C. R. Acad. Sci. Paris, Ser. I 334 (2002) 299–304

Équations aux dérivées partielles/*Partial Differential Equations*
(Physique mathématique/*Mathematical Physics*)

The multiconfiguration methods in quantum chemistry: Palais–Smale condition and existence of minimizers

Mathieu Lewin

CEREMADE, Université Paris IX Dauphine, place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16, France

Received 28 November 2001; accepted 17 December 2001

Note presented by Pierre-Louis Lions.

Abstract In this Note, we propose a new proof for the existence of a minimum in the multiconfiguration methods in Quantum Chemistry. We use a Palais–Smale condition with Morse-type information, whose proof is based on the Euler–Lagrange equations, written in a simple and useful way. ***To cite this article: M. Lewin, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 299–304.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Les méthodes de multiconfiguration en chimie quantique : condition de Palais–Smale et existence de minima

Résumé Dans cette Note, nous proposons une nouvelle preuve de l'existence d'un minimum pour les méthodes de multiconfiguration en Chimie Quantique. Nous utilisons une propriété de Palais–Smale (avec information de type Morse), dont la démonstration repose sur les équations d'Euler–Lagrange écrites sous une forme compacte aisément utilisable. ***Pour citer cet article : M. Lewin, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 299–304.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Version française abrégée

L'énergie quantique non relativiste de N électrons en présence de M noyaux ponctuels et immobiles est

$$\langle \Psi, \mathcal{H}\Psi \rangle = \sum_{i=1}^{N} \int_{\mathbb{R}^{3N}} \left(\frac{1}{2} \left| \nabla_{x_i} \Psi(x) \right|^2 + V(x_i) \left| \Psi(x) \right|^2 \right) \mathrm{d}x + \sum_{1 \leqslant i < j \leqslant N} \int_{\mathbb{R}^{3N}} \frac{|\Psi(x)|^2}{|x_i - x_j|} \, \mathrm{d}x.$$

Ici $x = (x_1, \ldots, x_N) \in (\mathbb{R}^3)^N$, $V(u) = -\sum_{m=1}^{M} \frac{z_m}{|u - \tilde{x}_m|}$ où $u \in \mathbb{R}^3$, et $\Psi \in \mathrm{H}^1_a(\mathbb{R}^{3N}, \mathbb{R})$ vérifie $\|\Psi\|_{\mathrm{L}^2} = 1$ (l'indice a signifie que Ψ est antisymétrique par rapport aux permutations de variables). Pour simplifier, nous avons négligé le spin et supposé que les Ψ sont à valeurs réelles. Nous noterons $Z = \sum_{m=1}^{M} z_m$ la charge totale des noyaux.

E-mail address: lewin@ceremade.dauphine.fr (M. Lewin).

S1631-073X(02)02252-5/FLA

Dans le modèle de multiconfiguration de rang K, on se limite à des fonctions d'onde qui sont des sommes finies de déterminants de Slater, formés à partir de K fonctions (ϕ_i) de $\mathrm{H}^1(\mathbb{R}^3)$ vérifiant $\int_{\mathbb{R}^3}\phi_i\phi_j=\delta_{ij}$:

$$\Psi=\sum_{I=\{i_1<i_2<\cdots<i_N\}\subset\{1,\dots,K\}} c_I\cdot|\phi_{i_1},\dots,\phi_{i_N}\rangle$$

où $|\phi_{i_1},\dots,\phi_{i_N}\rangle(x_1,\dots,x_N)=\frac{1}{\sqrt{N!}}\det(\phi_{i_k}(x_l))_{k,l}$ et $\sum_I {c_I}^2=1$. Ceci est équivalent à $\mathrm{rg}(\Psi):=\mathrm{rg}(\hat{\gamma}_\Psi)\leqslant K$, où $\hat{\gamma}_\Psi$ est l'opérateur de densité d'ordre 1 (*voir* [9,2]). Ce modèle est justifié par le fait que toute fonction d'onde peut être développée en une combinaison linéaire (éventuellement infinie) de déterminants de Slater.

