Ab initio molecular dynamics of metal surfaces

B G Walker[1]**, C Molteni**[2] **and N Marzari**[3]

[1] Theory of Condensed Matter Group, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK
[2] Department of Physics, King's College London, Strand, London WC2R 2LS, UK
[3] Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail: bgw21@cam.ac.uk

Abstract

We review in detail an *ab initio* methodology—Born–Oppenheimer molecular dynamics in the ensemble density-functional approach—able to handle molecular dynamics studies of metal surfaces in a computationally efficient and robust way. A summary of recent applications to the dynamical and thermodynamic properties of sp-bonded metals is presented, focusing on the cases of the Al(110) surface close to its premelting point and of the free surface of liquid Na.

1. Introduction

In the last two decades quantum mechanical computational techniques have become increasingly popular to describe condensed matter at the microscopic level, due to the availability of powerful computers and the development of novel algorithms that take advantage of this increased computational power. *Ab initio* or first-principles methods have thus emerged as a powerful tool to extract detailed electronic and structural information in condensed matter systems, without the need to resort to empirical parameters. Density-functional theory (DFT) is the *de facto* method of choice for large-scale applications, since it often provides an adequate level of accuracy, and requires a computational cost that scales more favourably with the number of atoms than other methods. Applications of such techniques are now common, and state-of-the-art research can take place not only on parallel and vector supercomputers, but also on modestly priced desktop workstations. DFT electronic-structure approaches can be easily combined with classical molecular dynamics techniques to provide an accurate description of thermodynamic properties and phase transitions, atomic dynamics and chemical reactions, and as a tool to coarsely sample the salient features of an unknown potential energy surface.

Dynamical and thermodynamic phenomena at metal surfaces, such as premelting and blocked melting, atomic diffusion, catalytic reactions and the corresponding activated pathways, are all within reach of current capabilities.

In this paper, we shall focus on applications of first-principles molecular dynamics to metal surfaces. Metal surfaces surround us in everyday life, and play a crucial role in industrial processes that range from catalysis to smelting and refining—understanding their microscopic behaviour and eventually designing novel functionalities is of paramount importance. In metals, the balance between energetic and entropic effects at surfaces is more subtle than in the case of semiconductors or insulators, leading to a rich phenomenology in the dependence of structure on temperature. Some of these effects cannot be reproduced without an accurate description of the electronic structure. Moreover, as one takes into account the electronic structure from first principles in an *ab initio* calculation, with no fitting to experimental data, phenomena that may not have been anticipated before the beginning of a simulation can emerge in a natural way.

We will describe in detail the methodology of one approach well suited for metals, namely Born–Oppenheimer molecular dynamics using ensemble density-functional theory. Illustrative results obtained with this approach will be described in the following sections, and alternative approaches mentioned.

2. Overview of methods

2.1. Molecular dynamics

In a molecular dynamics (MD) simulation the microscopic trajectory of each individual particle in the system is determined by the integration of Newton's equations of motion. In classical MD, the system is considered to be composed of massive, pointlike particles, with forces acting between them derived from empirical effective potentials. Most forms of *ab initio* MD make a similar assumption, considering the atomic nuclei as classical particles evolving according to the Newtonian equations of motion. However, the forces acting on the nuclei are deemed quantum mechanical in nature, and are derived from accurate electronic structure calculations. Since atomic nuclei are five orders of magnitude smaller than the atoms themselves, and each nucleon is two thousand times heavier than an electron, such a classical approximation (essentially a Born–Oppenheimer decomposition of the full wavefunction into an electronic wavefunction in the presence of an external electrostatic potential due to the nuclei) is often perfectly appropriate. At very low temperatures, quantum delocalization effects can gain relevance; this can be significant for the lightest elements, such as hydrogen. More importantly, the use of classical equations of motion implies that the vibrational degrees of freedom are not quantized, and follow classical Boltzmann statistics. Such an approximation can be relevant, since at room temperature every normal mode with frequency higher than ~ 300 cm^{-1} will be affected—while in reality high-frequency modes are often effectively in their zero-point motion ground state, in a classical MD simulation they exchange energy back and forth with all other degrees of freedom. Broadly speaking, the Debye temperature provides a reasonable cut-off above which a system can be considered to be accurately described by classical statistics. While path-integral techniques can be used to fully describe these quantum degrees of freedom, the vast majority of applications of quantum mechanical molecular dynamics techniques consider a classical evolution of the nuclei, adiabatically separated from the electrons, with a parametric evolution for the electronic wavefunction always in the ground state with respect to the instantaneous nuclear coordinates. First-principles molecular dynamics was first introduced by Car and Parrinello [1] in 1985; its essential features will be briefly described in section 2.3.

The Car–Parrinello approach introduces an extended Lagrangian to provide a simultaneous evolution of the physical (nuclear coordinates) and computational degrees of freedom (expansion coefficients for the electronic wavefunctions). Often an alternative approach is used (Born–Oppenheimer molecular dynamics), where the electronic degrees of freedom are self-consistently optimized at every ionic time step, trading computational efficiency and a coherent dynamical scheme for a simpler and more resilient algorithmic implementation. Detailed discussion of the practicalities involved in undertaking a molecular dynamics simulation can be found in [2]; for a fuller review of *ab initio* MD, see [3], and for some of the more subtle details involving extended Car–Parrinello Lagrangians see [4] and references therein.

2.2. Born–Oppenheimer molecular dynamics

As mentioned before, in Born–Oppenheimer MD simulations [5] use is made of the fact that the nuclei have masses significantly greater than the electrons. This mass difference means that typical nuclear frequencies are much smaller than electronic ones, so a separation between the nuclear and electronic motion is possible. Adiabaticity is assumed, whereby the electronic system relaxes instantaneously to the ground state for each configuration of the nuclei. The Born–Oppenheimer molecular dynamics (BOMD) scheme thus straightforwardly solves the static electronic structure problem for fixed nuclei at each MD step, and evolves the nuclei according to the classical Newtonian equations of motion.

For this purpose we introduce the Hamiltonian $\mathcal{H}_e$ for the electrons in the field of the classical nuclei (atomic units are used, $\hbar = m_e = e = 4\pi\epsilon_0 = 1$):

$$\mathcal{H}_e(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,i} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{I<J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \tag{1}$$

where $\mathbf{r}$ and $\mathbf{R}$ denote positions of electrons and nuclei respectively. Then the ground state electronic wavefunction Ψ_0 satisfies the stationary Schrödinger equation:

$$\mathcal{H}_e \Psi_0 = E_0 \Psi_0. \tag{2}$$

The force on a nucleus I at time t is given by the total derivative of the potential energy surface. The Hellmann–Feynman theorem [6, 7] states that it can be calculated simply as the expectation value of the 'force' operator:

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle = -\langle \Psi_0 | \nabla_I \mathcal{H}_e | \Psi_0 \rangle. \tag{3}$$

The forces are thus determined by the ground-state structure of the electronic system, and derive exclusively from a combination of the electrostatic repulsion between the nuclei and the electrostatic interactions between the nuclei and the electronic charge density (no other terms enter the gradient of the Hamiltonian with respect to ionic coordinates). All subtle quantum mechanical effects enter the calculation of the forces only through the ground state density they determine. The time step needed for an accurate integration of the equations of motion is completely determined by the nuclear dynamics; the computationally expensive step is the calculation of the electronic ground state at each MD step. Extrapolation techniques for the wavefunctions or for the charge density from one time step to the next [8] can greatly help in accelerating the convergence to self-consistency at each new step.

2.3. Car–Parrinello molecular dynamics

The Car–Parrinello MD scheme [1] reformulates the problem of the coupled electronic and nuclear systems, beginning with an extended Lagrangian in which the electronic degrees of

freedom are included as classical dynamical variables with fictitious masses $\{\mu_i\}$:

$$\mathcal{L}_{\text{CP}} = \frac{1}{2}\sum_I M_I \dot{\mathbf{R}}_I^2 + \frac{1}{2}\sum_i \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle - \langle \Psi_0 | \mathcal{H}_{\text{e}} | \Psi_0 \rangle + \{\text{constraints}\}. \tag{4}$$

The last term accounts for the holonomic constraint of orthonormality of the single-particle electronic orbitals $\{\psi_i\}$:

$$\int \psi_i^*(\mathbf{r}, t)\psi_j(\mathbf{r}, t)\, d\mathbf{r} = \delta_{ij}, \tag{5}$$

(we are considering the wavefunction Ψ_0 in the form of a Slater determinant of single-particle orbitals—as is the case in density-functional theory, described below, or in Hartree–Fock theory). Such an extended Lagrangian leads to the following equations of motion for the coupled electronic and nuclear degrees of freedom, which are propagated simultaneously:

$$\mu_i \ddot{\psi}_i(\mathbf{r}, t) = -\frac{\delta}{\delta \psi_i^*}\langle \Psi_0 | \mathcal{H}_{\text{e}} | \Psi_0 \rangle + \sum_k \Lambda_{ik}\psi_k(\mathbf{r}, t)$$
$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I}\langle \Psi_0 | \mathcal{H}_{\text{e}} | \Psi_0 \rangle. \tag{6}$$

In equations (6), $\{\Lambda_{ik}\}$ are the Lagrange multipliers introduced to account for the constraints of equation (5). If the fictitious masses $\{\mu_i\}$ for the electronic degrees of freedom are chosen to be much smaller than the nuclear masses, then a metastable separation between the evolution of the electronic orbitals and the (much slower) dynamics of the massive nuclei ensues, assuring that the nuclei act only adiabatically as a driving force for the electronic wavefunctions. Oscillations of the orbitals that are coherent with the instantaneous motion of the nuclei are favoured, and the orbitals will evolve remaining close to the ground state. In practical terms, there will always be a trade-off between the choice of smaller electronic orbital masses (which will give a more responsive dynamics and preserve a quasi-adiabatic behaviour but also require a small MD time step) and the requirement of keeping the time step for the integration of the equations of motion as large as possible, in order to decrease the computational cost to evolve the system for a given time.

