Quantum Optimal Control in Biophysics: studying optimization methods

Work presented as Master's thesis for Master degree in Quantitative Biotechnology

Author: Aitor Sánchez Mansilla
Thesis Director 1: Alberto Castro Barrigón
Thesis Director 2: Jesús Clemente Gallardo
Academic year: 2017/18

Abstract

The main purpose of this work is to study, computationally, the control of ionization of molecules upon radiation with high-intensity, short-duration (femtoseconds) laser pulses. This study is focused on the tailoring of the temporal shape of the pulse and it is placed in the context of previous works[1][2][3], that attempt to provide a theoretical counterpart to the rapid advances in the experimental control of ultrafast phenomena via laser pulse shaping. For that purpose, I have used Quantum Optimal Control Theory (QOCT), that deals with the problem of designing control fields that force the system to behave in some predefined *optimal way*. In this work QOCT is combined with Density Functional Theory (DFT), that is an alternative approach to the many-electron problem in a more suitable way for computational physics, specially for many-electron systems. The combination of those two theories is applied to the alanine molecule case placing the work in the context of bio-sciences. Finally, a comparison of different types of optimization algorithms that can be used to solve the QOCT problem is presented.

Contents

Chapter 1

Introduction

Laser sources reached in the 90s the femtosecond time scale –the scale of the chemical reactions–, and quickly surpassed this threshold to reach the attosecond time scale –the scale of the electronic motion in condensed matter systems–. This induced to think in a control or optimization of these physical processes by the appropriate modulation of the incident laser fields, giving birth to a new discipline: Quantum Optimal Control Theory.

The mathematical framework accompanying this emergent discipline has to deal, first, with the N-electron problem: The aim is to obtain the ground state configuration of a N-electron system considered in the presence of nuclei. Even considering the Born-Oppenheimer approximation –that assumes separated nuclear and electronic degrees of freedom–, the problem is still intractable for large electron numbers, as the computational cost grows exponentially with that number. The origin of the difficulties lies in the extremely large size of the fundamental object of the theory, the many-electron wave function, an object dependent on the coordinates of all the N electrons, and in the existence of the electron-electron Coulomb interaction.

Since the early times of Quantum Mechanics, there have been many solutions proposed for this problem. The first of them was the Thomas-Dirac-Fermi model (TDF) based on the Thomas-Fermi theory for an electron gas [5][6]. Its authors tried to model the exchange and correlation energy of the system in terms of the total electronic density $\rho(\mathbf{r})$, an object that only depends on one spatial coordinate, and which therefore greatly reduces the size of the full many-electron wave function. This attempt was unsuccessful, as proved by Teller, due to the incapacity of the model to stabilize molecules [7], because the energy of the molecule obtained by this method was greater than the sum of the energies of each atom. The TDF model is therefore only appropriate for a crude estimation of atomic properties.

One step forward was the formulation of the Hartree-Fock method (HF) and its variations (i.e. post-Hartree-Fock methods, that include more sophisticated correlation effects in place of a mean field). The HF approximations assumes that the wavefunction that summarizes the behavior of the system is a single Slater determinant of the one-electron wavefunctions. The minimization of the energy of this determinant leads to a self-consistent set of *HF orbital equations*, one for each one-electron wavefunction, that should be iterated in order to achieve a self-consistent solution (hence the name Self-Consistent Field (SCF) method). Still, HF is not enough to accurately predict many important properties of molecules, and materials, and post-HF methods (essentially, the inclusion of more Slater determinants in the space search for the optimal wave function) are required. Those methods, though can be very accurate, have the problem of the elevated computational cost due to the its fast scaling with the system size.

A further simplification to these methods is to consider classical Ehrenfest Dynamics for nuclei, as their movement is much more slower than the electronic motion. The model now consists in an array of nuclei that change its positions classically while the electronic cloud evolves in a quantum way and adapts instantly to the array's conformation. This simplifies

the nuclear part in the SCF method but the electronic part continues being a challenge in computational cost due to the big size of the electronic wavefunction.

Density Functional Theory (DFT) was born as an alternative solution in an attempt to simplify the calculations in computers, avoiding this drawback of wave-function based methods. This theory rescues, in some way, the idea of TDF of modelling the energies of the system as functionals[1] of the electronic density. This substitution achieves a huge reduction of degrees of freedom from $3N$ to 3 and as a consequence, a reduction in computational cost. DFT was introduced in 1964 by the two Hohenberg-Kohn theorems (HK) [8] and further developed by Kohn and Sham [9]. The extension to time-dependent systems to study the dynamics of many-body quantum systems is called Time-Dependent Density Functional Theory (TDDFT) and it was introduced in 1984 by Runge and Gross [10].

The other backbone of this work is Quantum Optimal Control Theory (QOCT)[11], which is the application to quantum systems of the more general Optimal Control Theory (OCT) –for general dynamical systems–, that establishes the necessary mathematical framework for the most recent advances in *coherent control*, the control of quantum systems with laser pulses and their optimization. The background of coherent control dates back to the study of the interaction between radiation and matter, i.e. *spectroscopy*: as radiation induces a reaction of matter, one may attempt to control this reaction by modifying the incident radiation. The invention of laser (Mainman[12], 1960) granted an ideal tool to manipulate matter at the microscopic scale. The laser sources are able to produce coherent emission and an almost monochromatic spectrum. The goal of coherent control, roughly speaking, is to find the form of the laser field that induces a determined reaction on the matter under study. For example, one may attempt to break a given bond of a molecules by tuning the laser frequency to the characteristic bond vibration.

However, it was soon realised that such simple-minded mono-chromatic laser pulses did not always do the job correctly. Several limitations had to be overcome –intramolecular vibrational redistribution[13], collective molecular motions, etc–, and it became clear that that the pulses had to be more complex than simply monochromatic. The improvement of *pulse shapers* allowing the creation of almost arbitrary laser pulses and the *adaptive feedback control*[14] (AFC) experimental scheme are the culminating point in this history, that led the way until the state of this discipline nowadays. The AFC scheme introduced, given a desired reaction in the matter, the concept of *performance measure*, i.e. the level of achievement of the task that laser has to produce. For example, if the goal is to induce a chemical reaction, experimentally one may perform a time-of-flight-spectroscopy of the products, obtaining the yield for the products that are wanted and the yield of those that are not. Then, the laser pulse can be changed and retried, so that a loop can be closed in order to optimize it. QOCT was developed alongside with those experimental advances as theoretical framework. In its essence, it consists of a set of equations (the *control equations*) that must be satisfied by the optimal external field, given a predefined *performance measure*.

The structure of this thesis is as follows: Chapter 2 is a brief introduction to the molecular modelling methods, including the tools and assumptions that are used in this work. In chapter 3, I summarize the main aspects of DFT and TDDFT. In chapter 4, I introduce QOCT as the application of OCT to quantum systems and in particular to the alanine molecule. Chapter 5 is a brief explanation of the algorithms used for optimization in the work. Finally, chapter 6 contains the results of the simulations and chapter 7 contains the conclusions reached. In particular, the computational tool for DFT and QOCT used in this work is the octopus-code (`octopus-code.org`).

[1] A functional F is a map which assigns a number $F[f]$ to a function f.

Chapter 2

Theoretical modelling of molecules

Molecules, as neutral groups of atoms held by chemical bonds, can be viewed as composed only of positively charged nuclei and negatively charged electrons. In a full quantum description, a system of K nuclei and N electrons is described by the Hamiltonian:

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{ne}(\mathbf{r}_1, ..., \mathbf{r}_N, \mathbf{R}_1, ..., \mathbf{R}_K) + \hat{V}_{ee}(\mathbf{r}_1, ..., \mathbf{r}_N) + \hat{V}_{nn}(\mathbf{R}_1, ..., \mathbf{R}_K)\,, \tag{2.0.1}$$

where $\mathbf{R}_1, .., \mathbf{R}_K$ are the position operators of the nuclei, and $\mathbf{r}_1, ..., \mathbf{r}_N$ are the position operators of the electrons. The first two terms are the kinetic energies of nuclei of electrons, respectively, and the rest are Coulombian interaction potentials between the two types of particles. These are defined as (α and β run over the K nuclei, whereas the latin index i runs over the electron coordinates):

$$\hat{T}_n = \sum_{\alpha=1}^{K} \frac{1}{2M_\alpha} \nabla_\alpha^2$$

$$\hat{T}_e = -\frac{1}{2} \sum_{i=1}^{N} \nabla_i^2$$

$$\hat{V}_{ne} = \sum_{i=1}^{N} \sum_{\alpha=1}^{K} \frac{-Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} = \sum_{i=1}^{N} v_{\text{nuc}}(\mathbf{r}_i)$$

$$\hat{V}_{ee} = \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

$$\hat{V}_{nn} = \sum_{\alpha \neq \beta} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}$$

From this point until the end of this chapter I will use $\mathbf{r} \equiv (\mathbf{r}_1, ..., \mathbf{r}_N)$ and $\mathbf{R} \equiv (\mathbf{R}_1, ..., \mathbf{R}_K)$ in order to use a lighter notation. The eigenstates of the system are the wavefunctions of well-defined energy [15]:

$$\hat{H}\Psi(\mathbf{r}, \mathbf{R}) = E_{\text{total}}\Psi(\mathbf{r}, \mathbf{R}) \tag{2.0.2}$$

The resolution of this equation presents the problem of a very high scalability in complexity and therefore in computational time with the size of the function Ψ, which depends on the $3\times(N+K)$ nuclear and electronic degrees of freedom. Hence, some approximations are necessary in order to define a tractable problem. In this chapter we will review those which are behind the computational tool we use in this thesis.

2.1 Separability assumption: Born-Oppenheimer approximation

The first approximation to be used in computational molecular modelling is the *Born-Oppenheimer (BO) approximation* [16], which states that the degrees of freedom of nuclei and electrons can

be separated. Thus, the molecular wavefunction (or molecular quantum state) can be broken into its electronic and nuclear components, related by the product ansatz:

$$\Psi_{\text{total}}(\mathbf{r},\mathbf{R}) = \psi_e(\mathbf{r};\mathbf{R}) \otimes \psi_n(\mathbf{R}) = \psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R})$$

where the ";" indicates that, though the electronic wavefunction ψ is a function of the electron positions $\mathbf{r}_i$, it depends parametrically on the positions of nuclei $\mathbf{R}_i$. This approximation is based on the assumption that electrons move much faster than nuclei, that are several thousand times heavier that them, adapting instantly to changes in the nuclei conformation without requiring a finite relaxation time. The problem, though still purely quantum, allows to be solved in two steps: one first step for the electrons only, considering the nuclei to be fixed points (classical particles), and one second step for the nuclear wavefunctions, that would move in a field that can be computed in that first step. Notice that this assumption means the absence of entanglement between the nuclear and electronic states.

The first step can be taken by considering the Hamiltonian (2.0.1) for the electronic wavefunction, but considering that the term T_n can be neglected for a fixed nuclear configuration as it is smaller than T_e by a factor $1/M_\alpha$. The term $\hat{V}_{nn}$ can also be taken out, because it is constant if the nuclei are clamped, so it gives no relevant information. The electronic Hamiltonian remains:

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} \tag{2.1.1}$$

We may then first solve for the eigenvalues and eigenfunctions of this purely electronic Hamiltonian (the nuclear positions are in this case mere parameters), obtaining for example the ground state electronic wave function $\psi_e(\mathbf{r};\mathbf{R})$. After this, we may then substitute the full wavefunction by the BO ansatz into equation (2.0.2):

$$\begin{aligned}\hat{H}\psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R}) = \hat{T}_n(\mathbf{R})\psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R}) + \hat{T}_e(\mathbf{r})\psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R}) + \\ +\hat{V}_{ne}(\mathbf{r};\mathbf{R})\psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R}) + \psi_e(\mathbf{r};\mathbf{R})\hat{V}_{nn}(\mathbf{R})\psi_n(\mathbf{R})\end{aligned} \tag{2.1.2}$$

Defining E_e as the eigenvalue of the electronic Hamiltonian (2.1.1), the terms can be grouped depending on which state the operators act. Operators $\hat{T}_e$ and $\hat{V}_{ee}$ act clearly only on electronic states as they do not depend in any way on nuclei. The interaction potential $\hat{V}_{ne}$ does not depend on nuclei because they are considered clamped in the electronic motion, so their positions are considered as parameters. These three terms conform the electronic Hamiltonian whose action is equal to the multiplication by the eigenvalue E_e for the electronic eigenstates ψ_e, resulting in the equation:

$$E_e(\mathbf{R})\psi_e(\mathbf{r};\mathbf{R})\psi_n(\mathbf{R}) + \psi_e(\mathbf{r};\mathbf{R})\left[\hat{T}_n(\mathbf{R}) + \hat{V}_{nn}(\mathbf{R})\right]\psi_n(\mathbf{R}) = E_{\text{total}}\psi_e(\mathbf{r};\mathbf{R})\psi_n \tag{2.1.3}$$

Then, assuming that the electronic ψ_e and nuclear ψ_n wavefunctions are normalized separately we can project onto each one of the two types of eigenstate, obtaining the equations for each one of them:

$$\hat{H}_e\psi_e(\mathbf{r};\mathbf{R}) = \left[\hat{T}_e(\mathbf{r}) + \hat{V}_{ne}(\mathbf{r};\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\right]\psi_e(\mathbf{r};\mathbf{R}) = E_e\psi_e(\mathbf{r};\mathbf{R}) \tag{2.1.4}$$

$$\hat{H}_n\psi_n(\mathbf{R}) = \left[\hat{T}_n(\mathbf{R}) + \hat{V}_{nn}(\mathbf{R})\right]\psi_n(\mathbf{R}) = E_n\psi_n(\mathbf{R}) \tag{2.1.5}$$

where E_n is the eigenstate of the nuclear Hamiltonian $\hat{H}_e$ In a similar fashion, we may obtain the corresponding *time-dependent* Schrödinger equations that provide the dynamics of the system:

$$i\hbar\frac{\partial}{\partial t}\psi_e(\mathbf{r};\mathbf{R}) = \left[\hat{T}_e(\mathbf{r}) + \hat{V}_{ne}(\mathbf{r};\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\right]\psi_e(\mathbf{r};\mathbf{R}) = \hat{H}_e\psi_e(\mathbf{r};\mathbf{R}) \tag{2.1.6}$$

$$i\hbar\frac{\partial}{\partial t}\psi_n(\mathbf{R}) = \left[\hat{T}_n(\mathbf{R}) + \hat{V}_{nn}(\mathbf{R})\right]\psi_n(\mathbf{R}) = \hat{H}_n\psi_n(\mathbf{R}) \tag{2.1.7}$$

The step of BO approximation takes care of *separation*, and its success in general depends on the quotient of the masses of both types of particles.

2.2 Quantum-classical assumption

The next assumption (Bornemann et al., 1996 [17][18]) leads us to the mixed *quantum-classical molecular dynamics* scheme by means of two basic assumptions. The first assumption consist in considering that the probability function $|\psi_n|^2$ of the nuclear state tends to an approximated δ function of small support, that allows to deal with them through classical dynamics. Bornemann proved that if this assumption is fulfilled then the error committed in the separation step along the temporal evolution, that is, the difference between the separated state and the true state is small [18]:

$$\|\psi_e(t) \otimes \psi_n(t) - \Psi(t)\| \to \mathcal{O}(\epsilon/L) \quad \text{if} \quad \psi_n \approx \delta_n \tag{2.2.1}$$

The success of this first assumption depends on the smallness of the variance $\epsilon(t)^2 \ll 1$ such that $\epsilon(t) \leq \epsilon \ll 1$, which is compared with the length of the system L. This variance gives an idea of the nuclear position uncertainty as the radius of the sphere where the nucleus can be found. The meaning of this is that, considering initially not entangled nuclei and electrons, we know that the temporal evolution will tend to entangle them, but the error that we commit depends on the ratio ϵ/L, that tends to 0 if the nuclear wavefunction is an approximate δ.

