Ab initio molecular dynamics with a finite-temperature density functional

1994 J. Phys.: Condens. Matter 6 1999

(http://iopscience.iop.org/0953-8984/6/10/017)

J. Phys.: Condens. Matter 6 (1994) 1999–2014. Printed in the UK

Ab initio molecular dynamics with a finite-temperature density functional

Matthew P Grumbach†‖, Detlef Hohl‡¶, Richard M Martin† and Roberto Car§

† Department of Physics and Materials Research Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

‡ National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

§ Institut Romand de Recherche Numerique en Physique des Materiaux (IRMMA), PHB-Ecublens, CH-1015, Lausanne, and Department of Condensed Matter Physics, University of Geneva, Geneva, CH-1211, Switzerland

Abstract. The unified method for molecular dynamics and density functional theory (MD/DF) introduced by Car and Parrinello is based on zero-temperature density functional theory. We have incorporated the finite-temperature extension of density functional theory proposed by Mermin into a consistent fictitious Lagrangian framework. Such an extension of the original MD/DF method is desirable for two rather different reasons. First this framework provides a general method to treat electronic states at finite temperature or in non-equilibrium excited states. Second it can alleviate certain practical problems that arise when Kohn–Sham DF methods in general and MD/DF in particular are applied to metallic and near-metallic systems. Our approach involves dynamically varying occupation numbers, which is important for states near the Fermi energy. We show that the added degrees of freedom of these states can be used to accelerate the convergence to the electronic ground state. In MD simulations this improved response of the electrons also leads to an increase in the rate of energy transfer from the ionic to the electronic degrees of freedom. Our method is illustrated by calculations on crystalline metallic carbon and simulations of liquid silicon.

1. Introduction

The original unified method for molecular dynamics and density functional theory (MD/DF) (Car and Parrinello 1985) has been used with great success on a variety of insulating (Hohl and Jones 1992), semiconducting (Car and Parrinello 1988, Hohl and Jones 1991) and metallic systems (Galli *et al* 1989, Fernando *et al* 1989, Qian *et al* 1990, Stich *et al* 1989a). This method is based on the Hohenberg–Kohn–Sham total energy for the ground state of a system of electrons (Hohenberg and Kohn 1964, Kohn and Sham 1965). In this paper we present a method based on the Mermin extension (Mermin 1965) of the Hohenberg–Kohn theorem that treats electronic systems at finite temperature.

Our first motivation for introducing the Mermin extension is that it offers the proper description of the inhomogeneous electron gas at finite temperatures. This would be appropriate for applications where the temperature of the electronic system is greater than the first excitation energy. Plasmas and liquid metals are typical examples. In principle

‖ Current address: Box 871504, Department of Physics and Astronomy, Arizona State University, Tempe, AZ 85287, USA.

¶ Permanent address: Institut für Festkörperforschung, Forschungszentrum Jülich, D-5170 Jülich, Germany.

0953-8984/94/101999+16$19.50

minimizing the free energy functional with respect to variations in the charge density will yield the exact equilibrium charge density. One problematic point in using this formalism is the evaluation of the temperature-dependent exchange–correlation functional, $E_{xc}^{T}[n]$. Practical means for doing this involve two approximations: a local density approximation (LDA) and an approximate expression for E_{xc}^{T} for the homogeneous electron gas (Perrot and Dharma-wardana 1984, Kanhere *et al* 1986). One can therefore envisage three schemes for simulating an electronic system at finite temperature: (i) use zero-temperature techniques, namely by requiring a ground state (i.e. $T = 0$) distribution of states and using $E_{xc}^{T=0}$; (ii) maintain the electrons in a thermal distribution of states but continue to use $E_{xc}^{T=0}$; (iii) maintain the electrons in a thermal distribution and use E_{xc}^{T}. In the context of ionic systems scheme (i) corresponds to the Born–Oppenheimer approximation, while schemes (ii) and (iii) correspond to the approximation that the electrons equilibrate to a finite temperature on time scales small compared to the ionic motion. In view of the approximate knowledge that we have of E_{xc}^{T} it is not clear which scheme will give more realistic ionic forces for use in an actual MD simulation. In this work we will offer an implementation of scheme (ii).

There has in fact been some debate about whether maintaining the electrons in the ground state is the physically most relevant procedure in systems with level crossings (Harris 1984, Dunlap 1984, Dunlap 1988). In such cases the energy surface for the ions has a cusp at the ionic configuration where the level crossing occurs, resulting in discontinuous forces. One motivation for introducing fractional occupation numbers is to eliminate this cusp. In principle this can be done by employing any energy functional in which the occupation numbers are included variationally (Weinert and Davenport 1992), although the free energy functional is the most natural choice.

There are also a number of technical problems which might be alleviated by the addition of extra states. Difficulties in finding the self-consistent solution of the Kohn–Sham (KS) equations for metals are well known. Common solution techniques exhibit slow convergence when applied to metallic and near-metallic systems using integral occupation of the ground state wavefunctions. Doubly occupying only the lowest $N_e/2$ states in a system of N_e electrons can lead to oscillatory behaviour of the highest occupied state. Below we will present evidence that iterative techniques also show slow convergence when applied to small gap systems. In the language of optimization theory, this is a problem of ill-conditioning, since the presence of nearly degenerate states introduces small eigenvalues into the Hessian matrix. The inclusion of extra states makes possible the development of algorithms that can treat this type of ill-conditioning (Gillan 1989, Arias *et al* 1992). In the appendix we explain in greater detail the origin of the conditioning problem related to metals and discuss specific strategies for overcoming it.

Total energy calculations on metals are also known to converge slowly with respect to the size of the k-point sampling grid. This is due to the complicated nature of the Fermi surface. It has been shown that using a smeared Fermi surface can improve the accuracy of a given k-point sampling grid (Fu and Ho 1983). The scheme described here using the free energy functional is one way of introducing this smearing.