We note here an important point relevant for applications to metallic systems. The gap in the single-particle energy spectrum determines the largest time step that will maintain the (metastable) adiabatic separation between the electronic and nuclear subsystems during the simulation. In metals there is no gap in the energy spectrum, leading to a severe breakdown of such separation. Several solutions to this problem have been proposed, from the coupling of thermostats to the electronic subsystem [9, 10], to the use of mixed Car–Parrinello/ensemble-DFT approaches [11, 12], to resorting to Born–Oppenheimer molecular dynamics [13, 14] or Born–Oppenheimer/ensemble DFT [15–17].

2.4. First-principles electronic structure: density-functional theory

To perform an *ab initio* molecular dynamics simulation, one requires a first-principles determination of the electronic structure at each time step. This can be obtained using one of a number of first-principles electronic structure methods, among which density-functional theory (DFT) is a popular choice for its accuracy and relatively low computational cost. An algorithmic implementation of DFT that is particularly suited to metallic systems is the *ensemble density-functional theory* (eDFT) approach, that will be described in section 2.7.

The original formulation of density-functional theory for application to quantum systems is embodied in two fundamental pieces of work: the *Hohenberg–Kohn theorems* [18] and the *Kohn–Sham mapping* [19]. Since these techniques have become popular in applications,

a number of comprehensive reviews of DFT and its implementations exist in the literature (e.g. [20, 21]), so only a brief outline covering the main points will be given here. The Hohenberg–Kohn theorems, published in 1964, together with the Kohn–Sham mapping, published in 1965, provide a computationally convenient (and in principle, exact) reformulation of the quantum mechanical problem of a system of N interacting electrons.

The first Hohenberg–Kohn theorem legitimizes the use of the electronic density $n(\mathbf{r})$ as a fundamental quantity to specify the state of a system, by showing that there is a one-to-one correspondence between the charge density of the system of electrons and the external potential $v(\mathbf{r})$ acting on that system. While there was earlier heuristic work focusing on the electron density as a variable for describing the properties—the Thomas–Fermi theory [22–24]—it did not provide qualitatively correct results except for the case of single atoms. The second Hohenberg–Kohn theorem demonstrates that the ground state energy of an electronic system is determined by a functional $E[n(\mathbf{r})]$ of the electronic density; for a given external potential, this functional is a minimum for the ground state density $n_0(\mathbf{r})$, and it gives the ground state energy of the system E_0. In this sense, the second Hohenberg–Kohn theorem represents the recasting of the Schrödinger equation into a variational principle on the charge density only. Denoting the external potential by $v_{\text{ext}}(\mathbf{r})$, the interaction of the electronic charge density with the external potential can be separated out in the total energy functional:

$$E[n(\mathbf{r})] = F[n(\mathbf{r})] + \int v_{\text{ext}}(\mathbf{r})n(\mathbf{r})\,d\mathbf{r}; \tag{7}$$

$F[n]$ is the density functional of DFT, and it can be shown that it does not depend explicitly on the external potential—hence it is a *universal* functional of the charge density. The Hohenberg–Kohn theorems allow us in principle to move away from a description based on the N-body wavefunction (dependent on $3N$ variables) to a description based on the density $n(\mathbf{r})$ alone, a function of only three variables. In practice, even though the density functional $F[n]$ is perfectly well defined, its explicit dependence on the charge density is not known. It is also very difficult to find accurate approximations based on the charge density alone—e.g. the quantum kinetic energy, that is contained in principle in $F[n]$, derives from the curvature of the many-body wavefunction, and such information is all but lost in the charge density itself. Notwithstanding these difficulties, approximations to $F[n]$ have been found that correctly describe several systems, such as sp-bonded metals. For recent applications, see [25–29].

The Kohn–Sham approach allows us to recover an accurate description of the main contributions to the total electronic energy (i.e. the quantum kinetic energy and the electrostatic Hartree contributions), with the help of a mapping of the true system of N *interacting* electrons onto a fictitious system of N *non-interacting* electrons, with the density $n(\mathbf{r})$ of the fictitious system being equal to that of the real one. A set of N orthonormal single-particle orbitals $\{\psi_i\}$ (the Kohn–Sham eigenstates) is thus introduced, such that the density (of both systems) is given by

$$n(\mathbf{r}) = \sum_{i=1}^{N} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}). \tag{8}$$

The universal energy functional $F[n(\mathbf{r})]$ is then decomposed as follows:

$$E[n] = T_s[n] + E_H[n] + \int v_{\text{ext}}(\mathbf{r})n(\mathbf{r})\,d\mathbf{r} + E_{\text{xc}}[n], \tag{9}$$

in which

$$T_s[n] = -\frac{1}{2}\sum_{i=1}^{N}\int \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})\,d\mathbf{r} \tag{10}$$

is the kinetic energy of the non-interacting electronic system; the Hartree term

$$E_{\mathrm{H}}[n] = \frac{1}{2} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{11}$$

takes account of the Coulomb interaction of the charge density, and the external potential $v_{\mathrm{ext}}(\mathbf{r})$ is that due to the nuclei,

$$v_{\mathrm{ext}}(\mathbf{r}) = -\sum_{i} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|} + \sum_{I<J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{12}$$

The final term E_{xc} in equation (9) is the *exchange–correlation* energy, operationally defined as the difference between the first three terms on the right-hand side of equation (9) and the exact energy of the system. The form of this universal exchange–correlation functional is unknown, and it contains most of the many-body complexity of the interacting electron gas. Still, its value in absolute terms is comparatively small, and, most importantly, simple approximations to it have proved to be reasonably accurate for a great variety of applications.

According to the second Hohenberg–Kohn theorem, the energy functional $E[n]$ in equation (9) must be minimized with respect to the density $n(\mathbf{r})$ (subject to the constraint of charge conservation) to find the ground state density and total energy. The minimum condition can be written as

$$\delta\left[T_{\mathrm{s}}[n] + E_{\mathrm{H}}[n] + \int v_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r})\,\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[n] - \mu\left(\int n(\mathbf{r})\,\mathrm{d}\mathbf{r} - N\right)\right] = 0, \tag{13}$$

where the Lagrange multiplier μ has been introduced to take account of charge conservation and orthonormality of the orbitals. Defining the Kohn–Sham effective potential $v_{\mathrm{KS}}(\mathbf{r})$,

$$v_{\mathrm{KS}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}\mathbf{r}' + v_{\mathrm{xc}}(\mathbf{r}), \tag{14}$$

in which

$$v_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n(\mathbf{r})]}{\delta n(\mathbf{r})}, \tag{15}$$

the minimization problem in equation (13) becomes

$$\frac{\delta T_{\mathrm{s}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} + v_{\mathrm{KS}}(\mathbf{r}) = \mu. \tag{16}$$

Finally, evaluating the functional derivative, the Euler–Lagrange equations for the constrained minimization problem, known as the *Kohn–Sham equations*, are obtained:

$$\left[-\tfrac{1}{2}\nabla^2 + v_{\mathrm{KS}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}). \tag{17}$$

A common approach to solving the Kohn–Sham equations is by iterative evolution of some 'reasonable' trial set of orbitals. Alternatively, the minimization problem they represent (that is finding the minimum of the total energy functional) can be solved by direct methods, such as the conjugate gradient technique [20].

As noted above, the form of the exchange–correlation functional in equation (9) is unknown, so an approximation is needed. A number of such approximations for the exchange–correlation energy functional have been published; popular and fairly reliable ones are the *local density approximation* (LDA), that was in fact originally suggested in the paper of Kohn and Sham [19], and the *generalized gradient approximation* (GGA) [30]. The local density approximation considers the exchange–correlation energy in each volume element $\mathrm{d}\mathbf{r}$ at $\mathbf{r}$ (where the density is $n(\mathbf{r})$) to be that of a homogeneous electron gas of density n:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n(\mathbf{r})] = \int \epsilon_{\mathrm{xc}}(n)n(\mathbf{r})\,\mathrm{d}\mathbf{r}, \tag{18}$$

where $\epsilon_{xc}(n)$ is the exchange–correlation energy per unit density of a homogeneous electron gas at constant density n. This has been obtained by interpolation between results from quantum Monte Carlo simulations [31] and high density asymptotic limits [32]. Despite its simplicity, it has been found that the local density approximation works surprisingly well in many practical applications. As their name suggests, the generalized gradient approximations include terms dependent on gradients of the charge density in the expression for the exchange–correlation functional.

The Hohenberg–Kohn–Sham density-functional theory formalism described above has become very popular for implementation in computational codes and a great number of practical applications have been made using it (for an extensive review of applications see [3]).

2.5. Plane waves and pseudopotentials

A number of additional techniques are often utilized when performing practical calculations within the density-functional theory described in section 2.4; an important group which generally go together are plane-wave basis-set expansions, fast Fourier transforms and pseudopotentials. These will be discussed briefly below but are reviewed in greater detail in [20].

