First principles simulations
M. Palummo, L. Reining, P. Ballone

To cite this version:

HAL Id: jpa-00251958
https://hal.archives-ouvertes.fr/jpa-00251958
Submitted on 1 Jan 1993

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
First principles simulations

M. PALUMMO, L. REINING* and P. BALLONE**

Dipartimento di Fisica, Università di Roma “Tor Vergata”, via della Ricerca Scientifica, 00133 Roma, Italy
* Laboratoire des Solides Irradiés, CEA-CNRS, URA 1380, Ecole Polytechnique, 91128 Palaiseau, France
** Institut für Festkörperforschung, Forschungszentrum Jülich, 5170 Jülich, Germany

Abstract

In this paper we outline the major features of the “ab-initio” simulation scheme of Car and Parrinello, focusing on the physical ideas and computational details at the basis of its efficiency and success. We briefly review the main applications of the method. We discuss the limitations of the standard scheme, as well as recent developments proposed in order to extend the reach of the method.

Moreover, we consider more in detail two specific subjects. First, we describe a simple improvement (Gradient Corrections) on the basic approximation of the “ab-initio” simulation, i.e. the Local Density Approximation. These corrections can be easily and efficiently included in the Car-Parrinello code, bringing computed structural and cohesive properties significantly closer to their experimental values. Finally, we discuss the choice of the pseudopotential, with special attention to the possibilities and limitations of the last generation of soft pseudopotentials.

I Introduction

The atomistic simulation of real materials requires an accurate model for the interatomic potential energy. The model has to be reliable, in order to describe the behavior of the system under a variety of thermodynamic conditions, external perturbations, and chemical transformations. Moreover, it has to be simple and computationally efficient, to allow for thousands of energy and forces evaluations as required by standard simulation schemes.

Idealized models (mainly pair potentials) have played a crucial role in the development of simulation techniques, and have provided a qualitative description for a variety of microscopic phenomena [1]. Quantitative comparison with experiments, however, was possible for only a well defined and limited set of materials.

With increasing importance of simulation, also the interest for more quantitative models, able to describe a larger class of systems has increased.

An important breakthrough came with the simulation scheme of Car and Parrinello (CP), proposed in 1985 [2]. The method is based on the Density Functional Theory (DFT) [3], usually within the Local Density Approximation (LDA), i.e. the standard framework for state of the art total energy and electronic structure computations. This basis provides a solid foundation for a reliable description of a large class of systems, including almost all the elements across the periodic table [4], [5]. The viability of the scheme as a Molecular Dynamics technique relies on an efficient strategy to update the electronic structure of the system as the atoms move.

The CP combination of the DFT-LDA framework with the Molecular Dynamics algorithm allowed, for the first time, the simultaneous description of the atomic dynamics and of the electronic
structure of a system, thus opening the possibility of simulating systems with strong and changing chemical bonds.

Despite the simplicity of the basic ideas, the success of the method is due to several judicious choices of physical ingredients and computational algorithms. In the following, we try to highlight these crucial features, to discuss their implications on the scope of the method, and, in particular, to point out the limitations they impose on the practical applicability of the "ab-initio" simulation scheme. We shall not attempt to provide a complete and detailed discussion of the method, for which we refer to two excellent papers by the original authors [6], [7].

For the sake of simplicity, we concentrate on the original formulation of the CP scheme. Important successive developments will be mentioned for completeness.

The last sections of the paper discuss two specific topics. In Sect. III we describe a simple improvement on the Local Density Approximation, which is the basic approximation underlying most of the ab-initio simulations performed until now. Finally, we discuss the choice of the pseudopotentials for CP simulations. Pseudopotentials are a basic and essential ingredient of the method, and also the reason for difficulties in studying important classes of materials like transition metals or oxides. Recently, several attempts have been made to improve the efficiency of pseudopotentials in the framework of the CP scheme. Our experience with these new pseudopotentials, described in Sec. IV, suggests there are only limited reasons for optimism.

