The tight-binding bond model

View the article online for updates and enhancements.

Related content

Recent citations

The tight-binding bond model

A P Sutton†, M W Finnis‡, D G Pettifor§ and Y Ohta‖

† Department of Metallurgy and Science of Materials, Oxford University, Parks Road, Oxford OX1 3PH, UK
‡ Theoretical Physics Division, AERE Harwell, Didcot OX11 0RA, UK
§ Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ, UK
‖ Department of Applied Physics, Nagoya University, Nagoya 464, Japan

Abstract. We present a tight-binding model of cohesion and interatomic forces which exploits the variational principle of density functional theory. The binding energy of a solid is expressed as a sum of four terms, each of which has a clear physical meaning. The first two terms are the covalent bond and promotion energies, which are found by solving the electronic Hamiltonian to obtain the density matrix. The remaining two terms describe changes in the total electrostatic and exchange–correlation energies on forming the solid from isolated atoms. The variational principle allows these two terms to be expressed as functionals of a superposition of frozen atomic charge densities. We show that they may then be approximated by a sum of pair potentials.

The importance of self-consistency in tight-binding models is discussed with particular attention to the evaluation of the bulk modulus and certain interatomic force constants by frozen-phonon calculations. It is shown that serious errors may arise in non-self-consistent models due to the violation of charge conservation and the neglect of variations in the potential caused by charge flow. We advocate local charge neutrality as the simplest approximation to self-consistency which overcomes these problems. This assumption leads to a remarkably simple expression for the force on an atom due to its neighbours, which is both physically transparent and computationally efficient.

These concepts are illustrated for three-dimensional solids by calculations of covalent bond energies in BCC, FCC and HCP transition metals using a canonical d-band model. Model one-dimensional calculations are also presented which illustrate the computation of covalent bond energies and interatomic forces at surfaces and interfaces, and the importance of local charge neutrality in the model.

1. Introduction

The calculation of the atomic and electronic structures of extended defects in crystalline solids is limited by the size and symmetry of the system. Improvements in computing power and numerical algorithms have recently enabled more ambitious first principles calculations to be done (Car and Parrinelo 1985, Payne *et al* 1987) but current calculations are limited to about sixty non-equivalent atoms in the computational cell. Consequently a variety of empirical and semi-empirical schemes has been developed to describe interatomic forces in solids. In some of these schemes, such as the embedded atom method (Stott and Zaremba 1980, Norskov 1982, Daw and Baskes 1984, Finnis and

Sinclair 1984, Ercolessi *et al* 1986), the local electronic structure enters into the inter-atomic forces only in some locally defined average sense and information about the local bonding is not available. A relatively simple scheme for relating the local atomic and electronic structures is the tight-binding method. The tight-binding theory of cohesion in solids has been extensively applied to describe the energies of transition metals and their alloys (Heine *et al* 1980, Pettifor 1983), sp-bonded elemental solids (Chadi and Cohen 1975, Harrison 1980, Paxton *et al* 1987) and compounds (Majewski and Vogel 1986) and pd-bonded compounds (Pettifor and Podloucky 1986).

Most of the earlier applications of the tight-binding method calculated the energies of atomic assemblies in arbitrary configurations, but the configuration had to be pre-scribed in advance. The use of tight-binding Hamiltonians to minimise the energy of atomic configurations required two developments. Firstly the total energy of the system is not simply the sum of occupied one-electron eigenvalues, or 'band energy'. In Chadi's pioneering work on calculating surface reconstructions (Chadi 1979, 1984) in sp-bonded solids the band energy is augmented by a short-range repulsive-pair potential. Chadi (1979) interpreted the pair potential in the Hartree approximation as representing the difference between the ion–ion and electron–electron electrostatic interaction energies.

This form for the total energy in tight-binding calculations has proved to be the most widely used for transition metals as well (see, for example, Legrand 1985) and we shall refer to this model as the tight-binding band model, or simply the band model. Secondly, a prescription for calculating the force on an atom had to be developed. Whereas this is a trivial problem for the pair potential component of the total energy, the derivative of the band energy is non-trivial and several schemes have been proposed (Chadi 1979, 1984, Moraitis and Gautier 1979, Sato and Masuda 1981, Boswarva and Esterling 1982, Treglia *et al* 1983, Masuda *et al* 1983, Legrand 1985, Sankey and Allen 1986, Pollmann *et al* 1986, Esterling *et al* 1987). Conceptually the simplest scheme is to differentiate the total density of states with the constraint that the total number of electrons in the system is conserved (Chadi 1979, Sato *et al* 1981, Boswarva *et al* 1982, Treglia *et al* 1983, Masuda *et al* 1983, Legrand 1985, Esterling *et al* 1987). This has been done sometimes by numerical differentiation involving a variety of simplifying assumptions about each local density of states (see, for example, Boswarva *et al* 1982, Legrand 1985, Esterling *et al* 1987). With this method one compares the local density of states before and after an atomic displacement at all sites that may be affected by the displacement. It is clear that many more atoms may be affected than merely those within the range of the Hamiltonian. A more efficient method is to write the band energy as the trace of the product of the density matrix and the Hamiltonian (see, for example, Williams *et al* 1982). The derivative of the band energy is then the trace of the product of the density matrix and the derivative of the Hamiltonian. To our knowledge this formula was first discovered by Coulson (1939) and it has been rediscovered many times (see, for example, Moraitis and Gautier 1979, Chadi 1984, Sankey *et al* 1986, Pollmann *et al* 1986). As noted by other authors (for example, Moraitis and Gautier 1979, Sankey and Allen 1986) the atoms that contribute explicitly to this force are only those within the range of the Hamiltonian. A further advantge of Coulson's formula is that the total force on an atom is broken down into contributions from each neighbouring atom, just like a force arising from a pairwise interaction, but, unlike a pairwise interaction, each contribution is a true many-body force which is generally non-central.

A detailed justification for the pair potential in the band model has not been presented. Indeed one may question whether the remaining terms in the Kohn–Sham total energy expression (Kohn and Sham 1965), after the band energy has been

subtracted, are representable by a pair potential. The absence of any form of self-consistency in the band model is also a cause for concern. In § 3 it is shown that there can be a significant error in the bulk modulus computed from the long-wavelength limit of phonons in the band model due to the absence of self-consistency. The same argument applies to force constants and thus by inference to inter-atomic forces.

In § 2 we present a model of cohesion and interatomic forces which we call the tight-binding bond (TBB) model. Following Harris (1985) and Foulkes (1987) we exploit the variational principle of density functional theory to express the total energy of the solid as a functional of an approximate charge density and a charge density obtained by solving the Schrödinger equation once. The variational principle tells us that the error in the total energy expression is second order in the error in the charge density. By choosing a superposition of atomic charge densities for the approximate charge density, as in Harris (1985) and Chelikowsky and Louie (1984), the total energy expression is transformed (Harris 1985) into a physically transparent form for the cohesive energy. In this form the cohesive energy is given by a sum of bond and promotion energies and two other terms. The latter two terms represent the changes in the total electrostatic interaction energy and the exchange–correlation energy as the atoms are condensed to form the solid. We show that they may be approximated by a sum of pair potentials. The bond and promotion energies may be calculated by solving the Schrödinger equation either in reciprocal space or real space. Motivated by the considerations of § 3 and by the desire to eliminate long-range Coulomb terms local charge neutrality is introduced into the model. This leads to a remarkably simple form for the force on an atom which is both physically appealing and computationally efficient. The total force on an atom is decomposed into individual bond forces from neighbouring atoms and contributions from the pair potential. Each bond force is dependent on the local atomic environment and it is also influenced by all atoms being charge neutral. Other workers (see, for example, Moraitis and Gautier 1979, Treglia *et al* 1983, Sayers 1984, Priester *et al* 1986) have also noted the importance of charge neutrality in tight-binding models.

Section 4 contains some exact results for linear chains obtained with the band and TBB models. It is argued that linear chains may be regarded as simple models for studying surfaces and interfaces. The relaxation of a semi-infinite chain as a model of a free surface is studied in the absence of local charge neutrality. The size dependence of finite chain models of a semi-infinite chain is studied both with and without local charge neutrality. It is shown that local charge neutrality can have a profound influence on conclusions about local bonding and interatomic forces. In § 5 we illustrate the above concepts for three-dimensional crystal structures by a study of bonding in transition metals with BCC, FCC and HCP structures, using a canonical d-band model. The σ-, π- and δ-character of the bonds is shown explicitly, and the effect of symmetry on nearest-neighbour forces is illustrated by the difference between HCP and FCC structures. We present our conclusions in § 6.

2. The tight-binding bond model

2.1. The total energy

Our starting point is the density functional scheme for the solid as developed by Harris (1985) and Foulkes (1987). In this scheme an approximate charge density, ρ^f, is assumed to be a reasonable approximation to the exact charge density in the solid. The effective

single particle potential of the Kohn–Sham equations (Kohn and Sham 1965) is expressed as a functional of ρ^f as follows.

$$\bar{V}(r) = v(r) + V^f(r) + \mu_{xc}^f(r) \tag{2.1}$$

Here $v(r)$ is the total ionic potential in the solid where each ion core is represented by a pseudopotential. V^f and μ_{xc}^f are the Hartree and exchange–correlation potentials corresponding to the charge density ρ^f. An approximate single-particle Hamiltonian $\bar{H}$ is constructed with this effective potential.

$$\bar{H} = -\tfrac{1}{2}\nabla^2 + \bar{V}(r). \tag{2.2}$$

This Hamiltonian is solved once and an output charge density $\rho^{out}(r)$ is constructed from its eigenstates. No self-consistency iterations are performed, but by exploiting the variational principle obeyed by the total energy it can be shown (Harris 1985, Foulkes 1987) that the leading corrections to the total energy are second order in the difference between ρ^f and the exact charge density, ρ^{sc}, in the solid. The approximate expression (Harris 1985, Foulkes 1987) for the total energy of the solid is as follows.

$$E \simeq \sum_n a_n\, \tilde{\varepsilon}_n - \int dr\, \rho^f(r)(\tfrac{1}{2}V^f(r) + \mu_{xc}^f(r)) + E_{xc}[\rho^f] + E_{ii}$$
$$+ O(\rho^{sc} - \rho^f)^2 + O(\rho^{sc} - \rho^{out})^2. \tag{2.3}$$

The eigenvalues of the Hamiltonian $\bar{H}$ are denoted by $\tilde{\varepsilon}_n$ and their occupation numbers by a_n. $E_{xc}[\rho^f]$ is the exchange–correlation energy for the charge density ρ^f. E_{ii} is the inter–ionic interaction. The choice of ρ^f is arbitrary but since the error in the total energy varies as the square of the error in the assumed charge density it is obviously desirable to assume a charge density that is not too remote from the self-consistent charge density. We can make contact with tight-binding theory by assuming that ρ^f is a superposition of atomic charge densities, as assumed by Harris (1985) in his studies of molecules. That is, we envisage that ρ^f is formed by condensing infinitely separated atoms to their final positions in the solid without allowing the atomic charge densities to respond during the condensation. Since V^f is linear in ρ^f it follows that the Hartree potential in the solid is a superposition of atomic Hartree potentials. Harris (1985) found that even when a strong covalent bond is formed in a homonuclear dimer such as C_2 the errors caused by this approximation, compared with the fully self-consistent solution, in the binding energy, bond length and vibration frequency were less than 21%, 7% and 17% respectively. If we were to make the Hartree approximation then the Hamiltonian, $\bar{H}$, in (2.2) would be of the usual tight-binding form (Callaway 1974). But we shall retain the exchange and correlation terms in the total energy expression and we shall show how they can be approximated by a sum of pair potentials. To do this, however, it is convenient to work with the binding energy of the solid rather than the total energy.

So far we have been working in the r-representation. Since we shall be using an atomic orbital representation we first express the total energy in a basis-independent form.