The second approximation step consist in the construction of the nuclear wavefunction for large masses $M \gg m$ for nuclei through the WKB approximation. Being an approximate δ, it can be expanded asymptotically as a quantum(for a nucleus α):

$$\psi_n^{QC}(\mathbf{R}_\alpha, t) = a(\mathbf{R}_\alpha, t) \exp\left(i\frac{S(\mathbf{R}_\alpha, t)}{\hbar}\right) + \mathcal{O}\left(\sqrt{\frac{m}{M_\alpha}}\right) \tag{2.2.2}$$

This step is called *quantum-classical approximation* and split the wavefunction into a "classical" probability amplitude $a(\mathbf{R}_\alpha, t)$ multiplied by an exponential phase factor dependent on the phase $S(\mathbf{R}_\alpha, t)$. These terms obey the equations:

$$\frac{\partial S}{\partial t} + \frac{1}{2M_\alpha}(\nabla_R S)^2 + \langle\psi_e \hat{V} \psi_e\rangle = 0 \quad \text{Hamilton-Jacobi equation} \tag{2.2.3}$$

$$\frac{\partial a^2}{\partial t} + \text{div}_R\left(a^2 \frac{\nabla_R S}{M_\alpha}\right) = 0 \quad \text{Continuinity equation} \tag{2.2.4}$$

So the phase S corresponds to a *classical action* –it fulfills a Hamilton-Jacobi equation– and the continuity equation for a ensures a local conservation of the probability density, which implies that an initially concentrated probability function remains, when $\epsilon \to 0$ concentrated along a classical path $\mathbf{R}_\alpha(t)$:

$$a^2(\mathbf{R}_\alpha, t) \to \delta(\mathbf{R}_\alpha - \mathbf{R}_\alpha(t)) \tag{2.2.5}$$

If S is the classical action then the theory states that the solution of the canonical equations:

$$\dot{\mathbf{R}}_\alpha = \frac{1}{M_\alpha}\mathbf{P}_\alpha, \quad \mathbf{R}_\alpha(0) = \mathbf{R}_{\alpha,0} \tag{2.2.6}$$

$$\dot{\mathbf{P}}_\alpha = -\langle\psi_e \nabla_R \hat{V} \psi_e\rangle, \quad \mathbf{P}_\alpha(0) = \nabla_q S(\mathbf{R}_{\alpha,0}, 0) \tag{2.2.7}$$

is fulfilled by

$$\mathbf{P}_\alpha(t) = \nabla_R S[\mathbf{R}_\alpha(t), t] \tag{2.2.8}$$

This allows to construct a relationship $S(.,0) \to S(\cdot, t)$ (here the $\cdot$ indicates that the position $\mathbf{R}_\alpha$ is fixed) while the relationship $\dot{\mathbf{R}}_\alpha(0) = M_n^{-1}\nabla_R S(\mathbf{R}_{\alpha,0}, 0)$ gives the map to the solution $\mathbf{R}_\alpha(0) \to \mathbf{R}_\alpha(t)$. This approximation can be affected by the presence of *caustics* (a quantum

phenomenon), that are surfaces formed by the focal points –in where two of more paths cross at time $t = t_{\max}$ and therefore the function S is multivalued– so we will refer only to the time interval $t < t_{\max}$. With this assumption, the correspondence of the phase and amplitude with the classical probability density and action can be proved and it leads to the classical variables $\mathbf{R}_\alpha, \mathbf{P}_\alpha$ by the relation (2.2.8) and the continuity equation for a, so the dynamics remain:

$$\dot{\mathbf{R}_\alpha} = \frac{\mathbf{P}_\alpha}{M_\alpha} \tag{2.2.9}$$

$$\dot{\mathbf{P}_\alpha} = -\nabla\langle\psi_e|\hat{V}|\psi_e\rangle \tag{2.2.10}$$

$$i\hbar\frac{\partial\psi_e}{\partial t} = \hat{H}_e\psi_e \tag{2.2.11}$$

According to [17] the error committed with these two approximations depends also on the mass ratio m/M_α:

$$\|\psi_e(t)\otimes\psi_n^{QC}(t) - \Psi(t)\| \to \mathcal{O}(\epsilon/L + \sqrt{m/M}) \;\text{ if }\; \psi_n \approx \delta_n \tag{2.2.12}$$

So, the consequence is that the approximations taken into account are valid in the cases in which the variance $\epsilon(t)$ is small enough compared with the dimension L of the system and $M \gg m$ that is fulfilled in the case of nuclei and electrons. These conditions can justify that our model reproduces faithfully the real system, that allows us to extract relevant conclusions. The next step is to consider is to simplify the last equation from the system above, that corresponds to the electron dynamics. This is done by Density Functional Theory which is explained in the next chapter.

2.3 LCAO and pseudo-potential approximation

Beyond the mere definition of the equations, one must consider the problem of representing them numerically in order to perform computations in practice. Perhaps the most common representation used in molecular modelling is the use of *linear combination of atomic orbitals* (LCAO), a technique for constructing molecular orbitals[19] in quantum chemistry. Each of the molecular orbitals ψ is built, by assumption, by a linear combination of the same number n of χ_i atomic orbitals (or basis functions) such as:

$$\psi = c_1\chi_1 + c_2\chi_2 + ... + c_n\chi_n \tag{2.3.1}$$

where the coefficients c_i have to be determined and are weighted as the contributions of each atomic orbital. These orbitals can be of several types:

- *Hydrogen-like orbitals*: Orbitals of one-electron systems but with the appropriate nuclear charge.
- *Slater type orbitals* (STOs): Orbitals introduced by Slater [20] with the following radial part:

$$R(r) = Nr^{n-1}e^{-\xi r} \qquad N = (2\xi)^n\sqrt{\frac{2\xi}{(2n)!}} \tag{2.3.2}$$

  where n is the principal quantum number (n=1,2,...), r is the distance from the electron to the nucleus and ξ is related to the effective charge determined by Slater's rules of the nucleus shielded in part by the core electrons. The angular part of STOs is often represented by the usual *spherical harmonics* $Y_l^m(\mathbf{r})$.

- *Gaussian type orbitals*[21]: Their radial part has a gaussian dependence on r. Cartesian GTOs have the form:

$$G_{ikj,\alpha} := N_{ijk,\alpha}(x - R_x)^i(y - R_y)^j(z - R_z)^k e^{-\alpha(r-R)^2} \text{ with } i + j + k = n \tag{2.3.3}$$

  where N is the normalization factor. They are not eigenstates of the angular momentum operator $\hat{L}$ but can be transformed into spherical GTOs:

$$G_{lmc,\alpha} = {}^{c}R_l^m(\mathbf{r})e^{-\alpha r^2} \quad G_{lms,\alpha} = {}^{s}R_l^m(\mathbf{r})e^{-\alpha r^2} \tag{2.3.4}$$

  where the "c" or "s" means if it has sine or cosine form.

There are several facts to be mentioned about these types of orbitals. STOs have the advantage of having a good physical interpretation, but the integrals that they implicate are needed to be solved numerically in SCF methods such as DFT, that involves an additional computational effort. On the contrary, GTOs have the advantage of the *Gaussian Product Theorem* which guarantees that the product of two GTOs centered on different atoms is a finite sum of GTOs centered on a point along the axis that connect them. This allows to compute integrals easily in contrast to STOs. In particular, the set of molecular orbitals used as a starting point by Octopus-code is a LCAO.

However, the main representation scheme of the octopus code is not atom-centered localized orbitals, but a numerical grid: functions are represented by the values that they take at the points of a normally equi-spaced rectangular regular grid in real space. This typically involves a larger dimension of the Hilbert space, but has the advantage of simplicity, and of the large sparsity of the operator matrices (i.e. the operators are very local in real space, and therefore almost all of the matrix elements are zero). The use of a grid spacing (and of other basis sets, such as plane waves) implies a difficulty, however, when dealing with the Coulomb potential functions, that become infinite when the distance between particles are zero. Atoms may therefore not be represented exactly on top of grid points, and when they approach the grid points there is a severe numerical problem. Because of this, and in order to avoid the treatment of core electrons that are typically inert, one uses pseudopotentials. Pseudo-potentials[22] are an attempt to replace the full atomic cores – non valence electrons and the nucleus – by an effective potential, such that the core states are eliminated and the valence electrons are described by pseudo-wavefunctions. These wavefunctions can be described by smooth functions in real space – or by a short Fourier series when using plane waves. The octopus software uses different types and formats of pseudo-potentials depending on the species.

Summarizing, the approximations and approaches taken into account until now are:

- Separation of the electronic and nuclear degrees of freedom through the Born-Oppenheimer approximation.
- Classic treatment of the nuclear motion.
- The pseudo-potential approximation in order to deal with the electron-nucleus interaction, and a grid-space representation to construct and represent molecular orbitals –although LCAO are also used internally in the code for auxiliary purposes–.

However, we are still left with a many-electron problem that is impossible to approach directly. The approach to this problem, shown in the next chapter, and which has been used for the calculation below, is Density Functional Theory (DFT).

Chapter 3

Density Functional Theory

Density Functional Theory is a modelling method born in the 60's. It is a versatile alternative to map the initial many-body problem onto a single-body problem by means of substituting the full interacting wave function by the electronic density as the key object. That idea was already present in Thomas-Dirac-Fermi Theory [5][6]. Yet DFT succeeds first in laying a firm theoretical basis for this substitution, and then by establishing a set of equations, that are similar in form to the HF equations. In summary, the goal of DFT is to re-formulate the quantum properties of many-body electronic systems like atoms, molecules and condensed phases in terms of the electronic density function $n(\mathbf{r})$. This function depends only on the three spatial coordinates $\mathbf{r} = (r_x, r_y, r_z)$, unlike the many-electron wavefunction $\psi(\mathbf{r}_1, ..., \mathbf{r}_N)$ that depends on the positions $\mathbf{r}_1, ..., \mathbf{r}_N$ of each electron, hence the reduction of degrees of freedom is from $3N$ to 3. This section is separated in two parts: the ground-state DFT and its time-dependent extension TD-DFT.

3.1 Ground-State Density Functional Theory

For this section we start from the Hamiltonian for N electrons in the presence of the nuclei: these are already considered to be a clamped array that causes an external coulombian potential $\hat{V}_{ne} = \hat{V}_{\text{ext}}$, so:

$$\hat{H} = \hat{T}_e + \hat{V}_{\text{ext}} + \hat{V}_{ee} \tag{3.1.1}$$

The terms above are the same as in the previous section:

$$\hat{T}_e = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 \tag{3.1.2}$$

$$\hat{V}_{\text{ext}} = \sum_{i=1}^{N} v_{\text{nuc}}(\mathbf{r}_i) = \sum_{i=1}^{N}\sum_{\alpha=1}^{K} \frac{-Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \tag{3.1.3}$$

$$\hat{V}_{ee} = \frac{1}{2}\sum_{i}\sum_{j\neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{3.1.4}$$

The eigenstates of the Hamiltonian are the wavefunctions ψ_k of well-defined energy E_k:

$$\hat{H}\psi_k(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, ..., \mathbf{r}_N\sigma_N) = E_k\psi_k(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, ..., \mathbf{r}_N\sigma_N) \tag{3.1.5}$$

Here, the index k runs over the set of many-electron quantum numbers. The index σ_i is the spin of electron i. Dealing with fermions, the solution can only be an antisymmetric function under exchange of two electron labels[23]:

$$\psi_k(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, ..., \mathbf{r}_i\sigma_i, ..., \mathbf{r}_j\sigma_j, ..., \mathbf{r}_N\sigma_N) = -\psi_k(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, ..., \mathbf{r}_j\sigma_j, ..., \mathbf{r}_i\sigma_i, ..., \mathbf{r}_N\sigma_N)$$

In order to build a many-electron wavefunction that fulfills the condition above, the most practical way is to build Slater's determinants from sets of one-electron orbitals $\phi_1, \ldots, \phi_N$:

$$\psi_k(\mathbf{r}_1\sigma_1, ..., \mathbf{r}_N\sigma_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1\sigma_1) & \phi_2(\mathbf{r}_1\sigma_1) & \cdots & \phi_N(\mathbf{r}_1\sigma_1) \\ \phi_1(\mathbf{r}_2\sigma_2) & \phi_2(\mathbf{r}_2\sigma_2) & \cdots & \phi_N(\mathbf{r}_2\sigma_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{r}_N\sigma_N) & \phi_2(\mathbf{r}_N\sigma_N) & \cdots & \phi_N(\mathbf{r}_N\sigma_N) \end{vmatrix} \equiv \langle \mathbf{r}_1\sigma_1, \ldots, \mathbf{r}_N\sigma_N | \phi_1, \cdots, \phi_N \rangle \tag{3.1.6}$$

Any many-electron wave-function can be written as a linear combination of these determinants.

The *one-electron spin density* corresponding to *any wavefunction* ψ, is defined as the object:

$$n_\sigma(\mathbf{r}) = N \sum_{\sigma_2,\ldots,\sigma_N} \int \mathrm{d}^3 r_2 ... \int \mathrm{d}^3 r_N \; |\psi_k(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, ..., \mathbf{r}_N\sigma_N)|^2 \tag{3.1.7}$$

which is the probability to find *one electron* with spin σ in volume d^3r. Here, the sum is over all $\sigma_i = \uparrow, \downarrow$ and notice that the integral is above all other electrons. The aim of DFT is to replace the many-electron wavefunction (3.1) by the *one-electron density*:

$$n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r}) \tag{3.1.8}$$

as the basic variable in order to deal with 3 degrees of freedom (r_x, r_y, r_z) in contrast to the 3N of the wavefunction, and writing all magnitudes in terms of it.

In order to do that, in the ground-state case, it was shown that the external potential, denoted as v_{ext}, is a unique functional of $n(\mathbf{r})$, so $v_{\text{ext}} = v_{\text{ext}}[n(\mathbf{r})]$[1]. This result is known as the **DFT theorem or first Hohenberg-Kohn theorem**. The theorem, in fact, establishes a one-to-one correspondence between the external potential and the electron density. Note that this does not directly imply a one-to-one correspondence with the ground state wavefunction, as this state can be degenerated. If we ignore this possibility of degeneration, however, we may also view the ground-state as a functional of the ground-state one-electron density in the way that $\psi_0 = \psi[n(\mathbf{r})]$.

In any case, what we may guarantee is that the ground-state's expected value of any operator, such as the ground state energy, is also a functional of $n(\mathbf{r})$:

$$E_0 = E[n(\mathbf{r})] = \langle \psi[n(\mathbf{r})] | \hat{T}_e + \hat{V}_{\text{ext}} + \hat{V}_{ee} | \psi[n(\mathbf{r})] \rangle \,, \tag{3.1.9}$$

as the value should not depend on which of the degenerate states is chosen.

The existence of this energy functional suggests a procedure to obtain the ground state density and energy: minimize the energy functional, by application of the minimum energy principle. The next step would therefore be to minimize the energy functional (3.1.9) by variational calculation in order to reach the correct ground-state energy, and thus, the ground-state electron density. The problem of this procedure is that the form of this functional is unknown.

The solution proposed by W. Kohn and L. J. Sham was the use of an auxiliary ideal system of non-interacting electrons, called **Kohn-Sham system**, whose density is however by construction identical to the density of the real one. As it is a system of non-interacting electrons,

[1] The [...] means functional dependence on a function, assigning a number to it

its wavefunction may be a single Slater determinant, and each of the orbitals that compose this determinant must fulfill a Schrödinger-type equation:

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right)\varphi_{k\sigma}(\mathbf{r}) = \hat{H}_{\text{KS}}[n(\mathbf{r}),\mathbf{r}]\varphi_{k\sigma} \tag{3.1.10}$$

$$= \epsilon_{k\sigma}\varphi_{k\sigma}(\mathbf{r})\,, \tag{3.1.11}$$

where $\varphi_{k\sigma}$ are the Kohn-Sham eigenfunctions for the non-interacting system, called **Kohn-Sham orbitals**, and $\epsilon_{k\sigma}$ are the Kohn-Sham energy levels.

The effective potential (sometimes called mean field potential, or Kohn-Sham potential) is:

$$v_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + \int \mathrm{d}^3r' \, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{\text{xc}}[n](\mathbf{r}) \tag{3.1.12}$$

that includes the "real" external potential (typically, the nuclear potential defined above), the potential due to self-repulsion (also known as Hartree potential, the second term in the previous equation) and the last term v_{xc} which is the called **exchange-correlation potential**. This term includes all the many-particle interactions omitted in all the other terms. Its functional form with respect to the density is unknown, and has to be approximated. Therefore, the resulting system is a non-interacting many-electron system with an extra term v_{xc} gathering all the unknown contributions to energy. The electron spin density in terms of Kohn-Sham states is written as:

$$n_\sigma(\mathbf{r}) = \sum_k \theta(\mu - \varepsilon_{k\sigma})|\varphi_{k\sigma}(\mathbf{r})|^2 \tag{3.1.13}$$

where the step-function θ ensures that all Kohn-Sham states with $\varepsilon_{k\sigma} < \mu$ are singly occupied and those with $\epsilon_{k\sigma} > \mu$ are empty such that the Pauli's principle is satisfied. The chemical potential μ is chosen to satisfy that the integral of the electron density over the whole space is the number of electrons:

$$\int \mathrm{d}^3r \; n(\mathbf{r}) = N \tag{3.1.14}$$

Note that the Kohn-Sham equations are non-linear. On first sight, they look like a eigenvalue problem (i.e. the problem of obtaining the lowest eigenvalues and eigenfunctions of the Kohn-Sham Hamiltonian). But the definition of the potential depends on the density –it is a functional of the density– which can only be computed if the solution eigenfunctions are known. In other words, the only way to solve the equations (3.1.11) and (3.1.13) is by iteration to self-consistency as they are interlinked.

The total energy of the system is:

$$E = T_{\text{KS}}[n_\uparrow, n_\downarrow] + \int \mathrm{d}^3r n(\mathbf{r})v_{\text{ext}}(\mathbf{r}) + \frac{1}{2}\int \mathrm{d}^3r \int \mathrm{d}^3r' \, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{\text{xc}}[n] \tag{3.1.15}$$

where the first term:

$$T_{\text{KS}}[n_\uparrow, n_\downarrow] = \sum_\sigma \sum_k \theta(\mu - \varepsilon_{k\sigma})\langle\varphi_{k\sigma}| - \frac{1}{2}\nabla^2|\varphi_{k\sigma}\rangle \tag{3.1.16}$$

is the non-interacting kinetic energy of the Kohn-Sham system. The term E_{xc} is the exchange and correlation (xc) energy functional. The xc potential used above can be deduced from it by functional differentiation:

$$v_{xc}[n](\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}\,. \tag{3.1.17}$$

It remains to establish an approximation for this energy term. There are nowadays hundreds of proposals. The first approximation that was proposed for this energy is the local spin density approximation (LSDA), proposed by Hohenberg and Kohn in their article[8]:

$$E_{\rm xc}^{\rm LSDA}[n_\uparrow, n_\downarrow] = \int {\rm d}^3r \; n(\mathbf{r}) e_{\rm xc}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})) \tag{3.1.18}$$

where $e_{\rm xc}(n_\uparrow, n_\downarrow)$ is the exchange-correlation energy per particle for an electron gas of uniform spin densities $n_\uparrow$, $n_\downarrow$. This energy density can be computed exactly and analytically for the exchange term, and with very high precision for the correlation part. The resulting approximation can in principle only be valid for systems that are close to the homogeneous electron gas. For example, simple metals. However, it has been successfully applied to other systems with surprisingly good results. Also, the LSDA is in fact a good starting point for other more sophisticated approximations. In any case, the LSDA has been used for all the DFT calculations shown below.