II The Car-Parrinello method

a) Density Functional Theory and the Local Density Approximation

The Car Parrinello method is a Molecular Dynamics scheme in which the potential energy of the system and the forces on the atoms are computed "ab-initio" within the Density Functional framework.

Given a set of atomic positions \( \{ \vec{R}_i, i = 1, N \} \), the basic theorems of Density Functional theory [3] state that the electronic ground state energy \( E \) of the system is the minimum of a universal functional of the electron density \( \rho(\vec{r}) \):

\[
E [\rho | \{ \vec{R}_i \}] = \text{Min}_\rho \left[ T + \frac{1}{2} \int \rho(\vec{r}) \varphi(\vec{r}) d\vec{r} + \int \rho(\vec{r}) V_{\text{ext}}(\vec{r}) d\vec{r} + E_{\text{XC}}[\rho] \right]
\] (1)

In the equation above \( T \) is the electron kinetic energy; \( \varphi \) is the Coulomb potential; \( V_{\text{ext}} \) is the "external" potential on the electron due to the atomic cores, and \( E_{\text{XC}} \) is the exchange and correlation energy.

The electronic energy \( E \), together with the direct core-core repulsion, defines one point of the Born-Oppenheimer surface, i.e. one point of the potential energy surface of the system.

In practical computations \( \rho \) is expressed in terms of the square of independent and orthonormal electron orbitals \( \{ \psi_i, i = 1, n \} \), where \( n \) is the number of electrons in the system:

\[
\rho(\vec{r}) = \sum_i |\psi_i(\vec{r})|^2
\] (2)

The kinetic energy \( T \) is obtained from these orbitals as:

\[
T = \sum_i < \psi_i | -\frac{1}{2} \nabla^2 | \psi_i >
\] (3)

and \( E_{\text{XC}} \) is usually computed within the well known Local Density Approximation:

\[
E_{\text{XC}}[\rho] = \int \rho(\vec{r}) \epsilon_{\text{XC}}(\rho(\vec{r})) d\vec{r}
\] (4)

where \( \epsilon_{\text{XC}}(\rho) \) is the exchange-correlation energy per particle of the uniform electron gas at density \( \rho \).

The minimum principle with respect to the density \( \rho \) is equivalent to the following Euler equation for the \( \psi \)'s:

\[
H_{KS} \psi_i(\vec{r}) = \left( -\frac{1}{2} \nabla^2 + \varphi + V_{\text{ext}} + \frac{\delta E_{\text{XC}}}{\delta \rho(\vec{r})} \right) \psi_i(\vec{r}) = \sum_j \Lambda_{ij} \psi_j(\vec{r})
\] (5)
which introduces a Schrödinger-like equation in terms of the so-called Kohn-Sham Hamiltonian \( H_{KS} \). The "eigenvalue" matrix \( \Lambda \) is indeed a matrix of Lagrange multipliers required to retain the orthonormality of the orbitals \( \{ \psi_i \} \).

The standard way of computing \( E \) is to choose a basis set \( \{ \gamma_j(\mathbf{r}), j = 1, M \} \), expand the \( \psi \)'s into this basis:

\[
\psi_i(\mathbf{r}) = \sum_j c_j^{(i)} \gamma_j(\mathbf{r})
\]

and diagonalize the Hamiltonian \( H_{KS} \). This process has to be repeated several times to achieve self-consistency between the density \( \rho \) and the potentials \( \varphi \) and \( \delta E_{XC}/\delta \rho(\mathbf{r}) \).

Depending on the material, the chosen basis set, and the accuracy in the self-consistency condition, the computation of a single \( E(\{ \vec{R}_i \}) \) by this standard route can require a non negligible computational effort.

