Linear scaling electronic structure methods

Stefan Goedecker*

Max-Planck Institute for Solid State Research, D-70569 Stuttgart, Germany

Methods exhibiting linear scaling with respect to the size of the system, the so-called O(N) methods, are an essential tool for the calculation of the electronic structure of large systems containing many atoms. They are based on algorithms that take advantage of the decay properties of the density matrix. In this article the physical decay properties of the density matrix will first be studied for both metals and insulators. Several strategies for constructing O(N) algorithms will then be presented and critically examined. Some issues that are relevant only for self-consistent O(N) methods, such as the calculation of the Hartree potential and mixing issues, will also be discussed. Finally some typical applications of O(N) methods are briefly described. [S0034-6861(99)00104-X]

CONTENTS

I. INTRODUCTION

The exact quantum-mechanical equations for many-electron systems are highly intricate. Any attempt to solve these equations analytically for real systems is doomed to fail. Numerical methods such as configuration-interaction-based methods (McWeeny, 1989; Fulde, 1995) or quantum Monte Carlo methods (Hammond, 1994; Nightingale, 1998) can, in principle, solve these many-electron equations but because of the extremely high numerical effort required, their applicability is rather limited in practice.

The bulk of all practical applications is therefore done within various independent-electron approximations such as the Hartree-Fock method (Szabo and Ostlund,

1982), density-functional methods (Parr and Yang, 1989), or tight-binding methods (Goringe, Bouler, and Hernandez, 1997; Majewski and Vogl, 1989). A comparison of the strength of different methods together with a selection of some interesting applications is given by Wimmer (1996). Even these approximate quantum-mechanical equations are still fairly complicated and in general not solvable by analytical methods. Finding efficient algorithms for solving the many-electron problem numerically within any of these approximations is imperative for the applicability of quantum mechanics to physics as well as to chemistry and materials science. Due to efforts in the past, satisfactory algorithms are now available and computational electronic structure methods are making very important contributions to our understanding of matter at the microscopic level. The 1998 Nobel prize awarded to W. Kohn and J. Pople is indicative of the importance of this approach.

Due to the constant increase in computer power and due to algorithmic improvements, the importance of computational methods continues to grow. Whereas computational methods nowadays mainly supplement experimentally obtained information, they are expected increasingly to supersede this information.

This article will concentrate on recently developed methods that allow us to calculate the total energy within various independent-electron methods for large systems. Practically all physical observables can be obtained from the total energy, for instance, in the form of derivatives with respect to certain external parameters. The reason why large systems containing many atoms are accessible with these algorithms is their linear scaling with respect to the number of atoms. In principle, linear scaling should also be obtainable for true many-electron methods. For a widely used approximate many-electron method, such an algorithm has indeed recently been reported (Ayala and Scuseria, 1998).

Traditional electronic structure algorithms calculate eigenstates associated with discrete energy levels. The reason for this is probably historical, since the prediction of these experimentally observed levels was the first big success of quantum mechanics. The disadvantage of this approach is that it leads to a diagonalization problem that has a cubic scaling in the computational effort. Direct diagonalization (Press *et al.*, 1986), which was the standard approach in the early days of the computa-

---

*Electronic address: goedeck@prr.mpi-stuttgart.mpg.de

tional electronic structure era, has a cubic scaling with respect to the size of the Hamiltonian matrix, i.e., with respect to the number of basis functions M_b. Iterative diagonalization schemes (Saad, 1996), preconditioned conjugate-gradient minimizations (Stich *et al.*, 1989; Teter *et al.*, 1989; Payne *et al.*, 1992), and the Car-Parrinello method (Car and Parrinello, 1985) for molecular dynamics simulations were a big algorithmic advance because of their improved scaling behavior. Their scaling was no longer proportional to the cube of the number of basis functions, but grew only like $M_b \log(M_b)$ if plane waves were used as a basis set. Nevertheless, these methods still have a cubic scaling with respect to the number of atoms N_{at}, which comes from the orthogonality requirement of the wave functions. The reason why this orthogonalization step scales cubically can easily be seen. As the system grows, each wave function extends over a larger volume and therefore has to be represented by a larger basis set, resulting in a longer vector. At the same time there are more such wave functions and each wave function has to be orthogonalized to all the others. Thus there are three factors that grow linearly, resulting in the postulated cubic behavior. The computer time T_{CPU} required for the calculation is thus given by

$$T_{CPU} = c_3 N_{at}^3, \qquad (1)$$

where c_3 is a prefactor. It has to be pointed out that Eq. (1) gives only the asymptotic scaling behavior. Within density-functional and Hartree-Fock calculations there are other terms with a lower scaling, which dominate for system sizes of less than a few hundred atoms due to their large prefactor. For plane-wave-type calculations, the fast Fourier transformations necessary for the application of the potential to the wave functions consume most of the computational time for small systems. Whereas for calculations using Gaussian-type orbitals (Hehre *et al.*, 1996), the calculation of the Hartree-Fock potential is the most time consuming. This cubic scaling is a major bottleneck nowadays, since in many problems of practical interest one has to do electronic structure calculations for systems containing many (a few hundred or more) atoms. Plainly, cubic scaling means that, if one doubles the number of atoms in the system, the required computer time will increase by a factor of 8. By enlarging the system one therefore rapidly reaches the limits of the most powerful computers.

So-called O(N) or low-complexity algorithms are therefore a logical next step of algorithmic progress, since they exhibit linear scaling with respect to the number of atoms:

$$T_{CPU} = c_1 N_{at}. \qquad (2)$$

Thus these methods offer the potential for calculating very large systems. The prefactors c_1 and c_3 depend on the approximation used for the many-electron problem. For a density-functional calculation with a large basis set, the prefactors are of course much larger than for a tight-binding calculation, where the number of degrees of freedom per atom is much smaller. The prefactor c_1

depends also on what O(N) method is used, but in general, the prefactor c_1 is always larger than c_3, assuming that the same independent-electron approximation is used in both the traditional and the O(N) version. There is, therefore, a so-called crossover point. For system sizes smaller than the crossover point the traditional cubic scaling algorithms are faster, whereas for larger systems the O(N) methods win. Tight-binding calculations are an ideal test environment for O(N) algorithms. Because of their rather small memory and CPU requirements one can easily treat systems comprising a very large number of atoms and venture into regions beyond the crossover point. Contrary to what one might naively think, the importance of O(N) algorithms will also increase as computers get faster. Whereas at present it is difficult to access the region beyond the crossover point, situated at some 100 atoms, using the density-functional framework, this will be easy with faster computers, and O(N) algorithms will be the algorithms of choice.

Even though O(N) algorithms contain many aspects of mathematics and computer science they have, nevertheless, deep roots in physics. Linear scaling is not obtainable by purely mathematical tricks, but it is based on an understanding of the concept of locality in quantum mechanics. Conversely, the need for constructing O(N) algorithms also served as an incentive for investigating locality questions more deeply and has thus led to a better understanding of this very fundamental concept. An algorithmic description of electronic structure in local terms can give a justification of the well-established concepts of bonds and lone electron pairs in empirical chemistry.

Since O(N) algorithms are based on a certain subdivision of a big system into smaller subsystems, techniques developed in this context might also be helpful in reaching another important goal, namely combining electronic structure methods of different accuracy, such as empirical tight binding and density-functional theory, in a single system.

II. LOCALITY IN QUANTUM MECHANICS

Locality in quantum mechanics means that the properties of a certain observation region comprising one or a few atoms are only weakly influenced by factors that are spatially far away from this observation region. This fundamental characteristic of insulators is well established within independent-electron theories (Heine, 1980) and it can even be carried over into the many-electron framework (Kohn, 1964).

Traditional chemistry is based on local concepts. Covalently bonded materials are described in terms of bonds and lone electron pairs. It is standard textbook knowledge that the properties of a bond are mainly determined by its immediate neighborhood. The decisive factors are what type of atoms and how many of them (the coordination number) are surrounding it. Second-nearest neighbors and other more distant atoms have a very small influence. As an example let us look at the total energy of a hydrocarbon chain molecule $C_n H_{2n+2}$.
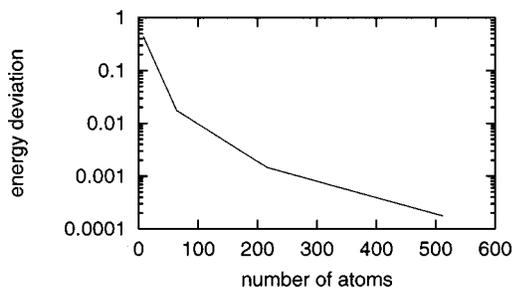
FIG. 1. The deviation of the total energy per silicon atom from its asymptotic bulk value as a function of the size of the periodic volume in which it is embedded. The calculation was done with a tight-binding scheme using exact diagonalization.

In this case each CH_2 subunit is, from an energetical point of view, practically an independent unit. As one adds one CH_2 subunit, the energy increases by an amount that is nearly independent of the chain length. Already the insertion of a CH_2 subunit into the smallest chain C_2H_6 gives an energy gain that agrees within 10^{-4} a.u. with the asymptotic value of the insertion energy for very long chains. This means that the electrons belonging to this inserted subunit no longer "see" the end of the chain for very short chain lengths. This example is a drastic illustration of a principle sometimes termed "nearsightedness" (Kohn, 1996). In other insulating materials the influence of the neighboring atoms decays more slowly. An example is shown in Fig. 1, where the total energy per silicon atom is plotted as a function of the size of its crystalline environment.

Even in metallic systems, where the elementary bond concept is no longer valid, locality still exists. This is supported by the well-known fact that the total charge density in a metal is given with reasonable accuracy by the superposition of the atomic charge densities. Since atomic charge densities decay rapidly, this implies that the charge density at the midpoint between two neighboring atoms is mainly determined by the two closest atoms and influenced very little by other more distant atoms. Another related example is given by V. Heine (1980), who points out that the magnetic moment of an iron atom embedded in an iron-aluminum alloy differs by less than 5% from the value for pure iron if the atom is locally surrounded by only eight iron atoms.

This locality is not at all reflected in standard electronic structure calculations, which are based on eigenorbitals extending over the whole system, requiring unnecessary computational effort and making the interpretation of the results more difficult. The simplistic bond concepts of empirical chemistry are certainly not adequate for electronic structure calculations aiming at high accuracy. Nevertheless, one might hope to incorporate some more general locality concepts into electronic structure calculations to make them both more intuitive and more efficient. In the following we shall therefore carefully examine the range of interactions in quantum-mechanical systems.

Self-consistent electronic structure methods require essentially two steps: the calculation of the potential from the electronic charge distribution and the determination of the wave function for a given potential. In non-self-consistent calculations such as tight-binding calculations, the first step is not needed.

The calculation of the potential consists usually of two parts: the exchange-correlation potential and the Coulomb potential. The exchange-correlation potential is a purely local expression in density-functional theory and can therefore be calculated with linear scaling. In the Hartree-Fock scheme one might first think that the exchange part is nonlocal, but a closer examination reveals (Sec. VIII.A) that it is local even in this case. The Coulomb potential, on the other hand, is very long range and needs proper treatment. A naive evaluation of the potential U arising from a charge distribution ρ by subdividing space into subvolumes ΔV and summing over these subvolumes,

$$U(\mathbf{r}_i) = \sum_j \frac{\rho(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} \Delta V,$$

would result in a quadratic scaling, since both indices i and j have to run over all grid points in the system. The Coulomb problem actually arises not only in the context of electronic structure calculations but also in classical calculations of Coulombic and gravitational systems such as galaxies of stars. Much effort has therefore been invested in this computational problem and several algorithms are known that solve the problem with linear scaling. These methods will be described in Sec. VIII.A.

The more interesting and more difficult part is to assess the role of locality for a given external potential. The appropriate quantity for studying this property is the density matrix. The one-particle density matrix F completely specifies our quantum-mechanical system within the independent-electron approximation, and all quantities of interest can easily be calculated from it. The central quantities in any electronic structure calculation, the kinetic energy E_{kin}, the potential energy E_{pot}, and the electronic charge density ρ, are given by

$$E_{kin} = -\frac{1}{2} \int \nabla_{\mathbf{r}}^2 F(\mathbf{r},\mathbf{r}') \bigg|_{\mathbf{r}=\mathbf{r}'} d\mathbf{r}', \tag{3}$$

$$E_{pot} = \int F(\mathbf{r}',\mathbf{r}') U(\mathbf{r}') d\mathbf{r}', \tag{4}$$

$$\rho(\mathbf{r}) = F(\mathbf{r},\mathbf{r}), \tag{5}$$

where $U(\mathbf{r}')$ is the potential. A related quantity, which will frequently be used throughout the article, is the band-structure energy E_{BS}, defined as

$$E_{BS} = E_{kin} + E_{pot}, \tag{6}$$

and the grand potential Ω

$$\Omega = E_{BS} - \mu N_{el}, \tag{7}$$

where μ is the chemical potential and N_{el} the number of electrons. Ω is by construction invariant under a constant potential offset. If one applies the shift $[U(\mathbf{r}) \rightarrow U(\mathbf{r}) + \text{const}]$, the potential energy will increase by

N_{el} const. In order to conserve the total number of electrons, μ also has to be shifted ($\mu \rightarrow \mu +$ const) and thus Ω remains constant.

Discretizing the Hamiltonian H, which is the sum of the kinetic and potential energies, as well as F with respect to a finite orthogonal basis $\phi_i(\mathbf{r})$, $i=1,\ldots,M_b$, one obtains

$$H_{i,j}= \int \phi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla_{\mathbf{r}}^2 + U(\mathbf{r})\right)\phi_j(\mathbf{r})d\mathbf{r}, \qquad (8)$$

$$F_{i,j}= \int \int \phi_i^*(\mathbf{r})F(\mathbf{r},\mathbf{r}')\phi_j(\mathbf{r}')d\mathbf{r}\,d\mathbf{r}', \qquad (9)$$

and the expressions for the central quantities become

$$E_{BS}= \mathrm{Tr}[FH], \qquad (10)$$

$$\Omega = \mathrm{Tr}[F(H-\mu I)], \qquad (11)$$

$$\rho(\mathbf{r})= \sum_{i,j} F_{i,j}\,\phi_i(\mathbf{r})\phi_j(\mathbf{r}), \qquad (12)$$

where Tr denotes the trace. It follows from Eq. (12) that the total number of electrons N_{el} in the system is given by

$$N_{el}= \mathrm{Tr}[F]. \qquad (13)$$

Evaluating the traces using the eigenfunctions Ψ_n of the Hamiltonian, one obtains immediately the well-known expressions for N_{el}, E_{BS}, Ω, and ρ within the context of conventional calculations which are based on diagonalization. Denoting the eigenvalues associated with the eigenfunctions Ψ_n by ϵ_n one obtains

$$N_{el}= \sum_n f(\epsilon_n), \qquad (14)$$

$$E_{BS}= \sum_n f(\epsilon_n)\epsilon_n, \qquad (15)$$

$$\Omega = \sum_n [f(\epsilon_n)-\mu]\epsilon_n = \sum_n f(\epsilon_n)\epsilon_n - \mu N_{el}, \qquad (16)$$

$$\rho(\mathbf{r})= \sum_n f(\epsilon_n)\,\Psi_n^*(\mathbf{r})\Psi_n(\mathbf{r}). \qquad (17)$$

The function f is the Fermi distribution,

$$f(\epsilon)= \frac{1}{1+\exp\left(\dfrac{\epsilon-\mu}{k_B T}\right)}, \qquad (18)$$

where k_B is Boltzmann's constant and T is the temperature. When we talk about temperature in this article, we always mean the electronic temperature, since we are not considering the motion of the ionic degrees of freedom that might be associated with a different ionic temperature. In Eqs. (14)–(17), as well as in the remainder of the article, we shall use the convention that all the subscripts indexing eigenvalues and eigenfunctions are combined orbital and spin indices, i.e., that we can put at most one electron in each orbital. This will eliminate bothersome factors of 2. The usual case of an unpolar-

ized system can easily be obtained by halving all sums over these indices and multiplying by 2.

In terms of the Hamiltonian H the density matrix is defined as the following matrix functional:

$$F=f(H). \qquad (19)$$

Since F is a matrix function of H, it has the same eigenfunctions Ψ_n as H:

$$H\Psi_n=\epsilon_n\Psi_n, \qquad (20)$$

$$F\Psi_n=f(\epsilon_n)\Psi_n. \qquad (21)$$

The density matrix can consequently be written as

$$F(\mathbf{r},\mathbf{r}')= \sum_n f(\epsilon_n)\Psi_n^*(\mathbf{r})\Psi_n(\mathbf{r}'), \qquad (22)$$

where n runs over all the eigenstates of the Hamiltonian. From the functional form of the Fermi distribution it follows that the eigenvalues $f(\epsilon_n)$ are always in the interval [0:1]. At zero temperature the density matrix of an insulating system containing N_{el} electrons will have N_{el} eigenvalues of value one, all others being zero. Thus the density matrix does not have full rank, but only rank N_{el}. Hence we can write it as

$$F(\mathbf{r},\mathbf{r}')= \sum_{n=occ} \Psi_n^*(\mathbf{r})\Psi_n(\mathbf{r}'), \qquad (23)$$

where n runs now only over the N_{el} occupied states. It is easy to see that $F(\mathbf{r},\mathbf{r}')$ is a projection operator, in this case

$$\int F(\mathbf{r},\mathbf{r}'')F(\mathbf{r}'',\mathbf{r}')d\mathbf{r}'' = F(\mathbf{r},\mathbf{r}'). \qquad (24)$$

A new set of N_{el} eigenfunctions $\Psi_n^{new}(\mathbf{r})$ can be obtained by any unitary transformation of all the N_{el} degenerate eigenfunctions $\Psi_n(\mathbf{r})$ associated with eigenvalues one,

$$\Psi_n^{new}(\mathbf{r})= \sum_{m=occ} U_{n,m}\Psi_m(\mathbf{r}), \qquad (25)$$

where U is a unitary N_{el}-by-N_{el} matrix. For crystalline periodic solids such a transformation can be used to generate the localized Wannier functions (Blount, 1962) from the extended eigenfunctions Ψ_n. We shall refer to any set of orthogonal exponentially localized orbitals that can be used to represent the density matrix according to Eq. (23) as Wannier functions. A method for constructing an optimally localized set of Wannier functions by the minimization of the total spread $\sum_n \langle r^2\rangle_n - \langle \mathbf{r}\rangle_n^2$ in a crystalline periodic solid has recently been shown by Marzari and Vanderbilt (1997). It has been well known in the chemistry community (Chalvet et al., 1976) that sets of maximally localized orbitals give excellent insight into the bonding properties of systems. In addition to the spread criterion used by Marzari and Vanderbilt (1997) there are still other criteria in common use in the chemistry community. They are all in a certain sense arbitrary, but usually lead to the same interpretation of
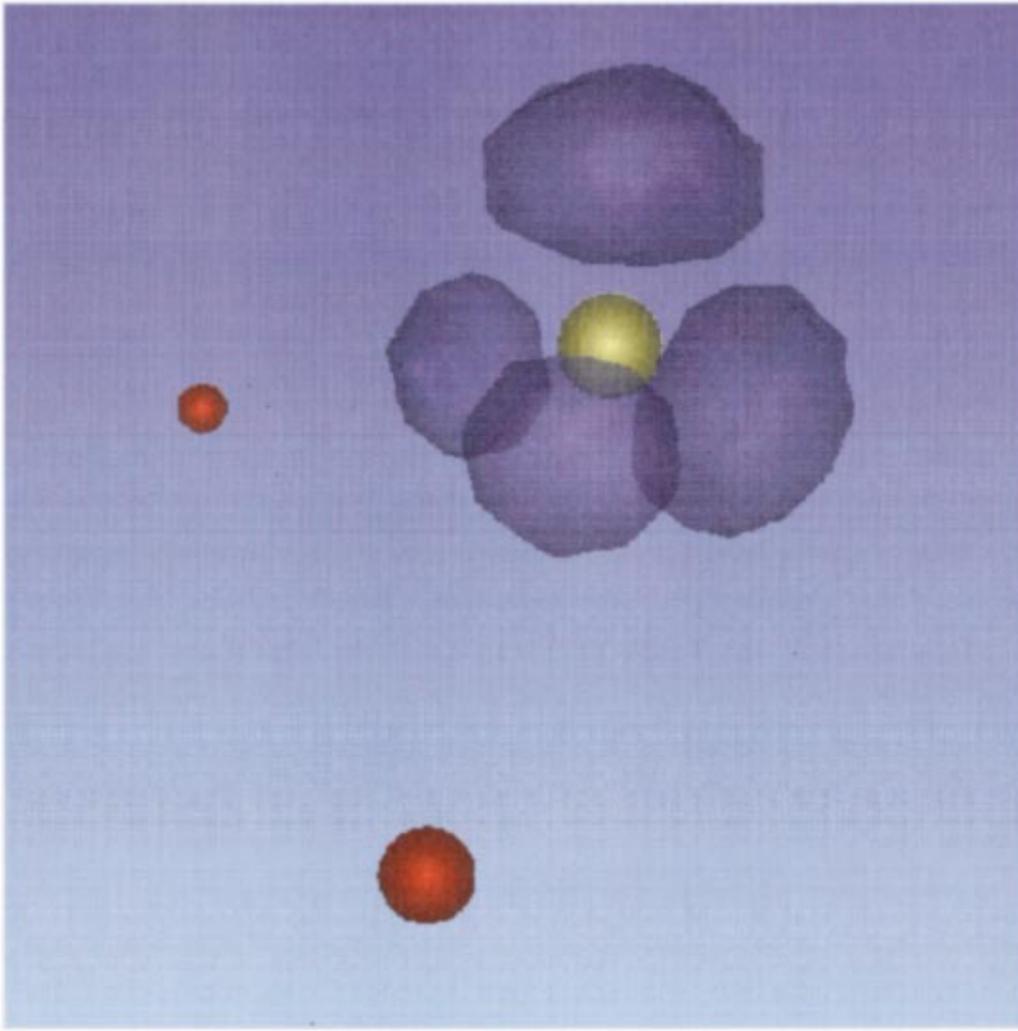
FIG. 2. A set of four Wannier orbitals for the water molecule. The oxygen nucleus in the center is shown as a yellow ball and the two hydrogen nuclei as red balls. The Wannier functions are indicated by dark blue clouds. One sees two Wannier functions along the lines connecting the central oxygen with the two hydrogen atoms, representing bonds, as well as two representing lone electron pairs. The centers of the four Wannier functions form a nearly tetragonal structure.

the bonding properties. Figure 2 shows the four Wannier functions for the water molecule.

The density matrix $F(\mathbf{r},\mathbf{r}')$ is a diagonally dominant operator, whose off-diagonal elements decay with increasing distance from the diagonal. The exact decay behavior depends on the material. We shall derive the decay properties within the theoretical framework of the description of periodic crystalline solids. For a periodic solid the density matrix is given by

$$F(\mathbf{r},\mathbf{r}') = \sum_n \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k}\, f(\epsilon_n(\mathbf{k})) \Psi_{n,\mathbf{k}}^*(\mathbf{r}) \Psi_{n,\mathbf{k}}(\mathbf{r}')$$

$$= \sum_n \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k}\, f(\epsilon_n(\mathbf{k})) u_{n,\mathbf{k}}^*(\mathbf{r})$$

$$\times u_{n,\mathbf{k}}(\mathbf{r}') e^{i\mathbf{k}(\mathbf{r}'-\mathbf{r})}, \qquad (26)$$

where $\Psi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}(\mathbf{r})}$ are the Bloch functions associated with the wave vector $\mathbf{k}$ and band index n. The integral is taken over the Brillouin zone (BZ) and V is the volume of the real-space primitive cell.

The Wannier functions W_n of the nth band in an insulating crystal are defined in the usual way:

$$W_n(\mathbf{r}-\mathbf{R}) = \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k}\, e^{-i\mathbf{kR}} \Psi_{n,\mathbf{k}}(\mathbf{r}). \qquad (27)$$

The Wannier functions are not uniquely defined. One can construct a different set of Bloch functions by multiplying them with a phase factor, $\Psi_{n,\mathbf{k}}(\mathbf{r}) \leftarrow e^{i\omega(\mathbf{k})} \Psi_{n,\mathbf{k}}(\mathbf{r})$, where $\omega(\mathbf{k})$ is an arbitrary function. This will obviously modify the Wannier functions. Further ambiguities arise in the case of degenerate bands (Blount, 1962). Because of these ambiguities in the construction of the Wannier functions it is advantageous to work with the density matrix where any phase factors cancel [Eq. (26)] and where degeneracies do not cause any problems since one sums over all the occupied bands.

We shall first discuss the decay properties of the density matrix in metallic systems. In this discussion we shall assume that metals behave essentially like jellium and

that exact results for jellium can be carried over to real metals.

The decay properties of the density matrix of a metallic system at zero temperature are well known (March *et al.*, 1967). Because the integral in Eq. (26) contains a discontinuity in the metallic case, the density matrix decays only algebraically with respect to the distance between $\mathbf{r}$ and $\mathbf{r}'$. The decay is given by

$$F(\mathbf{r},\mathbf{r}') \propto k_F \frac{\cos(k_F|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|^2}, \qquad (28)$$

where the Fermi wave vector k_F is related to the valence electron density by $N_{el}/V = k_F^3/3\pi^2$ in a non-spin-polarized system.

