



HAL
open science

Car–Parrinello Molecular Dynamics

Mauro Boero, Atsushi Oshiyama

► **To cite this version:**

Mauro Boero, Atsushi Oshiyama. Car–Parrinello Molecular Dynamics. Encyclopedia of Nanotechnology, Springer Netherlands, pp.1-10, 2015, 10.1007/978-94-007-6178-0_100946-1 . hal-02991687

HAL Id: hal-02991687

<https://hal.science/hal-02991687v1>

Submitted on 24 Nov 2020

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.

Car-Parrinello Molecular Dynamics

Mauro Boero¹ and Atsushi Oshiyama²

¹Institut de Physique et Chimie des Matériaux de Strasbourg (IPCMS), University of Strasbourg and CNRS, UMR 7504, 23 rue du Loess, F-67034, Strasbourg, France

²Department of Applied Physics, The University of Tokyo, 7-3-1 Hongo, Tokyo 113-8656, Japan

Synonyms

First Principles Molecular Dynamics, Ab Initio Molecular Dynamics, DFT-based Molecular Dynamics, CPMD

Definition

The Car-Parrinello molecular dynamics (CPMD) is an extension of the Lagrangean formalism of classical molecular dynamics in which the model potential describing the interaction among atoms is replaced by the total energy functional of the system as provided by the Density Functional Theory (DFT). The electronic wavefunctions are explicitly introduced as new dynamical variables. The simultaneous Euler-Lagrange equations of motion for both sets of dynamical variables, atomic coordinates and electronic wavefunctions, avoid the explicit minimization of the DFT total energy at each step of the dynamics. Instead, they introduce a fictitious dynamics of the wavefunctions representing an adiabatic updating on-the-fly of the electronic structure along the atomic dynamics.

Introduction

The main target in atomic-scale simulations is to reproduce in a realistic way physical and chemical events occurring in materials. Specifically, the scope of First Principles Molecular Dynamics (FPMD) is to study a system of *interacting* nuclei and electrons by recreating it on a computer in a way as close as possible to nature and by simulating its dynamics over a *physical length of time* relevant to the

properties of interest. The inherent complexity of the simulated systems, from solids to biological macromolecules, calls for methods able to go beyond the simple calculation of the electronic structure of a given set of coordinates \mathbf{R}_I representing the positions of atoms. This is exactly the idea that started the entire field of Molecular Dynamics (MD).

From an historical point of view, the MD approach was introduced by Alder and Wainwright [1] in the late 1950's to study the interactions of hard spheres. Many important insights concerning the behavior of simple liquids emerged from their studies, but due to the limitations of the computational facilities and the pioneering stage of the MD, it was only in 1964 that the first dynamical simulation could be done. That milestone case focused on liquid Ar with the interatomic interaction modeled by a truncated Lennard-Jones potential [2]. In a nutshell, any MD method is an iterative numerical scheme for solving some equations of motion (EOM), representing the physical evolution of the system under study. Modeling the interaction of atoms with an analytic potential $V(\mathbf{R}_I)$, especially when chemical bonds evolve in time and they are broken or formed is a hard task not yet solved apart from a very limited class of chemical species. On the other hand, the electronic structure for a general many-body system can be determined with a computationally reasonable workload by means of the density functional theory (DFT), originally proposed in the early 60s by Kohn, Hohenberg and Sham [3,4]. Its importance in the advancement of computational quantum chemistry and related fields was acknowledged by the Nobel Prize in Chemistry in 1998 awarded jointly to Walter Kohn and John A. Pople. Joining the two fields, MD and DFT, is exactly what the Car-Parrinello method is about, extending the range of both concepts [5,6].