Despite this drawback, and the initial skepticism \cite{3}, the DFT-LDA method has gained general acceptance as the basic tool to describe cohesive and electronic properties of condensed matter. A large number of computations has shown that DFT-LDA is highly successful in describing structural properties and the harmonic dynamics of a very large class of materials \cite{5}, \cite{4}. It is able to treat with the same accuracy and within a unified framework solids characterized by very different chemical bonding. Its reliability has been highlighted by the successful prediction of structural transformations of solids under pressure. Its range of applicability includes (although at a different level of accuracy) atoms, molecules and solids.

b) The standard Car-Parrinello scheme

This impressive record of flexibility and reliability makes DFT-LDA an attractive recipe to simulate real materials. Of course, to make the scheme practical, it is necessary to be able to compute several thousand times the energy \( E \) and its gradients with respect to the atomic positions for systems ranging in size from few \( 10^4 \) up to \( \sim 10^5 \) atoms. This computational feat is the specific achievement of the CP method.

In addition to the basic features of the standard method described above, the CP scheme adds two computational choices, which are also used in static DFT-LDA calculations. The total electron density is partitioned into a core and a valence charge, following the prescriptions of elementary chemistry. The simulation involves the valence charge only, while the effect of the core charge on the valence electrons is replaced by a frozen potential. This is added to the nuclear Coulomb potential to produce an ionic pseudopotential acting on the valence electrons. For most of the elements this approximation involves a negligible loss of accuracy, provided the choice of the pseudopotential is made with sufficient care (see Sec. IV below).

After this approximation is made, the "external potential" in the KS Hamiltonian is the sum of regular (i.e. continuous and non diverging) pseudopotentials centered on the atomic nuclei.

The second important choice concerns the basis set. In the standard Car Parrinello scheme the system under study is enclosed in a basic simulation box of fixed shape, periodically repeated in space. Then, the independent electron orbitals are expanded in plane waves:

\[
\psi_i(\mathbf{r}) = \sum_{\vec{G}} c^{(i)}_{\vec{G}} \exp(i\vec{G} \cdot \mathbf{r})
\]

where the \( \vec{G} \) are the vectors of the reciprocal lattice implied by the periodically repeated simulation box. The expansion is limited to a finite sum by including in eqn. (7) only the plane waves whose \( \vec{G} \) has modulus smaller than a cut off \( | \vec{G}_c | \). The choice of \( | \vec{G}_c | \) determines the spatial resolution of the basis set, and critically depends on the elements composing the system. Of course, elements with rapidly varying electron densities will require higher resolution and cutoff, reflected in a larger size for the plane wave expansion.

For classical systems the periodic repetition of a central cell does not pose serious problems, and periodic boundary conditions (pbc) are the standard choice to reduce size effects. For quantum systems pbc introduce more subtle size effects, that will be discussed below.

An immediate consequence of the plane waves choice is that the basis set will depend on the system volume. In fact, the density of \( \vec{G} \) vectors in reciprocal space is directly proportional to the volume of the real space cell. At fixed cutoff, therefore, the number of basis functions grows linearly with the system volume, even if the number of atoms or electrons remains constant. This is particularly a problem for finite systems (like clusters), and surface applications.
Even for homogeneous systems, a moderately large simulation cell will require a very large number of plane waves. For simple elements (the prototype is silicon), and for a system of about 100 atoms, the expansion (7) can require several thousand plane waves to accurately reproduce the orbitals $\psi$. The large size of the expansion is clearly related to the simplicity of the basis functions, not well adapted to the shape and symmetries of atomic-like orbitals.

The disadvantages of the plane waves, however, are more than offset by the possibility of using fast Fourier transforms to compute the energy and the KS Hamiltonian, (especially the kinetic energy and the Coulomb potential, particularly simple in reciprocal space) and by the simplifications achieved in the computation of the forces on the ions.

c) Starting the simulation

The first task in a Car-Parrinello simulation is to compute the energy $E$ for a given initial configuration of the atoms, ie to find the first point on the Born-Oppenheimer surface of the system. In principle, this could be achieved by the standard technique of diagonalization and iteration to self-consistency. It is easy to realize, however, that the problem becomes a formidable one as soon as the dimension $M$ of the basis set exceed a few hundred functions, since the dimension of the matrix $H_{KS}$ is equal to $M$ times the number of electrons in the system.

