Tight-binding modelling of materials

Rep. Prog. Phys. **60** (1997) 1447–1512. Printed in the UK PII: S0034-4885(97)67847-1

Tight-binding modelling of materials

C M Goringe†§, D R Bowler†‖ and E Hernández‡¶

† Department of Materials, Oxford University, Parks Road, Oxford OX1 3PH, UK
‡ Physics Department, Keele University, Keele, Staffordshire ST5 5BG, UK

Received 28 April 1997

Abstract

The tight-binding method of modelling materials lies between the very accurate, very expensive, *ab initio* methods and the fast but limited empirical methods. When compared with *ab initio* methods, tight-binding is typically two to three orders of magnitude faster, but suffers from a reduction in transferability due to the approximations made; when compared with empirical methods, tight-binding is two to three orders of magnitude slower, but the quantum mechanical nature of bonding is retained, ensuring that the angular nature of bonding is correctly described far from equilibrium structures. Tight-binding is therefore useful for the large number of situations in which quantum mechanical effects are significant, but the system size makes *ab initio* calculations impractical.

In this paper we review the theoretical basis of the tight-binding method, and the range of approaches used to exactly or approximately solve the tight-binding equations. We then consider a representative selection of the huge number of systems which have been studied using tight-binding, identifying the physical characteristics that favour a particular tight-binding method, with examples drawn from metallic, semiconducting and ionic systems. Looking beyond standard tight-binding methods we then review the work which has been done to improve the accuracy and transferability of tight-binding, and moving in the opposite direction we consider the relationship between tight-binding and empirical models.

§ To whom correspondence should be addressed. Present address: Australian Key Centre for Microscopy and Microanalysis, Madsen Building (F09), University of Sydney, NSW 2006, Australia. E-mail address: chris.goringe@cheerful.com

‖ Present address: Physics Department, Keele University, Keele, Staffordshire ST5 5BG, UK. E-mail address: d.bowler@keele.ac.uk

¶ Present address: Departamento de Física Teórica, Universidad de Valladolid, E-47011 Valladolid, Spain. E-mail address: ehe@boltzmann.fam.cie.uva.es

0034-4885/97/121447+66$59.50

Contents

1. Introduction

The huge increase in computational power available for scientific purposes in the past few decades has led to a proliferation of computational methods. In the field of atomistic modelling the early stages of the computer revolution, in which computational facilities were rare and expensive, led to the development of schemes which retained the essential quantum mechanical basis, but greatly reduced the computational time required. One such scheme, the tight-binding method, has proved to be of continuing interest for its combination of physical transparency and quantum mechanical sophistication, and (more pragmatically), its combination of computational speed and surprising accuracy. The significance of the original paper, by Slater and Koster (1954) can perhaps be best demonstrated by anecdotal evidence—in the shelves of the Radcliffe Science Library, Oxford, volume 94 of *Physical Review* stands out from its shelf for its broken spine, and falls open at table 1 on page 1503. The reason behind the development of the method is found in the limitations of computer power:

> 'All these complications add up to make the rigorous calculation of the matrix components of energy an almost impossible task. . . . The possibility is not excluded that eventually ways will be found to do this work by means of high speed computers. . . . No calculations which have yet been made by the LCAO method for crystals approach real rigour. . . '.

Perhaps the greatest testament to the value of the tight-binding method is that computing power which Slater and Koster could only speculate about is now routinely available, and rigorous *ab initio* calculations can be and are performed; interest in tight-binding is greater now than ever.

The structure of this review is as follows. Sections 2 and 3 respectively describe what tight-binding is and how it is done. Section 4 summarizes the approximations involved. Section 5 gives an overview of the wide range of scientific problems that have been tackled using tight-binding. Sections 6 and 7 consider the ways that tight-binding can be extended, to increase either accuracy or system size, and section 8 draws conclusions.

2. The tight-binding method

2.1. Overview

The tight-binding method works by writing the eigenstates of the Hamiltonian in an atomic-like basis set, $\{\phi_{i\alpha}\}$, and replacing the exact many-body Hamiltonian operator with a parametrized Hamiltonian matrix. The basis set is not, in general, explicitly constructed, but it is atomic-like in that it has the same symmetry properties as the atomic orbitals. In general only a small number of basis functions are used, those roughly corresponding to the atomic orbitals in the energy range of interest. For instance, when modelling diamond carbon, the 1s atomic orbitals can be reasonably neglected, and only the 2s and 2p orbitals considered. The eigenstates of the system are then obtained by solving the characteristic equation,

$$\hat{H}|\Psi_i\rangle = E_i|\Psi_i\rangle. \tag{1}$$

The justification of this approach is outlined in section 2.4.

In order to turn an electronic structure method into an atomistic method, total energy and atomic forces are required. The band energy of a crystal is evaluated by integrating the density of states, $n(E)$,

$$E_{\text{band}} = \int^{E_{\text{f}}} En(E)\,\mathrm{d}E \tag{2}$$

where E_{f} is the Fermi level. The method used to obtain E_{band}, and the derivatives thereof (and thus the band-energy contribution to the force) is what distinguishes the variety of tight-binding methods described below. All other parts of the energy are typically described by a pairwise function,

$$E_{\text{rep}} = \sum_{i \neq j} U_{ij} \tag{3}$$

an approximation discussed in section 2.4.

2.2. The tight-binding Hamiltonian

The starting point for any discussion of the tight-binding method for electronic and atomic-structure calculations must be Slater and Koster (1954). Indeed, the first three sections of that paper should still be considered essential reading for anyone with a serious interest in the subject. The problem was to find the electronic structure of an extended system. Bloch (1928) provided the formal mechanism for dealing with periodic systems, such as crystals, by means of the Bloch sum. Starting with an atomic orbital, $\phi_{i\alpha}(\boldsymbol{r} - \boldsymbol{R}_i)$, located on an atom at $\boldsymbol{R}_i$, the Bloch sum is taken over all the periodic images of this orbital,

$$N^{\frac{-1}{2}} \sum_{\boldsymbol{R}_i} \exp[\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{R}_i]\phi_{i\alpha}(\boldsymbol{r} - \boldsymbol{R}_i) \tag{4}$$

where N is the (infinite) number of periodic images. A Bloch sum is formed for each atomic orbital (labelled with the Greek character α, β, etc) on each atomic site (labelled with the Roman character i, j, etc) in the periodic-unit cell.

The atomic orbitals, $\phi_{i\alpha}(\boldsymbol{r} - \boldsymbol{R}_i)$, are not ideal for the purposes of analysis, as the orbitals on different atomic sites are not orthogonal to one another. Löwdin (1950) provided a scheme for creating an orthogonal set from a non-orthogonal one, in such a way as to preserve the symmetry properties of the original set. Löwdin functions, which are defined as

$$\psi_{i\alpha} = \sum_{i'\alpha'} S^{-1/2}_{i\alpha i'\alpha'} \phi_{i'\alpha'} \tag{5}$$

where S is the overlap matrix, have a greater extent in space than atomic orbitals, implying that the Hamiltonian matrix will have elements significantly different from zero between atoms which are second- or third-nearest neighbours. In Slater and Koster's scheme, Bloch sums are formed from these Löwdin functions, $\psi_{i\alpha}(\boldsymbol{r} - \boldsymbol{R}_i)$. These sums are used then as a set of basis functions with respect to which the Hamiltonian matrix elements,

$$H_{i\alpha j\beta} = N^{-1} \sum_{\boldsymbol{R}_i, \boldsymbol{R}_j} \exp[\mathrm{i}\boldsymbol{k} \cdot (\boldsymbol{R}_j - \boldsymbol{R}_i)] \times \int \psi^*_{i\alpha}(\boldsymbol{r} - \boldsymbol{R}_i) H \psi_{j\beta}(\boldsymbol{r} - \boldsymbol{R}_j)\,\mathrm{d}\boldsymbol{r} \tag{6}$$

can be evaluated as a function of $\boldsymbol{k}$. This can be simplified by noting that one of the two sums can be cancelled with the factor N^{-1}, giving a sum over the periodic images of one

of the two atomic sites,

$$H_{i\alpha j\beta} = \sum_{\boldsymbol{R}_j} \exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_i)] \times \int \psi^*_{i\alpha}(\boldsymbol{r} - \boldsymbol{R}_i) H \psi_{j\beta}(\boldsymbol{r} - \boldsymbol{R}_j)\,\mathrm{d}\boldsymbol{r}. \qquad (7)$$

The key idea in Slater and Koster's approach, and one which underlies all tight-binding calculations, is to *replace* the integral in this equation with a parameter which depends only upon the internuclear distance $|\boldsymbol{R}_i - \boldsymbol{R}_j|$ and the symmetry of the orbitals involved. The first approximation required is the 'two-centre approximation'. The integral is between two Löwdin orbitals, via the Hamiltonian operator, one part of which is the potential due to all atoms in the system. In the two-centre approximation, the potential part of the Hamiltonian is replaced by the potential due to only the two atoms, i and j, upon which the orbitals are located. The integral now depends only upon the form of the Löwdin functions, which have the same symmetry as the corresponding atomic orbitals, and the vector $(\boldsymbol{R}_j - \boldsymbol{R}_i)$ between the atoms. The Löwdin functions can be expanded as a sum of functions with well defined angular momenta with respect to the axis between the atoms, so a Löwdin function formed from an atomic p orbital is expanded as pσ and p$\pi_\pm$. The value of the integral (for a given interatomic separation) between these expanded terms can be written as a constant (depending only upon which Löwdin functions are involved) multiplied by an angular term (depending upon the symmetry of the functions and the direction cosines, k, l and m, of the vector $(\boldsymbol{R}_j - \boldsymbol{R}_i)$).

The Hamiltonian matrix elements can therefore be written

$$H_{i\alpha j\beta} = \sum_{\boldsymbol{R}_j, J} \exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_i)] h_{\alpha\beta J}(|\boldsymbol{R}_j - \boldsymbol{R}_i|) G_{\alpha\beta J}(k, l, m) \qquad (8)$$

where J notates the angular momentum of the bond, which takes the value zero (σ) if either α or β has s symmetry, zero or one (π) if both α and β have at least p symmetry, zero, one or two (δ) if both α and β have at least d symmetry and so on. $h_{\alpha\beta J}$ is the constant (for a given $(|R_j - R_i|)$) for this integral, and $G_{\alpha\beta J}(k, l, m)$ is the angular dependance, as taken from Slater and Koster (1954, table 1). The notation above is not identical to that of Slater and Koster, for reasons of consistency with later developments. Given the Hamiltonian matrix as a function of $\boldsymbol{k}$, the band structure is obtained by solving the single-particle Schrödinger equation, (1).

As discussed above, the basis functions do not need to be explicitly evaluated in the tight-binding method, as the only information required to find the electronic structure of the system is the Hamiltonian matrix elements. Since these are written in the parametrized form described above, the whole system is described by the parametrization scheme. This means that any tight-binding calculation is only as good as the parametrization used; the choice of parameters is a subject we shall return to throughout this paper.

2.3. From band structure to atomistics

Although Slater and Koster used the scheme only to investigate the band structure of periodic systems, later work has developed the tight-binding scheme into an atomistic method. In order to make this transition, three issues need to be considered.

First, a continuous functional form is needed for the distance dependance of the Hamiltonian matrix elements (whereas for bandstructure calculations only the values at discrete values of interatomic separation are required). Much early work assumed a r^{-2} dependance (Froyen and Harrison 1979), although this significantly underestimates the bulk modulus of many materials (Chadi 1979a). The choice of the distance dependance is an

integral part of the formulation of a good parametrization, and should be considered in that context.

Secondly, a total energy, rather than simply a band energy, must be obtained. This problem was first approached by Chadi (1979a), who considered the relaxation of the (110) surface of a range of diamond and zinc blende structure semiconductors. He proposed that the total energy could be written

$$E_{\mathrm{tot}} = E_{\mathrm{band}} + E_{\mathrm{rep}} \tag{9}$$

with the repulsive energy written as a sum of pair terms,

$$E_{\mathrm{rep}} = \sum_{i,j} U_{ij}. \tag{10}$$

Although Chadi chose to work with ΔE_{rep} rather than E_{rep}, and the particular form he chose for U is not widely used, the principle that the terms in the energy which are not included in the single-electron eigenstates can be approximated by a sum of pair terms is of great importance, and almost universally applied. The validity of this approximation is discussed in section 2.4 below.

Thirdly, it is often necessary to obtain atomic forces as well as energies. Assuming that a differentiable form for the repulsive pair energy, U, has been chosen, the difficulty lies on the differentiation of E_{bs}. How easy or difficult it is to obtain forces depends on the method used to obtain E_{bs}; this will be discussed in section 3.1 below.

2.4. The tight-binding bond formulation

We now need to return to the question of the relationship between tight-binding and density functional theory, and in particular, to the validity of the approximation (10) that the total energy of a system can be written as the sum of eigenstates of a Hamiltonian matrix plus a sum of pair terms. This question has been carefully dealt with in two papers, Sutton *et al* (1988) and Foulkes and Haydock (1989). The approaches are similar; herein we follow the notation of Sutton *et al* (1988), as it leads to the tight-binding bond model, a physically transparent form for the tight-binding energy.

The functional starting point is the formulation of Harris (1985) and Foulkes (1987). An approximate charge density, ρ^f, is chosen, and the single-particle Kohn–Sham (1965) potential is constructed;

$$\tilde{V} = v(\boldsymbol{r}) + V^f(\boldsymbol{r}) + \mu_{\mathrm{xc}}^f(\boldsymbol{r}) \tag{11}$$

where $v(\boldsymbol{r})$ is the ionic potential, and $V^f(\boldsymbol{r})$ and $\mu_{\mathrm{xc}}^f(\boldsymbol{r})$ the Hartree and exchange and correlation potentials due to the charge density ρ^f. The Kohn–Sham Hamiltonian is formed by adding the kinetic-energy operator; this is then solved and the eigenstates used to form an output-charge density, $\rho^{\mathrm{out}}(\boldsymbol{r})$.

In a self-consistent approach, this output-charge density would be used as a new input-charge density†, and the process repeated until the input and output densities were equal, at which point the self-consistent-charge density ρ^{sc} has been obtained. However, Harris and Foulkes instead exploited the variational principle to write

$$E \approx \sum_n a_n \tilde{\epsilon}_n - \int \mathrm{d}\boldsymbol{r}\,\rho(\boldsymbol{r})[\tfrac{1}{2}V^f(\boldsymbol{r}) + \mu_{\mathrm{xc}}^f(\boldsymbol{r})] + E_{\mathrm{xc}}[\rho^f] + E_{\mathrm{ii}} + \mathrm{O}(\rho^{\mathrm{sc}} - \rho^f)^2 + \mathrm{O}(\rho^{\mathrm{sc}} - \rho^{\mathrm{out}})^2. \tag{12}$$

† In fact, the output-charge density and the input-charge density are usually mixed to obtain a new input-charge density, but the principle remains the same.

The merit of this functional is that it is stationary; that is, errors in the energy are second order in errors in the assumed-charge density.

Writing (12) in basis-independent form,

$$E \approx \mathrm{Tr}[\rho^{\mathrm{out}} H] - \mathrm{Tr}\,\rho^f (V^f/2 + \mu_{\mathrm{xc}}^f) + E_{\mathrm{xc}}[\rho^f] + E_{\mathrm{ii}} \tag{13}$$

and taking the assumed-charge density to be the superposition of the free-atom-charge densities (following Harris (1985)), $\rho^f = \sum_i \rho_i$, (13) can be written

$$\begin{aligned} E \approx \mathrm{Tr}(\rho^{\mathrm{out}} - \rho^f)H + \mathrm{Tr}\sum_i \rho_i \left(\sum_{i \neq j} V_j^f/2 + v_j \right) + E_{\mathrm{ii}} + E_{\mathrm{xc}}[\rho^f] \\ - \sum_i E_{\mathrm{xc}}[\rho_i] + \sum_i (\mathrm{Tr}\,\rho_i (T + V_i^f/2 + v_i) + E_{\mathrm{xc}}[\rho_i]). \end{aligned} \tag{14}$$

Taking this line by line, we see that the first line is the sum of occupied eigenstates in the system (E_{band}) minus the sum of occupied eigenstates in the free atoms ('$E_{\mathrm{band}}^{\mathrm{free}}$'); thus it is the promotional energy (the increase in energy due to promotion of electrons from the lowest energy orbitals in isolation to form hybrid orbitals in the bulk) plus the bonding energy (the reduction in energy due to the mixing of these hybrids to form bonding and anti-bonding states). The second line is simply the change in total electrostatic energy due to the condensation from isolated atoms to the bulk; similarly, the third line is the change in total exchange and correlation energy. Finally, the fourth line is the total energy of the isolated atoms; this can be subtracted from both sides of (14) to give the binding energy in the tight-binding bond form:

$$\begin{aligned} E_{\mathrm{B}} \approx\ & \mathrm{Tr}(\rho^{\mathrm{out}} - \rho^f)H + \\ & \Delta E_{\mathrm{es}}[\rho^f] + \\ & \Delta E_{\mathrm{xc}}[\rho^f]. \end{aligned} \tag{15}$$

We can now turn to the question: Can the terms in the binding energy which are not included in the sum of occupied eigenvalues be approximated by a sum of pair terms? The change in electrostatic energy (line 2 in (15)) is already in such a form, so we need only consider the exchange and correlation term. This is not exactly a sum of pair terms, because the exchange–correlation-energy functional is nonlinear. Taking the correlation first, Lindholm and Lundqvist (1985) showed that

$$E_{\mathrm{c}} = \frac{1}{2} \iint \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{r}'\, \rho(\boldsymbol{r}) \frac{h_{\mathrm{c}}(\boldsymbol{r}, \boldsymbol{r}', \rho)}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho(\boldsymbol{r}') \tag{16}$$

where the function h_{c} is related to the pair correlation function. It is, however, only weakly dependant upon ρ (Lindholm and Lundqvist 1985, Hurley 1976, Kutzelnigg 1977); if we make the approximation that it is independant of ρ then the change in correlation energy in forming the bulk becomes

$$\Delta E_{\mathrm{c}} = \frac{1}{2} \sum_i \sum_{j \neq i} \iint \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{r}'\, \rho_i(\boldsymbol{r}) \frac{h_{\mathrm{c}}(\boldsymbol{r}, \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho_j(\boldsymbol{r}') \tag{17}$$

which is a sum of pair terms.

Finally, the exchange energy is written in the Hartree–Fock approximation (that is, that there exists an external potential which gives rise to a Hartree–Fock charge density, $\rho^{\mathrm{HF}}(\boldsymbol{r})$, which is equal to $\rho^f(\boldsymbol{r})$) as

$$E_{\mathrm{x}}[\rho] = -\frac{1}{2} \iint \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{r}'\, \frac{\rho^{\mathrm{HF}}(\boldsymbol{r}, \boldsymbol{r}') \rho^{\mathrm{HF}}(\boldsymbol{r}', \boldsymbol{r})}{|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{18}$$

Since $\rho^{\mathrm{HF}}(\boldsymbol{r})$ is equal to $\rho^{f}(\boldsymbol{r})$, we approximate $\rho^{\mathrm{HF}}(\boldsymbol{r},\boldsymbol{r}') \approx \rho^{f}(\boldsymbol{r},\boldsymbol{r}')$, which leads to a sum of pair terms for the change in exchange energy

$$\begin{aligned} \Delta E_{\mathrm{x}} &= -\tfrac{1}{2}\sum_{i\alpha}\sum_{j\beta,i\neq j}(\rho^{f})^{i\alpha i\alpha}(\rho^{f})^{j\beta j\beta}L_{i\alpha j\beta} \\ L_{i\alpha j\beta} &= \iint \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,\frac{\psi^{*}_{i\alpha}(\boldsymbol{r})\psi_{j\beta}(\boldsymbol{r})\psi^{*}_{j\beta}(\boldsymbol{r}')\psi_{i\alpha}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}. \end{aligned} \tag{19}$$

In summary, therefore, within well defined approximations, the Harris–Foulkes formulation of density functional theory can be written as a tight-binding Hamiltonian plus a sum of pair terms.

2.5. Self consistency

The tight-binding method is most commonly used in a non-self-consistent fashion; that is, the Hamiltonian matrix does not depend upon the distribution of electrons. This is not to say that electrons are non-interacting, as Fermi exclusion is imposed by the orthogonality of the eigenstates, but the Coulombic interactions are neglected. This approximation may break down in two circumstances; first, when charge transfer is important, for instance in materials with significant ionic bonding, and secondly when the solutions of the Hamiltonian give rise to spurious charge transfers which should be prevented by Coulombic screening. Three ways of introducing a degree of self-consistency are described, in order of their frequency of use: local charge neutrality (LCN), Hubbard U, and self-consistent tight-binding (SCTB), the last of which will be discussed as part of *ab initio* tight-binding, below.

We need first to consider how the charge on an atom is defined in the tight-binding-model. Writing the eigenstates of the Hamiltonian in terms of the atomic-like basis set,

$$\Psi_k = \sum_{i\alpha} c^{(k)}_{i\alpha}\psi_{i\alpha} \tag{20}$$

the number of electrons 'on' atom i can be written

$$Q_i = \sum_{k\alpha} |c^{(k)}_{i\alpha}|^2 f_k \tag{21}$$

where f_k notates the occupation of eigenstate k. The net charge on the atom is then obtained by the mismatch between this value and the ionic charge of the atom.

Because a charged particle attracts the opposite charge, it will gather about itself a 'screen' of equal and opposite charge, so that from the outside it appears neutral. Of particular interest is the distance scale over which this occurs, known as the screening length. In metallic materials, this length is shorter than the interatomic spacing, as a result of which no atom can carry any significant charge. This situation is described in tight-binding by the imposition of LCN. In this scheme, the on-site energies are adjusted site by site by a local offset,

$$H_{i\alpha i\alpha} = H^{0}_{i\alpha i\alpha} + \delta E_i \tag{22}$$

where δE_i is chosen such that the electron count on atom i is equal to the ionic charge. This value must be found iteratively, with the number of electrons on an atom obtained by solving the Hamiltonian, and the local offsets adjusted accordingly.

2.5.1. Hubbard U. In situations where it may not be appropriate to enforce LCN, because charge-transfer effects are significant, but bonding is still covalent, an approximation to self-consistency must be made. The simplest such approximation is the Hubbard U, in which an energy cost is introduced for charging an atom, normally by adding a term to the on-site energies of the Hamiltonian:

$$\delta E_i = U \times (Q_i - Z_i) \tag{23}$$

where the number of electrons on an atom, Q_i, is defined in (21) above, Z_i is the nuclear charge and U is the Hubbard U, which is generally in the range 1–20 V. Examples of the use of this simple formalism include small silicon clusters (Tománek and Schlüter 1986), charged silicon clusters (Tománek and Schlüter 1987), carbon clusters (Tománek and Schlüter 1991, Xu *et al* 1992), twist-grain boundaries in silicon (Kohyama and Yamamoto 1994a) and liquid carbon (Morris *et al* 1995). The major conclusions from these studies are that the calculation is relatively insensitive to the value of U, and that unusually coordinated semiconductor atoms require a Hubbard U to prevent unphysical charge transfers. Ho and co-workers concluded that the Hubbard U had little effect on their liquid-carbon results (Morris *et al* 1995), but their tests and results are far from conclusive.

3. Solving the tight-binding Hamiltonian—doing tight-binding

As available computer power has grown, the size of system that can be considered in the tight-binding method has also grown. As systems become larger, however, the issue of scaling becomes crucial. The number of computational operations required to diagonalize a matrix is proportional to the cube of the number of basis functions, and thus to the number of atoms; this behaviour is referred to as $\mathcal{O}(N^3)$ scaling. As a result, a thousand-fold increase in computer power only buys a ten-fold increase in system size. This reflects the possibility for eigenstates to be arbitrarily delocalized.

However, a chemical picture of bonding suggests that it is essentially local, that is, the behaviour of a particular bond depends only upon a local environment; quantum mechanical effects are, in general, short ranged. Conceptually, therefore, it should be possible to produce good approximations to 'exact' tight-binding which scale linearly with the system size ($\mathcal{O}(N)$ scaling). These break down naturally into two approaches, discussed below: the recursion methods, which formulate the problem in terms of continuous fractions, and minimization methods, which formulate the problem in terms of the minimization of an energy functional.

3.1. Matrix diagonalization

Once the Hamiltonian matrix elements are known, it is relatively easy to obtain both the eigenvalues and the eigenvectors of the matrix by solving (1) by direct matrix diagonalization. The bandstructure energy is then

$$E_{\text{band}} = \sum_{\boldsymbol{k}} \sum_{i} \epsilon_i f_i w_{\boldsymbol{k}} \tag{24}$$

where f_i is the occupation of ith eigenstate, and the $w_{\boldsymbol{k}}$ are the weights given to the special $\boldsymbol{k}$-points, such that $\sum_{\boldsymbol{k}} w_{\boldsymbol{k}} = 1$ (see Monkhurst and Pack 1976 for a discussion of special $\boldsymbol{k}$-point selection).

Alternatively, the eigenvectors

$$\Psi_m = \sum_{i\alpha} c_{i\alpha}^m \psi_{i\alpha} \tag{25}$$

can be obtained from the matrix diagonalization procedure and the density matrix evaluated

$$\rho_{i\alpha j\beta}(\boldsymbol{k}) = \sum_m c_{i\alpha}^{(m)*} c_{j\beta}^{(m)} f_m \tag{26}$$

in terms of which the bandstructure energy can be written

$$E_{\text{band}} = \sum_{\boldsymbol{k}} 2\,\text{Tr}[\rho \mathbf{H}] w_{\boldsymbol{k}}. \tag{27}$$

The derivative of bandstructure energy with respect to an atomic coordinate, λ, and hence the force, are given by

$$\begin{aligned} F_\lambda &= -\frac{\partial E_{\text{bs}}}{\partial \lambda} \\ &= -\sum_{\boldsymbol{k}} 2\,\text{Tr}\left[\rho \frac{\partial \mathbf{H}}{\partial \lambda}\right] w_{\boldsymbol{k}}. \end{aligned} \tag{28}$$

A number of the methods described below obtain the density matrix in real space rather than in reciprocal space. For completeness, therefore, we note that (27) and (28) then become

$$E_{\text{band}} = 2\,\text{Tr}[\rho \mathbf{H}] \tag{29}$$

and

$$F_\lambda = -2\,\text{Tr}\left[\rho \frac{\partial \mathbf{H}}{\partial \lambda}\right] \tag{30}$$

respectively.

Alternatively, by repeating the diagonalization of the Hamiltonian matrix for a range of values of $\boldsymbol{k}$, normally, along a high-symmetry line, a band-structure plot, such as that shown in figure 2, can be obtained.

Although the $\mathcal{O}(N^3)$ scaling of matrix diagonalization limits the number of atoms to a few hundred, direct diagonalization of the Hamiltonian is still the method of choice for calculating band structures and accurate densities of states. The band structures of the different polymorphs of a given material are frequently used as part of the data set to be fit during the construction of tight-binding models (see especially Vogl *et al* 1983). Even when band structures are not directly used in the fitting procedure, the success (or lack of it) in reproducing a band structure calculated with more accurate methods is a good measure of the quality of a tight-binding model.

3.2. Recursion methods

Recursion methods take advantage of the fact that the local density of states depends upon the local environment. The *total* band energy for a system of atoms can be obtained from the *total* density of states (DOS) for the system of atoms, $n(E)$:

$$E_{\text{band}} = \int^{E_f} E n(E)\,\mathrm{d}E. \tag{31}$$

However, this is a *global* property of the system; the *global* density of states can be decomposed into a set of *local* densities of states (LDOS) (Friedel 1954), thus:

$$n(E) = \sum_{i\alpha} n_{i\alpha}(E) \tag{32}$$

where $n_{i\alpha} = \sum_k c_{i\alpha}^{(k)} \delta(E - E_k) c_{i\alpha}^{(k)}$. For a finite basis set, the LDOS has a finite width, a definite mean, and a shape. These properties (and hence the LDOS) can be described by its *moments* (Heine 1980). The pth moment of the local density of states, $n_{i\alpha}$, is $\mu_{i\alpha}^{(p)}$, where

$$\mu_{i\alpha}^{(p)} = \int E^p n_{i\alpha}(E)\,\mathrm{d}E. \tag{33}$$

There is a useful identity (Ducastelle and Cyrot-Lackmann 1970) which enables a correspondence between the moment and powers of the Hamiltonian to be drawn:

$$\mu_{i\alpha}^{(p)} = \int E^p n_{i\alpha}(E)\,\mathrm{d}E = \langle i\alpha|\hat{H}^p|i\alpha\rangle \tag{34}$$

$$\mu_{i\alpha}^{(p)} = \sum_{j_1\beta_1,\ldots,j_{p-1}\beta_{p-1}} H_{i\alpha,j_1\beta_1} H_{j_1\beta_1,j_2\beta_2} \ldots H_{j_{p-1}\beta_{p-1},i\alpha} \tag{35}$$

which corresponds to hopping around a lattice along closed loops of length p, if $H_{i\alpha,j\beta}$ is the hopping parameter for orbital $i\alpha$ to $j\beta$. This shows a simple connection between the local bonding of an atom and its electronic structure.

