SOLUTIONS OF THE MULTICONFIGURATION DIRAC-FOCK EQUATIONS

ANTOINE LEVITT

ABSTRACT. The multiconfiguration Dirac-Fock (MCDF) model uses a linear combination of Slater determinants to approximate the electronic N-body wave function of a relativistic molecular system, resulting in a coupled system of nonlinear eigenvalue equations, the MCDF equations. In this paper, we prove the existence of solutions of these equations in the weakly relativistic regime. First, using a new variational principle as well as results of Lewin on the multiconfiguration nonrelativistic model, and Esteban and Séré on the single-configuration relativistic model, we prove the existence of critical points for the associated energy functional, under the constraint that the occupation numbers are not too small. Then, this constraint can be removed in the weakly relativistic regime, and we obtain non-constrained critical points, i.e. solutions of the multiconfiguration Dirac-Fock equations.

CONTENTS

1. INTRODUCTION

Consider an atom or molecule with N electrons. Nonrelativistic quantum mechanics dictates that, under the Born-Oppenheimer approximation, the electronic rest energy is given by the lowest fermionic eigenvalue of the N-body Hamiltonian. The complexity of this problem grows exponentially with N, and approximations are used to keep the

Date: June 15, 2013.

2010 *Mathematics Subject Classification.* Primary: 35Q40; Secondary: 49S05, 81V55.

Key words and phrases. multiconfiguration, relativistic quantum mechanics, variational methods.

Support from the grant ANR-10-BLAN-0101 of the French Ministry of Research is gratefully acknowledged.

problem tractable. Hartree-Fock theory uses the variational ansatz that the N-body wavefunction is a single Slater determinant. The optimization of the resulting energy over the orbitals gives rise to a nonlinear eigenvalue problem, which is solved iteratively.

It is well-known that this method overestimates the true ground state energy by a quantity known as the correlation energy, whose size can be significant in many cases of chemical interest [SO89]. This can be remedied by considering several Slater determinants, a technique known as multiconfiguration Hartree-Fock (MCHF) theory. This brings the model closer to the full N-body problem, and, in the limit of an infinite number of determinants, one recovers the true ground state energy.

Another source of errors is that the Hamiltonian used is non-relativistic. Indeed, in large atoms, the core electrons reach relativistic speeds (in atomic units, of the order of Z, compared with the speed of light $c \approx 137$). This causes a length contraction which affects the screening by the core electrons of the attractive potential of the nucleus. This has important consequences for the valence electrons and the chemistry of elements. Neglecting these effects leads to incorrect conclusions, and for instance fails to account for the difference in color between silver and gold [PD79].

For a fully relativistic treatment of the electrons, one should use quantum electrodynamics (QED). But this very precise theory is also extremely complex for all but the simplest systems. Therefore, physicists and chemists use approximate Hamiltonians to avoid working in the full Fock space of QED. The multiconfiguration Dirac-Fock (MCDF) model is obtained by using the Dirac operator in the multiconfiguration Hartree-Fock model. It incorporates relativistic effects into the multiconfiguration Hartree-Fock model, and has been used successfully in a number of applications [DFJ07, Gra07].

Although these models, and more complicated ones, are used routinely by physicists, many problems still remain in their mathematical analysis. The first rigorous proof of existence of ground states of the Hartree-Fock equations was given by Lieb and Simon [LS77] and later generalized to excited states by Lions [Lio87]. The multiconfiguration equations were studied by Le Bris [LB94], who proved existence in the particular case of doubly excited states. Friesecke later proved the existence of minimizers for an arbitrary number of determinants [Fri03a], and Lewin generalized his proof to excited states, in the spirit of the method of Lions [Lew04]. For relativistic models, Esteban and Séré proved existence of single-configuration solutions to the Dirac-Fock equations [ES99], and studied their non-relativistic limit [ES01]. To our knowledge, the present work is the first mathematical study of a relativistic multiconfiguration model.

The main mathematical difficulty of the multiconfiguration equations, apart from the increased algebraic complexity, is that one cannot simultaneously diagonalize the Fock operator and the matrix of Lagrange multipliers. Lewin rewrote the Euler-Lagrange equations in a vector formalism and used the same arguments as in the Hartree-Fock case [LS77, Lio87] to prove the existence of solutions.

The Dirac-Fock equations are considerably more difficult to handle than the Hartree-Fock equations. The main difficulty is that the Dirac operator is not bounded from below. This fact, which causes important problems already in the linear theory, complicates the search for solutions of the equations, because every critical point has an infinite Morse index. One can therefore no longer minimize the energy functional, or even use standard critical point theory. Esteban and Séré [ES99], later generalized by Buffoni, Esteban and Séré [BES06], used the concavity of the energy with respect to the negative directions of the free Dirac operator to reduce the problem to one whose critical points have a finite Morse index.

The MCDF model combines the two mathematical problems and adds the difficulty that, for the theory to make sense, the speed of light has to be above a constant that depends on a lower bound on the occupation numbers. Note that this difficulty with small occupation numbers is also encountered in numerical computations [ID93], and theoretical studies of the nonrelativistic evolution problem [BCMT10].

In this paper, we prove the existence of solutions, when the speed of light is large enough (weakly relativistic regime). We now describe our formalism.

2. Definitions

In atomic units, the Dirac operator is given by

$$D_c = -ic(\alpha \cdot \nabla) + c^2\beta. \tag{1}$$

In standard representation, α and β are 4×4 matrices given by

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}, \beta_k = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix},$$

where the σ_k are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The speed of light c has the physical value $c \approx 137$.

The operator D_c is self-adjoint on $L^2(\mathbb{R}^3, \mathbb{C}^4)$ with domain $H^1(\mathbb{R}^3, \mathbb{C}^4)$ and form domain $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$. It verifies the relativistic identity $D_c^2 = c^4 - c^2\Delta$. More precisely, it admits the spectral decomposition

$$D_c = P^+\sqrt{c^4 - c^2\Delta}\, P^+ - P^-\sqrt{c^4 - c^2\Delta}\, P^-, \tag{2}$$

where the projectors $P^\pm$ are given in the Fourier domain by

$$P^\pm(\xi) = \frac{1}{2}\left(1_{\mathbb{C}^4} + \pm\frac{c\alpha \cdot \xi + c^2\beta}{\sqrt{c^4 + c^2\xi^2}}\right). \tag{3}$$

We denote by

$$E = H^{1/2}(\mathbb{R}^3, \mathbb{C}^4) \tag{4}$$

the form-domain of D_c, and $E^\pm = P^\pm E$ the two positive and negative spectral subspaces.

We will use three scalar products in this paper:

$$\begin{aligned}
\langle \psi, \phi \rangle_{L^2} &= \int_{\mathbb{R}^3} \psi^* \phi, \\
\langle \psi, \phi \rangle_E &= \left\langle \psi, \sqrt{1-\Delta}\phi \right\rangle_{L^2}, \\
\langle \psi, \phi \rangle_c &= \left\langle \psi, \sqrt{1 - \frac{\Delta}{c^2}}\phi \right\rangle_{L^2},
\end{aligned}$$

with associated norms $\|\psi\|_{L^2}, \|\psi\|_E, \|\psi\|_c$. The purpose of this last norm is to simplify several estimates. It is related to the change of variables $d_c(\psi)(x) = c^{-3/2}\psi(\frac{x}{c})$ used in [ES01] in the sense that

$$\langle \Psi, \Phi \rangle_c = \langle d_c(\Psi), d_c(\Phi) \rangle_E$$

A molecule made of M nuclei with positions z_i and charges Z_i creates an attractive potential

$$V(x) = -\sum_{i=1}^{M} \frac{Z_i}{|x - z_i|}.$$

More generally, we consider a charge distribution $\mu \geq 0$ with $\mu(\mathbb{R}^3) = Z$, which creates a potential

$$V = -\mu \star \frac{1}{|x|}. \tag{5}$$

In the sequel, we shall always assume that $N < Z + 1$, which is the only case where we can prove existence of solutions to our equations. This assumption is made in existence proofs for the Hartree-Fock model to ensure that an electron cannot "escape to infinity", because it will then feel the effective attractive potential $\frac{(N-1)-Z}{|x|}$ [Lio87, LS77]. Mathematically, it is used to prove that second order information on Palais-Smale sequences implies that the Lagrange multipliers are not in the essential spectrum.

The Hamiltonian $D_c + V$ has a spectral gap around zero as long as

$$Z < \frac{2}{\pi/2 + 2/\pi} c.$$

This is related to the following Hardy-type inequality (see [Tix98, Her77, Kat66]) :

$$|\langle \psi, V\psi \rangle| \leq \frac{Z}{2} (\pi/2 + 2/\pi) \left\langle \psi, \sqrt{1 - \Delta} \psi \right\rangle \tag{6}$$

for all $\psi \in E^{\pm}$, a refinement of the Kato inequality

$$|\langle \psi, V\psi \rangle| \leq \frac{Z\pi}{2} \left\langle \psi, \sqrt{-\Delta} \psi \right\rangle \tag{7}$$

for all $\psi \in E$, which we will use extensively in this paper. We also recall the standard Hardy inequality:

$$\|V\phi\|_{L^2} \leq 2Z \|\nabla \phi\|_{L^2} \tag{8}$$

for all $\phi \in H^1$.