Introducing a finite electronic temperature T in a metal leads to a drastic change in this decay behavior. Instead of an algebraic decay one has a much faster exponential decay. As shown independently by Goedecker (1998a) and Ismail-Beigi and Arias (1999), the decay at low temperatures is then given by

$$F(\mathbf{r},\mathbf{r}') \propto k_F \frac{\cos(k_F|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|^2} \exp\left(-c \frac{k_B T}{k_F}|\mathbf{r}-\mathbf{r}'|\right), \qquad (29)$$

where c is a constant on the order of 1. We thus find oscillatory behavior with an exponentially damped amplitude. The decay rate depends linearly on temperature, and the oscillatory part is described by the wave vector k_F. The related correlation function at finite temperature exhibits the same temperature dependence of the decay rate with respect to temperature (Landau and Lifshitz, 1980). In an insulator finite temperature plays no role as long as the thermal energy $k_B T$ is much smaller than the gap, a condition that is usually fulfilled.

Let us next discuss the important case of an insulator with a band gap ϵ_{gap} at zero temperature. We shall first present some numerical results and then put forward some arguments to explain the qualitative features of the density matrix. Finally, we shall discuss in a more quantitative way the factors that determine the exact decay rate.

Numerical calculations of the density matrix or the related Wannier functions show an oscillatory behavior with a decaying amplitude. There is exactly one node per primitive cell, and logarithmic plots of the amplitude clearly reveal an exponential decay. For alkanes the decay of the density matrix calculated by the Hartree-Fock method has been studied and plotted on a logarithmic scale by Maslen *et al.* (1998). Interestingly, the decay depends also on the basis set used. Small low-quality basis sets lead to a larger band gap and consequently to a faster decay of the density matrix. For silicon treated by density-functional theory, logarithmic plots revealing the exponential decay of the Wannier functions have also been done for both grid-based basis sets (Goedecker, unpublished) and atomic basis sets (Stephan, 1998). Within the tight-binding method, the decay of the density matrix has also been studied numerically for crystalline and liquid carbon systems by Goedecker (1995) and for fullerenes by Itoh *et al.* (1996).

Let us now make plausible the exponential decay of the density matrix. The demonstration is based on the fact that one can express the Fourier components $\epsilon_n(\mathbf{R})$ of the band energy $\epsilon_n(\mathbf{k})$ through the Wannier functions $W_n(\mathbf{r})$,

$$\epsilon_n(\mathbf{R}) = \frac{V}{(2\pi)^3} \int_{BZ} \epsilon_n(\mathbf{k}) e^{-i\mathbf{k}\mathbf{R}} d\mathbf{k}$$

$$= \frac{(2\pi)^3}{V} \int_{space} W_n^*(\mathbf{r}') H W_n(\mathbf{r}'-\mathbf{R}) d\mathbf{r}', \qquad (30)$$

where $\mathbf{R}$ is a Bravais lattice vector. Now it is known that the band energy $\epsilon_n(\mathbf{k})$ is an analytic function (Blount, 1962). This is actually not surprising. The first and second derivatives of the band structure have physical meaning since they are related to the electron velocity and effective mass, so it is to be expected that higher derivatives exist as well. Since the Fourier transform of an analytic function decays faster than algebraically (see Appendix), there exists a decay constant γ and a normalization constant C such that

$$Ce^{-\gamma R} \geq \epsilon|_n(\mathbf{R})|$$

$$= \left| \frac{(2\pi)^3}{V} \int_{space} W_n^*(\mathbf{r}') H W_n(\mathbf{r}'-\mathbf{R}) d\mathbf{r}' \right|. \qquad (31)$$

It is reasonable to expect that $HW_n(\mathbf{r})$ will behave similarly to $W_n(\mathbf{r})$. In particular, we expect $W_n(\mathbf{r})$ to be small whenever $HW_n(\mathbf{r})$ is small, so we shall just drop H in Eq. (31). In addition, we shall define this modified integral not only for lattice vectors $\mathbf{R}$ but for arbitrary vectors $\mathbf{r}$ to obtain

$$Ce^{-\gamma r} \geq \left| \frac{(2\pi)^3}{V} \int_{space} W_n^*(\mathbf{r}') W_n(\mathbf{r}'-\mathbf{r}) d\mathbf{r}' \right|. \qquad (32)$$

If Eq. (32) holds, then one can use the mean value theorem to show that

$$Ce^{-\gamma r} \geq \left| \frac{(2\pi)^3}{V} \int_{space} W_n^*(\mathbf{r}') W_n(\mathbf{r}'-\mathbf{r}) d\mathbf{r}' \right|$$

$$= \left| \frac{(2\pi)^3}{V} \sum_{\mathbf{R}'} \int_{cell} W_n^*(\mathbf{r}'-\mathbf{R}') W_n(\mathbf{r}'-\mathbf{R}'-\mathbf{r}) d\mathbf{r}' \right|$$

$$= \left| \sum_{\mathbf{R}'} W_n^*(\mathbf{s}(\mathbf{r})-\mathbf{R}') W_n(\mathbf{s}(\mathbf{r})-\mathbf{R}'-\mathbf{r}) d\mathbf{r}' \right|$$

$$= |F(\mathbf{s}(\mathbf{r}),\mathbf{s}(\mathbf{r})-\mathbf{r})|, \qquad (33)$$

where the mean value $\mathbf{s}(\mathbf{r})$ is a vector within the primitive cell. Assuming that the density matrix has the same order of magnitude within each cell, one can neglect the dependence of $\mathbf{s}$ on $\mathbf{r}$ to obtain the final result

$$Ce^{-\gamma r} \geq |F(\mathbf{s},\mathbf{s}-\mathbf{r})|. \qquad (34)$$

The numerically observed nodal structure of the density matrix can be motivated in a very similar way. Because of the orthogonality of the Wannier functions we have

$$0 = \int_{space} W_n^*(\mathbf{r}') W_n(\mathbf{r}'-\mathbf{R}) d\mathbf{r}' \qquad (35)$$
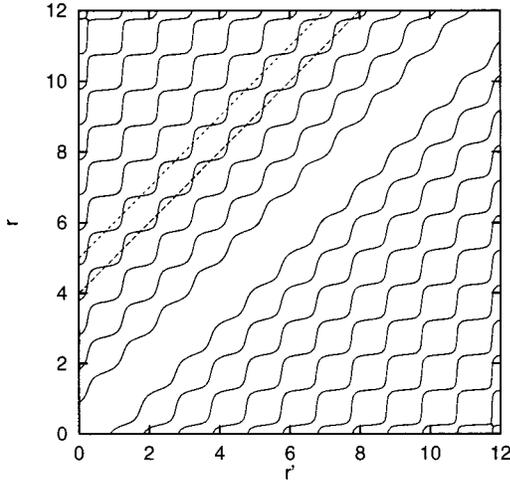
FIG. 3. The nodal structure of the density matrix $F(\mathbf{r},\mathbf{r}')$ for a one-dimensional model insulator with a bandwidth of 4 a.u. and a band gap of 2 a.u. The length of the primitive cell is 1. The nodes predicted by Eq. (36) are at the intersections with diagonal lines, two of which are shown by the dashed lines.

for any nonzero lattice vector $\mathbf{R}$. Doing the same sequence of transformation as in Eq. (33), one obtains

$$0 = F(\mathbf{s}(\mathbf{R}),\mathbf{s}(\mathbf{R})-\mathbf{R}). \qquad (36)$$

Hence there has to be one node in each cell. The numerically calculated nodal structure for a one-dimensional model insulator is shown in Fig. 3.

The next step is to examine in a more quantitative way which factors determine the rate of this exponential decay for an insulator with a band gap ϵ_{gap} at zero temperature.

Cloizeaux (1964) proved the exponential decay behavior of the zero-temperature density matrix, which is a projection operator. Considering the extension of the band energy $\epsilon_n(\mathbf{k})$ into the complex $\mathbf{k}$ plane, he found that the minimal distance of the branch points of $\epsilon_n(\mathbf{k})$ from the real axis determines the decay behavior. For the Wannier functions, which are closely related to the density matrix by Eq. (23), Kohn (1959) proved the same decay behavior in the case of a one-dimensional model crystal. In a later publication Kohn (1993) claimed that this distance to the real $\mathbf{k}$ axis should be related to the square root of the gap. Even though he did not present a derivation of this result, it was widely accepted as generally valid. Ismail-Beigi and Arias (1999) have shown, however, that Kohn's claim is not generally valid. They demonstrated that in the tight-binding limit the square-root behavior can be found under certain circumstances, but that different behaviors can be found as well. In the weak-binding limit, where the band structure can be obtained by perturbation theory from the band structure of the free-electron gas, they showed that the dependence is actually linear:

$$F(\mathbf{r},\mathbf{r}') \propto \exp(-\gamma|\mathbf{r}-\mathbf{r}'|), \quad \text{where} \quad \gamma = c\epsilon_{gap}a. \qquad (37)$$

Here, the lattice constant is denoted by a, and c is an unknown constant of the order of 1.

The dependence of the decay rate on the size of the band gap is a rather surprising relation. After all, it follows from Eq. (26) that only the properties of the occupied bands enter into the calculation of the density matrix, whereas the size of the gap is not directly related to the occupied states. In the following we shall give an intuitive explanation of the factors determining the decay rate. This explanation will again be based on Eq. (30) relating the band structure to the decay properties of the density matrix. As is known from complex analysis, the distance of the singularities from the real axis is comparable to the length over which one has very strong variations along the real axis of a complex function. Now, the long-range decay properties of a Fourier transform are exactly determined by the length Δk of such a region of strongest variation (see Appendix). One thus regains Cloizeaux's result that the decay rate is proportional to the distance of singularities from the real axis. Let us now explain the behavior found in the weak-binding limit by Ismail-Beigi and Arias (1999). In the weak-binding limit the effective mass establishes the connection between the gap and the important features of the occupied bands. The effective mass for the nth band at the point $\mathbf{k}_0$ is defined as (Kittel, 1963)

$$\frac{1}{m} = 1 + \frac{2}{3}\sum_{m \neq n} \frac{|\int \Psi^*_{n,\mathbf{k}_0}(\mathbf{r})\nabla\Psi_{m,\mathbf{k}_0}(\mathbf{r})d\mathbf{r}|^2}{\epsilon_n(\mathbf{k}_0)-\epsilon_m(\mathbf{k}_0)}. \qquad (38)$$

Since we are only interested in orders of magnitude, we have here averaged over the diagonal elements of the effective-mass tensor in order to obtain an effective mass that is a scalar quantity. For the weak-binding limit, a gap will open up at the boundaries of the Brillouin zone and this gap will be small. The effective mass is therefore small and proportional to $a^2\epsilon_{gap}$, where we have assumed that the dipole matrix elements $\int \Psi^*_{n,\mathbf{k}_0}(\mathbf{r})\nabla\Psi_{i,\mathbf{k}_0}(\mathbf{r})d\mathbf{r}$ are on the order of $1/a$. The band structure near the boundaries of the Brillouin zone is then given by

$$\frac{1}{2m}(\Delta k)^2 \propto \frac{1}{a^2\epsilon_{gap}}(\Delta k)^2, \qquad (39)$$

where Δk is the distance from the boundary, neglecting directional effects. Since the effective mass is small, the curvature of the band structure is large in this region. Hence this region is just the region with the strongest variation. As is well known (Ashcroft and Mermin, 1976), the perturbation theory arguments leading to Eq. (39) are valid within an energy range of the order of ϵ_{gap}. It then follows from Eq. (39) that the corresponding range of Δk is $\epsilon_{gap}a$, confirming the linear decay of the density matrix with respect to the size of the gap, i.e., $\gamma = c\epsilon_{gap}a$.

Let us next show how a square-root-like behavior $\gamma = c\sqrt{\epsilon_{gap}}$ can arise for real crystals with a large gap. In this case the effective mass is of the order of 1 at all stationary points $\mathbf{k}_0$ in the Brillouin zone. Assuming that it is then of the order of 1 over the whole Brillouin zone, the region of largest variation is just the Brillouin zone itself. The decay constant is therefore simply related to the lattice constant a:
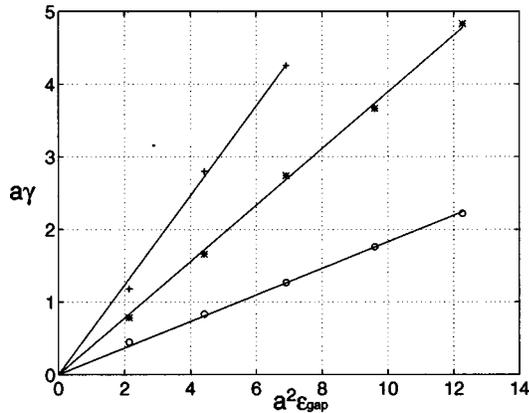
FIG. 4. The dependence of the decay constant γ on the gap. Plotted are the two dimensionless quantities $a\gamma$ vs $a^2\epsilon_{gap}$. The variation of the gap was obtained for a three-dimensional cubic model crystal by varying the strength of the potential. Circles refer to the [100], stars to the [110], and pluses to the [111] direction. This figure is reproduced with kind permission of the authors, from Ismail-Beigi and Arias (1999).

$$\gamma = c\,\frac{1}{a}. \tag{40}$$

In order to get the square-root dependence of the decay constant γ, one has to assume that

$$\epsilon_{gap} = C_{gap}\,\frac{1}{a^2}, \tag{41}$$

where C_{gap} is a constant that is not dependent (or only weakly dependent) on the material. Such a behavior has indeed been observed for certain classes of materials, where the tight-binding limit is the most appropriate one, such as ionic crystals (Harrison, 1980), but with a non-negligible variation of C_{gap} across different materials. A square-root behavior of γ can therefore be expected if one varies the lattice constant for a certain material, but the decay constants for different materials that happen to have the same gap are not necessarily comparable.

In practice the distinction between the tight-binding and weak-binding cases may not always be clear. Unless the region of strongest variation is really a very small fraction of the whole Brillouin zone, all the prefactors that were neglected in these considerations might be important enough to blur the differences. The importance of these prefactors can also be seen from the fairly strong directional dependence of the decay rate. Ismail-Beigi and Arias (1999) found such a strong directional dependence in numerical tests, confirming the linear dependence of the decay constant on the size of the gap (Fig. 4). Stephan and Drabold (1998) found the same behavior during tight-binding studies of carbon. So the statement in an early paper by Kohn (1964), that the decay length of the Wannier functions is of the order of the interatomic spacing, is for practical purposes probably in many cases the best available characterization of localization.
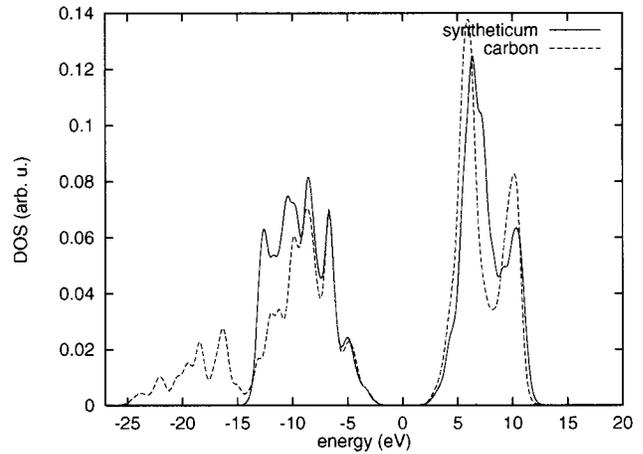
FIG. 5. Comparison of the density of states of carbon and syntheticum. As one can see, both have roughly the same gap. The lower part of the valence band, however, is drastically different. The valence band of syntheticum is roughly half as wide as that of carbon.

As a numerical illustration of this surprising result, that only a small part of the Brillouin zone in which one has the strongest variation determines the decay behavior of the Wannier functions, we compared the decay behavior of carbon in the diamond structure with "syntheticum" in the same structure. The artificial element "syntheticum" was computer generated within the tight-binding context in such a way that the top part of its conduction band as well as the gap is nearly identical to real carbon, whereas the lower part of the valence band is drastically different, as shown in Fig. 5. More precisely, carbon was characterized by the parameters of Goodwin (1991) and, to obtain syntheticum, ϵ_s was modified from -5.16331 to -1.16331, and $V_{ss\sigma}$ was modified from -4.43338 to -2.43338.

Figure 6 shows the decay behavior of the density matrix. As one can see, the decay behavior is very similar in
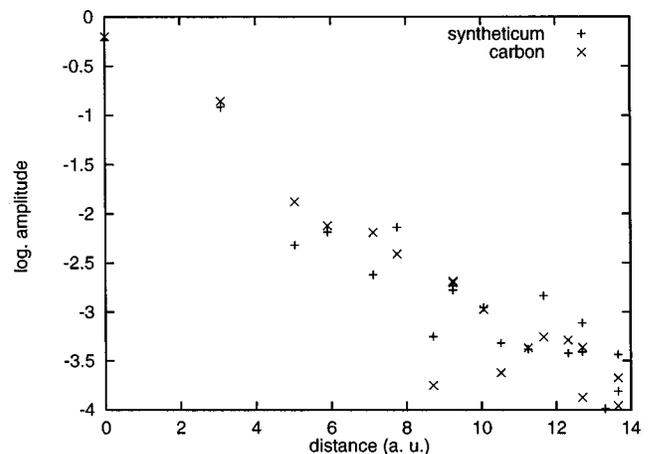
FIG. 6. Comparison of the decay behavior of the density matrices for carbon and syntheticum. They are both very similar. The moderate scattering comes from the fact that the density matrix does not decay equally fast in all directions.

both cases. We note that not only the gaps are similar but also the effective masses, since the density of states at the top of the valence band has the same behavior in both materials.

All the above arguments apply to simple and mainly periodic materials. Advanced electronic structure calculations, however, frequently study materials that are not in this class. The localization properties of such materials have not yet been studied systematically and so there is some uncertainty about which orbitals are localized and to what extent (Kohn, 1995). If the localization properties are unknown, one does better not to impose any localization constraints. In this case some of the discussed O(N) techniques still give a quadratic scaling, which also allows us to gain computational efficiency compared to the traditional cubically scaling algorithms.

III. BASIC STRATEGIES FOR O(N) SCALING

Most O(N) algorithms are built around the density matrix or its representation in terms of Wannier functions and take advantage of its decay properties. To obtain linear scaling, one has to cut off the exponentially decaying quantities when they are small enough. This introduces the concept of a localization region. Only inside this localization region is the quantity calculated; outside it is assumed to vanish. For simplicity the localization region is usually taken to be a sphere, even though the optimal shape might be different (Stephan and Drabold, 1998). In the tight-binding context the boundary of the localization region can either be defined by a geometric distance criterion or in terms of the number of ''hops,'' i.e., the number of steps one has to take along bonds connecting neighboring atoms to reach this boundary (Voter *et al.*, 1996). Different localization regions generally have significant overlaps. The localization regions thus do not form a partition of the computational volume, and one atom, in general, belongs to several localization regions.

In a numerical calculation the density operator $F(r,r')$ is discretized with respect to a basis. The basis set has to be chosen such that the matrix elements $F_{i,j}$ reflect the decay properties of the operator $F(r,r')$. This will obviously only be the case if the basis set consists of localized functions, such as atom-centered Gaussian-type basis functions. Sets of orthonormal basis functions usually facilitate the calculations. Unfortunately all currently used localized basis sets are nonorthogonal. In the context of the orthogonal tight-binding scheme (Goringe, Bowler, and Hernandez, 1997; Majewski and Vogl, 1989) one just assumes the existence of a basis set which is both atom centered and orthogonal. Since only the parametrized Hamiltonian matrix elements enter in the calculation, there is no need ever to explicitly construct such a basis set. In the following sections, we shall follow this practice and assume in all relevant parts that we are dealing with such a localized orthogonal basis set. The non-orthogonal case will be discussed in Sec. VII. Whenever we refer from now on

to a localization region, we actually mean the subset of all basis functions that are contained within this spatial localization region.

Obviously the size of the localization region needed to obtain a certain accuracy depends on the decay properties of the density matrix as well as on the selected accuracy threshold. It also depends on the quantity one wants to study. Generally, the total energy as well as derived quantities such as the geometric equilibrium configurations are surprisingly insensitive to finite localization regions, because these quantities are not strongly influenced by the exponentially small tails that are cut off by the introduction of a localization region. This insensitivity also holds true, though to a much lesser extent, for metals. As we have seen above, the introduction of a finite temperature leads to an exponential decay of the density matrix for a metal, which in turn justifies truncation. In a metal, the difference between the finite- and the zero-temperature total energy ΔE is proportional to the square of the temperature, $\Delta E \propto T^2$ (Ashcroft and Mermin, 1976) and thus rather small. There are, however, quantities that are very sensitive to finite localization regions. In the modern theory of polarization in solids (King-Smith and Vanderbilt, 1993), the polarization can be expressed in terms of the centers of the Wannier functions $\int W(\mathbf{r})\mathbf{r}W(\mathbf{r})d\mathbf{r}$. Using this formula (Fernandez *et al.*, 1997), one finds a strong influence from the tails of the Wannier functions because they are strongly weighted by the factor of $\mathbf{r}$ in the integral. Since the tails are much more influenced by the boundary of the localization region than the central part, this quantity is more sensitive to the size of the localization region.

There are even quantities that are not at all directly accessible by a solution given in terms of density matrices or Wannier functions. The Fermi surface in a metal, which can be calculated via the eigenvalues of the band structure $\epsilon_n(\mathbf{k})$, is such an example.

It is clear that one can gain significant computational efficiency only if the extent of the system is larger than the size of the localization region. The crossover point depends therefore on the decay properties of the density matrix of the system. However, it also depends on the dimensionality of the system. For a linear-chain molecule with a large band gap, it might be enough to have a localization region containing just two neighboring atoms on each side, hence just five atoms, and for systems larger than five atoms one might potentially gain computational efficiency by using an O(N) method. If one had a three-dimensional system with a comparable gap, then a spherical localization region extending out to the second neighbors would contain some 60 atoms and the crossover point would already be much larger. For a system with a small gap, such as silicon, or for metallic systems the crossover point would be even larger.

There are essentially six basic approaches to achieving linear scaling:

- The Fermi operator expansion (FOE) is based on Eq. (19). In this approach one finds a computable functional form of F as a function of H to build up

the density matrix. Two possible representations based on a Chebyshev expansion and a rational expansion will be discussed.

- The Fermi operator projection (FOP) is closely related to the FOE method. The computable form of F is not, however, used to construct the entire density matrix but to find the space spanned by the occupied states, i.e., the space corresponding to the eigenfunctions associated with the unit eigenvalues of the density matrix at zero temperature. These eigenfunctions can be considered as Wannier functions in the generalized sense defined above.

- In the divide-and-conquer (DC) method for the density matrix, the relevant parts of the density matrix are patched together from pieces that were calculated for smaller subsystems.

- In the density-matrix minimization (DMM) approach, one finds the density matrix by a minimization of an energy expression based on the density matrix.

- In the orbital minimization approach (OM), one finds a set of Wannier functions by minimization of an energy expression.

- The optimal basis density-matrix minimization scheme (OBDMM) contains aspects of both the OM and DMM methods. In addition to finding a density matrix with respect to the basis, one also finds an optimal basis by additional minimization steps. The number of basis functions has to be at least equal to the number of electrons in the system, but can be larger as well.

A major difference between these methods is whether they calculate the full density matrix or only its representation in terms of Wannier functions. The latter approach applies only to insulators, while the former is also applicable to systems with fractional occupation numbers [i.e., $f(\epsilon_n)$ is neither 1 nor 0] such as metals or systems at finite electronic temperature.

In the following each of these six approaches will be presented in detail.

A. The Fermi operator expansion

The Fermi operator expansion (FOE) (Goedecker and Colombo, 1994a; Goedecker and Teter, 1995) is the most straightforward approach for the calculation of the density matrix. The basic idea in this approach is to find a representation of the matrix function (19) that can be evaluated on a computer. Several such representations are possible. We shall discuss a Chebyshev and a rational representation.