A Brief Overview of Density Functional Theory: The CPMD Potential

The DFT is a formulation of the many-body quantum mechanics in terms of an electron density distribution, $\rho(\mathbf{x})$, which describes the ground state of a system composed of interacting electrons and point-like nuclei having positions $\{\mathbf{R}_I\}$ [7]. All along the text, atom-

ic units will be used for simplicity. In practice, single-particle wavefunctions $\psi_i(\mathbf{x})$ are used to express the many-body mathematical function $\rho(\mathbf{x})$. The dramatic simplification, is the fact that not even the specific analytic form of the complex function $\psi_i(\mathbf{x})$ matters, but only its square modulus, so that the electron density reads

$$\rho(\mathbf{x}) = \sum_{i=1}^{N^{occ}} f_i |\psi_i(\mathbf{x})|^2 \quad (1)$$

This expression is clearly a single Slater determinant constructed from wavefunctions representing all the N^{occ} occupied orbitals. The coefficients f_i are the (integer) occupation numbers, and they are equal to 1 in the case in which the spin is explicitly considered (spin-unrestricted) or equal to 2 if the spin is neglected and energy levels are considered as doubly-occupied (spin-restricted). Furthermore, the wavefunctions $\psi_i(\mathbf{x})$ are subject to the orthonormality constraint

$$\int \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) d^3x = \delta_{ij} \quad (2)$$

as in any quantum mechanics approach. The Kohn-Sham (KS) DFT total energy of the system in its ground state is then written as

$$E^{KS}[\{\psi_i\}] = E_k[\{\psi_i\}] + E_H[\rho] + E_{xc}[\rho] + E_{el}[\rho] + E_{II} \quad (3)$$

In Eq. (3) the first three terms on the right-hand side (E_k , E_H , E_{xc}) describe all the electron-electron interactions, the fourth term (E_{el}) refers to the electron-nucleus interaction and the fifth one (E_{II}) corresponds to the nucleus-nucleus interaction. More explicitly, E_k is the Schrödinger-like kinetic energy expressed in terms of the single-particle wavefunctions $\psi_i(\mathbf{x})$ as

$$E_k[\{\psi_i\}] = \sum_{i=1}^{N^{occ}} f_i \int \psi_i^*(\mathbf{x}) \left(-\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{x}) d^3x \quad (4)$$

It must be remarked that this expression for the kinetic energy does not depend on the density $\rho(\mathbf{x})$ but directly on the wavefunctions.

The second term, E_H , is the Hartree energy, i.e. the Coulomb electrostatic interaction between two charge distributions

$$E_H[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{x})\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d^3x d^3y \quad (5)$$

The exchange interaction and the electron correlations due to many-body effects are represented by the term $E_{xc}[\rho]$, whose exact analytical expression is unknown. There are good approximations derived from the homogeneous electron gas limit for the exchange interaction [7], the so-called local density approximation (LDA), whose name comes from the fact that an homogeneous distribution of interacting electrons is assumed, in which $\rho(\mathbf{x})$ depends just on the local point \mathbf{x} . Similarly, in the LDA version of the correlation energy [7], the explicit analytic form of the functional comes from a parameterization of the results of quantum Monte Carlo calculations. Due to the insufficiency of a simple LDA approximation for many real systems, non-local approximations including the gradient of the density, are often adopted and the exchange-correlation functional becomes $E_{xc}[\rho, \nabla \rho]$. In practical applications, however, the gradient enters only with its modulus, thus adding only a modest computational cost. The electrostatic interaction between electrons and nuclei, is then

$$E_{el}[\rho] = - \int \sum_{I=1}^M \frac{Z_I \rho(\mathbf{x})}{|\mathbf{x} - \mathbf{R}_I|} d^3x \quad (6)$$

where Z_I is the charge of the I^{th} nucleus. However, in practice, this expression “as is” is computationally expensive. In fact, two different length scales come into play: a small one for the core electrons, characterized by rapidly varying wavefunctions, especially in the region very close to the nucleus, and a longer one for the valence electrons that form chemical bonds and vary more smoothly. Clearly, the first one would dominate and add a computational workload that would make impractical simulations of large systems. To overcome this problem, one can observe that core electrons are generally inert and do not participate to chemical bonds. This crucial observation led to the use of pseudopotentials [6]. Namely, core electrons are

eliminated and a potential describing the core-valence interaction is built by fitting to the all-electron solutions of the Schrödinger or Dirac equation for the single atom of the chemical species considered. In a pseudopotential (PP) approach, the electron-nucleus interaction is rewritten as

$$E_{el}[\rho] = \int d^3x V_{ps}(\mathbf{x} - \mathbf{R}_I) \cdot \rho(\mathbf{x}) \quad (8)$$