To solve this problem, another, more direct route is available. The expansion (7) for the orbitals can be substituted into eq. (1), to provide an expression of the electron energy $E$ as a function of the expansion coefficients $\{c_G^{(i)}\}$. Then, the minimum condition can be seen as an optimization problem for the cost function $E(\{c_G^{(i)}\})$ with respect to the multitude of coefficients $\{c_G^{(i)}\}$. A simple strategy to achieve this optimization is:

i) start from an initial choice of the $\{c_G^{(i)}\}$.
ii) Compute the gradient $\partial E/\partial c_G^{(i)}$ of the energy with respect to the expansion coefficients.
iii) Exploit this information in a steepest descent or conjugate gradient routine to approach the minimum.

The practical feasibility of this approach depends on the following two conditions:

i) The evaluation of the gradient should be inexpensive.
ii) The surface $E(\{c_G^{(i)}\})$ has to be simple, ie it should not present local minima preventing the system from reaching the ground state.

An easy computation shows that, when the plane waves basis set is chosen, the gradient is simply and efficiently given by a Fourier transform:

$$\frac{\partial E}{\partial c_G^{(i)}} = \int H_{KS}(\vec{r})\psi_i(\vec{r}) \exp(-i\vec{G}\cdot\vec{r}) d\vec{r}$$

Moreover, the experience has shown that the surface $E(\{c_G^{(i)}\})$ is indeed very simple. For the systems of interest in a CP simulation (ie systems of low or no symmetry) the optimization proceeds rapidly, and ends up in the electronic ground state. Particularly efficient for this initial electronic optimization is the conjugate gradient method [8].

The direct minimization approach has several advantages over the standard method, the most important of which is that it does not require to compute and store the full KS Hamiltonian matrix. Instead, at each time, only one row of $H_{KS}$ is required to compute the gradient of $E$ and improve the solution.

The other major advantage is that only the $n$ occupied electronic eigenstates $\psi_i$ enter the computation, with an enormous reduction in complexity with respect to diagonalization methods computing the full spectrum of $H_{KS}$.

The direct minimization approach is closely related to standard computational algorithms to extract selected eigenstates of a very large matrix. These are always in advantage over brute force diagonalization as soon as:

i) the number of required eigenstates is much smaller than the matrix dimension;
ii) a good starting approximation to the required eigenvectors is available.

The evolution of the system toward the ground state in direct minimization schemes can be seen as the time evolution of the variables $\{c_G^{(i)}\}$ in the potential energy $E(\{c_G^{(i)}\})$. The steepest descent strategy, for instance, would correspond to:

$$c_G^{(i)} = -\frac{\partial E}{\partial c_G^{(i)}}$$
The point on the Born Oppenheimer surface computed by direct minimization defines also the forces acting on the atoms, via the local gradient of the potential energy.

In the standard CP scheme the forces on the atoms are given by the Hellman-Feynman expression:

$$\vec{F}_i = - \sum_i \int \psi_i^*(\vec{r}) \frac{\partial H_{KS}}{\partial \vec{R}_i} \psi_i(\vec{r}) d\vec{r} + \vec{F}_{\text{ion-ion}}$$

where $\vec{F}_{\text{ion-ion}}$ is the force due to the direct core-core repulsion. Again, $\partial H / \partial \vec{R}_i$ can be computed efficiently exploiting the fast Fourier transform. The simple and convenient expression (10), involving the gradient of $H_{KS}$ but not the explicit gradient of the $\psi$'s, is not an obvious result, and, indeed, it is a specific property of the plane wave basis set. Other possible choices of the basis set (like gaussians or atomic orbitals), more efficient for static calculations and extensively used in quantum chemistry, involve additional contributions to the forces (the Pulay forces) which are difficult and expensive to evaluate [5].