Expansion of wavefunctions and potentials in plane-wave basis sets is an approach having its roots in solid state theory. In a crystal the periodicity of the lattice means that the potential and charge density also have that same periodicity. When periodic boundary conditions are in effect, Bloch's theorem [33] allows expansion of the Kohn–Sham eigenstates $\{\psi_i\}$ in a discrete set of plane waves and the Kohn–Sham equations (17) to be written in their reciprocal space form:

$$\sum_{\mathbf{G}'} \left[\tfrac{1}{2} |\mathbf{k} + \mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + v_{KS}(\mathbf{G} - \mathbf{G}') \right] c_{i,\mathbf{k}+\mathbf{G}'} = \epsilon_i c_{i,\mathbf{k}+\mathbf{G}}, \tag{19}$$

where $c_{i,\mathbf{k}+\mathbf{G}}$ are expansion coefficients for the KS orbitals in the plane wave basis, $\mathbf{G}$ and $\mathbf{G}'$ are reciprocal lattice vectors for the system (determined by the periodic boundary conditions), and $v_{KS}(\mathbf{G} - \mathbf{G}')$ is the Fourier transform of the Kohn–Sham potential of equation (14).

As can be seen in equation (19), the kinetic energy operator is diagonal in reciprocal space; conversely, the potential energy acts diagonally in real space. The existence of efficient algorithms for performing conversions between the real and reciprocal spaces, the *fast Fourier transforms* (FFTs)[4], means it is computationally advantageous to evaluate the two parts of the energy in the spaces in which they are respectively diagonal and then convert between the two spaces by applying an FFT to obtain the final result.

Plane wave basis sets are not only useful for calculations where the system being considered possesses a full three-dimensional periodicity as in a pure crystal. They can also be used where there is a partial periodicity (such as for surfaces, where there is periodicity parallel to the surface) and even for inherently aperiodic systems such as molecules, provided the system of interest is embedded in a periodic 'supercell' that has to be large enough to reduce the interactions between periodic images of the system to an acceptable level.

The pseudopotential approach [34–36] exploits the natural separation in energy scales that exists between the 'valence' electrons, that are the loosely bound electrons participating in chemical bonding, and the 'core' electrons, that are much more tightly bound and largely unaffected by changes in the chemical environment. A pseudopotential replaces the combined effects of nucleus and (frozen) core electrons, and is constructed to reproduce the scattering of the valence electrons that the 'all-electron' core would give. Since the core electrons have

---

[4] A useful source of information about fast Fourier transform routines is the website http://www.fftw.org/

been effectively removed from the calculations, the pseudopotential is much smoother than the original unscreened Coulomb potential. The approach is critical when a plane wave basis set is used for the calculations. Since the valence electrons must have very high frequency spatial oscillations close to the nuclei to maintain orthogonality with the lower lying core electrons, the number of plane waves that would be required to accurately describe this region would be prohibitive. Replacing the nucleus and core electrons by a pseudopotential smoothes out the Coulombic divergence, and eliminates the need for the orthogonality oscillations. Various schemes have been devised to generate pseudopotentials that are both accurate and economical for use in *ab initio* electronic structure calculations—for a detailed description of the established algorithms for generating norm-conserving and 'ultrasoft' pseudopotentials, see [37–39].

2.6. Finite-temperature DFT extensions

The canonical, finite-temperature extension of density-functional theory was developed by Mermin [40] soon after the first papers by Hohenberg, Kohn, and Sham. A comprehensive description can be found in [41]. As noted earlier, the standard formulation of quantum mechanics for a system of N electrons uses the N-body wavefunction $|\Psi_N\rangle$ to describe the system. If the system can be represented in this way, it is said to be in a *pure state*. Alternatively, the theory can be formulated in terms of a *density operator*, defined by

$$\hat{\gamma} = |\Psi_N\rangle\langle\Psi_N|. \tag{20}$$

Expectation values of operators, for instance the energy of the system, are determined as traces:

$$E = \langle\hat{H}\rangle = \langle\Psi|\hat{H}|\Psi\rangle = \mathrm{Tr}(\hat{\gamma}\hat{H}) = \mathrm{Tr}(\hat{H}\hat{\gamma}). \tag{21}$$

If the system of interest is part of a larger system, and the total system cannot be described as a superposition of pure states, i.e. it cannot be characterized by a wavefunction, then it is said to be in a *mixed state*, and description using a density operator becomes mandatory. In this case, the *ensemble density operator* [41] is used:

$$\hat{\Gamma}_N = \sum_i w_i |\Psi_i\rangle\langle\Psi_i|. \tag{22}$$

It is a sum over all the pure states $\{|\Psi_i\rangle\}$ with associated weights $\{w_i\}$. The $\{w_i\}$ appearing in equation (22) represent a statistical distribution among the available pure states, resulting from a combination of the probabilistic nature of the wavefunction and an imperfect knowledge of the system. As with a system in a pure state, expectation values for systems in mixed states can be obtained using the density operator (so statements such as equation (21) also apply for systems in mixed states).

　　For studies at finite temperature, the system of interest will be coupled to a heat bath. As microscopic details of the interaction between the system of interest and the heat bath will not generally be known, specification of the exact Hamiltonian for the complete system is impossible, so the system of interest will be in a mixed state and use of an ensemble density operator is necessary. For finite-temperature considerations, free energies are the pertinent quantities, so an expression for entropy is needed. The one arising by analogy with classical statistical mechanics [42] is

$$S = -k_B \sum_i w_i \ln w_i \tag{23}$$

$$= -k_B \mathrm{Tr}(\hat{\Gamma}_N \ln \hat{\Gamma}_N). \tag{24}$$

If we consider the canonical ensemble, which is a mixture of pure states all having the same particle number N, the free energy of interest is the *Helmholtz free energy* ($A = E - TS$),

constructed in the density operator formalism as

$$A[\hat{\Gamma}_N] = \text{Tr}\,\hat{\Gamma}_N \left(\frac{1}{\beta} \ln \hat{\Gamma}_N + \hat{H} \right), \tag{25}$$

with the standard notation $\beta = 1/k_B T$. This free energy is a functional of the density operator, and a variational principle can be proved [42]:

$$A[\Gamma_N^0] \leqslant A[\Gamma_N], \tag{26}$$

in which Γ_N^0 is the equilibrium density operator. From here, the derivation of a canonical finite-temperature density-functional theory follows the Hohenberg–Kohn theory described in section 2.4 and a Kohn–Sham type mapping onto a non-interacting reference system can be made. In the finite-temperature development, we allow fractional occupation numbers $f_i \in [0, 1]$, so that the density is given by

$$n(\mathbf{r}) = \sum_{i=1}^{\infty} f_i |\psi_i(\mathbf{r})|^2. \tag{27}$$

Taking the Fermi–Dirac distribution for the occupancies,

$$f_i = f\left(\frac{\epsilon_i - \mu}{T} \right) = \left(1 + \exp\left(\frac{\epsilon_i - \mu}{k_B T} \right) \right)^{-1}, \tag{28}$$

the entropy term can be evaluated:

$$S = -k_B \sum_i [f_i \ln f_i + (1 - f_i) \ln(1 - f_i)]. \tag{29}$$

Accordingly, the free energy functional is

$$A[T; \{\psi_i\}, \{f_i\}] = T_s[n] + E_H[n] + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r})\, d\mathbf{r} + E_{\beta,\text{xc}}[n]$$
$$+ \frac{1}{\beta} \sum_i [f_i \ln f_i + (1 - f_i) \ln(1 - f_i)], \tag{30}$$

where the subscript β signifies that quantities are being considered at finite temperature. The canonical Mermin–Kohn–Sham equations that result from minimization of the functional in equation (30) are

$$\left[-\tfrac{1}{2}\nabla^2 + v_H(\mathbf{r}) + v_{\beta,\text{xc}}(\mathbf{r}) + v_{\text{ext}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \tag{31}$$

and the charge density is now given by equation (27).

2.7. Ensemble DFT

Ensemble density-functional theory is an extension of the DFT schemes described in the last section. We consider a set of generalized occupancies $\mathbf{f} = \{f_{ij}\}$, being the representation of the density operator $\hat{\Gamma}$ in the basis set of the orbitals $\{\psi_i\}$ [43]. Equation (30) becomes

$$A[T; \{\psi_i\}, \{f_{ij}\}] = \sum_{ij} f_{ji} \langle \psi_i | \hat{T}_e + \hat{V}_{\text{ext}} | \psi_j \rangle + E_H[n] + E_{\beta,\text{xc}}[n] - T S[\{f_{ij}\}] \tag{32}$$

$$n(\mathbf{r}) = \sum_{ij} f_{ji} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}), \tag{33}$$

where the matrix elements that only depend on the orbitals have been separated out and placed in the first term; $\hat{V}_{\text{ext}}$ appears there since in general it will be generated by an array of non-local pseudopotentials. The entropy term S appearing in equation (32) may be that derived from the

Fermi–Dirac distribution already described, or obtained from one of the generalized smearing methods to be described in section 2.9.