The N-body relativistic Hamiltonian is given by

$$H^N = \sum_{i=1}^{N} (D_{c,x_i} + V(x_i)) + \sum_{1 \leq i < j \leq N} \frac{1}{\left| x_i - x_j \right|}.$$

This Hamiltonian acts on $\bigwedge^N L^2_a(\mathbb{R}^3, \mathbb{C}^4)$, the fermionic N-body space. Its interpretation is problematic : in particular, its essential spectrum is all of $\mathbb{R}$, and it is not even known whether eigenvectors exists [Der12].

For a given $K \geq N$, the multiconfiguration ansatz is

$$\psi = \sum_{1 \leq i_1 < \cdots < i_N \leq K} a_{i_1, \ldots, i_N} \left| \psi_{i_1} \, \ldots \, \psi_{i_N} \right\rangle, \tag{9}$$

where

$$\left| \psi_{i_1} \, \ldots \, \psi_{i_N} \right\rangle (X_1, \ldots, X_N) = \frac{1}{\sqrt{N!}} \det(\psi_{i_k}(X_l))_{k,l}$$

with $X_l = (x_l, s_l), x_l \in \mathbb{R}^3, s_l \in \{1, 2, 3, 4\}$ are Slater determinants, and $a \in S, \Psi \in \Sigma$, where

$$S = \{a \in \mathbb{C}^{\binom{K}{N}}, \|a\|^2 = \sum_{1 \le i_1 < \dots < i_N \le K} \left| a_{i_1,\dots,i_N} \right|^2 = 1\}, \tag{10}$$

$$\begin{aligned} \Sigma &= \{\Psi \in E^K, \mathrm{Gram}\,\Psi = 1\}, \\ &= \{\Psi \in E^K, \left\langle \psi_i, \psi_j \right\rangle_{L^2} = \delta_{ij}\}. \end{aligned} \tag{11}$$

Our convention here and in the rest of this paper is to use lower case greek letters for orbitals $\psi \in E$, and upper case greek letters for vectors of orbitals $\Psi \in E^K$. We extend in a straightforward way the scalar products $\langle \cdot, \cdot \rangle_{L^2}, \langle \cdot, \cdot \rangle_E$ and $\langle \cdot, \cdot \rangle_c$ to the space E^K:

$$\langle \Psi, \Phi \rangle_* = \sum_{k=1}^{K} \langle \psi_k, \phi_k \rangle_* .$$

Following [Lew04], we define

$$\alpha_{i_1 \dots i_N} = \begin{cases} 0 & \text{if } \#(i_1 \dots i_N) < N, \\ \frac{\epsilon(\sigma)}{\sqrt{N!}} a_{i_{\sigma(1)},\dots,i_{\sigma(N)}} & \text{otherwise,} \end{cases}$$

where, for all $i_1, \dots, i_N$ with $\#(i_1 \dots i_N) = N$, σ is the unique permutation such that $i_{\sigma(1)} < \dots < i_{\sigma(N)}$.

With this definition,

$$\psi(X_1, \dots, X_N) = \sum_{1 \le i_1 \le N, \dots, 1 \le i_N \le N,} \alpha_{i_1,\dots,i_N} \psi_{i_1}(X_1) \dots \psi_{i_N}(X_N).$$

Then, substituting into the relativistic energy $\left\langle \psi, H^N \psi \right\rangle$, we obtain [Lew04]

$$\mathcal{E}(a, \Psi) = \left\langle \Psi, \left(D_c \Gamma_a + V \Gamma_a + W_{a,\Psi} \right) \Psi \right\rangle_{(L^2(\mathbb{R}^3, \mathbb{C}^4))^K}, \tag{12}$$

with the $K \times K$ Hermitian matrices

$$(\Gamma_a)_{i,j} = N \sum_{k_2 \dots k_N} \alpha^*_{i,k_2 \dots k_N} \alpha_{j,k_2 \dots k_N},$$

$$(W_{a,\Psi})_{i,j} = \frac{N(N-1)}{2} \sum_{k_3 \dots k_N} \sum_{k,l} \alpha^*_{i,k,k_3 \dots k_N} \alpha_{j,l,k_3 \dots k_N} \left(\psi_k^* \psi_l \star \frac{1}{|x|} \right).$$

The eigenvalues γ_i of Γ_a, for $a \in S$, satisfy $0 \le \gamma_i \le 1$, and are called occupation numbers. They measure the total weight of the corresponding orbital in the N-body wave function.

For reference, we define similarly the multiconfiguration Hartree-Fock energy

$$\mathcal{E}^{\mathrm{HF}}(a, \Phi) = \left\langle \Phi, \left(-\frac{1}{2} \Delta \, \Gamma_a + V \Gamma_a + W_{a,\Phi} \right) \Phi \right\rangle_{(L^2(\mathbb{R}^3, \mathbb{C}^2))^K}, \tag{13}$$

on $S \times \{\Phi \in (H^1(\mathbb{R}^3, \mathbb{C}^2))^K, \mathrm{Gram}\,\Phi = 1\}$.

One can define a group action on $S \times \Sigma$ that leaves $\mathcal{E}$ invariant : for any unitary matrix $U \in \mathcal{U}(K)$,

$$U \cdot (a, \Psi) = (a', U\Psi), \tag{14}$$

where a' is defined via the equivalent variables α' :

$$\alpha'_{i_1,\dots,i_N} = \sum_{j_1,\dots,j_N} (U^*)_{i_1,j_1} \dots (U^*)_{i_N,j_N} \alpha_{j_1,\dots,j_N},$$

where U^* is the adjoint of U. This group action is the multiconfiguration analogue of the well-known unitary invariance of the Hartree-Fock equations.

The MCDF equations, obtained as the Euler-Lagrange equations of $\mathcal{E}$ under the constraints $a \in S$ and $\Psi \in \Sigma$, are, for Ψ and a respectively,

$$H_{a,\Psi}\Psi = \Lambda\Psi, \tag{15}$$

$$\mathcal{H}_\Psi a = Ea, \tag{16}$$

where

$$H_{a,\Psi} = D_c\Gamma_a + V\Gamma_a + 2W_{a,\Psi} \tag{17}$$

is the Fock operator, and

$$(\mathcal{H}_\Psi)_{I,J} = \left\langle \psi_{i_1} \, \ldots \, \psi_{i_N} \left| H^N \right| \psi_{j_1} \, \ldots \, \psi_{j_N} \right\rangle \tag{18}$$

are the coefficients of the $\binom{K}{N} \times \binom{K}{N}$ matrix of the N-body Hamiltonian H^N in the basis of the Slater determinants. Our goal in this paper is to prove the existence of solutions to (15) and (16) by finding critical points of $\mathcal{E}$ on $S \times \Sigma$.

3. Strategy of proof

There are several major mathematical difficulties in the study of the MCDF model. Unlike in the single-configuration case, one can use the group action (14) to diagonalize Γ or Λ, but not both at the same time. Worse, because $W_{a,\Psi}$ does not in general commute with Γ, one can only prove that the Fock operator $H_{a,\Psi}$ has a spectral gap around 0 for values of c that depend on a lower bound on the eigenvalues of Γ. This gap is used centrally to prove the convergence of Palais-Smale sequences. Therefore, one needs a lower bound on Γ.

To obtain this lower bound, we consider the (formal) nonrelativistic limit of the multiconfiguration Dirac-Fock model, the multiconfiguration Hartree-Fock model. Let

$$I^K = \inf \left\{ \mathcal{E}^{\mathrm{HF}}(a, \Phi), a \in S, \Phi \in (H^1(\mathbb{R}^3, \mathbb{C}^2))^K, \operatorname{Gram} \Phi = 1 \right\} \tag{19}$$

be the ground-state energy of the nonrelativistic multiconfiguration method of rank $K \geq N$. I^N is the Hartree-Fock energy. I^K is non-increasing, and converges to I^∞, the Schrödinger energy. The behavior of I^K is not precisely known, but a result by Friesecke [Fri03b] shows that $I^{K+2} < I^K$. Therefore, $I^K < I^{K-1}$ at least for one every two K. When this strict inequality holds, every minimizer satisfies $\Gamma > 0$ in the sense of Hermitian matrices. Because of the compactness of these minimizers (implicitly proved in [Lew04]), there is a uniform bound $\gamma_0 > 0$ such that for every minimizer, $\Gamma_a \geq \gamma_0$ in the sense of Hermitian matrices.

Because there is no well-defined "ground state energy" in the relativistic case, we cannot use information of this type directly. Instead, we fix $\gamma < \gamma_0$, and use a min-max principle to look for solutions in the domain

$$S_\gamma = \{a \in S, \Gamma_a \geq \gamma\}.$$

By arguments inspired by [ES99, ES01, Lew04], we prove that the min-max principle yields solutions of $H_{a,\Psi}\Psi = \Lambda\Psi$, for c large enough. But these are only solutions of the equation $\mathcal{H}_\Psi a = Ea$ if the constraint is not saturated, *i.e.* if $\Gamma_a > \gamma$.