**d) Dynamics**

With the computation of the potential energy $E$ and its gradient with respect to the atomic positions, one can achieve the basic goal of performing one MD step by updating the configuration of the atoms by standard integration algorithms [1].

In principle, this could be repeated until a sufficient statistics is collected to describe the dynamics of the system and compute the appropriate thermodynamic averages.

In fact, the MD requirements are ideal for the iterative solutions described above. The slight update of the atomic positions during a time step will involve only a minor readjustment of the electronic configuration, and therefore the optimal coefficients at time $t$ will provide an excellent starting point to compute the $\{c_0^{(i)}\}$ at time $t + dt$.

In practice, however, two important problems arise. Optimizing the electronic structure at each time step implies that energy is constantly subtracted from the electronic degrees of freedom. This should not be a problem, because the motion of the atoms should account for the energy balance. In practice, however, the electronic energy minimization and the integration of the equations of motion for the ions cannot be performed with the same accuracy. As a result the method is unstable, and cannot be used for a MD simulation.

Moreover, the minimization process can be inefficient, since several steepest descent or conjugate gradient steps are required to reach the minimum, even starting very close to it.

The solution to this problem, devised by Car and Parrinello, is to move not only the atoms but also the coefficients $\{c_G^{(i)}\}$ according to Newton's equations of motion:

$$\mu c_G^{(i)} = - \frac{\partial E}{\partial c_G^{(i)}}$$

$$M \ddot{\vec{R}}_i = - \frac{\partial E}{\partial \vec{R}_i}$$

This equations can be derived from the classical Lagrangean:

$$L = \sum_i \int \mu |\dot{\psi}(\vec{r})|^2 d\vec{r} - E [\vec{R}_i | \{c_G^{(i)}\}] + \sum_{i,j} \Lambda_{ij} \left( \int \psi_i^*(\vec{r}) \psi_j(\vec{r}) d\vec{r} - \delta_{ij} \right)$$

In eq. (11) $\mu$ is a constant with the dimension of a mass. By choosing $\mu$ much smaller than $M$ the evolution of the $\{c_G^{(i)}\}$'s in their configuration space will be dominated by time scales significantly shorter than those of the atomic motion.

This strategy, odd at first, closely mimics the behavior of real systems. Eq. (11) treats the quantum evolution of the electrons as a classical one, and, of course, it does not describe the real electron dynamics. It retains, however, the separation in the time scales of the electron and atomic motion, and exploits the adiabatic principle to keep the system on the BO surface.

Experience has shown that, with a careful choice of $\mu$, the system of eqns. (11) and (12) is stable, and reliably describes the adiabatic motion of the atoms in the DFT-LDA potential energy.

Having outlined the basic features of the CP method, we turn to the discussion of what can be done by this method, and, more importantly, of what are the limitations of the standard formulation.

The crucial ingredient often determining the practical feasibility of a CP simulation is the pseudopotential and the related valence change density of the atoms composing the system. As mentioned
above, pseudopotentials (and valence charge densities) present widely different localization character in different elements. This localization is an intrinsic property of the atomic structure of the elements, and cannot be changed at will by the pseudopotential construction (see, however, sec. IV below). As a result, the cutoff $|G_c|$ and the number of plane waves required for a CP simulation can vary by orders of magnitude according to the elements composing the system.

Particularly favourable for the plane wave expansion are the simple metals, and the s-p semiconductors of the 3p, 4p, 5p groups of the periodic table. (Si, Ge, GaAs, P, Se, S, etc.), all characterized by soft pseudopotentials and smooth valence charge densities.

Instead, particularly difficult to describe in a plane wave basis set are the elements of the 2p and 3d groups of the periodic table, including elements like C, N, O, F, Fe, Ni and Cu that are among the cornerstones of chemistry and material science. The 3d transition elements, in particular, are practically outside the reach of the standard scheme, since an accurate description of their 3d valence states would require $\sim 10^6$ plane waves even for systems of moderate size.

