbf{G}' - \mathbf{G}) \cdot \mathbf{R}_\mu] \psi^*(\mathbf{k}_i + \mathbf{G}) \psi(\mathbf{k}_i + \mathbf{G}') \\ \times U_{\text{ps},l,\mathbf{k}_i + \mathbf{G},\mathbf{k}_i + \mathbf{G}'} \end{aligned} \quad (35)$$

In the local pseudopotential approximation, equation (35) is reduced as

$$F_{2, \text{local}} = -\nabla_{R_\mu} \Omega \sum_{\mathbf{v}, \mathbf{G}} \exp(i\mathbf{G} \cdot \mathbf{R}_\mathbf{v}) N^{-1} U_{\text{ps}}(\mathbf{G}) \rho(\mathbf{G}) = -i\Omega_{\text{at}} \sum_{\mathbf{G}} \mathbf{G} \exp(i\mathbf{G} \cdot \mathbf{R}_\mu) \times U_{\text{ps}}(\mathbf{G}) \rho(\mathbf{G}). \quad (36)$$

Summarising, the Hellmann–Feynman theorem in the momentum-space is

$$-\nabla_{R_\mu} E_{\text{total}} = -\nabla_{R_\mu} \sum_{\mathbf{v}} \frac{2Z^2}{|\mathbf{R}_\mu - \mathbf{R}_\mathbf{v}|} - i\Omega_{\text{at}} \sum_{i, l, \mathbf{G}, \mathbf{G}'} (\mathbf{G}' - \mathbf{G}) \exp[i(\mathbf{G}' - \mathbf{G}) \cdot \mathbf{R}_\mu] \times \psi^*(\mathbf{k}_i + \mathbf{G}) \psi(\mathbf{k}_i + \mathbf{G}') U_{\text{ps}, l, \mathbf{k}_i + \mathbf{G}, \mathbf{k}_i + \mathbf{G}'}. \quad (37)$$

Possible applications of the theorem are diverse. For example, we can study the equilibrium configuration of the surface atoms, the surface chemisorption, effects of the impurities and defects in the bulk and the surface, or the phonon modes of solids.

#### 4. Virial theorem

It is a well known fact of classical mechanics that the time average of the bounded motion in a  $r^n$  potential field gives the virial relation (Goldstein 1950)

$$\langle \text{kinetic energy} \rangle = \langle \text{potential energy} \rangle \quad (38)$$

Because the pseudopotential does not show the simple  $1/r$  behaviour, a trivial result is that the usual virial theorem does not hold for the pseudopotential Hamiltonian. For the all-electron Hamiltonian of the solid, Slater has shown (Slater 1972) that

$$\text{KE} = -\frac{1}{2}\text{PE} - \frac{1}{2} \sum_{\mu} \mathbf{R}_\mu \cdot \nabla_{R_\mu} E_{\text{total}}. \quad (39)$$

The core of his proof lies in the relation

$$\begin{aligned} \mathbf{r} \cdot \nabla_r \sum_{\mu} \frac{-2Z}{|\mathbf{r} - \mathbf{R}_\mu|} + \mathbf{R}_\mu \cdot \nabla_{R_\mu} \frac{-2Z}{|\mathbf{r} - \mathbf{R}_\mu|} &= 2Z \sum_{\mu} \frac{r^2 - \mathbf{r} \cdot \mathbf{R}_\mu}{|\mathbf{r} - \mathbf{R}_\mu|^3} + 2Z \sum_{\mu} \frac{R_\mu^2 - \mathbf{r} \cdot \mathbf{R}_\mu}{|\mathbf{r} - \mathbf{R}_\mu|^3} \\ &= \sum_{\mu} \frac{2Z}{|\mathbf{r} - \mathbf{R}_\mu|} = -U(r), \end{aligned} \quad (40)$$

where  $U(r)$  is the Coulomb potential between an electron and nucleus. For the pseudopotential case, we obtain

