



HAL
open science

Solutions of the multiconfiguration Dirac-Fock equations

Antoine Levitt

► **To cite this version:**

| Antoine Levitt. Solutions of the multiconfiguration Dirac-Fock equations. 2013. hal-00823185v1

HAL Id: hal-00823185

<https://hal.science/hal-00823185v1>

Preprint submitted on 16 May 2013 (v1), last revised 22 Jul 2014 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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 to these equations in the weakly relativistic regime. First, using results from Lewin on the multiconfiguration nonrelativistic model, and Esteban and Séré on the single-configuration relativistic model, we prove 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	1
2. Definitions	3
3. Strategy of proof	5
4. Results	7
5. Solutions of the constrained MCDF equations	8
5.1. The modified functional	8
5.2. Convergence of Palais-Smale sequences	9
5.3. Elimination of the negative directions	9
5.4. Convergence of Palais-Smale sequences for $I_{\nu,p}$	10
5.5. Existence of Palais-Smale sequences for $I_{\nu,p}$	11
5.6. Proof of Theorem 1	14
6. Nonrelativistic limit	14
6.1. Proof of Lemma 9	14
6.2. Proof of Theorem 2	16
Acknowledgements	17
References	17

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

Date: May 16, 2013.

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

Key words and phrases. multiconfiguration methods, relativistic quantum mechanics, critical points.

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

of this problem grows exponentially with N , and approximations are used to keep the 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 dates to Lieb and Simon [LS77], 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

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

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 value $c = 1/\alpha \approx 137$.

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

$$D_c = P_c^+ \sqrt{c^4 - c^2\Delta} P_c^+ - P_c^- \sqrt{c^4 - c^2\Delta} P_c^-, \quad (2)$$

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

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

We denote by $E = H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ the form-domain of D_c , and $E_c^\pm = P_c^\pm E$ the two positive and negative spectral subspaces.

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

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

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 $\frac{(N-1)-Z}{|x|}$. Mathematically, it is used to prove that second order information on Palais-Smale sequence 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 inequalities (see [Tix98, Her77, Kat66]) :

$$\langle \phi, V\phi \rangle \leq \frac{1}{2}(\pi/2 + 2/\pi) \langle \phi, \sqrt{1 - \Delta}\phi \rangle \quad (4)$$

for all $\phi \in E_1^\pm$, and

$$\langle \phi, V\phi \rangle \leq \frac{\pi}{2} \langle \phi, \sqrt{1 - \Delta}\phi \rangle \quad (5)$$

for all $\phi \in E$. These inequalities will be extensively used in our proof of existence of solutions.

The N -body relativistic Hamiltonian is given by

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

This Hamiltonian acts on $L_a^2(\mathbb{R}^{3N}, \mathbb{C}^4)$, 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

$$\psi = \sum_{1 \leq i_1 < \dots < i_N \leq K} a_{i_1, \dots, i_N} |\psi_{i_1} \dots \psi_{i_N}\rangle, \quad (6)$$

where

$$|\psi_{i_1} \dots \psi_{i_N}\rangle(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det(\psi_{i_k}(x_l))_{k,l}$$

are Slater determinants, and $a \in S$, $\Psi \in \Sigma$, where

$$S = \{a \in \mathbb{C}^{\binom{K}{N}}, \|a\|^2 = \sum_{1 \leq i_1 < \dots < i_N \leq K} |a_{i_1, \dots, i_N}|^2 = 1\}, \quad (7)$$

$$\Sigma = \{\Psi \in E^K, \text{Gram}_{L^2} \Psi = 1_K\}. \quad (8)$$

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,

$$\psi(x_1, \dots, x_N) = \sum_{1 \leq i_1 \leq N, \dots, 1 \leq i_N \leq N} \alpha_{i_1, \dots, i_N} \psi_{i_1}(x_1) \dots \psi_{i_N}(x_N).$$

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

$$\mathcal{E}(a, \Psi) = \left\langle \Psi, \left((D_c + V)\Gamma_a + W_{a, \Psi} \right) \cdot \Psi \right\rangle_{(L^2(\mathbb{R}^3, \mathbb{C}^4)^K)}, \quad (9)$$

with the $K \times K$ hermitian matrices

$$\begin{aligned} (\Gamma_a)_{i,j} &= N \sum_{k_2 \dots k_N} \alpha_{i,k_2 \dots k_N} \alpha_{j,k_2 \dots k_N}^*, \\ (W_{a,\Psi})_{i,j} &= \frac{N(N-1)}{2} \sum_{k_3 \dots k_N} \sum_{k,l} \alpha_{i,k,k_3 \dots k_N} \alpha_{j,l,k_3 \dots k_N}^* \left(\psi_k \psi_l^* \star \frac{1}{|x|} \right), \end{aligned}$$

where we denote by z^* the complex conjugate of z . The eigenvalues γ_i of Γ_a , for $a \in S$, satisfy $0 \leq \gamma_i \leq 1$, 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

$$\mathcal{E}^{\text{HF}}(a, \Phi) = \left\langle \Phi, \left(\left(-\frac{1}{2} \Delta + V \right) \Gamma_a + W_{a,\Phi} \right) \cdot \Phi \right\rangle_{(L^2(\mathbb{R}^3, \mathbb{C}^2))^K}, \quad (10)$$

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

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

$$U \cdot (a, \Psi) = (a', U\Psi), \quad (11)$$

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

$$\alpha'_{i_1, \dots, i_N} = \sum_{j_1, \dots, j_N} U_{i_1, j_1}^* \dots U_{i_N, j_N}^* \alpha_{j_1, \dots, j_N}. \quad (12)$$

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 ,

$$H_{a,\Psi} \cdot \Psi = \Lambda \Psi, \quad (13)$$

where

$$H_{a,\Psi} = (D_c \Gamma_a + V \Gamma_a + 2W_{a,\Psi}) \quad (14)$$

is the Fock operator, and, for a ,

$$\mathcal{H}_\Psi a = E a, \quad (15)$$

where

$$(\mathcal{H}_\Psi)_{I,J} = \left\langle \left| \psi_{i_1} \dots \psi_{i_N} \right\rangle, H^N \left| \psi_{j_1} \dots \psi_{j_N} \right\rangle \right\rangle \quad (16)$$

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 (13) and (15) by finding critical points of \mathcal{E} on $S \times \Sigma$.

3. STRATEGY OF PROOF

We can now see the major mathematical difficulties in the study of the MCDF model. One can use the group action (11) to diagonalize Γ_a or Λ , but not both at the same time. Furthermore, because $W_{a,\Psi}$ does not in general commute with Γ_a , 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 Γ_a . This gap is used centrally to prove the convergence of Palais-Smale sequences. Therefore, one needs a lower bound on Γ .