It can be shown that the free energy functional and the charge density do not change following a unitary transformation of the orbitals (this is because they are written as traces) and so they are *representation covariant*. We can then define the projected functional G, a well defined functional of the orbitals only, as

$$G[T; \{\psi_i\}] = \min_{\{f_{ij}\}} A[T; \{\psi_i\}, \{f_{ij}\}]. \tag{34}$$

This projected functional is *invariant* for any unitary transformation of the orbitals, and does not depend explicitly on the occupation matrix—it is as well conditioned as the total energy functional in the case of a semiconductor or insulator [43]. The definition of the projected functional G allows the construction of a practical minimization strategy. The minimization of the free energy functional A breaks down naturally into a two-loop nested scheme: there is an inner loop on the occupancies $\{f_{ij}\}$, where the functional G is calculated by bringing A to self-consistency with respect to a fixed set of orbitals $\{\psi_i\}$; this is followed by an outer step to update the orbitals. The strategy for performing the inner loop minimization on the occupancy matrix is modelled on the non-self-consistent problem; in that case, the required set of occupancies would be obtained by diagonalizing the Hamiltonian, applying the thermal distribution and then rotating back to the orbital representation. This non-self-consistent solution is used as the *search direction* in a direct line-minimization in the space of the $\{f_{ij}\}$ for updating the occupation matrix. We introduce the following notation for the matrix elements of the Hamiltonian that depend on just the orbitals or on the density respectively:

$$h_{ij} = \langle \psi_i | \hat{T}_s + \hat{V}_{\text{ext}} | \psi_j \rangle \tag{35}$$

$$v_{ij} = \langle \psi_i | \hat{v}_{\text{H}}[n] + \hat{V}_{\text{xc}}[n] | \psi_j \rangle. \tag{36}$$

After every update of the orbitals, the matrix elements $\{h_{ij}\}$ are calculated, and they remain unchanged during the inner loop on the occupancies, as the orbitals do not vary during the occupancy search. The charge density (for the mth iteration in the occupancy loop) is calculated:

$$\rho^{(m)}(\mathbf{r}) = \sum_{ij} f_{ji}^{(m)} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}). \tag{37}$$

This is the most computationally expensive part of the occupancy update, and it is efficiently calculated in the rotated representation in which the occupations are diagonal. From this density, the density-dependent terms $E_{\text{H}}^{(m)} + E_{\text{xc}}^{(m)}$, $V_{\text{H}}^{(m)} + v_{\text{xc}}^{(m)}$ and subsequently $V_{ij}^{(m)}$ are determined. The entropy term is obtained via the diagonalization of $\mathbf{f}$:

$$f_{ij}^{(m)} = \sum_l Y_{il}^{(m)\dagger} f_l^{(m)} Y_{lj}^{(m)}. \tag{38}$$

The Hamiltonian is updated and diagonalized:

$$H_{ij}^{(m)} = h_{ij} + V_{ij}^{(m)} = \sum_l Z_{il}^{(m)\dagger} \epsilon_l^{(m)} Z_{lj}^{(m)}. \tag{39}$$

In a non-self-consistent problem, the solution for $\mathbf{f}$ would be obtained by applying the thermal distribution to the eigenvalues of the Hamiltonian as

$$\tilde{f}_{ij}^{(m)} = \sum_l Z_{il}^{(m)\dagger} f_{\text{T}}(\epsilon_l^{(m)} - \mu) Z_{lj}^{(m)}, \tag{40}$$

where f_{T} is the thermal distribution being used. Although this is not the solution to the self-consistent problem here, it provides the search direction in the f_{ij} space: $\Delta \mathbf{f}^{(m)} = \tilde{\mathbf{f}}^{(m)} - \mathbf{f}^{(m)}$. Taking a step of size β along this search direction, the occupancies become

$$\mathbf{f}_\beta^{(m+1)} = \mathbf{f}^{(m)} + \beta \Delta \mathbf{f}^{(m)}. \tag{41}$$

This search is unconstrained as the trace of **f** is equal to N at both end-points $\beta = 0$ and 1, and by linearity at all points in between.

In the outer orbital loop, line minimization based on a conjugate gradient scheme with preconditioning works well. The use of a preconditioned formulation is driven by the broader spectrum of the Hamiltonian $\{f_i \epsilon_i\}$ (rather than the $\{\epsilon_i\}$ as would be the case for semiconductors/insulators), which contains values close to zero that cause significantly slower convergence than would be the case in a system having a gap in its energy spectrum. With the set of scaled variables $\tilde{c}_{in} = \sqrt{f_i} c_{in}$, the preconditioned gradients for the original variables would be

$$-\frac{\delta G}{\delta \psi_i^*} \longrightarrow -\frac{1}{f_i} \frac{\delta G}{\delta \psi_i^*} = -\hat{H} \psi_i. \tag{42}$$

To generalize to the non-diagonal case at hand, the gradients of G are calculated in the original representation, passed into the diagonal representation, preconditioned according to equation (42), and then transformed back into the original representation. The gradients in the original representation are

$$\mathbf{g}_i = -\frac{\delta G}{\delta \psi_i^*} = -\sum_j f_{ji} \hat{H} \psi_j, \tag{43}$$

and in the diagonal representation (defined by $\mathbf{f}' = f_{ii}' \delta_{ij} = \mathbf{U} \mathbf{f} \mathbf{U}^\dagger$) they are

$$\mathbf{g}_i' = -\frac{\delta G}{\delta \psi_i'^*} = -\sum_j f_{ji}' \hat{H} \psi_j' = -f_{ii}' \hat{H} \psi_i'. \tag{44}$$

Applying the preconditioning, the gradients in the diagonal representation are

$$\mathbf{G}_i' = -\hat{H} \psi_i' = -\hat{H} \left(\sum_m U_{im}^* \psi_m \right). \tag{45}$$

Finally, rotating back, the preconditioned gradients in the original non-diagonal representation are

$$\mathbf{G}_i = \sum_n U_{in}^{\dagger*} \mathbf{G}_n' = -\sum_m \left(\sum_n U_{in}^{\dagger*} U_{nm}^* \right) \hat{H} \psi_m = -\hat{H} \psi_i. \tag{46}$$

Use of the preconditioned gradients $\mathbf{G}_i$ to construct conjugate directions (details of which are provided in [44]) provides greatly improved convergence, updating the higher bands with the same speed as the lower ones, and provides a simple expression for the computation of the gradients. In addition, when a plane wave basis set is used in the computation, a standard kinetic preconditioning is applied, so that the procedure for updating the orbitals is said to be doubly preconditioned.

In earlier schemes, the occupancies and rotations of the orbitals within the occupied subspace were treated separately. The ensemble DFT approach described here treats them consistently as part of the same problem. Also, as the scheme is variational, the free energy is bound to decrease at each step (the importance of a consistent variational minimization of the free energy functional was stressed for instance in [45]). Due to these properties, the instabilities encountered in applying direct minimization approaches to metals (see section 2.8) are removed, providing a rapid and stable convergence to the ground state, and quick convergence of the Hellmann–Feynman forces, as is necessary for efficient structural relaxations and molecular dynamics simulations [15–17, 43].
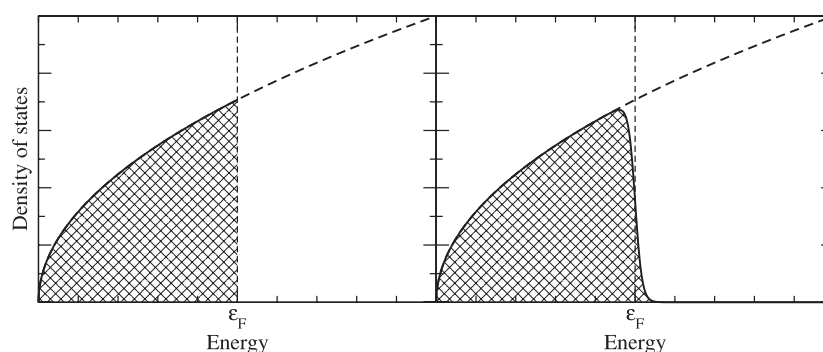
Figure 1. Schematic illustration of the density of states in a metal. The left-hand panel shows the DOS without smearing; the right-hand panel shows it with smearing.

2.8. Issues for calculations on metallic systems

One major problem in the case of metals arises from the occurrence of discontinuities when crossing the Fermi surface. When using periodic boundary conditions, all observables, such as the charge density of the system, become integrals over the first Brillouin zone in reciprocal space. These integrals are evaluated in practice by summation over a finite number of k-points, and the computational cost is linearly proportional to the number of points in the mesh. Systematic methods have been given [46, 47] for generating computationally expedient meshes of k-points. Unfortunately, when the function being integrated is discontinuous, the accuracy of a finite set of k-points in the Brillouin zone decreases dramatically.

This is exactly the situation in a metal at zero temperature: states inside the Fermi surface must be included in the integrals, and states outside, where the occupation is zero, do not contribute at all. The left-hand panel of figure 1 illustrates the integration of the density of states up to the Fermi energy required for calculating properties of the system; the sharp discontinuity in the occupation number at the Fermi energy means integration of a discontinuous function must be performed.

A solution is to introduce a finite electronic temperature, that results in a smearing of the Fermi surface, and greatly improves the accuracy of sampling in the Brillouin zone. The right-hand panel of figure 1 shows a metallic density of states that is broadened near the Fermi energy and demonstrates how the discontinuity there is smoothed. This will improve the accuracy of integrals evaluated using a coarser set of k-points.