To prove that this is the case, we take the nonrelativistic ($c \to \infty$) limit of the critical points found in the first step. By arguments similar to the ones in [ES01], we prove that these critical points converge, up to a subsequence, to a minimizer of the Hartree-Fock functional. Therefore, for c large, the constraint $\Gamma_a \geq \gamma$ is not saturated, and we obtain solutions of the MCDF equations.

In the rest of this paper, we will always assume that $I^K < I^{K-1}$, so that $\Gamma \geq \gamma_0$ on the nonrelativistic minimizers. $\gamma > 0$ is a fixed constant, taken to be less than γ_0. We also assume $N < Z + 1$.

First, for all $\Psi \in (L^2)^K$ such that $\mathrm{Gram}\,\Psi > 0$, we define the normalization

$$g(\Psi) = (\mathrm{Gram}\,\Psi)^{-1/2}\Psi. \tag{20}$$

This normalization was used in [ES01] to prove another variational principle for the relativistic “ground state”, which we shall not use here.

Define

$$\begin{aligned}\Sigma^+ &= \Sigma \cap (E^+)^K,\\ &= \left\{\Psi \in (E^+)^K, \mathrm{Gram}\,\Psi = 1\right\}.\end{aligned}$$

We will find solutions to our equations as a result of the following variational principle:

$$I_{c,\gamma} = \inf_{a\in S_\gamma, \Psi^+\in\Sigma^+} \sup_{\Psi^-\in(E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-)). \tag{21}$$

4. Results

Our first result is the well-posedness of our variational principle:

Theorem 1. *There are constants $K_1, K_2 > 0$ such that, for c large enough, there is a triplet $a_* \in S_\gamma, \Psi_*^+ \in \Sigma^+, \Psi_*^- \in (E^-)^K$ solution of the variational principle (21):*

$$\begin{aligned}\mathcal{E}(a_*, g(\Psi_*^+ + \Psi_*^-)) &= \max_{\Psi^-\in(E^-)^K} \mathcal{E}(a_*, g(\Psi_*^+ + \Psi^-)),\\ &= \min_{a\in S_\gamma,\Psi^+\in\Sigma^+} \max_{\Psi^-\in(E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-)).\end{aligned}$$

Denoting $\Psi_* = g(\Psi_*^+ + \Psi_*^-)$, Ψ_* is a solution of the equation $H_{a_*,\Psi_*}\Psi_* = \Lambda_*\Psi_*$ in Σ.

The Hermitian matrix of Lagrange multipliers Λ_* satisfies the estimates

$$(c^2 - K_1)\Gamma_* \leq \Lambda_* \leq (c^2 - K_2)\Gamma_*. \tag{22}$$

Furthermore, if $\Gamma_* > \gamma$, then a_* is a solution of $\mathcal{H}_{\Psi_*} a_* = I_{c,\gamma} a_*$.

We now study the nonrelativistic limit of these solutions, thanks to the control (22) on the Lagrange multipliers:

Theorem 2. *Let $c_n \to \infty$, and let (a_n, Ψ_n) be the solution of (21) obtained by Theorem 1 with $c = c_n$. Then, up to a subsequence,*

$$\begin{aligned}a_n &\to a,\\ \Psi_n &\to \begin{pmatrix}\Phi\\ 0\end{pmatrix}\end{aligned}$$

in H^1 norm, where $(a, \Phi) \in S_\gamma \times (H^1(\mathbb{R}^3, \mathbb{C}^2))^K$ is a minimizer of

$$I^K = \inf\left\{\mathcal{E}^{HF}(a, \Phi), a \in S, \Phi \in (H^1(\mathbb{R}^3, \mathbb{C}^2))^K, \mathrm{Gram}\,\Phi = 1\right\}. \tag{23}$$

The min-max level $I_{c,\gamma}$ satisfies the asymptotics

$$I_{c,\gamma} = Nc^2 + I^K + o_{c\to\infty}(1).$$

Since any minimizer of (23) must satisfy $\Gamma \geq \gamma_0 > \gamma$, we immediately obtain

Corollary 1. *For c large enough, there are solutions of the multiconfiguration Dirac-Fock equations* (15)-(16)*.*

The remainder of this paper is dedicated to the proof of Theorems 1 and 2.

For Theorem 1, we first begin with Lemma 1, a convergence result for Palais-Smale sequences of the functional $\mathcal{E}$ with Lagrange multipliers bounded away from the essential spectrum of $D_c\Gamma$. Then, at $(a, \Psi^+) \in S_\gamma \times \Sigma^+$ fixed, we study the variational principle

$$\sup_{\Psi^- \in (E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-))$$

in Lemma 2, under the condition that $\mathcal{E}(a, \Psi^+) < Nc^2$. We prove in Lemma 6 an upper bound on the asymptotic behavior of $I_{c,\gamma}$ which will enable us to restrict to this domain, and finally, we prove in Lemma 7 that Palais-Smale sequences with Morse-type information for the functional

$$\mathcal{F}_a(\Psi^+) = \sup_{\Psi^- \in (E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-))$$

satisfy the hypotheses of Lemma 1, and therefore are precompact. Their limit up to extraction is a solution of our min-max problem (21).

To prove Theorem 2, we use the estimates (22) on the Lagrange multipliers to prove the compactness of the sequence (a_n, Ψ_n), and the asymptotic behavior from Lemma 7 to show that the limit is a minimizer.

5. Proof of Theorem 1

Our first result is the convergence of Palais-Smale sequences with bounds on the Lagrange multipliers. The proof proceeds as in Lemma 2.1 of [ES99] for the single-configuration case.

5.1. **Palais-Smale sequences for the energy functional.**

Lemma 1. *For c large enough, if $(a_n, \Psi_n) \in S_\gamma \times \Sigma$ satisfies:*

(i) $H_{a_n,\Psi_n}\Psi_n - \Lambda_n\Psi_n = \Delta_n \to 0$ *in* $H^{-1/2}$ *with* Λ_n *Hermitian matrices,*

(ii) $\liminf \Lambda_n > 0$,

(iii) $\limsup c^2\Gamma_n - \Lambda_n > 0$,

then, up to extraction, $(a_n, \Psi_n) \to (a, \Psi)$ *in* $S_\gamma \times \Sigma$*, where* (a, Ψ) *is a solution of* $H_{a,\Psi}\Psi = \Lambda\Psi$*.*

Proof.

Step 1 : convergence in $H^{1/2}_{\mathbf{loc}}$. Let $\Psi \in E^K$, and $\Psi^\pm = P^\pm\Psi$. Using the inequality (6),

$$\begin{aligned}\left\langle \Psi^+, H_{a_n,\Psi_n}\Psi^+ \right\rangle &\geq \left\langle \Psi^+, \Gamma_n\sqrt{c^4 - c^2\Delta}\Psi^+ \right\rangle + \left\langle \Psi^+, \Gamma_n V\Psi^+ \right\rangle, \\ &\geq \left\langle \Psi^+, \Gamma_n\sqrt{c^4 - c^2\Delta}\Psi^+ \right\rangle - C_1\left\|\Psi^+\right\|_E^2, \\ &\geq (\gamma c^2 - C_1 c)\left\|\Psi^+\right\|_c^2,\end{aligned}$$

where $C_1 > 0$. Similarly,

$$\left\langle \Psi^-, H_{a_n,\Psi_n}\Psi^- \right\rangle \leq -(\gamma c^2 - C_2 c)\left\|\Psi^-\right\|_c^2,$$

with $C_2 > 0$.

Now, Ψ^+ and Ψ^- are orthogonal for the c scalar product, so

$$\|\Psi\|_c^2 = \left\|\Psi^+\right\|_c^2 + \left\|\Psi^-\right\|_c^2 = \left\|\Psi^+ - \Psi^-\right\|_c^2.$$

Denoting by $\|\cdot\|_c^*$ the dual norm of $\|\cdot\|_c$,

$$\begin{aligned}\left\|H_{a_n,\Psi_n}\Psi\right\|_c^* &\geq \frac{1}{\|\Psi\|_c}\left\langle \Psi^+ - \Psi^-, H_{a_n,\Psi_n}\Psi\right\rangle,\\ &= \frac{1}{\|\Psi\|_c}\left(\left\langle \Psi^+, H_{a_n,\Psi_n}\Psi^+\right\rangle - \left\langle \Psi^-, H_{a_n,\Psi_n}\Psi^-\right\rangle\right),\\ &\geq \frac{1}{\|\Psi\|_c}\left(c^2\gamma - c\max(C_1,C_2)\right)\left(\left\|\Psi^+\right\|_c^2 + \left\|\Psi^-\right\|_c^2\right),\\ &\geq h_0\|\Psi\|_c, \end{aligned} \tag{24}$$

with $h_0 > 0$ when c is large enough.