We prove the existence of solutions in two steps. First, for a given γ , we look for a critical point of \mathcal{E} in $S_\gamma \times \Sigma$, where

$$S_\gamma = \{a \in S, \det \Gamma_a > \gamma\}. \quad (17)$$

To ensure that Palais-Smale sequences do not accumulate at the border of S_γ (which would prevent their convergence), we penalize \mathcal{E} by $\theta_\gamma(\det \Gamma_a)$, where θ_γ is a smooth function from $(\gamma, +\infty)$ to \mathbb{R}^+ such that (see Figure 1)

$$\begin{cases} \theta_\gamma(x) \rightarrow +\infty & \text{as } x \rightarrow \gamma^+, \\ \theta_\gamma(x) \text{ is decreasing} & \text{on } (\gamma, 2\gamma), \\ \theta_\gamma(x) = 0 & \text{on } [2\gamma, +\infty). \end{cases} \quad (18)$$

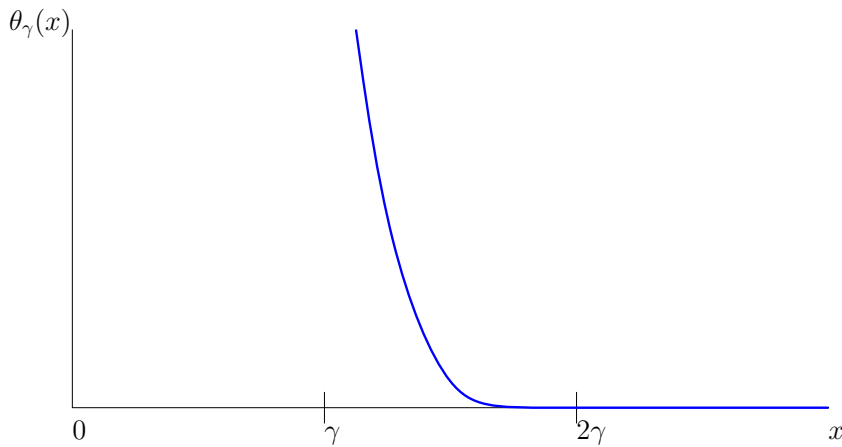


FIGURE 1. Plot of the θ function.

Since $\det \Gamma_a > \gamma$, under a lower bound on c that depends on γ , one can prove the existence of a critical point of $\mathcal{E}(a, \Psi) + \theta_\gamma(\det \Gamma_a)$ by the same techniques as in [Lew04, ES99]. This (a, Ψ) is a solution of (13), but is only a solution of (15) if $\det \Gamma_a > 2\gamma$.

To prove that this is the case, we examine the nonrelativistic case. There, we may assume that $I^K < I^{K-1}$, where

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

is the ground-state energy of the nonrelativistic multiconfiguration method of rank $K \geq N$. The cases where $I^K < I^{K-1}$ are not known exactly, 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 .

The inequality $I^K < I^{K-1}$ implies that $\det \Gamma$ is bounded from below on the minimizers of the nonrelativistic functional. The compactness of these minimizers (implicitly proved in [Lew04]) shows that the bound is uniform.

Because there is no well-defined “ground state energy” in the relativistic case, we cannot use information of this type directly. Rather, we take the nonrelativistic ($c \rightarrow \infty$) limit of the critical points found in the first step. By arguments similar to the ones of [ES01], we prove that these critical points converge, up to a subsequence, to a minimizer of the (penalized) Hartree-Fock functional. If $I^K < I^{K-1}$, $\det \Gamma$ is bounded from below by $\gamma_0 > 0$ on the minimizers. By taking $\gamma < \frac{1}{2}\gamma_0$, we obtain for c large enough critical points that satisfy $\det \Gamma_a > 2\gamma$, which are solutions of the MCDF equations (13)-(15).

4. RESULTS

Our first theorem is the existence of a critical point of $\mathcal{E}(a, \Phi) + \theta_\gamma(\det \Gamma_a)$ on $S_\gamma \times \Sigma$. This is a direct generalisation of the main theorem of [ES99], using the formalism of [Lew04].

We need the following definition to state our theorem. For all $\Psi \in \Sigma$, let $\mathcal{R}(K)$ be the set of hermitian matrices R of size K such that $0 < R < 1$. Then, for $(a_0, \Psi_0) \in S \times \Sigma$, let

$$D(a_0, \Psi_0) = \{(a, \Psi) \in S_\gamma \times E^K, \text{ there are } U \in \mathcal{U}(K), R \in \mathcal{R}(K) \text{ such that} \\ (a, \Psi) = U \cdot (a_0, R\Psi_0)\}, \quad (20)$$

where $U \cdot (a_0, R\Psi_0)$ is given by the group action (11).

Theorem 1. *Let $\gamma > 0, N < K + 1$. There are constants $c_0, K_1, K_2 > 0$ such that, for all $c \geq c_0$, there is a solution (a_c, Ψ_c) of $H_{a_c, \Psi_c} \cdot \Psi_c = \Lambda \Psi_c$ in $S_\gamma \times \Sigma$ that satisfies*

$$(c^2 - K_1)\Gamma_{a_c} \leq \Lambda \leq (c^2 - K_2)\Gamma_{a_c}. \quad (21)$$

If $\det \Gamma_{a_c} > 2\gamma$, there is $E \in \mathbb{R}$ such that $\mathcal{H}_{\Psi_c} a_c = E a_c$.

For any $(a_, \Psi_*^+) \in S_\gamma \times (\Sigma \cap (E_c^+)^K)$,*

$$\mathcal{E}(a_c, \Psi_c) + \theta(\det \Gamma_{a_c}) \leq \sup \left\{ \mathcal{E}(a, \Psi^+ + \Psi^-) + \theta(\det \Gamma_a), \right. \\ \left. (a, \Psi^+) \in D(a_*, \Psi_*^+), \Psi^- \in (E_c^-)^K, \text{Gram}(\Psi^+ + \Psi^-) \leq 1 \right\}. \quad (22)$$

Remarks. The constant c_0 can be made explicit as a function of γ . We do not do it because our main result (Corollary 1) is only valid for large c , with no explicit estimate on its minimum size.

Although there is no minimum of $\mathcal{E}(a, \Psi) + \theta_\gamma(\det \Gamma_a)$ on $S_\gamma \times \Sigma$, the property (22) makes (a, Ψ) a good candidate for a “ground state”. In particular, it can be used to prove that the nonrelativistic limit of the solutions obtained is a minimizer of the Hartree-Fock functional (Theorem 2 below).