Finally, the fifth and last term in right-hand side of Eq. (3) is simply the Coulomb interaction between two classical nuclei I and J and is written as

$$E_{II} = \sum_{I < J}^M \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (9)$$

where Z_I and Z_J are the net valence charge in a PP approach. The total energy E^{tot} of the ground state of such a system of interacting electrons and nuclei is obtained by minimizing the KS functional with respect to the single-particle orbitals $\psi_i(\mathbf{x})$, which, in practice, means solving the KS Schrödinger-like equations

$$\frac{\delta E^{KS}}{\delta \psi_i^*} \equiv H^{KS} \psi_i(\mathbf{x}) = \varepsilon_i \psi_i(\mathbf{x}) \quad (10)$$

The Basis Set Issue

A somehow arbitrary issue is the proper definition of $\psi_i(\mathbf{x})$. The answer is the selection of a proper basis set on which orbitals can be expanded. One possible choice is a localized basis set expressing the one-electron wavefunctions as

$$\psi_i(\mathbf{x}) = \sum_{k=1}^M c_i^k \phi_k(\mathbf{x}; \{\mathbf{R}_I\}) \quad (11)$$

and the number of analytic functions used, M , is also an indicator of the computational cost of the quantum calculation, in the obvious sense that the larger the basis set, the higher the computational workload. One of the most popular basis sets is represented by Gaussian-type orbitals (GTO)

$$\phi_k(\mathbf{x} - \mathbf{R}_I) = \phi_k(\mathbf{r}) = N_k \cdot r_x^{k_x} r_y^{k_y} r_z^{k_z} \cdot \exp(-\alpha_k \cdot r^2) \quad (12)$$

where $\mathbf{r} = \mathbf{x} - \mathbf{R}_I$. When such a basis set is used, the constants N_k , and α_k are kept fixed during the electronic structure calculation, whereas the coefficients c_i^k are allowed to vary until they are fully optimized [8]. It must be remarked that orbitals expanded in a localized basis set depend on the atomic positions \mathbf{R}_I . As a consequence, in any calculation in which the forces acting on the ions are required, the explicit derivatives of these wavefunctions with respect to \mathbf{R}_I must be computed, leading to non-Hellmann-Feynman force components known in the literature as Pulay forces [6,8]. An alternative basis set rather popular in physics is represented by plane waves (PW)

$$\psi_i(\mathbf{x}) = \sum_{\mathbf{G}=0}^{\mathbf{G}^{\max}} c_i(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{x}} \quad (13)$$

where the sum is truncated at a suitable cut-off \mathbf{G}^{\max} . In this case, no dependence on the atomic coordinates and no arbitrariness in the increase in the number of basis functions exist.

First Principles Molecular Dynamics

Until the early 1980s, few applications of DFT went beyond the static calculations of the electronic structure. Nonetheless, finite temperature and entropy effects are two of the dominant features in matter and their role is often far from negligible. In this respect, the FPMD has represented a huge step forward. In this particular combination of DFT and MD, the interactions among atoms, instead of being described by an analytical function $V(\{\mathbf{R}_I\})$ of the atomic coordinates

