

A mathematical and computational review of Hartree-Fock SCF methods in Quantum Chemistry

Pablo Echenique*^{†‡} and J. L. Alonso^{†‡}

[†] Departamento de Física Teórica, Universidad de Zaragoza,
Pedro Cerbuna 12, 50009, Zaragoza, Spain.

[‡] Instituto de Biocomputación y Física de los Sistemas Complejos (BIFI),
Edificio Cervantes, Corona de Aragón 42, 50009, Zaragoza, Spain.

(May 2007)

We present here a review of the fundamental topics of Hartree-Fock theory in Quantum Chemistry. From the molecular Hamiltonian, using and discussing the Born-Oppenheimer approximation, we arrive to the Hartree and Hartree-Fock equations for the electronic problem. Special emphasis is placed in the most relevant mathematical aspects of the theoretical derivation of the final equations, as well as in the results regarding the existence and uniqueness of their solutions. All Hartree-Fock versions with different spin restrictions are systematically extracted from the general case, thus providing a unifying framework. Then, the discretization of the one-electron orbitals space is reviewed and the Roothaan-Hall formalism introduced. This leads to an exposition of the basic underlying concepts related to the construction and selection of Gaussian basis sets, focusing in algorithmic efficiency issues. Finally, we close the review with a section in which the most relevant modern developments (specially those related to the design of linear-scaling methods) are commented and linked to the issues discussed. The whole work is intentionally introductory and rather self-contained, so that it may be useful for non experts that aim to use quantum chemical methods in interdisciplinary applications. Moreover, much material that is found scattered in the literature has been put together here to facilitate comprehension and to serve as a handy reference.

Keywords: Quantum Chemistry; introduction; Hartree-Fock; basis sets; SCF

Table of contents

- | | |
|---|--|
| 1. Introduction | 8. The Roothaan-Hall equations |
| 2. Molecular Hamiltonian and atomic units | 9. Introduction to Gaussian basis sets |
| 3. The Born-Oppenheimer approximation | 10. Modern developments: An introduction to linear-scaling methods |
| 4. The variational method | Appendix A: Functional derivatives |
| 5. The statement of the problem | Appendix B: Lagrange multipliers |
| 6. The Hartree approximation | |
| 7. The Hartree-Fock approximation | |

1 Introduction

In the hot field of computer simulation of biological macromolecules, available potential energy functions are often not accurate enough to properly describe complex processes such as the folding of proteins [1–7]. In order to improve the situation, it is convenient to extract ab initio information from quantum mechanical calculations with the hope of being able to devise less computationally demanding methods that can be used to tackle large systems. In this spirit, the effective potential for the nuclei calculated in the non-relativistic Born-Oppenheimer approximation is

*Corresponding author. Email: echenique.p@gmail.com

typically considered as a good reference to assess the accuracy of cheaper potentials [8–14]. The study of molecules at this level of theoretical detail and the design of computationally efficient approximations for solving the demanding equations that appear constitute the major part of the field called *quantum chemistry* [15,16]. In this work, we voluntarily circumscribe ourselves to the basic formalism needed for the ground-state quantum chemical calculations that are typically performed in this context. For more general expositions, we refer the reader to any of the thorough accounts in refs. [17–19].

In sec. 2, we introduce the molecular Hamiltonian and a special set of units (the atomic ones) that are convenient to simplify the equations. In sec. 3, we present in an axiomatic way the concepts and expressions related to the separation of the electronic and nuclear problems in the Born-Oppenheimer scheme. In sec. 4, we introduce the variational method that underlies the derivation of the basic equations of the Hartree and Hartree-Fock approximations, discussed in sec. 6 and 7 respectively. The computational implementation of the Hartree-Fock approximation is tackled in sec. 8, where the celebrated Roothaan-Hall equations are derived. In sec. 9, the main issues related to the construction and selection of Gaussian basis sets are discussed, and, finally, in sec. 10, the hottest areas of modern research are briefly reviewed and linked to the issues in the rest of the work, with a special emphasis in the development of linear-scaling methods.

2 Molecular Hamiltonian and atomic units

Since 1960, the international scientific community has agreed on an ‘official’ set of basic units for measurements: *Le Système International d’Unités*, or SI for short (see <http://www.bipm.org/en/si/> and ref. [20]). The meter (m), the kilogram (kg), the second (s), the ampere (A), the kelvin (K), the mole (mol), the joule (J) and the pascal (Pa) are examples of SI units.

Sticking to the SI scheme, the non-relativistic quantum mechanical Hamiltonian operator of a molecule consisting of N_N nuclei (with atomic numbers Z_α and masses M_α , $\alpha = 1, \dots, N_N$) and N electrons (i.e., the *molecular Hamiltonian*) is expressed as¹:

$$\hat{H} = - \sum_{\alpha=1}^{N_N} \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 - \sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_{\alpha \neq \beta} \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_\alpha|} - \sum_{i=1}^N \sum_{\alpha=1}^{N_N} \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_i|} + \frac{1}{2} \sum_{i \neq j} \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|}, \quad (1)$$

where \hbar stands for $h/2\pi$, being h Planck’s constant, m_e denotes the electron mass, e the proton charge, \mathbf{r}_i the position of the i -th electron, \mathbf{R}_α that of the α -th nucleus, ϵ_0 the vacuum permittivity and ∇_i^2 the Laplacian operator with respect to the coordinates of the i -th particle.

Although using a common set of units presents obvious communicative advantages, when circumscribed to a particular field of science, it is common to appeal to non-SI units in order to simplify the most frequently used equations by getting rid of some constant factors that always appear grouped in the same ways and, thus,

¹ Note that the non-relativistic molecular Hamiltonian does not depend on spin-like variables.

Table 1. Atomic units up to five significant digits. Taken from the National Institute of Standards and Technology (NIST) web page at <http://physics.nist.gov/cuu/Constants/>. Note that only four independent units are required in a mechanical-plus-electromagnetic system. The rest of them can be easily obtained from any such four. For example, using the units in the table, the relations $\hbar = 1$ and $1/(4\pi\epsilon_0) = 1$ result.

Unit of mass:	mass of the electron = $m_e = 9.1094 \cdot 10^{-31}$ kg
Unit of charge:	charge on the proton = $e = 1.6022 \cdot 10^{-19}$ C
Unit of length:	1 bohr = $a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.52918 \text{ \AA} = 5.2918 \cdot 10^{-11}$ m
Unit of energy:	1 hartree = $\frac{\hbar^2}{m_e a_0^2} = 627.51 \text{ kcal/mol} = 4.3597 \cdot 10^{-18}$ J

Table 2. Energy units conversion factors to five significant digits. Taken from the National Institute of Standards and Technology (NIST) web page at <http://physics.nist.gov/cuu/Constants/>. The table must be read by rows. For example, the value 4.1838, in the third row, fourth column, indicates that 1 kcal/mol = 4.1838 kJ/mol.

	1 hartree	1 eV	1 kcal/mol	1 kJ/mol	1 cm ⁻¹
1 hartree	1	27.211	627.51	262.54	219470
1 eV	$3.6750 \cdot 10^{-2}$	1	23.061	96.483	8065.5
1 kcal/mol	$1.5936 \cdot 10^{-3}$	$4.3363 \cdot 10^{-2}$	1	4.1838	349.75
1 kJ/mol	$3.8089 \cdot 10^{-4}$	$1.0364 \cdot 10^{-2}$	$2.3902 \cdot 10^{-1}$	1	83.595
1 cm ⁻¹	$4.5560 \cdot 10^{-6}$	$1.2398 \cdot 10^{-4}$	$2.8592 \cdot 10^{-3}$	$1.1962 \cdot 10^{-2}$	1

make the numerical values in any calculation of the order of unity. In the field of quantum chemistry, *atomic units* (see table 1), proposed in ref. [21] and named in ref. [22], are typically used. In these units, eq. (1) is substantially simplified to

$$\hat{H} = - \sum_{\alpha=1}^{N_N} \frac{1}{2M_\alpha} \nabla_\alpha^2 - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 + \frac{1}{2} \sum_{\alpha \neq \beta} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_\alpha|} - \sum_{i=1}^N \sum_{\alpha=1}^{N_N} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_i|} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|}. \quad (2)$$

Since all the relevant expressions in quantum chemistry are derived in one way or another from the molecular Hamiltonian, the simplification brought up by the use of atomic units propagates to the whole formalism. Consequently, they shall be the choice all throughout this work.

Apart from the atomic units and the SI ones, there are some other miscellaneous units that are often used in the literature: the *ångström*, which is a unit of length defined as $1 \text{ \AA} = 10^{-10}$ m, and the units of energy cm^{-1} (which reminds about the spectroscopic origins of quantum chemistry and, even, quantum mechanics), *electronvolt* (eV), *kilocalorie per mole* (kcal/mol) and *kilojoule per mole* (kJ/mol). The last two are specially used in the field of macromolecular simulations and quantify the energy of a mole of entities; for example, if one asserts that the torsional barrier height for H₂O₂ is ~ 7 kcal/mol, one is really saying that, in order to make a mole of H₂O₂ (i.e., $N_A \simeq 6.0221 \cdot 10^{23}$ molecules) rotate 180° around the O–O bond, one must spend ~ 7 kcal. For the conversion factors between the different energy units, see table 2.

Finally, to close this section, we rewrite eq. (2) introducing some self-explanatory notation that will be used in the subsequent discussion:

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{eN} + \hat{V}_{ee} , \quad (3a)$$

$$\hat{T}_N := - \sum_{\alpha=1}^{N_N} \frac{1}{2M_\alpha} \nabla_\alpha^2 , \quad (3b)$$

$$\hat{T}_e := - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 , \quad (3c)$$

$$\hat{V}_{NN} := \frac{1}{2} \sum_{\alpha \neq \beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} , \quad (3d)$$

$$\hat{V}_{eN} := - \sum_{i=1}^N \sum_{\alpha=1}^{N_N} \frac{Z_\alpha}{R_{\alpha i}} , \quad (3e)$$

$$\hat{V}_{ee} := \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}} . \quad (3f)$$

3 The Born-Oppenheimer approximation

To think of a macromolecule as a set of quantum objects described by a *wavefunction* $\Psi(X_1, \dots, X_{N_N}, x_1, \dots, x_N)$ dependent on the spatial and spin¹ degrees of freedom, $x_i := (\mathbf{r}_i, \sigma_i)$, of the electrons and on those of the nuclei, $X_\alpha := (\mathbf{R}_\alpha, \Sigma_\alpha)$, would be too much for the imagination of physicists and chemists. All the language of chemistry would have to be remade and simple sentences in textbooks, such as “rotation about this single bond allows the molecule to avoid steric clashes between atoms” or even “a polymer is a long chain-like molecule composed of repeating monomer units”, would have to be translated into long and counter-intuitive statements involving probability and ‘quantum jargon’. Conscious or not, we think of molecules as classical objects.

More precisely, we are ready to accept that electrons are quantum (we know of the interference experiments, electrons are light, we are accustomed to draw atomic ‘orbitals’, etc.), however, we are reluctant to concede the same status to nuclei. Nuclei are heavier than electrons (at least ~ 2000 times heavier, in the case of the single proton nucleus of hydrogen) and we picture them in our imagination as ‘classical things’ that move, bond to each other, rotate around bonds and are at precise points at precise times. We imagine nuclei ‘slowly moving’ in the field of the electrons, which, for each position of the first, immediately ‘adjust their quantum state’.

The formalization of these ideas is called *Born-Oppenheimer (BO) approximation* [23,24] and the confirmation of its being good for many relevant problems is a fact that supports our intuitions about the topic and that lies at the foundations of the vast majority of the images, the concepts and the research in quantum chemistry².

¹ One convenient way of thinking about functions that depend on spin-like variables is as an m -tuple of ordinary \mathbb{R}^{3N} functions, where m is the finite number of possible values of the spin. In the case of a one-particle wavefunction describing an electron, for example, σ can take two values (say, $-1/2$ and $1/2$) in such a way that one may picture any general *spin-orbital* $\Psi_i(x)$ as a 2-tuple $(\Phi_i^{-1/2}(\mathbf{r}), \Phi_i^{1/2}(\mathbf{r}))$. Of course, another valid way of imagining $\Psi_i(x)$ is simply as a function of four variables, three real and one discrete.

² There are many phenomena, however, in which the Born-Oppenheimer approximation is broken. For example, in striking a flint to create a spark, mechanical motion of the nuclei excites electrons into a

Like any approximation, the Born-Oppenheimer one may be either *derived* from the exact problem (in this case, the entangled behaviour of electrons and nuclei as the same quantum object) or simply *proposed* on the basis of physical intuition, and later confirmed to be good enough (or not) by comparison with the exact theory or with the experiment. Of course, if it is possible, the first way should be preferred, since it allows to develop a deeper insight about the terms we are neglecting and the specific details that we will miss. However, although in virtually every quantum chemistry book [18, 19, 26–28] hand-waving derivations up to different levels of detail are performed and the Born-Oppenheimer approximation is typically presented as unproblematic, it seems that the fine mathematical details on which these ‘standard’ approaches are based are far from clear [29–31]. This state of affairs does not imply that the final equations that will need to be solved are ill-defined or that the numerical methods based on the theory are unstable; in fact, it is just the contrary (see the discussion below), because the problems are related only to the precise relation between the concepts in the whole theory and those in its simplified version. Nevertheless, the many subtleties involved in a *derivation* of the Born-Oppenheimer approximation scheme from the exact equations suggest that the second way, that of *proposing* the approximation, be taken. Hence, in the following paragraphs, an *axiomatic* presentation of the main expressions, aimed mostly to fix the notation and to introduce the language, will be performed.

First of all, if we examine the Hamiltonian operator in eq. (2), we see that the term \hat{V}_{eN} prevents the problem from being separable in the nuclear and electronic coordinates, i.e., if we define $\underline{x} := (x_1, \dots, x_N)$ as the set of all electronic coordinates (spatial and spin-like) and do likewise with the nuclear coordinates \underline{X} , the term \hat{V}_{eN} prevents any wavefunction $\Psi(\underline{X}, \underline{x})$ solution of the *time-independent Schrödinger equation*,

$$\hat{H} \Psi(\underline{X}, \underline{x}) = \left(\hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{eN} + \hat{V}_{ee} \right) \Psi(\underline{X}, \underline{x}) = E \Psi(\underline{X}, \underline{x}) , \quad (4)$$

from being written as a product, $\Psi(\underline{X}, \underline{x}) = \Psi_N(\underline{X})\Psi_e(\underline{x})$, of an electronic wavefunction and a nuclear one. If this were the case, the problem would still be difficult (because of the Coulomb terms \hat{V}_{NN} and \hat{V}_{ee}), but we would be able to focus on the electrons and on the nuclei separately.

The starting point for the Born-Oppenheimer approximation consists in assuming that a less strict separability is achieved, in such a way that, for a pair of suitably chosen $\Psi_N(\underline{X})$ and $\Psi_e(\underline{x}; \underline{X})$, any wavefunction solution of eq. (4) (or at least those in which we are interested; for example, the eigenstates corresponding to the lowest lying eigenvalues) can be expressed as

$$\Psi(\underline{X}, \underline{x}) = \Psi_N(\underline{X})\Psi_e(\underline{x}; \underline{X}) , \quad (5)$$

where we have used a ‘;’ to separate the two sets of variables in the electronic part of the wavefunction in order to indicate that, in what follows, it is convenient to use the image that ‘from the point of view of the electrons, the nuclear degrees of freedom are fixed’, so that the electronic wavefunction depends ‘parametrically’ on them. In other words, that the \underline{X} are not quantum variables in eq. (6) below. Of course, it is just a ‘semantic’ semicolon; if anyone feels uncomfortable about it, she may drop it and write a normal comma.

plasma that then emits light [25].

Notably, in ref. [32], Hunter showed that any solution of the Schrödinger equation can in fact be written exactly in the form of eq. (5), and that the two functions, $\Psi_N(\underline{\mathbf{X}})$ and $\Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{X}})$, into which $\Psi(\underline{\mathbf{X}}, \underline{\mathbf{x}})$ is split may be interpreted as marginal and conditional probability amplitudes respectively. However, despite the insight that is gained from this treatment, it is of no practical value, since the knowledge of the exact solution $\Psi(\underline{\mathbf{X}}, \underline{\mathbf{x}})$ is required in order to compute $\Psi_N(\underline{\mathbf{X}})$ and $\Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{X}})$ in Hunter's approach.

In the Born-Oppenheimer scheme, an additional assumption is made in order to avoid this drawback: *the equations obeyed by the electronic and nuclear parts of the wavefunction are supposed to be known*. Hence, $\Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{X}})$ is assumed to be a solution of the time-independent *clamped nuclei Schrödinger equation*,

$$\left(\hat{T}_e + \hat{V}_{eN}(\underline{\mathbf{r}}; \underline{\mathbf{R}}) + \hat{V}_{ee}(\underline{\mathbf{r}}) \right) \Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{R}}) := \hat{H}_e(\underline{\mathbf{R}}) \Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{R}}) = E_e(\underline{\mathbf{R}}) \Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{R}}), \quad (6)$$

where the *electronic Hamiltonian operator* $\hat{H}_e(\underline{\mathbf{R}})$ and the *electronic energy* $E_e(\underline{\mathbf{R}})$ (both dependent on the nuclei positions) have been defined, and, since the nuclear spins do not enter the expression, we have explicitly indicated that Ψ_e depends parametrically on $\underline{\mathbf{R}}$ and not on $\underline{\mathbf{X}}$.

The common interpretation of the clamped nuclei equation is, as we have advanced at the beginning of the section, that the nuclei are much 'slower' than the electrons and, therefore, the latter can automatically adjust their quantum state to the instantaneous positions of the former. Physically, eq. (6) is just the time-independent Schrödinger equation of N particles (the electrons) of mass m_e and charge $-e$ in the *external electric field* of N_N *point charges* (the nuclei) of size eZ_α at locations \mathbf{R}_α . Mathematically, it is an eigenvalue problem that has been thoroughly studied in the literature and whose properties are well-known [33–38]. In particular, it can be shown that, in the case of neutral or positively charged molecules (i.e., with $Z := \sum_\alpha Z_\alpha \geq N$), the clamped nuclei equation has an infinite number of normalizable solutions in the discrete spectrum of $\hat{H}_e(\underline{\mathbf{R}})$ (*bound-states*) for every value of $\underline{\mathbf{R}}$ [39, 40].

These solutions must be regarded as the different *electronic energy levels*, and a further approximation that is typically made consists in, not only accepting that the electrons immediately 'follow' nuclear motion, but also that, for each value of the nuclear positions $\underline{\mathbf{R}}$, they are in the electronic ground-state¹, i.e., the one with the lower $E_e(\underline{\mathbf{R}})$.

Consequently, we define

$$E_e^{\text{eff}}(\underline{\mathbf{R}}) := E_e^0(\underline{\mathbf{R}}). \quad (7)$$

to be the *effective electronic field* in which the nuclei move, in such a way that, once we have solved the problem in eq. (6) and know $E_e^0(\underline{\mathbf{R}})$, the time-independent *nuclear Schrödinger equation* obeyed by $\Psi_N(\underline{\mathbf{X}})$ is:

$$\left(\hat{T}_N + \hat{V}_{NN}(\underline{\mathbf{R}}) + E_e^{\text{eff}}(\underline{\mathbf{R}}) \right) \Psi_N(\underline{\mathbf{X}}) := \hat{H}_N \Psi_N(\underline{\mathbf{X}}) = E_N \Psi_N(\underline{\mathbf{X}}), \quad (8)$$

¹ This is customarily assumed in the literature and it is supported by the general fact that electronic degrees of freedom are typically more difficult to excite than nuclear ones. Hence, in the vast majority of the numerical implementations of the theory, only the electronic ground-state is sought. We will see this in the forthcoming sections.

where the *effective nuclear Hamiltonian* \hat{H}_N has been implicitly defined.

Now, to close the section, we put together the main expressions of the Born-Oppenheimer approximation for quick reference and we discuss them in some more detail:

$$\hat{H}_e(\underline{\mathbf{R}}) \Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{R}}) := \left(\hat{T}_e + \hat{V}_{eN}(\underline{\mathbf{R}}) + \hat{V}_{ee} \right) \Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{R}}) = E_e(\underline{\mathbf{R}}) \Psi_e(\underline{\mathbf{x}}; \underline{\mathbf{R}}), \quad (9a)$$

$$E_e^{\text{eff}}(\underline{\mathbf{R}}) := E_e^0(\underline{\mathbf{R}}), \quad (9b)$$

$$\hat{H}_N \Psi_N(\underline{\mathbf{X}}) := \left(\hat{T}_N + V_{NN}(\underline{\mathbf{R}}) + E_e^{\text{eff}}(\underline{\mathbf{R}}) \right) \Psi_N(\underline{\mathbf{X}}) = E_N \Psi_N(\underline{\mathbf{X}}), \quad (9c)$$

$$\Psi(\underline{\mathbf{x}}, \underline{\mathbf{X}}) \simeq \Psi_e^0(\underline{\mathbf{x}}; \underline{\mathbf{R}}) \Psi_N(\underline{\mathbf{X}}), \quad E \simeq E_N. \quad (9d)$$

To start, note that the above equations are written in the logical order in which they are imagined and used in any numerical calculation. First, we assume the nuclei fixed at $\underline{\mathbf{R}}$ and we (hopefully) solve the clamped nuclei electronic Schrödinger equation (eq. (9a)), obtaining the electronic ground-state $\Psi_e^0(\underline{\mathbf{x}}, \underline{\mathbf{R}})$ with its corresponding energy $E_e^0(\underline{\mathbf{R}})$. Next, we repeat this procedure for all possible values¹ of $\underline{\mathbf{R}}$ and end up with an hyper-surface $E_e^0(\underline{\mathbf{R}})$ in $\underline{\mathbf{R}}$ -space. Finally, we add this function to the analytical and easily computable $V_{NN}(\underline{\mathbf{R}})$ and find the *effective potential* that determines the nuclear motion:

$$V_N^{\text{eff}}(\underline{\mathbf{R}}) := V_{NN}(\underline{\mathbf{R}}) + E_e^0(\underline{\mathbf{R}}). \quad (10)$$

It is, precisely, this effective potential that is called *Potential Energy Surface* (PES) (or, more generally, *Potential Energy Hyper-Surface* (PEHS)) in quantum chemistry and that is the central object through which scientists picture chemical reactions or conformational changes of macromolecules [41]. In fact, the concept is so appealing and the classical image so strong that, after ‘going quantum’, we can ‘go classical’ back again and think of nuclei as perfectly classical particles that move in the classical potential $V_N^{\text{eff}}(\underline{\mathbf{R}})$. In such a case, we would not have to solve eq. (9c) but, instead, integrate the Newtonian equations of motion. This is the basic assumption of every typical force field used for classical ground-state *molecular dynamics*, such as the ones in the popular CHARMM [42, 43], AMBER [44–46] or OPLS [47] packages.

Finally, we would like to remind the reader that, despite the hand-waving character of the arguments presented, up to this point, every computational step has a clear description and eqs. (9a) through (9c) could be considered as *definitions* involving a certain degree of notational abuse. To assume that the quantities obtained through this process are close to those that proceed from a rigorous solution of the time independent Schrödinger equation (eq. (4)) is where the approximation really lies. Hence, the more accurate eqs. (9d) are, the better the Born-Oppenheimer guess is, and, like any other one, if one does not trust in the heuristic grounds on which the final equations stand, they may be taken as axiomatic and judged a posteriori according to their results in particular cases².

¹ Of course, this cannot be done in practice. Due to the finite character of available computational resources, what is customarily done is to define a ‘grid’ in $\underline{\mathbf{R}}$ -space and compute $E_e^0(\underline{\mathbf{R}})$ in a finite number of points.

² Until now, two approximations have been done: the non-relativistic character of the objects studied and the Born-Oppenheimer approximation. In the forecoming, many more will be done. The a priori quantification of their goodness in large molecules is a formidable task and, despite the efforts in this direction, in the end, the comparison with experimental data is the only sound method for validation.

In quantum chemistry, the Born-Oppenheimer approximation is assumed in a great fraction of the studies and it allows the central concept of potential energy surface to be well-defined, apart from considerably simplifying the calculations. The same decision is taken in this work.

4 The variational method

There exists a mathematically appealing way of deriving the time independent Schrödinger equation (eq. (4)) from an extremal principle. To this end, we first define the functional (see appendix A) that corresponds to the expected value of the energy,

$$\mathcal{F}[\Psi] := \langle \Psi | \hat{H} | \Psi \rangle , \quad (11)$$

where the traditional *bra* and *ket* notation is read as

$$\langle \Psi | \hat{O} | \Psi \rangle := \int \Psi^*(x) \hat{O} \Psi(x) dx , \quad (12)$$

\hat{O} being any operator in the space of wavefunctions and x a dummy variable representing all possible coordinates on which Ψ depends. The norm of Ψ , in this notation is expressed as $\langle \Psi | \Psi \rangle = \int |\Psi(x)|^2 dx$, and we shall say that Ψ is *normalized* if $\langle \Psi | \Psi \rangle = 1$.

If we want to optimize the energy functional above restricting the search space to the normalized wavefunctions, the constrained-extremals problem that results can be solved via the *Lagrange multipliers method* (see appendix B) by constructing the *associated functional* $\tilde{\mathcal{F}}[\Psi]$, where we introduce a *Lagrange multiplier* λ to force normalization:

$$\tilde{\mathcal{F}}[\Psi] := \mathcal{F}[\Psi] + \lambda \left(\langle \Psi | \Psi \rangle - 1 \right) . \quad (13)$$

If we now ask that the functional derivative of $\tilde{\mathcal{F}}[\Psi]$ with respect to the complex conjugate Ψ^* of the wave function¹ be zero, i.e., we look for the stationary points of $\mathcal{F}[\Psi]$ conditioned by $\langle \Psi | \Psi \rangle = 1$, we obtain the eigenvalues equation for \hat{H} , i.e., the time-independent Schrödinger equation. Additionally, it can be shown, first, that, due to the self-adjointness of \hat{H} , the equation obtained from the stationarity condition with respect to Ψ (not Ψ^*) is just the complex conjugate and adds no new information.

Moreover, one can see that the reverse implication is also true [48], so that, if a given normalized wavefunction Ψ is a solution of the eigenvalue problem and belongs to the discrete spectrum of \hat{H} , then the functional in eq. (13) is stationary with respect to Ψ^* :

¹ A function of a complex variable z (or, analogously, a function on a space of complex functions) may be regarded as depending on two different sets of independent variables: either $\text{Re}(z)$ and $\text{Im}(z)$ or z and z^* . The choice frequently depending on technical issues.

$$\frac{\delta \tilde{\mathcal{F}}[\Psi]}{\delta \Psi^*} = 0 \iff \hat{H} \Psi = -\lambda \Psi := E \Psi \quad \text{and} \quad \langle \Psi | \Psi \rangle = 1 . \quad (14)$$

This result, despite its conceptual interest, is of little practical use, because it does not indicate an operative way to solve the Schrödinger equation different from the ones that we already knew. The equivalence above simply illustrates that mathematical variational principles are over-arching theoretical statements from which the differential equations that actually contain the details of physical systems can be extracted. Nevertheless, using similar ideas, we will derive another simple theorem which is indeed powerfully practical: the *Variational Theorem*.