L'énergie fondamentale de rang K est alors définie par

$$E_N^K=\inf\big\{\langle\Psi,\mathcal{H}\Psi\rangle,\ \Psi\in\mathrm{H}^1_a\big(\mathbb{R}^{3N}\big),\ \|\Psi\|_{\mathrm{L}^2(\mathbb{R}^{3N})}=1,\ \mathrm{rg}(\Psi)\leqslant K\big\}. \tag{1}$$

Pour des raisons algébriques, il n'existe pas de fonctions d'onde de rang $K=N+1$ et, lorsque $N=2$, elles sont toutes de rang pair. Si $K=N$, on retrouve le modèle de Hartree–Fock abondamment étudié dans la littérature [7,8,4]. Un cas particulier de $K=N+2$ est étudié par C. Le Bris dans [5] qui considère des sommes de deux déterminants et démontre l'inégalité $E_N^{N+2}<E_N^N$ (la généralisation $E_N^{K+2}<E_N^K$ pour tout $K\geqslant N$ peut être trouvée dans [3]). L'existence d'un minimum pour tout $K\geqslant N$ a récemment été obtenue par G. Friesecke dans [2]. Sa preuve, très intéressante, est basée sur des méthodes de localisation géométrique, un argument de type concentration-compacité de P.-L. Lions et l'inégalité $E_{N+1}^{K+1}<E_N^K$. Malheureusement, nous pensons que sa méthode n'est pas adaptée à la recherche de points critiques et c'est pourquoi nous avons cherché une autre preuve, dans l'esprit de [8].

Notre démonstration est basée sur les équations d'Euler–Lagrange, peu étudiées jusqu'à présent à cause de leur apparente complexité, mais que nous avons écrites sous une forme compacte aisément utilisable. Nous démontrons ainsi l'existence d'un minimum Ψ pour tout $K\geqslant N$. La formulation choisie pour le théorème 1 est motivée par une étude en cours [6] concernant l'existence de points critiques qui ne seraient pas des minima.

Si on pose $\alpha_{i_1,\dots,i_N}=0$ lorsque $\#\{i_1,\dots,i_N\}<N$ et $\alpha_{i_1,\dots,i_N}=\frac{\varepsilon(\sigma)}{\sqrt{N!}}c_{\{i_{\sigma(1)}<\cdots<i_{\sigma(N)}\}}$ sinon, on a alors $\Psi=\sum_{i_1,\dots,i_N}\alpha_{i_1,\dots,i_N}\phi_{i_1}\otimes\cdots\otimes\phi_{i_N}$. Avec $\Phi=(\phi_1,\dots,\phi_K)^{\mathrm{T}}$ et $c=(c_I)\in\mathbb{R}^{\binom{K}{N}}$, on obtient

$$\langle\Psi,\mathcal{H}\Psi\rangle=\mathcal{E}(c,\Phi):=\left\langle\left(\left(-\frac{\Delta}{2}+V\right)\Gamma+W_\Phi\right)\cdot\Phi,\ \Phi\right\rangle_{(\mathrm{L}^2(\mathbb{R}^3))^K}, \tag{2}$$

où les matrices Γ et W_Φ sont définies par $\Gamma_{i,j}=N\sum_{k_2,\dots,k_N}\alpha_{i,k_2,\dots,k_N}\alpha_{j,k_2,\dots,k_N}$ et $(W_\Phi)_{i,j}(x)=\frac{N(N-1)}{2}\sum_{k_3,\dots,k_N}\sum_{k,l}\alpha_{i,k,k_3,\dots,k_N}\alpha_{j,l,k_3,\dots,k_N}((\phi_k\phi_l)*\frac{1}{|r|})(x)$.

$\mathcal{E}$ est définie sur $\mathcal{M}_N^K=\big\{(c,\Phi)\in\mathbb{R}^{\binom{K}{N}}\times(\mathrm{H}^1(\mathbb{R}^3))^K,\ \sum_I {c_I}^2=1,\ \int_{\mathbb{R}^3}\phi_i\phi_j=\delta_{ij}\big\}$ par la formule (2). Nous noterons $d\mathcal{E}$ sa dérivée première, et $d^2_\Phi\mathcal{E}$ sa dérivée seconde par rapport à Φ, sur la variété Riemannienne $\mathcal{M}_N^K$. Remarquons que $\mathcal{E}$ est faiblement semi-continue inférieurement lorsqu'on l'étend à $\mathbb{R}^{\binom{K}{N}}\times(\mathrm{H}^1(\mathbb{R}^3))^K$.