Another technical problem arises during MD simulations. MD/DF simulations rely on the efficient thermal *decoupling* of the ionic and (fictitious) electronic classical subsystems $\{\boldsymbol{R}_I\}$ and $\{\psi_i\}$ (Pastore *et al* 1991). In systems with small optical gap one observes increasing transfer of thermal energy from the ionic to the electronic subsystem with decreasing gap size. This energy exchange can become rapid on the MD time scale and leads to large deviations of the wavefunctions from the ground state solution and a matching cooling of the ionic subsystem. The sum of ionic kinetic and potential energies

$\frac{1}{2}\sum_I M_I \dot{\boldsymbol{R}}_I^2 + E[\{\psi_i\},\{\boldsymbol{R}_I\}]$ is no longer conserved and the interatomic forces become unphysical. Several schemes have been proposed to treat this so-called "metals problem". A Nosé thermostat can be connected to the ionic system to maintain it at a constant temperature, despite energy loss to the electrons (Galli *et al* 1989). The excess energy that appears as fictitious kinetic energy of the electrons is removed either through periodic quenches (Galli *et al* 1989), or through another Nosé thermostat connected to the electronic system (Blöchl and Parrinello 1992). All these modified methods simulate the canonical ensemble and it would be very desirable to *prevent* the thermal equilibration in metallic systems rather than cope with it and thus be able to perform microcanonical MD/DF simulations.

Two processes are responsible for the unwanted energy transfer between the classical mechanical subsystems in small gap systems (Pastore *et al* 1991). (i) instabilities in the potential energy surface for the electrons that arise when the changing ionic configuration causes a level-crossing to occur and (ii) overlap of the characteristic frequencies of the ionic and electronic subsystems $\{\boldsymbol{R}_I\}$ and $\{\psi_i\}$ when the optical gap closes. Process (ii) is a simple consequence of basic classical mechanical rules (Pastore *et al* 1991). For insulating systems it can be dealt with by fine-tuning the relative masses μ_i and M_I of the electronic and ionic subsystems (Pastore *et al* 1991, Harris and Hohl 1990). For zero-gap systems, however, there is no choice of relative masses that will entirely separate the spectra of the two systems. Another possible motivation for introducing the free energy functional is that by addressing process (i) alone and removing discontinuities near the Fermi energy part of the energy transfer in metallic systems might be prevented. As we shall demonstrate, however, the use of fractional occupation numbers introduces new low frequency modes in the fictitious dynamics of the electrons, resulting in an increase in the energy transfer due to process (ii).

Schemes using fractional occupation numbers are, of course, well-established in LDA-KS studies of metallic systems and often involve occupations f_i according to a Fermi–Dirac distribution. A very similar scheme was used in the first MD/DF studies of a strongly metallic system (Fernando *et al* 1989, Qian *et al* 1990). The authors computed the eigenvalues of the occupied states in every MD time step and use them in a fixed Fermi distribution to assign occupations to the states. Pederson and Jackson (1991) have proposed using fractional occupation numbers in the zero-temperature DF by parametrizing the occupation numbers in terms of a Fermi function and then treat the parameter (called a 'pseudoenergy') as a variational parameter. Gillan (1989) has used a finite-temperature DF as a tool to eliminate discontinuities near the Fermi edge; this work introduces the free energy functional and treats the occupation numbers directly as variational parameters, but is concerned with static properties only and no MD is performed. The use of occupation numbers as variables of motion was suggested by Oguchi and Sasaki (1991), but they did not report a calculation in which the occupations actually vary dynamically. More recently the importance of using a properly variational energy functional was pointed out (Weinert and Davenport 1992) that includes an additional term in the occupation numbers f_i over the standard form. It was shown that a Fermi form of the f_i leads to the grand potential of finite-temperature thermodynamics. When MD is the goal, forces have to be derived from the properly variational functional. An MD study using the Mermin formulation of DF theory with a Fermi distribution for the f_i was performed by Wentzcovitch *et al* (1992). Not surprisingly, the authors found improved energy conservation using consistent expressions for force and potential energy.

In this work we present, apply and analyse a full implementation of the finite-temperature density functional within a fictitious Lagrangian framework. We adopt this framework as a numerical expedient, although in principle it could be used to simulate electronic systems at

finite temperature. In section 2 we show in detail how dynamically varying occupation numbers can be derived from a fictitious Lagrangian framework. Applications of this formalism to the electronic structure of liquid silicon and crystalline high-density carbon and also to molecular dynamics simulations of liquid silicon are presented in section 3. An appendix is also included in which we derive some simple properties of the total energy surface for the general case of fractionally occupied states.

2. Formalism

Incorporating fractional occupation numbers into the Hohenberg–Kohn–Sham expression for the total energy is straightforward (Janak 1978):

$$E^f[\{\psi_i\},\{\boldsymbol{R}_I\}] = 2\sum_{i=1}^{N} f_i \int \mathrm{d}\boldsymbol{r}\ \psi_i^*\left(-\frac{1}{2}\nabla^2\right)\psi_i + U[n] \tag{1}$$

$$U[n] = \int \mathrm{d}\boldsymbol{r}\ V(\boldsymbol{r})n(\boldsymbol{r}) + \frac{1}{2}\int\int \mathrm{d}\boldsymbol{r}\ \mathrm{d}\boldsymbol{r}'\ \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} + E_{\mathrm{xc}}[n] \tag{2}$$

$$n(\boldsymbol{r}) = 2\sum_{i=1}^{N} f_i|\psi_i|^2 \tag{3}$$

where we require

$$0 < f_i < 1 \tag{4}$$

and work with a fixed integral total number of electrons

$$2\sum_{i=1}^{N} f_i = N_e. \tag{5}$$

Here N is the total number of states included in the calculation. The ground state functional used in the original method can be recovered by setting $f_i = 1$ and $N = N_e/2$. Calculations based on this ground state functional will hereafter be referred to as the *standard method.* The appropriate thermodynamic potential for finite temperatures is the free energy (Mermin 1965, Callaway and March 1984):

$$\begin{aligned} F[\{\psi_i\},\{\boldsymbol{R}_I\},\{f_i\}] &= E^f[\{\psi_i\},\{\boldsymbol{R}_I\},\{f_i\}] - TS \\ &= E^f[\{\psi_i\},\{\boldsymbol{R}_I\},\{f_i\}] + 2T\sum_i (f_i \ln f_i + (1-f_i)\ln(1-f_i)). \end{aligned} \tag{6}$$