Let $\{\Psi_n\}$ be a basis of eigenstates of the Hamiltonian operator \hat{H} and $\{E_n\}$ their corresponding eigenvalues. Since \hat{H} is self-adjoint, the eigenstates Ψ_n can be chosen to be orthonormal (i.e., $\langle \Psi_m | \Psi_n \rangle = \delta_{mn}$) and any normalized wavefunction Ψ in the Hilbert space can be written as a linear combination of them¹:

$$|\Psi\rangle = \sum_n C_n |\Psi_n\rangle \quad \text{provided that} \quad \sum_n |C_n|^2 = 1 . \quad (15)$$

If we now denote by E_0 the lowest E_n (i.e., the energy of the ground-state)² and calculate the expected value of the energy on an arbitrary state Ψ such as the one in eq. (15), we obtain

$$\begin{aligned} \langle \Psi | \hat{H} | \Psi \rangle &= \sum_{m,n} C_m^* C_n \langle \Psi_m | \hat{H} | \Psi_n \rangle = \sum_{m,n} C_m^* C_n E_n \langle \Psi_m | \Psi_n \rangle \\ &= \sum_{m,n} C_m^* C_n E_n \delta_{mn} = \sum_n |C_n|^2 E_n \geq \sum_n |C_n|^2 E_0 = E_0 . \end{aligned} \quad (16)$$

This simple relation is the *Variational Theorem* and it states that any wavefunction of the Hilbert space has an energy larger than the one of the ground-state (the equality can only be achieved if $\Psi = \Psi_0$). However trivial this fact may appear, it allows a very fruitful ‘everything-goes’ strategy when trying to approximate the ground-state in a difficult problem. If one has a procedure for finding a promising guess wavefunction (called *variational ansatz*), no matter how heuristic, semi-empirical or intuitive it may be, one may expect that the lower the corresponding energy, the closer to the ground-state it is³. This provides a systematic strategy for improving the test wavefunction which may take a number of particular forms.

One example of the application of the Variational Theorem is to propose a family of normalized wavefunctions Ψ_θ parametrically dependent on a number θ and

¹ We assume here, for the sake of simplicity and in order to highlight the relevant concepts, that \hat{H} has only discrete spectrum. The ideas involved in a general derivation are the same, but the technical details and the notation are more complicated [49].

² Its existence is not guaranteed: it depends on the particular potential in \hat{H} . However, for the physically relevant cases, there is indeed a minimum energy in the set $\{E_n\}$.

³ Of course, this not necessarily so (and, in any case, it depends on the definition of ‘closer’), since it could happen that the $\langle \Psi | \hat{H} | \Psi \rangle$ landscape in the constrained subset of the Hilbert space in which the search is performed be ‘rugged’. In such a case, we may have very different wavefunctions (say, in the sense of the L^2 -norm) with similar energies $\langle \Psi | \hat{H} | \Psi \rangle$. The only ‘direction’ in which one can be sure that the situation improves when using the variational procedure is the (very important) energetic one. That one is also moving towards better values of any other observable is, in general, no more than a *bona fide* assumption.

calculate the θ -dependent expected value of the energy⁴:

$$E(\theta) := \langle \Psi_\theta | \hat{H} | \Psi_\theta \rangle . \quad (17)$$

Then, one may use the typical tools of one-variable calculus to find the minimum of $E(\theta)$ and thus make the best guess of the energy E_0 constrained to the family Ψ_θ . If the ansatz is cleverly chosen, this estimate could be rather accurate, however, for large systems that lack symmetry, it is very difficult to write a good enough form for Ψ_θ .

When dealing with a large number of particles, there exists another protocol based on the Variational Theorem that will permit us to derive the Hartree and Hartree-Fock equations for the electronic wavefunction Ψ_e (see secs. 6 and 7, respectively). The first step is to devise a restricted way (a function f with no free parameters) to express Ψ_e in terms of *one-electron wavefunctions*, also called *orbitals* and denoted¹ by $\{\psi_a(x)\}$, thus reducing the search space to a (typically small) subset of the whole Hilbert space:

$$\Psi_e(x_1, \dots, x_N) = f\left(\{\psi_a(x_i)\}\right) . \quad (18)$$

The second step consists in establishing a (possibly infinite) number of constraints on the one-electron functions²,

$$L_k\left(\{\psi_a(x_i)\}\right) = 0 . \quad (19)$$

With these two ingredients, we can now write the Lagrange functional that describes the constrained problem in terms of the orbitals ψ_a (see eq. (13)):

$$\tilde{\mathcal{F}}[\{\psi_a\}] = \left\langle f(\{\psi_a(x_i)\}) \left| \hat{H} \right| f(\{\psi_a(x_i)\}) \right\rangle + \sum_k \lambda_k L_k\left(\{\psi_a(x_i)\}\right) . \quad (20)$$

Finally, we take the derivatives of $\tilde{\mathcal{F}}[\{\psi_a\}]$ with respect to every $\psi_a(x)$ (normally, with respect to the complex conjugate $\psi_a^*(x)$, see footnote 1 in page 8) and we ask each one to be zero (see appendix A). This produces the final equations that must be solved in order to find the stationary one-electron orbitals.

Of course, these final equations may have multiple solutions. In the cases discussed in this work, there exist procedures to check that a particular solution (found computationally) is, not only stationary, but also minimal [50]. However, to assure that it is, not only locally minimal, but also globally (i.e., that is *optimal*), could be, in general, as difficult as for any other multi-dimensional optimization problem [51–53]. In the Hartree and Hartree-Fock cases, discussed in secs. 6 and 7

⁴ Note that, if the functions Ψ_θ were not normalized, then we should deal with the constrained problem as in (13), or, equivalently, we could include a dividing *overlap term* $\langle \Psi_\theta | \Psi_\theta \rangle$ in (17).

¹ In principle, there could be more orbitals than electrons, however, in both the Hartree and Hartree-Fock applications of this formalism, the index a runs, just like i , from 1 to N .

² Actually, both restrictions (the one at the level of the total wavefunction in eq. (18) and the one involving the one-particle ones in eq. (19)) are simply constraints (see appendix B). The distinction is not fundamental but operative, and it also helps us to devise variational ansatzs separating the two conceptual playgrounds.

respectively, the aufbau principle and a clever choice of the starting guess constitute particular techniques intended to alleviate this problem.

5 Statement of the problem

Assuming the Born-Oppenheimer approximation (see sec. 3 and eqs. (9)), the central problem that one must solve in quantum chemistry is *to find the ground-state of the electronic Hamiltonian for a fixed position $\underline{\mathbf{R}}$ of the nuclei*¹:

$$\hat{H} := \hat{T} + \hat{V}_{eN} + \hat{V}_{ee} := - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{\alpha=1}^{N_N} \frac{Z_{\alpha}}{R_{\alpha i}} + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}}. \quad (21)$$

As already remarked in sec. 3, this problem is well posed for neutral and positively charged molecules, and, in the same way in which the term \hat{V}_{eN} prevented the total wavefunction to be a product of an electronic and a nuclear part, the term \hat{V}_{ee} in the expression above breaks the separability in the one-electron variables x_i of the electronic time-independent Schrödinger equation associated to \hat{H} . Hence, a general solution $\Psi(\underline{\mathbf{x}})$ cannot be a product of orbitals and the search must be a priori performed in the whole Hilbert space. However, this is a much too big place to look for $\Psi(\underline{\mathbf{x}})$, since the computational requirements to solve the Schrödinger equation grow exponentially on the number of electrons.

Partially recognizing this situation, in the first days of quantum mechanics, Dirac wrote that,

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these equations leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation [49].

The description of the most popular approximate methods, which the great physicist envisaged to be necessary, will be the objective of the following sections. Two basic points responsible of the relative success of such an enterprise are the severe reduction of the space in where the ground-state is sought (which, of course, leads to only an approximation of it) and the availability of computers unimaginably faster than anything that could be foreseen in times of Dirac.

6 The Hartree approximation

One of the first and most simple approximations aimed to solve the problem posed in the previous section is due to Hartree in 1927 [21] (although the way in which the Hartree equations will be derived here, using the Variational Theorem, is due to Slater [54]). In this approximation, the total wavefunction is constrained to be a product (typically referred to as *Hartree product*) of N one-electron orbitals (see

¹ Since, from now on, we will only be dealing with the ‘electronic problem’, the notation has been made simpler by dropping superfluous subindices e where there is no possible ambiguity. As a consequence, for example, the electronic Hamiltonian is now denoted by \hat{H} , the electronic kinetic energy by \hat{T} and the electronic wavefunction by $\Psi(\underline{\mathbf{x}})$ (dropping the parametric dependence on $\underline{\mathbf{R}}$ in the same spirit).

eq. (18)), where the spin of the electrons and the antisymmetry (i.e., the Pauli exclusion principle) are not taken into account²:

$$\Phi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_{i=1}^N \phi_i(\mathbf{r}_i), \quad (22)$$

where the a index in the orbitals has been substituted by i due to the fact that each function is paired to a specific set of electron coordinates, consequently being the same number of both of them.

Also, the additional requirement that the one-particle wavefunctions be normalized is imposed (see eq. (19)):

$$\langle \phi_i | \phi_i \rangle = 1, \quad i = 1, \dots, N. \quad (23)$$

With these two ingredients, we can construct the auxiliary functional whose zero-derivative condition produces the solution of the constrained stationary points problem (see eq. (20)). To this effect, we introduce N Lagrange multipliers ε_i that force the normalization constraints¹:

$$\tilde{\mathcal{F}}[\{\phi_i\}] = \left\langle \prod_{i=1}^N \phi_i(\mathbf{r}_i) \left| \hat{H} \right| \prod_{i=1}^N \phi_i(\mathbf{r}_i) \right\rangle - \sum_{i=1}^N \varepsilon_i (\langle \phi_i | \phi_i \rangle - 1), \quad (24)$$

where the minus sign in the Lagrange multipliers term is chosen in order to get to the most common form of the final equations without having to define new quantities

This functional may be considered to depend on $2N$ independent functions: the N one-electron ϕ_i and their N complex conjugates (see footnote 1 in page 8). The *Hartree equations* are then obtained by imposing that the functional derivative of $\tilde{\mathcal{F}}$ with respect to ϕ_k^* be zero for $k = 1, \dots, N$. In order to obtain them and as an appetizer for the slightly more complicated process in the more used Hartree-Fock approximation, the functional derivative will be here computed in detail following the steps indicated in appendix A.

First, we write out² the first term in the right-hand side of eq. (24):

² We shall denote with capital Greek letters the wavefunctions depending on all the electronic variables, and with lowercase Greek letters the one-electron orbitals. In addition, by Ψ (or ψ), we shall indicate wavefunctions containing spin part (called *spin-orbitals*) and, by Φ (or ϕ), those that depend only on spatial variables.

¹ Note that the normalization of the total wavefunction is a consequence of the normalization of the one-electron ones and needs not to be explicitly asked.

² The limits in sums and products are dropped if there is no possible ambiguity.

$$\begin{aligned}
& \left\langle \prod_i \phi_i(\mathbf{r}_i) \left| \hat{H} \right| \prod_i \phi_i(\mathbf{r}_i) \right\rangle = \\
& - \frac{1}{2} \sum_i \left(\prod_{j \neq i} \langle \phi_j | \phi_j \rangle \right) \int \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) d\mathbf{r} \\
& - \sum_i \left(\prod_{j \neq i} \langle \phi_j | \phi_j \rangle \right) \int \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \left(\sum_{A=1}^{N_N} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right) d\mathbf{r} \\
& + \frac{1}{2} \sum_i \sum_{j \neq i} \left(\prod_{k \neq i, j} \langle \phi_k | \phi_k \rangle \right) \iint \frac{\phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' d\mathbf{r}, \quad (25)
\end{aligned}$$

where $d\mathbf{r}$ denotes the Euclidean \mathbb{R}^3 volume element $dx dy dz$.

Now, we realize that the products outside the integrals can be dropped using the constraints in eq. (23) (see the last paragraphs of appendix B for a justification that this can be done before taking the derivative). Then, using the previous expression and conveniently rearranging the order of the integrals and sums, we write out the first term in the numerator of the left-hand side of eq. (A1) that corresponds to an infinitesimal variation of the function ϕ_k^* :

$$\begin{aligned}
& \tilde{\mathcal{F}}[\phi_k^* + \epsilon \delta \phi_k^*] := \\
& \tilde{\mathcal{F}}[\phi_1, \phi_1^*, \dots, \phi_k, \phi_k^* + \epsilon \delta \phi_k^*, \dots, \phi_N, \phi_N^*] = \\
& - \frac{1}{2} \sum_i \int \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) d\mathbf{r} - \sum_i \int |\phi_i(\mathbf{r})|^2 \left(\sum_{A=1}^{N_N} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right) d\mathbf{r} \\
& + \frac{1}{2} \sum_i \int |\phi_i(\mathbf{r})|^2 \left(\int \frac{\sum_{j \neq i} |\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) d\mathbf{r} - \sum_i \epsilon_i \left(\int |\phi_i(\mathbf{r})|^2 d\mathbf{r} - 1 \right) \\
& - \frac{1}{2} \epsilon \int \delta \phi_k^*(\mathbf{r}) \nabla^2 \phi_k(\mathbf{r}) d\mathbf{r} - \epsilon \int \delta \phi_k^*(\mathbf{r}) \phi_k(\mathbf{r}) \left(\sum_{A=1}^{N_N} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right) d\mathbf{r} \\
& + \epsilon \int \delta \phi_k^*(\mathbf{r}) \phi_k(\mathbf{r}) \left(\int \frac{\sum_{i \neq k} |\phi_i(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) d\mathbf{r} - \epsilon \epsilon_k \int \delta \phi_k^*(\mathbf{r}) \phi_k(\mathbf{r}) d\mathbf{r}. \quad (26)
\end{aligned}$$

We subtract from this expression the quantity $\tilde{\mathcal{F}}[\{\phi_i(\mathbf{r}_i)\}]$, so that the first four terms cancel, and we can write

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{\tilde{\mathcal{F}}[\phi_k^* + \epsilon \delta \phi_k^*] - \tilde{\mathcal{F}}[\phi_k^*]}{\epsilon} = \\ \int \left[-\frac{1}{2} \nabla^2 \phi_k(\mathbf{r}) - \left(\sum_{A=1}^{N_N} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right) \phi_k(\mathbf{r}) \right. \\ \left. + \left(\int \frac{\sum_{i \neq k} |\phi_i(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \phi_k(\mathbf{r}) - \varepsilon_k \phi_k(\mathbf{r}) \right] \delta \phi_k^*(\mathbf{r}) d\mathbf{r} . \end{aligned} \quad (27)$$

Now, by simple inspection of the right-hand side, we see that the functional derivative (see eq. (A1)) is the part enclosed by square brackets:

$$\frac{\delta \tilde{\mathcal{F}}[\{\phi_i\}]}{\delta \phi_k^*} = \left(-\frac{1}{2} \nabla^2 + \hat{V}_e(\mathbf{r}) + \hat{V}_e^k(\mathbf{r}) - \varepsilon_k \right) \phi_k(\mathbf{r}) , \quad (28)$$

where the *nuclear potential energy* and the *electronic potential energy* have been respectively defined as¹

$$\hat{V}_N(\mathbf{r}) := - \sum_{A=1}^{N_N} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} , \quad (29a)$$

$$\hat{V}_e^k(\mathbf{r}) := \int \frac{\sum_{i \neq k} |\phi_i(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' . \quad (29b)$$

Finally, if we ask the functional derivative to be zero for $k = 1, \dots, N$, we arrive to the equations that the stationary points must satisfy, the *Hartree equations*:

$$\hat{\mathcal{H}}_k[\phi] \phi_k(\mathbf{r}) := \left(-\frac{1}{2} \nabla^2 + \hat{V}_N(\mathbf{r}) + \hat{V}_e^k(\mathbf{r}) \right) \phi_k(\mathbf{r}) = \varepsilon_k \phi_k(\mathbf{r}) , \quad (30)$$

for all $k = 1, \dots, N$.

Let us note that, despite the fact that the object $\hat{\mathcal{H}}_k[\phi]$ defined above is not a operator strictly speaking, since, as the notation emphasizes, it depends on the orbitals $\phi_{i \neq k}$, we will stick to the name *Hartree operator* for it, in order to be consistent with most of the literature.

Now, some remarks related to the Hartree equations are worth making. First, it can be shown that, if the variational ansatz in eq. (22) included the spin degrees of freedom of the electrons, all the expressions above would be kept, simply changing the orbitals $\phi_i(\mathbf{r}_i)$ by the spin-orbitals $\psi_i(\mathbf{r}_i, \sigma_i)$.

Secondly, and moving into more conceptual playgrounds, we note that the special structure of $\hat{V}_e^k(\mathbf{r})$ in eq. (29b) makes it mandatory to interpret the Hartree scheme as one in which each electron ‘feels’ only the average effect of the rest. In fact, if the *quantum charge density* $\rho_i(\mathbf{r}) := |\phi_i(\mathbf{r})|^2$ is regarded for a moment as a classical continuum distribution, then the potential produced by all the electrons but the k -th is precisely the one in eq. (29b). Supporting this image, note also the fact

¹ Compare the notation with the one in eqs. (3), here a subindex e has been dropped to distinguish the new objects defined.

that, if we write the joint probability density of electron 1 being at the point \mathbf{r}_1 , electron 2 being at the point \mathbf{r}_2 and so on (simply squaring eq. (22)),

$$\rho(\mathbf{r}_1, \dots, \mathbf{r}_N) := |\Phi(\mathbf{r}_1, \dots, \mathbf{r}_N)|^2 = \prod_{i=1}^N |\phi_i(\mathbf{r}_i)|^2 = \prod_{i=1}^N \rho_i(\mathbf{r}_i), \quad (31)$$

we see that, *in a probabilistic sense*, the electrons are independent (they could not be independent in a physical, complete sense, since we have already said that they ‘see’ each other in an average way).

Anyway, despite these appealing images and also despite the fact that, disguised under the misleading (albeit common) notation, these equations seem ‘one-particle’, they are rather complicated from a mathematical point of view. On the one hand, it is true that, whereas the original electronic Schrödinger equation in (9a) depended on $3N$ spatial variables, the expressions above only depend on 3. This is what we have gained from drastically reducing the search space to the set of Hartree products in eq. (22) and what renders the approximation tractable. On the other hand, however, we have paid the price of greatly increasing the mathematical complexity of the expressions, so that, while the electronic Schrödinger equation was one linear differential equation, the Hartree ones in (30) are N coupled non-linear integro-differential equations [55].

This complexity precludes any analytical approach to the problem and forces us to look for the solutions using less reliable iterative methods. Typically, in computational studies, one proposes a *starting guess* for the set of N orbitals $\{\phi_k^0\}$; with them, the Hartree operator $\hat{\mathcal{H}}_k[\phi^0]$ in the left-hand side of eq. (30) is constructed for every k and the N equations are solved as simple eigenvalue problems. For each k , the ϕ_k^1 that corresponds to the lowest ε_k^1 is selected and a new Hartree operator $\hat{\mathcal{H}}_k[\phi^1]$ is constructed with the $\{\phi_k^1\}$. The process is iterated until (hopefully) the n -th set of solutions $\{\phi_k^n\}$ differ from the $(n-1)$ -th one $\{\phi_k^{n-1}\}$ less than a reasonably small amount.

Many technical issues exist that raise doubts about the possible success of such an approach. The most important ones being related to the fact that a proper definition of the *Hartree problem* should be: *find the global minimum of the energy functional $\langle \Phi | \hat{H} | \Phi \rangle$ under the constraint that the wavefunction Φ be a Hartree product* and not: *solve the Hartree equations (30)*, whose solutions indeed include the global minimum sought but also all the rest of stationary points.

While the possibility that a found solution be a maximum or a saddle point can be typically ruled out [50, 55], as we remarked in sec. 4 and due to the fact that there are an infinite number of solutions to the Hartree equations [56], to be sure that any found minimum is the global one is impossible in a general case. There exists, however, one way, related to a theorem by Simon and Lieb [57, 58], of hopefully biasing a particular found solution of the Hartree equations to be the global minimum that we are looking for. They showed that, first, for neutral or positively charged molecules ($Z \geq N$), the Hartree global minimization problem has a solution (its uniqueness is not established yet [55]) and, second, that the minimizing orbitals $\{\phi_k\}$ correspond to the lowest eigenvalues of the $\hat{\mathcal{H}}_k[\phi]$ operators self-consistently constructed with them¹. Now, although *the reverse of the second*

¹ In quantum chemistry, where the number of electrons considered is typically small, the version of the Hartree equations that is used is the one derived here, with the Hartree operators depending on the index k in a non-trivial way. However, if the number of electrons is large enough (such as in condensed matter applications), is customary to add to the effective electronic repulsion in eq. (29b) the self-interaction of

part of the theorem is not true in general (i.e., from the fact that a particular set of orbitals are the eigenstates corresponding to the lowest eigenvalues of the associated Hartree operators, does not necessarily follow that they are the ones that minimize the energy) [55], in practice, the insight provided by Lieb and Simon's result is invoked to build each successive state in the iterative procedure described above, choosing the lowest lying eigenstates each time. In this way, although one cannot be sure that the global minimum has been reached, the fact that the found one has a property that the former also presents is regarded as a strong hint that it must be so (see also the discussion for the Hartree-Fock case in the next section).

This drawback and all the problems arising from the fact that an iterative procedure such as the one described above could converge to a fixed point, oscillate eternally or even diverge, are circumvented in practice by a clever choice of the starting guess orbitals $\{\phi_k^0\}$. If they are extracted, for example, from a slightly less accurate theory, one may expect that they could be 'in the basin of attraction' of the true Hartree minimum (so that the stationary point found will be the correct one) and close to it (so that the iterative procedure will converge). This kind of wishful thinking combined with large amounts of heuristic protocols born from many decades of trial-and-error-derived knowledge pervade and make possible the whole quantum chemistry discipline.

7 The Hartree-Fock approximation

The Hartree theory discussed in the previous section is not much used in quantum chemistry and many textbooks on the subject do not even mention it. Although it contains the seed of almost every concept underlying the *Hartree-Fock approximation* discussed in this section, it lacks an ingredient that turns out to be essential to correctly describe the behaviour of molecular species: *the indistinguishability of the electrons*. This was noticed independently by Fock [59] and Slater [54] in 1930, and it was corrected by proposing a variational ansatz for the total wavefunction that takes the form of a so-called *Slater determinant* (see eq. (33) below).

The most important mathematical consequence of the indistinguishability among a set of N quantum objects of the same type is the requirement that the total N -particle wavefunction must either remain unchanged (*symmetric*) or change sign (*antisymmetric*) when any pair of coordinates, x_i and x_j , are swapped. In the first case, the particles are called *bosons* and must have integer spin, while in the second case, they are called *fermions* and have semi-integer spin. Electrons are fermions, so the total wavefunction must be antisymmetric under the exchange of any pair of one-electron coordinates. This is a property that is certainly not met by the single Hartree product in eq. (22) but that can be easily implemented by forming linear combinations of many of them. The trick is to add all the possible Hartree products that are obtained from eq. (22) changing the order of the orbitals labels while keeping the order of the coordinates ones¹, and assigning to each term the sign of the permutation p needed to go from the natural order $1, \dots, N$ to the corresponding one $p(1), \dots, p(N)$. The sign of a permutation p is 1 if p can be written as a composition of an even number of two-element transpositions, and it is -1 if the number of transpositions needed is odd. Therefore, we define $\mathcal{T}(p)$ as

electron k with himself. In such a case, the Hartree operator is independent of k so that, after having achieved self-consistency, the orbitals ϕ_k turn out to be eigenstates corresponding to different eigenvalues of the same Hermitian operator, $\mathcal{H}[\phi]$, and, therefore, mutually orthogonal.

¹ It is immaterial whether the orbitals labels are kept and the coordinates ones changed or vice versa.

the minimum number² of transpositions needed to perform the permutation p , and we write the sign of p as $(-1)^{T(p)}$.

Using this, an antisymmetric wavefunction constructed from Hartree products of N different orbitals may be written as

$$\Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \sum_{p \in S_N} (-1)^{T(p)} \psi_{p(1)}(x_1) \cdots \psi_{p(N)}(x_N), \quad (32)$$

where the factor $1/\sqrt{N!}$ enforces normalization of the total wavefunction Ψ (if we use the constraints in eq. (34)) and S_N denotes the *symmetric group* of order N , i.e., the set of all permutations of N elements (with a certain multiplication rule).

The above expression is more convenient to perform the calculations that lead to the Hartree-Fock equations, however, there is also a compact way of rewriting eq. (32) which is commonly found in the literature and that is useful to illustrate some particular properties of the problem. It is the *Slater determinant*:

$$\Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(x_1) & \psi_2(x_1) & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \psi_2(x_2) & \cdots & \psi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(x_N) & \psi_2(x_N) & \cdots & \psi_N(x_N) \end{vmatrix}. \quad (33)$$

Now, having established the constraints on the form of the total wavefunction, we ask the Hartree-Fock one-electron orbitals to be, not only normalized, like we did in the Hartree case, but also mutually orthogonal:

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}, \quad i, j = 1, \dots, N. \quad (34)$$

Additionally note that, contrarily to what we did in the previous section, we have now used one-electron wavefunctions ψ_i dependent also on the spin σ (i.e., spin-orbitals) to construct the variational ansatz. A *general spin-orbital*¹ may be written as (see also footnote 1 in page 4)

$$\psi(x) = \varphi^\alpha(\mathbf{r}) \alpha(\sigma) + \varphi^\beta(\mathbf{r}) \beta(\sigma), \quad (35)$$

where the functions α and β correspond to the *spin-up* and *spin-down* eigenstates of the operator associated to the z -component of the one-electron spin. They are defined as

$$\begin{aligned} \alpha(-1/2) &= 0 & \beta(-1/2) &= 1 \\ \alpha(1/2) &= 1 & \beta(1/2) &= 0. \end{aligned} \quad (36)$$

² It can be shown that the parity of all decompositions of p into products of elementary transpositions is the same. We have chosen the minimum only for $T(p)$ to be well defined.

¹ Note that, if we had not included the spin degrees of freedom, the search space would have been half as large, since, where we now have $2N$ functions of \mathbf{r} (i.e., $\varphi_i^\alpha(\mathbf{r})$ and $\varphi_i^\beta(\mathbf{r})$, with $i = 1, \dots, N$), we would have had just N (the $\phi_i(\mathbf{r})$).

The formalism obtained when these general spin-orbitals are used is accordingly called General Hartree-Fock (GHF). The first part of the mathematical treatment in the following paragraphs shall be performed assuming this situation. The advantage of such a choice is that, later on, by imposing additional constraints to the spin part of the one-electron orbitals, we will be able to derive the basic equations for some other flavours of the Hartree-Fock theory, such as UHF, RHF and ROHF, in a very direct way.

Now, to calculate the expected value of the energy in a state such as the one in eqs. (32) and (33), let us denote the one-particle part (that operates on the i -th coordinates) of the total electronic Hamiltonian \hat{H} in eq. (21) by

$$\hat{h}_i := -\frac{\nabla_i^2}{2} - \sum_{\alpha=1}^{N_N} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_i|}, \quad (37)$$

in such a way that,

$$\langle \Psi | \hat{H} | \Psi \rangle = \sum_i \langle \Psi | \hat{h}_i | \Psi \rangle + \frac{1}{2} \sum_{i \neq j} \langle \Psi | \frac{1}{r_{ij}} | \Psi \rangle, \quad (38)$$

where $r_{ij} := |\mathbf{r}_j - \mathbf{r}_i|$.

We shall compute separately each one of the sums in the expression above. Let us start now with the first one: For a given i in the sum, and due to the structure of the electronic wavefunction in (32), the expected value $\langle \Psi | \hat{h}_i | \Psi \rangle$ is a sum of $(N!)^2$ terms of the form

$$\frac{1}{N!} (-1)^{\mathcal{T}(p)+\mathcal{T}(p')} \langle \psi_{p(1)}(x_1) \cdots \psi_{p(N)}(x_N) | \hat{h}_i | \psi_{p'(1)}(x_1) \cdots \psi_{p'(N)}(x_N) \rangle, \quad (39)$$

but, since \hat{h}_i operates only on x_i and due to the orthogonality of the spin-orbitals with different indices, we have that the only non-zero terms are those with $p = p'$. Taking this into account, all permutations p appear still as terms of the sum, and we see that every orbital ψ_j occurs depending on every coordinate x_i . Given a particular pair i and j , this happens in the terms for which $p(i) = j$ and one of such terms may be expressed as

$$\frac{1}{N!} \left(\prod_{k \neq j} \langle \psi_k | \psi_k \rangle \right) \langle \psi_j | \hat{h}_i | \psi_j \rangle, \quad (40)$$

where we have used that $(-1)^{2\mathcal{T}(p)} = 1$, and we have dropped the index i from \hat{h}_i noticing that the integration variables in $\langle \psi_j(x_i) | \hat{h}_i | \psi_j(x_i) \rangle$ are actually dummy.