THÉORÈME 1 (Condition de Palais–Smale avec information de type Morse). – *On suppose $Z>N-1$ et $K\geqslant N$. Soit $(c^n,\Phi^n)_{n\in\mathbb{N}}$ une suite de $\mathcal{M}_N^K$ telle que* :

(1) $(\mathcal{E}(c^n,\Phi^n))_{n\in\mathbb{N}}$ *est bornée* ;

(2) $\mathrm{d}\mathcal{E}(c^n,\Phi^n)\to 0$;

(3) *il existe $j\in\mathbb{N}$ et une suite $(\delta^n)_{n\in\mathbb{N}}$ de réels strictement positifs tendant vers* 0 *telle que, pour chaque n, $\mathrm{d}^2_\Phi\mathcal{E}(c^n,\Phi^n)$ a au plus j valeurs propres inférieures à $-\delta^n$.*

Alors, la suite des fonctions d'onde associées $(\Psi^n)_{n\in\mathbb{N}}$ est relativement compacte dans $\mathrm{H}^1_a(\mathbb{R}^{3N})$ et converge (à une sous-suite près) vers Ψ avec $\mathrm{rg}(\Psi)=K$ ou $K-1$.

To cite this article: M. Lewin, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 299–304

De plus, si $\Psi = \sum_{I\subset\{1,\dots,K'\}} c_I |\phi_{i_1},\dots,\phi_{i_N}\rangle$ *où* $K' = \mathrm{rg}(\Psi)$, *on a* (*avec des notations évidentes*) :

$$\left(\left(-\frac{\Delta}{2}+V\right)\Gamma + 2W_\Phi\right)\cdot\Phi + \Lambda\cdot\Phi = 0 \quad \textit{dans } \left(\mathrm{L}^2\left(\mathbb{R}^3\right)\right)^{K'}, \tag{3}$$

où $\Lambda > 0$ *est la matrice des multiplicateurs de Lagrange associés aux contraintes sur les* ϕ_i.

COROLLAIRE 2 (État fondamental pour les méthodes de multiconfiguration). – *On suppose* $Z > N-1$. *Pour tout* $K \geqslant N$, *il existe un minimum* Ψ *pour* E_N^K, *et on a* $\mathrm{rg}(\Psi) = K$ *ou* $K-1$.

Remarques. – (i) Lorsque $K = N$ (Hartree–Fock), on a $\Gamma = I_N$ et on peut ainsi diagonaliser Λ et Γ simultanément, ce qui permet de ramener (3) à un système d'équations scalaires en partie découplées, et simplifie la démonstration. Dans le cas général c'est impossible et dans les preuves nous diagonalisons Γ ou Λ selon le cas.

(ii) Le théorème 1 est une étape importante dans la recherche d'autres points critiques (travail en cours [6]). Les techniques exposées par exemple dans [4] permettent en effet d'obtenir des suites vérifiant les hypothèses du théorème, à partir de min-max appropriés.

(iii) On peut voir que si (c,Φ) est une solution de (3), alors les ϕ_i sont dans $\bigcap_{2\leqslant p<3} \mathrm{W}^{2,p}(\mathbb{R}^3)$ et sont analytiques réelles sur $\mathbb{R}^3\setminus\{\bar{x}_1,\dots,\bar{x}_M\}$. Si de plus $\Lambda > 0$, alors elles décroissent exponentiellement à l'infini.

---

1. Introduction

The non-relativistic quantum energy of N electrons interacting with M static nucleii is given by:

$$\langle\Psi,\mathcal{H}\Psi\rangle = \sum_{i=1}^{N}\int_{\mathbb{R}^{3N}}\left(\frac{1}{2}\left|\nabla_{x_i}\Psi(x)\right|^2 + V(x_i)\left|\Psi(x)\right|^2\right)\mathrm{d}x + \sum_{1\leqslant i<j\leqslant N}\int_{\mathbb{R}^{3N}}\frac{|\Psi(x)|^2}{|x_i-x_j|}\,\mathrm{d}x.$$