\mathbf{R}_I , is directly computed from the DFT total energy E^{DFT} , which is simultaneously a function of the electron wavefunctions and the atomic coordinates. The Born-Oppenheimer (BO) approximation [9] allows to disentangle the motions of the electrons and the nuclei, and each time the nuclei $\mathbf{R}_I(t)$ are displaced from given positions at time t to new positions $\mathbf{R}_I(t+\Delta t)$ at a subsequent time $t+\Delta t$, an optimization of the electronic structure has to be performed. Then the forces acting on the nuclei are estimated from the gradient of E^{DFT} with respect to the ionic position and the variables $\mathbf{R}_I(t)$ are updated to $\mathbf{R}_I(t+\Delta t)$ by solving via finite differences the Newton-like EOM

$$M_I \ddot{\mathbf{R}}_I = -\nabla_{\mathbf{R}_I} \min_{\{\psi_i\}} E^{KS}[\{\psi_i\}, \{\mathbf{R}_I\}] \quad (14)$$

This iterative procedure assumes that the electronic structure is recomputed and the full diagonalization of the Hamiltonian is performed at each time step t along the discrete trajectory $\{\mathbf{R}_I(t)\}$.

Car-Parrinello Molecular Dynamics

An alternative to this scheme, which has represented a real breakthrough in first principles dynamical simulations was proposed in 1985 by R. Car and M. Parrinello [5]. The scope (and driving force) was to overcome the two major efforts arising in FPMD: On one hand one has to integrate the equations of motion for the nuclear positions as in Eq. (14), which represent the long time scale part to the problem. On the other hand, one has to propagate dynamically the smooth time-evolving (ground state) electronic subsystem. The Car-Parrinello molecular dynamics (CPMD) is able to satisfy this second requirement in a numerically stable way and makes an acceptable compromise for the time step length of the nuclear motion. The formulation is an extension of a classical molecular dynamics Lagrangean in which the electronic degrees of freedom (wavefunctions) are added to the system, along with any other dynamical variable $q_\alpha(t)$, i.e. thermostats, barostats, reaction coordinates, etc.

$$\begin{aligned} \mathcal{L}^{CP} = & \frac{1}{2} \sum_I M_I \dot{\mathbf{R}}_I^2 + \sum_i \mu \int |\dot{\psi}_i(\mathbf{x})|^2 d^3x + \frac{1}{2} \sum_\alpha \eta_\alpha \dot{q}_\alpha^2 - E^{KS}[\rho, \{\mathbf{R}_I\}, q_\alpha] \\ & + \sum_{ij} \lambda_{ij} \left(\int d^3x \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) - \delta_{ij} \right) \end{aligned} \quad (15)$$

The first term on the right-hand side of Eq. (15) is the kinetic energy of the nuclei, the second one the fictitious kinetic energy of the electrons representing the update of the wavefunctions during the dynamics, the third one the kinetic term of any further dynamical variable used in the sense specified above, the fourth one is the DFT total energy, and the last addendum is the orthonormality constraint for the wavefunctions. The kinetic energy for the electronic degrees of freedom is the main novelty of the CPMD approach: A strategy to update on-the-fly the wavefunctions when ions undergo a dynamical displacement, avoiding expensive iterative diagonalization required by the BO approach at each time step. The Euler-Lagrange EOM are

$$\mu \ddot{\psi}_i(\mathbf{x}) = - \frac{\delta E^{KS}}{\delta \psi_i^*} + \sum_j \lambda_{ij} \psi_j(\mathbf{x}) \quad (16)$$

$$M_I \ddot{\mathbf{R}}_I = -\nabla_{\mathbf{R}_I} E^{KS} \quad (17)$$

$$\eta_\alpha \ddot{q}_\alpha = - \frac{\partial E^{KS}}{\partial q_\alpha} \quad (18)$$

The fictitious mass μ assigned to the orbitals $\psi_i(\mathbf{x})$ is the parameter that controls the speed of the updating of the wavefunctions with respect to the nuclear positions. For this reason, it determines the degree of adiabaticity of the two subsystems, electrons and nuclei.