We then have

$$\begin{aligned}\limsup_{n\to\infty}\|\Psi_n\|_c &\leq \limsup_{n\to\infty}\frac{1}{h_0}\left\|H_{a_n,\Psi_n}\Psi_n\right\|_c^*,\\ &\leq \limsup_{n\to\infty}\frac{1}{h_0}\left(\|\Delta_n\|_c^* + \|\Lambda_n\Psi_n\|_{L^2}\right).\end{aligned}$$

Therefore, Ψ_n is bounded in c norm, *i.e.* in $H^{1/2}$. Extracting a subsequence, again denoted by (a_n,Ψ_n), we may assume that $a_n \to a$, $\Gamma_n \to \Gamma$, $\Lambda_n \to \Lambda$, and $\Psi_n \to \Psi$ weakly in $H^{1/2}$, strongly in L^p_{loc}, $2 \leq p < 3$.

Since H_{a,Ψ_n} is self-adjoint from E^K to $(E^K)^*$ and bounded away from zero, it is invertible. Define Ψ'_n by

$$H_{a,\Psi_n}\Psi'_n = \Lambda\Psi_n.$$

Ψ'_n is bounded in $H^{1/2}$, and therefore precompact in L^p_{loc}, $2 \leq p < 3$.

We partially invert

$$\Psi'_n = (D_c\Gamma + V\Gamma)^{-1}(\Lambda\Psi_n - 2W_{a,\Psi_n}\Psi'_n).$$

From Young's inequality, $W_{a,\Psi_n}\Psi'_n$ is precompact in L^p_{loc}, $1 \leq p < 3$, so $\Lambda\Psi_n - 2W_{a,\Psi_n}\Psi'_n$ is precompact in L^2_{loc}. Therefore, Ψ'_n is precompact in $H^{1/2}_{\text{loc}}$. We extract again and impose $\Psi'_n \to \Psi$ in $H^{1/2}_{\text{loc}}$. But since

$$H_{a,\Psi_n}(\Psi_n - \Psi'_n) \to 0$$

in $H^{-1/2}$, from (24), $\Psi_n \to \Psi$ in $H^{1/2}_{\text{loc}}$.

Step 2 : convergence in $H^{1/2}$. We now have convergence of Ψ_n to Ψ in $H^{1/2}_{\text{loc}}$. Ψ satisfies

$$H_{a,\Psi}\Psi = \Lambda\Psi.$$

We now look at the convergence in $H^{1/2}$ by obtaining an approximate Euler-Lagrange equation satisfied by the error $\varepsilon_n = \Psi_n - \Psi$. We have the Euler-Lagrange equations satisfied by Ψ_n and Ψ:

$$\begin{aligned}(D_c\Gamma + V\Gamma + 2W_{a,\Psi_n})\Psi_n - \Lambda\Psi_n &= \Delta'_n,\\ (D_c\Gamma + V\Gamma + 2W_{a,\Psi})\Psi - \Lambda\Psi &= 0.\end{aligned}$$

with $\Delta'_n \to 0$ in $H^{-1/2}$. Subtracting and using the fact that $\varepsilon^n \to 0$ weakly in $H^{1/2}$ and strongly in $H^{1/2}_{\text{loc}}$, we get

$$L_n\varepsilon_n \to 0 \tag{25}$$

in $H^{-1/2}$, where

$$L_n = D_c\Gamma + 2W_{a,\Psi_n} - \Lambda \tag{26}$$

is the Hamiltonian "at infinity" seen by ε_n.

We now use a concavity argument to extract information on the positive and negative components $\varepsilon_n^\pm = P^\pm \varepsilon_n$ of ε_n separately.

Define the quadratic functional Q_n on $(E^-)^K$ by

$$Q_n(\delta^-) = \left\langle \varepsilon_n + \delta^-, L_n(\varepsilon_n + \delta^-) \right\rangle .$$

The second order terms are

$$\begin{aligned}\left\langle \delta^-, L_n \delta^- \right\rangle &= \left\langle \delta^-, (D_c\Gamma + 2W_{a,\Psi_n} - \Lambda)\delta^- \right\rangle , \\ &\leq -(c^2\gamma - C_2 c)\left\|\delta^-\right\|_c^2 - \langle \delta_n, \Lambda\delta_n \rangle . \end{aligned} \tag{27}$$

Since $\Lambda > 0$, we obtain that Q_n is strictly concave for n large.

The concavity allows us to write

$$\begin{aligned}\left\langle \varepsilon_n^+, L_n\varepsilon_n^+ \right\rangle &= Q_n(-\varepsilon_n^-) \\ &\leq Q_n(0) - \nabla Q_n(0)[\varepsilon_n^-], \\ &= \langle \varepsilon_n, L_n\varepsilon_n \rangle - 2\left\langle \varepsilon_n^-, L_n\varepsilon_n \right\rangle , \\ &\leq 3\|\varepsilon_n\|_E \|L_n\varepsilon_n\|_{E^*} . \end{aligned}$$

Hence

$$\limsup_{n\to\infty} \left\langle \varepsilon_n^+, L_n\varepsilon_n^+ \right\rangle \leq 0.$$

But

$$\left\langle \varepsilon_n^+, L_n\varepsilon_n^+ \right\rangle \geq \left\langle \varepsilon_n^+, (c^2\Gamma - \Lambda)\varepsilon_n^+ \right\rangle$$

Since $\Lambda < c^2\Gamma$, this implies convergence to 0 of ε_n^+ in L^2 and then in $H^{1/2}$. But, by (25), this implies that $L_n\varepsilon^- \to 0$ in $H^{-1/2}$ and therefore that $\left\langle \varepsilon_n^-, L_n\varepsilon_n^- \right\rangle \to 0$. By (27), we deduce $\varepsilon_n^- \to 0$ in $H^{1/2}$, which proves that $\Psi_n \to \Psi$ strongly in Σ. □

5.2. **The reduced functional.** For $(a, \Psi^+) \in S_\gamma \times \Sigma^+$, define the functional

$$F_{a,\Psi^+}(\Psi^-) = \mathcal{E}(a, g(\Psi^+ + \Psi^-))$$

on $(E^-)^K$. Our goal in this section is to prove

Lemma 2. *There is a constant $M > 0$ such that, for c large enough, for all $(a, \Psi^+) \in S_\gamma \times \Sigma^+$ with $\mathcal{E}(a, \Psi^+) \leq Nc^2$, the functional F_{a,Ψ^+} has a unique maximizer $h(a, \Psi^+)$ in $(E^-)^K$. The map h is smooth, and satisfies*

$$\left\|h(a, \Psi^+)\right\|_c \leq \frac{M_-}{c}. \tag{28}$$

We first begin with estimates on Ψ^+, for which we use the property $\mathcal{E}(a, \Psi^+) \leq Nc^2$.

Lemma 3 (A priori bounds on Ψ^+)**.** *There are $M_+, M_D > 0$ such that, for c large, if $(a, \Psi^+) \in S_\gamma \times \Sigma^+$ verifies $\mathcal{E}(\Psi^+) \leq Nc^2$, then*

$$\left\|\Psi^+\right\|_E \leq M_+, \tag{29}$$

$$D_c|_{\mathrm{Span}(\{\psi_i^+\})} \leq c^2 + M_D. \tag{30}$$

Proof.

$$\mathcal{E}(\Psi^+) = \left\langle \Psi^+, H_{a,\Psi^+}\Psi^+ \right\rangle,$$
$$\geq \left\langle \Psi^+, D_c\Gamma\Psi^+ \right\rangle - C\left\langle \Psi^+, \sqrt{-\Delta}\Psi^+ \right\rangle. \tag{31}$$

Here and in the rest of this paper, C denotes various positive constants independent of c. Since $\mathcal{E}(a,\Psi^+) < Nc^2$ and $\left\langle \Psi^+, \Gamma\Psi^+ \right\rangle = N$,

$$\left\langle \Psi^+, \left(\sqrt{c^4 - c^2\Delta} - c^2 - \frac{C}{\gamma}\sqrt{-\Delta}\right)\Gamma\Psi^+ \right\rangle \leq 0.$$

In the Fourier domain, we can write for all $0 < \alpha < c^2$ by the Cauchy-Schwarz inequality

$$\sqrt{c^4 + c^2|\xi|^2} \geq c^2\left(1 - \frac{\alpha}{c^2}\right) + c|\xi|\sqrt{1 - \left(1 - \frac{\alpha}{c^2}\right)^2},$$
$$= c^2 - \alpha + |\xi|\sqrt{2\alpha - \frac{\alpha^2}{c^2}}.$$

Therefore, we obtain

$$\left\langle \Psi^+, \left(-\alpha + \left(\sqrt{\alpha - \frac{\alpha^2}{c^2}} - \frac{C}{\gamma}\right)\sqrt{-\Delta}\right)\Gamma\Psi^+ \right\rangle \leq 0,$$

so

$$\left\langle \Psi^+, \sqrt{-\Delta}\Gamma\Psi^+ \right\rangle \leq \frac{N\alpha}{\sqrt{\alpha - \frac{\alpha^2}{c^2}} - \frac{C}{\gamma}}.$$

Taking $\alpha > \sqrt{C/\gamma}$ and c large, $\left\langle \Psi^+, \sqrt{-\Delta}\Gamma\Psi^+ \right\rangle$ is bounded independently of c. Since $\Gamma \geq \gamma$, so is $\left\|\Psi^+\right\|_E$, and (29) is proved.