An extension of the scheme, based on the ultrasoft pseudopotentials proposed by Vanderbilt [10], could provide the needed breakthrough, and open the possibility of simulating these important and difficult elements [11].

The second basic problem of the method is related to the subtle quantum mechanical size effects implicit in the expansion (7). In fact, this expansion is not the most general expression for the wave function of a quantum system periodically repeated in space. From elementary solid state theory, we know that the general solution of the Schrödinger equation in a periodic potential is a periodic function times a plane wave, and, therefore, the allowed independent electron states $\psi$ should be expanded as:

$$\psi_{i,(k)}(\vec{r}) = \sum_{\vec{G}} c_{\vec{G}}^{(i)}(k) \exp \left[ i(\vec{G} + \vec{k})\vec{r} \right]$$

where $\vec{k}$ is a vector in the first Brillouin zone. In computational terms this means we should solve eqn. (5) for several different $\vec{k}$ points in the BZ, and choose the occupied states among the different $\psi_{i,(k)}$ eigenstates in order to minimize the energy. This would multiply the computational effort, or require involved and expensive optimization techniques.

The distinction between the expansions (7) and (14) can be immaterial for large systems, for which the BZ is reduced to a very small volume around the origin, and the $\vec{k}$ vectors become all equivalent. However, it is never negligible for the system sizes that can be considered in practical CP simulations, and it is particularly important for metals, for which the eigenstates $\psi_{i,(k)}$ strongly depend on $\vec{k}$. This problem, mainly computational in nature, has until now prevented an extensive application of the CP scheme to metals.

Another problem specific to metals is that the separation of the time scales for the electron and nuclear motion vanishes for the electrons close to the Fermi surface. This could result in a loss of effectiveness of the adiabatic strategy to keep the system on the Born-Oppenheimer surface. This problem, often discussed, until now has not had much impact on the applications of the CP scheme, since a reliable simulation of metals is indeed prevented by the BZ sampling problem described above. A possible solution to the adiabatic problem in metals has recently been proposed by Blochl and Parrinello [12].

The importance of the computational considerations listed above is apparent from a review of the main applications of the CP method.

Because of the ease in the plane wave expansion, of limited problems with the BZ sampling, and also of the lack of simple and reliable alternatives to simulate these systems, most of the applications of the CP scheme have been devoted to semiconductors. Among the most representative applications, we list the simulation of amorphous and liquid silicon [9], [13] and carbon [14], [15], hydrogenated silicon [16].

Moreover it has been applied to finite systems [17], for which the inadequacies of potential approximations are particularly severe, and again the CP scheme provides the only possibility for a reliable investigation. Also for these systems there are no problems associated to the BZ summation.

Less computations have been devoted to metals. Among the few exceptions we mention Ref. [12], and the activity of the Argonne group [18].

### III Beyond Local Density: The Gradient Corrections

Besides the computational limitations described above, the Local Density Approximation, expressed by eqn. (4), remains the most basic source of error and uncertainties in the CP scheme.
Although very successful in many respects, the LDA has also several important limitations, that have been known since the beginning of its extensive application to study electronic properties of solids.

The most discussed one is the well known band gap problem [19], that, however, is only partly due to LDA. The eigenvalues $\Lambda$ of the KS equations, in general, cannot be interpreted as excitation energies. However, the similarity between the KS and Schroedinger equations is too tempting, and the comparison of KS eigenvalues with experimental excitation energies has often been attempted. This comparison has shown that DFT-LDA severely underestimate the energy of the first electronic excitation of semiconductors and insulators; the LDA further enhances the error which is already inherent in the application of DFT to an excited-state problem.

Another qualitative failure of LDA is its inability to correctly describe small negatively charged systems. This problem is directly related to the well known deficiency of LDA in describing the exchange-correlation potential [20].

While these difficulties may be not very important for MD applications, other problems are more directly related to the evaluation of the potential energy of the system, and therefore can affect the result of the "ab-initio" Molecular Dynamics.

