Eric Cancès

Claude Le Bris

On the convergence of SCF algorithms for the Hartree-Fock equations

ON THE CONVERGENCE OF SCF ALGORITHMS
FOR THE HARTREE-FOCK EQUATIONS

ERIC CANCÈS[1] AND CLAUDE LE BRIS[1]

Abstract. The present work is a mathematical analysis of two algorithms, namely the Roothaan and the level-shifting algorithms, commonly used in practice to solve the Hartree-Fock equations. The level-shifting algorithm is proved to be well-posed and to converge provided the shift parameter is large enough. On the contrary, cases when the Roothaan algorithm is not well defined or fails in converging are exhibited. These mathematical results are confronted to numerical experiments performed by chemists.

1. INTRODUCTION

The Hartree-Fock model [7, 15] is a standard tool for computing an approximation of the electronic ground state of a molecular system within the Born-Oppenheimer (or *clamped nuclei*) setting. The solution to the Hartree-Fock problem can be obtained either by directly minimizing the Hartree-Fock energy functional [5, 8, 14, 23] or by solving the associated Euler-Lagrange equations, the so-called Hartree-Fock equations [6, 17, 20, 21]. In any case, an iterative procedure has to be resorted to. Such an iterative procedure is often referred to as a self-consistent field (SCF) algorithm.

SCF algorithms for solving the Hartree-Fock equations are in general much more efficient than direct energy minimization techniques in term of computational effort. However, these algorithms do not *a priori* ensure the decrease of the energy and they may lead to convergence problems [22]. For instance, the famous Roothaan algorithm (see [20] and below) is known to sometimes give rise to stable oscillations between two states, none of them being a solution to the Hartree-Fock problem. This situation may occur even for simple chemical systems (see Ex. 9 below).

There are many articles dealing with SCF convergence issues *in the chemical literature*. The behavior of the Roothaan algorithm is notably investigated in [3, 9] and in [24, 25]. In [3, 9] convergence difficulties are demonstrated for elementary two-dimensional models; in [24, 25], a stability condition of the Roothaan algorithm in the neighbourhood of a minimum of the Hartree-Fock energy is given for the so-called closed-shell systems. More sophisticated SCF algorithms for solving the Hartree-Fock equations have also been proposed to improve the convergence using various techniques like for instance relaxation (or *damping* [6, 26]) or level-shifting [21]. Relaxation cures some convergence problems but many other remain. Numerical tests confirm that the level-shifting algorithm converges towards a solution to the Hartree-Fock equations for large enough shift parameters,

although there is no guarantee that the so-obtained stationary point of the Hartree-Fock energy functional is actually a minimum (even local); a perturbation argument is provided in [21] to prove this convergence in the neighborhood of a stationary point. Let us also mention the Direct Inversion in the Iteration Space (DIIS) algorithm [17], which is widely used in quantum chemistry calculations at the present time. Numerical tests show that this algorithm is very efficient in most cases, but that it sometimes fails. To our knowledge, the DIIS algorithm has not yet been studied from a theoretical point of view. Let us point out that all the convergence studies reported above concern the Hartree-Fock problem within the *local combination of atomic orbitals* (LCAO) approximation, *i.e.* within a Galerkin approximation rendering it finite dimensional. It follows that compactness properties, almost trivial in a finite dimensional setting, allow each author to oversimplify the proofs of convergence of the algorithms. In addition, it is to be remarked that all the studies mentioned so far rather provide *recipes* to cure the possible difficulties, and describe extensive programs of numerical experiment. The algorithms are rarely studied from a rigorous standpoint and when they are, it seems to us that there is room for improvements in the results. This is the reason why we have tried to contribute to such a study in the present work. Actually, we are only aware of *one* rigorous mathematical study dealing with algorithms for solving equations of quantum chemistry. This is the work by Auchmuty and Wenyao Jia [1] proving the convergence of a non-standard algorithm (unfortunately not used in practice as far as we know) for solving the *Hartree* equations. We strongly encourage the reader to point out to us any other reference of this kind.

Our purpose in this article will be to prove convergence results (a) on some of the algorithms that are actually in use in the community of computational chemistry, (b) that deal with the infinite dimensional case, thereby ensuring properties that are independent from the specific parameters fixed in the implementation, (c) that improve the existing results of the chemical literature, when applied to the finite dimensional setting. As will be seen, the results we shall obtain are likely to be improved in many respects. We hope our work will stimulate further research.

In the present paper, we focus on two algorithms that are *indeed* used in practice, namely the Roothaan and the level-shifting algorithms, considered as procedures for solving the original Hartree-Fock equations in infinite dimension. However, except Propositions 5 and 6 which are specific to the infinite dimension, all the results below, and notably Theorems 7 and 11 can be applied to the Hartree-Fock equations within the LCAO approximation (see also Rem. 8). In Section 3, we propose a new formulation of the Roothaan algorithm which is useful for the mathematical study (it provides us with a Lyapunov functional) and which in addition makes clear the risk of stable oscillations between two states. Examples of pathological situations are exhibited. In Section 4, we establish the convergence of the level-shifting algorithm, provided the level-shift is chosen high enough. Contrary to the local result proved in [21] by a perturbation argument, our result is global: we need not assume that the initial guess is close to a stationary point of the Hartree-Fock energy functional. As it is difficult to check that such an assumption is fulfilled, we therefore believe that our result improves that of [21] from the standpoint of applications. Although they are somewhat related to convergence issues, we leave apart in the present study questions of stability and symmetry breaking [4,16] as well as the problem of the choice of the initial guess. We hope to address these interesting questions in a future work.

Now, before turning to convergence studies, we briefly recall in Section 2 the main features of the Hartree-Fock model and of the associated SCF algorithms.

2. SCF algorithms for the Hartree-Fock equations

In the sequel, only spinless models are considered but most of the mathematical results detailed below can be extended *mutatis mutandis* to other models taking spin into account like the General Hartree-Fock (GHF), the Unrestricted Hartree-Fock (UHF), or the Restricted Hartree-Fock (RHF) models [22]. The spinless setting is thus chosen only for the sake of clarity. The rare situations when the spin plays an active role will be pointed out.

In Section 2.1, the Hartree-Fock model is presented, with a special attention to the density operator formulation (see the reason why in Rem. 4 below). Some standard SCF algorithms for solving the Hartree-Fock equations are described in Section 2.2.

2.1. Basics on the Hartree-Fock model

Computing the electronic ground state of a molecular system containing N electrons within the Born-Oppenheimer approximation consists in searching the fundamental eigenstate of the N-body Hamiltonian

$$H := -\sum_{i=1}^{N} \Delta_{x_i} + \sum_{i=1}^{N} V(x_i) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|},$$

in the space of admissible wave functions

$$\mathcal{H} := \bigwedge_{i=1}^{N} L^2(\mathbb{R}^3, \mathbb{C}).$$

The external potential V corresponds to the Coulomb potential generated by the nuclei; it is given by

$$V(x) := -\sum_{k=1}^{M} \frac{z_k}{|x - \bar{x}_k|}$$

for a molecular system consisting of M nuclei, the kth nucleus carrying a charge z_k and being located at $\bar{x}_k$. The variational formulation of the problem to solve reads

$$\inf \left\{ \langle \psi, H\psi \rangle, \quad \psi \in \mathcal{H}, \quad \int_{\mathbb{R}^{3N}} |\psi|^2 = 1 \right\}. \tag{1}$$

This problem is linear but, except for 1 or 2-electron molecular systems, it cannot be directly tackled with the standard tools of numerical analysis like finite-difference of finite-element methods because of the size of the space $\mathcal{H}$. The Hartree-Fock model is a variational approximation of the original problem (1). It consists in restraining the minimization set to the set of the so-called Slater determinants, namely to the set of the functions ψ which can be written as

$$\psi(x_1, \cdots, x_N) = \frac{1}{\sqrt{N!}} \det(\phi_i(x_j)),$$

where $\Phi := \{\phi_i\}_{1 \leq i \leq N}$ belongs to the set of molecular orbital configurations

$$\mathcal{W} := \left\{ \Phi = \{\phi_i\}_{1 \leq i \leq N}, \quad \phi_i \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} \phi_i \phi_j^* = \delta_{i,j}, \quad 1 \leq i, j \leq N \right\}.$$

Here and henceforth, the upper-script * denotes the complex conjugate. One obtains in this way the well-known Hartree-Fock minimization problem under its standard form

$$\inf \left\{ E^{HF}(\Phi), \quad \Phi \in \mathcal{W} \right\}, \tag{2}$$

where the Hartree-Fock energy functional reads

$$E^{HF}(\Phi) := \sum_{i=1}^{N} \int_{\mathbb{R}^3} |\nabla \phi_i|^2 + \sum_{i=1}^{N} \int_{\mathbb{R}^3} V|\phi_i|^2 + \frac{1}{2} \left(\sum_{i,j=1}^{N} D(|\phi_i|^2, |\phi_j|^2) - D(\phi_i \phi_j^*, \phi_i^* \phi_j) \right)$$

with $D(\rho, \rho') := \int_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho(x) \rho'(y)}{|x-y|} \, dx \, dy$. The mathematical properties of this minimization problem have been studied by Lieb and Simon [12] and by Lions [13]. The existence of a Hartree-Fock electronic ground state is

proved for positive ions ($Z := \sum_{k=1}^{M} z_k > N$) and neutral systems ($Z = N$). We are not aware of any general existence result for negative ions (the available existence proofs only work for $N < Z + 1$). On the other hand, there is a non-existence results for negative ions such that $N \geq 2Z + M$ [11] (this inequality holds for instance for the ion H^{2-}). As far as we know, uniqueness (of the density $\rho := \sum_{i=1}^{N} |\phi_i|^2$ at least) is an open problem probably of outstanding difficulty. In addition, the local structure of the Hartree-Fock energy functional in the neighborhoods of the stationary points is not known. These are serious limitations for convergence studies.

Let us remark that, if $U \in U(N)$ is a unitary $N \times N$ matrix, then for all $\Phi \in \mathcal{W}$, $U\Phi \in \mathcal{W}$, and both Φ and $U\Phi$ generate the same Slater determinant. This implies in particular that a minimizer of problem (2) is defined up to a unitary matrix. The Hartree-Fock problem (2) can be rewritten in a more intrinsic way which enables one to get rid of this gauge invariance, by means of density operators. Let us recall that the density operator $\mathcal{D}$ associated with a N-particle wave function ψ is the operator on $L^2(\mathbb{R}^3)$ defined by the Schwartz kernel

$$\tau_{\mathcal{D}}(x, y) = N \int_{\mathbb{R}^{3(N-1)}} \psi(x, x_2, \cdots, x_N) \, \psi(y, x_2, \cdots, x_N)^* \, \mathrm{d}x_2 \cdots \mathrm{d}x_N,$$

for any $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3$. The Schwarz kernel $\tau_{\mathcal{D}}$ is the so-called density matrix (also called 1-particle density matrix or reduced density matrix in the literature) associated with $\mathcal{D}$. The corresponding electronic density is denoted by $\rho_{\mathcal{D}}(x) := \tau_{\mathcal{D}}(x, x)$. The operator $\mathcal{D}$ is self-adjoint and trace-class (for an introduction to the theory of trace-class operators see for instance pp. 207–210 of [18]). It satisfies the following properties:

$$\mathrm{Tr}\,(\mathcal{D}) = N, \quad \text{and} \quad 0 \leq \mathcal{D} \leq 1.$$

The condition $\mathrm{Tr}\,(\mathcal{D}) = N$ is equivalent to $\int_{\mathbb{R}^3} \rho_{\mathcal{D}} = N$ and the non-negativity of $\mathcal{D}$ ensures that $\rho_{\mathcal{D}} \geq 0$. Slater determinants are characterized by density operators which are actually orthogonal projectors of rank N. More precisely, it is easy to see that there is a one-to-one correspondence between the set of finite energy Slater type density operators

$$\mathcal{P} = \left\{ \mathcal{D} \in \mathcal{L}^1, \quad \mathrm{Ran}(\mathcal{D}) \subset H^1(\mathbb{R}^3), \quad \mathcal{D}^2 = \mathcal{D} = \mathcal{D}^*, \quad \mathrm{Tr}\,(\mathcal{D}) = N \right\}$$

($\mathcal{L}^1$ denotes the vector space of trace class operators on $L^2(\mathbb{R}^3)$) and the set of molecular orbital configurations defined up to a unitary matrix. Indeed, on the one hand, the Hartree-Fock configuration $\Phi = \{\phi_i\}$ is associated with

$$\mathcal{D}_\Phi := \sum_{i=1}^{N} (\phi_i, \cdot)\phi_i,$$

which is such that $\mathcal{D}_{U\Phi} = \mathcal{D}_\Phi$ for any $U \in U(N)$ and for which

$$\tau_{\mathcal{D}_\Phi}(x, y) = \sum_{i=1}^{N} \phi_i(x)\phi_i^*(y), \qquad \rho_{\mathcal{D}_\Phi}(x) = \tau_{\mathcal{D}_\Phi}(x, x) = \sum_{i=1}^{N} |\phi_i|^2(x),$$

and on the other hand, a Slater type density operator $\mathcal{D}$ is associated with any $\Psi = U\Phi$ where $U \in U(N)$ and $\Phi = \{\phi_i\} \in \mathcal{W}$ is such that $\{\phi_i\}_{1 \leq i \leq N}$ is an orthonormal basis of $\mathrm{Ran}(\mathcal{D})$.