Here $x = (x_1,\dots,x_N)\in(\mathbb{R}^3)^N$, $V(u) = -\sum_{m=1}^{M}\frac{z_m}{|u-\bar{x}_m|}$ where $u\in\mathbb{R}^3$, and $\Psi\in\mathrm{H}^1_a(\mathbb{R}^{3N},\mathbb{R})$ is such that $\|\Psi\|_{\mathrm{L}^2}=1$. The subscript a on H^1_a indicates that we consider functions Ψ which are antisymetric under interchanges of variables: $\forall\sigma\in S_N,\ \Psi(x_1,\dots,x_N) = \epsilon(\sigma)\Psi(x_{\sigma(1)},\dots,x_{\sigma(N)})$. For the sake of simplicity, we restrict ourselves to real-valued functions and do not take the spin into account. In the sequel, we denote by $Z = \sum_{m=1}^{M} z_m$ the total nuclear charge.

In the multiconfiguration method of rank K, the set of admissible wave functions is limited to the Ψ which are a linear combination of Slater determinants built with K functions $\phi_1,\dots,\phi_K$ in $\mathrm{H}^1(\mathbb{R}^3)$ satisfying $\int_{\mathbb{R}^3}\phi_i\phi_j = \delta_{ij}$:

$$\Psi = \sum_{I=\{i_1<i_2<\cdots<i_N\}\subset\{1,\dots,K\}} c_I\cdot|\phi_{i_1},\dots,\phi_{i_N}\rangle$$

where $|\phi_{i_1},\dots,\phi_{i_N}\rangle(x_1,\dots,x_N) = \frac{1}{\sqrt{N!}}\det(\phi_{i_k}(x_l))_{k,l}$ and $\sum_I {c_I}^2 = 1$. This is equivalent (*see* [9,2]) to $\mathrm{rank}(\Psi) := \mathrm{rank}(\hat{\gamma}_\Psi)\leqslant K$ and $\mathrm{Range}(\hat{\gamma}_\Psi)\subset\mathrm{span}(\phi_i)$ where $\hat{\gamma}_\Psi$ is the *one-body density matrix*: this is the operator acting on $\mathrm{L}^2(\mathbb{R}^3)$ with kernel $\gamma_\Psi(x,y) = N\int_{\mathbb{R}^{3(N-1)}}\Psi(x,x_2,\dots,x_N)\Psi(y,x_2,\dots,x_N)\,\mathrm{d}x_2\cdots\mathrm{d}x_N$. Every wave function can be expressed as a (infinite) linear combination of Slater determinants, which justifies the model.

The N-body ground state energy of rank K is defined by

$$E_N^K = \inf\left\{\langle\Psi,\mathcal{H}\Psi\rangle,\ \Psi\in\mathrm{H}^1_a\left(\mathbb{R}^{3N}\right),\ \|\Psi\|_{\mathrm{L}^2(\mathbb{R}^{3N})}=1,\ \mathrm{rank}(\Psi)\leqslant K\right\}. \tag{1}$$

For algebraic reasons (*see* [2,6]), there does not exist N-body wave functions of rank $K = N+1$ and when $N=2$, all the Ψ have an even rank. When $K=N$, we obtain the well-known *Hartree–Fock*

model [7,8,4]. A simplification of the case $K = N + 2$ was studied by Le Bris [5] who considered the minimization over doubly excited configurations $\Psi = \alpha|\phi_1, \dots, \phi_N\rangle + \beta|\phi_1, \dots, \phi_{N-2}\phi_{N+1}\phi_{N+2}\rangle$. He proved the existence of a minimum and the inequality $E_N^{N+2} < E_N^N$ (the generalization $E_N^{K+2} < E_N^K$ for all $K \geqslant N$ may be found in [3]). The existence of a minimum for all $K \geqslant N$ was recently proved by Friesecke in [2]. He used a very interesting proof based on geometric localization methods, P.-L. Lions' concentration-compactness ideas and the inequality $E_{N+1}^{K+1} < E_N^K$. Unfortunately, we think that his method cannot easily be generalized to obtain saddle points, as in [8] for the Hartree–Fock case. That is why we have worked on a new proof, exposed in this Note.

Our method is based on the Euler–Lagrange equations (as in [8]), which were often neglected in the previous studies because of their apparent complexity. We express them in a very simple and useful way (3) and obtain the existence of a ground state Ψ for all $K \geqslant N$. Our formulation of Theorem 1 is motivated by a work in preparation [6] about the existence of saddle points.

