UNIVERSITÀ
DI TRENTO

Quantum molecular dynamics of the hydrogen molecule

Giuseppe Gambini

Contents

1 ABSTRACT

This project was made as a requirement for the course of Computational Physics held by professors Francesco Pederiva and Alessandro Roggero at the University of Trento.

In this report, we will treat the simple hydrogen molecule dynamics through different *ab initio* methods such as the Born-Oppenheimer Molecular Dynamics (BOMD), the Car-Parrinello Molecular Dynamics (CPMD) and the Conjugate Gradient Molecular Dynamics (CGMD). Among these, the most successful algorithm is certainly CPMD due to its degree of accuracy and its computational versatility. A large part of the discussion will be dedicated to this algorithm and how it differs from the other two. After a theoretical overview of the three algorithms, we will analyse a practical implementation for the hydrogen molecule. All the methods require the calculation of the electronic energy from first principles. In doing this, Hartree-Fock (HF) and Density Functional Theory (DFT) have been employed. The details and a discussion of these methods are included in the work.

2 THEORETICAL INTRODUCTION

2.1 *Ab initio* methods

The correct approach in modelling the dynamics in a quantum perspective involves solving the time-dependent Schrödinger equation

$$i\frac{\partial\Phi(\{\mathbf{R}_I\},\{\mathbf{r}_i\},t)}{\partial t} = \hat{H}\Phi(\{\mathbf{R}_I\},\{\mathbf{r}_i\},t) \tag{1}$$

where $\hat{H}$ is the Hamiltonian operator given by

$$\begin{aligned}\hat{H} &= -\sum_I \frac{\nabla_I^2}{2M_I} - \sum_i \frac{\nabla_i^2}{2} + \sum_{i<j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i,I}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|} + \sum_{I<J}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} \\ &= \hat{T}_N + \hat{T}_e + \hat{V}_{ee} + \hat{V}_{eN} + \hat{V}_{NN} = \hat{T}_N + \hat{H}_e\end{aligned}$$

Here atomic units (a.u.) are used and the nuclear masses are given in terms of the electron mass. The indices with capital letters refer to the nuclear coordinates, while the lowercase letters refer to electrons.
A possible approach to this problem in (1) is to solve the time-independent equation $\hat{H}_e\psi = E\psi$ (where $\hat{H}_e$ is the electronic Hamiltonian) at different geometries and obtain the potential energy surface (PES) given by $E(\{\mathbf{R}_I\})$. Once the PES is obtained, it is usually fit to some analytical functional form and employed as the potential in the time-dependent Schrödinger equation. Although these empirical potentials have been created to be as accurate as possible, this approach, known as *force field* approach, is not the most accurate, especially when dealing with extended systems. In fact, the transferability to systems or phase space regions[1] different from the ones to which they have been fitted is restricted and may lead to inaccuracy. Instead, the *ab initio molecular dynamics* method is preferred. In this approach the electrons are treated quantum mechanically, but nuclei are treated as classical particles. In other words, in an *ab initio* method, the positions of the nuclei $\{\mathbf{R}_I\}$ evolve according to the Newton's law as in classical *molecular dynamics* (MD), but no *force field* is required for the description of the behaviour of the bonds and the angles. The drawback of the inclusion of classical dynamics at the nuclear level is that the zero point vibrational energy and tunneling effects do not emerge in this approach. In addition, the tunneling effect might be important for systems involving hydrogen atoms since their nuclei are the lightest in nature.

A crucial aspect in every MD code is the choice of the Δt time step since it needs to be small enough to describe the timescales of interest. A good choice is usually a Δt of the order of femtoseconds, as the typical vibrational timescale of a molecule is around 10^{-11} to 10^{-14} s. This choice prevents us from sampling chemical processes such as bond breaking and reforming, which normally occur on a much larger timescale (about 10^{-6} to 10^{-10} s). The high chemical accuracy of *ab initio* methods prevents the access to long simulations times that would be affordable via standard molecular dynamics.

2.2 Classical limit recovered for nuclear motion

Contrarily to the rest of the text, in this section we will explicitly show the $\hbar$. The ideas presented are discussed in detail in the review article by Marx and Hutter[1] about *ab initio* molecular dynamics. We indicate with $\chi = \chi(\{\mathbf{R}_I\},t)$

[1] For instance, the stretching of a molecule, is usually carried out only for small deviations from the equilibrium conformation.

the nuclear wavefunction and with $\psi = \psi(\{\mathbf{r}_i\}, t)$ the electronic wavefunction. Henceforth we avoid the explicit notation with t, $\{\mathbf{R}_I\}$ and $\{\mathbf{r}_i\}$ dependencies.
Our aim is to find a justification for the classical treatment of the nuclei. We start by writing the total wavefunction $\Phi = \Phi(t, \{\mathbf{R}_I\}, \{\mathbf{r}_i\})$ as a product ansatz in which nuclear and electronic coordinates are treated separately:

$$\Phi \simeq \chi\psi \exp\left(\frac{i}{\hbar}\int_{t_0}^{t} dt'\ E_e(t')\right) \qquad\qquad E_e(t) = \int\left(\prod_{i,I} d\mathbf{r}_i d\mathbf{R}_I\right)\ \chi^*\psi^* H_e \psi\chi \tag{2}$$

Note that this ansatz differs significantly from the Born-Oppenheimer ansatz, in which the total wavefunction is expanded at any time as

$$\Phi = \sum_k^{\infty} \widetilde{\psi}_k(\{\mathbf{r}_i\}, \{\mathbf{R}_I\})\widetilde{\chi}_k(\{\mathbf{R}_I\})$$

where $\{\widetilde{\psi}_k\}$ is a complete orthogonal set for the electronic problem at clamped nuclei.
Inserting the ansatz of (2) in (1) and multiplying on the left by $\langle\psi|$ and $\langle\chi|$ leads to the following expression:

$$i\hbar\left\langle\psi\middle|\frac{\partial\psi}{\partial t}\right\rangle + i\hbar\left\langle\chi\middle|\frac{\partial\chi}{\partial t}\right\rangle - E_e(t) = \langle H\rangle$$

which is satisfied if the following differential equations hold:

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2\psi + \left[\int\left(\prod_I d\mathbf{R}_I\right)\ \chi^*(\hat{V}_{ee} + \hat{V}_{eN} + \hat{V}_{NN})\chi\right]\psi \\ i\hbar\frac{\partial\chi}{\partial t} &= -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2\chi + \left[\int\left(\prod_i d\mathbf{r}_i\right)\ \psi^*\hat{H}_e\psi\right]\chi \end{aligned} \tag{3}$$

The ansatz in (2) produces a *mean-field* picture of the nuclear and electronic dynamics. In fact, the movement of both electrons and nuclei is affected by the *effective potentials* in square brakets.

Classical mechanics for the nuclear motion can be extrapolated from the quantum equation in (3) in the limit $\hbar \to 0$ in which positions and momenta can be known with maximum accuracy. To prove this statement, we show that the ansatz $\chi = A(\{\mathbf{R}_I\}, t)\exp(iS(\{\mathbf{R}_I\}, t)/\hbar)$, where both the amplitude A and the action S are real quantities, leads to an equation for the action which is isomorphic to the Hamilton-Jacobi equation. We replace the ansatz for χ in the nuclear equation in (3) and isolate the real part:

$$\frac{\partial S}{\partial t} + \sum_I \frac{(\boldsymbol{\nabla}_I S)^2}{2M_I} + \int\left(\prod_i d\mathbf{r}_i\right)\ \psi^*\hat{H}_e\psi = \hbar^2\sum_I \frac{1}{2M_I}\frac{\nabla_I^2 A}{A} \tag{4}$$

In the classical limit $\hbar \to 0$, the term on the right side of (4) vanishes and we are left with a Hamilton-Jacobi equation (for more details see the dedicated section 5.1 in the appendix). The connecting transformation $\mathbf{P}_I = \boldsymbol{\nabla}_I S$ leads to the "classical" Hamiltonian (with classical momenta)

$$\hat{H}(\{\mathbf{R}_I\}, \{\mathbf{P}_I\}) = \sum_I \frac{\mathbf{P}_I^2}{2M_I} + \int\left(\prod_i d\mathbf{r}_i\right)\ \psi^*\hat{H}_e\psi$$

The integral on the electron coordinates is a function of the nuclear coordinates $\{\mathbf{R}_I\}$ since $\hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{eN} + \hat{V}_{NN}$. Now, the integral over the nuclear coordinates $\{\mathbf{R}_I\}$ in (3) in the electronic equation can be solved since the nuclear density $\chi^*\chi$ is just a point-like particle density, i.e., $\prod_I \delta(\mathbf{R}_I - \mathbf{R}_I(t))$. If we now write again the first equation for the electrons in (3), we obtain

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}_e(\{\mathbf{r}_i\}, \{\mathbf{R}_I\})\psi$$

It must be reminded that this equation contains an implicit parametrical dependence on the nuclear position $\{\mathbf{R}_I\}$. From a computational point of view, this differential equation, solved together with the classical equation for the nuclei, generates the so-called Ehrenfest dynamics. The alternative is to solve the *static* electronic problem at any time for a fixed set $\{\mathbf{R}_I\}$ and then evolve the nuclear positions according to the *mean field* $\langle\psi|\hat{H}_e|\psi\rangle$. This approach is known as Born-Oppenheimer molecular dynamics (BOMD).

	Ehrenfest	Born-Oppenhemier
Nuclear evolution	$M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I\langle\psi\|\hat{H}_e\|\psi\rangle$	$M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I \min_\psi\langle\psi\|\hat{H}_e\|\psi\rangle$
Electronic evolution	$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}_e\psi$	$\hat{H}_e\psi = E\psi$

Table 1. In BOMD the time-dependence of the electronic structure is given by the evolution of the nuclear coordinates and it is not implicitly inserted as in the differential equation of the Ehrenfest dynamics.

2.3 Born-Oppenheimer molecular dynamics (BOMD)

The central point of BOMD is the separation of the dynamics of nuclei and electrons due to the several orders of magnitude of difference between their masses. One important assumption is that electrons follow *adiabatically* the nuclear motion, meaning that the perturbation provided by the displacement of the nuclei is sufficiently slow to maintain the electronic configuration in its ground state.
The nuclear classical problem and the variational minimisation of the energy subject to the orthonormality constraint of the orbitals are often cast in the Lagrangian formalism:

$$\mathcal{L} = \frac{1}{2}\sum_I M_I\dot{\mathbf{R}}_I^2 - \langle\psi|\hat{H}_e|\psi\rangle + \sum_{jk}\Lambda_{jk}\left[\langle\varphi_j|\varphi_k\rangle - \delta_{jk}\rangle\right] \tag{5}$$

where φ_i are the electron orbitals and the total wavefunction ψ is the Slater determinant built from the whole set $\{\varphi_i\}$. We remind that, in general, the orbitals φ_i depend on both the electron $\{\mathbf{r}_i\}$ and parametrically on nuclear $\{\mathbf{R}_I\}$ coordinates[2]. The Euler-Lagrange equations of motion are reported below:

$$\begin{aligned}
&\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\varphi}_i^*} - \frac{\partial\mathcal{L}}{\partial\varphi_i^*} = 0 \implies \hat{H}_e\varphi_i = \sum_j\Lambda_{ij}\varphi_j \xrightarrow{\text{unitary transformation}} \hat{H}_e\varphi_i = \varepsilon_i\varphi_i \\
&\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{R}}_I} - \frac{\partial\mathcal{L}}{\partial\mathbf{R}_I} = 0 \implies M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I\min_\psi\langle\psi|\hat{H}_e|\psi\rangle + \sum_{jk}\Lambda_{jk}\boldsymbol{\nabla}_I\langle\varphi_j|\varphi_k\rangle
\end{aligned} \tag{6}$$

From the last equations we see that the nuclear dynamics is affected by the orthonormality constraint of the orbitals and the gradient of the variational energy. The forces can be computed with the Hellman-Feynman theorem:

$$\mathbf{F}_I = -\boldsymbol{\nabla}_I\min_\psi\langle\psi|\hat{H}_e|\psi\rangle = -\langle\psi_{\min}|\boldsymbol{\nabla}_I\hat{H}_e|\psi_{\min}\rangle \tag{7}$$

since the theorem is valid for the exact ground state and still holds also for variationally optimised wavefunctions.
For further details, a practical implementation of the BOMD for the hydrogen molecule is discussed in section 3.6.

2.4 Car-Parrinello molecular dynamics (CPMD)

In 1985 Car and Parrinello pioneered a method involving an extended Lagrangian with an artificial term that treats the electronic motion in classical terms. More specifically, Car and Parrinello assigned a *fictitious* time dependence to the electronic wave function and constructed a Lagrangian by adding a fictitious kinetic energy for the electronic degrees of freedom:

$$\mathcal{L} = \frac{1}{2}\sum_{i=1} m_i\langle\dot{\varphi}_i(\mathbf{r})|\dot{\varphi}_i(\mathbf{r})\rangle + \frac{1}{2}\sum_I M_I\dot{\mathbf{R}}_I^2 - \langle\psi|\hat{H}_e|\psi\rangle + \sum_{jk}\Lambda_{jk}\left[\langle\varphi_j|\varphi_k\rangle - \delta_{jk}\rangle\right] \tag{8}$$

Here m_i is a *fictitious* mass that should not be confused with the electron mass. Although m_i has the units of an energy per time squared, its effective role is the same of a mass, i.e., controlling the acceleration of the associated degrees of

[2] If a local basis set is used for representing the orbitals, the basis functions will depend on the quantities $|\mathbf{r}_i - \mathbf{R}_I|$. If a plane waves basis set is used instead, local dependencies are removed.

freedom. Note that the real kinetic energy of the electrons is already contained in the Hartree-Fock or Kohn-Sham energy $\langle\psi|\hat{H}_e|\psi\rangle$, hence, the fictitious kinetic energy seems to be a redundant term. For this reason, in a Car-Parrinello simulation we need to assure that this kinetic energy is kept well below the typical energy scale of the system.

The following equations of motion are obtained from the Lagrangian in (8):

$$\begin{aligned}
&\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\varphi}_i^*} - \frac{\partial\mathcal{L}}{\partial\varphi_i^*} = 0 \implies m_i\ddot{\varphi}_i(\mathbf{r}) = -\hat{H}_e\varphi_i + \sum_l \Lambda_{ij}\varphi_j(\mathbf{r}) \\
&\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{R}}_I} - \frac{\partial\mathcal{L}}{\partial\mathbf{R}_I} = 0 \implies M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I\langle\psi|\hat{H}_e|\psi\rangle + \sum_{jk}\Lambda_{jk}\boldsymbol{\nabla}_I\langle\varphi_j|\varphi_k\rangle
\end{aligned} \tag{9}$$

It must be pointed out that only the second equation has a real physical meaning. The equations of motion in (9) are clearly coupled because of the implicit dependence of $\hat{H}_e$ and φ_i on $\{\mathbf{R}_I\}$, thus, unlike BOMD, the electronic and nuclear degrees of freedom can exchange energy. This concept is better illustrated in the figure below:

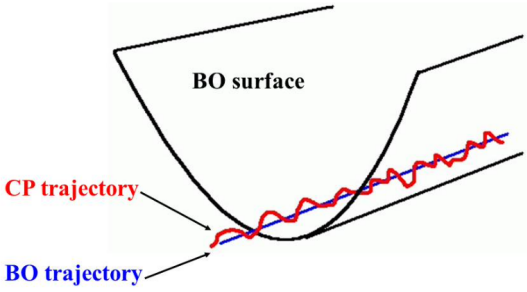


Figure 1. Schematic representation taken from the work of M.Boero and A.Oshiyama[2]. If the fictitious kinetic energy is relevant, the deviation of the CP trajectory from the BO surface becomes significant.