Let us now turn to the Hartree-Fock energy functional. Denoting by $\mathcal{D}_\Phi$ the density operator associated with the Hartree-Fock configuration $\Phi = \{\phi_i\}$, one has

$$E^{HF}(\Phi) = \mathcal{E}^{HF}(\mathcal{D}_\Phi),$$

with

$$\mathcal{E}^{HF}(\mathcal{D}) := \mathrm{Tr}\,(h\mathcal{D}) + \frac{1}{2}\,\mathrm{Tr}\,(\mathcal{G}(\mathcal{D}) \cdot \mathcal{D}).$$

where

$$h := -\Delta + V$$

and where for all $\phi \in H^1(\mathbb{R}^3)$ and all $x \in \mathbb{R}^3$

$$(\mathcal{G}(\mathcal{D}) \cdot \phi)(x) := \left(\rho_{\mathcal{D}} \star \frac{1}{|y|} \right)(x) \, \phi(x) - \int_{\mathbb{R}^3} \frac{\tau_{\mathcal{D}}(x,y)}{|x-y|} \phi(y) \, \mathrm{d}y.$$

The functional $\mathcal{G}$, describing what chemists are used to calling the bielectronic component of the energy, satisfies the following properties:

- for any $\mathcal{D} \in \mathcal{P}$, $\mathcal{G}(\mathcal{D})$ is non negative since $\forall \phi \in L^2(\mathbb{R}^3)$,

$$(\phi, \mathcal{G}(\mathcal{D}) \cdot \phi) = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \sum_{i=1}^{N} \frac{|\phi(x)\phi_i(y) - \phi(y)\phi_i(x)|^2}{|x-y|} \, \mathrm{d}x \, \mathrm{d}y \geq 0,$$

  $\Phi = \{\phi_i\}$ being such that $\mathcal{D}_\Phi = \mathcal{D}$;
- for any $\mathcal{D} \in \mathcal{P}$ and $\mathcal{D}' \in \mathcal{P}$,

$$\begin{aligned} \mathrm{Tr}\,(\mathcal{G}(\mathcal{D}) \cdot \mathcal{D}') &= \mathrm{Tr}\,(\mathcal{G}(\mathcal{D}') \cdot \mathcal{D}) \\ &= D(\rho_{\mathcal{D}}, \rho_{\mathcal{D}'}) - \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\tau_{\mathcal{D}}(x,y)\tau_{\mathcal{D}'}(x,y)^*}{|x-y|} \, \mathrm{d}x \, \mathrm{d}y \\ &= \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \sum_{i,j=1}^{N} \frac{|\phi_i(x)\phi_j'(y) - \phi_i(y)\phi_j'(x)|^2}{|x-y|} \, \mathrm{d}x \, \mathrm{d}y, \end{aligned}$$

  where $\Phi = \{\phi_i\} \in \mathcal{W}$, $\Phi' = \{\phi_i'\} \in \mathcal{W}$ are such that $\mathcal{D}_\Phi = \mathcal{D}$, $\mathcal{D}_{\Phi'} = \mathcal{D}'$. In particular,

$$\mathrm{Tr}\,(\mathcal{G}(\mathcal{D}) \cdot \mathcal{D}') \geq 0.$$

One also defines the Fock operator associated with the density operator $\mathcal{D}$ by

$$\mathcal{F}(\mathcal{D}) := h + \mathcal{G}(\mathcal{D})$$

so that for all $\phi \in H^1(\mathbb{R}^3)$ and all $x \in \mathbb{R}^3$,

$$(\mathcal{F}(\mathcal{D}) \cdot \phi)(x) = -\Delta\phi(x) + V(x)\phi(x) + \left(\rho_{\mathcal{D}} \star \frac{1}{|y|} \right)(x)\phi(x) - \int_{\mathbb{R}^3} \frac{\tau_{\mathcal{D}}(x,y)}{|x-y|} \phi(y) \, \mathrm{d}y.$$

Let us mention here that the essential spectrum of the operator $\mathcal{F}(\mathcal{D})$ is $[0, +\infty)$ for any $\mathcal{D} \in \mathcal{P}$ since the potential energy operator $V + \mathcal{G}(\mathcal{D})$ is Δ-compact. This result will be implicitly used below on several occasions.

Let us now state the minimization problem

$$\inf \left\{ \mathcal{E}^{HF}(\mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \right\}. \tag{3}$$

Lemmata 1 and 2 below provide characterizations of the stationary points of problems (2) and (3) respectively, and Lemma 3 states the connections between these two minimization problems.

Lemma 1. *The following statements are equivalent:*

1. Φ *is a stationary point of E^{HF} on $\mathcal{W}$.*

2. *There exists a hermitian matrix* $[\lambda_{ij}]$ *such that for any* $1 \le i \le N$

$$\mathcal{F}(\mathcal{D}_\Phi) \cdot \phi_i = \sum_{j=1}^{N} \lambda_{ij} \phi_j. \tag{4}$$

3. *There exists* $U \in U(N)$ *such that* $\Psi = U\Phi$ *satisfies*

$$\mathcal{F}(\mathcal{D}) \cdot \psi_i = -\epsilon_i \psi_i \tag{5}$$

for any $1 \le i \le N$ *with* $\epsilon_i \in \mathbb{R}$ *and* $\mathcal{D} = \mathcal{D}_\Phi = \mathcal{D}_\Psi$.
4. $\mathcal{F}(\mathcal{D}_\Phi)$ *and* $\mathcal{D}_\Phi$ *commute:*

$$\mathcal{F}(\mathcal{D}_\Phi)\mathcal{D}_\Phi - \mathcal{D}_\Phi \mathcal{F}(\mathcal{D}_\Phi) = 0. \tag{6}$$

Equations (4) (or equivalently (5)) are the so-called Hartree-Fock equations.

Lemma 2. *A necessary and sufficient condition for* $\mathcal{D}$ *being a stationary point of* $\mathcal{E}^{HF}$ *on* $\mathcal{P}$ *is*

$$\forall \delta \in \mathcal{L}^1, \qquad \left(\mathrm{Ran}(\delta) \subset H^1(\mathbb{R}^3), \quad \mathcal{D}\delta + \delta\mathcal{D} = \delta = \delta^*, \right.$$
$$\left. \mathrm{Tr}\, \delta = 0 \right) \Rightarrow \left(\mathrm{Tr}\, (\mathcal{F}(\mathcal{D})\delta) = 0 \right). \tag{7}$$

Lemma 3. *The functionals* E^{HF} *and* $\mathcal{E}^{HF}$ *defined on* $\mathcal{W}$ *and* $\mathcal{P}$ *respectively have the same stationary values. Moreover, if* Φ *is a stationary point of* E^{HF} *on* $\mathcal{W}$, *then* $\mathcal{D}_\Phi$ *is a stationary point of* $\mathcal{E}^{HF}$ *on* $\mathcal{P}$, *and if* $\mathcal{D}$ *is a stationary point of* $\mathcal{E}^{HF}$ *on* $\mathcal{P}$, *then any* $\Phi \in \mathcal{W}$ *spanning* $\mathrm{Ran}(\mathcal{D})$ *is a stationary point for* E^{HF}.

Remark 4. The duality between the molecular orbital and the density operator formulations has been early formulated (see in particular [14]) and extensively used by quantum chemists. In the present paper we have chosen to set up the algorithms and to state the convergence results within the density operator formalism because the expressions are simpler in this setting (at least to our point of view) and also because the convergence criteria implemented in quantum chemistry codes actually consider variations of the density operator between two successive iterations (and not variations of the molecular orbitals). However, the translation in the language of molecular orbitals is indicated throughout the text for three reasons: firstly, it is meaningful from a physical viewpoint in particular as far as the gap between occupied and unoccupied molecular orbitals is concerned, secondly, the molecular orbital formalism is widely used in the proofs, and thirdly we believe that it will make the article easier to read for the chemists and mathematicians who are more familiar with molecular orbitals than with density operators. □

Proof of Lemma 1. The Euler-Lagrange equations associated with the Hartree-Fock minimization problem (2) are the Hartree-Fock equations under the form (4). Therefore, statement 1 and 2 are equivalent. Besides, statements 2 and 3 are clearly equivalent since for any $\Phi \in \mathcal{W}$ and any $U \in U(N)$, $U\phi \in \mathcal{W}$ and $\mathcal{D}_\Phi = \mathcal{D}_{U\Phi}$. Now, we prove that statements 2 and 4 are equivalent too. Let Φ be a critical point of E^{HF} on $\mathcal{W}$; Φ satisfies the Hartree-Fock equations (4), the matrix $[\lambda_{ij}]$ being hermitian. A straightforward calculation shows that $\mathcal{F}(\mathcal{D}_\Phi)$ and $\mathcal{D}_\Phi$ commute. Now, let us consider $\Phi \in \mathcal{W}$ such that the density operator $\mathcal{F}(\mathcal{D}_\Phi)$ and $\mathcal{D}_\Phi$ commute. For any $1 \le i \le N$, the equation

$$(\mathcal{F}(\mathcal{D}_\Phi)\mathcal{D}_\Phi - \mathcal{D}_\Phi \mathcal{F}(\mathcal{D}_\Phi)) \cdot \phi_i = 0$$

can be rewritten as

$$\mathcal{F}(\mathcal{D}_\Phi)\phi_i = \sum_{j=1}^{N} (\phi_j, \mathcal{F}(\mathcal{D}_\Phi) \cdot \phi_i)\phi_j.$$

As $\mathcal{F}(\mathcal{D}_\Phi)$ is self-adjoint, the matrix $[(\phi_j, \mathcal{F}(\mathcal{D}_\Phi) \cdot \phi_i)]$ is hermitian and Φ satisfies the Hartree-Fock equations (4) with $\lambda_{ij} = (\phi_j, \mathcal{F}(\mathcal{D}_\Phi) \cdot \phi_i)$. $\qquad \square$

Proof of Lemma 2. Condition (7) simply expresses that at a critical point $\mathcal{D}$ on $\mathcal{P}$, the first order variation of the Hartree-Fock energy functional $\mathcal{E}^{HF}$ is zero on

$$T_\mathcal{D}\mathcal{P} := \left\{ \delta \in \mathcal{L}^1, \quad \mathrm{Ran}(\delta) \subset H^1(\mathbb{R}^3), \quad \mathcal{D}\delta + \delta\mathcal{D} = \delta = \delta^*, \quad \mathrm{Tr}\, \delta = 0 \right\}$$

which is the tangent subspace to $\mathcal{P}$ at $\mathcal{D}$. $\qquad \square$

Proof of Lemma 3. Let us consider a critical point Φ of the functional E^{HF} associated with the stationary value λ. Let $\delta \in \mathcal{L}^1$ satisfying $\mathrm{Ran}(\delta) \subset H^1(\mathbb{R}^3)$, $\mathcal{D}_\Phi\delta + \delta\mathcal{D}_\Phi = \delta = \delta^*$, and $\mathrm{Tr}\, \delta = 0$. Exploiting the fact that $\mathcal{D}_\Phi$ and $\mathcal{F}(\mathcal{D}_\Phi)$ commute, we obtain

$$\begin{aligned}
\mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_\Phi)\delta) &= \mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_\Phi)(\mathcal{D}_\Phi\delta + \delta\mathcal{D}_\Phi)) \\
&= 2\mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_\Phi)\mathcal{D}_\Phi\delta) \\
&= 2\mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_\Phi)\mathcal{D}_\Phi(\mathcal{D}_\Phi\delta + \delta\mathcal{D}_\Phi)) \\
&= 4\mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_\Phi)\mathcal{D}_\Phi\delta).
\end{aligned}$$

Thus $\mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_\Phi)\delta) = 0$. Therefore from Lemma 2, $\mathcal{D}_\Phi$ is a critical point of $\mathcal{E}^{HF}$ of energy λ. On the other hand, let us consider a critical point $\mathcal{D} \in \mathcal{P}$ of $\mathcal{E}^{HF}$ associated with the stationary value λ. Let $\psi \in \mathrm{Ker}(\mathcal{D}) \cap H^1(\mathbb{R}^3)$ and $\phi \in \mathrm{Ran}(\mathcal{D})$. The operator $\delta = (\phi, \cdot)\,\psi + (\psi, \cdot)\phi$ is of rank at most equal to 2 and satisfies $\mathrm{Ran}(\delta) \subset H^1(\mathbb{R}^3)$, $\mathcal{D}_\Phi\delta + \delta\mathcal{D}_\Phi = \delta = \delta^*$, and $\mathrm{Tr}\, \delta = 0$. Then $\mathrm{Tr}\, (\mathcal{F}(\mathcal{D})\delta) = 0$, which writes $(\psi, \mathcal{F}(\mathcal{D}) \cdot \phi) = 0$. Consequently, for any $\phi \in \mathrm{Ran}(\mathcal{D})$, $\mathcal{F}(\mathcal{D}) \cdot \phi \in \mathrm{Ran}(\mathcal{D})$, which means that any molecular orbital configuration Φ spanning $\mathcal{D}$ satisfies the Hartree-Fock equations (4). $\qquad \square$

2.2. On the algorithms

Let us now focus on the algorithms for solving the Hartree-Fock problem. In practice, this problem is in general approximated by a Galerkin method: the ϕ_i are spanned over a finite basis of atomic orbitals. This is the *linear combination of atomic orbitals* (LCAO) approximation (see [7], p. 17). For computational reasons that we do not detail here, the atomic orbitals χ_p are most often contracted Gaussian functions, namely finite sums of polynomials times Gaussian functions (see again [7], for instance). We recall from the introduction that, even if the SCF algorithms are always solved in finite dimension, they are considered in this paper as algorithms for solving the original problem in infinite dimension. Except Propositions 5 and 6 which are specific to the infinite dimension, all the results below can be applied to the Hartree-Fock problem within the LCAO approximation.