2. The results

If we use the convention $\alpha_{i_1,\dots,i_N} = 0$ if $\#\{i_1, \dots, i_N\} < N$ and $\alpha_{i_1,\dots,i_N} = \frac{\varepsilon(\sigma)}{\sqrt{N!}} c_{\{i_{\sigma(1)} < \dots < i_{\sigma(N)}\}}$ otherwise, we get $\Psi = \sum_{i_1,\dots,i_N} \alpha_{i_1,\dots,i_N} \phi_{i_1} \otimes \dots \otimes \phi_{i_N}$. With $\Phi = (\phi_1, \dots, \phi_K)^{\mathrm{T}}$ and $c = (c_I) \in \mathbb{R}^{\binom{K}{N}}$, we obtain

$$\langle \Psi, \mathcal{H}\Psi \rangle = \mathcal{E}(c, \Phi) := \left\langle \left(\left(-\frac{\Delta}{2} + V \right) \Gamma + W_\Phi \right) \cdot \Phi,\ \Phi \right\rangle_{(\mathrm{L}^2(\mathbb{R}^3))^K} \tag{2}$$

where the matrices Γ and W_Φ are defined by $\Gamma_{i,j} = N \sum_{k_2,\dots,k_N} \alpha_{i,k_2,\dots,k_N} \alpha_{j,k_2,\dots,k_N}$ and $(W_\Phi)_{i,j}(x) = \frac{N(N-1)}{2} \sum_{k_3,\dots,k_N} \sum_{k,l} \alpha_{i,k,k_3,\dots,k_N} \alpha_{j,l,k_3,\dots,k_N} ((\phi_k \phi_l) * \frac{1}{|r|})(x)$.

$\mathcal{E}$ is defined on $\mathcal{M}_N^K = \{(c, \Phi) \in \mathbb{R}^{\binom{K}{N}} \times (\mathrm{H}^1(\mathbb{R}^3))^K,\ \sum_I {c_I}^2 = 1,\ \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}\}$ by the formula (2). We shall denote by $d\mathcal{E}$ its first derivative, and $d_\Phi^2 \mathcal{E}$ its second derivative with regard to Φ, on the Riemannian manifold $\mathcal{M}_N^K$. It is easy to see that $\mathcal{E}$ is weakly lower semi-continuous when it is defined on $\mathbb{R}^{\binom{K}{N}} \times (\mathrm{H}^1(\mathbb{R}^3))^K$ (use Fatou's lemma for the term $\langle W_\Phi \cdot \Phi, \Phi \rangle$).

THEOREM 1 (Palais–Smale condition with Morse-type information). – *We assume $Z > N - 1$ and $K \geqslant N$. Let $(c^n, \Phi^n)_{n\in\mathbb{N}}$ be a sequence of $\mathcal{M}_N^K$ such that*

(1) *$(\mathcal{E}(c^n, \Phi^n))_{n\in\mathbb{N}}$ is bounded;*
(2) *$\mathrm{d}\mathcal{E}(c^n, \Phi^n) \to 0$;*
(3) *there exists $j \in \mathbb{N}$ and a sequence of positive real numbers $(\delta^n)_{n\in\mathbb{N}}$ with $\delta^n \to 0$, such that for every n, $\mathrm{d}_\Phi^2 \mathcal{E}(c^n, \Phi^n)$ has at most j eigenvalues below $-\delta^n$.*

Then the corresponding sequence of wave functions $(\Psi^n)_{n\in\mathbb{N}}$ is relatively compact in $\mathrm{H}_a^1(\mathbb{R}^{3N})$ and converges (up to a subsequence) to Ψ with $\mathrm{rank}(\Psi) = K$ or $K - 1$.

Moreover, if $\Psi = \sum_{I\subset\{1,\dots,K'\}} c_I |\phi_{i_1}, \dots, \phi_{i_N}\rangle$ with $K' = \mathrm{rank}(\Psi)$, then we have (with obvious notations):

$$\left(\left(-\frac{\Delta}{2} + V \right) \Gamma + 2W_\Phi \right) \cdot \Phi + \Lambda \cdot \Phi = 0 \quad \text{in } \left(\mathrm{L}^2(\mathbb{R}^3)\right)^{K'}, \tag{3}$$

where $\Lambda > 0$ is the Lagrange multipliers matrix, corresponding to the constraints on the ϕ_i.