The absence of an energy gap in metals also causes problems for application of the Car–Parrinello method to their study. As mentioned earlier, in the CP scheme, where a metastable adiabatic separation between the electronic and ionic subsystems is required, the largest time step that will maintain that separation is related to the energy gap. In the case of a metal, the absence of a gap means there is always significant overlap between the typical frequencies of the ionic and electronic subsystems, resulting in an immediate breakdown of the adiabatic separation. As mentioned before, solutions proposed include coupling thermostats to the electronic subsystem [9, 10], and the use of mixed Car–Parrinello/ensemble-DFT approaches [11, 12, 16]. A final cautionary note is that the Born–Oppenheimer separation requires further justification for application to metallic systems as the electronic excitations overlap with the ionic frequencies; Migdal's theorem [48, 49] is needed to show the validity of using the Born–Oppenheimer separation in a metal.

2.9. Generalized smearing schemes and cold smearing

Use of a finite-temperature formulation as a device to improve sampling of the Brillouin zone in calculations on metals, as discussed in the previous section, was first introduced by Fu and Ho [50]. The systematic errors introduced into a calculation by the use of such a finite-temperature 'smearing' of the Fermi surface can be corrected *a posteriori*, as has been extensively discussed in [45, 51]; while correction of the energies is straightforward, corrections for all other expectation values (e.g. forces) would require full linear-response calculations, which would be impractical for example in a molecular dynamics simulation.

Here we show how generalized smearing schemes can minimize the systematic errors introduced by a finite-temperature formulation. An extensive discussion of this topic can be found in chapter 4 of [44]. It was shown in [44, 51, 52] how finite-temperature formulations can be generalized by introducing arbitrary and general classes of free energies and entropies. The procedure is to choose the equilibrium thermal distribution (with suitable properties), and then derive the form of the entropy consistent with it. An inverse approach is thus pursued: instead of starting from a counting relation (the Pauli principle) and determining an equilibrium distribution (the Fermi–Dirac) from the requirement of maximizing the entropy (as in statistical mechanics), with the entropy itself a measure of the number of states available, the equilibrium distribution is arbitrarily chosen, and the form of the entropy that follows from the minimization requirement is derived. The choice is made that the (fictitious) thermal distribution can be written as

$$f(x) = \int_{-\infty}^{x} g(t)\,\mathrm{d}t, \tag{47}$$

where $g(t)$ is a broadening function normalized to 1 (or 2, in the case of spin degeneracy). This relation provides an operative definition of the fictitious temperature as an *integrated broadening*, with full freedom to choose the function g, as long as the usual physical constraints on the occupancies are satisfied. This relation can be propagated to the entropy; in the non-interacting picture the entropy is a linear combination of single-particle terms

$$S[\{f_i\}] = \sum_i S_i = \sum_i S(f_i), \tag{48}$$

and so a functional form for the entropy can be determined by integrating:

$$\frac{\mathrm{d}S}{\mathrm{d}f} = \frac{\epsilon - \mu}{\sigma} = -x \Rightarrow \frac{\mathrm{d}S}{\mathrm{d}x} = -x\frac{\mathrm{d}f}{\mathrm{d}x} \tag{49}$$

$$f(x) = \int_{-\infty}^{x} g(t)\,\mathrm{d}t \Rightarrow S_i = \int_{-\infty}^{x_i} -tg(t)\,\mathrm{d}t. \tag{50}$$

The previous relation provides an operative connection between the choice of a broadening function and the actual form of the entropy; as an example, if the following form for $g(t)$ is chosen (assuming spin-degeneracy),

$$g(t) = \frac{2}{(e^{\frac{t}{2}} + e^{-\frac{t}{2}})^2}, \tag{51}$$

the entropy that is derived according to the procedure above is

$$S = -2\sum_i [y_i \log y_i + (1 - y_i)\log(1 - y_i)], \tag{52}$$

with

$$y_i = \frac{f_i}{2} = \frac{1}{e^{\frac{\epsilon_i - \mu}{\sigma}} + 1} = \frac{1}{e^{-x_i} + 1}. \tag{53}$$

In addition to recovering this well known result, the freedom in choosing the broadening function can be exploited to identify arbitrary generalized free energies as functions of the thermal distributions. For instance, it is often convenient to employ in practical calculations [53] a Gaussian entropy,

$$S = \frac{1}{\sqrt{\pi}} \sum_i e^{-\left(\frac{\epsilon_i - \mu}{\sigma}\right)^2},$$ (54)

as obtained from the broadening:

$$g(t) = \frac{2}{\sqrt{\pi}} e^{-t^2}.$$ (55)

One problem of generalizing beyond the Fermi–Dirac distribution is that in the generalized schemes an important result is lost: the entropy is no longer expressible as a function of the occupation numbers f_i, but only as a function of their inverse under the equilibrium thermal distribution, namely the ϵ_i. It should be noted that the self-consistency condition $(d(E - TS)/df = 0)$ implies that

$$\frac{dS}{df_i} = -x_i,$$ (56)

i.e. the derivative of the entropy is equal to minus the normalized distance between the Fermi energy and the inverse of the occupation; this relation is useful if the steepest-descent directions on the occupancies are needed.

An alternative strategy for applying a 'thermal' smearing is to convolve the exact density of states with some smooth approximation $\tilde{\delta}(x)$ to the Dirac delta function. The approach taken by Methfessel and Paxton [54] in deriving a broadening function was along these lines and involved use of the first N terms in the expansion of the delta function in Hermite polynomials. The resulting smearing term has the form of a quadratic polynomial multiplied by e^{-x^2} and leads to a dependence of the free energy on the temperature that is quartic. A significant problem with this approach is that the occupation numbers are not positive definite, so negative occupancies are possible and such negative occupancies can lead to unphysical results. This is especially important in the case of metal surfaces, where states that are close to the Fermi energy, and thus possibly having negative Methfessel–Paxton occupancies, are also those most likely to spill into the vacuum, which leads to an overall 'negative' region of density. The flexibility in choosing broadening functions and consistent entropies allows a formulation combining the advantage of the Methfessel–Paxton scheme (i.e. systematic elimination of the quadratic errors) while avoiding its most troublesome consequence for self-consistent calculations (i.e. the introduction of negative occupation numbers). The *cold smearing* [44] is a choice of broadening function that remedies the problem of negative occupancies associated with the Methfessel–Paxton smearing. The derivation of the broadening functional is subject to these constraints:

(i) $\int_{-\infty}^{\infty} dx\, \tilde{\delta}(x) = 1$ normalization;

(ii) $\int_{-\infty}^{\infty} dx\, x \tilde{\delta}(x) = 0$ zero entropy at zero temperature;

(iii) $\int_{-\infty}^{\infty} dx\, x^2 \tilde{\delta}(x) = 0$ no second order term in entropy (cold smearing);

(iv) $\int_{-\infty}^{t} \tilde{\delta}(x)\, dx \geqslant 0$ positive occupancies.

The simplest form that can be used turns out to be a third degree polynomial:

$$\tilde{\delta}(x) = \frac{1}{\sqrt{2}} \left(ax^3 - x^2 - \frac{3}{2}ax + \frac{3}{2} \right) e^{-x^2}$$ (57)

where there is some freedom in the value of the parameter a that ensures the occupancies are positive definite; the value that has been used in the work to be described in section 4

is $a = -0.5634$. A less 'structured' and somewhat more elegant alternative expression was introduced in [15]:

$$\tilde{\delta}(x) = \frac{2}{\sqrt{\pi}} e^{-[x-(1/\sqrt{2})]^2}(2 - \sqrt{2}x). \tag{58}$$

It is interesting to note that the flexibility of generalized entropic formulations can be used to recover target finite-temperature (e.g. Fermi–Dirac) distributions, instead of the zero-temperature limit; a clear description of the approach is provided in [55].

2.10. Other work on metal surfaces

An impressive body of work has emerged on the chemistry and thermodynamics of metal surfaces in the past years, motivated primarily by their central importance as catalytic materials. Most of this work is based on static calculations and ionic relaxations, and so is not reviewed here in detail. Very robust iterative self-consistent techniques (e.g. [14, 56]) have nevertheless been central to the widespread application of density-functional approaches. For recent applications the reader is referred to [57–61] and references contained therein.

3. Metal surfaces at finite temperature

The thermodynamic stability of metal surfaces presents a very rich phenomenology [62, 63]: whereas the exotic reconstructions of semiconductor surfaces, driven by the strong covalent dangling bonds, are missing, the interplay of subtle energetic and entropic effects leads to a variety of different temperature-dependent behaviours. At lower temperatures surfaces can be either *relaxed* with respect to the ideal bulk geometry, or *reconstructed* in some simple way (e.g. in the *missing row* transition). The zero-temperature relaxation can often be understood in terms of simple models based on charge-density considerations [64]; for instance, the *contraction* of the inter-layer distance of the surface atoms found in the less packed, open surfaces (such as the fcc (110)) can be understood as a mechanism to increase the density surrounding the surface atoms, that 'dive' towards the bulk.

Higher temperatures enhance the imbalance between the different anharmonic contributions, and lead to a pattern of relaxations or de-relaxations of the inter-layer spacings [65–69]. In addition, entropic effects can lead to a *deconstruction* of the eventual reconstructed phase. At increasing temperatures the more open surfaces can start to develop defects (e.g. the formation of adatoms and vacancies), while the more close-packed surfaces, having a higher defect formation energy, can maintain ordered coherence up to the melting point. The (111) surfaces of fcc metals can be 'overheated', meaning they remain ordered even at temperatures above the bulk melting point. These different behaviours with temperature can be expected to strongly influence the free energy barriers and the mechanisms of diffusion for adatoms present on the surfaces. In addition, some peculiar phase transitions of the surface happen whenever the free energy for the formation of overlayer structures vanishes. Notably, a *roughening* transition [70] can occur, characterized by a diverging height–height correlation along the surface, that takes the form of a terraced landscape. This could be preceded by a *pre-roughening* transition [71, 72] in which disordering is confined to the surface overlayer and the long-range stability is preserved, in what is a *disordered flat* phase.