Let the eigenstates of $\bar{H}$ be $|n\rangle$. Then $\bar{H}$ can be written as

$$\bar{H} = \sum_n |n\rangle \tilde{\varepsilon}_n \langle n| \tag{2.4}$$

and the output density operator ρ^{out} can be written as

$$\rho^{out} = \sum_n a_n |n\rangle\langle n|. \tag{2.5}$$

The band energy $\Sigma_n a^n \bar{\varepsilon}_n$ is immediately seen to be $\mathrm{Tr}\, \rho^{\mathrm{out}} \tilde{H}$, where Tr denotes trace. Since the integration in (2.3) is also a trace we can rewrite (2.3) in a basis-independent form as follows.

$$E \simeq \mathrm{Tr}\, \rho^{\mathrm{out}} \tilde{H} - \mathrm{Tr}\, \rho^{\mathrm{f}}(V^{\mathrm{f}}/2 + \mu_{\mathrm{xc}}^{\mathrm{f}}) + E_{\mathrm{xc}}[\rho^{\mathrm{f}}] + E_{\mathrm{ii}}. \tag{2.6}$$

Following Harris (1985) we can transform this expression for the total energy into a much more physically transparent expression for the binding energy of the solid. This transformation takes advantage of the fact that ρ^{f} is a superposition of atomic charge densities.

$$\rho^{\mathrm{f}} = \sum_i \rho_i \tag{2.7}$$

where ρ_i is the free atomic valence charge density of the atom at site i. As we have already remarked it follows that

$$V^{\mathrm{f}} = \sum_i V_i^{\mathrm{f}} \tag{2.8}$$

where V_i^{f} is the Hartree potential of the non-interacting atomic charge density at site i. By adding and subtracting $\mathrm{Tr}\, \rho^{\mathrm{f}}\tilde{H}$ from the total energy and by writing $\tilde{H}$ as

$$\tilde{H} = T + V^{\mathrm{f}} + \mu_{\mathrm{xc}}^{\mathrm{f}} + v \tag{2.9}$$

where T is the kinetic energy, we find that the total energy can be expressed as

$$\begin{aligned}
E \simeq\ & \mathrm{Tr}(\rho^{\mathrm{out}} - \rho^{\mathrm{f}})\tilde{H} \\
& + \mathrm{Tr}\, \sum_i \rho_i \left(\sum_{j \neq i} V_j^{\mathrm{f}}/2 + v_j \right) + E_{\mathrm{ii}} \\
& + E_{\mathrm{xc}}[\rho^{\mathrm{f}}] - \sum_i E_{\mathrm{xc}}[\rho_i] \\
& + \sum_i \{ \mathrm{Tr}\, \rho_i(T + V_i^{\mathrm{f}}/2 + v_i) + E^{\mathrm{xc}}[\rho_i] \}.
\end{aligned} \tag{2.10}$$

The pseudopotential of the ionic core at site j is denoted by v_j. The last line in (2.10) is simply the sum of the total energies of the isolated atoms in their non-interacting states, i.e. before they are condensed to form the solid. When this term is taken over to the left-hand side of (2.10) the remaining terms on the right-hand side are equal to the binding energy, E_{B}. Here we have assumed that the free atom is not spin-polarised. In fact spin-polarisation can be an important contribution to the energy of an atom and, as noted by Harris (1985), it has to be added separately to the energies of the free atoms in order to be able to make comparisons with full local spin-density functional calculations or experimental measurements of binding energies of solids. Consider the terms on the second line. They are equal to the change in the total electrostatic energy, $\Delta E_{\mathrm{es}}[\rho^{\mathrm{f}}]$, of all valence electrons and ion cores when the atoms are condensed from infinity to make the solid. It is emphasised that ΔE_{es} includes electron–electron, electron–ion and ion–ion electrostatic interactions. Similarly the terms in the third line are the change in the exchange and correlation energy, $\Delta E_{\mathrm{xc}}[\rho^{\mathrm{f}}]$, in forming the charge density ρ^{f} from isolated atomic charge densities. Thus, the following simple expression is obtained for the binding energy of the solid.

$$E_{\mathrm{B}} \simeq \mathrm{Tr}(\rho^{\mathrm{out}} - \rho^{\mathrm{f}})\tilde{H} + \Delta E_{\mathrm{es}}[\rho^{\mathrm{f}}] + \Delta E_{\mathrm{xc}}[\rho^{\mathrm{f}}]. \tag{2.11}$$

This expression for the binding energy is equivalent to equation (3.7) of Harris (1985).

Since it is an exact rearrangement of the expression for the total energy the errors in (2.11) are again second order in $(\rho^{sc} - \rho^f)$. The physical meaning of each term in (2.11) is clear. The first term arises from the bonding and charge redistribution which takes place when the solid formed from its constituent atoms. The physical meaning of the other two terms has already been discussed. This is in direct contrast to (2.3), which is assumed by most authors, where the band energy is complemented by the unphysical ion–ion interaction energy minus the 'double counting term', and by the exchange–correlation energy minus the exchange–correlation potential terms.

2.2. The covalent bond energy and promotion energy contributions

In this section we express the binding energy in an atomic orbital representation. The orbitals we use are real atomic orbitals on each atom. The orbitals on atom j are denoted by $\varphi_{j\alpha}$ where α denotes the orbital type. Whereas the orbitals on a given atom form an orthonormal set for the isolated atom, in the solid the set of all atomic orbitals is not orthonormal. We assume that the atomic orbital basis set in the solid is a complete set. This is an approximation that can always be improved by enlarging the basis set (see Williams *et al* 1982). The overlap matrix **S** has elements defined by

$$S_{i\alpha j\beta} = \langle \varphi_{i\alpha} | \varphi_{j\beta} \rangle \tag{2.12}$$

and the Hamiltonian has matrix elements

$$\tilde{H}_{i\alpha j\beta} = \langle \varphi_{i\alpha} | \tilde{H} | \varphi_{j\beta} \rangle. \tag{2.13}$$

Ballentine and Kolar (1986) have introduced covariant and contravariant tensor notation into the matrix algebra associated with non-orthogonal basis functions. We agree with these authors that the clarity and simplicity which this notation introduces is considerable and we therefore adopt it here[†]. Corresponding to the complete set of atomic orbitals $\{\varphi_{i\alpha}\}$ (which is called the direct basis) there is a dual basis $\{\varphi^{i\alpha}\}$ that is constructed from the direct basis as follows

$$|\varphi^{i\alpha}\rangle = \sum_{j\beta} |\varphi_{j\beta}\rangle (S^{-1})^{j\beta i\alpha}. \tag{2.14}$$

To be consistent with the usual tensor notation in which matrix multiplication involves summation over repeated superscripted and subscripted indices, matrix elements of the inverse of the overlap matrix are written with superscripts. The direct and dual basis sets are 'bi-orthogonal'.

$$\langle \varphi^{i\alpha} | \varphi_{j\beta} \rangle = \delta^{i\alpha}_{j\beta} \tag{2.15}$$

where

$$\delta^{i\alpha}_{j\beta} = \begin{cases} 1 & \text{if } i\alpha = j\beta \\ 0 & \text{if } i\alpha \neq j\beta. \end{cases} \tag{2.16}$$

[†] This notation may be unfamiliar at first, but it has an exact analogue in crystallography (see, for example, Christian 1975) where the direct basis consists of three non-coplanar primitive-lattice vectors and the dual basis consists of the corresponding three reciprocal-lattice vectors. The analogue of the overlap matrix is the metric of the crystal lattice. The use of tensor notation in non-cubic crystallography is well established as in elegant and useful tool (see, for example, Bilby and Crocker 1965). Foulkes and Haydock (1986) have also chosen to use this notation in their analysis of expectation values of operators expressed in atomic orbital bases.

In general four types of matrix element of an operator exist depending on whether one uses basis functions from the direct basis only, the dual basis only or functions from either basis. For example the four Hamiltonian matrix elements are

$$\tilde{H}_{i\alpha j\beta} = \langle \varphi_{i\alpha} | \tilde{H} | \varphi_{j\beta} \rangle \qquad \tilde{H}^{i\alpha j\beta} = \langle \varphi^{i\alpha} | \tilde{H} | \varphi^{j\beta} \rangle$$
$$\tilde{H}^{i\alpha}{}_{j\beta} = \langle \varphi^{i\alpha} | \tilde{H} | \varphi_{j\beta} \rangle \qquad \tilde{H}_{i\alpha}{}^{j\beta} = \langle \varphi_{i\alpha} | \tilde{H} | \varphi^{j\beta} \rangle. \tag{2.17}$$

The matrices $\tilde{H}_{i\alpha j\beta}$ and $\tilde{H}^{i\alpha j\beta}$ are Hermitian but the matrices in the mixed representations are not Hermitian and instead they obey

$$\tilde{H}_{i\alpha}{}^{j\beta} = (\tilde{H}^{j\beta}{}_{i\alpha})^*. \tag{2.18}$$

One can transform matrix elements in one representation into those of another representation using the overlap matrix or its inverse

$$\tilde{H}^{i\alpha}{}_{j\beta} = \sum_{l\gamma} (S^{-1})^{i\alpha l\gamma} \tilde{H}_{l\gamma j\beta} = \sum_{l\gamma} \tilde{H}^{i\alpha l\gamma} S_{l\gamma j\beta} = \sum_{l\gamma} \sum_{p\mu} (S^{-1})^{i\alpha l\gamma} \tilde{H}_{l\gamma}{}^{p\mu} S_{p\mu j\beta}. \tag{2.19}$$

The identity operator is

$$I = \sum_{i\alpha} | \varphi^{i\alpha} \rangle \langle \varphi_{i\alpha} | \tag{2.20}$$

and it has matrix elements

$$I_{i\alpha j\beta} = S_{i\alpha j\beta} \qquad I^{i\alpha j\beta} = (S^{-1})^{i\alpha j\beta} \qquad I^{i\alpha}{}_{j\beta} = I_{j\beta}{}^{i\alpha} = \delta^{i\alpha}_{j\beta}. \tag{2.21}$$

See Ballentine and Kolar (1986) for further details about this notation. The significance of the dual basis set is realised when one considers the relation between the matrix elements of an operator in some atomic orbital basis and the expansion coefficients of its real-space representation in the same basis. These matrix elements and expansion coefficients are identical only with an orthonormal basis. An important quantity for our purposes is the atomic charge density ρ_i which is obtained as a sum over the orbitals on atom i.

$$\rho_i(r) = \sum_{\alpha} \langle r | \varphi_{i\alpha} \rangle \rho^{i\alpha i\alpha} \langle \varphi_{i\alpha} | r \rangle \tag{2.22}$$

where the $\rho^{i\alpha i\alpha}$ are expansion coefficients (Baraff and Schluter 1979) and they are constants. Since ρ^f is the superposition of such atomic charge densities we find

$$\rho^f(r) = \sum_{i\alpha} \langle r | \varphi_{i\alpha} \rangle (\rho^f)^{i\alpha i\alpha} \langle \varphi_{i\alpha} | r \rangle \tag{2.23}$$

where $(\rho^f)^{i\alpha i\alpha} = \rho^{i\alpha i\alpha}$ of (2.22). Then when $\rho^f(r)$ is expanded in the direct basis set $\{\varphi_{i\alpha}\}$ the expansion coefficients are matrix elements of ρ^f in the dual basis. It is important to note that ρ^f is diagonal in the real atomic orbital representation and that the expansion coefficients are constants that are independent of the local atomic environment.

We express $\mathrm{Tr}(\rho^{\mathrm{out}} - \rho^f)\tilde{H}$ in this atomic orbital representation as a sum over on-site terms and a sum over inter-site terms

$$\mathrm{Tr}(\rho^{\mathrm{out}} - \rho^f)\tilde{H} = \sum_{\substack{i \\ \alpha\beta}} [(\rho^{\mathrm{out}})^{i\alpha i\beta} - (\rho^f)^{i\alpha i\beta} \delta^{i\alpha}_{i\beta}]\tilde{H}_{i\beta i\alpha} + \sum_{\substack{i\alpha \, j\beta \\ i \neq j}} (\rho^{\mathrm{out}})^{i\alpha j\beta} \tilde{H}_{j\beta i\alpha}. \tag{2.24}$$

Consider the first term. In a perfect cubic crystal $\tilde{H}_{i\alpha i\beta} = \tilde{H}_{i\alpha i\beta} \delta^{i\beta}_{i\alpha}$. The diagonal elements $\tilde{H}_{i\alpha i\alpha}$ are the free atomic term values corrected by the crystal-field terms in the solid.

Thus this term is simply the promotion energy, i.e. the energy associated with the change of occupancy of the atomic orbitals on forming the solid from free atoms. For example in diamond formed from carbon atoms in s^2p^2 states the promotion energy is $(\bar{\varepsilon}_p - \bar{\varepsilon}_s)$ if the final state of hybridisation is s^1p^3 where $\bar{\varepsilon}_s$ and $\bar{\varepsilon}_p$ are the term values of the free carbon atom corrected by crystal-field terms. If the atomic environment around each site in the solid is distorted then $\tilde{H}_{i\alpha i\beta}$ are non-zero because of crystal-field terms. It is possible to diagonalise the part of the Hamiltonian $\tilde{H}$ associated only with site i and thus express the terms involving site i in (2.24) as a sum of promotion energies at site i but the new diagonal elements of $\tilde{H}$ will vary from site to site.

We call the second term in (2.24) the covalent bond energy of the solid. It is equal to the sum of the covalent energies of individual bonds between orbitals on different atoms. The covalent bond energy is part of the band energy.

$$E_{\text{cov}} = \tfrac{1}{2} \sum_{\substack{i\alpha\, j\beta \\ i \neq j}} 2(\rho^{\text{out}})^{i\alpha j\beta} \tilde{H}_{j\beta i\alpha} = E_{\text{band}} - \sum_{\substack{i \\ \alpha\beta}} (\rho^{\text{out}})^{i\alpha i\beta} \tilde{H}_{i\beta i\alpha}. \tag{2.25}$$

The covalent energy of the bond between atoms i and j is equal to

$$\sum_\alpha \sum_\beta 2(\rho^{\text{out}})^{i\alpha j\beta} \tilde{H}_{j\beta i\alpha}$$

where it has been assumed that the Hamiltonian matrix elements are real. The covalent bond energy of the solid is therefore equal to the sum of the covalent energies of all bonds and the factor of $\tfrac{1}{2}$ in front of the double sum of (2.25) compensates for the double counting of each bond. The off-diagonal density matrix element $(\rho^{\text{out}})^{i\alpha j\beta}$ is called the bond order between orbitals $\varphi^{i\alpha}$ and $\varphi^{j\beta}$ (Coulson 1939). It is equal to one half of the difference between the occupation of the bonding orbital $(1/2)^{1/2}(\varphi^{i\alpha} + \varphi^{j\beta})$ and the anti-bonding orbital $(1/2)^{1/2}(\varphi^{i\alpha} - \varphi^{j\beta})$. If $\tilde{H}_{i\alpha i\beta} = \tilde{H}_{i\alpha i\beta}\delta_{i\alpha}^{i\beta}$ and the atomic-orbital basis set were orthonormal the covalent bond energy of the solid could also be expressed as

$$E_{\text{cov}} = \sum_{i\alpha} \int^{E_F} n_{i\alpha}(E)(E - \tilde{H}_{i\alpha i\alpha}) \, dE \tag{2.26}$$

where $n_{i\alpha}(E)$ is the local density of states of orbital $\varphi_{i\alpha}$ (Allan 1970).