3.2 Time-Dependent Density Functional Theory

The aim of TDDFT is to extend the ideas of ground-state DFT in order to describe excitations, and in general any time-dependent phenomena. Here, the equation required to solve is the time-dependent Schrodinger's equation for a N-electron system:

$$i\frac{\partial}{\partial t}\psi(\mathbf{r}_1\sigma_1, ..., \mathbf{r}_N\sigma_N, t) = \hat{H}\psi(\mathbf{r}_1\sigma_1, ..., \mathbf{r}_N\sigma_N, t) \tag{3.2.1}$$

The time-dependent extension of the Hohenberg-Kohn theorem is the **Runge-Gross theorem**[10], that proves that there exists a one-to-one correspondence between the external potential and the one-electron density in the temporal evolution of the system.

In analogy with ground-state DFT, the TDDFT one-to-one correspondence permits to work with a fictitious non-interacting system of electrons, whose density is identical to the (time-dependent) density of the real one. The corresponding Kohn-Sham time-dependent equation for each Kohn-Sham state is:

$$i\frac{\partial}{\partial t}\varphi_{k\sigma}(\mathbf{r}, t) = \left[-\frac{\nabla^2}{2} + \int \frac{n(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} d^3r' + v_{\rm ext}(\mathbf{r}, t) + v_{\rm xc}(\mathbf{r}, t)\right] \varphi_{k\sigma}(\mathbf{r}, t) \tag{3.2.2}$$

$$= \hat{H}_{\rm KS}(\mathbf{r}, t)\varphi_{k\sigma} \tag{3.2.3}$$

Again, we have an xc term that must be approximated. The approximation of $v_{\rm xc}$ used by the octopus-code for these time-dependent propagations is the LSDA adiabatic version for TDDFT which assumes that the change of the exchange-correlation potential in time is adiabatic. This approach known as *adiabatic local density approximation* (ALDA) takes into consideration one density function per timestep:

$$v_{\rm xc}[n(\mathbf{r}, t)] \approx v_{xc}^{\rm gsDFT}[n_t(\mathbf{r})] \tag{3.2.4}$$

where $v_{xc}^{\rm gsDFT}$ means the exchange-correlation potential functional for ground-state DFT. This approximation was also the first to be used in TDDFT.

The temporal evolution events studied in this work are the ionization of molecules by laser pulses. Therefore, the external potential must include this external field, in addition to the

nuclear contribution. We consider strong laser fields[2] in the dipole approximation[3] $\varepsilon(t)$, and the expression (3.2.3) becomes:

$$i\frac{\partial}{\partial t}\varphi_{k\sigma}(\mathbf{r},t) = \left[-\frac{\nabla^2}{2} + \int \frac{n(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}d^3r' + \underbrace{\sum_{i=1}^{N}\sum_{\alpha=1}^{K}\frac{-Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha(t)|} + \varepsilon(t)\sum_{i=1}^{N}\mathbf{r}_i\cdot\boldsymbol{\pi}}_{V_{\text{ext}}} + v_{\text{xc}}(\mathbf{r},t)\right]\varphi_{k\sigma}(\mathbf{r},t) \tag{3.2.5}$$

$$= \hat{H}_{KS}[R(t),\mathbf{r},t]\varphi_{k\sigma} \tag{3.2.6}$$

where $\boldsymbol{\pi}$ is the unit polarization vector. Notice that the position of the nuclei $\mathbf{R}_\alpha$ is not required to be fixed – the combination of TDDFT with moving nuclei is discussed in the following section. To simplify the notation here $R(t)$ corresponds to the set of all nuclear positions $\mathbf{R}_\alpha(t)$. In the optimization of the laser that will be described below, the nuclei are considered to be fixed. However, it is necessary to consider them free to move with the purpose of observing possible fractures in the molecule.

One further issue that must be considered is the possibility of ionization of electrons. For that purpose, absorbing boundaries are placed surrounding the molecule (figure 3.1). This means that the number of particles N is not a constant of the system anymore. The absorbing boundaries consist in an absorbing potential:

$$V_{abs}(\mathbf{r}) = i\eta\sin^2\left(\frac{\mathrm{l}(\mathbf{r})\pi}{2L}\right), \tag{3.2.7}$$

being η an "absorbing height" that has to be set in order to be capable of absorbing all arriving electrons and $\mathrm{l}(\mathbf{r})$ is the distance from a point $\mathbf{r}$ to the frontier between the molecule region and the absorbing boundaries. The diameter of this sphere, initially, is chosen taking into account the size of the molecule in which the distance between the two furthest atoms is about 8 a.u and the condition that no electrons arrive to the boundaries in absence of a laser field. This last condition must be verified before the simulation. The use of absorbing boundaries is of course an approximation, but which nevertheless permits to get a rough idea of the number of electrons that are excited to the continuum.

3.3 Nuclear motion: Ehrenfest Dynamics

A laser field is able to trigger the cleavage of a molecule by any of its bonds. This cleavage implies the free movement of nuclei (or ions) with respect to other parts of the molecule, so we must deal not only with with the many-electron problem, but also with the movement of ions. One possible way to do this is the Ehrenfest dynamics, in which electrons are quantum particles while ions are considered classically, as explained in section 2. Here, the goal is to derive the equations of motion of nuclei in the framework of TDDFT, i.e. by considering the equations of motion for the system of fictitious electrons, instead of the true many-body Schrödinger equation for the real system. Fortunately, this can be done, as the force action on the ions can be written

[2] A laser whose electric field is comparable to the field felt by the electron due to the presence of the nuclei. In this regime the perturbation theory is not enough, and an explicitly time-dependent, non perturbative scheme, must be used.

[3] It is applied when the wavelength of the emitted photon is much larger than atomic dimension (≈ 1 Å). As a consequence, the factor $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$. The magnetic component can also be ignored as with non relativistic electrons $v/c << 1$

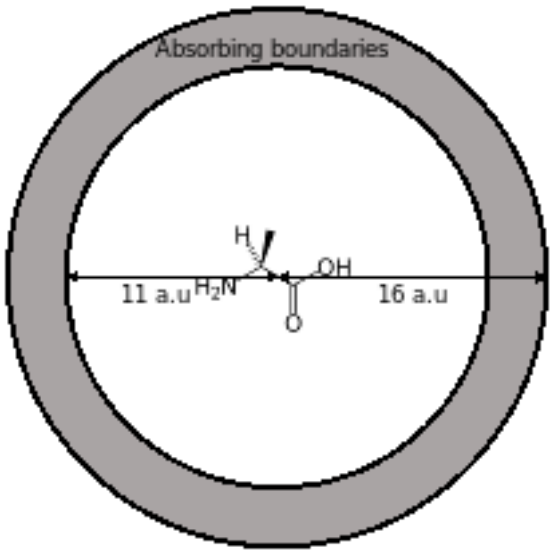

Figure 3.1: Simulation box used for the alanine molecule. The boundaries are placed in order to do not absorb in absence of laser.

in terms of the time-dependent density, or alternatively, in terms of the Kohn-Sham orbitals:

$$\mathbf{F}_\alpha[R(t),\varphi_{k\sigma}(t),t] = -\boldsymbol{\nabla}_\alpha V_{nn}(R(t)) + Z_\alpha \varepsilon(t)\boldsymbol{\pi} - \sum_{k\sigma}\langle\varphi_{k\sigma}(t)|\boldsymbol{\nabla}_\alpha \frac{Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha|}|\varphi_{k\sigma}(t)\rangle \tag{3.3.1}$$

The first term corresponds to the repulsive nucleus-nucleus coulombian interaction, with:

$$V_{\mathrm{nn}} = \sum_{\beta\neq\gamma}\frac{Z_\beta Z_\gamma}{|\mathbf{R}_\beta - \mathbf{R}_\gamma|} \tag{3.3.2}$$

The second term corresponds to the interaction of nuclei with the laser field in dipole approximation with amplitude $\varepsilon(t)$ and polarization unit vector $\boldsymbol{\pi}$. The third term is the force component due to interaction with the electron density. This interaction is written as the sum over all mean values of the coulombian interaction with respect all the KS orbitals $\varphi_{k\sigma}(t)$. These orbitals produce the solution density $n(\mathbf{r})$. In summary, the full set of equations that compose the Ehrenfest-TDDFT model are:

$$i\frac{d}{dt}\varphi_{k\sigma} = H_{\mathrm{KS}}[R(t),\mathbf{r}(t),t]\varphi_{k\sigma} \tag{3.3.3}$$

$$\dot{\mathbf{R}}_\alpha(t) = \frac{1}{M_\alpha}\mathbf{P}_\alpha(t) \tag{3.3.4}$$

$$\dot{\mathbf{P}}_\alpha(t) = \mathbf{F}_\alpha(t)[R(t),\varphi(t),t] \tag{3.3.5}$$

This system resembles to that seen in the chapter 2 (2.2.11), but with the Kohn-Sham system in place of the initial electronic Schrödinger's equation. The equations must to be solved simultaneously as they are coupled. The last two expressions are classical Newton equations, where M_α and P_α are the mass and the momentum of nucleus α, and H_{KS} in the first equation is the Kohn-Sham Hamiltonian from equation (3.2.6).

Chapter 4

Quantum Optimal Control Theory

Quantum Optimal Control Theory is the application to quantum systems of the mathematical framework called Optimal Control Theory, which deals with the problem of finding the **optimal** control function or set of parameters that maximizes or minimizes a target functional of the evolution of some *dynamical system*. A dynamical system can be defined by a set of time-dependent coordinates $y(t) \equiv y_1(t), ..., y_n(t)$, and by a dynamical law, an equation with the form:

$$\dot{y}(t) = f(y(t), u(t), t) \tag{4.0.1}$$

$$y(t_0) = y_0 \tag{4.0.2}$$

where f is the vector field which governs the dynamics. The set $u(t) \equiv u_1(t), ..., u_m(t)$ is the set of control functions, whose variation permits to modify the behaviour of the system. For a fixed value of these control functions, the knowledge of the state of the system at t_0 fully determines the behaviour of the system at $t > t_0$. Usually, the trajectory is constrained to a set of admissible trajectories $\mathcal{Y}$ such that $y \in \mathcal{Y}$, thus there exists a set of admissible controls $\mathcal{U}$ that keep the system moving into $\mathcal{Y}$. Supposing that we can set the values of controls $u \in \mathcal{U}$ and in order to get the objective trajectory, a target functional is defined as a *performance measure* that depends on the trajectory:

$$J[y(t), u(t)] = h(y(T)) + \int_{t_0}^{T} \mathrm{d}t \ \mathcal{L}(y(t), u(t), t) \tag{4.0.3}$$

This functional is separated into a part $h(y(T))$ that depends on the state at the final time T which can be fixed or not (sometimes it is wanted to be minimized also), and another time-dependent Lagrangian part $\int_{t_0}^{T} \mathrm{d}t \ \mathcal{L}(y(t), u(t), t)$ that depends on the full evolution of the system. In our case we will focus on the case in which the final time T is fixed. Into this framework, the *optimal control problem* can be formulated as the problem of finding the optimal set of controls $u* \in \mathcal{U}$ that drives the system in an optimal admissible trajectory $y^* \in \mathcal{Y}$.

$$\dot{y}^*(t) = f(y^*(t), u(t), t) \tag{4.0.4}$$

$$y^*(t_0) = y_0^* \tag{4.0.5}$$

such that the target functional is maximized, i.e.:

$$J[y^*(t), u^*(t)] = h(y^*(T)) + \int_{t_0}^{T} \mathrm{d}t \ \mathcal{L}(y^*(t), u^*(t), t) \geq J[y(t), u(t)] \tag{4.0.6}$$

for any $y(t), u(t)$. The necessary conditions for the controls $u(t)$ and the state trajectory were established by the Pontryagin's maximum[1] principle[24]. Given the *performance measure* (4.0.3)

[1] More frequently it is found as Pontryagin's minimum principle. Obviously the two problems are equivalent, as one transforms the minimization problem into the maximization one by simply changing the sign of the performance measure.

and the dynamical system described by (4.0.4) and (4.0.5), a Hamiltonian function (not to be mistaken with the Hamiltonian of the system, if it exists) is defined as:

$$\mathcal{H}(y(t), u(t), p(t), t) = \mathcal{L}(y(t), u(t), t) + p^{\mathrm{T}}(t) \cdot f(y(t), u(t), t) \tag{4.0.7}$$

where $\mathcal{L}$ is the Lagrangian and f the dynamical law defined above. The new set of defined variables $p(t) = p_1, ..., p_n$ is called *costate* or *momentum variable*. According to Pontryagin's principle[24], the necessary conditions for $u^*(t)$ and $y^*(t)$ for maximizing the *performance measure* with a fixed final time T are:

$$\dot{y}^*(t) = \nabla_p \mathcal{H}(y^*(t), u^*(t), p^*(t), t) \tag{4.0.8}$$

$$y^*(t_0) = y_0 \tag{4.0.9}$$

$$\dot{p}^*(t) = -\nabla_y \mathcal{H}(y^*(t), u^*(t), p^*(t), t) \tag{4.0.10}$$

$$p^*(T) = \nabla_y h(y(T)) \tag{4.0.11}$$

$$\mathcal{H}(y^*(t), u^*(t), p^*(t), t) \geq \mathcal{H}(y^*(t), u(t), p^*(t), t) \ \forall u \in \mathcal{U} \tag{4.0.12}$$

The initial condition for $p(t)$ (4.0.11) is given at the final time T, that implies that it will have to be propagated *backwards*.

The previous equations are general, for any type of dynamical system. Consider now in particular a quantum system such as the N electron system described below and its temporal evolution of the state $\psi(t)$ in the interval $t \in [0, T]$, driven by the action of a strong laser field $\varepsilon(t)$ that is our *control function*. The "dynamical law" that governs the dynamics is the time-dependent Schrödinger's equation:

$$f(\psi(t), \varepsilon(t), t) = -iH[\varepsilon(t)]\psi(t) \tag{4.0.13}$$

hence

$$i\frac{d}{dt}|\psi(t)\rangle = \hat{H}[\varepsilon(t)]|\psi(t)\rangle \tag{4.0.14}$$

The Hamiltonian $\hat{H}$ can be usually split assuming the dipole approximation:

$$\hat{H}[\varepsilon(t), t] = \hat{H}_0 + \varepsilon(t)\hat{D} \tag{4.0.15}$$

where H_0 is the Hamiltonian in absence of field, whereas $\hat{D}$ is the dipole coupling term to the electric field component of the laser. The *optimal control problem* in the case of this kind of quantum system consists in finding the **optimal laser pulse** $\varepsilon(t)$ which drives the system in such a way that the performance measure is optimized.

Rather then working with a continuous function $\varepsilon(t)$, it is numerically necessary to use a discretization, i.e. to describe the function with a set of parameters. The laser pulse can be described in a suitable way by its Fourier expansion in terms of the basis coefficients ε_i accompanying orthonormal basis functions (sin, cos) in a period T:

$$\varepsilon(t) = \frac{\varepsilon_0}{2} + \sum_{k=1}^{r/2} \left[\varepsilon_{2k} \sqrt{\frac{2}{T}} \cos \frac{2(2k)\pi}{T} t + \varepsilon_{2k-1} \sqrt{\frac{2}{T}} \sin \frac{2(2k-1)\pi}{T} t \right] \tag{4.0.16}$$

where t is the time, T is the period of one pulse and r is the number of degrees of freedom (minus one, corresponding to ε_0). However, these coefficients $\varepsilon_0, \varepsilon_1, ..., \varepsilon_r$ are limited in number

by the desired cut-off frequency ω_{max} with $\omega_k = \frac{2k\pi}{T}$. In addition, in order to represent a real physical laser, the function $\varepsilon(t)$ must satisfy the following conditions (or constraints):

$$\int_T \varepsilon(t)dt = 0 \tag{4.0.17}$$

$$\varepsilon(0) = \varepsilon(T) = 0 \tag{4.0.18}$$

The first condition follows from Maxwell's equations for a freely propagating pulse in dipole approximation, and allows us to remove the term ε_0 from (4.0.16). The second condition ensures that the pulse starts and finish in zero.

There is a further constraint that we will add to the optimization search. The energy per unit of area of a laser pulse is called the *fluence*, which integrates the intensity over the whole period:

$$\mathcal{F}[\varepsilon] = \int \varepsilon^2(t)\mathrm{d}t \tag{4.0.19}$$

As our laser pulse is described by a series of orthonormal functions, the expression for the fluence can also be written as:

$$\mathcal{F}[\varepsilon] = \sum_{i=1}^{r} \varepsilon_k^2 \tag{4.0.20}$$

In a real experiment, the intensity of a laser cannot be increased at will. There is usually a constraint over the maximum intensity, or over the total energy carried by the pulse, i.e. the fluence. In QOCT works, normally the goal is to maximize the performance measure, while at the same time maintain or minimize the fluence. In the runs below, the choice was to do a search over pulses with equal fluence. This condition implies a new constraint. In total, the number of degrees of freedom with the three conditions described above will be $r+1-3$ constraints $= r-2$.