While this seems to offer a promising way of reconstructing a function, most methods are unstable. However, the recursion method (Haydock 1980) is an important exception. This is a Green's function method for building the LDOS from a set of moments.

The LDOS is given in terms of the Green's function $G_{i\alpha,i\alpha}(Z)$ by:

$$n_{i\alpha}(E) = -\frac{1}{\pi} \lim_{\eta\to 0} \mathrm{Im}\{G_{i\alpha,i\alpha}(E + \mathrm{i}\eta)\}. \tag{36}$$

This form for the LDOS is useful, because the Green's function can be written in terms of a continued fraction related to the moments of the Hamiltonian, which is described in more detail in the appendix. The recursion method is therefore an efficient Green's-function method for the density of states. The band energy can be written in terms of a Green's function as:

$$E_{\mathrm{band}} = -\frac{1}{\pi} \sum_{i\alpha} \lim_{\eta\to 0} \mathrm{Im} \int^{E_f} \mathrm{d}E G_{i\alpha,i\alpha}(E + \mathrm{i}\eta) E \tag{37}$$

with $G_{i\alpha,i\alpha}$ obtained from equations (36) and (A3).

3.2.1. Recursion and forces. The derivation of forces from the recursion method is extremely helpful in understanding the formalism of the bond order potential (BOP) method, and also shows several of the problems with the recursion method. The obvious manner of obtaining the contribution to the force on an atom from the band energy is to differentiate the band energy with respect to the atomic coordinate.

However, this is extremely computer intensive for more than a few levels of recursion. An alternative is to use the density-matrix form of the force, (30). This is possible because the density-matrix element, $\rho_{i\alpha,j\beta}$, can be written in terms of a Green's-function element, $G_{i\alpha,j\beta}$:

$$\rho_{i\alpha,j\beta} = -\frac{1}{\pi} \lim_{\eta\to 0} \mathrm{Im} \int^{E_f} \mathrm{d}E G_{i\alpha,j\beta}(E + \mathrm{i}\eta). \tag{38}$$

The problem of how to find this Green's-function element is very important. The approach taken in the recursion method starts from bonding and anti-bonding states:

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|i\alpha\rangle + |j\beta\rangle) \\ |-\rangle &= \frac{1}{\sqrt{2}}(|i\alpha\rangle - |j\beta\rangle) \end{aligned} \tag{39}$$

From these two states, the Green's-function elements G_{++} and G_{--} can be found using the recursion technique. If the states $|+\rangle$ and $|-\rangle$ are expanded out, and G_{--} is subtracted from G_{++}, then the following expression for $G_{i\alpha,j\beta}$ is arrived at:

$$G_{i\alpha,j\beta} = \tfrac{1}{2}[G_{++} - G_{--}]. \tag{40}$$

The result just derived is known as the inter-site (IS) method, while that derived in the previous section is the site diagonal (SD) method, as it works with $G_{i\alpha,i\alpha}$. Both of these methods have severe problems—the problem with forces from the SD method has been alluded to above, while the IS method has appalling convergence for calculation of the bond energy (Glanville *et al* 1988), and a symmetry problem, which means that the energy and forces are not invariant for rotation about crystal axes (Inoue and Ohta 1987). These problems have led to a variety of different solutions, the first of which was the matrix recursion method (MRM) (Inoue and Ohta 1987, Paxton *et al* 1987, Paxton and Sutton 1989, Jones and Lewis 1984). In this method, the recursion coefficients are matrices rather than scalars (typically each matrix represents the orbitals on one atom), which removes the problem of rotational invariance; it does not, however, solve the problem of the IS convergence (Aoki 1993). For further information on this method, see the references given above. The improvement of the convergence of the IS method, and remedy for all the problems mentioned here, led to the development of BOP. Before that, an alternative moments approach will be examined.

3.2.2. The Fermi operator expansion method. The Fermi operator expansion (Goedecker and Colombo 1994, Goedecker and Teter 1995, Voter *et al* 1996) is an alternative, and conceptually simple, method of obtaining the density matrix, albeit at the expense of introducing a finite (and often significant) electronic temperature. The method starts by defining the Fermi matrix,

$$\boldsymbol{F}_{\mu,T} = f\left(\frac{\mathbf{H} - \mu}{kT}\right) \tag{41}$$

where $f(x) = 1/(1 + \exp(x))$, the Fermi function. The Fermi matrix plays the same role as the real-space density matrix in (29) and (30).

Now, the Fermi function only has to cover a finite width, that is the width of the density of states for the system in question (or the difference between the minimum and maximum eigenvalues of the Hamiltonian). Within this range, it can be represented by a polynomial in the energy,

$$f(x) = \sum_{p=0}^{n_{\mathrm{pl}}} C_p E^p \tag{42}$$

which means that the Fermi matrix $F_{\mu,T}$ can be represented as a polynomial in the Hamiltonian,

$$\boldsymbol{F}_{\mu,T} = \sum_{p=0}^{p=n_{\mathrm{pl}}} C_p \hat{H}^p. \tag{43}$$

This then gives the expression for one element of the Fermi matrix as:

$$\langle i\alpha|F_{\mu,T}|j\beta\rangle = \sum_{p=0}^{p=n_{\mathrm{pl}}} C_p \langle i\alpha|\hat{H}^p|j\beta\rangle \tag{44}$$

which is clearly a sum of moments of the Hamiltonian. Conceptually, then, the method works by fitting a polynomial to the Fermi function over the range of the eigenvalues. Then,

using the coefficients of this polynomial and moments of the Hamiltonian, elements of the finite-temperature density matrix, or the Fermi matrix, are constructed.

The order of the polynomial is given as $n_{\mathrm{pl}} \approx W/kT$ (Goedecker and Colombo 1994). As described, the method is $\mathrm{O}(N^2)$, as each element of the Fermi matrix will require n_{pl} operations, and there are N^2 elements. However, as discussed above, bonding in semiconductors and insulators (and, at high electronic temperature, metals) is local, and so the Fermi matrix can be truncated beyond a certain cut-off radius, leading to $\mathrm{O}(N)$ scaling.

If the simple moments of the Hamiltonian are used, then the method becomes unstable rather quickly. In practice, a Chebyshev polynomial is used (Goedecker and Teter 1995), which leads to a recursion relation for the coefficients:

$$p_{\mu,T}(H) = \frac{c_0}{2} + \sum_{j=1}^{n_{\mathrm{pl}}} c_j T_j(H) \tag{45}$$

and

$$\begin{aligned} T_0(H) &= I \\ T_1(H) &= H \\ T_{j+1} &= 2HT_j(H) - T_{j-1}(H) \end{aligned} \tag{46}$$

Once the Fermi matrix has been truncated, the forces calculated using (30) are not exactly equal to the derivative of energy (Voter *et al* 1996); the formalism derived therein does, however, give a force which is the exact derivative of the energy.

In some cases the error in energy due to the high electronic temperature may be significant; the Gillan scheme (1989) for extrapolating the $T = 0$ energy from a high-temperature energy will be discussed in section 5.1.7.

3.2.3. The BOP method. The BOP method (Pettifor 1989, Aoki 1993, Horsfield *et al* 1996) is an efficient method for obtaining the off-diagonal elements of the density matrix, $\rho_{i\alpha,j\beta}$ in terms of a single-site Green's function, $G_{i\alpha,i\alpha}$. It derives its name from the expression for the bond energy,

$$U_{\mathrm{bond}} = \sum_{i \neq j} h(R_{ij})\Theta_{ij} \tag{47}$$

where $h(R_{ij})$ are the hopping elements, and Θ_{ij} is the *bond order*, $\Theta_{i\alpha,j\beta} = 2\rho_{i\alpha,j\beta}$. This expression resembles a pair potential, but Θ_{ij} is environment dependent (Pettifor 1989). To assist in understand the mathematical formalism used in BOP, an earlier solution will be considered first.

3.2.4. The two-site BOP expansion. The two-site BOP expansion (Pettifor 1989, Pettifor and Aoki 1991, Aoki and Pettifor 1993) extends the simple result of (40) to a more general form. Consider the state:

$$|U_0^\lambda\rangle = \frac{1}{\sqrt{2}}[|i\alpha\rangle + \mathrm{e}^{\mathrm{i}\theta}|j\beta\rangle] \tag{48}$$

where $\theta = \cos^{-1}(\lambda)$. Here, λ defines a mixing between the two states; the previous states ($|+\rangle$ and $|-\rangle$) are given by $\lambda = -1$ and $\lambda = 1$. Now consider the Green's-function element G_{00}, and expand out $|U_0^\lambda\rangle$,

$$G_{00}^\lambda(Z) = \langle U_0^\lambda|\hat{G}(Z)|U_0^\lambda\rangle = \tfrac{1}{2}[\langle i\alpha|\hat{G}(Z)|i\alpha\rangle + \langle j\beta|\hat{G}(Z)|j\beta\rangle] + \lambda\langle i\alpha|\hat{G}(Z)|j\beta\rangle \tag{49}$$

A value for $G_{i\alpha,j\beta}$ can be found by differentiating (49) with respect to λ, which yields

$$G_{i\alpha,j\beta}(Z) = \frac{\partial G^{\lambda}_{00}(Z)}{\partial \lambda} \tag{50}$$

and

$$\begin{aligned} \Theta_{i\alpha,j\beta} &= \frac{\partial N^{\lambda}}{\partial \lambda} \\ N^{\lambda} &= -\frac{2}{\pi} \lim_{\eta \to 0} \mathrm{Im} \int^{E_f} \mathrm{d}E G^{\lambda}_{00}(E + \mathrm{i}\eta) \end{aligned} \tag{51}$$

where the physical interpretation of N^{λ} is the number of electrons in the state $|U^{\lambda}_0\rangle$. A further expansion of $G_{i\alpha,j\beta}$ in terms of recursion coefficients is described in the appendix.

The two-site BOP expansion, while it provides better convergence than (40), is still problematical. It is more slowly convergent for the bond energy than single-site recursion, which leads to a breakdown in equivalence between the single site (or site diagonal) and intersite expansions, and it is still affected by the symmetry problem referred to above. The solution to these problems was found in a reformulation of BOP, which will be described next.

3.2.5. The single-site BOP expansion. The single-site BOP expansion requires a reasonably high level of formal mathematics for rigorous proof and derivation; this is not the purpose of this paper. For rigorous definitions and derivations of this subject, the interested reader is referred to Aoki (1993) and Horsfield *et al* (1996). The discussion below is intended to describe enough of the formalism to enable a physical understanding of the method.

In the previous section, it was shown that the off-diagonal Green's-function element, $G_{i\alpha,j\beta}$ could be obtained from the composite state $|U^{\lambda}_0\rangle = \frac{1}{\sqrt{2}}[|i\alpha\rangle + \mathrm{e}^{\mathrm{i}\theta}|j\beta\rangle]$, where $\theta = \cos^{-1}(\lambda)$. From this Green's-function element the bond order between the orbitals $i\alpha$ and $j\beta$, $\Theta_{i\alpha,j\beta}$, can be obtained. Now consider a label which is different for every bond in the system, $\Lambda_{i\alpha,j\beta}$. This can be thought of as an element of a matrix Λ; the element can be given by the inner product between two vectors,

$$\Lambda_{i\alpha,j\beta} = (e^{\Lambda}_{i\alpha}|e^{\Lambda}_{j\beta}). \tag{52}$$

At this stage it is useful to introduce an *auxiliary space*. The vectors which span it will be denoted as $|e^0_{\nu})$, where the parenthesis is used to show the difference between this space and the space associated with the atomic orbitals.

The vectors $|e^{\Lambda}_{i\alpha})$ which were used above to construct the elements of the matrix Λ can be defined so as to have various useful properties: they exist in the space which is spanned by the orthonormal vectors $|e^0_{\nu})$; they are each associated with one orbital in the space spanned by the Hamiltonian, hence the index $i\alpha$; and they are only ever used in conjunction with an atomic orbital.

The generalization of the vector $|U^{\lambda}_0\rangle$ can be written:

$$|W^{\Lambda}_0\} = \sum_{i\alpha} |i\alpha\rangle|e^{\Lambda}_{i\alpha}) \tag{53}$$

where, again, the curly bracket is used to denote the difference between this vector and the others in the system. This is a generalization of the previous vector as it now links *all* the basis functions in the system, but will still be used as the starting vector for the recursion.

If this is used as the first vector for the recursion relation for G_{00}, then

$$\begin{aligned} G_{00}^{\Lambda} &= \{W_0^{\Lambda}|\hat{G}(Z)|W_0^{\Lambda}\} \\ &= \sum_{i\alpha,j\beta} (e_{i\alpha}^{\Lambda}|\langle i\alpha|\hat{G}(Z)|j\beta\rangle|e_{j\beta}^{\Lambda}) \\ &= \sum_{i\alpha,j\beta} G_{i\alpha,j\beta}(Z)\Lambda_{i\alpha,j\beta} \end{aligned} \tag{54}$$

This allows the central part of the one-site BOP expansion to be written:

$$G_{i\alpha,j\beta}(Z) = \frac{\partial G_{00}^{\Lambda}}{\partial \Lambda_{i\alpha,j\beta}}. \tag{55}$$

As the values of $\Lambda_{i\alpha,j\beta}$ and the vectors $|e_{i\alpha}^{\Lambda}\rangle$ have been left completely general, they can now be specified. If the value $\Lambda_{j\beta,k\gamma} = \delta_{i,j}\delta_{i,k}\delta_{\alpha,\beta}\delta_{\alpha,\gamma}$ is chosen (the notation relates to a piece of mathematical formalism described in Horsfield *et al* (1996)), then the following simple and central result can be derived:

$$G_{00}^{\Lambda} = G_{i\alpha,i\alpha} \tag{56}$$

Further applications of this result are touched on in the appendix, and are detailed in Horsfield *et al* (1996).

3.3. Minimization methods

The seminal paper by Car and Parrinello (1985) has provoked a revolution in computational condensed matter physics and chemistry over the last 12 years. They were the first to point out that the expansion coefficients of the Hamiltonian eigenstates could be regarded as the variables in a minimization problem. We will limit the discussion here to a brief description of the CP method as applied to tight-binding Hamiltonians, as was done by Laasonen and Nieminen (1990) and also by Khan and Broughton (1989).

The CP approach consists of extending the phase space associated with the ionic degrees of freedom by incorporating fictitious dynamical variables associated with the basis-set expansion coefficients. The extended Lagrangian that governs both the dynamics of the nuclei and the fictitious dynamics of the basis set coefficients is written as

$$\begin{aligned} \mathcal{L}(\{\boldsymbol{r}_i\},\{c_{i\alpha}\}) = {} & \tfrac{1}{2}\mu \sum_{n}^{\text{occ}} \dot{\boldsymbol{c}}_n \cdot \dot{\boldsymbol{c}}_n + \tfrac{1}{2}\sum_i m_i \dot{\boldsymbol{r}}_i^2 - 2\sum_{n}^{\text{occ}} \boldsymbol{c}_n \mathbf{H} \boldsymbol{c}_n \\ & + \sum_{nm}^{\text{occ}} \lambda_{nm}(\boldsymbol{c}_n \boldsymbol{S} \boldsymbol{c}_m - \delta_{nm}) - E_{\text{rep}}. \end{aligned} \tag{57}$$

In this equation, i labels ions, m and n label electronic states, and α labels the basis functions. μ is a parameter which plays the role of a mass, the fictitious mass of the $c_{i\alpha}$ coefficients. The vector $\boldsymbol{c}_n$ has as components the coefficients $c_{i\alpha}^{(n)}$ that define electronic state n. The dot over the variables represents a time derivative. The first two terms on the right-hand side are the fictitious kinetic energy associated with the wavefunction coefficients and the kinetic energy of the nuclei respectively. The third term is the band-structure energy, which affects both the dynamics of the nuclei and the $\{c_{i\alpha}\}$. The fourth term results from the need to impose orthogonality between the different eigenstates of the Hamiltonian; the λ_{mn} are the Lagrange multipliers corresponding to these constraints.

From (57) equations of motion can be derived for the $c_{i\alpha}$ coefficients and the ionic positions $\boldsymbol{r}_i$. These are given by

$$\mu\ddot{\boldsymbol{c}}_n = -4\mathbf{H}\boldsymbol{c}_n + 2\sum_m^{\text{occ}} \lambda_{nm}\boldsymbol{S}\boldsymbol{c}_m \tag{58}$$

and

$$m_i\ddot{\boldsymbol{r}}_i = -2\sum_n^{\text{occ}} \boldsymbol{c}_n(\nabla_{r_i}\mathbf{H})\boldsymbol{c}_n + \sum_{n\,m}^{\text{occ}} \lambda_{nm}\boldsymbol{c}_n(\nabla_{r_i}\mathbf{S})\boldsymbol{c}_m - \nabla_{r_i}E_{\text{rep}} \tag{59}$$

respectively. Equations (58) and (59) can be solved numerically using the standard algorithms of molecular dynamics (MD) (see for example Allen and Tildesley (1987) or Frenkel and Smit (1996)).

For a suitably chosen value of the fictitious mass μ, the dynamics of the ions and of the $c_{i\alpha}$ coefficients decouple. Then the eigenstates adiabatically adapt to the changing ionic configuration, and the ions follow a trajectory on the Born–Oppenheimer hypersurface of the extended phase space. For this to happen, the μ parameter has to be chosen much smaller than the ionic masses, and a sufficiently small time step must be used in the numerical integration of the equations of motion. This is one of the ways in which CPMD simulations are performed in the context of DFT plane-wave calculations, and similarly it could be used for tight-binding Hamiltonians, as indeed it has been used by Khan and Broughton (1989) in their study of Si clusters and surfaces. However, the timestep required for the integration of the equations of motion makes this approach unattractive compared with using the conjugate gradients (CG) method (Press *et al* 1992) to search for the electronic ground state, as proposed by Payne *et al* (1992).

One significant use of the extended Lagrangian method, first reported by Laasonen and Nieminen (1990), is the search for low-energy structures by combining the CP approach with simulated annealing. Laasonen and Nieminen investigated the low-energy configurations of small Si_n clusters, with n in the range of 2–10. Given the difficulties that conventional minimization methods encounter when applied to functions with many local minima, this is a very interesting application.

Both the extended Lagrangian and CG schemes, although more efficient than direct diagonalization of the Hamiltonian for large systems, suffer from the same drawback, as the cost of the computation scales as N^3, where N is the number of atoms in the system. The origin of this unfavourable scaling in minimization methods is the need to impose the orthogonality constraints between the occupied eigenstates of the Hamiltonian. The cost of this orthogonalization process is proportional to $N_{\text{occ}}M^2$, where N_{occ} is the number of occupied eigenstates, and M is the number of basis functions. It is in order to avoid this global orthogonalization that localized orbital methods (section 3.3.1) and truncated density-matrix methods (section 3.3.2) have been developed. Readers interested in linear-scaling methodologies should consult the recent review of Galli (1996), where a more-detailed account of such new methods is given.

3.3.1. The localized orbitals method. The method of localized orbitals (LO) was independently introduced by Mauri *et al* (1993) (MGC) and by Ordejón *et al* (1993). In both cases, a new object function was proposed, which, when minimized with respect to the expansion coefficients $c_{i\alpha}$, leads to the ground-state energy. The minimization of the new functional still requires $\mathcal{O}(N^3)$ operations, but if spatial localization is imposed on the orbitals, the scaling becomes $\mathcal{O}(N)$.

The method of Ordejón *et al* (1993) reduces to a particular case of the more general formalism of Mauri *et al* (1993), and therefore we will focus our attention on the latter method. Mauri *et al* introduced a new functional of $N/2$ orbitals $\{\psi\}$, where N is the number of electrons in the system. In the spirit of the CP methodology the $\{\psi\}$ are intended to be linear combinations of the eigenstates; the minimization discussed below is with respect to the coefficients of the $\{\psi\}$ in the atomic orbital-like $\{\phi\}$ basis.

The MGC functional has the form

$$\tilde{E}[\mathbf{Q}, \{\psi\}] = 2\,\mathrm{Tr}(\mathbf{QH}) - \eta[2\,\mathrm{Tr}(\mathbf{QS}) - N] \tag{60}$$

where $\mathbf{Q}$ is an $N/2 \times N/2$ matrix, and $\mathbf{H}$ and $\mathbf{S}$ are the Hamiltonian and overlap matrices in the basis of $\{\psi\}$ orbitals, respectively, and η is a parameter to be determined. Mauri *et al* (1993) showed that if the matrix $\mathbf{Q}$ is chosen as:

$$\mathbf{Q} = \sum_{n=0}^{\mathcal{N}} (\mathbf{I} - \mathbf{S})^n \tag{61}$$

where $\mathcal{N}$ is an odd number, and the value of the η parameter is appropriately chosen, then

$$\min \tilde{E}[\mathbf{Q}, \{\psi\}] = E_0 \tag{62}$$

i.e. the functional (60) has the electronic ground state as its absolute minimum.

Without any further constraints, the orbitals that minimize $\tilde{E}$ will extend throughout the volume of the whole system, and the computational cost of their calculation will still scale proportionally to N^3, but if the orbitals are constrained to remain localized within certain regions of space, as proposed by Galli and Parrinello (1992), then a linear scaling method results. However, in this case (62) is no longer correct, given the restriction that the imposition of localization constraints on the orbitals imposes on the degree of variational freedom, and thus it must be substituted by

$$\tilde{E}[\mathbf{Q}, \{\psi\}] \geqslant E_0. \tag{63}$$

In the particular case in which $\mathcal{N}$ is taken equal to 1 in (61), the functional proposed by Ordejón *et al* (1993) is obtained.

The LO method as described above is a viable alternative to diagonalization or constrained minimization techniques for solving the electronic-structure problem, having the advantage of being linear-scaling with the system size. However, while experience shows that the conventional functional $E[\{\psi\}]$ possesses a single, global minimum, this is not the case for the functional originally proposed by Mauri and co-workers. Indeed, it has been shown in several examples (Mauri and Galli 1994, Kim *et al* 1995) that (60) possesses local minima, which make the search for the ground state much more difficult, and sometimes only possible by using *a priori* knowledge of the system. It appears that constraining the number of orbitals to be equal to the number of occupied states and simultaneously imposing localization constraints results in the presence of these troublesome local minima. This situation has been solved by Kim *et al* (1995), who proposed a generalization of the original functional, involving a number of orbitals $M > N/2$. These authors showed that using a number of localized orbitals larger than the number of occupied states not only eliminated the appearance of local minima in the functional, but also improved significantly the estimates of ground-state properties such as bond distances and cohesive energies.

3.3.2. Density-matrix method. Rather closer contact may be made with the density functional roots of tight-binding by means of the density matrix. Writing the eigenstates as

in (25), the real-space density-matrix is defined as

$$\rho_{i\alpha,j\beta} = \sum_k c_{i\alpha}^{(k)*} c_{j\beta}^{(k)} f_k \tag{64}$$

which reduces, at zero temperature, to a sum over occupied states,

$$\rho_{i\alpha,j\beta} = \sum_k c_{i\alpha}^{(k)*} c_{j\beta}^{(k)}. \tag{65}$$

It is straightforward to demonstrate that the band energy and total number of electrons in the system are given by

$$E_{\text{band}} = 2\,\text{Tr}[\rho \mathbf{H}] \tag{66}$$

$$n_e = 2\,\text{Tr}[\rho]. \tag{67}$$

The density-matrix approach to tight-binding, initially proposed by Li *et al* (1993) (LNV), (with a very similar scheme being simultaneously proposed by Daw (1993)) is conceptually simple by comparison with the recursion methods described above. The ground-state density matrix (65) is obtained by minimizing the band energy (66) with respect to the elements of the density matrix, $\rho_{i\alpha,j\beta}$, subject to two constraints: (i) the density matrix must correspond to a set of orthogonal eigenstates and (ii) the electron number must be correct.

Condition (i) is an expression of the Fermi exclusion principle, and can be seen to be equivalent to the requirement that the density matrix be idempotent, that is, $\rho^2 = \rho$. This is satisfied by choosing ρ to be the result of a McWeeney (1960) transform applied to a 'trial' density matrix, σ,

$$\rho = 3\sigma^2 - 2\sigma^3. \tag{68}$$

If the band energy (66) is considered as a function of $\sigma_{i\alpha,j\beta}$, there exists a single local minimum, in which the idempotency condition is satisfied; at this local minimum ρ has the value which would be obtained from the eigenstates of the Hamiltonian, although the eigenstates themselves have not been evaluated.

The second condition can be fulfilled by shifting the energy zero in the Hamiltonian matrix to lie at the Fermi level, such that all occupied states have negative energy, and all unoccupied states have positive energy. Writing

$$\mathbf{H}' = \mathbf{H} - \mu \mathbf{I} \tag{69}$$

both conditions can therefore be satisfied by minimizing the grand potential,

$$\Omega = \text{Tr}[\rho \mathbf{H}'] \tag{70}$$

with respect to the elements $\sigma_{i\alpha,j\beta}$, for an appropriate choice of μ. In many circumstances μ can simply be chosen to lie in the bandgap; sometimes, however, it becomes necessary to allow μ to vary during a molecular dynamics run. In order to minimize Ω (70) must be differentiated with respect to the matrix elements of σ. Adopting the notation of LNV, this derivative is given by

$$\begin{aligned}(\nabla\Omega)_{i\alpha,j\beta} &= \frac{\partial\Omega}{\partial\sigma_{i\alpha,j\beta}} \\ &= (3(\sigma\mathbf{H}' + \mathbf{H}'\sigma) - 2(\sigma^2\mathbf{H}' + \sigma\mathbf{H}'\sigma + \mathbf{H}'\sigma^2))_{i\alpha,j\beta}\end{aligned} \tag{71}$$

Minimizing the grand potential (70) requires the repeated evaluation of Ω and $\nabla\Omega$, which involves matrix multiplication. As it stands, thereore, the method still scales as $\mathcal{O}(N^3)$. In order to obtain linear scaling, the density matrix needs to be made sparse in some manner.

The density matrix between basis functions falls away to zero as the separation between the atoms upon which they are sited increases. This decay is exponential for insulators, power law for metals (at 0 K) (Kohn 1959). The requirement of a sparse density matrix can therefore be fulfilled by allowing non-zero elements only between pairs of atoms within a given cut-off range of one another (Li *et al* 1993). Alternatively, non-zero elements can be allowed only between atoms within a given number of hops of one another, making contact with the recursion methods discussed above. The evaluation of the energy and gradient now requires sparse matrix multiplication; all operations scale as $\mathcal{O}(N)$.

As long as the number of non-zero elements of the density matrix per atom in the system remains constant with system size, all matrix multiplication operations (and hence the whole method) then scale linearly.

The energy minimization procedure is relatively straightforward, as along any line, σ^{search}, in the function space,

$$\sigma^{\text{new}} = \sigma^{\text{old}} + x\sigma^{\text{search}} \tag{72}$$

Ω is simply a cubic polynomial in x. The coefficients of the polynomial can be obtained directly from the value and first three derivatives at a given point (which can be obtained analytically) (Wang and Ho 1996), or the grand potential and first derivative may be obtained at two points (Goringe 1997), the latter method being more efficient in terms of total number of operations required.

As the density matrix method was originally proposed, the electron number was not constrained; instead, a fixed Fermi level, μ, was used. Although it is useful in some circumstances to be able to work at a fixed Fermi level, it is more common to constrain the electron number. This can be simply done by adding an outer iterative loop, in which the Fermi level is adjusted and minimization repeated until the electron number is correct. However, it is much more efficient to combine the electron correction step with the energy minimization, which has been discussed by Qiu *et al* (1994) and Goringe (1995, 1997).