Next, we use again the one-electron wavefunctions constraints in eq. (34) to remove the product of norms in brackets, and we realize that, for each j , there are as many terms like the one in the expression above as permutations of the remaining $N - 1$ orbital indices (i.e., $(N - 1)!$). In addition, we recall that every j must appear and perform the first sum in eq. (38), yielding

$$\begin{aligned} \sum_i \langle \Psi | \hat{h}_i | \Psi \rangle &= \sum_i (N-1)! \sum_j \frac{1}{N!} \langle \psi_j | \hat{h} | \psi_j \rangle = \\ N(N-1)! \sum_j \frac{1}{N!} \langle \psi_j | \hat{h} | \psi_j \rangle &= \sum_j \langle \psi_j | \hat{h} | \psi_j \rangle, \end{aligned} \quad (41)$$

where all factorial terms have canceled out.

The next step is to calculate the second sum in eq. (38). Again, we have that, for each pair (i, j) , $\langle \Psi | 1/r_{ij} | \Psi \rangle$ is a sum of $(N!)^2$ terms like

$$\frac{1}{N!} (-1)^{\mathcal{T}(p)+\mathcal{T}(p')} \langle \psi_{p(1)}(x_1) \cdots \psi_{p(N)}(x_N) | \frac{1}{r_{ij}} | \psi_{p'(1)}(x_1) \cdots \psi_{p'(N)}(x_N) \rangle. \quad (42)$$

For this expected value, contrarily to the case of \hat{h}_i and due to the two-body nature of the operator $1/r_{ij}$, not only do the terms with $p = p'$ survive, but also those in which p and p' differ over only a pair of values i and j , i.e., those for which $p(i) = p'(j)$, $p(j) = p'(i)$ and $p(k) = p'(k), \forall k \neq i, j$. The reason for this is that, even if $p(i) \neq p'(i)$ and $p(j) \neq p'(j)$, the integral $\langle \psi_{p(i)}(x_i) \psi_{p(j)}(x_j) | 1/r_{ij} | \psi_{p'(i)}(x_i) \psi_{p'(j)}(x_j) \rangle$ does not vanish.

Now, using that $1/r_{ij}$ operates only on x_i and x_j , the orthonormality conditions in eq. (34) and the fact that $(-1)^{2\mathcal{T}(p)} = 1$, we have that, when ψ_k depends on x_i and ψ_l depends on x_j , the $p = p'$ part of the corresponding terms in eq. (42) reads

$$\frac{1}{N!} \langle \psi_k \psi_l | \frac{1}{r} | \psi_k \psi_l \rangle, \quad (43)$$

where we have defined

$$\langle \psi_i \psi_j | \frac{1}{r} | \psi_k \psi_l \rangle := \sum_{\sigma, \sigma'} \iint \frac{\psi_i^*(x) \psi_j^*(x') \psi_k(x) \psi_l(x')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (44)$$

Next, we see that, for each pair (k, l) , and keeping $p = p'$, we can make $(N-2)!$ permutations among the $N-2$ indices of the orbitals on which $1/r_{ij}$ does not operate and still find the same expression (43). Therefore, for each pair (k, l) , we have a sum of $(N-2)!$ identical terms. In addition, if we perform the sum on i and j in eq. (38) and remark that the term in eq. (43) does not depend on the pair (i, j) (which is obvious from the suggestive notation above), we have that the $p = p'$ part of the second sum in eq. (38), which is typically called *Coulomb energy*, reads

$$\frac{1}{2} \sum_{i \neq j} (N-2)! \sum_{k \neq l} \frac{1}{N!} \langle \psi_k \psi_l | \frac{1}{r} | \psi_k \psi_l \rangle = \frac{1}{2} \sum_{k \neq l} \langle \psi_k \psi_l | \frac{1}{r} | \psi_k \psi_l \rangle, \quad (45)$$

where we have used that the sum $\sum_{i \neq j}$ is performed on $N(N-1)$ identical terms which do not depend neither on i nor on j .

On the other hand, in the case in which p and p' only differ in that the indices of the orbitals that depend on x_i and x_j are swapped, all the derivation above applies

except for the facts that, first, $(-1)^{\mathcal{T}(p)+\mathcal{T}(p')} = -1$ and, second, the indices k and l must be *exchanged* in eq. (45) (it is immaterial if they are exchanged in the bra or in the ket, since the indices are summed over and are dummy). Henceforth, the remaining part of the second sum in eq. (38), typically termed *exchange energy*, may be written as

$$-\frac{1}{2} \sum_{k \neq l} \langle \psi_k \psi_l | \frac{1}{r} | \psi_l \psi_k \rangle . \quad (46)$$

Finally, the expected value of the energy in the GHF variational state Ψ turns out to be

$$E^{\text{GHF}} := \langle \Psi | \hat{H} | \Psi \rangle = \sum_i \underbrace{\langle \psi_i | \hat{h} | \psi_i \rangle}_{h_i} + \frac{1}{2} \sum_{i,j} \left(\underbrace{\langle \psi_i \psi_j | \frac{1}{r} | \psi_i \psi_j \rangle}_{J_{ij}} - \underbrace{\langle \psi_i \psi_j | \frac{1}{r} | \psi_j \psi_i \rangle}_{K_{ij}} \right) , \quad (47)$$

where the *one-electron integrals* h_i have been defined together with the *two-electron integrals*, J_{ij} and K_{ij} , and the fact that $J_{ii} = K_{ii}, \forall i$ has been used to include the diagonal terms in the second sum.

Now, the energy functional above is the quantity that we want to minimize under the orthonormality constraints in eq. (34). So we are prepared to write the auxiliary functional $\tilde{\mathcal{F}}$, introducing N^2 Lagrange multipliers λ_{ij} (see eq. (20) and compare with the Hartree example in the previous section):

$$\tilde{\mathcal{F}} [\{\psi_i\}] = \sum_i h_i + \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}) - \sum_{i,j} \lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}) . \quad (48)$$

In order to get to the *Hartree-Fock equations* that the stationary orbitals ψ_k must satisfy, we impose that the functional derivative of $\tilde{\mathcal{F}} [\{\psi_i\}]$ with respect to ψ_k^* be zero. To calculate $\delta \tilde{\mathcal{F}} / \delta \psi_k^*$, we follow the procedure described in appendix A, using the same notation as in eq. (26). The variation with respect to each ψ_k^* shall yield the Hartree-Fock equations for the unconjugated ψ_i . The equations for the ψ_i^* are obtained either by differentiating $\tilde{\mathcal{F}} [\{\psi_i\}]$ with respect to each ψ_k or, if the λ_{ij} matrix is Hermitian (which will turn out to be the case), by simply taking the complex conjugate of both sides of the final equations in (53).

Now,

$$\begin{aligned}
& \lim_{\epsilon \rightarrow 0} \frac{\tilde{\mathcal{F}}[\psi_k^* + \epsilon \delta \psi_k^*] - \tilde{\mathcal{F}}[\psi_k^*]}{\epsilon} = \\
& \langle \delta \psi_k | \hat{h} | \psi_k \rangle + \sum_j \left(\langle \delta \psi_k \psi_j | \frac{1}{r} | \psi_k \psi_j \rangle - \langle \delta \psi_k \psi_j | \frac{1}{r} | \psi_j \psi_k \rangle \right) - \sum_j \lambda_{kj} \langle \delta \psi_k | \psi_j \rangle = \\
& \int \left[\hat{h} \psi_k(x) + \sum_j \left(\psi_k(x) \int \frac{|\psi_j(x')|^2}{|\mathbf{r} - \mathbf{r}'|} dx' - \psi_j(x) \int \frac{\psi_j^*(x') \psi_k(x')}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \right. \\
& \left. - \sum_j \lambda_{kj} \psi_j(x) \right] \delta \psi_k^*(x) dx, \tag{49}
\end{aligned}$$

where we have used the more compact notation $\int dx$ instead of $\sum_{\sigma} \int d\mathbf{r}$.

Then, like in the previous section, by simple inspection of the right-hand side, we see that the functional derivative is the part enclosed by square brackets (see eq. (A1)):

$$\frac{\delta \tilde{\mathcal{F}}[\{\psi_i\}]}{\delta \psi_k^*} = \left[\hat{h} + \sum_j \left(\hat{J}_j[\psi] - \hat{K}_j[\psi] \right) \right] \psi_k(x) - \sum_j \lambda_{kj} \psi_j(x), \tag{50}$$

where the *Coulomb* and *exchange operators* are respectively defined by their action on an arbitrary function $\varphi(x)$ as follows¹:

$$\hat{J}_j[\psi] \varphi(x) := \left(\int \frac{|\psi_j(x')|^2}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \varphi(x), \tag{51a}$$

$$\hat{K}_j[\psi] \varphi(x) := \left(\int \frac{\psi_j^*(x') \varphi(x')}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \psi_j(x). \tag{51b}$$

Therefore, if we define the GHF *Fock operator* as

$$\hat{F}^{\text{GHF}}[\psi] := \hat{h} + \sum_j \left(\hat{J}_j[\psi] - \hat{K}_j[\psi] \right), \tag{52}$$

we arrive to a first version of the *Hartree-Fock equations* by asking that the functional derivative in eq. (50) be zero:

$$\hat{F}^{\text{GHF}}[\psi] \psi_i(x) = \sum_j \lambda_{ij} \psi_j(x), \quad i = 1, \dots, N. \tag{53}$$

Now, in order to obtain a simpler version of them, we shall take profit from the fact that the whole problem is invariant under a unitary transformation among the one-electron orbitals.

¹ Like in the Hartree case in the previous section, the word *operator* is a common notational abuse if they act upon the very ψ_i on which they depend. This is again made explicit in the notation.

If we repeat the calculation in eq. (49) but varying ψ_k this time, instead of ψ_k^* , and use the following relation:

$$\int \psi_i^*(x) \hat{h} \psi_j(x) dx = \int [\hat{h} \psi_i^*(x)] \psi_j(x) dx, \quad (54)$$

we arrive to the GHF equations for the conjugated spin-orbitals:

$$\hat{F}^{\text{GHF}}[\psi] \psi_i^*(x) = \sum_j \lambda_{ji} \psi_j^*(x). \quad i = 1, \dots, N. \quad (55)$$

Then, we may subtract the complex conjugate of eq. (55) from eq. (53) yielding

$$\sum_j (\lambda_{ij} - \lambda_{ji}^*) \psi_j(x) = 0, \quad k = 1, \dots, N. \quad (56)$$

Therefore, since the set of the ψ_j is orthogonal and hence linearly independent, we have that the $N \times N$ matrix $\Lambda := (\lambda_{ij})$ of Lagrange multipliers is Hermitian:

$$\lambda_{ij} = \lambda_{ji}^*, \quad k, j = 1, \dots, N. \quad (57)$$

This actually means that we have a set of three equations that the stationary spin-orbitals satisfy, but only two of them are independent. These equations are the GHF equations for the ψ_i and ψ_i^* , in (53) and (55), respectively, and (57). Any pair of them could be in principle be chosen as the basic equations, however, in common practice the first and the last one of them are typically picked.

In any case, due to (57), a unitary matrix U exists that *diagonalizes* Λ ; in the sense that $\varepsilon := U^{-1} \Lambda U = U^+ \Lambda U$ is a diagonal matrix, i.e., $\varepsilon_{ij} = \delta_{ij} \varepsilon_i$. Using this unitary matrix U , we can transform the set of orbitals $\{\psi_i\}$ into a new one $\{\psi'_i\}$:

$$\psi_k(x) = \sum_j U_{kj} \psi'_j(x). \quad (58)$$

This transformation is physically legitimate since it only changes the N -electron wavefunction Ψ in an unmeasurable phase $e^{i\phi}$. To see this, let us denote by S_{ij} the (ij) -element of the matrix inside the Slater determinant in eq. (33), i.e., $S_{ij} := \psi_j(x_i)$. Then, after using the expression above, the (ij) -element of the new matrix S' can be related to the old ones via $S_{ki} = \sum_j U_{kj} S'_{ij}$, in such a way that $S = S' U^T$ and the desired result follows:

$$\Psi(\{\psi_i\}) = \frac{\det S}{\sqrt{N!}} = \frac{\det(S' U^T)}{\sqrt{N!}} = \frac{\det S' \det U^T}{\sqrt{N!}} = e^{i\phi} \Psi(\{\psi'_i\}). \quad (59)$$

Now, we insert eq. (58) into the first version of the Hartree-Fock equations in (53):

$$\hat{F}^{\text{GHF}}[U\psi'] \left(\sum_j U_{ij} \psi'_j(x) \right) = \sum_{j,k} \lambda_{ij} U_{jk} \psi'_k(x), \quad i = 1, \dots, N. \quad (60)$$

Next, we multiply by U_{li}^{-1} each one of the N expressions and sum in i :

$$\begin{aligned} \hat{F}^{\text{GHF}}[U\psi'] \left(\sum_{i,j} \underbrace{U_{li}^{-1} U_{ij}}_{\delta_{lj}} \psi'_j(x) \right) &= \sum_{i,j,k} \underbrace{U_{li}^{-1} \lambda_{ij} U_{jk}}_{\varepsilon_{lk} = \delta_{lk} \varepsilon_l} \psi'_k(x) \\ \implies \hat{F}^{\text{GHF}}[U\psi'] \psi'_l(x) &= \varepsilon_l \psi'_l(x), \quad l = 1, \dots, N. \end{aligned} \quad (61)$$

Although this new version of the Hartree-Fock equations can be readily seen as a *pseudo-eigenvalue problem* and solved by the customary iterative methods, we can go a step further and show that, like the N -particle wavefunction Ψ (see eq. (59)), the Fock operator $\hat{F}^{\text{GHF}}[\psi]$, as a function of the one-electron orbitals, is invariant under a unitary transformation such as the one in eq. (58). In fact, this is true for each one of the sums of Coulomb and exchange operators in eq. (52) separately:

$$\begin{aligned} \sum_j \hat{J}_j[U\psi'] \varphi(x) &= \sum_j \left(\int \frac{|\sum_k U_{jk} \psi'_k(x')|^2}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \varphi(x) = \\ &= \sum_j \left(\int \frac{\sum_{k,l} U_{jk}^* U_{jl} \psi'_k(x') \psi'_l(x')}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \varphi(x) = \\ &= \left(\int \frac{\sum_{j,k,l} U_{kj}^{-1} U_{jl} \psi'_k(x') \psi'_l(x')}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \varphi(x) = \\ &= \sum_k \left(\int \frac{|\psi'_k(x')|^2}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \varphi(x) = \sum_j \hat{J}_j[\psi'] \varphi(x), \quad \forall \varphi(x), \end{aligned} \quad (62)$$

where, in the step before the last, we have summed on j and l , using that $\sum_j U_{kj}^{-1} U_{jl} = \delta_{kl}$.

Performing very similar calculations, one can show that

$$\sum_j \hat{K}_j[U\psi'] \varphi(x) = \sum_j \hat{K}_j[\psi'] \varphi(x), \quad \forall \varphi(x), \quad (63)$$

and therefore, that $F^{\text{GHF}}[U\psi'] = F^{\text{GHF}}[\psi']$. In such a way that any unitary transformation on a set of orbitals that constitute a solution of the Hartree-Fock equations in (53) yields a different set that is also a solution of *the same* equations. For computational and conceptual reasons (see, for example, Koopmans' Theorem below), it turns out to be convenient to use this freedom and choose the matrix U in such a way that the Lagrange multipliers matrix is diagonalized (see eq. (56) and the paragraph below it). The particular set of one-electron orbitals $\{\psi'_i\}$ obtained

with this U are called *canonical orbitals* and their use is so prevalent that we will circumscribe the forecoming discussion to them and drop the prime from the notation.

Using the canonical orbitals, the *Hartree-Fock equations* can be written as

$$\hat{F}^{\text{GHF}}[\psi] \psi_i(x) = \varepsilon_i \psi_i(x) , \quad i = 1, \dots, N . \quad (64)$$

Many of the remarks related to these equations are similar to those made about the Hartree ones in (30), although there exist important differences due to the inclusion of the indistinguishability of the electrons in the variational ansatz. This is clearly illustrated if we calculate the joint probability density, associated to a wavefunction like the one in eq. (32), of the coordinates with label 1 taking the value x_1 , the coordinates with label 2 taking the value x_2 , and so on:

$$\begin{aligned} \rho^{\text{GHF}}(x_1, \dots, x_N) &= |\Psi(x_1, \dots, x_N)|^2 = \\ &= \frac{1}{N!} \sum_{p, p' \in S_N} (-1)^{\mathcal{T}(p)+\mathcal{T}(p')} \psi_{p(1)}^*(x_1) \cdots \psi_{p(N)}^*(x_N) \psi_{p'(1)}(x_1) \cdots \psi_{p'(N)}(x_N) . \end{aligned} \quad (65)$$

If we compare this expression with eq. (31), we see that the antisymmetry of Ψ has completely spoiled the statistical independence among the one-electron coordinates. However, there is a weaker quasi-independence that may be recovered: If, using the same reasoning about permutations that took us to the one-electron part $\sum_i \langle \Psi | \hat{h}_i | \Psi \rangle$ of the energy functional in page 18, we calculate the marginal probability density of the i -th coordinates taking the value x_i , we find

$$\rho_i^{\text{GHF}}(x_i) := \int \left(\prod_{k \neq i} dx_k \right) \rho^{\text{GHF}}(x_1, \dots, x_N) = \frac{1}{N} \sum_j |\psi_j(x_i)|^2 . \quad (66)$$

Now, since the coordinates indices are just immaterial labels, the actual probability density of finding *any* electron with coordinates x is given by

$$\rho^{\text{GHF}}(x) := \sum_i \rho_i^{\text{GHF}}(x) = \sum_i |\psi_i(x)|^2 , \quad (67)$$

which can be interpreted as a *charge density* (except for the sign), as, in atomic units, the charge of the electron is $e = -1$. The picture being consistent with the fact that $\rho^{\text{GHF}}(x)$ is normalized to the number of electrons N :

$$\int \rho^{\text{GHF}}(x) dx = N . \quad (68)$$

Additionally, if we perform the same type of calculations that allowed to calculate the two-electron part of the energy functional in page 19, we have that the two-body probability density of the i -th coordinates taking the value x_i and of the j -th coordinates taking the value x_j reads

$$\rho_{ij}^{\text{GHF}}(x_i, x_j) := \int \left(\prod_{k \neq i, j} dx_k \right) \rho^{\text{GHF}}(x_1, \dots, x_N) =$$

$$\frac{1}{N(N-1)} \left(\sum_{k,l} |\psi_k(x_i)|^2 |\psi_l(x_j)|^2 - \sum_{k,l} \psi_k^*(x_i) \psi_l^*(x_j) \psi_l(x_i) \psi_k(x_j) \right), \quad (69)$$

and, if we reason in the same way as in the case of $\rho_i^{\text{GHF}}(x_i)$, in order to get to the probability density of finding *any* electron with coordinates x at the same time that *any other* electron has coordinates x' , we must multiply the function above by $N(N-1)/2$, which is the number of immaterial (i, j) -labelings, taking into account that the distinction between x and x' is also irrelevant:

$$\rho^{\text{GHF}}(x, x') := \frac{N(N-1)}{2} \rho_{ij}^{\text{GHF}}(x, x') =$$

$$\frac{1}{2} \left(\sum_k |\psi_k(x)|^2 \sum_l |\psi_l(x')|^2 - \sum_{k,l} \psi_k^*(x) \psi_l^*(x') \psi_l(x) \psi_k(x') \right). \quad (70)$$

Finally, taking eq. (67) to this one, we have

$$\rho^{\text{GHF}}(x, x') = \frac{1}{2} \left(\rho^{\text{GHF}}(x) \rho^{\text{GHF}}(x') - \sum_{k,l} \psi_k^*(x) \psi_l^*(x') \psi_l(x) \psi_k(x') \right), \quad (71)$$

where the first term corresponds to independent electrons and the second one, called the *interference term*, could be interpreted as an exchange correction.

Although, in general, this is the furthest one may go, when additional constraints are imposed on the spin part of the one-electron wavefunctions (see the discussion about Restricted Hartree-Fock in the following pages, for example), the exchange correction in eq. (71) above vanishes for electrons of opposite spin, i.e., electrons of opposite spin turn out to be pairwise independent. However, whereas it is true that more correlation could be added to the Hartree-Fock results by going to higher levels of the theory and, in this sense, Hartree-Fock could be considered the first step in the ‘correlation ladder’, one should not regard it as an ‘uncorrelated’ approximation, since, even in the simplest case of RHF (see below), Hartree-Fock electrons (of the same spin) are statistically correlated. All of this has its roots in the *Pauli principle*, which states that no pair of electrons can share all the quantum numbers.

Let us now point out that, like in the Hartree case, the left-hand side of the Hartree-Fock equations in (64) is a complicated, non-linear function of the orbitals $\{\psi_i\}$ and the notation chosen is intended only to emphasize the nature of the iterative protocol that is typically used to solve the problem. However, note that, while the Hartree operator $\hat{\mathcal{H}}_k[\phi]$ depended on the index of the orbital ϕ_k on which it acted, the Fock operator in eq. (64) is the same for all the spin-orbitals ψ_i . This is due to the inclusion of the $i = j$ terms in the sum of the Coulomb and exchange two-electron integrals in eq. (47) and it allows to perform the iterative procedure solving only one eigenvalue problem at each step, instead of N of

them like in the Hartree case (see however the UHF and ROHF versions of the Hartree-Fock problem in what follows).

The one-particle appearance of eqs. (64) is again strong and, whereas the ‘eigenvalues’ ε_i are not the energies of the individual electrons, they are called *orbital energies* due to the physical meaning they receive via the well-known *Koopmans’ Theorem* [60].

To get to this result, let us multiply eq. (64) from the left by $\psi_i(x)$, for a given i , and then integrate over x . Using the definition of the Fock operator in eq. (52) together with the Coulomb and exchange ones in eqs. (51), we obtain

$$\langle \psi_i | \hat{F}^{\text{GHF}} | \psi_i \rangle = h_i + \sum_j (J_{ij} - K_{ij}) = \varepsilon_i, \quad i = 1, \dots, N, \quad (72)$$

where we have used the same notation as in eq. (47) and the fact that the one-electron orbitals are normalized.

If we next sum on i and compare the result with the expression in eq. (47), we found that the relation of the eigenvalues ε_i with the actual Hartree-Fock energy is given by

$$E^{\text{GHF}} = \sum_i \varepsilon_i - \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}). \quad (73)$$

Finally, if we assume that upon ‘removal of an electron from the k -th orbital’ the rest of the orbitals will remain unmodified, we can calculate the *ionization energy* using the expression in (47) together with the equations above:

$$\begin{aligned} \Delta E^{\text{GHF}} &:= E_{N-1}^{\text{GHF}} - E_N^{\text{GHF}} = \sum_{i \neq k} h_i - \sum_i h_i + \frac{1}{2} \sum_{i,j \neq k} (J_{ij} - K_{ij}) \\ &- \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}) = -h_k - \sum_j (J_{kj} - K_{kj}) = -\varepsilon_k, \end{aligned} \quad (74)$$

and this is Koopmans’ Theorem, namely, that *the k -th ionization energy in the frozen-orbitals approximation is ε_k .*

Moving now to the issue about the solution of the Hartree-Fock equations in (64), we must remark that the necessity of using the relatively unreliable iterative approach to tackle them stems again from their complicated mathematical form. Like in the Hartree case, we have managed to largely reduce the dimension of the space on which the basic equations are defined: from $3N$ in the electronic Schrödinger equation in (9a) to 3 in the Hartree-Fock ones. However, to have this, we have payed the price of dramatically increasing their complexity [55], since, while the electronic Schrödinger equation was one linear differential equation, the Hartree-Fock ones in (64) are N coupled non-linear integro-differential equations, thus precluding any analytical approach to their solution.

A typical iterative procedure¹ begins by proposing a *starting guess* for the set

¹ The process described in this paragraph must be taken only as an outline of the one that is performed in practice. It is impossible to deal in a computer with a general function as it is (a non-countable infinite set of numbers), and the problem must be discretized in some way. The truncation of the one-electron Hilbert space using a finite basis set, described in secs. 8 and 9, is the most common way of doing this.

of N spin-orbitals $\{\psi_i^0\}$. With them, the Fock operator $\hat{F}^{\text{GHF}}[\psi^0]$ in the left-hand side of eq. (64) is constructed and the set of N equations is solved as one simple eigenvalue problem. Then, the $\{\psi_i^1\}$ that correspond to the N lowest eigenvalues ε_i^1 are selected (see the discussion of the aufbau principle below) and a new Fock operator $\hat{F}^{\text{GHF}}[\psi^1]$ is constructed with them. The process is iterated until (hopefully) the n -th set of solutions $\{\psi_i^n\}$ differs from the $(n-1)$ -th one $\{\psi_i^{n-1}\}$ less than a reasonably small amount (defining the distance among solutions in some suitable way typically combined with a convergence criterium related to the associated energy change). When this occurs, the procedure is said to have converged and the solution orbitals are called *self-consistent*; also, a calculation of this kind is commonly termed *self-consistent field* (SCF).

Again, like in the Hartree case, many issues exist that raise doubts about the possible success of such an approach. The most important ones are related to the fact that a proper definition of the *Hartree-Fock problem* should be: *find the global minimum of the energy functional $\langle \Psi | \hat{H} | \Psi \rangle$ under the constraint that the wavefunction Ψ be a Slater determinant of one-electron spin-orbitals*, and not: *solve the Hartree-Fock equations* (64). The solutions of the latter are all the stationary points of the constrained energy functional, while we are interested only in the particular one that is the global minimum. Even ruling out the possibility that a found solution may be a maximum or a saddle point (which can be done [50,55]), one can never be sure that it is the global minimum and not a local one.

There exists, however, one way, related to the Hartree-Fock version of the theorem by Simon and Lieb [57, 58] mentioned in the previous section, of hopefully biasing a particular found solution of eqs. (64) to be the global minimum that we are looking for. They showed, first, that for neutral or positively charged molecules ($Z \geq N$), the Hartree-Fock global minimization problem has a solution (its uniqueness is not established yet [55]) and, second, that the minimizing orbitals $\{\psi_i\}$ correspond to the N lowest eigenvalues of the Fock operator $\hat{F}^{\text{GHF}}[\psi]$ that is self-consistently constructed with them. Therefore, although *the reverse is not true in general* [55] (i.e., from the fact that a particular set of orbitals are the eigenstates corresponding to the lowest eigenvalues of the associated Fock operator, does not necessarily follow that they are the optimal ones), the information contained in Simon and Lieb's result is typically invoked to build each successive state in the iterative procedure described above by keeping only the N orbitals that correspond to the N lowest eigenvalues ε_i . Indeed, by doing that, one is effectively constraining the solutions to have a property that the true solution does have, so that, in the worst case, the space in which one is searching is of the same size as the original one, and, in the best case (even playing with the possibility that the reverse of Simon and Lieb's theorem be true, though not proved), the space of solutions is reduced to the correct global minimum alone. This wishful-thinking way of proceeding is termed the *aufbau principle* [55], and, together with a clever choice of the starting-guess set of orbitals [61] (typically extracted from a slightly less accurate theory, so that one may expect that it could be 'in the basin of attraction' of the true Hartree-Fock minimum), constitute one of the many heuristic strategies that make possible that the aforementioned drawbacks (and also those related to the convergence of iterative procedures) be circumvented in real cases, so that, in practice, most of SCF calculations performed in the field of quantum chemistry do converge to the true solution of eqs. (64) in spite of the theoretical notes of caution.

Moreover, the iterative procedure is normally performed using not the spin-orbitals but the spatial ones. In this sense, the restricted versions of the Hartree-Fock problem, discussed below, are closer to the actual implementation of the theory in computer applications.

Now, to close this GHF part, let us discuss some points regarding the imposition of constraints as a justification for subsequently introducing three commonly used forms of the Hartree-Fock theory that involve additional restrictions on the variational ansatz (apart from those in eqs. (32) and (34)).