Here we have inserted the expression for the entropy of an ideal Fermi gas and are treating f_i as independent variables. Although we continue to use the zero-temperature form of E_{xc}, the development here does not depend on this specific choice. When the free energy is minimized with respect to f_i one recovers the Fermi function:

$$f_i = \frac{1}{1+\exp((H_{ii}-\epsilon_{\mathrm{F}})/T)} \tag{7}$$

The variables $\{f_i\}$ are difficult to use directly as dynamic variables since they must be constrained to the range [0,1]. We therefore parametrize each f_i with the form of a Fermi function:

$$f_i = \frac{1}{1+\exp((\epsilon_i-\epsilon_{\mathrm{F}})/T)} \tag{8}$$

where ϵ_{F} is chosen to satisfy equation (5) and T is identified with the temperature defined in equation (6). The constraint $0 < f_i < 1$ is now satisfied implicitly. Although other

parametrizations could be used here, this particular choice leads to a natural interpretation of the ϵ_i at the stationary point of the free energy. Note that since ϵ_i and ϵ_F only appear as differences, they are only defined within an additive constant. Let us next construct a fictitious Lagrangian, treating $\{\epsilon_i\}$ as classical degrees of freedom:

$$\mathcal{L} = \sum_i \mu_i \int \mathrm{d}\boldsymbol{r}\ |\dot{\psi}_i(\boldsymbol{r})|^2 + \frac{1}{2}\sum_I M_I \dot{\boldsymbol{R}}_I^2 + \frac{1}{2}\sum_i Q_i \dot{\epsilon}_i^2$$
$$-F[\{\psi_i\},\{\boldsymbol{R}\},\{\epsilon_i\}] + \sum_{ij}\Lambda_{ij}\left(\int \mathrm{d}\boldsymbol{r}\ \psi_i^*(\boldsymbol{r})\psi_j(\boldsymbol{r}) - \delta_{ij}\right). \tag{9}$$

This leads to equations of motion:

$$\mu_i \ddot{\psi}_i = -\frac{\delta F[\{\psi_i\}]}{\delta \psi_i^*} + \sum_j \Lambda_{ij}\psi_j = -f_i \hat{H}\psi_i + \sum_j \Lambda_{ij}\psi_j \tag{10}$$

$$Q_i \ddot{\epsilon}_i = -\frac{\partial F[\{\epsilon_i\}]}{\partial \epsilon_i} = g_i\left\{H_{ii} - \epsilon_i - \sum_j (H_{jj}-\epsilon_j)\frac{g_j}{\sum g}\right\} \tag{11}$$

$$M_I \ddot{\boldsymbol{R}}_I = -\nabla_{R_I} F \tag{12}$$

where we have introduced the definitions

$$g_i = g(\epsilon_i) = -\frac{\mathrm{d}f(\epsilon_i)}{\mathrm{d}\epsilon_i} \tag{13}$$

$$H_{ij} = \int \mathrm{d}\boldsymbol{r}\ \psi_i^*(\boldsymbol{r})\hat{H}\psi_j(\boldsymbol{r}). \tag{14}$$

The function $g(\epsilon)$ exhibits a peak of width $\sim k_B T$ and height $1/2T$ at $\epsilon = E_{\text{Fermi}}$.

The gradients of F that appear in the above equations can also be used in optimization algorithms such as steepest descent or conjugate gradient to minimize the free energy. At the stationary point the gradients vanish giving, for the electronic degrees of freedom,

$$f_i \hat{H}\psi_i = \sum_j \Lambda_{ij}\psi_j \tag{15}$$

$$\epsilon_i = H_{ii} + \text{constant}. \tag{16}$$

Thus at the minimum, the ϵ_i are equal to the diagonal matrix elements of the Hamiltonian. In the appendix it is shown that these stationary conditions also imply that the states near the Fermi energy are exact eigenstates of the Hamiltonian. In this case the H_{ii} are precisely the eigenvalues. When the electronic system is in a steady state, the Hellmann–Feynman theorem can be applied, guaranteeing that the forces defined in equation (12) are the physical forces acting on the ions.

3. Applications

3.1. Electronic minimization

To illustrate the utility of the free energy formalism, we have performed minimization calculations both on simple cubic carbon at very high pressures and on a liquid configuration of silicon. These two cases complement each other, since for the boundary conditions used in our calculations the crystalline phase of carbon exhibits an exact degeneracy at the Fermi level while the chosen configuration of silicon atoms exhibits a near degeneracy.

For the carbon calculation we have used a 64-atom supercell with lattice constant 9.03 bohr, which corresponds to a pressure near 30 Mbar. In this pressure regime, the

simple cubic structure is more stable than the diamond structure. To compute the free energy we include 32 extra states beyond the 128 needed for the ground state, and have used a broadening temperature of 1000 K. We have used a soft pseudopotential (Troullier and Martins 1991) with a cutoff of 40 Rydbergs and the Kleinman–Bylander (Kleinman and Bylander 1982) scheme for treating the non-local matrix elements. Only the Γ point is used in sampling the Brillouin zone.

For the calculations involving the standard formalism, the ψ_i are relaxed using a conjugate gradient algorithm (Stich *et al* 1989a). For the calculations involving the free energy formalism, the ψ_i were relaxed using the same conjugate gradient algorithm followed by one step of an iterative procedure defined by equation (31) in the appendix. The conjugate gradient step updates the ψ_i by eliminating components of excited states. The iterative procedure updates the ψ_i by mixing states that are within the occupied manifold, the effect of which is to bring states which are near the Fermi level closer to being eigenstates. The ϵ_i degrees of freedom were relaxed using a steepest descent procedure:

$$\epsilon_i^+ = \epsilon_i^0 - \Delta_i \left.\frac{\partial F}{\partial \epsilon_i}\right|_0 . \tag{17}$$

Since the gradient is proportional to $g(\epsilon_i)$, which is an exponentially varying function, it is not possible to choose a single step parameter, Δ_{const}, that would quickly converge all the ϵ_i. Instead we assign a different value to each:

$$\Delta_i = 0.005 * 2T/g(\epsilon_i). \tag{18}$$

In figure 1 we present the convergence of the total energy for the standard method and the free energy using the new method. For comparison we have also shown the convergence for the insulating diamond structure. It is clear that the efficiency of the standard CG algorithm is seriously degraded when an exact degeneracy is present at the Fermi energy. We find a substantial improvement in the convergence rate using the new method as compared with the standard formalism.