The off-diagonal density matrix elements $(\rho^{\text{out}})^{i\alpha j\beta}$ are related to Green function matrix elements defined by

$$\sum_{j\beta} (z S_{i\alpha j\beta} - \tilde{H}_{i\alpha j\beta}) G^{j\beta l\lambda}(z) = \delta_{i\alpha}^{l\lambda} \tag{2.27}$$

where z is a complex number in general. Then

$$(\rho^{\text{out}})^{i\alpha j\beta} = -\frac{2 \, \text{Im}}{\pi} \int^{E_F} G^{i\alpha j\beta}(E^+) \, dE \tag{2.28}$$

where the integration variable is the energy E plus an infinitesimally small positive imaginary part. The integration is taken to the Fermi energy E_F to ensure that only occupied states contribute to the bond order. The factor of 2 accounts for spin degeneracy in non-magnetic systems. It may be advantageous numerically to use an integration contour which is more remote from the real energy axis in the complex plane. See Williams *et al* (1982) for a discussion of such contours.

2.3. *The pair potential contribution*

We consider now the remaining terms, $\Delta E_{\text{es}}[\rho^f]$ and $\Delta E_{\text{xc}}[\rho^f]$ in the binding energy of

the solid, equation (2.11). We will argue that because $\rho^{\rm f}$ is a superposition of atomic charge densities these two terms may be approximated by a sum of two-body interactions, centred at atomic sites.

It is obvious that $\Delta E_{\rm es}[\rho^{\rm f}]$ is a sum of pair potentials because it is given by

$$\Delta E_{\rm es} = \sum_i \sum_{j \neq i} {\rm Tr}\, \rho_i (V_j^{\rm f}/2 + V_j) + E_{\rm ii}. \qquad (2.29)$$

This simple form for the change in the total electrostatic energy is a consequence of the fact that $\Delta E_{\rm es}$ is a functional of $\rho^{\rm f}$ and not $\rho^{\rm out}$.

Following Lindholm and Lundqvist (1985) we separate the exchange and correlation contributions to $\Delta E_{\rm xc}[\rho^{\rm f}]$. Lindholm and Lundqvist show that the correlation energy, $E_{\rm c}$, in an electron gas of density ρ is given exactly by

$$E_{\rm c} = \tfrac{1}{2} \iint {\rm d}\boldsymbol{r}\, {\rm d}\boldsymbol{r}'\, \rho(\boldsymbol{r}) \frac{h_{\rm c}(\boldsymbol{r}, \boldsymbol{r}', \rho)}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho(\boldsymbol{r}'). \qquad (2.30)$$

The function $h_{\rm c}$ is related to the pair correlation function and it arises from the Coulomb correlations only. In general $h_{\rm c}$ is a function of $\boldsymbol{r}$ and $\boldsymbol{r}'$ and also of the charge density $\rho(\boldsymbol{r})$. However, as pointed out by Lindholm and Lundqvist, many studies using configuration interaction rather than density functional theory to describe electron correlation indicate that the function $h_{\rm c}$ has only a weak dependence on the charge density which may be neglected to first order. This is the justification for the common practice of expressing the correlation energy as a sum of pair correlation energies (Hurley 1976, Kutzelnigg 1977). The change in the correlation energy associated with the condensation of atoms from infinity to form the solid with charge density $\rho = \rho^{\rm f}$ is then found to be

$$\Delta E_{\rm c} = \tfrac{1}{2} \sum_i \sum_{j \neq i} \iint \rho_i(\boldsymbol{r}) \frac{h_{\rm c}(\boldsymbol{r}, \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho_j(\boldsymbol{r}')\, {\rm d}\boldsymbol{r}\, {\rm d}\boldsymbol{r}' \qquad (2.31)$$

which is a sum of pairwise interactions centred at atomic sites.

In order to compute the exchange energy change, $\Delta E_{\rm x}[\rho^{\rm f}]$, we work with the Hartree–Fock approximation. That is, we assume that there exists an external potential in the Hartree–Fock equations such that the charge density, $\rho^{\rm HF}(\boldsymbol{r})$, constructed from the Hartree–Fock orbitals is equal to $\rho^{\rm f}(\boldsymbol{r})$. Then $E_{\rm x}[\rho^{\rm f}]$ is given by:

$$E_{\rm x}[\rho^{\rm f}] = -\tfrac{1}{2} \iint {\rm d}\boldsymbol{r}\, {\rm d}\boldsymbol{r}'\, \frac{\rho^{\rm HF}(\boldsymbol{r}, \boldsymbol{r}') \rho^{\rm HF}(\boldsymbol{r}', \boldsymbol{r})}{|\boldsymbol{r} - \boldsymbol{r}'|}. \qquad (2.32)$$

Since $\rho^{\rm HF}(\boldsymbol{r})$ is equal to $\rho^{\rm f}(\boldsymbol{r})$ by construction we approximate $\rho^{\rm HF}(\boldsymbol{r}, \boldsymbol{r}')$ by $\rho^{\rm f}(\boldsymbol{r}, \boldsymbol{r}')$. Then $E_{\rm x}[\rho^{\rm f}]$ simplifies to a sum of on-site terms and pair interactions and the on-site terms are cancelled exactly in the difference $\Delta E_{\rm x}[\rho^{\rm f}]$.

$$\Delta E_{\rm x} = -\tfrac{1}{2} \sum_{\substack{i\alpha\, j\beta \\ i \neq j}} (\rho^{\rm f})^{i\alpha i\alpha} (\rho^{\rm f})^{j\beta j\beta} (i\alpha j\beta | j\beta i\alpha) \qquad (2.33)$$

where

$$(i\alpha m\gamma | l\mu j\beta) = \iint {\rm d}\boldsymbol{r}\, {\rm d}\boldsymbol{r}'\, \varphi_{i\alpha}^*(\boldsymbol{r}) \varphi_{m\gamma}(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \varphi_{l\mu}^*(\boldsymbol{r}') \varphi_{j\beta}(\boldsymbol{r}'). \qquad (2.34)$$

We conclude that $\Delta E_{\rm es} + \Delta E_{\rm xc}$ may be approximated by a sum of pair interactions, centred at atomic sites. It follows that the binding energy of the solid, equation (2.11), may be expressed as a sum of bond energies and promotion energies. Each bond energy

is the sum of the covalent energy of the bond and the pair potential interaction arising from $\Delta E_{es} + \Delta E_{xc}$.

2.4. Interatomic forces

The force acting on the kth atom in the x direction can be obtained by differentiating the binding energy with respect to the x-coordinate of atom k, x_k.

$$\frac{\partial E_B}{\partial x_k} = \frac{\partial}{\partial x_k} \operatorname{Tr} \rho^{out} \tilde{H} - \frac{\partial}{\partial x_k} \operatorname{Tr} \rho^f \tilde{H} + \frac{\partial}{\partial x_k} (\Delta E_{es} + \Delta E_{xc}). \tag{2.35}$$

Since we have approximated $\Delta E_{es} + \Delta E_{xc}$ by a sum of pair potentials its derivative is trivial once the pair potential is known. To evaluate $(\partial/\partial x_k) \operatorname{Tr} \rho^f \tilde{H}$ it is advantageous to express the trace as

$$\frac{\partial}{\partial x_k} \operatorname{Tr} \rho^f \tilde{H} = \sum_{i\alpha} \frac{\partial (\rho^f)^{i\alpha i\alpha}}{\partial x_k} \tilde{H}_{i\alpha i\alpha} + (\rho^f)^{i\alpha i\alpha} \frac{\partial \tilde{H}_{i\alpha i\alpha}}{\partial x_k}. \tag{2.36}$$

Since the expansion coefficients, $(\rho^f)^{i\alpha i\alpha}$, are constants the first term is zero. To evaluate the derivative of the band energy, the first term in (2.35), we exploit the relation between the density matrix and the Green function given in (2.28), and the Hellmann–Feynman theorem. Thus,

$$\frac{\partial}{\partial x_k} \operatorname{Tr} \rho^{out} \tilde{H} = \tfrac{1}{2} \sum_{i\alpha} \sum_{j\beta} 2(\rho^{out})^{i\alpha}{}_{j\beta} \frac{\partial \tilde{H}^{j\beta}{}_{i\alpha}}{\partial x_k} = \frac{-2 \operatorname{Im}}{\pi} \sum_{i\alpha} \sum_{j\beta} \int^{E_F} G^{i\alpha}{}_{j\beta}(E^+) \, dE \, \frac{\partial \tilde{H}^{j\beta}{}_{i\alpha}}{\partial x_k} \tag{2.37}$$

or, using (2.19) and $(-\operatorname{Im}/\pi)(\Sigma_{i\alpha} H_{p\mu i\alpha} G^{i\alpha l\gamma}) = (-\operatorname{Im}/\pi) E G_{p\mu}{}^{l\gamma}$,

$$\frac{\partial}{\partial x_k} \operatorname{Tr} \rho^{out} \tilde{H} = -\frac{-2 \operatorname{Im}}{\pi} \sum_{i\alpha} \sum_{j\beta} \sum_{l\gamma} \sum_{p\mu} \int^{E_F} G^{i\alpha l\gamma} S_{l\gamma j\beta} \frac{\partial}{\partial x_k} [(S^{-1})^{j\beta p\mu} \tilde{H}_{p\mu i\alpha}] \, dE$$

$$= \frac{-2 \operatorname{Im}}{\pi} \sum_{i\alpha} \sum_{j\beta} \left(\int^{E_F} G^{i\alpha j\beta} \frac{\partial \tilde{H}_{j\beta i\alpha}}{\partial x_k} - \sum_{p\mu} E(S^{-1})^{i\alpha p\mu} G_{p\mu}{}^{j\beta} \frac{\partial S_{j\beta i\alpha}}{\partial x_k} \right) dE$$

$$= \frac{-2 \operatorname{Im}}{\pi} \sum_{i\alpha} \sum_{j\beta} \int^{E_F} G^{i\alpha j\beta} \left(\frac{\partial \tilde{H}_{j\beta i\alpha}}{\partial x_k} - E \frac{\partial S_{j\beta i\alpha}}{\partial x_k} \right) dE. \tag{2.38}$$

Equation (2.38) has also been derived by Foulkes (1987). The first expression, equation (2.37) is more compact than the second, although the derivative $\partial \tilde{H}^{j\beta}{}_{i\alpha}/\partial x_k$ contains the dependence of the overlap matrix **S** on x_k because

$$\tilde{H}^{j\beta}{}_{i\alpha} = \sum_{l\gamma} (S^{-1})^{j\beta l\gamma} \tilde{H}_{l\gamma i\alpha}. \tag{2.39}$$

If the atomic orbitals were an orthonormal set then there would be no distinction between the direct and dual bases and both (2.37) and (2.38) would become

$$\frac{\partial}{\partial x_k} \operatorname{Tr} \rho^{out} \tilde{H} = \tfrac{1}{2} \sum_{i\alpha} \sum_{j\beta} 2(\rho^{out})_{i\alpha j\beta} \frac{\partial \tilde{H}_{j\beta i\alpha}}{\partial x_k}. \tag{2.40}$$

This expression has been derived for orthonormal orbitals by several authors (see, for example, Moraitis and Gautier 1979, Chadi 1984, Sankey and Allen 1986, Pollmann *et al* 1986). To our knowledge the first derivation of (2.40) was in Coulson (1939).

Thus the derivative of the binding energy with respect to x_k, equation (2.35), can be written as

$$\frac{\partial E_B}{\partial x_k} = \frac{1}{2} \sum_{i\alpha} \sum_{j\beta} 2(\rho^{\text{out}})^{i\alpha}{}_{j\beta} \frac{\partial \bar{H}^{j\beta}{}_{i\alpha}}{\partial x_k} - \sum_{i\alpha} (\rho^{\text{f}})^{i\alpha i\alpha} \frac{\partial \bar{H}_{i\alpha i\alpha}}{\partial x_k} + \frac{\partial}{\partial x_k}(\Delta E_{\text{es}} + \Delta E_{\text{xc}}). \qquad (2.41)$$

Equation (2.41) can be simplified considerably by the following four approximations which lead to a physically transparent and readily computable expression, below, for the force. These approximations are consistent with the neglect of second-order contributions to the energy thus far.