From a computational point of view, the advantage of CPMD is that explicit diagonalisation of the Fock (or Fock-Kohn-Sham) matrix in the electronic self-consistent problem can be avoided.
Another important aspect regarding the equations in (9) is that the force $-\boldsymbol{\nabla}_I\langle\psi|\hat{H}_e|\psi\rangle$ acting on the nuclei is not computed in the exact variational minimum as in BOMD. Consequently, the Hellman-Feynman theorem cited in (7) is not strictly valid and the observables computed on a trajectory at every time step are not optimised in a variational sense, but derive from fictitious forces acting on the electronic degrees of freedom.
Another important observation about CPMD must be made: quantum dynamics imposes the form of the Lagrangian in (8) and the form of the $\hat{H}_e$ operator, but the equations of motion are solved as classical ones. Consequently, the Lagrange multipliers are devised in such a way as to conserve total energy. The multipliers can be computed, for instance, with the help of the iterative SHAKE method[3] or with other algorithms, such as the one presented in section 3.7, in which we provide a practical implementation of CPMD for the H_2 molecule.

2.5 Conjugate gradient molecular dynamics (CGMD)

The minimisation problem of the energy can be considered as an abstract numerical problem and any minimisation algorithm can be applied. For this type of methods, the equations of motion for the nuclei is the same reported in (6), while the electronic equation is replaced with a minimisation procedure.
We remind that, in a CPMD simulation, constraints are enforced using a Lagrangian multiplier in such a way as to conserve energy. In minimisation methods, instead, the goal is to lose energy as efficiently as possible to reach the ground state. This can be done with two main techniques:

- *Steepest descent*: let us consider a function $F(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ with a single minimum. Starting from a point $\mathbf{x}_0$, the best direction for the search of the minimum is $-\boldsymbol{\nabla}F|_{\mathbf{x}_0}$. Then it is possible to locate the minimum with a line minimisation. This process minimises only the value of the function along a particular direction in the multidimensional space. Therefore, multiple line minimisations have to be performed subsequently. In all these minimisations, each steepest descent vector is orthogonal to the previous one.
- *Conjugate gradient*: contrarily to the steepeset descent approach, the minimisation directions depend one on the other in such a way as to reduce by one the dimensionality of the vector space explored. In other words, all the minimisation directions are chosen so that the function remains minimised along the previous direction.

From a practical point of view, the conjugate gradient method is considered as superior to the steepest descent for reasons of convergence. In fact, the choice of taking the steepest descent vectors all orthogonal is not always guaranteed to be the best choice, especially when dealing with functions showing plateau regions.

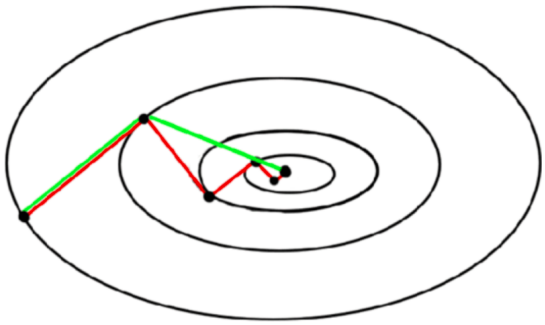


Figure 2. The steepest descent path is reported in red, while the conjugate gradient in green. The function (represented with the black contour lines) is quadratic. The convergence of the steepest descent method in two steps is guaranteed only in case of circular countour lines.

For a practical implementation of CGMD for H_2, see section 3.8.

2.6 Overview of the methods

The basic idea underlying every *ab initio* method is to compute the forces acting on the nuclei from electronic structure calculations performed "on–the–fly" as the trajectory is produced. In this context, both BOMD and CPMD are derived from the time-dependent mean–field approach that is obtained after separating the nuclear and electronic degrees of freedom. The conjugate gradient, which is also successfully employed in contexts other than *ab initio* simulations, also represents another efficient method driving the electronic minimisation. In the following table, we briefly summarise the three types of dynamics implemented in the report.

	BOMD	**CPMD**	**CGMD**
Nuclear problem	$M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I \min_\psi \langle\psi\|\hat{H}_e\|\psi\rangle$	$M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I \langle\psi\|\hat{H}_e\|\psi\rangle$	$M_I\ddot{\mathbf{R}}_I = -\boldsymbol{\nabla}_I \min_\psi \langle\psi\|\hat{H}_e\|\psi\rangle$
Electronic problem	$\hat{H}_e\psi = E\psi$	$m_i\ddot{\varphi}_i(\mathbf{r}) = -\hat{H}_e\varphi_i + \sum_l \Lambda_{ij}\varphi_j(\mathbf{r})$	$\min_\psi \langle\psi\|\hat{H}_e\|\psi\rangle$
self-consistency	✓	/	/
nuclear-electronic coupling	/	✓	/

3 METHODS AND RESULTS

3.1 Local basis formalism

The general choice for the spin formalism is to use different spatial orbitals for the *up* and *down* electrons, leading to a *spin-unrestricted* formalism. This formalism is able to capture the breaking of the spin symmetry (and consequently, the bond breaking) when molecules are highly stretched. If, as in the case of this project, the aim is to study the dynamics of an "associative" configuration of a H_2 molecule and not the dissociation, this approach is unnecessary. We can in fact impose the spatial orbitals to be the same for both spins, yielding less variational freedom and preventing the dissociation of the H_2 molecule. The basis Gaussian functions employed in the program $\chi_p(\mathbf{r})$ are centred around two nuclear positions $\mathbf{R}_A$ and $\mathbf{R}_B$. The most natural choice is represented by the atomic orbitals of the individual hydrogen atoms.

$$\varphi(\mathbf{r}) = \sum_p C_p \chi_p(\mathbf{r}) = \sum_{p=1}^{4} C_p^A \exp(-\alpha_p |\mathbf{r}-\mathbf{R}_A|^2) + C_p^B \exp(-\alpha_p |\mathbf{r}-\mathbf{R}_B|^2)$$
$$\mathbf{C} = (C_1^A ... C_4^A, C_1^B ... C_4^B)^\top \tag{10}$$

α_1	13.00773
α_2	1.962079
α_3	0.444529
α_4	0.1219492

The exponents $\{\alpha_p\}$ representing a STO-4G basis functions are suggested in the book by Thijssen[4]. Thanks to the Pauli principle, a single orbital can host the two electrons of H_2. Therefore, the *mean-field* many-body wavefunction is

$$\psi(\mathbf{r}_1, s_1, \mathbf{r}_2, s_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi(\mathbf{r}_1)\alpha(s_1) & \varphi(\mathbf{r}_2)\alpha(s_2) \\ \varphi(\mathbf{r}_1)\beta(s_1) & \varphi(\mathbf{r}_2)\beta(s_2) \end{vmatrix}$$

where $\alpha(s_i) = \begin{pmatrix} \delta_{i1} \\ \delta_{i2} \end{pmatrix}$ and $\beta(s_i) = \begin{pmatrix} \delta_{i2} \\ \delta_{i1} \end{pmatrix}$ are spinors. We can write the explicitly $\mathbf{R}_A$ and $\mathbf{R}_B$ for a diatomic molecule as

$$\mathbf{R}_A = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{R}_B = \begin{pmatrix} 0 \\ 0 \\ X \end{pmatrix} \implies |\mathbf{R}_A - \mathbf{R}_B| = X \tag{11}$$

The electronic Hamiltonian for the H_2 molecule reads

$$\hat{H}_e = -\frac{\nabla_1^2}{2} - \frac{\nabla_2^2}{2} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{1}{|\mathbf{R}_A - \mathbf{R}_B|} - \sum_{i=1}^{2} \sum_{I=A,B} \frac{1}{|\mathbf{r}_i - \mathbf{R}_I|} \tag{12}$$

The total wavefunction is a product of the same orbital φ orbital evaluated at two electron coordinates, consequently, the minimisation of the energy $\langle\psi|\hat{H}_e|\psi\rangle$ is equivalent to the minimisation of $\langle\varphi|\hat{\tilde{H}}_e|\varphi\rangle$, where $\hat{\tilde{H}}_e$ is a single electron Hamiltonian in which the dependence on the $\mathbf{r}_2$ degree of freedom has been averaged out:

$$\hat{\tilde{H}}_e = \langle\varphi(\mathbf{r}_2)|\hat{H}_e|\varphi(\mathbf{r}_2)\rangle \xrightarrow{\text{Removal of constant terms}} \hat{\tilde{H}}_e = -\frac{\nabla^2}{2} + \int d\mathbf{r}' \frac{|\varphi(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} - \sum_{I=A,B} \frac{1}{|\mathbf{r}-\mathbf{R}_I|}$$

where the removal of constant terms refers to all those terms in (12) which are constant with respect to $\mathbf{r}_1$. To avoid a heavy notation, henceforth we will refer to $\langle\varphi|\hat{\tilde{H}}_e|\varphi\rangle$ as $\langle\varphi|\hat{H}_e|\varphi\rangle$.

3.2 Matrix formalism

In all the *ab initio* methods we are going to discuss, the correct filling of matrices is fundamental and ubiquitous in the code. The variational principle imposes

$$\frac{\delta}{\delta\varphi^*}\left(\langle\varphi|\hat{H}_e|\varphi\rangle - \varepsilon \int d\mathbf{r}\ \varphi^*(\mathbf{r})\varphi(\mathbf{r})\right) = 0 \tag{13}$$

This problem can be turned into a generalised eigenvalue problem by replacing $\varphi(\mathbf{r})$ with the expansion in (10), multiplying by $\chi_p(\mathbf{r})$ and finally integrating over $d\mathbf{r}$

$$\sum_{pq}(\mathcal{H}_{pq} + \sum_{rs} C_r C_s \mathcal{Q}_{prqs}) C_q = \varepsilon \sum_{pq} \mathcal{S}_{pq} C_q \implies \mathcal{F}\mathbf{C} = \varepsilon \mathcal{S}\mathbf{C} \tag{14}$$

where the elements of the "one-body" matrix $\mathcal{H}$, the "two-body" tensor $\mathcal{Q}$ and the overlap matrix $\mathcal{S}$ are reported below:

$$\mathcal{H}_{pq} = \int d\mathbf{r}\ \chi_p(\mathbf{r}) \left(-\frac{\nabla^2}{2} - \sum_{I=A,B} \frac{1}{|\mathbf{r}-\mathbf{R}_I|} \right) \chi_q(\mathbf{r}) \qquad \mathcal{S}_{pq} = \int d\mathbf{r}\ \chi_p(\mathbf{r})\chi_q(\mathbf{r})$$

$$\mathcal{Q}_{pqrs} = \int d\mathbf{r}\ \chi_p(\mathbf{r})\chi_q(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \chi_r(\mathbf{r})\chi_s(\mathbf{r}')$$

The Fock matrix is then computed as

$$\mathcal{F}_{pq} = \mathcal{H}_{pq} + 2\sum_{rs} C_r C_s \left(\mathcal{Q}_{pqrs} - \frac{1}{2} Q_{prqs} \right) = \mathcal{H}_{pq} + \sum_{rs} C_r C_s \mathcal{Q}_{pqrs}$$

where the exchange term is exactly equal to half the Hartree (or direct) term. The energy is computed with

$$\begin{aligned} \langle \varphi | \mathcal{H}_e | \varphi \rangle &= 2\sum_{pq} C_p C_q \mathcal{H}_{pq} + \sum_{pqrs} C_p C_q C_r C_s \mathcal{Q}_{pqrs} + \frac{1}{X} = \\ &= \sum_{pq} C_p C_q (\mathcal{H}_{pq} + \mathcal{F}_{pq}) + \frac{1}{X} = \mathbf{C}^\top (\mathcal{H} + \mathcal{F}) \mathbf{C} + \frac{1}{X} \end{aligned} \tag{15}$$

The matrices have dimensions 8×8 and can be thought as composed of four submatrices, each one identified by the couple of nuclei on which the basis functions χ_p are centred:

$$\left| \begin{array}{c:c} AA & AB \\ \hdashline BA & BB \end{array} \right| \tag{16}$$

Here AA means that both the χ_p and χ_q are centred on the A nucleus, the other cases are similar. In the programs we will later describe, the Fock matrix is constantly updated at every step of nuclear dynamics due to the dependence of $\{\chi_p\}$ on the positions $\mathbf{R}_A$ and $\mathbf{R}_B$. The explicit analytical form of the matrix elements can be found in the appendix at 5.3.1.

3.3 Hybrid functionals

The single-determinant approximation of the Hartree-Fock wavefunction does not take into account Coulomb correlation. The Kohn-Sham theory still employs a single Slater determinant, but improves the energy using the *exchange and correlation* (XC) potential. Not only this represents an improvement in the estimate of the ground state energies, but also a better estimate of the forces driving the dynamics. The *local density approximation* (LDA), based on the assumption that the density can be treated locally as a free electron gas density, yields valuable estimates of the XC energy, which can be written as follows:

$$E_{xc}^{\mathrm{LDA}}[n] = \int d\mathbf{r}\ n(\mathbf{r}) \varepsilon_{xc}(n(\mathbf{r}))$$

If some part of the exchange energy is calculated from the HF theory (which provide an exact exchange), the functional is called *hybrid*. There exist many hybrid functionals known in literature (see the libxc website) involving both LDA and GGA (generalised gradient approximation) parts. In the present case of the H_2 molecule, the idea is to set up an hybrid functional in LDA that produces an estimate of the electronic ground state energy close to the most precise estimate available. From the literature[5], the most precise estimate comes from the full configuration interaction (full CI) calculation and is about -1.174360 a.u. at a distance of $X = 1.4$ a.u. . To try to replicate this energy, we devise the following hybrid functional

$$E_{xc}[n] = (1-a)E_x^{\mathrm{HF}}[n] + aE_x^{\mathrm{LDA}}[n] + E_c^{\mathrm{LDA}}[n] \qquad E_x^{\mathrm{LDA}}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int d\mathbf{r}\ [n(\mathbf{r})]^{4/3} \tag{17}$$

with a left as a parameter to be determined. The LDA functionals by Perdew-Zunger[6] and by Gunnarson-Lundqvist[7] have been tested in the program. We now clarify how this XC contribution has been implemented in the form of the

generalised eigenvalue problem (14). Based on the variational approach shown in (13), we compute the functional derivative of the XC energy with respect to the orbital φ^*. The chain rule imposes

$$\frac{\delta E_{xc}}{\delta \varphi^*(\mathbf{r})} = \int d\mathbf{r}' \, \frac{\delta E_{xc}}{\delta n(\mathbf{r}')} \frac{\delta n(\mathbf{r}')}{\delta \varphi^*(\mathbf{r})} = \frac{\delta E_{xc}}{\delta n(\mathbf{r})} \varphi(\mathbf{r}) = v_{xc}(\mathbf{r}) \varphi(\mathbf{r})$$

Now, proceeding as in the step between (13) and (14), we obtain a new generalised eigenvalue problem:

$$\sum_{pq} (\mathcal{H}_{pq} + \sum_{rs} C_r C_s \mathcal{Q}_{prqs} + \mathcal{V}^{xc}_{pq}) C_q = \varepsilon \sum_{pq} \mathcal{S}_{pq} C_q \qquad \mathcal{V}^{xc}_{pq} = \int d\mathbf{r} \; \chi_p(\mathbf{r}) v_{xc}(\mathbf{r}) \chi_q(\mathbf{r}) \tag{18}$$

The calculation of the integrals in (18) will be explained in the next section 3.4. Now, let us discuss the choice of the parameter a in (17). By direct comparison of the numerically obtained energetic profile representing $\langle \varphi | \hat{H}_e | \varphi \rangle$ in figure 3 with the full CI energy mentioned above, we conclude that the best choice of the parameter a is 0, meaning that the whole Hartree-Fock exchange energy has to be included. The reason of this is that the exchange contribution $E_x^{\mathrm{HF}}[n]$ in the case of H_2 is exactly equal to half the Hartree contribution $E_H[n]$. This does not happen in the homogeneous electron gas from which LDA is derived, consequently, $E_x^{\mathrm{LDA}}[n]$ does not fit the H_2 system properly.