Before reaching the melting temperature of the bulk system, it is possible that a self-wetting initiates on the surface, as a consequence of the half-balanced periodic potential that the surface ions experience, and of the increased anharmonicity around the equilibrium positions [73]. The phenomenology of the surface melting and pre-melting is quite complex [63]; different

surfaces, or different orientations of the same system, can exhibit a complete melting, where the thickness of the liquid phase increases as the bulk transition temperature is approached, or just a limited, blocked melting of the upper layers. The characterization of these processes is a current area of research, with focused experimental [74, 75] and theoretical [66, 67, 76] efforts.

4. Applications

Two applications will be discussed here, with the goal of showing the efficiency and microscopic understanding provided by the Born–Oppenheimer MD/ensemble-DFT first-principles approach we have described.

4.1. Thermal contraction and premelting of Al(110)

Several experimental studies of the finite-temperature properties of Al metal surfaces have appeared [77–81], together with theoretical works based on semi-empirical methods [76, 82], or restricted to zero-temperature total energy calculations [83] (that are in themselves a fundamental source of understanding). The Al(110) surface is unreconstructed, and displays a pattern of damped oscillatory interlayer relaxations. Experimental LEED (low-energy electron diffraction) measurements on Al(110) indicate a negative thermal expansion of the first interlayer distance, and a large positive expansion for the second interlayer distance, at variance with earlier computer simulations. For the case of Al(110), several experimental techniques have shown the appearance of a disordered, liquid layer at temperatures between 770 and 815 K [77–81], whereas the bulk melting temperature is 933 K. On Al(110) this *premelting* transition is preceded by an anomalous proliferation of adatoms on the surface, for which there is no reliable microscopic picture.

The study of Al(110) reported in [43] involved eDFT molecular dynamics simulations in a range of temperatures between 400 and 900 K. Besides reproducing the observed low-temperature pattern of interlayer relaxations, it provided a theoretical prediction of a *thermal contraction* for the outermost interlayer distance, for which solid experimental evidence has now been accumulated [84–87]. The origin of this anomalous contraction lies in the microscopic dynamics at the surface and in the layer below. As shown in figure 2, where the molecular dynamics trajectories for a sample run are plotted, the mean square displacements of atoms in the second layer in the direction perpendicular to the surface end up being larger than those of the atoms in the first layer. This strongly anisotropic dynamics, reported in figure 3, leads to a very anharmonic behaviour—in particular, the dependence of the mean square displacements in the second layer on temperature is stronger than linear. As the temperature increases, the second-layer atoms oscillate well beyond the Lindemann criterion for melting, creating a number of adatom/vacancy pairs that are responsible for the disordering and then melting of the surface.

4.2. Liquid metal surfaces

The second application of the BOMD/eDFT *ab initio* scheme that we will discuss is the study of liquid sodium surfaces [17, 88]. This work represents the first MD study of the surface of a typical liquid metal using a conventional version of DFT, though very recently MD simulations of lithium and sodium liquid surfaces by orbital-free DFT (which is an approximate form of DFT with reduced computational cost) have appeared in the literature [29]. Previously the
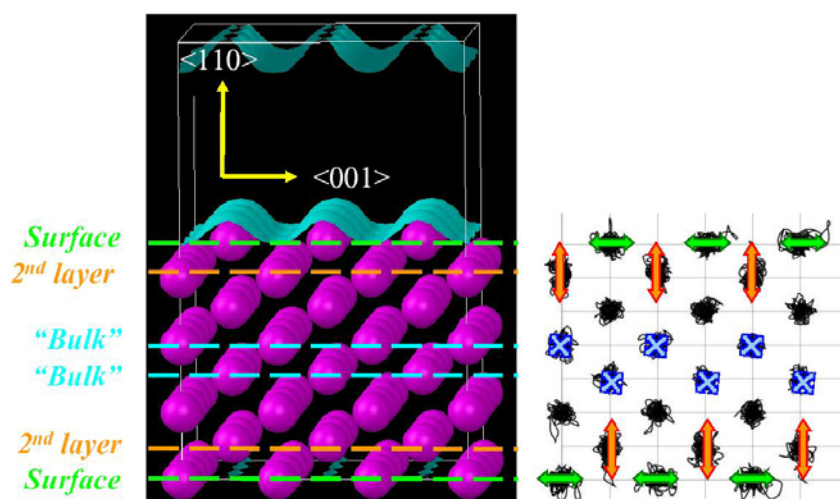
Figure 2. Left-hand panel: eight-layer Al(110) supercell used in the simulations. Right-hand panel: MD trajectories during a simulation, highlighting the anisotropic mean square displacements in the surface and in the second layer.
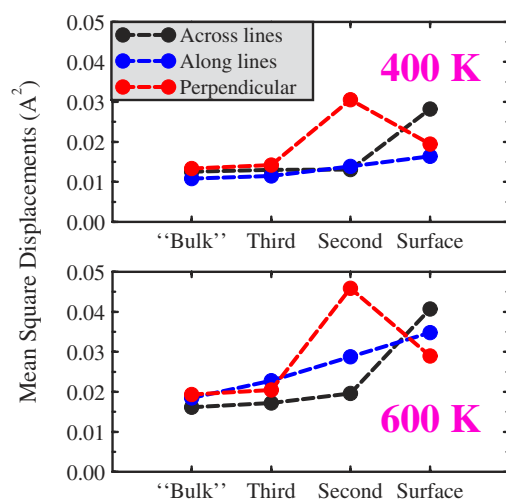

Figure 3. Layer-resolved mean square displacements at 400 and 600 K for the three principal crystallographic directions: across the widely spaced atomic lines on the Al(110) surface (i.e. the ⟨001⟩ directions), along those lines (⟨010⟩ directions), and perpendicular to the surface (⟨110⟩ directions).

only DFT work for liquid surfaces focused on silicon, which shows an unconventional metallic behaviour, retaining substantial covalent bonding character in the melt [89].

Simulations of liquid metal surfaces are motivated by the need to complement with microscopic insights the 'integrated' experimental information coming from recent x-ray reflectivity measurements[5]. The occurrence of a peak in the x-ray reflectivity at the liquid–vapour interface is consistent with a stratification of the ions on the liquid side into layers parallel to the interface. Such a peak was initially detected for Ga [90, 91], Hg [92, 93] and In [94], and also some alloys [95, 96].

The pure metals studied in those experiments have large surface tensions, which suppress thermal fluctuations, thus allowing the measurement of the microscopic surface structure.

---

[5] Internet page of the Harvard x-ray group: http://www.liquids.deas.harvard.edu/

Alkali metals have significantly lower surface tensions, which result in thermally induced surface roughening; disentangling the layering behaviour from roughening effects due to capillary waves has been a challenging task in the experiments. It is only very recently that data on a pure alkali metal, K [97], have been reported. Such measurements suggest that low-surface-tension metals also layer, with a lead-up to a peak rather than a peak in the x-ray reflectivity detected. On the other hand, for water, which has a surface tension comparable to that of K but is a dielectric, no evidence of the occurrence of a peak in the x-ray reflectivity was observed [98].

A layered profile in the atomic density across the vapour–liquid interface, as measured in metals, is markedly different to that expected for dielectrics, where, in traversing the liquid–vapour interface, the density profile decreases monotonically from the value characteristic of the bulk liquid to that of the vapour. This structural disparity highlights fundamental differences in the behaviour of the free surfaces of metals and dielectrics. In a metal the nature of the interactions changes dramatically across the vapour–liquid interface, going from metallic bonding, where the valence electrons are delocalized, in the liquid to van der Waals bonding, with the electrons localized on atoms, in the vapour phase. In a dielectric the nature of the interaction does not change across the liquid–vapour interface.

Although a few explanations have been proposed to elucidate the layering mechanisms, many open questions still remain. One explanation was put forward by Rice and co-workers [99–101]. Based on their 'effective Hamiltonian' calculations, they proposed that layering at the surfaces of liquid metals is caused by the rapid decay of the electronic density at the surface acting like a 'hard wall' against which the ions are packed (the formation of layers in a liquid interfaced to a rigid wall is a well known phenomenon [102]). A second explanation was derived from glue model simulations by Iarlori *et al* [103], and Di Tolla [63]. It is based on the idea that the under-coordinated ions at the surface attempt to get closer to the optimal coordination they would have if they were in the bulk, by contracting inward towards it. This results in an increased density in the outermost part of the interface, which then causes a density oscillation that propagates from the surface into the bulk. The explanations for the mechanism of layering mentioned above have their roots in the metallic character of the interactions. Recently it has been proposed that the formation of layers at free liquid surfaces may not be unique to metals, but might generally be seen at temperatures well below the critical temperature T_C in systems having a low ratio, T_M/T_C, of the melting temperature to the critical temperature [104–106]. The suggestion, supported by Monte Carlo simulations with pair potentials, is that the many-body nature of the metallic interactions does not necessarily play a leading role (aside from determining the shape of an effective pair potential). The layering would also be seen in dielectric materials, but since they usually have a larger T_M/T_C, the formation of layers is pre-empted by solidification and therefore cannot be observed.