The same scheme of proof would also yield the existence of an infinite number of critical points. But, even in the nonrelativistic limit, we are unable to prove that $\det \Gamma \geq 2\gamma$ for these critical points. Therefore, we only concern ourselves with “ground states”.

We now study the nonrelativistic limit of these “ground states”:

Theorem 2. *Let $\gamma > 0, N < Z + 1, c_n \rightarrow \infty$ with $c_n \geq c_0$ for all $n \in \mathbb{N}$, and (a_n, Ψ_n) the solution of (13) obtained by Theorem 1 with $c = c_n$. Then, up to a subsequence,*

$$(a_n, \Psi_n) \rightarrow \left(a, \begin{pmatrix} \Phi \\ 0 \end{pmatrix} \right) \text{ in } S_\gamma \times \Sigma, \text{ where } (a, \Phi) \text{ is a minimizer of}$$

$$I_\gamma^K = \inf \left\{ \mathcal{E}^{\text{HF}}(a, \Phi) + \theta_\gamma(\det \Gamma_a), a \in S_\gamma, \Phi \in (H^1(\mathbb{R}^3, \mathbb{C}^2))^K, \text{Gram } \Phi = 1 \right\}. \quad (23)$$

Our main result is now

Corollary 1. *If $I^K < I^{K-1}$, for c large enough, there are solutions of the multiconfiguration Dirac-Fock equations (13)-(15).*

Proof. Recall the definition (19)

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

Because $I^K < I^{K-1}$, $\det \Gamma$ is bounded from below on each minimizer of the multiconfiguration Hartree-Fock functional. From Theorem 1 of [Lew04], these minimizers form

a compact set, and therefore there is $\gamma_0 > 0$ such that $\det \Gamma \geq \gamma_0$ uniformly on the minimizers.

We now choose $\gamma = \frac{1}{4}\gamma_0$. By definition, $I_\gamma^K \geq I^K$. But any minimizer (a, Φ) of the variational problem (19) that defines I^K satisfies $\det \Gamma_a \geq 2\gamma$, and therefore achieves the value I^K in the variational problem (23) that defines I_γ^K . Therefore, we get $I_\gamma^K = I^K$. We obtain from Theorem 1 a sequence (a_n, Ψ_n) of solutions of (13). By Theorem 2, this sequence converges to a $\left(a, \begin{pmatrix} \Phi \\ 0 \end{pmatrix}\right)$, which satisfies

$$\mathcal{E}^{\text{HF}}(a, \Phi) + \theta_\gamma(\det \Gamma_a) = I_\gamma^K = I^K.$$

(a, Φ) is a minimizer of the nonrelativistic Hartree-Fock functional, and therefore $\det \Gamma_a \geq \gamma_0 > 2\gamma$. Since $a_n \rightarrow a$, for n large enough $\det \Gamma_{a_n} > 2\gamma$. By Theorem 1, (a_n, Φ_n) is then a solution of (13)-(15). \square

5. SOLUTIONS OF THE CONSTRAINED MCDF EQUATIONS

In this section we prove Theorem 1. We use the method of proof of Theorem 1.2 of [ES99], with the formalism of [Lew04]. The proof is essentially identical to the one in [ES99], with only technical modifications to accomodate the multiconfiguration formalism. We will only outline the main lemmas, and indicate which parts of the proof have to be modified in our case. We use centrally the fact that $\det \Gamma > \gamma$, and therefore the occupation numbers are uniformly bounded away from zero.

5.1. The modified functional. First, we regularize the potential $1/|x|$ appearing in the potential energy term (both the nucleus attraction and the electron repulsion). We make the replacement

$$\frac{1}{|x|} \rightarrow f_\nu \star \frac{1}{|x|},$$

where

$$f_\nu(x) = \frac{1}{\nu^3} f\left(\frac{x}{\nu}\right),$$

with f a fixed positive smooth function, normalized in L^1 . This substitution leaves unchanged the inequality (4), and is only used for the technical Lemma 4. In the limit $\nu \rightarrow 0$, we recover the original Coulomb potential.

We replace the normalization condition $\Psi \in \Sigma$ by a penalization term

$$\begin{aligned} \pi_p(\Psi) &= \text{tr} \left((\text{Gram } \Psi)^p (1 - \text{Gram } \Psi)^{-1} \right). \\ &= \sum_{k=1}^K f_p(\sigma_k), \end{aligned} \tag{24}$$

where the σ_k are the eigenvalues of $\text{Gram } \Psi$, and

$$f_p(x) = \frac{x^p}{1-x}$$

(see Figure 2.)

We now look for critical points of the functional

$$\mathcal{F}_{\nu,p}(a, \Psi) = \mathcal{E}_\nu(a, \Psi) - \pi_p(\Psi) + \theta_\gamma(\det \Gamma_a), \tag{25}$$

on $S_\gamma \times A$, where

$$A = \{\Psi \in E^K, 0 < \text{Gram } \Psi < 1\}. \tag{26}$$

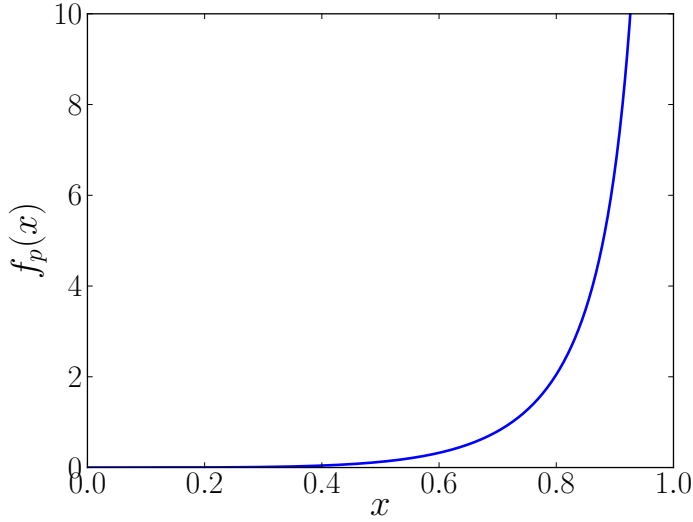


FIGURE 2. Plot of the f_p function.

The Euler-Lagrange equation with respect to Ψ is

$$H_{a,\Psi} \cdot \Psi = \nabla \pi_p(\Psi) \Psi. \quad (27)$$

Furthermore, if $\det \Gamma_a > 2\gamma$, then there is $E \in \mathbb{R}$ such that $\mathcal{H}_\Psi a = Ea$.