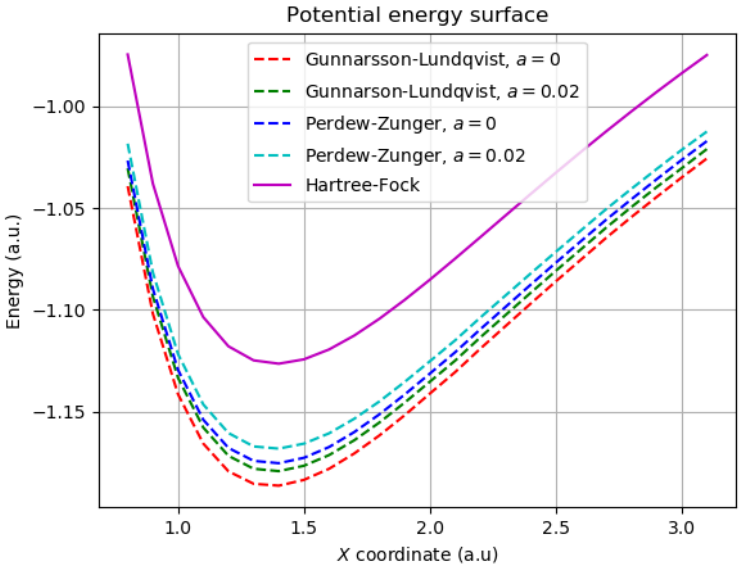


Figure 3. The Gunnarson-Lundqvist and Perdew-Zunger parametrisations of the exchange functional lead to an estimate of the ground state energy close to the full CI value of -1.174360 a.u. at $X = 1.4$ a.u.

From the differnce between the Hartree-Fock line and the other profiles computed with DFT in 3, we deduce that the correlation energy for the H_2 molecule represents about 4-5% of the total energy. This result is expected since the hydrogen molecule is known to be a highly correlated system.

3.4 Integrals of the exchange-correlation matrix

Contrarily to the other Fock-Kohn-Sham matrix elements, none of the integrals in $\mathcal{V}^{xc}$ are analytically known, thus they have to be performed numerically. The geometry of the hydrogen molecule with two centres at a distance X induces, in a natural way, the use of cylindrical coordinates $(x, y, z) \to (\rho \cos\theta, \rho \sin\theta, z)$ where z is the internuclear axis[3]. The integral over θ always gives a factor 2π since the density does not depend on the angle around the internuclear axis, but only on the distance ρ from it and on the z coordinate. Therefore, the advantage of cylindrical coordinates is that the integration is done in two dimensions (ρ and z) rather than in three (x, y, and z). Let us consider the integral on the right side of (18) and rewrite it using the notation $\mathbf{R} = (R^\rho, R^z, R^\theta)$. For simplicity, let us consider only the case of the

[3]We chose z as the internuclear axis instead of x to avoid a possible confusion deriving from the uppercase and lowercase formatting of x and X in the C++ code

north-east quadrant in (16), the other cases are analogous.

$$\mathcal{V}^{xc}_{pq} = 2\pi \int d\rho dz\ \rho v_{xc}(\rho,z) \exp\{-\alpha_p[(\rho - R^{\rho}_A)^2 + (z - R^z_A)^2]\} \exp\{-\alpha_q[(\rho - R^{\rho}_B)^2 + (z - R^z_B)^2]\} =$$
$$= 2\pi \exp\left\{-\frac{\alpha_p\alpha_q}{\alpha_p+\alpha_q}|\mathbf{R}_A - \mathbf{R}_B|^2\right\} \int d\rho dz\ \rho v_{xc}(\rho,z) \exp\{-(\alpha_p+\alpha_q)(\rho^2 + (z - R^z_C)^2)\} \qquad \mathbf{R}_C = \frac{\alpha_p\mathbf{R}_A + \alpha_q\mathbf{R}_B}{\alpha_p+\alpha_q}$$

Between the two steps, we replaced the product of two Gaussians with another Gaussian. We also got rid of the component of $\mathbf{R}_C$ along ρ: this simplification is allowed only in the case of diatomic molecules as the cylindrical coordinate transformation does not affect the $\mathbf{R}_A$ and $\mathbf{R}_B$ in (11).
The double integrations for the $\mathcal{V}^{xc}$ elements were performed using the Simpson method

$$\int_a^b dx\ f(x) \simeq \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] \tag{19}$$

in both integration variables ρ and z. Let us call N_ρ and N_z respectively the number of points in the ρ-grid and z-grid. The number of operations needed to get a single integral scales as $N_\rho \times N_z$. The number of different integrals in the form (18) is twenty[4], hence, filling the $\mathcal{V}^{xc}$ matrix requires $20 \times N_\rho \times N_z$ operations. It must be further considered that $\mathcal{V}^{xc}$ has to be filled several times in a single step of nuclear dynamics.
For this reason, to speed up the computation of integrals we decided to perform an *adaptive* Simpson integration. The method involves a sequential subdivision of the integration interval until the value of the integral is independent from the number of subdivisions. We report a version of this algorithm suggested by Lyness[8]: let us consider an interval $[a,b]$. We start by evaluating the integral using only the two extremes and the midpoint $(a+b)/2$ as in (19), then, we check the condition $|I(a,m) + I(m,b) - I(a,b)| < 15\varepsilon$ where I is the Simpson integral and m is the aforementioned midpoint. If the condition is not satisfied, the starting interval $[a,b]$ is split in two and the condition is checked over both the newly created intervals. At every iteration, the threshold ε is halved. The division is carried out until the condition $|I(a,m) + I(m,b) - I(a,b)| < 15\varepsilon$ is verified.

Algorithm 1 Adaptive integration

while $(|I(a,m) + I(b,m) - I(a,b)| < 15\varepsilon)$ **do**
  Redefine $[a,b] \to [a,m]$ or $[m,b]$
  Get $\int_a^m dx\ f(x)$ or $\int_m^b dx\ f(x)$
  Set $\varepsilon = \varepsilon/2$
end while

The above schematic reports the case of a one-dimensional integration. The extension to a bi-dimensional integration is straightforward.
In principle, the integration ranges are $(-\infty, +\infty)$ for z and $[0, +\infty)$ for ρ, however, the exponential character of the integrand functions of (18) yields numerically negligible results when evaluated far from the centre. For this reason, we introduce a cut-off distance depending on the range of the couple $\chi_p\chi_q$ in the integrand[5]. By direct inspection of the integrands and by controlling the convergence of the value of the integrals through adaptive integration, we concluded that the two integration intervals

$$\rho:\ [0.0,\ 2.0 + 1/(\alpha_p+\alpha_q)] \qquad z:\ [-1.5 - 1/(\alpha_p+\alpha_q),\ 2.5 + 1/(\alpha_p+\alpha_q)] \tag{20}$$

were sufficient to retrieve good estimates of all the 20 integrals mentioned above. Below, we show two of the integrands of (18) and the corresponding integration ranges on the axes.

[4] We have four Gaussians and hence 10 possible couples $\chi_p\chi_q$ without repetition, but we also have two type of integrals, namely, those with Gaussians centred on the same nucleus (AA or BB) and those centred on different nuclei (AB or BA).
[5] The presence of v_{xc} does not affect significantly the integration limits.

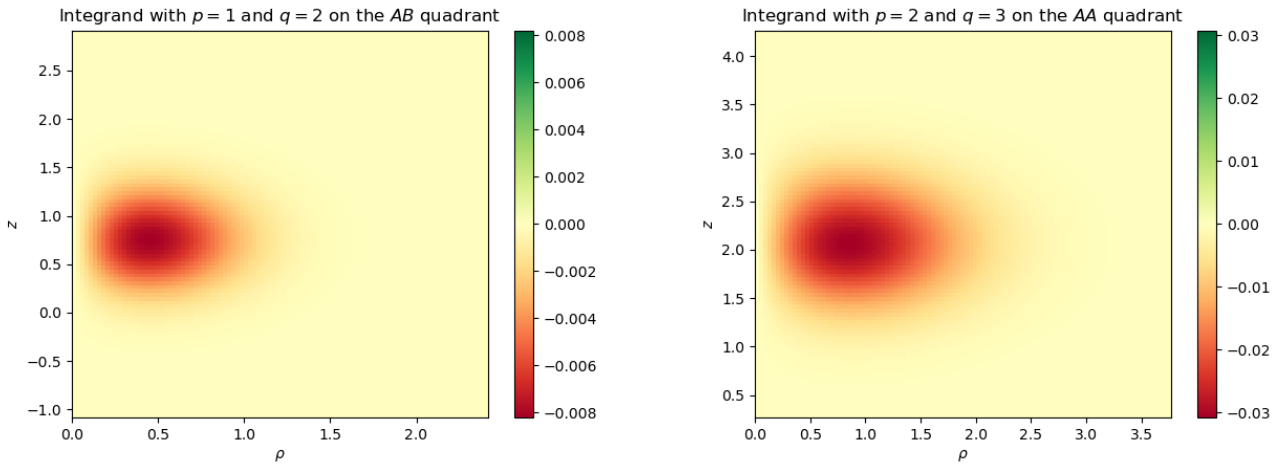


Figure 4. The ρ and z ranges correspond to the integration intervals used in the code. The integrands are plotted in a configuration with $X = 1.4$ a.u. as internuclear distance.

Although this method works for the specific problem, some observations must be pointed out.

- The integrands, i.e. the couple of Gaussians $\chi_p\chi_q$ multiplied by a complicated function of the density $v_{xc}(n)$ are "well behaved" and present no singularities, hence, adaptive integration is effective.
- Particular attention must be paid to the choice of the thresholds ε appearing in 1. The following set of ε constitutes a reasonable trade-off between speed and accuracy:

```
double eps[4][4] = {
    {1E-6, 1E-6, 1E-6, 1E-6},
    {1E-6, 1E-5, 1E-5, 1E-4},
    {1E-6, 1E-5, 1E-4, 1E-4},
    {1E-6, 1E-4, 1E-4, 1E-3}
};
```

  these thresholds are applied to each of the four submatrices in (16). The element `[0][0]` corresponds to the integral involving $\chi_0\chi_0$, namely, the fastest decaying couple of Gaussians. This couple leads to a small value of the integral (approximately 10^{-3}), so, to avoid large error in the estimate of the integral, we have to choose a threshold on a finer scale. The same line of reasoning has been applied also for the choice of the other thresholds.
- The Simpson's three-points rule can integrate exactly polynomials up to order three. In this type of quadrature, the points are uniformly spaced in the interval. The precision of the integration can be improved by letting the quadrature points and the weights unknown, for instance, with a Gauss quadrature. In this scheme, a polynomial of degree $2N-1$ can be integrated exactly using only N quadrature points.
- To avoid the cut-off approximation, we could have resort to a Gauss-Hermite (for the z integration) or Gauss-Laguerre (for the ρ integration) schemes, which are well suited for infinite integration ranges. These methods are present in the numerical integration library of gsl and we leave this as a future perspective for the improvement of the code.

3.5 Computation of the forces

The calculation of the forces is ubiquitous in the code as it governs the evolution of the dynamics of H_2. For a diatomic molecule in absence of external fields, the forces can only be directed along the internuclear axis[6] and must be computed as the derivative of the electronic energy $\langle\varphi|\hat{H}_e|\varphi\rangle$ with respect to the relative nuclear distance X. Whatever the type of dynamics (BOMD, CPMD, CGMD), the forces are computed in a Hellman-Feynman fashion:

$$F_X = -\frac{\partial\langle\varphi|\hat{H}_e|\varphi\rangle}{\partial X} = -\left\langle\varphi\left|\frac{\partial\mathcal{F}}{\partial X}\right|\varphi\right\rangle$$

[6]We assume that the molecule is not rotating, so that the nuclear motion is a vibration along the internuclear axis.

All the derivatives of the matrices $\mathcal{S}, \mathcal{H}$ and the tensor $\mathcal{Q}$ are analytically known and are reported in the section 5.3.2 of the appendix. The elements of $\mathcal{V}^{xc}$ represent the only exception: again, these elements have to be computed numerically. We insert the derivative in the integral:

$$\frac{\partial \mathcal{V}^{xc}_{pq}}{\partial X} = \int d\mathbf{r}\ \frac{\partial \chi_p(\mathbf{r})\chi_q(\mathbf{r})}{\partial X} v_{xc}(n) + \int d\mathbf{r}\ \chi_p(\mathbf{r})\chi_q(\mathbf{r}) \frac{\partial v_{xc}(n)}{\partial n(\mathbf{r})} \frac{\partial n(\mathbf{r})}{\partial X}$$

The derivative of v_{xc} with respect to the density is provided as an output by the Libxc library. As it was done for the elements of $\mathcal{V}^{xc}$, the integrals can again be performed in the cylindrical coordinates ρ and z.
Let us first focus on the first integral: the elements of the matrix $\partial \chi_p \chi_q / \partial X$ can be analytically computed:

$$\chi_p\chi_q = K_{pq}(X)\widetilde{\chi}_{pq}(X) \implies \frac{\partial \chi_p\chi_q}{\partial X} = \begin{cases} 0 \text{ in } AA \text{ quadrant} \\ \dfrac{\partial K_{pq}}{\partial X}\widetilde{\chi}_{pq} + K_{pq}\dfrac{\partial \widetilde{\chi}_{pq}}{\partial X} \text{ in } AB \text{ or } BA \text{ quadrant} \\ K_{pq}\dfrac{\partial \widetilde{\chi}_{pq}}{\partial X} \text{ in } BB \text{ quadrant} \end{cases}$$

where we defined conveniently

$$K_{pq}(X) = \exp\left\{-\frac{\alpha_p\alpha_q X^2}{\alpha_p+\alpha_q}\right\} \qquad \widetilde{\chi}_{pq}(X) = \exp\{-(\alpha_p+\alpha_q)[\rho^2 + (z-R_c^z)^2]\} \qquad R_C^z = \begin{cases} \dfrac{\alpha_p X}{\alpha_p+\alpha_q} \text{ in } AB \text{ quadrant} \\ \dfrac{\alpha_q X}{\alpha_p+\alpha_q} \text{ in } BA \text{ quadrant} \end{cases}$$

Although X lies on the same direction of the internuclear axis z, it must be pointed out that taking the derivative with respect to X is different from the taking the derivative with respect to z. In fact, the X dependence of $\chi_p\chi_q$ or n is just parametrical. Therefore, the integrand function is not guaranteed to have a symmetry axis as in the case of the integrands of $\mathcal{V}^{xc}$ in 4. To better explain this, we show the density, its derivative with respect to z and its derivative with respect to X. Finally, we also report one of the integrands of $\partial \mathcal{V}^{xc}/\partial X$. By direct inspection of the integrands, we chose the extremes of integration as in (20).