The *performance measure* functional introduced above (4.0.3) gathers the processes of optimization of the control function $u(t) = \varepsilon(t)$ in form of mathematical expression. In this case we will be concerned with the case in which it depends only on the state at the end of the pulse. For example, suppose we wish to carry the system to some final state ψ_f; the performance measure could be the projection onthe the subspace spanned by that state, i.e.:

$$J[\psi] = \langle\psi(T)|\underbrace{\psi_f\rangle\langle\psi_f|}_{\hat{O}}\psi(T)\rangle \tag{4.0.21}$$

that must be maximized. It can be seen as the expected value of an operator $\hat{O}$ or a projection on the target state. In addition to this functional, that encodes somehow the task that the control laser pulse must do, it is frequent (although we have not used it), to add a penalty function that avoids region of parameter space that are not desirable. For example, one may attempt to optimize the previous target while minimizing the fluence of the pulse, and in order to do that one may add a second term, ε-dependent, to the functional:

$$J[\psi, \varepsilon] = \langle\psi(T)|\underbrace{\psi_f\rangle\langle\psi_f|}_{\hat{O}}\psi(T)\rangle - \alpha\mathcal{F}[\varepsilon]\,, \tag{4.0.22}$$

where the penalty factor α is positive, and in this way prevents the fluence's value to become very large. However, in the runs shown below, the fluence is fixed by construction, and therefore there is no need to use this term.

Once a set of control parameters ε has been set, the value of this functional is determined by the evolution of the system: $\varepsilon \to \psi[\varepsilon] \to J[\psi[\varepsilon], \varepsilon]$. Therefore, we may define a function of the control parameters:

$$G[\varepsilon] = J[\psi[\varepsilon], \varepsilon]\,. \tag{4.0.23}$$

The problem therefore reduces to the optimization of this functional.

One may then apply Pontryagin's principle; the key result is the gradient of function G, which turns out to be given by:

$$\frac{\partial G}{\partial \varepsilon_m}(\varepsilon) = \frac{\partial J}{\partial \varepsilon_m}[\psi, \varepsilon]\bigg|_{\psi=\psi[\varepsilon]} + 2\mathrm{Im}\int_0^T \mathrm{d}t\ \left[\langle\chi(t)|\frac{\partial \hat{H}}{\partial \varepsilon_m}|\psi(t)\rangle\right] \tag{4.0.24}$$

The *costate* χ is a new auxiliary wave function that must be computed in order to get the gradient; the equation of motion for χ is:

$$i\frac{\partial}{\partial t}|\chi(t)\rangle = \hat{H}^\dagger|\chi(t)\rangle \tag{4.0.25}$$

$$|\chi(T)\rangle = \hat{O}|\psi(T)\rangle \tag{4.0.26}$$

Note the boundary value condition at the final time T, instead of at the initial time. Otherwise, these equations are identical to the one for the real wave function ψ. The expression (4.0.24), together with (4.0.25) and the equation of motion for ψ are usually referred as the *control equations*.

It remains to define the performance measure or target functional. In our particular case, the goal is to optimize a laser pulse of a fixed fluence (4.0.20) in order to make it to pull out as many electrons as possible from a system (ultimately, the L-alanine molecule, see figure 6.1). These electrons will be absorbed by a complex absorbing potential when they reach a certain distance from the center of the simulation box. The optimization will require changes of the frequency weights ε_i of the laser in order to get the characteristic frequencies of the molecule, but maintaining the initial fluence, so that the penalty function mentioned above is not needed anymore. The functional J can be defined in terms of the electron density since we wish to perform the calculations with TDDFT, and this is the only object that the theory truly provides (the real many-electron wave function is not provided by the theory). With this in mind, the target operator $\hat{O}$ is given by:

$$\hat{O} = -\int \mathrm{d}^3r\ \hat{n}(\mathbf{r}) \tag{4.0.27}$$

Therefore, the functional J is:

$$J[\psi] = \langle\psi(T)|\hat{O}|\psi(T)\rangle = -\int \mathrm{d}^3r\ n_T(\mathbf{r}) \tag{4.0.28}$$

where $n_T(\mathbf{r})$ is the electron density at the end of the pulse. In other words, by maximizing this functional we are maximizing the electronic leak or minimizing the final electron density that remains in the simulation box. Note that I do not include any term related to nuclei, as I only want to maximize the electronic escape. The movement of ions will be studied later in order to observe if the obtained optimized laser affects to the molecule's atomic structure.

Chapter 5

Optimization algorithms used for the Quantum Optimal Control calculations

The software used (the octopus code) can apply several optimization algorithms to find the maximum of the functions defined in the previous section. Numerically, they require series of propagations for ψ (in order to compute the value of the function) and χ (if the gradient is required).

The optimization algorithms can be divided into derivative-based and derivative-free (or gradient-based and free) methods. An important fact about this difference is that derivative methods must do *forward-backward* propagations per evaluation of J according to the initial condition at time T (equation 4.0.11) in order to obtain first the state χ at final time and after the *costate* ψ at the initial time. Derivative-free methods such as the BOBYQA scheme described below only require one unique *forward* propagation per iteration as they do not make use of gradient-involving equations, but rather make use of some interpolation processes in order to estimate the gradient.

In this work, the calculations have made use of the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm[25][26][27][28], in particular on its low-storage version (L-BFGS), that is a derivative-based quasi-Newton scheme; and on the BOBYQA (Bound Optimization BY Quadratic Approximation) algorithm[[29]] which is derivative-free. These methods will be shortly described here without too much detail.

5.1 The L-BFGS algorithm

The objective is to find the optimal control function $\varepsilon(t)$, parameterized by a set of r degrees of freedom, $\varepsilon(t) \equiv \boldsymbol{\varepsilon}$, which maximizes the functional $J(\varepsilon_1, ..., \varepsilon_r)$. The iteration procedure used by this algorithm is given in terms of the Hessian matrix, composed by the second derivatives (the sub-indices denote the iteration number):

$$\boldsymbol{\varepsilon}_{k+1} = \boldsymbol{\varepsilon}_k - \underbrace{\left(\nabla^2 J[\boldsymbol{\varepsilon}_k]\right)^{-1}}_{H_k^{-1}} \underbrace{\left(\boldsymbol{\nabla} J[\boldsymbol{\varepsilon}_k]\right)}_{g_k} \tag{5.1.1}$$

where $\boldsymbol{\varepsilon}_k$ refers to the vector $\boldsymbol{\varepsilon}$ in the k-th iteration. In quasi-Newton methods the inverse of Hessian matrix must be approximated due to the complexity of its calculus, in particular in the

BFGS scheme for the k-th iteration:

$$H_{k+1}^{-1} = H_k^{-1} - \frac{\boldsymbol{\delta}_k \boldsymbol{\gamma}_k^{\mathrm{T}} H_k^{-1} + H_k^{-1} \boldsymbol{\gamma}_k \boldsymbol{\delta}_k^{\mathrm{T}}}{\boldsymbol{\delta}_k^{\mathrm{T}} \boldsymbol{\gamma}_k} + \left(1 + \frac{\boldsymbol{\gamma}_k^{\mathrm{T}} H_k^{-1} \boldsymbol{\gamma}_k}{\boldsymbol{\delta}_k^{\mathrm{T}} \boldsymbol{\gamma}_k}\right) \frac{\boldsymbol{\delta}_k^{\mathrm{T}} \boldsymbol{\delta}_k}{\boldsymbol{\delta}_k^{\mathrm{T}} \boldsymbol{\gamma}_k} \tag{5.1.2}$$

where

$$\boldsymbol{\delta}_k = \boldsymbol{\varepsilon}_{k+1} - \boldsymbol{\varepsilon}_k$$
$$\boldsymbol{\gamma}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$$

This recursion for the inverse of Hessian matrix is shared by BFGS and L-BFGS. We can realise that the relevant term in (5.1.1) is $\mathbf{d}_k = H_k^{-1}\mathbf{g}_k$ which could be seen as the search direction gradient. It is assumed here that we have stored the m last updates of $\boldsymbol{\delta}_k$ and $\boldsymbol{\gamma}_k$. The steps to obtain the search direction in each iteration are:

- Define a sequence of vectors $\mathbf{q}_{k-m}, ..., \mathbf{q}_k$ where $\mathbf{q}_k = \mathbf{g}_k$ and $\mathbf{q}_i = \left(I - \frac{\boldsymbol{\gamma}_i \boldsymbol{\delta}_i^{\mathrm{T}}}{\boldsymbol{\gamma}_i^{\mathrm{T}} \boldsymbol{\delta}}{}_i\right) \mathbf{q}_{i+1}$. Define also $\alpha_i := \frac{\boldsymbol{\delta}_i}{\boldsymbol{\gamma}_i^{\mathrm{T}} \boldsymbol{\delta}_i} \mathbf{q}_{i+1}$ such that we can obtain recursively $\mathbf{q_i} = \mathbf{q}_{i+1} - \alpha_i \boldsymbol{\gamma}_i$.

- Define $\mathbf{z}_{k-m}, ..., \mathbf{z}_k$ as $\mathbf{z}_i = H_i^{-1}\mathbf{q}_i$. Define $\mathbf{z}_{k-m} = H_{k,0}^{-1}\mathbf{q}_{k-m}$ and $\beta_i := \frac{\boldsymbol{\gamma}_i^{\mathrm{T}}}{\boldsymbol{\gamma}_i^{\mathrm{T}} \boldsymbol{\delta}_i} \mathbf{z}_i$ as $\mathbf{z}_{i+1} = \mathbf{z}_i + (\alpha_i - \beta_i)\boldsymbol{\delta}_i$.

- The algorithm is:

  $\mathbf{q} = \mathbf{g}_k$
  For $i = k-1, k-2, ..., k-m$
  $\quad \alpha_i = \frac{\boldsymbol{\delta}_i}{\boldsymbol{\gamma}_i^{\mathrm{T}} \boldsymbol{\delta}_i} \mathbf{q}$
  $\quad \mathbf{q} = \mathbf{q} - \alpha_i \boldsymbol{\gamma}_i$
  $H_{k,0}^{-1} = \frac{\boldsymbol{\gamma_{k-1}} \boldsymbol{\delta}_{k-1}^{\mathrm{T}}}{\boldsymbol{\gamma}_{k-1}^{\mathrm{T}} \boldsymbol{\gamma}_{k-1}}$
  $\mathbf{z} = H_{k,0}^{-1}\mathbf{q}$
  For $i = k-m, k-m+1, ..., k-1$
  $\quad \beta_i = \frac{\boldsymbol{\gamma}_i^{\mathrm{T}}}{\boldsymbol{\gamma}_i^{\mathrm{T}} \boldsymbol{\delta}_i} \mathbf{z}$
  $\quad \mathbf{z} = \mathbf{z} + \boldsymbol{\delta}_i(\alpha_i - \beta_i)$
  Stop with $H_k \mathbf{g}_k = \mathbf{z}$

Notice that as we are maximizing, the matrix $H_{0,k}^{-1}$ should be negative definite. To summarize, the algorithm consists in a series of backward-forward propagations getting in first place the Hessian matrix by backward propagation from the last state and then doing a forward propagation to calculate the search direction. When the decrement in the functional is small enough such that $\Delta J < \epsilon$, being ϵ a threshold parameter decided by the user, the code stops giving the optimal parameters $(\varepsilon_1^*, ..., \varepsilon_r^*)$.

5.2 The BOBYQA algorithm

The BOBYQA method involves a quadratic approximation $Q[\boldsymbol{\varepsilon}]$ to the target functional $J[\boldsymbol{\varepsilon}]$ such that in the k-th iteration of the optimization process:

$$Q_k[\boldsymbol{\theta}_j] = J[\boldsymbol{\theta}_j];\ j = 1, 2, ..., m \tag{5.2.1}$$

The parameters $(\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_m)$ are the interpolation conditions and the integer m is selected from the set $[r+2, \frac{1}{2}(r+1)(r+2)]$. We let $\boldsymbol{\varepsilon}$ to be the parameter vector in the set $[\boldsymbol{\theta}_j]$, $j = 1, ..., m$ such that:

$$J[\boldsymbol{\varepsilon}_k] = \min\{J[\boldsymbol{\theta}_j] : j = 1, 2, ..., m\} \tag{5.2.2}$$

giving priority to an earlier evaluation of the least function value $J[\boldsymbol{\varepsilon}_k]$. A positive number Δ_k, called "trust region radius", is also available at the beginning of the k-th iteration. If the convergence criteria are satisfied then the termination occurs on the k-th iteration. Otherwise, a step $\mathbf{d}_k$ is calculated such that $||\mathbf{d}_k|| < \Delta_k$ and $\boldsymbol{\varepsilon}_{k+1} = \boldsymbol{\varepsilon}_k + \mathbf{d}_k$. These new values satisfy the restrictions such as ω_{max} or the fluence and do not belong to the set $\boldsymbol{\theta_1, ..., \theta_m}$. The process can be summarized as:

$$\boldsymbol{\varepsilon}_{k+1} = \begin{cases} \boldsymbol{\varepsilon}_k & J[\boldsymbol{\varepsilon}_k + \mathbf{d}_k] \leq J[\boldsymbol{\varepsilon}_k] \\ \boldsymbol{\varepsilon}_k + \mathbf{d}_k & J[\boldsymbol{\varepsilon}_k + \mathbf{d}_k] > J[\boldsymbol{\varepsilon}_k] \end{cases} \tag{5.2.3}$$

Then, one of the points $\boldsymbol{\theta_t}$ is replaced by $\boldsymbol{\varepsilon}_k + \mathbf{d}_k$. The values Δ_{k+1} and Q_{k+1} are generated for the next iteration and the process continues until the criterion of convergence is fulfilled.

Chapter 6

Results

We have used for this work the alanine molecule as an example, specifically the L-alanine isomer. I chose this molecule because it is the second smallest amino acid, so that the simulations are lighter, and also by its great importance in metabolism. The 3D structure was downloaded from the NIST database[1]. The first step was the calculation of the ionization potential, that is, the energy barrier to surpass by the most external electron in order to escape from the molecule. That is done by running a ground state DFT calculation with and without one electron and taking the difference in the total energy:

$$V_{\text{ionization}} = E(N-1) - E(N) = 0.351005\ a.u \tag{6.0.1}$$

This energy corresponds to 9.55 eV, an energy that is supposed to be big enough to keep alive the biological functionality of the molecule in not very extreme conditions (it is relatively hard to ionize this molecule, and therefore it is stable in this respect). Apart from the ionization potential, it is worth to know the intermediate molecular states from which electrons can "jump" to and from. They should be those between the HOMO (highest occupied molecular orbital) energy value, and the value of the HOMO plus the ionization potential. In fact, this should be zero, as the HOMO energy in exact DFT should exactly correspond to the ionization potential. But, as table (6.1) shows, the HOMO's energy does not correspond to the ionization potential in the DFT calculation. This is a particularity of DFT when using LDA or GGA approximations as it is collected in literature[30]: The HOMO level is upshifted from the true K-S value. Then, the intermediate molecular states would be those comprised in the interval [HOMO,0].

molecular level	**energy [a.u]**	**required energy [a.u]**
HOMO	-0.217917	0
LUMO	-0.045996	0.171921
unocc. 2	-0.035496	0.182421
unocc. 3	-0.009966	0.207951
unocc. 4	-0.001152	0.216765

Table 6.1: Intermediate molecular levels of the alanine molecule according DFT calculations.

In order to adjust carefully the parameters of the simulation –grid, time-step, time-pulse, box-size and box-shape, etc–, I made a very big approximation considering first a one-dimensional model of alanine. The model consists in one particle subjected to a soft-Coulomb potential:

$$v(x) = \frac{1}{\sqrt{\alpha^2 + x^2}} \tag{6.0.2}$$

[1] `https://cccbdb.nist.gov/geom3x.asp?method=13&basis=6`

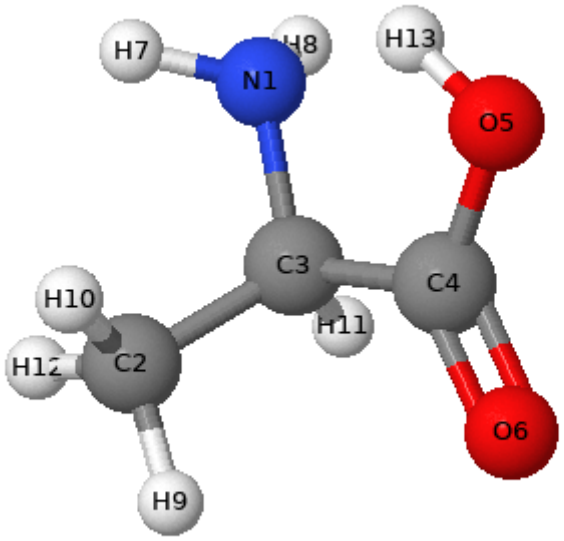

Figure 6.1: 3D diagram of the alanine molecule. Experimental coordinates from cccbdb.nist.gov.

where α is chosen to $v(x)$ be equal to the ionization potential (6.0.1). This value was obtained through a DFT calculation, but also could had been obtained from experimental data in the literature (from the website `cccbdb.nist.gov.`). In the DFT calculation of the 1D approximation, the value of the HOMO's energy coincided with the ionization potential. That reflects the true K-S HOMO energy, when the upshifting of LDA approximation is not present because there is only one particle. Still, an important issue that allows to overcome this difference is that the HOMO-LUMO gap almost coincides in both cases (table 6.2). The holding of this gap is the essential physical property for us, so we will not be interested in trying to give a complete physical meaning to K-S orbitals, that is still difficult according to the bibliography[31].

molecular level	**energy [a.u]**	**required energy [a.u]**
HOMO	-0.351005	0
LUMO	-0.186138	0.164867
unocc. 2	-0.112834	0.238171
unocc. 3	-0.069743	0.281262
unocc. 4	-0.031206	0.319799

Table 6.2: Intermediate molecular levels of the one dimensional model.