Now, using (31) again along with our new estimate (29), we have

$$\left\langle \Psi^+, D_c\Gamma\Psi^+ \right\rangle \leq Nc^2 + CM_+^2,$$
$$\left\langle \Psi^+, (D_c - c^2)\Gamma\Psi^+ \right\rangle \leq CM_+^2,$$
$$\operatorname{tr} A \leq CM_+^2,$$

where A is the $K \times K$ Hermitian matrix

$$A_{ij} = \Gamma_{ij}\left\langle \psi_i^+, (D_c - c^2)\psi_j^+ \right\rangle.$$

A is positive semi-definite and its trace is bounded by CM_+^2, so $A \leq CM_+^2$. Since $\Gamma \geq \gamma$, we conclude that

$$\left(\left\langle \psi_i^+, (D_c - c^2)\psi_j^+ \right\rangle\right)_{1\leq i,j\leq K} \leq \frac{CM_+^2}{\gamma},$$

and therefore (30) is proved. □

We now restrict our search for a maximizer to a neighborhood of zero.

Lemma 4 (A priori bounds on Ψ^-)**.** *There is a constant $M_- > 0$ such that, for c large enough, for all $(a,\Psi^+) \in S_\gamma \times \Sigma^+$ with $\mathcal{E}(a,\Psi^+) \leq Nc^2$,*

$$\sup_{\Psi^- \in (E^-)^K} F_{a,\Psi^+}(\Psi^-)$$

cannot be achieved outside a neighborhood of zero of size $\frac{M_-}{c}$ in the c norm.

Proof. Let $\Psi^- \in (E^-)^K$, $G = \mathrm{Gram}(\Psi^+ + \Psi^-)$, $\Psi = G^{-1/2}(\Psi^+ + \Psi^-)$. Using (30),

$$\begin{aligned}\langle \Psi, D_c\Gamma\Psi\rangle &= (c^2 + M_D)\langle \Psi, \Gamma\Psi\rangle + \left\langle \Psi, (D_c - c^2 - M_D)\Gamma\Psi\right\rangle,\\ &\leq N(c^2+M_D) + \left\langle G^{-1/2}\Psi^-, (D_c - c^2 - M_D)\Gamma G^{-1/2}\Psi^-\right\rangle,\\ &\leq N(c^2+M_D) - \gamma c^2 \left\| G^{-1/2}\Psi^- \right\|_c^2.\end{aligned}$$

On the other hand,

$$\begin{aligned}\left\langle \Psi, (V\Gamma + 2W_{a,\Psi})\Psi \right\rangle &\leq C\left\|G^{-1/2}\Psi^+\right\|_E^2 + C\left\|G^{-1/2}\Psi^-\right\|_E^2,\\ &\leq CM_+ + Cc\left\|G^{-1/2}\Psi^-\right\|_c^2.\end{aligned}$$

All together,

$$F_{a,\Psi^+}(\Psi^-) \leq Nc^2 + NM_D + CM_+ - \left(\gamma c^2 - Cc\right)\left\|G^{-1/2}\Psi^-\right\|_c^2.$$

But we also have

$$\begin{aligned}F_{a,\Psi^+}(0) &= \mathcal{E}(\Psi^+)\\ &\geq \left\langle \Psi^+, (D_c\Gamma + V\Gamma)\Psi^+\right\rangle,\\ &\geq Nc^2 - C\left\|\Psi^+\right\|_E^2,\\ &\geq Nc^2 - CM_+^2.\end{aligned}$$

Therefore,

$$F_{a,\Psi^+}(\Psi^-) \leq F_{a,\Psi^+}(0) + NM_D + 2CM_+^2 - \left(\gamma c^2 - Cc\right)\left\|G^{-1/2}\Psi^-\right\|_c^2.$$

So, in order to have $F_{a,\Psi^+}(\Psi^-) \leq F_{a,\Psi^+}(0)$, Ψ^- must satisfy

$$\begin{aligned}\left\|G^{-1/2}\Psi^-\right\|_c^2 &\leq O(1/c^2),\\ \left\|\Psi^-\right\|_c^2 &\leq O(1/c^2)(1 + \|\Psi_-\|_{L^2}^2),\\ &\leq O(1/c^2)(1 + \|\Psi_-\|_c^2),\end{aligned}$$

and therefore

$$\left\|\Psi^-\right\|_c^2 = O(1/c^2).$$

□

Restricting now to this domain, we prove that F_{a,Ψ^+} is strictly concave:

Lemma 5. *For c large, for all $(a, \Psi^+) \in S_\gamma \times \Sigma$ such that $\mathcal{E}(a, \Psi^+) \leq Nc^2$, for all Ψ^- in the region $\left\|\Psi^-\right\|_c \leq \frac{M_-}{c}$, for all $\Phi^- \in (E^-)^K$,*

$$F_{a,\Psi^+}(\Psi^-)''[\Phi^-, \Phi^-] \leq -\frac{\gamma c^2}{2}\left\|\Phi^-\right\|_c^2.$$

Proof. We have $g(\Psi^+ + \Psi^-) = (1 + \frac{1}{c}B(\Psi^-))(\Psi^+ + \Psi^-)$, where B is a matrix-valued function. In the region $\left\|\Psi^-\right\|_c \leq \frac{M_-}{c}$, B and its derivatives are bounded independently of c. Let Ψ^- be such that $\left\|\Psi^-\right\|_c \leq \frac{M_-}{c}$, and define $G = \mathrm{Gram}(\Psi^+ + \Psi^-)$. Then, for all $\Phi^- \in (E^-)^K$,

$$\begin{aligned}\frac{1}{2}F''_{a,\Psi^+}(\Psi^-)[\Phi^-,\Phi^-] &= \partial_\Psi \mathcal{E}(\Psi)\left[\frac{1}{c}B'(\Psi^-)[\Phi^-]\Phi^- + \frac{1}{2c}B''(\Psi^-)[\Phi^-,\Phi^-](\Psi^+ + \Psi^-)\right] \\ &\quad + \frac{1}{2}\partial^2_\Psi \mathcal{E}(\Psi)\left[G^{-1/2}\Phi^- + \frac{1}{c}B'(\Psi^-)[\Phi^-](\Psi^+ + \Psi^-)\right]^2, \\ &= \frac{1}{2}\partial^2_\Psi \mathcal{E}(\Psi)[\Phi^-,\Phi^-] + O\left(c\left\|\Phi^-\right\|_c^2\right).\end{aligned}$$

But we also have

$$\begin{aligned}\frac{1}{2}\partial^2_\Psi \mathcal{E}(\Psi)[\Phi^-,\Phi^-] &= -\left\langle \Phi^-, \sqrt{c^4 - c^2\Delta}\Gamma\Phi^-\right\rangle + O\left(\left\|\Phi^-\right\|_E^2\right), \\ &\leq \left(-\gamma c^2 + O(c)\right)\left\|\Phi^-\right\|_c^2.\end{aligned}$$

and so the result follows for c large.

□

Lemma 2 is now proved as a direct consequence of Lemmas 3, 4 and 5.

5.3. **Asymptotic behavior of $I_{c,\gamma}$.** In order to restrict to the domain $\mathcal{E}(a,\Psi^+) \leq Nc^2$, we prove that solutions of our min-max principle have to be in this domain for c large:

Lemma 6.

$$I_{c,\gamma} \leq Nc^2 + I^K + o_{c\to\infty}(1).$$

In particular, for c large enough,

$$\begin{aligned}I_{c,\gamma} &= \inf_{a\in S_\gamma, \Psi^+\in\Sigma^+} \ \sup_{\Psi^-\in(E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-)), \\ &= \inf_{\substack{a\in S_\gamma,\Psi^+\in\Sigma^+,\\ \mathcal{E}(a,\Psi^+)<Nc^2}} \ \sup_{\Psi^-\in(E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-)).\end{aligned} \tag{32}$$

Proof. Let $(a_*, \Phi_*) \in S_\gamma \times H^1(\mathbb{R}^3, \mathbb{C}^2)$ be a minimizer of the nonrelativistic multiconfiguration Hartree-Fock functional, and $\Psi_* = \begin{pmatrix}\Phi_*\\0\end{pmatrix}$. Set

$$\Psi_*^+ = g(P^+\Psi_*).$$

Ψ_*^+ belongs to Σ^+, and converges in H^1 to Ψ_* as $c \to \infty$. From the concavity inequality

$$\sqrt{1+|\xi|^2} \leq 1 + \frac{1}{2}|\xi|^2$$

in the Fourier domain, we get

$$\begin{aligned}\mathcal{E}(a_*, \Psi_*^+) &= \left\langle \Psi_*^+, (\sqrt{c^4 - c^2\Delta}\Gamma + V\Gamma + W_{a,\Psi_*^+})\Psi_*^+\right\rangle, \\ &\leq \left\langle \Psi_*^+, \left(c^2\Gamma - \frac{1}{2}\Delta\Gamma + V\Gamma + W_{a,\Psi_*^+}\right)\Psi_*^+\right\rangle, \\ &= Nc^2 + \mathcal{E}^{\mathrm{HF}}(a_*, \Psi_*) + o_{c\to\infty}(1), \\ &= Nc^2 + I^K + o_{c\to\infty}(1).\end{aligned}$$