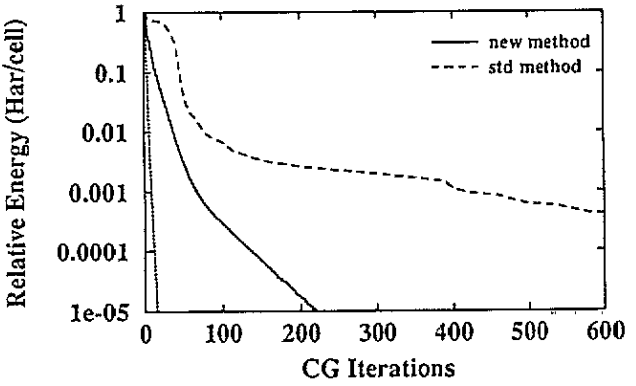

Figure 1. Comparison of the convergence behaviour of the standard (integrally occupied states) method (dashed line) and the fractional occupation method (solid line) when minimizing the electronic degrees of freedom for a 64-atom simple cubic carbon sample. Also shown is the convergence for 64-atom diamond sample using the standard method (dotted line). The count of iterations for the new method has been multiplied by 1.7 to account for increased computation time.

Some of the initial ψ_i consisted substantially of excited states of the Hamiltonian. The standard algorithm assigns full occupation, $f_i = 1$, to these states. The plateau seen early

on in figure 1 corresponds to the process of projecting out the correct, lower energy states that should be included in the occupied manifold. In the new method, these states were initially assigned a finite occupation, but the correct states were also included with an initial occupation of zero. By varying the occupation numbers rather than the entire state, the new method is able to efficiently remove the excited states.

Once the occupation numbers are approximately correct, both algorithms become limited by the process of unmixing the states near the Fermi level. Before the standard method reaches the minimum, the states near E_{Fermi} are not exactly degenerate. The algorithm attempts to reach the minimum by projecting out the subspace of unoccupied nearly-degenerate states from the subspace of occupied nearly-degenerate states. The near-degeneracy of these states makes the calculation badly conditioned, resulting in slow convergence. In the new method, the states near E_{Fermi} are all assigned nearly the same occupation, $f_i = 0.5$. In this case the minimum energy will be reached when the matrix H_{ij} is block diagonal, with the three sets of states corresponding to $f_i = 0, 0.5, 1$ forming the blocks. Since these sets of states are all separated by large gaps, the CG algorithm does not encounter the conditioning problem that besets the standard algorithm. In the final stage of the calculation we observe that the subspace mixing is responsible for the entire energy drop on each step. During this stage mixing of the occupied states results in small corrections to the occupation numbers, which then must be included in the Hamiltonian. This procedure must be continued until self-consistency is reached. Thus the convergence rate of the new method is limited by the efficiency of the mixing algorithm.

As another test of the algorithm we have performed a steepest descent minimization on a 64-atom supercell of silicon taken from a liquid simulation. The electronic structure for this configuration shows a gap of 0.01 eV at the Γ point, indicating nearly degenerate states at the Fermi level. We have used a standard pseudopotential (Bachelet *et al* 1982), with the Kleinman–Bylander form for the non-local matrix elements. The cutoff for the plane-wave basis is 12 Rydbergs, and only the Γ point is used for Brillouin zone summations. For the free energy formalism, we include 10 extra states beyond those needed for the ground state, and use a broadening temperature of 2000 K in the Fermi function. Each step in the calculation consists of one steepest descent step, one step of the iterative procedure defined by equation (31) in the appendix, and one steepest descent step for the ϵ_i parameters, equation (17). In addition a unitary transformation is applied to ψ_i once every 50 steps that exactly diagonalizes H_{ij}. The results of the comparison are presented in figure 2. Also shown is the convergence rate using the standard technique on a 64-atom unit cell in the diamond structure.

We find first of all that the simple steepest descent algorithm used successfully on the diamond structure is not efficient for the small gap structure. As the gap of the system decreases, the small eigenvalues of the Hessian matrix become even smaller, which leads to ill-conditioning. By including extra states with fractional occupations, the ill-conditioning is not eliminated (as happened in the high symmetry case above) but it does become tractable. Using the steepest descent procedure for effecting rotations within the occupied subspace, equation (31), gives an improved convergence as shown in the first 100 steps of figure 2. However we find that this procedure alone eventually leads to a convergence rate essentially identical to that of the standard method. We believe that the explanation for this is that initially the occupied states are very far from being eigenstates of the Hamiltonian and so the steepest descent procedure applied to this very non-linear problem is not efficient. To quickly move the wavefunctions into the vicinity of the minimum we have performed an exact diagonalization of H_{ij} on iteration number 55. After an initial increase in energy (for reasons discussed in the appendix) we see a substantial drop in energy followed by improved

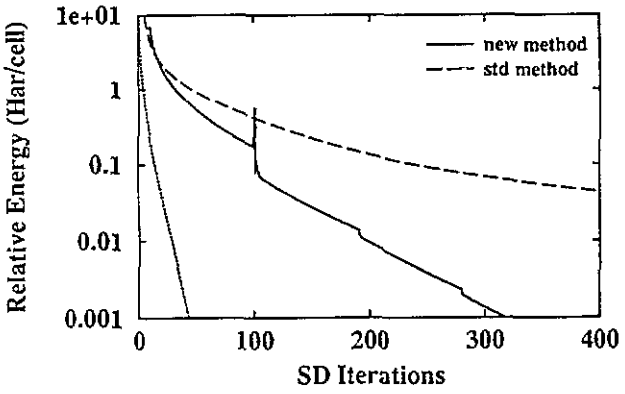

Figure 2. Comparison of the convergence behaviour of the standard (dashed line) method and the new method (solid line) when minimizing the electronic degrees of freedom for a 64-atom disordered silicon sample. Also shown is the convergence of the standard method applied to the 64-atom cell of silicon in the diamond structure (dotted line). The count of iterations for the new method has been multiplied by 1.8 to account for increased computation time.

convergence. We find that one diagonalization every 50 steps is sufficient to maintain a rapid rate of convergence.