$$\begin{aligned} \mathbf{r} \cdot \nabla_r \sum_{l, \mu} U_{\text{ps}, l}(|\mathbf{r} - \mathbf{R}_\mu|) + \sum_{l, \mu} \mathbf{R}_\mu \cdot \nabla_{R_\mu} U_{\text{ps}, l}(|\mathbf{r} - \mathbf{R}_\mu|) \\ = \sum_{l, \mu} \left( \frac{\partial U_{\text{ps}, l}(r')}{\partial r'} \right)_{r' = \mathbf{r} - \mathbf{R}_\mu} \left( \frac{r^2 - \mathbf{r} \cdot \mathbf{R}_\mu}{|\mathbf{r} - \mathbf{R}_\mu|} + \frac{R_\mu^2 - \mathbf{r} \cdot \mathbf{R}_\mu}{|\mathbf{r} - \mathbf{R}_\mu|} \right) \\ = \sum_{l, \mu} \left( \frac{r' \cdot \partial U_{\text{ps}, l}(r')}{\partial r'} \right)_{r' = \mathbf{r} - \mathbf{R}_\mu}. \end{aligned} \quad (41)$$

Comparing equations (40) and (41) and using the fact that the purely radial operator  $\mathbf{r} \cdot \nabla_r$  commutes with the angular momentum projection operator  $P$ , we arrive at the

conclusion that equation (39) is to be modified for the pseudopotential Hamiltonian as follows:

$$\begin{aligned} \text{KE} = & -\frac{1}{2}\text{PE} - \frac{1}{2} \sum_{\mu} \mathbf{R}_{\mu} \cdot \nabla_{\mathbf{R}_{\mu}} E_{\text{total}} \\ & + \frac{1}{2} \sum_{i, l, \mu} \int \psi_i^*(\mathbf{r}) \left( U_{\text{ps}, l}(\mathbf{r}') + \mathbf{r}' \cdot \frac{\partial U_{\text{ps}, l}(\mathbf{r}')}{\partial \mathbf{r}'} \right)_{\mathbf{r}' = \mathbf{r} - \mathbf{R}_{\mu}} \hat{P}_l \psi_i(\mathbf{r}) d^3 r. \end{aligned} \quad (42)$$

The difference between the all-electron case (equation 39) and the pseudopotential case (equation 42) originates primarily from the cancellation theorem. The correction term in equation (42), which is usually a large negative quantity, represents the reduction of the kinetic energy due to the repulsive pseudopotential near the nuclei. For example, with the Simons–Bloch pseudopotential in equation (20),  $-B_l/r^2$  enters the parentheses of the correction term in equation (42). Now the total energy is

$$\begin{aligned} E_{\text{total}} = \text{PE} + \text{KE} = & -\text{KE} - \sum_{\mu} \mathbf{R}_{\mu} \cdot \nabla_{\mathbf{R}_{\mu}} E_{\text{total}} \\ & + \sum_{i, l, \mu} \int \psi_i^*(\mathbf{r}) \left( U_{\text{ps}, l}(\mathbf{r}') + \mathbf{r}' \cdot \frac{\partial U_{\text{ps}, l}(\mathbf{r}')}{\partial \mathbf{r}'} \right)_{\mathbf{r}' = \mathbf{r} - \mathbf{R}_{\mu}} \hat{P}_l \psi_i(\mathbf{r}) d^3 r. \end{aligned} \quad (43)$$

In most periodic systems,  $\mathbf{R}_{\mu} \cdot \nabla_{\mathbf{R}_{\mu}} E_{\text{total}}$  of an atom not on the surface is automatically zero by symmetry (The force on individual atoms is zero.). In that case, equation (43) is useful for calculating the total energy at equilibrium (especially so if the pseudopotential is local), although it is not as powerful as in the all-electron case where  $E_{\text{total}}$  is simply the negative of the kinetic energy. We will not examine the trivial momentum-space expression of equation (43).

## Acknowledgments

This work was supported in part by National Science Foundation Grant DMR76-20647-A01 and the Division of Basic Energy Sciences, US Department of Energy.