From Lemmas 2 and 5, we now have

$$\begin{aligned} I_{c,\gamma} &\le F_{a_*,\Psi_*^+}(h(\Psi_*^+)), \\ &\le \mathcal{E}(a_*,\Psi_*^+) + F'_{a_*,\Psi_*^+}(0)[h(\Psi_*^+)] - \frac{\gamma c^2}{2}\left\|h(\Psi_*^+)\right\|_c^2, \\ &\le \mathcal{E}(a_*,\Psi_*^+) + C\left\|h(\Psi_*^+)\right\|_{L^2}, \\ &\le \mathcal{E}(a_*,\Psi_*^+) + O\left(\frac{1}{c}\right), \end{aligned}$$

so that

$$I_{c,\gamma} \le Nc^2 + I^K + o_{c\to\infty}(1).$$

Since $I^K < 0$ and, for all $a \in S_\gamma, \Psi^+ \in \Sigma$,

$$\sup_{\Psi^- \in (E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-)) \ge \mathcal{E}(a,\Psi^+),$$

(32) holds and the lemma is proved. □

5.4. **Borwein-Preiss sequences for the reduced functional.** Let

$$S'_\gamma = \{a \in S_\gamma, \inf_{\Psi^+\in\Sigma^+} \mathcal{E}(a,\Psi^+) < Nc^2\}.$$

For $a \in S'_\gamma$ fixed, we minimize the functional

$$\mathcal{F}_a(\Psi^+) = \mathcal{E}(a, g(\Psi^+ + h(a,\Psi^+)))$$

on the manifold

$$\Sigma_a^+ = \{\Psi^+ \in \Sigma^+, \mathcal{E}(a,\Psi^+) < Nc^2\}.$$

For all $\Psi^+ \in \Sigma_a^+, \Psi \in \Sigma$, define the tangent spaces

$$\begin{aligned} T_{\Psi^+}\Sigma_a^+ &= \{\Phi^+ \in (E^+)^K, \left\langle \phi_i^+, \psi_j^+ \right\rangle = 0 \text{ for all } i,j \in \{1,\dots,K\}\}, \\ T_\Psi \Sigma &= \{\Phi \in E^K, \left\langle \phi_i, \psi_j \right\rangle = 0 \text{ for all } i,j \in \{1,\dots,K\}\}. \end{aligned}$$

Lemma 7. *There are constants $K_1 > 0, K_2 > 0$ such that, for all c large enough, $a \in S'_\gamma$, if $\Psi_n^+ \in \Sigma_a^+$ is a Borwein-Preiss sequence for $\mathcal{F}_a$ on Σ_a^+, i.e. satisfies*

(i) $\mathcal{F}_a(\Psi_n^+) \to \inf_{\Psi^+\in\Sigma_a^+} \mathcal{F}_a(\Psi^+)$,

(ii) $\mathcal{F}'_a(\Psi_n^+)\Big|_{T_{\Psi_n^+}\Sigma_a^+} \to 0$ *in* $H^{-1/2}$,

(iii) There is a sequence $\beta_n \to 0$ such that the quadratic form $\Phi^+ \to \mathcal{F}''_a(\Psi_n^+)[\Phi^+,\Phi^+] + \beta_n\left\|\Phi^+\right\|_E^2$ is non-negative on $T_{\Psi_n^+}\Sigma_a^+$,

then, denoting $\Psi_n = g(\Psi_n^+ + h(a,\Psi_n^+))$,

(1) There is a sequence of Hermitian matrices Λ_n such that $H_{a,\Psi_n}\Psi_n - \Lambda_n\Psi_n = \Delta_n \to 0$ in $H^{-1/2}$,

(2) $\limsup \Lambda_n \le (c^2 - K_2)\Gamma$,

(3) $\liminf \Lambda_n \ge (c^2 - K_1)\Gamma$.

Proof. Define

$$k(\Psi^+,\Psi^-) = g(\Psi^+ + \Psi^-)$$

on $\Sigma_a^+ \times (E^-)^K$. From hypothesis *(ii)* and the definition of h, $(\Psi_n^+, h(\Psi_n^+))$ is a Palais-Smale sequence for $\mathcal{E}(a, k(\cdot))$. But $k'(\Psi_n^+, h(\Psi_n^+))$ is an isomorphism from $T_{\Psi_n^+}\Sigma_a^+ \times (E^-)^K$

to $T_{\Psi_n}\Sigma$, so that $\Psi_n = k(\Psi_n^+, h(\Psi_n^+))$ is a Palais-Smale sequence for $\mathcal{E}$ on Σ, and *(1)* is proved.

Upper bound on the Lagrange multipliers. Let us now prove the upper bound *(2)*. First, note that, since $a \in S'_\gamma$ and $\Psi_n^+ \in \Sigma_a^+$, the a priori estimates of Lemmas 2 and 3 hold.

Let $\delta_n \in T_{\Psi_n^+}$. Let $\Psi_n^- = h(\Psi_n^+)$, and $G_n = \mathrm{Gram}(\Psi_n^+ + \Psi_n^-) = 1 + \mathrm{Gram}\,\Psi_n^-$. For ε small enough, define the curve on Σ_a^+

$$\Psi_n^+(\varepsilon) = G_n^{1/2}\left(\frac{(G_n^{-1/2}\Psi_n^+)_1 + \varepsilon\delta_n}{\sqrt{1+\varepsilon^2}}, (G_n^{-1/2}\Psi_n^+)_2, \ldots, (G_n^{-1/2}\Psi_n^+)_K\right).$$

Define the associated

$$\begin{aligned} \Psi_n^-(\varepsilon) &= h(\Psi_n^+(\varepsilon)), \\ G_n(\varepsilon) &= \mathrm{Gram}(\Psi_n^+(\varepsilon) + \Psi_n^-(\varepsilon)) = 1 + \mathrm{Gram}\,\Psi_n^-(\varepsilon), \\ \Psi_n(\varepsilon) &= G_n^{-1/2}(\varepsilon)(\Psi_n^+(\varepsilon) + \Psi_n^-(\varepsilon)), \end{aligned}$$

and the infinitesimal increments

$$\begin{aligned} \Phi_n^+ &= \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Psi_n^+(\varepsilon)\right|_{\varepsilon=0}, \\ &= G_n^{1/2}(\delta_n, 0, \ldots, 0), \\ \Phi_n^- &= h'(\Psi_n)[\Phi_n^+]. \end{aligned}$$

Step 1. Our first step is a control on Φ_n^-.

Define

$$\begin{aligned} \mathcal{G}(\Psi^+, \Psi^-) &= F_{a,\Psi^+}(\Psi^-), \\ &= \mathcal{E}(a, g(\Psi^+ + \Psi^-)). \end{aligned}$$

Now, for all $\Psi^+ \in \Sigma_a^+$, $\Phi^- \in (E^-)^K$,

$$\partial_{\Psi^-}\mathcal{G}(\Psi^+, h(\Psi^+))[\Phi^-] = 0.$$

Differentiating with respect to Ψ^+, we get, for all $\Phi^+ \in T_{\Psi^+}\Sigma$,

$$\partial^2_{\Psi^-\Psi^+}\mathcal{G}(\Psi^+, h(\Psi^+))[\Phi^-, \Phi^+] + \partial^2_{\Psi^-\Psi^-}\mathcal{G}(\Psi^+, h(\Psi^+))[\Phi^-, h'(\Psi^+)[\Phi^+]] = 0,$$

and therefore, from the definition of $\mathcal{G}$,

$$-F''_{a,\Psi^+}(\Psi^-)[\Phi^-, h'(\Psi^+)[\Phi^+]] = \partial^2_{\Psi^-\Psi^+}\mathcal{G}(\Psi^+, h(\Psi^+))[\Phi^-, \Phi^+].$$

We now apply this to $\Psi^+ = \Psi_n^+$, $\Psi^- = \Psi_n^-$, $\Phi^+ = \Phi_n^+$ and $\Phi^- = \Phi_n^-$, and get

$$-F''_{a,\Psi_n^+}(\Psi_n^-)[\Phi_n^-, \Phi_n^-] = \partial^2_{\Psi^-\Psi^-}\mathcal{G}(\Psi_n^+, \Psi_n^-)[\Phi_n^-, \Phi_n^-]. \tag{33}$$

Using estimates similar to but slightly more complicated than those in [ES99], we estimate

$$\partial^2_{\Psi^-\Psi^-}\mathcal{G}(\Psi_n^+, \Psi_n^-)[\Phi_n^-, \Phi_n^-] \le O\left(\left\|\nabla\Phi_n^+\right\|_{L^2}\left\|\Phi_n^-\right\|_{L^2}\right),$$

where the notation O is for both c and n large.