In the proof of Theorem 1, we obtain a critical point $(a_{\nu,p}, \Psi_{\nu,p})$ at a fixed ν, p . Then, we pass to the limit $\nu \rightarrow 0, p \rightarrow \infty$. As $f'_p(x_p) \rightarrow 0$ as $p \rightarrow \infty$ unless $x_p \rightarrow 1$, a lower bound on the Lagrange multipliers Λ will imply that $\text{Gram } \Psi_p \rightarrow 1$, and we will recover a critical point of $\mathcal{E}(a, \Psi) + \theta_\gamma(\Gamma_a)$ in $S_\gamma \times \Sigma$.

5.2. Convergence of Palais-Smale sequences. First, we study the convergence of Palais-Smale sequences for \mathcal{E} with an upper bound on the Lagrange multipliers Λ . Because of our lower bound on the occupation numbers, the proof proceeds exactly as in Lemma 2.1 of [ES99], and we refer to this paper for details. The only technical modification is that the inequalities on Λ have to be taken with respect to the Γ metric. For instance, $\limsup \Lambda_n < c^2$ becomes $\limsup \Lambda_n - c^2 \Gamma_n < 0$.

Lemma 1. *Let $\gamma > 0, N < Z + 1$. Then there are constants $c_0, K_1 > 0$ such that, for $c \geq c_0, \nu_n \in (0, 1), p_n$ a non-decreasing sequence of integers, $(a_n, \Psi_n) \in S_\gamma \times A$ satisfying:*

- (1) $\liminf \text{Gram } \Psi_n > 0$,
- (2) $H_{\Psi_n} \Psi_n - \Lambda_n \Psi_n = \Delta_n \rightarrow 0$ in $(H^{-1/2})^K$,
- (3) $\limsup \Lambda_n - c^2 \Gamma_n < 0$,

we have, up to a subsequence,

- (1) $\liminf \Lambda_n - (c^2 - K_1) \Gamma_n \geq 0$,
- (2) $(a_n, \Psi_n) \rightarrow (a, \Psi)$ strongly in $S_\gamma \times A$,
- (3) If $p_n \rightarrow p, \nu_n \rightarrow \nu \in (0, 1)$, then (a, Ψ) is a solution of $H_\Psi \Psi = \Lambda \Psi$ in $S_\gamma \times A$, and $\mathcal{F}_{\nu,p}(a, \Psi) = \lim_{n \rightarrow \infty} \mathcal{F}_{\nu_n, p_n}(a_n, \Psi_n)$,
- (4) If $p_n \rightarrow \infty, \nu_n \rightarrow 0$, then (a, Ψ) is a solution of $H_\Psi \Psi = \Lambda \Psi$ in $S_\gamma \times \Sigma$, and $\mathcal{E}(a, \Psi) = \lim_{n \rightarrow \infty} \mathcal{F}_{\nu_n, p_n}(a_n, \Psi_n)$.

5.3. Elimination of the negative directions. Next, we use the concavity in the E_c^- directions to get rid of the infinite Morse index of the critical points.

Lemma 2. *Let $\gamma > 0, N < Z + 1$. Then there are $c_0, s > 0$ such that, for $c \geq c_0, \nu \in (0, 1), p \geq 1, (a, \Psi) \in S_\gamma \times A, \Phi^- \in (E_c^-)^K$,*

$$\partial_\Psi^2 \mathcal{F}_{\nu,p}(a, \Psi) \cdot [\Phi^-, \Phi^-] \leq -s \left\| \Phi^- \right\|_{E^K}^2 \quad (28)$$

Proof. The proof proceeds as in Lemma 2.2 of [ES99]. The penalization term $-\pi_p$ is concave, and therefore it suffices to compute the second derivative of \mathcal{E}_ν :

$$\partial_\Psi^2 \mathcal{E}_\nu(a, \Psi) \cdot [\Phi^-, \Phi^-] = \left\langle \Phi^-, (-\sqrt{c^4 - c^2 \Delta} + f_\nu \star V) \Gamma_a \Phi^- \right\rangle + \text{EE},$$

where EE regroups the terms arising from the electron-electron interaction. We can bound EE by the inequality (4) :

$$\text{EE} \leq C \left\| \Phi^- \right\|_{E^K}^2,$$

where C only depends on N and Z . Since $\det \Gamma_a > \gamma$, we have $\Gamma_a > \gamma$ in the sense of hermitian matrices, and

$$\partial_\Psi^2 \mathcal{E}_\nu(a, \Psi) \cdot [\Phi^-, \Phi^-] \leq (-c^2 \gamma + C) \left\| \Phi^- \right\|_{E^K}^2,$$

hence the result. □

Since, for a fixed Ψ^+ , $\mathcal{F}_{\nu,p}(a, \Psi^+ + \Psi^-) \rightarrow -\infty$ as $\det(1 - \text{Gram}(\Psi^+ + \Psi^-)) \rightarrow 0$, we can define the new functional

$$I_{\nu,p}(a, \Psi^+) = \sup_{\Psi^- \in (E_c^-)^K, \text{Gram}(\Psi^- + \Psi^+) < 1} \mathcal{F}_{\nu,p}(a, \Psi^+ + \Psi^-). \quad (29)$$

This functional is smooth on $S_\gamma \times A^+$, where $A^+ = A \cap (E_c^+)^K$.

5.4. Convergence of Palais-Smale sequences for $I_{\nu,p}$. We now look for critical points of $I_{\nu,p}$ on $S_\gamma \times A^+$. On this functional, second order information implies an upper bound on the Lagrange multipliers :

Lemma 3. *Let $\gamma > 0, N < Z + 1$ and d a positive integer. Then there are $c_0, K_2 > 0$ such that, if $c \geq c_0, \nu \in (0, 1), p \geq 2, M > 0, (a_n, \Psi_n^+) \in S_\gamma \times A^+$ are such that*

- (1) $\liminf \text{Gram} \Psi_n^+ > 0$,
- (2) $\limsup I_{\nu,p}(a_n, \Psi_n^+) \leq M$,
- (3) $\partial_{\Psi^+} I_{\nu,p}(a_n, \Psi_n^+) \rightarrow 0$ in $H^{-1/2}$,
- (4) *There is a sequence $\delta_n \rightarrow 0$ such that $\partial_{\Psi^+}^2 I_{\nu,p}(a_n, \Psi_n^+)[\Psi, \Psi] + \delta_n \|\Psi\|_{H^{1/2}}^2$ has a negative space of dimension at most d for all n ,*

then

$$\limsup \Lambda_n - (c^2 - K_2) \Gamma_n \leq 0.$$

Proof. By the same reasoning as in the proof of Lemma 4.3 of [ES99], there is a $C > 0$ such that for all $a \in S_\gamma, \Psi^+ \in A^+, \Phi^+ \in (E_c^+ \cap H^1(\mathbb{R}^3, \mathbb{C}^4))^K$ such that $\langle \Phi_i^+, \Psi_j^+ \rangle = 0 \forall i, j$,

$$\frac{1}{2} \partial_{\Psi^+}^2 I_{\nu,p}(a, \Psi^+) \cdot [\Phi^+, \Phi^+] \leq \frac{1}{2} \partial_{\Psi^+}^2 \mathcal{E}_\nu(a, \Psi^+) \cdot [\Phi^+, \Phi^+] - \langle \Phi, \Lambda \Phi \rangle + C \left\| \nabla \Phi^+ \right\|^2,$$

with $\Lambda = \nabla \pi_p(\Psi^+ + \Psi^-)$, where Ψ^- is the unique maximizer in the variational problem (29) defining $I_{\nu,p}(a, \Psi^+)$.