COROLLARY 2 (Ground state for multiconfiguration methods). – *Let $Z > N - 1$. For all $K \geqslant N$, there exists a minimum Ψ for E_N^K, with $\mathrm{rank}(\Psi) = K$ or $K - 1$.*

Remarks. – (i) Let us recall that any orthonormal basis of $\mathrm{Range}(\hat{\gamma}_\Psi)$ can be chosen to represent Ψ as a linear combination of Slater determinants. If $\Phi' = U \cdot \Phi$ where $U \in \mathcal{O}_K(\mathbb{R})$, then we have $\Gamma' = U\Gamma U^{\mathrm{T}}$, $W'_{\Phi'} = U W_\Phi U^{\mathrm{T}}$ and $\Lambda' = U\Lambda U^{\mathrm{T}}$. In the Hartree–Fock theory, one has $\Gamma = I_N$ and so one may suppose that Γ and Λ are both diagonal: this "crucial orthogonal invariance of the energy" (*see* [5]) simplifies the

arguments since the problem splits into N less coupled scalar equations. In our proofs, we suppose either Γ or Λ diagonal, according to what we need.

(ii) Theorem 1 is the first step towards the existence of saddle points (work [6] in preparation). One can obtain sequences which satisfy the assumptions of Theorem 1, thanks to the techniques developed for instance in [4]. They correspond to suitable min-maxing methods.

(iii) It is easy to see that if (c, Φ) is a solution of (3), then each ϕ_i is in $\bigcap_{2\leqslant p<3} \mathrm{W}^{2,p}(\mathbb{R}^3)$ and is real-analytic on $\mathbb{R}^3 \setminus \{\bar{x}_1, \ldots, \bar{x}_M\}$. If moreover $\Lambda > 0$, then it also has an exponential fall-off.

3. Proofs

Proof of Theorem 1. – Our method of proof is in the spirit of [8].

We take (c^n, Φ^n) as in the theorem and suppose that for all n, $\Gamma^n = \mathrm{diag}(\gamma_1^n, \ldots, \gamma_K^n)$, with $0 \leqslant \gamma_i^n \leqslant 1$ and $\gamma_1^n \geqslant \cdots \geqslant \gamma_K^n$. We may also assume that $c_I^n \to c_I$ for all $I \subset \{1, \ldots, K\}$, $|I| = N$. Thus each γ_i^n converges to some $\gamma_i \geqslant 0$. It is easy to see that, for all $i = 1, \ldots, K$, $\gamma_i^n \int_{\mathbb{R}^3} |\nabla \phi_i^n|^2$ is bounded. So either $\gamma_i = 0$ or $(\phi_i^n)_{n\in\mathbb{N}}$ is bounded in $\mathrm{H}^1(\mathbb{R}^3)$.

Step 1: Let us suppose for instance that $\gamma_K^n \to 0$. Since (ϕ_K^n) is bounded in $\mathrm{L}^2(\mathbb{R}^3)$, we may assume – by extracting subsequences if necessary – that $\sqrt{\gamma_K^n}\phi_K^n \to 0$ in $\mathrm{L}^2(\mathbb{R}^3)$ and $\sqrt{\gamma_K^n}\phi_K^n \rightharpoonup 0$ weakly in $\mathrm{H}^1(\mathbb{R}^3)$. We now show that this last convergence is strong.

Let $I \subset \{1, \ldots, K\}$ be such that $|I| = N$ and $K \in I$. Since $\mathcal{E}$ is quadratic with regard to the (c_I), we have

$$\mathcal{E}(c^n, \Phi^n) = \left(\sum_{i\in I} \int_{\mathbb{R}^3} \left|\nabla \phi_i^n\right|^2\right) \left(c_I^n\right)^2 + A_n \left(c_I^n\right)^2 + B_n c_I^n + C_n,$$

where $A_n (c_I^n)^2$ and $B_n c_I^n \to 0$: for instance one of the terms appearing in $A_n (c_I^n)^2$ is

$$\left(c_I^n\right)^2 \int_{\mathbb{R}^3} \frac{\phi_K^n(x)^2}{|x|}\mathrm{d}x \leqslant \int_{\mathbb{R}^3} \frac{(\sqrt{\gamma_K^n}\phi_K^n(x))^2}{|x|}\,\mathrm{d}x \to 0.$$