Beyond these energies the electrons should be absorbed by boundaries.

6.1 One-dimensional model

The implementation of the model in octopus-code consists into the setting of several parameters:

- Shape and size of the box: In this case it is a segment of 32 a.u length including absorbing boundaries of 5 a.u. See section 3.3 for further information about the boundaries.
- Grid's spacing and timestep: They are related as a division by two in spacing means a multiplication by two in amount of time. In 3D the multiplication is by eight. Here the spacing is 0.2 a.u and the time-step is 0.005 a.u.

- Potential: Here it is the user-defined soft-coulomb potential (6.0.2) with $\alpha = 4.53775$.
- Initial coordinates and laser field: Initial position in x=0 and an envelope sine function as the initial input laser field. It is randomized after.
- Optimization parameters such as the cut-off frequency (maximum component of frequency) of the final laser and the target functional which is the one described in chapter 4.

In principle, it is not a bad idea to set the cut-off frequency of the laser $\omega_{\max}$ in 0.5 $a.u$ that is greater than the ionization potential. This is because the spectrum of the optimized pulse is expected to have a big peak in sections corresponding to the excitation or resonance frequencies of the molecule. Moreover, a higher value of this parameter could cause the impossibility to realize the laser experimentally reaching the X-ray zone (≈ 1 $a.u$). The total time that will be considered is $T = 200$ a.u, that is of the order of *femtoseconds* (fs) and corresponds to the called ultrashort laser pulses.

The first step is to ensure that in absence of a laser field, a negligible part of the charge can be absorbed by the boundaries (figure 6.2). However, though it is negligible, a very small fraction of electronic charge reaches the boundaries. The explanation is that the electronic states are physically probability distributions whose tails are extended to infinite, so the probability of finding an electron in any point of space never turns to zero as there. Therefore, there is always a small tail of the distribution that can be observed because we are considering a finite simulation box.

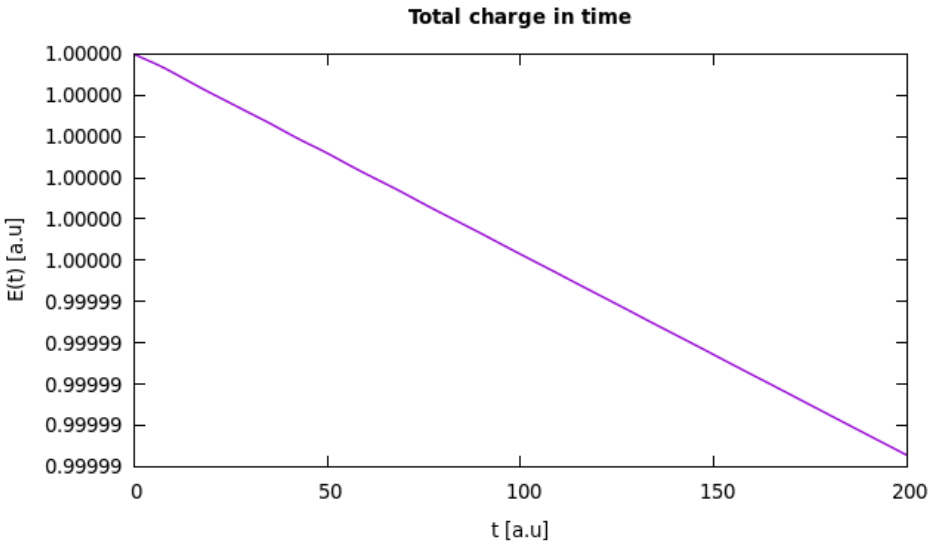

Figure 6.2: Loss of charge in absence of a laser field. This plot ensures us that the absorbing boundaries are well placed because electrons cannot reach them without external energy.

After this, three different random-pulses are generated holding the fluence in order to see if different initial conditions reach different maxima of the functional J. With $\omega_{\max} = 0.5$ we have 30 basis parameters $(\varepsilon_1, ..., \varepsilon_{30})$ that means $30-2 = 28$ degrees of freedom, while for $\omega_{\max} = 0.25$ the number of degrees of freedom is $14 - 2 = 12$. The first algorithm used to optimize the pulse is L-BFGS.

Setting a low initial fluence of 0.0225 (approximated intensity of 0.12 a.u) and $\omega_{\max} = 0.5$ (figure 6.4) is almost enough to get optimized pulses that are capable of pulling out almost the totality of the charge. To free the totality of electrons the fluence is increased until 0.04 a.u (approximated intensity of 0.22 a.u) (figure 6.5).

The frequency spectrum of the optimized laser pulse in both cases shows the presence of a big peak (or two) at the LUMO frequency and several of them in the lowest frequency region of the spectre. These frequencies are below the ionization energy of the system, so a possible explanation of this is a "multistep-ionization". Thanks to the big component of the laser in the HOMO-LUMO gap frequency the electrons receive the necessary energy in order to reach the LUMO. After that, the electrons must receive the necessary energy to jump between the rest of higher orbitals until arriving the ionization energy, and therefore managing to escape. Another simpler explanation could be that, being the ionization potential calculated in the 3D model, the results can differ notoriously as we will see in the 3D model QOCT calculation.

Another thing to underline is that, in these two first graphs, two of the lasers have reached the same solution. In the first one they arrived exactly to the same laser while in the second the purple laser is a mirror symmetric of the green laser. It is noted the importance of randomizing the initial conditions in order to reach different local minima. With this method and enough random starting points one can embrace all the space of minima and find the absolute minimum.

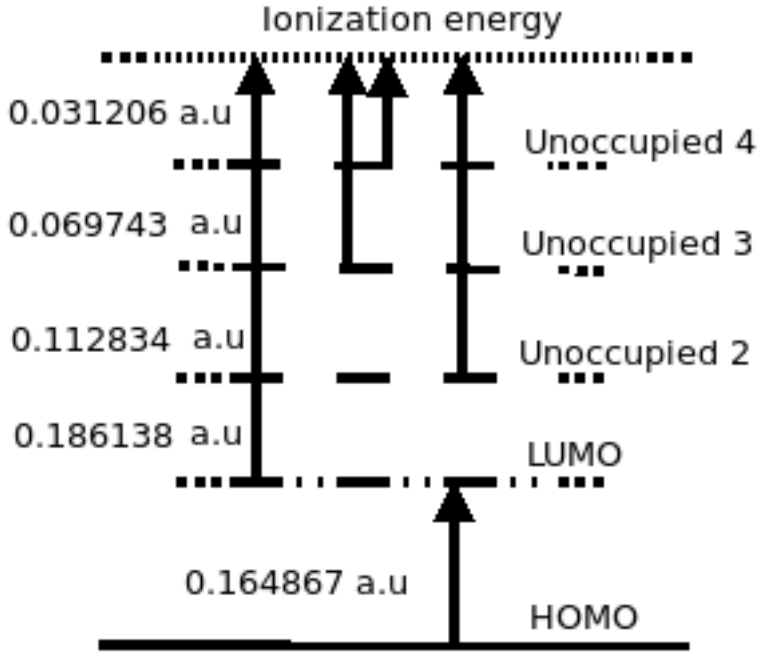

Figure 6.3: Orbital energy levels of the one dimensional model.

Additionally, the spectrum region $0.25 < \omega \leq 0.5$ is empty that means that there is no relevant contributions from these frequencies, so this region can be neglected conferring more agility to the calculus as it cuts in half the number of degrees of freedom but maintaining the same performance (figure 6.6). If ω_{max} is set to do not be higher than 0.1 $a.u$, several spectral components are not included in the optimization and the number of degrees of freedom is reduced to $6-2=4$, but as can it be seen (figure 6.7), the optimization is hard using only with this part of the spectrum. The results suggest us to set the threshold frequency in 0.25 $a.u$ ($\sim 1 \cdot 10^{16}$ Hz) as it can keep all of the information.

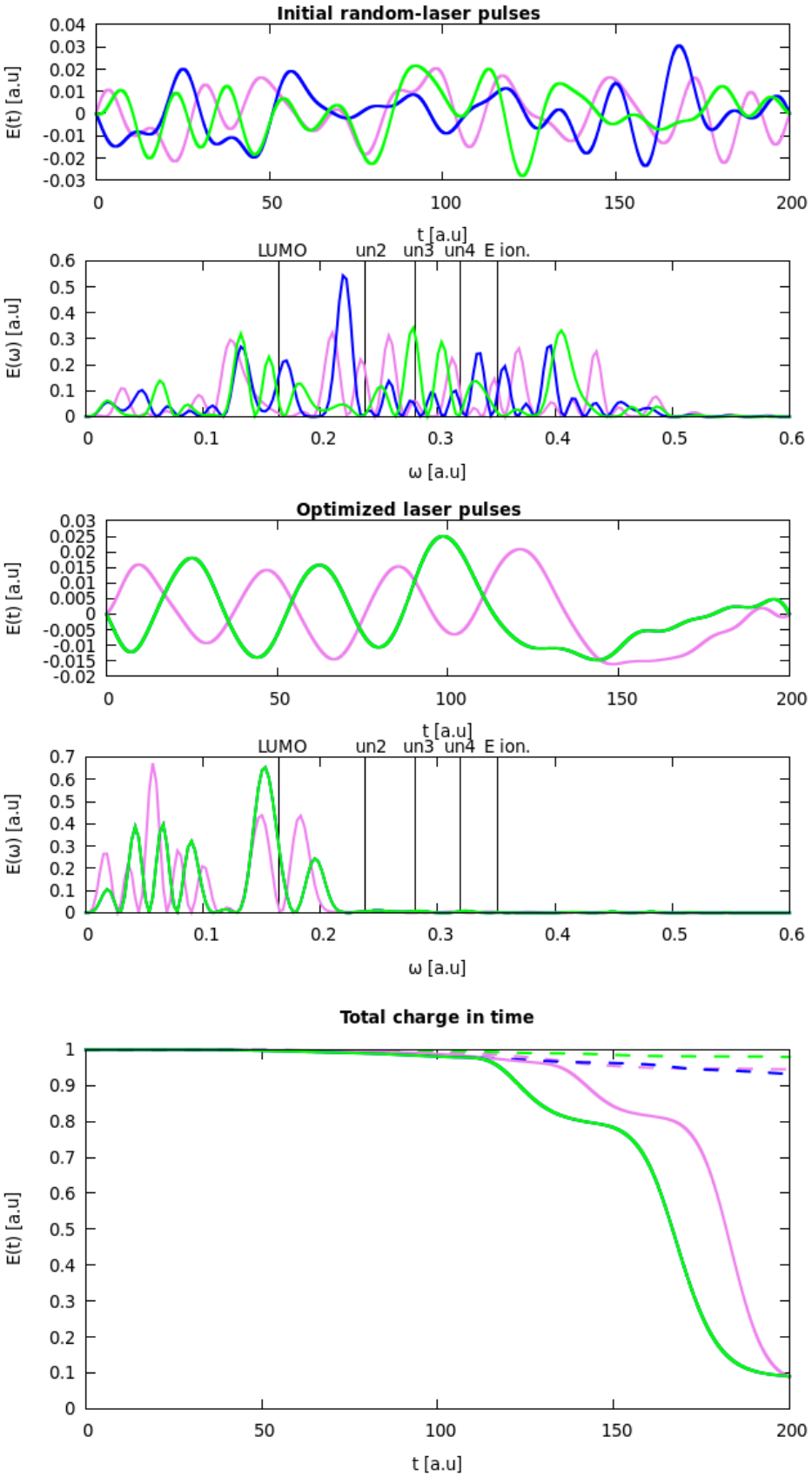

Figure 6.4: ***Top****: three different random initial lasers with fluence = 0.02250 and* $\omega_{max} = 0.5$ *a.u. Upper half: time-domain. Lower half: frequency-domain.*
Middle*: optimized lasers with the L-BFGS method, holding the initial fluence. Upper half: time-domain. Lower half: frequency-domain.*
Bottom*: loss of charge induced by the initial and optimized lasers. Lines of the same color mean loss of charge induced by the same initial laser. Dashed line: initial laser. Solid line: optimized laser.*

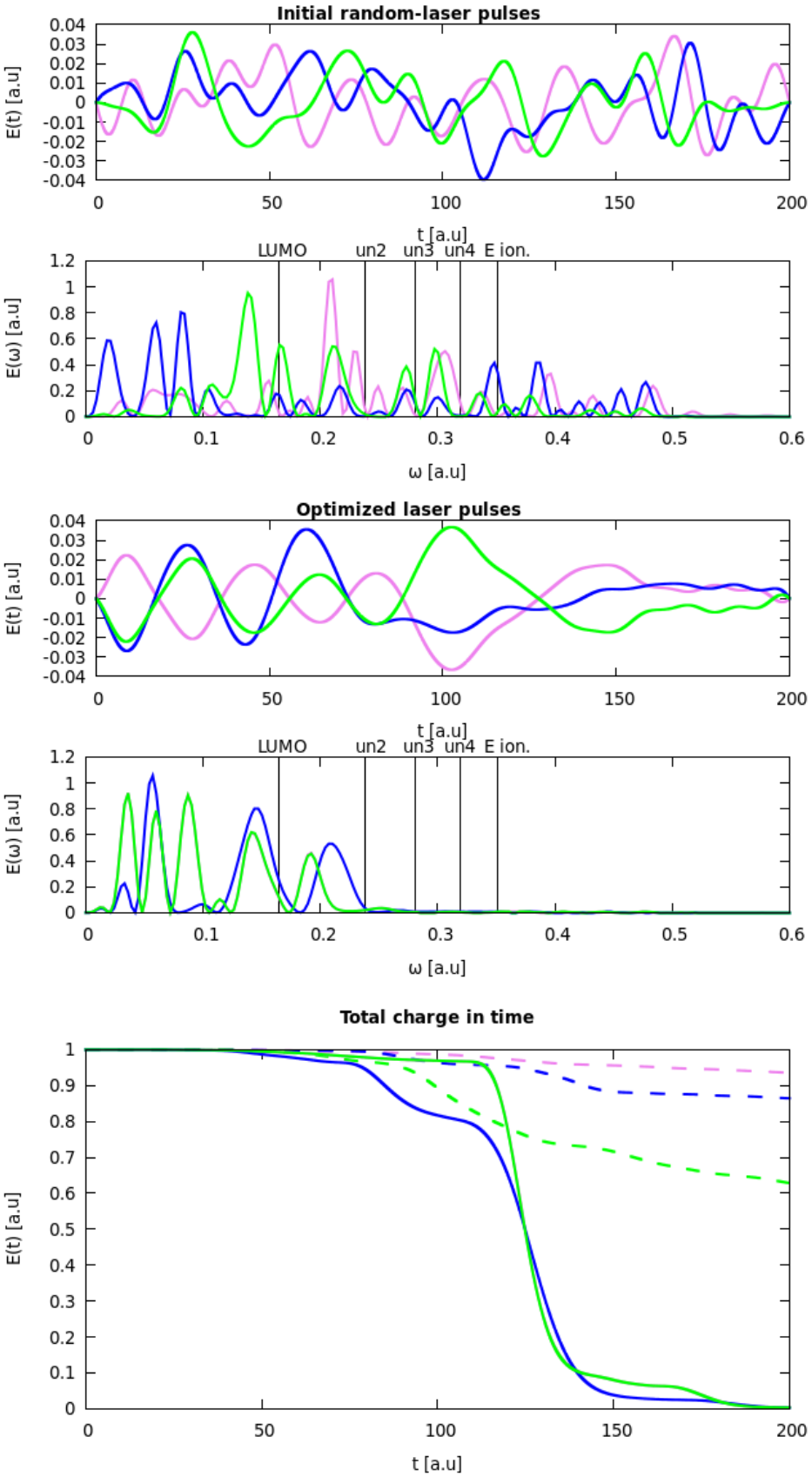

Figure 6.5: ***Top****: three different random initial lasers with fluence = 0.04000 and* $\omega_{max} = 0.5$ *a.u. Upper half: time-domain. Lower half: frequency-domain.*
Middle*: optimized lasers with the L-BFGS method, holding the initial fluence. Top: time-domain. Bottom: frequency-domain.*
Bottom*: loss of charge induced by the initial and optimized lasers. Lines of the same color mean loss of charge induced by the same initial laser. Dashed line: initial laser. Solid line: optimized laser.*