The algorithms under examination can be written in terms of density operators. In order to study their convergence, $\mathcal{P}$ is equipped by two distances d_0 and d_1 respectively inherited from the norms:

$$\|A\|_0 = (\mathrm{Tr}\, (A^*A))^{1/2}, \quad \text{and} \quad \|A\|_1 = (\mathrm{Tr}\, (A^*(-\Delta + 1)A))^{1/2}.$$

The norm $\|\cdot\|_0$ and $\|\cdot\|_1$ are the norms associated with Hilbert-Schmidt operators on $L^2(\mathbb{R}^3)$ and $H^1(\mathbb{R}^3)$ respectively. We have in particular for any $\Phi \in \mathcal{W}$, and $\Phi' \in \mathcal{W}$,

$$d_0(\mathcal{D}_\Phi, \mathcal{D}_{\Phi'}) = \|\mathcal{D}_\Phi - \mathcal{D}_{\Phi'}\|_0 = \left(2N - 2 \sum_{i,j=1}^N |(\phi_i, \phi_j')|^2 \right)^{1/2}.$$

Contrary to direct minimization algorithms, which directly tackle one of the minimization problems (2) or (3), self-consistent field methods for solving the Hartree-Fock equations consist in solving the nonlinear problem (6)

(or any equivalent problem, see Lemmata 1, 2 and 3) with an iterative fixed point procedure of the general form

$$(SCF) \qquad (\mathcal{D}_k)_{0 \le k \le n} \xrightarrow{\ 1\ } \widetilde{\mathcal{F}}_n \xrightarrow{\ 2\ } \mathcal{D}_{n+1}.$$

Step 1 consists in building a pseudo-Fock operator $\widetilde{\mathcal{F}}_n$ from the density operators $(\mathcal{D}_k)_{0 \le k \le n}$ computed in the previous iterations and step 2 consists in defining a new density operator $\mathcal{D}_{n+1}$ from $\widetilde{\mathcal{F}}_n$.

The *Roothaan algorithm* (also called *simple SCF* or *pure SCF* or *customary iterative procedure* in the literature) is natural when considering the Hartree-Fock equations under the form (5). It is defined by $\widetilde{\mathcal{F}}_n = \mathcal{F}(\mathcal{D}_n)$ and by the so-called *aufbau* principle which consists in taking for $\mathcal{D}_{n+1}$ a minimizer of

$$\inf \left\{ \mathrm{Tr}\,(\widetilde{\mathcal{F}}_n \mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \right\}$$

(issues concerning the existence and uniqueness of the solution to this minimization problem are examined below). In terms of molecular orbitals, the *aufbau* principle consists in taking $\mathcal{D}_{n+1} = \mathcal{D}_{\Phi_{n+1}}$, where Φ_{n+1} is any Hartree-Fock configuration obtained by choosing any set of orthonormal molecular orbitals ϕ_i^{n+1} corresponding to the N smallest eigenvalues (including multiplicity) $-\epsilon_i^{n+1}$ of $\widetilde{\mathcal{F}}_n$, that is to say in populating the molecular orbitals by starting with those of lowest energies. The Roothaan algorithm can be summarized in this way:

$$(Rth) \qquad \mathcal{D}_n \ \longrightarrow \ \widetilde{\mathcal{F}}_n = \mathcal{F}(\mathcal{D}_n) \ \xrightarrow{aufbau} \ \mathcal{D}_{n+1}.$$

The *level-shifting algorithm* generalizes the Roothaan algorithm. It is defined by

$$\widetilde{\mathcal{F}}_n = \mathcal{F}(\mathcal{D}_n) - b\mathcal{D}_n,$$

where b is a real non-negative parameter, and by the *aufbau* principle. The Roothaan algorithm is recovered by taking $b = 0$. If $\mathcal{D}$ is a critical point of the Hartree-Fock functional satisfying the *aufbau* principle, the spectrum of $\mathcal{F}(\mathcal{D}) - b\mathcal{D}$ is

$$\{-\epsilon_1 - b, -\epsilon_2 - b, \cdots, -\epsilon_N - b, -\epsilon_{N+1}, -\epsilon_{N+2}, \cdots\} \cup [0, +\infty[.$$

This means that the occupied energy levels have been shifted of magnitude $-b$ under the action of the operator $-b\mathcal{D}$. The level-shifting algorithm can be summarized by

$$(LS^b) \qquad \mathcal{D}_n^b \ \longrightarrow \ \widetilde{\mathcal{F}}_n = \mathcal{F}(\mathcal{D}_n^b) - b\mathcal{D}_n^b \ \xrightarrow{aufbau} \ \mathcal{D}_{n+1}^b.$$

Let us now focus a little longer on the *aufbau* principle. This way of populating the molecular orbitals is justified by the mathematical result stating that, at the Hartree-Fock minimum, the molecular orbitals are indeed populated according to the *aufbau* principle applied to the Fock operator [13]: the Lagrange multipliers appearing in (5) *are* the N lowest eigenvalues of the Fock operator. However, two types of difficulties may *a priori* appear when resorting to the *aufbau* procedure within an algorithm:

1. Existence problems: the problem

$$\inf \left\{ \mathrm{Tr}\,(\widetilde{\mathcal{F}}_n \mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \right\} \tag{8}$$

may admit no minimizer. This occurs when the pseudo-Fock operator $\widetilde{\mathcal{F}}_n$ has less that N eigenvalues (including multiplicity) smaller than or equal to the lower bound of the continuous spectrum;

2. Uniqueness problems: the minimization problem (8) may have more than one solution. This occurs when the Nth and the $(N+1)$th smallest eigenvalues of $\widetilde{\mathcal{F}}_n$ are equal, or in chemical language when there is no

gap between the highest occupied molecular orbital (HOMO) ϕ_N^{n+1} and the lowest unoccupied molecular orbital (LUMO) ϕ_{N+1}^{n+1}.

Let us point out that of course existence problems never come across in finite dimension (*i.e.* within the LCAO approximation), but that they may be present in infinite dimension as shown below in Proposition 6. On the other hand, uniqueness problem may even occur in finite dimension. They are in general related to the symmetries of the system. In the cases when the system does not exhibit any symmetry, numerical experiment shows that the eigenvalues of $\widetilde{\mathcal{F}}_n$ are generically non-degenerate for any n, whereas it may not be the case when the system does exhibit symmetries (consider for instance the spherical symmetry of the Hamiltonian in the atomic case $M = 1$).

We will say in the sequel that a SCF algorithm with initial guess $\mathcal{D}_0$ is *well posed* if it generates a sequence $(\mathcal{D}_n)_{n \in \mathbb{N}}$ defined in a unique way. In particular, when the construction of the updated density operator uses the *aufbau* principle, well-posedness implies that problem (8) has a unique solution for any $n \in \mathbb{N}$.

We will also say that a SCF algorithm of the form

$$(A) \qquad \mathcal{D}_n \quad \longrightarrow \quad \widetilde{\mathcal{F}}_n \quad \overset{aufbau}{\longrightarrow} \quad \mathcal{D}_{n+1}$$

with initial guess $\mathcal{D}_0$ is *uniformly well posed* if it is well posed and if in addition the following properties are fulfilled:

1. There exists $\epsilon > 0$ such that for any $n \in \mathbb{N}$, $\widetilde{\mathcal{F}}_n$ has at least N eigenvalues below $(\inf \sigma_c(\widetilde{\mathcal{F}}_n) - \epsilon)$, $\sigma_c(\cdot)$ denoting the continuous spectrum;
2. There exists $\gamma > 0$ such that for any $n \in \mathbb{N}$,

$$\inf \sigma \left(\widetilde{\mathcal{F}}_n + \sum_{i=1}^{N} \epsilon_i^{n+1}(\phi_i^{n+1}, \cdot)\phi_i^{n+1} \right) \geq -\epsilon_N^{n+1} + \gamma,$$

$\sigma(\cdot)$ denoting the spectrum, $-\epsilon_1^{n+1} \leq -\epsilon_2^{n+1} \leq \cdot \leq -\epsilon_N^{n+1}$ the N smallest eigenvalues of $\widetilde{\mathcal{F}}_n$, and $(\phi_i^{n+1})_{1 \leq i \leq N}$ an orthonormal set of associated eigenvectors.

In Theorem 11, the level-shifting algorithm (LS^b) with initial guess $\mathcal{D}_0$ is *proved* to satisfy the well-posedness property provided b is larger than some b_0 depending on $\mathcal{D}_0$. Consequently (see Th. 11 again), this algorithm enjoys good convergence properties. On the other hand, well-posedness remains so far an *assumption* for the Roothaan algorithm necessary to obtain the (poor) convergence results we shall establish in the forthcoming section (Th. 7). Indeed we unfortunately are not able to prove well-posedness for the Roothaan algorithm except for the special case (trivial form the standpoint of applications) when the initial guess $\mathcal{D}_0$ is a minimizer of the Hartree-Fock functional $\mathcal{E}^{HF}$ and when the molecular system is either a positive ion or a neutral system: $\mathcal{D}_0$ is then a fixed point of the algorithm and the uniform well-posedness is guaranteed because at any minimum $\mathcal{D}$ of the Hartree-Fock energy $\mathcal{E}^{HF}$ for a non-negatively charged molecular system, the Fock operator $\mathcal{F}(\mathcal{D})$ satisfies the following two properties:

1. $\mathcal{F}(\mathcal{D})$ has at least N negative eigenvalues including multiplicity $-\epsilon_1 \leq -\epsilon_2 \leq \cdots \leq -\epsilon_N < 0$ (see [13], pp. 39-40);
2. There exists a positive gap between $-\epsilon_N$ and the part of the spectrum above this eigenvalue (see [2]). Let us point out however that in models including spin, this assertion remains true for the GHF model (called "Unrestricted" in [2]) but is not proved for the UHF nor for the RHF models.

The above two properties of course motivate the definitions of well-posedness we have introduced.

To conclude this section, let us mention another SCF algorithm widely used in Hartree-Fock calculations, the so-called *Direct Inversion in the Iteration Space* (DIIS) algorithm for which the pseudo-Fock operator $\widetilde{\mathcal{F}}_n$ is build using not only the latest computed density operator $\mathcal{D}_n$, but all the density operators computed previously.

More precisely, the pseudo-Fock operator $\widetilde{\mathcal{F}}_n$ is a linear combination

$$\widetilde{\mathcal{F}}_n = \sum_{k=0}^{n} c_k^n \mathcal{F}(\mathcal{D}_k),$$

of the (n+1) Fock operators corresponding to the (n+1) density operators $\mathcal{D}_0, \cdots, \mathcal{D}_n$, the coefficients c_k^n being computed by solving the quadratic minimization problem

$$\inf \left\{ \| \sum_{k=0}^{n} c_k^n e_k \|_0^2, \quad \sum_{k=0}^{n} c_k^n = 1 \right\}, \qquad \text{with } e_k = \mathcal{F}(\mathcal{D}_k)\mathcal{D}_k - \mathcal{D}_k \mathcal{F}(\mathcal{D}_k)$$

(let us recall that $\| \cdot \|_0$ denotes the norm associated with Hilbert-Schmidt operators on $L^2(\mathbb{R}^3)$). The density operator $\mathcal{D}_{n+1}$ is then computed from the eigenvectors of the pseudo-Fock operator $\widetilde{\mathcal{F}}_n$ by using the *aufbau* principle. Let us notice that $e_k = 0$ if and only if $\mathcal{D}_k$ is a critical point of the Hartree-Fock functional $\mathcal{E}^{HF}$. We hope to address the question of convergence of the DIIS type algorithms in a future publication.

3. ON THE ROOTHAAN ALGORITHM

Let us begin the analysis of the Roothaan algorithm by considerations related to well-posedness in the sense we have defined in the previous section. In the Roothaan algorithm, well-posedness is fulfilled if and only if the minimization problem

$$\inf \{ \mathrm{Tr} \, (\mathcal{F}(\mathcal{D}_n)\mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \}$$

has a unique solution for any $n \in \mathbb{N}$. Although we are not able to deal with the uniqueness problem and therefore *postulate* uniqueness holds to go further (Th. 7), we are able to give definite conclusions on the existence problem.

Proposition 5. *For positive ions (Z > N), the minimization problem*

$$\inf \left\{ \mathrm{Tr} \, (\mathcal{F}(\widetilde{\mathcal{D}})\mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \right\}$$

has at least one solution for any $\widetilde{\mathcal{D}} \in \mathcal{P}$.