In principle, the target systems in which we are interested in our group and to which the theory developed in this work is meant to be applied are rather complex (short peptides, small ligands, etc.). They have many degrees of freedom and the different interactions that drive their behaviour typically compete with one another, thus producing complicated, ‘frustrated’ energy landscapes (see refs. [62–67], but note, however, that we do not need to think about macromolecules; a small molecule like CO₂ already has 22 electrons). This state of affairs renders the a priori assessment of the accuracy of any approximation to the exact equations an impossible task. As researchers calculate more and more properties of molecular species using quantum chemistry and the results are compared to higher-level theories or to experimental data, much empirical knowledge about ‘how good is Theory A for calculating Property X’ is being gathered. However, if the characterization of a completely new molecule that is not closely related to any one that has been previously studied is tackled with, say, the Hartree-Fock approximation, it would be very unwise not to ‘ask for a second opinion’.

All of this also applies, word by word, to the choice of the constraints on the wavefunction in variational approaches like the one discussed in this section: For example, it is impossible to know a priori what will be the loss of accuracy due to the requirement that the N -particle wavefunction Ψ be a Slater determinant as in eq. (32). However, in the context of the Hartree-Fock approximation, there exists a way of proceeding, again, partly based on wishful thinking and partly confirmed by actual calculations in particular cases, that is almost unanimously used to choose additional constraints which are expected to yield more *efficient* theories. It consists of imposing constraints to the variational wavefunction that are *properties that the exact solution to the problem does have*. In such a way that the obvious loss of accuracy due to the reduction of the search space is expected to be minimized, while the decrease in computational cost could be considerable.

This way of thinking is clearly illustrated by the question of whether or not one should allow that the one-electron spin-orbitals ψ_i (and therefore the total wavefunction Ψ) be complex valued. Indeed, due to the fact that the electronic Hamiltonian in eq. (21) is self-adjoint, the real and imaginary parts of any complex eigenfunction solution of the time independent Schrödinger equation in (9a) are also solutions of it [55]. Therefore, the ground-state, which is the exact solution of the problem that we are trying to solve, may be chosen to be real valued. Nevertheless, the exact minimum will not be achieved, in general, in the smaller space defined by the Hartree-Fock constraints in eqs. (32) and (34), so that there is no a priori reason to believe that allowing the Hartree-Fock wavefunction to take complex values would not improve the results by finding a lower minimum. In fact, in some cases, this happens [61]. Nevertheless, if one constrains the search to real orbitals, the computational cost is reduced by a factor two, and, after all, the whole formalism discussed in this work profits from the imposition of constraints (starting by the consideration of only one Slater determinant), all of which save some computational effort at the expense of a reduction in the accuracy. The search for the most efficient of these approximations constitutes the main part of the quantum chemistry field.

Apart from these ‘complex vs. real’ considerations, there exist three further restrictions that are commonly found in the literature and that affect the spin part of the one-electron orbitals ψ_i . The N -electron wavefunction Ψ of the GHF approximation (which is the one discussed up to now) is not an eigenstate of the total-spin

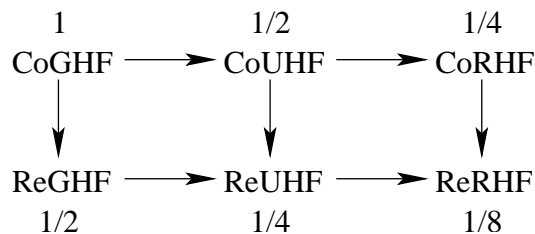


Figure 1. Schematic relation map among six types of Hartree-Fock methods discussed in the text: General Hartree-Fock (GHF), Unrestricted Hartree-Fock (UHF) and Restricted Hartree-Fock (RHF), in both their *complex* (Co) and *real* (Re) versions. The arrows indicate imposition of constraints; horizontally, in the spin part of the orbitals, and, vertically, from complex- to real-valued wavefunctions. Next to each method, the size of the search space relative to that in CoGHF is shown.

operator, \hat{S}^2 , nor of the z -component of it, \hat{S}_z [61]. However, since both of them commute with the electronic Hamiltonian in eq. (21), the true ground-state of the exact problem can be chosen to be an eigenstate of both operators simultaneously. Therefore three additional types of constraints on the spin part of the GHF wavefunction in (32) are typically made that force the variational ansatz to satisfy these ground-state properties and that should be seen in the light of the above discussion, i.e., as reducing the search space, thus yielding an intrinsically less accurate theory, but also as being good candidates to hope that the computational savings will pay for this.

The first approximation to GHF (in a logical sense) is called Unrestricted Hartree-Fock (UHF) and it consists of asking the orbitals ψ_i to be a product of a part $\phi_i(\mathbf{r})$ depending on the positions \mathbf{r} times a spin eigenstate of the one-electron \hat{s}_z operator, i.e., either $\alpha(\sigma)$ or $\beta(\sigma)$ (see eq. (36)). This is denoted by $\psi_i(x) := \phi_i(\mathbf{r})\gamma_i(\sigma)$, where γ_i is either the α or the β function. Now, if we call N_α and N_β the number of spin-orbitals of each type, we have that, differently from the GHF one, the UHF N -particle wavefunction Ψ is an eigenstate of the \hat{S}_z operator with eigenvalue $(1/2)(N_\alpha - N_\beta)$ (in atomic units, see sec. 2). However, it is not an eigenstate of \hat{S}^2 and, when this deviation results into a poor description of the observables in which we are interested, we talk about *spin contamination* [68]. Although the UHF wavefunction can be projected into pure \hat{S}^2 -states, the result is multideterminantal [61] and will not be considered here since it spoils many of the properties that render Hartree-Fock methods a low-cost choice.

Regarding the computational cost of the UHF approximation, it is certainly lower than that of GHF, since the search space is half as large: In the latter case, we had to consider $2N$ (complex or real) functions of \mathbb{R}^3 (the $\varphi_i^\alpha(\mathbf{r})$ and the $\varphi_i^\beta(\mathbf{r})$, see eq. (35)), while in UHF we only have to deal with N of them: the $\phi_i(\mathbf{r})$.

Now, if we introduce into the expression for the GHF constrained functional in (48) the following relations that hold for the UHF spin- and spatial orbitals¹

$$\langle \psi_i | \hat{h} | \psi_i \rangle = \langle \phi_i | \hat{h} | \phi_i \rangle, \quad (75a)$$

$$\langle \psi_i \psi_j | \frac{1}{r} | \psi_i \psi_j \rangle = \langle \phi_i \phi_j | \frac{1}{r} | \phi_i \phi_j \rangle, \quad (75b)$$

$$\langle \psi_i \psi_j | \frac{1}{r} | \psi_j \psi_i \rangle = \delta_{\gamma_i \gamma_j} \langle \phi_i \phi_j | \frac{1}{r} | \phi_j \phi_i \rangle, \quad (75c)$$

¹ Of course, the average values at both sides of the expressions are taken over different variables: over x and x' on the left-hand side, and over \mathbf{r} and \mathbf{r}' on the right-hand side. Also, let us remark that, although placing *functions* as arguments of the Kronecker's delta $\delta_{\gamma_i \gamma_j}$ is a bit unorthodox mathematically, it constitutes an intuitive (and common) notation.

we may perform a derivation analogous to the one performed for the GHF case, and get to a first version of the *UHF equations*

$$\hat{F}_i^{\text{UHF}}[\phi] \phi_i(\mathbf{r}) := \left(\hat{h} + \sum_j^N \hat{J}_j[\phi] - \sum_j^N \delta_{\gamma_i \gamma_j} \hat{K}_j[\phi] \right) \phi_i(\mathbf{r}) = \sum_j \lambda_{ij} \phi_j(\mathbf{r}), \quad (76a)$$

$$\hat{F}_i^{\text{UHF}}[\phi] \phi_i^*(\mathbf{r}) = \sum_j \lambda_{ji} \phi_j^*(\mathbf{r}), \quad (76b)$$

for all $i = 1, \dots, N$ and where the Coulomb and exchange operators dependent on the spatial orbitals ϕ_i are defined by their action on an arbitrary function $\varphi(\mathbf{r})$ as follows:

$$\hat{J}_j[\phi] \varphi(\mathbf{r}) := \left(\int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \varphi(\mathbf{r}), \quad (77a)$$

$$\hat{K}_j[\phi] \varphi(\mathbf{r}) := \left(\int \frac{\phi_j^*(\mathbf{r}') \varphi(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right) \phi_j(\mathbf{r}). \quad (77b)$$

Now, one must note that, differently from the GHF case, due to the fact that the exchange interaction only takes place between orbitals ‘of the same spin’, the *UHF Fock operator* $\hat{F}_i^{\text{UHF}}[\phi]$ depends on the index i . This precludes the solution of UHF as a *single* pseudoeigenvalue problem (c.f. eq. (64)) and makes necessary some further considerations in order to arrive to a more directly applicable form of the expressions:

First, although the equations for ϕ_i and ϕ_i^* in (76) can be combined in the same way as in the GHF case to yield the Hermiticity conditions $\lambda_{ij} = \lambda_{ji}^*$, there are fewer Lagrange multipliers in the UHF scheme than in the previous derivation. To see this, one must notice that the orthogonality constraints must be imposed *on the spin-orbitals, not on the spatial orbitals* (see eq. (34)). Therefore, since two UHF spin-orbitals $\psi_i = \phi_i \alpha$ and $\psi_j = \phi_j \beta$ are orthogonal no matter the value of $\langle \phi_i | \phi_j \rangle$ due to the different spin parts, the corresponding Lagrange multiplier λ_{ij} needs not to be included in the constrained functional from which the UHF equations come. This may be incorporated into the formalism by simply using that the matrix $\Lambda := (\lambda_{ij})$ in eqs. (76) presents the following block-diagonal form:

$$\Lambda^{\text{UHF}} := \begin{pmatrix} \Lambda^\alpha & 0 \\ 0 & \Lambda^\beta \end{pmatrix}, \quad (78)$$

where we have assumed (without loss of generality) that the UHF spin-orbitals are ordered in such a way that the α ones occur first, and 0 indicates a block of zeros of the appropriate size. (Of course, redundant constraints may be imposed on the orbitals by, for example in this case, including matrix terms in Λ that connect the α and β spaces. However, in order to know the exact freedom we have in the choice of the constraints, it is convenient to use the minimal number of Lagrange multipliers. If this approach were not followed, for example, the discussion below about the diagonalization of Λ^α and Λ^β would become much less direct.)

The next step consists in noticing that, despite the dependence of \hat{F}_i^{UHF} on the orbital index i in eqs. (76), there are actually only *two* different Fock operators:

one for the α orbitals and one for the β ones. Defining the sets of indices $A := \{i | 1 \leq i \leq N_\alpha\}$ and $B := \{i | N_\alpha + 1 \leq i \leq N_\alpha + N_\beta = N\}$, we can write these two α and β operators:

$$\hat{F}_\alpha^{\text{UHF}}[\phi] = \hat{h} + \sum_{j=1}^N \hat{J}_j[\phi] - \sum_{j \in A} \hat{K}_j[\phi], \quad (79a)$$

$$\hat{F}_\beta^{\text{UHF}}[\phi] = \hat{h} + \sum_{j=1}^N \hat{J}_j[\phi] - \sum_{j \in B} \hat{K}_j[\phi]. \quad (79b)$$

With them, and using the particular structure of the matrix Λ^{UHF} in (78), the original UHF equations in (76) are split into two disjoint sets of expressions that are only coupled through the Fock operators on the left hand sides, plus the Hermiticity condition:

$$\hat{F}_\alpha^{\text{UHF}}[\phi]\phi_i(\mathbf{r}) = \sum_{j \in A} \lambda_{ij}^\alpha \phi_j(\mathbf{r}), \quad \text{if } i \in A, \quad (80a)$$

$$\hat{F}_\beta^{\text{UHF}}[\phi]\phi_i(\mathbf{r}) = \sum_{j \in B} \lambda_{ij}^\beta \phi_j(\mathbf{r}), \quad \text{if } i \in B, \quad (80b)$$

$$\lambda_{ij} = \lambda_{ji}^*. \quad (80c)$$

The last step needed to arrive to the final form of the UHF equations (found by Pople and Nesbet [69] and named after them) is the diagonalization of both the Λ^α and Λ^β Hermitian matrices above. In order to achieve this, the orbitals ϕ_i must be transformed similarly to the GHF case. However, this is trickier than it was then, since not only the N -electron wavefunction Ψ and the Fock operators must remain invariant under the sought transformation, but also the UHF constraints must be kept.

As we saw before, any unitary transformation U in the set of spin-orbitals ψ_i like the one in eq. (58) is physically legitimate, since it changes the Slater determinant by only an unmeasurable phase and leave the Fock operators invariant. Then, if we write each spin-orbital as in (35):

$$\psi_i(x) = \varphi_i^\alpha(\mathbf{r}) \alpha(\sigma) + \varphi_i^\beta(\mathbf{r}) \beta(\sigma), \quad (81)$$

we can make use of (58), to obtain

$$\varphi_i^\alpha(\mathbf{r}) \alpha(\sigma) + \varphi_i^\beta(\mathbf{r}) \beta(\sigma) = \sum_j U_{ij} [\varphi_j'^\alpha(\mathbf{r}) \alpha(\sigma) + \varphi_j'^\beta(\mathbf{r}) \beta(\sigma)]. \quad (82)$$

By setting $\sigma = -1/2$ and $\sigma = 1/2$ in this expression, we see that any transformation U in the spin-orbitals ψ_i induces exactly the same transformation in their spatial components φ_i^α and φ_i^β ,

$$\varphi_i^\gamma(\mathbf{r}) = \sum_j U_{ij} \varphi_j'^\gamma(\mathbf{r}), \quad \gamma = \alpha, \beta. \quad (83)$$

Now, if we order the sets of spin and spatial orbitals: $\boldsymbol{\psi}^T := (\psi_1, \dots, \psi_N)$ and $(\boldsymbol{\varphi}^\gamma)^T := (\varphi_1^\gamma, \dots, \varphi_N^\gamma)$, with $\gamma = \alpha, \beta$, we can express the *UHF constraints* by saying that the $\boldsymbol{\varphi}^\gamma$ must have the form

$$(\boldsymbol{\varphi}^\alpha)^T = (\phi_1, \dots, \phi_{N_\alpha}, 0, \dots, 0), \quad (84a)$$

$$(\boldsymbol{\varphi}^\beta)^T = (0, \dots, 0, \phi_{N_\alpha+1}, \dots, \phi_{N_\alpha+N_\beta}). \quad (84b)$$

Then, since the fact that the α orbitals appear first constitutes no loss of generality, we must ask the transformed $\varphi_i'^\gamma$, with $\gamma = \alpha, \beta$, to have also the structure in (84) if we want to remain inside the UHF scheme. As a consequence, and due to the linear independence among the orbitals, not every unitary matrix U is allowed, but only those of the form

$$U^{\text{UHF}} := \begin{pmatrix} U^\alpha & 0 \\ 0 & U^\beta \end{pmatrix}, \quad (85)$$

using the same notation as in eq. (78).

This can be easily proved by focusing on a particular value for i and γ in eq. (83), say, $i \in A$ and $\gamma = \beta$. Due to the UHF constraints in (84), we know that the left-hand side of such an expression is zero and that only spatial orbitals with $j \in B$ appear in the sum on the right-hand side, yielding the relation $0 = \sum_{j \in B} U_{ij} \phi_j'(\mathbf{r})$. But the $\phi_j'(\mathbf{r})$, with $j \in B$, form an orthonormal, and therefore linearly independent set, so that the only possibility that such a relation can hold is that all coefficients U_{ij} be zero. By repeating this for all $i \in A$ and, then, for $\gamma = \alpha$, the result follows.

Fortunately, this limited freedom in the choice of U is still enough to independently diagonalize Λ^α and Λ^β in eq. (80) (which are both Hermitian) by suitably choosing the unitary submatrices U^α and U^β respectively.

This takes us to the final, diagonal form of the UHF equations, the *Pople-Nesbet equations* [69]:

$$\hat{F}_\alpha^{\text{UHF}}[\phi] \phi_i(\mathbf{r}) = \varepsilon_i^\alpha \phi_i(\mathbf{r}), \quad \text{if } i \in A, \quad (86a)$$

$$\hat{F}_\beta^{\text{UHF}}[\phi] \phi_i(\mathbf{r}) = \varepsilon_i^\beta \phi_i(\mathbf{r}), \quad \text{if } i \in B. \quad (86b)$$

Although these equations are coupled through the Coulomb term in the Fock operators on the left-hand side, at each step of the iterative SCF procedure, they can be solved as two independent eigenvalue problems. This has allowed to implement them in most quantum chemical packages and it has made UHF calculations now routine.

To close the UHF discussion, we shall now study the statistical properties of the probability densities appearing in this model. If we introduce the special form of the UHF orbitals in (84) into the general expression in (70), we can calculate the two-body probability density of finding *any* electron with coordinates x at the same time that *any other* electron has coordinates x' :

$$\rho^{\text{UHF}}(x, x') = \frac{1}{2} \left(\sum_{k,l} |\phi_k(\mathbf{r})|^2 |\phi_l(\mathbf{r}')|^2 \gamma_k(\sigma) \gamma_l(\sigma') - \sum_{k,l} \phi_k^*(\mathbf{r}) \phi_l^*(\mathbf{r}') \phi_l(\mathbf{r}) \phi_k(\mathbf{r}') \gamma_k(\sigma) \gamma_l(\sigma') \gamma_k(\sigma') \gamma_l(\sigma) \right). \quad (87)$$

If we compute this probability density for ‘electrons of the same spin’, i.e., for $\sigma = \sigma'$, we obtain¹

$$\begin{aligned} \rho^{\text{UHF}}(\mathbf{r}, \mathbf{r}'; \sigma = \sigma') &= \\ \frac{1}{2} \sum_{k,l \in I_\sigma} &\left(|\phi_k(\mathbf{r})|^2 |\phi_l(\mathbf{r}')|^2 - \phi_k^*(\mathbf{r}) \phi_l^*(\mathbf{r}') \phi_l(\mathbf{r}) \phi_k(\mathbf{r}') \right) = \\ \frac{1}{2} \left(\rho^{\text{UHF}}(\mathbf{r}, \sigma) \rho^{\text{UHF}}(\mathbf{r}', \sigma) - \sum_{k,l} \delta_{\gamma_k \sigma} \delta_{\gamma_l \sigma} \phi_k^*(\mathbf{r}) \phi_l^*(\mathbf{r}') \phi_l(\mathbf{r}) \phi_k(\mathbf{r}') \right), & \quad (88) \end{aligned}$$

where the following expression for the one-electron charge density has been used:

$$\rho^{\text{UHF}}(\mathbf{r}, \sigma) = \sum_k |\phi_k(\mathbf{r})|^2 \gamma_k(\sigma). \quad (89)$$

At this point, note that eq. (88) contains, like in the GHF case, the exchange correction to the first (independent electrons) term. Nevertheless, as we advanced, if we calculate the two-body ρ^{UHF} for ‘electrons of opposite spin’, i.e., for $\sigma \neq \sigma'$, we have that

$$\rho^{\text{UHF}}(\mathbf{r}, \mathbf{r}'; \sigma \neq \sigma') = \frac{1}{2} \rho^{\text{UHF}}(\mathbf{r}, \sigma) \rho^{\text{UHF}}(\mathbf{r}', \sigma'), \quad (90)$$

i.e., that *UHF electrons of opposite spin are statistically pairwise independent*.

There is another common approximation to GHF that is more restrictive than UHF and is accordingly called *Restricted Hartree-Fock* (RHF). Apart from asking the orbitals ψ_i to be a product of a spatial part times a spin eigenstate of the one-electron \hat{s}_z operator (like in the UHF case), in RHF, the number of ‘spin-up’ and ‘spin-down’ orbitals is the same, $N_\alpha = N_\beta$ (note that this means that RHF may only be used with molecules containing an even number of electrons), and each spatial wavefunction occurs twice: once multiplied by $\alpha(\sigma)$ and the other time by $\beta(\sigma)$. This is typically referred to as a *closed-shell* situation and we shall denote it by writing $\psi_i(x) := \phi_i(\mathbf{r}) \alpha(\sigma)$ if $i \leq N/2$, and $\psi_i(x) := \phi_{i-N/2}(\mathbf{r}) \beta(\sigma)$ if $i > N/2$; in such a way that there are $N/2$ different spatial orbitals denoted by $\phi_I(\mathbf{r})$, with, $I = 1, \dots, N/2$. Using the same notation as in (84), the *RHF constraints* on the spin-orbitals are

¹ Placing a function and a coordinate as arguments of the Kronecker’s delta $\delta_{\gamma_k \sigma}$ is even more unorthodox mathematically than placing two functions (in fact, $\delta_{\gamma_k \sigma}$ is exactly the same as $\gamma_k(\sigma)$), however, the intuitive character of the notation compensates again for this.

$$(\boldsymbol{\varphi}^\alpha)^T = (\phi_1, \dots, \phi_{N/2}, 0, \dots, 0), \quad (91a)$$

$$(\boldsymbol{\varphi}^\beta)^T = (0, \dots, 0, \phi_1, \dots, \phi_{N/2}). \quad (91b)$$

Due to these additional restrictions, we have that, differently from the GHF and UHF ones, the RHF N -particle wavefunction Ψ is an eigenstate of both the \hat{S}^2 and the \hat{S}_z operators, with zero eigenvalue in both cases [19, 61], just like the ground-state of the exact problem. I.e., *there is no spin-contamination in RHF*.

Regarding the computational cost of the RHF approximation, it is even lower than that of UHF, since the size of the search space has been reduced to one quarter that of GHF: In the latter case, we had to consider $2N$ (complex or real) functions of \mathbb{R}^3 (the $\varphi_i^\alpha(\mathbf{r})$ and the $\varphi_i^\beta(\mathbf{r})$, see eq. (35)), while in RHF we only have to deal with $N/2$ of them: the $\phi_I(\mathbf{r})$ (see fig. 1).

Now, using again the relations in (75), we can derive a first version of the *RHF equations*:

$$\hat{F}^{\text{RHF}}[\phi] \phi_I(\mathbf{r}) := \left[\hat{h} + \sum_J^{N/2} \left(2\hat{J}_J[\phi] - \hat{K}_J[\phi] \right) \right] \phi_I(\mathbf{r}) = \sum_J^{N/2} \lambda_{IJ} \phi_J(\mathbf{r}), \quad (92)$$

for all $I = 1, \dots, N/2$, where we have used that the *RHF Fock operator* $\hat{F}^{\text{RHF}}[\phi]$, differently from the UHF case, does not depend on the index I of the orbital on which it operates, and following the same steps as before, the minimal Lagrange multipliers matrix needed to enforce the orthogonality constraints among the *spinorbitals* in the RHF case is

$$\Lambda^{\text{RHF}} := \begin{pmatrix} \Lambda^{(N/2)}/2 & 0 \\ 0 & \Lambda^{(N/2)}/2 \end{pmatrix}, \quad (93)$$

where, this time, $\Lambda^{(N/2)} := (\lambda_{IJ})$ is an arbitrary $N/2 \times N/2$ Hermitian matrix, and the $1/2$ has been included in order to get to the classical RHF equations in (92) without irrelevant numerical factors.

Using the same arguments as for UHF, it is clear that, in order to ‘remain inside RHF’ upon an unitary transformation of the spin-orbitals ψ_i , not every unitary matrix U is allowed, but only those of the form

$$U^{\text{RHF}} := \begin{pmatrix} U^{(N/2)} & 0 \\ 0 & U^{(N/2)} \end{pmatrix}. \quad (94)$$

Finally, by suitable choosing the $N/2 \times N/2$ unitary block $U^{(N/2)}$, the Hermitian matrix $\Lambda^{(N/2)}$ can be diagonalized and the final, diagonal form of the *RHF equations* can be written:

$$\hat{F}^{\text{RHF}}[\phi] \phi_I(\mathbf{r}) := \left[\hat{h} + \sum_J^{N/2} \left(2\hat{J}_J[\phi] - \hat{K}_J[\phi] \right) \right] \phi_I(\mathbf{r}) = \varepsilon_I \phi_I(\mathbf{r}), \quad (95)$$

with $I = 1, \dots, N/2$.

As we can see, this version of the Hartree-Fock theory can be numerically solved as a *single* pseudoeigenvalue problem. This, together with the aforementioned small size of the RHF space, has made the RHF approximation (in its real-valued version) the first one cast into a computationally manageable form [79, 80] and the most used one in recent literature [14, 70–78] (for molecules with an even number of electrons).

Now, in order to investigate the statistical features of RHF, if we introduce the special form of the orbitals in (91) into the general expression in (70), we can calculate the RHF two-body probability density of finding *any* electron with coordinates x at the same time that *any other* electron has coordinates x' :

$$\begin{aligned} \rho^{\text{RHF}}(x, x') &= \frac{1}{2} \left(\sum_{K,L}^{N/2} |\phi_K(\mathbf{r})|^2 |\phi_L(\mathbf{r}')|^2 - \delta_{\sigma\sigma'} \sum_{K,L}^{N/2} \phi_K^*(\mathbf{r}) \phi_L^*(\mathbf{r}') \phi_L(\mathbf{r}) \phi_K(\mathbf{r}') \right) = \\ &= \frac{1}{2} \left(\rho(\mathbf{r}, \sigma) \rho(\mathbf{r}', \sigma') - \delta_{\sigma\sigma'} \sum_{K,L}^{N/2} \phi_K^*(\mathbf{r}) \phi_L^*(\mathbf{r}') \phi_L(\mathbf{r}) \phi_K(\mathbf{r}') \right). \end{aligned} \quad (96)$$

where the following expression for the RHF one-electron charge density has been used:

$$\rho^{\text{RHF}}(\mathbf{r}, \sigma) = \sum_K^{N/2} |\phi_K(\mathbf{r})|^2. \quad (97)$$

We notice that the situation is the same as in the UHF case: For RHF electrons with equal spin, there exists an exchange term in $\rho^{\text{RHF}}(x, x')$ that corrects the ‘independent’ part, whereas *RHF electrons of opposite spin are statistically pairwise independent*.

Finally, if we follow the same steps as for GHF, in page 26, we can relate the *RHF energy* to the eigenvalues ε_I and the two-electron spatial integrals:

$$E = 2 \sum_I^{N/2} \varepsilon_I - \sum_{I,J}^{N/2} \left(2 \langle \phi_I \phi_J | \frac{1}{r} | \phi_I \phi_J \rangle - \langle \phi_I \phi_J | \frac{1}{r} | \phi_J \phi_I \rangle \right). \quad (98)$$

To close this section, we shall discuss a fourth flavour of Hartree-Fock which is called Restricted Open-shell Hartree-Fock (ROHF). Compared to the rest of variants, ROHF is the most difficult to derive theoretically and it shall be described here only in an introductory manner. The ROHF wavefunction is somewhat between the RHF and the UHF ones, and both of them can be obtained as particular cases of the ROHF scheme (the RHF one provided that the molecule has an even number of electrons). In the (monodeterminantal) ROHF case, the one-electron spin orbitals ψ_i are constrained to be of two different types: $2N_D$ of them are *doubly occupied*, like in the RHF case, in such a way that they are formed by only N_D spatial orbitals ϕ_a , each one of them appearing once multiplied by $\alpha(\sigma)$ and once by $\beta(\sigma)$. The associated $2N_D$ spin-orbitals are said to belong to the *closed shell* part of the wavefunction. The remaining $N_S := N - 2N_D$ ones are *singly occupied*, like in the UHF case, and are said to belong to the *open shell*. Among them, N_α

are multiplied by an $\alpha(\sigma)$ spin part, and N_β by $\beta(\sigma)$.