In this respect, the most important problem is the overestimation by LDA of cohesive energies of molecules and solids. Systematic computations [5], [4] have indeed shown that LDA predicts cohesive energies $E_c$ significantly larger than the experimental values. The error is more severe for finite systems (molecules and clusters) for which the relative error can be as large as 30%. Again, this could be not very important in MD, where we do not deal with the absolute value of the cohesive energy, but with cohesive energy differences between closely related atomic configurations. However, it is difficult to imagine that such a large error in the potential energy does not appear anywhere in the dynamics and the thermodynamics of the system.

Even more serious than the cohesive energy overestimation for covalent and metallic systems, is the qualitative inadequacy of LDA in describing the weak covalent-van der Waals attraction among closed shell systems, and the hydrogen bonding. The first bonding characterize all the molecular solids, the second one is the major component of the bonding in water and many biological systems. For this materials, not only $E_c$ is overestimated, but also the computed interatomic distances, equilibrium structures, and phonon frequencies can be very far from the experimental values. To open the possibility of reliably describing these systems, it is important to find a simple and computationally efficient alternative to LDA.

The simplest addition to LDA, providing information on the spatial variation of the valence charge involves the gradients of $\rho(\vec{r})$. A systematic expansion of $E_{XC}$ in terms of density gradients was proposed very early in DFT [3], but did not gain general acceptance. Practical computations and theoretical considerations have shown that the expansion is not converging, and the first few terms do not provide any improvement of the results with respect to LDA.

More recently, semiempirical approaches following the general lines of the gradient expansion have appeared, and rapidly achieved a good degree of success. All of them approximate the exchange-correlation energy as:

$$E_{XC}[\rho] = E_{XC}^{LDA}[\rho] + \int f(\rho; \nabla \rho) d\mathbf{r}$$

(15)

where $f$ is a function of the density and its gradient. The function $f$ often includes free parameters adjusted on atomic properties.

The natural selection of practical computations has singled out the recipe of Becke [21] for the exchange energy, and the Perdew approximation for correlation [22].

We do not enter here into the details of the Gradient Correction theory, nor do we discuss the relative merits of the different recipes. We simply emphasize two basic points:

i) GC are semi-empirical in character, and should be considered as a simple and convenient recipe to approximate the potential energy of the system. They cannot be considered as a fundamental step beyond LDA. In fact, they do not improve the deep problems of the theory concerning the gap problem and the basic properties of the exchange correlation potential.

ii) GC are computationally efficient, especially in the plane wave basis set. Fast Fourier transform can be used to compute the derivatives required in the evaluation of $E_{XC}$ and $H_{\text{KS}}$. This few FFT are the extra effort required by GC with respect to the LDA computation. Their cost does not depend on the number of electrons in the system, and therefore their relative importance becomes rapidly negligible as the system size grows. Moreover, GC do not explicitly enter the computation of the forces on the ions, and therefore do not complicate the basic task of MD.

Systematic computations including GC are rapidly growing in number, and consistently show that indeed GC represent a systematic and substantial improvement on LDA to compute cohesive,
structural and vibrational properties of molecular and hydrogen bonded systems.

The first application of GC in a CP computation was presented in ref. [23], concerning the structural and dynamical properties of magnesium microclusters.

Recently, the Car-Parrinello group in Zürich and Lausanne presented several applications of the GC method, focused on the properties of ice, water clusters, and liquid water [24].

IV Soft Pseudopotentials

In this last section of the paper we discuss a crucial ingredient of every CP simulation, ie the choice of the pseudopotential. In particular, we would like to share our recent experience with the last generation of "soft" pseudopotentials.

The art of pseudopotential generation was well developed and successful well before the CP scheme appeared. A complete and still valid account is contained in Ref. [25].