Now, choosing Φ^+ composed of radial orbitals, we get from the proof of Lemma 4.4 of [ES99] that

$$\frac{1}{2}\partial_{\Psi^+}^2 \mathcal{E}_\nu(a, \Psi^+) \cdot [\Phi^+, \Phi^+] \leq \left\langle \Phi^+, \left(\sqrt{c^4 - c^2\Delta} + (N-1)f_\nu \star \frac{1}{|x|} - Zf_\nu \star V \right) \Gamma_a \Phi^+ \right\rangle$$

When Φ is composed of orbitals of the form $f(x/\lambda)$ for λ large, the term $\nabla\Phi$ becomes negligible, and the potential $(N-1)f_\nu \star \frac{1}{|x|} - Zf_\nu \star V$ can be approximated by $(Z-N+1)\frac{1}{|x|}$. One can find (see Lemma 4.5 of [ES99]) a well-chosen subspace $X_m \subset (E_c^+ \cap H^1(\mathbb{R}^3, \mathbb{C}^4) \cap (\psi_1^+, \dots, \psi_K^+)^\perp)^K$ of arbitrary dimension m such that

$$\frac{1}{2}\partial_{\Psi^+}^2 I_{\nu,p}(a, \Psi^+) \cdot [\Phi^+, \Phi^+] \leq \left\langle \Phi^+, ((c^2 - K_2)\Gamma_a - \Lambda)\Phi^+ \right\rangle$$

for all $\Psi^+ \in X_m$, with $K_2 > 0$ only depending on N, Z, γ and m . The result follows. \square

5.5. Existence of Palais-Smale sequences for $I_{\nu,p}$. We are looking for Palais-Smale sequences with second order information of $I_{\nu,p}$ on $S_\gamma \times A^+$. Note that, even after the sup in the E_c^- directions, this functional is still indefinite. We cannot use a minimization argument, and will use a generalization of the mountain pass theorem with second order information proved by Fang and Ghoussoub [FG92, Gho93].

Following [ES99], define $\Delta(\Psi^+) = \det(\text{Gram } \Psi^+)$. The problem we now face is that $I_{\nu,p}$ may stay positive as $\Delta(\Psi^+) \rightarrow 0$, and therefore violate the conditions of the mountain pass theorem (see Figure 3). To avoid this, we prove that, when $\det \text{Gram } \Psi^+$ is small, “increasing the mass” of Ψ^+ increases the value of $I_{\nu,p}$, in the following sense:

Lemma 4. *Let $\gamma > 0, N < Z + 1, \nu > 0$. There are $d(\nu), e(\nu) > 0$ such that, for all $p \geq 2$, if $(a, \Psi^+) \in S_\gamma \times A^+$ satisfies $d(\nu) \leq \Delta(\Psi^+) \leq 2d(\nu)$, then there is $X \in (E_c^+)^K$ such that*

$$\begin{cases} \partial_{\Psi^+} I_{\nu,p}(a, \Psi^+) \cdot X \geq e(\nu)\|X\|_{E^K}, \\ \Delta'(\Psi^+) \cdot X > 0. \end{cases} \quad (30)$$

This is the only place where the regularization of the potential $\frac{1}{|x|}$ is used.

We restrict our search to the domain $\Delta(\Psi^+) > 2d(\nu)$ by defining a new functional. Let $\alpha_\nu \in C^\infty((d(\nu), 1), \mathbb{R})$ be such that

$$\begin{cases} \alpha_\nu(x) = 0 & \forall x \geq 2d(\nu), \\ \alpha_\nu(x) > 0 & \forall x < 2d(\nu), \\ \alpha_\nu(x) \rightarrow -\infty & \text{as } x \rightarrow d(\nu), \end{cases} \quad (31)$$

and $\beta \in C^\infty(\mathbb{R}, \mathbb{R})$ such that

$$\begin{cases} \beta(t) = -1 & \forall t \in (-\infty, -1), \\ \beta(t) \leq 0 & \forall t \in [-1, 0], \\ \beta(t) = t & \forall t > 0. \end{cases} \quad (32)$$

We now define

$$J_{\nu,p}(a, \Psi^+) = \begin{cases} \beta(I_{\nu,p}(a, \Psi^+) + \alpha_\nu \circ \Delta(\Psi^+)) & \text{if } \Psi^+ \in A^+ \text{ and } \Delta(\Psi^+) > d(\nu) \\ -1 & \text{otherwise} \end{cases} \quad (33)$$

If (a, Ψ^+) is a critical point of $J_{\nu,p}$ with $J_{\nu,p}(\Psi^+) \geq 0$, then it must also be a critical point of $I_{\nu,p} + \alpha_\nu \circ \Delta$. By (30), this cannot be the case unless $\Delta(\Psi^+) > 2d(\nu)$, in which

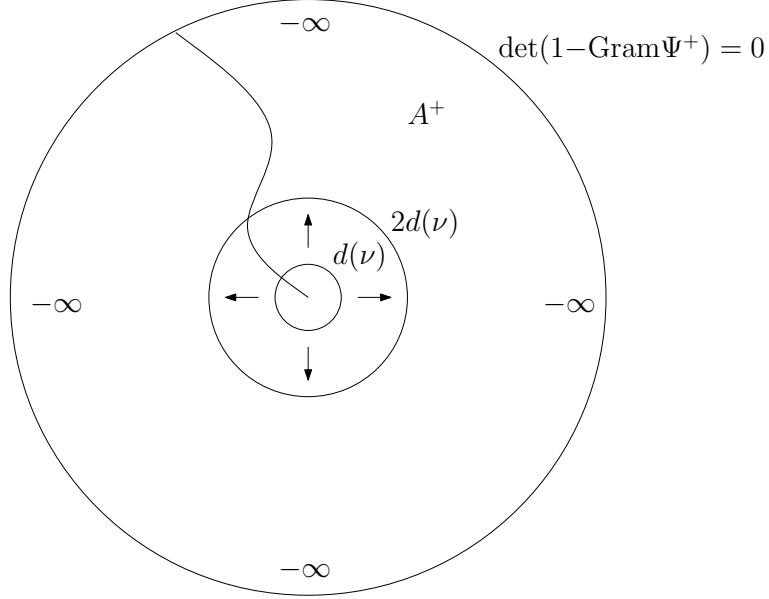


FIGURE 3. Schematic view of the minimization domain A^+ . Between $d(\nu)$ and $2d(\nu)$, the gradient field points towards increasing mass, and $I_{\nu,p} \rightarrow -\infty$ as $\det(1 - \text{Gram } \Psi^+) \rightarrow 0$. We obtain critical points using a mountain-pass type theorem.