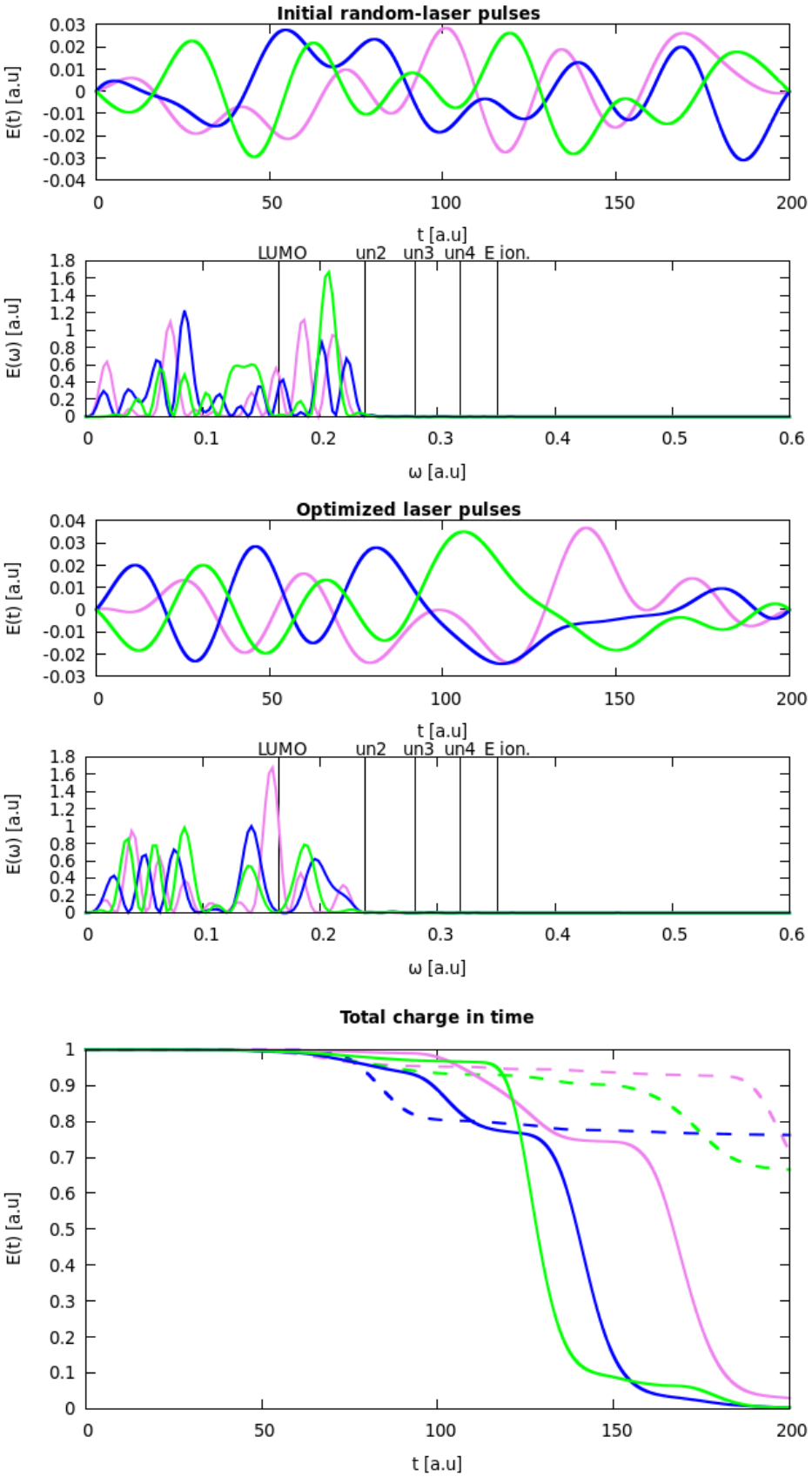

Figure 6.6: ***Top****: three different random initial lasers with fluence = 0.04000 and* $\omega_{max} = 0.25$ *a.u. Upper half: time-domain. Lower half: frequency-domain.*
Middle*: optimized lasers with the L-BFGS method, holding the initial fluence. Upper half: time-domain. Lower half: frequency-domain.*
Bottom*: loss of charge induced by the initial and optimized lasers. Lines of the same color mean loss of charge induced by the same initial laser. Dashed line: initial laser. Solid line: optimized laser.*

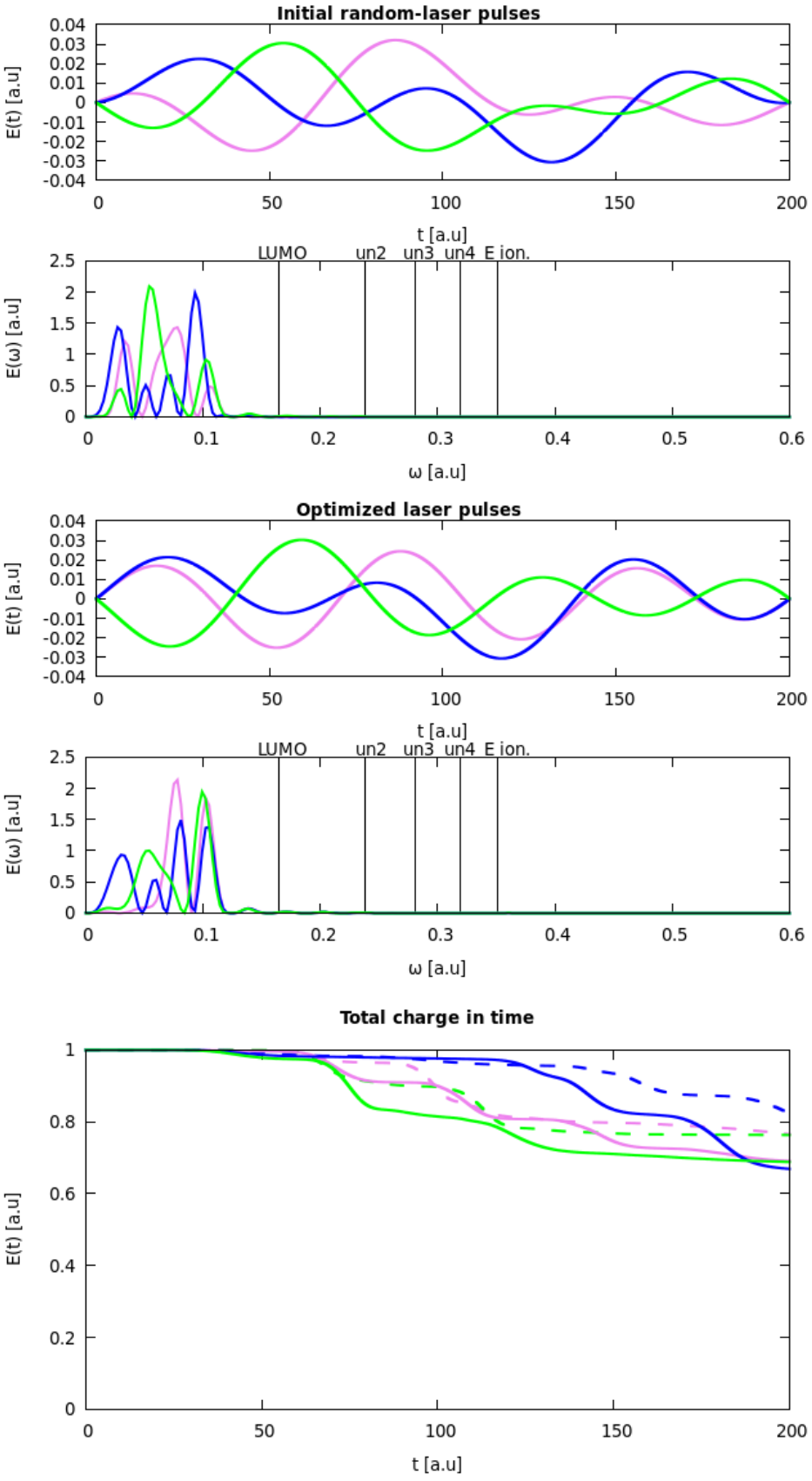

Figure 6.7: ***Top****: three different random initial lasers with fluence = 0.04000 and* $\omega_{max} = 0.1$ *a.u. Upper half: time-domain. Lower half: frequency-domain.*
Middle*: Optimized lasers with the L-BFGS method, holding the initial fluence. Upper half: time-domain. Lower half: frequency-domain.*
Bottom*: Loss of charge induced by the initial and optimized lasers. Lines of the same color mean loss of charge induced by the same initial laser. Dashed line: initial laser. Solid line: optimized laser.*

The second algorithm to be used is the BOBYQA derivative-free. The motivation for this is to compare its optimization power and number of iterations with L-BFGS. In all cases, the algorithm achieves the same optimization objectives as L-BFGS, so to summarise the results, only the case of $\omega_{\max} = 0.25$ $a.u$ is shown (figure 6.9). In order to compare both algorithms, it is a good idea to count the number of iterations required to converge for both of them. This can serve as a benchmark rather than the computational time that depends on the CPU. In the case of BFGS the *backward-forward* propagation is considered as two iterations, so we count the number of evaluations by algorithm.

The plot for $\omega_{\max} = 0.25$ (figure 6.9) shows that the derivative algorithm L-BFGS is more efficient as it needs less evaluations than BOBYQA to converge. For $\omega_{\max} = 0.5$ (figure 6.10), the difference in the convergence speed becomes more tangible, an indication that the BOBYQA algorithm is more affected by an increment of degrees of freedom.

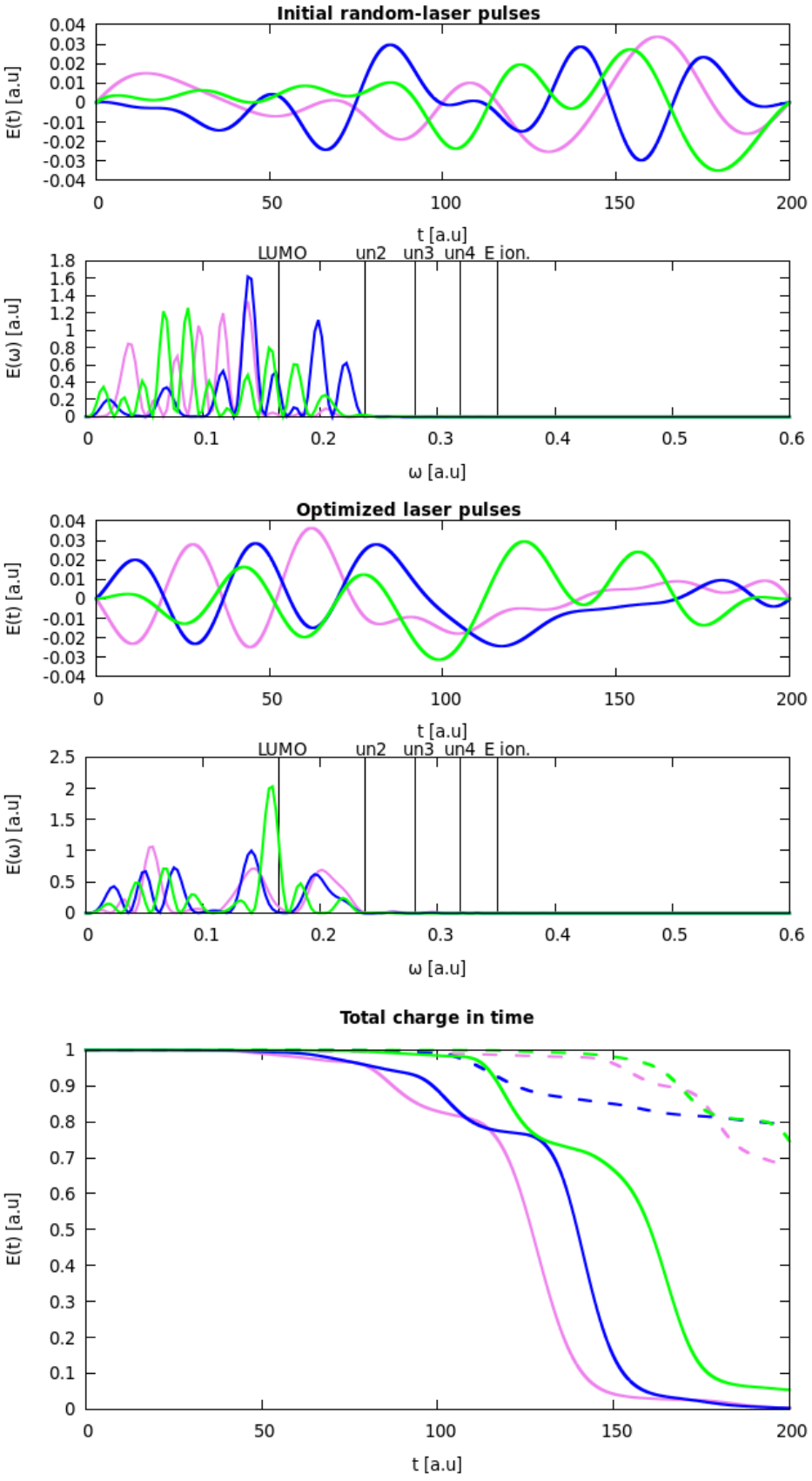

Figure 6.8: ***Top****: three different random initial lasers with fluence = 0.04000 and* $\omega_{max} = 0.25$ *a.u. Upper half: time-domain. Lower half: frequency-domain.*
Middle*: optimized lasers with the BOBYQA method, holding the initial fluence. Upper half: time-domain. Lower half: frequency-domain.*
Bottom*: loss of charge induced by the initial and optimized lasers. Lines of the same color mean loss of charge induced by the same initial laser. Dashed line: initial laser. Solid line: optimized laser.*

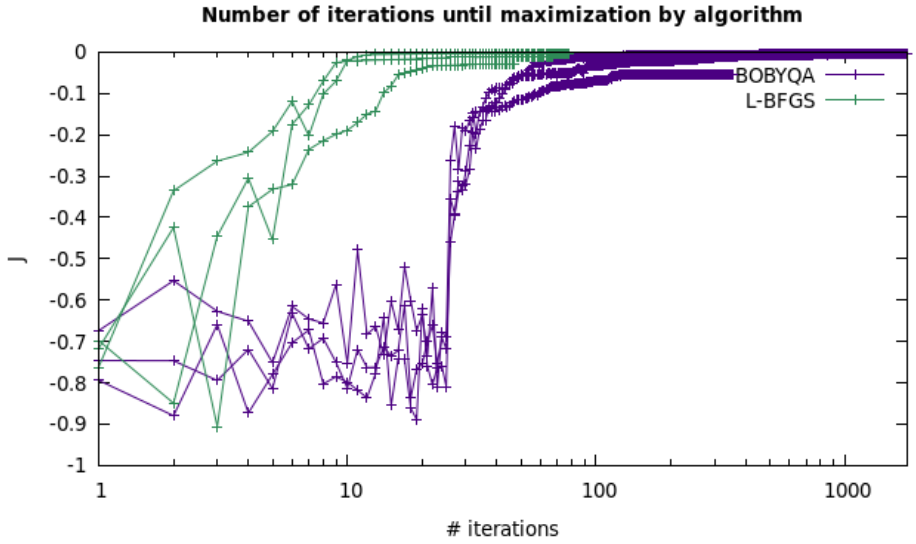

Figure 6.9: Number of function evaluations in the optimization for L-BFGS and BOBYQA algorithms for $\omega_{max} = 0.25$.

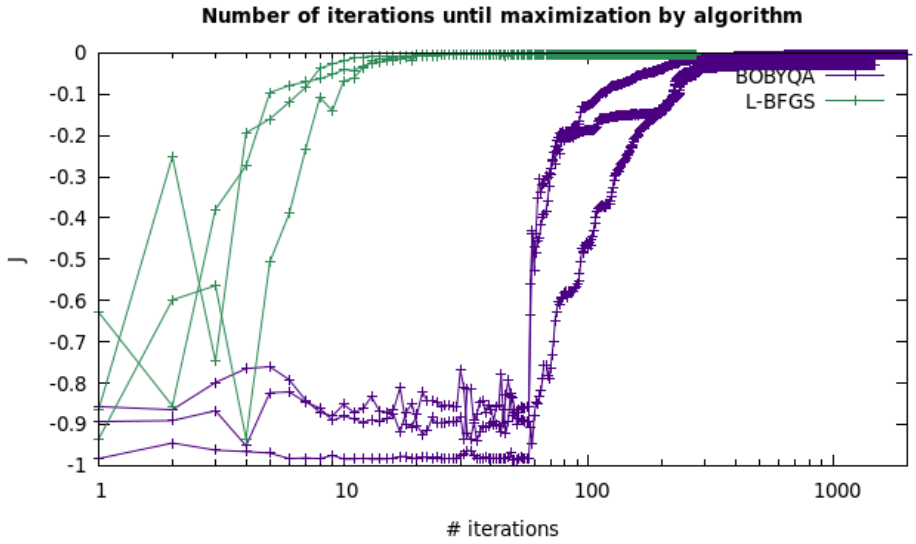

Figure 6.10: Number of function evaluations in the optimization for L-BFGS and BOBYQA algorithms for $\omega_{max} = 0.5$.

6.2 Ab initio alanine molecule

With the calculations above the work with the one-dimensional model is finished, providing an approximated value for the necessary intensity which we plan to use for the 3D model. The focus moves to the complete three-dimensional model whose electron density has been estimated by a DFT run (figures 6.11, 6.12). This time will not be possible to use the L-BFGS method due to a great error in the gradient calculation by the formulae given in the QOCT section, therefore only BOBYQA was used for the optimization. The implementation in octopus-code of the 3D model consists in the setting of these parameters:

- Shape and size of the box: In this case it is a spherical box of 32 a.u of diameter including absorbing boundaries of 3 a.u.
- Grid's spacing and timestep: Now the spacing is 0.435 a.u and the time-step is 0.04 a.u.
- Potential: Now pseudo-potentials defined on the grid are used for this purpose.
- Initial coordinates and laser field: Initial positions according to the centered experimental coordinates and an envelope sine function as the initial input laser field. As in the 1D case, it is randomized afterwards.
- Optimization parameters such as the cut-off frequency (maximum component of frequency) of the final laser and the target functional which is the one described in chapter 4.

Figure 6.13 shows the simulation with the same laser settings than in the one-dimensional model BOBYQA optimization. One can observe that the spectrum for optimized lasers has only one relevant component in a high-frequency zone, higher than the ionization energy. This could be due to the upshifting of the ionization potential while the molecule is being ionized, making more difficult to pull out inner electrons. An insight of this is that an electron has already been pulled out of the molecule when this happens (blue and purple, bottom graph of fig. 6.13), while in the case of the laser coloured in green one can observe that an entire unit charge has not been pulled out and therefore the spectrum has more contributions at lower frequencies.