## Appendix

When crystal wavefunctions include localised features which cannot be represented efficiently by plane waves (e.g. d orbitals in transition metals) it becomes advantageous to employ a 'mixed' representation where real-space basis functions such as Gaussians are present (Louie *et al* 1979);

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}} \psi(\mathbf{k}_i + \mathbf{G}) \exp[i(\mathbf{k}_i + \mathbf{G}) \cdot \mathbf{r}] + \sum_{\mu} C_{\mu}^{(i)} \phi_{\mu}^{(i)}(\mathbf{r}). \quad (\text{A1})$$

Here  $C_{\mu}^{(i)}$  denotes the variational expansion coefficient of the localised orbitals and  $\phi_{\mu}^{(i)}(\mathbf{r})$  denotes a Bloch basis function constructed from the localised real-space basis orbitals  $d_{\mu}(\mathbf{r})$ , where  $\mu$  indicates the central-field quantum numbers ( $n, l, m$ ). A Bloch transformation on  $d_{\mu}(\mathbf{r})$  yields

$$\phi_{\mu}^{(i)}(\mathbf{r}) = N^{-1/2} \sum_{\mathbf{R}_v} \exp(i\mathbf{k}_i \cdot \mathbf{R}_v) d_{\mu}(\mathbf{r} - \mathbf{R}_v). \quad (\text{A2})$$

The kinetic energy associated with the wavefunction in equation (A1) is

$$\langle \psi_i(\mathbf{r}) | -\nabla^2 | \psi_i(\mathbf{r}) \rangle = T_{LL} + T_{PP} + T_{PL}, \quad (\text{A3})$$

where  $T_{LL}$ ,  $T_{PP}$ , and  $T_{PL}$  denote the contributions of the localised-localised, plane wave-plane wave and plane wave-localised orbitals, respectively. To obtain convenient expressions for  $T_{LL}$  and  $T_{PL}$  ( $T_{PP}$  has been calculated in the text), we first Fourier transform the Bloch orbitals,  $\phi_\mu^{(i)}(\mathbf{r})$  as

$$\phi_\mu^{(i)}(\mathbf{r}) = \sum_{\mathbf{G}} S(\mathbf{G}) d_\mu(\mathbf{k}_i + \mathbf{G}) \exp(i(\mathbf{k}_i + \mathbf{G}) \cdot \mathbf{r}), \quad (\text{A4})$$

where  $d_\mu(\mathbf{k}_i + \mathbf{G})$  denotes the Fourier transform of the elemental localised basis function,

$$d_\mu(\mathbf{k}_i + \mathbf{G}) = \int \exp(-i(\mathbf{k}_i + \mathbf{G}) \cdot \mathbf{r}) d_\mu(\mathbf{r}) d^3r. \quad (\text{A5})$$

For simple localised orbitals (e.g. Gaussians, Slater orbitals, etc)  $d_\mu(\mathbf{k}_i + \mathbf{G})$  can be calculated analytically. Now the  $T_{LL}$  in equation (A3) is calculated using

$$\langle \phi_\mu^{(i)}(\mathbf{r}) | -\nabla^2 | \phi_\nu^{(i)}(\mathbf{r}) \rangle = \sum_{\mathbf{G}} (\mathbf{k}_i + \mathbf{G})^2 |S(\mathbf{G})|^2 d_\mu^*(\mathbf{k}_i + \mathbf{G}) d_\nu(\mathbf{k}_i + \mathbf{G}). \quad (\text{A6})$$

The  $T_{PL}$  contribution is calculated from the following relation

$$\langle \exp[i(\mathbf{k}_i + \mathbf{G}) \cdot \mathbf{r}] | -\nabla^2 | \phi_\mu^{(i)}(\mathbf{r}) \rangle = (\mathbf{k}_i + \mathbf{G})^2 S(\mathbf{G}) d_\mu(\mathbf{k}_i + \mathbf{G}). \quad (\text{A7})$$

If there are more than one kind of atoms present, the above expressions become a little complicated. Equation (A1) now reads

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}} \psi(\mathbf{k}_i + \mathbf{G}) \exp[i(\mathbf{k}_i + \mathbf{G}) \cdot \mathbf{r}] + \sum_{\mu\alpha} C_{\mu\alpha}^{(i)} \phi_{\mu\alpha}^{(i)}(\mathbf{r}), \quad (\text{A8})$$

where  $\alpha$  denotes different kinds of atoms (at different sites in the unit cell). Equations (A6) and (A7) become, in this case,

$$\langle \Phi_{\mu\alpha}^{(i)}(\mathbf{r}) | -\nabla^2 | \Phi_{\nu\beta}^{(i)}(\mathbf{r}) \rangle = \sum_{\mathbf{G}} (\mathbf{k}_i + \mathbf{G})^2 S_\alpha^*(\mathbf{G}) S_\beta(\mathbf{G}) d_{\mu\alpha}^*(\mathbf{k}_i + \mathbf{G}) d_{\nu\beta}(\mathbf{k}_i + \mathbf{G}), \quad (\text{A9})$$

and

$$\langle \exp[i(\mathbf{k}_i + \mathbf{G}) \cdot \mathbf{r}] | -\nabla^2 | \Phi_{\mu\alpha}^{(i)}(\mathbf{r}) \rangle = (\mathbf{k}_i + \mathbf{G})^2 S_\alpha(\mathbf{G}) d_{\mu\alpha}(\mathbf{k}_i + \mathbf{G}). \quad (\text{A10})$$

Matrix elements of other operators are calculated in a similar way.