1. The Chebyshev Fermi operator expansion

One of the most basic operations a computer can do is a matrix-times-vector multiplication. The simplest representation of the density matrix, requiring only this operation, would be a polynomial representation,

$$F \approx p(H) = c_0 I + c_1 H + c_2 H^2 + \cdots + c_{n_{pl}} H^{n_{pl}},$$

where I is the identity matrix. Unfortunately polynomials of high degree become numerically unstable. This instability can, however, be avoided by introducing a Chebyshev polynomial representation, which is a widely used numerical method (Press $et\ al.$, 1986):

$$p(H) = \frac{c_0}{2} I + \sum_{j=1}^{n_{pl}} c_j T_j(H). \tag{42}$$

Since the Chebyshev polynomials are defined only within the interval $[-1:1]$, we shall assume in the following that the eigenvalue spectrum of H falls within this interval. This can always be easily achieved by scaling and shifting of the original Hamiltonian. The Chebyshev matrix polynomials $T_j(H)$ satisfy the recursion relations

$$T_0(H) = I, \tag{43}$$

$$T_1(H) = H, \tag{44}$$

$$T_{j+1}(H) = 2H T_j(H) - T_{j-1}(H). \tag{45}$$

The expansion coefficients of the Chebyshev expansion can easily be determined. The eigenfunction representation [Eq. (21)] of F is

$$\langle \Psi_n | F | \Psi_m \rangle = f(\epsilon_n) \delta_{n,m}. \tag{46}$$

Evaluating the polynomial expansion in the same eigenfunction representation, we obtain

$$\langle \Psi_n | p(H) | \Psi_m \rangle = p(\epsilon_n) \delta_{n,m}, \tag{47}$$

where

$$p(\epsilon) = \frac{c_0}{2} + \sum_{j=1}^{n_{pl}} c_j T_j(\epsilon). \tag{48}$$

Comparing Eqs. (46) and (47), we see that the polynomial $p(\epsilon)$ has to approximate the Fermi distribution in the energy interval $[-1:1]$ where the scaled and shifted Hamiltonian has its eigenvalues. How to find the Chebyshev expansion coefficients for a scalar function is described in standard textbooks on numerical analysis (Press $et\ al.$, 1986). Actually it is not necessary to take the exact Fermi distribution. In practically all situations one is interested in the limit of zero temperature. Hence any function that approaches a step function in the limit of zero temperature can be used. For simulations of insulators, for instance, it is advantageous to take the function $f(\epsilon) = \frac{1}{2}\{1 - \mathrm{erf}[(\epsilon - \mu)/\Delta\epsilon]\}$ (shown in Fig. 7) since it decays faster to 0 respectively 1 away from the chemical potential. We shall use the term Fermi distribution in this broader sense. The energy resolution $\Delta\epsilon$ is chosen to be a certain fraction of the size of the gap (Goedecker and Teter, 1995). For metals, $\Delta\epsilon$ is chosen by considerations of numerical convenience. Large values of $\Delta\epsilon$ will give lower-accuracy results. However, as pointed out before, the convergence of the total energy with respect to $\Delta\epsilon$ is quadratic, and thus highly accurate total energies can be obtained with rather high values of $\Delta\epsilon$ (Goedecker and Teter, 1995). Small values of $\Delta\epsilon$ make the calculation numerically expensive. The detailed scaling behavior of the numerical effort in the limit of van-
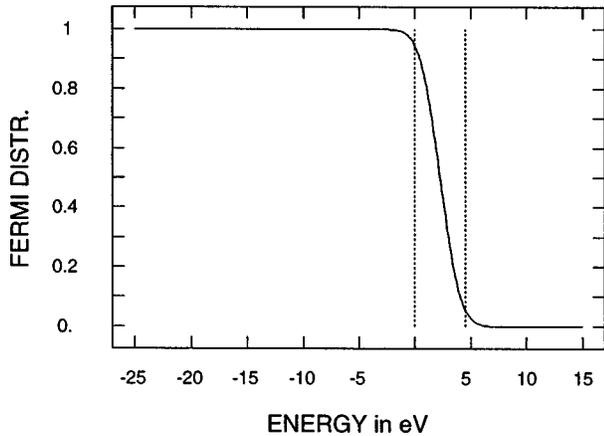
FIG. 7. The Fermi distribution as obtained by a Chebyshev fit of degree 40 in the case of a diamond structure. The band gap is in between the two vertical lines.

ishing gaps is analyzed in Sec. IV, where it is found that the increase in the size of the localization region is the limiting factor in all methods.

Even if one wants to study electronic properties in the limit of zero electronic temperature it is important that one uses a finite-temperature Fermi distribution for the Chebyshev fit. Using the zero-temperature step function introduces so-called Gibbs oscillations in the fit and spoils the Chebyshev fit over the whole interval.

A method for eliminating these Gibbs oscillations in the zero-temperature case is the so-called kernel polynomial method (Voter *et al.*, 1996; Silver *et al.*, 1996), which can be used as a starting point for an alternative derivation of the FOE method. The basic idea is to expand a delta function as a polynomial using damping factors to suppress large oscillations. This representation of an approximate delta function can then be integrated to obtain a smooth representation of the Fermi distribution. Used this way the kernel polynomial method is thus just another way to derive the expansion coefficients for the Chebyshev expansion (Kress *et al.*, 1998). In addition the kernel polynomial approach can also be used to smear out the density of states rather than the zero-temperature Fermi distribution, resulting in a method with practically identical computational requirements but some slightly different properties. One useful property is that the smeared density-of-states energy is an approximate lower bound to the energy, whereas the smeared Fermi energy is an approximate upper bound (Voter *et al.*, 1996).

Coming back to the original motivation for a polynomial representation, let us now show how the density matrix can be constructed using only matrix-times-vector multiplications. Let us denote by t_l^j the lth column of the Chebyshev matrix T_j. Now each column of these Chebyshev matrices satisfies the same recursion relations:

$$|t_l^0\rangle = |e_l\rangle, \tag{49}$$

$$|t_l^1\rangle = H|e_l\rangle,$$

$$|t_l^{j+1}\rangle = 2H|t_l^j\rangle - |t_l^{j-1}\rangle,$$

where e_l is a unit vector that has zeros everywhere except at the lth entry. So Eq. (49) demonstrates that we indeed need only matrix-vector multiplications. Once we have generated the lth columns of all the Chebyshev matrices, we can obtain the lth column f_l of the density matrix just by forming linear combinations:

$$|f_l\rangle = \frac{c_0}{2}|t_l^0\rangle + \sum_{j=1}^{n_{pl}} c_j|t_l^j\rangle. \tag{50}$$

As we have described the method so far it has a quadratic scaling instead of the linear scaling we finally want to achieve. If we have M_b basis functions, the density matrix is a $M_b \times M_b$ matrix which means that we have to calculate M_b full columns. For the calculation of each column, we have to do n_{pl} matrix-times-vector multiplications, each of which costs $M_b n_H$ operations, assuming the matrix H is a sparse matrix with n_H off-diagonal elements per row/column. So the total computational cost is $M_b^2 n_{pl} n_H$. The degree of the polynomial n_{pl} and the width n_H of the Hamiltonian are independent of the size of the system, whereas M_b is proportional to the number of atoms in the system. The overall scaling with respect to the number of atoms is therefore quadratic.

In order to do the correct shifting and scaling of the original Hamiltonian to map its eigenvalue spectrum on the interval $[-1:1]$ we have to know its lowest and highest eigenvalues ϵ_{min} and ϵ_{max}. In addition, we have to know the chemical potential μ. There are auxiliary matrix functions of H that can help us to determine these quantities. These functions of H can be built up in the same way as the density matrix. Since the recursive buildup of the Chebyshev matrices is the most costly part, the additional cost for evaluating other functions is negligible. To determine whether we have a vanishing density of states beyond an energy ϵ_{up}, we can, for instance, construct a Chebyshev fit $p_{up}(\epsilon)$ to a function that is zero (to within a certain tolerance) for energies below ϵ_{up}, but blows up for energies larger than ϵ_{up}. If $\text{Tr}[p_{up}(H)]$ does not vanish we have a nonvanishing density of states beyond ϵ_{up}. A similar procedure can be applied to determine a lower bound for the density of states. The determination of the chemical potential in an insulator can be done along the same lines as well (Bates *et al.*, 1998). Without any significant extra cost one can build up several Fermi distributions with different chemical potentials until one finds the correct chemical potential leading to charge neutrality. In a metallic system the search for the chemical potential can be accelerated since it is possible to predict with high accuracy how the number of electrons changes in response to a change in the chemical potential. From Eq. (13) it follows that

$$\frac{\partial N_{el}}{\partial \mu} = -\text{Tr}[p'(H)], \tag{51}$$

where p' is the derivative of the Chebyshev polynomial p that approximates the Fermi distribution. The Cheby-

shev expansion coefficients of p' can be calculated from the coefficients for p (Press *et al.*, 1986). Using the finite-difference approximation of Eq. (51),

$$\Delta \mu = \frac{\Delta N_{el}}{\text{Tr}[p'(H)]}, \tag{52}$$

it is possible to find the correction $\Delta \mu$ to the chemical potential which will nearly exactly eliminate an excess of ΔN_{el} electrons due to an incorrect initial chemical potential. The correct chemical potential in a metallic system can thus be found with very high accuracy with a few iterations.

The desired linear scaling can be obtained by introducing a localization region for each column, outside of which the elements are negligibly small. For the kth column, this localization region will be centered on the kth basis function. If we use atom-centered basis functions, then the localization region will consequently be centered on the atom to which this kth basis function belongs. We then have only to calculate that part of each column that corresponds to this localization region. This means that we can use a truncated Hamiltonian $H(k)$ which retains only the matrix elements corresponding to the basis functions contained within the localization region k. Denoting the number of basis functions in this region by M_{loc} (which might actually depend on the localization region k being considered), we then have an overall computational cost $M_b M_{loc} n_{pl} n_H$, which scales linearly. Let us stress that the size of the localization region is independent of the degree of the polynomial. If one uses, for instance, a polynomial of degree $n_{pl} = 50$, the recursion in Eq. (49) will extend over the 50 nearest-neighbor shells without localization constraint for a Hamiltonian coupling only nearest neighbors. The localization region, however, is typically much smaller, comprising just a few nearest-neighbor shells. Imposing a localization region introduces some subtleties. For instance, the eigenvalues of the truncated density matrix are no longer exactly given by $p(\epsilon_n)$ and F is no longer strictly symmetric. More importantly, strictly speaking, we can no longer use the trace notation, since we use different local Hamiltonians $H(k)$ to build up the different columns of the density matrix. The band-structure energy E_{BS} now has to be written as

$$E_{BS} = \sum_k \sum_j [p(H(k))]_{k,j} [H(k)]_{j,k}. \tag{53}$$

Another important quantity is the force. The force acting on the αth atom at position R_α is obtained by differentiating the total energy with respect to these positions. The total energy consists of the band-structure part and possibly other contributions. We shall only discuss the nontrivial part of the force arising from the differentiation of the band-structure energy E_{BS}. For simplicity let us assume that we have a simple polynomial expansion and not a Chebyshev expansion. Let us also assume that we calculate the full density matrix, i.e., that we do not truncate H by introducing a localization region. We then obtain

$$\frac{dE_{BS}}{dR_\alpha} = \frac{d}{dR_\alpha} \text{Tr}\left[H \sum_\nu c_\nu H^\nu \right] = \sum_\nu c_\nu \text{Tr}\left[\frac{\partial H^{\nu+1}}{\partial R_\alpha} \right]. \tag{54}$$

Let us consider, for instance, the term for which $\nu = 2$,

$$\frac{d \, \text{Tr}[H^3]}{dR_\alpha} = \text{Tr}\left[HH \frac{\partial H}{\partial R_\alpha} \right] + \text{Tr}\left[H \frac{\partial H}{\partial R_\alpha} H \right]$$
$$+ \text{Tr}\left[\frac{\partial H}{\partial R_\alpha} HH \right] = 3 \text{Tr}\left[HH \frac{\partial H}{\partial R_\alpha} \right], \tag{55}$$

where we used $\text{Tr}[AB] = \text{Tr}[BA]$. The final result for the force, which also holds in the case of a Chebyshev expansion, is thus

$$\frac{dE_{BS}}{dR_\alpha} = \text{Tr}\left[(p(H) + Hp'(H)) \frac{\partial H}{\partial R_\alpha} \right]. \tag{56}$$

For an insulator, the second term in the brackets $Hp'(H)$ is very small compared to the first term $p(H)$ at small but finite temperatures and it vanishes in the limit of zero temperature. The reason for this is that the eigenvalues of the matrix $p'(H)$ are $p'(\epsilon_n)$. Since at zero temperature $p'(\epsilon)$ is nonzero only at the chemical potential in the middle of the gap, all eigenvalues are zero and the matrix is identically zero. Nevertheless, retaining this term in numerical calculations is recommended because it leads to forces consistent with the total energy.

When we calculate only part of the density matrix, i.e., when we have a truncated Hamiltonian $H(k)$ following the energy expression (53), we cannot use the properties of the trace to simplify the force expression as we did in Eq. (55). The equation corresponding to Eq. (55) therefore reads

$$\sum_{k,j1,j2,k} [H(k)]_{k,j1} [H(k)]_{j1,j2} \left(\frac{\partial H(k)}{\partial R_\alpha} \right)_{j2,k}$$
$$+ [H(k)]_{k,j1} \left(\frac{\partial H(k)}{\partial R_\alpha} \right)_{j1,j2} [H(k)]_{j2,k}$$
$$+ \left(\frac{\partial H(k)}{\partial R_\alpha} \right)_{k,j1} [H(k)]_{j1,j2} [H(k)]_{j2,k}. \tag{57}$$

Similar results hold for all the other terms with different values of ν. For a Chebyshev expansion the situation is completely analogous, only the formulas are more complicated. The force formula has been worked out in this case by Voter *et al.* (1996) and is given by

$$\frac{dT_j(H)}{dR_\alpha} = \frac{dT_{j-2}(H)}{dR_\alpha} + \sum_{i=0}^{j-1} (1+k_i)(1+k_{j-1-i})$$
$$\times T_i(H) \frac{\partial H}{\partial R_\alpha} T_{j-1-i}(H), \tag{58}$$

where $k_j = 0$ if $j \leq 0$ and $k_j = 1$ otherwise. In the typical tight-binding context $\partial H / \partial R_\alpha$ is a very sparse matrix. If it contains n_D nonzero elements, we need of the order of $n_{pl}^2 n_D M_b$ operations to evaluate all the forces according to Eq. (58). The error incurred by using the approximate formula (56) is small if the localizarion region is large

enough. Since the approximate formula can be evaluated with order $n_{pl}n_D M_b$ operations, it might actually be preferable to do so. In a molecular dynamics simulation, the largest deviations in the conservation of the total energy come from events in which atoms enter or leave localization regions, and this kind of error is not taken into account by either force formula.

All the above force formulas were derived for the case in which we have a constant chemical potential and in which the polynomial representing the Fermi distribution thus does not change. Frequently, however, one wants to do simulations for a fixed number of electrons rather than for a fixed chemical potential. In this case one has to readjust the chemical potential for each new atomic configuration. The chemical potential is thus a function of all the atomic positions $\mu = \mu(R_\alpha)$, but the explicit functional form of this dependence is not known. The force formula can also be adapted to this case (Roberts and Clancy, 1998). Ignoring the above warnings and again using trace notation for simplicity, we have

$$E_{BS} = \text{Tr}[Hp(H - \mu I)] \qquad (59)$$

$$N_{el} = \text{Tr}[p(H - \mu I)] \qquad (60)$$

and consequently

$$\frac{dE_{BS}}{dR_\alpha} = \text{Tr}\left[(Hp' + p)\frac{\partial H}{\partial R_\alpha}\right] - \text{Tr}[(Hp')]\frac{\partial \mu}{\partial R_\alpha}, \qquad (61)$$

$$\frac{dN_{el}}{dR_\alpha} = \text{Tr}\left[p'\frac{\partial H}{\partial R_\alpha}\right] - \text{Tr}[p']\frac{\partial \mu}{\partial R_\alpha}. \qquad (62)$$

Since dN_{el}/dR_α has to be equal to zero, we can solve Eq. (62) for $\partial \mu / \partial R_\alpha$ and insert it into Eq. (61) to obtain the force under the constraint of a constant number of electrons.

Let us finally derive a force formula for the case in which a local charge neutrality condition is enforced (Kress *et al.*, 1998). The motivation for this approach is that in non-self-consistent tight-binding calculations one frequently finds an unphysically large transfer of charge between atoms. In a self-consistent calculation the electrostatic potential, built up by a charge transfer, is counteracting a further charge flow and thus limits charge transfer to reasonably small values. Some tight-binding schemes (Horsfield, Godwin, *et al.*, 1996) enforce a so-called local charge neutrality condition requiring that the total charge associated with an atom in a molecule or solid be equal to the charge of the isolated atom. This is done by determining a potential offset u_α for each atom α in the system which will ensure this neutrality. The total Hamiltonian H of the system is then given by $H_0 + U$, where H_0 is the Hamiltonian without any potential bias and U a diagonal matrix containing the atomic potential offsets u_α. The band-structure energy is given by

$$E_{BS} = \text{Tr}[(H_0 + U)p(H_0 + U)] - \sum_\alpha Q_\alpha u_\alpha, \qquad (63)$$

where the term containing the atomic valence charges Q_α has been subtracted to make the expression invariant under the application of a uniform potential bias to all atoms in the system. Expressed in terms of the density matrix, the local charge neutrality condition becomes

$$\sum_l p(H)_{\alpha,l;\alpha,l} = Q_\alpha. \qquad (64)$$

In Eq. (64) we have labeled the basis functions by a composite index, where α indicates on which atom the basis function is centered, and where l describes the character of the atom-centered basis function. If we have carbon atoms, for which $Q_\alpha = 4$, l would, for instance, denote the four orbitals $2s$, $2px$, $2py$, $2pz$. Using Eq. (64), Eq. (63) then simplifies to

$$E_{BS} = \text{Tr}[H_0 p(H_0 + U)]. \qquad (65)$$

Taking the derivative, we get

$$\frac{dE_{BS}}{dR_\alpha} = \sum_\beta \frac{\partial E_{BS}}{\partial u_\beta}\frac{\partial u_\beta}{\partial R_\alpha} + \frac{\partial E_{BS}}{\partial R_\alpha}, \qquad (66)$$

where

$$\frac{\partial E_{BS}}{\partial u_\beta} = \text{Tr}\left[H_0 p'(H)\frac{\partial H}{\partial u_\beta}\right]. \qquad (67)$$

As discussed above, the matrix $p'(H)$ is close to zero in an insulator at sufficiently low temperature and can often be neglected. The forces are therefore approximately given by

$$\frac{dE_{BS}}{dR_\alpha} = \text{Tr}\left[H_0 p(H)\frac{\partial H}{\partial R_\alpha}\right]. \qquad (68)$$

It has to be pointed out that to get sufficiently high accuracy the degree of the polynomial has to be higher than in the tight-binding case without local charge neutrality.

In general, the degree n_{pl} of the polynomial needed to represent the Fermi distribution is proportional to

$$n_{pl} \propto \frac{\epsilon_{max} - \epsilon_{min}}{\Delta\epsilon}. \qquad (69)$$

This follows from the fact that the nth-order Chebyshev polynomial has n roots and so a resolution that is roughly proportional to $1/n$. For the usual tight-binding Hamiltonians the ratio in Eq. (69) is not very large, and for silicon and carbon systems without gap states polynomials of degree 50 are sufficient if no local charge neutrality is enforced. In contexts other than tight-binding this ratio can, however, be fairly large and polynomial representation would become very inefficient.

2. The rational Fermi operator expansion

If very high order polynomials would be needed, a rational representation of the density matrix (Goedecker, 1995) is more efficient:

$$F = \sum_\nu \frac{w_\nu}{H - z_\nu}. \qquad (70)$$
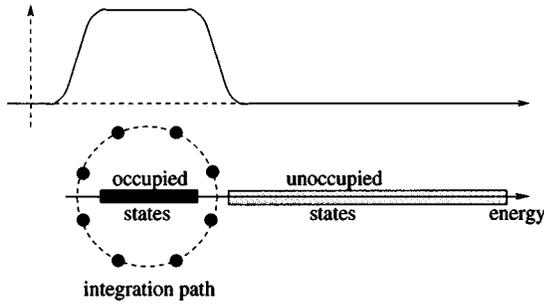
FIG. 8. A discretization of the contour integral in the complex energy plane of Eq. (71). The resulting Fermi distribution is shown on top.

As is well known, the function $f(\epsilon)$ given by

$$f(\epsilon) = \frac{1}{2\pi i} \oint \frac{dz}{\epsilon - z} \qquad (71)$$

is equal to 1 if ϵ is within the volume encircled by the contour integration path and zero otherwise. If the integration path contains the occupied states, as shown in Fig. 8, it can therefore be used as a zero-temperature Fermi distribution. The exact finite-temperature Fermi distribution $f(\epsilon_n(k))$ can be obtained by replacing the path in this contour integral by a sequence of paths that do not intersect the real axis (Goedecker, 1993; Nicholson and Zhang, 1997; Gagel, 1998). Actually, as already mentioned above, it is usually not necessary to have the exact Fermi distribution. The electronic temperature is just determined by the slope (and possibly some higher derivatives) of the distribution at the Fermi energy. Hence we shall also refer to such generalized distributions as Fermi distributions. A distribution of this type can be obtained by discretizing the zero-temperature contour integral from Eq. (71), as shown in Fig. 8.

In principle, any other set of z_ν's and w_ν's can be used as long as it satisfies

$$f(\epsilon) \approx \sum_{\nu=1}^{n_{pd}} \frac{w_\nu}{\epsilon - z_\nu}, \qquad (72)$$

where n_{pd} is the degree of the rational approximation. How can we now evaluate Eq. (70) on a computer? Denoting $I/(H - z_\nu)$ by F_ν we have

$$(H - z_\nu)F_\nu = I, \qquad (73)$$

$$F = \sum_\nu w_\nu F_\nu. \qquad (74)$$

So we first have to invert all the matrices $H - z_\nu$ and then need to form linear combinations of them. The inversion is equivalent to the solution of M_b linear systems of equations. This can be done using iterative techniques so that in the end everything can again be solved by matrix-times-vector multiplications. A rational approximation can represent the sharp variation near the chemical potential of a low-temperature Fermi distribution in a more efficient way than a Chebyshev approximation. Whereas in the Chebyshev case the degree of

the polynomial is given by Eq. (69), the degree of the rational approximation n_{pd} is given by

$$n_{pd} \propto \frac{\mu - \epsilon_{min}}{\Delta \epsilon}. \qquad (75)$$

Thus n_{pd} in contrast to n_{pl} does not depend on the largest eigenvalue ϵ_{max}. Once n_{pd} is of the order of magnitude given by Eq. (75) one has exponential convergence to the zero-temperature Fermi distribution. In the case, where the integration points and weights are obtained by discretizing the contour integral of Fig. 8, this exponential behavior is immediately comprehensible since an equally spaced integration scheme gives exponential convergence for periodic functions (Sloan and Joe, 1994). Since n_{pd} is usually reasonably small, the success of the method will hinge upon whether it is possible to solve the linear system of equations associated with each integration point with a small number of iterations. The number of iterations in an iterative method such as a conjugate-gradient scheme (Press *et al.*, 1986) is related to whether it is possible to find a good preconditioning scheme. For plane-wave calculations, a good preconditioner can be obtained from the diagonal elements and of the order of 10 iterations are required. In other schemes using Gaussians, for instance, it is not clear whether good preconditioners can be found. When the Hamiltonian depends on the atomic positions R_α, Eqs. (73) and (74) can be differentiated to obtain the derivative dF/dR_α, which is needed for the calculation of the forces.

B. The Fermi operator projection method

The FOE method is used to calculate the full density matrix. This can be inefficient if the number of basis functions per atom is very large. As was mentioned before, the density matrix at zero temperature does not have full rank. For an insulator, it can be constructed from N_{el} Wannier functions [Eq. (23)]. If one has a numerical representation of the zero-temperature density operator, which is actually a projection operator and which eliminates all components belonging to eigenvalues above the Fermi level, one can apply it to a set of trial Wannier functions $\tilde{V}_n$, $n = 1, \ldots, N_{el}$ to generate a set of orbitals that span the space of the Wannier functions. This is the basic idea of the Fermi operator projection (FOP) method. The numerical representation of the density operator can again either be a Chebyshev or rational one. We shall first discuss the rational case (Goedecker, 1995).

To do the projection with a rational representation, a system of equations analogous to Eqs. (73) and (74) has to be solved for each trial Wannier function $\tilde{V}_n$ and at each integration point ν:

$$(H - z_\nu)\tilde{W}_{n,\nu} = \tilde{V}_n, \qquad (76)$$

$$\tilde{W}_n = \sum_\nu w_\nu \tilde{W}_{n,\nu}. \qquad (77)$$
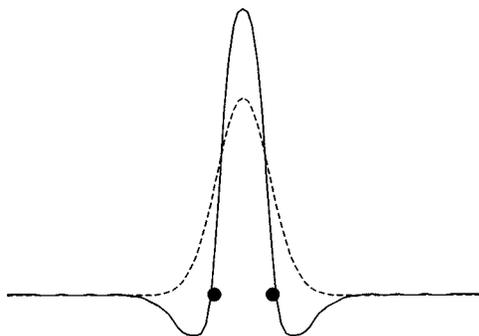
FIG. 9. The effect of applying the density operator, which is a projection operator in the eigenvalue space, to a Gaussian (dashed line) centered in the middle of a bond between two silicon atoms denoted by discs. The resulting function $\tilde{W}$ is shown by the solid line. The orthogonal Wannier function W obtained by symmetric orthogonalization is practically indistinguishable from $\tilde{W}$ on this scale. The calculation was done using density-functional theory with pseudopotentials.