3.2. Molecular dynamics simulation

We have also performed molecular dynamics simulations on liquid silicon using the new formalism. Special attention must be given to the integration of the forces acting on the ψ_i and ϵ_i degrees of freedom because they depend linearly on the functions $f(\epsilon)$ and $g(\epsilon)$. The first difficulty is that, since the forces vary exponentially, there is no single mass, μ_{const} or Q_{const}, that would be appropriate for all i. This is basically a conditioning problem which we can overcome in the same manner as used in the minimization calculation, namely by using state-dependent masses:

$$\mu_i = 300 \text{ au} * \overline{f}_i \tag{19}$$

$$Q_i = 10000 \text{ au} * \overline{g}_i/2T \tag{20}$$

where $\overline{f}_i$ and $\overline{g}_i$ are typical values expected during the simulation. We have chosen these equal to the values of f_i and g_i at the start of the simulation. Another more serious difficulty can arise during the course of the simulation: it is possible for a state with an initially small occupation number to evolve into a state with a significantly larger occupation. This means that a large force (proportional to the current value of the occupation number) will be acting on a small mass (assigned according to the initial value of the occupation), causing the Verlet algorithm to become unstable. We treat this instability by monitoring $f_i(t)/f_i(0)$ during the simulation. When this ratio becomes greater than 3.0 we alter the masses of the corresponding ψ_i and ϵ_i parameter to be commensurate with the current values of f_i and g_i, respectively. Strictly speaking this procedure is a violation of Lagrangian dynamics. However we find that the intervention is small and infrequently needed. In the simulation described here, only 6 alterations to individual masses were needed during the course of 1500 time steps.

The time step used is 5.5 au. The initial velocities, $\dot{\psi}_i$ and $\dot{\epsilon}_i$, are determined by performing quenches at the first two atomic positions. The iterative method of Ryckaert *et al* (1977) is used to maintain the holonomic constraints of orthonormality. Figure 3 shows

the time evolution of some of the eigenvalues, e_i, of the subspace Hamiltonian, H_{ij}, and also the time evolution of some of the diagonal elements of the subspace Hamiltonian, H_{ii}. A subspace diagonalization is performed at time $t = 0$, so initially $H_{ii} = e_i$. During the course of the simulation we observe two kinds of behaviour. States far from E_{Fermi} (not shown), which have equal occupation numbers, tend to mix substantially. This has no effect on the total energy or forces since this mixing is a symmetry of the system. States near E_{Fermi} undergo a lesser amount of mixing. This mixing, however, is not a symmetry of the system and does affect the calculated energy and forces. Were these mixing modes adiabatically decoupled from the ionic system, we could invoke averaging arguments to claim that the resulting ionic dynamics would closely approximate the exact Born–Oppenheimer dynamics. In the present simulation, however, the electronic and ionic systems *are* coupled, so that the quality of the forces deteriorates as the simulation proceeds. Figure 4 shows the time evolution of the ϵ_i parameter. As indicated by the equation of motion, ϵ_i tends to follow the corresponding value of H_{ii}. Figure 4 also shows the time evolution of the occupation number, f_i.

In parallel to the above, we have performed another simulation using only 128 states with equal (integral) occupations to compare with the above simulation. In particular the rate of fictitious kinetic energy transfer to the electronic wavefunctions ψ_i was examined. The time development of the fictitious kinetic energy of the two simulations is shown in figure 5. We find that the energy gain is faster when using the fractional occupation method. This will be further discussed below.

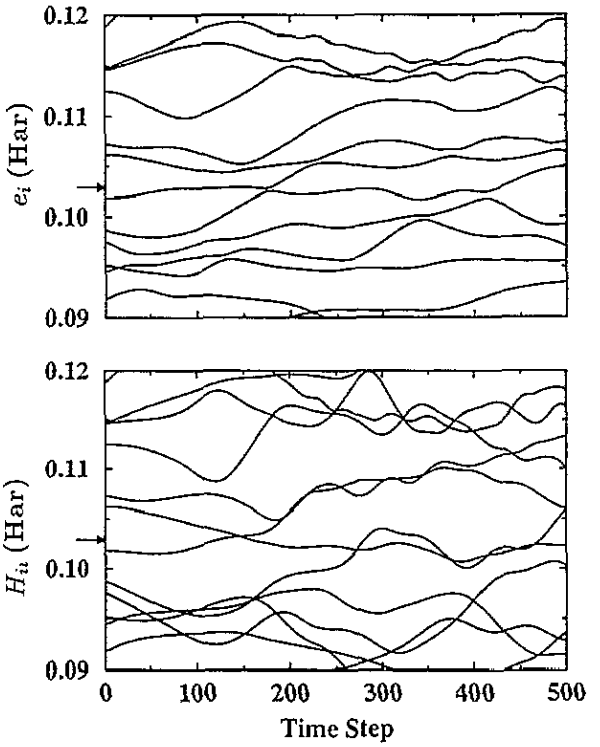

Figure 3. MD simulation for liquid silicon (see text). Time evolution of the eigenvalues, e_i, of H_{ij} and the diagonal matrix elements, H_{ii}. Arrows indicate the average value of E_{Fermi}.

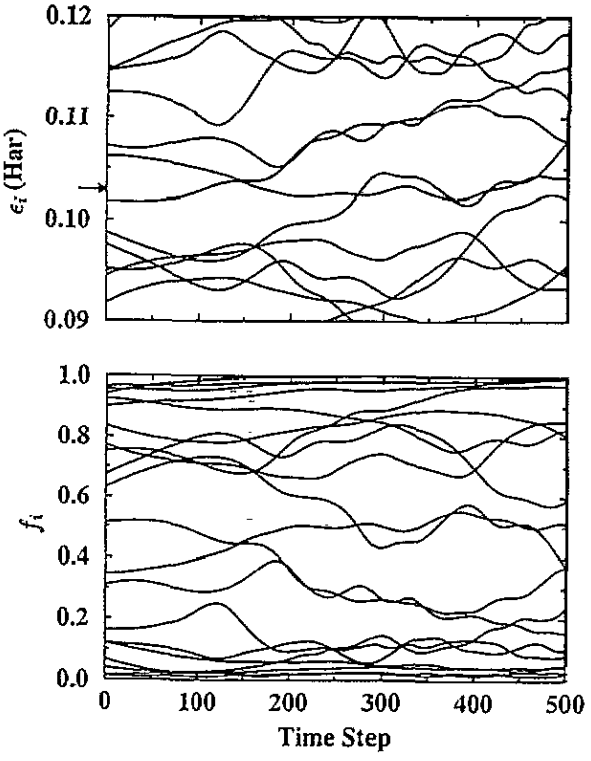

Figure 4. MD simulation for liquid silicon (see text). Time evolution of the parameters ϵ_i and occupation numbers f_i. Arrow indicates the average value of E_{Fermi}.