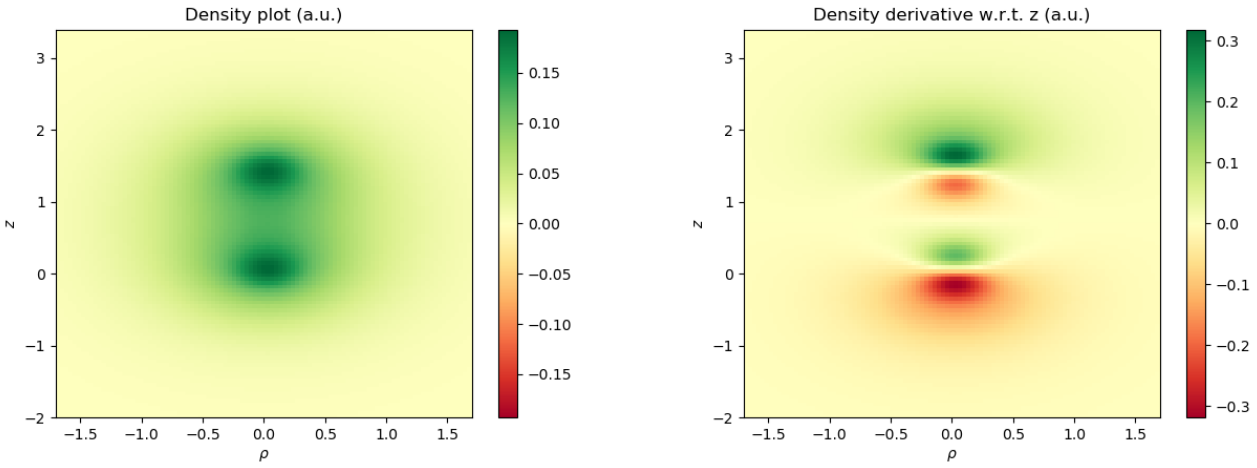


Figure 5. Density and its derivative along the z direction.

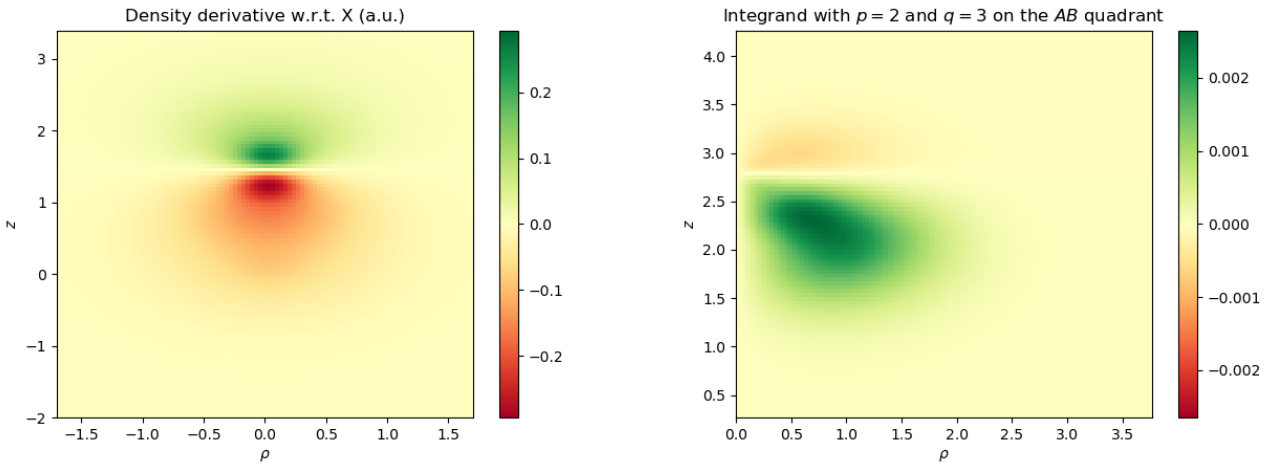
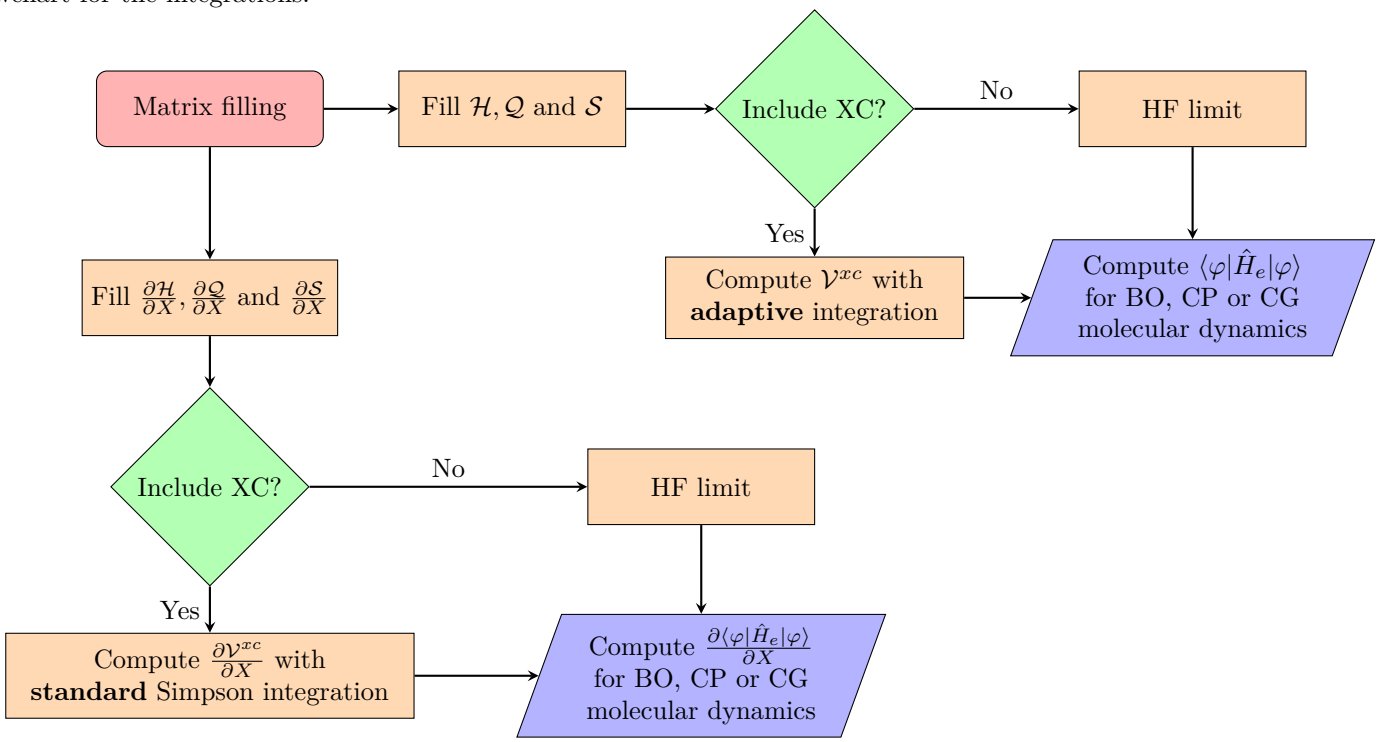


Figure 6. The symmetry of the z derivative is not present in the X derivative. This implies that there is no longer a symmetry axis in the integrand function reported in the last figure. The negative (unphysical) values of the distance ρ are reported just to show the symmetry of the density and its derivatives with respect to the $\rho = 0$ axis. As in 4, the plots are obtained at $X = 1.4$ a.u.

Remarkably, the scale of the values of the integrand function $\partial\mathcal{V}^{xc}/\partial X$ is considerably smaller than those of $\mathcal{V}^{xc}$ (about one order of magnitude as can be seen in the plots on the right of 4 and 6). Small values imply small integrals, which can eventually become comparable with the *a priori* thresholds ε of diagram 1. Consequently, the choice of the threshold of the adaptive integration is critical. A bad choice of the threshold could compromise the adaptive loop, which is not guaranteed to converge. For these reason, we decided to avoid the adaptive loop for the computation of the derivatives and perform a straightforward Simpson integration. This choice does not have major repercussions on the speed and efficiency of the code because, unlike the case of $\mathcal{V}^{xc}$, the calculation of the matrix $\partial\mathcal{V}^{xc}/\partial X$ occurs only once in every single step of nuclear dynamics. Before discussing the implementations of the three types of dynamics, we present a flowchart for the integrations.



3.6 Practical implementation of BOMD

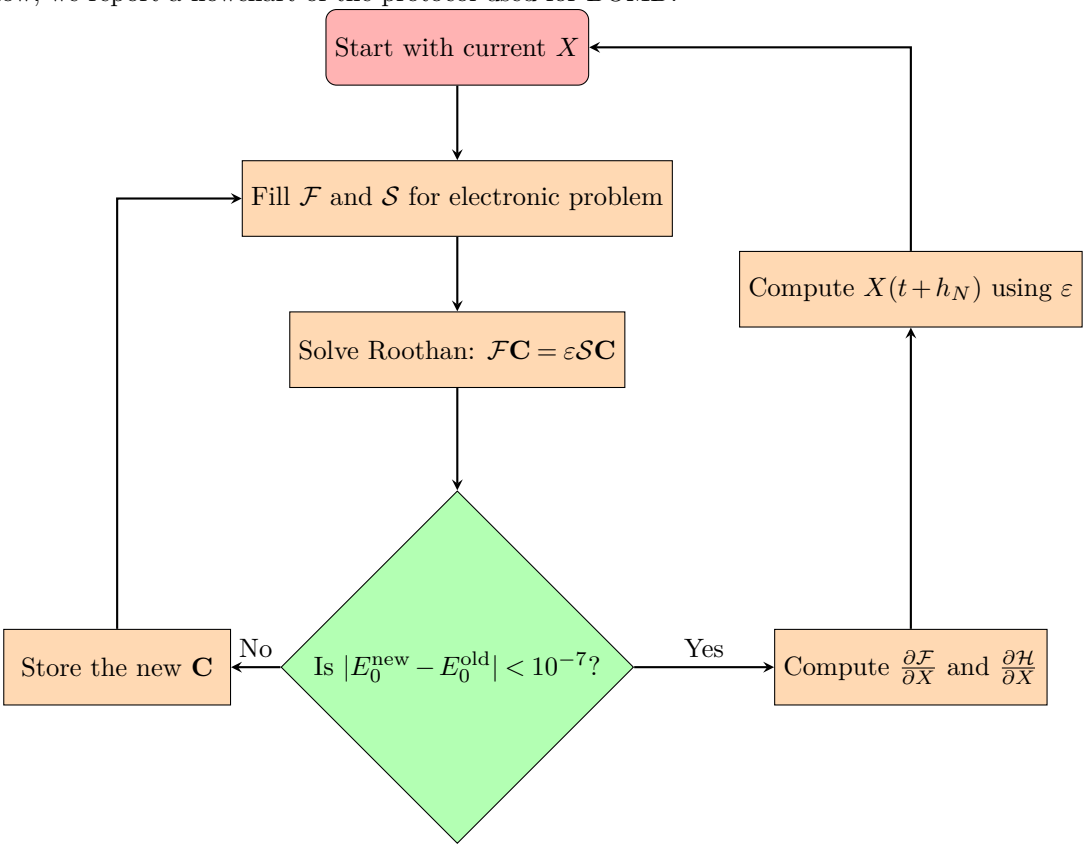
In this method we compute the gradient of the energy on the fly and use it in the propagation of a single time-step. A single time-step is relatively expensive since it requires the calculation of forces with *self-consistent* HF or DFT protocols. On the other hand, the electronic problem does not have an associated time scale, hence, the time step (which will be called h_N) can be quite large and it is usually of the order of femtoseconds. This choice still allows the observation of the oscillatory dynamics of the nuclei. In the case of H_2, the Lagrangian formalism leads to the following equation of motion:

$$\mathcal{L} = \frac{1}{2}\mu\dot{X}^2 - \langle\varphi|\hat{H}_e|\varphi\rangle + 2\varepsilon(\mathbf{C}^\top\mathcal{S}\mathbf{C} - 1) \Longrightarrow \begin{cases} \hat{H}_e\varphi = \varepsilon\varphi \\ \mu\ddot{X} = -\dfrac{\partial\langle\varphi|\hat{H}_e|\varphi\rangle}{\partial X} + 2\varepsilon\mathbf{C}^\top\dfrac{\partial\mathcal{S}}{\partial X}\mathbf{C} \end{cases} \qquad \mu = \frac{M_N}{2}$$

in which $\langle\varphi|\hat{H}_e|\varphi\rangle$ is computed with a Hartree-Fock or DFT self-consistent procedure. Note that the factor 2 in the constraint part of the Lagrangian is due to the double sum over the indices j and k in (5). The nuclear mass M_N expressed in atomic units is 1836.5 a.u. . The Hellman-Feynman force (see (7)) is used inside the Verlet evolution of X. The solution of the electronic equation $\hat{H}_e\varphi = \varepsilon\varphi$ can be recast into a generalised eigenvalue problem as in (14). This, however, cannot be solved with a single iteration as the Fock matrix depends on $\mathbf{C}$ itself. Therefore, the solution must be sought through subsequent iterations in which the newly computed eigenvector $\mathbf{C}$ is used to compute a new Fock matrix $\mathcal{F}$.

This procedure has been widely treated in the Computational Physics course, so we will not discuss the details.

Below, we report a flowchart of the protocol used for BOMD:



The computational cost of a BOMD strictly depends on the choice of the threshold for the difference $|E_0^{\text{new}} - E_0^{\text{old}}|$. That threshold determines the number of diagonalisations (which scale approximately as N^3, where N is the dimension of the square matrices) needed in one step of dynamics.

3.6.1 Equillibrium distance and different time-steps

One of the main observable is the "equilibrium" internuclear distance X. In order to provide a finite value, it is customary to introduce a damping coefficient γ_N in the nuclear motion, so that to "freeze" the dynamics and obtain a value of

X. This damped dynamics is clearly not physical for an isolated molecule, but it represents a way to circumvent the problem of performing averages over the periods of vibration.
Below, we report a comparison between the results obtained in the HF and DFT theories. The estimate of X produced are 1.38809 a.u. (0.7329 Å) in HF and 1.42012 a.u. (0.7498 Å) in DFT. On the right, three trajectories were compared with different time steps h_N. Comparable results are obtained as the time step does not enter in the self-consistent diagonalisation of the Fock-Kohn-Sham matrix, but only the Verlet integration.