The second assumption of Theorem 1 now yields the existence of a bounded sequence of Lagrange multipliers $(\beta^n)_{n\in\mathbb{N}}$ such that

$$\frac{\partial \mathcal{E}}{\partial c_I}(c^n, \Phi^n) = 2\left(\sum_{i\in I} \int_{\mathbb{R}^3} \left|\nabla \phi_i^n\right|^2\right)\left(c_I^n\right) + 2A_n c_I^n + B_n + \beta^n c_I^n \to 0.$$

Multiplying by c_I^n, we obtain $\left(\sum_{i\in I} \int_{\mathbb{R}^3} |\nabla \phi_i|^2\right)(c_I^n)^2 \to 0$ for all I containing K and this shows that $\sqrt{\gamma_K^n}\phi_K^n \to 0$ strongly in $\mathrm{H}^1(\mathbb{R}^3)$.

Let $K' \geqslant N$ be such that $\gamma_{K'+1} = 0$ and $\gamma_{K'} \neq 0$. We introduce $\tilde{c}_I^n = (\sum_{J\subset\{1,\ldots,K'\}} (c_J^n)^2)^{-1/2} c_I^n$ for all $I \subset \{1, \ldots, K'\}$, and $\widetilde{\Phi}^n = (\phi_1^n, \ldots, \phi_{K'}^n)$ so that $(\tilde{c}^n, \widetilde{\Phi}^n) \in \mathcal{M}_N^{K'}$. If $(\widetilde{\Psi}^n)$ are the corresponding wave functions, the previous arguments imply that we have $\Psi^n = \widetilde{\Psi}^n + r^n$ where $r^n \to 0$ in $\mathrm{H}^1_a(\mathbb{R}^{3N})$. The new sequence $(\tilde{c}^n, \widetilde{\Phi}^n)_{n\in\mathbb{N}}$ satisfies the assumptions (1)–(3) of Theorem 1 on $\mathcal{M}_N^{K'}$ and also $\widetilde{\Gamma}^n \geqslant \gamma I_{K'}$ with $\gamma > 0$. We now work with this new Palais–Smale sequence and omit the $\sim$ for simplicity.

Step 2: The first order condition also yields the existence of a $K' \times K'$ symetric matrix Λ^n such that

$$\frac{\partial \mathcal{E}}{\partial \Phi}(c^n, \Phi^n) = \left(\left(-\frac{\Delta}{2} + V\right)\Gamma^n + 2W_{\Phi^n}\right) \cdot \Phi^n + \Lambda^n \cdot \Phi^n \to 0. \tag{4}$$

We now prove that there exists a $\lambda > 0$ such that $\Lambda^n \geqslant \lambda I_{K'}$. Let us assume for this step that $\Lambda^n = \mathrm{diag}(\lambda_1^n, \ldots, \lambda_{K'}^n)$. The second order condition of Theorem 1 now implies

$$\int_{\mathbb{R}^3} |\nabla \phi|^2 + \int_{\mathbb{R}^3} \left(V + \rho_i^n * \frac{1}{|r|} + \frac{\lambda_i^n}{\gamma_{ii}^n} + \varepsilon_i^n\right)|\phi|^2 \geqslant 0$$

for all ϕ in a closed subspace of $\mathrm{H}^1(\mathbb{R}^3)$ of codimension at most $K+j$, where $\varepsilon_i^n \to 0$ and ρ_i^n (an explicit function of the c_I^n and the ϕ_i^n) is such that $\int_{\mathbb{R}^3} \rho_i^n = N-1 < Z$. We then use the results of [8] to obtain a $\delta > 0$ such that $\lambda_i^n/\gamma_{ii}^n + \varepsilon_i^n \geqslant \delta$ and finally we get (up to subsequences) $\lambda_i^n \geqslant \gamma_{ii}^n(\delta/2) \geqslant \gamma(\delta/2) = \lambda > 0$.

Since it is easy to see that Λ^n is bounded, we finally obtain $\Lambda^n \to \Lambda > 0$.