Thus the savings comes from the fact that one has to solve this equation (76) only for N_{el} right-hand sides, whereas one has M_b right-hand sides in Eq. (73). Obviously the solution of Eq. (76) has to be done not within the whole computational volume but only within the localization region to obtain linear scaling. The functions $\tilde{W}_n$ will now span our subspace unless one of our trial functions $\tilde{V}_n$ was chosen in such a way that it has zero overlap with the space of the occupied orbitals, which is highly unlikely. To obtain a set of valid Wannier functions W_n one still has to orthogonalize the orbitals $\tilde{W}_n$. Since the W_n's are localized, the overlap matrix is a sparse matrix and can be calculated with linear scaling. In the typical density-functional context, the inversion of this matrix is a rather small part, even if it is done with cubic scaling. In a tight-binding context it is much more important and linear scaling methods have been devised by Stephan and Drabold (1998) and by Challacombe (1999). The construction of the Wannier functions by projection according to Eqs. (76) and (77) is illustrated in Fig. 9 for the case of a silicon crystal. In this case one knows that the Wannier functions are bond centered and it is therefore natural to choose a set of bond-centered functions as an initial guess. In this example we used simple Gaussians. As shown in Fig. 9, the projection modifies the details of the Gaussian but does not significantly change its localization properties.

Chebyshev-based projection methods have been introduced in connection with other techniques by Sankey *et al.* (1994) and by Stephan and Drabold (1998).

C. The divide-and-conquer method

The original formulation of the divide-and-conquer (DC) method (Yang, 1991a, 1991b; Zhao and Yang, 1995) was based on a subdivision of the electronic density. To calculate the density at a certain point, an ordinary electronic structure calculation is done for a sub-

volume containing this point. Since the electronic density is only influenced by features in a rather small neighborhood, the density obtained in this way at that point is a good approximation to the density one would obtain by doing a calculation in the whole volume occupied by the molecule or solid under consideration. The more recent formulation of this theory (Yang and Lee, 1995) is, however, also based on the density matrix and we shall discuss this version in more detail. The idea is to calculate certain regions of the density matrix by considering subvolumes and then to generate the full density matrix by adding up these parts with the appropriate weights. How to calculate the density matrix for the subvolumes is, in principle, unspecified, but it is usually done using ordinary electronic structure calculations based on exact diagonalization. Let us illustrate the principle for synthesizing the density matrix from its subparts in a pictorial way by considering a one-dimensional chain molecule. For simplicity let us also assume that we have atom-centered basis functions.

First one divides the whole computational volume into subvolumes, which we shall call central regions. Three such central regions are displayed in Fig. 10 by the dark green, blue, and red colors. Around each central region one puts a buffer region denoted by light green, blue, and red. The sum of these two regions corresponds to the localization region of the other O(N) methods.

Once one has set up this subdivision, one does an electronic structure calculation within each localization region to obtain the density matrix belonging to this region. These different calculations are only coupled by the requirement that the Fermi level be the same in all the localization regions. Typically a very small temperature is used to stabilize the search for the overall Fermi level. From the density matrices obtained in this way, one cuts off all the corners, i.e., the regions where both matrix indices belong to basis functions in the buffer region. For the blue region, this series of transformations is shown in Fig. 11.

Using these cross-shaped blocks one then finally synthesizes the density matrix as shown in Fig. 12 by adding up the different contributions. The regions shown in dark colors which do not overlap with other regions have weight one, whereas the overlapping regions indicted by light colors each have weight one-half. Hence the sum of the weights is one in the overlap regions, as well.

The achievement of linear scaling in the divide-and-conquer and the Fermi operator expansion methods is conceptually related. In both methods certain parts of the density matrix are calculated independently. The main difference is that in the FOE method no calculated parts of the density matrix are discarded as in the divide-and-conquer method, as depicted in Fig. 11. The Fermi operator expansion method can thus be considered as a divide-and-conquer method in which only the central part of the density matrix, which is not contaminated by the boundary of the localization region, is calculated. The fact that in the divide-and-conquer method only a
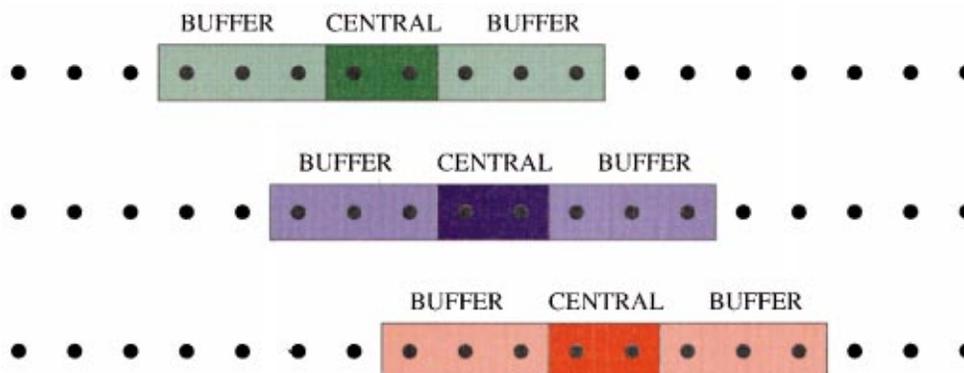
FIG. 10. The different subvolumes described in the text which are necessary for the application of the divide-and-conquer method to a chain molecule. The atoms are denoted by black dots.

small part of the density matrix obtained by costly diagonalizations is used, while the largest part associated with the buffer region is thrown away, results in a large prefactor [Eq. (2)] for this method. This is clearly a particularly serious disadvantage if large localization regions are needed, as will be discussed in more detail in Sec. IV.

The calculation of the forces acting on the atoms within the divide-and-conquer method is also described by Yang and Lee (1995). Their force formula is based on the Hellmann-Feynman theorem (Feynman, 1939) as well as some other terms, such as Pulay forces (Pulay, 1977), which arise from the use of atom-centered basis sets and auxiliary charge densities. As has been discussed in the case of the FOE method, the Hellmann-Feynman expression for the force [Eq. (56)] is not exactly consistent with the total energy expression in a nonvariational method, since it is based on the assumption that one is allowed to take traces. Even though the density matrix in the divide-and-conquer method is not calculated via a polynomial expansion, the analysis given for the FOE method also applies to the divide-and-conquer method since conceptually one can represent any matrix functional of H as a polynomial Taylor expansion. The total energy will consequently not have its minimum at exactly the same place where the Hellmann-Feynman forces vanish if both quantities are calculated with the divide-and-conquer method. For the FOE method, there is a simple analytic expression for

the calculation of the total energy, even in the case in which localization constraints are imposed [Eq. (53)]. One can therefore differentiate it without using the simplifications arising from the use of traces to obtain consistent forces. No such simple prescription can be written down for the divide-and-conquer method which would allow the calculation of consistent forces. Plainly this compatibility problem becomes negligible for large localization regions, and there are certainly practical applications where small inconsistencies of forces and energies are tolerable.

D. The density-matrix minimization approach

The density-matrix minimization (DMM) approach of Li, Nunes, and Vanderbilt (1993) is another method by which the full density matrix is constructed. In contrast to the FOE method one obtains the density matrix F in the limit of zero temperature, so no adjustable temperature parameter enters the calculation. The density matrix is obtained by minimizing the following functional for the grand potential Ω with respect to F:

$$\Omega = \mathrm{Tr}[(3F^2 - 2F^3)(H - \mu I)]. \tag{78}$$
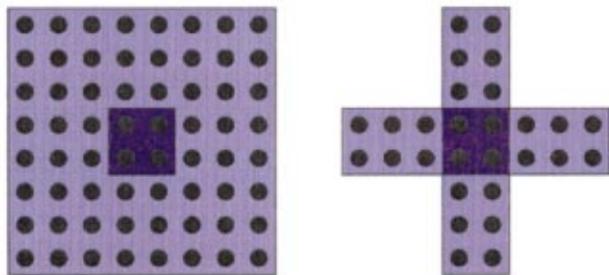
FIG. 11. Density matrix for a certain localization region. From the full density matrix (left-hand side) only the cross-shaped part (right-hand side) is used.
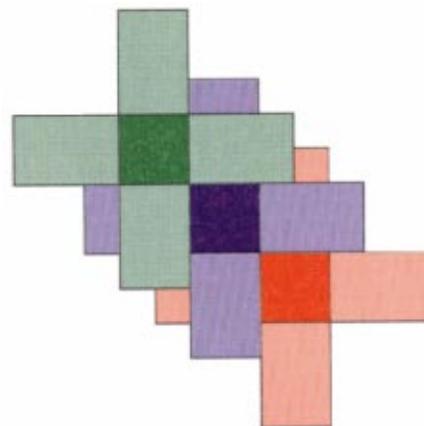
FIG. 12. Contributions to the full density matrix from the different localization regions. In this figure only the three contributions from the localization regions of Fig. 10 are shown.
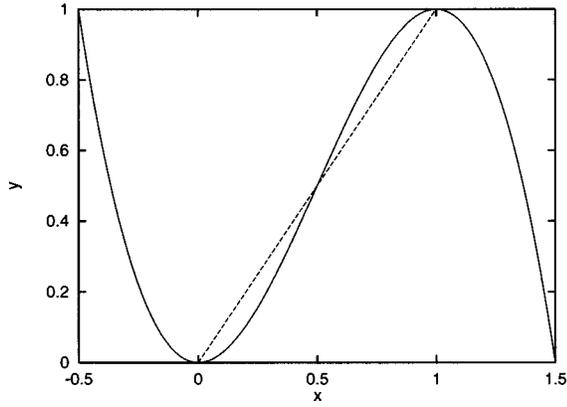
FIG. 13. The McWeeny (1960) purification function $y = 3x^2 - 2x^3$.

There is no constraint imposed during the minimization, so all the matrix elements of F are independent degrees of freedom. Nevertheless, the final density matrix will obey the correct constraint of being a projector if no localization constraints are imposed. This is related to the fact that the matrix $3F^2 - 2F^3$ is a purified version of F, as can be seen from Fig. 13. If F has eigenvalues close to zero or 1 then the purified matrix will have eigenvalues that are even closer to the same values. It is also clear from Fig. 13 that the eigenvalues of the purified matrix are contained in the interval $[0;1]$ as long as the eigenvalues of F are in the interval $[-1/2;3/2]$.

The gradient of Ω as given by Eq. (78) with respect to F is itself a matrix, and it is given by

$$\frac{\partial \Omega}{\partial F} = 3(FH' + H'F) - 2(F^2H' + FH'F + H'F^2), \quad (79)$$

where $H' = (H - \mu I)$. In order to verify that Eq. (78) defines a valid functional we have to show two things: first, that the grand potential expression (78) gives the correct result if we insert the exact density matrix F, and second, that the gradient (79) vanishes in this case. From Eq. (24) we see that the exact F is a projection operator, i.e, that $F^2 = F$. Therefore $(3F^2 - 2F^3) = F$ and the grand potential expression (78) indeed agrees with the correct result (11). Using in addition the fact that H' and the exact F commute [as follows from Eqs. (20) and (21)], it is also evident that the gradient in Eq. (79) vanishes. The gradient vanishes, however, not only for the ground-state density matrix F but also for any excited-state density matrix. In order to exclude the possibility of local minima, we have to verify that these stationary points are not minima. This can easily be done (Vanderbilt, 1998, private communication) using the fact that the functional is a cubic polynomial with respect to all its degrees of freedom. Let us suppose that there are two minima. Inspecting the functional along the line connecting these two minima, we would obviously again find these two minima, which is a contradiction because a cubic polynomial cannot have two minima. Thus we have proved by contradiction that the DMM functional has only one single minimum.

There is a second feature that is worrisome at first

sight with this functional. If the density matrix for an insulator at zero temperature is of the correct form (i.e., if the occupation numbers n_l are 0 or 1), the gradient (79) will vanish independently of the value of the chemical potential. This ambiguity, however, disappears as soon as one has fractional occupation numbers. Let us consider an approximate density matrix of the form

$$F = \sum_l n_l |\Psi_l\rangle\langle\Psi_l|. \quad (80)$$

Then it is easy to see that

$$\Omega = \sum_l (\epsilon_l - \mu)(3n_l^2 - 2n_l^3), \quad (81)$$

$$\frac{\partial \Omega}{\partial F} = \sum_l 6(\epsilon_l - \mu)n_l(1 - n_l)|\Psi_l\rangle\langle\Psi_l|. \quad (82)$$

The polynomial of Eq. (81) is the same as that shown in Fig. 13, and we see that components corresponding to eigenvalues larger than the chemical potential are damped until they vanish in the minimization process, whereas components corresponding to smaller eigenvalues are amplified until they reach the value one. Thus the chemical potential will determine the number of electrons to be found in the system, as it should. The above statements are actually only correct if all the n_l's are contained in the interval $[-1/2:3/2]$. If this is not the case then one can see from Fig. 13 that there can be runaway solutions, where some n_l tend to $\pm\infty$. When we implemented the scheme, however, we never encountered in practice such a runaway case.

Having convinced ourselves that the functional defined in Eq. (78) is well behaved, let us now estimate the number of iterations necessary to minimize it. As is well known, the error reduction per iteration step depends on the condition number κ, which is the ratio of the largest curvature a_{max} to the smallest curvature a_{min} at the minimum. These curvatures could be determined exactly by calculating the Hessian matrix at the minimum. Let us instead only derive an estimate of these curvatures by calculating the curvature along some representative directions. To do this let us now consider a ground-state density matrix where some fraction x of an excited state is mixed in

$$F(x) = \sum_{n=1}^{N_{el}} \Psi_n^*(\mathbf{r})\Psi_n(\mathbf{r}) - x\Psi_I^*(\mathbf{r})\Psi_I(\mathbf{r}) + x\Psi_J^*(\mathbf{r})\Psi_J(\mathbf{r}). \quad (83)$$

Here, the index I is a member of the N_{el} eigenstates below μ, and the index J refers to a state above μ. The expectation value of the orbital minimization functional for this density matrix is given by

$$\Omega(x) = \text{Tr}[(3F(x)^2 - 2F(x)^3)(H - \mu I)]$$
$$= \sum_{n=1}^{N_{el}} \epsilon_n + (3x^2 - 2x^3)(\epsilon_J - \epsilon_I) \quad (84)$$

and its second derivative by

$$\left.\frac{\partial^2 \Omega(x)}{\partial x^2}\right|_{x=0} = 6(\epsilon_J - \epsilon_I). \tag{85}$$

The largest curvature will be roughly $\epsilon_{max} - \epsilon_{min}$ and the smallest curvature of the order of the HOMO-LUMO[1] separation $\epsilon_{gap} = \epsilon_{N_{el}+1} - \epsilon_{N_{el}}$. The condition number is thus given by

$$\kappa = \frac{a_{max}}{a_{min}} \approx \frac{\epsilon_{max} - \epsilon_{min}}{\epsilon_{gap}}. \tag{86}$$

In the conjugate-gradient method, which is the most efficient method of minimizing the DMM functional, the error e_k decreases as follows (Saad, 1996):

$$e_k \propto \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k. \tag{87}$$

The error e_k is defined in this context as the length of the vector, which is the difference between the exact and approximate solutions at the kth iteration step. Under realistic conditions κ is large and the number of iterations n_{it} to achieve a certain accuracy is therefore

$$n_{it} \propto \sqrt{\kappa} = \sqrt{\frac{\epsilon_{max} - \epsilon_{min}}{\epsilon_{gap}}}. \tag{88}$$

This is an important result, since it indicates that in an insulator the number of iterations is independent of system size. This result is also confirmed by numerical tests.

The use of a conjugate-gradient scheme requires line minimizations along these conjugate directions. For arbitrary functional forms this has to be done by numerical techniques such as bisection (Press *et al.*, 1986). For the DMM functional, however, we have a cubic form along each direction. The four coefficients determining the cubic form can be calculated with four evaluations of the functional. Once these four coefficients are known, the minimum along this direction can easily be found.

Doing a series of minimization steps as outlined above will, in general, result in a density matrix that does not lead to the correct number of electrons. Thus one has to do some outer loops where one searches for the correct value of the chemical potential. For better efficiency, these two iteration loops can, however, be merged into one loop where one alternatingly minimizes the energy and adjusts the chemical potential (Qiu *et al.*, 1994).

The forces on the atoms are given by

$$\frac{d\Omega}{dR_\alpha} = \frac{\partial\Omega}{\partial F}\frac{\partial F}{\partial R_\alpha} + \frac{\partial\Omega}{\partial H}\frac{\partial H}{\partial R_\alpha}. \tag{89}$$

Since the method is variational, $\partial\Omega/\partial F$ vanishes at the solution and the force formula simplifies to

$$\frac{d\Omega}{dR_\alpha} = \frac{\partial\Omega}{\partial H}\frac{\partial H}{\partial R_\alpha} = \mathrm{Tr}\left[(3F^2 - 2F^3)\frac{\partial H}{\partial R_\alpha}\right], \tag{90}$$

which can easily be evaluated.
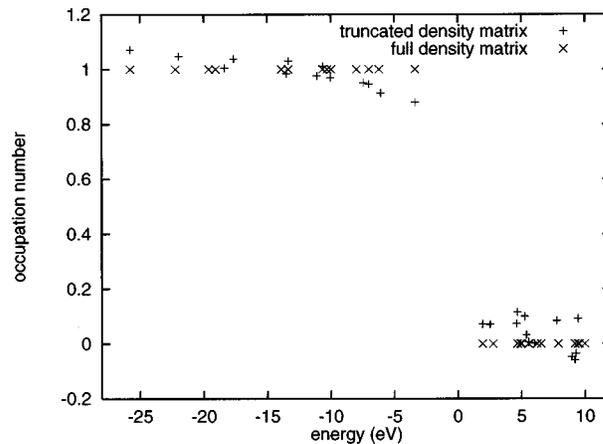
FIG. 14. An analysis of the eigenvectors of the full and truncated density matrix. In the case of the full density matrix the eigenvectors were chosen to be simultaneously eigenvectors of both F and H, and the eigenvalues with respect to F (occupation numbers) are plotted vs the eigenvalues with respect to H. For the truncated density matrix, the eigenvectors can no longer simultaneously diagonalize F and H. Therefore the eigenvalues with respect to F are plotted vs their energy expectation values with respect to H. Note that in the energy expression (78) the purified density matrix $3F^2 - 2F^3$ enters instead of F. The occupation numbers of the purified version are closer to zero or one.

The introduction of a localization region leads again to some subtleties. Whereas in the unconstrained case the eigenvalues of the final density matrix F will all be either zero or one, this is no longer the case when a localization region is introduced. Hence the truncated F is no longer a projection matrix but it is given by

$$F = \sum_{m=1}^{M_b} n_m \Psi_m^*(\mathbf{r})\Psi_m(\mathbf{r}), \tag{91}$$

where now Ψ_m are the eigenfunctions of the truncated F and the occupation numbers n_m are their eigenvalues. In a certain sense the localization constraint introduces a finite electronic temperature. This is actually not surprising, considering the relation between the temperature and the localization properties discussed in Sec. II. Figure 14 shows the energy expectation values of the eigenvectors of F versus the occupation numbers, for a crystalline Si cell of 64 atoms, where the localization region extends to the second-nearest neighbors. As one can see, the energy expectation values $\langle\Psi_m|H|\Psi_m\rangle$ of the eigenvectors of F are very close to the exact eigenvalues of H.

This close correspondence of the eigenvectors of F to the eigenvectors of H explains why the number of iterations needed to find the minimum does not increase as one introduces localization constraints. Equation (85) remains approximately valid if the occupation numbers

---

[1]Highest-occupied-molecular-orbital/lowest-unoccupied-molecular-orbital.

for the occupied states are close to 1 and if the occupation numbers for the unoccupied states are very small, as well as if the energy expectation values $\langle \Psi_m|H|\Psi_m\rangle$ are close to the exact eigenvalues of the Hamiltonian. These conditions are fulfilled as discussed above. Hence the condition number for the minimization process does not change appreciably in the truncated case.

All the arguments used to prove the absence of local minima remain valid in the truncated case as well. The force formula Eq. (90) remains equally valid.

An alternative derivation of this algorithm has been given by Daw (1993). He considers a differential equation which describes the evolution of a density matrix when the electronic temperature is cooled down from infinity to zero. The change of the density matrix during this process is equal to the gradient of Eq. (79).

E. The orbital minimization approach

The orbital minimization (OM) method (Mauri *et al.*, 1993; Ordejon *et al.*, 1993; 1995; Mauri and Galli, 1994, Kim *et al.*, 1995) also calculates the grand potential in the limit of zero temperature. In contrast to the previous methods, it does not calculate the density matrix directly but expresses it via the Wannier functions according to Eq. (23). These Wannier functions are obtained by minimizing the following unconstrained functional:

$$\Omega = 2\sum_n \sum_{i,j} c_i^n H'_{i,j} c_j^n - \sum_{n,m} \sum_{i,j} c_i^n H'_{i,j} c_j^m \sum_l c_l^n c_l^m,$$
(92)

where c_i^n is the expansion coefficient of the nth Wannier orbital with respect to the ith basis function and $H'_{i,j}$ are the matrix elements of the shifted Hamiltonian $H-\mu I$ with respect to the basis functions. In the original formulation (Mauri *et al.*, 1993; Ordejon *et al.*, 1993; 1995; Mauri and Galli, 1994) only N_{el} orbitals were included in the orbital sums in Eq. (92), i.e., $n=1\cdots N_{el}$, $m=1\cdots N_{el}$. In the formulation of Kim *et al.* (1995) more than N_{el} orbitals are included in the sums. The functional of Eq. (92) can be derived by considering the ordinary band-structure energy expression

$$E_{BS} = \sum_n \sum_{i,j} c_i^n H'_{i,j} c_j^n$$
(93)

and by incorporating the orthogonality constraint by a Taylor expansion of the inverse of the overlap matrix O between the occupied orbitals,

$$O_{n,m} = \sum_l c_l^n c_l^m,$$
(94)

up to first order. A family of related functionals can be obtained by Taylor expansions to higher order (Mauri and Galli 1994; Galli, 1996). Since these functionals do not offer any significant advantage and are not used in calculations we shall not discuss them. The gradient of the functional of Eq. (92) is given by

$$\frac{\partial\Omega}{\partial c_k^n} = 4\sum_j H'_{k,j} c_j^n - 2\sum_m \sum_j H'_{k,j} c_j^m \sum_l c_l^n c_l^m$$

$$-2\sum_m c_k^m \sum_{i,j} c_i^n H'_{i,j} c_j^m.$$
(95)

Let us first discuss this functional in the case when no localization constraint is imposed on the orbitals. It is easy to see that the functional (92) is invariant under unitary transformations of the occupied (i.e., N_{el} lowest) orbitals, so we can derive our results in terms of eigenorbitals rather than Wannier orbitals. The coefficients c_i^n are then the expansion coefficients of the eigenorbital. Using the fact that in this case $\sum_l c_l^n c_l^m = \delta_{n,m}$ and that $\sum_{i,j} c_i^n H'_{i,j} c_j^m = \delta_{n,m}(\epsilon_n-\mu)$, we obtain

$$\Omega = 2\sum_n \sum_{i,j} c_i^n H'_{i,j} c_j^n - \sum_{n,m} \sum_{i,j} c_i^n H'_{i,j} c_j^m \delta_{n,m}$$

$$= \sum_n \sum_{i,j} c_i^n H'_{i,j} c_j^n = \sum_n \epsilon_n - \mu N_{el},$$

which is the standard expression (11) for the grand potential. Similarly the gradient equation can be simplified, obtaining

$$\frac{\partial\Omega}{\partial c_k^n} = 4\sum_j H'_{k,j} c_j^n - 2\sum_m \sum_j H'_{k,j} c_j^m \delta_{n,m}$$

$$-2c_k^n \sum_m \delta_{n,m}(\epsilon_m-\mu)$$

$$= 2\sum_j H'_{k,j} c_j^n - 2c_k^n(\epsilon_n-\mu) = 0.$$
(96)

So the functional has indeed a vanishing gradient at the ground state and it gives the correct ground-state energy. As was the case for the DMM functional, the gradient vanishes not only for the set of ground-state orbitals but also for any set of excited states. Thus we have to verify that these stationary points are not local minima but saddle points. We do this by picking a certain direction along which the curvature is negative. In the OM case an excited state is described by a set of N_{el} orbitals Ψ_n where at least one index $n=I$, corresponding to an occupied orbital I, is replaced by an unoccupied orbital J. Let us now consider the variation of the grand potential $\Omega(x)$ under a transformation of the form $\Psi_I \rightarrow \cos(x)\Psi_I+\sin(x)\Psi_J$. One can show that the curvatures at these stationary points is given by

$$\left.\frac{\partial^2\Omega(x)}{\partial x^2}\right|_{x=0} = -4(\epsilon_J-\epsilon_I).$$
(97)

Since the unoccupied eigenvalue ϵ_J is higher in energy than the occupied one ϵ_I, the curvature is negative and we have indeed a saddle point. In the same way we can again show that the condition number is given by Eq. (86).