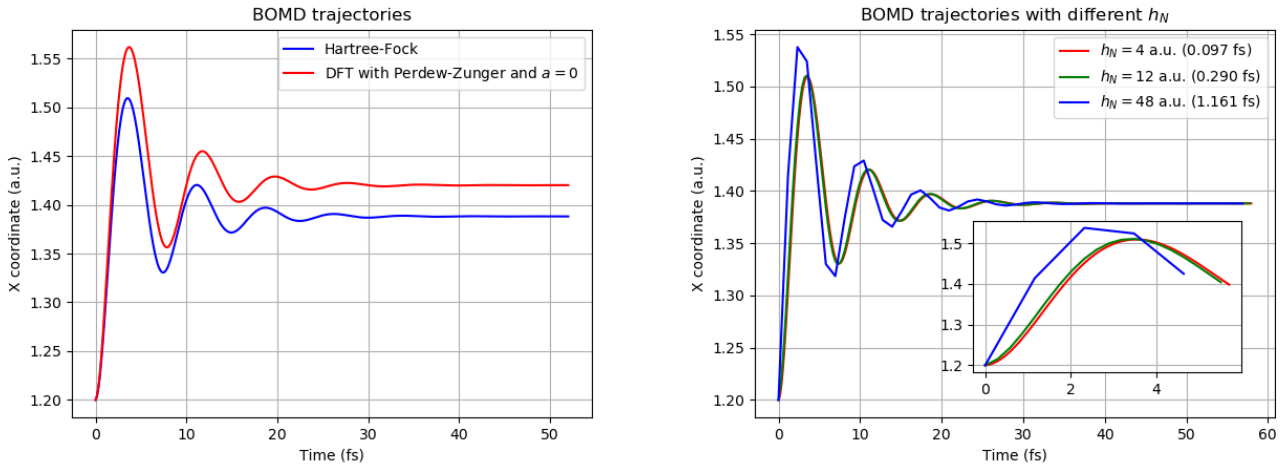


Figure 7. All the trajectories on the right were obtained with the same nuclear damping coefficient. On the left, the estimate produced in DFT is not clearly better than the estimate of HF since the experimental value is 0.7414 Å. This aspect will be discussed in the more detailed section 3.7.4. On the right, longer time steps do not produce a significant change in trajectory.

3.7 Practical implementation of CPMD

The Car-Parrinello Lagrangian in (8) written for the H_2 molecule reads

$$\mathcal{L} = \frac{1}{2}m\langle\dot{\varphi}|\dot{\varphi}\rangle + \frac{1}{2}\mu\dot{X}^2 - \langle\varphi|\hat{H}_e|\varphi\rangle + 2\lambda(1 - \mathbf{C}^\top\mathcal{S}\mathbf{C})$$

By replacing the orbital φ with its expansion of (10), we can derive the electronic equation of motion reported in (9) as an equation for the vector of coefficients $\mathbf{C}$. We also report the nuclear equation with the damping term already included.

$$\begin{aligned} m\ddot{\mathbf{C}}(t) &= -2(\mathcal{F} + \lambda\mathcal{S})\mathbf{C}(t) - \gamma\dot{\mathbf{C}}(t) \\ \mu\ddot{X} &= -\frac{\partial\langle\varphi|\hat{H}_e|\varphi\rangle}{\partial X} - 2\lambda\mathbf{C}^\top\frac{d\mathcal{S}}{dX}\mathbf{C} - \gamma_N\dot{X} \end{aligned} \tag{21}$$

Notice that an *electronic damping term* with γ has been included in the first equation of (21) with the aim of finding a stationary solution for $\mathbf{C}$. The meaning of this damping factor is completely different from the γ_N in nuclear motion and will be explained in the next subsection 3.7.1.
Some observations must be pointed out at this point:

- Contrarily to BOMD, in CPMD we need to introduce an *electronic time step* h, since the orbital φ is now representing an actual dynamical variable. In order to accomplish the adiabatic approximation, the evolution of $\mathbf{C}$ must occur on time-scales smaller than the nuclear ones. Therefore, we must choose $h \ll h_N$ and impose the dynamics of $\mathbf{C}$ in the first equation of (21) to occur between one nuclear step and another, namely, at fixed X. A practical way of doing this is to simply choose the number of electronic steps as h_N/h, so that the dynamics of $\mathbf{C}$ covers an interval of time exactly equal to h_N.
- The electronic damping does not lead to energy losses. In fact, the nuclear dynamics is only affected by the last value of $\mathbf{C}$ exiting the electronic cycle. In other words, the nuclear dynamics does not "see" what happens in time intervals smaller than h_N. When the X is updated, the electronic loop restarts with no memory of the preceding cycle.

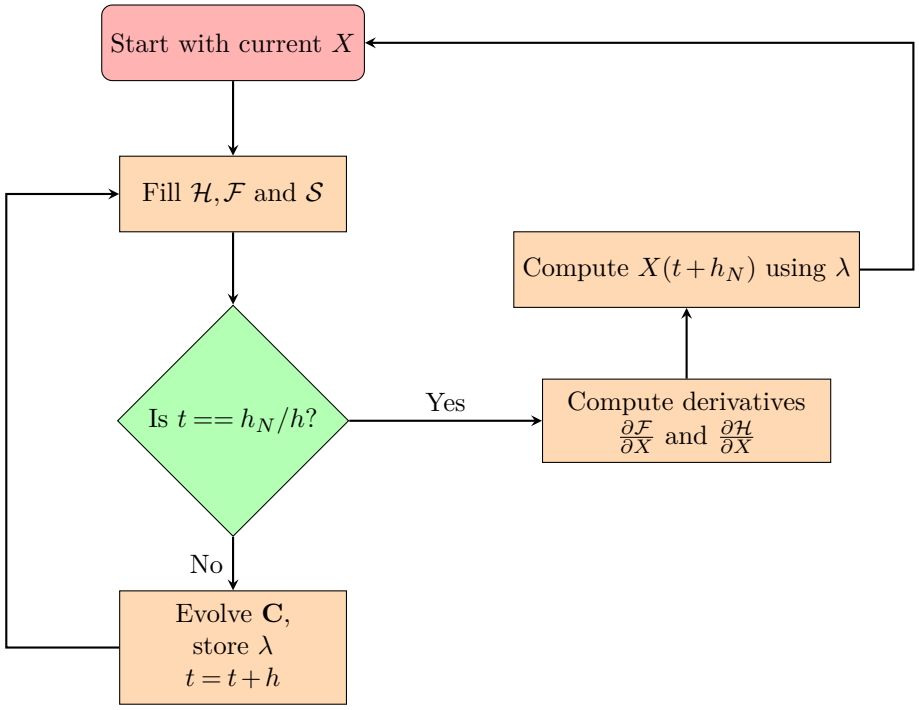
- The first equation of (21) contains the Lagrange multiplier λ. This multiplier guarantees that the normalisation $\langle\varphi|\varphi\rangle = 1$ is preserved along the trajectory. In many codes implementing Car-Parrinello dynamics, the problem of computing λ is solved with the iterative method SHAKE[3], which is usually employed in standard molecular dynamics programs. In the present code, we resort to another type of method requiring a partial update of $\mathbf{C}$ followed by a complete evolution.

Let us expand the last point of the list: we can first compute the partial result $\widetilde{\mathbf{C}}$ by neglecting the term $\lambda\mathcal{S}\mathbf{C}$ in (21). Then, subtracting the term $\lambda\mathcal{S}\mathbf{C}$ to $\widetilde{\mathbf{C}}$ leads to a second degree equation in λ. The lowest root of this equation is then taken as the Lagrange multiplier. Importantly, λ must be kept in memory since it is used in the nuclear step.

Algorithm 2 Evolution of $\mathbf{C}$

while $(t <= h_N)$ **do**
  Compute $\mathcal{F}$
  Find $\widetilde{\mathbf{C}}(t+h) = [2(m-h^2\mathcal{F})\mathbf{C}(t) - (m-h\gamma)\mathbf{C}(t-h)]/(m+h\gamma)$
  Define $\widetilde{\lambda} = \frac{2h^2\lambda}{m+\gamma h}$
  Solve $[\widetilde{\mathbf{C}}(t+h) - \widetilde{\lambda}\mathcal{S}\mathbf{C}(t)]^\top \mathcal{S}[\widetilde{\mathbf{C}}(t+h) - \widetilde{\lambda}\mathcal{S}\mathbf{C}(t)] = 1$ in $\widetilde{\lambda}$
  Set $\mathbf{C}(t+h) = \widetilde{\mathbf{C}}(t+h) - \widetilde{\lambda}\mathcal{S}\mathbf{C}(t)$
  Set $t = t+h$
end while

This algorithm for the update of $\mathbf{C}$ represents the most significant difference with the BOMD code. The remaining evolution of X is performed with the Verlet algorithm as in BOMD. As a summary, we present a flowchart of the code.



3.7.1 Error cancellation in Hellman-Feynman forces

In order to see the algorithm in 2 at work, we report in figure 8 the Hartree-Fock energy computed with (15) along the electronic cycle with different values of the electronic damping γ. The figure on the left has the purpose of showing the different oscillations provided in a single application of the algorithm in 2. We remind that this dynamics only concerns the evolution of $\mathbf{C}$ and not the distance X, which is kept fixed. Larger values of γ lead to a breakdown of the fluctuations of the values of $\mathbf{C}$, just like a larger classical damping coefficient leads to a more restrained motion. On the right of 8 we report the HF energy in a complete Car-Parrinello run. Based on this plot, we conclude that the choice of γ does not affect the energy and other observables here not shown.

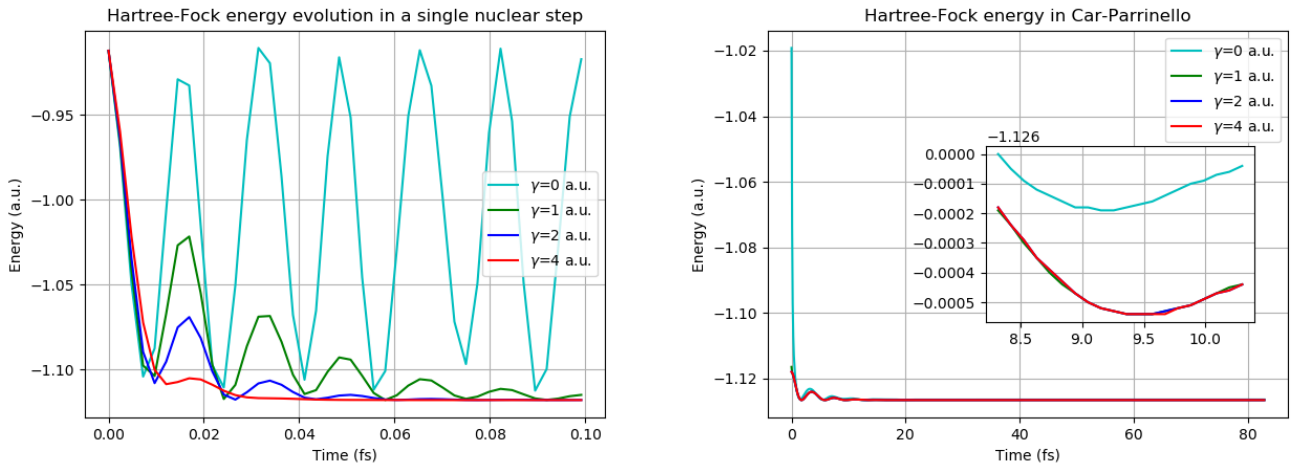


Figure 8. On the left, a temporal evolution of the HF energy in which $\mathbf{C}$ is updated using the algorithm in 2. The electronic step h is equal to 0.1 a.u., while the h_N is 4.3 a.u. (approximately 0.1 fs). This plot was taken at $X = 1.2$ a.u., i.e., the starting point of the trajectory represented on the right. On the right, the HF energies obtained along a whole CP trajectory. The differences within the four cases are minimal.

As highlighted on the right of 8, also the choice of $\gamma = 0$ produces valuable vectors $\mathbf{C}$ along the trajectory. This effect is known as *error cancellation* in Car-Parrinello simulations.
Let us explain the origin of this cancellation effect: when X is updated, the vector $\mathbf{C}$ is not instantly optimised, consequently, the algorithm 2 brings $\mathbf{C}$ towards the Born-Oppenheimer hypersurface[7], but then $\mathbf{C}$ starts oscillating around it according to the equation of motion in (21). Now, the forces governing the X evolution are computed with the last $\mathbf{C}$ exiting the cycle in 2. Every time $\mathbf{C}$ exits the cycle 2 along the whole trajectory, although it stays close to the Born-Oppenheimer hypersurface, it tends to have oscillations around it. Consequently, the Hellman-Feynman force at each nuclear step deviates from the optimal force of BOMD. This effect can be visualised in the diagrammatic figure 1 in section 2.4. Luckily, this oscillation around the Hellman-Feynman force cancels on average provided that it remains small at each time step, leading to a valuable trajectory.

3.7.2 Energy transfer

The possible *energy transfer* between the fictitious motion of the orbital and the classical motion of nuclear distances must always be kept under control in a Car-Parrinello simulation. Let us explain the concept of energy transferability: if the inertia of the orbital quantified by m is too large, it may happen that consecutive[8] oscillations are interrupted in a subspace of the $\mathbf{C}$-space that leads to a systematic error in the Hellman-Feynman force. Two strategies can be followed to fix this error:

- the fictitious mass m must be lowered. In fact, a higher mass implies less acceleration, so the response of the orbitals to a variation of X might be too slow to reach the variational minimum in time[9]. This effect can be seen on the left of figure 9
- increase γ, as an higher electronic damping limits the oscillations of $\mathbf{C}$ around the BO hypersurface. This can be seen in figure 9 on the right.

The trajectories with different m of the left plot in 9 are obtained with $\gamma = 1$ a.u., while the trajectories on the right are obtained with $m = 2$ a.u. .
On the left of 9, the CP forces oscillate around the BO force, however this does not produce an evident energy transfer along the whole trajectory. Thus, in a situation with $\gamma = 1$ a.u., the trajectory could be still considered reliable also with $m = 1000$ a.u. .
On the other hand, as can be seen on the left of 9, the $\gamma = 0$ a.u. line shows a small energy transfer visible from 20

[7] It is a surface of minimal variational energy in an eight-dimensional space.
[8] For instance, the oscillation of φ associated to $X(t)$ and the one of $X(t+h_N)$ are considered as consecutive oscillations.
[9] Remember that the Car-Parrinello trajectory moves in the $\mathbf{C}$-space according to (2) for a limited amount of time quantified by h_N at every nuclear step.

fs up to the end (in fact, the light blue line lies stably below the 0). As expected, when $\gamma \neq 0$ the oscillations of the orbital around the variational minimum are mitigated and consequently the CP forces are very close the ones of BO.