case $J_{\nu,p}$ coincides with $I_{\nu,p}$ in a neighborhood of Ψ^+ . Therefore, Palais-Smale sequences of $J_{\nu,p}$ at positive levels are also Palais-Smale sequences of $I_{\nu,p}$ at the same level. Since $J_{\nu,p} = -1$ on ∂A^+ , we are now in the setting of the mountain-pass theorem, as long as we can find a point (a, Ψ^+) with $I_{\nu,p}(a, \Psi^+) > 0$.

We now setup our min-max level. In [ES01], for a finite-dimensional subspace F of E_c^+ , using the dependence of $\mathcal{E}^{\text{HF}}(\Psi)$ only on the subspace spanned by Ψ , the min-max was defined over continuous deformations of

$$D(F) = \{ \Psi^+ \in F^N, \text{Gram } \Psi^+ \leq 1 \}$$

equivariant with respect to unitary transformation.

In our case, the $\mathcal{U}(K)$ symmetry has to act on a and Ψ both, by way of the group action $U \cdot (a, \Psi)$ defined in (11), and the definition of the deformed set has to be modified. Recall the definition (20) of $D(a_0, \Psi_0)$:

$$D(a_0, \Psi_0) = \{ (a, \Psi) \in S_\gamma \times E^K, \text{ there is } U \in \mathcal{U}(K), R \in \mathcal{R}(K) \text{ such that} \\ (a, \Psi) = U \cdot (a_0, R\Psi_0) \}.$$

We now define the set of deformations : a homotopy

$$h \in C \left([0, 1] \times \left(S_\gamma \times (E_c^+)^K \right), \left(S_\gamma \times (E_c^+)^K \right) \right)$$

is admissible if it preserves the boundary of A^+ and is equivariant with respect to the $\mathcal{U}(K)$ group action :

$$\begin{cases} h(\lambda, a, \Psi^+) \in S_\gamma \times \partial A^+ & \forall a \in S_\gamma, \Psi^+ \in \partial A^+, \\ h(\lambda, U \cdot (a, \Psi^+)) = U \cdot h(\lambda, (a, \Psi^+)) & \forall U \in \mathcal{U}(K), \lambda \in [0, 1], a \in S_\gamma, \Psi^+ \in (E_c^+)^K. \end{cases}$$

Let now $\mathcal{Q}(F)$ be the admissible deformations of $D(a_0, \Psi_0^+)$:

$$\mathcal{Q}(a_0, \Psi_0^+) = \{Q \subset S_\gamma \times A^+, \text{ there is } h \text{ admissible such that} \quad (34)$$

$$h(0, \cdot) = \text{Id}, h(1, D(a_0, \Psi_0^+)) = Q\}. \quad (35)$$

We can now define the min-max level

$$c_{\nu,p}(a_0, \Psi_0^+) = \inf_{Q \in \mathcal{Q}(a_0, \Psi_0^+)} \sup_{(a, \Psi^+) \in Q} J_{\nu,p}(a, \Psi^+). \quad (36)$$

It verifies the following properties:

Lemma 5. *Let $\gamma > 0, N < Z + 1$. Then there is a constant $c_0 > 0$ such that, for all $(a_0, \Psi_0^+) \in S_\gamma \times \Sigma \cap (E_c^+)^K$, there are $\underline{a} > 0, p_0$ such that, for $c \geq c_0, \nu \in (0, 1), p \geq p_0$,*

$$0 < \underline{a} \leq c_{\nu,p}(a_0, \Psi_0^+) \leq \sup \left\{ \mathcal{F}_{\nu,p}(a, \Psi^+ + \Psi^-), (a, \Psi^+) \in D(a_0, \Psi_0^+), \right. \\ \left. \Psi^- \in (E_c^-)^K, \text{Gram}(\Psi^+ + \Psi^-) \leq 1 \right\}.$$

Proof. The upper bound is immediately obtained by using the trivial deformation $h(\lambda, (a, \Psi^+)) = (a, \Psi^+)$ in (36). For the lower bound, it results from the intersection result of Lemma 5.4 in [ES01] that, for every path $Q \in \mathcal{Q}(a_0, \Psi_0^+)$, there is $(a, \Psi^+) \in Q$ with $\text{Gram} \Psi^+ = 1/2$. For such a point,

$$\begin{aligned} I_{\nu,p}(a, \Psi^+) &\geq \mathcal{E}_{\nu,p}(a, \Psi^+) - \pi_p(\Psi^+) \\ &\geq \langle \Psi^+, (D_c + V)\Gamma_a \Psi^+ \rangle - \left(\frac{1}{2}\right)^{p-1} \\ &\geq \frac{\gamma K \lambda_1(c)}{2} - \left(\frac{1}{2}\right)^{p-1}, \end{aligned}$$

for c large enough, where $\lambda_1(c)$ is the (positive) first eigenvalue of $\sqrt{c^4 - c^2 \Delta} + V$. Since $\lambda_1(c)$ is increasing with c , the result follows. \square

Using now arguments by Fang and Ghoussoub [FG92, Gho93], we obtain Palais-Smale sequences with second order information.

Lemma 6. *Let $\gamma > 0, N < Z + 1$. There is c_0 such that, for all $c \geq c_0, (a_0, \Psi_0^+) \in S_\gamma \times (\Sigma \cap (E_c^+)^K), \nu \in (0, 1), p \geq p_0$, there exists a sequence $(a_n, \Psi_n^+) \in S_\gamma \times A^+$, such that*

- (1) $I_{\nu,p}(a_n, \Psi_n^+) \rightarrow c_{\nu,p}(a_0, \Psi_0^+)$,
- (2) $\partial_\Psi I_{\nu,p}(\Psi_n^+) \rightarrow 0$,
- (3) $\liminf \text{Gram} \Psi_n^+ \geq d(\nu) > 0$,
- (4) *There is a sequence $\delta_n \rightarrow 0$ such that $\partial_{\Psi^+}^2 I_{\nu,p}(\Psi_n^+)[\Psi, \Psi] + \delta_n \|\Psi\|_{H^{1/2}}^2$ has a negative space of dimension at most K^2 ,*
- (5) $a_n \rightarrow a \in S_\gamma$, and $\partial_a I_{\nu,p}(a_n, \Psi_n) = 0$.