If we number the whole set of ROHF spatial orbitals ϕ_a , with $a = 1, \dots, N_D + N_\alpha + N_\beta$, in such a way that the doubly occupied ones occur first, with $a \in D := \{a | 1 \leq a \leq N_D\}$, then the alpha ones, with $a \in A := \{a | N_D + 1 \leq a \leq N_D + N_\alpha\}$, and finally the beta ones, with $a \in B := \{a | N_D + N_\alpha + 1 \leq a \leq N_D + N_\alpha + N_\beta\}$, we can express the *ROHF constraints* on the spin-orbitals using the same notation as in (84) and (91):

$$(\varphi^\alpha)^T = (\underbrace{\phi_1, \dots, \phi_{N_D}}_{N_D}, \underbrace{\phi_{N_D+1}, \dots, \phi_{N_D+N_\alpha}}_{N_\alpha}, \underbrace{0, \dots, 0}_{N_D+N_\beta}), \quad (99a)$$

$$(\varphi^\beta)^T = (\underbrace{0, \dots, 0}_{N_D+N_\alpha}, \underbrace{\phi_1, \dots, \phi_{N_D}}_{N_D}, \underbrace{\phi_{N_D+N_\alpha+1}, \dots, \phi_{N_D+N_\alpha+N_\beta}}_{N_\beta}), \quad (99b)$$

where the particular ordering has been chosen in order to facilitate the forthcoming calculations.

The general monodeterminantal ROHF wavefunction considered here and constructed using the above constraints has the same spin properties as the UHF one, i.e., it is an eigenstate of the the \hat{S}_z operator with eigenvalue $(1/2)(N_\alpha - N_\beta)$, but it is not an eigenstate of \hat{S}^2 . However, in the particular (and common) case in which all the N_S open shell orbitals are constrained to present parallel spin parts (either all α or all β), the ROHF wavefunction becomes an eigenstate of the \hat{S}^2 operator too, with eigenvalue $\frac{N_S}{2}(\frac{N_S}{2} + 1)$, thus avoiding the problem of spin contamination [81]. In order to construct wavefunctions with the same S^2 but lower S_z , several ROHF Slater determinants must be linearly combined. The subtleties arising from such a procedure are beyond the scope of this review; the interested reader may want to check references [81–84], which discuss this topic.

Regarding the size of the variational space in ROHF, it is somewhere between UHF and RHF, depending on the $2N_D/(N_\alpha + N_\beta)$ ratio.

Now, if the ROHF constraints in (99) are imposed on the GHF energy in eq. (47), the ROHF analogue in terms of the spatial orbitals ϕ_a can be calculated:

$$E^{\text{ROHF}}[\{\phi_a\}] = \sum_a f_a \langle \phi_a | \hat{h} | \phi_a \rangle + \frac{1}{2} \sum_{a,b} \left(f_a f_b \langle \phi_a \phi_b | \frac{1}{r} | \phi_a \phi_b \rangle - g_{ab} \langle \phi_a \phi_b | \frac{1}{r} | \phi_b \phi_a \rangle \right), \quad (100)$$

where the f_a are a sort of ‘occupation numbers’ that take the value $f_a = 2$ when $a \in D$ (i.e., when it corresponds to a closed shell orbital) and $f_a = 1$ otherwise. The matrix $g := (g_{ab})$ is defined as

$$g := \left(\begin{array}{c|cc} 2^D & 1 & \\ \hline & 1^\alpha & 0 \\ 1 & 0 & 1^\beta \end{array} \right), \quad (101)$$

being 2^D a $N_D \times N_D$ box of 2’s, 1^α and 1^β , $N_\alpha \times N_\alpha$ and $N_\beta \times N_\beta$ boxes of 1’s respectively. The off-diagonal blocks contain in all elements the number indicated

and are of the appropriate size.

Then, as we did for GHF, in order to derive the ROHF equations, we construct the functional for the conditioned stationary values problem, adding to the energy in (100) the Lagrange multipliers terms needed to enforce the orthonormality constraints:

$$\tilde{\mathcal{F}}[\{\phi_a\}] = E^{\text{ROHF}}[\{\phi_a\}] - \sum_{a,b} \lambda_{ab} \left(\langle \phi_a | \phi_b \rangle - \delta_{ab} \right). \quad (102)$$

Now, like in all restricted HF cases, special attention must be paid to the structure of the matrix $\Lambda^{\text{ROHF}} := (\lambda_{ab})$, since the requirement is that all *spin-orbitals* be orthogonal, not the spatial orbitals. In the ROHF case, this leads to explicitly impose the orthonormality conditions (of course) inside the three sets of ϕ_a : the doubly occupied, the alpha and the beta ones; but also between the doubly occupied and the alpha ones, and between the doubly occupied and the beta ones. The orthormality between the alpha and beta sets, however, needs not to be enforced, since the associated spin-orbitals are already orthogonal due to the different spin parts.

These considerations lead to the following form for the minimal Lagrange multipliers matrix:

$$\Lambda^{\text{ROHF}} := \begin{pmatrix} \Lambda^D & \Lambda^{D\alpha} & \Lambda^{D\beta} \\ (\Lambda^{D\alpha})^+ & \Lambda^\alpha & 0 \\ (\Lambda^{D\beta})^+ & 0 & \Lambda^\beta \end{pmatrix}, \quad (103)$$

where the notation used for the different blocks connecting the doubly occupied, alpha and beta shells is self-explanatory, and the fact that Λ^{ROHF} is Hermitian after reaching stationarity has been advanced.

Next, we impose the condition of zero functional derivative on the functional in (102), and obtain a first version of the *ROHF equations*:

$$\hat{F}_a^{\text{ROHF}}[\phi] \phi_a(\mathbf{r}) = \sum_b \lambda_{ab} \phi_b(\mathbf{r}), \quad a = 1, \dots, N_D + N_\alpha + N_\beta, \quad (104a)$$

$$\hat{F}_a^{\text{ROHF}}[\phi] \phi_a^*(\mathbf{r}) = \sum_b \lambda_{ba} \phi_b^*(\mathbf{r}), \quad a = 1, \dots, N_D + N_\alpha + N_\beta, \quad (104b)$$

where the *ROHF operator* $\hat{F}_a^{\text{ROHF}}[\phi]$ is defined as

$$\hat{F}_a^{\text{ROHF}}[\phi] := f_a \hat{h} + \sum_b \left(f_a f_b \hat{J}_b[\phi] - g_{ab} \hat{K}_b[\phi] \right), \quad (105)$$

and the Hermiticity property of the Lagrange multipliers matrix Λ^{ROHF} follows from conjugation and subtraction in eqs. (104).

Again, although the Fock operator depends on the index a of the orbital upon which it operates, this dependence presents a very particular structure, yielding *only three* different types of operators:

$$\hat{F}_D^{\text{ROHF}}[\phi] := 2\hat{h} + \sum_b \left(2f_b \hat{J}_b[\phi] - f_b \hat{K}_b[\phi] \right), \quad a \in D, \quad (106a)$$

$$\hat{F}_\alpha^{\text{ROHF}}[\phi] := \hat{h} + \sum_{b \in (D \cup A)} \left(\hat{J}_b[\phi] - \hat{K}_b[\phi] \right), \quad a \in A, \quad (106b)$$

$$\hat{F}_\beta^{\text{ROHF}}[\phi] := \hat{h} + \sum_{b \in (D \cup B)} \left(\hat{J}_b[\phi] - \hat{K}_b[\phi] \right), \quad a \in B. \quad (106c)$$

Note that, in the particular case that we had no open-shell orbitals, the operator for the doubly occupied ones does not reduce (as it should) to the RHF one in (92). This is because we hid a factor 2 in the definition of the Lagrange multipliers in the RHF derivation.

At this point, it would be desirable to continue with the same program that we followed in the UHF and RHF cases and diagonalize the matrix Λ^{ROHF} in order to arrive to a system of three pseudoeigenvalue equations, coupled only via the Fock operators. Nevertheless, this is not possible in ROHF, as we shall show in the following lines, and it is the root of ROHF being the most tricky flavour in the Hartree-Fock family. The obstruction to achieve this diagonalization comes from the fact, already used in the UHF case, that not every unitary transformation of the *spin-orbitals* is allowed if we want to remain inside the ROHF scheme, i.e., if we want that the transformed spin-orbitals satisfy the same ROHF constraints in (99) that the untransformed ones did.

To begin with, if we recall that the doubly occupied spatial orbitals must be orthogonal to both the alpha and the beta ones, we can use the same reasoning as in page 32 to show that the alpha-beta connecting parts of the allowed unitary matrix in this case must be zero. Hence, we can write

$$U^{\text{ROHF}} := \left(\begin{array}{cc|cc} U^D & U^{D\alpha} & & \\ U^{D\alpha} & U^\alpha & & \\ \hline & & U^{D\beta} & U^\beta \\ 0 & & U^{D\beta} & U^\beta \end{array} \right), \quad (107)$$

where the size of each block may be easily found from (99) and the notation is again self-explanatory.

Now, in order to obtain further restrictions to the form of U^{ROHF} , we write the transformation of the *two* spin-orbitals in the closed shell that correspond to *the same* spatial orbital ϕ_a , with $a \in A$. To this end, we use (107), (99), and the appropriate lines in the first and third lines of blocks in (107):

$$\phi_a(\mathbf{r}) = \sum_{b \in D} U_{ab}^D \phi'_b(\mathbf{r}) + \sum_{b \in A} U_{ab}^{D\alpha} \phi'_b(\mathbf{r}) = \sum_{b \in D} U_{ab}^{D\beta} \phi'_b(\mathbf{r}) + \sum_{b \in B} U_{ab}^{D\beta} \phi'_b(\mathbf{r}). \quad (108)$$

Using the orthogonality relations inside and among the three sets of spatial orbitals, we multiply the above expression by any $\phi_c(\mathbf{r})$, with $c \in D$ and integrate on \mathbf{r} . From this, the equality of U^D and $U^{D\beta}$ follows, and the corresponding sums in (108) subtract to zero, yielding

$$\sum_{b \in A} U_{ab}^{D\alpha} \phi'_b(\mathbf{r}) = \sum_{b \in B} U_{ab}^{D\beta} \phi'_b(\mathbf{r}), \quad (109)$$

which equates one vector in the linear space spanned by the alpha orbitals to another in the linear space spanned by the beta ones.

However, only the zero vector can belong to both spaces if we want the ROHF assumptions regarding the spin-orbitals to hold. To see this, we can use a *reductio ad absurdum* type of argument: Assume that both sides of (109) are different from zero. Then, we may perform a unitary transformation changing only the alpha and beta spaces, and with no elements connecting the two. This is allowed, since it does not change neither the N -electron wavefunction, nor the Fock operators, nor the ROHF constraints. Now, if we select the partial unitary transformations in the alpha and beta sets so that they turn the vectors at both sides of (109) into single elements in the bases of their respective spaces, we have that a single alpha orbital equals a beta one. Although this can happen in particular cases, we cannot ask it or we would be changing the fundamental assumptions made in (99). Therefore, both sides of (109) must be zero, and, since the alpha and beta sets are linearly independent, all coefficients must be zero too.

The proof that $U^{D\alpha}$ and $U^{D\beta}$ are also zero is performed using similar arguments, and the final form of U^{ROHF} satisfying all the restrictions reads

$$U^{\text{ROHF}} = \left(\begin{array}{cc|cc} U^D & 0 & & 0 \\ 0 & U^\alpha & & \\ \hline & & U^D & 0 \\ 0 & & 0 & U^\beta \end{array} \right). \quad (110)$$

Finally, if we write the associated matrix using the a indices, i.e., operating on the set of spatial orbitals with the doubly occupied ones *unrepeated*:

$$\tilde{U}^{\text{ROHF}} := \begin{pmatrix} U^D & & 0 \\ & U^\alpha & \\ 0 & & U^\beta \end{pmatrix}, \quad (111)$$

then, the transformed Λ^{ROHF} is related to the original one in (103) through simple matrix multiplication: $\Lambda' = \tilde{U}^+ \Lambda \tilde{U}$ (dropping the ROHF superindices). It is clear that such a restricted \tilde{U}^{ROHF} does not operate on the off-diagonal blocks of Λ^{ROHF} in (103) and, therefore, *the sought diagonalization is not possible*.

Using the fact that, however, the Λ^D , Λ^α and Λ^β do allow to be diagonalized, we can write the final, simplest possible form of the *ROHF equations* (forgetting their complex conjugate counterparts):

$$\hat{F}_D^{\text{ROHF}}[\phi]\phi_a(\mathbf{r}) = \varepsilon_a^D \phi_a(\mathbf{r}) + \sum_{b \in A} \lambda_{ab}^{D\alpha} \phi_b(\mathbf{r}) + \sum_{b \in B} \lambda_{ab}^{D\beta} \phi_b(\mathbf{r}), \quad a \in D, \quad (112a)$$

$$\hat{F}_\alpha^{\text{ROHF}}[\phi]\phi_a(\mathbf{r}) = \varepsilon_a^\alpha \phi_a(\mathbf{r}) + \sum_{b \in D} (\lambda_{ba}^{D\alpha})^* \phi_b(\mathbf{r}), \quad a \in A, \quad (112b)$$

$$\hat{F}_\beta^{\text{ROHF}}[\phi]\phi_a(\mathbf{r}) = \varepsilon_a^\beta \phi_a(\mathbf{r}) + \sum_{b \in D} (\lambda_{ba}^{D\beta})^* \phi_b(\mathbf{r}), \quad a \in B, \quad (112c)$$

where the superindices in the matrix elements are only written for visual convenience when comparing with (103).

In this final form, it is evident that the off-diagonal elements of the Lagrange multipliers matrix, i.e., those related to the orthogonality constraints between the closed and open shells, introduce a coupling in the right-hand side of the ROHF equations that completely spoils the possibility of casting them into pseudoeigenvalue ones. In the previous UHF and RHF versions, *all* orthogonality constraints were handled by simply choosing a special basis in the space spanned by the spatial orbitals in which the Lagrange multipliers matrix was diagonal. However, in the ROHF scheme, there is no such basis and we must deal with the problem in a different way, resulting into higher computational and theoretical difficulty.

Since the pioneering work by Roothaan [85], the solution of the ROHF problem has been attempted by distinct means, ranging from directly tackling the ROHF equations in (112) explicitly forcing the orthogonality constraints [86, 87], to the construction of a so-called *unified coupling operator* [85, 89, 90], which allows to turn the ROHF scheme into a single pseudoeigenvalue problem at the price of introducing certain ambiguities in the one-electron orbital energies [82, 88]. The details and subtleties involved in these issues are beyond the scope of a review of the fundamental topics such as this one. The interested reader may want to check the more specialized accounts in [81–84].

8 The Roothaan-Hall equations

The Hartree-Fock equations in the RHF form in expression (95) are a set of $N/2$ coupled integro-differential equations. As such, they can be tackled by finite-differences methods and solved on a discrete grid; this is known as *numerical Hartree-Fock* [91], and, given the present power of computers, it is only applicable to very small molecules.

In order to deal with larger systems, such as biological macromolecules, independently proposed by Roothaan [79] and Hall [80] in 1951, a different kind of discretization must be performed, not in \mathbb{R}^3 but in the Hilbert space \mathcal{H} of the one-electron orbitals. Hence, although the actual dimension of \mathcal{H} is infinite, we shall approximate any function in it by a finite linear combination of M different functions χ_a ¹. In particular, the one-electron orbitals that make up the RHF wavefunction, shall be approximated by

¹ In all this section and the next one, the indices belonging to the first letters of the alphabet, a, b, c, d , etc., run from 1 to M (the number of functions in the finite basis set); whereas those named with capital letters from I towards the end of the alphabet, I, J, K, L , etc., run from 1 to $N/2$ (the number of spatial wavefunctions ϕ_I , also termed the number of *occupied orbitals*).

$$\phi_I(\mathbf{r}) \simeq \sum_a^M c_{aI} \chi_a(\mathbf{r}) , \quad I = 1, \dots, N/2 , \quad M \geq \frac{N}{2} . \quad (113)$$

In both cases, numerical Hartree-Fock and discretization of the function space, the correct result can be only be reached asymptotically; when the grid is very fine, for the former, and when $M \rightarrow \infty$, for the latter. This exact result, which, in the case of small systems, can be calculated up to several significant digits, is known as the *Hartree-Fock limit* [92].

In practical cases, however, M is finite (often, only about an order of magnitude larger than $N/2$) and the set $\{\chi_a\}_{a=1}^M$ in the expression above is called the *basis set*. We shall devote the next section to discuss its special characteristics, but, for now, it suffices to say that, in typical applications, the functions χ_a are *atom-centred*, i.e., each one of them has non-negligible value only in the vicinity of a particular nucleus. Therefore, like all the electronic wavefunctions we have dealt with in the last sections, they parametrically depend on the positions $\underline{\mathbf{R}}$ of the nuclei (see sec. 3). This is why, sometimes, the functions χ_a are called *atomic orbitals*¹ (AO) (since they are localized at individual atoms), the ϕ_I are referred to as *molecular orbitals* (MO) (since they typically have non-negligible value in the whole space occupied by the molecule), and the approximation in eq. (113) is called *linear combination of atomic orbitals* (LCAO). In addition, since we voluntarily circumscribe to real-RHF, we assume that both the coefficients c_{aI} and the functions χ_a in the above expression are real.

Now, if we introduce the linear combination in eq. (113) into the Hartree-Fock equations in (95), multiply the result from the left by χ_b (for a general value of b) and integrate on \mathbf{r} , we obtain

$$\sum_a F_{ba} c_{aI} = \varepsilon_I \sum_a S_{ba} c_{aI} , \quad I = 1, \dots, N/2 , \quad b = 1, \dots, M , \quad (114)$$

where we denote by F_{ba} the (b, a) -element of the *Fock matrix*², and by S_{ba} the one of the *overlap matrix*, defined as

$$F_{ba} := \langle \chi_b | \hat{F}[\phi] | \chi_a \rangle \quad \text{and} \quad S_{ba} := \langle \chi_b | \chi_a \rangle , \quad (115)$$

respectively.

Note that we do not ask the χ_a in the basis set to be mutually orthogonal, so that the overlap matrix is not diagonal in general.

Next, if we define the $M \times M$ matrices $F[c] := (F_{ab})$ and $S := (S_{ab})$, together with the (column) M -vector $c_I := (c_{aI})$, we can write eq. (114) in matricial form:

$$F[c]c_I = \varepsilon_I S c_I . \quad (116)$$

¹ Some authors [18] suggest that, being strict, the term *atomic orbitals* should be reserved for the one-electron wavefunctions ϕ_I that are the solution of the Hartree-Fock problem (or even to the exact Schrödinger equation of the isolated atom), and that the elements χ_a in the basis set should be termed simply *localized functions*. However, it is very common in the literature not to follow this recommendation and choose the designation that appear in the text [18, 79, 94]. We shall do the same for simplicity.

² Note that the RHF superindex has been dropped from F .

Hence, using the LCAO approximation, we have traded the $N/2$ coupled integro-differential Hartree-Fock equations in (95) for this system of $N/2$ algebraic equations for the $N/2$ orbital energies ε_I and the $M \cdot N/2$ coefficients c_{aI} , which are called *Roothaan-Hall equations* [79, 80] and which are manageable in a computer.

Now, if we forget for a moment that the Fock matrix depends on the coefficients c_{aI} (as stressed by the notation $F[c]$) and also that we are only looking for $N/2$ vectors c_I while the matrices F and S are $M \times M$, we may regard the above expression as a M -dimensional *generalized eigenvalue problem*. Many properties are shared between this kind of problem and a classical eigenvalue problem (i.e., one in which $S_{ab} = \delta_{ab}$) [79], being the most important one that, due to the Hermiticity of $F[c]$, one can find an orthonormal set of M vectors c_a corresponding to real eigenvalues ε_a (where, of course, some eigenvalue could be repeated).

In fact, it is using this formalism how most of actual Hartree-Fock computations are performed, although the reader must also note that other approaches, in which the orthonormality constraints are automatically satisfied due to the choice of variables also exist in the literature [93]. The general outline of the iterative procedure is essentially the same as the one discussed in sec. 7: Choose a *starting-guess* for the coefficients c_{aI} (let us denote it by c_{aI}^0), construct the corresponding Fock matrix $F[c^0]$ ¹ and solve the generalized eigenvalue problem in eq. (116). By virtue of the aufbau principle discussed in the previous section, from the M eigenvectors c_a , keep the $N/2$ ones c_I^1 that correspond to the $N/2$ lowest eigenvalues ε_I^1 , construct the new Fock matrix $F[c^1]$ and iterate (by convention, the eigenvalues ε_a^n , for all n , are ordered from the lowest to the largest as a runs from 1 to M). This procedure ends when the n -th solution is close enough (in a suitable defined way) to the $(n - 1)$ -th one. Also, note that, after convergence has been achieved, we end up with M orthogonal vectors c_a . Only the $N/2$ ones that correspond to the lowest eigenvalues represent real one-electron solutions and they are called *occupied orbitals*; the $M - N/2$ remaining ones do not enter in the N -electron wavefunction (although they are relevant for calculating corrections to the Hartree-Fock results) and they are called *virtual orbitals*.

Regarding the mathematical foundations of this procedure, let us stress, however, that, whereas in the finite-dimensional GHF and UHF cases it has been proved that the analogous of Lieb and Simon's theorem (see the previous section) is satisfied, i.e., that the global minimum of the original optimization problem corresponds to the lowest eigenvalues of the self-consistent Fock operator, in the RHF and ROHF cases, contrarily, no proof seems to exist [55]. Of course, in practical applications, the positive result is assumed to hold.

Finally, if we expand F_{ab} in eq. (115), using the shorthand $|a\rangle$ for $|\chi_a\rangle$, we have

$$F_{ab} = \langle a | \hat{h} | b \rangle + \underbrace{\sum_{c,d} \left(\sum_J c_{cJ} c_{dJ} \right)}_{D_{cd}[c]} \underbrace{\left(2 \langle ac | \frac{1}{r} | bd \rangle - \langle ac | \frac{1}{r} | db \rangle \right)}_{G_{ab}^{cd}}, \quad (117)$$

where we have introduced the *density matrix* $D_{cd}[c]$, and also the matrix G_{ab}^{cd} , made up by the *two-electron four-centre integrals* $\langle ac | 1/r | bd \rangle$ (also called *electron repulsion integrals* (ERIs)) defined by

¹ Note (in eq. (117), for example) that the Fock matrix only depends on the vectors c_a with $a \leq N/2$.

$$\langle ac | \frac{1}{r} | bd \rangle := \iint \frac{\chi_a(\mathbf{r}) \chi_c(\mathbf{r}') \chi_b(\mathbf{r}) \chi_d(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' . \quad (118)$$

It is also convenient to introduce the *Coulomb* ($J_{ab}[c]$) and *exchange* ($K_{ab}[c]$) *matrices*

$$J_{ab}[c] := \sum_{c,d} D_{cd}[c] \langle ac | \frac{1}{r} | bd \rangle , \quad (119a)$$

$$K_{ab}[c] := \sum_{c,d} D_{cd}[c] \langle ac | \frac{1}{r} | db \rangle , \quad (119b)$$

in terms of which, the Fock operator in eq. (117) may be expressed as

$$F_{ab} = \langle a | \hat{h} | b \rangle + 2J_{ab}[c] - K_{ab}[c] . \quad (120)$$

After SCF convergence has been achieved, the *RHF energy* in the finite-dimensional case can be computed using the discretized version of eq. (98):

$$\begin{aligned} E &= 2 \sum_I \varepsilon_I - \sum_{I,J} \sum_{a,b,c,d} \left(2c_{aI}c_{bJ}c_{cI}c_{dJ} \langle ab | \frac{1}{r} | cd \rangle - c_{aI}c_{bJ}c_{cJ}c_{dI} \langle ab | \frac{1}{r} | dc \rangle \right) = \\ &= 2 \sum_I \varepsilon_I - \sum_{I,J} \sum_{a,b,c,d} c_{aI}c_{bJ}c_{cI}c_{dJ} \langle ab | \frac{1}{r} | cd \rangle = \\ &= 2 \sum_i \varepsilon_i - \sum_{a,b,c,d} D_{ac}[c] D_{bd}[c] \langle ab | \frac{1}{r} | cd \rangle , \end{aligned} \quad (121)$$

where a convenient rearrangement of the indices in the two sums has been performed from the first to the second line.

9 Introduction to Gaussian basis sets

In principle, arbitrary functions may be chosen as the χ_a to solve the Roothaan-Hall equations in the previous section, however, in eq. (117), we see that one of the main numerical bottlenecks in SCF calculations arises from the necessity of calculating the $O(M^4)$ four-centre integrals¹ $\langle ab | \frac{1}{r} | cd \rangle$, since the solution of the generalized eigenvalue problem in eq. (116) typically scales only like $O(M^3)$, and there are $O(M^2)$ two-centre $\langle a | \hat{h} | b \rangle$ integrals (see however the next section). Either if these integrals are calculated at each iterative step and directly taken from RAM memory (*direct SCF*) or if they are calculated at the first step, written to disk, and then read from there when needed (*conventional SCF*), an appropriate choice

¹ If the symmetry properties of the integrals are used, the precise number of ERIs is found to be $\frac{1}{8}M(M+1)(M^2+M+2)$ [95].

of the functions χ_a in the finite basis set is essential if accurate results are sought, M is intended to be kept as small as possible and the integrals are wanted to be computed rapidly. When one moves into higher-level theoretical descriptions and the numerical complexity scales with M even more unpleasantly, the importance of this choice greatly increases.

In this section, in order to support that study, we shall introduce some of the concepts involved in the interesting field of basis-set design. For further details not covered here, the reader may want to check refs. [18, 19, 96, 97].

The only analytically solvable molecular problem in non-relativistic quantum mechanics is the *hydrogen-like atom*, i.e., the system formed by a nucleus of charge Z and only one electron (H, He^+ , Li^{2+} , etc.). Therefore, it is not strange that all the thinking about atomic-centred basis sets in quantum chemistry is much influenced by the particular solution to this problem.

The spatial eigenfunctions of the Hamiltonian operator of an hydrogen-like atom, in atomic units and spherical coordinates, read¹

$$\phi_{nlm}(r, \theta, \varphi) = \sqrt{\left(\frac{2Z}{n}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} \left(\frac{2Z}{n}r\right)^l L_{n-l-1}^{2l+1}\left(\frac{2Z}{n}r\right) e^{-Zr/n} Y_{lm}(\theta, \varphi), \quad (122)$$

where n , l and m are the energy, total angular momentum and z -angular momentum quantum numbers, respectively. Their ranges of variation are coupled: all being integers, n runs from 1 to ∞ , l from 0 to $n-1$ and m from $-l$ to l . The function L_{n-l-1}^{2l+1} is a *generalized Laguerre polynomial* [98], for which it suffices to say here that it is of order $n-l-1$ (thus having, in general, $n-l-1$ zeros), and the function $Y_{lm}(\theta, \varphi)$ is a *spherical harmonic*, which is a simultaneous eigenfunction of the total angular momentum operator \hat{l}^2 (with eigenvalue $l(l+1)$) and of its z -component \hat{l}_z (with eigenvalue m).

The hope that the one-electron orbitals that are the solutions of the Hartree-Fock problem in many-electron atoms could not be very different from the ϕ_{nlm} above², together with the powerful chemical intuition that states that ‘atoms-in-molecules are not very different from atoms-alone’, is what mainly drives the choice of the functions χ_a in the basis set, and, in the end, the variational procedure that will be followed is expected to fix the largest failures coming from these too-simplistic assumptions.

Hence, it is customary to choose functions that are centred at atomic nuclei and that partially resemble the exact solutions for hydrogen-like atoms. In this spirit, the first type of AOs to be tried [94] were the *Slater-type orbitals* (STOs), proposed by Slater [99] and Zener [100] in 1930:

$$\chi_a^{\text{STO}}(\mathbf{r}; \mathbf{R}_{\alpha_a}) := \mathcal{N}_a^{\text{STO}} \tilde{Y}_{l_a m_a}^{c,s}(\theta_{\alpha_a}, \varphi_{\alpha_a}) |\mathbf{r} - \mathbf{R}_{\alpha_a}|^{n_a-1} \exp\left(-\zeta_a |\mathbf{r} - \mathbf{R}_{\alpha_a}|\right), \quad (123)$$

where $\mathcal{N}_a^{\text{STO}}$ is a normalization constant and ζ_a is an adjustable parameter. The

¹ For consistency with the rest of the text, the Born-Oppenheimer approximation has been also assumed here. So that the *reduced mass* $\mu := m_e M_N / (m_e + M_N)$ that should enter the expression is considered to be the mass of the electron $\mu \simeq m_e$ (recall that, in atomic units, $m_e = 1$ and $M_N \gtrsim 2000$).