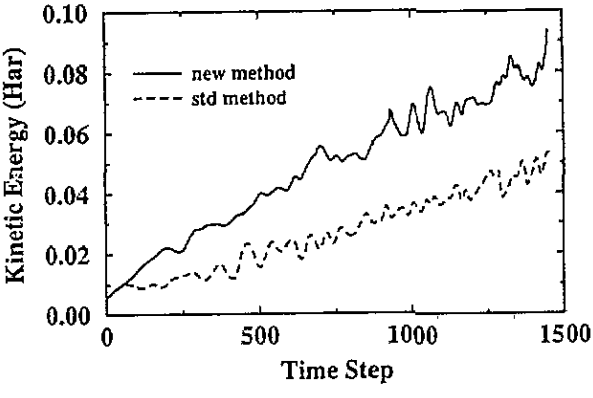

Figure 5. Time evolution of total kinetic energy of the electronic degrees of freedom during the liquid Si simulation. Solid line is the new method, dashed line the standard method with integrally-occupied states.

We have also examined the fictitious kinetic energy of the individual electronic states at a given time during the simulation. For the case of equal occupation numbers we see a uniform distribution of the kinetic energy among the states in the simulation. This can be understood as follows. At the beginning of the simulation, a sub-space diagonalization was performed to guarantee that the states were all exact eigenstates of the Hamiltonian. As the simulation progresses, however, there is no restoring force to maintain the states as

eigenstates of the instantaneous Hamiltonian, and in fact they consist of arbitrary unitary transformation of the eigenstates. The response of the electrons to resonant excitation from the ions is dominated by the response of the eigenstates near the Fermi level. Since the simulation states contain arbitrary linear combinations of these eigenstates, we expect a uniform response.

For the case of fractional occupations we find that states near E_{Fermi} gain a much greater amount of kinetic energy than states far from E_{Fermi}. This is a result of the fact that the total energy is no longer invariant with respect to unitary transformations of the states. The total energy is minimized when the states are exact eigenstates which implies that during a simulation they tend to oscillate (in the space of unitary transformations) around the exact eigenstates. In particular, the states near the Fermi level tend to remain near to eigenstates. Since these are also the states that dominate the electronic response to the ionic motion, we expect these states to gain a greater share of the kinetic energy.

Since the fractional occupation method does not suffer from the level switching instability present in the standard method, the increased rate of energy transfer seen in the new method must be explained in terms of the resonance mechanism mentioned in section 1. We would expect an increase in energy transfer if either the efficiency of each resonant mode increases or if the density of resonant modes increases. To investigate this process we have performed the following simulation on the ψ_i only, keeping the ions fixed. We take the instantaneous values of the ψ_i and of the ionic coordinates from step number 700 in the simulation discussed above. The ψ_i are then allowed to evolve dynamically starting with initial velocities of zero. The ions are held fixed during this simulation. After equilibrating the system for 300 steps, we begin recording the velocity–velocity autocorrelation function:

$$\Gamma(t) = \left\langle \frac{\sum_{i,G} \dot{c}^*_{iG}(t_0+t)\dot{c}^*_{iG}(t_0)}{\sum_{i,G} \dot{c}^*_{iG}(t_0)\dot{c}^*_{iG}(t_0)} \right\rangle_{t_0} . \tag{21}$$

By taking the Fourier transform of this quantity we are able to obtain the power spectrum of the vibrations that are excited during the evolution of the electronic system. A parallel simulation is also performed using the new method. In this case we also fix the ϵ_i degrees of freedom. Although the ionic positions were also taken from step number 700 of the simulation, the configuration is not exactly the same as above due to differences in the dynamics generated by the two methods. Our results are shown in figure 6. As compared to the standard calculation, the new calculation shows a major shift of spectral power to lower frequencies.

The origin of this shift can be traced back to the introduction of fractional occupation numbers in the total energy. Equation (29) in the appendix gives the characteristic frequencies that would be observed in a non-self-consistent simulation with small displacements. We can use this result to qualitatively demonstrate the origin of the shift in spectral power. For each of the two types of simulation (standard formalism and free energy formalism) we have taken the values of f_i and e_i from the fixed-ion simulation discussed above. These values were then used in equation (29) to generate a list of characteristic frequencies. A histogram was then constructed from this list to give a classical density of states, which is shown in figure 7. We find that the introduction of fractional occupation numbers results in a dramatic shift of the frequencies towards the lower end of the spectrum. The presence of additional modes at lower frequencies offers an explanation for the results of the previous simulations. Assuming that the additional modes are excited to an equal extent, we would expect to see a greater spectral weight at these frequencies. In fact they are not excited equally, as can be seen by comparing figures 6 and 7. The presence of a greater number of modes in the range of the phonon modes of liquid silicon ($\sim$ 2–20 THz)

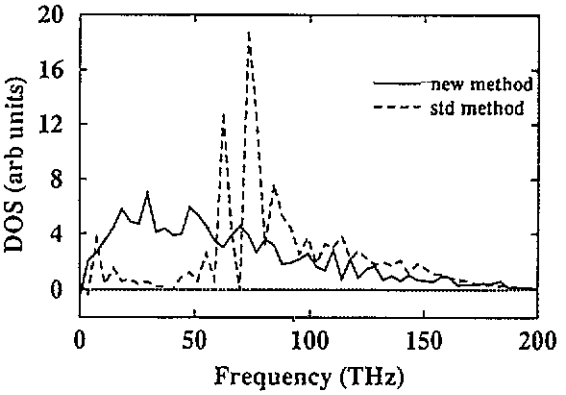

Figure 6. Classical power spectrum of electronic degrees of freedom for fixed ions using the new method (solid line) and the standard method (dashed line). The positions of the ions and the initial values for the c_{iG} were from the previous simulation.

would also result in a greater overall energy transfer from the ionic system to the electronic system.

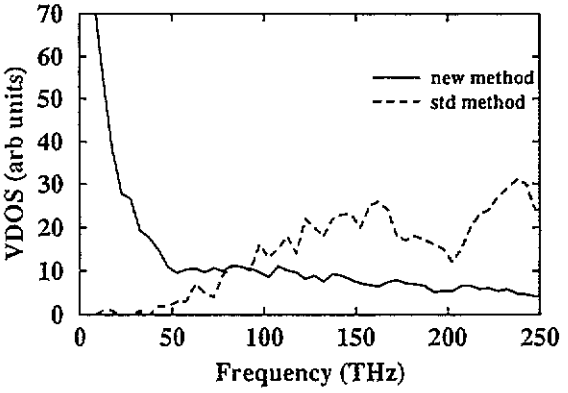

Figure 7. Density of states for fictitious electronic degrees of freedom as estimated by equation (29) of the appendix. Solid line indicates fractional occupation method, dashed line indicates integral occupation method.