It follows from Lemmas 3 and 1 that there is a critical point of $\mathcal{F}_{\nu,p}$ on $S_\gamma \times A$ at level $c_{\nu,p}(a_0, \Psi_0^+)$.

Although $c_{\nu,p}$ is defined using a particular (a_0, Ψ_0^+) , Sard's theorem forbids continuous families of critical points. Therefore, we have

Lemma 7. $c_{\nu,p}(a_0, \Psi_0^+)$ does not depend on (a_0, Ψ_0^+) .

To prove this, we need an infinite-dimensional version of Sard's theorem :

Lemma 8. *Let $F \in C^\infty(X, \mathbb{R})$, with X a separable Hilbert manifold. Then the set $\mathcal{C} = \{y \in \mathbb{R}, \exists x \in E, F(x) = y, dF(x) = 0, d^2 F(x) \text{ is Fredholm}\}$ is of measure zero.*

The proof of this theorem can be found in [FKN⁺74] (Corollary 2.1). The main idea is to use a Lyapunov-Schmidt reduction in a neighborhood of every critical point, which, by the classical Sard theorem, proves that the set of critical points in this neighborhood is of measure zero. Then, using the separability of X , one can piece together these informations and obtain the result.

We can now prove Lemma 7 :

Proof. Let $Y = c_{\nu,p}(S_\gamma \times (\Sigma \cap (E_c^+)^K))$. By the previous construction, for all $y \in Y$, there is a critical point of $\mathcal{F}_{\nu,p}$ at level y satisfying the inequality $\Lambda \leq (c^2 - K_2)\Gamma$. It follows that $d^2 \mathcal{E}$ is Fredholm on this critical point, and therefore, by Lemma 8, that Y is of measure zero. But since $c_{\nu,p}$ is continuous on $(S_\gamma \times \Sigma \cap (E_c^+)^K)$, Y is reduced to a point. \square

Note that we do not actually need this to prove our main theorem (Corollary 1). We could use a particular (a_0, Ψ_0^+) chosen to make the $c \rightarrow \infty$ limit converge to a nonrelativistic minimizer (see section 6). This lemma only makes the result cleaner.

5.6. Proof of Theorem 1. We can now prove Theorem 1. Take $p \geq p_0, \nu = 1/p$. By Lemma 6 then 3, there is a sequence that satisfies the hypotheses of Lemma 1, and therefore converges to (a_p, Ψ_p) .

Because $\nabla \pi_p(\Psi_p) = \Lambda_p > (c^2 - K_1)\Gamma_p$, and $\nabla \pi_p(\Psi_p) \rightarrow 0$ uniformly if $\liminf \text{Gram } \Psi_p < 1$, we conclude that $\lim \text{Gram } \Psi_p = 1$.

We now apply again Lemma 1 to the sequence (a_p, Ψ_p) , and obtain strong convergence to $(a, \Psi) \in S_\gamma \times \Sigma$. Because $\partial_a I_{\nu,p}(a_p, \Psi_p) = 0$, if $\det \Gamma_a > 2\gamma$, we obtain $\mathcal{H}_{\Psi} a = E a$.

6. NONRELATIVISTIC LIMIT

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

Lemma 9. *Let $\gamma > 0, c_n \rightarrow \infty, (a_n, \Psi_n) \in S_\gamma \times \Sigma$ solutions of*

$$H_{a_n, \Psi_n}^{c_n} \Psi_n = \Lambda_n \Psi_n$$

such that

$$(c_n^2 - K_1)\Gamma_n \leq \Lambda_n \leq (c_n^2 - K_2)\Gamma_n$$

for constants $K_1, K_2 > 0$.

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

6.1. Proof of Lemma 9.

Proof. First, we need an upper bound on Ψ_n in H^1 , which we obtain using the Hardy inequality.

$$\begin{aligned} c_n^4 \|\Gamma_n \Psi_n\|_{L^2}^2 + c_n^2 \|\Gamma_n \nabla \Psi_n\|_{L^2}^2 &= \|H_0 \Gamma_n \Psi_n\|_{L^2}^2 \\ &= \|(V\Gamma + 2W_{a_n, \Psi_n})\Psi_n - \Lambda_n \Psi_n\|_{L^2}^2 \\ &\leq c_n^4 \|\Gamma_n \Psi_n\|_{L^2}^2 + K_3 \|\nabla \Psi_n\|_{L^2}^2 + K_4 c_n^2 \|\Gamma_n \nabla \Psi_n\|_{L^2}^2, \end{aligned}$$

with constants $K_3, K_4 > 0$ independent of n . Therefore, because $\det \Gamma_n > \gamma$, 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 Dirac-Fock equations as

$$c_n \Gamma_n L \mathcal{X}_n + (V \Gamma_n + 2W_{a_n, \Psi_n}) \Phi_n = (\Lambda_n - c_n^2 \Gamma_n) \Phi_n, \quad (37)$$

$$c_n \Gamma_n L \Phi_n + (V \Gamma_n + 2W_{a_n, \Psi_n}) \mathcal{X}_n = (\Lambda_n + c_n^2 \Gamma_n) \mathcal{X}_n, \quad (38)$$

with the operator

$$L = -i \nabla \cdot \sigma \quad (39)$$

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

Because $\liminf \Lambda_n - c_n^2 \Gamma_n < -\infty$, using the Hardy inequality and the boundedness of Φ_n in H^1 , the first equation (37) yields

$$\|\Gamma_n L \mathcal{X}_n\|_{L^2} = \|\Gamma_n \nabla \mathcal{X}_n\|_{L^2} = O(1/c_n) \quad (40)$$

The second equation (38) gives

$$\begin{aligned} \mathcal{X}_n &= \frac{1}{2c} \left(\frac{1}{2} (\Gamma_n + \Lambda_n / c_n^2) \right)^{-1} \Gamma_n L \Phi_n + \frac{1}{c_n^2} O(\|\mathcal{X}_n\|_{H^1}) \\ &= \text{KB}(\Phi_n) + \frac{1}{c_n^2} O(\|\mathcal{X}_n\|_{H^1}) + O\left(\frac{1}{c_n^3}\right) \end{aligned} \quad (41)$$