² Note that the N -electron wavefunction of the exact ground-state of a non-hydrogen-like atom depends on $3N$ spatial variables in a way that cannot be written, in general, as a Slater determinant of one-electron functions. The image of single electrons occupying definite orbitals, together with the possibility of comparing them with the one-particle eigenfunctions of the Hamiltonian of hydrogen-like atoms, vanishes completely outside the Hartree-Fock formalism.

index α_a is that of the nucleus at which the function is centred, and, of course, in the majority of cases, there will be several χ_a^{STO} corresponding to different values of a centred at the same nucleus. The integers l_a and m_a can be considered quantum numbers, since, due to the fact that the only angular dependence is in $\tilde{Y}_{l_a m_a}^{c,s}$ (see below for a definition), the STO defined above is still a simultaneous eigenstate of the one-electron angular momentum operators \hat{l}^2 and \hat{l}_z (with the origin placed at \mathbf{R}_{α_a}). The parameter n_a , however, should be regarded as a ‘principal (or energy) quantum number’ only by analogy, since, on the one hand, it does not exist a ‘one-atom Hamiltonian’ whose exact eigenfunctions it could label and, on the other hand, only the leading term of the Laguerre polynomial in eq. (122) has been kept in the STO³.

Additionally, in the above notation, the fact that χ_a^{STO} parametrically depends on the position of a certain α_a -th nucleus has been stressed, and the functions $\tilde{Y}_{l_a m_a}^{c,s}$, which are called *real spherical harmonics* [101] (remember that we want to do real RHF), are defined in terms of the classical *spherical harmonics* $Y_{l_a m_a}$ by

$$\tilde{Y}_{l_a m_a}^c(\theta_{\alpha_a}, \varphi_{\alpha_a}) := \frac{Y_{l_a m_a} + Y_{l_a m_a}^*}{\sqrt{2}} \propto P_{l_a}^{m_a}(\cos \theta_{\alpha_a}) \cos(m_a \varphi_{\alpha_a}), \quad (124a)$$

$$\tilde{Y}_{l_a m_a}^s(\theta_{\alpha_a}, \varphi_{\alpha_a}) := -i \frac{Y_{l_a m_a} - Y_{l_a m_a}^*}{\sqrt{2}} \propto P_{l_a}^{m_a}(\cos \theta_{\alpha_a}) \sin(m_a \varphi_{\alpha_a}), \quad (124b)$$

where c stands for *cosine*, s for *sine*, the functions $P_{l_a}^{m_a}$ are the *associated Legendre polynomials* [98], and the spherical coordinates θ_{α_a} and φ_{α_a} also carry the α_a -label to remind that the origin of coordinates in terms of which they are defined is located at \mathbf{R}_{α_a} . Also note that, using that $\tilde{Y}_{l_a 0}^c = \tilde{Y}_{l_a 0}^s$, there is the same number of real spherical harmonics as of classical ones.

These χ_a^{STO} have some good physical properties. Among them, we shall mention that, for $|\mathbf{r} - \mathbf{R}_{\alpha_a}| \rightarrow 0$, they present a *cusp* (a discontinuity in the radial derivative), as required by Kato’s theorem [102]; and also that they decay at an exponential rate when $|\mathbf{r} - \mathbf{R}_{\alpha_a}| \rightarrow \infty$, which is consistent with the image that, an electron that is taken apart from the vicinity of the nucleus must ‘see’, at large distances, an unstructured point-like charge (see, for example, the STO in fig. 2). Finally, the fact that they do not present radial nodes (due to the aforementioned absence of the non-leading terms of the Laguerre polynomial in eq. (122)) can be solved by making linear combinations of functions with different values of ζ_a ¹.

Now, despite their being good theoretical candidates to expand the MO ϕ_I that make up the N -particle solution of the Hartree-Fock problem, these STOs have serious computational drawbacks: Whereas the two-centre integrals (such as $\langle a | \hat{h} | b \rangle$ in eq. (117)) can be calculated analytically, the four-centre ERIs $\langle ac | 1/r | bd \rangle$ can not [18, 94] if functions like the ones in eq. (123) are used. This fact, which was known as “the nightmare of the integrals” in the first days of computational quantum chemistry [94], precludes the use of STOs in practical ab initio calculations of large molecules.

³ If we notice that, within the set of all possible STOs (as defined in eq. (123)), every hydrogen-like energy eigenfunction (see eq. (122)) can be formed as a linear combination, we easily see that the STOs constitute a complete basis set. This is important to ensure that the Hartree-Fock limit could be actually approached by increasing M .

¹ This way of proceeding renders the choice of the exponent carried by the $|\mathbf{r} - \mathbf{R}_{\alpha_a}|$ part ($n_a - 1$ in the case of the STO in eq. (123)) a rather arbitrary one. As a consequence, different definitions may be found in the literature and the particular exponent chosen in actual calculations turns out to be mostly a matter of computational convenience.

A major step to overcome these difficulties that has revolutionised the whole field of quantum chemistry [55, 94] was the introduction of Cartesian Gaussian-type orbital (cGTO):

$$\chi_a^{\text{cGTO}}(\mathbf{r}; \mathbf{R}_{\alpha_a}) := \mathcal{N}_a^{\text{cGTO}} \left(r^1 - R_{\alpha_a}^1 \right)^{l_a^x} \left(r^2 - R_{\alpha_a}^2 \right)^{l_a^y} \left(r^3 - R_{\alpha_a}^3 \right)^{l_a^z} \exp \left(-\zeta_a |\mathbf{r} - \mathbf{R}_{\alpha_a}|^2 \right), \quad (125)$$

where the r^p and the $R_{\alpha_a}^p$, with $p = 1, 2, 3$, are the Euclidean coordinates of the electron and the α_a -th nucleus respectively, and the integers l_a^x , l_a^y and l_a^z , which take values from 0 to ∞ , are called *orbital quantum numbers*¹.

Although these GTOs do not have the good physical properties of the STOs (compare, for example, the STO and the GTO in fig. 2), in 1950, Boys [103] showed that all the integrals appearing in SCF theory could be calculated analytically if the χ_a had the form in eq. (125). The enormous computational advantage that this entails makes possible to use a much larger number of functions to expand the one-electron orbitals ϕ_i if GTOs are used, partially overcoming their bad short- and long-range behaviour and making the Gaussian-type orbitals the universally preferred choice in SCF calculations [18].

To remedy the fact that the angular behaviour of the Cartesian GTOs in eq. (125) is somewhat hidden, they may be linearly combined to form Spherical Gaussian-type orbital (sGTO):

$$\chi_a^{\text{sGTO}}(\mathbf{r}; \mathbf{R}_{\alpha_a}) := \mathcal{N}_a^{\text{sGTO}} \tilde{Y}_{l_a m_a}^{c,s}(\theta_{\alpha_a}, \varphi_{\alpha_a}) |\mathbf{r} - \mathbf{R}_{\alpha_a}|^{l_a} \exp \left(-\zeta_a |\mathbf{r} - \mathbf{R}_{\alpha_a}|^2 \right), \quad (126)$$

which are proportional to the real spherical harmonic $\tilde{Y}_{l_a m_a}^{c,s}(\theta_{\alpha_a}, \varphi_{\alpha_a})$, and to which the same remarks made in footnote 1 in page 45 for the STOs, regarding the exponent in the $|\mathbf{r} - \mathbf{R}_{\alpha_a}|$ part, may be applied.

The fine mathematical details about the linear combination that relates the Cartesian GTOs to the spherical ones are beyond the scope of this introduction. We refer the reader to refs. [104] and [101] for further information and remark here some points that will have interest in the subsequent discussion.

First, the cGTOs that are combined to make up a sGTO must have all the same value of $l_a := l_a^x + l_a^y + l_a^z$ and, consequently, this sum of the three orbital quantum numbers l_a^x , l_a^y and l_a^z in a particular Cartesian GTO is typically (albeit dangerously) referred to as the *angular momentum* of the function. In addition, apart from the numerical value of l_a , the spectroscopic notation is commonly used in the literature, so that cGTOs with $l_a = 0, 1, 2, 3, 4, 5, \dots$ are called *s*, *p*, *d*, *f*, *g*, *h*, \dots , respectively. Where the first four come from the archaic words *sharp*, *principal*, *diffuse* and *fundamental*, while the subsequent ones proceed in alphabetical order.

Second, for a given $l_a > 1$, there are more Cartesian GTOs $((l_a+1)(l_a+2)/2)$ than spherical ones $(2l_a+1)$, in such a way that, from the $(l_a+1)(l_a+2)/2$ functionally independent linear combinations that can be formed using the cGTOs of angular momentum l_a , only the angular part of $2l_a+1$ of them turns out to be proportional to a real spherical harmonic $\tilde{Y}_{l_a m_a}^{c,s}(\theta_{\alpha_a}, \varphi_{\alpha_a})$; the rest of them are proportional to real spherical harmonic functions with a different value of the angular momentum

¹ Since the harmonic-oscillator energy eigenfunctions can be constructed as linear combinations of Cartesian GTOs, we have that the latter constitute a complete basis set and, like in the case of the STOs, we may expect that the Hartree-Fock limit is approached as M is increased.

quantum number. For example, from the six different d-Cartesian GTOs, whose polynomial parts are x^2, y^2, z^2, xy, xz and yz (using an evident, compact notation), only five different spherical GTOs can be constructed: the ones with polynomial parts proportional to $2z^2 - x^2 - y^2, xz, yz, x^2 - y^2$ and xy [104]. Among these new sGTOs, which, in turn, are proportional (neglecting also powers of r , see footnote 1 in page 45) to the real spherical harmonics $\tilde{Y}_{20}, \tilde{Y}_{21}^c, \tilde{Y}_{21}^s, \tilde{Y}_{22}^c$ and \tilde{Y}_{22}^s , the linear combination $x^2 + y^2 + z^2$ is missing, since it presents the angular behaviour of an s-orbital (proportional to \tilde{Y}_{00}).

Finally, let us remark that, whereas Cartesian GTOs in eq. (125) are easier to be coded in computer applications than sGTOs [104], it is commonly accepted that these spurious spherical orbitals of lower angular momentum that appear when cGTOs are used do not constitute efficient choices to be included in a basis set [101] (after all, if we wanted an additional s-function, why include it in such an indirect and clumsy way instead of just designing a specific one that suits our particular needs?). Consequently, the most common practice in the field is to use Cartesian GTOs removing from the basis sets the linear combinations such as the $x^2 + y^2 + z^2$ above.

Now, even if the integrals involving cGTOs can be computed analytically, there are still $O(M^4)$ of them in a SCF calculation. For example, in the model dipeptide HCO-L-Ala-NH₂, which is commonly used to mimic an alanine residue in a protein [73, 76, 105, 106], there are 62 electrons and henceforth 31 RHF spatial orbitals ϕ_I . If a basis set with only 31 functions is used (this is a lower bound that will be rarely reached in practical calculations due to symmetry issues, see below), near a million of four-centre $\langle ac | 1/r | bd \rangle$ integrals must be computed. This is why, one must use the freedom that remains once the decision of sticking to cGTOs has been taken (namely, the choice of the *exponents* ζ_a and the *angular momentum* l_a) to design basis sets that account for the relevant behaviour of the systems studied while keeping M below the ‘pain threshold’.

The work by Nobel Prize John Pople’s group has been a major reference in this discipline, and their STO- n G family [107], together with the split-valence Gaussian basis sets, 3-21G, 4-31G, 6-31G, etc. [108–115], shall be used here to exemplify some relevant issues. However, note that most of the concepts introduced are also applicable to more modern basis sets, such as those by Dunning [116].

To begin with, let us recall that the short- and long-range behaviour of the Slater-type orbitals in eq. (123) is better than that of the more computationally efficient GTOs. In order to improve the physical properties of the latter, it is customary to linearly combine M_a Cartesian GTOs, denoted now by ξ_a^μ ($\mu = 1, \dots, M_a$), and termed *primitive Gaussian-type orbitals* (PGTO), having the same atomic centre \mathbf{R}_{α_a} , the same set of orbital quantum numbers, l_a^x, l_a^y and l_a^z , but different exponents ζ_a^μ , to make up a *contracted Gaussian-type orbitals* (CGTO), defined by

$$\chi_a(\mathbf{r}; \mathbf{R}_{\alpha_a}) := \sum_{\mu}^{M_a} g_a^\mu \xi_a^\mu(\mathbf{r}; \mathbf{R}_{\alpha_a}) = \left(r^1 - R_{\alpha_a}^1\right)^{l_a^x} \left(r^2 - R_{\alpha_a}^2\right)^{l_a^y} \left(r^3 - R_{\alpha_a}^3\right)^{l_a^z} \sum_{\mu}^{M_a} g_a^\mu \mathcal{N}_a^\mu \exp\left(-\zeta_a^\mu |\mathbf{r} - \mathbf{R}_{\alpha_a}|^2\right), \quad (127)$$

where the normalization constants \mathcal{N}_a^μ have been kept inside the sum because they typically depend on ζ_a^μ . Also, we denote now by M_C the number of contracted GTOs and, by $M_P := \sum_a M_a$, the number of primitive ones.

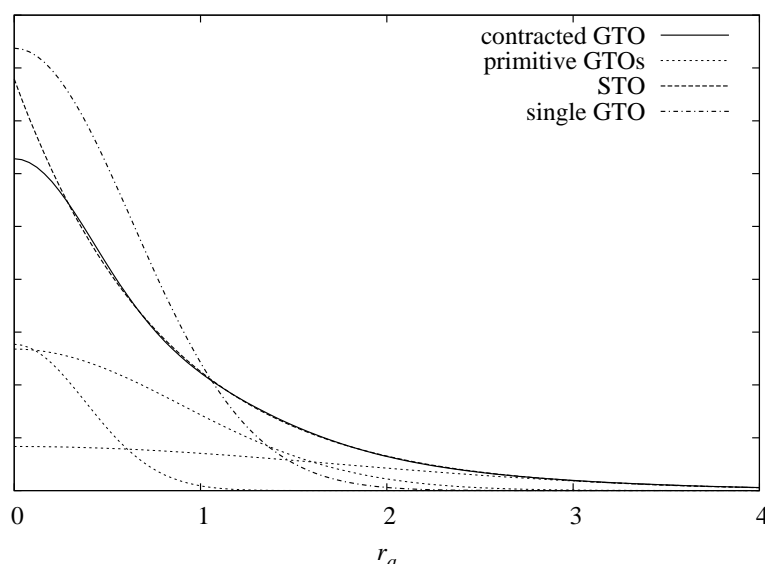


Figure 2. Radial behaviour of the 1s-contracted GTO of the hydrogen atom in the STO-3G basis set [107], the three primitive GTOs that form it, the STO that is meant to be approximated and a single GTO with the same norm and exponent as the STO. The notation r_a is shorthand for the distance to the α_a -th nucleus $|\mathbf{r} - \mathbf{R}_{\alpha_a}|$.

In the STO- n G family of basis sets, for example, n primitive GTOs are used for each contracted one, fitting the coefficients g_a^μ and the exponents ζ_a^μ to resemble the radial behavior of Slater-type orbitals [107]. In fig. 2, the 1s-contracted GTO (see the discussion below) of the hydrogen atom in the STO-3G basis set is depicted, together with the three primitive GTOs that form it and the STO that is meant to be approximated¹. We can see that the contracted GTO has a very similar behavior to the STO in a wide range of distances, while the single GTO that is also shown in the figure (with the same norm and exponent as the STO) has not.

Typically, the fitting procedure that leads to contracted GTOs is performed on isolated atoms and, then, the already mentioned chemical intuition that states that ‘atoms-in-molecules are not very different from atoms-alone’ is invoked to keep the linear combinations fixed from there on. Obviously, better results would be obtained if the contraction coefficients were allowed to vary. Moreover, the number of four-centre integrals that need to be calculated depends on the number of *primitive* GTOs (like $O(M_p^4)$), so that we have not gained anything on this point by contracting. However, the size of the variational space is M_C (i.e., the number of *contracted* GTOs), in such a way that, once the integrals $\langle ac | 1/r | bd \rangle$ are calculated (for non-direct SCF), all subsequent steps in the iterative self-consistent procedure scale as powers of M_C . Also, the disk storage (again, for non-direct schemes) depends on the number of *contracted* GTOs and, frequently, it is the disk storage and not the CPU time the limiting factor of a calculation.

An additional chemical concept that is usually defined in this context and that is needed to continue with the discussion is that of *shell*: Atomic shells in quantum chemistry are defined analogously to those of the hydrogen atom, so that each electron is regarded as ‘filling’ the multi-electron atom ‘orbitals’ according to Hund’s

¹ Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database at <http://www.emsl.pnl.gov/forms/basisform.html>, Version 02/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact Karen Schuchardt for further information.

Table 3. Exponents ζ_a^μ and contraction coefficients g_a^μ of the primitive Gaussian shells that make up the three different contracted ones in the STO-3G basis set for carbon (see ref. [107] and footnote 1 in page 48). The exponents of the 2s- and 2p-shells are constrained to be the same.

1s-shell		ζ_a^μ	2sp-shell	
ζ_a^μ	g_a^μ		g_a^μ (s)	g_a^μ (p)
71.6168370	0.15432897	2.9412494	0.15591627	-0.09996723
13.0450960	0.53532814	0.6834831	0.60768372	0.39951283
3.5305122	0.44463454	0.2222899	0.39195739	0.70115470

Table 4. Exponents ζ_a^μ and contraction coefficients g_a^μ of the primitive Gaussian shells that make up the three different constrained ones in the 6-31G basis set for carbon (see ref. [109] and footnote 1 in page 48). In the 1s-shell, there is only one contracted Gaussian shell made by six primitive ones, whereas, in the 2s- and 2p-valence shells, there are two CGSs, one of them made by three PGs and the other one only by a single PGS. The exponents of the 2s- and 2p-shells are constrained to be the same.

1s-shell		ζ_a^μ	2sp-shell	
ζ_a^μ	g_a^μ		g_a^μ (p)	g_a^μ (s)
3047.52490	0.0018347	7.8682724	-0.1193324	0.0689991
457.36951	0.0140373	1.8812885	-0.1608542	0.3164240
103.94869	0.0688426	0.5442493	1.1434564	0.7443083
29.21015	0.2321844			
9.286663	0.4679413	0.1687144	1.0000000	1.0000000
3.163927	0.3623120			

rules [117]. Hence, the *occupied shells* of carbon, for example, are defined to be 1s, 2s and 2p, whereas those of, say, silicon, would be 1s, 2s, 2p, 3s and 3p. Each shell may contain $2(2l + 1)$ electrons if complete, where $2l + 1$ accounts for the orbital angular momentum multiplicity and the 2 factor for that of electron spin.

Thus, using these definitions, all the basis sets in the aforementioned STO- n G family are *minimal*; in the sense that they are made up of only $2l + 1$ contracted GTOs for each completely or partially occupied shell, so that the STO- n G basis sets for carbon, for example, contain two s-type contracted GTOs (one for the 1s- and the other for the 2s-shell) and three p-type ones (belonging to the 2p-shell). Moreover, due to rotational-symmetry arguments in the isolated atoms, all the $2l + 1$ functions in a given shell are chosen to have the same exponents and the same contraction coefficients, differing only on the polynomial that multiplies the Gaussian part. Such $2l + 1$ CGTOs shall be said to constitute a *Gaussian shell* (GS), and we shall also distinguish between the *primitive* (PGS) and *contracted* (CGS) versions.

In table 3, the exponents ζ_a^μ and the contraction coefficients g_a^μ of the primitive GTOs that make up the three different shells in the STO-3G basis set for carbon are presented (see ref. [107] and footnote 1 in page 48). The fact that the exponents ζ_a^μ in the 2s- and 2p-shells are constrained to be the same is a particularity of some basis sets (like this one) which saves some computational effort and deserves no further attention.

Next, let us introduce a common notation that is used to describe the contraction scheme: It reads (*primitive shells*) / [*contracted shells*], or alternatively (*primitive shells*) \rightarrow [*contracted shells*]. According to it, the STO-3G basis set for carbon, for example, is denoted as (6s,3p) \rightarrow [2s,1p], or (6,3) \rightarrow [2,1]. Moreover, since for organic molecules it is frequent to have only hydrogens and the 1st-row atoms C, N and O (whose occupied shells are identical)¹, the notation is typically extended and the two groups of shells are separated by a slash; as in (6s,3p/3s) \rightarrow [2s,1p/1s] for STO-3G.

¹ In proteins, one may also have sulphur in cysteine and methionine residues.

The first improvement that can be implemented on a minimal basis set such as the ones in the STO- n G is the *splitting*, which consists in including more than one Gaussian shell for each occupied one. If the splitting is evenly performed, i.e., each shell has the same number of GSs, then the basis set is called *double zeta* (DZ), *triple zeta* (TZ), *quadruple zeta* (QZ), *quintuple zeta* (5Z), *sextuple zeta* (6Z), and so on; where the word *zeta* comes from the Greek letter ζ used for the exponents. A hypothetical TZ basis set in which each CGTO is made by, say, four primitive GTOs, would read (24s,12p/12s) \rightarrow [6s,3p/3s] in the aforementioned notation.

At this point, the already familiar intuition that says that ‘atoms-in-molecules are not very different from atoms-alone’ must be refined with another bit of chemical experience and qualified by noticing that ‘core electrons are less affected by the molecular environment and the formation of bonds than valence electrons’¹. In this spirit, the above evenness among different shells is typically broken, and distinct basis elements are used for the energetically lowest lying (*core*) shells than for the highest lying (*valence*) ones.

On one side, the contraction scheme may be different. In which case, the notation used up to now becomes ambiguous, since, for example, the designation (6s,3p/3s) \rightarrow [2s,1p/1s], that was said to correspond to STO-3G, would be identical for a *different* basis set in which the 1s-Gaussian shell of heavy atoms be formed by 4 PGSs and the 2s-Gaussian shell by 2 PGSs (in 1st-row atoms, the 2s- and 2p-shells are defined as valence and the 1s-one as core, while in hydrogen atoms, the 1s-shell is a valence one). This problem can be solved by explicitly indicating how many primitive GSs form each contracted one, so that, for example, the STO-3G basis set is denoted by (33,3/3) \rightarrow [2,1/1], while the other one mentioned would be (42,3/3) \rightarrow [2,1/1] (we have chosen to omit the angular momentum labels this time).

The other point at which the core and valence Gaussian shells may differ is in their respective ‘zeta quality’, i.e., the basis set may contain a different number of contracted Gaussian shells in each case. For example, it is very common to use a single CGS for the core shells and a multiple splitting for the valence ones. These type of basis sets are called *split-valence* and the way of naming their quality is the same as before, except for the fact that a capital V, standing for *valence*, is added either at the beginning or at the end of the acronyms DZ, TZ, QZ, etc., thus becoming VDZ, VTZ, VQZ, etc. or DZV, TZV, QZV, etc.

Pople’s 3-21G [113], 4-31G [108], 6-31G [109] and 6-311G [112] are well-known examples of split-valence basis sets that are commonly used for SCF calculations in organic molecules and that present the two characteristics discussed above. Their names indicate the contraction scheme, in such a way that the number before the dash represents how many primitive GSs form the single contracted GSs that is used for core shells, and the numbers after the dash how the valence shells are contracted, in much the same way as the notation in the previous paragraphs. For example, the 6-31G basis set (see table 4), contains one CGS made up of six primitive GSs in the 1s-core shell of heavy atoms (the 6 before the dash) and two CGS, formed by three and one PGSs respectively, in the 2s- and 2p-valence shells of heavy atoms and in the 1s-shell of hydrogens (the 31 after the dash). The 6-311G basis set, in turn, is just the same but with an additional single-primitive Gaussian shell of functions in the valence region. Finally, to fix the concepts discussed, let us mention that, using the notation introduced above, these two basis sets may be written as (631,31/31) \rightarrow [3,2/2] and (6311,311/311) \rightarrow [4,3/3], respectively.

Two further improvements that are typically used and that may also be incorpo-

¹ Recall that, for the very concept of ‘core’ or ‘valence electrons’ (actually for any label applied to a single electron) to have any sense, we must be in the Hartree-Fock formalism (see footnote 2 in page 44).

rated to Pople's split-valence basis sets are the addition of *polarization* [110,111] or *diffuse functions* [111,114,115]. We shall discuss them both to close both this section and the work.

Up to now, neither the contraction nor the splitting involved GTOs of larger angular momentum than the largest one among the occupied shells. However, the molecular environment is highly anisotropic and, for most practical applications, it turns out to be convenient to add these *polarization* (large angular momentum) Gaussian shells to the basis set, since they present lower symmetry than the GSs discussed in the preceding paragraphs. Typically, the polarization shells are single-primitive GSs and they are denoted by adding a capital P to the end of the previous acronyms, resulting into, for example, DZP, TZP, VQZP, etc., or, say, DZ2P, TZ3P, VQZ4P if more than one polarization shell is added. In the case of Pople's basis sets [110,111], these improvements are denoted by specifying, in brackets and after the letter G, the number and type of the polarization shells separating heavy atoms and hydrogens by a comma¹. For example, the basis set 6-31G(2df,p) contains the same Gaussian shells as the original 6-31G plus two d-type shells and one f-type shell centred at the heavy atoms, as well as one p-type shell centred at the hydrogens.

Finally, for calculations in charged species (specially anions), where the charge density extends further in space and the tails of the distribution are more important to account for the relevant behaviour of the system, it is common to augment the basis sets with *diffuse functions*, i.e., single-primitive Gaussian shells of the same angular momentum as some preexisting one but with a smaller exponent ζ than the smallest one in the shell. In general, this improvement is commonly denoted by adding the prefix *aug-* to the name of the basis set. In the case of Pople's basis sets, on the other hand, the insertion of a plus sign '+' between the contraction scheme and the letter G denotes that the set contains one diffuse function in the 2s- and 2p-valence shells of heavy atoms. A second + indicates that there is another one in the 1s-shell of hydrogens. For example, one may have the doubly augmented (and doubly polarized) 6-31++G(2d,2p) basis set.

10 Modern developments: An introduction to linear-scaling methods

The ground-breaking advances reviewed in the previous sections allow to routinely calculate, using Hartree-Fock SCF methods, physical properties of molecules of tens of atoms in present day computers. However, the simplest algorithms that can be devised to perform the limiting steps in such calculations are far from optimal. This large room for improvement, which may be enough to accommodate the exciting possibility of linearly scaling approaches that could open the door to thousands atoms computations, has been steering many innovative lines of research in the last years

To close this review, we shall briefly outline here some of the hottest areas of modern development related to the topics discussed, specially those aimed to the reduction of the in principle $O(M^4)$ complexity associated to the calculation of the $\langle ac | 1/r | bd \rangle$ ERIs in eq. (118), as well as the $O(M^3)$ cost of the diagonalization of the Fock operator in eq. (116)². For wider reviews on these topics, we recommend

¹ There also exists an old notation for the addition of a single polarization shell per atom that reads 6-31G** and that is equivalent to 6-31G(d,p).

² In principle, careful distinction must be made between the number of primitive GTOs M_P and the number of contracted ones M_C (see the previous section). However, since in this section we will be dealing only with approximate scalings without worrying much about the prefactor, the 'neutral' notation M has been chosen to denote a quantity which is linear on both M_P and M_C .

to the interested reader the accounts in refs. [118–120].

The first class of attempts to reduce the cost associated to the calculation of the $O(M^4)$ ERIs $\langle ac | 1/r | bd \rangle$ are those aimed to the improvement of the algorithms for analytically calculating them *without approximations*. There are basically two issues that render the construction of these algorithms a non trivial task: First, the fact that the only four-center ERIs that can be straightforwardly computed are the ones corresponding to a product of four s-type GTOs, while higher angular momentum ERIs may be obtained from them in a non unique way.