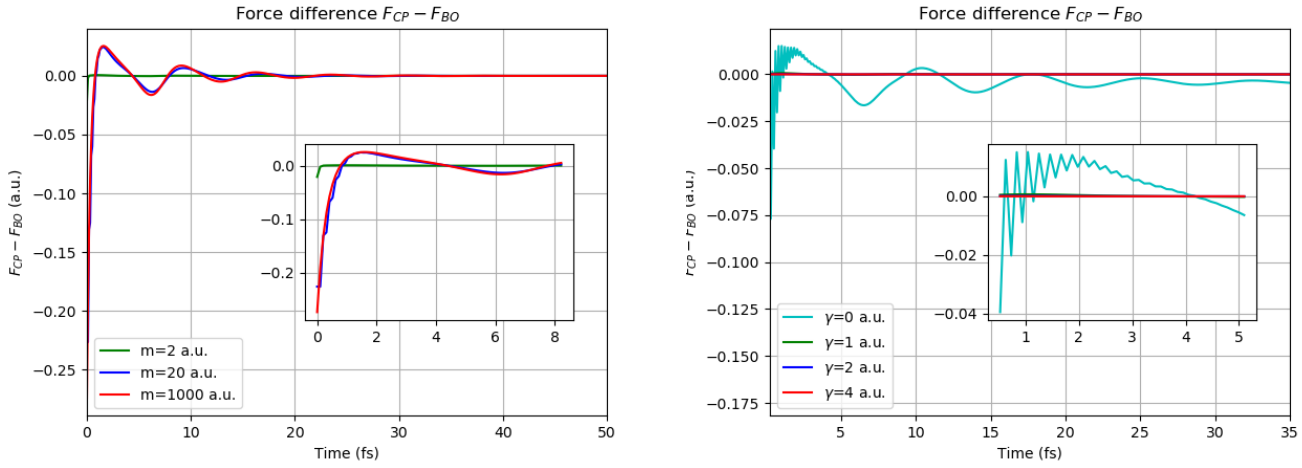


Figure 9. On the left, as can be seen in the zoom window, the trajectory with smaller fictitious mass ($m = 2$ a.u.) is practically the same in BOMD and CPMD. In the other cases ($m = 20$ and $m = 1000$ a.u.) the CP slightly deviates from the BO force. On the left, the effect of γ is to "cool down" the electrons in the orbital, so that $\mathbf{C}$ is close to its instantaneous minimum in the BO hypersurface.

In conclusion, the Car-Parrinello trajectory in the $\mathbf{C}$-space "oscillates" around the actual ground state. Although this implies the presence of some errors in the Hellman-Feynman forces, dynamical simulations can be carried out provided that those errors remain small at each time step.

3.7.3 Fictitious kinetic energy

The *fictitious* kinetic energy can be computed as the bra-ket product of the derivative[10] of φ:

$$\langle \dot{\varphi} | \dot{\varphi} \rangle = \sum_{p,q} \dot{C}_p \dot{C}_q S_{pq} + 2 \sum_{p,q} \dot{C}_p C_q \int d\mathbf{r}\ \chi_p \dot{\chi}_q + \sum_{p,q} C_p C_q \int d\mathbf{r}\ \dot{\chi}_p \dot{\chi}_q$$

The first of the three terms can be computed straightforwardly as a matrix-vector multiplication. Using the symmetry of $\mathcal{S}$ for the exchange of p and q, we can write the second term as:

$$2 \sum_{p,q} \dot{C}_p C_q \int d\mathbf{r}\ \chi_p \dot{\chi}_q = \sum_{p,q} \dot{C}_p C_q \left(\frac{dS}{dX} \right)_{pq} \dot{X}(t)$$

The integral in the third term requires a more subtle step: we can do the simplification reported below since the dependence on time is all contained within $X(t)$

$$\int d\mathbf{r}\ \dot{\chi}_p \dot{\chi}_q = |\dot{X}(t)|^2 \int d\mathbf{r}\ (\boldsymbol{\nabla} \chi_p)^\top \boldsymbol{\nabla} \chi_q = -|\dot{X}(t)|^2 \int d\mathbf{r}\ \chi_p \nabla^2 \chi_q$$

here we exploited the fact that the spatial derivative brings down the argument of the exponential in both χ_p exactly as a time-derivative would do.

The conserved energy associated to the Car-Parrinello Lagrangian in (8) is the sum of the nuclear kinetic energy (computed as $\mu \dot{X}^2$), the HF or DFT energy and the *fictitious* electronic kinetic energy. We report these quantities together in the next plot in 10.

The typical fluctuations of the nuclear kinetic energy and the HF energy are of the order of 10^{-2} a.u.. The conserved energy shows no global drift, but presents a small jump every period, meaning that the nuclear kinetic energy T_N and the fictitious energy T_f do no completely counterbalance the oscillations of the Hartree-Fock energy. This effect could be due to a poor estimate of the derivative $\dot{X}(t)$ (we used a simple two point formula).

[10] Also the basis functions χ_p depend on time through $X(t)$.

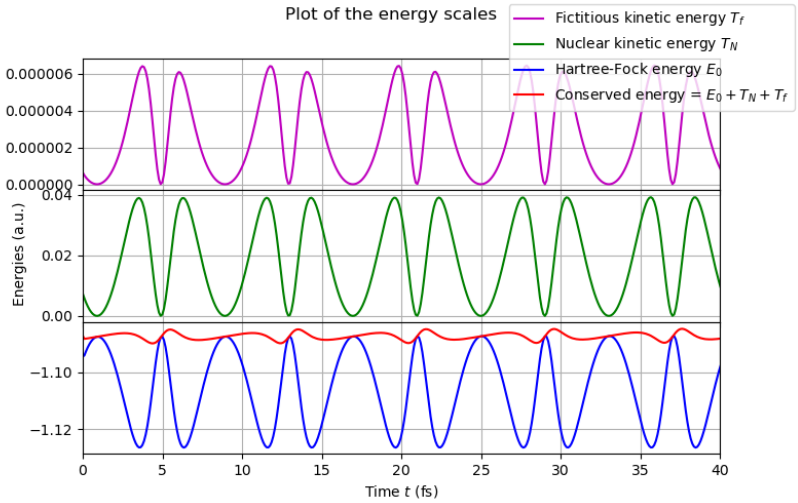


Figure 10. The conserved energy (or Hamiltonian) has relative fluctuations of the order 10^{-3} with no drift. These fluctuations are considerably smaller than the relative fluctuations of the electronic energy.

3.7.4 Equilibrium distance

We created an ensemble of trajectories with the same parameters $h, h_N, \gamma, \gamma_N, m$ but with different initial internuclear distance X and velocity. All the trajectories in the HF limit ended with an equilibrium value of 1.38809 a.u. and an energy of -1.12654 a.u. as shown in the figure below.

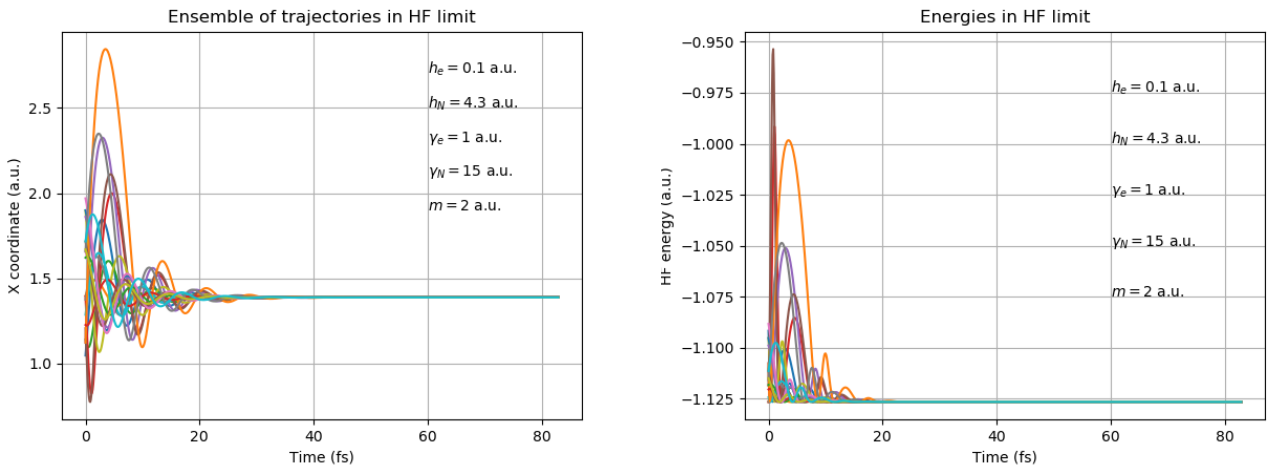


Figure 11. Many trajectories (with parameters reported in the figure) with different initial positions and velocities are reported. After a first section of equilibration, the relative distance X reaches a plateau at a value of 1.38809 a.u., which is equivalent to 0.733 Å. The experimental internuclear distance is 0.7414 Å. The ground state energy clearly differs considerably from the full CI result of 1.174 476 a.u.

Due to the high computational cost, we report single trajectories with DFT. The equilibrium distance for the Perdew-Zunger functional is 1.41948 a.u. (corresponding to 0.74948 Å) and 1.42407 a.u. with Gunnarson-Lundqvist (corresponding to 0.75190 Å). The estimate of the equilibrium distance can be adjusted with a better choice of the functional. Our choice was just an assumption based on the value of the full CI energy around 1.4 a.u. . This *a priori* assumption can be ameliorated with the inclusion of the exchange functional with a small a in the mixing equation in (17). The reason for this "failure" in estimating the equilibrium distance is that the XC functionals are not universal and even good functionals do not reproduce exactly all the observables. As we will point out in the next section 3.7.5, the functional adopted returns a valuable estimate of the vibrational frequency.

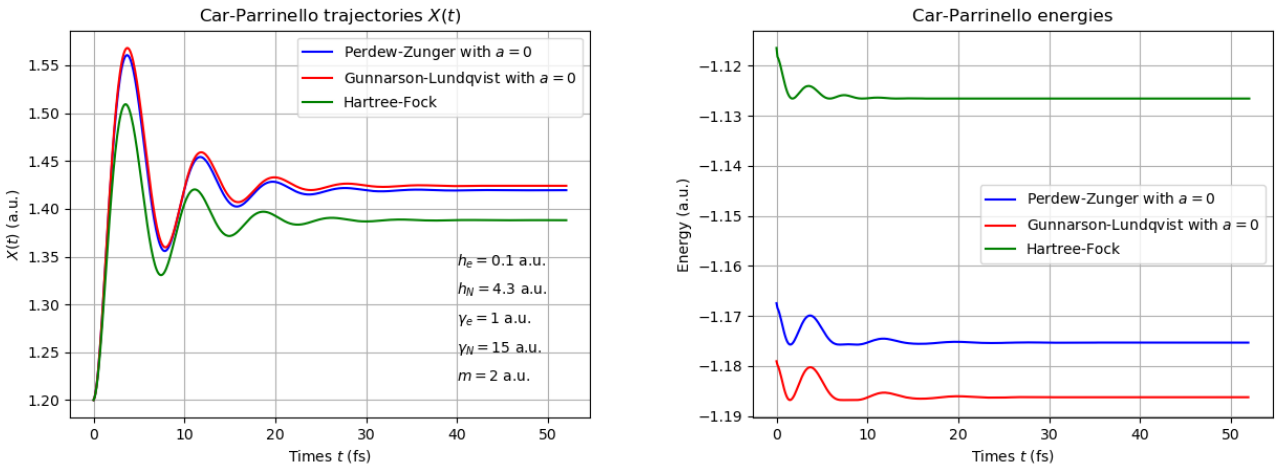


Figure 12. On the left, the "equilibrium" value is slightly different in the two cases with DFT and considerably different in the case of HF. The relative distance is about 2%. On the right, as expected from the assumption about the functional made in section 3.3, the plot of the Perdew-Zunger case shows an equilibrium value close to the full CI energy, i.e., -1.174360 a.u.

3.7.5 Vibrational frequency

We consider an HF and a DFT trajectory with $\gamma_N = 0$ and compute the typical frequency of the motion

$$\frac{d^2X(t)}{dt^2} = -\omega^2 X(t)$$

To this aim, we perform a discrete Fourier transform to capture the frequencies of the motion. The frequency of the main peak on the right of 13 is 13.2820×10^{13} Hz in the HF limit, while it gives 12.5346×10^{13} Hz in DFT. The experimental value for the hydrogen molecule is 12.48×10^{13} Hz.

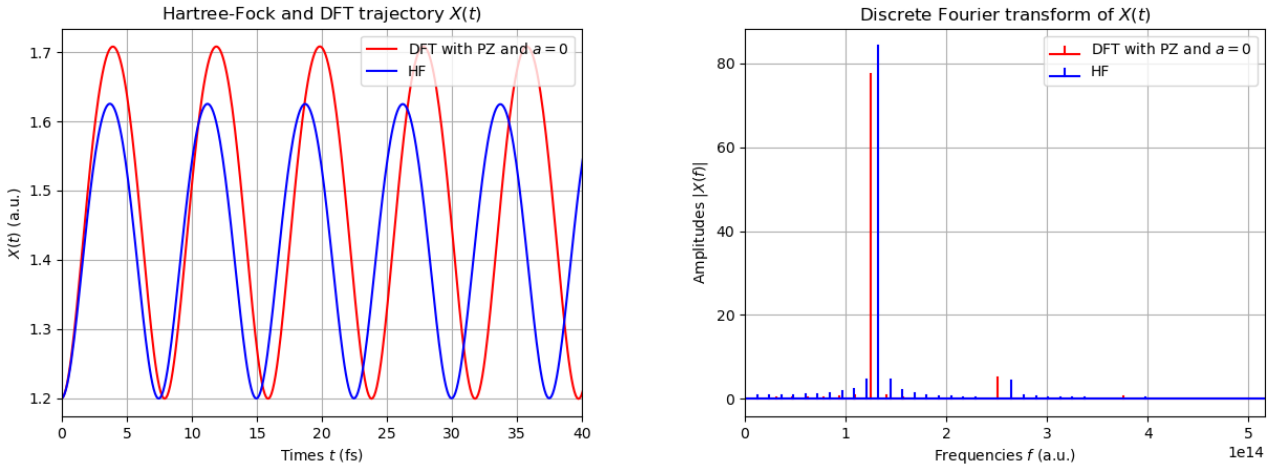


Figure 13. Using DFT with the Perdew-Zunger parametrisation, the estimate of the vibrational frequency is around 12.5346×10^{13} Hz, much closer to the experimental value.