Unlike the recursion methods, forces are easily calculated in the density-matrix method. As the approximate density matrix, ρ, is known, the Hellmann–Feynman forces can be calculated directly from (30). It is important to note that the Hellmann–Feynman forces are *exact* in the density-matrix method at any level of truncation; that is, as long as the energy has been minimized with respect to σ, the numerical and analytic derivatives of energy are identical. This point is graphically illustrated in the comparison of $\mathcal{O}(N)$ methods by Bowler *et al* (1997).

An important issue in any minimization method is the stability of the minimization algorithm. For the orthogonal tight-binding case discussed above, it has been found by all users that the McWeeny transformation results in a well behaved functional of which the local minimum is easily found. Tests on many thousands of randomly generated Hamiltonian matrices (Gillan, private communication) with a steepest descent minimization routine did not produce a single case in which the physical local minimum was not found. The extension of the density matrix method to non-orthogonal tight-binding (Nunes and Vanderbilt 1994), however, while easy to write down, leads to a more complex functional, and a more sophisticated minimization method must be adopted. White *et al* (1997) showed that the derivatives in Nunes and Vanderbilt are tensorially inconsistent and proposed a correct form which has the complication of involving the inverse of the overlap matrix. Interested readers are directed to White *et al* (1997) and references therein for further details.

4. Summary of approximations

The approximations made in the standard forms of orthogonal tight-binding can be summarized as follows:

- *The total energy is divided into the sum of single-electron eigenvalues plus the sum of pair terms.* The approximation is made that all terms in the energy which cannot be included in the single-electron effective Hamiltonian, in particular the exchange and correlation energy, can be described by a sum of pair terms.
- *The elements of the Hamiltonian matrix are taken to depend only upon the vector between two atoms.* That is, all three centre integrals in the Hamiltonian operator are neglected.
- *A minimal set of orthogonal, short-range basis functions, which span the occupied subspace, are assumed to exist.* A basis set can only be short range (atomic orbitals or linear muffin tin orbitals, for instance) or orthogonal (Löwdin orbitals). Generally in tight-binding it is assumed that a set can be chosen which is both.
- *Self consistency is neglected, or approximated by onsite terms.*

Wherever one of these approximations cannot be made, a more sophisticated technique is required. These are discussed in section 6 below.

Further approximations are made in particular techniques; these differ from the list above in that they can be systematically improved upon:

- *A limited set of k-points is used to integrate over the Brillouin zone.*
- *The recurring fraction is terminated after a finite number of levels.* This is equivalent to neglecting higher moments in the density of states.
- *The density matrix is taken to have a finite range.*

One of the final two approximations must be made if $\mathcal{O}(N)$ scaling is required, which will generally be the case for systems larger than around 100 atoms. The choice of which linear scaling approximation to make has been carefully investigated by Bowler *et al* (1997), but can be summarized as follows: for a non-metallic system, use the truncated density-matrix method, for a metallic system, use a recursion method. Figure 1 makes this point strikingly clear. Linear scaling methods have been applied to systems of several thousand atoms without supercomputing facilities. Should larger systems still be required, more approximate techniques based on tight-binding may be appropriate; these are discussed in section (7).

5. Applications of tight-binding

Throughout the following overview of calculations performed in the tight-binding approximation, the reader should note the wide range of systems in which the sheer number of atoms is such that *ab initio* methods are impractical, but in which the level of accuracy required is such that empirical methods are insufficient. As different versions of tight-binding tend to be appropriate for different chemical systems, the applications have been broken down into metals, semiconductors, fullerenes, and ionic materials.

5.1. Metals

To do justice to the modelling of metallic systems would require a complete review article; in the context of this work we can only hope to skim the surface. We will first consider some of the general issues involved in tight-binding studies of metals, then a number of variants on traditional tight-binding which are particularly relevant to metals, and finally

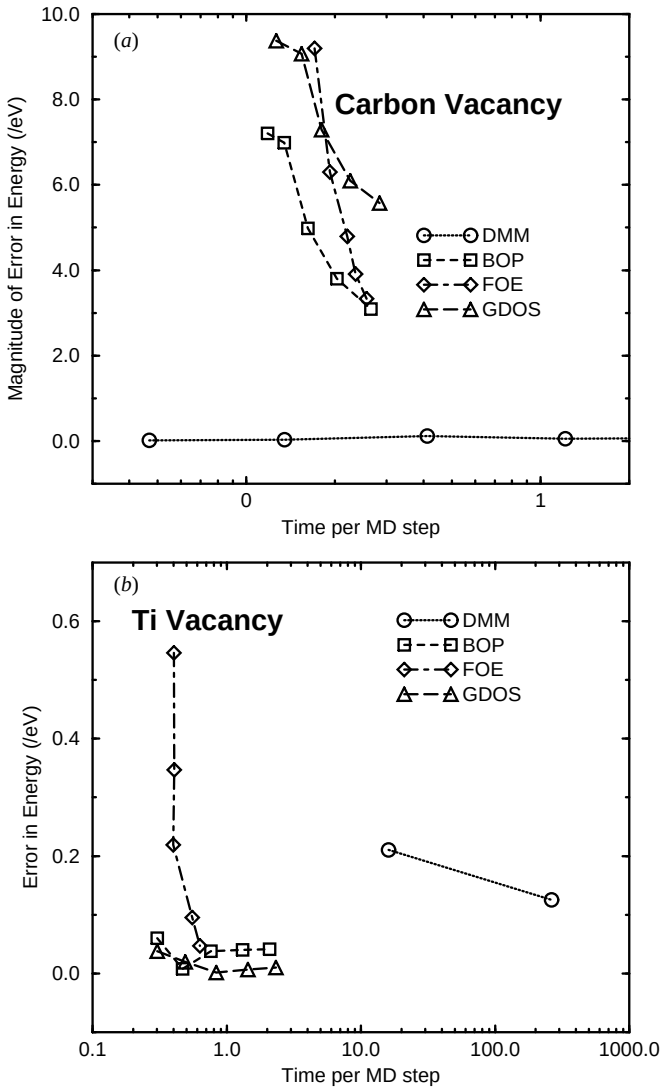

Figure 1. Error introduced by approximation plotted against computer time used with the density-matrix method and three recursion methods BOP, Fermi operator expansion (FOE) and global density of states (GDOS) (Horsfield 1996) for carbon (*a*) and titanium (*b*). Note that a logarithmic scale has to be used for titanium in order for the density matrix result to be included. From Bowler *et al* (1997), reproduced by kind permission.

a small selection of the issues in the electronic structure of metals to which tight-binding methodologies have been applied.

5.1.1. General issues. Metallic systems are characterized by the existence of delocalized states at the Fermi level. This statement has two important consequences for tight-binding calculations. First, states existing at the Fermi level imply incompletely filled bands, and therefore a Fermi surface which does not correspond to the Brillouin zone boundaries. As a consequence, the averaging of physical quantities over the Brillouin zone is much more

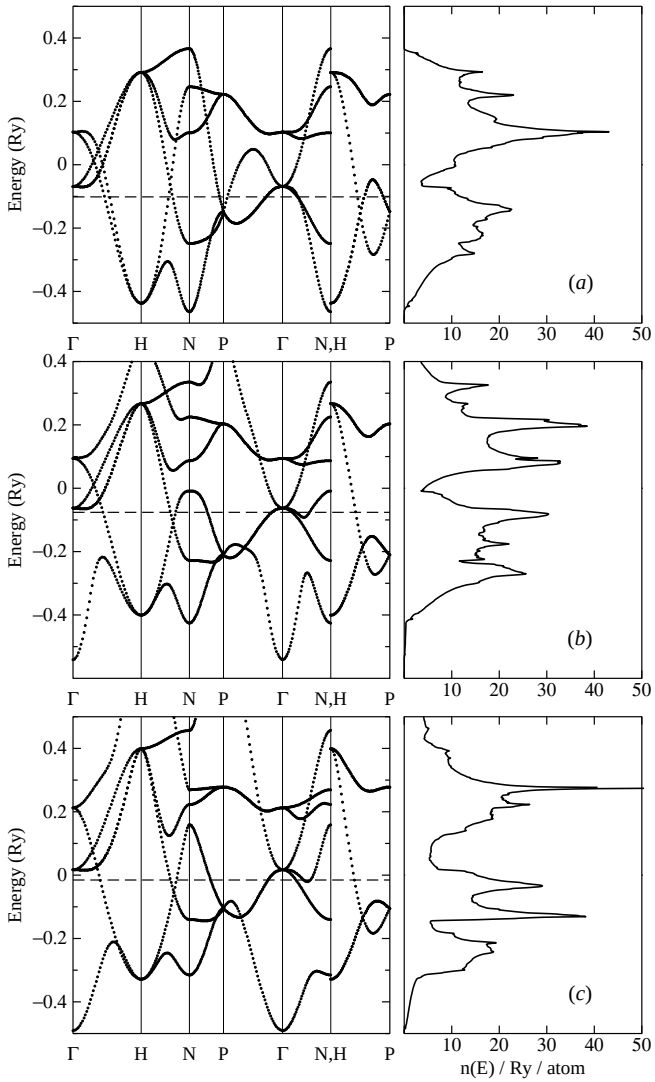

Figure 2. Energy bands and densities of states in Nb. (*a*) d-band tight-binding model, (*b*) spd-band tight-binding model, (*c*) local-density approximation. From Paxton (1996), reproduced by kind permission.

difficult than in the case of insulators; often an order of magnitude more *k*-points have to be used in a metal system than in an insulating one. Secondly, the delocalized states result in a long-range density matrix.

Because of the first of these, real-space-based methods have always been popular in metal tight-binding. Because of the second, density-matrix methods have proved very unsuccessful (this is particularly well illustrated in the comparison of recent linear-scaling techniques by Bowler *et al* (1997), see figure 1). Recursion methods, however, have had considerable success in tight-binding studies of metals. The generally accepted explanation for this success is that the energy of a metallic system is mainly determined by the width of the d-band, which depends only on the second moment of the density of states, and less

by the higher-order terms, which determine the detailed structure. Kress and Voter (1995) demonstrated that, unlike in the case of covalent materials, both cohesive and vacancy formation energy in metals converge rapidly with increasing numbers of moments.

For these reasons, pair-potential methods (reviewed by Carlsson (1990)) and low-moment approximations have been widely used in the study of metals. Moments methods have been described in detail in section 3.2, and low-order approximations are discussed in section 7. Cleri and Rosato (1993) have provided second-moment approximation parameters for fcc and hcp transition metals and Al and Pb simple metals; their paper also includes a useful discussion of the second-moment approximation.

There has been some discussion as to whether it is sufficient to use a d-band only model for transition metals, or whether the s and p states need to be included. Paxton (1996), in his study of Nb/Mo and Nb/Re interfaces, compared an orthogonal, d-band model for bulk Nb with a nonorthogonal, spd-band model. Although the latter reproduced the Nb bands a little better (by comparison with LDA, see figure 2), the elastic constants and the energy-volume curves were more accurately reproduced by the simpler d band model. Again, comparison between the d-band iron calculations of Pastor *et al* (1988) and the spd-band calculations of Bouarab *et al* (1996) (discussed in section 5.1.6 below) suggests that no clear advantage is gained by using the more complex method.

5.1.2. Moriarty generalized pseudopotential theory. Moriarty's generalized pseudopotential theory (GPT) was initially developed for filled and empty d-band metals (Moriarty 1972, 1977, 1982) and later extended to transition metals (Moriarty 1988). The total energy is written in terms of the atomic volume, Ω, and the n-body interatomic potentials, v_n;

$$E_{\rm tot}(R) = E_0(\Omega) + \tfrac{1}{2}\sum_{i,j} v_2(i,j) + \tfrac{1}{6}\sum_{i,j,k} v_3(i,j,k) + \tfrac{1}{24}\sum_{i,j,k,l} v_4(i,j,k,l) + \cdots . \qquad (73)$$

In the Moriarty approximation to GPT (MGPT) (Moriarty 1990), the interatomic potentials are evaluated analytically by making three, tight-binding-like approximations: first that the sp–d hybridization tails can be neglected; secondly that the d–d potentials are short ranged and need only be retained to fourth order; thirdly that the d-bands are canonical, and so the interatomic matrix elements can be described by simple analytic forms. In tight-binding language these approximations would be translated: only d-band matrix elements need be considered; the energy is well described in the fourth-moment approximation; the matrix elements may be written in a simple analytic form.

The resulting form for the four-body potential, v_4, in terms of six interatomic distances and twelve angles, is far from simple. It has, however, been implemented and applied to the cohesive, elastic, vibrational, thermal and melting properties of Mo with considerable success (Moriarty 1994), and more recently to shear strength, point defects and screw dislocations in Mo (Xu and Moriarty 1996). Data and subroutines for the Mo MGPT potential are available from the author (moriarty@sycamore.llnl.gov).

5.1.3. Tight-binding linear muffin tin orbitals (LMTO). LMTO form one of the more widely used atomic-like basis sets for *ab initio* calculations (see, for instance, Andersen 1975, 1984 and Skriver 1984). The fact that the LMTO basis set is minimal (that is, there is only a single basis function per atomic site and angular momentum quantum number lm) immediately makes close connection with the tight-binding method; however, the long range of the LMTO basis functions does not suit the semi-empirical short-range two-centre tight-binding approximation. Andersen and Jespen (1984) proposed a transformation of the LMTO basis set by mixing multipoles based on neighbouring sites into LMTO in order to shorten the

range of the Hamiltonian. Furthermore they showed that in the atomic-sphere approximation (ASA), the transformed basis set can be expressed as a two-centre Hamiltonian matrix with non-zero elements only between first- and second-nearest neighbours (in sc, bcc and fcc structures), and that the two-centre integrals can be written in an almost universal form.

The tight-binding–LMTO–ASA description of any material, therefore, is specified by a structure dependent universally scaling part and a number of potential dependant quantities. Within this approximation, therefore, it is extremely easy to deal with both homogeneous and heterogeneous bonding. Sluiter and Singh (1994) investigated the reliability and transferability of tight-binding parametrizations derived in this fashion for Ni and Ni–Al alloys. They found that sets of parameters produced from fcc, bcc and *A*15 LMTO calculations were almost indistinguishable (suggesting that such parameters are, indeed, highly transferable), and reproduced the major characteristics of the effective pair interaction (EPI) functions for nearest and second-nearest neighbours in $Ni_{0.5}Al_{0.5}$ alloys, whereas the 'standard' band-fitting parametrization (Papaconstantopoulos 1986) failed to do so. They concluded that, especially for alloys, tight-binding–LMTO-derived parameters exhibit greatly improved transferability over band-fitted parameters.

Other applications of tight-binding–LMTO include the copper-cluster studies of Lammers and Borstel (1994), which are discussed below. The majority of tight-binding–LMTO calculations, however, do not adopt the semiempirical methodology discussed above, but use Andersen and Jespen's (1984) mechanism to produce a short-range tight-binding-like basis for fully self-consistent *ab initio* calculations, taking them beyond the scope of this review.

5.1.4. Hybrid methods. A number of studies have been carried out in which tight-binding is used in a hybrid form with the effective medium approximation; see for instance Christensen and Cohen (1993) (iron clusters), Razee and Prasad (1993) (general random alloy clusters), Christensen and Jacobsen (1993) (copper clusters), Engdahl *et al* (1994) (surface alloying).

Hausleitner and Hafner (1992a, b, c) developed a hybrid tight-binding nearly free-electron method for nickel alloys, in which the sp-electron contribution to the energy was calculated in the nearly free-electron approximation, and the d-electron contributions in the tight-binding approximation. The Ni–Al phase diagram (Hausleitner and Hafner 1992c) is well reproduced by this method, as are the general trends in nickel metallic glasses (Hausleitner and Hafner 1992b, Kreuch and Hafner 1995).

5.1.5. Adsorption at metal surfaces. Adsorption of gas molecules at metal surfaces is of great importance in the study of catalysis; the catalytic conversion of carbon monoxide to carbon dioxide is an example of considerable commercial and public interest. Inverse-photoemission studies of CO adsorption on Cu(001) were interpreted in a tight-binding model by Tsuei and Johnson (1992). Pick has published a series of papers using self-consistent recursion tight-binding to model the adsorption of CO on Pd/W(110) and Pt/W(110) (Pick 1995a), Pd/W(110) and Pd/Ta(110) (Pick 1995b, c), Pd/Re(0001) (Pick 1995d), Pt/Ni(111) (Pick 1996) and magnetically live V(001) (Pick and Dreysse 1996).

Homogeneous adsorption, of obvious importance in the study of surfaces, has also been studied in tight-binding; Xu and Adams (1995) found both strong short-range anisotropy and long-range repulsion in W adatom–adatom interactions. Of more general interest is the study of homogeneous adsorption on fcc(111) surfaces, since there are two very similar adsorption sites; the epitaxial (fcc) site and the hcp site. Piveteau *et al* (1992a) showed that complete monolayers always favour the epitaxial (fcc) site; however, single Ir atoms

favour the hcp site (Piveteau *et al* 1992b). Further study of single and multiple adatoms (Papadia *et al* 1996) showed that the relative stability depends on the adcluster size and number of d-electrons per atom. By comparison with the bulk hcp–fcc transition, which occurs at 7.5 e/atom, the single adatom is more stable in the hcp site up to 8.2 e/atom, dimers are more stable in the hcp site up to 7.85 e/atom, and trimers up to 7.5–7.6 e/atom. The reason for the increased stability of the hcp sites for small adsorbed clusters was shown to be significant relaxations in the surface around the adsorbates.

Finally, the reaction of hydrogen at and in metal surfaces has been studied by Stauffer *et al* (1992) (Pd(001) and Ni(001)) and by Kang and Sohn (1996) (Pd(001)). Bifone (1993) studied the density of states in platinum catalysts both clean and hydrogen saturated, and accounted for the Knight shift and a number of key features of the NMR spectra, including the narrowing of the spectra after hydrogenation.

5.1.6. Small metal clusters. Lammers and Borstel (1994) used a tight-binding-LMTO method based on that of Andersen and Jespen (1984) (described above) in conjunction with a recursion solution of the Hamiltonian to study the electronic structure of small copper cluster. These clusters showed a number of interesting characteristics; in particular very large gaps in the density of states for clusters of size corresponding to the jellium model magic numbers $N = 2, 8, 18, 20, 34$ etc and a drop in the ionization potentials at these same magic numbers. Significantly for cluster calculations in metals in general, they also found high-symmetry structures to be Jahn–Teller unstable; the minimum energy structures showed high sphericity but low symmetry. Similar distortions from high-symmetry structures were found for nickel clusters by Menon and Connolly (1994). These results must cast some doubt upon calculations which assume a perfect bulk coordination of atoms.

A further cautionary tale is told by Rey *et al* (1993) in their study of the melting of transition metal clusters. They found the second-moment approximation to be inadequate for small clusters, and that the results obtained depended very sensitively on the empirical data used for the fitting of the parametrization. Zhao *et al* (1993), however, concluded that the second-moment approximation was adequate to study ionization potentials in Ni and NiB clusters. This underlines the truism that a good set of results in an empirical calculation may represent good fortune rather than a good parametrization.

Poteau and Spiegelmann (1993) studied the atomic structure of small sodium clusters, demonstrating the role of pentagonal-based 'seeds' in clusters with $N > 12$. This study was followed (Poteau *et al* 1994) by an analysis of the temperature dependance of geometries in Na_4, Na_8 and Na_{20} clusters. For the two smaller clusters they found a transformation from the solid phase to a liquid-like phase, whereas in Na_{20} the melting took place in two steps, with a preliminary phase in which the outer atoms behave in a liquid-like way (around the 'melting' point of the Na_4 and Na_8 cluster) while the inner icosahedral seed remained 'solid'.

Zhao *et al* (1995a) compared the bonding energies of vanadium clusters calculated in the tight-binding approximation with those calculated in the droplet model and with experimental results. Although the tight-binding calculations are a significant improvement on the classical model, bonding energies for clusters with an odd number of atoms were rather larger than the experimental results; they concluded that this error lay in the description of the delocalized 4s orbital.

Two tight-binding studies of iron clusters, the 3d-electron only calculations of Pastor *et al* (1988) and the 3d-, 4s- and 4p-electron calculations of Bouarab *et al* (1996), both show good agreement with the ionization potentials for clusters $N < 8$ and $N > 22$. In the case

of the larger clusters, the system is tending towards the bulk limit; the interest lies in the smaller clusters. Below $N = 8$ *ab initio* structures (Ballone and Jones 1995) were used by Bouarab *et al* (1996); however, the higher symmetry structures assumed by Pastor *et al* (1988) also acheived a good fit to experimental results; one might therefore assume that structural details are not important. However, between $N = 8$ and $N = 22$ the predicted ionization energies (assuming bcc-like structures) are qualitatively wrong, suggesting that structural details are indeed significant. This question has not yet been adequately resolved.

5.1.7. Finite electronic temperature. An important issue in the attempt to achieve linear scaling for metal systems is the long range of the density matrix at zero temperature. This range can be reduced considerably by working at a large electronic temperature (see, for instance, Bowler *et al* (1997)), which can be simply imposed in density-of-states methods such as those discussed in section (3.2), and rather less easily in density-matrix methods (section 3.3.2). However, the unphysically large electronic temperatures required (equivalent to 12 000 K in some cases) weaken the bonds by promoting electrons from bonding to antibonding states, with resulting errors in dynamics.

Horsfield and Bratkovsky (1996) applied the Gillan approximation (1989) to obtain the zero-temperature-band energy and forces from the high-temperature density matrix. Provided that the electronic temperature used is less than around 10% of the band width, they found this extrapolation to give considerably improved energies and forces, when compared with the exact zero-temperature result.

5.1.8. Magnetism. The Hubbard U approximation of section (2.5.1) can be extended, by making the U term spin dependant, to allow the modelling of magnetic materials in tight-binding. Krompiewski *et al* (1987) changed the on-site elements of the Hamiltonian to:

$$H_{ll}^{\alpha\alpha,\sigma} = \epsilon_\alpha + U^\alpha (n_{\alpha l,-\sigma} - n_{\alpha l}^{\text{para}}) \tag{74}$$

where $n_{\alpha l,\sigma}$ is the number of electrons in orbital α on atom l with spin σ and $n_{\alpha l}^{\text{para}}$ is the same but for the paramagnetic case. Using this formalism they modelled the magnetization of amorphous $Fe_{1-x}B_x$ with no external field and found good agreement with available experimental data. The method was later also applied to amorphous $Fe_{1-x}B_x$ and amorphous Fe (Krompiewski *et al* 1988) where the distribution function for local magnetic moments compared well with experimental hyperfine splitting measurements, and to amorphous $Fe_{1-x}Zr_xH_y$ (Krompiewski *et al* 1989), where good agreement with experiment was again found, and the effect of hydrogen on the magnetization was investigated.

The treatment was later expanded so that the local polarization direction, as well as magnitude, could vary (Krey *et al* 1990). The general Hamiltonian element now looks like this:

$$H_{l,\alpha,s;m,\beta,s'} = H^0_{l,\alpha,m\beta}\delta s, s' + \delta_{l,m}\delta_{\alpha,\beta}[(U_{l,\alpha}/2)(n_{l,\alpha,\uparrow} + n_{l,\alpha,\downarrow} - 2n^p_{l,\alpha})\delta_{s,s'} - \boldsymbol{h}_{l,\alpha} \times \sigma_{s,s'}]. \tag{75}$$

The local self-consistent effective field, $\boldsymbol{h}_{l,\alpha}$, is defined by:

$$h^i_{l,\alpha} = (U_{l,\alpha}/2) \sum_{\nu=1}^{N_{\text{el}}} \sum_{s,s'} (c^{(\nu)}_{l,\alpha,s})^* (\sigma_i)_{s,s'} c^{(\nu)}_{m,\beta,s'} \tag{76}$$

and σ_i is one of the three Pauli spin matrices. Using this extended theory, they examine the spin-glass states of amorphous Fe at different densities and amorphous $Fe_{1-x}Zr_xH_y$.

This method has also been applied to Mn, γ-Ni_3Mn and γ-FeMn (Süss and Krey 1993), though there is sufficient uncertainty about these systems to prevent clear comparisons with previous work.

This type of approach has been applied to transition metal monolayers (Stoeffler *et al* 1992, Mokrani *et al* 1992, Pick and Dreysse 1992, Moos *et al* 1996, Pizzagalli *et al* 1996), thin films (Pick and Dreysse 1993, Haroun *et al* 1994, Pick *et al* 1994, Dreysse *et al* 1994a, Dorantes-Dávila and Pastor 1995, Pastor and Dorantes-Dávila 1996), superlattices (Itoh *et al* 1993a, b, Tsymbal and Pettifor 1996), interfaces and surfaces (Perezdiaz and Muñoz 1994, Dreysse *et al* 1994b, Mpassimabiala *et al* 1996) and clusters (Zhao *et al* 1995a, b, Bouarab *et al* 1996, Andriotis *et al* 1996a, b).

The study of magnetism requires a very accurate description of the density of states at the Fermi level (see Stoner 1938). In their consideration of the applicability of tight-binding to the theory of magnetism, Haas *et al* (1996) make three important points. First, low-order moment approximations cannot reproduce the correct detailed behaviour of the density of states, and are therefore insufficient when the Fermi level lies in an energy regime with strong variations in the density of states. Secondly, an accurate parametrization is required with reliable scaling behaviour; fitting to a set of tight-binding–LMTO–ASA results at a number of lattice parameters was their solution. Thirdly, the calculation must be done self-consistently (previously discussed by Lorentz and Hafner (1995)); they argue that local-charge neutrality sufficiently describes the electrostatic self-consistency.

5.2. Semiconductors

The covalent bonding inherent in semiconductors makes them superb candidates for modelling with tight-binding. In this section, we will present an overview of those areas covered by researchers for the last 25 years, identifying for each the issues to be addressed in modelling each area, and selecting highlights to illustrate the research. Interested readers should also consult the recent review by Wang and Ho (1996).

As has been alluded to already, a truncated density-matrix method is particularly suitable for modelling semiconductors. In the perfect bulk structure, the localization of bonding makes any of the linear-scaling methods described ideal methods for calculations. Defects or surfaces, however, introduce narrow features in the density of states which require extremely high resolution, and therefore a large number of moments, as a result of which moments methods break down; however, these states are localized, so the density-matrix method retains its advantages. Such methods are clearly a method of choice for these systems.

5.2.1. Bulk properties. There are only a few issues to be addressed when modelling bulk properties. Foremost amongst these is the parametrization, which has received considerable attention. The earliest parametrizations were by Chadi and coworkers (Chadi and Cohen 1975, Chadi and Martin 1976, Chadi 1977, 1979a) who fitted to data from x-ray and photoemission band spectra, and *ab initio* band structure results and produced parametrizations for silicon and germanium, and for gallium arsenide and other III-V semiconductors. They acknowledged that their nearest-neighbour, orthogonal sp^3 parametrization could only reproduce the valence band, and fitted to this and elastic constants and phonons. The Chadi silicon parametrization (Chadi 1979b) is still widely used.

Froyen and Harrison (1979) compared the band structures produced by LCAO theory and free-electron theory, and using certain selected differences in energy from these two methods

found a universal formula for hopping elements in terms of a dimensionless coefficient $\eta_{ll'm}$,

$$V_{ll'm} = \eta_{ll'm}\frac{\hbar^2}{md^2} \tag{77}$$

where d is the distance between ions and m is the mass of the electron. This method works extremely well for semiconductors in the tetrahedral structure, but not for close-packed structures (Paxton *et al* 1987). The whole bond orbital model of Harrison is well described in his book on the subject (Harrison 1980), where he derives the d^{-2} scaling of hopping parameters.

One of the most successful sets of parameters for tetrahedral semiconductors is that of Vogl *et al* (1983). This is a classic paper for those wanting to model tetrahedral semiconductors, for as well as giving parameters for 12 of these, it also describes a general method for fitting to high-symmetry points of the band structure for new materials. They used an sp^3s^* basis in order to include the lowest conduction band in the fit. However, their parametrization does not include scaling rules for the hopping parameters (referring instead to the d^{-2} scaling of Harrison; it should be pointed out here that *assuming* d^{-2} scaling leads to very poor reproduction of elastic constants—the scaling should ideally be fitted to *ab initio* or experimental data).