In practice, it is necessary to assess the value of the cutoff,  $G_{\max}$  ( $G < G_{\max}$  in the summation) which will stabilise the summation in equation (A6). As the number of terms in equation (A6) does not affect the size of the secular equation to be solved variationally (the latter is determined by the number of terms in equation A1),  $G_{\max}$  can easily be increased to obtain good convergence.

## References

- Averill F W 1972 *Phys. Rev.* B **6** 3637-42  
 Ching W Y and Gallaway J 1974 *Phys. Rev.* B **9** 5115-21  
 Coldwell-Horsfall R A and Maradudin A A 1960 *J. Math. Phys.* **1** 395-404  
 Danese J B 1974 *J. Chem. Phys.* **61** 3071-80  
 DeCicco P D 1965 *PhD Thesis* Massachusetts Institute of Technology  
 Feynman R P 1939 *Phys. Rev.* **56** 340-3  
 Fuchs K 1935 *Proc. R. Soc.* **151** 585-602

- Goldstein H 1950 *Classical Mechanics* (Mass.: Addison-Wesley) 71
- Goroff I and Kleinman L 1970 *Phys. Rev. B* **1** 2514–81
- Harris F F and Monkhorst H J 1971 *Solid St. Commun.* **9** 1449–53
- Harrison W A 1966 *Pseudopotential in the Theory of Metals* (New York: Benjamin)
- Heine V and Weaire D 1970 *Solid State Physics* vol 24 ed Ehrenreich, Seitz and Turnbull (New York: Academic Press)
- Hellmann H 1937 *Einführung in die Quantenchemie* (Leipzig: Deuticke)
- Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864–71
- Ihm J and Cohen M L 1979 *Solid St. Commun.* **29** 711–4
- Janak J F 1974 *Phys. Rev. B* **9** 3985–8
- Janak J F, Moruzzi V L and Williams A R 1975 *Phys. Rev. B* **12** 1257–61
- Janak J F and Williams A R 1976 *Phys. Rev. B* **14** 4199–204
- Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133–8
- Louie S. G, Ho K M and Cohen M L 1979 *Phys. Rev. B* **19** 1774–82
- Moriarty J A 1974 *Phys. Rev. B* **10** 3075–91
- 1974 *Phys. Rev. B* **16** 2537–55
- Phillips J C and Kleinman L 1959 *Phys. Rev.* **116** 287–94
- Sabin J R, Worth J P and Trickey S B 1975 *Phys. Rev. B* **11** 3658–61
- Schaefer H F 1972 *The Electronic Structure of Atoms and Molecules* (Mass.: Addison-Wesley)
- Simons G and Bloch A N 1973 *Phys. Rev. B* **7** 2754–61
- Slater J C 1972 *J. Chem. Phys.* **57** 2389–96
- 1974 *The Self-Consistent Field for Molecules and Solids* vol 4 (New York: McGraw-Hill)
- Snow E C 1973 *Phys. Rev.* **8** 5391–7
- Wendel H and Martin R M 1978 *Phys. Rev. Lett.* **40** 950–3
- Wepfer G G, Euwema R N, Surrat G T and Wilhite D L 1974 *Phys. Rev. B* **9** 1670–3
- Wigner E and Seitz F 1933 *Phys. Rev.* **43** 804–10
- 1934 *Phys. Rev.* **46** 509–24
- Zunger A and Freeman A J 1977a *Phys. Rev. B* **15** 5049–65
- 1977b *Phys. Rev. B* **16** 906–24
- 1977c *Phys. Rev. B* **16** 2901–26
- 1978 *Phys. Rev. B* **17** 2030–42
- Zunger A and Cohen M L 1979 *Phys. Rev. B* **19** 568–82