Proof. For any $\mathcal{D} \in \mathcal{P}$,

$$\mathcal{F}(\mathcal{D}) \leq -\Delta + V + (\rho_\mathcal{D} \star \frac{1}{|x|}).$$

For positive ions one has in addition $\int_{\mathbb{R}^3} \rho_\mathcal{D} = N < Z$. Therefore, $\mathcal{F}(\mathcal{D})$ has an infinity of negative eigenvalues since this is the case for the operator $-\Delta + V + (\rho_\mathcal{D} \star \frac{1}{|x|})$ in view of Lemma II.1 in the article by Lions [13]. The latter result lays on the fact that at infinity $-\Delta + V + (\rho_\mathcal{D} \star \frac{1}{|x|})$ behaves as $-\Delta - \frac{Z-N}{|x|}$; a scaling argument enables one to conclude. $\qquad\square$

Proposition 6. *For any neutral (Z = N) or negatively charged (N > Z) system, there exists $\mathcal{D}_0 \in \mathcal{P}$ such that the minimization problem*

$$\inf \{ \mathrm{Tr} \, (\mathcal{F}(\mathcal{D}_0)\mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \} \tag{9}$$
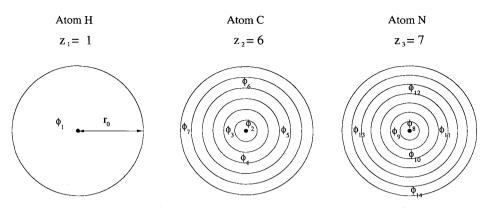
has no solution.

FIGURE 1. Supports of the ϕ_i for the HCN molecule.

Proof. Let us first of all prove that the self-adjoint operator

$$H_{a,\eta} = -\Delta - \frac{a}{|x|} 1_{|x| \leq \eta}$$

is non-negative when $a > 0$, $\eta \geq 0$ and $\eta < 1/a$. As the continuous spectrum of $H_{a,\eta}$ equals $[0, +\infty)$, it is sufficient to establish that $H_{a,\eta}$ has no bound states, that is to say no negative eigenvalues. This is a direct consequence of Bargmann's theorem for central potential Hamiltonians of the form $H = -\Delta + v(r)$ (see [19], p. 94), which states in particular that if

$$\int_0^{+\infty} r|v(r)|\,\mathrm{d}r < 1,$$

then H has no bound states. The above condition reads $\eta < 1/a$ for $H_{a,\eta}$.

Let us now consider $0 < r_0 < (2M\max(z_1, \cdots, z_k))^{-1}$ such that $|\bar{x}_k - \bar{x}_l| > 2\,r_0$, for any $1 \leq k < l \leq M$, and $\{\phi_i\}_{1 \leq i \leq N}$ such that for any $1 \leq i \leq N$,

- $\phi_i \in \mathcal{D}(\mathbb{R}^3)$, ϕ_i real-valued, non-negative and such that $\int_{\mathbb{R}^3} \phi_i^2 = 1$;

- for $1 \leq i \leq Z$, ϕ_i is supported in $\left\{ x \in \mathbb{R}^3, \ \frac{j-1}{z_k}r_0 \leq |x - \bar{x}_k| \leq \frac{j}{z_k}r_0 \right\}$, where j and k are the only integers such that $i = \sum_{l=1}^{k-1} z_l + j$ with $1 \leq j \leq z_k$, and ϕ_i is spherically symmetric with respect to $\bar{x}_k$. A graphical representation of this construction is provided below (see Fig. 1). In the case of negative ions the remaining electrons are distributed in the same way around a fictitious nucleus located at $\bar{x}_{M+1}$ such that for any $1 \leq l \leq M$, $|\bar{x}_{M+1} - \bar{x}_l| > 2\,r_0$.

Let us now consider the scalings of the ϕ_i defined in the following way:

$$\phi_i^\sigma(x) = \sigma^{3/2}\phi_i(\sigma(x - \bar{x}_k) + \bar{x}_k), \qquad 1 \leq i \leq Z,$$

for $\sigma \geq 1$, k being defined as above ($k = M + 1$ for $Z + 1 \leq i \leq N$ in the case of negative ions), and let us denote by $\mathcal{D}^\sigma = \mathcal{D}_{\Phi^\sigma}$. It is easy to prove that for any $\sigma \geq 1$,

- $\Phi^\sigma \in \mathcal{W}$;
- $\|\phi_i^\sigma\|_{L^{3/2}(\mathbb{R}^3)} = \sigma^{-1/2}\|\phi_i\|_{L^{3/2}(\mathbb{R}^3)}$;
- $W^\sigma = V + \left(\rho_{\mathcal{D}^\sigma} \star \frac{1}{|x|}\right)$ is such that

$$W^\sigma(x) \begin{cases} \geq -\frac{z_k}{|x - \bar{x}_k|} & \text{when} \quad |x - \bar{x}_k| < r_0, \quad 1 \leq k \leq M, \\ = 0 & \text{otherwise.} \end{cases}$$

(these are consequences of Gauss Theorem for spherically symmetric charge distributions).

Let $\phi \in H^1(\mathbb{R}^3)$. We have

$$
(\phi, \mathcal{F}(\mathcal{D}^\sigma)\phi) = \int_{\mathbb{R}^3} |\nabla \phi|^2 + \int_{\mathbb{R}^3} W^\sigma |\phi|^2 - \sum_{k=1}^N D(\phi_i^\sigma \phi^*, \phi_i^\sigma \phi)
$$

$$
\geq \sum_{k=1}^M \frac{1}{2M} \left[\int_{\mathbb{R}^3} |\nabla \phi|^2 - 2Mz_k \int_{|x-\bar{x}_k|<r_0} \frac{|\phi(x)|^2}{|x-\bar{x}_k|} \right] + \sum_{i=1}^N \frac{1}{2N} \left[\int_{\mathbb{R}^3} |\nabla \phi|^2 - 2ND(\phi_i^\sigma \phi^*, \phi_i^\sigma \phi) \right].
$$

On the one hand, r_0 has been chosen such that for any $1 \leq k \leq M$, Hardy inequality yields

$$
\int_{\mathbb{R}^3} |\nabla \phi|^2 - 2Mz_k \int_{|x-\bar{x}_k|<r_0} \frac{|\phi(x)|^2}{|x-\bar{x}_k|} \geq 0,
$$

and on the other hand one has for any $\phi \in H^1(\mathbb{R}^3)$ and $\psi \in H^1(\mathbb{R}^3)$,

$$
\begin{aligned}
|D(\phi\psi^*, \phi^*\psi)| &\leq \|(\phi\psi^* \star \frac{1}{|x|})\phi^*\psi\|_{L^1(\mathbb{R}^3)} \\
&\leq C_u \|\phi\psi^* \star \frac{1}{|x|}\|_{L^6(\mathbb{R}^3)} \|\phi^*\psi\|_{L^{6/5}(\mathbb{R}^3)} \\
&\leq C_u \|\frac{1}{|x|}\|_{L^{3,\infty}(\mathbb{R}^3)} \|\phi\psi^*\|_{L^{6/5}(\mathbb{R}^3)}^2 \\
&\leq C_u \|\psi\|_{L^{3/2}(\mathbb{R}^3)}^2 \|\phi\|_{L^6(\mathbb{R}^3)}^2 \\
&\leq C_u \|\psi\|_{L^{3/2}(\mathbb{R}^3)}^2 \|\nabla \phi\|_{L^2(\mathbb{R}^3)}^2,
\end{aligned}
$$

where C_u denotes any universal constant. Therefore, there exist $\sigma_0 \geq 1$ and $\alpha > 0$ such that

$$
(\phi, \mathcal{F}(\mathcal{D}^\sigma)\phi) \geq \sum_{i=1}^N \frac{1}{2N} \left[1 - 2NC_u\sigma^{-1}\|\phi_i\|_{L^{3/2}(\mathbb{R}^3)}^2 \right] \|\nabla \phi\|_{L^2(\mathbb{R}^3)}^2 \geq \alpha \|\nabla \phi\|_{L^2(\mathbb{R}^3)}^2
$$

for any $\sigma \geq \sigma_0$ and any $\phi \in H^1(\mathbb{R}^3)$. In particular, $\mathcal{F}(\mathcal{D}^{\sigma_0})$ has no non positive eigenvalues. Therefore, problem (9) has no solution since the continuous spectrum of $\mathcal{F}(\mathcal{D}^\sigma)$ is $[0, +\infty)$.   □

Let us now leave apart the cases when the Roothaan algorithm is not well defined and consider situations when the initial guess $\mathcal{D}_0$ is such that the Roothaan algorithm is *uniformly* well-posed.

Many observations on various chemical systems have confirmed that even under the assumption of uniform well-posedness, the Roothaan algorithm does not always converge. The following theorem gives hints to understand the behavior of this algorithm.

Theorem 7. *Let $\mathcal{D}_0 \in \mathcal{P}$ such that the Roothaan algorithm with initial guess $\mathcal{D}_0$ is uniformly well-posed. Let us consider the functional*

$$
\mathcal{E}(\mathcal{D}, \mathcal{D}') = \mathrm{Tr}\,(h\mathcal{D}) + \mathrm{Tr}\,(h\mathcal{D}') + \mathrm{Tr}\,(\mathcal{G}(\mathcal{D}) \cdot \mathcal{D}')
$$

defined on $\mathcal{P} \times \mathcal{P}$. The sequence $(\mathcal{D}_n)$ computed by the Roothaan algorithm is such that $\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1})$ decreases towards some stationary value $\lambda \in \mathbb{R}$ of the functional $\mathcal{E}$ and the sequence $(\mathcal{D}_{2n}, \mathcal{D}_{2n+1})$ (resp. $(\mathcal{D}_{2n+1}, \mathcal{D}_{2n+2})$) converges in $(\mathcal{P}, d_1) \times (\mathcal{P}, d_1)$ up to an extraction to a critical point of $\mathcal{E}$ associated to the stationary value λ.

In addition,

$$\sum_{n=0}^{+\infty} \|\mathcal{D}_{n+2} - \mathcal{D}_n\|_0^2 < +\infty. \tag{10}$$

Before giving the proof of Theorem 7, let us make a few comments. The above theorem shows that the Roothaan algorithm minimizes some functional $\mathcal{E}$ *which is not* the Hartree-Fock functional, but satisfies however for all $\mathcal{D} \in \mathcal{P}$,

$$\mathcal{E}(\mathcal{D}, \mathcal{D}) = 2\,\mathcal{E}^{HF}(\mathcal{D}).$$

For the sake of simplicity, let us assume for a start that the whole sequence $(\mathcal{D}_{2n}, \mathcal{D}_{2n+1})$ converges, *i.e.* that

$$\mathcal{D}_{2n} \underset{n\to+\infty}{\longrightarrow} \mathcal{D}, \qquad \mathcal{D}_{2n+1} \underset{n\to+\infty}{\longrightarrow} \mathcal{D}'. \tag{11}$$

We then face the following alternative:

- either $\mathcal{D} = \mathcal{D}'$ and the whole sequence $(\mathcal{D}_n)$ converges in $(\mathcal{P}, d_1)$ towards a critical point of the Hartree-Fock functional;
- or $\mathcal{D} \neq \mathcal{D}'$ and the sequence of the Hartree-Fock energies $(\mathcal{E}^{HF}(\mathcal{D}_n))$ oscillates between the two asymptotic values $\mathcal{E}^{HF}(\mathcal{D})$ and $\mathcal{E}^{HF}(\mathcal{D}')$ which are both greater than the Hartree-Fock ground state energy.

Unfortunately, we cannot prove the convergence of the whole sequence, not even in $(\mathcal{P}, d_0)$. However, in view of (11), we know that $\|\mathcal{D}_{n+2} - \mathcal{D}_n\|_0 \to 0$. Consequently, denoting $(\mathcal{D}, \mathcal{D}')$ one of the accumulation points of the sequence $(\mathcal{D}_{2n}, \mathcal{D}_{2n+1})$ one observes anyway *in numerical simulations:*

- either a convergence of the Hartree-Fock energy towards a stationary value, if $\mathcal{D} = \mathcal{D}'$ (the convergence of the density operators is not guaranteed);
- or in the other case, an oscillation between two states, possibly accompanied by slow drifts of the two states when (11) does not hold.

Remark 8. In finite dimension, the uniform well-posedness assumption resumes to the existence of a uniform gap between the Nth and the $(N+1)$th eigenvalues of the finite dimensional Fock operator. Indeed, statement 1 of the uniform well-posedness assumption is always true in finite dimension since there is no continuous spectrum. It is actually used in the proof of Theorem 7 in order to supply compactness, which is not necessary in finite dimension. Consequently, uniform well-posedness is a rather weak assumption, and we can conclude that the behavior of the Roothaan algorithm is rather well described by the conclusions of Theorem 7 and the above discussion. $\square$

We do not know whether it is possible in the general case to know *a priori* whether the Roothaan algorithm will converge or oscillate. However, we can make the following observation. Since $\mathcal{E}(\mathcal{D}, \mathcal{D}) = 2\mathcal{E}^{HF}(\mathcal{D})$, we have always

$$\inf\{\mathcal{E}(\mathcal{D}, \mathcal{D}'), \quad \mathcal{D} \in \mathcal{P}, \quad \mathcal{D}' \in \mathcal{P}\} \le 2\inf\{\mathcal{E}^{HF}(\mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\}.$$

Let us now assume that

$$\inf\{\mathcal{E}(\mathcal{D}, \mathcal{D}'), \quad \mathcal{D} \in \mathcal{P}, \quad \mathcal{D}' \in \mathcal{P}\} < 2\inf\{\mathcal{E}^{HF}(\mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\}. \tag{12}$$

If (12) is fulfilled and if the chosen initial guess $\mathcal{D}_0$ satisfies

$$\inf\{\mathcal{E}(\mathcal{D}_0, \mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\} < 2\inf\{\mathcal{E}^{HF}(\mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\},$$

then the Roothaan algorithm cannot converge to a solution to the Hartree-Fock equation since if one assumes that $\mathcal{D}_n \to \widetilde{\mathcal{D}}$ (in $(\mathcal{P}, d_0)$ for instance), one obtains

$$
\begin{aligned}
2\mathcal{E}^{HF}(\widetilde{\mathcal{D}}) &= \mathcal{E}(\widetilde{\mathcal{D}}, \widetilde{\mathcal{D}}) \\
&\leq \lim_{n \to +\infty} \mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) \\
&\leq \mathcal{E}(\mathcal{D}_0, \mathcal{D}_1) \\
&= \inf \{\mathcal{E}(\mathcal{D}_0, \mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\} \\
&< 2 \inf \{\mathcal{E}^{HF}(\mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\},
\end{aligned}
$$

which clearly highlights a contradiction. In particular, if $(\mathcal{D}_0, \mathcal{D}_0')$ is a minimizer of the functional $\mathcal{E}$, then the Roothaan algorithm oscillates (under the well-posedness assumption) between the two states $\mathcal{D}_0$ and $\mathcal{D}_0'$. In the following example, which concerns the Restricted Hartree-Fock (RHF) model, we exhibit a case when inequality (12) holds. We have not been able to build a similar example for a two electron system within the (general) Hartree-Fock model.