Goodwin *et al* (1989) addressed the problem of transferability: the ability of a parametrization to produce accurate results in a range of different environments. Their parametrization for silicon accurately fits both the tetrahedral phase of silicon and the close-packed, metallic phases. A similar parametrization, though with different scaling forms, has been developed by Sawada (1990). For more structural calculations, such as elastic constants, the fitting must be performed for the scaling parameters, fitting to elastic constants. A danger to be aware of with scaling is that of cut-off; the nearest neighbour (or second-nearest neighbour) distance which is easily defined in the perfect crystal can become extremely problematic when deformations or other crystal structures are considered. Simply put, the problem is that a small change in a position of the atoms may result in a new shell being included in Hamiltonian range interactions, particularly in close-packed or highly distorted structures. This is generally dealt with by introducing a smooth cut-off function which allows atoms to enter and leave the Hamiltonian range without a discontinuous change in energy (see, for instance, Xu *et al* (1992)).

Band-structure calculations which have been performed range from the earliest work of Slater and Koster (1954), who modelled diamond and InSb, and would have considered Ge had the *ab initio* calculations been available, to more recent calculations for new classes of semiconductor. Teng *et al* (1991) studied the effects of ordering on the band structure of III-V alloyed semiconductors. Hasbun *et al* (1987) studied $Al_xGa_{1-x}As$, using the Vogl *et al* parametrization, and found good agreement with experimental results for variations of x, while Capaz *et al* (1993) investigated the effect of pressure and composition on the band gap of $Al_xGa_{1-x}As$. Onwuagba (1994) studied cubic BN and BP using tight-binding–LMTO, and successfully reproduced previous results, while Jenkins and Dow (1989) modelled the electronic structure and doping of InN, $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$. Ekpenuma and Myles (1990) produced bandstructures for II-VI zinc-blende compounds and Sieranski and Szatkowski (1993) produced a model for the band structure of Zn_3P_2, and have extended it to Cd_3P_2, Cd_3As_2 and Zn_3As_2 (Sieranski *et al* 1994). Lee and Dow (1987) studied the band structure of $Pb_xSn_{1-x}Te$ for varying x.

Calculations of elastic constants and related quantities are harder to perform using tight-binding, and are generally more dependant on the quality of the fit. For instance, with nearest neighbour, orthogonal tight-binding, fitting to B, C_{11}, C' and C_{44} has proved extremely hard

for Si (Chadi 1978). Wang *et al* (1990) successfully reproduced the silicon anharmonic frequency shifts and phonon line widths as a function of temperature when compared to Raman spectroscopy. Work on the deformation potentials of III–V semiconductors has had reasonable success, when fitting to the X-point (Muñoz and Armelles 1993). Similar work has been performed on Zn-based II–VI compounds by Bertho *et al* (1991) who calculated the hydrostatic and uniaxial deformation potentials for ZnTe, ZnSe and ZnS using a self-consistent tight-binding model. They achieved good agreement with experiment, and correctly predicted the band offset in the strained ZnS/ZnSe system. A review of elastic constant calculations in semiconductor compounds covers *ab initio* and tight-binding calculations (Chen *et al* 1992).

Stability and strain have also been explored using tight-binding. Majewski and Vogl (1986) have used a universal tight-binding model for sp-bonded semiconductors and insulators, and successfully examined the stability of crystal types and phase transformations. Duran *et al* (1987) have successfully reproduced the experimental band bending at an ideal ZnSe/Ge interface, using a self-consistent tight-binding model. Xu (1993) calculated the bond lengths and lattice constants for unstrained Ge_xSi_{1-x} alloys and strained Ge_xSi_{1-x}/Si(001) and Ge_xSi_{1-x}/Ge(001) systems, and found good agreement for the lattice constants when compared with elastic theory. Using a method developed by Guinea, Sánchez-Dehesa and Flores (Guinea *et al* 1983, Sánchez-Dehesa *et al* 1984) for Schottky barriers at semiconductor surfaces (specifically for Si–Ag and Si–H–Ag), Platero *et al* (1986) examined band offsets at (100) junctions between GaAs–AlAs and InAs–GaSb. The method consists of tight-binding with a Hubbard U added, but one which varies from layer to layer (though normally it is only applied to the interface layers in the two materials under consideration) and is varied until certain charge and dipole conditions at the interface are met, and provides results which are within 0.2 eV of experimental measurements. The method has also been applied to GaAs–Ge(110) interfaces (Muñoz *et al* 1986), III–V and II–VI(110) heterojunctions (Muñoz *et al* 1987) and GaAs/AlAs superlattices (Saito and Ikoma 1992) with reasonable agreement with experiment.

5.2.2. Surfaces. As the surfaces of semiconductors are particularly accessible to experimental techniques, a wealth of investigations have taken place over the years. Many of these are summarized, along with calculations using different methods such as density functional theory and quantum chemical methods, in an excellent review by LaFemina (1992). Wilson *et al* (1990) conducted a study comparing tight-binding to empirical potentials for Si(110), Si(001) $-$ 2 $\times$ 1 Si(111) $-$ 2 $\times$ 1 and Si(113) $-$ 1 $\times$ 1, and showed that, while tight-binding produced results in good agreement with experiment and other theoretical work, the empirical potentials were unsuitable for surface calculations. This, and other results over many years has given great confidence in tight-binding studies of semiconductor surfaces.

The easiest and most commonly studied surfaces are those of group IV semiconductors. Chadi performed large amounts of the early work in this field, reviewed in Chadi (1989, 1994). He correctly predicted the Si(001) $-$ (2 $\times$ 1) buckled dimer reconstruction (Chadi 1979b) The Si(001) surface is well known, and recent studies have tended to study the more interesting features such as defects (Owen *et al* 1995) and steps (Chadi 1987), though there is a recent study of the electronic structure of the p(2 $\times$ 2) and c(4 $\times$ 2) reconstructions by Low and Ong (1994). Owen *et al* (1995) used tight-binding to obtain the structures and energetics for the single missing dimer defect on Si(001), and then used these structures to produce LDA charge densities which they compared with experiment.

Qian and Chadi (1987) predicted the surface energy for the dimer-adatom-stacking fault (DAS) model of the Si(111) $-$ (7 $\times$ 7) surface, which identified this model as having the lowest surface energy for the Si(111) surface. Mercer and Chou (1993) studied Si(111) and Ge(111) reconstructions, both unstrained and strained. They found that the DAS models were lower in energy for Si(111) than simpler (2 $\times$ 2) or c(2 $\times$ 8) structures, but that the Ge(111) surface goes to the simple c(2 $\times$ 8) structure, both of which results are in agreement with experiment. Interestingly, they found that 2% of strain was enough to destabilize the c(2 $\times$ 8) model with respect to the DAS models. Olmstead and Chadi (1986) studied the temperature dependence of optical absorption on the Si(111) $-$ (2 $\times$ 1) surface, finding good agreement with experimental evidence, and further corroboration for the π-bonded model for that surface. Chadi and Chelikowsky (1981) studied step formation on Si(111).

The (113) and related surfaces have proved to be very-low-energy surfaces, and may have implications for device technology. Chadi (1984a) worked on the structure of Si(211), (311) and (331); Feng *et al* (1996) have studied the structure of Si(113) and the disordered Si(113) surface (Wee *et al* 1996).

Group III–V and group II–VI surfaces are harder to model, as charge transfer effects become significant. Avery *et al* (1995) identified the GaAs(001)-β(2 $\times$ 4) surface structure by comparison between scanning tunnelling microscope (STM) images and tight-binding modelling of the kinking energy of the reconstruction. Figure 3 shows STM images of the GaAs(001) surface. Figure 3(*a*) shows the beta phase, which shows long-range order, while (*b*) shows the gamma phase, with high disorder. The beta phase was successfully identified by a calculation of kinking energy (shown in figure 4 schematically). The structure found, with an out-of-phase As dimer in the third layer, has a kinking energy of over 1 eV; the gamma phase is the same structure, but with a small amount of As added, which results in a *negative* kinking energy. Ren and Chang (1991) examined both the As rich (2 $\times$ 4) and the Ga rich (4 $\times$ 2) surfaces, and compared calculated dielectric functions with experimental data. Chan and Ong (1991) studied the electronic structure of the reconstructed InP(001) surface, and concluded that the (4 $\times$ 2) reconstruction is most likely to be the optimum structure.

Godin *et al* (1992) found that for six III–V and II–VI compounds, the structure and potential-energy surface of the (110) have a universal relation, scaled by lattice constant. Wang and Duke (1987) studied the cleavage surfaces of ZnS in the zincblende and wurtzite structures. For the ZnS(110) surface, they found good agreement for the reconstruction and electronic states with experiment. The $(10\bar{1}0)$ and $(11\bar{2}0)$ surfaces have no data available for comparison, but the reconstruction and electronic states have been calculated.

Chadi (1984b) found good evidence for a vacancy-induced (2 $\times$ 2) reconstruction on the Ga-terminated GaAs(111) surface, and a rather more complex one for the As-terminated (111) surface (Chadi 1986).

There is considerable interest in adsorbates at semiconductor surfaces, as these can give insight into surfactant effects, heterostructure growth and Schottky barrier formation, to name a few. Mailhiot *et al* (1984) studied Sb overlayers on the (110) surfaces of III–V semiconductors, finding a novel type of binding for the Sb chains (involving p^2 bonding). Schenter and LaFemina (1992) extended this study to the vibrational spectra of GaAs and InP(110) surfaces, both clean and with an Sb overlayer, with good, quantitative agreement with He scattering data. LaFemina *et al* (1990) studied the adsorption of column V on GaAs(110), and predicted a new structure, which does not seem to match LEED data. More discussion of this can be found in LaFemina (1992).

Allan and DeLerue (1992) studied samarium adsorbed on GaAs(110), and found good agreement with photoemission spectroscopy, while Wang *et al* (1994) modelled a

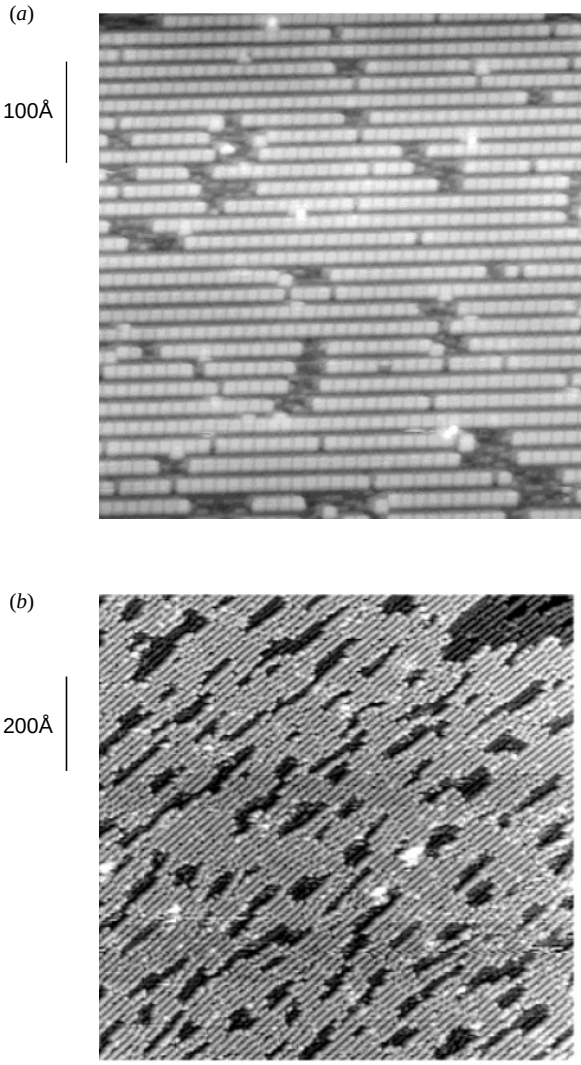

Figure 3. STM images of the GaAs(001)-(2 × 4) surface. (*a*) The beta phase, note that the long, dark lines are extremely straight. (*b*) The gamma phase; note the high degree of disorder. From Avery *et al* (1995), reproduced by kind permission.

magnesium overlayer on GaAs(110) and found electronic behaviour in good agreement with STM data. A review of alkali metals on GaAs(110) gives details of calculations and comparisons with experiment (Bechstedt and Scheffler 1993). The next stage of this area, formation of Schottky barriers, has been reviewed by Flores *et al* (1993).

5.2.3. Defects and dislocations. The study of defects and dislocations provides different challenges. For heterogeneous defects, there is the problem of fitting a reliable parametrization and the possibility of charge transfer, while for homogeneous defects, there is structural distortion and electronic structure which can be far from the bulk behaviour. Some of these problems are addressed in the papers below. The study of defects and

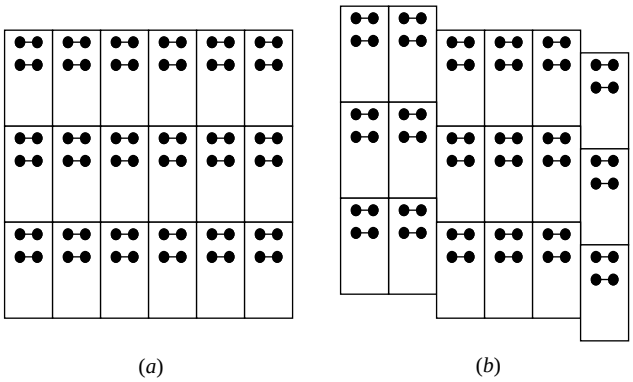

Figure 4. Schematic diagram of the kinking-energy calculation performed for the GaAs(001)-(2×4) surface. From Avery *et al* (1995), reproduced by kind permission.

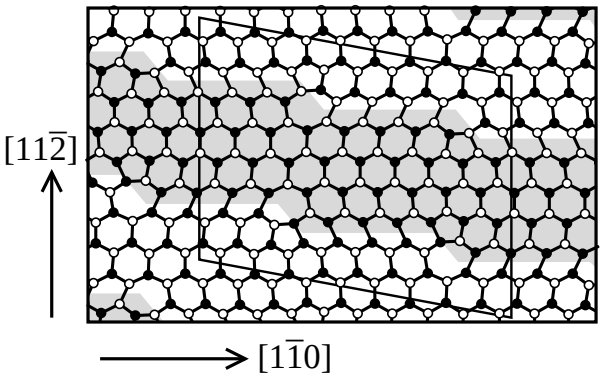

Figure 5. Supercell containing 864 silicon atoms used for simulating kinks and kink–soliton complexes in the 90° partial dislocation in silicon. From Nunes *et al* (1996), reproduced by kind permission.

dislocations is a field to which tight-binding is well suited, as it is frequently the case that the deviations from ideal bonding are large enough that empirical potentials are not sufficiently accurate, but the system size required to isolate the defect or dislocation makes the use of *ab initio* calculations prohibitively expensive. This is illustrated by the 864 atom calculations of kinks in the 90° partial dislocation in silicon of Nunes *et al* (1996), figure 5.

Defects are important for their electrical characteristics, which are also affected by the structural relaxations associated with them. Chadi (1990) studied the possible combinations of oxygen atoms and vacancies in Si which could form thermal donors. He fitted his parameters for Si–O and O–O interactions to α–SiO_2 and the O_2 dimer, and concluded the V–O_2, V–O_4 and an interstitial O_2 were the only candidates which matched experimental evidence. Rasband *et al* (1996) looked at defects in Si caused by B, and fitted to *ab initio* data. They found many interstitial structures, including new di-interstitials. Li and Myles (1991) used molecular dynamics to study the relaxation around a substitutional defect in semiconductors. They show results for impurities in GaP and Si, and compare with experiment. Wang *et al* (1991) studied the properties of point defects in Si, and obtained good agreement with first-principles data. They found that the Goodwin, Skinner, Pettifor parametrization for silicon described the Jahn–Teller distortion at a vacancy well. Song *et*

al (1993) also studied point defects in Si. They found that the common assumption that there is no relaxation is not justified.

Kobayashi *et al* (1982) studied substitutional defects in $Hg_{1-x}Cd_xTe$, both on cation and anion sites. They examined the chemical trends for a wide variety of substitutional defects, and found good agreement with available data. Zhu *et al* (1992) used tight-binding to confirm the identification of a deep-level trap. They found that oxygen in $In_xGa_{1-x}As_yP_{1-y}$ grown on GaAs gives a new level, which has been identified as an anion-site defect. Jenkins *et al* (1992) investigated N vacancies in $Al_xGa_{1-x}N$ for varying x. They explained electronic behaviour as a function of x, and successfully identified N-site Zn defects as the source of a deep level identified by extrinsic photoluminescence. Seong and Lewis (1995) used molecular dynamics to investigate point defects in GaAs for both Ga and As vacancies and antisites. They achieved good agreement with both experiment and *ab initio* calculations.

Examples of charge transfer in defects, and how the modelling has taken this into account, can be found in Sankey and Dow (1983) and Lee *et al* (1985). They studied tetrahedral-site interstitial impurities (Sankey and Dow 1983) and neutral and singly ionized substitutional impurities (Lee *et al* 1985) in silicon via a perturbation method. They applied the Hubbard U to the impurity only in determining the impurity potential. They obtained good agreement with experimental results for Al^+, S, Se and Te, as well as investigating the chemical trends for different impurities.

Dislocations and grain boundaries are of fundamental importance to crystal quality, and understanding the electronic and atomistic structure can aid characterization of these defects. While it is relatively easy to obtain atomistic structures from electron microscopy (as will be seen from the papers below), the electronic structure often only comes from modelling. Thomson and Chadi (1984) studied the electronic structure of a high-angle tilt grain boundary in Si. Using a structure based on electron microscopy data, they found no states in the gap, which they attributed to the lack of dangling bonds at the boundary. Bigger *et al* (1992) compared results from *ab initio* and tight-binding calculations for two 90° partial dislocation structures in silicon, and found that a nearest-neighbour parametrization successfully reproduced the more accurate results. Nunes *et al* (1996) modelled the structure and barriers to diffusion of kinks in the same dislocation. Liu *et al* (1995) have investigated elastic and electronic properties of edge dislocations in Si. Paxton and Sutton (1989) modelled three grain boundaries in Si (the $(11\bar{2})$, the $(1\bar{3}0)$ and the (111) boundaries) using the recursion method (dealt with in section 3.2). They found differences with experimental measurements of densities of states, and discuss localization of electronic states at grain boundaries. Kohyama and Yamamoto (1994a) modelled twist-grain boundaries in Si, and compared the results with tilt-grain boundaries. They found that the twist-grain boundaries have a much higher interfacial energy, which they attributed to the distortions which are inevitable for twist boundaries but are not present in tilt boundaries. They have also studied the origin of band tails, and other electronic structures, in these twist-grain boundaries (Kohyama and Yamamoto 1994b). Mauger *et al* (1987) studied the (211) $\Sigma = 3$ grain boundary in Ge. There has been little work on dislocations in III–V semiconductors, but Masudajindo (1994) studied misfit dislocations in ZnSe/GaAs heterostructures, and the electronic states associated with them.

If the substance being modelled is polar and the interface is also polar, then some form of charge transfer must be modelled. Kohyama *et al* studied non-polar (1990a) and polar (1991) versions of the $(122)\Sigma$ grain boundary in SiC using the self-consistent tight-binding model described in section 6.2.2. They found that charge transfer effects are important, and that, while all the structures considered are potentially stable, the polar interfaces are more stable when an atomic chemical potential for Si and C is considered.

5.2.4. Superlattices. Studies of superlattices are becoming more common, as they are used increasingly by the semiconductor industry. The key feature of interest for a superlattice is the electronic structure, for which it is being used.

Superlattices are gaining in importance in the semiconductor industry, because of their electronic structure (which can be tuned by varying the physical parameters of the superlattice). For this reason, as well as pure scientific interest, there is increasing interest in them. Providing that the constituent materials are well characterized, the issues that face the modeller are related to the interfaces between these materials. If there is significant charge transfer, then a self-consistent implementation of tight-binding may be required. Another issue at the interface is that of wrong bonds, and relaxation, which can be addressed in the usual manner for parametrizing. There is a review of the theory of semiconductor superlattice electronic structure, which covers simulation techniques other than tight-binding, and examines the subject in more detail than is possible here (Smith and Mailhiot 1990). Another review by Velasco and Garciamoliner (1994) studies the electronic properties of superlattices and quantum dots grown along [100] and [311] directions, and uses only tight-binding for the calculations.

Group IV superlattices generally consist of Si and Ge, and are used to obtain particular band gaps or optical properties. Ren *et al* (1992) studied the effect of uniaxial stress on Si_xGe_{1-x}/Ge(111) superlattices. By varying the value of x, and also applying external stress, they found that the Ge quantum wells could always be caused to luminesce, by forcing the Ge conduction band minimum to be direct gap. Tserbak *et al* (1993) used a three-centre tight-binding model to examine Si/Ge strained superlattices. They also found that, by applying appropriate strain, the superlattices could be made direct gap. Kumar and Singh (1988) studied the electronic structure of Ge/ZnSe, Ge/AlAs and Ge/GaAs(110) interfaces and superlattices, and achieved good results in comparison with available data.

Group III–V superlattices have more scope for variety, and are used as quantum wells as well as for the tailored electronic properties which are possible: Muñoz *et al* (1989) examined the electronic structure of AlAs–GaAs superlattices; Arriaga *et al* (1991) examined GaAs/GaP(001) superlattices, varying the thicknesses of the layers, and finding good agreement with optical transitions; Kumagai *et al* (1987) examined the direct/indirect gap and optical properties of GaP/AlP(001) superlattices, varying the period and number of layers. The electronic structure of $(AlAs)_k/(GaAs)_l/(AlAs)_m/(GaAs)_n(001)$ superlattices has been modelled by Fernandez-Alvarez *et al* (1996) for many combinations of k, l, m and n; and Osotchan *et al* (1994) have conducted an extensive study of a double barrier superlattice of $Al_xGa_{1-x}As/Al_yGa_{1-y}As/GaAs$, examining electronic structure as a function of slab thickness and alloy composition. They specifically examined a structure used for infra-red photodetectors.

Arriaga *et al* (1993a) studied the effects of period length and the composition of the top layer on the surface state of a semi-infinite GaAs/GaP(100) superlattice. This area is examined in more detail by Arriaga *et al* (1993b).

The band gap of II–VI semiconductors and their superlattices is in the energy range area where blue or green light may be generated, which is of technological importance. Marshall and Wilson (1994) examined the electronic structure of $(ZnSe)_m(ZnS)_n$ strained-layer superlattices. Shen *et al* (1990) modelled superlattices of GaAs and $(GaAs)_{1-x}(ZnSe)_x$ with ZnSe, and found that there was a possibility of band gaps in the blue-green region. Wang *et al* (1995) studied $(ZnSe)_n/(Ge)_m$ superlattices, and examined the optical and electronic properties and their dependence on period (n, m).

5.2.5. Amorphous and liquid semiconductors. Non-crystalline semiconductors have attracted a great deal of attention in recent years. From a theoretical perspective, the lack of periodicity in amorphous and liquid materials introduces structural and electronic properties which make them quite different from conventional crystalline systems. From a more practical point of view, amorphous materials have found numerous applications. This is the case of hydrogenated amorphous silicon, which is a cheap and easy-to-grow material used in solar cells, electrophotography, thin-film transistors, etc. Diamond-like amorphous carbon, to name another, has found applications as a coating material, given that it forms films of great hardness which are transparent and chemically inert. Tight-binding methods have also been used in the simulation of liquid semiconductors. Many of these materials melt at very high temperatures (carbon, for instance, melts at about 5000 K), which makes the experimental study of these systems rather difficult. Simulations are thus very useful to gain insight into the behaviour of liquid semiconductors.

The key problems for modelling amorphous semiconductors are producing the sample, and aspects of the parametrization, especially interaction cut-offs. The production of samples is discussed in some detail in many of the references given below; the cut-off is generally set at an arbitrary point which does not interfere with radial distribution functions, though the use of a smooth cut-off function as discussed in the section above on bulk calculations is sometimes used. A good review of the work to that date which is useful for an overview of the subject was written by Robertson (1983).

The most interesting question for a-Si is the presence of states in the band gap. Nichols and Winer (1988) studied the eigenstates at the band edge of amorphous silicon, using a continuous distorted network with only four-fold coordinated atoms. They found localized states at the band edges which were due to distortions in the bonding, but could not compare the degrees of contribution and localization due to dangling or floating bonds and distorted four-fold bonds. Biswas *et al* (1989) studied the effects of dangling and floating bonds, (i.e. three-fold and five-fold coordinate Si atoms) using Chadi's silicon parametrization, and found that dangling bonds have a large, localized contribution to the density of states in the band gap, while floating bonds are less localized and contribute less. The same effect was studied by Martin-Moreno and Verges (1989). Mercer and Chou (1991) studied a sample with a high density of floating bonds, using a simple nearest-neighbour orthogonal tight-binding scheme and a third-nearest-neighbour non-orthogonal scheme, and found that there was little difference between these parametrizations except with respect to localization of a dangling bond (which is to be expected given the different ranges of the Hamiltonian). The electronic structure of defects and impurities in a-Si has been investigated by Agrawal *et al* (1990). The effect of different sample preparation techniques on structural and electronic properties of a-Si were investigated by Servalli and Colombo (1993), who found that there was a clear dependence for both of these on the preparation technique, while Holender and Morgan (1991, 1992) studied the electronic structure of large samples (up to 13 824 atoms) using Chadi's parametrization, and investigated different models and means of producing samples. One such sample is shown in figure 6; the difference between this and a continuous distorted network is immediately apparent. It is equally apparent that calculations of this size are way beyond the reach of *ab initio* methods, and that tight-binding therefore provides the only mechanism by which quantum effects can be included. In a similar vein, Colombo and Maric (1995) studied the crystal to amorphous transition in Si due to defect formation at high temperatures and De Sandre *et al* (1996) studied the elastic constants in a-Si.

When a-Si is hydrogenated, the states in the band gap disappear. Holender *et al* (1993) hydrogenated previously generated samples of clean a-Si, and found that the band gap reappeared as three-fold and five-fold coordinate Si atoms were returned to four-fold

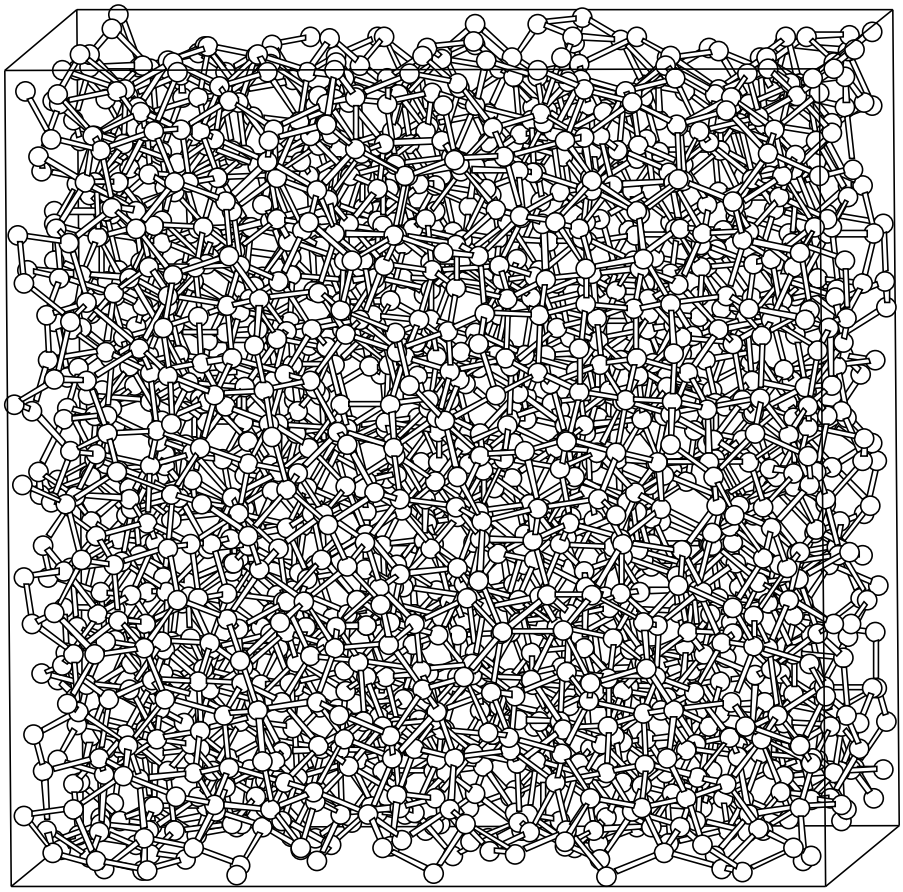

Figure 6. A sample of a-Si produced by a combination of anneals and quenches (Holender and Morgan 1991). From Holender and Morgan (1994), reproduced by kind permission. The original figure emphasized the coordination number by colour.

coordination. This finding was confirmed and expanded upon by Holender and Morgan (1994), who compared models of a-Si and a-Si:H, finding that it is the dangling bonds which give rise to the density of states in the gap for a-Si, but that these *localized* states do not yield conductivity. Lanzavecchia and Colombo (1996) studied hydrogen in amorphous Si, looking at both bonding and diffusion, and Stutzmann *et al* (1989) studied the structural and optical properties of amorphous, hydrogenated SiGe alloys. A thorough study of a-Si–C and hydrogenated a-Si–C has been detailed by Robertson (1992).