It is straightforward to give a Hamiltonian, instead of a Lagrangean, formulation of the CPMD method, via a simple Legendre transform after defining the momenta

$$\pi_i(\mathbf{x}) = \frac{\delta \mathbf{L}^{CP}}{\delta \dot{\psi}_i^*(\mathbf{x})} = \mu \dot{\psi}_i(\mathbf{x}) \quad \pi_i^*(\mathbf{x}) = \frac{\delta \mathbf{L}^{CP}}{\delta \dot{\psi}_i(\mathbf{x})} = \mu \dot{\psi}_i^*(\mathbf{x}) \quad (19)$$

$$\mathbf{p}_i = \nabla_{\dot{\mathbf{R}}_i} \mathbf{L}^{CP} = M_I \dot{\mathbf{R}}_I \quad (20)$$

$$\dot{\zeta}_\alpha = \frac{\partial \mathbf{L}^{CP}}{\partial \dot{q}_\alpha} = \eta_\alpha \dot{q}_\alpha \quad (21)$$

so that the Hamiltonian reads

$$H^{CP} = \sum_I \frac{\dot{\mathbf{p}}_I^2}{2M_I} + \sum_i \int \frac{\pi_i^*(\mathbf{x}) \pi_i(\mathbf{x})}{\mu} d^3x + \sum_\alpha \frac{\dot{\zeta}_\alpha^2}{2\eta_\alpha} - E^{KS}[\rho, \{\mathbf{R}_I\}, q_\alpha] - \sum_{ij} \lambda_{ij} \left(\int \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) d^3x - \delta_{ij} \right) \quad (22)$$

and the CPMD equations of motion (Fig. 1) will be given by the corresponding Hamilton EOMs.

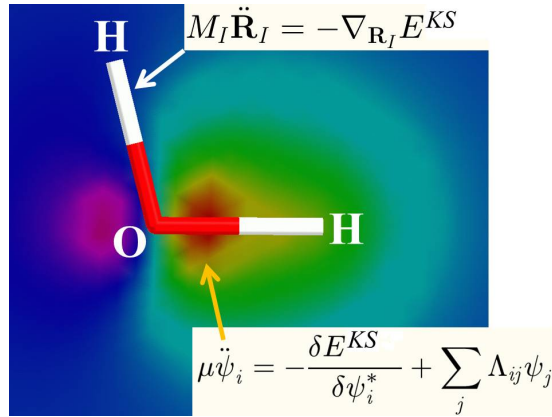


Fig. 1. Atomic structure (sticks; red = O, white = H) of a water molecule and electronic wavefunction of an O-H s-bond in terms of density map (blue: $|\psi(\mathbf{x})|^2=0$, red: $|\psi(\mathbf{x})|^2=\text{maximum}$). The two sets of dynamical variable evolve in time according to the equations of motions indicated, coupled via E^{KS} .

A rigorous mathematical proof of this scheme has been given by Bornemann and Schütte [10], showing that the CPMD trajectory $\{\mathbf{R}^{CP}(t)\}$ stays close to the BO one $\{\mathbf{R}^{BO}(t)\}$ and the upper bound is proportional to the square root of the fictitious mass μ

$$|\mathbf{R}^{CP}(t) - \mathbf{R}^{BO}(t)| < C \cdot \sqrt{\mu} \quad (23)$$

where C is a positive constant (Fig. 2).

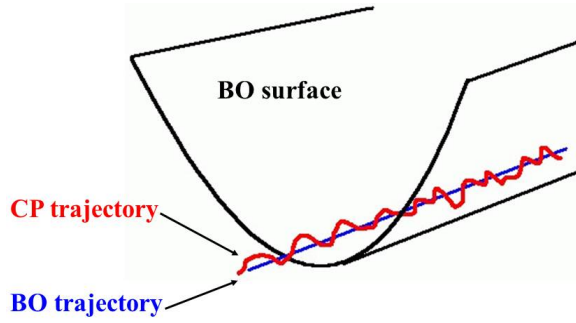


Fig. 2. Schematic representation of a Car-Parrinello trajectory (red line) with respect to a Born-Oppenheimer dynamics (blue line) on a given DFT-based potential energy surface.