(i) The non-orthogonality of the atomic-orbital basis set may often be neglected because the leading correction terms to the energy are second order in the overlap matrix (Pettifor 1977).

(ii) Three-centre terms may be neglected because the leading three-centre corrections to the energy are also of second order (Pettifor 1977).

(iii) Assume $\bar{H}_{i\alpha i\beta} = \bar{H}_{i\alpha i\alpha} \delta_{i\beta}^{i\alpha}$, i.e. the Hamiltonian matrix elements between different orbitals on the same atom may be neglected.

(iv) Each atom may be assumed to remain charge neutral by varying the on-site Hamiltonian matrix elements in such a way that the energy splittings between different orbitals on the same atom are preserved. This approximation ensures that contributions to the force in (2.41) from on-site terms in the first sum cancel those of the second sum, which leads to consistency with the force theorem (Pettifor 1976, 1978, Mackintosh and Andersen 1980).

With these approximations the derivative of the binding energy simplifies to

$$\frac{\partial E_B}{\partial x_k} = \sum_{j \neq k} \sum_{\alpha\beta} 2(\rho^{\text{out}})_{k\alpha j\beta} \frac{\partial \bar{H}_{j\beta k\alpha}}{\partial x_k} + \frac{\partial}{\partial x_k}(\Delta E_{\text{es}} + \Delta E_{\text{xc}}). \qquad (2.42)$$

The approximation of local charge neutrality when self-consistency is achieved is motivated by four considerations.

(i) In a metal where there is perfect screening any excess charge associated with an atom will be neutralised; similarly in a semiconductor where the band gap is much smaller than the widths of the valence and conduction bands the screening length is approximately the same as an inter-atomic separation and hence local charge neutrality will be a good approximation (Priester *et al* 1986).

(ii) In order to get the same answer for the bulk modulus of the solid when it is calculated by the method of long waves or by a homogeneous dilation it is essential that the charge density is treated within a self-consistent scheme and local charge neutrality is the simplest possible assumption. This point is discussed in § 3.

(iii) In binary systems there would be important long-range Coulomb terms if local charge neutrality were not required. As discussed by Pettifor and Podloucky (1986) it is almost always possible to find a localised orbital basis with respect to which each atom is neutral.

(iv) Charge neutrality within the TBB model leads to an internally consistent picture of the heats of formation of transition metal alloys (Pettifor 1987).

In the TBB model local charge neutrality is achieved by varying only on-site Hamiltonian matrix elements. This is an approximation since there will also be variations in the inter-site Hamiltonian matrix elements as the electronic charge is redistributed throughout the system (see Lannoo and Bourgoin 1981). Only the inter-site Hamiltonian

matrix elements in the TBB model depend explicitly on atomic positions. The explicit dependence of on-site Hamiltonian matrix elements on the positions of neighbouring atoms appears in our model in the pair potential because the on-site crystal-field terms contribute for example to ΔE_{es}. Thus, the raising of the d band centre of gravity in transition metals under compression is contained within the pair potential. Furthermore the requirement of local charge neutrality results in the on-site Hamiltonian matrix elements depending (in an unspecified way) on the positions of neighbouring non-equivalent atoms. This latter dependence is ignored in most tight-binding formulations, (for example, Chadi 1979, 1984), where the diagonal elements of $\bar{H}$ are not allowed to vary from one site to another in the solid. In that case we note that, since the total energy is simply the sum of the binding energy of the solid and the energies of the free atoms, the total energy can be represented as the band energy plus the same sum of pairwise interactions as in the TBB model plus a constant. However we emphasise that this is true only if the diagonal Hamiltonian matrix elements are not allowed to vary from one non-equivalent site to another. As soon as variations in the diagonal elements are allowed the 'constant' in the total energy becomes a site-dependent energy.

The energy splittings of on-site Hamiltonian matrix elements are assumed to be constant at each atomic site. Thus all of the diagonal Hamiltonian matrix elements at a particular site change by the same amount, $\delta \bar{H}_i$, although this amount will vary from site to site. It follows from this that there are no contributions to the force from the first two terms of (2.41) from on-site Hamiltonian matrix elements once local charge neutrality has been achieved. Thus the TBB model is consistent with the force theorem (Pettifor 1976, 1978, Mackintosh and Andersen 1980) in the approximation that self-consistency is adequately described by local charge neutrality. This theorem states that there is no contribution to the force on an atom arising from the self-consistent redistribution of electronic charge concomitant with a virtual atomic displacement. As has already been noted by several authors (Coulson 1939, Moraitis and Gautier 1979, Chadi 1984, Sankey and Allen 1986, Pollmann *et al* 1986) the first term in (2.42) has a very appealing structure. The contribution to the force from the covalent bond energy is broken down into contributions from individual bonds to an atom. Each covalent bond force consists of twice the bond order multiplied by the derivative of the corresponding Hamiltonian matrix element. Whereas the inter-site Hamiltonian matrix elements depend only on the relative position vector of the two bonded atoms the bond order is determined by all (inter-site and on-site) Hamiltonian matrix elements in the local atomic environment. Hence these covalent bond forces are not simply two-body interactions even though the total force on an atom can be expressed as a sum of forces from each neighbouring atom. The inter-site Hamiltonian matrix elements vary with the length, d, and direction of the relative position vector of the two atoms. The length dependence can be taken from canonical band theory, e.g. d^{-5} for d–d interactions in transition metals. The range of the Hamiltonian matrix elements determines the number of neighbours that contribute to the force on an atom. However, more distant atoms affect the force because they also influence the neighbouring bond orders. The angular dependence of the Hamiltonian matrix elements is taken from the tables in Slater and Koster (1954). The bond orders may be computed by diagonalising the Hamiltonian (see, for example, Chandi 1979, 1984) or by using the recursion method to calculate Green function matrix elements and (2.28) (Ohta *et al* 1987a, Paxton and Sutton 1987). The main advantage of the recursion method is that it is a real-space method and the relative influences of successive neighbour shells can be studied systematically.

The physical origin of the pair potential in (2.11) is the classical electrostatic inter-

action energy between neutral pseudo-atoms modified by the exchange–correlation energy difference $\Delta E_{xc}[\rho^f]$. Therefore it is distinct from the 'overlap interaction' which arises from non-orthogonality of the atomic orbitals (Slater 1968). In (2.11) the latter contribution to the binding energy appears in the covalent bond and promotion energies. If the covalent bond and promotion energies in the absence of overlap (i.e. $S = I$) are $\mathrm{Tr}(\rho^{out} - \rho^f)\tilde{H}$ then the first-order correction arising from $S = I + \Delta S$ is $-\mathrm{Tr}(\rho^{out} - \rho^f)\tilde{H}\Delta S$. This correction term consists of two- and three-centre interactions. On the reasonable assumption that bond orders and hopping integrals are of the same order as off-diagonal elements of the overlap matrix it is seen that the leading corrections arising from ΔS are the two-centre interactions which are second order in ΔS (see also Majewski and Vogl 1986). In some semi-empirical models (see, for example, Harrison 1985) the correction due to overlap appears as a repulsive pair potential rather than as a positive contribution to the bond and promotion energies. Strictly speaking, however, the overlap interaction is not representable as a sum of pair potentials in a defective crystal, because ρ^{out} is dependent on the local environment.

2.5. Atomic stress tensor

It is possible to calculate the local stress tensor at each atomic site by applying a virtual scaling of all inter-atomic distances. The local stress tensor is defined by the first-order response of the system to this scaling. The derivation follows exactly the same steps as that given in Born and Huang (1954) for inter-atomic forces described by pair potentials. It will be noted that there are also similarities to the derivation of the quantum mechanical stress theorem of Nielsen and Martin (1983, 1985a, 1985b) except that the stress tensor below cannot be defined at positions other than atomic sites. Satoko (1987) has also derived the atomic stress tensor in the tight-binding approximation and used it to study the stabilities of clusters of Si atoms.

Consider an assembly of atoms with arbitrary atomic positions for which the binding energy is given by (2.11). Apply a homogeneous virtual pure strain to all atoms in the assembly and let the I–J component of the virtual strain tensor be ε_{IJ} where I and J are Cartesian components x, y, or z. To first order in the strain tensor the change in energy of the system is equal to

$$\delta E^{(1)} = \sum_i \sum_{I,J} \Omega^i \sigma^i_{IJ} \varepsilon_{IJ} \tag{2.43}$$

where Ω^i is the volume associated with the atom at site i. The I–J component of the stress tensor at site i is represented in (2.43) by σ^i_{IJ}. If we make the same four approximations as in § 2.4 we obtain the following contribution to the stress tensor σ^i_{IJ} from the bond and promotion energies.

$$\frac{1}{\Omega^i} \sum_{j \neq i} \sum_{\alpha\beta} (\rho^{out})_{i\alpha j\beta} \frac{\partial \tilde{H}_{j\beta i\alpha}}{\partial R_{jI}} (R_{jJ} - R_{iJ}) \tag{2.44}$$

where R_{iJ} is the J-component of the position vector of the ith atom. The total stress tensor is obtained by adding the contribution from the pair potential which has a well known form (see, for example, Basinski *et al* 1971).

The form of the stress tensor contribution in (2.44) is very similar to the form for a pair potential interaction, namely it consists of moments of the forces exerted by neighbouring atoms on the central atom. Again we emphasise that in the absence of charge neutrality further contributions to the stress in (2.44) would arise from on-site

Hamiltonian matrix elements. It is anticipated that this stress tensor will be useful in characterising defects in the same way as the analogous stress tensor for pair potentials has been applied (see, for example, Egami and Vitek 1983).

3. The energy change to second order

It is well known (see, for example, Born and Huang 1954) that the elastic constants of a solid may either be obtained by calculating the change in the energy of the solid to second order in a homogeneous infinitesimal strain or from the long-wavelength limit of the phonon dispersion relations (the 'method of long waves' (Born and Huang 1954)). Let us consider the underlying physics of these two methods. In the case of a homogeneous strain applied to a perfect crystal with a monatomic basis all atoms remain equivalent and therefore there is no charge transfer from one atom to another. However, if there is a volume change associated with the strain the Fermi level changes and this gives rise to an important change in the energy which is second order in the strain (Heine *et al* 1980, p 97). On the other hand the passage of a phonon through a crystal destroys the equivalence of atomic sites and charge flows in such a way that the Fermi level remains constant at all times. As the charge flows the effective one-electron potential changes at each atom so as to prevent further charge flow in accordance with Le Chatelier's principle. This screening of the charge redistribution caused by the phonon can strongly affect the calculated energy of the phonon. It is clear from this discussion that self-consistency is crucial in a calculation of force constants. We shall show that neglect of self-consistency in the band model leads to an inconsistency in the bulk modulus computed by these two methods. By contrast the TBB model does not suffer from this inconsistency because of the requirement of local charge neutrality.

As mentioned in § 2.1 the total energy of a solid in the tight-binding band model consists of the band energy and a sum of repulsive pair potentials. Self-consistency is not normally a feature of the model. Let us evaluate the change in the energy of a solid to second order in a change in the Hamiltonian ΔH caused by some as yet unspecified atomic displacements. The change in the band energy is given by

$$
\Delta E_{\text{band}} = \Delta \int^{E_F} EN(E)\, dE
$$

$$
= \int^{E_F} E\,\Delta N(E)\, dE + \int_{E_F}^{E_F + \Delta E_F} E(N + \Delta N)\, dE \tag{3.1}
$$

where $N(E)$ is the total density of states. The condition of total charge conservation gives

$$
\int^{E_F} \Delta N\, dE + \int_{E_F}^{E_F + \Delta E_F} (N + \Delta N)\, dE = 0 \tag{3.2}
$$

and when this is inserted in (3.1) we obtain

$$
\Delta E_{\text{band}} = \int^{E_F} (E - E_F)\Delta N\, dE + \int_{E_F}^{E_F + \Delta E_F} (E - E_F)(N + \Delta N)\, dE. \tag{3.3}
$$

In fact (3.3) is exact regardless of the size of ΔH. Expanding the second term and

retaining only second-order terms we obtain

$$\Delta E_{\text{band}} = \int^{E_F} (E - E_F)\Delta N \, dE + \tfrac{1}{2}N(E_F)(\Delta E_F)^2. \tag{3.4}$$

We note that the first term in (3.3) is not merely a first-order term but is exact to all orders; in particular it includes terms of second order in the atomic displacements that correspond to force constants. The change in the Fermi energy, ΔE_F, follows from (3.2)

$$\Delta E_F = -\frac{1}{N(E_F)}\int^{E_F} \Delta N \, dE. \tag{3.5}$$

The calculation of ΔE_{band} due to a homogeneous dilatation entails changes in ΔN and ΔE_F that contribute to both terms in (3.4). The calculation of the second-order change in the band energy using the long-wavelength limit of the phonon dispersion relations in the band model would be based on the assumptions that the Fermi level is unchanged and that there is no self-consistent redistribution of charge. Therefore, unlike the TBB model, the diagonal elements of the Hamiltonian would be assumed to be constant. Thus only the first term of (3.4) would be evaluated and hence the two methods of calculating the change of the band energy do not agree. The second term of (3.4) may be regarded as arising from the condition of charge conservation which is violated by the long-wavelength phonon calculation in the band model. The resolution of the inconsistency entails the recognition of the self-consistent change of the electron potential which occurs so as to ensure a constant Fermi level throughout the system. In the TBB model this is done by varying the diagonal elements of the Hamiltonian in such a way as to ensure local charge neutrality. However it is not trivial to evaluate contributions to force constants due to changes in the diagonal Hamiltonain matrix elements because this would involve the inversion of large response matrices and further work is required to find appropriate simplifications.