Molteni *et al* (1993) have investigated liquid GaAs, and extended this study to quenching the melt to form amorphous GaAs. They also generated amorphous GaAs by quenching a liquid sample from 1600 K (Molteni *et al* 1994a), and find both good agreement with experimentally deduced structure, and an interesting comparison with a supercooled liquid at 800 K. They discuss the parametrization which they have used for liquid GaAs and amorphous GaAs (both of which present difficulties to modelling due to the demands of molecular dynamics with widely varying bond lengths and angles) thoroughly (Molteni *et al* 1994b). Seong and Lewis (1996) studied the effects of chemical disorder on the band gap and the density of a-GaAs.

Wang and Ho (1993) have reported the modelling of amorphous carbon networks produced by quenching liquid carbon during a tight-binding molecular dynamics simulation. In this way they produced a model structure which contained 74% fourfold coordinated atoms. Other properties derived from this sample, such as the radial distribution function, were in good agreement with experimental data.

Very frequently, simulations of liquids are performed as a preliminary step for the study of amorphous systems, i.e. as a seed for the production of amorphous structures by quenching, but in some other cases liquid semiconductors themselves have also been the focus of attention. For example, Virkkunen *et al* (1991) have reported a molecular dynamics simulation of liquid silicon at 1740 K. The interactions were described using the model developed by Goodwin *et al* (1989). The average coordination was found to be 6.4. This high coordination is characteristic of metalic systems. Indeed, the density of states showed no gap at the Fermi level, indicating that the model used correctly predicts liquid silicon to be a metal. Other structural and dynamical properties, such as the bond-angle distribution, phonon spectrum and diffusivity, were in good agreement with first principles calculations. This is in contrast with the predictions obtained using empirical potentials, such as the Stillinger–Webber model (1985), which gives a much poorer description of the liquid structure of silicon.

Wang *et al* (1993) have used a similar scheme to simulate liquid carbon at a range of densities. These authors used the model developed by Xu *et al* (1992), and also found good general agreement of their results with those of first-principles calculations. Rather different structures were observed in the low- and high-density regimes. At low densities (1.2 g cm^{-3}) the majority of atoms (68%) were found to be twofold coordinated. As the density was increased, the number of threefold coordinated atoms gradually increased and becomes dominant. In the high density regime (4.4 g cm^{-3}) it is the fourfold coordinated atoms that dominate in the structure. Radial and bond-angle distribution functions were also obtained, which reflected the structural changes observed as the density was increased.

5.3. Fullerenes

Ever since the discovery of C_{60} by Kroto *et al* (1985), there has been considerable interest in buckyball structures from both the experimental and theoretical perspectives. More recently, the discovery of nanotubes (Iijima (1991)) has also attracted a great deal of attention, not only due to the fact that, as for buckyballs, these are novel structures but also because they raise the possibility of new materials with important properties of flexibility, hardness and electronic conduction, due to which they have already begun to find technological applications.

There is a substantial amount of tight-binding work on the structural, electronic and dynamical properties of carbon fullerenes. One of the first applications of the non-orthogonal tight-binding scheme of Menon and Subbaswamy (1991) (described in more detail in section 6.1) was to the study of the dynamical properties, oxygen chemisorption and electron doping of C_{60}. They found that the frequency of the breathing and tangential vibrational modes were somewhat higher than the experimental values available for solid C_{60}, perhaps not surprisingly, since no attempt was made to fit the model specifically to the properties of C_{60}. However, the results obtained by them were in close agreement with other theoretical results. They found that the preferred adsorption site of an oxygen molecule was over the double bond, i.e. between two hexagons. The adsorption provoked a stretching of the C–C bond supporting the oxygen atom from 1.39 to 1.47 Å, resulting in a weakening of the bond. The doping of C_{60} with electrons resulted in an expansion of the molecular

radius and a softening of the normal modes.

In a series of papers, Ho and coworkers have performed an extensive study of many fullerene cages. They have studied the stability of all C_n clusters with even n in the range 20–94 (Zhang *et al* 1992a). They have also reported the minimum energy structures of C_n clusters for even n in the range 20–70 (Zhang *et al* 1992b) and 72–102 (Zhang *et al* 1993a). The determination of minimum-energy structures of fullerene cages is a problematic task, as these structures possess many potentially competing low-energy structures. In order to generate sensible starting guesses for each cluster, a *face-dual* network approach was taken, which has the advantage of avoiding the high-energy barriers between competing low-energy structures associated with the breaking and forming of bonds, thus facilitating the rapid location of low-energy structures.

Zhang *et al* (1992b, 1993a) found that with the exception of C_{22}, all the studied structures consisted of hexagons and 12 pentagons, and of all the clusters, only C_{60} and clusters C_{70} and larger obeyed the isolated pentagon rule. Most structures were found to have low symmetry, indeed some were found to have no symmetry at all, thus contradicting the view that C_n clusters are highly symmetric structures.

Once the most stable forms of C_n clusters ($20 \leqslant n \leqslant 102$) had been found, several properties, such as the heat of formation, the HOMO–LUMO gap and the fragmentation energy could be determined for each cluster. The heat of formation as a function of the cluster size is shown in figure 7; as a rule larger clusters are more stable than smaller ones. Thus, the reasons for the high abundance of clusters such as C_{60} cannot be explained only in terms of thermodynamic considerations; dynamical effects must play an important role. This conclusion is reinforced by the observation that C_{60} has a high fragmentation energy and a high HOMO–LUMO gap, ΔE_{HL}, when compared to other clusters. These properties are indicative of a less-reactive nature. Thus, clusters with high fragmentation energy and high HOMO–LUMO gaps are more likely to be kinetically favoured during the formation process. Zhang *et al* (1992a) showed that C_{60} has both the highest values of E_{frag} and ΔE_{HL} in the range of cluster sizes from 20 to 94, and thus should be (and indeed is) observed in high abundance. Conversely, C_{72}, which has a relatively high ΔE_{HL}, is not so abundant, because its low value of E_{frag} would tend to favour its fragmentation into smaller clusters.

The dynamical properties of some C_n clusters have also been studied. In particular, Wang *et al* (1992) and Zhang *et al* (1993b) have determined the vibrational spectra of C_{60} and C_{70}, and of C_{84} isomers respectively. The phonon dispersion curves and densities of states, as well as the effects of pressure on these, have been determined for solid C_{60} by Yu *et al* (1993, 1994). These authors used a combined tight-binding and force-field model to describe both the intra-molecular and inter-molecular interactions.

As well as the structural and dynamical properties of fullerenes, tight-binding has also been used to study the process of fragmentation. This has been done by Zhang *et al* (1993c), by Xu and Scuseria (1994) and by Kim and Tománek (1994). Zhang *et al* used molecular dynamics to study the fragmentation process for different cluster sizes. They found that the temperature of fragmentation depended almost linearly on the cluster size for small clusters, but it became more or less independent for sizes of C_{60} and larger. The large fragmentation temperatures found (between 4000 and 5500 K) made important the use of Fermi–Dirac statistics for the electrons. This was found to reduce somewhat the fragmentation temperatures. Xu and Scuseria found that the fragmentation process of C_{60} occurred via successive C_2 emission; only dimers were emitted from the fragmenting clusters. They were also able to observe the annealing of C_{60} metastable forms (the Stone–Wales isomer) back to buckminsterfullerene, a process having an activation barrier somewhat below the fragmentation barrier.

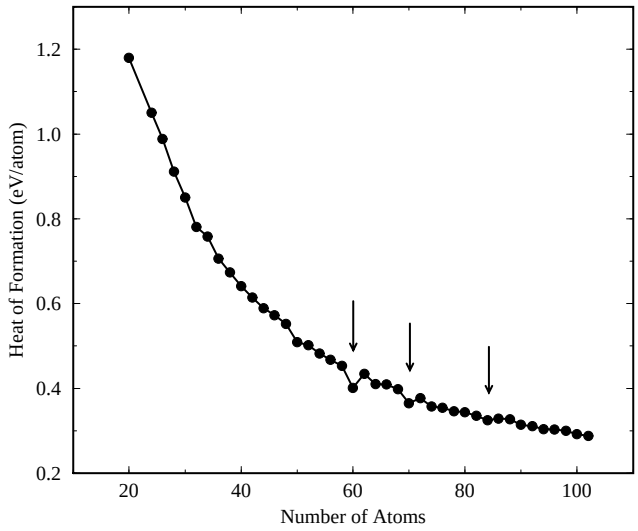

Figure 7. The heat of formation for carbon clusters as a function of cluster size. The heat of formation is given with respect to graphite. Data taken from Zhang *et al* (1992b, 1993a).

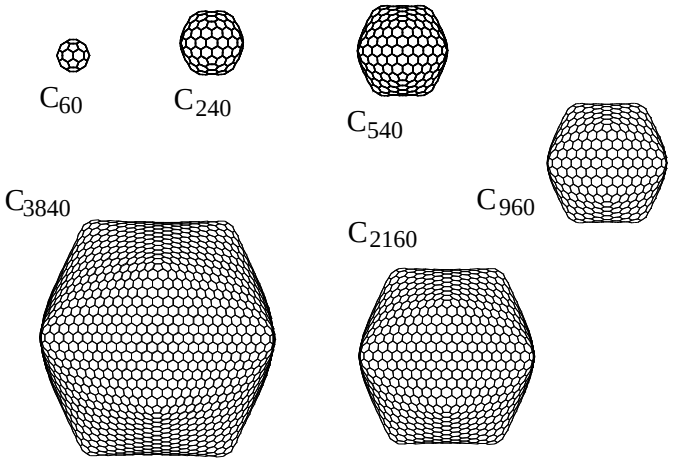

Figure 8. Minimum energy structure of the C_{60}, C_{240}, C_{540}, C_{960}, C_{2160} and C_{3840} buckminsterfullerenes. From Itoh *et al* (1996), reproduced by kind permission.

Large fullerenes have been often used as examples in which to put new linear scaling methods for electronic structure calculations to the test. This has been done for example by Galli and Mauri (1994), who performed molecular dynamics simulations of C_{60} clusters colliding on a diamond surface using the localized orbitals scheme (described in section 3.3.1). Also Itoh *et al* (1996) have used the localized orbitals scheme to determine the minimum energy structures of C_{60}, C_{240}, C_{540}, C_{960}, C_{2160} and C_{3840} (figure 8).

Recently, there has been considerable interest in fullerene analogues synthesized from materials other than carbon. Fowler *et al* (1996) have investigated the structures of $(BN)_n$ fullerene-like clusters, and have found important structural differences between these and their carbon counterparts.

Nanotubes also pose many interesting questions, both of a fundamental and practical nature. However, their larger sizes pose a bigger challenge for tight-binding calculations. Nevertheless, it is expected that tight-binding methods, especially linear-scaling approaches, will play an increasingly important role in the study of these systems. Some examples of the application of tight-binding models to the study of nanotubes are the work of Charlier *et al* (1996), who have investigated the effects of polygonization of the cross section of small carbon nanotubes on their electronic properties, Menon *et al* (1996), who used tight-binding to determine the vibrational frequencies of small nanotubes, Chico *et al* (1996), who studied the conductance of defective carbon nanotubes, and Carroll *et al* (1997), who performed structural simulations on closed-carbon nanotube tips.

5.4. Ionic materials

Traditionally the theoretical framework of tight-binding has been applied only to covalent and metalic systems. Ionic materials have been less studied because of the inability of conventional non-self-consistent tight-binding models to account for the charge-transfer effects present in these systems. More recently, however, self-consistent tight-binding models have been used successfully in the simulation of ionic materials. In this subsection we will focus our attention on a series of papers by Noguera and coworkers (Russo and Noguera 1992a, b, Goniakowski *et al* 1993, Noguera *et al* 1993). The motivation for their study was the need for a microscopic understanding of the acid/base properties of these oxide surfaces, which play an important role in the adhesion of polar polymers on these surfaces.

Given the difficulties of directly determining the pK of a given surface, it has been standard practice to judge the acid/base character of a given surface in terms of the pH of a solution put in contact with the surface. This quantity is refered to as the *zero point charge* (ZPC). The ZPC is clearly determined by the ability of the surface to adsorb protons and/or OH^- ions from the solution, so Russo and Noguera (1992a, b) undertook to investigate these processes. The substrate considered was MgO, and three surfaces were considered: a clean (100) surface, a clean (110) surface and a rough (100) surface. The aim was to gain an understanding of the H^+ and OH^- adsorption processes on these surfaces.

The calculations were performed on finite clusters of about 80 atoms in total which were embedded in a semi-infinite point-charge lattice. No relaxation of the surface was performed. According to the authors, surface relaxation would reduce the surface energy, but it would have little bearing on the adsorption energies of H^+ and OH^-; this is stated without justification. Calculations on three different surfaces were reported: (a) a clean (100) surface, (b) a clean (110) surface and (c) a rough (100) surface obtained by adding a cluster of six atoms above the surface occupying lattice positions. This cluster introduces the presence of two step sites and four corner sites on the rough surface.

Calculations were carried out first on the surfaces without adsorbate. The band gap in the projected density of states on the oxygen atoms was found to decrease progressively along the sequence bulk, (100), (110), step, corner. Two factors were found to influence the value of the band gap in the projected density of states: the first was the value of the Madelung field in the vicinity of the atom, and the second and less important was the coordination of the atom.

The total electron population by species varied rather little in going from the bulk to the surfaces (7.125 for O and 0.87 for Mg in the bulk, 7.141 and 0.865 in the (100) clean surface, 7.12 and 0.88 in the (110) surface, 7.16 and 0.8555 in the step and 7.01 and 0.915

in the corner). However, marked differences were found when the populations of individual orbitals were monitored. In general it was found that Mg atoms lose approximately a half of their bulk-orbital population to the missing O vacancy, but this loss is compensated by population increases in the remaining orbitals, which still participate in the bonding to neighbouring O atoms. The authors interpret this as a screening effect, which prevents the surface from bearing ionic charges much different from those found in the bulk.

Russo and Noguera (1992a) then investigated the adsorption of a proton onto the surfaces. They found that the most stable adsorption configuration in all cases was obtained with the proton located at a distance of 0.95 Å directly above a surface O atom, a value which is rather close to that obtained for the OH^- ion and the water molecule. The adsorption energy was found to be 8.04 eV on the (100) surface and 8.46 eV on the (110) surface.

Again, charge-transfer effects were observed in the screening of the bare proton by the electrons of the surface. The equilibrium charge of the adsorbed proton was found to be 0.3, i.e. decreased by 0.7 from that of the bare proton. About half of this charge transfer was from the O atom to which the proton is attached.

The situation is somewhat different for the adsorption of the OH^- ion. The distance between the O in the OH^- and the Mg ion in the surface is close to 2 Å, similar to the Mg–O bond distance found in the bulk. The adsorption energy is lower than for the proton, ranging from 1.56 eV on the clean (100) surface, to 2.74 eV on the on-bridge adsorption site. An electron transfer from the adsorbate to the surface is observed. The magnitude of transferred charge varies according to the adsorption site, but is around 0.2. However, this excess of charge is not localized on the Mg atom but rather is dissipated beyond it.

The general conclusion from these calculations is that the adsorption of H^+ is more favourable than that of OH^- ions, thus demonstrating the basic character of MgO surfaces. It was also revealing that covalent effects play a much more important role than anticipated on the basis of previous adsorption models, which considered the adsorption process on oxide surfaces to be driven essentially by electrostatic forces (Parks 1965).

In a subsequent paper, Goniakowski *et al* (1993) reported calculations of H^+ and OH^- adsorption on BaO (001), SrO (001), CaO (001), MgO (001), rutile-TiO_2 (110) and α-quartz SiO_2 (0001) surfaces, using a similar model to that employed by Russo and Noguera (1992a, b). These surfaces were chosen on the basis of their different acidity, as measured by the ZPC technique. This represented an attempt to corroborate if the self-consistent tight-binding model employed in their calculations was capable of reproducing the known trends observed experimentally. It was found that in the case of the rocksalt oxides, the proton adsorption was strong, with a large value of the adsorption energy and a significant degree of charge transfer from the surface to the proton. The adsorption energy decreased for TiO_2 and SiO_2 surfaces, though a significant degree of bonding was observed. This behaviour was correlated with the increasing covalent character of the material along the surface series: the more covalent the substrate, the less charge on the surface oxygen, and thus the weaker the bonding between the oxygen and the proton.

Again, a different behaviour was observed for the OH^- on the rocksalt surfaces; it was found that the adsorption of this ion was weaker than for protons. However, it was found to be much more favourable in the case of TiO_2 and SiO_2 surfaces, with an increased charge transfer and a more substantial modification on the projected density of states, evidencing a significant degree of bonding. The enhanced adsorption of OH^- on these surfaces was correlated with the smaller size of the cation on the substrate with respect to that in the rocksalt surfaces and their smaller electropositivity, which increases the hopping probabilities thus bringing the cation and hydroxyl levels closer in energy. The rocksalt surfaces were found to decrease their basic character in going from BaO to MgO, due to

the increasing difficulty of adsorbing a proton. The TiO_2 and SiO_2 surfaces, however, by their preference to adsorb OH^- over H^+, revealed their acid character.

6. Beyond orthogonal tight-binding: increasing accuracy

We have seen in previous sections that tight-binding is a popular approach for the simulation of materials. It has advantages over both empirical potentials and more accurate first-principle calculations. By bringing directly into play the electronic structure of the system, albeit in a very simplified way, tight-binding incorporates many-body quantum effects that are not included in empirical potentials. On the other hand, the simplified nature of the tight-binding approach enables it to be used in the study of large systems, not accessible with first-principles calculations due to the much higher computational demands of the latter. We have seen that appropriately parametrized tight-binding models are capable of accurately reproducing results obtained with first-principles calculations, at only a fraction of the cost.

However, in spite of its success and popularity, tight-binding shares with empirical potentials a serious drawback: its lack of transferability. A parametrization suitable for Si in the diamond structure, for example, is unlikely to be adequate for simulating liquid Si, unless information concerning the liquid phase is incorporated in the data used to construct the parametrization. The situation can be somewhat improved by including a wider range of empirical data in the fitting process, while at the same time increasing the number of disposable parameters in the model, in the hope of building in enough flexibility as to give a physically correct picture at the desired conditions of the simulation. This approach has been taken for example by Goodwin *et al* (1989) for Si, and by Xu *et al* (1992) for C. However, this approach is not systematic, as the mathematical form of the model is usually adopted in an *ad hoc* basis, without an underlying guiding principle. A different approach, which we will discuss in this section, consists of improving the transferability of tight-binding models by making them physically more plausible, without necessarily increasing the number of disposable parameters or widening the range of data used in the fit.

It will be convenient at this stage to review the causes for the lack of transferability in tight-binding, that is, the approximations listed in section 4. In standard tight-binding parametrizations it is assumed that the underlying basis set is orthogonal. As a justification for this neglect of the overlap, it is usually argued that the basis consists of orthogonalized functions constructed from non-orthogonal atomic-like orbitals via an orthogonalization transformation (for example Löwdin (1950)). The difficulty with this is that any orthogonalization transformation is valid only for a single configuration of the system; if the configuration changes, a different transformation is required to orthogonalize the atomic-like basis. Furthermore, orthogonalized functions have a wider extent than the equivalent non-orthogonal ones, because the orthogonalization process mixes functions centred on nearby atoms. Therefore, if the basis is assumed to consist of orthogonalized functions, the Hamiltonian matrix should be expected to have interactions between more distant neighbours than it would otherwise have. And yet, in conventional parametrizations it is common to adopt the *nearest-neighbour* approximation. Thus, conventional nearest-neighbour non-orthogonal tight-binding models cannot be expected to be transferable to different structures, and the failure to properly account for overlap effects is possibly the main reason for this lack of transferability. The effects of the neglect of non-orthogonality of the basis have been investigated by Mirabella *et al* (1994) for one-dimensional atomic chains, and by McKinnon and Choy (1995) for two- and three-dimensional lattices. These studies show that to properly account for the effects of the non-orthogonality, interactions extending beyond first-nearest neighbours are needed in the

Hamiltonian parametrization of orthogonal tight-binding models. Dorantes-Dávila *et al* (1993) and Dorantes-Dávila and Pastor (1995) have recently proposed an iterative scheme to incorporate in an orthogonal tight-binding framework the effects of interaction parameters and overlaps in a non-orthogonal tight-binding representation.

Another characteristic of conventional tight-binding models is the use of the two-centre approximation, which results in the neglect of three-centre integrals. As pointed out by Slater and Koster (1954), the contribution of these integrals is smaller than that from two-centre integrals, but is certainly not negligible (Pettifor 1977). Three-centre integrals are very sensitive to changes in the structure, and this sensitivity is partially lost when they are neglected. Most tight-binding models reported to date follow the practice of neglecting three-centre integrals, but there are exceptions to this rule, such as the first-principles tight-binding scheme of Sankey and Niklewski (1989), to be discussed later in this section.

Lack of self-consistency in tight-binding models also results in lack of transferability. In a self-consistent calculation with an atomic-like basis set, the on-site Hamiltonian matrix elements for atoms of the same species are different if the atoms are located in different environments. This response to the local environment is not reproduced in conventional tight-binding models. We have already discussed some of the most common approaches adopted in order to build some degree of self-consistency into tight-binding models. In this section we review other, more recent approaches used to make tight-binding self-consistent.

Finally, there is one remaining cause for the poor transferability of tight-binding parametrizations. We have seen earlier that the total tight-binding energy is expressed as a sum of a band-structure term and a repulsive term accounting for the ion–ion repulsion and the over counting of the electron–electron interaction energy in the band structure term:

$$E_{\mathrm{rep}} = E_{\mathrm{ii}} - E_{\mathrm{ee}}. \tag{78}$$

U_{rep} is generally assumed to take a pair-wise additive form. This is justified for E_{ii} if the ions are regarded to be point-like, but E_{ee} is many-body in nature, and, despite the arguments of section (2.4), cannot be expected to be well characterized by the same sum of pair potential terms in different configurations.

In order to obtain tight-binding parametrizations with improved transferability, at least some of the deficiencies of the conventional models listed above must be overcome. In this section we review the different approaches that have been considered to achieve this purpose.

6.1. Non-orthogonal tight-binding

Possibly the most natural extension to conventional tight-binding models consists of incorporating the overlap between the different basis functions, while retaining the empirical nature of the model. The eigenstates ψ_i and the eigenvalues E_i can be obtained by solving the generalized eigenvalue problem

$$(\mathbf{H} - E_i\mathbf{S})\psi_i = 0 \tag{79}$$

where $\mathbf{H}$ and $\mathbf{S}$ are the Hamiltonian and overlap matrices respectively in the atomic basis $\{\phi_\alpha\}$ representation.

If an explicit atomic basis (e.g. Slater-type orbitals) is used, the method that results is known as the *extended Hückel* method (Hoffmann 1963), popular in quantum chemistry, and forerunner of more sophisticated semiempirical methods. In extended Hückel calculations the overlap matrix elements

$$S_{\alpha\beta} = \int \mathrm{d}\boldsymbol{r}\, \phi_\alpha(\boldsymbol{r})\phi_\beta(\boldsymbol{r}) \tag{80}$$

are calculated explicitly. The on-site matrix elements are taken as the ionization potentials of the corresponding orbitals in the isolated atom, i.e. $H_{\alpha\alpha} = \epsilon_\alpha$, and the off-diagonal elements are obtained from these and the overlap using Mulliken's parametrization (Mulliken 1949a, b):

$$H_{\alpha\beta} = \tfrac{1}{2} K(\epsilon_\alpha + \epsilon_\beta) S_{\alpha\beta} \tag{81}$$

where K is an adjustable parameter.

An alternative to the extended Hückel method that does not require explicit consideration of a basis set, and is thus closer in spirit to the conventional tight-binding approach, was proposed by van Schilfgaarde and Harrison (1985, 1986). Their approach consists of assuming that Harrison's universal parametrization (Harrison 1980) gives the matrix elements $V_{ll'm}$ in a hypothetical orthogonal basis. The Hamiltonian matrix elements in the non-orthogonal atomic basis representation are related to these through the approximate expression

$$H_{ll'm} = V_{ll'm}\left(1 + \frac{1}{K} - S_2^2\right) \tag{82}$$

where S_2 is the non-orthogonality between two sp^3 hybrids†, given by

$$S_2 = \tfrac{1}{4}(S_{ss\sigma} - 2\sqrt{3}S_{sp\sigma} - 3S_{pp\sigma} + 3S_{pp\pi}). \tag{83}$$

The overlap matrix elements are obtained by inverting (81)

$$S_{ll'm} = \frac{2V_{ll'm}}{K(\epsilon_l + \epsilon_{l'})}. \tag{84}$$

van Schilfgaarde and Harrison used this scheme to study the properties of α-rhombohedral B (van Schilfgaarde and Harrison 1985) and Zn_3P_2 (van Schilfgaarde and Harrison 1986).

An approach based on that of van Schilfgaarde and Harrison has been developed by Menon and Subbaswamy (1991), among others. In this approach the universal matrix elements have an exponential dependence on the interatomic distance r,

$$V_{ll'm}(r) = V_{ll'm}(d_0)\mathrm{e}^{-\alpha(r-d_0)} \tag{85}$$

where d_0 is the equilibrium distance in a chosen bulk phase of the material. The Hamiltonian and overlap matrices in the non-orthogonal basis are obtained using (82) and (84), as done by van Schilfgaarde and Harrison. A repulsive pair potential is used, having the form

$$\chi(r) = \chi_0 \mathrm{e}^{-\beta(r-d_0')} \tag{86}$$

where $\beta = 4/r_0$, r_0 being one half of the dimer bond-length, and d_0' is the sum of the covalent radii of the interacting atoms. With this model Menon and Subbaswamy (1991) used the molecular dynamics technique to simulate C_{60}, studying the chemisorption of oxygen on the surface of the molecule, and the process of electron doping, analysing its effect on bond-lengths and vibrational frequencies of the molecule. They also used it to simulate Si_n clusters (Menon and Subbaswamy 1993) with n in the range 3–10. In order to simulate these clusters they introduced a bond-counting term, as done by Tománek and Schluter (1987). Later, Ordejón *et al* (1994) found that it was possible to improve the model of Menon and Subbaswamy by explicitly including the distance dependence of S_2 (83). This modification obviated the need for a bond-counting term, previously needed to reproduce *ab initio* results for the cohesive energies of the Si_n clusters. This model was further refined by Menon and Subbaswamy (1994) by incorporating into S_2 the missing term $3S_{pp\pi}$. The resulting model was tested for different bulk phases of Si and Si_n clusters.

† van Schilfgaarde and Harrison omitted the last term $3S_{pp\pi}$ in their formalism.

Although cluster results were well reproduced, some of the bulk phases (and in particular, the β-tin phase) were poorly described.

An approach not relying on the use of Harrison's universal matrix elements was adopted by Mattheiss and Patel (1981). Their approach consists of assuming a non-orthogonal atomic basis, and fitting the band-structure obtained from the resulting generalized eigenvalue calculation to results from accurate calculations. These authors developed a non-orthogonal parametrization for Si consisting of 26 parameters: the two orbital energies ϵ_{s} and ϵ_{p}, plus the 24 Hamiltonian and overlap integrals coupling each atom with its first-, second- and third-nearest neighbours. The values of these parameters were obtained by fitting the band structure at 12 points of the Brillouin zone, to results obtained by Chelikowsky and Cohen (1974) using an empirical non-local pseudopotential. This model was found to accurately reproduce the band structure up to energies of 6 eV above the valence-band maximum. Mattheiss and Patel used the thus parametrized model to study defect states produced by stacking faults in Si.

This approach was also adopted by Allen *et al* (1986), who fitted parameters to the band structures of bcc, fcc and sc lattices of Si obtained with LMTO calculations. They found that in order to accurately reproduce the LMTO bands using first- and second-nearest neighbour integrals in the fit, it was necessary to include the non-orthogonality between the orbitals centred on different atoms. The interesting point about this work is that when the values of the $H_{ll'm}$ and $S_{ll'm}$ parameters from their fits were plotted as a function of the interatomic distance, they were found to describe smooth curves. Furthermore, the values obtained by Mattheiss and Patel for the diamond structure were also found to align correctly on these curves, giving credence to the concept of transferability in tight-binding parameters.