3.8 Practical implementation of CGMD

The idea of applying this method in the context of *ab initio* calculations is treated in the review article by Payne *et al.*[9]. The general problem of an *ab initio* code is to locate the minimum of a function in a multidimensional space. Let $E[\mathbf{C}]$

be the energy $\langle\varphi|\hat{H}_e|\varphi\rangle$. The quantity we want to minimise is

$$F[\mathbf{C}] = E[\mathbf{C}] - 2\lambda\mathbf{C}^\top\mathbf{S}\mathbf{C} = 2\sum_{rs} C_r\mathcal{H}_{rs}C_s + \sum_{rstu} C_rC_sC_tC_t\mathcal{Q}_{rstu} + \frac{1}{X} - 2\lambda\sum_{rs}\mathcal{S}_{rs}C_rC_s$$

and its gradient with respect to $\mathbf{C}^*$ in the eight-dimensional $\mathbf{C}$-space is computed as

$$\begin{aligned}(\boldsymbol{\nabla}F[\mathbf{C}])_r &= 2\sum_{s}\mathcal{H}_{rs}C_s + 2\sum_{stu} C_sC_tC_u\mathcal{Q}_{rstu} - 2\lambda\sum_{s}\mathcal{S}_{rs}C_s \\ &= 2\mathcal{F}\mathbf{C} - 2\lambda\mathcal{S}\mathbf{C}\end{aligned}$$

where λ is a Lagrange multiplier that maintains the normalisation condition $\langle\varphi|\varphi\rangle = 1$. The function $F[\mathbf{C}]$ is a quartic in $\mathbf{C}$, however, it can be well approximated with a quadratic form close to the minimum (assuming that $\boldsymbol{\Delta}\mathbf{C}$ is sufficiently small).

$$\begin{aligned}F[\mathbf{C}+\Delta\mathbf{C}] &\simeq E[\mathbf{C}] - 2\lambda\mathbf{C}^\top\mathcal{S}\mathbf{C} + \Delta\mathbf{C}^T(2\mathcal{F}\mathbf{C} - 2\lambda\mathcal{S}\mathbf{C}) + \frac{1}{2}\Delta\mathbf{C}^T(2\mathcal{F} + 2(\boldsymbol{\nabla}\mathcal{F})\mathbf{C} - 2\lambda\mathcal{S})\Delta\mathbf{C} = \\ &= F[\mathbf{C}] - \Delta\mathbf{C}^T\mathbf{b} + \frac{1}{2}\Delta\mathbf{C}^T\mathbf{H}[\mathbf{C}]\Delta\mathbf{C}\end{aligned}$$

The minimisation problem is solved once we obtain $\boldsymbol{\nabla}F[\mathbf{C}] = 0$. This problem can be recast as a linear system $\mathbf{H}\Delta\mathbf{C} = \mathbf{b}$. In the same way as the self-consistent cycle and the algorithm in 2 were the "engines" of BOMD and CPMD, the iterative solution of $\mathbf{H}\Delta\mathbf{C} = \mathbf{b}$ drives the CGMD. The variational principle request is equivalent to imposing that the directional derivative of the function F is zero. This gives us the value of λ.

$$\frac{\delta}{\delta\varphi^*}\left(\langle\varphi|\hat{H}_e|\varphi\rangle - 2\lambda\int d\mathbf{r}\ \varphi^*(\mathbf{r})\varphi(\mathbf{r})\right) = 0 \implies \mathbf{C}^\top\boldsymbol{\nabla}F = 0 \Longleftrightarrow \lambda = \mathbf{C}^\top\mathcal{F}\mathbf{C}$$

3.8.1 Explanation of the algorithm

Henceforth we use a convenient notation indicating with $\{\mathbf{d}_k\}$ the conjugate directions and with $\{\mathbf{r}_k\}$ the "remainders" $\mathbf{b} - \mathbf{H}\Delta\mathbf{C}_k$. We remind that all these quantities are eight-dimensional vectors in the $\mathbf{C}$-space. Let us call $\Delta\mathbf{C}_0$ the first increment. We can identify the first direction wih $\mathbf{d}_0 = -\boldsymbol{\nabla}F$, perform a line minimisation along that direction and find $\Delta\mathbf{C}_1$. In order to find that new point we impose that the directional derivative at that point is zero:

$$\mathbf{d}_0^\top\boldsymbol{\nabla}F[\Delta\mathbf{C}_0 + \alpha_1\mathbf{d}_0] = 0 \Longrightarrow \mathbf{d}_0^\top(-\mathbf{b} + \mathbf{H}\Delta\mathbf{C}_0 + \alpha_1\mathbf{H}\mathbf{d}_0) = 0 \Longrightarrow \alpha_1 = \frac{\mathbf{d}_0^\top\mathbf{r}_0}{\mathbf{d}_0^\top\mathbf{H}\mathbf{d}_0}$$

Let us call $\Delta\mathbf{C}_1 = \Delta\mathbf{C}_0 + \alpha_1\mathbf{d}_0$ the newly computed point in the $\mathbf{C}$-space. Now we want to find the next point $\Delta\mathbf{C}_2 = \Delta\mathbf{C}_1 + \alpha_2\mathbf{d}_1$ by moving in a new direction $\mathbf{d}_1$ in which the gradient along $\mathbf{d}_0$ remains zero. If further, we impose to move in the opposite direction of the gradient of F as before, the condition to satisfy is $\mathbf{d}_0^\top\boldsymbol{\nabla}F[\Delta\mathbf{C}_1 + \alpha_2\mathbf{d}_1] = 0$. It can be shown that the requirement for the direction $\mathbf{d}_1$ is $\mathbf{d}_0^\top\mathbf{H}\mathbf{d}_1 = 0$. Below, a brief proof of this statement:

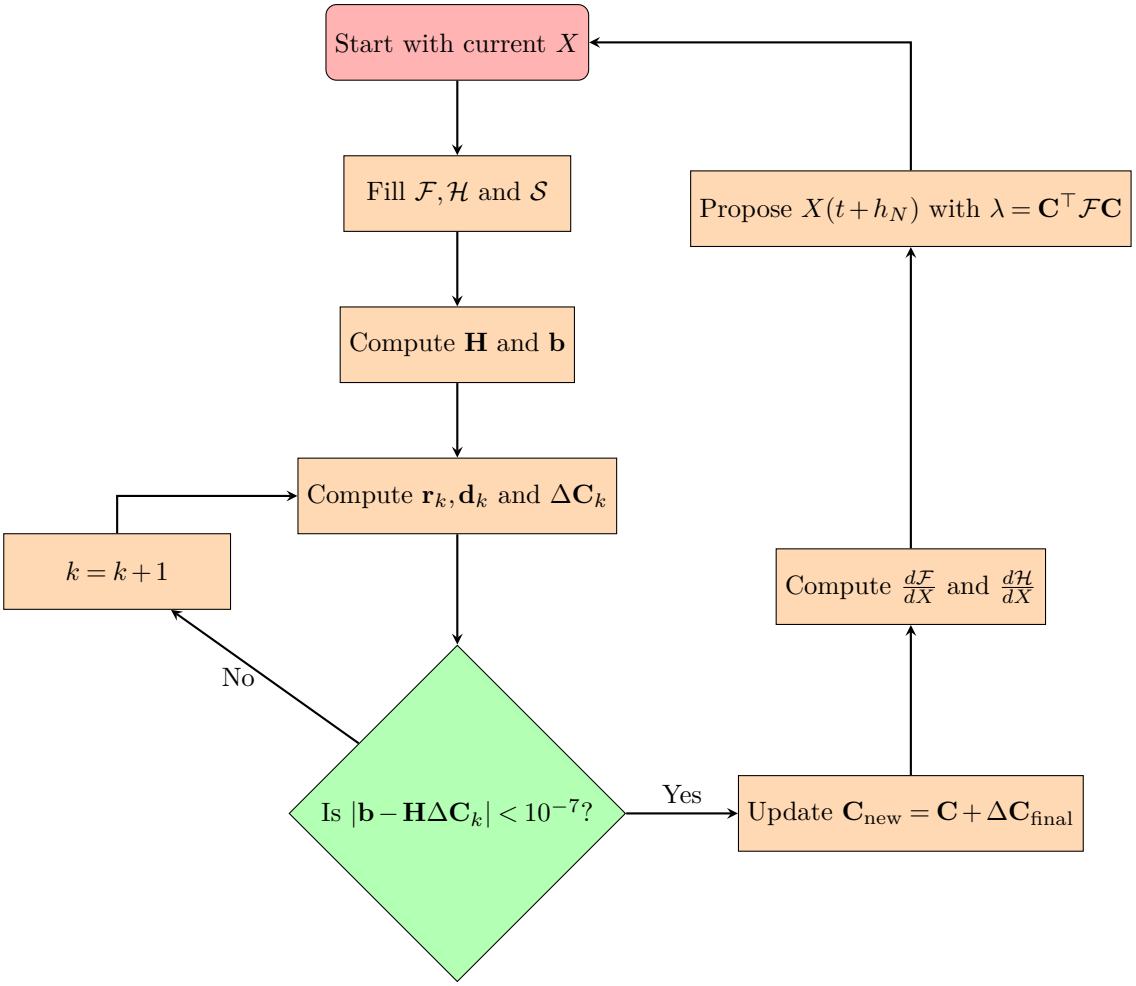
$$\begin{aligned}\mathbf{d}_0^\top\boldsymbol{\nabla}F[\Delta\mathbf{C}_1 + \alpha_2\mathbf{d}_1] &= \mathbf{d}_0^\top(-\mathbf{b} + \mathbf{H}\Delta\mathbf{C}_1 + \alpha_2\mathbf{H}\mathbf{d}_1) = \\ &= \mathbf{d}_0^\top\boldsymbol{\nabla}F(\Delta\mathbf{C}_1) + \alpha_2\mathbf{d}_0^\top\mathbf{H}\mathbf{d}_1 \underbrace{=}_{\text{if } \mathbf{d}_0^\top\mathbf{H}\mathbf{d}_1 = 0} \mathbf{d}_0^\top\boldsymbol{\nabla}F(\Delta\mathbf{C}_0 + \alpha_1\mathbf{d}_0) = 0\end{aligned}$$

The third direction, namely $\mathbf{d}_2$, will be designed such that it does not spoil the gradient along $\mathbf{d}_1$ and so on. In general, $\mathbf{d}_i^\top\mathbf{H}\mathbf{d}_j = 0$ must be satisfied for every $i \neq j$, and if such a condition is satisfied, the directions are said to be *conjugate*. It can be shown that the algorithm in 3 produces a sequence $\{\mathbf{d}_i\}$ of conjugate directions. Explicit demonstrations of the steps are reported in the dedicated section of the appendix 5.2.

Algorithm 3 Conjugate Gradient

while $(\mathbf{r}_{k-1}^\top\mathbf{r}_{k-1} > \varepsilon = 10^{-7})$ **do**
  $\alpha_k = \mathbf{r}_{k-1}^\top\mathbf{r}_{k-1}/\mathbf{d}_{k-1}^\top\mathbf{H}\mathbf{d}_{k-1}$
  $\Delta\mathbf{C}_k = \Delta\mathbf{C}_{k-1} + \alpha_k\mathbf{d}_{k-1}$
  $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k\mathbf{H}\mathbf{d}_{k-1}$
  $\beta_k = \mathbf{r}_k^\top\mathbf{r}_k/\mathbf{r}_{k-1}^\top\mathbf{r}_{k-1}$
  $\mathbf{d}_k = \mathbf{r}_k + \beta_k\mathbf{d}_{k-1}$
  $k = k+1$
end while

The final increment $\Delta\mathbf{C}_{\text{final}}$ exiting the algorithm in 3 must be added to the current $\mathbf{C}$ vector. We now present a flowchart of the CGMD.



Every time the above cycle is repeated, the Fock-Kohn-Sham matrix is filled using the newly computed $\mathbf{C}_{\text{new}}$. This is the key to achieve good performance of CGMD, as the conjugate gradient loop in the next cycle starts off from a configuration already close to the minimum[11].
In the next figure 14, we show a comparison between the three methods BOMD, CPMD and CGMD described in this report, together with the force difference with respect to the Born-Oppenheimer ones. On the left of 14, the trajectories are very similar, especially the BOMD and CGMD ones, as can be seen in the zoom. This fact is confirmed by the force difference plot. Thus, we conclude that the conjugate gradient technique is able to efficiently drive the dynamics through the $\mathbf{C}$-space.

[11] For this reason, we included in the code a little equilibration for $\mathbf{C}$ and then started the first conjugate gradient loop

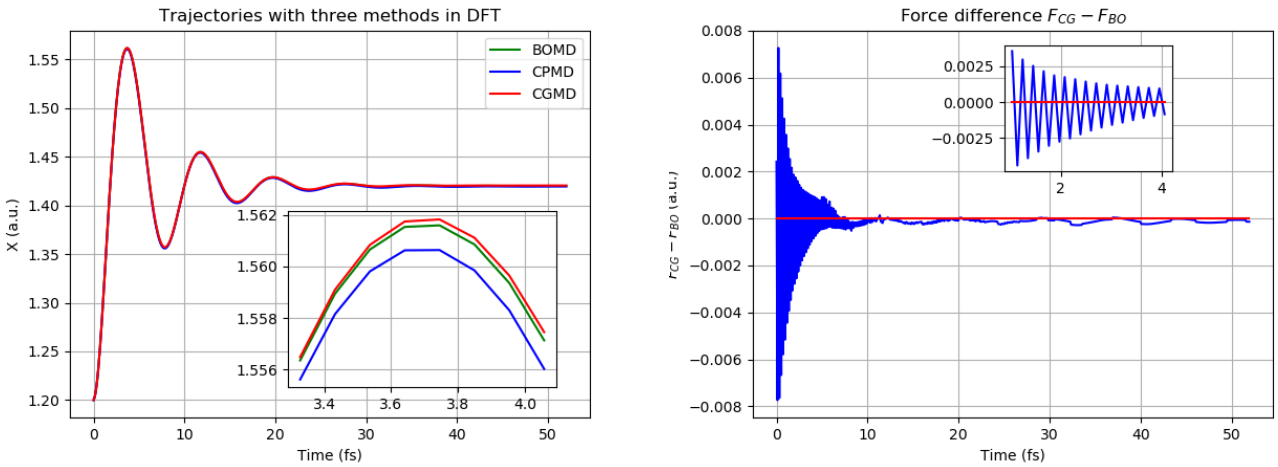


Figure 14. The functional used in the plot on the left is again Perdew-Zunger with $a = 0$. The fluctuations in the plot on the left do not present significant drift.

4 CONCLUSIONS

Among the presented methods, the algorithm by Car and Parrinello is certainly the most used in *ab initio* calculations. Its success derives principally from the application to more extended or periodic systems described in terms of plane-waves and it is not evident in the examined case of the isolated H_2 molecule. In this section we point out some qualitative and quantitative differences with the Born-Oppenheimer and conjugate gradient approaches:

- BOMD and CGMD do not need parameters chosen a priori, such as the fictitious mass m or the damping coefficient γ. Despite this, we have shown that the CPMD solution remains close to the other two, also with significant variations of the parameters.
- Unlike BOMD and CGMD in which the electronic problem is solved with the self-consistent Roothan equation $\mathcal{F}\mathbf{C} = \varepsilon\mathcal{S}\mathbf{C}$ and the linear system $\mathbf{H}\Delta\mathbf{C} = \mathbf{b}$, CPMD does not present any iterative process. This represents a great achievement in the treatment of more sophisticated molecules or extended systems for which an iterative solution would require a large amount of time to converge.

Lastly, a more quantitative comparison could be made by "projecting" the motion of BOMD and CGMD over the motion of Car-Parrinello. This can be done by computing the scalar product $\mathbf{C}^T\mathbf{S}\mathbf{C}$ as in figure 15.

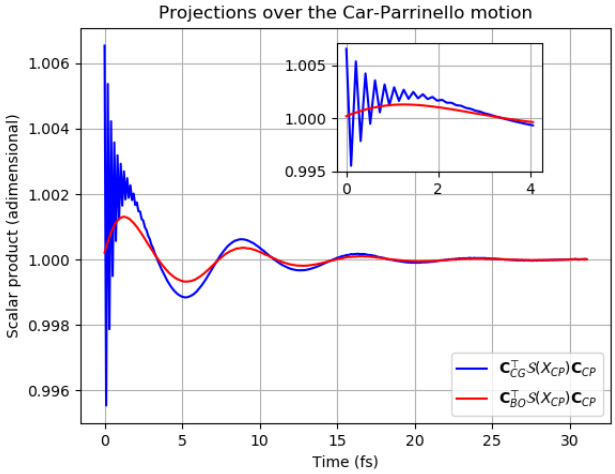


Figure 15. Projections of BOMD and CGMD over the motion of Car-Parrinello. The conjugate gradient projections show significant fluctuations, highlighting the different dynamics carried on the **C**-space.