To illustrate the inconsistency in the band model we have evaluated the bulk modulus within a canonical d-band model. The overlap parameters vary as the inverse fifth power of the inter-atomic distance R

$$dd\sigma = -6W\left(\frac{2}{5}\right)\left(\frac{S}{R}\right)^5$$

$$dd\pi = 4W\left(\frac{2}{5}\right)\left(\frac{S}{R}\right)^5$$

$$dd\delta = -1W\left(\frac{2}{5}\right)\left(\frac{S}{R}\right)^5. \tag{3.6}$$

S is fixed the Wigner–Seitz radius and W is an approximate band width. When a crystal is dilated homogeneously, the band energy within this model simply scales as $(S/R)^5$, as does the Fermi energy, measured from the centre of the d band, ε_d. In both the band and the TBB models ε_d remains constant because there is no charge redistribution. ε_d is chosen as the zero of energy. It is therefore straightforward to derive expressions for the pressure and bulk modulus in terms of the band energy

$$P = \tfrac{5}{3}E_{\text{band}}/V$$

$$B = \tfrac{40}{9}E_{\text{band}}/V. \tag{3.7}$$

V is the total volume for which the band energy is defined. Referring to (3.4) we can

consider B as the sum of two parts

$$B = B^{(1)} + B^{(2)} \tag{3.8}$$

corresponding to the first and second terms of (3.4). The second term of (3.4) gives the contribution

$$B^{(2)} = \frac{25}{9} \frac{N(E_F)E_F^2}{V}. \tag{3.9}$$

The contribution (3.9) is automatically included in B as given by the expression in (3.7). For a half-filled band $E_F = 0$ and this contribution vanishes, but in general it is not negligible.

We have evaluated the first term $B^{(1)}$ by calculating the density of states for a BCC crystal using the recursion method to 15 exact levels with the termination of Beer and Pettifor (1984). The integrations to the Fermi energy were carried out to high accuracy by quadratures on a rectangular contour in the upper half of the complex plane, a procedure we adopted mainly to avoid the singular behaviour of the density of states at the lower band edge. The density of states was calculated for the perfect crystal at three volumes, with dilations of 0, $+0.02$ and -0.02. The contributions $B^{(1)}$ and $B^{(2)}$ from the second derivatives of the term in (3.4) were then evaluated by numerical differencing. The resulting values of $B^{(1)}$ and $B^{(2)}$ are added together and compared with B according to (3.7) in figure 1. Also shown in figure 1 is the bulk modulus as obtained by a force constant summation out to ten shells of neighbours, where the force constants were obtained as in Finnis *et al* (1984). As anticipated, the force constant result corresponds to $B^{(1)}$ only. The deviations in the force constant result are due to the truncation to tenth neighbours together with the restriction to 15 levels in the continued fraction representation of the Green functions that enter the expressions for the force constants.

It is seen that the condition of charge conservation, expressed by the final term of (3.4), makes a very significant contribution to the bulk modulus, away from the centre
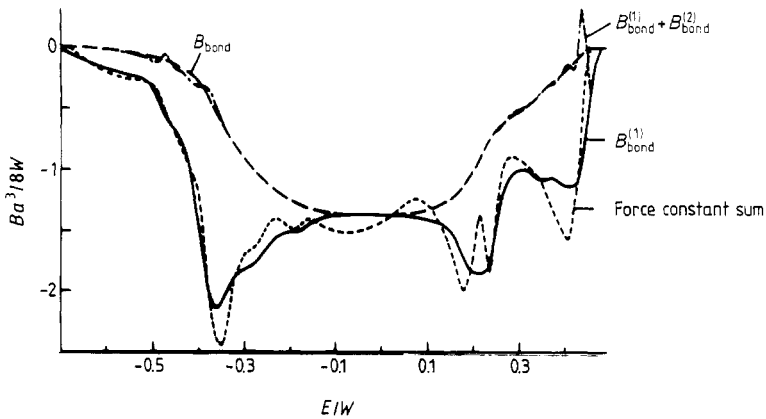

Figure 1. Bulk modulus in a BCC crystal. The bond energy contributions $B_{bond}^{(1)}$ and $B_{bond}^{(2)}$ calculated by homogeneous deformation and second differencing, compared to their theoretical total $B_{bond} = (40/9)U_{bond}/V$, for a BCC crystal of lattice parameter a with canonical tight-binding parameters. Also shown is the result of calculating B from the force constants out of ten shells of neighbours.

of the d band. By inference charge conservation, which is violated in the band model, is very significant in the calculation of tight-binding force constants. It is also very important in the calculation of inter-atomic forces where the first-order change in the band energy is not given by

$$\Delta E_{band} = \int^{E_F} E \, \Delta N(E) \, \mathrm{d}E \tag{3.10}$$

which violates charge conservation but by

$$\Delta E_{band} = \int^{E_F} (E - E_F) \Delta N(E) \, \mathrm{d}E. \tag{3.11}$$

Equation (3.11) is the first-order term of (3.4). It can be shown that ΔE_{band} is also equal to $\mathrm{Tr}\,\rho\Delta H$ for an orthonormal set. The requirement of local charge neutrality in the TBB model is the simplest condition that fulfils charge conservation.

4. Exact results for linear chain models

In this section we shall illustrate the use of (2.42) to predict relaxations in simple analytic models of free surfaces (§ 4.1) and grain boundaries (§ 4.2). In addition, in § 4.3, we present a study of the size dependence of self-consistent and non-self-consistent cluster calculations. The idea that interface problems may be reduced to linear chains is not new (see, for example, Lannoo 1980) and Pollmann and Pantelides (1978, 1980) have exploited this idea very effectively in their tight-binding studies of ideal heterojunctions. Each 'atom' in the chain corresponds to a layer of atoms parallel to the interface. For simplicity we associate only one atomic orbital with each atom in the three-dimensional bicrystal and we assume these orbitals form a complete orthonormal set. Within each layer one can construct two-dimensional Bloch sums of atomic orbitals, called 'layer orbitals', parallel to the interface. Each layer orbital is characterised by a wavevector, q, in the two-dimensional Brillouin zone of the interface, and an index denoting the layer it is associated with. Layer orbitals with different q-vectors or different layer indices are orthogonal to each other. Thus for a given q-vector an interface problem reduces to a linear chain of orthonormal layer orbitals. Conversely, solving the Hamiltonian of a linear chain of orthonormal atomic orbitals is equivalent to solving the Hamiltonian of a three-dimensional bicrystal at a particular q-vector. Thus, the simple one-dimensional models presented in this section are not unrelated to three-dimensional interface problems. At the same time we believe they contain some essential physics which, in our opinion, is overlooked by recent cluster calculations of grain boundaries (see, for example, Eberhart *et al* 1983, Briant and Messmer 1982).

4.1. The constant semi-infinite linear chain

In this section a constant semi-infinite linear chain of atoms is treated as a model of a free surface. Each atom in the chain is associated with a single atomic-like orbital and these orbitals are assumed to form a complete orthonormal set. In this simple model all on-site Hamiltonian matrix elements have the same, constant value which can be set equal to zero. The only non-zero off-diagonal Hamiltonian matrix elements are assumed to be between nearest neighbours and all such matrix elements are set equal to γ. This is not the self-consistent model discussed in § 2.4 but (2.42) still applies because the

diagonal Hamiltonian matrix elements in (2.41) are not allowed to vary in this model. The question of how much difference it makes to the bond orders appearing in (2.42) by requiring charge neutrality is addressed in § 4.3. It is possible to compute all Green function matrix elements for this model analytically and the derivation is given in detail in the Appendix. This derivation does not involve any Brillouin zone integrations and all Green functions are derived by suitable choices of 'perturbed and unperturbed' Hamiltonians and Dyson's equation (Lannoo 1980). Successive atoms starting with the surface atom are labelled 1, 2, 3, ... Since there is only one orbital associated with each atom each orbital may be referred to by a single (roman) subscript. Thus the Hamiltonian is defined by

$$H_{ii} = \quad 0 \qquad \text{for all } i$$

$$H_{ij} = \begin{cases} \gamma & \text{if } i \text{ and } j \text{ are nearest neighbours} \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

In the Appendix the Green function matrix element G_{jk} where $j \le k$ is shown to be

$$G_{ik}(E^+) = \frac{1}{\gamma} \frac{\sin j\theta}{\sin \theta} e^{ik\theta} \tag{4.2}$$

where

$$\cos \theta = E/2\gamma \tag{4.3}$$

and

$$G_{kj} = G_{jk}. \tag{4.4}$$

Using (2.28) the bond order ρ_{jk} can be readily computed from G_{jk}. In particular $\rho_{k,k+1}$ is given by

$$\rho_{k,k+1} = \frac{2\theta_F}{\pi} \left(\frac{\sin \theta_F}{\theta_F} - \frac{\sin(2k+1)\theta_F}{(2k+1)\theta_F} \right). \tag{4.5}$$

Inserting (4.5) into (2.42) gives the covalent bonding contribution to the force, f_k, on atom k

$$f_k = \frac{4\theta_F}{\pi} \left(\frac{\sin|2k-1|\theta_F}{|2k-1|\theta_F} - \frac{\sin(2k+1)\theta_F}{(2k+1)\theta_F} \right) \frac{\partial H_{k-1,k}}{\partial x_k} \tag{4.6}$$

where the Fermi wavevector θ_F is defined by

$$\cos \theta_F = E_F/2\gamma. \tag{4.7}$$

f_k is the covalent bonding contribution to the force which exists prior to any relaxation of atomic positions along the chain. Because of this we can write

$$\frac{\partial H_{k,k-1}}{\partial x_k} = - \frac{\partial H_{k,k+1}}{\partial x_k} \tag{4.8}$$

which will not be true once atoms start to be displaced. The total force acting on atom k is the force f_k and the force arising from the pair potential. The latter is also constant before atoms are displaced from their ideal constant spacing. Thus, apart from the surface atom, the initial total force acting on each atom (i.e. prior to any relaxation) is determined solely by the covalent bonding contribution. The total initial force on

the surface atom is determined by both the pair potential and the covalent bonding contribution. Equation (4.6) indicates that the initial total force on successive atoms causes Friedel oscillations in the atomic spacings. This is particularly clear for large values of k where f_k becomes

$$f_k \simeq -\frac{8\theta_F}{\pi} \sin \theta_F \frac{\cos 2k\theta_F}{2k\theta_F} \frac{\partial H_{k,k-1}}{\partial x_k}. \tag{4.9}$$

It is seen that the magnitude and sign of the oscillations depend on the band filling and the magnitude is greatest in a half-filled band. However, according to this model the charge on successive atoms also shows Friedel oscillations

$$\rho_{kk} = \frac{2\theta_F}{\pi}\left(1 - \frac{\sin 2k\theta_F}{2k\theta_F}\right). \tag{4.10}$$

Here $2\theta_F/\pi$ is the charge associated with each atom of an infinite linear chain. The Friedel oscillation is due entirely to the free surface. In § 4.3 we consider the effect of requiring local charge neutrality at all atoms in the chain on the bond orders giving rise to the Friedel oscillations in the inter-atomic spacings.

These results are strikingly similar to those obtained for a one-dimensional jellium model (Garcia-Moliner and Flores 1979). For a semi-infinite one-dimensional electron gas the density matrix is given by

$$\rho(x, x') = \frac{2\theta_F}{\pi}\left(\frac{\sin \theta_F(x - x')}{(x - x')\theta_F} - \frac{\sin \theta_F(x + x')}{(x + x')\theta_F}\right) \tag{4.11}$$

and the charge density by

$$\rho(x) = \rho(x, x) = \frac{2\theta_F}{\pi}\left(1 - \frac{\sin 2\theta_F x}{2\theta_F x}\right) \tag{4.12}$$

where $E_F = \theta_F^2$ is the dispersion relation between the free-electron Fermi energy and wavevector. Substituting $x = k + 1$ and $x' = k$ in (4.11) and (4.12) it is seen that they are identical to (4.5) and (4.10). This result may be surprising in view of the differences in the tight-binding and free-electron approximations. However, further analysis shows that this result does not hold in two or three dimensions because the Fermi surfaces in the free-electron and tight-binding models are not identical in two- or three-dimensional space. In one dimension the Fermi 'surface' is simply a pair of points along the wavevector in both models although the Fermi energies differ because of the different dispersion relations. Since density matrix elements can be expressed as integrations over occupied states in k-space the differing Fermi surfaces in two or three dimensions will give rise to differing density matrices even for an s-band tight-binding model.