Example 9. Let us consider the Hamiltonian

$$
H_a = -\Delta_{x_1} - \Delta_{x_2} + V_a(x_1) + V_a(x_2) + \frac{1}{|x_1 - x_2|},
$$

with

$$
V_a(x) = -\frac{1}{|x - a|} - \frac{1}{|x + a|},
$$

which models the hydrogen molecule H_2, the hydrogen atoms being located at points a and $-a$. We are searching the electronic ground state of this molecule within the Restricted Hartree-Fock (RHF) approximation. Let us recall that under the RHF approximation, the two electrons of the H_2 molecule are described by the same spatial molecular orbital $\phi \in H^1(\mathbb{R}^3)$, one electron being of *spin up*, the other electron being of *spin down* (see [7], p. 22). The SCF algorithm can thus be considered as an iterative procedure on ϕ for searching a solution to the Euler-Lagrange equation of the problem

$$
\inf \left\{ E^{RHF}(\phi), \quad \phi \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} |\phi|^2 = 1 \right\}
$$

where the RHF energy functional of the H_2 molecule reads

$$
E^{RHF}(\phi) = 2 \int_{\mathbb{R}^3} |\nabla \phi|^2 + 2 \int_{\mathbb{R}^3} V_a |\phi|^2 + D(|\phi|^2, |\phi|^2).
$$

We prove below that inequality (12) holds when $|a|$ is large enough and that for some initial guess ϕ_0, the Roothaan algorithm actually oscillates between two states. It is not a physically relevant case, for no Chemist would of course compute the dissociation of the hydrogen molecule within the RHF approximation. However, we find this example instructive because as far as we know it is the only situation to this day for which we can mathematically prove the existence of oscillations between two states for the Roothaan algorithm. It is also important to note that this behavior is not basis dependent since we reason about the original infinite dimensional Hartree-Fock equations. Let us now establish inequality (12), which also reads for this particular

problem

$$\inf\left\{E(\phi_1,\phi_2), \quad \phi_1 \in H^1(\mathbb{R}^3), \quad \phi_2 \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3}|\phi_1|^2 = 1, \quad \int_{\mathbb{R}^3}|\phi_2|^2 = 1\right\}$$
$$< \inf\left\{E^{RHF}(\phi), \quad \phi \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3}|\phi|^2 = 1\right\}.$$

with

$$E(\phi_1,\phi_2) = \int_{\mathbb{R}^3}|\nabla\phi_1|^2 + \int_{\mathbb{R}^3}V_a|\phi_1|^2 + \int_{\mathbb{R}^3}|\nabla\phi_2|^2 + \int_{\mathbb{R}^3}V_a|\phi_2|^2 + D(|\phi_1|^2,|\phi_2|^2).$$

Let us denote by $f(|a|)$ and $g(|a|)$ the left-hand side and the right-hand side respectively of the above inequality. The functions $|a| \mapsto f(|a|)$ and $|a| \mapsto g(|a|)$ are clearly continuous. Therefore, for establishing that the above inequality holds for $|a|$ large enough, it is sufficient to prove that

$$\limsup_{|a|\to+\infty} f(|a|) < \liminf_{|a|\to+\infty} g(|a|).$$

On the one hand, reasoning as in the proof of Proposition 6, we localize for instance ϕ_1 in $B_{|a|}(a)$ and ϕ_2 in $B_{|a|}(-a)$ $(B_R(x) := \{y \in \mathbb{R}^3, \quad |y - x| < R\})$ and we thus obtain

$$f(|a|) \leq 2\inf\left\{\int_{B_{|a|}(0)}|\nabla\phi|^2 - \int_{B_{|a|}(0)}\frac{|\phi|^2}{|x|}, \quad \phi \in H_0^1(B_{|a|}(0)), \quad \phi \text{ radial}, \quad \int_{\mathbb{R}^3}|\phi|^2 = 1\right\}.$$

Passing to the limit $|a| \to +\infty$, it follows that

$$\limsup_{|a|\to+\infty} f(|a|) \leq 2\lambda_1(-\Delta - \frac{1}{|x|}),$$

$\lambda_1(-\Delta - \frac{1}{|x|})$ denoting the first eigenvalue of the operator $-\Delta - \frac{1}{|x|}$ on $L^2(\mathbb{R}^3)$. On the other hand, denoting by ϕ_a a minimizer of the variational problem defining $g(|a|)$, we have, since ϕ_a, which is unique for a given $a \in \mathbb{R}^3$ up to a multiplying constant, satisfies $\phi_a(x) = \phi_a(-x)$,

$$g(|a|) = 2\int_{\mathbb{R}^3}|\nabla\phi_a|^2 + 2\int_{\mathbb{R}^3}V_a|\phi_a|^2 + D(|\phi_a|^2,|\phi_a|^2)$$
$$\geq 2\int_{\mathbb{R}^3}|\nabla\phi_a|^2 + 2\int_{\mathbb{R}^3}V_a|\phi_a|^2 + D(|\phi_a\,1_{x\cdot a>0}|^2,|\phi_a\,1_{x\cdot a>0}|^2) + D(|\phi_a\,1_{x\cdot a<0}|^2,|\phi_a\,1_{x\cdot a<0}|^2)$$
$$= 2\left(2\int_{x\cdot a>0}|\nabla\phi_a|^2 + 2\int_{x\cdot a>0}V_a|\phi_a|^2 + D(|\phi_a\,1_{x\cdot a>0}|^2,|\phi_a\,1_{x\cdot a>0}|^2)\right)$$
$$\geq 2\inf\left\{2\int_{x\cdot a>0}|\nabla\phi|^2 + 2\int_{x\cdot a>0}V_a|\phi|^2 + D(|\phi|^2,|\phi|^2), \quad \phi \in H^1(x\cdot a > 0), \quad \int_{x\cdot a>0}|\phi|^2 = \frac{1}{2}\right\}.$$

Therefore, it is now easy to check that

$$\liminf_{|a| \to +\infty} g(|a|) \geq 2 \inf \left\{ 2 \int_{\mathbb{R}^3} |\nabla \phi|^2 + 2 \int_{\mathbb{R}^3} \frac{|\phi|^2}{|x|} + D(|\phi|^2, |\phi|^2), \quad \phi \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} |\phi|^2 = \frac{1}{2} \right\}$$

$$> 2 \inf \left\{ 2 \int_{\mathbb{R}^3} |\nabla \phi|^2 - 2 \int_{\mathbb{R}^3} \frac{|\phi|^2}{|x|}, \quad \phi \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} |\phi|^2 = \frac{1}{2} \right\}$$

$$= 2\lambda_1 (-\Delta - \frac{1}{|x|}).$$

Finally,

$$\limsup_{|a| \to +\infty} f(|a|) \leq 2\lambda_1(-\Delta - \frac{1}{|x|}) < \liminf_{|a| \to +\infty} g(|a|),$$

which proves inequality (12) for $|a|$ large enough. Let us now consider a minimizer (ϕ_1, ϕ_2) of the variational problem defining $f(|a|)$ for a given $a \in \mathbb{R}^3$ with $|a|$ large enough for (12) being fulfilled. By a classical argument ϕ_1 (resp. ϕ_2) is the ground state of the operator $-\Delta + V_a + |\phi_2|^2 \star \frac{1}{|x|}$ (resp. $-\Delta + V_a + |\phi_1|^2 \star \frac{1}{|x|}$), *which is non-degenerate*. It is now straightforward to conclude that if ϕ_1 (or ϕ_2) is chosen as initial guess of the Roothaan procedure, the algorithm is (uniformly) well-posed and it oscillates between the two states ϕ_1 and ϕ_2 (up to multiplying constants whose choices leave unchanged the density operators $\mathcal{D}_{\phi_k} = 2(\phi_k, \cdot)\phi_k$).   □

For the reader who does not want to enter the technicalities of the above proof, who is only interested in the mathematical arguments and who does not want to bother with the sake of restricting himself to cases of chemical interest, we provide now an academic but convincing situation when oscillations of the Roothaan algorithm can be exhibited. In this academic situation, the proof is surprisingly simple. Let us indeed denote by $w > 0$ a positive minimizer of the Choquart-Pekar problem [10]

$$\inf \left\{ \int_{\mathbb{R}^3} |\nabla v|^2 - \frac{1}{2} D(|v|^2, |v|^2), \quad v \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} |v|^2 = 1 \right\}.$$

If we replace in Example 9 the point nuclei by smeared nuclei associated with the charge distribution w^2, one obtains that $\phi_1 = w(\cdot - a)$ and $\phi_2 = w(\cdot + a)$ with $a \in \mathbb{R}^3 \setminus \{0\}$ satisfy

$$\begin{cases} -\Delta \phi_1 + V_a \phi_1 + (|\phi_2|^2 \star \frac{1}{|x|})\phi_1 = \lambda \phi_1, \\ -\Delta \phi_2 + V_a \phi_2 + (|\phi_1|^2 \star \frac{1}{|x|})\phi_2 = \lambda \phi_2, \\ \phi_1 \in H^1(\mathbb{R}^3), \quad \phi_2 \in H^1(\mathbb{R}^3), \quad \phi_1 > 0, \quad \phi_2 > 0, \\ \int_{\mathbb{R}^3} |\phi_1|^2 = 1, \quad \int_{\mathbb{R}^3} |\phi_2|^2 = 1, \end{cases}$$

where $V_a = -w^2(\cdot - a) \star \frac{1}{|x|} - w^2(\cdot + a) \star \frac{1}{|x|}$ is the Coulomb potential generated by the smeared nuclei. Therefore (ϕ_1, ϕ_2) is a critical point of the functional

$$E(\phi_1, \phi_2) = \int_{\mathbb{R}^3} |\nabla \phi_1|^2 + \int_{\mathbb{R}^3} V_a |\phi_1|^2 + \int_{\mathbb{R}^3} |\nabla \phi_2|^2 + \int_{\mathbb{R}^3} V_a |\phi_2|^2 + D(|\phi_1|^2, |\phi_2|^2)$$

such that $\phi_1 \neq \phi_2$ and if ϕ_1 (or ϕ_2) is chosen as initial guess of the Roothaan procedure, the algorithm is well-posed (in view the non-degeneracy argument already mentioned in Ex. 9) and leads to an oscillation between the two states ϕ_1 and ϕ_2.

Let us now write the:

Proof of Theorem 7. Under the assumption of uniform well-posedness, it is clear that $\mathcal{D}_{n+1}$ is the only element in $\mathcal{P}$ which satisfies

$$\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) = \inf\{\mathcal{E}(\mathcal{D}_n, \mathcal{D}), \quad \mathcal{D} \in \mathcal{P}\}.$$

Thus

$$\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) \leq \mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n-1})$$

and

$$\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) \leq \mathcal{E}(\mathcal{D}_n, \mathcal{D}_n).$$

As $\mathcal{E}$ is symmetric, one easily gets

$$\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) \leq \mathcal{E}(\mathcal{D}_{n-1}, \mathcal{D}_n) \leq \cdots \leq \mathcal{E}(\mathcal{D}_0, \mathcal{D}_1) \leq \mathcal{E}(\mathcal{D}_0, \mathcal{D}_0) = 2\,\mathcal{E}^{HF}(\mathcal{D}_0).$$

Furthermore, the functional $\mathcal{E}$ is bounded from below by $-2\,N\,Z^2$. Indeed on the one hand $\mathrm{Tr}\,(\mathcal{G}(\mathcal{D}) \cdot \mathcal{D}') \geq 0$ and on the other hand

$$
\begin{aligned}
\mathrm{Tr}\,(h\mathcal{D}) &= \sum_{i=1}^{N}\left(\int_{\mathbb{R}^3} |\nabla\phi_i|^2 - \sum_{k=1}^{M} z_k \int_{\mathbb{R}^3} \frac{|\phi_i|^2(x)}{|x - \bar{x}_k|}\,\mathrm{d}x\right)\\
&\geq \sum_{i=1}^{N}\left(\int_{\mathbb{R}^3} |\nabla\phi_i|^2 - \sum_{k=1}^{M} z_k \left(\int_{\mathbb{R}^3} \frac{|\phi_i|^2(x)}{|x - \bar{x}_k|^2}\,\mathrm{d}x\right)^{1/2}\right)\\
&\geq \sum_{i=1}^{N}\left(\int_{\mathbb{R}^3} |\nabla\phi_i|^2 - 2\,Z\left(\int_{\mathbb{R}^3} |\nabla\phi_i|^2\right)^{1/2}\right)\\
&\geq \sum_{i=1}^{N}(\|\nabla\phi_i\|_{L^2} - Z)^2 - N\,Z^2\\
&\geq -N\,Z^2,
\end{aligned}
$$

for any $\mathcal{D} \in \mathcal{P}$, $\Phi = \{\phi_i\} \in \mathcal{W}$ being such that $\mathcal{D}_\Phi = \mathcal{D}$. Therefore the sequence $\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1})$ is decreasing towards some $\lambda \in \mathbb{R}$.