In analogy to the DMM functional, one can also show that in the formulation of Kim *et al.* (1995), the chemical potential μ determines the number of electrons by am-
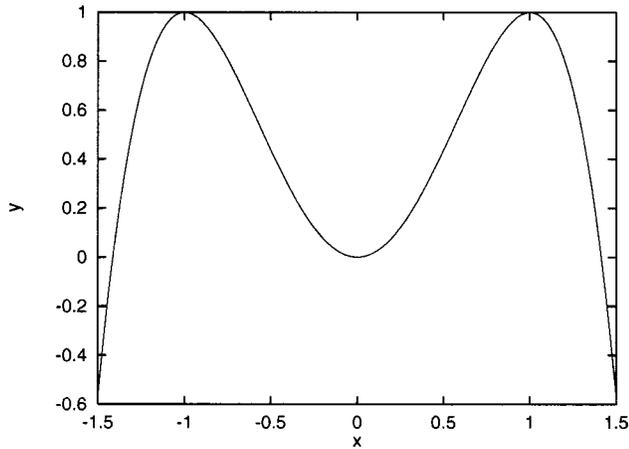
FIG. 15. The function $(2-x^2)x^2$ relevant in Eq. (98).

plifying components below μ and annihilating components above it. Considering a state consisting of a set of eigenfunctions Ψ_n of H, in which each eigenfunction is multiplied by a scaling factor a_n, the expectation value for the grand potential becomes

$$\Omega = \sum_n (2-a_n^2)a_n^2(\epsilon_n - \mu). \tag{98}$$

The relevant function $(2-x^2)x^2$ is shown in Fig. 15. One can see that the minimum of Eq. (98) is attained by $a_n = 0$ if $\epsilon_n > \mu$ and by $a_n = \pm 1$ if $\epsilon_n < \mu$. Again, this is only true if a_n is within a certain safety interval. Otherwise there can be runaway solutions. Infinitesimally close to the solution μ becomes ill defined in an insulator, as it should.

Whereas the DMM functional keeps all its good properties when one introduces a localization constraint, the OM functional loses most of them. The localization constraint is introduced in the OM functional by allowing each Wannier orbital to deviate from zero only within its own localization region. These localization regions are usually atom centered and contain a few shells of neighboring atoms. The basic idea of the OM functional, namely, of describing an electronic system by a set of Wannier functions with finite support, is already problematic. Orthogonality and finite support are mutually exclusive properties, and so the orbitals that one obtains in the minimization process are necessarily non-orthogonal. The true Wannier functions are, however, orthogonal. In addition, as we have seen in the DMM case, a density matrix that is truncated has full rank, i.e., none of its eigenvalues is exactly zero. Thus N_{el} Wannier orbitals are not sufficient to represent the density matrix in this case. The generalized formulation of Kim et al. (1995), in which more than N_{el} orbitals are used, alleviates this problem, but does not completely fix it unless the number of orbitals is equal to the number of basis functions M_b.

When implemented with localization constraints the orbital minimization functional exhibits the following problems:

• The functional has multiple minima (Ordejon et al.,

1995). Depending on the initial guess one thus obtains different answers, some of which are physically meaningless (Kim et al., 1995). As we have shown above in Eq. (97), the functional has no multiple minima in the nontruncated case. The analysis we used to show this was based on the eigenfunctions. Since the Wannier functions have no resemblance to the eigenfunctions this analysis cannot be carried over into the localized regime. For the density-matrix minimization method the absence of multiple minima could be proven using the fact that the DMM functional is cubic. The orbital minimization functional (92), however, is quartic with respect to its degrees of freedom and will thus in general have multiple minima. The problem of the multiple minima is alleviated by the formulation of Kim et al. (1995), but it is not completely removed since the functional still has quartic character. As a by-product of the multiple-minimum problem, the total energy cannot be conserved in molecular dynamics simulations, which is an important requirement. Here again, energy conservation is better in the Kim formulation but still far from perfect (Kim et al., 1995).

In practical applications of the orbital minimization approach to electronic structure methods, great care is usually taken to construct input guesses that correspond to the physical bonding properties of the molecule under consideration (Itoh et al., 1996). If the minimum that is closest in distance is always selected during the subsequent line minimizations then one will most likely end up in a physically reasonable minimum that reflects the bonding properties of the input guess (Stephan, 1998, private communication). This is especially true if the localization regions are large and if the topology of the total energy surface within a reasonably large region around the physical minimum is not too different from that in the nontruncated case. Such a procedure is, of course, not applicable to systems in which the exact bonding properties are unknown.

• The number of iterations is very large whenever any localization constraint is imposed together with tight convergence criteria. This is due to the deterioration of the condition number, a phenomenon that is easy to understand (Ordejon et al., 1995). Introducing a localization region destroys the strict invariance of the band-structure energy under unitary transformations among the occupied orbitals. When the localization region is large, this invariance will still approximately exist and one can find certain directions around the minimum where the energy varies extremely slowly and where the curvature is therefore much smaller than the smallest curvature ϵ_{gap} in the unconstrained case. Whereas directions where the curvature is strictly zero do not affect the condition number, these very small curvatures will have a negative effect on it [Eq. (86)], and the required number of iterations is consequently much larger in the constrained case than in the unconstrained case. Even though the condition number deteriorates with increasing localization region, the detri-

mental effect on the number of iterations will disappear at a certain point where the gradient due to these small curvatures becomes smaller than the numerical threshold determining the convergence criterion of the minimization procedure.

- The optimal localization regions would be centered on the centers of the Wannier functions. Since these centers are not known *a priori*, atom-centered localization regions are usually chosen. In this case the Wannier functions do not generally exhibit the correct symmetry (Ordejon *et al.*, 1995). As a consequence molecular geometries obtained from this functional can have broken symmetry as well. In a C_{60} molecule, for instance, there are only two equivalent sites. When treated with the orbital minimization functional they are each slightly different (Kim *et al.*, 1995).

- As follows from Eq. (98) there can be runaway solutions. We have encountered this problem in test cases with random numbers as input guesses. If one constructed a more sophisticated input guess, based on the bonding properties of the system, this would probably not occur. In the DMM method the possibility of runaway solutions also exists but is never found in practice, even with the most trivial input guess.

- If the method is used in the context of self-consistent calculations, where the electronic charge density is used to calculate the Hartree and exchange-correlation potentials, problems arise, since the total charge is not conserved during the minimization iteration (Mauri and Galli, 1994).

To overcome the competing requirements of orthogonality and localization, a related approach has recently been proposed by Yang (1997) in which the orbitals are allowed to be non-orthogonal. This approach of Yang has up to now not been applied in connection with a localization constraint.

F. The optimal basis density-matrix minimization method

Despite its many advantages in the tight-binding context, the DMM method has the big disadvantage that it is very inefficient if one needs very large basis sets (i.e., many basis functions per atom). Large basis sets are typically required in grid-based density-functional calculations. In this case it becomes impossible to calculate and store the full density matrix in the DMM method, even though it is a sparse matrix. From this point of view the Wannier-function-based methods are advantageous since they do not require the full density matrix. The basic idea of the optimal basis density-matrix minimization (OBDMM) method (Hierse and Stechel, 1994; Hernandez and Gillan, 1995) is first to contract the fundamental basis functions into a small number of new basis functions and then to set up the Hamiltonian and overlap matrix in this new small basis. A generalized version of the DMM method which can be applied to the non-orthogonal context (a subject that will be discussed later in the article) is then used to solve the electronic structure problem in this basis. The essential point is that one tries to do the contraction in an optimal way by minimizing the total energy with respect to the degrees of freedom determining the contracted basis functions Ψ_n. Formulated mathematically the density matrix is given by

$$F(\mathbf{r},\mathbf{r}') = \sum_{i,j} \Psi_i^*(\mathbf{r}) K_{i,j} \Psi_j(\mathbf{r}'). \tag{99}$$

The matrix K is a purified version of the the density matrix within the contracted basis L, and it is given by

$$K = 3LOL - 2LOLOL, \tag{100}$$

where O is the overlap matrix among the contracted orbitals. The main difference between the formulation of Hierse and Stechel (1994) and that of Hernandez and Gillan (1995) is that in the first formulation the number of contracted basis functions Ψ_i is equal to the number of electrons, whereas in the second approach it can be larger. In the formulation of Hernandez and Gillan (1995) the contracted basis set can, for instance, be chosen to have the size of a minimal basis set. The difference between this set and a standard minimal basis set in quantum chemistry is that it is optimally adapted to its chemical environment, since the contraction coefficients are not predetermined but found variationally. In practice, the full density matrix is found by a double-loop minimization procedure. In the inner loop one uses the ordinary DMM procedure to find the density matrix for a given contracted basis set. In the outer loop one searches for the optimally contracted basis functions Ψ_i for fixed L.

Unfortunately the minimization of the contracted basis functions Ψ_i is ill conditioned (Gillan *et al.*, 1998), and the number of iterations is therefore, at present, very large. As already explained before, ill conditioning occurs if the curvatures in the minimum along different directions are widely different. Three causes for the ill conditioning are reported by Gillan *et al.* (1998):

- Length scale ill conditioning:

  This problem is actually not related to the OBDMM algorithm itself but to the (uncontracted) basis functions that are taken to be so-called "blip" functions in the present implementation. This kind of problem can be found in all iterative electronic structure algorithms if grid-based basis functions such as finite elements are used. Its origin is easy to understand. Let us imagine that we are searching for the lowest state of jellium using a localized basis set associated with an equally spaced grid. By symmetry the solution is a constant vector, i.e., all basis functions have the same amplitude in the solution vector. Let us now assume that we explore the energy surface around the minimum along several directions. Let us first "go" in a direction where we add components in such a way that the sign of the amplitude of each neighboring basis function changes. This corresponds to a high-frequency plane wave and, since the kinetic energy of such a plane wave is large, the

total energy will rapidly increase if we add such a contribution to our solution vector. If on the other hand we add contributions that correspond to low-frequency plane waves, the energy will increase much more slowly. Since in grid-based methods the basis functions are usually narrow and since one can thus construct high-frequency functions, the condition number can be very bad. As one might suspect from the above explanation, the different curvatures can be estimated by means of a Fourier analysis. With this information one can then use preconditioning techniques to cure the length scale ill-conditioning problem. Such a scheme has been proposed by Bowler and Gillan (1998).

• Superposition ill conditioning:

This ill-conditioning problem is essentially identical to the ill-conditioning problem of the OM functional. If we have N_{el} contracted basis functions and no localization constraints, the total energy is invariant with respect to unitary transformations of these functions. The introduction of a localization constraint destroys this invariance but there is an approximate invariance left which manifests itself in very small curvatures in the minimum along certain directions.

• Redundancy ill conditioning:

This problem can only be found in the formulation of Hernandez and Gillan (1995), where the number of contracted basis functions is larger than the number of electrons. In this case one spans a space that contains not only the occupied orbitals but also some unoccupied. As was shown before in the context of the DMM functional, introducing a localization constraint will not assign zero occupation numbers but small occupation numbers to components corresponding to the unoccupied states in the unconstrained case. Since these components corresponding to the unoccupied states have very little weight, they have little influence on the total energy and one has again certain directions in which the total energy changes very slowly, resulting in very small curvatures.

Another open question is whether the OBDMM has local minima. The functional is a sixth-order polynomial with respect to the expansion coefficients of the contracted basis functions, as can be seen from Eqs. (99) and (100). The two overlap matrices in Eq. (100) each give a quadratic term, the two contracted orbitals in Eq. (99) a linear term. Minimization with respect to the contracted basis functions should therefore exhibit local minima. Local minima have, however, not been reported with this method so far. Perhaps the following density-matrix minimization step, which is free of local minima, saves the method from overall local minima.

IV. COMPARISON OF THE BASIC METHODS

It is certainly not possible to claim that a specific method is the best for all applications. Nevertheless, the methods presented so far differ in many respects and one can therefore clearly judge under which limiting circumstances certain methods will fail or perform well. In this section these methods will therefore be compared under several important aspects. The comparison will be done in two categories. The first category applies to electronic structure methods requiring a small number of degrees of freedom per atom. The tight-binding method belongs to the first category, requiring a few basis functions per atom (or just a few degrees of freedom in the case of semiempirical tight binding). But we shall also include the standard quantum chemistry methods in this first category, where one typically needs from a few up to a few dozen Gaussian-type basis functions per atom. The second category contains grid-based methods such as finite-difference schemes (Chelikowsky *et al.*, 1994), or those in which the basis functions can be associated with grid points such as in finite-element basis functions (White *et al.*, 1989) or blip basis functions (Hernandez *et al.*, 1997). In these methods one has typically many hundreds of degrees of freedom per atom. Even though the density matrix is sparse, O(N) methods that calculate the full density matrix cannot be applied to the second category of electronic structure methods. The memory requirements alone are already prohibitive. As pointed out before, we can expect that the localization region in a three-dimensional structure comprises on the order of 100 atoms. The density matrix will exhibit significant sparseness only for larger systems. Assuming that we have only 100 basis functions per atom, the storage of the density matrix would require about 1 gigabyte of memory, which is the upper limit of current workstations. The comparison in the class of large basis sets will therefore comprise only methods that are Wannier function based, namely, FOP, OM, and OBDMM. The comparison in the small basis set class will comprise FOE, DC, DMM, and OM, excluding two methods that are explicitly targeted at large basis sets, namely, FOP and OBDMM.

A. Small basis sets

The comparison of methods applicable to small basis sets is based on the following criteria:

• Scaling with respect to the size of the localization region:

The size of the localization region is taken as the number of atoms contained within it. Only the FOE method has linear scaling with respect to the size of the localization region. As one increases the size of this region the nonzero part of each column of the Chebyshev matrices increases linearly, implying a linear increase in the basic matrix-times-vector multiplication part. In the DMM method the CPU time increases quadratically, since the numerical effort for the basic matrix-times-matrix multiplications grows quadratically with respect to the number of off-diagonal elements of the matrix. Neglecting ill-conditioning problems, the orbital minimization method exhibits quadratic scaling, since the numerical effort for the calculation of the overlap matrix

[Eq. (94)] between the Wannier orbitals increases quadratically. As the localization region grows there are more matrix elements, and the calculation of each matrix element is more expensive since each orbital is described by a longer vector. Because the ill-conditioning problem becomes more severe for large localization regions, the number of iterations increases in reality in a way that is difficult to model, resulting in an effective scaling that is stronger than quadratic. The DC method has a cubic scaling with respect to the size of the localization region if each subvolume is treated with diagonalization schemes. Comparing the scaling behavior of all these methods, one can thus conclude that the FOE method will clearly perform best if large localization regions are needed. The FOE method is also the only method that can be faster than traditional cubically scaling algorithms if no localization constraints are imposed. In this case its overall scaling behavior is quadratic, whereas all other methods have a cubic scaling with a prefactor that is significantly larger than that for exact diagonalization.

- Scaling with respect to the accuracy:

A detailed comparison of the polynomial FOE method and the DMM method has recently been made by Baer and Head-Gordon (1997a, 1997b) for systems of different dimensionality. They conclude that in the one-dimensional case the DMM has the best asymptotic behavior, but its prefactor is much larger than that of the FOE method, so that the FOE method is more efficient in the relevant accuracy regime. In the two-dimensional case they have the same asymptotic behavior, but the FOE method has again a much smaller prefactor. In the most relevant three-dimensional case the FOE method has both the best asymptotic behavior and the best prefactor. These results are plausible after the preceding discussion of the scaling with respect to the size of the localization region. When one wants to improve the accuracy the most important factor is enlargement of the localization region. It is also clear that in higher dimensions the number of atoms within the localization region grows faster than in lower dimensions and that the scaling with respect to the number of atoms will thus become the decisive factor in three dimensions. In lower dimensions the number of iterations has higher relative importance, favoring the DMM method. A comparison of the FOE and DMM methods applied to quasi-two-dimensional huge tight-binding fullerenes by Bates *et al.* (1998) is also in agreement with the above statements. They found that the FOE and DMM methods gave nearly the same performance with a small advantage for the FOE method. As discussed before, the scaling of the OM and DC methods is stronger than quadratic with respect to the size of the localization region. It is therefore clear that the required numerical effort for increased accuracy will grow even faster for these methods than in the DMM method.

- Scaling with respect to the size of the gap:

In the FOE method the degree n_{pl} of the Chebyshev polynomial increases linearly with decreasing gap [Eq. (69)]. At the same time the density matrix decays more slowly. It follows from Eq. (37) that the linear extension of the localization region grows as ϵ_{gap}^{-1} in the applicable weak-binding limit. The volume of the localization region and the number of atoms contained in it consequently grow as ϵ_{gap}^{-3}. Taking into account the number of iterations [Eq. (69)], we find that the total numerical effort increases as ϵ_{gap}^{-4} in the FOE method. In the DMM method the number of iterations also increases with decreasing gap but more slowly, like $\epsilon_{gap}^{-1/2}$, as follows from Eq. (88). Taking into account the above discussion of the scaling properties of the DMM method with increasing localization region we obtain the overall scaling of $\epsilon_{gap}^{-13/2}$, which is higher than the scaling behavior of the FOE method. Obviously the scaling behavior of the OM and DC methods is worse. So in contrast to what one might first think the FOE method performs best in this limit. In three-dimensional metallic systems, the FOE method is thus expected to be the only method that will work efficiently at good accuracies.

- Finding a first initial guess:

No initial guess is required in the FOE and DC methods (except perhaps for the potential in a self-consistent calculation). In the DMM method an extremely simple and efficient input guess for the density matrix is just a diagonal matrix that sums up to the correct number of electrons. In the orbital minimization method this point is somewhat problematic. As mentioned above, a Wannier function represents typically a bond or lone electron pair. If one can draw the standard Lewis structure of a molecule, where bonds are denoted by lines and a lone electron pair by a pair of dots, one knows where the Wannier functions should be centered, and the Lewis formula can be the basis for the initial guess. This procedure cannot be used if the molecule is characterized by two or more Lewis structures that are resonating. Especially if the two Lewis structures correspond to an electron transfer over a distance larger than the range of the localization region, serious problems are to be expected with Wannier-function-based methods. In such a case it might be impossible not only to find an initial guess, but also to describe such a molecule by N_{el} localized Wannier functions.

- Number of iterations in electronic structure calculations:

In the variational methods (OM and DMM) the number of iterations depends on the condition number of the energy expression. As pointed out, the OM energy expression is ill conditioned under localization constraints, and therefore the required number of iterations is very large. Even for modest accu-

racy several hundred iterations are required (Mauri and Galli, 1994; Ordejon *et al.*, 1995). In the DMM method, on the other hand, the number of iterations is the same as one needs in ordinary [i.e., non $O(N)$] electronic structure calculations, namely, 20 to 30 (Nunes, 1998, private communication). The quantity that corresponds to the number of iterations in the FOE method is the degree n_{pl} of the polynomial. While it is difficult to compare the cost of one Chebyshev recursion step with the cost of a DMM minimization step, such a comparison can be made in the case of OM. In each orbital line minimization step one has to calculate the minimum of a quartic polynomial, which requires at least three applications of H to the wave function. One Chebyshev recursion step requires one application of H.

- Number of iterations in molecular dynamics simulations:

In molecular dynamics simulations as well as structural relaxation steps and self-consistent mixing schemes, the density matrix or the Wannier functions from the previous step are a good input guess for the next step. Good initial input guesses are beneficial in all methods except in the polynomial FOE method and the DC method. It is difficult to quantify the possible savings of such a reuse. To preserve the quality of the solution of the preceding step as an input guess in a molecular dynamics simulation, it may be necessary to make the time step smaller than the integration scheme would allow. How large the maximum time step can be also depends, of course, on the order and properties of the time integration scheme used to propagate the molecular dynamics simulation. Similar remarks apply to the case of structural relaxations. The decisive factor determining the number of iterations per molecular dynamics step is in this context again the condition number of the functional. With the DMM methods, of the order of 2 to 3 steps are needed both for accurate molecular dynamics simulations (Qiu, 1994) and for structural relaxations (Nunes, 1998, private communication). The smallest number of iterations that was used in molecular dynamics simulations with the orbital minimization method was 10, but at the price of very poor energy conservation.

- Crossover point for standard tight-binding systems:

The FOE method has the lowest reported crossover point for the standard carbon test system in the crystalline diamond structure. For the FOE method it is around 20 atoms (Goedecker, 1995), and for the DMM it is estimated (Li *et al.*, 1993) to be around 90 atoms. No crossover points were ever given for the OM and DC methods, and presumably they are much higher. All quoted crossover points for electronic structure calculations are for an accuracy of roughly 1% in the cohesive energies, but in the relevant publications not all computational details are listed to ensure that these numbers are really comparable in all respects. The low crossover point of the FOE method can be understood in terms of its scaling behavior with respect to the size of the localization region discussed above. For small systems the size of the localization region equals the size of the whole system. The FOE method therefore starts off with a quadratic scaling behavior, whereas all other methods start off with a cubic behavior. Consequently the crossover point for all other methods can only be for system sizes larger than the localization region, whereas the crossover point in the FOE method can already be at smaller system sizes if it is implemented efficiently.

In the context of molecular dynamics simulations the crossover points are different because some of the variational methods can benefit from good input guesses. For the FOE method the crossover point remains at 20 atoms, for the OM method Mauri and Galli (1994) quote 40 atoms, and for the DMM method Qiu *et al.* (1994) quote 60 atoms. Again no crossover point is given for the DC method. It has to be stressed, however, that the numbers for the FOE and DMM methods were for highly accurate molecular dynamics runs where the total energy was conserved, while in the benchmarks of Mauri and Galli (1994) no satisfactory energy conservation was obtained.

- Influence of the range of a sparse Hamiltonian matrix on the performance:

In the FOE method the numerical effort increases strictly linearly with respect to the number of nonzero elements per column, which depends cubically on the range of the Hamiltonian matrix. In the case of the DMM method it can be shown (Li *et al.*, 1993) that one has to calculate intermediate product matrices only up to a range that is the sum of the ranges of the density matrix and the Hamiltonian matrix. As long as the range of the Hamiltonian is small compared to the range of the density matrix, the number of operation increases only very weakly with respect to an increasing Hamiltonian range. The DMM method therefore outperforms the FOE method under such circumstances (Daniels and Scuseria, 1998). Hamiltonian matrices of relatively large range are found in the context of density-functional calculations using Gaussian basis sets. For tight-binding calculations, in contrast, the range of the Hamiltonian is usually small. The OM method shows the same behavior as the FOE method. The numerical effort increases linearly with respect to the number of nonzero elements per column of the Hamiltonian. In the DC method the numerical effort is independent of the bandwidth, but even in this case the DC method is not expected to outperform the FOE or DMM methods.

- Scaling with respect to the size of the basis set:

Let us now consider the case in which the number of atoms and all other relevant quantities, such as the

size of the localization region, are fixed and where we only increase the number of basis functions per atom m_b. Both the number of columns n and the number of off-diagonal elements per column m of the density matrix will then increase linearly with respect to m_b. We will also assume that the Hamiltonian is a sparse matrix with m_b off diagonal elements per column. In the DMM method the numerical effort will consequently grow cubically with respect to m_b, since the number of operations needed for the multiplication of two sparse matrices of linear dimension n with m off-diagonal elements per column is proportional to nm^2. The DC method scales cubically as well, since it is based on diagonalization. The FOE method likewise scales cubically, since three factors are increasing—the number of columns in the density matrix, the number of coefficients in each column, and the number of off-diagonal elements of the Hamiltonian matrix. Thus, in addition to the arguments showing the unrealistically large memory requirements of these methods when used with large basis sets, we also find a cubic scaling that prohibits the use of these algorithms in this context. In the OM method both the application of the Hamiltonian to the orbitals and the calculation of the overlap between orbitals scale quadratically with respect to m_b. This is due to the facts that both the number of nonzero coefficients c_i^n in the Wannier function expansion and the number of nonzero matrix elements in the Hamiltonian and overlap matrices increase linearly.

The assumption in the above discussion that all other relevant parameters remain constant when increasing m_b is rather unrealistic concerning the eigenvalue spectrum. In general, the highest eigenvalue ϵ_{max} will increase when basis functions are added. In addition to the aforementioned considerations of numerical effort for one step, one also has to take into account that the number of steps will increase due to a worsening of the condition number. Since n_{pl} in the FOE method increases faster [Eq. (69)] than the number of iterations n_{it} [Eq. (88)] in the DMM and OM methods, performance will deteriorate more in the FOE than in the DMM and OM methods. The DC method is insensitive to this effect.

• Memory requirements:

The DMM method requires the storage of the whole sparse density matrix. If one takes advantage of the fact that the matrix is symmetric, storage requirements can actually be cut into half. The OM method requires only the storage of the truncated Wannier orbitals and so the memory requirements are reduced by about 50% in the typical tight-binding context compared to the case in which one stores all the nonzero elements of the density matrix without taking advantage of its symmetry. If the formulation of Kim *et al.* (1995) is used, the gain can come down to

less. In both the DC and FOE methods only the subparts, the columns of the density matrix that are consecutively calculated, need to be stored. The storage requirements are therefore greatly reduced compared to the DMM and OM methods, roughly by a factor of N_{el}.