Sigalas and Papaconstantopoulos (1994) also fitted the Hamiltonian and overlap matrix elements to first-principles augmented plane-wave (APW) calculations. In their method, the tight-binding total energy has the following expression

$$E_{\mathrm{tot}} = \sum_{l} \epsilon_l + \frac{1}{2} \sum_{i \neq j} \left\{ \sum_{n} \frac{A_n}{r_{ij}^n} \right\} \mathrm{e}^{\lambda r_{ij}} + c. \tag{87}$$

The first term is the sum of eigenvalues of the occupied states obtained from solving the generalized eigenvalue problem. The Hamiltonian and overlap matrix elements are obtained by fitting the APW bands at a series of lattice parameters for different structures. The model retained only nearest-neighbour interactions, but since they applied this model to transition metals, it includes d as well as s and p orbitals. The second term is a repulsive pair potential, and the c parameter is a constant that accounts for the scale difference between the APW and tight-binding results. These authors determined the Hamiltonian and overlap integrals, as well as the A_n and λ parameters for the transition metals Rh, Pd, Ta, Ir and Au. For the first four metals, five A_n parameters were used ($n = 1, 5$), while nine ($n = -3, 5$) were required to obtain a good fit for Au. With these parametrizations they determined the vacancy formation energies and the elastic constants, obtaining good agreement with both experiments and first-principles calculations for all the systems investigated.

6.2. Self-consistent tight-binding

We have already discussed some of the modifications introduced in tight-binding models in order to build in some degree of self-consistency. Here we will further develop the approach so as to include Coulomb effects, and mention two recent approaches which are strongly related to some of the non-orthogonal tight-binding methods just discussed.

6.2.1. Hubbard U and Madelung sum. The methods described in section 2.5.1 treat the intra-atomic electron interactions by the addition of the Hubbard U, but neglect the interatomic electron interactions. One means of dealing with this is to perform a Madelung sum over a crystal of charges, and alter the Hubbard U:

$$U^* = U - \frac{\alpha e^2}{d} \tag{88}$$

where α is the Madelung constant for the crystal. The draw back for this method is that it breaks down when the crystal is not perfect. Harrison (1985) noted that the *effective* Hubbard U, U^*, could become negative for certain solids, usually as a result of the assumptions made in the Madelung sum breaking down. He defined a negative U^* to be zero accordingly. He studied covalent and ionic solids, the effect on band gaps, cohesive energy and correlated states with some accuracy. Some mention is made of self-consistency, though it is not used for the calculations described. The theory was later developed for elasticity in ionic solids (Straub and Harrison 1989) and dielectric screening in semiconductors (Harrison and Klepeis 1988), discussed in section 6.3.

Strehlow *et al* (1985) apply a similar method to III-V semiconductors, though they suggest that when U^* becomes negative, α, the Madelung constant, should be set to one. They also use three separate Us–U_{ss}, U_{sp} and U_{pp}. The method is extended to defects (Hanke *et al* 1986, Strehlow *et al* 1986 and Kühn *et al* 1987) by calculating the electrostatic interactions exactly for a small cluster around the defect (generally the four nearest neighbours) and embedding it in the Madelung field for the crystal. Reasonable agreement with such experimental and *ab initio* results as are available is obtained.

Bechstedt *et al* (1985) used a similar method to treat semiconductor surfaces, with an effective Hubbard U for the system:

$$U = \tfrac{1}{16}[U_{ss} + 6U_{sp} + 9U_{pp}]. \tag{89}$$

They investigated the stability of the Si(111) $-$ 1 $\times$ 1 surface by varying the height of the atoms above the surface. The approach was also applied to the (111) $-$ 2 $\times$ 1 and (100) $-$ 2 $\times$ 1 surfaces of C, Si, Ge and α-Sn (Bechstedt and Reichardt 1988a, b).

6.2.2. Hubbard U plus self-consistent Coulombic contributions. The previous section introduced the concept of intersite charge effects on the Hubbard U; however, it used a formalism where a perfect crystal shape is assumed, which limits the applicability of the method. This section extends this to a site-by-site consideration of the effects. The on-site term is changed according to the on-site occupancy as before, but with a sum over all charges in the system also added. This is significantly more time consuming, but is necessary in certain systems for accuracy.

Majewski and Vogl (1987) implemented self-consistent charge transfer, using a single, site-dependant U parameter, and included some non-orthogonality effects. The on-site energy is given by

$$\epsilon_{lI} = \epsilon_{lI}^{\text{atom}} - U_I(Z_I - Q_I) + \sum_{I \neq I'} (Z_{I'} - Q_{I'}) V(\boldsymbol{R}_I - \boldsymbol{R}_{I'}) + f_{lI}. \tag{90}$$

The U_I are taken from Harrison (1985); Z_I is the number of valence electrons an a free atom, while Q_I is the charge on an atom as given in (21). The third term is the electrostatic potential at site R_I due to all the other sites in the system (formulated to give the Coulomb form at large interatomic distances and the effect of charge overlap at extremely short

distances), and f_{lI} is an approximate non-orthogonality correction (Harrison 1980), given by

$$f_{lI} = -\tfrac{1}{2}(SH_{\rm el} + H_{\rm el}S)_{lI,lI}. \tag{91}$$

In their implementation, the crystal structure was used to simplify the above expression for f_{lI}. They use the following expression for the total energy of a solid:

$$\begin{aligned} E_{\rm tot} &= \sum_{n,\boldsymbol{k}}^{\rm occ} \epsilon_{n,\boldsymbol{k}} - E_{\rm el-el} + E_{\rm core-core} \\ E_{\rm el-el}^{\rm intra} &= \sum_{I} U_I Q_I^2 \\ E_{\rm core-core} - E_{\rm el-el}^{\rm inter} &= \sum_{I} \sum_{J \neq I} (Z_I Z_J - Q_I Q_J) V(R_I - R_J). \end{aligned} \tag{92}$$

The model is rather impressive, predicting bond lengths, phase changes with pressure, long-wavelength tranverse optical phonons and band structures for 62 cubic, sp-bonded binary non-metals with varying degrees of ionicity (e.g. from Si to NaCl). Their results are in good agreement with experiment, in terms of trends where there is not good numerical agreement. It does not, however, include a derivation of forces—this was performed in a reformulation of the method by Kohyama *et al* (1989). The model is applied to SiC (Kohyama *et al* 1990b) and a parametrization is derived. The results are reasonable, and there is an investigation of the functional forms for the overlap integrals. This parametrization is used in calculations of the atomic and electronic structure of the $\{122\}\Sigma = 9$ grain boundary in SiC (Kohyama *et al* 1990a). A model is proposed, based on previous structures for the same grain boundary in Si, and it is investigated. Those results which can be compared with experiment seem reasonable. Polar interfaces for the same grain boundary have also been studied (Kohyama *et al* 1991). They find that one of the two polar interfaces (i.e. an interface consisting of all Si–Si or all C–C bonds) is more stable than the non-polar interface (consisting of alternating Si–Si and C–C bonds).

6.2.3. Chemical hardness. Liu (1995) exploits the concept of the *chemical hardness* matrix, originally introduced by Parr and Pearson (1983) to construct a self-consistent tight-binding model for silicon. The chemical hardness matrix describes how the chemical potential of an isolated system changes as the number of electrons is changed. Liu's model consists of two parts: a non-orthogonal tight-binding scheme based on the extended Hückel method of Hoffmann (1963), similar to that used by van Schilfgaarde and Harrison (1985, 1986) and Menon and Subbaswamy (1991), and an iterative procedure to adjust the on-site matrix elements of the Hamiltonian according to the electron population of the corresponding atomic orbitals.

The overlap matrix elements are tabulated at a series of internuclear separations from a set of dimer calculations using an LCAO DFT calculation. These calculations also provide values of the off-diagonal Hamiltonian matrix elements, but these, rather than tabulated, are assumed to be given in terms of a generalized Hückel expression

$$H_{ll'm} = K_{ll'm}(H_{llm} + H_{l'l'm})S_{ll'm} \tag{93}$$

and the $K_{ll'm}$ scaling parameters are adjusted so that the $H_{ll'm}$ values thus calculated reproduce those resulting from the LCAO DFT calculation. Additional fitting of the $K_{ll'm}$ parameters was used so that the diamond–Si band structure was also reproduced.

Self-consistency is built into the method by making the on-site Hamiltonian matrix elements dependent on the orbital populations:

$$\epsilon_j = \epsilon_j^0 + \sum_k \eta_{jk}^{(1)} \Delta n_k + \tfrac{1}{2} \sum_{kl} \eta_{jkl}^{(2)} \Delta n_k \Delta n_l \tag{94}$$

where $\eta_{jk}^{(1)}$ and $\eta_{jkl}^{(2)}$ are the first- and second-order chemical hardness matrices, and Δn_k is the change in the occupation number of state k. The chemical hardness matrices are defined as:

$$\eta_{jk}^{(1)} = \frac{\partial \epsilon_j}{\partial n_k} \tag{95}$$

and

$$\eta_{jkl}^{(2)} = \frac{\partial^2 \epsilon_j}{\partial n_k \partial n_l} \tag{96}$$

respectively. These can be parametrized by fitting the results of first-principles atomic calculations in which the populations of the valence states are systematically changed to an expression such as (94). The resulting model was found capable of reproducing the valence bands and lowest two conduction bands of the band structure of diamond–Si with a high accuracy. Furthermore, the band structure of fcc–Si (which had not been used at all in the fitting procedure) was also reproduced with a high accuracy up to the Fermi level.

6.3. Tight-binding and electromagnetism

The preceding section has dealt with the effect of local electromagnetic fields on electrons, and how this has been implemented in tight-binding. In this section, the effect of external fields will be addressed, though somewhat perfunctorily as the area has received little attention to date.

A theory whereby the effect of external electromagnetic fields can be incorporated into tight-binding in a gauge-invariant manner, without adding extra parameters, has recently been described (Graf and Vogl 1995). The essential result of the theory is to alter the Hamiltonian matrix elements, as follows:

$$\begin{aligned} &\epsilon_{\alpha,\boldsymbol{R}} = \epsilon_{\alpha,\boldsymbol{R}}^0 - e\Phi(\boldsymbol{R}, t) \\ &h_{\alpha',\alpha}(\boldsymbol{R}' - \boldsymbol{R}) = h_{\alpha',\alpha}^0(\boldsymbol{R}' - \boldsymbol{R}) \exp\left[-\frac{\mathrm{i}e}{2\hbar c} (\boldsymbol{R}' - \boldsymbol{R}) \times [\mathbf{A}(\boldsymbol{R}', t) + \mathbf{A}(\boldsymbol{R}, t)] \right] \end{aligned} \tag{97}$$

where Φ and $\mathbf{A}$ are the electric and magnetic potentials in the usual way. From this formalism, they derive expressions for the effective mass tensor, the effective Landé g factor and the transverse and longitudinal dielectric functions.

A different method has been used in studies of transition metal alloys in external magnetic fields. The method (Yamada and Shimizu 1985) is similar to the Hubbard U studies of magnetism described earlier in section 5.1.8. The Hamiltonian on-site elements are altered as follows:

$$H_{i\sigma,i\sigma} = \epsilon_i - \sigma \mu_B \mathbf{H} \tag{98}$$

where $\mathbf{H}$ is the external magnetic field. The method is applied to the moment of YCo_2, and its metamagnetic transition, with results which are in reasonable qualitative agreement with available data. It has also been applied to the metamagnetic transition of $Hf(Fe, Co)_2$ (Yamada and Shimizu 1989) and high-field magnetization of $Y(Co_{1-x}\ M_x)_2$ (M = Ni or Fe) (Mitamura *et al* 1995) again with some qualitative success.

As mentioned above, the effect of an electric field on a semiconductor, and its dielectric constant, are discussed in Harrison and Klepeis (1988). The discussion is limited to the response to an applied potential, which modifies the potential experienced by atoms in the bulk. This is derived in terms of the bond-order formalism of Harrison, and used to find the dielectric function. The rest of the paper discusses screening when there is no external field.

6.4. Two-centre ab initio tight-binding

We saw in section 6.1 that first-principles calculations are sometimes used to generate data used in the generation of tight-binding parametrizations. Another possible approach consists of using first-principles calculations with an atomic-like basis set, not, as in the examples above, to obtain band structures that can then be used as input data to fit the Hamiltonian and overlap matrix elements, but rather as a means to calculate and tabulate them directly. This has been done recently by Porezag *et al* (1995). In their approach the tight-binding energy is still written as a sum of a band-structure term and a pair-repulsion term. Slater-type orbitals are used to construct an atomic basis:

$$\phi_\nu(\boldsymbol{r}) = \sum_{n,\alpha,l,m} a_{nm} r^{l+n} \mathrm{e}^{-\alpha r} Y_{lm}\left(\frac{\hat{\boldsymbol{r}}}{r}\right). \tag{99}$$

This basis is then used to solve a modified Kohn–Sham Hamiltonian equation,

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + V_{\mathrm{pa}}(\boldsymbol{r})\right]\phi_\nu(\boldsymbol{r}) = \epsilon_\nu^{\mathrm{pa}}\phi_\nu(\boldsymbol{r}) \tag{100}$$

where $V_{\mathrm{pa}}(\boldsymbol{r})$ is the *pseudo-atom* potential

$$V_{\mathrm{pa}}(\boldsymbol{r}) = V_{\mathrm{nuc}}(\boldsymbol{r}) + V_{\mathrm{H}}(\boldsymbol{r}) + V_{\mathrm{xc}}(\boldsymbol{r}) + \left(\frac{\boldsymbol{r}}{r_0}\right)^N. \tag{101}$$

V_{nuc}, V_{H} and V_{xc} are the nuclear, Hartree and exchange–correlation potentials respectively. The last term has the effect of concentrating the charge density closer to the nucleus; its addition renders the resulting orbitals ϕ_ν more suitable for bulk calculations. Porezag *et al* (1995) claim that good choices for these parameters are $N = 2$ and $r_0 = 2r_{\mathrm{cov}}$, where r_{cov} is the covalent radius of the element in question.

Once the $\{\phi_\nu\}$ have been obtained, the tight-binding parametrization is carried out. The overlap matrix elements are tabulated as a function of the internuclear distance directly. In order to obtain the Hamiltonian matrix elements, an effective potential V_{eff} is constructed for each structure from atomic contributions:

$$V_{\mathrm{eff}}(\boldsymbol{r}) = \sum_i V_0^{(i)}(|\boldsymbol{r} - \boldsymbol{r}_i|) \tag{102}$$

where $\boldsymbol{r}_i$ is the position of atom i, and $V_0^{(i)}$ is the Kohn–Sham potential due to this atom, i.e. (101) but without the term $(\boldsymbol{r}/r_0)^N$. Using this potential, the Hamiltonian matrix elements are tabulated from

$$\langle\phi_\nu^{(i)}|H|\phi_\mu^{(i)}\rangle = \begin{cases} \epsilon_\nu^{(i)} & \text{if } \mu = \nu,\ j = i \\ \langle\phi_\nu^{(i)}|T + V_0^{(i)} + V_0^{(j)}|\phi_\mu^{(i)}\rangle & \text{if } j \neq i \\ 0 & \text{otherwise} \end{cases} \tag{103}$$

where T is the kinetic-energy operator.

To completely specify the model the repulsive pair potential must be determined. This is obtained as

$$V_{\mathrm{rep}}(r) = E_{\mathrm{SC}}(r) - E_{\mathrm{BS}}(r) \tag{104}$$

where E_{BS} is the band-structure energy at separation r obtained by diagonalizing the parametrized Hamiltonian (103), and E_{SC} is the total energy obtained from a self-consistent calculation at the same atomic configuration. The repulsive potential is usually fitted to results for diatomic molecules, though in some cases this could lead to difficulties arising from level-crossings, in which case it is possible to fit using results from other structures.

Because of the ansatz adopted for the Hamiltonian matrix elements (103), only two-centre integrals appear in the formalism. This scheme has been applied to the study of carbon clusters, C_{60}, hydrocarbons, and for the diamond, graphite, sc, fcc and bcc bulk carbon structures, both crystalline (Porezag *et al* 1995) and amorphous (Köhler *et al* 1995), to Si and SiH clusters (Frauenheim *et al* 1995), and to c-BN and several borane structures (Widany *et al* 1996), testifying to the wide range of applicability of this method. The fact that it compares in accuracy for molecules, clusters and bulk phases indicates that, in spite of its simple nature, this method produces tight-binding parametrizations with good transferability properties.

6.5. Three-centre ab initio tight-binding

Although the tight-binding parametrization scheme of Porezag *et al* (1995) has proved successful in many applications, it should be remembered that three-centre integrals are neglected. Furthermore, the transferability of the pair potential, which is normally fitted to results obtained for diatomic molecules, is a possible source of error. These limitations are bypassed by a more sophisticated first-principles tight-binding method, due to Sankey and Niklewski (1989). These authors adopted the Harris functional (Harris 1985, Foulkes 1987) formalism as the underlying theory in their scheme. Their approach consists of adopting a basis set of slightly excited pseudo-atomic orbitals (PAO's). These are obtained from an electronic calculation in the free atom, where the core electrons are eliminated from the problem by means of pseudopotentials. A finite cut-off is imposed on the valence PAO's, $\phi_\nu(\boldsymbol{r})|_{r=r_c} = 0$, which has the effect of mixing in excited orbitals with a small weight within $r < r_c$. Once the basis has been constructed from calculations in the isolated atom(s), the task is to obtain interpolation tables for the different contributions to the Hamiltonian matrix in this basis. Some two-centre integrals, such as the kinetic-energy matrix elements, are tabulated as a function of internuclear distance by direct integration. But two-centre integrals of functions which depend on the electronic density, such as the matrix elements of the exchange–correlation energy and potential, cannot be tabulated in this way, because the actual value of the electron density in the volume of the integral is not known beforehand. Though in principle one could adopt the approach of calculating these matrix elements numerically once the electron density is known, this would be considerably slower than the interpolation of the remaining two-centre and three-centre integrals, and would thus become the most time-consuming task in the construction of the Hamiltonian matrix elements. Therefore Sankey and Niklewski adopted an approximate method, which allowed them to estimate these matrix elements in terms of the overlap matrix elements, which are tabulated beforehand just as the kinetic-energy matrix elements, and the electron-density matrix elements, which can be rapidly evaluated at the given configuration. This approximate procedure is explained in great detail in Sankey and Niklewski (1989).

Contrary to the approach taken by Porezag *et al* (1995), in this scheme three-centre

integrals are not neglected, but explicitly included. The three-centre integrals appearing in this formalism have the form $\langle\phi_{i\alpha}|V_k^{\mathrm{NA}} + V_k^{\mathrm{NL}}|\phi_{j\beta}\rangle$, where V_k^{NL} is the non-local part of the ionic potential centred on atom k, and V_k^{NA} is the *neutral-atom* potential generated by atom k, which has the form

$$V_k^{\mathrm{NA}}(\boldsymbol{r} - \boldsymbol{r}_k) = V_{\mathrm{loc}}(\boldsymbol{r} - \boldsymbol{r}_k) + \int \frac{n_k(\boldsymbol{r} - \boldsymbol{r}_k)}{|\boldsymbol{r} - \boldsymbol{r}'|}\,\mathrm{d}\boldsymbol{r}'. \tag{105}$$

Here, V_{loc} is the local part of the ionic potential, and the integral in the second term of the right-hand side of the above equation is the contribution to the total Hartree potential arising from the valence electron charge of atom k. One of the approximations implicitly assumed in this method is that the valence charge remains fixed at the value of the isolated atom, i.e. charge-transfer effects cannot be accounted for. Therefore, this method is not directly applicable to situations in which charge-transfer is important. Three-centre integrals require tabulation with respect to two distances only, since the influence of the third (angular) variable is worked out analytically (for details the reader should consult Sankey and Nicklewski 1989). Once the two-centre and three-centre integrals have been appropriately tabulated with respect to the necessary variables, the tight-binding parametrization is complete. In a tight-binding calculation the dominant step is solving the electron-structure problem. Constructing the tight-binding Hamiltonian from interpolation tables is a fast process, and therefore this scheme is not computationally more demanding than other non-orthogonal tight-binding schemes discussed earlier, though it certainly requires a greater effort in the tabulation of the matrix elements.

The scheme of Sankey and Niklewski has been applied widely. Sankey *et al* (1990) have used it to simulate Si_n clusters including $n = 2$–7. Drabold *et al* (1990a, b, 1991a) used this method in simulations of amorphous Si, and also to study C clusters, diamond and graphite (Drabold *et al* 1991b). Adams *et al* (1991) used it to study the vibrational modes of C_{60}. The method has also been used to study defects and their localization effects in a-Si (Fedders *et al* 1992) and a-Si:H (Fedders and Drabold 1993), the structure of Al clusters (Yang *et al* 1993), Si surfaces (Adams and Sankey 1991) and new forms of C (O'Keeffe *et al* 1992), among other applications.

6.5.1. Modified Sankey–Niklewski method. Tsai *et al* (1992) have modified the method of Sankey and Niklewski (1989) so as to allow this method to take into account charge transfer effects in an iterative way. The original method of Sankey and Nicklewski exploited the *neutral-atom approximation* to reduce drastically the number of three-centre integrals. However, in situations in which charged atoms are present, this approximation cannot be used directly, given the long range of the associated Coulomb potentials. Tsai *et al* used the Ewald technique, which consists of adding and subtracting to each charged atom a Gaussian charge distribution of opposite sign and equal total charge. Thus the system can be decomposed into a neutral-atom subsystem, containing the charged atoms and the neutralizing charge distribution, and the cancelling charge distribution. In the neutral-atom subsystem the Coulomb potentials are then short ranged, and can be calculated exactly as in the original method, while the potentials due to the Gaussian charge distribution can be easily calculated with standard reciprocal-space methods. In order to determine the self-consistent atomic charges, it is necessary to iteratively solve the Schrödinger equation updating the Kohn–Sham potential each time. To reach self-consistency in a more efficient way, Tsai *et al* use a simple mixing procedure, by which the population of orbital $\phi_{i\alpha}$ in the current iterative cycle, $n'_{i\alpha}$, is predicted to be

$$n'_{i\alpha} = (1 - \epsilon)n^0_{i\alpha} + \epsilon n_{i\alpha} \tag{106}$$

where ϵ is a mixing factor lying in the interval [0,1], and $n_{i\alpha}$ and $n^0_{i\alpha}$ are the populations obtained for this state in the two previous self-consistency iterations. This mixing procedure and rediagonalization is repeated until the differences $n'_{i\alpha} - n_{i\alpha}$ for all atomic orbitals fall below some predefined tolerance.

7. Beyond orthogonal tight-binding: increasing system size

Having discussed increasing accuracy, the other direction in which we may wish to extend tight-binding is to treat larger systems by means of a lower-order approximation. The formalism of the recursion methods (section 3.2) allows us to make close connection between an appoximation to the tight-binding method and the empirical Finnis and Sinclair (1984, 1986) and Tersoff (1988a, b) style potentials. A review of all the work which has been performed using such potentials is beyond the scope of this paper; the purpose here is to consider the *relationship* between these methods and the tight-binding formalism.

It is convenient to discuss two classes of materials to which these approaches have been applied; first, d-band metals, and secondly sp-bonded semiconductors.

7.1. Low-order moments approximations to d-band metals

The starting point for low-order moments approximations is the physical expectation that the energetics of transition metals are dominated by the width, shape and occupation of the d-band. Indeed, a great deal can be done by considering only the width, which is closely related to the second moment of the density of states. In the rectangular d-band approximation (Friedel 1969) a d-band of width W has a second moment of $W^2/12$. Recalling (35), the second moment is given by

$$\mu^2_{i\alpha} = \langle i\alpha|\hat{H}^2|i\alpha\rangle = \sum_{j\beta}(H_{i\alpha,j\beta})^2. \tag{107}$$

As the energy of the occupied d-states is proportional to the width of the d-band, we can make the second-moment approximation,

$$E_{\text{band}} = -A\sum_{i\alpha}\sqrt{\sum_{j\beta}(H_{i\alpha,j\beta})^2}. \tag{108}$$

This form leads very simply to the simplest possible form for the band energy that can still be considered to be based upon tight-binding, that of Finnis and Sinclair (FS) (1986),

$$\begin{aligned} E_{\text{band}} &= -A\sum_i F(\rho_i) \\ F(\rho_i) &= \sqrt{(\rho_i)} \\ \rho_i &= \sum_j \phi(R_{ij}) \\ \phi(r) &= (r-d)^2 + \beta(r-d)^3 d^{-1} \end{aligned} \tag{109}$$

where A, β and d are empirical parameters. To the energy in (109) is added a pairwise repulsive energy,

$$\begin{aligned} E_{\text{rep}} &= \tfrac{1}{2}\sum_{ij} V(R_{ij}) \\ V(r) &= \begin{cases} (r-c)^2(c_0 + c_1 r + c_2 r^2) & \text{if } r \leqslant c \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{110}$$

giving four further empirical parameters. Finnis and Sinclair showed that this functional form has two significant advantages over purely pairlike potentials such as those of Maeda *et al* (1982); first, that the Cauchy pressure, $\frac{1}{2}(C_{12} - C_{44})$ can take a non-zero value, and secondly that the vacancy formation energy can differ from the cohesive energy.

In passing, it should be noted that (109) is the same as the embedded atom form of Daw and Baskes (1984) except in the choice of *embedding function*, $F(\rho_i)$. The FS potential may therefore be considered to be an embedded atom potential with the choice of embedding function motivated by second-moment tight-binding.

The second moment alone, however, is not adequate to describe the properties of d-band metals. Ackland and Thetford (1987) showed that the FS potential is *attractive* at short ranges for V and Nb, and in general inaccurately reproduces compressive behaviour; they remedied this by simply adding a short-range correction to the repulsive term (110) to give

$$\mathrm{V_{AT}}(r) = \mathrm{V_{FS}}(r) + B(b_0 - r)^3 \mathrm{e}^{-\alpha r}. \tag{111}$$

A further problem, both with the FS and Ackland–Thetford forms (as well as a similar variant, due to Rebonato *et al* (1987)) was identified by Marchese *et al* (1988). The quasi-harmonic Grüneisen parameter, $\gamma_{\mathrm{QH}} = \mathrm{V}(\frac{\partial P}{\partial E})_{\mathrm{V}}$ is significantly underestimated; for V, Nb and Cr it is predicted to be negative, and for W, Ta and Fe, although it has the correct sign, it is at least a factor of three too small. The Grüneisen parameter describes, amongst other things, the change in phonon frequencies with volume. This failure suggests that FS-type potentials will have difficulty describing dynamic processes where there are local variations in density, such as in the vicinity of a defect.

The founding study for the inclusion of higher-moment terms, albeit from a density-functional rather than tight-binding background, is the work of Moriarty (1990), which was discussed in section 5.1.2. He showed that it is possible, in principle, to describe the total energy in terms of structure-independent, transferable interatomic potentials. The actual forms of these potentials, however, rapidly becomes very complex, as we have already seen to be the case for moments expansions in general. The link with tight-binding moments theory led, however, to the possibility of semi-empirical *functional forms* being found with a firm theoretical basis.

Carlsson (1991) therefore developed a semi-empirical model for W and Mo which included the fourth-moment terms,

$$\mu_4(i) = \sum_{jkl} \sum_{\alpha\beta\gamma\delta} h_{ij}^{\alpha\beta} h_{jk}^{\beta\gamma} h_{kl}^{\gamma\delta} h_{li}^{\delta\alpha}. \tag{112}$$

The second-moment contribution to the electronic energy is written in terms of the second-moment *matrix*,

$$\mu_2^{\alpha\beta}(i) = \sum_{j\gamma} h_{ij}^{\alpha\gamma} h_{ji}^{\gamma\beta} \tag{113}$$

as

$$E_{\mathrm{el}}^{(2)} = -\sum_i \mathrm{Tr}[\mu_2(\mathrm{i})^{1/2}]. \tag{114}$$

Carlsson showed that part of the fourth-moment contribution to the energy (about one half) is included in (114), that is, those paths of length four which are equivalent to two paths of length two,

$$\mu_4^{(2)}(\mathrm{i}) = \sum_{jl} \sum_{\alpha\beta\gamma\delta} h_{ij}^{\alpha\beta} h_{ji}^{\beta\gamma} h_{il}^{\gamma\delta} h_{li}^{\delta\alpha}. \tag{115}$$

The rest of the fourth-moment contribution was included as

$$E_{\rm el}^{(4)} = C\sum_i [\mu_4(i) - \mu_4^{(2)}(i)]/\mu_2(i)^{3/2} \tag{116}$$

with C an adjustable parameter.