But, by Lemma 5,

$$F''_{a,\Psi_n^+}(\Psi_n^-)[\Phi_n^-, \Phi_n^-] \le -\frac{\gamma c^2}{2}\left\|\Phi_n^-\right\|_c^2,$$

from which we conclude, from (33), that

$$\frac{\gamma c^2}{2}\left\|\Phi_n^-\right\|_c^2 \leq O\left(\left\|\nabla\Phi_n^+\right\|_{L^2}\left\|\Phi_n^-\right\|_{L^2}\right),$$

and therefore that

$$\left\|\Phi_n^-\right\|_c \leq \frac{1}{c^2} O\left(\left\|\nabla\Phi_n^+\right\|_{L^2}\right). \tag{34}$$

Step 2. We now write the Hessian of $\mathcal{E}$ along the curve $\Psi_n^+(\varepsilon)$.

First, we compute

$$\begin{aligned}\Phi_n &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Psi_n(\varepsilon)\bigg|_{\varepsilon=0}, \\ &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon} G_n^{-1/2}(\varepsilon)\left(\Psi_n^+(\varepsilon)+\Psi_n^-(\varepsilon)\right)\bigg|_{\varepsilon=0}, \\ &= (\delta_n, 0, \ldots, 0) + R_n^+ + R_n^-, \end{aligned} \tag{35}$$

where, using (34), we can estimate the remainder terms $R_n^\pm \in E^\pm$ as

$$\begin{aligned}\left\|R_n^+\right\|_c &= O\left(\frac{1}{c^4}\|\nabla\delta_n\|_{L^2}\right), \\ \left\|R_n^-\right\|_c &= O\left(\frac{1}{c^2}\|\nabla\delta_n\|_{L^2}\right). \end{aligned}$$

Using these estimates and the same arguments as in [Lew04], we can now compute

$$\begin{aligned}\mathcal{E}''(\Psi_n)[\Phi_n, \Phi_n] \leq \Gamma_{11}&\left(c^2\|\delta_n\|_c^2 + \left\langle \delta_n, (V + \rho_n \star \frac{1}{|x|})\delta_n\right\rangle\right) \\ &+ O\left(\frac{1}{c^2}\|\nabla\delta_n\|_{L^2}^2 + \frac{1}{c^2}\|\nabla\delta_n\|_{L^2}\|\delta_n\|_{L^2}\right),\end{aligned}$$

with $\int \rho_n = N-1$.

Now, let U be an arbitrary vector subspace of $H^1(\mathbb{R}^3, \mathbb{C}^4)$ consisting of functions of the form

$$\begin{pmatrix} f(|x|) \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

with dimension at least $K+1$. Let U_λ^+ be the positive projection of the dilation of U of a factor λ, *i.e.*

$$U_\lambda = P^+\left\{\psi\left(\frac{x}{\lambda}\right), \psi \in U\right\}.$$

U_λ^+ is also of dimension $K+1$ for c large enough, so we can find a function $\delta_n \in U_\lambda^+$ normalized in L^2 which is orthogonal to Ψ_n^+. For such a function,

$$\mathcal{E}''(\Psi_n)[\Phi_n, \Phi_n] \leq \left(c^2 - \eta\frac{Z-N-1}{\lambda} + O\left(\frac{1}{\lambda^2} + \frac{1}{\lambda c^2}\right)\right)\Gamma_{11},$$

with $\eta > 0$, where the O notation is understood for n, c and λ large. So, taking λ large enough independently of n and c, we get

$$\mathcal{E}''(\Psi_n)[\Phi_n, \Phi_n] \leq \left(c^2 - \kappa\right)\Gamma_{11}, \tag{36}$$

with $\kappa > 0$ independent of n and c.

Step 3. Using (35) again, we estimate

$$\langle \Phi_n, \Lambda_n \Phi_n \rangle = (\Lambda_n)_{11} + O\left(\frac{1}{c^4}\Lambda_n\right).$$

But we can obtain a very crude control on Λ_n thanks to the estimates in Lemmas 3 and 4 :

$$\begin{aligned}(\Lambda_n)_{ij} &= \left\langle \Psi_n^i, H_{a,\Psi_n}\Psi_n^j \right\rangle + o_{n\to\infty}(1),\\ &= c^2\Gamma_{ij} + O(\|\Psi_n\|_E^2) + o_{n\to\infty}(1),\end{aligned}$$

and therefore

$$\Lambda_n = c^2\Gamma + O(1). \tag{37}$$

Step 4. We now use the second order condition to conclude.

We have $\langle \Psi_n(\varepsilon), \Lambda_n \Psi_n(\varepsilon)\rangle = \operatorname{tr}\Lambda_n$, so, defining the Lagrangian $L_n(\Psi) = \mathcal{E}(\Psi) - \langle \Psi, \Lambda_n \Psi\rangle$, we get from the second order information *(iii)* that

$$\beta_n \left\|\Psi_n^+\right\|_E^2 \leq L_n''(\Psi_n)\left[\Phi_n, \Phi_n\right] + O\left(\left\|\frac{\mathrm{d}^2}{\mathrm{d}^2\varepsilon}\Psi_n(\varepsilon)\Big|_{\varepsilon=0}\right\|_{H^{1/2}} \left\|\mathcal{F}'(\Psi_n^+)\right\|_{H^{-1/2}}\right).$$

Therefore, from the Palais-Smale condition *(ii)*

$$\mathcal{E}''(\Psi_n)[\Phi_n, \Phi_n] \geq \langle \Phi_n, \Lambda_n \Phi_n\rangle + o_{n\to\infty}(1).$$

Finally, from (36) and (37), we obtain, for c and n large enough,

$$(\Lambda_n)_{11} \leq (c^2 - K_2)\Gamma_{11},$$

with $K_2 > 0$.

Using the group action (14), we could apply the same procedure to $(\tilde{a}, \widetilde{\Psi}_n^+) = U\cdot(a, \Psi_n^+)$ for any $U \in \mathcal{U}(K)$, and obtain

$$(U\Lambda_n U^*)_{11} \leq (c^2 - K_2)(U\Gamma U^*)_{11},$$

which proves our result

$$\Lambda_n \leq (c^2 - K_2)\Gamma.$$

Lower bound on the Lagrange multipliers. Let $A_n = (c^2 - K_2)\Gamma - \Lambda_n$. We know that, for n large enough, $A_n \geq 0$, and, from (37),

$$\begin{aligned}\operatorname{tr} A_n &= Nc^2 - NK_2 - \operatorname{tr}\Lambda_n,\\ &= O(1).\end{aligned}$$

So $A_n = O(1)$, and therefore $\Lambda_n \geq (c^2 - K_2)\Gamma - O(1)$. Because $\Gamma \geq \gamma > 0$, the result follows for c large. ☐

5.5. **Proof of Theorem 1.** For any $a \in S'_\gamma$, we can apply the Borwein-Preiss variational principle [BP87] to the functional $\mathcal{F}_a$ on Σ_a^+, and obtain a sequence Ψ_n^+ that satisfies the hypotheses of Lemma 7. The associated sequence Ψ_n satisfies the hypotheses of Lemma 1 so, the sequence (a, Ψ_n^+) converges up to extraction to a limit Ψ_a^+, solution of the min-max principle

$$\mathcal{F}_a(\Psi_a^+) = \min_{\Psi^+ \in \Sigma^+} \max_{\Psi^- \in (E^-)^K} \mathcal{E}(a, g(\Psi^+ + \Psi^-)).$$

We now take a minimizing sequence a_n for the continuous functional $F_a(\Psi_a^+)$ on S'_γ. The sequence $(a_n, \Psi_{a_n}^+)$ again verifies the hypotheses of Lemma 1, and therefore converges to (a_*, Ψ_*^+). The triplet $(a_*, \Psi_*^+, h(\Psi_*^+))$ is now a solution of the variational principle (21), and Theorem 1 is proved.

6. Proof of Theorem 2

6.1. **Nonrelativistic limit.** We begin with a lemma that is the multiconfiguration analogue of Theorem 3 of [ES01].

Lemma 8. *Let $c_n \to \infty, (a_n, \Psi_n) \in S_\gamma \times \Sigma$ solutions of*

$$H_{a_n,\Psi_n}\Psi_n = \Lambda_n \Psi_n \tag{38}$$

such that

$$({c_n}^2 - K_1)\Gamma_n \le \Lambda_n \le ({c_n}^2 - K_2)\Gamma_n$$

for constants $K_1, K_2 > 0$.