Green function matrix elements for the infinite linear chain are easily obtained from (4.2) by 'bonding' two semi-infinite linear chains together and using Dyson's equation. Thus we obtain the well known result (see, for example, Economou 1983) for an infinite linear chain

$$G_{jk} = \frac{i e^{i|j-k|\theta}}{2\gamma \sin \theta}. \tag{4.13}$$

Two sum rules for Green function matrix elements are readily illustrated with these Green functions.

(i) The band energy may be expressed in terms of site-diagonal Green functions only

or in terms of both site-diagonal and off-diagonal Green functions

$$E_{\text{band}} = \frac{-2\,\text{Im}}{\pi} \int_{2\gamma}^{E_F} \sum_i G_{ii}(E^+)E\,\mathrm{d}E = \frac{-2\,\text{Im}}{\pi} \int_{2\gamma}^{E_F} \sum_{ij} G_{ij}(E^+)H_{ji}\,\mathrm{d}E \tag{4.14}$$

which for (4.13) becomes

$$\frac{4\gamma \sin\theta_F}{\pi} = \frac{2}{\pi} \sum_k \frac{\sin|j-k|\theta_F}{|j-k|} H_{kj}. \tag{4.15}$$

Using (4.1) it is readily verified that (4.15) holds.

(ii) Following Terakura *et al* (1982) we may write

$$\mathrm{d}G/\mathrm{d}E = -G^2 \tag{4.16}$$

and integration yields

$$\rho_{jk}(E^+) = \frac{-2\,\text{Im}}{\pi} G_{jk}(E^+) = \frac{2\,\text{Im}}{\pi} \int_{2\gamma}^E \sum_l G_{jl}(E'^+)G_{lk}(E'^+)\,\mathrm{d}E'. \tag{4.17}$$

That is, the spectral density matrix (Baraff and Schluter 1979), whose site-diagonal terms are simply the local densities of states, is related to a sum of response functions. A response function is defined by

$$\chi_{mnpq}(E^+) = \frac{-2\,\text{Im}}{\pi} \int_{2\gamma}^E G_{mn}(E'^+)G_{pq}(E'^+)\,\mathrm{d}E'. \tag{4.18}$$

This response function determines the first-order change in the density matrix element, ρ_{mq}, when the Hamiltonian matrix element H_{np} is varied, i.e.

$$\delta\rho_{mq} = \chi_{mnpq}\delta H_{np} + O(\delta H_{np}^2). \tag{4.19}$$

It may be one, two, three or four centre (Finnis *et al* 1984). Using (4.13) and (4.18) the response function for the infinite linear chain is found to be

$$\chi_{mnpq}(2\gamma\cos\theta) = \frac{-1}{\pi\gamma} \int_0^\theta U_{z-1}(\cos\theta')\,\mathrm{d}\theta' \tag{4.20}$$

where $z = |m-n| + |p-q|$ and U_{z-1} is the Chebyshev polynomial of the second kind of order $z-1$

$$U_{z-1}(\cos\theta) = \sin z\theta/\sin\theta. \tag{4.21}$$

Thus the second sum rule, (4.17), becomes

$$\frac{-2\cos|j-k|\theta}{\sin\theta} = 2 \sum_{l=-\infty}^\infty \int_0^\theta \frac{\sin(|j-l|+|l-k|)\theta'\,\mathrm{d}\theta'}{\sin\theta'}. \tag{4.22}$$

By differentiating both sides of (4.22) with respect to θ and evaluating the resulting sum on the right-hand side it can be shown that the left- and right-hand sides of (4.22) differ at most by a constant.

4.2. Linear chain model of a grain boundary

In this one-dimensional model the grain boundary is represented by a single atom whose on-site Hamiltonian matrix element differs from all the other on-site Hamiltonian matrix

elements of an otherwise constant infinite linear chain. The Hamiltonian is therefore identical to that of the constant infinite linear chain, whose Green functions were given in (4.13), except the on-site Hamiltonian matrix element of atom number zero is set equal to U. The atoms are numbered from minus infinity to plus infinity. Thus the grain boundary is treated as a highly localised perturbation in the potential of the bicrystal and it has been assumed for simplicity that the hopping integrals between the boundary atom and the adjacent atoms are the same as those of the bulk chain. The latter assumption can be relaxed easily but the results are somewhat more cumbersome to write down (Sutton 1987). The physical origin of U may be ascribed to various sources. For example it could be caused by a layer of segragated impurities or by charge flowing to or from the boundary atom due to an expansion (or contraction) at the boundary. A simple application of Dyson's equation yields any required Green function matrix element. The unperturbed Green functions are taken to be those of the constant infinite linear chain, (4.13). Equation (2.28) then gives the density matrix elements for the grain boundary model. The change in the bond order between atoms n and $n + 1$ is thus found to be

$$\Delta\rho_{n,n+1} = \frac{-2}{\pi} \int_0^{\theta_F} \frac{2U\gamma \sin\theta \sin(2n + 1)\theta + U^2 \cos(2n + 1)\theta}{U^2 + 4\gamma^2 \sin^2\theta} \, d\theta. \tag{4.23}$$

We have not been able to evaluate this integral analytically but two limiting cases are easily solved. For $U/2\gamma \ll 1$ we find

$$\Delta\rho_{n,n+1} = \frac{-U}{\pi\gamma} \int_0^{\theta_F} \frac{\sin(2n + 1)\theta}{\sin\theta} \, d\theta + O\left(\frac{U}{2\gamma}\right)^2. \tag{4.24}$$

Equation (4.24) is simply the result of linear response theory

$$\Delta\rho_{n,n+1} = \chi_{n+1,0,0,n} U + O(U/2\gamma)^2. \tag{4.25}$$

That is, the first-order change in each bond order is determined by the appropriate response function given in (4.20).

For $U/2\gamma \gg 1$ we find

$$\Delta\rho_{n,n+1} = \frac{-2\theta_F}{\pi} \frac{\sin(2n + 1)\theta_F}{(2n + 1)\theta_F} \tag{4.26}$$

and by comparing this with (4.5) we see that this limit simply decouples the infinite linear chain into two free surfaces.

Recent cluster calculations (Briant and Messmer 1982) of the effect of an impurity on the charge density distribution at a grain boundary have emphasised the importance of the electronegativity difference between the host and impurity atoms. A more electronegative impurity is supposed to attract electrons from adjacent bonds in the host and thereby weaken host bonds. In our model the electronegativity difference is related to the value of U. However, (4.23) indicates clearly that the charge redistribution d epends not only on U but also on the degree of band filling. For example, in (4.24) it is seen that for a given small value of U the change in the bond order $\rho_{n,n+1}$ oscillates with band filling. We may conclude that the electronegativity difference is not the only factor controlling charge redistribution: the response of the host to the impurity depends equally strongly on the band structure of the solid and the degree of band filling. Therefore the conclusions of the cluster calculations referred to above are reliable only if the clusters are able to emulate accurately the electronic response of the infinite host.

This has recently been emphasised in a three-dimensional model system by Saqi and Pettifor (1987).

4.3. The finite linear chain

In the solid state physcs literature there are countless examples of modelling infinite solids by finite clusters. All these calculations are subject to the question of how accurately converged the results are as a function of the cluster size. In this section we present first an exact analytic study of the size dependence of bond orders in finite linear chains as a function of chain length, from the diatomic molecule to the semi-infinite chain. In the second part we present a study in which we have required local charge neutrality to see how this affects the conclusions of the first part.

Consider a linear chain of N atoms. As before there is one basis function for each atom and they are assumed to form a complete orthonormal set. For the present we set all on-site Hamiltonian matrix elements to zero and only the nearest-neighbour off-diagonal Hamiltonian matrix elements are set equal to γ; the rest are set equal to zero. Following a similar sequence of steps to that described in the Appendix it can be shown that

$$G_{kl} = \frac{1}{\gamma} \frac{\sin(N - l + 1)\theta \sin k\theta}{\sin \theta \sin(N + 1)\theta} \tag{4.27}$$

where $l \geq k$, $G_{lk} = G_{kl}$, and $E = 2\gamma \cos \theta$.

The poles of G_{kl} give the eigenvalues of the finite chain. They are at

$$\theta = \theta_p = p\pi/(N + 1) \qquad \text{where } 1 \leq p \leq N. \tag{4.28}$$

The bond order ρ_{kl} is given by a sum of residues of G_{kl} evaluated at the occupied eigenvalues

$$\rho_{kl} = \frac{2}{(N + 1)\gamma} \sum_{\substack{p \\ \text{occupied}}} \left[\sin\left(\frac{lp\pi}{N + 1}\right) \sin\left(\frac{kp\pi}{N + 1}\right) \Big/ \sin\left(\frac{p\pi}{N + 1}\right) \right]. \tag{4.29}$$

By taking the limit $N \to \infty$ the sum in (4.29) becomes an integral in the continuous variable θ of (4.28). This integral may be evaluated using the dispersion relation $E = 2\gamma \cos \theta$ and it may be confirmed that it gives the same results as (4.5) and (4.10) for the semi-infinite chain. We have used (4.29) to calculate the bond order between the first and second atoms of finite chains of lengths ranging from 2 to 50 atoms at various average numbers, ν, of electrons per atom. The results for $\nu = 1$ are shown in table 1. $\nu = 1$

Table 1. Bond orders between first and second atoms of finite linear chains compared with the semi-infinite limit ρ_{12}^∞. Average number of electrons per atom $= 1$.

N	ρ_{12}^N	$\rho_{12}^N - \rho_{12}^\infty$
2	1.0000	0.1512
3	0.7072	-0.1416
4	0.8944	0.0456
5	0.7886	-0.0602
6	0.8712	0.0224
7	0.8154	-0.0334
8	0.8620	0.0132
9	0.8276	-0.0212
10	0.8576	0.0088

Table 2. Non-self-consistent and self-consistent bond orders in finite chains of various lengths. Atoms are numbered from 1 to N where N is the number of atoms in the chain. The bond order between atoms n and $n + 1$ is represented by $(n, n + 1)$; this bond order equals 0.4502 in an infinite chain. The self-consistent charge on each atom $= 0.5 +$ or -0.0001 electrons. For each chain length the non-self-consistent bond order is in the left column.

	Number of atoms in chain													
	5		7		10		11		13		19		50	
(1, 2)	0.3608	0.5000	0.2368	0.5000	0.3236	0.4985	0.2957	0.4991	0.3416	0.4962	0.3151	0.4971	0.3125	0.4946
(2, 3)	0.7217	0.5001	0.5716	0.5000	0.6397	0.4924	0.6230	0.4952	0.6088	0.4826	0.5996	0.4863	0.5576	0.4765
(3, 4)			0.8084	0.4999	0.6580	0.4773	0.7093	0.4853	0.5505	0.4563	0.5908	0.4636	0.5205	0.4484
(4, 5)					0.4435	0.4501	0.5314	0.4644	0.3846	0.4342	0.4078	0.4377	0.3963	0.4367
(5, 6)					0.3130	0.4313	0.3220	0.4376	0.4100	0.4460	0.3483	0.4402	0.4149	0.4516
(6, 7)									0.5578	0.4660	0.4840	0.4620	0.5048	0.4600
(7, 8)											0.5996	0.4711	0.4933	0.4517
(8, 9)											0.5246	0.4607	0.4185	0.4430
(9, 10)											0.3757	0.4414	0.4246	0.4493

corresponds to a half-filled band in the limit of a semi-infinite chain. In this limit (4.5) indicates that $\rho_{12}^\infty = 0.8488$. It is seen that the value of ρ_{12}^N oscillates about ρ_{12}^∞ as N is increased but the magnitude of the oscillations diminishes with increasing N. We expect qualitatively similar results in two and three dimensions except the rate at which the oscillations decay with the number of atoms in the cluster should increase with increasing spatial dimensionality of the cluster.

The results described thus far have been derived for models that do not require local charge neutrality. As described in § 2.4 the requirement of local charge neutrality amounts to the introduction of a form of self-consistency in the tight-binding model. It is unlikely that the screening in linear molecules is perfect and therefore local charge neutrality can be regarded only as an approximation to the Kohn–Sham self-consistent charge distribution. Nevertheless it is very instructive to examine the effect that local charge neutrality has on the size dependence of cluster models and also on the picture of the bonding gained from such calculations. We have compared non-self-consistent and self-consistent bond orders in chains of length varying from 4 to 50 atoms for a variety of values of ν. In table 2 we compare successive nearest-neighbour bond orders for $\nu = 0.5$, for which the bond order in an infinite chain is 0.4502. For a given bond order, e.g. ρ_{23}, it is seen that the self-consistent solution oscillates with chain length much less than for the non-self-consistent solution. This shows that the size dependence of cluster calculations is much reduced by going to self-consistency. Secondly the Friedel oscillations in successive bond orders of a given chain are strongly damped in the self-consistent case. Therefore the oscillations in successive interatomic spacings predicted by (4.6) are much reduced by self-consistency. Thirdly for all chain lengths the non-self-consistent surface bond order, ρ_{12}, is less than the infinite limit. This would suggest that the surface bond order is reduced at this band filling compared with the bond order in the bulk of an infinite chain. However this picture is completely reversed by all the self-consistent solutions for ρ_{12}. The effect which self-consistency has on these one-dimensional models is large because the initial deviation from local charge neutrality can be very large particularly in 0.25 and 0.75 filled bands. These deviations are not expected to be so large in two- and three-dimensional models where charge flow and hence screening is facilitated by the greater connectivity of the clusters. However it seems reasonable to assume that in all dimensions the change in a bond order caused by requiring local charge neutrality will be of the same order as the deviation from local charge neutrality in the non-self-consistent solution. Therefore self-consistency could introduce a significant correction to the force on an atom.