The fact that the CPMD was a milestone step forward in realistic simulations of materials at various thermodynamics conditions can be easily seen by the number of publications in first principles molecular dynamics (FPMD) before and after 1985. Indeed, the original Car-Parrinello publication has more than 7000 citations in 2014 (source: ISI Web of Science), and, to acknowledge the importance of the method, the international Physics and Astronomy Classification Scheme (PACS) introduced in 1996 a new identification number, 71.15.Pd, to classify Car-Parrinello related publications. Since then, the method has been applied to a wide variety of materials, ranging from solids, to liquids and to biological systems [11,12].

Numerical Details

Although it is not restriction neither of DFT [13] nor of CPMD [14], PWs are often used as a convenient basis set for the coding of

CPMD, since they have several good properties: (i) Accuracy can be systematically improved in a fully variational way, (ii) PWs are independent from atomic positions (i.e. no Pulay forces [6]), (iii) PWs can be easily distributed in parallel processing. However, it must be observed that the fact that PWs they are not localized can lead to inefficiencies for small clusters or surfaces placed in a large simulation cell. The equations of motion are discretized by finite differences, via a Verlet, or velocity-Verlet, algorithm [15]. The ionic variables $\mathbf{R}_i(t)$ are updated at a rate Δt , while the electronic degrees of freedom are updated at a rate $\Delta t/\mu^{1/2}$, i.e.

$$\begin{aligned} \frac{\mu}{\Delta t^2} [c_i(\mathbf{G}, t + \Delta t) + c_i(\mathbf{G}, t - \Delta t) - 2c_i(\mathbf{G}, t)] = \\ - \sum_{\mathbf{G}'} \langle \mathbf{G} | H^{CP} | \mathbf{G}' \rangle c_i(\mathbf{G}') + \sum_j \Lambda_{ij} c_j(\mathbf{G}) \end{aligned} \quad (24)$$

For most of the applications, Δt and μ fall in the range 3-5 au and 300-600 au, respectively. Of course, the (quantum) time scale of electrons is dominating in this kind of approaches and simulation times are of the order of few tens or, at very best, hundreds of ps. As far as the system size is concerned, with N electrons and \mathbf{G}^{\max} PWs, \mathbf{G}^{\max} being integer, the scaling of the various parts composing the CPMD algorithm is $O(N \mathbf{G}^{\max})$ for the kinetic term, $O(N \mathbf{G}^{\max} \log \mathbf{G}^{\max})$ for the local potential and $O(N^2 \mathbf{G}^{\max})$ for both the non-local term and wavefunctions orthogonalization procedure.

Second-Generation Car-Parrinello Molecular Dynamics

The inherent high computational cost associated to the electronic structure calculations has limited the affordable time scales for several years. Only the most advanced high performing computer platforms recently available have allowed to increase the system size to about a thousand of atoms and simulation times towards hundreds of picosecond. Yet, many phenomena still call for a substantial boost. These are, for instance, diffusion in solids or, in the case of glasses generation from the melt, a less rapid cooling rate suitable to avoid

numerically induced structural problems. While linear scaling methods can be a viable way to access larger system sizes, they still have to face the problem of the simulation time scale. Moreover, the crossover point at which linear scaling methods become advantageous has remained fairly large, especially if high accuracy is needed. An interesting attempt at overcoming these limitations has been proposed in 2007 [16]. The basic idea is to join the advantages of both the BO approach and the CPMD; in a “nutshell”

	CPMD	BO
Conservation of constants of motion	Good	Convergence dependent
Electronic optimization	Not needed	Needed
Hamiltonian diagonalization	Not needed	Needed
Integration step Δt	Small	Large
Minimum of the BO surface	Approximate	Exact