After using the *Gaussian product rule* [19] (see the previous section for the notation used)

$$\begin{aligned} & \exp\left(-\zeta_a |\mathbf{r} - \mathbf{R}_{\alpha_a}|^2\right) \exp\left(-\zeta_b |\mathbf{r} - \mathbf{R}_{\alpha_b}|^2\right) = \\ & \underbrace{\exp\left(-\frac{\zeta_a \zeta_b}{\zeta_a + \zeta_b} |\mathbf{R}_{\alpha_a} - \mathbf{R}_{\alpha_b}|^2\right)}_{\mathcal{E}_{ab}} \exp\left(-(\zeta_a + \zeta_b) \left| \mathbf{r} - \frac{\zeta_a \mathbf{R}_{\alpha_a} + \zeta_b \mathbf{R}_{\alpha_b}}{\zeta_a + \zeta_b} \right|^2\right) \end{aligned} \quad (128)$$

which allows the four-centre integral to be turned into a two-center one, and whose absence for the case of STOs is the essential reason for their being non practical, we can turn the 6-dimensional ERI into a simple one-dimensional integral that can be readily calculated by different means [121,122]:

$$\begin{aligned} & \iint \frac{e^{-\zeta_a |\mathbf{r} - \mathbf{R}_{\alpha_a}|^2} e^{-\zeta_b |\mathbf{r}' - \mathbf{R}_{\alpha_b}|^2} e^{-\zeta_c |\mathbf{r} - \mathbf{R}_{\alpha_c}|^2} e^{-\zeta_d |\mathbf{r}' - \mathbf{R}_{\alpha_d}|^2}}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' = \\ & A_{abcd} \int_0^1 e^{-B_{abcd} x^2} dx, \end{aligned} \quad (129)$$

with

$$A_{abcd} := \mathcal{E}_{ac} \mathcal{E}_{bd} \frac{2\pi^{5/2}}{(\zeta_a + \zeta_c)(\zeta_b + \zeta_d)(\zeta_a + \zeta_c + \zeta_b + \zeta_d)^{1/2}}, \quad (130a)$$

$$B_{abcd} := \left(\frac{(\zeta_a + \zeta_c)(\zeta_b + \zeta_d)}{\zeta_a + \zeta_c + \zeta_b + \zeta_d} \right)^2 \left(\frac{\zeta_a \mathbf{R}_{\alpha_a} + \zeta_c \mathbf{R}_{\alpha_c}}{\zeta_a + \zeta_c} - \frac{\zeta_b \mathbf{R}_{\alpha_b} + \zeta_d \mathbf{R}_{\alpha_d}}{\zeta_b + \zeta_d} \right)^2. \quad (130b)$$

Such an integral is called a *fundamental* ERI and, as we said, ERIs involving GTOs with higher angular momentum than $l = 0$ are obtained from this fundamental one through iterative differentiations of eq. (129) with respect to the nuclear positions, leading to recurrence relations expressing ERIs of a given angular momentum as a function of the lower angular momentum ones. The particular flavour of these recurrence relations that is used is one of the matters in which the algorithms for calculating ERIs differ.

The second issue that renders the construction of algorithms for computing ERIs non-trivial is related to how the contraction of primitive GTOs is handled. In the geminal paper by Boys [103], the most naive procedure was suggested, namely, the conversion of each ERI of contracted GTOs into a quadruple sum of ERIs of primitive GTOs. However, this does not take profit, for example, from the fact, mentioned in the previous section, that all CGTOs in the same contracted Gaussian shell are formed by PGTOs with the same set of exponents ζ_a^μ . Much profit can be

taken from this and other constraints and, in fact, the cost-scaling profile of each algorithm with the contraction degree M_a of the CGTOs is strongly correlated to the moment at which the transformation between CGTOs and PGTOs is performed [122].

Among the most used of these analytical algorithms, we can mention the Pople-Hehre (PH) one [123], which additionally exploits the fact that for each four-center ERI of low angular momentum there is a privileged Cartesian axis system in which many primitive integrals vanish by symmetry; the *McMurchie-Davidson* (MD) approach [124], which avoids the rotation in PH thus being more efficient for high angular momentum ERIs; the *Obara-Saika-Schlegel* (OSS) algorithm [125, 126]; and a better defined and improved version of it: the *Head-Gordon-Pople* (HGP) algorithm [127]. Finally, if the moment at which the contraction is handled is chosen dynamically depending on the type of GTO appearing in the ERI, we have the PRISM modifications of MD and HGP: the MD-PRISM [121, 128, 129] and HGP-PRISM [122] algorithms, as well as a generalization of all the previous methods, called COLD PRISM [131].

Now, even if we implement any of these efficient methods for calculating the ERIs, there is still $O(M^4)$ of them. This scaling is simply too harsh for a too large class of applications. Therefore, the next natural step is to try to devise approximate methods that minimize the necessary decrease in accuracy at the same time that maximize the savings in computer time. Thanks to the particular characteristics of the ERIs, the Gaussian basis functions and the concrete physical problem intended to solve¹, this endeavour has been successfully pursued by many researchers and the ‘holy grail’ [18] of linear scaling with M is asymptotically approaching current calculations in large systems. Here, we shall discuss the basic issues that make this possible. For more in depth reviews, we point the readers to the accounts in references [119, 120, 132].

The first point to consider in order to reduce the $O(M^4)$ scaling attains the so-called *radial overlap*. If we take a look at eq. (127), we can see that, irrespective of the polynomial prefactor which contains all the angular dependence of the GTO, every function $\chi_a(\mathbf{r})$ contains a radial exponential part (actually, a sum of exponentials). Therefore, if we consider any product $\chi_a(\mathbf{r})\chi_b(\mathbf{r})$ of two GTOs, we will always find a multiplying sum of terms such as those depicted in the expression for the Gaussian product rule in (128). The exponential decay of all quantities \mathcal{E}_{ab} in those terms with the distance $|\mathbf{R}_{\alpha_a} - \mathbf{R}_{\alpha_b}|$ between the nuclei on which the GTOs are centred indicates that, among the $M(M+1)/2$ possible pair products $\chi_a(\mathbf{r})\chi_b(\mathbf{r})$, only $O(M)$ of them will be non-negligible. To see this, note that if we fix (say) a , the only values of b that will yield a non-negligible product $\chi_a(\mathbf{r})\chi_b(\mathbf{r})$ are those for which $|\mathbf{R}_{\alpha_a} - \mathbf{R}_{\alpha_b}|$ is ‘small’. Since atoms do not interpenetrate in most of the conformations that shall be studied, this can only happen for a number of different b ’s which is not $O(M)$ but a constant independent of the size of the molecule. There are M different possible values of a for which the above reasoning can be repeated, and the result follows.

As a consequence, if only $O(M)$ pairs $\chi_a(\mathbf{r})\chi_b(\mathbf{r})$ are non-negligible, then only $O(M^2)$ ERIs $\langle ac | 1/r | bd \rangle$ may in principle contribute in a significant way to the Fock operator in eq. (117) and not $O(M^4)$. (For similar estimations based on slightly different hypotheses, see [134, 135].) One must also note that the discussion is complicated by the fact that the ERIs do not appear alone, but contracted with the density matrix elements D_{cd} . This allows for further improvements of the scaling

¹ The locality of many-electron Quantum Mechanics [132], which is related to the *nearsightedness* concept introduced by Kohn [133], is the main property that allows to finally achieve linearity.

beyond $O(M^2)$ which are discussed later in this section.

Now, knowing that most of the ERIs are too small to be relevant, *we do not know which ones* and, if we calculated the $O(M^4)$ of them in order to spot the little ones, then we would have not gained anything. This simple argument shows the necessity of finding a set of *estimators* that allow us to selectively drop ERIs *without calculating them*. Of course, the number of estimators must also scale at worst like $O(M^2)$ in order for the scheme to be useful.

One of the first and simplest such estimators, was introduced by Almlöf et al. at the same time that they proposed the Direct SCF method [136]. In their scheme, each ERI $\langle ac|1/r|bd\rangle$ was approximated by the corresponding radial overlap factor $\mathcal{E}_{ab}\mathcal{E}_{cd}$ (note that there are only $O(M^2)$ numbers \mathcal{E}_{ab}). Although this estimator was relatively successful, it presented the important drawback of not being an upper-bound for the ERIs, thus rendering the control of errors an a priori impossible task. In order to overcome this problem, Häser and Ahlrichs [137] later proposed a different estimator which can be assured to be always greater than the associated ERI. Following them, after the proof of positive definiteness by Roothaan [79], the electrostatic interaction energy between two continuous charge distributions,

$$\int \frac{\rho_1(\mathbf{r})\rho_2(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' ,$$

can be easily shown to satisfy the properties of an inner product. Hence, if we choose $\rho_1 := \chi_a\chi_b$, and $\rho_2 := \chi_c\chi_d$, we can use the well-known Schwarz inequality to show that

$$\begin{aligned} \langle ac|\frac{1}{r}|bd\rangle &= \int \frac{\rho_1(\mathbf{r})\rho_2(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \leq \\ &\leq \left(\int \frac{\rho_1(\mathbf{r})\rho_1(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \right)^{1/2} \left(\int \frac{\rho_2(\mathbf{r})\rho_2(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \right)^{1/2} = \\ &\langle aa|\frac{1}{r}|bb\rangle^{1/2} \langle cc|\frac{1}{r}|dd\rangle^{1/2} . \end{aligned} \quad (131)$$

In such a way that, by calculating only the $M(M+1)/2$ different two-index ERIs in the last term, we can safely bound from above the whole set containing $O(M^4)$ of them.

After these seminal works, more sophisticated and tighter bounds have been developed through the years [138–140], their application being nowadays routine in Quantum Chemistry packages.

In a second generation of methods, the $O(M^2)$ formal scaling achieved in practical calculations [141] by using the above ideas has been recently attacked. To this end, more specifically physical properties of the problem are exploited (see footnote 1 in page 53), and different strategies are used to deal with the Coulomb and exchange parts of the Fock operator in (117). The main difference between the behaviours of these two contributions lies in the way in which the ERIs are contracted with the density matrix D_{cd} : In the ‘classical’ Coulomb part, the relative sizes of the ERIs $\langle ac|1/r|bd\rangle$ are largely correlated with the relative sizes of the associated elements D_{cd} [142], so that no decrease in the $\sim M^2$ scaling is expected from density matrix considerations. Differently, in the exchange terms, the fact that the elements D_{cd} couple ‘exchanged’ indices in the $\langle ac|1/r|bd\rangle$ ERIs produces cancellations which make this type of ‘quantum’ contributions rather short-range (for non-metallic

species) [142–144]. Hence, as for every short-range interaction, a number of terms scaling *linearly* with the system size is expected. Apart from the different treatment that this difference suggests, note that the separation of the Coulomb and exchange tasks allows for improved parallelization the computer codes [119].

For the exchange part, the aforementioned short-range behaviour allows to devise $O(M)$ algorithms just by intelligently ordering the loops in which the ERIs are calculated. The fine details of these methods are rather technical and they are beyond the scope of this review. We point the reader to [144–147] and references therein for further information.

In the Coulomb case, on the other hand, $O(M^2)$ terms still enter the sum in (117), and more physically-based approximations must be used. The *continuous fast multipole method* (CFMM) by White et al. [148], for example, is probably one of the most celebrated algorithms for calculating the Coulomb part of the Fock operator. It is a generalization for continuous charge distributions of the *fast multipole method* (FMM), introduced by Greengard and Rokhlin [149] in a ground-breaking paper and aimed for point-like charges. In both FMM and CFMM, a clever hierarchical tree-like division of the space into cells is performed¹, and the far away regions are approximated via truncated multipole expansions.

These two ingredients, which allow to calculate the Coulomb contribution in $O(M)$ steps for large systems, are common to most of the fast Coulomb algorithms². Despite this similarity, the room for improvement seems still large enough to accommodate a vigorous field with many publications appearing each year. Let us mention here, for example, the *generalized cell multipole method* (GCMM) by Kutteh et al., which can use moments higher than monopole [151]; the *quantum chemical tree code* (QCTC) by Challacombe et al. [152], which independently thresholds ‘bra’ and ‘ket’ distributions; and the *Gaussian very fast multipole method* (GvFMM) by Strain et al. [153], which benefits from the idea, introduced in [154] for point-like charges, of using a dynamical maximum angular momentum to further speed up the calculations

Note however, that the near-field contributions in these methods are still calculated without approximations and represent a great portion of the computer time. In this line, some modern algorithms are appearing to alleviate this part of the work, such as, for example, the *J matrix engine* by White and Head-Gordon [155], which uses and improves the ideas discussed in the first part of this section about analytically calculating the ERIs; or the method by Izmaylov et al. [156], which implements a hierarchy of screening levels to eliminate negligible integrals. According to recent reports [157], the combination of CFMM, with the J matrix engine technique and with the *Fourier transform Coulomb* (FTC) method by Füsti-Molnár and Pulay [158] is nowadays probably the fastest way for assembling the Coulomb matrix.

Now, once the construction process of the whole Fock matrix F_{ab} (via its Coulomb and exchange parts) has been cast into the form of an $O(M)$ algorithm, the importance of the second rate-limiting step in SCF procedures, the diagonalization of F_{ab} , comes into focus. Although the prefactor of the (in principle) $O(M^3)$ diagonalization step is very small and, for systems of less than a few thousands of atoms, the absolute time spent on it is smaller than the one needed for the formation of the Fock matrix [132, 141, 157, 159], it is clear that, in the long run, it will dominate the calculations in larger systems and will become the relevant bottleneck [159].

We shall point out the essentials regarding the methods aimed to reduce the

¹Basically, a truncation of a *Barnes and Hut* (BH) tree [150].

²For a review of different approaches, see [132].

scaling the diagonalization step. For more in depth reviews on the topic, we suggest to the reader the accounts in [132, 147, 159, 160].

The $O(M^3)$ complexity of classical diagonalization methods (such as the Givens-Householder one [161]) can be easily understood if we think that the core of the algorithm is just the multiplication of $M \times M$ matrices. Nevertheless, if the matrices multiplied are *sparse*, i.e., they have a number of non-negligible elements that scale not as $O(M^2)$ but as $O(M)$, then the product can be obtained in $O(M)$ steps. As a result of the already mentioned locality properties of many-body Quantum Mechanics (see footnote 1 in page 53), some matrices appearing in the SCF methods discussed in this review, namely, the density matrix D_{ab} and the Fock operator F_{ab} , are indeed sparse if the system is non-metallic, i.e., if it has a non-vanishing HOMO-LUMO gap [159, 162]. The idea behind most of the modern algorithms for achieving (or avoiding) diagonalization with $O(M)$ effort consists essentially in performing all operations using only these (local) sparse matrices, and avoiding (non-local) dense ones, such as the MO coefficients matrix c_{ab} .

The attempts to improve the scaling of the diagonalization steps fall basically in two groups [163]. In the first one, profit is taken from the use of MOs which, instead of being extended over the whole molecule, such as the canonical orbitals used in the previous sections, are *localized* in a small region of space. In these class of methods, diagonalization [164] or pseudo-diagonalization (annihilation of the occupied-virtual blocks of the Fock matrix) [165, 166] is still performed, and the $O(M)$ scaling is achieved because the representation of all operators in the basis of localized MOs is sparse. The second group of algorithms do not use the MOs as variables but the density matrix itself. Among them, two subfamilies of methods may be found: In the first one, the search for the optimal density matrix is simply treated as a standard optimization problem, being the score function the HF energy, and the variables the density matrix elements or a set of parameters of some suitable truncated expansion of it [171]. See, for example, the approaches by Ochsenfeld and Head-Gordon [167], by Salek et al. [168], by Millam and Scuseria [169], or by Ordejón et al. [170]. The other subfamily of density matrix-based methods use iterative procedures, in such a way that, at each step, the Fock operator is considered fixed and the equations are solved for the density matrix (much in the spirit of the MOs-based algorithms discussed in the previous sections). In this group, we can find, for example, the approach in [171] using the Lanczos algorithm [172, 173], or the method by Helgaker et al. [174].

These algorithms, combined with new strategies that also avoid diagonalization and improve SCF convergence properties, such as the one described in [175], represent the final step towards linear Hartree-Fock methods in Quantum Chemistry.

To close this section, although we have been concerned, up to now, with the calculation of the electronic ground-state given a fixed position of the nuclei, let us stress that it is also very common to use quantum chemical methods for finding the local energy minima of molecules. To this end, geometry optimizations must be performed and not only must we be able to compute the energy of the molecule, but also their derivatives with respect to the nuclear coordinates¹. Additionally, these derivatives are also needed to do *ab initio* molecular dynamics in the ground-state Born-Oppenheimer PES [176].

The most naive approach, namely, the computation of the gradient of the energy

¹Monte Carlo methods, in which the derivatives of the energy function are not needed, could also be used. However, although they are efficient (and often the only choice) for global optimization problems, most of the minimizations performed in Quantum Chemistry aim only for the closest local minimum. In such a case, methods which do need the derivatives, such as Newton-Raphson, steepest-descent or conjugate-gradient, usually perform better.

$E(\underline{\mathbf{R}})$ (using a simpler notation for it than $V_N^{\text{eff}}(\underline{\mathbf{R}})$ in (10)) by finite differences, is very inefficient for anything but the smallest molecules. To see this, one just need to notice that the gradient has as many components as the system degrees of freedom n . Hence, in order to obtain it, say, at a point $\underline{\mathbf{R}}_0$, we would have to compute $n + 1$ times a single point energy, in order to know $E(\underline{\mathbf{R}}_0)$ and $E(\underline{\mathbf{R}}_0 + \Delta\underline{\mathbf{R}}_i)$, being $\Delta\underline{\mathbf{R}}_i$, with $i = 1, \dots, n$, a small displacement in each of the nuclear degrees of freedom.

This drawback was overcome in the late 60s by Pulay and others (see [177, 178] and references therein) with the introduction of the so-called *analytical derivatives*, in which the gradient (and higher-order derivatives) are expressed, like the energy itself, just as a function of ERIs involving the wavefunction at the point $\underline{\mathbf{R}}_0$. This marked an inflexion point in the development of optimization and molecular dynamics algorithms that continues nowadays, as analytical derivatives are routinely introduced together with almost any new method for calculating the energy. In relation to the improvements reviewed in this work, for example, let us note that, in [179], the extension of the J matrix engine method to calculate the derivatives of the Coulomb part with respect of the nuclei coordinates is introduced; in [147], linear scaling exchange gradients are developed; analytic derivatives for the GvFMM are provided in [180]; the HGP algorithm for calculating the ERIs is extended to the computation of derivatives of the ERIs as well in [127]; and we may find similar developments for the FTC method [157] or for algorithms that achieve diagonalization with linear effort [167, 170].

Acknowledgments

This work has been supported by the research projects E24/3 and PM048 (Aragón Government), MEC (Spain) FIS2006-12781-C02-01 and MCyT (Spain) FIS2004-05073-C04-01. P. Echenique has been supported by a BIFI research contract and by a MEC (Spain) postdoctoral grant.

Appendix A: Functional derivatives

A *functional* $\mathcal{F}[\Psi]$ is a mapping that takes functions to numbers (in this work, only functionals in the real numbers are considered):

$$\begin{aligned} \mathcal{F} : \mathcal{G} &\longrightarrow \mathbb{R} \\ \Psi &\longmapsto \mathcal{F}[\Psi] \end{aligned}$$

For example, if the function space \mathcal{G} is the Hilbert space of square-integrable functions L^2 (the space of states of quantum mechanics), the objects in the domain of \mathcal{F} (i.e., the functions in L^2) can be described by infinite-tuples (c_1, c_2, \dots) of complex numbers and \mathcal{F} may be pictured as a function of infinite variables.

When dealing with function spaces \mathcal{G} that meet certain requirements¹, the limit on the left-hand side of the following equation can be written as the integral on the right-hand side:

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{F}[\Psi_0 + \varepsilon \delta\Psi] - \mathcal{F}[\Psi_0]}{\varepsilon} := \int \frac{\delta\mathcal{F}[\Psi_0]}{\delta\Psi}(x) \delta\Psi(x) dx, \quad (\text{A1})$$

¹ We will not discuss the issue further but let it suffice to say that L^2 does satisfy these requirements.

where x denotes a point in the domain of the functions in \mathcal{G} , and the object $(\delta\mathcal{F}[\Psi_0]/\delta\Psi)(x)$ (which is a function of x not necessarily belonging to \mathcal{G}) is called the *functional derivative* of $\mathcal{F}[\Psi]$ in the *point* Ψ_0 .

One common use of this functional derivative is to find stationary points of functionals. A function Ψ_0 is said to be an *stationary point* of $\mathcal{F}[\Psi]$ if:

$$\frac{\delta\mathcal{F}[\Psi_0]}{\delta\Psi}(x) = 0 . \quad (\text{A2})$$

In order to render this definition operative, one must have a method for computing $(\delta\mathcal{F}[\Psi_0]/\delta\Psi)(x)$. Interestingly, it is possible, in many useful cases (and in all the applications of the formalism in this work), to calculate the sought derivative directly from the left-hand side of eq. (A1). The procedure, in such a situation, begins by writing out $\mathcal{F}[\Psi_0 + \varepsilon\delta\Psi]$ and clearly separating the different orders in ε . Secondly, one drops the terms of zero order (by virtue of the subtraction of the quantity $\mathcal{F}[\Psi_0]$) and those of second order or higher (because they vanish when divided by ε and the limit $\varepsilon \rightarrow 0$ is taken). The remaining terms, all of order one, are divided by ε and, finally, $(\delta\mathcal{F}[\Psi_0]/\delta\Psi)(x)$ is identified out of the resulting expression (which must be written in the form of the right-hand side of eq. (A1)). For a practical example of this process, see secs. 6 and 7.

Appendix B: Lagrange multipliers and constrained stationary points

Very often, when looking for the stationary points of a function (or a functional), the search space is not the whole one, in which the derivatives are taken, but a certain subset of it defined by a number of constraints. An elegant and useful method for solving the constrained problem is that of the *Lagrange multipliers*.

Although it can be formally generalized to infinite dimensions (i.e., to functionals, see appendix A), here we will introduce the method in \mathbb{R}^N in order to gain some geometrical insight and intuition.

The general framework may be described as follows: we have a differentiable function $f(\mathbf{x})$ that takes points in \mathbb{R}^N to real numbers and we want to find the stationary points of f restricted to a certain subspace Σ of \mathbb{R}^N , which is defined by K constraints¹:

$$L_i(\mathbf{x}) = 0 \quad i = 1, \dots, K . \quad (\text{B1})$$

The points that are the solution of the constrained problem are those \mathbf{x} belonging to Σ where the first order variation of f would be zero if the derivatives were taken ‘along’ Σ . In other words, the points \mathbf{x} where the gradient ∇f has only components (if any) in directions that ‘leave’ Σ (see below for a rigorous formalization of these intuitive ideas). Thus, when comparing the solutions of the unconstrained problem to the ones of the constrained problem, three distinct situations arise (see fig. B1):

- (i) A point \mathbf{x} is a solution of the unconstrained problem (i.e. it satisfies $\nabla f(\mathbf{x}) = 0$) but it does not belong to Σ . Hence, it is not a solution of the constrained problem. This type of point is depicted as a white-filled circle in fig. B1.

¹ If the constraints are functionally independent, one must also ask that $K < N$. If not, Σ will be either a point (if $K = N$) or empty (if $K > N$).

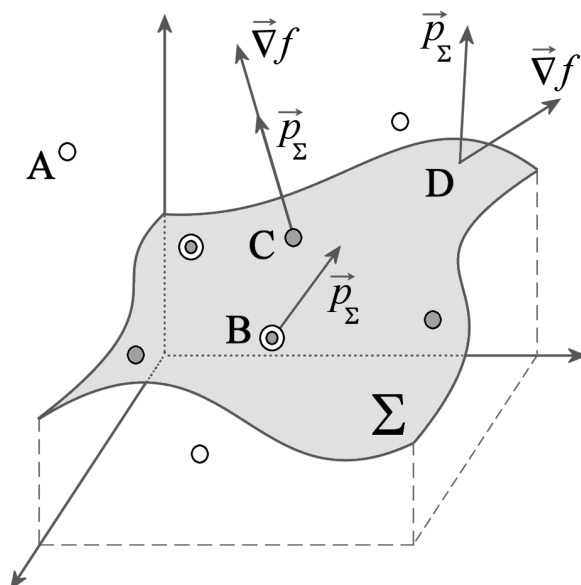


Figure B1. Schematic depiction of a constrained stationary points problem. Σ is the 2-dimensional search space, which is embedded in \mathbb{R}^3 . The *white-filled circles* are solutions of the unconstrained problem only, the *gray-filled circles* are solutions of only the constrained one and the *gray-filled circles inside white-filled circles* are solutions of both. A, B, C and D are examples of different situations discussed in the text.

- (ii) A point \mathbf{x} is a solution of the unconstrained problem (i.e. it satisfies $\nabla f(\mathbf{x}) = 0$) and it belongs to Σ . Hence, it is also a solution of the constrained problem, since, in particular, the components of the gradient in directions that do not leave Σ are zero. This type of point is depicted as a gray-filled circle inside a white-filled circle in fig. B1.
- (iii) A point \mathbf{x} is not a solution of the unconstrained problem (i.e., one has $\nabla f(\mathbf{x}) \neq 0$) but it belongs to Σ and the only non-zero components of the gradient are in directions that leave Σ . Hence, it is a solution of the constrained problem. This type of point is depicted as a gray-filled circle in fig. B1.

From this discussion, it can be seen that, in principle, no conclusions about the number (or existence) of solutions of the constrained problem may be drawn only from the number of solutions of the unconstrained one. This must be investigated for each particular situation.

In fig. B1, an schematic example in \mathbb{R}^3 is depicted. The constrained search space Σ is a 2-dimensional surface and the direction¹ in which one leaves Σ is shown at several points as a perpendicular vector \mathbf{p}_Σ . In such a case, the criterium that ∇f has only components in the direction of leaving Σ may be rephrased by asking ∇f to be parallel to \mathbf{p}_Σ , i.e., by requiring that there exists a number λ such that $\nabla f = -\lambda \mathbf{p}_\Sigma$. The case $\lambda = 0$ is also admitted and the explanation of the minus sign will be given in the following.

In this case, $K = 1$, and one may note that the perpendicular vector \mathbf{p}_Σ is precisely $\mathbf{p}_\Sigma = \nabla L_1$. Let us define \tilde{f} as

$$\tilde{f}(\mathbf{x}) := f(\mathbf{x}) + \lambda L_1(\mathbf{x}) . \tag{B2}$$

¹ Note that, only if $K = 1$, i.e., if the dimension of Σ is $N - 1$, there will be a vector perpendicular to the constrained space. For $K > 1$, the dimensionality of the vector space of the directions in which one 'leaves' Σ will be also larger than 1.

It is clear that, requiring the gradient of \tilde{f} to be zero, one recovers the condition $\nabla f = -\lambda \mathbf{p}_\Sigma$, which is satisfied by the points solution of the constrained stationary points problem. If one also asks that the derivative of \tilde{f} with respect to λ be zero, the constraint $L_1(\mathbf{x}) = 0$ that defines Σ is obtained as well.

This process illustrates the *Lagrange multipliers* method in this particular example. In the general case, described by eq. (B1) and the paragraph above it, it can be proved that the points \mathbf{x} which are stationary subject to the constraints imposed satisfy

$$\nabla \tilde{f}(\mathbf{x}) = 0 \quad \text{and} \quad \frac{\partial \tilde{f}(\mathbf{x})}{\partial \lambda_i} = 0 \quad i = 1, \dots, K, \quad (\text{B3})$$

where

$$\tilde{f}(\mathbf{x}) := f(\mathbf{x}) + \sum_{i=1}^K \lambda_i L_i(\mathbf{x}). \quad (\text{B4})$$

Of course, if one follows this method, the parameters λ_i (which are, in fact, the *Lagrange multipliers*) must also be determined and may be considered as part of the solution.

Also, it is worth remarking here that any two pair of functions, f_1 and f_2 , of \mathbb{R}^N whose restrictions to Σ are equal (i.e., that satisfy $f_1|_\Sigma = f_2|_\Sigma$) obviously represent the same constrained problem and they may be used indistinctly to construct the auxiliary function \tilde{f} . This fact allows us, after having constructed \tilde{f} from a particular f , to use the equations of the constraints to change f by another simpler function which is equal to f when restricted to Σ . This freedom is used to derive the Hartree and Hartree-Fock equations, in secs. 6 and 7, respectively.