Motivated by the above, we chose to study alkali metals (in particular sodium) using first-principles molecular dynamics. Because the alkali metals are prototypes of nearly-free-electron behaviour, they allowed us to study the surface layering mechanisms without the complications of more complex binding such as in Ga and Hg.

Our sodium simulations contribute to clarify the picture of the mechanisms of ionic layer formation at the surfaces of liquid metals that have been put forward in the literature. BOMD simulations were performed for liquid sodium surfaces with electronic structures determined by ensemble DFT [17, 88]. MD simulations of just over 50 ps were made using slabs of about 160 atoms plus vacuum regions of thickness ~11 Å, and using the periodic boundary conditions/supercell approach mentioned in section 2.5. Troullier–Martins pseudopotentials were used with the PW91 generalized gradient approximation [107] to the exchange and correlation functional; these pseudopotentials included the non-linear core correction [108].
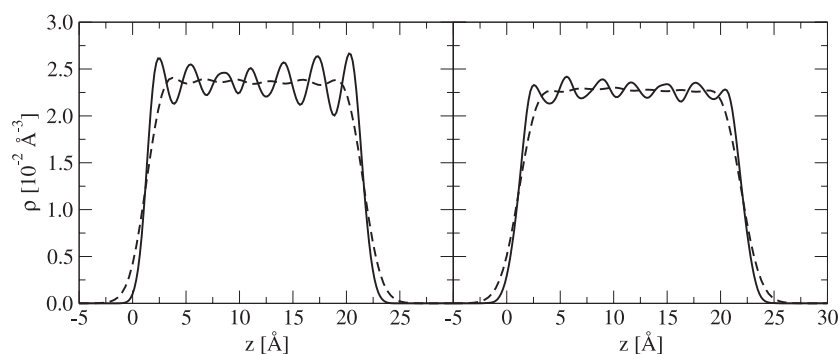
Figure 4. Density profiles averaged in planes parallel to the liquid–vapour interface for sodium from BOMD/eDFT simulations at temperature $T = 400$ K (left) and $T = 500$ K (right). The solid curve shows the ionic density profile, the dashed curve the valence electron density profile.

The Kohn–Sham orbitals were expanded in a plane wave basis set. Two different temperatures, $T = 400$ and 500 K, the first being just above the experimental sodium melting temperature ($T_M = 373$ K), the second higher, were examined, to observe the temperature dependence of the surface structure. Also, simulations using two geometrical cross-sections for the supercell (one consistent with cubic and the other with hexagonal symmetry) were made, in order to eliminate potential biasing of the in-plane surface structure by the supercell geometry. System size effects were also assessed.

We found that liquid surfaces of Na do layer, with an inter-layer spacing of about 3 Å, consistent with a close-packed arrangement of atoms against a 'hard wall'. Examples of the ionic and electronic density profiles (averaged in planes parallel to the surface) are shown in figure 4, for simulations at $T = 400$ and 500 K. The oscillations that can clearly be seen in those density profiles signify the formation of ionic layers at the surface, as found in the experimental results for the metals studied using x-ray reflectivity measurements. As illustrated in figure 4, stronger layering was observed in the simulations at lower temperatures.

A related goal of the sodium work was to understand the relevance of the Friedel oscillations for the layer formation mechanism. The Friedel oscillations [33] appear in the long-range tail of the screening response of the valence electrons in a metal to a local perturbation. Near the surface of a metal, Friedel oscillations are expected to appear as the valence electrons screen the disruption caused by the presence of the surface. It was recently suggested [109] that Friedel oscillations are responsible for fine-tuning of the interlayer relaxations at the $(10\bar{1}0)$ surface of *solid* Mg, so it is natural to ask what effect they might have at the surface of a *liquid* metal. The Friedel oscillations in sodium have a wavelength that is similar to the layer spacing; they are however a subtle effect, that is difficult to disentangle from 'hard wall' effects caused by the rapid decay of the electronic density at the surface. The electronic structure and dynamical information provided by first-principles MD calculations address some aspects of how the Friedel oscillations are affected by the structural and thermal disorder in the melt in relation to their possible role in tuning the surface layering.

Finally, we analysed whether any in-plane order is present at the liquid surfaces: we found mainly a fivefold coordination order (reminiscent of the icosahedral symmetry hypothesized for simple liquids) with hexatic regions which tend to become larger as thermal disorder is reduced by decreasing the temperature. In figure 5 two-dimensional plots of the ionic density in slices parallel to the interface (for slices near the left-most peaks in the density profiles of figure 4) are shown, for the two temperatures simulated. The plots shown provide an indication of the amount of structure present parallel to the surface.
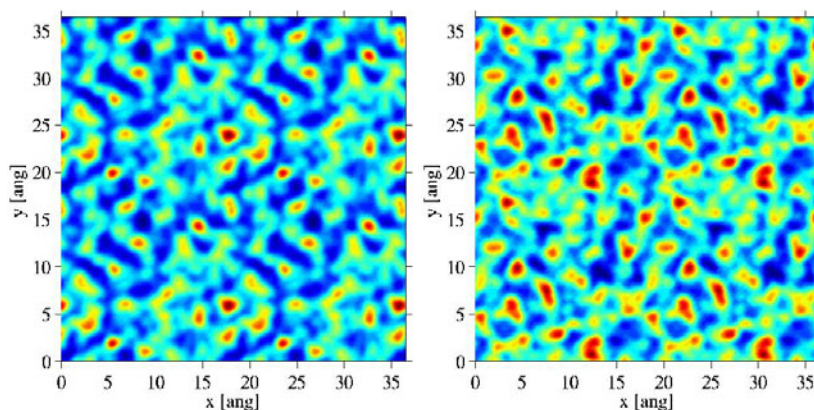
Figure 5. Two-dimensional ionic density plots for sodium from BOMD/eDFT *ab initio* simulations at $T = 400$ K (left) and $T = 500$ K (right), with the slices taken at approximately the positions of the left-most peaks in the ionic density profiles of figure 4. Red indicates the highest density; blue the lowest.

Using the *ab initio* BOMD/eDFT technique, we have obtained detailed and accurate microscopic information on the structure and dynamical behaviour of the liquid sodium surface at different temperatures. This has allowed us to look at the structure both along the normal to the surface, where we observe atomic layers, and in the plane of the surface.

5. Conclusions

We have reviewed an accurate and efficient method for performing *ab initio* dynamical studies of metallic systems: the Born–Oppenheimer molecular dynamics coupled with the ensemble density functional theory. To demonstrate the scope of the technique we have presented two applications where subtle electronic and dynamical effects, that are characteristic of metallic surfaces, were studied in detail using first-principles molecular dynamics. The peculiar structural and dynamical effects occurring at the Al(110) surface close to the premelting point, and in liquid Na surfaces, were described at length.

Acknowledgments

References

[1] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[2] Allen M P and Tildesley D J 1987 *Computer Simulations of Liquids* (New York: Oxford University Press)
[3] Marx D and Hutter J 2000 *Ab initio* molecular dynamics: theory and implementation *Modern Methods and Algorithms of Quantum Chemistry* (*NIC Series*) vol 1, ed J Grotendorst (Jülich: John von Neumann Institute for Computing) pp 301–449
[4] Tangney P and Scandolo S 2002 *J. Chem. Phys.* **116** 14
[5] Born M and Oppenheimer J R 1927 *Ann. Phys.* **84** 457
[6] Hellmann H 1937 *Einführung in die Quantenchemie* (Leipzig: Deuticke)
[7] Feynman R P 1939 *Phys. Rev.* **56** 340