5 APPENDIX

5.1 Hamilton-Jacobi formulation

In classical mechanics the action functional is defined as the integral

$$S[\mathbf{q}(t)] = \int_{t_1}^{t_2} dt\ \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$$

where $\mathbf{q}$ is a vector of $3N$ generalised coordinates. We can compute the variation δS:

$$\delta S = \sum_i^{3N} \int_{t_1}^{t_2} dt\ \left[\frac{\partial \mathcal{L}}{\partial q_i} \delta q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d(\delta q_i)}{dt} \right] \Longrightarrow \delta S = \sum_i^{3N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i \Big|_{t_1}^{t_2}$$

where the last equality follows from an integration by parts and from the Euler-Lagrange equations. If we consider a path with fixed endpoints in the space of the generalised coordinates, we get $\delta S = 0$ as expected, but if we slightly move the endpoint (and again compute the new action on the new path), we can define a derivative with respect to the endpoint: $\boldsymbol{\nabla}_{\mathbf{q}} S = \boldsymbol{\nabla}_{\dot{\mathbf{q}}} \mathcal{L}$. In an Hamiltonian description the vector $\boldsymbol{\nabla}_{\dot{\mathbf{q}}} \mathcal{L}$ is labelled with $\mathbf{p}$. We now remind that the Lagrangian is the total time derivative of the action, hence

$$\mathcal{L} = \frac{dS}{dt} = \sum_i^{3N} \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t} = \sum_i^{3N} p_i \dot{q}_i + \frac{\partial S}{\partial t} \implies \underbrace{\frac{\partial S(\mathbf{q}, t)}{\partial t} + H(\mathbf{q}, \boldsymbol{\nabla}_{\mathbf{q}} S, t) = 0}_{\text{Hamilton-Jacobi equation}} \tag{22}$$

the last equality follows from the Legendre transform involving $\mathcal{L}$ and H. A striking feature of the Hamilton-Jacobi formalism is that the classical limit ($\hbar \to 0$) of the Schrödinger equation is isomorphic to the equation in (22). To show this, let us consider a particle in a generic one-dimensional potential $V(x,t)$:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x,t)\psi(x,t)$$

The ansatz $\psi = \exp(iS(x,t)/\hbar)$ leads to an equation for $S(x,t)$ which, in the limit $\hbar \to 0$ reads

$$\frac{\partial S(x,t)}{\partial t} + \left(\frac{\partial S(x,t)}{\partial x} \right)^2 + V(x,t) = 0$$

5.2 Conjugate gradient

The first two steps reported in the diagram 3, namely

$$\alpha_k = \mathbf{r}_{k-1}^\top \mathbf{r}_{k-1} / \mathbf{d}_{k-1}^\top \mathbf{H} \mathbf{d}_{k-1}$$
$$\Delta \mathbf{C}_k = \Delta \mathbf{C}_{k-1} + \alpha_k \mathbf{d}_{k-1}$$

constitute a line minimisation along the direction $\mathbf{d}_{k-1}$. The third step: $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{H} \mathbf{d}_{k-1}$ is obtained by simply replacing the expression of the residual $\mathbf{r}_k$:

$$\begin{aligned} \mathbf{r}_k &= \mathbf{b} - \mathbf{H}\Delta \mathbf{C}_k = \\ &= \mathbf{b} - \mathbf{H}(\Delta \mathbf{C}_{k-1} + \alpha_k \mathbf{d}_{k-1}) = \\ &= \mathbf{r}_{k-1} - \alpha_k \mathbf{H} \mathbf{d}_{k-1} \end{aligned}$$

The fourth and the fifth steps encode the search of the *conjugate* direction. We remind that the condition to be satisfied is $\mathbf{d}_{k-1}^\top \mathbf{H} \mathbf{d}_k = 0$. By direct substitution of the fourth and fifth steps of 3, we obtain

$$\mathbf{d}_{k-1}^\top \mathbf{H} \left(\mathbf{r}_k^\top + \frac{\mathbf{r}_k^\top \mathbf{r}_k}{\mathbf{r}_{k-1}^\top \mathbf{r}_{k-1}} \mathbf{d}_{k-1}^\top \right) = 0 \Longrightarrow \mathbf{d}_{k-1}^\top \mathbf{H} \mathbf{r}_k = -\mathbf{d}_{k-1}^\top \mathbf{H} \mathbf{d}_{k-1} \frac{\mathbf{r}_k^\top \mathbf{r}_k}{\mathbf{r}_{k-1}^\top \mathbf{r}_{k-1}} \tag{23}$$

To demonstrate the equality above, we can compute the scalar product between one conjugate direction $\mathbf{d}_j$ and $\mathbf{r}_k - \mathbf{r}_{k-1} = -\alpha_k \mathbf{H} \mathbf{d}_{k-1}$. Due to the conjugation property $\mathbf{d}_j^\top \mathbf{H} \mathbf{d}_k = \delta_{jk}$, it follows that $\mathbf{d}_j^\top (\mathbf{r}_k - \mathbf{r}_{k-1}) = -\alpha_k \delta_{jk}$, but the only possibility is that $\mathbf{d}_j^\top \mathbf{r}_k = 0$ and $\mathbf{d}_j^\top \mathbf{r}_{k-1} = 0$ for every $j \neq k$. It can be shown that the space spanned by the

conjugate directions span$\{\mathbf{d}_0 ... \mathbf{d}_k\}$ is the same space spanned by the residuals span$\{\mathbf{r}_0 ... \mathbf{r}_k\}$, therefore the condition $\mathbf{d}_j^\top \mathbf{r}_{k-1} = 0$ implies also $\mathbf{r}_k^\top \mathbf{r}_{k-1} = 0$. Now, taking the expression for $\mathbf{r}_k$ presented in 3, one can directly derive the two equalities

$$\begin{aligned} &\mathbf{r}_k^\top \mathbf{r}_k = -\alpha_k \mathbf{d}_{k-1}^\top \mathbf{H} \mathbf{r}_k \\ &\mathbf{r}_{k-1}^\top \mathbf{r}_{k-1} = \alpha_k \mathbf{r}_{k-1}^\top \mathbf{H} \mathbf{d}_{k-1} = \alpha_k \mathbf{d}_{k-1}^\top \mathbf{H} \mathbf{d}_{k-1} \end{aligned} \tag{24}$$

The last equality follows from the fifth line in 3, namely, $\mathbf{d}_{k-1} = \mathbf{r}_{k-1} + \beta_{k-1}\mathbf{d}_{k-2}$ and the conjugation rule $\mathbf{d}_{k-2}^\top \mathbf{H} \mathbf{d}_{k-1} = 0$. The expressions in (24) can now be easily substituted in (23) to prove the equality.

5.3 Integrals involving the Gaussian basis functions

5.3.1 Matrix elements

The majority of the analytically solved integrals reported in this section are taken from the book by Thijssen[4]. The α_p represent the exponents of the Gaussian basis set employed in (10). We remind that the matrix elements are arranged according to the disposition in (16). Here we only report the AB case (north-east submatrix), the other cases are analogous. We use the convenient notation

$$\exp(-\alpha_p |\mathbf{r} - \mathbf{R}_A|^2) \exp(-\alpha_q |\mathbf{r} - \mathbf{R}_B|^2) = K_{pq} \exp(-(\alpha_p + \alpha_q)|\mathbf{r} - \mathbf{R}_C|^2)$$

and define the quantities

$$K_{pq} = \exp\left(-\frac{\alpha_p \alpha_q}{\alpha_p + \alpha_q} |\mathbf{R}_A - \mathbf{R}_B|^2\right) \qquad \mathbf{R}_C = \frac{\alpha_p \mathbf{R}_A + \alpha_q \mathbf{R}_B}{\alpha_p + \alpha_q}$$

The notation $\langle ... | ... \rangle$ will be used for the integrals involving the Gaussian functions. The elements of the overlap matrix $\mathcal{S}$:

$$\mathcal{S}_{pq} = \langle p, A | q, B \rangle = K_{pq} \left(\frac{\pi}{\alpha_p + \alpha_q}\right)^{3/2}$$

The one-body term (contained in the matrix $\mathcal{H}$):

$$\begin{aligned} &\langle p, A | -\nabla^2/2 | q, B \rangle = \frac{\alpha_p \alpha_q}{\alpha_p + \alpha_q} [3 - 2\ln(K_{pq})] \left(\frac{\pi}{\alpha_p + \alpha_q}\right)^{3/2} K_{pq} \\ &\langle p, A | -\frac{Z}{|\mathbf{r} - \mathbf{R}_{A/B}|} | q, B \rangle = -2\pi \frac{Z}{(\alpha_p + \alpha_q)} K_{pq} F_0\left((\alpha_p + \alpha_q)|\mathbf{R}_{A/B} - \mathbf{R}_C|^2\right) \end{aligned}$$

The two-body term collects the elements of the tensor $\mathcal{Q}$. As we did in the case of the square matrices, we can think of this tensor as composed of 16 subtensors. Here we represent an element of the subtensor $ABA'B'$.

$$\begin{aligned} \langle p, A; qB | \frac{1}{|\mathbf{r} - \mathbf{r}'|} | r, A'; s, B' \rangle = &\frac{2\pi^{5/2}}{(\alpha_p + \alpha_r)(\alpha_q + \alpha_s)\sqrt{(\alpha_p + \alpha_r + \alpha_q + \alpha_s)}} K_{pr}(\mathbf{R}_A, \mathbf{R}_{A'}) K_{qs}(\mathbf{R}_B, \mathbf{R}_{B'}) \\ &\times F_0\left(\frac{(\alpha_p + \alpha_r)(\alpha_q + \alpha_s)}{(\alpha_p + \alpha_r + \alpha_q + \alpha_s)} |\mathbf{R}_C(\alpha_p, \alpha_r, \mathbf{R}_A, \mathbf{R}_{A'}) - \mathbf{R}_P(\alpha_q, \alpha_s, \mathbf{R}_B, \mathbf{R}_{B'})|^2\right) \end{aligned}$$

with

$$F_0(t) = t^{-1/2} \int_0^{t^{1/2}} dy \; e^{-y^2}$$

5.3.2 Derivatives of the matrix elements

Derivative of the overlap matrix $\partial \mathcal{S} / \partial X$:

$$\frac{\partial \mathcal{S}_{pq}}{\partial X} = \frac{\partial}{\partial X} \langle p, A | q, B \rangle = -2 \frac{\alpha_p \alpha_q}{\alpha_p + \alpha_q} X \mathcal{S}_{pq}$$

One body terms:

$$\frac{\partial}{\partial X}\langle p,A|-\nabla^2/2|q,B\rangle = -4X\left(\frac{\alpha_p\alpha_q}{\alpha_p+\alpha_q}\right)^2\mathcal{S}_{pq}+\left[3\frac{\alpha_p\alpha_q}{\alpha_p+\alpha_q}-2\left(\frac{\alpha_p\alpha_q}{\alpha_p+\alpha_q}\right)^2X^2\right]\frac{\partial\mathcal{S}_{pq}}{\partial X}$$

The derivative is different in the case of functions centred on the same nucleus (AA or BB) and on different nuclei (AB or BA). Here we present the separated cases AA and AB.

$$\frac{\partial}{\partial X}\langle p,A|-\frac{Z}{|\mathbf{r}-\mathbf{R}_{A/B}|}|q,A\rangle = \sqrt{\frac{\alpha_p+\alpha_q}{\pi}}\mathcal{S}_{pq}F_0'(\alpha_pX^2+\alpha_qX^2)X(\alpha_p+\alpha_q)$$

$$\frac{\partial}{\partial X}\langle p,A|-\frac{Z}{|\mathbf{r}-\mathbf{R}_{A/B}|}|q,B\rangle = 2\sqrt{\frac{\alpha_p+\alpha_q}{\pi}}\frac{\partial\mathcal{S}_{pq}}{\partial X}(F_0(t_1)+F_0(t_2))+2\sqrt{\frac{\alpha_p+\alpha_q}{\pi}}\frac{\mathcal{S}_{pq}}{\alpha_p+\alpha_q}(F_0'(t_1)\alpha_p^2+F_0'(t_2)\alpha_q^2)X$$

with

$$F_0' = \frac{dF_0(t)}{dt} = \frac{e^{-t}-F_0(t)}{2t} \qquad t_1 = \frac{\alpha_p^2X^2}{\alpha_p+\alpha_q} \qquad t_2 = \frac{\alpha_q^2X^2}{\alpha_p+\alpha_q}$$

The last two-body term

$$\frac{\partial}{\partial X}\langle p,A;qB|\frac{1}{|\mathbf{r}-\mathbf{r}'|}|r,A';s,B'\rangle = \rho\frac{\partial\mathcal{S}_{pq}(A,B)}{\partial X}\mathcal{S}_{rs}(A',B')F_0(t)+$$
$$\rho\mathcal{S}_{pq}(A,B)\frac{\partial\mathcal{S}_{rs}(A',B')}{\partial X}F_0(t)+\rho\mathcal{S}_{pq}(A,B)\mathcal{S}_{rs}(A',B')F_0'(t)\frac{2t}{X}$$

with

$$\rho = 2\sqrt{\frac{(\alpha_p+\alpha_q)(\alpha_r+\alpha_s)}{\pi(\alpha_p+\alpha_q+\alpha_r+\alpha_s)}} \qquad t = \frac{(\alpha_p+\alpha_q)(\alpha_r+\alpha_s)}{\pi(\alpha_p+\alpha_q+\alpha_r+\alpha_s)}|\mathbf{R}_C-\mathbf{R}_C'|^2$$

5.4 Conversion table

$\hbar = m_e = e = 1$

Quantity	Atomic unit conversion
Lenght	Bohr radius, $a_0 = 0.528\times10^{-8}$ cm
Energy	Hartree, 1 H = 27.21 eV
Time	2.4188×10^{-17} s
Frequency	4.13×10^{16} Hz
Mass	$m_e = 9.109\times10^{-31}$ kg
Damping coefficient	3.7659×10^{-14} kg/s

References

1. Marx, D. & Hutter, J. Ab initio molecular dynamics: Theory and implementation. *Mod. methods algorithms quantum chemistry* **1**, 141 (2000).
2. Boero, M. & Oshiyama, A. Car–parrinello molecular dynamics (2015).
3. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. computational physics* **23**, 327–341 (1977).
4. Thijssen, J. *Computational physics* (Cambridge university press, 2007).
5. Corongiu, G. & Clementi, E. Energy and density analyses of the 1σu+ states in the h2 molecule from the united atom to dissociation. *The J. Phys. Chem. A* **113**, 14791–14799 (2009).
6. Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048 (1981).

7. Gunnarsson, O. & Lundqvist, B. I. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B* **13**, 4274 (1976).

8. Lyness, J. N. Notes on the adaptive simpson quadrature routine. *J. ACM (JACM)* **16**, 483–495 (1969).

9. Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. & Joannopoulos, a. J. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. modern physics* **64**, 1045 (1992).