• Parallel implementation:

Parallel computers and clusters of workstations are standard tools in the high-performance computing environment. The suitability of an algorithm for parallelization is therefore an important consideration. It is of course always possible to ''parallelize'' any program; the question is whether this can be done in a coarse- or fine-grained way, i.e., with a small or large communication-to-computation ratio. Only a coarse-grained parallel program will run efficiently on clusters of workstations with relatively slow communications as well as on a very large number of processors of a massively parallel computer. Both the Fermi operator expansion and the DC algorithms are intrinsically parallel approaches in that they subdivide a large computational problem into smaller subproblems which can be solved practically independently. For the FOE method (Goedecker and Hoisie, 1997) the calculations of the different columns of the density matrix are practically independent [Eq. (49)]. For the DC method the calculations of the different patches of the density matrix are practically independent as well. Both methods can therefore be implemented in a coarse-grained way. A program based on the FOE method won the 1993 Gordon Bell prize in parallel computing for its outstanding performance on a cluster of eight workstations, obtaining half of the peak speed of the whole configuration (Goedecker and Colombo, 1994b). Impressive speedups of up to 400 have been obtained with the FOE method on an 800-processor parallel machine (Kress *et al.*, 1998). Even though it is more difficult to implement the OM method in parallel, two such implementations have been reported. In the first scheme (Canning *et al.*, 1996) one associates with each processor a certain number of localized orbitals. This data structure is optimal for the application of the Hamiltonian to the orbitals, but requires communication for the calculation of the overlap between the orbitals. The second scheme (Itoh *et al.*, 1995) associates with a certain processor the coefficients of all the orbitals whose localization region has an intersection with a certain region of space. This data layout is optimal for the calculation of the overlap matrix, but requires communication for the application of the Hamiltonian. The optimal basis DMM method has also been parallelized (Goringe, Hernandez, *et al.*, 1997). Since the OBDMM is more complex than the other methods that have been implemented on parallel machines, three different parallelization paradigms are required.

• Quality of forces:

For variational (DMM and OM) methods, the force formula is particularly simple [Eq. (90)] since only the Hellmann-Feynman term survives. It has to be stressed, however, that this formula is only exact if one has succeeded in reducing to zero the gradient with respect to all variational quantities. If, in a simulation, the gradient is not reduced to zero within the required precision because too many iterations would be required, errors will creep into the calculated forces, making them inconsistent with the total energy. From this point of view the situation is easier in the FOE method. Since the Fermi operator expansion is not an iterative method (in the sense that one iterates until a certain accuracy criterion is met), the force formula of Voter [Eq. (58)] will always give forces consistent with the total energy. As has already been discussed, no consistent force formula exists for the DC method.

Consistent forces are a prerequisite for the conservation of the total energy in molecular dynamics simulations. Even with consistent forces, however, there are other factors that can cause deviations from perfect total energy conservation in molecular dynamics simulations, such as finite time steps and events where atoms enter or leave the localization region.

• Cases in which the methods become inefficient:

Cases in which different methods become inefficient have already implicitly been pointed out when discussing the previous criteria. Let us finally mention a special case in which the FOE method is inefficient. As discussed above, a small gap usually implies highly extended density matrices, and the FOE method is highly competitive. There can, however, be exceptions to this rule. If by symmetry restrictions there is practically no coupling between two well-localized states that are close together, their energy levels can be split by a very small amount. If the Fermi level falls in between these two levels, a very-high-degree polynomial is needed to separate them into an occupied and an unoccupied level. This scenario can be found in the case of a vacancy in a silicon crystal. A Jahn-Teller distortion leads to a very small splitting between an occupied and an unoccupied gap level. Using a high electronic temperature and a low-degree polynomial in the FOE method does not reproduce this Jahn-Teller distortion. A detailed study of this effect is given by Voter *et al.* (1996), showing that a polynomial of degree 200 is needed instead of the typical degree 50 for bulk silicon in the tight-binding context. From an energetic point of view it is frequently not necessary to track such Jahn-Teller distortions, since they lead to a rather small energy reduction. In molecular dynamics simulations of metallic systems this suppression of the gap opening can even be beneficial (Goedecker and Teter, 1995) since it leads to a smoother density of states around the Fermi level.

In summary, we see that performance depends critically on many parameters, which can change from one application to another. Claims of superior performance based on test runs of a particular system therefore have to be taken with caution.

B. Large basis sets

Whereas the methods that are mainly applicable in the context of small basis sets showed important differences under the various comparison criteria, the behavior of the FOP, OM, and OBDMM methods are quite similar under most of these criteria. The comparison of the methods that are applicable to large basis sets will therefore be based only on a smaller set of important criteria:

• Scaling with respect to the size of the basis set:

As pointed out before, the methods compared in this section all have a reasonable scaling with respect to the size of the basis set, thus allowing their use in the context of very large basis sets. In contrast to the same point within the small basis set framework, the number of nonzero matrix elements of the Hamiltonian is typically independent of the resolution of the grid, i.e., of the number of basis functions. The most important part of the FOP, OM, and optimal basis DMM algorithms, the application of the Hamiltonian matrix to a wave vector, therefore scales linearly.

• Finding a first initial guess:

As discussed in the comparison of small basis sets, it can be difficult to find an initial guess for Wannier-function-based methods. This problem does not exist in the OBDMM method if the number of orbitals is larger than the number of electrons. In this case the orbitals are just basis functions and, by analogy with the usual tight-binding or LCAO basis sets, it should always be possible to generate a physically motivated initial guess for these orbitals.

• Required number of iterations:

As mentioned above, both the OM and the OBDMM methods suffer from ill-conditioning problems and therefore require a frequently excessive number of iterations. No such ill-conditioning exists for the Fermi operator projection method.

• Cases in which the methods become highly inefficient:

None of the three methods has ever been applied to metallic systems, and they are all expected to fail in this case.

V. OTHER O(N) METHODS

The recursion method is a well-established method which also exhibits O(N) scaling. It is principally a method for calculating the density of states $D(\epsilon)$, but once the density of states is known, the band-structure energy can be calculated by integrating $\epsilon D(\epsilon)$ up to the

Fermi level. The recursion method has been described extensively (Haydock, 1980; Gibson *et al.*, 1993) and we shall therefore not review it. Let us just point out that within the original formulation of the recursion method only the diagonal elements of the density matrix could be calculated. For a calculation of the forces one would also need the off-diagonal elements, so the applicability of the recursion method is significantly reduced compared to the O(N) methods described in Sec. III, which all gave access to the forces. There have been several attempts to overcome this limitation (Aoki, 1993; Horsfield, Bratkovsky, Pettifor, and Aoki, 1996, Horsfield, Bratkovsky, Fearn *et al.*, 1996, Bowler *et al.*, 1997). In contrast to the methods of Sec. III, these bond order potential methods are fairly complex and difficult to implement. The basic idea in the bond order potential method is to calculate the off-diagonal elements of the density matrix as the derivatives of diagonal elements of a density matrix defined with respect to a transformed basis. Even though it is now possible to calculate forces within the bond order potential method, they are unfortunately not consistent with the total energy. In another version of the recursion method, the generalized density-of-states method (Horsfield, Bratkovsky, Pettifor, and Aoki, 1996; Horsfield, Bratkovsky, Fearn *et al.*, 1996) the exact forces can be calculated. It is, however, necessary to calculate some generalized moments H^k called interference terms from the recursion coefficients. This inversion is ill conditioned and becomes unstable if one tries to calculate more than 20 moments (Bowler *et al.*, 1997). With such a low number of moments it is not possible to describe many realistic systems (Kress and Voter, 1995), and the error one reaches when the instability sets in is frequently still much too large to be acceptable (Bowler *et al.*, 1997). So recursion-based methods seem not to be a general purpose tool for electronic structure calculations in which accurate energies and forces are required. Bond order potential methods can, however, give insight into basic bonding properties of crystalline solids (Pettifor, 1995). Since these methods, which are related to the recursion algorithm, require the calculation of all the diagonal elements of the density matrix, they are obviously not very efficient if a very large number of basis functions per atom is used, and they were indeed primarily proposed for tight binding or other schemes with a small number of basis functions. The only exception is a version proposed by Baroni and Giannozzi (1995), who suggested using delta functions as a basis in a density-functional-type calculation. With this basis set the diagonal elements of the density matrix are sufficient to determine the charge density, whereas for more general basis functions the off-diagonal elements are needed as well [Eq. (12)]. Because the number of delta functions has to be very large, even in the most favorable case of silicon, the crossover point was estimated to be around 1000 silicon atoms. This method is clearly not competitive with most other methods, in which the crossover point is much lower.

In addition to the recursion method, there are other O(N) methods that calculate the density of states and thus give information about the eigenvalue spectrum of a system. We shall not expand on these methods (Drabold and Sankey, 1993; Wang, 1994; Silver and Roeder, 1997) in this article, since our focus is on methods in which the primary quantity is the total energy. In principle, it would also be possible to derive total energies and forces from the spectrum by an integration up to the chemical potential. Attempts in this direction have, however, not been very successful up to now with the exception of the smeared density of states kernel polynomial method (Voter *et al.*, 1996), which is closely related to the FOE method. A broader discussion of these methods based on the density of states can be found in a review by Ordejon (1998).

Another approach to improving the scaling properties is based on so-called pseudo-diagonalization (Stewart *et al.*, 1982). This method is closely related to the well-known Jacobi method for matrix diagonalization. Whereas in the original Jacobi method rotation transformations are applied until all off-diagonal elements vanish, only the entries that couple occupied and unoccupied states are annihilated in the pseudo-diagonalization method. One obtains thus not the occupied eigenvectors of the Hamiltonian but an arbitrary set of vectors that span the same occupied space. In its original formulation (Stewart *et al.*, 1982) this method still had cubic scaling, but with a smaller prefactor than complete matrix diagonalization. Nearly linear scaling has been reported with this method (Stewart, 1996) if the Hamiltonian matrix is constructed with respect to a set of well-localized orbitals. In this way most of the elements in the block-coupling occupied and unoccupied states are zero at the start of the transformations. The annihilation of certain matrix elements during the rotation steps causes only a controlled spread of other rows and columns in the matrix, so at the end each column and row extends over a region comparable to the localization region in other O(N) methods.

A method proposed by Galli and Parrinello (1992) can be considered as a precursor of the orbital minimization method. The total energy is minimized with respect to a set of localized Wannier functions. In contrast to the orbital minimization method, however, one has the old conventional functional [Eq. (93)], which has to be minimized under the constraint of orthogonality, rather than having the orthogonalization constraint contained in the modified functional [Eq. (92)]. In this scheme it is necessary to calculate the inverse of the overlap matrix between the Wannier functions. From timing considerations this is not a serious drawback, since this part is only a small fraction of the total workload, even for large systems, and even if it is done with a scheme that scales cubically. There are, however, problems of numerical stability. As pointed out by Pandey *et al.* (1995), the overlap matrix becomes close to singular and the introduction of localization constraints can under these circumstances falsify the results.

Also vaguely related to the OM functional are methods in which certain constraints are included by a penalty function. Wang and Teter (1992) included the or-

thogonality constraint in this way, Kohn (1996) the idempotency condition of the zero-temperature density matrix.

O(N) implementations of electronic structure methods based on the multiple-scattering theory have also been reported. In the simplest version (Wang *et al.*, 1995) it is essentially a DC method, with the difference that within each localization region the calculation is done with the multiple-scattering method. A further development was to replace the buffer region by an effective medium (Abrikosov *et al.*, 1996, 1997). This can considerably reduce the prefactor of the calculation, especially in metallic systems where large localization regions are needed. For this class of methods, however, no force formulas have been reported, a deficiency that restricts their applicability.

A scheme that may lead to a reduced scaling behavior has also been proposed by Alavi *et al.* (1994). It is based on a direct representation under the form of a sparse matrix of the Mermin finite-temperature functional (Mermin, 1965), so it allows for a finite temperature, as does the FOE method.

As was already mentioned, the polynomial FOE method becomes inefficient for cases in which one has a large basis set, causing the highest eigenvalue to grow very large. This would necessitate a Chebschev polynomial of very high degree. An elegant method for overcoming this bottleneck within a polynomial method has been proposed by Baer and Head-Gordon (1998a, 1998b). They write the density matrix at a low temperature T as a telescopic expansion of differences of density matrices at higher temperatures Tq^j:

$$F_T = F_{Tq^n} + (F_{Tq^{n-1}} - F_{Tq^n}) + (F_{Tq^{n-2}} - F_{Tq^{n-1}})$$
$$+ \cdots + (F_{Tq^0} - F_{Tq^1}). \tag{101}$$

The factor q ($q>1$) determining the geometric sequence of temperatures is chosen from considerations of numerical convenience. As the temperature is lowered the numerical rank of each difference term becomes smaller and smaller, since the difference of two Fermi distributions of similar temperature is vanishing to within numerical precision over most of the spectrum. Hence it is necessary to find Chebyshev expansions only over successively smaller regions of the spectrum and it is also possible to calculate the traces (which in turn determine all physically relevant quantities) within spaces of smaller and smaller dimension.

Ordejon *et al.* (1995) proposed a method based on the orbital minimization method to calculate phonons from the electronic structure with linear scaling.

VI. SOME FURTHER ASPECTS OF O(N) METHODS

Hierse and Stechel (1994, 1996) examined whether non-orthogonal Wannier-like orbitals are transferable from one chemical environment to another similar one. If this were the case one could reuse a precalculated Wannier orbital as a very good initial guess. Such a property would thus reduce the prefactor of any method

regardless of its scaling behavior. Unfortunately, they could find reasonable transferability only under rather restrictive conditions. When they added CH_2 units to a C_nH_{2n+2} polymer they found good transferability of the orbitals associated with this building block within a density-functional scheme (Hierse and Stechel, 1994). But as soon as they started bending the polymer (Hierse and Stechel, 1996), the transferability was lost in the density-functional scheme. Only in a non-self-consistent Harris functional scheme were some efficiency gains still possible.

Hernandez *et al.* (1997) suggest a solution to the basis set problem in O(N) methods. As mentioned in Sec. III, O(N) techniques are difficult to reconcile with extended basis sets such as plane waves. Plane waves, however, have several important advantages and are therefore widely used in conventional [i.e., not O(N)] electronic structure calculations. One of their main advantages is that the accuracy can easily be improved by increasing one single parameter, namely the minimal wavelength, which corresponds to the resolution in real space. Hernandez *et al.* (1996) have proposed a basis set of "blip" functions which combines both advantages. It is localized and its resolution can systematically be improved. As an alternative to the "blip" functions one could also use finite-difference techniques (Chelikowsky *et al.*, 1994) or wavelets (Goedecker, 1998b; Goedecker and Ivanov, 1998; Lippert *et al.*, 1998; Arias, 1999) since they share the same advantages. As shown by Goedecker and Ivanov (1999), wavelets allow for a highly compact representation of both the density matrix and the Wannier functions, since they are localized in both real and Fourier space.

Much of the work of Ordejon *et al.* (1996) is also based on a new set of basis functions proposed by Sankey and Niklewski (1989). This basis set consists of atomic orbitals that are modified in such a way that they are strictly zero outside a certain spherical support region. This then gives rise to a Hamiltonian matrix that is strictly sparse. By tabulating these matrix elements it is possible to do density-functional calculations whose computational requirements are in between the requirements of traditional density-functional calculations and tight-binding calculations. Obviously one has to find a compromise between accuracy and speed. If the basis functions extend about a larger support region, one has a more accurate basis, but the numerical effort increases because the Hamiltonian is less sparse.

Horsfield and Bratkovsky (1996) have incorporated entropy terms in O(N) methods within the FOE algorithm. As soon as one has systems at nonzero temperatures, these terms should, in principle, be added; however, in most systems their effects are very small at room temperature. For computational convenience, temperatures much larger than room temperature can be used. Wentzcovitch *et al.* (1992) and Weinert and Davenport (1992) showed that the inclusion of entropy gives simplified force formulas, since only the Hellmann-Feynman term survives. The free energy A is given by

$$A = \text{Tr}[FH - k_B TS(F)], \tag{102}$$

where the entropy S is a matrix function of F,

$$S = -[F \ln(F) + (1-F)\ln(1-F)]. \tag{103}$$

Writing S as a Chebyshev polynomial in H and analyzing everything in terms of the eigenfunctions of H, they find that one has to do a Chebyshev fit to a distribution very similar to the Fermi distribution but with some additional features close to the chemical potential. Using a formalism by Gillan *et al.* (1998), they then extrapolate their results to zero temperature, obtaining faster convergence to the zero-temperature limit in this way. Let us stress again that with the FOE method it is possible to build up density matrices corresponding to several temperatures without significant extra cost. A set of generalized Fermi distributions that allow an efficient extrapolation to the zero-temperature limit by eliminating arbitrarily high powers of T has also been proposed by Methfessel and Paxton (1989).

As was mentioned in the Introduction, the fundamental cubic term in an electronic structure calculation based on orbitals comes from the orthogonalization requirement. In traditional pseudopotential calculations based on a Fourier-space formulation there is, however, a second very large cubic term, arising from the nonlocal part of the pseudopotential. This second cubic term can be eliminated by using pseudopotentials that are separable in real space (King-Smith *et al.*, 1991; Goedecker *et al.*, 1996; Hartwigsen *et al.*, 1998).

VII. NON-ORTHOGONAL BASIS SETS

Up to now we have always implicitly assumed that we are dealing with orthogonal basis sets, i.e., that

$$\int \phi_i^*(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r} = \delta_{i,j}. \tag{104}$$

Non-orthogonal basis sets give rise to an overlap matrix S,

$$S_{i,j} = \int \phi_i^*(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r}. \tag{105}$$

An orthogonality relation similar to Eq. (104) can be obtained by introducing the dual basis functions $\tilde{\phi}_i(\mathbf{r})$ given by

$$\tilde{\phi}_i(\mathbf{r}) = \sum_j S_{i,j}^{-1}\phi_j(\mathbf{r}), \tag{106}$$

where S^{-1} is the inverse of the overlap matrix S. It is then easy to verify that

$$\int \tilde{\phi}_i^*(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r} = \delta_{i,j}. \tag{107}$$

As has been mentioned in Sec. III, all realistic atom-centered localized basis sets are non-orthogonal. Within the tight-binding context, there are also many non-orthogonal schemes on the market. There is thus certainly a need to apply O(N) techniques for these

schemes. Most of the basic O(N) algorithms presented previously have been generalized to the non-orthogonal case, and we shall present these generalizations in this section. In the context of a non-orthogonal scheme one has to distinguish carefully between the eigenfunctions Ψ_n and the associated eigenvector $\mathbf{c}^n$ which contains the expansion coefficients c_i^n such that $\Psi_n(\mathbf{r}) = \Sigma_i c_i^n \phi_i(\mathbf{r})$. The eigenvector $\mathbf{c}^n$ satisfies the generalized eigenvalue equation

$$H\mathbf{c}^n = \epsilon_n S\mathbf{c}^n. \tag{108}$$

In the same way one has to distinguish carefully between the density-matrix operator $F(\mathbf{r},\mathbf{r}')$ and the density matrix $F_{i,j}$ itself. While Eq. (22) for the density-matrix operator remains the same,

$$F(\mathbf{r},\mathbf{r}') = \sum_n f(\epsilon_n)\Psi_n^*(\mathbf{r})\Psi_n(\mathbf{r}')$$

$$= \sum_n f(\epsilon_n)\sum_{i,j} c_i^{n*}c_j^n \phi_i^*(\mathbf{r})\phi_j(\mathbf{r}'), \tag{109}$$

the expression for the number of electrons [Eq. (13)] is modified to

$$N_{el} = \text{Tr}[FS]. \tag{110}$$

The expression for the band-structure energy (10), however, remains valid. In the definition of the density matrix $F_{i,j}$ [Eq. (9)] one has to use the dual basis functions instead of the ordinary:

$$F_{i,j} = \int \int \tilde{\phi}_i^*(\mathbf{r})F(\mathbf{r},\mathbf{r}')\tilde{\phi}_j(\mathbf{r}')d\mathbf{r}\,d\mathbf{r}'$$

$$= \sum_n c_i^{n*}c_j^n f(\epsilon_n). \tag{111}$$

This replacement can have important consequences for the locality of the density matrix $F_{i,j}$. If we have a set of localized orthogonal basis functions (the only known set of basis functions with this property are the Daubechies scaling functions), whose extension is less than the oscillation period of the density-matrix operator, then Eq. (9) ensures that $F_{i,j}$ will have the same decay properties as $F(\mathbf{r},\mathbf{r}')$. This does not necessarily hold true for Eq. (111). Even if the basis functions $\phi_i(\mathbf{r})$ are well localized, this is in general not true for their duals $\tilde{\phi}_i(\mathbf{r})$. If the duals have a very slow decay then this slow decay will be inherited by $F_{i,j}$ and it might not be possible to use O(N) algorithms for the calculation of $F_{i,j}$. Problems might, for instance, arise if high-quality Gaussian basis sets containing diffuse functions are used. Preliminary experience indicates that for small basis sets of moderate quality the duals are not so delocalized as to destroy the localization of $F_{i,j}$.

For the DC method the generalization to the non-orthogonal case is trivial. Since it is based on diagonalization within each subvolume, one only has to solve the generalized eigenvalue problem [Eq. (108)] instead of an ordinary one. In the density-functional context, the DC method has actually been used only with non-orthogonal orbitals.

The Fermi operator expansion method using a Chebyshev representation of the density matrix has been generalized in the non-orthogonal case by Stephan and Drabold (1998). It is easy to see that all the central equations of the FOE method remain correct if the Hamiltonian H is replaced by a modified Hamiltonian $\bar{H}$ (that is no longer Hermitian) given by

$$\bar{H} = S^{-1} H. \tag{112}$$

In particular, it remains true that the density matrix is given within arbitrary precision by

$$F \approx \frac{c_0}{2} I + \sum_{j=1}^{n_{pl}} c_j T_j(\bar{H}), \tag{113}$$

if a sufficiently large n_{pl} is used. The problem is how to find $\bar{H}$ efficiently. Even if S is a sparse matrix, the inverse S^{-1} is not sparse in general and $\bar{H}$ would be a full matrix as well, destroying immediately the linear scaling. However, it turns out that the matrix elements of $\bar{H}$ decay faster than the corresponding matrix elements of H (Gibson, 1993; Stephan and Drabold, 1998). One can therefore cut off the tight-binding Hamiltonian $\bar{H}$ at the same distance where one would usually cut off H. In this way all the matrices involved are reduced to sparse matrices and $\bar{H}$ can be constructed by solving the set of linear systems

$$S\bar{H} = H. \tag{114}$$

Since both the right-hand side in H and the solution vectors making up $\bar{H}$ are sparse, different systems of equations are only coupled by sub-blocks of S. So the big matrix inversion problem is decoupled into many small systems of equations and the scaling is therefore strictly linear.

If the FOP method is used in connection with a rational approximation, the generalization to the non-orthogonal case can be done straightforwardly and without any approximation (Goedecker, 1995):

$$F = \sum_{\nu} \frac{w_\nu}{H - z_\nu S}. \tag{115}$$

One has then to solve a system of linear equations, which is a generalization of Eq. (73):

$$(H - z_\nu S) F_\nu = I, \tag{116}$$

$$F = \sum_{\nu} w_\nu F_\nu. \tag{117}$$

A similar approach was adopted by Jayanthis et al. (1998). They formulated their method in terms of the Green's function. However, F_ν in Eq. (116) is a Green's function for a complex energy z_ν and the methods are essentially equivalent.

If the FOE method is used to calculate Wannier functions, the required projection operator F_p is slightly different from the density matrix and it is given by

$$F_p = \sum_{\nu} \frac{w_\nu S}{H - z_\nu S}. \tag{118}$$

The density-matrix minimization method has also been generalized to the non-orthogonal case (Nunes and Vanderbilt, 1994). Introducing a modified density matrix $\bar{F}$ defined as

$$\bar{F} = S^{-1} F S^{-1}, \tag{119}$$

one finds that the DMM functional (78) becomes

$$\Omega = \text{Tr}[(3\bar{F}S\bar{F} - 2\bar{F}S\bar{F}S\bar{F})(H - \mu I)]. \tag{120}$$

This has the advantage that one does not have to invert S if one minimizes directly with respect to $\bar{F}$. The calculation of the gradient of the DMM functional in the non-orthogonal case is a tricky point. The definition of the gradient is not as absolute as one might think. It is the direction of steepest descent per unit change of the variables, and one must therefore define a norm for the multidimensional space of variables before defining the gradient (Vanderbilt, 1998, private communication). Two different gradient expressions have been proposed by Nunes and Vanderbilt (1994) and by C. White et al. (1997), which correspond to two different choices of metric. The gradient of White et al. (1997) requires fewer minimization steps (Gillan et al., 1998), but each step is more expensive, since it requires the calculation of the inverse of the overlap matrix. From the point of view of overall numerical efficiency it is therefore not clear which gradient expression is more efficient.