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

$$\text{KB}(\Phi) = \frac{1}{2c} L \Phi. \quad (42)$$

Equation (41) gives $\|\mathcal{X}_n\|_{L^2} = \frac{1}{2c} \|L \Phi_n\|_{L^2} + O(1/c_n^2) = O(1/c_n)$, and then

$$\mathcal{X}_n = \text{KB}(\Phi_n) + O\left(\frac{1}{c_n^3}\right) \quad (43)$$

again in L^2 norm. Φ_n satisfies

$$\left(-\frac{1}{2} \Delta \Gamma_n + V \Gamma_n + 2W_{\Phi_n} \right) \Phi_n = (\Lambda_n - c_n^2 \Gamma_n) \Phi_n + h_n$$

$$\text{Gram } \Phi_n = 1 + o(1)$$

with $h_n \rightarrow 0$ in L^2 and therefore H^{-1} norm. (a_n, Φ_n) is a Palais-Smale sequence for the nonrelativistic functional, with control on the Lagrange multipliers $\limsup (\Lambda_n - c_n^2 \Gamma_n) < 0$ and non-degeneracy information $\det \Gamma_n > \gamma$. 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 (43) that

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

□

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 9 : up to a subsequence, it converges strongly in H^1 to $\left(a, \begin{pmatrix} \Phi \\ 0 \end{pmatrix}\right)$, with $\lim \mathcal{E}(a_n, \Psi_n) - Nc_n^2 = \mathcal{E}^{\text{HF}}(a, \Phi)$, so that $\mathcal{E}^{\text{HF}}(a, \Phi) + \theta_\gamma(\det \Gamma_a) \geq I_\gamma^K$. We now prove that $\mathcal{E}^{\text{HF}}(a, \Phi) + \theta_\gamma(\det \Gamma_a) \leq I_\gamma^K$.

By arguments of [Lew04], there is a $(a_*, \Phi_*) \in S_\gamma \times H^1(\mathbb{R}^3, \mathbb{C}^2)$ which is a minimizer of the nonrelativistic problem I_γ^K . Set

$$\Psi_{*,n} = \left(\text{Gram } P_{c_n}^+ \begin{pmatrix} \Phi_* \\ 0 \end{pmatrix} \right)^{-1/2} \left(P_{c_n}^+ \begin{pmatrix} \Phi_* \\ 0 \end{pmatrix} \right). \quad (44)$$

$\Psi_{*,n}$ belongs to $\Sigma \cap (E_{c_n}^+)^K$, and converges to $\begin{pmatrix} \Phi_* \\ 0 \end{pmatrix}$.

By the property of Theorem 1, the critical points (a_n, Ψ_n) satisfy

$$\mathcal{E}(a_n, \Psi_n) + \theta_\gamma(\det \Gamma_{a_n}) \leq \sup \left\{ \mathcal{E}(a, \Psi^+ + \Psi^-) + \theta_\gamma(\det \Gamma_a), \right. \\ \left. (a, \Psi^+) \in D(a_*, \Psi_{*,n}), \Psi^- \in (E_{c_n}^-)^K, \text{Gram}(\Psi^+ + \Psi^-) \leq 1 \right\} \quad (45)$$

Let $a \in S_\gamma, \Psi^\pm \in (E_{c_n}^\pm)^K$ be such that $\text{Gram}(\Psi^+ + \Psi^-) \leq 1$. By concavity of \mathcal{E} in the $E_{c_n}^-$ directions (see Lemma 2), for large n we have

$$\mathcal{E}(a, \Psi^+ + \Psi^-) \leq \mathcal{E}(a, \Psi^+) + \partial_\Psi \mathcal{E}(a, \Psi^+) \cdot \Psi^- - \frac{1}{4} \left\langle \Psi^-, \sqrt{c_n^4 - c_n^2 \Delta} \Gamma_a \Psi^- \right\rangle \\ \leq \mathcal{E}(a, \Psi^+) + M \left\| \Psi^- \right\|_{L^2} - c_n^2 \gamma \left\| \Psi^- \right\|_{L^2}^2,$$

where $M > 0$ is independent of n . For n large, we have for all $a \in S_\gamma$

$$\mathcal{E}(a_n, \Psi_n) \leq \sup_{(a, \Psi^+) \in D(a_*, \Psi_{*,n}), \text{Gram } \Psi^+ \leq 1} \mathcal{E}(a, \Psi^+), \quad (46)$$

By strict convexity of \mathcal{E} in the $E_{c_n}^+$ directions for n large, the sup is achieved at a point (a, Ψ_{\max}^+) such that $\text{Gram } \Psi_{\max}^+ = 1$. Therefore,

$$\mathcal{E}(a_n, \Psi_n) \leq \sup_{(a, \Psi^+) \in U \cdot (a_*, \Psi_{*,n}), U \in \mathcal{U}(K)} \mathcal{E}(a, \Psi^+) \\ \leq \mathcal{E}(a_*, \Psi_{*,n}). \quad (47)$$

We now use the inequality that, for all $\psi \in H^1(\mathbb{R}^3, \mathbb{C}^2)$,

$$\left\langle \psi, \sqrt{c_n^4 - c_n^2 \Delta} \psi \right\rangle \leq \left\langle \Psi, \left(c_n^2 - \frac{1}{2} \Delta \right) \Psi \right\rangle.$$

This is a simple consequence of the concavity inequality $\sqrt{1+x} \leq 1 + \frac{1}{2}x$ in Fourier domain.

Therefore,

$$\mathcal{E}^{\text{HF}}(a, \Phi) + \theta_\gamma(\det \Gamma_a) = \lim \mathcal{E}(a_n, \Psi_n) - Nc_n^2 + \theta_\gamma(\det \Gamma_{a_n}) \\ \leq \limsup \mathcal{E}(a_*, \Psi_{*,n}) - Nc_n^2 + \theta_\gamma(\det \Gamma_{a_*}) \\ \leq \mathcal{E}^{\text{HF}}(a_*, \Phi_*) + \theta_\gamma(\det \Gamma_{a_*}) \\ = I_\gamma^K.$$

□

ACKNOWLEDGEMENTS

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

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.
- [Der12] J. Dereziński. 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 Poincaré*, 2:941–961, 2001.
- [FG92] G. Fang and N. Ghoussoub. Second order information on Palais-Smale sequences in the mountain pass theorem. *Manuscripta mathematica*, 75(1):81–95, 1992.
- [FKN⁺74] S. Fučík, M. Kučera, J. Nečas, J. Souček, and V. Souček. Morse-Sard theorem in infinite dimensional Banach spaces and investigation of the set of all critical levels. *Časopis pro pěstování matematiky*, 99(3):217–243, 1974.
- [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.
- [Gho93] N. Ghoussoub. *Duality and perturbation methods in critical point theory*, volume 107. Cambridge University Press, 1993.
- [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