Step 3: Since $\gamma_i^n \geqslant \gamma > 0$, the $(\phi_i^n)_{n\in\mathbb{N}}$ are bounded in $\mathrm{H}^1(\mathbb{R}^3)$ and thus we may assume that ϕ_i^n converges weakly in $\mathrm{H}^1(\mathbb{R}^3)$ (and a.e. in $\mathbb{R}^3$) to some ϕ_i. We now prove that this convergence is strong.

Passing to the limit in (4), we get

$$\left(\left(-\frac{\Delta}{2}+V\right)\Gamma + 2W_\Phi\right)\cdot\Phi + \Lambda\cdot\Phi = 0.$$

This implies

$$\begin{aligned}\limsup_{n\to+\infty}\left\langle \Lambda^n\cdot\Phi^n, \Phi^n\right\rangle &= -\liminf_{n\to+\infty}\left\langle\left(\left(-\frac{\Delta}{2}+V\right)\Gamma^n + 2W_{\Phi^n}\right)\cdot\Phi^n, \Phi^n\right\rangle \\ &\leqslant -\left\langle\left(\left(-\frac{\Delta}{2}+V\right)\Gamma + 2W_\Phi\right)\cdot\Phi, \Phi\right\rangle = \left\langle \Lambda\cdot\Phi, \Phi\right\rangle,\end{aligned}$$

by using the same arguments as the one used to show that $\mathcal{E}$ is weakly lower semi-continuous on $\mathbb{R}^{\binom{K}{N}} \times (\mathrm{H}^1(\mathbb{R}^3))^K$. So $\Phi^n \to \Phi$ in $(\mathrm{L}^2(\mathbb{R}^3))^{K'}$, and then in $(\mathrm{H}^1(\mathbb{R}^3))^{K'}$. The convergence of (Ψ^n) follows.

Step 4: Finally one gets $K' = K$ or $K-1$ since there are no critical points of rank less than $K-1$ on $\mathcal{M}_N^K$ (use the same arguments as the one used in [5,3] to show the inequality $E_N^{K+2} < E_N^K$).

Proof of Corollary 2. – Let (c^n, Φ^n) be a minimizing sequence of E_N^K. We use the minimization principle of Borwein and Preiss (*see* [1,4]) to obtain a new minimizing sequence $(\hat{c}^n, \widehat{\Phi}^n)$ with $|\hat{c}_I^n - c_I^n| \to 0$ and $\|\widehat{\Phi}^n - \Phi^n\|_{(\mathrm{H}^1(\mathbb{R}^3))^K} \to 0$, which is the minimizer of the energy with a quadratic perturbation $\mathcal{E} + \frac{1}{n}Q_n$. This new minimizing sequence $(\hat{c}^n, \widehat{\Phi}^n)$ now satisfies the assumptions of Theorem 1, with $j = 0$.

Acknowledgements. I would like to thank Éric Séré for his constant attention and precious advice.

References

[1] J. Borwein, D. Preiss, A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, Trans. Amer. Math. Soc. 303 (1987) 517–527.
[2] G. Friesecke, The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions, Preprint, Mathematical Institute, University of Oxford, UK, 1999 (to appear in Arch. Rat. Mech. Anal.).
[3] G. Friesecke, On the infinitude of nonzero eigenvalues of the single-electron density matrix for atoms and molecules, Preprint, Mathematical Institute, University of Warwick, Coventry, UK, 2001 (submitted to Proc. Roy. Soc. London Ser. A).
[4] N. Ghoussoub, Duality and Perturbation Methods in Critical Point Theory, Cambridge University Press, 1993.
[5] C. Le Bris, A general approach for multiconfiguration methods in quantum molecular chemistry, Ann. Inst. H. Poincare Anal. Non Linéaire 11 (6) (1994) 441–484.
[6] M. Lewin, The multiconfiguration methods in quantum chemistry (in preparation).
[7] E.H. Lieb, B. Simon, The Hartree–Fock theory for Coulomb systems, Comm. Math. Phys. 53 (1977) 185–194.
[8] P.L. Lions, Solutions of Hartree–Fock equations for Coulomb systems, Comm. Math. Phys. 109 (1987) 33–87.
[9] P.O. Löwdin, Quantum theory of many-particle systems. I. Physical interpretations by mean of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction, Phys. Rev. 97 (6) (1955) 1474–1489.