[8]   Arias T A, Payne M C and Joannopoulos J D 1992 *Phys. Rev. Lett.* **69** 1077
[9]   Blöchl P E and Parrinello M 1992 *Phys. Rev.* B **45** 9413
[10] Blöchl P E 2002 *Phys. Rev.* B **65** 104303
[11] Vandevondele J and De Vita A 1999 *Phys. Rev.* B **60** 13241
[12] Stengel M and De Vita A 2000 *Phys. Rev.* B **62** 15283
[13] Alavi A, Kohanoff J, Parrinello M and Frenkel D 1994 *Phys. Rev. Lett.* **73** 2599
[14] Kresse G and Furthmüller J 1996 *Phys. Rev.* B **54** 11169
[15] Marzari N, Vanderbilt D, De Vita A and Payne M C 1999 *Phys. Rev. Lett.* **82** 3296
[16] Gygi F and Galli G 2002 *Phys. Rev.* B **65** 220102(R)
[17] Marzari N, Walker B G and Molteni C 2004 in preparation
[18] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[19] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[20] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045
[21] Pickett W E 1989 *Comput. Phys. Rep.* **9** 115
[22] Thomas L H 1927 *Proc. Camb. Phil. Soc.* **23** 542
[23] Fermi E 1927 *Rend. Accad. Lincei* **6** 602
[24] Fermi E 1928 *Rend. Accad. Lincei* **7** 342
[25] Smargiassi E and Madden P A 1994 *Phys. Rev.* B **49** 5220
[26] Smargiassi E and Madden P A 1995 *Phys. Rev.* B **51** 117
[27] Smargiassi E and Madden P A 1995 *Phys. Rev.* B **51** 129
[28] Govind N, Wang Y A and Carter E A 1999 *J. Chem. Phys.* **110** 7677
[29] Gonzalez L E, Gonzalez D J and Stott M J 2004 *Phys. Rev. Lett.* **92** 085501
[30] Perdew J P, Chevary J A, Vosko S H, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1992 *Phys. Rev.*
      B **46** 6671
[31] Ceperly D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566
[32] Perdew J and Zunger A 1981 *Phys. Rev.* B **23** 5048
[33] Ashcroft N W and Mermin N D 1979 *Solid State Physics* (Orlando, FL: Harcout College Publishers)
[34] Phillips J C 1958 *Phys. Rev.* **112** 685
[35] Phillips J C and Kleinman L 1959 *Phys. Rev.* **116** 287
[36] Heine V and Weaire D 1970 Pseudopotential theory of cohesion and structure *Solid State Physics* vol 24,
      ed H Ehrenreich, F Seitz and D Turnbull (New York: Academic) pp 249–463
[37] Hamann D R 1989 *Phys. Rev.* B **40** 2980
[38] Troullier N and Martins J L 1991 *Phys. Rev.* B **43** 1993
[39] Vanderbilt D 1990 *Phys. Rev.* B **41** 7892
[40] Mermin N D 1965 *Phys. Rev.* A **137** 1441
[41] Parr R G and Yang W 1989 *Density-Functional Theory of Atoms and Molecules* (New York: Oxford University
      Press)
[42] Feynman R P 1972 *Statistical Mechanics: a Set of Lectures* (Reading, MA: Benjamin)
[43] Marzari N, Vanderbilt D and Payne M C 1997 *Phys. Rev. Lett.* **79** 1337
[44] Marzari N 1996 *Ab initio* molecular dynamics for metallic systems *PhD Thesis* University of Cambridge
      Available online at http://nnn.mit.edu/phd
[45] Gillan M J 1989 *J. Phys.: Condens. Matter* **1** 689
[46] Chadi D J and Cohen M L 1973 *Phys. Rev.* B **8** 5747
[47] Monkhorst H J and Pack J D 1976 *Phys. Rev.* B **13** 5188
[48] Migdal A B 1958 *Sov. Phys.—JETP* **7** 996
[49] Chester G 1961 *Adv. Phys.* **10** 357
[50] Fu C-L and Ho K-M 1983 *Phys. Rev.* B **28** 5480
[51] De Vita A 1992 The energetics of defects and impurities in metals and ionic materials from first-principles
      PhD Thesis University of Keele
[52] de Gironcoli S 1995 *Phys. Rev.* B **51** 6773
[53] Elsässer C, Fähnle M, Chan C T and Ho K M 1994 *Phys. Rev.* B **49** 13975
[54] Methfessel M and Paxton A T 1989 *Phys. Rev.* B **40** 3616
[55] Verstraete M and Gonze X 2002 *Phys. Rev.* B **65** 035111
[56] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
[57] Stampfl C, Ganduglia-Pirovan M V, Reuter K and Scheffler M 2002 *Surf. Sci.* **500** 368
[58] Da Silva J L F, Stampfl C and Scheffler M 2004 *Phys. Rev. Lett.* **90** 066104
[59] Lundgren E, Kresse G, Klein C, Borg M, Andersen J N, De Santis M, Gauthier Y, Konvicka C, Schmid M and
      Varga P 2002 *Phys. Rev. Lett.* **88** 246103

[60] Mavrikakis M, Hammer B and Norskøv J K 1998 *Phys. Rev. Lett.* **81** 2819
[61] Dahl S, Logadottir A, Egeberg R C, Larsen J H, Chorkendorff I, Törnqvist E and Norskøv J K 1999 *Phys. Rev. Lett.* **83** 1814
[62] Bernasconi M and Tosatti E 1993 *Surf. Sci. Rep.* **17** 363
[63] Di Tolla F D 1995 Interplay of melting, wetting, overheating and faceting on metal surfaces: theory and simulation *PhD Thesis* Sissa/Isas
     Available online at http://www.sissa.it/
[64] Ditlevsen P D and Norskøv J K 1991 *Surf. Sci.* **254** 261
[65] Jayanthi C S, Tosatti E, Fasolino A and Pietronero L 1985 *Surf. Sci.* **152/153** 155
[66] Chen E T, Barnett R N and Landman U 1990 *Phys. Rev.* B **41** 439
[67] Barnett R N and Landman U 1991 *Phys. Rev.* B **44** 3226
[68] Yang L and Rahman T S 1991 *Phys. Rev. Lett.* **67** 2327
[69] Ditlevsen P D, Stoltze P and Norskøv J K 1991 *Phys. Rev.* B **44** 13002
[70] Mazzeo G 1992 Interplay between surface in-plane ordering and roughening *PhD Thesis* Sissa/Isas
[71] Rommelse K and den Nijs M 1987 *Phys. Rev. Lett.* **59** 2578
[72] Mazzeo G, Jug G, Levi A C and Tosatti E 1994 *Phys. Rev.* B **49** 7625
[73] Frenken J W M and van der Veen J F 1985 *Phys. Rev. Lett.* **54** 134
[74] van Pinxteren H 1994 Surfaces between melting and non-melting *PhD Thesis* FOM Amsterdam
[75] Molenbroek A 1995 Surface-melting instabilities at metal surfaces *PhD Thesis* FOM Amsterdam
[76] Stoltze P, Norskøv J K and Landman U 1988 *Phys. Rev. Lett.* **61** 440
[77] van der Gon A W D, Smith R J, Gay J M, O'Connor D J and van der Veen J F 1990 *Surf. Sci.* **227** 143
[78] van der Gon A W D, Frenkel D, Frenken J W M, Smith R J and Stoltze P 1991 *Surf. Sci.* **256** 385
[79] Dosch H, Höfer T, Peisl J and Johnson R L 1991 *Europhys. Lett.* **15** 527
[80] Pavlovska A, Tikhov M, Gun Y and Bauer E 1992 *Surf. Sci.* **278** 303
[81] Gester M, Kleinhesselink D, Ruggerone P and Toennies J P 1994 *Phys. Rev.* B **49** 5777
[82] Raphuthi A M, Wang X Q, Ercolessi F and Adams J B 1995 *Phys. Rev.* B **52** R5554
[83] Stumpf R and Scheffler M 1996 *Phys. Rev.* B **53** 4958
[84] Göbel H and von Blankenhagen P 1993 *Phys. Rev.* B **47** 2378
[85] Busch B W and Gustafsson T 1998 *Surf. Sci.* **415** L1074
[86] Busch B W and Gustafsson T 2000 *Phys. Rev.* B **61** 16097
[87] Mikkelsen A, Iruse J and Adams D L 1999 *Phys. Rev.* B **60** 7796
[88] Walker B G 2004 *Ab initio* molecular dynamics studies of liquid metal surfaces *PhD Thesis* University of Cambridge
[89] Fabricius G, Artacho E, Sánchez-Portal D, Ordejón P, Drabold D A and Soler J M 1999 *Phys. Rev.* B **60** R16283
[90] Kawamoto E H, Lee S, Pershan P S, Deutsch M, Maskil N and Ocko B M 1993 *Phys. Rev.* B **47** 6847
[91] Regan M J, Pershan P S, Magnussen O M, Ocko B M, Deutsch M and Berman L E 1996 *Phys. Rev.* B **54** 9730
[92] Magnussen O M, Ocko B M, Regan M J, Penanen K, Pershan P S and Deutsch M 1995 *Phys. Rev. Lett.* **74** 4444
[93] Di Masi E, Tostmann H, Ocko B M, Pershan P S and Deutsch M 1998 *Phys. Rev.* B **58** R13419
[94] Tostmann H, Di Masi E, Pershan P S, Ocko B M, Shpyrko O G and Deutsch M 1999 *Phys. Rev.* B **59** 783
[95] Regan M J, Pershan P S, Magnussen O M, Ocko B M, Deutsch M and Berman L E 1997 *Phys. Rev.* B **55** 15874
[96] Regan M J, Kawamoto E H, Lee S, Pershan P S, Maskil N, Deutsch M, Magnussen O M, Ocko B M and Berman L E 1996 *J. Non-Cryst. Solids* **205–7** 762
[97] Shpyrko O, Huber P, Grigoriev A, Pershan P, Ocko B, Tostmann H and Deutsch M 2003 *Phys. Rev.* B **67** 115405
[98] Shpyrko O G, Fukuto M, Pershan P S, Ocko B M and Kuzmenko I 2003 *Phys. Rev.* B submitted
[99] Harris J G, Gryko J and Rice S A 1987 *J. Chem. Phys.* **87** 3069
[100] Gomez M A and Rice S A 1994 *J. Chem. Phys.* **101** 8094
[101] Zhao M, Chekmarev D S, Cai Z and Rice S A 1997 *Phys. Rev.* E **56** 7033
[102] Rowlinson J S and Widom B 2002 *Molecular Theory of Capillarity* (New York: Dover)
[103] Iarlori S, Carnevali P, Ercolessi F and Tosatti E 1989 *Surf. Sci.* **211/212** 55
[104] Chacón E, Reinaldo-Falagán M, Velasco E and Tarazona P 2001 *Phys. Rev. Lett.* **87** 166101
[105] Tarazona P, Chacón E, Reinaldo-Falagán M and Velasco E 2002 *J. Chem. Phys.* **117** 3941
[106] Valesco E, Tarzona P, Reinaldo-Falagán M and Chacón E 2002 *J. Chem. Phys.* **117** 10777
[107] Perdew J P and Wang Y 1992 *Phys. Rev.* B **45** 13244
[108] Louie S G, Froyen S and Cohen M L 1982 *Phys. Rev.* B **26** 1738
[109] Cho J H, Ismail, Zhang Z and Plummer E W 1999 *Phys. Rev.* B **59** 1677