Let us now consider a sequence $(\Phi_n) = (\{\phi_i^n\})$ in $\mathcal{W}$ such that

- for any $n \in \mathbb{N}$, $\mathcal{D}_{\Phi_n} = \mathcal{D}_n$;
- for any $n \in \mathbb{N}$,

$$\mathcal{F}(\mathcal{D}_n)\phi_i^{n+1} = -\epsilon_i^{n+1}\phi_i^{n+1}, \qquad 1 \leq i \leq N,$$

with $-\epsilon_1^{n+1} \leq -\epsilon_2^{n+1} \leq \cdots \leq -\epsilon_N^{n+1}$.

Putting together the results obtained so far, we deduce

$$\sum_{i=1}^{N}(\|\nabla\phi_i^n\|_{L^2} - Z)^2 + \sum_{i=1}^{N}(\|\nabla\phi_i^{n+1}\|_{L^2} - Z)^2 - 2\,N\,Z^2 \leq \mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) \leq 2\,\mathcal{E}^{HF}(\mathcal{D}_0).$$

Consequently, for any $1 \leq i \leq N$, the sequence $(\phi_i^n)_{n \in \mathbb{N}}$ is bounded in $H^1(\mathbb{R}^3)$. Besides, the sequence $(-\epsilon_i^n)_{n \in \mathbb{N}}$ is bounded in $\mathbb{R}$ for any $1 \leq i \leq N$. Indeed, it is bounded from below by the fundamental eigenvalue of $-\Delta + V$

because of the positivity of $\mathcal{G}(\mathcal{D}_n)$ and it is bounded from above by $-\epsilon < 0$. Then, it is possible to extract from the sequences (ϕ_i^n) and (ϵ_i^n) subsequences which satisfy for any $1 \leq i \leq N$:

$$\phi_i^{n_k-1} \xrightarrow[k \to +\infty]{} \chi_i \qquad \phi_i^{n_k} \xrightarrow[k \to +\infty]{} \phi_i \qquad \phi_i^{n_k+1} \xrightarrow[k \to +\infty]{} \psi_i \qquad \phi_i^{n_k+2} \xrightarrow[k \to +\infty]{} \phi_i' \qquad (13)$$

$$-\epsilon_i^{n_k} \xrightarrow[k \to +\infty]{} \mu_i \leq -\epsilon \qquad -\epsilon_i^{n_k+1} \xrightarrow[k \to +\infty]{} \nu_i \leq -\epsilon \qquad -\epsilon_i^{n_k+2} \xrightarrow[k \to +\infty]{} \mu_i' \leq -\epsilon,$$

the convergences in (13) being fulfilled both in $H^1(\mathbb{R}^3)$ weak and in $L_{loc}^2(\mathbb{R}^3)$. In passing to the limit $k \to +\infty$ in $H^{-1}(\mathbb{R}^3)$ in the equations

$$\mathcal{F}(\mathcal{D}_{n_k-1})\phi_i^{n_k} = -\epsilon_i^{n_k}\phi_i^{n_k}, \qquad \mathcal{F}(\mathcal{D}_{n_k})\phi_i^{n_k+1} = -\epsilon_i^{n_k+1}\phi_i^{n_k+1},$$

$$\mathcal{F}(\mathcal{D}_{n_k+1})\phi_i^{n_k+2} = -\epsilon_i^{n_k+2}\phi_i^{n_k+2},$$

one obtains for all $1 \leq i \leq N$

$$-\Delta\phi_i + V\phi_i + \left(\sum_{i=1}^{N}|\chi_i|^2 \star \frac{1}{|x|}\right)\phi_i - \int_{\mathbb{R}^3} \frac{\sum_{i=1}^{N}\chi_i(.)\chi_i^*(y)}{|\cdot - y|}\phi(y)\,\mathrm{d}y = \mu_i\phi_i,$$

$$-\Delta\psi_i + V\psi_i + \left(\sum_{i=1}^{N}|\phi_i|^2 \star \frac{1}{|x|}\right)\psi_i - \int_{\mathbb{R}^3} \frac{\sum_{i=1}^{N}\phi_i(.)\phi_i^*(y)}{|\cdot - y|}\psi_i(y)\,\mathrm{d}y = \nu_i\psi_i,$$

$$-\Delta\phi_i' + V\phi_i' + \left(\sum_{i=1}^{N}|\psi_i|^2 \star \frac{1}{|x|}\right)\phi_i' - \int_{\mathbb{R}^3} \frac{\sum_{i=1}^{N}\psi_i(.)\psi_i^*(y)}{|\cdot - y|}\phi'(y)\,\mathrm{d}y = \mu_i'\phi_i'.$$

For passing to the limit in the terms involving convolutions, the following lemma has been applied (its proof is postponed until the end of the present section):

Lemma 10. *Let $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ be bounded sequences in $H^1(\mathbb{R}^3)$ which converge towards f and g respectively in $L_{loc}^2(\mathbb{R}^3)$. Let $(h_n)_{n\in\mathbb{N}}$ be a sequence in $L^2(\mathbb{R}^3)$ which weakly converge towards h and let $\phi \in L^2(\mathbb{R}^3)$. Then*

$$D(f_ng_n, h_n\phi) \xrightarrow[n \to +\infty]{} D(fg, h\phi).$$

Besides

$$\mu_i \int_{\mathbb{R}^3} |\phi_i|^2 = \int_{\mathbb{R}^3} |\nabla\phi_i|^2 + \int_{\mathbb{R}^3} V|\phi_i|^2 + D\left(\sum_{j=1}^{N} |\chi_j|^2, |\phi_i|^2\right) - \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\sum_{j=1}^{N} \chi_j(x)\chi_j^*(y)}{|x-y|} \phi_i(x)^*\phi_i(y)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_{\mathbb{R}^3} |\nabla\phi_i|^2 + \int_{\mathbb{R}^3} V|\phi_i|^2 + \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \sum_{j=1}^{N} \frac{|\phi_i(x)\chi_j(y) - \phi_i(y)\chi_j(x)|^2}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y$$

$$\leq \liminf_{k\to+\infty} \left(\int_{\mathbb{R}^3} |\nabla\phi_i^{n_k}|^2 + \int_{\mathbb{R}^3} V|\phi_i^{n_k}|^2 + \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \sum_{j=1}^{N} \frac{|\phi_i^{n_k}(x)\phi_j^{n_k-1}(y) - \phi_i^{n_k}(y)\phi_j^{n_k-1}(x)|^2}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y\right)$$

$$= \liminf_{k\to+\infty} \left(\int_{\mathbb{R}^3} |\nabla\phi_i^{n_k}|^2 + \int_{\mathbb{R}^3} V|\phi_i^{n_k}|^2 + D(\rho_{\mathcal{D}_{n_k-1}}, |\phi_i^{n_k}|^2)\right.$$

$$\left. - \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\tau_{\mathcal{D}_{n_k-1}}(x,y)}{|x-y|} \phi_i^{n_k}(x)^*\phi_i^{n_k}(y)\,\mathrm{d}x\,\mathrm{d}y\right)$$

$$= \liminf_{k\to+\infty} \left(-\epsilon_i^{n_k} \int_{\mathbb{R}^3} |\phi_i^{n_k}|^2\right) = \liminf_{k\to+\infty} (-\epsilon_i^{n_k}) = \mu_i.$$

Reasoning as in [13], we deduce from $\mu_i \leq -\epsilon < 0$ that $\int_{\mathbb{R}^3} |\phi_i|^2 \geq 1$, and therefore that $\int_{\mathbb{R}^3} |\phi_i|^2 = 1$ since

$$\int_{\mathbb{R}^3} |\phi_i|^2 \leq \liminf_{k\to+\infty} \int_{\mathbb{R}^3} |\phi_i^{n_k}|^2 = 1.$$

Thus, the N sequences $(\phi_i^{n_k})_{k\in\mathbb{N}}$ converge strongly in $L^2(\mathbb{R}^3)$. In particular $\Phi = \{\phi_i\} \in \mathcal{W}$. Besides, as $(\phi_i^{n_k-1})_{k\in\mathbb{N}}$ is bounded in $H^1(\mathbb{R}^3)$ and converges in $L^2_{\mathrm{loc}}(\mathbb{R}^3)$, and as $(\phi_i^{n_k})_{k\in\mathbb{N}}$ is bounded in $H^1(\mathbb{R}^3)$ and converges in $L^2(\mathbb{R}^3)$, we leave the reader check that for any $1 \leq i \leq N$

$$\int_{\mathbb{R}^3} V|\phi_i^{n_k}|^2 \xrightarrow[k\to+\infty]{} \int_{\mathbb{R}^3} V|\phi_i|^2$$

and

$$D\left(\sum_{j=1}^{N} |\phi_j^{n_k-1}|^2, |\phi_i^{n_k}|^2\right) - \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\sum_{j=1}^{N} \phi_j^{n_k-1}(x)\phi_j^{n_k-1}(y)^*}{|x-y|} \phi_i^{n_k}(x)^*\phi_i^{n_k}(y)\,\mathrm{d}x\,\mathrm{d}y$$

$$\xrightarrow[k\to+\infty]{} D\left(\sum_{j=1}^{N} |\chi_j|^2, |\phi_i|^2\right) - \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\sum_{j=1}^{N} \chi_j(x)\chi_j(y)^*}{|x-y|} \phi_i(x)^*\phi_i(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Therefore

$$
\begin{aligned}
\lim_{k \to +\infty} \int_{\mathbb{R}^3} |\nabla \phi_i^{n_k}|^2 &= \lim_{k \to +\infty} \left(-\epsilon_i^{n_k} - \int_{\mathbb{R}^3} V |\phi_i^{n_k}|^2 - D \left(\sum_{j=1}^N |\phi_j^{n_k-1}|^2, |\phi_i^{n_k}|^2 \right) \right. \\
&\qquad \left. + \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \sum_{j=1}^N \frac{\phi_j^{n_k-1}(x) \phi_j^{n_k-1}(y)^*}{|x-y|} \phi_i^{n_k}(x)^* \phi_i^{n_k}(y) \, \mathrm{d}x \, \mathrm{d}y \right) \\
&= \mu_i - \int_{\mathbb{R}^3} V|\phi_i|^2 - D\left(\sum_{j=1}^N |\chi_j|^2, |\phi_i|^2 \right) + \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \sum_{j=1}^N \frac{\chi_j(x) \chi_j(y)^*}{|x-y|} \phi_i(x)^* \phi_i(y) \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_{\mathbb{R}^3} |\nabla \phi_i|^2,
\end{aligned}
$$

which proves the convergence of $(\phi_i^{n_k})_{k \in \mathbb{N}}$ in $H^1(\mathbb{R}^3)$. In the same way, $(\phi_i^{n_k+1})_{k \in \mathbb{N}}$ and $(\phi_i^{n_k+2})_{k \in \mathbb{N}}$ converge towards ψ_i and ϕ_i' in $H^1(\mathbb{R}^3)$. In particular, $\Phi = \{\phi_i\}$, $\Psi = \{\psi_i\}$ and $\Phi' = \{\phi_i'\}$ are in $\mathcal{W}$ and $(\mathcal{D}_{n_k})$, $(\mathcal{D}_{n_k+1})$ and $(\mathcal{D}_{n_k+2})$ converge towards $\mathcal{D}_\Phi$, $\mathcal{D}_\Psi$ and $\mathcal{D}_{\Phi'}$ respectively in $(\mathcal{P}, d_1)$. Consequently,

$$
\mathcal{E}(\mathcal{D}_\Phi, \mathcal{D}_\Psi) = \mathcal{E}(\mathcal{D}_{\Phi'}, \mathcal{D}_\Psi) = \lambda.
$$

But we have also

$$
\begin{aligned}
\lambda &= \lim_{k \to +\infty} \mathcal{E}(\mathcal{D}_{n_k+1}, \mathcal{D}_{n_k+2}) \\
&= \lim_{k \to +\infty} [\inf \{ \mathcal{E}(\mathcal{D}_{n_k+1}, \mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \}] \\
&\leq \inf \{ \mathcal{E}(\mathcal{D}_\Psi, \mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \},
\end{aligned}
$$

since the for any $\mathcal{D}' \in \mathcal{P}$, the function $\mathcal{D} \mapsto \mathcal{E}(\mathcal{D}, \mathcal{D}')$ is continuous from $(\mathcal{P}, d_1)$ into $\mathbb{R}$. Finally,

$$
\mathcal{E}(\mathcal{D}_\Phi, \mathcal{D}_\Psi) = \mathcal{E}(\mathcal{D}_{\Phi'}, \mathcal{D}_\Psi) = \inf \{ \mathcal{E}(\mathcal{D}_\Psi, \mathcal{D}), \quad \mathcal{D} \in \mathcal{P} \}.
$$