Another generalization (Daniels et al., 1997, Millam and Scuseria, 1997) of the DMM method, which is similar in spirit to Stephan's generalization (Stephan and Drabold, 1998) of the FOE method, consists in first performing a transformation to an orthogonal basis set by finding the LU decomposition of the overlap matrix

$$S = U^T U, \tag{121}$$

where U is an upper triangular matrix. If, in addition, U is approximated as a sparse matrix with m off-diagonal elements, then Eq. (121) can be solved with a scaling proportional to nm^2, where n is the dimension of the matrices involved. Thus the scaling with respect to the size of the system is linear, as it should be.

The orbital minimization method can easily be generalized to the non-orthogonal case (Ordejon et al., 1996). The orbital overlap $\sum_l c_l^n c_l^m$ in the functional [Eq. (92)] has to be generalized to $\sum_{l,k} c_l^n S_{l,k} c_k^m$.

VIII. THE CALCULATION OF THE SELF-CONSISTENT POTENTIAL

We shall now discuss an issue that is relevant only in self-consistent electronic structure calculations, namely, the calculation of the potential arising from the electronic charge. This potential consists essentially of two parts, the electrostatic or Hartree potential and the exchange-correlation potential.

A. The electrostatic potential

The solution of Poisson's equation to find the electrostatic potential arising from a charge distribution ρ is a basic problem found in many branches of physics. Solution techniques are described in a wide variety of books and articles. We shall therefore only point out a few features that are important in the special context of O(N) electronic structure calculations.

If one is dealing with an electronic charge density that has only one length scale, Poisson's equation can be solved efficiently and with a scaling that is close to linear. Charge densities of this type are encountered in the context of pseudopotential calculations where one has eliminated the core electrons and where the characteristic length scale is the typical extension of an atomic valence electron. One can, for instance, use plane-wave or multigrid techniques (Briggs, 1987), which both have a scaling proportional to $n \log(n)$, where n is the number of grid points.

The situation becomes problematic when core electrons are included. In this case one could, in principle, still use the above-mentioned techniques with an increased resolution. One would still have linear scaling, but the prefactor would be so large as to make it completely impractical in terms of both timing and memory. Exactly the same arguments apply to the representation of the wave functions, and for this reason ordinary plane waves are not used for all-electron calculations.

A widely-used basis set for all-electron calculations are Gaussian-type basis sets (Boys, 1950). By varying the width of the Gaussians one can describe both core and valence electrons. The popularity of Gaussian-type basis functions comes from the fact that many important operations can be done analytically (Obara and Saika, 1986). One property that we shall use is that the product of two Gaussians is again a Gaussian, centered in between the two original Gaussian-type functions. The matrix elements of the electrostatic potential part of the Hamiltonian with respect to a set of Gaussian orbitals $g_i(\mathbf{r})$, $i = 1, \ldots, M_b$ are given by

$$H_{i,j} = \int d\mathbf{r}\, g_i(\mathbf{r}) \left(\int d\mathbf{r}' \sum_{k,l} \frac{F_{k,l} g_k(\mathbf{r}') g_l(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right) g_j(\mathbf{r}). \tag{122}$$

The elementary integral

$$\int d\mathbf{r} \int d\mathbf{r}' \frac{g_i(\mathbf{r}) g_j(\mathbf{r}) g_k(\mathbf{r}') g_l(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{123}$$

can also be calculated analytically (Obara and Saika, 1986). A straightforward evaluation of Eq. (122) would then result in a quartic scaling. There are, however, many well-known techniques (Challacombe et al., 1995) to reduce this scaling. The most obvious trick comes from the observation that there is a negligible contribution to the charge density ρ if the Gaussians g_l and g_k are centered far apart. Consequently the charge density is not a sum over M_b^2 product Gaussians G_m ($G_m = g_i g_j$), but only over M_a such Gaussians,

$$\rho(\mathbf{r}') = \sum_{k=1}^{M_b} \sum_{k=1}^{M_b} F_{k,l} g_k(\mathbf{r}') g_l(\mathbf{r}') \approx \sum_{m=1}^{M_a} c_m G_m(\mathbf{r}'). \tag{124}$$

The size of the auxiliary basis set M_a is proportional to M_b with a large prefactor, which depends on the ratio of the largest to the smallest extension of the Gaussians as well as on the accuracy target. In a similar way matrix elements $H_{i,j}$ become negligible if the basis functions g_i and g_j are centered very far apart. Using these two approximations, one obtains a quadratic scheme with a very large prefactor. An approximate quadratic scaling is also found in numerical tests (Strout and Scuseria, 1995).

Another widely used method consists in fitting the charge density by a set of M_a auxiliary Gaussians G_i. Even though we use the same symbols (M_a, G_i) as above, they now denote somewhat different quantities. The allowed number of auxiliary Gaussians M_a is now much smaller, namely, a few times M_b. The auxiliary functions themselves are no longer taken to be all the possible product functions, but are determined by empirical rules in such a way as to give the best possible fit in Eq. (124). The fitting involves the solution of an ill-conditioned system of linear equations and therefore has cubic scaling, but with a small prefactor. The evaluation of the matrix elements using this representation of the charge density then has quadratic scaling, if one again neglects small elements.

To obtain an even better scaling behavior requires the introduction of completely new concepts. One possibility is to build upon algorithms that solve the classical Coulomb problem for point particles. The classical Coulomb problem requires the calculation of the electrostatic potential arising from all the N particles with charge Z_j at all the positions $\mathbf{r}_i$

$$U(\mathbf{r}_i) = \sum_j \frac{Z_j}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{125}$$

By grouping particles into hierarchical groups and by describing their potential far away from such groups in a controlled, approximate way by multipoles, these fast algorithms allow us to evaluate Eq. (125) with linear instead of quadratic scaling. There are now several proposals (C. White et al., 1994; Strain et al., 1996; Perez-Jorda and Yang, 1997) for modifying one of these fast algorithms, the fast multipole method (Greengard, 1994) in such a way that it can also handle the continuous charge distributions arising in the context of electronic structure calculations. The basic idea is fairly straightforward. As we saw, the charge distribution is always given as a weighted sum of auxiliary Gaussians [Eq. (124)]. Now the electrostatic potential of such a Gaussian particle looks the same from a distance as the potential of a point particle. Concerning the far field, one can thus essentially take over the existing algorithms. To account for the nonpointlike nature of the Gaussian particles one has, however, to correct the near field. Since the calculation of these local corrections has linear scaling, the whole procedure has linear scaling as well. There are

two problems with this kind of approach. First, one has only linear scaling with respect to the size of the basis set if the volume covered by the basis set grows at least as fast as the size of the basis set. If one adds, for example, basis functions for a molecule of fixed size, to improve the accuracy of the basis set, one no longer has linear scaling. This is related to the fact that one can now apply the fast far-field treatment to a smaller number of Gaussian particle interactions. This behavior can be easily understood by considering the extreme case in which all Gaussian particles are centered very close together within a radius that is smaller than their width. In this case one, evidently, cannot use any more any far-field techniques. The second problem is closely akin to the first. If one adds extended Gaussians to the system, the efficiency deteriorates.

A method that scales strictly linearly with respect to the size of the basis set, independently of whether the volume is increased at the same time or not, and which can be applied within the context of any basis set, is based on wavelets (Goedecker and Ivanov, 1998a). As the input to this method, the charge density is needed on a real-space grid, which can have varying resolution. Thus near the core region of the atoms in a molecule the resolution can be arbitrarily increased. Using interpolating wavelets, one can map this charge density uniquely to a wavelet expansion, since a wavelet expansion can compactly describe nonuniform functions. In the wavelet basis one can then iteratively solve Poisson's equation,

$$\nabla^2 V = -4\pi\rho. \tag{126}$$

The matrix representing the Laplace operator ∇^2 is sparse and the matrix-times-vector multiplications needed for the iterative solution of Eq. (126) scale linearly. Using a preconditioning scheme in a basis of lifted wavelets, one finds that the condition number is independent of the size and of the maximal resolution of the wavelet expansion; the number of iterations is therefore constant. Thus one obtains an overall linear scaling.

B. The exchange correlation potential

Within the most popular versions of density-functional theory the exchange-correlation potential is a purely local function. For the local-density approximation (Parr and Yang, 1989) the exchange-correlation potential at a certain point depends only on the density at that point; for generalized gradient approximations (Becke 1988; Lee $et\ al.$, 1988; Perdew $et\ al.$, 1996) it also depends on the gradient of the density at that point. If one uses real-space methods such as finite differences or finite elements as well as plane-wave methods for which the calculation of the exchange-correlation potential is done on a real-space grid, as well, it is obvious that the numerical effort is linear with respect to the system size. If one uses more extended basis functions, such as Gaussian-type orbitals, it becomes more difficult to achieve linear scaling (Stratmann $et\ al.$, 1996).

For Hartree-Fock calculations the exchange energy

$$\sum_{i,j} \int \int d\mathbf{r}\, d\mathbf{r}' \frac{\Psi_i(\mathbf{r})\Psi_j(\mathbf{r})\Psi_i(\mathbf{r}')\Psi_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{127}$$

seems to be as nonlocal as the Coulomb potential. Using Eq. (23) one can, however, rewrite Eq. (127) to obtain

$$\int \int d\mathbf{r}\, d\mathbf{r}' \frac{F(\mathbf{r},\mathbf{r}')F(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{128}$$

showing that the exchange energy in an insulator is indeed a local quantity whose locality is determined by the decay properties of the density matrix. A linear method for evaluating exchange terms within Gaussian-type electronic structure calculations, based on the aforementioned locality properties, has been devised by Schwegler and Challacombe (1996). An alternative method based on the fast multipole method has been developed by Burant $et\ al.$ (1996).

IX. OBTAINING SELF-CONSISTENCY

To do a self-consistent electronic structure calculation, one needs to blend two ingredients. The first is the calculation of the density matrix in a fixed external potential, the main focus of this article. The second is the calculation of the potential from a given electronic charge density, which was discussed in the preceding Section (VIII). Even if both of these basic parts exhibit linear scaling, it is, however, not yet granted that one has overall linear scaling. It might happen that the number of times one has to repeat these two basic parts increases with the size of the system.

The easiest scheme for combining the calculation of the density matrix and the calculation of the potential is the so-called linear mixing scheme. Given an input charge density ρ_{in}, which determines the potential, one obtains [after the calculation of the density matrix for this potential via Eq. (12)] a new output density ρ_{out}. The new input density ρ_{in}^{new} is not the output density ρ_{out}, but rather a linear combination of the old input density and the output density

$$\rho_{in}^{new}(\mathbf{r}) = \rho_{in}(\mathbf{r}) + \alpha[\rho_{out}(\mathbf{r}) - \rho_{in}(\mathbf{r})]. \tag{129}$$

Here α is the mixing parameter. Overall linear scaling is endangered if one has to decrease α for reasons of numerical stability as the system becomes bigger and if one consequently needs a larger number of iterations.

The standard theory of mixing (Ho $et\ al.$, 1982; Dederichs and Zeller, 1983) is based on the dielectric response function in $\mathbf{k}$ space. Within this theory numerical instabilities arise if the dielectric response functions tend to infinity as k tends to zero. This happens in a metal but not in an insulator, where the dielectric response function always remains finite. Following the general philosophy of this paper to remain within a real-space formalism, we shall elucidate mixing from this real space perspective. The final conclusions are, of course, the same as those based on the Fourier-space theory.

Let us first consider a metal. We assume that we are doing a calculation for a one-dimensional metallic struc-

ture of length L. Let us also assume that due to deviations from the converged self-consistent charge density we transfer an incremental charge ΔQ_{in} from one end of the sample to the other. Using the basic formula for the potential in a capacitor, we get a constant electric field in the sample, giving rise to a potential difference of $\Delta U = L \Delta Q_{in}$ between the two ends. In a metal this potential difference will most likely be larger than the HOMO-LUMO separation (which vanishes for large systems), and we get a large charge transfer ΔQ_{out}. This charge transfer is related to the density of states at the Fermi level, $D(\mu)$, which in our one-dimensional case is the number of states per length unit and per energy unit. Hence the total charge transfer ΔQ_{out} is given by

$$\Delta Q_{out} \approx D(\mu) L \Delta U = D(\mu) L^2 \Delta Q_{in}. \tag{130}$$

If this induced charge transfer ΔQ_{out} is larger than the initial transfer ΔQ_{in}, then the charge transfer will exponentially increase in subsequent iterations and we have the numerical instability called "charge sloshing." To avoid this the mixing factor α has to be proportional to $1/[D(\mu) L^2]$. Doing the same analysis in a three-dimensional structure all the lateral dimensions cancel and we get the same result concerning α. Denoting the volume of our sample by v we find that α is proportional to $v^{-2/3}$. So α has to be decreased with increasing volume, and the number of iterations in the mixing scheme increases with increasing system size. Fortunately, and contrary to the implications of Annett (1995), this charge sloshing can be eliminated by state-of-the-art techniques (Kresse, 1996). One possibility (Kerker, 1981) is just to do the mixing in Fourier space and to have a k-dependent mixing parameter $\alpha(k) = \alpha_0 [k^2/(k^2 + k_0^2)]$:

$$\rho_{in}^{new}(\mathbf{k}) = \rho_{in}(\mathbf{k}) + \alpha_0 \frac{k^2}{k^2 + k_0^2} [\rho_{out}(\mathbf{k}) - \rho_{in}(\mathbf{k})]. \tag{131}$$

As we can see, long-wavelength components (corresponding to small k values) are now strongly damped by

$$\alpha_0 k^2 = \alpha_0 \left(\frac{2\pi}{\lambda} \right)^2 \tag{132}$$

and the dampening has the correct dimensional behavior with respect to the wavelength λ, which corresponds to the length L in our dimensional analysis above. Short-wavelength contributions are just damped by α_0, and this constant dampening sets in when k becomes comparable in magnitude to k_0. We know that for wavelengths of the order of the interatomic spacing a mixing parameter somewhat smaller than 1 works well, and so we can determine by these conditions the values of α_0 and k_0.

Let us next examine whether we can have charge sloshing in an insulator. We shall assume that the potential difference across the sample is not larger than the gap, in which case the discussion for the metallic case would apply instead. Again we consider a sample of length L. According to the modern theory of polarization in solids (King-Smith and Vanderbilt, 1993) a polar-

ization arises because the centers of the Wannier functions are displaced under the action of an electric field. Since the Wannier functions are exponentially localized, the charge that will build up at the two surfaces of our sample is mainly due to displacements of the Wannier function in the elementary cells of the crystal right on the surface, and the charge ΔQ_{out} is thus practically independent of the length of the sample. So the optimal mixing constant α is nearly independent of the size of the system and the number of iterations as well.

In conclusion, we see that linear scaling can also be obtained in the self-consistent case and that, even in a metal, charge sloshing problems can be overcome.

Mixing is the natural choice if the divide-and-conquer or the FOE methods are used in a self-consistent calculation. If methods based on minimization (density-matrix and orbital) are used, one can alternatively also obtain the ground state by a single minimization loop (Car and Parrinello, 1985; Štich *et al.*, 1989; Teter *et al.*, 1989; Payne *et al.*, 1992) without distinguishing between density-matrix optimization cycles and potential mixing cycles. As is not surprising after our discussion of mixing, one finds (Annett, 1995) that in an insulator the number of iterations does not depend upon whether one has a self-consistent type of calculation, where the potential is varying during each minimization step, or whether one has a fixed potential. In other words there is no charge sloshing. In metallic systems it is essential to have finite electronic temperature (Weinert and Davenport, 1992; Wentzcovitch *et al.*, 1992; Kresse, 1996), and therefore the minimization schemes cannot be applied straightforwardly in any case. Thus Annett's analysis (1995) showing that for metals the scaling is at least proportional to $N_{at}^{4/3}$ is irrelevant.

A completely different approach to the mixing problem has recently been proposed by Gonze (1996). He calculates the gradient of the total energy with respect to the potential. His gradient expression does not depend on the wave functions and could thus well be combined with O(N) schemes.

X. APPLICATIONS OF O(N) METHODS

This section is not intended to be a comprehensive review of all the applications to date using O(N) methods. It is rather intended as an illustration of the wide range of areas where O(N) methods have made possible the study of systems that were too big to be studied with conventional methods. In general, one can say that most large-scale tight-binding studies nowadays are done using O(N) methods. Systems comprising from a few hundred up to many thousands of atoms are typically studied. Treating such a large number of atoms with O(N) density-functional methods is much more difficult. For density-functional calculations, the benchmarking and verification aspect has usually dominated, whereas in tight-binding calculations the focus has been, in most cases, on solving challenging physical problems.

Questions concerning extended defects in crystalline materials were one of the main focuses of these tight-
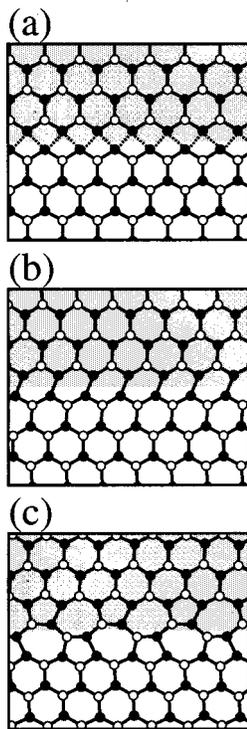
FIG. 16. (a) Symmetric reconstruction of the 90° partial dislocation in silicon. Shaded area indicates stacking fault; (b) the single-period symmetry-breaking reconstruction; (c) the double-period symmetry-breaking reconstruction. The figure is reproduced with the kind permission of the authors, from Nunes *et al.* (1998).

binding studies. Because several good tight-binding parameters are available for silicon, most studies were done on this material. The 90° partial dislocation in silicon was the subject of a series of tight-binding studies. The three structures that were examined are shown in Fig. 16. The energy difference between the structures (a) and (b) in Fig. 16 was studied both by Nunes *et al.* (1996) and by Hansen *et al.* (1995). Even though they used different tight-binding parameters and different O(N) algorithms (DMM and FOE), they both obtained exactly the same energy difference of 0.18 eV/Å in favor of structure (b). Later Benetto *et al.* (1997) discovered a new structure (c) that is even lower in energy. To validate their tight-binding results they did conventional density-functional calculations for smaller subsystems and found perfect agreement with the O(N) tight-binding results. This new structure is experimentally difficult to distinguish from structure (b), and so this result is a convincing illustration of the power of these new O(N) algorithms in materials science. All of these tight-binding calculations necessitated electronic structure calculations involving a few hundred atoms and would therefore have been unfeasible with standard algorithms.

Extended {311} defects in silicon systems containing more than 1000 atoms and their relation to point defects were studied by Kim *et al.* (1997) using the orbital minimization method in the improved version of Kim *et al.* (1995). An understanding of these processes is important for the fabrication of semiconductor devices, since defects have a strong influence on the diffusion properties of semiconductors. Unfortunately, the more realistic questions involving boron dopant atoms in addition to the bulk silicon atoms cannot be treated with current tight-binding models.

Ismail-Beigi and Arias (1998) examined the surface reconstruction properties of silicon nanobars, finding that the influence of edges in these small structures is strong enough to lead to surface reconstructions that are different from those in bulk silicon. They also both employed traditional density-functional calculations and O(N) FOE tight-binding calculations and found good agreement between both for small subsystems that are accessible to both approaches.

Roberts and Clancy (1998) simulated vacancy and interstitial diffusion processes in silicon using the FOE tight-binding method. The diffusion constants they obtained are in good agreement with similar calculations based on classical force fields and density-functional calculations. Compared to the density-functional calculations, they could also significantly enlarge both the number of atoms (216) and the simulation times. The diffusion constant predicted by all these simulations is, however, orders of magnitude larger than the experimental one, a fact for which no explanation is known until now.

Besides silicon there is another material for which several good tight-binding schemes are available, namely, carbon. Fullerene systems are therefore another focus of tight-binding studies.

Galli and Mauri (1994) did a molecular crash test of C_{60} fullerenes colliding with a diamond surface using the OM method. They found three different impact energy regimes, in which the impinging fullerenes (i) survived the collision undamaged, (ii) survived slightly damaged, or (iii) were completely destroyed. Even though the interaction region between the impinging fullerene and the surface did not comprise a very large number of atoms, their computational box contained more than 1000 atoms. The box had to be this large so that the phonons emitted during the collision would not be reflected back from the walls of the box during the time scale of the collision. This reflection of phonons is also a serious problem in classical force-field simulations of crack propagation, and for this reason systems comprising several million atoms are sometimes necessary (Zhou *et al.*, 1997). In the case of this molecular crash test most of the carbon atoms were propagating the phonons. Phonons are well described by classical force fields, and one could use this scheme for the majority of the atoms, while it would be necessary to use the more expensive tight-binding scheme only for the atoms in the collision region. Unfortunately such schemes, combining molecular methods of different speed and accuracy, have not yet been developed and thus a feature of many O(N) calculations is their overkill in a certain sense, treating a large number of essentially inactive atoms
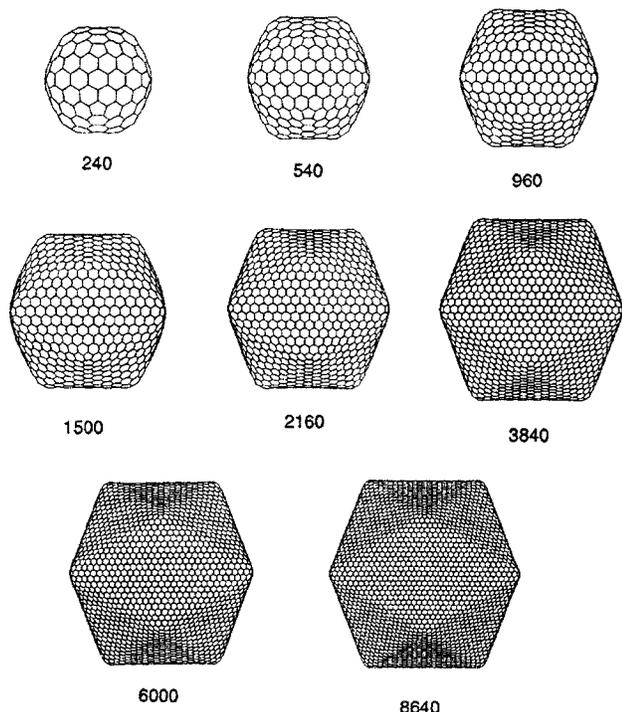
FIG. 17. Tight-binding equilibrium geometries of selected icosahedral fullerenes viewed along the C_2 symmetry axis. This figure is reproduced with the kind permission of the authors, from Xu and Scuseria (1996; reprinted from Chem. Phys. Lett. **262**, 219, with permission from Elsevier Science).

with highly accurate methods. Canning *et al.* (1997) examined thin films of C_{28} fullerenes with the same method, finding that thin superatom films can be formed.

The equilibrium geometries of large fullerenes such as C_{240} were also studied by several groups with O(N) techniques. The central question here is whether such large fullerenes have a spherical form or a polyhedrally faceted shape, where nearly flat polyhedral regions alternate with edges where the curvature is concentrated. York *et al.* (1996) used the original formulation of the DC method in terms of densities to do density-functional calculations of C_{240} and found spherical shapes. Itoh *et al.* (1995), on the other hand, using both empirical and *ab initio* tight binding in the context of the orbital minimization method, found polyhedral shapes. This result is also supported by Xu and Scuseria (1996), who investigated fullerenes up to C_{8640} using the DMM method. The optimized geometries they found for various large fullerenes are shown in Fig. 17.

Ajayan *et al.* (1998) used tight-binding FOE molecular dynamics to simulate irradiation-mediated knockout of carbon atoms from carbon nanotubes. In agreement with experimental observations they found that this atom removal leads to a shrinking of the diameter of the nanotube, but leaves the nanotube essentially intact until the diameter is practically zero. They could identify in their virtual 400-atom sample processes on the atomic level that are responsible for the rapid healing of the defects created by the removal of atoms.

Recently developed tight-binding parameters (Horsfield, Godwin, *et al.*, 1996) have also made it possible to study composite system, namely, hydrocarbons. This set of tight-binding parameters includes self-consistency by imposing a local charge neutrality requirement. Local charge neutrality is essential if different phases are studied because it prevents any unphysically large charge transfer between regions containing different phases arising from different chemical potentials in different phases. Using this new tight-binding scheme, Kress *et al.* (1998) studied the dissociation of CH_4 under high pressure using FOE molecular dynamics. Previous density-functional-based molecular dynamics studies by Ancilotto *et al.* (1997) were limited to very small system sizes of 16 CH_4 molecules and short simulation times of 2 ps. At variance with experimental findings, these density-functional simulations could not find a phase separation of methane into hydrogen and carbon. FOE molecular dynamics allowed the treatment of much larger systems of 128 molecules and much longer simulation times of 8 ps. After 4 ps a phase separation was indeed observed.

York *et al.* (1996) used the DC method in the context of the semiempirical AM1 method (Devar *et al.*, 1985) to calculate heats of formation, solvation free energies, and densities of states for protein and DNA systems containing up to 2700 atoms.

Daniels and Scuseria (1998) reported AM1 semiempirical benchmark calculations for up to 20 000 atoms using DMM, FOE, and pseudo-diagonalization methods.