The formal generalization of these ideas to functionals (see appendix A) is straightforward if the space \mathbb{R}^N is substituted by a functions space \mathcal{F} , the points \mathbf{x} by functions, the functions f , L_i and \tilde{f} by functionals and the requirement that the gradient of a function be zero by the requirement that the functional derivative of the analogous functional be zero.

Finally, let us stress something that is rarely mentioned in the literature: *There is another (older) method, apart from the Lagrange multipliers one, for solving a constrained optimization problem: simple substitution.* I.e., if we can find a set of $N - K$ independent *adapted coordinates* that parameterize Σ and we can write the score function f in terms of them, we would be automatically satisfying the constraints. Actually, in practical cases, the method chosen is a suitable combination of the two; in such a way that, if substituting the constraints in f is difficult, the necessary Lagrange multipliers are introduced to force them, and vice versa.

As a good example of this, the reader may want to check the derivation of the Hartree equations in sec. 6 (or the Hartree-Fock ones in sec. 7). There, we start by proposing a particular form for the total wavefunction Φ in terms of the one-electron orbitals ϕ_i (see eq. (22)) and we write the functional \mathcal{F} (which is the expected value of the energy) in terms of that special Φ (see eq. (24)). In a second step, we impose the constraints that the one-particle orbitals be normalized ($\langle \phi_i | \phi_i \rangle = 1, i = 1, \dots, N$) and force them by means of N Lagrange multipliers λ_i . Despite the different treatments, both conditions are constraints standing on the same footing. The only difference is not conceptual, but operative: for the first condition, it would be difficult to write it as a constraint; while, for the second one,

it would be difficult to define adapted coordinates in the subspace of normalized orbitals. So, in both cases, the easiest way for dealing with them is chosen.

References

- [1] Skolnick, J. *Proc. Natl. Acad. Sci. USA*, **102**, 2265–2266 (2005).
- [2] Snow, C. D., Sorin, E. J., Rhee, Y. M., and Pande, V. S. *Annu. Rev. Biophys. Biomol. Struct.*, **34**, 43–69 (2005).
- [3] Schueler-Furman, O., Wang, C., Bradley, P., Misura, K., and Baker, D. *Science*, **310**, 638–642 (2005).
- [4] Ginalski, K., Grishin, N. V., Godzik, A., and Rychlewski, L. *Nucleic Acids Research*, **33**, 1874–1891 (2005).
- [5] Bonneau, R. and Baker, D. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 173–189 (2001).
- [6] Hao, M.-H. and Scheraga, H. A. *Curr. Opin. Struct. Biol.*, **9**, 184–188 (1999).
- [7] Echenique, P. *Contemp. Phys.*, **48**, 81–108 (2007).
- [8] Morozov, A. V., Tsemekhman, K., and Baker, D. *J. Phys. Chem. B*, **110**, 4503–4505 (2006).
- [9] Jensen, F. *Ann. Rep. Comp. Chem.*, **1**, 1–17 (2005).
- [10] MacKerell Jr., A. R., Feig, M., and Brooks III, C. L. *J. Comp. Chem.*, **25**, 1400–1415 (2004).
- [11] Morozov, A. V., Kortemme, T., Tsemekhman, K., and Baker, D. *Proc. Natl. Acad. Sci. USA*, **101**, 6946–6951 (2004).
- [12] Bordner, A. J., Cavasotto, C. N., and Abagyan, R. A. *J. Phys. Chem. B*, **107**, 9601–9609 (2003).
- [13] Friesner, R. A. and Beachy, M. D. *Curr. Opin. Struct. Biol.*, **8**, 257–262 (1998).
- [14] Beachy, M., Chasman, D., Murphy, R., Halgren, T., and Friesner, R. *J. Am. Chem. Soc.*, **119**, 5908–5920 (1997).
- [15] Barden, C. J. and Schaffer III, H. F. *Pure Appl. Chem.*, **72**, 1405–1423 (2000).
- [16] Simons, J. *J. Chem. Phys.*, **95**, 1017–1029 (1991).
- [17] Levine, I. N. *Quantum Chemistry*, 5th edn, Prentice Hall, Upper Saddle River (1999).
- [18] Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester (1998).
- [19] Szabo, A. and Ostlund, N. S. *Modern Quantum Chemistry: Introduced to Advanced Electronic Structure Theory*, Dover Publications, New York (1996).
- [20] Taylor, B. N. Guide for the Use of the International System of Units (SI), NIST Special Publication 811, National Institute of Standards and Technology (1995).
- [21] Hartree, D. R. *Proc. Camb. Philos. Soc.*, **24**, 89 (1927).
- [22] Shull, H. and Hall, G. G. *Nature*, **184**, 1559 (1959).
- [23] Born, M. and Huang, K. *Dynamical Theory of Crystal Lattices*, appendices VII and VIII, Oxford University Press, London (1954).
- [24] Born, M. and Oppenheimer, J. R. *Ann. Phys. Leipzig*, **84**, 457–484 (1927).
- [25] Marder, M. P. *Condensed Matter Physics*, Wiley-Interscience, New York (2000).
- [26] Shida, T. *The Chemical Bond: A Fundamental Quantum-Mechanical Picture*, Springer Series in Chemical Physics, Springer-Verlag, Berlin (2006).
- [27] Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd edn, John Wiley & Sons, Chichester (2002).
- [28] Parr, R. G. and Yang, W. *Density-Functional Theory of Atoms and Molecules*, vol. 16 of *International series of monographs on chemistry*, Oxford University Press, New York (1989).
- [29] Sutcliffe, B. T. and Woolley, R. G. *Phys. Chem. Chem. Phys.*, **7**, 3664–3676 (2005).
- [30] Sutcliffe, B. T. *Adv. Quantum Chem.*, **28**, 65–80 (1997).
- [31] Sutcliffe, B. T. *J. Chem. Soc. Faraday Trans.*, **89**, 2321–2335 (1993).
- [32] Hunter, G. *Intl. J. Quant. Chem.*, **9**, 237–242 (1975).
- [33] Yserentant, H. On the electronic Schrödinger equation. Technical report, Universität Tübingen (2003). Available at: www.math.tu-berlin.de/~yserenta/.
- [34] Simon, B. *J. Math. Phys.*, **41**, 3523–3555 (2000).
- [35] Hunziger, W. and Sigal, I. M. *J. Math. Phys.*, **41**, 3448–3510 (2000).
- [36] Ruskai, M. B. and Solovej, J. P. *Lect. Notes Phys.*, **403**, 153–174 (1992).
- [37] Hunziker, W. *Helv. Phys. Acta*, **39**, 451–462 (1966).
- [38] Van Winter, C. *Mat. Fys. Skr. Dan. Vid. Selsk.*, **2**, 1–60 (1964).
- [39] Friesocke, G. *Arch. Rational Mech. Anal.*, **169**, 35–71 (2003).
- [40] Zhislin, G. M. *Trudy Moskovskogo matematicheskogo obschestva*, **9**, 81–120 (1960). (in Russian).
- [41] Hratchian, H. P. and Schlegel, H. B. in *Theory and Applications of Computational Chemistry: The First Forty Years*, C. Dykstra, G. Frenking, K. Kim and G. Scuseria (Eds), chapt. 10, Elsevier (2005).
- [42] Brooks, B. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. *J. Comp. Chem.*, **4**, 187–217 (1983).
- [43] MacKerell Jr., A. D., Brooks, B., Brooks III, C. L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. in *The Encyclopedia of Computational Chemistry*, P. v. R. Schleyer, P. R. Schreiner, N. L. Allinger, T. Clark, J. Gasteiger, P. Kollman and H. F. Schaefer III (Eds), pp. 217–277, John Wiley & Sons, Chichester (1998).
- [44] Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. R., Cheatham III, T. E., DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. *Comp. Phys. Commun.*, **91**, 1–41 (1995).
- [45] Ponder, J. W. and Case, D. A. *Adv. Prot. Chem.*, **66**, 27–85 (2003).
- [46] Cheatham III, T. E. and Young, M. A. *Biopolymers*, **56**, 232–256 (2001).
- [47] Jorgensen, W. L. and Tirado-Rives, J. *J. Am. Chem. Soc.*, **110**, 1657–1666 (1988).
- [48] Cohen-Tannoudji, C., Diu, B. and Laloë, F. *Quantum Mechanics*, Hermann and John Wiley & Sons, Paris (1977).
- [49] Dirac, P. A. M. *Proc. Roy. Soc. London*, **123**, 714 (1929).
- [50] Seeger, R. and Pople, J. A. *J. Chem. Phys.*, **66**, 3045–3050 (1977).
- [51] Marinari, E. and Parisi, G. *Europhys. Lett.*, **19**, 451–458 (1992).

- [52] Cerny, V. *J. Optimiz. Theory App.*, **45**, 41–51 (1985).
- [53] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. *Science*, **220**, 671–680 (1983).
- [54] Slater, J. C. *Phys. Rev.*, **35**, 210–211 (1930).
- [55] Cancès, E., DeFranceschi, M., Kutzelnigg, W., Le Bris, C., and Maday, Y. in *Handbook of numerical analysis. Volume X: Special volume: Computational chemistry*, P. Ciarlet and C. Le Bris (Eds), pp. 3–270, Elsevier (2003).
- [56] Lions, P. L. *Commun. Math. Phys.*, **109**, 33–97 (1987).
- [57] Lieb, E. H. and Simon, B. *J. Chem. Phys.*, **61**, 735–736 (1974).
- [58] Lieb, E. H. and Simon, B. *Commun. Math. Phys.*, **53**, 185–194 (1977).
- [59] Fock, V. *Z. Phys.*, **61**, 126 (1930).
- [60] Koopmans, T. *Physica*, **1**, 104 (1934).
- [61] Schlegel, H. B. and McDouall, J. J. W. in *Computational Advances in Organic Chemistry: Molecular Structure and Reactivity*, C. Ögretir and I. G. Csizmadia (Eds), pp. 167–185, Kluwer Academic, The Netherlands (1991).
- [62] Onuchic, J. N. and Wolynes, P. G. *Curr. Opin. Struct. Biol.*, **14**, 70–75 (2004).
- [63] Plotkin, S. S. and Onuchic, J. *Quart. Rev. Biophys.*, **35**, 111–167 (2002).
- [64] Dill, K. A. *Prot. Sci.*, **8**, 1166–1180 (1999).
- [65] Dobson, C. M., Šali, A., and Karplus, M. *Angew. Chem. Int. Ed.*, **37**, 868–893 (1998).
- [66] Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. *Proteins*, **21**, 167–195 (1995).
- [67] Bryngelson, J. D. and Wolynes, P. G. *Proc. Natl. Acad. Sci. USA*, **84**, 7524–7528 (1987).
- [68] Bally, T. and Borden, W. T. *Rev. Comp. Chem.*, **13**, 1–97 (1999).
- [69] Pople, J. A. and Nesbet, R. K. *J. Chem. Phys.*, **22**, 571–572 (1954).
- [70] Vargas, R., Garza, J., Hay, B. P., and Dixon, D. A. *J. Phys. Chem. A*, **106**, 3213–3218 (2002).
- [71] Láng, A., Csizmadia, I. G., and Perczel, A. *PROTEINS: Struct. Funct. Bioinf.*, **58**, 571–588 (2005).
- [72] Perczel, A., Farkas, Ö., Jákli, I., Topol, I. A., and Csizmadia, I. G. *J. Comp. Chem.*, **24**, 1026–1042 (2003).
- [73] Yu, C.-H., Norman, M. A., Schäfer, L., Ramek, M., Peeters, A., and van Alsenoy, C. *J. Mol. Struct.*, **567–568**, 361–374 (2001).
- [74] Elstner, M., Jalkanen, K. J., Knapp-Mohammady, M., and Suhai, S. *Chem. Phys.*, **263**, 203–219 (2001).
- [75] Baldoni, H. A., Zamarbide, G., Enriz, R. D., Jauregui, E. A., Farkas, Ö., Perczel, A., Salpietro, S. J., and Csizmadia, I. G. *J. Mol. Struct.*, **500**, 97–111 (2000).
- [76] Rodríguez, A. M., Baldoni, H. A., Suvire, F., Nieto Vázquez, R., Zamarbide, G., Enriz, R. D., Farkas, Ö., Perczel, A., McAllister, M. A., Torday, L. L., Papp, J. G., and Csizmadia, I. G. *J. Mol. Struct.*, **455**, 275–301 (1998).
- [77] Császár, A. G. and Perczel, A. *Prog. Biophys. Mol. Biol.*, **71**, 243–309 (1999).
- [78] Frey, R. F., Coffin, J., Newton, S. Q., Ramek, M., Cheng, V. K. W., Momany, F. A., and Schäfer, L. *J. Am. Chem. Soc.*, **114**, 5369–5377 (1992).
- [79] Roothaan, C. C. *J. Rev. Mod. Phys.*, **23**, 69–89 (1951).
- [80] Hall, G. G. *Proc. Roy. Soc. London Ser. A*, **205**, 541–552 (1951).
- [81] Cook, D. B. *Handbook of Computational Quantum Chemistry*, Dover Publications, Mineola, New York (2005).
- [82] Carbo, R. and Riera, J. M. *A General SCF Theory*, Lecture Notes in Chemistry, Springer-Verlag, New York (1977).
- [83] Hurley, A. C. *Introduction to the Electron Theory of Small Molecules*, Academic Press, New York (1976).
- [84] McWeeny, R. *Methods of Molecular Quantum Mechanics*, Academic Press (1992).
- [85] Roothaan, C. C. *J. Rev. Mod. Phys.*, **32**, 179–185 (1960).
- [86] Binkley, J. S., Pople, J. A. and Dobosh, P. A. *Mol. Phys.*, **28**, 1423–1429 (1974).
- [87] Peterson M. and Pourier R., *MONSTERGAUSS-92*, Department of Chemistry, University of Toronto and Memorial University of Newfoundland, St. John's, Newfoundland, Canada.
- [88] Plakhotin, B. N., Gorelik, E. V. and Breslavskaya, N. N. *J. Chem. Phys.*, **125**, 204110 (2006).
- [89] Hirao, K. *J. Chem. Phys.*, **60**, 3125–3133 (1974).
- [90] Guest, M. F. and Saunders, V. R. *Mol. Phys.*, **28**, 819–828 (1974).
- [91] Kobus, J. *Adv. Quantum Chem.*, **28**, 1–14 (1997).
- [92] Jensen, F. *Theo. Chem. Acc.*, **113**, 267–273 (2005).
- [93] Head-Gordon, M. and Pople, J. A. *J. Phys. Chem.*, **92**, 3063–3069 (1988).
- [94] Pople, J. A. *Rev. Mod. Phys.*, **71**, 1267–1274 (1999).
- [95] Brailsford, D. F. and Hall, G. G. *Intl. J. Quant. Chem.*, **5**, 657–668 (1971).
- [96] García de la Vega, J. M. and Miguel, B. in *Introduction to Advanced Topics of Computational Chemistry*, L. A. Montero, L. A. Díaz and R. Bader (Eds), chapt. 3, pp. 41–80, Editorial de la Universidad de la Habana (2003).
- [97] Helgaker, T. and Taylor, P. R. in *Modern Electronic Structure Theory. Part II*, D. R. Yarkony (Ed), pp. 725–856, World Scientific, Singapore (1995).
- [98] Abramowitz, M. and Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th edn, Dover, New York (1964).
- [99] Slater, J. C. *Phys. Rev.*, **36**, 57–54 (1930).
- [100] Zener, C. *Phys. Rev.*, **36**, 51–56 (1930).
- [101] Mathar, R. *J. Intl. J. Quant. Chem.*, **90**, 227–243 (2002).
- [102] Kato, T. *Commun. Pure Appl. Math.*, **10**, 151–177 (1957).
- [103] Boys, S. F. *Proc. Roy. Soc. London Ser. A*, **200**, 541–554 (1950).
- [104] Schlegel, H. B. and Frisch, M. J. *Intl. J. Quant. Chem.*, **54**, 83–87 (1995).
- [105] Echenique, P., Calvo, I., and Alonso, J. L. *J. Comp. Chem.*, **27**, 1748–1755 (2006).
- [106] Echenique, P. and Alonso, J. L. *J. Comp. Chem.*, **27**, 1076–1087 (2006).
- [107] Hehre, W. J., Stewart, R. F., and Pople, J. A. *J. Chem. Phys.*, **51**, 2657–2664 (1969).

- [108] Ditchfield, R., Hehre, W. J., and Pople, J. A. *J. Chem. Phys.*, **54**, 724–728 (1971).
- [109] Hehre, W. J., Ditchfield, R., and Pople, J. A. *J. Chem. Phys.*, **56**, 2257–2261 (1972).
- [110] Hariharan, P. C. and Pople, J. A. *Theor. Chim. Acta*, **28**, 213–222 (1973).
- [111] Frisch, M. J., Pople, J. A., and Binkley, J. S. *J. Chem. Phys.*, **80**, 3265–3269 (1984).
- [112] Krishnan, R., Binkley, J. S., Seeger, R., and Pople, J. A. *J. Chem. Phys.*, **72**, 650–654 (1980).
- [113] Binkley, J. S., Pople, J. A., and Hehre, W. J. *J. Am. Chem. Soc.*, **102**, 939–947 (1980).
- [114] Spitznagel, G. W., Clark, T., Chandrasekhar, J., and Schleyer, P. v. R. *J. Comp. Chem.*, **3**, 363–371 (1982).
- [115] Clark, T., Chandrasekhar, J., Spitznagel, G. W., and Schleyer, P. v. R. *J. Comp. Chem.*, **4**, 294–301 (1983).
- [116] Dunning Jr., T. H. *J. Chem. Phys.*, **90**, 1007–1023 (1989).
- [117] Woodgate, G. K. *Elementary Atomic Structure*, 2nd edn, Oxford University Press, USA (1983).
- [118] Shao, Y., Molnar L. F., Jung, Y., Kussmann, J., Ochsenfeld, C., Brown, S. T., Gilbert, A. T. B., Slipchenko, L. V., Levchenko, S. V., O’Neill, D. P., DiStasio, R. A., Lochan, R. C., Wang, T., Beran, G. J. O., Besley, N. A., Herbert, J. M., Lin, C. Y., Van Voorhis, T., Chien, S. H., Sodt, A., Steele, R. P., Rassolov, V. A., Maslen, P. E., Korambath, P. P., Adamson, R. D., Austin, B., Baker, J., Byrd, E. F. C., Dachsel, H., Doerksen, R. J., Dreuw, A., Dunietz, B. D., Dutoi, A. D., Furlani, T. R., Gwaltney, S. R., Heyden, A., Hirata, S., Hsu, C. P., Kedziora, G., Khallulin, R. Z., Klunzinger, P., Lee, A. M., Lee, M. S., Liang, W., Lotan, I., Nair, N., Peters, B., Proynov, E. I., Pieniazek, P. A., Rhee, Y. M., Ritchie, J., Rosta, E., Sherrill, C. D., Simmonett, A. C., Subotnik, J. E., Woodcock, H. L., Zhang, W., Bell, A. T., Chakraborty, A. K., Chipman, D. M., Keil, F. J., Warshel, A., Hehre, W. J., Schaefer, H. F., Kong, J., Krylov, A. I., Gill, P. M. W., Head-Gordon, M. *Phys. Chem. Chem. Phys.*, **8**, 3172–3191 (2006).
- [119] Challacombe, M., Schwegler, R. and Almlöf, J. in *Computational Chemistry: Review of Current Trends*, J. Leczszynski (Ed), pp. 53–107, World Scientific (1996).
- [120] Ochsenfeld, C., Kussmann, J. and Lambrecht, D. S. in *Reviews in Computational Chemistry*, K. B. Lipkowitz, T. R. Cundari (Eds), Vol. 23, pp. 1–82, John Wiley & Sons (2007).
- [121] Gill, P. M. W., Johnson, B. G. and Pople, G. A. *Intl. J. Quant. Chem.*, **40**, 745–652 (1991).
- [122] Gill, P. M. W. *Adv. Quant. Chem.*, **25**, 141–173 (1994).
- [123] Pople, J. A. and Hehre, W. J. *J. Comp. Phys.*, **27**, 161–168 (1978).
- [124] McMurchie, L. E. and Davidson, E. R. *J. Comp. Phys.*, **26**, 218–231 (1978).
- [125] Schlegel, H. B. *J. Chem. Phys.*, **90**, 5630–5634 (1989).
- [126] Obara, S. and Saika, A. *J. Chem. Phys.*, **89**, 1540–1559 (1988).
- [127] Head-Gordon, M. and Pople, J. A. *J. Chem. Phys.*, **89**, 5777–5786 (1988).
- [128] Gill, P. M. W., Head-Gordon, M. and Pople, G. A. *Intl. J. Quant. Chem., Symp.*, **23**, 269 (1989).
- [129] Gill, P. M. W., Head-Gordon, M. and Pople, G. A. *J. Chem. Phys.*, **94**, 5564 (1990).
- [130] Gill, P. M. W. and Pople, G. A. *Intl. J. Quant. Chem., Symp.*, **40**, 753 (1991).
- [131] Adams, T. R., Adamson, R. S. and Gill, P. M. W. *J. Chem. Phys.*, **107**, 124–131 (1997).
- [132] Goedecker, S. *Rev. Mod. Phys.*, **71**, 1085–1123 (1999).
- [133] Kohn, W. *Phys. Rev. Lett.*, **76**, 3168–3171 (1996).
- [134] Dyczmons, V. *Theoret. Chim. Acta*, **28**, 307–310 (1973).
- [135] Ahlrichs, R. *Theoret. Chim. Acta*, **33**, 157–167 (1974).
- [136] Almlöf, J., Faegri Jr., K. and Korsell, K. *J. Comp. Chem.*, **3**, 385–399 (1982).
- [137] Häser, M. and Ahlrichs, R. *J. Comp. Chem.*, **10**, 104–111 (1989).
- [138] Lambrecht, D. S. and Ochsenfeld, C. *J. Chem. Phys.*, **123**, 184101 (2005).
- [139] Gill, P. M. W., Johnson, B. G. and Pople, G. A. *Chem. Phys. Lett.*, **217**, 65–68 (1994).
- [140] Jiancheng, X. and Shouping, J. *Intl. J. Quant. Chem.*, **39**, 123–130 (1991).
- [141] Strout, D. L. and Scuseria, G. E. *J. Chem. Phys.*, **102**, 8448–8452 (1995).
- [142] Panas, I., Almlöf, J. and Feyerisen, M. W. *Intl. J. Quant. Chem.*, **40**, 797–807 (1991).
- [143] Termath, V. and Handy, N. C. *Chem. Phys. Lett.*, **230**, 17–24 (1994).
- [144] Burant, J. C., Scuseria, G. E. and Frisch, M. J. *J. Chem. Phys.*, **105**, 8969–8972 (1996).
- [145] Ochsenfeld, C., White, C. A. and Head-Gordon, M. *J. Chem. Phys.*, **109**, 1663–1669 (1998).
- [146] Schwegler, E., Challacombe, M. and Head-Gordon, M. *J. Chem. Phys.*, **106**, 9708–9717 (1997).
- [147] Ochsenfeld, C. *Chem. Phys. Lett.*, **327**, 216 (2000).
- [148] White, C. A., Johnson, B. G., Gill, P. M. W. and Head-Gordon, M. *Chem. Phys. Lett.*, **230**, 8–16 (1994).
- [149] Greengard, L. and Rokhlin, C. *J. Comp. Phys.*, **73**, 325–348 (1987).
- [150] Barnes, J. and Hut P. *Nature*, **104**, 446–449 (1986).
- [151] Kutteh, R., Aprà, E. and Nichols, J. *Chem. Phys. Lett.*, **238**, 173–179 (1995).
- [152] Challacombe, M., Schwegler, E. and Almlöf, J. *J. Chem. Phys.*, **104**, 4685 (1995).
- [153] Strain, M. C., Scuseria, G. E. and Frisch, M. J. *Science*, **271**, 51–53 (1996).
- [154] Petersen, H. G., Soelvason, D. and Perram, J. W. *J. Chem. Phys.*, **101**, 8870–8876 (1994).
- [155] White, C. A. and Head-Gordon, M. *J. Chem. Phys.*, **104**, 2620–2629 (1996).
- [156] Izmaylov, A. F., Scuseria, G. E. and Frisch, M. J. *J. Chem. Phys.*, **125**, 104103 (2006).
- [157] Füsti-Molnár, L. and Kong, J. *J. Chem. Phys.*, **122**, 074108 (2005).
- [158] Füsti-Molnár, L. and Pulay, P. *J. Chem. Phys.*, **117**, 7827 (2002).
- [159] Scuseria, G. E. *J. Phys. Chem. A*, **103**, 4782–4790 (1999).
- [160] Bowler, D. R., Miyazaki, T. and Gillan, M. J. *J. Phys.: Condens. Matter*, **14**, 2781–2798 (2002).
- [161] Ortega, J. in *Mathematical Methods for Digital Computers*, A. Ralston and H. S. Wilf (Eds), Vol. 2, p. 94, Wiley, New York (1957).
- [162] Maslen, P. E., Ochsenfeld, C., White, C. A., Lee, M. S. and Head-Gordon, M. *J. Phys. Chem. A*, **102**, 2215–2222 (1998).
- [163] Shao, Y., Saravanan, C., Head-Gordon, M. and White, C. A. *J. Chem. Phys.*, **118**, 6144–6151 (2003).
- [164] Mauri, F., Galli, G. and Car, R. *Phys. Rev. B*, **47**, 9973–9976 (1993).
- [165] Stewart, J. J. P. *Intl. J. Quant. Chem.*, **58**, 133–146 (1996).

- [166] Stewart, J. J. P., Császár, P. and Pulay, P. *J. Comp. Chem.*, **3**, 227–228 (1982).
- [167] Ochsenfeld, C. and Head-Gordon, M. *Chem. Phys. Lett.*, **270**, 399–405 (1997).
- [168] Salek, P., Høst, S., Thøgersen, L., Jørgensen, P., Manninen, P., Olsen, J., Jansík, B., Reine, S., Pawłowski, F., Tellgren, E., Helgaker, T. and Coriani, S. *J. Chem. Phys.*, **126**, 114110 (2007).
- [169] Millam, J. M. and Scuseria, G. E. *J. Chem. Phys.*, **106**, 5569–5577 (1997).
- [170] Ordejón, P., Drabold, D. A., Grumbach, M. P. and Martin, R. M. *Phys. Rev. B*, **48**, 14646–14649 (1993).
- [171] Liang, W., Saravanan, C., Shao, Y., Baer, R., Bell, A. T. and Head-Gordon, M. *J. Chem. Phys.*, **119**, 4117–4125 (2003).
- [172] Cullum, J. K. and Willoughby, R. A. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations. Vol 2*, Birkhäuser, Boston (1985).
- [173] Ericsson, T. and Ruhe, A. *Math. Comput.*, **35**, 1251 (1980).
- [174] Helgaker, T., Larsen, H., Olsen, J. and Jørgensen, P. *Chem. Phys. Lett.*, **327**, 397–403 (2000).
- [175] Thøgersen, L., Olsen, J., Yeager, D., Jørgensen, P., Salek, P. and Helgaker, T. *J. Chem. Phys.*, **121**, 16–27 (2004).
- [176] Marx, D. and Hutter, J. in *Modern Methods and Algorithms of Quantum Chemistry*, J. Grotendorst (Ed), Vol. 3, pp. 329–477, John von Neumann Institute for Computing, Jülich (2000).
- [177] Pulay, P. *Adv. Chem. Phys.*, **69**, 241 (1987).
- [178] Schlegel, H. B. *Adv. Chem. Phys.*, **67**, 249 (1987).
- [179] Shao, Y., White, C. A. and Head-Gordon, M. *J. Chem. Phys.*, **114**, 6572–6577 (2001).
- [180] Burant, J. C., Strain, M. C., Scuseria, G. E. and Frisch, M. J. *Chem. Phys. Lett.*, **248**, 42–49 (1996).