5. Covalent bond energies for canonical d bands in transition metals

To illustrate the concept of covalent bond energies in three-dimensional solids we have evaluated them for the three close-packed crystal structures BCC, FCC and HCP (ideal axial ratio) using the canonical d-band parameters given in (3.6), as a function of band filling. The covalent bond energy of the solid has the form given in (2.25) and since we are assuming the orbitals are orthonormal this may be written as

$$E_{cov} = \tfrac{1}{2} \sum_{\substack{i,j \\ i \neq j}} \sum_{\alpha\beta} 2\rho_{i\alpha j\beta} \bar{H}_{j\beta i\alpha}. \tag{5.1}$$

We shall restrict the discussion here to d band metals and hybridisation with s and p

electrons is ignored. There are five d orbitals on each atom and they have the following symmetry.

Orbital (α)	Symmetry
1	xy
2	yz
3	zx
4	$(x^2 - y^2)/2$
5	$(3z^2 - r^2)/2\sqrt{3}$

With the z axis along the bond the matrix elements of the Hamiltonian take the following simple form (Slater and Koster 1954).

$$\tilde{H}_{11} = \tilde{H}_{44} = \mathrm{dd}\delta$$

$$\tilde{H}_{22} = \tilde{H}_{33} = \mathrm{dd}\pi$$

$$\tilde{H}_{55} = \mathrm{dd}\sigma$$

$$\tilde{H}_{pq} = 0 \qquad p \neq q. \tag{5.2}$$

The covalent energy of each i–j bond becomes

$$E^{ij}_{\mathrm{cov}} = 2\mathrm{dd}\sigma\rho_{i5j5} + 2\mathrm{dd}\pi\,(\rho_{i2j2} + \rho_{i3j3}) + 2\mathrm{dd}\delta(\rho_{i1j1} + \rho_{i4j4}). \tag{5.3}$$

The matrix elements of $\tilde{H}$ for a general orientation of the bond can be obtained in terms of $\mathrm{dd}\sigma$, $\mathrm{dd}\pi$ and $\mathrm{dd}\delta$ by the transformations given in Slater and Koster (1954).

The calculation of the elements of ρ in (5.3) requires off-diagonal elements of the Green function. These were obtained by the recursion method. Following Heine *et al* (1980) and Finnis and Pettifor (1985) each off-diagonal matrix element was expressed as the difference between two diagonal Green function matrix elements. 15 exact levels of the continued fraction were used with the termination of Beer and Pettifor (1984). The results for the three structures are described in turn. For the BCC structure there are two contributing bonds, the first and second neighbours. In figure 2(*a*) we have plotted the $\mathrm{dd}\sigma$, $\mathrm{dd}\pi$ and $\mathrm{dd}\delta$ contributions to these covalent bond energies, i.e. the three terms in (5.3) weighted by the number of neighbours (6 and 8). It is clear that the nearest-neighbour σ- and π-bonds dominate the bonding energy, partly because of the R^{-5} decay of the tight-binding parameters, which by second neighbours are at about half of their first-neighbour values, and partly due to the smaller amplitude of the bond order.

The FCC structure is simpler, because only a single set of three tight-binding parameters is required for the 12 nearest neighbours. The contributions are shown in figure 2(*b*) and as for BCC, the σ- and π-terms dominate.

The HCP case differs because, although all 12 nearest neighbours are equidistant, and have the same tight-binding parameters, the basal-plane neighbours are not equivalent to the neighbours at $(a/2, a/2\sqrt{3}, c/2)$. Their bond orders must therefore be calculated separately. It is seen from figure 2(*c*) that the σ-, π- and δ-contributions from basal and non-basal planes are rather similar and not surprisingly resemble the FCC contributions. The differences between basal and non-basal covalent bond energies are significant on the scale of the energy differences between structures. We also speculate that they explain the lower than ideal c/a ratio of the HCP transition metals Ti and Zr. For band fillings between 1 and 2, the non-basal neighbours in the ideal structure have a lower covalent π-bond energy than the basal neighbours (see figure 2(*c*)). Although the matrix element $\mathrm{dd}\pi$ has the same value for each nearest-neighbour bond, the π-bond order $(\rho_{i2j2} + \rho_{i3j3})$ is larger between non-basal neighbours for the band-fillings 1–2 and 5–7. From (2.42) this means that the covalent bond force pulling the non-basal neighbours
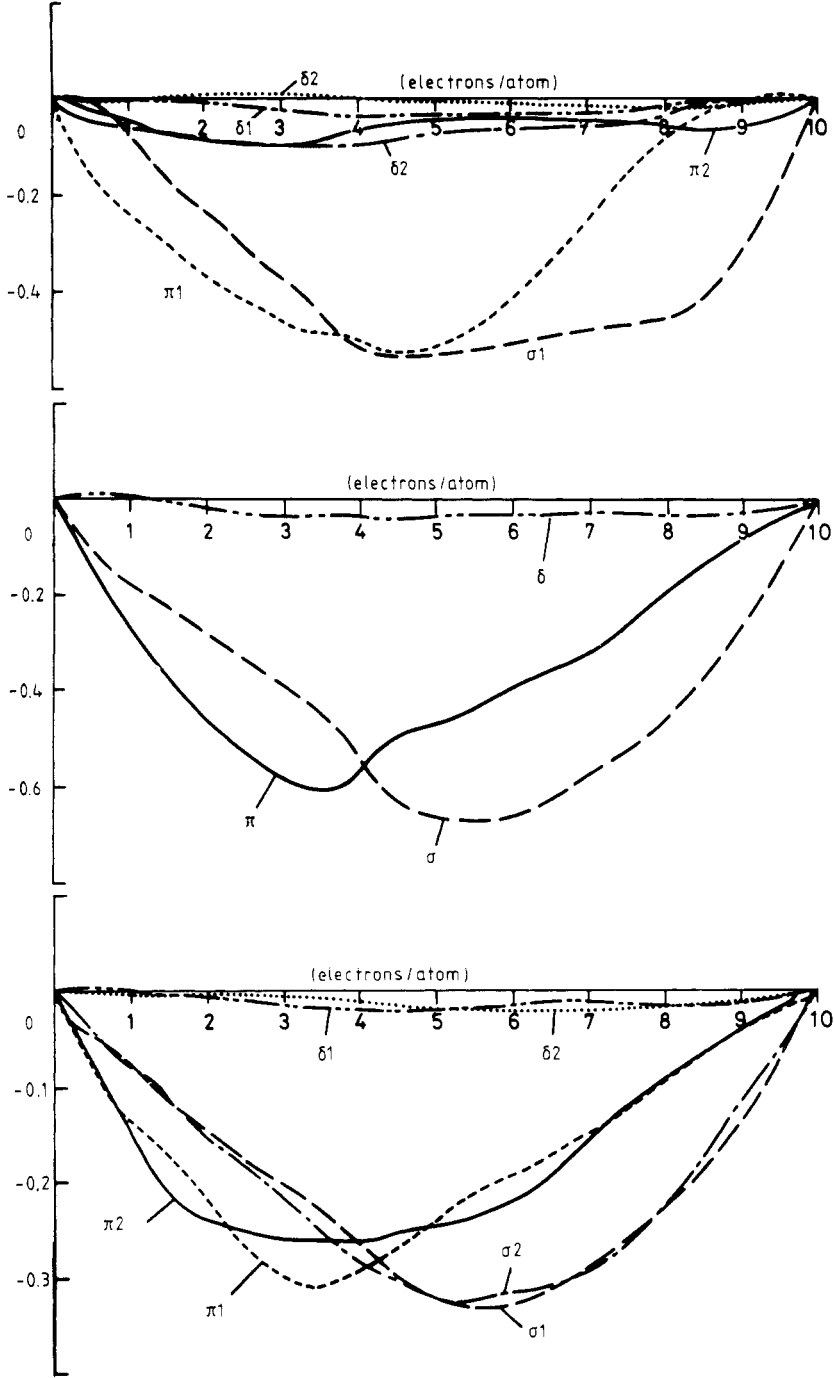
Figure 2. (*a*) Bond energies for a BCC structure. ddσ, ddπ and ddδ contributions to the bond energies, normalised by factors 6 and 8 for first and second neighbours, labelled 1 and 2 respectively. (*b*) Bond energies for a FCC structure. ddσ, ddπ and ddδ contributions to the bond energies, normalised by a factor 12 so that their sum gives the energy per atom. (*c*) Bond energies for an ideal HCP structure. ddσ, ddπ and ddδ contributions to the bond energies, normalised by a factor 6 so that their sum gives the energy per atom. 1 and 2 denote neighbours at $z = 0$ (basal plane) and $z = \pm c/2$. The unit of energy is W.

inwards is greater than that for the basal neighbours. Since the repulsive term is central it affects basal and non-basal neighbours alike. Thus the energy would be lowered by a reduction in c/a.

The covalent bond energies shown here can be summed from a given site to give the conventional site energy representation (see (2.26)). If we were interested only in calculating binding energies the on-site representation, (2.26), would be the more efficient method. However this is not true of the calculation of interatomic forces where the intersite representation, (2.42), is more efficiently calculated than the numerical derivative of the total density of states. In addition the covalent bond energy approach relates more closely to ideas of chemical bonding pioneered by Coulson (1939).

6. Conclusions

An expression for the binding energy of a solid was given in (2.11), which is correct to second order in the difference between the self-consistent charge density and a superposition of atomic charge densities. Each term in this expression has a clear physical meaning. The covalent bond and promotion energies are found by solving the electronic Hamiltonian to produce the output density matrix, (2.24). The covalent bond energy, which is part of the band energy, is simply the sum of the energies of all the covalent bonds in the solid. The two remaining terms in the binding energy represent the changes in the total electrostatic and exchange–correlation energies on condensing frozen atomic charge densities to form a fictitious solid with the same atomic positions as those in the real solid. To a good approximation these two terms may be represented by a sum of pair potentials.

The concept of covalent bond energies was illustrated for transition metals in § 5 using a canonical d-band model. In all three crystal structures examined the σ and π covalent bond energies dominated the δ covalent bond energies. An explanation for the lower than ideal value of the c/a ratio in HCP Ti and Zr was offered on the basis of the larger π-bond order between non-basal neighbours.

The requirement of local charge neutrality is the simplest approximation to self-consistency in the charge density which overcomes the problem discussed in § 3. Without any form of self-consistency in tight-binding models there can be large errors in the bulk modulus and certain force constants computed by frozen-phonon calculations. These errors arise because charge is not conserved, and because variations in the potential due to charge redistribution are not accounted for. The same argument applies to the calculation of interatomic forces. Indeed, model calculations in § 4.3 indicated that the value computed for an inter-atomic force can vary by the same order of magnitude as the deviaton from local charge neutrality. This is precisely what one would expect from the variational principle of density functional theory.

Model calculations in § 4.1 illustrated the computation of interatomic forces and relaxation for a semi-infinite linear chain model of a free surface. Two sum rules were also verified using an infinite linear chain. A linear chain model of a grain boundary was presented in § 4.2. This model, though very simple, showed that bond orders are affected by a perturbation in a way which is determined not only by the potential of the perturbation but also by the electronic response of the infinite bicrystal. In the case of a metal the electronic response of the bicrystal depends critically on the degree of band filling. Therefore, the success of cluster calculations of 'embrittlement' (see, for example, Briant and Messmer 1982) depends on how accurately the cluster is able to emulate the electronic response of the infinite host. Model one-dimensional calculations in § 4.3

indicated a marked dependence of the charge density on cluster size in non-self-consistent cases which was much reduced by requiring local charge neutrality.

We have used the TBB model to study bond energies and defect forces around a vacancy in transition metals (Ohta *et al* 1987b). We are currently employing this scheme to study grain boundaries in transition metals (Ohta *et al* 1987a) and in semiconductors (Paxton and Sutton 1987). In both cases we are using the recursion method to compute density matrix elements.

Acknowledgments

Appendix. Green function matrix elements for a constant semi-infinite linear chain

Consider the Hamiltonian defined by (4.1) and the Green function matrix element G_{jk}. We assume $k \geq j$, otherwise we evaluate G_{kj} and use $G_{jk} = G_{kj}$. Equation (4.2) may be derived without any Brillouin zone integrations by considering a sequence of 'perturbed' and 'unperturbed' Hamiltonians and using Dyson's equation to produce algebraic relations between corresponding Green functions. This method is based on Lannoo (1980).