4. Conclusions

We have shown how to incorporate finite-temperature density functional theory into a fictitious Lagrangian framework. This new method provides a means of introducing a thermal distribution of states into an MD/DF calculation, which is useful for simulations of thermal electronic systems such as liquid metals and plasmas.

We have investigated the problem of minimizing the electronic free energy, $F[\{\psi_i\}, \{f_i\}]$ with respect to the electronic degrees of freedom, ψ_i and f_i in small gap systems. Ill-conditioning in these calculations can be traced to unitary transformation among states near the Fermi level which only weakly affect the total energy. By including states above the Fermi level in the calculation we are able to attack the problem directly. We have applied this technique to a system with high symmetry and exact degeneracies (metallic carbon) and

also to a disordered system with near degeneracies (liquid silicon). In the high-symmetry case we find that the new method quickly finds all the degenerate states and assigns an appropriate occupation number to them. At this point states with different occupation number become separated by wide energy gaps, effectively eliminating the conditioning problem and resulting in rapid convergence, as shown in figure 1. On the other hand, in a disordered system with near accidental degeneracies merely assigning appropriate occupation numbers does not remove the conditioning problem. In this case we find that the convergence rate is determined by the efficiency of the technique used to obtain exact eigenstates of the states near the Fermi level. We have discussed several strategies for treating this and demonstrate that the convergence rate of the calculation can be improved, as illustrated in figure 2.

A full MD simulation on liquid silicon was carried out treating both the ionic and electronic systems dynamically. We have compared the new method (having dynamically varying occupation numbers) with the standard method (having fixed integral occupation numbers). For the particular case of liquid silicon, shown in figure 5, we find that the rate of energy transfer from the ionic to the electronic system is approximately doubled when the new method is used. We have analysed this behaviour and find that the introduction of fractional occupation numbers causes new low-energy modes to appear in the vibrational spectrum of the electronic degrees of freedom. These new modes provide additional coupling between the two systems, resulting in an increase in energy transfer. This rate of transfer can be handled by the same procedures used in current calculations on metals using thermostats to keep the ion and electron temperatures constant (Blöchl and Parrinello 1992). Only future tests incorporating such thermostats will determine whether or not the improved treatment of the electrons in metals using the fractional occupation method will offset any errors introduced by the thermostat method.

Acknowledgments

We thank Xiao-Ping Li and Giulia Galli for helpful discussions. Two of the authors (MPG and RMM) acknowledge the support of the National Science Foundation under Grant DM 89 20538. Computations were performed at the National Center for Supercomputing Applications, Urbana, IL.

Appendix. The role of unitary transformations

In this appendix we will explain how the total energy, $E^f[\{\psi_i\}]$, depends on unitary transformations of the occupied states and what this implies for both minimization and dynamical calculations.

The first step in this analysis is to express the wavefunctions in terms of another set of orthonormal functions:

$$\psi_i = \sum_{j=1}^{M} \alpha_{ij}\phi_j \tag{22}$$

where M is the size of the Hilbert space. For notational convenience we will also allow the index i to run from 1 to M, so that the number of states is equal to the size of the basis. In actual calculations the number of bands, N, is usually smaller than M. For most of the results below, the case of $i = 1, N$ where $N < M$ can be recovered by setting $f_i = 0$ for $N < i \leqslant M$. The constraint that $\{\psi_i\}$ be orthonormal now becomes the requirement that

the matrix α_{ij} be unitary. The first simplification that we will make is that the potential energy will be treated non-self-consistently:

$$U[\psi_i] = \sum_i f_i \int \psi_i^* V_{\text{eff}}(\boldsymbol{r}) \psi_i \, \mathrm{d}\boldsymbol{r} \tag{23}$$

where $V_{\text{eff}}(\boldsymbol{r})$ is fixed. We will also assume that the values of α_{ij} for $i \neq j$ are small. This simplifies the analysis in two ways. First, the total energy can then be written to second order in α as

$$\begin{aligned} E^f[\{\alpha_{ij}\}] &= \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{k=1}^{M} f_i \alpha_{ij} H_{jk} \alpha_{ik}^* \\ &= E^f[\{\phi_i\}] + \sum_{i=1}^{M} \sum_{j \neq i}^{M} [(f_i - f_j)\alpha_{ij} H_{ij} + (f_j - f_i)\alpha_{ij}^* \alpha_{ij} H_{ii}] \\ &\quad + \sum_{i=1}^{M} \sum_{j \neq i}^{M} \sum_{k \neq i,j}^{M} [f_i \alpha_{ik}^* \alpha_{ij} H_{kj} - f_j \alpha_{ik} \alpha_{jk}^* H_{ij}] + \mathcal{O}(\alpha^3). \end{aligned} \tag{24}$$

The set of variable $\{\alpha_{ij}\}$ is constrained by the condition that the matrix α be unitary. In principle the constrained first derivative could be found by evaluating the Jacobian. This can be avoided by noting that for small values of $\{\alpha_{ij}\}$ the equations of constraint simplify to:

$$\alpha_{ii} = 1 + \mathcal{O}(\alpha^2) \tag{25}$$

$$\alpha_{ij} = -\alpha_{ji}^* + \mathcal{O}(\alpha^2). \tag{26}$$

This is the second reason for considering only small values of $\{\alpha_{ij}\}$, since it allows us to treat the set of $N(N-1)/2$ variables $\{\alpha_{ij}, i < j\}$ as the independent degrees of freedom.

If we specify that $\{\phi_i\}$ for $i = 1, M$ be the exact eigenvectors of $\hat{H}$, so that $H_{ij} = e_i \delta_{ij}$, then the energy can be written:

$$E^f[\{\alpha_{ij}\}] = E^f[\{\phi_i\}] + \frac{1}{2} \sum_{i,j=1}^{M} (f_i - f_j)(e_j - e_i)\alpha_{ij}^* \alpha_{ij}. \tag{27}$$

We first note that in order for a minimum to exist, we must have $(f_i - f_j)(e_j - e_i) \geqslant 0$ for all i, j. Let us assume for the moment that this is true.