Then, up to a subsequence, $a_n \to a \in S_\gamma$, $\Psi_n \to \begin{pmatrix}\Phi\\0\end{pmatrix}$ in H^1, and $\mathcal{E}(a_n, \Psi_n) - Nc^2 \to \mathcal{E}^{HF}(a, \Phi)$

Proof. First, we need a uniform bound on Ψ_n in H^1.

$$\begin{aligned}\|D_c\Gamma_n\Psi_n\|_{L^2}^2 &= \left\langle \Gamma_n\Psi_n, (c^4 - c^2\Delta)\Gamma_n\Psi_n \right\rangle,\\ &= {c_n}^4\|\Gamma_n\Psi_n\|_{L^2}^2 + {c_n}^2\|\Gamma_n\nabla\Psi_n\|_{L^2}^2 .\end{aligned}$$

On the other hand,

$$\begin{aligned}\|D_c\Gamma_n\Psi_n\|_{L^2}^2 &= \left\|(V\Gamma + 2W_{a_n,\Psi_n})\Psi_n - \Lambda_n\Psi_n\right\|_{L^2}^2,\\ &\le {c_n}^4\|\Gamma_n\Psi_n\|_{L^2}^2 + C\|\nabla\Psi_n\|_{L^2}^2 + C{c_n}^2\|\Gamma_n\nabla\Psi_n\|_{L^2} .\end{aligned}$$

by the classical Hardy inequality, with $C > 0$. Therefore, Ψ_n is bounded in H^1.

We now write $\Psi_n = \begin{pmatrix}\Phi_n\\ \mathcal{X}_n\end{pmatrix}$, where $\Phi_n, \mathcal{X}_n \in H^1(\mathbb{R}^3, \mathbb{C}^2)$. We rewrite the equations (38) as

$$c_n\Gamma_n L\mathcal{X}_n + (V\Gamma_n + 2W_{a_n,\Psi_n})\Phi_n = (\Lambda_n - {c_n}^2\Gamma_n)\Phi_n, \tag{39}$$

$$c_n\Gamma_n L\Phi_n + (V\Gamma_n + 2W_{a_n,\Psi_n})\mathcal{X}_n = (\Lambda_n + {c_n}^2\Gamma_n)\mathcal{X}_n, \tag{40}$$

with the operator

$$L = -i\nabla \cdot \sigma \tag{41}$$

such that $L^2 = -\Delta$.

Because $\Lambda_n < ({c_n}^2 - K_2)\Gamma_n$, using the Hardy inequality and the boundedness of Φ_n in H^1, the first equation (39) yields

$$\|\Gamma_n L\mathcal{X}_n\|_{L^2} = \|\Gamma_n\nabla\mathcal{X}_n\|_{L^2} = O(1/c_n). \tag{42}$$

The second equation (40) gives

$$\mathcal{X}_n = \frac{1}{2c}\left(\frac{1}{2}\left(\Gamma_n + \Lambda_n/{c_n}^2\right)\right)^{-1}\Gamma_n L\Phi_n + \frac{1}{{c_n}^2}O(\|\mathcal{X}_n\|_{H^1})$$
$$= \mathrm{KB}(\Phi_n) + \frac{1}{{c_n}^2}O(\|\mathcal{X}_n\|_{H^1}) + O\left(\frac{1}{{c_n}^3}\right) \quad (43)$$

in L^2 norm, where the "kinetic balance" operator KB is given by

$$\mathrm{KB}(\Phi) = \frac{1}{2c}L\Phi. \quad (44)$$

Equation (43) gives $\|\mathcal{X}_n\|_{L^2} = \frac{1}{2c_n}\|L\Phi_n\|_{L^2} + O(1/{c_n}^2) = O(1/c_n)$, and then

$$\mathcal{X}_n = \mathrm{KB}(\Phi_n) + O\left(\frac{1}{{c_n}^3}\right) \quad (45)$$

again in L^2 norm. Φ_n satisfies

$$\left(-\frac{1}{2}\Delta\Gamma_n + V\Gamma_n + 2W_{\Phi_n}\right)\Phi_n = (\Lambda_n - {c_n}^2\Gamma_n)\Phi_n + \Delta_n$$
$$\mathrm{Gram}\,\Phi_n = 1 + o(1)$$

with $\Delta_n \to 0$ in L^2 and therefore H^{-1} norm. $(a_n, g(\Phi_n))$ is a Palais-Smale sequence for the nonrelativistic functional, with control on the Lagrange multipliers $(\Lambda_n - {c_n}^2\Gamma_n) < 0$ and non-degeneracy information $\Gamma_n \geq \gamma$. By the arguments in the proof of Theorem 1 of [Lew04], (a_n, Φ_n) converges, up to a subsequence, to (a, Φ) in H^1 norm, and it is easy to compute from (45) that

$$\langle \Psi_n, D_{c_n}\Gamma_n\Psi_n\rangle = N{c_n}^2 + \frac{1}{2}\langle\Phi_n, (-\Delta)\Gamma_n\Phi_n\rangle + o(1),$$

and the result follows. □

We are now ready to prove Theorem 2.

6.2. **Proof of Theorem 2.**

Proof. The sequence (a_n, Ψ_n) satisfies the hypotheses of Lemma 8 : up to a subsequence, it converges strongly in H^1 to $\left(a, \begin{pmatrix}\Phi\\0\end{pmatrix}\right)$, with $\lim \mathcal{E}(a_n, \Psi_n) - N{c_n}^2 = \mathcal{E}^{\mathrm{HF}}(a, \Phi)$. But since by Lemma 6 we have

$$\mathcal{E}(a_n, \Psi_n) = I_{c_n,\gamma} \leq Nc_n^2 + I^K + o_{c_n\to\infty}(1),$$

we obtain $\mathcal{E}^{\mathrm{HF}}(a, \Phi) = I^K$, hence the result. □

Acknowledgements

I would like to thank Éric Séré for his attention and helpful advice.

References

[BCMT10] C. Bardos, I. Catto, N. Mauser, and S. Trabelsi. Setting and analysis of the multi-configuration time-dependent Hartree-Fock equations. *Archive for Rational Mechanics and Analysis*, 198(1):273–330, 2010.

[BES06] B. Buffoni, M.J. Esteban, and E. Séré. Normalized solutions to strongly indefinite semilinear equations. *Adv. Nonlinear Stud.*, 6:323–347, 2006.

[BP87] J.M. Borwein and D. Preiss. A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions. *Trans. Amer. Math. Soc*, 303(51):7–527, 1987.

[Der12] J. Derezinski. Open problems about many-body Dirac operators. *Bulletin of International Association of Mathematical Physics*, 2012.

[DFJ07] K.G. Dyall and K. Faegri Jr. *Introduction to relativistic quantum chemistry.* Oxford University Press, 2007.

[ES99] M.J. Esteban and E. Séré. Solutions of the Dirac-Fock equations for atoms and molecules. *Communications in Mathematical Physics*, 203(3):499–530, 1999.

[ES01] M.J. Esteban and E. Séré. Nonrelativistic limit of the Dirac-Fock equations. *Annales Henri Poincare*, 2:941–961, 2001.

[Fri03a] G. Friesecke. The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Archive for Rational Mechanics and Analysis*, 169(1):35–71, 2003.

[Fri03b] G. Friesecke. On the infinitude of non-zero eigenvalues of the single-electron density matrix for atoms and molecules. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 459(2029):47–52, 2003.

[Gra07] I. Grant. *Relativistic Quantum Theory of Atoms and Molecules*. Springer, 2007.

[Her77] I. Herbst. Spectral theory of the operator $(p^2 + m^2)^{-1/2} - ze^2/r$. *Communications in Mathematical Physics*, 53(3):285–294, 1977.

[ID93] P. Indelicato and J.P. Desclaux. Projection operator in the multiconfiguration Dirac-Fock method. *Physica Scripta*, 1993(T46):110, 1993.

[Kat66] T. Kato. *Perturbation theory for linear operators.* Springer Verlag, 1966.

[LB94] C. Le Bris. A general approach for multiconfiguration methods in quantum molecular chemistry. *Annales de l'Institut Henri Poincaré. Analyse non linéaire*, 11(4):441–484, 1994.

[Lew04] M. Lewin. Solutions of the multiconfiguration equations in quantum chemistry. *Archive for Rational Mechanics and Analysis*, 171(1):83–114, 2004.

[Lio87] P.L. Lions. Solutions of Hartree-Fock equations for Coulomb systems. *Communications in Mathematical Physics*, 109(1):33–97, 1987.

[LS77] E.H. Lieb and B. Simon. The Hartree-Fock theory for Coulomb systems. *Comm. Math. Phys.*, 53(3):185–194, 1977.

[PD79] P. Pyykko and J.P. Desclaux. Relativity and the periodic system of elements. *Accounts of Chemical Research*, 12(8):276–281, 1979.

[SO89] A. Szabo and N.S. Ostlund. *Modern quantum chemistry.* McGraw-Hill New York, 1989.

[Tix98] C. Tix. Strict positivity of a relativistic Hamiltonian due to Brown and Ravenhall. *Bulletin of the London Mathematical Society*, 30(3):283–290, 1998.

Université Paris-Dauphine, CEREMADE, Place du Maréchal Lattre de Tassigny, 75775 Paris Cedex 16, France.

E-mail address: levitt@ceremade.dauphine.fr