Therefore $\mathcal{D}_\Phi = \mathcal{D}_{\Phi'}$, since in view of the assumption of uniform well-posedness and by continuity there is a gap at least equal to γ between the Nth smallest eigenvalue of $\mathcal{F}(\mathcal{D}_\Psi)$ and the part of the spectrum above this eigenvalue. Let us detail this point. On the one hand,

$$
\inf_{\substack{V \subset H^1(\mathbb{R}^3) \\ \dim V = N}} \sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} (v, \mathcal{F}(\mathcal{D}_\Psi)v) \leq \mu_N' = -\lim_{k \to +\infty} \epsilon_N^{n_k+2},
$$

and on the other hand, for any $V \subset H^1(\mathbb{R}^3)$ with $\dim V = N + 1$,

$$
\sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} (v, \mathcal{F}(\mathcal{D}_{n_k+1})v) \geq -\epsilon_N^{n_k+2} + \gamma.
$$

As V is finite dimensional, it follows that at the limit

$$
\sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} (v, \mathcal{F}(\mathcal{D}_\Psi)v) \geq \mu_N' + \gamma,
$$

and therefore

$$\inf_{\substack{V \subset H^1(\mathbb{R}^3) \\ \dim V = N+1}} \sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} (v, \mathcal{F}(\mathcal{D}_\Psi)v) \geq \inf_{\substack{V \subset H^1(\mathbb{R}^3) \\ \dim V = N}} \sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} (v, \mathcal{F}(\mathcal{D}_\Psi)v) + \gamma.$$

Finally, as $\mathcal{D}_{\Phi'} = \mathcal{D}_\Phi$, we obtain

$$-\Delta \psi_i + V\psi_i + (\rho_{\mathcal{D}_{\Phi'}} \star \frac{1}{|x|})\psi_i - \int_{\mathbb{R}^3} \frac{\tau_{\mathcal{D}_\Phi}(\cdot, y)}{|\cdot - y|} \psi_i(y)\, \mathrm{d}y = \nu_i \psi_i,$$

$$-\Delta \phi_i' + V\phi_i' + (\rho_{\mathcal{D}_\Psi} \star \frac{1}{|x|})\phi_i' - \int_{\mathbb{R}^3} \frac{\tau_{\mathcal{D}_\Psi}(\cdot, y)}{|\cdot - y|} \phi'(y)\, \mathrm{d}y = \mu_i' \phi_i',$$

and consequently $(\mathcal{D}_\Phi, \mathcal{D}_\Psi)$ is a critical point of the functional $\mathcal{E}$ associated with the stationary value λ. The convergence of the series $\sum_{n=0}^{+\infty} \|\mathcal{D}_{n+2} - \mathcal{D}_n\|_0^2$ is obtained by noticing that in view of the uniform well-posedness assumption,

$$\mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_{n+1})\mathcal{D}_n) \geq \mathrm{Tr}\, (\mathcal{F}(\mathcal{D}_{n+1})\mathcal{D}_{n+2}) + \gamma\|\mathcal{D}_{n+2} - \mathcal{D}_n\|_0^2.$$

Adding $\mathrm{Tr}\, (h\mathcal{D}_{n+1})$ to both sides leads to

$$\mathcal{E}(\mathcal{D}_n, \mathcal{D}_{n+1}) \geq \mathcal{E}(\mathcal{D}_{n+1}, \mathcal{D}_{n+2}) + \gamma\|\mathcal{D}_{n+2} - \mathcal{D}_n\|_0^2$$

for any $n \in \mathbb{N}$. We obtain the expected result by taking the sum for all $n \in \mathbb{N}$ of the above inequalities.

We conclude this section with the:

Proof of Lemma 10. Firstly, $(f_n g_n \star \frac{1}{|x|})$ is bounded in $L^\infty(\mathbb{R}^3)$ from Cauchy-Schwarz and Hardy inequalities. Furthermore, for any $x \in \mathbb{R}^3$ and any $R > 0$

$$\left| \int_{\mathbb{R}^3} \frac{f_n(y)g_n(y)}{|x - y|}\, \mathrm{d}y - \int_{\mathbb{R}^3} \frac{f(y)g(y)}{|x - y|}\, \mathrm{d}y \right| \leq \frac{1}{R} \int_{|x-y| \geq R} |f_n(y)g_n(y) - f(y)g(y)|\, \mathrm{d}y$$

$$+ \int_{|x-y| < R} \frac{|f_n(y)|}{|x - y|} |g_n(y) - g(y)|\, \mathrm{d}y$$

$$+ \int_{|x-y| < R} \frac{|g(y)|}{|x - y|} |f_n(y) - f(y)|\, \mathrm{d}y$$

$$\leq \frac{2}{R} \sup_{n \in \mathbb{N}} \|f_n\|_{L^2(\mathbb{R}^3)} \sup_{n \in \mathbb{N}} \|g_n\|_{L^2(\mathbb{R}^3)}$$

$$+ 2 \sup_{n \in \mathbb{N}} \|\nabla f_n\|_{L^2(\mathbb{R}^3)} \|g_n - g\|_{L^2(B_R(x))}$$

$$+ 2\|\nabla g\|_{L^2(\mathbb{R}^3)} \|f_n - f\|_{L^2(B_R(x))}$$

with $B_R(x) = \{y \in \mathbb{R}^3, \quad |x - y| < R\}$. Letting R, then n, go to infinity, it follows that $(f_n g_n \star \frac{1}{|x|}) \to (fg \star \frac{1}{|x|})$ pointwise. Next,

$$|D(f_n g_n, h_n \phi) - D(fg, h\phi)| = \left| \int_{\mathbb{R}^3} (f_n g_n \star \frac{1}{|x|})h_n \phi - \int_{\mathbb{R}^3} (fg \star \frac{1}{|x|})h\phi \right|$$

$$\leq \left| \int_{\mathbb{R}^3} \left[(f_n g_n \star \frac{1}{|x|})\phi - (fg \star \frac{1}{|x|})\phi \right] h_n \right| + \left| \int_{\mathbb{R}^3} (fg \star \frac{1}{|x|})\phi(h_n - h) \right|.$$

From Lebesgue convergence Theorem, $(f_n g_n \star \frac{1}{|x|})\phi$ converges towards $(fg \star \frac{1}{|x|})\phi$ in $L^2(\mathbb{R}^3)$. As (h_n) is bounded in $L^2(\mathbb{R}^3)$, the first term of the right hand side of the above inequality converges towards zero when n goes to infinity. And so does the second term since $(fg \star \frac{1}{|x|})\phi$ belongs to $L^2(\mathbb{R}^3)$ and (h_n) converges to h weakly in $L^2(\mathbb{R}^3)$.   □

4. CONVERGENCE OF THE LEVEL-SHIFTING ALGORITHM

As far as we know, convergence properties of the level-shifting algorithm have only been investigated through perturbation calculations [21]. In such studies, it is assumed in particular that the initial guess is close to a stationary state (only the local behavior is investigated) and that the variation of the density operator is small between two consecutive iterations, which is not the case if, for instance, the algorithm oscillates. The scope of studies such as [21] therefore seem to us rather limited. The following theorem is a proof of the global convergence of the level-shifting algorithm for large enough shift parameters.

Theorem 11. *For any initial guess $\mathcal{D}_0$, there exists a real positive constant b_0 such that for any level-shift parameter $b > b_0$:*

1. *the level-shifting algorithm with initial guess $\mathcal{D}_0$ is uniformly well-posed;*
2. *the sequence of the energies $\mathcal{E}^{HF}(\mathcal{D}_n^b)$ decreases towards some stationary value λ of the Hartree-Fock functional $\mathcal{E}^{HF}$;*
3. *the sequence $\{\mathcal{D}_n^b\}_{0 \le n \le +\infty}$ converges up to an extraction towards a stationary state of the Hartree-Fock functional in $(\mathcal{P}, d_1)$. In addition,*

$$\sum_{n=0}^{+\infty} \|\mathcal{D}_{n+1}^b - \mathcal{D}_n^b\|_0^2 < +\infty.$$

Remark 12. Theorem 11 shows in particular, that for large enough level-shift parameters, the level-shifting algorithm can also be considered as a direct energy minimization technique. This property has already been pointed out in [21]. Let us also notice that the parameter b_0 depends on the initial guess $\mathcal{D}_0$. Putting together the results obtained below, we can prove an estimate on the size of b_0, namely $b_0 \le N\alpha^2 + 4N\alpha - \lambda_1(-\Delta + V)$, where $\alpha = \sqrt{N}[1 + Z + 2(NZ^2 + \mathcal{E}^{HF}(\mathcal{D}_0))^{1/2}]$. This estimate is not sharp and there is probably room for improvement in it. This upper bound can be computed *a priori* before starting the algorithm and updated at each step by replacing $\mathcal{D}_0$ by $\mathcal{D}_n$ in the definition of α (it is an increasing function of the Hartree-Fock energy which decreases at each step).   □

Remark 13. For large enough shift parameters, the level-shifting algorithm always provides a solution to the Hartree-Fock equations, even if the Hartree-Fock *minimization* problem has no solution, which occurs for instance for negative ions such that $N > 2Z + M$ [11]. Of course, this solution is a critical point which is generically either a local (non global) minimum or a saddle point. The case of a local maximum is excluded for the level-shifting algorithm makes the Hartree-Fock energy decrease. As far as we know, SCF algorithms had not been used so far as means to establish constructive proofs of solutions to the Hartree-Fock equations.

□

The following lemma is useful for proving Theorem 11.

Lemma 14. *Let $\alpha > 0$ and $\gamma > 0$. There exists $b_0 > 0$ such that for any $b \ge b_0$:*

1. *if $\mathcal{D} \in \mathcal{P}$ with $\|\mathcal{D}\|_1 \le \alpha$, then $\mathcal{F}(\mathcal{D}) - b\mathcal{D}$ has at least N negative eigenvalues and there is a gap at least equal to γ between the N-th smallest eigenvalue $-\epsilon_N$ (including multiplicity) and the part of the spectrum above this eigenvalue;*
2. *if $\mathcal{D} \in \mathcal{P}$ and $\mathcal{D}' \in \mathcal{P}$, with $\|\mathcal{D}\|_1 \le \alpha$ and $\|\mathcal{D}'\|_1 \le \alpha$, then*

$$\text{Tr}\left((\mathcal{G}(\mathcal{D}) - \mathcal{G}(\mathcal{D}')) \cdot (\mathcal{D} - \mathcal{D}')\right) \le b\|\mathcal{D} - \mathcal{D}'\|_0^2.$$

Proof. The operator $\mathcal{D}$ being finite rank, we have the equality of the essential spectra

$$\sigma_{\text{ess}}(\mathcal{F}(\mathcal{D}) - b\mathcal{D}) = \sigma_{\text{ess}}(\mathcal{F}(\mathcal{D})) = [0, +\infty[.$$

Let $\Phi = \{\phi_i\} \in \mathcal{W}$ such that $\mathcal{D} = \mathcal{D}_\Phi$. For any $1 \leq i \leq N$, $\|\nabla\phi_i\|_{L^2(\mathbb{R}^3)} \leq \|\mathcal{D}\|_1 \leq \alpha$. Therefore, using Cauchy-Schwarz and Hardy inequalities, we obtain that for any $v \in \text{Span}(\phi_1, \cdots, \phi_N)$ such that $\|v\|_{L^2(\mathbb{R}^3)} = 1$,

$$(v, (\mathcal{F}(\mathcal{D}) - b\mathcal{D})v) = (v, \mathcal{F}(\mathcal{D}) \cdot v) - b$$

$$\leq \int_{\mathbb{R}^3} |\nabla v|^2 + \frac{1}{2}\sum_{i=1}^N \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{|\phi_i(x)v(y) - v(x)\phi_i(y)|^2}{|x-y|} \, dx\, dy - b$$

$$\leq N\alpha^2 + 4N\alpha - b.$$

Therefore,

$$-\epsilon_N = \inf_{\substack{V \subset H^1(\mathbb{R}^3) \\ \dim V = N}} \ \sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} \ (v, (\mathcal{F}(\mathcal{D}) - b\mathcal{D})v)$$

$$\leq \sup_{\substack{v \in \text{Span}(\phi_1, \cdots, \phi_N) \\ \|v\|_{L^2} = 1}} \ (v, (\mathcal{F}(\mathcal{D}) - b\mathcal{D})v)$$

$$\leq N\alpha^2 + 4N\alpha - b,$$

which implies that for $b > N\alpha^2 + 4N\alpha$, $\mathcal{F}(\mathcal{D}) - b\mathcal{D}$ has at least N negative eigenvalues. Moreover, let us denote by

$$-\epsilon_{N+1} = \inf_{\substack{V \subset H^1(\mathbb{R}^3) \\ \dim V = N+1}} \ \sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} \ (v, (\mathcal{F}(\mathcal{D}) - b\mathcal{D})v).$$

For any $V \subset H^1(\mathbb{R}^3)$ such that $\dim V = N+1$, there exists $\psi \in V$ such that $\psi \in \text{Span}(\phi_1, \cdots, \phi_N)^\perp$ and $\|\psi\|_{L^2(\mathbb{R}^3)} = 1$. Therefore

$$\sup_{\substack{v \in V \\ \|v\|_{L^2} = 1}} \ (v, (\mathcal{F}(\mathcal{D}) - b\mathcal{D})v) \geq (\psi, (\mathcal{F}(\mathcal{D}) - b\mathcal{D})\psi) = (\psi, \mathcal{F}(\mathcal{D})\psi) \geq \lambda_1(-\Delta + V).$$