For the case of equal occupation numbers ($f_i = f_j$) the terms that mix states i and j vanish, indicating a symmetry of the system. For the case of unequal occupation numbers ($f_i \neq f_j$ for $i \neq j$), then the total energy is minimized when $\alpha_{ij} = \delta_{ij}$, i.e. when the states, ψ_i, are equal to the eigenstates, ϕ_i. In the case where two states of unequal occupation are separated by a small energy gap, the second derivative, $\partial^2 E / \partial \alpha_{lm}^2$, becomes relatively small. This is the source of the conditioning problems in both the standard and free energy methods of optimization.

A simple fictitious Lagrangian can be introduced in order to examine dynamical quantities:

$$\mathcal{L} = \sum_i \mu_i \int \mathrm{d}\boldsymbol{r} \, |\dot{\psi}_i(\boldsymbol{r})|^2 - E^f[\{\alpha_{ij}\}]. \tag{28}$$

By applying the change of variables, (22), to the fictitious kinetic energy, we can write the Lagrangian solely in terms of the $\{\alpha_{ij}\}$, from which we can the obtain the characteristic frequencies of vibration:

$$\omega_{ij}^2 = \begin{cases} (f_i - f_j)(e_j - e_i)/(\mu_i + \mu_j) & \text{for } i, j \leqslant N \\ f_i(e_j - e_i)/\mu_i & \text{for } i \leqslant N, j > N. \end{cases} \tag{29}$$

Now we return to the energy expression in equation (24) and specify that ϕ_i be a set of orthonormal functions, but not necessarily eigenfunctions. In this general case the first derivative of the energy is

$$\frac{\partial E}{\partial \alpha_{lm}^*} = (f_l - f_m)H_{ml} + (f_m - f_l)\alpha_{lm}(H_{ll} - H_{mm}) + \sum_{j \neq l,m} [(f_m - f_l)\alpha_{jm}H_{jl} + (f_l - f_j)\alpha_{lj}H_{mj}]. \tag{30}$$

This result can be used to formulate an iterative procedure for minimizing the total energy by mixing states within the occupied subspace. Let ϕ_i for $i = 1, N$ be the current set of wavefunctions. Since we are using the current set of wavefunctions as the basis, the current values of α_{lm} are $\alpha_{lm}^0 = \delta_{lm}$. A simple steepest descent algorithm can then be used to find a set α_{lm}^+ closer to the minimum:

$$\alpha_{lm}^+ = \alpha_{lm}^0 - \Delta \left.\frac{\partial E}{\partial \alpha_{lm}}\right|_0 = \delta_{lm} - \Delta[(f_l - f_m)H_{ml}]. \tag{31}$$

A simple preconditioned steepest descent algorithm, which uses second derivative information, can also be devised:

$$\alpha_{lm} = \delta_{lm} - \Delta \frac{H_{ml}}{(H_{mm} - H_{ll})}. \tag{32}$$

These algorithms provide an iterative technique for finding the actual eigenstates of states with differing occupation. Thus the inclusion of extra states allows us to directly attack the conditioning problem by applying a different treatment to the soft modes of the system.

Note that during the course of an iterative minimization calculation the condition for the existence of a minimum, $(f_i - f_j)(e_j - e_i) \geqslant 0$, may not always be satisfied. When this is the case, applying a unitary transformation that exactly diagonalizes H would result in an increase in the energy. The iterative techniques have the advantage that we can set $\alpha_{ij} = 0$ whenever the existence criteria is not satisfied (Gillan 1989).

The preconditioned algorithm is essentially equivalent to the algorithm proposed by Gillan (Gillan 1989). We have observed that the preconditioned version is not stable in the sense that it does not always lower the energy. This may be an indication that the Hessian matrix is not diagonally dominant, in which case simple preconditioning is not applicable.

References

Arias T A, Payne M C and Joannopoulos J D 1992 *Phys. Rev. Lett.* **69** 1077
Bachelet G B, Hamann D R and Schluter M 1982 *Phys. Rev.* B **26** 4199
Blöchl P and Parrinello M 1992 *Phys. Rev.* B **45** 9413
Callaway J and March N H 1984 *Solid State Physics* vol 38 (New York: Academic) p 136
Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
——1988 *Phys. Rev. Lett.* **60** 204
Dunlap B 1984 *Phys. Rev.* A **29** 2902
——1988 *Chem. Phys.* **125** 89
Fernando W, Qian G X, Weinert M and Davenport J W 1989 *Phys. Rev.* B **40** 7985
Fu C-L and Ho K-M 1983 *Phys. Rev.* B **28** 5480
Galli G, Martin R M, Car R and Parrinello M 1990 *Phys. Rev. B* **42** 7470
Gillan M J 1989 *J. Phys.: Condens. Matter* **1** 689
Harris J 1984 *Phys. Rev.* A **29** 1648
Harris J and Hohl D 1990 *J. Phys.: Condens. Matter* **2** 5161
Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
Hohl D and Jones R O 1991 *Phys. Rev.* B **43** 3856

——1992 *Phys. Rev.* B **45** 8995
Janak J F 1978 *Phys. Rev.* B **18** 7165
Kanhere D G, Panat P V, Rajagopal A K and Callaway J 1986 *Phys. Rev.* A **33** 490
Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
Kleinman L and Bylander D M 1982 *Phys. Rev. Lett.* **48** 1425
Mermin N D 1965 *Phys. Rev.* **137** A1441
Oguchi T and Sasaki T 1991 *Prog. Theo. Phys. Suppl.* **103** 93
Pederson M R and Jackson K A 1991 *Phys. Rev.* B **43** 7312
Perrot F and Dharma-wardana MWC 1984 *Phys. Rev.* A **30** 2619
Qian G X, Weinert M, Fernando G W and Davenport J W 1990 *Phys. Rev. Lett.* **34** 1146
Pastore G, Smargiassi E and Buda F 1991 *Phys. Rev.* A **44** 6334
Ryckaert J P, Cicotti G, and Berendsen H J C 1977 *J. Comput. Phys.* **23** 327
Štich I, Car R and Parrinello M 1989 *Phys. Rev. Lett.* **63** 2240
Troullier N and Martins J L 1991 *Phys. Rev.* B **43** 1993
Weinert M and Davenport J W 1992 *Phys. Rev.* B **45** 13709
Wentzcovitch R M, Martins J M, and Allen P B 1992 *Phys. Rev.* B **45** 11372