The specific requirements of the CP scheme have stimulated the development of pseudopotential generation schemes optimized for the plane-wave expansion. This optimization can be achieved only by smoothing as much as possible the pseudopotential (giving rise to the so called "soft" pseudopotentials) and the associated valence charge density. This smoothing will decrease the spatial resolution required to describe the wave functions, and therefore will decrease the size of the expansion (7).

Of course, this process cannot be pushed too far, in order to conserve the essential features of the atomic structure and valence properties of the elements. It is therefore meaningful to investigate how much this process can be safely carried out, and what are the specific symptoms of the pseudopotential failure.

Among the different recipes recently proposed to produce soft pseudopotentials, we focus our attention to the Martins and Troullier (MT) scheme [26] because of its simplicity and ease of application in the standard CP code.

State of the art pseudopotentials are constructed in order to reproduce exactly the valence charge density of the atoms outside a sphere of radius $r_c$, much smaller than the typical interatomic distances the atom will explore in forming bonds. Within the sphere, the potential and the charge are modified to produce nodeless wave functions and smooth pseudopotentials.

The radius $r_c$ is the parameter controlling the reliability of the pseudopotential as well as its softness. A small $r_c$ implies a faithful reproduction of the valence charge over an extended region of space, and does not offer freedom in deciding the strength of the pseudopotential. Larger $r_c$, instead, offer this freedom, but also result in less reliable pseudopotentials.

The Martins Troullier scheme provides a series of conditions to extend the ability of the pseudopotential to reproduce valence properties even when the radius $r_c$ becomes comparable to half of a typical interatomic distance. The possibility of using such large radii would offer the possibility of drastically reduce the cutoffs of the plane wave expansion, and would bring oxides and some transition metals within the reach of the CP simulation, at least for systems of limited size.

Of course, the reduction in cutoff is due to the large $r_c$, and not to the details of the pseudopotential construction within $r_c$.

To test the quality of the MT scheme, we have performed systematic plane wave computations for a series of molecules, including several problematic elements like B, C, N, O and F. We carefully analyzed the performances of the pseudopotential as a function of $r_c$ and the ionicity of the molecular bond. The benchmark to judge the quality of the pseudopotential was provided by previous all-electron computations available in the literature [27].

The conclusions of our investigation are the following:

i) Non ionic systems (like $C_2$, $N_2$, etc.) are safely described by MT pseudopotentials with large radii $r_c$, requiring a limited expansion in plane waves. This is true even for traditionally difficult elements like O or F.

ii) With increasing ionicity of the molecular bond, however, the ability of the pseudopotential to mimic the real potential is more severely tested. With increasing $r_c$ the computed properties (cohesive energy, equilibrium distance and vibrational frequencies) progressively deviate from the correct values. With $r_c$ approaching half the interatomic distance, the pseudopotential scheme completely breaks down, and produces results in complete disagreement with more accurate computations.

These conclusions clearly indicate that, while some improvement can be achieved from a careful construction of the pseudopotential, this ingredient alone cannot provide the breakthrough required to allow an efficient and reliable simulation of several important materials.
V Conclusions

The discussion of the CP scheme presented in this paper is intended as a short introduction to the method for an audience mainly interested in material science applications. For this reason we have concentrated on the basic ideas, and outlined the specific (mainly computational) features of the method behind its success in describing a large variety of semiconductor systems, and its problems in metal applications. Often, we have sacrificed completeness to simplicity, and we also did not discuss several of the most recent developments of technical character.

This report photographs a rapidly evolving field, witnessing every day a succession of improvements and new applications.

As an example for the evolving boundaries of the method, we presented a simple introduction to the Gradient Corrected functionals and to the soft pseudopotential technique, focused on the specific point of view of the Car-Parrinello applications.

ACKNOWLEDGMENTS - We acknowledge warm hospitality at the Centre Européen de Calcul Atomique et Moléculaire (CECAM) in Paris. We also acknowledge several useful discussions with the CECAM resident and visiting scientists. One of us (M.P.) thanks the Laboratoire des Solides Irradiés for its support.

References