Thus $-\epsilon_{N+1} \geq \lambda_1(-\Delta + V)$. Therefore, for $b \geq N\alpha^2 + 4N\alpha - \lambda_1(-\Delta + V) + \gamma$, there is a gap at least equal to γ between the Nth smallest eigenvalue $-\epsilon_N$ and the part of the spectrum above $-\epsilon_N$. Statement 1 is therefore proved. For establishing statement 2, let us consider $\mathcal{D}$ and $\mathcal{D}'$ in $\mathcal{P}$ with $\|\mathcal{D}\|_1 \leq \alpha$, $\|\mathcal{D}'\|_1 \leq \alpha$, and Φ and Φ' in $\mathcal{W}$ such that $\mathcal{D} = \mathcal{D}_\Phi$, $\mathcal{D}' = \mathcal{D}_{\Phi'}$. Firstly,

$$\text{Tr}\left((\mathcal{G}(\mathcal{D}) - \mathcal{G}(\mathcal{D}')) \cdot (\mathcal{D} - \mathcal{D}'))\right) = D(\rho_\mathcal{D} - \rho_{\mathcal{D}'}, \rho_\mathcal{D} - \rho_{\mathcal{D}'}) - \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{|(\tau_\mathcal{D} - \tau_{\mathcal{D}'})(x,y)|^2}{|x-y|} \, dx\, dy$$

$$\leq D(\rho_\mathcal{D} - \rho_{\mathcal{D}'}, \rho_\mathcal{D} - \rho_{\mathcal{D}'}).$$

Secondly, let us write ϕ_i' as

$$\phi_i' = \sum_{j=1}^N (\phi_j, \phi_i') \, \phi_j + \psi_i,$$

with $\psi_i \in \mathrm{Span}(\phi_1, \cdots, \phi_N)^\perp$. Let $A_{jk} = \sum_{i=1}^N (\phi_j, \phi_i')(\phi_i', \phi_k)$. The matrix $[A_{jk}]_{1 \le j,k \le N}$ being hermitian, one can find a unitary U and a diagonal real matrix Δ such that $A = U^* \Delta U$. Denoting by $\tilde{\Phi} = U\Phi$, straightforward calculations lead to

$$(\rho_\mathcal{D} - \rho_{\mathcal{D}'})(x) = \sum_{i=1}^N (1 - \Delta_{ii})|\tilde{\phi}_i|^2(x) - 2\mathrm{Re}\left(\sum_{i,j=1}^N (\phi_i', \phi_j)\phi_j^*(x)\psi_i(x) \right) - \sum_{i=1}^N |\psi_i|^2(x),$$

$$0 \le \Delta_{ii} \le 1, \qquad \|\mathcal{D} - \mathcal{D}'\|_0^2 = 2\sum_{i=1}^N (1 - \Delta_{ii}) = 2\sum_{i=1}^N \|\psi_i\|_{L^2(\mathbb{R}^3)}^2.$$

Therefore, we deduce from Cauchy-Schwarz and Hardy inequalities that

$$D(\rho_\mathcal{D} - \rho_{\mathcal{D}'}, \rho_\mathcal{D} - \rho_{\mathcal{D}'}) \le (8N^2 + 4N)\alpha \|\mathcal{D} - \mathcal{D}'\|_0^2,$$

which concludes the proof of Lemma 14.   □

Proof of Theorem 11. Let $\alpha = \sqrt{N}[1 + Z + 2(NZ^2 + \mathcal{E}^{HF}(\mathcal{D}_0))^{1/2}]$ (which is well defined since $\mathcal{E}^{HF}(\mathcal{D}_0) \le -NZ^2$ by Cauchy-Schwarz and Hardy inequalities), b_0 such as defined in Lemma 14, and $b > b_0$. Let us consider the functional

$$\mathcal{E}^b(\mathcal{D}, \mathcal{D}') = \mathrm{Tr}\,(h\mathcal{D}) + \mathrm{Tr}\,(h\mathcal{D}') + \mathrm{Tr}\,(\mathcal{G}(\mathcal{D}) \cdot \mathcal{D}') + b\|\mathcal{D} - \mathcal{D}'\|_0^2.$$

The constant α is such that

$$\mathcal{E}^b(\mathcal{D}, \mathcal{D}') \le 2\mathcal{E}^{HF}(\mathcal{D}_0) \quad \Rightarrow \quad \|\mathcal{D}\|_1 \le \alpha \quad \|\mathcal{D}'\|_1 \le \alpha.$$

Therefore, the initial guess $\mathcal{D}_0$ satisfies $\|\mathcal{D}_0\|_1 \le \alpha$. Let us now assume that at step n, the induction hypotheses $\|\mathcal{D}_n^b\|_1 \le \alpha$ and $\mathcal{E}^{HF}(\mathcal{D}_n^b) \le \mathcal{E}^{HF}(\mathcal{D}_0)$ are fulfilled at step n. From Lemma 14, $\mathcal{F}(\mathcal{D}_n^b) - b\mathcal{D}_n^b$ has at least N negative eigenvalues and there is a gap between the N-th smallest eigenvalue and the part of the spectrum above this eigenvalue. Thus $\mathcal{D}_{n+1}^b$ is uniquely defined by the *aufbau* principle and is such that

$$\mathcal{E}^b(\mathcal{D}_{n+1}^b, \mathcal{D}_n^b) = \inf\left\{ \mathcal{E}^b(\mathcal{D}, \mathcal{D}_n^b), \quad \mathcal{D} \in \mathcal{P} \right\} \le \mathcal{E}^b(\mathcal{D}_n^b, \mathcal{D}_n^b) = 2\mathcal{E}^{HF}(\mathcal{D}_n^b) \le 2\mathcal{E}^{HF}(\mathcal{D}_0).$$

Therefore, we have in particular $\|\mathcal{D}_{n+1}^b\|_1 \le \alpha$. Moreover, the assertion

$$\mathcal{E}^b(\mathcal{D}_{n+1}^b, \mathcal{D}_n^b) \le \mathcal{E}^b(\mathcal{D}_n^b, \mathcal{D}_n^b),$$

is equivalent to

$$\mathcal{E}^{HF}(\mathcal{D}_{n+1}^b) - \frac{1}{2}\mathrm{Tr}\,\left((\mathcal{G}(\mathcal{D}_{n+1}^b) - \mathcal{G}(\mathcal{D}_n^b)) \cdot (\mathcal{D}_{n+1}^b - \mathcal{D}_n^b) \right) + b\|\mathcal{D}_{n+1}^b - \mathcal{D}_n^b\|_0^2 \le \mathcal{E}^{HF}(\mathcal{D}_n^b).$$

Therefore, using statement 2 of Lemma 14, one obtains

$$\mathcal{E}^{HF}(\mathcal{D}_{n+1}^b) + \frac{b}{2}\|\mathcal{D}_{n+1}^b - \mathcal{D}_n^b\|_0^2 \le \mathcal{E}^{HF}(\mathcal{D}_n^b). \tag{14}$$

In particular $\mathcal{E}^{HF}(\mathcal{D}_{n+1}^b) \le \mathcal{E}^{HF}(\mathcal{D}_n^b) \le \mathcal{E}^{HF}(\mathcal{D}_0)$ and the induction goes on. At this stage, we have proved that for $b > b_0$ (a) the level-shifting algorithm (LS^b) with initial guess $\mathcal{D}_0$ is uniformly well posed, (b) the

Hartree-Fock energy is a Lyapunov functional for this algorithm, and (c) using (14),

$$\sum_{n=0}^{+\infty} \|\mathcal{D}_{n+1}^b - \mathcal{D}_n^b\|_0^2 < +\infty.$$

As the Hartree-Fock functional is bounded from below, one has in addition

$$\lambda := \lim_{n \to +\infty} \mathcal{E}^{HF}(\mathcal{D}_n) \in \mathbb{R}.$$

Mimicking the proof of the convergence up to an extraction detailed in the proof of Theorem 7, it is easy to prove that λ is a stationary value of the functional $\mathcal{E}^b$ and that there exist extracted sequences $(\mathcal{D}_{n_k})_{k \in \mathbb{N}}$ and $(\mathcal{D}_{n_k+1})_{k \in \mathbb{N}}$ such that

$$\mathcal{D}_{n_k} \underset{k \to +\infty}{\longrightarrow} \mathcal{D}, \qquad \mathcal{D}_{n_k+1} \underset{k \to +\infty}{\longrightarrow} \mathcal{D}',$$

in $(\mathcal{P}, d_1)$, $(\mathcal{D}, \mathcal{D}')$ being a critical point of the functional $\mathcal{E}^b$. Furthermore, $\mathcal{D} = \mathcal{D}'$ since

$$\lim_{n \to +\infty} \|\mathcal{D}_{n+1}^b - \mathcal{D}_n^b\|_0 = 0,$$

and $\mathcal{D} = \mathcal{D}'$ is a critical point of the Hartree-Fock functional associated with the stationary value λ. This concludes the proof of Theorem 11. $\square$

REFERENCES

[1] G. Auchmuty and Wenyao Jia, Convergent iterative methods for the Hartree eigenproblem. *RAIRO Modél. Math. Anal. Numér.* **28** (1994) 575–610.

[2] V. Bach, E.H. Lieb, M. Loss and J.P. Solovej, There are no unfilled shells in unrestricted Hartree-Fock theory. *Phys. Rev. Lett.* **72** (1994) 2981–2983.

[3] V. Bonačić-Koutecký and J. Koutecký, General properties of the Hartree-Fock problem demonstrated on the frontier orbital model. II. Analysis of the customary iterative procedure. *Theoret. Chim. Acta* **36** (1975) 163–180.

[4] J.C. Facelli and R.H. Contreras, A general relation between the intrinsic convergence properties of SCF Hartree-Fock calculations and the stability conditions of their solutions. *J. Chem. Phys.* **79** (1983) 3421–3423.

[5] R. Fletcher, Optimization of SCF LCAO wave functions. *Mol. Phys.* **19** (1970) 55–63.

[6] D.R. Hartree, *The calculation of atomic structures.* Wiley (1957).

[7] W.J. Hehre, L. Radom, P.V.R. Schleyer and J.A. Pople, *Ab initio molecular orbital theory.* Wiley (1986).

[8] A. Igawa and H. Fukutome, A new direct minimization algorithm for Hartree-Fock calculations. *Progr. Theoret. Phys.* **54** (1975) 1266–1281.

[9] J. Koutecký and V. Bonačić, On convergence difficulties in the iterative Hartree-Fock procedure. *J. Chem. Phys.* **55** (1971) 2408–2413.

[10] E.H. Lieb, Existence and uniqueness of the minimizing solution of Choquard's nonlinear equation. *Stud. Appl. Math.* **57** (1977) 93–105.

[11] E.H. Lieb, Bound on the maximum negative ionization of atoms and molecules. *Phys. Rev. A* **29** (1984) 3018–3028.

[12] E.H. Lieb and B. Simon, The Hartree-Fock theory for Coulomb systems. *Comm. Math. Phys.* **53** (1977) 185–194.

[13] P.L. Lions, Solutions of Hartree-Fock equations for Coulomb systems. *Comm. Math. Phys.* **109** (1987) 33–97.

[14] R. McWeeny, The density matrix in self-consistent field theory. I. Iterative construction of the density matrix. *Proc. Roy. Soc. London Ser. A* **235** (1956) 496–509.

[15] R. McWeeny, *Methods of molecular Quantum Mechanics.* Academic Press (1992).

[16] J. Paldus, Hartree-Fock stability and symmetry breaking, in *Self Consistent Field Theory and Application.* Elsevier (1990) 1–45.

[17] P. Pulay, Improved SCF convergence acceleration. *J. Comput. Chem.* **3** (1982) 556–560.

[18] M. Reed and B. Simon, *Methods of modern mathematical physics. I. Functional analysis.* Academic Press (1980).

[19] M. Reed and B. Simon, *Methods of modern mathematical physics. IV. Analysis of operators.* Academic Press (1978).

[20] C.C.J. Roothaan, New developments in molecular orbital theory. *Rev. Modern Phys.* **23** (1951) 69–89.

[21] V.R. Saunders and I.H. Hillier, A "level-shifting" method for converging closed shell Hartree-Fock wave functions. *Int. J. Quantum Chem.* **7** (1973) 699–705.

[22] H.B. Schlegel and J.J.W. McDouall, Do you have SCF stability and convergence problems?, in *Computational Advances in Organic Chemistry*, Kluwer Academic (1991) 167–185.

[23] R. Seeger R. and J.A. Pople, Self-consistent molecular orbital methods. XVI. Numerically stable direct energy minimization procedures for solution of Hartree-Fock equations. *J. Chem. Phys.* **65** (1976) 265–271.

[24] R.E. Stanton, The existence and cure of intrinsic divergence in closed shell SCF calculations. *J. Chem. Phys.* **75** (1981) 3426–3432.

[25] R.E. Stanton, Intrinsic convergence in closed-shell SCF calculations. A general criterion. *J. Chem. Phys.* **75** (1981) 5416–5422.

[26] M.C. Zerner and M. Hehenberger, A dynamical damping scheme for converging molecular SCF calculations. *Chem. Phys. Lett.* **62** (1979) 550–554.