This picture contrast with the the one-dimensional model one, where the frequency of the laser does not have to surpass the ionization energy (maybe by the use of the 3D ionization potential). The real 3D system responds as expected with a necessary cut-off frequency higher than the ionization potential. By the other hand the intensity needed coincides in both models, maybe due to the fact that the energy levels are the same and the higher amount of degrees of freedom in the 3D case is compensated by a bigger characteristic length L of the system.

The comparison of the optimized laser field in purple (iteration #66 in the QOC procedure) with lasers from previous iterations (figure 6.14) could give information about the evolution of the peaks in the spectrum along the optimization process. The laser obtained in the 40th iteration has several peaks included one at the ionization energy. After the ionization of the alanine, the peaks are displaced to $\omega = 0.4$, an insight that more energy is needed for the next electron.

6.3 Displacement of nuclei

Now, we consider that the nuclei are left free and thus they evolve according to Ehrenfest dynamics (section 3.3). The purpose of this is to observe if the optimized laser (or lasers) obtained in previous calculations, produce any change in the atomic structure of the molecule. These effects may be not desired, such as the rupture of one or several bonds, losing the chemical structure and in consequence the functionality of the molecule. In this case, the time of exposition to

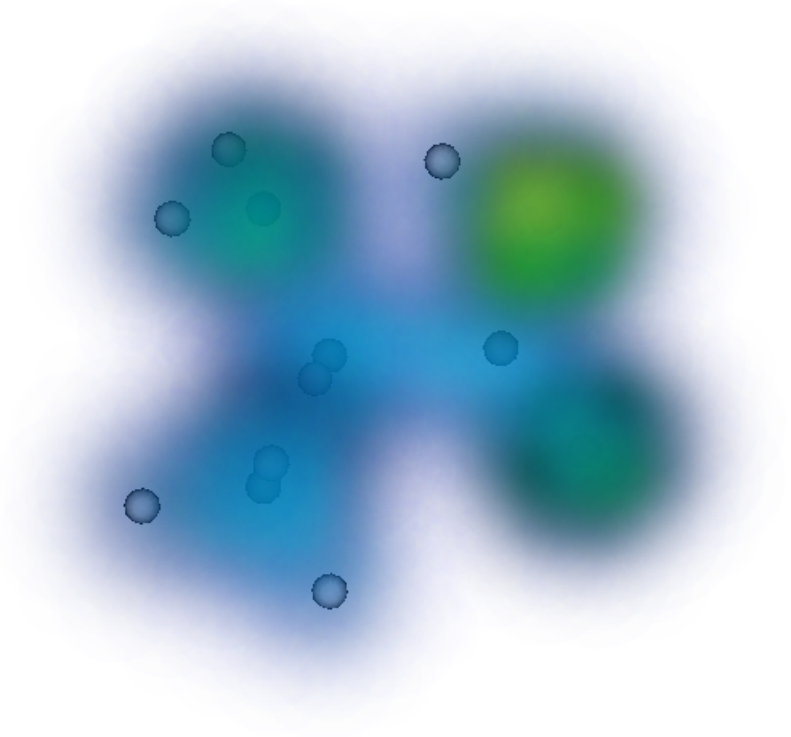

Figure 6.11: Electron density for the alanine molecule. The discs represent the positions of the ions. It can be observed that a bigger concentration of electrons in the N and O orbitals.

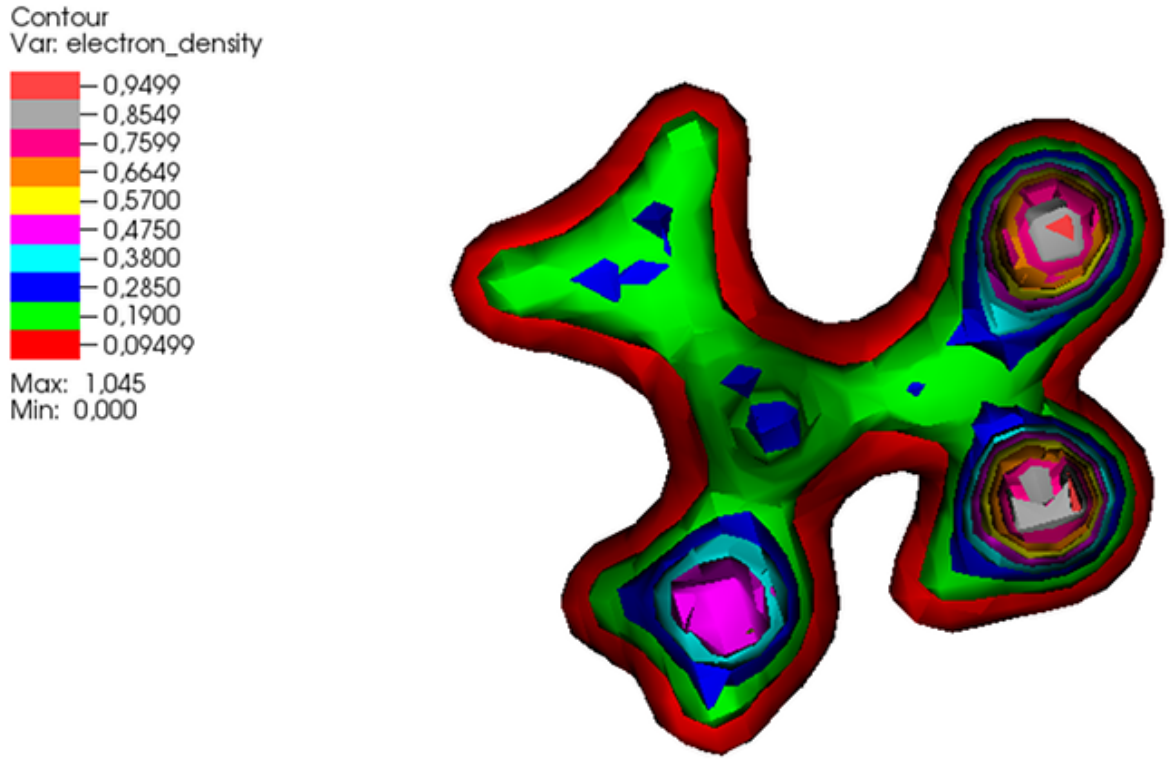

Figure 6.12: Electron density contour for the alanine molecule.

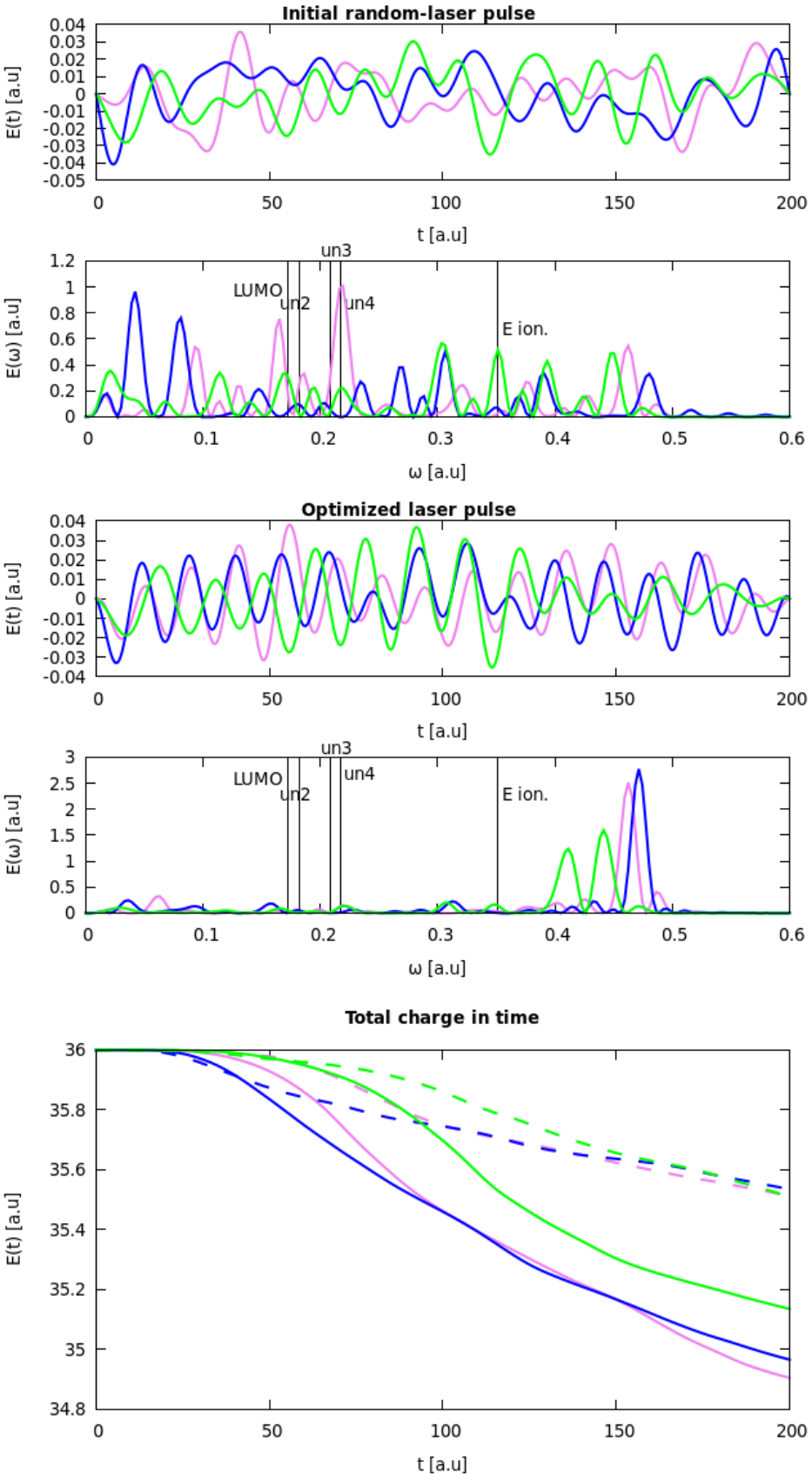

Figure 6.13: ***Top****: random initial laser with fluence = 0.04000 and* $\omega_{max} = 0.5$ *a.u. Upper half: time-domain. Lower half: frequency-domain.*
Middle*: optimize laser with the BOBYQA method, holding the initial fluence. Upper half: time-domain. Lower half: frequency-domain.*
Bottom*: loss of charge induced by the initial and optimized lasers. Lines of the same color mean loss of charge induced by the same initial laser. Dashed line: initial laser. Solid line: optimized laser.*

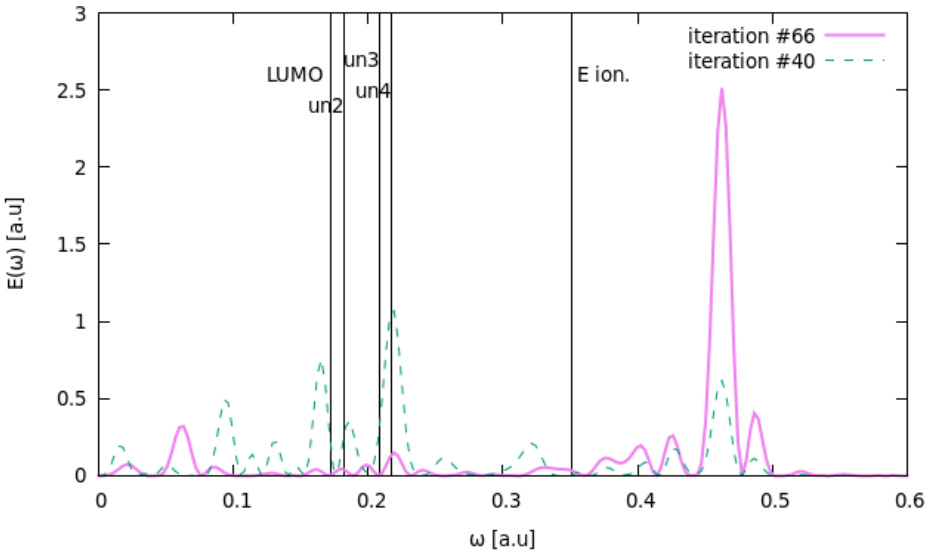

Figure 6.14: Spectrum of the final optimized laser (purple) and of a previous iteration's laser (green).

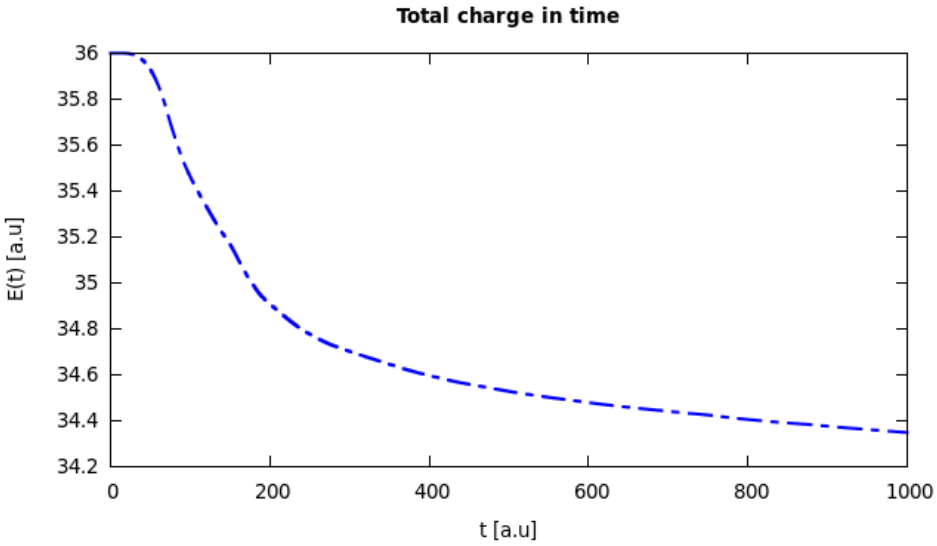

Figure 6.15: Loss of charge in presence of the optimized laser field with duration of 5 pulses.

the field –laser from fig. 6.14– is extended to 1000 femtoseconds –5 pulses– and the plot (6.15) shows that almost two units of electronic charge are pulled out.

The ionization power of the laser field decreases in time due to the upshifting of the ionization potential, so at the end one can observe that a longer exposition to it would not cause any considerable effect than the shown in figure 6.15. The only way to improve the results, assuming than the laser is optimal, would be an increment of the laser's intensity.

The coordinates of nuclei indicate that the molecule does not seem to be divided or cleaved in any way. Total displacement of all atoms (figures 6.16 6.18) shows, except for the carbon C4 (figure 6.1), a slowly increasing distance with respect to their equilibrium positions. However, the presence of an ondulatory component suggests that the atomic bonds are not broken, confirmed by singular velocities (figure 6.19) that oscillate between 0 and a maximum.

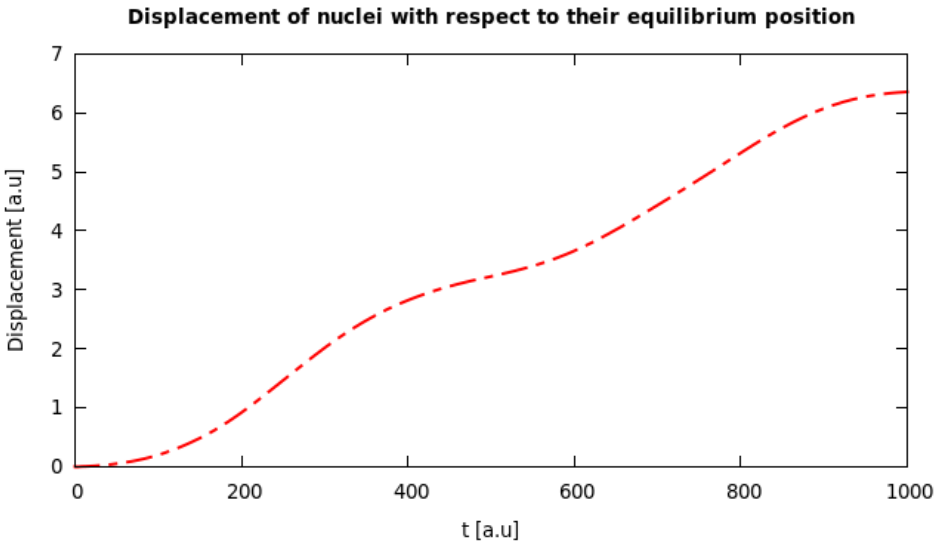

Figure 6.16: Total displacement of nuclei with respect to their equilibrium position.

Another question is if the molecule has been dropped out of boundaries. The dimension of the alanine molecule is around 8 $a.u$ of diameter and the maximum displacement of nuclei is 0.7 $a.u$ for hydrogen atoms. The boundaries start at 11 $a.u$ away from the center of the molecule, so it is impossible that the laser pulse had pulled it out beyond them as it is shown in figure 6.15, in which the electron's escape does not rise spontaneously at the end of the graph.