Results using this potential were mixed; although some bulk properties were considerably innacurate, this was the first semi-empirical model to correctly reproduce the W(100) $(\sqrt{2}\times\sqrt{2})R45^\circ$ surface. However, when a similar potential was produced by Foiles (1992) which more accurately fitted the bulk properties of W and Mo, the surface reconstructions could not be reproduced.

Xu and Adams (1994) showed that some improvement could be obtained by including the third-moment term,

$$E_{\rm el}^{(3)} = B\sum_i \frac{\mu_3(i)}{\mu_2(i)} \tag{117}$$

where

$$\mu_3(i) = \sum_{jk}\sum_{\alpha\beta\gamma} h_{ij}^{\alpha\beta} h_{jk}^{\beta\gamma} h_{ki}^{\gamma\alpha}. \tag{118}$$

Xu and Adams fitted the 10 parameters in their model to 12 bulk properties for Mo, W and V, and were able to obtain fairly good fits both to these properties, and to the surface structures of W(100).

The inclusion of higher-order terms in semi-empirical d-band models, giving closer approximations to the tight-binding model has been shown to be a systematic way of improving the angular characteristics of these models. These models, however, rapidly lose the two significant advantages they have over a formal low-order moment expansion of tight-binding, namely speed and forces. In terms of speed, Xu and Adams estimate that the inclusion of fourth-moment terms increase the computational power required by a factor of around 100. In terms of the forces, the difficulty of differentiating equations such as (112) have already been discused. The final word on this subject we leave with Moriarty (1990):

> 'In principle, such derivatives of the multi-ion potentials v_3 and v_4 may always be calculated analytically within the model GPT. In practice, we have succeeded in doing this for v_3 but not yet for v_4, where the result becomes algebraically complicated. . . .'

7.2. Low-order moments approximations to sp semiconductors

The form of (108) makes it clear that the second-moment formalism of section 7.1 cannot be applied to materials in which angular bonding is known to play a crucial role. The difficulty of obtaining forces from the higher-moment methods, discussed above, has meant that until recently it has not been possible to make close contact between tight-binding and the empirical potential of Tersoff (1988a, b). The inter-site formulation of the bond-order potential (Pettifor 1989, Horsfield *et al* 1996), however, provides a mechanism by which such contact can be made, and through which it is, in principle, possible to derive similar potentials. This approach was first described in Alinaghian *et al* (1993); further discussion is found in Horsfield *et al* (1996). The inter-site energy may be written

$$E_{\rm band}^{(ij)} = 2\sum_{\alpha\beta} \rho_{i\alpha,j\beta} H_{j\beta,i\alpha}. \tag{119}$$

Choosing the angular momentum quantization direction to be along the bond, and assuming $V_{\mathrm{sp}\sigma}(r_{ij}) = \sqrt{|h_{\mathrm{ss}\sigma}(r_{ij})h_{\mathrm{pp}\sigma}(r_{ij})|}$, the energy may be rewritten as a σ and a π energy

$$E_{\mathrm{band}}^{(ij)} = [h_{\mathrm{ss}\sigma}(r_{ij}) - h_{\mathrm{pp}\sigma}(r_{ij})]\rho_{i\sigma,j\sigma} + 2h_{\mathrm{pp}\pi}(r_{ij})\rho_{i\pi,j\pi} \tag{120}$$

in terms of the normalized orbitals, resolved such that the bond lies along the z direction,

$$\begin{aligned} |i\sigma\rangle &= \frac{\sqrt{|h_{\mathrm{ss}\sigma}(r_{ij})|}|is\rangle + \sqrt{h_{\mathrm{pp}\sigma}(r_{ij})}|iz\rangle}{\sqrt{|h_{\mathrm{ss}\sigma}(r_{ij}) + h_{\mathrm{pp}\sigma}(r_{ij})|}} \\ |j\sigma\rangle &= \frac{\sqrt{|h_{\mathrm{ss}\sigma}(r_{ij})|}|js\rangle - \sqrt{h_{\mathrm{pp}\sigma}(r_{ij})}|jz\rangle}{\sqrt{|h_{\mathrm{ss}\sigma}(r_{ij}) + h_{\mathrm{pp}\sigma}(r_{ij})|}}. \end{aligned} \tag{121}$$

The lowest-order approximation to the inter-site density matrix is given by

$$\rho_{i\alpha,j\beta} = -2\chi_{00,00}(E_f)\frac{H_{i\alpha,j\beta}}{b_1} \tag{122}$$

where b_1 is the recursion coefficient (see section 3.2), and the response function, χ, is given in terms of the Green's function, (A10). In this context, $\chi_{00,00}$ may be approximated to take the same value as it takes in a linear chain (Pettifor 1989)

$$\chi_{00,00} = \frac{1}{\pi}\left(\sin\phi_f - \frac{1}{3}\sin 3\phi_f\right) \tag{123}$$

with ϕ_f fixed by the number of electrons per atom, N,

$$N = \left(\frac{2\phi_f}{\pi}\right)\left(1 - \frac{\sin 2\phi_f}{2\phi_f}\right). \tag{124}$$

From the relationship between the recursion coefficient and the bond order, we rewrite (122)

$$\begin{aligned} \rho_{i\sigma,j\sigma} &= -2\chi_{00,00}(E_f)\frac{H_{i\sigma,j\sigma}}{\sqrt{\frac{1}{2}(\mu_{i\sigma}^{(2)} + \mu_{j\sigma}^{(2)})}} \\ \rho_{i\pi,j\pi} &= -2\chi_{00,00}(E_f)\frac{H_{i\pi,j\pi}}{\sqrt{\frac{1}{2}(\mu_{i\pi}^{(2)} + \mu_{j\pi}^{(2)})}} \end{aligned} \tag{125}$$

where $\mu_{i\sigma}^{(2)} = \langle i\sigma|\hat{H}^2|i\sigma\rangle$ and $\mu_{i\pi}^{(2)} = \frac{1}{2}(\mu_{ip_x}^{(2)} + \mu_{ip_y}^{(2)})$ and similar definitions.

The link to empirical potentials of the type due to Tersoff (1988a, b) becomes clearer when the bond-order terms are explicitly evaluated, using the symmetries in Slater and Koster (1954) table 1. The derivation will not be given here; see Horsfield *et al* (1996) and references therein. The result is a sum over pairs of neighbouring atoms, j and k, where the sum is taken for distinct values of i, j and k only:

$$\begin{aligned} \mu_{i\sigma}^{(2)} &= \sum [h_{\mathrm{ss}\sigma}(r_{ik}) - h_{\mathrm{pp}\sigma}(r_{ik}))]^2 g_\sigma(\theta) + [h_{\mathrm{ss}\sigma}(r_{ij}) - h_{\mathrm{pp}\sigma}(r_{ij})]^2 \\ \mu_{i\pi}^{(2)} &= \sum h_{\mathrm{pp}\pi}^2(r_{ij}) g_\pi(\theta) h_{\mathrm{pp}\pi}^2(r_{ik}) \end{aligned} \tag{126}$$

where the angular functions are given in terms of the angle θ_{jik}, and the ratios of hopping integrals, $\mathrm{p}_\sigma = h_{\mathrm{pp}\sigma}/|h_{\mathrm{ss}\sigma}|$ and $\mathrm{p}_\pi = |h_{\mathrm{pp}\pi}|/|h_{\mathrm{ss}\sigma}|$ (Alinaghian *et al* 1993) as

$$\begin{aligned} g_\sigma(\theta) &= \frac{(1 + \mathrm{p}_\sigma\cos\theta)^2}{(1 + \mathrm{p}_\sigma)^2} + \frac{\mathrm{p}_\sigma \mathrm{p}_\pi^2 \sin^2\theta}{(1 + \mathrm{p}_\sigma)^3} \\ g_\pi(\theta) &= \tfrac{1}{2}[(\mathrm{p}_\sigma/\mathrm{p}_\pi^2)(1 + \mathrm{p}_\sigma)\sin^2\theta + (1 + \cos^2\theta)] \end{aligned} \tag{127}$$

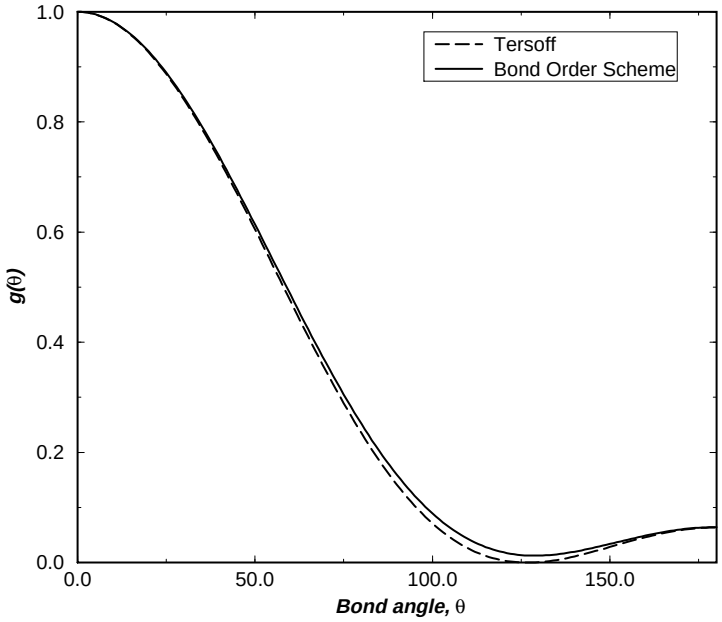

Figure 9. The angular functions g_σ in the Tersoff potential (1988a, b) and the bond-order potential (Pettifor 1989, Alinaghian *et al* 1993 and Horsfield *et al* 1996). From Horsfield *et al* (1996), reproduced by kind permission.

which Horsfield *et al* (1996) rewrote as

$$\begin{aligned} g_\sigma(\theta) &= a + b\cos\theta + c\cos 2\theta \\ g_\pi(\theta) &= d + e\cos 2\theta \end{aligned} \tag{128}$$

where a, b, c, d and e are constants depending upon the relative values of the hopping integrals, the form of which may be derived by comparison of (127) and (128).

Equations (120)–(128) provide an appropriate functional form for an empirical potential based upon the second-moment approximation. The success of Tersoff-style potentials (Tersoff 1988a, b) is due to the close similarity between this form and his; in particular it has been noted that the g_σ function used by Tersoff and that above are extremely similar, as shown in figure 9.

8. Conclusions

The tight-binding approximation provides a methodology for atomistic modelling in which the quantum mechanical nature of the chemical bond is explicitly included, but which is computationally far less demanding than *ab initio* density functional theory calculations. The price paid is a series of approximations which have the effect of restricting the transferability of a given parametrization. These were discussed in section 2.

A number of different methods of solving the tight-binding Hamiltonian have been developed. An issue of particular importance, especially with the increase in computer power available, is to take advantage of the physically short-ranged nature of quantum mechanical effects to obtain a linear scaling of computational time required with system size. Recursion-based methods have achieved this with some success for metallic systems; more

recently density-matrix minimization-based methods have extended this with great success to semiconducting materials. Several of these approaches were discussed in section 3.

The tight-binding method, in one form or another, has been successfully applied to metallic and semiconducting systems, including those with variable electronic character, such as the fullerene 'buckyballs' and 'buckytubes'. Some progress has also been made in the application of tight-binding to ionic materials. A selection of examples of the systems in which tight-binding has been used, and the particular issues associated with each system, were discussed in section 5.

Non-self-consistent, orthogonal tight-binding, solved by matrix diagonalization, lies in the gap in the modelling continuum between *ab initio* techniques and empirical potentials. As computer power has increased, the natural links between tight-binding and localized orbitals density functional theory have led to the development of *ab initio* tight-binding methods; at the same time the desire to produce linear-scaling density-function theory methods has led to what one might call tight-binding-like *ab initio* methods such as that of Hernandez *et al* (1996). The gap between tight-binding and density-functional theory can therefore be seen to be closing from both sides. The effort to improve the accuracy of tight-binding by eliminating the approximations involved was discussed in section 6.

Links have also been made in the opposite direction, with the development of low-order approximations to tight-binding which can be used to improve the functional form of empirical potentials. This work has been hindered by the extreme complexity of functional forms that rapidly results; recent developments in the BOP methods may provide a way forward. These issues were discussed in section 7.

Acknowledgments

We would like to thank all those who have contributed to our understanding of this field. In alphabetical order, these include Masato Aoki, Alex Bratkovsky, Mike Fearn, Mike Gillan, Paul Godwin, Stefan Goedecker, Martin Head-Gordon, Janusz Holender, Andrew Horsfield, Florian Kirchoff, Tony Paxton, David Pettifor, Adrian Sutton and David Vanderbilt.

Our tight-binding calculations were performed using the computing facilities of the Materials Modelling Laboratory, Oxford University.

Appendix. Further details for recursion methods

In this appendix, some of the more mathematical details of the recursion methods covered in section 3 are given. The intention is to describe the intricate workings, while still preserving physical insight. It is still recommended that the original references be consulted for exact details.

A.1. Recursion and continued fractions

In section 3.2, the local density of states for a single orbital, $n_{i\alpha}(E)$, was derived in terms of a Green's function. This Green's function can be written in terms of a continued fraction, whose components are related to the elements of the tridiagonalized Hamiltonian of the system (Haydock 1980); the element $G_{00}(Z)$ (where $G_{nm}(Z) = \langle U_n|\hat{G}(Z)|U_m\rangle$) can be

found from:

$$G_{00}(Z) = \cfrac{1}{Z - a_0 - \cfrac{b_1^2}{Z - a_1 - \cfrac{b_2^2}{Z - a_2 - \cfrac{b_3^2}{\ddots}}}}. \tag{A1}$$

The Lanczos recursion algorithm (Lanczos 1950) is an efficient scheme for tridiagonalizing a matrix. Consider a matrix H, which corresponds to the operator $\hat{H}$. Let the tridiagonal matrix be $\mathbf{H}'$, whose diagonal elements are given by a_n and whose off-diagonal elements are given by b_n. If the states which tridiagonalise the matrix are $|U_n\rangle$, then:

$$H'_{mn} = \langle U_m|\hat{H}|U_n\rangle = \begin{cases} a_n & \text{if } m = n \\ b_n & \text{if } m = n-1 \\ b_{n+1} & \text{if } m = n+1 \\ 0 & \text{otherwise.} \end{cases} \tag{A2}$$

The Lanczos recursion algorithm for tridiagonalizing a matrix is based on the following recursion relation:

$$\hat{H}|U_n\rangle = a_n|U_n\rangle + b_n|U_{n-1}\rangle + b_{n+1}|U_{n+1}\rangle \tag{A3}$$

and the condition that the tridiagonalizing states are orthonormal ($\langle U_n|U_m\rangle = \delta_{n,m}$). To apply this relation to a matrix, a starting state (which will generally be a tight-binding basis function) is chosen. This gives a_0 immediately from (A2) ($a_0 = \langle U_0|\hat{H}|U_0\rangle$). The construction requires that b_0 be zero, so the product $b_1|U_1\rangle$ can be found from (A3). From the orthonormality imposed on $|U_n\rangle$, the value of b_1 can be obtained, and $|U_1\rangle$ can be found. This then can be applied to the process again, giving a_1, b_2 and $|U_2\rangle$ and so on. If $|U_0\rangle$ is set equal to a basis function, $|\psi_{i\alpha}\rangle$, then $G_{i\alpha i\alpha}$ can be obtained from the continued fraction given in equation (A1), and from $G_{i\alpha i\alpha}$ the LDOS can be obtained from (36).

For an infinite system, there are an infinite number of levels in the continued fraction, so the exact values are replaced by estimated ones after a certain level. The simplest, and most widely used, approximation is to take $a_n = a_\infty, b_n = b_\infty$ for $n > N$, where N is the number of exact levels, and a_∞ and b_∞ are constants defining the band centre and width (Beer and Pettifor 1984). These constant terms can be summed exactly to give the square root terminator:

$$\begin{aligned} t(Z) &= \cfrac{1}{Z - a_\infty - \cfrac{b_\infty^2}{Z - a_\infty - \cfrac{b_\infty^2}{Z - a_\infty - \cfrac{b_\infty^2}{\ddots}}}} \\ &= \frac{1}{b_\infty}\left[\left(\frac{Z - a_\infty}{2b_\infty}\right) - \mathrm{i}\sqrt{1 - \left(\frac{Z - a_\infty}{2b_\infty}\right)^2}\right]. \end{aligned} \tag{A4}$$

Since the above may not bear any obvious relation to a moments method, it is useful to make the following observation:

$$\mu_{i\alpha}^{(n)} = \langle i\alpha|\hat{H}^n|i\alpha\rangle$$

$$= \langle U_0|\hat{H}^n|U_0\rangle$$
$$= \sum_{m_1 \ldots m_{n-1}} \langle U_0|\hat{H}|U_{m_1}\rangle\langle U_{m_1}|\hat{H}|U_{m_2}\rangle \ldots \langle U_{m_{n-1}}|\hat{H}|U_0\rangle. \tag{A5}$$

The first few moments defined by the above are:

$$\begin{aligned}
\mu_{i\alpha}^{(0)} &= 1 \\
\mu_{i\alpha}^{(1)} &= a_0 \\
\mu_{i\alpha}^{(2)} &= a_0^2 + b_1^2 \\
\mu_{i\alpha}^{(3)} &= a_0^3 + 2a_0 b_1^2 + a_1 b_1^2 \\
\mu_{i\alpha}^{(4)} &= a_0^4 + 3a_0^2 b_1^2 + 2a_0 a_1 b_1^2 + a_1^2 b_1^2 + b_1^2 b_2^2 + b_1^4.
\end{aligned} \tag{A6}$$

These equations can be inverted to give the recursion coefficients in terms of the moments. Every extra moment that is calculated allows another recursion coefficient to be found. However, such an inversion becomes numerically unstable beyond about 20 moments. It is often desirable to go higher than this in numbers of moments; in such a case (A3) is used, as it is stable for many levels.

A.2. The two-site BOP expansion

In section 3.2.4, the form for $G_{i\alpha,j\beta}$ of:

$$G_{i\alpha,j\beta}(Z) = \frac{\partial G_{00}^{\lambda}(Z)}{\partial \lambda} \tag{A7}$$

was derived. This expression can be further expanded, by using the chain rule and the dependence of G_{00}^{λ} on the recursion coefficients, $a_n^{\lambda}, b_n^{\lambda}$,

$$G_{i\alpha,j\beta} = \sum_{n=0}^{\infty} \frac{\partial G_{00}^{\lambda}}{\partial a_n^{\lambda}} \delta a_n^{\lambda} + \sum_{n=1}^{\infty} \frac{\partial G_{00}^{\lambda}}{\partial b_n^{\lambda}} \delta b_n^{\lambda} \tag{A8}$$

where $\delta a_n^{\lambda} = \partial a_n^{\lambda}/\partial\lambda$ and $\delta b_n^{\lambda} = \partial b_n^{\lambda}/\partial\lambda$. This gives an expression for the bond order,

$$\Theta_{i\alpha,j\beta} = \sum_{n=0}^{\infty} \frac{\partial N^{\lambda}}{\partial a_n^{\lambda}} \delta a_n^{\lambda} + \sum_{n=1}^{\infty} \frac{\partial N^{\lambda}}{\partial b_n^{\lambda}} \delta b_n^{\lambda} = -2\left[\sum_{n=0}^{\infty} \chi_{0n,n0}(E_f)\delta a_n^{\lambda} + \sum_{n=1}^{\infty} \chi_{0(n-1),n0} 2\delta b_n^{\lambda} \right]. \tag{A9}$$

The functions $\chi_{0m,n0}(E_f)$ are called *response functions* and are calculated from:

$$\chi_{0m,n0}(E_f) = \frac{1}{\pi} \lim_{\eta\to 0} \mathrm{Im} \int^{E_f} \mathrm{d}E\, G_{0m}^{\lambda}(E + \mathrm{i}\eta) G_{n0}^{\lambda}(E + \mathrm{i}\eta) \tag{A10}$$

and the Green's functions are derived from the recursion relation

$$(Z - a_n^{\lambda})G_{nm}^{\lambda}(Z) - b_n^{\lambda} G_{n-1,m}^{\lambda}(Z) - b_{n+1}^{\lambda} G_{n+1,m}^{\lambda}(Z) = \delta_{n,m} \tag{A11}$$

using the fact that $G_{0n} = G_{n0}$, and finding the starting value of G_{00}^{λ} from the continued fraction in (A1). By expanding out the derivatives of the recursion coefficients (δa_n and δb_n) with respect to the moments, the link back to moments methods can be seen:

$$\delta a_n^{\lambda} = \frac{\partial a_n^{\lambda}}{\partial \lambda} = \sum_{r=1}^{2n-1} \frac{\partial a_n^{\lambda}}{\partial \mu_{\lambda}^{(r)}} \frac{\partial \mu_{\lambda}^{(r)}}{\partial \lambda} = \sum_{r=1}^{2n-1} \frac{\partial a_n^{\lambda}}{\partial \mu_{\lambda}^{(r)}} \zeta_{i\alpha j\beta}^{r-1} \tag{A12}$$

$$\delta b_n^{\lambda} = \frac{\partial b_n^{\lambda}}{\partial \lambda} = \sum_{r=1}^{2n-1} \frac{\partial b_n^{\lambda}}{\partial \mu_{\lambda}^{(r)}} \frac{\partial \mu_{\lambda}^{(r)}}{\partial \lambda} = \sum_{r=1}^{2n-1} \frac{\partial b_n^{\lambda}}{\partial \mu_{\lambda}^{(r)}} \zeta_{i\alpha j\beta}^{r-1} \tag{A13}$$

where $\mu_\lambda^{(r)} = \langle U_0^\lambda | \hat{H}^r | U_0^\lambda \rangle$ and $\zeta_{i\alpha j\beta}^{r-1} = \langle i\alpha | \hat{H}^r | j\beta \rangle$, an *interference term.*

The physical interpretation for (A9) is as follows. The dependence of $\Theta_{i\alpha,j\beta}$ on the number of electrons in the bond is expressed in the functions $\chi_{0m,n0}$. These functions have a weak dependence on atomic coordination; this enters through the derivatives of the recursion coefficients, δa_n and δb_n.

A.3 The single-site BOP expansion

In section 3.2.5, the central part of the one-site BOP expansion was written:

$$G_{00}^\Lambda = G_{i\alpha,i\alpha} \tag{A14}$$

As G_{00}^Λ can clearly be written as a continued fraction using the recursion method as described above, then much of the formalism developed for the recursion method can be used here. Using the equivalent of (A8), the expression for an off-site Green's function element can be rewritten as

$$G_{i\alpha,j\beta} = \sum_{n=0}^{\infty} G_{0n}^\Lambda(Z) G_{n0}^\Lambda(Z) \frac{\partial a_n^\Lambda}{\partial \Lambda_{i\alpha,j\beta}} + 2 \sum_{n=1}^{\infty} G_{0,(n-1)}^\Lambda(Z) G_{n,0}^\Lambda(Z) \frac{\partial b_n^\Lambda}{\partial \Lambda_{i\alpha,j\beta}} \tag{A15}$$

where G_{0n}^Λ is given by the recursion relation shown in (A11). By integrating this as before (cf (A9)), the following expression for the bond order can be found:

$$\Theta_{i\alpha,j\beta} = -2 \left[\sum_{n=0}^{\infty} \chi_{0n,n0}^\Lambda \frac{\partial a_n^\Lambda}{\partial \Lambda_{i\alpha,j\beta}} + 2 \sum_{n=1}^{\infty} \chi_{0(n-1),n0}^\Lambda \frac{\partial b_n^\Lambda}{\partial \Lambda_{i\alpha,j\beta}} \right] \tag{A16}$$

with $\chi_{0m,n0}$ found from (A10).

The derivatives of the recursion coefficients, $\partial a_n^\Lambda / \partial \Lambda_{i\alpha,j\beta}$ and $\partial b_n^\Lambda / \partial \Lambda_{i\alpha,j\beta}$ can be evaluated simply and stably using a recurrence relation. There are a number of other issues which are too complicated for treatment here: the proof that within the new formalism the intersite and site-diagonal expressions for the energy are equivalent; the rotational invariance of the new formalism (obtained by working with moments averaged over the magnetic quantum number); and the truncation of the recursion series for the derivatives of the recursion coefficients. For a discussion of all of the above issues, the reader is again directed to Aoki (1993) and Horsfield *et al* (1996).