Sanchez-Portal *et al.* (1997) compared the experimental x-ray structure of a large DNA molecule comprising 650 atoms with the geometric structure obtained from a density-functional-based orbital minimization relaxation. They obtained a root-mean-square deviation from the experimental geometry of 0.23 Å. This was the first attempt to apply O(N) algorithms within density-functional schemes to realistic problems. Their method relies, however, on fairly drastic approximations, resulting in errors that are by far larger than the error one generally expects from a density-functional calculation. A similar study of a large biomolecule is reported by Lewis *et al.* (1997).

Applications of O(N) methods within density-functional theory that use basis sets large enough so that basis set errors do not dominate the density-functional error do not exist at present. However, benchmark calculations exploring the feasibility of such techniques have been published. Goringe *et al.* (1997) evaluated the performance of such schemes in calculations of a cell containing 6000 silicon atoms.

With the advent of faster computers and improved algorithms, the situation concerning sufficiently accurate density-functional O(N) schemes will certainly soon change. It is also interesting to note in this context that the 1998 version of the very popular Gaussian software package will contain O(N) algorithms.

Let us finally come back to a point briefly mentioned in the Introduction. The development of O(N) methods has also deepened our understanding of locality in

quantum-mechanical systems and has thereby fostered the development of theories based on a local picture rather than the conventional nonlocal Bloch-function picture. An example is the theory of polarization in crystalline materials by King-Smith and Vanderbilt (1993). Their theory is based on a local picture in terms of Wannier functions and allows for an intuitive understanding of these phenomena which were difficult to understand before.

XI. CONCLUSIONS

O(N) methods have become an essential part of most large-scale atomistic simulations based on either tight-binding or semiempirical methods. The physical foundations of O(N) methods are well understood. They are related to the decay properties of the density matrix. The use of O(N) methods within density-functional methods is not yet widespread. All the algorithms that would allow us to treat very large basis sets within density-functional theory have certain shortcomings. The OM and OBDMM methods suffer from ill-conditioning problems, and in both the OM and the FOP method detailed knowledge about the bonding properties is required to form the input guess. Thus some algorithmic progress will probably be necessary before these obstacles can be overcome. It is also unclear what the localization properties of very large complicated molecules are and whether perhaps a quadratic scaling rather than a linear scaling is the optimum one in certain cases. It is clear that the elimination of the cubic scaling bottleneck is a very important achievement and that it will pave the way for calculations of unprecedented size in the future. Such calculations will not only be beneficial to physics, but they will also nourish progress in many related fields such as chemistry, materials science, and biology. Even with O(N) algorithms it will not be possible in the foreseeable future to treat systems containing millions of atoms at a highly accurate density-functional level using large basis sets, as would be necessary for certain materials science applications. Such problems can only be approached if one succeeds in combining methods of different accuracy such as density-functional methods with classical force fields, applying the high-accuracy method only to regions where the low-accuracy method is expected to fail. Hybrid methods of this type will certainly be based on the same notions of locality as O(N) methods and will employ similar techniques.

ACKNOWLEDGMENTS

APPENDIX: DECAY PROPERTIES OF FOURIER TRANSFORMS

The density matrix in a periodic solid is defined in terms of a Fourier transform given by Eq. (26). The decay properties of the density matrix are thereby closely related to the decay properties of Fourier transforms. All the properties described in this paragraph are well known; for completeness we shall briefly outline them.

For simplicity let us consider a one-dimensional Fourier transform

$$g(r) = \int_{-\infty}^{\infty} e^{ikr} g(k) dk, \tag{A1}$$

where $g(k)$ is an integrable function tending rapidly to 0 for k tending to $\pm\infty$.

For any function $g(k)$ of this type, $g(r)$ will obviously tend to zero when r tends to infinity. In this case e^{ikr} is a very rapidly oscillating function. The product $e^{ikr} g(k)$ will therefore change sign very rapidly and thus the integral will tend to zero. The exact decay properties depend on how many derivatives are continuous. Let us consider first a function that is piecewise constant and has only a finite number of discontinuities. A function that falls into this class is the function $g(k)$, which is 1 in the interval $[-1:1]$ and zero everywhere else. Calculating the Fourier transform one finds $g(r) = 2\{[\sin(r)]/r\}$. Since any piecewise constant function $g(k)$ can be written as a linear combination of the above prototype function, its transform $g(r)$ will always decay like $1/r$. Using integration by parts we see that

$$r^l g(r) = \int_{-\infty}^{\infty} g(k) i^{-l} \left(\frac{\partial}{\partial k}\right)^l e^{ikr} dk$$

$$= (-i)^{-l} \int_{-\infty}^{\infty} e^{ikr} \left(\frac{\partial}{\partial k}\right)^l g(k) dk. \tag{A2}$$

If the lth derivative is integrable, then the integral will vanish for the reasons discussed above. So if we can do l integrations by parts, each transformation will accelerate the decay by one inverse power of r, and we can do such a transformation whenever our function has at least continuous first derivatives. Hence we arrive at the rule that if l derivatives of $g(k)$ are continuous, $g(r)$ will decay like $r^{-(l+1)}$.

If we have a function $g(k)$ that is analytic, i.e., one for which an infinite number of derivatives exists, then the transform will decay faster than any power of r. One then says that it decays exponentially instead of algebraically. This notion of exponential decay does not necessarily mean that it decays strictly like an exponential function. As an example we could take $g(k) = \exp(-k^2)$, where we know that the transform is again a Gaussian and thus decays faster than an ordinary exponential function. The rate of decay will be related to

the smallest length scale of $g(k)$. If the smallest length scale of $g(k)$ is k_{min}, then $g(r)$ will roughly decay like $\exp(-c|r|k_{min})$, where c is a constant of the order of 1. This follows from the fact that one will have an important cancellation of terms of opposite sign in the integral in Eq. (A1) only if several oscillations occur within the interval k_{min}.

Another qualitative feature of the Fourier transform is that it will have oscillations whenever $g(k)$ is shifted off center. The oscillation period is determined by this shift. As an example let us look at the Fourier transform of a shifted Gaussian $g(k) = \exp[-\frac{1}{2}(k-a)^2]$. The result is $g(r) = \sqrt{2\pi} \exp(iar) \exp(-\frac{1}{2}r^2)$, which is the transform of the unshifted Gaussian times an oscillatory term.

REFERENCES

Abrikosov, I., A. Niklasson, S. Simak, B. Johansson, A. Ruban, and H. Skriver, 1996, Phys. Rev. Lett. **76**, 4203.

Abrikosov, I., S. Simak, B. Johansson, A. Ruban, and H. Skriver, 1997, Phys. Rev. B **56**, 9319.

Ajayan, P., V. Ravikumar, and C. Charlier, 1998, Phys. Rev. Lett. **81**, 1437.

Alavi, A., J. Kohanoff, M. Parrinello, and D. Frenkel, 1994, Phys. Rev. Lett. **73**, 2599.

Ancilotto, F., G. Chiarotti, S. Scandolo, and E. Tosatti, 1997, Science **275**, 1288.

Annett, J., 1995, Comput. Mater. Sci. **4**, 23.

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

Arias, T., 1999, Rev. Mod. Phys. **71**, 267.

Ashcroft, N., and N. D. Mermin, 1976, *Solid State Physics* (Saunders, Philadelphia).

Ayala, P. Y., and G. E. Scuseria, 1999, J. Chem. Phys. **110**, 3660.

Baer, R., and M. Head-Gordon, 1997a, Phys. Rev. Lett. **79**, 3962.

Baer, R., and M. Head-Gordon, 1997b, J. Chem. Phys. **107**, 10003.

Baer, R., and M. Head-Gordon, 1998a, J. Chem. Phys. **109**, 10159.

Baer, R., and M. Head-Gordon, 1998b, Phys. Rev. B **58**, 15296.

Baroni, S., and P. Giannozzi, 1992, Europhys. Lett. **17**, 547.

Bates, K. R., A. D. Daniels, and G. E. Scuseria, 1998, J. Chem. Phys. **109**, 3308.

Becke, A., 1988, Phys. Rev. A **38**, 3098.

Bennetto, J., R. W. Nunes, and David Vanderbilt, 1997, Phys. Rev. Lett. **79**, 245.

Blount, E., 1962, in *Solid State Physics*, edited by F. Seitz and C. Turnbull (Academic, New York), Vol. 13, p. 305.

Bowler, D., M. Aoki, C. Goringe, A. Horsfield, and D. Pettifor, 1997, Modell. Simul. Mater. Sci. Eng. **5**, 199.

Bowler, D., and M. Gillan, 1998, Comput. Phys. Commun. **112**, 103.

Boys, S., 1950, Proc. R. Soc. London, Ser. A **200**, 542.

Briggs, W., 1987, *A Multigrid Tutorial* (Society for Industrial and Applied Mathematics, Philadelphia).

Burant, J., G. Scuseria, and M. Frisch, 1996, J. Phys. Chem. **105**, 8969.

Canning, A., G. Galli, and J. Kim, 1997, Phys. Rev. Lett. **78**, 4442.

Canning, A., G. Galli, F. Mauri, A. de Vita, and R. Car, 1996, Comput. Phys. Commun. **94**, 89.

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

Challacombe, M., E. Schwegler, and J. Almlöf, 1995, University of Minnesota Supercomputer Institute Research Report No. UMSI 95/186.

Challacombe, M., 1999, J. Chem. Phys. **110**, 2332.

Chalvet, O., R. Daudel, S. Diner, and J.-P. Malrieu, Eds., 1976, *Localization and Delocalization in Quantum Chemistry, Vols. I–IV* (Reidel, Dordrecht).

Chelikowsky, J., N. Troullier, and Y. Saad, 1994, Phys. Rev. Lett. **72**, 1240.

Daniels, A., J. Millam, and G. E. Scuseria, 1997, J. Chem. Phys. **107**, 425.

Daniels, A., and G. E. Scuseria, 1998, J. Chem. Phys. **110**, 1321.

Daw, M., 1993, Phys. Rev. B **47**, 10895.

Dederichs, P., and R. Zeller, 1983, Phys. Rev. B **28**, 5462.

des Cloizeaux, J., 1964, Phys. Rev. **135**, A685 and A698.

Devar, M., E. Zoebisch, E. Healy, and J. Stewart, 1985, J. Am. Chem. Soc. **107**, 3902.

Drabold, D., and O. Sankey, 1993, Phys. Rev. Lett. **70**, 3631.

Dickson, R., and A. Becke, 1993, J. Chem. Phys. **99**, 3898.

Feynman, R., 1939, Phys. Rev. **56**, 340.

Fernandez, P., A. Dal Corso, A. Baldereschi, and F. Mauri, 1997, Phys. Rev. B **55**, R1909.

Fulde, P., 1995, *Electron Correlations in Molecules and Solids*, Springer Series in Solid State Sciences No. 100 (Springer, New York).

Gagel, F., 1998, J. Comput. Phys. **139**, 399.

Galli, G., 1996, Curr. Opin. Solid State Mater. Sci. **1**, 864.

Galli, G., and F. Mauri, 1994, Phys. Rev. Lett. **73**, 3471.

Galli, G., and M. Parrinello, 1992, Phys. Rev. Lett. **69**, 3547.

Gibson, A., R. Haydock, and J. P. LaFemina, 1993, Phys. Rev. B **47**, 9229.

Gillan, M., 1989, J. Phys.: Condens. Matter **1**, 689.

Gillan, M., D. Bowler, C. Goringe, and E. Hernández, 1998, in *Proceedings of the Symposium on Complex Liquids*, 10–12 November 1997, Nagoya, Japan, edited by T. Fujiwara and M. Doi (World Scientific, Singapore).

Goedecker, S., 1993, Phys. Rev. B **48**, 17573.

Goedecker, S., 1995, J. Comput. Phys. **118**, 261.

Goedecker, S., 1998a, Phys. Rev. B **58**, 3501.

Goedecker, S., 1998b, *Wavelets and their Application for the Solution of Differential Equations* (Presses Polytechniques Universitaires et Romandes, Lausanne, Switzerland), ISBN 2-88074-398-2.

Goedecker, S., and L. Colombo, 1994a, Phys. Rev. Lett. **73**, 122.

Goedecker, S., and L. Colombo, 1994b, in *Proceedings of the 1994 ACM/IEEE Supercomputing Conference: Supercomputing 94*, Washington D.C. (Association for Computing Machinery, New York), p. 670.

Goedecker, S., and A. Hoisie, 1997 (Los Alamos National Laboratory, unclassified report No. LA-UR-97-1504.

Goedecker, S., and I. Ivanov, 1998a, Solid State Commun. **105**, 665.

Goedecker, S., and I. Ivanov, 1998b, Comput. Phys. **12**, 548.

Goedecker, S., and I. Ivanov, 1999, Phys. Rev. B **59**, 7270.

Goedecker, S., and M. Teter, 1995, Phys. Rev. B **51**, 9455.

Goedecker, S., M. Teter, and J. Hutter, 1996, Phys. Rev. B **54**, 1703.

Gonze, X., 1996, Phys. Rev. B **54**, 4383.

Goodwin, L., 1991, J. Phys.: Condens. Matter **3**, 3869.

Goringe, C., D. Bowler, and E. Hernandez, 1997, Rep. Prog. Phys. **60**, 1447.

Goringe, C., E. Hernandez, M. Gillan, and I. Bush, 1997, Comput. Phys. Commun. **102**, 1.

Greengard, L., 1994, Science **265**, 909.

Hammond, B. L., W. A. Lester, and P. J. Reynolds, 1994, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore).

Hansen, L., K. Stokbro, B. Lundquist, K. Jacobsen, and D. Deaven, 1995, Phys. Rev. Lett. **75**, 4444.

Harrison, W., 1980, *Electronic Structure and the Properties of Solids* (Dover, New York).

Hartwigsen, C., S. Goedecker, and J. Hutter, 1998, Phys. Rev. B **58**, 3641.

Haydock, R., 1980, Solid State Phys. **35**, 215.

Hehre, W. J., L. Radom, J. A. Pople, and P. V. R. Schleyer, 1996, *Ab Initio Molecular Orbital Theory* (Wiley, New York).

Heine, V., 1980, Solid State Phys. **35**, 1.

Hernandez, E., and M. Gillan, 1995, Phys. Rev. B **51**, 10157.

Hernandez, E., M. Gillan, and C. Goringe, 1996, Phys. Rev. B **53**, 7147.

Hernandez, E., M. Gillan, and C. Goringe, 1997, Phys. Rev. B **55**, 13485.

Hierse, W., and E. Stechel, 1994, Phys. Rev. B **50**, 17811.

Hierse, W., and E. Stechel, 1996, Phys. Rev. B **54**, 16515.

Ho, K-M., J. Ihm, and J. Joannopoulos, 1982, Phys. Rev. B **25**, 4260.

Horsfield, A., and A. Bratkovsky, 1996, Phys. Rev. B **53**, 15381.

Horsfield, A., A. Bratkovsky, M. Fearn, D. Pettifor, and M. Aoki, 1996, Phys. Rev. B **53**, 12694.

Horsfield, A., A. Bratkovsky, D. Pettifor, and M. Aoki, 1996, Phys. Rev. B **53**, 1656.

Horsfield, A., P. D. Godwin, D. G. Pettifor, and A. P. Sutton, 1996, Phys. Rev. B **54**, 15773.

Ismail-Beigi, S., and T. Arias, 1999 Phys. Rev. Lett. **82**, 2127.

Ismail-Beigi, S., and T. Arias, 1998, Phys. Rev. B **57**, 11923.

Itoh, S., P. Ordejon, and R. M. Martin, 1995, Comput. Phys. Commun. **88**, 173.

Itoh, S., P. Ordejon, D. Drabold, and R. Martin, 1996, Phys. Rev. B **53**, 2132.

Jayanthis, C., S. Wu, J. Cocks, N. Luo, Z. Xie, M. Menon, and G. Yang, 1998, Phys. Rev. B **57**, 3799.

Johnson, B., P. Gill, and J. Pople, 1993, J. Chem. Phys. **98**, 5612.

Kerker, G., 1981, Phys. Rev. B **23**, 3082.

Kim, J., F. Mauri, and G. Galli, 1995, Phys. Rev. B **52**, 1640.

Kim, J., J. Wilkins, F. Khan, and A. Canning, 1997, Phys. Rev. B **55**, 16186.

King-Smith, R., M. Payne, and J. Lin, 1991, Phys. Rev. B **44**, 13063.

King-Smith, R., and D. Vanderbilt, 1993, Phys. Rev. B **47**, 1651.

Kittel, C., 1963, *Quantum Theory of Solids* (Wiley, New York).

Kohn, W., 1959, Phys. Rev. **115**, 809.

Kohn, W., 1964, Phys. Rev. **133**, A171.

Kohn, W., 1973, Phys. Rev. B **7**, 4388.

Kohn, W., 1993, Chem. Phys. Lett. **208**, 167.

Kohn, W., 1995, Int. J. Quantum Chem. **56**, 229.

Kohn, W., 1996, Phys. Rev. Lett. **76**, 3168.

Kress, J., and A. Voter, 1995, Phys. Rev. B **52**, 8766.

Kress, J., S. Goedecker, A. Hoisie, H. Wassermann, O. Lubeck, L. Collins, and B. Holian, 1998, J. Comput.-Aided Mater. Des. **5**, 295.

Kresse, G., 1996, Phys. Rev. B **54**, 11169.

Landau, E. M., and L. P. Lifshitz, 1980, *Statistical Physics, Part 1*, 3rd edition (Pergamon, New York).

Lee, C., W. Yang, and R. Parr, 1988, Phys. Rev. B **37**, 785.

Lewis, J., P. Ordejon, and O. Sankey, 1997, Phys. Rev. B **55**, 6880.

Li, X.-P., W. Nunes, and D. Vanderbilt, 1993, Phys. Rev. B **47**, 10891.

Lippert, R., T. Arias, and A. Edelman, 1998, J. Comput. Phys. (in press).

Majewski, J., and P. Vogl, 1989, in *The Structure of Binary Compounds* (North-Holland, Amsterdam), p. 287.

March, N., W. Young, and S. Sampanthar, 1967, *The Many-Body Problem in Quantum Mechanics* (Cambridge University, Cambridge, England).

Marzari, N., and D. Vanderbilt, 1997, Phys. Rev. B **56**, 12847.

Maslen, P., C. Ochsenfeld, C. White, M. Lee, and M. Head-Gordon, 1998, J. Phys. Chem. **102**, 2215.

Mauri, F., and G. Galli, 1994, Phys. Rev. B **50**, 4316.

Mauri, F., G. Galli, and R. Car, 1993, Phys. Rev. B **47**, 9973.

McWeeny, R., 1960, Rev. Mod. Phys. **32**, 335.

McWeeny, R., 1989, *Methods of Molecular Quantum Mechanics* (Academic, New York).

Mermin, N., 1965, Phys. Rev. **137**, A1441.

Methfessel, M., and A. Paxton, 1989, Phys. Rev. B **40**, 3616.

Millam, J., and G. Scuseria, 1997, J. Chem. Phys. **106**, 5569.

Nicholson, D., and X.-G. Zhang, 1997, Phys. Rev. B **56**, 12805.

Nightingale, M. P., and C. J. Umrigar, 1998, Eds., *Quantum Monte Carlo Methods in Physics and Chemistry*, NATO ASI Series C, Mathematical and Physical Sciences, Vol. XXX (Kluwer Academic, Boston).

Nunes, R., 1998, private communication.

Nunes, R., J. Bennetto, and D. Vanderbilt, 1996, Phys. Rev. Lett. **77**, 1516.

Nunes, R., J. Bennetto, and D. Vanderbilt, 1998, Phys. Rev. B **58**, 12563.

Nunes, R., and D. Vanderbilt, 1994, Phys. Rev. B **50**, 17611.

Obara, S., and A. Saika, 1986, J. Chem. Phys. **84**, 3963.

Ordejon, P., 1998, Comput. Mater. Sci. **12**, 157.

Ordejon, P., E. Artacho, and J. Soler, 1996, Phys. Rev. B **53**, R10441.

Ordejon, P., D. Drabold, M. Grumbach, and R. Martin, 1993, Phys. Rev. B **48**, 14646.

Ordejon, P., D. A. Drabold, R. M. Martin, and M. P. Grumbach, 1995, Phys. Rev. B **51**, 1456.

Pandey, K., A. Williams, and J. Janak, 1995, Phys. Rev. B **52**, 14415.

Parr, R., and W. Yang, 1989, *Density-Functional Theory of Atoms and Molecules* (Oxford University, New York).

Payne, M., M. Teter, D. Allan, T. Arias, and J. Joannopoulos, 1992, Rev. Mod. Phys. **64**, 1045.

Perdew, J., K. Burke, and M. Ernzerhof, 1996, Phys. Rev. Lett. **77**, 3865.

Perez-Jorda, J., and W. Yang, 1997, J. Chem. Phys. **107**, 1218.

Pettifor, D., 1995, *Bonding and Structure of Molecules and Solids* (Clarendon, Oxford).

Press, W., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1986, *Numerical Recipes, The Art of Scientific Computing* (Cambridge University, Cambridge, England).

Pulay, P., 1977, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer (Plenum, New York), p. 153.

Qiu, S-Y., C. Wang, K. Ho, and C. Chan, 1994, J. Phys.: Condens. Matter **6**, 9153.

Roberts, B., K. Johnson, W. Luo, and P. Clancy, 1999, Chem. Eng. J. (in press).

Saad, Y., 1996, *Iterative Methods for Sparse Linear Systems* (PWS Publishing, Boston).

Sanchez-Portal, D., P. Ordejon, E. Artacho, and J. Soler, 1997, Int. J. Quantum Chem. **65**, 453.

Sankey, O., and D. Niklewski, 1989, Phys. Rev. B **40**, 3979.

Sankey, O., D. Drabold, and A. Gibson, 1994, Phys. Rev. B **50**, 1376.

Schwegler, E., and M. Challacombe, 1996, J. Chem. Phys. **105**, 2726.

Silver, R., and H. Roeder, 1997, Phys. Rev. E **56**, 4822.

Silver, R., H. Roeder, A. Voter, and J. Kress, 1996, J. Comput. Phys. **124**, 115.

Sloan, I., and S. Joe, 1994, *Lattice Methods for Multiple Integration* (Oxford University, New York).

Stephan, U., 1998, private communication.

Stephan, U., and D. Drabold, 1998, Phys. Rev. B **57**, 6391.

Stewart, J., 1996, Int. J. Quantum Chem. **58**, 133.

Stewart, J., P. Csaszar, and P. Pulay, 1982, J. Comput. Chem. **39**, 4997.

Štich, I., R. Car, M. Parrinello, and S. Baroni, 1989, Phys. Rev. B **39**, 4997.

Strain, M., G. Scuseria, and M. Frisch, 1996, Science **271**, 51.

Stratmann, R., G. Scuseria, and M. Frisch, 1996, Chem. Phys. Lett. **257**, 213.

Strout, D., and G. Scuseria, 1995, J. Chem. Phys. **102**, 8448.

Szabo, A., and N. Ostlund, 1982, *Modern Quantum Chemistry* (McGraw-Hill, New York).

Teter, M., M. Payne, and D. Allan, 1989, Phys. Rev. B **40**, 12255.

Vanderbilt, D., 1998, private communication.

Voter, A., J. Kress, and R. Silver, 1996, Phys. Rev. B **53**, 12733.

Wang, L.-W., and M. Teter, 1992, Phys. Rev. B **46**, 12798.

Wang, L.-W., 1994, Phys. Rev. B **49**, 10154.

Wang, Y., G. Stocks, W. Shelton, D. Nicholson, Z. Szotek, and W. Temmerman, 1995, Phys. Rev. Lett. **75**, 2867.

Weinert, M., and J. Davenport, 1992, Phys. Rev. B **45**, 11709.

Wentzcovitch, R., J. L. Martins, and P. Allen, 1992, Phys. Rev. B **45**, 11372.

White, S., J. Wilkins, and M. Teter, 1989, Phys. Rev. B **39**, 5819.

White, C., B. Johnson, P. Gill, and M. Head-Gordon, 1994, Chem. Phys. Lett. **230**, 8.

White, C., P. Maslen, M. Lee, and M. Head-Gordon, 1997, Chem. Phys. Lett. **276**, 133.

Wimmer, E., 1996, Mater. Sci. Eng., B **37**, 72.

Xu, C., and G. Scuseria, 1996, Chem. Phys. Lett. **262**, 219.

Yang, W., 1991a, Phys. Rev. Lett. **66**, 1438.

Yang, W., 1991b, J. Chem. Phys. **94**, 1208.

Yang, W., 1997, Phys. Rev. B **56**, 9294.

Yang, W., and T-S. Lee, 1995, J. Chem. Phys. **103**, 5674.

York, D., T.-S. Lee, and W. Yang, 1996, J. Am. Chem. Soc. **118**, 10940.

Zhao, Q., and W. Yang, 1995, J. Chem. Phys. **102**, 9598.

Zhou, S., D. Beazley, P. Lomdahl, and B. Holian, 1997, Phys. Rev. Lett. **78**, 479.