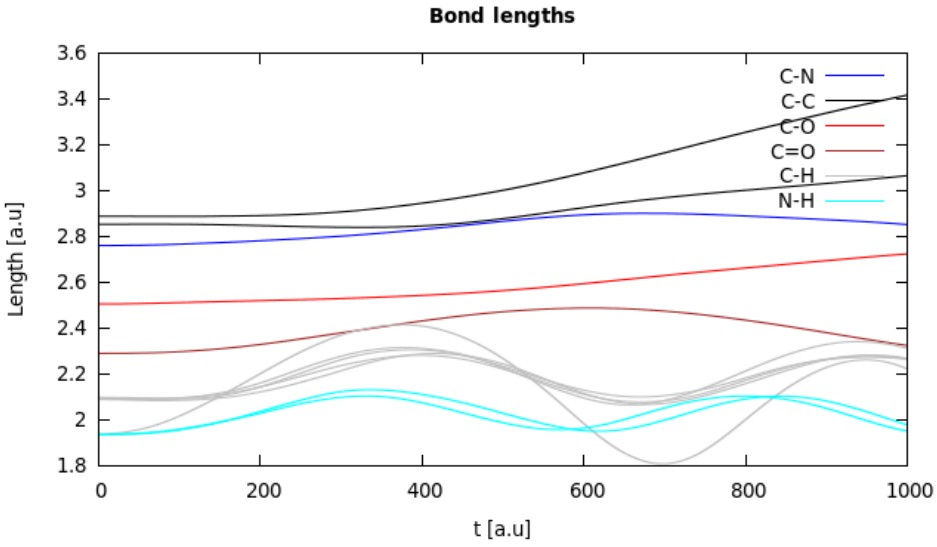

Figure 6.17: Evolution of length during the exposition time of the different types of bond that are present in the molecule.

Figure 6.17 shows the stretching of bonds during the exposition time. Frequency of stretching vibrations depends on the atomic masses and the strength of the bond. A heavier pair of particles has less bond frequency while if the type of bond is stronger, the frequency is higher. Each bond in principle vibrates with a characteristic frequency , for example one can appreciate that the two oxygen atoms vibrate with different frequency. That is because the oxygen O(6) is connected to the carbon C(4) by a double bond C=O, that has a higher characteristic frequency than the single bond C(4)-O(5). These cases are easily recognizable as their atoms are less affected by neighbors, but in the case of the central C-C bonds their movements are a mixture

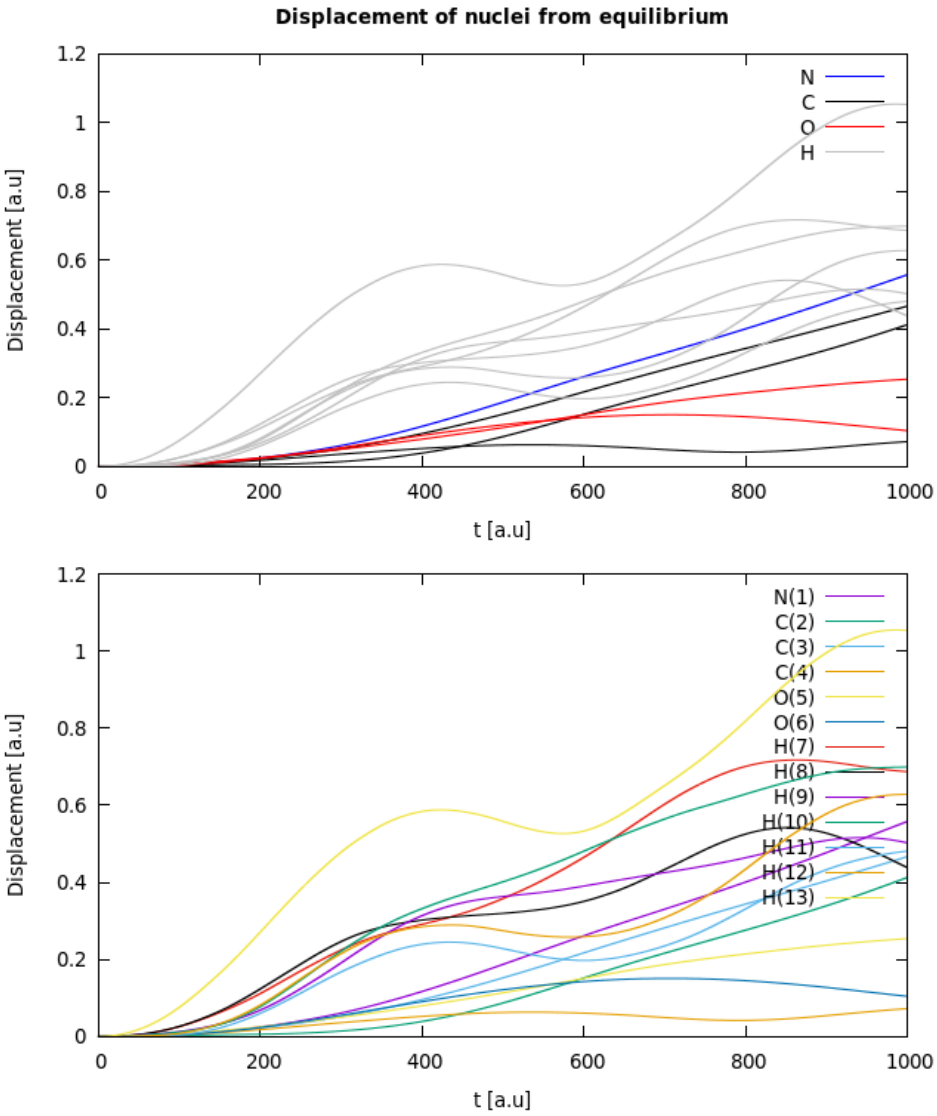

Figure 6.18: Displacement of nuclei with respect to their equilibrium position.

of their characteristic frequency and the interference with their adjacent bonds, acting as spatial constrains.

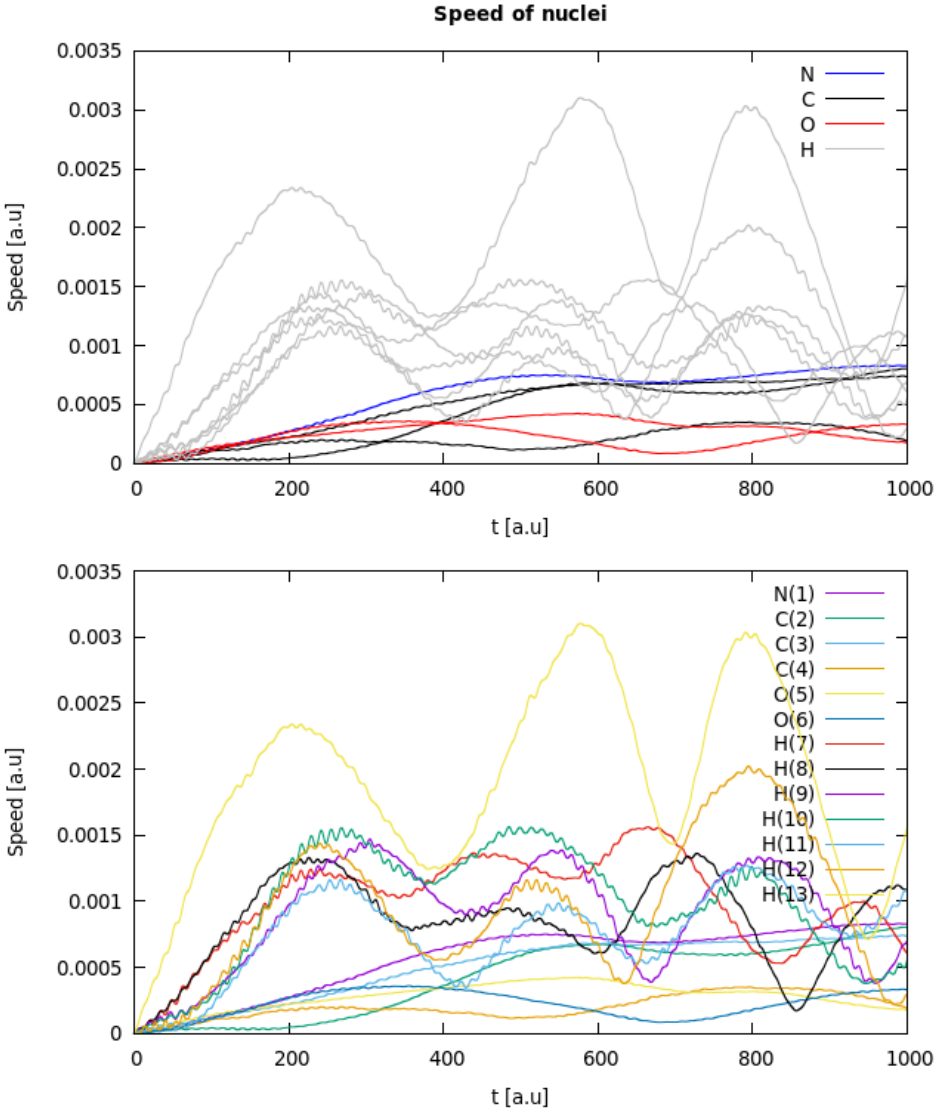

Figure 6.19: Module of velocity of all nuclei.

Chapter 7

Conclusions

This work has studied computationally the control of ionization of molecules upon radiation with high-intensity ultra-short laser pulses. The main conclusions reached have been:

- Still with some difficulties and requiring a good pre-analysis stage, it is feasible to make a theoretical analysis of the target molecule through the previously introduced tools such as TDDFT and Ehrenfest Molecular Dynamics and to implement an Optimal Control Theory solution within this framework.
- There is the possibility to propose different strategies and to implement different optimization methods in order to study the same problem.
- The work opens the door to several studies such as the implementation of geometry target functionals in order to break or to preserve determined bonds or the consideration of solvents surrounding the molecules. Moreover, an experimental realization of the experiment would be interesting.

The main numerical difficulty in the development of the work was to be concerned about the available time along with the possibilities of the resources. Some calculations with the L-BFGS algorithm in the 3D case could not be performed, as the error in the gradient calculated with the QOCT formulae by the L-BFGS method was much bigger than the calculated numerically. That caused some trouble in the process of optimization and the lose of some time. Unfortunately the required decrease in the timestep and mesh' spacing would cause that the simulation lasts for several months. On the other hand, the BOBYQA algorithm's procedure granted a valid optimized laser pulse that could be analysed in order to check the stability of the molecule, fulfilling the objectives set by the initial roadmap.

Bibliography

[1] Castro, A. (2009) "Femtosecond laser pulse shaping for enhanced ionization", EPL-Europhysics Letters **87**, 53001

[2] Kammerlander, D. ; Castro, A. ; Marques, Miguel A.L. (2017) "Optimization of the ionization time of an atom with tailored laser pulses: a theoretical study", Eur. Phys. J. B **90**, **91**

[3] Pueyo, A.G. ; Budagosky, J.A. ; Castro, A. (2016) "Optimal control with nonadiabatic molecular dynamics: Application to the Coulomb explosion of Sodium clusters", Phys. Rev. A **94**, 063421

[4] Keller, U. (2003) "Recent developments in compact ultrafast lasers". Nature (London) **424**: 831.

[5] Thomas, L.H. (1927) "The calculation of atomic fields". Proc. Camb. Phil. Soc. **23** (5): 542–548.

[6] Fermi, E. (1927) "Un Metodo Statistico per la Determinazione di alcune Prioprietà dell'Atomo". Rend. Accad. Naz. Lincei. **6**: 602–607.

[7] Teller, E.. (1962) "On the Stability of Molecules in the Thomas-Fermi Theory". Rev. Mod. Phys. **34**: 627.

[8] Hohenberg, P.; Kohn, W. (1964) "Inhomogeneous electron gas". Physical Review. **136** (3B): B864–B871.

[9] Kohn, W.; Sham, L.J. (1965) "Self-Consistent Equations Including Exchange and Correlation Effects". Physical Review. **140** (4A): A1133–A1138.

[10] Runge, E.; Grossm, E.K.U. (1984) "Density-Functional Theory for Time-Dependent Systems". Phys. Rev. Lett. **52** (12): 997–1000.

[11] Werschnik, J.; Gross, E.K.U. (2007) "Quantum Optimal Control Theory". J Phys B: At, Mol Opt Phys. **40** (18):R175.

[12] "Maiman Builds First Working Laser". Physics History: May 16, 1960. APS News **19**. May 2010.

[13] Bloembergen, N.; Zewail A.H.. (1984) "Energy redistribution in isolated molecules and the question of mode-selective laser chemistry revisited". The Journal of Physical Chemistry **88** (23): 5459–5465.

[14] Judson, R.S.; Rabitz H. (1992) "Teaching lasers to control molecules". Phys Rev Lett **68**: 1500–1503.

[15] Schrödinger, E. (1926) "An Undulatory Theory of the Mechanics of Atoms and Molecules". Phys. Rev. **28** (6): 1049–1070.

[16] Born, M.; Oppenheimer, J. R. (1927) "Zur Quantentheorie der Molekeln". *Annalen der Physik.* **389**(20): 457-484.

[17] Bornemann, F.A.; Nettesheim, P.; Schütte C. (1996) "Quantum-classical molecular dynamics as an approximation to full quantum dynamics". J. Chem. Phys. **105**.

[18] Bornemann, F.A. ; Schutte, C. (1999) "On the Singular Limit of the Quantum-Classical Molecular Dynamics Model". SIAM Journal on Applied Mathematics, **59** (4):1208-1224

[19] Mulliken; Robert S. (1932) "Electronic Structures of Polyatomic Molecules and Valence. II. General Considerations". Physical Review. **41** (1): 49–71.

[20] Slater, J.C. (1930) "Atomic Shielding Constants". Physical Review. **36**: 57.

[21] Boys, S.F., Proc. Royal Society, vol. **A200**, p. 542

[22] Hellmann, H. (1935) "A New Approximation Method in the Problem of Many Electrons", Journal of Chemical Physics, Karpow-Institute for Physical Chemistry, Moscow. **3**: 61

[23] Pauli, W. (1925) "Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren". Zeitschrift für Physik. **31**: 765–783.

[24] Pontryagin, L.S.; Boltyanskii V.G.; Gamkrelidze R.V.; Mishchenko E.F. (1962) "The Mathematical Theory of Optimal Processes" John Wiley & Sons.

[25] Broyden, C.G. (1970) "The convergence of a class of double-rank minimization algorithms", Journal of the Institute of Mathematics and Its Applications, **6**: 76–90.

[26] Fletcher, R. (1970) "A New Approach to Variable Metric Algorithms", Computer Journal, **13** (3): 317–322.

[27] Goldfarb, D. (1970) "A Family of Variable Metric Updates Derived by Variational Means", Mathematics of Computation, **24** (109): 23–26.

[28] Shanno, D.F. (1970) "Conditioning of quasi-Newton methods for function minimization", Mathematics of Computation, **24** (111): 647–656.

[29] Powell, M.J.D. (2009) "The BOBYQA algorithm for bound constrained optimization without derivatives" Department of Applied Mathematics and Theoretical Physics, Cambridge University. DAMTP 2009/NA06.

[30] Van Meer, R.; Gritsenko, V.; Baerends, E.J. "Physical Meaning of Virtual Kohn-Sham Orbitals and Orbital Energies: An Ideal Basis for the Description of Molecular Excitations".

[31] Stowasser, R. ; Hoffmann, R. (1999) "What Do the Kohn-Sham Orbitals and Eigenvalues Mean?" J. Am. Chem. Soc. 1999, **121**, 3414-3420

[32] Gordon, R.J. ; Rice, S.A. (1997) "Active control of the dynamics of atoms and molecules." Annual Review of Physical Chemistry **48**, no. 1: 601-641.

[33] Shapiro, M. ; Brumer, P. (2000) "Coherent control of atomic, molecular, and electronic processes." Advances in atomic, molecular, and optical physics **42**: 287-345.

[34] Møller, C.; Plesset, M.S. (1934) "Note on an Approximation Treatment for Many-Electron Systems". Phys. Rev. **46** (7): 618–622.

[35] Van Leeuwen, R. (1998) "Causality and Symmetry in Time-Dependent Density-Functional Theory". Phys. Rev. Lett. **80**, 1280.

Appendix A: Units

A.1 Fundamental Constants

- Planck's constant : $\hbar = 1 = \frac{h}{2\pi} = 1.05 \times 10^{-34} J \cdot s$
- Electron charge : $e = 1 = 1.60 \times 10^{-19} C$
- Electron mass : $m = 1 = 9.11 \times 10^{-31} kg$
- Permitivity : $\epsilon_0 = (4\pi)^{-1} = 8.85 \times 10^{-12} C^2 \cdot J^{-1} \cdot m^{-1}$
- Bohr radius : $a_0 = 1 = 4\pi\epsilon_0 \frac{\hbar^2}{mc^2} = 5.29 \times 10^{-11} m$
- Structure constant : $\frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} = \frac{1}{137}$
- Speed of light : $c = \alpha^{-1} = 137 = 3.00 \times 10^8 m \cdot s^{-1}$
- Permeability : $\mu_0 = 4\pi c^{-2} = 4\pi \times 10^{-7} N \cdot s^2 \cdot C^{-2}$

A.2 Atomic Units

One atomic unit is equivalent to the following quantities:

- Mass : $m = 9.11 \times 10^{-31} kg$
- Charge : $e = 1.60 \times 10^{-19} C$
- Length : $a_0 = 5.29 \times 10^{-11} m$
- Energy : $\frac{e^2}{a_0} = 27.2 eV$
- Speed : $v_0 = \alpha c = 2.19 \times 10^6 m \cdot s^{-1}$
- Time : $\frac{a_0}{v_0} = 2.42 \times 10^{-17} s = 24.2\ as$
- Frequency : $\frac{v_0}{2\pi a_0} = 6.58 \times 10^{15} Hz$
- Electric potential : $\frac{1}{4\pi\epsilon_0} \frac{e}{a_0} = 27.2 V$
- Electric field strength : $\frac{1}{4\pi\epsilon_0} \frac{e}{a_0^2} = 5.14 \times 10^9 V \cdot cm^{-1}$
- Light intensity : $\frac{1}{2} (\mu_0 c)^{-1} \left(\frac{1}{4\pi\epsilon_0} \frac{e}{a_0} \right)^2 = 3.5 \times 10^{16} W \cdot cm^{-2}$