These two approaches have nearly complementary features as sketched above. Following the CPMD formulation, it can be remarked that the Lagrangean formulation for the propagation of the wavefunctions is stable by construction, thus providing a reliable integration. This *stability* feature must then be preserved. Concerning the *efficiency*, large integration steps Δt are desirable and possibly a small, or better zero, deviation from the BO surface should be kept all along the dynamics to get a high *accuracy*. The mathematical result of this list of requirement resumes into a modified ionic EOM

$$M_I \ddot{\mathbf{R}}_I = -\frac{\partial E_{NSC}}{\partial \mathbf{R}_I} + \sum_{i,j} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle - 2 \frac{\partial \langle \psi_i |}{\partial \mathbf{R}_I} \left(\frac{\delta E_{NSC}}{\delta \langle \psi_i |} - \sum_j \Lambda_{ij} | \psi_j \rangle \right) \quad (25)$$

While the first term in the right hand of the equation is clear, the rest seems a bit puzzling at a first glance. Indeed, in the original formulation [16], the selected basis set is not PW, but a localized basis set as in equation (11). Hence, the electronic wavefunctions depend also on the atomic coordinates \mathbf{R}_I and the request of orthogonality at each step is relaxed to save time, meaning that the scalar product $\langle \psi_i | \psi_j \rangle$ is no longer vanishing. Analogously, the total energy E^{DFT} is not re-

optimized as in full self-consistent BO procedures and for this reason is indicated as non-self-consistent energy E_{NSC} . Wavefunctions are propagated according to an algorithm which resembles the original CPMD formulation in the sense that second order EOMs are used, but a damping term (first-order derivative) is present which reminds a sort of steepest-descent algorithm typical of the BO dynamics. The net result is the electron dynamics,

$$\mu \frac{d^2}{dt^2} |\psi_i\rangle + \gamma \frac{d}{dt} |\psi_i\rangle = - \frac{\delta E_{NSC}}{\delta \langle \psi_i |} + \sum_j \Lambda_{ij} |\psi_j\rangle \quad (26)$$

which is then solved via a predictor-corrector scheme. With no pretention of completeness, the procedure can be summarized as follows. On a first instance, in a localized basis set $\{|q\rangle\}$, the electronic wavefunctions are expanded as

$$|\psi_i\rangle = \sum_{q=1}^M C_{iq} |q\rangle \quad (27)$$

on the M functions composing the localized basis. Then the $N \times M$ matrix of the expansion coefficient is written as

$$\mathbf{C} = [C_{iq}] = \begin{pmatrix} C_{11} & \cdot & \cdot & C_{1M} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ C_{N1} & \cdot & \cdot & C_{NM} \end{pmatrix} \quad (28)$$

and the density matrix becomes $\mathbf{P} = \mathbf{C}\mathbf{C}^T = \mathbf{P}\mathbf{S}\mathbf{P}$. The $M \times M$ matrix indicated as \mathbf{S} is just given by the expansion coefficient and its matrix elements have the usual form

$$S_{qq'} = \sum_{i=1}^N C_{iq}^* C_{iq} \quad (29)$$

Hence, the (DFT) total energy can be rewritten as $E^{\text{tot}}[\mathbf{C}, \mathbf{R}_I]$ which can be used in a straightforward way to write a BO dynamics

$$M\ddot{\mathbf{R}}_I = -\nabla_{\mathbf{R}_I} \min \{E^{\text{tot}}[\mathbf{C}, \mathbf{R}_I]\} \quad \forall \mathbf{C}^T \cdot \mathbf{S} \cdot \mathbf{C} = \mathbf{1} \quad (30)$$

under the given constraint on \mathbf{C} which resumes in an implicit orthogonality condition. However it must be kept into account that: (i) Diagonalization and minimization of E^{tot} are required in BO; (ii) Hellman-Feynman forces are just one component since Pulay forces due to the local basis set are present. Residual force components appear due to non-self-consistency (NSC) of the approach. To take into account all the points above, the basic strategy can be summarized in four major points:

1. Propagate the electronic variables in time according to the CP original idea of updating on-the-fly to avoid expensive full diagonalization operations
2. Use a good propagation algorithm $\mathbf{C}(t_n)=f(\mathbf{C}(t_{n-1}),\dots,\mathbf{C}(t_{n-m}))$ depending on previous time steps $m \in [1, K]$ time steps
3. Select the appropriate number of steps K to keep $\mathbf{C}(t_n)$ as close as possible to the (electronic) ground state
4. Enforce convergence on the BO surface, correct this propagation CPMD-like afterwards

Point 3 corresponds to the first move in the numerical integration procedure and it can be identified as the “predictor” part directly deriving from a standard numerical integration of the CPMD type equations of motion. Point 5, instead, is the “corrector” needed afterwards to better converge the wavefunctions and to restore the neglected self-consistent loop. The use of not necessarily fully converged wavefunctions at the predictor propagation stage allows for large integration steps, thus resulting in a remarkable boost in the dynamics.

REFERENCES

- [1] Alder, B.J. , Wainwright, T.E.J.: Phase Transition for a hard Sphere System. *Chem. Phys.* **27**, 1208 (1957)
- [2] Rahman, A.: Correlation in the Motion of Atoms in Liquid Argon. *Phys. Rev.* **136**, A405 (1964)
- [3] Hohenberg, P., Kohn, W.: Inhomogeneous Electron Gas. *Phys. Rev.* **136**, B864 (1964)
- [4] Kohn, W., Sham, L. J.: Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140**, A1133 (1965)

- [5] Car, R., Parrinello, M.: Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **55**, 2471(1985)
- [6] Marx, D., Hutter, J.: *Ab initio molecular dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, New York, 2009). ISBN 978-0521898638
- [7] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989). ISBN 0-19-504279-
- [8] Hehre, W.J., Radom, L., Schleyer, P.v.R and Pople, J.A. *Ab Initio Molecular Orbital Theory* (John Wiley & Sons, New York, 1986). ISBN 978-0471812418
- [9] Born, M., Oppenheimer, J.R.: Zur Quantentheorie der Molekeln. *Annalen der Physik* **84**, 457 (1927)
- [10] Bornemann, F.A., Schütte, C.: A Mathematical Investigation of the Car-Parrinello Method. *Numerische Mathematik* **78**, 359 (1998)
- [11] Boero, M., Tateno, M.: *Quantum Theoretical Approaches to Proteins and Nucleic Acids in Oxford Handbook of Nanoscience and Technology, Vol. 1: Basic Aspects*, pp. 549-598 (Oxford University Press, New York, 2010). ISBN 978-0199533046
- [12] Boero, M.: *Reactive Simulations for Biochemical Processes*, in *Atomic-Scale Modeling of Nanosystems and Nanostructured Materials – Lect. Notes Phys.* **795**, pag. 81-98, Springer, Berlin Heidelberg 2010. ISBN 978-3-642-04650-6
- [13] Oshiyama, A., Iwata, J.: Large-Scale Electronic-Structure Calculations for nanomaterials in Density Functional Theory. *J. Phys.: Conf. Ser.* **302**, 012030 (2011)
- [14] Car, R., Parrinello, M.: The Unified Approach to Density Functional and Molecular Dynamics in Real Space. *Solid State Comm.* **62**, 403 (1987)
- [15] Pang, T.: *An Introduction to Computational Physics* (Cambridge University Press, New York, 2000). ISBN 978-1840858839
- [16] Kühne, T.D., Krack, M., Mohamed, F.R. Parrinello, M.: Efficient and Accurate Car-Parrinello-like Approach to Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.* **98**, 066401 (2007)

Cross References

[Ab Initio DFT Simulations of Nanostructures - 340082](#)
[Computer Modeling and Simulation of Materials - 340375](#)
[Electronic Structure Calculations - 340436](#)
[First Principles Calculations - 340462](#)
[Molecular Dynamics Method - 340660](#)
[Molecular Dynamics Simulations of Interactions Between Biological Molecules and Nanomaterials - 340661](#)
[Molecular Dynamics Simulations of Nano-Bio Materials - 340204](#)
[Surface Electronic Structure - 340183](#)