We begin with the following pair of unperturbed and perturbed Hamiltonians represented schematically in figure A1. The 'perturbed' Hamiltonian, H, is defined as that of
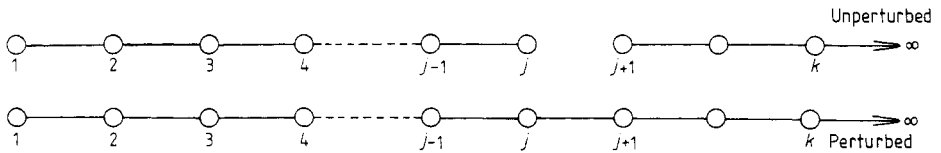

Figure A1. A schematic illustration of the Hamiltonians used to derive equation (A3).

the constant semi-infinite linear chain, given in (4.2). In the 'unperturbed' Hamiltonian, H^0, the bond between atoms j and $j + 1$ has been broken giving a finite chain of length j atoms which is decoupled from a semi-infinite chain of which the first atom is $j + 1$. The 'perturbation', V, is defined as $H - H^0$ and the only non-zero matrix elements are

$$V_{j,j+1} = V_{j+1,j} = \gamma. \tag{A1}$$

Dyson's equation relates the Greens function, G, for the perturbed system to that of the unperturbed Greens function, G^0, and the perturbation V.

$$G = G^0 + G^0 V G \tag{A2}$$

Using (A1) and (A2) it can be shown that

$$G_{jk} = \frac{G_{jj}^0 \gamma G_{j+1,k}^0}{1 - G_{jj}^0 \gamma^2 G_{j+1,j+1}^0}. \tag{A3}$$

G_{jj}^0 is the diagonal Green function matrix element for a terminating atom of the finite chain of j atoms and it is given by

$$G_{jj}^0 = D_{j-1}/D_j \tag{A4}$$

where D_j is the determinant of the following j by j matrix

$$\begin{pmatrix} E & -\gamma & 0 & & & \\ -\gamma & E & -\gamma & & & \\ 0 & -\gamma & & & & \\ & & & & -\gamma & 0 \\ & & & -\gamma & E & -\gamma \\ & & & 0 & -\gamma & E \end{pmatrix}$$

The energy E in this matrix is strictly $E + i\varepsilon$ where ε is an infinitesimal positive number. As is well known (see, for example Heine *et al* 1980) D_j satisfies a recurrence relation

$$D_j = ED_{j-1} - \gamma^2 D_{j-2}. \tag{A5}$$

Returning to (A3) $G_{j+1,j+1}^0$ is the diagonal element for the first atom of a constant semi-infinite chain. Using (A4) and (A5) in the limit of $j \to \infty$ it can be shown that

$$G_{j+1,j+1}^0 = (1/\gamma)\, e^{i\theta} \tag{A6}$$

where

$$\cos \theta = E/2\gamma. \tag{A7}$$

The remaining term, $G_{j+1,k}^0$ in (A3) is the same as G_{1n}^0 for a semi-infinite chain where $n = k - j$ and the index of the first atom of the chain is 1. To derive an expression for this matrix element consider the perturbed and unperturbed Hamiltonians shown schematically in figure A2. As before the perturbed Hamiltonian is simply the constant
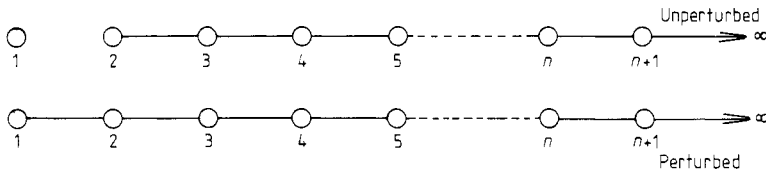

Figure A2. A schematic illustration of the Hamiltonians used to derive equation (A10).

semi-infinite chain. In the 'unperturbed' Hamiltonian the first atom of the chain has been decoupled, so the perturbation V now has the matrix elements

$$V_{12} = V_{21} = \gamma. \tag{A8}$$

Let the Green functions for these perturbed and unperturbed Hamiltonians be g and g^0 respectively. Then g_{1n} is identical to G_{1n}^0 for the semi-infinite chain. Since atom 1 is

decoupled and the diagonal Hamiltonian matrix elements are zero we have

$$g_{11}^0 = 1/E. \tag{A9}$$

Dyson's equation then yields

$$g_{1n} = g_{11}^0 \gamma g_{2n}^0 / (1 - g_{11}^0 g_{22}^0 \gamma^2) \tag{A10}$$

and since

$$g_{22}^0 = g_{11} = (1/\gamma) \, e^{i\theta} \tag{A11}$$

and

$$g_{2n}^0 = g_{1,n-1} \tag{A12}$$

equation (A10) can be expressed as

$$g_{1n} = [\gamma/(E - \gamma e^{i\theta})] \, g_{1,n-1} = e^{i\theta} g_{1,n-1} \tag{A13}$$

where we have used (A7). It follows from the recurrence relation, (A13), and (A6), that

$$g_{1n} = e^{in\theta}/\gamma. \tag{A14}$$

Equation (A3) can now be rewritten as

$$G_{jk} = [D_{j-1}/(D_j - \gamma D_{j-1} \, e^{i\theta})] \, e^{i(k-j)\theta} = (D_{j-1}/\gamma^j) \, e^{ik\theta} \tag{A15}$$

where use has been made of the recurrence relation, (A5), in the last equality. Chebyshev polynomials of the second kind satisfy a similar recurrence relation

$$U_j(\cos \theta) = 2 \cos \theta \, U_{j-1}(\cos \theta) - U_{j-2}(\cos \theta) \tag{A16}$$

where

$$U_j(\cos \theta) = \sin(j + 1)\theta/\sin \theta. \tag{A17}$$

Comparing (A5) and (A15) we see that

$$D_j(E) = \gamma^j U_j(\cos \theta) \tag{A18}$$

and therefore

$$G_{jk} = (1/\gamma)(\sin j\theta/\sin \theta) \, e^{ik\theta}. \tag{A19}$$

This expression may be confirmed directly by substituting it into $(E - H)G = I$.

References

Allan G 1970 *Ann. Phys., Paris* **5** 169
Ballentine L E and Kolar M 1986 *J. Phys. C: Solid State Phys.* **19** 981
Baraff G A and Schluter M 1979 *Phys. Rev.* B **19** 4965
Basinski Z S, Duesbery M S and Taylor R 1971 *Can. J. Phys.* **49** 2160
Beer N and Pettifor D G 1984 *Electronic Structure of Complex Systems* ed. P Phariseau and W M Temmerman (New York: Plenum) p 769
Bilby B A and Crocker A G 1965 *Proc. R. Soc.* A **288** 240
Born M and Huang K 1954 *Dynamical Theory of Crystal Lattices* (Oxford: OUP)
Boswarva I M and Esterling D M 1982 *J. Phys. C: Solid State Phys.* **15** L729
Briant C L and Messmer R P 1982 *Acta Metall.* **30** 1811

Callaway J 1974 *Quantum Theory of the Solid State* (New York: Academic) p 291

Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471

Chadi D J 1979 *Phys. Rev.* B **19** 2074

—— 1984 *Phys. Rev.* B **29** 785

Chadi D J and Cohen M L 1975 *Phys. Status Solidi* b **68** 405

Chelikowsy J R and Louie S G 1984 *Phys. Rev.* B **29** 3470

Christian J W 1975 *The Theory of Transformations in Metals and Alloys, Part I* (Oxford: Pergamon) p 33

Coulson C A 1939 *Proc. R. Soc.* A **169** 413

Daw M S and Baskes M I 1984 *Phys. Rev.* B **29** 6443

Eberhart M E, Johnson K H, Messmer R P and Briant C L 1983 *Atomistics of Fracture* ed. R M Latanision and J R Pickens (New York: Plenum) p 255

Economou E N 1983 *Green's Functions in Quantum Physics, Springer Series in Solid State Sciences* vol 7 (Berlin: Springer) p 79

Egami T and Vitek V 1983 *Amorphous Materials* ed. V Vitek (Warrendale PA: Metallurgical Society of the American Institute of Mining, Metallurgical and Petroleum Engineers) p 127

Ercolessi F, Tosatti E and Parrinello M 1986 *Phys. Rev. Lett.* **57** 719

Esterling D M, Som D K, Chatterjee A K and Boswarva I M 1987 *J. Phys. F: Met. Phys.* **17** 87

Finnis M W, Kear K L and Pettifor D G 1984 *Phys. Rev. Lett.* **52** 291

Finnis M W and Pettifor D G 1985 *The Recursion Method and its Applications, Springer Series in Solid State Sciences* vol 58, ed. D G Pettifor and D L Weaire (Berlin: Springer) p 120

Finnis M W and Sinclair J E 1984 *Phil. Mag.* A **50** 45

Foulkes M 1987 *PhD Thesis* University of Cambridge

Foulkes M and Haydock R 1986 *J. Phys. C: Solid State Phys.* **33** 6573

Garcia-Moliner F and Flores F 1979 *Introduction to the Theory of Solid Surfaces* (Cambridge: CUP) p 146

Harris J 1985 *Phys. Rev.* B **31** 1770

Harrison W A 1980 *Electronic Structure and the Properties of Solids* (San Francisco: Freeman)

—— 1985 *Phys. Rev.* B **31** 2121

Heine V, Haydock R, Bullett D W and Kelly M J 1980 *Solid State Physics* **35** (New York: Academic)

Hurley A C 1976 *Electron Correlation in Small Molecules* (London: Academic)

Kohn W and Sham L J 1965 *Phys. Rev.* **140** A113

Kutzelnigg W 1977 *Methods of Electronic Structure Theory* ed H F Schaefer III (New York: Plenum) p 129

Lannoo M 1980 *Electronic Structure of Crystal Defects and of Disordered Systems* (ed. F Gautier, M Gerl and P Guyot) (Les Ulis, France: Les Editions de Physique) p 45

Lannoo M and Bourgoin J 1981 *Point Defects in Semiconductors I, Springer Series in Solid State Sciences* vol 22 (Berlin: Springer) pp 115–21

Legrand B 1985 *Phil. Mag.* A **52** 83

Lindholm E and Lundqvist S 1985 *Phys. Scr.* **32** 220

Mackintosh A R and Andersen O K 1980 *Electrons at the Fermi Surface* ed. M Springford (Cambridge: CUP) § 5.3

Majewski J A and Vogl P 1986 *Phys. Rev. Lett.* **57** 1366

Masuda K, Yamamoto R and Doyama M 1983 *J. Phys. F: Met. Phys.* **13** 1407

Moraitis G and Gautier F 1979 *J. Phys. F: Met. Phys.* **9** 2025

Nielsen O H and Martin R M 1983 *Phys. Rev. Lett.* **50** 697

Nielsen O H and Martin R M 1985a *Phys. Rev.* B **32** 3780

—— 1985b *Phys. Rev.* B **32** 3792

Norskov J K 1982 *Phys. Rev.* B **26** 2875

Ohta Y, Finnis M W, Pettifor D G and Sutton A P 1987a to be published

—— 1987b *J. Phys. F: Met. Phys.* **17** L273

Payne M C, Bristowe P D and Joannopoulos J D 1987 *Phys. Rev. Lett.* **58** 1348

Paxton A T and Sutton A P 1987 *Phys. Rev. Lett.* submitted

Paxton A T, Sutton A P and Nex C M M 1987 *J. Phys. C: Solid State Phys.* **20** L263

Pettifor D G 1976 *Commun. Phys.* **1** 141

—— 1977 *J. Phys. F: Met. Phys.* **7** 613

—— 1978 *J. Chem. Phys. Lett.* **69** 2930

—— 1983 *Physical Metallurgy* ed. R W Cahn and P Haasen (Amsterdam: North-Holland) ch 3

—— 1987 *Solid State Phys.* **40** 43 (New York: Academic)

Pettifor D G and Podloucky R 1986 *J. Phys. C: Solid State Phys.* **19** 315

Pollmann J and Pantelides S T 1978 *Phys. Rev.* B **18** 5524

—— 1980 *Phys. Rev.* B **21** 709

Pollmann J, Kalla R, Kruger P, Mazur A and Wolfgarten G 1986 *Appl. Phys.* A **41** 21

Priester C, Allan G and Lannoo M 1986 *Phys. Rev.* B **33** 7386

Sankey O F and Allen R E 1986 *Phys. Rev.* B **33** 7164

Saqi M and Pettifor D G 1987 *Phil. Mag. Lett.* submitted

Sato A and Masuda K 1981 *Phil. Mag.* B **43** 1

Satoko C 1987 *Microclusters, Springer Series in Materials Science* vol 4 ed. S Sugano, Y Nishina and S Ohnishi
    (Berlin: Springer) p 142

Sayers C M 1984 *Phil. Mag.* B **50** 635

Slater J C 1968 *Quantum Theory of Matter* (New York: McGraw–Hill)

Slater J C and Koster G F 1954 *Phys. Rev.* **94** 1498

Stott M J and Zaremba E 1980 *Phys. Rev.* B **22** 1564

Sutton A P 1987 unpublished

Terakura K, Hamada N, Oguchi T and Asada T 1982 *J. Phys. F: Met. Phys.* **12** 1661

Treglia G, Desjonqueres M C and Spanjaard D 1983 *J. Phys. C: Solid State Phys.* **16** 2407

Williams A R, Feibelman P J and Lang N D 1982 *Phys. Rev.* B **26** 5433