References

Ackland G J and Thetford R 1987 *Phil. Mag.* A **56** 15

Adams G B and Sankey O F 1991 *Phys. Rev. Lett.* **67** 867

Adams G B, Page J B, Sankey O F, Sinha K, Menendez J and Huffman D R 1991 *Phys. Rev.* B **44** 4052

Agrawal B K, Agrawal S, Yadav P S and Negi J S 1990 *J. Phys.: Condens. Matter* **2** 6519

Alinaghian P, Gumbsch P, Skinner A J and Pettifor D G 1993 *J. Phys.: Condens. Matter* **5** 5795

Allan G and DeLerue C 1992 *J. Vac. Sci. Technol.* B **10** 1928

Allen M P and Tildesley D J 1987 *Computer Simulation of Liquids* (Clarendon Press: Oxford)

Allen P B, Broughton J Q and McMahan A K 1986 *Phys. Rev.* B **34** 859

Andersen O K 1975 *Phys. Rev.* B **12** 3060

——1984 *The Electronic Structure of Complex Systems* (New York: Plenum)

Andersen O K and Jespen O 1984 *Phys. Rev. Lett.* **53** 2571

Andriotis A N, Lathiotakis N and Menon M 1996a *Chem. Phys. Lett.* **260** 15

——1996b *Europhys. Lett.* **36** 37

Aoki M 1993 *Phys. Rev. Lett.* **71** 3842

Aoki M and Pettifor D G 1993 *Physics of Transition Metals* ed P M Oppeneer and J Kübler (World Scientific: Singapore)

Arriaga J, Garciamoliner F and Velasco V R 1993a *J. Phys.: Condens. Matter* **5** 5429

——1993b *Prog. Surf. Sci.* **42** 271
Arriaga J, Muñoz M C, Velasco V R and Garciamoliner F 1991 *Phys. Rev. B* **43** 9626
Avery A R, Goringe C M, Holmes D M, Sudijono J L and Jones T S 1995 *Phys. Rev. Lett.* **76** 3344
Ballone P and Jones R O 1995 *Chem. Phys. Lett.* **233** 632
Bechstedt F and Reichardt D 1988a *Surf. Sci.* **202** 58
——1988b *Surf. Sci.* **202** 83
Bechstedt F and Scheffler M 1993 *Surf. Sci. Rep.* **18** 145
Bechstedt F, Reichardt D and Enderlein R 1985 *Phys. State Solidi* b **131** 643
Beer N and Pettifor D G 1984 *Structure and Phase Stability of Alloys* ed P Phariseau and W Temmerman (New York: Plenum)
Bertho D, Boiron D, Simon A, Jouanin C and Priester C 1991 *Phys. Rev.* B **44** 6118
Bifone A 1993 *Solid State Comm.* **85** 409
Bigger J R K, McInnes D A, Sutton A P, Payne M C, Stich I, King-Smith R D, Bird D M and Clarke L J 1992 *Phys. Rev. Lett.* **69** 2224
Biswas R, Wang C Z, Chan C T, Ho K M and Soukoulis C M 1989 *Phys. Rev. Lett.* **63** 1491
Bloch F 1928 *Z. Phys.* **52** 555
Bouarab S, Vega A, Alonso J A and Iniguez M P 1996 *Phys. Rev.* B **54** 3003
Bowler D R, Aoki M, Goringe C M, Horsfield A P and Pettifor D G 1997 *Modelling Simul. Mater. Sci. Eng.* **5** 199
Capaz R B, Dearaujo G C, Koiller B and Vonderweid J P 1993 *J. Appl. Phys.* **74** 5531
Carlsson A E 1990 *Advances in Research and Application (Solid State Physics 43)* ed H Ehrenreich and D Tumbull (New York: Academic Press)
——1991 *Phys. Rev.* B **44** 6590
Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
Carroll D L, Redlich P, Ajayan P M, Charlier J C, Blase X, De Vita A and Car R 1997 *Phys. Rev. Lett.* **78** 2811
Chadi D J 1977 *Phys. Rev.* B **16** 790
——1978 *Phys. Rev. Lett.* **41** 1062
——1979a *Phys. Rev.* B **19** 2074
——1979b *J. Vac. Sci. Technol.* **16** 1290
——1984a *Phys. Rev.* B **29** 785
——1984b *Phys. Rev. Lett.* **52** 1911
——1986 *Phys. Rev. Lett.* **57** 102
——1987 *Phys. Rev. Lett.* **59** 1691
——1989 *Ultramicroscopy* **31** 1
——1990 *Phys. Rev.* B **41** 10 595
——1994 *Surf. Sci.* **300** 311
Chadi D J and Chelikowsky J R 1981 *Phys. Rev.* B **24** 4892
Chadi D J and Cohen M L 1975 *Phys. State Solidi* **68** 405
Chadi D J and Martin R M 1976 *Solid State Commun.* **19** 643
Chan B C and Ong C K 1991 *J. Phys. Chem. Solids* **52** 699
Charlier J C, Lambin P and Ebbesen T W 1996 *Phys. Rev.* B **54** 8377
Chelikowsky J R and Cohen M L 1974 *Phys. Rev.* B **10** 5095
Chen A B, Sher A and Yost W T 1992 *Semiconductors and Semimetals* **37** 1
Chico L, Benedict L X, Louie S G and Cohen M L 1996 *Phys. Rev.* B **54** 2600
Christensen O B and Cohen M L 1993 *Phys. Rev.* B **47** 13 643
Christensen O B and Jacobsen K W 1993 **5** 5591
Cleri F and Rosato V 1993 *Phys. Rev.* B **48** 22
Colombo L and Maric D 1995 *Europhys. Lett.* **29** 623
Daw M 1993 *Phys. Rev.* B **47** 10 895
Daw M S and Baskes M I 1984 *Phys. Rev.* B **29** 6443
De Sandre G, Colombo L and Bottani C 1996 *Phys. Rev.* B **54** 11 857
Dorantes-Dávila J, Vega A and Pastor G M 1993 *Phys. Rev.* B **47** 12 995
Dorantes-Dávila J and Pastor G M 1995 *Phys. Rev.* B **51** 16 627
Drabold D A, Dow J D, Fedders P A, Carlsson A E and Sankey O F 1990a *Phys. Rev.* B **42** 5345
Drabold D A, Dow J D, Fedders P A, Klemm S and Sankey O F 1991a *Phys. Rev. Lett.* **67** 2179
Drabold D A, Fedders P A, Sankey O F and Dow J D 1990b *Phys. Rev.* B **42** 5135
Drabold D A, Wang R, Klemm S, Sankey O F and Dow J D 1991b *Phys. Rev.* B **43** 5132
Dreysse H, Dorantes-Dávila J, Pick S and Pastor G M 1994a *J. Appl. Phys.* **76** 6328
Dreysse H, Vega A, Demangeat C and Balbas L C 1994b *Europhys. Lett.* **27** 165

Ducastelle F and Cyrot-Lackmann F 1970 *J. Phys. Chem. Solids* **31** 129
Duran J C, Muñoz A and Flores F 1987 *Phys. Rev.* B **35** 7721
Ekpenuma S N and Myles C W 1990 *J. Phys. Chem. Solids* **51** 93
Engdahl C, Stoltze P, Jacobsen K W, Norskov J K, Skriver H L and Alden M 1994 *J. Vac. Sci. Technol.* A **12** 1787
Fedders P A and Drabold D A 1993 *Phys. Rev.* B **47** 13 277
Fedders P A, Drabold D A and Klemm S 1992 *Phys. Rev.* B **45** 4048
Feng Y P, Wee T H, Ong C K and Poon H C 1996 *Phys. Rev.* B **54** 4766
Fernandez-Alvarez L, Monsivais G and Velasco V R 1996 *J. Phys.: Condens. Matter* **8** 8859
Finnis M W and Sinclair J E 1984 *Phil. Mag.* A **50** 45
——1986 *Phil. Mag.* A **53** 161
Flores F, Rincon R, Ortega J, Garciavidal F J and Perez R 1993 *Prog. Surf. Sci.* **42** 281
Foiles S M 1992 *MRS Symp. Proc.* **278** 339
Foulkes W M C 1987 *PhD Thesis* University of Cambridge
Foulkes W M C and Haydock R 1989 *Phys. Rev.* B **39** 12 520
Fowler P W, Heine T, Mitchell D, Schmidt R and Seifert G 1996 *J. Chem. Soc. Faraday Trans.* **92** 2197
Frauenheim T, Weich F, Köler T, Uhlmann S, Porezag D and Seifert G 1995 *Phys. Rev.* B **52** 11 492
Frenkel D and Smit B 1996 *Understanding Computer Simulation* (San Diego, CA: Academic)
Friedel J 1954 *Adv. Phys.* **3** 446
——1969 *The Physics of Metals* (Cambridge: Cambridge University Press) p 494
Froyen S and Harrison W A 1979 *Phys. Rev.* B **20** 2420
Galli G 1996 *Curr. Op. Solid State Mat. Sci.* **1** 864
Galli G and Mauri F 1994 *Phys. Rev. Lett.* **73** 3471
Galli G and Parrinello M 1992 *Phys. Rev. Lett.* **69** 3547
Gillan M J 1989 *J. Phys.: Condens. Matter* **1** 689
Glanville S, Paxton A T and Finnis M W 1988 *J. Phys. F: Met. Phys.* **18** 693
Godin T J, LaFemina J P and Duke C B 1992 *J. Vac. Sci. Technol.* A **10** 2059
Goedecker S and Colombo L 1994 *Phys. Rev. Lett.* **73** 122
Goedecker S and Teter M 1995 *Phys. Rev.* B **51** 9455
Goniakowski J, Bouette-Russo S and Noguera C 1993 *Surf. Sci.* **284** 315
Goodwin L, Skinner A J and Pettifor D G 1989 *Europhys. Lett.* **9** 701
Goringe C M 1995 *D Phil Thesis* University of Oxford
——1997 *Mod. Sim. Mat. Sci. Eng.* submitted
Graf M and Vogl P 1995 *Phys. Rev.* B **51** 4940
Guinea F, Sánchez-Dehesa J and Flores F 1983 *J. Phys. C: Solid State Phys.* **16** 6499
Haas H, Elsasser C and Fahnle M 1996 *J. Phys.: Condens. Matter* **8** 10 353
Hanke M, Strehlow R and Kühn W 1986 *Acta Phys. Pol.* A **69** 971
Haroun A, Chouairi A, Ouannasser S, Dreysse H, Fabricius G and Llois A M 1994 *Surf. Sci.* **309** 1087
Harris J 1985 *Phys. Rev.* B **31** 1770
Harrison W A 1980 *Electronic Structure and the Properties of Solids* (San Francisco, CA: Freeman)
——1985 *Phys. Rev.* B **31** 2121
Harrison W A and Klepeis J E 1988 *Phys. Rev.* B **37** 864
Hasbun J E, Singh V A and Roth L M 1987 *Phys. Rev.* B **35** 2988
Hausleitner C and Hafner J 1992a *Phys. Rev.* B **45** 115
——1992b *Phys. Rev.* B **45** 128
——1992c *Phys. Rev.* B **45** 1571
Haydock R 1980 *Solid State Phys.* **35** 216
Heine V 1980 *Solid State Phys.* **35** 1
Hernández E, Gillan M J and Goringe C M 1996 *Phys. Rev.* B **53** 7147
Hoffmann R 1963 *J. Chem. Phys.* **39** 1397
Holender J M and Morgan G J 1991 *J. Phys.: Condens. Matter* **3** 7241
——1992 *J. Phys.: Condens. Matter* **4** 4473
——1994 *Modelling Simul. Mater. Sci. Eng.* **2** 1
Holender J M, Morgan G J and Jones R 1993 *Phys. Rev.* B **47** 3991
Horsfield A P 1996 *Mater. Sci. Eng.* B **37** 219
Horsfield A P and Bratkovsky A M 1996 *Phys. Rev.* B **53** 15 381
Horsfield A P, Bratkovsky A M, Fearn M, Pettifor D G and Aoki M 1996 *Phys. Rev.* B **53** 12 694
Hurley A C 1976 *Electron Correlation in Small Molecules* (London: Academic)
Iijima S 1991 *Nature* **354** 56

Inoue J and Ohta Y 1987 *J. Phys. C: Solid State Phys.* **20** 1947
Itoh H, Inoue J and Maekawa S 1993a *Phys. Rev.* B **47** 5809
——1993b *J. Magn. Magn. Mater.* **126** 479
Itoh S, Ordejón P, Drabold D A and Martin R M 1996 *Phys. Rev.* B **53** 2132
Jenkins D W and Dow J D 1989 *Phys. Rev.* B **39** 3317
Jenkins D W, Dow J D and Tsai M H 1992 *J. Appl. Phys.* **72** 4130
Jones R and Lewis M W 1984 *Phil. Mag. B* **49** 95
Kang BS and Sohn K S 1996 *Physica* **1-2B** 160
Khan F S and Broughton J Q 1989 *Phys. Rev.* B **39** 3688
Kim J N, Mauri F and Galli G 1995 *Phys. Rev.* B **52** 1640
Kim S G and Tománek 1994 *Phys. Rev. Lett.* **72** 2418
Kobayashi A, Sankey O F and Dow J D 1982 *Phys. Rev.* B **25** 6367
Köhler T, Frauenheim T and Jungnickel G 1995 *Phys. Rev.* B **52** 11 837
Kohn W 1959 *Phys. Rev.* **115** 809
Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
Kohyama M and Yamamoto R 1994a *Phys. Rev.* B **49** 17 102
——1994b *Phys. Rev.* B **50** 8502
Kohyama M, Yamamoto R, Ebata Y and Kinoshita M 1989 *Phys. Status Solidi* b **152** 533
Kohyama M, Kose S, Kinoshita M and Yamamoto R 1990a *J. Phys.: Condens. Matter* **2** 7809
——1990b *J. Phys.: Condens. Matter* **2** 7791
Kohyama M, Kose S and Yamamoto R 1991 *J. Phys.: Condens. Matter* **39** 7555
Kress J D and Voter A F 1995 *Phys. Rev.* B **52** 8766
Kreuch G and Hafner J 1995 *J. Non-Cryst. Solids* **189** 227
Krey U, Krompiewski S and Krauss U 1990 *J. Magn. Magn. Mater.* **86** 85
Krompiewski S, Krauss U and Krey U 1989 *Phys. Rev.* B **39** 2819
Krompiewski S, Krey U and Ostermeier H 1987 *J. Magn. Magn. Mater.* **69** 117
Krompiewski S, Krey U, Krauss U and Ostermeier H 1988 *J. Magn. Magn. Mater.* **73** 5
Kroto H W, Heath J R, O'Brien S C, Curl R F and Smalley R E, 1985 *Nature* **318** 162
Kühn W, Strehlow R and Hanke M 1987 *Phys. Status Solidi* B **141** 541
Kumagai M, Takagahara T and Hanamura E 1987 *Sol. Stat. Commun.* **64** 659
Kumar A and Singh I 1988 *Phys. Status Solidi* b **148** 549
Kutzelnigg W 1977 *Methods of Electronic Structure Theory* ed H F Schaefer (New York: Plenum) p 129
La Femina J P 1992 *Surf. Sci. Rep.* **16** 133
La Femina J P, Duke C B and Mailhiot C 1990 *J. Vac. Sci. Technol.* B **8** 888
Laasonen K and Nieminen R M 1990 *J. Phys. C: Solid State Phys.* **2** 1509
Lammers U and Borstel G 1994 *Phys. Rev.* B **49** 17 360
Lanczos C 1950 *J. Res. Natl Bur. Stand.* **45** 225
Lanzavecchia S and Colombo L 1996 *Europhys. Lett.* **36** 295
Lee S and Dow J D 1987 *Phys. Rev.* B **36** 5968
Lee S, Dow J D and Sankey O F 1985 *Phys. Rev.* B **31** 3910
Lindholm E and Lundqvist S 1985 *Phys. Scr.* **32** 220
Liu F 1995 *Phys. Rev.* B **52** 10 677
Liu F, Mostoller M, Milman V, Chisholm M F and Kaplan T 1995 *Phys. Rev.* B **51** 17 192
Li W G and Myles C W 1991 *Phys. Rev.* B **43** 9947
Li X P, Nunes W and Vanderbilt D 1993 *Phys. Rev.* B **47** 10 891
Lorentz R and Hafner J 1995 *J. Magn. Magn. Mater.* **139** 209
Low K C and Ong C K 1994 *Phys. Rev.* B **50** 5352
Löwdin P 1950 *J. Chem. Phys.* **18** 365
Maeda K, Vitek V and Sutton A P 1982 *Acta Metall.* **30** 2001
Mailhiot C, Duke C B and Chadi D J 1984 *Phys. Rev.* B **53** 2114
Majewski J A and Vogl P 1986 *Phys. Rev. Lett.* **57** 1366
——1987 *Phys. Rev.* B **35** 9666
Marchese M, Jacucci G and Flynn C P 1988 *Phil. Mag. Lett.* **57** 25
Marshall T S and Wilson T M 1994 *Phys. Rev.* B **50** 15 034
Martin-Moreno L and Verges J A 1989 *Phys. Rev.* B **39** 3445
Masudajindo K 1994 *Superlatt. Microstruct.* **16** 359
Mattheiss L F and Patel J R 1981 *Phys. Rev.* B **23** 5384
Mauger A, Bourgoin J C, Allan G, Lannoo M, Bourret A and Billard L 1987 *Phys. Rev.* B **35** 1267

Mauri F and Galli G 1994 *Phys. Rev.* B **50** 4316
Mauri F, Galli G and Car R 1993 *Phys. Rev.* B **47** 9973
McKinnon B A and Choy T C 1995 *Phys. Rev.* B **52** 14 531
McWeeny R 1960 *Rev. Mod. Phys.* **32** 335
Menon M and Connolly J 1994 *Phys. Rev.* B **50** 8903
Menon M and Subbaswamy K R 1991 *Phys. Rev. Lett.* **67** 3487
——1993 *Phys. Rev.* B **47** 12 754
——K R 1994 *Phys. Rev.* B **50** 11 577
Menon M, Richter E and Subbaswamy K R 1996 *J. Chem. Phys.* **104** 5875
Mercer J L and Chou M Y 1991 *Phys. Rev.* B **43** 6768
——1993 *Phys. Rev.* B **48** 5374
Mirabella D A, Aldao C M and Deza R R 1994 *Phys. Rev.* B **50** 12 152
Mitamura H, Sakakibara T, Goto T and Yamada H 1995 *J. Magn. Magn. Mater.* **144** 821
Mokrani A, Dreysse H, Bouarab S and Demangeat C 1992 *J. Magn. Magn. Mater.* **113** 201
Molteni C, Colombo L and Miglio L 1993 *Europhys. Lett.* **24** 659
——1994a *Phys. Rev.* B **50** 4371
——1994b *J. Phys.: Condens. Matter* **6** 5243
Monkhurst H J and Pack J D 1976 *Phys. Rev.* B **13** 5188
Moos T H, Hubner W and Bennemann K H 1996 *Solid State Comm.* **98** 639
Moriarty J A 1972 *Phys. Rev.* B **5** 2066
——1977 *Phys. Rev.* B **16** 2537
——1982 *Phys. Rev.* B **26** 1754
——1988 *Phys. Rev.* B **38** 3199
——1990 *Phys. Rev.* B **42** 1609
——1994 *Phys. Rev.* B **49** 12 431
Morris J R, Wang C Z and Ho K H 1995 *Phys. Rev.* B **52** 4138
Mpassimabiala B, Moraitis G, Demangeat C and Mokrani A 1996 *Surf. Sci.* **352** 907
Mulliken R S 1949a *J. Chem. Phys.* **46** 497
——1949b *J. Chem. Phys.* **46** 675
Muñoz A, Sánchez-Dehesa J and Flores F 1986 *Europhys. Lett.* **2** 385
Muñoz A, Durán J C and Flores F 1987 *Surf. Sci. Lett.* **181** L200
Muñoz M C, Velasco V R and Garciamoliner F 1989 *Phys. Rev.* B **39** 1786
Muñoz M C and Armelles G 1993 *Phys. Rev.* B **48** 2839
Nichols C S and Winer K 1988 *Phys. Rev.* B **38** 9850
Noguera C, Goniakowski J and Bouette-Russo S 1993 *Surf. Sci.* **287** 188
Nunes R W and Vanderbilt D 1994 *Phys. Rev.* B **50** 17 611
Nunes R W, Bennetto J and Vanderbilt D 1996 *Phys. Rev. Lett.* **77** 1516
O'Keeffe M, Adams G B and Sankey O F 1992 *Phys. Rev. Lett.* **68** 2325
Olmstead M A and Chadi D J 1986 *Phys. Rev.* B **33** 8402
Onwuagba B N 1994 *Sol. Stat. Comm.* **89** 289
Ordejón P, Drabold D A, Grumbach M P and Martin R M 1993 *Phys. Rev.* B **48** 14 646
Ordejón P, Lebedenko D and Menon M 1994 *Phys. Rev.* B **50** 5645
Osotchan T, Chin V W L, Vaughan M R, Tansley T L and Goldys E M 1994 *Phys. Rev.* B **50** 2409
Owen J H G, Bowler D R, Goringe C M, Miki K and Briggs G A D 1995 *Surf. Sci. Lett.* **341** L1042
Papaconstantopoulos D A 1986 *Handbook of the Band Structure of Elemental Solids* (New York: Plenum)
Papadia S, Piveteau B, Spanjaard D and Desjonqueres M C 1996 *Phys. Rev.* B **54** 14 720
Parks G A 1965 *Chem. Rev.* **65** 177
Parr R G and Pearson R G 1983 *J. Am. Chem. Soc.* **105** 7512
Pastor G M and Dorantes-Dávila J 1996 *Mater. Sci. Eng.* A **217** 286
Pastor G M, Dorantes-Dávila J and Benneman K H 1988 *Chem. Phys. Lett.* **148** 459
Paxton A T 1996 *J. Phys. D: Appl. Phys.* **29** 1689
Paxton A T and Sutton A P 1989 *Acta Metall.* **37** 1693
Paxton A T, Sutton A P and Nex C M M 1987 *J. Phys. C: Solid State Phys.* **20** L263
Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045
Perezdiaz J L and Muñoz M C 1994 *J. Appl. Phys.* **75** 6470
Pettifor D G 1977 *J. Phys. F: Met. Phys.* **7** 613
——1989 *Phys. Rev. Lett.* **63** 2480
Pettifor D G and Aoki M 1991 *Phil. Trans. R. Soc. London* A **334** 439

Pick S 1995a *Chem. Phys. Lett.* **239** 84
——1995b *Surf. Sci.* **333** 736
——1995c *J. Phys.: Condens. Matter* **7** 7729
——1995d *J. Phys. C: Solid State Phys.* **99** 15 375
——1996 *Surf. Sci.* **352** 300
Pick S and Dreysse H 1992 *Phys. Rev.* B **46** 5802
——1993 *Phys. Rev.* B **48** 13 588
——1996 *Solid State Commun.* **100** 67
Pick S, Dorantes-Dávila J, Pastor G M and Dreysse H 1994 *Phys. Rev.* B **50** 993
Piveteau B, Desjonqueres M C and Spanjaard D 1992a *J. Physique* **2** 1677
——1992b *Phys. Rev.* B **46** 7121
Pizzagalli L, Stoeffler D and Gautier F 1996 *Phys. Rev.* B **54** 12 216
Platero G, Sánchez-Dehesa J, Tejedor C and Flores F 1986 *Surf. Sci.* **168** 553
Porezag D, Frauenheim T, Köhler, Seifert G and Kaschner R 1995 *Phys. Rev.* B **51** 12 947
Poteau R and Spiegelmann F 1993 *J. Chem. Phys.* **98** 6540
Poteau R, Spiegelmann F and Labastie P 1994 *Z. Phys.* D **30** 57
Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes, The Art Of Scientific Computing* 2nd edn (Cambridge: Cambridge University Press)
Qian G-X and Chadi D J 1987 *Phys. Rev.* B **35** 1288
Qiu S-Y, Wang C Z, Ho K M and Chan C T 1994 *J. Phys.: Condens. Matter* **6** 9153
Rasband P B, Clancy P and Thompson M O 1996 *J. Appl. Phys.* **79** 8998
Razee S S A and Prasad R 1993 *Phys. Rev.* B **48** 1361
Rebonato R, Welch D O, Hatcher R D and Bilello J C 1987 *Phil. Mag.* A **55** 655
Ren S F and Chang Y C 1991 *Phys. Rev.* B **44** 13 573
Ren S Y, Dow J D and Yang G L 1992 *Phys. Rev.* B **45** 6628
Rey C, Gallego L J, Garciarodeja J, Alonso J A and Iniguez M P 1993 *Phys. Rev.* B **48** 8253
Robertson J 1983 *Adv. Phys.* **32** 361
——1992 *Phil. Mag.* **B 66** 615
Russo S and Noguera C 1992a *Surf. Sci.* **262** 245
——1992b *Surf. Sci.* **262** 259
Saito T and Ikoma T 1992 *Superlatt. Microstruct.* **12** 81
Sánchez-Dehesa J, Flores F and Guinea F 1984 *J. Phys. C: Solid State Phys.* **17** 2039
Sankey O F and Dow J D 1983 *Phys. Rev.* B **27** 7641
Sankey O F and Niklewski D J 1989 *Phys. Rev.* B **40** 3979
Sankey O F, Niklewski D J, Drabold D A and Dow J D 1990 *Phys. Rev.* B **41** 12 750
Sawada S 1990 *Vacuum* **41** 612
Schenter G K and LaFemina J P 1992 *J. Vac. Sci. Technol. A* **10** 2429
Seong H and Lewis L J 1995 *Phys. Rev.* B **52** 5675
——1996 *Phys. Rev.* B **53** 4408
Servalli G and Colombo L 1993 *Europhys. Lett.* **22** 107
Shen J, Dow J D and Ren S Y 1990 *J. Appl. Phys.* **67** 3761
Sieranski K and Szatkowski J 1993 *Solid State Commun.* **88** 663
Sieranski K, Szatkowski J and Misiewicz J 1994 *Phys. Rev.* B **50** 7331
Sigalas M M and Papaconstantopoulos D A 1994 *Phys. Rev.* B **49** 1574
Skriver H L 1984 *The LMTO Method* (New York: Springer)
Slater P C and Koster G F 1954 *Phys. Rev.* **94** 1498
Sluiter M H F and Singh P P 1994 *Phys. Rev.* B **49** 10 918
Smith D L and Mailhiot C 1990 *Rev. Mod. Phys.* **62** 173
Song E G, Kim E, Lee Y H and Hwang Y G 1993 *Phys. Rev.* B **48** 1486
Stauffer L, Habar M and Dreysse H 1992 *Surf. Sci.* **270** 1116
Stillinger F and Weber T 1985 *Phys. Rev.* B **31** 5262
Stoeffler D, Fabricius G, Llois A M and Dreysse H 1992 *Surf. Sci.* **27** 632
Stoner E C 1938 *Proc. R. Soc.* A **165** 372
Straub G K and Harrison W A 1989 *Phys. Rev.* B **39** 10 325
Strehlow R, Hanke M and Kühn W 1985 *Phys. Status Solidi* b **131** 631
Strehlow R, Kühn W and Hanke M 1986 *Phys. Status Solidi* b **134** 257
Stutzmann M, Street R A, Tsai C C, Boyce J B and Ready S E 1989 *J. Appl. Phys.* **66** 569
Süss F and Krey U 1993 *J. Magn. Magn. Mater.* **125** 351

Sutton A P, Finnis M W, Pettifor D G and Ohta Y 1988 *J. Phys. C: Solid State Phys.* **21** 35
Teng D, Shen J, Newman K E and Gu B L 1991 *J. Phys. Chem. Solids* **52** 1109
Tersoff J 1988a *Phys. Rev.* B **37** 6991
——1988b *Phys. Rev.* B **38** 9902
Thomson R E and Chadi D J 1984 *Phys. Rev.* B **29** 889
Tománek D and Schlüter M A 1986 *Phys. Rev. Lett.* **56** 1055
——1987 *Phys. Rev.* B **36** 1208
——1991 *Phys. Rev. Lett.* **67** 2331
Tsai M H, Sankey O F and Dow J D 1992 *Phys. Rev.* B **46** 10 464
Tserbak C, Polatoglou H M and Therodorou G 1993 *Phys. Rev.* B **47** 7104
Tsuei K D and Johnson P D 1992 *Phys. Rev.* B **45** 13 827
Tsymbal E Y and Pettifor D G 1996 *Phys. Rev.* B **54** 15 314
van Schilfgaarde M and Harrison W A 1985 *J. Phys. Chem. Solids* **46** 1093
——1986 *Phys. Rev.* B **33** 2653
Velasco V R and Garciamoliner F 1994 *Prog. Surf. Sci.* **46** 211
Virkkunen R, Laasonen K and Nieminen R M 1991 *J. Phys.: Condens. Matter* **3** 7455
Vogl P, Hjalmarson H J and Dow J D 1983 *J. Phys. Chem. Solids* **44** 365
Voter A F, Kress J D and Silver R N 1996 *Phys. Rev.* B **53** 12 733
Wang C Z, Ho K M and Chan C T 1993 *Phys. Rev.* B **47** 14 835
Wang C Z and Ho K M 1993 *Phys. Rev. Lett.* **71** 1184
——1996 *Adv. Chem. Phys.* **93** 651
Wang C Z, Chan C T and Ho K M 1990 *Phys. Rev.* B **42** 11 276
——1991 *Phys. Rev. Lett.* **66** 189
Wang C Z, Chan C T and Ho K M 1992 *Phys. Rev.* B **46** 9761
Wang E G, DeCarpigny J N and Allan G 1994 *J. Phys.: Condens. Matter* **6** 4999
Wang E G, Chen C F and Ting C S 1995 *J. Appl. Phys.* **78** 1832
Wang Y R and Duke C B 1987 *Phys. Rev.* B **36** 2763
Wee T H, Feng Y P, Ong C K and Poon H C 1996 *J. Phys.: Condens. Matter* **8** 6511
White C A, Maslen P, Lee M S and Head-Gordon M 1997 *Preprint*
Widany J, Frauenheim T, Köhler T, Sternberg M, Porezag D, Jungnickel G and Seifert G 1996 *Phys. Rev.* B **53** 4443
Wilson J H, Todd J D and Sutton A P 1990 *J. Phys.: Condens. Matter* **2** 10 259
Xu C and Scuseria G E 1994 *Phys. Rev. Lett.* **72** 669
Xu C H, Wang C Z, Chan C T and Ho K M 1992 *J. Phys. C: Solid State Phys.* **4** 6047
Xu W and Adams J B 1994 *Surf. Sci.* **301** 371
——1995 *Surf. Sci.* **339** 241
Xu W and Moriarty J A 1996 *Phys. Rev.* B **54** 6941
Xu Z Z 1993 *J. Phys.: Condens. Matter* **5** 9077
Yamada H and Shimizu M 1985 *J. Phys. F: Met. Phys.* **15** L175
——1989 *J. Phys.: Condens. Matter* **1** 2597
Yang S H, Drabold D A, Adams J B and Sachdev A 1993 *Phys. Rev.* B **47** 1567
Yu J, Kalia R K and Vashishta P 1993 *Appl. Phys. Lett.* **63** 3152
——1994 *Phys. Rev.* B **49** 5008
Zhang B L, Xu C H, Wang C Z, Chan C T and Ho K M 1992a *Phys. Rev.* B **46** 7333
Zhang B L, Wang C Z, Ho K M, Xu C H and Chan C T 1992b *J. Chem. Phys.* **97** 5007
——1993a *J. Chem. Phys.* **98** 3095
Zhang B L, Wang C Z and Ho K M 1993b *Phys. Rev.* B **47** 1643
Zhang B L, Wang C Z, Chan C T and Ho K M 1993c *Phys. Rev.* B **48** 11 381
Zhao J J, Han M and Wang G H 1993 *Phys. Rev.* B **48** 15 297
Zhao J J, Chen X S, Sun Q, Liu F Q, Wang G H and Lain K D 1995a *Physica* **215B** 377
Zhao J J, Chen X S, Sun Q, Liu F Q and Wang G H 1995b *Europhys. Lett.* **32** 113
Zhu Q S, Hiramatsu K, Sawaki N and Akasaki I 1992 *Phys. Status Solidi* b **172** 647